

DIPLOMARBEIT

Effiziente numerische Optionsbewertung in Lévy-Modellen mittels \mathcal{H} -Matrix-Approximation

angefertigt am
Institut für Numerische Simulation

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

Dezember 2008

Von
Tim Osadnik
aus
Bottrop

Inhaltsverzeichnis

1	Einleitung	1
2	Modelle zur Optionsbewertung	7
2.1	Definitionen und Bezeichnungen	7
2.2	Black-Scholes-Modell	10
2.2.1	Marktannahmen	10
2.2.2	Die Black-Scholes-Gleichung	11
2.2.3	Die Black-Scholes Formel	11
2.2.4	Amerikanische Optionen im Black-Scholes Modell	12
2.3	Das Merton-Modell	14
2.3.1	Marktannahmen	14
2.3.2	Der Aktienwert und die Dynamik des Optionspreises	15
2.3.3	Eine Optionspreisformel	17
2.3.4	Geschlossene Lösungen für das Merton Modell	19
2.4	Kou-Modell	20
2.4.1	Beschreibung des Modells	21
2.4.2	Eigenschaften des Modells	22
2.5	VG- und CGMY-Modell	24
2.5.1	Lévy-Prozesse	24
2.5.2	VG-Modell	26
2.5.3	CGMY-Modell	27
2.5.4	Die Parameter des CGMY-Prozesses	27
2.5.5	Eigenschaftenvergleich der Modelle	29
2.6	Mehrdimensionale PIDE	30
2.7	Lévy-Copulas	32
2.7.1	Definition von Lévy-Copulas	32
2.7.2	Beispiele von Lévy-Copulas	35
2.7.3	Clayton-Lévy-Copula	37
2.8	Die Dichte des zweidimensionalen Lévy-Prozesses	38
3	Numerik der PIDE	41
3.1	Transformationen und Lokalisierung	42
3.2	Zeitdiskretisierung	44
3.3	Ortsdiskretisierung	45
3.3.1	Stabilität der Ortsdiskretisierung	45
3.4	Diskretisierung des Integraloperators	46
3.4.1	Kollokationsverfahren	47
3.5	Lösung der Gleichungssysteme	49
3.5.1	Das SOR-Verfahren	49
3.5.2	Der Brennan-Schwartz-Algorithmus	50
3.6	Diskretisierung der mehrdimensionalen PIDE	51
3.6.1	Zeitdiskretisierung	52

3.6.2	Ortsdiskretisierung	53
3.6.3	Diskretisierung des Integraloperators	53
3.6.4	Lösung der Gleichungssysteme	54
4	Hierarchische Matrizen	57
4.1	Niedrigrangmatrizen	57
4.1.1	Effiziente Darstellung	58
4.1.2	Niedrigrangapproximation	58
4.2	Matrixpartitionierung	59
4.3	Cluster-Bäume und Block-Cluster-Bäume	62
4.3.1	Cluster Bäume	62
4.3.2	Analyse einer einfachen Partition	63
4.3.3	Block-Cluster Bäume	66
4.4	Hierarchische Matrizen	68
4.4.1	Die Menge der hierarchischen Matrizen	68
4.4.2	Matrix-Vektor-Multiplikation	69
4.4.3	Invertierung von \mathcal{H} -Matrizen	70
5	Numerische Ergebnisse	73
5.1	Vorbemerkungen	73
5.2	Bewertung von Plain-Vanilla-Optionen	75
5.2.1	Merton-Modell	75
5.2.2	Kou-, VG- und CGMY-Modell	79
5.2.3	Amerikanische Optionen	80
5.3	Bewertung von Basket-Optionen	83
5.3.1	Europäische Optionen	85
5.3.2	Amerikanische Optionen	89
5.4	Fazit	90
6	Zusammenfassung und Ausblick	93
	Literaturverzeichnis	95

Kapitel 1

Einleitung

Mittlerweile handelt es sich bei einem Großteil der an den Finanzmärkten gehandelten Produkte um Finanzderivate. Wichtige Beispiele für Derivate sind Aktienoptionen. Im einfachen Fall von sogenannten Plain-Vanilla-Optionen handelt es sich um einen Vertrag, der dem Käufer das Recht einräumt, ein Wertpapier zu einem festgelegten Wert zu kaufen oder zu verkaufen. Man unterscheidet dabei zwischen Europäischen Optionen, die ein Ausüben der Option nur zu einem bestimmten Zeitpunkt erlauben, und Amerikanischen Optionen, die ein Ausüben innerhalb eines gewissen Zeitraumes gestatten. Komplexe Optionen hängen nicht nur von einem, sondern von mehreren Wertpapieren ab. Die bekannteste und am weitesten verbreitete Klasse solcher Optionen sind Basket-Optionen. Die Bewertung von Optionen erfolgt dabei entweder durch Angebot und Nachfrage an den Märkten oder, im Fall von wenig gehandelten Optionen, durch mathematische Modelle.

Die Anfänge der Optionspreisbestimmung liegen in der Arbeit [BS73] von Black und Scholes, in der gezeigt wurde, dass der faire Preis einer Europäischen Option durch die Lösung einer speziellen partiellen Differentialgleichung (PDE) gegeben ist. Ihre Arbeit basiert auf der Annahme, dass der Kurs des Wertpapiers, welches der Option zugrunde liegt, einer geometrischen Brownschen Bewegung folgt. Allerdings ist es nicht möglich, durch eine geometrische Brownsche Bewegung alle Eigenschaften eines Aktienkurses wiederzugeben, s. [Sch05]. Insbesondere Sprünge im Kursverlauf des Wertpapiers können durch Verwendung einer Brownschen Bewegung nur unzureichend dargestellt werden. Eine erste Erweiterung dieses rein stetigen Modells stellte die Arbeit [Mer76] von Merton dar. Er nimmt darin an, dass sich der Kurs der Aktie aus einer stetigen und einer Sprungkomponente zusammensetzt. Dies führt dazu, dass zur Bewertung einer Option nicht mehr nur eine PDE, sondern eine partielle Integro-Differentialgleichung (PIDE) gelöst werden muss. Die Sprungkomponente im Merton-Modell wird durch einen zusammengesetzten Poisson-Prozess, dessen Sprungintensitäten lognormalverteilt sind, modelliert. Ausgehend von dieser Arbeit sind in den letzten Jahren zahlreiche weitere Modelle entwickelt worden, die man in der Klasse der Sprung-Diffusions-Prozesse zusammenfasst. Zu dieser Klasse gehören unter anderem das Modell von Kou [Kou02], das Varianz-Gamma-Modell (VG-Modell) [MCC98] und das CGMY-Modell von Carr, Geman, Madan und Yor [CGMY02]. Diese Modelle verwenden komplexere und flexiblere Prozesse zur Modellierung des Sprungteils.

Durch eine geeignete Erweiterung der Modelle können sie auch zur Bewertung von Basket-Optionen, also von Optionen, deren Wert von mehreren Wertpapieren abhängt, verwendet werden. Die Kurse der verschiedenen Wertpapiere werden in diesem Fall durch eine mehrdimensionale Brownsche Bewegung sowie einen mehrdimensionalen Sprungprozess dargestellt. Die Verwendung dieser Prozesse führt dann zu einer mehrdimensionalen PDE bzw. PIDE, die zur Bewertung der Basket-Optionen zu lösen ist. Während für die Bewertung von Basket-Optionen im Merton-Modell eine mehrdimensionale Lognormalverteilung zur Modellierung des Sprungteils verwendet werden kann, s. [CF08], ist es im Falle des VG-Modells und des CGMY-Modells nicht offensichtlich, wie mehrdimensionale Sprung-Diffusions-Prozesse und deren Dichten modelliert werden können, wenn man bestimmte

Eigenschaften des Prozesses erhalten will. Ein Ansatz zur Lösung dieses Problems sind die sogenannten *Lévy-Copulas*, s. [KT06]. Lévy-Copulas sind Funktionen, mit denen es möglich ist, mehrdimensionale Lévy-Dichten, die bestimmte Eigenschaften, wie z.B. die Abhängigkeit einzelner Komponenten, aufweisen, aus eindimensionalen Lévy-Dichten zu generieren.

Wird das Black-Scholes- oder das Merton-Modell verwendet, können im Falle von Europäischen Optionen die auftretenden PDEs bzw. PIDEs analytisch gelöst werden. Falls keine analytische Lösung existiert, muss man auf numerische Verfahren zur Lösung der PDE bzw. PIDE zurückgreifen. Da in einem Bankhaus sehr viele Optionen in sehr kurzen Abständen bewertet werden müssen, ist es von großer Bedeutung, ein robustes, effizientes und speichersparendes Verfahren zur Durchführung der Berechnungen zur Verfügung zu haben. In der Regel löst man die PIDE durch Anwendung von Finite-Differenzen- oder Finite-Elemente-Verfahren. Die Bewertung einer Option im Black-Scholes-Modell führt dann zu einem linearen Gleichungssystem (LGS), das unter Berücksichtigung der Tridiagonalgestalt der Systemmatrix des LGS mit optimalem Aufwand $\mathcal{O}(N)$ gelöst werden kann. Dabei bezeichnet N die Anzahl an Diskretisierungspunkten, die das Verfahren verwendet. Bei PIDEs, wie sie in Sprung-Diffusions-Prozessen auftreten, führt der Integralteil der Gleichung zu einer vollbesetzten Systemmatrix. Diese führt dann wiederum zu einem Aufwand in der Größenordnung $\mathcal{O}(N^3)$ bei Anwendung eines direkten und $\mathcal{O}(N^2)$ bei Anwendung eines iterativen Löser. Eine weitere Verringerung des Aufwands wie im Black-Scholes-Fall ist nun nicht mehr ohne Weiteres möglich, da die Tridiagonaleigenschaft der Systemmatrix durch das Integral verlorengegangen ist.

Eine Möglichkeit, diesen hohen Aufwand zu verringern, wird in [MvPS04] von Matache, von Petersdorff und Schwab vorgestellt. Sie schlagen in ihrer Arbeit eine Wavelet-Galerkin-Diskretisierung des Integrals vor. Dabei wird die Matrix, die bei der Galerkin-Diskretisierung des Integrals entsteht, in einer Wavelet-Basis dargestellt. Diese Darstellung führt zu einer Matrix, die sehr viele kleine Einträge enthält, die vernachlässigt werden können. Das Gleichungssystem wird dann nach einer Vorkonditionierung der Matrix mittels einer Wavelet-Basis durch den Einsatz eines GMRES-Verfahrens iterativ gelöst. Das Gesamtverfahren führt dann zu einer Lösung, die in optimaler Komplexität erreicht werden kann. Allerdings beschränkt sich die Arbeit auf eindimensionale Probleme.

In [CV05] wird von Cont und Voltchkova eine rein explizite Behandlung des Integrals durch ein schnelles Kollokationsverfahren vorgeschlagen. Ein großer Vorteil dieses Ansatzes ist die Tatsache, dass die Abkoppelung des Integrals auf diese Art dazu führt, dass nur noch eine PDE in jedem Zeitschritt gelöst werden muss. Diese PDE führt dann wieder zu einem Gleichungssystem, dessen Systemmatrix Tridiagonalgestalt hat und daher mit $\mathcal{O}(N)$ Aufwand gelöst werden kann. Zur Berechnung der rechten Seite des LGS bleibt allerdings eine Matrix-Vektor-Multiplikation durchzuführen, die den Gesamtaufwand dann wieder auf die Größenordnung $\mathcal{O}(N^2)$ zurückfallen lässt.

Ein solcher quadratischer Aufwand führt insbesondere im Fall von Basket-Optionen zu langen Rechenzeiten, da sich der Aufwand mit der Dimension potenziert. Zur Bewertung von Basket-Optionen betrachten wir hier beispielhaft den Fall einer Option auf zwei Wertpapiere. Im Sprung-Diffusions-Modellen führt dies zu einer zweidimensionalen PIDE. Nach der Diskretisierung der PIDE ergibt sich ein Gleichungssystem, das durch einen direkten Löser nur mit einem Aufwand in der Größenordnung $\mathcal{O}(N^6)$ zu lösen ist, der sich durch Verwendung eines iterativen Löser auf $\mathcal{O}(N^4)$ verringern lässt. Selbst eine explizite Behandlung des Integralteils führt im zweidimensionalen Fall nicht zu einer Systemmatrix in Tridiagonalgestalt. Es ist zwar möglich, das Gleichungssystem durch die Verwendung von Mehrgitterverfahren wie sie in [Hac85] vorgestellt werden, mit einem Aufwand in der

Größenordnung $\mathcal{O}(N^2)$ zu lösen, der Aufwand zur Auswertung des Integral bleibt jedoch davon unbeeinträchtigt bei $\mathcal{O}(N^4)$.

In dieser Arbeit stellen wir ein neues numerisches Verfahren vor, mit dem es möglich ist, die PIDEs, die bei der Bewertung von Europäischen und Amerikanischen Plain-Vanilla-Optionen bzw. Basket-Optionen auf zwei Wertpapieren auftreten, mit optimalem Aufwand zu lösen. Dazu wenden wir unser Wissen nach zum ersten Mal im Finanzbereich eine \mathcal{H} -Matrix-Approximation zur Behandlung des Integralteils an. Dabei handelt es sich um ein Verfahren, bei dem die Matrix zunächst hierarchisch in Blöcke unterteilt wird. Dann erfolgt eine Approximation der einzelnen Blöcke durch Niedrigrangmatrizen, s. [Hac99, Beb08]. Die dabei nötigen Approximationen und \mathcal{H} -Matrix-Operationen führen wir mit Hilfe des Softwarepakets *AHMED* [Beb] von Prof. Bebendorf durch. Um eine separate Behandlung des Integralteils durch \mathcal{H} -Matrizen zu ermöglichen, folgen wir dem Ansatz aus [CV05] und verwenden eine Finite-Differenzen-Diskretisierung der PIDE und eine explizite Behandlung des Integrals.

Die Zeitdiskretisierung erfolgt über ein θ -Schema, das es möglich macht, verschiedene Zeitdiskretisierungsverfahren zu wählen. Um eine schnellere Konvergenz als in [CV05] zu erreichen, verwenden wir anstatt einer impliziten Euler-Diskretisierung das Crank-Nicolson-Verfahren. Dies führt insgesamt zu einem balancierten Gesamtverfahren, da die Konvergenzordnung der verwendeten Ortsdiskretisierung ebenfalls eine Konvergenzrate von 2 aufweist. Um Instabilitäten zu vermeiden, die bei der Anwendung eines reinen Crank-Nicolson-Verfahrens auftauchen, erweitern wir das Verfahren auf ein Rannacher-Zeitschritt-Verfahren. Dabei werden zunächst einige Zeitschritte mittels des impliziten Euler-Verfahrens ausgeführt und dann auf das Crank-Nicolson-Verfahren gewechselt. Dieses Verfahren wird auch in [Toi08] für PIDEs aus dem Finanzbereich angewendet, um Oszillationen in der Lösung der Gleichungen zu verhindern.

Zur Validierung unseres Verfahrens betrachten wir zunächst das Merton-Modell, da in diesem Fall die Konvergenz des Verfahrens gegen die analytische Lösung untersucht werden kann. Darüber hinaus wenden wir unser Verfahren auf komplexere Modelle, für die keine analytische Lösung mehr bekannt ist, an. Dazu verwenden wir die Doppel-Log-Exponential-Dichte des Kou-Modells [Kou02], sowie die Lévy-Dichten aus dem VG- und CGMY-Modell, s. [MCC98, CGMY02]. Die Lévy-Dichten wie sie im VG- und CGMY-Modell auftreten sind jedoch singular im Nullpunkt und erfordern daher besondere Aufmerksamkeit. Die Singularitäten werden, dem Ansatz aus [CV05] folgend, abgeschnitten und die dadurch fehlenden infinitesimalen Sprünge ersetzen wir durch eine Verstärkung der Brownschen Bewegung.

Bei der Bewertung von Plain-Vanilla-Optionen approximieren wir nur die Kollokationsmatrix des Integralteils mit \mathcal{H} -Matrizen, da das Lösen des Gleichungssystems in diesem Fall mit linearem Aufwand erfolgen kann, weil es sich bei der Systemmatrix um eine Tridiagonalmatrix handelt. Dies ist durch die explizite Behandlung des Integrals möglich. Durch die \mathcal{H} -Matrix-Approximation wird in allen Fällen eine erhebliche Verkürzung der benötigten Rechenzeit gegenüber der direkten Ausführung der Matrix-Vektor-Multiplikation erreicht, ohne dabei die Genauigkeit der Lösung zu beeinträchtigen. Die Verbesserung in diesen beiden Bereichen zeigen wir im Merton-, Kou-, VG- und CGMY-Modell bei der Bewertung Europäischer und Amerikanischer Optionen. Darüber hinaus zeigen wir, dass die Approximation zu einem geringeren Speicherverbrauch für die Behandlung der Kollokationsmatrix führt.

Aufbauend auf der Bewertung von Plain-Vanilla-Optionen werden wir zwei Modelle für zweidimensionale Basket-Optionen vorstellen. Zum einen werden wir den mehrdimensionalen

nen Sprungprozess wie in [CF08] mittels einer mehrdimensionalen Lognormalverteilung modellieren. Der zweite Ansatz folgt dann der Arbeit [KT06] von Kallsen und Toivanen. Hier wird vorgeschlagen, den mehrdimensionalen Sprungprozess mittels Lévy-Copulas zu modellieren. Dabei werden die eindimensionalen Randverteilungen des mehrdimensionalen Sprungprozesses durch VG- oder CGMY-Dichten vorgegeben und daraus dann die entsprechende mehrdimensionale Lévy-Dichte generiert.

Bei der Bewertung einer Basket-Option auf zwei Wertpapiere reicht die Approximation der Kollokationsmatrix allein nicht aus, um den Aufwand der Berechnung hinreichend zu verringern. Dies folgt aus der Tatsache, dass das lineare Gleichungssystem selbst bei expliziter Behandlung des Integrals nicht in linearer Zeit gelöst werden kann. Wir werden daher die Matrix des Gleichungssystems ebenfalls durch eine \mathcal{H} -Matrix approximieren und das System dann durch Invertierung der Matrix lösen. Wir werden auch hier feststellen, dass die Verwendung der \mathcal{H} -Matrix-Approximation zu einer erheblichen Verbesserung des Rechenaufwands beiträgt. Dazu wird es allerdings nötig sein, eine geometrische Bedingung einzuführen, die es ermöglicht, die \mathcal{H} -Matrix-Approximation besser zu steuern. Die Bewertung ist dann bis auf logarithmische Faktoren mit einem Aufwand in der Größenordnung $\mathcal{O}(N^2)$ möglich.

Um den geringen Aufwand zu gewährleisten, dürfen die System- und die Kollokationsmatrix nicht explizit aufgestellt werden. Daher ist es notwendig, das *AHMED* Softwarepaket mit Funktionen zu erweitern, die die Generierung der einzelnen Matrixeinträge ermöglichen. Darüberhinaus wurden Routinen und Klassen zur Erfassung und Abspeicherung der Geometrie des Integrationsbereichs hinzugefügt.

Die Hauptbeiträge dieser Arbeit lassen sich wie folgt zusammenfassen:

- Umsetzung eines Finite-Differenzen-Verfahrens mit Rannacher-Zeitschrittverfahren und expliziter Behandlung des Integralteils zur numerischen Lösung der PIDEs, die bei der Bewertung von Optionen im Merton-, Kou-, VG- und CGMY-Modell entstehen
- Umsetzung eines Finite-Differenzen-Verfahrens mit expliziter Behandlung des Integrals zur numerischen Lösung der PIDEs, die bei der Bewertung von Basket-Optionen auf zwei Wertpapieren im Merton Modell und mittels Lévy-Copulas entsteht
- Reduktion von Rechenzeit und Speicherbedarf durch \mathcal{H} -Matrix-Approximation der Kollokationsmatrix und der Systemmatrix des LGS bei der Bewertung von Basket-Optionen
- Erweiterung des \mathcal{H} -Matrix Softwarepakets *AHMED* durch folgende Routinen:
 - Berechnung der Matrixeinträge der Kollokations- und Systemmatrix
 - Einlesen und Speichern der Geometrie des Integrationsbereichs
 - Hinzufügen einer geometrischen Bedingung bei der Generierung der \mathcal{H} -Matrix
- Durchführung numerischer Experimente zum Vergleich der Genauigkeit und Rechenzeit unseres Ansatzes mit der direkten Auswertung des Integrals in den vier Modellen bzgl. Europäischer und Amerikanischer Optionen in einer und zwei Dimensionen

Die vorliegende Arbeit gliedert sich wie folgt:

Kapitel 2 befasst sich mit den verschiedenen Modellen zur Bewertung von Optionen. Zu Beginn wird kurz das Modell von Black und Scholes mit den entsprechenden Annahmen an den Markt sowie die Black-Scholes-Formel zur Bewertung von Europäischen Optionen

vorge stellt. Wir werden außerdem die Bewertung Amerikanischer Optionen diskutieren. Danach betrachten wir Sprung-Diffusions-Modelle, beginnend mit dem Merton-Modell. Anschließend behandeln wir das Kou-Modell, sowie, nach einem kurzen Exkurs über Lévy-Pozesse, das VG- und CGMY-Modell. Im zweiten Teil des Kapitels wenden wir uns der Bewertung von Basket-Optionen zu. Die Modellierung erfolgt hier zunächst mittels einer zweidimensionalen Lognormalverteilung und dann mit Hilfe der Lévy-Copulas. Im ersten Abschnitt über Copulas fassen wir die Definition und Eigenschaften von Lévy-Copulas und den Zusammenhang zwischen Lévy-Maßen, ihren eindimensionalen Randverteilungen und den Lévy-Copulas kurz zusammen. Im letzten Abschnitt stellen wir dann die Clayton-Copula-Familie und die daraus erzeugte mehrdimensionale Lévy-Dichte vor.

Die Diskretisierung der PIDE wird dann in Kapitel 3 beschrieben. Wir diskutieren zunächst die Lokalisierung des Gebiets, auf dem wir die Gleichung lösen. Im zweiten Abschnitt befassen wir uns mit der Zeitdiskretisierung und stellen das Crank-Nicolson-Verfahren und das Rannacher-Verfahren vor. Der nächste Abschnitt befasst sich dann mit der Finite-Differenzen-Diskretisierung der Ableitungen im Ort. Wir geben die entsprechenden Differenzenquotienten an und untersuchen ihre Stabilität. Die Berechnung des Integrals behandeln wir dann im vierten Abschnitt. Dazu beschreiben wir das verwendete Kollokationsverfahren sowie die Berechnung der Matrixeinträge der Kollokationsmatrix mittels Quadraturverfahren. Wir werden ebenfalls die iterativen und direkten Löser, die zur Lösung des LGS eingesetzt werden, angeben. Der zweite Teil des Kapitels befasst sich mit der Diskretisierung der zweidimensionalen PIDE in Ort und Zeit. Hier werden wir ebenfalls die Berechnung der Einträge der Kollokationsmatrix des zweidimensionalen Integrals mittels der Simpson-Quadratur-Formel erläutern.

Kapitel 4 gibt dann anhand von [Beb08] einen Überblick über das Prinzip der \mathcal{H} -Matrix-Approximation und deren Eigenschaften. Wir zeigen, wie durch eine hierarchische Unterteilung der Matrix in einzelne Blöcke und deren Approximation durch Niedrigrangmatrizen eine Struktur entsteht, mit Hilfe derer man die gängigen Matrixoperationen mit beinahe linearem Aufwand durchführen kann. Dies erläutern wir anhand der Matrix-Vektor-Multiplikation und der Matrix-Invertierung.

Die numerischen Ergebnisse präsentieren wir dann in Kapitel 5. Dabei vergleichen wir den Aufwand, die Rechenzeit und den Speicherbedarf der \mathcal{H} -Matrix-Approximation mit der direkten Behandlung des Integralteils. Diesen Vergleich führen wir für alle vier Modelle sowie für Europäische und Amerikanische Optionen durch. Der zweite Abschnitt des Kapitels behandelt dann die Effizienz der Approximation bezüglich Basket-Optionen mit zwei zugrunde liegenden Wertpapieren.

In Kapitel 6 geben wir eine Zusammenfassung dieser Arbeit und einen Ausblick auf mögliche Erweiterungen.

Danksagungen

An dieser Stelle möchte ich mich bei allen bedanken, die das Entstehen dieser Arbeit möglich gemacht haben. Mein besonderer Dank gilt dabei Prof. Griebel für die Überlassung des Themas und die gute Betreuung. Ein besonderer Dank gebührt ebenso Prof. Bebendorf für die Übernahme des Koreferates und für die Bereitstellung des AHMED-Softwarepakets sowie seine Hilfe bei allen diesbezüglich entstandenen Problemen und Fragen. Ich möchte auch Dr. Thomas Gerstner und Markus Holtz für die Anregungen und große Unterstützung bei der Erstellung dieser Arbeit danken.

Kapitel 2

Modelle zur Optionsbewertung

In diesem Kapitel werden wir die Modelle, welche wir später zur Bewertung von Optionen verwenden werden, vorstellen. Darunter das erste, grundlegende Modell von Black, Scholes und Merton, welches den modernen Optionshandel ermöglichte. Der grundlegende Unterschied der einzelnen Modelle besteht in der Darstellung des Aktienkurses, der der Option zu Grunde liegt. Black und Scholes verwenden in ihrer Arbeit dazu eine geometrische Brownsche Bewegung. Diese Darstellung ermöglicht es jedoch nicht, alle Eigenschaften des Kurses ausreichend darzustellen. Daher betrachten wir die Erweiterung dieses Modells auf ein Sprung-Diffusionsmodell durch Merton aus dem Jahr 1973. Da dieses Modell die Grundlage für die Betrachtung des Aktienkurses als Sprung-Diffusionsprozess legt, werden wir uns mit diesem Modell ausführlicher befassen. Es folgen zwei komplexere Modelle aus dem Jahr 2002. Zum einen das Modell von Kou, beschrieben in [Kou02], das eine Doppel-Exponential Verteilung der Sprünge zugrunde legt, zum anderen das wahrscheinlich flexibelste Modell von Carr, Geman, Madan und Yor (kurz CGMY-Modell).

Doch bevor wir die verschiedenen Modelle zur Bewertung von Optionen vorstellen werden in diesem Kapitel zunächst einige Begriffe und Eigenschaften im Zusammenhang mit Optionen vorgestellt. Die Beschreibungen der Optionen orientiert sich hauptsächlich an [Hul06].

2.1 Definitionen und Bezeichnungen

Definition 2.1.1 [OPTION]

Durch den Kauf einer Option erwirbt der Käufer das Recht, aber nicht die Pflicht, einen zugrunde liegenden Basiswert zu oder vor einem festgelegten Zeitpunkt T zu einem speziellen Preis K , dem Ausübungspreis, zu kaufen (Call-Option oder Call) oder zu verkaufen (Put-Option oder Put).

Der Ausübungspreis K , der Zeitpunkt T sowie der zugrunde liegende Basiswert S müssen vor dem Kauf der Option festgelegt werden. Der Wert V zum Zeitpunkt t einer solchen Option wird sowohl durch die oben gewählten Parameter, als auch durch den aktuellen Preis des Basiswertes $S(t)$ bestimmt. Die vollständige Bezeichnung des Optionspreises lautet daher $V = V(S(t), K, T, t)$. Im weiteren Verlauf werden wir jedoch die kürzere Bezeichnung $V = V(S, t)$ verwenden und dabei den Ausübungspreis K und den Fälligkeitszeitpunkt T als fest gewählte Parameter behandeln.

Definition 2.1.2 [EUROPÄISCHE, AMERIKANISCHE OPTION]

Erwirbt der Käufer mit der Option das Recht, die Option nur zum Fälligkeitszeitpunkt T auszuüben, so spricht man von einer Europäischen Option. Ist eine Ausübung hingegen einmalig innerhalb des gesamten Zeitraums bis zum Zeitpunkt T möglich, so spricht man von einer Amerikanischen Option.

Definition 2.1.3 [BASKET-OPTION]

Bei einer Basket Option handelt es sich um eine Option, deren Preis von mehreren Basiswerten abhängt.

Die Art des Basiswerts der einer Option zu Grunde liegt kann sehr stark variieren. Er kann von Aktienpreisen über Rohstoffpreise bis hin zu Regenmengen reichen. Im folgenden werden wir uns mit Optionen beschäftigen, deren Basiswert eine dividendengeschützte Aktie ist, d.h. eine Aktie, für die keine Dividende ausgeschüttet wird. Die Ergebnisse lassen sich jedoch auf Basiswerte mit stetiger oder diskreter Dividendenausschüttung übertragen. Neben den schon vorgestellten Standard-Optionen (oder Plain-Vanilla-Optionen) werden noch andere Typen von Optionen (Exotische Optionen) an den Börsen gehandelt. Ihr Wert hängt beispielsweise nicht nur vom Preis des Basiswertes zum Zeitpunkt der Ausübung ab, sondern von Teilen oder dem gesamten Kursverlauf. Beispiele hierfür sind *asiatische Optionen* oder *Bermuda-Optionen*.

Eine wichtige Aufgabe der Finanznumerik ist es für solche Optionen einen fairen Preis zu ermitteln, da eine geschlossene analytische Lösung für dieses Problem nur in wenigen Fällen existiert. Die errechneten Preise müssen nicht mit den Preisen der an den Börsen gehandelten Optionen übereinstimmen, da diese aus Angebot und Nachfrage am Markt entstehen. Ein fairer Preis ist jedoch für alle Beteiligten ein sehr wichtiges Hilfsmittel im systematischen Handel mit Optionen.

Unter bestimmten Annahmen an die Modelle zur Darstellung der Kursverläufe des Basiswerts lassen sich diese Preise über numerische Verfahren bestimmen. Eine der wichtigsten Annahmen ist in diesem Fall die Arbitragefreiheit des Marktes.

Definition 2.1.4 [ARBITRAGEMÖGLICHKEIT]

Eine Arbitragemöglichkeit ist eine Handelsstrategie, die keine Anfangsinvestition benötigt und mit positiver Wahrscheinlichkeit einen Gewinn ergibt, ohne das Risiko eines Verlustes zu beinhalten.

Definition 2.1.5 [ARBITRAGEFREIER MARKT]

Ein arbitragefreier Markt bezeichnet einen Finanzmarkt, in dem keine Arbitragemöglichkeiten bestehen.

Beispiel 2.1.6

Nehmen wir an, dass eine Aktie an einer Börse A und einer Börse B gehandelt wird. Der Preis der Aktie an der Börse A beträgt 99 €, der Preis an der Börse B beträgt jedoch 101 €. Falls keine Transaktionskosten anfallen, ist es möglich die Aktie für 99 € zu kaufen und zu einem Preis von 101 € direkt wieder zu verkaufen. Es entsteht ein risikoloser Gewinn von 2 € pro Aktie. In einem realen Markt tauchen solche Preisunterschiede allerdings nicht auf, da diese durch Konkurrenz der Händler untereinander entfernt werden.

Die Annahme des arbitragefreien Marktes liefert das *Gesetz des einheitlichen Preises*, nach dem alle Finanzderivate mit derselben Auszahlung einen einheitlichen Preis besitzen müssen.

Im weiteren Verlauf der Arbeit werden im Zusammenhang mit Optionen die Bezeichnungen aus Tabelle 2.1 verwendet.

Sollte aus dem jeweiligen Kontext nicht hervorgehen, um welche Art von Option es sich handelt, kennzeichnen die unteren Indizes p bzw. c , ob es sich um eine Put- oder Call-Option handelt, die oberen Indizes eur bzw. am , ob es sich um eine Europäische oder Amerikanische Option handelt. V_c^{eur} bezeichnet also eine Europäische Call-Option.

Der Wert einer Option zum Ende der Laufzeit lässt sich sehr einfach bestimmen. Hierfür betrachten wir exemplarisch eine Europäische Call-Option. Sollte der Kurs des Basiswertes

Bezeichnung	Bedeutung
$S(t) \in \mathbb{R}^+$	Preis des zugrunde liegenden Basiswertes zum Zeitpunkt t
$T \in \mathbb{R}^+$	Verfallsdatum, Fälligkeitszeitpunkt, Laufzeit
$K \in \mathbb{R}^+$	Ausübungspreis, Strike
$r \in \mathbb{R}^+$	Risikoloser Zinssatz
$\sigma \in \mathbb{R}^+$	Volatilität, (jährliche) Schwankungsbreite des Kurses
$V(S, t) \in \mathbb{R}$	Wert einer Option zum Zeitpunkt t bei aktuellem Kurs $S = S(t)$

Tabelle 2.1: Bezeichnungen im Zusammenhang mit Optionen

zum Zeitpunkt T über dem Ausübungspreis K der Option liegen, so wird der Halter diese ausüben. Er kauft die Aktie zum Preis K und erhält nach sofortigem Wiederverkauf einen Gewinn von $S - K$. Falls der Preis der Aktie kleiner als der Ausübungspreis ist, wird der Halter sein Recht nicht ausüben, da er die Aktie am Markt günstiger erwerben kann. Die Option verfällt und ist daher wertlos. Eine analoge Aussage lässt sich für eine entsprechende Put-Option treffen, man erhält daher für den Zeitpunkt T folgenden Zusammenhang:

$$\begin{aligned} V_c(S, T) &= \max\{S - K, 0\} = \mathcal{P}_c && \text{bei Call-Optionen,} \\ V_p(S, T) &= \max\{K - S, 0\} = \mathcal{P}_p && \text{bei Put-Optionen.} \end{aligned} \quad (2.1)$$

Im weiteren Verlauf der Arbeit werden wir die verkürzte Schreibweise

$$a^+ := \max\{a, 0\}$$

verwenden. Den Wert einer Option zum Fälligkeitszeitpunkt T in Abhängigkeit vom Aktienkurs S nennt man *Auszahlungs-* oder *Payoff-Funktion* $\mathcal{P} = \mathcal{P}(S)$. Handelt es sich bei der Option um eine Put-Option so gilt $\mathcal{P} = \mathcal{P}_p$, bei einer Call-Option entsprechend $\mathcal{P} = \mathcal{P}_c$. In Abbildung 2.1 sind die Auszahlungsfunktionen für eine Europäische Call- und eine Put-Option dargestellt.

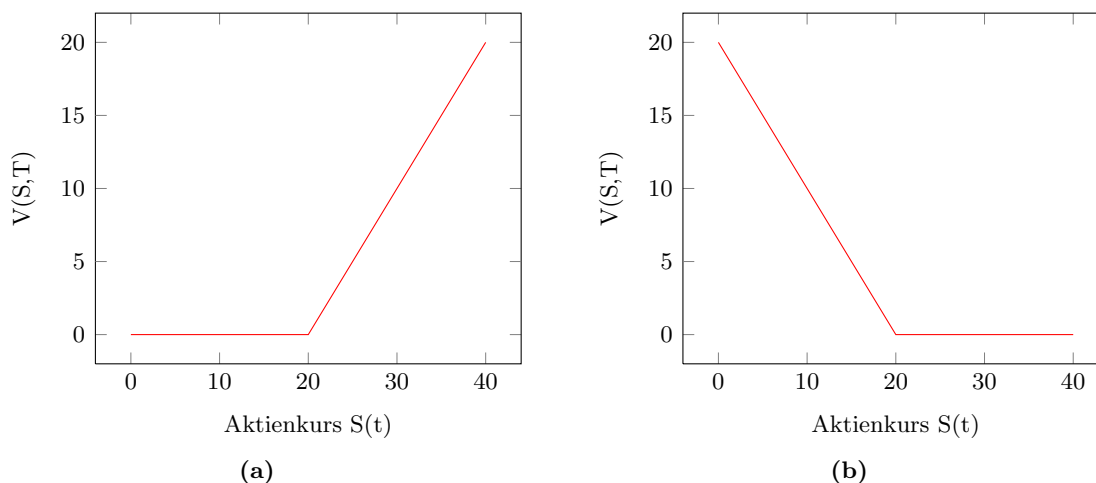


Abbildung 2.1: Auszahlungsfunktion einer Call-Option (a) und einer Put-Option (b) bei einem Ausübungspreis von $K=20$

2.2 Black-Scholes-Modell

Das bekannteste und am weitesten verbreitete Modell zur Bewertung von Optionen ist das Black-Scholes-Modell, welches wir in diesem Abschnitt vorstellen wollen.

Im Jahr 1973 wurde durch das von F. Black und M. Scholes veröffentlichte Black-Scholes-Modell [BS73] der Grundstein zum weltweiten Handel mit Optionen gelegt. Erstmals war es mit diesem Modell möglich, den fairen Preis einer Option auf eine dividendengeschützte Aktie zu berechnen. An der Entwicklung des Modells war ebenfalls R. Merton beteiligt, der jedoch ein eigenes Papier [Mer73] veröffentlichte. Aus diesem Grund findet man das Modell auch unter dem Namen Black-Scholes-Merton-Modell. Im Jahre 1997 erhielten R. Merton und M. Scholes den Nobelpreis für Wirtschaftswissenschaften für dieses Modell, F. Black war zu diesem Zeitpunkt bereits verstorben.

2.2.1 Marktannahmen

Im Black-Scholes-Modell werden folgende Annahmen an den Markt getroffen:

1. Es gibt keine Transaktionskosten.
2. Für die Aktien werden keine Dividenden gezahlt.
3. Der risikolose Zinssatz r ist konstant über die Zeit.
4. Alle Transaktionen sind jederzeit in beliebigem Umfang ohne Rückwirkung auf die Kursentwicklung möglich. Dabei sind Leerkäufe erlaubt.
5. Alle Wertpapiere sind beliebig teilbar.
6. Alle Parteien haben Zugang zu allen Informationen.
7. Es gibt keine Arbitragemöglichkeiten.
8. Der Aktienkurs $S(t)$ folgt der linearen stochastischen Differentialgleichung

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \quad (2.2)$$

mit konstanten Parametern $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}^+$. Dabei bezeichnet $W(t) \sim N(0, t)$ einen Wiener Prozess bzw. eine Standard-Brownsche Bewegung, wobei $N(0, t)$ eine Gauß'sche Normalverteilung mit Mittelwert 0 und Varianz t ist. Dies bedeutet, dass der Basiswert durch eine geometrische Brownsche Bewegung modelliert wird, wobei man den Parameter μ als Drift bezeichnet, den Parameter σ als Volatilität. Die grundlegenden Definitionen aus der Wahrscheinlichkeitstheorie findet man in [Bau02].

Bis auf die letzten zwei lassen sich diese Annahmen auch verallgemeinern. Es ist durchaus möglich Transaktionskosten zu berücksichtigen oder stochastische Volatilität zuzulassen. Diese Verallgemeinerungen führen dabei in den meisten Fällen zu komplizierteren Modellen. Eine Verallgemeinerung der letzten Annahme, d.h. die Verwendung allgemeinerer Prozesse zur Darstellung des Aktienkurses führt unter anderem zu den Sprung-Diffusionsprozessen, die wir später in diesem Kapitel vorstellen.

2.2.2 Die Black-Scholes-Gleichung

Black und Scholes haben in ihrer Arbeit [BS73] gezeigt, dass es möglich ist, ein dynamisches, risikoloses Portfolio bestehend aus α Aktien und einem Geldbetrag G zu erstellen, welches unabhängig vom Kurs der Aktie zum Zeitpunkt T die gleiche Auszahlung wie eine Europäische Option liefert. Aus der Arbitragefreiheit des Marktes folgt dann, dass der Wert der Option gleich der Anfangsinvestition in das Portfolio sein muss, d.h.

$$V(S, t) = \alpha S(t) + G.$$

Mit der Konstruktion dieses Portfolios und partiellen Differentialgleichungen als wichtigstem technischen Hilfsmittel zeigen Black und Scholes, dass der Preis einer Europäischen Option bei konstantem Zins r und konstanter Volatilität σ der *Black-Scholes-Gleichung*

$$\mathcal{L}V := \frac{\partial}{\partial t}V + \frac{\sigma^2}{2}S^2 \frac{\partial^2}{\partial S^2}V + rS \frac{\partial}{\partial S}V - rV = 0 \quad (2.3)$$

genügen muss. Da in dieser Arbeit keine Anforderungen an das zugrunde liegende Wertpapier gemacht werden, lässt sich diese Formel auf viele verschiedene Aktientypen übertragen, solange sie vom europäischen Typ sind, keine Dividenden ausschütten und einer geometrischen Brownschen Bewegung folgen. Diese unterschiedlichen Optionstypen realisiert man durch unterschiedliche Anfangsbedingungen für die obige partielle Differentialgleichung (2.3). Es existieren also mehrere Lösungen für die Black-Scholes-Gleichung, die nur von der Anfangsbedingung abhängen.

2.2.3 Die Black-Scholes Formel

Die von Black und Scholes aufgestellte Gleichung (2.3) besitzt eine analytische Lösung. Der Wert einer Europäischen Call-Option lässt sich darstellen als

$$V_c(S(0), t) = S(0)N(d_1) - Ke^{-r(T-t)}N(d_2). \quad (2.4)$$

Dabei ergeben sich d_1 und d_2 aus

$$d_1 = \frac{\ln\left(\frac{S(0)}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \quad (2.5)$$

und

$$d_2 = d_1 - \sigma\sqrt{T-t}. \quad (2.6)$$

$N(x)$ bezeichnet dabei den Wert der kumulativen Normalverteilung mit Mittelwert 0 und Varianz 1

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy. \quad (2.7)$$

Der Wert einer Europäischen Put-Option ist durch

$$V_p(S(0), t) = Ke^{-r(T-t)}N(-d_2) - S(0)N(-d_1) \quad (2.8)$$

gegeben. Da sich der Wert einer Put- bzw. Call-Option über die Put-Call Parität jeweils aus dem anderen berechnen lässt, werden wir von nun an nur noch die Bewertung Europäischer Put-Optionen betrachten.

2.2.4 Amerikanische Optionen im Black-Scholes Modell

Wie in Kapitel 1 beschrieben hat der Halter einer Amerikanischen Option jederzeit vor dem Verfallsdatum das Recht, diese auszuüben. Dieses Ausübungsrecht erschwert die Bewertung einer dieser Optionen und führt unter anderem dazu, dass zur Berechnung des Preises grundsätzlich keine geschlossenen Formeln existieren, weshalb man immer auf ein numerisches Verfahren zurückgreifen muss. Auf Grund der Tatsache, dass im Fall einer Amerikanischen Call-Option ein vorzeitiges Ausüben nie sinnvoll ist, s. [Hul06], werden wir uns in diesem Abschnitt nur mit Amerikanischen Put-Optionen befassen. Für diese gilt:

Satz 2.2.1

Für eine Amerikanische Put-Option gibt es immer einen Aktienkurs $S(t) > 0$, für den ein vorzeitiges Ausüben zum Zeitpunkt $t < T$ optimal ist.

Der Zeitpunkt, zu dem der Halter seine Option ausübt, ist im Vorfeld unbekannt und führt zu einem instationären freien Randwertproblem. Dieses ergibt sich wie folgt: Wir wissen aus Satz 2.2.1, dass im Falle einer Amerikanischen Put-Option für alle Zeitpunkte $t < T$ ein Kurs S existiert, so dass eine sofortige Ausübung der Option optimal ist. Die Auszahlung beträgt in diesem Fall $K - S > 0$. Falls für einen Kurs S die Ausübung optimal sein sollte, so muss dies auch für alle Kurse $\tilde{S} < S$ gelten, da die Auszahlung mit fallendem Kurs steigt. Wir erhalten damit für jeden Zeitpunkt $t < T$ einen eindeutigen Kurs $S_f(t)$, für den gilt, dass ein Ausüben der Option optimal ist, falls $S < S_f(t)$, und, dass die Option noch gehalten werden sollte, falls $S_f(t) < S$. Die Kurse $S_f(t)$ nennt man *optimale Ausübungspreise* der Option, sie definieren den *freien Rand* S_f . Dieser freie Rand teilt nun das Gebiet $\Omega := [0, \infty] \times [0, T]$, auf welchem wir unseren Optionspreis $V(S, t)$ ermitteln wollen, in zwei Teile. Der erste Teil umfasst das Gebiet $\Omega_1 := \{(S, t) \in \Omega : S \leq S_f(t)\}$ und der andere das Gebiet $\Omega_2 := \{(S, t) \in \Omega : S > S_f(t)\}$. Nach Satz 2.2.1 wissen wir, dass

$$0 < S_f(t) < K \tag{2.9}$$

gelten muss. Da die Option im Gebiet Ω_1 direkt ausgeübt würde, hat sie dort den Wert $V(S, t) = K - S$ und erfüllt nicht mehr die Black-Scholes Gleichung, denn durch Einsetzen in die Gleichung erhält man

$$\mathcal{L}V := \frac{\partial}{\partial t}V + \frac{\sigma^2}{2}S^2 \frac{\partial^2}{\partial S^2}V + rS \frac{\partial}{\partial S}V - rV = -rK < 0. \tag{2.10}$$

Im Gebiet Ω_2 gilt jedoch $V(S, t) > \mathcal{P}(S)$ und der Wert der Option erfüllt die Black-Scholes Gleichung.

Um das freie Randwertproblem eindeutig lösen zu können, müssen noch Randdaten auf dem freien Rand vorgegeben werden. Da für die Kurse auf dem freien Rand ein Ausüben optimal ist, gilt

$$V(S_f(t), t) = (K - S_f(t))^+. \tag{2.11}$$

Es lassen sich ebenfalls zwei weitere wichtige Eigenschaften des freien Randes zeigen ([Kwo98], Kapitel 4.1) :

$$\frac{\partial V}{\partial S}(S_f(t), t) = -1 \tag{2.12}$$

und

$$\lim_{t \rightarrow T} S_f(t) = K. \tag{2.13}$$

Abschließend kann man das freie Randwertproblem zur Bewertung einer Amerikanischen Option wie folgt zusammenfassen:

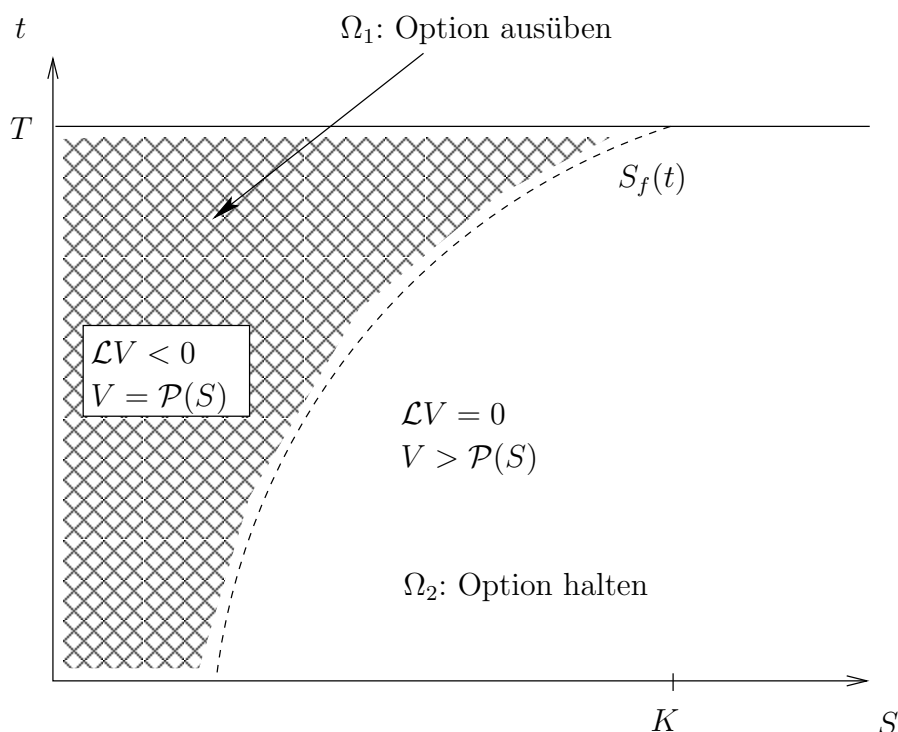


Abbildung 2.2: Amerikanische Option als freies Randwertproblem

Satz 2.2.2 [AMERIKANISCHE PUT-OPTION]

Der Wert $V(S, t)$ einer Amerikanischen Put-Option erfüllt

$$\begin{aligned} \mathcal{L}V &:= \frac{\partial}{\partial t}V + \frac{\sigma^2}{2}S^2 \frac{\partial^2}{\partial S^2}V + rS \frac{\partial}{\partial S}V - rV = 0 && \text{für } S > S_f \text{ und } 0 \leq t < T, \\ V(S, t) &= (K - S)^+ && \text{für } S \leq S_f \text{ und } 0 \leq t < T, \end{aligned}$$

mit den Randdaten

$$V(S, t) = 0 \quad \text{für } S \rightarrow \infty \text{ und } 0 \leq t < T,$$

den Enddaten

$$V(S, T) = (K - S)^+ \quad \text{für } S \geq 0$$

und der Bedingung, dass V und $\frac{\partial V}{\partial S}(S_f, t)$ stetig sind.

In der Formulierung des Problems in Satz (2.2.2) kommt der freie Rand noch explizit vor. Dieser ist aber wie schon erwähnt unbekannt. Daher bietet es sich an, anstatt einer direkten Lösung des Problems erst eine Umformulierung vorzunehmen. Die nun folgende Darstellung, welche dann auch verwendet wird, um das Problem zu lösen, nennt man die *linear komplementäre Darstellung* des Problems.

Satz 2.2.3 [LINEAR KOMPLEMENTÄRES PROBLEM]

Der Wert $V(S, t)$ einer Amerikanischen Put-Option erfüllt die linear komplementäre Be-

dingung

$$\begin{aligned} \mathcal{L}V &\geq 0 && \text{für } S > 0, \\ V(S, t) &\geq (K - S)^+ && \text{für } S > 0, \\ \mathcal{L}V(V(S, t) - (K - S)^+) &= 0 && \text{für } S > 0 \end{aligned} \tag{2.14}$$

mit der Anfangsbedingung

$$V(S, 0) = (K - S)^+. \tag{2.15}$$

2.3 Das Merton-Modell

Die Arbeiten von Black und Scholes führten zu einem großen Durchbruch im Bereich der Optionspreisberechnungen und des Optionshandels. Für die Gültigkeit der Black-Scholes-Formel muss man allerdings von den in Abschnitt 2.2.1 gemachten Voraussetzungen an das Marktmodell ausgehen. Eine strittige Annahme ist aber, dass sich der Kurs des Wertpapiers wie eine geometrische Brownsche Bewegung, also wie ein kontinuierlicher stochastischer Prozess, verhält. Merton hat in [Mer76] ein Modell entwickelt, in dem der Aktienkurs einem Sprung-Diffusionsprozess folgt. Die Grundlagen des Modells und die Herleitung der partiellen Integro-Differentialgleichung (PIDE), die in diesem Abschnitt erläutert werden stammen aus [Mer76].

2.3.1 Marktannahmen

1. Es gibt keine Transaktionskosten.
2. Für die Aktien werden keine Dividenden gezahlt.
3. Der risikolose Zinssatz r ist konstant über die Zeit.
4. Alle Transaktionen sind jederzeit in beliebigem Umfang ohne Rückwirkung auf die Kursentwicklung möglich. Dabei sind Leerkäufe erlaubt.
5. Alle Wertpapiere sind beliebig teilbar.
6. Alle Parteien haben Zugang zu allen Informationen.
7. Es gibt keine Arbitragemöglichkeiten.
8. Der Aktienkurs $S(t)$ folgt einem Prozess, der durch die stochastische partielle Differentialgleichung

$$\frac{dS(t)}{S(t)} = (\alpha - \lambda\kappa)dt + \sigma dW(t) + dq(t) \tag{2.16}$$

bestimmt wird. Hierbei gibt α den zu erwartenden Aktienertrag und σ die Volatilität der Aktie an. Bei dem zusätzlichen Prozess $q(t)$ handelt es sich um einen Poisson-Prozess, der unabhängig von dem Wiener-Prozess $W(t)$ ist. Die Wahrscheinlichkeiten dieses Poisson-Prozesses lassen sich als

$$\begin{aligned} P(\text{Das Ereignis tritt im Zeitintervall } [t, t+h] \text{ nicht ein}) &= 1 - \lambda h \\ P(\text{Das Ereignis tritt im Zeitintervall } [t, t+h] \text{ ein}) &= \lambda h \end{aligned} \tag{2.17}$$

beschreiben. Dabei ist λ die durchschnittliche Anzahl an Ereignissen pro Zeitintervall und $\kappa = E[Y - 1] = [Y] - 1$, wobei $(Y - 1)$ die prozentuale Veränderung des Preises für den Fall ist, dass das Ereignis eintritt. Wir bezeichnen mit $E[\cdot]$ den Erwartungswertoperator über die Zufallsvariable Y .

Hier wird noch einmal deutlich, dass alle Marktannahmen bis auf die Modellierung des Aktienkurses aus dem Black-Scholes-Modell beibehalten wurden.

2.3.2 Der Aktienwert und die Dynamik des Optionspreises

Die Gesamtdynamik des Aktienkurses setzt sich aus zwei Komponenten zusammen. Zum einen gibt es die normalen Preisschwankungen. Sie entstehen zum Beispiel aus Unterschieden bei Angebot und Nachfrage oder durch neue Informationen am Markt, die minimale Veränderungen im Aktienpreis hervorrufen. Dieser Part der Dynamik wird durch eine Standard geometrische Brownsche Bewegung dargestellt. Sie besitzt konstanten Drift, konstante Volatilität und fast sicher stetige Pfade.

Zum anderen gibt es die Veränderungen im Preis, die durch wichtige neue Informationen hervorgerufen werden. Diese verursachen Sprünge im Kurs der Aktie, treten allerdings nur zu diskreten, zufälligen Zeitpunkten auf. Diese Komponente des Preises wird durch einen Sprungprozess, in der Regel durch einen Poisson Prozess, beschrieben. Falls ein Poisson-Ereignis eintritt, beschreibt die Zufallsvariable Y die Auswirkung des Ereignisses auf den Aktienkurs. Das heißt, wenn $S(t)$ den Aktienkurs zum Zeitpunkt t widerspiegelt, ist der Kurs zum Zeitpunkt $(t + h)$ gegeben durch $S(t + h) = S(t)Y$. Dabei wird vorausgesetzt, dass es sich bei der Zufallsvariable Y um eine Variable mit kompaktem Träger handelt, dass $Y \geq 0$ gilt und dass die Zufallsvariablen $\{Y\}$ voneinander unabhängig sind. In (2.16) beschreibt also $\sigma dW(t)$ die normalen marginalen Preisschwankungen und $dq(t)$ die Preissprünge. Wählt man $\lambda = 0$, so erhält man die gleiche Formel wie im Black-Scholes-Modell. Wir können die Gleichung (2.16) daher umschreiben in

$$\frac{dS}{S} = \begin{cases} (\alpha - \lambda\kappa)dt + \sigma dW(t) & , \text{ falls das Poisson Ereignis nicht eintritt} \\ (\alpha - \lambda\kappa)dt + \sigma dW(t) + (Y - 1) & , \text{ falls das Poisson Ereignis eintritt} \end{cases} \quad (2.18)$$

wobei mit Wahrscheinlichkeit 1 nur ein Poisson Ereignis an einem Zeitpunkt stattfindet. Der entstehende Pfad ist dann größtenteils kontinuierlich mit einigen diskreten Sprüngen, die unterschiedliche Vorzeichen und Größen aufweisen. Falls die Parameter α , λ , κ , und σ konstant gewählt werden, kann man das Verhältnis von $S(t)$ und $S(0)$ wie folgt ausdrücken:

$$\frac{S(t)}{S(0)} = \exp\left(\left(\frac{\alpha - 1}{\sigma^2} - \lambda\kappa\right)t + \sigma W(t)\right) X_n. \quad (2.19)$$

$W(t)$ ist dabei eine normalverteilte Zufallsvariable mit Erwartungswert 0 und Varianz t , $X_n = 1$ für $n = 0$ und $X_n = \prod_{j=1}^n Y_j$, wobei die Y_j unabhängig und gleich verteilt sind. Die Anzahl n der Sprünge ist Poisson-verteilt mit Parameter λt .

Wir kommen nun zur Darstellung des Kurses der entsprechenden Option. Dazu wird angenommen, dass der Optionspreis V als zweimal differenzierbare Funktion des Aktienkurses geschrieben werden kann. Folgt nun der Aktienkurs der in (2.18) beschriebenen Dynamik kann der Optionspreis in einer ähnlichen Weise als

$$\frac{dV(S, t)}{V(S, t)} = (\alpha_{opt} - \lambda\kappa_{opt})dt + \sigma_{opt}dW(t) + dq_{opt}(t) \quad (2.20)$$

beschrieben werden, wobei α_{opt} und σ_{opt} den erwarteten Ertrag bzw. die Volatilität der Option wiedergeben. $q_{opt}(t)$ ist ein unabhängiger Poisson-Prozess mit Parameter λ . $\kappa_{opt} =$

$E[Y_{opt} - 1]$, wobei $(Y_{opt} - 1)$ wieder die prozentuale Veränderung des Preises im Falle des Poisson-Ereignisses angibt.

Mit Hilfe des Ito-Lemmas für Sprungprozesse erhält man folgende Beziehungen

$$\begin{aligned} \alpha_{opt} &= \frac{1}{2}\sigma^2 S^2 V_{SS}(S, t) + (\alpha - \lambda\kappa)SV_S(S, t) + V_t(S, t) \\ &\quad + \lambda E[V(SY, t) - V(S, t)]/V(S, t) \end{aligned} \quad (2.21)$$

$$\sigma_{opt} = V_S(S, t)\sigma \frac{S}{V(S, t)}. \quad (2.22)$$

Die Indizes an $V(S, t)$ bezeichnen die partiellen Ableitungen. Des weiteren ist der Poisson-rozess $q_{opt}(t)$ des Optionspreises abhängig vom Poisson-Prozess $q(t)$ des Aktienpreises. Das heißt, es tritt nur dann ein Poisson Ereignis für die Option auf, wenn auch eines für den Aktienkurs aufgetreten ist. Wenn in einem solchen Fall die Zufallsvariable Y den Wert y annimmt hat die Zufallsvariable Y_{opt} den Wert $\frac{V(Sy, t)}{V(S, t)}$. Obwohl die beiden Prozesse voneinander abhängig sind besteht keine lineare Abhängigkeit, da der Optionspreis V nicht linear von S abhängt.

Wir betrachten nun ein Portfolio bestehend aus einer Aktie, einer Option und einer risikolosen Geldanlage mit Zinssatz r . Das Portfolio enthält diese in Anteilen w_1 , w_2 und w_3 mit $\sum_{i=1}^3 w_i = 1$. Wenn P der Wert des Portfolios ist, dann kann seine Dynamik als

$$\frac{dP}{P} = (\alpha_P - \lambda\kappa_P)dt + \sigma_P dW + dq_P \quad (2.23)$$

beschrieben werden, wobei α_P und σ_P den erwarteten Ertrag bzw. die Volatilität des Portfolios wiedergeben, falls das Poisson Ereignis nicht eintritt. $q_P(t)$ ist ein unabhängiger Poisson Prozess mit Parameter λ , $\kappa_{opt} = E[Y_P - 1]$, wobei $(Y_P - 1)$ die prozentuale Veränderung des Portfoliowertes im Falle des Poisson Ereignisses angibt.

Aus (2.16) und (2.20) erhält man

$$\alpha_P = w_1(\alpha - r) + w_2(\alpha_{opt} - r) + r \quad (2.24)$$

$$\sigma_P = w_1\sigma + w_2\sigma_{opt} \quad (2.25)$$

$$Y_P - 1 = w_1(Y - 1) + w_2(V(SY, t) - V(S, t))/V(S, t), \quad (2.26)$$

wobei w_3 durch $w_3 = 1 - w_1 - w_2$ ersetzt wurde. Im Black-Scholes-Fall ($\lambda = 0$) ist es möglich das Portfolio risikolos zu gestalten, indem man $w_1 = w_1^*$ und $w_2 = w_2^*$ so wählt, dass $w_1^*\sigma + w_2^*\sigma_{opt} = 0$ gilt. In diesem Fall muss aus Gründen der Arbitragefreiheit der erwartete Ertrag des Portfolios gleich dem risikolosen Zinssatz r sein. Dies impliziert mit (2.24) und (2.25)

$$\frac{\alpha - r}{\sigma} = \frac{\alpha_{opt} - r}{\sigma_{opt}}. \quad (2.27)$$

Aus (2.21) (mit $\lambda = 0$), (2.22) und (2.27) folgt dann die im vorigen Abschnitt vorgestellte Black-Scholes Gleichung für Optionspreise

$$\mathcal{L}V := \frac{\partial}{\partial t}V + \frac{\sigma^2}{2}S^2 \frac{\partial^2}{\partial S^2}V + rS \frac{\partial}{\partial S}V - rV = 0.$$

Leider lassen sich durch die Anwesenheit des Poisson-Prozesses $dq(t)$ die Gewichte nicht so wählen, dass man das Sprungrisiko ausschaltet. Dies liegt daran, dass die Portfolio-Optimierung ein linearer Prozess ist, der Preis der Option jedoch nicht linear vom Kurs des Wertpapiers abhängt. Allerdings kann man sich die Eigenschaften des Portfoliowertes

ansehen, wenn man die Gewichte wie im Black-Scholes-Fall wählt. Wenn P^* der Wert des Portfolios im Falle der Black-Scholes Gewichtung ist, dann folgt aus (2.23)

$$\frac{dP^*}{P^*} = (\alpha_P^* - \lambda\kappa_P^*)dt + dq_P^* \quad (2.28)$$

Hierbei ist zu bemerken, dass der Wert des Portfolios ein reiner Sprungprozess ist, da der kontinuierliche Teil der Aktienkursbewegung durch die Wahl der Parameter nicht mehr vorkommt. (2.28) kann nun analog zu (2.18) umgeschrieben werden in

$$\frac{dP^*}{P^*} = \begin{cases} (\alpha_P^* - \lambda\kappa_P^*)dt & , \text{ falls das Poisson Ereignis nicht eintritt} \\ (\alpha_P^* - \lambda\kappa_P^*)dt + (Y_P^* - 1) & , \text{ falls das Poisson Ereignis eintritt.} \end{cases} \quad (2.29)$$

Es ist nun leicht zu sehen, dass der Wert des Portfolios größtenteils vorhersagbar ist und $(\alpha_P^* - \lambda\kappa_P^*)$ beträgt. Allerdings gibt es in jedem Zeitintervall der Länge $\frac{1}{\lambda}$ durchschnittlich einen Sprung. Wir können noch eine weitere qualitative Charakteristik des Portfoliowertes herleiten. Aus (2.22) und (2.26) folgt

$$Y_P^* - 1 = w_2^* \frac{V(SY, t) - V(S, t) - V_S(S, t)(SY - S)}{V(S, t)}. \quad (2.30)$$

2.3.3 Eine Optionspreisformel

Wie wir im vorigen Abschnitt gesehen haben, gibt es keine Möglichkeit, ein risikoloses Portfolio zu konstruieren, weshalb die Idee von Black und Scholes auf dieses Problem nicht anwendbar ist. Nach [Sam65] ist es jedoch möglich, eine Formel für den Optionspreis zu bestimmen, wenn er als Funktion vom Aktienkurs und der verbleibenden Zeit bis zum Fälligkeitszeitpunkt dargestellt wird. Sei $g(S, \tau)$ der Ausgleich, der den erwarteten Wert der Option widerspiegelt, wobei S der aktuelle Kurs des Wertpapiers und $\tau = T - t$ die noch verbleibende Zeit bis zum Ausübungszeitpunkt ist. Dann erhalten wir aus (2.21), dass der Wert V in Abhängigkeit von τ anstatt von t der partiellen Differentialgleichung

$$0 = \frac{1}{2}\sigma^2 S^2 V_{SS} + (\alpha - \lambda\kappa)SV_S - V_\tau - g(S, \tau)V + \lambda E[V(SY, \tau) - V(S, \tau)] \quad (2.31)$$

mit den Randbedingungen

$$\begin{aligned} V(0, \tau) &= 0 \\ V(S, 0) &= (S - K)^+ \end{aligned} \quad (2.32)$$

genügen muss.

Eine weitere Näherung an das Problem folgt dem Ansatz, dass das „Capital Asset Pricing Modell“ aus [BS73] eine gültige Beschreibung für die Erträge des Ausgleichs ist. Im vorangehenden Abschnitt wurde gezeigt, dass die Dynamik des Optionswertes aus zwei Komponenten besteht, einem kontinuierlichen Teil und einem Sprungteil. Letzterer beschreibt die Sprünge, die durch wichtige neue Informationen ausgelöst werden. Falls diese Informationen branchen- oder firmenbezogen sind so haben sie keine Auswirkungen auf den gesamten Markt. Wenn die Ursache der Sprünge solche Informationen sind, so repräsentiert der Ertrag der Aktie ein „nicht systematisches Risiko“, das heißt die Sprünge sind nicht mit dem Markt korreliert. Wenn wir uns (2.28) ansehen, wird klar, dass die Sprungkomponente die einzige Quelle für Unsicherheiten im Ertrag ist. Wenn nun das „Capital Asset Pricing Modell“ gilt, muss der erwartete Ertrag gleich dem risikolosen Zinssatz sein,

d.h. $\alpha_P^* = r$. Nach (2.24) folgt daraus allerdings $w_1^*(\alpha - r) + w_2^*(\alpha_{opt} - r) = 0$, oder man ersetzt w_1^* und w_2^* und erhält

$$\frac{\alpha - r}{\sigma} = \frac{\alpha_{opt} - r}{\sigma_{opt}} \quad (2.33)$$

Dies impliziert aber zusammen mit (2.21) und (2.22), dass der Optionspreis V

$$0 = \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - V_\tau - rV + \lambda E[V(SY, \tau) - V(S, \tau)] \quad (2.34)$$

mit den Randbedingungen (2.32) genügen muss. Formal hat diese Gleichung dieselbe Form wie (2.31), sie hängt aber weder von α noch von $g(S, \tau)$ ab, sondern es taucht wie in der Black-Scholes-Formel nur der risikolose Zinssatz r auf. Darüberhinaus reduziert sich (2.34) im Fall $\lambda = 0$ auf die Black-Scholes-Gleichung. Es bleibt noch zu erwähnen, dass die Sprünge, obwohl sie ein komplett nicht-systematisches Risiko darstellen trotzdem einen Einfluss auf den Ausgleichs-Optionspreis haben. Man kann also einen fairen Preis für die Option nur unter Berücksichtigung des Sprungteils angeben.

Mit der Definition des Erwartungswertes

$$E = \int_0^\infty Y g(Y) dY \quad (2.35)$$

lässt sich (2.34) umschreiben in

$$\begin{aligned} V_\tau &= \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - rV + \lambda \left(\int_0^\infty V(SY, \tau) g(Y) dY - V(S, \tau) \right) \\ &= \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - (r + \lambda)V + \lambda \int_0^\infty V(SY, \tau) g(Y) dY. \end{aligned} \quad (2.36)$$

$g(Y)$ ist hierbei die Wahrscheinlichkeitsdichtefunktion des Sprungprozesses. Näheres zu Dichtefunktionen und deren Eigenschaften findet sich in [Bau02]. Im weiteren Verlauf werden wir die Gleichung (2.16) ersetzen durch

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) + dq(t). \quad (2.37)$$

Einen Beweis für die Übereinstimmung der beiden Darstellungen findet man in [Gla04]. Bei der Bewertung einer Amerikanischen Option in Sprung-Diffusions-Modellen treten die gleichen Schwierigkeiten wie im Black-Scholes-Modell auf. Das Ausübungsrecht des Halters führt zu einem unbekanntem freien Rand und dieser wiederum auf ein partielles Integro-Differential-Ungleichungssystem. Die Herleitung dieser Ungleichung verläuft analog zu Abschnitt 2.2.4 und wird daher hier nicht wiederholt. Letztendlich erhalten wir für Amerikanische Put-Optionen das linear komplementäre Problem

$$\begin{aligned} \mathcal{L}_{SD} V &\geq 0 && \text{für } S > 0, \\ V(S, t) &\geq (K - S)^+ && \text{für } S > 0, \\ \mathcal{L}_{SD} V(V(S, t) - (K - S)^+) &= 0 && \text{für } S > 0 \end{aligned} \quad (2.38)$$

mit der Anfangsbedingung

$$V(S, 0) = (K - S)^+ \quad (2.39)$$

und dem Operator

$$\mathcal{L}_{SD} = -V_\tau \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - rV + \lambda \left(\int_0^\infty V(SY, \tau) g(Y) dY - V(S, \tau) \right). \quad (2.40)$$

2.3.4 Geschlossene Lösungen für das Merton Modell

Leider ist es im Gegensatz zur Black-Scholes-Gleichung (2.3) nicht einmal für Europäische Optionen möglich, für die von Merton entwickelte PIDE (2.34) eine allgemeine geschlossene Formel zur Lösung anzugeben, da das nicht-systematische Risiko des Sprungteils bei der Berechnung des fairen Preises berücksichtigt werden muss. Merton gibt allerdings in [Mer76] zwei Spezialfälle an, in denen eine analytische Lösung gefunden werden kann. Der eine Fall ist der, dass im Falle eines Sprungs der Kurs des Wertpapiers auf 0 fällt. Diese Variante wird als plötzlicher Ruin bezeichnet. Die andere Möglichkeit ist die, dass es sich bei der Verteilung der Sprungintensitäten um eine Lognormalverteilung handelt. Im folgenden Abschnitt werden die Lösungsformeln für diese beiden Fälle angegeben.

Plötzlicher Ruin

Die erste Möglichkeit für die Formulierung einer analytischen Lösung ist die des plötzlichen Ruins. Wenn das Poisson-Ereignis eintritt fällt, der Aktienkurs auf 0. Dies bedeutet, dass die Zufallsvariable Y , die die Änderung im Falle des Poisson-Ereignisses angibt, mit Wahrscheinlichkeit 1 gleich 0 ist. Die prozentuale Änderung des Aktienkurses liegt dann bei $(Y - 1) = -1$, spiegelt also den völligen Wertverlust wieder. In diesem Fall ist der Wert einer Europäischen Call-Option mit verbleibender Laufzeit $\tau = T - t$

$$\begin{aligned} V_C(S, \tau) &= e^{-\lambda\tau} BS_C(Se^{\lambda\tau}, \tau, K, \sigma^2, r) \\ &= e^{-\lambda\tau} \left(Se^{\lambda\tau} N(d_1) - Ke^{-r\tau} N(d_2) \right) \\ &= SN(d_1) - Ke^{-(r+\lambda)\tau} N(d_2) \end{aligned} \quad (2.41)$$

mit

$$\begin{aligned} d_1 &= \frac{\ln(Se^{\lambda\tau}/K) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \\ &= \frac{\ln(S/K) + (r + \lambda + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \end{aligned}$$

und

$$d_2 = d_1 - \sigma\sqrt{\tau}.$$

Damit lässt sich (2.41) schreiben als

$$V_C(S, \tau) = BS_C(S, \tau, K, \sigma^2, r + \lambda). \quad (2.42)$$

$BS_C(S, \tau, K, \sigma^2, r + \lambda)$ ist hierbei die Lösung der Black-Scholes-Gleichung für die angegebenen Parameter und $N(\cdot)$ die Normalverteilung. Diese Lösung ist bis auf die höhere Zinsrate $\tilde{r} = r + \lambda$ identisch mit der Standard Black-Scholes Lösung. In [Mer73] wird gezeigt, dass der Optionspreis eine wachsende Funktion der Zinsrate ist. Somit ist eine Option auf eine Aktie mit positiver Wahrscheinlichkeit auf einen plötzlichen Ruin teurer als eine Option auf eine Aktie, die diese Möglichkeit nicht berücksichtigt.

Lognormalverteilte Sprünge

Im zweiten Fall wird vorausgesetzt, dass die Zufallsvariable Y , welche die Kursänderung im Sprungfall angibt, lognormalverteilt ist. Das heißt, für die Wahrscheinlichkeitsdichtefunktion $g(Y)$ gilt

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right). \quad (2.43)$$

Der Preis einer Call-Option kann unter dieser Bedingung wie folgt geschrieben werden:

$$V_C(S, \tau) = e^{-r\tau} \sum_{n=0}^{\infty} P(n \text{ Sprünge}) E_0 [\max(0, S_T - K) | n \text{ Sprünge}] \quad (2.44)$$

$$= \sum_{n=0}^{\infty} \left[\frac{e^{-\lambda\tau} (\lambda\tau)^n}{n!} \right] E_{0, X_n} [BS_C(SX_n e^{-\lambda\kappa\tau}, \tau, K, \sigma^2, r)] \quad (2.45)$$

$$= \sum_{n=0}^{\infty} \left[\frac{e^{-\tilde{\lambda}\tau} (\tilde{\lambda}\tau)^n}{n!} \right] [SN(d_{1,n}) - e^{-r_n\tau} KN(d_{2,n})] \quad (2.46)$$

$$= \sum_{n=0}^{\infty} \left[\frac{e^{-\tilde{\lambda}\tau} (\tilde{\lambda}\tau)^n}{n!} \right] BS_C(S, \tau, K, v_n^2, r_n), \quad (2.47)$$

wobei

E_{0, X_n} = Erwartungswert bezüglich der Verteilung der X_n zum Zeitpunkt 0,

$$S_T = S \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right),$$

$$\tilde{\lambda} = \lambda(1 + \kappa),$$

$$\kappa = E[Y - 1] = \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right) - 1,$$

$$d_{1,n} = \frac{\ln(S/K) + \left(r_n + \frac{1}{2}v_n^2\right)\tau}{v_n\sqrt{\tau}},$$

$$d_{2,n} = d_{1,n} - v_n\sqrt{\tau},$$

$$r_n = r - \lambda\kappa + \frac{n(\log(1 + \kappa))}{\tau},$$

$$v_n^2 = \sigma^2 + \frac{n\sigma_J^2}{2},$$

μ_J = Erwartungswert der Sprungverteilung,

σ_J^2 = Varianz der Sprungverteilung.

Wir erhalten also auch im Fall der lognormalverteilten Sprünge eine Anwendung der Black-Scholes-Formel mit veränderten Parametern. Zur Berechnung des Optionspreises muss eine unendliche, jedoch konvergente, mit einer Poisson-Verteilung gewichtete Summe ausgewertet werden.

2.4 Kou-Modell

Ein weiteres Modell zur Bewertung von Optionen ist das in [Kou02] von S. G. Kou beschriebene. Im Gegensatz zu Merton geht Kou davon aus, dass die Sprünge des Wertpapiers nicht lognormalverteilt sind, sondern einer Doppel-Exponential-Verteilung folgen. Die übrigen Marktannahmen bleiben erhalten, es ändert sich also nur die Modellierung des Aktienkurses.

2.4.1 Beschreibung des Modells

Der Kurs der Aktie wird bei Kou durch folgende stochastische partielle Differentialgleichung beschrieben

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) + d \left(\sum_{i=0}^{N(t)} (V_i - 1) \right). \quad (2.48)$$

Hierbei entspricht, wie auch schon in den vorangegangenen Modellen μ dem erwarteten Ertrag der Aktie und σ der Volatilität. $W(t)$ ist eine Standard Brownsche Bewegung, $N(t)$ ein Poisson Prozess mit Parameter λ und die V_i sind unabhängige, gleich verteilte, positive Zufallsvariablen für die gilt, dass $Y = \ln(V)$ eine asymmetrische Doppel-Exponential-Verteilung hat. Die Wahrscheinlichkeitsdichtefunktion der Zufallsvariablen Y lautet in diesem Fall

$$g_Y(x) = p\eta_1 e^{-\eta_1 x} 1_{\{x \geq 0\}} + q\eta_2 e^{\eta_2 x} 1_{\{x < 0\}}, \quad \eta_1 > 1, \quad \eta_2 > 0. \quad (2.49)$$

Es gilt weiterhin $p, q \geq 0$, $p + q = 1$. p und q repräsentieren die Wahrscheinlichkeiten für Aufwärts- bzw. Abwärtssprünge. Mit anderen Worten

$$\ln(V) = Y \stackrel{d}{=} \left\{ \begin{array}{l} \zeta^+, \text{ mit Wahrscheinlichkeit } p \\ -\zeta^-, \text{ mit Wahrscheinlichkeit } q \end{array} \right\}, \quad (2.50)$$

wobei ζ^+ und ζ^- exponential verteilte Zufallsvariablen mit Erwartungswert $\frac{1}{\eta_1}$ bzw. $\frac{1}{\eta_2}$ sind und $\stackrel{d}{=}$ Gleichheit in der Verteilung bedeutet. In der Gleichung (2.49) sind alle Zufallsvariablen unabhängig und der Einfachheit halber werden die Drift μ und die Volatilität als konstant angenommen. Diese Einschränkungen lassen sich jedoch leicht zur Entwicklung eines allgemeineren Modells aufheben.

Die Lösung der Differentialgleichung (2.49) liefert uns folgende Dynamik des Wertpapierpreises

$$S(t) = S(0) \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W(t) \right) \prod_{i=1}^{N(t)} V_i. \quad (2.51)$$

Dabei gilt $E[Y] = \frac{p}{\eta_1} - \frac{q}{\eta_2}$, $Var[Y] = qp \left(\frac{1}{\eta_1} + \frac{1}{\eta_2} \right)^2 + \left(\frac{p}{\eta_1^2} + \frac{q}{\eta_2^2} \right)$ und

$$\begin{aligned} E[V] &= E[e^Y] \\ &= q \frac{\eta_2}{\eta_2 + 1} + p \frac{\eta_1}{\eta_1 - 1}, \quad \eta_1 > 1, \quad \eta_2 > 0. \end{aligned} \quad (2.52)$$

Wenn der Kurs des zugrunde liegenden Wertpapiers der in (2.51) beschriebenen Dynamik folgt, ergibt sich für den Wert V einer entsprechenden Europäischen Option die PIDE

$$V_t = -\frac{1}{2} \sigma^2 S^2 V_{SS} - (r - \lambda \alpha) S V_S + (r + \lambda) V - \lambda \int_{\mathbb{R}^+} V(Sy, t) g_Y(y) dy, \quad (2.53)$$

siehe [CT04] und [Toi08].

2.4.2 Eigenschaften des Modells

„volatility smile“ und „leptokurtic feature“

Obwohl jedes Modell ungenau ist, kann man Eigenschaften angeben, die zur annähernd realistischen Nachbildung eines Aktienkurses erfüllt sein sollten. Hierzu gehören das „asymmetric leptokurtic feature“ und der „volatility smile“. Ersteres besagt, dass die Verteilung des Ertrages der Aktie nach links verzogen ist und einen höheren Peak und breitere Tails als die Normalverteilung besitzt. Die zweite Eigenschaft besagt, dass die implizite Volatilität des Kurses eine streng konvexe Funktion bezüglich des Ausübungspreises ist.

Der „volatility smile“ kann gezeigt werden indem die implizite Volatilität in Abhängigkeit vom Ausübungspreis berechnet wird. Dies ergibt im Falle des Kou-Modells eine streng konvexe Funktion, im Falle des Black-Scholes-Modells jedoch eine konstante Volatilität. Das „leptokurtic feature“ ist hingegen schwieriger zu zeigen, weshalb wir dies an einem Beispiel betrachten. Mit Hilfe von (2.51) ergibt sich für den Wert der Aktie über ein Zeitintervall Δt

$$\begin{aligned} \frac{\Delta S(t)}{S(t)} &= \frac{S(t + \Delta t)}{S(t)} - 1 \\ &= \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) \Delta t + \sigma (W(t + \Delta t) - W(t)) + \sum_{i=N(t)+1}^{N(t+\Delta t)} Y_i \right) - 1. \end{aligned}$$

Die leere Summe ist dabei gleich 0. Wenn der Zeitraum Δt klein genug ist, z.B. bei der Beobachtung über einen Handelstag, kann die Verteilung des Wertes approximiert werden, indem man die Terme höherer Ordnung als Δt ignoriert und $e^x \approx 1 + x + x^2/2$ setzt. Es gilt dann :

$$\frac{\Delta S(t)}{S(t)} \approx \mu \Delta t + \sigma Z \sqrt{\Delta t} + B \cdot Y, \quad (2.54)$$

wobei Z und B Standard- bzw. Bernoulli-verteilte Zufallsvariablen sind mit $P(B = 1) = \lambda \Delta t$ und $P(B = 0) = 1 - \lambda \Delta t$. Die Dichte die sich nun für die rechte Seite von (2.54) ergibt ist in Abbildung 2.1 gegen eine Normalverteilung mit gleichem Erwartungswert und gleicher Varianz aufgetragen.

Die Parameter sind $\Delta t = 1/250$ (ein Handelstag), $\sigma = 0.2$, $\mu = 0.15$, $\lambda = 10$, $p = 0.3$, $\eta_1 = 50$ und $\eta_2 = 25$. Sie stammen aus [Kou02] und spiegeln annähernd das Verhalten amerikanischer Aktien wieder. Das „leptokurtic feature“ ist in der Abbildung 2.4.2 leicht zu erkennen, der Peak der Normalverteilung liegt bei ca. 25, derjenige der Dichte aus dem Kou-Modell liegt bei ca. 31. Es ist auch direkt zu sehen, dass die Tails breiter sind als die der Normalverteilung, insbesondere auf der linken Seite des Graphen.

Geschlossene Lösungen im Kou Modell

Die beiden Eigenschaften aus dem vorangehenden Abschnitt können ebenso durch das Merton-Modell verwirklicht werden. Geschlossene Lösungen existieren jedoch nur für die Bewertung Europäischer Optionen. Im Black-Scholes-Modell existieren auch geschlossene Lösungen für pfadabhängige Optionen, die jedoch sehr aufwendig in der Auswertung sind, was dazu führt, dass in diesem Fall eine numerische Lösung des Problems schneller ist. Eine geschlossene Lösung für pfadabhängige Optionen im Sprung-Diffusions-Modell von Merton ist nicht möglich. Dies liegt vor allem an der analytischen Traktabilität der Pfade. Es ist

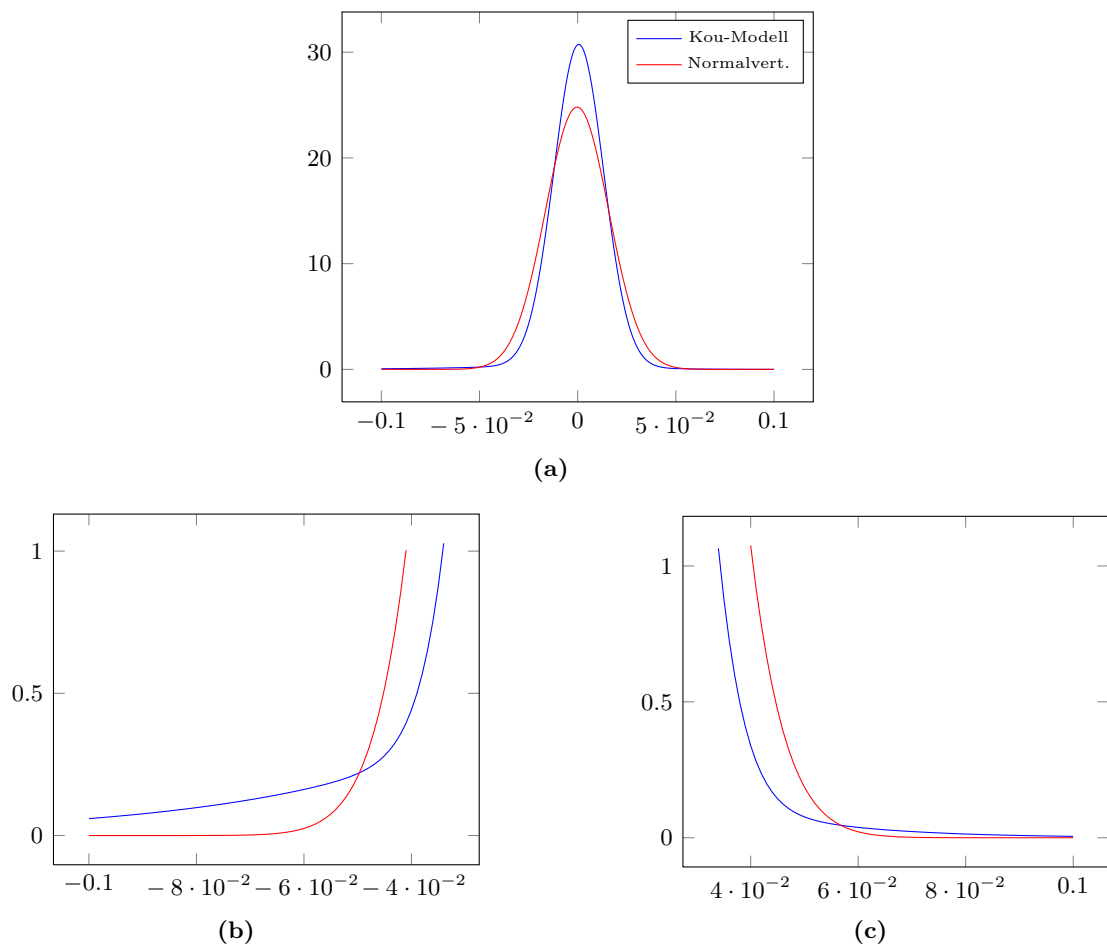


Abbildung 2.3: a) Vergleich der Dichten aus dem Kou-Modell und im Falle einer Normalverteilung insgesamt, b) linke Tails, c) rechte Tails

wichtig, die Anfänge der Pfade der Aktienkurse zu untersuchen, bis sie auf eine Schranke treffen. Ein Problem dabei ist der „overshoot“. Damit wird der Fall bezeichnet, bei dem der Kurs über die Schranke springt. Um eine geschlossene Lösung für pfadabhängige Optionen anzugeben, muss man die Verteilung dieses „overshoot“ kennen. Was allerdings im Merton-Modell nicht möglich ist. Auf Grund der Eigenschaft der Gedächtnislosigkeit der Doppel-Exponential-Verteilung ist dies jedoch im Kou-Modell machbar. Aufgrund der Traktabilität ist es also im Falle einer Doppel-Exponential-Verteilung möglich, analytische geschlossene Lösungen für Europäische und pfadabhängige Optionen anzugeben. Die genaue Herleitung dieser Formeln findet man in [Kou02] und [KW04].

2.5 VG- und CGMY-Modell

Die vorangehenden Sprung-Diffusions-Modelle haben gemeinsam, dass sie beide einen kontinuierlichen Teil und einen Sprungteil verwenden, der durch einen Poisson-Prozess gesteuert wird. Der kontinuierliche Teil soll die infinitesimalen Bewegungen des Aktienkurses abdecken, der Sprungteil die außergewöhnlichen großen Ereignisse. Es entstand die Frage, ob es nicht möglich ist, beides durch einen Sprungprozess mit unendlicher Aktivität zu simulieren. Carr, Geman, Madan und Yor entwickelten hierzu das CGMY-Modell. Darin wird nur ein Sprungprozess zur Simulation des Aktienkurses verwendet, welcher durch geeignete Parameterwahl endliche oder unendliche Aktivität und Varianz realisieren kann. Wir wollen dieses Modell hier vorstellen und orientieren uns dabei hauptsächlich an [CGMY02] und [Alm05]. Das CGMY-Modell stellt eine Erweiterung des Varianz-Gamma-Modells (VG-Modell) dar, welches wir zuerst beschreiben werden. Die Beschreibung orientiert sich hauptsächlich an [MCC98] und [Sch05]. Es sei auch gleich zu Beginn gesagt, dass für die Optionspreisberechnung in beiden Modellen keine geschlossenen Lösungen existieren, weder für Europäische noch für Amerikanische Optionen.

2.5.1 Lévy-Prozesse

Um den VG- und den CGMY-Prozess besser charakterisieren zu können, folgt hier eine kleine Einleitung zu Lévy-Prozessen.

Definition 2.5.1 [CHARAKTERISTISCHE FUNKTION]

Die Charakteristische Funktion Φ einer Zufallsvariable X oder ihrer Verteilung ist die Fourier-Stieltjes-Transformation der Verteilungsfunktion $F(x) = P(X \leq x)$, $x \in \mathbb{R}$

$$\Phi_X(z) = E[\exp(izX)] = \int_{-\infty}^{\infty} \exp(izx) dF(x), \quad (2.55)$$

wobei $E[\cdot]$ den Erwartungswertoperator bezeichnet.

Wichtige Eigenschaften der charakteristischen Funktion sind $\Phi(0) = 1$ und $|\Phi(z)| \leq 1$ für alle $z \in \mathbb{R}$. Die charakteristische Funktion existiert immer und ist stetig. Eine weitere wichtige Eigenschaft ist, dass die Verteilungsfunktion F eindeutig durch Φ bestimmt wird.

Definition 2.5.2 [UNENDLICH OFT TEILBAR]

Eine Zufallsvariable X oder die zu ihr gehörende Verteilung nennt man unendlich oft teilbar, wenn für alle $n \in \mathbb{N}$ Zufallsvariablen $Y_1^{(n)} \dots Y_n^{(n)}$ existieren, so dass X als

$$X \stackrel{d}{=} \sum_{i=1}^n Y_i^{(n)} \quad (2.56)$$

geschrieben werden kann. Dabei meint $\stackrel{d}{=}$ Gleichheit in Verteilung.

Ein weiterer wichtiger Begriff ist der *charakteristische Exponent* einer Zufallsvariable oder ihrer Verteilung. Damit bezeichnet man die logarithmierte charakteristische Funktion $\psi(z) = \log \Phi(z)$, welche der Lévy-Kintchine Formel

$$\psi(z) = i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{-\chi}^{\infty} (\exp(izx) - 1 - izx1_{\{|x|<1\}})\nu(dx) \quad (2.57)$$

genügt. Dabei sind $\gamma, \sigma \in \mathbb{R}$ und ν ist ein Maß auf $\mathbb{R} \setminus \{0\}$ mit

$$\int_{-\infty}^{\infty} \inf\{1, x^2\}\nu(dx) = \int_{-\chi}^{\infty} (1 \wedge x^2)\nu(dx) < \infty. \quad (2.58)$$

Ein Lévy-Prozess lässt sich also über den charakteristischen Exponenten bestimmen, dieser wiederum ist durch das Triplett $(\gamma, \sigma^2, \nu(dx))$ eindeutig bestimmt. Dieses Triplett wird als das Lévy-Triplett des Prozesses bezeichnet.

Definition 2.5.3 [LÉVY-PROZESS]

Ein reellwertiger stochastischer Prozess $X = (X_t)_{t \geq 0}$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ ist ein Lévy-Prozess, wenn er folgende Eigenschaften besitzt:

1. $X_0 = 0$ \mathbb{P} -fast sicher.
2. Die Zuwächse des Prozesses sind unabhängig.
3. Die Zuwächse sind stationär und zeitlich unabhängig.
4. X_t ist stochastisch stetig, d.h. $\forall t \geq 0$ und $\epsilon \geq 0$ gilt $\lim_{s \rightarrow t} \mathbb{P}(|X_s - X_t| > \epsilon) = 0$.
5. X besitzt die Cádłág-Eigenschaft, d.h. mit Wahrscheinlichkeit 1 sind die Pfade rechtstetig für $t \geq 0$ und besitzen einen linksseitigen Grenzwert für $t > 0$.

Durch das Lévy-Triplett wird deutlich, dass der Lévy-Prozess aus drei Teilen besteht, dem deterministischen Teil γ , dem brownischen Teil σ^2 und dem Sprungteil, der durch $\nu(dx)$ beschrieben wird. Nach [Sat99] kann man diese Prozesse in drei Klassen unterteilen.

1. Klasse A, falls $\sigma^2 = 0$ und $\nu(\mathbb{R}) < \infty$,
2. Klasse B, falls $\sigma^2 = 0$, $\nu(\mathbb{R}) = \infty$ und $\int_{(|x| \leq 1)} |x|\nu(dx) < \infty$
3. Klasse C, falls $\sigma^2 \neq 0$ oder $\int_{(|x| \leq 1)} |x|\nu(dx) = \infty$

Die Pfade verlaufen in jeder der drei Klassen anders. $\sigma^2 = 0$ bedeutet, dass der Prozess nur endlich viele Sprünge in jedem Intervall macht, also von endlicher Aktivität ist. Gilt $\nu(\mathbb{R}) < \infty$ oder $\nu(\mathbb{R}) = \infty$ und $\int_{(|x| \leq 1)} |x|\nu(dx) < \infty$, so ist der Prozess von endlicher Variation. Prozesse aus der Klasse A sind daher von endlicher Aktivität. Für Prozesse aus den Klassen A und B gilt, dass sie in jedem endlichen Intervall von endlicher Variation sind, für diejenigen aus Klasse C gilt, dass sie in jedem Intervall von unbegrenzter Variation sind.

Nach diesen Vorbemerkungen kommen wir nun zur Vorstellung des VG- und des CGMY-Modells.

2.5.2 VG-Modell

Das Varianz-Gamma-Modell basiert auf der Annahme, dass der Aktienkurs einem Prozess folgt, dessen Zuwächse Varianz-Gamma verteilt sind. Wir werden diesen Prozess hier etwas genauer vorstellen. Allgemein hat ein Aktienkurs in exponentiellen Lévy-Modellen folgende Form:

$$S(t) = S(0) \exp(\mu t + X_t) \quad (2.59)$$

Hierbei ist μ der erwartete Ertrag der Aktie und X_t ein Lévy-Prozess. Im Falle des VG-Modells lässt sich der Prozess unter dem risikoneutralen Maß dann als

$$S(t) = S(0) \exp(rt + X_t(\sigma, \nu, \theta) + \omega t) \quad (2.60)$$

schreiben, wobei $\omega = \frac{1}{\nu} \ln((1 - \theta\nu - \sigma^2\nu/2))$ und X_t die Differenz zweier Gamma-Prozesse ist.

Die charakteristische Funktion des Varianz-Gamma-Prozesses ist durch

$$\phi_{VG}(z; \sigma, \nu, \theta) = (1 - iz\theta\nu + \frac{1}{2}\sigma^2\nu z^2)^{-1/\nu} \quad (2.61)$$

gegeben. Es gibt für den Varianz-Gamma Prozess zwei unterschiedliche Parametrisierungen. Zum einen die (σ, ν, θ) -Variante, welche wir zur Darstellung der charakteristischen Funktion (2.61) verwendet haben, und die (C, G, M) Variante, welche wir im Folgenden verwenden werden, da sich der Prozess auf diese Weise besser mit dem CGMY-Prozess vergleichen lässt. Nach [MCC98] kann man die Lévy-Dichte des Prozesses dann als

$$\nu_{VG} = \begin{cases} C \frac{\exp(Gx)}{|x|} dx & x < 0, \\ C \frac{\exp(Mx)}{|x|} dx & x > 0, \end{cases} \quad (2.62)$$

schreiben. Die Parameter C, G und M ergeben sich wie folgt:

$$\begin{aligned} C &= \frac{1}{\nu} > 0, \\ G &= \left(\sqrt{\frac{1}{4}\theta^2\nu^2 + \frac{1}{2}\sigma^2\nu} - \frac{1}{2}\theta\nu \right)^{-1} > 0, \\ M &= \left(\sqrt{\frac{1}{4}\theta^2\nu^2 + \frac{1}{2}\sigma^2\nu} + \frac{1}{2}\theta\nu \right)^{-1} > 0. \end{aligned}$$

Das Lévy-Triplett ist dann gegeben durch $(\gamma, 0, \nu_{CG})$, mit

$$\gamma = \frac{-C(G(\exp(-M) - 1) - M(\exp(-G) - 1))}{MG}.$$

In der CGM -Parametrisierung ergibt sich für die charakteristische Funktion

$$\phi_{VG}(z; C, G, M) = \left(\frac{GM}{GM + (M - G)iz + z^2} \right)^C.$$

Den Erwartungswert, die Varianz, die Schiefe und die Kurtosis des Prozesses werden wir am Ende des Kapitels im Vergleich zum CGMY-Prozess und zur Brownschen Bewegung angeben, ebenso die Bedeutung der einzelnen Parameter für den verwendeten Prozess. Die

PIDE, die die Dynamik des Optionspreises beschreibt, dessen zugrunde liegendes Wertpapier einem VG-Prozess folgt, ist gegeben durch

$$V_t = \frac{\sigma^2}{2} V_{xx} + \left(r - \frac{\sigma^2}{2}\right) V_x - rV + \int_{\mathbb{R}^+} (V(t, x+y) - V(t, x) - (e^y - 1)V_x(t, x)) \nu_{VG}(dy). \quad (2.63)$$

Sie ergibt sich aus einem Spezialfall des Ito Lemmas und der Variablensubstitution $x = \log(S)$. Die genaue Herleitung der Formel findet man in [Rai00].

2.5.3 CGMY-Modell

Das CGMY-Modell ist eine Generalisierung des im vorangegangenen Abschnitts vorgestellten VG-Modells. In der Lévy-Dichte des Prozesses wird ein weiterer Parameter eingeführt, um die Feinstruktur des Prozesses zu steuern. Das Modell wurde im Jahr 2002 von Carr, Geman, Madan und Yor in [CGMY02] vorgestellt. Der CGMY-Prozess besitzt also vier Parameter und die dazu gehörige Lévy-Dichte lautet:

$$\nu_{CGMY} = \begin{cases} C \frac{\exp(-G|x|)}{|x|^{Y+1}} & x < 0 \\ C \frac{\exp(-M|x|)}{|x|^{Y+1}} & x > 0 \end{cases}, \quad (2.64)$$

wobei $C > 0$, $G \geq 0$, $M \geq 0$ und $Y < 2$. Die Bedingung $Y > 2$ ist notwendig, da sonst kein gültiges Lévy-Maß vorliegen würde. Das Lévy-Triplett des CGMY Prozesses lautet $(\gamma, 0, \nu_{CGMY})$, mit

$$\gamma = C \left(\int_0^1 \exp(-Mx) x^{-Y} dx - \int_{-1}^0 \exp(Gx) |x|^{-Y} dx \right). \quad (2.65)$$

Die PIDE zur Berechnung des Optionspreises für den Fall, dass der Kurs des Wertpapiers einem CGMY-Prozess folgt, ist beinahe identisch mit (2.63), allerdings wird die Dichte ν_{CGMY} des CGMY-Prozesses anstatt ν_{VG} verwendet.

2.5.4 Die Parameter des CGMY-Prozesses

Der CGMY-Prozess ist mit vier Parametern der flexibelste der bislang vorgestellten Möglichkeiten, einen Aktienkurs zu simulieren, weshalb es angebracht ist, die Funktion der einzelnen Parameter etwas genauer zu betrachten. Mit Hilfe der vier Parameter lassen sich die Eigenschaften des Prozesses festlegen, insbesondere die Schiefe (Skewness) und die Kurtosis, wobei es sich hierbei um die Maße für die Asymmetrie bzw. für die Breite der Tails der Verteilungsfunktion handelt.

Parameter C

Der Parameter C kann als Steuerung für die Gesamtaktivität des Prozesses angesehen werden. Hält man alle anderen Parameter konstant und integriert über alle Pfade, welche eine minimale Aktivität überschreiten so zeigt sich, dass die Aktivität dieser Pfade durch Änderungen in C gesteuert werden kann. Möchte man zum Beispiel einen Prozess mit zufälliger Aktivität generieren, so kann man C als unabhängigen, positiven Prozess wählen. Im Sonderfall $G = M$ ist die Verteilung des CGMY-Prozesses symmetrisch. Madan et al. haben in [MCC98] gezeigt, dass in diesem Fall die Kurtosis der Verteilung durch C bestimmt wird. Dieser Fall wird ebenfalls in [Kop95] untersucht, des weiteren wird dort eine alternative Schreibweise für die charakteristische Funktion vorgestellt.

Parameter G, M

Die Parameter G und M sind für den exponentiellen Abfall des linken bzw. rechten Tails verantwortlich. Im Fall $G \neq M$ führt dies zu einer verzerrten Verteilung (skewed distribution) des Prozesses. Für $G < M$ ist der linke Tail breiter als der rechte, was, wie schon vorher angemerkt, mit den Eigenschaften der risikoneutralen Verteilung übereinstimmt, welche bei Optionspreisen impliziert wird. Wenn G und M von der risikoneutralen Verteilung impliziert werden, so gibt die Differenz das Maß zwischen Abfallen und Wachsen des Prozesses an. Die Summe hingegen bestimmt den Preis einer großen Bewegung gegenüber einer kleinen.

In der statistischen Verteilung hingegen gibt die Differenz die relative Frequenz von Fall zu Anstieg und die Summe die Wahrscheinlichkeit großer Bewegungen gegenüber kleinen an.

Parameter Y

Der Parameter Y ist verantwortlich für die Feinstruktur des Prozesses. So ist es zum Beispiel von Interesse, ob die Kurssprünge nach oben und unten eine vollständig monotone Dichte besitzen oder ob der Prozess endliche oder unendliche Aktivität besitzt. Dies lässt sich über Y steuern, weshalb wir daher diese Eigenschaften hier kurz beschreiben werden.

Vollständige Monotonie Eine vollständig monotone Lévy-Dichte (VM) setzt das Vorkommen von großen Sprüngen in Beziehung zu dem von kleinen Sprüngen, indem unter anderem vorausgesetzt wird, dass große Sprünge seltener vorkommen als kleine. Dies bedeutet, dass $\nu(x)$ bei wachsendem $|x|$ abfällt, oder dass die Ableitung der Dichte positiv ist für $x < 0$ und negativ, falls $x > 0$. Das Vorzeichen der ersten Ableitung wechselt also in 0 ist ansonsten jedoch konstant. Für vollständige Monotonie muss diese Eigenschaft ebenfalls für alle weiteren Ableitungen gelten. Eine Beschreibung einiger vollständig monotoner Dichten sind in [GMY01] und [CGMY02] zu finden.

Endliche Variation In der Theorie zur Bewertung von Optionen sind Prozesse mit endlicher Variation (EV) und endlicher Aktivität (EA) wichtig, um die Änderungen vom statistischen zum risikoneutralen Maß zu erklären, weil sie größere Flexibilität zwischen den lokalen Eigenschaften der Martingalkomponenten der beiden Maße ermöglichen. Bei einem Prozess mit endlicher Variation, z.B. der Brownschen Bewegung, ist die Volatilität, und somit die lokale Martingalkomponente, invariant bezüglich der Maßänderung. Bei Sprungprozessen mit unendlicher Variation bedeutet die Äquivalenz der Maßänderung, dass die Differenz zwischen der risikoneutralen und der statistischen Lévy-Dichte von endlicher Variation ist. Dies folgt entweder aus der Tatsache, dass sie beide denselben Exponenten aufweisen, oder dass sie beide im selben Maß von unendlicher Variation sind. Falls beide von endlicher Variation sind gilt dies natürlich auch für die Differenz, eine parametrische Einschränkung ist daher in diesem Fall nicht nötig. Diese Beobachtungen sind vor allem wichtig, wenn man Optionsdaten betrachtet, die zeigen, dass die risikoneutrale Dichte wesentlich größer ist als die statistische.

Endliche Aktivität Prozesse von endlicher Aktivität sind dann besonders wichtig, wenn man die Wertpapiere nach ihrem Aktivitätslevel gruppieren möchte. Die Anwendung von Prozessen mit unendlicher Aktivität wird danach als erste Approximation für Märkte mit hohem Aktivitätslevel angesehen.

Diese drei Eigenschaften werden über den Parameter Y gesteuert. In Tabelle 2.2 ist festgehalten, welche Wahl von Y welche Eigenschaft bedingt.

Wert von Y	Eigenschaften des Prozesses
$Y < -1$	nicht VM, EA
$-1 < Y < 0$	VM, EA
$0 < Y < 1$	VM, EA, EV
$1 < Y < 2$	VM, UV, UA

Tabelle 2.2: Zusammenhang zwischen der Wahl von Y und den Prozesseigenschaften um 0

2.5.5 Eigenschaftenvergleich der Modelle

Nachdem nun einige Modelle vorgestellt wurden, werden wir in den folgenden Tabellen 2.3-2.6 noch einmal Erwartungswert, Varianz, Schiefe und Kurtosis der verschiedenen Verteilungsfunktionen angeben.

Verteilungsfunktion	Erwartungswert
$N(\mu, \sigma^2)$	μ
$\text{Poisson}(\lambda)$	λ
$\text{Doppel-Log-Exp}(p, q, \eta_1, \eta_2)$	$\frac{p}{\eta_1} - \frac{q}{\eta_2}$
$\text{VG}(\sigma, \nu, \theta)$	θ
$\text{VG}(C, G, M)$	$C(G - M)/(MG)$
$\text{CGMY}(C, G, M, Y)$	$C(M^{Y-1} - G^{Y-1})\Gamma(1 - Y)$

Tabelle 2.3: Erwartungswerte der Verteilungsfunktionen

Verteilungsfunktion	Varianz
$N(\mu, \sigma^2)$	σ^2
$\text{Poisson}(\lambda)$	λ
$\text{Doppel-Log-Exp}(p, q, \eta_1, \eta_2)$	$pq \left(\frac{1}{\eta_1} + \frac{1}{\eta_2} \right)^2 + \frac{p}{\eta_1^2} + \frac{q}{\eta_2^2}$
$\text{VG}(\sigma, \nu, \theta)$	$\sigma^2 + \nu\theta^2$
$\text{VG}(C, G, M)$	$C(G^2 + M^2)/(MG)^2$
$\text{CGMY}(C, G, M, Y)$	$C(M^{Y-2} + G^{Y-2})\Gamma(Y - 2)$

Tabelle 2.4: Varianz der Verteilungsfunktionen

Verteilungsfunktion	Schiefe
$N(\mu, \sigma^2)$	0
$\text{Poisson}(\lambda)$	$\frac{1}{\sqrt{\lambda}}$
$\text{VG}(\sigma, \nu, \theta)$	$\frac{\theta\nu(3\sigma^2 + 2\nu\theta^2)}{(\sigma^2 + \nu\theta^2)^{3/2}}$
$\text{VG}(C, G, M)$	$2C^{-1/2}(G^3 - M^3)/(G^2 + M^2)^{3/2}$
$\text{CGMY}(C, G, M, Y)$	$\frac{C(M^{Y-3} + G^{Y-3})\Gamma(3 - Y)}{(C(M^{Y-2} + G^{Y-2})\Gamma(2 - Y))^{3/2}}$

Tabelle 2.5: Schiefe der Verteilungsfunktionen

Verteilungsfunktion	Kurtosis
$N(\mu, \sigma^2)$	3
$\text{Poisson}(\lambda)$	$3 + \frac{1}{\lambda}$
$\text{VG}(\sigma, \nu, \theta)$	$3(1 + 2\nu - \nu\sigma^4(\sigma^2 + \nu\theta^2)^{-2})$
$\text{VG}(C, G, M)$	$3(1 - 2C^{-1}(G^4 + M^4)/(M^2 + G^2)^2)$
$\text{CGMY}(C, G, M, Y)$	$3 + \frac{C(M^{Y-4} + G^{Y-4})\Gamma(4-Y)}{(C(M^{Y-2} + G^{Y-2})\Gamma(2-Y))^2}$

Tabelle 2.6: Kurtosis der Verteilungsfunktionen

2.6 Mehrdimensionale PIDE

Die bisher vorgestellten Modelle zur Bewertung von Optionen befassen sich nur mit Optionen, denen ein einzelnes Wertpapier zu Grunde liegt. Allerdings kommen in der Praxis sehr häufig Optionen vor, die von zwei oder mehr Kursen abhängen. Auch für diese Optionen muss ein Verfahren zur effizienten Bewertung gefunden werden. Sowohl im Black-Scholes-Modell, als auch im Merton-Modell ist eine Verallgemeinerung auf Basket-Optionen weitestgehend unproblematisch. Nimmt man zum Beispiel eine zweidimensional Basketoption, so kann man die auftretenden Sprünge durch eine zweidimensional Lognormalverteilung realisieren. Dabei ist es durch geeignete Parameterwahl sehr einfach verschiedene Eigenschaften der Sprungprozesse wie zum Beispiel ihre Korrelation abzubilden. Wesentlich schwieriger wird dies im Falle von Sprungprozessen, welchen ein CGMY- oder VG-Prozess zu Grunde liegt. Ein Ansatz um die mehrdimensionalen Lévyprozesse mit den gewünschten Eigenschaften zu realisieren sind *Lévy-Copulas*. Diese sind Funktionen, die die mehrdimensionale Dichte eines Prozesses mit gegebenen Randverteilungen eindeutig bestimmen und dabei Eigenschaften wie z.B. den Grad der Abhängigkeit der Sprungprozesse voneinander abzubilden. Wir werden hier kurz die grundlegenden Prinzipien und Eigenschaften dieser Copulas erläutern und die entsprechende mehrdimensionale Dichte anhand einer speziellen Familie von Lévy-Copulas angeben. Bei der Beschreibung der Copulas richten wir uns in diesem Kapitel hauptsächlich nach [KT06] und [FRS06].

Die Kurse der einzelnen Wertpapiere sind wie im eindimensionalen Fall aus einer Diffusions- und einer Sprungkomponente zusammengesetzt. Wir gehen hier davon aus, dass die Sprungintensitäten zwar identisch verteilt, jedoch nicht für alle Aktienkurse identisch sind. Wir erhalten in der n-dimensionalen PIDE ein n-dimensionales Integral, weshalb wir daher die Änderungen eines einzelnen Kurses als

$$\frac{dS_i}{S_i} = (\mu_i - \lambda\kappa)dt + \sigma_i dW_i + dq_i \quad i = 1, \dots, n \quad (2.66)$$

schreiben können, wobei μ_i und σ_i die Drift bzw. die Volatilität des einzelnen Wertpapiers angeben. W_i bezeichnet eine Brownsche Bewegung und q_i einen Poissonprozess. Wir bezeichnen den Korrelationskoeffizienten zwischen W_i und W_j mit ρ_{ij} , d.h.

$$dW_i dW_j = \rho_{ij} dt \quad i, j = 1, \dots, n, \quad i \neq j. \quad (2.67)$$

Durch Anwendung des mehrdimensionalen Ito-Lemmas erhalten wir dann folgende Dynamik für die Option:

$$dV = \frac{\partial V}{\partial t} dt + \sum_{i=1}^n \frac{\partial V}{\partial S_i} dS_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i \sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} dt. \quad (2.68)$$

Für das mehrdimensionale Black-Scholes-Modell kann wie im eindimensionalen Modell über eine Portfoliokonstruktion gezeigt werden, dass der erwartete Ertrag des Wertpapiers gleich dem risikolosen Zinssatz r ist. Im Merton-Modell folgt diese Annahme wiederum aus dem „Capital Asset Pricing Modell“. Wir führen nun noch die Transformation $\tau = T - t$ durch und erhalten dann die Differentialgleichung

$$\begin{aligned} V_\tau &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i \sigma_j S_i S_j V_{S_i S_j} + (r - \lambda \kappa) \sum_{i=1}^n S_i V_{S_i} - rV \\ &+ \lambda \mathbb{E}[V(\mathbf{S} + \Delta \mathbf{S}, \tau) - V(\mathbf{S}, \tau)]. \end{aligned} \quad (2.69)$$

Dabei verwenden wir die verkürzte Schreibweise $\mathbf{S} = (S_1, \dots, S_n)^\tau$. $\Delta \mathbf{S} = (\Delta S_1, \dots, \Delta S_n)^\tau$ gibt die Änderung im Aktienkurs an. Es ergibt sich daher

$$\begin{aligned} \mathbf{S} + \Delta \mathbf{S} &= \mathbf{S} \mathbf{Y} \\ &= ((S_1 Y_1, \dots, S_n Y_n)). \end{aligned} \quad (2.70)$$

Dabei ist \mathbf{Y} die mehrdimensionale Verteilungsfunktion der Sprünge der einzelnen Kurse. Wir können daher den Erwartungswert in der Differentialgleichung auf folgende Weise umschreiben:

$$\lambda \mathbb{E}[V(\mathbf{S} + \Delta \mathbf{S}, \tau) - V(\mathbf{S}, \tau)] = -\lambda V(\mathbf{S}, \tau) + \lambda \mathbb{E}[V(\mathbf{S} + \Delta \mathbf{S}, \tau)] \quad (2.71)$$

$$= -\lambda V(\mathbf{S}, \tau) + \lambda \mathbb{E}[V(\mathbf{S} \mathbf{Y}, \tau)] \quad (2.72)$$

$$= -\lambda V(\mathbf{S}, \tau) + \lambda \int_{\mathbb{R}^n} V(\mathbf{S} \mathbf{Y}, \tau) g(\mathbf{Y}) d\mathbf{Y}. \quad (2.73)$$

$g(\mathbf{Y})$ ist dann die mehrdimensionale Dichte der Sprungverteilung, also im Merton-Modell die mehrdimensionale Lognormalverteilung. Wir erhalten dann folgende mehrdimensionale PIDE:

$$\begin{aligned} V_\tau &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i \sigma_j S_i S_j V_{S_i S_j} + (r - \lambda \kappa) \sum_{i=1}^n S_i V_{S_i} - (r + \lambda) V \\ &+ \lambda \int_{\mathbb{R}^n} V(\mathbf{S} \mathbf{Y}, \tau) g(\mathbf{Y}) d\mathbf{Y}. \end{aligned} \quad (2.74)$$

Wir führen wie im eindimensionalen Fall zwei Transformationen durch. Wir logarithmieren den Preis und die Integrationsvariable und erhalten dann abschließend die Gleichung

$$\begin{aligned} V_\tau &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i \sigma_j V_{x_i x_j} + \sum_{i=1}^n (r - \lambda \kappa - \frac{\sigma_i^2}{2}) V_{x_i} - (r + \lambda) V \\ &+ \lambda \int_{\mathbb{R}^n} V(\mathbf{x} + \mathbf{y}, \tau) f(\mathbf{y}) d\mathbf{y}. \end{aligned} \quad (2.75)$$

Wie eingangs schon erwähnt, ist es in manchen Modellen einfach, die Dichtefunktion $f(\mathbf{y})$ anzugeben und damit auch Abhängigkeiten der Prozesse untereinander darzustellen. Das zweidimensionale Merton-Modell verwendet dabei die zweidimensionale Lognormalverteilung, dieser Ansatz zur Bewertung zweidimensionaler Optionen wird z.B. in [CF08] gewählt. Im Fall der exponentiellen Lévy-Modelle gilt dies nicht. Wir werden daher den Begriff der Lévy-Copulas einführen, um die mehrdimensionale Lévy-Dichte zu erzeugen.

2.7 Lévy-Copulas

Copulas ermöglichen die Trennung der Abhängigkeiten eines Zufallsvektors von seinen eindimensionalen Randverteilungen. Es gibt dabei grundsätzlich zwei Arten, die Copulas zu betrachten. Einerseits ermöglichen sie eine vollständige Charakterisierung der möglichen Abhängigkeitsstrukturen eines Zufallsvektors mit vorgegebenen Randverteilungen. Andererseits können Copulas dazu genutzt werden, eine mehrdimensionale Verteilung mit vorgegebenen Abhängigkeiten aus eindimensionalen Verteilungen zu erzeugen.

Grundsätzlich ist ein mehrdimensionaler Lévy-Prozess $X = (X_t)_{t \in \mathbb{R}}$ eindeutig durch die Verteilung von X_t zum Zeitpunkt t bestimmt. Man könnte also die Abhängigkeiten durch Copulas C_t von X_t ausdrücken, die hängen allerdings von t ab, und es ist nicht klar, welche Copulas einen d -dimensionalen Lévy-Prozess ergeben. Da X ebenso durch sein Lévy-Triplett $(\gamma, \Sigma, \nu(dx))$ bestimmt wird, ist es naheliegend dieses Triplett zur Generierung einer zeitunabhängigen Copula zu nutzen. Dieser Ansatz wird in [KT06] verfolgt. Der Parameter γ hat dabei keinen Einfluss auf die darzustellenden Abhängigkeiten, und die des Brownschen Teil werden durch die Kovarianzmatrix Σ ausgedrückt. Wir müssen also nur noch den nicht stetigen Teil des Prozesses zu beschreiben.

Doch zuvor führen wir noch einige Schreibweisen ein. Von nun an sei $\overline{\mathbb{R}} := (-\infty, \infty]$ und

$$\operatorname{sgn} x := \begin{cases} 1 & \text{für } x \geq 0 \\ -1 & \text{für } x < 0 \end{cases} . \quad (2.76)$$

Für $a, b \in \mathbb{R}^d$ schreiben wir $a \leq b$, wenn diese Ungleichung komponentenweise erfüllt ist. In diesem Fall beschreibt $(a, b]$ ein halboffenes Intervall von \mathbb{R}^d :

$$(a, b] := (a_1, b_1] \times \dots \times (a_d, b_d] . \quad (2.77)$$

Definition 2.7.1 [F-VOLUMEN]

Sei $F : S \rightarrow \overline{\mathbb{R}}$ für eine Teilmenge $S \subset \overline{\mathbb{R}}^d$. Für $a, b \in S$ mit $a \leq b$ sei das F-Volumen von $(a, b]$ definiert durch

$$V_F((a, b]) := \sum_{u \in \{a_1, b_1\} \times \dots \times \{a_d, b_d\}} (-1)^{N(u)} F(u), \quad (2.78)$$

wobei $N(u) := \#\{k : u_k = a_k\}$.

Bemerkung: 2.7.2: Wählt man $F(u) := \prod_{i=1}^d u_i$, dann ist das F-Volumen gleich dem Lebesgue-Volumen des Intervalls.

Definition 2.7.3 [D-ANSTEIGEND]

Eine Funktion $F : S \rightarrow \overline{\mathbb{R}}$ für eine Teilmenge $S \subset \overline{\mathbb{R}}^d$ wird d -ansteigend genannt, wenn $V((a, b]) \geq 0$ für alle $a, b \in S$ mit $a \leq b$ und $\overline{(a, b]}$ gilt.

2.7.1 Definition von Lévy-Copulas

Definition 2.7.4 [LÉVY-COPULA]

Eine Funktion $F : \overline{\mathbb{R}}^d \rightarrow \overline{\mathbb{R}}$ heißt Lévy-Copula, falls sie folgende Eigenschaften besitzt:

1. $F(u_1, \dots, u_d) \neq \infty$, für $(u_1, \dots, u_d) \neq (\infty, \dots, \infty)$,
2. $F(u_1, \dots, u_d) = 0$, falls $u_i = 0$ für mindestens ein $i \in \{1, \dots, d\}$,
3. F ist d -ansteigend,

4. $F^{\{i\}}(u) = u$ für alle $i \in \{1, \dots, d\}$, $u \in \mathbb{R}$.

Für die vierte Eigenschaft benötigen wir eine weitere Definition.

Definition 2.7.5 [I-MARGIN]

Sei $F : \overline{\mathbb{R}}^d \rightarrow \overline{\mathbb{R}}$ eine d -ansteigende Funktion mit Eigenschaft 2. Für jede nicht leere Indexmenge $I \subset \{1, \dots, d\}$ ist das I -Margin von F eine Funktion $F^I : \overline{\mathbb{R}}^I \rightarrow \overline{\mathbb{R}}$, definiert durch

$$F^I((u_i)_{i \in I}) := \lim_{\xi \rightarrow \infty} \sum_{(u_j)_{j \in I^C} \in \{-\xi, \infty\}^{I^C}} F(u_1, \dots, u_d) \prod_{j \in I^C} \operatorname{sgn} u_j, \quad (2.79)$$

wobei $I^C := \{1, \dots, d\} \setminus I$.

Für den Fall $I = \{1\}$ und $d = 2$ erhalten wir dann $F^{\{1\}} = F(u, \infty) - \lim_{\xi \rightarrow -\infty} F(u, \xi)$. Eine wichtige Eigenschaft von Lévy-Copulas ist ihre Lipschitz-Stetigkeit.

Lemma 2.7.6

Sei F eine Lévy-Copula und $u, v \in \mathbb{R}^d$. Dann gilt

$$|F(v_1, \dots, v_d) - F(u_1, \dots, u_d)| \leq \sum_{i=1}^d |v_i - u_i|. \quad (2.80)$$

Im weiteren Verlauf dieses Abschnitts werden wir noch die folgende Notation benötigen:

$$\mathcal{I}(x) := \begin{cases} (x, \infty), & x \geq 0, \\ (-\infty, x], & x < 0. \end{cases}$$

In der gleichen Weise, wie man die Verteilung eines Zufallsvektors durch seine Verteilungsfunktion charakterisieren kann, ist es möglich, das Lévy-Maß des Lévy-Prozesses durch seine tail-Integrale zu repräsentieren.

Definition 2.7.7 [TAIL-INTEGRAL]

Sei X ein \mathbb{R}^d -wertiger Lévy-Prozess mit Lévy-Maß ν . Das tail-Integral von X ist die Funktion $U : (\mathbb{R}^d \setminus 0) \rightarrow \mathbb{R}$, definiert durch

$$U(x_1, \dots, x_d) := \prod_{i=1}^d \operatorname{sgn}(x_i) \nu \left(\prod_{j=1}^d \mathcal{I}(x_j) \right). \quad (2.81)$$

Da das tail-Integral nur auf $\mathbb{R}^d \setminus 0$ definiert ist, ist das entsprechende Lévy-Maß dadurch nicht eindeutig bestimmt. Wir werden jedoch sehen, dass ein Lévy-Maß eindeutig durch sein tail-Integral und seine margin-Integrale bestimmt ist.

Definition 2.7.8

Sei X ein \mathbb{R}^d -wertiger Lévy-Prozess und sei $I \subset \{1, \dots, d\}$ nicht leer. Das I -marginale tail-Integral U^I von X ist das tail-Integral des Prozesses $X_I := (X^i)_{i \in I}$. Um die Notation zu verkürzen schreiben wir das eindimensionale margin-Integral als $U_i := U^{\{i\}}$.

Der nun folgende Satz spiegelt den Zusammenhang zwischen dem Lévy-Prozess und seinen margin-tail-Integralen wider.

Satz 2.7.9

Sei X ein \mathbb{R}^d -wertiger Lévy-Prozess. Seine marginal-tail-Integrale

$$\{U^I : I \subset \{1, \dots, d\}, I \text{ nicht leer}\}$$

sind eindeutig durch sein Lévy-Maß ν bestimmt. Andererseits ist sein Lévy-Maß eindeutig durch die Menge seiner marginal-tail-Integrale bestimmt.

Beweis. \Rightarrow : Nach Lemma 11.10 aus [Sat99] ist das Lévy-Maß von X^I für jede nicht leere Indexmenge $I \subset \{1, \dots, d\}$ gegeben durch

$$\nu^I(A) = \nu(\{x \in \mathbb{R}^d : (x_i)_{i \in I} \in A \setminus \{0\}\}), \quad A \in \mathcal{B}(\mathbb{R}^{|I|}).$$

Die margin-tail-Integrale sind eindeutig durch die margin-Lévy-Maße bestimmt.

\Leftarrow : Es ist ausreichend zu zeigen, dass $\nu((a, b])$ für alle $a, b \in \mathbb{R}^d$ mit $a \leq b$ und $0 \notin (a, b]$ eindeutig durch die tail-Integrale bestimmt ist. Wir zeigen durch eine Induktion über $k = 1, \dots, k$, dass $\nu^I(\prod_{i \in I}(a_i, b_i])$ eindeutig durch die tail-Integrale bestimmt ist, wobei $a, b \in \mathbb{R}$ mit $a \leq b$ und $a_i b_i \leq 0$ für mindestens k Indizes und jede nicht leere Indexmenge $I \subset \{1, \dots, d\}$ mit $0 \notin \prod_{i \in I}(a_i, b_i]$ gilt.

Im Fall $k = 0$ liefern uns die Definitionen 2.7.7 und 2.7.8

$$\nu^I\left(\prod_{i \in I}(a_i, b_i]\right) = (-1)^{|I|} V_{UI}\left(\prod_{i \in I}(a_i, b_i]\right).$$

Seien $a, b \in \mathbb{R}^d$ nun so gewählt, dass $a_i b_i \leq 0$ für höchstens k Indizes. Um die Notation zu vereinfachen, nehmen wir an, dass $a_i b_i \leq 0$ für $i = 1, \dots, k$. Sei $I \subset \{1, \dots, d\}$ eine nicht leere Indexmenge mit $0 \notin \prod_{i \in I}(a_i, b_i]$. Nach Induktionsannahme ist $\nu^I(\prod_{i \in I}(a_i, b_i])$ eindeutig durch die tail-Integrale bestimmt, falls $k \notin I$. Nehmen wir an $k \in I$. Wenn $a_k = 0$ ist, dann gilt

$$\nu^I\left(\prod_{i \in I}(a_i, b_i]\right) = \lim_{\alpha \searrow 0} \nu^I\left(\prod_{i \in I, i < k}(a_i, b_i] \times (\alpha, b_k] \times \prod_{i \in I, i > k}(a_i, b_i]\right)$$

und die Aussage ergibt sich für die rechte Seite durch die Induktionsannahme. Wenn $a_k \neq 0$ ist, dann gilt

$$\begin{aligned} \nu^I\left(\prod_{i \in I}(a_i, b_i]\right) &= \nu^{I \setminus \{k\}}\left(\prod_{i \in I \setminus \{k\}}(a_i, b_i]\right) \\ &\quad - \lim_{\beta \searrow b_k; c \nearrow \infty} \nu^I\left(\prod_{i \in I, i < k}(a_i, b_i] \times (\beta, c] \times \prod_{i \in I, i > k}(a_i, b_i]\right) \\ &\quad - \lim_{c \searrow -\infty} \nu^I\left(\prod_{i \in I, i < k}(a_i, b_i] \times (c, a_k] \times \prod_{i \in I, i > k}(a_i, b_i]\right), \end{aligned}$$

und die rechte Seite ist ebenfalls eindeutig bestimmt. \square

Der folgende Satz ist der wichtigste Teil dieses Kapitels. Er beschreibt den Zusammenhang zwischen Lévy-Copulas und Lévy-Prozessen.

Satz 2.7.10

Dieser Satz wird auch als Sklars Theorem für Lévy-Copulas bezeichnet. Es gilt:

1. Sei $X = (X_1, \dots, X_d)$ ein \mathbb{R}^d -wertiger Lévy-Prozess. Dann existiert eine Lévy-Copula F , so dass die tail-Integrale von X folgender Bedingung genügen:

$$U^I((x_i)_{i \in I}) = F^I(U^I((x_i)_{i \in I})), \quad (2.82)$$

wobei $I \subset \{1, \dots, s\}$ eine nicht leere Indexmenge ist und alle $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})$. Die Lévy-Copula F ist eindeutig auf $\prod_{i=1}^d \overline{\text{Ran}U_i}$.

2. Sei F eine d -dimensionale Lévy-Copula und $U_i, i = \{1, \dots, d\}$ die tail-Integrale von reellwertigen Lévy-Prozessen. Dann gibt es einen \mathbb{R}^d -wertigen Lévy-Prozess X , dessen Komponenten die tail-Integrale U_1, \dots, U_d haben und dessen margin-tail-Integrale Gleichung (2.82) für eine nicht leere Indexmenge $I \subset \{1, \dots, s\}$ und alle $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})$ erfüllen. Das Lévy-Maß ν von X ist eindeutig durch F und $U_i, i = 1, \dots, d$ bestimmt.

Beweis. Einen vollständigen Beweis der obigen Aussage findet man in [KT06]. □

Definition 2.7.11

Wir nennen die Lévy-Copula aus Aussage 1 in Satz 2.7.10 Lévy-Copula des Lévy-Prozesses X .

Im nächsten Abschnitt werden wir nun einige einfache Beispiele von Lévy-Copulas kennenlernen. Es sollte hier noch erwähnt werden, dass die Verwendung von Lévy-Copulas nicht auf Lévy-Prozesse beschränkt ist. Man kann dieses Prinzip ebenso auf alle Arten von Markov-Prozessen anwenden, die sich lokal wie ein Lévy-Prozess verhalten, was in diesem Fall heißt, dass man seine Dynamik mit Hilfe einer Drift, einer Kovarianzmatrix und einem Lévy-Maß beschreiben kann.

2.7.2 Beispiele von Lévy-Copulas

In diesem Abschnitt werden wir zwei Extremfälle im Bereich der Lévy-Copulas vorstellen, zum einen den Fall der vollständigen Abhängigkeit der mehrdimensionalen Lévy-Prozesse, zum anderen die vollständige Unabhängigkeit dieser Prozesse. Dazu benötigen wir zunächst ein einfaches Kriterium für die Unabhängigkeit von Lévy-Prozessen.

Lemma 2.7.12

Die Komponenten X^1, \dots, X^d eines \mathbb{R}^d -wertigen Lévy-Prozesses sind genau dann unabhängig, wenn ihre Brownschen Teile unabhängig sind und wenn die tail-Integrale des Lévy-Prozesses für alle $I \subset \{1, \dots, d\}$ mit $\#I > 2$ und alle $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$ die Bedingung $U^I(x_i)_{i \in I} = 0$ erfüllen.

Wir können nun die Lévy-Copula zu Lévy-Prozessen mit unabhängigen Komponenten angeben.

Satz 2.7.13

Die Komponenten X^1, \dots, X^d eines \mathbb{R}^d -wertigen Lévy-Prozesses X sind genau dann unabhängig, wenn ihre Brownschen Teile unabhängig sind und X eine Lévy-Copula der Form

$$F_{\perp}(x_1, \dots, x_d) := \sum_{i=1}^d x_i \prod_{j \neq i} 1_{\{\infty\}}(x_j) \tag{2.83}$$

besitzt.

Beweis. Offensichtlich wird durch (2.83) eine Lévy-Copula definiert.

⇐: Sei $I \subset \{1, \dots, d\}$ mit $\#I \geq 2$. Aus Definition 2.7.5 folgt, dass $F^I((u_i)_{i \in I}) = 0$ für alle $(u_i)_{i \in I} \in \mathbb{R}^I$. Daher ist nach (2.82) auch $U^I((x_i)_{i \in I}) = 0$ für alle $(x_i)_{i \in I} \in \mathbb{R}^I \setminus \{0\}$. Zusammen mit Lemma 2.7.12 erhalten wir dann die Aussage.

⇒: Es gilt $U^I((x_i)_{i \in I}) = F^I(U^I((x_i)_{i \in I})) = 0$ für alle $I \subset \{1, \dots, d\}$ mit $\#I > 2$ und alle

$(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$ (nach Lemma 2.7.12). F ist also die Lévy-Copula zum Lévy-Prozess X . \square

Nun folgt noch die Lévy-Copula für den Fall vollständig abhängiger Komponenten. Für diese Copula werden allerdings noch einige Definitionen benötigt. Zunächst sei daran erinnert, dass man eine Teilmenge S von \mathbb{R}^d *geordnet* nennt, wenn für je zwei Vektoren $u, v \in S$ entweder $u_k \leq v_k$ für $k = 1, \dots, d$ oder $u_k \geq v_k$ für $k = 1, \dots, d$ gilt. Eine Teilmenge $S \subset \mathbb{R}^d$ heißt *streng geordnet*, wenn für je zwei Vektoren $u, v \in S$ entweder $u_k < v_k$ für $k = 1, \dots, d$ oder $u_k > v_k$ für $k = 1, \dots, d$ gilt. In der folgenden Definition setzen wir

$$K := \{x \in \mathbb{R}^d : \operatorname{sgn} x_1 = \dots = \operatorname{sgn} x_d\}. \quad (2.84)$$

Definition 2.7.14

Sei X ein \mathbb{R}^d -wertiger Lévy-Prozess. Die Sprunganteile werden als vollständig unabhängig bezeichnet, falls es eine streng geordnete Teilmenge $S \subset K$ gibt, so dass $\Delta X_t := X_t - X_{t-} \in S$, $t \in \mathbb{R}$. Diese Aussage gilt bis auf eine Nullmenge von Pfaden.

In einer streng geordneten Menge ist offensichtlich jedes Element schon durch Angabe einer Koordinate eindeutig bestimmt. Daher sind, wenn die Komponenten eines Lévy-Prozesses vollständig unabhängig voneinander sind, die Sprünge aller Komponenten nur durch Angabe des Sprungs einer einzelnen Komponente vollständig festgelegt. Verfügt der Lévy-Prozess über keinen stetigen Martingalteil, so kann man sogar die kompletten Trajektorien aller Komponenten durch eine einzelne Trajektorie vollständig bestimmen. Definition 2.7.14 ist also eine sinnvolle Definition für die Dynamik vollständig abhängiger Lévy-Prozesse. Die Bedingung $\Delta X_t \in K$ besagt, dass die einzelnen Komponenten *komonoton* sind, d.h. sie springen immer in dieselbe Richtung.

Für jeden \mathbb{R}^d -wertigen Lévy-Prozess X mit Lévy-Maß ν und für jedes $B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ ist die Anzahl an Sprüngen mit Größe $y \in B$ innerhalb des Zeitintervalls $[0, t]$ eine Poissonverteilte Zufallsvariable mit Parameter $t\nu(B)$. Daher kann man die Definition 2.7.14 auf folgende Weise umformulieren:

Definition 2.7.15

Sei X ein \mathbb{R}^d -wertiger Lévy-Prozess mit Lévy-Maß ν . Seine Sprungkomponenten sind vollständig abhängig, falls es eine streng geordnete Teilmenge $S \in K$ gibt, so dass $\nu(\mathbb{R}^d \setminus S) = 0$.

Der folgende Satz charakterisiert nun die vollständige Abhängigkeit der Sprungkomponenten in der Lévy-Copula Formulierung.

Satz 2.7.16

Sei X ein \mathbb{R}^d -wertiger Lévy-Prozess. Der Träger seines Lévy-Maßes ν sei eine geordnete Menge $S \subset K$. Dann ist die complete dependence Lévy-Copula

$$F_{\parallel}(x_1, \dots, x_d) := \min\{|x_1|, \dots, |x_d|\} 1_K(x_1, \dots, x_d) \prod_{i=1}^d \operatorname{sgn} x_i \quad (2.85)$$

die Lévy-Copula von X .

Wenn hingegen F_{\parallel} die Lévy-Copula eines Lévy-Prozesses X ist, dann ist der Träger des Lévy-Maßes von X eine geordnete Teilmenge von K . Wenn darüber hinaus noch die tail-Integrale U_i von X^i stetig sind und der Bedingung $\lim_{x \rightarrow 0} U_i(x) = \infty$ genügen, dann sind die Sprungkomponenten von X vollständig abhängig.

Der Beweis zu diesem Satz gestaltet sich etwas schwieriger als der Beweis für die Lévy-Copula bei unabhängigen Sprungkomponenten. Er kann in [KT06] nachgelesen werden.

2.7.3 Clayton-Lévy-Copula

Nachdem wir nun zwei einfache Beispiele für Lévy-Copulas kennengelernt haben, wird in diesem Abschnitt die Clayton-Lévy-Copula vorgestellt. Dabei handelt es sich um eine Familie von Lévy-Copulas mit zwei Parametern zur Steuerung der Abhängigkeit der Sprungkomponenten. Wir werden uns in diesem Abschnitt mit der zweidimensionalen Clayton-Lévy-Copula befassen, da wir auch später mit dieser eine Basket-Option auf zwei Wertpapiere bewerten wollen.

Neben den beiden im vorherigen Abschnitt vorgestellten Fällen gibt es noch ein allgemeineres Verfahren, um Lévy-Copulas zu erzeugen. Dies geschieht unter Verwendung von streng monoton steigenden, stetigen Funktionen. Die so entstehenden Copulas werden in einem Satz aus [KT06] beschrieben:

Satz 2.7.17

Sei $\phi : [-1, 1] \rightarrow [-\infty, \infty]$ eine streng monoton ansteigende, stetige Funktion, für die $\phi(1) = \infty$, $\phi(0) = 0$ und $\phi(-1) = -\infty$ gilt. Ihre Ableitungen bis zur Ordnung d existieren und erfüllen auf dem Intervall $(-1, 0)$ die Bedingung

$$\frac{\partial \phi(e^x)}{\partial x^d} \geq 0, \quad \frac{\partial \phi(-e^x)}{\partial x^d} \leq 0, \quad x \in (-\infty, 0). \quad (2.86)$$

Sei

$$\tilde{\phi}(u) := 2^{d-2}(\phi(u) - \phi(-u)) \quad u \in [-1, 1]. \quad (2.87)$$

Dann definiert

$$F(u_1, \dots, u_d) := \phi \left(\prod_{i=1}^d \tilde{\phi}^{-1}(u_i) \right) \quad (2.88)$$

eine Lévy-Copula.

Man hat nun die Möglichkeit, unter Verwendung verschiedener Funktionen ϕ unterschiedliche Lévy-Copulas zu erzeugen. Eine dieser Copulas, die unter anderem in [FRS06] verwendet wird, ist die sogenannte *Clayton-Lévy-Copula*. Dazu wählen wir

$$\phi(x) := \eta(-\log|x|)^{-\frac{1}{\theta}} 1_{\{x \geq 0\}} - (1 - \eta)(-\log|x|)^{-\frac{1}{\theta}} 1_{\{x < 0\}} \quad (2.89)$$

mit $\theta > 0$ und $\eta \in (0, 1)$. Dann erhalten wir

$$\begin{aligned} \tilde{\phi}(x) &= 2^{2-d}(-\log|x|)^{-\frac{1}{\theta}} \operatorname{sgn} x, \\ \tilde{\phi}^{-1} &= \exp(-|2^{2-d}u^{-\theta}|) \operatorname{sgn} u. \end{aligned}$$

Mit diesen Funktionen erhalten wir nun die Lévy-Copula

$$F(u_1, \dots, u_d) = 2^{2-d} \left(\sum_{i=1}^d |u_i|^{-\theta} \right)^{-\frac{1}{\theta}} (\eta 1_{\{u_1 \dots u_d \geq 0\}} - (1 - \eta) 1_{\{u_1 \dots u_d < 0\}}). \quad (2.90)$$

Diese definiert eine Zwei-Parameter-Familie von Lévy-Copulas, die man die Clayton-Familie der gewöhnlichen Copulas (Clayton family of ordinary copulas) nennt. F ist für alle $\theta > 0$ und $\eta \in [0, 1]$ eine Lévy-Copula. Die Rolle der Parameter wird im Falle einer zweidimensionalen Clayton-Copula besonders deutlich. Bei $d = 2$ wird aus (2.90) die Copula

$$F(u, v) = (|u|^{-\theta} + |v|^{-\theta})^{-\frac{1}{\theta}} (\eta 1_{\{uv \geq 0\}} - (1 - \eta) 1_{\{uv < 0\}}). \quad (2.91)$$

In dieser Darstellung kann man leicht sehen, dass der Parameter η die Abhängigkeit der Vorzeichen der beiden Sprungkomponenten angibt. Bei $\eta = 0$ haben die Sprünge immer dasselbe Vorzeichen, $\eta = 1$ führt dazu, dass ein positiver Sprung in einer Komponente immer zusammen mit einem negativen Sprung in der anderen Komponente einhergeht. Werte zwischen 0 und 1 führen dazu, dass beide Fälle in unterschiedlichen Ausprägungen auftreten können. Der Parameter θ ist für die Abhängigkeit der absoluten Werte der Sprünge verantwortlich. Genauer gesagt konvergiert F für $\eta = 1$ und $\theta \rightarrow 0$ gegen die unabhängige Lévy-Copula (2.83). Für $\eta = 1$ und $\theta \rightarrow \infty$ konvergiert F gegen die complete dependence Copula (2.85).

Auf diese Art und Weise kann man, wie schon erwähnt, viele verschiedene Familien von Lévy-Copulas erzeugen, je nachdem wie viele Informationen über die Abhängigkeitsstrukturen der einzelnen Sprungkomponenten untereinander vorhanden sind. Wir werden uns dabei auf die Clayton-Copula beschränken. Weitere Familien von Copulas und deren Eigenschaften sind in [Ket06] nachzulesen.

Die Copula (2.91) wird auch als bidirektionale Clayton-Copula bezeichnet, wobei bidirektional den Einfluss von η auf die Copula bezeichnet. Da wir in den numerischen Experimenten, welche im letzten Kapitel beschrieben werden, grundsätzlich $\eta = 0.5$ wählen werden, werden wir von nun, um die Notation zu vereinfachen, die Clayton-Copula

$$F(u, v) = (|u|^{-\theta} + |v|^{-\theta})^{-\frac{1}{\theta}} \quad (2.92)$$

betrachten.

2.8 Die Dichte des zweidimensionalen Lévy-Prozesses

Da wir nun eine Lévy Copula zur Verfügung haben, fehlt noch die Dichtefunktion des zweidimensionalen Lévy-Prozesses, also die genaue Charakterisierung des zweidimensionalen Lévy-Maßes. Nach [FRS06] kann man die Dichte des Lévy-Maßes durch Differentiation bestimmen. Die Dichtefunktion $k(x_1, x_2)$ hat danach die Form

$$\nu(dx_1, dx_2) = \partial_1 \partial_2 F|_{\xi_1=U_1(x_1), \xi_2=U_2(x_2)} \nu_1(dx_1) \nu_2(dx_2), \quad (2.93)$$

wobei ν_1 und ν_2 die Dichten der eindimensionalen Lévy-Prozesse sind. Setzen wir für F die Clayton-Copula ein, dann erhalten wir die Dichte

$$k(x_1, x_2) = (1 + \theta)(|u|^{-\theta} + |v|^{-\theta})^{-\frac{1}{\theta}-2} |v|^{-\theta-1} |u|^{-\theta-1} f_1(x_1) f_2(x_2), \quad (2.94)$$

wobei $u = U_1(x_1)$, $v = U_2(x_2)$ und f_1 bzw. f_2 die eindimensionalen Lévy-Dichten sind. Um die Dichtefunktion zum Beispiel für das Kollokationsverfahren auszuwerten, muss man demnach die tail-Integrale der eindimensionalen Dichten bestimmen. Wenn für die entsprechende Dichte eine Stammfunktion existiert, ist diese Auswertung unproblematisch, verwenden wir allerdings eine kompliziertere Dichte, wie sie zum Beispiel beim CGMY-Prozess vorkommt, so müssen diese Integrale durch Quadraturverfahren bestimmt werden.

Nun stehen uns Modelle zur Bewertung Plain-Vanilla- und Basket-Optionen zur Verfügung, wobei der zugrunde liegende Basiswert einem Sprung-Diffusions-Prozess folgt. Der nächste Schritt ist die numerische Umsetzung, also die Diskretisierung der Probleme. Diese werden wir im nächsten Kapitel erläutern.

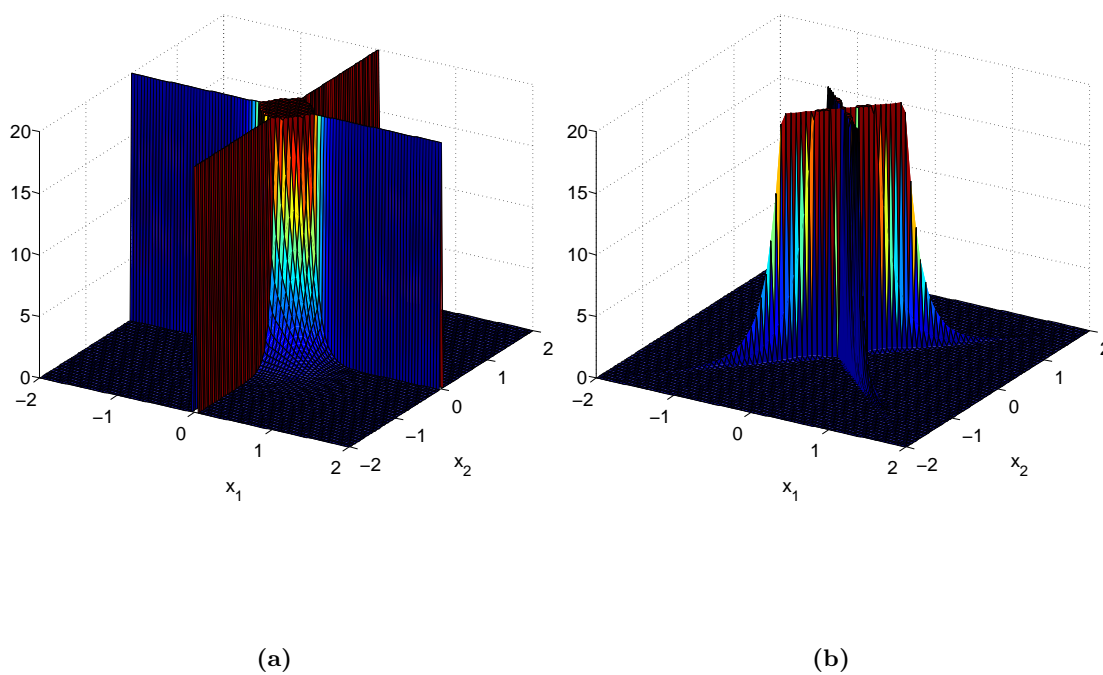


Abbildung 2.4: Dichte der Clayton-Copula mit CGMY-Dichte als margin-tail-Integralen ($C=1, G=M=4, Y=1$) und a) unabhängigen tails ($\theta = 0.5$) b) abhängigen tails ($\theta = 10$)

Kapitel 3

Numerik der PIDE

Zur Bewertung einer Option gibt es grundsätzlich zwei Ansätze. Zum einen gibt es die Möglichkeit, die Option mittels Monte-Carlo-Simulation zu bewerten, zum anderen gibt es den in dieser Arbeit vorgestellten PIDE-Ansatz. Die Monte-Carlo-Simulation liefert im Falle einer Plain-Vanilla-Option eine Konvergenzrate von $\frac{1}{2}$. Solche Simulationsverfahren sind sowohl für Europäische (s. [Gla04]) als auch für Amerikanische Optionen (s. [LS01]) möglich, sofern die Verteilung der Sprünge nicht zu komplex ist. Das heißt in unserem Fall, dass die Darstellung der Aktienkurse im Merton-Modell noch sehr einfach zu realisieren ist, wohingegen die Simulation eines CGMY-Prozesses schon größere Schwierigkeiten bereitet, da dazu die inverse Verteilungsfunktion benötigt wird. Diese inverse Funktion ist allerdings nur für die Normalverteilung bekannt.

Unser Ziel sollte es also sein ein Verfahren zu finden, welches unterschiedlichste Sprungverteilungen berücksichtigt und eine Konvergenzrate $> \frac{1}{2}$ gewährleistet. Dies ist durch eine Diskretisierung der in Kapitel 2 vorgestellten PIDEs und Lösen des entstehenden Gleichungssystems möglich.

In den letzten Jahren wurden verschiedene Arbeiten erstellt, welche sich mit der Lösung und Diskretisierung der PIDEs, wie sie bei der Bewertung von Europäischen und Amerikanischen Optionen entstehen, befassen. Dabei sind unter anderem [Toi08], [Alm05], [IT05], [FV02], [CV05], [DFV05] und [MvPS04] zu nennen. In [CV05] stellen Cont und Voltchkova ein Finite-Differenzen Schema zur Diskretisierung der PIDE vor. Zur numerischen Lösung des Problems führen sie ein Operatorsplitting durch, welches den lokalen und den nicht lokalen Teil voneinander trennt. Durch die dann vorgenommene implizite Behandlung des lokalen und die explizite Behandlung des nicht lokalen Teils wird eine effiziente numerische Lösung des Problems möglich. Die Gleichung wird dann für das Merton-Modell und das Varianz-Gamma Modell gelöst.

Jari Toivanen widmet sich in [Toi08] der Diskretisierung der PIDE unter dem Kou-Modell für Amerikanische und Europäische Optionen. Es wird ein implizites Finite-Differenzen-Verfahren entwickelt, welches den Differentialteil der Gleichung auf einem nicht-uniformen Gitter löst. Um das Verfahren zu stabilisieren, wird ein Rannacher-Zeitschrittverfahren verwendet. Für die Berechnung des Integralteils werden einfache rekursive Formeln hergeleitet. Die Bewertung Amerikanischer Optionen wird sowohl mittels eines Operatorsplitting- als auch eines Penalty-Verfahrens vorgenommen.

Ikonen und Toivanen beschreiben in [IT05] ein Finite-Differenzen Verfahren zur Bewertung Amerikanischer Optionen. In der Arbeit wird ein Runge-Kutta-Verfahren zweiter Ordnung mit dem Crank-Nicolson-Verfahren verglichen und gezeigt, dass ersteres kleinere Fehler und geringere Oszillationen verursacht. Darüber hinaus wird eine Neuformulierung des Algorithmus von Brennan und Schwartz mit der Verwendung eines projektiven SOR-Verfahrens und der Methode von Elliott und Ockendon verglichen.

In [FV02] wird ein implizites Penalty-Verfahren zur Bewertung Amerikanischer Optionen vorgestellt und mit Projektionsmethoden verglichen. Dabei wird das Konvergenzverhalten für den Fall untersucht, dass Zeit- und Ortsschrittweite gegen 0 gehen.

In [Alm05] wird der Brennan-Schwartz-Algorithmus auf Amerikanische Optionen unter CGMY-Prozessen angewendet. Auch hier wird zunächst ein Operatorsplitting in einen Integro- und einen Differentialteil durchgeführt und dann die entstehende Gleichung gelöst. Zur Beschleunigung der Berechnungen wird ein FFT-Verfahren verwendet.

D'Halluin, Forsyth und Vetzal untersuchen in [DFV05] die Stabilität verschiedener expliziter Verfahren. Darüber hinaus wird ein Fixpunkt-Iterationsschema vorgestellt, welches global konvergent ist. Der Integralteil wird hier ebenfalls mittels FFT behandelt.

In [MvPS04] wird eine Wavelet-Galerkin Diskretisierung für die PIDE und ein θ Zeitschrittverfahren vorgestellt und Europäische Optionen mit zugrunde liegenden CGMY- und VG-Prozessen berechnet. In [MNS] wird das Verfahren auf Amerikanische Optionen erweitert. Die Lösung wird hier durch ein Finite-Elemente Verfahren ermittelt, welchem eine Spline-Wavelet-Basis zugrunde liegt. Die entstandene Matrix wird komprimiert und dann in annähernd linearem Aufwand invertiert, ohne die Genauigkeit der Lösung zu beeinflussen.

Wir werden hier hauptsächlich dem Operatorsplitting aus [CV05] folgen und Integralteil bzw. Differentialteil gesondert behandeln. Im eindimensionalen Fall werden die Amerikanischen Optionen mit Hilfe des Brennan-Schwartz-Algorithmus gelöst, im zweidimensionalen Fall mittels eines projektiven SOR-Verfahrens. Europäische Optionen werden grundsätzlich mittels eines SOR-Verfahrens bewertet. Die Grundlagen der \mathcal{H} -Matrix-Approximation, welche zur Bearbeitung des Integrals verwendet wird, werden wir im nächsten Kapitel vorstellen.

Der grundsätzliche Aufbau der Diskretisierung und die verwendeten Bezeichnungen orientieren sich hauptsächlich an [Rei06].

3.1 Transformationen und Lokalisierung

Wir haben nun grundsätzlich zwei PIDEs, die diskretisiert werden müssen. Zum einen Gleichung (2.36) aus dem Merton-Modell und Gleichung (2.63) aus dem VG- bzw. CGMY-Modell. Betrachten wir zunächst die Gleichung

$$V_\tau = \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - (r + \lambda)V + \lambda \int_{\mathbb{R}^+} V(SY, \tau)g(Y)dY \quad (3.1)$$

aus dem Merton Modell. Wir überführen sie zunächst in einen logarithmischen Preis, d.h wir substituieren $x = \log(S)$ und erhalten

$$V_\tau = \frac{1}{2}\sigma^2 V_{xx} + (r - \lambda\kappa - \frac{\sigma^2}{2})V_x - (r + \lambda)V + \lambda \int_{\mathbb{R}^+} V(e^x Y, \tau)g(Y)dY. \quad (3.2)$$

Darüberhinaus führen wir die Substitution $y = \log(Y)$ durch, d.h. $Y = e^y$ und $dY = e^y dy$. Damit ergibt sich

$$V_\tau = \frac{1}{2}\sigma^2 V_{xx} + (r - \lambda\kappa - \frac{\sigma^2}{2})V_x - (r + \lambda)V + \lambda \int_{\mathbb{R}^+} V(x + y, \tau)f(y)dy, \quad (3.3)$$

wobei $f(y) = g(e^y)e^y$ ist. Die ursprüngliche Lognormalverteilung der Sprünge ist damit in eine Normalverteilung übergegangen. Denn aus der Dichte

$$g(Y) = \frac{1}{\sqrt{2\pi}\sigma_J Y} \exp\left(-\frac{(\ln Y - \mu_J)^2}{2\sigma_J^2}\right) \quad (3.4)$$

wurde durch die Transformation die neue Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_J} \exp\left(-\frac{(x - \mu_J)^2}{2\sigma_J^2}\right). \quad (3.5)$$

Dabei bezeichnen σ_J und μ_J Varianz bzw. Erwartungswert der Sprungdichte. Der Parameter $\kappa = E[Y - 1]$ ist nun gegeben durch $\kappa = \exp(\mu_J - \sigma_J^2/2) - 1$.

Nun wollen wir versuchen, die PIDE aus dem VG- bzw. CGMY-Modell auf eine ähnliche Form zu bringen. Wenn dies geschehen ist, müssen beide bei der Diskretisierung nicht mehr getrennt behandelt werden. Die ursprüngliche Gleichung, in logarithmischem Preis und transformierter Zeit, hat die Form

$$\begin{aligned} 0 &= \frac{\sigma^2}{2} V_{xx} + \left(r - \frac{\sigma^2}{2}\right) V_x - rV_\tau + \int_{\mathbb{R}^+} (V(\tau, x+y) - V(\tau, x) - (e^y - 1)V_x(\tau, x)) \nu(dy) \\ &= \frac{\sigma^2}{2} V_{xx} + \left(r - \frac{\sigma^2}{2}\right) V_x - rV_\tau \\ &\quad - V_x \int_{\mathbb{R}^+} (e^y - 1) \nu(dy) - V \int_{\mathbb{R}^+} \nu(dy) + \int_{\mathbb{R}^+} (V(\tau, x+y)) \nu(dy). \end{aligned}$$

Nun ersetzen wir $\kappa = E_\nu[e^y - 1] = \int_{\mathbb{R}^+} (e^y - 1) \nu(dy)$ und $\lambda_\nu = \int_{\mathbb{R}^+} \nu(dy)$ und erhalten die PIDE

$$\begin{aligned} 0 &= \frac{\sigma^2}{2} V_{xx} + \left(r - \frac{\sigma^2}{2}\right) V_x - rV_\tau - \kappa V_x - \lambda_\nu V + \int_{\mathbb{R}^+} \nu(dy) (V(\tau, x+y)) \nu(dy). \\ &= \frac{1}{2} \sigma^2 V_{xx} + \left(r - \kappa - \frac{\sigma^2}{2}\right) V_x - (r + \lambda_\nu) V + \int_{\mathbb{R}^+} V(x+y, \tau) \nu(y) dy. \end{aligned} \quad (3.6)$$

Vergleicht man diese nun mit (3.3), stellt man leicht fest, dass sie die gleiche Struktur haben und daher grundsätzlich gleich diskretisiert werden. Natürlich muss bei der Diskretisierung des Integrals auf die unterschiedlichen Dichten Rücksicht genommen werden. Wir werden im weiteren Verlauf nur die Form (3.3) behandeln. Es ist also das Problem

$$V_\tau = \frac{1}{2} \sigma^2 V_{xx} + \left(r - \lambda\kappa - \frac{\sigma^2}{2}\right) V_x - (r + \lambda) V + \lambda \int_{\mathbb{R}^+} V(x+y, \tau) f(y) dy, \quad (3.7)$$

mit den Randbedingungen

$$V(x, \tau) = 0 \quad \text{für } x \rightarrow -\infty \quad (3.8)$$

$$V(x, \tau) = (Ke^{-r\tau} - e^x)^+ \quad \text{für } x \rightarrow +\infty \quad (3.9)$$

und der Anfangsbedingung

$$V(x, 0) = (K - e^x)^+ \quad (3.10)$$

zu lösen. Um diese Problem numerisch lösen zu können, muss das unendliche Gebiet $\Omega = \mathbb{R}$ auf ein endliches Gebiet $\Omega_e = [x_{min}, x_{max}]$ beschränkt werden. Da das Zeitintervall $I_T = [0, T]$ schon endlich ist, bleibt es unverändert. Daher lösen wir das Problem (3.7)-(3.10) nicht auf $\Omega \times I_T$, sondern auf $\Omega_e \times I_T$. Wir erhalten

$$V_\tau = \frac{1}{2} \sigma^2 V_{xx} + \left(r - \lambda\kappa - \frac{\sigma^2}{2}\right) V_x - (r + \lambda) V + \lambda \int_{\mathbb{R}^+} V(x+y, \tau) f(y) dy \quad \text{in } \Omega_e \times I_T \quad (3.11)$$

mit den Randbedingungen

$$V(x, \tau) = 0 \quad \text{für } x = x_{min} \quad (3.12)$$

$$V(x, \tau) = (Ke^{-r\tau} - e^{x_{max}})^+ \quad \text{für } x = x_{max}. \quad (3.13)$$

Die Anfangsbedingung (3.10) bleibt erhalten.

3.2 Zeitdiskretisierung

Um zu einer semidiskreten Schreibweise des Problems zu gelangen, definieren wir uns zunächst den Differentialoperator

$$\mathcal{L}(V) := \frac{1}{2}\sigma^2 V_{xx} + (r - \lambda\kappa - \frac{\sigma^2}{2})V_x - (r + \lambda)V + \lambda \int_{\mathbb{R}^+} V(x + y, \tau) f(y) dy. \quad (3.14)$$

Damit lässt sich unser Problem schreiben als

$$V_\tau - \mathcal{L}(V) = 0. \quad (3.15)$$

Der Arbeitsweise aus [CV05] folgend wählen wir nun eine Zerlegung dieses Operators in einen Differentialteil und einen Integralteil, wodurch wir die Operatoren

$$\mathcal{L}_{Diff} := \frac{1}{2}\sigma^2 V_{xx} + (r - \lambda\kappa - \frac{\sigma^2}{2})V_x - (r + \lambda)V \quad (3.16)$$

und

$$\mathcal{L}_{Int} := \lambda \int_{\mathbb{R}^+} V(x + y, \tau) f(y) dy \quad (3.17)$$

erhalten.

Es gilt dann $\mathcal{L} = \mathcal{L}_{Diff} + \mathcal{L}_{Int}$. Wir wählen nun eine äquidistante Diskretisierung des Zeitintervalls I_T wie folgt: Sei $\Delta\tau_n := n\Delta\tau$ für $n = 0, \dots, N$ eine äquidistante Einteilung mit Schrittweite $\Delta\tau = T/N$. Wir bezeichnen mit V^n die numerische Approximation des n -ten Zeitschritts, d.h. $V^n(x) = V(x, n\Delta\tau)$. Die Finite-Differenzen-Diskretisierung führt dann zu folgendem θ -Zeitschema

$$\frac{V^{n+1} - V^n}{\Delta\tau} - \theta \mathcal{L}_{Diff}(V^{n+1}) - (1-\theta) \mathcal{L}_{Diff}(V^n) - \theta \mathcal{L}_{Int}(V^{n+1}) - (1-\theta) \mathcal{L}_{Int}(V^n) = 0. \quad (3.18)$$

Von der Wahl von θ hängt nun die Form des Zeitschrittverfahrens ab. Für $\theta = 0$ erhält man das *explizite Eulerverfahren*, für $\theta = 1$ das *implizite Eulerverfahren* und für $\theta = \frac{1}{2}$ das *Crank-Nicolson-Verfahren*. Bei einer Aufteilung des Operators, so wie wir sie durchgeführt haben, ist es möglich, für den Differential- und den Integralteil unterschiedliche Schemata zu wählen. Cont und Voltchkova haben in [CV05] gezeigt, dass eine explizite Behandlung des Integrals keinen Einfluss auf die Konvergenz des Gesamtverfahrens hat. Der Vorteil dieser Methode ist, dass die Matrix des zu lösenden Gleichungssystems nur dünn besetzt ist und das System daher einfacher gelöst werden kann, da man in jedem Zeitschritt statt einer PIDE nur eine PDE lösen muss. Wir werden daher ebenfalls eine explizite Behandlung des Integralteils vornehmen. Um eine höhere Konvergenzrate des Gesamtverfahrens zu erzielen werden wir ein Crank-Nicolson-Verfahren zur Zeitdiskretisierung verwenden. Dabei ist allerdings zu beachten, dass das Verfahren nicht uneingeschränkt stabil ist. Es gilt:

Satz 3.2.1

Das Crank-Nicolson-Verfahren ist hinreichend stabil unter der Beschränkung, dass $\lambda\Delta\tau$ hinreichend klein ist.

Der Beweis hierzu findet man in [DFV05]. Wir erhalten daher eine Bedingung für die Größe der Zeitschritte, allerdings auch eine höhere Konvergenzordnung.

Rannacher Zeitschrittverfahren

Um Oszillationen zu verhindern, welche insbesondere in der Nähe des Ausübungspreises auftreten können, werden wir kein reines Crank-Nicolson verwenden, sondern ein Mischverfahren, und zwar das *Rannacher-Zeitschritt-Verfahren*. Dieses Verfahren wird ebenfalls in [Toi08] verwendet, um Oszillationen in der Lösung zu vermeiden. Mit dem Rannacher-Verfahren ergibt sich die Matrixformulierung

$$\begin{aligned} [\mathbf{I} - \theta_n \mathbf{A}]V^{n+1} &= [\mathbf{I} + (1 - \theta_n)\mathbf{A}]V^n + \lambda\Delta\tau\mathbf{B}V^n, \\ \theta_n &= \begin{cases} 1, n = 0, \dots, n_{Rann} \\ \frac{1}{2}, n = n_{Rann}, \dots, N \end{cases}. \end{aligned}$$

Es genügen hier ein paar Schritte mit dem impliziten Eulerverfahren, um die Anfangsbedingungen hinreichend zu glätten und Oszillationen zu vermeiden. In den numerischen Experimenten im letzten Kapitel werden wir 10 implizite Schritte ausführen und danach das Crank-Nicolson-Verfahren verwenden.

Da das Integral nur explizit behandelt wird, erscheint der Integralteil nur auf der rechten Seite der Gleichung. Dies führt zu einer Tridiagonalmatrix \mathbf{A} und einem Gleichungssystem, welches mit linearem Aufwand gelöst werden kann. Dazu wird bei Europäischen Optionen ein SOR-Verfahren verwendet und bei Amerikanischen der Brennan-Schwartz-Algorithmus mit LR-Zerlegung aus [IT05]. Im folgenden Abschnitt werden wir die Ortsdiskretisierung mittels Finiter-Differenzen vorstellen.

3.3 Ortsdiskretisierung

Wir wählen eine äquidistante Unterteilung des Gebiets Ω_e . Sei dazu $x_i = x_{min} + i\Delta x$ für $i = 0, \dots, M$ eine solche Unterteilung mit Maschenweite $\Delta x = \frac{x_{max} - x_{min}}{M}$. Darüber hinaus sei V_i^n die numerische Approximation von $V(i\Delta x, n\Delta\tau)$. Für die Approximation der ersten Ableitung gibt folgende Varianten:

$$V_x(i\Delta x, n\Delta\tau) \approx \frac{V_{i+1}^n - V_i^n}{\Delta x} \quad (\text{Vorwärts-Differenzenquotient}), \quad (3.19)$$

$$V_x(i\Delta x, n\Delta\tau) \approx \frac{V_i^n - V_{i-1}^n}{\Delta x} \quad (\text{Rückwärts-Differenzenquotient}), \quad (3.20)$$

$$V_x(i\Delta x, n\Delta\tau) \approx \frac{V_{i+1}^n - V_{i-1}^n}{2\Delta x} \quad (\text{zentraler Differenzenquotient}). \quad (3.21)$$

Die Approximation der zweiten Ableitung erfolgt durch den zentralen Differenzenquotienten

$$V_{xx}(i\Delta x, n\Delta\tau) \approx \frac{V_{i-1}^n - 2V_i^n + V_{i+1}^n}{\Delta x^2}. \quad (3.22)$$

3.3.1 Stabilität der Ortsdiskretisierung

Das vorliegende Problem enthält sowohl erste als auch zweite Ableitungsterme und ist daher ein Konvektions-Diffusions-Problem. Die Schwierigkeit dabei ist, dass konvektive Terme, also die ersten Ableitungen, zu Differenzenquotienten führen, die nicht uneingeschränkt stabil sind oder eine geringere Konsistenzordnung als die anderen Quotienten aufweisen. Im einfachen Fall der Black-Scholes-Gleichung (2.3) tritt dieses Problem nicht auf, da eine Transformation existiert, mit deren Hilfe sich das Problem in die Wärmeleitungsgleichung überführen lässt. Diese führt dazu, dass die konvektiven Terme wegfallen und nur noch eine zweite Ableitung im Ort vorkommt. Für die hier vorliegende PIDE ist

eine solche Transformation leider nicht bekannt und man kann das Problem nicht in eine reine Diffusionsgleichung transformieren. Daher ist es nötig, auch die ersten Ableitungen geeignet zu diskretisieren.

Die einseitigen Differenzenquotienten (3.19) und (3.20) besitzen die Konsistenzordnung 1, wohingegen der zentrale Differenzenquotient (3.21) die Konsistenzordnung 2 besitzt ([HB02]). Es scheint also einleuchtend, den zentralen Differenzenquotienten zu verwenden, um eine höhere Konsistenzordnung zu erzielen. Dabei kann es jedoch zu Instabilitäten kommen, da der zentrale Differenzenquotient nicht uneingeschränkt stabil ist.

Definition 3.3.1 [STABILITÄT]

Ein Differenzenverfahren heißt stabil, wenn positive Konstanten C und h^* existieren, so dass für alle $0 < h < h^*$ die Matrix der Diskretisierung invertierbar ist mit $\|L_H^{-1}\|_\infty \leq C$.

Die Stabilität bzw. der Nachweis der Stabilität ist mit der Monotonieeigenschaft der Matrix verbunden (s. [HB02] Abschnitt 83 und 84). Für unser Problem führt die Frage nach der Stabilität des zentralen Differenzenquotienten zu folgender Bedingung:

$$\begin{aligned} \frac{\sigma^2}{2} &> r - \lambda\kappa - \frac{\sigma^2}{2} \\ \Leftrightarrow \sigma^2 &> r - \lambda\kappa. \end{aligned}$$

Sofern diese erfüllt ist, ist das Verfahren stabil. Es bleibt zu beachten, dass diese Beschränkung nur im Fall der zentralen Differenzenquotienten besteht. Bei der Verwendung der einseitigen Quotienten ist es möglich, in Abhängigkeit des Vorzeichens des Konvektionsterms entweder die Vorwärts- oder die Rückwärtsdifferenzen zu verwenden. Somit kann in diesem Fall Stabilität immer gewährleistet werden. Eine Möglichkeit hierfür ist die sogenannte *Upwind Diskretisierung*. Die erste Ableitung würde dann folgendermaßen diskretisiert

$$\begin{aligned} V_x(i\Delta x, n\Delta\tau) &\approx \frac{V_{i+1}^n - V_i^n}{\Delta x} && \text{für } r - \lambda\kappa - \frac{\sigma^2}{2} < 0 \\ V_x(i\Delta x, n\Delta\tau) &\approx \frac{V_i^n - V_{i-1}^n}{\Delta x} && \text{für } r - \lambda\kappa - \frac{\sigma^2}{2} > 0. \end{aligned}$$

Den umgekehrten Fall bezeichnet man als *Downwind Diskretisierung*. Da in unserem Fall die Koeffizienten der PIDE konstant sind, ist es möglich, für festgelegte Parameter eine festgelegte Diskretisierung zu wählen. Es gilt hier also, die uneingeschränkte Stabilität der Upwind oder Downwind Diskretisierung gegen die höhere Konsistenzordnung des zentralen Differenzenquotienten abzuwägen. Wir werden jedoch feststellen, dass für die in den numerischen Experimenten gewählten Parameter der zentrale Quotient stabil ist. Daher werden wir von nun an den zentralen Differenzenquotienten zur Approximation der ersten Ableitung wählen, um die höhere Konsistenzordnung ausnutzen zu können. Eine Untersuchung der Auswirkungen der Upwind- bzw. Downwind- Diskretisierung auf die Stabilität und Konvergenz des Finite-Differenzen-Verfahrens findet sich in [Rei06].

3.4 Diskretisierung des Integraloperators

Eines der Probleme bei der Lösung der vorgestellten PIDE ist die Auswertung des Integraloperators \mathcal{L}_{Int} . Bei den verwendeten Dichten in den verschiedenen Modellen (Lognormal, Doppel-log-exponential, CGMY) ist dieser Operator nicht explizit gegeben und muss daher diskretisiert werden. Für die Diskretisierung des Integrals gibt es unterschiedliche Möglichkeiten. Um später die \mathcal{H} -Matrix Approximation gut anwenden zu können, werden wir ein

Kollokationsverfahren wählen, denn dieses Verfahren ermöglicht eine direkte Auswertung des Integrals, ohne diese vorher lösen zu müssen, und ist daher für unsere explizite Behandlung des Integrals gut geeignet.

3.4.1 Kollokationsverfahren

Das Kollokationsverfahren beruht auf der Grundidee, dass man einen Operator

$$A\phi = h \tag{3.23}$$

durch Bestimmung einer Näherungslösung in einem unendlich dimensionalen Unterraum löst, die an endlich vielen Punkten, den Kollokationspunkten, mit der eigentlichen Lösung übereinstimmt.

Definition 3.4.1 [KOLLOKATIONSVERFAHREN]

Seien X und Y Banachräume mit $Y \subset C[a, b]$ und $A : X \rightarrow Y$ ein beschränkter linearer Operator. Für $n \in \mathbb{N}$ seien $X_n \subset X$ und $Y_n \subset Y$ weitere Unterräume mit $\dim X_n = \dim Y_n = n$ sowie n Kollokationspunkte $x_1, \dots, x_n \in [a, b]$ gegeben, so dass Y_n bezüglich dieser Punkte unisolvent ist, d.h. die Interpolationsaufgabe mit dem Unterraum Y_n und den Kollokationspunkten ist eindeutig lösbar. Dann approximiert das Kollokationsverfahren die Lösung von (3.23) durch ein Element $u_n \in X_n$ mit

$$(Au_n)(x_j) = f(x_j) \quad \text{für} \quad j = 1 \dots n \tag{3.24}$$

Sei $X_n = \text{span}\{\phi_1 \dots \phi_n\}$. Dann lassen sich die u_n als Linearkombination

$$u_n = \sum_{k=1}^n \gamma_k \phi_k \tag{3.25}$$

darstellen. Daher ist (3.23) äquivalent zum linearen Gleichungssystem

$$\sum_{k=1}^n \gamma_k (A\phi_k)(x_j) = f(x_j) \quad \text{für} \quad j = 1 \dots n$$

mit den Koeffizienten $\gamma_1, \dots, \gamma_n$.

Wir betrachten nun das Integral aus (3.3)

$$\mathcal{I}(x) = \int_{-\infty}^{\infty} V(x+y)f(y)dy. \tag{3.26}$$

Sei $\mathcal{I}_i = \mathcal{I}(x_i)$ das Integral am Gitterpunkt x_i und $V_{i+j} = V(x_{i+j}) = V(x_{min} + (i+j)\Delta x)$ der Optionspreis an der Stelle x_{i+j} . Dann können wir das Integral in diskretisierter Form darstellen als

$$\mathcal{I}_i = \sum_{j=-\frac{M}{2}}^{\frac{M}{2}} V_{i+j} f_j \tag{3.27}$$

wobei

$$f_j = \int_{x_j - \frac{\Delta x}{2}}^{x_j + \frac{\Delta x}{2}} f(y)dy. \tag{3.28}$$

Grundsätzlich müssen bei der Berechnung der V_{i+j} zwei Fälle unterschieden werden. Zum einen der Fall $i+j = l \leq M$, d.h. $x_{i+j} \in \Omega_e$. In diesem Fall ist keine besondere Behandlung notwendig, da man direkt $V_l = V_{i+j}$ verwenden kann. Der zweite Fall tritt ein, falls $x_{i+j} \notin \Omega_e$, d.h. $i+j < 0$ oder $i+j > M$. In diesem Fall wird V_{i+j} durch die entsprechende Randbedingung approximiert. Bei der Bewertung einer Europäischen Put-Option heißt das

$$\begin{aligned} V_{i+j} &= 0 && \text{für } i+j > M, \\ V_{i+j} &= (e^{-r\tau}K - e^{x_{i+j}})^+ && \text{für } i+j < 0. \end{aligned}$$

Die Art der Approximation ist sinnvoll, da die f_j für betragsmäßig große j schnell abfallen. Dieses Vorgehen wird unter anderem in [dFL04] vorgeschlagen und verwendet.

Approximation der f_j für Verteilungsfunktionen mit endlicher Variation

Bei der Berechnung der f_j müssen ebenfalls zwei Fälle unterschieden werden. Bei den Lévy-Dichten, wie sie bei der Bewertung von Optionen im VG- oder CGMY-Modell entstehen, muss die Singularität im Nullpunkt berücksichtigt werden. Wir betrachten zunächst den einfachen Fall der nicht-singulären Dichten. Existieren für die verwendete Dichte keine geschlossenen Formeln für die Integration, so muss das Integral (3.28) mit Hilfe einer Quadraturformel approximiert werden. Wir werden im numerischen Teil dieser Arbeit dazu die Simpsonregel verwenden.

Simpsonregel

$$\int_a^b f(x)dx = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) \quad (3.29)$$

Der Fehler der bei dieser Approximation auftritt, beträgt

$$\int_a^b f(x)dx - \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) = (\Delta x)^5 \frac{1}{90} f^{(4)}(\zeta). \quad (3.30)$$

In Kapitel 5 sehen wir, dass die Approximation durch das Simpson-Verfahren ausreichend ist, um Konvergenz zu gewährleisten.

Approximation der f_j für Verteilungsfunktionen mit unendlicher Variation

Wenn für die dem Modell zugrunde liegende Verteilungsfunktion $\nu(\mathbb{R}) = \infty$ gilt, so kann man das Kollokationsverfahren nicht direkt anwenden. Eine Möglichkeit, dieses Problem zu umgehen wird in [CV05] beschrieben. Cont und Voltchkova schlagen darin vor, die Verteilungsfunktion durch eine Funktion mit endlicher Variation zu approximieren. Dazu wird ein kleines Gebiet um die Singularität herausgeschnitten. Die infinitesimalen Sprünge, die durch das Herausschneiden unberücksichtigt bleiben, werden durch eine Verstärkung der Brownschen Bewegung ersetzt. Das Integral hat dann die Form

$$\mathcal{I} = \int_{\mathbb{R} \setminus [-\epsilon, \epsilon]} V(x+y) f(y) dy. \quad (3.31)$$

Die Sprünge, deren Größe kleiner als ϵ ist, müssen dann durch einen Parameter $\sigma(\epsilon)$ ausgeglichen werden, der die Brownsche Bewegung leicht verstärkt. Im Gegensatz zu den Verteilungsfunktionen mit endlicher Variation werden wir hier nicht die Simpsonregel zur Kollokation verwenden. Da eine Verwendung der Simpson-Regel die Singularität der Sprungdichte

nicht hinreichend auflösen kann werden wir die Quadraturroutinen aus der gsl_integriation Bibliothek verwenden.

3.5 Lösung der Gleichungssysteme

Mit den vorgestellten Diskretisierungsmethoden ist es nun möglich, eine vollständige Diskretisierung der PIDE (3.7) anzugeben. Diese hat mit dem vorgestellten θ -Schema und der Verwendung des zentralen Differenzenquotienten die Form

$$\begin{aligned}
 & [1 + \theta(\alpha + \beta + r + \lambda)\Delta\tau]V_i^{n+1} - \theta\Delta\tau\beta V_{i+1}^{n+1} - \theta\Delta\tau\alpha V_{i-1}^{n+1} \\
 = & [1 - (1 - \theta)(\alpha + \beta + r + \lambda)\Delta\tau]V_i^n + (1 - \theta)\Delta\tau\beta V_{i+1}^n + (1 - \theta)\Delta\tau\alpha V_{i-1}^n \\
 & + \lambda\Delta\tau \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} V_n(x_i + y_j)f_j
 \end{aligned} \tag{3.32}$$

mit

$$\begin{aligned}
 \alpha &= \frac{\sigma^2}{2\Delta x^2} - \frac{r - \lambda\kappa - \frac{\sigma^2}{2}}{2\Delta x}, \\
 \beta &= \frac{\sigma^2}{2\Delta x^2} + \frac{r - \lambda\kappa - \frac{\sigma^2}{2}}{2\Delta x}.
 \end{aligned}$$

Um eine etwas übersichtlichere Darstellung der Diskretisierung zu erhalten führen wir nun eine Matrixschreibweise von (3.32) ein, die zu einer kompakteren Darstellung führt. Dazu definieren wir Matrizen \mathbf{A} und \mathbf{B} , so dass gilt

$$[\mathbf{A}V^n]_i = \Delta\tau\alpha V_{i-1}^n - (\alpha + \beta + r + \lambda)\Delta\tau V_i^n + \Delta\tau\beta V_{i+1}^n, \tag{3.33}$$

$$[\mathbf{B}V^n]_i = \sum_j b_{ij}V_j^n = \sum_{j=-\frac{M}{2}}^{\frac{M}{2}} V^n(x_i + y_j)f_j. \tag{3.34}$$

Damit erhalten wir dann die Matrixformulierung

$$[\mathbf{I} - \theta\mathbf{A}]V^{n+1} = [\mathbf{I} + (1 - \theta)\mathbf{A}]V^n + \lambda\Delta\tau\mathbf{B}V^n. \tag{3.35}$$

3.5.1 Das SOR-Verfahren

Wir gehen davon aus, dass die Matrix \mathbf{A} die folgende Form hat

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & a_{23} & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1n-2} & a_{n-1n-1} & a_{n-1n} \\ & & & a_{nn-1} & a_{nn} \end{pmatrix}. \tag{3.36}$$

Die rechte Seite des Gleichungssystems $\mathbf{A}x = b$ sei der Vektor $b = (b_1 \dots b_n)^\top$. Der folgende Algorithmus löst das Gleichungssystem mit dem SOR-Verfahren. Hierbei bezeichnet

der obere Index j den aktuellen Zeitpunkt, der untere Index i den Ort, der obere Index it den Iterationsschritt im SOR-Verfahren und $\|\cdot\|_2$ die diskrete l_2 -Norm. Da es sich bei der Matrix um eine Tridiagonalmatrix handelt, müssen beim SOR-Verfahren nur die Einträge auf den drei Diagonalen berücksichtigt werden. Daher ist in diesem Fall ein Lösen des Systems in linearem Aufwand auch unter Verwendung des SOR-Verfahrens möglich.

```

setze  $z^{(0)} = y^{(j)}$ 
wiederhole
  für  $i = 1, 2 \dots M - 1$ 
    
$$r_i = b_i - \sum_{k=1}^{i-1} a_{ik} z_k^{(it+1)} - \sum_{k=i}^{M-1} a_{ik} z_k^{(it)}$$

    setze  $it = it + 1$ 
  solange  $\|z^{(it)} - z^{(it-1)}\|_2 > \varepsilon$ 
  setze  $y^{(j+1)} = z^{(it)}$ 
    
```

Algorithmus 1: Das SOR Verfahren.

3.5.2 Der Brennan-Schwartz-Algorithmus

Eine direkte Möglichkeit das entstehende Gleichungssystem für Amerikanische Optionen zu lösen, ist der von Brennan und Schwartz in [BS97] zur Bewertung Amerikanischer Optionen vorgestellte und nach ihnen benannte Algorithmus. Dieser ist effizienter als das projektive SOR-Verfahren, kann aber bei komplexeren Modellen und Optionen nur noch eingeschränkt verwendet werden. Beim Brennan-Schwartz-Algorithmus wird die Iterationsmatrix des Zeitschrittverfahrens A zunächst mittels einer LR-Zerlegung in eine obere und eine untere Dreiecksmatrix zerlegt. Das Gleichungssystem $Ay = b$ kann dann mit Hilfe von Vorwärts- und Rückwärtssubstitution gelöst werden. Da die Bandstruktur der Matrix bei der LR-Zerlegung erhalten bleibt, sind diese Schritte in linearer Zeit durchführbar. Um die Ungleichung zu erfüllen, wird der Lösungsvektor innerhalb der Rückwärtssubstitution komponentenweise auf die Hindernisfunktion projiziert.

Wie beim projektiven SOR steigt die Zeit, die dieses Verfahren benötigt, nur linear in der Anzahl der Variablen, allerdings müssen nicht mehrere Iterationen durchgeführt werden, um zur Lösung zu gelangen. Darüber hinaus genügt es, die LR-Zerlegung nur einmal durchzuführen, da die Matrix A bei äquidistantem Gitter nicht über die Zeit variiert, d.h. in einem Zeitschritt muss nur jeweils eine Vorwärts- und eine Rückwärtssubstitution und die Projektion durchgeführt werden.

Die folgende Formulierung des Brennan-Schwartz-Algorithmus verwendet eine LR-Zerlegung, bei der auf der Diagonalen der oberen Dreiecksmatrix Einsen stehen. Hierbei bezeichnet a_{ij} die Elemente der Iterationsmatrix, l_{ij} die Elemente der unteren und u_{ij} die Elemente der oberen Dreiecksmatrix.

<p>LR-Zerlegung:</p> $l_{1,1} = a_{1,1}$ $u_{1,2} = a_{1,2}/l_{1,1}$ <p style="padding-left: 20px;">wiederhole für $i = 2, 3 \dots M - 1$</p> $l_{i,i-1} = a_{i,i-1}$ $l_{i,i} = a_{i,i} - l_{i,i-1}u_{i-1,i}$ $u_{i,i+1} = a_{i,i+1}/l_{i,i}$ $l_{M,M-1} = a_{M,M-1}$ $l_{M,M} = a_{M,M} - l_{M,M-1}u_{M-1,M}$ <p>Vorwärtssubstitution:</p> $z_1 = b_1/l_{1,1}$ <p style="padding-left: 20px;">wiederhole für $i = 2, 3 \dots M$</p> $z_i = (b_i - l_{i,i-1}z_{i-1})/l_{i,i}$ <p>Rückwärtssubstitution:</p> $y_M = z_M$ $y_M = \max(y_M, g_M)$ <p style="padding-left: 20px;">wiederhole für $i = M, M - 1 \dots 1$</p> $y_i = z_i - u_{i,i+1}y_{i+1}$ $z_i = \max(y_i, g_i)$

Algorithmus 2: Der Brennan-Schwartz Algorithmus

3.6 Diskretisierung der mehrdimensionalen PIDE

Im vorigen Kapitel haben wir die mehrdimensionale PIDE

$$\begin{aligned}
 V_\tau = & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i \sigma_j V_{x_i x_j} + \sum_{i=1}^n (r - \lambda \kappa - \frac{\sigma_i^2}{2}) V_{x_i} - (r + \lambda) V \\
 & + \lambda \int_{\mathbb{R}^n} V(\mathbf{x} + \mathbf{y}, \tau) f(\mathbf{y}) d\mathbf{y}.
 \end{aligned}
 \tag{3.37}$$

hergeleitet. Für eine Put-Option wählen wir die Anfangsbedingungen

$$V(\mathbf{x}, 0) = (K - \max\{e^{x_1}, \dots, e^{x_d}\})^+.
 \tag{3.38}$$

Da wir uns in Kapitel 5 nur mit zweidimensionalen Basket-Optionen befassen werden, werden wir hier die Diskretisierung auch direkt auf die zweidimensionale PIDE anwenden. Wir erhalten dann die Gleichung

$$\begin{aligned}
 V_\tau = & \frac{1}{2} \rho \sigma_1 \sigma_2 V_{x_1 x_2} + (r - \lambda \kappa - \frac{\sigma_1^2}{2}) V_{x_1} + (r - \lambda \kappa - \frac{\sigma_2^2}{2}) V_{x_2} - (r + \lambda) V \\
 & + \lambda \int_{\mathbb{R}^2} V(\mathbf{x} + \mathbf{y}, \tau) f(\mathbf{y}) d\mathbf{y}.
 \end{aligned}
 \tag{3.39}$$

mit den Anfangsbedingungen

$$V(\mathbf{x}, 0) = (K - \max\{e^{x_1}, e^{x_2}\})^+.
 \tag{3.40}$$

Lokalisierung

Wie im eindimensionalen Fall muss das unendliche Gebiet auf dem unsere PIDE lebt eingeschränkt werden. Dazu schränken wir unser unendliches Gebiet $\Omega^2 = \mathbb{R}^2$ ein auf das Gebiet $\Omega_e^2 = [x_{min}, x_{max}] \times [x_{min}, x_{max}]$. Das ohnehin endliche Zeitintervall $I_T = [0, T]$ bleibt erhalten. Damit ist das Problem (3.39) auf dem Gebiet $\Omega_e^2 \times I_T$ zu lösen. Die Anfangsbedingungen (3.40) bleiben erhalten, wir müssen allerdings noch Randbedingungen angeben. Für $x_1 = x_{max}$ bzw. $x_2 = x_{max}$ gilt aufgrund der Struktur der Put-Option offensichtlich $V((x_1, x_2), \tau) = 0$. Es bleiben also noch zwei weitere Ränder. Gilt $x_2 = x_{min}$, so hängt der Wert der Option nur noch vom Kurs der ersten Aktie ab. Das Problem reduziert sich an diesem Rand also auf die Bewertung einer eindimensionalen Option mit Volatilität σ_1 und Laufzeit τ . Das Gleiche gilt für den Fall $x_2 = x_{min}$. Wir erhalten damit also die Randbedingungen

$$V((x_{max}, x_2), \tau) = 0, \quad (3.41)$$

$$V((x_1, x_{max}), \tau) = 0, \quad (3.42)$$

$$V((x_{min}, x_2), \tau) = V_{1d}(x_2, \sigma_2, \tau) \quad (3.43)$$

$$V((x_1, x_{min}), \tau) = V_{1d}(x_1, \sigma_1, \tau), \quad (3.44)$$

dabei bezeichnet $V_{1d}(x_i, \sigma_i, \tau)$ den entsprechenden Wert einer eindimensionalen Option mit Laufzeit τ und Volatilität σ_i , die restlichen Parameter werden aus der zweidimensionalen PIDE übernommen. Man muss also in jedem Zeitschritt den Wert von zwei eindimensionalen Optionen berechnen.

3.6.1 Zeitdiskretisierung

Sei $\Delta\tau_n := n\Delta\tau$ für $n = 1 \dots, N$ eine äquidistante Unterteilung des Zeitintervalls I_T mit Schrittweite $\Delta\tau = \frac{T}{N}$. V^n bezeichnet die numerische Approximation von $V(\mathbf{x}, n\Delta\tau)$. Wir führen analog zum eindimensionalen Problem ein Operatorsplitting durch und erhalten mittels der Finite-Differenzen-Diskretisierung das θ -Schema

$$\frac{V^{n+1} - V^n}{\Delta\tau} - \theta \mathcal{L}_{Diff}(V^{n+1}) - (1-\theta) \mathcal{L}_{Diff}(V^n) - \theta_J \mathcal{L}_{Int}(V^{n+1}) - (1-\theta_J) \mathcal{L}_{Int}(V^n) = 0 \quad (3.45)$$

mit

$$\mathcal{L}_{Diff} = \frac{1}{2} \rho \sigma_1 \sigma_2 V_{x_1 x_2} + (r - \lambda \kappa - \frac{\sigma_1^2}{2}) V_{x_1} + (r - \lambda \kappa - \frac{\sigma_2^2}{2}) V_{x_2} - (r + \lambda) V \quad (3.46)$$

und

$$\mathcal{L}_{Int} = \lambda \int_{\mathbb{R}^2} V(\mathbf{x} + \mathbf{y}, \tau) f(\mathbf{y}) d\mathbf{y}. \quad (3.47)$$

Dabei gilt wie im eindimensionalen Fall $\mathcal{L} = \mathcal{L}_{Diff} + \mathcal{L}_{Int}$. Bei der Lösung der zweidimensionalen PIDE werden wir den Integralteil explizit behandeln und für den Differentialteil die implizite Variante wählen. Die Ortsdiskretisierung erfolgt mittels Finiter-Differenzen und einem Kollokationsverfahren für den Integralteil, welches die Simpson-Kubaturformel verwendet.

3.6.2 Ortsdiskretisierung

Zur Diskretisierung der ersten Ableitungen sowie der zweiten Ableitungen, die nur in eine Richtung erfolgen, können wir die Differenzenquotienten aus dem eindimensionalen Fall übernehmen. Eine Neuerung stellt jedoch die gemischte zweite Ableitung dar. Sei nun $\Delta(x_1, x_2) = (x_{min} + i\Delta x, x_{min} + j\Delta x)$ für $i, j = 0, \dots, M$ eine äquidistante Unterteilung von Ω_e^2 mit Maschenweite $\Delta x = \frac{x_{max} - x_{min}}{M}$. Wir werden im weiteren Verlauf die verkürzte Schreibweise $V(x_{min} + i\Delta x, x_{min} + j\Delta x, n\Delta\tau) = V_{i,j}^n$ verwenden. Der Differenzenquotient für die gemischte Ableitung ergibt sich dann wie folgt:

$$\begin{aligned} [V_{i,j}^n]_{x_1 x_2} &= \left[[V_{i,j}^n]_{x_1} \right]_{x_2} \\ &= \left[\frac{V_{i+1,j}^n - V_{i-1,j}^n}{2\Delta x} \right]_{x_2} \\ &= \frac{V_{i+1,j+1}^n - V_{i-1,j+1}^n - V_{i+1,j-1}^n + V_{i-1,j-1}^n}{4(\Delta x)^2}. \end{aligned} \quad (3.48)$$

Dieser zentrale Differenzenquotient hat wie auch die anderen vorgestellten zentralen Quotienten die Ordnung 2.

Stabilität der Ortsdiskretisierung

Eine geschlossene Formel für die Berechnung des Optionspreises ist leider nur im Fall der mehrdimensionalen Black-Scholes-Gleichung möglich, d.h. wenn die Kurse der zugrunde liegenden Wertpapiere einer geometrischen Brownschen Bewegung folgen. In diesem Fall ist wieder eine Transformation der Gleichung auf die Wärmeleitungsgleichung möglich, es treten dann also weder der konvektive Term noch der reaktive in der zu lösenden PDE auf. Folgen die Kurse jedoch wie in unserem Fall einem Sprung-Diffusionsprozess, so sind keine geschlossenen Lösungen bekannt. Ebenso existiert keine entsprechende Transformation, weshalb wir die konvektiven Terme ebenso wie die reaktiven berücksichtigen müssen. Es ergeben sich also hier die selben Probleme wie bei der Diskretisierung des eindimensionalen Problems. Die Stabilitätsaussagen für die verwendeten Differenzenquotienten sind ebenfalls gleich, daher wird hier auf eine erneute Diskussion dieser Argumente verzichtet.

3.6.3 Diskretisierung des Integraloperators

Betrachten wir nun das Integral

$$\int_{\mathbb{R}^2} V(\mathbf{x} + \mathbf{y}, \tau) f(\mathbf{y}) d\mathbf{y} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V(x_1 + y_1, x_2 + y_2) f(y_1, y_2) dy_1 dy_2. \quad (3.49)$$

Zur numerischen Auswertung muss das unendliche Integrationsgebiet $\Omega_I = \mathbb{R}^2$ beschränkt werden. Wir berechnen das Integral daher auf dem endlichen Gebiet $\Omega_{I_e} = [y_{min}, y_{max}] \times [y_{min}, y_{max}]$. Sei dazu $\Delta(y_1, y_2) = (y_{min} + i\Delta y, y_{min} + j\Delta y)$ für $i, j = 0, \dots, M$ eine äquidistante Unterteilung von Ω_{I_e} mit Maschenweite $\Delta y = \frac{y_{max} - y_{min}}{M}$.

Wir erhalten dann folgende Näherung des Integrals an der Stelle $(x_{min} + i\Delta x, x_{min} + j\Delta x)$:

$$\mathcal{I}_{i,j} = \sum_{k=-\frac{M}{2}}^{\frac{M}{2}} \sum_{l=-\frac{M}{2}}^{\frac{M}{2}} V_{i+k,j+l} f_{k,l}, \quad (3.50)$$

wobei

$$f_{k,l} = \int_{y_k - \Delta y}^{y_k + \Delta y} \int_{y_l - \Delta y}^{y_l + \Delta y} f(u, v) du dv. \quad (3.51)$$

Für die Wahrscheinlichkeitsdichte wird im zweidimensionalen Fall entweder die zweidimensionale Lognormalverteilung oder eine zweidimensionale Clayton-Copula-Dichte verwendet, wie sie in Kapitel 4 beschrieben wird. Wir wenden uns nun der Berechnung der $V_{i+k,j+l}$ und der $f_{k,l}$ zu.

Berechnung der $V_{i+k,j+l}$

Wie im eindimensionalen Fall müssen auch hier zwei grundsätzliche Fälle unterschieden werden. Wenn $i + l \leq M$ und $j + l \leq M$ liegt der gesuchte Punkt im Gebiet Ω_c^2 und kann direkt verwendet werden. Sollte einer der beiden Indizes jedoch die Grenzen des Gebiets überschreiten, so müssen wir den Optionswert an diesen Stellen mittels der Randbedingungen fortsetzen, d.h. im Fall einer Put-Option

$$V_{i+k,j+l} = 0 \quad \text{für } i + k > M, \quad (3.52)$$

$$V_{i+k,j+l} = 0 \quad \text{für } j + l > M, \quad (3.53)$$

$$V_{i+k,j+l} = V_i^{1d} \quad \text{für } i + l < 0 \text{ und } i > j, \quad (3.54)$$

$$V_{i+k,j+l} = V_j^{1d} \quad \text{für } j + l < 0 \text{ und } j > i. \quad (3.55)$$

Dabei bezeichnen V_i^{1d} bzw. V_j^{1d} den Wert der entsprechenden eindimensionalen Randoption.

Berechnung der $f_{k,l}$

Da hier ein zweidimensionales Integral zu berechnen ist reicht eine Quadraturformel wie sie im eindimensionalen Fall verwendet wird nicht mehr aus. Entsprechende Formeln für zweidimensionale Integrale nennt man Kubaturformeln. Wir werden zur Berechnung der $f_{k,l}$ die Simpson-Kubaturformel verwenden.

Simpson-Kubatur

Die Simpson-Regel zur approximativen Berechnung zweidimensionaler Integrale lautet

$$I_{h^2} = \int_{-h}^h \int_{-h}^h f(x, y) dx dy = \frac{h^2}{9} [f(-h, -h) + f(-h, h) + f(h, h) + f(h, -h) + 4(f(-h, 0) + f(h, 0) + f(0, h) + f(0, -h)) + 16f(0, 0)]. \quad (3.56)$$

Die Fehlerdarstellung der zweidimensionalen Simpson-Regel lautet

$$I_{h^2} - \int_{-h}^h \int_{-h}^h f(x, y) dx dy = \frac{4h^2}{180} (f_{xxxx}(0, 0)h^4 + f_{yyyy}(0, 0)h^4 + \mathcal{O}(h^5)). \quad (3.57)$$

3.6.4 Lösung der Gleichungssysteme

Mit Hilfe der vorgestellten Diskretisierungen der PDE und des Integralteils können wir nun abschließend die Matrixformulierung der Diskretisierung vorstellen. Die entsprechende Formulierung für unser zweidimensionales Problem lautet

$$[I - \theta \mathbf{A}] V^{n+1} = [I + (1 - \theta) \mathbf{A}] V^n + \lambda \Delta \tau \mathbf{B} V^n. \quad (3.58)$$

Die Matrix \mathbf{A} setzt sich in diesem Fall aus dem Neunpunkte-Stern

$$\begin{bmatrix} -\gamma\Delta\tau & \beta_2\Delta\tau & \gamma\Delta\tau \\ \alpha_1\Delta\tau & -(\alpha_1 + \beta_1 + \alpha_2 + \beta_2 + r + \lambda)\Delta\tau & \beta_1\Delta\tau \\ \gamma\Delta\tau & \alpha_2\Delta\tau & -\gamma\Delta\tau \end{bmatrix} \quad (3.59)$$

zusammen, wobei

$$\begin{aligned} \alpha_i &= \frac{\sigma_i^2}{2\Delta x^2} - \frac{r - \lambda\kappa - \frac{1}{2}\sigma_i^2}{2\Delta x} & \text{für } i = 1, 2, \\ \beta_i &= \frac{\sigma_i^2}{2\Delta x^2} + \frac{r - \lambda\kappa - \frac{1}{2}\sigma_i^2}{2\Delta x} & \text{für } i = 1, 2 \end{aligned} \quad (3.60)$$

und

$$\gamma = \frac{\sigma_1\sigma_2}{2\Delta x^2}. \quad (3.61)$$

Die Einträge der Matrix \mathbf{B} sind so gewählt, dass

$$[\mathbf{B}V^n]_{i,j} = \sum_{k=-\frac{M}{2}}^{\frac{M}{2}} \sum_{l=-\frac{M}{2}}^{\frac{M}{2}} V_{i+k,j+l} f_{k,l} \quad (3.62)$$

gilt. Wie vorher schon erwähnt werden wir in den numerischen Experimenten $\theta = 1$ wählen, um ein implizites Eulerschema zu erhalten. Somit entsteht das zu lösende Gleichungssystem

$$[I - \mathbf{A}]V^{n+1} = V^n + \lambda\Delta\tau\mathbf{B}V^n. \quad (3.63)$$

Leider ist die Matrix \mathbf{A} des Gleichungssystems im zweidimensionalen Fall keine Tridiagonalmatrix, d.h das Lösen des Gleichungssystems ist nicht mehr ohne Weiteres in linearer Zeit realisierbar.

Wir haben nun die Diskretisierung des ein- und zweidimensionalen Problems kennengelernt. In beiden Fällen muss eine Matrix-Vektor-Multiplikation durchgeführt werden, um den Wert der Option zu bestimmen. Im eindimensionalen Fall sorgt dieser Integralterm dafür, dass der Aufwand der Berechnung quadratisch in der Anzahl der Ortsdiskretisierungsschritte ansteigt, da man das Gleichungssystem mit linearem Aufwand lösen kann. Im nächsten Kapitel werden nun die \mathcal{H} -Matrizen vorgestellt. Wie wir sehen werden, ist es mit Hilfe der \mathcal{H} -Matrix-Approximation möglich, die gewünschte Matrix-Vektor Multiplikation in linearer Zeit durchzuführen. Dies gilt ebenso für das Invertieren der Gleichungssystemmatrix \mathbf{A} im zweidimensionalen Fall. Durch Anwendung der \mathcal{H} -Matrix Approximation lassen sich also, wie wir in Kapitel 5 sehen werden, beide Probleme bis auf logarithmische Faktoren mit linearem Aufwand lösen.

Kapitel 4

Hierarchische Matrizen

Wie wir gesehen haben, führen die vorgestellten Sprung-Diffusionsmodelle zu partiellen Integro-Differentialgleichungen. Durch die Verwendung des im vorigen Abschnitts beschriebenen Kollokationsverfahrens beschränkt sich die Auswertung des Integrals auf eine Matrix-Vektor-Multiplikation. Diese muss in jedem einzelnen Zeitschritt des Verfahrens durchgeführt werden und erhöht damit den Aufwand erheblich. Im eindimensionalen Fall führt dies beispielsweise zu einem Gesamtaufwand in der Größenordnung $\mathcal{O}(N^2)$, wobei N die Anzahl an Gitterpunkten angibt. Wenn wir die nötige Multiplikation in linearer oder beinahe linearer Komplexität durchführen könnten, wäre das eindimensionale Problem insgesamt ebenfalls in linearer Zeit lösbar. Um dies zu erreichen stellen wir die Kollokationsmatrix des Integralteils als \mathcal{H} -Matrix dar. Im eindimensionalen Fall reicht dies aus, da das Lösen des eigentlichen Gleichungssystems ebenfalls mit linearem Aufwand möglich ist, weil es sich bei der entsprechenden Matrix um eine Tridiagonalmatrix handelt. Bei der Bewertung einer Basket-Option ist dies nicht der Fall und das Lösen des Gleichungssystems hat selbst einen Aufwand von $\mathcal{O}(N^4)$. Wir werden daher in diesem Fall die Systemmatrix des Gleichungssystems ebenfalls durch eine \mathcal{H} -Matrix approximieren und das Gleichungssystem durch die Invertierung dieser \mathcal{H} -Matrix lösen.

In diesem Kapitel werden die Grundprinzipien Niedrigrangapproximation und Matrixpartitionierung, die den \mathcal{H} -Matrizen zu Grunde liegen, erläutert. Dabei wird hauptsächlich Bezug genommen auf die ersten beiden Kapitel von [Beb08] und [Hac07].

4.1 Niedrigrangmatrizen

Die beiden wichtigsten Grundlagen für die Effizienz von \mathcal{H} -Matrizen sind Niedrigrangmatrizen und Matrixpartitionierung. Wir befassen uns in diesem Abschnitt zunächst mit den Grundlagen der Niedrigrangmatrizen und im gehen im nächsten Abschnitt dann zu Matrixpartitionierung über.

Seien $m, n \in \mathbb{N}$. Dann bezeichnen wir die Menge der Matrizen $A \in \mathbb{C}^{m \times n}$ mit höchstens k unabhängigen Zeilen oder Spalten als

$$\mathbb{C}_k^{m \times n} := \{A \in \mathbb{C}^{m \times n} : \text{rang}(A) \leq k\}.$$

Hierbei ist es wichtig zu erwähnen, dass der entstehende Raum $\mathbb{C}_k^{m \times n}$ kein linearer Raum ist, da der Rang der Summe zweier Rang- k Matrizen normalerweise nur durch $2k$ beschränkt ist und nicht durch k . Für den Rang der Unterblöcke einer solchen Rang- k Matrix gilt jedoch ebenfalls die Beschränkung durch k .

4.1.1 Effiziente Darstellung

Eine wichtige Eigenschaft einer Matrix $A \in \mathbb{C}_k^{m \times n}$ ist die Möglichkeit, sie in effizienter Weise als Produkt zweier Matrizen niedrigerer Dimension darzustellen. Man nennt diese Darstellung die äußere Produktform.

Satz 4.1.1

Eine Matrix $A \in \mathbb{C}^{m \times n}$ liegt in $\mathbb{C}_k^{m \times n}$ genau dann, wenn es Matrizen $U \in \mathbb{C}^{m \times k}$ und $V \in \mathbb{C}^{k \times n}$ gibt, so dass

$$A = UV^H$$

gilt.

Wir können eine solche Matrix daher mit Hilfe von Vektoren $u_i, v_i, i = 1 \dots k$ darstellen. Diese benötigen nur $k \cdot (n + m)$ Speichereinheiten anstatt der $m \cdot n$ Speichereinheiten, die die ursprüngliche Darstellung benötigt. Neben der Speicherreduktion vereinfacht diese Darstellung ebenfalls die Matrix-Vektor Multiplikation:

$$Ax = UV^H x = U(V^H x).$$

Um zum Beispiel die Operation $y := y + Ax$ komponentenweise durchzuführen, kann man A auch auf folgende Art und Weise mit x multiplizieren:

1. $z := V^H x \in \mathbb{C}^k$
2. $y := y + Uz$

Man braucht dabei anstatt der üblichen $2mn$ Operationen, die bei der vollen Darstellung von A benötigt werden, nur $2k(m+n) - k$ Operationen. Die Kosten für andere Operationen können in [Hac07] Abschnitt 2.3 nachgelesen werden.

Für die Herleitung der genauen Definition einer Niedrigrangmatrix nehmen wir an, dass $m = n$ gilt. Es kann vorkommen, dass eine Darstellung der Matrix in ihrer äußeren Produktform nicht vorteilhaft ist. Betrachten wir dazu eine Matrix mit vollem Rang. Die Darstellung in der Produktform ersetzt in der Komplexität der Matrix eine Dimension durch $2k$, d.h. statt $mn = m^2$ haben wir eine Komplexität von $k(m+n) = 2km$. Im Falle eine Matrix mit vollem Rang ($k = m$) bedeutet dies allerdings eine Komplexität von $2m^2$, welche doppelt so hoch ist wie in der komponentenweisen Darstellung. Daher charakterisiert die folgende Definition diejenigen Matrizen, für die eine Darstellung in der äußeren Produktform sinnvoll ist.

Definition 4.1.2 [NIEDRIGRANGMATRIX]

Eine Matrix $A \in \mathbb{C}_k^{m \times n}$ wird als Niedrigrangmatrix bezeichnet, falls

$$k(m+n) < n \cdot n \tag{4.1}$$

gilt.

Es werden also alle Niedrigrangmatrizen in der äußeren Produktform dargestellt, für alle anderen wird die komponentenweise Darstellung gewählt.

4.1.2 Niedrigrangapproximation

Obwohl Matrizen üblicherweise vollen Rang haben, ist es möglich, diese durch Niedrigrangmatrizen zu approximieren. Der folgende Satz sagt aus, dass die Bestapproximation durch eine Matrix aus $\mathbb{C}_k^{m \times n}$ zu einer gegebenen Matrix $A \in \mathbb{C}^{m \times n}$, $m \geq n$, durch die Singulärwertzerlegung $A = U\Sigma V^H$, mit $U^H U = I_n = V^H V$, und einer Diagonalmatrix $\Sigma \in \mathbb{R}^{n \times n}$ mit Einträgen $\sigma_1 \geq \dots \sigma_n \geq 0$, gegeben ist.

Satz 4.1.3 [BESTAPPROXIMATION]

Sei $A = U\Sigma V^H$ die Singulärwertzerlegung von $A \in \mathbb{C}^{m \times n}$, $m \geq n$. Dann gilt für $k \in \mathbb{N}$, mit $k \leq n$

$$\min_{M \in \mathbb{C}_k^{m \times n}} \|A - M\| = \|A - A_k\| = \|\Sigma - \Sigma_k\|. \quad (4.2)$$

Hierbei ist $A_k := U\Sigma_k V^H \in \mathbb{C}_k^{m \times n}$ und $\Sigma_k := \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0) \in \mathbb{R}^{n \times n}$. Der auftretende Fehler in der 2-Norm bzw. der Frobenius Norm ist dabei

$$\|A - A_k\|_2 = \sigma_{k+1} \quad \text{und} \quad \|A - A_k\|_F = \sqrt{\sum_{i=k+1}^n \sigma_i^2}. \quad (4.3)$$

Für den Beweis des Satzes sei hier auf [Hac07], Anhang C, bzw. [Mir60] verwiesen. Um die äußere Produktform der Approximation zu erhalten, muss dann noch U oder V mit Σ_k multipliziert werden. Wie man in (4.3) sehen kann ist es möglich, den Fehler der Approximation durch Betrachten der Singulärwertzerlegung zu bestimmen. Die Information über den Fehler kann auf zwei unterschiedliche Weisen genutzt werden. Zum einen kann man, wenn der maximale Rang k einer Approximation bekannt ist, den damit verbundenen Fehler bestimmen. Wenn zum anderen der relative Fehler ϵ der Approximation A_k gegeben ist, zum Beispiel durch

$$\|A - A_k\|_2 < \epsilon \|A_k\|_2, \quad (4.4)$$

dann ist auch der für die Approximation benötigte Rang k gegeben, und zwar durch

$$k(\epsilon) := \min\{k \in \mathbb{N} : \sigma_{k+1} < \epsilon \sigma_1\}.$$

Um eine optimale Approximation zu finden, ist es also nötig, die Singulärwertzerlegung einer Matrix zu berechnen. Falls die Matrix A in der äußeren Produktform vorliegt und niedrigen Rang hat, ist dies in wesentlich weniger Operationen als in der komponentenweisen Darstellung möglich. Bei allgemeinen Matrizen aus $\mathbb{C}^{m \times n}$ werden $21mn^2$ Operationen benötigt, bei einer Matrix in der Produktdarstellung mit Rang k hat die Zerlegung jedoch nur eine Komplexität von $\mathcal{O}(k^2(m+n))$. Eine genaue Herleitung dieses Aufwandes findet man in [Beb08].

4.2 Matrixpartitionierung

Die zweite Grundlage für die \mathcal{H} -Matrizen ist die Matrixpartitionierung. Wir werden hier kurz vorstellen, was unter einer Partition zu verstehen ist, wann eine solche zulässig ist und werden zeigen, dass eine Tensorprodukt-Partition keine sinnvolle Wahl gegenüber einer hierarchischen Partition ist.

Normalerweise können nur Teilblöcke einer Matrix $X \in \mathbb{C}^{m \times n}$ die aus einer geeigneten Partition $I \times J$ der Matrix entstehen durch Niedrigrangmatrizen approximiert werden.

Definition 4.2.1 [PARTITION]

Seien $I, J \subset \mathbb{N}$. Eine Teilmenge $P \subset \mathcal{P}(I \times J)$ der gesamten Teilmengen von $I \times J$ nennt man eine Partition, wenn

$$I \times J = \bigcup_{b \in P} b. \quad (4.5)$$

Dabei gilt für alle $b_1, b_2 \in P$ mit $b_1 \neq b_2$ $b_1 \cap b_2 = \emptyset$. Die Element $b \in P$ nennt man Index Blöcke oder kurz Blöcke.

Das Ziel besteht nun darin, einen Algorithmus zu finden, mit dem man eine Matrix A mit Hilfe einer Partition P so zerlegen kann, dass jeder Block A_b , $b \in P$ entweder klein ist, oder durch eine Matrix mit niedrigem Rang approximiert werden kann. Dabei müssen zwei Dinge beachtet werden. Einerseits muss die Partitionierung fein genug sein, damit möglichst viele Blöcke approximiert werden können, andererseits dürfen es nicht zu viele Blöcke werden, da sonst der Aufwand für arithmetische Operationen zu hoch wird. Die Frage, ob ein Block b approximiert werden kann, hängt häufig mit dem zugrunde liegenden Problem zusammen. Da wir uns hier nicht auf ein spezielles Problem beschränken wollen, führen wir eine abstrakte Bedingung ein, die die Frage nach der Zulässigkeit eines Block klärt, die *Zulässigkeitsbedingung*. Diese muss folgende Eigenschaften besitzen:

- Wenn b zulässig ist, dann fallen die Singulärwerte von A_b exponentiell ab.
- Die Bedingung kann für jeden Block $t \times s \in P$ mit höchstens $\mathcal{O}(|t| + |s|)$ Operationen überprüft werden.
- Wenn b zulässig ist, dann ist auch jede Teilmenge $b' \subset b$ zulässig.

Definition 4.2.2 [ZULÄSSIGKEIT]

Eine Partition P heißt zulässig, wenn jeder Block $t \times s \in P$ entweder zulässig oder klein ist, d.h. die Kardinalitäten $|t|$ und $|s|$ genügen der Bedingung $\min\{|t|, |s|\} \leq n_{min}$ für ein gegebenes $n_{min} \in \mathbb{N}$.

Die nächste Frage ist die nach der Art der Partition. Wir betrachten hier den Fall $I = J$, d.h. die Matrix wird in Zeilen und Spalten identisch unterteilt. Es gibt zwei grundsätzliche Möglichkeiten, eine Matrix für eine Partitionierung zu unterteilen. Die erste ist die Tensorprodukt-Partition. Dabei wird die Matrix durch einfügen senkrechter und waagerechter Grenzen, welche durch die gesamte Matrix verlaufen, unterteilt. Ein Beispiel einer solchen Unterteilung ist in Abb. 4.1 gezeigt. Die nächste Verfeinerungsstufe wird durch das Hinzufügen weiterer Unterteilungslinien erreicht.

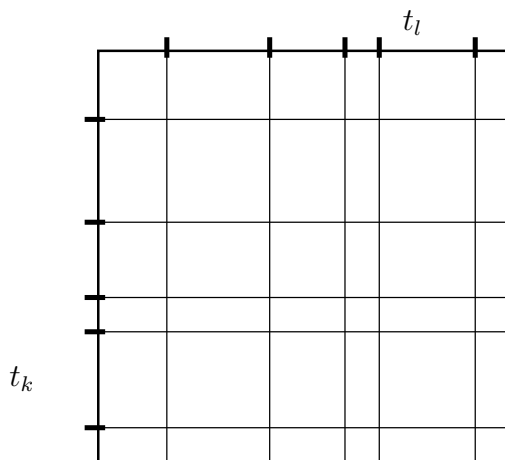


Abbildung 4.1: Tensorprodukt Partition

Im Gegensatz dazu gibt es die hierarchische Partition. Diese zeichnet sich dadurch aus, Verfeinerungen durch Unterteilung schon erstellter Blöcke erfolgen. Abb. 4.2 zeigt ein Beispiel einer hierarchischen Partition.

Zu Beginn dieses Kapitels haben wir gesagt, dass die benötigten Operationen in beinahe linearer Zeit zu bearbeiten sind. Damit dies möglich ist, muss diese Komplexitätsgrenze

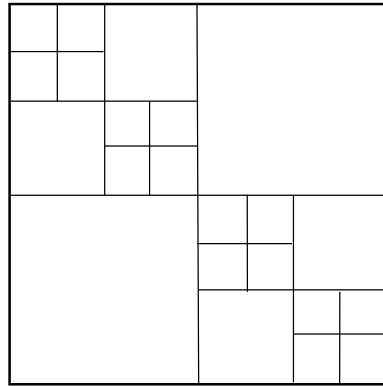


Abbildung 4.2: Hierarchische Partition

auch für den Speicherbedarf der Struktur gelten, unter der wir unsere Matrix abspeichern. Die Tensorprodukt-Partition erfüllt diese Bedingung nicht. In [Beb08] wird dazu beispielhaft der Speicherbedarf einer Matrix untersucht, welche folgenden Bedingungen genügt:

1. Jeder Diagonalenblock ist vollbesetzt.
2. Alle anderen Blöcke sind zulässig und als Rang 1 Matrix gespeichert, d.h. $A_{ts} = uv^H$, wobei $u \in \mathbb{C}^t$ und $v \in \mathbb{C}^s$.

Zunächst betrachten wir die Tensorprodukt Partition

$$P = P_I \times P_J = \{t_k \times t_l : t_k, t_l \in P_I\}, \quad (4.6)$$

wobei $P_I := \{t_k, k = 1, \dots, p\}$ eine Partition der Indexmenge I ist. Die Diagonalelemente von A benötigen $\sum_{k=1}^p |t_k|^2$ Speichereinheiten. Die Blöcke außerhalb der Diagonalen benötigen

$$\sum_{k \neq l} |t_k| + |t_l| = 2 \sum_{k=1}^p \sum_{\substack{l=1 \\ l \neq k}}^p |t_k| = 2(p-1) \sum_{k=1}^p |t_k| = 2(p-1)|I| \quad (4.7)$$

Speichereinheiten. Mit Hilfe der Cauchy-Schwartz-Ungleichung ergibt sich ein minimaler Speicheraufwand von $|I|^2/p + 2(p-1)|I|$ Einheiten. Bei optimaler Wahl von $p = \sqrt{|I|/2}$ ergibt dies immer noch einen Speicherbedarf von der Ordnung $|I|^{3/2}$ und ist somit nicht ausreichend für die Komplexität die wir erreichen wollen.

Wählen wir für die hierarchische Partition die in Abb. 4.2 dargestellte Form, lässt sich die Matrix allerdings mit beinahe linearer Komplexität speichern. Nehmen wir an, die Matrix A besitzt $\mathbb{N} \ni n = 2^p$ Einträge. Wir nehmen nun an, dass die Verfeinerung der Partition einer bestimmten Rekursion folgt, und weiterhin, dass t im l -ten Level generiert worden, dann unterteilen wir $t = \{i_1, \dots, i_{2^{p-l}}\}$ in zwei gleich große Teile

$$t_1 = \{i_1, \dots, i_{2^{p-l-1}}\} \quad \text{und} \quad t_2 = \{i_{2^{p-l-1}+1}, \dots, i_{2^{p-l}}\} \quad (4.8)$$

und erhalten so eine 2×2 Unterteilung des ursprünglichen Blocks. Wir postulieren, dass bei der Tensorprodukt Partition die Blöcke außerhalb der Diagonalen als Rang 1 Matrizen dargestellt sind. Wenn wir diese Verfeinerung auf der Diagonalen fortführen erhalten wir auf Level 3 die Partition aus Abb. 4.2. Sei N_p^{st} der Speicherbedarf der hierarchischen Partition bei 2^p Einträgen. Die Blöcke außerhalb der Diagonalen brauchen $4 \cdot 2^{p-1} = 2^{p+1}$ Speichereinheiten. Es ergibt sich die Rekursion

$$N_p^{st} = 2N_{p-1}^{st} + 2^{p+1} \quad (4.9)$$

mit $N_0^{st} = 1$. Folgt man der Rekursion, ergibt sich ein Speichervolumen von $N_p^{st} = (2p + 1)2^p = 2n \log_2 n + n$ Einheiten. Damit haben wir einen Speicheraufwand in der gewünschten Komplexität erhalten.

4.3 Cluster-Bäume und Block-Cluster-Bäume

Wir haben nun gesehen, dass wir mit Hilfe einer hierarchischen Partition und einer Niedrigrangapproximation der Blöcke eine Matrix mit beinahe linearem Speicherbedarf abspeichern können. Die gängigen Operationen, wie zum Beispiel die Matrix-Vektor-Multiplikation, können ebenfalls in beinahe linearer Zeit durchgeführt werden. Es fehlt jetzt noch eine Struktur, um die Partition zu speichern. Diese muss ebenfalls in fast linearer Zeit zu erstellen sein und wird in diesem Abschnitt vorgestellt.

Wir durchsuchen nicht alle Partitionen der von $I \times J$ nach einer Optimalen, da der Aufwand dafür zu hoch ist. sondern betrachten nur Partitionen, welche durch iterative Teilung von I, J entstehen. Dies garantiert zwar keine optimale Partition, man kann jedoch zeigen, dass die entstehende Partition unter realistischen Annahmen von ausreichend niedriger Komplexität ist. Um die Indexmengen zu partitionieren, brauchen wir eine Regel, um sie zu unterteilen. Diese führt zu sogenannten Cluster-Bäumen (siehe auch [HN89]).

4.3.1 Cluster Bäume

Definition 4.3.1 [CLUSTER-BAUM]

Ein Baum $T_I = (V, E)$ mit Knoten V und Kanten E wird Cluster-Baum einer Menge $I \subset \mathbb{N}$ genannt, wenn folgende Bedingungen gelten:

1. I ist die Wurzel von T .
2. $\emptyset \neq t = \bigcup_{t' \in S(t)} t'$ für alle $t \in V \setminus \mathcal{L}(T_I)$.
3. Der Grad $\deg t := |S(t)| \geq 2$ aller Knoten $t \in V \setminus \mathcal{L}(T_I)$ ist nach unten beschränkt.

Dabei ist die Menge der Söhne $S(t) := \{T' \in V : (t, t') \in E\}$ paarweise disjunkt und $\mathcal{L}(T_I) := \{t \in V : S(t) = \emptyset\}$ bezeichnet die Menge der Blätter von T_I .

Mit den *Level* eines Knoten $t \in T_I$ bezeichnen wir seinen Abstand zur Wurzel des Baumes. Bedingung 2 impliziert, dass jedes Level des Baumes eine Partition von I enthält.

Bemerkung: 4.3.2: Es ist sinnvoll, im weiteren Verlauf davon auszugehen, dass die Cluster eine minimale Größe $n_{min} > 1$ haben, da die äußere Produktform für sehr kleine Blöcke nicht effizient ist. Eine minimal Clustergröße führt dann zu einer minimalen Blockgröße in der Partition.

Da die Anzahl an Blättern $|\mathcal{L}(T_I)|$ durch $|I|/n_{min}$ beschränkt ist, lässt sich zeigen, dass ein Cluster Baum nur linear viel Speicherplatz benötigt. Dies ist wichtig, da wir immer darauf achten müssen unsere Komplexitätsgrenze nicht zu überschreiten und insgesamt einen beinahe linearen Aufwand zu garantieren. Daher benötigen wir folgendes Lemma.

Lemma 4.3.3

Sei $q := \min_{t \in T \setminus \mathcal{L}(T_I)} \deg t \geq 2$. Dann gilt für die Anzahl an Knoten in T_I

$$|T_I| \leq \frac{q|\mathcal{L}(T_I)| - 1}{q - 1} \leq 2|\mathcal{L}(T_I)| - 1. \quad (4.10)$$

Die Anzahl der Blätter in einem solchen Baum ist linear von der Anzahl an Indizes $|I|$ abhängig. Um allerdings eine logarithmische Tiefe des Baumes gewährleisten zu können, müssen alle von S erzeugten Cluster von vergleichbarer Größe sein. Wir werden daher balancierte Cluster-Bäume verwenden.

Definition 4.3.4 [BALANCIERTER BAUM]

Ein Baum T_I ist balanciert, wenn

$$R := \min_{t \in T_I \setminus \mathcal{L}(T_I)} \{|t_1| \setminus |t_2|, t_1, t_2 \in S(t)\} \quad (4.11)$$

unabhängig von $|I|$ nach unten beschränkt ist.

Für einen solchen balancierten Baum gilt dann, dass die Tiefe $L(T_I) := \max_{t \in T_I} \text{Level } t + 1$ des Baumes nur logarithmisch von der Anzahl an Indizes $|I|$ abhängt, denn man kann folgendes Lemma zeigen: (siehe [Beb08]).

Lemma 4.3.5

Sei T_I ein balancierter Cluster Baum mit $q := \min_{t \in T_I \setminus \mathcal{L}(T_I)} \text{deg } t \geq 2$. Dann gilt für die Tiefe

$$L(T_I) \leq \log_{\zeta}(|I|/n_{\min}) + 1 \sim \log |I| \quad (4.12)$$

und $|t| \leq |I|\zeta^{-l}$, wobei l den Level von t angibt und $\zeta := R(q - 1) + 1$.

Die balancierten Cluster-Bäume genügen also unseren Bedingungen an die Komplexität der Struktur. Wir werden uns hier nicht mit der expliziten Konstruktion eines Cluster-Baumes befassen. Eine Methode einen solchen balancierten Baum zu erstellen, findet sich ebenfalls in [Beb08] in Abschnitt 1.4.1. Der wichtigste Teil der Konstruktion besteht darin, einen Algorithmus anzugeben, der die Blöcke so zu unterteilt, dass am Ende ein balancierter Cluster-Baum entsteht. In besagtem Abschnitt wird eine Methode vorgestellt, welche darauf basiert, eine Hauptrichtung in einer Punktmenge zu finden. Diese Methode wird auch *principal component analysis* genannt. Dabei wird der Fall einer Finiten-Element-Diskretisierung eines elliptischen Operators betrachtet, bei der jedem Index ein Teil des Rechengebiets zugewiesen wird. Die dort vorgestellte Methode führt dann zu balancierten Cluster-Bäumen. Aus dem so gewonnenen Cluster-Baum T_I kann dann eine zulässige Partition P von $I \times J$ erstellt werden. Diesen Vorgang werden wir im nächsten Abschnitt an einem kurzen Beispiel vorstellen.

Zum Abschluss dieses Abschnitts möchten wir noch auf eine weitere Eigenschaft von Cluster-Bäumen aufmerksam machen, welche für spätere Abschätzungen benötigt wird. Für einen Cluster-Baum T_I gilt:

$$\sum_{t \in T_I} |t| \leq L(T_I)|I|, \quad (4.13)$$

da jedes der $L(T_I)$ Level des Baumes aus disjunkten Teilmengen von I besteht.

4.3.2 Analyse einer einfachen Partition

Die in diesem Abschnitt gezeigte Partitionierung besitzt eine minimale Anzahl an Blöcken, wobei die generierten, zulässigen Blöcke maximale Größe haben. Wir werden die Zulässigkeitsbedingung

$$\text{diam} X_t \leq \eta \text{dist}(X_t, X_s) \quad (4.14)$$

für den Block $t \times s$ verwenden. Die Grundidee besteht nun darin, einen gegebenen Block $t \times s$ in einen maximalen zulässigen Block $t \times \tau$ und einen nicht zulässigen Teil $t \times (s \setminus \tau)$

zu unterteilen. Dabei bezeichnet $\tau := \mathcal{F}(t) \cap s$ und

$$\mathcal{F}(t) := \{j \in J : \text{diam} X_t \leq \eta \text{ dist}(X_t, X_s)\}$$

das *Far-Field* von $t \subset I$. Während $t \times \tau$ (4.14) erfüllt und daher P abgespeichert wird ist dies für den übrigen Teil $t \times (s \setminus \tau)$ von $t \times s$ nicht der Fall. Das Verfahren wird nun rekursiv auf die Unterblöcke $t' \times (s \setminus \tau)$, $t' \in S(t)$ angewendet. Wenn $S(t) = \emptyset$ gilt, wird der Block als nicht zulässiger, aber kleiner Block der Partition hinzugefügt.

Bemerkung: 4.3.6: Die Berechnung von $\mathcal{F} \cap s$ benötigt $\mathcal{O}(|t| \cdot |s|)$ Operationen, was zu einer quadratischen Gesamtkomplexität führen würde. Daher ersetzen wir $\mathcal{F}(t)$, $t \in T_I$ durch die Menge

$$\tilde{\mathcal{F}}(t) := \left\{ j \in J : \text{diam} X_t \leq \frac{\eta}{\eta + 1} \text{ dist}(z_t, X_t) \right\}, \quad (4.15)$$

wobei $z_t \in \mathbb{R}^d$ ein zufällig gewählter, fester Punkt ist. Offensichtlich ist $\tilde{\mathcal{F}}(t) \subset \mathcal{F}(t)$. Die Berechnung von $\tilde{\mathcal{F}}(t) \cap s$ kann mit $\mathcal{O}(|t| + |s|)$ Operationen berechnet werden. Somit ergibt sich Algorithmus 3 zur Lösung des Problems.

```

procedure Partition( $t, s, \text{var} P$ )
  berechne  $\tau := \tilde{\mathcal{F}}(t) \cap s$ 
  wenn  $\tau \neq \emptyset$  dann setze  $P := P \cup \{t \times \tau\}$ 
  wenn  $t \in \mathcal{L}(T_I)$  oder  $|s \setminus \tau| \leq n_{\text{min}}$  dann setze
     $P := P \cup \{t \times (s \setminus \tau)\}$ 
  ansonsten wiederhole für  $t' \in S(t)$ 
    Partition( $t', s \setminus \tau, P$ )
    
```

Algorithmus 3: Maximal zulässige Partition

Wendet man den Algorithmus an, so erhält man eine zulässige Partition P von $I \times J$, falls P zuvor mit $P = \emptyset$ initialisiert wurde. Hierbei sei noch bemerkt, dass es nicht nötig ist, den Cluster-Baum T_I zuvor zu konstruieren, da der Algorithmus ihn nur *hinabsteigt*, was durch die Anwendung von S ohne das vorherige Speichern des Baumes erreicht werden kann.

Wir werden nun die Kosten betrachten, die der Algorithmus bei Anwendung in quasi-uniformen Gittern verursacht. Da für jeden Cluster $t \in T_I$ zumindest die Blöcke $t \times \tau$ und $t \times (s \setminus \tau)$ zu P hinzugefügt werden, erhalten wir für die maximale Anzahl an Blöcken in P

$$c_{sp}(P) := \max_{t \in T_I} |\{s \subset J : t \times s \in P\}|, \quad (4.16)$$

die zu einer gegebenen Menge an Spaltenindizes t gehört, durch 2 nach unten beschränkt ist.

Lemma 4.3.7

Die Anzahl $|P|$ an generierten Blöcken ist beschränkt durch $2|T_I| \sim |I|$

Beweis. Da C_{sp} durch 2 beschränkt ist, erhalten wir

$$|P| = \sum_{t \in T_I} |\{s \subset J : t \times s \in P\}| \leq c_{sp}(P) |T_I| \leq 2|T_I|. \quad (4.17)$$

Die Behauptung folgt aus Lemma (4.3.3). □

Um den Aufwand der Konstruktion von P abschätzen zu können, benötigen wir zunächst eine Schranke für die Größe des *near-fields* $\tilde{\mathcal{N}}(t) = J \setminus \tilde{\mathcal{F}}(t)$. Wir nehmen dazu an, dass die Mengen X_i die folgenden Eigenschaften besitzen:

- Quasi-Uniformität, d.h. es gibt eine Konstante $c_U > 0$, so dass

$$\max_{i \in I} \mu(X_i) \leq c_U \mu(X_i), \quad i \in I \quad (4.18)$$

- *Shape-Regularity*, d.h. es gibt eine Konstante $c_R > 0$, so dass

$$\mu(X_i) \geq c_R (\text{diam} X_i)^m, \quad i \in I \quad (4.19)$$

Dabei ist $\mu(M)$ das Maß einer m -dimensionalen Mannigfaltigkeit $M \subset \mathbb{R}^d$. Wir nehmen weiterhin an, dass unser Gebiet *Omega* eine m -dimensionale Mannigfaltigkeit ist, d.h. es gibt eine Konstante $c_\Omega > 0$, so dass für alle $z \in \Omega$ gilt:

$$\mu(\Omega \cap B_r(z)) \leq c_\Omega r^m \quad \text{für alle } r > 0. \quad (4.20)$$

Aus (4.18) und (4.19) erhalten wir nun die Aussage, dass es eine Konstante $c_S > 0$ gibt, so dass

$$\max_{i \in I} \text{diam} X_i \leq c_S \min_{i \in I} \text{diam}(X_i) \quad (4.21)$$

gilt. Eine weitere Eigenschaft, die wir benötigen werden, ist die geometrische Balance des Cluster-Baumes, d.h es gibt Konstanten $c_g, c_G > 0$, so dass für alle Level $\ell = 0, \dots, L(T_I) - 1$

$$(\text{diam}(X_t))^m \leq c_g 2^{-\ell} \quad \text{und} \quad \mu(X_t) \geq \frac{2^{-\ell}}{c_G} \quad \text{für alle } t \in T_I^{(\ell)} \quad (4.22)$$

gilt. Die Erstellung des Cluster-Baumes mittels der *principal component analysis* garantiert diese Eigenschaft. Wenden wir uns nun wieder dem *Near-Field* zu.

Lemma 4.3.8

Sei T_I ein geometrisch-balancierter Cluster Baum, d.h. (4.22) gilt. Dann erhalten wir unter den oben gemachten Annahmen die Abschätzung

$$|\tilde{\mathcal{N}}(t)| \leq c \left(1 + \frac{1}{\eta}\right)^m |t| \quad \text{für alle } t \in T_I. \quad (4.23)$$

Beweis. Den vollständigen Beweis findet man in [Beb08], Abschnitt 1.4.2. Dazu werden hauptsächlich die obigen Eigenschaften des Cluster-Baumes verwendet. \square

Satz 4.3.9

Die Anzahl an Operationen und die Speichermenge, die bei der Anwendung von Algorithmus 3 auf $I \times J$ benötigt werden, sind von der Ordnung $\eta^{-m}|I| \log |I|$.

Beweis. Algorithmus 3 steigt den Cluster-Baum T_I hinab. In jedem Knoten $t \in T_I$ werden höchstens

$$c(|t| + |\tilde{\mathcal{N}}(t)|) \quad (4.24)$$

Operationen benötigt, da der Algorithmus nur auf die Blöcke $t \times s$ angewendet wird, für die $s \subset \tilde{\mathcal{N}}(t)$ gilt. Aus Lemma 4.3.8 und 4.13 folgt dann die Behauptung. \square

Der Vorteil dieser Partition ist der, dass die Anzahl an Blöcken nicht von η abhängt, sondern dass sie maximale Größe haben. Dies ist dann von Vorteil, wenn nur die Partition benötigt wird, z.B. um die Matrix zu speichern oder mit einem Vektor zu multiplizieren. Der Block τ wird dabei höchstwahrscheinlich kein Knoten in T_I sein. Das macht es schwierig, rekursive Block-Algorithmen zu definieren, welche auf einer rekursiven Unterteilung der Zeilen und Spalten beruhen. Wenn wir also solche Algorithmen, die auf einer hierarchischen Struktur von Block-Partitionen basieren, realisieren wollen, so brauchen wir eine Partition, die in Form von Blättern eines sogenannten Block-Cluster-Baumes konstruiert wurde.

4.3.3 Block-Cluster Bäume

Wir haben also gesehen, dass die hierarchische Partition und die Speicherung dieser als Cluster Baum die Anforderung an die Komplexität erfüllen, welche wir uns gesetzt hatten. Dabei enthielt jedes Level unseres Baumes T_I eine Partition von I . Unser Ziel ist es jedoch eine zulässige Partition von $I \times J$ anzugeben. Wir brauchen also einen Cluster-Baum $T_{I \times J}$ zur Indexmenge $I \times J$, welchen wir aufgrund der Struktur seiner Knoten *Block-Cluster-Baum* nennen werden.

Wir werden auch hier nicht die gesamte Menge an möglichen Partitionen betrachten, sondern nur die, welche durch sukzessive Unterteilung der Zeilen und Spalten entstehen. Seien T_I und T_J zwei Cluster Bäume für I und J mit dazu gehörigen Abbildungen S_I und S_J . Wir betrachten dann einen Block Cluster Baum, der durch die folgende Abbildung $S_{I \times J}$ bestimmt wird:

$$S_{I \times J}(t \times s) := \begin{cases} \emptyset, & \text{falls } t \times s \text{ zulässig ist oder } S_I(t) = \emptyset \text{ oder } S_J(s) = \emptyset, \\ S_I(t) \times S_J(s), & \text{sonst.} \end{cases} \quad (4.25)$$

Für die Tiefe des resultierenden Baumes gilt offensichtlich

$$L(T_{I \times J}) \leq \min\{L(T_I), L(T_J)\}. \quad (4.26)$$

Seine Blätter bilden eine zulässige Partition $P := \mathcal{L}(T_{I \times J})$ von $I \times J$. Jeder seiner Unterbäume mit Wurzel $(t \times s) \in T_{I \times J}$ ist ebenfalls ein Block-Cluster-Baum und wird mit $T_{t \times s}$ bezeichnet. Seine Blätter bilden eine zulässige Partition von $t \times s$.

Ein Maß für die Komplexität einer Partition ist die sogenannte Dünnheitskonstante. Bei hierarchischen Matrizen bezeichnet sie das Maximum an Blöcken, die zu einer bestimmten Zeile bzw. Spalte gehören.

Definition 4.3.10 [DÜNNHEITSKONSTANTE (SPARSITY CONSTANT)]

Seien T_I und T_J Cluster Bäume der Indexmengen I und J und sei $T_{I \times J}$ ein Block-Cluster-Baum zu $I \times J$. Für einen Cluster $t \in T_I$ bezeichnen wir die maximale Anzahl von Blöcken $t \times s \in T_{I \times J}$ mit

$$c_{sp}^r(T_{I \times J}, t) := |\{s \subset J : t \times s \in T_{I \times J}\}|.$$

Auf die gleiche Weise definieren wir für einen Cluster $s \in T_J$

$$c_{sp}^c(T_{I \times J}, s) := |\{t \subset I : t \times s \in T_{I \times J}\}|.$$

Die Dünnheitskonstante c_{sp} des Block-Cluster-Baums ist dann definiert als

$$c_{sp}(T_{I \times J}) := \max \left\{ \max_{t \in T_I} c_{sp}^r(T_{I \times J}, t), \max_{s \in T_J} c_{sp}^c(T_{I \times J}, s) \right\}.$$

Die meisten Algorithmen im Zusammenhang mit hierarchischen Matrizen können blockweise angewendet werden. In diesem Fall sind die Gesamtkosten des Algorithmus' die Summe über die Kosten der einzelnen Blöcke. Wir nehmen nun an, die Kosten pro Block sind durch eine Konstante $c > 0$ beschränkt. In diesem Fall gilt

$$\sum_{t \times s \in T_{I \times J}} c \leq c \sum_{t \in T} |\{s \subset J : t \times s \in T_{I \times J}\}| \leq cc_{sp}^r |T_I| \leq cc_{sp} |T_I|. \quad (4.27)$$

Wenn man die Rollen von t und s vertauscht, erhält man auf die gleiche Weise die Abschätzung $\sum_{t \times s \in T_{I \times J}} c \leq cc_{sp} |T_J|$. Wenn die Kosten des Algorithmus für jeden Block $t \times s \in P$ durch $c(|t| + |s|)$ beschränkt sind, dann können die Gesamtkosten abgeschätzt werden durch

$$\begin{aligned} \sum_{t \times s \in T_{I \times J}} c(|t| + |s|) &= c \sum_{t \in T'_I} \sum_{\{s \in T_J : t \times s \in T_{I \times J}\}} |t| + c \sum_{s \in T'_J} \sum_{\{t \in T_I : t \times s \in T_{I \times J}\}} |s| \\ &\leq cc_{sp} \left(\sum_{t \in T'_I} |t| + \sum_{s \in T'_J} |s| \right) \leq cc_{sp} L(T_{I \times J}) [|I| + |J|] \\ &\leq cc_{sp} [L(T_I) |I| + L(T_J) |J|] \end{aligned} \quad (4.28)$$

nach 4.13. T'_I und T'_J sind dabei Unterbäume von T_I bzw. T_J , die benötigt werden um $T_{I \times J}$ zu konstruieren, das heißt

$$T'_I := \{t \in T_I : \text{es gibt ein } t' \subset t \text{ und } s' \in T_J \text{ mit } t' \times s' \in T_{I \times J}\}.$$

Lemma 4.3.11

Seien T_I und T_J Cluster Bäume zu den Indermengen I und J . Für die Anzahl Blöcke in einer Partition $\mathcal{L}(T_{I \times J})$ und für die Anzahl an Knoten in $T_{I \times J}$ gilt:

$$|\mathcal{L}(T_{I \times J})| \leq |T_{I \times J}| \leq 2c_{sp} \min\{|\mathcal{L}(T_I)|, |\mathcal{L}(T_J)|\}. \quad (4.29)$$

Falls $|t| \geq n_{min}$ für alle $T \in T_I \cup T_J$, dann gilt $|\mathcal{L}(T_{I \times J})| \leq |T_{I \times J}| \leq 2c_{sp} \min\{|I|, |J|\} / n_{min}$.

Beweis. Wendet man (4.27) auf $|T_{I \times J}| = \sum_{t \times s \in T_{I \times J}} 1$ an, so erhält man

$$|T_{I \times J}| \leq c_{sp} \min\{|T_I|, |T_J|\}.$$

Mit Lemma 4.3.3 ergibt sich dann $|T_I| \leq 2|\mathcal{L}(T_I)|$ und $|T_J| \leq 2|\mathcal{L}(T_J)|$. □

Lemma 4.3.12

Seien T_I und T_J balancierte Cluster-Bäume. Unter Berücksichtigung der Bedingung (ii) an die Zulässigkeitsbedingung ist die Anzahl an benötigten Operationen für die Konstruktion eines Block-Cluster-Baumes von der Ordnung $c_{sp}(|I| \log |I| + |J| \log |J|)$.

Beweis. Mit (4.28) und Lemma 4.3.5. □

Die Zeit, eine zulässige Matrix-Partitionierung zu erstellen, kann im Vergleich zu den übrigen Berechnungen vernachlässigt werden. Unter Verwendung einer zulässigen Partition werden wir nun im nächsten Abschnitt die Menge der hierarchischen Matrizen definieren.

4.4 Hierarchische Matrizen

In diesem Abschnitt seien T_I und T_J wieder binäre Cluster-Bäume zu den Indexmengen I bzw. J . Weiterhin wird angenommen, dass der Block-Cluster-Baum $T_{I \times J}$ unter Verwendung einer Zulässigkeitsbedingung generiert wurde, die den Bedingungen aus dem vorangegangenen Abschnitt genügen. Wir werden hier die Menge der hierarchischen Matrizen auf die gleiche Weise definieren, wie Hackbusch dies schon in [Hac99] getan hat, siehe auch [Hac07]. Die Elemente dieser Menge können mit logarithmisch-linearer Komplexität gespeichert werden und bieten eine effiziente Repräsentation von vollbesetzten Matrizen. Darüberhinaus kann durch Kombination der hierarchischen Partition und die Struktur von Niedrigrangmatrizen eine approximative Algebra definiert werden, welche auf divide-and-conquer Versionen der üblichen Block-Operationen beruhen. Die daraus entstehenden effizienten Versionen von Matrix-Addition und Multiplikation können verwendet werden, um schnelle Varianten von höheren Matrixoperationen wie Inversion oder LR-Zerlegung zu definieren. Näheres zu diesen anspruchsvollen Operationen findet man in [Beb08]. Wir werden uns in diesem Abschnitt auf die Matrix-Vektor-Multiplikation und die Invertierung beschränken, da diese bei der Lösung des Integralteils unserer Gleichung verwendet wird.

4.4.1 Die Menge der hierarchischen Matrizen

Definition 4.4.1 [HIERARCHISCHE MATRIZEN]

Die Menge der hierarchischen Matrizen auf einem Block Cluster Baum $T_{I \times J}$ zu einer zulässigen Partition $P := \mathcal{L}(T_{I \times J})$ und blockweisem Rang k ist definiert als

$$\mathcal{H}(T_{I \times J}, k) := \{A \in \mathbb{C}^{I \times J} : \text{Rang} A_b \leq k \text{ für alle zulässigen } b \in P\}.$$

Der Einfachheit halber werden Elemente aus $\mathcal{H}(T_{I \times J}, k)$ kurz \mathcal{H} -Matrizen genannt.

Für eine effiziente Behandlung der zulässigen Blöcke sollte die äußere Produktform aus Abschnitt 3.3.1 als Darstellung verwendet werden. Darüberhinaus sollte man nicht den maximalen Rang k der Darstellung verwenden, sondern den tatsächlichen Rang des Blocks als Anzahl von Zeilen und Spalten verwenden. Die komponentenweise Abspeicherung von nicht zulässigen Blöcken wird die Effizienz der Darstellung erhöhen.

Lemma 4.4.2

Sei $A \in \mathcal{H}(T_{I \times J}, k)$. Dann gilt :

1. Jede Teilmatrix A_b $b \in T_{I \times J}$ gehört zu $\mathcal{H}(T_b, k)$.
2. Die Transponierte A^T und die Hermitsche A^H gehören zu $\mathcal{H}(T_{J \times I}, k)$, falls die Zulässigkeitsbedingung symmetrisch ist, das heißt $s \times t$ ist zulässig genau dann, wenn $t \times s$ zulässig ist.

Wir haben schon zu Beginn dieses Kapitels gesehen, dass der Speicherbedarf eines zulässigen Blocks $b = t \times s \in \mathcal{L}(T_{I \times J})$ einer Matrix $A \in \mathcal{H}(T_{I \times J}, k)$

$$N_{st}(A_b) = k(|t| + |s|)$$

beträgt. Ein nicht zulässiger Block A_b , $b \in P$ wird komponentenweise abgespeichert, was zu einem Speicherbedarf von $|t||s|$ Einheiten führt. Da $\min\{|t|, |s|\} \leq n_{\min}$ (siehe Bemerkung 4.3.2) erhalten wir

$$|t||s| = \min\{|t|, |s|\} \max\{|t|, |s|\} \leq n_{\min}(|t| + |s|). \quad (4.30)$$

Wir benötigen also höchstens $\max\{k, n_{\min}\}(|t| + |s|)$ Speichereinheiten, um einen Block A_b , $b \in P$ abzuspeichern. Unter Verwendung von (4.28) erhalten wir folgenden Satz.

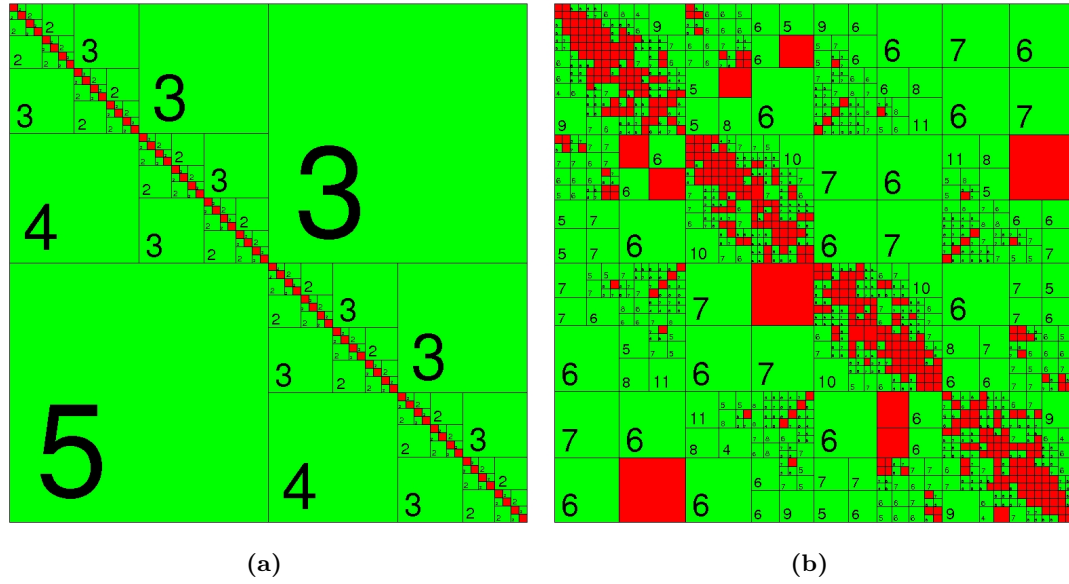


Abbildung 4.3: \mathcal{H} -Matrix mit blockweisem Rang für ein (a) eindimensionales Integral und (b) zweidimensionales Integral

Satz 4.4.3

Sei c_{sp} die Dünnekonstante der Partition P . Die Anzahl an Speichereinheiten N_{st} für $A \in \mathcal{H}(T_{I \times J}, k)$ sind beschränkt durch

$$N_{st}(A) \leq c_{sp} \max\{k, n_{min}\} (L(T_I)|I| + L(T_J)|J|).$$

Falls T_I und T_J balancierte Cluster Bäume sind, dann gilt

$$N_{st}(A) \sim \max\{k, n_{min}\} (|I| \log |I| + |J| \log |J|).$$

Bemerkung: 4.4.4: Obwohl das Konzept der \mathcal{H} -Matrix*Approximation eher auf vollbesetzte Matrizen abzielt, sind dünnbesetzte Matrizen A , die auf zulässigen Blöcken verschwinden, ebenfalls in $\mathcal{H}(T_{I \times I}, n_{min})$. Da die Größe der Cluster, die zu einem der nicht zulässigen Blöcke gehören, gleich oder niedriger als n_{min} ist, ist der Rang der Blöcke A_b nicht größer als n_{min} .

4.4.2 Matrix-Vektor-Multiplikation

Um das Integral aus unserer PIDE mittels des vorgestellten Kollokationsverfahrens berechnen zu können, muss eine Matrix-Vektor-Multiplikation (3.34) durchgeführt werden. Das Verfahren, das bei der \mathcal{H} -Matrix-Vektor-Multiplikation verwendet wird und dessen Aufwand werden wir hier kurz vorstellen. Die Kosten, um eine \mathcal{H} -Matrix $A \in \mathcal{H}(T_{I \times J}, k)$ oder deren hermitesche Transponierte A^H mit einem Vektor x zu multiplizieren, entstehen aus den Kosten der blockweisen Matrix-Vektor-Multiplikation:

$$Ax = \sum_{t \times s \in P} A_{ts} x_s \quad \text{und} \quad A^H x = \sum_{t \times s \in P} (A_{ts})_H x_t. \quad (4.31)$$

Da jeder zulässige Block $t \times s$ in der äußeren Produktform $A_{ts} = UV^H$, $U \in \mathbb{C}^{t \times k}$, $V \in \mathbb{C}^{k \times s}$ vorliegt, sind höchstens $2k(|t| + |s|)$ Operationen nötig, um das Matrix-Vektor-Produkt

$A_{ts}x_s = UV^H x_s$ und $(A_{ts})^H x_s = VU^H x_s$ zu bestimmen. Wenn $t \times s$ nicht zulässig sein sollte, ist A_{ts} komponentenweise abgespeichert und es gilt $\min\{|t|, |s|\} \leq n_{min}$. Wie in (4.30) sieht man, dass in diesem Fall

$$2|t||s| \leq 2n_{min}(|t| + |s|) \quad (4.32)$$

Operationen nötig sind. Mit derselben Argumentation wie in Satz 4.4.3 erhalten wir folgenden Satz.

Satz 4.4.5

Sei $A_{ts}x_s = UV^H x_s$ und $c \in \mathbb{C}^J$. Dann gilt für die Anzahl N_{MV} die für die Ausführung der Multiplikation Ax benötigt wird, folgende Schranke

$$N_{MV} \leq 2c_{sp} \max\{k, n_{min}\}(L(T_I)|I| + L(T_J)|J|). \quad (4.33)$$

Sind T_I und T_J balancierte Cluster-Bäume, so gilt

$$N_{MV} \sim \max\{k, n_{min}\}(|I| \log |I| + |J| \log |J|). \quad (4.34)$$

Bemerkung: 4.4.6: Der Aufwand, der zur Berechnung des Matrix-Vektor-Produkts benötigt wird, ist allerdings vom maximalen Rang k der Blöcke abhängig. Nur wenn wir davon ausgehen, dass k nur logarithmisch von der Anzahl der Freiheitsgrade abhängt, können wir einen logarithmisch-linearen Gesamtaufwand für die Multiplikation erwarten. Diese Abhängigkeit beruht allerdings auf der Art des zu lösenden Problems. Wir werden im nächsten Abschnitt, welcher die numerischen Ergebnisse enthält, sehen, dass dies in unserem Fall zutrifft.

Die Matrix-Vektor-Multiplikation ist exakt bis auf Maschinengenauigkeit. Dies gilt nicht für andere Matrixoperationen wie die LR-Zerlegung oder Inversion, welche in [Beb08] vorgestellt werden. Da wir im weiteren Verlauf nicht nur die Multiplikation, sondern ebenso die Inversion einer \mathcal{H} -Matrix benötigen werden, wird auch diese hier vorgestellt. Weitere Anwendungsmöglichkeiten wie die parallel Matrix-Vektor-Multiplikation, die Berechnung der LR-Zerlegung von \mathcal{H} -Matrizen oder die Matrix-Matrix-Multiplikation sind in [Beb08] zu finden.

4.4.3 Invertierung von \mathcal{H} -Matrizen

Wir nehmen nun an, dass jeder Block A_{tt} , $t \in T_I$, von $A \in \mathcal{H}(T_{(i \times j)}, k)$ invertierbar ist. Der folgende divide-and-conquer Algorithmus aus [Hac99] nutzt die Blockunterteilung

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

von A entsprechend seines Block-Cluster-Baums aus. Für die exakte Inverse von A gilt

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}S^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S^{-1} \\ -S^{-1}A_{21}A_{11}^{-1} & S^{-1} \end{bmatrix}, \quad (4.35)$$

wobei $S := A_{22} - A_{21}A_{11}^{-1}A_{12}$ das Schur-Komplement von A_{11} in A bezeichnet. Für die Berechnung der Inversen von A müssen also die Matrizen A_{11} und S invertiert werden, diese sind aber nur ungefähr halb so groß wie die ursprüngliche Matrix A . Bei der Berechnung der \mathcal{H} -Matrix-Inversen C von A werden die in (4.35) auftretenden Multiplikationen und Additionen durch die entsprechenden \mathcal{H} -Matrix Versionen der Operationen ersetzt.

```

procedure invertH( $t, A, var C$ )
  wenn  $t \in \mathcal{L}(T_I)$  emphdann ist  $C_{tt} := A_{tt}^{-1}$  die Inverse
  ansonsten
  seien  $t_1, t_2$  die Söhne von  $t$ 
  invertH( $t_1, A, C$ )
   $T_{t_1t_2} = T_{t_1t_2} - C_{t_1t_1}A_{t_1t_2}$ 
   $T_{t_2t_1} = T_{t_2t_1} - A_{t_2t_1}C_{t_1t_1}$ 
   $A_{t_2t_2} = A_{t_2t_2} - A_{t_2t_1}T_{t_1t_2}$ 
  invertH( $t_2, A, C$ )
   $C_{t_1t_2} = C_{t_1t_2} - T_{t_1t_2}C_{t_2t_2}$ 
   $C_{t_2t_1} = C_{t_2t_1} - C_{t_2t_2}T_{t_2t_1}$ 
   $C_{t_1t_1} = C_{t_1t_1} - T_{t_1t_2}C_{t_2t_1}$ 

```

Algorithmus 4: \mathcal{H} -Matrix Inversion

Darüberhinaus benötigt man eine temporäre Matrix $T \in \mathcal{H}(T_{(I \times I)}, k)$, die zusammen mit C initialisiert wird. Ruft man nun `invertH` mit den Parametern I , A und C , dann erhält man die gewünschte Approximation C der Inversen A^{-1} , wobei A zerstört wird.

Die Komplexität der Berechnung der Inversen richtet sich nun nach den Kosten der \mathcal{H} -Matrix-Multiplikation. Für den folgenden Satz nehmen wir an, dass jede Summe von zwei Matrizen mit Rang k auf eine Matrix mit Rang k gerundet wird. Es sei hier bemerkt, dass diese Annahme im Einzelfall überprüft werden muss, da sonst der Approximationsfehler unkontrollierbar würde.

Satz 4.4.7

Die Berechnung der \mathcal{H} -Matrix Inversen $C \in \mathcal{H}(T_{I \times I}, k)$ von $A \in \mathcal{H}(T_{I \times I}, k)$ benötigt unter Verwendung von Algorithmus 4 $\mathcal{O}(k^2|I|(\log |I|)^2)$ Operationen.

Mit der Matrix-Vektor-Multiplikation und der Invertierung von \mathcal{H} -Matrizen ist es nun möglich die PIDEs, die bei der Optionspreisbewertung entstehen effizient zu lösen. Im nächsten Kapitel werden wir uns mit numerischen Ergebnissen dazu befassen. Wir werden darin die Konvergenz und den Aufwand der klassischen Behandlung der Gleichungen mit der \mathcal{H} -Matrix Approximation vergleichen.

Kapitel 5

Numerische Ergebnisse

In diesem Kapitel werden wir die numerischen Ergebnisse der in den vorangegangenen Kapiteln besprochenen Verfahren vorstellen. Die Ergebnisse, die mittels \mathcal{H} -Matrix Approximation berechnet wurden, sind unter Verwendung des Softwarepakets *AHMED* (another software library on hierarchical matrices for elliptic differential equations) von Prof. Dr. Bebendorf entstanden.

Zu Beginn werden wir uns mit der Konvergenz und der benötigten Rechenzeit der beiden Verfahren, also der direkten Behandlung des Integrals und der \mathcal{H} -Matrix Approximation, gegen die geschlossene Lösung des Merton-Modells befassen. In späteren Experimenten werden wir zur Überprüfung der Konvergenz die Lösung der direkten Behandlung auf einem feineren Gitter verwenden.

Des Weiteren werden wir die Konvergenzgeschwindigkeit im Kou-, VG- und CGMY-Modell untersuchen. Die Untersuchung erfolgt in allen Modellen für die Bewertung von Europäischen und Amerikanischen Optionen. Im zweiten Abschnitt des Kapitels befassen wir uns dann mit der Bewertung von Basket Optionen im zweidimensionalen Merton- und Copula-Modell.

Insgesamt werden wir die benötigte Rechenzeit und die erreichte Genauigkeit, sowie den Speicherbedarf der beiden Verfahren vergleichen. Darüberhinaus werden wir die Auswirkung der Parameterwahl in der Zulässigkeitsbedingung der \mathcal{H} -Matrix Approximation auf die Konvergenzgeschwindigkeit untersuchen.

5.1 Vorbemerkungen

Wir werden in diesem Kapitel die Konvergenz, also das Verhältnis von Aufwand zu Genauigkeit der beiden Verfahren untersuchen. Der Aufwand wird dabei auf zwei unterschiedliche Arten gemessen. Die erste Variante verwendet die Anzahl an Diskretisierungspunkten M des Gebiets Ω als ein Maß für den Aufwand. Da in diesem Fall die Konvergenzgeschwindigkeit nicht von der Wahl der Darstellung der Kollokationsmatrix abhängt, sollten beide Verfahren ähnliche Werte liefern. Die zweite Möglichkeit, die wir hier betrachten werden ist das Messen der Rechenzeit als Maß des Aufwands. Bei der Zeitmessung innerhalb der Berechnung fließen folgende Faktoren mit ein :

$$\begin{aligned} & \text{Aufstellen der Matrizen (Gleichungssystem+Kollokation)} \\ + & \text{ Lösen des Gleichungssystems} \\ + & \text{ Berechnung des Integrals (Matrix-Vektor-Multiplikation)} \\ \hline = & \text{ Gesamtrechenzeit } \Theta \end{aligned}$$

Wir werden die Konvergenz der Verfahren bezüglich dieser Gesamtrechenzeit betrachten. Die Genauigkeit der Lösungen der PIDE werden über den relativen Fehler e_{rel} in der diskreten L_2 -Norm angegeben. Sei dazu M_{max} die Anzahl der Ortsdiskretisierungspunkte der Referenzlösung V_{Ref} . Die numerische Lösung der PIDE V_{PIDE} , die auf einem größeren

Gitter berechnet wurde, wird auf das feine Gitter mit M_{max} Punkten interpoliert. Der Fehler zwischen dem Optionspreis und der Referenzlösung ergibt sich dann wie folgt:

$$\begin{aligned} \|V_{PIDE} - V_{Ref}\|_{L_2} &\approx \frac{1}{\sqrt{M_{max}}} \|V_{PIDE} - V_{Ref}\|_{l_2} \\ &= \frac{1}{\sqrt{M_{max}}} \left(\sum_{i=1}^{M_{max}} (V_{PIDE}^i - V_{Ref}^i)^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (5.1)$$

Dabei ist $V^i = V(S_i, 0) = V(e^{x_i}, 0)$ der Optionswert bei entsprechendem Startkurs S_i des zugrunde liegenden Wertpapiers. Der relative Fehler e_{rel} ergibt sich nun durch

$$\begin{aligned} e_{rel}^l &= \frac{\|V_{PIDE_l} - V_{Ref}\|_{L_2}}{\|V_{Ref}\|_{L_2}} \\ &\approx \frac{\frac{1}{\sqrt{M_{max}}} \|V_{PIDE_l} - V_{Ref}\|_{l_2}}{\frac{1}{\sqrt{M_{max}}} \|V_{Ref}\|_{l_2}} \\ &= \frac{\left(\sum_{i=1}^{M_{max}} (V_{PIDE_l}^i - V_{Ref}^i)^2 \right)^{\frac{1}{2}}}{\left(\sum_{i=1}^{M_{max}} (V_{Ref}^i)^2 \right)^{\frac{1}{2}}}. \end{aligned} \quad (5.2)$$

Dabei gibt l die Stufe der Verfeinerung an, unser Gitter hat dann die Maschenweite $\frac{x_{max} - x_{min}}{2^l}$. Da wir die Konvergenzraten bezüglich der Rechenzeit Θ und der Anzahl an Ortsdiskretisierungspunkten M bestimmen wollen, benötigen wir noch ein Maß für die Veränderung der Genauigkeit bezüglich der Veränderung des Aufwandes. Wir definieren dazu die Konvergenzrate der Stufe l mit

$$\rho_{\Theta}^l = \frac{\log(e_{rel}^{l-1}/e_{rel}^l)}{\log(\Theta^l/\Theta^{l-1})}, \quad (5.3)$$

wenn wir die Konvergenz bezüglich der Rechenzeit messen und mit

$$\rho_M^l = \frac{\log(e_{rel}^{l-1}/e_{rel}^l)}{\log(M^l/M^{l-1})}, \quad (5.4)$$

wenn wir die Konvergenz bezüglich der Anzahl an Gitterpunkten messen. Dabei ist M^l gleich der Anzahl an Ortsdiskretisierungspunkten auf der Verfeinerungsstufe l .

Da die Dichten der Sprungkomponenten sehr schnell abfallen, ist es ausreichend das Integrationsgebiet auf das Intervall $[-4, 4]$ zu beschränken. Das Gebiet $\Omega = \mathbb{R}$ wird dann entsprechend auf das Gebiet $\Omega_e = [\log(K) - 4, \log(K) + 4]$ beschränkt. Die Beschränkung des Gebiets ist von K abhängig, da in [Rei06] gezeigt wurde, dass eine Verschiebung des Gebiets zugunsten von K zu einem stabileren Verfahren führt. Die Untersuchung der Konvergenz gegen die analytische Lösung im Merton-Modell wird zeigen, dass diese Einschränkungen keinen Einfluss auf die Konvergenz haben.

Die Parameter, die bei der Lösung der PIDEs verwendet werden können den entsprechenden Tabellen in den einzelnen Abschnitten dieses Kapitels entnommen werden. Die Konstruktion der Block-Cluster-Bäume mittels des $\mathcal{A}\mathcal{H}\mathcal{M}\mathcal{E}\mathcal{D}$ -Softwarepakets geschieht über die Zulässigkeitsbedingung

$$\text{dist}(X_t, X_s)_{\eta_{\mathcal{H}}} \geq \min\{\text{diam}(X_t), \text{diam}(X_s)\}, \quad (5.5)$$

wobei $t, s \in T_I$ Blöcke aus dem Cluster-Baum sind. Die *dist*-Funktion berechnet hier den Abstand der Mittelpunkte der Cluster X_t und X_s . Wir werden später auch die Auswirkung des Parameters $\eta_{\mathcal{H}}$ auf die Gesamtrechenzeit untersuchen. Grundsätzlich gilt: alle Werte und Graphen, die mit dem Index \mathcal{H} versehen sind entstanden durch Verwendung des Approximationsverfahrens, alle Werte denen dieser Index fehlt, stammen aus dem klassischen Verfahren.

Parameter der \mathcal{H} -Matrix-Approximation

Im folgenden sind die Parameter der \mathcal{H} -Matrix-Approximation so gewählt worden, dass eine quadratische Konvergenz des Gesamtverfahrens sichergestellt ist. Der maximale Rang der Blöcke der \mathcal{H} -Matrix ist auf 20 beschränkt und der relative Fehler der Approximation so gewählt, dass er die Genauigkeit der Approximation nicht beeinträchtigt. Wir zeigen in den numerischen Experimenten, dass diese Parameterwahl den Aufwand des Gesamtverfahrens ausreichend verringert.

5.2 Bewertung von Plain-Vanilla-Optionen

5.2.1 Merton-Modell

Da das Merton-Modell für Europäische Optionen das einzige Sprung-Diffusions-Modell ist, für das eine geschlossene Lösung existiert, werden wir unsere Modelle zunächst in diesem Modell testen und die Konvergenz gegen die geschlossene Lösung untersuchen. Die hier angegebenen Konvergenzraten und Graphen beziehen sich auf eine Europäische Put-Option mit den Parametern aus Tabelle 5.1. Die Parameter stammen aus [dFL04] und spiegeln die Kurse eines Wertpapiers aus dem *S&P 500* Index wieder. σ_J und μ_J bezeichnen dabei die Varianz und den Erwartungswert des Sprungprozesses.

Parameter	Wert
S	100
K	100
T	0.25
r	0.1
μ_J	-0.9
σ_J	0.45
λ	0.1
$\eta_{\mathcal{H}}$	0.4

Tabelle 5.1: Parameter im Merton-Modell

Wir wählen hier $\eta_{\mathcal{H}}$ sehr klein. Das hat den Vorteil, dass ein breites Gebiet um die Diagonale der Kollokationsmatrix exakt dargestellt wird. Ein weiterer Vorteil ist, dass am Rand der Matrix, also auch am Rand des Integrationsgebiets, kleinere Blöcke auftreten (s. Abb 5.1). In dieser Abbildung sehen wir die Unterteilung und den Rang der Blöcke der \mathcal{H} -Matrix, die bei der Approximation der Kollokationsmatrix des Merton-Modells entsteht.

Der Vorteil der kleineren Blöcke ist, dass man den Aufbau der \mathcal{H} -Matrix besser steuern kann. Dies ist im Merton Modell wichtig, da die Dichte keinen symmetrischen Aufbau bezüglich der Diagonalen der Kollokationsmatrix besitzt. Der Peak der Dichte ist abhängig vom Erwartungswert nach links oder rechts verschoben, s. Abb. 5.2.

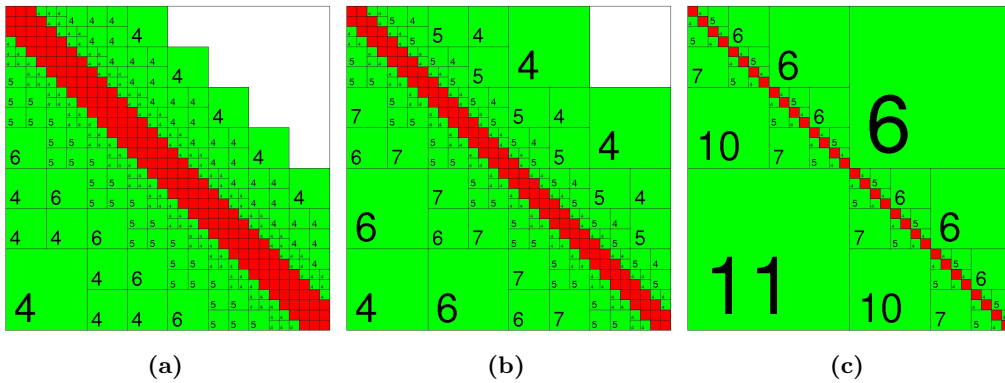


Abbildung 5.1: \mathcal{H} -Matrix für a) $\eta = 0.4$ b) $\eta = 0.8$ c) $\eta = 1.1$

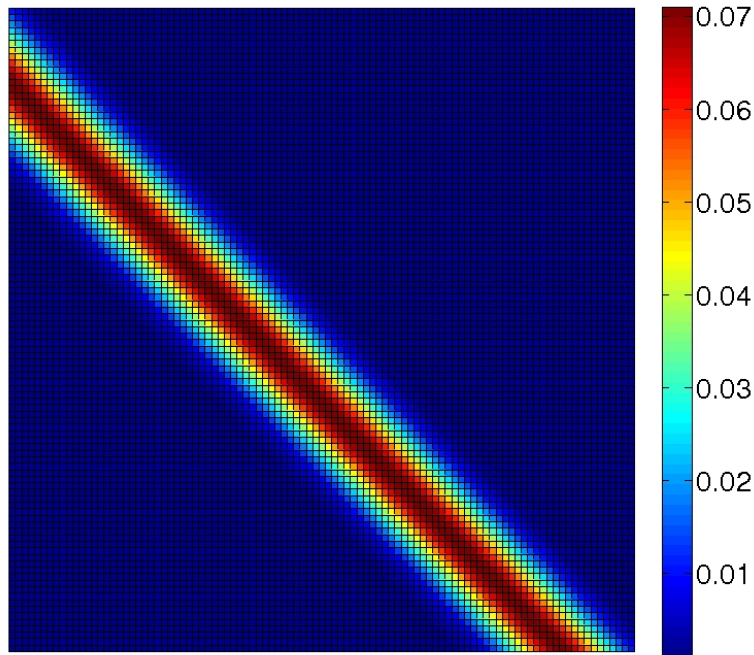


Abbildung 5.2: Einträge der Kollokationsmatrix im Merton-Modell

Dieser Aufbau wird von der Zulässigkeitsbedingung nicht ohne Weiteres berücksichtigt. Wir wählen daher $\eta_{\mathcal{H}} = 0.4$ und fügen eine geometrische Bedingung in den Aufbau der \mathcal{H} -Matrix ein. Diese führt dazu, dass Blöcke, deren Cluster X_t und X_s einen Mindestabstand voneinander aufweisen, von vornherein auf Null gesetzt und somit nicht behandelt werden. Der Mindestabstand der Cluster ist hier gleichbedeutend mit der Tatsache, dass der Block einem Gebiet entspricht, das am Rand des Integrationsgebiets liegt. Da die Dichte schnell abfällt, können diese Gebiete bei der Berechnung des Integrals vernachlässigt werden. Wir

wählen bei unseren Experimenten die Bedingung

$$\text{dist}(X_t, X_s) < 5, \quad (5.6)$$

um einen linearen Aufwand des Gesamtverfahrens gewährleisten zu können. Dabei ist $X_t \times X_s \in T_{I \times J}$ ein Block der \mathcal{H} -Matrix. Wir erhalten durch diese Methode die in Abb. 5.3 dargestellten Konvergenzplots.

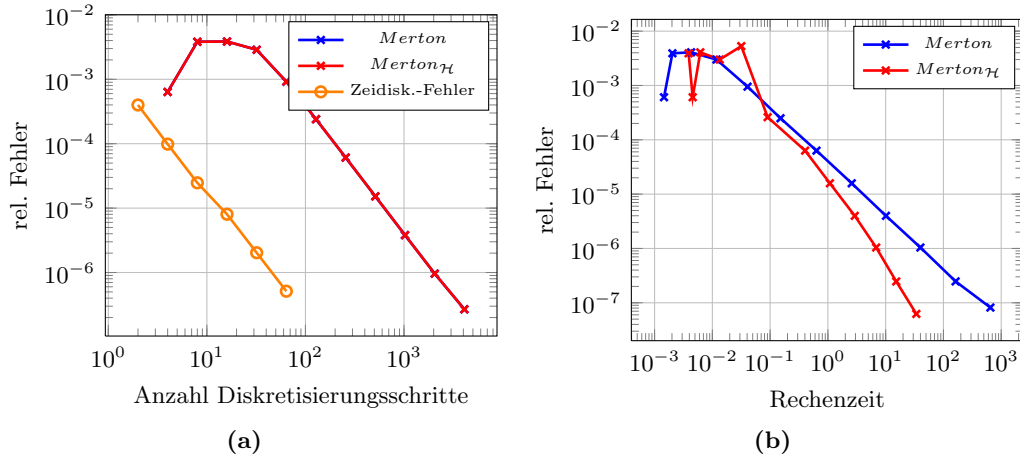


Abbildung 5.3: Konvergenz gegen die geschlossene Lösung, dabei wurde der relative Fehler gegen a) Aufwand in Diskretisierungspunkten b) Aufwand in Rechenzeit aufgetragen

Zusätzlich zum Fehler der Ortsdiskretisierung ist in Abb. 5.3 (a) der Fehler der Zeitdiskretisierung aufgetragen. Dieser liegt deutlich unter dem Fehler der Ortsdiskretisierung und konvergiert mit einer Rate von 2. Diese Konvergenzrate wird durch die Verwendung des Crank-Nicolson-Verfahrens erreicht und entspricht daher den Erwartungen an das Verfahren. Da der Zeitfehler auch in den weiteren Verfahren sehr klein ist und durch die Approximation nicht beeinflusst wird, verzichten wir von nun auf eine explizite Untersuchung dieses Fehlers.

Betrachten wir die Konvergenzgraphen des Ortsfehlers in Abb. 5.3 (a), so wird deutlich, dass die Approximation die Genauigkeit des Verfahrens nicht beeinflusst. Die Konvergenzrate bezüglich der Ortsverfeinerung liegt bei 2. Dies entspricht der Konvergenzordnung der verwendeten Differenzenquotienten. Die Graphen in Abb 5.3 (b) machen den Vorteil der Approximation deutlich. Der Graph der klassischen Methode hat eine Konvergenzrate von 1, da der Rechenaufwand durch die Matrix-Vektor-Multiplikation quadratisch ansteigt. Wird die Kollokationsmatrix jedoch durch eine \mathcal{H} -Matrix approximiert, verringert sich dieser Aufwand und wir erzielen eine wesentlich bessere Konvergenzgeschwindigkeit. Die genauen Werte für den relativen Fehler und die Konvergenzraten sind in den Tabellen 5.2 und 5.3 ausführlich dargestellt. Wir sehen, dass die Konvergenzrate des klassischen Verfahrens bezüglich der Rechenzeit einen Wert von ca. 1 erreicht. Die Rate bei Verwendung der \mathcal{H} -Matrix-Approximation hingegen steigt bis auf einen Wert von 2 an. Wir sehen also, dass durch die Approximation eine Lösung der PIDE mit optimalem Aufwand zu berechnen ist.

Wir wollen nun die Auswirkung der Wahl von $\eta_{\mathcal{H}}$ betrachten. Wir haben dazu die Approximation sowohl mit $\eta_{\mathcal{H}} = 0.4$ als auch mit $\eta_{\mathcal{H}} = 1.1$ durchgeführt und in Abb. 5.4 (a) die Ergebnisse verglichen. Wählt man den Parameter $\eta_{\mathcal{H}} = 1.1$, so ist es nicht mehr möglich, die Randgebiete der Dichte unberücksichtigt zu lassen, da die Größe der Blöcke

M	Θ	e_{rel}	ρ_M	ρ_Θ
128	$1.5080 \cdot 10^{-1}$	$2.4906 \cdot 10^{-4}$	1.93	1.02
256	$5.9264 \cdot 10^{-1}$	$6.3198 \cdot 10^{-5}$	1.98	1.00
512	$2.6542 \cdot 10^0$	$1.5823 \cdot 10^{-5}$	2.00	0.92
1,024	$1.0553 \cdot 10^1$	$3.9847 \cdot 10^{-6}$	1.99	1.00
2,048	$4.1891 \cdot 10^1$	$1.0414 \cdot 10^{-6}$	1.94	0.97
4,096	$1.7080 \cdot 10^2$	$2.4739 \cdot 10^{-7}$	2.07	1.02
8,192	$6.9089 \cdot 10^2$	$6.2457 \cdot 10^{-8}$	1.99	0.98

Tabelle 5.2: Konvergenz ohne Approximation

M	$\Theta^{\mathcal{H}}$	$e_{rel}^{\mathcal{H}}$	$\rho_M^{\mathcal{H}}$	$\rho_\Theta^{\mathcal{H}}$
256	$3.9305 \cdot 10^{-1}$	$6.3198 \cdot 10^{-5}$	2.04	1.02
512	$1.0694 \cdot 10^0$	$1.5823 \cdot 10^{-5}$	2.00	1.38
1,024	$2.9440 \cdot 10^0$	$3.9843 \cdot 10^{-6}$	1.99	1.36
2,048	$6.7370 \cdot 10^0$	$1.0411 \cdot 10^{-6}$	1.94	1.62
4,096	$1.5988 \cdot 10^1$	$2.4692 \cdot 10^{-7}$	2.08	1.67
8,192	$3.3400 \cdot 10^1$	$6.2408 \cdot 10^{-8}$	1.98	1.87

Tabelle 5.3: Konvergenz mit \mathcal{H} -Matrix Approximation

dies verhindert. Verwendet man diesen Ansatz, so ist die Berechnung der Lösung auf sehr groben Gittern zwar schneller möglich, die Konvergenzrate bezüglich der Rechenzeit nimmt jedoch mit feineren Gittern immer weiter ab. Die anfangs schnellere Bewertung rührt daher, dass auf der Diagonalen weniger vollbesetzte Blöcke berücksichtigt werden müssen. Steigt jedoch die Anzahl der Einträge an, so sorgen die in immer größerer Zahl auftretenden Nulleinträge der Matrix dafür, dass die Komplexität des Algorithmus' auf das Niveau des klassischen Verfahrens abfällt, s. Abb 5.4 (a). Die Verminderung der Rate resultiert aus dem prozentualen Anteil an ausgewerteten Matrixeinträgen. Die Nulleinträge am Rand der Matrix führen dazu, dass mehr Einträge bei der Approximation berücksichtigt werden, s. Abb 5.4 (b). Wir sehen hier deutlich, dass der prozentuale Anteil an ausgewerteten Einträgen bei schlechter Wahl von $\eta_{\mathcal{H}}$ ab einem gewissen Punkt gleichbleibt. Dies ist gleichbedeutend mit einem quadratischen Anwachsen des Aufwands von diesem Punkt an, da für eine lineare Wachstumsrate eine weitere Verminderung des ausgewerteten Anteils nötig wäre. Den optimalen Verlauf sehen wir bei der Wahl $\eta_{\mathcal{H}} = 0.4$ und gleichzeitiger Anwendung der geometrischen Abstandsbedingung.

Wir können an diesem ersten Beispiel sehen, dass eine Verwendung der \mathcal{H} -Matrizen die Rechengeschwindigkeit stark erhöht. Bei ungeschickter Nutzung führt die Approximation allerdings bei sehr feinen Gittern nur noch zu einer Verbesserung in der Konstanten und nicht in der Konvergenzrate selbst. Nutzt man jedoch die Struktur der Sprungdichte aus und überträgt diese in den Konstruktionsprozess der Approximation der Kollokationsmatrix, so ist eine deutlicher Verbesserung in der Konvergenzrate festzustellen.

Speicherbedarf

Wir werden nun exemplarisch für das Merton-Modell den Speicherbedarf für die Darstellung der Kollokationsmatrix in den beiden Verfahren vergleichen. Bei direkter Ausführung der Matrix-Vektor-Multiplikation wird die Matrix explizit aufgestellt. Dies führt zu

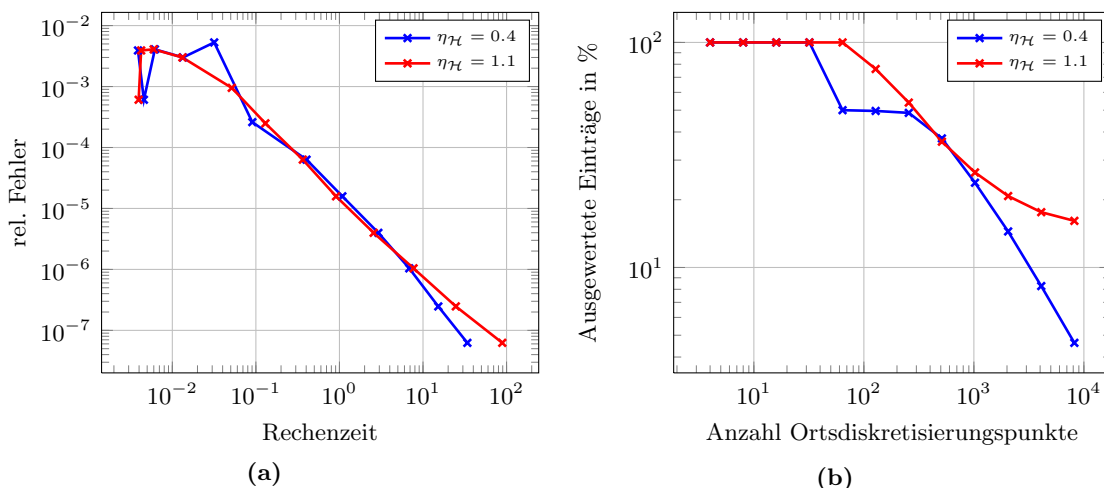


Abbildung 5.4: a) Vergleich der Rechenzeit bei $\eta_{\mathcal{H}} = 0.4$ und $\eta_{\mathcal{H}} = 1.1$ b) Vergleich der ausgewerteten Matrixeinträge in Prozent

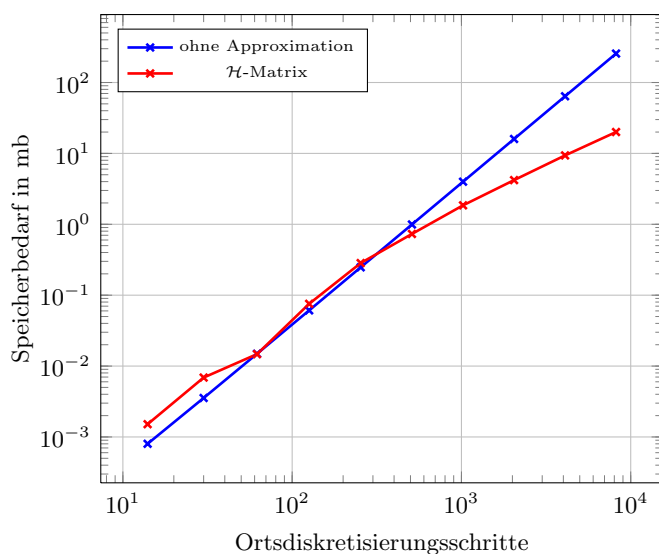


Abbildung 5.5: Speicherbedarf der Kollokationsmatrix im Mertonmodell ($\eta_{\mathcal{H}} = 0.4$)

einem Speicherbedarf in der Größenordnung $\mathcal{O}(N^2)$, wobei N die Anzahl an Gitterpunkten ist. In Abb 5.5 sehen wir, dass die Verwendung der \mathcal{H} -Matrizen zu einem geringeren Speicherbedarf führt. Bei kleinen Matrizen, also wenigen Gitterpunkten, verbraucht die Approximation noch mehr Speicher als die direkte Darstellung, da der Speicherbedarf der algebraischen Struktur über dem der eigentlichen Matrix liegt. Die Steigerung des Bedarfs nimmt jedoch bei steigender Matrixgröße ab und liegt schließlich bis auf logarithmische Faktoren bei einer Größenordnung von $\mathcal{O}(N)$.

5.2.2 Kou-, VG- und CGMY-Modell

Wir kommen nun zu der Bewertung von Plain-Vanilla-Optionen im Kou-, VG- und CGMY-Modell. Wir untersuchen in diesem Kapitel die Konvergenz der Lösung der PIDEs, wie sie bei der Bewertung von Plain-Vanilla-Put-Optionen entstehen. Die verwendeten Para-

meter sind in den Tabellen 5.4 und 5.5 aufgelistet. Die für das Kou-Modell verwendeten stammen aus [KW04]. Die Parameter für das VG- und CGMY-Modell sind der Arbeit [MvPS04] von Matache, von Petersdorff und Schwab entnommen. In beiden Arbeiten werden mit diesen Parametern ebenfalls Europäische Plain-Vanilla-Optionen bewertet. Die Referenzwerte zur Bestimmung der Konvergenzraten wurden auf einem Gitter mit 8192 Ortsdiskretisierungspunkten und 500 Zeitschritten berechnet.

Parameter	Wert
η_1	3.0465
η_2	3.0775
p	0.3445
$\eta\mathcal{H}$	0.4

Tabelle 5.4: Parameter im Kou-Modell

Parameter	Wert
C	1
G	1.8
M	2.5
Y	0 (VG-Modell)
	0.2 (CGMY-Modell)
ϵ	0.001
$\eta\mathcal{H}$	0.4

Tabelle 5.5: Parameter im VG- und CGMY-Modell

In den hier verwendeten Modellen weisen die Dichten der Sprungprozesse eine Struktur auf, die von der Zulässigkeitsbedingung der \mathcal{H} -Matrix-Approximation sehr gut behandelt werden kann. Alle Dichten besitzen eine Singularität im Nullpunkt; dies entspricht einer Singularität auf der Diagonalen der Kollokationsmatrix, s. Abb. 5.6.

Auf Grund dieser Struktur ist es im Kou, VG- und CGMY-Modell nicht nötig, bei der Approximation auf eine geometrische Bedingung zurückzugreifen. In Abb. 5.7 sind die Konvergenzraten bezüglich der Rechenzeit für das Kou-, VG- und CGMY-Modell aufgetragen. Da, wie auch im Merton-Modell, kein Verlust in der Genauigkeit bei Verwendung der Approximation eintritt betrachten wir für das Kou- und VG-Modell nur die Konvergenzrate im Verhältnis zur Rechenzeit, wie sie bei Verwendung der \mathcal{H} -Matrix-Approximation erreicht wird. Im Fall des CGMY-Modells sind noch einmal die Konvergenzgraphen beider Verfahren abgebildet. Wie auch im Merton-Modell wird in den hier betrachteten Modellen eine Steigerung der Konvergenzrate erreicht. In Tabelle 5.6 sind die entsprechenden Raten der Modelle zu finden. Die direkte Matrix-Vektor-Multiplikation erreicht eine Konvergenzrate von 1, bei Verwendung der \mathcal{H} -Matrix-Approximation sehen wir, dass Konvergenzraten von bis zu 1.9 erreicht werden. Insgesamt beobachten wir auch hier ein Ansteigen der Konvergenzrate. Dies ist durch die logarithmischen Faktoren, die bei der \mathcal{H} -Matrix-Approximation auftreten, bedingt.

5.2.3 Amerikanische Optionen

In diesem Abschnitt wenden wir uns der Bewertung von Amerikanischen Optionen zu. Die Lösung der partiellen Integro-Differential-Ungleichung erfolgt durch den Brennan-

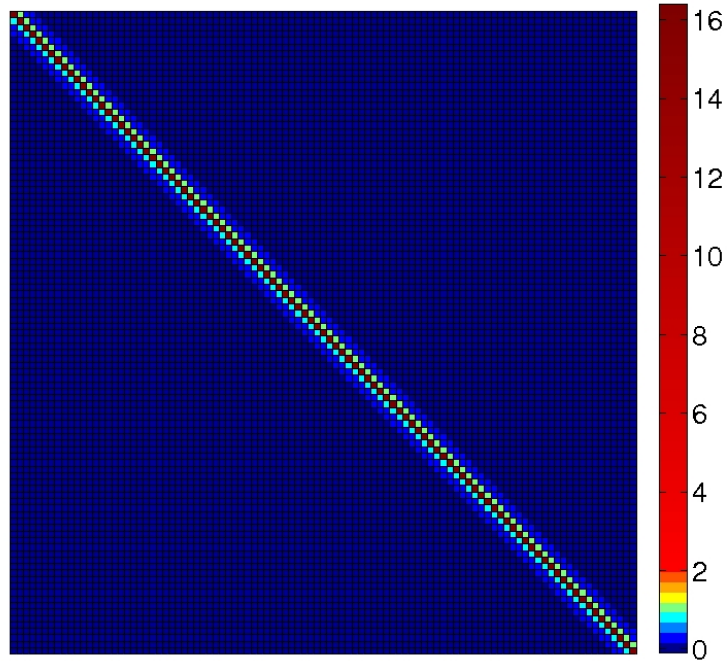


Abbildung 5.6: Einträge der Kollokationsmatrix im CGMY-Modell mit den Parametern aus Tab. 5.5

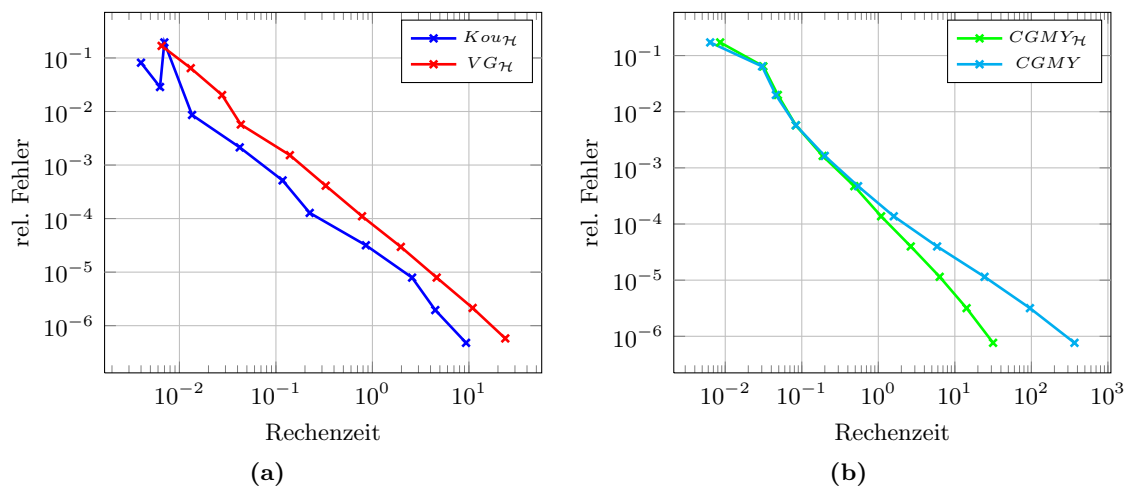


Abbildung 5.7: (a) Konvergenz im Kou- bzw. VG-Modell (b) Vergleich der Konvergenz im CGMY-Modell mit und ohne Approximation

Schwartz-Algorithmus. Die Parameter, die zur Bewertung verwendet werden, sind identisch mit denen, die bei der Bewertung von Europäischen Optionen verwendet wurden. Sie sind den Tabellen 5.1, 5.4 und 5.5 zu entnehmen. Es ist nicht zu erwarten, dass sich die Konvergenzrate gegenüber der Bewertung Europäischer Optionen verschlechtert, da die Projektion

M	$\rho_{\Theta}^{Kou_{\mathcal{H}}}$	$\rho_{\Theta}^{VG_{\mathcal{H}}}$	$\rho_{\Theta}^{CGMY_{\mathcal{H}}}$	ρ_{Θ}^{CGMY}
4	0.00	0.00	0.00	0.00
8	2.30	1.38	0.73	0.56
16	-18.59	1.55	2.09	2.13
32	4.72	2.83	2.02	1.90
64	1.23	1.13	1.57	1.47
128	1.39	1.53	1.34	1.27
256	2.16	1.52	1.53	1.17
512	1.04	1.42	1.41	0.94
1,024	1.26	1.54	1.46	0.83
2,048	2.53	1.72	1.61	0.92
4,096	1.92	1.88	1.87	0.99

Tabelle 5.6: Konvergenzraten der unterschiedlichen Modelle mit \mathcal{H} -Matrix Approximation

innerhalb des Brennan-Schwartz-Algorithmus' keinen Einfluss auf die Kollokationsmatrix hat.

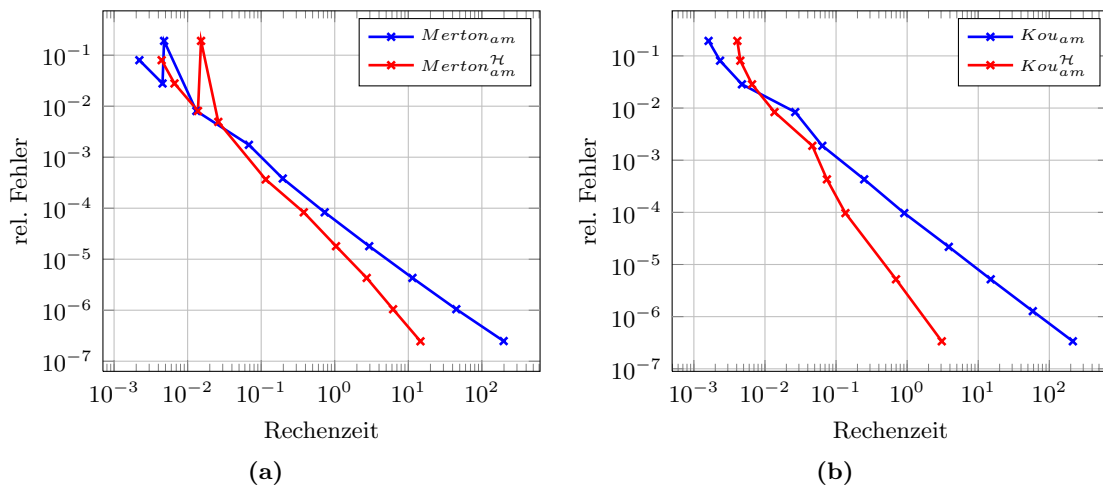


Abbildung 5.8: Vergleich der Konvergenz bei einer Amerikanischen Option (a) im Merton-Modell, (b) im Kou-Modell

In Abb. 5.8 sind die Konvergenzgraphen im Merton- und Kou-Modell abgebildet. Wir können hier wie im Fall Europäischer Optionen bei Verwendung der \mathcal{H} -Matrix-Approximation eine höhere Konvergenzgeschwindigkeit feststellen. Wir sehen ebenfalls, dass die Konvergenzrate im Merton-Modell nicht so schnell ansteigt, wie dies im Kou-Modell der Fall ist. Diese Tatsache folgt aus der Struktur der Dichte der Lognormalverteilung. Wie auch bei der Bewertung Europäischer Optionen verwenden wir bei der Approximation der Kollokationsmatrix im Merton-Modell die geometrische Bedingung, um eine effiziente Darstellung gewährleisten zu können.

In Abb. 5.9 sind die Konvergenzraten im VG- und CGMY-Modell beider Verfahren dargestellt. Man kann hier sehr gut erkennen, dass die Konvergenzrate des klassischen Verfahrens sehr schnell auf einen Wert nahe 1 gelangt und diesen dann hält. Die Konvergenzrate unter Verwendung der \mathcal{H} -Matrix-Approximation steigt hingegen bei stärkerer Verfeinerung an und nähert sich einer Rate von 2. Dieses Anwachsen der Rate ist wie auch in den an-

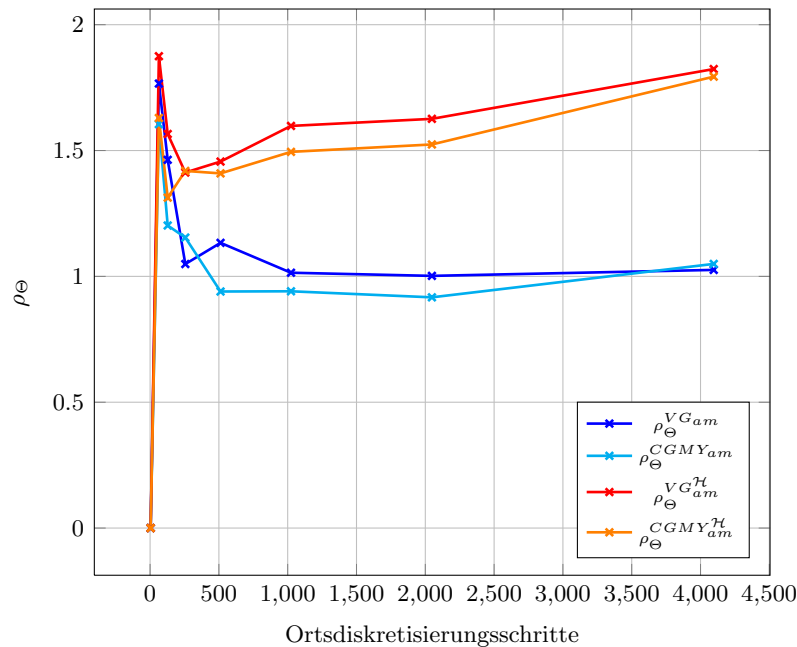


Abbildung 5.9: Vergleich der Konvergenzraten bei der Bewertung Amerikanischer Optionen im VG- und CGMY-Modell

deren Modellen den logarithmischen Faktoren im Aufwand der \mathcal{H} -Matrix-Approximation zuzuschreiben

In den Tabellen 5.7 und 5.8 sind die Konvergenzraten der Verfahren noch einmal aufgelistet. Wir sehen, dass durch die Approximation Konvergenzraten von bis zu 1.8 erreicht werden, wohingegen das klassische Verfahren in allen Modellen nur eine Konvergenzrate von 1 liefert.

M	ρ_{Θ}^{Merton}	$\rho_{\Theta}^{Merton^{\mathcal{H}}}$	ρ_{Θ}^{Kou}	$\rho_{\Theta}^{Kou^{\mathcal{H}}}$
4	0	0	0	0
64	0.93	0.77	1.7	1.22
128	1.43	1.75	1.09	3.13
256	1.16	1.24	1.15	2.49
512	1.1	1.51	1.03	1.43
1,024	1.06	1.49	1.06	1.69
2,048	1.03	1.71	1.03	1.78
4,096	0.97	1.81	1.03	1.85

Tabelle 5.7: Konvergenzrate im Merton- und Kou-Modell bei der Bewertung Amerikanischer Put-Optionen

5.3 Bewertung von Basket-Optionen

In diesem Abschnitt werden wir uns mit der Bewertung von Basket-Optionen auf zwei Wertpapiere befassen. Dazu müssen wir eine zweidimensionale PIDE lösen, wie sie in Kapitel 3 vorgestellt wurde. Die Fehlermessung erfolgt analog zum eindimensionalen Fall

M	ρ_{Θ}^{VG}	$\rho_{\Theta}^{VG^{\mathcal{H}}}$	ρ_{Θ}^{CGMY}	$\rho_{\Theta}^{CGMY^{\mathcal{H}}}$
4	0	0	0	0
64	1.6	1.63	1.77	1.87
128	1.2	1.31	1.46	1.57
256	1.15	1.42	1.05	1.41
512	0.94	1.41	1.13	1.46
1,024	0.94	1.49	1.01	1.6
2,048	0.92	1.52	1	1.63
4,096	1.05	1.79	1.03	1.82

Tabelle 5.8: Konvergenzrate im VG- und CGMY-Modell bei der Bewertung Amerikanischer Put-Optionen

in der diskreten L_2 -Norm, auch die Konvergenzrate wird wie im eindimensionalen Fall bestimmt. Als Anfangsbedingung wählen wir in allen Fällen die Payoff-Funktion

$$\mathcal{P}(S_1, S_2) = (K - \max\{S_1, S_2\})^+ \quad (5.7)$$

für Europäische und Amerikanische Put-Optionen. Die Parameter, die wir für die Bewertung der Basket-Optionen verwenden sind den Tabellen 5.9 und 5.10 zu entnehmen. Die in Tabelle 5.10 nicht aufgeführten Parameter werden aus Tabelle 5.9 übernommen. Die Parameter σ_1 und σ_2 bestimmen die Varianz der Lognormalverteilung des jeweiligen Sprungprozesses. Die Parameter μ_1 und μ_2 bestimmen deren Erwartungswert. Die Korrelation der Brownschen Komponente bzw. der Sprungkomponente werden durch die Parameter ρ_B und ρ_J bestimmt. Die Referenzwerte zur Bestimmung der Konvergenzraten wurden auf einem Gitter mit 15876 Ortsdiskretisierungspunkten und 160 Zeitschritten berechnet.

Parameter	Wert
S_1	100
S_2	100
K	100
T	0.25
r	0.1
σ_1	0.15
σ_2	0.15
ρ_B	0.5
μ_{J_1}	-0.9
μ_{J_2}	-0.9
σ_{J_1}	0.45
σ_{J_2}	0.45
ρ_J	0.5
λ	0.1
$\eta_{\mathcal{H}}$	0.4

Tabelle 5.9: Parameter im 2d Merton-Modell

Auch bei der Bewertung von Basket-Optionen müssen wir auf die geometrische Bedingung zur Steuerung der \mathcal{H} -Matrix-Approximation zurückgreifen. Da die Systemmatrix des Gleichungssystems sehr viele Nulleinträge aufweist, muss eine komplette Auswertung ver-

Parameter	Wert
C	0.5
G	1.2
M	1.2
Y	0.2
θ_C	0.5
η_C	0.5
$\eta_{\mathcal{H}}$	0.4

Tabelle 5.10: Parameter im 2d Copula-Modell

hindert werden, da der Aufwand sonst auf die Größenordnung $\mathcal{O}(N^4)$ ansteigt. Als Bedingung, die zur Auswertung eines Blocks erfüllt sein muss, wählen wir für die Systemmatrix

$$\text{dist}(X_t, X_s) < 3 \quad (5.8)$$

und für die Approximation der Kollokationsmatrix

$$\text{dist}(X_t, X_s) < 4. \quad (5.9)$$

In den numerischen Experimenten zeigt sich, dass diese Wahl der Parameter zu der gewünschten Verbesserung der Konvergenzrate bezüglich der Rechenzeit führt.

5.3.1 Europäische Optionen

Zunächst befassen wir uns mit der Lösung der PIDEs wie sie bei der Bewertung Europäischer Optionen im zweidimensionalen Merton- bzw. Copula-Modell entstehen. In Abb. 5.10 sind die Konvergenzgraphen im Merton-Modell mit und ohne \mathcal{H} -Matrix-Approximation angegeben. Auch hier können wir keine Verschlechterung in der Genauigkeit des Ergebnisses feststellen, s. Abb. 5.10 (a). Beide Verfahren erreichen eine Konvergenzrate von 1, wie sie bei der Lösung einer zweidimensionalen PIDE mittels eines Finite-Differenzen-Verfahrens zu erwarten ist. In Abb. 5.10 (b) ist die Konvergenzgeschwindigkeit der beiden Verfahren aufgetragen. Wir sehen, dass die Konvergenzrate des klassischen Verfahrens auf $\frac{1}{2}$ abfällt, wohingegen bei Verwendung der \mathcal{H} -Matrix-Approximation keine Verschlechterung festzustellen ist. Die Ergebnisse der Lösung der PIDE sind in den Tabellen 5.11 und 5.12 aufgelistet. Wir sehen, dass sich bei Verwendung der \mathcal{H} -Matrix-Approximation die Konvergenzrate bezüglich der Diskretisierungspunkte und die Konvergenzrate bezüglich der Rechenzeit kaum unterscheiden, beide liegen ungefähr bei 1. Ohne die Approximation beobachten wir jedoch beinahe eine Halbierung der Rate, wenn man bei der Betrachtung von den Diskretisierungspunkten zur Rechenzeit wechselt.

Die Ergebnisse aus Abb. 5.10 und Tabelle 5.12 sind durch die Approximation von Kollokationsmatrix und Systemmatrix des linearen Gleichungssystems entstanden. Anders als im Fall einer eindimensionalen PIDE lässt sich das Gleichungssystem der zweidimensionalen PIDE nicht mehr mit linearem Aufwand lösen. In Abb. 5.11 (a) wird die Konvergenz bezüglich der Rechenzeit bei ausschließlicher Approximation der Kollokationsmatrix und der Approximation von Kollokations- und Systemmatrix verglichen. Wir sehen hier, dass die Approximation der Kollokationsmatrix nicht ausreicht. Der Aufwand fällt auf die Größenordnung $\mathcal{O}(N^4)$ des klassischen Verfahrens zurück und unterscheidet sich von dieser nur in der Konstanten. Nur durch die Approximation beider Matrizen ist die Lösung mit einem Aufwand in der Größenordnung $\mathcal{O}(N^2)$ möglich.

M	Θ	e_{rel}	ρ_M	ρ_Θ
256	$1.3622 \cdot 10^0$	$4.8428 \cdot 10^{-2}$	0.79	0.73
484	$3.2246 \cdot 10^0$	$2.9372 \cdot 10^{-2}$	0.87	0.58
900	$6.2192 \cdot 10^0$	$1.6533 \cdot 10^{-2}$	0.96	0.87
1,600	$1.4911 \cdot 10^1$	$9.2547 \cdot 10^{-3}$	1.00	0.66
3,136	$4.3327 \cdot 10^1$	$4.6973 \cdot 10^{-3}$	1.01	0.64
4,900	$9.7282 \cdot 10^1$	$2.9051 \cdot 10^{-3}$	1.03	0.59
7,396	$2.5337 \cdot 10^2$	$1.9366 \cdot 10^{-3}$	0.99	0.42
10,000	$5.0085 \cdot 10^2$	$1.4000 \cdot 10^{-3}$	1.02	0.48
14,400	$1.2512 \cdot 10^3$	$8.9245 \cdot 10^{-4}$	1.09	0.49

Tabelle 5.11: Konvergenz ohne Approximation im zweidimensionalen Merton-Modell

M	$\Theta^{\mathcal{H}}$	$e_{rel}^{\mathcal{H}}$	$\rho_M^{\mathcal{H}}$	$\rho_\Theta^{\mathcal{H}}$
256	$1.6475 \cdot 10^0$	$4.8428 \cdot 10^{-2}$	0.79	0.69
484	$2.6233 \cdot 10^0$	$2.9372 \cdot 10^{-2}$	0.87	1.07
900	$4.7892 \cdot 10^0$	$1.6533 \cdot 10^{-2}$	0.96	0.95
1,600	$9.6886 \cdot 10^0$	$9.2547 \cdot 10^{-3}$	1.00	0.82
3,136	$2.1594 \cdot 10^1$	$4.6973 \cdot 10^{-3}$	1.01	0.85
4,900	$3.2479 \cdot 10^1$	$2.7051 \cdot 10^{-3}$	1.11	1.35
7,396	$5.5266 \cdot 10^1$	$1.9466 \cdot 10^{-3}$	0.92	0.82
10,000	$7.4027 \cdot 10^1$	$1.4200 \cdot 10^{-3}$	1.01	1.08
14,400	$1.1038 \cdot 10^2$	$9.1245 \cdot 10^{-4}$	1.08	1.11

Tabelle 5.12: Konvergenz mit \mathcal{H} -Matrix Approximation im zweidimensionalen Merton-Modell

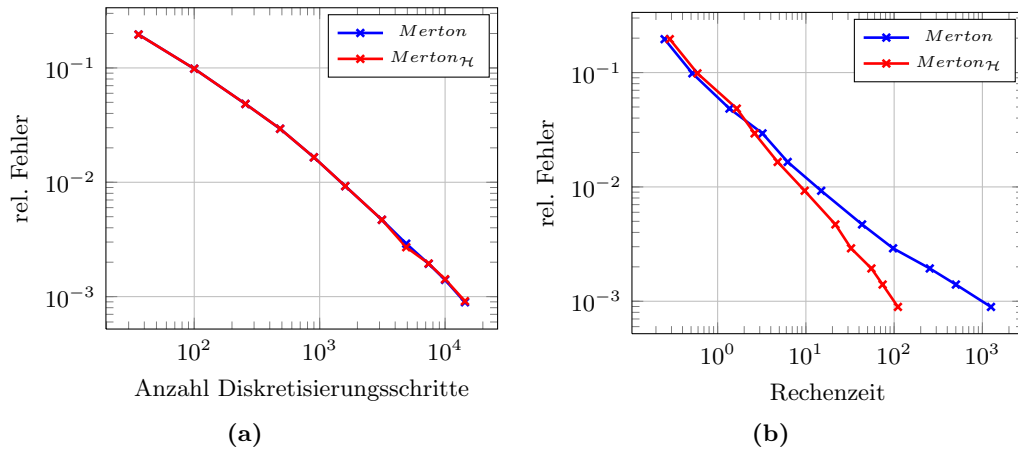


Abbildung 5.10: Konvergenz im zweidimensionalen Merton-Modell, dabei wurde der relative Fehler gegen a) Aufwand in Diskretisierungspunkten b) Aufwand in Rechenzeit aufgetragen

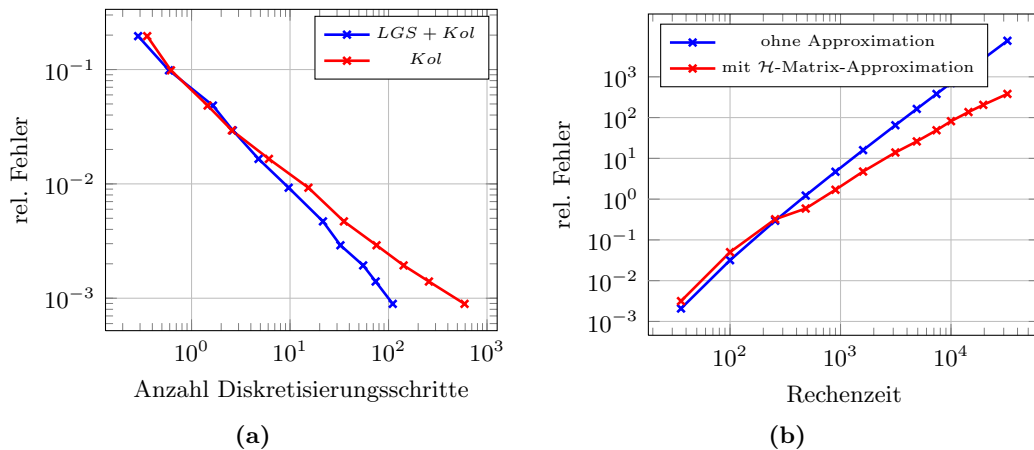
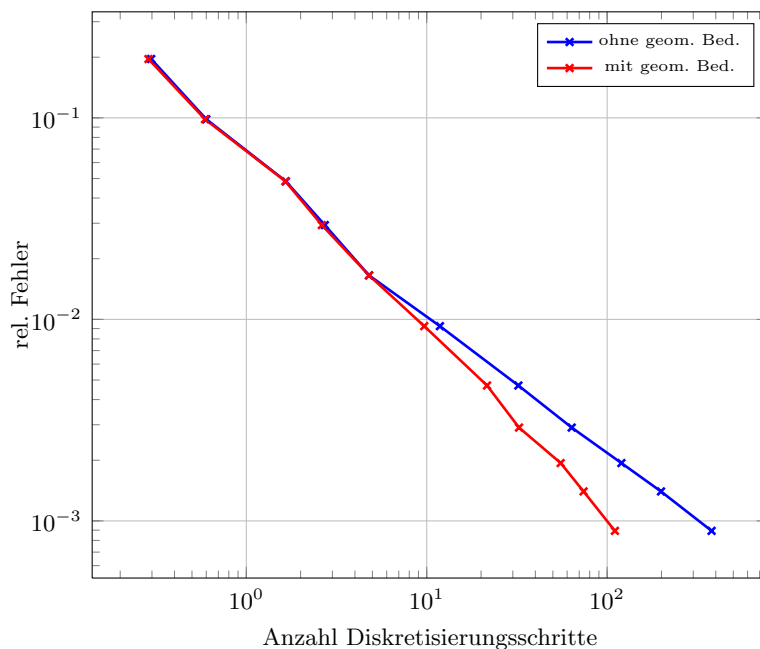


Abbildung 5.11: (a) Vergleich der Konvergenzrate bei Approximation der Kollokationsmatrix und Approximation der System- und Kollokationsmatrix (b) Vergleich des Speicherbedarfs ohne Approximation und bei Approximation beider Matrizen

Die hohe Konvergenzrate in den zweidimensionalen Modellen kann nur durch das Einfügen der oben schon erwähnten geometrischen Bedingung erreicht werden. In Abb. 5.12 sehen wir den Unterschied in der Konvergenzgeschwindigkeit mit und ohne Anwendung der geometrischen Bedingung. Die Konvergenzrate nimmt auf feineren Gittern ab, wenn man die Approximation der System- und Kollokationsmatrix nicht steuert. Dies liegt hauptsächlich an der großen Menge an Nulleinträgen in der Systemmatrix. Diese weist nur Einträge in der Nähe der Diagonalen auf, weshalb es sinnvoll ist, die Auswertung der Matrixeinträge auf dieses Gebiet zu einzuschränken.

Die Beschränkung der zu approximierenden Blöcke innerhalb der Kollokationsmatrix beruht auf der Tatsache, dass auch die mehrdimensionale Dichte am Rand des Integrationsgebiets schnell abnimmt. Dies wird durch die Beschränkung der Approximation ausgenutzt.

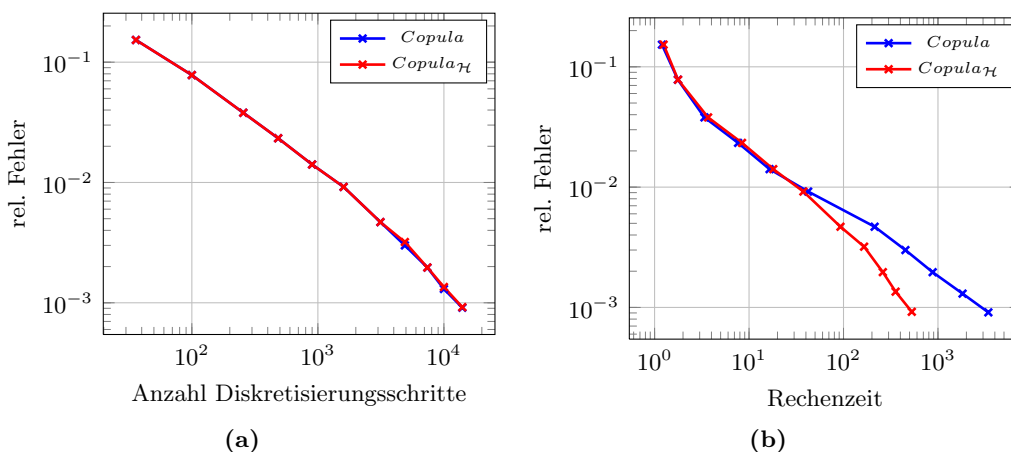
In Abb. 5.13 sehen wir die Konvergenz gegen die Lösung im zweidimensionalen Copula-Modell. Auch hier ist keine Verschlechterung in der Genauigkeit durch die Anwendung der \mathcal{H} -Matrix-Approximation festzustellen. In Abb. 5.13 (b) wird deutlich, dass die Approxi-



(a)

Abbildung 5.12: Vergleich der Konvergenzgeschwindigkeit mit und ohne geometrische Bedingung im zweidimensionalen Merton-Modell

mation der Kollokationsmatrix und der Systemmatrix des linearen Gleichungssystems die Konvergenzrate bezüglich der Rechenzeit erhöht. Wie auch im Merton-Modell erreichen wir hier eine Konvergenzrate von 1 gegenüber einer Rate von $\frac{1}{2}$ bei direkter Ausführung der Matrix-Vektor-Multiplikation und iterativem Lösen des Gleichungssystems. Bei gleicher Konvergenzordnung stellen wir im Copula-Modell eine wesentlich höhere Rechenzeit als im Merton-Modell fest. Dies liegt daran, dass die Berechnung der einzelnen Matrixeinträge der Kollokationsmatrix im Copula-Modell aufwendiger ist. Da in jedem Summanden der Simpson-Regel zwei Integrale ausgewertet werden müssen, erhöht sich die Zeit, die zur Aufstellung der Kollokationsmatrix benötigt wird, im klassischen Modell um den Faktor 3, bei Verwendung der Approximation um den Faktor 5.



(a)

(b)

Abbildung 5.13: Konvergenz im zweidimensionalen Copula-Modell, dabei wurde der relative Fehler gegen (a) Aufwand in Diskretisierungspunkten (b) Aufwand in Rechenzeit aufgetragen

M	Θ	e_{rel}	ρ_M	ρ_Θ
256	$3.3721 \cdot 10^0$	$3.7984 \cdot 10^{-2}$	0.77	1.11
484	$7.6566 \cdot 10^0$	$2.3319 \cdot 10^{-2}$	0.77	0.59
900	$1.6499 \cdot 10^1$	$1.4106 \cdot 10^{-2}$	0.81	0.65
1,600	$4.2217 \cdot 10^1$	$9.1880 \cdot 10^{-3}$	0.75	0.46
3,136	$2.1346 \cdot 10^2$	$4.6864 \cdot 10^{-3}$	1.00	0.42
4,900	$4.5234 \cdot 10^2$	$2.9985 \cdot 10^{-3}$	1.00	0.59
7,396	$8.7860 \cdot 10^2$	$1.9627 \cdot 10^{-3}$	1.03	0.64
10,000	$1.8212 \cdot 10^3$	$1.3024 \cdot 10^{-3}$	1.36	0.56
14,000	$3.4144 \cdot 10^3$	$9.0994 \cdot 10^{-4}$	1.07	0.57

Tabelle 5.13: Konvergenz ohne Approximation im zweidimensionalen Copula-Modell

M	$\Theta^{\mathcal{H}}$	$e_{rel}^{\mathcal{H}}$	$\rho_M^{\mathcal{H}}$	$\rho_\Theta^{\mathcal{H}}$
256	$3.6841 \cdot 10^0$	$3.7984 \cdot 10^{-2}$	0.80	0.99
484	$8.4242 \cdot 10^0$	$2.3319 \cdot 10^{-2}$	0.86	0.71
900	$1.8038 \cdot 10^1$	$1.4106 \cdot 10^{-2}$	0.89	0.77
1,600	$3.7628 \cdot 10^1$	$9.1880 \cdot 10^{-3}$	0.86	0.74
3,136	$9.2874 \cdot 10^1$	$4.6864 \cdot 10^{-3}$	1.00	0.79
4,900	$1.6520 \cdot 10^2$	$3.1985 \cdot 10^{-3}$	0.94	0.82
7,396	$2.6080 \cdot 10^2$	$1.9649 \cdot 10^{-3}$	1.08	1.03
10,000	$3.5614 \cdot 10^2$	$1.3524 \cdot 10^{-3}$	1.07	1.06
14,000	$5.2819 \cdot 10^2$	$9.1994 \cdot 10^{-4}$	1.05	0.99

Tabelle 5.14: Konvergenz mit \mathcal{H} -Matrix Approximation im zweidimensionalen Copula-Modell

Die genauen Ergebnisse der Lösung der PIDE aus dem Copula-Modell sind in den Tabellen 5.13 und 5.14 aufgelistet. Hier wird noch einmal deutlich, dass die Verringerung der Konvergenzrate nur bei direkter Behandlung des Integrals und iterativem Lösen des LGS eintritt. Verwendet man die \mathcal{H} -Matrix-Approximation, bleibt die Rate auch bei der Messung bezüglich der Rechenzeit auf 1.

5.3.2 Amerikanische Optionen

Im letzten Abschnitt lösen wir nun die zweidimensionale partielle Integro-Differentialgleichung, wie sie bei der Bewertung Amerikanischer Optionen im Merton- bzw. Copula-Modell entsteht. Wie auch bei der Bewertung Europäischer Optionen stellt sich die erwartete Verbesserung in der Konvergenzgeschwindigkeit durch die \mathcal{H} -Matrix-Approximation ein. In Abb. 5.14 sind die Konvergenzgraphen der beiden Verfahren für das Merton- und das Copula-Modell abgebildet. In beiden Modellen sehen wir nach einer kurzen Phase, in der beide Verfahren gleich schnell konvergieren, dass das klassische Verfahren auf eine Konvergenzrate von $\frac{1}{2}$ bezüglich der Rechenzeit zurückfällt. Die identische Konvergenz der beiden Verfahren zu Beginn ist durch die Tatsache bedingt, dass im Fall kleiner Gleichungssysteme nicht das Lösen des Systems oder das Aufstellen der Matrizen, sondern andere Faktoren zeitlich dominieren. Durch die \mathcal{H} -Matrix-Approximation kann eine Konvergenzrate von 1 erreicht werden. Damit werden auch Amerikanische Optionen in optimaler Zeit gelöst. In Tabelle 5.15 sind die Konvergenzraten der beiden Verfahren bezüglich der Re-

chenzeit aufgelistet. Wir sehen hier, dass durch die Approximation eine beinahe doppelt so hohe Konvergenzrate erzielt werden kann.

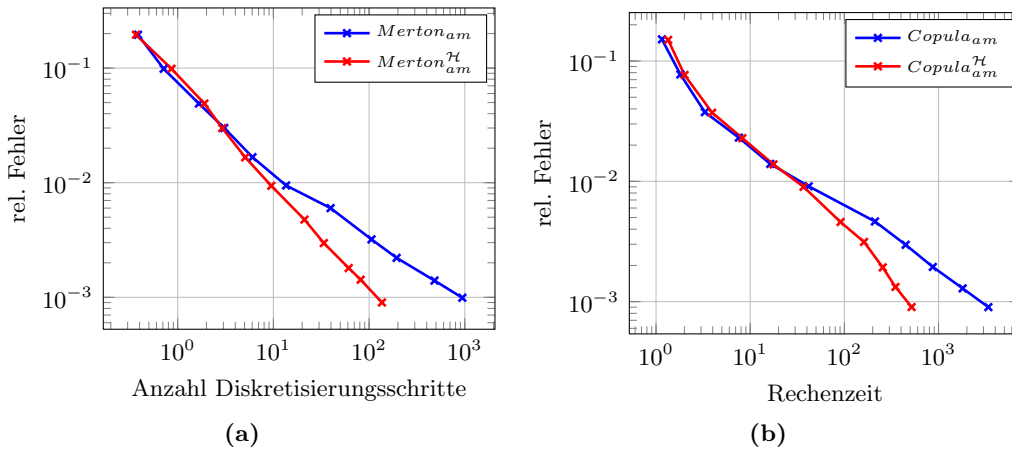


Abbildung 5.14: (a) Vergleich der Konvergenzrate bei der Bewertung Amerikanischer Optionen im zweidimensionalen Merton-Modell (b) Vergleich des Speicherbedarfs ohne Approximation und bei Approximation beider Matrizen

M	ρ_{Θ}^{Merton}	$\rho_{\Theta}^{Merton^{\mathcal{H}}}$	ρ_{Θ}^{Copula}	$\rho_{\Theta}^{Copula^{\mathcal{H}}}$
36	2.52	0.15	0.00	0.00
100	1.11	0.84	2.50	1.67
256	0.82	0.91	1.44	1.19
484	0.80	1.07	0.94	0.99
900	0.86	1.03	0.65	0.77
1,600	0.70	0.95	0.46	0.74
3,136	0.43	0.89	0.42	0.79
4,900	0.64	1.01	0.59	0.82
7,396	0.61	0.91	0.64	1.03
10,000	0.50	0.95	0.56	1.06
14,000	0.52	0.95	0.57	0.99

Tabelle 5.15: Konvergenz mit \mathcal{H} -Matrix Approximation

5.4 Fazit

Wir haben in diesem Kapitel gesehen, dass es möglich ist, durch Verwendung der \mathcal{H} -Matrix-Approximation Plain-Vanilla- und Basket-Optionen in optimaler Zeit zu bewerten. Die schnellere Konvergenzgeschwindigkeit gegenüber der direkten Behandlung der Kollokationsmatrix wurde in allen getesteten Modellen erreicht. Bei der Bewertung von Basket-Optionen ist es allerdings zusätzlich notwendig, die Systemmatrix des linearen Gleichungssystems zu approximieren. Da die Matrix, im Gegensatz zum eindimensionalen Fall keine Tridiagonalmatrix ist, ist selbst bei Verwendung eines iterativen Löser eine Lösung des LGS nur durch einen Aufwand in der Größenordnung $\mathcal{O}(N^4)$ zu erhalten. Durch die Approximation ist es möglich, eine Plain-Vanilla-Option in Sprung-Diffusions-Modellen

mit einem Aufwand in der Größenordnung $\mathcal{O}(N)$ und eine Basket-Option auf zwei Wertpapiere mit einem Aufwand in der Größenordnung $\mathcal{O}(N^2)$ zu bewerten.

Um die hohe Konvergenzgeschwindigkeit in allen Modellen garantieren zu können, ist es notwendig, die \mathcal{H} -Matrix-Approximation durch zusätzliches Einfügen einer geometrischen Bedingung an die Bearbeitung der \mathcal{H} -Matrix-Blöcke zu optimieren. Eine Anwendung der Approximation ohne Anpassung an die verwendete Sprungdichte führt unter Umständen dazu, dass das Verfahren wieder auf eine Rechenzeit in der Größenordnung des klassischen Modells zurückfällt.

Kapitel 6

Zusammenfassung und Ausblick

In dieser Arbeit wurde ein effizientes numerisches Verfahren zur Lösung von PIDEs vorgestellt, wie sie bei der Bewertung von Europäischen und Amerikanischen Plain-Vanilla- und Basket-Optionen in Sprung-Diffusions-Modellen entstehen. Dafür wurde unseres Wissens nach erstmalig in der Finanznumerik eine \mathcal{H} -Matrix-Approximation zur Behandlung des Integralteils der Gleichung eingesetzt.

Die Bewertung der Plain-Vanilla-Optionen fand in dem Sprung-Diffusions-Modell von Merton, dem Modell von Kou, sowie dem VG- und CGMY-Modell statt. Zur Diskretisierung der PIDEs wurde ein Finite-Differenzen-Verfahren mit zentralen Differenzenquotienten verwendet. Da diese eine Konvergenzordnung von 2 aufweisen, war es notwendig, ein Crank-Nicolson-Verfahren zur Zeitdiskretisierung zu verwenden, da dieses die gleiche Konvergenzordnung besitzt. Durch die Crank-Nicolson-Diskretisierung hat das Verfahren also identische Konvergenzraten in der Orts- und Zeitdiskretisierung. Um Oszillationen in der Lösung der PIDE zu verhindern, wurde das Crank-Nicolson-Verfahren auf ein Rannacher Zeitschritt-Verfahren erweitert. Der Integralteil der Gleichung wurde ausschließlich explizit behandelt, was dazu führte, dass sich die Auswertung des Integrals auf die Anwendung eines Kollokationsverfahrens beschränkte. Die aus dem Kollokationsverfahren resultierende Matrix-Vektor-Multiplikation wurde dann mit Hilfe der \mathcal{H} -Matrix-Approximation durchgeführt.

Zur Bewertung der Basket-Optionen benutzten wir das zweidimensionale Merton-Verfahren, sowie die noch relativ neue Variante der Lévy-Copulas. Dabei wurde die mehrdimensionale Lévy-Dichte mit Hilfe sogenannter Lévy-Copulas aus den eindimensionalen Dichten eines CGMY-Prozesses generiert.

Wir konnten feststellen, dass sowohl bei der Bewertung von Plain-Vanilla-Optionen, als auch bei der Bewertung von Basket-Optionen eine Verbesserung der Konvergenzrate durch die \mathcal{H} -Matrix-Approximation erzielt werden konnte. In Kapitel 5 haben wir gezeigt, dass die Bewertung von Plain-Vanilla-Optionen mit optimalem Aufwand in der Größenordnung $\mathcal{O}(N)$ möglich ist. Die Bewertung von Basket-Optionen auf zwei Wertpapiere war ebenfalls mit optimalem Aufwand in der Größenordnung $\mathcal{O}(N^2)$ möglich. Wir haben gezeigt, dass es bei der Bewertung von Basket-Optionen notwendig ist, auch die Systemmatrix des LGS zu approximieren, wenn man die optimale Konvergenz erreichen will. Ebenso wurde gezeigt, dass man bei der Approximation der Kollokationsmatrizen die Struktur der verwendeten Dichte berücksichtigen muss. Dies wurde durch das Einfügen einer geometrischen Bedingung zur Approximation der \mathcal{H} -Matrix-Blöcke möglich. Bei Nichtbeachtung dieser Strukturen kann der Fall auftreten, dass die Konvergenzrate auch bei Verwendung der Approximation auf die der direkten Behandlung abfällt.

Insgesamt haben wir also gezeigt, dass man mit Hilfe der \mathcal{H} -Matrix-Approximation ein numerisches Verfahren implementieren kann, mit dem man sowohl Plain-Vanilla als auch Basket-Optionen in Sprung-Diffusionsmodellen mit optimalem Aufwand bewerten kann.

Neben den in dieser Arbeit vorgestellten Anwendungen der \mathcal{H} -Matrix-Approximation wäre es interessant zu untersuchen, wie sich die Anwendung der Approximation auf die Berechnung von Optionskennziffern, den sogenannten Greeks auswirkt. Eine weitere Möglichkeit der Erweiterung besteht darin, das Verfahren auf Basket-Optionen mit mehr als zwei zugrunde liegenden Wertpapieren zu erweitern. Dabei ist darauf zu achten, dass der mit der Dimension exponentiell ansteigende Aufwand, der bei klassischen Finite-Differenzen-Verfahren entsteht, vermieden wird. Dies kann z.B. durch die Anwendung von Dünngitterverfahren geschehen. Es sollte ebenfalls getestet werden, wie sich das Approximations-Verfahren auswirkt, wenn man die Finite-Differenzen-Diskretisierung durch eine Finite-Element-Diskretisierung ersetzt und die Behandlung des Integrals nicht durch ein Kollokationsverfahren, sondern mit einer Galerkin-Diskretisierung durchführt.

Literaturverzeichnis

- [Alm05] ALMENDRAL, A.: *Numerical Valuation of American Options Under the CGMY Process*. in: A.E. Kyprianou, W. Schoutens, P. Wilmott (Eds.), *Exotic Option Pricing and Advanced Lévy Models*, 2005.
- [Bau02] BAUER, H.: *Wahrscheinlichkeitstheorie*. deGruyter, 2002.
- [Beb] BEBENDORF, M.: *AHMED - Another software library on hierarchical matrices for elliptic differential equations*. Erhältlich unter : <http://bebendorf.ins.uni-bonn.de/AHMED.html>.
- [Beb08] BEBENDORF, M.: *Hierarchical Matrices*. Springer, 2008.
- [BS73] BLACK, F. und M. SCHOLES: *The Pricing of Options an Liabilities*. *Journal of Political Economy*, 81, 1973.
- [BS97] BRENNAN, M. und E. SCHWARTZ: *The Valuation of American Put Options*. *Journal of Finance*, 32:449–462, 1997.
- [CF08] CLIFT, S. und P. FORSYTH: *Numerical Solution of Two Asset Jump Diffusion Models for Option Valuation*. *Applied Numerical Mathematics*, 58(6):743–782, 2008.
- [CGMY02] CARR, P., H. GEMAN, D. MADAN und M. YOR: *The Fine Structure of Asset Returns: An Empirical Investigation*. *Journal of Business*, 75(2):305–332, 2002.
- [CT04] CONT, R. und P. TANKOV: *Financial Modelling with Jump Diffusion Processes*. Chapman & Hall/CRC Financial Mathematical Series, 2004.
- [CV05] CONT, R. und E. VOLTCHKOVA: *A Finite Difference Scheme for Option Pricing in Jump Diffusion and Eponential Levy Models*. *SIAM Journal of Numerical Analysis*, 43:1596–1626, 2005.
- [dFL04] D’HALLUIN, Y., P. FORSYTH und G. LABAHN: *A Penalty Method for American Options with Jump Diffusion*. *Numerische Mathematik*, 97(2):321–352, 2004.
- [DFV05] D’HALLUIN, Y., P. FORSYTH und K. VETZAL: *Robust numerical methods for contingent claims under jump diffusion processes*. *IMA Journal of Numerical Analysis*, 25, 2005.
- [FRS06] FARKAS, W., N. REICH und C. SCHWAB: *Anisotropic Stable Lévy Copula Processes - Analytical and Numerical Aspects*. Research Report No. 2006-08, Eidgenössische Technische Hochschule Zürich, 2006.
- [FV02] FORSYTH, P. und K. VETZAL: *Quadratic Convergence for Valuing American Options using a Penalty Method*. *SIAM Journal on Scientific Computing*, 23, 2002.

- [Gla04] GLASSERMANN, P.: *Monte Carlo Methods in Financial Engineering*. Springer, 2004.
- [GMV01] GEMAN, H., D. MADAN und M. YOR: *Time Changes for Levy-Processes*. *Mathematical Finance*, 11:19–37, 2001.
- [Hac85] HACKBUSCH, W.: *Multi-grid methods and applications*. Springer, 1985.
- [Hac99] HACKBUSCH, W.: *A Sparse Matrix Arithmetic Based on \mathcal{H} -Matrices. Part I: Introduction to \mathcal{H} -Matrices*. *Computing*, 62(2):89–108, 1999.
- [Hac07] HACKBUSCH, W.: *Hierarchische Matrizen - Algorithmen und Analysis*, 2007.
- [HB02] HANKE-BOURGEOIS, M.: *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Teubner, 2002.
- [HN89] HACKBUSCH, W. und Z. NOWAK: *On the Fast Matrix Multiplication in the Boundary Element Method by Panel Clustering*. *Num. Math*, 54(4):463–491, 1989.
- [Hul06] HULL, J.: *Optionen, Futures und andere Derivate*. Prentice Hall, 2006.
- [IT05] IKONEN, S. und J. TOIVANEN: *Pricing American Options using LU Decomposition*. *Applied Mathematical Sciences*, 1(51):2529 – 2551, 2005.
- [Ket06] KETTLER, P.: *Lévy-Copula-Driven Financial Processes*. Pure Mathematics, Preprint #23, Department of Mathematics, University of Oslo, 2006.
- [Kop95] KOPONEN, I.: *Analytic Approach to the Problem of Convergence of Truncated Levy Flights Towards the Gaussian Stochastic Process*. *Physical Review*, 52:1197–1199, 1995.
- [Kou02] KOU, S.: *A Jump-Diffusion Model for Option Pricing*. *Management Science*, 48:1086–1101, 2002.
- [KT06] KALLSEN, J. und P. TANKOV: *Characterization of dependence of multidimensional Lévy processes using Lévy copulas*. *Journal of Multivariate Analysis*, 97(7):1551 – 1572, 2006.
- [KW04] KOU, S. und HUI WANG: *Option Pricing Under a Double Exponential Jump Diffusion Model*. *Manage. Sci.*, 50(9):1178–1192, 2004.
- [Kwo98] KWOK, Y.: *Mathematical Models of Financial Derivates*. Springer Finance, 1998.
- [LS01] LONGSTAFF, FRANCIS A. und S. SCHWARTZ, EDUARDO: *Valuing American Options by Simulation : A Simple Least-Squares Approach*. *The Review of Financial Studies*, 14, 2001.
- [MCC98] MADAN, D., P. CARR und E. CHANG: *The Variance Gamma Process and Option Pricing*. *European Finance Review*, 2:79–105, 1998.
- [Mer73] MERTON, R.: *Theory of Rational Option Pricing*. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.

-
- [Mer76] MERTON, R.: *Option Pricing when Underlying Stocks are Discontinuous*. Journal of Financial Economics, 3:125–144, 1976.
- [Mir60] MIRSKY, L.: *Symmetric Gauge Functions and Unitarily Invariant Norms*. Quart. J. Math. Oxford(2), 11:50–59, 1960.
- [MNS] MATAACHE, A., P.-A. NITSCHKE und C. SCHWAB: *Wavelet Galerkin Pricing of American Options on Levy Driven Assets*. Research report / Seminar für Angewandte Mathematik No. 2003-06, ETH Zürich.
- [MvPS04] MATAACHE, A., T. VON PETERSDORFF und C. SCHWAB: *Fast Deterministic Pricing of Options on Levy Driven Assets*. Modélisation mathématique et analyse numérique, 38(2):37–51, 2004.
- [Rai00] RAIBLE, S.: *Levy Processes in finance, Theory, Numerics and Empirical Facts*. Doktorarbeit, Albert-Ludwigs-Universität Freiburg i. Br., 2000.
- [Rei06] REIFERSCHIED, M.: *Numerische Simulation von Sprung-Diffusions-Prozessen zur Optionspreisbewertung*. Diplomarbeit, 2006.
- [Sam65] SAMUELSON, P.: *Rational Theory of Warrant Pricing*. Industrial Management Review, 6:13–31, 1965.
- [Sat99] SATO, K.: *Levy-Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.
- [Sch05] SCHOUTENS, W.: *Lévy Processes in Finance*. Wiley Series in probability and statistics, 2005.
- [Toi08] TOIVANEN, J.: *Numerical Valuation of European and American Options under Kou's Jump-Diffusion Model*. SIAM Journal on Scientific Computing, 30:1949–1970, 2008.