Technische Universität München

Zentrum Mathematik

Dynamische Portfoliooptimierung

Diplomarbeit

von

Ulrich Mauthner

Themenstellerin:	Prof. Dr. C. Klüppelberg
Betreuerin:	Prof. Dr. C. Klüppelberg
Abgabetermin:	15. Juni 2003

Hiermit erkläre ich, daß ich die Diplomarbeit selbständig angefertigt und nur die angegebenen Quellen verwendet habe. München, 15. Juni 2003

Danksagung

Bedanken möchte ich mich bei allen, die zum Gelingen dieser Diplomarbeit beigetragen haben:

Meiner Themenstellerin Frau Professor Klüppelberg, die den Fortgang der Arbeit mit hilfreichen und motivierenden Ratschlägen begleitet hat,

den Mitarbeitern am Lehrstuhl M4 für das allgemein gute Verhältnis untereinander

meiner Familie für die Unterstützung und das Verständnis für mich während der letzten Monate

und meinen Komilitonen für ihre Freundschaft während der Studienzeit und hoffentlich auch darüber hinaus

Inhaltsverzeichnis

1 Stochastische Prozesse **2**

1.1 Grundlagen 2

1.1.1 Der stochastische Prozess 2

1.1.2 Die Filtration 4

1.1.3 Martingale 5

1.1.4 Stoppzeiten 9

1.2 Lévy-Prozesse 11

1.2.1 Beispiele 12

1.2.2 Strukturresultate 13

1.3 Markov-Prozesse und Diffusionen 14

1.3.1 Markov-Prozesse 14

1.3.2 Diffusionen und Itô-Diffusionen 17

1.4 Semimartingale 18

1.4.1 Prozesse mit endlicher Variation 18

1.4.2 Lokalisierung 19

1.4.3 Semimartingale- Definition und Beispiele 21

2 Stochastische Integration **23**

2.1 Allgemeines zur Integration 23

2.1.1 Überlegungen aus der Funktionalanalysis 23

2.1.2 Probleme bei der Integration stochastischer Prozesse . 26

2.2 Das stochastische Integral 28

2.2.1 Die ucp-Konvergenz 28

2.2.2 Definition für einfache vorhersehbare Prozesse als Integranden 29

2.2.3 Erweiterung auf Integranden in $\mathbb{L}$ 30

2.3 Eigenschaften und Anwendungen des stochastischen Integrals . 31

2.3.1 Die quadratische Variation eines Semimartingals 32

2.3.2 Die Itô-Formel 33
2.3.3 Integrale mit speziellen Integratoren 34
2.3.4 Stochastische Differentialgleichungen 34

3 Modelle in der Portfoliooptimierung 36
3.1 Das allgemeine Modell 36
3.1.1 Der Preisprozeß 36
3.1.2 Der Handelsprozeß φ und abgeleitete Größen 37
3.1.3 Das Bewertungsfunktional 41
3.1.4 Zusammenfassung des allgemeinen Modells 42
3.2 Beispiele 42
3.2.1 Das klassische Modell von Black/Scholes/Merton ohne Transaktionskosten 43
3.2.2 Das Modell von Black/Scholes/Merton mit Transaktionskosten 44
3.2.3 Ein Modell von Kallsen [8] 45

4 Stochastische Kontrolltheorie und die HJB-Gleichung 47
4.1 Stochastische Kontrolltheorie 47
4.1.1 Die Problemstellung 47
4.1.2 Die Lösungsidee 49
4.1.3 Die HJB-Gleichung und ihre klassischen Lösungen . . . 52
4.1.4 Schwache Lösungen (Viscosity Solutions) 54
4.2 Kontrolltheorie für Lévy-Prozesse 57
4.2.1 Berechnung von $\sum_{i=1}^{n+1} \int_t^{t+h} \partial_i \Phi(y_{s-}^v) dy_s^{v,i}$ 58
4.2.2 Berechnung von $\frac{1}{2} \sum_{i,j=1}^{n+1} \int_t^{t+h} \partial_{i,j} \Phi(y_{s-}^v) d[y^{v,i}, y^{v,j}]$ . . . 59
4.2.3 Berechnung von $\sum_{t<s\leq t+h} \{\Phi(y_s) - \Phi(y_{s-}) - \Phi'(y_{s-})\Delta y_s\}$ 60
4.2.4 Berechnung von $A\Phi$ 60
4.2.5 Berechnung der verbleibenden Sprungterme 62
4.2.6 Zusammenfassung 64
4.3 Anwendung auf die Portfoliooptimierung 65
4.3.1 Beispiel 1: Black-Scholes-Merton ohne Transaktionskosten 65
4.3.2 Beispiel 2: Black-Scholes-Merton mit Transaktionskosten 68
4.3.3 Beispiel 3: Das Modell von Kallsen 71
4.4 Bewertung der Kontrolltheorie 72

5 Die Martingalmethode **73**
5.1 Spezifizierung des allgemeinen Modells aus 3.1.4 74
5.1.1 Der Preisprozeß S 74
5.1.2 Die zulässigen Handelsprozesse φ 74
5.1.3 Das Bewertungsfunktional 74
5.2 Dualität, Vollständigkeit und Arbitragefreiheit 75
5.2.1 Das Dualitätsprinzip 75
5.2.2 Vollständigkeit und Arbitragefreiheit 85
5.3 Die Martingalmethode mit Lévy-Prozessen als Preisprozesse . 87
5.4 Anwendung auf die Portfoliooptimierung 91
5.4.1 Beispiel 1: Black-Scholes-Merton ohne Transaktionskosten 91
5.4.2 Beispiel 2: Black-Scholes-Merton mit Transaktionskosten 93
5.4.3 Beispiel 3: Das Modell von Kallsen 93
5.5 Bewertung der Martingalmethode 96

Einleitung

Eines der bekanntesten Probleme in der Finanzmathematik ist das Finden einer optimalen Investmentstrategie für einen Investor, dem verschiedene risikolose und risikobehaftete Anlagemöglichkeiten zur Verfügung stehen. Das Problem ist -so formuliert- recht allgemeiner Natur, und daher verwundert es nicht, daß eine Vielzahl verschiedener mathematischer Modelle benutzt werden, um es darzustellen. Je nachdem, welche Umstände des realen Marktes in die Modellierung einbezogen werden sollen (und auch je nachdem, welche Modelle vom mathematischen Standpunkt aus lösbar sind), werden Investitionsentscheidungen und Aktienkurse nur zu diskreten oder kontinuierlichen Zeitpunkten betrachtet, wird gehandelt oder das Anfangsportfolio bis zum Schluß beibehalten usw.

Es gibt aber auch eine Reihe von Gemeinsamkeiten in den Modellen: Es werden z.B. immer stochastische Prozesse zur Modellierung der Aktienkurse und auch der Investitionen verwendet und es wird ebenfalls immer eine Informationsstruktur durch eine Filtration von σ-Algebren (Stichwort:„Handelsentscheidungen nur aufgrund vergangener und gegenwärtiger Kurse") benutzt.

Eine erschöpfende Behandlung des Themas Portfoliooptimierung ist in einer Diplomarbeit sicher nicht möglich. Ziel dieser Arbeit ist daher, einen Überblick über die in der einschlägigen Literatur verwendeten Modelle zu geben sowie die verschiedenen Lösungsmethoden darzustellen, die ihnen zugrundeliegenden Ideen zu erläutern und sie schließlich auch anhand von Beispielen anzuwenden. Diese Arbeit beschränkt sich dabei auf den Fall eines zeitstetigen Aktienmarktes mit endlich vielen verschiedenen Anlagemöglichkeiten.

Der Aufbau der Arbeit ist folgender: Im ersten und zweiten Kapitel werden die mathematischen Grundlagen, d.h. die Theorie der stochastischen Prozesse und der stochastischen Integration dargestellt und auch auf die ent-

sprechende Literatur hingewiesen. Im folgenden dritten Kapitel wird ein allgemeines Modell der zeitstetigen Portfoliooptimierung aufgestellt, auf das die im weiteren vorgestellten Modelle zurückgeführt werden. Im vierten und fünften Kapitel werden dann die beiden bekannten Lösungsmethoden, die stochastische Kontrolltheorie und die sogenannte Martingalmethode behandelt und anhand von Beispielen erläutert.

Kapitel 1

Stochastische Prozesse

In diesem Kapitel sollen grundlegende Eigenschaften von stochastischen Prozessen dargestellt werden. In ihm wie auch in den folgenden Kapiteln werden elementare Begriffe der Wahrscheinlichkeitstheorie (z.B. σ -Algebren, fast sichere oder stochastische Konvergenz u.ä.), wie sie z.B. in Bauer [3] dargestellt werden, als bekannt vorausgesetzt.

Der Aufbau der Kapitel 1 und 2 folgt im Wesentlichen Protter [17], Beweise werden, falls sie zu lang oder zu kompliziert für eine Wiedergabe erscheinen, skizziert bzw. ausgelassen (in diesem Fall wird dann auf entsprechende Stellen in der Literatur verwiesen).

1.1 Grundlagen

1.1.1 Der stochastische Prozess

Wir gehen aus von einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, der vollständig ist, d.h. es gilt $B \subset A \in \mathcal{F}, P(A) = 0 \implies B \in \mathcal{F}$. Weiterhin existiere eine vollständig geordnete Indexmenge I (in unserem Fall ist das immer $\mathbb{R}_0^+$, $\mathbb{N}_0$ oder ein Intervall $[t_0, t_1] \subset \mathbb{R}_0^+$), deren Elemente im folgenden mit t bezeichnet werden (in Anlehnung an die Interpretation der Indexmenge als stetige oder diskrete Zeitachse).

Definition 1.1.1 *Eine Abbildung $X : \Omega \times I \longrightarrow \mathbb{R}^d$ heißt ($\mathbb{R}^d$-wertiger) stochastischer Prozess (kurz: Prozess), falls $X(., t)$ für alle $t \in I$ messbar bezüglich $\mathcal{F}$ ist. Die Funktionen $X(\omega, .)$ heißen Pfade des stochastischen Prozesses. Ein Prozess X mit $I = \mathbb{R}_0^+$ heißt messbar, falls X als Abbildung*

von $\Omega \times \mathbb{R}_0^+$ *nach* $\mathbb{R}^d$ $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_0^+)$*-*$\mathcal{B}(\mathbb{R}^d)$*-messbar ist, eine analoge Definition gilt bei* $I = [t_0, t_1]$.

Je nach Zusammenhang verwendet man auch kurz die Schreibweise X_t , um die Zufallsvariable $X(., t)$ zu bezeichnen. Den Prozess kann man auch als durch die Elemente von I indizierte Kollektion von Zufallsvariablen auffassen und bezeichnet ihn dementsprechend mit $(X_t)_{t \in I}$.

Definition 1.1.2 *Zwei Prozesse* X *und* Y *heißen ununterscheidbar, falls* $P(\{\omega | X(\omega, .) = Y(\omega, .)\}) = 1$. X *heißt Modifikation von* Y*, falls* $P(\{\omega | X(\omega, t) = Y(\omega, t)\}) = 1$ *für alle* $t \in I$.

Wie aus der Definition unmittelbar ersichtlich ist, folgt aus der Ununterscheidbarkeit zweier Prozesse, daß sie Modifikationen voneinander sind.

Die umgekehrte Folgerung gilt aber, falls $I = [0, \infty)$ oder $I = [t_0, t_1]$, im Allgemeinen nicht, denn $\{\omega | X(\omega, .) \neq Y(\omega, .)\} = \bigcup_{t \in I} \{\omega | X(\omega, t) \neq Y(\omega, t)\})$ ist als Vereinigung von überabzählbar vielen Nullmengen selbst nicht unbedingt eine Nullmenge. Allerdings ist unter bestimmten Regularitätsvoraussetzungen auch die umgekehrte Folgerung gültig. Hierzu beschränkt man sich auf die Betrachtung von (rechts- oder linksseitig) stetigen Prozessen:

Definition 1.1.3 *Ein Prozess* X *heißt* P *-f.s. stetig (rechtsseitig stetig, linksseitig stetig), falls*

$$P(\{\omega | X(\omega, .)(\text{linksseitig, rechtsseitig}) \text{ stetig auf ganz } I\}) = 1$$

Man sagt, er besitzt P *-f.s. linksseitige (rechtsseitige) Grenzwerte, falls*

$$P(\{\omega | \lim_{\substack{s \to t \\ s<t \quad (s>t)}} X(\omega, s) \text{ existiert } \forall t \in I\}) = 1$$

Definition 1.1.4 *Ein Prozess* X *heißt càdlàg (**c**ontinue **à** **d**roite, **l**imites **à** **g**auche), falls er* P*-f.s. rechtsseitig stetig ist und linksseitige Grenzwerte besitzt. Entsprechend heißt er càglàd, falls er* P*-f.s. linksseitig stetig ist und rechtsseitige Grenzwerte besitzt. Die Menge der càdlàg-(càglàd-)Prozesse wird auch mit* $\mathbb{D}$ *(*$\mathbb{L}$*) bezeichnet. Für càdlàg-(bzw. càglàd-)Prozesse definiert man* X_- *(*X_+*) als* $X_-(\omega, t) := \lim_{\substack{s \to t \\ s<t}} X(\omega, s)$ *(*$X_+(\omega, t) := \lim_{\substack{s \to t \\ s>t}} X(\omega, s)$*) sowie* $\Delta X := X - X_-$ *(*$\Delta X := X_+ - X_-$*).*

Theorem 1.1 *Seien X und Y zwei rechts- oder linksstetige Prozesse, $I = [0,\infty)$ oder $I = [t_0, t_1]$, und sei X eine Modifikation von Y . Dann sind X und Y ununterscheidbar.*
Beweis: *Wir nehmen an, daß X und Y rechtsstetig sind, der linksstetige Fall ist analog. Sei $A := \{\omega | X(\omega, .)$ besitzt rechtsseitige Unstetigkeitsstellen $\}$, $B := \{\omega | Y(\omega, .)$ besitzt rechtsseitige Unstetigkeitsstellen $\}$. Sei weiter $N := \bigcup_{t \in \mathbb{Q}_0^+ \cap I} \{\omega | X(\omega, t) \neq Y(\omega, t)\}$. Dann gilt nach Voraussetzung $P(N) = P(A) = P(B) = 0$, also $P(N \cup A \cup B) = 0$. Aus Stetigkeitsgründen gilt aber auch für $t \notin \mathbb{Q}_0^+ \cap I$ $\{\omega | X(\omega, t) \neq Y(\omega, t)\} \cap \complement A \cap \complement B \subseteq N$, also mit $N' := \bigcup_{t \in \mathbb{I}} \{\omega | X(\omega, t) \neq Y(\omega, t)\}$ auch $N' \cap \complement A \cap \complement B \subseteq N$ und schließlich $P(N') = P(N' \cap \complement A \cap \complement B) + P(N' \cap (A \cup B)) \leq P(N) + P(A \cup B) = 0$.* ■

Insbesondere sind also càdlàg- (bzw. càglàd-) Prozesse genau dann ununterscheidbar, wenn sie Modifikationen voneinander sind. Eine ähnlicher Zusammenhang (mit der gleichen Beweistechnik) besteht auch zwischen Prozessen und messbaren Prozessen: Eine $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_0^+)$-$\mathcal{B}(\mathbb{R}^d)$-messbare Abbildung von $\Omega \times \mathbb{R}_0^+$ (d.h. ein messbarer Prozess) nach $\mathbb{R}^d$ ist immer auch ein Prozess, umgekehrt ist ein Prozess, der càdlàg oder càglàd ist, ein messbarer Prozess (analog für $I = [t_0, t_1]$).

1.1.2 Die Filtration

Um sinnvolle Modelle für die Portfoliooptimierung aufstellen zu können, ist es notwendig, den Wahrscheinlichkeitsraum mit einer (zeitlich geordneten) „Informationsstruktur" zu versehen, denn die Entscheidungen über Käufe oder Verkäufe dürfen nur vom Marktverlauf in der Vergangenheit abhängen. Man bedient sich dabei der Tatsache, daß man eine σ -Algebra auch als eine (i.A. unvollständige) Information auffassen kann, die man zur Verfügung hat: Man kennt nicht das konkrete $\omega \in \Omega$, aber man weiß zumindest für jede Menge $A \in \mathcal{F}$, ob $\omega \in A$ oder $\omega \notin A$. Interpretiert man σ -Algebren in diesem Sinne, so gilt $\mathcal{F} \subset \mathcal{G} \implies \mathcal{F}$ enthält weniger Information als $\mathcal{G}$. Diese Idee führt zur folgenden

Definition 1.1.5 *Eine Menge $(\mathcal{F}_t)_{t \in I}$ von Teil-σ-Algebren von $\mathcal{F}$ heißt Filtration, falls $(t \leq s) \implies \mathcal{F}_t \subseteq \mathcal{F}_s$ gilt. Eine Filtration heißt rechtsstetig, falls $\bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$ $\forall t \in I$. Ein Prozess X heißt adaptiert (an $(\mathcal{F}_t)_{t \in I}$), falls $X(., t)$ $\mathcal{F}_t$ -messbar ist $\forall t$.*

Im Folgenden gehen wir (im Fall $I = [0, \infty)$, $\mathbb{N}_0$ und $[t_0, t_1]$) außerdem davon aus, daß $\mathcal{F}_0$ bzw. $\mathcal{F}_{t_0}$(und damit alle $\mathcal{F}_t$, $t \geq 0$ bzw. $t \geq t_0$) vollständig ist, d.h. alle Teilmengen von P-Nullmengen enthält. Hat man einen Prozess X gegeben, so kann man in naheliegender Weise eine dazu passende Filtration definieren:

Definition 1.1.6 *Die Filtration $(\mathcal{G}_t)_{t \in I}$ definiert durch $\mathcal{G}_t := \sigma(X_s, s \leq t)$ heißt die vom Prozess X erzeugte Filtration; sie ist offenbar die (bezüglich der Teilmengenrelation) kleinste Filtration, bezüglich der X adaptiert ist.*

Die Klasse der adaptierten càdlàg-(bzw. càglàd-)Prozesse bezeichnen wir mit $\mathbb{D}([0, \infty), \mathbb{R}^d)$ (bzw. $\mathbb{L}([0, \infty), \mathbb{R}^d)$) oder auch, falls die Dimension aus dem Zusammenhang klar ist, als $\mathbb{D}$ (bzw. $\mathbb{L}$). Eine weitere Verschärfung des Begriffs der Messbarkeit bietet folgende

Definition 1.1.7 *Ein Prozess X mit Indexmenge $[0, \infty)$ heißt progressiv messbar, falls für alle t $X|_{[0,t]}$ als Abbildung von $[0, t] \times \Omega$ nach $\mathbb{R}^d$ messbar ist bezüglich $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ und $\mathcal{B}(\mathbb{R}^d)$. Eine analoge Definition existiert für $I = [t_0, t_1]$*

Ein progressiv messbarer Prozess X mit $I = [0, \infty)$ ist offenbar adaptiert und er ist auch messbar wegen

$$X(.,.)^{-1}(A) = \bigcup_{t \in \mathbb{N}} (X|_{[0,t]}(.,.))^{-1}(A)$$

denn die Mengen auf der rechten Seite sind nach Voraussetzung in $\mathcal{B}([0, t]) \otimes \mathcal{F}_t \subseteq \mathcal{B}([0, \infty)) \otimes \mathcal{F}$. Umgekehrt braucht ein messbarer Prozess nicht progressiv messbar zu sein. Ein adaptierter Prozess, der càdlàg oder càglàd ist, ist auch progressiv messbar, der Beweis hierzu ist wieder ähnlich zu dem von Theorem 1.1. Analoge Aussagen gelten im Fall $I = [t_0, t_1]$

1.1.3 Martingale

Eine zentrale Rolle in der Theorie der stochastischen Prozesse spielt der Begriff des Martingals:

Definition 1.1.8 *Ein reellwertiger, adaptierter Prozess $(X_t)_{t \in I}$ heißt Martingal (Submartingal, Supermartingal) bezüglich der Filtration $\mathcal{F}_t$, wenn gilt:*

1. $X_t \in \mathcal{L}^1(dP) \quad \forall t \in I$

2. $E[X_t|\mathcal{F}_s] \overset{f.s.}{=} X_s$, *falls* $t \leq s$ *(bzw. „$\leq$ ", „$\geq$ "). Äquivalent hierzu ist die Bedingung* $E[X_t - X_s|\mathcal{F}_s] = 0$

Sofern sogar $X_t \in \mathcal{L}^2(dP) \quad \forall t \in I$ gilt, so kann der nach $\mathcal{F}_s$ bedingte Erwartungswert auch als Projektion auf den Unterraum der $\mathcal{F}_s$ -messbaren Funktionen (mit $\mathcal{L}^2(dP)$ als zugrundeliegendem Hilbertraum und $\langle X, Y\rangle := E[XY]$ als zugehörigem Skalarprodukt) betrachtet werden. Mit dieser Interpretation liegt es nahe, nach einer Zufallsvariable zu suchen, so daß für beliebiges t X_t die Projektion dieser Zufallsvariable auf den Unterraum der $\mathcal{F}_t$ -messbaren Zufallsvariablen ist.

Definition 1.1.9 *Eine Zufallsvariable Y schließt ein Martingal X ab, falls $Y \in \mathcal{L}^1$ und $X_t = E[Y|\mathcal{F}_t], \quad 0 \leq t < \infty$.*

Definition 1.1.10 *Eine Familie $(U_\alpha)_{\alpha \in A}$ von Zufallsvariablen heißt gleichmäßig integrierbar (uniformly integrable), wenn*

$$\lim_{n\to\infty} \sup_{\alpha \in A} \int_{\{|U_\alpha| \geq n\}} |U_\alpha| dP = 0.$$

Sie heißt p-integrierbar, falls

$$\sup_{\alpha \in A} \int_\Omega |U_\alpha|^p dP = \sup_{\alpha \in A} ||U_\alpha||^p_{L^p} < \infty$$

($1 \leq p < \infty$, wichtigster Fall: $p = 2$, in diesem Fall spricht man auch von quadratischer Integrierbarkeit).

Wie die folgende Rechnung zeigt, impliziert die p-Integrierbarkeit für $p > 1$ die gleichmäßige Integrierbarkeit (dabei sei q wie üblich definiert über $\frac{1}{p} + \frac{1}{q} = 1$):

$$
\begin{aligned}
&\lim_{n\to\infty} \sup_{\alpha} \int_{\Omega} |U_\alpha| 1_{\{|U_\alpha|\geq n\}} dP \\
\leq\ & \lim_{n\to\infty} \sup_{\alpha} (\int_{\Omega} |U_\alpha|^p dP)^{\frac{1}{p}} \cdot (\int_{\Omega} 1^q_{\{|U_\alpha|\geq n\}})^{\frac{1}{q}} \\
=\ & \lim_{n\to\infty} \sup_{\alpha} ||U_\alpha||_{\mathcal{L}^p} \cdot (\int_{\Omega} 1_{\{|U_\alpha|\geq n\}})^{\frac{1}{q}} \\
=\ & \lim_{n\to\infty} \sup_{\alpha} ||U_\alpha||_{\mathcal{L}^p} \cdot (P(|U_\alpha| \geq n))^{\frac{1}{q}} \\
\leq\ & \lim_{n\to\infty} \sup_{\alpha} ||U_\alpha||_{\mathcal{L}^p} \cdot (\frac{||U_\alpha||^p_{L^p}}{n^p})^{\frac{1}{q}} \\
\leq\ & \lim_{n\to\infty} \sup_{\alpha} ||U_\alpha||^{1+\frac{p}{q}}_{\mathcal{L}^p} \cdot n^{-\frac{p}{q}} \\
\leq\ & (\sup_{\alpha} ||U_\alpha||^{1+\frac{p}{q}}_{\mathcal{L}^p}) \cdot \lim_{n\to\infty} n^{-\frac{p}{q}} \\
=\ & 0
\end{aligned}
$$

Aus der gleichmäßigen Integrierbarkeit wiederum folgt die Integrierbarkeit, denn $\lim_{n\to\infty} \sup_{\alpha\in A} \int_{\{|U_\alpha|\geq n\}} |U_\alpha| dP = 0$ impliziert, daß es ein n gibt, so daß $\sup_{\alpha\in A} \int_{\{|U_\alpha|\geq n\}} |U_\alpha| dP < \infty$. Dann aber gilt $\sup_{\alpha\in A} \int_{\Omega} |U_\alpha| dP \leq \sup_{\alpha\in A} \int_{\{|U_\alpha|\geq n\}} |U_\alpha| dP + n < \infty$.

Es gilt außerdem auch im Fall eines Martingals X für $s \leq t$ $||X_s||_{\mathcal{L}^p} = ||E[X_t|\mathcal{F}_s]||_{\mathcal{L}^p} \leq ||X_t||_{\mathcal{L}^p}$. Somit ist ein Martingal genau dann p-integrierbar, falls $\lim_{t\to\infty} ||X_t||_{\mathcal{L}^p} < \infty$.

Zur Frage nach der Existenz eines Abschlusses eines Martingals erhält man folgendes Ergebnis (ohne Beweis):

Theorem 1.2 *Sei X ein rechtsstetiges gleichmäßig integrierbares Martingal. Dann existiert $Y = \lim_{t\to\infty} X_t$ fast sicher, $E[|Y|] < \infty$, und Y schließt X ab.*

Insbesondere besitzen damit alle Martingale mit $\lim_{t\to\infty} ||X_t||_{\mathcal{L}^p} < \infty$ für $p > 1$ einen Abschluß.

Ein weiteres interessantes Ergebnis (ohne Beweis, nach Protter [17]) ist das folgende

Theorem 1.3 *Zu einem Martingal X existiert eine (bis auf Ununterscheidbarkeit eindeutige) Modifikation, die càdlàg ist.*

Wir können also, sofern wir zwei Prozesse, die Modifikationen voneinander sind, als gleichwertig betrachten, bei einem Martingal immer davon ausgehen, daß es càdlàg ist.

Da der bedingte Erwartungswert für einen stochastischen Prozeß X je nach verwendetem Wahrscheinlichkeitsmaß verschieden sein kann, hängt auch die Martingaleigenschaft ganz wesentlich vom verwendeten Wahrscheinlichkeitsmaß ab. Insbesondere kann man bei einem gegebenen Prozeß versuchen, ein neues (nach Möglichkeit äquivalentes) Wahrscheinlichkeitsmaß zu finden, unter dem er zum Martingal wird. Ein wichtiges Ergebnis in diesem Zusammenhang ist folgendes

Theorem 1.4 *Es sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und ein zu P äquivalentes Wahrscheinichkeitsmaß Q gegeben. Sei weiter der Prozeß Z gegeben durch $Z_t := E_P[\frac{\partial Q}{\partial P}|\mathcal{F}_s]$. Dann gilt: Ein adaptierter càdlàg-Prozeß M ist genau dann ein Q-Martingal, wenn MZ ein P-Martingal ist.*

Beweis: Es gilt: M ist Q-Martingal

$$\begin{aligned}
&\iff E_Q[X_t - X_s|\mathcal{F}_s] = 0 \quad \forall s \leq t \\
&\iff \int_A (X_t - X_s) dQ = 0 \quad \forall A \in \mathcal{F}_s, \forall s \leq t \\
&\iff \int_A (X_t - X_s) \frac{\partial Q}{\partial P} dP = 0 \quad \forall A \in \mathcal{F}_s, \forall s \leq t \\
&\iff \int_A X_t \frac{\partial Q}{\partial P} dP = \int_A X_s \frac{\partial Q}{\partial P} dP \quad \forall A \in \mathcal{F}_s, \forall s \leq t \\
&\iff \int_A X_t E[\frac{\partial Q}{\partial P}|\mathcal{F}_t] dP = \int_A X_s E[\frac{\partial Q}{\partial P}|\mathcal{F}_s] dP \quad \forall A \in \mathcal{F}_s, \forall s \leq t \\
&\iff \int_A X_t Z_t dP = \int_A X_s Z_s dP \quad \forall A \in \mathcal{F}_s, \forall s \leq t \\
&\iff \int_A ((XZ)_t - (XZ)_s) dP = 0 \quad \forall A \in \mathcal{F}_s, \forall s \leq t \\
&\iff E_P[((XZ)_t - (XZ)_s)|\mathcal{F}_s] = 0 \quad \forall s \leq t
\end{aligned}$$

was wiederum genau dann der Fall ist, wenn XZ ein P-Martingal ist. ■

Eine analoge Aussage gilt für Super- und Submartingale.

1.1.4 Stoppzeiten

Definition 1.1.11 *Eine nichtnegative Zufallsvariable T heißt Stoppzeit, falls* $\{\omega | T(\omega) \leq t\} \in \mathcal{F}_t \quad \forall t \in I$.

Zu einer Stoppzeit und einem gegebenen Prozess $(X_t)_{t \in I}$ definiert man den gestoppten Prozess X^T als $X^T(\omega, t) := X(\omega, \min(t, T(\omega)))$ und die Zufallsvariable X_T als $X_T(\omega) := X_{T(\omega)}(\omega)$.

Stoppzeiten spielen in den Beweisen, die zur stochastischen Integration führen, eine große Rolle. Die folgenden zwei Beispiele von oft verwendeten Stoppzeiten illustrieren den Begriff:

Beispiel 1.1.12 *Sei X ein adaptierter càdlàg-Prozeß und $\Lambda \in \mathbb{R}^d$ eine offene Menge. Dann ist $T(\omega) := \inf\{t > 0 | X(\omega, t) \in \Lambda\}$ eine Stoppzeit.*

Beweis: siehe Protter [17] ■

Beispiel 1.1.13 *Sei X ein adaptierter càdlàg-Prozeß und $\Lambda \in \mathbb{R}^d$ eine abgeschlossene Menge. Dann ist $T(\omega) := \inf\{t > 0 | X(\omega, t) \in \Lambda$ oder $X_-(\omega, t) := \lim_{\substack{s \to t \\ s < t}} X(\omega, s) \in \Lambda\}$ eine Stoppzeit.*

Beweis: siehe Protter [17] ■

Weiterhin sind Vielfache einer Stoppzeit sowie Summen, Minima und Maxima von zwei Stoppzeiten wieder Stoppzeiten. So wie die σ -Algebra $\mathcal{F}_t$ die Information zum Zeitpunkt t angibt, so ist man auch daran interessiert, die Information bis zu einer zufallsabhängigen Stoppzeit T mit Hilfe einer σ -Algebra darzustellen. Man benutzt folgende Definition:

Definition 1.1.14 *Sei T eine Stoppzeit. Dann ist das Mengensystem $\mathcal{F}_T := \{\Lambda \in \mathcal{F} | \Lambda \cap \{T \leq t\} \in \mathcal{F}_t \quad \forall t\}$ eine σ -Algebra (direkt aus der Definition ersichtlich) und wird σ -Algebra zur Stoppzeit T genannt.*

Die Interpretation dieser σ -Algebra als „Information bis T " wird durch das nächste Theorem klar:

Theorem 1.5 *Es gilt (mit den Bezeichnungen wie oben): $\mathcal{F}_T = \sigma(X_T; X$ adaptiert und càdlàg)*
Beweis: *Sei $\mathcal{G} := \sigma(X_T; X_T$ adaptiert und càdlàg) . Sei $\Lambda \in \mathcal{F}_T$ gegeben. Dann ist der durch $X_t := 1_\Lambda 1_{\{t \geq T\}}$ definierte Prozess adaptiert und càdlàg,*

und außerdem gilt $X_T = 1_\Lambda$ *, somit auch* $\Lambda \in \mathcal{G}$ *. Um* $\mathcal{G} \subset \mathcal{F}_T$ *zu zeigen, reicht es, zu zeigen, daß* X_T $\mathcal{F}_T$ *-messbar ist für alle adaptierten càdlàg-Prozesse* X*. Sei dazu für festes, aber beliebiges* t $\phi : \{\omega | T(\omega) \leq t\} \longrightarrow \mathbb{R}_0^+ \times \Omega$ *definiert durch* $\phi(\omega) := (T(\omega), \omega)$ *. Da* X *als càdlàg-Prozeß* $\mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{F} - \mathcal{B}(\mathbb{R}^d)$ *-messbar ist und* ϕ *eine meßbare Abbildung von* $(\{\omega | T(\omega) \leq t\}, \mathcal{F}_t \cap \{\omega | T(\omega) \leq t\})$ *nach* $(\mathbb{R}_0^+ \times \Omega, \mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{F})$ *, ist auch* $X \circ \phi$ *eine* $(\mathcal{F}_t \cap \{\omega | T(\omega) \leq t\}$ *-* $\mathcal{B}(\mathbb{R}^d)$ *-)meßbare Abbildung. Andererseits gilt aber auch* $X \circ \phi = X_T|_{\{\omega | T(\omega) \leq t\}}$*. Somit gilt für eine beliebige Borelmenge* $A \in \mathcal{B}(\mathbb{R}^d)$*:* $X_T^{-1}(A) \cap \{\omega | T(\omega) \leq t\} = X_T|^{-1}_{\{\omega | T(\omega) \leq t\}}(A) = (X \circ \phi)^{-1}(A) \in \mathcal{F}_t \cap \{\omega | T(\omega) \leq t\} \subset \mathcal{F}_t$ *, also* $X_T^{-1}(A) \in \mathcal{F}_T$ *(Die letzte Teilmengenrelation folgt daraus, daß* $\{\omega | T(\omega) \leq t\} \in \mathcal{F}_t$ *nach Definition des Begriffs Stoppzeit).* ■

Bemerkung 1.1.15 *Im Beweis wird verwendet, daß ein càdlàg-Prozeß auch als Abbildung von* $\mathbb{R}_0^+ \times \Omega$ *nach* $\mathbb{R}^d$ *, ausgestattet mit den* σ *-Algebren* $\mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{F}$ *und* $\mathcal{B}(\mathbb{R}^d)$ *meßbar ist, wohingegen die Definition des Begriffs Prozeß nur Meßbarkeit von* $X(., t)$ *bezüglich* $\mathcal{F}$ *und* $\mathcal{B}(\mathbb{R}^d)$ *für alle* t *verlangt. Es liegt also ein ähnlicher Unterschied wie zwischen den Begriffen „ununterscheidbar" und „Modifikation" vor, und auch der (hier ausgelassene) Beweis der Äquivalenz der beiden Begriffe für càdlàg-Prozesse verwendet zunächst die Abzählbarkeit von* $\mathbb{Q}_0^+$ *und anschließend die Stetigkeit des Prozesses.*

Das folgende Theorem (ohne Beweis) zeigt schließlich, daß die „Projektionseigenschaft" des Martingals auch mit Stoppzeiten anstelle von festen Zeiten gültig ist:

Theorem 1.6 (Doob's optional sampling theorem) *Sei* X *ein rechtsstetiges Martingal mit Abschluß,* T *und* S *zwei Stoppzeiten, so daß* $S \leq T$ *f.s. Dann gilt* $E[X_T | \mathcal{F}_S] = X_{T \wedge S}$ *f.s.*

Mit diesem Theorem folgt weiter:

Theorem 1.7 *Sei* X *ein gleichmäßig integrierbares rechtsseitig stetiges Martingal, und* T *eine Stoppzeit. Dann ist auch der gestoppte Prozeß* X^T *ein gleichmäßig integrierbares, rechtsseitig stetiges Martingal.*

Beweis: siehe Protter [17]. ■

Bemerkung 1.1.16 *Im folgenden ist, sofern nicht ausdrücklich anders erwähnt,* $I = [0, \infty)$ *, auch wenn die Aussagen in der Regel (ggf. mit naheliegender Modifikation) auch im Fall* $I = \mathbb{N}_0$ *oder* $I = [t_0, t_1]$ *gültig sind.*

1.2 Lévy-Prozesse

Eine wichtige Klasse von Prozessen, die in vielen Bereichen, insbesondere in der Finanzmathematik, Anwendung findet, ist die der Levy-Prozesse, d.h. Prozesse mit unabhängigen, stationären Zuwächsen:

Definition 1.2.1 *Ein adaptierter ($\mathbb{R}^d$-wertiger) Prozeß X heißt Lévy-Prozeß, wenn gilt:*

1. $X_0 = 0$ *fast sicher*
2. $X_t - X_s$ *ist unabhängig von* $\mathcal{F}_s, \quad 0 \leq s < t < \infty$
3. $X_t - X_s \stackrel{d}{=} X_{t-s} - X_0 (= X_{t-s}), \quad 0 \leq s < t < \infty$
4. $X_t \xrightarrow[t \to s]{p} X_s, \quad 0 \leq s, t < \infty$

Bemerkung 1.2.2 *Die Bedingung 1 in der obigen Definition ist nicht wesentlich; hat man einen adaptierten Prozeß X , der die Bedingungen 2 bis 4 erfüllt, so ist $\widetilde{X} := (X_t - X_0)_{t \geq 0}$ ein Lévy-Prozeß. Die Konvergenzbedingung in 4 ist für stetige Prozesse erfüllt, da für diese ja sogar $X_t \xrightarrow[t \to s]{f.s.} X_s$, $0 \leq s, t < \infty$. Wie man unmittelbar aus der Definition erkennen kann, sind außerdem auch Vielfache von Lévy-Prozessen und Summen untereinander unabhängiger Lévy-Prozesse wieder Lévy-Prozesse.*

Bemerkung 1.2.3 *Bedingung 2 ist, falls $X_t \in \mathcal{L}^1(dP)$, $t \geq 0$, äquivalent zu $E[X_t - X_s | \mathcal{F}_s] = E[X_t - X_s]$. Somit folgt in diesem Fall, daß $(X_t - E[X_t])_{t \geq 0}$ ein Martingal ist.*

Lévy-Prozesse sind sehr fruchtbar in der Anwendung, da die Bedingungen von Definition 1.2.1 die Klasse der zu betrachtenden Prozesse noch nicht zu sehr einschränken, man aber sehr starke Resultate bezüglich der Struktur solcher Prozesse hat. Zunächst allerdings betrachten wir einige prominente Beispiele von Lévy-Prozessen.

1.2.1 Beispiele

Die Brownsche Bewegung

Definition 1.2.4 *Ein adaptierter, $\mathbb{R}^d$-wertiger Prozeß B heißt (d-dimensionale) Brownsche Bewegung (oder auch Wiener-Prozeß), falls er die folgenden Bedingungen erfüllt, wobei C eine beliebige nichtsinguläre Matrix ist:*

1. $B_0 = 0$ *fast sicher*
2. $B_t - B_s$ *ist unabhängig von* $\mathcal{F}_s$, $0 \leq s < t < \infty$
3. $B_t - B_s \stackrel{d}{=} N(0, (t-s)C)$, $0 \leq s < t < \infty$

Man kann zeigen, daß ein solcher Prozeß eine Modifikation besitzt, die P-f.s. stetig ist. Es handelt sich also (ggf. nach Modifizierung, die die beiden ersten Bedingungen nicht berührt) um einen Lévy-Prozeß. Eine Verallgemeinerung dieser Definition ist möglich, indem man Bedingung 1 wegläßt.

Die Brownsche Bewegung ist einer der am besten untersuchten stochastischen Prozesse, siehe z.B. [18]. Daher ist es auch nicht verwunderlich, daß sie in den Anfängen der Portfoliooptimierung regelmäßig zur Modellierung eingesetzt wurde. Allerdings ist ihre Praxistauglichkeit beschränkt, da die Normalverteilungsannahme 3 eine starke, bisweilen zu starke, Voraussetzung ist.

Der (zusammengesetzte) Poissonprozeß

Definition 1.2.5 *Sei $(X_i)_{i\in\mathbb{N}}$ eine iid-Folge von $\exp(\lambda)$-verteilten reellwertigen Zufallsvariablen, und (Y_i) eine weitere, von $(X_i)_{i\in\mathbb{N}}$ unabhängige Folge von $\mathbb{R}^d$-wertigen iid Zufallsvariablen. Sei $N := (\sum_{i\in\mathbb{N}} 1_{\{\omega|\sum_{k=1}^{i} X_i(\omega)\leq t\}})_{t\geq 0}$. Dann heißt der Prozeß $Z := (\sum_{l=1}^{N_t} Y_i)_{t\geq 0}$ zusammengesetzter Poissonprozess.*

Wie man zeigen kann, besitzt dieser Prozeß unabhängige und stationäre Zuwächse, und ist auch stochastisch stetig. Da außerdem $Z_0 = 0$ fast sicher, handelt es sich auch hier um einen Lévy-Prozeß.

Der Poissonprozeß hat darüber hinaus weitere charakteristische Eigenschaften: Er ist offenbar stückweise konstant, rechtsstetig und besitzt linksseitige Grenzwerte, er ist also ein càdlàg-Prozeß.

1.2.2 Strukturresultate

Wie in Bemerkung 1.2.2 bereits erwähnt, sind auch Summen unabhängiger Lévy-Prozesse wieder Lévy-Prozesse. Da auch der (nicht zufallsabhängige Prozeß $(at)_{t\geq 0}$ ein Lévy-Prozeß ist, sind alle Prozesse der Form

$$(at + B_t + Z_t)_{t\geq 0}$$

mit einer Brownschen Bewegung B und einem von B unabhängigen Poissonprozeß Z Lévy-Prozesse. Wie in Sato [19] gezeigt wird, läßt sich jeder Lévy-Prozeß als Verallgemeinerung eines solchen Prozesses darstellen. Genauer gesagt gilt:

Theorem 1.8 (Lévy-Khintchine-Darstellung) *1. Die charakteristische Funktion von L_t kann man darstellen als*

$$E[\exp(isL_t)] = \exp(t\Psi(s)),\ \ s \in \mathbb{R}^d,$$

mit

$$\Psi(s) = ia_L^\top s - \frac{1}{2}s^\top\beta^\top\beta s + \int_{\mathbb{R}^d}(e^{is^\top x} - 1 - is^\top x \cdot 1_{|x|\leq 1})\nu_L(dx), \tag{1.1}$$

wobei $a \in \mathbb{R}^d$, $\beta^\top\beta$ eine positiv definite und symmetrische $d \times d$-Matrix ist, und ν ein Maß ist, das $\nu(\{0\}) = 0$ und $\int_{\mathbb{R}^d}(|x|^2 \wedge 1)\nu(dx) < \infty$ erfüllt (sogenannte Lévy-Khintchine-Darstellung). Das Tripel (a, β, ν) ist für einen gegebenen Lévy-Prozeß eindeutig und bestimmt seine endlichdimensionale Verteilung.

2. Weiterhin kann man L darstellen als:

$$L_t = at + \beta B_t + \sum_{0<s\leq t}\Delta L(s)\cdot 1_{\{|\Delta L(s)|>1\}} + \int_0^t\int_{|x|\leq 1} x(M(dx,ds) - \nu(dx)ds) \tag{1.2}$$

wobei für $\omega \in \Omega$ $\Delta L(t,\omega) := L(t,\omega) - L(t-,\omega)$, $t \geq 0$, und $M(A,\omega) = \#\{(t, \Delta L(t,\omega)) \in A\}$ für jede Borelmenge A aus $[0,\infty) \times (\mathbb{R}^d \setminus \{0\})$.

3. Falls $\int_{\mathbb{R}^d}|x|\nu(dx) < \infty$, hat L die Darstellung

$$L_t = \gamma t + \beta B_t + \sum_{0<s\leq t}\Delta L_s, \quad t \geq 0,$$

mit $\gamma = a - \int_{|x|\leq 1} x\nu(dx)$.

1.3 Markov-Prozesse und Diffusionen

Eine weitere Klasse von Prozessen ist die der Markov-Prozesse, die der stochastischen Kontrolltheorie (Kapitel 4) zugrundeliegen. Die wesentliche Eigenschaft der Markov-Prozesse ist, daß ihr zukünftiges Verhalten nur vom gegenwärtigen Zustand, nicht aber vom gesamten Verlauf in der Vergangenheit abhängt.

1.3.1 Markov-Prozesse

Die Markov-Eigenschaft

Definition 1.3.1 *Ein $\mathbb{R}^d$-wertiger adaptierter Prozeß X mit der von ihm selbst erzeugten Filtration $(\mathcal{F}_t)_{t\geq 0}$ besitzt die Markov-Eigenschaft, wenn*

$$P(X_t \in B|\mathcal{F}_s) = P(X_t \in B|X_s) \quad \forall s \leq t \quad P-f.s. \tag{1.3}$$

Dabei ist $P(A|\mathcal{G}) := E\{1_A(.)|\mathcal{G}\}$ für $A \in \mathcal{F}$, $\mathcal{G} \subseteq \mathcal{F}$.

Anschaulich bedeutet diese Eigenschaft: Der gegenwärtige Zustand des Prozesses beinhaltet die ganze für die Zukunft relevante Information aus dem bisherigen Verlauf des Prozesses.

Bemerkung 1.3.2 *Es wurden bereits Martingale und Lévy-Prozesse vorgestellt. Die wesentliche Eigenschaft der Martingale war, daß $E[X_t - X_s|\mathcal{F}_s] = 0$ für $s \leq t$ war. Eine der beiden wesentlichen Eigenschaften von Lévy-Prozessen war die Unabhängigkeit der Zuwächse, d.h. $E[X_t - X_s|\mathcal{F}_s] = E[X_t - X_s]$ (bei Integrierbarkeit). Aus der Martingaleigenschaft folgt die Unabhängigkeit der Zuwächse, denn $E[X_t - X_s] = E[E[X_t - X_s|\mathcal{F}_s]] = E[0] = 0 = E[X_t - X_s|\mathcal{F}_s]$. Eine weitere Abschwächung ist die Forderung $E[X_t - X_s|\mathcal{F}_s] = E[X_t - X_s|X_s]$, die aus $E[X_t - X_s|\mathcal{F}_s] = E[X_t - X_s]$ folgt wegen $E[X_t - X_s|X_s] = E[E[X_t - X_s|\mathcal{F}_s]|X_s] = E[E[X_t - X_s]|X_s] = E[X_t - X_s] = E[X_t - X_s|\mathcal{F}_s]$ (alle Gleichungen dieser Bemerkung gelten $P-f.s.$). Bei entsprechender Integrierbarkeit ist dies eine typische Eigenschaft von Markov-Prozessen.*

Bemerkung 1.3.3 *Die Markov-Eigenschaft 1.3 kann leicht allgemein für vollständige metrische Zustandsräume Σ und vollständig geordnete Indexmengen (z.B. $\mathbb{N}$) formuliert werden. Bei der Indexmenge $\mathbb{N}$ und einem abzählbaren Zustandsraum führt dies zum Begriff der Markovkette.*

Die Übergangsverteilung

Bei Markovketten kann das Verhalten des Prozesses allein durch die Übergangsmatrix (und die Anfangsverteilung) für viele Zwecke ausreichend beschrieben werden. Es ist daher naheliegend, eine analoge Größe für zeitstetige reellwertige Prozesse zu definieren. Es ergeben sich allerdings einige „technische" Schwierigkeiten: In einem diskreten Zustandsraum Σ konnte man i.A. von $P(X_s = i) > 0 \quad \forall i \in \Sigma$ (oder zumindest für i in einer Teilmenge von Σ) ausgehen. Im Fall $\Sigma = \mathbb{R}^d$ wird dagegen $P(X_s = x) = 0$ die Regel sein. Daher muß man die intuitive Formulierung der bedingten Wahrscheinlichkeit als Quotient ($P(A|B) = \frac{P(A \cap B)}{P(B)}$) ersetzen durch die bedingte Erwartung einer Indikatorfunktion, ausgewertet an einem geeigneten ω:

Definition 1.3.4 (Bedingte Wahrscheinlichkeit) *Sei ein beliebiger Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und $\mathbb{R}^d$ (bzw. $\mathbb{R}^m$)-wertige Zufallsvariablen X, Y gegeben. Die bedingte Wahrscheinlichkeit, daß Y in einer Menge $B \in \mathcal{B}(\mathbb{R}^d)$ ist gegeben $X = x \in \mathbb{R}^m$ ist dann definiert durch*

$$P(Y \in B | X = x) = E[1_{Y^{-1}(B)}(.)|\sigma(X)](\omega) \quad \textit{mit } X(\omega) = x \tag{1.4}$$

Diese Größe ist wohldefiniert, da $E[1_{Y^{-1}(B)}(.)|\sigma(X)]$ $\sigma(X)$-meßbar sein muß und somit für alle ω mit $X(\omega) = x$ denselben Wert hat.

Diese Definition ist offenbar nicht ganz unproblematisch: Da bedingte Erwartungen nur $P - f.s.$ eindeutig sind, muß die bedingte Wahrscheinlichkeit keineswegs für alle x ein Wahrscheinlichkeitsmaß sein, wie es die Vorstellung nahelegt (näheres siehe z.B. Bauer [3], darin wird die Frage behandelt, ob man geeignete Versionen der bedingten Wahrscheinlichkeit finden kann). Wir beschränken uns im weiteren darauf, zusätzliche Regularitätsbedingungen bei der Definition der Übergangsverteilung (dem Analogon zur Übergangsmatrix) zu fordern.

Definition 1.3.5 *Die Übergangsverteilung $\widehat{P}$ zu einem $\mathbb{R}^d$-wertigen Prozeß $(X_t)_{t \geq 0}$ ist definiert durch*

$$\widehat{P}(s, x, t, B) = P(X_t \in B | X_s = x), \quad s, t \in [0, \infty), s < t, x \in \mathbb{R}^d, B \in \mathcal{B}(\mathbb{R}^d) \tag{1.5}$$

sofern Versionen der bedingten Erwartung existieren, so daß gilt:

1. *Die Funktion $\widehat{P}(s,x,t,.)$ ist für festes s,t,x ein Wahrscheinlichkeitsmaß*

2. *$\widehat{P}(s,.,t,B)$ ist für festes s,t,B meßbar (bezüglich der Borel-σ-Algebren auf $\mathbb{R}$ bzw. $\mathbb{R}^d$)*

3. *Es gelten die Chapman-Kolmogorov-Gleichungen:*

$$\widehat{P}(s,x,t,B) = \int_{y\in\mathbb{R}^d} \widehat{P}(r,y,t,B)\widehat{P}(s,x,r,dy) \tag{1.6}$$

Bemerkung 1.3.6 *Die ersten beiden Forderungen bedeuten gerade, daß $\widehat{P}(s,.,t,.)$ ein Übergangskern ist. Die dritte Forderung ist ein kontinuierliches Analogon zur Formel von Bayes. Weiterhin ist die Definition nicht auf Prozesse mit der Markov-Eigenschaft beschränkt, wird aber im weiteren nur im Zusammenhang mit solchen verwendet.*

Markov-Prozesse

Definition 1.3.7 *Ein Prozeß X, der die Markov-Eigenschaft besitzt und der eine Übergangsverteilung im Sinne der Definition 1.3.5 besitzt, heißt Markov-Prozeß. Gilt zusätzlich $\widehat{P}(s,x,t,B) = \widehat{P}(0,x,t-s,B)\ \forall s,t,x,B$, so heißt der Prozeß autonomer oder zeithomogener Markov-Prozeß.*

Bei gewöhnlichen zeitinhomogenen Differentialgleichungen kann man eine äquivalente homogene Gleichung aufstellen, indem man die Zeit als zusätzliche Variable auffaßt. Eine analoge Erweiterung des Zustandsraums um eine Dimension führt auch bei Markov-Prozessen zum Erfolg: Ist X ein Markov-Prozeß, so ist der Prozeß $(X_t,t)_{t\geq 0}$ ebenfalls ein Markov-Prozeß, aber zeithomogen (vgl. Øksendal [15]). Im weiteren werden wir daher den zeithomogenen Fall betrachten, sofern dies vorteilhaft erscheint. Außerdem ist unter dem Begriff Markov-Prozeß im folgenden immer ein $\mathbb{R}^d$-wertiger Markov-Prozeß mit Indexmenge $I=[0,\infty)$ gemeint, sofern nichts anderes erwähnt ist.

Wir definieren schließlich für Funktionen Φ, die auf $I\times\mathbb{R}^d$ definiert sind, einen zum Markovprozeß assoziierten Operator A mit dem durch die Übergangsverteilung erzeugten Erwartungswert E_{tx} durch

$$A\Phi(t,x) = \lim_{h\to 0^+} h^{-1}[E_{tx}\Phi(t+h,x(t+h)) - \Phi(t,x)] \tag{1.7}$$

sofern der Grenzwert existiert. In Fleming [7] wird die Frage behandelt, für welche Φ der Grenzwert existieren muß, damit A den Markovprozeß (probabilistisch) eindeutig bestimmt.

1.3.2 Diffusionen und Itô-Diffusionen

Bestimmte Markov-Prozesse, sogenannte Itô-Diffusionen, können als Lösungen von stochastischen Differentialgleichungen dargestellt werden. Diese Prozesse, zu denen auch die Brownsche Bewegung gehört, sind die Grundlage der Kontrolltheorie.

Diffusionen

Definition 1.3.8 *Ein $\mathbb{R}^d$-wertiger Markov-Prozeß X mit Indexmenge I heißt Diffusion, wenn folgende Bedingungen erfüllt sind:*

1. *Für $\epsilon > 0$ gilt*

$$\lim_{h\to 0^+} h^{-1} \int_{|x-y|>\epsilon} \widehat{P}(t,x,t+h,dy) = 0, \quad t \in I, x \in \mathbb{R}^d$$

2. *Es gibt Funktionen $a_{ij}(t,x)$, $f_i(t,x)$, $i,j = 1,2,\ldots,d$, so daß für $\epsilon > 0$*

$$\lim_{h\to 0^+} h^{-1} \int_{|x-y|\le\epsilon} (y_i - x_i)\widehat{P}(t,x,t+h,dy) = f_i(t,x), \quad t \in I, x \in \mathbb{R}^d$$
$$\lim_{h\to 0^+} h^{-1} \int_{|x-y|\le\epsilon} (y_i - x_i)(y_j - x_j)\widehat{P}(t,x,t+h,dy) = a_{ij}(t,x), \quad t \in I, x \in \mathbb{R}^d$$

Die Funktionen $f = (f_i)_{i=1,2,\ldots,d}$ und $a = (a_{ij})_{i,j=1,2,\ldots,d}$ werden auch als lokaler Driftkoeffizient bzw. als lokale Kovarianzmatrix bezeichnet, denn falls $\lim_{h\to 0^+} h^{-1} \int_{|x-y|>\epsilon} |x-y|^2 \widehat{P}(t,x,t+h,dy) = 0$ gilt, sind hf und ha offenbar Approximationen zum Erwartungswert und der Kovarianzmatrix von $(X_{t+h} - X_t)$. Im zeithomogenen Fall hängen f und a nicht von t ab. Für den Operator A (vgl. 1.7) gilt nach Fleming [7] mit zweimal stetig differenzierbaren Funktionen Φ:

$$A\Phi = \Phi_t + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\Phi_{x_i x_j} + \sum_{i=1}^{d} f_i \Phi_{x_i}$$

Itô-Diffusionen

Definition 1.3.9 *Ein $\mathbb{R}^d$-wertiger Prozeß X mit Indexmenge $[0,\infty)$ heißt Itô-Diffusion, wenn er eine Differentialgleichung der Form*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad t \geq s, X_s = x \tag{1.8}$$

erfüllt, wobei B eine m-dimensionale Brownsche Bewegung ist, und $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ sowie $\sigma : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ Funktionen, die Eindeutigkeit und Existenz der Lösung der Differentialgleichung garantieren. Der Prozeß heißt zeithomogen, falls b und σ nicht von t abhängen.

Ist ein Prozeß X eine Itô-Diffusion, so erfüllt der Prozeß $\widetilde{X} := (t, X_t)_{t\geq 0}$ offenbar die Differentialgleichung

$$d\widetilde{X}_t = \begin{pmatrix} dt \\ dX_t \end{pmatrix} = \begin{pmatrix} dt \\ b(t, X_t)dt + \sigma(t, X_t)dB_t \end{pmatrix} = \begin{pmatrix} 1 \\ b(\widetilde{X}_t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma(\widetilde{X}_t) \end{pmatrix} dB_t$$

und die neuen Koeffizienten garantieren wieder die Eindeutigkeit der Lösung. Es genügt also, zeithomogene Itô-Diffusionen zu betrachten. Eine wichtige Erkenntnis ist nun, daß Itô-Diffusionen Diffusionen sind und daß sich die charakteristischen Größen f und a aus den Koeffizienten der Differentialgleichung ergeben:

Theorem 1.9 (Fleming [7], S.135) *Ein Prozeß, der eine Itô-Diffusion ist, ist auch eine Diffusion. Mit den Notationen aus Definition 1.3.8 und 1.3.9 gilt:*

$$a_{ij}(t, x) = \sum_{l=1}^{d} \sigma_{il}(t, x)\sigma_{jl}(t, x), \quad i, j = 1, 2, \ldots, d$$
$$b_i(t, x) = f_i(t, x), \quad i = 1, 2, \ldots, d$$

1.4 Semimartingale

1.4.1 Prozesse mit endlicher Variation

Definition 1.4.1 *Eine Funktion $f : I \longrightarrow \mathbb{R}^d$, $I \subseteq \mathbb{R}$ ist von beschränkter Variation, falls $\sup\limits_{\pi\in\Pi_I} \sum_i |f(t_{i+1}) - f(t_i)| < \infty$, wobei Π_I die Menge aller Partitionen von I ist und t_i über $\pi = \bigcup_i [t_i, t_{i+1})$ definiert sind. f heißt auf $I' \subset I$ von beschränkter Variation, falls $f|_{I'}$ von beschränkter Variation ist.*

Definition 1.4.2 *Ein Prozeß X ist von beschränkter Variation auf $I \subseteq \mathbb{R}$, wenn $P(\{\omega | X(\omega, .)$ ist von beschränkter Variation auf $I\}) = 1$. Ein Prozeß mit Indexmenge $[0, \infty)$ heißt FV-Prozeß, wenn er auf allen kompakten Intervallen von beschränkter Variation ist. Die Menge aller FV-Prozesse bezeichnen wir mit $FV([0, \infty), \mathbb{R}^d)$ oder kurz FV.*

Offenbar ist es (wegen $I_1 \subseteq I_2 \implies \Pi_{I_1} \oplus \Pi_{I_2 \setminus I_1} \subseteq \Pi_{I_2} \implies \sup\limits_{\pi \in \Pi_{I_1}} \sum_i |f(t_{i+1}) - f(t_i)|) \leq \sup\limits_{\pi \in \Pi_{I_2}} \sum_i |f(t_{i+1}) - f(t_i)|$) hinreichend dafür, daß X ein FV-Prozeß ist, zu zeigen, daß X beschränkte Variation auf jedem Intervall $[0, a_n)$ mit einer beliebigen, gegen ∞ konvergierenden Folge a_n besitzt (bei der Begründung in Klammern ist die direkte Summe so zu interpretieren, daß man aus zwei Partitionen auf $I_1 bzw.$ $I_2 \setminus I_1$ eine auf I_2 bildet, indem man die Punkte aus beiden Partitionen zusammennimmt).

Beispiel 1.4.3 *Ein Poissonprozeß ist ein Prozeß mit beschränkter Variation auf jedem kompakten Intervall, da er stückweise konstant ist und nur endlich viele Sprünge besitzt. Dagegen ist die Brownsche Bewegung von unbeschränkter Variation (was nicht ganz trivial ist, Beweis siehe Resnick [18]).*

Die Eigenschaft, unbeschränkte Variation zu besitzen, wird bei der Definition des stochastischen Integrals (mit Brownscher Bewegung als Integrator) Probleme bereiten.

Schließlich führen wir an dieser Stelle noch den Begriff des Variationsprozesses ein:

Definition 1.4.4 *Sei A ein Prozeß mit beschränkter Variation. Dann heißt der Prozeß $(|A_t|)_{t \geq 0} := (\sup_{n \geq 1} \sum\limits_{k=1}^{2^n} |A_{\frac{tk}{2^n}} - A_{\frac{t(k-1)}{2^n}}|)_{t \geq 0}$ (totaler) Variationsprozeß zu A.*

1.4.2 Lokalisierung

Hat man einen beliebigen Prozeß X gegeben, so kann man zu ihm Stoppzeiten definieren, so daß der gestoppte Prozeß „günstigere" Eigenschaften aufweist als der ursprüngliche. Definiert man z. B.(bei gegebenem stetigen reellwertigen Prozeß X und einer beliebigen positive Zahl K) $T(\omega) := \inf\{t \geq 0 | X(\omega, t) \geq K\}$, so ist der gestoppte Prozeß X^T offenbar durch K

beschränkt. Wählt man nun eine Folge $(T_n)_{n\in\mathbb{N}}$ von Stoppzeiten, so daß $T_n \overset{f.s.}{\to} \infty$, so gilt $X^{T_n} \to X$ punktweise in $\Omega \times \mathbb{R}$. Man kann somit den Prozeß X durch ggf. „einfachere" Prozesse approximieren (im Sinne punktweiser Konvergenz). Diese Technik ist in vielen Beweisen unverzichtbar: Man kann nun Eigenschaften von X zeigen, indem man nachweist, daß die gestoppten Prozesse diese besitzen und die Eigenschaft unter punktweiser Konvergenz erhalten bleibt. Diese Idee ist die Grundlage der folgenden

Definition 1.4.5 *Sei $\mathcal{P}$ irgendeine Eigenschaft, die ein Prozeß besitzen kann. Dann sagt man, der Prozeß besitze diese Eigenschaft lokal, falls es eine Folge $(T_n)_{n\in\mathbb{N}}$ von Stoppzeiten mit $T_n \overset{f.s.}{\to} \infty$ gibt, so daß $X^{T_n} 1_{\{T_n>0\}}$ diese Eigenschaft besitzt $\forall n \in \mathbb{N}$.*

Beispiel 1.4.6 *Sei T_n eine Folge von Stoppzeiten, die f.s. gegen unendlich geht. Dann heißt ein Prozeß X lokal quadratintegrierbar, falls*

$\sup\limits_t \int_\Omega |X^{T_n}(.,t)1_{\{T_n>0\}}(.)|^2 dP < \infty \ \forall n \in \mathbb{N}$.

Weiterhin gibt es noch den Begriff des lokalen Martingals:

Definition 1.4.7 *Ein Prozeß X heißt lokales Martingal, wenn er lokal ein gleichmäßig integrierbares Martingal ist.*

Offenbar ist also ein lokales Martingal auch lokal ein Martingal, aber nicht unbedingt umgekehrt. Nach Protter [17] (Kapitel I.5,Theorem **46**) gilt aber auch die Umkehrung, sofern der betrachtete Prozeß adaptiert und càdlàg ist.

An dieser Stelle stellt sich auch die Frage nach dem Zusammenhang zwischen der (globalen) Eigenschaft, ein Martingal zu sein, und der Eigenschaft, ein lokales Martingal zu sein. Ein Martingal ist natürlich auch lokal ein Martingal (mit $T_n := n \ \forall n \in \mathbb{N}$) und somit, sofern es ein adaptierter càdlàg-Prozeß ist, auch ein lokales Martingal. Die umgekehrte Folgerung benötigt zusätzliche Integrierbarkeitsanforderungen. Hierzu zunächst folgende auch später verwendete

Definition 1.4.8 *Sei X ein beliebiger Prozeß. Dann ist $X^*(\omega,t) := \sup\limits_{s\le t} |X_s(\omega)|$ und $X^*(\omega) := \sup\limits_s |X_s(\omega)|$.*

Mit diesen Bezeichnungen gilt folgendes

Theorem 1.10 *Sei X ein lokales Martingal, und sei $E[X_t^*] < \infty$ $\forall t \geq 0$ Dann ist X ein Martingal. Wenn sogar $E[X^*] < \infty$, dann ist X darüber hinaus auch gleichmäßig integrierbar.*

Beweis: Sei T_n die in der Definition verlangte Folge von Stoppzeiten für X. Sei $s \leq t$. Dann gilt mit Theorem 1.6 $E[X_t^{T_n}|\mathcal{F}_s] = E[X_{t\wedge T_n}|\mathcal{F}_s] = X_{(t\wedge T_n)\wedge s} = X_{s\wedge T_n} = X_s^{T_n}$. Da $X_t^{T_n} \leq X_t^*$, kann man dominierte Konvergenz anwenden und somit die Grenzwertbildung und den Erwartungswert vertauschen. Dies ergibt die Gleichung $E[X_t|\mathcal{F}_s] = X_s$. Da natürlich $X_t \in \mathcal{L}^1(dP)$, folgt, daß X ein Martingal ist. Wenn $E[X^*] < \infty$, so gilt $\lim\limits_{n\to\infty} \sup\limits_{t\geq 0} \int |X_t| 1_{\{|X_t|\geq n\}} dP \leq \lim\limits_{n\to\infty} \sup\limits_{t\geq 0} \int |X^*| 1_{\{|X^*|\geq n\}} dP = \lim\limits_{n\to\infty} \int |X^*| 1_{\{|X^*|\geq n\}} dP = 0$ (wiederum dominierte Konvergenz), also ist X dann gleichmäßig integrierbar. ■

In Protter [17] werden noch weitere hinreichende Kriterien dafür aufgestellt, daß ein lokales Martingal ein Martingal ist, unter anderem ist dies der Fall, falls der quadratischen Variationsprozeß (siehe Kapitel Stochastische Integration) für alle festen Zeitpunkte t endlichen Erwartungswert besitzt.

1.4.3 Semimartingale- Definition und Beispiele

Die bis jetzt betrachteten Prozesse sind noch nicht allgemein genug, um alle Möglichkeiten der stochastischen Integration auszuschöpfen. Die Brownsche Bewegung, die früher als einziger Prozeß zur Modellierung in der Portfoliooptimierung genutzt wurde, läßt sich (wie auch der zusammengesetzte Poissonprozeß) als Spezialfall der allgemeineren Klasse der Lévy-Prozesse auffassen. Daneben gibt es die Klasse der Prozesse mit beschränkter Variation und die Klasse der càdlàg-Martingale, die wiederum in der Klasse der lokalen Martingale enthalten ist. Eine letzte Erweiterung zu einer Klasse, die alle lokalen Martingale sowie alle Prozesse mit endlicher Variation (und, wie man sehen wird, damit auch alle Levy-Prozesse) beinhaltet, betrachten wir nun.

Definition 1.4.9 *Ein adaptierter càdlàg-Prozeß X heißt zerlegbar, falls es Prozesse M, A gibt mit $M_0 = A_0 = 0$, so daß*

$$X_t = X_0 + M_t + A_t \tag{1.9}$$

wobei M ein lokal quadratintegrierbares lokales Martingal ist und A ein FV-Prozeß.

Definition 1.4.10 *Ein adaptierter càdlàg-Prozeß X heißt Semimartingal, falls es Prozesse N, B gibt mit $N_0 = B_0 = 0$, so daß*

$$X_t = X_0 + N_t + B_t, \tag{1.10}$$

wobei N ein lokales Martingal ist und B ein FV-Prozeß.

Beispiel 1.4.11 *Die Summe einer Brownschen Bewegung und eines zusammengesetzten Poissonprozesses ist zerlegbar, da die Brownsche Bewegung ein càdlàg-Martingal, also auch ein lokales Martingal und lokal quadratintegrierbar ist (mit $T_n = n$ gilt $\sup_t E[|X_t^{T_n}|^2] = n < \infty$), und da ein zusammengesetzter Poissonprozeß beschränkte Variation auf kompakten Intervallen besitzt.*

Es läßt sich sogar zeigen, daß jeder Lévy-Prozeß zerlegbar, also insbesondere auch ein Semimartingal ist (siehe Protter[17], Kapitel I, Theorem 40).

Ein zerlegbarer Prozeß ist also ein Semimartingal. Es gilt aber auch (nach Protter[17], Kapitel III, Theorem 1) folgendes

Theorem 1.11 *Folgende Aussagen über einen adaptierten càdlàg-Prozeß X sind äquivalent*

(i) *X ist zerlegbar*

(ii) *X ist ein Semimartingal*

(iii) *Zu $\beta > 0$ gibt es Prozesse M, A mit $M_0 = A_0 = 0$, wobei M ein lokales Martingal mit durch β beschränkten Sprüngen ist und A ein FV-Prozeß, so daß $X_t = X_0 + M_t + A_t$ $\forall t$.*

Im nächsten Kapitel werden wir nun Semimartingale dazu benutzen, um ein stochastisches Integral zu definieren.

Kapitel 2

Stochastische Integration

2.1 Allgemeines zur Integration

Wir wollen nun ein stochastisches Integral definieren, um das Problem der Portfoliooptimierung mathematisch modellieren zu können. Ein solches Integral sollte diverse Eigenschaften erfüllen, die man intuitiv mit dem Integralbegriff verknüpft.

Forderung 2.1 *1. Für zeitlich abschnittsweise konstante Integranden soll das Integral elementar als Summe zu berechnen sein*

2. Es soll linear im Integranden sein.

3. Es sollte bezüglich des Integranden (und einer geeigneten Metrik) stetig sein.

Diese noch etwas vage formulierten Forderungen werden in den nächsten Abschnitten präzisiert.

2.1.1 Überlegungen aus der Funktionalanalysis

Aus funktionalanalytischer Sicht handelt es sich bei einem Integral um eine (von einem festen Integrator induzierte) Abbildung von einem metrischen Raum (in unserem Fall der Raum der stochastischen Prozesse) in einen anderen metrischen Raum. Diese Abbildung ist durch die obigen Forderungen auf einem Unterraum (der idealerweise dicht im ganzen Raum bezüglich der

verwendeten Metrik ist) festgelegt, und sie ist linear und stetig. Die folgenden Überlegungen klären die Frage, inwiefern eine Abbildung durch diese Forderungen bereits eindeutig bestimmt ist.

Definition 2.1.1 *Sei X ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\rho : X \to \mathbb{R}$ heißt Fréchet-Metrik, wenn folgende Bedingungen erfüllt sind:*

1. $\rho(x) \geq 0$ *und* $\rho(x) = 0 \Longleftrightarrow x = 0$
2. $\rho(x) = \rho(-x)$
3. $\rho(x + y) \leq \rho(x) + \rho(y)$

Offenbar ist die von einer Fréchet-Metrik induzierte Abbildung $d : X \times X \to \mathbb{R}, \quad d(x_1, x_2) = \rho(x_1 - x_2)$ eine Metrik, und jede Norm ist eine Fréchet-Metrik.

Lemma 2.2 *Gilt für eine Fréchet-Metrik ρ die Ungleichung $\rho(ax) \leq \max(1, |a|)\rho(x) \forall a \in \mathbb{R}, x \in X$, so gilt für konvergente Folgen x_n, y_n:*

$$\lim_{n\to\infty}(ax_n + y_n) = a \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n \tag{2.1}$$

Dabei ist der Grenzwertbegriff der von der von ρ induzierten Metrik d_X erzeugte.

Beweis: Seien x und y die Grenzwerte der Folgen x_n und y_n. Dann gilt:

$$\begin{aligned}
&\lim_{n\to\infty} d_X(ax + y, ax_n + y_n) \\
&= \lim_{n\to\infty} \rho(a(x - x_n) + (y - y_n)) \\
&\leq \lim_{n\to\infty} \rho(a(x - x_n)) + \lim_{n\to\infty} \rho(y - y_n) \\
&\leq \max(1, |a|) \lim_{n\to\infty} \rho(x - x_n) + \lim_{n\to\infty} \rho(y - y_n) \\
&= 0
\end{aligned}$$

■

Theorem 2.3 *Sei (X, d_X) ein metrischer $\mathbb{R}-$ Vektorraum, (Y, d_Y) ein vollständiger metrischer $\mathbb{R}-$ Vektorraum mit Metriken, die von Fréchet-Metriken ρ_X, ρ_Y erzeugt werden, die zusätzlich die Bedingung des Lemmas 2.2 erfüllen. Sei*

U ein (bzgl. d_X)dichter Unterraum von X, und sei $S : U \to Y$ eine lineare, bezüglich $d_X|_U$ und d_Y stetige Abbildung. Dann gibt es eine eindeutig bestimmte lineare stetige (bezüglich d_X und d_Y) Abbildung $T : X \to Y$, so daß $T|U = S$.

Beweis: Zunächst zur Eindeutigkeit: Seien T_1, T_2 zwei solche Abbildungen. Dann ist $T_1 - T_2$ stetig und $T_1 - T_2|U = 0$. Sei nun $x \in X$ beliebig. Dann gibt es nach Voraussetzung eine Folge x_n in U mit $d_X(x_n, x) \overset{n\to\infty}{\to} 0 \implies d_Y((T_1 - T_2)x_n, (T_1 - T_2)x) = d_Y(0, (T_1 - T_2)x) \overset{n\to\infty}{\to} 0 \implies (T_1 - T_2)x = 0$, also ist $T_1 = T_2$.

Nun zur Existenz: für $x \in X$ sei x_n eine Folge in U mit $d_X(x_n, x) \to 0$. Dann ist x_n Cauchy-Folge, also auch Sx_n (Sei $\epsilon > 0$ gegeben, dann gibt es wegen der Stetigkeit ein $\delta > 0$, so daß $d_X(x_n, x_m) = \rho_X(x_n - x_m) \leq \delta \implies d_Y(Sx_n, Sx_m) = \rho_Y(S(x_n - x_m)) \leq \epsilon$. Sucht man nun ein N mit $n, m \geq N \implies d_Y(Sx_n, Sx_m) \leq \epsilon$, so nimmt man dasjenige N mit $n, m \geq N \implies d_X(x_n, x_m) \leq \delta$). Somit besitzt die Folge Sx_n einen (eindeutig bestimmten) Grenzwert. Für jede andere Folge $(x_n^*)_{n\in\mathbb{N}}$ mit $d_X(x_n^*, x) \to 0$ gilt

$$\begin{aligned} & d_Y(\lim_{n\to\infty} Sx_n, \lim_{n\to\infty} Sx_n^*) \\ & \quad \leq \quad d_Y(\lim_{n\to\infty} Sx_n, Sx_n) + d_Y(Sx_n, Sx_n^*) + d_Y(Sx_n^*, \lim_{n\to\infty} Sx_n^*) \end{aligned}$$

Wir definieren somit $Tx := \lim_{n\to\infty} Sx_n$.

Es bleibt noch zu zeigen, daß $T|U = S$ ist und daß T tatsächlich stetig und linear ist. Es gilt: zu $u \in U$ ist $x_n := u \quad \forall n$ eine approximierende Folge, somit ist $Tu = \lim_{n\to\infty} Sx_n = \lim_{n\to\infty} Su = Su$. Weiterhin gilt: Ist x_n^i eine approximierende Folge von x^i, $i = 1, 2$, so ist $ax_n^1 + x_n^2$ wegen Lemma 2.2 eine approximierende Folge von $ax^1 + x^2$. Somit gilt $T(ax^1 + x^2) = \lim_{n\to\infty} S(ax_n^1 + x_n^2) = \lim_{n\to\infty}(aSx_n^1 + Sx_n^2) = a \lim_{n\to\infty} Sx_n^1 + \lim_{n\to\infty} Sx_n^2 = aTx_1 + Tx_2$. Dabei gilt das vorletzte Gleichheitszeichen wieder aufgrund des Lemmas.

Zur Stetigkeit: Sei $\epsilon > 0$ gegeben. Dann gibt es ein $\delta > 0$, so daß $d_Y(Sx, 0) \leq \epsilon \quad \forall x \in U$ mit $d_X(x, 0) \leq \delta$. Sei nun $x \in X$ mit $d_X(x, 0) \leq \frac{\delta}{2}$.

Dann gilt

$$\begin{aligned}
& d_Y(Tx,0) \\
&= \rho_Y(Tx) = \rho_Y(\lim_{n\to\infty} Sx_n) \\
&\leq \lim_{n\to\infty} \rho_Y(Sx_n) + \lim_{n\to\infty} \rho_Y(\lim_{n\to\infty} Sx_n - Sx_n) \\
&= \lim_{n\to\infty} \rho_Y(Sx_n) = \lim_{n\to\infty} d_Y(Sx_n,0) \\
&= \lim_{n\to\infty} d_Y(Sx_n,0) \leq \epsilon
\end{aligned}$$

Das letzte Ungleichheitszeichen gilt, da es wegen $d_X(x,0) \leq \frac{\delta}{2}$ und $d_X(x_n,x) \to 0$ ein N geben muß, so daß $d_X(x_n,0) \leq \delta \quad \forall n \geq N$, woraus wiederum $d_Y(Sx_n,0) \leq \epsilon \quad \forall n \geq N$ folgt. ■

Bemerkung 2.1.2 *Sind die beiden Metriken von einer Norm induziert, so sind die Bedingungen des Theorems an die Metriken offensichtlich erfüllt.*

Bemerkung 2.1.3 *Die zusätzliche Bedingung des Lemmas 2.2 ist, wie aus dem Beweis hervorgeht, nur für die Linearität der Fortsetzung nötig.*

Das Ziel einer stochastischen Integrationstheorie ist es also, geeignete Teilklassen X, Y von stochastischen Prozessen, Metriken auf X, Y und eine Abbildung auf einem Unterraum U von X zu finden, so daß die Bedingungen des obigen Theorems erfüllt sind. Dabei sollte die Teilklasse X der Integranden möglichst umfassend sein, die vorausgesetzte Abbildung S der Vorstellung eines Integrals entsprechen und der von der Metrik erzeugte Konvergenzbegriff sollte mit den aus der Stochastik bekannten Konvergenzbegriffen(f.s. Konvergenz, Konvergenz in Wahrscheinlichkeit usw.) verwandt sein.

2.1.2 Probleme bei der Integration stochastischer Prozesse

Bei der Definition eines stochastischen Integrals ist es naheliegend, sich zunächst am klassischen Integralbegriff zu orientieren. Die Einführung dieser Integrale ist in vielen Büchern gut beschrieben und wird hier nur als Motivation in wenigen Sätzen skizziert. Beispiele für Integrale auf $\mathbb{R}$ sind das Riemann-Integral und das Lebesgue-Integral.

Das Riemann-Integral verwendet (in der Notation von Theorem 2.3) als Unterraum U den Raum der reellwertigen, stückweise konstanten Funktionen

mit kompaktem Träger $K \subseteq \mathbb{R}$, als Metrik die von der Supremumsnorm für Funktionen induzierte Metrik, und als endgültigen Definitionsbereich den Raum der durch Elemente von U (bezüglich der Supremumsnorm) beliebig gut approximierbaren Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Man erhält dadurch das totale Integral ($Y = \mathbb{R}$); das Integral mit Unter- und Obergrenze kann man dann mit Hilfe von Indikatorfunktionen ableiten.

Das Lebesgue-Integral verwendet den gleichen Unterraum U, allerdings verwendet es als Metrik die von der (zuvor unabhängig vom Integralbegriff definierten) L^1-Norm erzeugte Metrik und als X die Klasse der Funktionen, die durch Elemente aus U beliebig genau (bezüglich der L^1-Norm) approximierbaren Funktionen. Durch die Wahl dieser Metrik erhält man ein größeres X, also einen allgemeineren Integralbegriff.

Ein erster Unterschied eines stochastischen Integrals zu diesen Integralbegiffen ist das Ersetzen von dx bzw. $dP(x)$ (bei maßdefinierenden Funktionen) durch ein noch genauer zu definierendes $dB(\omega, t)$, das von einem Prozeß B erzeugt wird. Wählen wir in einem ersten Ansatz den Raum der (für festes ω) stückweise stetigen Prozesse als U und die von der Supremumsnorm erzeugte Metrik. Dann ist es naheliegend,

$$\int u(\omega, t) dB(\omega, t) := \sum_i u(\omega, \frac{t_i + t_{i+1}}{2})(B(\omega, t_{i+1}) - B(\omega, t_i))$$

für $u \in U$ zu wählen, wobei t_i die (von ω abhängigen) Grenzen der Intervalle sind, auf denen $u(\omega, .)$ konstant ist. Protter [17] (Kapitel I.7) zeigt nun, daß man eine Folge $u_n \in U$ mit $||u_n|| = 1$ (grob gesagt $u(\omega, \frac{t_i+t_{i+1}}{2}) = sgn(B(\omega, t_{i+1}) - B(\omega, t_i))$)wählen kann, so daß $\sup_{n\in\mathbb{N}} \int u_n(\omega, t) dB(\omega, t) \geq Variation(B(\omega, .))$. Damit kann nach dem Satz von Banach-Steinhaus der so definierte Integraloperator nicht stetig auf U bzgl. der Supremumsnorm sein, wenn man einen Prozeß B mit Pfaden von unbegrenzter Variation (wie z.B. die Brownsche Bewegung) hat. Protter führt weiter aus, daß auch die Abschwächung der verwendeten Metrik auf die zur Konvergenz in Wahrscheinlichkeit gehörende Metrik keine Stetigkeit garantiert. Im nächsten Abschnitt wird eine andere Wahl von U und der verwendeten Metrik getroffen, die zum Erfolg führt. Der wesentliche Unterschied wird dabei sein, daß man vorhersehbare Prozesse verwendet (die die Variation der Brownschen Bewegung nicht „berücksichtigen" können).

2.2 Das stochastische Integral

In diesem Abschnitt wird nun mit der Technik von Theorem 2.3 ein für unsere Zwecke hinreichend allgemeines Integral eingeführt. Im ersten Unterabschnitt wird die Frage nach der verwendeten Metrik beantwortet, im zweiten der verwendete Unterraum U (und auch der Bildraum Y) vorgestellt, und im dritten schließlich die Frage der Stetigkeit des Integraloperators auf U geklärt.

2.2.1 Die ucp-Konvergenz

Definition 2.2.1 *Eine Folge H_n von Prozessen konvergiert ucp (uniformly on compacts in probability) gegen einen Prozeß H, wenn $(H^n - H)^*_t = \sup_{0\leq s\leq t} |H^n_s - H_s|$ für alle t in Wahrscheinlichkeit gegen 0 konvergiert für $n \to \infty$.*

Diese Konvergenz leitet sich aus bekannten Konvergenzbegriffen der Stochastik ab. Im folgenden Theorem wird eine zu dieser Konvergenz kompatible Metrik vorgestellt:

Theorem 2.4 *Eine Folge H_n von Prozessen konvergiert genau dann ucp gegen einen Prozeß H, wenn $d(H_n, H) \to 0$ mit $d(X,Y) = \sum_{k=1}^{\infty} 2^{-k} E[\min(1, (X-Y)^*_k)]$.*

Beweis: Es gelte $H_n \to H$ ucp. Dann gilt für alle $\delta \leq 1$ $E[\min(1, (H^n - H)^*_k)] \leq P(\sup_{0\leq s\leq k} |H^n_s - H_s| \geq \delta) + \delta P(\sup_{0\leq s\leq k} |H^n_s - H_s| < \delta) \leq P(\sup_{0\leq s\leq k} |H^n_s - H_s| \geq \delta) + \delta$

Sei nun $\epsilon > 0$ gegeben. Wähle $K \in \mathbb{N}$ so, daß $\sum_{k=K+1}^{\infty} 2^{-k} \leq \frac{\epsilon}{2}$. Wähle $\delta = \frac{\epsilon}{4K}$. Wegen der stochastischen Konvergenz von $\sup_{0\leq s\leq k} |H^n_s - H_s|$ gegen 0 gibt es ein $N \in \mathbb{N}$, so daß $P(\sup_{0\leq s\leq k} |H^n_s - H_s| \geq \delta) \leq \delta$ für alle $k \in$

$\{1, \ldots, K\}$ und alle $n \geq N$. Dann gilt aber für $n \geq N$:

$$\begin{aligned}
d(H_n, H) &= \sum_{k=1}^{\infty} 2^{-k} E[\min(1, (H^n - H)_k^*)] \\
&\leq \sum_{k=1}^{K} E[\min(1, (H^n - H)_k^*)] + \sum_{k=K+1}^{\infty} 2^{-k} \\
&\leq \sum_{k=1}^{K} (P(\sup_{0 \leq s \leq k} |H_s^n - H_s| \geq \delta) + \delta) + \frac{\epsilon}{2} \\
&\leq \sum_{k=1}^{K} (\delta + \delta) + \frac{\epsilon}{2} \\
&\leq 2K\delta + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

Es gelte umgekehrt $d(H_n, H) \underset{n \to \infty}{\to} 0$. Dann folgt mit der Ungleichung von Tschebyscheff $P(\sup_{0 \leq s \leq t} |H_s^n - H_s| \geq \epsilon) \leq P(\sup_{0 \leq s \leq \lceil t \rceil} |H_s^n - H_s| \geq \epsilon) \leq \frac{E[\min(1, (H^n - H)_{\lceil t \rceil}^*)]}{\epsilon} \leq \frac{1}{\epsilon} 2^{\lceil t \rceil} d(H_n, H) \underset{n \to \infty}{\to} 0$ ■

Diese Metrik ist offensichtlich von der Fréchet-Metrik $\rho(X) = \sum_{k=1}^{\infty} 2^{-k} E[\min(1, X_k^*)]$ erzeugt, und es gilt $\rho(aX) \leq \max(1, |a|)\rho(X)$ (Beweis durch Nachrechnen), sie erfüllt also die Anforderungen des Theorems 2.3.

2.2.2 Definition für einfache vorhersehbare Prozesse als Integranden

Definition 2.2.2 *Ein Prozeß H heißt einfach vorhersehbar (simple predictable), wenn es Stoppzeiten $T_1, T_{2,\ldots,} T_{n+1}$ gibt mit $0 = T_1 \leq T_2 \leq \cdots \leq T_{n+1} < \infty$ und Zufallsvariablen H_i mit $|H_i| < \infty$ fast sicher und H_i $\mathcal{F}_{T_i}$-meßbar, $i = 1, \ldots n$, so daß*

$$H(\omega, t) = H_0(\omega) 1_{\{0\}}(t) + \sum_{i=1}^{n} H_i(\omega) 1_{(T_i(\omega), T_{i+1}(\omega)]}(t) \tag{2.2}$$

Den Vektorraum der einfachen vorhersehbaren Prozesse bezeichnen wir auch als $\mathbb{S}$.

Definition 2.2.3 *Sei X ein Semimartingal, H ein einfacher vorhersehbarer Prozeß mit Darstellung 2.2. Dann heißt der Prozeß Z, definiert durch*

$$\begin{aligned} Z(\omega,t) := J_X(H)(\omega,t) &:= \int_0^t H(\omega,t)dX(\omega,t) \\ &:= H_0X_0 + \sum_{i=1}^{n} H_i(\omega)(X^{T_{i+1}}(\omega,t) - X^{T_i}(\omega,t)) \end{aligned} \tag{2.3}$$

das stochastische Integral von H bezüglich X.

Bemerkung 2.2.4 *Um die Wohldefiniertheit des Integrals zu beweisen, ist es noch notwendig, zu zeigen, daß sich für jede Darstellung der Form 2.2 derselbe Prozeß Z ergibt. Dies kann man erreichen, indem man zuerst zeigt, daß eine Verfeinerung der (Zufalls)Partition das gleiche Integral ergibt, und anschließend zeigt, daß es zu zwei Darstellungen desselben Prozesses Verfeinerungen ihrer Partitionen gibt, die übereinstimmen (Näheres siehe Protter [17]).*

Protter [17] zeigt nun, daß die so definierte Abbildung $J_X : \mathbb{S} \to \mathbb{D}$ stetig ist bezüglich ucp-Konvergenz (bzw. bezüglich der zugehörigen Metrik). Weiter zeigt er auch, daß $\mathbb{D}$ bezüglich der ucp-Konvergenz vollständig ist.

2.2.3 Erweiterung auf Integranden in $\mathbb{L}$

Wie in den beiden letzten Unterabschnitten ausgeführt, sind fast alle Voraussetzungen des Theorems 2.3 gegeben: Wir kennen den Unterraum $U = \mathbb{S} \subseteq \mathbb{L} = X$, wir haben einen Bildraum $Y = \mathbb{D}$, wir haben passende Metriken und die Information, daß $\mathbb{D}$ vollständig bezüglich dieser Metriken ist. Schließlich haben wir auch eine stetige lineare Abbildung von $\mathbb{S}$ nach $\mathbb{D}$. Um das Integral auf $\mathbb{L}$ zu erweitern, fehlt also nur noch der Nachweis, daß $\mathbb{S}$ dicht liegt in $\mathbb{L}$.

Theorem 2.5 (Protter, II.10) *$\mathbb{S}$ ist dicht in $\mathbb{L}$ bezüglich ucp*

Beweis: Sei $Y \in \mathbb{L}$ gegeben. Definiere Stoppzeiten $R_n := \inf\{t : |Y_t| > n\}$. Dann sind $Y^{R_n} 1_{\{R_n>0\}}$ beschränkt und konvergieren gegen Y in ucp. Man kann also $Y \in \mathbb{L}$ annehmen. Wir definieren den Prozeß Z durch $Z_t := \lim_{\substack{u\to t \\ u>t}} Y_u$. Dann ist $Z \in \mathbb{D}$. Wir definieren weiterhin für beliebiges $\epsilon > 0$

die Stoppzeiten $T_0^\epsilon := 0, T_{n+1}^\epsilon := \inf\{t : t > T_n^\epsilon \text{ und } |Z_t - Z_{T_n^\epsilon}| > \epsilon\}$. Da Z ein càdlàg-Prozeß ist, gilt $T_n^\epsilon \to \infty$ fast sicher (daß die Folge monoton steigend ist, ist aus der Definition ersichtlich, falls sie einen endlichen Grenzwert $T(\omega)$ besäße, könnte man einen Widerspruch zur Existenz eines linksseitigen Grenzwertes konstruieren). Sei weiter $Z^\epsilon := \sum_n Z_{T_n^\epsilon} 1_{[T_n^\epsilon, T_{n+1}^\epsilon)}$. Dann sind die Z^ϵ wie auch Y beschränkt und konvergieren für $\epsilon \to 0$ in der Supremumsnorm gegen Z. Sei $U^\epsilon := Y_0 1_{\{0\}} + \sum_n Z_{T_n^\epsilon} 1_{(T_n^\epsilon, T_{n+1}^\epsilon]}$. Dann ist U^ϵ beschränkt und in $\mathbb{L}$, und U^ϵ konvergiert ucp gegen $Y_0 1_{\{0\}} + Z_- = Y$. Schließlich definieren wir noch $Y^{n,\epsilon} := Y_0 1_{\{0\}} + \sum_{i=1}^n Z_{T_i^\epsilon} 1_{(T_i^\epsilon, T_{i+1}^\epsilon]}$, womit wir Y beliebig gut in ucp annähern können durch geeignete Wahl von n und ϵ. ■

Nun sind alle Voraussetzungen erfüllt, um ein stochastisches Integral mit Integranden in $\mathbb{L}$ und einem beliebigem Semimartingal X als Integrator zu definieren:

Definition 2.2.5 *Die nach Theorem 2.3 eindeutige stetige lineare Fortsetzung des Integraloperators J_X auf $\mathbb{L}$ heißt stochastisches Integral bezüglich X. Sie wird ebenfalls als $J_X(H) = \int H dX = H \cdot X$ bezeichnet.*

Bemerkung 2.2.6 *Diese Integraldefinition kann, wie in Protter [17] ausgeführt wird, auf Integranden, die bezüglich der σ-Algebra, die von den Prozessen aus $\mathbb{L}$ (aufgefaßt als Abbildungen von $\Omega \times [0, \infty)$ nach $\mathbb{R}$) erzeugt wird, meßbar sind (sogenannte vorhersehbare Prozesse), erweitert werden.*

Bemerkung 2.2.7 *Analog zur Schreibweise des gewöhnlichen Integrals definiert man für ein (offenes, abgeschlossenes oder halboffenes) beschränktes Intervall $A \subseteq [0, T]$) $\int_A H dX := \int_{[0,T]} (1_A H) dX := (\int (1_A H) dX)_T$. Für $A = (0, T]$ schreibt man auch $\int_{0+}^T H dX$, offenbar gilt dann $\int_{[0,T]} H dX = H_0 X_0 + \int_{0+}^T H dX$.*

2.3 Eigenschaften und Anwendungen des stochastischen Integrals

Um stochastische Integrale in der Anwendung nutzen zu können, muß man Methoden finden, sie zu berechnen oder umzuformen. Beim klassischen Integralbegriff ist (neben der Linearität des Integrals) der Hauptsatz der

Differential- und Integralrechnung und die mit seiner Hilfe abgeleiteten Regeln (partielle Integration und Substitution, die mit Hilfe der Produkt- bzw. Kettenregel aufgestellt werden) von großen Nutzen.

Bei stochastischen Integralen gibt es ebenfalls ein starkes Hilfsmittel: Die Itô-Formel.

2.3.1 Die quadratische Variation eines Semimartingals

Definition 2.3.1 *Seien X, Y Semimartingale. Dann heißt der Prozeß $[X,Y] := XY - \int X_- dY - \int Y_- dX$ der quadratische Kovariationsprozeß von X und Y. Der Prozeß $[X,X]$ heißt quadratischer Variationsprozeß.*

Die quadratische Variation wird zur Aufstellung der Itô-Formel benötigt. Im folgenden wird gezeigt, daß man diesen Prozeß sehr anschaulich als Grenzwert darstellen kann.

Definition 2.3.2 *Eine Zufallspartition ist eine endliche Folge von Stoppzeiten $T_1, T_2, \ldots, T_k$, so daß $0 = T_0 \leq T_1 \leq T_2 \leq \cdots \leq T_k < \infty$ fast sicher gilt.*

Theorem 2.6 *Der quadratische Variationsprozeß eines Semimartingals X ist càdlàg, monoton steigend und adaptiert. Außerdem gilt:*

1. $[X,X]_0 = X_0^2$ *und* $\Delta[X,X] = (\Delta X)^2$
2. *Sei* $(\sigma_n)_{n\in\mathbb{N}} = ((T_k^n)_{k=0,1,2,\ldots,k_n})_{n\in\mathbb{N}}$ *eine Folge von Zufallspartitionen, so daß* $T_{k_n}^n \underset{n\to\infty}{\to} \infty$ *f.s. und* $\lim_{n\to\infty} \sup_k |T_{k+1}^n - T_k^n| = 0$ *fast sicher. Dann gilt*
$$X_0^2 + \sum_i (X^{T_{i+1}^n} - X^{T_i^n})^2 \overset{ucp}{\to} [X,X] \tag{2.4}$$
3. *Für jede Stoppzeit T gilt:* $[X^T, X] = [X, X^T] = [X^T, X^T] = [X,X]^T$

Beweis: siehe Protter [17] ■

2.3.2 Die Itô-Formel

Theorem 2.7 (Itô-Formel) *Sei X ein Semimartingal und f eine Funktion aus $\mathcal{C}^2$. Dann ist $(f(X_t))_{t\geq 0}$ wieder ein Semimartingal, und es gilt:*

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s^c + \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\} \tag{2.5}$$

Dabei ist $[X,X]^c$ der stetige Anteil von $[X,X]$.

Beweis: siehe Protter [17] ■

Man erkennt in der Formel bereits die Verwandtschaft mit der bekannten Taylor-Entwicklung, allerdings kommen hier einige Terme hinzu. Die Formel ist intuitiv verständlich, sofern es sich bei X um ein Semimartingal mit $\Delta X \equiv 0$ handelt. In diesem Fall fällt die Summe auf der rechten Seite weg, und als einziger Zusatzterm gegenüber der Taylorformel bleibt das zweite Integral übrig, das offenbar von der unendlichen Variation des Integrators herrührt:

Korollar 2.8 *Sei X ein stetiges Semimartingal und f eine Funktion aus $\mathcal{C}^2$. Dann ist $(f(X_t))_{t\geq 0}$ wieder ein Semimartingal, und es gilt:*

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s \tag{2.6}$$

Gruppiert man die Terme der allgemeinen Itô-Formel um,so erhält man

$$f(X_t) - \sum_{0\leq s\leq t} \Delta(f(X))_s = (\int_{0+}^{t} f'(X_{s-})dX_s - \sum_{0<s\leq t} f'(X_{s-})\Delta X_s) + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s^c \tag{2.7}$$

In dieser Form kann man die linke Seite als den stetigen Anteil von $f(X)$ und den ersten Term der rechten als Integral über $f'(X_{s-})$ mit dem stetigen Teil von X als Integrator (also auch als stetigen Anteil des ursprünglichen Integrals) interpretieren. Die Verallgemeinerung der Itô-Formel für stetige

vektorwertige Semimartingale $X = (X^i)_{i=1,2,\ldots,d}$ (d.h. vektorwertige Prozesse, deren Komponenten stetige Semimartingale sind) und reellwertiges f ergibt (ohne Beweis):

$$f(X_t) - f(X_0) = \sum_i \int_{0+}^{t} \partial_i f(X_{s-}) dX_s^i + \frac{1}{2} \sum_{i,j} \partial_{ij} f(X_{s-}) d[X^i, X^j]_s \quad (2.8)$$

2.3.3 Integrale mit speziellen Integratoren

Es ist nun naheliegend, zu fragen, welchen Einfluß bestimmte Eigenschaften (z.B. Martingaleigenschaft oder lokale Martingaleigenschaft) des Integrators auf das Integral haben. Protter zeigt hierzu folgende Ergebnisse, die ohne Beweis angegeben werden:

Theorem 2.9 *Sei X ein lokal quadratintegrierbares lokales Martingal und $H \in \mathbb{L}$. Dann ist $H \cdot X$ ebenfalls ein lokal quadratintegrierbares lokales Martingal.*

Theorem 2.10 *Sei X ein lokales Martingal und $H \in \mathbb{L}$. Dann ist $H \cdot X$ ebenfalls ein lokales Martingal.*

Mit einem weiter verallgemeinerten Integralbegriff kann man noch weitere Aussagen dieses Typs treffen.

2.3.4 Stochastische Differentialgleichungen

Stochastische Differentialgleichungen (genauer gesagt sind es Integralgleichungen, die abkürzend in Differentialform geschrieben werden) werden eingesetzt, um die zeitliche Entwicklung von Prozessen (z.B. in der Portfoliooptimierung) zu beschreiben. Bei Differentialgleichungen ist neben der Berechnung einer Lösung (oft mit Hilfe der Itô-Formel) auch die Frage der Eindeutigkeit dieser Lösung von Interesse. Kann man eine Lösung nicht explizit angeben, so ist auch die Existenzfrage wichtig. Das folgende Theorem ist für unsere Zwecke ausreichend:

Theorem 2.11 *Sei Z ein Semimartingal mit $Z_0 = 0$, X_0 eine endliche und $\mathcal{F}_0$-meßbare Zufallsvariable und sei $f : \mathbb{R}_+^0 \times \Omega \times \mathbb{R} \to \mathbb{R}$ eine Funktion, die die beiden folgenden Bedingungen erfüllt:*

1. *für festes x ist $f(.,.,x) \in \mathbb{L}$*

2. $|f(t,\omega,x) - f(t,\omega,y)| \leq K(\omega)|x-y|$ *mit einer endlichen Zufallsvariable K*

Dann besitzt die Gleichung

$$X_t = X_0 + \int_0^t f(s,.,X_{s-})dZ_s \tag{2.9}$$

eine eindeutige Lösung und diese ist ein Semimartingal.

Beweis: siehe Protter [17] ■

Ist die Funktion f nicht von ω abhängig, so genügt es also, daß f als Funktion von t càglàd ist und bezüglich x lipschitzstetig ist mit Lipschitzkonstante K unabhängig von t. Die in folgenden Kapiteln auftauchenden Differentialgleichungen werden diese Bedingungen immer erfüllen.

Eine wichtige Differentialgleichung, deren Lösung man explizit angeben kann, ist folgende:

$$Z_t = 1 + \int_0^t Z_{s-}dX_s \tag{2.10}$$

Hier ist $f(t,\omega,x) = x$, also bezüglich t konstant und bezüglich x lipschitzstetig mit Lipschitzkonstante 1. Es existiert also eine eindeutige Lösung, nämlich

$$Z_t = \exp(X_t - \frac{1}{2}[X,X]_t) \prod_{0<s\leq t} (1+\Delta X_s)\exp(-\Delta X_s + \frac{1}{2}(\Delta X_s)^2) \tag{2.11}$$

Dieses Z bezeichnet man auch als stochastisches Exponential und schreibt $Z = \mathcal{E}(X)$.

Kapitel 3

Modelle in der Portfoliooptimierung

3.1 Das allgemeine Modell

Um ein mathematisches Modell für die Portfoliooptimierung aufzustellen, betrachtet man die Kurse der zur Verfügung stehenden Anlagen als stochastische Prozesse in einem Wahrscheinlichkeitsraum mit passender Filterung. Weiterhin benötigt man Prozesse, die die Stückzahlen der gehaltenen Aktien und den Konsum von erworbenem Reichtum beschreiben (diese beiden Prozesse kann man als Investor beeinflussen). Allerdings werden nicht alle Prozesse zulässig sein (aus naheliegenden Gründen), weshalb man Beschränkungen für diese Prozesse formulieren muß. Schließlich muß man noch ein Funktional für den erhaltenen Portfolioverlauf definieren, das als Bewertung dient.

3.1.1 Der Preisprozeß

In der Portfoliooptimierung geht man davon aus, daß ein Investor eine „sichere" Anlagemöglichkeit, deren Preisentwicklung nicht zufallsabhängig ist und verschiedenen riskante Anlagemöglichkeiten (z.B. Aktien) hat. Die Kurse dieser Anlagen (in Geldeinheiten) werden durch einen stochastischen Prozeß $S : [0, \infty) \times \Omega \to \mathbb{R}^{d+1}$ beschrieben. Dabei ist $S_i(t, .)$, $i = 1, 2, \ldots, d$ der (zufallsabhängige) Kurs einer Einheit der Aktie i zum Zeitpunkt t, und $S_0(t, \omega) = \widetilde{S}_0(t)$ der deterministische Wert einer Einheit eines Bankguthabens

zum Zeitpunkt t. Oft verwendet man $\widetilde{S}_0(t) = \exp(rt)$ (stetige Verzinsung mit dem sicheren Zinssatz r).

Welche Eigenschaften sollte nun der Prozeß S besitzen?

Da man ihn in den folgenden Anwendungen als Integrator verwenden will, kommen nur Semimartingale in Frage. Eine weitere sinnvolle Forderung ist sicherlich, daß er (komponentenweise) nichtnegativ ist, denn weder Aktienkurse noch der Wert eines Bankguthabens können negativ werden. Diese beiden Forderungen werden z.B. von allen Prozessen S der Form $(S_i = \exp(\sum_j \sigma_{ij} Y_j))_{i=1,2,\ldots,d}$, Y $\mathbb{R}^m$-wertiges Semimartingal, $\sigma \in \mathbb{R}^{d\times m}$, erfüllt ($\mathcal{C}^2$-Transformationen von Semimartingalen sind wieder Semimartingale, siehe Itô-Formel).

Wir gehen also im weiteren davon aus, daß wir als Preisprozeß einen Prozeß S auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ mit der von S erzeugten Filtration $(\mathcal{F}_t)_{t\geq 0}$ haben, dessen Komponenten nichtnegative Semimartingale sind.

Bemerkung 3.1.1 *Zwar kann man sowohl mit Aktien („short selling") als auch mit dem Bankguthaben Schulden eingehen, dies äußert sich aber in einer negativen Anzahl von gehaltenen Einheiten, der Wert einer Einheit ist nach wie vor positiv.*

Bemerkung 3.1.2 *Wir gehen davon aus, daß der Preis der Aktie für beliebige Zeiten definiert ist. Das Vorhandensein eines endlichen Zeithorizonts wird in der Regel durch die zulässigen Handelsstrategien und das Zielfunktional modelliert (s.u.).*

3.1.2 Der Handelsprozeß φ und abgeleitete Größen

Als nächstes sind die Investitionsentscheidungen als stochastische Prozesse zu modellieren. Diese Prozesse sollten auf jeden Fall adaptiert (sonst könnte man Informationen aus der Zukunft in die Strategie einfließen lassen) und integrabel (also im allgemeinsten Fall vorhersehbar im Sinne des Kapitels über stochastische Integration) sein.

Wir verwenden einen adaptierten vorhersehbaren Prozeß $\varphi = (\varphi^i)_{i=0,1,2,\ldots,d}$, den Handelsprozeß (auch Handelsstrategie genannt), und interpretieren φ_t^i als die Anzahl der in Anlage i zum Zeitpunkt t gehaltenen Aktien (bzw. im Fall $i = 0$ die Anzahl der Einheiten im Bankguthaben). Je nachdem, ob man

einen endlichen oder unendlichen Zeithorizont betrachtet, verwendet man als Indexmenge für φ entweder $[0, T]$ oder $[0, \infty)$.

Es ist intuitiv klar, daß ein Modell, in dem beliebige Handelsprozesse möglich sind, die Realität unzureichend abbildet. Man benötigt daher Einschränkungen. Da diese in den vorhandenen Modellen unterschiedlich sind, wird in unserem allgemeinen Modell nur gefordert, daß $\varphi \in \Phi$ gelten muß, wobei Φ eine im jeweiligen Modell zu spezifizierende Klasse von adaptierten und vorhersehbaren Handelsprozessen ist.

Bemerkung 3.1.3 *Die Menge Φ ist oftmals dadurch gegeben, daß man zunächst Prozesse φ' einführt, an diese bestimmte Anforderungen stellt (kurz: $\varphi' \in \Phi'$) und dann Φ als Menge derjenigen Handelsprozesse φ definiert, die man aus den $\varphi' \in \Phi'$ durch Anwendung einer bestimmten Transformation $T : \Phi' \to \Phi$ gewinnt. Diese Technik wird vorzugsweise verwendet, wenn die Angabe von φ' den Bezug des Modells zur Realität klarer macht und/oder die Menge Φ'' einfacher zu beschreiben ist als Φ (vgl. Beispiele)*

Bemerkung 3.1.4 *Im folgenden wird immer vorausgesetzt, daß φ adaptiert und vorhersehbar ist.*

Um die Klasse der zulässigen Handelsprozesse zu beschreiben (und auch um einen Handelsprozeß zu bewerten), ist es sinnvoll, verschiedene, aus den Größen φ und S abgeleitete, Prozesse zu definieren:

Der Vermögensprozeß

Der Vermögensprozeß $(V^{\varphi,i})_{i=0,1,\ldots,d}$ zu einem bestimmten Handelsprozeß φ ergibt sich als

$$V_t^{\varphi,i} = \varphi_t^i S_t^i, \quad i = 0, 1, 2, \ldots, d \tag{3.1}$$

und der Gesamtvermögensprozeß V^φ ist $V^\varphi = \sum_{i=0}^d V^{\varphi,i}$, kurz auch V genannt.

Bemerkung 3.1.5 *Oft fordert man $V^\varphi \geq const$ für zulässige φ, dabei kann entweder const festgelegt sein oder aber ggf. von φ abhängen. Ist $const = 0$, so schließt dies die Möglichkeit aus, Schulden zu machen (zumindest im Hinblick auf das gesamte Vermögen)*

Der Investitionsprozeß ohne Transaktionskosten

Kann man φ beliebig (abgesehen von den bereits erwähnten „technischen" Einschränkungen) wählen, so muß man im Lauf der Zeit offenbar entweder zusätzliche Mittel aufwenden oder es bleiben Mittel übrig: In einem zeitdiskreten Modell ($t = 0, 1, 2, \ldots, T$) gilt: $V_t^{\varphi,i} - V_{t-1}^{\varphi,i} = \varphi_t^i S_t^i - \varphi_{t-1}^i S_{t-1}^i = \varphi_t^i S_t^i - \varphi_{t-1}^i S_t^i + \varphi_{t-1}^i S_t^i - \varphi_{t-1}^i S_{t-1}^i = (\varphi_t^i - \varphi_{t-1}^i) S_t^i + \varphi_{t-1}^i (S_t^i - S_{t-1}^i)$, d.h. der Vermögenszuwachs in Anlage i setzt sich zusammen aus Zuwachs aufgrund von Neuinvestition in Anlage i ($(\varphi_t^i - \varphi_{t-1}^i) S_t^i$) und Zuwachs aufgrund der Preisänderung am Markt $\varphi_{t-1}^i (S_t^i - S_{t-1}^i)$. Der Zuwachs $I_t - I_{t-1}$ des gesamten Investitionsvolumens I_t (mit $I_0 := 0$) beträgt somit (in Abwesenheit von Transaktionskosten) $I_t - I_{t-1} := \sum_{i=0}^d (\varphi_t^i - \varphi_{t-1}^i) S_t^i = (\varphi_t^\top S_t - \varphi_{t-1}^\top S_{t-1}) - \sum_{i=0}^d \varphi_{t-1}^i (S_t^i - S_{t-1}^i)$. Ist er positiv, so investiert man, ist er negativ, so verbraucht man Geld.

Die Verallgemeinerung dieser Größe auf den zeitstetigen Fall lautet (unter Beibehaltung der Bezeichnung I_t, ggf. auch als I^φ bezeichnet):

$$I_t := \varphi_t^\top S_t - \int_0^t \varphi^\top dS, \quad t \geq 0 \tag{3.2}$$

Analog ist auch die Investition in Aktie i definiert:

$$I_t^i := \varphi_t^{i\top} S_t^i - \int_0^t \varphi^{i\top} dS^i, \quad t \geq 0 \tag{3.3}$$

Bemerkung 3.1.6 *In dieser Definition ist $I_0 = 0$, d.h. die anfängliche Investition $\varphi_0^\top S_0$ wird nicht zu I gezählt.*

Bemerkung 3.1.7 *Oft fordert man $I_t = 0$ für zulässige φ. Diese Bedingung bezeichnet man auch als Selbstfinanzierbarkeitsbedingung.*

Der Investitionsprozeß unter Transaktionskosten

Will man Transaktionskosten modellieren, bietet sich eine Modifizierung des eben eingeführten Investitionsprozesses an. Man orientiert sich dabei wieder am diskreten Fall: $(\varphi_t^i - \varphi_{t-1}^i) S_t^i$ waren hier die Kosten für die Aktien selbst. Die hinzukommenden Transaktionskosten hängen im Allgemeinen von der Aktiensorte sowie vom Handelsvolumen (in Geldeinheiten) $(\varphi_t^i - \varphi_{t-1}^i) S_t^i$, nicht aber direkt von der Stückzahl oder dem Preis ab (tun sie dies dennoch,

so könnte man auch das in naheliegender Weise modellieren). Im diskreten Fall ergibt sich damit die Neuinvestition $I_t^i - I_{t-1}^i$ in Aktie i zu

$$\begin{aligned}
&I_t^i - I_{t-1}^i \\
&= (1 + f_i((\varphi_t^i - \varphi_{t-1}^i)S_t^i))(\varphi_t^i - \varphi_{t-1}^i)S_t^i \\
&= (1 + f_i(\varphi_t^i S_t^i - \varphi_{t-1}^i S_{t-1}^i - \varphi_{t-1}^i(S_t^i - S_{t-1}^i)))(\varphi_t^i S_t^i - \varphi_{t-1}^i S_{t-1}^i - \varphi_{t-1}^i(S_t^i - S_{t-1}^i))
\end{aligned}$$

wobei f_i die Transaktionskosten für Aktie i in Abhängigkeit vom Transaktionsvolumen darstellt (der gewöhnliche Investitionsprozeß ergibt sich mit $f = 0$). Ein Problem ergibt sich nun beim Auflösen der Gleichungen für $t = 1, 2, \ldots, T$ nach I_T: Ist f_i konstant, so fallen die Terme $\varphi^i S^i$ heraus („Teleskopsumme"). Beim Übergang zum zeitstetigen Fall ist das hilfreich, da man die (für allgemeines φ gar nicht definierte) Größe $d(\varphi^i S^i)$ bzw. daraus gebildete stochastische Integrale vermeidet. Der Fall $f_i = 0$ ergibt dabei den Spezialfall des Investitionsprozesses ohne Transaktionskosten, andere konstante f_i ergeben im Endeffekt lediglich mit dem entsprechenden Faktor multiplizierte Preisprozesse. Will man daher I für allgemeinere $(f_i)_{i=0,1,\ldots,d}$ definieren, so muß man fordern, daß $\varphi^i S^i$ ein Semimartingal ist (z.B. indem man fordert, daß φ selbst ein Semimartingal ist). Eine weitere Schwierigkeit ist, daß das Argument von f_i im stetigen Fall ein stochastisches Differential wird. Eine Lösung für dieses Problem ist, den Term $(\varphi_t^i - \varphi_{t-1}^i)$ als $\frac{(\varphi_t^i - \varphi_{t-1}^i)}{t-(t-1)}$ aufzufassen, d.h. als Differenzenquotient, und als Argument im stetigen Fall $\frac{d\varphi}{dt}$ zu verwenden. Diese Definition setzt Differenzierbarkeit von φ voraus.

Die üblichen Modelle mit Transaktionskosten fordern daher stetige Differenzierbarkeit, womit ebenfalls gesichert ist, daß φ ein Semimartingal ist. Die Größe I ergibt sich dann als:

$$I_t = \sum_{i=0}^{d} \int_{0+}^{t} (1 + f_i(\frac{d\varphi_s^i}{ds} S_s^i)) S_s^i d\varphi_s^i \tag{3.4}$$

Dabei werden wieder die einzelnen Summanden bei Bedarf unter der Bezeichnung I^i verwendet. Oft wählt man f_i stückweise konstant auf $(-\infty, 0)$ und $[0, \infty)$ (d.h. Transaktionskosten proportional zum Handelsvolumen, aber ggf. unterschiedlich für An- und Verkauf).

Bemerkung 3.1.8 *Oft fordert man $I_t \leq 0$ für zulässige φ, d.h. es ist nur Verbrauch möglich, keine Neuinvestition.*

Bemerkung 3.1.9 *Sofern nicht anders erwähnt, versteht man im folgenden unter I immer den Investitionsprozeß ohne Transaktionskosten.*

Bemerkung 3.1.10 *In jedem sinnvollen Modell sind die Funktionen $x \to (1+f_i(x))x$ streng monoton steigend (ansonsten wären u.U. ein Aktienpaket mit höherem Marktwert billiger als ein anderes) und insbesondere sind diese Funktionen invertierbar. Diese Umkehrfunktionen seien mit g_i bezeichnet, d.h. es gilt $g_i(\frac{dI_s^i}{ds}) = \frac{d\varphi_s^i}{ds}$.*

Der Portfolioprozeß

Eine weitere abgeleitete Größe, die bisweilen alternativ zum Handelsprozeß verwendet wird, ist der Portfolioprozeß π. Er ist definiert über

$$\pi^i = \frac{V^{\varphi,i}}{V^\varphi}, \quad i = 0, 1, 2, \ldots, d \tag{3.5}$$

und beschreibt den Anteil des Vermögens in Aktie i am Gesamtvermögen. Offenbar genügt es, π^i für $i = 1, 2, \ldots, d$ zu kennen, π^0 ergibt sich dann über $\sum_{i=0}^d \pi^i = 1$.

3.1.3 Das Bewertungsfunktional

Die letzte noch fehlende Komponente des Modells ist die Bewertung eines zulässigen Kontrollprozesses. Als Grundlage hierzu bietet sich der Vermögens- und Investitionsprozeß an.

Will man den Investitionsprozeß zur Bewertung heranziehen und dabei den Konsum, d.h. negative Investition, zeitlich unterschiedlich bewerten (z.B. mit einem Abzinsungsfaktor) oder mit einer Nutzenfunktion gewichten, so stößt man wiederum auf ähnliche Probleme, wie sie bereits bei der Definition des Investitionsprozesses unter Transaktionskosten auftraten. Fordert man wieder von φ die stetige Differenzierbarkeit, so kann man als Zielfunktional

$$E[\int_0^T e^{-\beta s} U(-\frac{dI_s}{ds}) ds] \tag{3.6}$$

wobei ggf. $T = \infty$ ist, verwenden.

Verwendet man den Vermögensprozeß (genauer gesagt, seinen Wert zum Endzeitpunkt), so umgeht man diese Schwierigkeiten, indem man als Zielfunktional

$$E[U(V_T^{\varphi})] \tag{3.7}$$

wählt.

Dabei ist $U : \mathbb{R} \to \mathbb{R}$ in beiden Fällen eine monoton steigende, konkave Funktion , die man auch als Nutzenfunktion („utility function") bezeichnet (oft wird z.B. $U(x) = 1_{[0,\infty)} \ln x$ oder $U(x) = 1_{[0,\infty)} \frac{x^\gamma}{\gamma}$ verwendet). Sie erlaubt eine variable Gewichtung des erworbenen Reichtums. Der Faktor $e^{-\beta s}$, $\beta \geq 0$ dient als Abzinsung (im zweiten Fall ist er nicht nötig, da nur der Endzeitpunkt betrachtet wird; auch im ersten Fall kann man $\beta = 0$ wählen, wenn man bereits den Preisprozeß abgezinst hat). Unter Umständen verwendet man auch die Summe dieser beiden Funktionalen mit möglicherweise verschiedenen Nutzenfunktionen und verschiedenen Abzinsungsfaktoren, d.h.

$$E[\int_0^T e^{-\beta_1 s} U_1(-\frac{dI_s}{ds}) ds + U_2(V_T^{\varphi})] \tag{3.8}$$

3.1.4 Zusammenfassung des allgemeinen Modells

Gegeben sei ein nichtnegatives, $\mathbb{R}^{d+1}$-wertiges Semimartingal S, dessen nullte Komponente deterministisch sei und eine Menge Φ von adaptierten, vorhersehbaren $\mathbb{R}^{d+1}$-wertigen Prozessen φ. Weiterhin sei ein Funktional $F : \Phi \to \mathbb{R}$ mit der in Abschnitt 3.1.3 beschriebenen Form gegeben.

Das zu lösende Problem lautet dann:

$$\text{Finde } \varphi^* \in \Phi, \text{ so daß } F(\varphi^*) = \sup_{\varphi \in \Phi} F(\varphi)$$

3.2 Beispiele

In diesem Abschnitt werden einige Modelle, die zum Teil auch in späteren Kapiteln vorkommen, vorgestellt und auf das oben eingeführte allgemeine Modell zurückgeführt. Wie aus dem vorigen Abschnitt hervorgeht, sind die wesentlichen Merkmale des konkreten Modells die Wahl des Preisprozesses S, die Menge Φ der zulässigen Kontrollprozesse sowie die Wahl des Bewertungsfunktionals F. Die vorgestellten Modelle finden sich z.B. in [13], [11] und[12].

3.2.1 Das klassische Modell von Black/Scholes/Merton ohne Transaktionskosten

Dieses Modell wurde in den Anfangszeiten der Portfoliooptimierung gern verwendet, da es relativ leicht zu behandeln ist. Es hat allerdings den Nachteil, daß man durch die Festlegung auf die Brownsche Bewegung als Preisprozeß stark eingeschränkt ist und Transaktionen nichts kosten.

Der Preisprozeß S

Es gilt

$$\begin{aligned} dS_t^0 &= S_t^0 r dt \\ dS_t^i &= S_t^i(\alpha^i dt + \sum_{l=1}^{m} \sigma^{il} dw_t^l), \quad i = 1, 2, \ldots, d \end{aligned}$$

mit einer m-dimensionalen Standard Brownschen Bewegung w, einem festen Zinssatz r, einer (vektorwertigen) Driftkonstante α und einer (matrixwertigen) Varianzkonstante σ. Damit ist $S_t^0 = S_0^0 e^{-rt}$ und S eine geometrische Brownsche Bewegung.

Die zulässigen Handelsprozesse

Es wird gefordert, daß das Gesamtvermögen nicht negativ wird und daß kein Geld nachinvestiert wird, d.h. $V \geq 0$ und $I = 0$. Somit gilt

$$\Phi = \{\varphi | V^\varphi \geq 0, I^\varphi = 0\}$$

Das Bewertungsfunktional

Man verwendet das Funktional $F(\varphi) = E[V_T^\varphi]$. Will man den Erwartungswert unter Beschränkung der Varianz des Endvermögens optimieren, so kann man die entsprechende Forderung $Var(V_T^\varphi) \leq C$ in der Definition von Φ aufstellen.

3.2.2 Das Modell von Black/Scholes/Merton mit Transaktionskosten

Der Preisprozeß S

Es wird derselbe Preisprozeß verwendet wie im Modell ohne Transaktionskosten.

Die zulässigen Handelsprozesse

Die zulässigen Handelsprozesse werden durch Beschränkungen für den Investitionsprozeß unter Transaktionskosten beschrieben. Daher nehmen wir φ als differenzierbar an. Die Funktion f, mit deren Hilfe der Investitionsprozeß definiert ist, ist gegeben durch:

$$f_0 = 0$$
$$f_i(x) = \begin{cases} \lambda^{0i} & \text{falls} \quad x \geq 0 \\ -\lambda^{i0} & \text{falls} \quad x < 0 \end{cases}$$

Dabei interpretiert man λ^{0i} als Kosten pro Aktieneinheit beim Kauf der Aktie i und λ^{i0} als Kosten beim Verkauf (d.h. $-\lambda^{i0}$ als Einnahmen beim Verkauf). Ökonomisch sinnvoll ist sicherlich die Forderung $|\lambda^{i0}| < 1$ und $|\lambda^{0i}| < 1$ (da ansonsten die Gebühr das Transaktionsvolumen übersteigen würde) sowie $\lambda^{0i} \geq -\lambda^{i0}$ (ansonsten wären die Gebühreneinnahmen beim Verkauf größer als die Gebührenausgaben beim Ankauf, man könnte also allein durch Kaufen und sofortiges Wiederverkaufen Geld verdienen). Beide Forderung sind beispielsweise bei Gültigkeit von $0 \leq \lambda^{i0}, \lambda^{0i} < 1$ (Interpretation: sowohl Kauf als auch Verkauf kosten etwas) trivialerweise erfüllt.

Man fordert außerdem, daß nichts nachinvestiert wird.

Somit gilt $\Phi = \{\varphi |\ \varphi \text{ stetig differenzierbar}, I^\varphi \leq 0\}$

Das Bewertungsfunktional

Man bewertet nur bis zum Austritt des Prozesses $(V^{\varphi,i})_{i=0,1,\ldots,d}$ aus $G := \{x | x^0 + \sum_{i=1}^d \min\{(1-\lambda^{i0})x^i, (1+\lambda^{0i})x^i\} \geq 0\}$. Die Wahl dieses G kann man wie folgt begründen: Man fordert, daß bei einer sofortigen Liquidierung der Aktienkonten ein nichtnegatives Vermögen übrigbleibt (diese Forderung ist zwar „modelltechnisch" nicht konsistent mit der Annahme einer stetig differenzierbaren Handelsstrategie, dies stellt aber kein mathematisches Problem

dar, da die Überlegung nur in den Bewertungsvorgang einfließt, nicht aber zur Bestimmung der zulässigen Handelsstrategien herangezogen wird). Die dazu nötige Investition in Aktie i beträgt dann $(1+\lambda^{0i})(-x^i)$, falls x^i negativ ist. Ist x^i positiv, ist die Investition $(1-\lambda^{i0})(-x^i)$. Die Fallunterscheidung umgeht man durch Verwendung des Terms $-\min\{(1+\lambda^{0i})x^i, (1-\lambda^{0i})x^i\}$, der in jedem Fall mit der Investition übereinstimmt. Das Gesamtvermögen nach der Liquidierung ist dann die Summe aus dem Bankvermögen und den negativen Investitionen.

Als bewertendes Funktional verwenden wir $F(\varphi) = E[\int_0^\tau e^{-\beta_1 s} U_1(-\frac{dI_s}{ds})ds + U_2(V_\tau^\varphi)]$ mit $\tau := \inf\{t \geq 0 : V_t^\varphi \notin G\}$

3.2.3 Ein Modell von Kallsen [8]

Der Preisprozeß S

Als Preisprozeß S wird ein d-dimensionaler exponentieller Lévy-Prozeß verwendet, d.h. es gilt für $i = 1, 2, \ldots, d$:

$$S_t^i = \exp(L_t^i), \tag{3.9}$$

mit einem d-dimensionalen Lévy-Prozeß L.

Bemerkung 3.2.1 *Die Komponenten von L können in diesem Modell untereinander abhängig sein. Ist L eine Brownsche Bewegung mit Drift, so ist L auf jeden Fall eine Lineartransformation eines anderen Lévy-Prozesses mit unabhängigen Kompnenten, d.h. die Abhängigkeitsstruktur ist relativ einfach. Will man dies auch für allgemeine Lévy-Prozesse erreichen, so fordert man $L = \sigma\widetilde{L}$ mit einem Prozeß $\widetilde{L}$, dessen Komponenten untereinander unabhängig sind.*

Die zulässigen Handelsprozesse

In diesem Modell wird wie im Black-Scholes-Modell ohne Transaktionskosten vorausgesetzt, daß

$$\varphi_t^\top Y_t = \varphi_0^\top Y_0 + \int_{0+}^t \varphi_{s-}^\top dY_s, \quad t \geq 0 \tag{3.10}$$

gilt, d.h. $I_t = 0$, wobei I hier wieder der Investitionsprozeß ohne Transaktionskosten ist. Ebenso wird wiederum verlangt, daß $V^\varphi \geq 0$ gilt, d.h. daß keine Schulden gemacht werden können.

Somit gilt $\Phi = \{\varphi | V^\varphi \geq 0,\ I^\varphi = 0\}$

Das Bewertungsfunktional

Als Bewertungsfunktional verwendet dieses Modell $F(\varphi) = E[U(V_T^{\varphi})]$.

Kapitel 4

Stochastische Kontrolltheorie und die HJB-Gleichung

In diesem Kapitel wird gezeigt, wie man die Ergebnisse der stochastischen Kontrolltheorie auf das Portfoliooptimierungsproblem anwenden kann. Dies wird nicht in allen Fällen möglich sein: man wird gegenüber dem vorgestellten allgemeinen Modell z.B. die Klasse der möglichen Preisprozesse einschränken müssen.

Wir formulieren zunächst die Problemstellung der stochastischen Kontrolltheorie, leiten dann die HJB-Gleichungen (=Hamilton-Jacobi-Bellman-Gleichungen) ab, definieren den Begriff einer „schwachen" Lösung (viscosity solution), und zeigen abschließend, wie man die erhaltenen Ergebnisse auf die Portfoliooptimierung anwenden kann. Die Darstellung des ersten Abschnitts folgt im Wesentlichen Fleming [7] und Øksendal [15].

4.1 Stochastische Kontrolltheorie

Als Grundlage der folgenden Überlegungen dient der Abschnitt 1.3 über Markov-Prozesse und Diffusionen.

4.1.1 Die Problemstellung

Endlicher Zeithorizont

Wir gehen von einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t_0 \leq t \leq t_1}, P)$ mit einer d-dimensionalen Standard Brownschen Bewegung w aus. Weiter-

hin seien Zeitpunkte t_0, t_1 gegeben sowie Funktionen $f : \overline{Q_0} \times U \to \mathbb{R}^n$ und $\sigma : \overline{Q_0} \times U \to \mathbb{R}^{n\times d}$ mit $Q_0 := [t_0, t_1) \times \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$ abgeschlossen, so daß $f(.,.,v)$ und $\sigma(.,.,v)$ stetig differenzierbar sind für alle $v \in U$ und es eine Konstante C gibt, so daß gilt:

$$\begin{aligned} |f_t| + |f_x| \leq C, \quad |\sigma_t| + |\sigma_x| \leq C \\ |f(t,x,v)| \leq C(1 + |x| + |v|), \quad |\sigma(t,x,v)| \leq C(1 + |x| + |v|) \end{aligned} \tag{4.1}$$

Dabei ist $|.|$ die euklidische Norm bzw. die Operatornorm.

Sei weiter ein Anfangszeitpunkt $t \in [t_0, t_1)$ und ein Startwert $x \in \mathbb{R}^n$ gegeben. Ein progressiv meßbarer Prozeß $u : [t, t_1] \times \Omega \to U$ heißt zulässig, falls $E[\int_t^{t_1} |u_s|^m ds] < \infty$ für $m \in \mathbb{N}$ (dies gilt z.B. für kompaktes U). Dann hat die Differentialgleichung

$$dx = f(s, x_s, u_s)ds + \sigma(s, x_s, u_s)dw_s, \quad t \leq s \leq t_1, x_t = x \tag{4.2}$$

mit einem zulässigen progressiv meßbaren Prozeß u nach Fleming [7] eine eindeutige progressiv meßbare Lösung. Dabei kann man w ohne Beschränkung der Allgemeinheit mit unabhängigem Komponenten wählen, ggf. ändere man σ entsprechend. Um zu optimieren, benötigen wir noch eine Bewertung des von u erzeugten Prozesses $(x_s)_{t\leq s\leq t_1}$. Sei dazu eine offene Menge $O \subseteq \mathbb{R}^n$ gegeben, $Q := [t_0, t_1) \times O$ und $\tau := \inf\{s : (s, x_s) \notin Q\}$. Weiterhin seien stetige Funktionen $L : Q \times U \to \mathbb{R}$ und $\psi : ([t_0, t_1) \times \partial O) \cup (\{t_1\} \times \overline{O}) \to \mathbb{R}$ mit

$$\begin{aligned} |L(t,x,v)| \leq C(1 + |x|^k + |v|^k) \\ |\psi(t,x)| \leq C(1 + |x|^k) \end{aligned}$$

mit geeigneten Konstanten C, k gegeben. Sei weiter

$$J(t,x;u) = E[\int_t^\tau L(s, x_s, u_s)ds + \psi(\tau, x_\tau)]$$

Dann heißt eine zulässige Kontrolle u^* optimal, falls $J(t,x;u^*) = \sup_u J(t,x;u)$. Betrachtet man das Supremum als Funktion von t und x, so spricht man von der Wertfunktion W. Es zeigt sich, daß man unter bestimmten Bedingungen optimale Kontrollen finden kann, indem man zuerst eine Differentialgleichung für W aufstellt, diese löst und dann mit Hilfe von W u^* bestimmt. Diese Zusammenhänge werden in den nächsten Unterabschnitten behandelt.

Unendlicher Zeithorizont

Gegenüber endlichem Zeithorizont ergeben sich in diesem Fall einige Änderungen: Die Differentialgleichungskoeffizienten und L, ψ werden nun als zeitunabhängig (genauer: nicht explizit zeitabhängig) angenommen, wodurch die partiellen Ableitungen nach t in den Bedingungen an f und σ wegfallen. Als Startwert wird 0 angenommen, τ kann nun ∞ sein, und als Bewertungsfunktional verwendet man

$$J(x;u) = E[\int_0^\tau e^{-\beta s} L(x_s, u_s) ds + 1_{[0,\infty)}(\tau) e^{-\beta\tau} \psi(x_\tau)]$$

Als zulässig bezeichnen wir einen progressiv messbaren Prozeß, der neben der Ungleichung $E[\int_0^t |u_s|^m ds] < \infty$ für $m \in \mathbb{N}, t \in [0,\infty)$ auch $E_x[\int_0^\tau e^{-\beta s} |L(x_s^u, u_s)| ds] < \infty$ erfüllt.

4.1.2 Die Lösungsidee

In diesem Unterabschnitt wird die Idee, die der HJB-Gleichung und den Beweisen zu den folgenden Theoremen zugrundeliegen, im Fall des endlichen Zeithorizonts vorgestellt. Für eine rigorose Herleitung sei auf Fleming[7] verwiesen.

Im folgenden wird der von der Steuerung u erzeugte Prozeß als x^u bezeichnet. Zu Anfangswerten t, x sei die dazugehörige optimale Strategie mit $u^{*t,x}$ bezeichnet.

Es gilt:

$$W(t,x) \geq E[\int_t^{t+h} L(s, x_s^v, v) ds + W(t+h, x_{t+h}^v)]$$

Diese Ungleichung erhält man, indem man für die rechte Seite eine Kontrolle wählt, die bis zum Zeitpunkt h konstant v ist, und danach die optimale Strategie für die (zufallsabhängigen!) Anfangsdaten $(t+h, x_{t+h}^v)$. Dabei muß bei exakter Behandlung natürlich sichergestellt werden, daß ein solches Vorgehen wieder eine zulässige Kontrolle erzeugt. Auch muß $h < \tau$ gefordert werden, ansonsten muß in der obigen Ungleichung ein Term mit ψ dazugefügt werden, der dann möglichst mit $h \to 0$ verschwindet. Wir gehen nun von der Gültigkeit der Ungleichung aus. Umordnen, Multiplikation mit h^{-1} und Grenzwertbildung ergeben dann (wiederum unter geeigneten Regularitätsbedingungen)

$$\lim_{h\to 0} h^{-1} E[W(t+h, x_{t+h}^v) - W(t,x)] + L(t,x,v) \leq 0$$

bzw. kurz

$$A^v W(t,x) + L(t,x,v) \leq 0$$

wobei wir analog zur Definition des Operators A (vgl. 1.7) für eine Steuerung u den Operator A^u durch $(A^u\Phi)(t,x) := \lim_{h\to 0} h^{-1}E[\Phi(t+h, x^u_{t+h}) - \Phi(t,x)]$ definieren, sofern der Grenzwert existiert. Der Operator kann mit Hilfe der Itô-Formel berechnet werden, wenn man $y^v_s = (x^v_s, s) =: (y^{v,i}_s)_{i=1,2,\ldots,n+1}$ betrachtet:

$$\begin{aligned}
&\Phi(y^v_{t+h}) - \Phi(t,x) \\
&= \sum_{i=1}^{n+1} \int_t^{t+h} \partial_i \Phi(y^v_s) dy^{v,i}_s + \frac{1}{2} \sum_{i,j=1}^{n+1} \int_t^{t+h} \partial_{i,j} \Phi(y^v_s) d[y^{v,i}, y^{v,j}] \\
&= \int_t^{t+h} \Phi_t(y^v_s) ds + \sum_{i=1}^{n} \int_t^{t+h} \Phi_{x_i}(y^v_s) dx^{v,i}_s + \frac{1}{2} \int_t^{t+h} \Phi_{tt} d[s,s] \\
&\quad + \sum_{i=1}^{n} \Phi_{tx_i}(y^v_s) d[x^{v,i}, s] + \frac{1}{2} \sum_{i,j=1}^{n} \int_t^{t+h} \Phi_{x_i x_j}(y^v_s) d[x^{v,i}, x^{v,j}]
\end{aligned}$$

Verwendet man nun, daß $d[s,s] = d[w^k, s] = 0$ (vgl. Protter [17] Theorem II.28), $d[w_k, w_l] = \delta_{lk} ds$ und $d[x^{v,i}, s] = f_i(y^v_s) d[s,s] + \sum_{k=1}^d \sigma_{ik}(y^v_s) d[w_k, s]$

usw. ist erhält man:

$$
\begin{aligned}
&\int_t^{t+h} \Phi_t(y_s^v)ds + \sum_{i=1}^n \int_t^{t+h} \Phi_{x_i}(y_s^v)dx_s^{v,i} \\
&\quad + \frac{1}{2}\sum_{i,j=1}^n \int_t^{t+h} \Phi_{x_ix_j}(y_s^v)d[x^{v,i}, x^{v,j}] \\
&= \sum_{i=1}^n \int_t^{t+h} (\Phi_t(y_s^v) + \Phi_{x_i}(y_s^v)f_i(y_s^v, v))ds \\
&\quad + \sum_{i=1}^n \sum_{l=1}^d \int_t^{t+h} \Phi_{x_i}(y_s^v)\sigma_{il}(y_s^v, v)dw_l \\
&\quad + \frac{1}{2}\sum_{i,j=1}^n \sum_{k,l=1}^d \int_t^{t+h} \Phi_{x_ix_j}(y_s^v)\sigma_{ik}(y_s^v, v)\sigma_{jl}(y_s^v, v)d[w_k, w_l] \\
&\quad + \sum_{i,j=1}^n \sum_{k=1}^d \int_t^{t+h} \Phi_{x_ix_j}(y_s^v)f_j(y_s^v, v)\sigma_{ik}(y_s^v, v)d[s, w_k] \\
&\quad + \frac{1}{2}\sum_{i,j=1}^n \int_t^{t+h} \Phi_{x_ix_j}(y_s^v)f_j(y_s^v, v)f_i(y_s^v, v)d[s, s] \\
&= \sum_{i=1}^n \int_t^{t+h} (\Phi_t(y_s^v) + \Phi_{x_i}(y_s^v)f_i(y_s^v, v))ds \\
&\quad + \frac{1}{2}\sum_{i,j=1}^n \sum_{k=1}^d \int_t^{t+h} \Phi_{x_ix_j}(y_s^v)a_{ij}(y_s^v, v)ds \\
&\quad + \sum_{i=1}^n \sum_{l=1}^d \int_t^{t+h} \Phi_{x_i}(y_s^v)\sigma_{il}(y_s^v, v)dw_l
\end{aligned}
$$

Dabei ist $a = \sigma\sigma^\top$, f, σ die Koeffizienten der Differentialgleichung. Der letzte Term ist ein Martingal, verschwindet also beim Bilden des Erwartungswertes. Multiplikation mit h^{-1} und Grenzwertbildung liefert dann:

$$A^v\Phi = \Phi_t + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t, x, v)\Phi_{x_ix_j} + \sum_{i=1}^n f_i(t, x, v)\Phi_{x_i}$$

Wählt man $v = u^*(t)$ (die rechte Seite ist zufallsabhängig, also ist wieder eine genauere Betrachtung nötig), so erwartet man, daß die zusammenge-

setzte Strategie gegen u^* konvergiert (siehe Bemerkung 4.1.1), also die Ungleichung $A^v W(t,x) + L(t,x,v) \leq 0$ zur Gleichung wird. Es gilt damit

$$\sup_{v \in U}(A^v W(t,x) + L(t,x,v)) = 0 \quad \text{und}$$
$$u_t^* \in \{v^* \in U | A^{v^*} W(t,x) + L(t,x,v^*) = \sup_{v \in U}(A^v W(t,x) + L(t,x,v)\}$$

Diese Menge wird auch als $\arg\max[A^v W(t,x) + L(t,x,v)]$ bezeichnet. Definiert man weiter für $(t,x) \in \overline{Q_0}, p \in \mathbb{R}^n$, symmetrisches $M \in \mathbb{R}^{n \times n}$

$$\mathcal{H}(t,x,p,M) := \sup_{v \in U}[f(t,x,v)p + \frac{1}{2} \operatorname{tr} a(t,x,v)M + L(t,x,v)], \tag{4.3}$$

so erhält man die partielle Differentialgleichung 2. Ordnung

$$W_t = -\mathcal{H}(t,x,D_x W, D_x^2 W), \quad (t,x) \in Q \tag{4.4}$$

Die Randbedingungen sind

$$W(t,x) = \psi(t,x), \quad (t,x) \in \partial^* Q \tag{4.5}$$

Bemerkung 4.1.1 *Der exakte Beweis der HJB-Gleichung beruht unter anderem darauf, daß die der Dynamik zugrundeliegende Brownsche Bewegung als Markov-Prozeß „gedächtnislos" ist und daher stückweise optimale Kontrollen auch global optimal sind. Somit kann man den Versuch machen, lokal optimale Strategien zu einer global optimalen zusammenzusetzen. Andererseits geht aus dieser Überlegung aber auch hervor, daß diese Technik nur für Markov-Prozesse erfolgversprechend ist, nicht jedoch für allgemeine Semimartingale.*

4.1.3 Die HJB-Gleichung und ihre klassischen Lösungen

Endlicher Zeithorizont

Wie in Fleming [7] gezeigt wird, gilt mit den Bezeichnungen der Unterabschnitte 4.1.2 und 4.1.1 folgendes Theorem:

Theorem 4.1 *Sei W eine Funktion, die bezüglich x zweimal, bezüglich t einmal auf Q stetig differenzierbar ist, auf $\overline{Q}$ stetig und polynomial beschränkt ist und die die Gleichungen 4.4 und 4.5 erfüllt. Dann gilt:*

1. *$W(t,x) \geq J(t,x;u)$ für alle zulässigen u und alle (t,x)*
2. *Gibt es einen zulässigen Kontrollprozeß u^*, so daß*

$$u_s^* \in \arg\max[f(s,x_s^{u^*},v)D_xW(s,x_s^{u^*}) + \frac{1}{2}\operatorname{tr} a(s,x_s^{u^*},v)D_x^2W(s,x_s^{u^*}) + L(s,x_s^{u^*},v)] \tag{4.6}$$

für Lebesgue $\otimes P$- fast alle $(s,\omega) \in [t,\tau] \times \Omega$, dann gilt $W(t,x) = J(t,x,u^*)$. Dabei ist τ die Austrittszeit zu den Anfangswerten t,x und der Steuerung u^* und $[t,\tau] \times \Omega = \{(s,\omega)|t \leq s \leq \tau(\omega)\}$.

Beweis: siehe Fleming [7], Seite 163. ■

Aus der Theorie der partiellen Differentialgleichungen weiß man, daß z.B. folgende Bedingungen die Existenz und Eindeutigkeit einer Lösung mit den im Theorem geforderten Eigenschaften garantieren:

Bedingung 4.2
1. *Es gibt ein $c > 0$, so daß $\sum_{i,j=1}^n a_{ij}(t,x,v)\xi_i\xi_j \geq c|\xi|^2$ $\forall (t,x,v) \in Q_0 \times U, \xi \in \mathbb{R}^n$.*
2. *U ist kompakt.*
3. *O ist beschränkt und ∂O ist eine $\mathcal{C}^3$-Mannigfaltigkeit.*
4. *a, f, L sind auf $\overline{Q} \times U$ bezüglich t einmal, bezüglich x zweimal stetig differenzierbar.*
5. *ψ ist auf $\overline{Q_0}$ dreimal stetig differenzierbar.*

Auch andere Bedingungskombinationen, die hier nicht aufgeführt werden, garantieren Existenz und Eindeutigkeit.

Unendlicher Zeithorizont

In diesem Fall wird die mit $e^{-\beta t}$ abgezinste Wertfunktion als nicht von t abhängig vorausgesetzt und dementsprechend vereinfacht sich auch die HJB-Gleichung. Nach Fleming [7] gilt folgendes Theorem:

Theorem 4.3 *Sei $W : \overline{O} \to \mathbb{R}$ eine auf O zweimal stetig differenzierbare, auf $\overline{O}$ stetige und polynomial beschränkte Funktion, die folgende Gleichungen erfüllt:*

1. $\beta W(x) + \mathcal{H}(x, DW(x), D^2W(x)) = 0, \quad x \in O$

2. $W(x) = \psi(x), \quad x \in \partial O$

Dann gilt:

1. $W(x) \geq J(x\,;u)$ *für alle zulässigen* u *mit* $\liminf_{t_1\to\infty} E_x[1_{[t_1,\infty)}(\tau)W(x_{t_1})] \leq 0$ *und alle* x*, wobei* x_t *und* τ *die zum Steuerungsprozeß* u *gehörenden Größen sind.*

2. Gibt es einen zulässigen Kontrollprozeß u^**, so daß*

$$u_s^* \in \arg\max[f(x_s, v)DW(x_s) + \frac{1}{2}\operatorname{tr} a(x_s, v)D^2W(x_s) + L(x_s, v)]$$

für Lebesgue$\otimes P$*-fast alle* $(s, \omega) \in [0, \tau)\times\Omega$ *und* $\lim_{t_1\to\infty} E_x[1_{[t_1,\infty)}(\tau)W(x_{t_1})] = 0$*, wobei* x_s *und* τ *die zur Steuerung* u^* *gehörenden Größen sind, dann ist* $W(x) = J(x; u^*)$

Beweis: siehe Fleming [7], Seite 172. ■

4.1.4 Schwache Lösungen (Viscosity Solutions)

Im letzten Unterabschnitt wurde gezeigt, daß eine hinreichend glatte Lösung der HJB-Gleichung, zu der es eine entsprechende Steuerung gibt, die Wertfunktion sein muß. Es wurden auch Kriterien genannt, die die Existenz einer solchen Lösung garantieren. In Fällen, in denen es keine solche gibt, kann es aber zumindest eine Funktion geben, die einem verallgemeinerten Lösungsbegriff genügt. Unter Umständen ist dann diese Funktion die Wertfunktion. Wie auch im letzten Unterabschnitt werden hier nur die Ergebnisse von Fleming [7] angegeben.

Definition 4.1.2 *Sei* $\mathcal{S}^n$ *die Menge aller symmetrischen* $n \times n$*-Matrizen. Dann heißt*

$$D^{+(1,2)}W(t,x) = \{(p,q,A) \in \mathbb{R}^n \times \mathbb{R}\times\mathcal{S}^n :$$
$$\lim \sup_{\substack{(h,y)\to 0\\(t+h,x+y)\in Q}} \frac{W(t+h, x+y) - W(t,x) - qh - p^\top y - \frac{1}{2}y^\top Ay}{|h| + |y|^2} \leq 0\}$$

die Menge der Superdifferentiale im Punkt (t,x) *und*

$$D^{-(1,2)}W(t,x) = \{(p,q,A) \in \mathbb{R}^n \times \mathbb{R} \times \mathcal{S}^n :$$
$$\lim \inf_{\substack{(h,y)\to 0 \\ (t+h,x+y)\in Q}} \frac{W(t+h,x+y) - W(t,x) - qh - p^\top y - \frac{1}{2}y^\top Ay}{|h| + |y|^2} \geq 0\}$$

die Menge der Subdifferentiale im Punkt (t,x).

Offenbar ist bei einer hinreichend glatten Funktion das Tripel $(D_xW, \partial_t W, D_x^2 W)$ sowohl in der Menge der Super- als auch der Subdifferentiale.

Definition 4.1.3 *Eine Funktion* W *heißt schwache Sublösung (viscosity subsolution) der HJB-Gleichung, wenn*

$$q + \mathcal{H}(t,x,p,A,W(t,x)) \geq 0 \quad \forall (p,q,A) \in D^{+(1,2)}W(t,x) \quad \forall (t,x) \in Q \tag{4.7}$$

Eine Funktion W *heißt schwache Superlösung (viscosity supersolution), wenn*

$$q + \mathcal{H}(t,x,p,A,W(t,x)) \leq 0 \quad \forall (p,q,A) \in D^{-(1,2)}W(t,x) \quad \forall (t,x) \in Q \tag{4.8}$$

Eine Funktion, die zugleich schwache Sub- und Superlösung ist, heißt schwache Lösung (viscosity solution)

Im weiteren sind nun (wie im klassischen Fall) folgende Fragen von Bedeutung:

1. Unter welchen Bedingungen existiert eine schwache Lösung zur HJB-Gleichung und wann ist sie eindeutig (ggf. mit Randbedingungen)?

2. Wann ist eine schwache Lösung mit passenden Randbedingungen auch die Wertfunktion des zugrundeliegenden Kontrollproblems?

Im folgenden gehen wir nun von einem endlichen Zeithorizont aus.

Fleming [7] zeigt, daß die Wertfunktion eines Kontrollproblems, wenn sie stetig ist, auch eine schwache Lösung der HJB-Gleichung ist. Sofern also W stetig ist und eine Funktion, die Lösung der entsprechenden HJB-Gleichung sein und zusätzlich die Randbedingungen erfüllen soll, durch diese Forderungen bereits eindeutig festgelegt ist, handelt es sich bei dieser Lösung auch

um W. Im folgenden sind also noch folgende zwei Fragen zu klären: Welche Bedingungen sind hinreichend für die Stetigkeit von W, und unter welchen Umständen ist die Lösung der HJB-Gleichungen (mit Randbedingungen) eindeutig?

Nach Fleming gilt zur Frage der Stetigkeit folgendes:

Theorem 4.4 *Seien folgende Voraussetzungen erfüllt:*

1. *U ist kompakt, f und σ seien stetig und bezüglich t und x stetig differenzierbar*

2. *Für geeignete Konstanten C_1, C_2 gilt $|f_t| + |f_x| \leq C_1$, $|\sigma_t| + |\sigma_x| \leq C_1$, $|f(t,0,v)| + |\sigma(t,x,v)| \leq C_2$*

3. *Es gilt mit geeigneten Konstanten C_3, k: $|L(t,x,v)| \leq C_3(1+|x|^k)$*

4. *Es gibt eine dreimal stetig differenzierbare Funktion $g : \overline{Q_0} \to \mathbb{R}$, die selbst beschränkt ist und deren Ableitungen bis zur Ordnung drei beschränkt sind , die auf $[t_0, t_1) \times \partial O$ mit ψ übereinstimmt, auf $\{t_1\} \times \overline{O}$ kleiner oder gleich ψ ist und die außerdem*

$$g_t(t,x) + \mathcal{H}(t,x,D_x g(t,x), D_x^2 g(t,x)) \geq 0, \quad \forall (t,x) \in \overline{Q_0}$$

erfüllt.

Dann ist W stetig auf $\overline{Q}$
Beweis: *siehe Fleming [7], S.219* ■

Nun bleibt noch die Frage der Eindeutigkeit einer schwachen Lösung mit erfüllten Randbedingungen zu klären (die Existenz ist ja durch die Stetigkeit der Wertfunktion bereits gesichert). Es gilt:

Theorem 4.5 *Sei 4.1 erfüllt, sei außerdem U kompakt sowie f, σ, L, L_x, L_t beschränkt auf $Q_0 \times U$. Für zwei beschränkte schwache Lösungen der HJB-Gleichung (ohne Randbedingungen) V und W gilt dann:*

$$\sup_{\overline{Q_0}}(|W - V|) = \sup_{\mathbb{R}^n}(|W(t_1,.) - V(t_1,.)|)$$

Insbesondere ist dann eine Lösung, die zusätzlich die Randbedingungen erfüllt, eindeutig bestimmt.
Beweis: *siehe Fleming [7] S.249* ■

Im allgemeinen sind natürlich auch ohne die Voraussetzungen dieses Theorems schwache Lösungen der HJB-Gleichungen sinnvolle Kandidaten für die Wertfunktion. Je nach der Art des konkreten Modells muß man dann im Einzelfall Überlegungen und Beweise der Theoreme auf die spezielle Situation anpassen.

4.2 Kontrolltheorie für Lévy-Prozesse

Wie bereits erwähnt, kann die Überlegung der stochastischen Kontrolltheorie auch angewandt werden, wenn statt der Brownschen Bewegung w ein beliebiger Markov-Prozeß in der Differentialgleichung $dx = f(s, x_{s-}, u_s)ds + \sigma(s, x_{s-}, u_s)dw_s, \quad t \leq s \leq t_1, x_t = x$ verwendet wird (Die Verwendung von x_{s-} statt x_s sichert die Lösbarkeit der Differentialgleichung, für die Brownsche Bewegung als stetigen Prozeß spielt diese Unterscheidung keine Rolle). In diesem Kapitel wird daher versucht, die HJB-Gleichung für Lévy-Prozesse abzuleiten.

Wenn w nicht als Brownsche Bewegung, sondern als allgemeiner Lévy-Prozeß angenommen wird, kann man auch den Prozeß $A_t = t$ als erste Komponente zu einem vorhandenen Lévy-Prozeß hinzufügen, ebenso f als neue erste Spalte zu σ und erhält dann für die Differentialgleichung die Form $dx = \widetilde{\sigma}(s, x_{s-}, u_s)d\widetilde{w}_t$ (bzw. $x = \widetilde{\sigma}(., x_{.-}, u_.) \cdot \widetilde{w}$) mit $\widetilde{\sigma} = (f, \sigma)$ und $\widetilde{w} = (A, w)$. Wir nehmen daher ohne Beschränkung der Allgemeinheit $f = 0$ an. Die Dimension von x sei n, die von w sei m.

Das Prinzip der dynamischen Programmierung, d.h. die Ungleichung

$$W(t, x) \geq E[\int_t^{t+h} L(s, x_s^v, v)ds + W(t + h, x_{t+h}^v)]$$

gilt bei Verwendung von Lévy-Prozessen nach wie vor, da zu seiner Begründung nur die Markov-Eigenschaft der Brownschen Bewegung verwendet wurde. Ebenso ergibt sich nach wie vor die HJB-Gleichung als

$$\sup_{v \in U}(A^v W(t, x) + L(t, x, v)) = 0$$

allerdings muß nun der Operator $A^v\Phi := \lim_{h\to 0} h^{-1}E[\Phi(t+h, x_{t+h}^u) - \Phi(t, x)]$ im Fall von Lévy-Prozesse als der DGL zugrundeliegenden Prozessen berechnet werden. Es sei also nun w ein Lévy-Prozeß mit charakteristischem Tripel

$(\alpha, \beta\beta^\top, \nu)$. Dann gilt (wiederum mit der Abkürzung $y = (s, x)$)

$$\begin{aligned}
&\Phi(y^v_{t+h}) - \Phi(t,x) \\
&= \sum_{i=1}^{n+1} \int_t^{t+h} \partial_i \Phi(y^v_{s-}) dy^{v,i}_s + \frac{1}{2} \sum_{i,j=1}^{n+1} \int_t^{t+h} \partial_{i,j} \Phi(y^v_{s-}) d[y^{v,i}, y^{v,j}] \\
&\quad + \sum_{t<s\leq t+h} \{\Phi(y_s) - \Phi(y_{s-}) - \Phi'(y_{s-})\Delta y_s\}
\end{aligned}$$

Die drei Terme der rechten Seite werden nun berechnet. Dabei sei $w_t = at + \beta B_t + N_t$ die übliche Zerlegung von Levy-Prozessen in einen linearen deterministischen Teil, eine lineare Transformation einer d-dimensionalen Standard-Brownschen Bewegung und den Rest.

4.2.1 Berechnung von $\sum_{i=1}^{n+1} \int_t^{t+h} \partial_i \Phi(y^v_{s-}) dy^{v,i}_s$

Es gilt

$$\sum_{i=1}^{n+1} \int_t^{t+h} \partial_i \Phi(y^v_{s-}) dy^{v,i}_s = \int_t^{t+h} \partial_t \Phi(y^v_{s-}) ds + \int_t^{t+h} \nabla \Phi(y^v_{s-}) dx^v_s$$

Der erste Term der rechten Seite ist bereits in der gewünschten Form. Für den zweiten ergibt sich weiter

$$\begin{aligned}
&\int_t^{t+h} \nabla \Phi(y^v_{s-}) dx^v_s \\
&= \int_t^{t+h} \nabla \Phi(y^v_{s-}) d(\sigma(y^v_{\cdot-}, v) \cdot w) \\
&= \int_t^{t+h} \nabla \Phi(y^v_{s-}) \sigma(y^v_{s-}, v) dw_s \\
&= \int_t^{t+h} \nabla \Phi(y^v_{s-}) \sigma(y^v_{s-}, v) a ds + \int_t^{t+h} \nabla \Phi(y^v_{s-}) \sigma(y^v_{s-}, v) \beta dB_s \\
&\quad + \int_t^{t+h} \nabla \Phi(y^v_{s-}) \sigma(y^v_{s-}, v) dN_s
\end{aligned}$$

4.2.2 Berechnung von $\frac{1}{2}\sum_{i,j=1}^{n+1}\int_t^{t+h}\partial_{i,j}\Phi(y_{s-}^v)d[y^{v,i},y^{v,j}]$

Es gilt

$$\begin{aligned}
&\frac{1}{2}\sum_{i,j=1}^{n+1}\int_t^{t+h}\partial_{i,j}\Phi(y_{s-}^v)d[y^{v,i},y^{v,j}]\\
&= \frac{1}{2}\int_t^{t+h}\partial_{tt}\Phi(y_{s-}^v)d[s,s]+\sum_{i=1}^{n}\partial_{ti}\Phi(y_{s-}^v)d[x^{v,i},s]\\
&\quad +\frac{1}{2}\sum_{i,j=1}^{n}\int_t^{t+h}\partial_{ij}\Phi(y_{s-}^v)d[x^{v,i},x^{v,j}]
\end{aligned}$$

Die ersten beiden Terme der rechten Seite verschwinden wegen $d[s,s] = d[x^{v,i},s]=0$. Der letzte ergibt weiter

$$\begin{aligned}
&\frac{1}{2}\sum_{i,j=1}^{n}\int_t^{t+h}\partial_{ij}\Phi(y_{s-}^v)d[x^{v,i},x^{v,j}]\\
&= \frac{1}{2}\sum_{i,j=1}^{n}\int_t^{t+h}\partial_{ij}\Phi(y_{s-}^v)d[\sigma^{i\cdot}(y_{s-}^v,v)\cdot w,\sigma^{j\cdot}(y_{s-}^v,v)\cdot w]\\
&= \frac{1}{2}\sum_{i,j=1}^{n}\int_t^{t+h}\partial_{ij}\Phi(y_{s-}^v)d[\sum_{l=1}^{m}\sigma^{il}(y_{s-}^v,v)\cdot w^l,\sum_{k=1}^{m}\sigma^{jk}(y_{s-}^v,v)\cdot w^k]\\
&= \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\int_t^{t+h}\sigma^{il}(y_{s-}^v,v)\partial_{ij}\Phi(y_{s-}^v,v)\sigma^{jk}(y_{s-}^v,v)d[w^l,w^k]\\
&= \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\int_t^{t+h}\sigma^{il}(y_{s-}^v,v)\partial_{ij}\Phi(y_{s-}^v,v)\sigma^{jk}(y_{s-}^v,v)d[(a^lt)_{t\geq 0}\\
&\quad +\sum_{r=1}^{d}\beta^{lr}B^r+N^l,(a^kt)_{t\geq 0}+\sum_{p=1}^{d}\beta^{kp}B^p+N^k]
\end{aligned}$$

Nutzt man nun die Linearität des Variationsprozesses aus, so erhält man für festes i,j,k,l $(d+2)^2$ Variationsprozesse (alle Kombinationen der Summanden in $d[(a^lt)_{t\geq 0}+\sum_{r=1}^{d}\beta^{lr}B^r+N^l,(a^kt)_{t\geq 0}+\sum_{p=1}^{d}\beta^{kp}B^p+N^k]$). Alle Kombinationen der Form $d[s,s]d[s,B^k],d[s,N^k],d[B^k,N^l]$ sind als Kombinationen je eines „quadratic pure jump"-Prozesses (vgl. Protter [17], S.63 sowie Theorem II.28) und eines stetigen Prozessses 0. Kombinationen der

Form $d[B^r, B^p]$ ergeben 0 für $r \neq p$ und dt für $r = p$. Insgesamt ergibt sich dann

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}\int_t^{t+h} \beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v, v)\partial_{ij}\Phi(y_{s-}^v, v)\sigma^{jk}(y_{s-}^v, v)dt \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\int_t^{t+h} \beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v, v)\partial_{ij}\Phi(y_{s-}^v, v)\sigma^{jk}(y_{s-}^v, v)d[N^l, N^k] \\
&= \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}\int_t^{t+h} \beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v, v)\partial_{ij}\Phi(y_{s-}^v, v)\sigma^{jk}(y_{s-}^v, v)dt \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h} \beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v, v)\partial_{ij}\Phi(y_{s-}^v, v)\sigma^{jk}(y_{s-}^v, v)\Delta N_s^l \Delta N_s^k
\end{aligned}
$$

denn nach Protter [17], Theorem 28 und S.63 ist $[N^l, N^k]_t = \sum_{0<s\leq t} \Delta N^l \Delta N^k$

4.2.3 Berechnung von $\sum_{t<s\leq t+h}\{\Phi(y_s) - \Phi(y_{s-}) - \Phi'(y_{s-})\Delta y_s\}$

Es gilt

$$
\begin{aligned}
\Delta y_s &= \Delta(\sigma(y_{.-}, u)\cdot w)_s \\
&= \sigma(y_{s-}, u_s)\Delta w_s \\
&= \sigma(y_{s-}, u_s)\Delta N_s
\end{aligned}
$$

4.2.4 Berechnung von $A\Phi$

Wir setzen nun die berechneten Terme in $\frac{1}{h}E[\Phi(y_{t+h}^v) - \Phi(t,x)]$ ein und bilden den Grenzwert. Man erhält für den Term $\sum_{i=1}^{n+1}\int_t^{t+h}\partial_i\Phi(y_{s-}^v)dy_s^{v,i}$:

$$
\begin{aligned}
&\lim_{h\to 0}\frac{1}{h}E[\sum_{i=1}^{n+1}\int_t^{t+h}\partial_i\Phi(y_{s-}^v)dy_s^{v,i}] \\
&= \lim_{h\to 0}\frac{1}{h}E[\int_t^{t+h}\partial_t\Phi(y_{s-}^v)ds] + \lim_{h\to 0}\frac{1}{h}E[\int_t^{t+h}\nabla\Phi(y_{s-}^v)\sigma(y_{s-}^v, v)ads] \\
&\quad + \lim_{h\to 0}\frac{1}{h}E[\int_t^{t+h}\nabla\Phi(y_{s-}^v)\sigma(y_{s-}^v, v)\beta dB_s] + \lim_{h\to 0}\frac{1}{h}E[\int_t^{t+h}\nabla\Phi(y_{s-}^v)\sigma(y_{s-}^v, v)dN_s] \\
&= \partial_t\Phi(t,x) + \nabla\Phi(t,x)\sigma(t,x,v)a + \lim_{h\to 0}\frac{1}{h}E[\int_t^{t+h}\nabla\Phi(y_{s-}^v)\sigma(y_{s-}^v, v)dN_s]
\end{aligned}
$$

Für den Term $\frac{1}{2}\sum_{i,j=1}^{n+1}\int_t^{t+h}\partial_{i,j}\Phi(y_{s-}^v)d[y^{v,i},y^{v,j}]$ erhält man

$$\begin{aligned}
&\lim_{h\to 0}\frac{1}{h}E[\frac{1}{2}\sum_{i,j=1}^{n+1}\int_t^{t+h}\partial_{i,j}\Phi(y_{s-}^v)d[y^{v,i},y^{v,j}]] \\
&= \lim_{h\to 0}\frac{1}{h}E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}\int_t^{t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v,v)\partial_{ij}\Phi(y_{s-}^v,v)\sigma^{jk}(y_{s-}^v,v)dt] \\
&\quad + \lim_{h\to 0}\frac{1}{h}E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v,v)\partial_{ij}\Phi(y_{s-}^v,v)\sigma^{jk}(y_{s-}^v,v)\Delta N_s^l\Delta N_s^k] \\
&= \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v) \\
&\quad + \lim_{h\to 0}\frac{1}{h}E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v,v)\partial_{ij}\Phi(y_{s-}^v,v)\sigma^{jk}(y_{s-}^v,v)\Delta N_s^l\Delta N_s^k]
\end{aligned}$$

Der dritte Term $\lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\leq t+h}\{\Phi(y_s)-\Phi(y_{s-})-\Phi'(y_{s-})\Delta y_s\}]$ kann vorerst nicht weiter vereinfacht werden.

Um A weiter berechnen zu können, stellen wir nun folgende Forderung auf:

Forderung 4.6 *N ist ein zusammengesetzter Poissonprozeß*

In diesem Fall ist $\int_t^{t+h}\nabla\Phi(y_{s-}^v)\sigma(y_{s-}^v,v)dN_s=\sum_{t<s\leq t+h}\Phi'(y_{s-})\Delta y_s$, und man erhält

$$\begin{aligned}
&\lim_{h\to 0}h^{-1}E[\Phi(t+h,x_{t+h}^u)-\Phi(t,x)] \\
&= \partial_t\Phi(t,x)+\nabla\Phi(t,x)\sigma(t,x,v)a \\
&\quad +\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v) \\
&\quad + \lim_{h\to 0}\frac{1}{h}E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^v,v)\partial_{ij}\Phi(y_{s-}^v,v)\sigma^{jk}(y_{s-}^v,v)\Delta N_s^l\Delta N_s^k] \\
&\quad + \lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\leq t+h}(\Phi(y_{s-}+\Delta y_s)-\Phi(y_{s-}))]
\end{aligned}$$

Die Terme der ersten beiden Zeilen entsprechen im Fall $\beta = id$ genau den Termen der HJB-Gleichung im Fall der Brownschen Bewegung (hier konnte ohne Beschränkung der Allgemeinheit $\beta = id$ gewählt werden, indem man ggf. σ geeignet modifizierte). Die beiden letzten Terme bedürfen der weiteren Untersuchung.

4.2.5 Berechnung der verbleibenden Sprungterme

Sei M der zu N gehörende Zählprozeß, d.h. M_t gibt die Anzahl der Sprünge bis zum Zeitpunkt t an. Weiter seien μ das Maß der Sprunghöhen und P eine Zufallsvariable (unabhängig von M) mit dieser Verteilung. Dann gilt $P(M_t = k) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$. Somit ist

$$
\begin{aligned}
&E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\Delta N^l_s\Delta N^k_s]\\
&= P(M_{t+h}-M_t=0)E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\\
&\quad \Delta N^l_s\Delta N^k_s|M_{t+h}-M_t=0]\\
&\quad +P(M_{t+h}-M_t=1)E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\\
&\quad \Delta N^l_s\Delta N^k_s|M_{t+h}-M_t=1]\\
&\quad +P(M_{t+h}-M_t>1)E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\\
&\quad \Delta N^l_s\Delta N^k_s|M_{t+h}-M_t>1]\\
&= P(M_{t+h}-M_t=1)E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\\
&\quad \Delta N^l_s\Delta N^k_s|M_{t+h}-M_t=1]\\
&\quad +P(M_{t+h}-M_t>1)E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\\
&\quad \Delta N^l_s\Delta N^k_s|M_{t+h}-M_t>1]
\end{aligned}
$$

Wegen $P(M_{t+h}-M_t>1) = \sum_{i=2}^{\infty} e^{-\lambda h}\frac{(\lambda h)^k}{k!} \leq Ch^2$ und

$E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^{v},v)\partial_{ij}\Phi(y_{s-}^{v},v)\sigma^{jk}(y_{s-}^{v},v)\Delta N_s^l\Delta N_s^k|M_{t+h}-M_t>1]=\mathcal{O}(1)$ muß nur der Term in der vorletzten Zeile weiter berechnet werden. Für y_{s-}^{v} kann dabei (t,x) gesetzt werden, da zwischen (t,x) und y_{s-}^{v} ja gerade kein Sprung stattfindet. Es ergibt sich also weiter

$$\lim_{h\to 0}\frac{1}{h}E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y_{s-}^{v},v)\partial_{ij}\Phi(y_{s-}^{v},v)\sigma^{jk}(y_{s-}^{v},v)\Delta N_s^l\Delta N_s^k]$$

$$=\lim_{h\to 0}\frac{1}{h}P(M_{t+h}-M_t=1)E[\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{t<s\leq t+h}\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v)$$

$$\Delta N_s^l\Delta N_s^k|M_{t+h}-M_t=1]$$

$$=\lim_{h\to 0}\frac{1}{h}\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}e^{-\lambda h}\lambda hE[\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v)P^lP^k]$$

$$=\lim_{h\to 0}\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}e^{-\lambda h}\lambda\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v)E[P^lP^k]$$

$$=\frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\lambda E[P^lP^k]\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v)$$

Die Größe $\lambda E[P^lP^k]$ ergibt weiter

$$\begin{aligned}&\lambda E[P^lP^k]\\ &=\int\lambda x^lx^k\mu(dx)\\ &=\int x^lx^k\nu(dx)\end{aligned}$$

mit dem Lévy-Maß ν von w.

Die eben gemachtenÜberlegungen sind auch beim anderen Sprungterm anwendbar, d.h. man erhält

$$\lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\leq t+h}(\Phi(y_{s-}+\Delta y_s)-\Phi(y_{s-}))]$$

$$\begin{aligned}&=\lim_{h\to 0}\frac{1}{h}P(M_{t+h}-M_t=1)E[\sum_{t<s\leq t+h}(\Phi(t,x+\sigma(t,x,v)\Delta N_s)-\Phi(t,x))|M_{t+h}-M_t=1]\\ &=\lambda E[(\Phi(t,x+\sigma(t,x,v)P)-\Phi(t,x))]\\ &=\int(\Phi(t,x+\sigma(t,x,v)y)-\Phi(t,x))\nu(dy)\end{aligned}$$

4.2.6 Zusammenfassung

Der HJB-Operator A im Fall eines Lévy-Prozesses w, der aus einer deterministischen Komponente, einer Brownschen Bewegung und einem zusammengesetztem Poisson-Prozeß besteht, ist gegeben durch

$$\begin{aligned}
A^v\Phi &= \lim_{h\to 0} h^{-1}E[\Phi(t+h, x^u_{t+h}) - \Phi(t,x)] \\
&= \partial_t\Phi(t,x) + \nabla\Phi(t,x)\sigma(t,x,v)a \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n}\sum_{k,l=1}^{m}\sum_{r=1}^{d}(\int x^l x^k \nu(dx))\beta^{lr}\beta^{kr}\sigma^{il}(t,x,v)\partial_{ij}\Phi(t,x,v)\sigma^{jk}(t,x,v) \\
&\quad + \int(\Phi(t, x+\sigma(t,x,v)y) - \Phi(t,x))\nu(dx)
\end{aligned}$$

sofern für Φ die Gleichungen

$$\lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\le t+h}(\Phi(y_{s-}+\Delta y_s) - \Phi(y_{s-}))] = \lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\le t+h}(\Phi(y_t+\Delta y_s) - \Phi(y_t))]$$

und

$$\begin{aligned}
&\lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\le t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_{s-},v)\partial_{ij}\Phi(y^v_{s-},v)\sigma^{jk}(y^v_{s-},v)\Delta N^l_s\Delta N^k_s] \\
&= \lim_{h\to 0}\frac{1}{h}E[\sum_{t<s\le t+h}\beta^{lr}\beta^{kr}\sigma^{il}(y^v_t,v)\partial_{ij}\Phi(y^v_t,v)\sigma^{jk}(y^v_t,v)\Delta N^l_s\Delta N^k_s]
\end{aligned}$$

gelten. Sind Φ, Φ_{ij}, σ Lipschitz-stetig, so ist hierfür z.B.

$$\lim_{h\to 0}\frac{1}{h}E[\sum_k |y^v_{\tau_k-} - y^v_t|] = 0$$

und

$$\lim_{h\to 0}\frac{1}{h}E[\sum_k |y^v_{\tau_k-} - y^v_t||\Delta N|^2] = 0$$

mit den Sprungzeiten τ_k hinreichend.

Die HJB-Gleichung ist dann

$$\sup_{v\in U}(A^v\Phi(t,x) + L(t,x,v)) = 0 \tag{4.9}$$

4.3 Anwendung auf die Portfoliooptimierung

Es soll nun versucht werden, die stochastische Kontrolltheorie auf die drei Beispiele des Modellierungskapitels anzuwenden. Dabei wird jeweils versucht, das Modell solange wie möglich allgemein zu halten, und ggf. einzuschränken. Die Notation gleicht der des Modellierungskapitels.

4.3.1 Beispiel 1: Black-Scholes-Merton ohne Transaktionskosten

Zur Erinnerung: Gegeben sind Preisprozesse der Form

$$dS^i = S^i(\alpha^i dt + \sum_{l=1}^{m} \sigma^{il} dw^l), \quad i = 0, 1, 2, \ldots, d$$

wobei die erste Spalte der Matrix σ 0 sei, es sind beliebige selbstfinanzierende Handelsstrategien mit nichtnegativem Vermögensverlauf zulässig, das Zielfunktional ist der erwartete Endnutzen $E[U(V_T)]$. Jede dieser Handelsstrategien korrespondiert offenbar zu einem Portfolioprozeß π. Die Differentialgleichung des Vermögensprozesses lautet dann

$$\begin{aligned}
dV &= \sum_{i=0}^{d} \varphi^i dS^i = \sum_{i=0}^{d} \frac{V^i}{S^i} dS^i \\
&= \sum_{i=0}^{d} \frac{\pi^i V}{S^i} dS^i = V \sum_{i=0}^{d} \pi^i \frac{dS^i}{S^i} \\
&= V \sum_{i=0}^{d} \pi^i (\alpha^i dt + \sum_{l=1}^{m} \sigma^{il} dw^l) \\
&= V(\pi^\top \alpha) dt + V(\pi^\top \sigma) dw
\end{aligned}$$

Man wählt also als kontrollierten Prozeß den Vermögensprozeß und als Kontrollprozeß den Portfolioprozeß π. Die Koeffizientenfunktionen f und σ der Kontrolltheorie lauten somit

$$\begin{aligned}
f(x, v) &= x(v^\top \alpha) \\
\sigma(x, v) &= x(v^\top \sigma)
\end{aligned}$$

Wählt man als Zielfunktional wiederum $E[U(V_T)]$ und eine Menge zulässiger Kontrollen U, so ergibt sich als HJB-Operator

$$\begin{aligned}\mathcal{H}(x,p,M) &= \sup_{\pi\in U}[f(x,\pi)p + \frac{1}{2}\operatorname{tr} a(x,\pi)M + L(x,\pi)] \\ &= \sup_{\pi\in U}[(\pi^\top\alpha)xp + \frac{1}{2}x^2(\pi^\top\sigma\sigma^\top\pi M)]\end{aligned}$$

Um weiter rechnen zu können, muß nun U konkret gewählt werden; wir wählen als Beispiele $U = \{\pi|\sum_{i=0}^d \pi^i = 1\}$ und $U = \{\pi|\pi^0 + \pi^1 = 1, 0 \leq \pi^1 \leq 1\}$.

Beispiel 1.1: $U = \{\pi|\sum_{i=0}^d \pi^i = 1\}$

Um die Optimierung im Fall $U = \{\pi|\sum_{i=0}^d \pi^i = 1\}$ durchführen zu können, schreiben wir das Optimierungsproblem in anderer Form. Dazu definieren wir zunächst $\widetilde{\sigma}$ als die Matrix, die aus σ durch Weglassen der Nullzeile entsteht, $\widetilde{\alpha}$ der Vektor α ohne erste Komponente und $\widetilde{\pi}$ der Vektor π ohne erste Komponente. Weiter sei abkürzend $C := x^2M\widetilde{\sigma}\widetilde{\sigma}^\top$ und $\mathbf{1}$ der $d+1$- oder d-dimensionale Vektor (je nach Zusammenhang) mit Einsen als Einträgen.Offensichtlich ist dann

$$\begin{aligned}\mathcal{H}(x,p,M) &= \sup_{\widetilde{\pi}\in\mathbb{R}^d}((\widetilde{\pi}^\top\widetilde{\alpha} + (1-\mathbf{1}^\top\widetilde{\pi})\alpha^0)xp + \frac{1}{2}\widetilde{\pi}^\top C\widetilde{\pi}) \\ &= \sup_{\widetilde{\pi}\in\mathbb{R}^d}(\alpha^0xp + \widetilde{\pi}^\top(\widetilde{\alpha} - \mathbf{1}\alpha^0)xp + \frac{1}{2}\widetilde{\pi}^\top C\widetilde{\pi})\end{aligned}$$

Dann muß der Punkt $\widetilde{\pi}$, an dem die Funktion ihr Supremum annimmt, die Gleichung $(\widetilde{\alpha} - \mathbf{1}\alpha^0)xp + \widetilde{\pi}^\top C = 0$ erfüllen. Umformen ergibt $\widetilde{\pi}^\top = xp(\alpha^0\mathbf{1}-\widetilde{\alpha})C^{-1} = \frac{p}{xM}(\alpha^0\mathbf{1}-\widetilde{\alpha})(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}$. Damit sind in der optimalen Strategie die Verhältnisse der riskanten Anlagen untereinander zeitlich konstant (da der Vorfaktor $\frac{p}{xM}$ für alle entsprechenden Komponenten gleich ist).

Der HJB-Operator für negatives M ist dann gegeben durch

$$\begin{aligned}&\mathcal{H}(x,p,M) \\ &= \alpha^0xp + xp(\alpha^0\mathbf{1}-\widetilde{\alpha})C^{-1}(\widetilde{\alpha} - \mathbf{1}\alpha^0)xp + \frac{1}{2}xp(\alpha^0\mathbf{1}-\widetilde{\alpha})C^{-1}CC^{-1}(\alpha^0\mathbf{1}-\widetilde{\alpha})^\top xp \\ &= \alpha^0xp - \frac{1}{2}x^2p^2(\alpha^0\mathbf{1}-\widetilde{\alpha})C^{-1}(\alpha^0\mathbf{1}-\widetilde{\alpha})^\top \\ &= \alpha^0xp - \frac{1}{2M}p^2(\alpha^0\mathbf{1}-\widetilde{\alpha})(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\alpha^0\mathbf{1}-\widetilde{\alpha})^\top\end{aligned}$$

Die HJB-Gleichung lautet somit

$$-W_t W_{xx} = \alpha^0 x W_x W_{xx} - \frac{1}{2} W_x^2 (\alpha^0 \mathbf{1} - \widetilde{\alpha})(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\alpha^0 \mathbf{1} - \widetilde{\alpha})^\top \tag{4.10}$$

zusammen mit den Randbedingungen $W(T,.) = U_1$. Gibt es daher eine Lösung W der Differentialgleichung (4.10) mit $W_{xx} < 0$ im gesamten Definitionsbereich, so ist diese Funktion auch eine Lösung der eigentlichen HJB-Gleichung. Die Vermutung $W_{xx} < 0$ ist aufgrund der Konkavität der Nutzenfunktion zumindest plausibel. Die allgemeine Lösung dieser HJB-Gleichung analytisch anzugeben, ist nicht möglich, da es sich um eine nichtlineare partielle Differentialgleichung zweiter Ordnung handelt. Kann man sie numerisch lösen, so ergibt sich mit der Lösung der Differentialgleichung unmittelbar auch die optimale Handelsstrategie als

$$\pi_t = (1 - \frac{W_x(V_{t-})}{V_{t-} W_{xx}(V_{t-})} \mathbf{1}^\top (\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\alpha^0 \mathbf{1} - \widetilde{\alpha}), \frac{W_x(V_{t-})}{V_{t-} W_{xx}(V_{t-})} (\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\alpha^0 \mathbf{1} - \widetilde{\alpha}))$$

Bemerkung 4.3.1 *Die optimale Strategie besteht aus einer zeitlich konstanten Aufteilung des Vermögens innerhalb der riskanten Anlagen. Allerdings variiert der Gesamtanteil im riskanten Bereich aufgrund des Vorfaktors* $\frac{W_x(V_{t-})}{V_{t-} W_{xx}(V_{t-})}$, *den man ökonomisch auch als das Verhältnis von globalem* ($\frac{W_x(V_{t-})}{V_{t-}}$) *zu lokalem* ($W_{xx}(V_{t-})$) *Wertzuwachs auffasssen kann.*

Beispiel 1.2: $U = \{\pi | \pi^0 + \pi^1 = 1, 0 \leq \pi^1 \leq 1\}$

Wählen wir $d = 1$ und $U = \{\pi | \pi^0 + \pi^1 = 1, 0 \leq \pi^1 \leq 1\}$, so ergibt sich

$$\sup_{\pi^1 \in [0,1]} [(\pi^0 \alpha^0 + (1 - \pi^0)\alpha^1) xp + \frac{1}{2} x^2 \pi^\top \begin{pmatrix} 0 \\ \sigma^{11} \end{pmatrix} \begin{pmatrix} 0 & \sigma^{11} \end{pmatrix} \pi M]$$

$$= \sup_{\pi^1 \in [0,1]} [(1 - \pi^1)\alpha^0 xp + \pi^1 \alpha^1 xp + \frac{1}{2} x^2 (\pi^1)^2 (\sigma^{11})^2 M]$$

$$= \begin{cases} \max[\alpha^0 xp, \alpha^1 xp + \frac{1}{2} x^2 (\sigma^{11})^2 M] \\ [(1 - \frac{(\alpha^0 - \alpha^1)xp}{x^2(\sigma^{11})^2 M})\alpha^0 xp + \frac{(\alpha^0 - \alpha^1)xp}{x^2(\sigma^{11})^2 M} \alpha^1 xp + \frac{1}{2} x^2 (\frac{(\alpha^0 - \alpha^1)xp}{x^2(\sigma^{11})^2 M})^2 (\sigma^{11})^2 M] \end{cases}$$

falls $\frac{(\alpha^0 - \alpha^1)p}{x(\sigma^{11})^2 M} \notin [0,1]$ oder $M \geq 0$

sonst

$$= \begin{cases} \max[\alpha^0 xp, \alpha^1 xp + \frac{1}{2} x^2 (\sigma^{11})^2 M] & \text{falls } \frac{(\alpha^0 - \alpha^1)p}{x(\sigma^{11})^2 M} \notin [0,1] \text{ oder } M \geq 0 \\ [\alpha^0 xp - \frac{(\alpha^0 - \alpha^1)^2 p^2}{2(\sigma^{11})^2 M}] & \text{sonst} \end{cases}$$

In einem (allerdings eher trivialen) Fall läßt sich diese HJB-Gleichung explizit lösen: Ist $\sigma = 0$, d.h. der Kursverlauf der Anlage 1 ebenfalls deterministisch, so ergibt sich als HJB-Operator $\mathcal{H}(x,p,M) = \sup_{\pi^1 \in [0,1]}[(1-\pi^1)\alpha^0 xp + \pi^1\alpha^1 xp + \frac{1}{2}x^2(\pi^1)^2(\sigma^{11})^2 M] = \max(\alpha^0 xp, \alpha^1 xp) = \max(\alpha^0,\alpha^1)xp$. Der letzte Schritt benutzt dabei, daß nur positives Vermögen möglich ($x \geq 0$) ist und die Wertfunktion monoton steigend im Anfangsvermögen ($p \geq 0$) ist. Die Gleichung lautet dann

$$-W_t = \max(\alpha^0,\alpha^1)xW_x$$

Lösungen sind die Funktionen $W(t,x) = Cx\exp(-\max(\alpha^0,\alpha^1)t)$, $C \geq 0$ (wie auch anschaulich zu erwarten war) und als Strategie das ausschließliche Investment in die Anlage mit höherer Rendite.

4.3.2 Beispiel 2: Black-Scholes-Merton mit Transaktionskosten

Da φ in diesem Modell differenzierbar ist, ist auch I^i differenzierbar. Wir fordern $\frac{dI_s^i}{dt} \in U$ mit einer geeigneten Menge U und daß $\frac{dI_s^i}{dt}$ der Integrierbarkeitsforderung $E[\int_0^t |\frac{dI_s^i}{ds}|^m ds] < \infty$ für $m \in \mathbb{N}, t \in [0,\infty)$ und $E_x[\int_0^\tau e^{-\beta_1 s}|U_1(-\frac{dI_s}{ds})|ds] < \infty$ der Kontrolltheorie genügt, wobei U_1 und β_1 die Grössen des Bewertungsfunktionals und τ die Austrittszeit des assoziierten Vermögensprozesses aus G ist.

Die Differentialgleichung für die i-te Komponente des Vermögensprozesses V^φ lautet (mit den Bezeichnungen aus Abschnitt 3.1.2):

$$\begin{aligned}
dV_s^{\varphi,i} &= d(\varphi_s^i S_s^i) \\
&= \varphi_s^i dS_s^i + S_s^i d\varphi_s^i \\
&= \varphi_s^i dS_s^i + g_i(\frac{dI_s^i}{ds})ds \\
&= V_s^{\varphi,i}\frac{dS_s^i}{S_s^i} + g_i(\frac{dI_s^i}{ds})ds \\
&= V_s^{\varphi,i}(\alpha^i ds + \sum_{l=1}^m \sigma^{il}dw_s^l) + g_i(\frac{dI_s^i}{ds})ds \\
&= (V_s^{\varphi,i}\alpha^i + g_i(\frac{dI_s^i}{ds}))ds + \sum_{l=1}^m (V_s^{\varphi,i}\sigma^{il})dw_s^l
\end{aligned}$$

Es bietet sich daher an, statt des Handelsprozesses φ oder des Portfolioprozesses den Investitionsprozeß I als Kontrollprozeß zu nutzen, die zugehörige Handelsstrategie muß dann in einem zweiten Schritt ermittelt werden. In der Schreibweise der Kontrolltheorie erhält man also als Koeffizientenfunktionen

$$f_i((V_s^j)_{j=0,1,\ldots,d}, (\frac{dI_s^i}{ds})_{j=0,1,\ldots,d}) = V_s^{\varphi,i}\alpha^i + g_i(\frac{dI_s^i}{ds})$$
$$\sigma^{il}((V_s^j)_{j=0,1,\ldots,d}, (\frac{dI_s^i}{ds})_{j=0,1,\ldots,d}) = V_s^{\varphi,i}\sigma^{il}$$

mit den durch den Preisprozeß S vorgegebenen Konstanten. Offenbar erfüllen diese Koeffizientenfunktionen sämtliche in der Kontrolltheorie gestellten Forderungen (Differenzierbarkeit, Lipschitzstetigkeit usw.). Die Bewertungsfunktion L ist damit $L((V_s^j)_{j=0,1,\ldots,d}, (\frac{dI_s^i}{ds})_{j=0,1,\ldots,d}) = U_1(-\sum_{i=0}^d \frac{dI_s^i}{ds})$, ψ ist 0.

Der HJB-Operator ist somit

$$\begin{aligned}
&\mathcal{H}(x,p,M)\\
&= \sup_{u:=(u^0,u^1,\ldots,u^d)\in U}[f(x,u)p + \frac{1}{2}\operatorname{tr} a(x,u)M + L(x,u)]\\
&= \sup_{u\in U}[\sum_{i=0}^d (x^i\alpha^i + g_i(u^i))p^i + \frac{1}{2}\operatorname{tr}(x^i(\sigma\sigma^\top)^{ij}x^j)_{i,j=0,1,2,\ldots,d}M + U_1(-\sum_{i=0}^d u^i)]\\
&= \sum_{i=0}^d x^i\alpha^i p^i + \frac{1}{2}\operatorname{tr}(x^i(\sigma\sigma^\top)^{ij}x^j)_{i,j=0,1,2,\ldots,d}M + \sup_{u\in U}[(g_i(u_i))^\top_{i=0,1,\ldots,d}p + U_1(-\sum_{i=0}^d u^i)]
\end{aligned}$$

woraus die HJB-Gleichung folgt, wenn man U kennt. Die Randbedingungen sind offenbar $W|_{\partial Q} = 0$. Hat man die HJB-Gleichung gelöst (ggf. numerisch), so erhält man die optimale Strategie durch Einsetzen von $V_{s-}, W_x(V_{s-})$ und $W_{xx}(V_{s-})$ für x, p und M.

Bemerkung 4.3.2 *Aus der Form des HJB-Operators ist bereits zu sehen, daß die optimale Strategie nicht von M, also nicht von $W_{xx}(V_{s-})$ abhängt.*

Beispiel 2.1 $d = 1, U = \{x | x^0 + x^1 \leq 0, x^0 \geq -1, x^1 \geq -1\}$

In diesem Fall existiert nur eine riskante Anlage,man darf nicht nachinvestieren und die Konsumrate (Konsum pro Zeiteinheit) ist beschränkt. Diese Bedingungen sichern die Kompaktheit von U und sind auch vom ökonomischen

Standpunkt gesehen sinnvoll. Weiterhin wählen wir $\lambda^{10} = 0$ und $\lambda^{01} > 0$ beliebig, d.h. Kosten fallen nur beim Kauf an. Damit ist $g_0 = id$ und $g_1 = ((1+\lambda^{01})^{-1}\mathbf{1}_{[0,\infty)} + \mathbf{1}_{(-\infty,\infty)})id$. Es gilt außerdem $Q := \{x | x^0 + x^1 \geq 0, x^0 + (1+\lambda^{01})x^1 \geq 0\}$. Es ergibt sich somit für den noch zu berechnenden Teil des HJB-Operators

$$\begin{aligned}
&\sup_{u\in U}[u^0p^0 + g_1(u^1)p^1 + U_1(-(u^0+u^1))] \\
&= \max(\sup_{u\in U, u^1\geq 0}(u^0p^0 + u^1(1+\lambda^{01})^{-1}p^1 + U_1(-(u^0+u^1))), \\
&\quad \sup_{u\in U, u^1<0}(u^0p^0 + u^1p^1 + U_1(-(u^0+u^1))))
\end{aligned}$$

Für das erste Supremum ergibt sich weiter

$$\begin{aligned}
&\sup_{u\in U, u^1\geq 0}(u^0p^0 + u^1(1+\lambda^{01})^{-1}p^1 + U_1(-(u^0+u^1))) \\
&= \sup_{s\in[-1,0]}\sup_{u^1\in[0,1-s]}((s-u^1)p^0 + u^1(1+\lambda^{01})^{-1}p^1 + U_1(-s)) \\
&= \sup_{s\in[-1,0]}(sp^0 + U_1(-s) + \sup_{u^1\in[0,1-s]}u^1((1+\lambda^{01})^{-1}p^1 - p^0)) \\
&= \sup_{s\in[-1,0]}(sp^0 + U_1(-s) + \max(0,(1-s)((1+\lambda^{01})^{-1}p^1 - p^0)))
\end{aligned}$$

und für das zweite Supremum

$$\begin{aligned}
&\sup_{u\in U, u^1<0}(u^0p^0 + u^1p^1 + U_1(-(u^0+u^1))) \\
&= \sup_{s\in[-2,0]}\sup_{u^1\in[-1,\min(0,1+s)]}((s-u^1)p^0 + u^1p^1 + U_1(-s)) \\
&= \sup_{s\in[-2,0]}(sp^0 +_1 (-s) + \sup_{u^1\in[-1,\min(0,1+s)]}u^1(p^1 - p^0)) \\
&= \sup_{s\in[-2,0]}(sp^0 + U_1(-s) + \max((p^0-p^1), \min(0,1+s)(p^1-p^0))
\end{aligned}$$

Diese beiden Ausdrücke können nun mit Fallunterscheidungen für p^0, p^1 berechnet werden. Da dies mathematisch problemlos, aber sehr platzaufwendig und wenig informativ ist, wird hier darauf verzichtet. Man sieht aber an dieser Stelle, daß der HJB-Operator vermutlich (wegen der Fallunterscheidungen) nicht stetig differenzierbar in p ist, der Begriff der schwachen Lösung dürfte daher vonnöten sein.

4.3.3 Beispiel 3: Das Modell von Kallsen

Zur Erinnerung: Das Modell von Kallsen unterscheidet sich vom klassischen Black-Scholes-Merton Modell nur durch die Wahl des Preisprozesses. Da die gewöhnlichen und die stochastischen Exponentiale von Lévy-Prozessen die gleiche Klasse darstellen, gehen wir von

$$dS_t^i = S_t^i dw_t^i, \quad i = 0, 1, 2, \ldots, d$$

aus, wobei w ein $(d+1)$-dimensionaler Lévy-Prozeß mit Tripel (α, β, ν) sei. Für den Vermögensprozeß gilt dann wieder

$$\begin{aligned} dV &= \sum_{i=0}^{d} \varphi^i dS^i = \sum_{i=0}^{d} \frac{V^i}{S^i} dS^i \\ &= \sum_{i=0}^{d} \frac{\pi^i V}{S^i} dS^i = V \sum_{i=0}^{d} \pi^i \frac{dS^i}{S^i} \\ &= V \sum_{i=0}^{d} \pi^i dw^i \\ &= V \pi^\top dw \end{aligned}$$

Die Koeffizientenfunktion σ in der Notation von Abschnitt 4.2 ist somit $\sigma(t,x,v) = xv^\top$, der HJB-Operator ist dann

$$\begin{aligned}
A^v\Phi &= \sup_{v\in U}[\partial_t\Phi(t,x) + \nabla\Phi(t,x)xv^\top a \\
&\quad + \frac{1}{2}\partial_{xx}\Phi(t,x,v)x^2\sum_{k,l=1}^{m}\sum_{r=1}^{d}\beta^{lr}\beta^{kr}v^l v^k \\
&\quad + \frac{1}{2}\partial_{xx}\Phi(t,x,v)x^2\sum_{k,l=1}^{m}\sum_{r=1}^{d}(\int y^l y^k \nu(dy))\beta^{lr}\beta^{kr}v^l v^k \\
&\quad + \int(\Phi(t,x+xv^\top y) - \Phi(t,x))\nu(dx)] \\
&= \sup_{v\in U}[\partial_t\Phi(t,x) + \Phi_x(t,x)xv^\top a \\
&\quad + \frac{1}{2}\Phi_{xx}(t,x,v)x^2 v^\top\beta\beta^\top v \\
&\quad + \frac{1}{2}\Phi_{xx}(t,x,v)x^2\sum_{k,l=1}^{m}\sum_{r=1}^{d}(\int y^l y^k \nu(dy))\beta^{lr}\beta^{kr}v^l v^k \\
&\quad + \int(\Phi(t,x(1+v^\top y)) - \Phi(t,x))\nu(dx)]
\end{aligned}$$

Die Lösung der zugehörigen Differentialgleichung ist nur numerisch möglich, da bereits die Supremumsbildung nicht analytisch durchgeführt werden kann.

4.4 Bewertung der Kontrolltheorie

Die Kontrolltheorie leistet in vielen Fällen gute Dienste, da in ihr die meisten Optimierungsprobleme dargestellt werden können, insbesondere auch solche mit Transaktionskosten.

Ein Nachteil besteht darin, daß man Markov-Prozesse als Preisprozesse verwenden muß. Ein weiterer, deutlich gravierenderer Nachteil besteht darin, daß bei Anwendung der Kontrolltheorie zuerst ein Optimierungsproblem gelöst werden muß, was unter Umständen analytisch gar nicht möglich ist (z.B. für allgemeine Lévy-Prozesse) und danach eine unter Umständen ebenfalls recht komplizierte partielle Differentialgleichung gelöst werden muß.

Kapitel 5

Die Martingalmethode

Der im letzten Kapitel vorgestellte Lösungsansatz aus der Kontrolltheorie leistet in vielen Fällen gute Dienste, er hat jedoch auch einige Nachteile: Zum ersten muß der zugrundeliegende Preisprozeß auf jeden Fall ein Markov-Prozeß sein, um die Kontrolltheorie anwenden zu können (wir haben uns sogar auf die Brownsche Bewegung beschränkt). Zum zweiten muß zur Lösung des Problems eine partielle Differentialgleichung zuerst aufgestellt (mit einem vorher zu lösenden Optimierungsproblem beim Bestimmen des Operators $\mathcal{H}$), dann gelöst werden, was unter Umständen recht schwierig sein kann. Ende der 70er/Anfang der 80er Jahre des letzten Jahrhunderts wurde daher von Harrison/Kreps [9] und Harrison/Pliska [10] ein anderer Ansatz vorgestellt, die sogenannte Martingalmethode. Es wird das in der Optimierung in verschiedenen Zusammenhängen verwendete Prinzip der Dualität aufgegriffen: Die Lösung des dualen Problems liefert relativ unmittelbar auch die Lösung des ursprünglichen Problems. In bestimmten einfachen Fällen (in sogenannten vollständigen und arbitragefreien Märkten) ist (vgl. Pliska [16]) dann das duale Problem lediglich eine eindimensionale Optimierungsaufgabe, also sehr einfach zu lösen. Auch in anderen Fällen kann es einfacher zu lösen sein als das ursprüngliche Problem oder auch Beiträge zur genaueren Untersuchung des Problems bieten (vgl. [14]).

Der Begriff Martingalmethode gründet sich auf die Erkenntnis, daß die Kontrollprozesse des dualen Problems durch Dichten bezüglich des Wahrscheinlichkeitsmaßes P gegeben sind, die den Preisprozeß zu einem Martingal oder Supermartingal machen, vgl. die folgenden Ausführungen.

5.1 Spezifizierung des allgemeinen Modells aus 3.1.4

Wie im letzten Kapitel ist auch hier eine Behandlung des Modells aus 3.1.4 nicht in voller Allgemeinheit möglich. In diesem Abschnitt wird daher das allgemeine Modell soweit spezifiziert, wie es für den Rahmen der Martingalmethode nötig erscheint, in einzelnen Zusammenhängen wird eine weitere Spezifizierung notwendig sein.

5.1.1 Der Preisprozeß S

Der Preisprozeß S unterliegt gegenüber dem allgemeinen Modell vorerst keinen weiteren Einschränkungen, ist also ein beliebiges positives Semimartingal. Den Fall des endlichen Zeithorizonts kann man „simulieren", indem man Y als konstant nach dem Endzeitpunkt vorausgesetzt. Oft betrachten wir den diskontierten Preisprozeß $\widehat{S} := \frac{1}{S^0} S$ mit $\widehat{S}^0 \equiv 1$.

5.1.2 Die zulässigen Handelsprozesse φ

In der Regel wird verlangt, daß φ selbstfnanzierend und der assoziierte Vermögensprozeß entweder nichtnegativ oder zumindest von unten beschränkt ist. Gemäß der Notation in [20] wird der Vermögensprozeß auch mit dem Buchstaben X bezeichnet. Es gilt folgende

Definition 5.1.1 *Ein vorhersehbarer Prozeß φ heißt a-zulässig für $a \geq 0$, wenn $\varphi \cdot S \geq -a$. Er heißt zulässig, falls es ein $a \geq 0$ gibt, für das er zulässig ist. Für $x \in \mathbb{R}$ ist $\mathfrak{X}^b(x) := \{x + \varphi \cdot Y | \varphi$ ist zulässig und selbstfinanzierend$\}$ und $\mathfrak{X}(x) := \{x + \varphi \cdot Y | \varphi$ ist 0-zulässig und selbstfinanzierend$\}$. Statt $\mathfrak{X}(1)$ schreibt man auch kurz $\mathfrak{X}$, ebenso bei $\mathfrak{X}^b$*

5.1.3 Das Bewertungsfunktional

Als Bewertungsfunktional verwendet man $E[U(X_T)]$. Die Funktion U wird als stetig differenzierbar, strikt konkav und streng monoton wachsend vorausgesetzt. Sie ist entweder auf $\mathbb{R}$ oder auf $[0, \infty)$ definiert, je nachdem ob negatives Vermögen (bzw. negative Konsumrate) zugelassen werden.

5.2 Dualität, Vollständigkeit und Arbitragefreiheit

Dieser Abschnitt orientiert sich an den Arbeiten von Kramkov/Schachermayer [14] und Schachermayer [20]. Wir führen zunächst einige Begriffe ein, die im folgenden benötigt werden.

Definition 5.2.1 *Die Größe*

$$AE(U) := \limsup_{x\to\infty} \frac{xU'(x)}{U(x)}$$

heißt die asymptotische Elastizität der Funktion U.

Definition 5.2.2 *Ein Wahrscheinlichkeitsmaß Q, das äquivalent ist zum (bzw. absolutstetig bezüglich des) ursprünglichen Wahrscheinlichkeitsmaß P, heißt äquivalentes (absolutstetiges) lokales Martingalmaß, falls S unter Q ein lokales Martingal ist. Ist S unter Q ein Martingal, so heißt Q äquivalentes (absolutstetiges) Martingalmaß.*

Für den Fall eines lokal beschränkten Preisprozesses S ist diese Definition nach Schachermayer [20] äquivalent dazu, daß alle $X \in \mathfrak{X}(1)$ lokale Q-Martingale sind. Die Menge der äquivalenten bzw. absolutstetigen lokalen Martingalmaße zu einem gegebenen Preisprozeß Y wird $\mathcal{M}^e$ bzw. $\mathcal{M}^a$ genannt.

In den folgenden Überlegungen wird öfters $\mathcal{M}^e \neq \emptyset$ oder $\mathcal{M}^e = \{\mathcal{Q}\}$ als Voraussetzung verwendet. Am Ende des Abschnitts werden diese Bedingungen untersucht, und es wird dargestellt, inwiefern sie mit der Existenz einer sogenannten Arbitrage (risikoloser Gewinn) bzw. der Vollständigkeit des Marktes (jedes Endvermögen kann erreicht werden) in Verbindung stehen und somit ökonomisch interpretierbar sind.

5.2.1 Das Dualitätsprinzip

In diesem Unterabschnitt wird das Dualitätsprinzip, das Kernstück der Martingalmethode, vorgestellt. Die Darstellung folgt den Arbeiten von Schachermayer [20] und Kramkov/Schachermayer [14]; zunächst betrachten wir den Fall, daß der Vermögensprozeß nicht negativ werden darf.

Wir gehen davon aus, daß der deterministische Preisprozeß S^0 eine Konstante ist (ggf. definieren wir zum gegebenen Preisprozeß S den diskontierten Preisprozeß $\widehat{S} := \frac{1}{S^0}S$ und verwenden ihn im weiteren), und daß es ein äquivalentes Maß Q gibt, unter dem S zum lokalen Martingal wird, also $M^e \neq \emptyset$. Weiterhin verwenden wir die Bezeichnung $u(x) := \sup_{X \in \mathfrak{X}(x)} E[U(X_T)]$.

Definition 5.2.3 *Sei $V(y) := \sup_{x>0}[U(x) - xy]$ für $y > 0$ und $\mathfrak{Y}(y) := \{Y | Y \geq 0, Y_0 = 0, XY$ ist Supermartingal für alle $X \in \mathfrak{X}(1)\}$. Dann heißt die Optimierungsaufgabe:*

$$\textit{Finde } \widehat{Y} \in \mathfrak{Y}, \textit{ so daß } E[V(\widehat{Y}_T)] = \inf_{Y \in \mathfrak{Y}(y)} E[V(Y_T)]$$

das zum eben vorgestellten Problem duale Optimierungsproblem und man definiert $v(y) := \inf_{Y \in \mathfrak{Y}(y)} E[V(Y_T)]$ (unabhängig davon, ob das Infimum tatsächlich angenommen wird)

Bemerkung 5.2.4 *Die Funktion V ist monoton fallend (wegen $U(x) - xy \leq U(x) - xy'$ für $y > y'$, $x > 0$) und konvex, denn*

$$\begin{aligned}
V(\lambda y_1 + (1-\lambda)y_2) &= \sup_{x>0}[U(x) - x(\lambda y_1 + (1-\lambda)y_2)] \\
&= \sup_{x>0}[\lambda(U(x) - xy_1) + (1-\lambda)(U(x) - xy_2)] \\
&\leq \sup_{x>0}[\lambda(U(x) - xy_1)] + \sup_{x>0}[(1-\lambda)(U(x) - xy_2)] \\
&= \lambda \sup_{x>0}[U(x) - xy_1] + (1-\lambda) \sup_{x>0}[U(x) - xy_2] \\
&= \lambda V(y_1) + (1-\lambda)V(y_2)
\end{aligned}$$

Ist U stetig differenzierbar, strikt konkav und erfüllt $U'(0) = \infty, U'(\infty) = 0$, so gilt offenbar $V(y) = U(U'^{-1}(y)) - U'^{-1}(y)y$.

Im weiteren werden wir folgenden Fragen nachgehen:

1. Für welche x und y sind $u(x)$ bzw. $v(y)$ endlich, bestehen zwischen ihnen Zusammenhänge und ggf. welche?

2. Wann wird das Supremum bzw. das Infimum in den Optimierungsproblemen angenommen, sind die optimalen $\widehat{X}$ bzw. $\widehat{Y}$ eindeutig und kann man die Kandidaten für optimale $\widehat{X}$ bzw. $\widehat{Y}$ eingrenzen?

3. Kann man das duale Problem einfacher lösen als das ursprüngliche? Wenn ja, wie genau?

Letztenendes ist für unsere Zwecke die dritte Frage von Bedeutung, die anderen beiden werden aber auf dem Weg dorthin zu stellen sein.

Eine Neuformulierung des Problems

Um die eben gestellten Fragen zu beantworten, stellen wir zunächst noch eine abstrakte Version des Problems dar. Sei dazu

$$\mathcal{C}(x) := \{g \in L^0_+(\Omega, \mathcal{F}, P) | 0 \leq g \leq X_T \text{ für ein } X \in \mathfrak{X}(x)\}$$
$$\mathcal{D}(y) := \{h \in L^0_+(\Omega, \mathcal{F}, P) | 0 \leq g \leq Y_T \text{ für ein } Y \in \mathfrak{Y}(y)\}$$

d.h. wir betrachten statt der Vermögensprozesse nur die (nichtnegativen) Endvermögen sowie nichtnegative Zufallsvariablen, die von diesen dominiert werden. Die Optimierungsaufgaben

$$\text{Finde } \widehat{g} \in \mathcal{C}(x), \text{ so daß } EU(\widehat{g}) = \sup_{g \in \mathcal{C}(x)} EU(g)$$
$$\text{Finde } \widehat{h} \in \mathcal{D}(y), \text{ so daß } EV(\widehat{h}) = \inf_{h \in \mathcal{D}(y)} EV(h)$$

sind dann äquivalent zum ersten Teil des eigentlichen Optimierungsproblems, da die Funktionen U und V monoton steigend bzw. fallend sind, zu jedem $g \in \mathcal{C}(x)$ ein Endvermögen $X_T \geq g$ eines Prozesses $X \in \mathfrak{X}(x)$ existiert, und mit diesem X_T $E[U(g)] \leq E[U(X_T)]$ gilt.

Bemerkung 5.2.5 *Es gilt $\mathcal{C}(x) = x\mathcal{C}(1)$, $\mathcal{D}(y) = y\mathcal{D}(1)$, $\mathfrak{X}(x) = x\mathfrak{X}(1)$ sowie $\mathfrak{Y}(y) = y\mathfrak{Y}(1)$, denn zu jedem $X \in \mathcal{C}(1)$ ist mit Anfangsvermögen x der Prozeß xX durch die Handelsstrategie $x\varphi$ zu erreichen (wobei φ die zu X gehörende Handelsstrategie ist). Für $\mathfrak{Y}$ gilt dieses Argument sinngemäß auch, denn die Supermartingaleigenschaft eines Prozesses bleibt unter der Multiplikation mit einer Konstanten natürlich erhalten. für die Mengen $\mathcal{C}(x)$ und $\mathcal{D}(y)$ ist der Zusammenhang dann eindeutig. Im weiteren werden daher oft Ergebnisse für $\mathfrak{X} = \mathfrak{X}(1)$ usw. angegeben, die dann leicht auf $\mathfrak{X}(x)$ usw. verallgemeinert werden können.*

Um diese Abstraktion nutzen zu können, muß zuerst eine duale Relation zwischen $\mathcal{C}$ und $\mathcal{D}$ hergestellt werden:

Proposition 5.1 (Schachermayer [14]) *Es existiere ein Wahrscheinlichkeitsmaß $Q \sim P$, so daß alle $X \in \mathfrak{X}$ lokale Q-Martingale sind. Dann haben die eben definierten Mengen $\mathcal{C}$ und $\mathcal{D}$ folgende Eigenschaften:*

1. *$\mathcal{C}$ und $\mathcal{D}$ sind konvex, abgeschlossen bezüglich der Topologie der Konvergenz in Wahrscheinlichkeit und haben die Eigenschaft $g \in \mathcal{C}$, $0 \leq g^* \leq g \implies g^* \in \mathcal{C}$ sowie $h \in \mathcal{D}$, $0 \leq h^* \leq h \implies h^* \in \mathcal{D}$*

2. *$g \in \mathcal{C}$ genau dann wenn $E[gh] \leq 1 \quad \forall h \in \mathcal{D}$ und $h \in \mathcal{D}$ genau dann wenn $E[gh] \leq 1 \quad \forall g \in \mathcal{C}$*

3. *Die konstante Zufallsvariable 1 ist in $\mathcal{C}$*

Die Konvexität folgt aus der Beobachtung, daß es zu $g_1, g_2 \in \mathcal{C}$ $X^1, X^2 \in \mathfrak{X}$ gibt, so daß $0 \leq g_i \leq X_T^i$, $i = 1, 2$, somit $0 \leq \lambda g_1 + (1-\lambda) g_2 \leq \lambda X_T^1 + (1-\lambda) X_T^2$ ist und der Prozeß $\lambda X^1 + (1-\lambda) X^2$ wieder in $\mathfrak{X}$ liegt (Erzeugung durch Linearkombination der entsprechenden Handelsstrategien). Auch die letzte Behauptung ist offensichtlich, da ein sicheres Endvermögen von 1 entsteht, wenn das gesamte Vermögen in der sicheren Anlage verbleibt. Der zweite Teil der ersten Aussage ist weniger offensichtlich, und die zweite Aussage stellt den Kern der Proposition dar. Sie wird in Kramkov/Schachermayer [14] bewiesen, wobei jeweils die eine Richtung der Äquivalenzen aufgrund der Definition von $\mathfrak{X}$ und $\mathfrak{Y}$ offensichtlich ist.

Das erste Dualitätstheorem für $\mathcal{C}$ und $\mathcal{D}$

Theorem 5.2 (Kramkov/Schachermayer [14]) *Gegeben seien Mengen $\mathcal{C}$ und $\mathcal{D}$ positiver Zufallsvariablen, die die Aussagen der Proposition 5.1 erfüllen sowie das dazugehörige Optimierungsproblem. Die Funktion U habe die Eigenschaften $\lim_{x \to 0} U'(x) = \infty$, $\lim_{x \to \infty} U'(x) = 0$. Weiterhin existiere ein $x > 0$ mit $u(x) < \infty$. Dann gilt:*

1. *$u(x) < \infty$ für alle $x > 0$ und es gibt ein $y_0 > 0$, so daß $v(y)$ endlich ist für $y > y_0$. Die Funktionen u und v sind konjugiert zueinander, d.h.*

$$v(y) = \sup_{x>0} [u(x) - xy], \quad y > 0$$
$$u(x) = \inf_{y>0} [v(y) + xy], \quad x > 0$$

Die Funktion u ist stetig differenzierbar auf $(0,\infty)$ und v ist strikt konvex auf (y_0,∞). Weiterhin gilt:

$$\lim_{x\to 0} u'(x) = \infty, \quad \lim_{y\to\infty} v'(y) = 0$$

2. *Ist $v(y) < \infty$, so gibt es ein eindeutiges $\widehat{h} \in \mathcal{D}(y)$, für das das Infimum angenommen wird, d.h. $E[V(\widehat{h})] = v(y)$*

Beweis: (skizziert, vollständige Version siehe Kramkov/Schachermayer [14])

1. Zunächst zur zweiten Aussage: Man kann zeigen, daß es zu jeder Folge $(g^n)_{n\in\mathbb{N}}$ nichtnegativer Zufallsvariablen eine Folge $h^n \in \operatorname{conv}(g^n, g^{n+1}, \dots)$ gibt, die fast sicher gegen eine nichtnegative numerische (d.h. evtl. mit Werten $= \infty$) Zufallsvariable h konvergiert. Sei nun g^n eine Minimalfolge in $\mathcal{D}(y)$, d.h. $\lim_{n\to\infty} E[V(g^n)] = v(y)$ und $h^n \in \operatorname{conv}(g^n, g^{n+1}, \dots)$ eine fast sicher gegen $\widehat{h}$ konvergente Folge. Dann ist wegen der Konvexität von V $E[V(h^n)] = E[V(\sum_{i\geq n} \lambda_i g^i)] \leq \sum_{i\geq n} \lambda_i E[V(g^i)] \leq \sum_{i\geq n} \lambda_i \sup_{m\geq n} E[V(g^m)] = \sup_{m\geq n} E[V(g^m)] \leq \sup_{m\geq (n-k)} E[V(g^m)]$ und damit natürlich auch $\sup_{m\geq n} E[V(h^m)] \leq \sup_{m\geq n} E[V(g^m)]$, woraus wiederum $v(y) \leq \liminf_{n\to\infty} E[V(h^n)] \leq \limsup_{n\to\infty} E[V(h^n)] \leq \limsup_{n\to\infty} E[V(g^n)] = \lim_{n\to\infty} E[V(g^n)] = v(y)$, also $\lim_{n\to\infty} E[V(h^n)] = v(y)$ folgt. Nun ist für die Existenz der Optimallösung noch $\widehat{h} \in \mathcal{D}(y)$ (mit Hilfe der stochastischen Abgeschlossenheit) zu zeigen sowie $E[V(\widehat{h})] = E[\lim_{n\to\infty} V(h^n)] = \lim_{n\to\infty} E[V(h^n)]$ (mit Hilfe des Fatou-Lemmas und ähnlicher Aussagen). Die Eindeutigkeit folgt aus folgender Überlegung: Seien $\widehat{h}_1$ und $\widehat{h}_2$ zwei Optimallösungen. Dann ist $\frac{1}{2}(\widehat{h}_1+\widehat{h}_2) \in \mathcal{D}(y)$ und wegen der strikten Konvexität von V $E[V(\frac{1}{2}(\widehat{h}_1+\widehat{h}_2))] < E[\frac{1}{2}V(\widehat{h}_1) + \frac{1}{2}V(\widehat{h}_2)] = v(y)$ im Widerspruch zur Optimalität von $\widehat{h}_1$ und $\widehat{h}_2$.

2. Nun zur ersten Aussage: Die zentrale Behauptung ist, daß u und v konjugiert zueinander sind. Dabei genügt es, eine der beiden Relationen zu zeigen, die andere läßt sich dann allein aufgrund der Dualität folgern. Ebenso sind die anderen Behauptungen Folgerungen der Dualitätsrelation von u und v. Um $v(y) = \sup_{x>0}[u(x) - xy], \quad y > 0$ zu zeigen, definiert man zunächst $\mathcal{B}_n := \{g \in L^0(\Omega, \mathcal{F}, P) | 0 \leq g \leq n\}$ und zeigt dann

$$\sup_{g\in\mathcal{B}_n} \inf_{h\in\mathcal{D}(y)} E[U(g) - gh] = \inf_{h\in\mathcal{D}(y)} \sup_{g\in\mathcal{B}_n} E[U(g) - gh]$$

Die Vertauschung von Supremums- und Infimumsbildung ist möglich nach dem sogenannten Minimax-Theorem. An dieser Stelle geht die Kompaktheit/Abgeschlossenheit der Mengen $\mathcal{B}_n$ und $\mathcal{D}(y)$ ein. Um diesen Zusammenhang zu nutzen, verwendet man unter anderem $\mathcal{B}_n \subseteq \mathcal{C}(n)$ (da die konstante Zufallsvariable n in $\mathcal{C}(n)$ ist). Für die linke Seite der Gleichung gilt:

$$\begin{aligned}
\lim_{n\to\infty} \sup_{g\in\mathcal{B}_n} \inf_{h\in\mathcal{D}(y)} E[U(g) - gh] &= \sup_{n\in\mathbb{N}} \sup_{g\in\mathcal{B}_n} (E[U(g)] - \sup_{h\in\mathcal{D}(y)} E[gh]) \\
&= \sup_{x>0} \sup_{g\in\mathcal{C}(x)} (E[U(g) - xy]) \\
&= \sup_{x>0} (u(x) - xy)
\end{aligned}$$

Andererseits gilt auch

$$\begin{aligned}
\sup_{g\in\mathcal{B}_n} E[U(g) - gh] &= \sup_{g\in\mathcal{B}_n} \int [U(g(\omega)) - g(\omega)h(gw)]dP(\omega) \\
&= \sup_{g, 0\le g(\omega)\le n} \int [U(g(\omega)) - g(\omega)h(w)]dP(\omega) \\
&= \int \sup_{0\le g(\omega)\le n} [U(g(\omega)) - g(\omega)h(\omega)]dP(\omega) \\
&= \int V^n(h)dP
\end{aligned}$$

mit $V^n(y) := \sup_{0\le x\le n}[U(x) - xy]$ und damit

$$\lim_{n\to\infty} \inf_{h\in\mathcal{D}(y)} \sup_{g\in\mathcal{B}_n} E[U(g) - gh] = \lim_{n\to\infty} \inf_{h\in\mathcal{D}(y)} E[V^n(h)]$$

Nun ist noch zu zeigen, daß dieser Term $\inf_{h\in\mathcal{D}(y)} E[V(h)] = v(y)$ ist. Wegen $V^n \le V$ gilt auf jeden Fall $\lim_{n\to\infty} \inf_{h\in\mathcal{D}(y)} E[V^n(h)] \le v(y)$. Um „$\ge$" zu zeigen, wählt man eine Folge h^n, so daß $\lim_{n\to\infty} \inf_{h\in\mathcal{D}(y)} E[V^n(h)] = \lim_{n\to\infty} E[V^n(h^n)]$ (z.B. durch Wahl von h^n mit $|\inf_{h\in\mathcal{D}(y)} E[V^n(h)] - E[V^n(h^n)]| \le \frac{1}{n}$). Hierauf wählt man eine fast sicher konvergente Folge von Zufallsvariablen $f^n \in \operatorname{conv}(h^n, h^{n+1}, \dots)$. Deren Grenzwert h liegt wegen der stochastischen Konvergenz wieder in $\mathcal{D}(y)$, und es gilt (da V konvex ist und $V^n(.)$ monoton steigend in n) $E[V^n(f^n)] = E[V^n(\sum_{i\ge n} \lambda_i h^i)] \le \sum_{i\ge n} \lambda_i E[V^n(h^i)] \le \sum_{i\ge n} \lambda_i \sup_{i\ge n} E[V^n(h_i)] =$

$\sup_{i\geq n} E[V^n(h_i)] \leq \sup_{i\geq n} E[V^i(h^i)]$, woraus wiederum (mit Hilfe des Fatou-Lemmas) $\lim_{n\to\infty} \inf_{h\in\mathcal{D}(y)} E[V^n(h)] = \lim_{n\to\infty} E[V^n(h^n)]$

$= \limsup_{n\to\infty} E[V^n(h^n)] \geq \liminf_{n\to\infty} E[V^n(f^n)] \geq E[\liminf_{n\to\infty} V^n(f^n)] = E[V(h)] \geq v(y)$ folgt.

■

Das zweite Dualitätstheorem für $\mathcal{C}$ und $\mathcal{D}$

Theorem 5.3 *Ist zusätzlich zu den Bedingungen des Theorems 5.2 $AE(U) < 1$, so gilt auch:*

1. *$v(y) < \infty$ für alle $y > 0$. u und v sind stetig differenzierbar auf $(0, \infty)$, u' und $-v'$ sind streng monoton fallend und es gilt:*

$$\lim_{x\to\infty} u'(x) = 0, \quad \lim_{y\to 0} v'(y) = -\infty$$

   sowie $AE(u)_+ \leq AE(U)_+ < 1$

2. *Es gibt ein eindeutiges $\widehat{g}(x) \in \mathcal{C}(x)$, für das das Supremum angenommen wird, d.h. $E(U(\widehat{g}(x))) = u(x)$. Ist $\widehat{h}(y)$ die (eindeutige) Lösung des dualen Problems für $y = u'(x)$, so gilt:*

$$\widehat{g}(x) = -V'(\widehat{h}(y)) = U'^{-1}(\widehat{h}(y))$$

   und $E(\widehat{g}(x)\widehat{h}(y)) = xy$.

3. *Es gilt:*

$$u'(x) = E[\frac{\widehat{g}(x)U'(\widehat{g}(x))}{x}], \quad v'(y) = E[\frac{\widehat{h}(y)V'(\widehat{h}(y))}{y}]$$

Beweis: (skizziert, vollständige Version siehe Kramkov/Schachermayer [14])

1. Man zeigt zunächst, daß für eine konvergente Folge y^n mit Grenwert y auch $\widehat{h}(y^n)$ in Wahrscheinlichkeit gegen $\widehat{h}(y)$ sowie $V(\widehat{h}(y^n))$ gegen $V(\widehat{h}(y))$ in $\mathcal{L}^1$ konvergiert. Dann benutzt man, daß die asymptotische Elastizität von U kleiner als 1 ist: Es folgt aus dieser Tatsache nämlich die Existenz von Konstanten C und y_0 mit $|V'(y)|y = -V'(y)y < CV(y)$ für $0 < y < y_0$. Hieraus wiederum folgert man die $\mathcal{L}^1$-Konvergenz von $V'(\widehat{h}(y^n))\widehat{h}(y^n)$ gegen $V'(\widehat{h}(y))\widehat{h}(y)$.

2. Dieser Zusammenhang wird dann dazu verwendet, um $yv'(y) = E[\widehat{h}(y)V'(\widehat{h}(y))]$ zu beweisen: Es gilt

$$\lim_{\lambda\to 1}\frac{v(y)-v(\lambda y)}{\lambda-1} = -\lim_{h\to 0}\frac{v(y+hy)-v(y)}{hy}y = -y\lim_{h\to 0}\frac{v(y+h)-v(y)}{h} = -yv'(y)$$

sofern der Grenzwert existiert. Daher will man (mit $I(y) := -V'(y) = U'^{-1}(y)$)

$$\begin{aligned}\limsup_{\lambda\to 1,\lambda>1}\frac{v(y)-v(\lambda y)}{\lambda-1} &\le E[\widehat{h}(y)I(\widehat{h}(y))]\\ \liminf_{\lambda\to 1,\lambda>1}\frac{v(y)-v(\lambda y)}{\lambda-1} &\ge E[\widehat{h}(y)I(\widehat{h}(y))]\end{aligned}$$

zeigen, was die Gültigkeit von $v_r'(y) = E[\frac{\widehat{h}(y)V'(\widehat{h}(y))}{y}]$ mit der rechtsseitigen Ableitung v_r' von v beweist. Die im ersten Schritt bewiesene Stetigkeit der rechten Seite und die Konvexität von v beweist dann die stetige Differenzierbarkeit von v und somit die Formel. Die erste der beiden Abschätzungen erhält man durch

$$\begin{aligned}\limsup_{\lambda\to 1,\lambda>1}\frac{v(y)-v(\lambda y)}{\lambda-1} &= \limsup_{\lambda\to 1,\lambda>1}\frac{E[V(\widehat{h}(y))]-E[V(\widehat{h}(\lambda y))]}{\lambda-1}\\ &\le \limsup_{\lambda\to 1,\lambda>1}\frac{1}{\lambda-1}E[V(\frac{1}{\lambda}\widehat{h}(\lambda y))-V(\widehat{h}(\lambda y))]\\ &\le \limsup_{\lambda\to 1,\lambda>1}\frac{1}{\lambda-1}E[(\frac{1}{\lambda}-1)\widehat{h}(\lambda y)V'(\frac{1}{\lambda}\widehat{h}(\lambda y))]\\ &= E[\widehat{h}(y)I(\widehat{h}(y))]\end{aligned}$$

Dabei folgt die erste Ungleichung aus der Beobachtung, daß $\frac{1}{\lambda}\widehat{h}(\lambda y)$ in $\mathcal{D}(y)$ liegt (wegen $\widehat{h}(\lambda y) \in \mathcal{D}(\lambda y) = \lambda\mathcal{D}(y)$), und die zweite aus der Ungleichung $V(x+h) - V(x) \ge hV'(x)\ \forall x \forall h$. Die abschließende Gleichung bedarf eines gesonderten Beweises, der ähnlich wie die in

Schritt 1 verläuft. Die zweite Abschätzung

$$\begin{aligned}
\liminf_{\lambda\to 1,\lambda>1} \frac{v(y)-v(\lambda y)}{\lambda-1} &= \liminf_{\lambda\to 1,\lambda>1} \frac{E[V(\widehat{h}(y))]-E[V(\widehat{h}(\lambda y))]}{\lambda-1} \\
&\geq \liminf_{\lambda\to 1,\lambda>1} \frac{E[V(\widehat{h}(y))]-E[V(\lambda\widehat{h}(y))]}{\lambda-1} \\
&\geq \liminf_{\lambda\to 1,\lambda>1} \frac{1}{\lambda-1} E[(1-\lambda)\widehat{h}(y)V'(\lambda\widehat{h}(y))] \\
&= E[\widehat{h}(y)I(\widehat{h}(y))]
\end{aligned}$$

basiert analog dazu auf der Beobachtung, daß $\lambda\widehat{h}(y) \in \lambda\mathcal{D}(y) = \mathcal{D}(\lambda y)$ und auf der eben erwähnten Ungleichung für konvexe Funktionen.

3. Als nächstes wird nun die zweite Aussage bewiesen: Man gibt sich daher Zahlen x und y mit $x = -v'(y)$ vor. Zu zeigen ist als erstes, daß $-V'(\widehat{h}(y))$ zu $\mathcal{C}(x)$ gehört: Es reicht aus, die Ungleichung $E[hI(\widehat{h}(y))] \leq xy = -yv'(y) = E[\widehat{h}(y)I(\widehat{h}(y))]$ für alle $h \in \mathcal{D}(y)$ zu zeigen (die beiden Gleichungen folgen aus dem eben bewiesenen). Sei dazu für beliebiges, aber festes $h \in \mathcal{D}(y)$ $h_\delta := (1-\delta)\widehat{h}(y) + \delta h, \quad \delta \in (0,1)$. Dann gilt wegen der Optimalität von $\widehat{h}(y)$ und der Konvexität von V (bzw. der daraus folgenden Monotonie von V'): $0 \leq E[V(h_\delta)] - E[V(\widehat{h}(y))] = E[\int_{h_\delta}^{\widehat{h}(y)} I(z)dz] \leq E[I(h_\delta)(\widehat{h}(y)-h_\delta)] = E[I(h_\delta)(\delta(\widehat{h}(y)-h))] = \delta E[I(h_\delta)(\widehat{h}(y)-h)]$ und somit $E[I((1-\delta)\widehat{h}(y))\widehat{h}(y)] \geq E[I((1-\delta)\widehat{h}(y)+\delta h)\widehat{h}(y)] \geq E[I(h_\delta)h]$. Bildet man nun den Grenzwert für $\delta \to 0$, so erhält man mit Hilfe des Fatou-Lemmas und dem Theorem von der monotonen Konvergenz die gewünschte Ungleichung. Nun zur Optimalität: Sei $g \in \mathcal{C}$ beliebig. Dann ist $E[g\widehat{h}(y)] \leq xy$ und $U(g) = (U(g(\omega)))_{\omega\in\Omega} = (\inf_{y\geq 0} V(y)+yg(\omega))_{\omega\in\Omega} \leq (V(\widehat{h}(\omega))+\widehat{h}(\omega)g(\omega))_{\omega\in\Omega} = V(\widehat{h}) + \widehat{h}g$ (Anmerkung: Diese beiden Ungleichungen gelten auch für jedes andere $h \in \mathcal{D}(y)$). Damit folgt dann $E[U(g)] \leq v(y) + xy \leq E[V(\widehat{h}(y))+\widehat{h}(y)I(\widehat{h}(y))] \leq E[U(I(\widehat{h}(y)))] = E[U(\widehat{g}(x))]$. Die erste Ungleichung ist unmittelbar einsichtig aus dem im vorigen Satz erwähnten Zusammenhang, die zweite folgt aus dem vorher Bewiesenen. Die darauffolgende Ungleichung folgt aus $V(\widehat{h}(y)(\omega)) + \widehat{h}(y)(\omega)I(\widehat{h}(y)(\omega)) \leq \sup_{x>0} V(x) + xI(\widehat{h}(y)(\omega)) = U(I(\widehat{h}(y)(\omega)))$. Die Eindeutigkeit der optimalen Lösung folgt aus der strikten Konkavität von U: Wären $\widehat{g}_1, \widehat{g}_2$

verschiedene Optimallösungen, so wäre $\frac{1}{2}\widehat{g}_1 + \frac{1}{2}\widehat{g}_2$ ebenfalls in $\mathcal{C}(x)$ und $E[U(\frac{1}{2}\widehat{g}_1 + \frac{1}{2}\widehat{g}_2)] > \frac{1}{2}E[U(\widehat{g}_1)] + \frac{1}{2}E[U(\widehat{g}_2)]$.

4. Die Aussagen $E(\widehat{g}(x)\widehat{h}(y)) = xy$ und $u'(x) = E[\frac{\widehat{g}(x)U'(\widehat{g}(x))}{x}]$ ergeben sich nun aus $v'(y) = E[\frac{\widehat{h}(y)V'(\widehat{h}(y))}{y}]$, indem man die Relationen $u'(x) = y$, $x = -v'(y)$, $\widehat{g}(x) = -V'(\widehat{h}(y))$ und $\widehat{h}(y) = U'(\widehat{g}(x))$ benutzt.

■

Ergebnisse für das ursprüngliche Problem

Die Aussagen der beiden Theoreme für $\mathcal{C}$ und $\mathcal{D}$ ergeben nun entsprechende Aussagen für das ursprüngliche Optimierungsproblem. Die folgenden Theoreme listen die wichtigsten auf:

Theorem 5.4 *Es gebe ein äquivalentes lokales Martingalmaß, die Nutzenfunktion U sei streng monoton steigend, streng konkav, stetig differenzierbar und erfülle $\lim_{x\to 0} U'(x) = \infty$, $\lim_{x\to\infty} U'(x) = 0$ sowie $AE(U) < 1$. Weiterhin gebe es ein x, so daß $\sup_{X\in\mathfrak{X}(x)} E[U(X_T)] < \infty$. Dann gilt:*

Es gibt eine eindeutige Lösung $\widehat{X}(x) \in \mathfrak{X}(x)$ des Optimierungsproblems. Sie ist gegeben durch die Gleichung

$$\widehat{X}_T(x) = I(\widehat{Y}_T(y))$$

mit $y = u'(x)$ wobei $\widehat{Y}(y)$ die (ebenfalls eindeutige) Lösung des dualen Optimierungsproblems ist und die Tatsache, daß $\widehat{X}(x)\widehat{Y}(y)$ ein gleichmäßig integrierbares Martingal ist.

Dieses Theorem legt folgende Lösungstechnik nahe: Man sucht zu gegebenem x das passende y über die Gleichung $y = u'(x)$, löst das dazugehörige duale Problem (da man nur $\widehat{Y}_T(y)$ benötigt, kann man auch die mit den Mengen $\mathcal{C}$ und $\mathcal{D}$ umformulierte Version des Problems verwenden) und erhält $\widehat{X}_T(x)$ sowie $\widehat{X}_T(x)\widehat{Y}_T(y)$. Da der Prozeß $\widehat{X}(x)\widehat{Y}(y)$ ein gleichförmig integrierbares Martingal ist, gilt $(\widehat{X}(x)\widehat{Y}(y))_t = E[\widehat{X}_T(x)\widehat{Y}_T(y)|\mathcal{F}_t]$, woraus man wiederum u.U. $\widehat{X}_t(x)$ gewinnen und damit eine Handelsstrategie aufstellen kann. Dabei stellt sich allerdings das Problem, daß man u' nicht kennt, also das zu x passende y nicht ohne weiteres finden kann. Dieses Problem kann man umgehen, indem man sich ein beliebiges y vorgibt und das dazu passende

x mit Hilfe des Prozesses $\widehat{X}_t$ erhält. Ein weiteres Problem ist die Lösbarkeit des dualen Problems und die Berechnung der bedingten Erwartungen. Diese Probleme werden im folgenden Spezialfall wesentlich „entschärft".

Theorem 5.5 *Es seien die Voraussetzungen des vorhergehenden Theorems erfüllt und zusätzlich gebe es genau ein äquivalentes lokales Martingalmaß Q. Dann gilt:*

$$\widehat{X}_T(x) = I(y\frac{dQ}{dP})$$

mit $-x = v'(y) = E[\frac{dQ}{dP}V'(y\frac{dQ}{dP})]$ und $\widehat{X}(x)$ ist ein gleichmäßig integrierbares Martingal unter Q

Wie man sieht, kann man in diesem Fall das Problem relativ leicht lösen. Als erstes bestimmt man die Dichte von Q bezüglich P und berechnet $E[\frac{dQ}{dP}V'(y\frac{dQ}{dP})]$ in Abhängigkeit von y. Dann wählt man y zu x passend und erhält unmittelbar $\widehat{X}_T(x)$, woraus man wiederum durch Bilden der bedingten Erwartung den Prozeß $\widehat{X}(x)$ erhält und die entsprechende Handelsstrategie aufstellen kann.

5.2.2 Vollständigkeit und Arbitragefreiheit

Um die Theoreme des letzten Unterabschnittes nutzen zu können, muß man bestimmen können, ob ihre Voraussetzungen im jeweiligen Fall zutreffen. Die Voraussetzungen $\lim_{x\to 0} U'(x) = \infty$ und $\lim_{x\to\infty} U'(x) = 0$ sind einfach nachzuprüfen, ebenso, ob $AE(U) < 1$ ist. Die Voraussetzung $u(x) < \infty$ für ein x ist zwar nicht ohne weiteres nachzuprüfen, trifft sie nicht zu, so wäre allerdings das gestellte Problem nicht mehr sinnvoll. Wir beschäftigen uns daher mit der letzten verbleibenden Voraussetzung, der Existenz (und ggf. Eindeutigkeit) eines äquivalenten lokalen Martingalmasses. Es gilt nach Delbaen/Schachermayer [4] folgender Zusammenhang:

Theorem 5.6 *Sei S ein lokal beschränktes Semimartingal. Dann existiert ein äquivalentes lokales Martingalmaß genau dann, wenn es keine Folge $X^n \in \mathfrak{X}^b$, $X_T^n \to X_T$ f.s. gibt, für die $(x - X^n)_+$ in der Supremumsnorm gegen 0 konvergiert und für die $P(X_T > 0) > 0$ gilt.*

Die ökonomische Interpretation dieser Aussage ist unmittelbar einzusehen: Gäbe es eine solche konvergente Folge, so könnte man eine Strategie, die zumindest die Möglichkeit eines Gewinnes verspricht, ohne ein Risiko aufzuweisen, beliebig gut simulieren (im Sinne der Beschränkung des Risikos mit einer beliebig vorgegebenen Schranke). Eine solche Strategie bezeichnet man als Arbitrage (daher der Begriff Arbitragefreiheit). Wir können daher davon ausgehen, daß es für einen ökonomisch sinnvollen Preisprozeß immer ein äquivalentes lokales Martingalmaß gibt.

Die Frage der Eindeutigkeit eines äquivalenten Martingalmasses ist ebenfalls von Bedeutung: Kann man die Eindeutigkeit des Martingalmasses zeigen, so wird das duale Problem recht einfach lösbar. Aussagen über die Eindeutigkeit des Martingalmasses sind in Spezialfällen bekannt, z.B. bei Verwendung der Brownschen Bewegung als Preisprozeß. Die Frage der Eindeutigkeit ist verbunden mit dem Begriff der Vollständigkeit:

Definition 5.2.6 *Ein Preisprozeß S (bzw. der durch S beschriebene Aktienmarkt) heißt vollständig, wenn es zu jeder von unten beschränkten Zufallsvariable X eine Zahl x und einen zulässigen Handelsprozeß φ gibt mit $X = x + \int \varphi dS$.*

Man kann zeigen, daß die Eigenschaft der Vollständigkeit mit der Eindeutigkeit des Martingalmasses im Wesentlichen äquivalent ist.

Bemerkung 5.2.7 *Die in der Literatur vorhandenen Aussagen zu diesem Thema sind relativ speziell, oft nur für Brownsche Bewegung als Preisprozeß, auch ist der Nachweis der Vollständigkeit oft schwierig, so daß sich keine echte Vereinfachung der uns eigentlich interessierenden Optimierungsaufgabe ergibt. Aus diesen Gründen wird an dieser Stelle nicht weiter auf Details dieses Zusammenhangs eingegangen.*

Bemerkung 5.2.8 *Da man zur Lösung des dualen Problems ohnehin ein „optimales" äquivalentes lokales Martingalmaß finden muß, dienen die Aussagen über Vollständigkeit und Arbitragefreiheit eher dem grundsätzlichen Verständnis des Problems als der konkreten Lösung.*

5.3 Die Martingalmethode mit Lévy-Prozessen als Preisprozesse

Wir spezifizieren nun das verwendete Modell weiter gegenüber dem am Anfang dieses Kapitels beschriebenen: Wir verwenden nun exponentielle Lévy-Prozesse (d.h. Prozesse der Form $S^i = e^{L^i}, i = 1, 2, \ldots, d$, L^i untereinander unabhängige Lévy-Prozesse) als Preisprozesse. Nach Kallsen [12] sind das genau die stochastischen Exponentiale von Lévy-Prozessen, d.h. Prozesse der Form $S^i = \mathcal{E}(L^i)$. Weiterhin lassen wir nur Handelsprozesse mit nichtnegativem Vermögensprozeß zu und betrachten einen endlichen Zeitraum $[0, T]$. Ein erster Schritt zur Lösung ist es, die „duale" Menge $\mathfrak{Y}$ zu betrachten. Nach unserer Definition ist ein Prozeß Y mit $Y_0 = y$ in $\mathfrak{Y}(y)$, falls XY für alle $X \in \mathfrak{X}(1)$ ein Supermartingal ist. Diese Eigenschaft ist offenbar schwer nachzuprüfen. Für lokal beschränkte Preisprozesse S allerdings gilt (wie bereits erwähnt): Ist S ein lokales Martingal unter dem äquivalenten Maß Q, so ist jedes $X \in \mathfrak{X}(1)$ ebenfalls ein lokales Martingal unter Q. Diesen Zusammenhang benutzt das folgende

Lemma 5.7 *Sei S ein lokal beschränktes, nichtnegatives Semimartingal sowie Q ein zum ursprünglichen Wahrscheinlichkeitsmaß P äquivalentes Wahrscheinlichkeitsmaß, unter dem S ein lokales Martingal ist . Dann ist der Prozeß $(E[\frac{dQ}{dP}|\mathcal{F}_t])_{0 \leq t \leq T}$ ein Element der Menge $\mathfrak{Y}$.*

Beweis: Nach Kramkov/Schachermayer [14] ist jedes $X \in \mathfrak{X}(1)$ ein lokales Martingal unter Q. Somit gibt es eine gegen T konvergente aufsteigende Folge von Stoppzeiten T_n, so daß X^{T_n} ein gleichmäßig integrierbares Martingal unter Q. Damit ist wiederum $X^{T_n} E[\frac{dQ}{dP}|\mathcal{F}_.]$ ein Martingal unter P (vgl. 1. Kapitel). Da $X^{T_n} E[\frac{dQ}{dP}|\mathcal{F}_.]$ nichtnegativ ist, kann man das Lemma von Fatou

anwenden, um nachzuweisen, daß $XE[\frac{dQ}{dP}|\mathcal{F}_.]$ ein Supermartingal ist:

$$\begin{aligned}
&E[X_t E[\frac{dQ}{dP}|\mathcal{F}_t]|\mathcal{F}_s] \\
&= E[\limsup_{n\to\infty} X_t^{T_n} E[\frac{dQ}{dP}|\mathcal{F}_t]|\mathcal{F}_s] \\
&\geq \limsup_{n\to\infty} E[X_t^{T_n} E[\frac{dQ}{dP}|\mathcal{F}_t]|\mathcal{F}_s] \\
&= \limsup_{n\to\infty} X_s^{T_n} E[\frac{dQ}{dP}|\mathcal{F}_s] \\
&= X_s E[\frac{dQ}{dP}|\mathcal{F}_s]
\end{aligned}$$

Da X beliebig war, ist die Behauptung gezeigt. ■

Korollar 5.8 *Sei S ein lokal beschränktes Semimartingal, betrachtet werde das zugehörige Optimierungsproblem, für das die Voraussetzungen des Theorems 5.4 erfüllt seien, sei Q ein äquivalentes Maß, unter dem S ein lokales Martingal ist. Sei weiter ein Anfangsvermögen x und eine Zahl y_0 mit $x = E[\frac{dQ}{dP}U'^{-1}(y_0\frac{dQ}{dP})]$. Dann gilt:*

$$u(x) \leq E[U(U'^{-1}(y_0\frac{dQ}{dP}))]$$

Beweis:

$$\begin{aligned}
u(x) &= \inf_{y>0}[v(y) + xy] \\
&\leq \inf_{y>0}[E[V(y\frac{dQ}{dP})] + xy] \\
&= \inf_{y>0}[E[U(U'^{-1}(y\frac{dQ}{dP})) - y\frac{dQ}{dP}U'^{-1}(y\frac{dQ}{dP})] + xy] \\
&= \inf_{y>0}[E[U(U'^{-1}(y\frac{dQ}{dP}))] + yE[x - \frac{dQ}{dP}U'^{-1}(y\frac{dQ}{dP})]] \\
&\leq E[U(U'^{-1}(y_0\frac{dQ}{dP}))]
\end{aligned}$$

■

Bemerkung 5.3.1 *Die erste Ungleichung in der eben getätigten Abschätzung läßt sich nicht verändern, wenn man nur ein äquivalentes lokales Martingalmaß kennt. Die zweite (d.h. die spezielle Wahl von y) erscheint willkürlich, erweist sich aber als zweckmäßig (siehe Kallsen [12])*

Findet man somit eine Handelsstrategie zum Anfangsvermögen x, deren Endvermögen gerade $U'^{-1}(y_0 \frac{dQ}{dP})$ beträgt, so hat man die optimale Handelsstrategie gefunden.

Um dieses Ergebnis für Lévy-Prozesse anwenden zu können, ist folgendes Lemma nützlich:

Lemma 5.9 *Ein Lévy-Prozeß mit beschränkten Sprüngen ist lokal beschränkt.*

Beweis: Sei L ein gegebener Lévy-Prozeß, C eine obere Schranke für seine Sprünge. Wir definieren eine aufsteigende Folge von Stoppzeiten durch $T_0 := 0, T_n := \min\{\inf\{t \geq 0, |L_t| \geq n\}, T\}$. Dann ist $L^{T_n} \leq n + C$ (denn $|L_t^{T_n}| \leq n$ $\forall t < T_n$ und $|L_{T_n}^{T_n}| \leq |L_{T_n-}^{T_n}| + \Delta L_{T_n} \leq n + C$). Es bleibt zu zeigen, daß $T_n \to T$ $P - f.s.$ Angenommen $P(\sup_n T_n < T) > 0$. Dann gibt es wegen $P(\sup_n T_n < T) = P(\bigcup_{m \in \mathbb{N}} \{\omega | \sup_n T_n < T - \frac{1}{m}\}) = \lim_{n\to\infty} P(\sup_n T_n < T - \frac{1}{m})$ ein $m \in \mathbb{N}$ mit $P(\sup_n T_n < T - \frac{1}{m}) > 0$. Da fast alle Pfade càdlàg sind, seien o.B.d.A. die zu $\omega \in \{\omega | \sup_n T_n(\omega) < T - \frac{1}{m}\}$ gehörenden Pfade càdlàg. ω sei nun fest, aber beliebig aus dieser Menge gewählt. Da der zu ω gehörende Pfad links- und rechtsseitige Grenzwerte besitzen, gibt es um den Zeitpunkt $\sup_n T_n(\omega)$ ein Intervall $(\sup_n T_n(\omega) - \epsilon, \sup_n T_n(\omega) + \epsilon)$ mit $\epsilon > 0$, in dem der Betrag von L beschränkt ist (o.B.d.A. durch $n_0 \in \mathbb{N}$). Wählen wir nun ein $n_1 > n_0$, so daß $T_{n_1}(\omega) \in (\sup_n T_n(\omega) - \frac{\epsilon}{2}, \sup_n T_n(\omega))$, so enthält nach Definition jede Umgebung von $T_{n_1}(\omega)$ Zeitpunkte, an denen $|L| \geq n_1$ ist. Dann gibt es aber auch in $(\sup_n T_n(\omega) - \epsilon, \sup_n T_n(\omega) + \epsilon)$ einen Zeitpunkt mit $|L| \geq n_1 \geq n_0$, im Widerspruch zur Voraussetzung. ■

Um die optimale Handelsstrategie für allgemeine Lévy-Prozesse angeben zu können, hier noch zwei Resultate von Kallsen [12], die auf der eben entwickelten Technik beruhen: Man sucht ein äquivalentes Martingalmaß und erhält mit etwas Glück bereits das Optimum.

Theorem 5.10 *Sei $S^i = \mathcal{E}(L^i)$, wobei $L = (L^i)_{i=1,2,\ldots,d}$ ein Lévy-Prozeß mit Triplett (α, β^2, ν) sei. Sei $U(x) := \ln x$. Es existiere weiter ein $\gamma \in \mathbb{R}^d$ mit*

1. $\nu(\{x \in \mathbb{R}^d | 1 + \gamma^\top x \leq 0\}) = 0$

2. $\int |\frac{x}{(1+\gamma^\top x)} - 1_{(0,1)}(x)|\nu(dx) < \infty$

3. $\alpha - \beta\beta^\top\gamma + \int(\frac{x}{(1+\gamma^\top x)} - 1_{(0,1)}(||x||))\nu(dx) = 0$

Dann ist $\varphi_t^i := \frac{\gamma^i}{S_{t-}^i} x\mathcal{E}(\gamma^\top L)_{t-}, i = 1, 2, \ldots, d$, $\varphi_t^0 := \int_0^t \varphi_s^\top dS_s - \sum_{i=1}^d \varphi_t^i S_t^i$ *eine optimale Handelsstrategie.*

Theorem 5.11 *Sei* $S^i = \mathcal{E}(L^i)$, *wobei* $L = (L^i)_{i=1,2,\ldots,d}$ *ein Lévy-Prozeß mit Triplett* (α, β^2, ν) *sei. Sei* $U(x) := \frac{x^{1-p}}{1-p}$ *für* $p \in (0,1)$. *Es existiere weiter ein* $\gamma \in \mathbb{R}^d$ *mit*

1. $\nu(\{x \in \mathbb{R}^d | 1 + \gamma^\top x \leq 0\}) = 0$

2. $\int |\frac{x}{(1+\gamma^\top x)^p} - 1_{(0,1)}(x)|\nu(dx) < \infty$

3. $\alpha - p\beta\beta^\top\gamma + \int(\frac{x}{(1+\gamma^\top x)^p} - 1_{(0,1)}(||x||))\nu(dx) = 0$

Sei außerdem $a := \frac{1-p}{2}(\gamma^\top\beta)^2 + \frac{1}{p}\int(\frac{1+p\gamma^\top x}{(1+\gamma^\top x)^p} - 1)\nu(dx)$ *und* $A := (t)_{t\geq 0}$. *Dann ist* $\varphi_t^i := \frac{\gamma^i}{S_{t-}^i} x e^{-at}\mathcal{E}(\gamma^\top L + aA)_{t-}, i = 1, 2, \ldots, d$ *und* $\varphi_t^0 := \int_0^t \varphi_s^\top dS_s - \sum_{i=1}^d \varphi_t^i S_t^i$ *eine optimale Handelsstrategie.*

Bemerkung 5.3.2 *Sowohl die numerierten Voraussetzungen als auch die Aussage des zweiten Terms ergeben die entsprechenden Voraussetzungen und Aussagen des ersten Theorems als Spezialfall* $p = 1$. *Die ersten beiden Voraussetzungen sind eher technischer Natur, die dritte ist die wesentliche Voraussetzung, um ein äquivalentes lokales Martingalmaß und eine optimale Handelsstrategie zu finden.*

Bemerkung 5.3.3 *Offenbar gilt* $\frac{\varphi_t^i S_{t-}^i}{\varphi_t^\top S_{t-} - \varphi_t^0 S_t^0} = \frac{\gamma^i}{\gamma}$, *d.h. die optimale Strategie weist eine zeitlich konstante Aufteilung des „riskant".angelegten Vermögensteils in die einzelnen Aktien auf, wie es bereits in der Kontrolltheorie der Fall war.*

5.4 Anwendung auf die Portfoliooptimierung

5.4.1 Beispiel 1: Black-Scholes-Merton ohne Transaktionskosten

Zur Erinnerung: Gegeben sind Preisprozesse der Form

$$dS^i = S^i(\alpha^i dt + \sum_{l=1}^{m} \sigma^{il} dw^l), \quad i = 0, 1, 2, \ldots, d$$

wobei die erste Spalte der Matrix σ 0 sei, es sind beliebige selbstfinanzierende Handelsstrategien mit nichtnegativem Vermögensverlauf zulässig, das Zielfunktional ist der erwartete Endnutzen $E[U(V_T)]$. Der Preisprozeß S ist damit gegeben als $S^i = \mathcal{E}(L^i)$ mit Lévy-Prozessen L^i, deren charakteristisches Triplett (als vektorwertiger Prozeß) $(\alpha, \sigma\sigma^\top, 0)$ ist. Dann ist die erste und zweite Forderung der beiden Theoreme 5.10 und 5.11 offensichtlich für beliebiges γ erfüllt. Die dritte Forderung wird zu

$$\alpha = p\sigma\sigma^\top\gamma$$

und schließlich $\gamma = \frac{1}{p}(\sigma\sigma^\top)^{-1}\alpha$. Weiter ist dann

$$\begin{aligned} a &= \frac{1-p}{2}(\gamma^\top\sigma\sigma^\top)(\gamma^\top\sigma\sigma^\top)^\top \\ &= \frac{1-p}{2}p^2\alpha^\top\alpha \end{aligned}$$

Die zugehörige Handelsstrategie ist

$$\varphi_t^i := \frac{\gamma^i}{S_t^i} x e^{-at} \mathcal{E}(\gamma^\top L + aA)_t, \quad i = 1, 2, \ldots, d$$

Dieser Term soll nun berechnet in Abhängigkeit von S allein ausgedrückt werden unter Verwendung der Gleichungen $\mathcal{E}(\lambda B) = \exp(\lambda B - \frac{\lambda^2}{2}t)$ für eine Standard-Brownsche Bewegung B und $\mathcal{E}(X + Y + [X, Y]) = \mathcal{E}(X)\mathcal{E}(Y)$. Es

gilt also:

$$
\begin{aligned}
\mathcal{E}(\gamma^\top L + aA)_t &= \mathcal{E}(\gamma^\top \sigma B + (a+\gamma^\top\alpha)A)_t \\
&= \mathcal{E}(\sum_{l=1}^m (\gamma^\top\sigma)^l B^l + (a+\gamma^\top\alpha)A))_t \\
&= \exp(\sum_{l=1}^m (\gamma^\top\sigma)^l B^l + (a+\gamma^\top\alpha - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t) \\
&= \exp(\sum_{l=1}^m \sum_{i=1}^d \gamma^i\sigma^{il}B^l - (\frac{1}{2}\sum_{l=1}^m\sum_{i=1}^d \gamma^i(\sigma^{il})^2)t + (\frac{1}{2}\sum_{l=1}^m\sum_{i=1}^d \gamma^i(\sigma^{il})^2)t \\
&\quad + (a+\gamma^\top\alpha - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t) \\
&= (\prod_{i=1}^d \exp(\sum_{l=1}^m \sigma^{il}B^l - \frac{1}{2}\sum_{l=1}^m(\sigma^{il})^2 t + \alpha^i t)^{\gamma^i}) \exp((\frac{1}{2}\sum_{l=1}^m\sum_{i=1}^d \gamma^i(\sigma^{il})^2)t \\
&\quad + (a - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t) \\
&= (\prod_{i=1}^d \mathcal{E}(\sum_{l=1}^m \sigma^{il}B^l + \alpha^i A)_t^{\gamma^i}) \exp((\frac{1}{2}\sum_{l=1}^m\sum_{i=1}^d \gamma^i(\sigma^{il})^2)t \\
&\quad + (a - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t) \\
&= (\prod_{i=1}^d (S_t^i)^{\gamma^i}) \exp((\frac{1}{2}\sum_{l=1}^m\sum_{i=1}^d \gamma^i(\sigma^{il})^2)t + (a - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t) \\
&= \exp((\frac{1}{2}\sum_{i=1}^d \gamma^i(\sigma\sigma^\top)^{ii})t + (a - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t) \prod_{i=1}^d (S_t^i)^{\gamma^i}
\end{aligned}
$$

Insgesamt ist damit

$$
\varphi_t^i = \frac{\gamma^i}{S_t^i}\prod_{i=1}^d (S_t^i)^{\gamma^i} x \exp((\frac{1}{2}\sum_{i=1}^d \gamma^i(\sigma\sigma^\top)^{ii} - \frac{1}{2}\gamma^\top\sigma\sigma^\top\gamma)t), \quad i = 1,2,\ldots,d
$$

und

$$
\varphi_t^0 = \int_0^t \varphi_s^\top dS_s - \sum_{i=1}^d \varphi_t^i S_t^i
$$

5.4.2 Beispiel 2: Black-Scholes-Merton mit Transaktionskosten

Wie bereits aus dem darstellenden Teil hervorgeht, sind Modelle mit Transaktionskosten nicht mit Hilfe der Martingalmethode behandelbar.

5.4.3 Beispiel 3: Das Modell von Kallsen

Dieses Modell unterscheidet sich vom klassischen Black-Scholes-Merton-Modell nur durch die Verwendung von $S^i = \exp(L^i)$ als Preisprozeß, wobei $L = (L^i)_{i=1,2,\ldots,d}$ ein Lévy-Prozeß mit Triplett (α, β^2, ν) ist. Um die Theoreme 5.10 oder 5.11 anwenden zu können, muß das charakteristische Triplett $(\widetilde{\alpha}, \widetilde{\beta}^2, \widetilde{\nu})$ von $\widetilde{L}$ mit $\mathcal{E}(\widetilde{\mathcal{L}}) = \exp(L)$ berechnet werden. Dieses ergibt sich nach Kallsen [12] zu

$$\begin{aligned}
\widetilde{\alpha} &= \alpha + \frac{\beta^2}{2} + \int ((e^x - 1)1_{(-1,1)}(e^x - 1) - x1_{(-1,1)}(x))\nu(dx) \\
\widetilde{\beta}^2 &= \beta^2 \\
\widetilde{\nu}(G) &= \int 1_G(e^x - 1)\nu(dx)
\end{aligned}$$

Ist dabei ν durch eine Dichte f_ν bezüglich des Lebesgue-Maßes gegeben, so besitzt auch $f_{\widetilde{\nu}}$ eine Dichte. Diese erhält man mit folgender Berechnung:

$$\begin{aligned}
\widetilde{\nu}(G) &= \int 1_G(e^x - 1)\nu(dx) \\
&= \int 1_{G\cap(-1,\infty)}(e^x - 1)f_\nu(x)dx \\
&= \int 1_{G\cap(-1,\infty)}(y)f_\nu(\ln(1+y))\frac{1}{1+y}dy
\end{aligned}$$

Die Dichte von $\widetilde{\nu}$ ist somit $1_{(-1,\infty)}(y)f_\nu(\ln(1+y))\frac{1}{1+y}$. Um über die Aussagen der Theoreme 5.10 oder 5.11 gehende Berechnungen zu machen, muß der verwendete Lévy-Prozeß genauer bestimmt werden.

Beispiel 3.1: Ein Prozeß mit gleichverteilten, beschränkten Sprüngen, $d = 1$

Als Beispiel sei nun $f_\nu = \mathbf{1}_{(-1,1)}$. Berechnet man $\widetilde{\alpha}$ und $\widetilde{\nu}$, so erhält man:

$$\begin{aligned}
\widetilde{\alpha} - (\alpha + \frac{\beta^2}{2}) &= \int (e^x - 1) 1_{(-1,1)} (e^x - 1) \nu(dx) \\
&= \int_{-\infty}^{\ln 2} (e^x - 1) \nu(dx) \\
&= \int_{-1}^{1} (e^x - 1) dx \\
&= e - \frac{1}{e} - 2
\end{aligned}$$

sowie

$$f_{\widetilde{\nu}}(x) = 1_{(\frac{1}{e}-1, e-1)}(x) \frac{1}{1+x}$$

Die dritte Voraussetzung des Theorems lautet somit

$$\widetilde{\alpha} - p\widetilde{\beta}^2 \gamma + \int (\frac{x}{(1+\gamma x)^p} - 1_{(0,1)}(x)) 1_{(\frac{1}{e}-1, e-1)}(x) \frac{1}{1+x} dx = 0$$

Dieses Integral kann nur numerisch ausgewertet werden (in Abhängigkeit von γ). Danach ist (ebenfalls numerisch) die Lösung γ der Gleichung zu bestimmen. Gleichzeitig mit dieser Berechnung überprüft man auch die Existenz des Integrals $\int (\frac{x}{(1+\gamma x)^p} - 1_{(0,1)}(x)) 1_{(\frac{1}{e}-1, e-1)}(x) \frac{1}{1+x} dx$, d.h. die zweite Bedingung der Theoreme 5.10 bzw. 5.11. Die erste Bedingung ist, wenn man γ berechnet hat, sehr einfach nachzuprüfen: Es muß $(\frac{1}{e} - 1, e - 1) \cap \{x | 1 + \gamma x \leq 0\} = \emptyset$ sein. Diese Bedingung ist explizit als Ungleichung für γ anzugeben: Ist $\gamma < 0$, so ist $\{x | 1 + \gamma x \leq 0\} = [-\gamma, \infty)$, d.h. es muß $-\gamma \geq e - 1$ bzw. $\gamma \leq 1 - e$ sein. Ist $\gamma \geq 0$, so ergibt sich $\{x | 1 + \gamma x \leq 0\} = (-\infty, \gamma]$, d.h. γ erfüllt die erste Bedingung der beiden Theoreme nicht. Zusammenfassend ist die erste Bedingung also für $\gamma \leq 1 - e$ erfüllt.

Beispiel 3.2: Ein Prozeß mit deterministischer Sprunghöhe, $d = 1$

Als weiteres Beispiel wählen wir $\nu = \lambda\delta_{-1}$. Berechnet man nun $\widetilde{\alpha}$ und $\widetilde{\nu}$, so erhält man

$$\begin{aligned}\widetilde{\alpha} - (\alpha + \frac{\beta^2}{2}) &= \int ((e^x - 1)1_{(-1,1)}(e^x - 1) - x1_{(-1,1)}(x))\nu(dx) \\ &= \lambda(e^{-1} - 1)\end{aligned}$$

und

$$\begin{aligned}\widetilde{\nu}(G) &= \int 1_G(e^x - 1)\nu(dx) \\ &= \lambda 1_G(e^{-1} - 1)\end{aligned}$$

also $\widetilde{\nu} = \lambda\delta_{e^{-1}-1}$. Die dritte Voraussetzung des Theorems lautet damit:

$$\widetilde{\alpha} - p\widetilde{\beta}^2\gamma + \lambda(\frac{e^{-1} - 1}{(1 + \gamma(e^{-1} - 1))^p} - 1) = 0$$

bzw.

$$\widetilde{\alpha}(1 + \gamma(e^{-1} - 1))^p - p\widetilde{\beta}^2\gamma(1 + \gamma(e^{-1} - 1))^p + \lambda(e^{-1} - 1 - (1 + \gamma(e^{-1} - 1))^p) = 0$$

Wählen wir zur weiteren Vereinfachung $p = 1$, so ergibt sich

$$\begin{aligned}\widetilde{\alpha}(1 + \gamma(e^{-1} - 1)) - \widetilde{\beta}^2\gamma(1 + \gamma(e^{-1} - 1)) + \lambda(e^{-1} - 1 - (1 + \gamma(e^{-1} - 1))) = 0 \\ \gamma^2[(1 - e^{-1})\widetilde{\beta}^2] + \gamma[\widetilde{a}(e^{-1} - 1) - \widetilde{\beta}^2 - \lambda(e^{-1} - 1)] + [\widetilde{a} + \lambda(e^{-1} - 2)] = 0\end{aligned}$$

Die Lösungen dieser quadratischen Gleichung sind

$$\begin{aligned}\gamma = &\frac{\widetilde{a}(1 - e^{-1}) + \widetilde{\beta}^2 + \lambda(e^{-1} - 1)}{2(1 - e^{-1})\widetilde{\beta}^2} \\ &\frac{\pm\sqrt{[\widetilde{a}(e^{-1} - 1) - \widetilde{\beta}^2 - \lambda(e^{-1} - 1)]^2 - 4[(1 - e^{-1})\widetilde{\beta}^2][a + \lambda(e^{-1} - 2)]}}{2(1 - e^{-1})\widetilde{\beta}^2}\end{aligned}$$

Die zweite Bedingung des Theorems ist ohnehin erfüllt, da ν diskret ist. Die erste Bedingung wiederum lautet $\widetilde{\nu}(\{x \in \mathbb{R}^d | 1 + \gamma^\top x \leq 0\}) = 0$ bzw. $1 + \gamma(e^{-1} - 1) > 0$ und somit $\gamma < \frac{e}{e-1}$. Hat man konkrete Parameter α, β und γ vorgegeben, so ist direkt zu sehen, ob diese Bedingung erfüllt ist, und die optimale Handelsstrategie ergibt sich daraus.

5.5 Bewertung der Martingalmethode

Wie aus den zuletzt behandelten Beispielen hervorgeht, weist die Martingalmethode folgende Vorteile auf:

Sie erlaubt für Preisprozesse, die von Lévy-Prozessen erzeugt werden, einen relativ einfachen Lösungsweg, sowie die Erkenntnis, daß in diesem Fall die Verhältnisse der Vermögenswerte in den einzelnen riskanten Anlagen konstant sind, lediglich das Verhältnis des Anlagewerts im Bankguthaben und dem Rest variiert ggf. Für eine exponentielle Brownsche Bewegung kann man die explizite Lösung (die mit der in der Kontrolltheorie auch tatsächlich übereinstimmt) finden.

Sie hat ferner den Vorteil, daß sie im Prinzip für beliebige Semimartingale anwendbar ist, und daß mit ihr die bisweilen recht unangenehmen partiellen Differentialgleichungen der Kontrolltheorie vermieden werden.

Allerdings kann man mit der Martingalmethode Modelle mit Transaktionskosten nicht behandeln, und für allgemeine Semimartingale kann die Bestimmung des äquivalenten Martingalmaßes schwierig oder nicht durchführbar sein.

Zusammenfassung

In der vorliegenden Arbeit wurden zunächst die mathematischen Grundlagen der Portfoliooptimierung in Grundzügen dargestellt.

Im ersten Kapitel wurden neben Grundbegriffen einige Prozessklassen, die im weiteren Verlauf benötigt wurden, eingeführt, unter anderem Markov- und Lévy-Prozesse.

Im zweiten Kapitel wurde der Weg zur Definition eines stochastischen Integrals beschrieben, die Itô-Formel vorgestellt sowie der Begriff der stochastischen Differentialgleichung erläutert.

Darauf aufbauend wurde im dritten Kapitel ein allgemein gehaltenes Modell der Portfoliooptimierung aufgestellt sowie beschrieben, wie mit diesem einige in der Literatur verwendete Modelle dargestellt werden können.

Im vierten und fünften Kapitel wurden die beiden gängigen Lösungsmethoden des Problems vorgestellt:

Im vierten Kapitel wurde die stochastische Kontrolltheorie verwendet, um die Optimierungsaufgabe zu lösen: Als Preisprozesse können im allgemeinsten Fall Markov-Prozesse verwendet werden. Es wird (aufbauend auf der Markov-Eigenschaft) das Prinzip der dynamischen Programmierung verwendet, um eine partielle Differentialgleichung, die HJB-Gleichung, für die Wertfunktion aufzustellen. Kann man diese lösen (evtl. im schwachen Sinne), so erhält man die Wertfunktion, mit deren Hilfe man wiederum die zugehörige optimale Handelsstrategie bestimmt. Die Forderung, daß die Preisprozesse Markov-Prozesse sein sollen, ist eine „Schwäche" der Lösungstechnik, eine weitere ist die unter Umständen aufwendige Lösung der HJB-Gleichung.

Im fünften Kapitel wurde eine andere Lösungsmethode, die sogenannte Martingalmethode vorgestellt. Man stellt ein zum eigentlichen Problem duales Optimierungsproblem auf, das oft einfacher, im Fall eines vollständigen Marktes sogar wesentlich einfacher als das ursprüngliche Problem zu lösen ist. Eine Beschränkung an den Preisprozeß findet hier nur insofern statt, als

man die Existenz eines äquivalenten Martingalmaßes fordert; diese Forderung ist aber im Wesentlichen äquivalent zur Arbitragefreiheit des Marktes, stellt also keine allzu große Einschränkung der Modellwahl dar. Allerdings kann die Bestimmung der äquivalenten Martingalmaße bisweilen aufwendig sein.

Literaturverzeichnis

[1] Akian, M., Menaldi, J.L., Sulem, A. (1996), On an investment-consumption model with transaction costs, SIAM Journal Control and Optimization **34** (1), 329-364

[2] Alt, H. W. (1999), Lineare Funktionalanalysis, Springer Verlag Berlin

[3] Bauer, H. (1991), Wahrscheinlichkeitstheorie, de Gruyter Berlin

[4] Delbaen, F,Schachermayer, W. (1994), A general version of the fundamental Theorem of asset pricing, Mathematische Annalen **300**, 463-520

[5] Emmer, S., Klüppelberg, C. (2002), Optimal Portfolios when stock prices follow an exponential Lévy-Process, Finance & Stochastics. To appear

[6] Fleming, W., Rishel, R. (1975), Deterministic and Stochastic Optimal Control, Springer Verlag Berlin

[7] Fleming, W., Soner, H. (1993), Controlled Markov Processes and Viscosity Solutions, Springer Verlag Berlin

[8] Goll, T., Kallsen, J. (2000), Optimal portfolios for logarithmic utility, Stochastic Processes and Their Applications **89**, 31-48

[9] Harrison, M., Kreps, D. (1979), Martingales and arbitrage in multiperiod securities markets, Journal of Economic Theory **20**, 381-408

[10] Harrison, M., Pliska, S. (1981), Martingales and stochastic integrals in the theory of continous trading, Stochastic Processes and Their applications **11**, 215-260

[11] Kabanov, Y., Klüppelberg, C. (2000), A geometric approach to portfoliooptimization in models with transaction costs. Submitted.

[12] Kallsen, J. (2000), Optimal portfolios for exponential Lévy processes, Mathematical Methods of Operations Research **51** , 357-374

[13] Korn, Ralf (1997), Optimal portfolios, World Scientific

[14] Kramkov, D., Schachermayer, W. (1999), Acondition on the asymptotic elasticity of utility functions and optimal investment in incomplete markets, The annals of applied probability **9**, 904-950

[15] Øksendal, B. (2000), Stochastic Differential Equations, Springer Verlag Berlin

[16] Pliska, M. (1986) A stochastic calculus model of continous trading: Optimal portfolios, Mathematics of Operations Research **11**, 371-384

[17] Protter, P. (1992), Stochastic Integration and Differential Equations, Springer Verlag Berlin

[18] Resnick, S. (1992), Adventures in Stochastic Processes, Birkhäuser Boston

[19] Sato, K. (1999), Lévy Processes and Infinitely Divisible Distributions, Cambridge University Press

[20] Schachermayer, W. (2001), Optimal Investment in Incomplete Markets when Wealth may Become Negative, Annals of Applied Probability **11**, S. 694-734.

[21] Schweizer, M. (1994), Approximating Random Variables by Stochastic Integrals, Annals of Probability **22**, 1536-1575

[22] Shiryaev, A. (2000), Essentials of Stochastic Finance, World Scientific, Singapur