

DIPLOMARBEIT

Optionspreisbewertung mit dünnen Gittern

Angefertigt am

Institut für Numerische Simulation

Vorgelegt der

Mathematisch-Naturwissenschaftlichen Fakultät der

Rheinischen Friedrich-Wilhelms-Universität Bonn

März 2005

Von

Thomas Mertens

aus

Stuttgart

Inhaltsverzeichnis

1	Einleitung	6
2	Dünne Gitter	11
2.1	Einleitung	11
2.1.1	Vorteile dünner Gitter	11
2.2	Aufbau und Eigenschaften dünner Gitter	13
2.2.1	Multi-Level Splitting	13
2.2.2	Eigenschaften der Unterräume	16
2.2.3	Volle versus dünne Gitter	18
2.2.4	Eigenschaften dünner Gitter und der Vergleich mit vol- len Gittern	19
2.3	Approximation von Funktionen auf dünnen Gittern	20
2.3.1	Allgemeinere Funktionenräume	22
2.3.2	Verallgemeinerung der Ansatzfunktionen	26
2.4	Adaptive Verfeinerung	30
2.4.1	Rechnen auf adaptiven dünnen Gittern	32
3	Optionen	34
3.1	Einführung	34
3.1.1	Grundbegriffe und verschiedene Arten von Optionen	35
3.1.2	Betrachtung von Auszahlungen	36
3.1.3	Verwendung von Optionen	38
3.2	Bewertung von Optionen	38
3.2.1	Bezeichnungen und Notationen	39
3.2.2	Annahmen	39
3.2.3	Herleitung	41
3.2.4	Alternative Herleitung	43
3.2.5	Numerik bei Aktienoptionen	45
3.2.6	Erweiterungen des Modells	47
3.3	Die "Greeks"	47
3.3.1	Einige wichtige Greeks	48

3.3.2	Die Berechnung der Greeks mit der Lösung der transformierten Gleichung	50
3.4	Europäische Optionen auf "Aktienkörbe"	51
3.4.1	Transformation auf die Wärmeleitungsgleichung	51
3.4.2	Existenz, Positivität und Eindeutigkeit der Lösung zur Black-Scholes Gleichung	52
3.4.3	Anzahl der Dimensionen	53
3.5	Randbedingungen und Gebietsgröße	54
3.5.1	Abschätzungen zur Gebietsgröße und den Randbedingungen	54
3.5.2	Randbedingungen	56
4	Zinsderivate	60
4.1	Bonds	60
4.1.1	Bewertung von Bonds	61
4.1.2	Randbedingungen	63
4.2	Cross-Currency Swaps	63
4.2.1	Swaps	63
4.2.2	Vorzeitig auflösbarer Differential Swap	65
4.2.3	Ein erweitertes Vasicek-Modell	65
4.2.4	Bewertung der Swaps	67
5	Verwendeter Code und seine Wirkungsweise	69
5.1	Generelle Beschreibung	69
5.2	Problemstellung	70
5.2.1	Vorteile einer adaptiven Diskretisierung in Raumzeit	70
5.3	Übersicht über bisher bekannte Verfahren	71
5.4	Normäquivalenzen	72
5.5	Petrov-Galerkin und Finite Differenzen	73
5.5.1	Diskretisierung des Differentialoperators	74
5.6	Adaptivität	75
5.6.1	Adaptive dünne Gitter versus modifizierte Dünngitter	75
5.6.2	Fehlerschätzer	77
5.7	Sonstige Eigenschaften	77
5.7.1	Löser für Gleichungssysteme	77
5.7.2	Vorkonditionierung	78
5.7.3	Gebiete	78
5.7.4	Parallelisierung	78
5.7.5	Fehlermessung	78
5.7.6	Randbedingungen	79

6	Numerische Ergebnisse	83
6.1	Einleitung	83
6.2	Wahl der Standardeinstellungen	83
6.3	Adaptive versus reguläre resp. volle Gitter	84
6.3.1	Glatte Lösung	85
6.3.2	Unglatte Startdaten und Randkorrekturen	88
6.3.3	Unglatte rechte Seite	91
6.4	Fehlerschätzer	92
6.4.1	Interpolet- versus Lifting-Fehlerindikator	92
6.5	Konvergenzraten für verschiedene Zeitdiskretisierungen	97
6.5.1	Glatte rechte Seite	98
6.5.2	Unglatte Anfangsdaten	98
6.6	Konvergenz des BiCGStab2	100
6.6.1	Verhalten des BiCGStab2 für verschiedene Levels	100
6.6.2	Raten für verschiedene Differentialoperatoren	101
6.6.3	Veränderung der Diskretisierung	103
6.7	Speicherverbrauch	103
6.8	Vergleich der verschiedenen Verfahren	104
7	Anwendungen	105
7.1	Basket-Optionen	105
7.1.1	Gebietsgröße	106
7.1.2	Anfangsbedingungen	106
7.1.3	Randbedingungen	108
7.1.4	Randbedingungen für die transformierte Lösung	114
7.1.5	Konvergenz für Basket-Optionen	115
7.1.6	Greeks für Basket-Optionen	118
7.2	Zinsderivate	119
7.3	Cross-Currency Swaps	120
8	Zusammenfassung und Ausblick	124
A	Aufschlüsselung des DAX	126

Kapitel 1

Einleitung

In dieser Diplomarbeit wird ein numerisches Verfahren vorgestellt, das das Lösen von linearen parabolischen partiellen Differentialgleichungen zweiter Ordnung in mehreren Dimensionen erlaubt. Interessant ist diese Fragestellung, da sich Anwendungen in vielen Wissenschaften, darunter physikalische Chemie, Physik, Ingenieurs- und Finanzwissenschaften finden. Mit letzteren werden wir uns in dieser Arbeit befassen. Das Computerprogramm, das das numerische Verfahren umsetzt, wurde auf die Berechnung von Optionen auf Aktienkörbe, sogenannten Basket-Optionen, und auf Zinsderivate angepasst.

Generell gibt es mehrere Möglichkeiten, Basket-Optionen zu berechnen. Das Problem lässt sich, wie wir später sehen werden, als Erwartungswert über spätere Auszahlungen formulieren. Dieser Erwartungswert kann in seiner üblichen Darstellung als Integral geschrieben werden, das sich dann diskretisieren und numerisch auswerten lässt. Als zweite Möglichkeit, allerdings auch am wenigsten effektivste, kann man die Pfade der geometrischen Brownschen Bewegung möglichst oft simulieren und mitteln, wonach sich nach dem Gesetz der großen Zahlen der Erwartungswert approximieren lässt. Eine dritte Möglichkeit ergibt sich aus der Verknüpfung des Integrals für den Erwartungswert mit einer partiellen Differentialgleichung. Die Lösung der Differentialgleichung ist mit dem gesuchten Erwartungswert identisch. Sie ist jedoch komplex, da sie neben einer ersten Ableitung in der Zeit auch Diffusions-, Konvektions- und Reaktionsterme im Ort mit nicht konstanten Koeffizienten beinhaltet. Für die Diskretisierung steht uns ein Spektrum an Möglichkeiten zur Verfügung. Zum einen sind Finite-Elemente-Verfahren gut geeignet, zum anderen können die Ableitungen direkt mit Finiten Differenzen diskretisiert werden. Wir werden die Literatur zu diesen beiden Möglichkeiten vergleichen und uns dann auf die Finiten Differenzen konzentrieren.

Um die Problemstellung zu vereinfachen, wollen wir die ortsabhängigen Koeffizienten eliminieren. Es gibt verschiedene Möglichkeiten, diese partielle

Differentialgleichung zu transformieren, so dass sie einfacher zu diskretisieren ist. Kürzlich hat Reisinger ([25]) in seiner Dissertation vorgestellt, wie eine Art der Transformation aussehen könnte. Allerdings führt dies nach wie vor zu nicht konstanten Koeffizienten. In dieser Diplomarbeit wird eine neue Methode entwickelt: Der Standard-Ansatz zur Transformation auf die Wärmeleitungsgleichung, wie er beispielsweise bei Seydel ([27]) beschrieben ist, wird auf mehrere Dimensionen erweitert. Die Transformation beinhaltet dann zusätzlich die Drehung des Raumes, allerdings nicht in der Zeitkoordinate.

Da wir nur auf beschränkten Gebieten rechnen können, müssen wir das zugrunde liegende unendliche Gebiet abschneiden. Um weiterhin Existenz und Eindeutigkeit der Lösung garantieren zu können, müssen wir am Rand künstliche Randbedingungen implementieren. Dabei machen wir auf dem Rand einen Fehler, der aber über das Gebiet hinweg kontrolliert werden kann (sogar im Gebiet abfällt). Die entsprechenden Abschätzungen dazu werden wir später behandeln. Wir werden bekannte Randdaten verwenden, weitere Möglichkeiten testen und diese miteinander vergleichen.

Ein wichtiger Punkt beim Lösen der Gleichung wird die Behandlung des Verhältnisses von Raum und Zeit sein. Die einfachste Variante besteht darin, das elliptische Problem, das sich auf jeder Zeitscheibe ergibt, zu lösen und dann mittels eines Zeitschrittverfahrens fortzuschreiten. In dieser Arbeit soll die Gleichung jedoch in Raumzeit gelöst werden. Da wir die Zeit somit als weitere Dimension betrachten, erhöht sich der Rechenaufwand. Wir werden adaptive Gitter verwenden, was sich durch die Glattheit unserer Lösung in Raum und Zeit als nützlich erweist, um gegen den sogenannten "Fluch der Dimensionen" anzukämpfen. Er bezeichnet den exponentiellen Anstieg der Anzahl der Gitterpunkte in der Dimension. Dieser Zuwachs ist intuitiv einleuchtend: nimmt man in jeder Dimensionsrichtung n Punkte, so erhält man in der Fläche n^2 Punkte, im Raum n^3 und allgemein n^d Punkte.

Der Ausgangspunkt des Verfahrens besteht in der Darstellung der Lösung in Multiskalen-Schreibweise. Die Grundlagen dazu finden sich bei Yserentant ([32]). Eine Funktion wird in die Summe von Basisfunktionen mit unterschiedlicher Trägergröße zerlegt. Die Skalierung der Basisfunktion wird eindimensional durch ihre Höhe bestimmt, die in dem jeweiligen Gitterpunkt gespeichert werden kann. Ein Großteil der Aufgabe wird sein, das Gitter so zu wählen, dass wir möglichst aufwendige Probleme rechnen können. Zenger ([33]) hat in seiner grundlegenden Arbeit die Wahl des Gitteraufbaus so getroffen, dass für genügend glatte Funktionen asymptotisch die gleiche Approximationsrate wie bei vollen Gittern (die dem "Fluch" unterworfen sind) entsteht, obwohl seine dünnen Gitter mit lediglich $O(n(\log n)^{d-1})$ Punkten auskommen.

Nun könnte man dagegen einwenden, dass es sich nicht lohne, viel Energie

in die Reduktion der Anzahl der Gitterpunkte zu stecken, da der Fortschritt der Computertechnologie die Berechnung großer Probleme bald möglich machen wird. Dieser Einwand übersieht allerdings, dass gerade mit größeren Problemen, das heißt mit größerem n und d , der Unterschied zwischen der Anzahl der Gitterpunkte bei vollen versus dünnen Gittern anwächst. Mit anderen Worten: je besser die Rechner sind, desto mehr lohnt es sich auch, dünne Gitter einzusetzen!

Wir werden die Anzahl der Gitterpunkte weiter reduzieren, indem wir ein adaptives Verfahren verwenden. Neue Gitterpunkte werden in jedem Verfeinerungsschritt dabei nur noch dort hinzugefügt, wo die Funktion noch nicht hinreichend gut approximiert worden ist, das heißt nur an den Stellen, die den Gesamtfehler reduzieren.

Wir werden allerdings feststellen, dass die echte Dünngitterverfeinerung nicht ausreicht, um Optionspreisaufgaben rechnen zu können. Da die Anfangsdaten bei Optionsaufgaben nicht stetig differenzierbar sind, erfüllt das Problem nicht die erforderlichen Glattheitsvoraussetzungen. Dies resultiert in einem nicht konvergenten Verfahren. Aus diesem Grund stellen wir hier eine neue Möglichkeit des Gitteraufbaus, ein modifiziertes Dünngitter, vor, das auch im Falle von Punkt singularitäten zu einem konvergenten Verfahren führt, indem lokal (beinahe) volle Gitter zugelassen werden. Wie wir später sehen werden, sind die unglatten Stellen bei Optionsaufgaben niederdimensional, so dass mit modifizierten Dünngittern die gleiche Konvergenzrate erzielt werden kann wie mit dünnen Gittern im glatten Fall. Eine alternative Möglichkeit besteht in der künstlichen Randglättung. Wie wir später sehen werden, leidet dann die Konvergenzrate allerdings unter der Balance zweier Fehler.

Eine Herausforderung für Optionsbewertungen ist die Berechnung der sogenannten Greeks. Es handelt sich dabei um Sensitivitäten des Optionswertes — hier also der Lösung der Gleichung — auf Veränderungen des darunterliegenden Wertes, in der Zeit oder auf Änderungen der Parameter. Numerisch stellt sich dabei das Problem, dass eine erste oder zweite Ableitung der numerischen Lösung berechnet werden muss, die gegen die erste bzw. zweite Ableitung der Lösung des Modells konvergieren soll. Damit allerdings eine zweite Ableitung gegen die "wirkliche" Ableitung mit Ordnung zwei konvergieren kann, muss das numerische Verfahren eine Konvergenzordnung von mindestens vier bereitstellen können. Wir wollen hier an Approximationsgenauigkeit gewinnen, indem wir Basisfunktionen mit einem höheren Grad an polynomialer Exaktheit verwenden. Aus diesem Grund verwenden wir Interpolants als Basisfunktionen. Wir werden später sehen, dass sich mit unserem Verfahren der Grad der Konsistenzordnung der Finiten Differenzen (durch das Verwenden entsprechender Sterne) so einstellen lässt, dass wir die nöti-

gen Ordnungen erzielen können.

Koster ([22]) hat in seiner Dissertation ein ähnliches Verfahren für elliptische Probleme vorgestellt. Neu ist hier, dass wir die Approximationseigenschaften aus der Arbeit von Koster nun für parabolische Gleichungen zeigen. Die Adaptivität findet in Raumzeit statt, was uns bei der Berechnung von Optionen erhebliche Einsparungen gegenüber einem Zeitschrittverfahren erlaubt. Der Grund dafür liegt in der Art der Gleichung: betrachten wir das Integrationsproblem, so sehen wir, dass die Lösung entsteht, indem man die Anfangsdaten gegen den Gauß-Kern integriert. Damit wissen wir von der Lösung, dass sie auch bei unglatten Anfangsdaten (hier $C^{0,1}$) nach einem ε -Schritt in der Zeit analytisch ist. Der Gauß-Kern glättet über die Zeit alle Unebenheiten aus, wobei in der Umgebung einer Unglattheit hohe Ableitungen durchaus auftreten können. Wir wollen ein in Raumzeit adaptives Verfahren einsetzen und können so die Glattheit optimal ausnutzen. Verschiedene Fehlerschätzer für das adaptive Verfahren werden später diskutiert.

Wir möchten hier noch einmal stichpunktartig die wichtigsten Neuerungen dieser Arbeit vorstellen.

- Behandlung von parabolischen Gleichungen mit dünnen Gittern und glatten Ansatzfunktionen
- Einführung modifizierter Dünngitter
- Korrektur unglatter Startdaten
- Transformation der multivariaten Black-Scholes Gleichung auf die Wärmeleitungsgleichung
- Vergleich verschiedener Randdaten

Die Arbeit ist wie folgt gegliedert: Kapitel 2 führt dünne Gitter ein und behandelt Aussagen zu deren Eigenschaften. Kapitel 3 beschäftigt sich mit Optionen. Wir stellen dort das entsprechende Modell zur Bewertung des fairen Wertes einer Option vor und diskutieren verschiedene Arten von Optionen. Wir interessieren uns insbesondere für "Basket-Optionen" und zeigen, wie diese numerisch bewertet werden können. Dazu präsentieren wir eine Transformation und führen Randwerte ein. In Kapitel 4 stellen wir zwei Arten von Zinsderivaten vor, deren numerische Berechnung von besonderem Interesse ist. Kapitel 5 zeigt, wie wir die theoretischen Ergebnisse aus den Kapiteln 2 bis 4 numerisch umgesetzt haben. Das Verfahren wird in Kapitel

6 an Modellproblemen vorgestellt, bevor wir dann die Ergebnisse zu den Anwendungen in Kapitel 7 präsentieren. Kapitel 8 schließt ab und zeigt mögliche Anknüpfungspunkte an diese Arbeit auf.

Ich möchte mich bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben. An vorderster Stelle ist hier mein Betreuer, Herr Prof. Dr. Michael Griebel, zu nennen, der mein Interesse auf dünne Gitter lenkte und mir während der Erstellung der Arbeit mit zahlreichen Hilfestellungen zur Seite stand. Herrn Prof. Dr. Rolf Krause danke ich für die Übernahme des Zweitgutachtens. Ebenfalls bedanken möchte ich mich bei allen Kollegen am Institut für Numerische Simulation (INS) der Universität Bonn für die angenehme Arbeitsatmosphäre und besonders bei Dr. Marcel Arndt, Dr. Thomas Gerstner und Daniel Oeltz, die stets bereit waren, über auftretende Probleme zu diskutieren. Hilfreiche Kommentare erhielt ich von Prof. Ulrich Doraszelski, PhD. Dankbar bin ich auch den Kollegen am INS in Bonn und den Mitarbeitern am Department of Economics an der Harvard University für das Bereitstellen der Rechner. Ich freue mich, mit Christian Feuersänger und Allan Zulficar Zimmerkollegen gehabt zu haben, mit denen das Arbeiten Spaß machte. Dr. Marcel Arndt gebührt besonderer Dank für das Korrekturlesen. Weiterhin bin ich meinen Eltern und meinem Bruder Stefan dankbar, dass sie mich während des Studiums in allen Belangen unterstützt haben.

Cambridge, MA im März 2005

Kapitel 2

Dünne Gitter

2.1 Einleitung

Dünne Gitter sind seit der grundlegenden Arbeit von Zenger ([33]) auf die meisten numerischen Methoden zum Lösen von Differentialgleichungen und Integrationsproblemen übertragen worden. So wurden sie für Finite Elemente und Galerkin-Verfahren, Kollokation, Finite Differenzen, Finite Volumen, Spektralverfahren und für die Kombinationstechnik eingesetzt. In diesem Kapitel werden wir die Entstehung und Eigenschaften dünner Gitter betrachten.

2.1.1 Vorteile dünner Gitter

Der Grund, warum dünne Gitter für die numerische Simulation attraktiv geworden sind, liegt auf der Hand. Die Anzahl der Gitterpunkte und der damit einhergehende Rechen- und Speicheraufwand stellen eine Beschränkung der Rechenkapazität dar. Der konventionelle Ansatz, volle Gitter, benutzt $N = O(n^d)$ Gitterpunkte auf dem Gebiet, wobei n die Anzahl der Punkte in jeder der d Dimensionen bezeichnet. Dieser exponentielle Zuwachs des Aufwandes ist unter dem Stichwort "Fluch der Dimension" bekannt. Das Ziel dünner Gitter ist nun, diesen Fluch zumindest teilweise brechen zu können. In der Tat läßt sich die Anzahl der Gitterpunkte auf $N = O(n(\log n)^{d-1})$ reduzieren. In der Energienorm kommt das Verfahren sogar mit einer Anzahl von $N = O(n)$, also unabhängig von der Dimension, aus und ist damit bestens geeignet für hochdimensionale Problemstellungen. Wie eingangs erwähnt, macht sich der Unterschied in der Anzahl der Gitterpunkte gerade bei starken Verfeinerungen und großen Problemen bemerkbar.

Im folgenden wollen wir den Aufbau dünner Gitter genauer untersuchen und Abschätzungen zu Approximationsgenauigkeiten finden. Der Ansatz, wie er bereits in der ursprünglichen Arbeit von Zenger ([33]) präsentiert wurde,

sieht zunächst ein Splitting der Funktion in Unterräume von Ansatzfunktionen verschiedenen Levels vor. Entgegen dem Ansatz von Yserentant ([32]) werden nun anisotrope Verfeinerungen (d.h. unterschiedliche Verfeinerungsgrade in den Richtungen) zugelassen, die erheblich bessere Approximationseigenschaften im Kosten-Nutzen Verhältnis haben, wenn die Funktion entsprechend glatt ist, das heißt in diesem Fall beschränkte gemischte Ableitungen besitzt. Der konventionelle Ansatz mit isotropen Verfeinerungen (also gleiche Maschenweite in jede Richtung) ist besser auf seine Approximationseigenschaften untersucht. Besonders gut sieht man die Vorteile dünner Gitter am Beispiel der Kantensingularität.

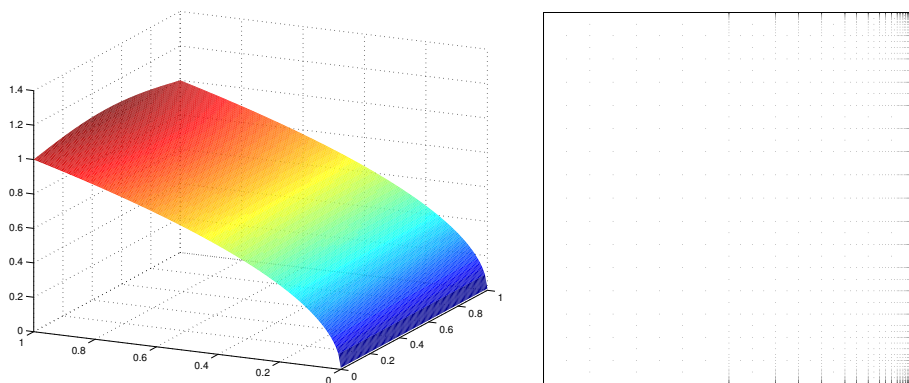


Abbildung 2.1: Singularität auf der Kante und zugehöriges adaptives Gitter

Ist die Funktion in einer Richtung konstant, so sollte das Verfahren sogar mit Ordnung $N = O(\log n)$ Punkten auskommen, da in die Richtung, in der die Funktion konstant ist, nicht verfeinert werden muss. Eine solche Funktion ist auf der linken Seite der Abbildung 2.1 zu sehen. Das adaptive Gitter muss allerdings in unserem Falle auch an den Rändern leicht verfeinern, da der Rand von dem Innern der Funktion getrennt werden muss. Wir kommen auf diesen Punkt später zurück.

Für allgemeine Funktionen mit beschränkten gemischten Ableitungen erreicht man mit der hierarchischen Basis (die "Hütchenfunktion") eine Approximationsgenauigkeit von $O(n^{-2}(\log n)^{d-1})$ in der L_2 - und L_∞ -Norm. Für allgemeine Funktionen aus Besov-Räumen ist bekannt, dass die Funktionen genauso gut (in manchen Fällen bis auf einen logarithmischen Faktor, der von der Dimension abhängen kann) approximiert werden. Das Verfahren lässt sich auch auf nichtglatte Funktionen anwenden, wenn die Verfeinerung adaptiv ist. Wir werden entsprechende Messungen dazu im Kapitel über den verwendeten Code und in den numerischen Ergebnissen sehen.

Anschließend werden wir betrachten, wie sich Gleichungen — insbesonde-

re partielle Differentialgleichungen — auf dünnen Gittern lösen lassen. Hier sollen Aussagen zu dem verwendeten Kollokations- bzw. Finite Differenzen-Verfahren getroffen werden.

2.2 Aufbau und Eigenschaften dünner Gitter

2.2.1 Multi-Level Splitting

Wir beziehen uns im folgenden auf den Raum der Funktionen mit beschränkten gemischten Ableitungen¹ bis zur Ordnung \mathbf{r}

$$X^{\mathbf{r},d}(\bar{\Omega}) = \{u^{(d)} : \bar{\Omega} \subset \mathbb{R}^d \rightarrow \mathbb{R}, D^{\mathbf{a}}u \in C^0(\bar{\Omega}), \mathbf{a} \leq \mathbf{r}\}$$

auf dem Einheitswürfel $\bar{\Omega} = [0, 1]^d$, wobei

$$D^{\mathbf{a}}u := \frac{\partial^{|\mathbf{a}|_1} u}{\partial x_1^{a_1} \cdots \partial x_d^{a_d}}$$

die gemischten Ableitungen der mehrdimensionalen Funktion u sind. Mit $X_0^{\mathbf{r},d}(\bar{\Omega})$ bezeichnen wir analog die Teilmenge obigen Raumes, in dem die Funktionen auf dem Rand verschwinden. Auf dem Raum werden die Seminormen

$$|u|_{\mathbf{a},\infty} := \|D^{\mathbf{a}}u\|_{\infty} \quad \text{bzw.} \quad |u|_{\mathbf{a},2} := \|D^{\mathbf{a}}u\|_2 = \left(\int_{\bar{\Omega}} |D^{\mathbf{a}}u|^2 d\mathbf{x} \right)^{\frac{1}{2}}$$

eingeführt. Für die Multiindizes werden Normen, Multiplikationen und Relationen eingeführt, die teilweise in obiger Definition bereits verwendet wurden:

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &:= (a_1 b_1, \dots, a_d b_d) & \gamma \cdot \mathbf{a} &:= (\gamma a_1, \dots, \gamma a_d) \\ \mathbf{a} \leq \mathbf{b} &\iff \forall_{1 \leq j \leq d} a_j \leq b_j & 2^{\mathbf{a}} &:= (2^{a_1}, \dots, 2^{a_d}) \\ |\mathbf{a}|_{\infty} &= \max_{1 \leq j \leq d} a_j & |\mathbf{a}|_1 &= \sum_i a_i \end{aligned}$$

Diese Operationen mit den Multiindizes erlauben eine kompakte Schreibweise der Entstehung dünner Gitter.

Für Level $\mathbf{l} = (l_1, \dots, l_d) \in \mathbb{N}^d$ erhalten wir eine Familie von Gittern

$$\{\Omega_{\mathbf{l}} : \mathbf{l} \in \mathbb{N}^d\}$$

¹Wie bereits von Schiekhofer ([26]) erläutert, müssen im Fall der Finiten Differenzen stetige gemischte Ableitungen angenommen werden, da an Punkten ausgewertet wird. Dies macht in L_p -Räumen keinen Sinn, da die Aussagen nur "fast überall" gelten.

als Partitionen auf $\bar{\Omega}$. Die Maschenweite sei dabei

$$\mathbf{h}_1 := (h_{l_1}, \dots, h_{l_d}) := 2^{-1},$$

sodass Punkte

$$\mathbf{x}_{1,\mathbf{i}} := (x_{l_1,i_1}, \dots, x_{l_d,i_d}) := \mathbf{i} \cdot \mathbf{h}_1, \quad \mathbf{0} \leq \mathbf{i} \leq 2^{\mathbf{1}}$$

entstehen, die durch ihr Level \mathbf{l} und ihren Ort \mathbf{i} eindeutig bestimmt sind. Die Gitter sind also auf jedem Level rechteckig und in jede Raumrichtung äquidistant. Lediglich die Maschenweiten in die verschiedenen Richtungen können voneinander abweichen.

Wir bestimmen nun eine 1D-Ansatzfunktion $\phi : [0, 1] \rightarrow \mathbb{R}$, die durch Dilatation und Translation den Gitterpunkten zugewiesen wird. Anschließend liefert das Tensorprodukt unsere mehrdimensionale Ansatzfunktion

$$\phi_{1,\mathbf{i}}(\mathbf{x}) = \prod_{j=1}^d \phi_{l_j,i_j}(x_j) = \prod_{j=1}^d \phi\left(\frac{x_j - i_j \cdot h_j}{h_j}\right).$$

Die Implikationen für das Lösen von Gleichungen werden wir später diskutieren. Nimmt man nun nur die ungeraden Indizes

$$\mathbf{I}_1 := \{\mathbf{i} \in \mathbb{N}^d : \mathbf{1} \leq \mathbf{i} \leq 2^{\mathbf{1}} - \mathbf{1}, i_j \text{ ungerade für alle } 1 \leq j \leq d\},$$

so haben die Ansatzfunktionen auf jedem Level disjunkte Träger. Nehmen wir als Ansatzfunktion die beispielsweise die hierarchische Basis, so entsteht auf natürliche Weise (siehe ([32]) oder ([33])) eine Basis, die (nach Vervollständigung bezüglich der H^1 -Norm) den Sobolev-Raum $H_0^1(\bar{\Omega})$ aufspannt.

Seien dazu

$$V_1 = \text{span} \{ \phi_{1,\mathbf{i}}(\mathbf{x}) : i_j = 1, \dots, 2^{l_j} - 1, 1 \leq j \leq d, \mathbf{l} \in \mathbb{N}^d \}$$

die Räume der Ansatzfunktionen, bei denen die Träger nicht disjunkt sind. Anhand derer können wir die Differenzräume

$$W_1 = V_1 \setminus \bigoplus_{j=1}^d V_{1-\mathbf{e}_j}$$

definieren, wobei \mathbf{e}_j der Einheits-Multiindex ist. Nun erhalten wir, wie oben angekündigt, den Raum V durch

$$V = \sum_{l_1=1}^{\infty} \cdots \sum_{l_d=1}^{\infty} W_{(l_1, \dots, l_d)} = \bigoplus_{\mathbf{l} \geq \mathbf{1}} W_{\mathbf{l}}, \quad (2.1)$$

der nach Vervollständigung bezüglich der H^1 -Norm identisch ist mit dem Sobolev-Raum $H_0^1(\bar{\Omega})$.

Das bedeutet, die Räume W_1 lassen sich durch die Basis darstellen:

$$W_1 = \text{span} \{ \phi_{1,i}(\mathbf{x}) \in V_1 : i_j = 1, \dots, 2^{l_j} - 1 \text{ mit } i_j \text{ ungerade} \} \leq j \leq d \}$$

Nach Erhalt der Basis können wir nun jede Funktion $u \in H_0^1(\bar{\Omega})$ und daher auch jede Funktion $u \in X_0^{\mathbf{r},d}(\bar{\Omega})$ eindeutig darstellen als

$$u(\mathbf{x}) = \sum_1 \sum_{i \in \mathbf{I}_1} u_{1,i} \phi_{1,i}(\mathbf{x}).$$

Formal kann das dünne Gitter Approximationsschema als Optimierungsproblem aufgefasst werden. Genaueres ist in Bungartz und Griebel ([8]) mit Bezug auf DeVore ([12]) zu finden.

Nun wenden wir uns dem Ergebnis des betrachteten Vorgehens zu. Wählen wir als eindimensionale Mutterfunktion $\phi(x)$ zum Beispiel das "Hütchen", so erhalten wir die hierarchische Basis als Ansatzfunktionen.

$$\phi(x) = \begin{cases} 1 - |x| & \text{für } x \in [0, 1] \\ 0 & \text{sonst} \end{cases}$$

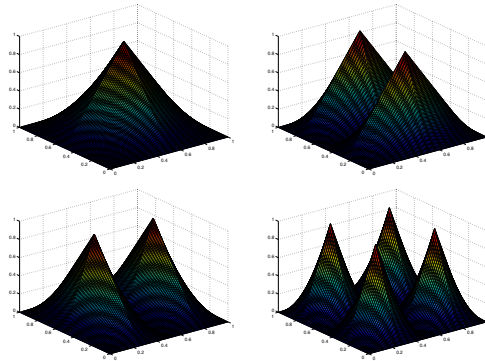


Abbildung 2.2: Zweidimensionale hierarchische Basis auf Leveln 1 und 2

Die Approximation von glatten Funktionen sei hier exemplarisch anhand der Parabel in Abbildung 2.3 dargestellt. Sichtbar ist hier, dass die Koeffizienten mit zunehmendem Level abfallen. Genaueren Abschätzungen werden wir uns unten widmen.

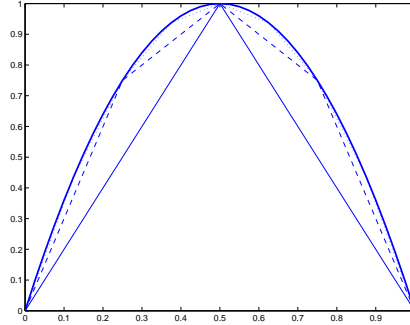


Abbildung 2.3: Approximation der nach unten geöffneten Parabel mit hierarchischen Ansatzfunktionen

2.2.2 Eigenschaften der Unterräume

Zunächst beobachten wir, dass jeder Unterraum W_1 auf Level l 2^{l-1} Punkte besitzt. Die Anzahl der Punkte für ein bestimmtes Level hängt vom Verfahren ab — je nachdem ob volle oder dünne Gitter verwendet werden. Wir werden dies unten im Vergleich der beiden Verfahren diskutieren.

Zunächst jedoch widmen wir uns den Koeffizienten. Wir betrachten diese am Beispiel der hierarchischen Basis $\phi_{1,i}(\cdot)$, die bereits oben eingeführt wurde. Wie man leicht nachvollziehen kann, stellt sich ein Koeffizient $u_{1,i}$ dadurch dar, dass die benachbarten Punkte ihren hierarchischen Anteil beisteuern. Der Koeffizient errechnet sich letztlich an jedem Punkt durch

$$u_{1,i} = \left(\prod_{j=1}^d \begin{bmatrix} -\frac{1}{2} & 1 & -\frac{1}{2} \end{bmatrix}_{x_{l_j, i_j}, h_j} \right) u(\mathbf{x}).$$

Kontinuierlich läßt sich das dadurch darstellen, dass "gegen ein Hütchen" integriert wird. Dies wird im folgenden Satz verdeutlicht:

Satz 2.2.1 (Kontinuierliche Darstellung der Hierarchisierung)

Sei $\psi_{1,i}(\mathbf{x}) := \prod_{j=1}^d 2^{-l_j+1} \cdot \phi_{l_j, i_j}(x_j)$ die umskalierte Ansatzfunktion. Dann berechnet sich der Koeffizient für jedes $u \in X_0^{\mathbf{r},d}(\bar{\Omega})$ durch

$$u_{1,i} = \int_{\Omega} \psi_{1,i}(\mathbf{x}) \cdot D^2 u(\mathbf{x}) d\mathbf{x}. \quad (2.2)$$

Beweis: Integriert man (2.2) partiell, so verschwindet die Funktion an den Rändern. Der Hauptsatz der Differential- und Integralrechnung liefert dann das gewünschte Ergebnis. Die Rechnung findet sich in ([7]), Lemma 3.

□

Dieser Satz wird später wichtig sein, wenn wir die Regularisierung des Randes diskutieren, um C^0 -Anfangsdaten behandeln zu können.

Um ein Kriterium für die Verfeinerung im Falle von Adaptivität zu erhalten, macht man sich meist den Abfall der Koeffizienten zu Nutze (siehe z.B. DeVore ([12])). Zu zeigen ist allerdings noch, wie sich die hierarchischen Koeffizienten mit zunehmendem Level verhalten. Hier sollte — zumindest in der Konstanten — die Glattheit der Funktion u eingehen.

Lemma 2.2.1 *Sei $u \in X_0^{\mathbf{r},d}(\bar{\Omega})$. Dann gilt für die zugehörigen Koeffizienten*

$$\begin{aligned} |u_{1,i}| &\leq 2^{-d} \cdot 2^{-2 \cdot |l_1|} \cdot |u|_{\mathbf{2},\infty}, \\ |u_{1,i}| &\leq 2^{-d} \cdot \left(\frac{2}{3}\right)^{d/2} \cdot 2^{-(3/2) \cdot |l_1|} \cdot |u|_{\text{supp}(\phi_{1,i})}|_{\mathbf{2},2}, \end{aligned} \quad (2.3)$$

wobei $\text{supp}(\cdot)$ den Träger bezeichnet.

Beweis: Siehe Bungartz und Griebel ([8]). □

Aggregiert man dieses Ergebnis über alle Koeffizienten auf einem Level (d.h. man läuft über alle Orte), so erhält man

Lemma 2.2.2 *Für die Komponenten $u_1 = \sum_{i \in \mathbf{I}} u_{1,i} \cdot \phi_{1,i}(\mathbf{x}) \in W_1$ gilt*

$$\begin{aligned} \|u_1\|_{\infty} &\leq 2^{-d} \cdot 2^{-|l_1|} \cdot |u|_{\mathbf{2},\infty}, \\ \|u_1\|_2 &\leq 3^{-d} \cdot 2^{-|l_1|} \cdot |u|_{\mathbf{2},2}, \\ \|u_1\|_E &\leq \frac{1}{2 \cdot 12^{(d-1)/2}} \cdot 2^{-|l_1|} \cdot \left(\sum_{j=1}^d 2^{2 \cdot l_j}\right)^{1/2} \cdot |u|_{\mathbf{2},\infty}, \\ \|u_1\|_E &\leq \sqrt{3} \cdot 3^{-d} \cdot 2^{-|l_1|} \cdot \left(\sum_{j=1}^d 2^{2 \cdot l_j}\right)^{1/2} \cdot |u|_{\mathbf{2},2}, \end{aligned}$$

wobei die Energienorm definiert ist als

$$\|u\|_E := \left(\int_{\Omega} \sum_{j=1}^d \left(\frac{\partial u(\mathbf{x})}{\partial x_j} \right)^2 d\mathbf{x} \right)^{1/2}$$

Beweis: Siehe ([8]). □

2.2.3 Volle versus dünne Gitter

Nachdem wir nun eine Aufteilung des Funktionenraums V in Unterräume betrachtet haben, wenden wir uns den verschiedenen Arten zu, Funktionen mit endlichen Summen zu approximieren. Von nun an soll nur noch eine endliche Menge an Unterräumen verschiedenen Levels betrachtet werden. Nach Lemma 2.2.2 ist der Fehler klein, wenn man das Level groß genug wählt. Dass wir zudem auch noch die Anzahl der aktiven Ortspunkte einschränken können, werden wir später anhand des adaptiven Algorithmus sehen. Wir führen nun zwei verschiedene Arten von endlich-dimensionalen Funktionenräumen ein:

Definition 2.2.1 *Der endlich-dimensionale Raum*

$$V_n^{(\infty)} = \bigoplus_{|\mathbf{l}|_\infty \leq n} W_{\mathbf{l}}$$

ist der Raum der stückweise d -linearen Funktionen auf dem vollen Gitter mit äquidistanter Maschenweite $h_n = 2^{-n}$.

Der Raum

$$V_n^{(1)} = \bigoplus_{|\mathbf{l}|_1 \leq n+d-1} W_{\mathbf{l}}$$

ist der entsprechende Raum auf dem dünnen Gitter.

Ein dünnes Gitter ist dabei wie folgt definiert:

Definition 2.2.2 **Dünnes Gitter** *Ein dünnes Gitter Ω^{dg} wird als Vereinigung der Gitter*

$$\Omega^{dg} := \bigcup_{|\mathbf{l}|_1 \leq n+d-1} \Omega_{\mathbf{l}}$$

gebildet.

Für dünne Gitter benutzen wir also — gegenüber dem Ansatz, der auf volle Gitter führt — weniger Unterräume. Die Summe der Levelindizes (d.h. die 1-Norm des Multiindex) soll kleiner als das vorgegebene Level sein. Beim entsprechenden vollen Gitter verwenden wir die Maximum-Norm des Levelindex. Dies ist in Abbildung 2.4 deutlich gemacht.

Das entstehende Gitter hat nun die für dünne Gitter charakteristische Form, die in Abbildung 2.5 im Vergleich zum vollen Gitter dargestellt ist.

Die nächste Frage, die sich nun aufdrängt, betrifft die Eigenschaften des entstehenden dünnen Gitters. Was können wir zur Komplexität sagen? Wie ist sie im Vergleich zum vollen Gitter? Diesen Fragen werden wir uns im nächsten Abschnitt zuwenden.

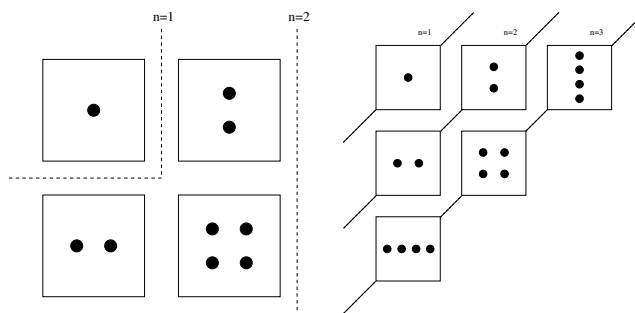


Abbildung 2.4: Unterräume zur Erzeugung eines vollen Gitters vom Level 2 und dünnes Gitter des Levels 3

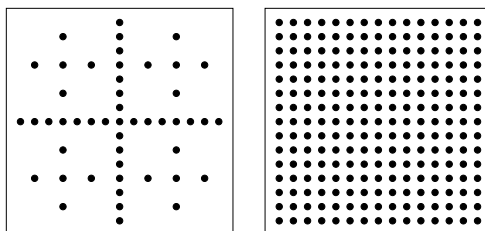


Abbildung 2.5: Dünnes Gitter und entsprechendes volles Gitter jeweils auf Level 4

2.2.4 Eigenschaften dünner Gitter und der Vergleich mit vollen Gittern

Als Maß für die Komplexität des Verfahrens wollen wir die Anzahl der Gitterpunkte betrachten, die notwendig sind, um einen Funktionenraum aufzubauen. Im folgenden Abschnitt zeigen wir, wie gut wir eine Funktion mit diesen Punkten approximieren können. Wir werden dabei die entscheidenden Vorteile dünner Gitter gegenüber vollen Gittern sehen.

$|V_n^\infty| = (2^n - 1)^d = O(2^{d \cdot n}) = O(h_n^{-d})$ ist offensichtlich die Anzahl der Punkte eines vollen Gitters. Wie eingangs bereits beschrieben, sieht man an dieser Gleichung, was mit dem "Fluch der Dimensionen" gemeint ist. Die Anzahl der benötigten Punkte für ein fixes Level steigt exponentiell mit der Dimension.

Die Anzahl der Punkte auf einem dünnen Gitter zu bestimmen erfordert genaueres Hinsehen.

Lemma 2.2.3 *Die Anzahl der Freiheitsgrade im Innern eines dünnen Git-*

ters wird durch

$$\begin{aligned}
|V_n^{(1)}| &= \sum_{i=0}^{n-1} 2^i \cdot \binom{d-1+i}{d-1} \\
&= (-1)^d + 2^n \cdot \sum_{i=0}^{d-1} (-2)^{d-1-i} \binom{n+d-1}{i} \\
&= 2^n \cdot \left(\frac{n^{d-1}}{(d-1)!} + O(n^{d-2}) \right)
\end{aligned} \tag{2.4}$$

beschrieben. Damit haben wir — wie eingangs erwähnt — die Gittergröße $|V_n^{(1)}| = O(h_n^{-1} \cdot |\log_2 h_n|^{d-1})$.

Beweis: Man wendet die Definition von $|V_n^{(1)}|$ an und nutzt aus, dass man die Größe der W_1 kennt. Die Rechnung kann in Bungartz und Griebel ([8]) nachgelesen werden. \square

Es kann sogar leicht eine Bijektion von der Menge der Punkte $\mathbf{x}_{1,i}$ auf die Menge der natürlichen Zahlen angegeben werden. Dazu muss man wissen, dass alle Räume W_1 mit $|\mathbf{l}|_1 = c$ gleich viele, nämlich 2^{c-d} , Punkte besitzen. Nun ist interessant, wieviele dieser Unterräume in d Dimensionen auf Level $|\mathbf{l}|_1 - d + 1$ existieren. Diese Frage wird mit dem folgenden Lemma beantwortet:

Lemma 2.2.4 *Um zu berechnen, wieviele Unterräume $U_d(n)$ auf einem Level $n = |\mathbf{l}|_1 + d - 1$ verbraucht werden, kann eine geschlossene Formel angegeben werden:*

$$U_d(n) = \left\{ \begin{array}{ll} 1 & \text{für } n \leq 2 \\ \sum_{z=1}^{\min(n-2,d)} \binom{d}{z} \binom{n-2}{z-1} & \text{sonst} \end{array} \right\}$$

Wendet man die Formel des Lemmas an, kann man die Unterräume über alle Levels summieren, wobei jeder Unterraum auf dem jeweiligen Level die gleiche Anzahl an Punkten hat.

2.3 Approximation von Funktionen auf dünnen Gittern

Da wir nun die verschiedenen Komplexitäten verglichen haben, können wir uns dem spannendsten Teil der Fragestellung widmen. Wie gut approximieren

die Verfahren? Kann man mit dünnen Gittern annähernd so gute Ergebnisse erzielen wie mit vollen Gittern, obwohl weniger Punkte spendiert werden?

Dazu betrachten wir folgenden Satz:

Satz 2.3.1 *Die Approximationen $u_n^{(\infty)}$ auf einem vollen und $u_n^{(1)}$ auf einem dünnen Gitter haben folgende Approximationsfehler an $u \in X_0^{\mathbf{r},q}(\bar{\Omega})$:*

$$\|u - u_n^{(\infty)}\|_{\infty} \leq \frac{d}{6^d} \cdot 2^{-2n} \cdot |u|_{\mathbf{2},\infty} = O(h_n^2) \quad (2.5)$$

$$\|u - u_n^{(\infty)}\|_2 \leq \frac{d}{9^d} \cdot 2^{-2n} \cdot |u|_{\mathbf{2},2} = O(h_n^2) \quad (2.6)$$

$$\|u - u_n^{(\infty)}\|_E \leq \frac{d^{3/2}}{2 \cdot 3^{(d-1)/2} \cdot 6^{d-1}} \cdot 2^{-n} \cdot |u|_{\mathbf{2},\infty} = O(h_n) \quad (2.7)$$

$$\|u - u_n^{(\infty)}\|_E \leq \frac{d^{3/2}}{\sqrt{3} \cdot 9^{d-1}} \cdot 2^{-n} \cdot |u|_{\mathbf{2},2} = O(h_n), \quad (2.8)$$

beziehungsweise

$$\|u - u_n^{(1)}\|_{\infty} \leq \frac{1}{8^d} \cdot 2^{-2n+1} \cdot A(d, n) |u|_{\mathbf{2},\infty} = O(h_n^2 \cdot n^{d-1}) \quad (2.9)$$

$$\|u - u_n^{(1)}\|_2 \leq \frac{1}{12^d} \cdot 2^{-2n+1} \cdot A(d, n) |u|_{\mathbf{2},2} = O(h_n^2 \cdot n^{d-1}) \quad (2.10)$$

$$\|u - u_n^{(1)}\|_E \leq \frac{d}{3^{(d-1)/2} \cdot 4^{d-1}} \cdot 2^{-(n+1)} \cdot |u|_{\mathbf{2},\infty} = O(h_n) \quad (2.11)$$

$$\|u - u_n^{(1)}\|_E \leq \frac{d}{\sqrt{3} \cdot 4^{d-1}} \cdot 2^{-n} \cdot |u|_{\mathbf{2},2} = O(h_n), \quad (2.12)$$

wobei

$$A(d, n) = \sum_{k=0}^{d-1} \binom{n+d-1}{k} = \frac{n^{d-1}}{(d-1)!} + O(n^{d-2})$$

Das Verfahren der dünnen Gitter erlaubt also bezüglich der L_{∞} - und L_2 -Norm eine Approximation zweiter Ordnung.

Definition 2.3.1 *Die Approximationsordnung $\alpha(X)$ für ein Verfahren bezüglich dem Funktionenraum X ist definiert als:*

$$\alpha(X) := \sup \{ \alpha : \sigma_{n,2}(u) = O(h_n^{\alpha} \|u\|_X), \forall u \in X \},$$

wobei

$$\sigma_{n,2}(u) := \inf_{I:|I|=2^n} \inf_{\{u_{(1,i)}\}_{(1,i) \in I}} \|u - \sum_{(1,i) \in I} u_{(1,i)} \phi_{(1,i)}\|_2$$

und Λ eine Indexmenge in Level und Ort ist.

Zunächst wollen wir den Koeffizientenabfall bei der Approximationen der Funktionen in $X^{r,d}(\bar{\Omega})$ untersuchen. Wir werden sehen, wie die Koeffizienten mit zunehmendem Level gegen Null streben und können somit später die Größe des Absolutbetrages der Koeffizienten als Kriterium für adaptive Verfeinerungen verwenden.

Satz 2.3.2 Koeffizientenabfall

Sei $u \in X_0^{r,d}(\bar{\Omega})$ in hierarchischer Darstellung gegeben. Dann erfüllen die hierarchischen Koeffizienten $u_{1,i}$ folgende Eigenschaften:

$$|u_{1,i}| \leq 2^{-d} \cdot 2^{-2 \cdot |l_1|} \cdot |u|_{\mathbf{2},\infty} \quad (2.13)$$

$$|u_{1,i}| \leq 2^{-d} \cdot \left(\frac{2}{3}\right)^{d/2} \cdot 2^{-3/2 \cdot |l_1|} \cdot |u|_{\text{supp}(\phi_{1,i})}_{\mathbf{2},2}, \text{ sowie} \quad (2.14)$$

$$\|u_1\|_{\infty} \leq 2^{-d} \cdot 2^{-2 \cdot |l_1|} \cdot |u|_{\mathbf{2},\infty} \quad (2.15)$$

$$\|u_1\|_2 \leq 3^{-d} \cdot 2^{-2 \cdot |l_1|} \cdot |u|_{\mathbf{2},2} \quad (2.16)$$

$$\|u_1\|_E \leq \frac{1}{2 \cdot 12^{(d-1)/2}} \cdot 2^{-2 \cdot |l_1|} \cdot \left(\sum_{j=1}^d 2^{2 \cdot l_j}\right)^{1/2} \cdot |u|_{\mathbf{2},\infty} \quad (2.17)$$

$$\|u_1\|_E \leq \sqrt{3} \cdot 3^{-d} \cdot 2^{-2 \cdot |l_1|} \cdot \left(\sum_{j=1}^d 2^{2 \cdot l_j}\right)^{1/2} \cdot |u|_{\mathbf{2},\infty} \quad (2.18)$$

Beweis: siehe Bungartz und Griebel ([8]). □

Die Koeffizienten in der hierarchischen Darstellung fallen also exponentiell mit dem Level der Verfeinerung ab. Als Konstante gilt hier die Seminorm der Funktion. Es ist dabei durchaus erlaubt, sich die Glattheit lokal vorzustellen. Mit zunehmendem Level wird der Träger der Basisfunktionen kleiner, sodass die Koeffizienten für glatte Funktionen schnell gegen Null streben.

2.3.1 Allgemeinere Funktionenräume

Bisher haben wir uns auf den Funktionenraum $X^{r,d}(\bar{\Omega})$ beschränkt. Wir betrachten in diesem Abschnitt nun auch die allgemeinere Approximation von Funktionen in Sobolev-Räumen beziehungsweise den entsprechenden Besov-Räumen. Dass die Approximation auch in diesen Funktionenräumen funktioniert, ist für das weitere Vorgehen erheblich, denn — wie wir später sehen werden — werden unsere Anfangsdaten an einer Stelle nicht differenzierbar sein. Ziel soll es sein, eine Abschätzung zu erhalten, die uns die Approximation von Punktsingularitäten — wie im Falle der stückweise differenzierbaren Anfangsdaten — erlaubt.

Als Ansatzfunktionen werden nun Wavelets verwendet, die per Tensorprodukt auf mehrere Dimensionen verallgemeinert werden. Die dadurch gewonnene Biorthogonalität erlaubt die Schreibweise, dass der Überschuss bezüglich der Wavelet-Darstellung, der wie üblich mit $u_{(\mathbf{l}, \mathbf{i})}$ bezeichnet wird, und das Skalarprodukt $\langle \tilde{\phi}, u \rangle$ gleich groß sind. Die Funktion $\tilde{\phi}$ ist dabei zur Ansatzfunktion ϕ dual. Diese sollte im Raum L_2 sein, um eine L_2 -stabile Multiskalendarstellung zu erhalten, die man für die entsprechenden Normabschätzungen benötigt.

Von besonderer Wichtigkeit ist hier die L_2 -Stabilität der univariaten Ansatzfunktion auf dem jeweiligen Level. Diese ist bereits von der soeben erwähnten Form.

Lemma 2.3.1 *Auf dem Level $l \geq l_0$ gilt mit levelunabhängigen Äquivalenzkonstanten*

$$\left\| \sum_i u_{l,i} \cdot \phi_{l,i} \right\| \sim 2^{-l} \sum_i |u_{l,i}|^2 \quad (2.19)$$

Der Vorfaktor 2^{-l} ist ein Vorfaktor, der von der Skalierung der Basisfunktionen abhängt.

Wir betrachten zunächst die Approximation von Funktionen in Sobolev-Räumen. Dies sieht man am besten, wenn man den Sobolev-Raum über die globale Fourier-Reihe definiert.

Definition 2.3.2 *Eine Funktion u ist genau dann im Sobolev-Raum \mathcal{H}^s , wenn gilt:*

$$\|u\|_{\mathcal{H}^s} := \left(\sum_{\mathbf{k}} |\hat{u}_{\mathbf{k}}|^2 \cdot (1 + |\mathbf{k}|_{\infty})^{2s} \right)^{1/2} < \infty, \quad u(\mathbf{x}) = \sum_{\mathbf{k}} \hat{u}_{\mathbf{k}} \cdot e^{i\mathbf{k} \cdot \mathbf{x}} \quad (2.20)$$

Die Fourier-Koeffizienten fallen bekanntlich mit zunehmender Frequenz umso schneller ab, je glatter die Funktion ist. Ist die lokale Glattheit über das gesamte Gebiet $\bar{\Omega}$ in etwa gleich, so wird das Gitter, das die beste Approximation liefert, voll sein. Wie zum Beispiel in deVore ([12]), bezeichnet man die oben eingeführte Art als lineare Approximation, da die Best-Approximation dadurch entsteht, dass u mit Hilfe des Skalarprodukts orthogonal auf den Raum der Ansatzfunktionen projiziert werden kann.

Wir bezeichnen nun mit der Indexmenge $I_L^{(\infty)} := \{(\mathbf{l}, \mathbf{i}) : |\mathbf{l}|_{\infty} = L\}$ die Menge der Indizes auf dem vollen Gitter des Levels L .

Satz 2.3.3 *Sei $u \in \mathcal{H}^s$ und sei $0 \leq t < s \leq N$, wobei N die Anzahl der verschwindenden Momente der Wavelets bezeichnet. Dann gilt mit einer Konstanten C*

$$\inf_{\{u_{\mathbf{l},\mathbf{i}}\}_{(\mathbf{l},\mathbf{i}) \in I_L^{(\infty)}}} \left\| u - \sum_{(\mathbf{l},\mathbf{i}) \in I_L^{(\infty)}} u_{\mathbf{l},\mathbf{i}} \phi_{\mathbf{l},\mathbf{i}} \right\|_{\mathcal{H}^t} \leq C \cdot \left| I_L^{(\infty)} \right|^{(t-s)/d} \|u\|_{\mathcal{H}^s}. \quad (2.21)$$

Analog gilt dies auch in höheren Dimensionen für anisotrope Wavelets (siehe Koster ([22])). Auf der rechten Seite geht dann allerdings noch die Menge der Punkte in einem logarithmischen Faktor ein. Bungartz ([6]) hat die Abschätzungen für die hierarchische Basis in der L_2 -Norm berechnet. Hier kommt dann ein weiterer logarithmischer Faktor hinzu.

Wir haben also gesehen, wie die Approximation mit Wavelet-Basen auch auf Sobolev-Räume übertragbar ist. Natürlich geraten wir in Schwierigkeiten, wenn wir uns auf schwache Ableitungen beschränken, denn bei der Methode der Finiten Differenzen nehmen wir Punktauswertungen vor. Sobald eine Funktion nur noch in einem der L_p -Räume lebt, gelten die Aussagen nur noch "fast überall".

Wenn wir, wie wir später sehen werden, jedoch Finite Differenzen auf Funktionen anwenden, die an endlich vielen Stellen nicht differenzierbar sind, so entstehen Punktsingularitäten in den Ableitungen, die über den Differentialoperator einen Pollutionseffekt auf den Gesamtfehler haben können. Dabei kann die Konvergenzrate erheblich verschlechtert werden, im schlimmsten Falle sogar Divergenz entstehen.

Mit den Sobolev-Räumen messen wir die Glattheit über dem gesamten Raum. Wir sind hier aber speziell daran interessiert, die endlich vielen Stellen lokal in den Griff zu bekommen. Oder anders ausgedrückt: Wir möchten ein Maß für die Glattheit finden, das die lokalen Unglattheiten nur schwach gewichtet, um bestimmen zu können, ob und mit welcher Rate eine Funktion approximierbar ist. Dieses Maß wird vor allem für die Fehlerabschätzungen im Falle von adaptiven Verfahren wichtig sein, denn hier sollen gerade wenig Punkte in glatten Bereichen verwendet werden, während man unglatte Stellen besser auflöst.

Wie zum Beispiel in DeVore ([12]) zu finden ist, sind Besov-Räume die geeigneten Räume für Aussagen über adaptive Verfahren.

Definition 2.3.3 *Eine Funktion u ist genau dann in einem Besov-Raum $\mathcal{B}_{p,q}^s$, wenn*

$$\|u\|_{\mathcal{B}_{p,q}^s} := \|u\|_{L_p} + \left(\int_{\mathbb{R}^d} |\mathbf{h}|^{-(qs+d)} \|\Delta_{\mathbf{h}}^m u\|_{L_p}^q d\mathbf{h} \right)^{1/q} < \infty, \quad s < m \in \mathbb{N}, \quad (2.22)$$

wobei Δ den Differenzen-Operator $\Delta_{\mathbf{h}}$ in Richtung $\mathbf{h} \in \mathbb{R}^d$

$$\Delta_{\mathbf{h}}^m u(\mathbf{x}) := u(\mathbf{x} + \mathbf{h}) - u(\mathbf{x})$$

bezeichnet.

Der Sobolev-Raum \mathcal{H}^s kann auch über die $\mathcal{B}_{2,2}^s$ -Norm definiert werden. Insofern stellen Besov-Räume eine Verallgemeinerung der Sobolev-Räume dar. Nun erhalten wir eine Abschätzung für isotrop verfeinerte Wavelets:

Satz 2.3.4 *Seien die isotrop verfeinerten Wavelets $\phi_{\mathbf{1},i}$ für $r \in]-\gamma, \gamma[$ bezüglich des Sobolev-Raumes \mathcal{H}^r -stabil. Die dualen Wavelets $\tilde{\phi}$ haben N verschwindende Momente. Gelte für die Parameter $t \in]-\gamma, \gamma[$ und $t \leq s \leq N$ und sei $\phi_{\mathbf{1},i} \in \mathcal{B}_{\tau,\tau}^r$ für ein $r > s$ und $\tau = ((s-t)/d + 1/2)^{-1}$. Dann sind äquivalent:*

$$(i) \quad u \in \mathcal{B}_{\tau,\tau}^s \quad (2.23)$$

$$(ii) \quad \sum_{(\mathbf{1},i)} \left(2^{|\mathbf{l}|\infty \cdot t} \left| \left\langle \tilde{\phi}_{\mathbf{1},i}, u \right\rangle \right| \right)^\tau < \infty \quad (2.24)$$

$$(iii) \quad \sum_{m=1}^{\infty} \left(2^{m \cdot (s-t)/d} \sigma_{2^m, \mathcal{H}^t}^{iso}(u) \right)^\tau < \infty \quad (2.25)$$

Das Theorem besagt in Aussage ()), dass die Koeffizienten der Wavelet-Darstellung von u abfallen müssen, wenn u im entsprechenden Besov-Raum lebt. Es ist hier sogar möglich, die Koeffizienten für eine gegebene Funktion nachzurechnen und so die Norm und den Approximationsfehler bei isotroper Verfeinerung zu bestimmen.

Beispiel 2.3.1 Punktsingularität

Koster ([22]) zeigt, dass die d -dimensionale Funktion $u(\mathbf{x}) = |\mathbf{x}|^\beta$ für beliebig kleines $\epsilon > 0$ im Raum $\mathcal{H}^{d/2+\beta-\epsilon}$ liegt. Ferner wird dort gezeigt, dass u für $\beta > -d/2$ in jedem Besov-Raum $\mathcal{B}_{\tau,\tau}^s$ liegt und daher die Konvergenzordnung nur von der L_2 -Norm der Wavelets abhängt. Sie kann mit $\sigma_{n, \mathcal{H}^t}^{iso}(u) \leq C \cdot n^{N/d} \|u\|_{\mathcal{B}_{\mathcal{H}^0}^N}$ beziffert werden.

Konvergenzordnungen, die man für glatte Funktionen erhält, können also auch auf bestimmte nichtglatte Funktionen übertragen werden, wenn adaptive Verfahren verwendet werden. Genauere Ausführungen zu den Funktionenräumen finden sich in DeVore ([12]).

2.3.2 Verallgemeinerung der Ansatzfunktionen

Bisher haben wir als Ansatzfunktionen lediglich die hierarchische Basis verwendet. Es gibt allerdings eine ganzen Reihe von Funktionen, die sich für Approximationen eignen. Wir wollen im folgenden eine Reihe von Funktionen kurz beschreiben und uns dann mit Interpolets beschäftigen, die eine Grundlage des Codes bilden, der für die Ergebnisse dieser Arbeit verwendet wurde. Die Ansatzfunktionen unterscheiden sich in der Glattheit und in der Größe des Trägers. Wichtig für die Approximation ist dabei, bis zu welcher Ordnung Monome (und daher auch Polynome) exakt approximiert werden können. Die Konvergenzordnungen für glatte Funktionen sind dann auf einfache Weise aus der Taylor-Reihenentwicklung abzulesen. Hier betrachten wir zwei Arten von Anfangsfunktionen mit beliebig hoher polynomialer Exaktheit:

Hierarchische Lagrange-Interpolation

Lagrange-Interpolation ist die einfachste Art, die hierarchische Basis zu verallgemeinern. Man gibt sich einen Polynomgrad p vor, mit dem eine 1D-Basisfunktion bestimmt wird. Verwendet man das Tensorprodukt, so läßt sich die multivariate Ansatzfunktion errechnen, wobei der Polynomgrad in jede Richtung unterschiedlich sein kann. Interessant ist nun, dass man hier die hierarchische Struktur der Multiskalenanalyse nutzen kann, um die Ansatzfunktionen in einem Punkt zu fixieren. Um ein Polynom vom Grad p zu erzeugen, braucht man $p + 1$ Bedingungen. Diese werden von den hierarchischen Vorgängern geliefert, da diese Knoten ja bereits eingeführt wurden. Man braucht also, um ein Polynom vom Grad p zu erzeugen, ein Mindestlevel von $l = p + 1$. Dann aber benötigt man weiterhin nur einen Knoten pro Freiheitsgrad.

Die Abschätzungen, die man durch die Erhöhung des Polynomgrades erhält, spiegeln in erster Linie das Wegfallen der ersten $p + 1$ Terme der Taylor-Reihe wider. Dies gilt natürlich nur unter der Voraussetzung, dass u genügend glatt (in diesem Falle $p + 1$ mal stetig differenzierbar) ist.

Interpolets

Interpolets entstehen aus einem Interpolationsprozess, der erstmals in der Arbeit von Deslauriers und Dubuc ([11]) auftaucht. Seien Daten $w(s)$ für $s \in \mathbb{Z}$ gegeben. Wir suchen eine Funktion $y : \mathbb{R} \rightarrow \mathbb{R}$, die möglichst glatt ist und die Funktionswerte von w an den gegebenen Stellen erfüllt. Man nimmt $2k - 1$ gegebene Werte und bildet das Lagrange-Polynom. Aus dem gewonnenen Polynom $p(x)$ liest man die Werte an den Stellen $\mathbb{Z} + \frac{1}{2}$ ab und verwendet sie

für den nächsten Interpolationsschritt, der aus den jeweils $2k - 1$ Werten um s die Werte an den Stellen $\mathbb{Z} + \frac{1}{4}$ ermittelt und als Werte für y übernommen. Iteriert man dieses Verfahren, erhält man die gesamte Funktion y , die rein aus den diskreten Anfangsdaten interpoliert wurde. Da die Funktion linear von den Anfangsdaten abhängt, läßt sich eine Fundamentalfunktion $F : \mathbb{R} \rightarrow \mathbb{R}$ finden, die das Verfahren vereinfacht. Es gilt dann nämlich:

$$y(x) = \sum_{s \in \mathbb{Z}} w(s) F(x - s). \quad (2.26)$$

Die Interpolet-Mutterfunktion ist nun schlicht die Funktion F , also $\phi \equiv F$ und kann dann nach Dilatation und Translation wieder wie oben verwendet werden. Die Mutterfunktion besitzt eine Glattheit, die offensichtlich von k abhängt. Sie wird, wie wir gleich sehen werden, dementsprechend auch den Grad der polynomialen Exaktheit bestimmen.

Nun wenden wir uns den Eigenschaften der Interpolets zu.

Satz 2.3.5 Skalierungsgleichung *Die Funktion ϕ erfüllt die Skalierungsgleichung*

$$\phi(x) = \sum_{s \in \mathbb{Z}} h_s \phi(2x - s), \quad (2.27)$$

wobei die Koeffizienten gegeben sind durch $h_0 = 1$ und $h_s = h_{-s}$ ($s \in \mathbb{Z}$) und

$$\begin{pmatrix} 1^0 & 3^0 & \dots & (2k-1)^0 \\ 1^2 & 3^2 & \dots & (2k-1)^2 \\ \vdots & \vdots & \vdots & \vdots \\ 1^{2k-2} & 3^{2k-2} & \dots & (2k-1)^{2k-2} \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ h_3 \\ \vdots \\ h_{2k-1} \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Alle weiteren Koeffizienten sind 0.

Träger *Wie unter Punkt 1 gesehen, ist der Träger kompakt und beschränkt auf das Intervall $\text{supp } \phi =] -2k + 1, 2k - 1[$.*

Interpolationseigenschaft *Die Funktion ϕ erfüllt $\phi(s) = \delta(s)$, ($s \in \mathbb{Z}$) ($\delta(\cdot)$ beschreibt das Dirac-Funktional) und ist daher interpolierend in den diskreten Werten \mathbb{Z} . Daraus folgt, dass sich die Monome für $0 \leq i \leq 2k$ auf einfache Weise darstellen lassen:*

$$x^i = \sum_{s \in \mathbb{Z}} s^i \phi(x - s). \quad (2.28)$$

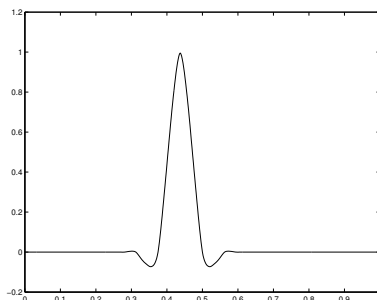


Abbildung 2.6: Interpolet auf Level 5.

Polynomiale Exaktheit *Wie gerade gesehen, ist der Grad der polynomialen Exaktheit $2k$.*

Die Interpolets lassen sich auf einfache Weise auf das Intervall einschränken. Für genauere Ausführungen siehe Koster ([22]). Die dann entstandenen Interpolets auf dem Intervall lassen sich über das gewöhnliche und mehrfach erwähnte Tensorprodukt auf beliebige Dimensionen erweitern.

Wir werden nun auf die Approximationseigenschaften der Interpolets eingehen. Sie unterscheiden sich nicht wesentlich von denen der Lagrange-Polynome. Ausschlaggebend für die Approximation ist vor allem der Grad an polynomialer Exaktheit, der sowohl bei Interpolets wie auch bei durch Lagrange-Interpolation verallgemeinerter hierarchischer Basis selbst gewählt werden kann.

Wie man bei der Definition der Interpolets bereits ahnt, ergeben sich für glatte Funktionen bessere Approximationseigenschaften, wenn man Interpolets als Ansatzfunktionen wählt. Daher werden wir die folgenden Abschätzungen unter stärkeren Annahmen treffen müssen, bekommen dafür aber auch bessere Approximationen. Zu überlegen ist hier auch, ob sich — je nach Aufgabenstellung — ein Verfahren lohnt, das auf einer Art h-p-Methode basiert. Der Grad polynomialer Exaktheit ist in unserem Falle zwar wählbar, wird allerdings vom Verfahren selbst nicht justiert. Die Approximationseigenschaften werden im ersten Lemma zunächst für den eindimensionalen Fall vorgestellt und später wieder über die Tensorproduktstruktur auf beliebige Dimensionen erweitert.

Lemma 2.3.2 Abfall hierarchischer Überschüsse

Sei nun $u \in \mathcal{C}^{N+1}([0, 1])$, $1 \leq N \leq l$. Dann gilt für den hierarchischen Überschuss der Interpolets mit polynomialer Exaktheit $N = 2n$ und einer

Zwischenstelle $\xi \in [x_{l,i} - (n-1)h_l, x_{l,i} + n \cdot h_l]$

$$v^I(x_{l,i}) := u(x_{l,i}) - u^I(x_{l,i}) = \frac{1}{N!} \cdot D^N u(\xi) \cdot \prod_{k=1}^N (x_{l,i} - x_{l,i-N+2k-1}). \quad (2.29)$$

Damit hat die Approximation mit Interpolets die Ordnung

$$|v^I(x_{l,i})| \leq \frac{1}{N!} \cdot |D^N u(\xi)| \cdot h_l^N \cdot 2^{N \cdot (N-1)/2-1}. \quad (2.30)$$

Beweis: Man setzt die verwendeten Punkte auf das höchste Level und nimmt die ungeraden Indizes. Eine ausführliche Fassung kann in Bungartz und Griebel ([8]) gefunden werden. \square

Die Verallgemeinerung auf den mehrdimensionalen Fall funktioniert analog zu der Approximation höherer Ordnung, die meist in der Literatur zu finden ist.

Lemma 2.3.3 Sei nun $u \in X_0^{\mathbf{N},d}$. Dann gelten die folgenden Abschätzungen für die hierarchischen Überschüsse:

$$|v_{1,i}^I| \leq \left(\frac{1}{2}\right)^d \cdot \prod_{j=1}^d \frac{2^{N_j \cdot (N_j-1)/2}}{N!} \cdot 2^{-|\mathbf{l} \cdot \mathbf{N}|_1} \cdot |u|_{\mathbf{N},\infty} \quad (2.31)$$

$$|v_{1,i}^I| \leq \left(\frac{1}{6}\right)^{d/2} \prod_{j=1}^d \frac{2^{N_j \cdot (N_j-1)/2}}{N!} 2^{-|\mathbf{l} \cdot \mathbf{N}|_1 + |\mathbf{l}|_1/2} \cdot \|D^{\mathbf{N}} u(\mathbf{x})|_{\text{supp } \phi_{1,i}}\|_2. \quad (2.32)$$

Lifting-Interpolets

Die Interpolets sind bereits in ihrer eindimensionalen Form nicht für alle Werte von k stabil bezüglich \mathcal{H}^1 , weil die dualen Pendanten zu den Interpolets Dirac-Funktionale sind. Damit theoretische Aussagen getroffen werden können, ist es jedoch von Vorteil, wenn diese Eigenschaft erfüllt ist. Nun ist es jedoch möglich, ein Lifting-Schema auf die Interpolets anzuwenden, das sie stabil werden lässt. Wir werden die Nützlichkeit der Lifting-Interpolets später sehen, wenn wir unterschiedliche Verfeinerungskriterien für den adaptiven Algorithmus testen. Zusätzlich werden wir die Lifting-Interpolets zum Vorkonditionieren des Gleichungssystems verwenden. Bei der Diskretisierung des Differentialoperators beschränken wir uns jedoch auf Interpolets.

Aus einer Multiskalenbasis $\phi_{1,i}$ von Interpolets lassen sich die Lifting-Interpolets $\bar{\phi}_{1,i}$ dadurch gewinnen, dass sie aus unterschiedlicher Leveln linear kombiniert werden.

Definition 2.3.4 Mit $I_l := \{i : 0 \leq i \leq 2^l - 1\}$ definiert man die eindimensionalen Lifting-Interpolets $\bar{\phi}_{l,i}$ als

$$\bar{\phi}_{l,i} := \begin{cases} \phi_{s,i} + \sum_{s \in I_{l-1}} K_{s,i}^l \phi_{(l-1,s)} & l > l_0 \\ \phi_{(l,i)} & l = l_0 \end{cases}, \quad (2.33)$$

wobei die Koeffizienten $K_{s,i}^l$ so gewählt werden, dass die entstehenden Lifting-Interpolets \bar{M} verschwindende Momente haben.

2.4 Adaptive Verfeinerung

Adaptive Verfahren wurden entwickelt, um ein besseres Verhältnis von Aufwand zu Genauigkeit zu erhalten. Dies ist insbesondere der Fall, wenn die Funktion nicht sehr glatt ist. Wie bereits oben angesprochen, kann man die gleiche Approximationsrate für nicht glatte Funktionen wie für glatte Funktionen erhalten, wenn man adaptive Verfahren verwendet. Im Falle von sehr glatten Funktionen (gemeint ist: kleine Ableitungen) wird das uniforme Gitter ideal sein. Gerade aber im Falle von Punkt- oder Kantensingularitäten bei ansonsten glatter Funktion kann die Approximationsrate wie bei regulärer Verfeinerung erreicht werden. Wir haben dies bereits im Abschnitt 2.3.1 über allgemeine Funktionenräume angesprochen. Funktionen mit niedriger Sobolev-, aber hoher Besovregularität können mit Hilfe von adaptiven Verfahren (annähernd) gleich gut approximiert werden.

Bevor wir auf die Details der Adaptivität mit Interpolets bzw. Lifting-Interpolets zu sprechen kommen, beschreiben wir zunächst, was wir mit einem adaptiven Verfahren meinen und wo der Unterschied zur Best-n-Term-Approximation liegt: Zunächst gibt es die Möglichkeit, regulär zu verfeinern. Diese haben wir oben bereits diskutiert. Die optimale nichtlineare Approximation mit gegebenen Ansatzfunktionen besteht darin, das Gitter, das heißt den Träger der Ansatzfunktionen, so anzupassen, dass der Fehler auf dem Gebiet minimal wird. Dies führt direkt auf die Best-n-Term Approximation. Die dritte Möglichkeit der Approximation wollen wir uns nun ansehen — nämlich die adaptive Verfeinerung. Ziel ist es einen Algorithmus zu finden, der aus einem gegebenem Fehlerschätzer, einer Fehlertoleranz ε und einer Verfeinerungsstrategie (hier: dyadisch — analog zu der regulären Verfeinerung) ein Gitter liefert, das dem der Best-n-Term Approximation nahe kommt. Der adaptive Algorithmus sieht dann wie folgt aus.

Jeder Knoten erhält ein Flag, das anzeigt, ob er zu der Menge der "guten" oder der "schlechten" Intervalle (\mathcal{G} bzw. \mathcal{B}) gehört. In jedem Schritt wird geprüft, ob der Fehler an einem Knoten größer als ε ist. Ist das Krite-

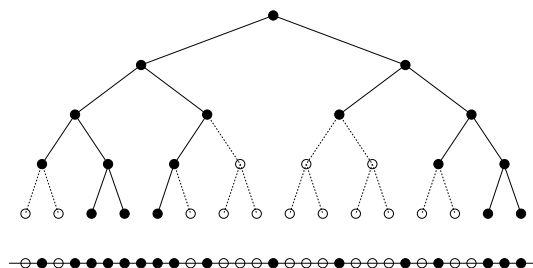


Abbildung 2.7: Baumdarstellung des adaptiven Gitters in einer Dimension

rium für alle oder zumindest hinreichend viele Knoten erfüllt, terminiert der Algorithmus. Ein genereller Schritt des Algorithmus sieht wie folgt aus:

Gegeben sei ein Intervall $\mathcal{I} \in \mathcal{B}$. Man sucht nun die zwei Söhne \mathcal{J}_1 und \mathcal{J}_2 und prüft, ob der Fehlerschätzer für sie ein Ergebnis kleiner als das Fehlerkriterium ε liefert. Ist das Fehlerkriterium für \mathcal{J}_1 erfüllt, so wird dieses Intervall der Menge der "guten" Intervalle \mathcal{G} hinzugefügt, andernfalls der Menge der "schlechten" Intervalle \mathcal{B} . Analog verfährt man mit \mathcal{J}_2 .

Diese Prozedur wird so lange durchlaufen, bis keine (oder hinreichend wenige) Intervalle in der Menge \mathcal{B} verbleiben. Graphisch lässt sich die adaptive Methode für Multilevel-Methoden anhand einer Baumstruktur verdeutlichen, bei der nur die Äste in dicken Linien weitergeführt werden, deren Fehlerbeitrag zu groß ist (siehe Abbildung 2.7). Die Äste, die auf einem Level abbrechen, erfüllen auf diesem Level zum ersten Mal das Fehlerkriterium.

Entscheidend für ein gutes Verfahren ist ein guter Fehlerschätzer. Der einfachste Fehlerschätzer ist, den Absolutwert des hierarchischen Koeffizienten zu verwenden. Die Abschätzung über das Abfallverhalten stellt das Terminieren des Algorithmus sicher. Sind jedoch die Voraussetzungen für das Abfallen der Koeffizienten, nämlich beschränkte gemischte zweite Ableitungen, nicht erfüllt, kann das Verfahren eventuell das Fehlerkriterium nie erfüllen.

Beispiel 2.4.1 Sei $f : [0, 1] \rightarrow \mathbb{R}$. Dann fallen die hierarchischen Koeffizienten für die Funktion

$$f(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1/2 \\ 2 & \text{für } 1/2 \leq x \leq 1 \end{cases} \quad (2.34)$$

links des Sprunges nicht ab, sondern bleiben konstant bei 1, unabhängig vom jeweiligen Level.

Natürlich kann man andere Fehlerschätzer bauen, wie zum Beispiel gewichtete Normen. Das Problem kann allerdings nicht überwunden werden:

zu jedem noch so aufwendig gebauten Fehlerschätzer gibt es eine Funktion, für die der Fehlerschätzer versagt.

Es seien hier noch ein paar Arten der Fehlerschätzer genannt, die später auch im Code verwendet werden können. Ein denkbar einfacher Fehlerschätzer wäre es, die 1-Norm des Index $|\mathbf{I}|_1$ zu verwenden. Diese führt dann direkt wieder auf ein dünnes Gitter vom Level $|\mathbf{I}|_1 \leq \varepsilon$. Insofern stellen adaptive Verfeinerungen eine Verallgemeinerung regulärer dünner Gitter dar. Als Fehlerschätzer kann man ansonsten neben dem Absolutwert des hierarchischen Koeffizienten alle möglichen Normen, wie die diskrete L_2 -, \mathcal{H}^1 - oder \mathcal{H}^s -Norm, heranziehen.

In dem Code, der dieser Arbeit zugrunde liegt, können für die Adaptivität zwei verschiedene Arten von Koeffizienten herangezogen werden. Zum einen kann die Funktion in einer Interpolet-Basis dargestellt werden, zum anderen allerdings auch mit den zugehörigen Lifting-Interpolets. Dabei können durchaus unterschiedliche Ergebnisse für die Konvergenz erzielt werden, wie wir im Abschnitt über die numerischen Ergebnisse sehen werden.

Es gibt verschiedene Arten von adaptiven Verfeinerungen in einer Multiskalenumgebung. Man kann einerseits uniform verfeinern, oder im anderen Extrem rein mit dünnen Gittern. Dazwischen liegt aber auch noch ein Verfahren, das es erlaubt, nicht alle Punkte zu verwenden, sondern sich bestimmte Unterräume auszuwählen. Wie wir später sehen werden, erlaubt diese Methode die Voraussetzungen an die Funktion so abzuschwächen, dass Singularitäten mit weniger Aufwand als der uniformen adaptiven Verfeinerung gerechnet werden können, die mit dünnen Gittern nicht approximierbar sind.

2.4.1 Rechnen auf adaptiven dünnen Gittern

Operiert man mit dünnen Gittern, so lässt es sich nicht vermeiden, dass Funktionen, die auf unterschiedlichen adaptiven Gittern leben, miteinander verknüpft werden. Die einfachsten Operationen sind dabei Addition und Subtraktion. Sie treten bereits auf, wenn man einen Differentialoperator erstellt, bei dem man auf adaptive Gitter zugreift, die anhand unterschiedlicher Funktionen erstellt wurden. Wir untersuchen beispielhaft den einfachen Fall der Addition auf unterschiedlichen adaptiven dünnen Gittern. Wir beschränken uns auf den Fall der eindimensionalen Funktionen, da auch hier das Tensorprodukt erneut herangezogen werden kann, um den multivariaten Fall analog zu beschreiben.

Lemma 2.4.1 *Gelte für Dünngitterapproximationen $u_1^{\varepsilon_1}$ an die Funktion u_1 und $u_2^{\varepsilon_2}$ an die Funktion u_2 $\|u_1 - u_1^{\varepsilon_1}\| \leq \eta_1$ bzw. $\|u_2 - u_2^{\varepsilon_2}\| \leq \eta_2$. Dann gilt*

für $u_3(x) := u_1(x) \pm u_2(x)$ die Abschätzung

$$\|u_3 - (u_1^{\varepsilon_1} \pm u_2^{\varepsilon_2})\| \leq \varepsilon_1 + \varepsilon_2. \quad (2.35)$$

Der Beweis des Lemmas ist unmittelbar klar. Wie Multiplikation und Division richtig durchgeführt werden, ist in Schiekofer ([26]) beschrieben. Der einfachste Weg ist dabei aber meist die Rücktransformation auf die nodale Basis.

Kapitel 3

Optionen

3.1 Einführung

Der Durchbruch für den Optionshandel fand im Jahre 1973 statt — nicht nur, dass Optionen in diesem Jahr zum ersten Male an der Börse gehandelt wurden (eingeführt am Chicago Board Options Exchange), auch das Geheimnis ihrer Bewertung unter der Annahme eines ”perfekten Marktes” wurde entschlüsselt. Zeitgleich arbeiteten Fischer Black und Myron Scholes sowie Robert Merton an Modellen, die den Wert einer Option bestimmen. Die Aussagen sind so fundamental, dass sie bis heute Bestand haben und die Methodik noch heute in beinahe allen Artikeln über Optionsbewertungen auftaucht.

Wie wir sehen werden, ist die Grundidee recht einfach. Der Preis der Option zum Endzeitpunkt ist für alle Werte des zugrunde liegenden Wertpapiers bekannt. Da der Preis der Option gegeben einem Aktienwert an diesem Tag bekannt ist, kann der Wert der Option unter Annahme einer stochastischen Bewegung als der Erwartungswert über die diskontierte Auszahlung berechnet werden. Die Feynman-Kac Formel liefert dann den Zusammenhang zwischen dem durch Integration berechneten Erwartungswert und einer partiellen Differentialgleichung. Wir wollen uns in dieser Arbeit dem Lösen der partiellen Differentialgleichung widmen, wie wir unten noch genauer ausführen werden.

Die numerische Simulation des Modells erfordert das Eingrenzen des Gebietes, das als der Raum \mathbb{R}^n modelliert wurde. Um eine eindeutige Lösung zu erhalten, müssen an den durch das ”Abschneiden” generierten künstlichen Rändern neue Randbedingungen eingeführt werden. Es gibt dabei zahlreiche Kandidaten, die hierfür in Frage kommen. Wir werden uns im Abschnitt 3.5 damit eingehender befassen. Zunächst soll dort gezeigt werden, dass im

parabolischen Fall — in dem wir uns bewegen — der Fehler im Innern des Gebietes durch den Fehler auf dem Rand kontrolliert werden kann. Die Lösung des numerischen Problems kann also nicht beliebig entarten, wenn künstliche Randbedingungen eingeführt werden.

3.1.1 Grundbegriffe und verschiedene Arten von Optionen

Definition 3.1.1 (Option) *Eine Call- (Put-)Option ist das Recht, das dem Halter einräumt, einen zugrundeliegenden Basiswert an einem Zeitpunkt in der Zukunft (Ausübungszeitpunkt) zu einem festgelegten Preis zu kaufen (oder verkaufen).*

Dabei unterscheiden sich Optionen in der Art der zugrundeliegenden Werte, die frei wählbar sind. Am häufigsten werden Aktien, Währungen und Rohstoffe zugrunde gelegt, wobei auch ausgefallener Basisgrößen wie Auszahlungen bei bestimmten Wetterlagen (z.B. Regentage, Sonnenstunden, Lufttemperatur oder Windgeschwindigkeit, siehe Auer ([1])) in Frage kommen. Die Wahl dieser Basisgrößen ist insofern nicht verwunderlich, als Optionen dazu dienen, Risiken durch Schwankungen in den Basiswerten durch sogenanntes "Hedging" zu minimieren, wie wir später noch sehen werden.

- **Europäische Optionen:** Ausübungszeitpunkt und -preis werden beim Schreiben der Option festgelegt und sind fester Bestandteil des Vertrages.
- **Amerikanische Optionen:** Es wird ein Ausübungszeitraum festgelegt, innerhalb dessen der Halter die Option zu einem festen Preis ausüben kann.
- **Asiatische Optionen:** Der Ausübungspreis hängt von dem Durchschnitt der Preise im vorangegangenen Zeitraum ab.
- **"Barrier Options":** Die Option wird ausgelöst, wenn eine bestimmte obere oder untere Schranke im Basiswert erreicht ist.
- **"Lookback Options":** Der Ausübungspreis hängt vom Maximum oder Minimum der vergangenen Preise ab.

Um die Begrifflichkeiten für die spätere Modellbildung einzuführen, betrachten wir nun zwei Definitionen, die sich mit dem Handel auf Finanzmärkten beschäftigen.

Definition 3.1.2 (Kurze Positionen) *Man geht bezüglich eines Wertpapiers "kurz" (engl. "short selling"), wenn man Wertpapiere verkauft, die man (noch) nicht besitzt. Auf Finanzmärkten ist dies — in einem gewissen rechtlichen Rahmen — erlaubt. Das Portfolio enthält demnach eine "negative Anzahl" an Wertpapiereinheiten.*

Definition 3.1.3 (Arbitrage) *Arbitrage ist das risikolose Ausnutzen unterschiedlicher Preise für ein und dasselbe Gut auf unterschiedlichen Märkten.*

3.1.2 Betrachtung von Auszahlungen

Um ein Gefühl für die Bewertung von Optionen zu bekommen, wenden wir uns einem Beispiel zu. Hier wird die Wirkungsweise von Optionen anhand einer europäischen Call-Option verdeutlicht.

Beispiel 3.1.1 *In diesem Beispiel soll unter Vernachlässigung der Zinsen veranschaulicht werden, wie die Bewertung von Optionen in seinen Grundzügen vorgenommen werden kann.*

Seien der Ausübungszeitpunkt genau in zwei Monaten und der Ausübungspreis auf 100€ festgelegt. Sei weiterhin die Wahrscheinlichkeit jeweils 50%, dass das Wertpapier zum Ausübungszeitpunkt entweder 80€ oder 120€ wert ist. In letzterem Falle kann der Halter die Option auslösen und sofort wieder verkaufen. Er erwirtschaftet damit einen Gewinn von 120€-100€ = 20€. Im anderen Falle wird der Halter die Option nicht auslösen, so dass der Erwartungswert seines Gewinns bei

$$\frac{1}{2} \times 0€ + \frac{1}{2} \times 20€ = 10€$$

liegt. Auch wenn dieses Modell nicht die Wirklichkeit abbildet, wird der Preis doch etwa bei 10€ liegen.

Was wir in dem Beispiel sehen konnten, war der Hebeleffekt, den die Option auf Gewinnschwankungen hat. Nehmen wir an, die Aktie sei heute 100€ wert. Dann liegt die Gewinnspanne bei $\pm 20\%$, also 20%. Die Option hingegen hat eine Gewinnspanne von 100%.

Mit dem Beispiel 3.1.1 erschließen wir leicht die Auszahlungsfunktion für europäische Optionen im kontinuierlichen Fall zum Zeitpunkt T . Der Halter einer Call-Option wird die Option nur einlösen, wenn der Ausübungspreis E niedriger als der aktuelle Marktwert ist. Dann ist sein Gewinn in diesem Zeitpunkt gleich der Differenz des Ausübungspreises zum aktuellen

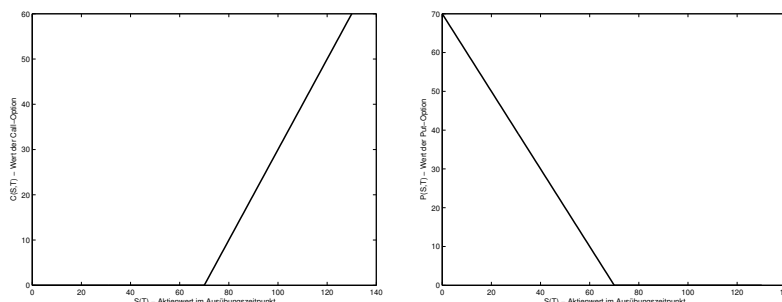


Abbildung 3.1: Auszahlungsfunktion zum Ausübungszeitpunkt für eine Call-Option (links) und eine Put-Option (rechts).

Marktwert. Anderenfalls verfällt die Option ohne Auszahlung. Wäre diese Auszahlungsfunktion nicht exakt so beschaffen, ergäben sich direkte Arbitragemöglichkeiten, die wir in einem reibungsfreien Markt nicht zulassen. Die Auszahlungsfunktion läßt sich also schreiben als

$$C(S, T) = \max(S - E, 0). \quad (3.1)$$

Die analogen Überlegungen lassen sich für Put-Optionen anstellen und führen dann zur Auszahlungsfunktion

$$P(S, T) = \max(E - S, 0). \quad (3.2)$$

Die beiden Funktionen sind in Abbildung 3.1 veranschaulicht.

Beispiel 3.1.2 (Die Put-Call Parität) Die Put-Call Parität ist ein Beispiel, in dem Put- und Call-Optionen in einem Portfolio so miteinander verknüpft sind, dass jegliches Risiko eliminiert wird.

Sei Π der Gesamtwert an gehaltenen Finanzanlagen, die aus einem Put, einem Call und einem Call ("short") bestehen. Put und Call haben dabei dasselbe Enddatum T und denselben Ausübungspreis E .

$$\Pi(t) = S(t) + P(t) - C(t)$$

und folglich gilt zum Endzeitpunkt T

$$\Pi(T) = S(T) + \max(E - S(T), 0) - \max(S(T) - E, 0)$$

Damit gilt: $\Pi(T) = E$.

Da der Wert zum Ausübungszeitpunkt deterministisch ist, kann das Portfolio zu jedem Zeitpunkt (bei bekannter Zinsrate) eindeutig bewertet werden. Es ergibt sich, indem der Wert im Ausübungszeitpunkt T , nämlich E , auf den Zeitpunkt t diskontiert wird. Somit:

$$\Pi(t) = S(t) + P(t) - C(t) = Ee^{-\int_t^T r(s)(T-s)ds}$$

3.1.3 Verwendung von Optionen

Optionen werden häufig entweder geschrieben, weil sie sich aufgrund des Hebeleffektes (siehe Beispiel 3.1.1) sehr gut für Spekulationen eignen, oder weil sie durch "Hedging" Risiken minimieren können.

Definition 3.1.4 (Hedging) *Hedging ist der Versuch, die Sensitivität eines Portfolios auf Änderungen eines Wertes im Portfolio zu reduzieren, indem andere Werte mit gegenläufigen Bewegungen gekauft werden.*

Diesen Vorgang werden wir anhand eines Beispiels sehen:

Beispiel 3.1.3 (Hedging) *Der "Protective Put" ist ein einfaches Beispiel für eine Hedging-Strategie: Investoren haben häufig das Bedürfnis, sich gegen niedrige Aktienkurse abzusichern. Dazu können andere Aktien gekauft werden, die negativ mit dem Portfolio korreliert sind. Dies ist allerdings mit Risiko verbunden, da die Korrelationen nicht immer bekannt sind. Der "Protective Put" dagegen sichert den niedrigen Aktienkurs mit Hilfe einer Option auf diese Aktie ab. Dazu kauft man zu der Aktie $S(t)$ einen Put $P(t)$, die man im Portfolio $\Pi(t)$ hält. Warum dies eine Hedging-Strategie ist, die gegen niedrige Kurse absichert, sieht man, wenn man das Portfolio zum Ausübungszeitpunkt betrachtet*

$$\begin{aligned}\Pi(T) &= P(T) + S(T) = \max\{E - S(T), 0\} + S(T) \\ &= \max\{S(T), E\}.\end{aligned}\tag{3.3}$$

Allerdings ist diese Hedging-Strategie mit Kosten verbunden. Im Falle, dass der Aktienkurs $S(T)$ (zum Endzeitpunkt) den Ausübungspreis E übersteigt, ist das Portfolio $\Pi(T) = S(T)$ wertlos. In diesem Falle war der Kauf des Puts $P(T)$ wertlos. Die entstandenen Kosten von $P(t_0)$ bezeichnet man daher als Hedging-Kosten.

3.2 Bewertung von Optionen

In einem sehr stark simplifizierten Fall haben wir die Bewertung einer Option bereits gesehen (Beispiel 3.1.1). Der faire Wert einer Option kann als ein abdiskontierter Erwartungswert über den Gewinn zum Ausübungszeitpunkt aufgefasst werden. Um einen Erwartungswert berechnen zu können, muss ein stochastischer Prozess angenommen werden (in unserem Beispiel war die Wahrscheinlichkeit für jedes Ereignis $1/2$). Ist nun auch noch die Zinsrate bekannt, kann der faire Wert einer Option berechnet werden. Damit der faire Wert einer Aktie in dem Modell auch wirklich angenommen wird, wird ein

perfekter Markt — also insbesondere die Abwesenheit von Arbitragemöglichkeiten — vorausgesetzt.

Empirisch hat sich gezeigt, dass sich der Aktienmarkt sehr gut durch eine geometrische Brownsche Bewegung beschreiben lässt. Theoretisch passt dieses Modell für den Finanzmarkt zur Annahme über effiziente Märkte. Es kann nämlich gezeigt werden, dass sich ein perfekter Finanzmarkt wie ein Markov-Prozess verhält. Diese Aussage ist als "Efficient Market Hypothesis" — die Hypothese über effiziente Märkte — bekannt.

Das Modell, das die gerade beschriebenen Eigenschaften besitzt, wurde von Fischer Black und Myron Scholes ([5]) bzw. Robert Merton ([24]) entwickelt. Berühmt geworden ist das Modell nicht nur, weil man damit den fairen Wert von Optionen bestimmen kann und weil die Lösung eine geschlossene Form besitzt, sondern auch, weil es einer empirischen Überprüfung sehr gut standhält. Wir werden dieses Modell mit seinen Annahmen und der geschlossenen Lösung im folgenden diskutieren.

3.2.1 Bezeichnungen und Notationen

In der Literatur zu Optionsbewertungen hat sich weitestgehend eine einheitliche Notation durchgesetzt. Diese ist in diesem Abschnitt aufgelistet und sollen im folgenden verwendet werden.

S	Wert des zugrundeliegenden Wertpapiers
V	Optionswert — bezeichnet Put- oder Call-Option
C	Optionswert eines Calls
P	Optionswert eines Puts
σ	Volatilität
E	Ausübungspreis
T	Zeitpunkt, an dem die Option ausläuft
r	Zinsrate

Darüber hinaus verwendete Symbole werden im Text erklärt.

3.2.2 Annahmen

Die Annahmen decken sich mit den in der Volkswirtschaftslehre üblichen Annahmen an perfekte Märkte. Neuere Versuche, psychologische Aspekte in das Modell zu integrieren, sind nicht Gegenstand dieser Arbeit. Diese Annahmen des klassischen Black-Scholes Modells werden weitestgehend für den Rest der Arbeit verwendet.

- Die sogenannte "Hypothese über effiziente Märkte"¹ wird angenommen. Daraus folgt, dass der Preis einer Aktie als ein konstanter Trend μ mit lognormal verteiltem weißem Rauschen modelliert wird. Die zugehörige stochastische Differentialgleichung

$$\frac{dS}{S} = \mu dt + \sigma dX \quad (3.4)$$

enthält neben dem Preis S die Volatilität σ und einen Wiener Prozess² dX . Wie sich der Aktienkurs dann bewegen kann, wollen wir an 100 Durchläufen sehen, die mit einer Volatilität von $\sigma = 0.3$, Zinsrate $r = 0.1$ und $T = 1$ berechnet wurden. Die Zeitschrittweite betrug 0.01.

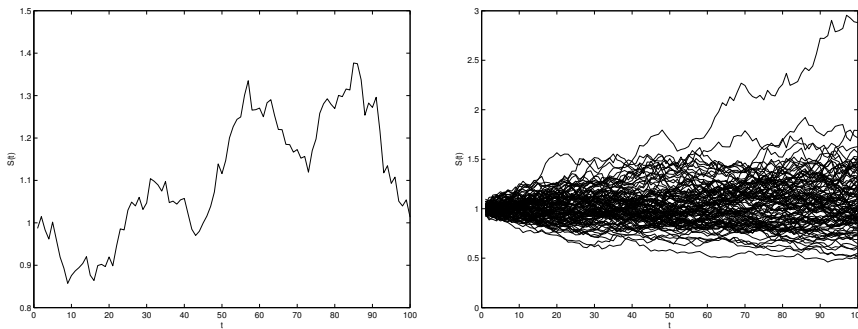


Abbildung 3.2: Realisierung einer einzigen Brownschen Bewegung (links) und nach 100 Durchläufen (rechts).

- Zinsrate und Volatilität sind vollständig bekannte konstante Funktionen in der Zeit.
- Es wird ein perfekter Markt angenommen, also insbesondere
 - gibt es keine Arbitragemöglichkeiten.
 - ist der Handel frei von Transaktionskosten.

¹Hypothese über effiziente Märkte:

- Alle Informationen, d.h. vergangene Preise und Erwartungen für die Zukunft, sind vollständig im Aktienpreis enthalten.
- Die Märkte reagieren unmittelbar auf neue Information.

Aus diesen beiden Annahmen folgt, dass der Aktienpreis durch einen Markov-Prozess beschrieben werden kann.

²Wiener Prozess:

dX ist eine normal verteilte Zufallsvariable mit Erwartungswert 0 und Varianz dt .

– existieren keine Informationsasymmetrien.

- Die Aktie kann beliebig geteilt werden.
- Es ist erlaubt, eine kurze Position zu halten (siehe Definition 3.1.2).
- Für die Halter der Aktien werden keine Dividenden ausgeschüttet.
- Der Handel der Aktien findet kontinuierlich statt.

3.2.3 Herleitung

Die zwei entscheidenden Faktoren für die Herleitung der Black-Scholes partiellen Differentialgleichung sind Itô's Lemma und die Möglichkeit, stochastische Funktionen durch geeignete Subtraktion in deterministische Funktionen zu verwandeln.

Itô's Lemma ist allgemeiner formuliert, als es hier verwendet wird. Es ist eine Art Verallgemeinerung der Taylor-Entwicklung auf stochastische Variablen.

Lemma 3.2.1 (Itô) *Sei eine stochastische Differentialform gegeben durch*

$$dG = A(G, t)dX + B(G, t)dt, \quad (3.5)$$

wobei dX die Zufallsvariable darstellt. Sei $f(G, t)$ gegeben und f zweimal stetig differenzierbar nach G und einmal stetig differenzierbar nach t . Dann gilt:

$$df = A \frac{df}{dG} dX + \left(B \frac{df}{dG} + \frac{1}{2} A^2 \frac{d^2 f}{dG^2} \right) dt. \quad (3.6)$$

Das andere Konzept ist, eine Funktion f , die zu diesem Zwecke differenzierbar sein muss, in eine deterministische Differentialform zu verwandeln.

Beispiel 3.2.1 *Sei*

$$g = f - \frac{\partial f}{\partial S} S. \quad (3.7)$$

Dann ist die Differentialform dg deterministisch, wie die folgende Rechnung zeigt:

$$\begin{aligned} dg &= \sigma S \frac{\partial f}{\partial S} dX + \left(\mu S \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \frac{\partial f}{\partial t} \right) dt - \frac{\partial f}{\partial S} (\sigma S dX + \mu S dt) \\ &= \left(\frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \frac{\partial f}{\partial t} \right) dt \end{aligned} \quad (3.8)$$

Wendet man diese beiden Konzepte auf Optionen an (siehe auch ([30])), so erhält man die Black-Scholes partielle Differentialgleichung. Wir werden hier die wichtigsten Schritte nachzeichnen.

Wir starten mit der Brownschen Bewegung aus Gleichung (3.4), die vom Typ (3.5) ist. Der Optionspreis sei dargestellt durch $V(S, t)$. An dieser Stelle müssen wir noch nicht spezifizieren, ob wir von einer Call- oder einer Put-Option sprechen, denn die Unterschiede werden sich lediglich in den Randdaten widerspiegeln. Gehen wir nun davon aus, dass der Optionswert $V(S, t)$ einmal nach t und zweimal nach S stetig differenzierbar ist. Dann sind die Voraussetzungen von Itô's Lemma erfüllt und wir können analog zu (3.6) schreiben:

$$dV = \sigma S \frac{\partial V}{\partial S} dX + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt. \quad (3.9)$$

Nun sehen wir also den stochastischen Prozess, nach dem sich der Optionspreis in unserem Modell bewegt.

Wir werden nun ein Portfolio Π konstruieren, das aus Teilen aus der Option und aus dem zugrundeliegenden Wert besteht. Der Anteil des Wertpapiers sei zunächst nicht weiter spezifiziert und mit Δ bezeichnet³

$$\Pi = V - \Delta S. \quad (3.10)$$

Π ändert sich in einem Zeitschritt um

$$d\Pi = dV - \Delta dS, \quad (3.11)$$

wobei Δ während des Schrittes konstant gehalten wird. Setzen wir nun (3.4) und (3.9) in (3.11) ein, so erhalten wir den stochastischen Prozess für das Portfolio

$$d\Pi = \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dX + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} - \mu \Delta S \right) dt. \quad (3.12)$$

Wie wir im Faktor vor dem stochastischen Anteil dX sehen, können wir das Risiko vollständig eliminieren, indem wir das Delta in jedem Zeitschritt exakt der partiellen Ableitung

$$\Delta = \frac{\partial V}{\partial S} \quad (3.13)$$

³Wir werden später sehen, warum hier der Buchstabe Δ gewählt wird. Mit der Konstruktion unseres Portfolios gehen wir einer besonderen Hedging-Strategie nach, die das Risiko eliminiert.

zu Beginn jeden Zeitschritts anpassen. Wir erhalten dann eine deterministische Bewegung für den Wert unseres Portfolios

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \quad (3.14)$$

Nun können wir das mächtige Instrument der Arbitragefreiheit anwenden, das wir in unserem Modell vorausgesetzt haben. Der Gewinn, den wir mit unserem Portfolio in einem Zeitschritt machen, ist deterministisch, wie wir gerade eben gesehen haben. Da jedoch alle Arbitragemöglichkeiten ausgeschlossen sind, muss in einem perfekten Markt ohne Transaktionskosten die Rendite aller risikolosen Positionen gleich sein. Das bedeutet, dass die Änderung des Wertes unseres Portfolios gleich dem Ertrag ist, den man erhielte, wenn man das Geld statt in unser Portfolio zu investieren zum Marktzins anlegte.

$$d\Pi = r\Pi dt = rV dt - rS \frac{\partial V}{\partial S} dt \quad (3.15)$$

Das war der entscheidende Durchbruch auf dem Weg zur gesuchten Gleichung. Fügen wir nämlich (3.15) und (3.14) zusammen, so erhalten wir die **Black-Scholes partielle Differentialgleichung**

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0. \quad (3.16)$$

3.2.4 Alternative Herleitung

Cox-Ross-Rubinstein Man kann die Black-Scholes Gleichung (3.16) auch über ein diskretes Modell herleiten. Man stelle sich den zugrunde liegenden Wert S so vor, dass er innerhalb eines Zeitschritts mit einer Wahrscheinlichkeit p um den Faktor u steigt und mit einer Wahrscheinlichkeit $1 - p$ um den Faktor d fällt. Iteriert man nun dieses Vorgehen, wobei ein Wert, der einmal steigt und einmal fällt, derselbe ist, wie wenn er einmal gefallen und dann einmal gestiegen wäre, so erhält man eine Wahrscheinlichkeitsverteilung über alle möglichen Ereignisse. Diese lassen sich in einem Binomialbaum darstellen (siehe Abbildung 3.3).

Der Optionspreis zum Zeitpunkt $T - m$ entspricht also einem m -stufigen Bernoulliexperiment. So ist auch die Formel von Cox, Ross und Rubinstein (siehe ([9])) zu interpretieren.

Satz 3.2.1 (Cox-Ross-Rubinstein) *Der Preis (bei vollständiger Arbitragefreiheit) einer Call-Option zum Zeitpunkt $T - m$ ist gegeben durch*

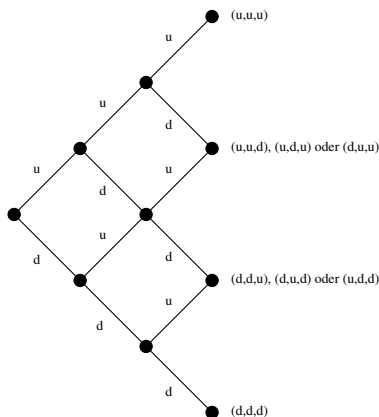


Abbildung 3.3: Binomialbaum aus dem Cox-Ross-Rubinstein Modell

folgende Formel

$$C_{T-m} = S_{T-m} \cdot \sum_{j=a}^m \binom{m}{j} q^j (1-q)^{m-j} - K(1+r)^{-m} \sum_{j=a}^m \binom{m}{j} (\hat{p})^j (1-\hat{p})^{m-j} \quad (3.17)$$

für $j = 1, \dots, T$, wobei

$$\hat{p} = \frac{1+r-d}{u-d}, \quad a = \min \{0 \leq j \leq m : u^j d^{m-j} S_{T-m} > K\} \quad (3.18)$$

$$q = \frac{\hat{p}u}{1+r} \quad (3.19)$$

Macht man nun den Übergang zum kontinuierlichen Modell (siehe zum Beispiel ([29]) oder ([19])), so erhält man wiederum die Black-Scholes partielle Differentialgleichung (3.16).

Feynman-Kac Alternativ zur Herleitung über Itô's Lemma kann man die partielle Differentialgleichung auch dadurch erhalten, dass man den Erwartungswert direkt betrachtet. Vernachlässigen wir nun den Zinssatz für einen Moment (er kann später wieder problemlos hinzugenommen werden). Die Brownsche Bewegung für den zugrunde liegenden Wert ist nach wie vor dieselbe wie in (3.4). Sei der Optionswert nun gegeben als der Erwartungswert

$$V(S, t) = E_{S,t}[V_{end}(S(T))]. \quad (3.20)$$

Dann sind die Voraussetzungen des Theorems von Feynman und Kac⁴ erfüllt, so dass sich der Optionswert als Lösung der partiellen Differentialgleichung

$$\frac{\partial V}{\partial t} + \mu \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} = 0 \quad (3.21)$$

mit Endbedingung

$$V(S, T) = V_{end}(S) \quad (3.22)$$

ergibt. Nimmt man den Zinssatz hinzu, landet man wieder bei der Black-Scholes Gleichung (3.16).

3.2.5 Numerik bei Aktienoptionen

Prinzipiell gibt es drei grundlegend verschiedene Möglichkeiten, den Preis von Aktienoptionen mit dem Black-Scholes Modell numerisch zu berechnen. Zum einen kann die partielle Differentialgleichung (3.16) gelöst werden, oder das Problem kann mittels der Feynman-Kac Formel auf ein Integrationsproblem transformiert werden, das dann mit Integrationsverfahren gelöst werden kann. Als dritte Möglichkeit wird das Black-Scholes Verfahren approximativ über Binomialbäume gelöst. Wir simulieren die Pfade mit einem Monte-Carlo oder Quasi-Monte-Carlo Verfahren. Die Beschränkung beim Lösen der partiellen Differentialgleichung war bislang, dass nur in wenigen Dimensionen gerechnet werden konnte. Die zunehmend besseren Rechner, aber vor allem auch verbesserte numerische Verfahren wie dünne Gitter, erlauben eine Erhöhung der berechenbaren Dimensionen. Monte-Carlo- und Quasi-Monte-Carlo Verfahren können zwar eine größere Anzahl von Dimensionen bewältigen, leiden dann allerdings stark an der langsamen Konvergenz.

Insgesamt ist es wünschenswert, das Modell über die partielle Differentialgleichung zu lösen. Erstens ist die Konvergenz bei weitem schneller als bei Monte-Carlo Verfahren und zweitens ist die Lösung dann auf dem gesamten Gebiet vorhanden, was die Berechnung der Greeks — die Ableitungen in verschiedene Richtungen — erlaubt. Dies ist ein wesentlicher Vorteil, da Anleger anhand den Werten dieser Ableitungen das Risiko ihres Portfolios bestimmen. Zur einfacheren Berechnung wollen wir die ortsabhängigen Koeffizienten entfernen. Dies gelingt mit folgender Transformation:

Satz 3.2.2 *Die Black-Scholes partielle Differentialgleichung (3.16) lässt sich auf die Wärmeleitungsgleichung*

$$-\frac{\partial \tilde{V}(\tilde{S}, \tilde{t})}{\partial \tilde{t}} + \Delta \tilde{V}(\tilde{S}, \tilde{t}) = 0 \quad (3.23)$$

⁴siehe zum Beispiel ([4])

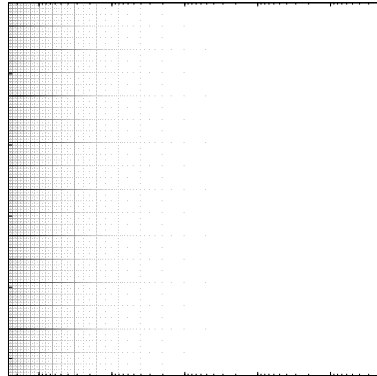


Abbildung 3.4: Reguläres dünnes Gitter auf Level 9 durch Transformation gradiert. Die Ortsachse ist horizontal, die Zeitachse vertikal.

transformieren.

Beweis: Die Transformationen, die die Black-Scholes Gleichung auf die Wärmeleitungsgleichung transformieren, sind gegeben durch:

$$V(S, t) = e^{\gamma\tilde{S} + \delta\tilde{t}} \tilde{V}(\tilde{S}, \tilde{t}) \quad (3.24)$$

$$S = e^{\tilde{S}} \quad (3.25)$$

$$t = T - \frac{2}{\sigma^2} \tilde{t} \quad (3.26)$$

Einsetzen liefert das gewünschte Ergebnis. Die Randbedingungen ändern sich analog (siehe auch ([27])). Wir werden später eine multivariate Transformation einführen. \square

Wie wir gerade gesehen haben, wird das Gebiet von $[0, \infty)$ auf $(-\infty, \infty)$ erweitert. Da wir nur auf endlichen Gebieten rechnen können, muss nun also auf beide Seiten "abgeschnitten" werden, um numerische Berechnungen starten zu können. Wird die Wärmeleitungsgleichung mit adäquaten Randdaten berechnet, so erhält man die Lösung der Black-Scholes Gleichung durch Rücktransformation. Das Gitter, das dabei entsteht, ist nun gradiert (siehe Abbildung 3.4). Die Ballung von Punkten ist nicht unbedingt erwünscht. Da wir jedoch adaptive Gitter verwenden werden, ist dadurch keine Einbuße bemerkbar.

3.2.6 Erweiterungen des Modells

Zahlreiche Erweiterungen des Modells führen auf hochdimensionale Probleme. Darunter fallen auch diese Arten⁵:

Optionen auf Baskets Wie wir im Abschnitt 3.4 sehen werden, erhöht sich die Dimension mit der Anzahl der zugrundeliegenden Wertpapiere. In der Praxis können auf diese Art hochdimensionale Probleme entstehen, wenn man beispielsweise an Optionen auf den DAX⁶ denkt, der 30 Werte beinhaltet. Mit dieser Art von Optionen werden wir uns eingehender beschäftigen.

Erweiterung bislang konstanter Parameter Die Zinsrate und die Volatilität werden, wie oben beschrieben, im klassischen Black-Scholes Modell als konstant angenommen. Neuerdings werden Modelle entwickelt, die den Unterschied zwischen kurz- und langfristiger Zinsrate berücksichtigen. Die Volatilität wird selbst als stochastische Variable modelliert. Die bisherigen Parameter sind dann Funktionen mehrerer Veränderlicher.

Pfadabhängige Optionen Wie bereits erwähnt, werden Optionen angeboten, deren Wert sich nach vergangenen Werten der zugrundeliegenden Wertpapiere richten (siehe Abschnitt 3.1.1). Diese Optionen kann man als Anreihung eindimensionaler Optionen betrachten, deren Endwert der jeweilige Startwert der nächsten Option ist. So wird der Pfad in mehrere einzelne Optionen aufgeteilt.

Lieferung einer bestimmten Qualität Ein Vertrag, der zur Lieferung einer bestimmten Ware festgelegter Qualität verpflichtet, kann als Option auf mehrere Veränderliche interpretiert werden. Der Lieferant hat nämlich die Wahl der verschiedenen Bezugsquellen, die die Qualität garantieren, wird aber zum Ausübungszeitpunkt die billigste Bezugsquelle wählen.

3.3 Die "Greeks"

Von besonderem Interesse, vor allem für das Hedging, sind die Änderungen der Optionswerte, wenn Parameter oder Variablen geändert werden. Berechnet werden diese Änderungen durch die entsprechenden Ableitungen,

⁵siehe auch Barraquand ([4])

⁶Deutscher Aktienindex

die dann einen griechischen Buchstaben zugewiesen bekommen. Daher rührt auch ihr Name — die sogenannten "Greeks".

Die Greeks können in zwei Arten unterschieden werden: zum einen werden Ableitungen in Richtung einer Variablen berechnet, zum anderen in Richtung eines Parameters. Für die numerische Berechnung ist dieser Unterschied bedeutend. Während für die Ableitungen nach einer Variablen nur ein Operator auf die Lösung angewandt werden muss, kann die Ableitung in Richtung eines Parameters nur durch den Vergleich mehrerer Lösungen errechnet werden. Daher steigt der Aufwand bei der Berechnung letzterer Greeks beträchtlich.

Wir haben bereits zwei extreme Arten des Hedgings kennengelernt: zum einen haben wir die Put-Call Parität (siehe Beispiel 3.1.2), zum anderen haben wir in der Herleitung der Black-Scholes partiellen Differentialgleichung das "Delta-Hedging" diskutiert. Beide führten jeweils zur vollständigen Elimination des Risikos. Sie sind freilich grundverschieden. Man betrachte dazu die Transaktionen, die vonnöten sind, um das Portfolio im risikofreien Zustand zu halten. Im ersten Falle werden die drei Positionen einmalig gekauft und dann bis zum Ausübungszeitpunkt gehalten. Das Delta-Hedging jedoch ist auf eine dynamische Anpassung angewiesen. In jedem Zeitpunkt wird das Delta berechnet und das Portfolio dann kontinuierlich angepasst.

Es gibt allerdings eine ganze Bandbreite an Hedging-Strategien. Man kann nicht nur die Effekte erster, sondern auch zweiter und höherer Ordnung, hedgen. Dazu betrachten wir nun die wichtigsten "Greeks".

3.3.1 Einige wichtige Greeks

Hier soll eine Liste der bekanntesten Kennzahlen für Optionswerte gegeben werden.

Ableitungen nach Variablen

Delta $\Delta = \frac{\partial V}{\partial S}$

Delta misst die Sensitivität des Aktienwertes, wenn sich der Aktienkurs minimal verändert. Die entsprechende Hedging-Strategie haben wir eingangs erwähnt und in der Black-Scholes Analysis praktisch umgesetzt. Es wird hier versucht, Risiken aus Effekten erster Ordnung zu eliminieren zu können.

Gamma $\Gamma = \frac{\partial^2 V}{\partial S^2}$

Gamma misst, wie stark das Delta auf Änderungen des Aktienwertes reagiert. Gamma beschreibt die Krümmung des Aktienwertes in Richtung des zugrunde liegenden Wertes. Das Gamma wird für den Investor

dann interessant, wenn er Effekte zweiter Ordnung kontrollieren möchte.

Theta $\Theta = \frac{\partial V}{\partial t}$

Theta gibt an, wie sich der Optionswert in der Zeit bewegen wird. Für Hedging-Strategien ist es interessant, ob sich der Wert der Option in Kürze stark ändern wird. Ein sehr kleines Θ zeigt an, dass der Zeitpunkt des Hedgings nicht allzu große Auswirkungen hat. Ist Θ jedoch sehr groß, so muss der Investor möglicherweise schnell handeln. Das Θ spielt vor allem eine große Rolle, wenn man Transaktionskosten berücksichtigt, da dann die Anzahl der Transaktionen einen Einfluss auf den Gewinn hat.

Der Zusammenhang zwischen den Ableitungen nach Variablen ergibt sich aus der Black-Scholes Gleichung (3.16):

$$\Theta + rS\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma - rV = 0 \quad (3.27)$$

Demnach kann ein Optionschreiber Θ und Γ nicht unabhängig voneinander wählen.

Ableitung nach Parametern

Die Kennzahlen Rho und Vega geben an, wie unterschiedlich der Kurs der Option sich in einer leicht veränderten Umgebung ergeben würde.

Rho $\rho = \frac{\partial V}{\partial r}$

Die Zinsrate wird in unseren Anwendungen als konstant angenommen. Dies muss aber nicht immer so sein. Man kann die Modelle durchaus dahingehend erweitern, dass sie Schwankungen im Diskontfaktor berücksichtigen. Die Ergebnisse bringen für Aktienoptionen — aus theoretischer Sicht — keine substantiellen Veränderungen mit sich. In der Praxis kann die Zinsrate schwanken und man möchte wissen, ob sich daraus abweichende optimale Hedging-Strategien ergeben.

Vega $\Lambda = \frac{\partial V}{\partial \sigma}$

Andere Modelle erlauben, σ als Variable zu betrachten, die einem eigenen stochastischen Prozess — eventuell gekoppelt an den des zugrunde liegenden Wertes — folgt. Dann macht es durchaus Sinn, das Λ zu berechnen. In der Praxis wird die implizierte Volatilität aus den Marktdaten gewonnen. Da man Veränderungen in der Volatilität also sofort

sieht, möchte man gerne den Einfluss auf die optimale Strategie erfahren. Dazu reicht aber das Wissen um den Trend der Volatilität nicht aus. Es muss ein exakter Wert für die Volatilität bestimmt werden, was sich als Schwierigkeit herausstellt, da dieser von dem betrachteten Zeitraum abhängt. Ebenfalls sind Einschätzungen über den zukünftigen Verlauf der Volatilität nur eingeschränkt möglich.

Mit diesen Kennzahlen werden wir uns hier nicht weiter befassen.

3.3.2 Die Berechnung der Greeks mit der Lösung der transformierten Gleichung

Ziel dieses Abschnitts ist es, die Greeks aus der Lösung der Wärmeleitungsgleichung berechnen zu können, ohne vorher eine Rücktransformation durchzuführen. Es sollen also Operatoren gefunden werden, die direkt auf die Lösung angewendet werden können und dann die Werte der Greeks liefern. Im folgenden Satz zeigen wir, dass sich die Greeks direkt aus der Lösung des transformierten Problems berechnen lassen:

Satz 3.3.1 (Die Greeks für das transformierte Problem) *Die Greeks aus der Lösung der Black-Scholes Gleichung (3.16) können aus der Lösung der Wärmeleitungsgleichung mit entsprechenden Anfangsdaten $\tilde{V}(\tilde{S}, \tilde{t})$ gewonnen werden, indem folgende Operatoren angewendet werden:*

- *Delta:*

$$\Delta = \frac{\partial V}{\partial S} = e^{\gamma(\tilde{S}-1)+\delta\tilde{t}}(\gamma\tilde{V}(\tilde{S}, \tilde{t}) + \frac{\partial\tilde{V}}{\partial\tilde{S}}) \quad (3.28)$$

- *Gamma:*

$$\Gamma = \frac{\partial^2 V}{\partial S^2} = \frac{\partial\Delta}{\partial S} = e^{\gamma(\tilde{S}-2)+\delta\tilde{t}}(\gamma^2\tilde{V}(\tilde{S}, \tilde{t}) + 2\gamma\frac{\partial\tilde{V}}{\partial\tilde{S}}(\tilde{S}, \tilde{t}) + \frac{\partial^2\tilde{V}}{\partial\tilde{S}^2}) \quad (3.29)$$

- *Theta:*

$$\Theta = \frac{\partial V}{\partial t} = -\frac{\sigma^2}{2}e^{\gamma\tilde{S}+\delta\tilde{t}}(\delta\tilde{V}(\tilde{S}, \tilde{t}) + \frac{\partial\tilde{V}}{\partial\tilde{t}}) \quad (3.30)$$

Beweis: Die entscheidenden Transformationen der Black-Scholes Gleichung auf die Wärmeleitungsgleichung haben wir bereits im Beweis zu Satz 3.2.2 gesehen. Nach Anwendung der Kettenregel lassen sich die Operatoren auf einfache Weise berechnen. \square

3.4 Europäische Optionen auf "Aktienkörbe"

Wir werden uns in dieser Arbeit auf die Berechnung europäischer Optionen beschäftigen. In diesem Abschnitt wird die Analysis für europäische Optionen auf Aktienkörbe, sogenannte Basket-Optionen, diskutiert.

Es sei $\mathbf{S}(t)$ der endlich-dimensionale Vektor, der die Zustandsvariablen (hier: die Werte der darunterliegenden Aktien) beschreibt. Zum Zeitpunkt $t = a$ sind alle Werte $S_i(\tau)$ bekannt für $\tau \leq a$. Für $\tau > a$ ist $\mathbf{S}(\tau)$ ein stochastischer Prozess, hier per Konstruktion durch einen Markov-Prozess. Dieser sei der Einfachheit halber ein Diffusions-Prozess, der durch eine stochastische Differentialgleichung beschrieben werden kann:

$$d\mathbf{S} = \mathbf{A}(\mathbf{S}(t), t)dt + B(\mathbf{S}(t), t)d\mathbf{W} \quad (3.31)$$

Dabei ist der Vektor $\mathbf{A}(\mathbf{S}(t), t) = a_i(s_1, \dots, s_n, t)$ die Drift des Prozesses. \mathbf{W} ist eine d-dimensionale Brownsche Bewegung.

Häufig wird auch folgende Darstellung verwendet, die sich durch folgende Variablen ergibt $\mu_{S_i} := \frac{a_i}{S_i}$ und $\sigma_{ij} := \frac{b_{ij}}{S_i}$:

$$\frac{dS_i}{S_i} = \mu_{S_i}(S_1, \dots, S_d, t) + \sum_{j=1}^d \sigma_{ij}(S_1, \dots, S_d, t)dw_j$$

Die Matrix Σ ist dabei die Volatilitätsmatrix.

Analog zum eindimensionalen Fall läßt sich die mehrdimensionale Black-Scholes Differentialgleichung für eine Option mit Wert V herleiten, die dann numerisch zu lösen ist.

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \sigma_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + r \sum_{i=1}^d S_i \frac{\partial V}{\partial S_i} - rV = 0 \quad (3.32)$$

3.4.1 Transformation auf die Wärmeleitungsgleichung

Die Black-Scholes partielle Differentialgleichung ist vom Euler-Typ und läßt sich daher auf die bekannte Art und Weise transformieren. Die eindimensionale Transformation findet man in der einschlägigen Literatur (zum Beispiel Seydel ([27])). Dass es auch eine Transformation für mehrdimensionale Problemstellungen gibt, wollen wir hier zeigen:

Lemma 3.4.1 *Die Black-Scholes partielle Differentialgleichung (3.32) läßt sich auf die Wärmeleitungsgleichung*

$$\tilde{V}_t - \Delta \tilde{V} = 0 \quad (3.33)$$

transformieren.

Beweis: Mit den Transformationen

$$S_i = K \cdot e^{\hat{S}_i} \quad (3.34)$$

$$t = T - 2\tilde{t} \quad (3.35)$$

$$V(S, t) = K e^{\gamma \hat{\mathbf{S}} + \delta \tilde{t}}(\hat{\mathbf{S}}, \tilde{t}) \quad (3.36)$$

wollen wir den Konvektions- und Reaktionsteil beseitigen und lediglich konstante Koeffizienten behalten. Wählen wir γ und δ so, dass

$$2 \cdot \sum_j \gamma_j \sigma_{ij} + 2 \cdot r - \sigma_{ii} = 0 \quad \forall_{1 \leq i \leq d} \quad (3.37)$$

$$\delta = \sum_{i,j} \gamma_i \gamma_j \sigma_{ij} + \sum_i (2 \cdot r - \sigma_{ii}) \gamma_i, \quad (3.38)$$

so erhalten wir die Gleichung

$$-\frac{\partial \tilde{V}(\hat{\mathbf{S}}, \tilde{t})}{\partial \tilde{t}} + \sum_{i,j} \sigma_{ij} \frac{\partial^2 \tilde{V}(\hat{\mathbf{S}}, \tilde{t})}{\partial \hat{S}_i \partial \hat{S}_j} = 0 \quad (3.39)$$

Mit einer letzten Transformation

$$\tilde{\mathbf{S}} = A^{-1} \cdot \hat{\mathbf{S}} \quad (3.40)$$

verschwinden nun auch die Koeffizienten und die Wärmeleitungsgleichung resultiert. \square

3.4.2 Existenz, Positivität und Eindeutigkeit der Lösung zur Black-Scholes Gleichung

Um einfache Resultate für die Existenz, Positivität und Eindeutigkeit von Lösungen der Black-Scholes partiellen Differentialgleichung zu erhalten, transformieren wir die Gleichung auf die Wärmeleitungsgleichung. Bezeichne \mathcal{L} den Wärmeleitungsoperator.

Die Existenz einer Lösung für

$$\mathcal{L}V = 0 \quad (3.41)$$

$$V(S, 0) = V_0(S) \quad (3.42)$$

ist bekannt, denn die Wärmeleitungsgleichung besitzt die Fundamentallösung $\tilde{V} : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$

$$\tilde{V}(\mathbf{S}, t) = \begin{cases} \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|\mathbf{S}|^2}{4t}} & t > 0 \\ 0 & \text{sonst} \end{cases} \quad (3.43)$$

Die Fundamentallösung wird gegen die Anfangsdaten gefaltet, um die Lösung des Anfangswertproblems zu erzeugen.

Die beiden folgenden Resultate liefern uns nun Positivität und Existenz der Black-Scholes Differentialgleichung.

Satz 3.4.1 Sei $\mathcal{L}V \leq 0$ in $\Omega_0 = \mathbb{R}^d \times (0, T]$ und

$$V(\mathbf{S}, t) \geq -Be^{\beta|\mathbf{S}|^2} \text{ in } \Omega = \mathbb{R}^d \times [0, T] \quad (3.44)$$

für positive Konstanten B, β . Sei nun $V_0(\mathbf{S}) \geq 0$ in \mathbb{R}^d , dann gilt: $V(\mathbf{S}, t) \geq 0$ in Ω .

Beweis: Siehe Friedman ([15]), Kapitel 2, Theorem 9 □

Satz 3.4.2 Es gibt höchstens eine Lösung des Problems (3.41) das

$$|\tilde{V}(\mathbf{S}, t)| \leq Be^{\beta|\mathbf{S}|^2} \quad (3.45)$$

erfüllt.

Beweis: Aus dem Maximumprinzip folgt, dass es genügt zu zeigen: wenn $\tilde{V}_0 \equiv 0$, dann ist $\tilde{V} \equiv 0$. Das folgt aber direkt aus dem vorherigen Satz, wenn man ihn auf \tilde{V} und $-\tilde{V}$ anwendet. □

3.4.3 Anzahl der Dimensionen

Die Frage ist, wie viele Dimensionen wir berechnen wollen. Prinzipiell gilt hier natürlich: "je mehr, desto besser". Interessant ist allerdings, ob es in der Anzahl der Dimensionen besonders wichtige Schwellen gibt. Dies ist eindeutig zu bejahen. Basket Optionen werden häufig auf bestehende Einteilungen in der Unternehmenslandschaft geschrieben. Möchte man sich beispielsweise gegen Konjunkturschwankungen absichern, bietet sich die Versicherungsbranche an, die meist azyklisch zur Konjunktur schwankt.

Die wohl wichtigste Gruppierung in Deutschland ist der DAX mit seinen 30 Werten. Damit haben wir den ersten Schwellenwert gefunden, der in Zukunft als Ziel für numerische Verfahren dienen wird. Der momentane Stand der Forschung erlaubt jedoch noch keine Berechnung von Problemen in 30 Dimensionen (plus Zeit). Diese dreißig Werte lassen sich allerdings in vier bzw. fünf (je nach Spezifizierungsgrad) Untergruppen aufsplitten. Geht man nun davon aus, dass die Werte innerhalb der Untergruppen sehr stark miteinander korreliert sind, so kann eine Option auf den DAX als Basket Option auf die vier resp. fünf Branchen berechnet werden.

Dieser Schwellenwert von vier bis fünf Branchen soll mit dieser Arbeit erreicht werden. Wollen wir eine Option auf den DAX berechnen, führen wir zunächst die gerade beschriebene Dimensionsreduktion durch⁷. Mit größeren und schnelleren Computern und mit verbesserten Algorithmen lassen sich in ein paar Jahren sicher auch 30 Dimensionen in Angriff nehmen. Approximativ sollten wir aber über die Dimensionsreduktion bereits jetzt brauchbare Ergebnisse erzielen können. Einen ähnlichen Ansatz wählt Reisinger (siehe Reisinger ([25])), der jedoch von Beginn an eine Kombitechnik — das heißt, Dimensionsreduktion und dünne Gitter — vorschlägt. Reisinger kann gleich große Probleme lösen wie wir, rechnet allerdings keine Greeks, auf deren Bestimmung wir Wert legen.

3.5 Randbedingungen und Gebietsgröße

3.5.1 Abschätzungen zur Gebietsgröße und den Randbedingungen

Zunächst ist es für unsere späteren Überlegungen wichtig zu wissen, wie sich ein Fehler am künstlichen Rand, den wir durch Abschneiden des Gebietes erzeugen müssen, ins Innere des Gebietes ausbreitet. Wir betrachten erst das Verhalten von Lösungen der Black-Scholes Gleichung (3.16) auf unendlichem Gebiet. Dann widmen wir uns beschränkten Gebieten und betrachten den Einfluss des Fehlers durch die künstlichen Randdaten auf das Innere.

Lemma 3.5.1 *Sei $D \in (0, \infty)^d$ ein offenes Gebiet, $V, W \in (\bar{D} \cap (0, \infty)^d) \times [0, T]$ und seien die ersten zwei örtlichen sowie die erste zeitliche Ableitung stetig in $D \times [0, T]$. Sei weiterhin*

$$-\frac{\partial V}{\partial t} + \tilde{\mathcal{L}}V \leq 0 \quad \text{in } D \times (0, T) \quad (3.46)$$

$$-\frac{\partial W}{\partial t} + \tilde{\mathcal{L}}W \geq 0 \quad \text{in } D \times (0, T) \quad (3.47)$$

$$V \geq W \quad \text{in } t = 0 \text{ und in } (\partial D \cap (0, \infty)^d) \times (0, T) \quad (3.48)$$

wobei der Operator \mathcal{L} der elliptische Teil des Black-Scholes Operators ist, das heißt

$$\tilde{\mathcal{L}} := \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \sigma_{ij} S_i S_j \frac{\partial}{\partial S_i \partial S_j} + \sum_{i=1}^d r S_i \frac{\partial}{\partial S_i} - r \quad (3.49)$$

⁷Im Anhang ist eine Gruppierung des DAX aufgeführt.

und gelte für Konstanten $C, k \geq 0$

$$V(\mathbf{S}, t) - W(\mathbf{S}, t) \geq -Ce^{k \cdot \sum_{i=1}^d (\log S_i)^2} \quad \forall (\mathbf{S}, t) \in D \times (0, T) \quad (3.50)$$

Dann gilt: $V \geq W$ in $D \times (0, T)$

Beweis: Mit der Transformation der Variablen $\tilde{S}_i = \log S_i$ erhalten wir einen uniformen parabolischen Operator. Friedman ([15]) liefert dann das gewünschte Resultat. \square

Dieses Lemma wird uns helfen, die Entwicklung der Lösung aus den Anfangsdaten zu verstehen.

Satz 3.5.1 Sei $V(\mathbf{S}, t)$ die Lösung der Black-Scholes Gleichung mit Anfangsdaten $V(\mathbf{S}, t) = V_0(\mathbf{S})$. Gelte für $\nu_i, \kappa_i \geq 0$, $i = 0, \dots, d$

$$-\nu_0 + \sum_{i=1}^d \nu_i S_i \leq V_0(\mathbf{S}) \leq \kappa_0 + \sum_{i=1}^d \kappa_i S_i \quad \forall \mathbf{S} \in (0, \infty)^d \quad (3.51)$$

Dann gelten die beiden folgenden Abschätzungen:

$$-\nu_0 e^{-r \cdot t} + \sum_{i=0}^d \nu_i S_i \leq V(\mathbf{S}, t) \quad (\mathbf{S}, t) \in (0, \infty)^d \times (0, T) \quad (3.52)$$

$$V(\mathbf{S}, t) \leq \kappa_0 e^{-r \cdot t} + \sum_{i=1}^d \kappa_i S_i \quad (\mathbf{S}, t) \in (0, \infty)^d \times (0, T) \quad (3.53)$$

Beweis: Definiere

$$u_1(\mathbf{S}, t) := -\nu_0 e^{-r \cdot t} + \sum_{i=1}^d \nu_i S_i \quad (3.54)$$

$$u_2(\mathbf{S}, t) := -\kappa_0 e^{-r \cdot t} + \sum_{i=1}^d \kappa_i S_i \quad (3.55)$$

Dann gilt:

$$\tilde{\mathcal{L}}u_1 \geq 0 \quad \text{und} \quad \tilde{\mathcal{L}}u_2 \leq 0 \quad (3.56)$$

Beweis: Der Beweis folgt direkt aus Lemma 3.5.1. \square

Nun widmen wir uns dem abgeschnittenen Gebiet $\Omega := [0, S_1^{max}] \times \dots \times [0, S_d^{max}]$ mit Rand Γ . Kangro und Nicolaidis ([21]) zeigen, dass der Randfehler im Gebiet beschränkt ist. Sie geben eine Schranke für den Fehler in folgendem Satz an:

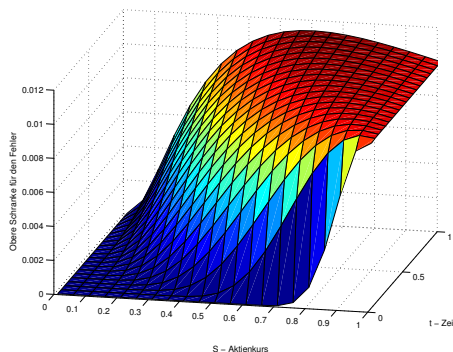


Abbildung 3.5: Obere Schranke für den Fehler, wenn rechts ein konstanter Fehler von 0.01 angenommen wird.

Satz 3.5.2 für zwei Lösungen V, W der Differentialgleichung (mit unterschiedlichen Randwerten) können wir die Differenz punktweise abschätzen durch:

$$|V(\mathbf{S}, t) - W(\mathbf{S}, t)| \leq \sup_{(\mathbf{S}^*, t^*) \in \Gamma \times (t, T)} \left\{ \frac{|V(\mathbf{S}^*, t^*) - W(\mathbf{S}^*, t^*)|}{z((\mathbf{S}^*, t^*))} \right\} z((\mathbf{S}, t)) \quad (3.57)$$

für alle $(\mathbf{S}, t) \in \Omega \times (0, T)$, wobei für nicht negative Konstanten C_i

$$z((\mathbf{S}, t)) = \sum_{i=1}^d C_i \cdot \frac{1}{\sqrt{T + \varepsilon_i - t}} e^{\left(-\gamma \frac{(\log \frac{S_i}{k_i S_i^{max}})^2}{T + \varepsilon_i - t} \right)} \quad (3.58)$$

mit $\varepsilon_i \geq 0$, $\gamma_i \geq 0$ und $k_i \geq 1$ definiert ist.

Dieser etwas längliche Ausdruck ist in Abbildung 3.5 für einen konstanten Randfehler veranschaulicht.

3.5.2 Randbedingungen

Um die Lösung numerisch berechnen zu können, müssen wir uns ein endliches Gebiet wählen. Die numerische Lösung des Problems auf diesem Gebiet soll die Black-Scholes Formel eingeschränkt auf das Gebiet approximieren. Damit das Problem gut gestellt ist, müssen wir Randbedingungen vorgeben. Unter welchen Umständen wir eine eindeutige Lösung für das abgeänderte Problem erhalten, wollen wir im folgenden untersuchen. Die Randbedingungen werden in diesem Abschnitt in allgemeiner Form präsentiert und später im Abschnitt 7.1.3 spezifiziert. Wir werden dann auch untersuchen, welche Art von Randbedingung für unsere Zwecke am besten geeignet ist.

Im folgenden werden wir einen allgemeinen Operator betrachten, um somit den Fall der Black-Scholes Differentialgleichung und ihrer transformierten Gleichung gleichzeitig zu behandeln. Wir untersuchen hier die folgenden drei Problemstellungen: in jedem der Fälle gilt im Gebiet $\Omega \times (0, T)$ mit $\Omega := [S_1^{max} \times \dots \times S_d^{max}]$ die Gleichung

$$-\frac{\partial V(\mathbf{S}, t)}{\partial t} + P(\mathbf{S})V(\mathbf{S}, t) = 0, \quad (3.59)$$

wobei der Operator $P(\mathbf{S}, t)$ entweder den elliptischen Teil des Wärmeleitungsoperators oder des Black-Scholes Operators bezeichnet. Die Aussagen im folgenden sind für beide Operatoren gültig. Das Randwertproblem habe die Startdaten

$$V(\mathbf{S}, t) = V_0(\mathbf{S}) \quad (3.60)$$

Nun verwenden wir verschiedene Randbedingungen, nämlich

Dirichlet-Randdaten Die Randbedingungen geben die Funktionswerte

$$V(\mathbf{S}, t) = V^D(\mathbf{S}, t) \quad \text{für } (\mathbf{S}, t) \in \partial\Omega \times (0, T) \quad (3.61)$$

auf dem Rand vor.

Neumann-Rand Hier sind die Ableitungen in Normalenrichtung auf dem Rand vorgegeben als

$$\frac{\partial V(\mathbf{S}, t)}{\partial S_i} = V_i^N(\mathbf{S}, t) \quad \text{für } (\mathbf{S}, t) \in \partial\Omega \times (0, T) \text{ für alle } i \quad (3.62)$$

Neumann-Bedingung zweiter Ordnung Und entsprechend

$$\frac{\partial^2 V(\mathbf{S}, t)}{\partial S_i^2} = V_i^{NN}(\mathbf{S}, t) \quad \text{für } (\mathbf{S}, t) \in \partial\Omega \times (0, T) \text{ für alle } i \quad (3.63)$$

Aussagen zu Dirichlet-Randwerten

Existenz und Eindeutigkeit des Problems mit parabolischer Gleichung (3.59) ist bekannt und kann in Friedman ([15]) nachgelesen werden. Wir wollen aber Aussagen über die Regularität der Lösung treffen. Die Funktion $V^D(\mathbf{S})$ in (3.61) auf den Rändern kann nicht beliebig gewählt werden. Welche Regularität wir erwarten können, sagt uns der folgende Satz aus Flyer und Swarztrauber ([14]):

Satz 3.5.3 *Die Lösung des Anfangswertproblems, das durch die Gleichungen (3.59), (3.60) und (3.61) gegeben ist, hat eine Lösung $u(x, t)$ in C^{2k} , wenn das Gebiet eine C^∞ Mannigfaltigkeit der Dimension $(d-1)$ ist und die Randbedingungen folgende Kompatibilitätsbedingungen erfüllen:*

$$L^k V_0 = \frac{\partial^k V^D}{\partial t^k} \quad k = 0, 1, 2, \dots \quad \in [0, 0] \times \partial\Omega \quad (3.64)$$

Aussagen zu Neumann-Randwerten

Existenz und Eindeutigkeit können für das obige Problem aus einem Theorem aus ([15]) gewonnen werden, der ebenfalls angibt, wie glatt die Lösung sein wird. Daraus lassen sich dann die entsprechenden Kompatibilitätsbedingungen gewinnen. Für den folgenden Satz brauchen wir folgende Bedingung:

Satz 3.5.4 *Seien die Koeffizienten des Operators (3.32) hinreichend glatt in $\bar{\Omega} \times [0, T]$ und sei der Rand ∂D genügend glatt. Dann hat das Problem eine eindeutige Lösung. Weiterhin ist die Lösung $V(\mathbf{S}, t)$ in C^p für eine beliebige gegebene ganze Zahl $p > 0$, wenn der Rand und die Koeffizienten genügend glatt sind.*

Aussagen zu Neumann-Randwerten zweiter Ordnung

Existenz und Eindeutigkeit für das Neumann-Problem mit zweiten Ableitungen (3.63) lassen sich im eindimensionalen Fall auf das vorige Problem zurückführen. Wir werden dies am Beispiel des Black-Scholes Operators (3.32) berechnen. Nötig ist eine höhere Regularität, denn man differenziert die Gleichung nach dem Aktienwert S und erhält

$$-V_t S + \sigma^2 S V_{SS} + \frac{1}{2} \sigma^2 S^2 V_{SSS} + r V_S + r S V_{SS} - d V_S = 0 \quad (3.65)$$

Definiert man nun $\tilde{V} := V_S$, so erhält man das folgende Problem

$$-\tilde{V}_t + \frac{1}{2} \sigma^2 S \tilde{V}_{SS} + (\sigma^2 + r) S^2 \tilde{V}_S = 0 \quad (3.66)$$

mit Startdaten $V_S(S, 0) = V_{0S}$ und Neumann-Randdaten $V(S, t) = V^N(t)$ für $S \in \partial\Omega$. Die ursprünglich gesuchte Lösung $V(S, t)$ läßt sich nun wieder zurückgewinnen durch

$$V(S, t) = \int_{S_0}^S \tilde{V}(\xi, t) d\xi + c(t). \quad (3.67)$$

Partielle Integration und Leibniz-Regel führen auf die Form

$$\begin{aligned} \int_{S_0}^S -\tilde{V}_t + \frac{1}{2} \sigma^2 \xi^2 \tilde{V}_{SS} + (\sigma + r) \xi \tilde{V}_S d\xi \\ - \dot{c}(t) + \frac{1}{2} \sigma^2 S_0^2 \tilde{V}_S(S_0, t) + r S_0 \tilde{V}(S_0, t) = 0 \end{aligned} \quad (3.68)$$

Verwendet man nun Gleichung (3.66), so kann man den zeitabhängigen Koeffizienten über

$$c(t) = \int_{t_0}^t \frac{1}{2} \sigma^2 S_0^2 \tilde{V}_S(S_0, t) + r S_0 \tilde{V}(S_0, t) dt + k(S) \quad (3.69)$$

Die Konstante $k(S)$ lässt sich über die Startdaten $V_0(S)$ fixieren.

Damit sind Existenz und Eindeutigkeit für das eindimensionale Problem gesichert.

Kapitel 4

Zinsderivate

Für Aktienoptionen haben wir bereits gesehen, wie man die partielle Differentialgleichung zur Bewertung des jeweiligen Derivats herleiten kann. Im wesentlichen bleiben die Methoden die gleichen, wenn Zinsderivate betrachtet werden. Im folgenden werden wir nach einer Einführung in die Welt der Zinsderivate zwei bedeutende Arten von Finanzprodukten betrachten, Bonds und Swaps. Für beide Arten werden wir uns interessante Spezialfälle suchen, die große Relevanz an den Finanzmärkten haben, aber auch von numerischem Interesse sind. Wir werden hier sogenannte Floors betrachten, die auf eine partielle Differentialgleichung mit glatten Anfangsdaten, konstanten Koeffizienten, aber unglatter rechter Seite führen. Im letzten Teil werden wir numerische Ergebnisse für dieses Produkt sehen und zeigen können, dass unser numerisches Verfahren unter den auftretenden Schwierigkeiten nicht leiden wird, das heißt, dass die Konvergenzrate erhalten bleibt.

4.1 Bonds

Definition 4.1.1 (Bond) *Ein Bond ist ein Vertrag, bei dem der eine Vertragspartner vorab dafür bezahlt, dass er zum Verfallsdatum T einen festgelegten Betrag Z erhält. Es kann dabei vereinbart werden, dass kontinuierlich oder zu diskreten Zeitpunkten Zahlungen fällig werden.*

Spannend ist für uns in diesem Zusammenhang, wie groß der faire Wert für die Vorab-Zahlung ist. Zur Illustration der Problemstellung betrachten wir ein Modellproblem mit stark stilisierten Fakten.

Beispiel 4.1.1 *Sei die Zinsrate r konstant und bekannt, dann läßt sich der Wert des Bonds $V(t)$, der eine kontinuierliche Dividende von $K(t)$ abwirft,*

über seine Änderung bewerten. Die Änderung des Gegenwertes für den Bond beläuft sich auf

$$\left(\frac{dV}{dt} + K(t) \right) dt \quad (4.1)$$

pro Zeitschritt dt . Aus Arbitragegründen muss dies gleich dem Ertrag sein, den man bei Anlage zum Marktzins erwirtschaften würde. Daher gilt:

$$\frac{dV}{dt} + K(t) = r(t)V \quad (4.2)$$

Die Lösung dieser gewöhnlichen Differentialgleichung ist bekannt:

$$V(t) = e^{-\int_t^T r(\tau) d\tau} \left(Z + \int_t^T K(s) e^{\int_s^T r(\tau) d\tau} ds \right) \quad (4.3)$$

4.1.1 Bewertung von Bonds

Zunächst gibt man sich eine stochastische Bewegung für die Zinsrate vor, wie wir dies auch bereits bei Aktienoptionen getan haben. Der Zinssatz ist hier analog in der Form der stochastischen Differentialgleichung

$$dr = u(r, t)dt + w(r, t)dX \quad (4.4)$$

gegeben. Anders als bei Aktienoptionen kann man hier nicht gegen einen dahinterliegenden Wert hedgen und so ein risikofreies Portfolio generieren. Daher geben wir uns zwei Bonds vor, V_1 und V_2 , die unterschiedliche Ablaufdaten T_1 und T_2 , aber die gleiche kontinuierliche Dividendenzahlung $K(t)$ haben. Das zu betrachtende Portfolio ergibt sich nun aus

$$\Pi = V_1 - \Delta V_2. \quad (4.5)$$

Wenden wir nun Itô's Lemma an, um die Änderung des Portfolios zu bestimmen, so erhalten wir

$$\begin{aligned} d\Pi = & \frac{\partial V_1}{\partial t} dt + K(t)dt + \frac{\partial V_1}{\partial r} dr + \frac{1}{2} w^2 \frac{\partial^2 V_1}{\partial r^2} dt \\ & - \Delta \left(\frac{\partial V_2}{\partial t} dt + K(t)dt + \frac{\partial V_2}{\partial r} dr + \frac{1}{2} w^2 \frac{\partial^2 V_2}{\partial r^2} dt \right). \end{aligned} \quad (4.6)$$

Wählt man den bisher freien Parameter Δ als

$$\Delta = \frac{\partial V_1 / \partial r}{\partial V_2 / \partial r}, \quad (4.7)$$

so erhält man durch Einsetzen in (4.6), dass für das Portfolio gilt:

$$d\Pi = r\Pi dt. \quad (4.8)$$

Sammelt man nun die Terme, so trifft man auf die Gleichung

$$\begin{aligned} \left(\frac{\partial V_1}{\partial t} + \frac{1}{2}w^2 \frac{\partial^2 V_1}{\partial r^2} - rV_1 + K(t) \right) / \frac{\partial V_1}{\partial r} \\ = \left(\frac{\partial V_2}{\partial t} + \frac{1}{2}w^2 \frac{\partial^2 V_2}{\partial r^2} - rV_2 + K(t) \right) / \frac{\partial V_2}{\partial r}, \end{aligned} \quad (4.9)$$

wobei die linke Seite eine Funktion von T_1 , die rechte Seite jedoch eine Funktion von T_2 ist. Das bedeutet, dass die Gleichung unabhängig von den Ablaufdaten T_1 und T_2 ist und wir erhalten die Form

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2}w^2 \frac{\partial^2 V}{\partial r^2} - rV + K \right) / \frac{\partial V}{\partial r} = w(r, t)\lambda(r, t) - u(r, t) \quad (4.10)$$

mit dem Marktpreis für Risiko $\lambda(r, t) = \frac{\mu-r}{\sigma}$. (4.10) lässt sich direkt in die partielle Differentialgleichung

$$\frac{\partial V}{\partial t} + \frac{1}{2}w^2 \frac{\partial^2 V}{\partial r^2} + (u - \lambda w) \frac{\partial V}{\partial r} - rV + K = 0 \quad (4.11)$$

umschreiben. Die Endbedingung ist dabei gegeben durch

$$V(r, T) = Z \quad (4.12)$$

Definition 4.1.2 (Eindimensionaler Floor) *Ein eindimensionaler Floor ist ein Darlehen zum fließenden Zinssatz (d.h. zum jeweils gültigen Marktzinssatz) mit der zusätzlichen Klausel, dass die Zinsrate nicht unter einen bestimmten Wert \bar{r} fallen darf.*

Wir erkennen sofort, dass die Dividendenzahlung in (4.11) die Struktur

$$K(t) = \max(r, \bar{r}) \quad (4.13)$$

erhält. Floors koppeln unter Umständen verschiedene Bonds. Ein Investor kauft von einem Broker d verschiedene Bonds, deren Zinssraten sich unabhängig voneinander bewegen (die Bonds können in verschiedenen Ländern laufen, während die Zahlung in der nationalen Währung fixiert ist). Weil er sich bereits durch anderweitiges Hedging gegen Schwankungen abgesichert hat, will der Investor, dass der Zinssatz nicht unter ein bestimmtes Level

fällt. Um seine Floors bewerten zu können, löst folgende mehrdimensionale partielle Differentialgleichung:

$$\frac{\partial V}{\partial t} + \sum_{i=1}^d \frac{1}{2} w_i^2 \frac{\partial^2 V_i}{\partial r_i^2} + \sum_{i=1}^d (u_i - \lambda_i w_i) \frac{\partial V}{\partial r_i} - r_i V_i + \max\left(\frac{1}{d} \sum_{i=1}^d r_i, \bar{r}\right) = 0 \quad (4.14)$$

Die Idee ist hierbei, dass der Wechselkurs direkt in den Zinswert miteinfließt. Die Volatilitäten müssen dementsprechend angepasst werden. Wir werden die Bewertung dieser Art von Floors später numerisch simulieren.

4.1.2 Randbedingungen

Überall, wo

$$\sum_{i=1}^d r_i < \bar{r} \quad (4.15)$$

gilt, wird der Zinssatz \bar{r} als Zinsrate für die Diskontierung des Endwertes verwendet. Der Wert steigt dort also exponentiell an. Die Dirichlet-Daten sind:

$$V(\bar{\mathbf{r}}, t) = Z \cdot e^{\bar{r} \cdot t} \quad \text{für kleine Zinsraten} \quad (4.16)$$

$$\lim_{\max r_i \rightarrow \infty} V(\mathbf{r}, t) = 0 \quad (4.17)$$

Die Endbedingung ist durch die konstante Zahlung Z gegeben, die unabhängig von der Zinsrate gezahlt wird.

4.2 Cross-Currency Swaps

4.2.1 Swaps

Wir werden uns nun mit einem weiteren Zinsderivat, den sogenannten Swaps beschäftigen. Wir haben gesehen, dass sich die Zinsrate ähnlich wie ein Aktienwert bewegt. Allerdings gibt es die Möglichkeit, Geld langfristig zu einem vorab festgelegten Zinssatz anzulegen. Analog zu Optionen lässt sich ein größerer Hebeleffekt erzielen, wenn man lediglich die Differenz zwischen fließender und freier Zinsrate betrachtet. Diese Hebelwirkung wird durch einen Swap erreicht:

Definition 4.2.1 (Swap) *Ein Swap ist die Vereinbarung, zu einem festen Termin die Zinsen auf einen festgelegten Basiswert Z zu tauschen.*

Eine besonders interessante Art von Swaps entsteht, wenn die Zahlungen in unterschiedlichen Währungen fällig werden. Dieses Zinsderivat erfreut sich bei Investoren zunehmender Beliebtheit (siehe zum Beispiel Hutton [20]). So können die Anleger mit Zinsraten unterschiedlicher Länder spekulieren oder sich gegen die Schwankungen absichern. Interessant ist diese Art des Swaps vor allem, weil in der Berechnung ein weiterer unsicherer Faktor, nämlich der Wechselkurs, mit ins Spiel kommt. Da wir die Zinsraten in den Ländern als fließend annehmen, erhalten wir also drei stochastische Größen: die beiden Zinsraten sowie den Wechselkurs. Speziell betrachten wir hier den Differential Swap, das heißt, dass nationale und ausländische Zinsraten ausgetauscht werden, die aber auf denselben nationalen Basiswert Z gezahlt werden.

Wir betrachten im folgenden den Austausch von LIBOR-Raten¹. Die LIBOR-Rate L gilt immer für einen bestimmten Zeitraum (zum Beispiel 3 oder 6 Monate), der hier mit dem Intervall $[t, M]$ bezeichnet wird. Wir bezeichnen den Zero-Coupon Bond (ein Bond, der keinen Zins abwirft, aber billig gehandelt wird) für diese Periode mit $P(t, M)$.

Die LIBOR-Rate per anno läßt sich dann aus den gegebenen Daten berechnen durch

$$L(t, M) := \frac{1}{\delta} \times \frac{1 - P(t, M)}{P(t, M)} \quad \text{mit} \quad (4.18)$$

$$\delta := \frac{\text{Anzahl der Tage in } [t, M]}{\text{Basistage pro Jahr}}, \quad (4.19)$$

wobei die Anzahl der Basistage pro Jahr beispielsweise bei 365 für Pfund Sterling und 360 für US Dollars liegen.

Jeweils zum Ende der jeweiligen LIBOR-Perioden $[t_{j-1}, t_j]$ fallen Zahlungen p_{kj} an, die vor dem Erreichen des Zeitpunktes t_j beglichen werden. k kann dabei den Wert d für die hiesige Zahlung und f für die Zahlung im Ausland annehmen. Für Cross-Currency Swaps sehen die Zahlungen wie folgt aus:

$$p_j := Z\delta_j[k_d(L_d(t_{j-1}, t_j) + m) - k_f L_f(t_{j-1}, t_j)] \quad (4.20)$$

Die Endbedingungen jeder Periode bestehen in dem Barwert der verbleibenden Zahlungen, so zum Beispiel für die Periode $[t_{j-2}, t_{j-1}]$

$$V_j(t_{j-1}^-) = P_d(t_{j-1}, t_j)p_{dj} - S(t_{j-1})P_f(t_{j-1}, t_j)p_{fj}, \quad (4.21)$$

wobei $S(t)$ den Wechselkurs zum Zeitpunkt t bezeichnet.

¹London Interbank Offered Rate

4.2.2 Vorzeitig auflösbarer Differential Swap

Die Gegenseite hat hier die Möglichkeit, den Swap zu Beginn jeder Zinsperiode aufzulösen. Die Kosten für die Beendigung X werden in der ausländischen Währung fällig. Aus der einfachen europäischen Auszahlungsstruktur wurde hier eine amerikanische mit endlich vielen Auszahlungszeitpunkten, also bermudisch. In der vorletzten Periode (wenn zum letzten Mal die Entscheidung des Auflörens ansteht) haben wir folgende Endbedingung

$$V(t_{N-1}^-) = \min XS(t_{N-1}), P_d(t_{N-1}, t_N)p_N \quad (4.22)$$

Gehen wir in der Zeit weiter rückwärts, so sehen wir, dass die Endbedingung (4.22) weiterhin bis auf den Barwert der noch folgenden Zahlungen erfüllt ist. Wir erhalten somit

$$V(t_{N-1}^-) = \min XS(t_{N-1}), P_d(t_{N-1}, t_N)p_N \quad (4.23)$$

4.2.3 Ein erweitertes Vasicek-Modell

Die Erweiterung des Vasicek-Modells, die wir im folgenden betrachten werden, stammt ursprünglich von Simon Babbs. Wir verwenden allerdings die Version aus der Arbeit von Dempster und Hutton ([10]), die ebenfalls Cross-Currency Swaps bewerten.

Das Grundmodell besteht aus stochastischen Differentialgleichungen für die drei Faktoren

$$\frac{dP_d(t, T)}{P_d(t, T)} = [r_d(t) + \theta_d(t)\sigma_d(t, T)]dt + \sigma_d(t, T)dZ_d(t) \quad (4.24)$$

$$\frac{dP_f(t, T)}{P_f(t, T)} = [r_f(t) + \theta_f(t)\sigma_f(t, T)]dt + \sigma_f(t, T)dZ_f(t) \quad (4.25)$$

$$\frac{dS(t)}{S(t)} = [r_d(t) - r_f(t) + \theta_S(t)\sigma_S(t)]dt + \sigma_S(t)dZ_S(t), \quad (4.26)$$

wobei die σ_d , σ_f und σ_S die Bond- und Wechselkursvolatilitäten darstellt. Die stochastischen Terme Z_d , Z_f und Z_S sind Wiener Prozesse, die imperfekt über die Matrix $\Sigma := \{dZ_k dZ_l\}_{k,l=d,f,S} = \{\rho_{kl}\}_{k,l=d,f,S}$ korreliert sind. Wir schreiben nun die Prozesse in unabhängige Wiener Prozesse um:

$$dZ_k(t) = \sum_{j \in \{d,f,S\}} \alpha_{kj}(t)dW_j(t) \quad k = d, f, S \quad (4.27)$$

Die Matrix $\{\alpha_{kj}\}_{k,j \in \{d,f,S\}}$ ist dabei die Wurzel aus der Korrelationsmatrix Σ .

Die separable erweiterte Vasicek-Bedingung fordert, dass sich die deterministische Volatilität in der folgenden Form schreiben lässt:

$$\sigma_k(t, T) = [G_k(T) - G_k(t)]\lambda_k(t) \quad k = d, f \quad (4.28)$$

Dabei eliminieren wir jegliche Pfadabhängigkeit in der Zinsrate, die sich dann schreiben lässt als:

$$dr_k(t) = \left\{ \mu_k(t) - \frac{G_k''(t)}{G_k'(t)} [\mu_k(t) - r_k(t)] \right\} - G_k'(t)\lambda_k(t)[\theta_k(t)dt + dZ_k(t)],$$

wobei $k = d, f$ und

$$\mu_k(t) = F_k(0, t) + G_k'(t) \int_0^t [G_k(t) - G_k(s)]\lambda_k^2(s) ds$$

$F(a, b)$ ist dabei die aktuelle Rate zwischen den Zeitpunkten a und b .

Die Zinsrate ist also von der allgemeinen Form

$$dr(t) = (\alpha(t) - \beta(t)r(t))dt + \sigma dZ(t), \quad (4.29)$$

was einem Modell vom Typ Vasicek entspricht. α/β ist dabei der langfristige Erwartungswert, β die Rate, mit der die Zinsrate zu ihrem langfristigen Mittel zurückkehrt, und σ die Standardabweichung von r .

Wir können hier zusätzlich annehmen, dass die Funktionen G_k und λ_k folgende Form haben

$$G_k(t) = \frac{1 - e^{-\phi_k t}}{\phi_k} \quad k = d, f \quad (4.30)$$

$$\lambda_k(t) = e^{\phi_k t} \kappa_k(t) \quad k = d, f. \quad (4.31)$$

ϕ_k sind hier beliebige Konstanten und κ_k eine Kalenderfunktion.

Wie gerade gesehen kehrt der kurzfristige Zinssatz mit einer konstanten Rate $-G_k''(t)/G_k'(t)$ zu seinem Mittelwert zurück. Wie stark der kurzfristige Zinssatz variiert, steht in der Kalenderfunktion $\kappa_k(t)$, die hier gleich $G_k'(t)\lambda_k(t)$ ist.

Wendet man nun das Radon-Nikodym Theorem und den Satz von Girsanov an, so lassen sich die Variablen in folgender Form schreiben

$$P_d(t, T) = \frac{P_d(0, T)}{P_d(0, t)} e^{(G_d(T) - G_d(t))(X_d(t) - \int_0^t \lambda_d^2(s) (\frac{G_d(T) + G_d(t)}{2} - G_d(H)) ds)} \quad (4.32)$$

$$P_f(t, T) = \frac{P_f(0, T)}{P_f(0, t)} e^{(G_f(T) - G_f(t))(X_f(t) - \int_0^t h_f(t, T, H, s) ds)} \quad (4.33)$$

$$S(t) = \frac{P_f(0, t) \cdot S(0)}{P_d(0, t)} \cdot e^{-G_d(t)X_d(t) + G_f(t)X_f(t) + X_S(t) - \frac{1}{2} \int_0^t h_S(t, H, s) ds} \quad (4.34)$$

mit den Funktionen

$$h_f(t, T, H, s) := \lambda_f(s) \left(\frac{\sigma_f(s, T) + \sigma_f(s, t)}{2} + \sigma_S(s) \rho_{fS}(s) - \sigma_d(s, H) \rho_{df}(s) \right) \quad (4.35)$$

$$h_S(t, H, s) := \sigma_f^2(s, t) + \sigma_S^2(s) - \sigma_d^2(s, t) + 2\sigma_d(s, t)\sigma_d(s, H) + 2\sigma_f(s, t)\rho_{fS}(s)\sigma_S(s) - 2\sigma_f(s, t)\rho_{df}(s)\sigma_d(s, H) - 2\sigma_S(s)\rho_{dS}(s)\sigma_d(s, H) \quad (4.36)$$

wobei die Zustandsvariablen als

$$X_d(t) := \sum_{j \in \{d, f, S\}} \int_0^t \alpha_{dj}(s) \lambda_d(s) (dW_j(s) + \theta_j(s) ds) \quad (4.37)$$

$$X_f(t) := \sum_{j \in \{d, f, S\}} \int_0^t \alpha_{fj}(s) \lambda_f(s) (dW_j(s) + \theta_j(s) ds) \quad (4.38)$$

$$X_S(t) := \sum_{j \in \{d, f, S\}} \int_0^t [\alpha_{dj}(s) G_d(s) \lambda_d(s) - \alpha_{fj}(s) G_f(s) \lambda_f(s) + \alpha_{Sj}(s) \sigma_S(s)] (dW_j(s) + \theta_j(s) ds) \quad (4.39)$$

definiert sind.

4.2.4 Bewertung der Swaps

Die stochastischen Prozesse X_d , X_f und X_S haben zum Zeitpunkt $t \in [0, H]$ folgende Varianzen und Kovarianzen:

$$\text{var}[X_d(t)] = \int_0^t \lambda_d^2(s) ds \quad (4.40)$$

$$\text{var}[X_f(t)] = \int_0^t \lambda_f^2(s) ds \quad (4.41)$$

$$\text{var}[X_S(t)] = \int_0^t H_{SS}(s) ds \quad (4.42)$$

$$\text{cov}[X_d(t), X_f(t)] = \int_0^t H^{df}(s) ds \quad (4.43)$$

$$\text{cov}[X_d(t), X_S(t)] = \int_0^t H^{dS}(s) ds \quad (4.44)$$

$$\text{cov}[X_f(t), X_S(t)] = \int_0^t H^{fS}(s) ds, \quad (4.45)$$

wobei die Argumente der Integrale definiert sind als

$$\begin{aligned} H^{SS}(s) &:= G_d^2(s)\lambda_d^2(s) + G_f^2(s)\lambda_f^2(s) + \sigma_S^2(s) \\ &\quad - 2\rho_{df}(s)G_d(s)\lambda_d(s)G_f(s)\lambda_f(s) + 2\rho_{dS}(s)G_d(s)\lambda_d(s)\sigma_S(s) \\ &\quad - 2\rho_{fS}(s)G_f(s)\lambda_f(s)\sigma_S(s) \end{aligned} \quad (4.46)$$

$$H^{df}(s) := \rho_{df}(s)\lambda_d(s)\lambda_f(s) \quad (4.47)$$

$$H^{dS}(s) := \lambda_d(s)[G_d(s)\lambda_d(s) - \rho_{df}(s)G_f(s)\lambda_f(s) + \rho_{dS}(s)\sigma_S(s)] \quad (4.48)$$

$$H^{fS}(s) := \lambda_f(s)[\rho_{df}(s)G_d(s)\lambda_d(s) - G_f(s)\lambda_f(s) + \rho_{fS}(s)\sigma_S(s)] \quad (4.49)$$

$$(4.50)$$

Wir können nun den Wert eines Cross-Currency Swaps berechnen:

Satz 4.2.1 *Der Wert des Swaps im Inland $V := V(X_d, X_f, X_S, t)$ sei zweimal stetig differenzierbar. Dann kann der normalisierte Wert*

$$V^*(t) := \frac{V(t)}{P_d(t, T)} \quad (4.51)$$

durch die Lösung der folgenden Differentialgleichung beschrieben werden:

$$\begin{aligned} \frac{\partial V^*}{\partial t} + \frac{1}{2}\lambda_d^2 \frac{\partial^2 V^*}{\partial X_d^2} + \frac{1}{2}\lambda_f^2 \frac{\partial^2 V^*}{\partial X_f^2} + \frac{1}{2}H^{SS} \frac{\partial^2 V^*}{\partial X_S^2} \\ + H^{df} \frac{\partial^2 V^*}{\partial X_d \partial X_f} + H^{dS} \frac{\partial^2 V^*}{\partial X_d \partial X_S} + H^{fS} \frac{\partial^2 V^*}{\partial X_f \partial X_S} = 0 \end{aligned} \quad (4.52)$$

Die Randbedingungen, die wir hier künstlich implementieren müssen, werden wir im Abschnitt über die numerischen Ergebnisse der Anwendungen diskutieren.

Kapitel 5

Verwendeter Code und seine Wirkungsweise

5.1 Generelle Beschreibung

Ziel eines jeden numerischen Verfahrens sollte es sein, einen möglichst geringen Rechenaufwand zu haben. Dazu stehen bei gegebenem Verfahren prinzipiell zwei Möglichkeiten zur Verfügung: Zunächst können schnelle Löser gefunden werden, zum anderen kann eine geschickte Auswahl an Gitterpunkten vorgenommen werden, die die Funktion an unglatten Stellen höher auflöst. Wir werden hier unser adaptives Verfahren für dünne Gitter und für modifizierte Dünngitter beschreiben. Für die Adaptivität benötigen wir einen lokalen Fehlerschätzer, der die Unglattheit misst. Wir haben aber im Kapitel 2 über dünne Gitter gesehen, dass wir die hierarchischen Überschüsse als Indikatoren verwenden können, da diese für glatte Funktionen abfallen.

Die entsprechenden Normabschätzungen, die wir für den elliptischen Fall kennen (siehe Koster ([22])), werden hier nur teilweise behandelt. Eine exakte Konvergenztheorie für den parabolischen Fall steht noch aus.

Wir werden in unserem Code Interpolets verwenden, die wir bereits vorgestellt haben. Ferner ist die Diskretisierung des Differentialoperators von zentraler Bedeutung. Dieser soll mit möglichst hoher Konsistenzordnung stabil diskretisiert werden und schnell auszuwerten sein. Die Lösung des dann entstehenden Gleichungssystems sollte in möglichst geringer Zeit berechnet werden können.

Wir werden uns in diesem Abschnitt den Methoden zuwenden, die für unsere Beispiele Anwendung gefunden haben.

5.2 Problemstellung

Generell wollen wir partielle Differentialgleichungen mit Operatoren vom Typ

$$\mathcal{L} = -\frac{\partial}{\partial t} + \sum_{i=1}^d \sum_{j=1}^d a_{i,j}(\mathbf{x}) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial}{\partial x_i} + c(\mathbf{x}). \quad (5.1)$$

lösen.

Wir haben es also mit einem parabolischen Differentialoperator zu tun. Der große Unterschied unserer Diskretisierungen des Operators und der Lösung der Gleichung zu den in der Literatur gewöhnlicherweise verwendeten Verfahren besteht darin, dass wir in Raum und Zeit gleichzeitig diskretisieren.

Wir werden zunächst die Motivation für diese Vorgehensweise darlegen und dann die verschiedenen Diskretisierungen und ihre Konsistenzordnungen betrachten. Anschließend untersuchen wir die Stabilität des verwendeten Verfahrens.

5.2.1 Vorteile einer adaptiven Diskretisierung in Raumzeit

Die einfachste Möglichkeit, parabolische Gleichungen zu lösen, ist es, lediglich im Raum zu lösen und dann ein explizites Schrittweitenverfahren (z.B. Euler-Verfahren) zu verwenden. Pro Zeitschritt muss dann ein elliptisches Problem gelöst werden, das via dem Zeitschrittverfahren auf der rechten Seite des Problems in der nächsten Zeitscheibe eingeht. Dieses Verfahren lässt sich auf einfache Weise verbessern, indem man ein implizites Zeitschrittverfahren verwendet.

Auf diesen beiden Grundideen haben sich viele verschiedene Verfahren gebildet. Es werden adaptive Verfahren in Ort, h-p-Verfahren, adaptive Verfahren in der Zeit und auch adaptive Verfahren in Ort und Zeit verwendet. Dass jedoch die Gleichung in Raumzeit diskretisiert werden soll, ist relativ neu und wurde bisher nur in sehr wenigen Arbeiten (zum Beispiel Reisinger ([25])) verwendet.

Die Motivation für das Rechnen in Raumzeit besteht darin, dass die gesuchte Funktion glatt in Raum und Zeit ist, so dass adaptive Verfahren die Komplexität essentiell senken können. Die Raumzeit-Diskretisierung betrachtet die Zeit als eine weitere Dimension, die aber keine besondere Bedeutung im Vergleich mit den Ortsdimensionen für das Verfahren hat. Die Dimension der Problemstellung ist damit um eine Dimension höher als beim Zeitschrittverfahren. Die Rechenkomplexität steigt dadurch — zum Beispiel im Fall

des vollen Gitters um den Faktor n (Anzahl der Punkte pro Raumrichtung). Rechnet man in zwei Dimensionen mit einer Verfeinerung vom Level 10, so ist $n = 2^{10} = 1024$ und das volle Gitter hat 2^{20} Unbekannte. Erweitert man nun die Dimensionen um eine weitere, so erhöht sich die Anzahl der Unbekannten um den Faktor 1024 auf 2^{30} , was über einer Milliarde Gitterpunkte entspricht. Dies übersteigt die Rechen- und Speicherkapazitäten heutiger Rechner bei weitem.

Um die Rechenkomplexität auch für diese Problemstellungen in den Griff zu bekommen, haben wir bereits dünne Gitter eingeführt. Dünne Gitter geben uns überhaupt erst die Möglichkeit, die Glattheit in Raumzeit nutzen zu können. Dann aber sind mit Hilfe von adaptiven Verfahren gute Ergebnisse zu erwarten, da die Ansatzfunktionen anisotrop gleichzeitig in alle d Richtungen adaptiv wirken können. Trennt man Raum und Zeit, so ist die Schrittweite in die Zeitrichtung vorgegeben und unabhängig von den Schrittweiten im Ort.

Wir werden bei unseren Optionspreisaufgaben allerdings an Grenzen der dünnen Gitter stoßen. Daher führen wir in diesem Kapitel modifizierte Dünn-
gitter ein.

5.3 Übersicht über bisher bekannte Verfahren

Wir haben im Kapitel 2 über dünne Gitter erörtert, warum Interpolets für unsere Problemstellung, nämlich Funktionen mit hoher Regularität gut approximieren zu können, sehr gut geeignet sind, stellt sich aber immer noch die Frage, wie wir die partielle Differentialgleichung lösen wollen. Wichtig dabei ist, dass bei unserem Verfahren die Diskretisierung der Operatoren von der Wahl der Ansatzfunktionen unabhängig ist. Wir können so in jedem Teilschritt eine geeignete Methode suchen. Finite Differenzen sind dabei eine günstige Alternative, um die Diskretisierung des Differentialoperators vorzunehmen. Sie benötigt einen deutlich geringeren Rechenaufwand als beispielsweise eine Galerkin-Diskretisierung. Auch die rechte Seite läßt sich dabei einfacher auswerten, da in unserem Verfahren keine adaptiven Quadraturen benötigt werden.

In der Vergangenheit wurden bereits verschiedene Methoden vorgestellt, um partielle Differentialgleichungen auf dünnen Gittern zu lösen.

Zunächst einmal stehen Finite Elemente beziehungsweise Galerkin Verfahren zur Verfügung. Diese wurden in den Anfängen der dünnen Gitter nach der ursprünglichen Arbeit von Zenger ([33]) verwendet. Bungartz ([6]) löst so die dreidimensionale Poisson-Gleichung

$$\Delta u = f \tag{5.2}$$

auf dünnen Gittern mit Dirichlet-Randdaten. Balder ([2]) und Balder, Rüde, Schneider und Zenger ([3]) lösen so erstmalig parabolische Gleichungen.

Eine neue Vielfalt von Möglichkeiten eröffnete die Arbeit von Griebel ([18]), in der er die Verwendung von Finite Differenzen Methoden untersucht. Die Konsistenzordnung ist dabei im Falle dünner Gitter genau so hoch wie für die entsprechenden vollen Gitter. Schiekofer ([26]) wendete dann Finite Differenzen auf elliptische Problemstellungen an, wobei er allgemeine Gebietsformen erlaubte.

Koster ([22]) untersuchte elliptische Differentialgleichungen und löste die Navier-Stokes Gleichungen mit Zeitschrittverfahren unter Verwendung von Interpolets. Sein Verfahren bildet die Grundlage für diese Arbeit, die Koster's Ansatz auf parabolische Differentialgleichungen erweitert. Koster hat einen Ansatz gewählt, bei dem sich ein Petrov-Galerkin Verfahren, hier ein Kollokationsverfahren, von einem Finite Differenzen-Verfahren nur in dem univariaten Algorithmus (die Anwendung des Operators entlang einer Linie) unterscheidet.

Eine weitere Möglichkeit der Berechnung von parabolischen Gleichungen auf dünnen Gittern zeigt Reisinger ([25]), der eine Kombitechnik anwendet.

5.4 Normäquivalenzen

Die Grundlage für eine Konvergenztheorie sind die Normäquivalenzen. Diese sind grundsätzlich von der Form

$$\left\| \sum_{\mathbf{1},\mathbf{i}} u_{\mathbf{1},\mathbf{i}} \phi_{\mathbf{1},\mathbf{i}} \right\|_{\mathcal{H}^s}^2 \sim \sum_{\mathbf{1},\mathbf{i}} |u_{\mathbf{1},\mathbf{i}}|^2 \|\phi_{\mathbf{1},\mathbf{i}}\|_{\mathcal{H}^s}^2 \quad (5.3)$$

Wir erhalten aus der L_2 -Stabilität (siehe Lemma 2.3.1) die folgende Normäquivalenz (siehe Koster ([22])):

Satz 5.4.1 Normäquivalenz

Sei die univariate Multiskalenfunktion L_2 -stabil. Für $s \in (-M, \gamma)$ mit $\gamma = \sup\{s | \phi \in \mathcal{H}^s\}$, M der Anzahl der verschwindenden Momente und $u = \sum_{|\mathbf{1}|=l_0}^L \sum_{\mathbf{i}} u_{\mathbf{1},\mathbf{i}} \cdot \phi_{\mathbf{1},\mathbf{i}}$ gilt

$$\left\| \sum_{(\mathbf{1},\mathbf{i})} u_{\mathbf{1},\mathbf{i}} \phi_{\mathbf{1},\mathbf{i}} \right\|_{\mathcal{H}^s}^2 \sim \sum_{(\mathbf{1},\mathbf{i})} 4^{s \cdot \|\mathbf{1}\|_\infty} |u_{\mathbf{1},\mathbf{i}}|^2 \cdot 2^{-\|\mathbf{1}\|_1}. \quad (5.4)$$

Wir verwenden in unserem Programm Interpolets. Wir haben zwar den Vorteil sehr guter Approximation und ein gutes Verhalten in unseren Algorithmen, aber die Normäquivalenzen gelten nur sehr eingeschränkt. Daher

werden wir in Fällen, in denen wir die Abschätzungen benötigen, wie etwa bei der Fehlerschätzung und der Vorkonditionierung, auf Lifting-Interpolets transformieren. Dort gelten die entsprechenden Normäquivalenzen. Wir werden an den entsprechenden Stellen darauf zurückkommen.

Koster ([22]) beziffert die Sobolev-Regularität von Interpolets vom Grad $N = 4$ mit $\gamma = 2.171$. Für die zugehörigen Lifting-Interpolets nennt er eine Zahl von $\bar{\gamma} = 0.29$.

5.5 Petrov-Galerkin und Finite Differenzen

Der als Grundlage dieser Diplomarbeit verwendete Code hat die Möglichkeit, entweder ein Petrov-Galerkin Verfahren oder ein Finite Differenzen Verfahren zu verwenden. Der Petrov-Galerkin Ansatz führt in unserem Fall zu einem Kollokationsverfahren, da die dualen Multiskalenfunktionen aus dem Dirac-Funktional gewonnen werden. Es werden also Punktauswertungen vorgenommen. Das eindimensionale Poisson-Problem läßt sich wie folgt diskretisieren:

$$\left\langle \tilde{\phi}_{(\mathbf{l}, \mathbf{i})}, \sum_{(\mathbf{l}, \mathbf{i}) \in \mathbf{T}} \partial_{xx} \phi_{\mathbf{l}, \mathbf{i}} u_{\mathbf{l}, \mathbf{i}} \right\rangle = \left\langle \tilde{\phi}_{(\mathbf{l}, \mathbf{i})}, f \right\rangle \quad \forall (\mathbf{l}, \mathbf{i}) \in \mathbf{T}, \quad (5.5)$$

wobei sich die Lösung u darstellen läßt als $u = \sum_{(\mathbf{l}, \mathbf{i}) \in \mathbf{T}} \phi_{\mathbf{l}, \mathbf{i}} u_{\mathbf{l}, \mathbf{i}}$. Die dualen Multiskalenfunktionen $\tilde{\phi}_{(\mathbf{l}, \mathbf{i})}$ sind dabei Linearkombinationen von Dirac-Funktionalen. Weil aber alle hierarchischen Väter in \mathbf{T} enthalten sind (die adaptiven Gitter erfüllen die sogenannte Kegelbedingung in der Art wie wir sie aufgebaut haben), können wir schließen, dass es eine Bijektion zwischen dem diskreten Gitter und der Indexmenge \mathbf{T} gibt (Koster ([22]): Theorem 3.1). Damit gilt $\partial_{xx} u = f$ punktweise auf dem diskreten Gitter, und wir erhalten eine Kollokationsmethode. Gegenüber der Finite-Differenzen Methode hat das Kollokationsverfahren jedoch einige Nachteile. Zunächst haben wir geringe Einbußen im Kosten-Nutzen-Verhältnis. Außerdem haben wir starke Einschränkungen bei der Diskretisierung, da das Petrov-Galerkin Verfahren zu einem zentralen Finite-Differenzen Schema führt, das nicht stabil ist.

Aus diesem Grund wollen wir das Finite-Differenzen Verfahren verwenden. Finite Differenzen haben den Vorteil, dass die Diskretisierung einfach zu handhaben und sehr flexibel ist. Eine ausführliche Diskussion findet sich in Hutton ([20]). Auf dünnen Gittern stellt sich dabei jedoch die Herausforderung, dass das Finite-Differenzen Schema nicht direkt übertragbar ist, weil die Gitter um einen Knoten nicht äquidistant entfernt sind. Schiekofner ([26]) schlägt daher vor, die Funktion zunächst auf die nodale Basis zu

transformieren, dann auf diesem Teilgitter Finite Differenzen anzuwenden und anschließend wieder auf die gewünschte Basis-Darstellung zu wechseln.

Es genügt allerdings, gewisse Punkte einzufügen. Dazu zeigt Koster ([22]) folgendes Lemma:

Lemma 5.5.1 *Sei \mathbf{T} die Indexmenge für die Interpolet-Darstellung mit Interpolets vom Grad $N \geq 4$ auf von unserem Code generierten adaptiven Gittern (d.h. die Kegelbedingung ist erfüllt). Dann unterscheidet sich der Abstand zum linken Nachbarn vom Abstand zum rechten Nachbarn maximal um den Faktor 4 für alle Punkte in unserem adaptiven Gitter.*

Damit hält sich der Aufwand also stark in Grenzen.

5.5.1 Diskretisierung des Differentialoperators

Konsistenz

Koster ([22]) zeigt, dass wir den numerischen Fehler der diskreten Lösung $\tilde{u}_{\mathbf{T}}$ auf dem adaptiven Gitter $\Omega_{\mathbf{T}}$ durch einen Interpolationsfehler (in der \mathcal{H}^s -Norm) und einen Konsistenzfehler (in der \mathcal{H}^{s-t} -Norm) abschätzen können:

$$\|u - \tilde{u}_{\mathbf{T}}\|_{\mathcal{H}^s} \lesssim \|u - \mathcal{I}_{\mathbf{T}}u\|_{\mathcal{H}^s} + \|A_{\mathbf{T}}\mathcal{I}_{\mathbf{T}}u - \mathcal{I}_{\mathbf{T}}Au\|_{\mathcal{H}^{s-t}} \quad (5.6)$$

Der Parameter t stammt aus der $(s, s-t)$ -Stabilität, die wie folgt definiert ist:

Definition 5.5.1 *Ein Operator $A_{\mathbf{T}}$ heißt $(s, s-t)$ -stabil, wenn es eine von \mathbf{T} unabhängige Konstante $C_s > 0$ gibt, so dass*

$$C_s \|\phi_{\mathbf{T}}\|_{\mathcal{H}^s} \leq \|A_{\mathbf{T}}\phi_{\mathbf{T}}\|_{\mathcal{H}^{s-t}} \quad \forall \phi_{\mathbf{T}} \in \text{span}\{\phi_{\mathbf{l},\mathbf{i}} \mid (\mathbf{l}, \mathbf{i}) \in \mathbf{T}\} \quad (5.7)$$

Wir erhalten hier für unseren Differentialoperator folgende Konsistenz:

Satz 5.5.1 *Sei die diskrete Lösung $\tilde{u} = \sum_{\mathbf{l},\mathbf{i}} u_{\mathbf{l},\mathbf{i}} \phi_{\mathbf{l},\mathbf{i}}$ ($\mathbf{l}, \mathbf{i}) \in \mathbf{T}$ mit dem Verfahren berechnet worden. Das univariate Finite-Differenzen Schema habe die Konsistenzordnung $K-t$. Das Gitter von \tilde{u} enthalte alle Aufsetzpunkte für das Finite-Differenzen Schema. Die zugehörige Indexmenge \mathbf{T} enthalte alle Indizes, die*

$$2^{|\mathbf{l}|_1 K + l_i t} \cdot \sup_{\mathbf{x} \in \text{supp}\{\phi_{\mathbf{l},\mathbf{i}}\}} |D^{K-1}u(\mathbf{x})| > \delta \quad (5.8)$$

erfüllen, sowie deren Söhne in i -ter Koordinatenrichtung. Sei L das maximale Level von \mathbf{T} . Dann gilt:

$$\|D^{t\mathbf{e}_i} - A_{\mathbf{T}}\tilde{u}\|_{\Omega_{\mathbf{T}}} \lesssim L^{d-1}\delta \quad (5.9)$$

Nachdem wir die Konsistenz auf dünnen Gittern gezeigt haben, wissen wir, dass wir beliebige Finite-Differenzen Schemen direkt übertragen können und die gleichen Konvergenzordnungen erreichen wie im Falle voller Gitter, wenn der Operator nicht instabil wird.

Stabilität

Mit der Stabilität der Operatormatrix meinen wir folgendes:

Definition 5.5.2 *Der Operator $A_{\mathbf{T}}$ heißt stabil, falls es eine von \mathbf{T} unabhängige Konstante $C > 0$ gibt, so dass*

$$C\|\phi_{\mathbf{T}}\|_{\Omega_{\mathbf{T}}} \leq \|A_{\mathbf{T}}\phi_{\mathbf{T}}\|_{\Omega_{\mathbf{T}}} \quad \forall \phi_{\mathbf{T}} \in \text{span}\{\phi_{\mathbf{l},\mathbf{i}} \mid (\mathbf{l}, \mathbf{i}) \in \mathbf{T}\} \quad (5.10)$$

Zu den Stabilitätsaussagen führen wir hier keine Resultate auf und verweisen auf die in der Literatur aufgeführten Ergebnisse. Wir werden allerdings in der Diskussion über den Löser des Gleichungssystems nochmals auf die Stabilität zu sprechen kommen.

5.6 Adaptivität

Wie wir bereits erwähnt haben, adaptieren unsere Gitter in Raumzeit. Dass dadurch wirklich ein konvergentes Verfahren für die Lösung parabolischer partieller Differentialgleichungen entsteht, wird in dieser Arbeit theoretisch nicht gezeigt. Die entsprechenden Normäquivalenzen für den asymmetrischen Operator gibt es leider nicht. Dass wir jedoch Konvergenz mit unserem Verfahren erzielen können, werden wir in späteren Abschnitten zeigen.

Wir hatten im Abschnitt über dünne Gitter bereits gezeigt, dass wir durch das Weglassen von Gitterpunkten gegenüber dem klassischen vollen Gitter höhere Regularitätsanforderungen haben. Allerdings sind diese für Optionspreisaufgaben nicht erfüllt, da die Startwerte lediglich in $C^{0,1}$ sind.

5.6.1 Adaptive dünne Gitter versus modifizierte Dünn-gitter

Um die fehlende Regularität in den Anfangsdaten zu kompensieren, können wir eine andere Art der Adaptivität verwenden. Wir erlauben dabei mehr Gitterpunkte in unglatten Regionen und dünne Gitter in glatten Gebieten. Lokal kann sogar ein beinahe volles Gitter entstehen. Damit erhalten wir Konvergenz auch für Randdaten, die Knicke oder Sprünge haben.

Um die Interpolets aufzusetzen, müssen wir ein volles Gitter mindestens auf Level 3, das heißt mit Maschenweite 2^{-3} verwenden. Ein dünnes Gitter entsteht dann zwischen den Punkten des vollen Gitters, indem wir zu jedem Punkt die Söhne zunächst in die eine Richtung, dann in die andere einfügen. Wir sehen die Entstehung der dünnen Gitter in unserem Code in Abbildung 5.1.

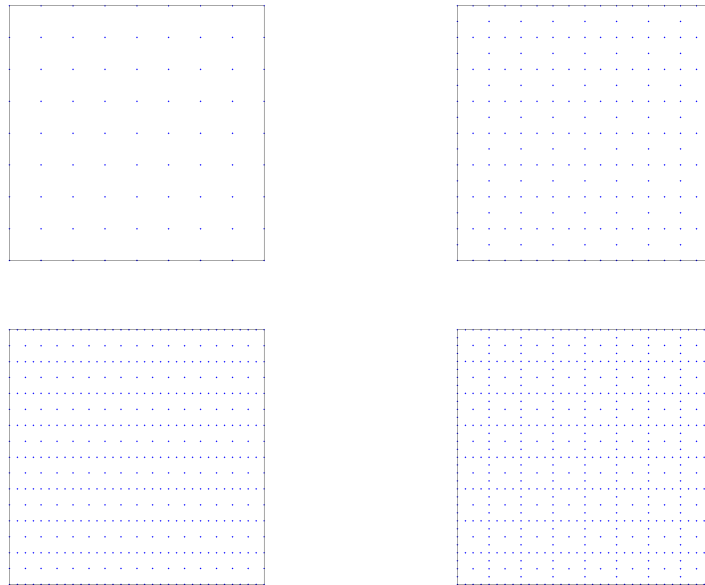


Abbildung 5.1: Oben links sehen wir ein volles Gitter vom Level 3, oben rechts ein darauf aufgesetztes dünnes Gitter vom Level 4. Verfeinern wir die Knoten aus dem Gitter oben rechts zunächst einmal in die eine Richtung, entsteht das Gitter unten links. Verfeinern wir dann die Punkte aus dem dünnen Gitter vom Level 4 in die andere Richtung, so entsteht das Gitter unten rechts — ein dünnes Gitter vom Level 5.

Um auch mit unglatten Randdaten oder rechten Seiten umgehen zu können, erlauben wir dem adaptiven Verfahren, auf weitere Gitterpunkte zuzugreifen zu können. Unsere Methode sieht vor, dass wir wie zuvor zunächst alle Söhne in die eine Richtung einfügen. Im nächsten Schritt werden jedoch sowohl die bestehenden als auch die neu hinzugefügten Gitterpunkte in die andere Richtung verfeinert. Im ersten Schritt erhalten wir das gleiche Gitter auf Level 4 (das aufgesetzte dünne Gitter). Im nächsten Schritt wird das Gitter jedoch feiner, wie wir in Abbildung 5.2 sehen können. Wir werden diese Art im folgenden als "modifizierte Dünngitter" bezeichnen.

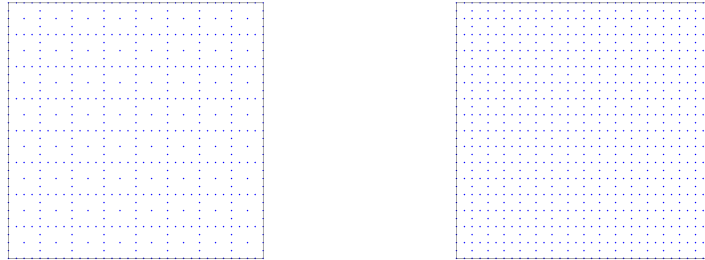


Abbildung 5.2: In der linken Abbildung sehen wir nochmals das dünne Gitter auf Level 5, rechts ist das modifizierte Dünngitter auf Level 5 dargestellt.

5.6.2 Fehlerschätzer

Wie wir im Abschnitt über dünne Gitter gesehen haben, können wir die Absolutbeträge der Koeffizienten für die Ansatzfunktionen als Fehlerschätzer verwenden. Natürlich gibt es Funktionen, die dann nicht optimal approximiert werden können.

In Bezug auf die Normabschätzungen haben wir gesehen, dass wir stabilere Ergebnisse haben, wenn wir Lifting-Interpolets verwenden. Wir können den Absolutbetrag der Koeffizienten entweder in der Lifting-Basis oder den Interpolets als Kriterium heranziehen. Wir werden sehen, dass — wie zu erwarten ist — die Lifting-Basis zu besseren Resultaten führt.

5.7 Sonstige Eigenschaften

5.7.1 Löser für Gleichungssysteme

Wir verwenden hier den Löser BiCGStab2. Sein Vorteil besteht darin, dass er auch für asymmetrische Operatoren, wie wir sie im Fall parabolischer Differentialgleichungen vorfinden, noch annehmbare Konvergenzraten aufweist. Allerdings sehen wir, dass für den Black-Scholes Operator keine Stabilität im Sinne von Definition 5.5.2 erreicht werden kann. Der Grund hierfür liegt darin, dass wir es mit einem degeneriert parabolischen Operator zu tun haben. Dass wir dadurch starke Einbußen in der Konvergenz des BiCGStab2 haben, sehen wir im nächsten Kapitel. Die Transformation auf die Wärmeleitungsgleichung hilft uns jedoch über diese Hürde hinweg.

5.7.2 Vorkonditionierung

Prinzipiell haben wir zwei verschiedene Arten der Vorkonditionierung implementiert. Wir können einerseits in Interpolet-Darstellung diagonal vorkonditionieren oder zuvor auf Lifting-Darstellung wechseln. Da wir nur im Falle der Lifting-Interpolets auf die Normäquivalenzen zugreifen können, macht sich die Transformation in einer schnelleren Konvergenz des BiCGStab2 bemerkbar. Der Grund hierfür ist, dass die L_2 -Stabilität durch das entsprechende Vorkonditionieren eine (beinahe) maschenweitenunabhängige Konditionszahl der vorkonditionierten Operatormatrix garantiert. Numerische Ergebnisse zum Lösen von Gleichungssystemen sehen wir im nächsten Kapitel.

5.7.3 Gebiete

Die Form des Gebietes ist auf Intervalle beschränkt. Die Ausdehnung, sprich die Länge, ist für jede Dimension getrennt einstellbar.

5.7.4 Parallelisierung

Der Code ist vollständig parallelisiert. Die Prozesse werden zwischen den verschiedenen Prozessoren aufgeteilt. Die numerischen Ergebnisse wurden auf Computern mit zwei Prozessoren und Rechnern mit vier Prozessoren berechnet.

5.7.5 Fehlermessung

Zur Fehlermessung stehen verschiedene Möglichkeiten zur Auswahl. Wir können den Fehler in verschiedenen Normen messen, gegen verschiedene Referenzlösungen und auf verschiedenen (Teil-)Gebieten. Zunächst einmal ist nach der Norm zu unterscheiden. Hierzu gehen wir zur nodalen Basis (der Knotendarstellung) $u(\cdot)$ über und berechnen die diskrete L_1 -, L_2 - und L_{\max} -Norm durch die folgenden Formeln:

$$\|u\|_{1,d} = \sum_{i=1}^N |u(\mathbf{x}_i)| \quad (5.11)$$

$$\|u\|_{2,d} = \sqrt{\sum_{i=1}^N (u(\mathbf{x}_i))^2} \quad (5.12)$$

$$\|u\|_{\max,d} = \max_i \{|u(\mathbf{x}_i)|\} \quad (5.13)$$

Besonders für unsere Optionspreisaufgaben ist es von großer Bedeutung, auf welchem Gebiet wir den Fehler messen. Wir wollen jegliche Effekte, die durch den freien Rand entstehen können, eliminieren. Da wir das Gebiet in der Zeitdimension beliebig lang wählen können, wollen wir in bestimmten Fällen auf der mittleren Zeitscheibe den Fehler messen. Numerische Ergebnisse zu dieser Möglichkeit sehen wir im nächsten Kapitel.

Haben wir dann den Fehler auf der mittleren Zeitscheibe berechnet, interessieren wir uns vor allem für die Region um den Ausübungspreis. Wie vorher bereits diskutiert, werden wir künstliche Randbedingungen implementieren müssen. Damit ist der Fehler am Rand gegenüber der wirklichen Lösung recht groß. Diesen Effekt wollen wir aber nur messen, wenn wir die Randbedingungen testen. Den wirklichen Fehler um den Ausübungspreis erhalten wir entweder, indem wir das Gebiet für die Fehlermessung weiter einschränken oder gegen eine mit dem Verfahren berechnete Referenzlösung messen, da sich dann die Fehler am Rand gegenseitig aufheben. Ersteres wollen wir hier nicht verfolgen. Bei einer diskreten Referenzlösung (auf feinerem Level) haben wir dieselben Randbedingungen und der Fehler am Rand wird eliminiert. Damit können wir die Konvergenz des Verfahrens messen — müssen aber auf die Wahl der Randbedingungen vertrauen. Diese können wir nur gegen eine analytische Lösung testen. Fehlermessungen sind sowohl für Messungen gegen analytische wie auch gegen diskrete Referenzlösungen implementiert.

Für die Fehlermessung werden wir die Lösung zunächst weiter verfeinern und die diskrete Funktion auf das feinere Gitter interpolieren. Dann messen wir den Fehler auf dem feineren Gitter, um genauere Werte zu erhalten.

5.7.6 Randbedingungen

Welche Anfangsdaten?

Mit den adaptiven modifizierten Dünngittern können wir Anfangsdaten aus dem Raum C^0 verwenden — sowohl für Dirichlet-, als auch für Neumann-Randdaten. Für echte dünne Gitter sind wir auf Anfangsdaten beschränkt, die stark differenzierbar sind. Das heißt, sie müssen so oft stetig differenzierbar sein, wie vom Operator gefordert wird.

Der Grund hierfür liegt in der Randbehandlung. Wir teilen die Funktion u mit ihren Randdaten in zwei Teile auf. Wir schreiben die Funktion als

$$u(\mathbf{x}, t) = u^b(\mathbf{x}, t) + u^i(\mathbf{x}, t), \quad (5.14)$$

wobei u^i die Funktion auf dem inneren Teil des Gebietes und u^b die Darstellung des Randes in Interpoletdarstellung bezeichnen. Anstatt wie gewohnt

die Differentialgleichung $\mathcal{L}u = f$ mit Randdaten u^b zu berechnen, können wir nun die Gleichung

$$\mathcal{L}u^i = f - \mathcal{L}u^b \quad (5.15)$$

lösen. Dabei haben wir die Randbedingungen direkt berücksichtigt. Wie wir aber an der rechten Seite der Gleichung sehen, muss der Rand stark differenzierbar sein.

Ist dies nicht der Fall, erhalten wir für adaptive dünne Gitter im allgemeinen keine Konvergenz. Wir werden im Kapitel über die numerischen Ergebnisse darauf zurückkommen. Aber auch in diesem Fall lassen sich Korrekturen anwenden, die Konvergenz auch für diesen Fall erzwingen. Diese Korrekturen sind nach der Art von Gietl und Zenger ([17]) und werden hier erstmals vorgestellt.

Korrekturen im regulären Fall

Wie wir bereits im Kapitel (3) über Optionen gesehen haben, sind die Glattheitsvoraussetzungen, die wir gerade diskutiert haben, nicht erfüllt. Die Lösung hat zum Endzeitpunkt einen Knick und ist daher nur $C^{0,1}$. Berechnen wir die Lösung numerisch mit dünnen Gittern, so zeichnet sich eine Singularität ab, die mit zunehmendem Level stärker wird. Numerische Experimente dazu sehen wir im Abschnitt 6.3.2.

Wir können die Anfangsdaten an der Stelle so manipulieren, dass die numerische Lösung wieder glatt ist. Dies erfordert sehr viel Geduld, da wir zwei Fehler gegeneinander balancieren müssen: die Änderung der Anfangsdaten muss gegen die gedämpfte Singularität abgewogen werden. Das Optimum ist dann erreicht, wenn die Fehler von gleicher Größe sind.

Im folgenden werden wir uns zwei verschiedene Möglichkeiten ansehen, wie die Anfangsdaten abgeändert werden können.

Glättung mit Peano-Kern

Zunächst betrachten wir ein Lemma, das uns erlaubt, später die Größe der Randkorrektur zu berechnen.

Lemma 5.7.1 *Integriert man eine Funktion u gegen die Hütchenfunktion mit Grundseite $2h$ und Höhe $1/h$, so erhält man den Stern der zentralen Finiten Differenzen mit dem Muster $[1 \ -2 \ 1]$.*

Beweis:

$$\int_{+\infty}^{-\infty} u''(x)h_h(x)dx = \underbrace{h'_h(x)u'(x)}_{=0} \Big|_{s-h}^{s+h} + \int_{s-h}^{s+h} h'_h(x)u'(x)dx$$

$$\begin{aligned}
&= \int_{s-h}^s h'_h(x)u'(x)dx + \int_s^{s+h} h'_h(x)u'(x)dx \\
&= h'_h(x)u(x)|_{s-h}^s + h'_h(x)u(x)|_s^{s+h} - \underbrace{\int_{s-h}^s h''_h(x)u(x)dx}_{=0} - \underbrace{\int_s^{s+h} h''_h(x)u(x)dx}_{=0} \\
&= \frac{1}{h^2}(u(s-h) - 2u(s) + u(s+h))
\end{aligned}$$

Aus diesem Ergebnis lässt sich folgern: $\int_{-\infty}^{+\infty} h_h(x)u(x)dx = 0$ für u konstant.
□

Offensichtlich gilt: $\int_{-\infty}^{+\infty} h_h(x)u(x)dx = 0$ falls u linear ist, denn $a(x-h) + b - 2(ax+b) + a(x+h) = 0$

In einem ersten Schritt integrieren wir also die Anfangsdaten $u_0(x) = \max\{x-s, 0\}$ gegen einen Peano-Kern:

$$G_1(y) := \int_{\Omega} u_0(x)(x-y)dx \quad (5.16)$$

Nun möchten wir wissen, wie stark wir die Knickstelle korrigieren müssen. Zur Berechnung dessen verwenden wir Lemma (5.7.1):

$$\int_{s-h}^{s+h} h_h(x)G_{lev}(x)dx = \frac{G_{lev}(s+h) - 2G_{lev}(s) + G_{lev}(s-h)}{h^2} = \frac{h}{3} \quad (5.17)$$

Glättung mit Gauss-Kern

Eine Glättung mit dem Gauss-Kern ist eine einfach zu implementierende Aufgabe. Da der Erwartungswert für einen Prozess, der einer geometrischen Brownschen Bewegung unterliegt, bereits die Integration gegen einen Gauss-Kern ist, können wir von einem späteren Zeitpunkt starten, wenn die Anfangsdaten ausgeglättet sind. Die Lösung ist bereits nach einem ε -Schritt in der Zeit analytisch, kann jedoch um die Knickstelle noch eine hohe Krümmung in Ortsrichtung aufweisen.

Analog zur Glättung mit dem Peano-Kern versucht man den Glättungsfehler gegenüber der Singularität zu äquilibrieren.

Rand auf das maximale Level verfeinern

Eine Methode, die sicher funktioniert und auch einfach zu implementieren ist, besteht darin, den Rand auf das höchste Level zu setzen. Hierbei ist sofort klar, dass natürlich drastische Verluste in der Effizienz zu verzeichnen sind.

Wie wir im nächsten Kapitel sehen werden, wird die Anzahl der Gitterpunkte dadurch noch weiter erhöht, dass mehrere Kopien aller Randpunkte im Inneren aufgesetzt werden müssen, damit der Träger der Interpolets definiert ist.

Kapitel 6

Numerische Ergebnisse

6.1 Einleitung

Bis jetzt haben wir den Aufbau dünner Gitter und die Modelle gesehen, die wir lösen wollen. Im letzten Kapitel haben wir den Code beschrieben, dessen Eigenschaften wir nun an verschiedenen Modellproblemen testen wollen. In erster Linie geht es darum, das beste Verfahren für unsere Problemstellungen zu finden. Wir werden einzelne Bausteine des Verfahrens variieren, während die restlichen Bausteine unverändert bleiben. So wollen wir herausfinden, welche Einstellungen am besten funktionieren.

Nachdem wir die Standardeinstellungen des Verfahrens aufgelistet haben, wenden wir uns den ersten Modellproblemen zu. Wir beginnen mit glatten Problemen, Gleichungen mit unglatten rechten Seiten und Problemen mit unglatten Anfangsdaten. In letzterem Fall wollen wir auch die Randkorrektur testen.

Im darauffolgenden Abschnitt untersuchen wir den Fehlerschätzer. Haben wir eine passende Norm als Adaptivitätskriterium gefunden, können wir die unterschiedlichen Diskretisierungen untersuchen. Übrig bleibt dann noch der Test des BiCGStab2, bevor wir das für unsere Zwecke geeignete Verfahren herausstellen können.

6.2 Wahl der Standardeinstellungen

Wir haben eine Menge an Parametern, die wir für die Rechnungen konstant halten wollen. Wenn nichts anderes im Text erwähnt ist, werden dies die Standardeinstellungen sein. Wir wollen im folgenden die Effekte zeigen, die auftreten, wenn man von den Standardeinstellungen abweicht.

Die folgenden Resultate sind mit einer Zeitdiskretisierung von der Konsistenzordnung zwei und einer Ortsdiskretisierung der Ordnung vier gerechnet, wenn nichts anderweitig genannt wird. Wie alle folgenden Resultate verwenden wir Interpolets mit einem polynomialen Exaktheitsgrad der Ordnung vier. Um sie auf dem niedrigsten Level aufstellen zu können, wird ein volles Gitter auf Level drei als Startgitter verwendet. Davon ausgehend verwenden wir im Normalfall das adaptive dünne Gitter. Wir unterscheiden hier zwischen vollen Gittern (mit $N = O(n^d)$ Punkten), regulären dünnen Gitter (wie im Kapitel 2 eingeführt), adaptiven Gittern (meint hier adaptive dünne Gitter) und modifizierten Dünngittern. Wir zeigen in diesem Kapitel, dass die Verfeinerungsstrategie der dünnen Gitter nicht ausreicht, um unglatte Probleme zu lösen. Wir werden dann modifizierte Dünngitter verwenden. Für beide Arten der adaptiven Gitter verwenden wir den L_∞ -Fehlerschätzer. Um den Fehler zu berechnen, verfeinern wir das Gitter zunächst zweimal, bevor wir gegen eine analytisch gegebene oder eine diskrete Referenzlösung zu messen. Wir messen den Fehler in der L_1 -, L_2 - und L_∞ -Norm.

Der Begriff "Freiheitsgrade" soll im folgenden die Punkte des Gitters betrachten, die nicht durch den Dirichlet-Rand vorgegeben sind. Die Summe der inneren Punkte, der freien Randpunkte und gegebenenfalls der Neumann-Randpunkte bildet dann die "Anzahl der Freiheitsgrade" (oder "dof" für "degrees of freedom"). Der Begriff "Gitterpunkte" (manchmal kurz " $|G|$ ") bezeichnet alle Punkte, die im Gitter sind, also auch die durch Dirichlet-Ränder vorgegebenen Punkte.

In manchen Fällen messen wir die Konvergenzrate des Verfahrens. Damit meinen wir folgendes: erhöhen wir die Anzahl der Gitterpunkte von 10^1 auf 10^2 , so würde sich der Fehler bei einer Rate r von, sagen wir, 10^{-a} auf 10^{-a-r} verringern. Diese Art der Messung ist aus unseren Graphen einfach abzulesen, indem man den Trend des Fehlerplots einzeichnet und dann die Steigung über Werte an der vertikalen Achse abliest.

Wir lassen den BiCGStab2 im folgenden (mit Ausnahme des Abschnitts 6.6 über die Konvergenz des Lösers) stets auskonvergieren. In unserem Fall wird so lange gelöst, bis das Residuum die Schranke von 10^{-14} unterschritten hat.

6.3 Adaptive versus reguläre resp. volle Gitter

Im Kapitel 2 haben wir den Unterschied zwischen vollen, regulären dünnen und adaptiven Gittern gesehen. Im folgenden wollen wir numerische Ergebnisse zu den verschiedenen Gitterarten betrachten. Die Anzahl der Gitterpunkte ist für den Fall der vollen und der regulären dünnen Gitter bekannt.

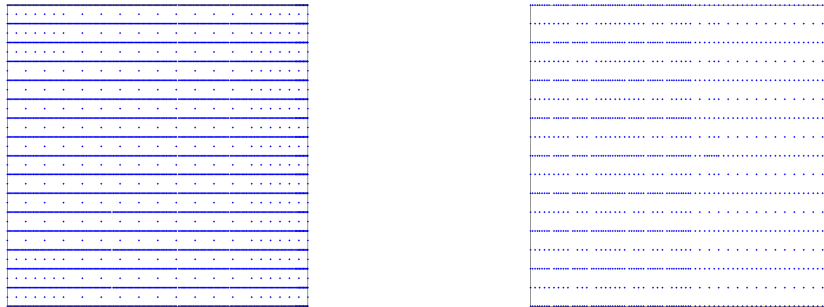


Abbildung 6.1: Links: modifiziertes Dünngitter für die Lösung aus (6.3). Rechts: adaptives dünnes Gitter.

Hier geht es also nur darum, die Konvergenzraten zu betrachten. Wir werden sehen, dass reguläre dünne Gitter nicht als Grundlage für alle Problemstellungen herangezogen werden können, für die volle Gitter funktionieren. In diesem Fall soll dann aber die Adaptivität der im vorigen Kapitel entwickelten modifizierten Dünngitter die fehlende Regularität auffangen und die gleichen Konvergenzraten wie im glatten Fall liefern.

Gerechnet wurde die Wärmeleitungsgleichung, gelöst wurde jeweils

$$Lu := u_t - \Delta u = f. \quad (6.1)$$

Die rechte Seite wurde als $f \equiv L\bar{u}$ gesetzt, wobei $\bar{u}(t, x)$ eine im gesamten Raum-Zeit-Gebiet definierte Referenzlösung ist. Als Anfangsbedingung setzen wir

$$u(0, x) = \bar{u}(0, x), \quad (6.2)$$

auf dem Rand Γ entsprechend $\bar{u}|_{\Gamma}$.

6.3.1 Glatte Lösung

Die Konvergenzraten auf der glatten Lösung dienen uns als Referenzmarke. Die hier erzielten Ergebnisse versuchen wir für geringere Regularitätsanforderungen zu erreichen. Die Lösung sei hier analytisch gegeben durch

$$u(t, x) = e^{-t} \cdot x \cdot (1 - x). \quad (6.3)$$

Damit kennen wir auch die passende rechte Seite $f = Lu$.

Beispielhaft führen wir nachfolgend ein adaptives Gitter für obige Problemstellung (Abbildung 6.1) und die entsprechenden Tabellen auf.

$ G $	dof	l1	l2	lmax	cgiter
81	56	2.477e-05	2.846e-05	4.928e-05	14
289	240	6.362e-06	7.266e-06	1.286e-05	34
1089	992	1.600e-06	1.827e-06	3.257e-06	85
4225	4032	4.005e-07	4.575e-07	8.175e-07	134
16641	16256	1.001e-07	1.144e-07	2.045e-07	238
66049	65280	2.504e-08	2.861e-08	5.114e-08	625
263169	261632	6.261e-09	7.151e-09	1.279e-08	639
1050625	1047552	1.564e-09	1.786e-09	3.188e-09	1021

Tabelle 6.1: Glatter Fall volles Gitter: Fehlerentwicklung (zu Lösung (6.3)).

$ G $	dof	l1	l2	lmax	cgiter
81	56	2.476e-05	2.839e-05	4.928e-05	14
225	176	6.362e-06	7.263e-06	1.286e-05	33
577	480	1.600e-06	1.827e-06	3.258e-06	73
1409	1216	4.004e-07	4.574e-07	8.175e-07	131
3329	2944	1.001e-07	1.144e-07	2.045e-07	221
7681	6912	2.504e-08	2.861e-08	5.123e-08	434
17409	15872	6.269e-09	7.157e-09	1.281e-08	663
38913	35840	1.564e-09	1.786e-09	3.207e-09	1071
86017	79872	3.854e-10	4.400e-10	7.904e-10	2277
188417	176128	7.287e-11	8.361e-11	1.550e-10	4166

Tabelle 6.2: Glatter Fall reguläres dünnes Gitter: Fehlerentwicklung (zu Lösung (6.3)).

$ G $	dof	l1	l2	lmax	cgiter
289	240	6.361e-06	7.263e-06	1.286e-05	35
504	426	1.594e-06	1.820e-06	3.269e-06	63
1314	1143	2.978e-07	3.602e-07	6.755e-07	191
2178	1905	1.002e-07	1.145e-07	2.045e-07	237
4851	4288	2.529e-08	2.899e-08	5.178e-08	372
9530	8447	6.454e-09	7.403e-09	1.309e-08	558
43876	34490	1.316e-09	1.597e-09	3.185e-09	501
69631	56927	3.197e-10	4.394e-10	1.243e-09	735
145990	96251	3.039e-10	4.309e-10	1.203e-09	545
200126	120271	2.201e-10	2.612e-10	4.975e-10	512

Tabelle 6.3: Glatter Fall adaptiv: Fehlerentwicklung für das modifizierte Dünngitter (zu Lösung (6.3)).

$ G $	dof	l_1	l_2	l_{\max}	cgiter
225	176	6.362e-06	7.263e-06	1.286e-05	33
416	337	1.596e-06	1.823e-06	3.262e-06	86
968	805	3.028e-07	3.632e-07	6.785e-07	180
1654	1387	1.004e-07	1.147e-07	2.044e-07	238
3564	3007	2.531e-08	2.899e-08	5.133e-08	383
7047	5966	6.605e-09	7.631e-09	1.516e-08	543
20558	17261	1.126e-09	1.378e-09	3.746e-09	533
23818	16734	1.370e-09	1.625e-09	3.147e-09	446
51837	41883	3.185e-10	4.379e-10	1.199e-09	773
55468	29113	4.981e-10	5.727e-10	9.644e-10	546
117960	79048	4.721e-10	5.846e-10	1.316e-09	522

Tabelle 6.4: Glatter Fall adaptiv: Fehlerentwicklung für das dünne Gitter (zu Lösung (6.3)).

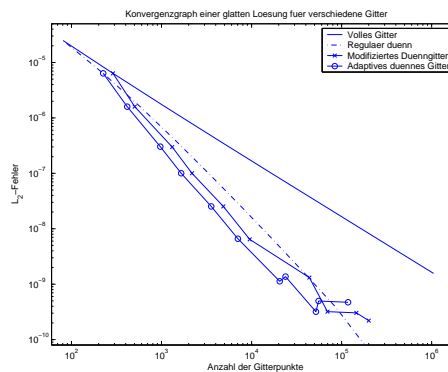


Abbildung 6.2: Konvergenzgraphen im glatten Fall für verschiedene Gittertypen.

Betrachten wir die Fehler, die gegen die gegebene analytische Referenzlösung berechnet wurden, sehen wir unterschiedliche Konvergenzraten für die verschiedenen Gittertypen (siehe Abbildung 6.2).

Abbildung 6.2 zeigt, dass die Konvergenzrate für adaptive Gitter im glatten Fall dadurch nichts verliert, dass modifizierte Dünngitter statt adaptiven dünnen Gittern eingesetzt werden. Die Konvergenzrate ist stets zwei außer für volle Gitter, wo die Rate bei eins liegt. Dies stimmt mit den theoretischen Überlegungen überein. Die Theorie sagt uns, dass die Adaptivität für glatte Probleme keine Vorteile bringt. Bis jetzt konnten wir die Konvergenzrate von zwei (für eine Diskretisierung der Ordnung zwei in der Zeit und vier im Ort) und dass diese Rate von allen Arten von dünnen Gittern erreicht wird. Warum wir die modifizierten Dünngitter eingeführt haben, sehen wir im nächsten Abschnitt, wenn wir keine Konvergenz für die dünnen Gitter mehr erhalten.

6.3.2 Unglatte Startdaten und Randkorrekturen

Entstehung einer Singularität

Wir betrachten nun verschiedene Arten an Unglattheiten. Zunächst wollen wir nun einen Knick in den Anfangsdaten zulassen, später betrachten wir dann unglatte rechte Seiten.

Versuchen wir also ganz naiv die Wärmeleitungsgleichung aus (6.1) mit homogener rechter Seite zu lösen, wobei die Anfangsdaten eine Knickstelle aufweisen. Speziell wurden hier die Startdaten als

$$f(x, 0) = 1 - |1 - 2x| \quad (6.4)$$

gewählt. Diese erfüllen nicht die Anforderung der starken Differenzierbarkeit, die wir im Abschnitt 5.7.6 erwähnt hatten. Dementsprechend bildet sich eine Singularität (siehe Abbildung 6.3, vergleiche mit wirklicher Lösung: Abbildung 6.4), die aus der zweiten zentralen finiten Differenz stammt, denn an der Knickstelle \hat{x} gilt:

$$\lim_{h \rightarrow 0} \frac{f(\hat{x} + h) - 2f(\hat{x}) + f(\hat{x} - h)}{h^2} = \lim_{h \rightarrow 0} \frac{-4h}{h^2} = -\infty \quad (6.5)$$

Die Finite Differenz ist an dieser Stelle im Grenzwert nicht bestimmt. Die Levelabhängigkeit des Peaks ist in Abbildung 6.3 zu erkennen.

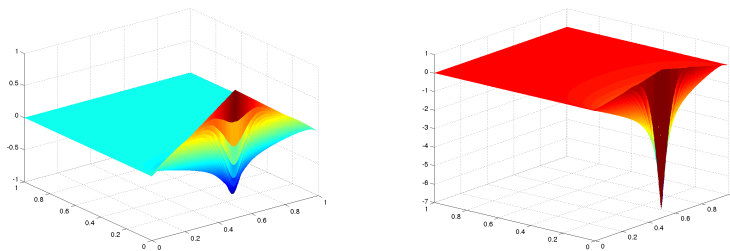


Abbildung 6.3: Berechnung von (6.4) auf den Leveln 7 und 12.

Verstärkt wird der Effekt dadurch, dass die Knickstelle auf einem Gitterpunkt liegt. Verschiebt man die Anfangsfunktion neben den Gitterpunkt (hier um $h/2$), so erhält man eine bessere Approximation, aber keine Konvergenz für das Lösen der partiellen Differentialgleichung. Ursächlich für die Singularität schien die Anwendung des Operators auf die Randdaten zu sein. Abbildung 6.5 zeigt, wie die Verschiebung von einem Gitterpunkt weg die Singularität abschwächen kann.

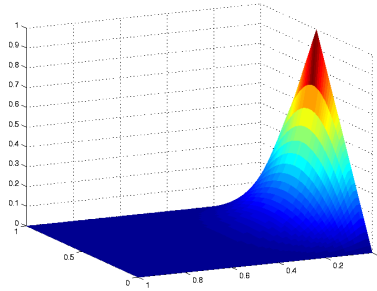
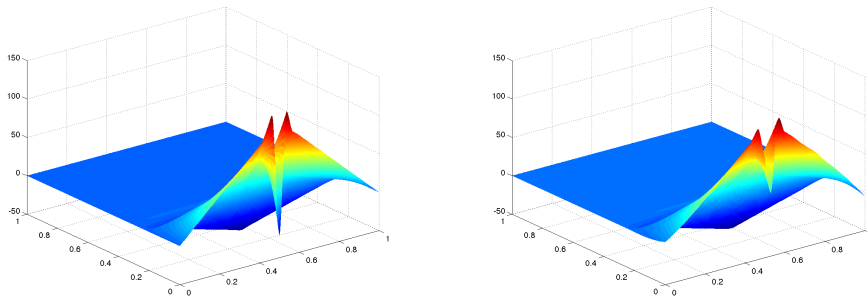


Abbildung 6.4: Tatsächliche Lösung des Randwertproblems

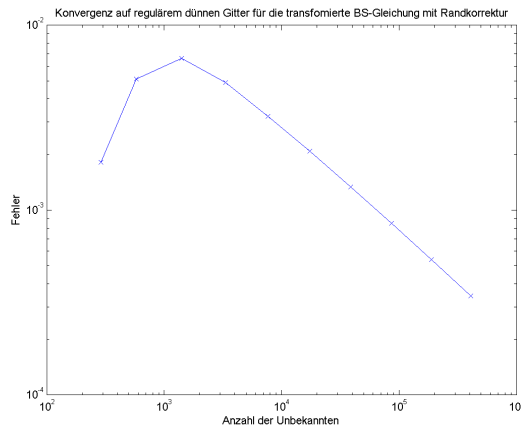
Abbildung 6.5: Anwendung des Laplace-Operators auf Level 5 auf die in der Interpolet-Basis dargestellten Randfunktion (entspricht u^b aus Abschnitt 5.7.6): links ohne Verschiebung, rechts mit einer Verschiebung um 0.02.

Zur Berechnung der Werte wurde das transformierte Black-Scholes Problem von einem Gitter der Größe $[-5, 5]$ im Ort und $[0, 1]$ in der Zeit auf das Einheitsintervall transformiert. Der Ausübungspreis liegt bei 0.3. Tabelle 6.7 zeigt, dass das adaptive dünne Gitter für das transformierte Black-Scholes Problem nicht konvergiert. Das bedeutet, dass die Adaptivität die Singularität nicht auflösen kann. Verwenden wir hingegen modifizierte Dünngitter, ebenfalls bei einer Diskretisierung der Ordnung vier in Ort und Zeit, so stellt sich Konvergenz mit einer Rate von vier ein. Hier sehen wir die Vorzüge des abgeänderten Gitteraufbaus.

Randkorrekturen

Im Abschnitt 5.7.6 hatten wir verschiedene Methoden zur Glättung der Randdaten theoretisch abgehandelt. Dort wurde bereits gezeigt, dass die Verschiebung um eine (levelabhängige) Konstante der Glättung mit dem Peano-Kern gleichkommt, da die Startdaten hier stückweise linear sind. Zu zeigen bleibt,

dass auch die Glättung mit dem Gauss-Kern erfolgreich ist. Hier wenden wir uns unserer Standardanwendung, der Berechnung der eindimensionalen europäischen Call-Option zu. Wir bezeichnen die analytische Lösung der Black-Scholes Gleichung mit $V(S, t)$. Wie bereits in Abschnitt 5.7.6 diskutiert, können wir $V(S, \varepsilon(\mathbf{1}))$ als geglättete Startdaten einsetzen.



Der Grund, warum die Korrektur nicht gleich in den konvergenten Bereich gelangt, liegt an der Levelabhängigkeit der Konstanten. Für kleine Levels ist die Konstante so groß, dass die Reduktion der Singularität durch die Verschiebung kaum weiteren Effekt hat. Verringert man also die Konstante von Level 3 auf Level 4, so erhöht sich die Singularität stärker als der Fehler durch die Verschiebung verringert werden kann. Es ist ebenfalls interessant zu sehen, dass sich die Konvergenzrate deutlich verschlechtert hat. Daher werden wir später nicht mehr mit Randkorrekturen arbeiten, sondern uns auf modi-

Adaptives dünnes Gitter			Modifiziertes Dünngitter		
$ G $	L_2 -Fehler	Raten	$ G $	L_2 -Fehler	Raten
426	1.427e-02	2.780e+00	681	1.078e-03	-1.231e+00
692	1.700e-02	-3.487e+00	1096	5.998e-04	-1.263e+00
1361	1.400e-02	-9.204e+00	1428	4.294e-04	-3.657e+00
2300	1.322e-02	-6.544e+01	1558	3.123e-04	-3.476e+00
3603	1.313e-02	-5.505e+00	1886	1.607e-04	-1.015e+00
5743	1.207e-02	-1.751e+00	2215	1.365e-04	-4.683e-01
6419	1.132e-02	5.748e+00	2303	1.341e-04	-2.754e-01
8754	1.195e-02	-1.680e+00	2959	1.251e-04	-6.667e+00
9565	1.134e-02	5.985e+00	3368	5.278e-05	-7.611e+00
13126	1.195e-02	-1.813e+00	6594	1.919e-05	-2.953e+00
14436	1.134e-02	4.299e+01	9157	7.276e-06	-2.742e+00
20420	1.143e-02	1.175e+00	12539	3.073e-06	-4.498e-01
24183	1.320e-02	-1.665e+00	17627	2.637e-06	-3.547e+00
29946	1.161e-02	0.000e+00	24899	7.745e-07	0.000e+00

Tabelle 6.7: Direkter Vergleich des dünnen Gitters mit dem modifizierten Dünngitter, wenn die rechte Seite unglatt ist.

fizierte Dünngitter verlassen, da wir dort die gewünschten Konvergenzraten erhalten. Dass dies so ist, sehen wir in Tabelle 6.7.

Die letzte Methode, die es noch zu testen gilt, löst den Rand nicht adaptiv auf, sondern setzt die Randfunktion auf einem Gitter des höchsten Levels auf. Das dabei entstehende Gitter in Abbildung 6.6 zeigt, dass hier nicht die gewünschte Konvergenz erzielt werden kann.

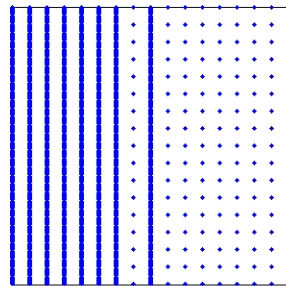


Abbildung 6.6: Adaptives dünnes Gitter, wenn der Rand auf das höchste Level verfeinert wird. Die Punkte im Inneren werden benötigt, um die Interpolets aufzusetzen.

6.3.3 Unglatte rechte Seite

Wir werden in diesem Abschnitt zwei Versuche machen: gerechnet wird das Problem

$$-\frac{\partial u}{\partial t}(x, t) + \Delta u(x, t) = f_i(x, t) \quad i = 1, 2, \quad (6.6)$$

mit unglatten Startdaten (wie bei Optionspreisen), wobei die rechte Seite jeweils unglatt ist. Wir testen hier einerseits eine Funktion mit Knick und andererseits mit Sprung. Ziel der Aufgabe ist dabei herauszufinden, ob die Konvergenzgeschwindigkeit gegenüber dem glatten Fall beeinträchtigt wird oder nicht. Interessant ist überdies, ob die Adaptivität die Knick- resp. Sprungstelle entsprechend findet und auflöst. Da wir Unglattheiten auflösen müssen, verwenden wir modifizierte Dünngitter.

Wir betrachten die Funktionen

$$f_1(x, t) = \max\{t - 1/2, 0\} \quad (6.7)$$

$$f_2(x, t) = \begin{cases} 0 & \text{wenn } t < 1/2 \\ 1 & \text{wenn } t \geq 1/2 \end{cases} \quad (6.8)$$

Die Rechnungen wurden mit den Standard-Einstellungen und einer Zeit- und Ortsdiskretisierung der Ordnung 4 gerechnet.

G	dof	l1	l2	lmax
289	240	3.312e-04	1.375e-03	2.578e-02
331	276	3.317e-04	9.001e-04	1.881e-02
682	590	5.347e-05	1.041e-04	2.129e-03
880	776	2.890e-05	4.778e-05	9.766e-04
1677	1509	1.011e-05	1.658e-05	4.883e-04
2296	2082	4.592e-06	9.354e-06	1.130e-04
3229	2930	3.143e-06	6.105e-06	4.872e-05
4237	3860	1.060e-06	1.841e-06	2.427e-05
5133	4706	5.342e-07	9.435e-07	1.146e-05
6138	5653	2.770e-07	5.991e-07	5.126e-06
7407	6889	1.348e-07	2.169e-07	4.036e-06
9550	8928	5.645e-08	9.410e-08	9.085e-07
11651	10957	2.651e-08	4.839e-08	3.634e-07
14284	13544	8.039e-09	1.232e-08	4.512e-08
21973	21204	2.711e-10	4.887e-10	2.647e-09

G	dof	l1	l2	lmax
289	240	4.189e-03	6.342e-03	2.348e-02
480	405	1.710e-03	2.620e-03	2.157e-02
901	779	3.852e-03	8.922e-03	6.461e-02
1549	1355	2.415e-04	4.101e-04	2.426e-03
2435	2147	1.423e-04	2.933e-04	1.574e-03
3275	2917	3.509e-03	9.043e-03	6.989e-02
3523	3159	2.792e-05	4.314e-05	4.883e-04
4060	3678	4.525e-06	8.917e-06	1.105e-04
4941	4514	4.225e-06	6.606e-06	4.614e-05
6050	5603	1.990e-06	3.120e-06	2.211e-05
7147	6672	5.271e-07	9.963e-07	1.181e-05
9651	9085	5.436e-07	8.301e-07	5.491e-06
12230	11632	8.873e-08	1.838e-07	4.043e-06
15253	14561	5.097e-08	9.604e-08	9.439e-07
19610	18841	8.240e-08	1.121e-07	4.290e-07
23521	22752	2.836e-09	4.731e-09	3.853e-08
27632	26863	1.975e-09	3.063e-09	1.463e-08
31642	30873	1.721e-09	3.260e-09	1.782e-08
34459	33690	3.463e-10	5.151e-10	2.366e-09

Tabelle 6.8: Konvergenztabelle für unglatte rechte Seiten: oben f_1 , unten f_2

Die Ergebnisse zeigen, dass kein Verlust in der Konvergenzgeschwindigkeit gegenüber dem glatten Fall zu verzeichnen ist. Das heißt, wir erhalten Konvergenz der Ordnung 4. In Abbildung 6.7 sehen wir auf der rechten Seite das modifizierte Dünngitter der Lösung, wenn wir die Sprungstelle einfügen und zusätzlich einen Knick in den Startdaten (wie bei Black-Scholes) auf den Punkt 0.5 legen. Die Konvergenzrate wird dabei allerdings nicht beeinträchtigt.

6.4 Fehlerschätzer

6.4.1 Interpolet- versus Lifting-Fehlerindikator

Wie wir bereits im Kapitel zum verwendeten Code diskutiert haben, können wir den Fehlerschätzer entweder auf die Koeffizienten in der Interpolet-Basis oder auf die Wavelet-Basis nach Lifting anwenden. In diesem Abschnitt soll untersucht werden, welche der zwei Methoden wirksamer ist. Da der Fehlerschätzer die Anzahl der hinzuzufügenden Punkte bestimmt, wirkt er direkt auf die Konvergenzrate: werden "unnötige" Gitterpunkte hinzugefügt, also Punkte, die nicht zur Fehlerreduktion beitragen, so verschlechtert sich die

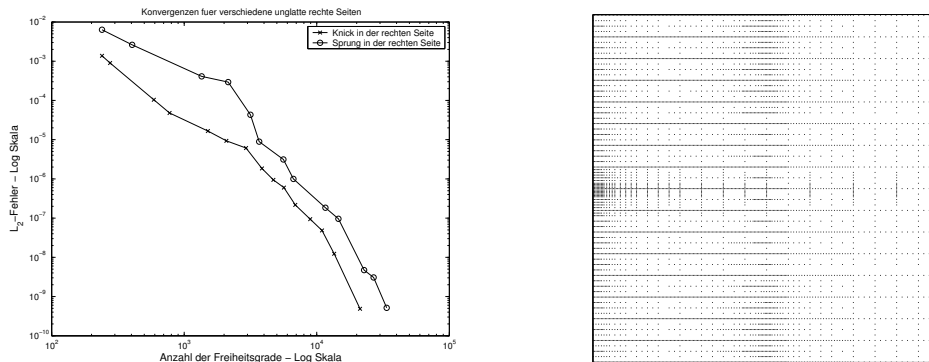


Abbildung 6.7: Konvergenzgraph für Sprung und Knick in der rechten Seite und zugehöriges modifiziertes Dünngitter.

Konvergenzrate.

Auf der anderen Seite muss der Fehlerschätzer unglatte Stellen schnell erkennen können und sie dann entsprechend auflösen. Wie wir hier sehen werden, macht es wenig Unterschied, welcher Fehlerschätzer für glatte Fälle verwendet wird. Für unglatte Anfangsdaten ergeben sich dann aber Unterschiede.

Wir werden im folgenden die modifizierten Dünngitter verwenden. Für die adaptiven dünnen Gitter ergeben sich jedoch die gleichen Resultate.

Das Modellproblem ist wie in (6.1), nur dass hier die Lösung

$$u(t, x) = e^{-5 \cdot (t+x)} \quad (6.9)$$

verwendet wird.

Messen wir nun den Fehler durch Betrachten der Koeffiziente in der Lifting-Wavelet Basis, so erhalten wir die aus Abschnitt 6.3.3 bekannten Raten β , da wir mit Konsistenzordnungen von 4 in Raum und Zeit diskretisieren:

p	∞	8	4	3	2	1
β	3.1318	3.1197	3.1976	2.8141	3.1382	3.2549

Ähnliche Raten erhalten wir aber auch, wenn wir den Koeffizienten aus der Interpolet-Basis heranziehen:

p	∞	8	4	3	2	1
β	2.9315	3.3446	3.1763	3.0883	3.1843	3.0145

Die Rechnung ohne Lifting erzielt zwar gute Raten, braucht aber lange zum Rechnen und erreicht sehr schnell das maximale Level.

Um nun einen Unterschied zwischen den unterschiedlichen Fehlerschätzern zu erhalten, müssen wir wie im Abschnitt über das Liften beim Fehlerschätzen die Regularität der Lösung herabsetzen.

Hierzu verwenden wir die transformierte zweidimensionale Black-Scholes Gleichung. Als Parameter verwenden wir eine Volatilität $\sigma = 0.1$ und eine Zinsrate $r = 0.04$. Der Ausübungspreis liegt bei $E = 0.3$. Dabei erhalten wir unglatte Anfangsdaten. Die Konvergenz des Verfahrens für verschiedene Fehlerschätzer sieht nun anders aus.

Im Falle nichtglatter Daten erhalten wir mit Lifting:

p	∞	8	4
β	2.0050	1.7565	1.5506

Wir rechneten hier mit einer Konsistenzordnung von 4 im Ort und 2 in der Zeit. Die Raten liegen auch hier bei 4, wenn wir mit Konsistenzordnungen von jeweils 4 diskretisieren, wie wir im nächsten Abschnitt sehen werden. Das bedeutet, dass wir den Fehlerschätzer benötigen, der am stärksten auf Unglattheiten reagiert. Für die Normen L_3 und abwärts stellt sich erratic behavior ein, wie wir in Abbildung 6.8 sehen können. Die Konvergenztabeln sind in Tabelle 6.12 zu sehen.

L_∞ Fehlerschätzer:			L_8 Fehlerschätzer:		
G	L2-Fehler	Rate	G	L2-Fehler	Rate
1289	5.355e-06	1.289e+03	1043	1.033e-05	-3.317e+00
1717	1.945e-06	-3.532e+00	1475	3.273e-06	-2.862e+00
2438	6.764e-07	-3.013e+00	1922	1.534e-06	-3.460e+00
3239	2.399e-07	-3.648e+00	2674	4.894e-07	-3.388e+00
4333	1.121e-07	-2.614e+00	3567	1.844e-07	-2.878e+00
5957	3.976e-08	-3.258e+00	4685	8.413e-08	-3.282e+00
7824	1.576e-08	-3.395e+00	6305	3.174e-08	-7.990e-02
10650	6.051e-09	-3.104e+00	8376	3.103e-08	-6.178e+00
14209	2.153e-09	-3.585e+00	11299	4.881e-09	-4.017e+00
18750	9.402e-10	-2.987e+00	14933	1.593e-09	-2.473e+00
25672	3.740e-10	-2.934e+00	19536	8.194e-10	-2.582e+00

L_4 Fehlerschätzer:			L_3 Fehlerschätzer:		
G	L2-Fehler	Rate	G	L2-Fehler	Rate
893	1.880e-05	-3.206e+00	860	2.464e-05	-3.374e+00
1251	6.378e-06	-3.760e+00	1143	9.435e-06	-4.117e+00
1623	2.396e-06	-2.734e+00	1425	3.806e-06	-2.803e+00
2212	1.028e-06	-3.639e+00	1932	1.621e-06	-3.587e+00
2835	4.167e-07	-3.356e+00	2624	5.408e-07	-3.384e+00
3821	1.530e-07	-3.153e+00	3435	2.174e-07	-3.082e+00
5059	6.316e-08	-3.501e+00	4501	9.450e-08	-3.178e+00
6739	2.315e-08	-3.267e+00	6083	3.629e-08	-3.793e+00
8874	9.419e-09	-2.708e+00	7958	1.310e-08	-2.972e+00
11891	4.264e-09	-4.344e+00	10342	6.012e-09	-3.240e+00
15655	1.291e-09	-1.821e+00	13805	2.359e-09	-3.158e+00
20325	8.027e-10	-1.367e+00	17887	1.041e-09	-1.323e+00
27238	5.379e-10	-1.382e+00	23657	7.189e-10	-2.213e+00

L_2 Fehlerschätzer:			L_1 Fehlerschätzer:		
G	L2-Fehler	Rate	G	L2-Fehler	Rate
654	4.382e-05	-2.717e+00	385	3.742e-04	-3.124e+00
893	1.880e-05	-2.948e+00	505	1.603e-04	-4.210e+00
1197	7.926e-06	-3.867e+00	622	6.667e-05	-3.240e+00
1511	3.220e-06	-2.832e+00	802	2.926e-05	-2.202e+00
1923	1.627e-06	-3.402e+00	1063	1.574e-05	-4.276e+00
2616	5.709e-07	-3.419e+00	1349	5.681e-06	-3.771e+00
3461	2.192e-07	-3.163e+00	1691	2.423e-06	-3.022e+00
4347	1.066e-07	-3.182e+00	2077	1.302e-06	-1.316e+00
5905	4.023e-08	-7.157e-01	2854	8.568e-07	-6.115e+00
7723	3.320e-08	-6.258e+00	3627	1.979e-07	-2.853e+00
9872	7.144e-09	-3.081e+00	4581	1.016e-07	-2.958e+00
13089	2.996e-09	-3.451e+00	5877	4.864e-08	-3.851e+00
16981	1.220e-09	-2.770e+00	7591	1.815e-08	-3.268e+00
21721	6.169e-10	-2.525e+00	9532	8.626e-09	-2.797e+00
			12101	4.425e-09	-4.665e+00
			15611	1.349e-09	-3.699e+00
			19845	5.552e-10	0.000e+00

Tabelle 6.10: Konvergenztabelle für glatte Anfangsdaten mit unterschiedlichen Normen als Fehlerschätzer.

L_{max} Fehlerschätzer			L_4 Fehlerschätzer		
G	L2-Fehler	Rate	G	L2-Fehler	Rate
681	1.078e-03	-1.231e+00	854	4.671e-03	5.415e-03
1096	5.998e-04	-1.263e+00	1023	4.675e-03	-2.018e+02
1428	4.294e-04	-3.657e+00	1034	5.403e-04	7.158e-01
1558	3.123e-04	-3.476e+00	1376	6.630e-04	-1.977e+01
1886	1.607e-04	-1.015e+00	1429	3.140e-04	-6.400e-01
2215	1.365e-04	-4.683e-01	1780	2.729e-04	4.432e+00
2303	1.341e-04	-2.754e-01	2087	5.524e-04	-5.893e+00
2959	1.251e-04	-6.667e+00	2648	1.358e-04	-3.154e+00
3368	5.278e-05	-7.611e+00	2894	1.026e-04	8.295e-01
3667	2.763e-05	6.009e+00	3837	1.297e-04	1.798e+308
3952	4.332e-05	-2.764e+00	3837	1.297e-04	-9.437e+00
4898	2.393e-05	-7.438e-01	4648	2.123e-05	8.237e-01
6594	1.919e-05	-2.953e+00	7085	3.005e-05	7.587e-01
9157	7.276e-06	-2.742e+00	9581	3.778e-05	-6.465e+00
12539	3.073e-06	-4.498e-01	12137	8.191e-06	2.726e-01
17627	2.637e-06	-3.547e+00	18181	9.145e-06	8.488e-01
24899	7.745e-07	1.500e+00	23859	1.152e-05	-5.700e+00

L_8 Fehlerschätzer		
G	L2-Fehler	Rate
434	1.884e-03	-2.853e+00
523	1.106e-03	2.526e+00
927	4.696e-03	-5.543e+00
1320	6.621e-04	-4.457e+00
1618	2.673e-04	-1.877e+00
2337	1.341e-04	-1.022e-01
3665	1.280e-04	-2.204e+00
4924	6.677e-05	-4.510e+00
6491	1.920e-05	-8.795e-01
9391	1.388e-05	-4.679e+00
12695	3.386e-06	-1.792e+00

Tabelle 6.12: Konvergenztabelle für unglatte Anfangsdaten mit unterschiedlichen Normen als Fehlerschätzer.

Als Beispiel für erratisches Verhalten soll hier die L_3 -Norm herangeführt werden.

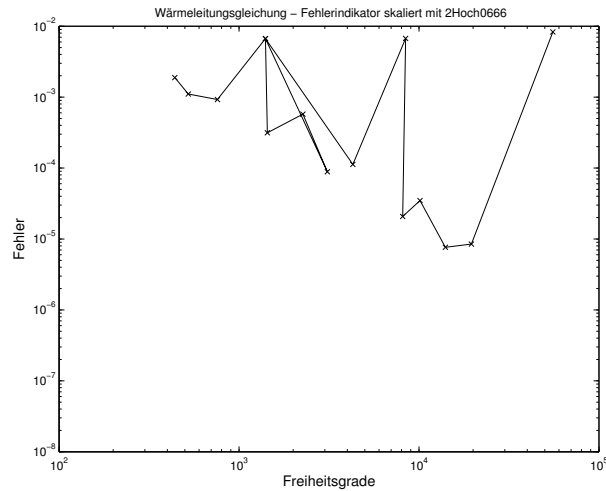


Abbildung 6.8: Erratisches Verhalten ab einer Norm mit $p = 3$.

Wenden wir uns nun dem nichtglatten Fall im Falle eines Fehlerschätzers zu, der nicht auf Lifting Koeffizienten beruht. Dann erhalten wir keine Konvergenz — egal welche Norm wir für den Fehlerschätzer erhalten.

6.5 Konvergenzraten für verschiedene Zeitdiskretisierungen

Zunächst betrachten wir die Konvergenzraten für den glatten Fall. Wie gewöhnlich rechnen wir hier die Wärmeleitungsgleichung und messen gegen eine analytisch berechnete Referenzlösung $u(x, t)$. Anschließend wenden wir uns unglatten Anfangsdaten und rechten Seiten zu und werden feststellen, dass unser Verfahren die gleichen Konvergenzraten produzieren wird. Der Fehlerschätzer springt jeweils auf L_∞ -Abweichungen an.

Wir werden in diesem Abschnitt mit adaptiven Verfahren stets modifizierten Dünngitter bezeichnen. Da wir unglatte Anfangsdaten verwenden werden, würden wir andernfalls keine Konvergenz erhalten.

Zeit	Ort	voll	regulär	adaptiv
1	2	0.55	0.91	0.8940
1	4	0.50	0.84	0.9140
2	2	1.05	1.60	1.9073
2	4	1.00	1.92	2.0104
4	2	1.01	1.69	2.6517
4	4	1.94	3.26	3.2941

Tabelle 6.13: Konvergenzraten im glatten Fall für volle und reguläre Gitter sowie modifizierte Dünngitter.

6.5.1 Glatte rechte Seite

Wir betrachten die Lösung auf dem adaptiven dünnen Gitter und auf dem vollen Gitter. Die Referenzlösung wurde hier mit

$$u(x, t) = \sin(\pi t) \cdot x(1 - x) \quad (6.10)$$

gewählt und die rechte Seite dementsprechend. Wir starten auf Level 4 und erhöhen es bis Level 12. Die gemessenen Konvergenzraten finden wir in Tabelle 6.13.

Die Ergebnisse zeigen, dass die Raten, die wir aufgrund der Theorie erwarten würden, auch wirklich erzielt werden. Die Beschränkung für die Konvergenzrate ist jeweils die kleinere der beiden Konvergenzordnungen. Dies ist nicht weiter überraschend, wenn man sich vor Augen führt, dass wir in Raumzeit keinen Unterschied zwischen den Dimensionen machen. Die Adaptivität kann die Rate zwar verbessern, allerdings reagiert sie mit unserem Fehlerschätzer nicht wesentlich auf die unterschiedlichen Diskretisierungsordnungen. In Abbildung 6.9 sehen wir, dass keine großen Unterschiede zwischen den Gittern bestehen, also insbesondere keine Vergrößerungen in eine der Richtungen auftritt. Gerechnet wurde hier das Modellproblem mit Lösung (6.9).

Dass im Falle einer Orts- und Zeitdiskretisierung der Konsistenzordnung 4 im regulären Fall nicht wirklich eine Konvergenzordnung von 4 erreicht wird, liegt am logarithmischen Faktor, den wir uns mit den dünnen Gittern eingefangen haben. Wie bereits vorher erläutert, ist der Faktor nur in der Präasymptotik zu sehen und würde mit steigendem Level verschwinden.

6.5.2 Unglatte Anfangsdaten

Als Standard-Modell für unglatte Anfangsdaten verwenden wir hier erneut das transformierte Black-Scholes Problem, da wir hierfür auf einfache Weise eine Referenzlösung analytisch angeben können.

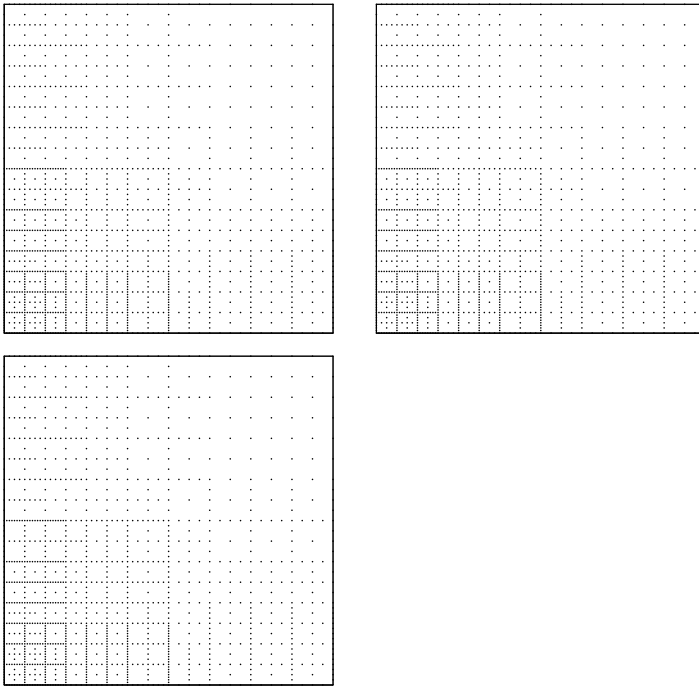


Abbildung 6.9: Oben links ist das Gitter für Konsistenzordnungen Zeit 2, Ort 2; oben rechts sehen wir Zeit 4, Ort 2; und unten links entsprechend Zeit 2, Ort 4

Legen wir den Knick auf den Gitterpunkt an der Stelle $x = 0.5$ bzw. daneben, so erhalten wir für den Fall modifizierter Dünngitter die Konvergenzraten aus Tabelle 6.14.

Zeit	Ort	Rate adaptiv	Zeit	Ort	Rate adaptiv
1	2	1.0882	1	2	2.3526
1	4	1.1305	1	4	1.1743
2	2	2.0657	2	2	2.0705
2	4	1.9911	2	4	2.0679
4	2	2.2447	4	2	2.2890
4	4	3.1623	4	4	3.1857

Tabelle 6.14: Konvergenzraten für unglatte Startdaten unter Verwendung von modifizierten Dünngittern. Die unglatte Stelle liegt auf einem Gitterpunkt (links) und neben einem Gitterpunkt (rechts).

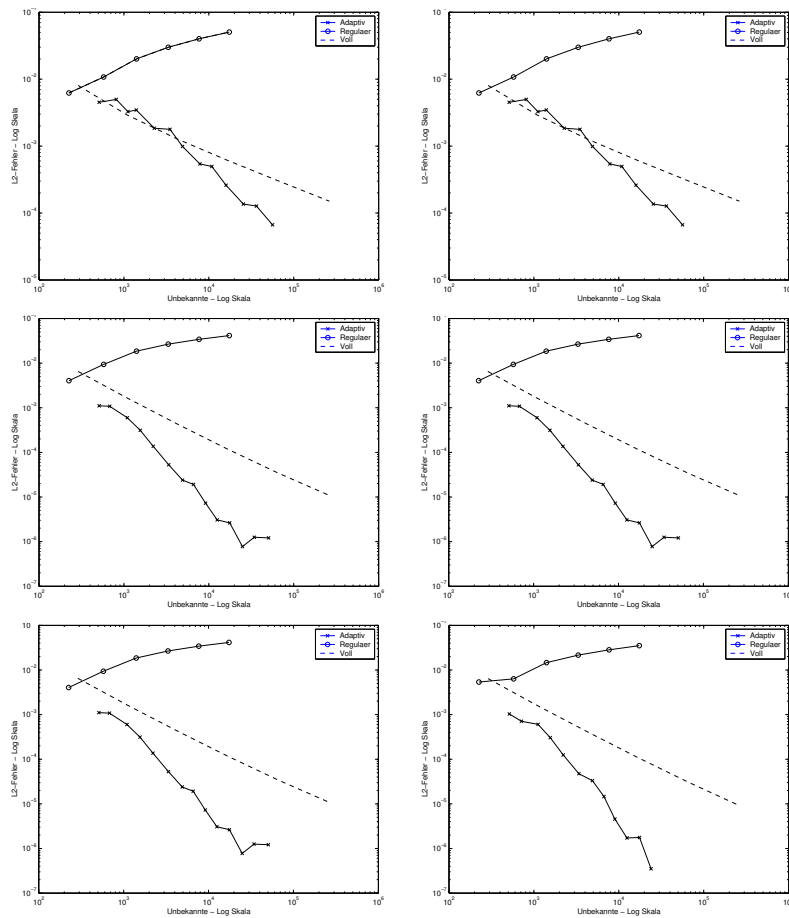


Abbildung 6.10: Konvergenzgraphen für modifizierte Dünngitter bei verschiedenen Diskretisierungen. Obere Reihe: Konsistenzordnung 1 in der Zeit, Mitte: Ordnung 2 in der Zeit, unten Ordnung vier. Die linke Spalte zeigt Ordnung 2 im Ort, rechts Ordnung 4.

6.6 Konvergenz des BiCGStab2

6.6.1 Verhalten des BiCGStab2 für verschiedene Levels

Gerechnet wurde hier das folgende Modellproblem mit einer Diskretisierung der Ordnung vier in der Zeit und vier im Ort:

$$u(t, x) = \sinh(\pi t) \cdot \sin(\pi t) \cdot \sinh(\pi x) \cdot \sin(\pi x). \quad (6.11)$$

Der Operator und die Problemstellungen sind analog zu (6.1).

Aus Abbildung 6.11 können wir nicht wirklich erkennen, ob wir einen echten Mehrgitter-Löser gewonnen haben. Das bedeutet, wir machen weitere

Tests, um zu sehen, ob die Konvergenz des Lösers unabhängig von der Maschenweite, also dem Level, ist. In der Abbildung 6.12 sehen wir für zwei verschiedene Diskretisierungen beim Rechnen des transformierten Black-Scholes Problems, dass die Konvergenzrate mit zunehmendem Level abnimmt, und wir demnach keinen maschenweitenunabhängigen Löser haben.

6.6.2 Raten für verschiedene Differentialoperatoren

Im folgenden schauen wir auf die Konvergenzraten des Lösers für das Gleichungssystem. Wie wir sehen werden verschlechtert sich die Rate erheblich, sobald wir nicht konstante Koeffizienten einführen, die die Differentialgleichung an einer Kante degenieren lassen, das heißt, dass die parabolische Struktur nicht mehr erhalten bleibt.

Es sollen ausschließlich parabolische partielle Differentialgleichungen mit verschiedenen Operatoren gelöst werden. Die Randwerte werden aus der Testfunktion als Dirichlet-Randdaten übernommen. Als rechte Seite nehmen wir die Funktion, die wir analytisch aus der Anwendung des Operators auf die Referenzlösung erhalten.

Die partielle Differentialgleichung soll dann auf verschiedenen Leveln regulär gelöst werden. Gemessen wird dann die Zahl der BiCGStab2-Schritte, die notwendig ist, um das Gleichungssystem bis auf eine Genauigkeit von 10^{-14} zu lösen. In den Graphen werden wir dann die Zahl der Schritte gegen die Anzahl der Unbekannten auftragen. Weiterhin werden wir den Fehler der Lösung (gegenüber der Referenzlösung) gegen die Anzahl der Unbekannten auftragen.

Als Standard-Referenzlösung dient

$$u(x, t) = (t + 1)(t - 1)e^{ax} \quad (6.12)$$

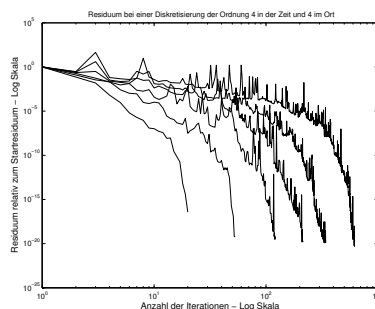


Abbildung 6.11: Relativer Fehler des BiCG im glatten Fall mit einer Diskretisierung von Konsistenz 4 in Ort und Zeit unter Verwendung von modifizierten Dünngittern.

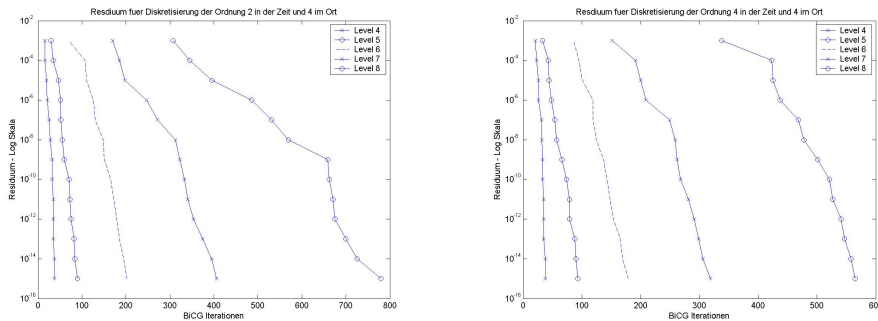


Abbildung 6.12: Links: Konvergenz des BiCGStab2 für eine Diskretisierung von zweiter Ordnung in der Zeit und vierter Ordnung im Ort. Rechts: Konsistenzordnung jeweils vier in Ort und Zeit.

Wir betrachten nun den allgemeinen Differentialoperator, den wir im folgenden variieren werden

$$L = -\frac{\partial}{\partial t} + a \cdot \frac{\partial^2}{\partial x^2} + b \cdot \frac{\partial}{\partial x} + c \quad (6.13)$$

Erster Versuch Reine Diffusion mit konstanten Koeffizienten erzeugt durch die Parameter $a = 1/2 \cdot \sigma^2$, $b = 0$, $c = 0$

Zweiter Versuch Diffusion mit konstanten Koeffizienten und Reaktionsterm. Parameter: $a = 1/2 \cdot \sigma^2$, $b = 0$, $c = -r$

Dritter Versuch Diffusion und Konvektion, jeweils mit konstanten Koeffizienten mit den Parametern: $a = 1/2 \cdot \sigma^2$, $b = r$, $c = 0$

Vierter Versuch Diffusion, Konvektion und Reaktion, jeweils mit konstanten Koeffizienten erzeugt durch die Parameter: $a = 1/2 \cdot \sigma^2$, $b = r$, $c = -r$

Die Konvergenz für die bisherigen Problemstellungen ist schnell. Im folgenden wollen wir austesten, was mit der Konvergenzrate geschieht, wenn wir die Koeffizienten wie in der Black-Scholes Gleichung zulassen, so dass das Problem degeneriert.

Black-Scholes Operator Parameter: $a = 1/2 \cdot \sigma^2 \cdot x^2$, $b = r \cdot x$, $c = -r$

Ortsabhängige Diffusion Parameter: $a = 1/2 \cdot \sigma^2 \cdot x^2$, $b = r$, $c = -r$

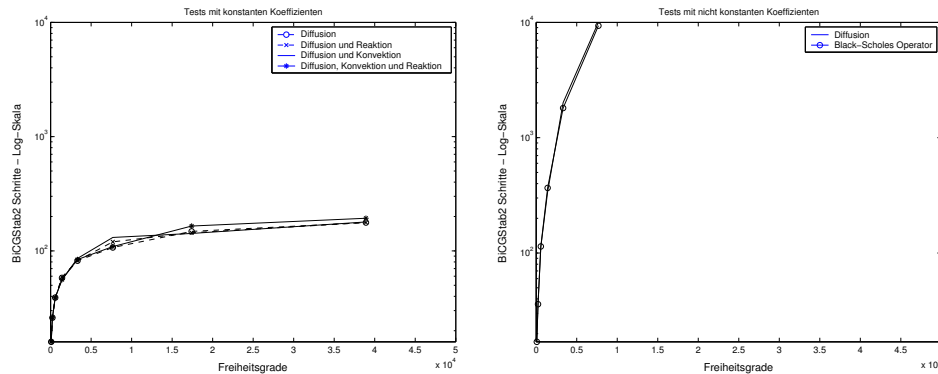


Abbildung 6.13: Konvergenz des BiCGStab2 mit konstanten (links) und nicht konstanten (rechts) Koeffizienten.

6.6.3 Veränderung der Diskretisierung

Da die Konvergenzrate des Lösers bei nicht-konstanten Koeffizienten vor der Diffusion nicht befriedigend ist, $\nabla A(\mathbf{S}, t) \nabla V(\mathbf{S}, t)$ (div-grad)

Allerdings stellt sich dabei keine Verbesserung ein, wie die nachfolgenden Tabellen zeigen.

Iterations	G	L_2	Iterations	G	L_2
19	81	5.873989e-07	14	81	1.372803e-06
45	225	3.352810e-08	34	225	6.570498e-08
77	577	2.204596e-09	101	577	4.275262e-09
110	1409	2.308447e-10	220	1409	2.229293e-10
159	3329	1.149113e-10	356	3329	6.116237e-11
182	7681	5.936306e-11	420	7681	3.345526e-11

Tabelle 6.15: Konstante Koeffizienten, Konvektion, Diffusion und Reaktion. Links: bisherige Diskretisierung, rechts: div-grad

Hier sehen wir, dass die ursprünglich gewählte Diskretisierung besser funktioniert. Der Unterschied wird größer, wenn wir zeitabhängige Koeffizienten wählen.

6.7 Speicherverbrauch

Wir messen hier den Speicherverbrauch für verschiedene Level und Dimensionen. Dabei wird jeweils ein reguläres dünnes Gitter verwendet. Die Ergebnisse wurden mit dem Programm "top" gemessen und sind in Tabelle 6.16 aufgeführt.

Dimension	Level	$ G $	Speicherverbrauch
2	3	81	40 MB
	4	225	40 MB
	5	577	41 MB
	6	1409	41 MB
3	3	729	41 MB
	4	2673	42 MB
	5	8289	46 MB
	6	23489	53 MB
4	3	6561	53 MB
	4	29889	75 MB
	5	107649	135 MB
	6	343809	312MB
5	3	59049	174 MB
	4	321489	767 MB
	5	1312929	1.620 MB
	6	4643649	5.108 MB
6	3	531441	1.5 GB
	4	3365793	6.2 GB
	5	15333057	20.7 GB

Tabelle 6.16: Speicherverbrauch für Rechnungen auf regulären dünnen Gittern.

6.8 Vergleich der verschiedenen Verfahren

Bevor wir zu den Anwendungen im nächsten Kapitel kommen, fassen wir noch einmal dieses Kapitel zusammen. Wir haben gesehen, dass die Konvergenzraten für die glatten Probleme im regulären wie im Fall der adaptiven dünnen Gitter und des modifizierten Dünngitters die gewünschte Konvergenzrate (abhängig von der Konsistenzordnung der Diskretisierung) erzielen. Die schnellste Konvergenz wurde erreicht, indem wir mit Konsistenzen der Ordnung vier in Ort und Zeit arbeiteten. Allerdings haben auch festgestellt, dass das Verfahren auf regulären dünnen Gittern für nichtglatte Anfangsdaten nicht konvergiert. Die Knickstelle kann zwar mit verschiedenen Verfahren repariert werden, aber die Konvergenzrate aus dem glatten Fall wird dabei nicht reproduziert. Das modifizierte Dünngitter-Verfahren hingegen löst die Unglattheit schnell genug auf, um die Konvergenzrate zu bewahren. Dabei ist der Fehlerschätzer von Bedeutung. Wir haben verschiedene Arten von Fehlerschätzern mit verschiedenen Gewichten getestet und festgestellt, dass wir mit einem stark greifenden L_∞ -Fehlerschätzer den besten Erfolg haben. Dieser funktioniert dann besonders gut, wenn er auf die Koeffizienten in Lifting-Basis angewendet wird. Dies steht im Einklang mit der Theorie, wie wir im letzten Kapitel bereits ausgeführt haben.

Besonders gut funktioniert unser Verfahren für konstante Koeffizienten. Werden die Faktoren vor der Diffusion ortsabhängig, so kann der Löser trotz Vorkonditionierer in Lifting-Wavelets kaum schnelle Konvergenz erzielen, im Gegensatz zum Fall konstanter Koeffizienten.

Kapitel 7

Anwendungen

Nachdem wir gesehen haben, welche Arten von Gleichungen wir berechnen können und welche Konvergenzraten wir mit welchen Einstellungen erzielen, wenden wir uns den Anwendungen zu. Zunächst betrachten wir Basket-Optionen, deren Herleitung wir im Abschnitt 3.4 beschrieben haben. Wir beschränken uns auf die Berechnung der transformierten Gleichung, da wir hier schnellere Konvergenzen für den Löser des Gleichungssystems erhalten. Interessant wird sein, wie hoch wir die Dimension steigern können, bis wir an Speichergrenzen stoßen. Anschließend simulieren wir den fairen Wert von Zinsderivaten, in unserem Falle "Floors". Wie wir bereits im vorigen Abschnitt gesehen haben, verlieren wir keine Geschwindigkeit in der Konvergenz, wenn wir eine unglatte rechte Seite einfügen. Dies werden wir uns hier zunutze machen. Zuletzt werden wir ein Modell für "Cross-Currency Swaps" numerisch lösen, das drei Zustandsvariablen beinhaltet.

Wir verwenden im Falle der Optionen und Bonds modifizierte Dünngitter. Lediglich für Cross-Currency Swaps können wir adaptive dünne Gitter verwenden, da dort die Daten glatt sind. Wir wenden in diesem Kapitel das Verfahren an, das sich im letzten Kapitel als geeignet herausgestellt hat (siehe Abschnitt 6.8).

7.1 Basket-Optionen

Die partielle Differentialgleichung ist in diesem Modell durch die mehrdimensionale Black-Scholes Gleichung (3.32) gegeben. Wir haben die Transformation auf die Wärmeleitungsgleichung in Abschnitt 3.4.1 beschrieben. Nun wollen wir zunächst das Gebiet, auf dem wir rechnen werden, beschreiben, Anfangs- und Randdaten diskutieren und uns dann der Berechnung der Basket-Optionen in verschiedenen Dimensionen zuwenden. Haben wir dann

eine Lösung gefunden, zeigen wir die Konvergenz für die Greeks auf.

7.1.1 Gebietsgröße

Wir haben im Kapitel über Optionen gesehen, dass der Fehler, den wir auf dem Rand machen, im Inneren des Gebietes abfällt. Kangro und Nicolaidis ([21]) haben dazu die Abschätzung aus Abschnitt 3.5.1 angegeben. Da wir insbesondere am Optionspreis in der Nähe des Ausübungspreises interessiert sind, spielt die Gebietsgröße eine wichtige Rolle. Der Ausübungspreis muss hinreichend weit vom Rand entfernt sein, damit die Lösung in der interessanten Region nicht vom Randfehler verfälscht wird. Als Faustregel geben Kangro und Nicolaidis an, dass die Ausbreitung des Gebietes dreimal dem Ausübungspreis entsprechen soll.

Allerdings brauchen wir in diesem Fall nicht unbedingt flexible Gebietsgrößen, obwohl unser Programm diese Möglichkeit bietet. Dass wir die transformierte Gleichung auf dem Einheitsintervall rechnen können, sehen wir am folgenden Lemma.

Lemma 7.1.1 *Der Wärmeleitungsoperator ist invariant unter linearen Gebietstransformationen. Der einzige Bezugspunkt ist somit der Ausübungspreis. Der Operator wird durch die Form*

$$\mathcal{L} = -\frac{\partial}{\partial \hat{t}} + a^2 \cdot \frac{\partial^2}{\partial \hat{S}^2} \quad a \in \mathbb{R} \quad (7.1)$$

ersetzt.

Beweis: Einsetzen der Transformation $\hat{S} = a\tilde{S} + b$ liefert direkt das gesuchte Resultat. \square

Es ist allerdings zu beachten, dass wir bei der transformierten Black-Scholes Gleichung den Ausübungspreis auf den Ursprung verschoben haben und wir das Gebiet links und rechts abschneiden müssen. Wir haben hier jedoch den Vorteil, dass wir die Gebietsgröße relativ einfach erreichen können, weil wir durch den Logarithmus ein gradiertes Gitter erhalten. Wir werden dazu im nächsten Abschnitt eine Abbildung sehen. Wollen wir die Gleichung auf dem Einheitsintervall rechnen, können wir das tun, indem wir den Ausübungspreis wieder in den positiven Abschnitt schieben und das Gebiet entsprechend Lemma 7.1.1 strecken.

7.1.2 Anfangsbedingungen

Um eine eindeutige Lösung zu erhalten, müssen sowohl Anfangs- als auch Randdaten spezifiziert werden. Die Anfangsdaten sind durch das Modell be-

stimmt und werden in diesem Abschnitt diskutiert.

Reguläre Black-Scholes Gleichung

Das Auszahlungsdiagramm für den Fall einer eindimensionalen europäischen Call-Option wurde bereits in Abbildung 3.1 dargestellt. Wollen wir Basket-Optionen betrachten, ändert sich die grundlegende Idee nicht, außer, dass der Optionspreis durch die Summe der Einzelpreise ersetzt wird. Die Anfangsdaten sind damit bestimmt durch

$$V(\mathbf{S}, 0) = \max \left[\sum_{i=1}^d \omega_i S_i - K, 0 \right]. \quad (7.2)$$

Ohne Beschränkung der Allgemeinheit können wir die Gewichte auf eins normieren. Dass wir dabei keine Einschränkungen haben, sehen wir daran, dass wir den Raum auch nach der Berechnung drehen und in die verschiedenen Richtungen strecken können. Die Berechnung auf dem Einheitsintervall bleibt also auch mit diesen Anfangsdaten möglich.

Transformierte Black-Scholes Gleichung

Da wir die transformierte Gleichung berechnen wollen, müssen wir auch die Anfangsdaten transformieren. Für den eindimensionalen Fall sehen wir die transformierten Anfangsdaten in Abbildung 7.1 links. Rechts in der Abbildung sehen wir die Lösung der Black-Scholes Differentialgleichung zu verschiedenen Zeitpunkten. Je weiter wir vom Ausübungszeitpunkt entfernt sind, desto glatter (und größer) ist der Wert der Option.

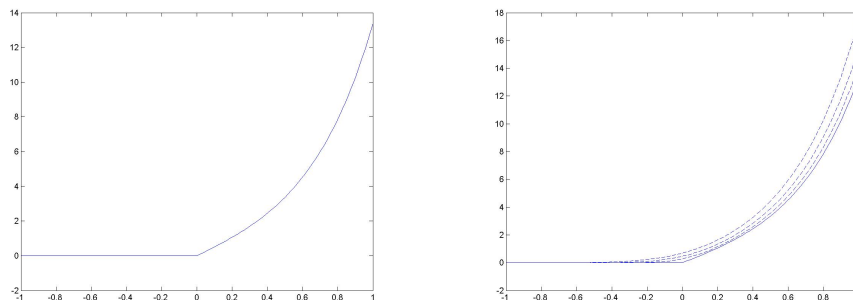


Abbildung 7.1: Links: Anfangsdaten des eindimensionalen transformierten Black-Scholes Problems. Rechts: die Lösung zu verschiedenen Zeitpunkten.

7.1.3 Randbedingungen

Wir können die Randbedingungen frei wählen, jedoch immer unter der Prämisse, die tatsächliche Lösung des Modells auf unserem Gebiet gut zu approximieren. Wie wir im Abschnitt 3.5.2 gesehen haben, erhalten wir für glatte Randdaten eine eindeutige Lösung. Deshalb ist unsere einzige Forderung, die Daten am Rand möglichst gut an die Lösung des Modells anzupassen.

Kwok ([23]) stellt eine Reihe von Eigenschaften des Optionswertes zusammen, die auch für unsere Randdaten gelten müssen. Zunächst ist der Optionspreis stets nicht-negativ. Das sehen wir schon an dem einfachen Arbitrageargument, dass eine Option nur von Nutzen, niemals schädlich sein kann, und ein negativer Preis daher eine Arbitragemöglichkeit bieten würde. Weiterhin muss jede Option ihrem intrinsischen Wert entsprechen, das heißt für Call- bzw. Put-Optionen, dass

$$C(S, t) \geq e^{-rt} \max\{S - K, 0\} \text{ bzw. } P(S, t) \geq e^{-rt} \max\{K - S, 0\} \quad (7.3)$$

gilt.

Wir wollen uns in den Berechnungen um Call-Optionen kümmern. Alle Berechnungen und Konvergenzraten sind aber direkt auf Put-Optionen übertragbar und stellen daher keine Einschränkung dar. Für Call-Optionen gilt $C(0, t) = 0$ für alle Zeitpunkte $t \leq T$. Weiterhin nutzen wir die Put-Call-Parität (3.1.2), um den Optionswert nach oben und unten abzuschätzen. Der Optionspreis ist daher beschränkt durch S (obere Schranke) und $S - Ee^{rt}$ (untere Schranke). Die Abschätzung ist scharf in dem Sinne, dass die erste Ungleichung in $S(t) = 0$ und die zweite Abschätzung asymptotisch für $S(t) \rightarrow \infty$ angenommen werden. Im ersten Falle ist dies offensichtlich, da $V(0, t) = 0$. Der zweite Fall folgt aus der Put-Call Parität. Diese Kriterien müssen von allen Randwerten erfüllt werden, die wir im folgenden vergleichen werden.

Wir werden nun verschiedene Arten von Dirichlet und Neumann-Rändern vergleichen. Dabei werden wir kurz auf den Zusammenhang mit dem Modell eingehen und anschließend Grafiken für den eindimensionalen Fall sehen, wie weit der künstliche Randwert vom tatsächlichen Rand entfernt ist. Wir können dies im eindimensionalen Fall tun, da wir über die Black-Scholes Formel eine analytische Lösung des Optionsproblems besitzen.

Dirichlet-Ränder

Exakte Daten Die Black-Scholes Formel liefert eine geschlossene Lösung des Optionswertes für Call- respektive Put-Optionen.

$$C(S, t) = SN(d_1) - Ee^{-r(T-t)}N(d_2) \quad (7.4)$$

$$P(S, t) = Ee^{-r(T-t)}N(-d_2) - SN(-d_1) \quad (7.5)$$

wobei

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-1/2 \cdot s^2} ds$$

Sie kann nicht nur als Referenzlösung herangezogen werden, sondern ihre Randwerte lassen sich auch direkt implementieren. Allerdings ist eine geschlossene Formel nur für den eindimensionalen Fall möglich. Anderenfalls wäre eine numerische Behandlung des Problems schließlich nicht notwendig.

Die Formel kann nun an beliebigen Punkten (S, t) ausgewertet werden. Dabei ist lediglich die Berechnung der kumulativen Normalverteilung $N(\cdot)$ nicht trivial. Sie kann jedoch mit Hilfe des Moro-Verfahrens schnell ausgewertet werden (siehe ([16])). Die Approximation ist dabei bis auf mindestens acht Stellen genau und damit für unsere Zwecke exakt genug (eine Approximation auf 8 Stellen entspräche einer Genauigkeit auf den Cent genau, wenn der Aktienwert bei einer Million Euro liegen würde).

Das Moro-Verfahren ist schnell zu implementieren. Wir bestimmen zunächst die Werte

$$\begin{aligned} A0 &= 0.398942270991 & A1 &= 0.020133760596 \\ A2 &= 0.002946756074 & B1 &= 0.217134277847 \\ B2 &= 0.018576112465 & B3 &= 0.000643163695 \\ C0 &= 1.398247031184 & C1 &= -0.360040248231 \\ C2 &= 0.022719786588 & D0 &= 1.460954518699 \\ D1 &= -0.305459640162 & D2 &= 0.038611796258 \\ D3 &= -0.003787400686, \end{aligned}$$

um dann die kumulative Normalverteilung über folgende Formeln zu ermitteln: Falls $x \leq 1$, 87

$$N = 0,5 + x \cdot (A0 + (A1 + A2 \cdot x^2) \cdot x^2) / (1 + (B1 + (B2 + B3 \cdot x^2) \cdot x^2) \cdot x^2)$$

wenn $1,87 < x < 6$

$$N = 1 - ((C0 + (C1 + C2 \cdot x) \cdot x) / (D0 + (D1 + (D2 + D3 \cdot x) \cdot x) \cdot x))^{16}$$

und $N = 1$ andernfalls.

Damit haben wir für den eindimensionalen Fall eine Referenzlösung gewonnen und können uns nun um die verschiedenen Methoden zum Aufsetzen von Randdaten kümmern. Die obere und untere Schranke aus der Put-Call-Parität lassen sich als Randdaten verwenden.

Konstant Im ersten Fall können also die Randdaten

$$V(0, t) = 0 \quad (7.6)$$

$$V(S^{max}, t) = S^{max} \quad (7.7)$$

implementiert werden.

Der mehrdimensionale Fall wird analog aufgebaut, wobei an den Kanten, an denen eine Variable bei 0, eine andere bei S_i^{max} ist, eine Sprungstelle entsteht.

Exponentieller Anstieg Aus der Put-Call Parität (3.1.2) erhält man die zweite Abschätzung, die sich in den Randdaten

$$V(0, t) = 0 \quad (7.8)$$

$$V(S^{max}, t) = S^{max} - Ee^{-r(T-t)} \quad (7.9)$$

ausdrückt.

Mehrdimensional sieht die Randfunktion nun anders aus. Für jeden Randpunkt (S^R, t) setzen wir

$$V(S^R, t) = V(S^R, 0) \cdot e^{-r(T-t)} \quad (7.10)$$

Neumann-Bedingungen zweiter Ordnung als Dirichlet Randdaten Tavella und Randall ([28]) beschreiben in ihrem Buch, dass Neumann-Ränder zweiter Ordnung geeignet sind, um den Optionswert am Rand zu beschreiben. Die Idee ist, dass die zweite Ableitung am Rand verschwindet, und dadurch keine Krümmung des Wertes mehr vorhanden ist. Der Preis ist also am Rand linear.

Windcliff, Forsyth und Vetzal ([31]) stellen fest, dass diese Bedingung äquivalent zu Dirichlet-Randdaten ist. Damit sind Existenz und Eindeutigkeit der Lösung garantiert. Ihr Ansatz ist einfach: der Optionswert am Rand ist linear - daher läßt sich der Wert am Rand als

$$V(0, t) = b_l(t) \quad (7.11)$$

$$V(S^{max}, t) = a(t)S^{max} + b_r(t) \quad (7.12)$$

darstellen. Die Parameter sind dabei wählbar, lassen sich aber gegenüber der Referenzlösung gut testen und sich mit diesen Werten dann auf den mehrdimensionalen Fall übertragen.

Neumann-Ränder

Man kann allerdings die Linearitätsbedingung am Rand analog zu (7.11) auch direkt als Neumann-Rand implementieren.

Neumann-Bedingungen erster Ordnung Für erste Ableitungen am Rand erhalten wir:

$$\frac{\partial V}{\partial S} V(0, t) = b_l(t) \quad (7.13)$$

$$\frac{\partial V}{\partial S} V(S^{max}, t) = b_r(t) \quad (7.14)$$

Neumann-Bedingungen zweiter Ordnung Den einfachsten Fall für den Anwender erhält man, wenn man Neumann-Randbedingungen zweiter Ordnung

$$\frac{\partial^2 V}{\partial S^2} V(0, t) = 0 \quad (7.15)$$

$$\frac{\partial^2 V}{\partial S^2} V(S^{max}, t) = 0 \quad (7.16)$$

implementiert.

Quasi-Monte Carlo Zur Berechnung des Optionspreises zum Zeitpunkt t können wir ein Quasi-Monte Carlo Verfahren verwenden (siehe ([16])). Dazu schreiben wir den Optionspreis als Integrationsproblem, das uns den Erwartungswert der Auszahlung gegeben einem Aktienwert liefert. Für den eindimensionalen Fall sieht es folgendermaßen aus:

$$V(S, t) = e^{-r(T-t)} \int_0^1 \left(S(T-t) e^{(r-\frac{1}{2}\sigma^2)\cdot(T-t)+\sigma\sqrt{T-t}N^{-1}(x)} - E \right)^+ dx, \quad (7.17)$$

wobei $N^{-1}(\cdot)$ die inverse Normalverteilung bezeichnet, die wir mit Hilfe eines weiteren Moro-Verfahrens analog zum Optionspreis bestimmen können und $(\cdot)^+$ die Funktion $\max(\cdot, 0)$ ersetzt. Die Erweiterung zum mehrdimensionalen Fall ist offensichtlich: wir ersetzen die Maximum-Funktion durch die Integration von der Nullstelle ab (also nur den positiven Teil des Integranden) und schreiben statt des eindimensionalen Gauß-Kerns den mehrdimensionalen Kern und integrieren über alle Richtungen.

Das Integral wird nun durch das arithmetische Mittel des an verschiedenen Stellen ausgewerteten Integranden ersetzt. Die Wahl der Auswertungsstellen ist beim Quasi-Monte Carlo Verfahren deterministisch vorgegeben. In

unseren Berechnungen verwenden wir die Van-der-Corput Folge mit der n -ten Primzahl als Basis in der n -ten Dimension. Das bedeutet, dass wir für den ersten Aktienwert die Basis $p_1 = 2$, für den zweiten $p_2 = 3$, für den dritten $p_3 = 5$, usw. verwenden. Aus der Basis können dann die Folgenglieder mit dem folgenden Algorithmus berechnet werden:

$$\begin{aligned}
 x &= 0 \\
 \text{für } i &= 1 \dots d \\
 z &= 1 - x_i \\
 v &= 1/p_i \\
 &\text{während } z < v + 10^{-11} \\
 v &= v/p_i \\
 x_i &= x_i + (p_i + 1) \cdot v - 1
 \end{aligned}$$

Das Moro-Verfahren kann implementiert werden mit dem folgenden Algorithmus. Mit den Werten

$$\begin{array}{ll}
 E0 = 2.50662823884 & E1 = -18.61500062529 \\
 E2 = 41.39119773534 & E3 = -25.44106049637 \\
 F0 = -8.47351093090 & F1 = 23.08336743743 \\
 F2 = -21.06224101826 & F3 = 3.13082909833 \\
 G0 = 0.3374754822726147 & G1 = 0.9761690190917186 \\
 G2 = 0.1607979714918209 & G3 = 0.0276438810333863 \\
 G4 = 0.0038405729373609 & G5 = 0.0003951896511919 \\
 G6 = 0.0000321767881768 & G7 = 0.0000002888167364 \\
 G8 = 0.0000003960315187 &
 \end{array}$$

können wir die inverse Normalverteilung mit folgendem Algorithmus berechnen:

$$\begin{aligned}
 p &= x - 0.5 \\
 \text{Wenn } |p| &< 0.42 \\
 N_{inv} &= \frac{p(((E3p^2 + E2)p^2 + E1)p^2 + E0)}{(((F3p^2 + F2)p^2 + F1)p^2 + F0)p^2 + 1}
 \end{aligned}$$


```

sonst
  falls  $p < 0$ 
     $r = x$ 
  sonst
     $r = 1 - x$ 
   $r = \log(-\log(r))$ 
   $r = r(G3 + r(G4 + r(G5 + r(G6 + r(G7 + rG8))))))$ 
   $r = G0 + r(G1 + r(G2 + r))$ 
falls  $p < 0$ 
   $N_{inv} = -r$ 
sonst
   $N_{inv} = r$ 

```

Resultate der Experimente zu den Randbedingungen

In den Tests wählen wir eine Zinsrate von $r = 0.02$ und eine Volatilität $\sigma = 0.1$. Die Ergebnisse sind in Abbildung 7.2 und Tabelle 7.2 zu sehen.

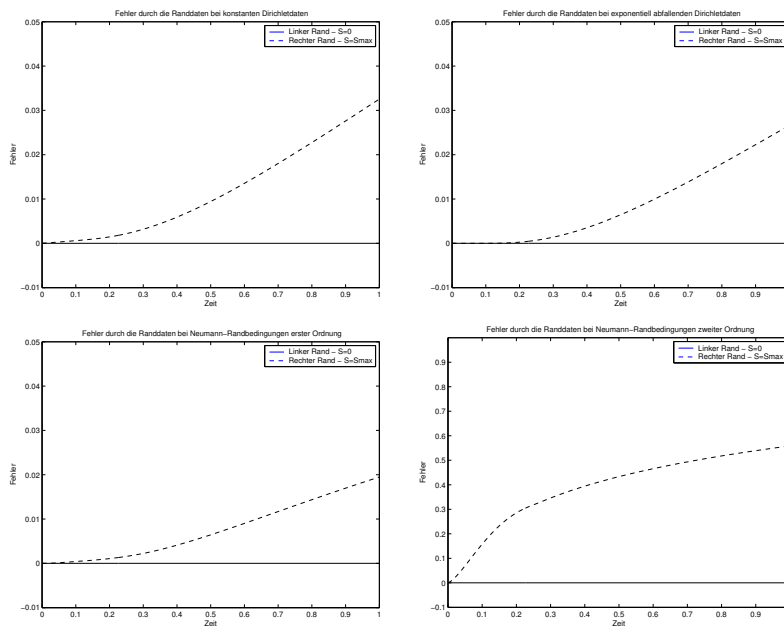


Abbildung 7.2: Oben links sehen wir den Fehler am Rand für konstante Randdaten, oben rechts für exponentielle Daten. Die Neumann-Randdaten sehen wir unten links (erste Ordnung) und rechts (zweite Ordnung).

7.1.4 Randbedingungen für die transformierte Lösung

Dirichlet-Daten

Die Aussagen aus dem vorherigen Abschnitt gelten überwiegend weiterhin. Hier werden nun die Randfunktionen aufgelistet wie sie sich nach Transformation präsentieren.

Konstant

$$\tilde{V}(\tilde{S}^{min}, \tilde{t}) = 0 \quad (7.18)$$

$$\tilde{V}(\tilde{S}^{max}, \tilde{t}) = e^{(1-\gamma)\tilde{S}^{max} - \delta\tilde{t}} \quad (7.19)$$

Exponentiell

$$\tilde{V}(\tilde{S}^{min}, \tilde{t}) = 0 \quad (7.20)$$

$$\tilde{V}(\tilde{S}^{max}, \tilde{t}) = e^{(1-\gamma)\tilde{S}^{max} - \delta\tilde{t}} - e^{-\gamma\tilde{S}^{max} - (\delta+q)\tilde{t}} \quad (7.21)$$

Exakte Lösung Wir geben die Referenzlösung über die Black-Scholes Formel nach Transformation mit (3.4.1) an, die sich folgendermaßen darstellen läßt:

$$\tilde{V}(\tilde{S}, \tilde{t}) = (e^{\tilde{S}}N(\tilde{d}_1) - e^{-q\tilde{t}}N(\tilde{d}_2)) \cdot (e^{(1/2 \cdot (q-1)) \cdot \tilde{S} + (1/4 \cdot (q-1)^2 + q) \cdot \tilde{t}}) \quad (7.22)$$

$$\tilde{d}_1 = \frac{\tilde{S} + (q+1)\tilde{t}}{\sqrt{2\tilde{t}}} \quad (7.23)$$

$$\tilde{d}_2 = \frac{\tilde{S} + (q-1)\tilde{t}}{\sqrt{2\tilde{t}}} \quad (7.24)$$

$$q = \frac{2r}{\sigma^2} \quad (7.25)$$

Neumann Ränder Die Neumann-Ränder zweiter Ordnung sind in diesem Zusammenhang nicht sinnvoll, da ebenfalls Terme erster Ordnung durch Nachdifferenzieren übrig bleiben.

Der Neumann-Rand erster Ordnung wird durch die Transformation in eine gewöhnliche Differentialgleichung umgewandelt, die es dann zu lösen gilt. Die Lösung läßt sich dann als Dirichlet-Randdatum implementieren. Für den Fall des europäischen Vanilla-Calls erhalten wir folgende Differentialgleichung

$$\frac{\partial \tilde{V}(\tilde{S}, \tilde{t})}{\partial \tilde{S}} = -e^{\gamma(\tilde{S}-1) + \delta\tilde{t}} \gamma \tilde{V}(\tilde{S}, \tilde{t}) \quad (7.26)$$

E	Konstant	Exponentiell	Neumann 1	Neumann 2	QMC
0,6	3.5159e-002	3.2370e-002	1.6757e-001	7.2179e-002	2.0122e-004
0,3	6.1016e-003	4.7464e-003	1.8540e-001	4.3868e-002	1.6113e-004
0,15	1.2340e-003	5.9548e-004	1.8766e-001	2.1025e-002	1.6378e-004
0,075	3.9258e-004	6.6695e-005	1.8772e-001	1.3084e-002	1.7011e-004
0,0375	1.8369e-004	3.8785e-005	1.8764e-001	5.9903e-003	1.8436e-004

Tabelle 7.2: Originale Black-Scholes Gleichung: Die Tabelle stellt den L_2 -Fehler auf dem ganzen Gebiet für verschiedene Randvorgaben und Ausübungspreise (und damit Gebietsgrößen) dar.

Resultate zu den Tests

Das Quasi-Monte Carlo Verfahren verwendete 100000 Durchläufe pro Gitterpunkt und war damit sehr teuer. Wie wir in der Tabelle sehen, sind alle Verfahren bis auf Quasi-Monte Carlo stark von der Gebietsgröße abhängig.

Für die transformierte Black-Scholes Gleichung ändern sich die Randwerte nicht, da der Ausübungspreis auf die Nullstelle verschoben ist und die Gebietsgröße einem enorm großen Ursprungsgebiet entspricht. Der Fehler, der auf der linken Seite gemacht wird, indem man das Gebiet abschneidet, ist geringer als die Rechengenauigkeit der exakten Lösung und damit unerheblich.

7.1.5 Konvergenz für Basket-Optionen

Den Konvergenzgraphen für den eindimensionalen Fall haben wir bereits im letzten Kapitel gesehen. Nun wollen wir die Ergebnisse für die höherdimensionalen Baskets vorstellen. Während wir im eindimensionalen Fall gegen eine analytische Referenzlösung gemessen haben, werden wir hier auf zuvor (auf feinerem Gitter) berechnete Lösungen zurückgreifen müssen. Wie zuvor bereits diskutiert, schalten wir damit auch die Randeffekte aus. Dies sehen wir beispielsweise am Fehlerplot (Abbildung 7.3) für den zweidimensionalen Fall. Wir messen hier den Fehler auf der Zeitscheibe $\tilde{t} = 0.5$.

Im fünfdimensionalen Fall (gemeint sind fünf Dimensionen plus eine Zeitdimension) können wir aufgrund von Speicherbeschränkungen nur auf einer Zeitscheibe messen.

Wir rechnen nun also die transformierte Black-Scholes Differentialgleichung mit Finiten Differenzen. Wir verwenden eine Konsistenzordnung von vier in Ort und Zeit. Die Volatilitätsmatrix enthält 0.4 auf der Diagonalen und 0.1 auf der Nebendiagonalen und ist somit positiv definit, symmetrisch und voll besetzt. Die Randbedingungen sind exponentielle Dirichlet-Daten, was der unteren Schranke der Put-Call-Parität entspricht.

Betrachten wir zunächst einmal die Konvergenzgraphen für das zweidi-

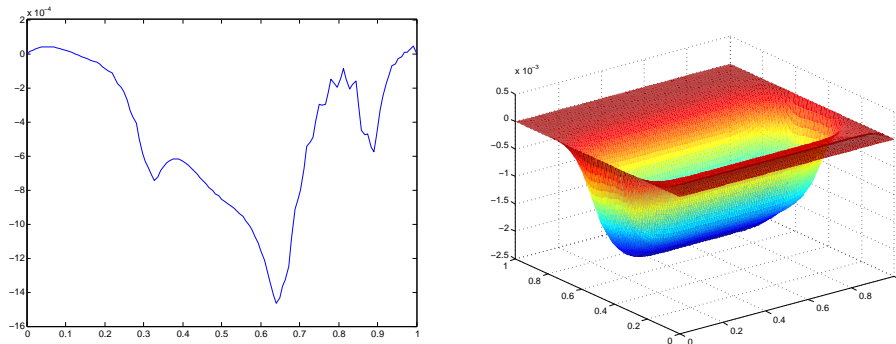


Abbildung 7.3: Fehlermessung auf einer Zeitscheibe für ein modifiziertes Dünngitter im eindimensionalen (links) und zweidimensionalen (rechts) transformierten Black-Scholes-Problem.

mensionale und das dreidimensionale Problem in Abbildung 7.4. Dort sehen wir, dass wir bereits beim dreidimensionalen Problem Schwierigkeiten bekommen, in den asymptotischen Bereich zu gelangen, wo der logarithmische Term keine Rolle mehr spielt.

Die weiteren Ergebnisse sind in Tabelle 7.3 dargestellt. Wir erhalten hier annäherungsweise die gewünschte Approximationsrate von vier. Der logarithmische Faktor der dünnen Gitter schlägt in der Präasymptotik noch zu Buche. Wir können genau gleich viele Dimensionen berechnen wie Reisinger ([25]). Entgegen der Arbeit von Reisinger wollen wir uns mit den Greeks befassen. Es bleibt im nächsten Abschnitt noch zu zeigen, dass wir auch hier die gewünschten Konvergenzraten erzielen können.

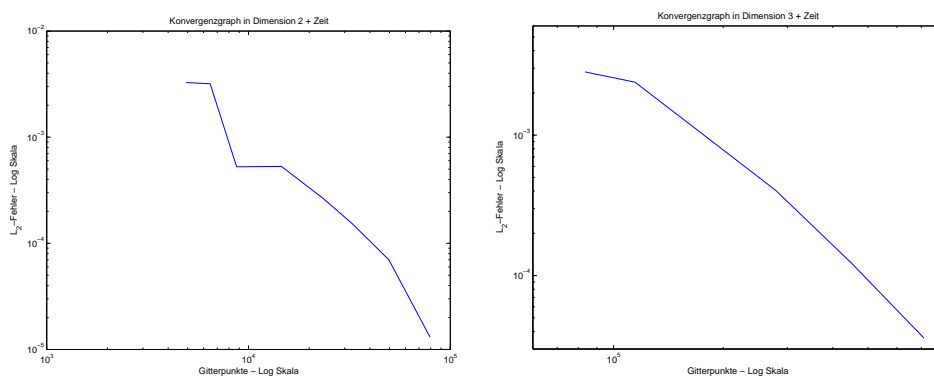


Abbildung 7.4: Die Konvergenzgraphen neigen sich beide zu einer Rate vierter Ordnung, die aber aufgrund der begrenzten Rechenkapazität nicht ganz erreicht werden kann.

Dimension 2 plus Zeit:				
G	dof	l1	l2	lmax=
4913	3600	3.270e-03	5.810e-03	4.811e-02
6460	4845	3.184e-03	5.192e-03	3.121e-02
8738	6705	5.262e-04	9.475e-04	2.756e-02
14563	11418	5.301e-04	7.639e-04	1.568e-02
23257	18681	2.680e-04	3.891e-04	1.007e-02
32507	26490	1.544e-04	2.581e-04	7.693e-03
49607	40890	7.009e-05	1.005e-04	2.944e-03
79418	66576	1.307e-05	2.396e-05	1.221e-03
Dimension 3 plus Zeit:				
G	dof	l1	l2	lmax
83521	54000	2.822e-03	5.255e-03	4.801e-02
114444	76275	2.385e-03	3.871e-03	3.087e-02
170221	116100	1.082e-03	2.593e-03	3.303e-02
278375	193950	4.055e-04	6.116e-04	1.499e-02
450523	320132	1.224e-04	1.961e-04	9.092e-03
711926	517756	3.592e-05	9.045e-05	5.934e-03
Dimension 4 plus Zeit:				
G	dof	l1	l2	lmax
1419857	810000	2.417e-03	4.710e-03	4.725e-02
2004504	1177875	1.119e-03	2.013e-03	2.998e-02
2161720	1282500	5.138e-04	1.217e-03	2.997e-02
2942887	1775250	9.775e-04	2.491e-03	3.237e-02
5161161	3175970	3.022e-04	5.314e-04	1.024e-02
5966405	3665250	2.995e-04	5.305e-04	1.030e-02
6094143	3746250	2.671e-04	4.727e-04	1.030e-02
Dimension 5 plus Zeit:				
G	dof	l1	l2	lmax
3365793	924385	1.101e-04	4.590e-04	3.714e-03
4521352	1312034	4.136e-05	2.070e-04	1.838e-03
10248325	4142817	8.409e-06	4.456e-05	7.269e-04

Tabelle 7.3: Konvergenztabelle für das transformierte Black-Scholes Problem in verschiedenen Dimensionen.

7.1.6 Greeks für Basket-Optionen

Als Ziel hatten wir uns nicht nur gesetzt, Basket-Optionen und andere Modelle, die auf eine lineare parabolische Differentialgleichung führen, zu lösen, sondern auch die Sensitivitäten bezüglich einzelner Variablen approximieren zu können. Bereits in der Einleitung hatten wir erwähnt, dass, wenn wir ein Verfahren haben, das vierte Ordnung konvergent ist, wir selbst die Gamma-Sensitivität als zweite Ableitung mit einer Konvergenzordnung von zwei approximieren können. Die ersten Ableitungen in Zeit- und Ortsrichtung, das sogenannte Theta Θ und das Delta Δ , sollten sogar mit einer Konvergenzordnung von drei berechnet werden können. Wir werden uns die entsprechenden Graphen und Approximationstabellen dazu ansehen.

Wir haben dies hier am Beispiel der eindimensionalen Funktion berechnet. Die Konvergenzordnung ist in allen Dimensionen, wie auch im vorigen Abschnitt deutlich wurde, jedoch die gleiche. Damit gelten folgende Ordnungen auch im allgemeinen Fall.

Abbildung 7.5 zeigt die Konvergenz von Delta und Theta, sowie von Gamma. Die zugehörigen Tabellen sehen wir in Tabelle 7.4. Für das Hedgen ist das Berechnen der Sensitivitäten von besonderem Interesse, da wir zur Erstellung eines risikominimalen Portfolios das Delta wissen müssen. Wenn wir die Veränderung des Deltas um den gesuchten Wert betrachten wollen, schauen wir auf das Gamma. Diese Sensitivitäten können mit dem jetzigen Verfahren bestens berechnet werden.

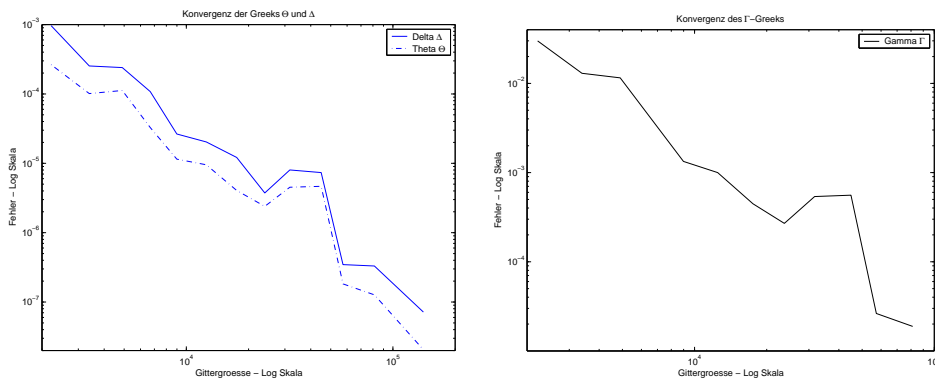


Abbildung 7.5: Auf der linken Seite sehen wir die Konvergenz der Greeks mit den ersten Ableitungen. In der anderen Grafik sehen wir die Konvergenz des Gamma-Greeks. Beide Grafiken zeigen die Konvergenzraten, die wir eingangs erwähnt hatten.

Delta Δ :		Theta Θ :		Gamma Γ :	
$ G $	L_2 -Fehler	$ G $	L_2 -Fehler	$ G $	L_2 -Fehler
2215	9.661043e-04	2215	2.641611e-04	2215	2.995402e-02
3389	2.531039e-04	3389	1.010734e-04	3389	1.297700e-02
4896	2.396567e-04	4896	1.120247e-04	4896	1.153602e-02
6686	1.081456e-04	6686	3.280030e-05	6686	3.808964e-03
8990	2.641813e-05	8990	1.147200e-05	8990	1.333385e-03
12487	2.033732e-05	12487	9.512694e-06	12487	9.996199e-04
17527	1.213446e-05	17527	4.057910e-06	17527	4.467976e-04
23962	3.736767e-06	23962	2.391429e-06	23962	2.693243e-04
31688	8.021943e-06	31688	4.520419e-06	31688	5.379120e-04
44942	7.350007e-06	44942	4.664117e-06	44942	5.577002e-04
57334	3.466158e-07	57334	1.827040e-07	57334	2.628536e-05
81315	3.308452e-07	81315	1.269269e-07	81315	1.881890e-05
140413	7.166833e-08	140413	2.053262e-08		

Tabelle 7.4: Konvergenztabelle für die eindimensionalen Greeks. Wir sehen, dass wir ein Verfahren dritter Ordnung für Θ und Δ und zweiter Ordnung für Γ erreicht haben.

7.2 Zinsderivate

In diesem Abschnitt betrachten wir Zinsderivate, im speziellen sogenannte Floors. Es wird hier eine gewisse Zinszahlung, also ein Mindestzins, auf ein Darlehen festgeschrieben. Im Modell drückt sich diese Nebenbedingung dadurch aus, dass wir eine unglatte rechte Seite erhalten. Wir haben die entsprechende Gleichung (4.14) bereits gesehen. In diesem Abschnitt wollen wir nun die numerischen Ergebnisse der Berechnung des fairen Wertes für Floors präsentieren.

Zunächst schauen wir auf die Parameter, die wir für das Modell zugrunde gelegt haben. Die Koeffizienten aus Gleichung (4.14) wurden wie folgt gewählt:

$$u_i = 0.2 \qquad \lambda_i = 1.0 \qquad (7.27)$$

$$w_i = 0.1 \qquad \bar{r} = 0.02 \qquad (7.28)$$

$$Z = 1.0 \qquad (7.29)$$

Als Gebiet wurde wieder das Einheitsintervall verwendet, wobei die Gleichung um den Faktor 10 skaliert wurde, so dass das Floorlevel \bar{r} nach Skalierung auf 0.2 lag. Diese Umskalierung bedeutet allerdings auch, dass wir ein sehr feines Gitter haben, da wir den interessantesten Bereich des Derivats um den Faktor 10 vergrößert haben.

Die Anfangsbedingungen sind durch den Basiswert Z vorgegeben. Als Randbedingungen haben wir die Dirichlet-Werte gewählt, die exponentiell an den Rändern abklingen. Der Grund ist, dass die Nebenbedingung, die für Floors für geringe Zinswerte bindet (der Marktzins liegt unter \bar{r}), sich zu früheren Zeitpunkten weniger auswirkt, da die Wahrscheinlichkeit, bis zum

Endzeitpunkt größere Zinsraten zu erreichen, wächst. Für größere Zinswerte spielt die Bedingung kaum eine Rolle und die Diskontierung greift. Wir diskontieren hier mit der jeweils realisierten Zinsrate am Rand.

Die Einstellungen für den Code waren die üblichen: wir haben Finite Differenzen Sterne der Konsistenzordnung 4 für alle Ableitungen verwendet. Weiterhin wurden Lifting-Interpolets als Adaptivitätskriterium und zum Vorkonditionieren verwendet.

Dass wir mit Knicken in der rechten Seite umgehen können, haben wir für den eindimensionalen Fall bereits im vorigen Kapitel gesehen. Für die verschiedenen Dimensionen erreichen wir die in Tabelle 7.5 und Abbildung 7.6 zusammengefassten Ergebnisse.

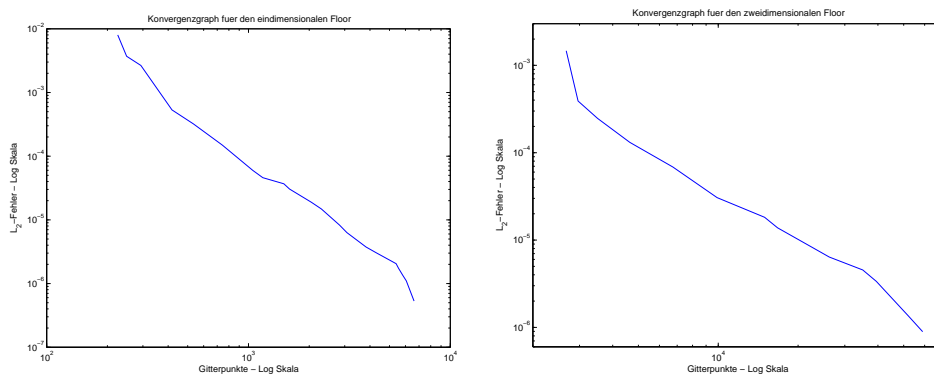


Abbildung 7.6: Links ist der eindimensionale Floor geplottet, rechts der zweidimensionale.

Wir sehen hier für die Floors wieder die gleiche Konvergenzrate von vier wie im vorigen Kapitel. Wir können somit eine ganze Klasse von Zinsderivaten berechnen, da sich die Modelle meist nur in der Art der Randbedingungen oder der rechten Seite unterscheiden.

7.3 Cross-Currency Swaps

Als letzte Anwendung werden wir einen Cross-Currency Swap berechnen. Wir wollen das Modell lösen, das im Abschnitt 4.2 eingeführt wurde.

Wir nehmen hier einen Auszahlungswert von $Z = 1$ an, der uns die Endbedingung bestimmt. Als Randbedingungen verwenden wir Neumann-Ränder, die für diese Art von Problemen geeignet sind (siehe Ekvall ([13])). Bestimmen wir also nun die Koeffizienten der Differentialgleichung, um das Modell vollständig zu spezifizieren.

Dimension 1 plus Zeit:				
G	dof	l1	l2	lmax
225	176	1.627e-03	8.015e-03	1.209e-01
249	197	6.146e-04	3.704e-03	7.980e-02
293	233	1.125e-03	2.653e-03	5.614e-02
418	348	1.194e-04	5.327e-04	2.470e-02
534	447	1.270e-04	3.208e-04	1.587e-02
742	631	5.115e-05	1.494e-04	6.132e-03
1052	907	2.742e-05	5.959e-05	3.881e-03
1178	1022	2.642e-05	4.593e-05	2.643e-03
1497	1309	2.167e-05	3.685e-05	1.536e-03
1598	1396	1.726e-05	3.042e-05	1.560e-03
2049	1809	5.777e-06	1.887e-05	9.728e-04
2298	2039	4.598e-06	1.484e-05	6.557e-04
2827	2527	3.980e-06	8.276e-06	6.558e-04
3074	2747	3.295e-06	6.303e-06	3.820e-04
3826	3429	2.304e-06	3.744e-06	2.266e-04
4300	3867	1.036e-06	3.034e-06	2.260e-04
5392	4880	6.532e-07	2.052e-06	1.255e-04
5542	5022	5.436e-07	1.736e-06	1.252e-04
5817	5277	4.122e-07	1.349e-06	3.938e-05
6056	5494	3.693e-07	1.102e-06	4.272e-05
6617	6018	2.532e-07	5.309e-07	1.297e-05
Dimension 2 plus Zeit:				
G	dof	l1	l2	lmax
2673	1680	1.745e-04	1.467e-03	4.801e-02
2961	1878	5.482e-05	3.908e-04	2.655e-02
3509	2244	3.970e-05	2.473e-04	2.056e-02
4639	3072	2.892e-05	1.315e-04	1.432e-02
6744	4639	1.728e-05	6.849e-05	1.046e-02
9898	7068	1.059e-05	3.062e-05	5.546e-03
14955	10995	8.308e-06	1.822e-05	3.506e-03
16735	12346	6.430e-06	1.381e-05	3.434e-03
22567	16901	4.648e-06	8.266e-06	2.543e-03
26221	19917	3.795e-06	6.395e-06	1.791e-03
35052	26922	2.374e-06	4.542e-06	1.116e-03
39425	30394	1.676e-06	3.366e-06	7.334e-04
52084	40600	7.070e-07	1.354e-06	3.702e-04
58951	46200	4.978e-07	8.931e-07	2.846e-04
Dimension 3 plus Zeit:				
G	dof	l1	l2	lmax
29889	14896	6.119e-05	5.025e-04	2.526e-02
30577	15247	3.924e-05	2.828e-04	1.687e-02
31489	15733	2.837e-05	1.698e-04	1.167e-02
33217	16678	1.969e-05	1.355e-04	1.168e-02
35193	17677	1.561e-05	8.993e-05	7.014e-03
36633	18487	1.533e-05	8.878e-05	7.014e-03
38344	19459	1.298e-05	7.820e-05	7.035e-03
41762	21421	1.168e-05	7.459e-05	7.041e-03
44098	22690	8.336e-06	6.534e-05	6.788e-03
46666	24238	6.987e-06	5.932e-05	6.296e-03
49734	26047	5.541e-06	3.359e-05	2.822e-03

Tabelle 7.5: Konvergenz der Lösung auf dem modifizierten Dünngitter in verschiedenen Dimensionen.

Aus Gleichung (4.46) sehen wir die Koeffizienten der Matrix, die hier zeitabhängig und wie folgt definiert sind:

$$\lambda_d = 1 - e^{0.2 \cdot (T-t)} \quad (7.30)$$

$$\lambda_f = 1 - e^{0.3 \cdot (T-t)} \quad (7.31)$$

$$G_d(s) = e^{-0.2 \cdot (T-t)} \quad (7.32)$$

$$G_f(s) = 2 \cdot e^{-0.3 \cdot (T-t)} \quad (7.33)$$

$$\sigma(s) = 1 \quad (7.34)$$

$$\rho_{df}(s) = 0.5 + 0.25 \sin(T - s) \quad (7.35)$$

$$\rho_{dS}(s) = 0.5 + 0.25 \sin\left(\frac{T - s}{3}\right) \quad (7.36)$$

$$\rho_{fS}(s) = 0.5 + 0.25 \sin\left(\frac{T - s}{4}\right) \quad (7.37)$$

Mit diesen Werten läßt sich die partielle Differentialgleichung aufstellen und lösen. Da die Anfangs- und Randdaten glatt sind, können wir adaptive dünne Gitter verwenden. Der Cross-Currency Swap ist auf drei Ortsdimensionen und eine Zeitdimension festgelegt.

Hier wird lediglich gezeigt, dass das Programm in der Lage ist, diese Zinsderivate zu berechnen. Die mittlere Zeitscheibe wird dann in eine Datei geschrieben und kann als Startdaten für den nächsten Zeitschritt eingelesen werden. Eine Rechnung mit mehreren Ausübungszeitpunkten steht noch aus. Wir messen hier den ersten Zeitschritt und lösen bis zum ersten Datum, an dem der Swap aufgelöst werden kann. Die Fehlermessung erfolgte gegen eine Referenzlösung auf einem feinerem Gitter.

Die Tabelle 7.6 und die entsprechende Abbildung 7.7 zeigen eine weniger gekrümmte Kurve als bei den vorigen Anwendungen. Die Konvergenzordnung von vier wird asymptotisch erreicht. Der Grund für den Unterschied in dem Graphen liegt darin, dass wir hier keine Unglattheit auflösen müssen. Während vorher die modifizierten Dünngitter die unglatte Stelle erst auflösen mussten, können die adaptiven dünnen Gitter hier direkt den Gesamtfehler reduzieren.

$ G $	dof	l1	l2	lmax
29889	13377	1.758e-04	5.966e-04	3.737e-03
29889	13377	1.758e-04	5.966e-04	3.737e-03
36666	17010	7.482e-05	1.512e-04	7.505e-04
37962	17598	2.734e-05	5.447e-05	4.359e-04
45792	21966	4.017e-05	7.025e-05	2.603e-04
48888	23520	2.078e-05	3.579e-05	1.583e-04
50670	24402	1.649e-05	2.751e-05	1.069e-04
55926	27734	1.313e-05	2.153e-05	8.029e-05
72702	36778	8.474e-06	1.302e-05	5.028e-05
75798	38430	7.212e-06	1.090e-05	3.985e-05
92709	47579	4.408e-06	6.779e-06	2.330e-05
128058	67222	1.660e-06	2.623e-06	9.422e-06
160485	85051	7.901e-07	1.205e-06	4.176e-06
173175	92415	7.098e-07	1.095e-06	3.854e-06
212892	115032	3.874e-07	6.038e-07	2.060e-06
239352	130292	2.669e-07	4.145e-07	1.407e-06
287595	157749	3.648e-07	5.482e-07	2.935e-06
313470	172302	1.148e-07	1.753e-07	6.221e-07
394546	218050	7.302e-08	1.130e-07	4.023e-07
440590	245366	5.114e-08	7.908e-08	2.692e-07
539380	301124	2.922e-08	4.550e-08	1.607e-07
590212	330678	2.104e-08	3.217e-08	1.157e-07

Tabelle 7.6: Konvergenztabelle für den ersten Zeitschritt im Cross-Currency Swap.

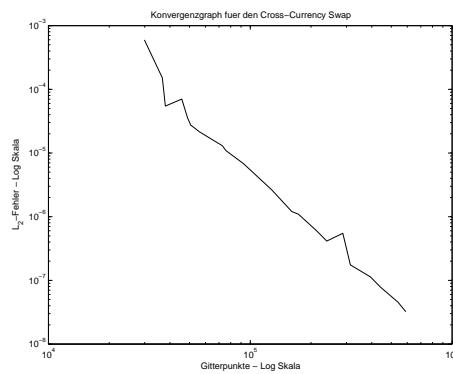


Abbildung 7.7: Konvergenzplot des Cross-Currency Swaps.

Kapitel 8

Zusammenfassung und Ausblick

Wir haben in der vorliegenden Arbeit gesehen, wie wir Basket-Optionen, Floors und Cross-Currency Swaps bewerten können. In jedem der drei Fälle galt es, die Lösung einer partiellen Differentialgleichung numerisch zu approximieren. Da es sich in jedem der drei Fälle um multivariate Probleme handelt, haben wir die Methode der dünnen Gitter gewählt, um Lösungen berechnen zu können. Als problematisch stellte sich jedoch heraus, dass das Optionspreisproblem die Glattheitsanforderungen der dünnen Gitter nicht erfüllte. Um diese Schwierigkeit zu meistern, haben wir unseren adaptiven Gittern erlaubt, sich lokal wie volle Gitter zu verhalten, und so modifizierte Dünngitter geschaffen. Als Alternative haben wir eine Randkorrektur angegeben, die ebenfalls zu Konvergenz, wenn auch zu einer Verringerung in der Rate, führte.

Die für den multivariaten Fall erstmals vorgestellte Transformation hilft uns dabei, Schwierigkeiten mit dem Löser des Gleichungssystems wegen ortsabhängiger Koeffizienten in der Diffusion zu umgehen.

Da wir im Falle der Basket-Optionen nicht nur die Lösung, sondern auch die Sensitivitäten berechnen wollen, wählten wir Interpolets als Ansatzfunktionen, die uns eine glatte Lösung und sehr gute Approximationseigenschaften lieferten. Zusammen mit Finite-Differenzen-Sternen von hoher Konsistenzordnung erhalten wir so ein Verfahren, das vierter Ordnung konvergent ist. Damit sind wir in der Lage, selbst Sensitivitäten mit zweiten Ableitungen schnell und gut approximieren zu können.

Da wir die aus den jeweiligen Modellen entstehenden partiellen Differentialgleichungen nur auf endlichen Gebieten lösen können, mussten wir künstliche Randbedingungen implementieren. Wir haben verschiedene Varianten vorgestellt und miteinander verglichen.

Insgesamt konnten wir zeigen, dass mit der vorgestellten Transformation ein Verfahren vorliegt, das die gewünschten Eigenschaften vollständig erfüllt.

Das Verfahren kann noch auf weitere Problemstellungen im Finanzbereich ausgeweitet werden. An erster Stelle sollten Cross-Currency Swaps in mehreren Zeitschritten gerechnet werden. Als nächster Schritt ist das Berechnen amerikanischer Optionen angedacht. Der Unterschied zu europäischen Optionen liegt in der Behandlung freier Ränder. Damit ist der Kreis der Anwendungen stark vergrößert. Langfristig soll noch eine Optimierungsroutine implementiert werden, um die Bellman Gleichung (in kontinuierlicher Zeit) berechnen zu können. Diese Gleichung führt auf eine parabolische Differentialgleichung (linear, zweiter Ordnung), wobei jedoch zuvor eine Parameterfunktion maximiert werden muss. Die zwei Schritte, nämlich der Optimierungsprozess und die Lösung der Differentialgleichung lassen sich getrennt behandeln. Das Anwendungsgebiet sind hier Aufgaben, die sich nicht auf gewöhnliche Differentialgleichungen zurückführen lassen, und sowohl feste als auch freie Ränder haben können.

Anhang A

Aufschlüsselung des DAX

Gewichtung Stand: 27. Januar 2005

Aktie	Gewichtung
Konsumgüter, Pharma und Chemie	
Adidas-Salomon	1.15 %
Altana	0.70 %
BASF	6.26 %
Bayer	3.75 %
Fresenius Medical Care	0.46 %
Henkel KGaA	0.89 %
Metro	1.25 %
Finanzdienstleistungen	
Allianz	6.90 %
Bayerische Hypo- und Vereinsbank	2.23 %
Commerzbank	1.73 %
Deutsche Bank	7.71 %
Deutsche Boerse	1.16 %
Muenchener Rueckvers.-Ges.	3.54 %
Schering	2.01 %
Verkehr, Logistik und Schwerindustrie	
BMW	2.37 %
Continental	1.66 %
Daimler-Chrysler	6.35 %
Deutsche Lufthansa	0.94 %
Deutsche Post	1.87 %
MAN	0.68 %
Thyssen-Krupp	1.52 %
Tui	0.62 %
Volkswagen	1.74 %
Technologie	
Deutsche Telekom	9.37 %
E.ON	10.32 %
Infineon Technologies	0.96 %
Linde	0.86 %
RWE	4.11 %
SAP	5.21 %
Siemens	11.67 %

Literaturverzeichnis

- [1] **Auer, J. (2003):** *Wachstumsmarkt Wetterderivate*. DEUTSCHE BANK RESEARCH, NR. 255, 7. FEBRUAR 2003.
- [2] **Balder, R. (1994):** *Adaptive Verfahren für elliptische und parabolische Differentialgleichungen auf dünnen Gittern*. DISSERTATION, INSTITUT FÜR INFORMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN.
- [3] **Balder, R., U. Rüde, S. Schneider und C. Zenger (1994):** *Sparse grid and extrapolation methods for parabolic problems*. IN PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTATIONAL METHODS IN WATER RESOURCES, HEIDELBERG 1994 (A. PETERS ET AL., EDS), KLUWER ACADEMIC, DORDRECHT, SEITEN 1383-1392.
- [4] **Barraquand, J. (1993):** *Numerical Valuation of High Dimensional Multivariate European Securities*. RESEARCH REPORT 26, DIGITAL EQUIPMENT CORPORATION, PARIS RESEARCH LABORATORY.
- [5] **Black, F. und M. Scholes (1973):** *The Pricing of Options and Corporate Liabilities*. THE JOURNAL OF POLITICAL ECONOMY, VOL. 81, No. 3, S. 637-654.
- [6] **Bungartz, H.J. (1992)** *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*. DISSERTATION, TECHNISCHE UNIVERSITÄT MÜNCHEN.
- [7] **Bungartz, H.-J. und M. Griebel (1997)** *A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives*. SFB-REPORT 256, NUMMER 254, UNIVERSITÄT BONN, 1997.
- [8] **Bungartz, H.-J. und M. Griebel (2004)** *Sparse Grids*. ACTA NUMERICA (2004), PP. 1-121.
- [9] **Cox, J. C., S. Ross und M. Rubinstein (1979):** *Option Pricing: A simplified approach*. JOURNAL OF FINANCIAL ECONOMICS, VOL. 7, S. 229-263.

- [10] **Dempster, M. und J.P. Hutton (1997):** *Numerical Valuation of Cross-Currency Swaps and Swaptions*. IN MATHEMATICS OF DERIVATIVE SECURITIES, M.A.H. DEMPSTER AND S.R. PLISKA (EDS.), CAMBRIDGE UNIVERSITY PRESS.
- [11] **Deslauriers, G. und S. Dubuc (1989):** *Symmetric iterative interpolation processes*. CONSTRUCTIVE APPROXIMATION (1989), 5, 49-68.
- [12] **DeVore, R. (1998):** *Nonlinear approximation*. ACTA NUMERICA, S. 51-150, CAMBRIDGE UNIVERSITY PRESS.
- [13] **Ekvall, N. (1993):** *Two Finite Difference Schemes for Evaluation of Contingent Claims with Three Underlying State Variables*. TECHNICAL REPORT 6520, EKONOMISKA FORSKNING SINSTITUTET, STOCKHOLM, NORWAY, 1993.
- [14] **Flyer, N. und P. N. Swarztrauber (2002):** *The Convergence of Spectral and Finite Difference Methods for Initial-Boundary Value Problems*. SIAM JOURNAL OF SCIENTIFIC COMPUTING, VOL. 23, NR. 5, S. 1731-1751.
- [15] **Friedman, A. (1964):** *Partial Differential Equations of parabolic type*. PRENTICE-HALL.
- [16] **Gerstner, T. (2003):** *Unterlagen zum Praktikum "Numerische Verfahren der Finanzmathematik"*. KOPIE INSTITUT FÜR ANGEWANDTE MATHEMATIK DER UNIVERSITÄT BONN.
- [17] **Gietl, H. und C. Zenger (1978):** *Improved Difference Schemes for the Dirichlet Problem of Poisson's Equation in the Neighbourhood of Corners*. NUMERISCHE MATHEMATIK (1986), VOL. 30, 315-332.
- [18] **Griebel, M. (1998):** *Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences*. COMPUTING, 61(2):151-179.
- [19] **Hull, J. (2000):** *Options, Futures and other Derivative Securities*. PRENTICE HALL, UPPER SADDLE RIVER.
- [20] **Hutton, J.P. (1995):** *Fast Pricing of Derivative Securities*. PHD THESIS, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ESSEX.
- [21] **Kangro, R. und R. Nicolaidis (2000):** *Far Field Boundary Conditions for Black-Scholes Equations*. SIAM JOURNAL ON NUMERICAL ANALYSIS, VOL. 38, NO. 4, S. 1357-1368.

- [22] **Koster, F. (2001):** *Multiskalen- basierte Finite- Differenzen Verfahren auf adaptiven dünnen Gittern.* DISSERTATION, RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN, 1998.
- [23] **Kwok, Y.-K. (1998):** *Mathematical Models of Financial Derivatives.* SPRINGER-VERLAG, 1998.
- [24] **Merton, R. (1973):** *Theory of Rational Option Pricing.* BELL JOURNALS OF ECONOMICS AND MANAGEMENT SCIENCE, 4, S. 141-183
- [25] **Reisinger, C. (2004):** *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben.* DISSERTATION, RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG.
- [26] **Schiekofer, T. (1998):** *Die Methode der Finiten Differenzen auf dünnen Gittern zur Lösung elliptischer und parabolischer partieller Differentialgleichungen.* DISSERTATION, RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN.
- [27] **Seydel, R. (2002):** *Tools for Computational Finance.* SPRINGER VERLAG.
- [28] **Tavella, D. und C. Randall (2000):** *Pricing Financial Instruments.* JOHN WILEY AND SONS.
- [29] **Wilmott, P., S. Howison und J. Dewynne (1993):** *Option Pricing.* OXFORD, FINANCIAL PRESS.
- [30] **Wilmott, P., S. Howison und J. Dewynne (1995):** *The mathematics of financial derivatives.* CAMBRIDGE UNIVERSITY PRESS.
- [31] **Windcliff, H., P.A. Forsyth und K.R. Vetzal (2003):** *Analysis of the stability of the linear boundary condition for the Black-Scholes equation.* ERSCHEINT IM JOURNAL OF COMPUTATIONAL FINANCE.
- [32] **Yserentant, H. (1986):** *On the Multi-Level Splitting of Finite Element Spaces.* NUMERISCHE MATHEMATIK (1986), 49, 379-412.
- [33] **Zenger, C. (1991):** *Sparse Grids.* IN PARALLEL ALGORITHMS FOR PARTIAL DIFFERENTIAL EQUATIONS (W. HACKBUSCH, ED.), VOL. 31 OF NOTES ON NUMERICAL FLUID MECHANICS, VIEWEG VERLAG.