



Institut für Numerische und Angewandte Mathematik

Bewertung von Flexibilität in freien Stromhandelssystemen unter Berücksichtigung von Sprüngen

Diplomarbeit

betreut von

Prof. Dr. Martin Burger und Magnus Wobben

eingereicht von

Oleg Reichmann

Münster, 29.08.2008

Danksagung

Ich danke

Martin Burger und Magnus Wobben für die Betreuung und Unterstützung bei der Erstellung dieser Arbeit und die Beantwortung meiner zahlreichen Fragen.

Martin Drohmann für die zahlreichen fruchtbaren Diskussionen und das Korrekturlesen dieser Arbeit.

Arvind Sarin für das Korrekturlesen dieser Arbeit.

Julia Bröcker für das Korrekturlesen dieser Arbeit.

Meinen Eltern und meiner Schwester, die mich immer unterstützt haben.

Inhaltsverzeichnis

I. Abbildungsverzeichnis	i
II. Abkürzungsverzeichnis	iii
III. Tabellenverzeichnis	iv
1 Einleitung	1
2 Mathematische Grundlagen	2
2.1 Stochastische Grundlagen	2
2.2 Halbgruppentheorie	12
2.3 Pseudo-Differentialoperatoren	21
2.4 Lévy-Prozesse und Halbgruppen	24
2.5 Bewertung von contingent claims im NIG-Modell	27
3 Einführung in Energiemärkte	29
3.1 Grundlagen	29
3.2 Wesentliche Eigenschaften des Spotkurses	30
4 Das Kluge-Modell	34
4.1 Beschreibung	34
4.2 Herleitung der PIDE für den europäischen Put	36
5 Mathematische Analyse des Kluge-Modells	38
5.1 Existenz einer Lösung	38
5.1.1 A_0 erzeugt analytische Halbgruppe	40
5.1.2 A erzeugt analytische Halbgruppe	42
5.1.3 $A + K$ erzeugt analytische Halbgruppe	42
5.1.4 Lipschitz-Stetigkeit von $t \mapsto (A + K)(t)$	44
5.1.5 Lipschitz-Stetigkeit der rechten Seite	45
5.1.6 Existenzresultat	45
5.2 Maximumprinzip der Differentialgleichung	46

6	Kalibrierung des Kluge-Modells	49
6.1	Kalibrierung des Prozesses unter dem realen Maß	49
6.2	Kalibrierung unter einem risikoneutralen Maß	53
6.2.1	Forwardbewertung	55
6.2.2	Umsetzung der hourly Forward Curve	56
7	Auswertung des Kluge-Modells	59
7.1	Diskretisierung	59
7.2	Fehleranalyse	62
7.3	Umsetzung	67
7.4	Numerische Analyse	68
8	Feller Prozesse vom NIG-Typ	72
8.1	Das FPNIG-Modell	72
8.2	Bewertung von contingent claims im FPNIG-Modell	74
8.3	Kalibrierung und Implementierung	77
9	Modellvergleich	80
9.1	Optionsbewertung im Lucia-Schwartz-Modell	80
9.2	Vergleich der Optionspreise	81
10	Fazit	84
	Anhang	85
A	Halbgruppentheorie	85
A.1	Hilfssätze der Halbgruppentheorie	85
A.2	Fundamentallösung für zeitabhängige Operatoren	85
A.3	Satz von Hille-Yosida	88
B	Hilfsresultate	92
B.1	Bessel Funktionen	92
B.2	Hilfssätze der Stochastik	93
B.3	Lucia-Schwartz Modell	95
B.4	Herleitung der Randbedingungen	97
IV.	Literaturverzeichnis	98

Abbildungsverzeichnis

2.1	Beispielpfade arithmetisch Brownscher Bewegungen	3
2.2	Beispielpfade geometrisch Brownscher Bewegungen	4
2.3	Beispielpfad für das Lucia-Schwartz Modell	4
2.4	Beispielpfad für das Merton Modell	5
2.5	NIG-Dichte mit folgenden Parameteren (a) $\alpha = 0.01, \beta = 0, \mu = 0,$ $\delta = 0.1$ und (b) $\alpha = 0.001, \beta = 0, \mu = 0, \delta = 1$	11
2.6	NIG-Dichte mit folgenden Parameteren (a) $\alpha = 3, \beta = 2.99, \mu = 0,$ $\delta = 0.4$ und (b) $\alpha = 3, \beta = -2.99, \mu = 0, \delta = 0.4$	11
2.7	Intergrationsweg Γ	19
3.1	Beispiel für Angebots- und Nachfragekurven	30
3.2	Tagessaisonalität Beispielhaft vom 3.03.07 bis zum 12.03.07 dargestellt .	31
3.3	Beispiel für Angebots- und Nachfragekurven	32
3.4	Spotkurs der Jahre 2005-2007 an der EEX	32
4.1	Deterministische Komponente im Kluge-Modell	35
4.2	Realisation von X_t und Y_t mit $\kappa = 1.7892, \sigma = 0.209, \mu_J = 0.9, \eta =$ $15/365, \beta = 9.4$	35
4.3	Simulation des Kluge-Prozesses	35
6.1	Getrennte desaisonalisierte Preise	51
6.2	Vergleich der Forwardkurve mit den tatsächlichen Preisrealisationen . . .	56
6.3	Kalibrierte Tagesforwardpreise	57
6.4	Kalibrierte Stundenforwardpreise bei einem konstanten MPR $\lambda = -0.94,$ $\eta^* = 0.08$	57
6.5	Kalibrierte Stundenforwardpreise bei variablem MPR	58
7.1	Putpreise im Kluge-Modell	69
7.2	Preisunterschiede bei Variation von η^* und β	69
7.3	Vergleich der Putpreise für unterschiedliche Randwerte	70
8.1	Darstellung des Parameters β in Abhängigkeit von x . $\chi = 2, \beta_0 = 0, x_0 = 0$ und $\epsilon = 0.9$	73

8.2	Simulation eines Pfades für $\alpha = 3, \beta$ wie oben, $\mu = -0.004$ und $\delta = 0.139$	73
8.3	Simulation eines Pfades für $a_0 = 1.1, a_1 = 0.8, a_2 = 0.25, \mu = -0.004$ und $\delta = 0.139$	74
8.4	Preise aus dem Kluge-Modell (blau) und die kalibrierten FPNIG-Preise für unterschiedliche ω	78
9.1	Puts im Lucia Schwartz-Modell	82
9.2	Relative Unterschiede zwischen Optionspreisen im LS-Modell und im Kluge-Modell	82
9.3	Vergleich von Optionen im LS-Modell und FPNIG-Modell	82
9.4	Vergleich von Optionspreisen im FPNIG-Modell und im Kluge-Modell	83

Abkürzungsverzeichnis

EEX	European Energy Exchange
FDM	Finite Differenzen Methode
FPNIG	Feller Process vom NIG-Typ
hFC	Hourly Forward Curve
kWh	Kilowattstunde
LS	Lucia-Schwartz
MAD	Mean Absoulte Error
ML	Maximum Likelihood
MPR	Marktpreis des Risikos
MSE	Mean Square Error
MW	Megawatt
MWh	Megawattstunde
OTC	Over the Counter
OU	Ornstein-Uhlenbeck
PDE	Partielle Differentialgleichung
PIDE	Partielle Differential- und Intergralgleichung
PDO	Pseudo-Differentialoperator
SDE	Stochastische Differentialgleichung

Tabellenverzeichnis

6.1	Kalibrierte Saisonparameter	50
6.2	Kalibrierte Parameter des Diffusionsprozesses	53
6.3	Kalibrierte Parameter des Spikeprozesses	53
6.4	Risikoneutrale Parameter	58
7.1	EOC	71
8.1	Kalibrierte Parameter im FPNIG-Modell	78

Kapitel 1

Einleitung

Mit der Liberalisierung der Strommärkte in Europa und insbesondere in Deutschland entstand erstmals die Möglichkeit Strom in großen Mengen zu handeln. Diese Möglichkeit wurde nicht nur von Erzeugern und Nachfragern, sondern auch von Dritten genutzt und spiegelt sich heute in einem OTC- und Börsenhandel sowohl für Spot- als auch für Terminmarktprodukte wieder. Vor der Liberalisierung bestand weder die Möglichkeit noch die Notwendigkeit Strom zu handeln, da auf dem Markt ausschließlich vertikal integrierte Versorger auftraten.

Mit der Entstehung solcher Märkte geht die Frage nach der Bewertung der dort gehandelten Produkte einher. Für die Bewertung von Terminmarktprodukten werden i.d.R. leicht angepasste Aktienmodelle verwendet. Dies ist bei Stundenoptionen auf Grund der besonderen Charakteristika von Strom, den sogenannten stylized facts, nicht mehr möglich. Im Moment werden Stundenoptionen an der Börse nicht gehandelt, obwohl sie für Kraftwerke, die selten im Geld sind, eine durchaus attraktive Einnahmequelle darstellen und auch für Nachfrager insbesondere für einzelne Peakstunden von Interesse sein könnten. Ein Versuch, Spotoptionen vom asiatischen Typ zu handeln, scheiterte an der Nord Pool Börse. Als möglicher Grund dafür kann u.U. die Schwierigkeit bei der Bewertung gesehen werden.

Ziel dieser Arbeit ist es in zwei Modellen, welche die Besonderheiten des Strommarktes berücksichtigen, Stundenoptionen zu bewerten sowie die Existenz und Eindeutigkeit einer Lösung für europäische Optionen nachzuweisen. Die Arbeit ist wie folgt gegliedert: In Kapitel 2 werden die notwendigen mathematischen Grundlagen gelegt, insbesondere wird auf die Halbgruppentheorie und die Theorie der Pseudo-Differentialoperatoren eingegangen. Kapitel 3 behandelt die ökonomischen Grundlagen und Kapitel 4-7 beschäftigen sich mit dem Kluge-Modell. In Kapitel 8 wird das Feller Process of NIG Type Modell (FPNIG) behandelt und schließlich werden die verschiedenen Ansätze in Kapitel 9 verglichen.

Kapitel 2

Mathematische Grundlagen

2.1 Stochastische Grundlagen

Im Verlauf dieser Arbeit wird die Dynamik der Strompreisentwicklung mit Hilfe einer stochastischen Differentialgleichung beschrieben. Um Optionen auf dieses bewerten zu können, sind einige Grundlagen notwendig, die im Folgenden dargelegt werden sollen. Mit Hilfe von Itô's Lemma wird aus der stochastischen Differentialgleichung, welche die Entwicklung des Spotkurses für Strom beschreibt, eine partielle Differentialgleichung für den Wert einer Option auf Strom hergeleitet. Zunächst muss hierfür die Definition eines stochastischen Prozesses gegeben werden.

Sei im Folgenden (Ω, F, P) ein vollständiger Wahrscheinlichkeitsraum und $\{F_t\}_{t \geq 0}$ eine Filtration auf diesem Raum ¹. Ein stochastischer Prozess $X = \{X(t, \omega) | t \in \mathbb{T}, \omega \in \Omega\}$ auf dem Wahrscheinlichkeitsraum ist eine Abbildung, die jedem Zeitpunkt und Elementarereignis eine reelle Zahl zuordnet. Es gibt zwei Möglichkeiten einen stochastischen Prozess anzusehen, einerseits als Funktion, die jedem Elementarereignis ω einen Zeitpfad $X(\cdot, \omega)$ zuordnet und andererseits als Funktion, die jedem Zeitpunkt t eine Zufallsvariable $X(t, \cdot)$ zuordnet.

Ein besonders wichtiger stochastischer Prozess ist der Wiener Prozess. Er wird in vielen Bereichen benutzt, um die stochastische Entwicklung von Vorgängen zu beschreiben. Ein typisches Anwendungsfeld ist die Beschreibung zufälliger Bewegungen von Teilchen in der Physik.

Definition 2.1 (Wiener Prozess). *Ein stochastischer Prozess W ist eine Brownsche Bewegung oder ein Wiener Prozess, falls folgende drei Bedingungen erfüllt sind:*

1. $W(0) = 0$ f.s..
2. $W(t)$ ist eine stetige Funktion von t .

¹Wir betrachten im Folgenden nur rechtstetige Filtrationen und an die Filtration adaptierte Prozesse.

3. W hat unabhängige und normalverteilte Zuwächse. Konkret gilt für beliebige $0 = t_0 < \dots < t_n$ mit $Y_i = W(t_i) - W(t_{i-1})$ für $i = 1, \dots, n$

- Y_1, Y_2, \dots, Y_n sind unabhängig,
- $\mathbb{E}(Y_i) = 0 \forall i$,
- $\mathbb{V}(Y_i) = t_i - t_{i-1} \forall i$.

Mit Hilfe der Brownschen Bewegung lässt sich nun eine große Klasse von Prozessen, die sogenannten $\hat{\text{Ito}}$ -Prozesse, definieren. Ein stochastischer Prozess (X) wird als $\hat{\text{Ito}}$ -Prozess bezeichnet, falls er folgende SDE erfüllt:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t,$$

wobei W_t eine Brownsche Bewegung ist. Die Funktionen $\mu(X_t, t)$ und $\sigma(X_t, t)$ müssen folgende Bedingung erfüllen:

$$\int_0^t (\sigma^2(x, \tau) + |\mu(x, \tau)|) d\tau < \infty \quad \forall x, t. \quad (2.1)$$

Für $\hat{\text{Ito}}$ -Prozesse lässt sich folgendes, für die Optionsbewertung fundamentale, Lemma zeigen.

Satz 2.1 ($\hat{\text{Ito}}$ s Lemma). Sei X ein $\hat{\text{Ito}}$ -Prozess und $g \in C^{2,1}(\mathbb{R} \times \mathbb{R}_+)$ eine Funktion der Zeit und von X , dann gilt

$$dg(X_t, t) = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t + \frac{\sigma^2}{2} \frac{\partial^2 g}{\partial x^2} dt. \quad (2.2)$$

$\hat{\text{Ito}}$ s Lemma stellt das stochastische Pendant zur Kettenregel in der Analysis dar. Die Kettenregel ist nicht direkt übertragbar, da ein Wiener Prozess nirgends differenzierbar ist und auch quadratische Variation ungleich Null hat. Die Differentialschreibweise in (2.2) ist zwar sehr anschaulich, jedoch mathematisch nicht exakt, da dW_t nicht wohldefiniert ist (vgl. [31, Kapitel 15]). Demnach entspricht (2.2) der intuitiven Veranschaulichung von

$$\begin{aligned} g(X_t, t) - g(X_0, 0) &= \int_0^t \frac{\partial g}{\partial s}(s, X_s) + \frac{\partial g}{\partial x}(s, X_s) ds \\ &+ \frac{1}{2} \int_0^t \sigma^2(s, X_s) \frac{\partial^2 g}{\partial x^2}(s, X_s) ds + \int_0^t \frac{\partial g}{\partial x}(s, X_s) \sigma(s, X_s) dW_s. \end{aligned}$$

Nun soll die Klasse der $\hat{\text{Ito}}$ -Prozesse mit drei Beispielen aus der Optionsbewertungstheorie veranschaulicht werden.

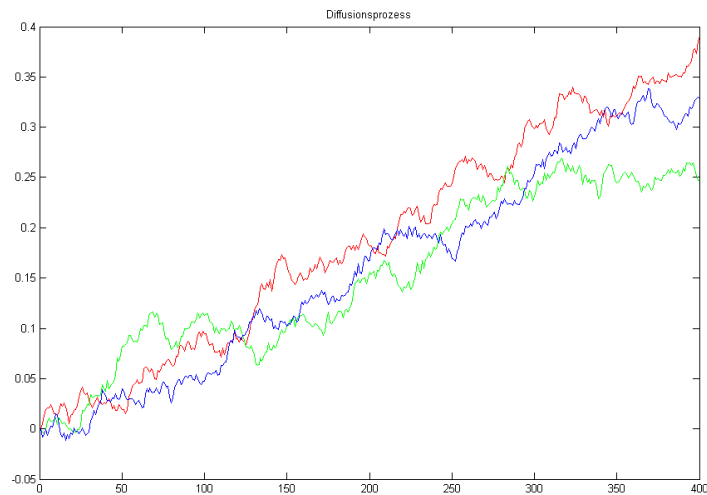


Abbildung 2.1: Beispielpfade arithmetisch Brownscher Bewegungen

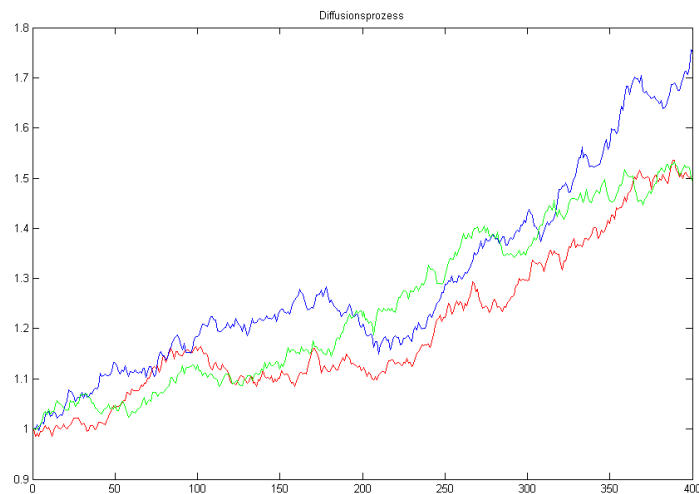


Abbildung 2.2: Beispielpfade geometrisch Brownscher Bewegungen

Beispiel 2.1. Zunächst betrachten wir einen stochastischen Prozess S , dessen Dynamik folgendermaßen beschrieben wird:

$$dS_t = \mu dt + \sigma dW_t,$$

Der Prozess S wird als arithmetisch Brownsche Bewegung bezeichnet. Einige Beispielpfade dieses Prozesses sind in Abbildung 2.1 dargestellt.

Beispiel 2.2. Für die Abbildung von Aktienkursentwicklungen wird als Standardmodell eine geometrisch Brownsche Bewegung benutzt. Eine geometrisch Brownsche Bewegung

(S) wird über folgende SDE definiert:

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

Gegenüber der arithmetisch Brownschen Bewegung hat sie den Vorteil, dass negative Preise ausgeschlossen werden können. Bei der geometrisch Brownschen Bewegung folgt die Entwicklung der Returns einer arithmetisch Brownschen Bewegung. In Abbildung 2.2 sind einige Beispielpfade dargestellt.

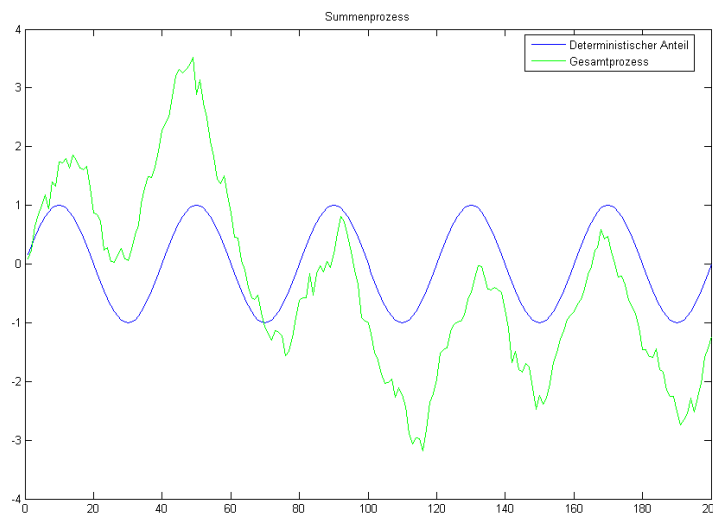


Abbildung 2.3: Beispielpfad für das Lucia-Schwartz Modell

Beispiel 2.3. Bei der Modellierung von Commodities ist oft das Einbeziehen einer saisonalen Komponente notwendig. Das einfachste Modell für die Modellierung der Entwicklung von Strompreisen ist mit Hilfe eines OU-Prozesses mit Mean Reversion. Dabei wird die Dynamik der Strompreisentwicklung durch folgende SDE beschrieben:

$$dX_t = (\mu(t) - X_t)dt + \sigma dW_t. \quad (2.3)$$

Die Funktion $\mu(t)$ wird als Mean Reversion Niveau bezeichnet und oft als periodische Funktion modelliert. Ein Beispielpfad ist in Abbildung 2.3 dargestellt. Dieses Modell wird im Folgenden als LS-Modell bezeichnet (vgl. [24]).

Eine Erweiterung von Itô-Prozessen stellt das Einbeziehen von Sprüngen dar. Dies kann mit Hilfe von Jump-Diffusion Prozessen geschehen. Dabei werden Sprünge in der Entwicklung des Underlyings mit Hilfe eines Poissonprozesses miteinbezogen. Die Dynamik eines Underlyings (S) unter der Berücksichtigung von Sprüngen wird folgender-

maßen beschrieben:

$$dS_t = \mu(S_{t-}, t)dt + \sigma(S_{t-}, t)dW_t + S_{t-}dN_t(J_t - 1). \quad (2.4)$$

Dabei ist N_t ein Poissonprozess mit Intensität λ und $\{J_t\}_{t \geq 0}$ ist eine Familie von Zufallsvariablen, welche die Verteilung der Sprunghöhe beschreibt (typische Verteilungen sind beispielsweise die Normalverteilung und die Exponentialverteilung). Durch das Subtrahieren von eins in (2.4) erreichen wir, dass der Stand des Prozesses S_t nach einem Sprung J -mal dem Stand des Prozesses vor dem Sprung entspricht. Alternativ kann die Entwicklung des Prozesses mit obigen Bezeichnungen auch folgendermaßen aufgeschrieben werden:

$$dS_t = \mu(S_{t-}, t)dt + \sigma(S_{t-}, t)dW_t + S_{t-} \left(d \sum_{i=0}^{N_t} (J_i - 1) \right)$$

Beispiel 2.4. Ein Beispiel für einen Jump-Diffusion Prozess stellt der Merton Prozess dar. Dieser wird über folgende SDE beschrieben:

$$dS_t = \mu S_{t-} dt + \sigma S_{t-} dW_t + S_{t-} \left(d \sum_{i=0}^{N_t} (J_i - 1) \right).$$

Einige Beispielpfade sind in Abbildung 2.4 dargestellt.

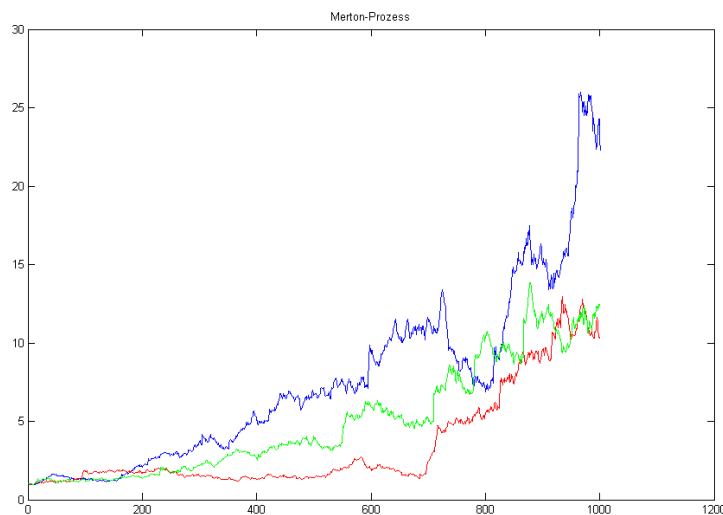


Abbildung 2.4: Beispielpfad für das Merton Modell

Für viele Anwendungen ist das Modellieren von mehreren Risikoquellen, dargestellt durch mehrere Jump-Diffusion Prozesse, unerlässlich. In Kapitel 4 werden wir den Kluge-Prozess näher untersuchen. Dieser ist ein Beispiel für einen Summenprozess, also einen Prozess, dessen Entwicklung durch mehrere Prozesse beschrieben wird. Seine Dynamik

lässt sich folgendermaßen darstellen:

$$\begin{aligned} S_t &= \exp(f(t) + X_t + Y_t) \\ dX_t &= -\alpha X_t dt + \sigma dW_t \\ dY_t &= -\beta Y_{t-} dt + J_t dN_t, \end{aligned}$$

wobei $\{J_t\}_{t \geq 0}$ eine Familie von exponentialverteilten Zufallsvariablen ist.

Für viele Prozesse ist aus Komplexitätsgründen die Darstellung und Rechnung mit SDEs kaum möglich. Eine Alternative stellt die Rechnung mit charakteristischen Funktionen dar. Im Folgenden wollen wir uns mit einer Klasse von Prozessen beschäftigen, für die dies besonders einfach möglich ist. Diese Klasse sind die Lévy-Prozesse. Um einen Lévy-Prozess zu definieren, sind jedoch einige Vorbereitungen notwendig.

Definition 2.2 (Fouriertransformierte). *Sei $f \in L^1(\mathbb{R})$, dann können die dazugehörige Fourier-Transformierte \hat{f} und die inverse Fouriertransformierte \tilde{f} folgendermaßen definiert werden:*

$$\begin{aligned} \hat{f}(\xi) &= \int_{\mathbb{R}} f(x) e^{-ix\xi} dx, \\ \tilde{f}(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} f(\xi) e^{i\xi x} d\xi. \end{aligned}$$

Die Fouriertransformation stellt auf dem Schwarzraum $S(\mathbb{R}) \subset L^1(\mathbb{R})$ der schnell fallenden Funktionen einen Automorphismus dar. Es lassen sich dort folgende Identitäten zeigen:

$$f = \hat{\hat{f}} = \tilde{\tilde{f}}. \quad (2.5)$$

Des Weiteren gelten folgende Rechenregeln:

$$\begin{aligned} \widehat{(D^{(\alpha)} f)}(\xi) &= i^\alpha \xi^\alpha \hat{f}(\xi), \\ \widehat{(x^{(\alpha)} f)}(\xi) &= i^\alpha (D^\alpha \hat{f})(\xi), \\ \widehat{f * g} &= \hat{f} \hat{g} \end{aligned}$$

für $f, g \in S(\mathbb{R})$. Nach dem Satz von Plancherel existiert eine Fortsetzung der Fouriertransformation für L^2 -Funktionen.

Nun können die charakteristische Funktion und die momenterzeugende Funktion definiert werden. Diese spielen bei der Bewertung in den beiden untersuchten Modelle eine wichtige Rolle. Die charakteristische Funktion stellt im Wesentlichen die Verallgemeinerung der Fouriertransformation auf Zufallsvariable dar.

Definition 2.3. *Für eine Zufallsvariable X auf \mathbb{R}^d wird ihre charakteristische Funktion*

$\Phi_X : \mathbb{R}^d \rightarrow \mathbb{R}$ definiert durch:

$$\Phi_X(z) = \mathbb{E}[e^{\langle iz, X \rangle}] = \int_{\mathbb{R}^d} e^{\langle iz, x \rangle} \mu_X(dx) \quad \text{für } \forall z \in \mathbb{R}^d.$$

Eine charakteristische Funktion charakterisiert die Verteilung der zugehörigen Zufallsvariable vollständig. Zwei Zufallsvariablen mit der gleichen charakteristischen Funktion haben auch die gleiche Verteilung. Es lässt sich elementar nachrechnen, dass folgende Eigenschaften gelten:

- $|\Phi_X(z)| \leq 1$
- $\Phi_X(-u) = \overline{\Phi_X(u)}$.

Mit Hilfe der charakteristischen Funktion lassen sich auch die Momente einer Zufallsvariable bestimmen.

Satz 2.2. 1. Falls $\mathbb{E}[|X|^n] < \infty$ gilt, dann folgt, dass Φ_X n -mal stetig differenzierbar in $z = 0$ ist. Die Momente von X lassen sich folgendermaßen darstellen:

$$\mathbb{E}[X^k] = \frac{1}{i^k} \frac{\partial^k \Phi_X}{\partial z^k}(0) \quad \forall k = 1, \dots, n.$$

2. Seien (X_1, \dots, X_n) unabhängige Zufallsvariablen, dann gilt für die charakteristische Funktion von $S_n = \sum_{i=1}^n X_i$

$$\Phi_{S_n}(z) = \prod_{i=1}^n \Phi_{X_i}(z). \quad (2.6)$$

Für den Kluge-Prozess ist die Eigenschaft (2.6) von großem Nutzen.

Definition 2.4. Sei X eine Zufallsvariable auf dem \mathbb{R}^d , dann wird die momenterzeugende Funktion M_X folgendermaßen definiert:

$$M_X(z) = \mathbb{E}[e^{\langle z, X \rangle}].$$

Im Gegensatz zur charakteristischen Funktion ist die momenterzeugende Funktion nicht immer wohldefiniert. Ist dies der Fall, dann gilt

$$M_X(u) = \Phi_X(-iu).$$

Falls die momenterzeugende Funktion in einer ϵ -Umgebung von Null definiert ist, dann folgt

$$\mathbb{E}[X^n] = \frac{\partial^n M_X}{\partial u^n}(0).$$

Dies wird im Folgenden benutzt, um eine explizite Formel für den Forwardpreis F_t^T im Kluge-Modell zu erhalten. Dieser ist nichts anderes als der erwartete Spotpreis, also das erste Moment der Zufallsvariable Strompreis S_T zu einem bestimmten Zeitpunkt t . Als Nächstes definieren wir den charakteristischen Exponenten einer Zufallsvariable. In der Literatur findet sich keine einheitliche Darstellung, wir folgen der Definition in [11].

Definition 2.5 (Charakteristischer Exponent). *Sei X eine Zufallsvariable auf dem \mathbb{R}^d , dann wird die Abbildung $\psi(\xi) : \mathbb{R}^d \rightarrow \mathbb{R}$ definiert durch:*

$$e^{-\psi(\xi)} = \mathbb{E}[e^{i\langle \xi, X \rangle}]$$

als charakteristischer Exponent der Zufallsvariable bezeichnet.

Nun können wir einen Lévy-Prozess definieren.

Definition 2.6 (Lévy-Prozess). *Sei $X = (X(t), t \geq 0)$ ein reellwertiger stochastischer Prozess auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , dann bezeichnen wir X als Lévy-Prozess, falls gilt*

1. $X(0) = 0$ f.s..
2. X hat unabhängige Zuwächse, d.h. für alle $n \in \mathbb{N}$ und alle $0 \leq t_1 \dots \leq t_n < \infty$ sind die Zufallsvariablen $(X(t_{j+1}) - X(t_j), 1 \leq j \leq n)$ unabhängig.
3. Die Zuwächse von X sind stationär, d.h. die Verteilung von $X(t+s) - X(t)$ ist unabhängig von t .
4. Die Pfade von X sind rechts-stetig auf $[0, \infty)$ und besitzen f.s. von links einen Limes für alle $t > 0$. Solche Prozesse werden als Cadlag-Prozesse bezeichnet².
5. X ist stochastisch stetig, d.h. für alle $t \geq 0$ und $\epsilon \geq 0$ gilt

$$\lim_{s \rightarrow t} P[|X(s) - X(t)| > \epsilon] = 0.$$

Diese Klasse von Prozessen enthält unter anderem den Wiener Prozess und den Poissonprozess. Der Vorteil der Lévy-Prozesse ist, dass sie einerseits eine sehr große Klasse von stochastischen Prozessen darstellen, andererseits aber eine eindeutige Repräsentation durch den charakteristischen Exponenten zulassen. Dies spielt bei der Bewertung von Derivaten auf Basis von Lévy-Prozessen eine wichtige Rolle.

Definition 2.7. *Als charakteristischer Exponent eines Lévy-Prozesses X wird analog zu Definition 2.5 die Funktion $\psi(\xi)$ bezeichnet mit*

$$e^{-t\psi(\xi)} = \mathbb{E}[e^{i\xi X(t)}].$$

²Cadlag ist ein französisches Akronym und steht für „continu à droite, limite à gauche“, was nichts anderes bedeutet als „rechts-stetig mit linksseitigem Limes“.

Der charakteristische Exponent eines Lévy-Prozesses besitzt, wie der folgende Satz zeigt, eine besonders einfache Darstellung.

Satz 2.3 (Lévy-Khintchine). [11, Satz 2.5] und [14, Theorem 3.2] Sei X ein Lévy-Prozess, dann lässt sich der charakteristische Exponent $\psi(\xi)$ darstellen als:

$$\psi(\xi) = \frac{1}{2}\sigma^2\xi^2 - i\mu\xi - \int_{\mathbb{R}} (e^{ix\xi} - 1 - ix\xi\mathbf{I}_{[-1,1]}(x))\nu(dx), \quad (2.7)$$

wobei $\mu, \sigma \in \mathbb{R}$ sind und $\nu(dx)$ ein Maß auf \mathbb{R} ist mit

$$\nu(\{0\}) = 0 \quad \text{und} \quad \int_{\mathbb{R}} \min(|x|^2, 1)\nu(dx) < \infty. \quad (2.8)$$

Die Darstellung (2.7) ist eindeutig.

Aus $\min(|x|^2, \epsilon) \leq \min(|x|^2, 1)$ folgt sofort mit (2.8) die Integrierbarkeit von $\nu(dx)$ auf $(-\epsilon, \epsilon)^c$. Zu beachten ist jedoch, dass die Endlichkeit des Maßes $\nu(dx)$ auf ganz \mathbb{R} nicht gefordert wird. Anschaulich entspricht dies, wie sich zeigen wird, der Möglichkeit einen Lévy-Prozess mit unendlich vielen beliebig kleinen Sprüngen zu definieren³. Das Tripel $(\sigma^2, \mu, \nu(dx))$ beschreibt einen Lévy-Prozess vollständig und wird als Lévy-Tripel bezeichnet. Für den charakteristischen Exponenten eines Lévy-Prozesses lässt sich folgender Satz zeigen:

Satz 2.4. Sei $\psi(\xi)$ der charakteristische Exponent eines Lévy-Prozesses X , dann ist $\psi(\xi)$ stetig für alle $\xi \in \mathbb{R}$ und es gilt folgende nützliche Abschätzung:

$$|\psi(\xi)| \leq C(1 + |\xi|^2). \quad (2.9)$$

Beweis. siehe [3, S.30f]. □

Beispiel 2.5. Der charakteristische Exponent einer arithmetisch Brownschen Bewegung mit Drift μ und Standardabweichung σ ist

$$\psi(\xi) = \frac{1}{2}t\sigma^2 - i\mu\xi.$$

Das zugehörige Lévy-Tripel ist demzufolge $(\sigma^2, \mu, 0)$. Aufgrund dieser Tatsache wird der Parameter σ als Diffusionsparameter eines Lévy-Prozesses bezeichnet. Der Parameter μ lässt sich als linearer Trend interpretieren (vgl. [14, S.83]).

Beispiel 2.6. Der charakteristische Exponent eines Poissonprozesses mit Intensität λ ist

$$\psi(\xi) = \int_{\mathbb{R}} (e^{ix\xi} - 1 - ix\xi\mathbf{I}_{[-1,1]}(x))\lambda\delta(1) dx,$$

³Die Existenz von unendlich vielen großen Sprüngen ist auf Grund der cadlag-Eigenschaft des Prozesses ausgeschlossen.

wobei $\delta(1)$ für das Diracmaß steht.

Das zugehörige Lévy-Tripel ist $(0, 0, \lambda\delta(1))$. Daraus wird klar, dass das Maß $\nu(dx)$ den Sprunganteil eines Lévy-Prozesses beschreibt. Für einen Lévy-Prozess X und eine Teilmenge $A \subset \mathbb{R} \setminus B_\epsilon(0)$ (mit $\epsilon > 0$ beliebig) folgen Sprünge, deren Größe ein Element aus A ist, auf dem Zeitintervall $[0, 1]$ einem Poissonprozess mit Intensität $\int_A \nu(dx)$. Grundsätzlich ergeben sich daraus zwei Gruppen von Lévy-Prozessen. Einerseits Prozesse mit unendlicher Aktivität und andererseits Prozesse mit endlicher Aktivität. Die erste Gruppe besitzt in jedem beliebig kleinen Zeitintervall eine unendliche Anzahl an beliebig kleinen Sprüngen, d.h. $\nu(dx)$ ist nicht integrierbar. Solche Prozesse brauchen in der Regel keinen Diffusionsanteil um eine nichttriviale Dynamik zu beschreiben. Bei der zweiten Gruppe von Prozessen werden Sprünge als seltene Ereignisse angesehen und die Dynamik wird im Wesentlichen über den Diffusionsparameter abgebildet. Die Lévy-Dichte wird in solchen Fällen integrierbar modelliert.

Beispiel: Der NIG-Prozess

Nun wird ein spezieller Lévy-Prozess, der NIG-Prozess, näher untersucht, der für die Modellierung der Entwicklung von Commoditypreisen besonders geeignet scheint⁴. In Kapitel 8 werden wir diesen Prozess zum FPNIG verallgemeinern, um noch besser die Charakteristika von Strom abbilden zu können.

Zunächst wird die NIG-Verteilung beschrieben, um danach einen NIG-Prozess definieren zu können. Die Dichte der NIG-Verteilung kann explizit angegeben werden:

$$f(x; \mu, \alpha, \beta, \delta) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right) \frac{K_1\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\sqrt{\delta^2 + (x - \mu)^2}}, \quad (2.10)$$

wobei $\mu, \alpha, \beta, \delta \in \mathbb{R}$ mit $\alpha > |\beta|$ und $\delta > 0$ gilt. K_1 ist die modifizierte Bessel Funktion 3. Art der Ordnung 1 (siehe Abschnitt B.1). Die Dichtefunktion lässt sich in diesem Fall zwar noch explizit angeben, jedoch ist die obige Definition nicht besonders anschaulich. Bei vielen Lévy-Prozessen ist es nicht möglich die Dichte der Verteilung explizit anzugeben. In den meisten Fällen ist die charakteristische Funktion wesentlich einfacher handhabbar. Die charakteristische Funktion einer NIG-verteilten Zufallsvariable ist gegeben durch:

$$\Phi(z) = \exp(\mu iz) \frac{\exp\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}{\exp\left(\delta\sqrt{\alpha^2 - (\beta + iz)^2}\right)}. \quad (2.11)$$

⁴vgl. [8].

Somit ergibt sich der charakteristische Exponent als:

$$\psi(\xi) = -i\mu\xi + \delta \left[\sqrt{\alpha^2 - (\beta^2 + i\xi)^2} - \sqrt{\alpha^2 - \beta^2} \right]. \quad (2.12)$$

Eine wichtige Eigenschaft der NIG-Verteilung ist die im folgenden Satz formulierte Abgeschlossenheit unter der Faltung.

Satz 2.5 (Faltungseigenschaft der NIG-Verteilung). *Seien X und Y zwei Zufallsvariablen mit $X \sim NIG(\mu_1, \alpha, \beta, \delta_1)$ und $Y \sim NIG(\mu_2, \alpha, \beta, \delta_2)$ dann gilt:*

$$X + Y \sim NIG(\mu_1 + \mu_2, \alpha, \beta, \delta_1 + \delta_2).$$

Beweis. Der Beweis folgt sofort aus der Darstellung der charakteristischen Funktion und (2.6). □

Die Momente einer NIG-verteilten Zufallsgröße X lassen sich elementar bestimmen:

$$\begin{aligned} \mathbb{E}(X) &= \mu + \delta \frac{\beta}{\gamma} & \mathbb{V}(X) &= \delta \frac{\alpha^2}{\gamma^3} \\ \mathbb{S}(X) &= 3 \frac{\beta}{\alpha \sqrt{\delta \gamma}} & \mathbb{K}(X) &= 3 + 3 \left(1 + 4 \left(\frac{\beta}{\alpha} \right)^2 \right) \frac{1}{\delta \gamma}, \end{aligned}$$

wobei $\gamma = \sqrt{\alpha^2 - \beta^2}$ ist. Hiermit lassen sich die Parameter der NIG-Verteilung gut interpretieren. So bestimmt α die Dicke der Tails und β bestimmt das Ausmaß der Asymmetrie. Für $\beta = 0$ ist die Verteilung symmetrisch. Der Parameter δ kann als Skalierungsfaktor interpretiert werden und μ beschreibt den linearen Trend.

Die Verteilung besitzt semi-heavy Tails, d.h. für große und kleine x fällt die Dichtefunktion langsamer als die Dichtefunktion der Normalverteilung⁵. Die Geschwindigkeit des Abfalls wird von $\alpha + \beta$ und $\alpha - \beta$ gesteuert. Es gilt

$$f(x; \mu, \alpha, \beta, \delta) \sim |x|^{-\frac{3}{2}} \exp((-\alpha + \beta)x) \quad \text{für } x \rightarrow +\infty \quad (2.13)$$

und

$$f(x; \mu, \alpha, \beta, \delta) \sim |x|^{-\frac{3}{2}} \exp((\alpha + \beta)x) \quad \text{für } x \rightarrow -\infty. \quad (2.14)$$

In den Abbildungen 2.5 und 2.6 sind die Dichtefunktionen NIG-verteilter Zufallsvariablen für unterschiedliche Parameterkombinationen dargestellt. Wie die Abbildung veranschaulicht, ist die Verteilung äußerst flexibel und lässt dem Modellierer große Freiräume.

Nun sind wir in der Lage einen NIG-Prozess zu definieren.

Definition 2.8. *Ein stochastischer Prozess X heißt NIG-Prozess, falls*

⁵Bei Verteilungen mit heavy Tails fällt die Dichte Funktion nur noch polynomial für große und kleine x .

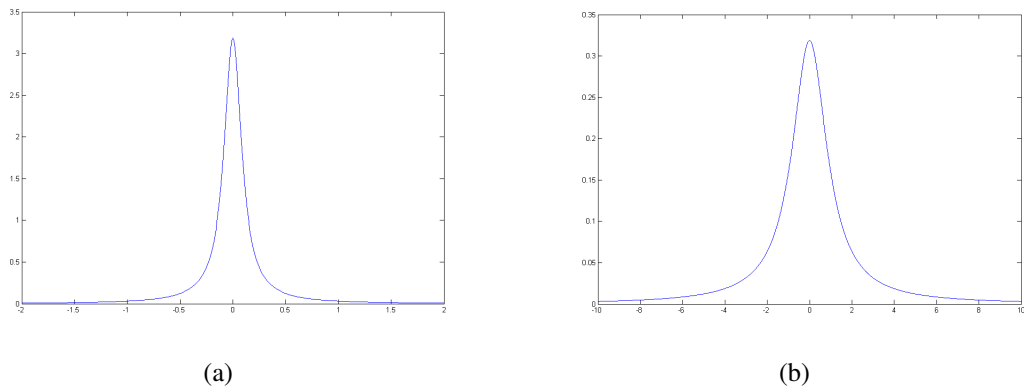


Abbildung 2.5: NIG-Dichte mit folgenden Parameteren (a) $\alpha = 0.01, \beta = 0, \mu = 0, \delta = 0.1$ und (b) $\alpha = 0.001, \beta = 0, \mu = 0, \delta = 1$

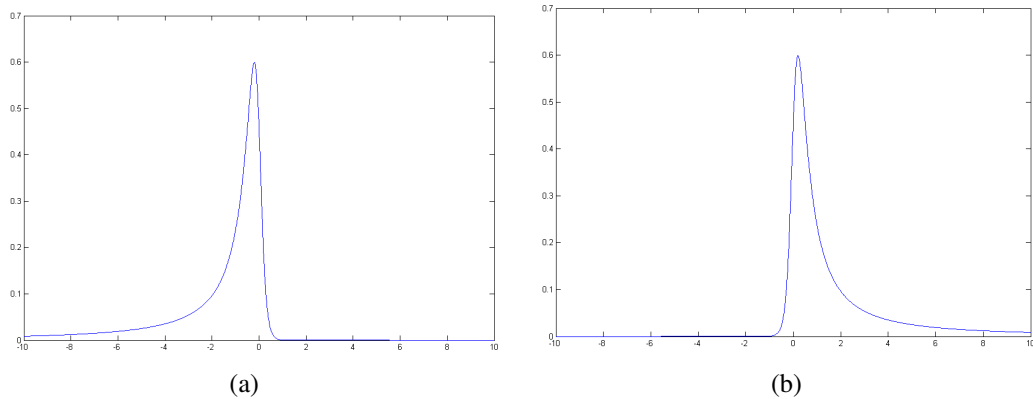


Abbildung 2.6: NIG-Dichte mit folgenden Parameteren (a) $\alpha = 3, \beta = 2.99, \mu = 0, \delta = 0.4$ und (b) $\alpha = 3, \beta = -2.99, \mu = 0, \delta = 0.4$

1. $X_0 = 0$ f.s.,
2. $X_t \sim NIG(\mu t, \alpha, \beta, t\delta)$ mit μ, α, β und δ wie in (2.10).

Insbesondere lässt sich nachrechnen, dass jeder NIG-Prozess ein Lévy-Prozess mit charakteristischem Exponenten (2.12) ist (vgl.[4]). Der NIG-Prozess ist ein reiner Jump-Prozess ohne Diffusionsanteil. Das Lévy-Tripel ist gegeben durch $(\gamma, 0, \nu_{NIG}(dx))$ mit

$$\gamma = \mu + \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x) dx$$

und

$$\nu_{NIG}(dx) = \frac{\delta\alpha \exp(\beta x) K_1(\alpha |x|)}{\pi |x|} dx.$$

Um die besonderen Eigenschaften von Strom zu berücksichtigen, werden wir die Klasse der Levy-Prozesse auf Feller-Prozesse vom Levy-Typ erweitern, bei denen die Parameter des Lévy-Tripels ortsabhängig definiert werden. Einen solchen Prozess, den FPNIG, werden wir in Abschnitt 8 untersuchen.

Nachdem wir uns mit der Modellierung der für diese Arbeit wesentlichen Prozesse be-

schäftigt haben, sollen nun einige Grundlagen für die Bewertung von Optionen in den Modellen dargestellt werden. Wie sich zeigen wird, führt die Bewertung von Optionen auf PDEs (PIDEs) und PDOs. Sowohl bei theoretischen Überlegungen als auch bei der numerischen Umsetzung ist die im Folgenden dargestellte Halbgruppentheorie von großem Nutzen.

2.2 Halbgruppentheorie

Die Idee hinter der Halbgruppentheorie ist, eine PDE als gewöhnliche DGL auf einem Banachraum aufzufassen und sich der Methoden zur Lösung gewöhnlicher DGL zu bedienen, um diese zu lösen. Mit Hilfe der Halbgruppentheorie soll die Existenz einer Lösung einer im Verlauf der Arbeit wichtigen PIDE gezeigt werden.

Beispiel 2.7. [29, S.395]

Es wird folgende PDE betrachtet:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \text{ für } x \in (0, 1) \text{ und } t > 0.$$

Mit Randwerten:

$$u(1, t) = u(0, t) = 0 \text{ für } t > 0$$

und Anfangswerten:

$$u(x, 0) = u_0(x) \text{ für } x \in [0, 1].$$

Dieses Problem lässt sich auf entsprechenden Räumen als gewöhnliche Differentialgleichung auffassen. Sei $X = L^2(0, 1)$, $A = \frac{d^2}{dx^2}$ und $D(A) = \{u \in H^2(0, 1) \mid u(1) = u(0) = 0\}$. Dann lässt sich die obige PDE schreiben als gewöhnliche DGL der folgenden Art:

$$\frac{\partial u}{\partial t} = Au \quad u(0) = u_0. \quad (2.15)$$

Wäre nun A eine Matrix und kein allgemeiner Operator und $u = u(t)$, dann ließe sich die Lösung als $\exp(At)u_0$ schreiben. Was $\exp(At)$ ist, wird im Folgenden erläutert. Wir folgen dabei den Erläuterungen aus [29], [7] und [2].

Definition 2.9. Sei X ein Banachraum. Eine Familie $\{T(t)\}_{t \geq 0}$ von beschränkten linearen Operatoren auf X heißt stark-stetige oder C_0 -Halbgruppe, falls

1. $T(t + s) = T(t)T(s)$,
2. $T(0) = I$,
3. $\forall x \in X$ ist $[0, \infty) \ni t \mapsto T(t)x$ stetig.

Nun zeigen wir eine Abschätzung für die Norm einer C_0 -Halbgruppe.

Satz 2.6. [29, Satz 12.7] Sei $\{T(t)\}_{t \geq 0}$ eine C_0 -Halbgruppe beschränkter linearer Operatoren auf einem Banachraum X . Dann existiert der Limes

$$\omega_0 = \lim_{t \rightarrow \infty} \frac{\log \|T(t)\|}{t},$$

wobei dieser auch $-\infty$ sein kann. Für alle $\gamma > \omega_0$ existiert ein M_γ mit $\|T(t)\| \leq M_\gamma \exp(\gamma t)$.

Beweis. siehe Anhang A.1. □

Es wird ω_0 als der Typ der Halbgruppe bezeichnet.

Als nächstes definieren wir den infinitesimalen Erzeuger einer C_0 -Halbgruppe. Dieser wird eine wesentliche Rolle beim Existenzbeweis für die Lösung einer PIDE im Kluge-Modell spielen. Es wird sich zeigen, dass für die Lösbarkeit einer gewöhnlichen Differentialgleichung der Art von (2.15) entscheidend ist, ob der Operator infinitesimaler Erzeuger einer Halbgruppe ist oder nicht.

Definition 2.10. Sei $\{T(t)\}_{t \geq 0}$ eine C_0 -Halbgruppe beschränkter linearer Operatoren auf einem Banachraum X , dann ist der infinitesimale Erzeuger A der Halbgruppe folgendermaßen definiert:

$$Ax = \lim_{h \downarrow 0} \frac{T(h)x - x}{h}. \tag{2.16}$$

Wobei der Definitionsbereich von A die Menge aller $x \in X$ ist, für die der Grenzwert existiert.

Satz 2.7. [29, Lemma 12.11] Sei A der infinitesimale Erzeuger einer C_0 -Halbgruppe $\{T(t)\}_{t \geq 0}$, dann gilt

1. $\forall x \in X$

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_t^{t+h} T(s)x \, ds = T(t)x \tag{2.17}$$

2. $\forall x \in X$ und alle $t > 0$

$$\int_0^t T(s)x \, ds \in D(A) \quad \text{und} \quad A\left(\int_0^t T(s)x \, ds\right) = T(t)x - x \tag{2.18}$$

3. $\forall x \in D(A)$ ist $T(t)x \in D(A)$. Außerdem ist die Funktion $[0, \infty) \ni t \mapsto T(t)x$ differenzierbar und es gilt:

$$\frac{d}{dt}T(t)x = AT(t)x = T(t)Ax \tag{2.19}$$

4. $\forall x \in D(A)$

$$T(t)x - T(s)x = \int_s^t T(\tau)Ax \, d\tau = \int_s^t AT(\tau)x \, d\tau \quad (2.20)$$

Beweis. Ad 1:

Dies folgt aus der Stetigkeit der Abbildung $[0, \infty) \ni t \mapsto T(t)x$.

Ad 2:

Wir betrachten

$$\begin{aligned} \frac{T(h) - I}{h} \int_0^t T(s)x \, ds &= \frac{1}{h} \int_0^t (T(s+h) - T(s))x \, ds \\ &= \underbrace{\frac{1}{h} \int_t^{t+h} T(s)x \, ds}_{\xrightarrow{\lim_{h \downarrow 0}} T(t)x} - \underbrace{\frac{1}{h} \int_0^h T(s)x \, ds}_{\xrightarrow{\lim_{h \downarrow 0}} x}. \end{aligned}$$

Daraus folgt Teil 2, da $\lim_{h \downarrow 0} \frac{T(h) - I}{h} = A$.

Ad 3:

Wir betrachten:

$$\frac{T(t+h)x - T(t)x}{h} = \frac{T(h) - I}{h} T(t)x = T(t) \frac{T(h)x - x}{h}$$

und

$$\frac{T(t)x - T(t-h)x}{h} = \frac{T(h) - I}{h} T(t-h)x = T(t-h) \frac{T(h)x - x}{h}.$$

Betrachten wir nun den Limes $h \downarrow 0$, so ergibt sich die Aussage.

Ad 4:

Folgt sofort aus 3 mit dem Hauptsatz der Differential- und Integralrechnung. \square

Als nächstes zeigen wir eine notwendige Eigenschaft eines infinitesimalen Erzeugers. Diese wird uns dabei helfen, Operatoren daraufhin zu überprüfen, ob sie eine C_0 -Halbgruppe erzeugen.

Satz 2.8. [29, Satz 12.12] Sei A der infinitesimale Erzeuger einer C_0 -Halbgruppe auf X . Dann ist $D(A)$ dicht in X und A ist abgeschlossen.

Beweis. Nach (2.17) gilt:

$$x = \lim_{h \downarrow 0} \frac{1}{h} \int_0^h T(s)x \, ds.$$

Dabei ist die rechte Seite für alle h mit $h > 0$ nach (2.18) in $D(A)$ enthalten. Da x beliebig war, ist $D(A)$ in X dicht. Sei nun $(x_n)_{n \in \mathbb{N}} \in D(A)$, $x_n \rightarrow x$ und $Ax_n \rightarrow y$. Zu zeigen

ist $x \in D(A)$ und $Ax = y$. Dazu betrachten wir nach (2.20)

$$T(h)x_n - x_n = \int_0^h T(s)Ax_n ds.$$

Für $n \rightarrow \infty$ folgt auf Grund der Stetigkeit von $t \mapsto T(t)x$ und der Beschränktheit von $T(t)$ auf jedem endlichen Intervall:

$$T(h)x - x = \int_0^h T(s)y ds.$$

Teilen wir nun beide Seiten durch h und lassen h gegen Null laufen, so folgt die Aussage aus (2.17) und (2.16). \square

Kommen wir nun zurück zu den partiellen Differentialgleichungen. Als nächstes zeigen wir, dass sich eine Lösung einer PDE, geschrieben in der Form

$$\frac{du}{dt} = Au + f(t), \quad u(0) = u_0, \tag{2.21}$$

wobei A der infinitesimale Erzeuger einer C_0 -Halbgruppe ist, mit Hilfe der zugehörigen Halbgruppe ausdrücken lässt.

Satz 2.9. [29, Satz 12.14] Sei X ein Banachraum, $u_0 \in D(A)$, $f \in C([0, T]; X)$ und $u \in C^1([0, T]; X) \cap C([0, T]; D(A))$, dann lässt sich u folgendermaßen darstellen:

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s) ds \tag{2.22}$$

Beweis. Sei $g(s) = T(t-s)u(s)$. Dann gilt

$$\begin{aligned} \frac{dg}{ds} &\stackrel{(2.19)}{=} -AT(t-s)u(s) + T(t-s)\frac{du}{dt}(s) \\ &\stackrel{(2.21)}{=} -AT(t-s)u(s) + T(t-s)(Au(s) + f(s)) \stackrel{(2.19)}{=} T(t-s)f(s). \end{aligned}$$

Durch Integration folgt:

$$g(t) - g(0) = u(t) - T(t)u_0 = \int_0^t T(t-s)f(s) ds.$$

Und das war zu zeigen. \square

Die durch (2.22) definierte Lösung wird schwache Lösung des Problems (2.21) genannt. Unter unterschiedlichen Voraussetzungen an f ändert sich die Regularität der schwachen Lösung (vgl.[26]). Unter bestimmten Voraussetzungen lässt sich zeigen, dass eine schwache Lösung eine Lösung im klassischen Sinne ist.

Wir wollen uns jetzt mit der Frage befassen, wann ein Operator A der Erzeuger einer analytischen Halbgruppe ist.

Satz 2.10 (Hille-Yosida-Theorem). [29, Satz 12.17] Sei A ein Operator auf einem Banachraum X . Dann ist A genau dann der Erzeuger einer C_0 -Halbgruppe mit $\|T(t)\| \leq M \exp(\omega t)$, wenn

1. $D(A)$ dicht in X und A abgeschlossen ist.
2. Alle $\lambda \in \mathbb{R}$ mit $\lambda > \omega$ in der Resolventenmenge enthalten sind und $\|R_\lambda(A)^n\| \leq \frac{M}{(\lambda - \omega)^n}$ für alle $n \in \mathbb{N}$ gilt.

Beweis. siehe Anhang A.3. Aus dem Beweis wird klar, dass alle $\lambda \in \mathbb{C}$ mit $\operatorname{Re}(\lambda) > \omega$ in der Resolventenmenge enthalten sind. Es folgt ebenfalls

$$T(t) = \lim_{n \rightarrow \infty} \left(I - \frac{t}{n} A \right)^{-n}.$$

Zusammen mit (2.19) erklärt dies auch die Schreibweise $T(t) = \exp(At)$. □

Damit haben wir zwar ein Verfahren, um zu überprüfen, ob ein Operator der infinitesimale Erzeuger einer analytischen Halbgruppe ist, jedoch ist dieses nicht einfach handhabbar, da alle natürlichen Potenzen von $R_\lambda(A)$ betrachtet werden müssen. Wichtige Spezialfälle der obigen Situation sind Halbgruppen $\{T(t)\}_{t \geq 0}$, für die die Abschätzung mit $M = 1$ für $n = 1$ und somit auch für alle $n \in \mathbb{N}$ gilt und jene, für die zusätzlich $\omega = 0$ gilt.

Definition 2.11. Eine C_0 -Halbgruppe heißt *Quasikontraktionshalbgruppe*, falls $\|R_\lambda(A)^n\| \leq \frac{1}{(\lambda - \omega)^n}$ für ein $\omega \in \mathbb{R}$. Sie heißt *Kontraktionshalbgruppe*, falls $\|R_\lambda(A)^n\| \leq 1$.

Wir definieren auch eine besondere Kontraktionshalbgruppe, die bei der Analyse des FPNIG-Modells eine besondere Rolle spielen wird.

Definition 2.12 (Feller-Halbgruppe). Wir bezeichnen eine C_0 -Halbgruppe $\{T(t)\}_{t \geq 0}$ auf $C_\infty(\mathbb{R}; \mathbb{R})$ als *Feller-Halbgruppe*, falls diese eine Kontraktion ist und positivitätserhaltend, d.h. aus $u > 0$ folgt $Tu > 0$.

Nun wird ein Satz formuliert, der es ermöglicht, einfacher als in Satz 2.10 zu überprüfen, ob ein Operator eine Quasikontraktionshalbgruppe auf einem Hilbertraum erzeugt.

Satz 2.11 (Lumer-Phillips). [29, Satz 12.22] und [26, Satz 1.4.3] Sei A ein linearer Operator auf einem Hilbertraum H , dann erzeugt A eine Quasikontraktionshalbgruppe, falls gilt:

1. $D(A)$ ist dicht in H .
2. $\Re(x, Ax) \leq w(x, x)$ für alle $x \in D(A)$.
3. Es gibt ein $\lambda_0 > \omega$, so dass $A - \lambda_0 I$ surjektiv ist.

Beweis. Wir zeigen, dass aus den obigen Bedingungen die Voraussetzungen von Satz 2.10 folgen. Für $\lambda > \omega$ gilt:

$$\|(A - \lambda I)x\| \|x\| \stackrel{\text{Ca.-Schw.}}{\geq} |\Re(x, (\lambda I - A)x)| \geq (\lambda - \omega) \|x\|^2. \quad (2.23)$$

Also gilt $\|(A - \lambda I)x\| \geq (\lambda - \omega) \|x\|$ und somit ist $A - \lambda I$ injektiv. Falls wir zeigen können, dass $A - \lambda I$ surjektiv ist für alle $\lambda > \omega$, so existiert die Resolvente und es gilt folgende Abschätzung für die Norm $\frac{\|R_\lambda(A)y\|}{\|y\|} \leq (\lambda - \omega)^{-1}$. Da die Resolvente für λ_0 existiert und beschränkt ist, also auch abgeschlossen, ist $A - \lambda_0 I$ abgeschlossen und somit auch A . Um nun die Surjektivität zu zeigen, beweisen wir, dass die Menge Λ offen und abgeschlossen in (ω, ∞) ist:

$$\Lambda = \{\lambda \in (\omega, \infty) \mid (A - \lambda I) \text{ ist surjektiv}\}.$$

Da wir wissen, dass Λ nicht leer ist, würde dies den Beweis vollenden.

Sei nun $\lambda \in \Lambda$ beliebig. Dann wissen wir, dass λ in der Resolventenmenge enthalten ist, da $A - \lambda I$ injektiv ist. Weiterhin ist die Resolventenmenge offen, da A abgeschlossen ist. Betrachten wir jetzt den Schnitt der reellen Achse mit einer Umgebung von λ , die in der Resolventenmenge enthalten ist, so ist diese sicherlich in Λ . Also ist Λ offen. Um zu zeigen, dass Λ auch abgeschlossen ist, wählen wir eine beliebige Folge $(\lambda_n)_{n \in \mathbb{N}}$ in Λ mit $\lambda_n \rightarrow \lambda > \omega$ und zeigen $\lambda \in \Lambda$. Dazu betrachten wir ein beliebiges $y \in H$. Zu diesem y existiert eine Folge $(x_n)_{n \in \mathbb{N}}$ aus $D(A)$ mit $Ax_n - \lambda_n x_n = y$, da λ_n in der Resolventenmenge enthalten ist. Es wird nun gezeigt, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Dazu wird zunächst bewiesen, dass die Folge der x_n beschränkt ist.

$$(\lambda_n - \omega) \|x_n\| \stackrel{(2.23)}{\leq} \|Ax_n - \lambda_n x_n\| = \|y\|$$

Es gibt also eine Konstante $C > 0$ mit $\|x_n\| < C$. Die Differenz aus x_n und x_m lässt sich nun wie folgt abschätzen.

$$(\lambda_m - \omega) \|x_n - x_m\| \stackrel{(2.23)}{\leq} \|(A - \lambda_m I)(x_n - x_m)\| \leq |\lambda_n - \lambda_m| \|x_n\|$$

Da x_n beschränkt ist, lässt sich jetzt für alle $\epsilon > 0$ ein N finden d.d. $\|x_n - x_m\| < \epsilon$ für alle $n, m > N$ gilt. Sei x nun der Grenzwert der Folge. Dann gilt $Ax_n \rightarrow y + \lambda x$. Da A abgeschlossen ist, folgt $x \in D(A)$ und $Ax - \lambda x = y$. Und y war aber beliebig gewählt, somit ist $A - \lambda I$ surjektiv und somit $\lambda \in \Lambda$. Also ist Λ offen und abgeschlossen. \square

Kommen wir nun zu einer Klasse von Halbgruppen, die in unserem Spezialfall mit einem Operator $A = A(t)$ von besonderem Interesse sind.

Definition 2.13. Eine C_0 -Halbgruppe $\{T(t)\}_{t \geq 0}$ auf X heißt *analytische Halbgruppe* falls folgende Bedingungen erfüllt sind:

1. Es existiert ein $\theta \in (0, \frac{\pi}{2})$ d.d. $T(t) \in \mathcal{L}(X)$ auf $\Delta_\theta = \{0\} \cup \{t \in \mathbb{C} \mid |\arg(t)| < \theta\}$ ausgeweitet werden kann und die Eigenschaften aus Definition 2.10 für alle $t \in \Delta_\theta$ erfüllt sind.
2. Für $t \in \Delta_\theta$ ist $T(t)$ analytisch in t .

Als nächstes wollen wir ein Kriterium finden, welches es ermöglicht, einen Operator daraufhin zu überprüfen, ob er eine analytische Halbgruppe erzeugt.

Satz 2.12. [29, Satz 12.31] Sei A abgeschlossen und $D(A)$ dicht in X , dann erzeugt A genau dann eine analytische Halbgruppe, wenn ein $\omega \in \mathbb{R}$ existiert d.d. alle λ mit $\operatorname{Re}(\lambda) > \omega$ in der Resolventenmenge von A enthalten sind und es eine Konstante C gibt mit $\|R_\lambda(A)\| \leq \frac{C}{|\lambda - \omega|}$ für $\operatorname{Re}(\lambda) > \omega$.

Beweis. Ist dies der Fall, so ist die Resolvente sogar größer. Sie enthält alle λ mit $|\arg(\lambda - \omega)| \leq \frac{\pi}{2} + \delta$ für ein $\delta > 0$. In diesem Fall lässt sich die Halbgruppe folgendermaßen darstellen:

$$T(t) = \frac{1}{2\pi i} \int_\Gamma e^{\lambda t} (\lambda I - A)^{-1} d\lambda,$$

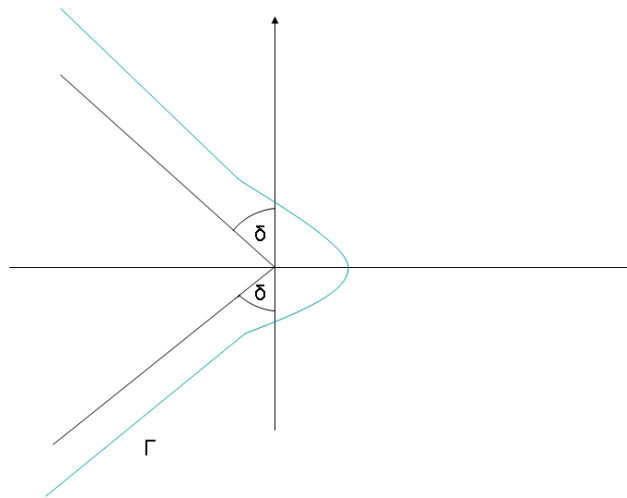
wobei Γ ein Weg von $-e^{i\phi}\infty$ nach $e^{i\phi}\infty$ ist, d.d. Γ vollständig in der Menge $\{|\arg(\lambda - \omega)| \leq \phi\}$ enthalten ist. Dabei ist $\phi \in (\frac{\pi}{2}, \frac{\pi}{2} + \delta)$. Der Integrationsweg ist in Abbildung 2.7 dargestellt. \square

Für den Operator, der für den Existenzbeweis im Kluge-Modell benötigt wird, sind die Voraussetzungen von Satz 2.12 besonders einfach zu zeigen. Dies ist die Aussage des folgenden Satzes.

Satz 2.13. Sei H ein Hilbertraum und A ein selbstadjungierter Operator mit dichtem Definitionsbereich, der negativ semidefinit ist und für den $A - \lambda I$ surjektiv für $\lambda > 0$ gilt. Dann erzeugt A eine analytische Halbgruppe.

Beweis. A erfüllt die Voraussetzungen von Satz 2.11 und erzeugt damit eine C_0 -Halbgruppe. Aus Satz 2.10 folgt, dass alle λ mit $\operatorname{Re}(\lambda) > 0$ in der Resolventenmenge enthalten sind. Es bleibt also für die Voraussetzungen des Satzes 2.12 noch zu zeigen, dass $\|R_\lambda(A)\| \leq \frac{1}{|\lambda|}$ gilt. Dies folgt aus:

$$\begin{aligned} \|(A - \lambda I)x\| \|x\| &\geq |((A - \lambda I)x, x)| = \sqrt{(\operatorname{Re}((A - \lambda I)x, x))^2 + (\operatorname{Im}((A - \lambda I)x, x))^2} \\ &\geq \sqrt{(\operatorname{Re}(\lambda))^2 \|x\|^4 + (\operatorname{Im}(\lambda))^2} = |\lambda| \|x\|^2. \end{aligned}$$

Abbildung 2.7: Integrationsweg Γ

Dies kann für λ aus der Resolventenmenge zu $\|R_\lambda(A)\| \leq \frac{1}{|\lambda|}$ umgestellt werden. \square

Die wesentlichen Grundlagen für das Führen des Existenzbeweises im Kluge-Modell wurden bereits dargestellt. Die Strategie bei dem Beweis ist es, den Operator A zu zerlegen und zu zeigen, dass ein Teil des Operators eine analytische Halbgruppe erzeugt. Dann werden wir zeigen, dass durch die Hinzunahme des vorher nicht betrachteten Teils diese Eigenschaft nicht verloren geht. Um eine solche Aussage zeigen zu können, brauchen wir ein Theorem, das uns sagt, inwieweit ein Erzeuger einer analytischen Halbgruppe „gestört“ werden kann, um weiterhin eine analytische Halbgruppe zu erzeugen.

Satz 2.14. [29, Satz 12.37] und [26, 3.2.1] Sei A der infinitesimale Erzeuger einer analytischen Halbgruppe. Dann gilt für alle B mit:

1. B ist abgeschlossen und $D(B) \supseteq D(A)$,
2. $\|Bu\| \leq a\|Au\| + b\|u\|$ für $u \in D(A)$ mit $a < \frac{1}{C(A)+1}$, wobei C die Konstante aus der Abschätzung für die Norm von A aus Satz 2.12 ist,

dass $A + B$ infinitesimaler Erzeuger einer analytischen Halbgruppe ist.

Beweis. Da A eine analytische Halbgruppe erzeugt, existieren $\omega \in \mathbb{R}$ und $C > 0$ mit $\|R_\lambda(A)\| \leq \frac{C}{|\lambda-\omega|}$ für $\operatorname{Re}(\lambda) > \omega$. $BR_\lambda(A)$ ist beschränkt und die Norm lässt sich folgen-

dermaßen abschätzen:

$$\begin{aligned} \|BR_\lambda(A)u\| &\stackrel{n.V.}{\leq} a\|AR_\lambda(A)u\| + b\|R_\lambda(A)u\| \\ &= a\|(A - \lambda I + \lambda I)R_\lambda(A)u\| + b\|R_\lambda(A)u\| \\ &\leq a\left(1 + \frac{C|\lambda|}{|\lambda - \omega|}\right)\|u\| + \frac{bC}{|\lambda - \omega|}\|u\| \end{aligned}$$

Nun findet sich für alle $\epsilon > 0$ ein ω' d.d. $\|BR_\lambda\| \leq a(1 + C + \epsilon)$ für $\operatorname{Re}(\lambda) > \omega'$. Somit ist klar, dass für ein bestimmtes ω' die Norm von $BR_\lambda(A)$ echt kleiner 1 ist. Um den Beweis zu vollenden, nutzen wir folgende Identität, die durch Inversion beider Seiten nachzurechnen ist.

$$R_\lambda(A + B) = R_\lambda(A)(I + BR_\lambda(A))^{-1}$$

Nun wissen wir aber aus der Definition der Neumann-Reihe, dass sich die Norm des zweiten Termes auf der rechten Seite für ω' gegen eine Konstante abschätzen lässt. Also folgt

$$R_\lambda(A + B) \leq \frac{M'}{|\lambda - \omega'|} \quad \text{für } \operatorname{Re}(\lambda) > \omega'.$$

Somit sind die Voraussetzungen von Satz 2.12 gezeigt. \square

Schließlich müssen wir noch klären, wie die Zeitabhängigkeit des infinitesimalen Erzeugers die ganze Theorie beeinflusst.

Satz 2.15. [2, Satz 1.2.1] *Wir betrachten das folgende Problem auf einem Banachraum X und auf dem Zeitintervall $[0, T]$*

$$\frac{du}{dt} = A(t)u + f(t), \quad u(0) = u_0. \quad (2.24)$$

Dann besitzt dieses Problem eine Lösung, falls folgende Bedingungen gelten:

1. $A(t)$ ist ein Element in $\mathcal{L}(D(A(t)), X)$ mit $D(A(t))$ konstant für alle t und dicht in X .
2. $A(t)$ erzeugt eine analytische Halbgruppe für alle t .
3. Für f gilt $f \in C^\rho([0, T], X)$ mit $\rho \in (0, 1)$.
4. Die Abbildung $t \mapsto A(t)$ Hölder-stetig ist mit demselben ρ wie f .
5. Und schließlich gilt $u_0 \in X$.

Sind diese Bedingungen erfüllt, so lässt sich die Lösung des Problems mit Hilfe der Formel für Variation der Konstanten angeben. Dabei ist die Fundamentallösung ein Evolutionsoperator. Für Details siehe [2] S.45ff.

Beweis. Der Beweis findet sich in [2] auf den Seiten 43ff und 66f. □

Unter einer Lösung verstehen wir eine Funktion $u \in C([0, T]; X) \cap C^1((0, T]; X) \cap C((0, T]; D(A))$ mit $u(t) \in D(A)$ für $t \in (0, T]$, welche die Differentialgleichung dort identisch erfüllt und außerdem noch der Anfangsbedingung genügt. Des Weiteren folgt aus Satz 2.15, dass die Lösung u zusätzlich noch weitere Regularitätsvoraussetzungen erfüllt und zwar lässt sich zeigen, dass $u \in C^\rho((0, T]; D(A)) \cap C^{1+\rho}((0, T]; D(A))$.

Die Definition eines parabolischen Evolutionsoperators wird in Anhang A.2 gegeben, an dieser Stelle soll folgendes Beispiel die Idee veranschaulichen.

Beispiel 2.8. Sei A der Erzeuger einer analytischen Halbgruppe auf X und A unabhängig von t , dann ist $U_A(t, s) = \exp(A(t - s))$ ein parabolischer Evolutionsoperator auf dem entsprechenden Definitionsbereich und Zeitintervall.

Dies zeigt den Zusammenhang mit dem Fall zeitunabhängiger Operatoren. Die Eindeutigkeit einer Lösung des obigen Problems lässt sich mit der Eindeutigkeit des Evolutionsoperators begründen. Um die Existenz einer Lösung zu beweisen, muss die Existenz eines Evolutionsoperators gezeigt werden. Dieser wird in [2, S. 53-66] explizit konstruiert.

2.3 Pseudo-Differentialoperatoren

Die Theorie der Pseudo-Differentialoperatoren (PDO) wird bei der Bewertung von Optionen im FPNIG-Modell eine wesentliche Rolle spielen. Im Folgenden sollen die hierfür notwendigen Grundlagen dargelegt werden.

Definition 2.14. Eine Funktion $\psi \in C^\infty(\Omega \times \mathbb{R})$ mit $\Omega \subset \mathbb{R}$ heißt Symbolfunktion der Ordnung $m \in \mathbb{R}$, falls die Ungleichung

$$|D_\xi^\alpha D_x^\beta \psi(x, \xi)| \leq C_{\alpha, \beta} \left(\sqrt{1 + |\xi|^2} \right)^{m - \alpha}$$

für alle Indizes $\alpha, \beta \in \mathbb{N}$, sowie alle $(x, \xi) \in \Omega \times \mathbb{R}$ erfüllt ist. Die Konstanten $C_{\alpha, \beta}$ sind dabei von (x, ξ) unabhängig. Die Klasse aller Symbolfunktionen der Ordnung m wird als $S_{1,0}^m(\Omega \times \mathbb{R})$ oder einfach $S_{1,0}^m(\Omega)$ bzw. S^m bezeichnet. Weiter setzen wir:

$$S^{-\infty}(\Omega) = \bigcap_{m \in \mathbb{R}} S_{1,0}^m(\Omega).$$

Beispiel 2.9.

$$\psi(x, \xi) = \sum_{k=1}^m a_k(x) \xi^k$$

gehört für $m \in \mathbb{N}_0$ und $a_k \in C^\infty(\Omega)$ zu $S_{1,0}^m(\Omega)$.

Nun soll ein PDO und das dazugehörige Symbol definiert werden. Wir betrachten hierzu eine Funktion $u \in S(\Omega)$ für $\Omega \subset \mathbb{R}$ und einen Differentialoperator $P(x, D)$ mit C^∞ -Koeffizienten. Dann kann u mit (2.5) dargestellt werden als:

$$u(x) = \frac{1}{2\pi} \int e^{ix\xi} \hat{u}(\xi) d\xi.$$

Wenden wir nun den Differentialoperator $P = \sum_{\alpha \leq m} a_\alpha(x) D^\alpha$ mit $D = i^{-1} \frac{\partial}{\partial x}$ auf u an, so folgt

$$P(x, D)u(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \psi(x, \xi) \hat{u}(\xi) d\xi, \quad (2.25)$$

wobei $\psi(x, \xi) = \sum_{\alpha \leq m} a_\alpha(x) \xi^\alpha \in C^\infty(\Omega \times \mathbb{R})$.

Die Theorie der PDO studiert Operatoren der Form (2.25), wobei ψ nicht unbedingt ein Polynom in ξ sein muss, sondern z.B. der Symbolklasse $S_{1,0}^m$ angehören kann.

Definition 2.15 (Pseudo-Differentialoperator). *Sei $\psi \in S_{1,0}^m(\Omega)$. Dann definieren wir auf $S(\Omega)$ den linearen Operator $P(x, D)$ durch (2.25). Wir bezeichnen $P(x, D)$ als den zur Symbolfunktion $\psi(x, \xi)$ gehörigen PDO. In der Literatur wird auch oft die Schreibweise $\psi(x, D)$ für $P(x, D)$ verwendet. Da die Fouriertransformation ein Automorphismus auf dem Schwarzraum ist und $\psi \in S_{1,0}^m(\Omega)$ gilt, konvergiert das Integral in (2.25) absolut und für alle kompakten Mengen $K \subset \Omega$ gleichmäßig in x .*

Beispiel 2.10. *Jeder lineare Differentialoperator mit C^∞ -Koeffizienten ist ein PDO.*

Im Folgenden und insbesondere in Abschnitt 8 nutzen wir folgende Schreibweise:

$$\psi_{(\beta)}^{(\alpha)}(x, \xi) = \partial_\xi^\alpha \partial_x^\beta \psi(x, \xi).$$

Definition 2.16. *Sei $U \subset \mathbb{R}$ offen, $0 \in \bar{U}$ und $m \in \mathbb{R}$. Dann definieren wir analog zu Definition 2.14 die Elemente ψ in $S_{1,0}^m(\mathbb{R} \times (\mathbb{R} + i\bar{U}))$ folgendermaßen:*

1. $\psi \in C^\infty(\mathbb{R} \times \mathbb{R})$ und ψ lässt sich mit seinen Ableitungen analytisch in ξ auf $\Im \xi \in U$ fortsetzen.
2. Es gilt folgende Abschätzung für die Ableitungen von ψ :

$$\left| \psi_{(\beta)}^{(\alpha)} \right| \leq C_{\alpha,\beta} \left(\sqrt{1 + |\xi|^2} \right)^{m-\alpha},$$

wobei die Konstanten $C_{\alpha,\beta}$ unabhängig von $(x, \xi) \in \mathbb{R} \times (\mathbb{R} + i\bar{U})$ sind.

Bei der Berechnung von Optionen im FPNIG-Modell werden wir eine Approximation des zu einem PDO inversen Operators benötigen. Hierzu werden die zwei folgenden Definitionen benötigt.

Definition 2.17. Wir bezeichnen $a \in IS_{1,0}^m(\mathbb{R} \times \mathbb{R})$ genau dann, wenn $a \in S_{1,0}^m(\mathbb{R} \times \mathbb{R})$ und a invertierbar ist mit folgender Schranke:

$$|a(x, \xi)^{-1}| \leq C \left(\sqrt{1 + |\xi|^2} \right)^{-m},$$

wobei C unabhängig von (x, ξ) ist.

Definition 2.18 (Asymptotische Entwicklung). Wir sagen, dass das Symbol $a \in S_{1,0}^m(\mathbb{R} \times \mathbb{R})$ eine asymptotische Entwicklung

$$a \sim \sum_{j=0}^{\infty} a_j \tag{2.26}$$

zulässt, falls für alle $N \in \mathbb{N}$

$$r_N = a - \sum_{j=0}^{N-1} a_j \in S_{1,0}^{m-N}(\mathbb{R} \times \mathbb{R})$$

gilt.

Die Summanden auf der rechten Seite der Gleichung werden für große ξ klein. Somit kann unter gewissen Voraussetzungen an die asymptotische Entwicklung von a , der Term r_N vernachlässigt werden. Die Summe der a_j kann dann als Approximation von a verstanden werden. Der folgende Satz sagt uns, dass durch die Vorgabe der asymptotischen Entwicklung der entsprechende Operator „fast“ eindeutig beschrieben wird.

Satz 2.16. [18, Proposition 18.1.3.] Seien $a_j \in S_{1,0}^{m-j}(\mathbb{R} \times \mathbb{R})$, $j = 0, 1, \dots$, dann existiert ein $a \in S_{1,0}^m(\mathbb{R} \times \mathbb{R})$, d.d. (2.26) erfüllt ist. Das Symbol ist eindeutig bis auf Elemente aus $S^{-\infty}(\mathbb{R} \times \mathbb{R})$ festgelegt.

Nun sind alle Grundlagen dargelegt, um den Hauptsatz für die Anwendung zu formulieren.

Satz 2.17. [6, Theorem 16.8] Sei $a \in IS_{1,0}^m(\mathbb{R} \times \mathbb{R})$ und sei für ein $s \in \mathbb{R}$ der Operator $A = a(x, D) : H^s(\mathbb{R}) \rightarrow H^{s-m}(\mathbb{R})$ invertierbar, dann ist der zu A inverse Operator ein PDO mit Symbol $b \in S_{1,0}^{-m}$. Das Symbol b lässt sich folgendermaßen asymptotisch entwickeln:

$$b = \sum_{j=0}^{\infty} b_{-m-j}, \tag{2.27}$$

wobei

$$b_{-m} = \frac{1}{a} \in S_{1,0}^{-m}(\mathbb{R} \times \mathbb{R}) \tag{2.28}$$

und die $b_{-m-j} \in S_{1,0}^{-m-j}$ induktiv folgendermaßen definiert sind:

$$b_{-m-j} = -a^{-1} \sum_{k < j, k+\alpha=j} (\alpha!)^{-1} a^{(\alpha)}(b_{-m-k})_{(\alpha)}. \quad (2.29)$$

Bevor wir zum Abschluss dieses Kapitels die unerwartete Verbindung zwischen Halbgruppentheorie und der Theorie der PDO herstellen, brauchen wir die folgenden zwei Definitionen.

Definition 2.19. Sei $m \in (0, 2]$, $a \in S_{1,0}^m(\mathbb{R} \times (\mathbb{R} \times i\bar{U}))$ und für alle $x \in \mathbb{R}$ sei $a(x, \psi)$ der charakteristische Exponent eines Lévy-Prozesses. Dann nutzen wir die Schreibweise $a \in FLS_{1,0}^m(\mathbb{R} \times (\mathbb{R} \times i\bar{U}))$.

Definition 2.20. Sei $A : D(A) \rightarrow C_\infty(\mathbb{R}; \mathbb{R})$ ein linearer Operator, $D(A) \subset C_\infty(\mathbb{R}; \mathbb{R})$. Dann erfüllt $(A, D(A))$ ein positives Maximumsprinzip genau dann, wenn für alle $u \in D(A)$ aus $u(x_0) = \sup_{x \in \mathbb{R}} u(x) \geq 0$ folgt, dass $Au(x_0) \leq 0$ ist.

Nun kann gezeigt werden, dass ein PDO unter bestimmten Voraussetzungen eine Feller-Halbgruppe erzeugt.

Satz 2.18. [6, Theorem 3.5] und [19, Theorem 4.5.3] Sei $a \in FLS^m(\mathbb{R} \times (\mathbb{R} \times i\bar{U}))$, dann besitzt a einen Abschluss und der Abschluss erzeugt eine Feller-Halbgruppe auf $C_\infty(\mathbb{R})$.

Beweis. Es genügt folgende drei Bedingungen zu zeigen.

1. $D(a(x, D)) = S(\mathbb{R})$ ist dicht in C_∞ .
2. $a(x, D)$ erfüllt ein positives Maximumsprinzip auf seinem Definitionsbereich.
3. Für $\lambda > 0$ ist das Bild von $a(x, D) + \lambda$ dicht in $C_\infty(\mathbb{R})$.

□

2.4 Lévy-Prozesse und Halbgruppen

Nun kann die Frage nach dem Zusammenhang von Halbgruppentheorie und der Theorie der Lévy-Prozesse beantwortet werden.

Definition 2.21. Für einen Lévy-Prozess X definieren wir den Übergangoperator $T_{t,s}$ für $f \in C_\infty(\mathbb{R})$ folgendermaßen:

$$T_{t,s}f(x) = \mathbb{E}[f(X_t)|X_s = x].$$

Auf Grund der zeitlichen Homogenität lässt sich dies zu $T_{t-s} = T_{t-s,0} = T_{t,s}$ umschreiben.

Satz 2.19. [3, Theorem 3.1.9] Die durch einen Lévy-Prozess definierten Übergangsoperatoren erzeugen eine Feller Halbgruppe.

Beweis. Wegen der Markov-Eigenschaft von Lévy-Prozessen und Eigenschaft 5 aus Definition 2.6 erzeugen die Übergangsoperatoren $\{T(t)\}_{t \geq 0}$ eine C_0 -Halbgruppe. Definieren wir die Operatoren $\{T(t)\}_{t \geq 0}$ auf dem Banachraum $C_\infty(\mathbb{R})$, ausgestattet mit der Supremumsnorm, so bilden die Operatoren eine Kontraktionshalbgruppe. Die Feller-Eigenschaft der Halbgruppe folgt sofort aus der Definition. \square

Nun wollen wir den infinitesimalen Erzeuger dieser Halbgruppe herleiten. Es ist klar, dass dieser bei der Bewertung von Derivaten eine wesentliche Rolle spielen wird. Hierbei ist der Übergang in den Fourierraum sehr nützlich. Wir betrachten für $f \in S(\mathbb{R})$ (vgl. [3, Theorem 3.3.3] und [30, Lemma 31.7]):

$$T_t f(x) = \mathbb{E}[f(X_t) | X_0 = x] = \int_{\mathbb{R}} f(x+y) \mu_t(dy),$$

wobei μ_t das Wahrscheinlichkeitsmaß zum Zeitpunkt t ist. Mit Hilfe der inversen Fouriertransformation können wir nun folgendermaßen umformen:

$$T_t f(x) = \frac{1}{2\pi} \mathbb{E} \left[\int_{\mathbb{R}} e^{iu(x+y)} \hat{f}(u) du \right]. \quad (2.30)$$

Um den Satz von Fubini anwenden zu können, reicht es folgende Abschätzung zu berücksichtigen:

$$\left| \int_{\mathbb{R}} e^{iux} \mathbb{E}[e^{iuy}] \hat{f}(u) du \right| \leq \int_{\mathbb{R}} |e^{iux} \mathbb{E}[e^{iuy}]| |\hat{f}(u)| du \leq \int_{\mathbb{R}} |\hat{f}(u)| du \leq \infty. \quad (2.31)$$

Deswegen folgt:

$$T_t f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} \mathbb{E}[e^{iuy}] \hat{f}(u) du = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} e^{-t\psi(u)} \hat{f}(u) du.$$

Somit ist $T(t)$ ein PDO mit Symbol $e^{-t\psi(u)}$. Auch der infinitesimale Erzeuger des Prozesses ist ein PDO, wie folgende Rechnung für $f \in S(\mathbb{R})$ zeigt.

$$\begin{aligned} Af(x) &= \lim_{t \rightarrow 0} \frac{T_t f(x) - f(x)}{t} = \frac{1}{2\pi} \lim_{t \rightarrow 0} \int_{\mathbb{R}} e^{iux} \frac{e^{-t\psi(u)} - 1}{t} \hat{f}(u) du \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} (-\psi(u)) \hat{f}(u) du. \end{aligned}$$

Integration und Limesbildung dürfen in obiger Rechnung auf Grund des Theorems von

der dominierten Konvergenz und folgender Abschätzung vertauscht werden:

$$\int_{\mathbb{R}} \left| e^{iux} \frac{e^{-t\psi(u)} - 1}{t} \hat{f}(u) \right| du \leq \int_{\mathbb{R}} |\psi(u) \hat{f}(u)| du$$

$$\stackrel{(2.9)}{\leq} C \int_{\mathbb{R}} (1 + |u|^2) |\hat{f}(u)| du < \infty.$$

Die letzte Abschätzung folgt, da $(1 + |u|^2)\hat{f}(u) \in S(\mathbb{R})$ gilt. Der Operator A wird als Erzeuger des Prozesses X bezeichnet. Ist $-\psi(x) \in S_{1,0}^m(\mathbb{R} \times (\mathbb{R} \times \bar{U}))$, so bezeichnen wir X als Prozess der Ordnung m und exponentiellem Typ U .

Somit würde sich die Bewertung von Derivaten europäischen Typs auf die Berechnung eines zu (2.30) analogen Integrals reduzieren. Die Schwierigkeit, die beim FPNIG hinzukommt ist, dass sich die Halbgruppe nicht explizit angeben lässt.

Um die Theorie etwas zu veranschaulichen, soll nun beispielhaft eine Halbgruppe hergeleitet werden.

Beispiel 2.11 (Halbgruppe für einen gewichteten Poissonprozess). Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von unabhängig identisch verteilten Zufallsvariablen mit Dichte $h(x)$ und sei N_t ein Poissonprozess mit Intensität λ . Sei weiter $S_n := X_1 + \dots + X_n$ die n -te Partialsumme. Der gewichtete Poissonprozess ist gegeben durch:

$$Y_t = \sum_{n=1}^{\infty} S_n \mathbf{1}_{[N_t=n]}.$$

Nun lassen sich die Elemente der Halbgruppe folgendermaßen bestimmen:

$$(T_t f)(x) = \mathbb{E}^x[f(Y_t)] = \sum_{n=0}^{\infty} \mathbb{E}[f(x + S_n)] P(N_t = n)$$

$$= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mathbb{E}[f(x + S_n)] = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} L^n f(x),$$

mit $L^n f(x) = \mathbb{E}[f(x + X_1 + \dots + X_n)] = \mathbb{E}[f(x + S_n)]$.

Formal können wir dies zu

$$T_t f(x) = (e^{At} f)(x)$$

umschreiben, wobei

$$A f(x) = \lambda(L - I)f(x) = \lambda \int_{\mathbb{R}} (f(x + y) - f(x)) h(y) dy.$$

Somit ist der infinitesimale Erzeuger formal hergeleitet. Um nun A auch rigoros herleiten zu können, nutzen wir, dass sich der charakteristische Exponent eines gewichteten

Poissonprozesses explizit angeben lässt (vgl. [14, Proposition 3.4]):

$$\psi(\xi) = \lambda(1 - \phi_{X_i}(\xi)),$$

wobei $\phi_{X_i}(\xi)$ die charakteristische Funktion von X_i ist. Wir können also nach obiger Rechnung den infinitesimalen Erzeuger A der Halbgruppe der Übergangsoperatoren für $f \in S(\mathbb{R})$ folgendermaßen berechnen:

$$\begin{aligned} Af(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} \lambda(\phi_{X_i}(u) - 1) \hat{f}(u) du \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} \lambda \left(\int_{\mathbb{R}} e^{iuy} h(y) dy - 1 \right) \hat{f}(u) du \\ &= \frac{1}{2\pi} \int_{\mathbb{R} \times \mathbb{R}} \lambda e^{iu(x+y)} \hat{f}(u) h(y) dy du - \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} \hat{f}(u) du \\ &\stackrel{(2.31)}{=} \frac{1}{2\pi} \int_{\mathbb{R} \times \mathbb{R}} \lambda e^{iu(x+y)} \hat{f}(u) du h(y) dy - \lambda f(x) \\ &= \lambda \int_{\mathbb{R}} f(x+y) h(y) dy - \lambda f(x) = \lambda \int_{\mathbb{R}} (f(x+y) - f(x)) h(y) dy \end{aligned}$$

Damit haben wir unsere formale Rechnung nun rigoros bestätigt.

Der infinitesimale Erzeuger der Halbgruppe der Übergangsoperatoren eines Lévy-Prozesses lässt sich nicht nur als PDO herleiten, sondern auch als PIDE. Dabei ist er für alle $f \in C_{\infty}^2$ definiert und hat für einen Lévy-Prozess X mit charakteristischem Tripel $(\mu, \sigma^2, \nu(dx))$ folgende Gestalt:

$$Af(x) = \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial x^2}(x) + \mu \frac{\partial f}{\partial x}(x) + \int_{\mathbb{R}} \left(f(x+y) - f(x) - y \frac{\partial f}{\partial x}(x) \mathbf{1}_{[-1,1]}(y) \right) \nu(dy)$$

Für $f \in S(\mathbb{R})$ lässt sich die Äquivalenz der beiden Darstellungen zeigen (vgl. [3, Theorem 3.3.3]). Wir haben uns für die Darstellung mit Hilfe der PDOs entschieden, da diese für die Verallgemeinerung in Kapitel 8 die richtige Darstellung ist.

2.5 Bewertung von contingent claims im NIG-Modell

Nun sollen im NIG-Modell, welches in Abschnitt 2.1 definiert wurde, Optionen bewertet werden. Es lässt sich zeigen, dass der NIG-Prozess die Ordnung 1 hat und den exponentiellen Typ $[\lambda_-, \lambda_+]$ (vgl. (2.13) und (2.14)), wobei

$$\lambda_- = -\alpha + \beta \quad \lambda_+ = \alpha + \beta.$$

Des Weiteren lässt sich elementar nachrechnen, dass für alle ξ mit $\Im\xi \in [\lambda_-, \lambda_+]$

$$\Re\psi(\xi) \geq c(1 + |\xi|) - C \quad (2.32)$$

mit Konstanten $c, C > 0$ gilt (siehe [11, Lemma 3.6]). Nun sind wir in der Lage, endfällige Optionen im NIG-Modell zu bewerten. Im Folgenden betrachten wir Optionen auf ein Underlying (S) mit folgender risikoneutraler Dynamik

$$S_t = \exp(X_t),$$

wobei X ein NIG-Prozess ist. Der Wert ($F(S_t, t)$) eines contingent claim ist der vom Fälligkeitszeitpunkt (T) auf den aktuellen Zeitpunkt (t) diskontierte risikoneutrale Erwartungswert des Auszahlungsprofils. Oder formal

$$F(S_t, t) = e^{-r\tau} \mathbb{E}[g(X_T) | X_t = x] \quad (2.33)$$

mit $x = \ln S_t$ und $\tau = T - t$. Für einfache Optionen europäischen Typs mit Auszahlungsprofil $g(x) \in C_\infty(\mathbb{R})$ ist die Bewertung klar. Es ist nach (2.30) folgendes Integral zu bestimmen:

$$F(S_t, t) = \frac{1}{2\pi} e^{-r\tau} \int_{\mathbb{R}} e^{ix\xi} e^{-\tau\psi(\xi)} \hat{g}(\xi) d\xi. \quad (2.34)$$

Dies schließt jedoch Puts und Calls nicht ein, da in beiden Fällen das Auszahlungsprofil nicht in $C_\infty(\mathbb{R})$ enthalten ist.

Für messbare Auszahlungsprofile mit einem $\sigma \in (\lambda_-, \lambda_+)$ d.d. $g_\sigma(x) = e^{\sigma x} g(x)$ in $L_1(\mathbb{R})$ enthalten ist, lässt sich unter Benutzung von (2.32) eine analoge Bewertungsgleichung herleiten (vgl. [23]):

$$F(S_t, t) = \frac{1}{2\pi} \int_{-\infty+i\sigma}^{+\infty+i\sigma} e^{ix\xi} e^{-\tau(r+\psi(\xi))} \hat{g}(\xi) d\xi. \quad (2.35)$$

Diese lässt sich für einen europäischen Call zu

$$F_{\text{Call}}(S_t, t) = -\frac{K}{2\pi} \int_{-\infty+i\sigma}^{+\infty+i\sigma} \frac{\exp(i\xi \ln(S_t/K) - \tau(r + \psi(\xi)))}{\xi(\xi + i)} d\xi \quad (2.36)$$

vereinfachen. Sowohl in (2.34) als auch in (2.35) lässt sich der Wert der Option als PDO auffassen:

$$F(S_t, t) = \exp(-\tau(r + \psi(D_x)))g(x). \quad (2.37)$$

Kapitel 3

Einführung in Energiemärkte

3.1 Grundlagen

Um die Dynamik des Strompreises in Deutschland adäquat modellieren zu können, ist ein Verständnis fundamentaler Eigenschaften des Strommarktes unerlässlich.

Vor der Liberalisierung agierten auf dem deutschen Strommarkt vertikal integrierte Akteure, die ein Gebietsmonopol inne hatten. Dies bedeutet, dass die Nachfrager nicht zwischen verschiedenen Stromanbietern auswählen konnten, sondern von ihrem lokalen Monopolisten versorgt wurden. Die Preise für Strom entstanden nicht im Wettbewerb, sondern wurden im Rahmen einer Kosten-Plus-Methode von der Regulierungsbehörde festgelegt. Es wurde den Gebietsmonopolisten zu ihren Kosten eine gewisse Marge zugestanden¹. In einem solchen System gab es keine Notwendigkeit für aktiven Handel mit Strom.

Seit der Liberalisierung Ende der neunziger Jahre dürfen alle Stromanbieter im ganzen Bundesgebiet Strom anbieten und die Preise ergeben sich aus Angebot und Nachfrage. Im Zuge der Liberalisierung entstand die Nachfrage nach vorher nicht existenten Termin- und Spotmärkten. Des Weiteren entstanden mehrere Handelsstufen, einerseits der Großhandelsmarkt und andererseits der Einzelkundenmarkt. Auf dem Großhandelsmarkt agieren Produzenten, große Industriekunden, Weiterverteiler und Dritte, die kein physisches Interesse am Gut haben. Der Handel wird in Deutschland über die EEX in Leipzig oder OTC abgewickelt. Alternativ wäre auch ein Handel in einem Poolssystem vorstellbar². Auf dem Endkundenmarkt agieren die (kleinen) Endverbraucher und die Weiterverteiler (z.B. Stadtwerke). Wir wollen uns im Folgenden nur mit dem Großhandelsmarkt beschäftigen.

Der Großhandelsmarkt kann unterteilt werden in einen Spot- und einen Terminmarkt. Auf

¹Es fand ebenfalls eine Preisdifferenzierung nach den unterschiedlichen Nachfragergruppen statt. So war der Strom für Haushalte oder den Handel teurer als für Industriebetriebe.

²In einem Poolsystem wird der Handel durch einen Poolmanager organisiert. Es agieren nur Produzenten und (physische) Nachfrager von Strom auf diesem Markt.

dem Terminmarkt werden Stromlieferungen für einen in der Zukunft liegenden Zeitraum gehandelt. In der Regel wird ab einer Wartezeit von mehreren Tagen bis zur Lieferung von einem Terminmarktkontrakt gesprochen. Spotkontrakte können bis 75 Minuten vor Lieferung an der Börse abgeschlossen werden³.

An der EEX werden Standardprodukte wie Monats-, Quartals- und Jahresfuture angeboten. Dies sind Kontrakte, die in einem in der Zukunft liegenden Zeitraum (Monat, Quartal oder Jahr) durchgehend eine Menge von einem MW liefern. Bei OTC Geschäften können alle möglichen Konditionen vereinbart werden. Zu beachten ist, dass in den Kontrakten sowohl eine physische als auch eine finanzielle Erfüllung vereinbart werden kann, d.h. entweder bekommt der Käufer eines solchen Kontraktes eine tatsächliche Stromlieferung oder er bekommt eine Ausgleichszahlung in entsprechender Höhe. Die Nutzung von Terminkontrakten hat sowohl für den Käufer als auch für den Verkäufer den großen Vorteil der Preissicherheit. Beiden Parteien ist bekannt, wie groß ihre Einnahmen bzw. ihre Ausgaben für Strom in der Zukunft sein werden. Deswegen ist es nicht verwunderlich, dass ein Großteil des in Deutschland gehandelten Stromes über Terminmarktkontrakte abgewickelt wird.

Strom, der nicht über den Terminmarkt vermarktet werden konnte oder sollte, wird auf dem Spotmarkt angeboten. Dementsprechend ist der Spotmarkt ein Markt für Restkapazitäten. Der Spotmarkt kann in einen Day-Ahead-Markt und einen Intra-Day-Handel unterteilt werden. Auf dem Day-Ahead-Markt wird, wie der Name bereits andeutet, Strom für Lieferungen am Folgetag gehandelt, während im Intra-Day-Handel Strom für Lieferungen am selben Tag und bis kurz vor Fälligkeit gehandelt wird. Spot Strom kann genauso wie alle Terminmarktprodukte OTC oder an der Börse gehandelt werden. Wir werden nur Börsenprodukte betrachten, da für den OTC Handel keine Marktpreise bekannt sind. Weiterhin werden wir uns auf den Day-Ahead-Markt beschränken, da die dort gehandelten Volumina den Intra-Day-Handel bei weitem übersteigen und die Day-Ahead-Preise als Richtpreise für viele Produkte verwendet werden.

3.2 Wesentliche Eigenschaften des Spotkurses

Der Spotkurs an der EEX bildet sich für alle Stunden einzeln am Tag vor Lieferung durch eine beidseitige Auktion von Seiten der Nachfrager und Anbieter. Der Kurs war bis Mitte dieses Jahres nach unten durch Null beschränkt und der Price-cap lag bei 3000 Euro/MWh. In Abbildung 3.1 sind die Angebots- und Nachfragekurve exemplarisch für eine Stunde am 27.07.2007 abgebildet. Der Spotpreis für diese Stunde entspricht dem Schnittpunkt der beiden Kurven. Um den Spotkurs auf vernünftige Art und Weise zu modellieren, sind besondere Eigenschaften des Commodities Strom, die sogenannten stylized facts, zu

³OTC wäre es auch vorstellbar noch kurzfristiger Kontrakte zu handeln.

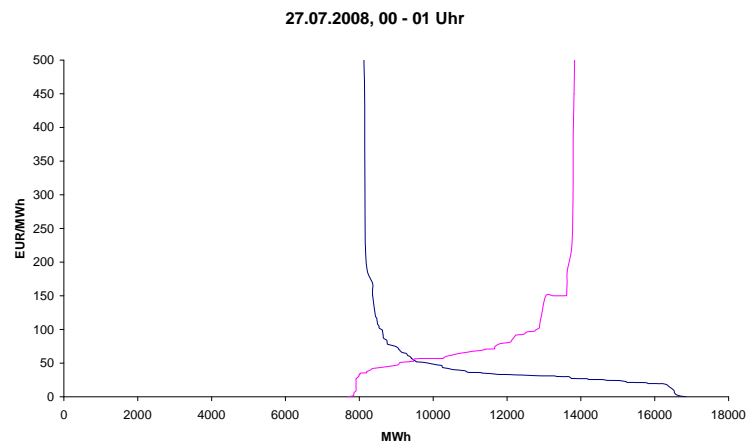


Abbildung 3.1: Beispiel für Angebots- und Nachfragekurven

berücksichtigen. Strom ist in Deutschland (zumindest im Moment) ein nicht in relevanten Mengen lagerfähiges Gut. Es ist demzufolge nicht möglich, große Mengen Strom zu speichern, er muss bei Erzeugung sofort verbraucht werden. Daraus folgt, dass der an der EEX für verschiedene Stunden eines Tages gehandelte Strom verschiedene Produkte darstellt, zwischen denen keine Arbitrageüberlegungen (wie dies bei Aktien der Fall ist) möglich sind. Des Weiteren können Bewertungsmethoden für Derivate, die sich auf das Halten oder short-Verkaufen des Underlyings stützen, bei Strom nicht angewandt werden. Weiterhin unterliegt der Strompreis, wie der Abbildung 3.2 zu entnehmen ist, mehreren

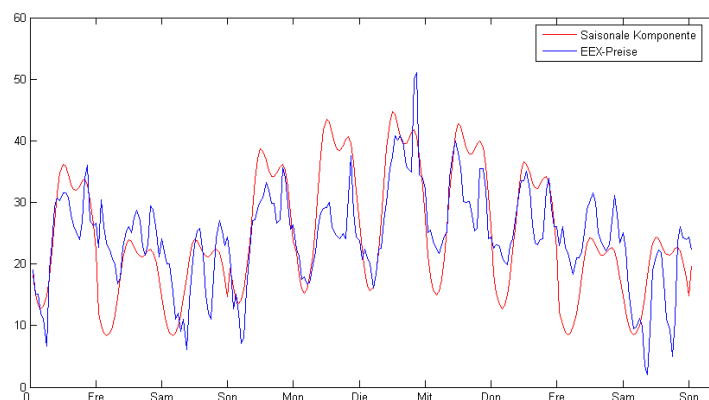


Abbildung 3.2: Tagessaisonalität Beispielhaft vom 3.03.07 bis zum 12.03.07 dargestellt

Saisonalitäten. So ist ganz deutlich eine Tagessaisonalität mit einem Hochpunkt bei 12 Uhr und dem charakteristischen Höcker um 19.00 zu erkennen. Dies rührt daher, dass der über den Tag schwankenden Nachfrage nach Strom ein (weitestgehend) konstantes An-

gebot gegenübersteht.

Des Weiteren sehen wir eine klare Wochensaisonalität mit einem Tiefpunkt am Sonntag und höheren Preisen unter der Woche. Die Erklärung hierfür ist analog zur Tagessaison. Nicht so deutlich zu erkennen sind die Jahressaisonalitäten, dies ist eine besondere Eigenschaft des deutschen Marktes. Dies ist auf dem Nord Pool Markt wesentlich einfacher zu sehen und auch in Südeuropa wesentlich stärker ausgeprägt, da in diesen Märkten die Temperaturunterschiede zwischen den Jahreszeiten größer sind. In Südeuropa sind im Sommer höhere Preise auf Grund der höheren Nachfrage nach Kühlung zu beobachten, während im Norden im Winter höhere Preise wegen des höheren Heizbedarfes zu beobachten sind. Diese Saisonalitäten lassen sich ebenfalls über die Angebotsseite begründen. Für den deutschen Markt wäre u.U. eine stochastische Modellierung des jährlichen Saisoneinflusses sinnvoll.

Aus Abbildung 3.2 ist ebenfalls zu erkennen, dass neben dem deterministischen Saisoneinfluss auch eine stochastische Komponente die Preisentwicklung beeinflusst. Die Preise weichen immer wieder zwar von dem deterministischen Niveau ab, kehren aber auch auf das Niveau zurück. Dies hat seine Ursache in kurzfristigen Verschiebungen der Nachfrage- und Angebotskurven. Die Eigenschaft der Rückkehr zu einem deterministischen Niveau wird als Mean Reversion bezeichnet. Ähnliche Phänomene sind bei Zinzentwicklungen zu beobachten und dementsprechend kann die Art diese zu modellieren auch übernommen werden.

Eine weitere wichtige Eigenschaft von Spotpreisen in Deutschland sind die sogenannten Spikes. Dies sind starke Preisanstiege, gepaart mit einem schnellen Preisrückgang auf ein normales Niveau. In Abbildung 3.3 ist die Angebots- und Nachfragesituation für einen Spike abgebildet. Es ist deutlich zu erkennen, dass sich der Schnittpunkt beider Kurven im unelastischen Bereich sowohl der Nachfrage- als auch der Angebotskurve liegt. Der unelastische Bereich der Nachfragekurve stellt jene Nachfrager dar, die bereit sind für einen beliebig hohen Preis (unter dem Price-cap) Strom zu beziehen. Dies hat den Grund, dass ein bestimmter Teil der Nachfrager auf Stromlieferungen angewiesen ist und nicht auf Preissignale reagieren kann. Beispiele hierfür sind industrielle Fertigungsprozesse, deren Unterbrechung sehr hohe Kosten nach sich zieht oder Stadtwerke, die mit Haushalten Festpreise vereinbaren und somit zu jedem Preis Strom liefern müssen. Der unelastische Teil der Angebotskurve stellt die begrenzten Produktionskapazitäten dar. Ab einer bestimmten Ausbringungsmenge können Kapazitäten nur noch sehr teuer hinzugefügt werden⁴. Verschiebt sich nun der Schnittpunkt von Angebot und Nachfrage in den unelastischen Bereich, so kann dies zu sehr hohen Preisausschlägen führen. Gründe für eine solche Verschiebung können auf Seiten der Anbieter Kraftwerksausfälle sein oder auf Seiten der Nachfrager plötzliche Nachfragesprünge.

Das Besondere an Spikes im Gegensatz zu Sprüngen ist die schnelle Rückkehr (i.d.R.

⁴Strategisches Bieten großer Energieversorger kann ebenfalls nicht ausgeschlossen werden.

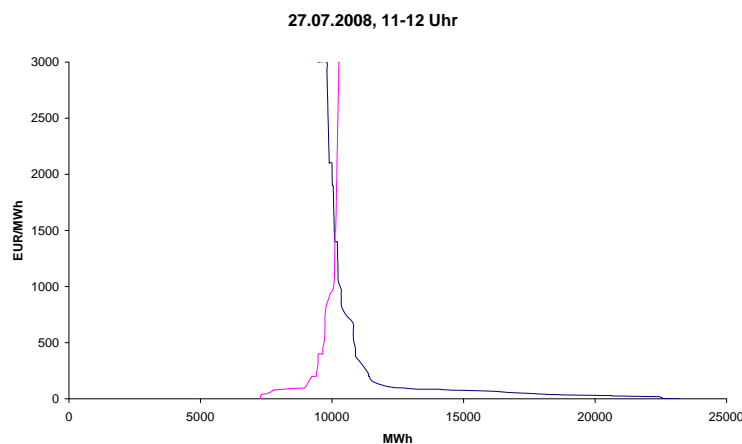


Abbildung 3.3: Beispiel für Angebots- und Nachfragekurven

innerhalb weniger Stunden) auf ein normales Niveau nach einer solchen Preisspitze. Aus Abbildung 3.4 ist zu erkennen, dass Spikes seltene Ereignisse sind, die aufgrund der kurzen Historie des Spotmarktes schwierig zu modellieren und kalibrieren sind. Andererseits ist jedoch klar, dass ein Spike auf Grund des enormen Unterschieds zu normalen Preisen ein sehr wichtiges Ereignis darstellt und daher bei einer vernünftigen Modellierung des Strompreises unbedingt berücksichtigt werden sollte. Es lässt sich ebenfalls beobachten, dass Spikes nur bei einem an sich bereits erhöhten Preisniveau auftreten. Der Grund hierfür ist, dass bei einem niedrigen Preisniveau in der Regel noch hinreichend Kapazität im Markt vorhanden ist, um einen Kraftwerksausfall oder einen Nachfragesprung ohne extrem teure Kapazitäten abzufedern.

Ein Spike stellt einen einzelnen hohen Preisausschlag dar, im Gegensatz dazu sind auch sogenannte Hochpreisphasen zu beobachten. Darunter ist ein Zeitraum zu verstehen, in dem ein höheres Preisniveau im Vergleich zu Durchschnittspreisen vorherrscht. Diese treten in Deutschland vor allem im Sommer bei einer Knappheit an Kühlwasser und somit eingeschränktem Angebot auf dem Strommarkt auf.

Schließlich lässt sich auf Strommärkten beobachten, dass sich die Volatilität, also die Schwankung der Preise um das deterministische Niveau, im Zeitablauf ändert. Es wäre zu klären, ob diese Änderungen deterministischer oder stochastischer Natur sind.

Die wesentlichen Charakteristika deutscher Spotpreise sind im Folgenden nochmal aufgelistet, die meisten können auf anderen Märkten ebenfalls beobachtet werden. Ein vernünftiges Spotpreismodell sollte einen Großteil dieser Eigenschaften abbilden.

- Saisonaler Einfluss
- Stochastische Komponente

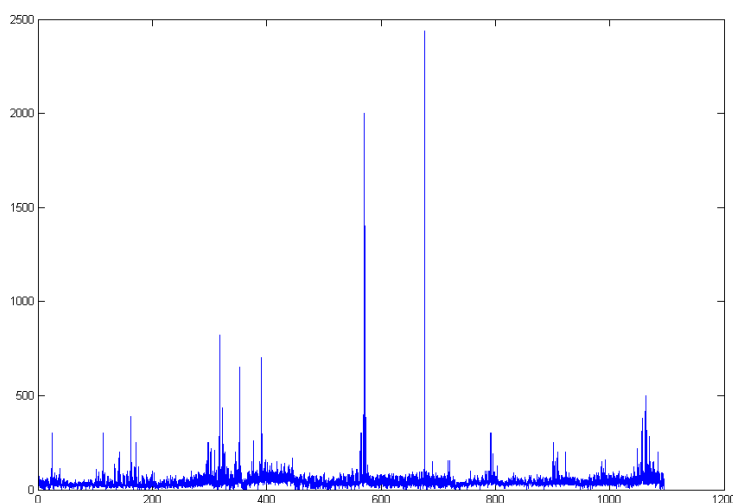


Abbildung 3.4: Spotkurs der Jahre 2005-2007 an der EEX

- Mean Reversion
- Spikes
- Hochpreisphasen
- Nicht-konstante Volatilität
- Nicht-Negativität⁵

Prinzipiell gibt es nun zwei Möglichkeiten, ein stochastisches Modell für die Beschreibung der Spotpreisentwicklung zu nutzen. Die erste Möglichkeit wäre, einzeln das Angebot und die Nachfrage zu modellieren und daraus die sich im Gleichgewicht ergebenden Preise zu berechnen. Der große Vorteil solcher Modelle liegt darin, dass a priori Wissen über die Beschaffenheit von Angebot und Nachfrage gut berücksichtigt werden kann. Der wesentliche Nachteil liegt in der großen Anzahl an exogenen und endogenen Variablen und dem somit äußerst hohen Rechenaufwand. Die zweite Möglichkeit, die in dieser Arbeit verwendet wird, ist es, den Spotkurs direkt zu modellieren. Hier kann das a priori Wissen nicht so gut wie bei der ersten Klasse von Modellen eingebracht werden, jedoch sind die Modelle wesentlich einfacher handhabbar.

⁵Die Nicht-Negativität ist seit Mitte dieses Jahres an der EEX aufgehoben.

Kapitel 4

Das Kluge-Modell

4.1 Beschreibung

Für die Modellierung des Spotkurses nutzen wir ein in [22] entwickeltes Modell. Dieses ist in der Lage Spikes abzubilden. Außerdem gehen stylized facts, wie die verschiedenen Saisonalitäten, die Mean Reversion und die Nichtnegativität, ein. Andere Aspekte, wie die variable Volatilität sowie die Hochpreisphasen, bleiben unberücksichtigt. Das Spikerisiko müsste, um der Realität Rechnung zu tragen, als nicht konstant modelliert werden. So gibt es beispielsweise Nachts keine Spikes. Die Berücksichtigung all dieser Aspekte würde spätestens bei der Bewertung von Derivaten auf Grund der Hochdimensionalität numerisch kaum zu bewältigen sein. Das Modell für den Spotkurs S_t ist ein Summenmodell, bestehend aus zwei stochastischen Prozessen (X_t und Y_t) und einer deterministischen Komponente $f(t)$.

$$S_t = \exp(f(t) + X_t + Y_t) \quad (4.1)$$

$$dX_t = -\alpha X_t dt + \sigma dW_t \quad (4.2)$$

$$dY_t = -\beta Y_{t-} dt + J_t dN_t \quad (4.3)$$

Dabei ist W_t ein Wiener Prozess, N_t ein Poissonprozess mit Intensität η und $(J_t)_{t \geq 0}$ ist eine Familie von unabhängigen und identisch exponentialverteilten Zufallsvariablen mit Parameter μ_J . Die Prozesse W_t , J_t und N_t sind stochastisch unabhängig. Die Exponentialdarstellung wird gewählt, um negative Preise auszuschließen. Der Prozess X_t beschreibt zusammen mit der saisonalen Komponente f die alltägliche Entwicklung des Spotkurses, während der Prozess Y_t die Spikes abbildet.

Die Funktion $f(t)$ wird als Summe einiger periodischer Funktionen, die die verschiedenen Saisonalitäten berücksichtigen gewählt. Dabei läuft t von 0 bis 365:

$$f(t) = f_{Jahr}(t) + f_{Woche}(t) + f_{Tag}(t)$$

Die deterministische Komponente, kalibriert an EEX Daten von 2005-2007, ist in Abbildung 4.1 dargestellt (vgl. Kapitel 6). Der Prozess X_t ist für die „normalen“ Schwan-

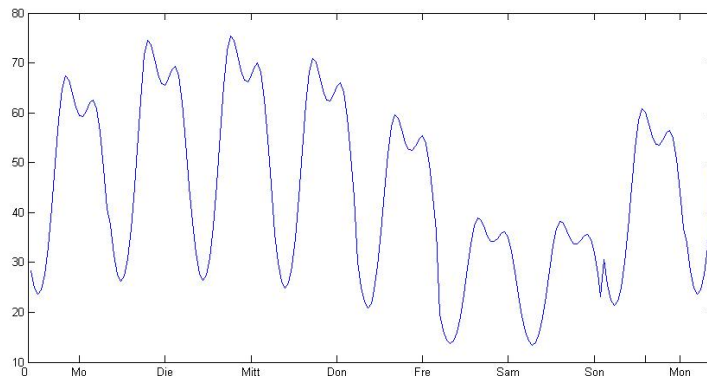


Abbildung 4.1: Deterministische Komponente im Kluge-Modell

kungen um das deterministische Niveau verantwortlich. Der Parameter κ bestimmt dabei die Geschwindigkeit der Mean Reversion und σ bestimmt die Stärke des stochastischen Einflusses. Wir erwarten ein relativ geringes κ .

Der Prozess Y_t ist für die Abbildung der Spikes verantwortlich. Der Intensitätsparameter des Poissonprozesses η steuert die Häufigkeit der Spikes, während der Parameter der Exponentialverteilung μ_J die Höhe der Spikes reguliert. Der Parameter β beeinflusst die Mean Reversion Geschwindigkeit des Spike Prozesses. Da ein Spike in der Regel nach wenigen Stunden wieder verschwunden ist, muss $\beta \gg \kappa$ gewählt werden. Eine Beispielrealisation von X_t und Y_t ist der Abbildung 4.2 zu entnehmen. Wie zu erwarten war, ist der Prozess Y_t meistens Null und hat manchmal große Ausschläge. Der Prozess X_t ist ständig aktiv. An dieser Stelle wird der Vorteil der Modellierung mit einem Summenprozess deutlich, denn würde die ganze Dynamik mit einem Prozess beschrieben, müsste für die normale Entwicklung des Prozesses und das Auftreten von Spikes mit derselben Mean Reversion Geschwindigkeit operiert werden. Dies würde unweigerlich zu einer unrealistischen Abbildung des Spotpreises führen. In Abbildung 4.3 ist ein Beispielpfad des Gesamtprozesses abgebildet. Die Pfade bilden die wesentlichen Charakteristika ab und weisen keine wesentlichen strukturellen Unterschiede zu den beobachteten Preisen auf. Um den Prozess analytisch besser handhaben zu können, ist eine äquivalente Schreibweise sehr hilfreich, welche den deterministischen Anteil im Diffusionsprozess als deterministischen Trend berücksichtigt. Dazu brauchen wir folgenden einfachen Satz:

Satz 4.1. Sei X ein OU-Prozess mit

$$dX_t = -\alpha X_t dt + \sigma dW_t$$

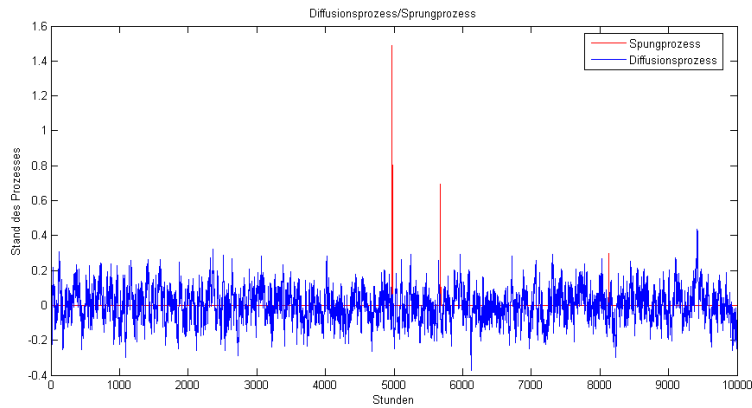


Abbildung 4.2: Realisation von X_t und Y_t mit $\kappa = 1.7892$, $\sigma = 0.209$, $\mu_J = 0.9$, $\eta = 15/365$, $\beta = 9.4$

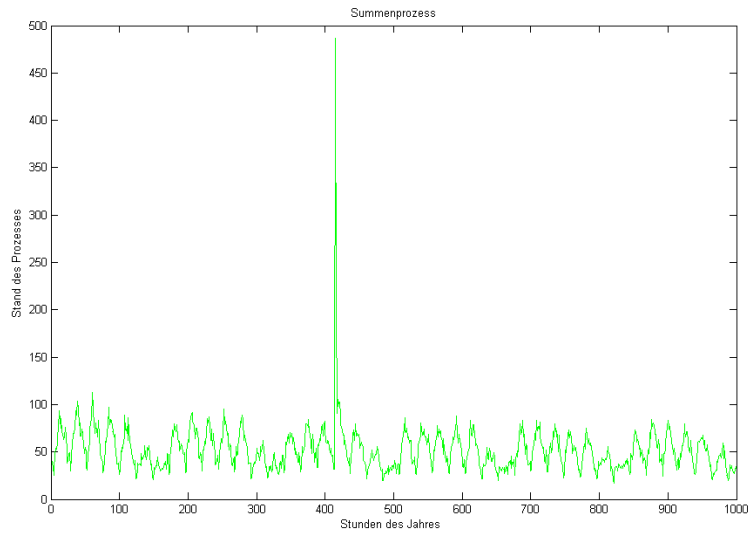


Abbildung 4.3: Simulation des Kluge-Prozesses

und $f : [0, \infty) \rightarrow \mathbb{R}$ mit $f \in C^2((0, \infty)) \cap C([0, \infty))$, dann folgt für $\tilde{X}_t = f(t) + X_t$

$$d\tilde{X}_t = \alpha(\mu(t) - \tilde{X}_t)dt + \sigma dW_t \quad \text{mit} \quad \mu(t) = f(t) + \frac{1}{\alpha} f'(t)$$

Beweis. siehe Anhang. □

Somit lässt sich die Entwicklung des Spotkurses umschreiben zu

$$S_t = \exp(\tilde{X}_t + Y_t) \tag{4.4}$$

$$d\tilde{X}_t = \alpha(\mu(t) - \tilde{X}_t)dt + \sigma dW_t \tag{4.5}$$

$$dY_t = -\beta Y_{t-} dt + J_t dN_t \tag{4.6}$$

mit $\mu(t)$ wie oben. Im Folgenden schreiben wir X_t für \tilde{X}_t .

4.2 Herleitung der PIDE für den europäischen Put

Nun sind wir in der Lage, eine PIDE für den europäische Put im Kluge-Modell herzuleiten (vgl. [22, S.87f]). Das Auszahlungsprofil und somit der Wert eines europäischen Puts mit Basispreis K bei Fälligkeit ist

$$V(x, y, T) = \max(K - \exp(f(T) + x + y), 0).$$

Sein Wert zu einem früheren Zeitpunkt lässt sich als erwartetes diskontiertes Auszahlungsprofil berechnen oder formal aufgeschrieben:

$$V(x, y, t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}} [\max(K - \exp(f(T) + x + y), 0)].$$

Dabei ist zu berücksichtigen, dass der Erwartungswert unter einem risikoneutralen Maß berechnet werden muss. Die direkte Übertragung der Bewertungsmethoden, die aus Kapitalmarktmodellen bekannt sind, ist nicht möglich, da der modellierte Strommarkt unvollständig ist. Einerseits ist das Underlying nicht handelbar (keine Leerverkäufe und keine Speichermöglichkeit) und andererseits sind die Preispfade nicht stetig. Somit ergibt sich, dass kein eindeutiges risikoneutrales Maß existiert (siehe [20]). Die Auswahl eines Bewertungsmaßes wird in Kapitel 6 besprochen.

Es folgt sofort aus der obigen Darstellung, dass $e^{-rt}V(x, y, t)$ ein Martingal ist. Desweiteren gilt:

$$d(e^{-rt}V(x, y, t)) = -e^{-rt}(-r + dV).$$

Nun wird zunächst der Term dV abgeschätzt. Sei $V = V(x, y, t) \in C^{2,1,1}$, dann folgt mit Itô's Lemma und [14, Proposition 8.14]:

$$\begin{aligned} dV(X_t, Y_t, t) &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial x} dX_t + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} dt - \beta Y_t \frac{\partial V}{\partial y} dt \\ &\quad + (V(X_t, Y_t + J_t, t) - V(X_t, Y_t, t)) dN_t \end{aligned}$$

Um die erwartete Änderung von dV zu bestimmen, ist es notwendig den Term $\mathbb{E} [(V(X_t, Y_t + J_t, t) - V(X_t, Y_t, t)) dN_t]$ abzuschätzen. Es lässt sich zeigen (vgl. [22, S.96 f] und [14, Proposition 8.13]), dass dieser Term mit

$$\eta \mathbb{E}[V(X_t, Y_t + J_t, t) - V(X_t, Y_t, t)] \tag{4.7}$$

wächst. Oder anders formuliert ist die Differenz aus beiden Termen ein Martingal. Der Term in (4.7) wird als Kompensation des Sprungprozesses bezeichnet. Es folgt also ins-

gesamt:

$$rV = \frac{\partial V}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} + \alpha(\mu(t) - x) \frac{\partial V}{\partial x} - \beta y \frac{\partial V}{\partial y} + \eta \mathbb{E}[V(x, x + J, t) - V(x, y, t)]$$

Oder wenn der Erwartungswert ausgeschrieben wird:

$$\begin{aligned} rV = & \frac{\partial V}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} + \alpha(\mu(t) - x) \frac{\partial V}{\partial x} - \beta y \frac{\partial V}{\partial y} \\ & + \eta \int_{-\infty}^{\infty} ((V(x, x + z, t) - V(x, y, t))) f_J(z) dz, \end{aligned}$$

wobei f_J die Dichte der exponentialverteilten Zufallsvariable J ist. Somit ist eine PIDE für den fairen Wert eines europäischen Puts hergeleitet. Der europäische Call lässt sich über die Put-Call-Parität bewerten.

Kapitel 5

Mathematische Analyse des Kluge-Modells

Zunächst wird in diesem Kapitel ein Existenzresultat für die PIDE, welche den Wert eines europäischen Puts beschreibt, formuliert. Dieses wird mit Hilfe der Halbgruppentheorie bewiesen. Im zweiten Teil dieses Abschnitts wird ein Maximumsprinzip für die PIDE gezeigt.

5.1 Existenz einer Lösung

Um im Laufe der Arbeit einen Optionswert für europäische Optionen zu bestimmen, ist zunächst zu klären, ob diese überhaupt im Kluge-Modell bewertbar sind, d.h. ob die zugehörige PIDE überhaupt eine Lösung besitzt. Im Folgenden wird die Existenz einer Lösung, der in Abschnitt 4.2 hergeleiteten PIDE, im Sinne des Satzes (2.15) gezeigt. Wobei in diesem Fall $X = L^2(\Omega)$ ist und $D(A) = H^{2,1}(\Omega)$, wie sich später zeigen wird. Der Beweis nutzt die in Kapitel 2 erläuterte Halbgruppentheorie. Formulieren wir zunächst unser Problem:

$$\begin{aligned} & \frac{\partial u}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2} - \beta y \frac{\partial u}{\partial y} + (\mu(t) - x) \frac{\partial u}{\partial x} - ru \\ & + \eta \int_{\mathbb{R}} (u(x, y + z, t) - u(x, y, t)) f_J(z) dz = 0 \end{aligned} \quad (5.1)$$

auf $[0, T] \times \Omega = [0, T] \times [x_{\min}, x_{\max}] \times [0; \infty)$

und

$$u(x, y, T) = \max(K - \exp(g(T) + x + y), 0) \quad \text{für } \forall x, y \in \Omega \quad (5.2)$$

sowie

$$u(x_{\max}, y, t) = g(x_{\max}, y, t) \quad (5.3)$$

und

$$u(x_{\min}, y, t) = g(x_{\min}, y, t). \quad (5.4)$$

Dabei ist g eine glatte und samt ihrer Ableitungen integrierbare Funktion (vgl. Abschnitt 5.1.5). Als erstes führen wir eine Variablentransformation der Art $\tau = T - t$ durch. Somit ist nun ein Anfangswertproblem und kein Endwertproblem wie zuvor zu lösen. Der nächste Schritt ist eine weitere Variablentransformation, die den Beweis wesentlich vereinfacht. Zeigen wir die Existenz einer Lösung für die transformierte Gleichung, so ist durch die Rücktransformation eine Lösung unseres Ausgangsproblems gefunden. Die Transformation lautet folgendermaßen: $\tilde{y} = y \exp(-\beta\tau)$. Nun wird die Differentialgleichung für die Funktion $w(x, \exp(-\beta\tau)y, \tau) = u(x, y, \tau)$ hergeleitet. Es gilt:

$$\begin{aligned} \frac{\partial w(x, \tilde{y}, \tau)}{\partial x} &= \frac{\partial u(x, y, \tau)}{\partial x} \\ \exp(-\beta\tau) \frac{\partial w(x, \tilde{y}, \tau)}{\partial \tilde{y}} &= \frac{\partial u(x, y, \tau)}{\partial y} \\ -\beta y \exp(-\beta\tau) \frac{\partial w(x, \tilde{y}, \tau)}{\partial \tilde{y}} + \frac{\partial w(x, \tilde{y}, \tau)}{\partial \tau} &= \frac{\partial u(x, y, \tau)}{\partial \tau}. \end{aligned}$$

Der Integralterm wird wie folgt transformiert:

$$\begin{aligned} &\eta \int_{\mathbb{R}} (u(x, y + z, \tau) - u(x, y, \tau)) f_J(z) dz \\ &= \eta \int_{\mathbb{R}} (w(x, (y + z) \exp(-\beta\tau), \tau) - w(x, y \exp(-\beta\tau), \tau)) f_J(z) dz \\ &\iff \eta \int_{\mathbb{R}} (u(x, y + z, \tau) - u(x, y, \tau)) f_J(z) dz \\ &= \eta \int_{\mathbb{R}} (w(x, \tilde{y} + \tilde{z}, \tau) - w(x, \tilde{y}, \tau)) f_J(\exp(\beta\tau)\tilde{z}) \exp(\beta\tau) d\tilde{z}, \\ &\text{wobei } \tilde{z} = \exp(-\beta\tau)z \text{ ist.} \end{aligned}$$

Also lautet die transformierte Differentialgleichung

$$\frac{\partial w}{\partial \tau} = A(\tau)(w(\tau)) + K(\tau)(w(\tau))$$

mit

$$A(\tau)(w(\tau)) = \frac{\sigma^2}{2} \frac{\partial^2 w}{\partial x^2}(\tau) + \alpha(\mu(\tau) - x) \frac{\partial w}{\partial x}(\tau) \quad (5.5)$$

und

$$\begin{aligned} &K(\tau)(w(\tau)) \\ &= r w + \eta \int_{\mathbb{R}} (w(x, \tilde{y} + \tilde{z}, \tau) - w(x, \tilde{y}, \tau)) f_J(\exp(\beta\tau)\tilde{z}) \exp(\beta\tau) d\tilde{z}. \end{aligned} \quad (5.6)$$

Als letzte Transformation bereinigen wir die PIDE noch um die Randwerte am x -Rand, um mit Hilfe des Gauß'schen Integralsatzes die Selbstadjungiertheit eines Teiloperators zu zeigen. Wir definieren: $v(x, \tilde{y}, \tau) = w(x, \tilde{y}, \tau) - g(x, \tilde{y}, \tau)$. Daraus folgt:

$$\frac{\partial v}{\partial \tau} = \frac{\partial w}{\partial \tau} - \frac{\partial g}{\partial \tau}.$$

Also gilt

$$\frac{\partial v}{\partial \tau} = A(\tau)(v(\tau)) + K(\tau)(v(\tau)) - A(\tau)(g(\tau)) - K(\tau)(g(\tau)) - \frac{\partial g}{\partial \tau}(\tau).$$

Dies lässt sich umschreiben zu

$$\frac{\partial v}{\partial \tau} = A(\tau)(v(\tau)) + K(\tau)(v(\tau)) + f(x, y, \tau)$$

mit

$$f(x, y, \tau) = -K(\tau)(g(x, y, \tau)) - A(\tau)(g(x, y, \tau)) - \frac{\partial g}{\partial \tau}(x, y, \tau). \quad (5.7)$$

Für den Beweis brauchen wir noch einen weiteren Operator. Wir definieren:

$$A_0 v = \frac{\sigma^2}{2} \frac{\partial^2 v}{\partial x^2}. \quad (5.8)$$

Um die Notation zu vereinfachen schreiben wir im Folgenden t für τ und y für \tilde{y} . Der Beweis für die Existenz einer Lösung obiger Gleichung gliedert sich jetzt wie folgt:

1. Es wird gezeigt, dass der Operator A_0 eine analytische Halbgruppe erzeugt.
2. Durch ein Störungsargument wird gezeigt, dass $A(t) = A_0 + (A(t) - A_0)$ für alle $t \in [0, T]$ eine analytische Halbgruppe erzeugt.
3. Es wird gezeigt, dass $(A + K)(t)$ für alle $t \in [0, T]$ eine analytische Halbgruppe erzeugt.
4. Schließlich wird gezeigt, dass $t \rightarrow (A + K)(t)$ Lipschitz-stetig ist.
5. Es wird gezeigt, dass $f(\cdot, x) : [0, T] \mapsto L^2(\Omega)$ Lipschitz-stetig ist für alle $x \in L^2(\Omega)$.

Im Folgenden bezeichnen wir mit $H_0^{2,0}(\Omega)$ den entsprechenden Sobolevraum mit homogenen x -Rändern.

5.1.1 A_0 erzeugt analytische Halbgruppe

Lemma 5.1. *Der Operator $A_0 : H_0^{2,0}(\Omega) \rightarrow L^2(\Omega)$, definiert durch (5.8), erzeugt eine analytische Halbgruppe.*

Beweis. Nach den Satz 2.13 genügt es zu zeigen:

1. $D(A_0)$ ist dicht im Ausgangsraum, also $L^2(\Omega)$.
2. A_0 ist selbstadjungiert.
3. A_0 ist negativ semidefinit.
4. $A_0 - \lambda I$ ist surjektiv für $\lambda > 0$.

Ad 1:

$D(A_0) = H_0^{2,0}(\Omega)$, also enthält $D(A_0)$ den Raum $H_0^2(\Omega)$ und ist somit dicht in $L^2(\Omega)$.

Ad 2:

Sei $u, v \in D(A_0)$

$$\begin{aligned} (A_0 u, v)_{L^2(\Omega)} &= \int_{\Omega} \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial^2 x} v \, dx \\ &\stackrel{G.I.S.}{=} - \int_{\Omega} \frac{\sigma^2}{2} \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \, dx \stackrel{G.I.S.}{=} \int_{\Omega} \frac{\sigma^2}{2} u \frac{\partial^2 v}{\partial^2 x} \, dx = (u, A_0 v)_{L^2(\Omega)}. \end{aligned}$$

Des Weiteren gilt $D(A_0) = D(A_0^*)$ also ist 2 gezeigt.

Ad 3:

Sei $u \in D(A_0)$

$$(A_0 u, u)_{L^2(\Omega)} = \int_{\Omega} \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial^2 x} u \, dx \stackrel{P.I.}{=} - \int_{\Omega} \frac{\sigma^2}{2} \frac{\partial u}{\partial x} \frac{\partial u}{\partial x} \, dx = -\frac{\sigma^2}{2} \left\| \frac{\partial u}{\partial x} \right\|_2^2 \leq 0.$$

Ad 4:

Um Surjektivität zu zeigen, muss nachgewiesen werden, dass für alle $f(x, y) \in L^2(\Omega)$ und alle $\lambda > 0$ ein $u \in D(A_0)$ existiert d.d. $(A_0 - \lambda I)u = f$ gilt. Um das zu zeigen, bedienen wir uns der Theorie gewöhnlicher Differentialgleichungen.

Beh: $A_0 - \lambda I$ ist surjektiv für $\lambda > 0$.

Wir betrachten zuerst die homogene Differentialgleichung

$$A_0 u = \lambda u.$$

Bekanntlich bilden $e^{\sqrt{\lambda}x}$ und $e^{-\sqrt{\lambda}x}$ eine Basis des homogenen Lösungsraumes. Die Lösung des inhomogenen Problems lässt sich mit Hilfe der Variation der Konstanten bestimmen. In diesem speziellen Fall lässt sich $R_{\lambda}(A_0)$ sogar explizit angeben.

Somit ist nach den Ausführungen im Grundlagenteil gezeigt, dass A_0 eine analytische Halbgruppe erzeugt. \square

5.1.2 A erzeugt analytische Halbgruppe

Lemma 5.2. *Der Operator $A(t) : H_0^{2,0}(\Omega) \rightarrow L^2(\Omega)$ definiert durch (5.5) (mit $t \in [0, T]$ beliebig) erzeugt eine analytische Halbgruppe.*

Beweis. Es genügt nach Satz 2.14 zu zeigen:

1. $D(A_0) \subseteq D(A(t) - A_0)$.
2. $\|(A(t) - A_0)u\|_2 \leq C(\|u\|_2 \rho^\gamma + \|A_0 u\|_2 \rho^{1-\gamma})$ für ein $\gamma \in (0, 1)$, $\rho > 0$ beliebig klein und $u \in D(A_0)$.
3. $A(t) - A_0$ ist abgeschlossen.

Ad 1:

Die Aussage folgt daraus, dass $D(A(t) - A_0) = H_0^{1,0}(\Omega) \supset H_0^{2,0}(\Omega) = D(A_0)$ gilt.

Ad 2: Sei $u \in H_0^{2,0}(\Omega)$, dann folgt

$$\begin{aligned} \|(A(t) - A_0)u\|_2 &= \|(\mu(t) - x) \frac{\partial u}{\partial x}\|_2 \leq C \left\| \frac{\partial u}{\partial x} \right\|_2^{\frac{1}{2}} \stackrel{\text{P.I. und Ca.-Schw.}}{\leq} C \left\| \frac{\partial^2 u}{\partial x^2} \right\|_2 \|u\|_2^{\frac{1}{2}} \\ &\leq C \|u\|_2^{\frac{1}{2}} \|A_0 u\|_2^{\frac{1}{2}} \stackrel{\text{Young-Ungl.}}{\leq} C (\|u\|_2 \rho^{\frac{1}{2}} + \rho^{-\frac{1}{2}} \|A_0 u\|_2). \end{aligned}$$

Also haben wir die notwendige Abschätzung mit $\gamma = 0.5$ gefunden.

Ad 3:

Es bleibt zu zeigen, dass $A(t) - A_0$ abgeschlossen ist. Dies folgt sofort aus der Beschränktheit von $A(t) - A_0$ auf $H_{1,0}(\Omega)$, für $t \in [0, T]$ beliebig und fest. \square

5.1.3 $A + K$ erzeugt analytische Halbgruppe

Lemma 5.3. *Der Operator $(A + K)(t) : H_0^{2,0}(\Omega) \rightarrow L^2(\Omega)$ definiert durch (5.5) und (5.6) (mit $t \in [0, T]$ beliebig) erzeugt eine analytische Halbgruppe.*

Beweis. Es genügt nach Satz 2.14 zu zeigen:

1. $D(K(t)) \supseteq D(A(t))$,
2. $\|K(t)u\|_2 \leq b\|u\|_2 \forall u \in D(A)$,
3. $K(t)$ abgeschlossen.

Ad 1:

Die Aussage folgt daraus, dass $D(K(t)) = L^2(\Omega)$ für alle $t \in [0, T]$ gilt.

Ad 2:

Im Folgenden wird die Norm des Integraloperator abgeschätzt. Sei $u \in H_0^{2,0}(\Omega)$, dann folgt

$$\begin{aligned}
\|K(t)u\|_2^2 &\leq \|ru\|_2^2 + \|\eta \int_0^\infty (u(x, y+z) - u(x, y)) f_J(e^{\beta t} z) e^{\beta t} dz\|_2^2 \\
&\leq r\|u\|_2^2 + \eta \int_{x_{\min}}^{x_{\max}} \int_0^\infty \left(\int_0^\infty (u(x, y+z) - u(x, y)) \right. \\
&\quad \left. \times \sqrt{f_J(e^{\beta t} z) e^{\beta t}} \sqrt{f_J(e^{\beta t} z) e^{\beta t}} dz \right)^2 dy dx \\
&\stackrel{\Delta\text{-Ung.}}{\leq} r\|u\|_2^2 + \eta \int_{x_{\min}}^{x_{\max}} \int_0^\infty \int_0^\infty |u(x, y+z) - u(x, y)|^2 f_J(e^{\beta t} z) e^{\beta t} dz \\
&\quad \times \underbrace{\int_0^\infty f_J(e^{\beta t} z) e^{\beta t} dz}_{\leq L} dy dx \\
&\leq r\|u\|_2^2 + 2L\eta \int_{x_{\min}}^{x_{\max}} \int_0^\infty \int_0^\infty |u(x, y+z)|^2 f_J(e^{\beta t} z) e^{\beta t} dz dy dx \\
&\quad + 2L\eta \underbrace{\int_{x_{\min}}^{x_{\max}} \int_0^\infty \int_0^\infty |u(x, y)|^2 f_J(e^{\beta t} z) e^{\beta t} dz dy dx}_{\leq L\|u\|_2^2} \\
&\leq r\|u\|_2^2 + 2L\eta \int_0^\infty f_J(e^{\beta t} z) e^{\beta t} \int_0^\infty \int_{x_{\min}}^{x_{\max}} |u(x, y+z)|^2 dx dy dz + C'\|u\|_2^2 \\
&\stackrel{\tilde{y}=z+y}{\leq} (C' + r)\|u\|_2^2 + 2L\eta \int_0^\infty f_J(e^{\beta t} z) e^{\beta t} \int_z^\infty \int_{x_{\min}}^{x_{\max}} |u(x, \tilde{y})|^2 dx d\tilde{y} dz \\
&\leq C\|u\|_2^2.
\end{aligned}$$

C ist unabhängig von t , da das Zeitintervall beschränkt ist.

Ad 3:

Die Aussage folgt daraus, dass $L^2(\Omega)$ der Definitionsbereich von $K(t)$ ist und $K(t)$ beschränkt ist (für $t \in [0, T]$ beliebig). \square

5.1.4 Lipschitz-Stetigkeit von $t \mapsto (A + K)(t)$

Lemma 5.4. *Die Abbildung $t \mapsto (A + K)(t)$, wobei A und K wie in 5.1.3 definiert sind, ist Lipschitz-stetig.*

Beweis. Wir zeigen die Lipschitz-Stetigkeit von $(A + K)(t)$ direkt, indem wir die Norm von $(A + K)(t) - (A + K)(t')$ nach oben abschätzen.

$$\begin{aligned}
 & \| (A + K)(t) - (A + K)(t') \|_{\mathcal{L}(H_0^{2,0}(\Omega), L^2(\Omega))}^2 = \sup_{u \in D(A+K), \|u\|_{H_0^{2,0}}=1} \| (\mu(t) - \mu(t')) \frac{\partial u}{\partial x} \\
 & + \eta \int_0^\infty (u(x, y+z) - u(x, y)) (f_J(e^{\beta t} z) e^{\beta t} - f_J(e^{\beta t'} z) e^{\beta t'}) dz \|_2^2 \\
 & \stackrel{\text{MWS}}{\leq} \sup_{u \in D(A+K), \|u\|_{H_0^{2,0}}=1} |t - t'|^2 \|\mu'(t)\|_\infty^2 \left\| \frac{\partial u}{\partial x} \right\|_2^2 \\
 & + \eta \int_{x_{\min}}^{x_{\max}} \int_0^\infty \left(\int_0^\infty |u(x, y+z) - u(x, y)| |t - t'| \right. \\
 & \quad \left. \times \underbrace{\left(\sup_{t \in [0, T]} \frac{d}{dt} f_J(e^{\beta t} z) e^{\beta t} \right)}_{=h(z)} dz \right)^2 dy dx \\
 & \stackrel{\Delta\text{-Ung.}}{\leq} \sup_{u \in D(A+K), \|u\|_{H_0^{2,0}}=1} |t - t'|^2 \|\mu'(t)\|_\infty^2 \left\| \frac{\partial u}{\partial x} \right\|_2^2 \\
 & + \eta \int_{x_{\min}}^{x_{\max}} \int_0^\infty \int_0^\infty |t - t'|^2 |u(x, y+z) - u(x, y)|^2 h(z) dz \underbrace{\int_0^\infty h(z) dz}_{=L} dy dx \\
 & \leq \sup_{u \in D(A+K), \|u\|_{H_0^{2,0}}=1} C_1 |t - t'|^2 \left\| \frac{\partial u}{\partial x} \right\|_2^2 \\
 & + 2L |t - t'|^2 \int_0^\infty h(z) \int_0^\infty \int_{x_{\min}}^{x_{\max}} |u(x, y+z)|^2 dx dy dz \\
 & + 2L |t - t'|^2 \int_0^\infty h(z) \int_0^\infty \int_{x_{\min}}^{x_{\max}} |u(x, y)|^2 dx dy dz \\
 & \leq C |t - t'|^2
 \end{aligned}$$

Es wurde die Beschränktheit der Ableitung von $\mu(t)$ in t auf $[0, T]$ und die Integrierbarkeit von $h(z)$ benutzt. Bei der letzten Abschätzung ist die Argumentation analog zu Abschnitt 5.1.3. \square

5.1.5 Lipschitz-Stetigkeit der rechten Seite

Lemma 5.5. $f(\cdot, x) : [0, T] \rightarrow L^2(\Omega)$ definiert wie in (5.7) ist Lipschitz-stetig.

Beweis. Die Lipschitz-Stetigkeit von $f(\cdot, x)$ kann unter Benutzung der Glattheit von g gezeigt werden. Dazu ist zunächst folgende Abschätzung notwendig.

$$\begin{aligned} \|f(t, x, y) - f(t', x, y)\|_2 &\leq \underbrace{\|K(t)g(t, x, y) - K(t')g(t', x, y)\|_2}_{A1} \\ &\quad + \underbrace{\|A(t)g(t, x, y) - A(t')g(t', x, y)\|_2}_{A2} \\ &\quad + \underbrace{\left\| \frac{\partial g}{\partial t}(t, x, y) - \frac{\partial g}{\partial t}(t', x, y) \right\|_2}_{A3} \end{aligned}$$

Die einzelnen Terme lassen sich jetzt folgendermaßen abschätzen. $A1 \leq C_1 |t - t'|$, falls wir $\|\sup_{t \in [0, T]} \frac{d}{dt} g(t, x, y)\|_2 < \infty$ fordern. $A2$ lässt sich gegen $C_2 |t - t'|$ abschätzen, falls wir $\|\sup_{t \in [0, T]} \frac{d}{dt} (\mu(t) - x) \frac{\partial g}{\partial x}\|_2 < \infty$ und $\|\sup_{t \in [0, T]} \frac{d}{dt} \frac{\partial^2 g}{\partial x^2}\|_2 < \infty$ fordern. $A3$ schließlich lässt sich gegen $C_3 |t - t'|$ abschätzen, falls wir $\|\sup_{t \in [0, T]} \frac{\partial^2 g}{\partial t^2}\|_2 < \infty$ annehmen. Fordern wir nun, dass g diese Voraussetzungen erfüllt, also beispielsweise $g \in C^2([0, T]; H^2(\Omega))$, so folgt insgesamt:

$$\|f(t, x, y) - f(t', x, y)\|_2 \leq (C_1 + C_2 + C_3) |t - t'| = C |t - t'|.$$

□

5.1.6 Existenzresultat

In diesem Abschnitt wird der Existenzsatz formuliert. Der Beweis für die Existenz einer Lösung der PIDE folgt im wesentlichen aus den Resultaten der vorangegangenen Abschnitte.

Satz 5.1. *Die Differentialgleichung*

$$\frac{\partial v}{\partial t}(x, y, t) = A(t)(v(x, y, t)) + K(t)(v(x, y, t)) + f(t, x, y)$$

mit A wie in Abschnitt 5.1.2, K wie in Abschnitt 5.1.3 und f wie in Abschnitt 5.1.5, sowie der Anfangsbedingung

$$v(x, y, 0) = \max(K - \exp(x + y + f(T)), 0)$$

und homogenem Rand in x , besitzt eine eindeutige Lösung v mit $v \in C([0, T], L^2(\Omega)) \cap C^1((0, T], L^2(\Omega)) \cap C((0, T], H_0^{2,0}(\Omega))$.

Beweis. Nach Satz 2.15 genügt es folgende Bedingungen zu zeigen:

1. $D((A + K)(t))$ ist konstant für alle t und dicht in $L^2(\Omega)$.
2. $(A + K)(t)$ erzeugt eine analytische Halbgruppe für alle $t \in [0, T]$.
3. Die Abbildung $t \rightarrow f(\cdot, x)$ ist Lipschitz-stetig.
4. Die Abbildung $t \rightarrow (A + K)(t)$ ist Lipschitz-stetig.
5. Und schließlich gilt $u_0 \in L^2(\Omega)$.

1 und 2 wurden in Abschnitt 5.1.3 gezeigt, 3 wurde nach Abschnitt 5.1.5 gezeigt, 4 wurde in Abschnitt 5.1.4 gezeigt und 5 gilt offensichtlich. Somit ist die Existenz einer Lösung bewiesen. Die Eindeutigkeit folgt aus Satz A.2. \square

Daraus ergibt sich durch Rücktransformation die Existenz einer Lösung des ursprünglichen Problems.

Satz 5.2. *Das Problem (5.1)-(5.4) besitzt eine eindeutige Lösung u mit $u \in C([0, T], L^2(\Omega)) \cap C^1((0, T], L^2(\Omega)) \cap C((0, T], H^{2,1}(\Omega))$.*

Beweis. \square

Die Existenz einer Lösung für den europäischen Call folgt aus der Put-Call-Parität. Anzumerken ist, dass der Beweis analog auch für Neumannrandwerte am x -Rand durchgeführt werden kann. Die Wahl der Randwerte ist bei der Bewertung von Optionen oft ein schwieriges Problem. Bei der numerischen Auswertung werden wir sowohl Neumann- als auch Dirichletränder analysieren.

5.2 Maximumprinzip der Differentialgleichung

Im Folgenden wird gezeigt, dass die oben hergeleitete Differentialgleichung für den europäischen Put (5.1)-(5.4) ein Maximumprinzip erfüllt. Im Allgemeinen erfüllen parabolische Probleme mit einem Integraloperator kein Maximumprinzip mehr, da dieses durch den Integraloperator zerstört wird. Jedoch hat der Integraloperator in unserem Fall so viel Struktur, dass das Maximumprinzip erhalten bleibt.

Lemma 5.6. *Sei $u \in C^1([0, T]; C^{2,1}(\Omega)) \cap C^1([0, T]; H^{2,1}(\Omega))$ die eindeutige Lösung von (5.1) mit Endwerten (5.2) und Randwerten (5.3), (5.4), dann hat u kein positives globales Maximum im Inneren des Gebietes Ω .*

Beweis. Beweis durch Widerspruch:

Annahme: Es existiert ein positives Maximum im Inneren des Gebietes Ω im Punkt $M = (\bar{x}, \bar{y}, \bar{t})$, daraus folgt

$$\frac{\partial u}{\partial t}(\bar{x}, \bar{y}, \bar{t}) = 0, \quad \frac{\partial u}{\partial x}(\bar{x}, \bar{y}, \bar{t}) = 0 \quad \text{und} \quad \frac{\partial u}{\partial y}(\bar{x}, \bar{y}, \bar{t}) = 0$$

Des Weiteren gilt

$$\frac{\partial^2 u}{\partial x^2}(\bar{x}, \bar{y}, \bar{t}) \leq 0.$$

Der Integraloperator lässt sich folgendermaßen abschätzen:

$$\begin{aligned} K(\bar{t})(u(\bar{x}, \bar{y}, \bar{t})) &= \eta \int_{\mathbb{R}} (u(\bar{x}, \bar{y} + z, \bar{t}) - u(\bar{x}, \bar{y}, \bar{t})) f_J(z) dz \\ &= \frac{\eta}{\mu_J} \int_0^\infty (u(\bar{x}, \bar{y} + z, \bar{t}) - u(\bar{x}, \bar{y}, \bar{t})) e^{-\frac{z}{\mu_J}} dz \leq 0 \end{aligned}$$

Die letzte Abschätzung folgt, da das globale Maximum in $(\bar{x}, \bar{y}, \bar{t})$ angenommen wird und somit $u(\bar{x}, \bar{y} + z, \bar{t}) \leq u(\bar{x}, \bar{y}, \bar{t})$ gilt. Also haben wir insgesamt:

$$\underbrace{\frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2}(\bar{x}, \bar{y}, \bar{t}) + K(\bar{t})(u(\bar{x}, \bar{y}, \bar{t}))}_{\leq 0} = \underbrace{ru(\bar{x}, \bar{y}, \bar{t})}_{> 0}.$$

Dies ist ein Widerspruch. □

Es gibt also kein positives Maximum im Inneren des Gebietes. Genauso lässt sich zeigen, dass es kein negatives Minimum im Inneren gibt. Als nächstes zeigen wir, dass ein positives Maximum nur auf einem bestimmten Teil des Randes liegen kann.

Lemma 5.7. *Sei u wie in Lemma 5.6, dann hat u auf $R = \{0\} \times \mathring{\Omega} \cup \{0\} \times \{0\} \times (x_{\min}, x_{\max})$ kein positives globales Maximum.*

Beweis. Beweis durch Widerspruch.

1. Annahme : Es existiert ein positives Maximum auf $\{0\} \times \mathring{\Omega}$

$$\frac{\partial u}{\partial t}(\bar{x}, \bar{y}, \bar{t}) + \underbrace{\frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2}(\bar{x}, \bar{y}, \bar{t}) + K(\bar{t})(u(\bar{x}, \bar{y}, \bar{t}))}_{\leq 0} = \underbrace{ru(\bar{x}, \bar{y}, \bar{t})}_{> 0}$$

Daraus folgt $\frac{\partial u}{\partial t} > 0$. Dies ist aber ein Widerspruch, da für ein Maximum, auch wenn es am Rand liegt, die partiellen Ableitungen in alle Richtungen nicht positiv sein müssen.

2. Annahme : Es existiert ein positives Maximum auf $\{0\} \times \{0\} \times (x_{\min}, x_{\max})$

$$\frac{\partial u}{\partial t}(\bar{x}, \bar{y}, \bar{t}) + \underbrace{\frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2}(\bar{x}, \bar{y}, \bar{t}) + K(\bar{t})(u(\bar{x}, \bar{y}, \bar{t}))}_{\leq 0} = \underbrace{ru(\bar{x}, \bar{y}, \bar{t})}_{> 0}.$$

Daraus folgt $\frac{\partial u}{\partial t} > 0$. Dies ist auf analoge Weise zu oben ein Widerspruch. \square

Satz 5.3 (Maximumprinzip). *Die eindeutige Lösung $u \in C^1([0, T]; C^{2,1}(\Omega)) \cap C^1([0, T]; H^{2,1}(\Omega))$ von (5.1) mit Endwerten (5.2) und Randwerten (5.3), (5.4) nimmt negative globale Minima und positive globale Maxima nur bei Fälligkeit ($t = T$) oder an den x -Rändern an.*

Beweis. Die Aussage für positive Maxima ergibt sich sofort aus den beiden vorhergehenden Lemmata. Die Argumentation für ein negatives Minimum verläuft analog. \square

Aus dem Maximumprinzip für die Differentialgleichung folgt einerseits die L^∞ -Stabilität der Lösung und auch die Eindeutigkeit. Gäbe es nämlich zwei solcher Lösungen u, v die (5.1)-(5.4) erfüllen, so würde dies auch für $w = u - v$ mit homogenen Rand- und Endwerten gelten. Die Funktion w müsste auf dem gesamten Gebiet $[0; T] \times \Omega$ konstant Null sein. Wäre dies nämlich nicht der Fall und gäbe es OE ein $P = (x, y, t)$ mit $u(P) > 0$, dann gäbe es einen Punkt an dem ein positives (lokales) Maximum angenommen wird. Dies ist ein Widerspruch, da Rand- und Endwerte homogen sind. Also muss $u = v$ gelten.

Das oben bewiesene Maximumprinzip gilt nur für starke Lösungen. Es lässt sich jedoch auch für schwache Lösungen durch Approximation mit glatten Funktion die L^∞ -Stabilität zeigen.

Kapitel 6

Kalibrierung des Kluge-Modells

Wir kalibrieren das Modell an Daten der Jahre 2005-2007. Dabei wird zunächst die Strompreisdynamik unter dem realen Maß bestimmt. Diese ist jedoch für die Bewertung von Derivaten ungeeignet. Daher kalibrieren wir danach unseren Prozess unter einem Bewertungsmaß \mathbb{Q} , indem wir die Parameter mit Hilfe von Forwarddaten anpassen. Da sich eine explizite Darstellung des Forwards im Kluge-Modell gewinnen lässt, können wir mit realen Forwarddaten einen Fit durchführen. Dafür müssen bestimmte Annahmen darüber getroffen werden, welche Risiken bepreist werden.

6.1 Kalibrierung des Prozesses unter dem realen Maß

Zunächst bereinigen wir die Daten um die saisonale Komponente $f(t)$. Die Funktion $f(t)$ wird als Summe von periodischen Funktionen gewählt. Dabei läuft t von 0 bis 365:

$$f(t) = f_{Jahr}(t) + f_{Woche}(t) + f_{Tag}(t)$$

mit

$$\begin{aligned} f_{Jahr}(t) &= j_1 \sin(2\pi/365t) + j_2 \cos(2\pi/365t) \\ &\quad + j_3 \sin(4\pi/365t) + j_4 \sin(4\pi/365t) + j_5 t + j_6 \end{aligned}$$

$$f_{Woche}(t) = \sum_{i=1}^7 \mathbf{1}_{N(i)}(t) d_i$$

und

$$\begin{aligned}
 f_{Tag}(t) = & (t_1 \sin(2\pi t) + t_2 \cos(2\pi t))\mathbf{1}_{\{N(1), \dots, N(5)\}}(t) \\
 & + (t_3 \sin(2\pi t) + t_4 \cos(2\pi t))\mathbf{1}_{N(6)}(t) + (t_5 \sin(2\pi t) + t_6 \cos(2\pi t))\mathbf{1}_{N(7)}(t) \\
 & + t_7 \mathbf{1}_{M_i}(t) + t_8 \mathbf{1}_{F_e}(t) + t_9 \mathbf{1}_{N_a}(t).
 \end{aligned}$$

Die Funktionen $\mathbf{1}_{N(i)}$ sind Indikatorfunktionen für den i -ten Wochentag, die Funktion $\mathbf{1}_{M_i}$ ist eine Indikatorfunktion für die Mittagsstunde, $\mathbf{1}_{F_e}$ für die Feierabendstunde um 19 Uhr und $\mathbf{1}_{N_a}$ für die Nachtstunden zwischen 4 Uhr und 5 Uhr¹. Die Funktion f wird mit Hilfe der `lsqcurvefit`-Routine in Matlab gefittet. Um Spikes aus der Kalibrierung zu eliminieren, wird der Fit iterativ durchgeführt. Dabei wird zunächst ein Fit mit den echten Daten durchgeführt. Danach werden die Daten an den Stellen, an denen Ausreißer auftreten, korrigiert und der Fit wird wiederholt (vgl. [17, S.18]). Dem Algorithmus muss einerseits ein Abbruchkriterium und andererseits eine Toleranzschwelle für die Korrektur der Daten übergeben werden. Als Abbruchkriterium wählen wir eine Schranke für die Differenz der L^2 -Norm der saisonalen Funktion zweier aufeinanderfolgenden Schätzungen (0.01). Als Toleranzschwelle wird ein Vielfaches der Standardabweichung der desaisonalisierten Daten gewählt (3.5). Die Tagessaisonalität und der restliche Saisoneinfluss werden einzeln kalibriert. Die Ergebnisse sind stabil und in Tabelle 6.1 angegeben. Als Nächstes

Tabelle 6.1: Kalibrierte Saisonparameter

Parameter \ i	1	2	3	4	5	6	7	8	9
j_i	-0.12	0.14	-0.06	0.03	0.00	3.00	-	-	-
d_i	0.96	1.04	1.04	1.01	0.93	0.67	0.41	-	-
t_i	-0.12	-0.28	-0.08	-0.10	-0.18	0.05	0.11	-0.42	0.19

wurden die Parameter der beiden Prozesse Y und X geschätzt. Um die Schätzung durchführen zu können, müssen wir in unseren Daten die beiden Prozesse trennen, um aus einem beobachteten Preis S_t den Anteil X_t und Y_t abzuleiten. Dazu wird ein rekursiver Filter auf die desaisonalisierten Preise Z angewendet. Hierbei wird zunächst die Standardabweichung der Zuwächse von Z bestimmt. Danach werden Spikes an den Stellen, an denen die Zuwächse von Z ein Vielfaches (c) der Standardabweichung überschreiten, definiert. Im nächsten Schritt wird der zugehörige Spikeprozess Y erzeugt. Dazu müssen die Parameter μ_J und β übergeben werden. Diese werden so gewählt, dass sie nahe an den weiter unten erklärten Schätzungen dieser Parameter liegen. Nun wird das Verfahren mit $Z_t / \exp(Y_t)$ wiederholt. Die Schwelle für die Identifikation der Spikes wird so gewählt, dass der Spikeprozess die typischen Charakteristika aufweist. In Abbildung 6.1 sind die

¹Im Laufe der Arbeit werden wir eine differenzierbare Saisonkurve benötigen, daher rechnen wir an entsprechender Stelle mit einer hinreichend glatten Approximation der Funktion $f(t)$ und identifizieren diese wieder mit $f(t)$.

getrennten Prozesse dargestellt. Nun können die Parameter der einzelnen Prozesse geschätzt werden.

Die Verteilung der Zuwächse von X zwischen zwei Zeitpunkten (in unserem Fall von

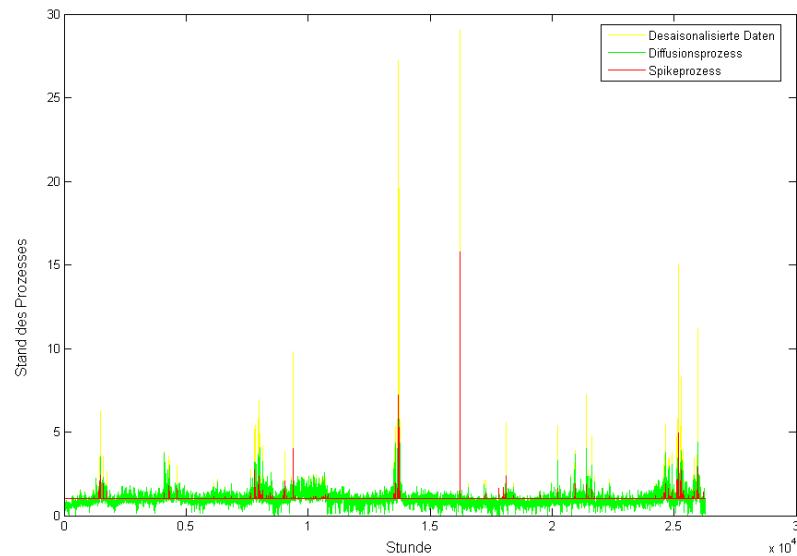


Abbildung 6.1: Getrennte desaisonalisierte Preise

einer Stunde zur nächsten) lässt sich explizit durch Lösen einer stochastischen Differentialgleichung angeben. Somit können die Parameter des Prozesses mit einem ML-Ansatz ermittelt werden. Wir wissen, dass der Prozess X_t folgender SDE genügt:

$$dX_t = -\alpha X_t dt + \sigma dW_t$$

Es lässt sich zeigen (siehe Satz B.2), dass der Stand des Prozesses zum Zeitpunkt t bei bekanntem Stand des Prozesses (x_s) zum Zeitpunkt $s < t$ eine normalverteilte Zufallsvariable ist mit:

$$X_t \sim N \left(x_s e^{-\alpha(t-s)}, \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha(t-s)}) \right) \quad (6.1)$$

Nun können wir mit dem ML-Ansatz beginnen. Wir möchten den Wert der Dichte der gemeinsamen Verteilung unserer Beobachtungen maximieren. Seien $\{x_0, \dots, x_n\}$ die Beobachtungen und $f_{X_1, \dots, X_n}(x_1, \dots, x_n)$ die gemeinsame Dichte, dann gilt:

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1|X_0=x_0}(x_1) \cdots f_{X_n|X_{n-1}=x_{n-1}}(x_n)$$

auf Grund der Markov Eigenschaft des zu Grunde liegenden Prozesses. Da sich die bedingte Dichte wegen (6.1) folgendermaßen darstellen lässt

$$f_{X_t|X_s=x_s}(x) = \sqrt{\frac{\alpha}{\pi\sigma^2(1-e^{-2\alpha(t-s)})}} \exp\left(-\frac{\alpha(x-e^{-\alpha(t-s)}x_s^2)}{\sigma^2(1-e^{-2\alpha(t-s)})}\right),$$

folgt für die log-Likelihood Funktion mit äquidistanten Zeitschritten $\Delta t = t_{i+1} - t_i$

$$L(\sigma, \alpha) = \frac{1}{2} \sum_{i=1}^n \ln\left(\frac{\alpha}{\pi\sigma^2(1-e^{-2\alpha\Delta t})}\right) - \frac{\alpha}{\sigma^2} \sum_{i=1}^n \frac{(x_i - e^{-\alpha\Delta t}x_{i-1})^2}{1 - e^{-2\alpha\Delta t}}.$$

Oder vereinfacht:

$$L(\sigma, \alpha) = \frac{n}{2} \ln\left(\frac{\alpha}{\pi\sigma^2(1-e^{-2\alpha\Delta t})}\right) - \frac{\alpha}{\sigma^2(1-e^{-2\alpha\Delta t})} \sum_{i=1}^n (x_i - e^{-\alpha\Delta t}x_{i-1})^2. \quad (6.2)$$

Mit einer Taylorapproximation erster Ordnung für $(1-e^{-2\alpha\Delta t})$ folgt $1-e^{-2\alpha\Delta t} = 2\alpha\Delta t + o(\Delta t^2)$. Falls wir Terme höherer Ordnung vernachlässigen, folgt

$$L(\sigma, \alpha) \approx \tilde{L}(\sigma, \alpha) = \frac{n}{2} \ln\left(\frac{1}{2\pi\sigma^2\Delta t}\right) - \frac{1}{2\sigma^2\Delta t} \sum_{i=1}^n (x_i - e^{-\alpha\Delta t}x_{i-1})^2.$$

Dies kann nach α differenziert werden

$$\frac{\partial \tilde{L}}{\partial \alpha} = -\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - e^{-\alpha\Delta t}x_{i-1})e^{-\alpha\Delta t}x_{i-1}$$

und es ergibt sich als Schätzer für α :

$$\hat{\alpha} = -\frac{1}{\Delta t} \ln \frac{\sum_{i=1}^n (x_i x_{i-1})}{\sum_{i=1}^n x_{i-1}^2}.$$

Falls wir (6.2) nach σ differenzieren, erhalten wir:

$$\frac{\partial L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2\alpha}{\sigma^3} \sum_{i=1}^n \frac{(x_i - e^{-\alpha\Delta t}x_{i-1})^2}{1 - e^{-2\alpha\Delta t}}.$$

Somit folgt mit $\nabla L = 0$ im Maximum

$$\hat{\sigma}^2 = \frac{2\alpha}{n(1-e^{-2\alpha\Delta t})} \sum_{i=1}^n (x_i - e^{-\alpha\Delta t}x_{i-1})^2.$$

Die hinreichenden Bedingungen für ein Maximum in σ und α lassen sich elementar nachrechnen. Nun haben wir also einen Schätzer für die Parameter des Diffusionsprozesses bestimmt. Die Güte des Schätzers ist auf Grund der Approximation der Likelihoodfunk-

tion bei der Herleitung eines Schätzers für α nicht klar. In [9] wurde jedoch mit anderen Methoden derselbe Schätzer hergeleitet.

Da wir die Verteilung der Zuwächse kennen, ist es sinnvoll die Schätzung durchzuführen und danach alle Ereignisse, die außerhalb einer bestimmten Bandbreite (Vielfaches von Standardabweichungen) liegen, zu eliminieren. Damit stellen wir sicher, dass wirklich nur der Diffusionsprozess kalibriert wird, ohne die Spikeanteile zu berücksichtigen. Nach zehn Iterationen konvergiert das Verfahren.

Zusätzlich wurde die Maximierung der log-Likelihoodfunktion mit Hilfe mit der `fminsearch`-Routine unter Benutzung der Simplex-Suchmethode anstatt des ML-Schätzers durchgeführt. Dabei wurden in jedem Schritt die Parameter durch Maximierung der log-Likelihoodfunktion bestimmt, um danach analog zum ML-Algorithmus unwahrscheinliche Ereignisse herauszufiltern. Die zugehörigen geschätzten Parameter sind in Tabelle 6.2 abgebildet.

Als nächstes ist der Sprungprozess zu kalibrieren. Dieser hat drei zu bestimmende Pa-

Tabelle 6.2: Kalibrierte Parameter des Diffusionsprozesses

	α	σ
ML	1.81	0.53
Simplex	1.82	0.54

parameter β , λ und μ . Für die Schätzung von λ können wir die Häufigkeit von beobachteten Spikes durch die Anzahl der Beobachtungen teilen und erhalten so einen Schätzer. Die Sprunghöhe wird mit einem ML-Schätzer berechnet. Dieser ist für den Parameter der Exponentialverteilung als die mittlere Sprunghöhe gegeben. Die Mean Reversion Geschwindigkeit des Spikeprozesses kann über die Autokorrelationsfunktion geschätzt werden. Diese ist für den Spikeprozess folgendermaßen bestimmt:

$$\rho(t) = e^{-\beta t}.$$

Die empirische Autokorrelation ist folgendermaßen definiert:

$$ACF(k) = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}, \quad (6.3)$$

wobei \bar{Y} das arithmetische Mittel der Zeitreihe ist. Nun lässt sich β bestimmen, indem die theoretische Autokorrelationsfunktion an die empirische gefittet wird. Wir haben für den least-squares Fit die `lsqcurvefit`-Routine in Matlab verwendet. Die kalibrierten Parameter sind in Tabelle 6.3 dargestellt. Anzumerken ist, dass die Trennung der Prozesse und die Schätzung der Parameter des Spikeprozesses auf Grund der Seltenheit der Spikes nicht unproblematisch sind. Daher wird es interessant sein zu sehen, wie sensibel die Optionspreise auf Variationen der Sprungparameter reagieren.

Tabelle 6.3: Kalibrierte Parameter des Spikeprozesses

μ_J	λ	β
0.3621	0.1293	8.2371

6.2 Kalibrierung unter einem risikoneutralen Maß

Da Strommärkte unvollständig sind, gibt es mehr als ein risikoneutrales Maß unter dem Optionen bewertet werden könnten ohne Arbitrage zu generieren. Es stellt sich die Frage, wie das „richtige“ Maß ausgewählt, oder anders formuliert, wie eine risikoneutrale Dynamik für den Prozess definiert werden kann. Eine Möglichkeit dies zu tun, wäre Annahmen über die Nutzenfunktionen der Marktteilnehmer zu treffen und daraus ein Bewertungsmaß abzuleiten. Da uns jedoch risikoneutral bewertete Produkte zur Verfügung stehen, werden wir diese zur Bewertung nutzen. Dies geschieht, indem wir die Dynamik des Prozesses an durch den Markt bewertete Produkte anpassen. Wir verändern die Parameter des Prozesses so, dass sie die realen Preise erklären und das Bewertungsmaß zum realen Maß äquivalent ist. Unter all den Maßen, die diese Voraussetzungen erfüllen, wählen wir eines, das „nahe“ am realen Maß liegt. Da uns keine Stundenoptionen oder andere kurzfristige Spotkontrakte zur Verfügung stehen, müssen wir auf den Futuremarkt ausweichen, um die risikoneutralen Parameter zu fitten.

An der EEX ist der Kontrakt mit der kürzesten Lieferperiode der Monatsfuture, so dass mit diesen Kontrakten ein tages- oder wochenabhängiger MPR kaum anzupassen sein wird. Des Weiteren ist es fraglich, ob das Sprungrisiko auf Grund der starken Glättung, die aus der großen Lieferperiode resultiert, vernünftig abzubilden sein wird.

Alternativ steht uns auch eine hourly Forward Curve zur Verfügung². Diese bildet, von einem festen Zeitpunkt aus, die (fiktiven) Forwardpreise für jede Stunde in einem zukünftigen Zeitraum ab. Diese Preise werden unter Ausnutzung aller möglichen am Markt gehandelter Kontrakte (OTC und EEX) erstellt und bilden somit die Erwartungen des Marktes gut ab.

Es ist zu erwarten, dass Stunden, in denen mit erhöhtem Spikerisiko zu rechnen ist, also Mittagsstunden und u.U. frühe Abendstunden, einen Forwardpreis aufweisen werden, der wesentlich über dem erwarteten Spotpreis liegt. Analog ist die Argumentation für Off-Peak-Stunden.

Zunächst ist jedoch zu klären, auf welche Faktoren unseres Prozesses der Maßwechsel einen Einfluss hat. Für den Prozess Y beschränken wir uns auf die Variation der Sprungintensität η und lassen die Sprunghöhe μ_J konstant. Dieses Vorgehen stellt durchaus eine gewisse Beliebigkeit dar und ist damit zu begründen, dass i.d.R. das Produkt beider Faktoren relevant ist (vgl.[25]). Für den Diffusionsprozess geht der Maßwechsel, wie üblich,

²Eine hFC ist ein gängiges Werkzeug für die Bewertung von Stromprodukten. Sie wurde uns freundlicherweise von einem europäischen Energieunternehmen zur Verfügung gestellt.

über den Marktpreis des Risikos ein. Insgesamt folgt:

$$dX_t = \alpha(\theta(t) - X_t)dt + \sigma dW_t^*, \quad dY_t = -\beta Y_t + J_t dN_t^*,$$

wobei N^* ein Poissonprozess mit Intensität η^* , W^* ein Wienerprozess unter \mathbb{Q} und $\theta(t) = -\lambda(t)\sigma/\alpha$ ist. Der Faktor $\lambda(t)$ wird als Marktpreis des Risikos bezeichnet. Durch diese Wahl der Parameter wird nur der Preis des Diffusionsrisikos als variabel über die Zeit angesehen, während das Sprungrisiko konstant bepreist wird. Alternativ wäre es auch vorstellbar, die Sprungintensität variabel zu definieren und das Diffusionsrisiko als konstant anzusehen. Obige Darstellung lässt sich mit Satz 4.1 auf die gewohnte Darstellung von X bringen, wobei $f(t)$ nun als risikoneutrale Saison zu interpretieren ist. Wir werden mehrere Varianten für die Funktion $\lambda(t)$ untersuchen. Zunächst soll nun der Forward im Kluge-Modell bewertet werden.

6.2.1 Forwardbewertung

Zunächst leiten wir die momenterzeugende Funktion für den Spikeprozess her.

Satz 6.1. [22, Lemma 3.4.2] Seien $\{J_1, J_2, \dots\}$ unabhängig und identisch verteilte Zufallsvariablen mit momenterzeugender Funktion $\Phi_J(\theta) = \mathbb{E}[e^{\theta J}]$, die auf einer Teilmenge $\Theta \subset \mathbb{C}$, die die Null enthält, definiert ist. Seien weiter $\{\tau_1, \tau_2, \dots\}$ die zufälligen Sprungzeitpunkte eines Poissonprozesses N_t mit Intensität η , dann hat der Spikeprozess mit Anfangsbedingung $Y_0 = 0$ und

$$Y_t = \sum_{i=1}^{N_t} e^{-\beta(t-\tau_i)} J_i$$

folgende momenterzeugende Funktion

$$\Phi_Y(\theta, t) = \exp\left(\eta \int_0^t \Phi_J(\theta e^{-\beta s}) - 1 ds\right) \quad \forall \theta \in \Theta. \quad (6.4)$$

Beweis. siehe Satz B.3. □

Die Gleichung (6.4) lässt sich weiter vereinfachen, da sich die momenterzeugende Funktion der Exponentialverteilung berechnen lässt:

$$\Phi_J(x) = \frac{1}{1 - x\mu_J}, \quad x\mu_J < 1.$$

Setzen wir dies in (6.4) ein, so erhalten wir

$$\begin{aligned} \Phi_Y(\theta, t) &= \exp\left(\eta \int_0^t \frac{1}{1 - \theta e^{-\beta s} \mu_J} - 1 ds\right) \\ &= \exp\left(\frac{\eta}{\beta} [\ln(1 - \theta e^{-\beta s} \mu_J)]_0^t\right) = \left(\frac{1 - \theta e^{-\beta t} \mu_J}{1 - \theta \mu_J}\right)^{\frac{\eta}{\beta}}. \end{aligned}$$

Kommen wir nun zu der momenterzeugenden Funktion des Diffusionsprozesses. Dafür brauchen wir einen Hilfssatz, der uns etwas über $\mathbb{E}[e^{\theta X}]$ aussagt, wobei $X \sim N(\mu, \sigma^2)$.

Satz 6.2. Sei $X \sim N(\mu, \sigma^2)$, dann gilt $\mathbb{E}[e^{\theta X}] = \exp(\theta\mu + \frac{1}{2}\theta^2\sigma^2)$.

Beweis. siehe Satz B.4. □

Somit folgt aus (6.1) und dem vorangegangenen Satz

$$\Phi_X(\theta, t) = \mathbb{E}e^{\theta X_t} = \exp(\theta x_0 e^{-\alpha t} + \theta^2 \frac{\sigma^2}{4\alpha} (1 - e^{-2\alpha t})).$$

Nun können wir die Formel für den Forward angeben.

Sei zunächst $P_t = \ln(S_t) = f(t) + X_t + Y_t$, dann folgt

$$\begin{aligned} \mathbb{E}[e^{\theta P_t}] &= \exp\left(\theta f(t) + \theta X_0 e^{-\alpha t} + \theta^2 \frac{\sigma^2}{4\alpha} (1 - e^{-2\alpha t}) + \theta Y_0 e^{-\beta t} \right. \\ &\quad \left. + \eta \int_0^t \Phi_J(\theta e^{-\beta s}) - 1 ds\right). \end{aligned}$$

Dies folgt aus der Unabhängigkeit von X und Y . Der Term $\theta Y_0 e^{-\beta t}$ ergibt sich daraus, dass wir uns bei obiger Rechnung auf $Y_0 = 0$ beschränkt haben.

Der Forwardpreis ist nun folgendermaßen gegeben:

$$\begin{aligned} F_t^T = \mathbb{E}[S_T | X_t, Y_t] &= \exp\left(f(T) + X_t e^{-\alpha(T-t)} + \frac{\sigma^2}{4\alpha} (1 - e^{-2\alpha(T-t)}) \right. \\ &\quad \left. + Y_t e^{-\beta(T-t)} + \eta \int_0^{T-t} \Phi_J(e^{-\beta s}) - 1 ds\right). \end{aligned}$$

Vereinfacht gilt

$$\begin{aligned} F_t^T &= \exp\left(f(T) + X_t e^{-\alpha(T-t)} + \frac{\sigma^2}{4\alpha} (1 - e^{-2\alpha(T-t)}) + Y_t e^{-\beta(T-t)}\right) \\ &\quad \times \left(\frac{1 - e^{-\beta(T-t)} \mu_J}{1 - \mu_J}\right)^{\frac{\eta}{\beta}} \end{aligned}$$

für $\mu_J < 1$.

6.2.2 Umsetzung der hourly Forward Curve

Uns wurde eine hFC für das Jahr 2007 zur Verfügung gestellt. Diese bildet den erwarteten Strompreis unter einem Bewertungsmaß mit bekannten Preisen zu einem vorherigen Zeitpunkt ab. Der Preis zu diesem Zeitpunkt geht also als Bedingung in den Erwartungswert ein. Zur Erstellung der Forward Curve wird eine Vielzahl von Preisinformationen des Marktes benutzt, diese werden auf die einzelnen Stunden umgerechnet. Daher sollte die Kurve ein gutes Abbild „des“ Bewertungsmaßes des Marktes darstellen. Wie in

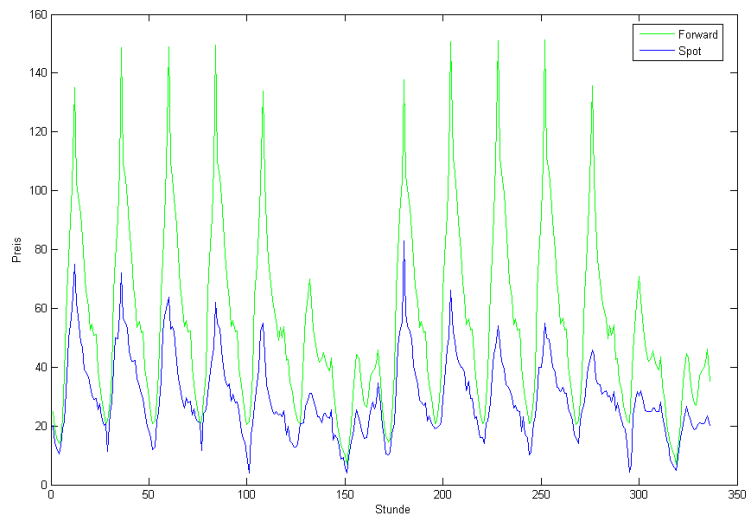
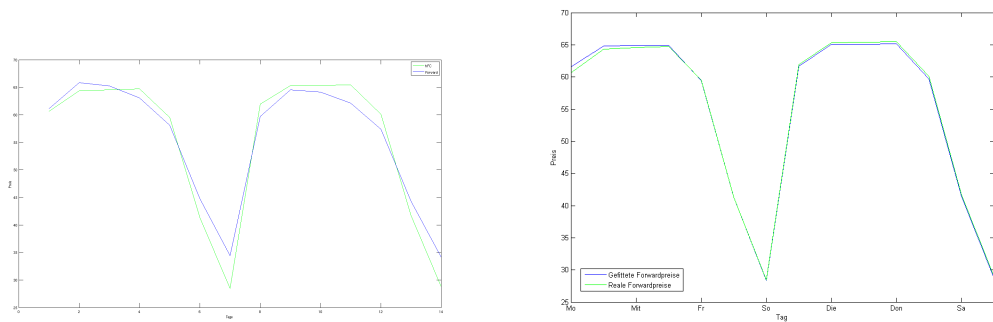


Abbildung 6.2: Vergleich der Forwardkurve mit den tatsächlichen Preisrealisationen



(a) Konstanter MPR $\lambda = -6.76, \eta^* = 7.92$

(b) Variabler MPR

Abbildung 6.3: Kalibrierte Tagesforwardpreise

Abbildung 6.2 zu erkennen liegen etwa um 12 Uhr an Wochentagen die Forwardpreise deutlich über den tatsächlichen Preisrealisationen, während nachts kaum ein Unterschied zu erkennen ist. Ähnlich ist dies am Wochenende. Ein direkter Schluß auf die Höhe der Risikoprämien ist nicht möglich, da hierfür die erwarteten Spotpreise mit den Forwardpreisen verglichen werden müssten.

Das Einbeziehen eines Marktpreises des Risikos kann auch als Korrektur der saisonalen Komponente interpretiert werden. Wir kalibrieren im Folgenden das Modell risikoneutral an der hFC vom 21.05.2007 bis zum 3.06.2007. Eine solche Rechnung ist in der Praxis durchaus realistisch, da eine hFC in der Regel für einen Zeitraum von 1-2 Wochen im Voraus erstellt wird und den Händlern zur Bewertung von Produkten dient. Es müssten für die Bewertung von Optionen zwei Anpassungen durchgeführt werden. Einerseits müsste die deterministische Komponente auf den konkreten Bewertungszeitraum angepasst werden. Bei einer solchen Anpassung würden beispielsweise erwartete Kraftwerksrevisionen berücksichtigt. Andererseits muss der Prozess, wie oben beschrieben, risikoneutral ka-

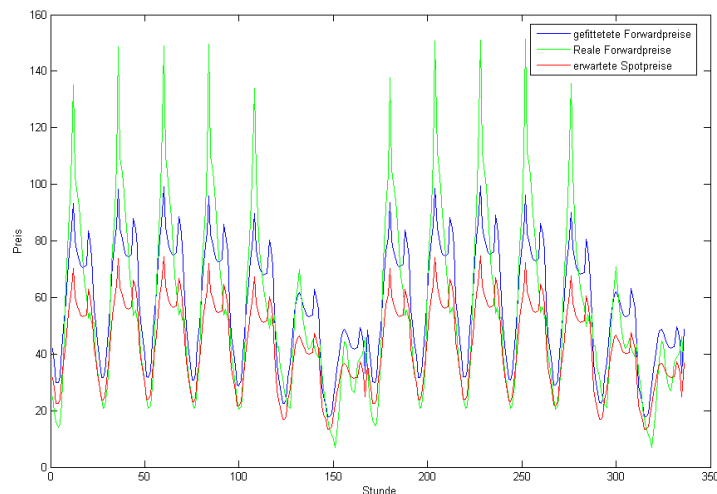
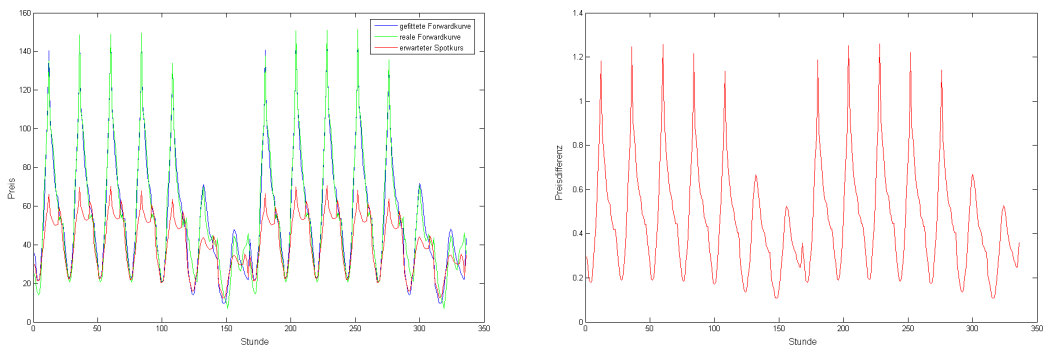


Abbildung 6.4: Kalibrierte Stundenforwardpreise bei einem konstanten MPR $\lambda = -0.94$, $\eta^* = 0.08$

libriert werden. Für die erste Anpassung wäre ein Prognosemodell für die realen Preise notwendig. Da uns ein solches Modell nicht zur Verfügung steht, nutzen wir für beide Anpassungen die hFC.

Zunächst wurde der MPR als konstant angenommen. Somit sind zwei freie Parameter zu bestimmen, einerseits λ und andererseits η^* . Diese wurden mit Hilfe eines least-squares Fit angepasst. Wie aus Abbildung 6.4 zu erkennen ist, wird zwar zu einem Teil das Risiko eingepreist, d.h. die gefitteten Preise bilden die Forwardpreise besser ab als die erwarteten Spotpreise, jedoch kann von einem guten Fit nicht gesprochen werden. Ebenfalls deutlich zu erkennen sind die unterschiedlichen Risikoprämien zu den verschiedenen Stunden, die sich im variierenden Abstand zwischen erwarteten Spotkursen und Forwardpreisen äußern. Die Begründung für die unterschiedlichen Risikoprämien zu verschiedenen Tageszeiten liegt in den unterschiedlichen Risiken. So schwanken Preise um 3 Uhr nachts wesentlich weniger als Preise um die Mittagszeit, daher sind die Marktteilnehmer bereit einen höheren Preis zu bezahlen um sich diesem Risiko nicht auszusetzen. In Abbildung 6.3 (a) ist der Fit für Tagesfutures desselben Zeitraumes dargestellt. Hier kann durch einen konstanten MPR ein besseres Ergebnis erzielt werden. Danach wurde der Fit mit einem variablen MPR durchgeführt. Dabei wurde der MPR als Teil der saisonalen Komponente angesehen und die Funktionen f_{Tag} und f_{Jahr} wurden entsprechend angepasst. Für die Funktion f_{Jahr} wurde nur der konstante Anteil angepasst, während alle Parameter von f_{Tag} kalibriert wurden. Wie zu erwarten war, sind hier die Ergebnisse wesentlich besser. In Abbildung 6.5 (a) ist die angepasste Forwardkurve dargestellt. Der MAE beträgt 0.20 Euro/h und MSE 1.62 Euro/h. Für die Tagesforwards wurde die Funktion f_{Woche} angepasst. Der Fit ist in Abbildung 6.3 (b) dargestellt. Der MSE beträgt 0.15 Euro/Tag und der MAE 0.34 Euro/Tag. Als nächstes ist zu klären, wie groß der Unterschied zwischen



(a) Vergleich mit erwarteten Spotkursen

(b) Preisdifferenz aus der Einbeziehung von Spikes

Abbildung 6.5: Kalibrierte Stundenforwardpreise bei variablem MPR

dem gewöhnlichen Diffusionsmodell und dem Kluge-Modell ist. In Abbildung 6.5 (b) ist die Differenz aus Forwardpreisen im Kluge-Modell und Preisen mit den selben Parametern im Lucia-Schwartz Modell dargestellt. Wie aus der Abbildung zu erkennen ist, führt das Einbeziehen der Spikes zu einer Erhöhung der Forwardpreise um bis zu 1.30 Euro. Die angepassten Parameter für das Kluge-Modell sind in Tabelle 6.4 dargestellt. Nun ha-

Tabelle 6.4: Risikoneutrale Parameter

Par. \ i	1	2	3	4	5	6	7	8	9
t_i	-0.16	-0.59	-0.11	-0.312	-0.24	-0.07	0.344	-0.43	0.39
η^*	0.11	-	-	-	-	-	-	-	-
\dot{j}_i	-	-	-	-	-	3.14	-	-	-

ben wir den Kluge-Prozess risikoneutral kalibriert und können uns der Optionsbewertung widmen. Anzumerken ist, dass eine interessante Erweiterung das Einbeziehen einer zeitabhängigen Sprungintensität darstellen würde. Jedoch sollte auch berücksichtigt werden, dass mit steigender Anzahl an zu bestimmenden Parametern das Trennen der Risikoquellen schwieriger wird.

Kapitel 7

Auswertung des Kluge-Modells

7.1 Diskretisierung

Zunächst ist zu sagen, dass für die numerische Umsetzung die Darstellung der Spotpreisdynamik in Form von (4.1)-(4.3) wesentlich besser geeignet ist als (4.4)-(4.6). Zwar ist diese für die analytische Untersuchung der PIDE sehr nützlich, jedoch hat die erste Darstellung wesentliche Vorteile bei der Umsetzung mit Hilfe der Finite Differenzen Methode. Die großen Änderungen in $f(t)$ führen dazu, dass die Ortsvariablen auf einem großen Gebiet diskretisiert werden müssten, um eine vernünftige Diskretisierung zu gewährleisten. Dies führt zu einem hohen Rechenaufwand und bei äquidistanter Diskretisierung zu einer großen Anzahl an Datenpunkten für geringe Stände des Prozesses. Außerdem ist es nicht besonders praxisrelevant, da es wesentlich interessanter wäre, genaue Informationen für Preise in der Nähe der real erwarteten Saison zu erhalten. Daher bedienen wir uns einer alternativen Herleitung der PIDE analog zu [21]. Wir sehen den Put als Funktion der Zeit und der beiden Prozesse X_t und Y_t an, wobei

$$dX_t = -\alpha X_t dt + \sigma dW_t \quad \text{und} \quad dY_t = -\beta Y_t dt + J_t dN_t$$

gilt. Die Saisonalität spielt lediglich bei dem Auszahlungsprofil eine Rolle:

$$V(0, x, y) = \max(K - \exp(f(T) + x + y), 0). \quad (7.1)$$

Der Optionspreis für einen bestimmten Preis S_τ bei einer Restlaufzeit von τ wäre, falls wir annehmen, dass zum aktuellen Zeitpunkt der Jump-Anteil Null ist, $V(\tau, \log(S_\tau) - f(T - \tau), 0)$. Ohne diese Annahme ist es nicht klar, zu welchen x - und y -Koordinaten ein beobachteter Preis S_τ gehört. Eine solche Annahme ist jedoch nicht besonders restriktiv, da es beispielsweise auch bei Banken üblich ist, bei großen Preissprüngen Produkte vorübergehend auf Grund der fehlenden Bewertbarkeit vom Handel auszusetzen.

Analog zur Herleitung im Grundlagenteil lässt sich auch hier eine PIDE für den Putpreis

angeben:

$$\frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} - \alpha x \frac{\partial V}{\partial x} - \beta y \frac{\partial V}{\partial y} - rV + \eta \int_{-\infty}^{\infty} ((V(x, y + z, t) - V(x, y, t))) f_J(z) dz = \frac{\partial V}{\partial \tau} \quad (7.2)$$

auf $\Omega = [0, T] \times [x_{\min}, x_{\max}] \times [0, \infty)$. Weiterhin müssen Randbedingungen in x -Richtung vorgegeben werden. Diese sind bei Problemen der Optionsbewertung nicht ad hoc klar. Wir nutzen mehrere Ansätze, indem wir uns einerseits Neumann- und andererseits Dirichletränder vorgeben. Eine mögliche Wahl für die Neumannränder ist

$$\frac{\partial V}{\partial x}(x_{\min}, y, \tau) = \frac{\partial V}{\partial x}(x_{\max}, y, \tau) = 0 \quad \forall \tau, y \in [0, T] \times [0, \infty) \quad (7.3)$$

und für Dirichletränder geben wir uns die entsprechende Lösung aus dem LS-Modell vor (vgl. Abschnitt 9.1). Es wird die Umsetzung für Neumannränder erläutert; für Dirichletränder verläuft diese analog. Es muss also die PIDE (7.2) mit (7.1) als Endbedingung und (7.3) als Randbedingung auf Ω gelöst werden. Wir werden dies mit Hilfe der Finite Differenzen Methode tun.

Dazu wird zunächst der Definitionsbereich der Lösung $V(x, y, \tau)$ beschränkt. Gesucht ist nur eine Lösung auf $\tilde{\Omega} = [0, T] \times [x_{\min}, x_{\max}] \times [0, y_{\max}]$, wobei y_{\max} hinreichend groß gewählt werden sollte. Auf Grund der Nichtlokalität des Integralterms in (7.2) muss $V(x, y, \tau)$ für $y > y_{\max}$ vorgegeben werden. Zuletzt muss auch die Sprunggröße nach oben beschränkt werden. Dies bedeutet, dass im Integralterm nicht mit der Dichtefunktion einer exponentialverteilten Zufallsvariable multipliziert wird, sondern die Dichte nur auf einem endlichen Intervall lebt. Als Definitionsbereich für die abgeschnittene Dichtefunktion \tilde{f}_J wählen wir $[0, I_{\max}]$, mit $\int_0^{I_{\max}} f_J(x) dx \approx 1$.

Nun können die Orts- und Zeitvariablen diskretisiert werden. Die Diskretisierung wird äquidistant vorgenommen:

$$x_i = x_{\min} + i\Delta x \quad i = 0, \dots, N_x, \quad \Delta x = (x_{\max} - x_{\min})/N_x \quad (7.4)$$

$$y_j = j\Delta y \quad j = 0, \dots, N_y, \quad \Delta y = (y_{\max})/N_y \quad (7.5)$$

$$t_k = k\Delta t \quad k = 0, \dots, L, \quad \Delta t = T/L. \quad (7.6)$$

Dabei ist $v_{i,j}^k$ die Lösung in dem Gitterpunkt (x_i, y_j, t_k) . Für $j \notin [0, N_y]$ muss diese entsprechend fortgesetzt werden. Nun können wir uns der Diskretisierung der unterschiedlichen Integral- und Differentialterme widmen.

Die Ortsableitungen erster Ordnung werden mit Hilfe des Upwind-Verfahrens implizit diskretisiert. Die implizite Upwind Diskretisierung führt zu einem unbedingt stabilen Verfahren und entspricht dem Informationsfluss der Differentialgleichung. Upwind bedeutet ‘‘gegen den Wind‘‘, dies soll andeuten, dass in die entgegengesetzte Richtung zum Fluss der Charakteristiken diskretisiert wird. Bei der Ableitung erster Ordnung in x -Richtung

muss zusätzlich beachtet werden, dass sich die Richtung der Charakteristiken für verschiedene x -Werte ändert. Wann dies der Fall ist, hängt nicht vom Zeitschritt ab und kann einmal zu Beginn der Rechnung bestimmt werden. Somit diskretisieren wir die Ortsableitung in x -Richtung folgendermaßen:

$$\alpha x_i \frac{\partial v}{\partial x}(x_i, y_j, t_k) \approx \min(\alpha x_i, 0) \frac{v_{i+1,j}^k - v_{i,j}^k}{\Delta x} + \max(\alpha x_i, 0) \frac{v_{i,j}^k - v_{i-1,j}^k}{\Delta x}. \quad (7.7)$$

Anzumerken ist, dass das Upwind-Verfahren ein einseitiges Diskretisierungsverfahren darstellt, da höchstens einer der beiden Summanden aus (7.7) von Null verschieden ist. Die Ableitung erster Ordnung in y -Richtung wird analog diskretisiert:

$$\beta y_j \frac{\partial v}{\partial y}(x_i, y_j, t_k) \approx \beta y_j \frac{v_{i,j}^k - v_{i,j-1}^k}{\Delta y}. \quad (7.8)$$

Hier verwenden wir nur den Rückwärtsdifferenzenquotienten, da $y_j \geq 0$ für alle j gilt. Der Rückwärtsdifferenzenquotient ist einem zentralen Differenzenquotienten vorzuziehen, da β groß ist und ansonsten die Diagonaldominanz der Systemmatrix nicht sichergestellt werden könnte.

Für den Diffusionsterm wurde ein zentraler Differenzenquotient genutzt und die Diskretisierung wurde ebenfalls implizit vorgenommen:

$$\frac{\sigma^2}{2} \frac{\partial^2 v}{\partial x^2}(x_i, y_j, t_k) \approx \frac{\sigma^2}{2} \frac{v_{i+1,j}^k - 2v_{i,j}^k + v_{i-1,j}^k}{(\Delta x)^2}.$$

Der Integralterm wurde durch eine endliche Summe approximiert, wobei die Schrittweite für die Diskretisierung des Integrals analog zur Schrittweite in y -Richtung gewählt wurde. Im Gegensatz zu den Differentialtermen, wurde der Integralterm explizit diskretisiert, um eine dünnbesetzte Systemmatrix zu erhalten. Der Integralterm wird mit Hilfe der Trapezregel folgendermaßen approximiert:

$$\int_0^\infty (V(x_i, y_j + z, t_k) - V(x_i, y_j, t_k)) f_J(z) dz \approx \sum_{n=0}^P v_{i,j+n}^{t_{k-1}} \nu_n, \quad (7.9)$$

wobei $P = I_{max}/\Delta y$ ist. Die $v_{i,j+n}^k$ werden für $j+n > N_y$ gleich Null gesetzt. Dies entspricht der Vorstellung, dass bei großen Spikes die Produkte nicht bewertet und gehandelt werden können. Weiterhin ist ν_n das Integral über die Dichte der Sprungverteilung auf dem entsprechenden Intervall. Wir wählen

$$\nu_n = \eta \int_{\Delta y(n-\frac{1}{2})}^{\Delta y(n+\frac{1}{2})} f_J(z) dz \quad \forall n > 0 \quad (7.10)$$

und $\nu_0 = \eta \int_0^{\frac{1}{2}} f_J(z) dz$.

Nun können wir das Differenzenschema angeben. Wir verwenden folgende Konventionen, um die Schreibweise zu vereinfachen:

$$\begin{aligned} (D^v u)_{i,j} &= \frac{\sigma^2}{2} \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} - \alpha x_i \frac{u_{i+1,j} - u_{i,j}}{\Delta x} - \beta j \frac{u_{i,j} - u_{i,j-1}}{\Delta y} \\ (D^r u)_{i,j} &= \frac{\sigma^2}{2} \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} - \alpha x_i \frac{u_{i,j} - u_{i-1,j}}{\Delta x} - \beta j \frac{u_{i,j} - u_{i,j-1}}{\Delta y} \\ (Ju)_{i,j} &= \sum_{n=0}^P \nu_n u_{i,j+n} - \gamma u_{i,j}, \end{aligned}$$

mit $\gamma = \sum_{n=0}^P \nu_n$. Für homogene Neumannränder in x und eine Fortsetzung mit Null in y lautet das Differenzenschema folgermaßen:

$$u_{i,j}^0 = h(x_i, y_j) \quad \forall (i, j) \in [0, \dots, N_x] \times [0, \dots, N_y] \quad (7.11)$$

$$\frac{u_{i,j}^{k+1} - u_{i,j}^k}{\Delta t} = \mathbf{1}_{x_i \leq 0} (D^v u^{k+1})_{i,j} + \mathbf{1}_{x_i > 0} (D^r u^{k+1})_{i,j} + (Ju^k)_{i,j} \quad (7.12)$$

$$\forall (i, j, k) \in [1, \dots, N_x - 1] \times [0, \dots, N_y] \times [0, \dots, L - 1].$$

An den Rändern in x ergibt sich auf Grund der Neumannbedingung folgende Gleichungen.

Randwerte in x :

$$\frac{u^{k+1} - u^k}{\Delta t} = \frac{\sigma^2}{2} \frac{u_{1,j}^{k+1} - u_{0,j}^{k+1}}{(\Delta x)^2} - \alpha x_{\min} \frac{u_{1,j}^{k+1} - u_{0,j}^{k+1}}{\Delta x} - \beta j \Delta y \frac{u_{0,j}^{k+1} - u_{0,j-1}^{k+1}}{\Delta y} + (Ju^k)_{0,j} \quad (7.13)$$

$$\forall (j, k) \in [0, \dots, N_y] \times [0, \dots, L - 1],$$

$$\begin{aligned} \frac{u^{k+1} - u^k}{\Delta t} &= \frac{\sigma^2}{2} \frac{u_{N_x-1,j}^{k+1} - u_{N_x,j}^{k+1}}{(\Delta x)^2} - \alpha x_{\max} \frac{u_{N_x,j}^{k+1} - u_{N_x-1,j}^{k+1}}{\Delta x} \\ &- \beta j \Delta y \frac{u_{N_x,j}^{k+1} - u_{N_x,j-1}^{k+1}}{\Delta y} + (Ju^k)_{N_x,j} \quad \forall (j, k) \in [0, \dots, N_y] \times [0, \dots, L - 1], \end{aligned} \quad (7.14)$$

Für den y -Rand ergibt sich auf Grund der trivialen Fortsetzung folgende Bedingung.

Randwerte in y :

$$u_{i,j}^k = 0 \quad \forall j > N_y. \quad (7.15)$$

7.2 Fehleranalyse

Im Folgenden sollen die Fehler, die bei der numerischen Lösung der Differentialgleichung auftreten, analysiert werden. Die erste Vereinfachung, die vorgenommen wird, ist die Beschränkung des Bereichs auf dem die Differentialgleichung erfüllt sein muss. Wir lösen nicht auf Ω , sondern auf $\tilde{\Omega}$ und setzen die Lösung für größere y -Werte fort. Eine Fortsetzung ist auf Grund der Nicht-Lokalität des Integralterms notwendig. Der Fehler,

der durch das Beschränken des Definitionsbereichs entsteht, soll nun abgeschätzt werden. Wir vergleichen eine Lösung v des Problems (7.1)-(7.2) mit der abgeschnittenen Lösung v_A des folgenden Problems:

$$\frac{\partial V}{\partial \tau} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} - \alpha x \frac{\partial V}{\partial x} - \beta y \frac{\partial V}{\partial y} + \eta \int_{-\infty}^{\infty} ((V(x, y + z, \tau) - V(x, y, \tau))) f_J(z) dz = rV \quad (7.16)$$

auf $\tilde{\Omega} = [0, T] \times (-\infty, \infty) \times [0, y_{\max}]$ und $v(x, y, t) = 0$ für $y > y_{\max}$.

Weiterhin gilt:

$$V(0, x, y) = \max(K - \exp(f(T) + x + y), 0) = h(x, y). \quad (7.17)$$

Satz 7.1. Sei v_A die Lösung von (7.16)-(7.17) und v die Lösung von (7.1)-(7.2) mit unbeschränktem Definitionsbereich in x , dann gilt mit $0 < k < \mu_J$ beliebig:

$$|v(x, y, t) - v_A(x, y, t)| \leq C \|h\|_{\infty} e^{-k(y_{\max} - y)}, \quad \forall (x, y) \in (-\infty, \infty) \times [0, y_{\max}].$$

Beweis. Der Beweis benutzt die wahrscheinlichkeitstheoretische Darstellung des Optionspreises, wir folgen dabei der Beweistechnik in [15, Proposition 5]. Weiterhin benutzen wir, dass sich für alle Ereignisse aus unserem Wahrscheinlichkeitsraum der Stand des Prozesses Y nach oben gegen den Stand eines gewichteten Poissonprozesses \tilde{Y} ohne Mean Reversion abschätzen lässt. Der Optionspreis lässt sich darstellen als:

$$\begin{aligned} v(x, y, t) &= e^{-r\tau} \mathbb{E}[h(X, Y, T) | X_t = x, Y_t = y] \\ v_A(x, y, t) &= e^{-r\tau} \mathbb{E}[h(X, Y, T) \mathbf{1}_{M_t^y < y_{\max}} | X_t = x, Y_t = y], \end{aligned}$$

wobei $M_t^y = \sup_{\tau \in [t, T]} Y_{\tau}$ und $\tilde{M}_t^y = \sup_{\tau \in [t, T]} \tilde{Y}_{\tau}$ ist. Nun können wir folgendermaßen abschätzen:

$$\begin{aligned} |v(x, y, t) - v_A(x, y, t)| &= e^{-r\tau} \left| \mathbb{E}[h(x, y, t) \mathbf{1}_{M_t^y \geq y_{\max}}] \right| \leq e^{-r\tau} \left| \mathbb{E}[h(x, y, t) \mathbf{1}_{\tilde{M}_t^y \geq y_{\max}}] \right| \\ &\leq e^{-r\tau} \|h\|_{\infty} Pr(\tilde{M}_t^y \geq y_{\max}). \end{aligned} \quad (7.18)$$

Die Levy-Dichte des gewichteten Poissonprozesses ist $\nu(x) = \eta f_J(x)$ und somit gilt $\int_1^{\infty} e^{kx} f_J(x) dx < \infty$ für $0 < k < \mu_J$. Daraus folgt $\mathbb{E}[e^{k\tilde{Y}_t}] < \infty$ und es lässt sich mit [30, Theorem 25.18] zeigen, dass $\mathbb{E}[e^{k\tilde{M}_t^0}] < \infty$ gilt. Nun kann die Markov-Ungleichung angewendet werden und es folgt:

$$Pr(\tilde{M}_{\tau}^0 \geq y_{\max}) = Pr(e^{k\tilde{M}_{\tau}^0} \geq e^{ky_{\max}}) \leq \mathbb{E}[e^{k\tilde{M}_{\tau}^0}] e^{-ky_{\max}}.$$

Nun fehlt uns noch der Übergang von \tilde{M}_{τ}^y zu \tilde{M}_{τ}^0 . Es gilt, da $y \geq 0$ und \tilde{Y} ein Levy-

Prozess ist:

$$\widetilde{M}_\tau^y = \sup \left| \widetilde{Y}_\tau + y \right| = \sup \left| \widetilde{Y}_\tau \right| + y.$$

Daraus folgt:

$$Pr(\widetilde{M}_\tau^y \geq y_{max}) = Pr(\widetilde{M}_\tau^0 \geq y_{max} - y) \leq \mathbb{E}[e^{k\widetilde{M}_\tau^0}]e^{-k(y_{max}-y)}.$$

Die Aussage des Satzes ergibt sich somit aus der letzten Abschätzung und (7.18). \square

Da wir uns in der Regel für den Optionswert an der Stelle $y = 0$ interessieren, ist dies ein zufriedenstellendes Resultat. Das Lösen der Differentialgleichung auf einem beschränkten Bereich in x ist auf Grund des lokalen Charakters der Differentialgleichung in diesen Variablen nicht problematisch. Nun soll separat der Fehler durch das Abschneiden des Integralterms betrachtet werden. Nehmen wir nun an, dass unsere Lösung $v \in L^\infty(\mathbb{R}_+ \times \mathbb{R} \times [0, T])$ mit einer Schranke U versehen ist, so können wir folgendermaßen abschätzen:

$$\begin{aligned} & \left| \int_0^\infty (v(x, y+z, \tau) - v(x, y, \tau)) f_J(z) dz - \int_0^{I_{max}} (v(x, y+z, \tau) - v(x, y, \tau)) f_J(z) dz \right| \\ &= \left| \int_{I_{max}}^\infty (v(x, y+z, \tau) - v(x, y, \tau)) f_J(z) dz \right| \\ &\leq 2U \int_0^\infty f_J(z) dz = 2Ue^{-\eta I_{max}}. \end{aligned}$$

Somit fällt der Fehler exponentiell mit der maximalen Sprunghöhe. Falls die Lösung v in y eine globale Lipschitz-Bedingung erfüllt, kann sogar eine stärkere Abschätzung mit Hilfe der Lipschitzkonstanten gezeigt werden.

Nun soll die Konvergenz, Monotonie, Konsistenz und Stabilität des Differenzenverfahrens untersucht werden. Wir zeigen zunächst die Konsistenz des Differenzschemas für die Lösung von:

$$\begin{aligned} \frac{\partial V}{\partial \tau} &= \underbrace{\frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} - \alpha x \frac{\partial V}{\partial x} - \beta y \frac{\partial V}{\partial y} - rV}_{=DV} \\ &\quad + \underbrace{\eta \int_0^{I_{max}} ((V(x, y+z, \tau) - V(x, y, \tau))) f_J(z) dz}_{=JV} \end{aligned} \quad (7.19)$$

auf $\Omega = [0, T] \times [x_{min}, x_{max}] \times [0, y_{max}]$ und $v(x, y, \tau) = 0$ für $y > y_{max}$, und entsprechenden Anfangs- und Randwerten.

Satz 7.2 (Konsistenz). *Es gilt für ein beliebiges $v \in C_0^3(\mathbb{R} \times \mathbb{R}_+ \times [0, T])$ und (x_i, y_j, t_k) beliebig¹:*

$$\left| \frac{v_{i,j}^{k+1} - v_{i,j}^k}{\Delta t} - (D^v v)_{i,j}^{k+1} - (Jv)_{i,j}^k - \left(\frac{\partial v}{\partial t} - (D + J)v \right) (x_i, y_j, t_k) \right| = r_{i,j}^k(\Delta x, \Delta y, \Delta t) \rightarrow 0$$

für $(\Delta x, \Delta y, \Delta t) \rightarrow 0$. Es existiert außerdem ein $c > 0$ mit

$$|r_{i,j}^k(\Delta x, \Delta y, \Delta t)| \leq c(\Delta x + \Delta y + \Delta t).$$

Beweis. Mit Taylorentwicklung folgt:

$$\left| \frac{v_{i,j}^{k+1} - v_{i,j}^k}{\Delta t} - \frac{\partial v}{\partial t}(x_i, y_j, t_k) \right| \leq \frac{\Delta t}{2} \left\| \frac{\partial^2 v}{\partial t^2} \right\|_{\infty}.$$

Der Differentialterm lässt sich folgendermaßen abschätzen:

$$\begin{aligned} |(Dv)_{i,j}^{k+1} - (Dv)(x_i, y_j, t_k)| &= |(Dv)_{i,j}^{k+1} - (Dv)(x_i, y_j, t_{k+1}) + \Delta t(Dv)(x_i, y_j, \tilde{t})| \\ &\leq C_1 \Delta t + C_2 \Delta x + C_3 \Delta y. \end{aligned}$$

Schließlich gilt für den Integralterm:

$$\begin{aligned} & |(Jv)_{i,j}^k - (Jv)(x_i, y_j, t_k)| \\ &= \left| \sum_{n=0}^P v_{i,j+n}^k \nu_n - \eta \int_0^{I_{max}} v(x_i, y_j + z, t_k) f_J(z) dz \right| \\ &= \left| \sum_{n=0}^P \eta \int_{\max(n-1/2, 0)\Delta y}^{(n+1/2)\Delta y} (v(x_i, y_j + y_n, t_k) - v(x_i, y_j + z, t_k)) f_J(z) dz \right| \\ &= \left| \sum_{n=0}^P \eta \int_{(\max(n-1/2, 0))\Delta y}^{(n+1/2)\Delta y} (y_n - z) \frac{\partial v}{\partial y}(x_i, y_j + \xi, t_k) f_J(z) dz \right| \\ &\leq \frac{\Delta y}{2} \gamma \left\| \frac{\partial v}{\partial y} \right\|_{\infty} \end{aligned}$$

Aus den drei Abschätzungen folgt die Konsistenz des Verfahrens und die gewünschte Abschätzung. □

Als Nächstes widmen wir uns der Stabilität des Differenzschemas. Hierzu schreiben

¹Wir nehmen im Folgenden $x_{\min} + i\Delta x < 0$ an, somit muss die Ableitung erster Ordnung mit einem Vorwärtsdifferenzenquotienten diskretisiert werden. Für den anderen Fall verläuft die Argumentation analog.

wir das Differenzenschema folgendermaßen um:

$$v_{i+1,j}^{k+1}a\Delta t + v_{i,j}^{k+1}(b+1)\Delta t + v_{i-1,j}^{k+1}c\Delta t + v_{i,j-1}^{k+1}d\Delta t = (1 - \Delta t\gamma)v_{i,j}^k + \Delta t \sum_{n=0}^P \nu_n u_{i,j+n}^k. \quad (7.20)$$

Dabei ist

$$\begin{aligned} a &= -\frac{\sigma^2}{(2\Delta x)^2} + \alpha \frac{(x_{\min} + i\Delta x)}{\Delta x} < 0 \\ b &= \frac{\sigma^2}{(\Delta x)^2} - \alpha \frac{(x_{\min} + i\Delta x)}{\Delta x} + \beta j > 0 \\ c &= -\frac{\sigma^2}{(2\Delta x)^2} < 0 \\ d &= -\beta j \leq 0. \end{aligned}$$

Satz 7.3 (Stabilität). *Das Differenzenverfahren (7.11)-(7.15) ist für $\Delta t \leq 1/\gamma$ stabil.*

Beweis. Die Systemmatrix des Verfahrens ist eine M-Matrix, da alle Nebendiagonaleinträge nicht positiv sind² und das M-Matrizen-Kriterium für $w = (1, \dots, 1)$ erfüllt ist (vgl. [28, Property 1.19]). Das Kriterium besagt, dass eine Matrix A mit nicht positiven Nebendiagonaleinträgen genau dann eine M-Matrix ist, falls ein Vektor w existiert d.d. $Aw > 0$ gilt. Somit ist das aufgestellte LGS eindeutig lösbar. Nun wird die Beschränktheit der Lösung mittels vollständiger Induktion über die Zeitschritte gezeigt. Für $k = 0$ folgt aus $v_{i,j}^0 = h(x_i, y_j)$ die Abschätzung $\|v^0\|_\infty \leq \|h\|_\infty$. Wir nehmen an $\|v^k\|_\infty \leq \|h\|_\infty$ gelte für ein beliebiges $k \in \mathbb{N}$ und es ist die Aussage für $k + 1$ zu zeigen. Es existieren i_0 und j_0 mit $|v_{i_0,j_0}^{k+1}| = \|v^{k+1}\|_\infty$ und $|v_{i,j}^{k+1}| \leq |v_{i_0,j_0}^{k+1}|$ für alle $i, j \in \mathbb{N}$. Daraus folgt

$$\begin{aligned} \|v^{k+1}\|_\infty &= |v_{i_0,j_0}^{k+1}| = |v_{i_0,j_0}^{k+1} a\Delta t + |v_{i_0,j_0}^{k+1}|(b+1)\Delta t + |v_{i_0,j_0}^{k+1}|c\Delta t + |v_{i_0,j_0}^{k+1}|d\Delta t \\ &\leq |v_{i_0+1,j_0}^{k+1}| a\Delta t + |v_{i_0,j_0}^{k+1}|(b+1)\Delta t + |v_{i_0-1,j_0}^{k+1}|c\Delta t + |v_{i_0,j_0-1}^{k+1}|d\Delta t \\ &\stackrel{\Delta\text{-Ung.}}{\leq} |v_{i_0+1,j_0}^{k+1} a\Delta t + v_{i_0,j_0}^{k+1}(b+1)\Delta t + v_{i_0-1,j_0}^{k+1}c\Delta t + v_{i_0,j_0-1}^{k+1}d\Delta t|. \end{aligned}$$

Mit (7.20) folgt daraus:

$$\|v^{k+1}\|_\infty \leq (1 - \Delta t\gamma)\|v^k\|_\infty + \Delta t\gamma\|v^k\|_\infty \leq \|h\|_\infty$$

und somit ist $\|v^{k+1}\|_\infty \leq \|h\|_\infty$ gezeigt. \square

Die Intuition hinter der Stabilitätskonstante ist, dass $\gamma\Delta t$ der erwarteten Anzahl an Sprüngen innerhalb eines Zeitschritts entspricht. Wird nun Δt zu groß gewählt, so erwarten wir mehr als einen Sprung pro Zeitschritt und das Verfahren kann nicht vernünftig sein. Außerdem lässt sich für das Differenzenschema eine Monotonieeigenschaft zeigen.

²An den Rändern in x gilt dies eine analoge Abschätzung.

Satz 7.4 (Monotonie). *Seien u^n und v^n zwei Lösungen des Differenzenschemas (7.11)-(7.15) zu Anfangswerten g und h und es gelte $g(x, y) \geq h(x, y)$ für alle x und y , dann folgt $u^n > v^n$ für alle $n \in \mathbb{N}$.*

Beweis. Wir definieren $w^n = u^n - v^n$. Wir wollen $w^n \geq 0$ für alle $n \in \mathbb{N}$ zeigen. Der Beweis verläuft analog zum Stabilitätsbeweis. Für $n = 0$ ist die Aussage klar. Wir nehmen an, die Aussage gelte für ein beliebiges n und es ist die Aussage für $n + 1$ zu zeigen. Wir nehmen $\inf_{i,j \in \mathbb{Z}} w_i^{n+1} < 0$ an. Wieder muss das Infimum für ein (i_0, j_0) angenommen werden. Somit folgt:

$$\begin{aligned} \inf_{i,j \in \mathbb{Z}} w_{i,j}^{k+1} &= w_{i_0,j_0}^{k+1} = w_{i_0,j_0}^{k+1} a \Delta t + w_{i_0,j_0}^{k+1} (b+1) \Delta t + w_{i_0,j_0}^{k+1} c \Delta t + w_{i_0,j_0}^{k+1} d \Delta t \\ &\geq w_{i_0+1,j_0}^{k+1} a \Delta t + w_{i_0,j_0}^{k+1} (b+1) \Delta t + w_{i_0-1,j_0}^{k+1} c \Delta t + w_{i_0,j_0-1}^{k+1} d \Delta t \\ &= (1 - \Delta t \gamma) w_{i_0,j_0}^k + \Delta t \sum_{n=0}^P \nu_n w_{i_0,j_0+n}^k \geq 0. \end{aligned}$$

Dies widerspricht jedoch unserer Annahme und somit ist $\inf_{i,j \in \mathbb{Z}} w_{i,j}^{n+1} \geq 0$ gezeigt. \square

Aus Konsistenz und Stabilität folgt insgesamt bei hinreichender Glattheit der Lösung die Konvergenz des Verfahrens.

7.3 Umsetzung

Das Programm wurde in Matlab implementiert. Nach der Diskretisierung aller Integral- und Differentialterme erhalten wir in jedem Zeitschritt das folgende LGS:

$$Av^k = (B_1 + B_2)v^{k+1} + b^k,$$

wobei v^k die Approximation der Lösung zum Zeitpunkt t_k ist. Die Vektoren v^k und b^k haben die Größe $N_x \cdot N_y$ und A , B_1 und B_2 sind Matrizen der Größe $(N_x \cdot N_y, N_x \cdot N_y)$. In v^k sind die Lösungswerte blockweise angeordnet, sodass zu jedem y -Wert nacheinander die Lösungen zu allen x -Werten angegeben sind. Analog müssen die Matrizen A , B_1 und B_2 sowie b^k aufgebaut sein. Die Matrix A ist dabei die Summe der Matrizen A_1 , A_2 , A_3 und A_4 , welche die Diskretisierung der einzelnen Differentialterme darstellt. Dabei stellen die Matrizen A_1 und A_2 die Diskretisierung der Ableitungen in x -Richtung dar und sind Tridiagonalmatrizen. Die Matrix A_3 stellt die Ableitung in y -Richtung dar. Sie ist ebenfalls eine Bandmatrix mit nichttrivialer Diagonale und von Null verschiedenen Einträgen auf der N_x -ten unteren Nebendiagonale. Die Matrix A_4 bildet einen Teil der Zeitableitung ab und ist eine reine Diagonalmatrix. Falls wir Neumannränder in x -Richtung annehmen, müssen die erste und letzte Zeile in A_1 entsprechend angepasst werden. Rechnen wir mit Dirichleträndern, so enthält der Vektor b^k die entsprechenden Werte.

Die Matrix B_1 enthält die Diskretisierung des Integralterms. Diese wird wie in (7.9) und (7.10) vorgenommen. Die Matrix B_2 stellt den zweiten Teil der Zeitdiskretisierung dar. Somit ist das zu lösende System aufgestellt. Jedoch muss eine Matrix der Größe $(N_x N_y, N_x N_y)$ invertiert werden. Dies ist bei vernünftiger Diskretisierung numerisch sehr aufwendig. Um den Rechenaufwand zu reduzieren, wurde das in jedem Zeitschritt zu lösende LGS weiter vereinfacht. Da sich die y -Abhängigkeit im impliziten Teil der Diskretisierung auf die Ableitung erster Ordnung beschränkt, die wir mit einem Rückwärtsdifferenzenquotient diskretisieren, können wir das große System in eine Folge kleiner Systeme zerlegen. Dabei wird in jedem Zeitschritt zu jedem y -Wert einzeln und jeweils allen x -Werten das Problem gelöst. In jedem Zeitschritt bestimmen wir zunächst den Vektor $(v_{0,0}^k, \dots, v_{N_x,0}^k)$. Dieser hängt nur von v^{k-1} ab, da die Ableitung in y -Richtung entfällt. Danach bestimmen wir den Vektor $(v_{0,1}^k, \dots, v_{N_x,1}^k)$. Dieser hängt nur von v^{k-1} und dem bereits bestimmten Vektor $(v_{0,0}^k, \dots, v_{N_x,0}^k)$ ab. Für die weiteren Teilvektoren wird analog vorgegangen. Diese Zerlegung ist möglich, da wir für die Diskretisierung der Ableitung in y -Richtung das Upwind-Verfahren nutzen und der Integralterm explizit diskretisiert wird. Die Systemmatrizen in den Teilsystemen sind von dem Zeitschritt unabhängig und können vor Beginn der Rechnung invertiert und gespeichert werden. Während der Bewertung beschränkt sich die Rechnung auf Matrixmultiplikationen und der Aufwand ist linear in der Zeit. Die reduzierten N_y Systemmatrizen A^k haben die Größe (N_x, N_x) und sind die Summe aus Matrizen, die eine zu A_1 , A_2 und A_4 identische Struktur aufweisen, sowie einer Diagonalmatrix \tilde{A}_3^k mit Diagonaleintrag $-\beta(j-1)$. Dies ergibt sich sofort aus der Umstellung des obigen Gleichungssystems. Die rechte Seite wird ähnlich vereinfacht.

7.4 Numerische Analyse

Nachdem das Kluge-Modell kalibriert wurde und die PIDE für die Bewertung eines europäischen Puts diskretisiert wurde, soll nun das Modell numerisch ausgewertet werden. Im Folgenden betrachten wir Optionswerte an der Stelle $y = 0$. Eine Annahme über den Stand der einzelnen Prozesse ist notwendig, da die Variablen X_t und Y_t , im Gegensatz zu S_t , nicht beobachtbar sind. Ökonomisch kann diese Annahme damit begründet werden, dass bei dem Auftreten von Spikes, Produkte vorübergehend nicht bewertet werden. Dies ist im Aktienumfeld durchaus üblich. In Abbildung 7.1 (a) ist die Preisoberfläche in $\exp(x)$ und $\exp(y)$ dargestellt. Deutlich ist sowohl der typische Putpreisverlauf als auch die Mean Reversion Eigenschaft zu erkennen. In Abbildung 7.1 (b) sind Puts mit unterschiedlichen Restlaufzeiten abgebildet, wobei der Bewertungszeitpunkt Montag der 21.05.2007 ist. Mit steigender Restlaufzeit nimmt die Sensitivität der Optionspreise bezüglich des aktuellen Standes des Prozesses deutlich ab, dies ist der Mean Reversion Effekt. Als nächstes wurde die Sensitivität des Optionspreises bezüglich des Parameters η^*

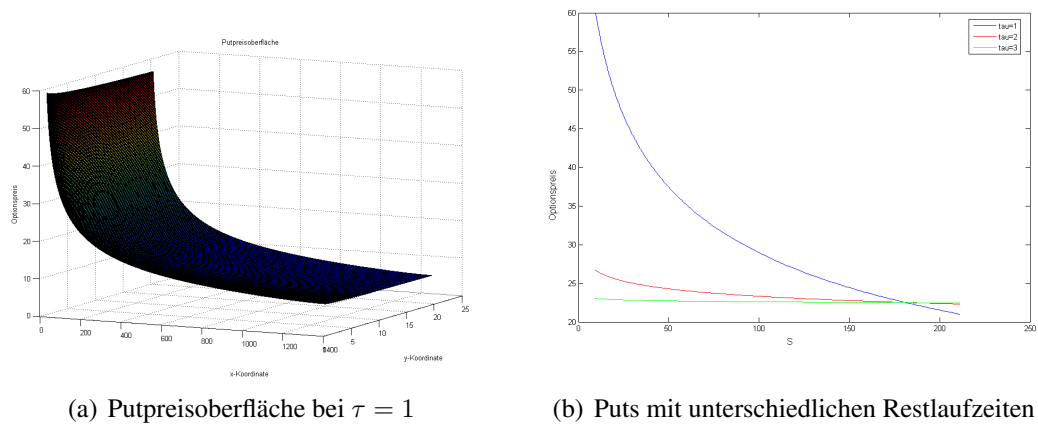


Abbildung 7.1: Putpreise im Kluge-Modell

betrachtet. Wie aus Abbildung 7.2 (a) zu erkennen ist, beeinflusst der Parameter η^* die Optionspreise in großem Maße. Dies bestätigt, dass Spikes bei der Bewertung von Optionen auf dem Strommarkt nicht vernachlässigt werden dürfen und zeigt aber auch, dass eine genaue Bestimmung des Parameters η^* unerlässlich ist. Analoge Aussagen lassen sich, wie aus 7.2 (b) zu erkennen ist, für β treffen. Die Intuition hinter höheren Optionspreisen bei steigendem β und sinkendem η^* ist, dass die Wahrscheinlichkeit für das Vorhandensein eines Spikes bei Fälligkeit der Option zurückgeht.

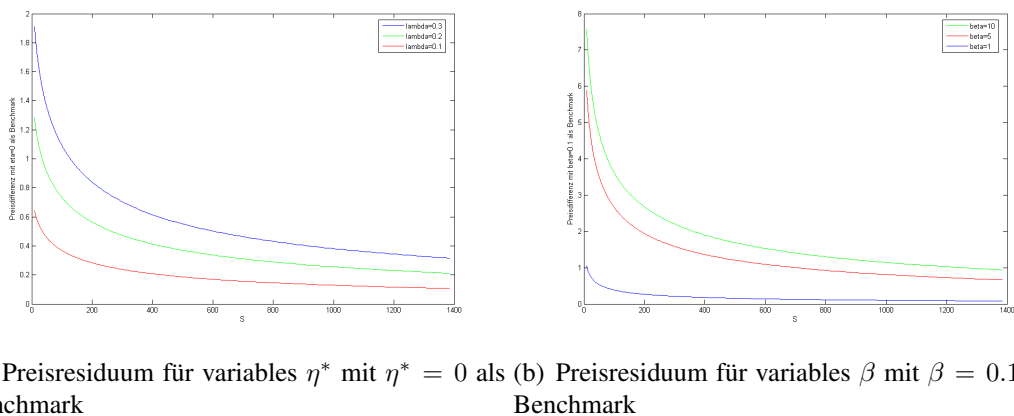
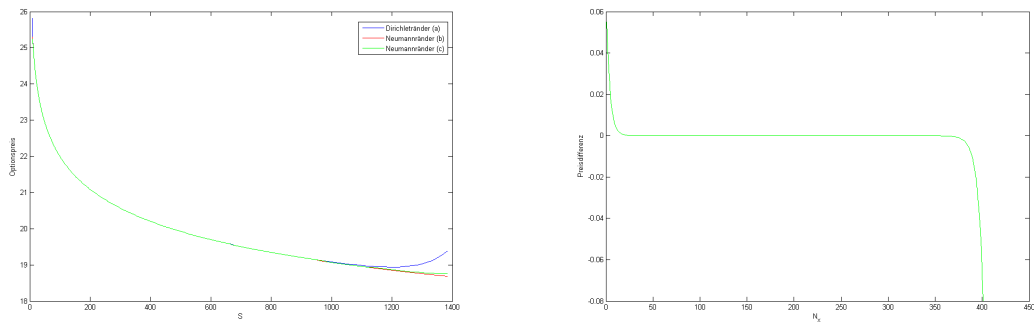


Abbildung 7.2: Preisunterschiede bei Variation von η^* und β

Nun sollen die verschiedenen Ansätze für die Randwerte am x -Rand verglichen werden. Es werden Dirichletränder mit LS-Randbedingungen wie in Gleichung (9.1)(i), Neumannränder mit $\partial v / \partial x = 0$ als Randbedingungen (ii) und Neumannränder mit $\partial^2 V / \partial (e^x)^2 = 0$ als Randbedingung (iii) untersucht. Für die Variante (iii) muss zunächst eine Bedingung für die Randwerte in x hergeleitet werden. Die Bedingung am x -Rand lautet (vgl. Anhang B.4):

$$\frac{\partial^2 v}{\partial x^2} = \frac{\partial v}{\partial x}.$$

In Abbildung 7.3 sind die Werte europäischer Putoptionen berechnet mit den verschie-



(a) Preisvergleich bei einer Restlaufzeit von $\tau = 2$ für unterschiedliche Randwerte (i) blau, (ii) grün und (iii) rot (b) Residuum der Optionspreise aus (ii) und (iii)

Abbildung 7.3: Vergleich der Putpreise für unterschiedliche Randwerte

denen Ansätze dargestellt. Die Preiskurve aus Ansatz (iii) weist einen gewohnten Verlauf auf. Während bei Ansatz (i) an beiden Rändern ein starker Anstieg des Wertes zu beobachten ist. Dies ist damit zu erklären, dass als Randbedingung die analytische Lösung aus dem LS-Modell benutzt wird. Im LS-Modell werden Spikes nicht miteingepreist und somit ist der erwartete Wert eines Puts höher als im Kluge-Modell. Durch den Diffusionsanteil der PIDE bildet sich die beobachtete Randschicht. Die Vorgabe von Dirichletrandwerten ist bei diesem Problem nicht sinnvoll, da die Charakteristiken ins Innere des Gebietes zeigen. Ein Unterschied zwischen den Lösungen aus Variante (ii) und (iii) ist auf den ersten Blick kaum zu erkennen. Daher wurde zusätzlich das Residuum dargestellt. Die Lösung aus Variante (ii) steigt an beiden Rändern langsamer als die Lösung aus (iii), da die Lösung am Rand konstant gesetzt wird (homogener Neumannrand). Aus ökonomischer Sicht sind die Varianten (ii) und (iii) sinnvoll interpretierbar. Eine eindeutige Aussage über das Randverhalten ist im Gegensatz zu vielen physikalischen Anwendungen nicht möglich.

Nun soll die Konvergenz des Verfahrens numerisch untersucht werden. Aus Abschnitt 7.2 ist bekannt, dass das Verfahren linear in $(\Delta x + \Delta y + \Delta t)$ konvergiert. Dies soll nun numerisch bestätigt werden. Dafür berechnen wir die Empirical Order of Convergence (EOC) des Verfahrens. Dieser ist folgendermaßen definiert:

$$EOC = \frac{\ln(\|u - u_{2h}\| / \|u - u_h\|)}{\ln 2}.$$

Dabei bezeichnet u die analytische Lösung des Problems. Da keine analytische Lösung des Problems bekannt ist, approximieren wir diese durch eine numerische Lösung auf einem sehr feinen Gitter. Die feine Lösung wurde auf einem 192×192 -Gitter berechnet und es wurden 400 Zeitschritte verwendet. In Tabelle 7.1 ist die Konvergenzordnung für verschiedene Gitterweiten dargestellt. Dabei wurden die fett gedruckten Werte für die Berechnung des EOC verdoppelt. Wie deutlich zu erkennen ist, ist die Konvergenz in x wesentlich schneller als in t und y . Somit ist eine feine Diskretisierung in dieser Variablen

Tabelle 7.1: EOC

EOC \ Gitter	x	y	t
1.11	24	48	100
1.27	48	48	100
0.27	96	24	100
0.14	96	48	100
0.30	96	48	10
0.06	96	48	20

äußerst sinnvoll.

Kapitel 8

Feller Prozesse vom NIG-Typ

Nachdem in den vorangegangenen Kapiteln ein OU-Modell für die Modellierung von Strompreisen vorgestellt wurde, soll im Folgenden ein alternatives Modell vorgestellt werden. Dieses modelliert den Spotpreis als Feller Prozess vom NIG-Typ (FPNIG). Dabei wird der Effekt der Mean Reversion nicht wie zuvor durch einen Mean Reversion Parameter, sondern durch eine ortsabhängige Definition der Parameter des Prozesses erreicht.

8.1 Das FPNIG-Modell

Um die Dynamik von Strompreisen zu modellieren, ist die Klasse der NIG-Verteilungen leider zu klein, da Spikes und Mean Reversion nicht abgebildet werden können. Um diese Eigenschaften miteinzubeziehen, könnten beispielsweise OU-Prozesse oder allgemeinere Feller-Prozesse betrachtet werden .

Wir werden den zweiten Ansatz in Anlehnung an [6] verfolgen. Die Idee bei diesem Ansatz ist es, die Verteilung der Zuwächse eines NIG-Prozesses ortsabhängig zu definieren, somit wäre die Unabhängigkeit der Inkremente des Prozesses und auch die Stationarität aufgehoben. Indem die Wahrscheinlichkeit für einen großen Sprung nach unten mit dem Stand des Prozesses steigt, ist es möglich, Spikes und Mean Reversion abzubilden. Wir werden eine approximative Bewertungsformel für Optionen auf ein Underlying, welches einem FPNIG folgt, herleiten, wobei diese eine zu (2.37) analoge Darstellung haben wird. Zunächst soll der Prozess definiert werden. Dies ist am einfachsten über den charakteristischen Exponenten möglich. Ganz allgemein gesprochen, formalisieren wir die in der Praxis oft angewandte Methode konstante Koeffizienten in einem Modell zu variieren. Der charakteristische Exponent des FPNIG stellt für jedes feste x den charakteristischen Exponenten eines NIG-Prozesses dar, jedoch hängen die Koeffizienten von x ab.

Definition 8.1 (FPNIG). Sei $\mu, \delta, \alpha, \beta \in C_b^\infty(\mathbb{R})$, δ und α positiv und μ und δ reell,

weiterhin seien Konstanten $C, c > 0$ gegeben d.d. für alle $x \in \mathbb{R}$

$$\delta(x) > 0, \quad \alpha(x) - |\beta(x)| > c, \quad |\mu(x)| \leq C$$

gilt. Wir definieren nun $\psi(x, \xi)$ als

$$\psi_{NIG}(x, \xi) = -\mu(x)i\xi + \delta(x) \left[\sqrt{\alpha(x)^2 - (\beta(x) + i\xi)^2} - \sqrt{\alpha(x)^2 - \beta(x)^2} \right]$$

und bezeichnen einen stochastischen Prozess X mit Erzeuger $-\psi(x, D_x)$ als FPNIG. Der Prozess hat die Ordnung 1 und den exponentiellen Typ $[\lambda_-, \lambda_+]$ für alle $\lambda_- < \lambda_+$ mit

$$\sup_x \{-\alpha(x) + \beta(x)\} < \lambda_- < 0 < \lambda_+ < \inf_x \{\alpha(x) + \beta(x)\}.$$

Für unsere Zwecke genügt eine relativ einfache Ortsabhängigkeit.

Beispiel 8.1 (Parameterwahl). Wir wählen α, δ und μ konstant. Für β wählen wir folgende Darstellung:

$$\beta(x) = -\frac{2\chi}{\pi} \arctan(\epsilon(x - x_0)) + \beta_0,$$

mit $\epsilon > 0, \chi > 0$ und $|\beta_0 \pm \chi| < \alpha$. In Abbildung 8.1 ist die Funktion $\beta(x)$ für eine bestimmte Parameterwahl dargestellt. Bei einer solchen Wahl der Parameter wird sich eine rechtsschiefe Verteilung der Zuwächse für hohe x -Werte und eine linksschiefe Verteilung für niedrige x -Werte ergeben. Betrachten wir nun durch diesen Prozess generierte Pfade

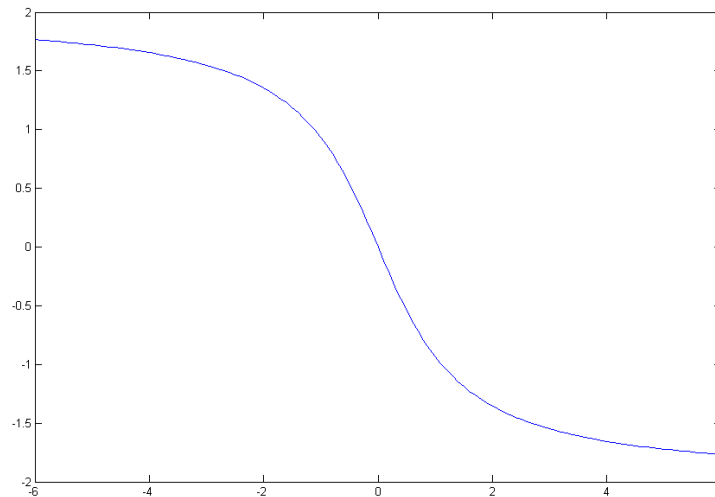


Abbildung 8.1: Darstellung des Parameters β in Abhängigkeit von x . $\chi = 2, \beta_0 = 0, x_0 = 0$ und $\epsilon = 0.9$.

(Abbildung 8.2), so stellen wir fest, dass zwar starke Preisausschläge nach oben auftreten, der Prozess jedoch danach zu langsam auf ein normales Niveau zurückkehrt ¹. Dies

¹Die Simulation des Prozesses wurde mit Hilfe des Programms `randraw` durchgeführt. Dieses kann

ist dadurch zu erklären, dass nach einem starken Preisausschlag nach oben der Prozess zwar mit hoher Wahrscheinlichkeit nach unten springt, jedoch der Tail der Sprungverteilung zu dünn ist und somit nur kleine Sprünge auftreten. Somit verbleibt der Prozess zu lange auf einem hohen Niveau.

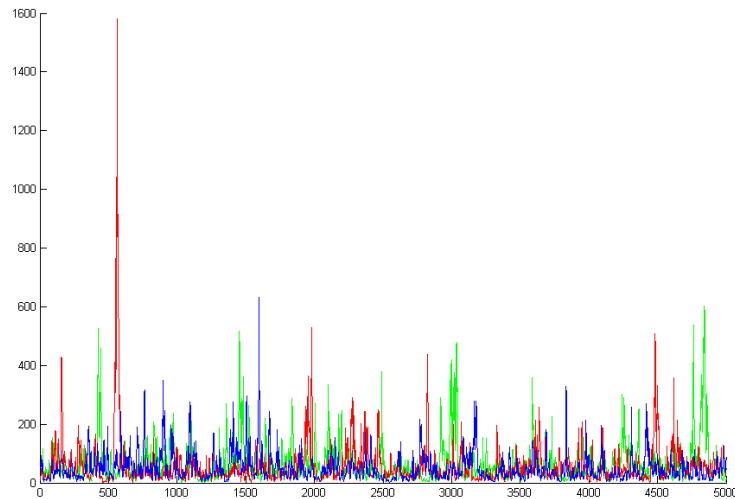


Abbildung 8.2: Simulation eines Pfades für $\alpha = 3$, β wie oben, $\mu = -0.004$ und $\delta = 0.139$

Um dieses Problem zu beheben, ist es notwendig einen weiteren Parameter zu variieren.

Beispiel 8.2 (Parameterwahl). *Es ist naheliegend den Parameter zu verändern, der die Dicke der Tails steuert, also den Parameter α . Der Abfall des rechten Tails wird durch $\alpha - \beta$ gesteuert und der Abfall des linken Tails durch $\alpha + \beta$. Wir wählen für α folgende Darstellung:*

$$\alpha = a_0 + \frac{a_1}{a_2 + x^2}$$

Somit fällt der linke Tail der Dichteverteilung für große x exponentiell nur mit $a_0 + \beta(x)$, wobei $\beta(x) < 0$ gilt. Der rechte Tail fällt mit $a_0 - \beta(x)$. Analog gilt dies für kleine Werte von x . In Abbildung 8.3 sind einige Beispielpfade eines solchen Prozesses abgebildet. Wie deutlich zu erkennen ist, wurde das Problem beseitigt und die Realisationen scheinen die stylised facts gut abbilden zu können.

Für die approximative Bewertungsformel wird es wichtig sein, den Betrag der Ableitungen des charakteristischen Exponenten abschätzen zu können. Hierfür ist folgende Abschätzung der Ableitungen von β sehr nützlich:

$$|\beta^{(s)}(x)| \leq C_s \epsilon^s, \quad \text{für } s \in \mathbb{N} \quad (8.1)$$

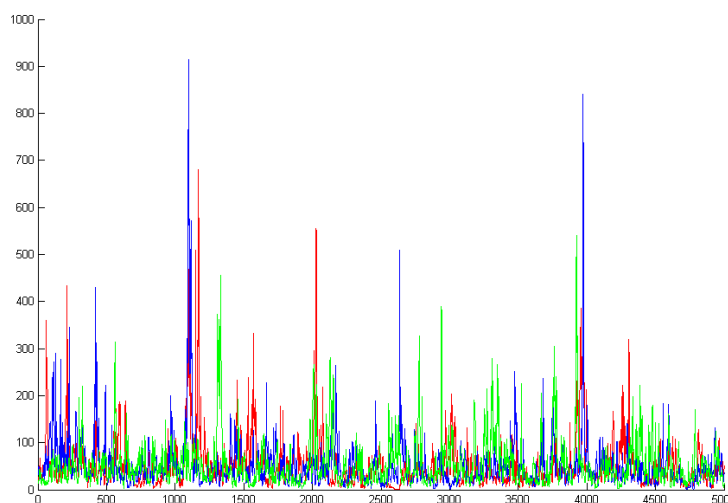


Abbildung 8.3: Simulation eines Pfades für $a_0 = 1.1$, $a_1 = 0.8$, $a_2 = 0.25$, $\mu = -0.004$ und $\delta = 0.139$

und C_s unabhängig von ϵ . Daraus lässt sich folgende Schranke für die Ableitungen des Symbols aus Beispiel 8.1 ableiten:

$$\left| \psi(x, \xi)_{(\kappa)}^{(\gamma)} \right| \leq C \epsilon^\kappa \alpha^{-\gamma}.$$

Für die Parameterwahl in Beispiel 8.2 folgt mit $\alpha_0 = \min \alpha(x) \gg \beta_0 = \max |\beta(x)|$

$$\left| \psi_{(\kappa)}^{(\gamma)}(x, \xi) \right| \leq C \alpha_0^{1-\gamma-\min(1,\kappa)}.$$

8.2 Bewertung von contingent claims im FPNIG-Modell

Wie im Kluge-Modell betrachten wir nun Optionen auf ein Underlying (S) mit

$$S_t = e^{X_t + f(t)},$$

wobei X ein FPNIG mit Erzeuger $-\psi(x, D)$ ist. Analog zu Kapitel 7 berücksichtigen wir das Auszahlungsprofil nur in der Endbedingung. Sei nun $v(t, X_t)$ der Preis eines contingent claims mit Auszahlungsprofil $g(x) \in C_\infty(\mathbb{R})$. Es lässt sich analog zu Abschnitt 2.5 zeigen, dass v folgende Pseudo-Differentialgleichung (vgl. [12, Abschnitt 2.3] und [11, Lemma 14.1]):

$$(\partial_t - (r + \psi(x, D_x))) v(t, x) = 0 \quad (8.2)$$

mit

$$v(T, x) = g(x) \quad (8.3)$$

als Randbedingung erfüllt. Diese Pseudo-Differentialgleichung lässt sich mit Hilfe der Halbgruppentheorie lösen. Nach Satz 2.18 erzeugt der Abschluss des Operators $\psi(x, D)$ eine Feller Halbgruppe auf $C_\infty(\mathbb{R})$. Die Lösung von (8.2) und (8.3) lässt sich demzufolge mit Hilfe der Formel für die Variation der Konstanten ausdrücken:

$$v(t, x) = e^{-rt} T_\tau g(x),$$

wobei $\{T_\tau\}_{\tau \geq 0}$ die von $-\psi(x, D)$ erzeugte Halbgruppe ist. Es gilt also:

$$e^{-r\tau} T_\tau g = \exp[-\tau(r + \psi(x, D))] g.$$

Leider ist es für den FPNIG nicht möglich die Halbgruppe explizit anzugeben, daher müssen wir uns eines Tricks bedienen. Aus Satz 2.12 ist eine Darstellung der Halbgruppe mit Hilfe des Erzeugers bekannt. Um diese nutzen zu können, muss gezeigt werden, dass $r + \psi(x, D)$ eine analytische Halbgruppe auf $C_\infty(\mathbb{R})$ erzeugt. Dies wurde in [11, Lemma 4.1] getan. Somit kann die Lösung der Pseudo-Differentialgleichung folgendermaßen dargestellt werden:

$$e^{-r\tau} T_\tau g = \frac{1}{2\pi i} \int_{L_\phi} e^{\lambda\tau} (\lambda + r + \psi(x, D))^{-1} g \, d\lambda, \quad (8.4)$$

wobei L_ϕ ein Weg mit $\lambda = \lambda(\sigma)$ von $e^{-i\phi}\infty$ nach $e^{i\phi}\infty$ ist. Dabei wählen wir $\arg \lambda(\sigma) = -\phi$ für $\sigma < 0$ und $\arg \lambda(\sigma) = \phi$ für $\sigma > 0$, ϕ wurde wie in Satz 2.12 gewählt. Im Folgenden wird eine asymptotische Entwicklung für das Symbol der Resolvente in (8.4) hergeleitet.

Satz 8.1. *Die Resolvente in (8.4) ist ein PDO und das zugehörige Symbol besitzt eine zu (2.27)-(2.29) analoge Darstellung.*

Beweis. Für den Operator $P(\lambda) = (\lambda + r + \psi(x, D))$ müssen die Voraussetzungen von Satz 2.17 gezeigt werden. Dass $P(\lambda)$ von H^s nach H^{s-1} abbildet, folgt aus [5, Theorem A.8] unter Ausnutzung der Tatsache, dass der Besovraum $B_{2,2}^s$ dem Sobolevraum H^s entspricht. Dass $P(\lambda)$ invertierbar ist, folgt aus der Tatsache, dass nach obiger Argumentation $r + \psi(x, D)$ eine analytische Halbgruppe erzeugt. Somit enthält die Resolvente nach Satz 2.12 einen Sektor und der Integrationsweg L_ϕ verläuft innerhalb dieses Sektors. Die Abschätzung aus Definition 2.17 folgt aus folgendem Argument. Es genügt $|P(\lambda)|$ für ξ mit $\Re(\xi) \rightarrow \infty$ zu betrachten. Es gilt mit $\xi = a + ib$:

$$\sqrt{\alpha(x)^2 - (\beta(x) + i\xi)^2} = \sqrt{\alpha(x)^2 - \beta(x)^2 + a^2 - b^2 - 2i\xi\beta(x) + 2iab}. \quad (8.5)$$

Somit verhält sich dieser Term für große ξ wie $\Re(\xi)$ und es gilt insgesamt:

$$|\psi(\xi, x)| \sim |i\mu(x)\Re(\xi) + \Re(\xi)| = \sqrt{\mu(x)^2 + (\Re(\xi))^2}. \quad (8.6)$$

Die benötigte Abschätzung folgt aus der Beschränktheit des Imaginärteils von ξ (siehe auch [12, Lemma 4.1]). Somit sind die Voraussetzungen von Satz 2.17 gezeigt. \square

Also ist der zu $\lambda + r + \psi(x, D)$ inverse Operator ein PDO dessen Symbol (b) der Klasse $S_{1,0}^{-1}(\mathbb{R} \times \mathbb{R})$ angehört und die asymptotische Entwicklung (2.27)-(2.29) zulässt. Nun sollen zunächst die beiden ersten Terme aus (2.27) hergeleitet und in (8.4) eingesetzt werden. Das erste Glied der asymptotischen Entwicklung von b ist

$$(\lambda + r + \psi(x, \xi))^{-1}.$$

Setzen wir dies in (8.4) ein und vertauschen die Integrationsreihenfolge, was auf Grund der absoluten Konvergenz des Integranden möglich ist, lässt sich das erste Glied der Reihe folgendermaßen darstellen:

$$\frac{1}{4\pi^2 i} \int_{-\infty}^{+\infty} e^{ix\xi} \int_{L_\phi} e^{\lambda\tau} (\lambda + r + \psi(x, \xi))^{-1} \hat{g}(\xi) d\lambda d\xi.$$

Mit Hilfe des Residuensatzes lässt sich dies umschreiben zu

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(ix\xi - \tau(r + \psi(x, \xi))) \hat{g}(\xi) d\xi, \quad (8.7)$$

da der Integrand im inneren Integral in $-r - \psi(x, \xi)$ einen Pol erster Ordnung besitzt. Das zweite Glied in der asymptotischen Entwicklung von b ergibt sich über die Kettenregel zu

$$b_{-m-1} = (\lambda + r + \psi(x, \xi))^{-3} \psi^{(1)}(x, \xi) \psi_{(1)}(x, \xi).$$

Dies setzen wir wieder in (8.4) ein, vertauschen die Integrationsreihenfolge und erhalten:

$$\frac{1}{4\pi^2 i} \int_{L_\phi} \int_{-\infty}^{+\infty} \frac{\exp(\lambda\tau + i\xi x) \psi^{(1)}(x, \xi) \psi_{(1)}(x, \xi)}{(\lambda + r + \psi(x, \xi))^{-3}} \hat{g}(\xi) d\xi d\lambda.$$

Mit einem analogen Argument zu (8.7) lässt sich obiger Term zu

$$\frac{\tau^2}{4\pi} \int_{-\infty}^{+\infty} \exp(ix\xi - \tau(r + \psi(x, \xi))) \psi^{(1)}(x, \xi) \psi_{(1)}(x, \xi) \hat{g}(\xi) d\xi \quad (8.8)$$

vereinfachen.

Es wurde also der Integraloperator für die beiden ersten Glieder der asymptotischen Entwicklung von b hergeleitet, wir werden sehen, dass dies für unsere Zwecke genügt. Analog können wir die Rechnung für einen weitergefassten Raum der Auszahlungsprofile

durchführen. Es genügt $g(x)$ stückweise stetig anzunehmen und folgende Abschätzung zu fordern:

$$g(x) \leq C e^{-w_- x} \quad \text{für } x > 0 \quad \text{und} \quad g(x) \leq C e^{-w_+ x} \quad \text{für } x < 0,$$

wobei w_- und w_+ so zu wählen sind, dass $\lambda_- < w_- < w_+ < \lambda_+$. Die intuitive Erklärung hinter dieser Forderung ist die Aussage, dass die erwartete Auszahlung der Option endlich sein muss, also für große x langsamer steigen muss als die Dichte abfällt. Dabei beschreiben die Parameter λ_- und λ_+ die Dicke der Tails. Für einen europäischen Call oder Put reduziert sich diese Bedingung auf Grund der Linearität des Auszahlungsprofils, auf $\lambda_- < -1 < 0 < \lambda_+$.

Der Integrationsweg in (8.8) und (8.7) verschiebt sich entsprechend. Für einen europäischen Put würden wir beispielsweise über $x \in (-\infty, \infty) \times i\sigma$ mit $\sigma \in (0, \lambda_+)$ integrieren. Eine Verschiebung ist notwendig, um die Wohldefiniertheit der Fouriertransformation zu gewährleisten.

Wir approximieren den Wert der Option, indem wir nach den ersten beiden Gliedern der asymptotischen Entwicklung der Resolvente abschneiden. Nun können wir zur Kalibrierung des FPNIG übergehen.

8.3 Kalibrierung und Implementierung

Das Modell wurde mit Hilfe von Optionspreisen aus dem Kluge-Modell kalibriert. Dabei wurde ein least-squares Fit in Matlab durchgeführt. Zu beachten ist, dass eine Driftrestriktion für die risikoneutrale Kalibrierung des FPNIG, also eine Einschränkung an das Wachstum des Underlyings unter dem Bewertungsmaß, nicht eingehalten werden muss. Dies hat den Grund, dass der Stromkurs unter einem risikoneutralen Maß kein Martingal sein muss, da keine Arbitrageüberlegungen zwischen verschiedenen Stromlieferungen in verschiedenen Perioden möglich sind. Für die Bewertung einer Option mit Auszahlungsprofil $g(x)$ muss das folgendes Integral bestimmt werden:

$$\begin{aligned} v(x, \tau) &\approx \frac{1}{2\pi} \int_{-\infty+i\sigma}^{+\infty+i\sigma} \exp(ix\xi - \tau(r + \psi(x, \xi))) \hat{g}(\xi) d\xi \\ &\quad + \frac{\tau^2}{4\pi} \int_{-\infty+i\sigma}^{+\infty+i\sigma} \exp(ix\xi - \tau(r + \psi(x, \xi))) \psi^{(1)}(x, \xi) \psi_{(1)}(x, \xi) \hat{g}(\xi) d\xi \\ &= K(x, \tau)(g), \end{aligned}$$

mit $\psi(x, \xi)$ wie oben. Um das Integral zu bestimmen, muss zunächst die Fouriertransformierte des Auszahlungsprofils berechnet werden. Für einen europäischen Put ist diese

folgendermaßen gegeben:

$$\begin{aligned} \hat{g}(\xi) &= \int_{-\infty}^{\infty} (K - e^{f(t)+x})^+ e^{-ix\xi} dx = \int_{-\infty}^{\ln(K)-f(T)} (K - e^{f(T)+x}) e^{-ix\xi} dx \\ &= \frac{K e^{-i\xi(\ln K - f(T))}}{-i\xi} - \frac{\exp(f(T)(\ln K - f(T))(1 - i\xi))}{1 - i\xi}, \end{aligned}$$

wobei $f(t)$ wie immer die logarithmierte Saisonalität ist und $\xi \in \mathbb{R} \times i\sigma$ mit $\sigma \in (0, \lambda_+)$ gilt.

Die zwei zu bestimmenden Integrale wurden mit Hilfe der Trapezregel diskretisiert. Die

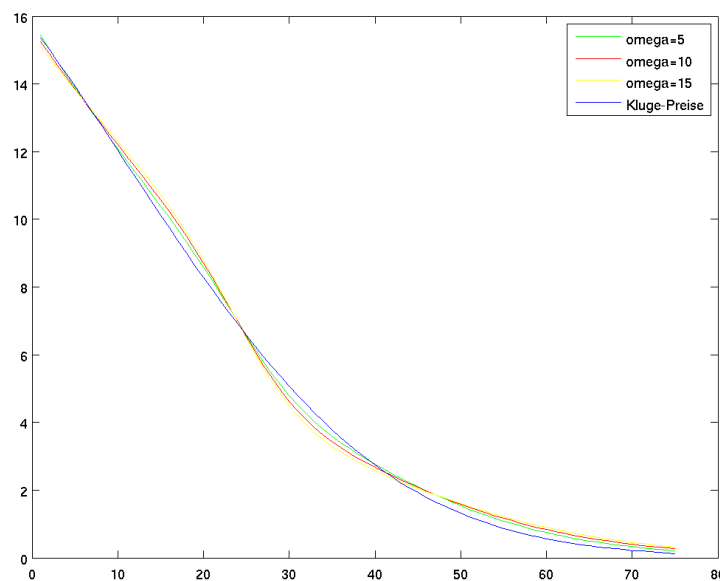


Abbildung 8.4: Preise aus dem Kluge-Modell (blau) und die kalibrierten FPNIG-Preise für unterschiedliche ω

Bestimmung der Parameter des Prozesses aus Optionspreisen ist ein schlecht gestelltes inverses Problem (vgl. [14, Kapitel 13]). In unserem Fall bedeutet dies, dass einerseits die Parameter sehr sensibel auf kleine Änderungen der Optionspreise reagieren und andererseits die optimale Parameterkombination nicht eindeutig sein muss. Somit führt ein least-squares Fit unweigerlich zu nicht sinnvollen Ergebnissen. Um dieses Problem zu beheben, benötigen wir Informationen über die ungefähre Lage der Parameter. Diese können durch Simulation des Prozesses erworben werden. Das besser gestellte Problem lautet nun:

$$\min_{\theta \in \Theta} \|K(x, \tau, \theta)(g) - y\|_2^2 + \omega \|\theta^0 - \theta\|_2^2.$$

Dabei ist Θ die Menge der zulässigen Parameterkombinationen, y sind die zu x und τ korrespondierenden Optionswerte aus dem Kluge-Modell und θ^0 ist eine Parameterkombination die vernünftige Optionspreise liefert. Der Parameter ω regelt, wie stark der Vektor θ^0

Tabelle 8.1: Kalibrierte Parameter im FPNIG-Modell

Glättungspar. \ Kalibrierte Par.	μ	δ	χ	ϵ	a_0	β_0	a_1
5	-0.07	1.33	2.29	1.97	3.98	0.57	1.15
10	-0.06	1.43	2.05	1.77	3.58	0.55	1.27
15	-0.06	1.51	1.93	1.64	3.34	0.55	1.42
20	-0.06	1.57	1.86	1.55	3.16	0.55	1.56
25	-0.06	1.54	1.82	1.93	3.08	0.55	1.84

in die Optimierung eingeht, für $\omega = 0$ erhalten wir unser ursprüngliches Problem. Das Modell wurde mit at-the-money² Puts kalibriert. Die Optimierung wurde in Matlab mit Hilfe der Optimierungstoolbox durchgeführt. Dabei wurde die `fmincon`-Routine unter Benutzung des active-set line search Algorithmus benutzt. Es wurden unterschiedliche Parameterkombinationen als freie Parameter in den Optimierungsalgorithmus eingegeben. Als signifikant erweisen sich die Parameter μ , δ , χ , ϵ , a_0 , β_0 und a_1 . Die Optimierung unter Benutzung der realen Saisonkomponente ist nicht gelungen, da die Größe der Risikoprämien vom Prozess nicht erklärt werden konnte. Deswegen wurde der Prozess mit Hilfe der risikoneutralen Saisonkomponente gefittet. Die optimierten Werte unter Benutzung verschiedener Parameter ω sind in Tabelle 8.1 dargestellt. In Abbildung 8.4 sind die Preise aus dem Kluge-Modell und die daran kalibrierten FPNIG Preise abgebildet. Der MSE beträgt in Abhängigkeit von ω zwischen 0.06 Euro und 0.23 Euro. Der Realteil des zweiten Terms der asymptotischen Entwicklung ist von der Größenordnung 10^{-16} , so dass damit heuristisch das Vernachlässigen der nachkommenden Terme begründet ist. Nun ist das FPNIG-Modell kalibriert und wir können zum Vergleich der Modelle übergehen.

²Wir definieren die moneyness in Bezug auf die Forwardpreise und nicht auf den aktuellen Stand des Prozesses. Der aktuelle Stand des Prozesses bildet auf Aktienmärkten einen vernünftigen Referenzpunkt, während auf Strommärkten auf Grund der Saisonalitäten der Forwardpreis wesentlich besser geeignet ist.

Kapitel 9

Modellvergleich

9.1 Optionsbewertung im Lucia-Schwartz-Modell

Um die numerischen Ergebnisse auswerten zu können, vergleichen wir nun die Modelle untereinander. Wir nutzen zusätzlich das in der Einleitung vorgestellte LS-Modell, in dem Spikes nicht in die Bewertung eingehen. Somit würden wir erwarten, dass ein Call im Kluge-Modell teurer ist, da zusätzlich zum Diffusionsrisiko ein Spikerisiko bewertet wird. Eine solche Aussage in Bezug auf das FPNIG-Modell lässt sich a priori nicht treffen. Im LS-Modell ergeben sich bei der Kalibrierung dieselben Parameter wie beim Diffusionsprozess im Kluge-Modell. Der Strompreis lässt sich durch folgende stochastische Differentialgleichung beschreiben:

$$S_t = \exp(f(t) + X_t)$$

$$dX_t = -\alpha X_t + \sigma dW_t$$

Es lässt sich zeigen, dass $\ln S_T$ normalverteilt mit Erwartungswert $m(T) = f(T) + (\ln S_t - f(t))e^{-\alpha(T-t)}$ und Varianz $v(T) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(T-t)})$ (vgl. Anhang B.3). Somit kann mit Hilfe von Satz 6.2 sowohl der Call als auch der Forward bewertet werden. Sei der Einfachheit halber $s(t) = e^{f(t)}$ dann folgt

$$F_t^T = \mathbb{E}[S_T | S_t] = \exp(m(T) + \frac{1}{2}v(T)) = s(T) \left(\frac{S_t}{s(t)}\right)^{e^{-\alpha(T-t)}} \exp\left(\frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha(T-t)})\right).$$

Für einen europäischen Call und Put gilt:

$$\begin{aligned} C(t, S_t, T, K) &= e^{-r(T-t)} (\mathbb{E}[(S_T - K)^+ | S_t]) \\ &= e^{-r(T-t)} (\mathbb{E}[S_T \mathbf{1}_{S_T \geq K} | S_t] - K \mathbb{E}[\mathbf{1}_{S_T \geq K} | S_t]) \\ &= e^{-r(T-t)} (\mathbb{E}[S_T | S_t] N(d_1) - K N(d_2)) \\ &= e^{-r(T-t)} (F_t^T N(d_1) - K N(d_2)) \end{aligned}$$

$$P(t, S_t, T, K) = e^{-r(T-t)}(KN(-d_2) - F_t^T N(-d_1)) \quad (9.1)$$

$$\text{mit } d_1 = \frac{-\ln K + m(T) + v(T)}{\sqrt{v(T)}} \quad \text{und} \quad d_2 = \frac{-\ln K + m(T)}{\sqrt{v(T)}}.$$

Die Gleichungen für d_1 und d_2 werden in Anhang B.3 hergeleitet. In der Black-Scholes Welt lässt sich $e^{-t(T-t)}F_t^T$ zu S_t vereinfachen. Dies ist bei Strommarktmodellen nicht möglich, da keine Arbitragebeziehung zwischen Strompreisen zu verschiedenen Zeitpunkten besteht. Aus den Bewertungsformeln für den Call und dem Future wird klar, dass für große Restlaufzeiten (großes $T - t$) der Einfluss des aktuellen Preises sehr gering ist. Betrachten wir hierfür die einzelnen Terme, die in den Preis eingehen

$$m(T) \sim f(T) \quad \text{und} \quad v(T) \sim \frac{\sigma^2}{2\alpha},$$

so folgt für den Forward:

$$F_t^T \sim s(T) \exp\left(\frac{\sigma^2}{4\alpha}\right).$$

Für den europäischen Call lässt sich eine ähnliche Abschätzung herleiten. Mit

$$d_1 \sim \frac{-\ln K + f(T) + \frac{\sigma^2}{2\alpha}}{\sqrt{\sigma^2 2\alpha}} = \tilde{d}_1 \quad \text{und} \quad d_2 \sim \frac{-\ln K + f(T)}{\sqrt{\sigma^2 2\alpha}} = \tilde{d}_2$$

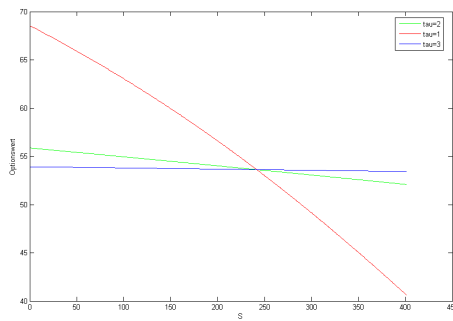
folgt

$$C(t, S_t, T, K) \sim e^{-r(T-t)}(s(T) \exp\left(\frac{\sigma^2}{4\alpha}\right)N(\tilde{d}_1) - KN(\tilde{d}_2)).$$

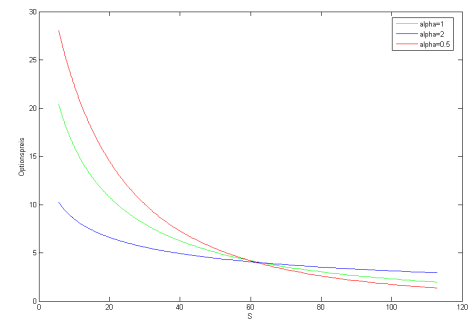
Somit wären der Callpreis und der Putpreis im Grenzfall unabhängig von dem heutigen Kurs (S_t). In Abbildung 9.1 (a) sind Putpreise zu verschiedenen Restlaufzeiten und in Abhängigkeit vom Strike abgetragen. Es ist ganz klar zu erkennen, dass mit steigender Restlaufzeit der Einfluss des aktuellen Preises abnimmt. Dies liegt daran, dass mit längerer Laufzeit damit zu rechnen ist, dass ein Preis der über dem Saisoneinfluss liegt durch die Mean Reversion Eigenschaft des Prozesses geglättet wird. Wie schnell dieser Prozess vor sich geht, hängt von α ab und ist in Abbildung 9.1 (b) demonstriert. Für das Kluge-Modell hatten wir ein ähnliches Verhalten beobachtet.

9.2 Vergleich der Optionspreise

Im Folgenden sollen die Optionspreise aus den verschiedenen Modellen verglichen werden. In Abbildung 9.2 sind die relativen Unterschiede zwischen dem LS-Modell und dem Kluge-Modell dargestellt ($P_{\text{Kluge}}/P_{\text{LS}}$). Wie deutlich zu erkennen ist, werden Calls im LS-



(a) Unterschiedlichen Restlaufzeiten



(b) Unterschiedliche Parameterwahlen für α

Abbildung 9.1: Puts im Lucia Schwartz-Modell

Modell deutlich niedriger bewertet als im Kluge-Modell. Dies rührt daher, dass das Spikerisiko nicht mit einbezogen wird. Die Abweichungen treten verstärkt bei deep-out-of-the-money Optionen auf. Da bei solchen Optionen das Spikerisiko einen wesentlichen Beitrag zum Wert der Option leistet. In Abbildung 9.3 sind Putpreise im LS-Modell und im

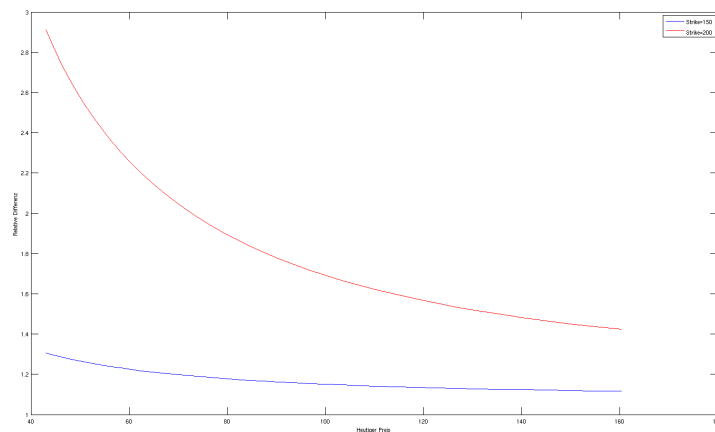


Abbildung 9.2: Relative Unterschiede zwischen Optionspreisen im LS-Modell und im Kluge-Modell

FPNIG-Modell für unterschiedliche Strikes dargestellt. Ein Basispreis von 50 entspricht einer at-the-money Option, während bei einem Basispreis von 110 von einer deep-in-the-money Option gesprochen werden kann. In beiden Fällen werden Puts im FPNIG-Modell niedriger bewertet als im LS-Modell. Bei den deep-in-the-money Puts ist dies stärker ausgeprägt, somit würden deep-out-of-the-money Calls im FPNIG-Modell deutlich höher bewertet als im LS-Modell. Dies ist nicht verwunderlich, da das Modell mit Preisen aus dem Kluge-Modell kalibriert wurde und somit Spikes mit in die Bewertung eingehen.

Zuletzt wollen wir das Kluge-Modell und das FPNIG-Modell vergleichen. Für at-the-money Optionen ist ein Vergleich nicht sinnvoll, da diese für die Kalibrierung verwendet wurden. Deswegen vergleichen wir das Verhalten an den Tails. Wir betrachten hierfür

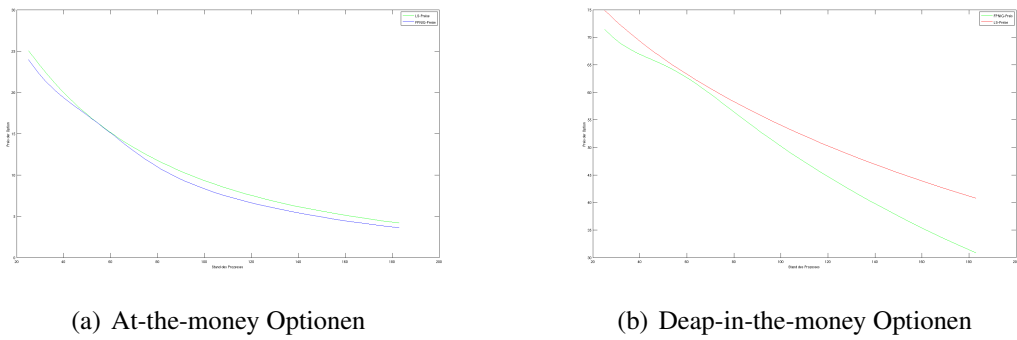


Abbildung 9.3: Vergleich von Optionen im LS-Modell und FPNIG-Modell

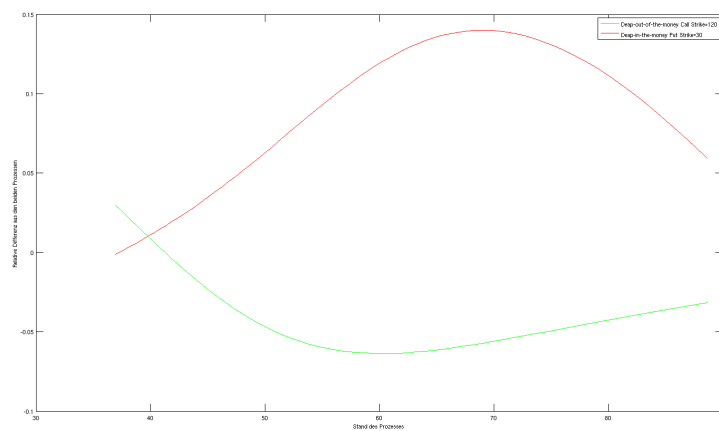


Abbildung 9.4: Vergleich von Optionspreisen im FPNIG-Modell und im Kluge-Modell

die relativen Differenzen aus den Preisen ($(P_{\text{Kluge}} - P_{\text{NIG}}) / P_{\text{Kluge}}$). In Abbildung 9.4 sind diese für einen deep-out-of-the-money Call und einen deep-in-the-money Put dargestellt. Wie deutlich zu erkennen ist, wird der Call im Kluge-Modell deutlich niedriger bewertet als im FPNIG-Modell. Dies kann mit den semi-heavy Tails der NIG-Verteilung erklärt werden. Analog ist die Erklärung für den deep-in-the-money Put. Um ein Urteil über die Güte der Modelle zu treffen, ist ein Vergleich mit realen Marktdaten notwendig. Diese stehen uns leider nicht zur Verfügung.

Der wesentliche Vorteil der Bewertung im FPNIG-Modell gegenüber dem Kluge-Modell ist die Rechengeschwindigkeit. Die Bewertung ist auf einer gewöhnlichen CPU innerhalb von wenigen Sekunden möglich und ist außerdem unabhängig von der Restlaufzeit einer Option. Wesentliche Nachteile sind die recht schwierige Kalibrierung sowie die schwer zu interpretierenden Parameter des Prozesses. Des Weiteren ist es bei unserer Wahl der Funktionen $\alpha(x)$ und $\beta(x)$ schwer, sehr geringe Abhängigkeiten von Basispreisen, die einer hohen Mean Reversion Geschwindigkeit entsprechen, abzubilden.

Kapitel 10

Fazit

Ziel dieser Arbeit war es Stundenoptionen auf Strom in geeigneten Modellen zu bewerten.

Die Bewertung von Optionen in den untersuchten Modellen führt auf das Lösen von PIDEs bzw. PDOs. Es konnte die Existenz einer Lösung der PIDE für europäische Putoptionen in einem der Modelle gezeigt werden. Mit Hilfe der Finite Differenzen Methode wurde die entsprechende PIDE diskretisiert und der Optionswert bestimmt. Es zeigte sich bei der Untersuchung der Optionspreise, dass eine Trennung der Risikoprämien nur mit Hilfe von Forwardpreisen kaum möglich ist. Daher ergeben sich je nach Annahmen über die Bepreisung der Risikoquellen auch stark unterschiedliche Preise. Danach wurden europäische Optionen in einem verallgemeinerten Levy-Modell bewertet. Die Bewertung führt auf das Lösen von eines PDO und wurde mit Hilfe einer approximativen Formel durchgeführt. Die Kalibrierung wurde mit Hilfe von Optionspreisen aus dem ersten Modell durchgeführt. Dabei ist ein großer Vorteil des Levy-Modells in der wesentlich niedrigeren Rechenzeit zu sehen.

Eine mögliche Erweiterungen des Kluge-Modells wäre durch die Hinzunahme einer nicht konstanten Volatilität möglich. Mögliche Erweiterungen des Levy-Modells wären einerseits durch die Variation anderer Parameter möglich, desweiteren wären Alternativen zur benutzten Kalibrierungsmethode eine interessante Frage.

In beiden Fällen ist ein Abgleich mit realen Preisen sehr interessant. Dies war nicht möglich, da zur Zeit keine börsennotierten Stundenoptionen auf Strom gehandelt werden.

Eine weitere Anwendung der Ergebnisse der Arbeit ist die Bewertung von Swingkontrakten auf Stundenbasis, da diese sich in ein Portfolio von europäischen Optionen und Forwardkontrakten zerlegen lassen (vgl.[21]).

Anhang A

Halbgruppentheorie

A.1 Hilfssätze der Halbgruppentheorie

Satz A.1. [29, Satz 12.7] Sei $\{T(t)\}_{t \geq 0}$ eine C_0 -Halbgruppe beschränkter linearer Operatoren auf einem Banachraum X . Dann existiert der Limes

$$\omega_0 = \lim_{t \rightarrow \infty} \frac{\log \|T(t)\|}{t},$$

wobei dieser auch $-\infty$ sein kann und für alle $\gamma > \omega_0$ existiert ein M_γ mit $\|T(t)\| \leq M_\gamma \exp(\gamma t)$.

Beweis. Es folgt aus der Halbgruppeneigenschaft direkt mit Dreiecksungleichung die Subadditivität von $\log(\|T(t)\|)$. Aus der Stetigkeit von $t \mapsto T(t)x$ für alle x , folgt die Beschränktheit von $\|T(t)x\|$ auf jedem beschränkten Intervall. Nach dem Satz von Banach-Steinhaus folgt daraus die Beschränktheit von $\|T(t)\|$ auf jedem beschränkten Intervall. Damit folgt mit Hilfe von [29, Lemma 12.6] die Existenz eines ω_0 mit obigen Eigenschaften. Für alle $\gamma > \omega_0$ gibt es nun ein t_0 d.d. für alle $t \geq t_0$ $\frac{\log \|T(t)\|}{t} < \gamma$ gilt also auch $\|T(t)\| < \exp(\gamma t)$. Die Aussage des Satzes folgt für

$$M_\gamma = \max(1, \sup_{t \in [0, t_0]} \|T(t)\| \exp(-\gamma t)). \quad \square$$

A.2 Fundamentallösung für zeitabhängige Operatoren

Um die Fundamentallösung des Problems (2.24) definieren zu können, sind einige Vorbereitungen notwendig. Mit Hilfe der Evolutionsoperatoren lässt sich auch die Eindeutigkeit einer Lösung zeigen.

Definition A.1.

$$J_\Delta = \{(t, s) \in J \times J; s \leq t\} \tag{A.1}$$

und

$$J_{\Delta}^* = \{(t, s) \in J_{\Delta}; s < t\}.$$

Sei φ eine Funktion auf J_{Δ}^* und ψ eine Funktion auf J , dann nutzen wir folgende Konventionen:

$$\psi\varphi(t, s) = \psi(t)\varphi(t, s), \quad \varphi\psi(t, s) = \varphi(t, s)\psi(s)$$

Nun können wir einen parabolischen Evolutionsoperator definieren.

Definition A.2. Sei X ein Banachraum und J ein Zeitintervall. $A(t)$ sei für alle $t \in J$ ein stetiger Operator auf X . Sei weiter $D(A)$ konstant und dicht in X . Ist nun F ein Banachraum und dicht in X , d.d. $D(A) \subset F$ gilt, so wird eine Abbildung $U : J_{\Delta} \rightarrow \mathcal{L}(X)$ parabolischer Evolutionsoperator oder Fundamentallösung mit regularisierendem Unterraum F genannt, falls

1.

$$U \in C(J_{\Delta}, \mathcal{L}(X)) \cap C(J_{\Delta}^*, \mathcal{L}(X, F))$$

2.

$$U(t, t) = 1, \quad U(t, s) = U(t, \tau)U(\tau, s) \quad s \leq \tau \leq t, \quad (t, s) \in J_{\Delta}$$

3.

$$R(U(t, s)) \subset D(A(t)) \subset F, \quad (t, s) \in J_{\Delta}^*$$

4.

$$AU \in C(J_{\Delta}^*, \mathcal{L}(X))$$

5.

$$\sup_{(t,s) \in I_{\Delta}^*} (t-s) \|AU(t, s)\|_{\mathcal{L}(X)} < \infty, \quad I \subset \subset J$$

6.

$$U(\cdot, s) \in C^1(J \cap (s, \infty), \mathcal{L}(X)), \quad s \in J$$

7.

$$\partial_1 U = AU$$

8.

$$U(t, \cdot) \in C^1(J \cap (-\infty, t), \mathcal{L}(F, X)), \quad t \in J$$

9.

$$\partial_2 U \supset UA$$

Aus dem nun folgenden Satz ergibt sich die Eindeutigkeit des Evolutionsoperators und somit die der Lösung.

Satz A.2. [2, Satz 2.1.2] Sei $f \in C([s, \infty), X)$ und u eine Lösung des folgenden Anfangswertproblems:

$$\frac{\partial u}{\partial t} = A(t)u + f(t), \quad t \in (s, \infty), \quad u(s) = x. \quad (\text{A.2})$$

Sei weiter $V \in C(J_\Delta, \mathcal{L}(X))$ mit $V(t, t) = 1$ und

$$V(t, \cdot) \in C^1((s, t), \mathcal{L}(F, X))$$

mit $\partial_2 V \supset VA$ für $t \in (s, \infty)$. Dann gilt:

1. u lässt sich folgendermaßen darstellen:

$$u(t) = V(t, s)u(s) + \int_s^t V(t, \tau)f(\tau) d\tau \quad (\text{A.3})$$

2. Angenommen es existiere ein parabolischer Evolutionsoperator U für A . Dann hat das Anfangswertproblem (A.2) höchstens eine Lösung und diese ist durch (A.3) mit $V=U$ gegeben.

3. Es gibt höchstens einen parabolischen Evolutionsoperator für A .

Beweis. Ad 1:

Es gilt $u(t) \in D(A) \subset F$ für $t \in (s, \infty)$ und $u(s) = x \in X$, da u eine Lösung ist. Daraus folgt, dass für $t \in (s, \infty)$ $V(t, \cdot)u \in C^1((s, t)) \cap C([s, t], X)$ gilt. Weiter gilt analog zu Satz 2.9 $\partial_2(Vu)(t, \tau) = Vf(t, \tau)$ für $s < \tau < t$ und somit mit dem Hauptsatz der Differential- und Integralrechnung:

$$V(t, t')u(t') - V(t, s')u(s') = \int_{t'}^{s'} V(t, \tau)f(\tau) d\tau$$

Die Aussage folgt nun für $t' \rightarrow t$ und $s' \rightarrow s$

Ad 2:

Folgt sofort aus 1, da sich zwei mögliche Lösungen mit Hilfe von (A.3) mit $V = U$ darstellen ließen.

Ad 3:

Seien U_j $j \in \{1, 2\}$ zwei Evolutionsoperatoren für A mit regularisierenden Unterräumen F_j . Sei $s \in \mathbb{R}_+$ beliebig und $x \in X$, dann folgt aus Definition A.1 und (1), dass $U_j(\cdot, s)x$ Lösungen des Anfangswertproblems (A.2) mit $f = 0$ sind. Aus (2) folgt, dass die beiden Lösungen übereinstimmen, also folgt $U_1 = U_2 = U$ und der Schnitt von F_1 und F_2 ist ein regularisierender Unterram für U . Die Funktion (A.2) lässt sich bereits definieren, wenn nur die lokale Integrierbarkeit von f gefordert wird. Sie wird analog zu Satz 2.9 als schwache Lösung bezeichnet. \square

A.3 Satz von Hille-Yosida

Satz A.3 (Satz von Hille-Yosida). *Sei A ein Operator auf einem Banachraum X . Dann ist A genau dann der Erzeuger einer C_0 -Halbgruppe mit $\|T(t)\| \leq M \exp(\omega t)$, wenn*

1. $D(A)$ dicht in X und A abgeschlossen ist.
2. Alle $\lambda \in \mathbb{R}$ mit $\lambda > \omega$ in der Resolventenmenge enthalten sind und $\|R_\lambda(A)^n\| \leq \frac{M}{(\lambda - \omega)^n}$ für alle $n \in \mathbb{N}$ gilt.

Beweis. „ \Leftarrow “ Die Notwendigkeit der ersten Bedingung folgt aus dem Satz 2.8. Um die Notwendigkeit der zweiten Bedingung einzusehen, betrachten wir die Funktion

$$I_n(\lambda)x = \frac{1}{(n-1)!} \int_0^\infty t^{n-1} \exp(-\lambda t) T(t)x \, dt.$$

$I_n(\lambda)$ ist für $\operatorname{Re}(\lambda) > \omega$ wohldefiniert und die Norm lässt sich folgendermaßen abschätzen:

$$\|I_n(\lambda)\| \leq \frac{1}{(n-1)!} \left| \int_0^\infty t^{n-1} \exp(-\operatorname{Re}(\lambda)t) M \exp(\omega t) \, dt \right|.$$

Durch wiederholte partielle Integration erhalten wir:

$$\|I_n(\lambda)\| \leq \left| \frac{M}{(\operatorname{Re}(\lambda) - \omega)^n} [\exp((\omega - \operatorname{Re}(\lambda))t)]_0^\infty \right| = \frac{M}{(\operatorname{Re}(\lambda) - \omega)^n}$$

Für $n > 1$ und $x \in D(A)$ folgt:

$$\begin{aligned} I_n(\lambda)Ax &= \frac{1}{(n-1)!} \int_0^\infty t^{n-1} \exp(-\lambda t) T(t)Ax \, dt \\ &\stackrel{(2.19)}{=} \frac{1}{(n-1)!} \int_0^\infty t^{n-1} \exp(-\lambda t) \frac{d}{dt}(T(t)x) \, dt \\ &\stackrel{P.I.}{=} -\frac{1}{(n-2)!} \int_0^\infty t^{n-2} \exp(-\lambda t) T(t)x \, dt \\ &\quad + \frac{\lambda}{(n-1)!} \int_0^\infty t^{n-1} \exp(-\lambda t) T(t)x \, dt \\ &= -I_{n-1}(\lambda)x + \lambda I_n(\lambda)x. \end{aligned}$$

Für $n = 1$ folgt mit partieller Integration $I_1(\lambda)Ax = -x + \lambda I_1(\lambda)x$. Wir können obige

Rechnung auch für $\frac{T(h)-I}{h}I_n(\lambda)x$ durchführen.

$$\begin{aligned}
 & \lim_{h \rightarrow 0} \frac{T(h) - I}{h} I_n(\lambda)x \\
 &= \lim_{h \rightarrow 0} \frac{1}{h(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} T(t+h)x \, dt - \frac{1}{h(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} T(t)x \, dt \\
 &= \lim_{h \rightarrow 0} \frac{1}{h(n-1)!} \int_h^\infty (t-h)^{n-1} e^{-\lambda t + \lambda h} T(t)x \, dt - \frac{1}{h(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} T(t)x \, dt \\
 &= \lim_{h \rightarrow 0} \frac{e^{\lambda h} - 1}{h(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} T(t)x \, dt - \frac{e^{\lambda h}}{h(n-1)!} \int_0^h t^{n-1} e^{-\lambda t} T(t)x \, dt \\
 &\quad - \frac{e^{\lambda h}}{(n-2)!} \int_h^\infty t^{n-2} e^{-\lambda t} T(t)x \, dt + o(h) \\
 &= -I_{n-1}(\lambda) + \lambda I_n(\lambda).
 \end{aligned}$$

Es folgt also:

$$AI_n(\lambda)x = -I_{n-1}(\lambda)x + \lambda I_n(\lambda)x, \quad x \in X, \quad n > 1$$

und

$$AI_1(\lambda)x = -x + \lambda I_1(\lambda)x, \quad x \in X.$$

Insgesamt erhalten wir $I_n(\lambda)(A - \lambda I) = (A - \lambda I)I_n(\lambda) = -I_{n-1}(\lambda)$ für $n > 1$ und $I_1(\lambda)(A - \lambda I) = (A - \lambda I)I_1(\lambda) = -I$. Aus der zweiten Aussage folgt die Existenz von $R_\lambda(A)$ und aus der ersten folgt $I_n(\lambda) = (-R_\lambda(A))^n$ und somit die behauptete Normabschätzung.

„ \Rightarrow “ Nun wird gezeigt, dass obige Bedingungen hinreichend sind. Hierzu konstruieren wir die Operatorhalbgruppe explizit. Sei also

$$U_n(t) = \left(I - \frac{t}{n}A\right)^{-n}$$

Aus der zweiten Bedingung folgt, dass $U_n(t)$ für $\frac{n}{t} > \omega$ wohldefiniert ist. Wir wollen jetzt zeigen, dass für $t \rightarrow 0$ $U_n(t) \rightarrow I$ in der Operatornorm konvergiert. Dafür reicht es zu zeigen, dass $(I - \frac{t}{n}A)^{-1}$ in der Operatornorm gegen I konvergiert. Sei zunächst $u \in D(A)$ dann gilt

$$\|(I - \frac{t}{n}A)^{-1}u - u\| = \|((I - \frac{t}{n}A)^{-1} - I)u\| = \frac{t}{n} \|(I - \frac{t}{n}A)^{-1}Au\| \leq Ct\|Au\|,$$

was für $t \rightarrow 0$ gleichmäßig in n gegen Null konvergiert. Sei nun $u \in X$, sei weiter

$u_k \in D(A)$ mit $\lim_{k \rightarrow \infty} u_k = u$ dann folgt:

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} \lim_{t \rightarrow 0} \left\| \left(I - \frac{t}{n} A \right)^{-1} u_k - u_k \right\| = \lim_{t \rightarrow 0} \lim_{k \rightarrow \infty} \left\| \left(I - \frac{t}{n} A \right)^{-1} u_k - u_k \right\| \\ &= \lim_{t \rightarrow 0} \left\| \left(I - \frac{t}{n} A \right)^{-1} u - u \right\| = 0. \end{aligned} \quad (\text{A.4})$$

Die beiden Grenzwerte dürfen vertauscht werden, da der Term gleichmäßig beschränkt in t und k ist. Denn:

$$\left\| \left(I - \frac{t}{n} A \right)^{-1} u_k - u_k \right\| \leq \frac{n}{t} \left\| \left(A - \frac{n}{t} I \right)^{-1} \right\| \|u_k\| + \|u_k\| \leq K \quad (\text{A.5})$$

wobei K unabhängig von t und k ist. Als nächstes zeigen wir, dass $U_n(t)$ für $n \rightarrow \infty$ in der Operatornorm konvergiert. Dafür betrachten wir wieder $U_n(t)u - U_m(t)u$ auf einer dichten Teilmenge und zwar $D(A^2) = (A - \lambda)^{-1}D(A)$ mit $\operatorname{Re}(\lambda) > \omega$. Es gilt $\frac{dU}{dt} = A \left(I - \frac{t}{n} A \right)^{-n-1}$. Dies folgt aus folgender Rechnung mit Hilfe der Kettenregel.

$$\begin{aligned} &\left\| \lim_{h \rightarrow 0} \frac{R_{\frac{t+h}{n}}(A) - R_{\frac{t}{n}}(A)}{h} - \frac{A}{n} \left(I - \frac{t}{n} A \right)^{-2} \right\| \\ &= \left\| \lim_{h \rightarrow 0} \frac{\left(R_{\frac{t+h}{n}} \left(I - \frac{t}{n} A \right)^2 - I + \frac{t}{n} A - A \frac{h}{n} \right) \left(I - \frac{t}{n} A \right)^{-2}}{h} \right\| \\ &= \left\| \lim_{h \rightarrow 0} \frac{\left(I - \frac{t+h}{n} A \right)^{-1} \left(\left(I - \frac{t}{n} A \right)^2 - \left(I - \frac{t+h}{n} A \right) \left(I - \frac{t}{n} A + Ah \right) \right) \left(I - \frac{t}{n} A \right)^{-2}}{h} \right\| \\ &= \left\| \lim_{h \rightarrow 0} \frac{\left(I - \frac{t+h}{n} A \right)^{-1} \frac{A^2 h^2}{n^2} \left(I - \frac{t}{n} A \right)^{-2}}{h} \right\| \\ &= 0 \end{aligned}$$

Die Norm dieses Terms geht für $h \rightarrow 0$ wegen der Abschätzung für $R_\lambda(A)$ gleichmäßig in n gegen Null und für alle $u \in D(A^2)$. Somit ist $U_n(t)$ differenzierbar mit obiger Ableitung.

$$\begin{aligned} U_n(t)u - U_m(t)u &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^{t-\epsilon} \frac{d}{ds} [U_m(t-s)U_n(s)u] ds \\ &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^{t-\epsilon} \left[-\frac{\partial U_m(t-s)}{\partial s} U_n(s)u + U_m(t-s) \frac{\partial U_n(s)}{\partial s} u \right] ds \\ &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^{t-\epsilon} \left[-A \left(I - \frac{t-s}{m} A \right)^{-m-1} \left(I - \frac{s}{n} A \right) \right. \\ &\quad \left. + \left(I - \frac{t-s}{m} A \right)^{-m-1} \left(I - \frac{t-s}{m} A \right) A \right] \left(I - \frac{s}{n} A \right)^{-n-1} u ds \\ &= \int_0^t \left(\frac{s}{n} - \frac{t-s}{m} \right) \left(I - \frac{t-s}{m} A \right)^{-m-1} \left(I - \frac{s}{n} A \right)^{-n-1} A^2 u ds. \end{aligned} \quad (\text{A.6})$$

Die letzte Abschätzung lässt sich durchführen, da $R_\lambda(A)$ und A kommutieren. Auf Grund

der Beschränktheit von $R_\lambda(A)$ lässt sich folgende Abschätzung finden:

$$\|U_n(t)u - U_m(t)u\| \leq C\|A^2u\| \int_0^t \frac{s}{n} + \frac{t-s}{m} ds = \frac{Ct^2}{2} \left(\frac{1}{n} + \frac{1}{m} \right) \|A^2u\|$$

Also existiert der Grenzwert in der Operatornorm für alle $u \in X$ mit analoger Argumentation zu (A.4) und (A.5) und dem Satz von Banach-Steinhaus. Wir bezeichnen diesen Grenzwert mit $T(t)$. Auf Grund von $\|U_n(t)\| \leq M(1 - \frac{t}{n}\omega)^{-n}$ folgt $\|T(t)\| \leq M \exp(\omega t)$. Es bleibt also noch zu zeigen, dass $\{T(t)\}_{t \geq 0}$ eine Halbgruppe bilden und dass A der infinitesimale Erzeuger ist. Für t aus einem endlichen Intervall ist die Konvergenz in (A.6) gleichmäßig in t . Aus der Stetigkeit von $U_n(t)u$ für alle n folgt die Stetigkeit von $T(t)u$ für alle $u \in D(A^2)$. Da $D(A^2)$ dicht und $T(t)$ für t aus einem endlichen Intervall beschränkt ist, folgt die Stetigkeit von $T(t)u$ für alle $u \in X$. $T(0) = I$ folgt daraus, dass $U_n(0) = I$ für alle n gilt. Aus (A.6) folgt mit $n = m$

$$\frac{d}{ds} [U_n(t-s)U_n(s)u] = \frac{2s-t}{n} \left(I - \frac{t-s}{n} A \right)^{-n-1} \left(I - \frac{s}{n} A \right)^{-n-1} A^2 u$$

für $u \in D(A^2)$ für $n \rightarrow \infty$ folgt, dass $T(t-s)T(s)u$ unabhängig von s für $u \in D(A^2)$ ist, da $D(A^2)$ dicht ist, folgt die Aussage auch für X . Daraus folgt die Halbgruppeneigenschaft, da ja $T(0) = I$ ist. Weiter gilt $\frac{d}{dt} T(t)u|_{t=0} = Au$ für alle $u \in D(A)$. Somit ist der infinitesimale Erzeuger von $\{T(t)\}_{t \geq 0}$ wegen (2.19) eine Erweiterung von A . Auf Grund der Wachstumsabschätzung für $T(t)$ und der bereits bewiesenen Notwendigkeit des Satzes, folgt dass die Resolvente des infinitesimalen Erzeugers für $\lambda > \omega$ existiert, diese muss aber mit der Resolvente von A übereinstimmen und somit stimmt A mit dem infinitesimalen Erzeuger überein. \square

Anhang B

Hilfsresultate

B.1 Bessel Funktionen

Wir folgen den Ausführungen in [3, S.289f]. Die Besselgleichung der Ordnung ν lautet für $\nu \in \mathbb{R}$:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - \nu^2)y = 0$$

für $x \in \mathbb{R}$. Die Lösungen dieser Differentialgleichung lassen sich folgendermaßen ausdrücken:

$$y(x) = C_1 J_\nu + C_2 J_{-\nu}(x),$$

wobei $\nu \notin \mathbb{Z}$, C_1 und C_2 beliebige Konstanten sind. Die Funktion J_ν ist die Besselfunktion erster Art gegeben durch:

$$J_\nu(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (x/2)^{\nu+2n}}{n! \Gamma(\nu + n + 1)}.$$

Die modifizierte Besselfunktion erster Ordnung lautet:

$$I_\nu(x) = e^{-\nu\pi/2} J_\nu(ix).$$

Nun können wir auch die modifizierte Besselfunktion der Ordnung 3 definieren. Diese ist gegeben durch :

$$K_\nu(x) = \left(\frac{\pi}{2}\right) \frac{I_{-\nu}(x) - I_\nu(x)}{\sin(\nu\pi)}.$$

Die modifizierte Besselfunktion dritter Art lässt sich auch folgendermaßen darstellen:

$$K_\nu(x) = \frac{1}{2} \int_0^\infty u^{\nu-1} \exp\left(-\frac{1}{2}x \left(u + \frac{1}{u}\right)\right) du.$$

Für $\nu = \frac{1}{2}$ lässt sich dies zu

$$K_{1/2}(x) = \int_0^\infty \exp\left(-\frac{1}{2}x\left(u^2 + \frac{1}{u^2}\right)\right) du$$

und für $\nu = 1$

$$K_1(x) = x \int_1^\infty e^{-xu} \sqrt{u^2 - 1} du$$

vereinfachen. Für Details siehe auch [1, Kapitel 9.6].

B.2 Hilfssätze der Stochastik

Satz B.1. Sei X_t ein OU-Prozess mit

$$dX_t = -\alpha X_t dt + \sigma dW_t$$

und $f: [0, \infty) \rightarrow \mathbb{R}$ mit $f \in C^2([0, \infty))$ dann folgt für $Y_t = f(t) + X_t$

$$dY_t = \alpha(\mu(t) - Y_t)dt + \sigma dW_t \quad \text{mit} \quad \mu(t) = f(t) + \frac{1}{\alpha}f'(t)$$

Beweis. Der Beweis folgt sofort aus Itos-Lemma. Denn es gilt nach (2.2)

$$dY_t = f'(t)dt + dX_t = \alpha\left(\frac{f'(t)}{\alpha} + f(t) - f(t) - X_t\right)dt + \sigma dW_t$$

und das war gerade zu zeigen. □

Satz B.2 (Herleitung der Verteilung von X_t). Wir wissen, dass der Prozess X_t folgender SDE genügt:

$$dX_t = -\alpha X_t dt + \sigma dW_t$$

Falls wir den Prozess X'_t definieren als

$$dX'_t = e^{\alpha t} X_t$$

dann folgt für dX'_t mit Itos-Lemma:

$$dX'_t = e^{\alpha t}(\sigma dW_t).$$

Diese lässt sich einfach durch Integration lösen, also

$$X'_t - X'_0 = \sigma \int_0^t e^{\alpha s} dW_s.$$

Daraus folgt aber für unseren Diffusionsprozess

$$X_t = X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s. \quad (\text{B.1})$$

Somit ist X_t für ein festes $x_0 = X_0$ normalverteilt mit

$$X_t \sim N(X_0 e^{-\alpha t}, \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha t}))$$

Dies folgt aus der Verteilung für X_t , da $\int_0^t f(s) dW_s$ ein Martingal ist und für das zweite Moment folgende Gleichung gilt (Vgl. [14] S. 258 f.)

$$E[(\int_0^t f(s) dW_s)^2] = E[\int_0^t f(s)^2 ds].$$

Also ist die bedingte Verteilung der Zufallsvariable X_t , falls der Stand zum Zeitpunkt $s < t$ x_s ist, gegeben durch:

$$X_t \sim N(x_s e^{-\alpha(t-s)}, \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha(t-s)})) \quad (\text{B.2})$$

Satz B.3 (Momenteerzeugende Funktion von Y_t). [22, Lemma 3.4.2] Seien $\{J_1, J_2, \dots\}$ unabhängig und identisch verteilte Zufallsvariablen mit momenterzeugender Funktion $\Phi_J(\theta) = \mathbb{E}e^{\theta J}$, die auf einer Teilmenge $\Theta \subset \mathbb{C}$, die die Null enthält definiert ist. Seien weiter $\{\tau_1, \tau_2, \dots\}$ die zufälligen Sprungzeitpunkte eines Poissonprozesses N_t mit Intensität λ , dann hat der Spikeprozess mit Anfangsbedingung $Y_0 = 0$ und

$$Y_t = \sum_{i=1}^{N_t} e^{-\beta(t-\tau_i)} J_i$$

folgende momenteerzeugende Funktion

$$\Phi_Y(\theta, t) = \exp(\lambda \int_0^t \Phi_J(\theta e^{-\beta s}) - 1 ds), \quad \text{für } \forall \theta \in \Theta.$$

Beweis. Wir beweisen den Satz, indem wir zunächst die momenteerzeugende Funktion nach dem ersten Sprung betrachten und diese bedingte Erwartung auf eine GDL für $\Phi_{Y_t}(\theta, t)$ zurückführen.

Sei also $\tau_1 = s$, dann gilt

$$\mathbb{E}[e^{\theta Y_t} | \tau_1 = s] = \mathbb{E} \exp(\theta e^{-\beta(t-s)} J_s) \mathbb{E} \exp(\theta \sum_{i=2}^{N_t} e^{-\beta(t-\tau_i)} J_{\tau_i} | \tau_1 = s).$$

Dies gilt, da die Zeitpunkte der Sprünge τ_i , sowie J_s und J_t (für $s \neq t$) unabhängig sind

und somit die Erwartungswerte getrennt werden dürfen. Auf Grund der Homogenität der Inkremente des Poissonprozesses kann die bedingte Erwartung als unbedingte Erwartung geschrieben werden, wobei die Summe bei $i = 1$ beginnt und der Endzeitpunkt auf $t - s$ gesetzt wird. Es folgt

$$\mathbb{E}[e^{\theta Y_t} | \tau_1 = s] = \mathbb{E}[\exp(\theta e^{-\beta(t-s)} J)] \Phi_Y(\theta, t - s) = \Phi_J(\theta e^{-\beta(t-s)}) \Phi_Y(\theta, t - s). \quad (\text{B.3})$$

Nun können wir die unbedingte Verteilung bestimmen. Wir wissen, dass $\tau_1 \sim \text{Exp}(\lambda)$ ist also gilt

$$\Phi_Y(\theta, t) = \mathbb{E}[\mathbb{E}[e^{\theta Y_t} | \tau_1]] = \int_0^t \mathbb{E}[e^{\theta Y_t} | \tau_1 = s] \lambda e^{-\lambda s} ds.$$

Mit (B.3) folgt

$$\Phi_Y(\theta, t) = \int_0^t \Phi_J(\theta e^{-\beta(t-s)}) \Phi_Y(\theta, t-s) \lambda e^{-\lambda s} ds \stackrel{\text{Sub.}}{=} \int_0^t \Phi_J(\theta e^{-\beta s}) \Phi_Y(\theta, s) \lambda e^{-\lambda(t-s)} ds \quad (\text{B.4})$$

Falls wir diese Integralgleichung nach t differenzieren, folgt mit Kettenregel:

$$\begin{aligned} \frac{\partial \Phi_Y}{\partial t}(\theta, t) &= \Phi_J(\theta e^{-\beta t}) \Phi_Y(\theta, t) \lambda - \lambda \int_0^t \Phi_J(\theta e^{-\beta s}) \Phi_Y(\theta, s) \lambda e^{-\lambda(t-s)} ds \\ &\stackrel{(\text{B.4})}{=} \lambda \int_0^t (\Phi_J(\theta e^{-\beta s}) - 1) \Phi_Y(\theta, s) ds. \end{aligned}$$

Es folgt, da $Y_0 = 0$ gilt $\Phi_Y(\theta, 0) = 1$ und somit

$$\Phi_Y(\theta, t) = \exp\left(\lambda \int_0^t (\Phi_J(\theta e^{-\beta s}) - 1) ds\right) \quad (\text{B.5})$$

□

Satz B.4. Sei $X \sim N(\mu, \sigma^2)$ dann gilt $\mathbb{E}[e^{\theta X}] = \exp(\theta \mu + \frac{1}{2} \theta^2 \sigma^2)$.

Beweis.

$$\begin{aligned} \mathbb{E}[e^{\theta X}] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{\theta x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = e^{m\theta + \frac{1}{2}\theta^2\sigma^2} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2\theta))^2}{2\sigma^2}} dx \\ &= e^{m\theta + \frac{1}{2}\theta^2\sigma^2} \end{aligned}$$

□

B.3 Lucia-Schwartz Modell

Der Strompreis lässt sich durch folgende stochastische Differentialgleichung beschreiben:

$$S_t = \exp(f(t) + X_t),$$

$$dX_t = -\alpha X_t + \sigma dW_t.$$

Es gilt für den natürlichen Logarithmus des Preises:

$$\ln S_T = f(T) + X_T$$

Dann gilt nach (B.1):

$$\ln S_T = f(T) + X_t e^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(t-s)} dW_s.$$

Oder mit $X_t = \ln S_t - f(t)$

$$\ln S_T = f(T) + (\ln S_t - f(t))e^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(t-s)} dW_s.$$

Es ist also $\ln S_T$ normalverteilt mit Erwartungswert $m(T) = f(T) + (\ln S_t - f(t))e^{-\alpha(T-t)}$ und Varianz $v(T) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(T-t)})$. Es lassen sich nun einfach mit Hilfe von Satz 6.2 sowohl der Call als auch der Forward bewerten. Sei der Einfachheit halber $s(t) = e^{f(t)}$ dann folgt

$$F_t^T = \mathbb{E}[S_T | S_t] = \exp(m(T) + \frac{1}{2}v(T)) = s(T) \left(\frac{S_t}{s(t)}\right)^{e^{-\alpha(T-t)}} \exp\left(\frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha(T-t)})\right).$$

Für den Call gilt:

$$\begin{aligned} C(t, S_t, T, K) &= e^{-r(T-t)} (\mathbb{E}[(S_T - K)^+ | S_t]) \\ &= e^{-r(T-t)} (\mathbb{E}[S_T \mathbf{1}_{S_T \geq K} | S_t] - K \mathbb{E}[\mathbf{1}_{S_T \geq K} | S_t]) \\ &= e^{-r(T-t)} (\mathbb{E}[S_T | S_t] N(d_1) - K N(d_2)) \\ &= e^{-r(T-t)} (s(T) \left(\frac{S_t}{s(t)}\right)^{e^{-\alpha(T-t)}} \exp\left(\frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha(T-t)})\right) N(d_1) - K N(d_2)) \\ \text{mit } d_1 &= \frac{-\ln K + m(T) + v(T)}{\sqrt{v(T)}} \quad \text{und} \quad d_2 = \frac{-\ln K + m(T)}{\sqrt{v(T)}}. \end{aligned}$$

Die Gleichungen für d_1 und d_2 leiten sich wie folgt her. Der erste Erwartungswert in der obigen Gleichung lässt sich mit einem Lemma aus der Wahrscheinlichkeitstheorie vereinfachen.

Lemma B.1. Sei $\ln X \sim N(\mu, \sigma^2)$ dann gilt

$$\mathbb{E}[X \mathbf{1}_{X \geq K}] = \mathbb{E}[X] N\left(\frac{-\ln K + \mu + \sigma^2}{\sigma}\right).$$

Wir können uns also der Indikatorfunktion im Erwartungswert entledigen und erhalten

sofort die Gleichung für d_1 . Die Gleichung für d_2 erhalten wir wie folgt:

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{S_T \geq K} | S_t] &= \mathbb{E}[\mathbf{1}_{\ln S_T \geq \ln K} | S_t] \stackrel{\text{Symmetrie}}{=} \mathbb{E}[\mathbf{1}_{\ln S_T \leq -\ln K + 2m(T)} | S_t] \\ &= N\left(\frac{-\ln K + 2m(T) - m(T)}{\sqrt{v(T)}}\right).\end{aligned}$$

Aus der Herleitung wird klar, dass für große Restlaufzeiten (großes $T - t$) der Einfluss des aktuellen Preises sehr gering ist. Betrachten wir hierfür die einzelnen Terme, die in den Preis eingehen:

$$m(T) \sim f(T) \quad \text{und} \quad v(T) \sim \frac{\sigma^2}{2\alpha}.$$

Also folgt für den Forward:

$$F_t^T \sim s(T) \exp\left(\frac{\sigma^2}{4\alpha}\right).$$

Für den europäischen Call lässt sich eine ähnliche Abschätzung herleiten. Mit

$$d_1 \sim \frac{-\ln K + f(T) + \frac{\sigma^2}{2\alpha}}{\sqrt{\frac{\sigma^2}{2\alpha}}} = \tilde{d}_1 \quad \text{und} \quad d_2 \sim \frac{-\ln K + f(T)}{\sqrt{\frac{\sigma^2}{2\alpha}}} = \tilde{d}_2$$

folgt

$$C(t, S_t, T, K) = e^{-r(T-t)} (s(T) \exp\left(\frac{\sigma^2}{4\alpha}\right) N(\tilde{d}_1) - KN(\tilde{d}_2)).$$

Somit wäre der Callpreis im Grenzfall unabhängig von dem heutigen Kurs (S_t).

B.4 Herleitung der Randbedingungen

Es wird angenommen, dass die Funktion $V(x, y, t)$ mit

$$V(\tilde{x}, y, t) = V(e^x, y, t) = v(x, y, t), \tag{B.6}$$

sich für große und kleine \tilde{x} linear in \tilde{x} verhält. Also gilt

$$\frac{\partial^2 V}{\partial \tilde{x}^2} = 0$$

Wird nun die Ableitungen von v berechnet und benutzt dass (B.6) gilt, so folgt

$$\begin{aligned}\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x} &= e^x \frac{\partial V(\tilde{x}, y, t), t}{\partial \tilde{x}} + e^x \frac{\partial^2 V(\tilde{x}, y, t)}{\partial \tilde{x}^2} - e^x \frac{\partial V(\tilde{x}, y, t)}{\partial \tilde{x}} \\ &= 0\end{aligned}$$

Literaturverzeichnis

- [1] M. Abramowitz, I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, Applied Mathematics Series 55, Washington D.C., 1972.
- [2] H. Amann. *Linear and Quasilinear Parabolic Problems*, Volume I, Birkhäuser, Berlin, 1995.
- [3] D. Applebaum *Lévy Processes and Stochastic Calculus*, Cambridge University Press, New York, 2004.
- [4] O.E. Barndorff-Nielsen *Processes of Normal Inverse Gaussian Type*, Finance and Stochastics (1998), Vol. 2, No. 1, **41-68**.
- [5] O.E. Barndorff-Nielsen, S.Z. Levendorskii *Feller Processes of Normal Inverse Gaussian Type* (2001), Quantitative Finance, Vol.1, No. 3, **318-331**.
- [6] O.E. Barndorff-Nielsen, S.Z. Levendorskii *Feller Processes of Normal Inverse Gaussian Type*, Working Paper,
Preprint: <http://www.cls.dk/caf/wp/wp-76.pdf>.
- [7] A. Bellini-Morante *Applied semigroups and evolution equations*, Oxford University Press, London, 1978.
- [8] F.E. Benth, J. Saltyte-Benth *The Normal Inverse Gaussian Distribution and Spot Price Modelling in Energy Markets* (2004), International Journal of Theoretical and Applied Finance, Vol. 7, No. 2, **177-192**.
- [9] B. Bibby, M.Sorensen *Martingale estimation functions for discretely observed diffusion processes* (1995), Bernoulli, Vol.1, No. 1/2, **17-39**.
- [10] N. Branger, C. Schlag *Zinsderivate*, Springer, Heidelberg, 2004.
- [11] S.I. Boyarchenko, S.Z. Levendorskii *Non-Gaussian Merton-Black-Scholes Theory*, World Scientific, London, 2002.

- [12] S.I. Boyarchenko, S.Z. Levendorskii *Barrier options and touch-and-out options under regular Lévy processes of exponential type* (2002), The Annals of Applied Probability, Vol. 12, No. 4, **1261-1298**.
- [13] M. Briani, C. La Chioma, R. Natalini *Convergence of numerical Schemes for Viscosity Solutions to Integro-Differential degenerate parabolic problems arising in financial theory*, Numerische Mathematik, No.1 Vol. 98, **607-646**.
- [14] R. Cont, P. Tankov *Financial Modelling with jump processes*, Chapman & Hall, New York, 2004.
- [15] R. Cont, E. Voltchkova *Finite Difference Methods for option pricing in Jump diffusion and exponential levy Models* (2003), Rapport Interne 513, CMAP, Ecole Polytechnique,
Preprint: <http://www.cmap.polytechnique.fr/~rama/papers/eccomas1.ps>.
- [16] R. Cont, E. Voltchkova *A Finite Difference Methods for option pricing in Jump diffusion and exponential levy Models* (2005), SIAM Journal of Numerical Analysis, Vol. 43, No. 4, **1596-1626**.
- [17] P. J. Huber *Robust Statistics*, John Wiley & Sons, Hoboken, 2004.
- [18] L. Hörmander *The Analysis of Linear Partial Differential Operators III*, Springer, Berlin, 1985.
- [19] N. Jacob *Pseudo Differential Operators and Markov Processes Volume I Fourier Analysis and Semigroups*, Imperial College Press, Berlin, 2001.
- [20] M. Jansen und M. Wobben *Preisbildung auf den Märkten für Elektrizität und Erdgas - Ein Blick auf 2007 und die kommenden Jahre* (2008), erscheint in Energiepolitische Tagesfragen, Nr.10.
- [21] M. Kjaer *Pricing of some Path-Dependent Options on Equities and Commodities* (2006), Dissertation, Univesity of Göteborg
Preprint: http://gupea.ub.gu.se/dspace/bibstream/2077/2911/1/Kjaer_avh_fulltext.pdf.
- [22] T. Kluge *Pricing Swing Options and other Electricity Derivatives* (2006), Dissertation, University of Oxford,
Preprint: <http://eprints.maths.ox.ac.uk/246/01/kluge.pdf>.
- [23] A. L. Lewis *A simple Option Formula for General Jump-Diffusion and other exponential Levy Processes*, Working Paper,
Preprint: <http://optioncity.net/pubs/ExpLevy.pdf>

- [24] J. Lucia, E. Schwartz *Electricity Prices and Power Derivates: Evidence from the Nordic Power Exchange* (2002), Review of Derivates Research, No.1, Vol. 5, **5-50**.
- [25] L. Pan *The jump-risk premia implicit in options evidence from an integrated time-series study* (2002), The Journal of Financial Economics, Vol.63, **3-50**.
- [26] A. Pazy *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, Berlin, 1983.
- [27] P.E. Protter *Stochastic Integration and Differential Equations*, Springer, Berlin, 2005.
- [28] A. Quarteroni, R. Sacco, F. Saleri *Numerical Mathematics*, Springer, Berlin, 2000.
- [29] M. Renardy, R.C. Rogers *A Introduction to Partial Differential Equations*, Springer, New York, 2004.
- [30] K.I. Sato *Lévy Processes and infinitely divisible Distributions*, University Press, Cambridge, 1999.
- [31] S. E. Shreve *Stochastic Calculus for Finance II*, Springer, Berlin, 2005.

Abschlussklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und nur mit den angegebenen Quellen und Hilfsmitteln erstellt habe.

Münster, 29.08.2008

Oleg Reichmann