

INAUGURAL-DISSERTATION

ZUR
ERLANGUNG DER DOKTORWÜRDE
DER
NATURWISSENSCHAFTLICH-MATHEMATISCHEN
GESAMTFAKULTÄT
DER
RUPRECHT-KARLS-UNIVERSITÄT
HEIDELBERG

vorgelegt von
Diplom-Ingenieur Christoph Reisinger
aus Pasching/Linz (Österreich)

Tag der mündlichen Prüfung: 8. 6. 2004

**Numerische Methoden
für
hochdimensionale parabolische Gleichungen
am Beispiel von
Optionspreisaufgaben**

Gutachter: Prof. Dr. Gabriel Wittum
Prof. Dr. Dr. h.c. mult. Willi Jäger

Zusammenfassung

Zunehmend an Bedeutung in einer Vielzahl von Anwendungsfeldern gewinnen hochdimensionale Aufgabenstellungen, oft in der Form von partiellen Differentialgleichungen. Kernpunkt dieser Arbeit sind Diskretisierungsverfahren speziell für diese Problemklasse, wobei *en passant* auch auf die untrennbar verbundenen Aspekte iterativer Gleichungslösung und speziell auch der Modellreduktion (hier in Form von asymptotischer Analysis zur Reduzierung der Dimension) eingegangen wird. Als Anwendungsbeispiel werden Optionspreisaufgaben in bis zu dreißig Dimensionen studiert, da viele Produkte am Markt durch die Anzahl der Faktoren direkt auf hochdimensionale (Un-)Gleichungen führen und außerdem die oft unabdingbare stochastische Modellierung von Marktdaten die Dimension weiter hochtreibt.

Die Basis dazu bilden dünne Gitter. Mittels der Kombinationstechnik wird die Dünngitter-Lösung aus Finite-Differenzen-Lösungen auf einer Familie anisotroper kartesischer Gitter extrapoliert, wodurch die Zahl der Freiheitsgrade entscheidend reduziert und der Algorithmus auf natürliche Weise parallelisiert werden kann. Aus einer geeigneten Fehlerdarstellung am kartesischen Gitter werden (Dünngitter-)Fehlerabschätzungen in einer geschlossenen Form für beliebige Dimensionen abgeleitet und ein zusätzlicher multivariater Extrapolationsschritt für eine höhere Konvergenzordnung motiviert. Dies erlaubt die numerische Differentiation mit hinreichender Genauigkeit. Optimale Komplexität des Gesamtalgorithmus sowie eine effiziente Lastverteilung werden durch robuste Mehrgitterverfahren mit Block-Glätten und angepassten Transferoperatoren an freien Rändern (bei Amerikanischen Optionen) erzielt.

Am Beispiel von Basket-Optionen wird schließlich demonstriert, wie Probleme, deren Dimensionalität den mit dünnen Gittern behandelbaren Rahmen (ungefähr sechs) übersteigt, mittels Hauptkomponentenanalyse und asymptotischer Analysis durch Aufgaben drastisch reduzierter Dimension mit hoher Genauigkeit approximiert werden können.

Abstract

This thesis deals with the analysis and application of discretisation schemes tailored for high-dimensional parabolic equations, which are of increasing importance in a wide range of applications, and covers *en passant* the inevitably connected issues of iterative solvers and especially model reduction (here in the form of asymptotic analysis for the reduction of the dimension). These methods are applied to equations and variational inequalities (American options) from derivative pricing with up to thirty stochastic factors.

The approach is based on sparse grids. For the combination technique closed form error estimates in arbitrary dimensions are derived from a suitable representation of the discretisation error on anisotropic Cartesian grids, which also motivates an additional multivariate extrapolation step for higher order convergence. This allows to estimate sensitivities with sufficient accuracy. Robust multigrid methods with block smoothing and adapted grid transfer at free boundaries ensure optimal complexity and efficient load balancing.

Asymptotic formulae are derived for the Black-Scholes price of large basket options, where the direct numerical treatment on sparse grids is no longer feasible. In a more general context, the price is extrapolated with astonishing accuracy from two-dimensional numerical solutions.

Vorwort

Die vorliegende Arbeit entstand im Rahmen des BMBF-Projekts „Analytische Methoden und effiziente numerische Algorithmen zur Bewertung derivativer Finanzinstrumente“ zwischen dem Interdisziplinären Zentrum für Wissenschaftliches Rechnen (IWR) der Universität Heidelberg und der Dresdner Bank AG.

Die hier vorgestellten mathematischen Konzepte wurden im dafür entwickelten Programmpaket \mathcal{SG} , das auf parallelen Modulen der von der Arbeitsgruppe Technische Simulation entwickelten \mathcal{UG} -Toolbox [BBJ⁺97] basiert und diese um hochdimensionale Datenstrukturen sowie Lösungsverfahren erweitert, praktisch umgesetzt.

Es ist eine angenehme Pflicht, den Menschen zu danken, die hinter dem Erfolg dieser Arbeit stehen:

An erster Stelle gebührt der Dank meinem Betreuer Prof. Dr. Gabriel Wittum für die überspringende wissenschaftliche Begeisterung und die freundschaftliche Führung, aber auch für unvergleichliche Rahmenbedingungen am Institut. Prof. Dr. Dr. h.c. mult. Willi Jäger danke ich sehr herzlich für richtungsweisende Diskussionen und seine uneingeschränkte Unterstützung in allen Belangen.

Tief empfundener Dank gilt weiters meinen Kollegen Dr. Stefan Lang, dessen Handschrift die parallele Programmentwicklung trägt, Dr. Nicolas Neuss für viele erleuchtende Diskussionen in mathematischen Belangen aller Art und Dr. Oliver Sterz, in dessen unerschöpflicher \LaTeX -Trickkiste ich mich mit Freuden bediente. Dr. Jörg Kampen hat zum erfolgreichen gemeinsamen BMBF-Projekt viele der „analytischen Methoden“ beigesteuert und mich in diversen Gesprächen in Aspekte der Finanzmathematik eingeführt. In diesem Zusammenhang bedanke ich mich auch herzlich bei Dr. Jürgen Linde von der Dresdner Bank AG, der mir die Anforderungen aus der Praxis nähergebracht und die in dieser Arbeit verwendeten Daten bereitgestellt hat. Ein großer Dank geht auch an Jürgen Schumacher vom Institut für Informatik der Universität Bonn für die wertvollen Vergleichsergebnisse aus Monte-Carlo-Simulationen und an Stefan Friedel für die technische Unterstützung bei den parallelen Rechnungen am IWR-Cluster HELICS.

Eine große Hilfe war die konstruktive Kritik von Dr. Jens Eberhard, Dr. Peter Frolkovič und Dr. Markus Wabro, wenngleich die von ihm überlassenen Monte-Carlo-Referenzdaten (Gesamtbevölkerung: 29 972, davon 12 047 Franzosen, 5 070 Monegassen,...) von begrenztem Nutzen waren.

Nicht zuletzt bin ich allen Freunden am IWR — namentlich meinen lieben Bürokollegen, Sport- und Bergkameraden Dirk Feuchter, Ingo Heppner und Jochen Hittler — dankbar verbunden.

Heidelberg, im Jänner 2004

Christoph Reisinger

Inhaltsverzeichnis

Einleitung	1
1 Optionspreismodelle	7
1.1 Optionen	8
1.1.1 Kontrakttypen	8
1.1.2 Wozu Optionen?	9
1.1.3 Beispiele	9
1.2 Modellierung von Basiswert und Option	10
1.2.1 Eigenschaften von Finanz-Zeitreihen	10
1.2.2 Das Black-Scholes-Modell — Aktienoptionen	11
1.2.3 Stochastische Volatilitätsmodelle	13
1.2.4 Amerikanische und Exotische Optionen	15
1.2.5 Wechselkurs-Optionen	16
1.3 Mehrdimensionale Black-Scholes-Gleichung und Transformationen	17
1.3.1 Logarithmische Koordinaten	17
1.3.2 Problemangepasste Koordinaten	18
1.3.3 Hauptachsentransformation	19
2 Diskretisierung auf gradierten kartesischen Gittern	21
2.1 Diskretisierung	22
2.1.1 Ortsdiskretisierung	22
2.1.2 Zeitdiskretisierung	23
2.1.3 Numerische Beispiele und Konvergenzdaten	24
2.2 Gradierte Gitter	27
2.3 Extrapolation und Sensitivitäten	30
2.3.1 Richardson-Extrapolation	30
2.3.2 Numerische Differentiation	31
3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme	33
3.1 Multigrid	34
3.1.1 Iterative Verfahren	34
3.1.2 Der Basis-Mehrgitter-Algorithmus	37
3.1.3 Spezielle Anforderungen	38
3.2 Multigrid für anisotrope Probleme	39
3.2.1 Ebenenglätter	40
3.2.2 Numerische Tests	41
3.3 Multigrid für Hindernisprobleme	45
3.3.1 Der projizierte Multigrid-Zyklus	45

3.3.2	Gittertransfer	46
3.3.3	Volles Mehrgitter	47
3.3.4	Robuste Glätter	48
3.3.5	Kaskadisches Mehrgitter	49
4	Konvergenzabschätzungen für die Kombinationstechnik	51
4.1	Das Prinzip dünner Gitter	52
4.2	Die Kombinationstechnik	55
4.3	Konvergenzabschätzungen	57
4.4	Verbesserte Konvergenzabschätzungen	63
5	Anwendung auf Optionspreisaufgaben	67
5.1	Fehleranalyse	68
5.2	Der Vergleich mit vollen Gittern	69
5.2.1	Ein 3-Faktor-Modell für die FX-Option	69
5.2.2	Basket-Option mit zwei Komponenten — gradierte dünne Gitter	70
5.3	Komplexitätsbetrachtungen	71
5.4	Aspekte der parallelen Lösung	72
5.5	Höhere Dimensionen	75
5.5.1	Konstruktion der Testbeispiele	75
5.5.2	Numerische Ergebnisse	77
6	Multivariate Extrapolation und eine Kombinationstechnik höherer Ordnung	81
6.1	Multivariate Extrapolation	82
6.2	Eine Kombination höherer Ordnung	87
6.2.1	Das Prinzip	87
6.2.2	Konvergenzabschätzungen	89
6.3	Numerische Ergebnisse	90
6.3.1	FX-Option	90
6.3.2	Basket mit zwei Komponenten	91
6.3.3	Basket mit fünf Komponenten	92
7	Dimensionsreduktion für hochdimensionale parabolische Gleichungen	95
7.1	Motivation	96
7.2	Hauptkomponentenanalyse	99
7.2.1	Das Basket-Modell	100
7.2.2	Eigenmoden	100
7.2.3	Kombination	103
7.3	Asymptotik erster Ordnung	106
7.3.1	Analytische Asymptotik für Black-Scholes	106
7.3.2	Verbesserte numerische Approximation	109
7.4	Vergleich	112
	Diskussion und Ausblick	113
A	Herleitung der asymptotischen Entwicklung	117

Einleitung

Börsenrauschen

„Es ist nicht so, dass an Finanzmärkten keine Intelligenz versammelt wäre — man merkt nur nichts davon. Zumindest die Kursverläufe an der Londoner Börse zwischen 1998 und 2000 unterscheiden sich in ihren statistischen Eigenschaften durch nichts von den Daten, die zwei Forscher vom Santa Fé Institute durch reine Zufallsprozesse erzeugt haben. Dieses Ergebnis versetzt dem ohnehin wankenden Glauben an die Rationalität ökonomischen Handelns einen weiteren Stoß.“

Randnotiz in der *Frankfurter Allgemeinen Sonntagszeitung* vom 28. 9. 2003,
offenbar Bezug nehmend auf den Artikel von
J. D. Farmer, P. Patelli, I. I. Zovko
“The Predictive Power of Zero Intelligence in Financial Markets”

Seit den bahnbrechenden Arbeiten von Black, Scholes [BS73] und Merton [Mer73] aus den Siebziger-Jahren sind die Methoden zur Bewertung von Finanzderivaten in bemerkenswerter Weise fortgeschritten und haben völlig neue Anforderungen an Stochastik, Analysis und Numerik gestellt. Diese Entwicklungen lassen sich grob in zwei Hauptrichtungen einteilen.

Zum einen stellte man schnell fest, dass wesentliche Eigenschaften der beobachteten Marktdaten durch ursprüngliche Modelle nicht oder nur unzureichend wiedergegeben werden, sodass eine

- realistischere Modellierung der treibenden Prozesse

unabdingbar ist. Zum anderen verlangt eine ständig wachsende Fülle an standardisierten und exotischen Produkten die

- Modellierung und (numerische) Bewertung komplexer Kontrakttypen.

Die Eigenheiten maßgeschneiderter Konstrukte erstaunen in der Tat, wenn man Produkte wie *Hawaii*-, *Himalaya*- oder *Holiday-Optionen* angeboten bekommt. Großer Beliebtheit erfreuen sich insbesondere Optionen auf Indizes, da sie eine Absicherung gegenüber Schwankungen eines repräsentativen Querschnitts erlauben. In typischen Fällen wie dem DAX 30 oder dem S&P 500 hat man es unter Umständen mit Dutzenden oder gar Hunderten von Komponenten zu tun.

Auf Modellierungsseite sind vereinfachende Annahmen wie die Konstanz von Marktparametern nicht haltbar. Die Brisanz der Problematik wird unter anderem dadurch belegt, dass erst vor wenigen Wochen der Stochastiker Robert Engle, dessen Arbeiten zu ARCH-Modellen für ökonomische Zeitreihen sich genau aus dieser Forderung entwickelt haben, mit dem Nobelpreis bedacht wurde. Eine Fülle an weiteren Prozessen wurde unterdessen vorgeschlagen, die von der traditionellen logarithmischen Normalverteilung der Kurs-Inkrementen abweichen und wesentliche Charakteristika, die unter den Schlagworten *Heavy Tails*, *Skewness* oder *Smile* prominent geworden sind, besser erfassen.

Gemein ist diesen beiden Herausforderungen — komplexe Kontrakte mit mehreren Komponenten zu preisen und die stochastische Natur von Parametern zu berücksichtigen —, dass sie oftmals als mehrdimensionale partielle Differentialgleichungen zu formulieren sind, unabhängig davon, ob die multivariaten stochastischen Faktoren direkt als Basiswerte in den Kontrakt eingehen oder als Hintergrundprozesse relevante Marktdaten beschreiben.

Die vorliegende Arbeit widmet sich der numerischen Lösung solcher parabolischer Gleichungen mit der von der Praxis vorgegebenen hohen Dimensionalität. Weitere Anwendungsbeispiele, bei denen hochdimensionale Aufgabenstellungen eine zentrale Rolle spielen, finden sich oft auch in ganz anderen Bereichen wie der Quantentheorie — etwa mit Anwendungen in der Molekulardynamik — oder der strukturierten Populationsdynamik.

Die fundamentale Schwierigkeit bei der Lösung derartiger Problemstellungen besteht darin, dass die Dimension der diskreten Räume (typischerweise nach einer Diskretisierung des kontinuierlichen Modells auf einem kartesischen Gitter) exponentiell mit der Raumdimension D ansteigt, was unter dem bezeichnenden Namen „Fluch der Dimension“ gehandelt wird: Zur Lösung mit einer Genauigkeit ϵ sind

$$N \sim \epsilon^{-D/p}$$

Freiheitsgrade notwendig, wobei p die Approximationsordnung beschreibt.

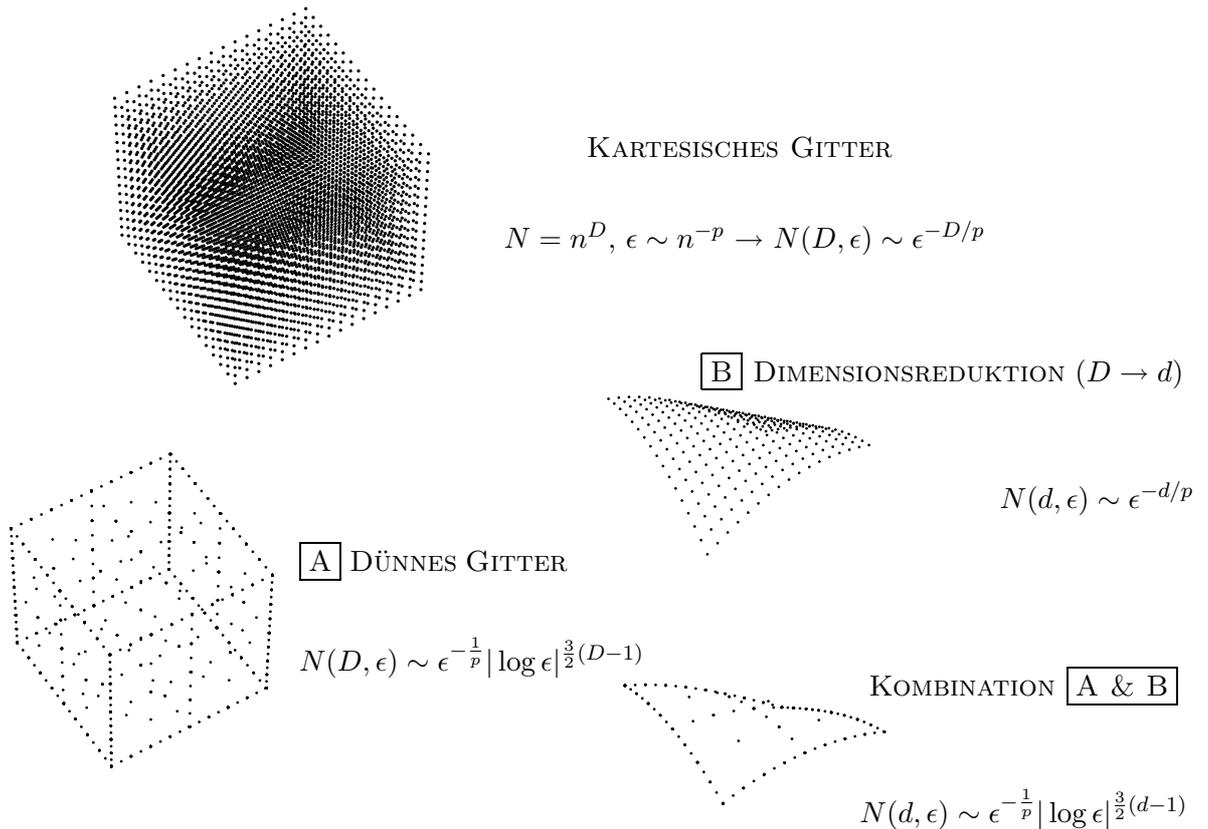


Abbildung 0.1: „Fluch der Dimension“: Die Anzahl N an Freiheitsgraden zur Darstellung der Lösung mit einer Genauigkeit ϵ steigt exponentiell in der Dimension D des Problems. Zwei Strategien: Diskretisierung auf dünnen Gittern und analytische Dimensionsreduktion.

Einleitung

Wenn jetzt eine verlangte Genauigkeit ϵ fix vorgegeben ist, beispielsweise eine Fehlerschranke für den Optionspreis, bleiben noch zwei Parameter, an denen man drehen kann — p und D — und daraus ergeben sich zwei Strategien:

- p mit der Dimension hochtreiben oder
- D reduzieren.

Um von ansteigender Ordnung p zu profitieren, muss naturgemäß ein hohes Maß an Glätte vorausgesetzt werden, wie sie jedoch durch die Glättungseigenschaft parabolischer Operatoren oft gegeben ist. Eine Variante, die auf multivariater Extrapolation aus einer Hierarchie anisotroper Gitter basiert, führt auf sog. *dünne Gitter* mit einer ϵ -Komplexität von der Art

$$N \sim \epsilon^{-1/p} |\ln \epsilon|^{\frac{3}{2}(D-1)}.$$

Es ist ein fundamentales Resultat dieser Arbeit, dass sich solche Approximationsaussagen auf Konvergenzabschätzungen für die *Kombinationstechnik* entkoppelter Gleichungen übertragen.

Die Problemdimension — auf den ersten Blick fest vorgegeben — kann mitunter durch die Wahl geeigneter Koordinaten beeinflusst werden. Dies funktioniert insbesondere dann, wenn die hochdimensionale Lösung durch wenige Hauptkomponenten, die sich aus der Struktur des Differenzialoperators ergeben, wenn nicht exakt, so zumindest mit für praktische Zwecke hinreichender Genauigkeit dargestellt werden kann. Die hier entwickelte Technik geht noch einen Schritt weiter und leitet mittels asymptotischer Analysis eine allgemeinere Darstellung durch niedrigdimensionale Funktionen ab.

Dies führt in natürlicher Weise auf die folgende Struktur der Darstellung:

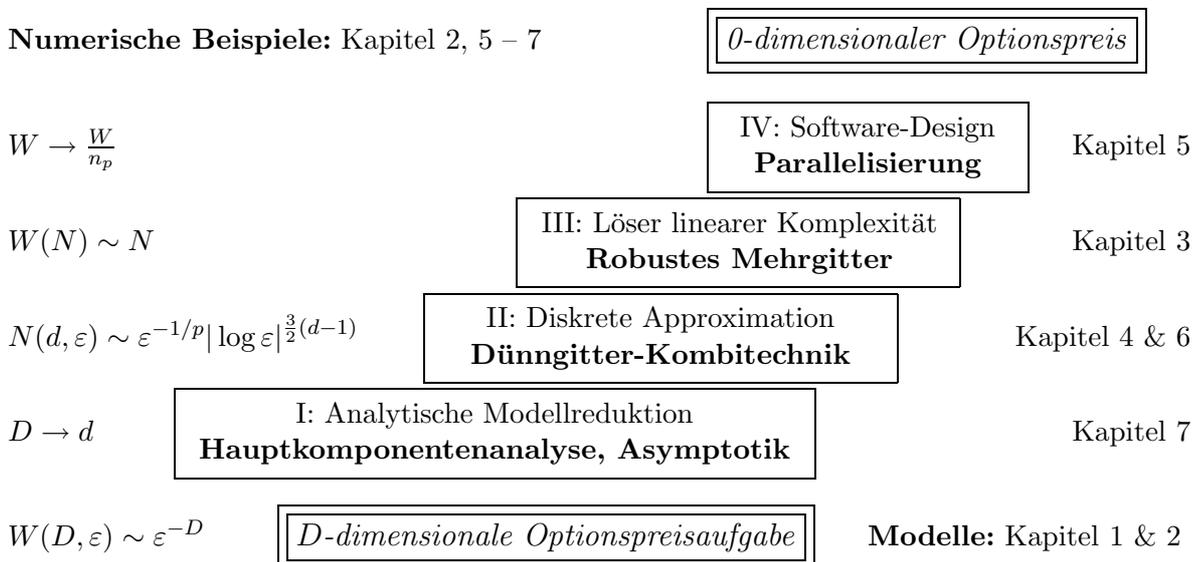


Abbildung 0.2: Der Lösungsweg führt stets über eine, mehrere oder alle der folgenden Stufen:

- I: Zurückführung auf niedrigdimensionale Probleme mittels asymptotischer Analysis.
- II: Diskrete Approximation mit Ansatzräumen (der Dimension N) auf dünnen Gittern.
- III: (Multigrid-)Lösung der diskretisierten Gleichungen mit optimaler Komplexität W .
- IV: Effiziente Implementierung auf einem Parallelrechner mit n_p Prozessoren.

Kapitel 1 stellt die wesentlichen finanzmathematischen Konzepte vor, die der Bewertung von Optionen zugrunde liegen, wobei für Details des Standardstoffes auf die einschlägige Literatur verwiesen wird. Verschiedene für die analytische und numerische Lösung günstige Formulierungen werden für die mehrdimensionalen Black-Scholes-Gleichung abgeleitet.

In **Kapitel 2** werden elementare Diskretisierungen mittels finiter Differenzen sowie impliziter Einschrittverfahren an niedrigdimensionalen Modellen für FX-Optionen und kleine Baskets illustriert, wobei geeignete Gradierungen des Gitters und die Approximation von Ableitungen über (Richardson-)Extrapolation und numerische Differentiation betrachtet werden.

Die Lösung der diskretisierten Systeme mittels Mehrgitterverfahren wird in **Kapitel 3** diskutiert. Im Vordergrund stehen dabei robuste Glätter für anisotrope Gleichungen, wie sie sowohl für dünne Gitter als auch für die hier wichtigen degeneriert parabolischen Operatoren charakteristisch sind. Ein zweiter Schwerpunkt ist die Übertragung auf lineare Komplementaritätsprobleme (Amerikanische Optionen) mittels projizierter Glätter und einseitiger Transferoperatoren am freien Rand.

Die so gewonnenen Lösungen dienen als Bausteine für die Kombinationstechnik dünner Gitter. **Kapitel 4** entwickelt aus der Fehlerdarstellung auf diesen kartesischen Gittern eine Dünngitter-Fehlerabschätzung in geschlossener Form in allgemeinen Dimensionen. Der Beweis offenbart an entscheidender Stelle den verblüffenden Extrapolationsmechanismus.

Eine Diskussion der parallelen Umsetzung und numerische Beispiele, welche einerseits die Komplexitätsvorteile gegenüber klassischen Diskretisierungen verdeutlichen und andererseits die Grenzen der Methode (im Sinne der Dimensionalität) ausloten, finden sich in **Kapitel 5**.

Eine Verallgemeinerung des der Kombinationstechnik zugrundeliegenden Schemas, die in **Kapitel 6** entwickelt wird, vereinigt multivariate Extrapolation mit der Kombinationstechnik, um auf dem dünnen Gitter eine Lösung vierter Ordnung zu erhalten.

Das abschließende **Kapitel 7** demonstriert Zugänge zur Dimensionsreduktion am Beispiel von Basket-Optionen. Für das Black-Scholes-Modell kann analytisch eine asymptotische Entwicklung des Optionspreises für stark korrelierte Aktien angegeben werden. In allgemeinerem Kontext ist eine Zurückführung des hochdimensionalen Problems auf zweidimensionale numerische Aufgaben möglich, die sogar noch eine bessere Approximation liefert.

Einleitung

1 Optionspreismodelle

Überblick

Derivative Finanzinstrumente (*Derivate*) sind abgeleitete Produkte von *Basiswerten* (*Referenzwerten*). Dieses einführende Kapitel dient dazu, die elementaren Prinzipien der Optionspreistheorie zu erläutern und verschiedene Modelle vorzubereiten.

Nach der Begriffsbildung von Abschnitt 1.1 wird in 1.2, ausgehend von ein paar Bemerkungen zu den oftmals beobachteten Eigenschaften von Finanz-Zeitreihen, der Weg vom stochastischen Prozess für den Basiswert zum fairen Preis einer Europäischen Option skizziert.

Nach einem Ausflug zu Amerikanischen und Exotischen Optionen werden in 1.3 mehrere äquivalente Formulierungen der klassischen mehrdimensionalen Black-Scholes-Gleichung abgeleitet, die sich für die numerische Lösung als günstig erweisen.

1.1 Optionen

Eine *Option* mit *Ausübungspreis* K gibt dem Käufer das Recht, einen Basiswert (z. B. eine Aktie) zu einem oder mehreren vorher festgelegten Zeitpunkten in der Zukunft zum Ausübungspreis zu kaufen oder zu verkaufen. Dementsprechend spricht man von einer Kaufs- (*Call*) oder Verkaufsoption (*Put*). Im Gegensatz zu Termingeschäften (*Futures, Forwards*) besteht für den Käufer der Option keine Verpflichtung zur Ausübung, während deren Verkäufer zum Handel gezwungen ist, sollte sich Ersterer für die Ausübung entscheiden (man spricht hier übrigens von *Long* bzw. *Short Positions*).

Verschiedene Optionstypen unterscheiden sich neben der Art des Basiswerts durch die Möglichkeiten der Ausübung, dem Auszahlungsprofil und weiteren technischen Absprachen, die den Kontrakt erweitern.

1.1.1 Kontrakttypen

Eine Vielzahl der gehandelten Standardoptionen gehört einem der folgenden Muster an:

- **Europäische Optionen** können nur zu einem fixen Zeitpunkt T , der *Expiration*, ausgeübt werden.
- Bei **Amerikanischen Optionen** kann die Ausübung jederzeit bis zur Expiration erfolgen. Entscheidend ist hier also auch die Wahl eines optimalen Ausübungszeitpunkts.
- Ähnlich den Amerikanischen sind **Bermuda-Optionen**, die zu vorgeschriebenen (diskreten) Zeitpunkten $\{T_1, \dots, T_N\}$ während ihrer Laufzeit ausgeübt werden dürfen.
- Die Auszahlung **Asiatischer Optionen** hängt vom Mittelwert des Kurses über einen bestimmten Zeitraum ab.

Einem ähnlichen Prinzip gehorchen **Lookback-Optionen**, wobei an die Stelle des Mittelwerts das Maximum bzw. Minimum des Kurses tritt.

- Der Wert einer **Barriere-Option** hängt davon ab, ob der Basiswert während der Laufzeit eine gewisse Schranke durchbricht und verfällt dann entweder (*Knock Out*) oder wird dadurch erst aktiv (*Knock In*). Begriffe wie *Down and Out Call* oder *Up and In Put* etc. erklären sich von selbst.

Weitere Typen

Man kann sich leicht beliebig komplexe Kontrakte zusammenbasteln und eine beeindruckende Fülle sogenannter *Exotischer Optionen* wird tatsächlich gehandelt. Für eine Vielzahl weiterer Beispiele (die im Folgenden nicht weiter betrachtet werden) siehe z. B. [Hul00] und auch www.global-derivatives.com.

1.1.2 Wozu Optionen?

Die Motivation, mit Optionen zu handeln, kann vielfältig sein.

1. **Spekulation:** Für sich betrachtet haben Optionen einen stark spekulativen Charakter, da sie die Schwankungen des Basiswerts um den Ausübungspreis widerspiegeln („Hebeleffekt“).
2. **Arbitrage:** Darunter versteht man im Wesentlichen das Erzielen eines risikolosen Profits aus dem Handel mit Finanzgütern. Es wird in der Regel davon ausgegangen, dass dies nur über sehr kurze Zeiträume möglich sein kann (*No-Arbitrage-Prinzip*), da die Preisanpassung als Ergebnis von Angebot und Nachfrage (*Markteffizienz*) diese Möglichkeit schnell schließt.
3. **Hedging:** Das Ausschalten (die Reduktion) von Risiko durch Ausnutzen der Korrelation von unterschiedlichen Positionen. In diesem Sinne werden Optionen als „Versicherung“ gegen ungünstige Marktentwicklungen verwendet.

Die Punkte 2. und 3. bilden ein fundamentales Modellierungsprinzip von Finanzmärkten: Die Hypothese, dass kein risikofreies Portfolio einen instantanen Gewinn bringen kann, liefert Bedingungen an einen fairen Optionspreis.

1.1.3 Beispiele

Vielfältig wie die Regelung der Auszahlung sind auch die zugrundegelegten Werte selbst. Es seien nur zwei wichtige Beispielklassen erwähnt, die in dieser Arbeit gepreist werden.

Aktien-, Basket- und Indexoptionen

Klassisches Beispiel ist die Option auf eine einzelne Aktie. Eine naheliegende Variante sind Optionen auf Aktienkörbe, wo also eine gewichtete Summe von Aktienkursen betrachtet wird. Ähnlich verhält es sich bei den Optionen auf Indizes — etwa dem DAX 30, FTSE 100, S&P 500 etc. oder Branchenindizes wie dem NASDAQ.

Die genannten Optionen werden sowohl in Amerikanischer als auch Europäischer Form gehandelt.

Wechselkursoptionen

Man sichert sich heute einen gewissen Wechselkurs für ein bestimmtes Nominal einer Fremdwährung zu einem zukünftigen Termin. Die Bedeutung liegt offensichtlich in der Absicherung zukünftiger Geschäfte gegen Kursschwankungen.

1.2 Modellierung von Basiswert und Option

1.2.1 Eigenschaften von Finanz-Zeitreihen

Natürlicher Ausgangspunkt für die stochastische Modellierung von Optionspreisen sind die am Markt beobachteten Zeitreihen der Basiswerte [Con01]. Deren Eigenschaften, die über die Annahme einer konkreten Verteilung hinaus einer Klasse von Daten (Instrumente, Märkte, Zeiträume) gemein sind, bezeichnet man gerne als *Stylized Empirical Facts*.

Wenn $S(t)$ der Kurs zum Zeitpunkt t ist und $Y(t) := \ln S(t)$, dann ist es sinnvoll und gebräuchlich, den *Log Return*

$$r(t, \Delta t) := Y(t + \Delta t) - Y(t)$$

zu betrachten.

Verteilung

Die große Unbekannte ist nun die Verteilung der Zuwächse r nach einem Zeitintervall Δt . Frühe Arbeiten legen durchwegs eine Normalverteilung zugrunde, was noch immer gängige Praxis ist und eine Reihe von praktischen Vorteilen aufweist. In einer Vielzahl von Fällen stellt man jedoch qualitative Abweichungen fest: Insbesondere über kurze Zeiträume werden sogenannte *Fat Tails* beobachtet, die durch ein Potenzgesetz besser erfasst werden. Viele Arbeiten haben sich seitdem mit dem *Fitting* von parametrischen Modellen an Marktdaten beschäftigt, siehe z. B. [BN97].

Korrelationen

1. Man beobachtet, dass die **lineare Autokorrelationsfunktion**

$$A(\tau) := \langle r(t, \Delta t), r(t + \tau, \Delta t) \rangle$$

mit dem Skalarprodukt $\langle \cdot, \cdot \rangle$ sehr schnell (typischerweise im Bereich von Minuten) gegen 0 geht. Das erklärt sich aus dem, was man *Markteffizienz* nennt: gäbe es eine signifikante, länger andauernde Autokorrelation, könnte man aus der Prognose Gewinn machen, was der Korrelation entgegenwirken würde. Dazwischen sieht man also nur die kurze Reaktionszeit des Marktes.

2. Im Hinblick auf den multivariaten Fall ist die **Querkorrelation**

$$\begin{aligned} \sigma_{ij} &:= \langle r_i(t, T), r_j(t, T) \rangle \\ &=: \sigma_i \sigma_j \rho_{ij} \end{aligned}$$

zwischen den einzelnen Komponenten von Interesse, wobei $\sigma_i = \sqrt{\langle r_i(t, T), r_i(t, T) \rangle}$ die Volatilität des i -ten Prozesses ist und die letzte Gleichung als Definition für die Korrelationen ρ_{ij} fungiert.

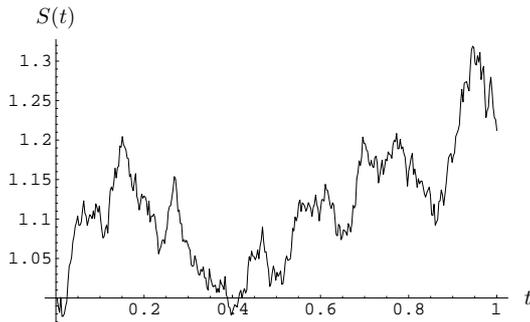


Abbildung 1.1: Pfad einer geometrischen Brownschen Bewegung (1.1) mit $\mu = 0.1$, $\sigma = 0.2$.

3. Im Gegensatz zu 1. stellt man aber sehr wohl eine positive Korrelation von nichtlinearen abgeleiteten Größen fest (die Inkremente sind also nicht unabhängig). Diesen Effekt kann man schon an den Quadraten festmachen:

$$A_2(\tau) := \langle r(t, \Delta t)^2, r(t + \tau, \Delta t)^2 \rangle$$

ist ein Maß für die *Volatilität* des Prozesses und zeigt typischerweise ein langsames Abfallen mit $\tau^{-\beta}$, $\beta \in [0.2, 0.4]$. Man beobachtet sog. *Cluster* hoher und niedriger Volatilität, was unter dem Terminus **Volatility Clustering** bekannt ist.

4. Derartige Maße für die Volatilität sind zumeist negativ mit den Kursänderungen korreliert (sogenanntes **Leverage**).

1.2.2 Das Black-Scholes-Modell — Aktienoptionen

Geometrische Brownsche Bewegung

In den ursprünglichen Arbeiten von Black, Scholes [BS73] und Merton [Mer73] wird der Wertentwicklung einer Aktie eine *geometrische Brownsche Bewegung*

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) \tag{1.1}$$

mit konstanter *Volatilität* σ und *Drift* μ zugrundegelegt. $W(t)$ ist hier ein *Wiener-Prozess* (wegen der physikalischen Anschauung auch *Brownsche Bewegung*), der stetig ist und unabhängige, normalverteilte Inkremente

$$W(t + \tau) - W(t) \sim \mathcal{N}(0, \sqrt{\tau})$$

hat. Es sei bemerkt, dass (1.1) in der mathematisch sinnvollen integrierten Form zu interpretieren ist, der die Integration über die Brownsche Bewegung W (*Ito-Integral*) zugrundeliegt. Abb. 1.1 zeigt eine mögliche Realisierung von (1.1) für spezielles μ und σ . Die postulierte logarithmische Normalverteilung

$$\ln S(t) \sim \mathcal{N}\left(S(0) + (\mu - \sigma^2/2)t, \sigma\sqrt{t}\right)$$

der Zuwächse deckt sich jedoch z. T. nicht mit den beobachteten Marktdaten. Darüberhinaus schließt die Stetigkeit Marktsprünge aus. Trotz dieser Mängel werden leicht modifizierte Modelle (in denen etwa die Volatilität durch ein zeitabhängiges $\sigma(t)$ ersetzt ist) wegen ihrer guten Handhabbarkeit in der Praxis häufig eingesetzt.

Der Arbitrage-freie Optionspreis

Es wird im Folgenden kurz der Weg vom zugrundeliegenden Prozess (1.1) für S zum Preis einer (Europäischen) Option auf S skizziert, unter Berufung auf Resultate der stochastischen Analysis, die in der einschlägigen Literatur — siehe etwa [Øks98], [Irl98], [Eth02] für eine rigorose Abhandlung oder [WHD95] für einen heuristischen Zugang — ausführlich behandelt sind. Es wird davon ausgegangen, dass die anfallenden Transaktionskosten vernachlässigbar sind, keine Dividenden ausgeschüttet werden und der Zinssatz (*Short Rate*) deterministisch und konstant gleich r ist.

Der Preis der Option ist dann dadurch bestimmt, dass es eine *selbstfinanzierende Handelsstrategie* gibt, mit der die Forderung an den Aussteller der Option am Ende der Laufzeit exakt *repliziert* werden kann. Der Ertrag aus dem Verkauf der Option wird so in Aktien und risikolose Bonds investiert — unter ständiger Umschichtung, aber ohne weiteren Zuschuss von Mitteln —, dass am Ende die verlangte Auszahlung bestritten werden kann.

Es ist jetzt die zentrale Aussage des *Fundamentalsatzes der Preistheorie*, dass Arbitragefreiheit äquivalent ist mit der Existenz eines (eindeutigen) äquivalenten (sogenannten *risikoneutralen*) Wahrscheinlichkeitsmaßes P^* , unter dem (1.1) ein Martingal ist. Der Preis der Option u ist dann der diskontierte Erwartungswert der Auszahlung $g(S)$ zum Zeitpunkt T unter diesem Maß (vgl. *Satz von Girsanov*):

$$u = E^* \left\{ e^{-r(T-t)} g(S(T)) \right\}. \quad (1.2)$$

Die Feynman-Kac-Formel stellt schließlich den Zusammenhang zwischen diesem Erwartungswert und der parabolischen Gleichung

$$\frac{\partial u}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} + r \left(S \frac{\partial u}{\partial S} - u \right) = 0 \quad \forall S > 0, t \in [0, T] \quad (1.3)$$

für u her. Zur Expiration ist der Wert der Option klarerweise

$$u(S, T) = g(S) := \begin{cases} (S - K)_+ := \max(S - K, 0) & \text{für einen Call,} \\ (K - S)_+ := \max(K - S, 0) & \text{für einen Put.} \end{cases}$$

Ausgehend von dieser Endbedingung wird die Evolution der Lösung auf der positiven Achse rückwärts bis $t = 0$ verfolgt. Man beachte, dass (1.3) eine Anti-Diffusionsgleichung ist, die rückwärts in der Zeit gelöst wird, also „korrekt gestellt“ ist. Eine Randbedingung für $S = 0$ ist nicht notwendig, da (1.3) dort zu einer gewöhnlichen Differentialgleichung degeneriert.

Durch Transformation von (1.3) auf die Wärmeleitungsgleichung können vermöge deren Greenscher Funktion analytische Formeln angegeben werden:

- Call:

$$u = SN(d_1) - Ke^{-rt}N(d_2),$$

- Put:

$$u = Ke^{-rt}N(-d_2) - SN(-d_1), \quad (1.4)$$

wobei

$$d_{1/2} = \frac{\ln(S/K) + (r \pm \sigma^2/2)t}{\sigma\sqrt{t}} \quad \text{und} \\ N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-s^2/2} ds$$

die Verteilungsfunktion der Standard-Normalverteilung ist.

Black-Scholes — multivariat

Im Fall multivariater Prozesse

$$\frac{dS_i(t)}{S_i(t)} = \mu_i(t) dt + dW_i(t), \quad i \in \{1, \dots, d\},$$

wobei W_i korrelierte Wiener-Prozesse mit

$$\langle dW_i, dW_j \rangle = \sigma_i \sigma_j \rho_{ij} dt$$

sind, führen ähnliche Argumente wie im eindimensionalen Fall auf eine d -dimensionale partielle Differentialgleichung der Gestalt

$$\frac{\partial u}{\partial t} = \mathcal{L}_{BS} u \quad \text{mit} \tag{1.5}$$

$$\mathcal{L}_{BS} u := -\frac{1}{2} \sum_{i,j=1}^d \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 u}{\partial S_i \partial S_j} - r \sum_{i=1}^d S_i \frac{\partial u}{\partial S_i} + ru. \tag{1.6}$$

Die Überführung in ein Anfangswertproblem durch Zeitumkehr ergibt beispielsweise für eine Put-Option auf einen Basket aus d Aktien das Anfangswertproblem

$$\begin{aligned} \frac{\partial u}{\partial t} + \mathcal{L}_{BS} u &= 0 & \forall (\mathbf{S}, t) \in \mathbb{R}_+^n \times (0, T), \\ u(\mathbf{S}, 0) = u_0(\mathbf{S}) &= \left(K - \sum_{i=1}^d \mu_i S_i \right)_+ & \forall \mathbf{S} \in \mathbb{R}_+^d. \end{aligned} \tag{1.7}$$

Für $S_i = 0$ gelten wieder natürliche Randbedingungen, da alle Koeffizienten auf diesen Rändern verschwinden. Für $|\mathbf{S}| \rightarrow \infty$ können oftmals asymptotische Werte als Dirichlet-Daten auf einem hinreichend großen Gebiet herangezogen werden (hier etwa $u = 0$). Eine weitere Möglichkeit ist die Transformation auf ein endliches Gebiet durch Einführung von Variablen der Gestalt

$$x_i = \frac{S_i}{K + S_i} \in [0, 1].$$

Auf dem Würfel $[0, 1]^d$ gelten wieder natürliche Randbedingungen. In [ZL03] werden solche Formulierungen als Ausgangspunkt für die Eindeigkeitstheorie von Lösungen genommen.

Bemerkung. Die Gleichung (1.7) ist invariant unter Skalierung der Koordinaten, sodass o. B. d. A. das durch Abschneiden künstlich begrenzte Gebiet als Einheitswürfel betrachtet werden kann. Einziger Referenzwert ist der Ausübungspreis K .

1.2.3 Stochastische Volatilitätsmodelle

Da sich die Volatilität in der Praxis nicht als konstant erweist, sondern selbst zufälligen Schwankungen unterliegt, wurden (erfolgreich) Versuche unternommen, sie in Abhängigkeit eines stochastischen Hintergrundprozesses zu modellieren:

$$dS(t)/S(t) = \mu dt + f(Y(t)) dW(t),$$

1 Optionspreismodelle

wobei f eine positive Funktion ist und Y z. B. wieder als Diffusionsprozess gewählt wird. Gestützt auf die Beobachtung, dass Volatilitäten über längere Zeiträume Schwankungen um einen Mittelwert ausführen (*Mean Reversion*), wird gerne ein Ornstein-Uhlenbeck Prozess der Art

$$dY(t) = \alpha(m - Y(t)) dt + \beta dZ(t) \quad (1.8)$$

mit Mittelwert m und Wiener-Prozess $Z(t)$ angesetzt, der zum Aktien-Prozess (typischerweise negativ) korreliert ist:

$$\langle dW(t), dZ(t) \rangle = \rho dt.$$

Da sich die Volatilität nicht mit am Markt gehandelten Produkten *Hedgen* lässt, ist das risikoneutrale Wahrscheinlichkeitsmaß nicht eindeutig. Vielmehr gibt es eine Familie von Maßen $P^{*(\gamma)}$, unter denen der Preisprozess

$$\begin{aligned} dS(t) &= rS(t) dt + f(Y(t))S(t) dW^*(t), \\ dY(t) &= \left[\alpha(m - Y(t)) - \beta \left(\rho \frac{\mu - r}{f(Y(t))} + \gamma \sqrt{1 - \rho^2} \right) \right] dt + \beta dZ_t^* \end{aligned}$$

die Martingaleigenschaft hat. Der hierdurch nicht eindeutig bestimmte Optionspreis

$$u = E^{*(\gamma)} \left\{ e^{-r(T-t)} g(S(T)) \right\} \quad (1.9)$$

hängt gleichermaßen von der Wahl von γ ab, das „durch den Markt“ bestimmt wird. $\gamma = \gamma(y)$ wird daher als *Marktpreis des Volatilitätsrisikos* bezeichnet.

PDE-Zugang

Die (1.9) entsprechende Feynman-Kac-PDE lautet

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{1}{2} f(Y)^2 S^2 \frac{\partial^2 u}{\partial S^2} + r \left(S \frac{\partial u}{\partial S} - u \right) + \\ \rho \beta S f(Y) \frac{\partial^2 u}{\partial S \partial Y} + \frac{1}{2} \beta^2 \frac{\partial^2 u}{\partial Y^2} + \alpha(m - Y) \frac{\partial u}{\partial Y} - \beta \Lambda \frac{\partial u}{\partial Y} = 0, \end{aligned}$$

wobei

$$\Lambda = \rho \frac{\mu - r}{f(Y)} + \gamma \sqrt{1 - \rho^2}.$$

Bemerkung (Allgemeiner Rahmen). Allgemein ergeben sich aus der Modellierung der Basiswerte und Parameter in Abhängigkeit von (geometrischen) Brownschen Bewegungen parabolische Gleichungen vom Typ

$$\frac{\partial u}{\partial t} + \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial u}{\partial x_i} + ru = 0, \quad (1.10)$$

wobei typischerweise

$$a_{ij}(\mathbf{x}) = 0 \quad \text{für } \mathbf{x} \in \Gamma, \quad (1.11)$$

$$b_i(\mathbf{x}) = 0 \quad \text{für } \mathbf{x} \in \Gamma \text{ oder zumindest} \quad (1.12)$$

$$b_i(\mathbf{x}) n_i(\mathbf{x}) \geq 0 \quad \text{für } \mathbf{x} \in \Gamma \quad (1.13)$$

mit der Außennormalen $n_i(\mathbf{x})$, wobei Γ ein Teil des Randes ist. Diese Bedingungen drücken die Beschränktheit der Prozesse aus und garantieren unter gängigen Voraussetzungen die Eindeutigkeit der Lösung ohne Randbedingungen (auf Γ) [ZL03].

Asymptotik

Bei Sichtung der Marktdaten im Hinblick auf (1.8) findet man hohe Frequenzen α : Umkehrperioden von $1\frac{1}{2}$ Tagen (im Vergleich zu Laufzeiten von Monaten) sind typisch für Indizes wie den S&P 500 und führen zum *Volatility Clustering* (vgl. 1.2.1, 3.).

Darauf aufbauend betreiben Fouque, Papanicolaou und Sircar [FPS00] eine Asymptotik im kleinen Parameter $\epsilon = 1/\alpha$. Mit $\nu^2 = \beta^2/(2\alpha)$ werden

$$\begin{aligned} dS^\epsilon(t) &= rS^\epsilon(t) dt + f(Y^\epsilon(t))S^\epsilon(t) dW^*(t) \\ dY^\epsilon(t) &= \left[\frac{1}{\epsilon}(m - Y^\epsilon(t)) - \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\Lambda(Y^\epsilon(t)) \right] dt + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} dZ^*(t) \end{aligned}$$

bzw. die PDE

$$\begin{aligned} \frac{\partial u^\epsilon}{\partial t} + \frac{1}{2}f(Y)^2 S^2 \frac{\partial^2 u^\epsilon}{\partial S^2} + r \left(S \frac{\partial u^\epsilon}{\partial S} - u^\epsilon \right) + \\ \rho \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} S f(Y) \frac{\partial^2 u^\epsilon}{\partial S \partial Y} + \frac{1}{2} \frac{\nu^2}{\epsilon} \frac{\partial^2 u^\epsilon}{\partial Y^2} + \alpha(m - y) \frac{\partial u^\epsilon}{\partial Y} - \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} \Lambda \frac{\partial u^\epsilon}{\partial Y} = 0 \end{aligned}$$

für kleine ϵ betrachtet. Es ergeben sich die Korrekturen

$$u^\epsilon = u^0 - \sqrt{\epsilon}(T - t) \left(v S^2 \frac{\partial^2 u^0}{\partial S^2} + w S^3 \frac{\partial^3 u^0}{\partial S^3} \right) + O(\epsilon)$$

zum Black-Scholes Preis u^0 , bzw. im multivariaten Fall

$$u_\epsilon(S, t) = u^0(S, t) + \sqrt{\epsilon}(T - t) \left(\sum_{i,j=1}^d v_{ij} S_i S_j \frac{\partial^2 u^0}{\partial S_i \partial S_j} + \sum_{i,j,k=1}^d w_{ijk} S_i S_j S_k \frac{\partial^3 u^0}{\partial S_i \partial S_j \partial S_k} \right) + O(\epsilon).$$

Bemerkung. Der Preis kann also bis auf Terme der Ordnung ϵ durch den Black-Scholes-Preis und dessen Ableitungen bis zur dritten Ordnung ausgedrückt werden. Die numerischen Verfahren in den folgenden Kapiteln werden genau dieser Anforderung gerecht.

Das Kalibrierungsproblem reduziert sich auf die Parameter v und w (v_{ij} und w_{ijk}) anstatt der unbekanntten Funktion γ .

1.2.4 Amerikanische und Exotische Optionen

Amerikanische Optionen

Durch die freie Wahl des Ausübungszeitpunkts wird aus (1.2) das *Optimale-Stopp-Problem*

$$u = \sup_{0 \leq \tau \leq t} E^* \left\{ e^{-(\tau-t)} g(S(T)) \right\}.$$

An die Stelle der PDE (1.7) tritt im Black-Scholes-Fall das *lineare Komplementaritätsproblem*

$$\frac{\partial u}{\partial t} + \mathcal{L}_{BS} u \leq 0, \quad (1.14)$$

$$u \geq g, \quad (1.15)$$

$$\left(\frac{\partial u}{\partial t} + \mathcal{L}_{BS} u \right) \cdot (u - g) = 0 \quad (1.16)$$

1 Optionspreismodelle

mit dem Black-Scholes-Operator aus (1.6). Für eine heuristische Ableitung aus Arbitrage-Argumenten siehe etwa [WHD95], wo sich ebenso eine Formulierung als äquivalentes *freies Randwertproblem* findet, das aus der *Smooth Fit Condition*, z. B.

$$\begin{aligned}u(S_f(t), t) &= K - S_f(t), \\ \frac{\partial u}{\partial S}(S_f(t), t) &= -1\end{aligned}$$

am freien Rand $S_f(t)$ für einen (eindimensionalen) Put, abgeleitet werden kann. Für mehrdimensionale Probleme gelten die naheliegenden Verallgemeinerungen.

Die Menge $\{u = g\}$ wird als *Exercise Region* (*Ausübungsbereich*), das Komplement als *Continuation Region* bezeichnet.

Weitere Beispiele

Eine Vielzahl weiterer Optionen lässt sich durch partielle Differentialgleichungen mit modifizierten Anfangsbedingungen (*Digital Options*, *Best-Of-Options*, etc.) oder angepassten Randbedingungen (z. B. *Barriere-Optionen*) modellieren. Bei pfadabhängigen Optionen sind mitunter zusätzliche Variablen einzuführen (etwa bei *Asiatischen Optionen* mit kontinuierlichem Abtasten der Mittelwert oder bei *Lookbacks* das Minimum bzw. Maximum). Geldflüsse oder für den Wert der Option relevante Ereignisse zu diskreten Zeitpunkten lassen sich in der Regel direkt einarbeiten (z. B. bei *Dividenden* oder *Bermuda-Optionen*).

Modelle und numerische Methoden für solche Kontrakte finden sich unter anderem in den Arbeiten von Oosterlee [OFG04], Wilmott, Howison, Dewynne [WDH93] und Forsyth [PFV99].

1.2.5 Wechselkurs-Optionen

Wir betrachten den Wechselkurs $S(t)$ zum Zeitpunkt t zwischen zwei Währungen in Märkten mit Zins r_d (*Domestic*) und r_f (*Foreign*):

$$dS(t) = (r_d - r_f)S(t)dt + \sigma_S S(t)dW(t).$$

Für kurze Laufzeiten (etwa unter 3 Jahre) können die Zinsen in guter Näherung als konstant angesehen werden und man erhält für den Preis die eindimensionale (Rückwärts-)Gleichung

$$\frac{\partial u}{\partial t} - \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 u}{\partial S^2} - (r_d - r_f)S \frac{\partial u}{\partial S} + r_d u = 0. \quad (1.17)$$

Für längere Laufzeiten jedoch müssen die Zinssätze in den jeweiligen Märkten auch als stochastische Prozesse betrachtet werden:

$$dr_i(t) = \mu_i(r_i(t), t)dt + \nu_i(r_i(t), t)dW_i(t),$$

wobei gängige Modelle von der Gestalt

$$\begin{aligned}\mu_i(r_i, t) &= \kappa_i(\theta_i(t) - r_i) \\ \nu_i(r_i, t) &= r_i^{\beta_i} \sigma_i(t)\end{aligned}$$

mit positiven Parametern κ_i , β_i und einem zeitabhängigen Reversionsniveau θ_i sind, das zur Kalibrierung der Zinsstruktur herangezogen werden kann.

1.3 Mehrdimensionale Black-Scholes-Gleichung und Transformationen

Die entsprechende partielle Differentialgleichung lautet ($\mathbf{x} = (S, r_d, r_f)$)

$$\frac{\partial u}{\partial t} - \frac{1}{2} \sum_{i,j=1}^3 \nu_i \nu_j \rho_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^d \mu_i \frac{\partial u}{\partial x_i} + \rho_{31} \nu_3 \nu_1 \frac{\partial u}{\partial x_3} + x_2 u = 0, \quad (1.18)$$

wobei $\nu_1 = \sigma_S S$ und ρ_{ij} die Korrelationen der einzelnen Wiener-Prozesse sind. Anfangsbedingung ist jeweils für einen Put

$$u(\mathbf{x}, 0) = (K - x_1)_+.$$

1.3 Mehrdimensionale Black-Scholes-Gleichung und Transformationen

Als hochdimensionales Modellproblem dient im Folgenden oft das Black-Scholes-Modell

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{1}{2} \sum_{i,j=1}^d \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 u}{\partial S_i \partial S_j} - r \sum_{i=1}^d S_i \frac{\partial u}{\partial S_i} + ru = 0 & \quad \forall (\mathbf{S}, t) \in \mathbb{R}_+^d \times (0, T), \\ u(\mathbf{S}, 0) = g(\mathbf{S}) := \left(K - \sum_{i=1}^d \mu_i S_i \right)_+ & \quad \forall \mathbf{S} \in \mathbb{R}_+^d \end{aligned}$$

für eine Option auf einen Basket $\sum_{i=1}^d \mu_i S_i$ mit positiven Gewichten μ_i .

Verschiedene Formulierungen können für die Analysis und numerische Lösung nützlich sein.

1.3.1 Logarithmische Koordinaten

Eine gängige Transformation ist der sogenannte *Log Price*

$$y_i := \ln S_i, \quad 1 \leq i \leq d,$$

der auf die Gleichung

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{1}{2} \sum_{i,j=1}^n \sigma_i \sigma_j \rho_{ij} \frac{\partial^2 u}{\partial y_i \partial y_j} - \sum_{i=1}^n \left(r - \frac{1}{2} \sigma_i^2 \right) \frac{\partial u}{\partial y_i} + ru = 0 & \quad \forall (\mathbf{y}, t) \in \mathbb{R}^n \times (0, T), \\ u(\mathbf{y}, 0) = g(e^{\mathbf{y}}) = \left(K - \sum_{i=1}^d \mu_i e^{y_i} \right)_+ & \quad \forall \mathbf{y} \in \mathbb{R}^d \end{aligned}$$

führt.

Vor- und Nachteile:

- + konstante Koeffizienten
- das „doppelt unendliche“ Gebiet (durch die Zuordnung $\log :]0, \infty[\rightarrow]-\infty, \infty[$) erzeugt ein zusätzliches Abschneideproblem
- der nichtlineare Knick in der Anfangsbedingung kann nicht exakt am Diskretisierungsgitter erfasst werden

1.3.2 Problemangepasste Koordinaten

Eine Möglichkeit, um die Probleme aus 1.3.1 zu vermeiden, ist die Wahl des Basket selbst als Koordinate und die Einführung von „Winkeln“ um diese Richtung:

$$\begin{aligned} x_1 &= \sum_{k=1}^d S_k, \\ x_i &= \frac{S_{i-1}}{\sum_{k=i-1}^d S_k}, \quad 1 < i \leq d \end{aligned}$$

(die Gewichte wurden o. B. d. A. auf 1 gesetzt). Umgekehrt ergibt sich

$$\begin{aligned} S_1 &= x_1 x_2, \\ S_i &= x_1 x_{i+1} \prod_{k=2}^i (1 - x_k), \quad 2 < i < d, \\ S_d &= x_1 \prod_{k=2}^d (1 - x_k). \end{aligned}$$

Die Transformation ist ein Diffeomorphismus $]0, 1[^d \rightarrow]0, d[\times]0, 1[^{d-1}$, unstetig auf $[0, 1]^{d-2} \times \{0\} \times \{0\}$. Der Operator

$$\mathcal{L}_{BS} = \sum_{i,j=1}^d a_{ij}(\mathbf{S}) \frac{\partial^2}{\partial S_i \partial S_j} + \sum_{i=1}^d b_i(\mathbf{S}) \frac{\partial}{\partial S_i} + c(\mathbf{S})$$

transformiert zu

$$\hat{\mathcal{L}}_{BS} = \sum_{i,j=1}^d \alpha_{ij}(\mathbf{x}) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d \beta_i(\mathbf{x}) \frac{\partial}{\partial x_i} + \gamma(\mathbf{x})$$

mit

$$\begin{aligned} \alpha_{ij}(\mathbf{x}) &= \sum_{k,l=1}^d a_{kl}(\mathbf{S}(\mathbf{x})) \frac{\partial x_i}{\partial S_k} \frac{\partial x_j}{\partial S_l}, \\ \beta_i(\mathbf{x}) &= \sum_{k=1}^d b_k(\mathbf{S}(\mathbf{x})) \frac{\partial x_i}{\partial S_k} + \sum_{k,l=1}^d a_{kl}(\mathbf{S}(\mathbf{x})) \frac{\partial^2 x_i}{\partial S_k \partial S_l}, \\ \gamma(\mathbf{x}) &= c(\mathbf{S}(\mathbf{x})). \end{aligned}$$

Langwierige, einfache Rechnung ergibt

$$\begin{aligned} \alpha_{11}(\mathbf{x}) &= x_1^2 \sum_{k,l=1}^d \rho_{kl} f_{1k}(\mathbf{x}) f_{1l}(\mathbf{x}), \\ \alpha_{1j}(\mathbf{x}) &= x_1 x_j (1 - x_j) \sum_{k,l=1}^d (\rho_{k,j-1} - \rho_{kl}) f_{1k}(\mathbf{x}) f_{jl}(\mathbf{x}) \quad \forall 1 < j \leq d, \\ \alpha_{j1}(\mathbf{x}) &= \alpha_{1j}(\mathbf{x}) \quad \forall 1 \leq j \leq d, \\ \alpha_{ij}(\mathbf{x}) &= x_i (1 - x_i) x_j (1 - x_j) \sum_{k,l=1}^d [\rho_{kl} - \rho_{i-1,l} - \rho_{k,j-1} + \rho_{i-1,j-1}] f_{ik}(\mathbf{x}) f_{jl}(\mathbf{x}) \quad \forall 1 < i, j \leq d \end{aligned}$$

1.3 Mehrdimensionale Black-Scholes-Gleichung und Transformationen

mit

$$f_{ik}(\mathbf{x}) := \begin{cases} x_{k+1} \prod_{j=i+1}^k (1 - x_j) & i < k < d, \\ \prod_{j=i+1}^k (1 - x_j) & i < k = d, \\ x_{k+1} & i = k < d, \\ 1 & i = k = d, \\ 0 & i > k \end{cases}$$

aus sukzessiver Anwendung der Beziehungen

$$\begin{aligned} f_{1k}(\mathbf{x})x_1 &= S_k \quad \forall k, \\ f_{ik}(\mathbf{x})(1 - x_i)S_{i-1} &= x_i S_k \quad \forall 1 < i \leq k \end{aligned}$$

sowie

$$\sum_{k=1}^d f_{ik}(\mathbf{x}) = \sum_{k=i}^d f_{ik}(\mathbf{x}) = 1.$$

Ähnlich folgt

$$\begin{aligned} \beta_1(\mathbf{x}) &= rx_1, \\ \beta_i(\mathbf{x}) &= x_i(1 - x_i) \sum_{k,l=1}^d [-2\rho_{i-1,i-1}x_i + (2x_i - 1)(\rho_{k,i-1} + \rho_{l,i-1}) + 2(1 - x_i)\rho_{kl}] f_{ik}(\mathbf{x})f_{il}(\mathbf{x}) \end{aligned}$$

für alle $i > 1$.

Bemerkung. Alle Koeffizienten a_{ij} , b_j ($i \geq 1, j > 1$) verschwinden für $x_j \rightarrow \{0, 1\}$, a_{i1} und b_1 verschwinden zudem für $x_1 \rightarrow 0$.

Die Setzung von Randbedingungen ist daher nicht notwendig (und nicht sinnvoll). Das transformierte Gebiet ist nur für $x_1 \rightarrow \infty$ unbeschränkt.

Vor- und Nachteile:

- + gute Approximation der Hauptkomponente
- + exakte Erfassung der Anfangsbedingung (die nur von x_1 abhängt) auf einem geeignet gewählten kartesischen Gitter
- aufwändige Auswertung der Koeffizienten

1.3.3 Hauptachsentransformation

Ausgangspunkt sind die logarithmischen Koordinaten aus 1.3.1, welche die stochastische Differentialgleichung

$$dY_i = \mu_i dt + dW_i$$

mit

$$\langle dY_i, dY_j \rangle = \sigma_{ij} dt$$

erfüllen. Da $\Sigma := (\sigma_{ij})$ symmetrisch positiv definit ist, gibt es eine orthogonale Transformation \mathbf{Q} auf Diagonalform:

$$\mathbf{Q}\Sigma\mathbf{Q}^T =: \text{diag}(\lambda_i) =: \Lambda.$$

1 Optionspreismodelle

Mit

$$\mathbf{Z} := \mathbf{QY}$$

folgt

$$\langle dZ_i, dZ_j \rangle = \lambda_i \delta_{ij} dt.$$

Die Black-Scholes-Gleichung transformiert zu

$$\frac{\partial u}{\partial t} - \frac{1}{2} \sum_{i=1}^d \lambda_i \frac{\partial^2 u}{\partial z_i^2} - \sum_{i=1}^d \sum_{j=1}^d q_{ij} \left(r - \frac{1}{2} \sigma_j^2 \right) \frac{\partial u}{\partial z_i} + ru = 0 \quad \forall (\mathbf{z}, t) \in \mathbb{R}^d \times (0, T), \quad (1.19)$$

$$u(\mathbf{z}, 0) = g(e^{\mathbf{Q}^T \mathbf{z}}) = \left(K - \sum_{i=1}^d \mu_i e^{\sum_{j=1}^d q_{ji} z_j} \right)_+ \quad \forall \mathbf{z} \in \mathbb{R}^d \quad (1.20)$$

bzw. nach der Translation

$$\begin{aligned} \mathbf{x} &:= \mathbf{z} - t\mathbf{b}, \\ b_i &:= \sum_{j=1}^d q_{ij} \left(r - \frac{1}{2} \sigma_j^2 \right) \end{aligned}$$

zu

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^d \lambda_i \frac{\partial^2 u}{\partial x_i^2} - ru \quad \forall (\mathbf{x}, t) \in \mathbb{R}^d \times (0, T), \quad (1.21)$$

$$u(\mathbf{x}, 0) = g(e^{\mathbf{Q}^T \mathbf{x}}) = \left(K - \sum_{i=1}^d \mu_i e^{\sum_{j=1}^d q_{ji} x_j} \right)_+ \quad \forall \mathbf{x} \in \mathbb{R}^d. \quad (1.22)$$

Bemerkung. Durch die Substitution $v = e^{rt}u$ kann noch der „Reaktionsterm“ ru in (1.21) zum Verschwinden gebracht werden und man erhält die Wärmeleitungsgleichung. Die Lösung in einem Punkt \mathbf{S}_0 hat daher eine Darstellung über deren Greensche Funktion:

$$u(\mathbf{S}_0, t) = \frac{e^{-rt}}{\sqrt{2^d t^d \pi^d \det \Sigma}} \int_{\mathbb{R}^d} g(\mathbf{S}(\mathbf{y}, t)) e^{-\frac{1}{2t} \sum_{i=1}^d (x_i(\mathbf{S}_0, t) - y_i)^2 / \lambda_i} d\mathbf{y}. \quad (1.23)$$

Vor- und Nachteile:

- wie in 1.3.1 bei logarithmischen Koordinaten
- + keine Terme erster Ordnung
- + diagonale Diffusion
- + Lösungsdarstellung mittels Greenscher Funktion (\rightarrow Integrationsproblem)
- + Ausgangspunkt für Hauptkomponentenanalyse (Kapitel 7)

2 Diskretisierung auf gradierten kartesischen Gittern

Überblick

Die numerische Lösung der linearen parabolischen Probleme ist im Prinzip den wohlbekannten Standardmethoden zugänglich und wird „nur“ durch den enormen Aufwand in hohen Dimensionen undurchführbar. Daher wird nur kurz die später verwendete Diskretisierung der „Ortskoordinaten“ (also des elliptischen Teils des Differentialoperators) mit finiten Differenzen (Abschnitt 2.1.1) und das implizite Zeitschrittverfahren (Abschnitt 2.1.2) beschrieben.

Sehr nützlich ist in diesem Zusammenhang die zeitabhängige Gradierung des Gitters, die Rücksicht auf die fehlende Glattheit am — für Optionspreisaufgaben typischen — Knick der Anfangsbedingung nimmt. Große Bedeutung kommt in der Praxis den Ableitungen der Lösung (den Sensitivitäten des Optionspreises nach den zugrundeliegenden Kursen) zu, die durch Richardson-Extrapolation und numerische Differentiation approximiert werden. Zur Illustration dienen Wechselkursderivate (3-Faktor-Modell) und Aktienoptionen (2-Faktor-Modell).

Die für die hier diskutierten impliziten Verfahren erforderliche Lösung der linearen diskretisierten Systeme ist Gegenstand von Kapitel 3.

2.1 Diskretisierung

2.1.1 Ortsdiskretisierung

Die Diskretisierung der d -dimensionalen Gleichungen vom Typ

$$\frac{\partial u}{\partial t} + \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial u}{\partial x_i} + ru = 0$$

mit fast überall positiv definitem $(a_{ij})_{1 \leq i,j \leq d}$ kann sehr einfach gehalten werden. Da die natürliche Beschränkung der Koordinaten typischerweise Intervalle sind (entweder zwischen den Schranken eines beschränkten Prozesses, einer beschränkten Transformierten oder eines geeignet abgeschnittenen Intervalls), bietet sich auf dem quaderförmigen Gebiet ein kartesisches Gitter an.

Die hier betrachteten Modelle sind durchwegs diffusionsdominant, mit Ausnahme mancher Gebietsränder, wo die Gleichung degeneriert. Im Black-Scholes-Fall ist etwa für $x_i, x_j \rightarrow 0$

$$\begin{aligned} a_{ij}(\mathbf{x}) &\sim x_i x_j, \\ b_i(\mathbf{x}) &\sim x_i. \end{aligned}$$

Wir werfen einen Blick auf die *Gitter-Peclet-Zahl*. Für die eindimensionale Black-Scholes-Gleichung errechnet man beispielsweise

$$\text{Pe} = \frac{hrx}{\frac{1}{2}\sigma^2 x^2} \leq 2 \frac{r}{\sigma^2},$$

wobei der kritische Wert im kleinsten Gitterpunkt bei $x = h$ angenommen wird. Für typische Parameter bei Aktienoptionen (etwa $r = 0.05$, $\sigma = 0.4$) ist $Pe < 2$, sodass auch zentrale Differenzen für die Konvektion stabil sind.

In den folgenden Testbeispielen wählen wir eine Diskretisierung der ersten und zweiten Ableitungen mittels zentraler Differenzenquotienten, was (für glatte Daten) ein Verfahren zweiter Ordnung ergibt. Finite Differenzen haben gegenüber finiten Elementen den Vorteil, dass (beim Fehlen von gemischten Ableitungen) die Sterne kompakter sind. Insbesondere in höheren Dimensionen d ist der Unterschied von 3^d zu $2d + 1$ Verbindungen im Matrix-Graph wesentlich. Schon bei fünfdimensionalen Problemen ist selbst für grobe Gitter ein 243-Punkte-Stern kritisch, während elf Nachbarn leicht zu handhaben sind. Eine denkbare Alternative in komplexeren Fällen mit dominantem Drift wären finite Volumen oder unstetige finite Elemente.

2.1.2 Zeitdiskretisierung

Ausgangspunkt ist die Semi-Diskretisierung

$$\frac{\partial \mathbf{u}_h}{\partial t} + \mathbf{A}_h \mathbf{u}_h = \mathbf{0},$$

wobei sich \mathbf{A}_h aus den Diskretisierungsmatrizen \mathbf{C}_h , \mathbf{D}_h , \mathbf{R}_h für Konvektion, Diffusion und Reaktion zusammensetzt, ergänzt durch entsprechende Randbedingungen. Für zeitabhängige Koeffizienten hängt \mathbf{A}_h offensichtlich von t ab.

θ -Schema

Als klassisches Beispiel verwenden wir ein Einschritt-Schema der Gestalt

$$\frac{\mathbf{u}_h^{m+1} - \mathbf{u}_h^m}{\delta t} + \theta \mathbf{A}_h \mathbf{u}_h^{m+1} + (1 - \theta) \mathbf{A}_h \mathbf{u}_h^m = \mathbf{b}_h$$

bzw.

$$(\mathbf{I} + \theta \delta t \mathbf{A}_h) \mathbf{u}_h^{m+1} = (\mathbf{I} - (1 - \theta) \delta t \mathbf{A}_h) \mathbf{u}_h^m + \delta t \mathbf{b}_h$$

bzw.

$$(\mathbf{I} + \theta \delta t \mathbf{A}_h) \mathbf{d}_h^m = \delta t \mathbf{A}_h \mathbf{u}_h^m + \delta t \mathbf{b}_h, \quad (2.1)$$

$$\mathbf{u}_h^{m+1} = \mathbf{u}_h^m + \mathbf{d}_h^m, \quad (2.2)$$

wobei hauptsächlich der unbedingt stabile Fall $\theta = 0.5$ (*Crank-Nicolson*) von zweiter Ordnung interessant ist. Randbedingungen müssen entsprechend eingearbeitet werden, im Fall zeitabhängiger Daten sind die Schritte in naheliegender Weise anzupassen.

Fractional-Step- θ

Beim Fractional-Step- θ -Verfahren (siehe etwa [Ran99]) handelt es sich um ein dreistufiges Schema der Gestalt

$$\begin{aligned} [\mathbf{I} + \alpha \theta \delta t \mathbf{A}_h^{m-1}] \mathbf{u}_h^{m-1+\theta} &= [\mathbf{I} - \beta \theta \delta t \mathbf{A}_h^{m-1}] \mathbf{u}_h^{m-1} + \theta \delta t \mathbf{b}_h^{m-1}, \\ [\mathbf{I} + \beta \theta' \delta t \mathbf{A}_h^{m-\theta}] \mathbf{u}_h^{m-\theta} &= [\mathbf{I} - \alpha \theta' \delta t \mathbf{A}_h^{m-1+\theta}] \mathbf{u}_h^{m-1+\theta} + \theta' \delta t \mathbf{b}_h^{m-1+\theta}, \\ [\mathbf{I} + \alpha \theta \delta t \mathbf{A}_h^m] \mathbf{u}_h^m &= [\mathbf{I} - \beta \theta \delta t \mathbf{A}_h^{m-\theta}] \mathbf{u}_h^{m-\theta} + \theta \delta t \mathbf{b}_h^{m-\theta}. \end{aligned}$$

2 Diskretisierung auf gradierten kartesischen Gittern

Bei der Wahl

$$\begin{aligned}\theta &= 1 - \sqrt{2}/2, \\ \theta' &= 1 - 2\theta = \sqrt{2} - 1\end{aligned}$$

ist das Verfahren von zweiter Ordnung und stark A-stabil für alle $\alpha \in (1/2, 1]$, $\beta = 1 - \alpha$. Eine ausgezeichnete Rolle spielt dabei der Wert

$$\alpha = \frac{1 - 2\theta}{1 - \theta},$$

da dann $\alpha\theta = \beta\theta'$ und die Systemmatrizen die gleiche Form haben. In diesem Fall lässt sich das Schema als drei hintereinander ausgeführte θ -Schritte mit unterschiedlicher Weite τ_i und unterschiedlichem θ_i schreiben, und zwar mit

$$\begin{aligned}\theta_1 &= \sqrt{2} - 1, & \tau_1 &= (\sqrt{2} - 1)\delta t, \\ \theta_2 &= \theta_3 = 2 - \sqrt{2}, & \tau_2 &= \tau_3 = (1 - \sqrt{2}/2)\delta t.\end{aligned}$$

Glattheit

Aufgrund der fehlenden Glattheit der Anfangsdaten ist die Lösung nicht gleichmäßig analytisch zum Anfang hin. Vielmehr ist die Lösung am Knick nicht einmal gleichmäßig stetig in der Zeitkoordinate (für $t = 0$, siehe 7.3.1). Schwab [vPS02] schlägt daher ein *Discontinuous-Galerkin-Verfahren* in der Zeit vor, bei dem das Zeitgitter zum Anfang hin exponentiell verfeinert und dann in jedem Zeitschritt die Ordnung erhöht wird.

In den später betrachteten numerischen Beispiel kann jedoch auch für die oben dargestellten Verfahren und sogar für BDF-Formeln hoher Ordnung kein Ordnungsverlust beobachtet werden.

2.1.3 Numerische Beispiele und Konvergenzdaten

Diskretisierungsfehler und Diskretisierungsordnung

Ausgegangen wird von einer Fehlerasymptotik in der Gitterweite h von der Form

$$u - u_h = c(h)h^p$$

mit Ordnung p und $c(h) \leq \bar{c}$, c stetig in 0, $c(0) \neq 0$. Dann gilt einerseits

$$\lim_{h \rightarrow 0} \frac{u_{2h} - u_h}{u_{2h} - u} = \lim_{h \rightarrow 0} \frac{c(2h)2^p - c(h)}{c(2h)2^p} = \frac{2^p - 1}{2^p}$$

und andererseits

$$\lim_{h \rightarrow 0} \frac{u_{4h} - u_{2h}}{u_{2h} - u_h} = \lim_{h \rightarrow 0} 2^p \frac{2^p c(4h) - c(2h)}{2^p c(2h) - c(h)} = 2^p,$$

also

$$\log_2 \left| \frac{u_{4h} - u_{2h}}{u_{2h} - u_h} \right| \rightarrow p \quad \text{für } h \rightarrow 0. \quad (2.3)$$

Für eine Folge von Gitterstufen n mit $h_n = 2^{-n}$ und dazugehörigen Gitterlösungen u_n betrachten wir daher als Schätzer für den absoluten Fehler

$$\delta_n := \frac{2^p}{2^p - 1} (u_{n+1} - u_n)$$

und für die Ordnung

$$o_n := \log_2 \left| \frac{u_n - u_{n-1}}{u_{n+1} - u_n} \right|. \quad (2.4)$$

FX-Option (3 Faktoren)

Ein interessantes Testbeispiel mit drei Faktoren stellt das Modell (1.18) dar. Der Einfachheit wegen wird $\beta_i = 0$ (*Hull-White-Modell*) und die Prozesse als unkorreliert angenommen ($\rho_{ij} = \delta_{ij}$), sodass sich die Gleichung (1.18) auf

$$\frac{\partial u}{\partial t} - \frac{1}{2} \sigma_1^2 x_1^2 \frac{\partial^2 u}{\partial x_1^2} - \frac{1}{2} \sum_{i=2,3} \sigma_i^2 \frac{\partial^2 u}{\partial x_i^2} - (x_2 - x_3) x_1 \frac{\partial u}{\partial x_1} - \sum_{i=2,3} \kappa_i (\theta_i - x_i) \frac{\partial u}{\partial x_i} + x_2 u = 0$$

reduziert.

Die Parameter seien $\sigma_i = 0.15$, $\kappa_i = 0.5$, $\theta_i = 0.045$, $K = 0.95$, $T = 10$ a, die heutigen Kurse $S = 0.9$, $r_d = r_f = 0.05$. Das Berechnungsgebiet ist $[0, 4K] \times [0, 0.1]^2$. Für $x_1 = 0$ gelten natürliche Randbedingungen, für $x_1 = 4K$ werden homogene Dirichletwerte gesetzt. Auf den anderen Rändern werden homogene Neumann-Daten angenommen.

Tabelle 2.1 zeigt die Konvergenzdaten der Finite-Differenzen-Diskretisierung für eine punktweise Auswertung am heutigen Kurs.

n	u_n	δ_n	o_n
0	0.076196	3.75 e-2	1.71
1	0.113716	1.32 e-2	2.08
2	0.125224	2.71 e-3	2.01
3	0.12794	6.76 e-4	2.00
4	0.128612	1.68 e-4	
5	0.12878		

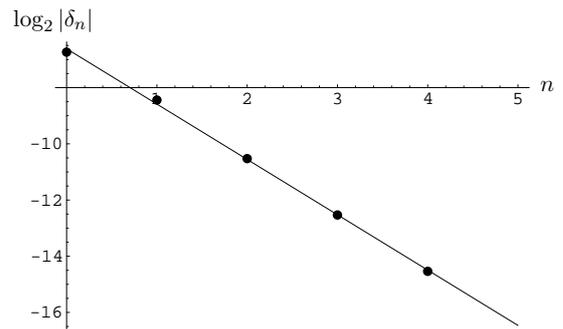


Tabelle 2.1: 3-Faktor-Modell für die Europäische FX-Option. Punktweise Konvergenz (ausgewertet am heutigen Kurs) der Lösung mit finiten Differenzen auf Kartesischen Gittern mit $(2^{n+2} + 1) \cdot (2^n + 1) \cdot (2^n + 1)$ Punkten (maximal 138 303) und 64 Zeitschritten des Fractional-Step-Schemas.

Black-Scholes-Basket (2 Faktoren)

Als weiteres Beispiel wird hier ein kleiner Basket aus zwei unkorrelierten Aktien gebracht (mehr oder weniger stark korrelierte Assets werden in nachfolgenden Kapiteln noch betrachtet und stellen keine weitere Schwierigkeit dar). Die verwendeten Daten sind $\sigma_1 = 0.5$, $\sigma_2 = 0.3$, $r = 0.05$, $T = 1$ a, $K = 0.25$, $\mu_1 = \mu_2 = 0.5$.

Die Black-Scholes-Gleichung wird in der transformierten Form aus 1.3.2 gerechnet, die im speziellen Fall so aussieht:

$$\begin{aligned} \frac{\partial u}{\partial t} - \sum_{i,j=1}^2 \alpha_{ij}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^2 \beta_i(\mathbf{x}) \frac{\partial u}{\partial x_i} + ru &= 0 & \forall (\mathbf{x}, t) \in]0, \infty[\times]0, 1[\times]0, T[, \\ u(\mathbf{x}, 0) &= (K - x_1)_+ & \forall \mathbf{x} \in]0, \infty[\times]0, 1[, \end{aligned}$$

2 Diskretisierung auf gradierten kartesischen Gittern

wobei

$$\begin{aligned}
 \alpha_{11}(\mathbf{x}) &= x_1^2 [\sigma_1^2 x_2^2 + \sigma_2^2 (1 - x_2)^2], \\
 \alpha_{12}(\mathbf{x}) = \alpha_{21}(\mathbf{x}) &= -x_1 x_2 (1 - x_2) [\sigma_2^2 (1 - x_2) - \sigma_1^2 x_2], \\
 \alpha_{22}(\mathbf{x}) &= x_2^2 (1 - x_2)^2 [\sigma_1^2 + \sigma_2^2], \\
 \beta_1 &= r x_1, \\
 \beta_2 &= x_2 (1 - x_2) [2(1 - x_2) \sigma_2^2 - 2\sigma_1^2 x_2].
 \end{aligned}$$

Für $x_1 = 0$, $x_2 \in \{0, 1\}$ gelten natürliche Randbedingungen, für $x_1 = 1$ wurde die Lösung auf $u = 0$ abgeschnitten.

Die Konvergenzdaten für eine Europäische und eine Amerikanische Option sind in den Tabellen 2.2 und 2.3 gesammelt. Die Lösungen sind in den Abbildungen 2.1 und 2.2 dargestellt.

n	0	1	2	3	4	5	6	7
$u_n \cdot 100$	2.12318	1.96816	2.72356	2.90175	2.94102	2.9506	2.95298	2.95357
$\delta_n \cdot 100$	-1.55 e-1	7.55 e-1	1.78 e-1	3.93 e-2	9.59 e-3	2.83 e-3	5.94 e-4	
o_n	-2.28	2.08	2.18	2.03	2.01	2.00		

Tabelle 2.2: Europäische Basket-Option mit 2 Faktoren: Konvergenz der Finite-Differenzen-Lösung am Kartesischen Gitter mit $(2^{n+2} + 1) \cdot (2^n + 1)$ Punkten (maximal 66 177) und 64 Zeitschritten des Fractional-Step-Schemas.

n	0	1	2	3	4	5	6	7
$u_n \cdot 100$	2.17917	2.47286	3.26258	3.43387	3.4715	3.48147	3.48398	3.48461
$\delta_n \cdot 100$	2.94 e-1	7.90 e-1	1.71 e-1	3.76 e-2	9.97 e-3	2.51 e-3	6.24 e-4	
o_n	-1.43	2.20	2.19	1.92	1.99	2.01		

Tabelle 2.3: Amerikanische Basket-Option mit 2 Faktoren: Konvergenz der Finite-Differenzen-Lösung am Kartesischen Gitter mit $(2^{n+2} + 1) \cdot (2^n + 1)$ Punkten (maximal 66 177) und 64 Zeitschritten des Fractional-Step-Schemas.

Im Fall der Amerikanischen Option wird die mit finiten Differenzen und Fractional-Step- θ diskretisierte Form von (1.14) bis (1.16), in Teilschritten

$$\begin{aligned}
 (\mathbf{I} + \theta \delta t \mathbf{A}_h) \mathbf{d}_h^m &\leq \delta t \mathbf{A}_h \mathbf{u}_h^m, \\
 \mathbf{d}_h^m &\geq \mathbf{c}_h - \mathbf{u}_h^m, \\
 ((\mathbf{I} + \theta \delta t \mathbf{A}_h) \mathbf{d}_h^m - \delta t \mathbf{A}_h \mathbf{u}_h^m) \star (\mathbf{d}_h^m - \mathbf{c}_h + \mathbf{u}_h^m) &= 0
 \end{aligned}$$

(siehe dazu (2.1) und (2.2) für den Fall linearer Gleichungen), mit den in Kapitel 3 beschriebenen Mehrgitterverfahren für Hindernisprobleme gelöst.

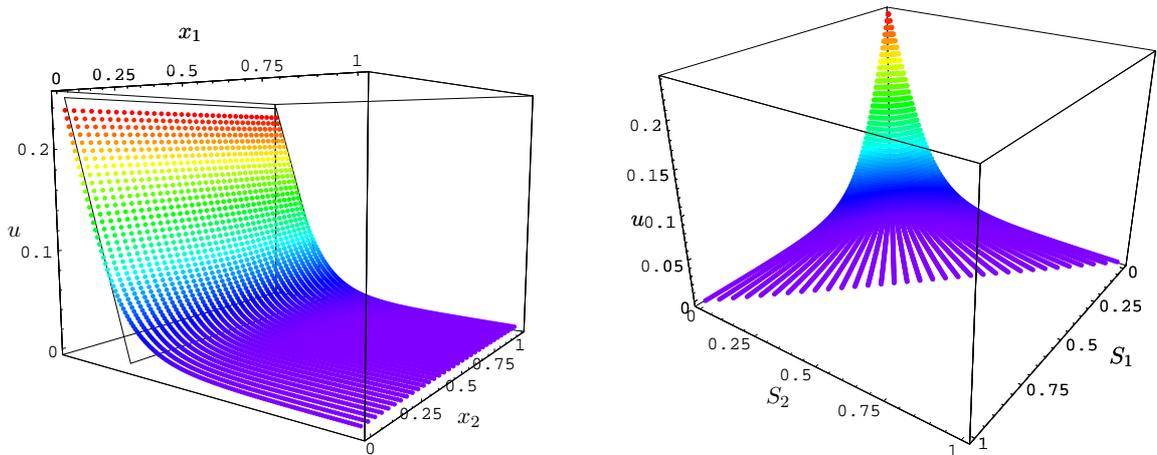


Abbildung 2.1: Preis u des Europäischen Put im transformierten x_1 - x_2 -System und in Abhängigkeit der Aktienkurse S_1, S_2 . Die hinter der Lösung sichtbare Anfangsbedingung hängt nur von x_1 ab und wird am Gitter exakt erfasst. Daher bleibt trotz des Knicks die zweite Ordnung der Diskretisierung erhalten.

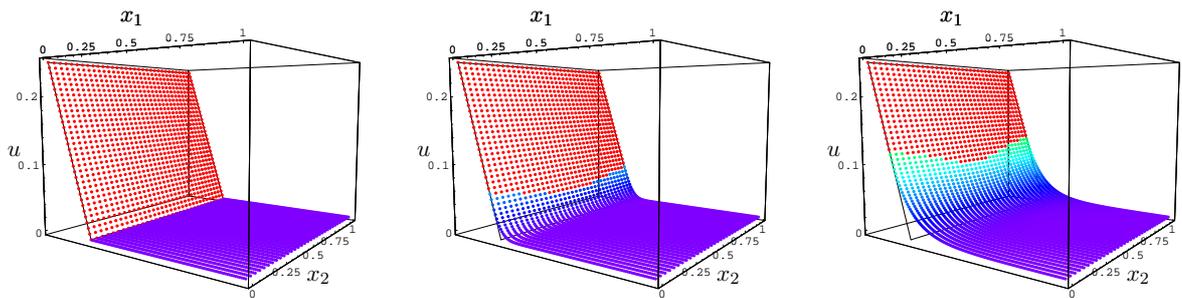


Abbildung 2.2: Preis u des Amerikanischen Put im transformierten x_1 - x_2 -System zu den Zeiten $t = 0, t = 0.1$ und $t = 1$ vor der Expiration. Die rote Zone stellt den Ausübungsbereich dar, der freie Rand ist eine glatte Funktion von x_2 .

2.2 Gradierte Gitter

Typisch für Optionspreisaufgaben sind stückweise glatte Anfangsbedingungen, die zwischen diesen Bereichen Knicke aufweisen. Eine lokale Verfeinerung entlang dieser Knicke wird sinnvoll sein, wobei aber die Erweiterbarkeit auf dünne Gitter (und da insbesondere die Kombinationstechnik) im Auge behalten werden muss. Adaptivität, wie sie bei dünnen finiten Elementen auf natürliche Weise möglich ist [Bun92, Bun98], kann bei der Kombinationstechnik nicht direkt eingesetzt werden, da die Fehlerentwicklung auf den einzelnen Gittern für die Extrapolation nicht gestört werden darf.

Daher werden die Gitter hier mit geeignet gewählten Funktionen in die einzelnen Richtungen gradiert, was sich leicht überträgt. Dies ist insbesondere dann sinnvoll, wenn die Fehleranteile in die einzelnen Koordinaten separieren, was gut zum Tensorprodukt-Ansatz der Kombinationstechnik und der dort angenommenen Fehlerdarstellung passt (siehe Abschnitt 4.2 oder z. B. auch [GT95, Gar98]).

Es wird zunächst der eindimensionale Fall dargestellt und die anderen Koordinaten unterdrückt, da die Gradierung ja nur eindimensional wirkt. Das Intervall sei $[0, 1]$.

Gradierungsfunktion

Die äquidistanten Gitterpunkte $x_i = i/2^n \in [0, 1]$, $0 \leq i \leq 2^n$, werden ersetzt durch $H(x_i)$ mit einer streng monotonen (hinreichend glatten) Gradierungsfunktion $H : [0, 1] \rightarrow [0, 1]$. $h(x) := H'(x)$ ist dann ein Maß für die Gitterweite im Punkt $H(x)$ (Abb. 2.3).

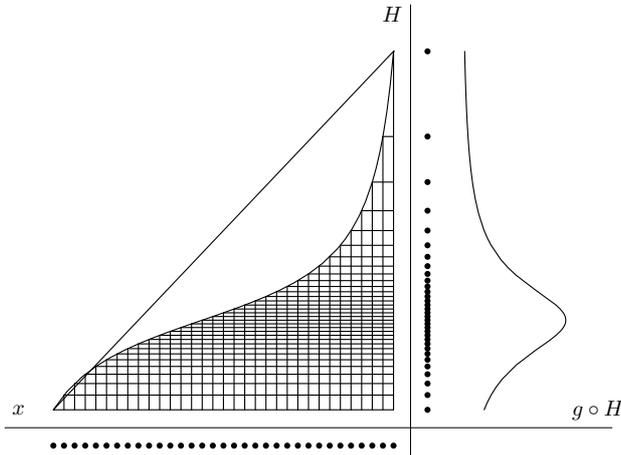


Abbildung 2.3: Gradierung eines äquidistanten Ausgangsgitters mit einer Funktion $H: h = H'$ gibt die lokale Gitterweite im transformierten Gitter an, $1/h = g \circ H$ mit $g = (H^{-1})'$ beschreibt die Dichte der Punkte.

Anstatt das Gitter zu transformieren, kann man auch eine Variablentransformation der Differentialgleichung mit der Inversen der Gradierungsfunktion $G := H^{-1}$ durchführen:

$$v(y) := u(H(y)) \Leftrightarrow u(x) = v(G(x)) \quad \forall x, y \in [0, 1].$$

Die Transformation der Differentialgleichung benutzt die Beziehungen ($G' := g$)

$$\begin{aligned} u'(x) &= v'(G(x))g(x), \\ u''(x) &= v''(G(x))g^2(x) + v'(G(x))g'(x). \end{aligned}$$

Die Gradierungsfunktion soll (neben technischen Anforderungen an Glattheit und Bijektivität) vor allem die Eigenschaft haben, dass das transformierte Gitter um den Knick in K am dichtesten ist und diese Gradierung über einen Parameter gesteuert werden kann.

Eine geeignete Funktion ist zum Beispiel

$$G_\lambda(x) := \frac{\arctan \lambda(x - K) + \arctan K\lambda}{\arctan \lambda(1 - K) + \arctan K\lambda} \tag{2.5}$$

(siehe Abbildung 2.3 für $\lambda = 7$), streng monoton für alle $0 \leq K \leq 1$, $\lambda > 0$ und bijektiv auf $[0, 1]$. Man sieht direkt, dass $G_\lambda(0) = 0$, $G_\lambda(1) = 1$. Außerdem hat G_λ einen Wendepunkt in K , sodass

$$g_\lambda(x) := G'_\lambda(x) = \frac{\lambda / (\arctan \lambda(1 - K) + \arctan K\lambda)}{1 + \lambda^2(x - K)^2}$$

in K ein Maximum hat. Der Parameter λ steuert die „Verdichtung“ des Gitters um K . Die Umkehrfunktion ist

$$H_\lambda(y) := K + \frac{1}{\lambda} \tan ((\arctan \lambda(1 - K) + \arctan K\lambda) y - \arctan K\lambda).$$

Aufgrund der Glättungseigenschaft parabolischer Gleichungen ist zu erwarten, dass eine starke Gradierung nur zu Beginn sinnvoll ist. Die oben vorgeschlagene Gradierungsfunktion kann über die Zeitabhängigkeit des Parameters λ gesteuert werden,

$$G(x, t) := G_{\lambda(t)}(x),$$

wobei Funktionen der Art

$$\lambda(t) := \lambda_0 \left(\frac{\tau}{\tau + t} \right)^\alpha \quad (2.6)$$

mit $\lambda_0, \tau, \alpha > 0$ die folgende gewünscht Eigenschaft haben: $\lambda(0) = \lambda_0$ gibt die maximale Gradierung an, die Parameter τ und α steuern das Abklingverhalten.

Beispiel

Als Beispiel dient wieder der Zwei-Faktor-Basket von oben. Hier zeigt sich ein weiterer Nutzen der Transformation aus 1.3.2: da der Knick orthogonal zu x_1 ist, kann die Gradierung ausschließlich in diese Richtung erfolgen.

Um den Effekt zu veranschaulichen, wird zuerst ein reduziertes eindimensionales Modell mit gemittelter *Basket-Volatilität* $\sigma = 0.4$ gerechnet (siehe dazu 7.2.1, (7.2)), für das man durch Probieren leicht günstige Parameter für die Gradierung findet.

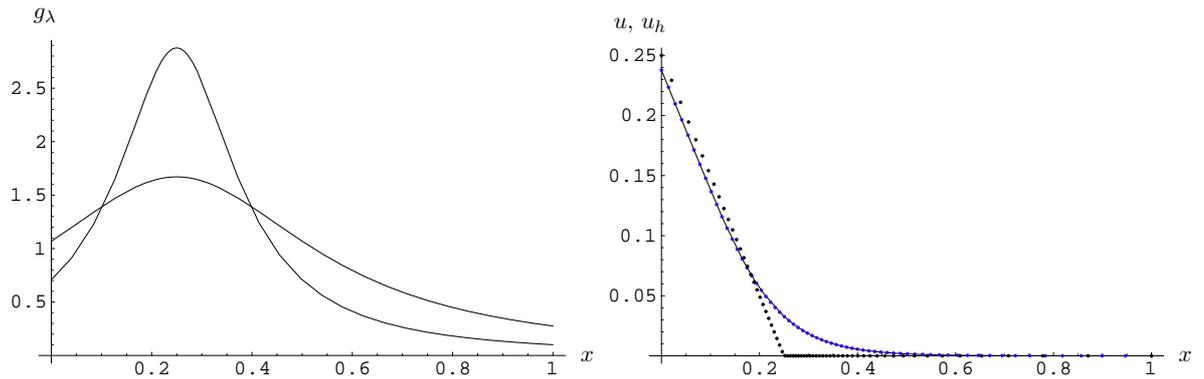


Abbildung 2.4: Links: Punktichte g_λ für $\lambda = 7$ (scharfes Maximum) und $\lambda = 3$. Rechts: Eindimensionale analytische Lösung $u(x, 1)$ zusammen mit der numerischen Lösung $u_h(x_i, 1)$ und Anfangsbedingung $u_h(x_i, 0)$, $0 \leq i \leq 64$, auf einem bewegten Gitter mit der Gradierungsfunktion nach (2.5) und (2.6) mit $\lambda(0) = 7$, $\lambda(1) = 3$ und dem Parameter $\alpha = 1$.

Diese werden dann für das zweidimensionale Problem verwendet. Abbildung 2.5 zeigt, dass durch geschickte zeitabhängige Gradierung des Gitters im zweidimensionalen Modell bei gleicher Genauigkeit ein bis zwei Verfeinerungsstufen eingespart werden können. Die dazugehörige Lösung ist in Abbildung 2.6 zusammen mit dem Gitter zu sehen.

2 Diskretisierung auf gradierten kartesischen Gittern

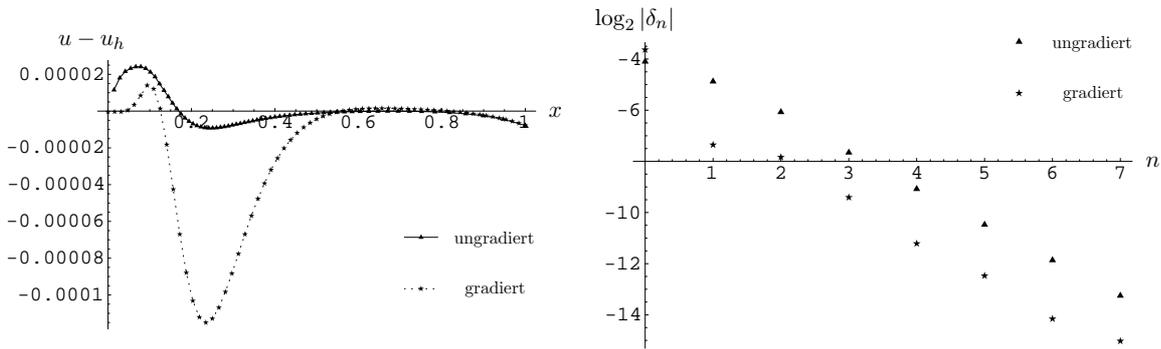


Abbildung 2.5: Links: Der Fehler der eindimensionalen numerischen Lösung u_h (im Vergleich zur analytischen Black-Scholes-Lösung u) für ein ungradiertes und ein bewegtes Gitter mit der Gradierungsfunktion nach (2.5) und (2.6) mit $\lambda(0) = 7$, $\lambda(1) = 3$ und dem Parameter $\alpha = 1$. Rechts: Der Fehler im zweidimensionalen Modell auf den derart gradierten Gittern im Vergleich zu den ungradierten mit der gleichen Anzahl von $(2^{n+2} + 1) \cdot (2^n + 1)$ Punkten.

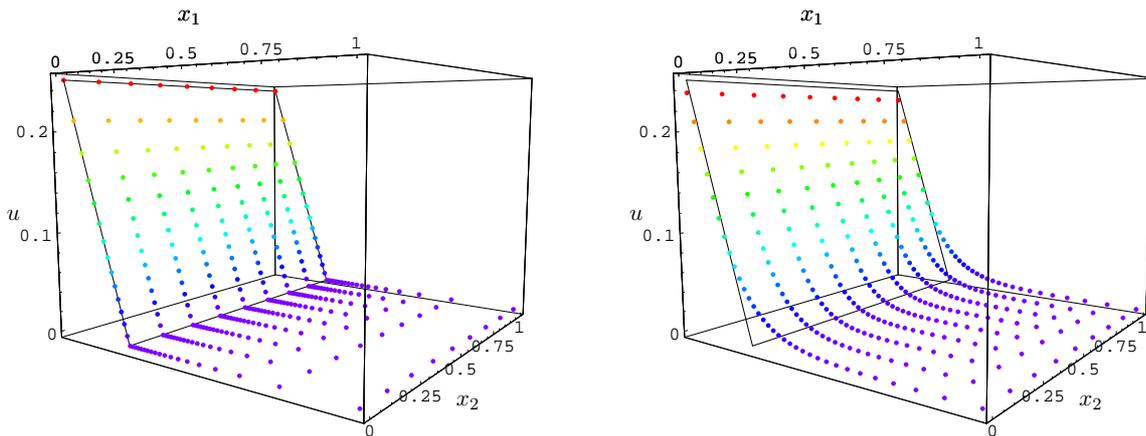


Abbildung 2.6: Preis u des Europäischen Put im transformierten x_1 - x_2 -System zu den Zeiten $t = 0$ und $t = 1$ vor der Expiration. Die Gradierung in x_1 -Richtung nach (2.5) mit $\lambda(t)$ aus (2.6) nimmt mit der Zeit von $\lambda(0) = 7$ auf $\lambda(1) = 3$ ab.

2.3 Extrapolation und Sensitivitäten

Über den direkten Nutzen verbesserter Lösungsgenauigkeit hinaus ist eine höhere Konvergenzordnung der Diskretisierung wichtig für die numerische Differentiation zur Bestimmung der praktisch wichtigen Sensitivitäten Δ und Γ . Mit geringem Aufwand kann höhere Ordnung durch Extrapolation gewonnen werden, die Ableitung erhält man wie üblich durch Differentiation geeigneter interpolierender Polynome.

2.3.1 Richardson-Extrapolation

Ausgangspunkt ist eine Fehlerasymptotik von der Form

$$u_h(\mathbf{x}, t) - u(\mathbf{x}, t) = c_2(\mathbf{x}, t)h^2 + c_4(h, \mathbf{x}, t)h^4$$

mit $|c_4(h, \mathbf{x}, t)| \leq \bar{c}$, c_4 stetig und $c_4(0, \mathbf{x}, t) \neq 0$. Für die extrapolierte Lösung

$$\hat{u}(\mathbf{x}, t) := \frac{4}{3}u_{h/2}(\mathbf{x}, t) - \frac{1}{3}u_h(\mathbf{x}, t)$$

ergibt sich bekanntlich

$$\hat{u}_h(\mathbf{x}, t) - u(\mathbf{x}, t) = \hat{c}(\mathbf{x}, t)h^4 + o(h^4),$$

also vierte Ordnung.

Beispiel

Für den oben betrachteten Basket mit zwei Komponenten ist die (vierte) Konvergenzordnung sehr gut in Tabelle 2.4 zu sehen.

n	0	1	2	3	4	5	6
$u_n \cdot 100$	1.91649	2.97537	2.96115	2.9541	2.95379	2.95377	2.95377
$\delta_n \cdot 100$	1.06	-1.42 e-2	-7.05 e-3	-3.11 e-4	-1.86 e-5	-1.15 e-6	
o_n	6.22	1.01	4.50	4.06	4.01		

Tabelle 2.4: Europäische Basket-Option mit 2 Faktoren: Konvergenz der Richardson-Extrapolation am Kartesischen Gitter mit $(2^{n+2} + 1) \cdot (2^n + 1)$ Punkten. Alle Daten wie oben.

2.3.2 Numerische Differentiation

Da durch die Lösung mittels finiter Differenzen zunächst nur eine Skelettlösung (auf den Gitterpunkten) gegeben ist, muss sie mit entsprechender Ordnung q interpoliert werden. Die verwendeten Lagrange-Polynome sind für den eindimensionalen Fall für $q = 3$ in Abbildung 2.7 dargestellt.

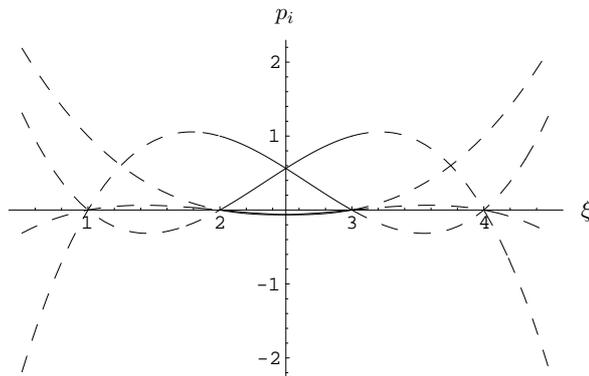


Abbildung 2.7: Die Interpolation der Lösung im Intervall $[2, 3]$ geschieht durch die Lagrange-Basispolynome mit Stützstellen $\{1, 2, 3, 4\}$. An den Gebietsrändern werden die Punkte entsprechend ins Innere verschoben bzw. — falls dies nicht möglich ist — die Ordnung reduziert.

Im d -dimensionalen Fall konstruiert man die Lösung aus Produkten der eindimensionalen Funktionen. Seien $\mathbf{i} = (i_1, \dots, i_d) \in \times_{k=1}^d \{1, \dots, n_k\}$ die Indizes der Nachbarn und p_{k,i_k} die entsprechenden Basisfunktionen in Richtung k vom Grad n_k , die am i_k -ten Punkt in dieser Richtung nicht verschwinden.

2 Diskretisierung auf gradierten kartesischen Gittern

Dann hat eine Interpolation über die entsprechende Tensorproduktbasis die Form

$$\hat{u}(\mathbf{x}) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \prod_{k=1}^d p_{k,i_k}(\xi_k) \cdot u_{\mathbf{i}} = \sum_{i_1=1}^{n_1} p_{1,i_1}(\xi_1) \cdots \sum_{i_d=1}^{n_d} p_{d,i_d}(\xi_d) u_{\mathbf{i}}.$$

Die Mehrfachsummen können rekursiv definiert und ausgewertet werden:

$$\begin{aligned} u_{\mathbf{i},k} &:= u_{\mathbf{i}}, \\ u_{\mathbf{i},k-1} &:= \sum_{i_k=1}^{n_k} p_{k,i_k}(\xi_k) u_{\mathbf{i},k} \quad d \geq k \geq 1. \end{aligned}$$

Der interpolierte Wert ist dann $u_{\mathbf{i},0}$. Die Ableitungen werden in naheliegender Weise durch Differentiation der Basispolynome gewonnen:

$$\frac{\partial u_h}{\partial x_j} = \sum_{i_1, \dots, i_d} p'_{j,i_j}(\xi_j) \prod_{k \neq j} p_{k,i_k}(\xi_k) \cdot u_{\mathbf{i}}, \quad \frac{\partial^2 u_h}{\partial x_j^2} = \sum_{i_1, \dots, i_d} p''_{j,i_j}(\xi_j) \prod_{k \neq j} p_{k,i_k}(\xi_k) \cdot u_{\mathbf{i}} \quad \text{etc.}$$

Beispiel

Für die durch Richardson-Extrapolation gewonnene numerische Lösung aus 2.3.1 wird in Abb. 2.8 und Tab. 2.5 die Konvergenz der numerischen Ableitungen angegeben.

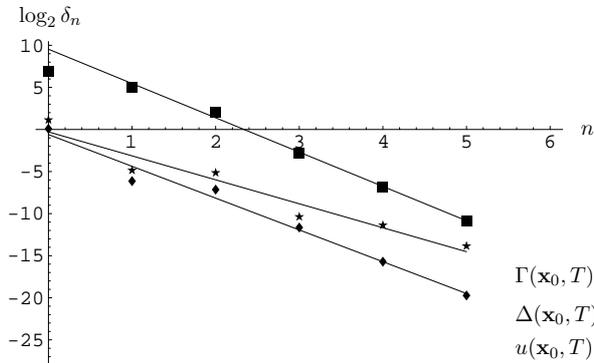


Abbildung 2.8: Europäische Basket-Option mit 2 Faktoren: Punktweise Konvergenz im aktuellen Kurs \mathbf{x}_0 der Richardson-extrapolierten Finite-Differenzen-Lösung und der ersten beiden Ableitungen (Δ und Γ) nach x_1 . Die durch Regression geschätzten Konvergenzordnungen sind 3.8, 2.9 und 4.1. Für die zweite Ableitung stellt man also Superkonvergenz fest.

n	0	1	2	3	4	5	6
$u'_n \cdot 100$	-35.4917	-37.6837	-37.6491	-37.6774	-37.6767	-37.6763	-37.6762
$u''_n \cdot 100$	567.676	445.926	412.776	416.816	416.964	416.973	416.973
$\delta'_n \cdot 100$	-219	3.46	-2.84	7.55 e-2	3.78 e-2	6.81 e-3	
$\delta''_n \cdot 100$	-1.22 e+4	-3.32 e+3	4.04 e+2	1.48 e+1	8.46 e-1	5.30 e-2	
o'_n	5.98	0.29	5.23	1.00	2.47		
o''_n	1.88	3.04	4.77	4.13	4.00		

Tabelle 2.5: Europäische Basket-Option mit 2 Faktoren: Konvergenz der ersten (u'_n) und zweiten (u''_n) numerischen Ableitung der Richardson-extrapolierten Lösung nach x_1 .

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

Überblick

Zunächst werden in Abschnitt 3.1 Grundkonzepte der iterativen Lösung großer schwach besetzter Systeme eingeführt und die speziellen Anforderungen für die betrachtete Problemklasse aufgezeigt. In 3.2 werden sodann Glätter für einen Mehrgitteralgorithmus analysiert, der im Hinblick auf anisotrope Gitter und Entartung der Gleichung robust ist, wobei vor allem Linien- und Ebenenglätter sowie ihre Entsprechungen in höheren Dimensionen eine entscheidende Rolle spielen. Die gewonnenen Erkenntnisse werden in 3.3 auf lineare Komplementaritätsprobleme übertragen, die zusätzlich einen angepassten Gittertransfer am freien Rand und eine gute Anfangsapproximation desselben durch eine hierarchische Näherung verlangen (volles, kaskadisches Mehrgitter).

3.1 Multigrid

Dieses Kapitel widmet sich der schnellen Lösung von linearen Gleichungen

$$\mathbf{A}_h \mathbf{u}_h = \mathbf{b}_h$$

($\mathbf{u}_h, \mathbf{b}_h \in \mathbb{R}^N$, $\mathbf{A}_h \in \mathbb{R}_N^N$) und Komplementaritätsproblemen (mit einem Hindernis $\mathbf{c}_h \in \mathbb{R}^N$)

$$\mathbf{A}_h \mathbf{u}_h \leq \mathbf{b}_h, \quad (3.1)$$

$$\mathbf{u}_h \geq \mathbf{c}_h, \quad (3.2)$$

$$(\mathbf{A}_h \mathbf{u}_h - \mathbf{b}_h) \star (\mathbf{u}_h - \mathbf{c}_h) = \mathbf{0}, \quad (3.3)$$

wie sie nach einer impliziten Diskretisierung der entsprechenden parabolischen (Un-)Gleichungen entstehen. Hier bezeichnet \star die komponentenweise Multiplikation.

In Vorbereitung auf die hier betrachteten Dünngitter-Diskretisierungen interessiert speziell der Fall kartesischer Gitter mit $n_k = 2^{l_k} + 1$ ($l_k \in \mathbb{N}_0$) Punkten (Knotenfreiheitsgraden) in den einzelnen Raumrichtungen $k = 1, \dots, d$, wobei die Gitterweiten $h_k = 2^{-l_k}$ stark unterschiedlich sein können. Die Gesamtzahl der Unbekannten ist demnach $N = n_1 \cdot \dots \cdot n_d$.

3.1.1 Iterative Verfahren

Betrachtet werden Verfahren der Gestalt

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{N} \left(\mathbf{b} - \mathbf{A} \mathbf{u}^k \right)$$

(2. Normalform), $k \in \mathbb{N}_0$, wobei der Startwert \mathbf{u}^0 beliebig gewählt sei. Dies schließt formal auch den Fall eines Hindernisproblems ein, wobei dann $\mathbf{N} = \mathbf{N}(\mathbf{u}^k, \mathbf{c})$ ein nichtlinearer Operator ist. Im linearen Fall kann man die *Konvergenzrate* einführen als

$$\rho := \varrho(\mathbf{M}),$$

wobei $\mathbf{M} := \mathbf{I} - \mathbf{A} \mathbf{N}$ und ϱ der *Spektralradius* ist. Für $\rho < 1$ ist das Verfahren konvergent.

Für die experimentelle Beurteilung des Konvergenzverhaltens eignen sich Raten der Gestalt

$$\rho_k := \frac{\|\mathbf{r}^k\|_2}{\|\mathbf{r}^{k-1}\|_2}$$

($k > 0$) mit den Residuen $\mathbf{r}^k := \mathbf{b} - \mathbf{A}\mathbf{u}^k$.

Für lineare Komplementaritätsprobleme lassen sich dementsprechend projizierte Residuen \mathbf{r}^+ gemäß

$$r_j^+ := \begin{cases} r_j & u_j > c_j, \\ \min(r_j, 0) & u_j = c_j, \end{cases}$$

betrachten, wobei $\mathbf{u} \geq \mathbf{c}$ vorausgesetzt wird. Die Indizes, für die der erste Fall zutrifft, bilden die *inaktive Menge*, letztere die *aktive Menge*. Wir messen also den Fehler am Residuum der Gleichung (inaktive Menge) bzw. an der Verletztheit der ersten Ungleichung (aktive Menge). Es wird davon ausgegangen, dass die Iterierten (durch geeignete Projektion) immer in der zulässigen Menge $\{\mathbf{u} \geq \mathbf{c}\}$ liegen. Als dazugehörige Rate ergibt sich

$$\rho_k = \frac{\|\mathbf{r}^{+,k}\|_2}{\|\mathbf{r}^{+,k-1}\|_2}.$$

Der entscheidende Nachteil klassischer Iterationsverfahren ist die Abhängigkeit der Konvergenzrate von der *Konditionszahl*

$$\kappa(\mathbf{A}) := \varrho(\mathbf{A})\varrho(\mathbf{A}^{-1})$$

der Systemmatrix \mathbf{A} , die im Zusammenhang mit Diskretisierungen elliptischer und parabolischer Gleichungen typischerweise gemäß

$$\kappa(\mathbf{A}_h) = \mathcal{O}(h^{-2}) \tag{3.4}$$

wächst, wobei h die Gitterweite ist.

Bemerkung. Die Ordnung in (3.4) ist bestimmt durch die (zweite) Ordnung des Differentialoperators, unabhängig von der Dimension d . Da aber die Anzahl der Unbekannten mit $N \sim h^{-d}$ ansteigt, wächst die Kondition in höheren Dimensionen sehr viel schwächer mit N .

Diese Aussage gilt jedoch **nicht** für anisotrope Gitter, wo die kleinste der Gitterweiten $h = \min_{1 \leq k \leq d} h_k$ in den einzelnen Richtungen entscheidend ist. Dieser im Hinblick auf dünne Gitter wesentliche Fall macht eine Form der (Multilevel-)Präkonditionierung unabdingbar.

Einfache Iterationen

Eine systematische Übersicht und Analyse von iterativen Verfahren findet man bei [Hac93]. Als später verwendetes Beispiel sei hier nur die *Gauß-Seidel-Iteration*

$$u_i^{k+1} = \frac{1}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij} u_j^{k+1} - \sum_{j > i} a_{ij} u_j^k \right), \quad i = 1, \dots, N, \tag{3.5}$$

bzw. die etwas allgemeinere *SOR-Iteration* (*Successive Overrelaxation*)

$$u_i^{k+1} = (1 - \omega) u_i^k + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij} u_j^{k+1} - \sum_{j > i} a_{ij} u_j^k \right) \tag{3.6}$$

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

mit einem Relaxationsparameter $\omega \in [1, 2)$ angeführt.

Eine naheliegende Erweiterung auf Hindernisprobleme, bei der in jedem Gauß-Seidel-Schritt die Lösung auf die zulässige Menge projiziert wird, wurde erstmals von Cryer [Cry71] betrachtet:

$$u_i^{k+\frac{1}{2}} = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij} u_j^{k+1} - \sum_{j>i} a_{ij} u_j^k \right), \quad (3.7)$$

$$u_i^{k+1} = \max \left(u_i^{k+\frac{1}{2}}, c_i \right). \quad (3.8)$$

Leider sind die Raten von einfachen Iterationen dieses Typs, etwa des Gauß-Seidel-Verfahrens, von der asymptotischen Form

$$\rho = 1 - \mathcal{O}(\kappa^{-1}) = 1 - \mathcal{O}(h^2).$$

Gradientenverfahren

Eine wesentliche Beschleunigung stellen CG-Verfahren dar, die typischerweise Konvergenzraten von der Form

$$\rho = 1 - \mathcal{O}(\kappa^{-\frac{1}{2}}) = 1 - \mathcal{O}(h)$$

aufweisen und idealerweise nach endlich vielen Schritten in der exakten Lösung landen, was aber praktisch selten relevant ist.

Stabile Erweiterungen von der klassischen Variante für symmetrische Probleme auf den unsymmetrischen Fall stellt unter anderem das Verfahren Bi-CGSTAB [vdV92] dar, welches im Folgenden zur Lösung von hinreichend groben (gut konditionierten) Systemen verwendet wird.

Eine Erweiterung für restringierte Aufgabenstellungen wurde für den symmetrischen Fall sowohl von Braess et al. [HBS03] als auch von Dostál und Schöberl [DS02] angegeben. Methoden für den unsymmetrischen Fall sind bis dato noch ausständig.

Wir orientieren uns an der von Kőcvara et al. [KZ94] für den symmetrischen Fall vorgeschlagenen hybriden Iteration SSOR-PCG, in der jeweils m projizierte SSOR-Schritte (zur Bestimmung des freien Randes) mit s (vorkonditionierten) CG-Schritten am Residuum der inaktiven Gleichungen (zur Beschleunigung der Konvergenz) abwechseln. Seien dazu

$$\begin{aligned} \mathcal{I}(\mathbf{u}) &:= \{i : u_i = c_i\} \text{ bzw.} \\ \mathcal{A}(\mathbf{u}) &:= \{1, \dots, N\} \setminus \mathcal{I}(\mathbf{u}) \end{aligned}$$

die inaktive bzw. aktive Menge und $P_{\mathcal{A}(\mathbf{u})} : \mathbb{R}^N \rightarrow \mathbb{R}^{|\mathcal{A}(\mathbf{u})|}$ die Projektion auf $\mathcal{A}(\mathbf{u})$. Das Inkrement aus der Krylov-Raum-Iteration wird dabei so gedämpft, dass die Korrektur nicht aus der zulässigen Menge fällt. Aus der Konvergenz der projizierten Gauß-Seidel-Iteration [Cry71] und der Abstiegseigenschaft des CG-Verfahrens im Bezug auf das zugeordnete restringierte Minimierungsproblem kann die Konvergenz des Zwei-Schritt-Verfahrens bewiesen werden [KZ94].

Eine Erweiterung auf den nicht-symmetrischen Fall liegt auf der Hand:

Algorithmus 3.1. SORP-Bi-CGSTAB

```

procedure SORP-Bi-CGSTAB(A, b, c, u)
begin
  while  $err \geq \epsilon$  do begin
    Schritt 1:
    for  $i = 1, \dots, m$  do begin
       $\mathbf{u} \leftarrow \text{SORP}(\mathbf{A}, \mathbf{b}, \mathbf{c}, \mathbf{u});$ 
    end
    Schritt 2:
     $\mathbf{r} = \mathbf{b} - \mathbf{A} \mathbf{u};$ 
    while  $err_2 > \epsilon_2$  do begin
       $\mathbf{d} \leftarrow \text{Bi-CGSTAB}(\mathbf{P}_{\mathcal{A}} \mathbf{A} \mathbf{P}_{\mathcal{A}}^T, \mathbf{P}_{\mathcal{A}} \mathbf{r}, \mathbf{d});$ 
    end
     $\mathbf{u} \leftarrow \mathbf{u} + \gamma \mathbf{P}_{\mathcal{A}}^T \mathbf{d}$       ( $\gamma \leq 1, \mathbf{u} + \gamma \mathbf{P}_{\mathcal{A}}^T \mathbf{d} \geq \mathbf{c}$ )
  end
return SORP-Bi-CGSTAB

```

3.1.2 Der Basis-Mehrgitter-Algorithmus

Eine klassische Mehrgitter-Einführung stammt von Hackbusch [Hac93]. Hier seien nur kurz die wesentlichen Komponenten zur späteren Referenz zusammengestellt. Ausgangspunkt ist die Defekt-Korrektur-Form: Mit dem Residuum

$$\mathbf{r}_h = \mathbf{b}_h - \mathbf{A}_h \mathbf{u}_h,$$

gilt für die Korrektur \mathbf{d}_h aus

$$\mathbf{A}_h (\mathbf{u}_h + \mathbf{d}_h) = \mathbf{b}_h$$

die Gleichung

$$\mathbf{A}_h \mathbf{d}_h = \mathbf{r}_h.$$

Da das Residuum nach ein paar Iterationsschritten — etwa eines Gauß-Seidel-Verfahrens — glatt ist, können Residuum und Korrektur auf einem größeren Gitter dargestellt werden. Es werden daher zunächst zwei Gitterebenen betrachtet, welche durch die Transferoperatoren Prolongation \mathbf{I}_h^H (z. B. multilineare Interpolation) und Restriktion (z. B. punktweise — *Injektion* — oder besser Transponierte der Prolongation) aufeinander abgebildet werden.

In einem Zweigitter-Schritt wird nach Anwendung des Glätters

$$\mathbf{u}_h \leftarrow \mathcal{S}(\mathbf{A}_h, \mathbf{b}_h, \mathbf{u}_h)$$

das (glatte) Residuum restringiert,

$$\mathbf{r}_H = \mathbf{I}_H^h \mathbf{r}_h,$$

und näherungsweise die Grobgitterkorrektur als Lösung von

$$\mathbf{A}_H \mathbf{d}_H = \mathbf{r}_H$$

berechnet, und zwar mit einer geeigneten Grobgittermatrix \mathbf{A}_H , z. B. nach dem Galerkinansatz

$$\mathbf{A}_H = \mathbf{I}_H^h \mathbf{A}_h \mathbf{I}_h^H.$$

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

Schließlich wird die Korrektur prolongiert,

$$\mathbf{d}_h = \mathbf{I}_h^H \mathbf{r}_H,$$

und zur alten Iterierten addiert,

$$\mathbf{u}_h \leftarrow \mathbf{u}_h + \mathbf{d}_h.$$

Durch Nachweis der *Glättungs-* und der *Approximationseigenschaft* kann (für bestimmte Probleme) Konvergenz mit h -unabhängiger Rate nachgewiesen werden, was die lineare Komplexität des Lösungsverfahrens für dünn besetzte Systeme garantiert. Eine rekursive Anwendung der Schritte auf die Grobgitterlösung mit direktem Löser auf der untersten Stufe liefert einen Mehrgitterschritt. Als Stufenindex wird daher l statt h , H verwendet.

Algorithmus 3.2. V-Zyklus

```

procedure V( $l$ ,  $\mathbf{A}^l$ ,  $\mathbf{b}^l$ ,  $\mathbf{u}^l$ )
begin
  if  $l = 1$  then
     $\mathbf{u}^1 \leftarrow (\mathbf{A}^1)^{-1} \mathbf{b}^1$ ;
  else begin
     $\mathbf{u}^l \leftarrow \mathcal{S}(\mathbf{A}^l, \mathbf{b}^l, \mathbf{u}^l)$ ; (Vorglättung)
     $\mathbf{r}^l = \mathbf{b}^l - \mathbf{A}\mathbf{u}^l$ ;
     $\mathbf{r}^{l-1} = I_l^{l-1} \mathbf{r}^l$ ;
     $\mathbf{d}^{l-1} \leftarrow \mathbf{V}(l-1, \mathbf{A}^l, \mathbf{r}^{l-1})$ ;
     $\mathbf{d}^l = I_{l-1}^l \mathbf{d}^{l-1}$ ;
     $\mathbf{u}^l = \mathbf{u}^l + \mathbf{d}^l$ ;
     $\mathbf{u}^l \leftarrow \mathcal{S}(\mathbf{A}^l, \mathbf{b}^l, \mathbf{u}^l)$ ; (Nachglättung)
  end
return V-Zyklus

```

Die Kunst besteht im komplexen Anwendungsfall in der problemangepassten (besser noch: robusten, d. h. über eine weite Klasse von Problemen geeigneten) Wahl der Komponenten: des Glätters, der Gitterhierarchie, der Transferoperatoren und des Grobgitterlösers.

3.1.3 Spezielle Anforderungen

Zweierlei Eigenheiten der betrachteten Modelle und Diskretisierungen verlangen eine spezielle Berücksichtigung im Mehrgitteralgorithmus: die Glättung bei anisotroper Diffusion sowie der Gittertransfer bei freien Rändern.

Anisotrope Gleichungen

Die Anisotropien stammen offensichtlich von drei Faktoren:

1. globale Anisotropien durch gestreckte Gitter (dünne Gitter),
2. globale Anisotropien durch unterschiedlich starke Diffusion in die einzelnen Richtungen (etwa durch verschiedene Volatilitäten in (1.6) oder den Abfall im Spektrum in (1.21)),
3. lokale Anisotropien am Rand (wie in (1.10) zusammen mit (1.11),(1.12)).

Klarerweise sind 1. und 2. auf diskreter Ebene weitgehend gleichbedeutend. Diese Problemklasse ist in der Literatur sehr ausführlich abgehandelt. Beispielsweise findet sich in [LM00] eine systematische und ausführliche numerische Analyse für verschiedene Punkt-, Linien- und Ebenenglätter am Beispiel der dreidimensionalen Poisson-Gleichung oder bei [Oos97] die Anwendung von Ebenenglätttern für die dreidimensionalen anisotropen Navier-Stokes-Gleichungen.

Daher liegt das Augenmerk hier auf dem letzten Aspekt.

Lineare Komplementaritätsprobleme

Wenn an die Stelle der Gleichung ein Komplementaritätsproblem tritt, muss dem Gittertransfer zusätzliche Beachtung geschenkt werden, da (vor allem nahe der Lösung) durch die Grobgitterkorrektur die Komplementaritätsbedingung nicht verletzt werden darf (das Verfahren endet sonst typischerweise in einem Zyklus).

Die wohl einfachste Alternative ist punktweise Korrektur vom Grobgitter (Injektion), wie man es in der illustrativen Arbeit von Brandt und Cryer [BC83] findet, die (wie das Gros der nachfolgenden Arbeiten) projizierte SOR-Schritte als Glätter verwenden und für das Beispiel des porösen Damms gute Raten erzielen. Für variierende Koeffizienten ist die Konvergenz allerdings unzufriedenstellend.

Hoppe, Kornhuber [HK87] und Wittum [KSW90] schränken die Transferoperatoren durch einseitige Extrapolation am freien Rand auf die aktive bzw. inaktive Menge ein, sodass unter Beibehaltung der Ordnung die Grobgitterkorrektur an der Aufsatzmenge ausgeschaltet wird.

Ganz umgangen wird diese Problematik durch eine von Braess et al. in [HBS03] vorgeschlagene kaskadische Technik, die ohne Restriktion auskommt. Schwierigkeiten treten selbst bei einfachen Problemen allerdings dann auf, wenn man die Lösung sehr genau — unterhalb des Diskretisierungsfehlers — wissen möchte, wo die Mehrgitterkonvergenz verloren geht. Für Extrapolations- und Dünngitterverfahren ist das ein wesentlicher Punkt.

Vorarbeiten bei Optionspreisaufgaben

Köcvara et al. [SMK98] wenden auf (univariate) Amerikanische Optionen das kombinierte Verfahren aus [KZ94] an, siehe 3.1.1. Die CG-Beschleunigung der projizierten Gauß-Seidel-Iteration funktioniert gut für hinreichend wohlkonditionierte Systeme, sodass es in der für unsymmetrische Systeme modifizierten Form (Algorithmus 3.1) einen passablen Grobgitterlöser abgibt.

Clarke und Parrott [CP99] haben wohl als erste Mehrgitter für Amerikanische Optionspreisaufgaben angewandt. Sie kombinieren für ein zweidimensionales Modell (eine Aktie mit stochastischer Volatilität) den globalen projizierten Punkt-Gauß-Seidel mit Linienglättung auf der inaktiven Menge.

Oosterlee [Oos03] untersucht das gleiche Problem und beschleunigt die Methode von Brandt und Cryer für dieses Anwendungsbeispiel durch Iteranden-Rekombination.

3.2 Multigrid für anisotrope Probleme

Es ist wohlbekannt — und wird auch durch die folgenden numerischen Tests illustriert — dass das Mehrgitterverfahren aus 3.1 mit Standardkomponenten nicht robust im Bezug auf Anisotropien ist. Verschiedene Zugänge sind denkbar, diesem Problem zu begegnen:

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

1. *Semi-Coarsening*, d. h. Vergrößerung nur in Richtung der Anisotropie,
2. *Blockglätter*, d. h. Blöcke mit schwacher Kopplung werden gemeinsam invertiert, sowie
3. *Matrix-abhängiger Gittertransfer*.

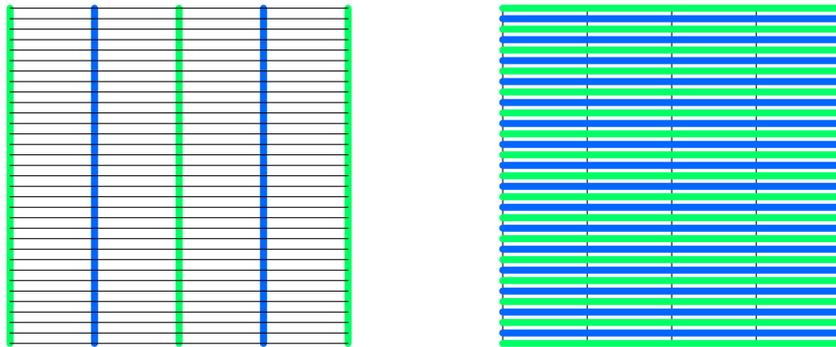


Abbildung 3.1: Schematische Darstellung zum alternierenden Linienglätter.

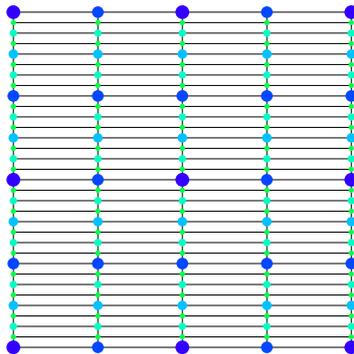


Abbildung 3.2: Schematische Darstellung zum Semi-Coarsening.

Semi-Coarsening ist adäquat im Hinblick auf die gestreckten Gitter, eine Berücksichtigung der inhomogenen anisotropen Diffusion ist aber schwierig und würde überdies die reguläre Struktur zerstören. Block-Glätter sind hier angemessen, wenn man bedenkt, dass die Koeffizienten orthogonal zum Rand gegen Null gehen und das Herausschneiden von Ebenen — etwa im Rahmen eines Block-Gauß-Seidel-Schrittes — die unzureichende Glättung abfangen kann. Matrix-abhängiger Gittertransfer hingegen scheint nicht notwendig, da die Sprünge in den (diskretisierten) Koeffizienten moderat sind.

Methode der Wahl sind demnach (im allgemeinen Fall) alternierende Blockglätter.

3.2.1 Ebenenglätter

Auf dem Tensorprodukt-Gitter ist die Indexmenge der (skalaren) Unbekannten

$$\mathcal{I} = \{\mathbf{1} \leq \mathbf{i} \leq \mathbf{n}\},$$

wobei \mathbf{i} ein Multiindex und $\mathbf{n} = (n_1, \dots, n_d)$ die Knotenzahl des Gitters in den verschiedenen Richtungen ist (was eine Gesamtzahl an $N = n_1 \cdot \dots \cdot n_d$ Unbekannten ergibt). Darauf können $(d - m)$ -dimensionale Blöcke

$$\mathcal{I}_{k_1, \dots, k_m}^{j_1, \dots, j_m} := \{\mathbf{1} \leq \mathbf{i} \leq \mathbf{n} : i_{k_1} = j_1, \dots, i_{k_m} = j_m\}$$

definiert werden, d. h. die Koordinaten der m Richtungen k_1, \dots, k_m werden zu j_1, \dots, j_m fixiert. Die Zahl der Blöcke ist dann $M = (n_1 \cdot \dots \cdot n_m) / (n_{k_1} \cdot \dots \cdot n_{k_m})$. Der Einfachheit halber seien für fixe $k_1, \dots, k_m, j_1, \dots, j_m$ die Blöcke mit $\mathcal{I}_1, \dots, \mathcal{I}_M$, bezeichnet, sodass $\mathcal{I} = \mathcal{I}_1 \dot{\cup} \dots \dot{\cup} \mathcal{I}_M$.

Diese Partitionierung korrespondiert mit Vektor- und Matrixblöcken, \mathbf{u}_i , \mathbf{b}_i und \mathbf{A}_{ij} , $1 \leq i, j \leq M$, und definiert einen Block-Gauß-Seidel-Algorithmus der Form

$$\mathbf{u}_i^{k+1} = \mathbf{A}_{ii}^{-1} \left(\mathbf{b}_i - \sum_{j < i} \mathbf{A}_{ij} \mathbf{u}_j^{k+1} - \sum_{j > i} \mathbf{A}_{ij} \mathbf{u}_j^k \right). \quad (3.9)$$

Für zwei und drei Dimensionen ist der Fall $m = 1$ als Linienglätter, der Fall $m = 2$ als Ebenenglätter bekannt.

Im allgemeinen Kontext kommen den Fällen $m = d$ (direkter Löser), $m = d - 1$ (Hyper-Ebenenglätter), $m = 1$ (Linienglätter) und $m = 0$ (Punktglätter) Sonderstellungen zu, die noch herausgearbeitet werden.

Wenn es nicht anders betont wird, werden (nicht-triviale) Blocksysteme genauso wie die Grobgittersysteme im Folgenden mit Bi-CGSTAB gelöst.

3.2.2 Numerische Tests

Betrachtet wird als erstes das hinreichend komplexe Modellproblem

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{1}{2} \sum_{i=1}^d \sigma_i^2 S_i^2 \frac{\partial^2 u}{\partial S_i^2} - r \sum_{i=1}^d S_i \frac{\partial u}{\partial S_i} + ru &= 0 & \forall (\mathbf{S}, t) \in \mathbf{R}_+^d \times (0, T), \\ u(\mathbf{S}, 0) = g(\mathbf{S}) &:= \max \left(K - \sum_{i=1}^d S_i, 0 \right) & \forall \mathbf{S} \in [0, 1]^d \end{aligned}$$

(Black-Scholes-Gleichung für unkorrelierte Assets), das genau die diskutierten Eigenschaften der Entartung an den Gebietsrändern (und zwar an den Ebenen durch den Ursprung) aufweist.

Als Testdaten werden für $d = 2, 3, 4$ Dimensionen für die Praxis typische Werte verwendet: $\sigma_i = 0.4$, $r = 0.05$, $T = 1$, $K = 0.25$. Die Gitter werden zunächst uniform und isotrop gewählt. Für eine Diskretisierung mit zentralen Differenzen wird jeweils ein impliziter Euler-Schritt über das ganze Zeitintervall gerechnet.

Ausgehend von einer Startlösung \mathbf{u}^0 (der Anfangsbedingung) werden zu den Iterierten \mathbf{u}^k gemittelte Raten

$$\bar{\rho}_k = \left(\frac{\|\mathbf{b} - \mathbf{A}\mathbf{u}^k\|}{\|\mathbf{b} - \mathbf{A}\mathbf{u}^0\|} \right)^{\frac{1}{k}} \quad (3.10)$$

betrachtet. In der Praxis zählt allerdings in erster Linie die *Time to Solution*, gemessen an den Arbeitseinheiten $W(\epsilon)$, die zur Reduktion des Fehlers um einen kleinen Faktor ϵ notwendig sind. Als Standard-Einheit für den Aufwand bietet sich ein einzelner $V(1, 1)$ -Zyklus

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

mit (Punkt-)Gauß-Seidel-Glätter auf dem entsprechenden Gitter an. Zum Effizienzvergleich ist also nicht eine über die Iterationen gemittelte Rate nach (3.10) sinnvoll, sondern eine, die den Aufwand pro Zyklus mit einbezieht,

$$\hat{\rho} = \bar{\rho} \frac{1}{W}, \quad (3.11)$$

wobei W die mittlere CPU-Zeit einer Iteration in Standard-Einheiten, also gemessen am Referenz- $V(1, 1)$ -Zyklus ist.

Es sei bemerkt, dass diese Größen architektur- und implementierungsabhängig sind.

2-Faktor-Basket

In zwei Dimensionen sind Punkt- und Linienglätter zu vergleichen. Die Blocksysteme sind tridiagonal und können mit einem schnellen direkten Tridiagonal-Löser mit Aufwand proportional zur Linienlänge invertiert werden, sodass insgesamt lineare Komplexität für jeden Zyklus erhalten bleibt. Das sieht man auch in Abb. 3.3, in der der relative Aufwand W pro Schritt im Wesentlichen konstant über die Gitterstufen ist.

Darüberhinaus ist festzustellen, dass beim Linienglätter im Gegensatz zum Punktglätter die asymptotische h -Unabhängigkeit der Raten erhalten bleibt.

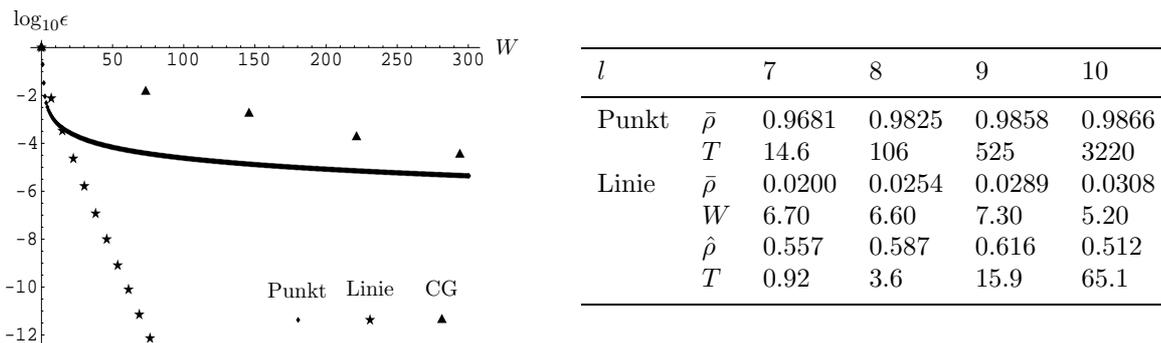


Abbildung 3.3: 2D Black-Scholes: **Links:** Logarithmischer Fehler in Abhängigkeit der Recheneinheiten für Punkt-Gauß-Seidel-Glätter und alternierenden Linienglätter auf einem Gitter mit 263169 Punkten. **Rechts:** Gemittelte Konvergenzrate $\bar{\rho}$, Lösungszeit T (s) für eine Fehlerreduktion von 10^{-12} sowie Aufwand W pro Zyklus und relative Konvergenzrate $\hat{\rho}$ für Gitterstufen $l = 7, \dots, 10$ (16641 bis 1050625 Punkte).

3-Faktor-Basket

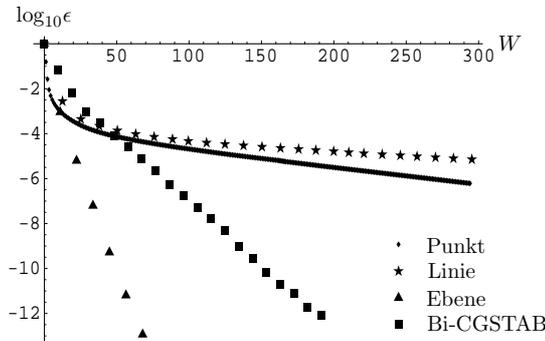
Zusätzlich zu den Möglichkeiten im zweidimensionalen Fall ist hier Ebenenglättung eine Alternative. Da im Gegensatz zu den Linien zweidimensionale Blöcke nicht mehr in optimaler Zeit mit einem direkten Löser invertiert werden können, muss man sich über einen iterativen Löser und eine sinnvolle Fehlerreduktion für die Teilsysteme Gedanken machen.

Zu untersuchen ist zunächst der Einfluss der Genauigkeit ϵ_2 der Blocklösung (Bi-CGSTAB) auf die Konvergenzraten und den Gesamtaufwand zur Lösung mit vorgegebener Genauigkeit $\epsilon = 10^{-12}$. Die Daten finden sich in Tabelle 3.1. Auch über dieses Beispiel hinaus zeigt sich in der Regel, dass eine Fehlerreduktion auf weniger als 1/10 für hinreichende Glättung nicht notwendig (und damit nicht sinnvoll) ist. Im folgenden wird daher $\epsilon_2 = 0.1$ gesetzt.

Tabelle 3.1: 3D Black-Scholes, Ebenen-Glätter auf 65^3 -Gitter: Mittlere Rate $\bar{\rho}$, gemittelte CPU-Zeit \bar{t} pro Iteration und Gesamtzeit T (s) zur Lösung für $\epsilon_2 = 10^{-k/3}$.

k	1	2	3	4	5	10
$\bar{\rho}$	0.085	0.019	0.0070	0.0066	0.0066	0.0075
\bar{t}	9.6	11.7	13.1	14.6	16.5	27.5
T	115	82	79	87	99	165

Damit können die verschiedenen in drei Dimensionen möglichen Blockglätter verglichen werden. In den Experimenten aus Abbildung 3.4 zeigt sich, dass der Linienglätter keine Mehrgittereffizienz zeigt. Die verbesserten Raten gegenüber dem Punktglätter werden durch den höheren Aufwand wettgemacht. Der Faktor (≈ 11) rührt vom 3-fachen Aufwand durch die alternierenden Schritte und von der Komplexität des Tridiagonal-Lösers. Einzig der alternierende Ebenenglätter zeigt robuste Raten, wobei nicht von einem asymptotischen Befund gesprochen werden kann.



	l	4	5	6	7
Punkt	$\bar{\rho}$	0.74	0.93	0.98	
	$\bar{\rho}$	0.2276	0.6710	0.9031	
	W	11.6	11.1	12.3	
Ebene	$\hat{\rho}$	0.88	0.96	0.99	
	$\bar{\rho}$	0.0006	0.026	0.0070	0.0158
	W	6.5	5.4	11.3	
	$\hat{\rho}$	0.31	0.51	0.65	

Abbildung 3.4: 3D Black-Scholes: **Links:** Logarithmischer Fehler in Abhängigkeit der Zeiteinheiten W für Punkt-Gauß-Seidel-Glätter, alternierende Linien- und Ebenenglätter auf einem Gitter mit 274625 Punkten. **Rechts:** Gemittelte Konvergenzrate $\bar{\rho}$, Zeiteinheit W und relative Konvergenzrate $\hat{\rho}$ für Gitterstufen $l = 4, \dots, 7$ (4913 bis 2146691 Punkte).

Als weiteres Beispiel dafür dient das transformierte System aus 1.3.2 mit polynomialen Koeffizienten, das zu allen Rändern von $[0, 1]^3$ (außer am Dirichlet-Rand $x_1 = 1$) entartet. Über die Volatilitäten $\sigma_1 = 0.01$, $\sigma_2 = 0.1$, $\sigma_3 = 0.5$, die den für Aktien relevanten Bereich abdecken, wird eine zusätzliche Anisotropie eingeführt. Für die Ergebnisse siehe Tabelle 3.2.

l	4	5	6	7
$\bar{\rho}$	$3.8 \cdot 10^{-4}$	$5.3 \cdot 10^{-2}$	$2.9 \cdot 10^{-2}$	$6.1 \cdot 10^{-2}$
it	2	3	4	5
T	0.25	3.3	45	591

Tabelle 3.2: 3D transformierte Gleichung: Mittlere Rate $\bar{\rho}$, Iterationszahl it und CPU-Zeit T (s) für eine Fehlerreduktion von 10^{-6} auf Gitterstufen $l = 4, \dots, 7$ (4913 bis 2146691 Punkte) für den Ebenenglätter.

Wie schon im vorangegangenen Beispiel stellt man fest, dass die lineare Komplexität verlorengeht, da die Blocksysteme nicht mit optimalem Aufwand gelöst werden. Die Anzahl der

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

CG-Iterationen nimmt proportional mit der Zahl der Punkte in jede Richtung zu, was der Wurzel der Konditionszahl der Blocksysteme entspricht.

Im Allgemeinen (d. h. in d Dimensionen) sind bei n Punkten in jede von d Richtungen $d \cdot n$ Block-Systeme pro alternierendem Glättungsschritt zu lösen, die jeweils n^{d-1} Unbekannte haben. Da sich aus der Gitterweite $1/n$ eine Konditionszahl proportional n^2 ergibt, sind $\mathcal{O}(n)$ CG-Schritte mit einem Aufwand von je $\mathcal{O}(n^{d-1})$ Operationen erforderlich. Insgesamt steigt also der Aufwand pro Iteration gemäß $\mathcal{O}(n^{d+1})$, also bei insgesamt $N = n^d$ Unbekannten für eine Fehlerreduktion von ϵ asymptotisch nach

$$W(N, \epsilon) \sim |\log \epsilon| \cdot N^{(d+1)/d}. \quad (3.12)$$

Das superlineare Verhalten wird in höheren Dimensionen zunehmend schwächer.

Um insgesamt lineare Konvergenz zu erhalten, wird der CG-Block-Löser durch einen inneren Multigrid-Löser mit Linienglättung in den Ebenen ersetzt. Ein V-Zyklus reduziert den Fehler ausreichend.

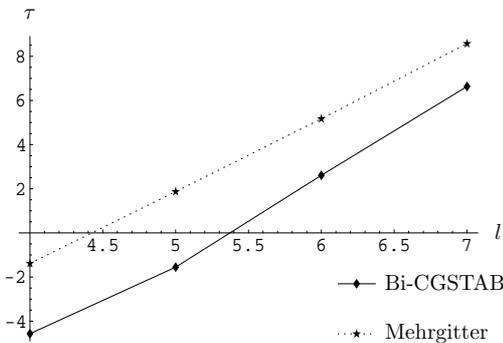


Abbildung 3.5: 3D Black-Scholes, Ebenen-Gauß-Seidel-Glättung mit Bi-CGSTAB und Multigrid (mit Linienglättung) als Ebenen-Löser: $\tau = -\frac{\log_2 T}{\log_{10} \epsilon}$ mit CPU-Zeit T bei Gitterstufe l (4913 bis 2146691 Punkte) für eine Fehlerreduktion $\epsilon = 10^{-12}$.

Den Daten in Abbildung 3.5 entnimmt man — in Übereinstimmung mit (3.12) — eine asymptotische Ordnung von etwa $4/3$ für den Aufwand des Mehrgitterverfahrens mit CG als Blocklöser, und lineare Ordnung für das rekursiv angewandte Mehrgitter. Für praktische Zwecke ist allerdings der zusätzliche Aufwand zum Aufbau der Mehrgitterhierarchie auf den Blöcken zu groß.

Höhere Dimensionen

Schon in vier Dimensionen zeigt sich, dass für praktisch berechenbare Systeme ein einfaches CG-Verfahren dem Mehrgitter-Algorithmus mit Block-Glättung auf dreidimensionalen Hyperebenen überlegen ist (Tabelle 3.3).

Da die Kondition des Systems proportional $h^{-2} = N^{2/d}$ ist, ergibt sich für das CG-Verfahren ein asymptotischer Aufwand wie schon in (3.12). Die vorteilhafte asymptotische Komplexität des Mehrgitterzugangs wirkt sich für praktikable Gitterweiten nicht aus.

Anisotrope Gitter

Die vorangegangenen Abschnitte beschäftigten sich ausschließlich mit (in alle Richtungen) äquidistanten Gittern. Dünne Gitter setzen sich jedoch aus stark anisotropen Gittern zusammen, sodass die Punkte mitunter auf wenigen niederdimensionalen Ebenen liegen, in denen

l		3	4	5
Bi-CGSTAB	T	0.36	18.5	809
Multigrid	T	1.06	33.5	1292
	$\bar{\rho}$	$3 \cdot 10^{-7}$	$6.7 \cdot 10^{-5}$	$7 \cdot 10^{-4}$

Tabelle 3.3: 4D Black-Scholes: Lösungszeit T (s) für Stufen $l = 3, 4, 5$ (6561, 83521 und 1185921 Punkte) für Bi-CGSTAB-Löser und Mehrgitter mit alternierendem 3D-Hyperebenen-Glätter sowie die mittlere Mehrgitterrate $\bar{\rho}$.

sehr kleine Gitterweiten auftreten, sodass die Kondition der Blocksysteme beliebig schlecht sein kann. In diesem Fall ist nur ein nach dem Schema der letzten Abschnitte auf niedrigerdimensionale Blocksysteme rekursiv angewandtes Mehrgitterverfahren robust.

Tabelle 3.4 zeigt für einen derartigen Algorithmus, angewandt auf die dreidimensionale Black-Scholes-Gleichung, die Lösungszeit für verschiedene dreidimensionale kartesische Gitter mit jeweils $N = (2^{l_1} + 1) \cdot (2^{l_2} + 1) \cdot (2^{l_3} + 1)$ Punkten, wobei $l_1 + l_2 + l_3 = 15$. Solche Familien von Gittern sind essenziell für die Dünngitter-Kombitechnik.

l_1	l_2	l_3	N	T	T/N
0	0	15	131076	106.88	0.000815405
0	3	12	73746	29.71	0.000402869
0	6	9	66690	21.46	0.000321787
1	2	12	61455	27.89	0.000453828
1	5	9	50787	21.7	0.000427275
2	2	11	51225	25.54	0.000498585
2	5	8	42405	20.12	0.000474472
3	4	8	39321	19.91	0.000506345
4	4	7	37281	33.61	0.000901532
5	5	5	35937	17.58	0.000489189

Tabelle 3.4: 3D Black-Scholes: CPU-Zeit T für rekursives Mehrgitter auf anisotropen Gittern mit Stufen l_1, l_2, l_3 ($n_i = 2^{l_i} + 1$ Punkten) in den einzelnen Richtungen und insgesamt $N = n_1 n_2 n_3$ Unbekannten.

Im Idealfall wäre T/N konstant über die gesamte Klasse von Gittern. Die Auswirkung auf die Lastverteilung bei der Parallelisierung der Kombinationstechnik wird in Abschnitt 5.4 diskutiert.

3.3 Multigrid für Hindernisprobleme

Die Beobachtungen aus 3.2 führen in Kombination mit den in 3.1.1 vorgestellten Komponenten anhand der nun folgenden Überlegungen auf robuste Mehrgitterverfahren für anisotrope lineare Komplementaritätsprobleme.

3.3.1 Der projizierte Multigrid-Zyklus

Im Mehrgitter-Sinne ordnet man (3.1)–(3.3) ein Grobgittersystem der Gestalt

$$\mathbf{A}_H \mathbf{d}_H \leq \mathbf{b}_H \quad := \quad \bar{I}_H^h(\mathbf{A}_h \mathbf{u}_h - \mathbf{b}_h), \quad (3.13)$$

$$\mathbf{d}_H \geq \mathbf{c}_H \quad := \quad \hat{I}_H^h(\mathbf{c}_h - \mathbf{u}_h), \quad (3.14)$$

$$(\mathbf{A}_H \mathbf{d}_H - \mathbf{b}_H) \star (\mathbf{d}_H - \mathbf{c}_H) = \mathbf{0} \quad (3.15)$$

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

zu, wobei \bar{I}_H^h, \hat{I}_H^h Restriktionsoperatoren sind und die Grobgitterkorrektur mit einer passenden Prolongation \bar{I}_h^H auf das Feingitter gebracht wird.

Problem (3.13) bis (3.15) definiert eine Mehrgitter-Iteration der folgenden Struktur (wobei wieder die Indizes h, H durch die Mehrgitterstufe l ersetzt werden):

Algorithmus 3.3. Projizierter V-Zyklus

```

procedure PV( $l, \mathbf{A}^l, \mathbf{b}^l, \mathbf{c}^l, \mathbf{u}^l$ )
begin
  if  $l = 1$  then
     $\mathbf{u}^1 \leftarrow \text{SORP-Bi-CGSTAB}(\mathbf{A}^1, \mathbf{b}^1, \mathbf{c}^1)$ ;
  else begin
     $\mathbf{u}^l \leftarrow \mathcal{S}(\mathbf{A}^l, \mathbf{b}^l, \mathbf{c}^l, \mathbf{u}^l)$ ;
     $\mathbf{r}^l = \mathbf{b}^l - \mathbf{A}\mathbf{u}^l$ ;
     $\mathbf{r}^{l-1} = \bar{I}_l^{l-1}\mathbf{r}^l$ ;
     $\mathbf{c}^{l-1} = \hat{I}_l^{l-1}(\mathbf{c}^l - \mathbf{u}^l)$ ;
     $\mathbf{d}^{l-1} \leftarrow \text{PV}(l-1, \mathbf{A}^{l-1}, \mathbf{r}^{l-1}, \mathbf{c}^{l-1}, \mathbf{d}^{l-1})$ ;
     $\mathbf{d}^l = \bar{I}_{l-1}^l\mathbf{d}^{l-1}$ ;
     $\mathbf{u}^l = \mathbf{u}^l + \mathbf{d}^l$ ;
     $\mathbf{u}^l \leftarrow \mathcal{S}(\mathbf{A}^l, \mathbf{b}^l, \mathbf{c}^l, \mathbf{u}^l)$ ;
  end
return Projizierter V-Zyklus

```

3.3.2 Gittertransfer

Eine Minimalforderung an die Restriktion besteht offenbar darin, dass für die exakte diskrete Lösung (\mathbf{u}_h) von (3.1)–(3.3)

$$\begin{aligned}
 \bar{I}_H^h(\mathbf{A}_h\mathbf{u}_h - \mathbf{b}_h) &\geq \mathbf{0}, \\
 \hat{I}_H^h(\mathbf{c}_h - \mathbf{u}_h) &\leq \mathbf{0}, \\
 \bar{I}_H^h(\mathbf{A}_h\mathbf{u}_h - \mathbf{b}_h) \star \hat{I}_H^h(\mathbf{c}_h - \mathbf{u}_h) &= \mathbf{0},
 \end{aligned} \tag{3.16}$$

d. h. $\mathbf{d}_H = \mathbf{0}$ das Grobgittersystem (3.13)–(3.15) löst. Dies garantiert, dass \mathbf{u}_h ein Fixpunkt der Iteration ist. Insbesondere (3.16) ist essenziell und im allgemeinen, z. B. auch für „*Full Weighting*“, nicht erfüllt. Injektion wiederum, die klarerweise (3.16) erfüllt, führt durch die verlorene Approximationsordnung auf dürftige Konvergenz.

In Anlehnung an [HK87] und [KSW90] wird daher einseitige Extrapolation am freien Rand eingeführt:

$$\left(\bar{I}_H^h\mathbf{u}_h\right)_i = \begin{cases} \left(I_H^h\bar{\mathbf{u}}_h^A\right)_i & \text{wenn } i \in \mathcal{A}, \\ \left(I_H^h\bar{\mathbf{u}}_h^I\right)_i & \text{wenn } i \in \mathcal{I} \end{cases} \tag{3.17}$$

mit multilinearer I_H^h und

$$\left(\bar{\mathbf{u}}_h^{A/I}\right)_i = \begin{cases} u_i & \text{wenn } i \in \mathcal{A}/\mathcal{I}, \\ 0 & \text{sonst.} \end{cases}$$

Die entsprechende Prolongation ist dann

$$\left(\bar{I}_h^H\mathbf{u}_H\right)_i = \begin{cases} 0 & \text{wenn } c_i = 0, \\ I_h^H\mathbf{u}_H(x_h) & \text{sonst.} \end{cases} \tag{3.18}$$

(3.17) und (3.18) trennen aktive and inaktive Menge und erhalten die Komplementarität über die Gitterebenen hindurch. Die Grobgitterkorrektur hält die Position des freien Randes fest. Das Hindernis wird punktweise mittels Injektion \hat{I}_h^H restringiert.

Im Folgenden werden verschiedene Transferoperatoren verglichen, wobei der hybride Glätter aus Abschnitt 3.3.4 verwendet wird. Deutlich sieht man in Abb. 3.6 (links) die verlorene Approximationsordnung der Injektion. Bei einseitigen (projizierten) bilinearen Transferoperatoren erhält man grob die gleiche asymptotische Konvergenz wie im linearen (unrestringierten) Fall. Da der freie Rand nur am feinsten Gitter (durch den Glätter) korrigiert wird, geht in der Anfangsphase der Mehrgittereffekt verloren. Teilweise wird das durch global bilineare Prolongation behoben, allerdings wird die exakte Position des freien Randes ebenfalls erst spät erfasst, was an den Sprüngen des projizierten Residuums ersichtlich ist.

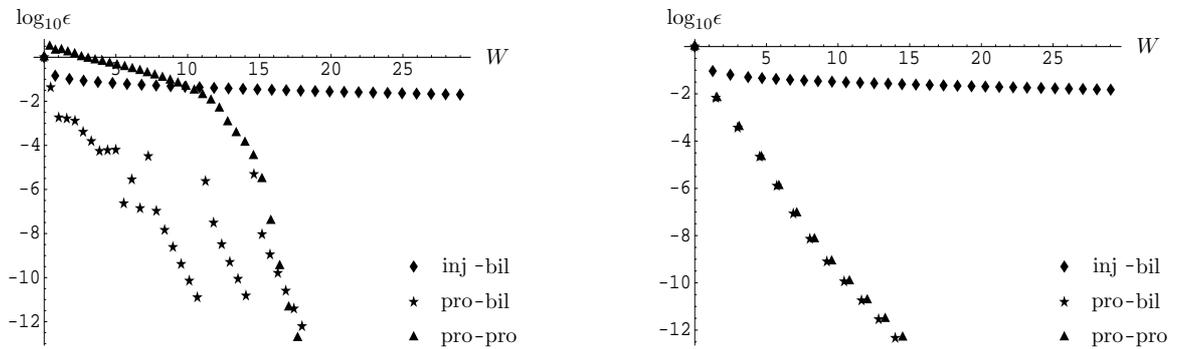


Abbildung 3.6: 2D Black-Scholes, Amerikanische Option: Fehlerreduktion ϵ bei 263169 Punkten in Abhängigkeit der CPU-Zeit W für V(1,1)-Zyklen (**links**) und mit drei vorgeschalteten F(1,1)-Schritten (**rechts**): Injektion und bilineare Interpolation, einseitige (projizierte) Restriktion und bilineare Interpolation, einseitige Restriktion und Prolongation.

Dies zeigt, wie wichtig eine gute Anfangsapproximation der aktiven Menge ist, welche wiederum hierarchisch von den größeren Gittern gewonnen werden kann.

3.3.3 Volles Mehrgitter

Motiviert durch 3.3.2 wird von einer Startlösung auf dem größten Gitter eine Folge von Approximationen auf den feineren Stufen bis hinauf zu L konstruiert.

Algorithmus 3.4. Projiziertes Volles Mehrgitter

```

procedure PFM( $\mathbf{A}^L, \mathbf{b}^L, \mathbf{c}^L, \mathbf{u}^L$ )
begin
     $\mathbf{u}^1 \leftarrow \text{LCP}(\mathbf{A}^1, \mathbf{b}^1, \mathbf{c}^1)$ ;
    for  $l = 2, \dots, L$  begin
         $\mathbf{u}^l = I_{l-1}^l \mathbf{u}^{l-1}$ ;
        for  $k = 1, \dots, n_l$  begin
             $\mathbf{u}^l \leftarrow \text{PV}(l, \mathbf{A}^l, \mathbf{b}^l, \mathbf{c}^l, \mathbf{u}^l)$ ;
        end
    end
return Projiziertes Volles Mehrgitter
    
```

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

Da die Topologie der aktiven Menge sehr einfach ist (siehe Abb. 3.7 in drei Dimensionen), reichen wenige volle Mehrgitterschritte aus. Sobald der freie Rand gefunden wurde, reduziert sich das Problem auf ein lineares auf der inaktiven Menge. Die Raten stimmen (bei einseitiger Restriktion) für bilineare und einseitige Prolongation weitgehend überein und entsprechen annähernd dem linearen Fall (siehe vorne Abb. 3.6).

3.3.4 Robuste Glätter

Der projizierte Gauß-Seidel-Algorithmus wird gerne als Glätter für Hindernisprobleme verwendet (siehe z. B. schon [BC83]), aber im vorliegenden Fall ist die Konvergenz schwach. Gründe und Auswege dafür wurden schon im vorangegangenen Abschnitt für das lineare Problem besprochen.

Zwei Zugänge liegen auf der Hand, die Idee der Blockglätter auf Hindernisprobleme zu übertragen. Zum einen könnte man die Hindernisprobleme blockweise lösen:

$$\mathbf{u}_i^{k+1} = \mathcal{A}_{ii}^{-1}(\mathbf{b}_i - \sum_{j<i} \mathbf{A}_{ij} \mathbf{u}_j^{k+1} - \sum_{j>i} \mathbf{A}_{ij} \mathbf{u}_j^k, \mathbf{c}_i). \quad (3.19)$$

$\mathcal{A}_{ii}^{-1}(\mathbf{b}_i, \mathbf{c}_i)$ soll hier die Lösung des auftretenden Komplementaritätsproblems am Block bezeichnen. Es kann dazu die kombinierte SORP-Bi-CGSTAB Iteration aus 3.1.1 herangezogen werden. Eine andere Wahl wurde in [CP99] vorgeschlagen. Hier wird für ein zweidimensionales Problem mit stochastischer Volatilität ein projizierter punktwiser Gauß-Seidel-Schritt durch einen Schritt mit Linienglättung auf der inaktiven Menge in Richtung der Anisotropie beschleunigt. Dem Beispiel folgen wir in der Form von alternierenden Blockglättern auf der inaktiven Menge als zweitem Schritt.

Tabelle 3.5 zeigt für dasselbe Beispiel wie in 3.3.2, dass dieser hybride Glätter leicht bessere Raten hat, was übrigens auch die generell beobachtete Tendenz ist. Der zusätzliche Bonus, dass im Glättungsschritt keine Komplementaritätprobleme gelöst werden müssen, ist entscheidend für die Effizienz.

l		5	6	7	8	9
Linie	T	0.31	1.3	4.9	23	138
	$\bar{\rho}$	0.047	0.03	0.043	0.059	0.099
Hybrid	T	0.26	0.92	3.8	16	85
	$\bar{\rho}$	0.018	0.028	0.041	0.043	0.051

Tabelle 3.5: 2D Amerikanische Option: CPU-Zeit T (s) und mittlere Konvergenzrate $\bar{\rho}$ für eine Fehlerreduktion von 10^{-12} auf Gitterstufen $l = 5, 6, 7, 8, 9$ (1089 bis 263169 Punkte) für projizierten Linienglätter und hybriden Glätter.

Wir gehen jetzt noch einmal zurück zum transformierten System aus 1.3.2 und vergleichen die Konvergenz mit dem linearen Fall (siehe Tab. 3.2). Die ersten drei Schritte waren stets $F(1, 1)$, die restlichen $V(1, 1)$ bei einseitiger Restriktion und Prolongation. Tabelle 3.6 zeigt vergleichbare Resultate wie Tabelle 3.2 für den linearen Fall. Der freie Rand (rücktransformiert auf natürliche Koordinaten) ist in Abb. 3.7 dargestellt.

l	4	5	6	7
$\bar{\rho}$	$1.8 \cdot 10^{-4}$	$4.6 \cdot 10^{-3}$	$2.4 \cdot 10^{-2}$	$8.9 \cdot 10^{-2}$
it	2	3	4	6
T	0.28	3.5	42	867

Tabelle 3.6: 3D transformiertes Hindernisproblem: Mittlere Konvergenzrate $\bar{\rho}$, Zahl der Iterationen it und CPU-Zeit T für eine Fehlerreduktion von 10^{-6} auf Gitterstufen $l = 4, \dots, 7$ (4913 bis 2146691 Punkte) für den hybriden Glätter.

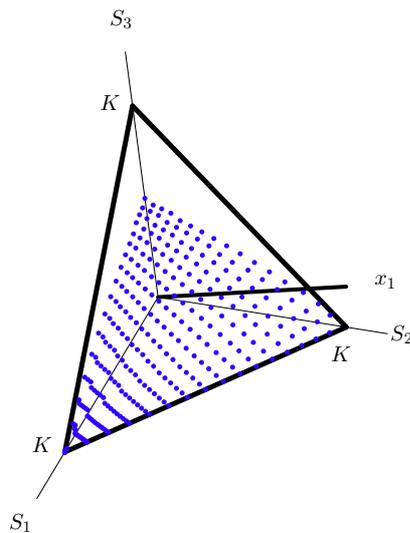


Abbildung 3.7: Der freie Rand im Raum, der von drei unkorrelierten Aktien mit $\sigma_1 = 0.01$, $\sigma_2 = 0.1$, $\sigma_3 = 0.5$ aufgespannt wird. Der Ausübungsbereich für einen Amerikanischen Put ($T = 1$, $r = 0.05$) auf den Basket mit Ausübungspreis K liegt unter der Fläche, die im transformierten Koordinatensystem nach 1.3.2 als Funktion $x_1 = x_1(x_2, x_3)$ dargestellt werden kann.

3.3.5 Kaskadisches Mehrgitter

In [HBS03] bewerben Braess et al. einen kaskadischen Mehrgitter-Algorithmus für Variationsungleichungen. Die Idee ist attraktiv, da das Problem der Restriktion des freien Randes auf größere Ebenen umgangen wird: Ausgehend vom Grobgitter wird die Lösung auf feinere Gitter (schließlich bis hinauf zu L) prolongiert und dort hinreichend glättet.

Algorithmus 3.5. Projiziertes Kaskadisches Multigrid

```

procedure PCM( $A^L, b^L, c^L, u^L$ )
begin
 $u^1 \leftarrow$  SORP-Bi-CGSTAB( $A^1, b^1, c^1$ );
for  $l = 2, \dots, L$  begin
     $u^l = I_{l-1}^l u^{l-1}$ ;
    while  $\|r^l\| > \epsilon_l \|r_0^l\|$  begin
         $u^l \leftarrow \mathcal{S}(A^l, b^l, c^l, u^l)$ ;
    end
end
return Projiziertes Kaskadisches Multigrid
    
```

Hier ist r_0^l das Ausgangsresiduum auf Stufe l , unmittelbar nach der Prolongation.

Da die gröberen Gitter nur einmal miteinbezogen werden, ist es wesentlich, dass die langwelligen Fehleranteile dann schon genügend gedämpft werden. Daher wird in [HBS03] das

3 Robuste Mehrgitterverfahren für anisotrope Gleichungen und Hindernisprobleme

folgende Abbruchkriterium auf Stufe l vorgeschlagen:

$$\epsilon_l = \epsilon_0 0.4^{L-l}. \quad (3.20)$$

Die Anzahl der Glättungsschritte erweist sich dann als weitgehend konstant über die Gitter-

l	N_l	$\ \mathbf{r}_0^{+,l}\ $	$\ \mathbf{r}^{+,l}\ $	n_l
3	81	$1.94 \cdot 10^1$	$2.67 \cdot 10^{-4}$	3
4	289	$7.98 \cdot 10^0$	$1.33 \cdot 10^{-3}$	3
5	1089	$3.31 \cdot 10^0$	$1.18 \cdot 10^{-3}$	3
6	4225	$1.30 \cdot 10^0$	$1.04 \cdot 10^{-3}$	3
7	16641	$4.94 \cdot 10^{-1}$	$6.52 \cdot 10^{-3}$	2
8	66049	$1.93 \cdot 10^{-1}$	$2.03 \cdot 10^{-3}$	2
9	263169	$9.15 \cdot 10^{-2}$	$5.94 \cdot 10^{-4}$	2

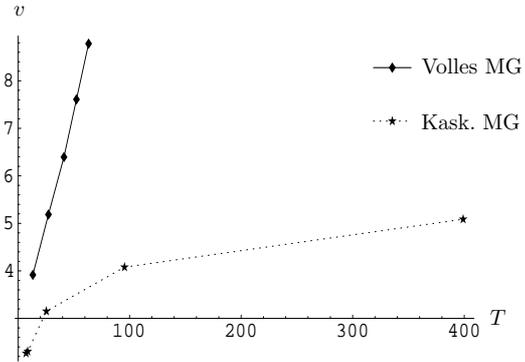


Abbildung 3.8: 2D Black-Scholes-Modell, Amerikanische Option: **Rechts:** CPU-Zeit T für v gültige Stellen mit vollem und kaskadischem Mehrgitter (513^2 Punkte). **Links:** Anfangsfehler $\|\mathbf{r}_0^{+,l}\|$, Iterationen n_l und Endfehler $\|\mathbf{r}^{+,l}\|$ auf allen Stufen l mit N_l Freiheitsgraden.

ebenen (Abbildung 3.8 für $\epsilon_0 = 0.1$). Das führt auf lineare Komplexität für den kaskadischen Algorithmus bis hinunter auf einen Fehler in der Größenordnung des Diskretisierungsfehlers (siehe Abb. 3.8, rechts), dann geht der Mehrgittereffekt verloren.

Für Extrapolationsmethoden, wie sie hier benötigt werden, reicht das leider nicht aus.

4 Konvergenzabschätzungen für die Kombinationstechnik

Überblick

In den vorangegangenen Kapiteln wurden klassische Diskretisierungsmethoden für parabolische Gleichungen und deren effiziente Mehrgitterlösung diskutiert, wobei festzustellen war, dass in mehr als drei Raumdimensionen der exponentiell steigende Aufwand für hinreichend exakte Lösungen in vernünftiger Zeit undurchführbar wird.

In Abschnitt 4.1 wird zunächst erläutert, wie der „Fluch der Dimension“ durch Approximation in sog. *Dünngitterräumen* durchbrochen werden kann, was mit jüngsten Anwendungen auf Optionspreisaufgaben untermauert wird.

Für die Kombinationstechnik, einem darauf basierenden Extrapolationsalgorithmus, werden in Abschnitt 4.3 Konvergenzabschätzungen in einer geschlossenen Form für allgemeine Dimensionen bewiesen, was das mathematische Herzstück dieser Arbeit darstellt.

4.1 Das Prinzip dünner Gitter

Betrachtet man in d Dimensionen ein kartesisches Gitter mit n Punkten in jede Richtung, so läßt sich etwa durch multilineare Ansatzfunktionen (für hinreichend glatte Funktionen) ein Approximationsfehler von $\epsilon = \mathcal{O}(n^{-2})$ für den Preis von n^d Freiheitsgraden erreichen. Ansätze höherer Ordnung drücken den Aufwand (bei vorausgesetzter Glattheit), kommen aber um das prinzipielle Problem exponentiell ansteigenden Aufwands nicht herum.

Dünne Gitter eignen sich unter noch stärkeren Regularitätsvoraussetzungen zur Approximation in höheren Raumdimensionen mit einer substantiell geringeren Anzahl an Freiheitsgraden. Dem Ansatz liegt eine Approximation über hierarchische Tensorproduktbasen zugrunde. Sei dazu V ein vollständiger (separabler) Raum, der eine Darstellung als (abgeschlossene) direkte Summe von endlichdimensionalen Räumen der Gestalt

$$V^{(n_1, \dots, n_d)} := V_{n_1}^{(1)} \times \dots \times V_{n_d}^{(d)}$$

mit $n_j \in \mathbb{N}_0$ hat. Ein volles Gitter entspricht dann einer Folge von Unterräumen

$$V^{(n)} := \bigoplus_{n_k \leq n} V^{(n_1, \dots, n_d)},$$

für das dünne Gitter wählt man hingegen

$$\hat{V}^{(n)} := \bigoplus_{\sum_k n_k \leq n} V^{(n_1, \dots, n_d)}.$$

Betrachten wir als Spezialfall $V = H_0^1([0, 1]^d)$ und

$$V_{n_i}^{(i)} := \left\{ v : [0, 1] \rightarrow \mathbb{R} \mid v \text{ ist stückweise linear mit Stützstellen } \frac{j}{2^{n_i}}, 0 \leq j \leq 2^{n_i} \right\}.$$

Die Abbildungen 4.1 und 4.2 zeigen die entsprechenden Gitter für $d = 2$.

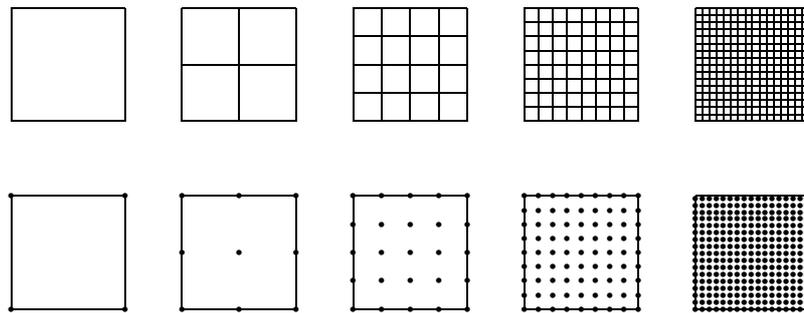


Abbildung 4.1: Zweidimensionale Vollgitterhierarchie, dargestellt sind die Freiheitsgrade der hierarchischen Basis als Knoten. Die Anzahl steigt in d Dimensionen mit 2^{nd} .

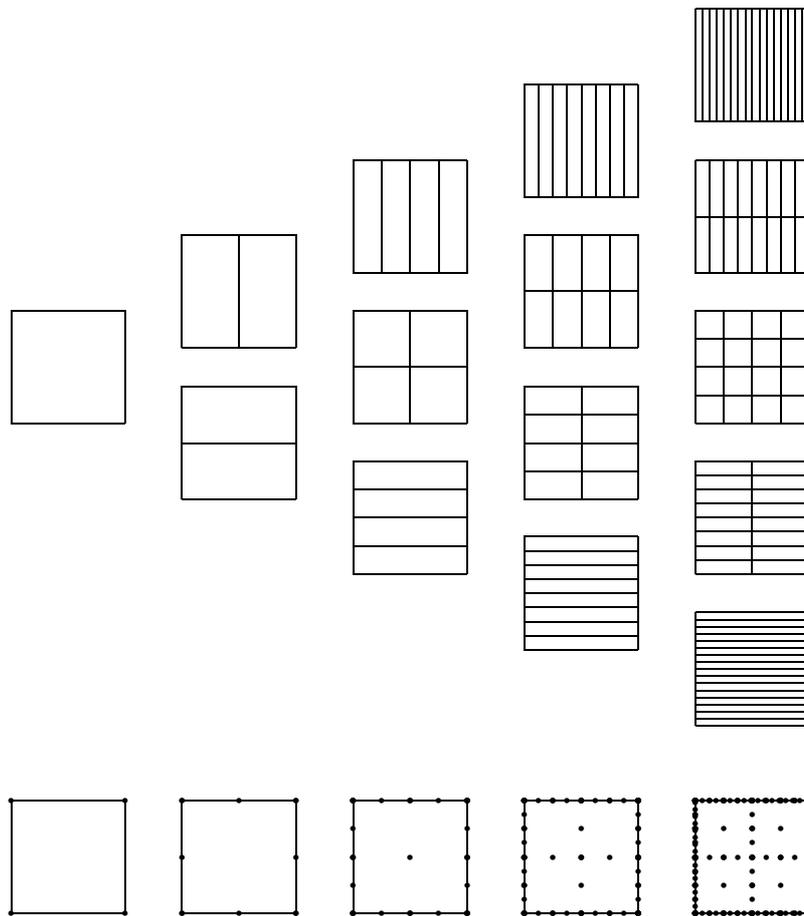


Abbildung 4.2: Zweidimensionale Dünngitterhierarchie. Der Raum auf jeder Stufe $n \geq 0$ entspricht einer Familie von Gittern. Die Anzahl der Knoten pro Gitter steigt in d Dimensionen nur mit $\mathcal{O}(2^n)$, die der Gitter mit $\mathcal{O}(n^{d-1})$.

4 Konvergenzabschätzungen für die Kombinationstechnik

Es ist einsichtig, dass es auf Stufe n

$$N = \binom{n+d-1}{d-1} \quad (4.1)$$

Gitter mit jeweils genau 2^n Elementen und $\mathcal{O}(2^n)$ Knoten gibt. Die Anzahl der Unbekannten steigt also nur nach

$$N_n^{(d)} = \mathcal{O}(n^{d-1}2^n). \quad (4.2)$$

Bungartz [Bun92] zeigt für den entsprechenden multilinearen Interpolanden u_n die Abschätzung

$$\|u - u_n\|_\infty \leq \frac{1}{6^d} \left\| \frac{\partial^{2d} u}{\partial x_1^2 \dots \partial x_d^2} \right\|_\infty \left(1 + \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i} \right) h_n^2 \quad (4.3)$$

für Funktionen mit derart beschränkten gemischten Ableitungen. Der Dünngitterraum $\hat{V}^{(n)}$ hat damit die Approximationseigenschaft

$$d_\infty(u, \hat{V}^{(n)}) = \mathcal{O}(h_n^2 (\ln_2 h_n^{-1})^{d-1})$$

mit $h_n = 2^{-n}$ für hinreichend glatte u (d. h. mit beschränkten gemischten zweiten Ableitungen). Die Verschlechterung zum vollen Raum $V^{(n)}$ besteht nur in logarithmischen Faktoren.

Dieses Approximationsresultat und Varianten davon können auf mehrere Weisen zur effizienten Lösung von Aufgabenstellungen hoher (in der Praxis: moderater) Dimension angewandt werden. Über die reine Approximation (*Data Mining*) hinaus sind die Resultate zur Integration und zur Lösung partieller Differentialgleichungen nützlich. Im Folgenden ein kurzer Streifzug mit speziellem Fokus auf Anwendungen aus der Finanzmathematik.

Dünngitter-Integration

Dünne Gitter tauchen schon in den Sechziger-Jahren bei Smolyak [Smo63] zur Integration von Funktionen in Tensorprodukträumen auf. Es ist offensichtlich, wie das Approximationsresultat (4.3) auf entsprechende Produkte von Quadraturformeln zu übertragen ist.

Anwendungen auf Probleme der Finanzmathematik wurden von Griebel und Gerstner [GG98] behandelt. Im Fall von Europäischen Optionen kann dazu die Integraldarstellung (1.23) des Black-Scholes-Preises mittels Greenscher Funktion herangezogen werden.

In [GG03] verwenden sie dazu eine Heuristik für die Verteilung der Freiheitsgrade auf die Raumrichtungen („dimensionsadaptive“ dünne Gitter).

Dünngitter-Finite Elemente

Bungartz zeigt Approximationsresultate wie (4.3) für Elemente erster [Bun92] und höherer Ordnung [Bun98] und überträgt diese Abschätzungen über einen Standard-Galerkin-Ansatz vermöge klassischer Orthogonalitätsprinzipien auf Finite-Elemente-Konvergenzresultate der Gestalt

$$\|u - u_n\|_1 \leq c(d) |\log h_n|^{d-1} h_n^p. \quad (4.4)$$

Entscheidender Nachteil ist die fehlende Lokalität des diskreten Operators, was sich in der Besetztheitsstruktur der Steifigkeitsmatrix niederschlägt.

Petersdorff und Schwab [vPS02] wählen einen ähnlichen Zugang mit Wavelet-Basen. Durch Approximation der Koeffizienten mit Polynomen und spezielle rekursive Auswertung können

sie (wie auch [Bun92] für die Laplace-Gleichung, aber in einem allgemeineren Kontext) trotzdem lineare Komplexität für die Matrix-Vektor-Multiplikation erreichen. Außerdem erhalten sie selbst bei reduzierter Regularität der Anfangsbedingung ($u_0 \in H^\epsilon$ mit $\epsilon > 0$, wie es für Optionspreisaufgaben typisch ist) durch die Glättungseigenschaft des Operators für Elemente der Ordnung p eine L^2 -Fehlerasymptotik der Form $\mathcal{O}(h^{\theta_0 p + \delta})$, wobei $\delta = p/((p+1)d-1)$ ist und $\theta_0 \in (0, 1]$ mit der Regularität des Problems zusammenhängt.

4.2 Die Kombinationstechnik

In [GSZ92] wird eine auf der Dünngitterhierarchie basierende Extrapolationstechnik eingeführt, die dieselben asymptotischen Eigenschaften im Hinblick auf (4.2) und — wie wir sehen werden — auch (4.3) aufweist, bei der aber nur Gleichungen auf kartesischen Gittern gelöst werden müssen. Der Vorzug der Methode vom praktischen Standpunkt besteht darin, dass die einzelnen Probleme völlig entkoppelt parallel mit einer Methode der Wahl gerechnet werden können. Für die Konvergenz der Dünngitterlösung ist dabei die Fehlerasymptotik der Skelettlösungen auf den einzelnen Gittern, die in Abbildung 4.2 für $d = 2$ dargestellt sind, entscheidend.

Der zweidimensionale Fall

Zur Motivation dient der zweidimensionale Fall, da hier übersichtlich das Prinzip deutlich wird. Es ist naheliegend, bei Konvergenz zweiter Ordnung von einer Fehlerasymptotik der Gestalt

$$u_{h_1, h_2} = u + \gamma_1(h_1)h_1^2 + \gamma_2(h_2)h_2^2 + \gamma_{12}(h_1, h_2)h_1^2h_2^2$$

mit den zwei unabhängigen Gitterweiten h_1 und h_2 auszugehen. Eine solche Abschätzung muss von der verwendeten Diskretisierung am kartesischen Gitter kommen und wird z. B. von Griebel et al. für eine Finite-Differenzen-Diskretisierung der zweidimensionalen Laplace-Gleichung über Fourierreihen gezeigt [BGRZ94].

Betrachtet man die Gitterlösungen in Abb. 4.2 spaltenweise ($l = 0, \dots, n$ in Spalte n),

$$u_{2^{-l}, 2^{-n+l}} = u + \gamma_1(2^{-l})2^{-2l} + \gamma_2(2^{-n+l})2^{-2(n-l)} + \gamma_{12}(2^{-l}, 2^{-n+l})2^{-2n},$$

so fällt auf, dass der gemischte Term nur mehr das — auf all diesen Gittern gleiche — Produkt der Gitterweiten $h_1 h_2 = 2^{-n}$ enthält. Die Terme niedrigerer Ordnung können durch Summation

$$\begin{aligned} S_n &:= \sum_{l=0}^n u_{2^{-l}, 2^{-n+l}} \\ &= (n+1)u + \sum_{l=0}^n \gamma_1(2^{-l})2^{-2l} + \sum_{l=0}^n \gamma_2(2^{-l})2^{-2l} + 2^{-2n} \sum_{l=0}^n \gamma_{12}(2^{-l}, 2^{-n+l}) \end{aligned}$$

und einen anschließenden Extrapolationsschritt

$$\begin{aligned} u_n &:= S_{n+1} - S_n \\ &= u + \gamma_1(2^{-n-1})2^{-2n-2} + \gamma_2(2^{-n-1})2^{-2n-2} \\ &\quad + 2^{-2n-2} \sum_{l=0}^{n+1} \gamma_{12}(2^{-l}, 2^{-n-1+l}) - 2^{-2n} \sum_{l=0}^n \gamma_{12}(2^{-l}, 2^{-n+l}) \end{aligned} \tag{4.5}$$

4 Konvergenzabschätzungen für die Kombinationstechnik

zum Verschwinden gebracht werden.

Sind $\gamma_1, \gamma_2, \gamma_{12}$ sinnvollerweise betragsmäßig beschränkt, etwa durch K , gilt also

$$\begin{aligned} |u - u_n| &\leq K4^{-n-1} + K4^{-n-1} + K(n+2)4^{-n-1} + K(n+1)4^{-n} \\ &= K4^{-n} \left(2 + \frac{5}{4}n \right). \end{aligned} \quad (4.6)$$

Der allgemeine Fall

In einer allgemeinen Dimension d betrachten wir diskrete Lösungen auf einer Hierarchie von (anisotropen) kartesischen Gittern, charakterisiert durch die (zunächst äquidistanten) Gitterweiten $\mathbf{h} = (h_1, \dots, h_d)$ in den einzelnen Raumrichtungen. Die dazugehörige Gitterlösung $u_{\mathbf{h}}$ wird wechselweise auch als $u_{\mathbf{i}}$ mit Multiindex \mathbf{i} bezeichnet, wobei dann $h_k = 2^{-i_k}$ gemeint ist. Der Betrag für einen solchen Multiindex sei $|\mathbf{i}| := \sum_{k=1}^d i_k$.

Die Kombinationstechnik fußt dort ähnlich dem zweidimensionalen Fall auf der Annahme einer Fehlerasymptotik der Form

$$u - u_{\mathbf{h}} = \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^2 \cdot \dots \cdot h_{j_m}^2, \quad (4.7)$$

wobei die γ_{j_1, \dots, j_m} beschränkt sein sollen, d. h. $|\gamma_{j_1, \dots, j_m}| \leq K$.

Eine natürliche Erweiterung der Kombinationsformel (4.5) ist die iterierte Bildung von Differenzen benachbarter Spalten des Gittertableaus. Für eine Funktion natürlicher Zahlen $f : \mathbb{N}_0 \rightarrow \mathbb{R}$ sei also

$$\Delta_n f : n \rightarrow f(n+1) - f(n) \quad \forall n \in \mathbb{N}_0.$$

Dann definiert sich die Kombinationslösung auf Stufe n als

$$u_n := \Delta_n^{d-1} \sum_{|\mathbf{i}|=n} u_{\mathbf{i}}. \quad (4.8)$$

Iterierte Anwendung ergibt nach der binomischen Formel ($m \geq 0$)

$$\Delta_n^m f(n) = \sum_{j=0}^m (-1)^{(m-j)} \binom{m}{j} f(n+j) \quad \forall n \in \mathbb{N}_0$$

und somit läßt sich (4.8) schreiben als

$$u_n = \sum_{l=n}^{n+d-1} a_{l-n} \sum_{|\mathbf{i}|=l} u_{\mathbf{i}} \quad (4.9)$$

mit

$$a_i := (-1)^{d-1-i} \binom{d-1}{i} \quad 0 \leq i \leq d-1. \quad (4.10)$$

Griebel et al. zeigen in [GSZ92] unter Verwendung von (4.9), (4.10) und unter der Voraussetzung (4.7) punktweise Fehlerabschätzungen für den zweidimensionalen Fall,

$$|u - u_n| \leq K h_n^2 \left(1 + \frac{5}{4} \log_2 h_n^{-1} \right), \quad (4.11)$$

(vergleiche (4.6) mit einer unterschiedlichen Indizierung der Gitter) und den dreidimensionalen Fall,

$$|u - u_n| \leq Kh_n^2 \left(1 + \frac{65}{32} \text{ld } h_n^{-1} + \frac{25}{32} (\log_2 h_n^{-1})^2 \right). \quad (4.12)$$

Der dortige Zugang ist zwar prinzipiell übertragbar auf beliebige (feste) höhere Dimensionen, was angesichts der exponentiell ansteigenden Anzahl an Fehlertermen, die getrennt abgeschätzt werden müssen, allerdings schon in drei Dimensionen ein mühsames Unterfangen ist.

Daher ist es erfreulich, dass im folgenden Abschnitt eine allgemeine Abschätzung in geschlossener Form für beliebige Dimensionen angegeben werden kann.

4.3 Konvergenzabschätzungen

In diesem Abschnitt wird der folgende zentrale Satz, ein Pendant zu (4.11) und (4.12) in allgemeinen Dimensionen, hergeleitet.

Satz 4.1 (Konvergenz zweiter Ordnung). Für alle Lösungen $u_{\mathbf{h}}$ gelte punktweise eine Fehlerasymptotik der Form

$$u - u_{\mathbf{h}} = \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^2 \cdot \dots \cdot h_{j_m}^2, \quad (4.13)$$

wobei

$$|\gamma_{j_1, \dots, j_m}| \leq K \quad \forall 1 \leq m \leq d \quad \forall \{j_1, \dots, j_m\} \subset \{1, \dots, d\}.$$

Dann erfüllt die Kombinationslösung

$$u_n = \Delta_n^{d-1} \sum_{|\mathbf{i}|=n} u_{\mathbf{i}} \quad (4.14)$$

die Fehlerabschätzung

$$|u - u_n| \leq \frac{2K}{(d-1)!} \left(\frac{5}{2} \right)^{d-1} (n + 2(d-1))^{d-1} 4^{-n}. \quad (4.15)$$

Vor dem eigentlichen Beweis werden noch vorbereitende Resultate gebracht. Die folgende elementare Formel für Differenzen von Produkten — eine diskrete Form der Leibniz-Regel — wird dabei gelegentlich hilfreich sein.

Lemma 4.2. Seien $f, g : \mathbb{N}_0 \rightarrow \mathbb{R}$, $f_n := f(n)$, $g_n := g(n)$. Dann gilt

$$\Delta_n^k (f_n g_n) = \sum_{j=0}^k \binom{k}{j} \Delta_n^{k-j} f_{n+j} \Delta_n^j g_n \quad \forall k \in \mathbb{N}_0. \quad (4.16)$$

Beweis. Induktion nach k . $k = 0$ ist klar, für $k = 1$ gilt

$$\Delta_n (f_n g_n) = f_{n+1} g_{n+1} - f_n g_n = f_{n+1} (g_{n+1} - g_n) + (f_{n+1} - f_n) g_n = f_{n+1} \Delta_n g_n + \Delta_n f_n g_n. \quad (4.17)$$

4 Konvergenzabschätzungen für die Kombinationstechnik

Induktionsschritt $k \rightarrow k + 1$:

$$\begin{aligned}
\Delta_n^{k+1}(f_n g_n) &= \Delta_n \Delta_n^k(f_n g_n) \\
&= \sum_{j=0}^k \binom{k}{j} \Delta_n \left[\Delta_n^{k-j} f_{n+j} \Delta_n^j g_n \right] \\
(4.17) \quad &\stackrel{=}{=} \sum_{j=0}^k \binom{k}{j} \left[\Delta_n^{k-j} f_{n+j+1} \Delta_n^{j+1} g_n + \Delta_n^{k-j+1} f_{n+j} \Delta_n^j g_n \right] \\
&= \sum_{j=0}^k \binom{k}{j} \Delta_n^{k-j} f_{n+j+1} \Delta_n^{j+1} g_n + \sum_{j=0}^k \binom{k}{j} \Delta_n^{k-j+1} f_{n+j} \Delta_n^j g_n \\
&= \sum_{j=1}^{k+1} \binom{k}{j-1} \Delta_n^{k-j+1} f_{n+j} \Delta_n^j g_n + \sum_{j=0}^k \binom{k}{j} \Delta_n^{k-j+1} f_{n+j} \Delta_n^j g_n \\
&= \sum_{j=1}^k \left[\binom{k}{j-1} + \binom{k}{j} \right] \Delta_n^{k-j+1} f_{n+j} \Delta_n^j g_n + \binom{k}{0} \Delta_n^{k+1} f_n \Delta_n^0 g_n \\
&\quad + \binom{k}{k} \Delta_n^0 f_{n+k+1} \Delta_n^{k+1} g_n \\
&= \sum_{j=0}^{k+1} \binom{k+1}{j} \Delta_n^{k+1-j} f_{n+j} \Delta_n^j g_n
\end{aligned}$$

□

Da die Zahl der Gitter auf Stufe n in d Dimensionen durch (4.1) gegeben ist, liefert das folgende Lemma 4.3 (mit der Umbenennung $n + d - 1 \rightarrow n$) eine notwendige Bedingung für die Konsistenz der Kombinationsformel (d. h. die Summe der Koeffizienten ist 1):

Lemma 4.3 (Konsistenz). $\forall d \in \mathbb{N}, \forall n \in \mathbb{N}, n > d - 1$ gilt:

$$\Delta_n^{d-1} \binom{n}{d-1} = 1$$

Beweis. Für alle $n > d - 1, 0 \leq k \leq d - 1$ gilt

$$\Delta_n^k \binom{n}{d-1} = \binom{n}{d-k-1},$$

wie man induktiv ($k \rightarrow k + 1$) aus

$$\Delta_n^{k+1} \binom{n}{d-1} = \Delta_n \binom{n}{d-k-1} = \binom{n+1}{d-k-1} - \binom{n}{d-k-1} = \binom{n}{d-k-2}$$

sieht. Insbesondere ergibt sich

$$\Delta_n^{d-1} \binom{n}{d-1} = \binom{n}{0} = 1.$$

□

Für den Konvergenzbeweis ist noch ein technisches Hilfsresultat vonnöten, das zunächst etwas unmotiviert scheint, aber im Satz 4.1 an einer zentralen Stelle eingehen wird.

Grob gesprochen wird dort F_n die Fehlerentwicklung auf einer Spalte im Gittertableau sein, wobei δ_n dem Fehler einer bestimmten Ordnung ($\delta_l \sim s_l 2^{-l}$) und der binomische Faktor der Anzahl solcher Fehlerterme entspricht. Entscheidend ist der zweite Teil des Lemma, demnach in der Kombinationslösung nur mehr der Term höchster Ordnung übrigbleibt.

Lemma 4.4 (Darstellungsformel). *Seien $\delta_l \in \mathbb{R} \forall l \in \mathbb{N}_0$ beliebig,*

$$F_n := \sum_{l=0}^n \delta_l \binom{n-l+d-1}{d-1}.$$

1. Für $0 \leq k < d$ gilt

$$\Delta_n^k F_n = G_n^k + H_n^k,$$

wobei

$$G_n^k := \begin{cases} 0 & k = 0 \\ \sum_{j=1}^k \delta_{n+j} \binom{d-j-1}{k-j} & k \geq 1 \end{cases}$$

$$H_n^k := \sum_{l=0}^n \delta_l \binom{n-l+d-1}{d-k-1}$$

2.

$$\Delta_n^d F_n = \delta_{n+d}$$

Beweis. 1. Induktion nach k : Klarerweise ($k = 0$)

$$\Delta_n^0 F_n = F_n = H_n^0 = G_n^0 + H_n^0.$$

Da F_n von der Gestalt $\sum_{l=0}^n f(l, n)$ ist, gilt

$$\Delta_n F_n = f(n+1, n+1) - \sum_{l=0}^n (f(l, n+1) - f(l, n)),$$

also

$$\begin{aligned} \Delta_n F_n &= \delta_{n+1} + \sum_{l=0}^n \delta_l \left[\binom{n-l+d}{d-1} - \binom{n-l+d-1}{d-1} \right] \\ &= \delta_{n+1} + \sum_{l=0}^n \delta_l \binom{n-l+d-1}{d} \\ &= G_n^1 + H_n^1, \end{aligned}$$

der Fall $k = 1$.

4 Konvergenzabschätzungen für die Kombinationstechnik

Für $k > 1$ ergibt sich im Induktionsschritt

$$\begin{aligned}
 \Delta_n G_n^k &= \sum_{j=1}^k (\delta_{n+j+1} - \delta_{n+j}) \binom{d-j-1}{k-j} \\
 &= \delta_{n+k+1} - \delta_{n+1} \binom{d-2}{k-1} + \sum_{j=2}^k \delta_{n+j} \left[\binom{d-j}{k-j+1} - \binom{d-j-1}{k-j} \right] \\
 &= \sum_{j=2}^{k+1} \delta_{n+j} \binom{d-j-1}{k-j+1} - \delta_{n+1} \binom{d-2}{k-1} \\
 &= G_n^{k+1} - \delta_{n+1} \binom{d-2}{k} - \delta_{n+1} \binom{d-2}{k-1} \\
 &= G_n^{k+1} - \delta_{n+1} \binom{d-1}{k}.
 \end{aligned}$$

Für H_n^k gilt wiederum

$$\begin{aligned}
 \Delta_n H_n^k &= \delta_{n+1} \binom{d-1}{d-k-1} + \sum_{l=0}^n \delta_l \left[\binom{n-l+d}{d-k-1} - \binom{n-l+d-1}{d-k-1} \right] \\
 &= \delta_{n+1} \binom{d-1}{k} + \sum_{l=0}^n \delta_l \binom{n-l+d-1}{d-k-2} \\
 &= H_n^{k+1} + \delta_{n+1} \binom{d-1}{k}.
 \end{aligned}$$

Damit ($k \rightarrow k+1$) fñgt sich alles zusammen als

$$\Delta_n^{k+1} F_n = \Delta_n \Delta_n^k F_n = \Delta_n (G_n^k + H_n^k) = \Delta_n G_n^k + \Delta_n H_n^k = G_n^{k+1} + H_n^{k+1}.$$

2. Nach 1. ergibt sich für $k = d-1$

$$\Delta_n^{d-1} F_n = \sum_{j=1}^{d-1} \delta_{n+j} + \sum_{l=0}^n \delta_l$$

und damit

$$\Delta_n^d F_n = \sum_{j=1}^{d-1} (\delta_{n+j+1} - \delta_{n+j}) + \delta_{n+1} = \delta_{n+d}.$$

□

Beweis für Satz 4.1. Wegen Lemma 4.3 kann man schreiben

$$u - u_n = \Delta_n^{d-1} \sum_{|\mathbf{i}|=n} (u - u_{\mathbf{i}}).$$

Dabei ist

$$u - u_{\mathbf{i}} = \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m} (2^{-i_{j_1}}, \dots, 2^{-i_{j_m}}) 4^{-i_{j_1}} \dots 4^{-i_{j_m}}.$$

Wir betrachten nun getrennt die einzelnen Fehlerterme aus (4.13), o. B. d. A. sei für einen beliebigen Term mit m Faktoren $j_k = k$ für $1 \leq k \leq m$:

$$\begin{aligned}
 \sum_{|\mathbf{i}|=n} \gamma_{1,\dots,m}(h_{i_1}, \dots, h_{i_m}) h_{i_1}^2 \dots h_{i_m}^2 &= \sum_{|\mathbf{i}|=n} \gamma_{1,\dots,m}(2^{-i_1}, \dots, 2^{-i_m}) 4^{-\sum_{k=1}^m i_k} \\
 &= \sum_{l=0}^n \underbrace{\sum_{\sum_{k=1}^m i_k=l} \gamma_{1,\dots,m}(2^{-i_1}, \dots, 2^{-i_m})}_{=:s_l} 4^{-l} \sum_{\sum_{k=m+1}^d i_k=n-l} 1 \\
 &= \sum_{l=0}^n s_l 4^{-l} \binom{n-l+d-m-1}{d-m-1},
 \end{aligned}$$

wobei s_l eigentlich noch von m , ja sogar von der Gestalt des Fehlerterms $\gamma_{1,\dots,m}$ abhängt, was wir im Hinterkopf behalten. Aus Lemma 4.4, 2. ($d \rightarrow d-m$, $\delta_l \rightarrow s_l 4^{-l}$) folgt die zentrale Identität

$$\Delta_n^{d-m} \sum_{|\mathbf{i}|=n} \gamma_{1,\dots,m}(h_{i_1}, \dots, h_{i_m}) h_{i_1}^2 \dots h_{i_m}^2 = s_{n+d-m} 4^{-(n+d-m)}. \quad (4.18)$$

Außerdem gilt für alle $j \geq 0$

$$\Delta_n^j 4^{-n} = \left(-\frac{3}{4}\right)^j 4^{-n},$$

wie man sofort induktiv ($j \rightarrow j+1$) aus

$$\Delta_n^{j+1} 4^{-n} = \left(-\frac{3}{4}\right)^j (4^{-(n+1)} - 4^{-n}) = \left(-\frac{3}{4}\right)^{j+1} 4^{-n}$$

sieht. Damit folgt aus Lemma 4.2

$$\begin{aligned}
 \Delta_n^{m-1} [s_{n+d-m} 4^{-(n+d-m)}] &= \sum_{j=0}^{m-1} \binom{m-1}{j} \Delta_n^j s_{n+d-m} \Delta_n^{m-1-j} 4^{-(n+d-m+j)} \\
 &= \sum_{j=0}^{m-1} \binom{m-1}{j} \Delta_n^j s_{n+d-m} \left(\frac{-3}{4}\right)^{m-1-j} 4^{-(n+d-m+j)} \\
 &= 4^{-(n+d-1)} \sum_{j=0}^{m-1} \binom{m-1}{j} (-3)^{m-1-j} \Delta_n^j s_{n+d-m} \\
 &= 4^{-(n+d-1)} (\Delta_n - 3)^{m-1} s_{n+d-m} \\
 &= 4^{-(n+d-1)} \sum_{i=0}^{m-1} s_{n+d-m+i} \binom{m-1}{i} (-4)^{m-1-i} \\
 &= 4^{-(n+d-1)} \sum_{i=0}^{m-1} s_{n+d-i-1} \binom{m-1}{i} (-4)^i.
 \end{aligned}$$

Es sei bemerkt, dass bis zu diesem Punkt keine Abschätzungen getroffen wurden und daher

$$\Delta_n^{d-1} \sum_{|\mathbf{i}|=n} \gamma_{1,\dots,m}(h_{i_1}, \dots, h_{i_m}) h_{i_1}^2 \dots h_{i_m}^2 = 4^{-(n+d-1)} \sum_{i=0}^{m-1} s_{n+d-i-1} \binom{m-1}{i} (-4)^i$$

4 Konvergenzabschätzungen für die Kombinationstechnik

mit

$$s_l = \sum_{\sum_{k=1}^m i_k = l} \gamma_{1,\dots,m}(2^{-i_1}, \dots, 2^{-i_m})$$

eine exakte Darstellung der m -faktorigen Fehlerterme in der Kombinationslösung ist. Der Clou dabei ist, dass die Faktoren h_i hier nicht mehr getrennt auftreten, sondern nur mehr in der maximalen Ordnung $h_1 \cdot \dots \cdot h_n = 4^{-n}$, was die (zweite) Ordnung der Kombinationslösung sicherstellt. Entscheidend war dabei der Einsatz von Lemma 4.4 in (4.18), wonach sich die gemischten Fehlerterme niedriger Ordnung auf wundersame Weise wegheben.

In der Folge müssen noch die Fehlerkoeffizienten abgeschätzt werden, deren Summation auf die polynomialen Terme in n führt. Da aus $|\gamma_{1,\dots,m}| \leq K$

$$|s_l| \leq K \binom{l+m-1}{m-1} \Rightarrow \max_{i=0}^{m-1} |s_{n+d-i-1}| \leq K \binom{n+d+m-2}{m-1}$$

folgt und damit

$$\begin{aligned} |(\Delta_n - 3)^{m-1} s_{n+d-m}| &\leq \max_{i=0}^{m-1} |s_{n+d-i-1}| \sum_{i=0}^{m-1} \binom{m-1}{i} 4^i & (4.19) \\ &\leq K \binom{n+d+m-2}{m-1} 5^{m-1}, \end{aligned}$$

gilt insgesamt

$$\begin{aligned} \left| \Delta_n^{d-1} \sum_{|\mathbf{i}|=n} \gamma_{1,\dots,m}(h_{i_1}, \dots, h_{i_m}) h_{i_1}^2 \dots h_{i_m}^2 \right| &= 4^{-(n+d-1)} |(\Delta_n - 3)^m s_{n+d-m}| \\ &\leq 4^{-(n+d-1)} 5^{m-1} K \binom{n+d+m-2}{m-1} & (4.20) \end{aligned}$$

$$\leq 4^{-(n+d-1)} 5^{d-1} K \binom{n+2(d-1)}{d-1} \quad (4.21)$$

$$< 4^{-n} \left(\frac{5}{4}\right)^{d-1} \frac{K}{(d-1)!} (n+2(d-1))^{d-1}. \quad (4.22)$$

Da $2^d - 1$ Fehlerterme auftreten, ergibt sich zusammen

$$|u - u_n| < \left(\frac{5}{2}\right)^{d-1} \frac{2K}{(d-1)!} (n+2(d-1))^{d-1} 4^{-n}.$$

□

Betrachten wir zunächst die Aussage des Satzes 4.1 für kleine Dimensionen. Für $d = 1$, wo das dünne Gitter gleich dem vollen Gitter ist, reduziert sich die Abschätzung auf $|u - u_n| \leq 2K4^{-n}$ (die Ersetzung von $2^d - 1$ durch 2^d am Ende erklärt den unnötigen Faktor 2). Im Fall $d = 2$ ist der führende Fehlerterm durch (4.22) gegeben, sodass wir in der höchsten Potenz von n

$$|u - u_n| \sim K \frac{5}{4} n 4^{-n}$$

erhalten, was sich mit dem Resultat (4.11) deckt. Analog ergibt sich in $d = 3$ Dimensionen

$$|u - u_n| \sim K \frac{1}{2} \left(\frac{5}{4}\right)^2 4^{-n} = K \frac{25}{32} n^2 4^{-n}$$

in Übereinstimmung mit (4.12). Die Terme niedrigerer Ordnung differieren durch Unterschiede bei der Indizierung der Gitter. Ab (4.21) wurden diese Terme überdies zur Vereinfachung des Resultats nicht gesondert abgeschätzt.

Interessanter ist natürlich, wie sich (4.15) für große d verhält. Hält man n fest, so kann man die Fakultät nach der Stirling-Formel

$$k! \approx \sqrt{2\pi k} \left(\frac{k}{e}\right)^k$$

(siehe etwa [BS91] oder [Heu91] für einen Beweis) abschätzen und sieht, dass die dimensionsabhängigen Terme asymptotisch mit $(5e)^d / \sqrt{d}$ gegen Unendlich streben.

4.4 Verbesserte Konvergenzabschätzungen

Es drängt sich die Frage auf, ob man die Dimensionsabhängigkeit der Konstanten in (4.15), Satz 4.1, durch sorgfältigeres Abschätzen weiter drücken kann. Ausgangspunkt müssen die Ungleichungen (4.20) bis (4.22) sein. Da die s_l — indem sie von den unbekanntem Fehlerkoeffizienten abhängen — allgemein sind, werden sie dazu aus den Differenzen separiert.

Korollar 4.5 (zu Satz 4.1). 1. Unter den Voraussetzungen von Satz 4.1 gilt für $d \geq 2$ auch die etwas schärfere Fehlerabschätzung

$$|u - u_n| \leq 2K \left(\frac{5}{2}\right)^{d-1} \left(1 + (n + d - 1) \frac{1 + \ln(d-1)}{d-1}\right)^{d-1} 4^{-n}.$$

2. Darüberhinaus gilt die Abschätzung

$$|u - u_n| \leq \frac{K}{(d-1)!} \left(\frac{5}{4}\right)^{d-1} n^{d-1} 4^{-n} \quad \forall n \geq N \quad (4.23)$$

für hinreichend großes N .

3. Ist zusätzlich der Fehlerterm $\gamma_{1,\dots,d}$ stetig in $\mathbf{0} \in \mathbb{R}^d$ mit $\gamma_0 := \gamma_{1,\dots,d}(0, \dots, 0) \neq 0$, so kann sogar die asymptotische Näherung

$$|u - u_n| \sim \gamma_0 \left(\frac{3}{4}\right)^{d-1} \frac{n^{d-1}}{(d-1)!} 4^{-n} \quad \text{für } n \rightarrow \infty \quad (4.24)$$

gezeigt werden ($f(n) \sim g(n) :\Leftrightarrow \lim_{n \rightarrow \infty} f(n)/g(n) = 1$).

Beweis. 1. Ein Binomialkoeffizient der Gestalt von (4.20) bzw. (4.21) kann umgeschrieben werden ($l = n + d - 1$, $k = m - 1$) als

$$\binom{l+k}{k} = \frac{\prod_{j=1}^k (l+k)}{\prod_{j=1}^k k} = \prod_{j=1}^k \left(1 + \frac{l}{k}\right).$$

4 Konvergenzabschätzungen für die Kombinationstechnik

Nach der Ungleichung vom arithmetischen und geometrischen Mittel gilt

$$\sqrt[k]{\prod_{j=1}^k \left(1 + \frac{l}{k}\right)} < \frac{1}{k} \left[k + l \left(\sum_{j=1}^k \frac{1}{j} \right) \right],$$

weilers

$$\begin{aligned} \sum_{j=1}^k \frac{1}{j} &< 1 + \int_1^k \frac{dx}{x} \\ &= 1 + \ln k, \end{aligned}$$

weil Erstere eine Untersumme für das Integral ist. Insgesamt gibt das

$$\binom{l+k}{k} < \left[1 + l \frac{1 + \ln k}{k} \right]^k$$

bzw.

$$\binom{n+d+m-2}{m-1} < \left[1 + (n+d-1) \frac{1 + \ln(m-1)}{m-1} \right]^{m-1}.$$

Der Rest folgt analog.

2. Für hinreichend große n ist der Fehlerterm (4.21) für $m = d$ dominant und mit

$$\lim_{n \rightarrow \infty} \binom{n+2d-2}{d-1} / n^{d-1} = \frac{1}{(d-1)!}$$

folgt die Behauptung, da es nur einen d -faktorigen Fehlerterm gibt.

3. Ist keine Information über s_l verfügbar, kann man (4.19) nur mit dem Absolutbetrag der einzelnen Summanden abschätzen und handelt sich den Faktor 5^d von Satz 4.1 ein. Wenn die s_l hingegen eine gemeinsame Asymptotik haben, können die wechselnden Vorzeichen in der binomischen Summe ausgenutzt werden. In der Tat gilt unter der Stetigkeitsannahme für $\gamma_{1,\dots,m}$

$$\lim_{n \rightarrow \infty} \frac{s_n}{\binom{n+m-1}{m-1}} = \gamma_{1,\dots,m}(0, \dots, 0) =: \gamma_0. \quad (4.25)$$

Das gilt es zunächst zu zeigen (wenn keine Verwechslung möglich ist, wird der Index von $\gamma_{1,\dots,m}$ weggelassen, also ab jetzt schlicht γ). Sei dazu $\epsilon > 0$ beliebig und

$$c := \sup_{0 \leq h_1, \dots, h_m \leq 1} |\gamma(h_1, \dots, h_m) - \gamma_0|.$$

Wegen der Stetigkeit von γ kann n_0 so gewählt werden, dass

$$\forall i_k \geq n_0, 1 \leq k \leq m : |\gamma(2^{-i_1}, \dots, 2^{-i_m}) - \gamma_0| \leq \frac{\epsilon}{2}$$

und danach n so groß, dass

$$N_0 := \binom{n - kn_0 + m - 1}{m-1} \geq \left(1 - \frac{\epsilon}{c}\right) \binom{n+m-1}{m-1} =: \left(1 - \frac{\epsilon}{c}\right) N.$$

Letzteres geht, weil

$$\binom{n - kn_0 + m - 1}{m - 1} / \binom{n + m - 1}{m - 1} = \prod_{j=1}^{m-1} \left(1 - \frac{kn_0}{n + j}\right) \rightarrow 1 \text{ für } n \rightarrow \infty.$$

Die Idee ist jetzt zu zeigen, dass der Beitrag der Summanden, die nicht in einer ϵ -Umgebung von γ_0 liegen, vernachlässigt werden kann. Das legt eine Aufspaltung in

$$s_n = \sum_{\substack{\sum_k i_k = n \\ \forall k : i_k \geq n_0}} \gamma(2^{-i_1}, \dots, 2^{-i_m}) + \sum_{\substack{\sum_k i_k = n \\ \exists k : i_k < n_0}} \gamma(2^{-i_1}, \dots, 2^{-i_m})$$

nahe. Wegen

$$\sum_{\substack{\sum_k i_k = n \\ \forall k : i_k \geq n_0}} 1 = \sum_{\substack{\sum_k i_k = n - mn_0 \\ \forall k : i_k \geq 0}} 1 = N_0$$

gilt schließlich

$$s_n - \gamma_0 N = \sum_{\substack{\sum_k i_k = n \\ \forall k : i_k \geq n_0}} (\gamma(2^{-i_1}, \dots, 2^{-i_m}) - \gamma_0) + \sum_{\substack{\sum_k i_k = n \\ \exists k : i_k < n_0}} (\gamma(2^{-i_1}, \dots, 2^{-i_m}) - \gamma_0)$$

und damit

$$\begin{aligned} |s_n - \gamma_0 N| &\leq \frac{\epsilon}{2} N + c(N - N_0) \\ &\leq \frac{\epsilon}{2} N + c \frac{\epsilon}{2c} N \\ &= \epsilon N. \end{aligned}$$

Division durch N liefert die Behauptung (4.25).

Daher gilt asymptotisch

$$(\Delta_n - 3)^{m-1} s_{n+d-m} \sim \gamma_0 N \sum_{i=0}^{m-1} \binom{m-1}{i} (-4)^{m-1-i} = \gamma_0 \binom{n+m-1}{m-1} (-3)^{m-1}$$

bzw. exakter

$$\forall \epsilon > 0 \exists N \forall n \geq N : |(\Delta_n - 3)^{m-1} s_{n+d-m}| \leq \gamma_0 \binom{n+m-1}{m-1} 3^{m-1} (1 + \epsilon). \quad (4.26)$$

Letzteres gilt nicht für $\gamma_0 = 0$. In diesem Fall, der ungünstig wechselnde Vorzeichen der s_l nicht ausschließt, gilt aber die stärkere Aussage

$$\forall \epsilon > 0 \exists N \forall n \geq N : |(\Delta_n - 3)^{m-1} s_{n+d-m}| \leq \epsilon \binom{n+m-1}{m-1} 5^{m-1},$$

die bzgl. diesem Term auf höhere Konvergenzordnung führt und hier nicht weiter betrachtet wird.

In allen anderen Fällen ist asymptotisch der Fehlerterm für $m = d$ entscheidend, was die Aussage liefert.

□

4 Konvergenzabschätzungen für die Kombinationstechnik

In Anbetracht von (4.23) und (4.24) stellt man fest, dass die Vorfaktoren von 4^{-n} mit $d \rightarrow \infty$ gegen 0 streben. Das ist aber wohlgemerkt nur eine asymptotische Aussage, die unter anderem dadurch erreicht wurde, dass $(n + d - 1)/n \rightarrow 1$, also n insbesondere im Vergleich zu d groß ist, sodass man in höheren Dimensionen immer später in diesen asymptotischen Bereich kommt.

In dieser Verschiebung um $d - 1$ versteckt sich auch die Tatsache, dass die Kombinationslösung eine linear (mit d) ansteigende Zahl an Gitterstufen enthält.

5 Anwendung auf Optionspreisaufgaben

Überblick

Es sollen jetzt die vorangegangenen theoretischen Resultate an ausgewählten Optionspreisaufgaben verschiedener Dimension quantitativ betrachtet werden. Im Vordergrund stehen dabei einerseits (in 5.2) der Effizienzgewinn im Vergleich zu vollen Gittern für die Fälle, in denen dort eine Lösung überhaupt möglich ist, sowie andererseits (in 5.3) die maximalen Dimensionen, die bei paralleler Rechnung auf aktuellen Rechnerarchitekturen noch eine sinnvolle Approximation auf dünnen Gittern erlauben. Die Fehlerasymptotik für diese Beispiele wird in 5.5 quantitativ anhand von praxisrelevanten Beispielen untersucht. Dabei tun sich naturgemäß auch Fragen zur optimalen Lastverteilung auf, die in 5.4 beantwortet werden.

5.1 Fehleranalyse

Die Fehleranalyse im Hinblick auf Gitterkonvergenz erweitert die Vorgehensweise für volle Gitter um die Schätzung der logarithmischen Faktoren. Für eine Fehlerdarstellung der Form

$$u - u_h = c(h)(\log_2 h)^q h^p \quad (5.1)$$

(mit stetigem c , $c(h) \leq \bar{c}$, $c(0) \neq 0$) gilt dennoch

$$\lim_{h \rightarrow 0} \frac{u_{4h} - u_{2h}}{u_{2h} - u_h} = \lim_{h \rightarrow 0} 2^p \frac{2^p c(4h)(\log_2 h + 2)^q - c(2h)(\log_2 h + 1)^q}{2^p c(2h)(\log_2 h + 1)^q - c(h)(\log_2 h)^q} = 2^p$$

und daraus kann wie in (2.3) p als Grenzwert von

$$o_n := \log_2 \left| \frac{u_n - u_{n-1}}{u_{n+1} - u_n} \right| \quad (5.2)$$

bestimmt werden. Aus (5.1) folgt weiters

$$\begin{aligned} u - u_h &= c(0)(\log_2 h)^q h^p + \mathcal{O}(|\log_2 h|^q h^{p+1} + |\log_2 h|^{q-1} h^p) \\ u_{2h} - u_h &= c(0)(2^p - 1)(\log_2 h)^q h^p + \mathcal{O}(|\log_2 h|^q h^{p+1} + |\log_2 h|^{q-1} h^p) \end{aligned}$$

und damit

$$\log_2 |u_{2h} - u_h| \sim a + q \log_2 |\log_2 h| + p \log_2 h$$

bzw. konkret für eine Folge von Gitterlösungen u_n

$$\log_2 |u_n - u_{n+1}| \sim a + q \log_2 n - pn, \quad (5.3)$$

$$\log_2 |u - u_n| \sim a + q \log_2 n - pn. \quad (5.4)$$

Aus den logarithmischen Differenzen können also die Ordnungen p und q durch lineare Regression mit den Ansatzfunktionen 1 , n und $\log_2 n$ geschätzt werden. Eine genaue Identifizierung von q aus den numerischen Daten ist allerdings schwierig, da die logarithmische Singularität schwächer ist als ein Faktor $h^{-\epsilon}$ für jedes $\epsilon > 0$.

5.2 Der Vergleich mit vollen Gittern

Wir kehren zurück zu den numerischen Beispielen aus 2.1 und 2.2 und vergleichen die Konvergenz auf dünnen Gittern mit den dortigen Resultaten.

5.2.1 Ein 3-Faktor-Modell für die FX-Option

Für die FX-Option aus 2.1.3 ist der punktweise Fehler in Abb. 5.1 im Vergleich zum vollen Gitter dargestellt. Im Sinne von (5.3) stellt sich als Ordnung $p = 1$, $q = -3.8$ anstatt der erwarteten $p = 2$, $q = d - 1 = 2$ heraus. Vernachlässigt man den logarithmischen Term, liefert Regression $p = 1.81$, für die letzten 6 Werte alleine $p = 1.99$. Dies zeigt, wie sensitiv der geschätzte Exponent des Logarithmus von den Daten abhängt.

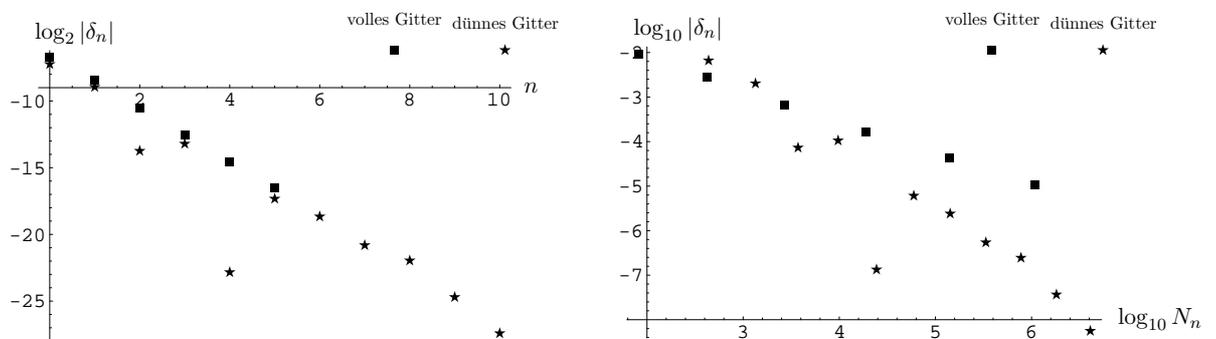


Abbildung 5.1: Links: Fehler δ_n für das dünne Gitter im Vergleich zum vollen Gitter auf gleicher Stufe n . Rechts: Der Vergleich des Fehlers in Abhängigkeit der Zahl der Gitterpunkte N_n zeigt die überlegene ϵ -Komplexität.

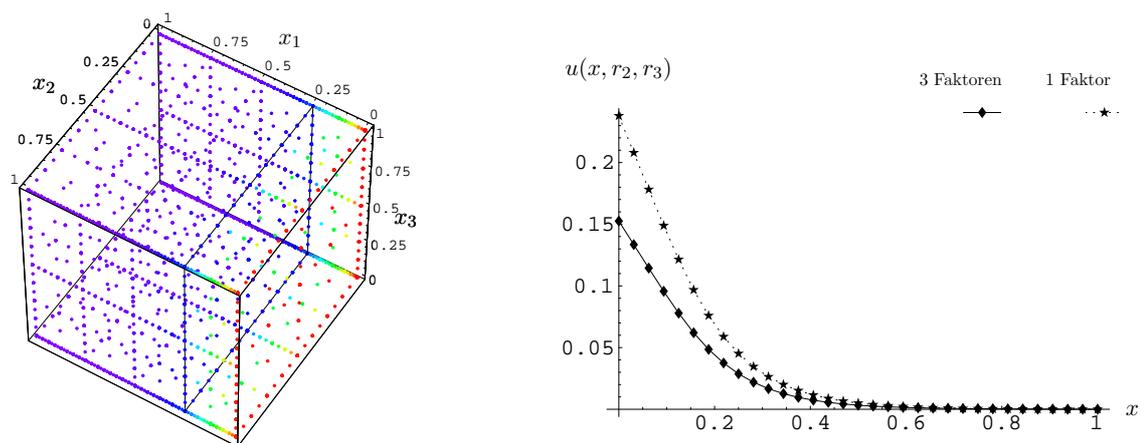


Abbildung 5.2: Links: Dünnes Lösungsgitter für das 3-Faktor-Modell für die FX-Option mit 3726 Freiheitsgraden. Die Punkte sind nach der Lösung eingefärbt. Rechts: Vergleich der Lösung an der Mittellinie mit dem eingeschränkten eindimensionalen Modell aus (1.17).

5.2.2 Basket-Option mit zwei Komponenten — gradierte dünne Gitter

In Abbildung 5.3 ist die gradierte Dünngitter-Lösung für die zweidimensionale Basket-Option aus 2.1.3 dargestellt, welche eine vergleichbare Genauigkeit mit der numerischen Lösung am gradierten vollen Gitter aus Abschnitt 2.2 aufweist.

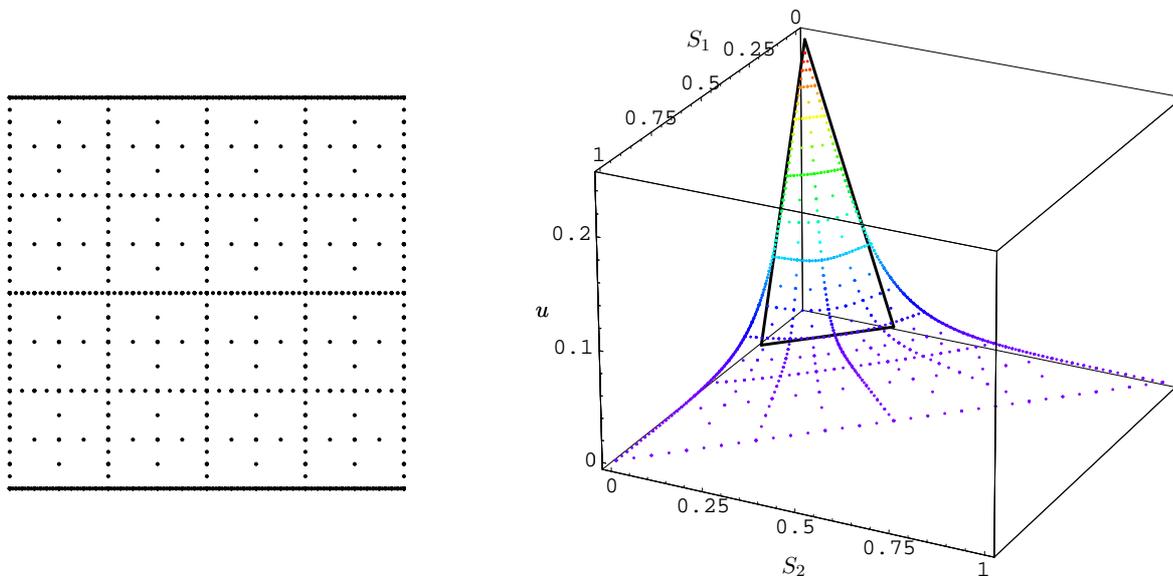


Abbildung 5.3: Links: Das dünne Berechnungsgitter mit zwei zusätzlichen Verfeinerungen in Richtung x_1 und 687 Punkten. Rechts: Die Lösung u am transformierten gradierten Gitter in Abhängigkeit von den Ursprungskordinaten S_1 und S_2 .

Abbildung 5.4 vergleicht den Fehler am gradierten vollen und dünnen Gitter über mehrere Verfeinerungsstufen hinweg. Die geschätzte Ordnung (bei Vernachlässigung des logarithmischen Faktors) ist 1.84.

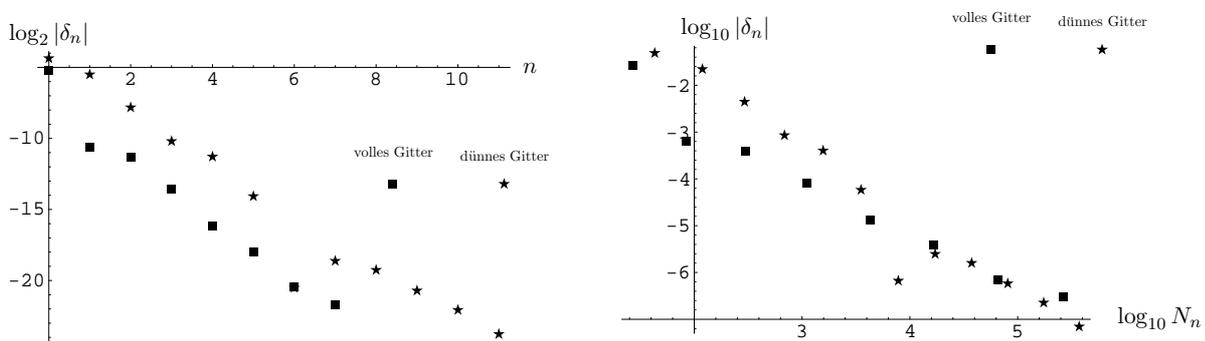


Abbildung 5.4: Links: Fehler δ_n für das gradierte dünne Gitter im Vergleich zum gradierten vollen Gitter auf gleicher Stufe n . Rechts: Vergleich des Fehlers in Abhängigkeit der Zahl der Gitterpunkte N_n . Die überlegene ϵ -Komplexität zeigt sich erst im Grenzbereich.

5.3 Komplexitätsbetrachtungen

Der vorangegangene Abschnitt zeigt, dass der „Fluch der Dimension“ gegenüber einem kartesischen Gitter stark abgeschwächt ist. Obwohl die Größe des dünnen Gitters bis auf logarithmische Faktoren einem eindimensionalen entspricht, ist die Zahl der Freiheitsgrade schon in moderaten Dimensionen dennoch enorm. In Abbildung 5.5 sind die Dünngitterpunkte für fünf Dimensionen abgezählt und im Vergleich zum eindimensionalen Fall gezeichnet. Asymptotisch

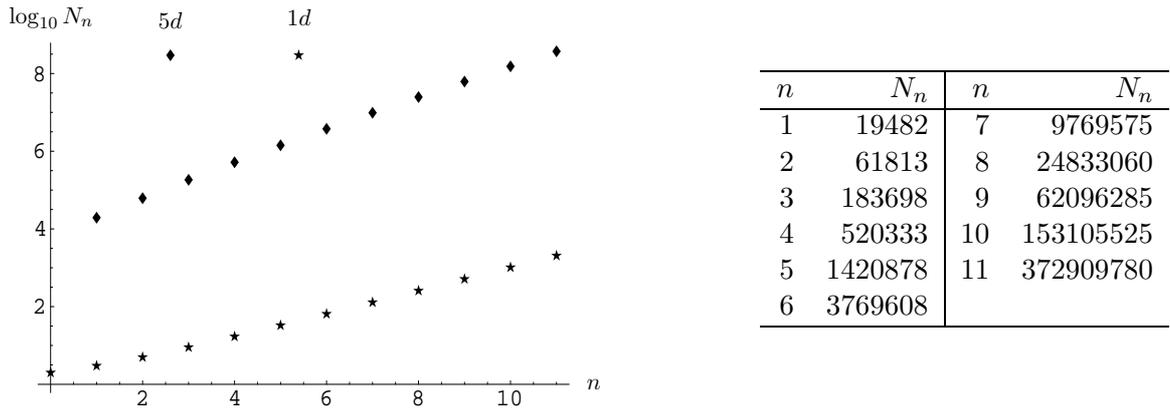


Abbildung 5.5: Anzahl der Freiheitsgrade N_n auf der n -ten Stufe einer Hierarchie fünfdimensionaler dünner Gitter (**rechts**) und der Vergleich zum eindimensionalen Fall (**links**).

verdoppeln sich die Freiheitsgrade bei Verfeinerung zwar nur, allerdings ist der Vorfaktor, der von den größten Gittern herrührt, ausgesprochen groß. Dieses Verhalten geht auch aus der Verschiebung in der logarithmischen Graphik in Abb. 5.5 deutlich hervor.

Um die behandelbare Problemgröße (bzw. -dimension) abzuschätzen, betrachten wir in den einzelnen Dimensionen die größten Gitter, die zumindest einen inneren Punkt besitzen, was auf der d -ten Stufe der Fall ist. Nach (4.1) müssen dann Lösungen auf

$$\binom{2d-1}{d-1}$$

Gittern berechnet werden, deren größtes — wie man sich leicht überlegt — durch d -malige Verfeinerung in eine einzige Richtung aus dem Ausgangsgitter mit 2^d Punkten hervorgeht, also $(2^d + 1)2^{d-1}$ Punkte besitzt. Das kleinste Gitter auf dieser Stufe wäre übrigens das (einzige) mit dem inneren Punkt (also drei in jede Richtung) mit 3^d Punkten insgesamt. Tabelle 5.1

d	1	2	3	4	5	6	7	8	9	10
ν_d	1	3	10	35	126	462	1716	6435	24310	92378
m_d	3	10	36	136	528	2080	8256	32896	131328	524800

Tabelle 5.1: Anzahl der Gitter ν_d auf der größten (= d -ten) Stufe mit innerem Punkt sowie die maximale Punktezahl m_d eines Gitters auf dieser Stufe in d Dimensionen.

verdeutlicht dies. Dazu kommen natürlich noch die Beiträge der größeren Gitter für die Kombinationslösung, die asymptotisch insgesamt etwa noch einmal soviel ausmachen wie die Zahl der Freiheitsgrade von der höchsten Stufe. Für Knotenfreiheitsgrade mit Dirichlet-Randdaten

sind diese Gitter die größten, die die Differentialgleichung überhaupt berücksichtigen — auf größeren Stufen liegen alle Punkte am Rand und die Lösung wird nur aus den Randbedingungen interpoliert.

Darüberhinaus ist anzumerken, dass die Zahl der Nicht-Null-Elemente pro Zeile im Allgemeinen mit 3^d exponentiell in der Raumdimension ansteigt. Man führe sich also vor Augen, dass z. B. schon in sieben Dimensionen pro Zeile bis zu $3^7 = 2187$ Einträge stehen. Die größte Dünngitterstufe, die gerade einmal einen inneren Punkt hat (siehe oben), fordert schon Gitter mit bis zu $2^6(2^6 + 1) = 3365$ Elementen, sodass in der Matrix mehr als 6 Millionen Werte abgespeichert werden müssen. Dies relativiert den in zwei oder drei Dimensionen intuitiven Begriff von „dünn Besetztheit“. Pro weiterer Dimension erhöht sich diese Zahl um ca. einen Faktor 12. Trotz der dramatischen Reduktion der Freiheitsgrade durch die Kombinationstechnik stoßen wir schon in moderaten Dimensionen an eine Grenze.

Für die parallel zu führende Diskussion der Fehlerentwicklung siehe die Kommentare in und um Korollar 4.5.

Dies alles führt zu der Einsicht, dass die erreichbare Grenze — trotz der zweifellos dramatischen Reduktion der Freiheitsgrade durch die Kombinationstechnik — unter zehn Dimensionen liegen sollte und schon in moderaten Dimensionen Parallelisierung unabdingbar sein wird.

5.4 Aspekte der parallelen Lösung

Ein zweiter Blick offenbart, wie denn die große Anzahl der Freiheitsgrade zustandekommt. Tabelle 5.2 gibt — wieder für fünf Dimensionen — eine Statistik über die Punkte der zu kombinierenden Gitter. Die Gesamtzahl der Unbekannten kommt (bedingt durch die Kombi-

n	M_n	m_n	ν_n	n	M_n	m_n	ν_n	n	N_n	m_n	ν_n
1	32	32	1	6	42363	528	126	11	6042330	16400	1001
2	240	48	5	7	122125	1040	210	12	15185610	32784	1365
3	1120	80	15	8	337755	2064	330	13	37600980	65552	1820
4	4200	144	35	9	904745	4112	495	14	91913985	131088	2380
5	13890	272	70	10	2362620	8208	715	15	222166875	262160	3060

Tabelle 5.2: Gesamtzahl der Freiheitsgrade M_n , maximale Zahl der Punkte auf einem einzelnen Gitter m_n und die Anzahl ν_n der involvierten Gitter auf Stufe n für fünf Dimensionen.

nationsformel) durch Addition von jeweils fünf Stufen zustande.

Fazit ist, dass sich das dünne Gitter aus einer großen Anzahl relativ kleiner kartesischer Gitter zusammensetzt. Da die Zahl der Gitter insbesondere größer ist als die der Prozessoren, die in der Regel zur Verfügung stehen werden, legt dies eine einfache Form der Parallelisierung nahe: die Gitter werden nach einem bestimmten Schlüssel auf die Prozessoren verteilt, dort die entsprechenden Gleichungen diskretisiert und gelöst, und die einzige Kommunikation ist für das Sammeln der Teilergebnisse zur Auswertung der Kombinationsformel für die interessierenden Punkte nötig.

Dies wirft die Frage einer gleichmäßigen Lastverteilung auf, die in diesem Zusammenhang darin besteht, die ganze Schar an kartesischen Gitter so auf die Prozessoren zu verteilen, dass alle möglichst gleichzeitig mit der Lösung fertig werden. Diese Aufgabe wird dadurch erschwert, dass die einzelnen Gleichungen, selbst wenn sich die Zahl der Unbekannten recht gleichmäßig verteilen lässt, durch die extremen Anisotropien der Gitter starke Unterschiede im

Lösungsverhalten aufweisen. Vom linearen (Mehrgitter-)Löser wird über die standardmäßige Bedeutung von linearer Komplexität hinaus verlangt, dass diese Skalierungseigenschaft über die gesamte Problemklasse an anisotropen kartesischen Gittern gleichmäßig erfüllt ist. Es sei bemerkt (siehe schon die Bemerkungen im Kapitel 3 zur iterativen Lösung), dass selbst die relativ kleinen Probleme in höheren Dimensionen sehr schlecht konditioniert sein können, wenn in wenige Richtungen (eventuell nur eine) sehr kleine Gitterweiten auftreten. Da solche Gitter immer vorhanden sind, darf nicht der Fall eintreten, wo ein oder wenige Prozessoren deutlich länger mit der Lösung beschäftigt sind als andere.

Gelingt es, einen derart robusten Löser zu konstruieren, kann man bei einer geeigneten Verteilung auf statistische Hilfe hoffen, sofern die Zahl der Gitter pro Prozessor nicht zu klein ist. Die Verteilung der Rechenzeiten auf 200 Prozessoren für das im nächsten Abschnitt diskutierte (fünfdimensionale) Beispiel der Black-Scholes-Gleichung in Hauptachsen, gelöst mit Mehrgitter mit Semi-Coarsening auf 372 909 780 Punkten, zeigt Abbildung 5.6. Die Gitter wurden ihrer standardmäßigen Numerierung folgend — also im Prinzip zufällig — auf die Prozessoren verteilt. Durch die Beschaffenheit des Abzählens war nur sichergestellt, dass die Differenz der Freiheitsgrade auf den Prozessoren in der Größenordnung der maximalen Gittergröße bleibt. Für den *Speed-Up* entscheidend ist das Verhältnis der längsten Rechenzeit eines

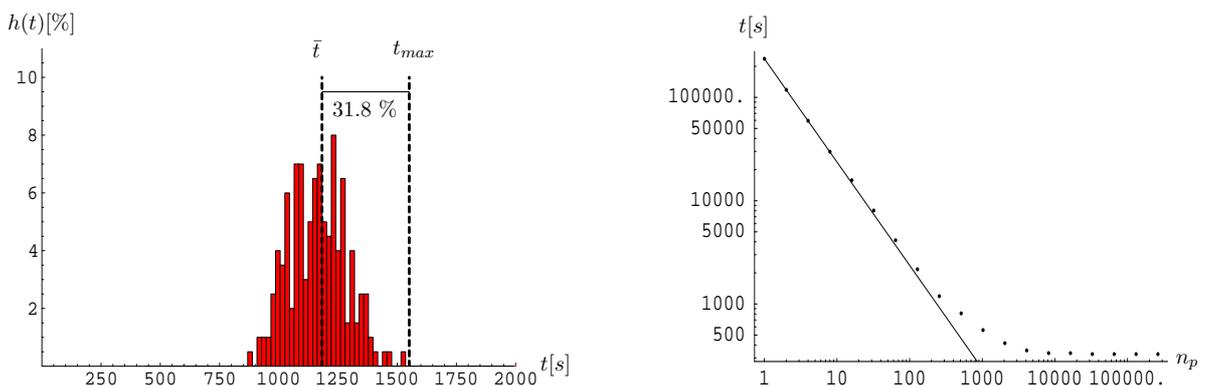


Abbildung 5.6: Verteilungsdichte $h(t)$ der Rechenzeiten t bei 200 Prozessoren mit dem Mittelwert \bar{t} (**links**) und die Skalierung bei steigender Anzahl n_p der Prozessoren (**rechts**, siehe auch Tab. 5.3).

Einzelprozessors, welche die parallele Lösungszeit bestimmt, zum Mittelwert der Einzelzeiten, die einer idealen Skalierung des sequentiellen Algorithmus entspricht. Hier ist dieser Faktor 1.32 (siehe Abb. 5.6, links).

Zum Vergleich wurden die Gitter (fiktiv) dermaßen auf die Prozessoren verteilt, dass der Reihe nach immer der mit den bislang wenigsten Freiheitsgraden das nächste zugeteilt bekam. Das Verhältnis der längsten zur durchschnittlichen Rechenzeit ist dann sogar 1.33, wobei die Zahl der Unbekannten auf den Prozessoren um maximal 10% abweicht (im Vergleich zu 17% bei der primitiven Verteilung).

Die Ungleichverteilung der Rechenzeiten hat mehrere Gründe. Zum einen ist selbst bei diesem einfachen Modellproblem die „ h -Unabhängigkeit“ asymptotisch, aber im betrachteten Bereich eben nicht exakt erfüllt, sodass für die gleiche Fehlerreduktion die Zahl der benötigten Zyklen nicht immer identisch ist. Zum anderen wächst der Aufwand pro Zyklus zwar grundsätzlich linear in der Anzahl der Unbekannten, aber nicht gleichmäßig über die ganze Klasse an Gittern hinweg, da z. B. Semi-Coarsening unterschiedlich große Grobgitterprobleme

5 Anwendung auf Optionspreisaufgaben

erzeugt. Demzufolge ist bei einem der extrem anisotropen Gitter der Aufwand pro Iteration proportional zu $N \cdot (1 + 1/2 + 1/4 + \dots) \approx 2N$, während er bei einem isotropen Gitter nur $N \cdot (1 + 1/4 + 1/16 + \dots) \approx 3/2N$ ausmacht. Doch selbst eine pedantische Buchhaltung und zahlenmäßig identische Verteilung aller arithmetischen Operationen könnte durch die unvermeidliche Variation in den Raten keine bessere Lastverteilung garantieren.

Tabelle 5.3 listet quantitativ die in Abb. 5.6 dargestellten Daten zur Skalierung der parallelen Lösungszeit mit der Anzahl der verwendeten Prozessoren, die durch Kombination der Rechenzeiten für die Einzelgitter ermittelt wurden. Dieses Vorgehen ist gerechtfertigt, da die Kommunikation zwischen den Prozessoren vernachlässigt werden kann. Ist die Zahl der Prozessoren verhältnismäßig (d. h. im Vergleich zu der Zahl der betrachteten Gitter) klein, skaliert das Verfahren nahezu perfekt, erst bei einer Prozessorzahl in der gleichen Größenordnung wie der der Gitter geht naturgemäß Effizienz verloren und die minimale Zeit ist klarerweise beschränkt durch die längste Lösungszeit eines Einzelgitters. In solchen Fällen könnte man die einzelnen Berechnungen weiter parallelisieren, was für die von uns betrachteten praktischen Fälle jedoch nicht relevant ist.

$\log_2 n_p$	1	2	3	4	5	6	7	
	8	9	10	11	12	13	14,15	≥ 16
$t[s]$	2.37 e+5	11.8 e+5	5.94 e+4	2.98 e+4	1.58 e+4	8.04 e+3	4.14 e+3	
	2.17 e+3	1.19 e+3	8.11 e+2	5.59 e+2	4.18 e+2	3.55 e+2	3.34 e+2	3.27 e+2

Tabelle 5.3: Skalierung der Rechenzeit t bei steigender Anzahl n_p an Prozessoren

Die parallelen Testrechnungen wurden am HEidelberg LInux Cluster System (HELICS, <http://helics.iwr.uni-heidelberg.de>) durchgeführt. Die technischen Daten sind wie folgt:

- Knoten
 - 256 commodity of the shelf PC
 - Dual AMD Athlon 1,4 Ghz
 - 2 GB RAM
 - Myrinet NIC M3F-PCI64B-2 (66 Mhz / 64 bit PCI)
- Netz
 - Myrinet 2 GBit, full bisection, Clos
 - Fibre optic
 - Software layer: GM
 - 100 MBit + 10 MBit Ethernet
- Software
 - Debian Woody Kernel 2.4.21
 - MPI: mpich-gm von Myricom
 - Compiler: g++

Eine Überschlagsrechnung gemäß Tabelle 5.1 ergibt ca. 200 Mio. Punkte für das achtdimensionale Grobgitter. Es stellt sich heraus, dass dieses Problem das größte ist, das noch in den Speicher passt, wobei die kritische Größe der Speicherbedarf für die Matrizen der Einzelprobleme ist. Eine sinnvolle Approximation war in acht Dimensionen jedoch nicht mehr möglich.

Zur Beobachtung der Gitterkonvergenz betrachten wir daher ein fünfdimensionales Beispiel.

5.5 Höhere Dimensionen

Zur Problematik der steigenden Anzahl an Gitterpunkten tritt bei zunehmender Dimension die wachsende Anzahl derer Nachbarn. Insbesondere bei der Diskretisierung mit finiten Elementen (bzw. bei gemischten Ableitungen auch für finite Differenzen) führt dies zu einer exponentiell anwachsenden Zahl an Nicht-Null-Einträgen in den Zeilen der „dünn besetzten“ Systemmatrix mit dem daraus folgenden Speicher- bzw. Rechenzeitaufwand. Im Fall von diagonaler Diffusion kann das durch finite Differenzen (oder auch durch finite Volumen) vermieden werden. Tabelle 5.4 führt den eklatanten Unterschied für $d = 1, \dots, 10$ vor Augen.

d	1	2	3	4	5	6	7	8	9	10
FE	3	9	27	81	243	729	2187	6561	19683	59049
FD	3	5	7	9	11	13	15	17	19	21

Tabelle 5.4: Anzahl der (Nicht-Null-)Elemente pro Matrixzeile bei Diskretisierung des d -dimensionalen Laplace-Operators mit finiten Elementen bzw. finiten Differenzen.

5.5.1 Konstruktion der Testbeispiele

Diese Beobachtung überzeugt schnell, für die folgenden numerischen Tests in höheren Dimensionen die transformierte Gleichung (1.21),

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^d \lambda_i \frac{\partial^2 u}{\partial x_i^2} - ru \quad \forall (\mathbf{x}, t) \in \mathbb{R}^d \times (0, T)$$

mit Anfangsbedingung

$$u(\mathbf{x}, 0) = \left(K - \sum_{i=1}^d \mu_i e^{\sum_{j=1}^d q_{ji} x_j} \right)_+ \quad \forall \mathbf{x} \in \mathbb{R}^d \quad (5.5)$$

heranzuziehen.

Da die Gleichung (mit Gitter-basierten Methoden) nur auf einem endlichen Gebiet gelöst werden kann, stellt sich die Abschneideproblematik. Aus praktischen Gründen (hauptsächlich vorgegeben durch den Tensorproduktansatz der Diskretisierung) wird in einem Würfel

$$W = [\underline{x}_1, \bar{x}_1] \times \dots \times [\underline{x}_d, \bar{x}_d]$$

mit hinreichend kleinen bzw. großen \underline{x}_i bzw. \bar{x}_i gerechnet, auf dem asymptotische Randbedingungen gesetzt werden müssen. Da, wie schon erwähnt, insbesondere in höheren Dimensionen viele der Dünngitterpunkte auf dem Rand liegen, sind Dirichlet-Randdaten ungünstig, da ein großer Teil der Freiheitsgrade in die Interpolation der Lösung aus den — hier noch dazu künstlichen und damit falschen — Randwerten investiert wird. Aus diesem Grund werden homogene Neumann-Randbedingungen verwendet, die bei der Diskretisierung mit finiten Differenzen am Rand nicht direkt gesetzt, sondern in die Differenzengleichung zur Elimination der fiktiven Punkte außerhalb des Gebiets verwendet werden, nach der Art (in einer Dimension)

$$\frac{u_{-1} - u_1}{h} = 0 \quad \rightsquigarrow \quad \frac{u_{-1} - 2u_0 + u_1}{h^2} = \frac{-2u_0 + 2u_1}{h^2} \quad \text{am Randpunkt 0 etc.}$$

5 Anwendung auf Optionspreisaufgaben

Die Verwendung von Neumann-Randbedingungen ergibt hier Sinn, da x_1 die Hauptkomponente ist, die annähernd dem Basket entspricht und damit näherungsweise die Anfangsbedingung beschreibt, sodass u in x_2, \dots, x_d nur schwach variiert. Aus (5.5) sieht man weiters, dass wegen der positiven Korreliertheit $q_{1i} > 0$,

$$\lim_{x_1 \rightarrow \pm\infty} \frac{\partial}{\partial x_1} u(\mathbf{x}, 0) = 0.$$

Durch dieses Vorgehen wird am Rand die Differentialgleichung tangenzial richtig diskretisiert.

Ein anderer Zugang wäre die Abbildung auf ein endliches Gebiet. Da die Koeffizienten der Gleichung durch das Wesen der Transformation am neuen (endlichen) Rand verschwinden, müssen bzw. können dort keine Randbedingungen spezifiziert werden. Die Wahl der Transformation bestimmt implizit eine Gradierung des Gitters in den ursprünglichen Koordinaten. Dieser Weg wird für die numerischen Rechnungen in Kapitel 7 eingeschlagen und hat den Vorteil, dass zum Diskretisierungsfehler nicht noch zusätzlich der Abschneidefehler kontrolliert werden muss.

Als Testbeispiele lassen sich leicht (hypothetische, aber nichtsdestoweniger realitätsnahe) Baskets aus verschiedenen Komponenten zu beliebiger Größe zusammensetzen. Zur Verfügung standen geschätzte Werte für die Volatilitäten und Korrelationen der Aktien im DAX (für Details zur Gewinnung der Daten siehe im letzten Kapitel Abschnitt 7.1, insbesondere Tabelle 7.1). Für ein Beispiel mit fünf Faktoren wurde eine Gruppe von Aktien aus der Finanzindustrie ausgewählt und mit dem gleichen relativen Gewicht wie im DAX zu einem Basket kombiniert. Die für die Bewertung von Derivaten relevanten Größen sind in Tabelle 5.5 zusammengetragen. Das Spektrum der Kovarianzmatrix findet sich in Tabelle 5.6. Als Zinssatz wurden 5% ange-

Aktie	i	μ_i	σ_i	$\rho_{ij}, 1 \leq j \leq 5$				
Deutsche Bank	1	38.1	0.518	1.00	0.79	0.82	0.91	0.84
Hypo-Vereinsbank	2	6.5	0.648	0.79	1.00	0.73	0.80	0.76
Commerzbank	3	5.7	0.623	0.82	0.73	1.00	0.77	0.72
Allianz	4	27.0	0.570	0.91	0.80	0.77	1.00	0.90
Münchner Rück	5	22.7	0.530	0.84	0.76	0.72	0.90	1.00

Tabelle 5.5: Für einen Basket aus fünf Finanzwerten die Gewichte μ_i sowie Volatilitäten σ_i und Korrelationen ρ_{ij} der Aktien.

i	λ_i	$q_{ij}, 1 \leq j \leq 5$				
1	1.4094	0.41185	0.49201	0.46551	0.45490	0.40580
2	0.1125	-0.06536	0.46195	-0.82753	0.17662	0.25756
3	0.1006	-0.21914	0.73713	0.20074	-0.38797	-0.46668
4	0.0388	-0.66364	0.03029	0.22253	-0.25437	0.66666
5	0.0213	-0.58109	-0.01584	0.09316	0.73936	-0.32675

Tabelle 5.6: Das Eigensystem des Basket aus Tabelle 5.5.

nommen ($r = 0.05$). Der Ausübungspreis ist $K = 0.25$, die Expiration $T = 1$ (Jahr). Die Auswertung erfolgt jeweils punktweise im aktuellen Kurs $(S_1, S_2, S_3, S_4, S_5) = (K, K, K, K, K)$. Das Gebiet wurde auf $x_1 \in]4 \cdot x_1^0, 0[$ und $x_k \in]-0.5, 0.5[$ abgeschnitten, wobei $x_1^0 < 0$ die x_1 -Koordinate des aktuellen Kurses ist, in dem ausgewertet werden soll.

5.5.2 Numerische Ergebnisse

Termingeschäft

Um den Einfluss der (fehlenden) Glattheit der Startwerte zu sehen, wird zuerst ein Termingeschäft mit analytischer Anfangsbedingung

$$u(\mathbf{x}, 0) = K - \sum_{i=1}^d \mu_i e^{\sum_{j=1}^d q_j i x_j}$$

statt (5.5) berechnet. Die punktweise Konvergenz der Zeitdiskretisierung mit Fractional-Step- θ ist zahlenmäßig in Tab. 5.7 anhand der Differenzen $\delta^m = u^{m+1} - u^m$ für die Lösungen u^m mit 2^m Zeitschritten aufgelistet und in Abb. 5.7 graphisch dargestellt. Die Daten für die

m	0	1	2	3	4	5
u^m	-0.01196035	-0.01204348	-0.01206793	-0.01207428	-0.01207589	-0.01207628
δ^m	-8.31 e-5	-2.44 e-5	-6.35 e-6	-1.61 e-6	-3.95 e-7	
o^m	1.77	1.94	1.98	2.02		

Tabelle 5.7: Konvergenz des Fractional-Step- θ -Schemas für ein dünnes Gitter auf Stufe $n = 9$.

Ortsdiskretisierung auf dem dünnen Gitter mit zentralen Differenzen und $\delta_n = u_{n+1} - u_n$ finden sich in Tab. 5.8 und wieder in Abb. 5.7. Das feinste Gitter hat 372 909 780 Punkte.

n	0	1	2	3	4	5
	6	7	8	9	10	
u_n	242.9280	-0.3996812	-0.06898902	-0.02084201	-0.01383344	-0.01244910
	-0.01215905	-0.01209219	-0.01207589	-0.01207184	-0.01207096	
δ_n	-2.43 e+2	3.31 e-1	4.81 e-2	7.01 e-3	1.38 e-3	2.90 e-4
	6.69 e-5	1.63 e-5	4.05 e-6	8.74 e-7		
o_n	9.52	2.78	2.78	2.34	2.25	2.12
	2.04	2.01	2.21			

Tabelle 5.8: Punktweise Konvergenz am dünnen Gitter für 17 Zeitschritte (siehe Tabelle 5.7).

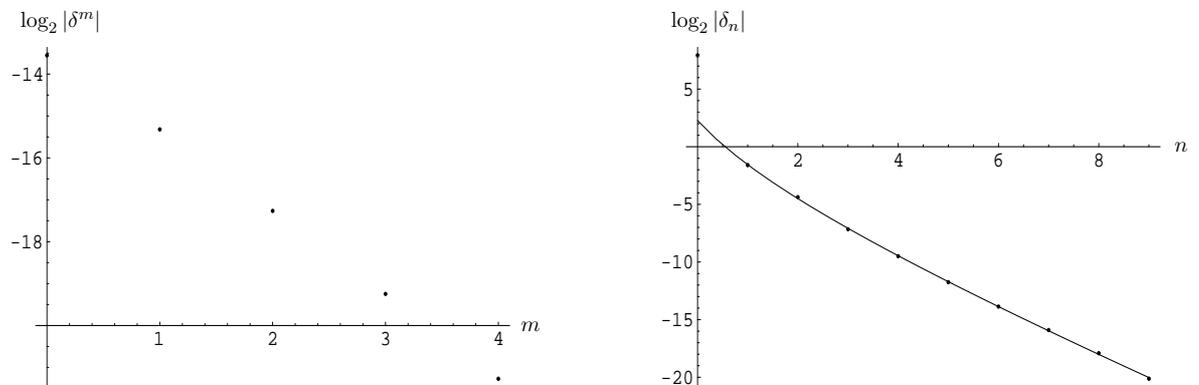


Abbildung 5.7: Termingeschäft: Konvergenz in Ort und Zeit. **Links:** $\delta^m = u^{m+1} - u^m$ für 2^m Zeitschritte und $n = 9$. **Rechts:** $\delta_n = u_{n+1} - u_n$ auf Dünngitter-Level n für $m = 4$.

5 Anwendung auf Optionspreisaufgaben

Die asymptotische Kurve (5.3) in Abbildung 5.7 wurde durch Regression bestimmt. Es ergibt sich

$$\begin{aligned} p &= 1.166, \\ q &= -3.109 \end{aligned}$$

anstelle der theoretisch erwarteten $p = 2$, $q = d - 1 = 4$, obwohl man in Tabelle 5.8 anhand von o_n aus (5.2) klar die zweite Ordnung sieht.

Zur weiteren Illustration gehen wir dazu von der Asymptotik des Fehlers $|u - u_n|$ aus und „bestimmen“ die unbekannte Lösung u so, dass sich aus (5.4) $p = 2$ ergibt, was auf $u = -0.0120709572$ und $q = 3.312$ führt. Umgekehrt wird $q = 4$ für $u = -0.0120709608$ erreicht und dann ist $p = 2.202$.

Dies zeigt, wie sensitiv der geschätzte Exponent des logarithmischen Faktors von den Daten abhängt. Für kleine Gitter können überdies die polynomialen Terme niedrigerer Ordnung nicht vernachlässigt werden.

Dünnes Raum-Zeit-Gitter

Wenn die Diskretisierungsordnung in der Zeit wie im Ort von zweiter Ordnung ist und man überdies davon ausgeht, dass die für die Kombinationstechnik unterstellte Fehlerasymptotik auf die Zeitschrittweite τ ausgedehnt werden kann zu

$$\begin{aligned} u - u_{\mathbf{h},\tau} = & \sum_{\substack{m=1 \\ \{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}}^d [\gamma_{j_1, \dots, j_m, 0}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^2 \cdot \dots \cdot h_{j_m}^2 \\ & + \gamma_{j_1, \dots, j_m, 1}(h_{j_1}, \dots, h_{j_m}, \tau) h_{j_1}^2 \cdot \dots \cdot h_{j_m}^2 \cdot \tau^2], \end{aligned}$$

scheint ein $(d + 1)$ -dimensionales dünnes Gitter in Ort und Zeit sinnvoll. Der asymptotische Aufwand sollte von $\mathcal{O}(2^m 2^n n^{d-1})$ (bei $\tau_m = 2^{-m}$) auf $\mathcal{O}(2^n n^d)$ sinken.

Praktisch wird für verschiedene Zeitschritte $\tau_l = 2^{-l}$ die Differentialgleichung auf dünnen (Orts-)Gittern der Stufe $n - l$ gelöst und gemäß der $(d + 1)$ -dimensionalen Kombinationsformel kombiniert.

Die Konvergenzdaten für das zuvor betrachtete Beispiel stehen in Tabelle 5.9.

n	0	1	2	3	4	5
u_n	-0.00930727	-0.01130481	-0.01181459	-0.01198279	-0.01204242	-0.0120624
δ_n	-2.00 e-3	-5.10 e-4	-1.68 e-4	-5.96 e-5	-2.00 e-5	
o_n	1.97	1.60	1.50	1.58		

Tabelle 5.9: Punktweise Konvergenz am $(5 + 1)$ -dimensionalen dünnen Raum-Zeit-Gitter.

Europäische Option

Die Konvergenzdaten für die Europäische Option auf dem dünnen Gitter finden sich in Tab. 5.10 und Abb. 5.8. Aufgrund der fehlenden Glattheit der Anfangsbedingung ist die Ordnung o_n (Tab. 5.10) nicht mehr so klar ersichtlich wie zuvor. Umsomehr ist es bemerkenswert, dass p und q durch Regression sehr genau bestimmt werden.

n	0	1	2	3	4	5	6
u_n	0.05450104	0.04008010	0.04500332	0.04264528	0.04333669	0.04342236	0.04340526
δ_n	-1.44 e-2	4.92 e-3	-2.36 e-3	6.91 e-4	8.57 e-5	- 1.71 e-5	
o_n	1.55	1.06	1.77	3.01	2.32		

Tabelle 5.10: Europäische Option: Punktweise Konvergenz am dünnen Gitter (33 Zeitschritte).

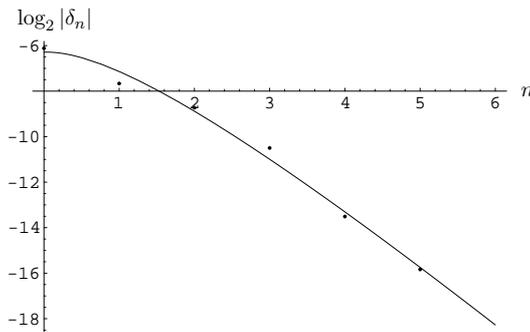


Abbildung 5.8: 5-dimensionaler Basket, Europäische Option: Punktweise Konvergenz am dünnen Gitter mit einer maximalen Anzahl von 372 909 780 Punkten. Die asymptotische Kurve (5.3) wurde durch Regression bestimmt. Es ergibt sich $p = 2.085$, $q = 3.115$.

Amerikanische Option

Die Daten für die Amerikanische Option für die Diskretisierung auf dem dünnen Gitter mit zentralen Differenzen stehen in Tab. 5.11 und Abb. 5.9.

n	0	1	2	3	4	5	6
u_n	0.05487201	0.04195511	0.04640899	0.04440458	0.04495280	0.04501128	0.04502635
δ_n	1.29 e-2	4.45 e-3	2.00 e-3	5.48 e-4	5.85 e-5	1.51 e-5	
o_n	1.54	1.15	1.87	3.23	1.96		

Tabelle 5.11: Amerikanische Option: Punktweise Konvergenz am dünnen Gitter.

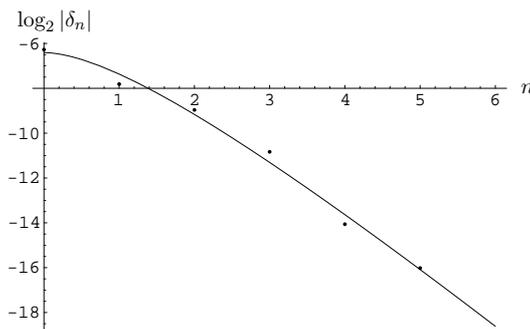


Abbildung 5.9: 5-dimensionaler Basket, Amerikanische Option: Punktweise Konvergenz am dünnen Gitter mit einer maximalen Anzahl von 372 909 780 Punkten. Die asymptotische Kurve (5.3) wurde durch Regression bestimmt. Es ergibt sich $p = 2.066$, $q = 2.915$.

5 Anwendung auf Optionspreisaufgaben

6 Multivariate Extrapolation und eine Kombinationstechnik höherer Ordnung

Überblick

Ziel dieses Kapitels ist es, den Extrapolationsmechanismus der Kombinationstechnik weitergehend auszuschöpfen, um gleichzeitig die Ordnung zu erhöhen. In 6.1 wird ein multivariates Extrapolationsschema abgeleitet, das gemischt quadratische Terme aus der Fehlerasymptotik auf kartesischen Gittern eliminiert. Im Zusammenspiel mit Konvergenzaussagen für die Kombination von derartigen Lösungen höherer Ordnung (6.2) entsteht in 6.3 eine verallgemeinerte Dünngitter-Lösung vierter Ordnung aus Skelettlösungen zweiter Ordnung auf den Kartesischen Gittern. Numerische Konvergenzbeispiele und aus der Lösung durch numerische Differentiation gewonnene Ableitungen illustrieren dies.

6.1 Multivariate Extrapolation

Ziel dieses Abschnitts ist es, eine Kombinationsformel höherer Ordnung vorzubereiten, wobei zunächst ein Extrapolationsschema vierter Ordnung entwickelt wird.

Bemerkung (Wieso nicht Richardson?). *Es stellt sich die Frage, ob man ein Extrapolationsverfahren direkt auf die Kombinationslösung anwenden kann. Wie die folgenden Zeilen für die Richardson-Extrapolation zeigen, scheitert das an den logarithmischen Faktoren in der Fehlerentwicklung*

$$u - u_h = c \cdot |\log h|^{d-1} h^p + \mathcal{O}\left(|\log h|^{d-2} h^p\right).$$

Der gewohnte Ansatz bei Ordnung p

$$\begin{aligned} 2^p(u_h - u) - (u_{2h} - u) &= 2^p |\log h|^{d-1} h^p - |\log 2h|^{d-1} (2h)^p + \mathcal{O}\left(|\log h|^{d-2} h^p\right) \\ &= (2h)^p \left(|\log h|^{d-1} - |\log 2h|^{d-1}\right) + \mathcal{O}\left(|\log h|^{d-2} h^p\right) \\ &= -\frac{2^p(d-1)}{\theta} |\log \theta h|^{d-2} h^p + \mathcal{O}\left(|\log h|^{d-2} h^p\right) \end{aligned}$$

(mit $\theta \in [1, 2]$ aus dem Mittelwertsatz) reduziert für $d \geq 2$ nur den Exponenten des Logarithmus. Die Ordnung p bleibt. Man könnte allenfalls versuchen, durch mehrere (und zwar d) derartige Schritte alle Logarithmen zum Verschwinden zu bringen.

Hier wird hingegen die Reihenfolge von Extrapolation und Kombination umgedreht. Durch Extrapolation auf klassischen kartesischen Gittern werden zunächst Lösungen höherer Ordnung konstruiert und dann erst kombiniert.

Bemerkung (Dünne Gitter höherer Ordnung). *Es ist zu beachten, dass ein Diskretisierungsverfahren, das auf einer Folge voller Gitter (etwa punktweise) mit einer Ordnung größer als zwei konvergiert, zusammen mit der Kombinationsformel nicht notwendigerweise auf ein Verfahren der entsprechenden Ordnung auf dem dünnen Gitter führt. Schuld daran sind Fehlerterme gemischter Ordnung in den einzelnen Raumrichtungen, die zwar zusammen von hoher, einzeln betrachtet aber mitunter nur von zweiter Ordnung sind.*

Betrachtet man nämlich in zwei Dimensionen einen Ausdruck der Gestalt

$$h_1^p h_2^2$$

mit $p > 2$, so wird daraus durch Kombination

$$\begin{aligned} \Delta_n \sum_{i=0}^n 2^{-pi} 4^{-(n-i)} &= 2^{-p(n+1)} + \sum_{i=0}^n 2^{-pi} \left[4^{-(n+1-i)} - 4^{-(n-i)} \right] \\ &= 2^{-p(n+1)} - \frac{3}{4} \sum_{i=0}^n 2^{-pi} 4^{-(n-i)} \\ &= 2^{-p(n+1)} - 3 \cdot 4^{-(n+1)} \sum_{i=0}^n 2^{(2-p)i} \\ &= 2^{-p(n+1)} - 3 \cdot 4^{-(n+1)} \frac{1 - 2^{(n+1)(2-p)}}{1 - 2^{(2-p)}} \\ &= \mathcal{O}(4^{-n}), \end{aligned}$$

obwohl man am vollen Gitter die Ordnung $p + 2$ sehen würde.

Durch einen vorgeschalteten Extrapolationsschritt auf den einzelnen Gittern des Kombinationstableaus soll sichergestellt werden, dass tatsächlich alle gemischten Fehlerterme bis zur gewünschten Ordnung verschwinden.

Arbeiten in Richtung systematischer multivariater Extrapolation sind z. B. in [BGR] zu finden, wo unter anderem das unter dem Namen *Splitting Extrapolation* von Schüller und Lin [SL85] in zwei Dimensionen eingeführte Schema

$$u_{h_1, h_2}^{extra} = \frac{4}{3} u_{h_1/2, h_2} + \frac{4}{3} u_{h_1, h_2/2} - \frac{5}{3} u_{h_1, h_2}$$

betrachtet wird. Die resultierende Ordnung $\mathcal{O}(h_1^2 h_2^2 + h_1^4 + h_2^4)$ ist aber aus besagten Gründen unzureichend.

Zur Bestimmung der für die Extrapolation benötigten Gitter und der dazupassenden Koeffizienten orientieren wir uns daher zunächst systematisch an quadratischen Fehlertermen der Gestalt

$$\begin{aligned} u - u_{\mathbf{h}} &= \sum_{j_1=1}^d \beta_{j_1} h_{j_1}^2 + \sum_{j_1, j_2=1}^d \beta_{j_1, j_2} h_{j_1}^2 h_{j_2}^2 + \dots + \beta_{1, \dots, d} h_1^2 \dots h_d^2 + R(\mathbf{h}) \\ &= \sum_{m=0}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \beta_{j_1, \dots, j_m} h_{j_1}^2 \dots h_{j_m}^2 + R(\mathbf{h}), \end{aligned} \quad (6.1)$$

die allesamt herausextrapoliert werden müssen. Für ein Tableau, das aus dieser notwendigen Bedingung bestimmt wird, kann dann tatsächlich Konvergenz vierter Ordnung bewiesen werden.

In die Extrapolation werden alle Gitter miteinbezogen, die in jede Richtung höchstens einmal — also insgesamt bis zu d mal — verfeinert wurden (siehe Abbildung 6.1 für zwei Dimensionen). Das führt auf Gitterlösungen der Gestalt

$$u_{\mathbf{h}}^{(i_1, \dots, i_k)} := u_{\mathbf{h}(i_1, \dots, i_k)}, \text{ wobei } h_j^{(i_1, \dots, i_k)} := \begin{cases} h_j/2 & j \in \{i_1, \dots, i_k\}, \\ h_j & \text{sonst.} \end{cases} \quad (6.2)$$

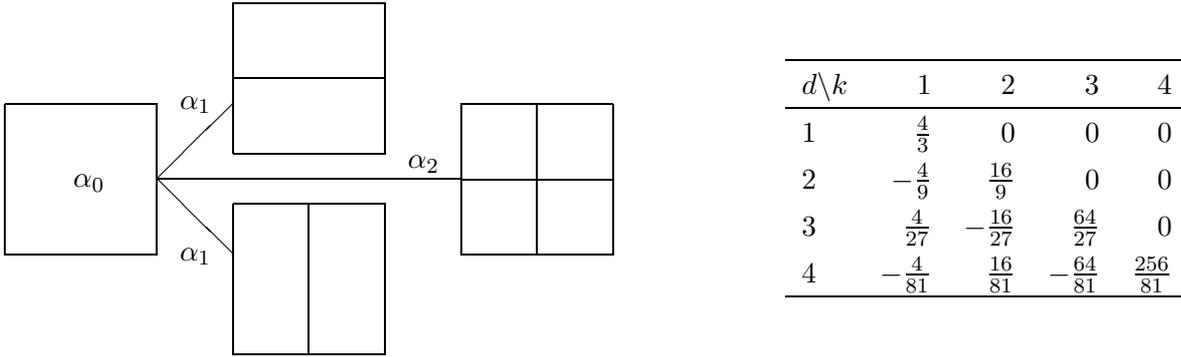


Abbildung 6.1: Links: Multivariates Extrapolationsschema nach (6.2) in zwei Dimensionen. Rechts: Die Gewichte α_k für niedrige Dimensionen d .

Diese Hierarchie korrespondiert in natürlicher Weise mit den Fehlertermen.

Der naheliegende Ansatz für die Extrapolationslösung lautet dann

$$\begin{aligned}
 \tilde{u}_{\mathbf{h}} &= \alpha_0 u_{\mathbf{h}} + \alpha_1 \sum_{i_1=1}^d u_{\mathbf{h}}^{(i_1)} + \alpha_2 \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^d u_{\mathbf{h}}^{(i_1, i_2)} + \dots + \alpha_d u_{\mathbf{h}}^{(1, \dots, d)} \\
 &= \sum_{k=0}^d \alpha_k \sum_{\substack{\{i_1, \dots, i_k\} \\ \subset \{1, \dots, d\}}} u_{\mathbf{h}}^{(i_1, \dots, i_k)}. \tag{6.3}
 \end{aligned}$$

Die Koeffizienten α_i müssen so bestimmt werden, dass alle quadratischen Terme herausfallen. Die innere Summe in (6.3) besteht aus $\binom{d}{k}$ Summanden. Eine minimale Konsistenzforderung ist also zusätzlich

$$\sum_{k=0}^d \alpha_k \binom{d}{k} = 1. \tag{6.4}$$

Was jetzt bevorsteht ist die mühsame und sehr technische Auswertung einer Hierarchie von multivariaten Fehlertermen (beschrieben durch die Indizes j_1, \dots, j_m) auf einer Hierarchie von anisotropen Gittern (beschrieben durch i_1, \dots, i_k). Da sich jeder einzelne Fehlerterm bei der Extrapolation herauskürzen muss, halten wir j_1, \dots, j_m fest und summieren über $\{i_1, \dots, i_k\} \subset \{1, \dots, d\}$.

Setzt man die Definition (6.2) in die Entwicklung (6.1) ein, erhält man

$$u - u_{\mathbf{h}}^{(i_1, \dots, i_k)} = \sum_{m=1}^d \sum_{\substack{j_1, \dots, j_m \\ \text{pw. versch.}}} 4^{-N(\mathbf{i}, \mathbf{j})} \beta_{j_1, \dots, j_m} h_{j_1}^2 \dots h_{j_m}^2 + R(\mathbf{h}),$$

wobei $N(\mathbf{i}, \mathbf{j}) := |\{i_1, \dots, i_k\} \cap \{j_1, \dots, j_m\}|$. Sei für fixe j_1, \dots, j_m

$$N_l := \sum_{\substack{\{i_1, \dots, i_k\} \subset \{1, \dots, d\} \\ |\{j_1, \dots, j_m\} \cap \{i_1, \dots, i_k\}| = l}} 1$$

(die Häufigkeit von $N(\mathbf{i}, \mathbf{j}) = l$), dann können wir die Summe über alle \mathbf{i} — d. h. über alle Gitter mit k Verfeinerungen — teilweise ausführen:

$$\sum_{\substack{\{i_1, \dots, i_k\} \\ \subset \{1, \dots, d\}}} \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} 4^{-N(\mathbf{i}, \mathbf{j})} \beta_{j_1, \dots, j_m} h_{j_1}^2 \dots h_{j_m}^2 = \sum_{l=0}^d N_l 4^{-l} \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \beta_{j_1, \dots, j_m} h_{j_1}^2 \dots h_{j_m}^2.$$

Es ist klar, dass $N(\mathbf{i}, \mathbf{j}) \leq \min(m, k)$, außerdem $N(\mathbf{i}, \mathbf{j}) \geq m + k - d$. Ansonsten treten genau

$$N_l = \binom{m}{l} \binom{d-m}{k-l}$$

Schnittmengen der Mächtigkeit l auf. Demnach können wir schreiben

$$\sum_{\substack{\{i_1, \dots, i_k\} \\ \subset \{1, \dots, d\}}} \left(u - u^{(i_1, \dots, i_k)} \right) = \sum_{m=1}^d \left(\sum_{l=\max(0, m+k-d)}^{\min(m, k)} 4^{-l} N_l \right) \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \beta_{j_1, \dots, j_m} h_{j_1}^2 \dots h_{j_m}^2 + R(\mathbf{h})$$

sowie unter der Konsistenzbedingung (6.4)

$$u - \tilde{u}_{\mathbf{h}} = \sum_{k=1}^d \alpha_k \sum_{\substack{\{i_1, \dots, i_k\} \\ \subset \{1, \dots, d\}}} \left(u - u^{(i_1, \dots, i_k)} \right).$$

Der m -te Fehlerterm ($1 \leq m \leq d$) verschwindet also, falls

$$\sum_{k=0}^d \alpha_k \sum_{l=\max(0, m+k-d)}^{\min(m, k)} 4^{-l} \binom{m}{l} \binom{d-m}{k-l} = 0. \quad (6.5)$$

(6.4) und (6.5) bilden ein $(d+1)$ -dimensionales lineares Gleichungssystem für die $d+1$ Unbekannten α_k , $0 \leq k \leq d$. Die Lösungen für kleine d in Abb. 6.1 legen die Vermutung

$$\alpha_k = \frac{(-4)^k}{(-3)^d} \quad 0 \leq k \leq d \quad (6.6)$$

nahe. Wir beweisen nun, dass die α_k aus (6.6) wirklich auf eine Extrapolationsformel vierter Ordnung führen und gehen dazu von einer Fehlerasymptotik der Form

$$\begin{aligned} u - u_{\mathbf{h}} &= \sum_{j=1}^d \beta_j(h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_d) h_j^2 \\ &+ \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^4 \dots h_{j_m}^4 \end{aligned} \quad (6.7)$$

aus. Die rein quadratischen Anteile sind in der ersten Summe abgetrennt, die höheren Ordnungen in der zweiten zusammengefasst.

6 Multivariate Extrapolation und eine Kombinationstechnik höherer Ordnung

Lemma 6.1. Sei $\alpha_k := (-4)^k/(-3)^d$, $0 \leq k \leq d$. Dann gilt für beliebige $\beta : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$, $i_j \in \{0, 1\}$, $1 \leq j \leq d$:

$$\sum_{k=0}^d \alpha_k \sum_{i_1+\dots+i_d=k} 4^{-i_1} \beta(2^{-i_2}, \dots, 2^{-i_d}) = 0.$$

Beweis. Einsetzen ergibt

$$\begin{aligned} \sum_{k=0}^d \alpha_k \sum_{i_1+\dots+i_d=k} 4^{-i_1} \beta(h_2, \dots, h_d) &= \sum_{k=0}^d \frac{(-4)^k}{(-3)^d} \sum_{i_1+\dots+i_d=k} 4^{-i_1} \beta(h_2, \dots, h_d) \\ &= \frac{1}{(-3)^d} \sum_{k=0}^d (-1)^k \sum_{i_1+\dots+i_d=k} 4^{i_2+\dots+i_d} \beta(h_2, \dots, h_d) \\ &= \frac{1}{(-3)^d} \sum_{k=0}^d (-1)^k \sum_{i_1+\dots+i_d=k} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d}. \end{aligned} \quad (6.8)$$

Da die letzte Summe nicht von h_1 abhängt, gilt für alle $0 < k < d$

$$\sum_{i_1+\dots+i_d=k} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d} = \sum_{i_2+\dots+i_d=k-1} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d} + \sum_{i_2+\dots+i_d=k} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d}$$

und am Anfang bzw. Ende der Summe

$$\sum_{i_1+\dots+i_d=0} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d} = \sum_{i_2+\dots+i_d=0} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d}$$

bzw.

$$\sum_{i_1+\dots+i_d=d} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d} = \sum_{i_2+\dots+i_d=d-1} \frac{\beta(h_2, \dots, h_d)}{h_2 \cdot \dots \cdot h_d}.$$

Daher ist (6.8) eine Teleskopsumme. □

Satz 6.2. Sei

$$\tilde{u}_{\mathbf{h}} = \sum_{k=0}^d \alpha_k \sum_{\substack{\{i_1, \dots, i_k\} \\ \subset \{1, \dots, d\}}} u_{\mathbf{h}}^{(i_1, \dots, i_k)}$$

wie in (6.2), (6.3) mit $\alpha_k = (-4)^k/(-3)^d$ für $0 \leq k \leq d$, wobei für $u_{\mathbf{h}}$

$$\begin{aligned} u - u_{\mathbf{h}} &= \sum_{j=1}^d \beta_j(h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_d) h_j^2 \\ &\quad + \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^4 \dots h_{j_m}^4 \end{aligned}$$

und $|\gamma_{j_1, \dots, j_m}| \leq K$ für alle j_1, \dots, j_m vorausgesetzt wird. Dann gilt

$$u - \tilde{u}_{\mathbf{h}} = \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \tilde{\gamma}_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^4 \cdots h_{j_m}^4,$$

wobei $|\tilde{\gamma}_{j_1, \dots, j_m}| \leq \frac{5^d}{3^d} K$.

Beweis. Es wird zuerst die Normierungsbedingung (6.4) gezeigt:

$$\sum_{k=0}^d \alpha_k \binom{d}{k} = \frac{1}{(-3)^d} \sum_{k=0}^d (-4)^k \binom{d}{k} = \frac{1}{(-3)^d} (1 - 4)^d = 1.$$

Der Rest ergibt sich aus Lemma 6.1 sowie aus

$$\begin{aligned} u - \tilde{u}_{\mathbf{h}} &= \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \sum_{k=0}^d \alpha_k \sum_{\substack{\sum_l i_l = k \\ i_l \in \{0, 1\}}} \gamma_{j_1, \dots, j_m}(h_{j_1+i_{j_1}}, \dots, h_{j_m+i_{j_m}}) h_{j_1+i_{j_1}}^4 \cdots h_{j_m+i_{j_m}}^4 \\ &= \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} h_{j_1}^4 \cdots h_{j_m}^4 \underbrace{\sum_{k=0}^d \frac{(-4)^k}{(-3)^d} \sum_{\substack{\sum_l i_l = k \\ i_l \in \{0, 1\}}} \gamma_{j_1, \dots, j_m}(h_{j_1} 2^{-i_{j_1}}, \dots, h_{j_m} 2^{-i_{j_m}}) 16^{-(i_{j_1} + \dots + i_{j_m})}}_{=:\tilde{\gamma}_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m})} \end{aligned}$$

mit

$$|\tilde{\gamma}_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m})| \leq \sum_{k=0}^d K \frac{4^k}{3^d} \binom{d}{k} = K \frac{5^d}{3^d}.$$

□

6.2 Eine Kombination höherer Ordnung

6.2.1 Das Prinzip

Wir wenden diese Extrapolationsformel nun auf alle Gitter aus dem Kombinationstableau (4.9) an. Um auf so einem Gitter auf Stufe l eine Extrapolationslösung aufzubauen, greifen wir auf größere Gitter, nämlich auf jene auf den d zuvorliegenden Stufen zu. Das werden im Allgemeinen sowohl Gitter sein, die schon in der Kombinationsformel enthalten sind, als auch solche, die das nicht sind. Die „Bandweite“, also die Anzahl der l , die miteinbezogen werden müssen, verdoppelt sich demnach von d auf $2d$. Aus der Kombinationsformel

$$u_n = \sum_{l=n}^{n+d-1} a_{l-n} \sum_{|j|=l} u_j$$

wird

$$\begin{aligned}
 \hat{u}_n &= \sum_{l=n}^{n+d-1} a_{l-n} \sum_{|\mathbf{j}|=l} \tilde{u}_{\mathbf{j}} \\
 &= \sum_{l=n}^{n+d-1} a_{l-n} \sum_{|\mathbf{j}|=l} \sum_{i=0}^d \alpha_i \sum_{\substack{i_k \in \{0,1\} \\ |\mathbf{i}|=i}} u_{\mathbf{j}+\mathbf{i}} \\
 &= \sum_{j=0}^{d-1} a_j \sum_{|\mathbf{j}|=n+j} \sum_{i=0}^d \alpha_i \sum_{\substack{i_k \in \{0,1\} \\ |\mathbf{i}|=i}} u_{\mathbf{j}+\mathbf{i}}.
 \end{aligned}$$

Wir ordnen jetzt nach den Gitterlösungen um. Wenn $N(\mathbf{k})$ die Anzahl der Nicht-Null-Einträge von \mathbf{k} ist, dann gibt es genau $\binom{N(\mathbf{k})}{i}$ Zerlegungen der Form $\mathbf{k} = \mathbf{i} + \mathbf{j}$ mit $i_k \in \{0, 1\}$, $|\mathbf{i}| = i$, sofern $N(\mathbf{k}) \geq i$, sonst 0. Durch Einsetzen von

$$\sum_{|\mathbf{j}|=n+j} \sum_{\substack{i_k \in \{0,1\} \\ |\mathbf{i}|=i}} u_{\mathbf{j}+\mathbf{i}} = \sum_{\substack{i_k \in \{0,1\}, j_k \geq 0 \\ |\mathbf{i} + \mathbf{j}| = n+i+j}} u_{\mathbf{j}+\mathbf{i}} = \sum_{|\mathbf{k}|=n+i+j} \binom{N(\mathbf{k})}{i} u_{\mathbf{k}}$$

erhält man dann mit den Substitutionen $i = k - j$ ($0 \leq i = k - j \leq d \Rightarrow k - d \leq j \leq k$) und $l = n + k$

$$\begin{aligned}
 \hat{u}_n &= \sum_{j=0}^{d-1} a_j \sum_{i=0}^d \alpha_i \sum_{|\mathbf{k}|=n+i+j} \binom{N(\mathbf{k})}{i} u_{\mathbf{k}} \\
 &= \sum_{k=0}^{2d-1} \sum_{j=\max(0, k-d)}^{\min(k, d-1)} \sum_{|\mathbf{k}|=n+k} a_j \alpha_{k-j} \binom{N(\mathbf{k})}{k-j} u_{\mathbf{k}} \\
 &= \sum_{l=n}^{n+2d-1} \sum_{|\mathbf{l}|=l} \sum_{j=\max(0, l-n-d)}^{\min(l-n, d-1)} a_j \alpha_{l-n-j} \binom{N(\mathbf{l})}{l-n-j} u_{\mathbf{l}}.
 \end{aligned}$$

Endergebnis ist also die Formel

$$\boxed{\hat{u}_n = \sum_{l=n}^{n+2d-1} \sum_{|\mathbf{l}|=l} b_{\mathbf{l}} u_{\mathbf{l}}} \tag{6.9}$$

mit entsprechend gewählten Koeffizienten

$$b_{\mathbf{l}} = \sum_{j=\max(0, l-n-d)}^{\min(l-n, d-1)} a_j \alpha_{l-n-j} \binom{N(\mathbf{l})}{l-n-j}. \tag{6.10}$$

Bemerkung. Für symmetrische Diskretisierungen (d. h. nur Fehlerterme gerader Ordnung) liefert die Kombinationsformel (6.9) mit (6.10) aus einer Hierarchie von Lösungen mit quadratischer Fehlerasymptotik eine Lösung mit Genauigkeit vierter Ordnung (im Vergleich zu

zweiter Ordnung bei der herkömmlichen Kombinationsformel). Der Preis, der dafür zu bezahlen ist, besteht im „Mitschleppen“ von d weiteren Stufen. Der Aufwand zur Berechnung dieser Lösungen ist vergleichsweise klein. Es sei jedoch bemerkt, dass in höheren Dimensionen eine Berechnung der Lösung selbst auf der niedrigsten Stufe schon zu aufwändig sein kann (in fünf Dimensionen etwa müssen schon zehn Stufen betrachtet werden).

6.2.2 Konvergenzabschätzungen

Korollar 6.3 (zu Satz 4.1). Für alle Lösungen $u_{\mathbf{h}}$ gelte punktweise eine Fehlerasymptotik der Form

$$u - u_{\mathbf{h}} = \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^p \cdot \dots \cdot h_{j_m}^p, \quad (6.11)$$

wobei

$$|\gamma_{j_1, \dots, j_m}| \leq K \quad \forall 1 \leq m \leq d \quad \forall \{j_1, \dots, j_m\} \subset \{1, \dots, d\}.$$

Dann erfüllt die Kombinationslösung u_n aus Satz 4.1 die Fehlerabschätzung

$$|u - u_n| \leq \frac{2K}{(d-1)!} \left(\frac{2^p + 1}{2^{p-1}} \right)^{d-1} (n + 2d - 2)^{d-1} 2^{-pn}.$$

Beweis. Dem Beweis von Satz 4.1 folgend, ergeben sich schrittweise die Modifikationen

$$\Delta_n^{d-m} \sum_{|\mathbf{i}|=n} \gamma_{1, \dots, m}(h_{i_1}, \dots, h_{i_m}) h_{i_1}^p \cdot \dots \cdot h_{i_m}^p = s_{n+d-m} 2^{-p(n+d-m)}$$

sowie

$$\Delta_n^j 2^{-pn} = \left(-\frac{2^p - 1}{2^p} \right)^j 2^{-pn},$$

sodass

$$\begin{aligned} \left| \Delta_n^{d-1} \sum_{|\mathbf{i}|=n} \gamma_{1, \dots, m}(h_{i_1}, \dots, h_{i_m}) h_{i_1}^p \cdot \dots \cdot h_{i_m}^p \right| &= 2^{-p(n+d-1)} \left| (\Delta_n - 2^p - 1)^{m-1} s_{n+d-m} \right| \\ &< 2^{-pn} \left(\frac{2^p + 1}{2^p} \right)^{d-1} \frac{K}{(d-1)!} (n + 2d - 2)^{d-1} \end{aligned}$$

und die Abschätzung wie gehabt folgt. □

Korollar 6.3 und Satz 6.2 ergeben nun die Konvergenzordnung der modifizierten Kombinationsformel.

Satz 6.4 (Konvergenz vierter Ordnung). Für alle Lösungen $u_{\mathbf{h}}$ gelte punktweise eine Fehlerasymptotik der Form

$$\begin{aligned} u - u_{\mathbf{h}} &= \sum_{j=1}^d \beta_j(h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_d) h_j^2 \\ &+ \sum_{m=1}^d \sum_{\substack{\{j_1, \dots, j_m\} \\ \subset \{1, \dots, d\}}} \gamma_{j_1, \dots, j_m}(h_{j_1}, \dots, h_{j_m}) h_{j_1}^4 \cdot \dots \cdot h_{j_m}^4, \end{aligned}$$

6 Multivariate Extrapolation und eine Kombinationstechnik höherer Ordnung

wobei

$$|\gamma_{j_1, \dots, j_m}| \leq K \quad \forall 1 \leq m \leq d \quad \forall \{j_1, \dots, j_m\} \subset \{1, \dots, d\}.$$

Dann erfüllt die Kombinationslösung

$$\hat{u}_n = \sum_{l=n}^{n+2d-1} \sum_{|\mathbf{l}|=l} b_{\mathbf{l}} u_{\mathbf{l}},$$

mit

$$b_{\mathbf{l}} = \sum_{j=\max(0, l-n-d)}^{\min(l-n, d-1)} a_j \alpha_{l-n-j} \binom{N(\mathbf{l})}{l-n-j}, \quad \alpha_j = (-4)^j / (-3)^d, \quad a_j = (-1)^{d-1-i} \binom{d-1}{i}$$

und $N(\mathbf{l})$ der Anzahl der Nicht-Null-Elemente im Multiindex \mathbf{l} , die folgende Fehlerabschätzung:

$$|u - \hat{u}_n| \leq \frac{10}{3} \frac{K}{(d-1)!} \left(\frac{85}{24}\right)^{d-1} (n+2d-2)^{d-1} 16^{-n}.$$

Beweis. Korollar 6.3 liefert für $p = 4$ die entsprechende Ordnung mit einem Faktor $(17/8)^{d-1}$, von Satz (6.2) kommt zusätzlich $(5/3)^d$. \square

Verbesserte Abschätzungen und asymptotische Aussagen können analog zu Korollar 4.5 abgeleitet werden.

6.3 Numerische Ergebnisse

Zur experimentellen Konvergenzanalyse werden wieder die Beispiele aus Kapitel 5 betrachtet.

6.3.1 FX-Option

Für das 3-Faktor-Modell der FX-Option aus 5.2.1 ist die Konvergenz der numerischen Lösung und der ersten beiden Abbildungen davon nach x_1 (wie in 2.3.2) in Abb. 6.2 zu sehen. Zur Schätzung der Konvergenzordnung wurde der logarithmische Faktor extrahiert und

$$|u_{n+1} - u_n| - 4 \log_2 n \sim a - pn$$

betrachtet.

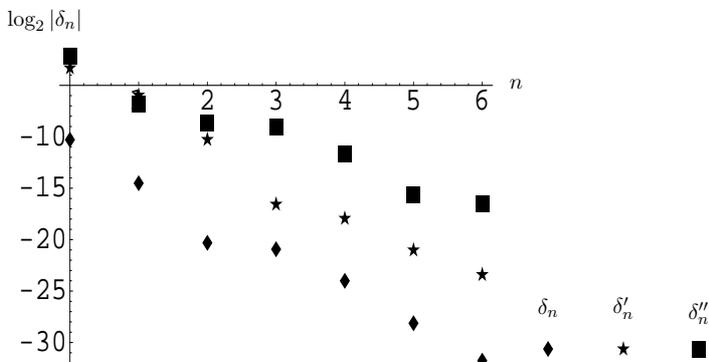


Abbildung 6.2: 3-Faktor Modell für die FX-Option: Eine Regression der vom logarithmischen Faktor bereinigten Werte für $\delta_n = u_{n+1} - u_n$ ab $n = 3$ gibt die Ordnungen 4.20, 2.90 und 3.18 für Lösung, erste und zweite Ableitung nach x_1 .

6.3.2 Basket mit zwei Komponenten

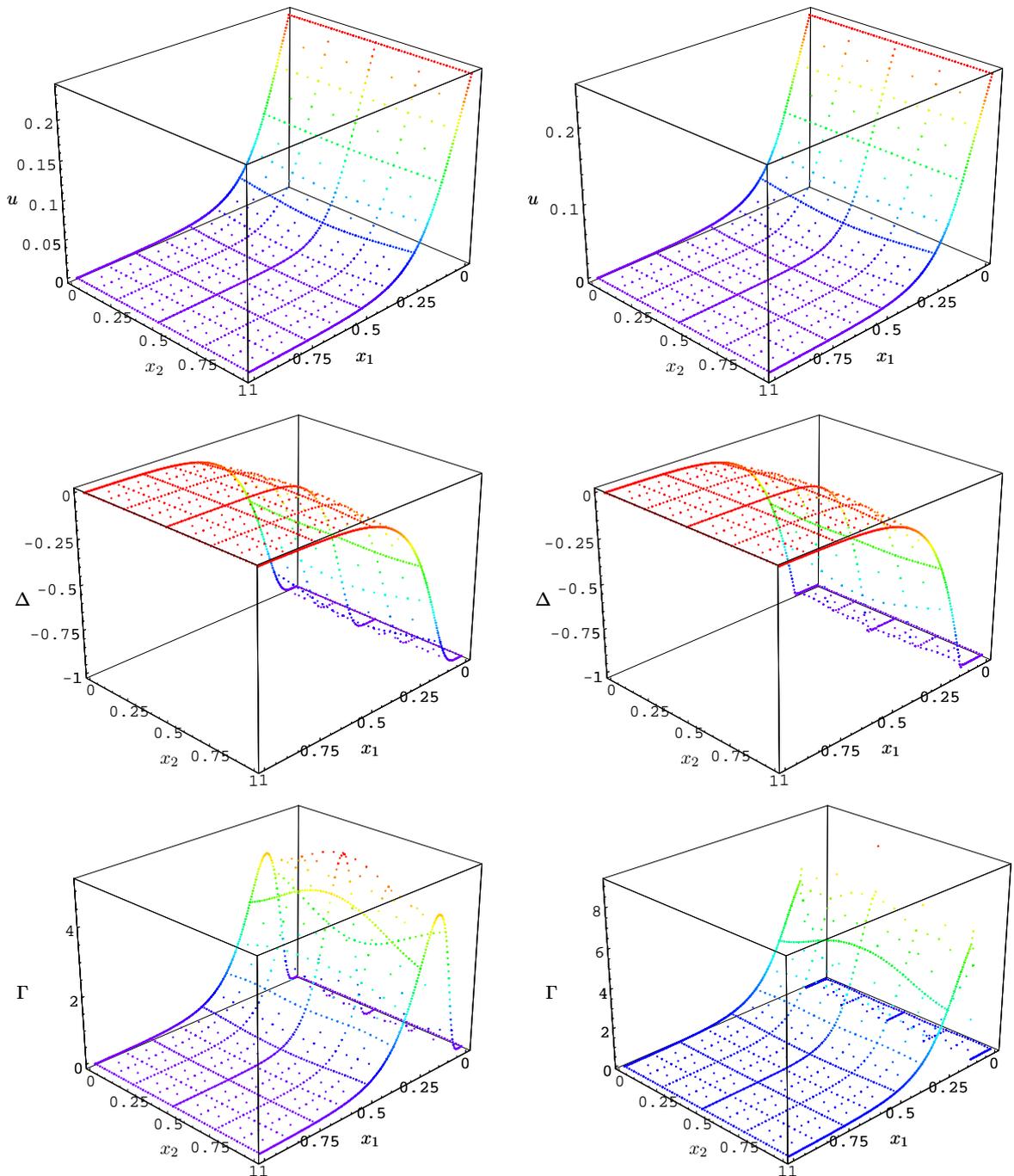


Abbildung 6.3: Put auf einen Basket aus zwei Aktien wie in Abb. 2.1, 2.2, dargestellt im transformierten System x_1 - x_2 aus 1.3.2. Gerechnet wurde auf einem dünnen Gitter vierter Ordnung mit insgesamt 4210 Punkten, verteilt auf 22 Gittern. Dargestellt sind jeweils die numerische Lösung u und deren erste und zweite Ableitung, $\Delta = \frac{\partial u}{\partial x_1}$, $\Gamma = \frac{\partial^2 u}{\partial x_1^2}$ für eine Europäische Option (**links**) und die Amerikanische (**rechts**). Der Sprung der zweiten Ableitung am freien Rand wird am dünnen Gitter überraschend gut aufgelöst.

6.3.3 Basket mit fünf Komponenten

Es werden noch einmal die Beispiele aus 5.5 berechnet. Die Daten sind wie dort, lediglich wurde das Gebiet vergrößert: deshalb $K = 0.125$, $x_1 \in]8 \cdot x_1^0, 0[$, sonst $x_k \in]-1, 1[$. Der Preis muss daher im Vergleich zu 5.5.2 mit 2 skaliert werden.

Angegeben sind jeweils die Gitterlösung u'_n und die ersten beiden Ableitungen u'_n und u''_n nach x_1 , dazu $\delta_n = u_{n+1} - u_n$ und die geschätzte Ordnung o_n bzw. analog für u'_n, u''_n .

Termingeschäft

n	0	1	2	3	4	5
u_n	-1.015471	-0.01105000	-0.01194737	-0.01205911	-0.01206908	-0.01207037
u'_n	1.400938	1.090555	0.7794913	0.8507590	0.8568695	0.8573149
u''_n	-6.626567	-2.53794	-2.782764	-2.875418	-2.895245	-2.897572
δ_n	1.00	-8.97 e-4	-1.12 e-4	-9.96 e-6	-1.29 e-6	
δ'_n	-3.10 e-1	-3.11	7.13 e-2	6.11 e-3	4.45 e-4	
δ''_n	4.09	-2.45 e-1	-9.27 e-2	-1.98 e-2	-2.33 e-3	
o_n	10.1	3.01	3.49	2.95		
o'_n	-3.16 e-3	2.13	3.54	3.78		
o''_n	4.06	1.40	2.22	3.09		

Tabelle 6.1: Konvergenz der Kombinationstechnik vierter Ordnung beim Termingeschäft.

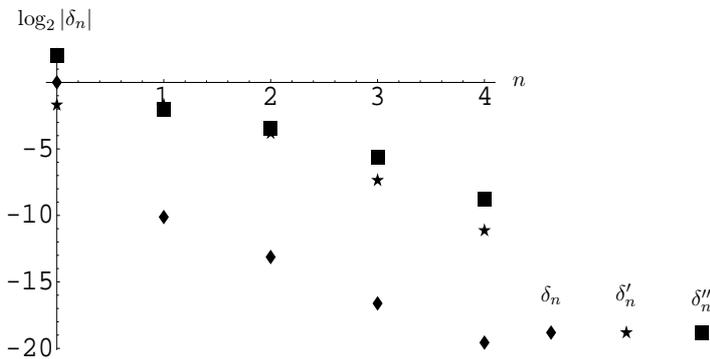


Abbildung 6.4: Konvergenz der Kombinationstechnik vierter Ordnung beim Termingeschäft für den Optionspreis und die ersten beiden Ableitungen nach x_1 . Trotz der glatten Daten wird der asymptotische Bereich nicht erreicht.

Europäische und Amerikanische Option

n	0	1	2	3	4	5
u_n	0.01974383	0.02770947	0.02091797	0.02170517	0.02193318	0.02196671
u'_n	-0.5122855	-0.5909403	-0.5268881	-0.5046863	-0.5268322	-0.5228420
u''_n	7.961965	4.258356	5.900518	6.224795	6.093612	6.061680
δ_n	7.97 e-3	6.79 e-3	7.87 e-4	2.28 e-4	3.35 e-5	
δ'_n	-7.87 e-2	6.41 e-2	2.22 e-2	-2.21 e-2	3.99 e-3	
δ''_n	-3.70	1.64	3.24 e-1	-1.31 e-3	-3.19 e-2	
o_n	0.230	3.11	1.79	2.77		
o'_n	0.296	1.53	3.64 e-3	2.47		
o''_n	1.17	2.34	1.31	2.04		

Tabelle 6.2: Konvergenz der Kombinationstechnik vierter Ordnung, Europäische Option.

n	0	1	2	3	4	5
u_n	0.02034060	0.02808817	0.02166512	0.02235362	0.02263813	0.0226578
u'_n	-0.5366711	-0.6093635	-0.5518479	-0.5292749	-0.5512550	-0.5474903
u''_n	8.833452	4.968524	6.511755	6.985718	6.768850	6.766430
δ_n	7.75 e-3	-6.42 e-3	6.88 e-4	2.85 e-4	1.97 e-5	
δ'_n	-7.27 e-2	5.75 e-2	2.26 e-2	-2.20 e-2	3.76 e-3	
δ''_n	-3.86	1.54	4.74 e-1	-2.17 e-1	-2.42 e-3	
o_n	0.270	3.22	1.27	3.85		
o'_n	0.338	1.35	0.038	2.55		
o''_n	1.32	1.70	1.13	6.49		

Tabelle 6.3: Konvergenz der Kombinationstechnik vierter Ordnung, Amerikanische Option.

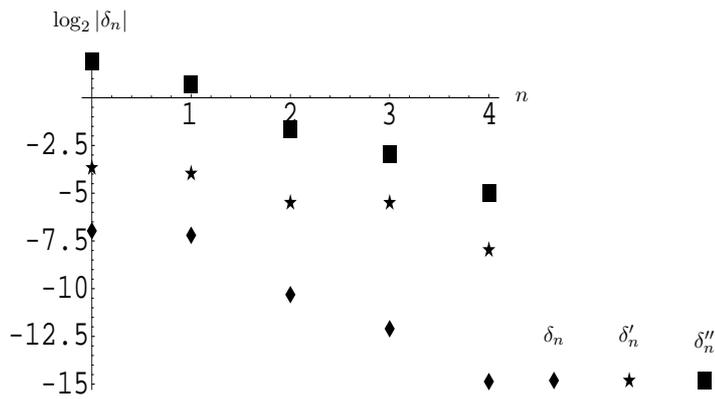


Abbildung 6.5: Punktweise Konvergenz der Kombinationstechnik vierter Ordnung, Europäische Option. Für praktikable Gittergrößen ist keine wesentlich Genauigkeits- bzw. Ordnungsverbesserung für die Lösung im Vergleich zum Verfahren zweiter Ordnung (Abb. 5.8) sichtbar. Dennoch können die Sensitivitäten für praktische Zwecke hinreichend genau berechnet werden.

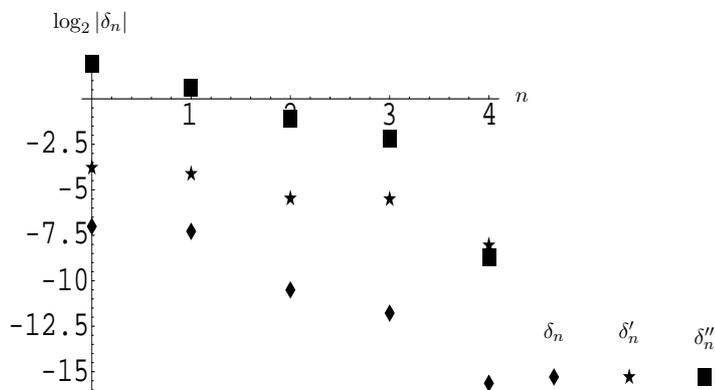


Abbildung 6.6: Punktweise Konvergenz der Kombinationstechnik vierter Ordnung, Amerikanische Option. Die numerischen Befunde sehen ähnlich aus wie beim Europäischen Gegenstück.

6 *Multivariate Extrapolation und eine Kombinationstechnik höherer Ordnung*

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

Überblick

Das letzte Kapitel hat gezeigt, dass durch die direkte Lösung von partiellen Differentialgleichungen auf dünnen Gittern Aufgabenstellungen moderater Dimensionalität ($d \leq 6$) behandelt werden können. Da wichtige Probleme aus der Praxis diesen Rahmen oft überschreiten, wird im Folgenden eine Methode entwickelt und untersucht, mit der hochdimensionale Probleme durch solche niedrigerer Dimension approximiert werden.

Nach einer kurzen Motivation anhand der Daten in 7.1 werden in Abschnitt 7.2 durch Analyse des Kontrakts und der Kovarianzmatrix Hauptkomponenten identifiziert. Die Ergebnisse dieser Dimensionsreduktion können durch einen asymptotischen Ansatz in den kleinen Eigenwerten noch wesentlich verbessert werden. 7.3 zeigt, wie auf diese Weise große Baskets aus einer Folge von zweidimensionalen numerischen Problemen — bzw. im Black-Scholes-Fall sogar analytisch — sehr genau extrapoliert werden können.

7.1 Motivation

Spektrum

Schon an den verwendeten realen Daten aus Kapitel 5 war augenfällig, dass die betrachteten Kurs-Zeitreihen stark korreliert sind. Dies ist besonders ausgeprägt, wenn man Werte aus gleichen Sektoren heranzieht. Betrachtet man beispielsweise einen typischen Aktienindex wie den DAX (siehe Tabelle 7.1), so sieht man Blöcke aus Technologiewerten, Pharma-Firmen, Finanzdienstleistern etc.

Es ist zu erwarten dass das System in erster Näherung auf solche Hauptkomponenten reduziert wird, wobei die wesentliche Dynamik durch den Basket selbst wiedergegeben wird. Es ist gängige Praxis, den Index auf diese einzige Komponente zu reduzieren. Weitere Anteile sollten sich zur besseren Approximation durch störungstheoretische Ansätze berücksichtigen lassen.

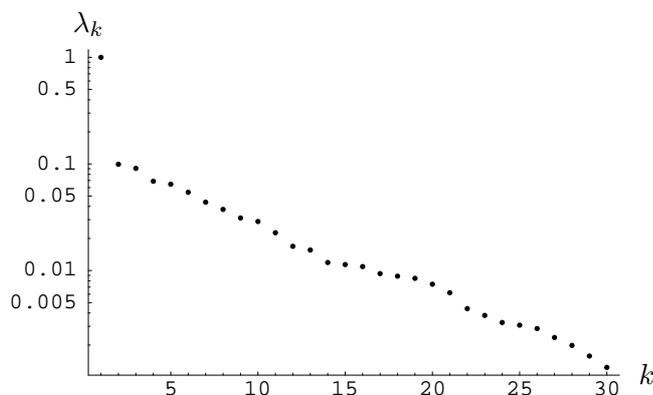


Abbildung 7.1: Abfallend geordnetes Spektrum $\{\lambda_k\}$ der Kovarianzmatrix der 30 Aktien im DAX vom 16. 1. 2003. Die Volatilitäten $\sigma_t = \sqrt{\frac{1}{T} \sum_{k=0}^{T-1} \varepsilon_{t-k}^2}$ und Korrelationen sind aus den täglichen logarithmischen Kurs-Inkrementen ε_t aus einem Zeitraum von $T = 250$ Tagen geschätzt. Der größte Eigenwert wurde auf 1 skaliert.

Abbildung 7.1 zeigt die 30 Eigenwerte der Kovarianzmatrix des DAX.¹ Der größte Eigen-

¹Dankenswerterweise bereitgestellt von Dr. Jürgen Linde aus der Gruppe Risk Control der Dresdner Bank

Aktie	Gewichtung	Kurs	rel. Vola.
Adidas-Salomon	45422500	82.60	0.3079
Allianz	266565625	59.71	0.5696
Altana	140400000	44.70	0.3304
BASF	571651410	38.30	0.4020
BMW	622227918	29.71	0.4377
Bayer	730341920	15.82	0.4906
Commerzbank	542206626	8.06	0.6234
Daimler Crysler	1012803481	29.25	0.5141
Deutsche Bank	621568446	45.65	0.5184
Deutsche Börse	111802880	10.55	0.3750
Deutsche Post	1112800000	10.55	0.4100
Deutsche Telekom	4197752425	12.10	0.5271
E.On	692000000	40.69	0.4688
Fresenius	70000000	48.00	0.6659
Henkel	59387625	57.11	0.1991
Hypo-Vereinsbank	521735101	9.62	0.6480
Infineon	720880604	7.36	0.7921
Linde	119262134	31.42	0.3427
Lufthansa	381600000	9.10	0.5203
MAN	140974350	15.30	0.5632
Metro	324109563	22.43	0.5179
MLP	108640686	7.89	0.8345
Münchner Rück	178667287	81.34	0.5301
RWE	523405000	23.30	0.3844
SAP	314895878	90.31	0.5501
Schering	196500000	39.76	0.3139
Siemens	890374001	43.75	0.5381
Thyssen Krupp	1223858305	6.98	0.6005
TUI	178038299	12.01	0.4624
Volkswagen	320289940	32.35	0.5430

Tabelle 7.1: DAX-Werte: Die Gewichtung bezeichnet die Anzahl der Aktien im DAX nach der Marktkapitalisierung vom 5. 2. 2003, die Kurse stammen vom 16. 4. 2003, die Volatilitäten wurden geschätzt aus 250 Tages-Kursen unmittelbar vor dem 16. 1. 2003.

wert, dessen Eigenvektor übrigens ausschließlich positive Einträge hat und daher in Richtung des Basket zeigt, unterscheidet sich um eine Größenordnung von den anderen, die — wie anhand der logarithmischen Skala deutlich wird — exponentiell abfallen.

In [LCPB99] wird die Problematik der Schätzung von Korrelationen großer Baskets am Beispiel des S&P500 diskutiert. Da die (symmetrische) Kovarianzmatrix in diesem Fall 12250 verschiedene Einträge hat, sind sehr lange Zeitreihen für eine verlässliche Schätzung insbesondere auch der kleinen Eigenwerte notwendig. Anhand von täglichen Daten aus dem Zeitraum 1991-1996 (1309 Tage) stellen Laloux et al. fest, dass abgesehen von den 10 größten Eigenwerten das Spektrum empirisch dem einer positiv definiten Zufallsmatrix gleicht, also keine relevanten Informationen enthält.

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

Diese Befunde liefern Argumente dafür, dass die relevanten Größen durch eine kleine Anzahl an Hauptkomponenten beschrieben werden können. Die folgenden Abschnitte widmen sich der Identifizierung dieser Koordinaten, der Asymptotik in den kleinen Eigenwerten sowie der quantitativen Analyse niedrigdimensionaler Approximationen.

Referenzergebnisse

Als Referenz für hochdimensionale Aufgaben werden die Ergebnisse von Monte-Carlo- und Quasi-Monte-Carlo-Verfahren herangezogen.² Der Preis, der aus hinreichend vielen ($N=40\,000\,000$) Durchläufen einer Monte Carlo Simulation für den Black-Scholes-Preis einer Europäischen Put-Option auf den DAX mit Laufzeit 1 Jahr, Zinssatz $r = 0.05$ und $K = 1$ *At-the-Money* bestimmt wurde, ist

$$u = 0.135323.$$

In Tabelle 7.2 ist die Konvergenzhistorie zusammen mit den Standardabweichungen σ und der CPU-Zeit t angeführt. Der Erwartungswert des Fehlers geht mit $N^{-\frac{1}{2}}$ gegen 0.

N	10^5	10^6	10^7	$2 \cdot 10^7$	$4 \cdot 10^7$
u	0.135083	0.135251	0.135343	0.135372	0.135323
$\sigma(u)$	0.169494	0.169487	0.169565	0.169576	0.169545
$\sigma(u)/\sqrt{N}$	5.36 e-4	1.69 e-4	5.36 e-5	3.79 e-5	2.68 e-5
t [sec]	7.7	76.9	768	1537	—

Tabelle 7.2: Konvergenz des Monte-Carlo-Verfahrens (MC) für den Put auf den DAX.

Neben dem vollen DAX wird noch einmal der schon in Abschnitt 5.5.2 numerisch bewertete Finanz-Basket betrachtet, für den der Vollständigkeit wegen in Tabelle 7.3 noch einmal die Daten zusammengetragen sind. Wie zu erwarten ist, sind die Finanzwerte stark korreliert,

Aktie	i	μ_i	σ_i	ρ_{ij}				
Deutsche Bank	1	38.1	0.518	1.00	0.79	0.82	0.91	0.84
Hypo-Vereinsbank	2	6.5	0.648	0.79	1.00	0.73	0.80	0.76
Commerzbank	3	5.7	0.623	0.82	0.73	1.00	0.77	0.72
Allianz	4	27.0	0.570	0.91	0.80	0.77	1.00	0.90
Münchner Rück	5	22.7	0.530	0.84	0.76	0.72	0.90	1.00

Tabelle 7.3: „Finanz-Basket“: Fünf Aktien aus der Finanzindustrie mit den angegebenen Volatilitäten σ_i und Korrelationen ρ_{ij} bilden mit den Gewichten μ_i den Referenzwert für eine Basket-Option.

das Spektrum

$$\{1.409, 0.113, 0.101, 0.0388, 0.0213\}$$

zeigt wieder einen dominanten Eigenwert. Der dazugehörige Eigenvektor

$$(0.41185, 0.49201, 0.46551, 0.45490, 0.40580)$$

²Freundlicherweise durchgeführt von Jürgen Schumacher vom Institut für Informatik der Universität Bonn

hat Komponenten fast gleicher Größe.

Der Preis einer Europäischen Put-Option auf diesen Basket, wieder mit Laufzeit 1 Jahr, Zinssatz $r = 0.05$ und $K = 1$ *At-the-Money*, ist

$$u = 0.175866,$$

wie aus den Ergebnissen der Quasi-Monte-Carlo-Simulation in Tabelle 7.4 hervorgeht.

	N	10^5	10^6	10^7	10^8	10^9	$2 \cdot 10^9$
MC	u	0.177141	0.176319	0.176167	0.175982	0.175897	0.17587
	$\sigma(u)$	0.206036	0.206749	0.207117	0.207086	0.207046	0.207051
	$\sigma(u)/\sqrt{N}$	6.52 e-4	2.07 e-4	6.55 e-5	2.07 e-5	6.55 e-6	4.63 e-6
	t [sec]	0.03	0.22	2.13	21.1	211	686
QMC	u	0.176039	0.175902	0.175871	0.175867	0.175866	0.175866
	t [sec]	0.05	0.54	5.47	57.6	576	1161

Tabelle 7.4: Konvergenz von Monte-Carlo-Verfahren (MC) und Quasi-Monte-Carlo-Verfahren (QMC, Halton-Sequenz) für eine Europäische Option auf den Finanz-Basket.

Greensche Funktion

Eine wichtige Folgerung aus der Transformation (1.21) in 1.3.3 ist die Darstellung des Black-Scholes-Preises einer Europäischen Option vermöge der Greenschen Funktion der Wärmeleitungsgleichung:

$$u(\mathbf{x}, t) = \frac{e^{-rt}}{(2t\pi)^{d/2} \sqrt{\lambda_1 \cdot \dots \cdot \lambda_d}} \int_{\mathbb{R}^d} g(\mathbf{x}') e^{-\frac{1}{2t} \sum_{i=1}^d (x'_i - x_i)^2 / \lambda_i} d\mathbf{x}', \quad (7.1)$$

wobei $x_i = \sum_{j=1}^d q_{ij} (\log S_j + t(r - \sigma_j^2/2))$ von den Ausgangsvariablen S_i abhängt und

$$g(\mathbf{x}) = \left(K - \sum_{i=1}^d \mu_i e^{\sum_{j=1}^d q_{ji} x_j} \right)_+ = \max \left(K - \sum_{i=1}^d \mu_i e^{\sum_{j=1}^d q_{ji} x_j}, 0 \right)$$

die Anfangsbedingung ist. Diese Integralform drückt die Lösung im Eigensystem der Kovarianzmatrix aus und wird als Ausgangspunkt für die Asymptotik dienen.

7.2 Hauptkomponentenanalyse

Zum Preisen von Basket-Optionen wird die Dynamik in der Praxis oft auf eine einzelne Komponente reduziert, nämlich die, die durch den Basket gegeben ist. Die Beobachtungen aus 7.1 legen hingegen nahe, den Eigenvektor zum größten Eigenwert zu betrachten. In beiden Fällen kann eine Black-Scholes-artige Formel als analytische Lösung des reduzierten eindimensionalen Problems herangezogen werden.

In diesem Abschnitt werden die beiden Zugänge verglichen. Manche Aussagen beziehen sich speziell auf den Fall positiv korrelierter Aktien, d. h. $\rho_{ij} > 0$, $1 \leq i, j \leq d$.

7.2.1 Das Basket-Modell

Einem Basket $\sum_{i=1}^d S_i$ von Aktien mit Volatilitäten σ_i entspricht eine *Basket-Volatilität*

$$\sigma_b = \frac{\sum_{i=1}^d \mu_i S_i \sigma_i}{\sum_{i=1}^d \mu_i S_i}, \quad (7.2)$$

vgl. [AG00]. Legt man ein derart reduziertes Modell zugrunde, ist die analytische Lösung für den Optionspreis (wie in 1.2.2, (1.4))

$$u(S, t) = e^{-rt} KN(-d_2) - SN(-d_1)$$

mit

$$d_1 = \frac{\ln S/K + t(r + 1/2\sigma_b^2)}{\sqrt{t}\sigma_b},$$

$$d_2 = \frac{\ln S/K + t(r - 1/2\sigma_b^2)}{\sqrt{t}\sigma_b}.$$

Für eine große Zahl stark korrelierter Aktien liefert das passable Näherungen.

Beispiele

- Finanz-Basket: Die *Basket-Vola* ist $\sigma_b = 0.3851$, der entsprechende Black-Scholes-Preis für den Put $u = 0.125818$, absoluter Fehler 0.050048, relativer Fehler 28.5%.
- DAX: $\sigma_b = 0.500668$, $u = 0.169406$, absoluter Fehler 0.0262967, relativer Fehler 19.4%.

7.2.2 Eigenmoden

Ein anderer Zugang zur Reduktion des Systems wird durch die Eigenwertzerlegung der Kovarianzmatrix motiviert (siehe schon Gleichung (1.21)):

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^d \lambda_i \frac{\partial^2 u}{\partial x_i^2} - ru \quad \forall (\mathbf{x}, t) \in \mathbb{R}^d \times (0, T), \quad (7.3)$$

$$u(\mathbf{x}, 0) = \left(K - \sum_{i=1}^d \mu_i e^{\sum_j q_j i x_j} \right)_+ \quad \forall \mathbf{x} \in \mathbb{R}^d. \quad (7.4)$$

Die Theorie von Perron und Frobenius zeigt eine gewisse Verknüpfung zwischen den zwei Wegen auf.

Satz 7.1 (Perron, Frobenius). Sei $\mathbb{R}_n^n \ni \mathbf{A} \geq \mathbf{0}$ irreduzibel, dann gilt:

1. $\lambda_1 := \varrho(\mathbf{A}) > 0$ ist einfacher Eigenwert und
2. zu λ_1 gehört ein positiver Eigenvektor $\mathbf{x}_1 > \mathbf{0}$.

Für einen Beweis siehe [Hac93]. In der Folge seien die Voraussetzungen von Satz 7.1 stets erfüllt.

Eindimensionale (analytische) Näherung

Lemma 7.2. Die Funktion $x : z_1 \rightarrow \sum_{i=1}^d \mu_i \exp(\sum_{j=1}^d q_{ji} z_j)$ ist bei $q_{1i} > 0$ für alle z_2, \dots, z_d streng monoton in z_1 , insbesondere injektiv.

Beweis.

$$\frac{\partial x}{\partial z_1} = \sum_{i=1}^d \mu_i q_{1i} \exp\left(\sum_{j=1}^d q_{ji} z_j\right) > 0$$

□

Für $\lambda_i \ll \lambda_1 \forall i > 1$ gehen wir aus von der Näherung

$$\frac{\partial u}{\partial t} = \frac{1}{2} \lambda_1 \frac{\partial^2 u}{\partial x^2} - ru \quad \forall (x, t) \in \mathbb{R} \times (0, T), \quad (7.5)$$

$$u(x, 0) = \left(K - \sum_{i=1}^d \tilde{\mu}_i e^{q_{1i} x}\right)_+ \quad \forall x \in \mathbb{R}. \quad (7.6)$$

Dabei ist $\tilde{\mu}_i = \mu_i \exp\left(\sum_{1 < j \leq d} q_{ji} x_j\right)$ mit fixen $x_j = x_j(\mathbf{S}, t)$ für $j > 1$. Der Zusammenhang zu den Ausgangsvariablen S_i wird über

$$S_i(t) = \exp\left(\sum_{j=1}^d q_{ji} x_j - t(r - \sigma_i^2/2)\right)$$

hergestellt. Die Lösung von Gleichung (7.5), (7.6) kann — im Wesentlichen wie im eindimensionalen Black-Scholes-Fall in 7.2.1 — analytisch mittels Faltung mit der Greenschen Funktion (7.1) durch die Fehlerintegrale

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\xi^2}{2}} d\xi$$

dargestellt werden ($\lambda_1 =: \sigma^2$):

$$u(x, t) = \frac{e^{-rt}}{\sqrt{2\pi t \sigma}} \int_{-\infty}^{\infty} \left(K - \sum_{i=1}^d \mu_i e^{q_{1i} x} e^{\sum_{j=2}^d q_{ji} x_j}\right)_+ e^{-\frac{(x-x_1)^2}{2t\sigma^2}} dx \quad (7.7)$$

$$= \frac{e^{-rt}}{\sqrt{2\pi t \sigma}} \int_{-\infty}^{\bar{x}} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)}\right)_+ e^{-\frac{(x-x_1)^2}{2t\sigma^2}} dx, \quad (7.8)$$

wobei \bar{x} der Punkt des Knicks ist, also die Gleichung

$$K = \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(\bar{x}-x_1)} \quad (7.9)$$

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

löst. Die Transformation $x = x_1 + \xi\sqrt{t}\sigma$ ($\bar{x} = x_1 + \bar{\xi}\sqrt{t}\sigma$) liefert

$$\begin{aligned} u(x, t) &= \frac{e^{-rt}}{\sqrt{2\pi}} \int_{-\infty}^{\bar{\xi}} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}\sqrt{t}\sigma\xi} \right) e^{-\frac{\xi^2}{2}} d\xi \\ &= e^{-rt} K \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\bar{\xi}} e^{-\frac{\xi^2}{2}} d\xi - \sum_{i=1}^d \mu_i S_i(0) \frac{e^{-rt}}{\sqrt{2\pi}} \int_{-\infty}^{\bar{\xi}} e^{-\frac{\xi^2}{2} + q_{1i}\sqrt{t}\sigma\xi} d\xi \\ &= e^{-rt} K N(\bar{\xi}) - \sum_{i=1}^d \mu_i S_i(0) e^{-rt} e^{\frac{q_{1i}^2 t \sigma^2}{2}} N(\bar{\xi} - q_{1i}\sqrt{t}\sigma). \end{aligned}$$

Aus

$$S_i(0) = e^{\sum_{j=1}^d q_{ji} x_j} = \exp\left(\sum_{j=1}^d q_{ji} x_j - t(r - \sigma_i^2/2)\right) e^{t(r - \sigma_i^2/2)} = S_i e^{t(r - \sigma_i^2/2)}$$

folgt

$$u(x, t) = e^{-rt} K N(\bar{\xi}) - \sum_{i=1}^d \mu_i S_i \exp\left(-rt + \frac{q_{1i}^2 t \lambda_1}{2} + t\left(r - \frac{\sigma_i^2}{2}\right)\right) N(\bar{\xi} - q_{1i}\sqrt{t\lambda_1}), \quad (7.10)$$

wobei $\bar{\xi}$ implizit durch

$$K = \sum_{i=1}^d \mu_i S_i \exp\left(q_{1i}\sqrt{t\lambda_1}\bar{\xi} + t\left(r - \frac{\sigma_i^2}{2}\right)\right) \quad (7.11)$$

bestimmt wird. Aus der Monotonie folgt wieder die eindeutige Lösbarkeit.

Beispiele

- Finanz-Basket: $u = 0.180648$, absoluter Fehler 0.004782, relativer Fehler 2.7%.
- DAX: $u = 0.149418$, absoluter Fehler 0.0140951, relativer Fehler 10.4%.

Mehrdimensionale numerische Näherung

Zur Berücksichtigung weiterer Hauptkomponenten betrachten wir approximative n -dimensionale Probleme, die dadurch entstehen, dass nur die ersten n Eigenwerte berücksichtigt werden, d. h. der Parameter-Vektor $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_d)$ wird ersetzt durch $\boldsymbol{\lambda}^{(n)}$, wobei

$$\lambda_i^{(n)} := \begin{cases} \lambda_i & 1 \leq i \leq n, \\ 0 & \text{sonst.} \end{cases} \quad (7.12)$$

Die entsprechende Lösung

$$u^{(n)}(t, \mathbf{x}) := u(t, \mathbf{x}, \boldsymbol{\lambda}^{(n)})$$

erfüllt die Anfangswertaufgabe

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^n \lambda_i \frac{\partial^2 u}{\partial z_i^2} - ru \quad \forall (\mathbf{z}, t) \in \mathbb{R}^n \times (0, T), \quad (7.13)$$

$$u(z_1, \dots, z_n, 0) = \left(K - \sum_{i=1}^d \tilde{\mu}_i e^{\sum_{j \leq n} q_{ji} z_j} \right)_+ \quad \forall \mathbf{z} \in \mathbb{R}^n. \quad (7.14)$$

Dabei ist $\tilde{\mu}_i = \mu_i \exp\left(\sum_{n < j \leq d} q_{ji} z_j\right)$ mit fixen $z_j = z_j(\mathbf{S}, t)$ für $j > n$.

Beispiel

Durch Diskretisierung auf dünnen Gittern (wie in den Beispielen aus Abschnitt 5.5) wurden für den Finanz-Basket die Werte aus Tabelle 7.5 mit hinreichender Genauigkeit bestimmt. Für hochdimensionale Probleme ist nicht zu erwarten, dass durch Hinzunahme weniger wei-

n	1	2	3	4	5
$u(S, T)$	0.1806	0.1796	0.1777	0.1764	0.1758
rel. Fehler	2.7 %	2.2 %	1.1 %	0.38 %	-

Tabelle 7.5: Näherungen der Dimension $n = 1, \dots, 4$ nach (7.13) und (7.14) für das fünfdimensionale Problem des Finanzbasket.

terer Komponenten eine wesentliche Verbesserung gegenüber der eindimensionalen Näherung erzielt werden kann.

7.2.3 Kombination

Bei der numerischen Behandlung des Systems (7.13), (7.14) treten zwei Probleme auf:

- die Abschneideproblematik, also die Begrenzung auf ein endliches Gebiet (verbunden mit der Setzung passender Randbedingungen) und
- die genaue Erfassung der nicht glatten Anfangsbedingung auf einem kartesischen Gitter, was die Approximationsordnung beschränkt.

Beides wird wesentlich erleichtert durch die Transformation

$$\begin{aligned} x_1 &= \sum_{i=1}^d \mu_i S_i = \sum_{i=1}^d \mu_i \exp\left(\sum_{j=1}^d q_{ji}(z_j - tb_j)\right) = \sum_{i=1}^d \tilde{\mu}_i(t) e^{q_{1i} z_i}, \\ x_k &= z_k \quad \forall k > 1, \end{aligned}$$

wobei wieder $b_i = \sum_{j=1}^d q_{ij}(r - \sigma_j^2/2)$ und $\tilde{\mu}_i(t) := \exp\left(\sum_{j=2}^d q_{ji} z_j - t(r - \sigma_i^2/2)\right)$. Die Umkehrbarkeit ist unter der Voraussetzung $q_{1i} > 0$ gegeben, wie man direkt aus der Gestalt der Jacobi-Matrix

$$\begin{pmatrix} \sum_{i=1}^d \tilde{\mu}_i(t) q_{1j} e^{q_{1i} z_i} & * & * & \cdots & * \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

sieht, da $\sum_{i=1}^d \tilde{\mu}_i q_{1j} e^{q_{1i} z_i} > 0$. Da nur x_1 nicht-trivial transformiert wird, ergibt sich

$$\frac{\partial u}{\partial z_k} = \frac{\partial u}{\partial x_1} \frac{\partial x_1}{\partial z_k} + \begin{cases} 0 & k = 1, \\ \frac{\partial u}{\partial x_k} & \forall k > 1 \end{cases}$$

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

und

$$\frac{\partial^2 u}{\partial z_k^2} = \frac{\partial^2 u}{\partial x_1^2} \left(\frac{\partial x_1}{\partial z_k} \right)^2 + \frac{\partial u}{\partial x_1} \frac{\partial^2 x_1}{\partial z_k^2} + \begin{cases} 0 & k = 1, \\ \frac{\partial^2 u}{\partial x_k^2} + \frac{\partial^2 u}{\partial x_1 \partial x_k} \frac{\partial x_1}{\partial z_k} & \forall k > 1. \end{cases}$$

Daher

$$\sum_{k=1}^d \lambda_k \frac{\partial^2 u}{\partial z_k^2} = \frac{\partial^2 u}{\partial x_1^2} \sum_{k=1}^d \lambda_k \left(\frac{\partial x_1}{\partial z_k} \right)^2 + \sum_{k>1} \lambda_k \frac{\partial x_1}{\partial z_k} \frac{\partial^2 u}{\partial x_1 \partial x_k} + \sum_{k>1} \lambda_k \frac{\partial^2 u}{\partial x_k^2} + \frac{\partial u}{\partial x_1} \sum_{k=1}^d \lambda_k \frac{\partial^2 x_1}{\partial z_k^2}.$$

Vom praktischen Standpunkt ist zu bemerken, dass die Umkehrtransformation $\mathbf{x} \rightarrow \mathbf{z}$ nicht analytisch ausgewertet werden kann, vielmehr muss man

$$z_1 \rightarrow \sum_{i=1}^d \tilde{\mu}_i(t) e^{q_{1i} z_1}$$

numerisch invertieren. Wegen Lemma 7.2 kann die Gleichung schnell mit ein paar (eindimensionalen) Newton-Schritten aufgelöst werden.

Die transformierte Gleichung hat die Gestalt

$$\frac{\partial u}{\partial t} - a_1(\mathbf{x}) \frac{\partial^2 u}{\partial x_1^2} - \sum_{k=1}^d a_k(\mathbf{x}) \frac{\partial^2 u}{\partial x_1 \partial x_k} - \frac{1}{2} \sum_{k=1}^d \lambda_k \frac{\partial^2 u}{\partial x_k^2} - b(\mathbf{x}) \frac{\partial u}{\partial x_1} + ru = 0, \quad (7.15)$$

wobei

$$a_1(\mathbf{x}, t) = \frac{1}{2} \sum_{i=1}^d \lambda_i \left(\frac{\partial x_1}{\partial z_i} \right)^2, \quad (7.16)$$

$$a_k(\mathbf{x}, t) = \frac{1}{2} \lambda_k \frac{\partial x_1}{\partial z_k} \quad 1 < k \leq d, \quad (7.17)$$

$$b(\mathbf{x}, t) = \frac{1}{2} \sum_{i=1}^d \lambda_i \frac{\partial^2 x_1}{\partial z_i^2} - \frac{\partial x_1}{\partial t} \quad (7.18)$$

und

$$\frac{\partial x_1}{\partial z_i} = \sum_{j=1}^d \mu_j q_{ij} e^{\sum_{k=1}^d q_{kj} (z_k - t b_k)}, \quad \frac{\partial^2 x_1}{\partial z_i^2} = \sum_{j=1}^d \mu_j q_{ij}^2 e^{\sum_{k=1}^d q_{kj} (z_k - t b_k)}.$$

Für a_1 sieht man

$$\begin{aligned} a_1(\mathbf{x}, t) &= \frac{1}{2} \sum_{i=1}^d \lambda_i \left(\frac{\partial x_1}{\partial z_i} \right)^2 = \frac{1}{2} \sum_{i=1}^d \lambda_i \left(\sum_{j=1}^d \mu_j q_{ij} S_j \right)^2 = \frac{1}{2} \sum_{j,k=1}^d \left(\sum_{i=1}^d q_{ij} q_{ik} \lambda_i \right) \mu_j \mu_k S_j S_k \\ &= \frac{1}{2} \sum_{j,k=1}^d \rho_{jk} \sigma_j \sigma_k \mu_j \mu_k S_j S_k. \end{aligned}$$

Darüberhinaus

$$\frac{\partial x_1}{\partial t} = - \sum_{i=1}^d \mu_i \left(r - \frac{1}{2} \sigma_i^2 \right) S_i$$

und

$$\sum_{i=1}^d \lambda_i \frac{\partial^2 x_1}{\partial z_i^2} = \sum_{i,j=1}^d \lambda_i \mu_j q_{ij}^2 S_j = \sum_{j=1}^d \mu_j \sigma_j^2 S_j,$$

also

$$b(\mathbf{x}, t) = r \sum_{j=1}^d \mu_j S_j.$$

Im Vergleich mit dem eindimensionalen Modell aus 7.2.1,

$$\frac{\partial u}{\partial t} - \frac{1}{2} \sigma_b^2(S_1, \dots, S_d) S^2 \frac{\partial^2 u}{\partial S^2} - r S \frac{\partial u}{\partial S} + ru = 0$$

mit $S = \sum_{i=1}^d \mu_i S_i$, identifiziert man eine andere (kleinere) *Basket-Volatilität* (die hier außerdem nicht als konstant betrachtet wird) und den Drift.

Der folgende Satz zeigt, dass das asymptotische Verhalten in Richtung x_1 dasselbe wie für die eindimensionale Black-Scholes-Gleichung ist.

Satz 7.3. Es existieren Konstanten \bar{a}_1 , \underline{a}_1 und \bar{a}_k ($k \geq 2$), sodass

1. $0 < \underline{a}_1 \leq \frac{a_1(\mathbf{x})}{x_1^2} \leq \bar{a}_1 \quad \forall x_k \quad \forall k > 1,$
2. $\frac{a_k(\mathbf{x}, t)}{x_1} \leq \bar{a}_k \quad \forall \mathbf{x} \quad \forall k \geq 1$ und $\lim_{x_k \rightarrow \infty} a_k(\mathbf{x}, t) < \infty \quad \forall x_j \quad \forall j > 1, j \neq k$ fix.

Die Grenzwerte hängen von x_j , $j \neq k$, ab.

Beweis.

1. Klarerweise erhält man wegen der positiven Definitheit der Kovarianzmatrix

$$\begin{aligned} \lambda_d x_1^2 &= \lambda_d \left(\sum_{k=1}^d \mu_k S_k(\mathbf{x}, t) \right)^2 \leq \\ &= \sum_{i,j=1}^d \mu_i \mu_j \rho_{ij} \sigma_i \sigma_j S_i(\mathbf{x}, t) S_j(\mathbf{x}, t) \\ &\leq \lambda_1 \left(\sum_{k=1}^d \mu_k S_k(\mathbf{x}, t) \right)^2 = \lambda_1 x_1^2. \end{aligned}$$

2. Sind überdies λ_1 und μ_i so skaliert, dass $\sum_{i=1}^d \mu_i = 1$ und $\lambda_1 = 1$, gilt

$$a_k(\mathbf{x}, t) = \lambda_k \sum_{j=1}^d \mu_j q_{kj} S_j(\mathbf{x}, t) \leq x_1$$

bzw. Äquivalentes durch Umskalierung. Der zweite Teil folgt sofort, wenn x_1 festgehalten wird.

□

Die Motivation für diese Wahl der Koordinaten war, die Richtungen schwacher Diffusion zu vernachlässigen. Zu diesem Zweck wird die Entwicklung nach einem passenden Index n abgeschnitten und die Koordinaten x_k für $k > n$ festgehalten. In den Raumrichtungen x_1, \dots, x_n wird die Gleichung

$$\frac{\partial u}{\partial t} - \sum_{k=1}^n a_k(\mathbf{x}, t) \frac{\partial^2 u}{\partial x_1 \partial x_k} - \frac{1}{2} \sum_{k=2}^n \lambda_k \frac{\partial^2 u}{\partial x_k^2} - b(\mathbf{x}) \frac{\partial u}{\partial x_1} + ru = 0 \quad \forall (\mathbf{x}, t) \in]0, \infty[\times \mathbb{R}^{n-1} \times (0, T) \quad (7.19)$$

mit Koeffizienten

$$a_1(\mathbf{x}, t) = \frac{1}{2} \sum_{i=1}^d \mu_i \mu_j \rho_{ij} \sigma_i \sigma_j S_i(\mathbf{x}, t) S_j(\mathbf{x}, t), \quad (7.20)$$

$$a_k(\mathbf{x}, t) = \frac{1}{2} \lambda_k \sum_{i=1}^d \mu_i q_{ki} S_i(\mathbf{x}, t) \quad 1 < k \leq n, \quad (7.21)$$

$$b(\mathbf{x}) = rx_1 \quad (7.22)$$

und Anfangsbedingung

$$u(\mathbf{x}, 0) = (K - x_1)_+$$

gelöst. Für $n = d$ erhält man natürlich das volle Modell zurück.

7.3 Asymptotik erster Ordnung

Sei u eine Lösung von (7.3), $u^{(n)}$ eine von (7.13). Dann ist

$$u(\mathbf{x}, t) = u^{(n)}(\mathbf{x}, t) + R(\lambda_{n+1} + \dots)$$

mit $R(\lambda_{n+1} + \dots) \rightarrow 0$ für $\lambda_i \rightarrow 0$, $i > n$.

Aus den vorangegangenen Beispielen ist ersichtlich, dass schon $n = 1$ eine gute Näherung liefert, während durch Hinzunahme einiger weiterer Richtungen kaum eine Verbesserung möglich ist. Die Komponenten, die zu $\lambda_2, \dots, \lambda_d$ gehören, können nicht ganz vernachlässigt werden und sollen in linearer Näherung mitgenommen werden.

Vorausgesetzt, dass die Lösung nach den Parametern differenzierbar ist, kann man entwickeln

$$u(\mathbf{x}, t) = u^{(n)}(\mathbf{x}, t) + \sum_{j=n+1}^d \lambda_j \left. \frac{\partial u}{\partial \lambda_j}(\mathbf{x}, t) \right|_{\boldsymbol{\lambda}^{(n)}} + R(\lambda_2^2 + \dots). \quad (7.23)$$

mit $\boldsymbol{\lambda}^{(n)}$ gemäß (7.12).

Entscheidend für eine solche Vorgehensweise ist die Regularität der Lösung in den Parametern, wobei natürlich insbesondere der singuläre Punkt $\lambda = 0$ von Interesse ist.

7.3.1 Analytische Asymptotik für Black-Scholes

Ausgangspunkt kann im Black-Scholes-Fall die Darstellung der Lösung mittels der Greenschen Funktion (7.1) sein.

Der eindimensionale Fall

Es ist illustrativ, die Asymptotik zunächst am eindimensionalen Fall zu studieren. Für $\epsilon := \sqrt{\lambda t}$ bekommt man

$$\begin{aligned} u_\epsilon(x, t) &= \frac{1}{\sqrt{2\pi\epsilon}} \int_{\mathbb{R}} g(x') e^{-\frac{1}{2}\left(\frac{x'-x}{\epsilon}\right)^2} dx' \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(x + s\epsilon) e^{-\frac{s^2}{2}} ds \end{aligned}$$

mit $x' = x + s\epsilon$.

Für stetige g , die gegen ∞ nicht zu schnell wachsen, gilt klarerweise

$$\lim_{\epsilon \rightarrow 0} u_\epsilon(x, t) = g(x). \quad (7.24)$$

Dies drückt lediglich die Konsistenz mit der Anfangsbedingung aus. Hier ist konkret

$$g(x) = \left(K - \sum_{i=1}^d \mu_i \exp\left(\sum_{j=1}^d q_{ij} x_j\right) e^{q_{1j}(x-x_1)} \right)_+,$$

sodass (7.24) erfüllt ist.

Ist g zweimal stetig differenzierbar in x , so gilt darüberhinaus

$$\begin{aligned} \frac{u_\epsilon(x, t) - g(x)}{\epsilon^2} &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{g(x + \epsilon s) - g(x)}{\epsilon^2} e^{-\frac{s^2}{2}} ds \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left[\frac{g'(x)s}{\epsilon} + \frac{1}{2} g''(x)s^2 \right] e^{-\frac{s^2}{2}} ds + R(\epsilon) \\ &= \frac{g'(x)}{\sqrt{2\pi\epsilon}} \underbrace{\int_{\mathbb{R}} s e^{-s^2} ds}_0 + \frac{1}{2\sqrt{2\pi}} g''(x) \underbrace{\int_{\mathbb{R}} s^2 e^{-\frac{s^2}{2}} ds}_{\sqrt{2\pi}} + R(\epsilon) \\ &= \frac{1}{2} g''(x) + R(\epsilon) \end{aligned}$$

mit $\lim_{\epsilon \rightarrow 0} R(\epsilon) = 0$.

Typisch für Optionspreisaufgaben sind Anfangsbedingungen, die stückweise glatt und dazwischen einseitig differenzierbar sind. In solchen Punkten x gilt

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{u_\epsilon(x, t) - g(x)}{\epsilon} &= \frac{1}{\sqrt{2\pi}} \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}} \frac{g(x + \epsilon s) - g(x)}{\epsilon} e^{-\frac{s^2}{2}} ds \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^0 g'(x-) s e^{-\frac{s^2}{2}} ds + \int_0^{\infty} g'(x+) s e^{-\frac{s^2}{2}} ds \right) \\ &= \frac{1}{\sqrt{2\pi}} (g'(x+) - g'(x-)). \end{aligned}$$

Für diese beiden praktisch interessierenden Fälle kann man schreiben

$$u(x, t) = g(x) + \lambda^\alpha \frac{\partial^\alpha}{\partial \lambda^\alpha} u_\lambda(x, t) \Big|_{\lambda=0} + R(\lambda)$$

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

mit $R(\lambda)/\lambda^\alpha \rightarrow 0$ ($\lambda \rightarrow 0$) sowie der gebrochenen Ableitung

$$\frac{\partial^\alpha}{\partial \lambda^\alpha} u_\lambda(x, t) := \lim_{\lambda \rightarrow 0} \frac{u_\lambda(x, t) - u_0(x, t)}{\lambda^\alpha},$$

wobei hier

$$\alpha = \begin{cases} 1 & \text{g zweimal differenzierbar in } \mathbf{x}, \\ \frac{1}{2} & \text{g stetig und einseitig differenzierbar.} \end{cases}$$

Bemerkung. Da λ nur in der Kombination $\epsilon = \sqrt{\lambda t}$ eingeht, gilt die gleiche Asymptotik auch für kleine Zeiten t . Insbesondere ist die Lösung am Knick zu $t = 0$ nicht nach t differenzierbar.

Der mehrdimensionale Fall

Für die Praxis relevant ist der mehrdimensionale Fall, wo ein Eigenwert die anderen dominiert. Wir halten daher λ_1 fest und betrachten die Asymptotik für kleine λ_k , $k \geq 2$. Diese Situation ist qualitativ anders, weil die Gleichung durch die kleinen Parameter nicht mehr singular gestört ist.

In der Tat gilt die folgende Entwicklung, deren Herleitung sich in Anhang A findet.

Satz 7.4 (Asymptotische Entwicklung). Sei $u(\mathbf{S}, t)$ der Black-Scholes-Preis eines Put auf einen Basket mit d Komponenten S_i mit Volatilitäten σ_i , Gewichten μ_i , Zins r , Ausübungspreis K und Laufzeit t . Die Eigenwerte der Kovarianzmatrix seien λ_i , die Eigenvektoren $\mathbf{q}_i = (q_{ij})_{1 \leq j \leq d}$, $1 \leq i \leq d$. Es wird angenommen, dass $\mathbf{q}_1 > \mathbf{0}$ und dass die Projektion $\bar{\mathbf{S}}$ von \mathbf{S} auf den Knick der Anfangsbedingung entlang des ersten Eigenvektors,

$$\sum_{j=1}^d q_{ij} (\log(\bar{S}_j/S_j) + t(r - \sigma_j^2/2)) = \begin{cases} \bar{\xi} \sqrt{\lambda_1 t} & i = 1, \\ 0 & 1 < i \leq d, \end{cases}$$

wobei $\bar{\xi}$ in (7.28) als Punkt des Knicks definiert ist, nicht orthogonal auf den k -ten Eigenvektor für all $k > 1$ steht:

$$\sum_{i=1}^d \mu_i \bar{S}_i q_{ki} \neq 0. \quad (7.25)$$

Dann gilt

$$u(\mathbf{S}, t) = u_0(\mathbf{S}, t) + t \sum_{k=2}^d \lambda_k u_k(\mathbf{S}, t) + o(\lambda_k) \quad (7.26)$$

mit $\lim_{\lambda \rightarrow 0} o(\lambda)/\lambda = 0$. Dabei sind

$$\begin{aligned} u_0(\mathbf{S}, t) &:= e^{-rt} KN(\bar{\xi}) - \sum_{i=1}^d \tilde{\mu}_i(t) S_i N(\bar{\xi} - q_{1i} \sqrt{t \lambda_1}), \\ u_k(\mathbf{S}, t) &:= -\frac{1}{2} \sum_{i=1}^d q_{ik}^2 \tilde{\mu}_i(t) S_i N(\bar{\xi} - q_{1i} \sqrt{\lambda_1 t}) + \frac{e^{-\frac{1}{2} \bar{\xi}^2}}{\sqrt{8\pi t \lambda_1}} \frac{\left(\sum_{i=1}^d \hat{\mu}_i(t) q_{ki} S_i \right)^2}{\sum_{i=1}^d \hat{\mu}_i(t) q_{1i} S_i}, \end{aligned} \quad (7.27)$$

wobei

$$\begin{aligned} \tilde{\mu}_i(t) &:= \mu_i \exp(q_{1i}^2 \lambda_1 t/2 - t \sigma_i^2/2), \\ \hat{\mu}_i(t) &:= \mu_i \exp(q_{1i} \sqrt{\lambda_1 t} \bar{\xi} - t \sigma_i^2/2) \end{aligned}$$

und $\bar{\xi}$ implizit (eindeutig) definiert ist durch

$$K = \sum_{i=1}^d \mu_i S_i \exp(q_{1i} \sqrt{t \lambda_1} \bar{\xi} + t(r - \sigma_i^2/2)). \quad (7.28)$$

Bemerkung. Die Näherungsformel (7.26) erfordert die numerische Auswertung der Verteilungsfunktion der Normalverteilung, die auch in der eindimensionalen Black-Scholes-Formel vorkommt, an d Stellen.

Dazu kommt hier die Bestimmung von $\bar{\xi}$, was der Rolle von $\ln K$ in der Black-Scholes-Formel entspricht. Aufgrund der Monotonie und Glattheit kann die Gleichung natürlich sehr schnell mit ein paar Newton-Schritten gelöst werden.

Beispiele

Es werden wieder die zwei Beispiele aus den vorangegangenen Abschnitten betrachtet. Die lineare Approximation bringt die folgende Verbesserung:

- Finanz-Basket: $u = 0.175254$, absoluter Fehler 6.1210^{-4} , relativer Fehler 0.35%.
- DAX: $u = 0.130133$, absoluter Fehler $u = 0.00519$, relativer Fehler 3.84%.

Zum späteren Vergleich schauen wir uns noch für das fünfdimensionale Beispiel die Störterme an.

Tabelle 7.6: Analytische Korrekturen u_k aus (7.27) zum 1-dimensionalen Modell (7.10) des 5d-Problems aus Abschnitt 7.2.2.

k	2	3	4	5
u_k	-0.0238	-0.0058	-0.0343	-0.0376
$\lambda_k u_k$	-2.68 e-3	-5.82 e-4	-1.33 e-3	-8.02 e-4

7.3.2 Verbesserte numerische Approximation

Eine Extrapolationsformel

Anstatt analytisch die Ableitung der Lösung nach den Eigenwerten auszurechnen und eine lineare Entwicklung in den λ_k aufzustellen, wird hier durch numerische Differentiation eine Approximationsformel gleicher Ordnung konstruiert. Ausgangspunkt dazu ist Gleichung (7.19). Im Sinne der Konsistenz mit der Asymptotik in 7.3.1 werden allerdings nicht die vereinfachten Koeffizienten (7.20)-(7.22) herangezogen, sondern in der Ausgangsform der Gleichungen (7.15)-(7.18), die nur im Eigensystem ausgedrückt ist, die höheren Eigenwerte vernachlässigt. Die resultierende Gleichung

$$\frac{\partial u}{\partial t} - a_1^{(n)}(\mathbf{x}, t) \frac{\partial^2 u}{\partial x_1^2} - \sum_{k=1}^n a_k^{(n)}(\mathbf{x}, t) \frac{\partial^2 u}{\partial x_1 \partial x_k} - \frac{1}{2} \sum_{k=1}^n \lambda_k \frac{\partial^2 u}{\partial x_k^2} - b^{(n)}(\mathbf{x}, t) \frac{\partial u}{\partial x_1} + ru = 0 \quad (7.29)$$

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

auf $\mathbb{R}^+ \times \mathbb{R}$ mit

$$\begin{aligned} a_1^{(n)}(\mathbf{x}, t) &= \frac{1}{2} \sum_{i=1}^n \lambda_i \left(\frac{\partial x_1}{\partial z_i} \right)^2, \\ a_k^{(n)}(\mathbf{x}, t) &= \frac{1}{2} \lambda_k \frac{\partial x_1}{\partial z_k} \quad \forall k > 1, \\ b^{(n)}(\mathbf{x}, t) &= \frac{1}{2} \sum_{i=1}^n \lambda_i \frac{\partial^2 x_1}{\partial z_i^2} - \frac{\partial x_1}{\partial t} \end{aligned}$$

ist äquivalent mit (7.13). Dabei deutet der Index n an, dass es sich um n -dimensionale Approximationen handelt. Die Lösung sei $u^{(n)}$.

Sei jetzt $\boldsymbol{\lambda}^{(n,j)}$ für $j > n$ definiert als abgeschnittener Parametervektor

$$\lambda_i^{(n,j)} := \begin{cases} \lambda_i & 1 \leq i \leq n \vee i = j, \\ 0 & \text{sonst} \end{cases}$$

und

$$u^{(n,j)}(\mathbf{x}, t) := u(\mathbf{x}, t, \boldsymbol{\lambda}^{(n,j)})$$

eine Lösung von

$$\begin{aligned} \frac{\partial u}{\partial t} - \sum_{k=1}^n a_k^{(n,j)}(\mathbf{x}, t) \frac{\partial^2 u}{\partial x_1 \partial x_k} - \frac{1}{2} \sum_{k=1}^n \lambda_k \frac{\partial^2 u}{\partial x_k^2} - b^{(n,j)}(\mathbf{x}, t) \frac{\partial u}{\partial x_1} + ru \\ - a_j^{(n,j)}(\mathbf{x}, t) \frac{\partial^2 u}{\partial x_1 \partial x_j} - \frac{1}{2} \lambda_j \frac{\partial^2 u}{\partial x_j^2} = 0 \quad \forall (\mathbf{x}, t) \in]0, \infty[\times \mathbb{R}^n \times (0, T), \end{aligned}$$

wobei die $a_k^{(n,j)}$ und $b^{(n,j)}$ aus a_k und b wieder so hervorgehen, dass alle $\lambda_i = 0$ gesetzt werden außer für $1 \leq i \leq n$ oder $i = j$. Dann gilt

$$\left. \frac{\partial u}{\partial \lambda_j}(\mathbf{x}, t) \right|_{\boldsymbol{\lambda}^{(n)}} = \frac{u^{(n,j)}(\mathbf{x}, t) - u^{(n)}(\mathbf{x}, t)}{\lambda_j} + R(\lambda_j)$$

mit $\lim_{\lambda_j \rightarrow 0} R(\lambda_j) = 0$ und Einsetzen in (7.23) ergibt die Näherungsformel

$$u(\mathbf{x}, t) = (1 - d + n)u^{(n)}(\mathbf{x}, t) + \sum_{j=n+1}^d u^{(n,j)}(\mathbf{x}, t) + o(\lambda_{n+1} + \dots).$$

mit $\lim_{\lambda_{n+1} \rightarrow 0} o(\lambda_{n+1} + \dots) = 0$. Der Clou dabei ist, dass die volle d -dimensionale Lösung bis auf Terme zweiter Ordnung durch ein n -dimensionales und $d - n$ Probleme der Dimension $n + 1$ approximiert wird. Die Lösung hat sozusagen eine niedrige *Superposition Dimension*, kann also als Summe von niedrigdimensionalen Funktionen dargestellt werden. Es müssen dazu nur entkoppelte Gleichungen gelöst werden.

Im Falle von Basket-Optionen, wo eine Komponente dominiert (Stichwort: *Spectral Gap*), reicht durchwegs $n = 1$, sodass nur zweidimensionale Probleme gelöst werden müssen.

Numerische Lösung

Bei der Diskretisierung dieser zweidimensionalen Probleme mit den in Kapitel 2 skizzierten oder anderen Methoden (z. B. auf dünnen Gittern) hängt der Approximationsfehler von 6 Faktoren ab: den Gitterweiten h_1 und h_2 in den beiden Richtungen, der Zeitschrittweite τ , aber auch den beiden Intervallen $[0, \bar{x}_1]$ und $[\underline{x}_2, \bar{x}_2]$, auf die das Gebiet begrenzt wird. Da $d-1$ Probleme mit völlig unterschiedlicher Skalierung zu lösen sind und es kaum möglich ist, systematisch insbesondere den Abschneidefehler in den Griff zu bekommen, ist die quantitative Kontrolle des Gesamtfehlers höchst mühsam.

Daher wird noch eine weitere Transformation $\xi : \mathbb{R}^+ \times \mathbb{R} \rightarrow [0, 1]^2$ auf das Einheitsquadrat nachgeschaltet, etwa

$$\xi_1 = \frac{x_1}{K + x_1}, \quad \xi_2 = \frac{1}{\pi} \arctan \lambda x_2 + \frac{1}{2}$$

mit einem $\lambda > 0$. Anders betrachtet hat man den Abschneidefehler mit dem Diskretisierungsfehler verknüpft, indem die Urbilder der „äußersten inneren Punkte“ mit jeder Verfeinerung weiter gegen ∞ wandern. Gleichzeitig hat man eine bestimmte Gradierung des Gitters in den Ausgangskordinaten festgelegt.

Beispiele

Um die numerisch genäherte Richtungsableitung entlang der Eigenvektoren mit dem analytischen, exakten Wert zu vergleichen, untersuchen wir zuerst ein zweidimensionales Beispiel. Hier sind die Eigenwerte $\lambda_1 = 0.431454$, $\lambda_2 = 0.024396$ erwartungsgemäß deutlich getrennt.

Aktie	i	μ_i	σ_i	$\rho_{ij}, 1 \leq j \leq 2$
BMW	1	0.384	0.438	1 0.890
Daimler	2	0.616	0.514	0.890 1

Tabelle 7.7: Daten für einen Basket aus einer BMW- und einer Daimler-Crysler-Aktie mit den entsprechenden Gewichten μ_i . Restliche Parameter wie sonst.

Der aktuelle Kurs wurde auf $(K, K) = (1, 1)$ skaliert.

Die analytische eindimensionale Näherung $u^{(1)} = 0.160787$ wird durch die lineare Näherung mit der Richtungsableitung $u_2 = 0.0858361$ nach Satz 7.4 auf 0.158693 korrigiert.

Die numerisch gewonnene, hinreichend (d. h. auf sechs Nachkommastellen) exakte Lösung für den aktuellen Kurs ist 0.158723. Die analytische Approximation weist demnach einen relativen Fehler von 0.019% auf.

Alternativ kann die Richtungsableitung durch numerische Differentiation approximiert werden. Tabelle 7.8 zeigt die Konvergenz des Differenzenquotienten zur Richtungsableitung.

k	3	2	1	0	-1	-2	-3
$u(\mathbf{x}, t, \delta_k)$	0.145869	0.152882	0.156718	0.158723	0.159748	0.160266	0.160526
$\frac{u(\mathbf{x}, t, \delta_k) - u_0}{\delta_k}$	-0.0764	-0.0810	-0.0834	-0.0846	-0.0853	-0.0856	-0.0859

Tabelle 7.8: Der kleinere Eigenwert wird — ausgehend vom obigen Beispiel — schrittweise halbiert bzw. verdoppelt: $\delta_k = \lambda_2 \cdot 2^{-k}$, $-3 \leq k \leq 3$. Aufgelistet sind die entsprechenden Lösungen sowie die genäherte Richtungsableitung für diese Folge von Schrittweiten.

7 Dimensionsreduktion für hochdimensionale parabolische Gleichungen

Nächstes Beispiel ist der Finanz-Basket. Die Korrekturen zur eindimensionalen Näherung sind in Tabelle 7.9 zusammengetragen.

k	2	3	4	5
$u^{(1,k)}$	0.179264	0.180095	0.178896	0.179344
$u^{(1,k)} - u_0$	-1.3810^{-3}	-5.5210^{-4}	-1.7510^{-3}	-1.0110^{-3}

Tabelle 7.9: Finanz-Basket: Numerische Korrekturen $u^{(1,k)} - u_0$ zum 1-dimensionalen Modell.

Im Vergleich zu Tabelle 7.6 zeigt sich vor allem für λ_2 eine dramatische Abweichung. Es stellt sich heraus, dass die numerische Approximation deutlich näher, ja sogar beeindruckend nahe bei der exakten Lösung liegt:

- Finanz-Basket: $u = 0.175972$, absoluter Fehler 0.000106, relativer Fehler 0.06%.
- DAX: $u = 0.135393$, absoluter Fehler 0.000073, relativer Fehler 0.05%.

7.4 Vergleich

Abschließend kann die Qualität der verschiedenen Näherungen für die zwei betrachteten Modellbeispiele mit fünf bzw. dreißig Dimensionen verglichen und am Aufwand bemessen werden.

Basiswert	Finanz-Basket ($d = 5$)		DAX ($d = 30$)	
exakter Preis	0.175866		0.135323	
	Lösung	rel. Fehler	Lösung	rel. Fehler
1d-Black-Scholes	0.125818	28.5 %	0.169406	19.4 %
1d-PCA	0.180648	2.7 %	0.149418	10.4 %
Asymptotik	0.175254	0.35 %	0.130133	3.84 %
2d-Extrapolation	0.175972	0.06 %	0.135393	0.05 %

Tabelle 7.10: Von den eindimensionalen Näherungen, die analytisch über Black-Scholes-artige Formeln ausgewertet werden können, liefert die Hauptkomponente der Kovarianzmatrix (aus 7.2.2) eine deutlich bessere Approximation als das univariate Black-Scholes-Modell für den Basket (siehe 7.2.1). Die analytische Entwicklung des Optionspreises (aus 7.3.1) nach den kleinen Eigenwerten bringt eine weitere signifikante Korrektur. Eine erstaunlich gute Approximation der Lösung ist schließlich durch Extrapolation aus Problemen möglich, die auf Ebenen durch den aktuellen Kurs definiert sind (siehe 7.3.2). Dies erfordert in d Dimensionen die Lösung einer eindimensionalen und von $d - 1$ zweidimensionalen Diffusionsgleichungen, was numerisch keinerlei Schwierigkeiten bereitet. Außerdem ist erfreulich, dass dieser Zugang nicht von der einfachen Gestalt der Anfangsbedingung, sondern nur von der Regularität der Lösung in den Parametern Gebrauch macht und daher auf andere, komplexe Kontrakte ausgedehnt werden kann. Die günstige Komplexität, die (mit der Zahl der Teilprobleme) nur linear in der Dimension wächst, zusammen mit der Möglichkeit der Bestimmung der Sensitivitäten ohne Mehraufwand sowie der Erweiterbarkeit auf Amerikanische Kontrakte, liefert deutliche Vorteile gegenüber stochastischer Simulation.

Diskussion und Ausblick

Diskussion und Ausblick

Die numerischen Experimente dieser Arbeit entwerfen das folgende Bild: Obwohl die in der numerischen Analysis aus Kapitel 4 nachgewiesene überlegene ϵ -Komplexität dünner Gitter im Vergleich zu herkömmlichen Diskretisierungen eindrucksvoll bestätigt wird, ist die direkte numerische Lösung von Optionspreisaufgaben mit mehr als sechs Faktoren im allgemeinen Fall, d. h. ohne Ausnutzung einer speziellen Struktur der Lösung, selbst durch massiv parallele Rechnung nicht mit sinnvoller Genauigkeit möglich. Es ist zu erwarten, dass durch die Wahl von Volumenfreiheitsgraden (etwa finite Volumen oder unstetige finite Elemente) die Grenze eventuell minimal nach oben verschoben werden kann.

Bei Aufgaben noch höherer Dimension kommt man nicht umhin, unter Ausnutzung der speziellen Struktur des Modells eine analytische Reduktion durchzuführen. Dies wurde anhand der Black-Scholes-Gleichung exerziert und hat gezeigt, dass in diesem Fall für praktisch relevante Fragen Monte-Carlo-Verfahren ausgestochen werden können. Diese Fortschritte fordern heraus, die Methodik weiterzuentwickeln und auf eine weitere Modellklasse und Produktpalette sowie über den Finanzbereich hinaus zu verallgemeinern.

Das Prinzip der Kombinationstechnik ist — abgesehen von der starken geometrischen Beschränkungen des Gebietes, im wesentlichen auf Würfel — relativ allgemein anwendbar und nicht auf lineare parabolische Probleme beschränkt. Vielmehr ist unter der Voraussetzung der in Kapitel 4 diskutierten Fehlerasymptotik auf den zugrundeliegenden Tensorproduktgittern eine direkte Anwendung auf nichtlineare Modelle (wie etwa das Leland-Modell zur Berücksichtigung von Transaktionskosten) und sogar auf partielle Integro-Differentialgleichungen (wie sie zum Beispiel von Jump-Diffusion-Prozessen herrühren) denkbar.

Im Bezug auf die Dimensionsreduktion war das betrachtete Black-Scholes-Modell insofern speziell, als eine Transformation auf eine Gleichung mit konstanten Koeffizienten existiert, die einerseits eine einfache (analytische) Beschreibung des Drift ergibt und andererseits mit der Wahl einer Eigenvektorbasis des Diffusionstensors direkt eine Hauptkomponentenanalyse ermöglicht. Im allgemeinen Fall eines parabolischen Operators mit variablen Koeffizienten existieren noch immer lokale orthogonale Koordinaten, welche die Diffusion diagonalisieren, während sich die Terme erster Ordnung über Charakteristiken, die zwar im hochdimensionalen Raum verlaufen, aber durch gewöhnliche Differentialgleichungen der entsprechenden Dimension bestimmt werden, berücksichtigen lassen. Eine derartige Weiterentwicklung wäre insbesondere im Hinblick auf Modelle mit lokaler oder stochastischer Volatilität von essenzieller Bedeutung.

Im Übrigen treten hochdimensionale Optionspreisaufgaben nicht nur bei Aktien-Baskets auf. Im Risikomanagement wesentlich sind Zinsderivate, wobei hier dem LIBOR-Marktmodell eine große praktische Bedeutung zukommt und Dimensionen in der Größenordnung von 80 alltäglich sind. Doch abermals sind die LIBORen stark korreliert, sodass unter gängigen Modellvereinfachungen die in Kapitel 7 diskutierten Methoden zur Dimensionsreduktion greifen.

Über den Rahmen der Optionsbewertung und Finanzmathematik hinaus können die hier vorgestellten Techniken als Ausgangspunkt für eine spannende Entwicklung analytischer und numerischer Methoden für hochdimensionale Aufgaben mit unterschiedlichem Hintergrund dienen. Als Beispiel seien neben der Beschreibung quantenmechanischer Mehr-Teilchen-Systeme über die u. U. hochdimensionale Schrödingergleichung Fragestellungen der Populationsdynamik erwähnt, wo die Kopplung von intern strukturierten Zustandsräumen (etwa Masse, Alter, Nährstoffgehalt etc. von Individuen) mit externer (räumlicher) Strukturierung in einem turbulenten Rührkessel zu ausgesprochen anspruchsvollen numerischen Aufgaben führt.

Diskussion und Ausblick

A Herleitung der asymptotischen Entwicklung

Es wird hier der Beweis von Satz 7.4 nachgeliefert, der die Entwicklung des Black-Scholes-Preises eines Basket-Put nach den Eigenwerten der Kovarianzmatrix angibt.

Entscheidend für die analytische Integration ist die Nullmenge der Anfangsbedingung, die durch Lemma A.1 beschrieben wird.

Lemma A.1. *Sei*

$$f(x, y) := \sum_{i=1}^d a_i(x) e^{b_i y} - c$$

mit $a_i : \mathbb{R} \rightarrow]0, \infty[$ stetig, streng monoton wachsend und surjektiv, nicht alle b_i haben gleiches Vorzeichen und $c > 0$. Dann gilt:

$$\exists x_0 \in \mathbb{R} : \begin{cases} f(x, \cdot) \text{ hat genau zwei Nullstellen } \underline{y}(x) < \bar{y}(x) \text{ für } x \in]-\infty, x_0[, \\ f(x_0, \cdot) \text{ hat genau eine Nullstelle und} \\ f(x, \cdot) \text{ hat keine Nullstellen für } x \in]x_0, \infty[. \end{cases}$$

Für $x < x_0$ und $\underline{y}(x) < y < \bar{y}(x)$ ist $f(x, y) < 0$, sonst ist überall $f(x, y) \geq 0$. Ferner gilt $\lim_{x \rightarrow -\infty} \underline{y}(x) = -\infty$, $\lim_{x \rightarrow -\infty} \bar{y}(x) = \infty$.

Beweis. Aus

$$\lim_{y \rightarrow \pm\infty} f(x, y) = \infty$$

(da $\forall i : a_i > 0$ und $\exists i, j : b_i > 0, b_j < 0$) und

$$\frac{\partial^2 f}{\partial y^2} = \sum_{i=1}^d a_i(x) b_i^2 e^{b_i y} > 0$$

folgt, dass $\frac{\partial f}{\partial y}(x, \cdot)$ für alle x genau eine Nullstelle, also $f(x, \cdot)$ maximal zwei Nullstellen hat. Für

$$h(x) := \min_{y \in \mathbb{R}} f(x, y)$$

gilt

$$\begin{aligned} \lim_{x \rightarrow -\infty} h(x) &= -c, \\ \lim_{x \rightarrow \infty} h(x) &= \infty, \end{aligned}$$

und h ist streng monoton wachsend, da für $x_1 \neq x_2$ und $y_1 = \operatorname{argmin} f(x_1, \cdot)$, $y_2 = \operatorname{argmin} f(x_2, \cdot)$ aus $h(x_1) \geq h(x_2)$

$$f(x_1, y_1) = h(x_1) \geq h(x_2) = f(x_2, y_2) > f(x_2, y_1)$$

und daraus (wegen der strengen Monotonie von $f(\cdot, y_1)$) $x_1 > x_2$ folgt. Die Nullstelle x_0 von h hat die gewünschte Eigenschaft. \square

A Herleitung der asymptotischen Entwicklung

Beweis für Satz 7.4, Asymptotische Entwicklung. Gleichung (7.26) stellt eine lineare Näherung der Black-Scholes-Lösung (7.1) in den Eigenwerten $\lambda_2, \dots, \lambda_d$ dar. Der entscheidende Schritt ist daher die Berechnung der Ableitungen von (7.1) in die Richtungen λ_k , $k > 1$, ausgewertet für $\lambda_2 = \dots = \lambda_d = 0$. Diese Richtungsableitungen können aus der Gleichung mit $\lambda_i = 0$ für $i \notin \{1, k\}$, also zweidimensionalen Problemen auf den x_1 - x_k -Ebenen gewonnen werden. Bei festen x_i , $i \notin \{1, k\}$, erhält man für die Lösungen dieser Gleichungen aus (7.1) die Darstellung

$$u^{(1,k)}(\mathbf{x}, t) = \frac{e^{-rt}}{2t\pi\sqrt{\lambda_1\lambda_k}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(K - \sum_{i=1}^d \mu_i e^{q_{1i}x} e^{q_{ki}y} e^{\sum_{j \notin \{1,k\}} q_{ji}x_j} \right)_+ e^{-\frac{1}{2t} \left[\frac{(x-x_1)^2}{\lambda_1} + \frac{(y-x_k)^2}{\lambda_k} \right]} dy dx.$$

Interessant ist das Verhalten für kleine Parameter $\epsilon := \sqrt{t\lambda_k}$, und um diese Abhängigkeit auszudrücken wird die Lösung im Folgenden als $u_\epsilon^{(k)}$ bezeichnet. Mit $\sigma := \sqrt{\lambda_1}$ und mit der Beziehung zu den Ausgangsvariablen

$$S_i(t) = \exp \left(\sum_{j=1}^d q_{ji}x_j - t \left(r - \frac{\sigma_i^2}{2} \right) \right)$$

ergibt sich

$$u_\epsilon^{(k)}(\mathbf{x}, t) = \frac{e^{-rt}}{2\pi\sqrt{t}\sigma\epsilon} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}(y-x_k)} \right)_+ e^{-\frac{1}{2} \left[\frac{(x-x_1)^2}{t\sigma^2} + \frac{(y-x_k)^2}{\epsilon^2} \right]} dy dx.$$

Wegen Lemma A.1, angewandt auf

$$f(x, y) := \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}(y-x_k)} - K$$

(wegen $\mathbf{q}_1 > \mathbf{0}$ und $\mathbf{q}_k \perp \mathbf{q}_1$ sind die dortigen Voraussetzungen erfüllt), kann man das Integral folgendermaßen auf den Träger des Integranden, der (mit den Bezeichnungen aus Lemma A.1) „links“ von x_0 und dort jeweils zwischen den beiden Nullstellen $\underline{y}(x)$ und $\bar{y}(x)$ liegt, beschränken:

$$\begin{aligned} u_\epsilon^{(k)}(\mathbf{x}, t) &= \frac{e^{-rt}}{2\pi\sqrt{t}\sigma\epsilon} \int_{-\infty}^{x_0} \int_{\underline{y}(x)}^{\bar{y}(x)} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}(y-x_k)} \right)_+ e^{-\frac{1}{2} \left[\frac{(x-x_1)^2}{t\sigma^2} + \frac{(y-x_k)^2}{\epsilon^2} \right]} dy dx \\ &= \frac{e^{-rt}}{2\pi\sqrt{t}\sigma} \int_{-\infty}^{x_0} \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}\epsilon\eta} \right)_+ e^{-\frac{1}{2} \left[\frac{(x-x_1)^2}{t\sigma^2} + \eta^2 \right]} d\eta dx, \quad (\text{A.1}) \end{aligned}$$

wobei η durch $y = x_k + \eta\epsilon$ definiert ist, $\underline{\eta}_\epsilon(x) := (\underline{y}(x) - x_k)/\epsilon$ und $\bar{\eta}_\epsilon(x) := (\bar{y}(x) - x_k)/\epsilon$.

Die eindimensionale Lösung für $\lambda_2 = \dots = \lambda_d = 0$ hat die Darstellung

$$u_0(\mathbf{x}, t) = \frac{e^{-rt}}{\sqrt{2\pi t}\sigma} \int_{-\infty}^{\bar{x}} \underbrace{\left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} \right)}_{=:g(x)} e^{-\frac{1}{2t} \frac{(x-x_1)^2}{\sigma^2}} dx$$

(siehe (7.7) bis (7.9) und das dort eingeführte \bar{x}). Es gilt

$$\bar{x} = \max\{x : g(x) \geq 0\} = \max\{x : f(x, x_k) \leq 0\} \leq \max\{x : \exists y : f(x, y) \leq 0\} = x_0.$$

Der Fall $\bar{x} = x_0$ impliziert

$$\frac{\partial f}{\partial y}(\bar{x}, x_k) = \sum_{i=1}^d \mu_i S_i(0) q_{ki} e^{q_{1i}(\bar{x}-x_1)} = 0,$$

da $f(\bar{x}, \cdot)$ in x_k dann ein Minimum hat. Im Folgenden wird angenommen, dass dies nicht der Fall ist, was mit Bedingung (7.25) äquivalent ist.

Für $\bar{x} < x_0$ kann man das Integral über x in (A.1) aufspalten in $] -\infty, \bar{x}]$ und $[\bar{x}, x_0]$, um die gleiche Grenze wie bei u_0 zu bekommen, und erhält

$$\begin{aligned} \frac{u_\epsilon^{(k)} - u_0}{\epsilon^2} &= \frac{e^{-rt}}{\sqrt{2\pi t \sigma \epsilon^2}} \int_{-\infty}^{\bar{x}} \left\{ \frac{1}{\sqrt{2\pi}} \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}\epsilon\eta} \right) e^{-\frac{\eta^2}{2}} d\eta \right. \\ &\quad \left. - \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} \right) \right\} e^{-\frac{1}{2t} \frac{(x-x_1)^2}{\sigma^2}} dx \\ &\quad \underbrace{\hspace{15em}}_{=: A_\epsilon(\mathbf{S}, t)} \\ &+ \frac{e^{-rt}}{2\pi \sqrt{t \sigma \epsilon^2}} \int_{\bar{x}}^{x_0} \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}\epsilon\eta} \right) e^{-\frac{1}{2} \left[\frac{(x-x_1)^2}{t\sigma^2} + \eta^2 \right]} d\eta dx. \\ &\quad \underbrace{\hspace{15em}}_{=: B_\epsilon(\mathbf{S}, t)} \end{aligned} \quad (\text{A.2})$$

Widmen wir uns zunächst B_ϵ und führen die Integration über η aus:

$$\int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} e^{-\frac{1}{2}\eta^2} d\eta = \sqrt{2\pi} \left[N(\bar{\eta}_\epsilon(x)) - N(\underline{\eta}_\epsilon(x)) \right],$$

$$\int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} e^{q_{ki}\epsilon\eta} e^{-\frac{1}{2}\eta^2} d\eta = \sqrt{2\pi} e^{\frac{1}{2}\epsilon^2 q_{ki}^2} \left[N(\bar{\eta}_\epsilon(x) - \epsilon q_{ki}) - N(\underline{\eta}_\epsilon(x) - \epsilon q_{ki}) \right].$$

Wegen $f(\bar{x}, x_k) = 0$ ist entweder $\underline{y}(\bar{x}) = x_k$ oder $\bar{y}(\bar{x}) = x_k$.

Wir betrachten zuerst den Fall $\bar{y}(\bar{x}) = x_k$. Dann gilt für $\bar{x} \leq x \leq x_0$: $\underline{y}(x) < x_k$ und daher

$$\lim_{\epsilon \rightarrow 0} N(\underline{\eta}_\epsilon(x)) = \lim_{\epsilon \rightarrow 0} N\left(\frac{\underline{y}(x) - x_k}{\epsilon}\right) = 0 \text{ sowie } \lim_{\epsilon \rightarrow 0} N(\underline{\eta}_\epsilon(x) - \epsilon q_{ki}) = 0$$

gleichmäßig in x . Zu untersuchen bleibt daher

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \int_{\bar{x}}^{x_0} \left(KN(\bar{\eta}_\epsilon(x)) - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{\frac{1}{2}\epsilon^2 q_{ki}^2} N(\bar{\eta}_\epsilon(x) - \epsilon q_{ki}) \right) e^{-\frac{1}{2} \frac{(x-x_1)^2}{t\sigma^2}} dx \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \int_{\bar{x}}^{x_0} N\left(\frac{\bar{y}(x) - x_k}{\epsilon}\right) \underbrace{\left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} \right)}_{=: g(x)} e^{-\frac{1}{2} \frac{(x-x_1)^2}{t\sigma^2}} dx \\ &\quad \underbrace{\hspace{15em}}_{=: h(x)} \end{aligned} \quad (\text{A.3})$$

$$= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{z_\epsilon}^0 N(z) \frac{h(\bar{y}^{-1}(x_k + \epsilon z))}{\bar{y}'(\bar{y}^{-1}(x_k + \epsilon z))} dz \quad (\text{A.4})$$

A Herleitung der asymptotischen Entwicklung

mit der Transformation $z := (\bar{y}(x) - x_k)/\epsilon$, $x = \bar{x} \Leftrightarrow z = 0$, $x = x_0 \Leftrightarrow z = z_\epsilon := (\bar{y}(x_0) - x_k)/\epsilon < 0$ sowie

$$\begin{aligned} x &= \bar{y}^{-1}(x_k + \epsilon z), \\ dx &= \epsilon \bar{y}'(\bar{y}^{-1}(x_k + \epsilon z))^{-1} dz. \end{aligned}$$

Man sieht der Reihe nach

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \bar{y}'(\bar{y}^{-1}(x_k + \epsilon z)) &= \bar{y}'(\bar{x}), \\ \lim_{\epsilon \rightarrow 0} h(\bar{y}^{-1}(x_k + \epsilon z)) &= 0 \quad (\text{da } g(\bar{x}) = 0), \\ \lim_{\epsilon \rightarrow 0} \frac{h(\bar{y}^{-1}(x_k + \epsilon z))}{\epsilon} &= e^{-\frac{1}{2} \frac{(x_1 - \bar{x})^2}{t\sigma^2}} \lim_{\epsilon \rightarrow 0} \frac{g(\bar{y}^{-1}(x_k + \epsilon z))}{\epsilon} \\ &= e^{-\frac{1}{2} \frac{(x_1 - \bar{x})^2}{t\sigma^2}} z \frac{g'(\bar{x})}{\bar{y}'(\bar{x})}, \end{aligned}$$

und da $z_\epsilon \rightarrow -\infty$ für $\epsilon \rightarrow 0$, erhält man im Grenzwert für (A.4)

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{z_\epsilon}^0 N(z) \frac{h(\bar{y}^{-1}(x_k + \epsilon z))}{\bar{y}'(\bar{y}^{-1}(x_k + \epsilon z))} dz = e^{-\frac{1}{2} \frac{(x_1 - \bar{x})^2}{t\sigma^2}} \frac{g'(\bar{x})}{\bar{y}'(\bar{x})} \int_{-\infty}^0 z N(z) dz,$$

somit

$$\lim_{\epsilon \rightarrow 0} B_\epsilon(\mathbf{S}, t) = \frac{e^{-rt}}{\sqrt{2\pi t \lambda_1}} e^{-\frac{1}{2} \frac{(x_1 - \bar{x})^2}{t\sigma^2}} \frac{g'(\bar{x})}{\bar{y}'(\bar{x})} \int_{-\infty}^0 z N(z) dz. \quad (\text{A.5})$$

Wir bleiben beim Fall $\bar{y}(x) = x_k$ und betrachten A_ϵ aus (A.2). Für konstantes c gilt

$$\begin{aligned} c &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c e^{-\frac{\eta^2}{2}} d\eta \\ &= \frac{c}{\sqrt{2\pi}} \int_{-\infty}^{\bar{\eta}_\epsilon(x)} e^{-\frac{\eta^2}{2}} d\eta + \frac{1}{\sqrt{2\pi}} \int_{\bar{\eta}_\epsilon(x)}^{\infty} c e^{-\frac{\eta^2}{2}} d\eta + \frac{c}{\sqrt{2\pi}} \int_{\bar{\eta}_\epsilon(x)}^{\infty} e^{-\frac{\eta^2}{2}} d\eta. \end{aligned} \quad (\text{A.6})$$

Mit der Wahl $c := K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)}$ kann der mittlere Teil aus (A.6) unter das η -Integral von A_ϵ in (A.2) gebracht werden. Der erste Summand in (A.6) führt auf

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} N\left(\frac{\underline{y}(x) - x_k}{\epsilon}\right)$$

und geht auf $]-\infty, \bar{x}]$ gleichmäßig (in x) gegen 0, da $\underline{y}(x) - x_k < 0$. Der letzte Term, eingesetzt in (A.2), ergibt

$$-\frac{e^{-rt}}{\sqrt{2\pi t \sigma \epsilon^2}} \int_{-\infty}^{\bar{x}} \left(K - \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} \right) e^{-\frac{1}{2t} \frac{(x-x_1)^2}{\sigma^2}} N\left(\frac{x_k - \bar{y}(x)}{\epsilon}\right) dx,$$

da $1 - N(\bar{\eta}_\epsilon(x)) = N(-\bar{\eta}_\epsilon(x))$. Dieser Ausdruck führt wie oben (A.3) im Grenzübergang auf (A.5) mit umgekehrtem Vorzeichen.

Analog ergibt sich im Fall $\underline{y}(\bar{x}) = x_k$, dass sich die Summe aus erstem und letztem Term in (A.6), eingesetzt in (A.2), mit $B_\epsilon(\mathbf{S}, t)$ wegheben.

Zusammen ergibt das

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{u_\epsilon^{(k)} - u_0}{\epsilon^2} &= \lim_{\epsilon \rightarrow 0} \frac{e^{-rt}}{2\pi\sqrt{t}\sigma\epsilon^2} \int_{-\infty}^{\bar{x}} \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} (1 - e^{q_{ki}\epsilon\eta}) e^{-\frac{1}{2t} \frac{(x-x_1)^2}{\sigma^2} - \frac{\eta^2}{2}} d\eta dx \\ &=: \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \int_{-\infty}^{\bar{x}} \sum_{i=1}^d h_i(x) \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} (1 - e^{q_{ki}\epsilon\eta}) e^{-\frac{\eta^2}{2}} d\eta dx, \end{aligned} \quad (\text{A.7})$$

wobei die letzte Gleichung als Definition der h_i fungiert.

Aus der Taylor-Entwicklung von exp,

$$\exp(\epsilon q_{ki}\eta) = 1 + \epsilon q_{ki}\eta + \frac{\epsilon^2 q_{ki}^2 \eta^2}{2} + \frac{\epsilon^3 q_{ki}^3 \eta^3}{6} e^{\theta \epsilon q_{ki}\eta}$$

mit $\theta \in [0, 1]$, ergibt sich

$$\begin{aligned} \frac{1 - e^{\epsilon q_{ki}\eta}}{\epsilon^2} &= -\frac{q_{ki}}{\epsilon} \eta - \frac{q_{ki}^2}{2} \eta^2 + R_\epsilon(\eta), \\ |R_\epsilon(\eta)| &\leq \epsilon \frac{q_{ki}^3}{6} \eta^3 \exp(\epsilon q_{ki}\eta). \end{aligned} \quad (\text{A.8})$$

Die Terme in (A.8) werden der Ordnung von ϵ nach untersucht:

Ordnung ϵ^{-1} :

$$\frac{1}{\epsilon} \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \eta e^{-\frac{\eta^2}{2}} d\eta = \frac{e^{-\frac{1}{2}\eta_\epsilon(x)^2} - e^{-\frac{1}{2}\bar{\eta}_\epsilon(x)^2}}{\epsilon}$$

und damit in (A.7)

$$\begin{aligned} -\lim_{\epsilon \rightarrow 0} \sum_{i=1}^d q_{ki} \int_{-\infty}^{\bar{x}} h_i(x) \frac{1}{\epsilon} \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \eta e^{-\frac{\eta^2}{2}} d\eta dx &= \\ &= \sum_{i=1}^d q_{ki} \int_{-\infty}^{\bar{x}} h_i(x) \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left[e^{-\frac{1}{2} \left(\frac{\bar{\eta}_\epsilon(x)}{\epsilon} \right)^2} - e^{-\frac{1}{2} \left(\frac{\underline{\eta}_\epsilon(x)}{\epsilon} \right)^2} \right] dx. \end{aligned}$$

Im Fall $\boxed{\bar{y}(\bar{x}) = x_k}$ liefert nur der erste Term in den eckigen Klammern im Grenzwert einen Beitrag, es gilt

$$\frac{1}{\epsilon} \int_{-\infty}^{\bar{x}} h_i(x) e^{-\frac{1}{2} \left(\frac{\bar{y}(x) - x_k}{\epsilon} \right)^2} dx = \int_{-\infty}^0 \frac{h_i(\bar{y}^{-1}(x_k + \epsilon z))}{\bar{y}'(\bar{y}^{-1}(x_k + \epsilon z))} e^{-\frac{1}{2} z^2} dz \rightarrow \frac{\sqrt{2\pi} h_i(\bar{x})}{2 \bar{y}'(\bar{x})}$$

für $\epsilon \rightarrow 0$. Dabei ist

$$h_i(x) = \frac{e^{-rt}}{2\pi\sqrt{t}\lambda_1} \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{-\frac{1}{2t} \frac{(x-x_1)^2}{\lambda_1^2}}$$

und $\bar{y}'(\bar{x})$ errechnet sich aus

$$\begin{aligned} f(x, \bar{y}(x)) &= \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}(\bar{y}(x) - x_k)} - K = 0 \quad \forall x \in] - \infty, x_0[\Rightarrow \\ \frac{\partial f(x, \bar{y}(x))}{\partial x} &= \sum_{i=1}^d \mu_i S_i(0) e^{q_{1i}(x-x_1)} e^{q_{ki}(\bar{y}(x) - x_k)} [q_{1i} + q_{ki} \bar{y}'(x)] = 0 \end{aligned}$$

A Herleitung der asymptotischen Entwicklung

zu

$$\bar{y}'(\bar{x}) = -\frac{\sum_{i=1}^d \mu_i S_i(0) q_{1i} e^{q_{1i}(x-x_1)}}{\sum_{i=1}^d \mu_i S_i(0) q_{ki} e^{q_{ki}(x-x_1)}}.$$

Der Beitrag zu (A.7) ist demnach

$$\frac{1}{\sqrt{8\pi t \lambda_1}} e^{-\frac{1}{2}\bar{\xi}^2} \frac{\left(\sum_{i=1}^d \hat{\mu}_i(t) q_{ki} S_i\right)^2}{\sum_{i=1}^d \hat{\mu}_i(t) q_{1i} S_i} \quad (\text{A.9})$$

mit $\bar{x} - x_1 = \bar{\xi} \sqrt{\lambda_1 t}$ und

$$\hat{\mu}_i(t) = \exp\left(q_{1i} \sqrt{\lambda_1 t} \bar{\xi} - t\sigma_i^2/2\right). \quad (\text{A.10})$$

Das gleiche Resultat ergibt sich für $\boxed{\underline{y}(\bar{x}) = x_k}$.

Ordnung ϵ^0 :

$$-\frac{1}{2} q_{ki}^2 \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \eta^2 e^{-\frac{\eta^2}{2}} d\eta \rightarrow -\sqrt{\frac{\pi}{2}} q_{ki}^2 \quad \text{für } \epsilon \rightarrow 0 \quad \left(\int_{-\infty}^{\infty} \eta^2 e^{-\frac{\eta^2}{2}} d\eta = \sqrt{2\pi}\right),$$

da für $x < \bar{x}$ auch $\underline{y}(x) < x_k < \bar{y}(x)$ gilt und daher $\lim_{\epsilon \rightarrow 0} \underline{\eta}_\epsilon(x) = -\infty$ und $\lim_{\epsilon \rightarrow 0} \bar{\eta}_\epsilon(x) = \infty$. In dieser Ordnung ergibt sich daher zusammen

$$\begin{aligned} & -\frac{1}{2} \frac{e^{-rt}}{\sqrt{2\pi t \sigma}} \sum_{i=1}^d q_{ik}^2 \mu_i S_i(0) \int_{-\infty}^{\bar{x}} e^{q_{1i}(x-x_1)} e^{-\frac{1}{2t} \frac{(x-x_1)^2}{\sigma^2}} dx \\ &= -\frac{1}{2} e^{-rt} \sum_{i=1}^d q_{ik}^2 \mu_i S_i(0) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\bar{\xi}} e^{q_{1i} \xi \sqrt{t} \sigma} e^{-\frac{\xi^2}{2}} d\xi \\ &= -\frac{1}{2} e^{-rt} \sum_{i=1}^d q_{ik}^2 \mu_i S_i(0) e^{\frac{q_{1i}^2 t \sigma^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\bar{\xi}} e^{-\frac{1}{2}(\xi - q_{1i} \sigma \sqrt{t})^2} d\xi \\ &= -\frac{1}{2} e^{-rt} \sum_{i=1}^d q_{ik}^2 \mu_i S_i(0) e^{\frac{q_{1i}^2 t \sigma^2}{2}} N(\bar{\xi} - q_{1i} \sigma \sqrt{t}) \\ &= -\frac{1}{2} e^{-rt} \sum_{i=1}^d q_{ik}^2 \mu_i S_i(0) \exp(q_{1i}^2 \sigma^2 t/2) N(\bar{\xi} - q_{1i} \sigma \sqrt{t}) \\ &= -\frac{1}{2} e^{-rt} \sum_{i=1}^d q_{ik}^2 \mu_i S_i \exp(q_{1i}^2 \lambda_1 t/2 + t(r - \sigma_i^2/2)) N(\bar{\xi} - q_{1i} \sqrt{\lambda_1 t}) \\ &= -\frac{1}{2} e^{-rt} \sum_{i=1}^d q_{ik}^2 \tilde{\mu}_i S_i N(\bar{\xi} - q_{1i} \sqrt{\lambda_1 t}) \end{aligned} \quad (\text{A.11})$$

mit

$$\tilde{\mu}_i(t) = \exp\left(q_{1i}^2 \lambda_1 t/2 + t(r - \sigma_i^2/2)\right). \quad (\text{A.12})$$

Das **Restglied** geht gleichmäßig gegen 0,

$$\epsilon \int_{\underline{\eta}_\epsilon(x)}^{\bar{\eta}_\epsilon(x)} \eta^3 e^{\epsilon q_{ki} \eta} e^{-\frac{\eta^2}{2}} d\eta \rightarrow 0 \quad \text{für } \epsilon \rightarrow 0,$$

sodass die Aussage des Satzes folgt. \square

Literaturverzeichnis

- [AG00] AVELLANEDA, M. und R. GAMBA: *Conquering the Greeks in Monte Carlo: Efficient Calculation of the Market Sensitivities and Hedge-Ratios of Financial Assets by Direct Numerical Simulation*. Courant Institute of Mathematical Sciences, New York University, 2000.
- [BBJ⁺97] BASTIAN, P., K. BIRKEN, K. JOHANNSEN, S. LANG, N. NEUSS, H. RENTZ-REICHERT und C. WIENERS: *UG — A flexible software toolbox for solving partial differential equations*. Computing and Visualization in Science, 1, 1997.
- [BC83] BRANDT, A. und C. W. CRYER: *Multigrid Algorithms for the Solution of Linear Complementarity Problems Arising from Free Boundary Problems*. SIAM J. Sci. Stat. Comput., 4(4), December 1983.
- [BCN04] BRIANI, M., C. LA CHIOMA und R. NATALINI: *Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory*. To appear in: Numerische Mathematik, 2004.
- [BGR] BUNGARTZ, H., M. GRIEBEL und U. RÜDE: *Extrapolation, Combination, and Sparse Grid Techniques for Elliptic Boundary Value Problems*. Institut für Informatik, Technische Universität München.
- [BGRZ94] BUNGARTZ, H.-J., M. GRIEBEL, D. RÖSCHKE und C. ZENGER: *Pointwise Convergence of the Combination Technique for the Laplace Equation*. East-West J. Num. Math., 2:21–45, 1994.
- [BN97] BARNDORFF-NIELSEN, O.-E.: *Normal Inverse Gaussian Distributions and the modelling of stock returns*. Scandinavian J. Statistics, 24:1–13, 1997.
- [BRSZ] BALDER, R., U. RÜDE, S. SCHNEIDER und C. ZENGER: *Sparse Grid and Extrapolation Methods for Parabolic Problems*. Institut für Informatik, Technische Universität München.
- [BS73] BLACK, F. und M. SCHOLES: *The pricing of options and corporate liabilities*. J. Pol. Econ., 81:637–659, 1973.
- [BS91] BRONSTEIN, I. N. und K. A. SEMENDJAJEW: *Taschenbuch der Mathematik*. B. G. Teubner, 25 Auflage, 1991.
- [Bun92] BUNGARTZ, H.-J.: *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*. Doktorarbeit, Technische Universität München, 1992.

Literaturverzeichnis

- [Bun98] BUNGARTZ, H.-J.: *Finite Elements of Higher Order on Sparse Grids*. Habilitationsschrift, Technische Universität München, 1998.
- [Con01] CONT, R.: *Empirical properties of asset returns: stylized facts and statistical issues*. *Quantitative Finance*, 1:223–236, 2001.
- [CP99] CLARKE, N. und K. PARROTT: *Multigrid for American Option Pricing with Stochastic Volatility*. *Appl. Math. Finance*, 6:177–197, 1999.
- [Cry71] CRYER, C. W.: *The solution of a quadratic programming problem using systematic overrelaxation*. *SIAM J. Control*, 9:385–392, 1971.
- [DS02] DOSTÁL, Z. und J. SCHÖBERL: *Modified proportioning for bound constraint quadratic programming with the rate of convergence and finite termination*. Technischer Bericht, Technical University Ostrava, 2002.
- [Eth02] ETHERIDGE, A.: *A Course in Financial Calculus*. Cambridge University Press, 2002.
- [FPS00] FOUQUE, J.-P., G. PAPANICOLAOU und R. SIRCAR: *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press, 2000.
- [Gar98] GARCKE, J.: *Berechnung von Eigenwerten der stationären Schrödingergleichung mit der Kombinationstechnik*. Diplomarbeit, Institut für Angewandte Mathematik, Universität Bonn, 1998.
- [GG98] GERSTNER, T. und M. GRIEBEL: *Numerical Integration using Sparse Grids*. *Numerical Algorithms*, 18:209–232, 1998.
- [GG03] GERSTNER, T. und M. GRIEBEL: *Dimension-Adaptive Tensor-Product Quadrature*. *Computing*, 71(1):65–87, 2003.
- [GSZ92] GRIEBEL, M., M. SCHNEIDER und C. ZENGER: *A combination technique for the solution of sparse grid problems*. In: GROEN, P. DE und R. BEAUWENS (Herausgeber): *Iterative Methods in Linear Algebra*. IMACS, Elsevier, North Holland, 1992.
- [GT95] GRIEBEL, M. und V. THURNER: *The efficient solution of fluid dynamics problems by the combination technique*. *Int. J. Num. Meth. for Heat and Fluid Flow*, 1995.
- [Hac93] HACKBUSCH, W.: *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Studienbücher Mathematik. B. G. Teubner, 2. Auflage, 1993.
- [HBS03] H. BLUM, D. BRAESS und F. T. SUTTMEIER: *A Cascadic Multigrid Algorithm for Variational Inequalities*. *Computing and Visualization in Science*, 2003. accepted.
- [Heu91] HEUSER, H.: *Lehrbuch der Analysis*. B. G. Teubner, 9. Auflage, 1991.
- [HK87] HOPPE, R. und R. KORNHUBER: *Multi-grid methods for the two phase Stefan problem*. Technischer Bericht 171, TU Berlin, 1987.
- [Hul00] HULL, J. C.: *Options, Futures, & Other Derivatives*. Prentice-Hall International, Vierte Auflage, 2000.

- [Irl98] IRLE, A.: *Finanzmathematik, Die Bewertung von Derivaten*. Teubner Studienbücher, 1998.
- [Kor97] KORNHUBER, R.: *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Advances in Numerical Mathematics. B. G. Teubner Stuttgart, 1997.
- [KSW90] KAWOHL, B., J. STARÁ und G. WITTUM: *Analysis and Numerical Studies of a Shape Design Problem*, 1990. Preprint. Sonderforschungsbereich 123, Stochastische Mathematische Modelle.
- [KZ94] KÖCVARA, M. und J. ZOWE: *An iterative two-step algorithm for linear complementarity problems*. Numerische Mathematik, 68:95–106, 1994.
- [Lan01] LANG, S.: *Parallele Numerische Simulation instationärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*. Technischer Bericht 110, Institut für Wasserbau, Universität Stuttgart, 2001.
- [LCPB99] LALOUX, L., P. CIZEAU, M. POTTERS und J.-P. BOUCHAUD: *Random Matrix Theory*. Risk, 12(69), 1999.
- [LM00] LLORENTE, I. M. und N. D. MELSON: *Behavior of Plane Relaxation Methods as Multigrid Smoothers*. Electronic Transactions on Numerical Analysis, 10:92–114, 2000.
- [Mer73] MERTON, R. C.: *Theory of Rational Option Pricing*. Bell Journal of Economics and Management Science, 4:141–183, 1973.
- [MvPS02] MATACHE, A.-M., T. VON PETERSDORFF und C. SCHWAB: *Fast Deterministic Pricing of Options on Lévy Driven Assets*. Research Report No. 2002-11, ETH Zürich, Seminar für Angewandte Mathematik, July 2002.
- [OFG04] OOSTERLEE, C. W., J. C. FRISCH und F. J. GASPAR: *TVD, WENO and blended BDF discretisations for Asian options*. To appear in Computing and Visualization in Science, 2004.
- [Øks98] ØKSENDAL, B.: *Stochastic Differential Equations*. Springer, 5. Auflage, 1998.
- [Oos97] OOSTERLEE, C. W.: *A GMRES-based plane smoother in multigrid to solve 3D anisotropic fluid flow problems*. Journ. Comp. Physics, 130:41–53, 1997.
- [Oos03] OOSTERLEE, C. W.: *On multigrid for linear complementarity problems with application to American-style options*. Electr. Trans. Num. Analysis, 15:165–185, 2003.
- [Pfl97] PFLAUM, C.: *Convergence of the Combination Technique for Second-Order Elliptic Differential Equations*. SIAM J. Numer. Anal., 34(6):2431–2455, December 1997.
- [PFV99] P. FORSYTH, R. ZVAN und K. VETZAL: *A finite element approach to the pricing of discrete lookbacks with stochastic volatility*. Applied Mathematical Finance, 1999.
- [PZ99] PFLAUM, C. und A. ZHOU: *Error Analysis of the Combination Technique*. Numerische Mathematik, 84:327–350, December 1999.

Literaturverzeichnis

- [Ran99] RANNACHER, R.: *Finite Element Methods for the Incompressible Navier-Stokes Equations*. Preprint 99-37, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, September 1999. SFB 359.
- [SL85] SCHÜLLER, A. und Q. LIN: *Efficient High Order Algorithms for Elliptic Boundary Value Problems Combining Full Multigrid Techniques and Extrapolation Methods*. Arbeitspapiere der GMD, Gesellschaft für Mathematik und Datenverarbeitung, Dezember 1985.
- [SMK98] SIDDIQI, A. H., P. MANCHANDA und M. KÖCVARA: *An iterative two step algorithm for American option pricing*. Technischer Bericht 241, Universität Erlangen, Institute of Applied Mathematics, 1998.
- [Smo63] SMOLYAK, S. A.: *Quadrature and interpolation formulas for tensor products of certain classes of functions*. Dokl. Akad. Nauk SSSR, (148):1042–1043, 1963. Russisch, Engl. Übers.: Soviet Math. Dokl. 4:240-243.
- [vdV92] VORST, H. v. D.: *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*. SIAM J. Sci. Stat. Comp., 13(2):631–644, 1992.
- [VO03] VÁSQUEZ, A. A. und C. W. OOSTERLEE: *Numerical Valuation of Options with Jumps in the Underlying*. Submitted to Applied Numerical Mathematics, 2003.
- [vPS01] PETERSDORFF, T. VON und C. SCHWAB: *Wavelet-discretizations of parabolic integro-differential equations*. Research Report No. 2001-07, ETH Zürich, Seminar für Angewandte Mathematik, August 2001.
- [vPS02] PETERSDORFF, T. VON und C. SCHWAB: *Numerical Solution of Parabolic Equations in High Dimensions*. Technischer Bericht, Isaac Newton Institute, Cambridge, 2002.
- [WDH93] WILMOTT, P., J. N. DEWYNNE und S. D. HOWISON: *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, 1993.
- [WHD95] WILMOTT, P., S. D. HOWISON und J. N. DEWYNNE: *The Mathematics of Financial Derivatives*. Cambridge University Press, 1995.
- [Zen90] ZENGER, C.: *Sparse Grids*. In: HACKBUSCH, W. (Herausgeber): *Parallel Algorithms for Partial Differential Equations*, Band 31 der Reihe *Notes on Numerical Fluid Dynamics*, 1990. Proceedings of the Sixth GAMM-Seminar.
- [Zhu03] ZHU, Y.-L.: *Three-factor interest rate models*. Communications in Mathematical Sciences, 1:557–573, 2003.
- [ZL03] ZHU, Y.-L. und J. LI: *Multi-factor Financial Derivatives on Finite Domains*. Communications in Mathematical Sciences, 1:343–359, 2003.

Für einen umfassenden Überblick über Veröffentlichungen im Kontext „Dünne Gitter“ siehe

<http://bibliothek.iam.uni-bonn.de/duennbib.html>

(*Sparse Grid Bibliography*).