Vorlesungsskript „Lévy-Prozesse und stochastische Kontrolltheorie”

Christoph Kühn

Sommersemester 2008

letzte Aktualisierung: 17. Juni 2022

Inhaltsverzeichnis

1 Einführung

Sei $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ ein filtrierter Wahrscheinlichkeitsraum, der die „üblichen Voraussetzungen" erfülle. $T \in \mathbb{R}_+$ ist der endliche Zeithorizont des Modells. Sei $W = (W_t^1, \ldots, W_t^m)_{t\in[0,T]}$ eine Standard m-dimensionale Brownsche Bewegung auf $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ (dies bedeutet insbesondere, dass $W^1, \ldots, W^m$ stochastisch unabhängig sind).

Definition 1.1. *Sei $k \in \mathbb{N}$, $U \subset \mathbb{R}^k$. Mit $\mathcal{U}_0$ bezeichnen wir die Menge der vorhersehbaren Prozesse mit Werten in U. $\nu \in \mathcal{U}_0$ wird* **Kontrollprozess** *oder* **Kontrolle** *genannt.*

Seien

$$b: \quad [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n \quad (t,x,u) \mapsto b(t,x,u)$$

und

$$\sigma: \quad [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n,m} \quad (t,x,u) \mapsto \sigma(t,x,u)$$

$\mathcal{B}([0,T] \times \mathbb{R}^n \times U) - \mathcal{B}(\mathbb{R}^n)$-messbare bzw. $\mathcal{B}([0,T] \times \mathbb{R}^n \times U) - \mathcal{B}(\mathbb{R}^{n,m})$-messbare Funktionen, die folgende **Lipschitz-Bedingung** erfüllen: es existiert ein $K \in \mathbb{R}_+$ s.d.

$$\begin{aligned} &||b(t,x,u) - b(t,y,u)|| + ||\sigma(t,x,u) - \sigma(t,y,u)|| \\ &\leq K||x-y||, \quad \forall (t,x,y,u) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times U, \end{aligned} \tag{1.1}$$

wobei $||A|| := \sqrt{\mathrm{Spur}\,(AA^\top)} = \sqrt{\sum_{i=1}^n \sum_{j=1}^m A_{ij}^2}$ die Hilbert-Schmidt Norm von $A \in \mathbb{R}^{n,m}$ bezeichnet.

Ein Kontrollprozess $\nu \in \mathcal{U}_0$ heißt **zulässig**, wenn

$$E\left(\int_0^T \left(||b(t,x,\nu_t)||^2 + ||\sigma(t,x,\nu_t)||^2\right)\, dt\right) < \infty, \quad \forall x \in \mathbb{R}^n. \tag{1.2}$$

Die Menge der zulässigen Kontrollprozesse wird mit $\mathcal{U}$ bezeichnet.

1.1 Starke Lösungen von stochastischen Differentialgleichungen

Theorem 1.2. *Für jeden zulässigen Kontrollprozess $\nu \in \mathcal{U}$ besitzt die stochastische Differentialgleichung (SDE)*

$$dX_t^i = b^i(t, X_t, \nu_t)\, dt + \sum_{j=1}^m \sigma^{ij}(t, X_t, \nu_t)\, dW_t^j, \quad i = 1, \ldots, n, \quad X_0 = x. \tag{1.3}$$

zu gegebenem Startwert $x \in \mathbb{R}^n$ eine eindeutige starke Lösung, d.h. es existiert ein (bis auf Ununterscheidbarkeit) eindeutiger adaptierter, stetiger Prozess $X = (X^1, \ldots, X^n)$, so dass

$$X^i = x^i + \int_0^{\cdot} b^i(t, X_t, \nu_t)\, dt + \sum_{j=1}^m \int_0^{\cdot} \sigma^{ij}(t, X_t, \nu_t)\, dW_t^j, \quad i = 1, \ldots, n. \tag{1.4}$$

Für die eindeutige Lösung X gilt

$$E\left(\sup_{t\in[0,T]} ||X_t||^2\right) < \infty. \tag{1.5}$$

Bemerkung 1.3. *Grob gesprochen bedeutet ,,starke Lösung", dass wir X als einen (stetigen) adaptierten Prozess auf dem* **ex ante gegebenen** *filtrierten Wahrscheinlichkeitsraum finden, während man bei einer ,,schwachen Lösung" den Wahrscheinlichkeitsraum möglicherweise noch vergrößern muss und dieser damit Teil der Lösung wird.*

Betrachte dazu als Beispiel die SDE

$$dX_t = \operatorname{sign}(X_t)\, dW_t \qquad \textit{mit} \quad X_0 = 0, \tag{1.6}$$

wobei $\operatorname{sign}(y) = 1$ *für* $y \geq 0$ *und* $\operatorname{sign}(y) = -1$ *für* $y < 0$. *Die Funktion* sign *ist offenbar nicht Lipschitz, Theorem 1.2 ist also nicht anwendbar. Bzgl. der Filtration* $(\mathcal{F}_t^W)_{t\in[0,T]}$, *die nur von der Brownschen Bewegung W erzeugt wird, besitzt die SDE auch keine Lösung (ohne Beweis, aber siehe Erklärung unten). Gibt man sich aber statt W, den Prozess X als Brownsche Bewegung vor und definiert*

$$\widetilde{W} := \operatorname{sign}(X) \bullet X,$$

dann gilt

$$\operatorname{sign}(X) \bullet \widetilde{W} = \operatorname{sign}(X) \bullet (\operatorname{sign}(X) \bullet X) = (\operatorname{sign}^2(X)) \bullet X = X. \tag{1.7}$$

Bzgl. der von X erzeugten Filtration $(\mathcal{F}_t^X)_{t\in[0,T]}$ *ist der Integrand* $\operatorname{sign}(X)$ *vorhersehbar und $\widetilde{W}$ damit ein lokales Martingal. Da zudem* $[\widetilde{W}, \widetilde{W}] = (\operatorname{sign}^2(X)) \bullet [X, X] = [X, X]$ *gilt, folgt mit dem Theorem von Lévy, dass $\widetilde{W}$ (wie W) eine Standard Brownsche Bewegung ist. Daher kann man wegen (1.7) sagen, dass (1.6) trotzdem eine Lösung hat (schwache Lösung). $\widetilde{W}$ ist an die von X erzeugte Filtration adaptiert, aber nicht umgekehrt.*

Wäre X eine starke Lösung von (1.6) auf der Filtration $(\mathcal{F}_t^W)_{t\in[0,T]}$, *dann gölte auch* $\operatorname{sign}(X) \bullet X = \operatorname{sign}(X) \bullet (\operatorname{sign}(X) \bullet W) = \operatorname{sign}^2(X) \bullet W = W$. *$W$ wäre somit an die von $|X|$ erzeugte Filtration* $(\mathcal{F}_t^{|X|})_{t\in[0,T]}$ *adaptiert (ohne Beweis, aber plausible, die Vorzeichenfunktion ist i.W. symmetrisch, da die Menge $\{X = 0\}$ nicht in das Integral eingeht, damit kann man aus* $\operatorname{sign}(X) \bullet X$ *das Vorzeichen von X nicht beobachten). Da die von $|X|$ erzeugte Filtration echt kleiner als die von X erzeugte Filtration ist, ist X nicht an* $(\mathcal{F}_t^W)_{t\in[0,T]}$ *adaptiert, ein Widerspruch. Also kann es keine starke Lösung auf der Filtration* $(\mathcal{F}_t^W)_{t\in[0,T]}$ *geben. Für einen formalen Beweis siehe z.B. Karatzas, Shreve [Brownian Motion and Stochastic Calculus].*

Bemerkung 1.4. *Man beachte, dass die Integrale auf der rechten Seite von (1.4) unter den Voraussetzungen (1.1) und (1.2) für alle stetigen, adaptierten Prozesse X existieren.*

Bemerkung 1.5. *Die dynamische Kontrolle besteht also in der Beeinflussung der lokalen Driftrate und der lokalen Volatilität des Prozesses X, wobei dies i.A. nicht unabhängig voneinander geschehen kann. ν^i, $i = 1, \ldots, k$ nennt man auch die* **Kontrollvariablen** *und X^i, $i = 1, \ldots, n$ die* **Zustandsvariablen**.

Beispiel 1.6 (Dynamischer Handel)**.** *Seien die Preisprozesse der verfügbaren Wertpapiere gegeben durch*

$$S_t^0 = \exp(rt) \qquad \textit{und} \qquad dS_t^i = S_t^i\left[\mu^i\,dt + \sum_{j=1}^m \widetilde{\sigma}^{i,j}\,dW_t^j\right], \quad i = 1,\ldots,d.$$

Spezifiziere wie üblich durch den vorhersehbaren Prozess φ^i die Anzahl an Wertpapieren vom Typ i, die eine Investorin in ihrem Portfolio hält. Wenn $(\varphi^0,\varphi^1,\ldots,\varphi^d)$ selbstfinanzierend ist, dann gilt für den dazugehörigen Vermögensprozess X:

$$\begin{aligned} dX_t &= \sum_{i=0}^d \varphi_t^i\,dS_t^i = \sum_{i=1}^d\left[\varphi_t^iS_t^i\mu^i\,dt + \sum_{j=1}^m \varphi_t^iS_t^i\widetilde{\sigma}^{i,j}\,dW_t^j\right] + \frac{\left(X_t - \sum_{i=1}^d \varphi_t^iS_t^i\right)}{S_t^0}S_t^0 r\,dt \\ &= rX_t\,dt + \sum_{i=1}^d\left[\nu_t^iX_t(\mu^i - r)\,dt + \sum_{j=1}^m \nu_t^iX_t\widetilde{\sigma}^{i,j}\,dW_t^j\right], \end{aligned} \tag{1.8}$$

wobei

$$\nu_t^i := \frac{\varphi_t^iS_t^i}{X_t} \tag{1.9}$$

der Anteil des Gesamtvermögens ist, das zum Zeitpunkt t im i-ten Wertpapier investiert ist. Die Substitution (1.9) macht deutlich, dass für die Beschreibung der Vermögensdynamik die Zustände der Wertpapierpreise nicht zwingend benötigt werden. (1.8) passt in den Rahmen von (1.3) mit $n = 1$ und $k = d$: Der Vermögensprozess ist die eindimensionale Zustandsvariable, die Anzahl der Kontrollvariablen ist um 1 *kleiner als die Anzahl der verfügbaren Wertpapiere. Die* **stochastischen Gewinnmöglichkeiten** *um den Zeitpunkt t sind nämlich unabhängig von den Wertpapierpreisen zum Zeit t, da aktuelle Wertpapierpreise die Drift und die Volatilität gleichermaßen beinflussen. Dies war eine wichtige Beobachtung von Robert C. Merton und führt bei der großen Anzahl verfügbarer Wertpapiere zu einer enormen Reduktion der benötigten Zustandsvariablen (die Anzahl der Kontrollvariablen wird zwar nicht reduziert, aber wir werden sehen, dass für die Handhabbarkeit eines Nutzen-Optimierungsproblems die Anzahl der Zustandvariablen kritischer ist).*

Statt dem ,,Anteil des Gesamtvermögens investiert im i-ten Wertpapier" wird oft auch der ,,Eurobetrag investiert im i-ten Wertpapier" als Kontrollvariable gewählt, also

$$\widetilde{\nu}_t^i := \varphi_t^iS_t^i \tag{1.10}$$

(was zweckmäßig ist, hängt von der Nutzenfuktion ab, in die man das kontrollierte Vermögen einsetzen möchte). Entscheidend ist, dass auch bei der Substitution (1.10) die Wertpapierpreise eliminiert und somit als Zustandsvariablen nicht benötigt werden.

Die Lipschitz-Stetigkeit (1.1) erscheint eine natürliche Bedingung zu sein, um Existenz und Eindeutigkeit einer Lösung zu gewährleisten. Betrachte dazu folgende deterministische Beispiele (d.h. $\sigma = 0$)

Beispiel 1.7. *Offenbar besitzt die Differentialgleichung*

$$\frac{dX_t}{dt} = X_t^2, \quad X_0 = 1.$$

auf $[0,1)$ *die (eindeutige) Lösung*

$$X_t = \frac{1}{1-t}, \quad t \in [0,1),$$

die jedoch bei 1 *explodiert. Man beachte, dass die Lipschitz-Bedingung (1.1) nicht erfüllt ist. Für jedes* $M \in (1,\infty)$ *kann man Theorem 1.2 jedoch auf die abgewandelte Differentialgleichung* $dX_t = (X_t^2 \wedge M)\,dt$ *anwenden. Solange die Lösung* X^M *der abgewandelten Differentialgleichung dem Betrag nach kleiner als* $\sqrt{M}$ *bleibt, ist sie auch die eindeutige Lösung der urspründlichen Differentialgleichung. Dies ist für* $t \leq 1 - 1/\sqrt{M}$ *der Fall. Lässt man nun* M *gegen* ∞ *laufen, ergibt dies Existenz und Eindeutigkeit für* $t \in [0,1)$.

Beispiel 1.8. *Die Differentialgleichung*

$$\frac{dX_t}{dt} = 3X_t^{2/3}, \quad X_0 = 0.$$

besitzt mehrere Lösungen und zwar ist für jedes $a \in \mathbb{R}_+ \cup \{+\infty\}$ *der Prozess*

$$X_t := \begin{cases} 0 & : \quad \text{für } t \leq a\,, \\ (t-a)^3 & : \quad \text{für } t > a \end{cases}$$

eine Lösung. Man beachte, dass die Lipschitz-Bedingung (1.1) nicht erfüllt ist.

Lemma 1.9 (Gronwall)**.** *Seien* $f, g : [0,T] \to \mathbb{R}$ *integrierbare Funktionen und* $C \in \mathbb{R}_+$, *s.d.*

$$f(t) \leq g(t) + C \int_0^t f(s)\,ds, \quad \forall t \in [0,T]. \tag{1.11}$$

Dann gilt

$$f(t) \leq g(t) + C \int_0^t e^{C(t-s)} g(s)\,ds, \quad \forall t \in [0,T].$$

Ist g *konstant* $G \in \mathbb{R}$, *dann folgt* $f(t) \leq Ge^{Ct}, \quad \forall t \in [0,T]$.

Beweis. Setze $F(t) := \int_0^t f(s)\,ds$ und $h(t) := F(t)e^{-Ct}$. Dann gilt unter (1.11)

$$h'(t) = f(t)e^{-Ct} - CF(t)e^{-Ct} \leq g(t)e^{-Ct}.$$

$h(0) = 0$ und Integration liefern

$$F(t) = e^{Ct}h(t) = e^{Ct} \int_0^t h'(s)\,ds \leq \int_0^t e^{C(t-s)} g(s)\,ds, \quad \forall t \in [0,T].$$

Eingesetzt in (1.11) ergibt dies

$$f(t) \leq g(t) + CF(t) \leq g(t) + C \int_0^t e^{C(t-s)} g(s)\, ds, \quad \forall t \in [0,T].$$

□

Bemerkung 1.10. *Um die Notationen zu vereinfachen, nehmen wir im folgenden an, dass $n = m = 1$. Der allgemeine Fall gilt natürlich analog.*

Beweis von Satz 1.2. Eindeutigkeit: Seien X und $\widetilde{X}$ Lösungen von (1.3), also nach Definition insbesondere adaptiert und stetig. Es gilt

$$X_t - \widetilde{X}_t = \int_0^t \left(b(s, X_s, \nu_s) - b(s, \widetilde{X}_s, \nu_s) \right) ds + \int_0^t \left(\sigma(s, X_s, \nu_s) - \sigma(s, \widetilde{X}_s, \nu_s) \right) dW_s.$$

Wir können o.B.d.A. annehmen, dass X und $\widetilde{X}$ beschränkt sind. Andernfalls lokalisiert man die lokal beschränkten Prozesse und zeigt Gleichheit zunächst nur auf den stochastischen Intervallen $[\![0, T_n]\!]$ mit $T_n := \inf\{t \geq 0 \mid |X_t| \vee |\widetilde{X}_t| > n\} \wedge T$. Für festes $n \in \mathbb{N}$ müssen b und σ auf $]\!]T_n, T]\!]$ gleich Null gesetzt werden (dies sei durch den Kontrollprozess möglich). Die abgestoppten, dem Betrage nach durch n beschränkten Prozsse X^{T_n} und $\widetilde{X}^{T_n}$ sind dann Lösungen der modifizierten SDE (1.3), die die Bedingungen (1.1) und (1.2) nach wie vor erfüllt. Wenn für alle $n \in \mathbb{N}$ Gleichheit bis auf Ununterscheidbarkeit gilt, müssen auch die gesamten Prozesse bis auf Ununterscheidbarkeit gleich sein.

Mit der Cauchy-Schwarz'schen Ungleichung, der Jensenschen Ungleichung und der Itô-Isometrie folgt

$$\begin{aligned}
&E\left(X_t - \widetilde{X}_t\right)^2 \\
&\leq 2E\left(\int_0^t \left(b(s, X_s, \nu_s) - b(s, \widetilde{X}_s, \nu_s)\right) ds\right)^2 + 2E\left(\int_0^t \left(\sigma(s, X_s, \nu_s) - \sigma(s, \widetilde{X}_s, \nu_s)\right) dW_s\right)^2 \\
&\leq 2TE\left(\int_0^t \left(b(s, X_s, \nu_s) - b(s, \widetilde{X}_s, \nu_s)\right)^2 ds\right) + 2E\left(\int_0^t \left(\sigma(s, X_s, \nu_s) - \sigma(s, \widetilde{X}_s, \nu_s)\right)^2 ds\right) \\
&\leq 2(T+1)K^2 \int_0^t E\left(X_s - \widetilde{X}_s\right)^2 ds,
\end{aligned}$$

wobei in die letzte Ungleichung die Lipschitz-Bedingung (1.1) und Fubini eingeht. Nun wenden wir für $f(t) := E\left(X_t - \widetilde{X}_t\right)^2$, $g = 0$ und $C := 2(T+1)K^2$ das Lemma von Gronwall an. Wegen der angenommenen Beschränktheit von X und $\widetilde{X}$ gilt $f(t) < \infty$ für alle $t \in [0,T]$. Es gilt zudem $f(t) \leq C \int_0^t f(s)\, ds$, was mit dem Lemma $f = 0$ nach sich zieht. Da X und $\widetilde{X}$ als Lösungen von (1.3) stetige Pfade besitzen, folgt Ununterscheidbarkeit.

Existenz: Wir werden eine Lösung durch eine Variation des Picard'schen Iterationsverfahren (für gewöhnliche Differentialgleichungen) konstruieren. Setze $X^{(0)} := x$ und

$$X_t^{(n+1)} := x + \int_0^t b(s, X_s^{(n)}, \nu_s)\, ds + \int_0^t \sigma(s, X_s^{(n)}, \nu_s)\, dW_s, \quad n = 0, 1, 2, \ldots \tag{1.12}$$

(Beispiel: $b(t,x,u) = \alpha x$ und $\sigma(t,x,u) = 0 \rightsquigarrow X_t^{(1)} = x + \alpha x t$, $X_t^{(2)} = x + \int_0^t \alpha(x + \alpha x s)\, ds = x + \alpha x t + \frac{1}{2} x \alpha^2 t^2$ und allgemein $X_t^{(n)} = x \sum_{k=0}^n \frac{(\alpha t)^k}{k!} \to x \exp(\alpha t)$ für $n \to \infty$)

Durch Induktion nach n zeigt man, dass die Integrale auf der rechten Seite von (1.12) wohldefiniert sind. Hierfür braucht man Voraussetzung (1.2), die Lipschitzstetigkeit von b und σ im Argument $X^{(n)}$ und die Tatsache, dass $X^{(n)}$ als stetiger adaptierter Prozess lokal beschränkt ist.

Schritt 1: Man zeige zunächst, dass

$$E\left(\sup_{0 \le s \le t} \left(X_s^{(n)} \right)^2 \right) < \infty \quad \forall n \in \mathbb{N}. \tag{1.13}$$

Man führe wieder eine Induktion nach n durch. Es gilt

$$\begin{aligned}
& E\left(\sup_{0 \le s \le t} \left(X_s^{(n+1)} - x \right)^2 \right) \\
& \le 2E\left(\sup_{0 \le s \le t} \left(\int_0^s b(u, X_u^{(n)}, \nu_u)\, du \right)^2 \right) + 2E\left(\sup_{0 \le s \le t} \left(\int_0^s \sigma(u, X_u^{(n)}, \nu_u)\, dW_u \right)^2 \right) \\
& \le 2tE\left(\int_0^t b^2(s, X_s^{(n)}, \nu_s)\, ds \right) + 8E\left(\int_0^t \sigma^2(s, X_s^{(n)}, \nu_s)\, ds \right) \\
& \le 2(t+4)K^2 \int_0^t E\left(X_s^{(n)} \right)^2\, ds + 2tE\left(\int_0^t b^2(s, 0, \nu_s)\, ds \right) + 8E\left(\int_0^t \sigma^2(s, 0, \nu_s)\, ds \right) \\
& \le 2t(t+4)K^2 E\left(\sup_{0 \le s \le t} \left(X_s^{(n)} \right)^2 \right) + 2tE\left(\int_0^t b^2(s, 0, \nu_s)\, ds \right) + 8E\left(\int_0^t \sigma^2(s, 0, \nu_s)\, ds \right).
\end{aligned}$$

In die zweite Ungleichung gehen die Itô-Isometrie und die Doobschen-Ungleichung für quadratintegrierbare Martingale ein. In die dritte Ungleichung geht die Lipschitz-Stetigkeit von b und σ ein. Die Erwartungswerte $E\left(\int_0^t b^2(s, 0, \nu_s)\, ds \right)$ und $E\left(\int_0^t \sigma^2(s, 0, \nu_s)\, ds \right)$ sind wegen Voraussetzung (1.2) endlich (man beachte, dass es wegen der Lipschitz-Stetigkeit bei dieser Voraussetzung egal ist, ob man sie für ein x oder für alle x macht).

Schritt 2: Wir wollen die Differenz

$$X^{(n+1)} - X^{(n)} = M + B$$

abschätzen, wobei

$$M_t := \int_0^t \left(\sigma(s, X_s^{(n)}, \nu_s) - \sigma(s, X_s^{(n-1)}, \nu_s) \right)\, dW_s$$

$$B_t := \int_0^t \left(b(s, X_s^{(n)}, \nu_s) - b(s, X_s^{(n-1)}, \nu_s) \right) ds.$$

M ist ein lokales Martingal. Wegen

$$E\left(\int_0^t \left(\sigma(s, X_s^{(n)}, \nu_s) - \sigma(s, X_s^{(n-1)}, \nu_s) \right)^2 ds \right) \leq K^2 \int_0^t E\left(\left(X_s^{(n)} - X_s^{(n-1)} \right)^2 ds \right) \overset{\text{Schritt 1}}{<} \infty$$

(mit der Lipschitz-Bedingung (1.1) und Fubini) ist M ein quadratintegrierbares Martingal und mit der Doobschen Ungleichung für quadratintegrierbare Martingale folgt

$$\begin{aligned} E\left(\sup_{s \in [0,t]} M_s^2 \right) &\overset{\text{Doob}}{\leq} 4E\left(M_t^2 \right) \\ &\overset{\text{Itô-Isometrie}}{=} 4E\left(\int_0^t \left(\sigma(s, X_s^{(n)}, \nu_s) - \sigma(s, X_s^{(n-1)}, \nu_s) \right)^2 ds \right) \\ &\leq 4K^2 \int_0^t E\left(\left(X_s^{(n)} - X_s^{(n-1)} \right)^2 ds \right). \end{aligned} \tag{1.14}$$

Mit der Cauchy-Schwarz'schen Ungleichung und der Jensenschen Ungleichung folgt für den Anteil von endlicher Variation

$$B_t^2 \leq t \int_0^t \left(b(s, X_s^{(n)}, \nu_s) - b(s, X_s^{(n-1)}, \nu_s) \right)^2 ds \tag{1.15}$$

und da die rechte Seite von (1.15) monoton steigend ist

$$\begin{aligned} E\left(\sup_{s \in [0,t]} B_s^2 \right) &\leq tE\left(\int_0^t \left(b(s, X_s^{(n)}, \nu_s) - b(s, X_s^{(n-1)}, \nu_s) \right)^2 ds \right) \\ &\leq tK^2 \int_0^t E\left(X_s^{(n)} - X_s^{(n-1)} \right)^2 ds. \end{aligned} \tag{1.16}$$

Setze

$$\Delta^n(t) := E\left(\sup_{s \in [0,t]} (X_s^{(n+1)} - X_s^{(n)})^2 \right), \quad n = 0, 1, 2, \ldots$$

Aus (1.14) und (1.16) folgt

$$\Delta^n(t) \leq C \int_0^t \Delta^{n-1}(s)\, ds, \quad n = 1, 2, 3, \ldots, t \in [0, T] \tag{1.17}$$

für eine geeignete Konstante C. Es folgt aus (1.2)

$$\Delta^0(t) \leq CE\left(\int_0^t b^2(s, x, \nu_s)\, ds + \int_0^t ||\sigma(s, x, \nu_s)||^2\, ds \right) =: \widetilde{C} < \infty.$$

Iteriert man (1.17) und nutzt schließlich die Monotonie von $\Delta^0(u)$ in u aus, so erhält man

$$\Delta^n(t) \le C\int_0^t \Delta^{n-1}(s)\,ds \le C^2\int_0^t\int_0^s \Delta^{n-2}(u)\,du\,ds \le \ldots \le \frac{(Ct)^n}{n!}\Delta^0(t) \tag{1.18}$$

(da $\int_0^t\int_0^{t_n}\int_0^{t_{n-1}}\ldots\int_0^{t_2} dt_1\ldots dt_{n-1}\,dt_n = \frac{t^n}{n!}$). Mit der Markovschen Ungleichung folgt

$$\begin{aligned}\sum_{n=1}^{\infty} P\left(\sup_{s\in[0,t]}|X_s^{(n+1)}-X_s^{(n)}|>2^{-n}\right) &= \sum_{n=1}^{\infty} P\left(\sup_{s\in[0,t]}(X_s^{(n+1)}-X_s^{(n)})^2>4^{-n}\right)\\ &\le \sum_{n=1}^{\infty} 4^n\Delta^n(t)\\ &\le \Delta^0(t)\sum_{n=1}^{\infty}\frac{(4Ct)^n}{n!} \le \widetilde{C}e^{4Ct}<\infty.\end{aligned}$$

Aus dem Lemma von Borel-Cantelli folgt

$$P(\sup_{t\in[0,T]}|X_t^{(n+1)}-X_t^{(n)}|>2^{-n} \text{ für unendlich viele } n\in\mathbb{N})=0$$

Demnach ist für P-fast alle $\omega\in\Omega$ $(X^{(n)}(\omega))_{n\in\mathbb{N}}$ eine Cauchy-Folge bzgl. der Supremumsnorm. Wegen der Vollständigkeit der Supremumsnorm existiert ein Prozess $X\in\mathbb{D}$ mit $P(\sup_{t\in[0,T]}|X_t^{(n)}-X_t|\to 0,\quad n\to\infty)=1$. Da Konvergenz ,,gleichmäßig in Wahrscheinlichkeit" der Integranden, Konvergenz ,,gleichmäßig in Wahrscheinlichkeit" der Integrale impliziert, folgt aus der Lipschitz-Stetigkeit der Funktionen b und σ im zweiten Argument, dass

$$t\mapsto x+\int_0^t b(s,X_s^{(n)},\nu_s)\,ds+\int_0^t\sigma(s,X_s^{(n)},\nu_s)\,dW_s \tag{1.19}$$

für $n\to\infty$ gegen

$$t\mapsto x+\int_0^t b(s,X_s,\nu_s)\,ds+\int_0^t\sigma(s,X_s,\nu_s)\,dW_s$$

konvergiert. Da aber (1.19) der Prozess X^{n+1} ist, der für $n\to\infty$ auch ,,gleichmäßig in Wahrscheinlichkeit" gegen X konvergiert, müssen die beiden Grenzprozesse bis auf Ununterscheidbarkeit übereinstimmen, d.h. X ist starke Lösung der SDE (1.3).

Schritt 3: Bleibt zu zeigen, dass die eindeutige Lösung die Integrierbarkeitseigenschaft (1.5) erfüllt. Hierzu mache man sich zunächst klar, dass für zwei Prozesse Y und Z die Dreiecksungleichung

$$\begin{aligned}&\sqrt{E\left(\sup_{0\le s\le t}(Y_s+Z_s)^2\right)}\\ &\le\sqrt{E\left(\sup_{0\le s\le t}Y_s^2\right)+2E\left(\sup_{0\le s\le t}|Y_s|\sup_{0\le s\le t}|Z_s|\right)+E\left(\sup_{0\le s\le t}Z_s^2\right)}\\ &\le\sqrt{E\left(\sup_{0\le s\le t}Y_s^2\right)}+\sqrt{E\left(\sup_{0\le s\le t}Z_s^2\right)}.\end{aligned} \tag{1.20}$$

gilt (in die zweite Ungleichung gehen die Cauchy-Schwarz'sche Ungleichung und $\left(\sup_{0\le s\le t}|Y_s|\right)^2 = \sup_{0\le s\le t} Y_s^2$, $\left(\sup_{0\le s\le t}|Z_s|\right)^2 = \sup_{0\le s\le t} Z_s^2$ ein). Aus der iterativen Anwendung von (1.20) und (1.18) folgt

$$\begin{aligned}
\sqrt{E\left(\sup_{0\le s\le t}\left(X_s^{(n)}\right)^2\right)} &\le x + \sum_{k=1}^{n} \sqrt{E\left(\sup_{0\le s\le t}\left(X_s^{(k)} - X_s^{(k-1)}\right)^2\right)} \\
&\le x + \sum_{k=1}^{n} \sqrt{\frac{(Ct)^k}{k!}\Delta^0(t)} \\
&\le x + \sum_{k=1}^{\infty} \sqrt{\frac{(Ct)^k}{k!}\Delta^0(t)} < \infty, \quad \forall n \in \mathbb{N}.
\end{aligned}$$

Mit Fatou folgt

$$E\left(\sup_{0\le s\le t}(X_s)^2\right) \le \liminf_{n\to\infty} E\left(\sup_{0\le s\le t}\left(X_s^{(n)}\right)^2\right) < \infty$$

und somit (1.5). □

Beispiel 1.11 (Ornstein-Uhlenbeck Prozess)**.** *Betrachte für $\alpha > 0$ den Mean-Reverting Prozess*

$$X_t = X_0 - \alpha \int_0^t X_s\, ds + W_t. \tag{1.21}$$

Die Lösung von (1.21) ist offenbar gegeben durch

$$X_t = X_0 e^{-\alpha t} + \int_0^t e^{\alpha(s-t)}\, dW_s, \quad \forall t \in [0,T]. \tag{1.22}$$

Man rechnet dies nach durch (beachte, dass $e^{\alpha(s-t)} - 1 = -\alpha \int_s^t e^{\alpha(s-u)}\, du$). Für X aus (1.22) gilt:

$$\begin{aligned}
X_t &= X_0 e^{-\alpha t} + \int_0^t \left(e^{\alpha(s-t)} - 1\right) dW_s + \int_0^t dW_s \\
&= X_0 e^{-\alpha t} - \alpha \int_{\mathbb{R}}\int_{\mathbb{R}} 1_{(0\le s\le u\le t)} e^{\alpha(s-u)}\, du\, dW_s + \int_0^t dW_s \\
&\overset{Fubini}{=} X_0 e^{-\alpha t} - \alpha \int_{\mathbb{R}}\int_{\mathbb{R}} 1_{(0\le s\le u\le t)} e^{\alpha(s-u)}\, dW_s\, du + \int_0^t dW_s \\
&= X_0\left(1 - \alpha\int_0^t e^{-\alpha u}\, du\right) - \alpha\int_0^t \left(\int_0^u e^{\alpha(s-u)}\, dW_s\right) du + \int_0^t dW_s \\
&= X_0\left(1 - \alpha\int_0^t e^{-\alpha u}\, du\right) - \alpha\int_0^t \left(X_u - X_0 e^{-\alpha u}\right) du + \int_0^t dW_s \\
&= X_0 - \alpha\int_0^t X_u\, du + W_t.
\end{aligned}$$

Die dritte Gleichheit folgt aus dem Satz von Fubini für stochastische Integrale. Für die fünfte Gleichheit benutzt man die Definition von X in (1.22) an der Stelle u.

2 Stochastisches Kontrollproblem und Hamilton-Jacobi-Bellmann Gleichung

Seien

$$f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}, \quad (t,x,u) \mapsto f(t,x,u)$$

und

$$g : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto g(x)$$

stetige Funktionen (Stetigkeit ist eigentlich eine unnötig harte Voraussetzung und kann durch Messbarkeit ersetzt werden, wir machen sie hier aber, um die Lücken in den folgenden Beweisen so klein wie möglich zu halten).

Sei $x \in \mathbb{R}^n$. Wir wollen eine zulässige Kontrolle ν finden, so dass ν zusammen mit dem zum Zeitpunkt 0 in x gestarteten und mit ν kontrollierten Prozess X den Ausdruck

$$E\left(\int_0^T f(s, X_s, \nu_s)\, ds + g(X_T)\right). \tag{2.1}$$

maximiert.

Beispiel 2.1. *Betrachte das optimale Konsum- und Investitionsproblem in einem friktionslosen Finanzmarkt mit den Preisprozessen aus Beispiel 1.6. Der Vermögensprozess ist dann der eindimensionale Kontrollprozess und $x \in \mathbb{R}$ das vorgegebene Startkapital der Investorin. ν_t^i, $i = 1, \dots, d$, soll den Anteil des Gesamtvermögens, das zum Zeitpunkt t im i-ten Wertpapier investiert ist, spezifizieren. Zudem konsumiert die Investorin. Der akkumulierte Konsum muss absolutstetig bzgl. der verstrichenen Zeit sein, also $C_t = \int_0^t c_s\, ds$, wobei $(c_t)_{t\in[0,T]}$ die Konsumrate ist. $c := \nu^{d+1}$ ist eine zusätzliche Kontrollvariable. Der Vermögensprozess ist dann formal als Lösung der SDE*

$$dX_t = \sum_{i=0}^d \varphi_t^i\, dS_t^i - c_t\, dt = rX_t\, dt + \sum_{i=1}^d \left[\nu_t^i X_t(\mu^i - r)\, dt + \sum_{j=1}^m \nu_t^i X_t \widetilde{\sigma}^{i,j}\, dW_t^j\right] - c_t\, dt$$

mit $X_0 = x$ definiert. Zudem setzen wir

$$f(t,x,\nu) = \exp(-\delta t)(\nu^{d+1})^\gamma \quad \textit{und} \quad g(x) = \exp(-\delta T)x^\gamma.$$

$u(y) = y^\gamma$, $\gamma \in (0,1)$ ist die Powernutzenfunktion und δ die Zeitpräferenz der Investorin. δ fällt oft mit dem risikolosen Zins r zusammen, was aber nicht zwingend ist.

Um das Optimierungsproblem (2.1) lösen zu können, führen wir nun zunächst eine dynamische Version von (2.1) ein. Für festes $t \in [0,T]$ können wir $X = X^{t,x,\nu}$ auch als Lösung der SDE

$$dX_s = b(s, X_s, \nu_s)\, ds + \sigma(s, X_s, \nu_s)\, dW_s, \quad s \in [t,T],\ X_t = x.$$

definieren, d.h. der Prozess startet erst bei t. Die "eingeschränkte" SDE besitzt mit Satz 1.2 auch eine eindeutige starke Lösung. Der Prozess $X = X^{t,x,\nu}$ hängt also formal von seinem Startzeitpunkt, -wert und der Kontrolle ab.

Die **Gewinnfunktion** $\mathcal{J}$ (u.a. in Abhängigkeit von der gewählten Kontrolle) ist definiert als

$$\mathcal{J}(t,x,\nu) := E\left(\int_t^T f(s, X_s^{t,x,\nu}, \nu_s)\,ds + g(X_T^{t,x,\nu}) \mid \mathcal{F}_t\right).$$

Man beachte, dass $\mathcal{J}(t,x,\nu)$ auch gegeben den Wert von x in t i.A. eine $\mathcal{F}_t$-messbare Zufallsvariable ist und keine Konstante (wieso ?)

Die **Wertfunktion** V ist definiert als

$$\begin{aligned} V(t,x) &:= \operatorname{ess\,sup}_{\nu\in\mathcal{U}} \mathcal{J}(t,x,\nu) \\ &= \operatorname{ess\,sup}_{\nu\in\mathcal{U}} E\left(\int_t^T f(s, X_s^{t,x,\nu}, \nu_s)\,ds + g(X_T^{t,x,\nu}) \mid \mathcal{F}_t\right). \end{aligned} \tag{2.2}$$

Das essentielle Supremum in (2.2) wird offenbar über eine maximumsstabile Menge von Zufallsvariablen gebildet. Seien $\nu^{(1)}, \nu^{(2)} \in \mathcal{U}$. Setze

$$\nu^{(3)} := \nu^{(1)} 1_{(t,T]\times A} + \nu^{(2)} 1_{(t,T]\times(\Omega\setminus A)},$$

wobei

$$\begin{aligned} A := \Big\{ \quad & E\left(\int_t^T f(s, X_s^{t,x,\nu^{(1)}}, \nu_s^{(1)})\,ds + g(X_T^{t,x,\nu^{(1)}}) \mid \mathcal{F}_t\right) \\ & \geq E\left(\int_t^T f(s, X_s^{t,x,\nu^{(2)}}, \nu_s^{(2)})\,ds + g(X_T^{t,x,\nu^{(2)}}) \mid \mathcal{F}_t\right) \Big\} \in \mathcal{F}_t. \end{aligned}$$

Es gilt

$$\begin{aligned} & E\left(\int_t^T f(s, X_s^{t,x,\nu^{(3)}}, \nu_s^{(3)})\,ds + g(X_T^{t,x,\nu^{(3)}}) \mid \mathcal{F}_t\right) \\ &= 1_A E\left(\int_t^T f(s, X_s^{t,x,\nu^{(1)}}, \nu_s^{(1)})\,ds + g(X_T^{t,x,\nu^{(1)}}) \mid \mathcal{F}_t\right) \\ &\quad + 1_{\Omega\setminus A} E\left(\int_t^T f(s, X_s^{t,x,\nu^{(2)}}, \nu_s^{(2)})\,ds + g(X_T^{t,x,\nu^{(2)}}) \mid \mathcal{F}_t\right) \\ &= E\left(\int_t^T f(s, X_s^{t,x,\nu^{(1)}}, \nu_s^{(1)})\,ds + g(X_T^{t,x,\nu^{(1)}}) \mid \mathcal{F}_t\right) \\ &\quad \vee E\left(\int_t^T f(s, X_s^{t,x,\nu^{(2)}}, \nu_s^{(2)})\,ds + g(X_T^{t,x,\nu^{(2)}}) \mid \mathcal{F}_t\right). \end{aligned} \tag{2.3}$$

Bemerkung 2.2. *Unter den gemachten Voraussetzungen existieren Versionen der bedingten Erwartungswerte und essentiellen Suprema, so dass die Ausnahmenullmengen* $\{\mathcal{J}(t, x, \nu) > V(t, x)\}$ *zwar von* ν *nicht aber von* t *und* x *abhängen. Beweisidee: Konstruierte die Größen zunächst nur für alle rationalen* t, x*. Die abzählbare Vereinigung der Ausnahemnullmengen bleibt eine Ausnahmenullmenge (abhängig von der Kontrolle* ν*). Wegen Stetigkeit von* $X^{t,x,\nu}$ *in* (t, x) *und von* f *und* g *kann nun alles auf die Versionen für rationale* t, x *zurückgeführt werden.*

Wir setzen voraus, dass $f, g \geq 0$ und $V(0, x) < \infty$ für alle $x \in \mathbb{R}^n$.

Bemerkung 2.3. *Im Prinzip ist man nur an der Zahl* $V(0, x)$ *für ein festes* $x \in \mathbb{R}^n$ *interessiert – also an dem Wert des Optimierungsproblems zum Startzeitpunkt mit Startwert* $X_0 = x$*. Zur Bestimmung dieses Wertes erweist es sich allerdings als nützlich, Aussagen über den ,,Zwischenwert" des Optimierungsproblems zum Zeitpunkt* $t \in (0, T)$ *unter der Bedingung* $X_t = x'$ *zu machen (wobei* x' *alle möglichen Werte durchläuft, die* X_t *annehmen kann). Dazu dient die etwas allgemeinere Definition der Wertfunktion (2.2).* $V(t, x)$ *kann interpretiert werden als die erwartete Auszahlung (wenn die laufenden Auszahlungen erst bei* t *beginnen) unter der Bedingung* $X_t = x$ *und unter der Annahme, dass man sich zwischen* t *und* T *optimal verhält. Die Kontrolle vor* t *spielt offenbar keine Rolle für* $V(t, x)$ *(sie beeinflusst aber i.A. den Zustand* x*, in dem man zum Zeitpunkt* t *den Prozess* X *neu startet)*

Theorem 2.4 (Dynamisches Programmierprinzip)**.** *Sei* $t \in [0, T]$ *und* θ *eine* $[t, T]$*-wertige Stoppzeit. Es gilt*

$$V(t, x) = \text{ess sup}_{\nu \in \mathcal{U}} E\left(\int_t^\theta f(s, X_s^{t,x,\nu}, \nu_s)\, ds + V(\theta, X_\theta^{t,x,\nu}) \mid \mathcal{F}_t\right). \tag{2.4}$$

Zudem existieren deterministische Versionen von $V(t, x)$ *(d.h. für die optimale Kontrolle hängt die Gewinnfunktion von der Information* $\mathcal{F}_t$ *nur über den Wert des Prozesses* X *in* t *ab).*

Beweisskizze. Sei $\nu \in \mathcal{U}$ und θ eine $[t, T]$-wertige Stoppzeit. Es lassen sich Versionen von Lösungsfamilien $(X^{t,x,\nu})_{t \in [0,T],\ x \in \mathbb{R}}$ finden, so dass außerhalb einer ,,globalen" P-Nullmenge gilt

$$X_s^{t,x,\nu} = X_s^{\theta, X_\theta^{t,x,\nu}, \nu}, \quad \forall s \geq \theta \tag{2.5}$$

(die P-Nullmenge kann von ν abhängen, nicht aber von s, t, x und θ). Anschaulich erscheint (2.5) klar: die Zuwächse des Prozesses $X^{t,x,\nu}$ um die Zeitpunkte s mit $s \geq \theta$ hängen nur von ν_s, s, dW_s und dem aktuellen Wert $X_s^{t,x,\nu}$ des Prozesses ab. Also kann man auch statt $X^{t,x,\nu}$ eine erst bei θ gestartete SDE nehmen, wobei man unter Nutzung der Information $\mathcal{F}_\theta$ die SDE mit dem Startwert $\widetilde{x} = X_\theta^{t,x,\nu}$ wählt, d.h. in die Lösungsfamilie $X^{\widetilde{t},\widetilde{x},\nu}$ für $(\widetilde{t}, \widetilde{x})$ den zufälligen Wert $(\theta, X_\theta^{t,x,\nu})$ einsetzt (streng mathematisch werden wir (2.5) hier aber nicht beweisen).

Bezeichne mit $\widetilde{V}(t,x)$ die rechte Seite von (2.4). Für alle $\nu \in \mathcal{U}$ gilt

$$\begin{aligned}
\mathcal{J}(t,x,\nu) &= E\left(\int_t^\theta f(s,X_s^{t,x,\nu},\nu_s)\,ds \mid \mathcal{F}_t\right) \\
&\quad + E\left(\int_\theta^T f(s,X_s^{t,x,\nu},\nu_s)\,ds + g(X_T^{t,x,\nu}) \mid \mathcal{F}_t\right) \\
&\overset{(2.5)}{=} E\left(\int_t^\theta f(s,X_s^{t,x,\nu},\nu_s)\,ds \mid \mathcal{F}_t\right) \\
&\quad + E\left(\int_\theta^T f(s,X_s^{\theta,X_\theta^{t,x,\nu},\nu},\nu_s)\,ds + g(X_T^{\theta,X_\theta^{t,x,\nu},\nu}) \mid \mathcal{F}_t\right) \\
&= E\left(\int_t^\theta f(s,X_s^{t,x,\nu},\nu_s)\,ds \mid \mathcal{F}_t\right) \\
&\quad + E\left(\underbrace{E\left(\int_\theta^T f(s,X_s^{\theta,X_\theta^{t,x,\nu},\nu},\nu_s)\,ds + g(X_T^{\theta,X_\theta^{t,x,\nu},\nu}) \mid \mathcal{F}_\theta\right)}_{=\mathcal{J}(\theta,X_\theta^{t,x,\nu},\nu)} \mid \mathcal{F}_t\right) \\
&= E\left(\int_t^\theta f(s,X_s^{t,x,\nu},\nu_s)\,ds \mid \mathcal{F}_t\right) + E\left(\mathcal{J}(\theta,X_\theta^{t,x,\nu},\nu) \mid \mathcal{F}_t\right). \qquad (2.6)
\end{aligned}$$

Da für alle (s,y,ν) gilt $V(s,y) \geq \mathcal{J}(s,y,\nu)$ folgt aus (2.6), dass

$$\mathcal{J}(t,x,\nu) \leq E\left(\int_t^\theta f(s,X_s^{t,x,\nu},\nu_s)\,ds \mid \mathcal{F}_t\right) + E\left(V(\theta,X_\theta^{t,x,\nu},\nu) \mid \mathcal{F}_t\right) \qquad (2.7)$$

und damit

$$P(V(t,x) \leq \widetilde{V}(t,x)) = 1. \qquad (2.8)$$

Die umgekehrte Richtung werden wir nur skizzieren: Sei $(s_m,y_m)_{m\in\mathbb{N}}$ eine dichte Folge in $[t,T]\times\mathbb{R}^n$. Wegen der in (2.3) gezeigten Maximumsstabilität folgt die Existens einer (von m abhängigen) Folge $(\widetilde{\nu}^k)_{k\in\mathbb{N}}$ mit

$$\mathcal{J}(s_m,y_m,\widetilde{\nu}^k) \uparrow V(s_m,y_m), \quad P-\text{f.s.}, \quad k\uparrow\infty.$$

Somit existiert zu jedem (s_m,y_m) und jedem $\varepsilon_m > 0$ eine ε_m-optimale („fast" optimale) Strategie ν^m im folgenden Sinne:

$$P\left(\mathcal{J}(s_m,y_m,\nu^m) \geq V(s_m,y_m) - \varepsilon_m\right) \geq 1-\varepsilon_m$$

(Da $\mathcal{J}(s_m,y_m,\widetilde{\nu}^k)$ und $V(s_m,y_m)$ Zufallsvariablen sind, kommt $\mathcal{J}(s_m,y_m,\widetilde{\nu}^k)$ i.A. nicht gleichmäßig in ω an $V(s_m,y_m)$ ran). Da man ε_m so wählen kann, dass auch $\sum_{m\in\mathbb{N}}\varepsilon_m$ klein ist, lassen sich zu jedem $\varepsilon > 0$ Strategien finden, so dass

$$P\left(\mathcal{J}(s_m,y_m,\nu^m) \geq V(s_m,y_m) - \varepsilon \quad \forall m\in\mathbb{N}\right) \geq 1-\varepsilon. \qquad (2.9)$$

Neben $\varepsilon > 0$ werde nun auch $\mu \in \mathcal{U}$ fixiert. Wegen der Lipschitz-Stetigkeit der Koeffizienten b und σ ist die Lösung der SDE stetig im Startwert (s, y) bzgl. der Norm $\sqrt{E(\sup_{u\in[t,T]} ||X_u||^2)}$ (gleichmäßig in der Kontrolle ν). Da zudem f und g als stetig vorausgesetzt wurden, kann gezeigt werden, dass die Wertfunktion $V(s, y)$ stetig in s und y ist. Damit kann aus den Strategien in (2.9) eine zulässige Kontrolle ν^ε „zusammengeklebt" werden, die

$$P\left(\mathcal{J}(\theta, X_\theta^{t,x,\mu}, \nu^\varepsilon) \geq V(\theta, X_\theta^{t,x,\mu}) - 2\varepsilon\right) \geq 1 - 2\varepsilon \tag{2.10}$$

erfüllt. ν^ε kann von der Form

$$\nu^\varepsilon = \mu 1_{[\![0,\theta]\!]} + \nu^m 1_{]\!]\theta,T]\!]} \quad \text{mit } m = \arg\min\left\{\widetilde{m} = 1, \ldots, m_0 \mid |\theta - s_{\widetilde{m}}| + |X_\theta^{t,x,\mu} - y_{\widetilde{m}}|\right\}$$

gewählt werden, d.h. ν^ε stimmt vor θ mit der Kontrolle μ überein. Diese Wahl ist möglich, da nur die Werte von ν^ε auf dem Intervall $]\!]\theta, T]\!]$ in $\mathcal{J}(\theta, X_\theta^{t,x,\mu}, \nu^\varepsilon)$ eingehen. Da andererseits ν^ε mit μ auf dem stochastischen Intervall $]\!]t, \theta]\!]$ mit μ übereinstimmt, folgt aus (2.6)

$$\begin{aligned}
E(V(t,x)) &\geq E(\mathcal{J}(t, x, \nu^\varepsilon)) \\
&\overset{(2.6)}{=} E\left(\int_t^\theta f(s, X_s^{t,x,\nu^\varepsilon}, \nu_s^\varepsilon)\, ds\right) + E\left(\mathcal{J}(\theta, X_\theta^{t,x,\nu^\varepsilon}, \nu^\varepsilon)\right) \\
&\overset{\nu^\varepsilon=\mu \text{ auf }]\!]t,\theta]\!]}{=} E\left(\int_t^\theta f(s, X_s^{t,x,\mu}, \mu_s)\, ds\right) + E\left(\mathcal{J}(\theta, X_\theta^{t,x,\mu}, \nu^\varepsilon)\right) \\
&\overset{(2.10) \text{ und } \mathcal{J}\geq 0}{\geq} E\left(\int_t^\theta f(s, X_s^{t,x,\mu}, \mu_s)\, ds\right) + E\left(V(\theta, X_\theta^{t,x,\mu})\right) - h(\varepsilon),
\end{aligned}$$

wobei

$$h(\varepsilon) := 2\varepsilon + \sup_{A\in\mathcal{F},\ P(A)\leq 2\varepsilon} E\left(1_A V(\theta, X_\theta^{t,x,\mu})\right).$$

Zunächst lässt man zu festem $\mu \in \mathcal{U}$ ε gegen 0 laufen. Integrierbarkeit von $V(\theta, X_\theta^{t,x,\mu})$ vorausgesetzt, folgt $h(\varepsilon) \to 0$ für $\varepsilon \to 0$* und damit

$$E(V(t,x)) \geq E\left(\int_t^\theta f(s, X_s^{t,x,\mu}, \mu_s)\, ds\right) + E\left(V(\theta, X_\theta^{t,x,\mu})\right).$$

Nun kann man auf der rechten Seite das Supremum über alle $\mu \in \mathcal{U}$ bilden und schließen

$$E(V(t,x)) \geq \sup_{\mu\in\mathcal{U}} \left(E\left(\int_t^\theta f(s, X_s^{t,x,\mu}, \mu_s)\, ds\right) + E\left(V(\theta, X_\theta^{t,x,\mu})\right)\right) = E(\widetilde{V}(t,x))$$

und mit (2.8) $P(V(t,x) = \widetilde{V}(t,x)) = 1$. $\square$

*Für eine integrierbare Zufallsvariable Y gilt:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall A \in \mathcal{F} \quad P(A) \leq \delta \implies E(Y 1_A) \leq \varepsilon$$

(siehe z.B. Satz 7.37 in [4]).

Bemerkung 2.5. *Nicht für alle Optimierungsprobleme gibt es ein dynamisches Programmierprinzip. Betrachte z.B.*

$$V(t,x) := \sup_{\nu \in \mathcal{U}} \left(E\left(X_T^{t,x,\nu}\right) - \lambda \mathrm{Varianz}(X_T^{t,x,\nu}) \right). \tag{2.11}$$

für ein $\lambda > 0$ (auch hier würden Erwartungswert und Varianz des **optimal** *kontrollierten Prozesses durch die Information $\mathcal{F}_t$ nicht verändert, d.h. $E(\dots \mid \mathcal{F}_t) = E(\dots)$ und* $\mathrm{Varianz}(\dots \mid \mathcal{F}_t) = \mathrm{Varianz}(\dots)$ *und wir können direkt ohne Bedingen auf die Information $\mathcal{F}_t$ mit dem gewöhnlichen Supremum arbeiten*). (2.11) könnte z.B. ein Portfoliooptimierungsproblem für einen Erwartungswert/Varianz-Optimierer in einem Markt mit einer risikobehafteten Aktie mit Preisprozess $(S_t)_{t\in[0,T]}$ und risikolosem Zins $r = 0$ sein, also*

$$\sup_{\nu \in \mathcal{U}} \left(E\left(x + \int_0^T \nu_s \, dS_s\right) - \lambda \mathrm{Varianz}\left(x + \int_0^T \nu_s \, dS_s\right)\right). \tag{2.12}$$

Das Problem (2.11) ist **zeitinkonsistent** *im folgenden Sinne: Nimmt man die für den Zeitpunkt 0 und den Startwert x optimale dynamische Kontrolle $\widehat{\nu} = (\widehat{\nu}_s)_{s\in[0,T]}$ und betrachtet zum späteren Zeitpunkt $t > 0$ das erneuerte Optimierungsproblem mit zufälligem Startwert $X_t^{0,x,\widehat{\nu}}$, dann stellt man fest, dass die Kontrolle $\widehat{\nu}$ (eingeschränkt auf das Intervall $[t,T]$) i.A. nicht mehr optimal für das erneuerte Problem ist.*

Besitzt die Aktie eine nicht-verschwindene Drift, dann besteht in (2.12) nämlich die Möglichkeit, Gewinne oder Verluste aus der Vergangenheit "tendenziell" auszugleichen, was für den Optimierer interessant ist, da es die Gesamtvarianz reduziert. Bei einem neu gestarteten Problem entfällt dieser Anreiz. Beispiel ?

Grundlage für das Dynamische Programmierprinzip ist, dass die Zielfunktion das sog. **Unabhängigkeitsaxiom** *erfüllt. Sei $\succ$ eine Präferenzordnung auf der Menge der eindimensionalen Wahrscheinlichkeitsverteilungen†. $\succ$ erfüllt das Unabhängigkeitsaxiom (***independence axiom***), wenn für alle Verteilungsfunktionen F^1, F^2, F^3 und $\lambda \in (0,1]$ die*

---

*Zur Erinnerung: $\mathrm{Varianz}(Y \mid \mathcal{G}) := E((Y - E(Y \mid \mathcal{G}))^2 \mid \mathcal{G})$. Es gilt die Zerlegung

$$\mathrm{Varianz}(Y) = E\left(\mathrm{Varianz}(Y \mid \mathcal{G})\right) + \mathrm{Varianz}(E(Y \mid \mathcal{G})).$$

†$\succ$ ist also eine binäre Relation auf der Menge der Wahrscheinlichkeitsverteilungen, d.h. sie ordnet jedem Paar den Wert „wahr" oder „falsch" zu. Die Interpretation von „$F^1 \succ F^2$" ist, dass der Agent eine zufällige Auszahlung mit Verteilung F^1 einer zufälligen Auszahlung mit Verteilung F^2 (im strikten Sinne) vorziehen würde. Eine Präferenzordnung $\succ$ muss die folgenden Eigenschaften erfüllen:

- Asymmetrie: Für alle F^1, F^2 gilt

$$F^1 \succ F^2 \quad \Longrightarrow \quad F^2 \not\succ F^1.$$

- Negative Transitivität: Für alle F^1, F^2, F^3 gilt

$$F^1 \succ F^2 \quad \Longrightarrow \quad (F^1 \succ F^3 \quad \text{oder} \quad F^3 \succ F^2).$$

Implikation

$$F^1 \succ F^2 \implies \lambda F^1 + (1-\lambda)F^3 \succ \lambda F^2 + (1-\lambda)F^3 \tag{2.13}$$

gilt.

Für $A \in \mathcal{F}_1$ kann dann, gegeben eine Kontrolle vor 1, die optimale Kontrolle nach 1 *separat auf den Mengen A und $\Omega \setminus A$ gefunden werden. Die Teilprobleme sind lediglich über die Kontrolle vor* 1 *miteinander gekoppelt, da diese beide Teilprobleme beinflusst (vor* 1 *weiß man noch nicht, ob A oder $\Omega \setminus A$ eintritt). Dazu werden F^1 und F^2 aus (2.13) auf bedingte Verteilungen der Endauszahlung gegeben, dass das Ereignis A eintritt, angewandt und F^3 auf bedingte Verteilungen gegeben $\Omega \setminus A$. Das Unabhängigkeitsaxiom besagt, dass eine Verbesserung der bedingten Verteilung gegeben A stets eine Verbesserung der absoluten Verteilung nach sich zieht – unabhängig davon was die bedingte Verteilung gegeben $\Omega \setminus A$ ist.*

Beim Mittelwert-Varianzproblem ist (2.13) offenbar nicht erfüllt, da sich die Varianz als mittlere quadratische Abweichung vom Erwartungswert bestimmt, der über den Gesamtraum gebildet wird. Dazu betrachte man für die Präferenzordnung

$$F^1 \succ F^2 \quad :\Leftrightarrow \quad \text{Varianz}(X_1) < \text{Varianz}(X_2), \quad \textit{wobei } X_1 \sim F^1,\ X_2 \sim F^2$$

das Beispiel: $F^1(x) = 1_{(x \geq 1)}$, $F^2(x) = \frac{1}{2}1_{(x\geq 0)} + \frac{1}{2}1_{(x \geq 1)}$ und $F^3(x) = 1_{(x\geq 0)}$ (die erste und dritte ZV ist also deterministisch mit Erwartungswert 1 *bzw.* 0 *und die zweite ist Bernoulli-verteilt mit Parameter* 1/2*). Es gilt also $F^1 \succ F^2$. Wähle nun $\lambda = 1/2$. $\frac{1}{2}F^1 + \frac{1}{2}F^3$ ist Bernoulli-verteilt mit Parameter* 1/2 *und $\frac{1}{2}F^2 + \frac{1}{2}F^3$ ist Bernoulli-verteilt mit Paramter* 1/4*. Da letztere Verteilung die kleinere Varianz hat (Varianz der Bernoulli-Verteilung $= p(1-p)$), gilt $\frac{1}{2}F^2 + \frac{1}{2}F^3 \succ \frac{1}{2}F^1 + \frac{1}{2}F^3$ und damit (2.13) offenbar nicht.*

2.1 Die Hamilton-Jacobi-Bellmann (HJB) Gleichung

Sei $H : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n,n} \to \mathbb{R}$

$$H(t,x,p,A) := \sup_{u \in U} \left\{ b(t,x,u)^\top p + \frac{1}{2}\text{Spur}[\sigma(t,x,u)\sigma(t,x,u)^\top A] + f(t,x,u) \right\} \tag{2.14}$$

Definiere den Operator zweiter Ordnung $\mathcal{L}^u$

$$\mathcal{L}^u \varphi(t,x) := b(t,x,u)^\top D\varphi(t,x) + \frac{1}{2}\text{Spur}[\sigma(t,x,u)\sigma(t,x,u)^\top D^2\varphi(t,x)],$$

wobei D und D^2 den Gradienten bzw. die Hesse-Matrix der Abbildung $x \mapsto \varphi(t,x)$ bezeichnet.

Die Wertfunktion $V(t,x)$ ist in unserem Modellrahmen immer deterministisch (hängt also nur von x und nicht noch zusätzlich von ω ab). Setzt man zusätzlich eine gewisse Glattheit der Funktion $V : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ voraus (die aber nicht in jeder Anwendung gegeben sein muss !), dann löst V eine gewisse PDE. Dies soll nun in zwei Schritten gezeigt werden.

Proposition 2.6. *Nehme an, die Wertfunktion V aus (2.2) sei in $C^{1,2}([0,T),\mathbb{R}^n)$ und $f(\cdot,\cdot,u)$ sei für alle $u \in U \subset \mathbb{R}^k$ stetig. Dann gilt für alle $(t,x) \in [0,T) \times \mathbb{R}^n$*

$$\frac{\partial V}{\partial t}(t,x) + H(t,x,DV(t,x),D^2V(t,x)) \leq 0$$

Beweis. Seien $(t,x) \in [0,T) \times \mathbb{R}^n$ und $u \in U$ fest. Betrachte den mit der konstanten Kontrolle $\nu = u$ bei t in x gestarteten Prozess $X = X^{t,x,u}$. Für $h > 0$ definiere die Stoppzeit

$$\theta_h := \inf\{s > t \mid s - t > h \text{ oder } ||X_s - x|| > 1\}.$$

Für jedes $\omega \in \Omega$ gibt es natürlich ein $\widetilde{h}(\omega) > 0$, s.d. $\theta_h = t + h$ für $h \leq \widetilde{h}(\omega)$. Aus dem dynamischen Programmierprinzip (2.4) folgt

$$\begin{aligned}
0 \quad &\leq \quad V(t,x) - E\left(\int_t^{\theta_h} f(s,X_s^{t,x,u},u)\,ds + V(\theta_h, X_{\theta_h}^{t,x,u})\right) \\
&\overset{\text{Itô}}{=} \quad -E\left(\int_t^{\theta_h}\left(\frac{\partial V}{\partial t}(s,X_s) + (\mathcal{L}^u V)(s,X_s,u) + f(s,X_s,u)\right)\,ds\right) \\
&\qquad -E\left(\int_t^{\theta_h} DV(s,X_s)^\top \sigma(s,X_s,u)\,dW_s\right).
\end{aligned}$$

Da DV auf dem stochastischen Intervall $]\!]t,\theta_h]\!]$ beschränkt ist, gilt

$$E\left(\int_t^{\theta_h} DV(s,X_s)^\top \sigma(s,X_s,u)\,dW_s\right) = 0$$

und wir erhalten

$$E\left(\frac{1}{h}\int_t^{\theta_h}\left(\frac{\partial V}{\partial t}(s,X_s) + (\mathcal{L}^u V)(s,X_s) + f(s,X_s,u)\right)\,ds\right) \leq 0.$$

Aus $V \in C^{1,2}([0,T),\mathbb{R}^n)$ und der Stetigkeit von X folgt zunächst, dass obige Zufallsvariable für $h \to 0$ punktweise gegen

$$\frac{\partial V}{\partial t}(t,x) + (\mathcal{L}^u V)(t,x) + f(t,x,u)$$

konvergiert. Da die Zufallsvariablen gleichmäßig in h beschränkt sind, folgt Konvergenz der Erwartungswerte und damit

$$\frac{\partial V}{\partial t}(t,x) + (\mathcal{L}^u V)(t,x) + f(t,x,u) \leq 0.$$

□

Proposition 2.7. *Nehme an, die Wertfunktion erfüllt $V \in C^{1,2}([0,T),\mathbb{R}^n)$ und die Funktion H sei stetig. Dann gilt auch die „Umkehrung" von Proposition 2.6 also*

$$\frac{\partial V}{\partial t}(t,x) + H(t,x,DV(t,x),D^2V(t,x)) \geq 0 \tag{2.15}$$

Beweis. Sei $(t_0, x_0) \in [0, T) \times \mathbb{R}^n$. Nehme an, dass

$$\frac{\partial V}{\partial t}(t_0, x_0) + H(t_0, x_0, DV(t_0, x_0), D^2V(t_0, x_0)) < 0. \tag{2.16}$$

Dies wollen wir zu einem Widerspruch führen. Insbesondere wollen wir zeigen, dass unter der Annahme (2.16) das Supremum auf der rechten Seite von (2.4) für eine geeignete Stoppzeit θ strikt kleiner als $V(t, x)$ wäre, also $\widetilde{V}(t_0, x_0) < V(t_0, x_0)$.

Ein Problem ist, dass wir noch nicht wissen, dass das Supremum $V(t_0, x_0)$ durch eine dynamische Kontrolle ν angenommen wird. Wir wissen nur, dass es beliebig gut approximiert wird, so dass wir uns etwas Luft verschaffen müssen.

Definiere dazu die $C^{1,2}([0, T), \mathbb{R}^n)$-Funktion

$$\varphi(t, x) := V(t, x) + \sum_{i=1}^{n} |x^i - x_0^i|^3 + (t - t_0)^2.$$

Der Korrekturterm ist so gewählt, dass er im Punkt (t_0, x_0) die Funktion und alle in (2.15) vorkommenden Ableitungen nicht verändert, d.h.

$$(\varphi - V)(t_0, x_0) = 0, \quad (D\varphi - DV)(t_0, x_0) = 0, \quad \frac{\partial(\varphi - V)}{\partial t}(t_0, x_0) = 0 \tag{2.17}$$

und

$$(D^2\varphi - D^2V)(t_0, x_0) = 0. \tag{2.18}$$

Wegen (2.17)/(2.18) und der (geforderten) Stetigkeit von H existiert ein $\eta > 0$ mit $t_0 + \sqrt{\eta} < T$ und

$$\begin{aligned} h(t, x) &:= \frac{\partial \varphi}{\partial t}(t, x) + H(t, x, D\varphi(t, x), D^2\varphi(t, x)) \\ &< 0, \quad \forall t \geq t_0, x \in \mathbb{R}^n \text{ mit } (t - t_0)^2 + \sum_{i=1}^{n} |x^i - x_0^i|^3 \leq \eta. \end{aligned} \tag{2.19}$$

Sei nun für $\varepsilon := \eta/2 > 0$ $\nu \in \mathcal{U}$ eine ε-optimale Strategie in dem Sinne, dass

$$E\left(\mathcal{J}(t_0, x_0, \nu)\right) \geq V(t_0, x_0) - \varepsilon.$$

Zu $X = X^{t_0, x_0, \nu}$ definiere die Stoppzeit

$$\theta := \inf\{s > t_0 \mid (s, X_s) \notin \mathcal{N}_\eta\}$$

mit

$$\mathcal{N}_\eta := \{(t, x) \mid (t - t_0)^2 + \sum_{i=1}^{n} |x^i - x_0^i|^3 \leq \eta\}.$$

θ nimmt Werte im Intervall $(t_0, t_0+\sqrt{\eta}]$ an. Wegen der Stetigkeit von X gilt $P((\theta, X_\theta) \in \partial\mathcal{N}) = 1$ und damit

$$(\varphi - V)(\theta, X_\theta) = \eta = 2\varepsilon. \tag{2.20}$$

Es gilt

$$\begin{aligned}
&\int_{t_0}^{\theta} f(s, X_s, \nu_s)\, ds + V(\theta, X_\theta) - V(t_0, x_0) \\
&\overset{(2.20)}{=} \int_{t_0}^{\theta} f(s, X_s, \nu_s)\, ds + \varphi(\theta, X_\theta) - \varphi(t_0, x_0) - 2\varepsilon \\
&\overset{\text{Itô}}{=} \int_{t_0}^{\theta} \frac{\partial\varphi}{\partial t}(s, X_s)\, ds + \int_{t_0}^{\theta} \frac{\partial\varphi}{\partial x}(s, X_s) b(s, X_s, \nu_s)\, ds + \frac{1}{2}\int_{t_0}^{\theta} \frac{\partial^2\varphi}{\partial x^2}(s, X_s)\sigma^2(s, X_s)\, ds \\
&+ \int_{t_0}^{\theta} f(s, X_s, \nu_s)\, ds + \int_{t_0}^{\theta} \frac{\partial\varphi}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s - 2\varepsilon \\
&\leq \int_{t_0}^{\theta} \frac{\partial\varphi}{\partial t}(s, X_s) + H(s, X_s, D\varphi(s, X_s), D^2\varphi(s, X_s))\, ds \\
&+ \int_{t_0}^{\theta} \frac{\partial\varphi}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s - 2\varepsilon \\
&\overset{(2.19)}{\leq} \int_{t_0}^{\theta} \frac{\partial\varphi}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s - 2\varepsilon.
\end{aligned} \tag{2.21}$$

Die vorletzte Ungleichung folgt aus dem Supremum über $u \in \mathcal{U}$ in der Definition von H. Die letzte Ungleichung folgt aus (2.19). Da $\frac{\partial\varphi}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)$ auf $]\!]t_0, \theta]\!]$ beschränkt ist, fällt bei der Erwartungswertbildung der dW_t-Term weg und es folgt

$$\begin{aligned}
E\left(\mathcal{J}(t_0, x_0, \nu)\right) \geq V(t_0, x_0) - \varepsilon \quad &\overset{(2.21)}{\geq} \quad E\left(\int_{t_0}^{\theta} f(s, X_s, \nu_s)\, ds + V(\theta, X_\theta)\right) + \varepsilon \\
&\overset{(2.7)}{\geq} \quad E\left(\mathcal{J}(t_0, x_0, \nu)\right) + \varepsilon,
\end{aligned}$$

was natürlich ein Widerspruch ist. □

Wir haben damit gezeigt:

Theorem 2.8. *Unter den Voraussetzungen der Propositionen 2.6 und 2.7 erfüllt die Wertfunktion die* **Hamilton-Jacobi-Bellmann (HJB-) Gleichung**

$$\frac{\partial V}{\partial t}(t, x) + H(t, x, DV(t, x), D^2V(t, x)) = 0, \quad \forall t \in (0, T),\ x \in \mathbb{R}^n.$$

2.2 Verifikationssatz

Um die Optimalität einer Strategie zu beweisen, braucht man natürlich gerade die Umkehrung von Satz 2.8: jede Lösung der HJB-Gleichung ist die Wertfunktion. Diese Aussage nennt man Verifikationssatz. Hat man erstmal die Wertfunktion gefunden, lässt sich die optimale Strategie einfach aus (2.14) gewinnen, indem man für jedes Paar (t, x) das k-dimensionale Maximierungsproblem

$$\sup_{u \in U} \left\{ b(t,x,u)^\top DV(t,x) + \frac{1}{2}\mathrm{Spur}[\sigma(t,x,u)\sigma(t,x,u)^\top D^2V(t,x)] + f(t,x,u) \right\}$$

löst.

Theorem 2.9. *Sei $v \in C^{1,2}([0,T), \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$. Erfüllen v und f quadratische Wachstumsbedingungen, d.h.*

$$|v(t,x)| + |f(t,x,u)| \leq C(1 + |x|^2), \quad \forall (t,x,u) \in [0,T] \times \mathbb{R}^n \times U.$$

Nehme an $v(T,x) = g(x)$, $\forall x \in \mathbb{R}^n$ und es existiert für alle $(t,x) \in [0,T) \times \mathbb{R}^n$ ein Maximierer $\widehat{u}(t,x)$ der Funktion $u \mapsto \mathcal{L}^u v(t,x) + f(t,x,u)$, so dass

$$\begin{aligned} 0 &= \frac{\partial v}{\partial t}(t,x) + H(t,x,Dv(t,x),D^2v(t,x)) \\ &= \frac{\partial v}{\partial t}(t,x) + \mathcal{L}^{\widehat{u}(t,x)}v(t,x) + f(t,x,\widehat{u}(t,x)) \end{aligned}$$

und die SDE

$$dX_s = b(s, X_s, \widehat{u}(s, X_s))\, ds + \sigma(s, X_s, \widehat{u}(s, X_s))\, dW_s, \quad X_t = x. \tag{2.22}$$

zu jedem Startzeitpunkt $t \in [0,T]$ und Startwert $x \in \mathbb{R}^n$ eine eindeutige starke Lösung $X = X^{t,x}$ besitzt[‡]*. Nehme zudem an, dass $\widehat{\nu}$ mit $\widehat{\nu}_s := \widehat{u}(s, X_s)$ eine zulässige Kontrolle definiert. Dann gilt $v = V$ und $\widehat{\nu}$ ist eine optimale Kontrolle.*

Beweis. Seien $(t,x) \in [0,T) \times \mathbb{R}^n$ und $\nu \in \mathcal{U}$ ein beliebiger Kontrollprozess. Zu dem zugehörigen Prozess $X = X^{t,x,\nu}$ definiere die Stoppzeiten

$$\theta_m := \inf\left\{ s > t \mid \int_t^s \left(\frac{\partial v}{\partial x}(u, X_u)\sigma(u, X_u, \nu_u) \right)^2 du = m \right\} \wedge T, \quad m \in \mathbb{N}.$$

Da $v \in C^{1,2}([0,T), \mathbb{R}^n)$ ist $\frac{\partial v}{\partial x}$ auf kompakten Mengen der Gestalt $[0, T-\varepsilon] \times [-K,K]^n$ beschränkt. Damit ist sichergestellt, dass $P(\theta_m \to T, \quad m \to \infty) = 1$. **Achtung**: Es ist nicht klar, ob $(\theta_m)_{m \in \in \mathbb{N}}$ lokalisierend ist, also ob $\int_t^T \left(\frac{\partial v}{\partial x}(u, X_u)\sigma(u, X_u, \nu_u)\right)^2 du$ eine P-f.s. endliche Zufallsvariable ist, da $\frac{\partial v}{\partial x}$ bei T explodieren könnte. Es folgt (siehe [6])

$$E\left(\int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s \right)^2 = E\left(\int_t^{\theta_m} \left(\frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s) \right)^2 ds \right) \leq m < \infty$$

[‡]Formal kann man (2.22) als eine SDE auffassen, die nicht kontrolliert wird, d.h. bei der die Koeffizienten in (1.3) nur von t und x abhängen. Die Koeffizienten sind durch die Abbildungen $(t,x) \mapsto b(t,x,\widehat{u}(t,x))$ und $(t,x) \mapsto \sigma(t,x,\widehat{u}(t,x))$ gegeben.

und

$$E\left(\int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s \mid \mathcal{F}_t\right) = 0. \tag{2.23}$$

Mit der Itô-Formel und

$$\frac{\partial v}{\partial t} + \mathcal{L}^u v + f(\cdot, \cdot, u) \leq \frac{\partial v}{\partial t} + H(\cdot, \cdot, Dv, D^2 v) = 0 \quad \forall u \in U \tag{2.24}$$

gilt

$$\begin{aligned}
&v(t,x) \\
&\overset{\text{Itô}}{=} v(\theta_m, X_{\theta_m}) - \int_t^{\theta_m} \frac{\partial v}{\partial t}(s, X_s)\, ds - \int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s) b(s, X_s, \nu_s)\, ds \\
&\quad -\frac{1}{2}\int_t^{\theta_m} \frac{\partial^2 v}{\partial x^2}(s, X_s)\sigma^2(s, X_s, \nu_s)\, ds - \int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s \\
&= v(\theta_m, X_{\theta_m}) - \int_t^{\theta_m} \left(\frac{\partial v}{\partial t} + \mathcal{L}^{\nu_s} v\right)(s, X_s)\, ds - \int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s \\
&\geq v(\theta_m, X_{\theta_m}) - \int_t^{\theta_m} \left(\frac{\partial v}{\partial t} + H(\cdot, \cdot, Dv, D^2 v)\right)(s, X_s)\, ds \\
&+ \int_t^{\theta_m} f(s, X_s, \nu_s)\, ds - \int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s \\
&= v(\theta_m, X_{\theta_m}) + \int_t^{\theta_m} f(s, X_s, \nu_s)\, ds - \int_t^{\theta_m} \frac{\partial v}{\partial x}(s, X_s)\sigma(s, X_s, \nu_s)\, dW_s.
\end{aligned}$$

Zusammen mit (2.23) folgt

$$v(t,x) \geq E\left(v(\theta_m, X_{\theta_m}) + \int_t^{\theta_m} f(s, X_s, \nu_s)\, ds \mid \mathcal{F}_t\right), \quad \forall m \in \mathbb{N}. \tag{2.25}$$

Nun wollen wir zum Limes $m \to \infty$ übergehen. Wegen der Stetigkeit von X gilt $(\theta_m, X_{\theta_m}) \to (T, X_T)$ punktweise. Zusammen mit der Stetigkeit von v auf ganz $[0,T] \times \mathbb{R}^n$ folgt, dass die Zufallsvariable auf der rechten Seite von (2.25) punktweise gegen

$$v(T, X_T) + \int_t^T f(s, X_s, \nu_s)\, ds = g(X_T) + \int_t^T f(s, X_s, \nu_s)\, ds$$

konvergiert. Zudem dürfen unter den gemachten Voraussetzungen bedingte Erwartungswertbildung und Limesbildung vertauscht werden, da

$$\begin{aligned}
&\left| v(\theta_m, X_{\theta_m}) + \int_t^{\theta_m} f(s, X_s, \nu_s)\, ds \right| \\
&\leq C\left(1 + X_{\theta_m}^2 + T + \int_t^T |X_s^2|\, ds\right) \\
&\leq C\,(1+T)\left(1 + \sup_{s \in [t,T]} |X_s^2|\right) \in L^1(\Omega, \mathcal{F}, P)
\end{aligned}$$

wegen (1.5). Somit gilt

$$v(t,x) \geq E\left(g(X_T) + \int_t^T f(s, X_s, \nu_s)\, ds \mid \mathcal{F}_t\right), \quad \forall \nu \in \mathcal{U}.$$

Wir müssen noch zeigen, dass es eine zulässige Kontrolle gibt, für die Gleichheit gilt. Unter den gemachten Voraussetzungen funktioniert dies völlig analog, indem wir ausnutzen, dass die Strategie $\widehat{\nu}$ (2.24) mit Gleichheit erfüllt. □

Bemerkung 2.10. *Der Beweis von Theorem 2.9 baut nicht auf Theorem 2.4 auf. In vergleichbaren Situationen muss das dynamische Programmierprinzip, das in seiner größeren Allgemeinheit (keine Glattheitannahme an die Wertfunktion) oft schwieriger zu zeigen ist, also nicht vorab bewiesen werden.*

Als Nebenprodukt von Theorem 2.9 erhält man, dass die optimale Kontrolle Markov ist, also nur von der Zeit und den aktuellen Werten der Zustandsvariablen abhängt.

3 Anwendungen

3.1 Portfoliooptimierung

Betrachte den Black-Scholes Markt, der aus zwei Wertpapieren, einem risikolosen Bond und einer risikobehafteten Aktie, besteht.

$$S_t^0 = \exp(rt) \quad \text{und} \quad S_t^1 = \exp\left(\mu t + \sigma W_t - \frac{1}{2}\sigma^2 t\right)$$

Der vorhersehbare Prozess $\pi = (\pi_t)_{t\in[0,T]}$ soll hier den **Anteil des Vermögens, den der Investor in die Aktie investiert** bezeichnen. Der verbleibende Anteil $1 - \pi_t$ des Vermögens wird dann in den risikolosen Bond investiert (man beachte, dass der Wertebereich von π nicht auf $[0, 1]$ beschränkt ist). Zudem kann der Investor während der Laufzeit $[0, T]$ konsumieren. Der vorhersehbare stochastische Prozess $(c_t)_{t\in[0,T]}$ bezeichne die Lebesgue-Rate mit der der Investor konsumieren möchte. Unter der Strategie (π, c) besitzt der Vermögensprozess $X = X^{t,x,(\pi,c)}$ die Dynamik

$$\begin{aligned}
dX_t &= \pi_t X_t \frac{dS_t^1}{S_t^1} + (1-\pi_t) X_t \frac{dS_t^0}{S_t^0} - c_t\, dt \\
&= \pi_t X_t (\mu\, dt + \sigma\, dW_t) + (1-\pi_t) X_t r\, dt - c_t\, dt \\
&= [X_t(r + (\mu - r)\pi_t) - c_t]\, dt + \sigma X_t \pi_t\, dW_t
\end{aligned}$$

Wir betrachten folgendes Optimierungsproblem

$$V(t,x) = \sup_{\pi,c} E\left(\int_t^T e^{-\delta s} U_1(c_s)\, ds + e^{-\delta T} U_2(X_T)\right)$$

wobei U_1, U_2 Nutzenfunktionen sind. Der Paramter δ modelliert die Zeitpäferenz der Investorin. Sie kann sich vom riskolosen Zins r unterscheiden. π soll die Bedingung

$$E\left[\int_0^T |\pi_t|^2\, dt\right] < \infty.$$

erfüllen.

Man sieht, dass dieses Problem in den Rahmen des stochastisches Kontrollproblems aus Kapitel 2 fällt, wobei der Kontrollprozess (π, c) offenbar zweidimensional ist.

Die HJB Gleichung zu diesem Problem lautet

$$\begin{aligned} 0 &= \frac{\partial v(t,x)}{\partial t} \\ &+ \sup_{\pi\in\mathbb{R},\ c\in\mathbb{R}_+} \left\{ [x(r+(\mu-r)\pi)-c]\frac{\partial v(t,x)}{\partial x} + \frac{1}{2}\sigma^2 x^2 \pi^2 \frac{\partial^2 v(t,x)}{\partial x^2} + e^{-\delta t}U_1(c)\right\} \end{aligned} \tag{3.1}$$

und

$$v(T,x) = e^{-\delta T}U_2(x).$$

Bemerkung 3.1. *In unserer Formulierung des Problems ist $V(t,x)$ in Nutzeneinheiten ,,zum Zeitpunkt 0" zu verstehen (der optimale erwartete Nutzen wird also durch* $\exp(-\delta t)$ *auf den Zeitpunkt* 0 *diskontiert). Alternativ kann man ein ähnliches Problem formulieren, bei dem $V(t,x)$ als Nutzeneinheiten zum Zeitpunkt t zu verstehen ist (Übung).*

Schritt 1: Das optimale Aktieninvestment und die optimale Konsumrate zum Zeitpunkt t lassen sich nun ,,gegeben die noch unbekannte Wertfunktion v" bestimmen. Unter der Bedingung, dass $x \mapsto v(t,x)$ konkav ist, lauten die hinreichenden "first order conditions" für das Optimum in (3.1)

$$\widehat{\pi} = -\frac{\mu - r}{\sigma^2}\frac{\frac{\partial v(t,x)}{\partial x}}{x\frac{\partial^2 v(t,x)}{\partial x^2}} \tag{3.2}$$

und

$$\widehat{c} = (U_1')^{-1}\left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right). \tag{3.3}$$

Schritt 2: Setzt man (3.2) und (3.3) in (3.1) ein, so ergibt dies die PDE

$$\begin{aligned} 0 &= \frac{\partial v(t,x)}{\partial t} + \left[rx - \frac{(\mu-r)^2}{\sigma^2}\frac{\frac{\partial v(t,x)}{\partial x}}{\frac{\partial^2 v(t,x)}{\partial x^2}} - (U_1')^{-1}\left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right)\right]\frac{\partial v(t,x)}{\partial x} \\ &+\frac{1}{2}x^2\frac{(\mu-r)^2}{\sigma^2}\left(\frac{\frac{\partial v(t,x)}{\partial x}}{x\frac{\partial v^2(t,x)}{\partial x^2}}\right)^2\frac{\partial^2 v(t,x)}{\partial x^2} + e^{-\delta t}U_1\left((U_1')^{-1}\left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right)\right) \qquad (3.4) \\ &= \frac{\partial v(t,x)}{\partial t} + \left[rx - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{\frac{\partial v(t,x)}{\partial x}}{\frac{\partial^2 v(t,x)}{\partial x^2}} - (U_1')^{-1}\left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right)\right]\frac{\partial v(t,x)}{\partial x} \\ &+e^{-\delta t}U_1\left((U_1')^{-1}\left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right)\right). \qquad (3.5) \end{aligned}$$

Diese PDE muss nun gelöst werden.

Bemerkung 3.2. *Die Konkavität der Wertfunktion in x musste für die Resultate in Kapitel 2 (insbesondere für das dynamische Programmierprinzip) nicht vorausgesetzt werden. Meistens kann man aber nur mit ihr zeigen, dass das Supremum in der HJB-Gleichung endlich ist und angenommen wird.*

Beispiel 3.3. *Um die Lösung explizit bestimmen zu können, betrachten wir den Spezialfall, dass $U_1(x) = U_2(x) = \frac{1}{\gamma}x^\gamma$, $\gamma \in (0,1)$. In diesem Fall vereinfacht sich die PDE (3.4) zu*

$$\begin{aligned}
&\frac{\partial v(t,x)}{\partial t} + \left(rx - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{\frac{\partial v(t,x)}{\partial x}}{\frac{\partial^2 v(t,x)}{\partial x^2}} - \left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right)^{\frac{1}{\gamma-1}} \right)\frac{\partial v(t,x)}{\partial x} \\
&+e^{-\delta t}\frac{1}{\gamma}\left(e^{\delta t}\right)^{\frac{\gamma}{\gamma-1}}\left(\frac{\partial v(t,x)}{\partial x}\right)^{\frac{\gamma}{\gamma-1}} \\
&= \frac{\partial v(t,x)}{\partial t} + \left(rx - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{\frac{\partial v(t,x)}{\partial x}}{\frac{\partial^2 v(t,x)}{\partial x^2}} \right)\frac{\partial v(t,x)}{\partial x} \\
&+e^{\delta t/(\gamma-1)}\frac{1-\gamma}{\gamma}\left(\frac{\partial v(t,x)}{\partial x}\right)^{\frac{\gamma}{\gamma-1}} \\
&\overset{!}{=} 0
\end{aligned} \tag{3.6}$$

mit Endbedingung $v(T,x) = e^{-\delta T}\frac{x^\gamma}{\gamma}$. Bei der Gestalt der Nutzenfunktionen und der Dynamik des Aktienkurses liegt es nahe, es mit dem Ansatz

$$v(t,x) = e^{-\delta t} f(t)\frac{x^\gamma}{\gamma}$$

zu versuchen. Setzt man dies mit den entsprechenden Ableitungen in (3.6) ein, so erhält man die **gewöhnliche** *Differentialgleichung*

$$e^{-\delta t}\frac{x^\gamma}{\gamma}\left(f'(t) - f(t)\left(\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{\gamma}{\gamma-1} - r\gamma + \delta\right) - (\gamma-1)(f(t))^{\frac{\gamma}{\gamma-1}} \right) = 0 \tag{3.7}$$

mit Endbedingung $f(T) = 1$. (3.7) lässt sich schreiben als

$$f'(t) = a_1 f(t) + a_2 (f(t))^{\frac{\gamma}{\gamma-1}}, \quad f(T) = 1,$$

wobei $a_1 = \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{\gamma}{\gamma-1} - r\gamma + \delta$ und $a_2 = \gamma - 1$. Setze $h(t) := (f(t))^{\frac{1}{1-\gamma}}$. Die ODE wird zu der linearen ODE

$$\begin{aligned}
h'(t) &= f'(t)\frac{1}{1-\gamma}(f(t))^{\frac{\gamma}{1-\gamma}} \\
&= \left(a_1 f(t) + a_2 (f(t))^{\frac{\gamma}{\gamma-1}}\right)\frac{1}{1-\gamma}(f(t))^{\frac{\gamma}{1-\gamma}} \\
&= a_1\frac{1}{1-\gamma}f(t)^{\frac{1}{1-\gamma}} + a_2\frac{1}{1-\gamma} \\
&= a_1\frac{1}{1-\gamma}h(t) + a_2\frac{1}{1-\gamma} \\
&= a_1\frac{1}{1-\gamma}h(t) - 1
\end{aligned} \tag{3.8}$$

mit der Endbedingung $h(T) = 1$. (3.8) besitzt die eindeutige Lösung

$$h(t) = \exp\left(-\frac{a_1}{1-\gamma}(T-t)\right)\left(1 - \frac{1-\gamma}{a_1}\right) + \frac{1-\gamma}{a_1}.$$

Setzt man dies wieder in v ein, so erhält man

$$v(t,x) = e^{-\delta t}\left[\exp\left(-\frac{a_1}{1-\gamma}(T-t)\right)\left(1 - \frac{1-\gamma}{a_1}\right) + \frac{1-\gamma}{a_1}\right]^{1-\gamma}\frac{x^\gamma}{\gamma}.$$

Setzt man die Wertfunktion in die Ausdrücke für die optimale Strategie in (3.2) und (3.3) ein, so erhält man

$$\widehat{\pi}_t = -\frac{\mu - r}{\sigma^2}\frac{\frac{\partial v(t,x)}{\partial x}}{x\frac{\partial^2 v(t,x)}{\partial x^2}} = \frac{\mu - r}{\sigma^2}\frac{1}{1-\gamma}$$

und

$$\widehat{c}_t = (U_1')^{-1}\left(e^{\delta t}\frac{\partial v(t,x)}{\partial x}\right) = (f(t))^{\frac{1}{\gamma-1}}X_t.$$

Für $\mu \neq 0$, $\delta = r = 0$ ist $a_1 < 0$, die Funktion $f(t)$ somit monoton fallend und damit ist $(f(t))^{\frac{1}{\gamma-1}}$ monoton steigend. Der Konsum ist also proportional zum aktuellen Vermögen und wächst ceteris paribus mit der Zeit. Dies liegt daran, dass bei einer langen Restlaufzeit als günstig empfundene Gewinnmöglichkeiten bestehen, die bei sofortigem Konsum des Vermögens nicht realisert werden könnten.

Der Vermögensprozess $X = X^{t,x,(\widehat{\pi},\widehat{c})}$ zu der optimalen Strategie erfüllt die SDE

$$dX_t = X_t\left[r + \frac{(\mu - r)^2}{\sigma^2}\frac{1}{1-\gamma} - (f(t))^{\frac{1}{\gamma-1}}\right]dt + X_t\sigma\frac{\mu - r}{\sigma^2}\frac{1}{1-\gamma}dW_t$$

Die Koeffizienten unterscheiden sich von der Standard-Geometrischen Brownschen Bewegung nur um zeitabhängige aber deterministische Faktoren. Mit dem Verifikationssatz Theorem 2.9 folgt, dass $(\widehat{\pi}, \widehat{c})$ eine optimale Kontrolle ist.

4 Ein Steilkurs in Lévy-Prozessen

Definition 4.1 (Lévy-Prozess). *Ein adaptierter Prozess $X = (X_t)_{t\in\mathbb{R}_+}$ mit càdlàg Pfaden, Werten in $\mathbb{R}^d$ und $X_0 = 0$ ist ein* **Lévy-Prozess***, wenn er folgende Eigenschaften erfüllt*

(i) Unabhängige Zuwächse: für alle $n \in \mathbb{N}$, $0 \leq t_1 \leq \ldots \leq t_n$ sind die Zufallsvariablen $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ stochastisch unabhängig.

(ii) Stationäre Zuwächse: für alle $0 \leq s \leq t$ haben die Zufallsvariablen $X_t - X_s$ und X_{t-s} die gleiche Verteilung.

Bemerkung 4.2. *Man nennt einen Prozess, der die Bedingungen aus Definition 4.1 erfüllt auch einen* **intrinsischen** *Lévy-Prozess. Allgemeiner kann man einen Lévy-Prozess bzgl. eine Filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ definieren. (ii) wird dann durch die Forderung ersetzt, dass für alle $0 \leq t_1 \leq t_2$ die Zufallsvariable $X_{t_2} - X_{t_1}$ stochastisch unabhängig von der σ-Algebra $\mathcal{F}_{t_1}$ ist. Für die von X erzeugte Filtration $(\mathcal{F}_t^X)_{t\in\mathbb{R}_+}$ sind die Definitionen dann äquivalent.*

Bemerkung 4.3. *Wenn X càdlàg Pfade besitzt und Bedingung (ii) erfüllt, kann man daraus*

(iii) Stochastische Stetigkeit:

$$\lim_{h\to 0} P\left(||X_{t+h} - X_t|| > \varepsilon\right) = 0, \quad \forall \varepsilon > 0, \quad \forall t \in \mathbb{R}_+$$

folgern. Offensichtlich folgt aus der Rechtsstetigkeit von X Eigenschaft (iii) mit der Einschränkung $h \to 0$ und $h > 0$. Die Konvergenz gilt dann aber auch für $h < 0$, da $X_{t+h} - X_t$ und $X_t - X_{t-h}$ mit (ii) die gleiche Verteilung haben.

(iii) bedeutet, dass der Prozess zu einem **vorgegebenen** *Zeitpunkt $t \in \mathbb{R}_+$ nur mit Wahrscheinlichkeit 0 springt. Da aber $\mathbb{R}_+$ überabzählbar ist, bedeutet dies* **nicht***, dass der Prozess während der gesamten Zeit nur mit Wahrscheinlichkeit 0 springt.*

Bemerkung 4.4. *Jeder Lévy-Prozess X ist bzgl. seiner natürlichen Filtration $(\mathcal{F}_t^X)_{t\in\mathbb{R}_+}$ ein Semimartingal. Dies folgt aus der Lévy-Itô Zerlegung (Theorem 4.20). Ohne Benutzung des Theorems können wir die Aussage bereits plausible machen, wenn wir zusätzlich $E(|X_t|) < \infty$ fordern (was aber nicht jeder Lévy-Prozess erfüllt). Es gilt dann $E(X_t) = tE(X_1)$ (wieso ?) und der Prozess $t \mapsto X_t - E(X_t)$ ist wegen (i) ein $(\mathcal{F}_t^X)_{t\in\mathbb{R}_+}$-Martingal.*

Definition 4.5 (Unendliche Teilbarkeit). *Eine Verteilungsfunktion F auf dem $\mathbb{R}^d$ heißt* **unendlich teilbar***, wenn es zu jedem $n \in \mathbb{N}$ unabhängig, identisch verteilte Zufallsvariablen $Y_1, \ldots, Y_n$ gibt, so dass die Summe $Y_1 + \ldots + Y_n$ die Verteilungsfunktion F besitzt.*

Theorem 4.6. *Sei $X = (X_t)_{t\in\mathbb{R}_+}$ ein Lévy-Prozess. Dann besitzt für jedes $t \in \mathbb{R}_+$ der Zufallsvektor X_t eine unendlich teilbare Verteilungsfunktion. Umgekehrt existiert zu jeder unendlich teilbaren Verteilungsfunktion F ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ auf dem ein Lévy-Prozess $X = (X_t)_{t\in\mathbb{R}_+}$ definiert ist, so dass X_1 wie F verteilt ist.*

Proof. „⇒" Sei X ein Lévy-Prozess, $t > 0$ und $n \in \mathbb{N}$. Offenbar gilt

$$X_t = \sum_{k=1}^{n} \left(X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}} \right)$$

und die Zufallsvariablen $X_{\frac{t}{n}}, X_{\frac{2t}{n}} - X_{\frac{1t}{n}}, \dots, X_t - X_{\frac{(n-1)t}{n}}$ sind i.i.d. Also ist die Verteilung von X_t unendlich teilbar.

„⇐" Siehe z.B. Theorem 7.10 in Sato [10] (Den Beweis werden wir hier nicht führen, aber es wird später klar werden, wie X konstruiert werden kann). □

Bemerkung 4.7. *Außer der Einpunkt-Verteilung ist* **keine** *Wahrscheinlichkeitsverteilung, die nur Werte aus einer kompakten Menge annimmt, unendlich teilbar.*

Wenn eine d-dimensionale Verteilungsfunktion unendlich teilbar ist, müssen auch die Randverteilungen unendlich teilbar sein. Daher reicht es, den eindimensionalen Fall zu beweisen. Nehme also an, dass $P(Y \in [a,b]) = 1$ und $Y = Y_1 + \dots, Y_n$, wobei Y_i i.i.d. Es gilt $P(Y_1 > b/n)^n = P(Y_1 > b/n, \dots, Y_n > b/n) \leq P(Y > b) = 0$ und $P(Y_1 < a/n)^n = P(Y_1 < a/n, \dots, Y_n < a/n) \leq P(Y < a) = 0$. Es folgt $P(Y_1 \in [a/n, b/n]) = 1$ und damit $\mathrm{Var}(Y_1) \leq \left(\frac{b-a}{n}\right)^2$. Für die Varianz von Y bedeutet dies $\mathrm{Var}(Y) = n\mathrm{Var}(Y_1) \leq \frac{(b-a)^2}{n}$. Da diese Überlegung für jedes $n \in \mathbb{N}$ gilt, folgt $\mathrm{Var}(Y) = 0$.

Alle auf kompakten Zeitintervallen beschränkten Lévy-Prozesse sind also lineare Funktionen in der Zeit, d.h. $X_t = \gamma t$.

Proposition 4.8 (Charakteristische Funktion eines Lévy-Prozesses)**.** *Sei $X = (X_t)_{t \in [0,T]}$ ein $\mathbb{R}^d$-wertiger Lévy-Prozess. Dann existiert eine stetige Funktion $\psi : \mathbb{R}^d \to \mathbb{C}$, die wir* **charakteristischen Exponenten** *des Prozesses X nennen wollen, so dass*

$$E(\exp(iu^\top X_t)) = e^{t\psi(u)}, \quad \forall u \in \mathbb{R}^d,\ t \in \mathbb{R}_+. \tag{4.1}$$

Die Verteilung von X_t, $t \in \mathbb{R}_+$, ist durch die Verteilung von X_1 eindeutig bestimmt.

Beweisskizze. Definiere für jedes $t \in \mathbb{R}_+$ die charakteristische Funktion des Zufallsvektors X_t

$$\Psi_{X_t}(u) := E(\exp(iu^\top X_t)), \quad \forall u \in \mathbb{R}^d.$$

Sei $s \leq t$. Da $X_t = X_s + X_t - X_s$ und die Zufallsvektoren X_s und $X_t - X_s$ stochastisch unabhängig sind und letzterer die gleiche Verteilung wie X_{t-s} besitzt, gilt

$$\begin{aligned}
\Psi_{X_t}(u) &= E[\exp(iu^\top X_t)] \\
&= E[\exp(iu^\top X_s + iu^\top (X_t - X_s))] \\
&= E[\exp(iu^\top X_s)] E[\exp(iu^\top (X_t - X_s))] \\
&= E[\exp(iu^\top X_s)] E[\exp(iu^\top X_{t-s})] \\
&= \Psi_{X_s}(u) \Psi_{X_{t-s}}(u), \quad \forall u \in \mathbb{R}^d.
\end{aligned} \tag{4.2}$$

Aus (4.2) folgt, dass für alle $q \in \mathbb{N}$ gilt

$$\Psi_{X_1}(u) = \left(\Psi_{X_{1/q}}(u)\right)^q, \quad \forall u \in \mathbb{R}^d. \tag{4.3}$$

Man kann zeigen, dass für charakteristische Funktionen von unendlich teilbaren Verteilungsfunktionen gilt $\Psi_{X_1}(u) \neq 0$ für alle $u \in \mathbb{R}$ (siehe z.B. Lemma 7.5 in Sato [10]). Es existiert eine eindeutige stetige Funktion $f : \mathbb{C} \to \mathbb{C}$ mit $f(0) = 0$ und $\exp(f(u)) = \Psi_{X_1}(u)$ für alle $u \in \mathbb{R}$. f wird der ,,ausgezeichnete Logarithmus" von Ψ_{X_1} genannt, d.h. $\ln(\Psi_{X_1}) := f$. Des weiteren existiert eine eindeutige stetige Funktion $g : \mathbb{C} \to \mathbb{C}$ mit $g(0) = 1$ und $g(u)^q = \Psi_{X_1}(u)$ für alle $u \in \mathbb{R}$ (die q-te Faltungswurzel von Ψ_{X_1} ist also als stetige Funktion eindeutig definiert). Zwischen f und g besteht der Zusammenhang $g(u) = \exp(f(u)/q)$ (siehe Lemma 7.6 in [10] für diese Aussagen). Zusammen mit (4.3) und der Stetigkeit von $u \mapsto \Psi_{X_{1/q}}(u)$ folgt

$$\Psi_{X_{1/q}}(u) = \exp\left(\frac{\ln\left(\Psi_{X_1}(u)\right)}{q}\right), \quad \forall u \in \mathbb{R}^d. \tag{4.4}$$

Wiederum aus (4.2) folgt, dass für alle $p \in \mathbb{N}$ gilt

$$\Psi_{X_{p/q}}(u) = \left(\Psi_{X_{1/q}}(u)\right)^p, \quad \forall u \in \mathbb{R}^d.$$

Wir gehen analog zu oben vor. Statt der q-ten Wurzel aus Ψ_{X_1} ziehen wir die p-te Wurzel aus $\Psi_{X_{p/q}}$ und erhalten

$$\Psi_{X_{1/q}}(u) = \exp\left(\frac{\ln\left(\Psi_{X_{p/q}}(u)\right)}{p}\right), \quad \forall u \in \mathbb{R}^d. \tag{4.5}$$

Aus (4.4) und (4.5) folgt für alle $p, q \in \mathbb{N}$

$$\exp\left(\frac{\ln\left(\Psi_{X_{p/q}}(u)\right)}{p}\right) = \exp\left(\frac{\ln\left(\Psi_{X_1}(u)\right)}{q}\right), \quad \forall u \in \mathbb{R}^d.$$

Bildet man von beiden Seiten die p-te Potenz ergibt dies

$$\exp\left(\ln\left(\Psi_{X_{p/q}}(u)\right)\right) = \exp\left(\frac{p}{q}\ln\left(\Psi_{X_1}(u)\right)\right), \quad \forall u \in \mathbb{R}^d, \tag{4.6}$$

wobei die linke Seite mit $\Psi_{X_{p/q}}(u)$ übereinstimmt.

Aus Eigenschaft (iii) in Bemerkung 4.3 und majorisierter Konvergenz folgt, dass für jedes u die Abbildung $t \mapsto \Psi_{X_t}(u)$ stetig ist. Damit überträgt sich (4.6) auch auf die irrationalen t und es folgt

$$\Psi_{X_t}(u) = e^{t \ln(\Psi_{X_1}(u))}, \quad \forall t \in \mathbb{R}_+,\ u \in \mathbb{R}^d.$$

Setze $\psi(u) := \ln(\Psi_{X_1}(u))$. Da sich ψ aus Ψ_{X_1} eindeutig bestimmt, ist die Verteilung von X_t durch die Verteilung von X_1 eindeutig festgelegt. □

Theorem 4.9 (Lévy-Khintchine Formel)**.** *Sei F eine eindimensionale Wahrscheinlichkeitsverteilung und*

$$\psi(u) := \ln(E(\exp(iuY))), \quad u \in \mathbb{R}$$

der charakteristische Exponent von F (wobei $Y \sim F$)[§]. F ist genau dann unendlich teilbar, wenn es ein Tripel (γ, σ, Π) gibt, wobei $\gamma \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$ und Π ein Maß auf $\mathbb{R} \setminus \{0\}$ mit $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$, so dass gilt

$$\psi(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux1_{\{|x|\leq 1\}}\right) \Pi(dx), \quad \forall u \in \mathbb{R}. \tag{4.7}$$

Zudem ist das Tripel durch die Wahrscheinlichkeitsverteilung eindeutig gegeben.

Definition 4.10. *Π aus (4.7) wird Lévy-Maß und (γ, σ, Π) Lévy-Tripel der unendlich teilbaren Verteilungsfunktion F genannt.*

Bemerkung 4.11. *Der Term $iux1_{\{|x|\leq 1\}}$ in (4.7) sichert die Integrierbar bzgl. des Maßes Π. Dies sieht man an einer Taylorentwicklung der Funktion $x \mapsto e^{iux}$ im Nullpunkt und der geforderten Π-Integrierbarkeit von $x \mapsto 1 \wedge x^2$. Man beachte jedoch, dass, wenn Π ein endliches Maß ist, der Term $\int_{\mathbb{R}} iux1_{\{|x|\leq 1\}}\Pi(dx)$ mit dem Term $i\gamma u$ zusammengefasst werden kann und damit in der Darstellung verschwindet.*

Auf einen vollständigen Beweis von Theorem 4.9 verzichten wir. Das folgende Beispiel macht jedoch eine Richtung des Theorems (aus der Darstellung (4.7) folgt die unendliche Teilbarkeit) für wesentliche Spezialfälle deutlich. Die umgekehrte Richtung (unendliche Teilbarkeit impliziert Darstellung (4.7)) folgt aus der nicht bewiesen Richtung von Theorem 4.6 und Theorem 4.20 (Lévy-Itô Zerlegung), das wir ohne Benutzung von Theorem 4.9 beweisen werden (kommt später).

Beispiele 4.12. *(i)* **Normalverteilung**. *Sei $Y \sim \mathcal{N}(\mu, \sigma^2)$. Es gilt*

$$E\left(\exp(iuY)\right) = \exp\left(i\mu u - \frac{1}{2}\sigma^2 u^2\right), \quad \forall u \in \mathbb{R}.$$

Die n-te Wurzel beträgt

$$\exp\left(i\frac{\mu}{n}u - \frac{1}{2}\left(\frac{\sigma}{\sqrt{n}}\right)^2 u^2\right),$$

was die Verteilungsfunktion der Normalverteilung mit Erwartungswert $\frac{\mu}{n}$ und Varianz $\frac{\sigma^2}{n}$ ist.

($\sigma = 0$ liefert natürlich die Einpunkt-Verteilung)

[§] $\ln(\ldots)$ ist wiederum als der ,,ausgezeichnete Logarithmus" der stetigen komplexwertigen Funktion $u \mapsto E(\exp(iuY))$ zu verstehen.

(ii) **Zusammengesetzte Poisson-Verteilung**. *Sei N eine Poisson-verteilte Zufallsvariable mit Parameter $\lambda > 0$ und $(\xi_k)_{k\in\mathbb{N}}$ eine i.i.d. Folge von Zufallsvariablen mit Verteilungsfunktion F, die unabhängig von N ist. Es gilt*

$$\begin{aligned} E\left(e^{iu\sum_{k=1}^N \xi_k}\right) &= \sum_{n\geq 0} E\left(e^{iu\sum_{k=1}^n \xi_k}\right) e^{-\lambda}\frac{\lambda^n}{n!} \\ &= \sum_{n\geq 0}\left(\int_{\mathbb{R}} e^{iux}\,F(dx)\right)^n e^{-\lambda}\frac{\lambda^n}{n!} \\ &= e^{-\lambda}e^{\lambda\int_{\mathbb{R}} e^{iux}\,F(dx)} \underbrace{\sum_{n\geq 0} e^{-\lambda\int_{\mathbb{R}} e^{iux}\,F(dx)}\left(\lambda\int_{\mathbb{R}} e^{iux}\,F(dx)\right)^n \frac{1}{n!}}_{=1} \\ &= e^{\lambda\int_{\mathbb{R}}\left(e^{iux}-1\right)F(dx)}. \end{aligned} \tag{4.8}$$

Aus der Darstellung folgt sofort, dass die n-te Wurzel der charkteristischen Funktion die charakteristische Funktion einer zufälligen Summe mit Poisson-Parameter $\frac{\lambda}{n}$ ist (und der gleichen Verteilung der Summanden ξ_k). Damit ist $\sum_{k=1}^N \xi_k$ unendlich teilbar. Das Lévy-Tripel ist gegeben durch $\gamma = \lambda\int_{0<|x|\leq 1} xF(dx)$, $\sigma = 0$ und $\Pi(dx) = \lambda F(dx)$.

(iii) **Gammaverteilung**. *Betrachte für $\alpha, r > 0$ das Wahrscheinlichkeitsmaß mit Dichte*

$$f_{\alpha,r}(x) = \begin{cases} \frac{\alpha^r}{\Gamma(r)}x^{r-1}e^{-\alpha x} & : \text{für } x > 0 \\ 0 & : \text{für } x \leq 0. \end{cases}$$

α wird als Größenparameter und r als Formparameter bezeichnet. $\Gamma(r) := \int_0^\infty t^{r-1}\exp(-t)\,dt$ bezeichnet die Gamma-Funktion (für $n \in \mathbb{N}$ gilt $\Gamma(n) = (n-1)!$).

Für $r \in \mathbb{N}$ ist die Gammaverteilung die Verteilung der Summe von r unabhängigen, exponentialverteilten Zufallsvariablen mit gleichem Parameter α. Der Zeitpunkt Y des r-ten Sprungs eines Poisson-Prozesses N mit Parameter α ist also eine gammaverteilte Zufallsvariable. Für alle $t > 0$ gilt nämlich

$$P(Y \leq t) = P(N_t \geq r) = 1 - e^{-\alpha t}\sum_{k=0}^{r-1}\frac{(\alpha t)^k}{k!} = \int_0^t \frac{\alpha^r}{(r-1)!}x^{r-1}e^{-\alpha x}\,dx,$$

wobei sich die letzte Gleichung durch Differentiation nach t ergibt (die Produktregel führt zu einer Teleskopsumme).

Wir rechnen zunächst die Laplace-Transformierte der Gammaverteilung aus. Für *alle* $\lambda \in \mathbb{R}_+$ *gilt*

$$\begin{aligned}
\int_0^\infty e^{-\lambda x} f_{\alpha,r}(x)\,dx &= \frac{\alpha^r}{(\alpha+\lambda)^r\,\Gamma(r)} \int_0^\infty (\alpha+\lambda)((\alpha+\lambda)x)^{r-1} e^{-(\alpha+\lambda)x}\,dx \\
&\overset{y:=(\alpha+\lambda)x}{=} \frac{\alpha^r}{(\alpha+\lambda)^r\,\Gamma(r)} \int_0^\infty y^{r-1} e^{-y}\,dy \\
&= \frac{\alpha^r}{(\alpha+\lambda)^r} \\
&= \frac{1}{\left(1+\frac{\lambda}{\alpha}\right)^r}. \qquad (4.9)
\end{aligned}$$

Auf dem Gebiet $\mathbb{G} := \{a+bi \mid a \in \mathbb{R}_-, b \in \mathbb{R}\}$ *sind die Funktionen* $(a+bi) \mapsto \int_0^\infty e^{(a+bi)x} f_{\alpha,r}(x)\,dx$ *und* $(a+bi) \mapsto \frac{1}{\left(1-\frac{a+bi}{\alpha}\right)^r}$ *endlich und holomorph. Wegen (4.9) ist die komplexe Null Häufungspunkt der Menge, wo die beiden Funktionen übereinstimmen. Da beide Funktionen holomorph sind, folgt mit dem Identitätssatz, dass sie auf ganz* $\mathbb{G}$ *übereinstimmen. Insbesondere folgt für die charakteristische Funktion*

$$\int_0^\infty e^{iux} f_{\alpha,r}(x)\,dx = \frac{1}{\left(1-\frac{iu}{\alpha}\right)^r} \qquad (4.10)$$

Mit

$$\frac{1}{\left(1-\frac{iu}{\alpha}\right)^r} = \left[\frac{1}{\left(1-\frac{iu}{\alpha}\right)^{r/n}}\right]^n$$

folgt die unendliche Teilbarkeit der Gammaverteilung (die n*-te Wurzel der Verteilung ist sogar wieder eine Gamma-Verteilung, nun mit Parameter* $(\alpha, r/n)$*).*

Für das Lévy-Tripel erhalten wir

$$\gamma = \int_0^1 x\Pi(dx), \qquad (4.11)$$

$\sigma = 0$ *und*

$$\Pi(dx) = \begin{cases} rx^{-1}e^{-\alpha x}\,dx & : \text{für } x > 0 \\ 0 & : \text{für } x \le 0. \end{cases}$$

Dies kann man mit folgendem Lemma zeigen

Lemma 4.13 (Frullani Integral)**.** *Für alle* $\alpha, r > 0$ *und* $z \in \mathbb{C}$*, s.d.* $\mathrm{Re}(z) \le 0$ *gilt*

$$\frac{1}{(1-z/\alpha)^r} = \exp\left(-\int_0^\infty (1-\exp(zx))rx^{-1}\exp(-\alpha x)\,dx\right).$$

Mit dem Lemma folgt für den charakteristischen Exponenten

$$\psi(u) = r \int_0^\infty \left(e^{iux} - 1\right) \frac{1}{x} e^{-\alpha x}\, dx = -r \ln\left(1 - \frac{iu}{\alpha}\right), \quad \forall u \in \mathbb{R}. \tag{4.12}$$

Die Wahl von a in (4.11) garantiert nun, dass der Ausdruck $iu1_{|x|\le 1}$ *in dem Integral nach* Π *kompensiert wird und daher im mittleren Term von (4.12) nicht mehr vorkommt.*

Nach Theorem 4.6 existiert ein Lévy-Prozess $(X_t)_{t\in\mathbb{R}_+}$*, so dass* X_1 *gammaverteilt ist. Ein solcher Prozess wird* **Gamma-Prozess** *genannt.*

Wie man in (ii) sieht, muss die Lévy-Khintchine Formel für einen zusammengesetzten Poisson-Prozess die Gestalt $\psi(u) = \lambda \int_{\mathbb{R}} \left(e^{iux} - 1\right) F(dx)$ *mit* $F(\mathbb{R}) = 1$ *besitzen. Insbesondere ist das Lévy-Maß eines zusammengesetzten Poisson-Prozesses endlich, also* $\Pi(\mathbb{R}) < \infty$*. Der Gamma-Prozess ist demnach* **kein** *zusammengesetzter Poisson-Prozess. Er besitzt unendlich viele Sprünge. Der Gamma-Prozess ist nicht-fallend, besitzt also trotz der unendlich vielen Sprünge endliche Variation (Offenbar ist jeder Lévy-Prozess* X *mit* $P(X_1 \ge 0) = 1$ *nicht-fallend).*

(i) und (ii) aus Beispiel 4.12 zusammen mit Bemerkung 4.11 und der Tatsache, dass die Faltung zweier unendlich teilbaren Verteilungsfunktionen wieder unendlich teilbar ist, zeigen eine Richtung von Theorem 4.9 (aus der Darstellung (4.7) folgt die unendliche Teilbarkeit) für den Spezialfall, dass Π ein endliches Maß ist, also $\int_{\mathbb{R}} 1\, \Pi(dx) < \infty$.

4.1 Poisson-Zufallsmaß

Definition 4.14 (Poisson-Zufallsmaß)**.** *Sei* $(\Omega, \mathcal{F}, P)$ *ein Wahrscheinlichkeitsraum,* $E \subset \mathbb{R}^n$ *und* μ *ein* σ*-endliches Maß auf* $\mathcal{B}(E)$[¶]*. Ein Poisson-Zufallsmaß* M *auf* $\mathcal{B}(E)$ *mit Intensitätsmaß* μ *ist eine Abbildung*

$$M : \Omega \times \mathcal{B}(E) \to \mathbb{N}_0 \cup \{+\infty\}, \quad (\omega, A) \mapsto M(\omega, A)$$

mit den Eigenschaften

(i) Für P*-fast alle* $\omega \in \Omega$ *ist* $M(\omega, \cdot)$ *ein* $\mathbb{N}_0 \cup \{+\infty\}$*-wertiges Radon-Maß*[‖] *auf* E *und für alle* $A \in \mathcal{B}(E)$ *ist* $M(\cdot, A)$ *eine* $\mathbb{N} \cup \{+\infty\}$*-wertige Zufallsvariable.*

(ii) Für jedes $A \in \mathcal{B}(E)$ *mit* $\mu(A) < \infty$ *ist* $M(\cdot, A)$ *eine Poisson-verteilte Zufallsvariable mit Parameter* $\mu(A)$*, d.h.*

$$P(M(\cdot, A) = k) = \exp(-\mu(A)) \frac{(\mu(A))^k}{k!}, \quad \forall k \in \mathbb{N}_0.$$

Für $A \in \mathcal{B}(E)$ *mit* $\mu(A) = \infty$ *gilt* $P(M(\cdot, A) = \infty) = 1$.

[¶] μ σ-endlich $:\Leftrightarrow \exists (E_n)_{n\in\mathbb{N}} \subset \mathcal{B}(E)$ mit $E = \bigcup_{n\in\mathbb{N}} E_n$ und $\mu(E_n) < \infty\ \forall n \in \mathbb{N}$.

[‖] D.h. kompakte Mengen haben endliche Masse.

(iii) Seien $m \in \mathbb{N}$ und $A_1, \ldots, A_m \in \mathcal{B}(E)$ disjunkt. Dann sind die Zufallsvariablen $M(\cdot, A_1), \ldots, M(\cdot, A_m)$ stochastisch unabhängig.

Bemerkung 4.15. *Im Folgenden ist $E = \mathbb{R}_+ \times (\mathbb{R}^d \setminus \{(0, \ldots, 0)\}$, wobei die erste Komponente die Zeit modelliert.*

Definition 4.16. *Sei $X = (X_t)_{t \in \mathbb{R}_+}$ ein $\mathbb{R}^d$-wertiger stochastischer Prozess mit càdlàg Pfaden. Definiere das zufällige Maß $J_X : \Omega \times \mathcal{B}(\mathbb{R}_+ \times (\mathbb{R}^d \setminus \{(0, \ldots, 0)\})) \to \mathbb{N}_0 \cup \{+\infty\}$*

$$J_X(A) = \#\{t \mid \Delta X_t \neq 0 \text{ und } (t, \Delta X_t) \in A\}, \quad \forall A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d).$$

J_X wird als Sprungmaß des Prozesses X bezeichnet ($\#\{\ldots\}$ ist die Anzahl der Elemente einer Menge, $\Delta X_t \neq 0$ bedeutet natürlich nur, dass in mindestens einer der d Komponenten ein Sprung stattfindet).

Da zu festem $\varepsilon > 0$ jeder Prozess mit càdlàg Pfaden auf dem Zeitintervall $[0, T]$ höchstens endlich viele Sprünge besitzt, die dem Betrag nach größer als ε ist, ist J_X ein σ-endliches Maß, das auf $[0, T] \times (\mathbb{R} \setminus [-\varepsilon, \varepsilon])^d$ endlich ist.

Definition 4.17 (Lévy-Maß)**.** *Sei $X = (X_t)_{t \in \mathbb{R}_+}$ ein $\mathbb{R}^d$-wertiger Lévy-Prozess. Das Maß $\nu : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ definiert durch*

$$\nu(B) := E\left(\#\left\{t \in [0,1] \mid \Delta X_t \neq 0,\ \Delta X_t \in B\right\}\right), \quad B \in \mathcal{B}(\mathbb{R}^d)$$

wird als Lévy-Maß bezeichnet.

(Man beachte, dass das Lévy-Maß Π aus Theorem 4.9 bzgl. einer Verteilung und nicht bzgl. eines Lévy-Prozesses definiert war. Ein Zusammenhang zwischen den beiden Maßen wird in Bemerkung 4.26 hergestellt werden)

Proposition 4.18. *Sei $X = (X_t)_{t \in \mathbb{R}_+}$ ein zusammengesetzter Poisson-Prozess, d.h. $X_t = \sum_{i=1}^{N_t} \xi_i$, wobei N ein Poisson-Prozess mit Parameter $\lambda > 0$ ist und $(\xi_i)_{i \in \mathbb{N}}$ eine i.i.d. Folge von Zufallsvariablen mit Verteilungsfunktion F, die unabhängig von N ist. Dann ist J_X ein Poisson-Zufallsmaß auf $\mathbb{R}_+ \times \mathbb{R}^d$ mit Intensitätsmaß $\mu(dt \times dx) = \lambda\, dt F(dx)$.*

Proof. Ad Eigenschaft (ii): Sei $A = (t_1, t_2] \times B$. Mit einer analogen Rechnung wie in

Beispiel 4.12(ii) gilt**

$$
\begin{aligned}
&E\left(\exp\left(iu\sum_{j=N_{t_1}+1}^{N_{t_2}} 1(\xi_j \in B)\right)\right)\\
&= \sum_{n\geq 0} E\left(\exp\left(iu\sum_{j=1}^{n} 1(\xi_j \in B)\right)\right) e^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^n}{n!}\\
&= \sum_{n\geq 0}\left(e^{iu}P(\xi_1\in B)+1-P(\xi_1\in B)\right)^n e^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^n}{n!}\\
&= e^{-\lambda(t_2-t_1)}e^{\lambda(t_2-t_1)((e^{iu}-1)P(\xi_1\in B)+1)}\\
&\quad \underbrace{\sum_{n\geq 0} e^{-\lambda(t_2-t_1)((e^{iu}-1)P(\xi_1\in B)+1)}\left(\lambda(t_2-t_1)((e^{iu}-1)P(\xi_1\in B)+1)\right)^n\frac{1}{n!}}_{=1}\\
&= \exp\left(\lambda(t_2-t_1)(e^{iu}-1)P(\xi_1\in B)\right), \quad \forall u\in\mathbb{R}.
\end{aligned}
$$

Dies ist die charakteristische Funktion der Poisson-Verteilung zum Parameter $\lambda(t_2 - t_1)P(\xi_1 \in B)$.

Ad Eigenschaft (iii): Nun wollen wir Unabhängigkeit von $J_X(A_1)$ und $J_X(A_2)$ für $A_1 \cap A_2 = \emptyset$ zeigen. Betrachte zunächst den Fall $A_1 = (t_1, t_2] \times B_1$ und $A_2 = (t_1, t_2] \times B_2$ mit $B_1 \cap B_2 = \emptyset$. Wir berechnen die charakteristische Funktion der Zufallsvariablen $(J_X(A_1), J_X(A_2))$.

$$
\begin{aligned}
&E\left(\exp\left(iuJ_X(A_1)+ivJ_X(A_2)\right)\right)\\
&= E\left(\exp\left(iu\sum_{j=N_{t_1}+1}^{N_{t_2}} 1(\xi_j\in B_1)+iv\sum_{j=N_{t_1}+1}^{N_{t_2}} 1(\xi_j\in B_2)\right)\right)\\
&= \sum_{n\geq 0}\left(e^{iu}P(\xi_1\in B_1)+e^{iv}P(\xi_1\in B_2)+1-P(\xi_1\in B_1)-P(\xi_1\in B_2)\right)^n e^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^n}{n!}\\
&= \exp\left(\lambda(t_2-t_1)\left[(e^{iu}-1)P(\xi_1\in B_1)+(e^{iv}-1)P(\xi_1\in B_2)\right]\right)\\
&= \exp\left(\lambda(t_2-t_1)(e^{iu}-1)P(\xi_1\in B_1)\right)\exp\left(\lambda(t_2-t_1)(e^{iv}-1)P(\xi_1\in B_2)\right)\\
&= E\left(\exp\left(iuJ_X(A_1)\right)\right)E\left(\exp\left(ivJ_X(A_2)\right)\right), \quad \forall u,v\in\mathbb{R}.
\end{aligned}
\tag{4.13}
$$

Damit sind $J_X(A_1)$ und $J_X(A_2)$ unabhängig. Für $A_1 = (s_1, s_2] \times B_1$ und $A_2 = (t_1, t_2] \times B_2$ mit $0 \leq s_1 \leq s_2 \leq t_1 \leq t_2 < \infty$ gilt die Aussage wegen der Unabhängigkeit der Zuwächse eines zusammengesetzen Poisson-Prozessen. Wegen der Additivität von J_X gilt die Aussage für den durchschnittsstabilen Erzeuger $\{(t_1, t_2] \times B \mid B \in \mathcal{B}(\mathbb{R}^d \setminus \{(0, \ldots, 0)\})\}$ von $\mathcal{B}(\mathbb{R}_+ \times (\mathbb{R}^d \setminus \{(0, \ldots, 0)\}))$ und somit (wegen σ-Additivität von J_X und μ) auch für die gesamte σ-Algebra. □

** $\sum_{n=0}^{\infty} b^n \frac{a^n}{n!}\exp(-a) = \sum_{n=0}^{\infty}\frac{(ba)^n}{n!}\exp(-ba)\exp(ba)\exp(-a) = \exp((b-1)a)$

Ein zusammengesetzter Poisson-Prozess X kann also geschrieben werden als

$$X_t = \sum_{s \leq t} \Delta X_s = \int_{[0,t] \times \mathbb{R}^d} x J(ds \times dx).$$

Proposition 4.19. *Für einen zusammengesetzten Poisson-Prozess X mit endlicher Varianz gilt*

$$E(X_t) = \lambda t \int_{\mathbb{R}} x F(dx), \quad \operatorname{Var}(X_t) = \lambda t \int_{\mathbb{R}} x^2 F(dx).$$

Proof. Folgt mit Darstellung (4.8) der charakteristischen Funktion:

$$\begin{aligned}
\frac{\partial E\left(e^{iu\sum_{k=1}^{N_t} \xi_k}\right)}{\partial u} &= \frac{\partial \exp\left(\lambda t \int_{\mathbb{R}} \left(e^{iux} - 1\right) \, F(dx)\right)}{\partial u} \\
&= \lambda t i \int_{\mathbb{R}} x e^{iux} \, F(dx) \exp\left(\lambda t \int_{\mathbb{R}} \left(e^{iux} - 1\right) \, F(dx)\right), \quad \forall u \in \mathbb{R},
\end{aligned}$$

und damit

$$\frac{\partial^2 E\left(e^{iu\sum_{k=1}^{N_t} \xi_k}\right)}{\partial u^2}\Bigg|_{u=0} = -\lambda t \int_{\mathbb{R}} x^2 \, F(dx) - \lambda^2 t^2 \left(\int_{\mathbb{R}} x \, F(dx)\right)^2.$$

Es folgt

$$E(X_t) = \frac{1}{i} \frac{\partial E\left(e^{iu\sum_{k=1}^{N_t} \xi_k}\right)}{\partial u}\Bigg|_{u=0} = \lambda t \int_{\mathbb{R}} x \, F(dx)$$

und

$$E(X_t^2) = -\frac{\partial^2 E\left(e^{iu\sum_{k=1}^{N_t} \xi_k}\right)}{\partial u^2}\Bigg|_{u=0} = \lambda t \int_{\mathbb{R}} x^2 \, F(dx) + \lambda^2 t^2 \left(\int_{\mathbb{R}} x \, F(dx)\right)^2.$$

Für die Varianz bedeutet dies

$$\operatorname{Varianz}(X_t) = E(X_t^2) - (E(X_t))^2 = \lambda t \int_{\mathbb{R}} x^2 \, F(dx).$$

□

Theorem 4.20 (Lévy-Itô Zerlegung)**.** *Sei $X = (X_t)_{t \in \mathbb{R}_+}$ ein $\mathbb{R}^d$-wertiger Lévy-Prozess mit Lévy-Maß ν aus Definition 4.17. Es gilt*

(i) ν ist ein σ-endliches Maß auf $\mathbb{R}^d \setminus \{0\}$ für das gilt

$$\int_{|x| \leq 1} |x|^2 \, \nu(dx) < \infty \quad \textit{and} \quad \int_{|x| \geq 1} \nu(dx) < \infty.$$

(ii) Das Sprungmaß von X, das wir mit J_X bezeichnen, ist ein zufälliges Poisson-Maß auf $\mathbb{R}_+ \times \mathbb{R}^d$ mit Intensität $\widetilde{\nu}(dx \times dt) := \nu(dx)dt$.

(iii) Es existiert ein $\gamma \in \mathbb{R}^d$ und eine d-dimensionale Brownsche Bewegung $(B_t)_{t\in\mathbb{R}_+}$ mit Kovarianzmatrix $A \in \mathbb{R}^{d\times d}$ s.d.[††]

$$X_t = \gamma t + B_t + X_t^1 + \lim_{\varepsilon\downarrow 0} \widetilde{X}_t^\varepsilon \tag{4.14}$$

wobei

$$X_t^1 := \int_{0\le s\le t,\ |x|>1} x J_X(ds \times dx)$$

$$\begin{aligned}\widetilde{X}_t^\varepsilon &:= \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} x(J_X(ds \times dx) - \widetilde{\nu}(ds \times dx)) \\ &= \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} x\widetilde{J}_X(ds \times dx)\end{aligned}$$

mit $\widetilde{J}_X := J_X - \widetilde{\nu}$ (die Differenz ist wohldefiniert). Die Prozesse X^1, $\widetilde{X}^\varepsilon$ und B sind unabhängig voneinander. Die Konvergenz in (4.14) gilt P-f.s. und für $T \in \mathbb{R}_+$ gleichmäßig in $t \in [0, T]$.

(Offenbar sind γ und A durch den Lévy-Prozess eindeutig bestimmt)

Definition 4.21. *(γ, A, ν) wird das* **charakteristische Tripel** *des Lévy-Prozesses X genannt.*

Zum Beweis von Theorem 4.20 benötigen wir folgendes Lemma.

Lemma 4.22. *Sei $(X_t, Y_t)_{t\in\mathbb{R}_+}$ ein Lévy-Prozess. Wenn Y ein zusammengesetzter Poisson-Prozess ist und X und Y niemals zusammen springen, d.h. $P(\Delta X_t \Delta Y_t = 0,\ \forall t \in \mathbb{R}_+) = 1$, dann sind die Prozesse stochastisch unabhängig voneinander.*

Beweisskizze. Schritt 1: Zunächst zeigen wir, dass für alle $s \le t$ die Zufallsvariablen $X_t - X_s$ und $Y_t - Y_s$ unabhängig voneinander sind. O.B.d.A. $s = 0$ und $t = 1$. Für alle $u, v \in \mathbb{R}$ betrachte die Prozesse

$$M_t = \frac{\exp(iuX_t)}{E[\exp(iuX_t)]} \quad \text{und} \quad N_t = \frac{\exp(ivY_t)}{E[\exp(ivY_t)]}, \quad t \ge 0.$$

[††] Eine d-dimensionale Brownsche Bewegung ist ein Prozess, der sich als lineare Transformation einer d-dimensionalen Standard-Brownschen Bewegung $\widetilde{B}$ schreiben lässt, d.h. $B = C\widetilde{B}$ für eine Matrix C (die Komponenten von $\widetilde{B}$ sind stochastisch unabhängig). Die Zuwächse von B sind also *multivariat* normalverteilt (was mehr ist als normalverteilte Komponenten !) Für die Kovarianzmatrix gilt $A = C^\top C$. Man beachte, dass $B = 0$ nicht ausgeschlossen ist, was bedeuten würde, dass der Brownsche Anteil in (4.14) verschwindet.

Man beachte, dass wegen der unendlichen Teilbarkeit von X_t und Y_t die charakteristischen Funktionen $E[\exp(iuX_t)]$ und $E[\exp(ivY_t)]$ nicht verschwinden können, siehe Lemma 7.5 in Sato [10]. Für $s \leq t$ gilt

$$\frac{\exp(iuX_t)}{E[\exp(iuX_t)]} = \frac{\exp(iuX_s)}{E[\exp(iuX_s)]} \frac{\exp(iu(X_t - X_s))}{E[\exp(iu(X_t - X_s))]}$$

und $X_t - X_s$ ist stochastisch unabhängig von $(X_u, Y_u)_{0 \leq u \leq s}$. Damit ist M (und analog N) ein komplexwertiges Martingal bzgl. der von X und Y erzeugten Filtration (d.h. $\mathrm{Re}(M)$ und $\mathrm{Im}(M)$ sind reellwertige Martingale). Zudem sind M und N auf kompakten Zeitintervallen beschränkt.

N hat endliche und sogar integrierbare Variation auf $[0,1]$. Für jedes n folgt aus der Martingaleigenschaft beider Prozesse

$$\begin{aligned}
E(M_1 N_1) - 1 &= E\left[\sum_{k=1}^{n} \left(M_{k/n} N_{k/n} - M_{(k-1)/n} N_{(k-1)/n}\right)\right] \\
&= \underbrace{E\left[\sum_{k=1}^{n} M_{(k-1)/n} \left(N_{k/n} - N_{(k-1)/n}\right)\right] + E\left[\sum_{k=1}^{n} N_{(k-1)/n} \left(M_{k/n} - M_{(k-1)/n}\right)\right]}_{=0} \\
&\quad + E\left[\sum_{k=1}^{n} \left(M_{k/n} - M_{(k-1)/n}\right) \left(N_{k/n} - N_{(k-1)/n}\right)\right] \\
&= E\left[\sum_{k=1}^{n} \left(M_{k/n} - M_{(k-1)/n}\right) \left(N_{k/n} - N_{(k-1)/n}\right)\right] \\
&= E\left[\int_0^1 \left(M_{[(sn+1)-]/n} - M_{[(sn)-]/n}\right) dN_s\right],
\end{aligned}$$

wobei $[\ldots]$ die Gaußklammer bezeichnet und

$$[u-] = \begin{cases} u - 1 : \text{für } u \in \mathbb{N} \\ [u] \quad : \text{sonst.} \end{cases}$$

Für $n \to \infty$ konvergiert der Integrand punktweise gegen ΔM (hierzu beachte man, dass $[(sn)-]/n < s \leq [(sn+1)-]/n$ für alle $s > 0$), was wegen der Stetigkeit des Integrals

$$E\left[\int_0^1 \left(M_{[(sn+1)-]/n} - M_{[(sn)-]/n}\right) dN_s\right] \to E\left[\int_0^1 \Delta M_s \, dN_s\right] = E\left[\sum_{s \leq t} \Delta M_s \Delta N_s\right] = 0$$

Also gilt $E(M_1 N_1) = 1$ und damit

$$E\left[\exp(iuX_1)\exp(ivY_1)\right] = E\left[\exp(iuX_1)\right] E\left[\exp(ivY_1)\right], \quad \forall u, v \in \mathbb{R}.$$

D.h. X_1 und Y_1 sind unabhängig.

Schritt 2: Seien $n \in \mathbb{N}$ und $t_1 \leq \ldots \leq t_n$. Aus Schritt 1 und der Unabhängigkeit der Zuwächse von (X, Y) folgt die gemeinsame Unabhängigkeit der $2n$ Zufallsvariablen

$$X_{t_1}, Y_{t_1}, X_{t_2} - X_{t_1}, Y_{t_2} - Y_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}, Y_{t_n} - Y_{t_{n-1}}$$

(die charakteristische Funktion dieses $2n$-dimensionalen Zufallsvektors lässt sich faktorisieren). Damit sind insbesondere die beiden Zufallsvektoren $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ und $(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n})$ voneinander stochastisch unabhängig. □

Lemma 4.23. *Sei $E \subset \mathbb{R}^n$ und M ein Poisson-Zufallsmaß auf $\mathcal{B}(E)$ mit Intensität μ. Für jede Menge $B \in \mathcal{B}(E)$ und jede messbare Funktion $f : E \to \mathbb{C}$ mit $\mu(B) < \infty$ und $\int_B e^{f(x)} \mu(dx) < \infty$ gilt*

$$E\left[\exp\left(\int_B f(x) M(dx)\right)\right] = \exp\left[\int_B \left(e^{f(x)} - 1\right) \mu(dx)\right]$$

Proof. Für Elementarfunktionen, d.h. $f = \sum_{k=1}^m \alpha_k 1_{A_k}, \alpha_k \in \mathbb{C}, A_k \in \mathcal{B}(E), A_1, \ldots, A_m$ disjunkt, gilt die Aussage mit der gleichen Rechnung wie in (4.13). □

Beweis von Theorem 4.20 ohne Benutzung von Theorem 4.9. [‡‡]

1. Schritt: Für $\varepsilon > 0$ definiere

$$\widehat{X}_t^\varepsilon \quad := \quad \int_{0 \leq s \leq t,\ \varepsilon \leq |x| \leq 1} x J_X(ds \times dx)$$

und $R^\varepsilon := X - \widehat{X}^\varepsilon$. $\widehat{X}^\varepsilon$ ist ein zusammengesetzter Poissonprozess. Offenbar ist $(\widehat{X}^\varepsilon, R^\varepsilon)$ wieder ein Lévy-Prozess. Nehme ein beliebiges $u \neq 0$ und $t > 0$. Es gilt $E(\exp(iuX_t)) \neq 0$ (siehe Lemma 7.5 in Sato [10]) und damit

$$|E(\exp(iuX_t))| > 0.$$

Da mit Lemma 4.22 die Prozesse $\widehat{X}^\varepsilon$ und R^ε unabhängig sind, gilt

$$E(\exp(iuX_t)) = E(\exp(iu\widehat{X}_t^\varepsilon))E(\exp(iuR_t^\varepsilon))$$

Da $|E(\exp(iuR_t^\varepsilon))| \leq E(|\exp(iuR_t^\varepsilon)|) = 1$ ist $|E(\exp(iu\widehat{X}_t^\varepsilon))|$ gleichmäßig in ε von der Null entfernt. Aus Lemma 4.23 folgt

$$\left|\exp\left(t \int_{\varepsilon \leq |x| \leq 1} \left(e^{iux} - 1\right) \nu(dx)\right)\right| \geq C > 0$$

was

$$-\infty < \ln(C) \leq \mathrm{Re}\left(t \int_{\varepsilon \leq |x| \leq 1} \left(e^{iux} - 1\right) \nu(dx)\right) = t \int_{\varepsilon \leq |x| \leq 1} (\cos(ux) - 1)\, \nu(dx)$$

[‡‡] Zur besseren Lesbarkeit ist der Beweis für $d = 1$ aufgeschrieben. Man kann ihn jedoch auch mehrdimensional lesen. $|x|$ ist irgendeine zu $\max_{i=1,\ldots,d} |x_i|$ äquivalente Norm. Integrale wie $\int_{0 \leq s \leq t,\ \varepsilon \leq |x| \leq 1} x J_X(ds \times dx)$ sind vektorwertige Zufallsvariablen.

und damit

$$\int_{\varepsilon\leq|x|\leq 1}\left(|x|^2\wedge 1\right)\nu(dx)\leq \widetilde{C}\int_{\varepsilon\leq|x|\leq 1}(1-\cos(ux))\,\nu(dx)\leq \widehat{C}<\infty \tag{4.15}$$

für geeignete Konstanten $\widehat{C}$ und $\widetilde{C}$ (unabhängig von ε) nach sich zieht. Da $t\mapsto\int_{0\leq s\leq t,\ |x|>1} xJ_X(ds\times dx)$ ein zusammengesetzter Poissonprozess ist, ist zudem die erwartete Anzahl der Sprünge von X im Zeitintervall $[0,1]$, die dem Betrage nach größer als 1 sind, endlich, also $\nu((1,\infty))<\infty$. Aus (4.15) für $\varepsilon\to 0$ folgt

$$\int_{\mathbb{R}}\left(|x|^2\wedge 1\right)\nu(dx)<\infty. \tag{4.16}$$

2. Schritt: Definiere

$$Y^n:=\widetilde{X}^{1/(n+1)}-\widetilde{X}^{1/n}=\int_{0\leq s\leq t,\ 1/(n+1)\leq|x|<1/n} x\left(J_X(ds\times dx)-\widetilde{\nu}(ds\times dx)\right)$$

$(Y^n)_{n\in\mathbb{N}}$ sind kompensierte zusammengesetzte Poissonprozesse. Mit Proposition 4.19 gilt $E(Y_t^n)=0$ und $\operatorname{Var}(Y_t^n)=t\int_{1/(n+1)\leq|x|<1/n}|x|^2\nu(dx)$. Zudem sind Y^n voneinander unabhängig. Daher gilt wegen (4.16)

$$\begin{aligned}
\operatorname{Var}\left(\sum_{k=n}^{m}Y_t^k\right) &= \sum_{k=n}^{m}\operatorname{Var}\left(Y_t^k\right)\\
&= t\int_{1/(m+1)\leq|x|<1/n}|x|^2\nu(dx)\\
&\leq t\int_{\mathbb{R}}\left(|x|^2\wedge 1\right)\nu(dx)<\infty,\quad \forall n,m\in\mathbb{N},\ n\leq m.
\end{aligned}\tag{4.17}$$

Definiere die Prozesse $Z^n:=\sum_{k=1}^{n}Y^k$. Für jedes $t\in\mathbb{R}_+$ ist $(Z_t^n)_{n\in\mathbb{N}}$ wegen (4.17) eine $L^2(\Omega,\mathcal{F},P)$-Cauchy Folge. Da Z^n bzgl. der Filtration $\mathcal{F}_t^X:=\sigma(X_s,\ s\leq t,\mathcal{N})$ quadratintegrierbare Martingale sind, folgt mit der Doobschen Ungleichung

$$\begin{aligned}
E\left(\sup_{s\in[0,t]}(Z_s^n-Z_s^m)^2\right) &\leq 4E\left((Z_t^n-Z_t^m)^2\right)\\
&= 4t\int_{1/(m+1)\leq|x|<1/n}|x|^2\nu(dx)\\
&\leq 4t\int_{\mathbb{R}}\left(|x|^2\wedge 1\right)\nu(dx)<\infty,\quad \forall n,m\in\mathbb{N}.
\end{aligned}\tag{4.18}$$

Für quadratintegrierbare Martingale auf $\mathbb{R}_+$ definiere die Metrik

$$d(M,M'):=\inf\left\{\varepsilon>0\mid E\left(\sup_{t\in[0,1/\varepsilon]}(M_t-M_t')^2\right)\leq\varepsilon\right\}.$$

Aus (4.18) folgt, dass die Folge von Martingalen $(Z^n)_{n\in\mathbb{N}}$ bzgl. der Metrik d eine Cauchy-Folge ist. Wegen der Vollständigkeit des Raumes der quadratintegrierbaren Martingale ausgestattet mit der Metrik d gibt es einen $\mathcal{F}^X$-adaptierten Grenzprozess Z mit $d(Z^n, Z) \to 0,\ n \to \infty$. Der Grenzprozess $\lim_{\varepsilon\to 0} \widetilde{X}^\varepsilon := Z$ existiert also (und ist ein Martingal). Hierzu wird benutzt, dass $\mathcal{F}^X$ die "usual conditions" erfüllt (auf Pfaden, wo Z^n nicht konvergiert, kann Z identisch Null gesetzt werden). Mit Kolmogorov's three series theorem konvergiert die Folge $(\sup_{0\le s\le t} |Z^n_s - Z_s|)_{n\in\mathbb{N}}$ für alle $t \in \mathbb{R}_+$ auch P-f.s. gegen Null.

Der Prozess $X^c := X - X^1 - \lim_{\varepsilon\to 0} \widetilde{X}^\varepsilon$ ist nach Konstruktion stetig und mit Lemma 4.22 unabhängig von X^1 und unabhängig von allen $\widetilde{X}^\varepsilon$, $\varepsilon > 0$, und damit unabhängig von $\lim_{\varepsilon\to 0} \widetilde{X}^\varepsilon$.

3. Schritt: Es bleibt zu zeigen, dass X^c eine Brownsche Bewegung mit Drift ist. X^c ist offenbar ein Lévy-Prozess, d.h. die Zuwächse sind schonmal unabhängig. Sei $s < t$. Für $n \in \mathbb{N}$ definiere

$$\xi_{nk} := X^c_{s+\frac{k}{n}(t-s)} - X^c_{s+\frac{k-1}{n}(t-s)}, \quad k = 1, \dots, n.$$

Also

$$X^c_t - X^c_s = \sum_{k=1}^n \xi_{nk}.$$

Die Zufallsvariablen $\xi_{n1}, \dots, \xi_{nn}$ sind i.i.d. (die Verteilung hängt nur vom ersten, nicht jedoch vom zweiten Index ab) und aus der Stetigkeit von X^c folgt

$$\max_{k=1,\dots,n} |\xi_{nk}| \to 0, \quad n \to \infty, \quad P - \text{f.s.}$$

Mit dem zentralen Grenzwertsatz von Feller/Lévy (siehe z.B. Theorem 4.15 in Kallenberg [3]) folgt, dass $X^c_t - X^c_s$ normalverteilt ist* □

Corollary 4.24. *Ein stetiger Prozess X mit unabhängigen Zuwächsen sowie $E(X_t) = 0$ und $\mathrm{Var}(X_t) = t$ für alle $t \in \mathbb{R}_+$ ist eine Standard Brownsche Bewegung.*

Proof. Folgt aus der Tatsache, dass in diesem Spezialfall $X^c = X$. Zudem muss man beachten, dass für die Anwendung des Grenzwertsatz von Feller/Lévy nicht gebraucht wird, dass $\xi_{nk_1} \overset{\mathcal{D}}{=} \xi_{nk_2}$. □

Wenn man die Brownsche Bewegung mathematisch elegant einführen möchte und dabei Verständlichlichkeit nicht besonders wichtig ist, würde man also auf die Forderung, dass die Zuwächse normalverteilt sind, verzichten, da sich dies bereits aus den anderen Forderungen, die man an eine Brownsche Bewegung stellt, ableiten lässt.

Corollary 4.25 (Korollar der Lévy-Itô Zerlegung)**.** *Jeder Lévy-Prozess X ist ein Semimartingal bzgl. der Filtration $\mathcal{F}^X_t := \sigma(X_s,\ s \le t,\ \mathcal{N})$, wobei $\mathcal{N}$ alle Nullmengen von $(\Omega, \mathcal{F}, P)$ bezeichnet.*

---

*Im mehrdimensionalen Fall wende man das Argument auf beliebige Linearkombinationen $\sum_{i=1}^d \lambda_i((X^c)^i_t - (X^c)^i_s)$ an, die dann auch normalverteilt sein müssen. Mit dem Cramer-Wold Device folgt, dass der Vektor $X^c_t - X^c_s$ *multivariat* normalverteilt ist.

Proof. Dies folgt aus der Darstellung (4.14). Zunächst machen wir die Feststellung, dass die Prozesse X^1 und $\widetilde{X}^\varepsilon$ adaptiert sind (Sprünge von X sind mit dem Informationsverlauf $\mathcal{F}^X$ beobachtbar und Kompensation deterministisch). Der Prozess $t \mapsto \gamma t$ ist es sowieso. Damit sind alle 4 Summanden in (4.14) adaptiert. Der Prozess $t \mapsto \gamma t + X_t^1$ kann für die Semimartingal-Zerlegung als Prozess von endlicher Variation und $B + \lim_{\varepsilon \to 0} \widetilde{X}^\varepsilon$ als Martingalanteil gewählt werden. □

(Die ensprechende Aussage für die Filtration $\mathcal{F}^0(X) := \sigma(X_s,\ s \le t)$ ohne Nullmengenerweiterung gilt genauso, im Beweis muss aber ein klein bisschen anders argumentiert werden, wieso ?)

Nun können wir ohne wesentlichen zusätzlichen Aufwand die Lévy-Khintchine Formel beweisen.

Beweis von Theorem 4.9. Sei F eine unendlich teilbare Verteilungsfunktion. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ auf dem ein Lévy-Prozess $X = (X_t)_{t \in \mathbb{R}_+}$ definiert ist, so dass X_1 wie F verteilt ist (vgl. Theorem 4.6).

Die Lévy-Itô Zerlegung von X zeigt, dass für jedes $t \in \mathbb{R}$ die Zufallsvariable $X_t^c + X_t^1 + \widetilde{X}_t^\varepsilon$ für $\varepsilon \to 0$ gegen X_t konvergiert. Dies zieht Konvergenz der charakteristischen Funktionen nach sich. Da die Komponenten X_t^c, X_t^1 und $\widetilde{X}_t^\varepsilon$ unabhängig sind, gilt

$$\begin{aligned} E(\exp(iu(X_t^c + X_t^1 + \widetilde{X}_t^\varepsilon))) &= \exp\left(it\gamma u - \frac{1}{2}t\sigma^2 u^2\right) \times \exp\left(t \int_{\mathbb{R}} \left(e^{iux} - 1\right) \nu(dx)\right) \\ &\quad \times \exp\left(t \int_{\varepsilon \le |x| \le 1} \left(e^{iux} - 1 - iux\right) \nu(dx)\right), \quad \forall u \in \mathbb{R} \end{aligned}$$

und wegen (4.16) konvergiert der Ausdruck für $\varepsilon \to$ gegen

$$\exp\left(it\gamma u - \frac{1}{2}t\sigma^2 u^2 + t \int_{\mathbb{R}} \left(e^{iux} - 1 - iux1_{\{|x| \le 1\}}\right) \nu(dx)\right), \quad u \in \mathbb{R}. \tag{4.19}$$

(4.19) muss dann die charakteristische Funktion der Zufallsvariablen X_t sein. Mit $t = 1$ folgt, dass der charakteristische Exponent von F die Darstellung (4.7) besitzt. □

Bemerkung 4.26. *Das Tripel (γ, σ, Π) aus Theorem 4.9 ist eindeutig durch F gegeben (ohne Beweis, siehe Beweis von Theorem 8.1(ii) in Sato [10]). Daher stimmt das Lévy-Tripel (γ, σ, Π) der Verteilung von X_1 mit dem charakteristischen Tripel (γ, A, ν) des Lévy-Prozesses X überein.*

Aus Theorem 4.20 (Lévy-Itô Zerlegung) werden wir einige (Pfad-)Eigenschaften eines Lévy-Prozesses in Abhängigkeit von seinem charakteristischen Tripel (γ, A, ν) herleiten.

Proposition 4.27. *Ein reellwertiger Lévy-Prozess ist von endlicher Variation (besitzt also mit Wahrscheinlichkeit 1 endliche Variation), wenn für sein charakteristisches Tripel (γ, A, ν) gilt*

$$A = 0 \quad \textit{und} \quad \int_{|x| \le 1} |x| \nu(dx) < \infty. \tag{4.20}$$

Wenn (4.20) nicht erfüllt ist, dann besitzt X mit Wahrscheinlichkeit 1 *unendliche Variation.*

Proof. 1. Teil der Aussage: Setze man (4.20) voraus. In diesem Fall kann der Grenzübergang im Beweis von Theorem 4.20 auch ohne Kompensation des Sprungmaßes J_X durch ν durchgeführt werden und der (zufällige) Gaußsche Anteil verschwindet. Wir werden zeigen, dass

$$X_t = bt + \int_{0\le s\le t,\ |x|>1} x J_X(ds\times dx) + \lim_{\varepsilon\downarrow 0} X_t^\varepsilon \tag{4.21}$$

für ein $b\in\mathbb{R}^d$, wobei $X_t^\varepsilon = \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} xJ_X(ds\times dx)$. Die ersten beiden Terme sind offenbar von endlicher Variation. Für jedes $\varepsilon>0$ ist X^ε ein zusammengesetzter Poisson-Prozess mit Variation

$$\text{Variation}(X^\varepsilon)_t = \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} |x|J_X(ds\times dx).$$

Sei $(\varepsilon_n)_{n\in\mathbb{N}}$ eine Nullfolge. Mit Fubini (für Übergangskerne)† folgt

$$\begin{aligned} E\left(\text{Variation}\left(X^{\varepsilon_{n+1}} - X^{\varepsilon_n}\right)_t\right) &= E\left(\int_{0\le s\le t,\ \varepsilon_{n+1}\le|x|<\varepsilon_n} |x|J_X(ds\times dx)\right) \\ &= t\int_{\varepsilon_{n+1}\le|x|<\varepsilon_n} |x|\nu(dx). \end{aligned}$$

Da die rechte Seite summierbar in $n\in\mathbb{N}$ ist, existiert wegen der Vollständigkeit des Raumes der adaptierten Prozesse mit càdlàg Pfaden und endlicher Variation ausgestattet mit der Variationsnorm $d(A,A') := E\left(\text{Variation}(A-A')_t\right)$ ein reellwertiger Grenzprozess $\lim_{\varepsilon\downarrow 0} X^\varepsilon$ mit càdlàg Pfaden und endlicher Variation gegen den X^ε in der Variationsnorm konvergiert, also

$$E\left(\text{Variation}(\lim_{\varepsilon\downarrow 0} X^\varepsilon)_t\right) = t\int_{0<|x|\le 1} |x|\nu(dx) < \infty.$$

---

†Es mag nützlich sein, J_X als Übergangskern von $(\Omega,\mathcal{F})$ nach $(\mathbb{R}_+\times\mathbb{R}^d, \mathcal{B}(\mathbb{R}_+\times\mathbb{R}^d))$ zu betrachten und das Maß

$$P\otimes J_X : \mathcal{F}\otimes\mathcal{B}(\mathbb{R}_+\times\mathbb{R}^d)\to[0,\infty],\ A\mapsto \int\int 1_A(\omega,s,x)J_{X.(\omega)}(ds,dx)P(d\omega)$$

zu definieren. $P\otimes J_X$ nennt man das Produkt von P und J_X. Offenbar gilt für alle $B\in\mathcal{B}(\mathbb{R}_+\times\mathbb{R}^d)$

$$P\otimes J_X(\Omega\times B) = \widetilde{\nu}(B),$$

wobei $\widetilde{\nu}$ das Intensitätsmaß von J_X ist (siehe Theorem 4.20(ii)). Der Satz von Fubini für Übergangskerne besagt, dass

$$\int f\, d(P\otimes J_X) = \int\int f(\omega,s,x)\, J_{X.(\omega)}(ds,dx)P(d\omega)$$

Man wendet ihn nun auf die Funktion $f(\omega,s,x) = |x|1_{0\le s\le t,\ \varepsilon_{n+1}\le|x|<\varepsilon_n}$ an. Da f nicht von ω abhängt, fällt die linke Seite mit $\int |x|1_{0\le s\le t,\ \varepsilon_{n+1}\le|x|<\varepsilon_n}\widetilde{\nu}(ds,dx) = t\int |x|1_{\varepsilon_{n+1}\le|x|<\varepsilon_n}\nu(dx)$ zusammen.

Aus Theorem 4.20 folgt (4.21) mit $b = \gamma - \int_{|x|\le 1} x\nu(dx)$.

Die Variation des Lévy-Prozesses nach Herausnahme der großen Sprünge ist also sogar integrierbar.

2. Teil der Aussage: Man setze $P(\text{Variation}(X)_t < \infty) > 0$ voraus. Betrachte die Lévy-Itô-Zerlegung (4.14). Sei $\varepsilon > 0$. Die Variation eines Prozesses X mit càdlàg Pfaden ist stets größer als die Variation, die nur von den Sprüngen von X, die dem Betrage nach grösser als ε sind, erzeugt wird, also

$$\text{Variation}(X)_t \ge \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} |x| J_X(ds\times dx) \tag{4.22}$$

(die rechte Seite von (4.22) ist für alle Pfade ω endlich). Es folgt

$$\begin{aligned} &\text{Variation}(X)_t \\ &\ge \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} |x| J_X(ds\times dx) \\ &= t\int_{\varepsilon\le|x|\le 1} |x|\nu(dx) + \int_{0\le s\le t,\ \varepsilon\le|x|\le 1} |x|\left(J_X(ds\times dx) - \nu(dx)\,ds\right) \end{aligned} \tag{4.23}$$

Die Varianz der Zufallsvariablen auf der rechten Seite ist mit Proposition 4.19

$$t\int_{\varepsilon<|x|\le 1} |x|^2\nu(dx)$$

Wegen $\int_{0<|x|\le 1} |x|^2\nu(dx) < \infty$ konvergiert

$$\int_{0\le s\le t,\ \varepsilon\le|x|\le 1} |x|\left(J_X(ds\times dx) - \nu(dx)\,ds\right) \tag{4.24}$$

mit Kolmogorovs "three series theorem" fast sicher gegen eine endliche Zufallsvariable (analog zum Beweis der Lévy-Itô Zerlegung).

Nehme nun an, $\int_{0<|x|\le 1} |x|\nu(dx) = \infty$. Dies bedeutet, dass $\int_{\varepsilon\le|x|\le 1} |x|\nu(dx)$ für $\varepsilon \to 0$ gegen unendlich konvergiert, was wegen (4.23) und der Konvergenz von (4.24) gegen eine $\mathbb{R}$-wertige Zufallsvariable mit $P(\text{Variation}(X)_t < \infty) > 0$ nicht vereinbar ist. Also gilt

$$\int_{0<|x|\le 1} |x|\nu(dx) < \infty. \tag{4.25}$$

Wenn nun aber (4.25) gilt, dann lässt sich X wegen Theorem 4.20 schreiben als

$$X_t = (\gamma - \int_{|x|\le 1} x\nu(dx))t + \sigma B_t + \int_{0\le s\le t,\ x\in\mathbb{R}^d} xJ_X(ds\times dx).$$

Mit den Argumenten aus dem Beweis des 1. Teils ist $\int_{0\le s\le t,\ x\in\mathbb{R}^d} xJ_X(ds\times dx)$ ein Prozess von endlicher Variation (hat also mit Wahrscheinlichkeit 1 endliche Variation). Also muss wegen der Annahme $P(\text{Variation}(X)_t < \infty) > 0$ der Prozess $t \mapsto \sigma B_t + (\gamma - \int_{|x|\le 1} x\nu(dx))t$ zumindest mit positiver Wahrscheinlichkeit endliche Variation haben, was $\sigma = 0$ nach sich zieht. □

Bemerkung 4.28. *Es gilt*

$$E(\text{Variation}(X)_t) \geq E\left(\underbrace{\int_{0\leq s\leq t,\ 0<|x|<1} |x| J_X(ds\times dx)}_{=\sum_{0<s\leq t}|\Delta X_s|}\right) = t\int_{0<|x|<1} |x|\nu(dx). \quad (4.26)$$

Man beachte jedoch, dass selbst $P(\text{Variation}(X)_t < \infty) = 1$ *i.A.* **nicht** $E(\text{Variation}(X)_t) < \infty$ *nach sich zieht, so dass (4.26) die Zerlegung in (4.23) nicht ersetzen kann.*

Bemerkung 4.29. *Das Ereignis* $\left\{\int_{0\leq s\leq t,\ x\in\mathbb{R}^d} |x|J_X(ds\times dx) = \infty\right\} \in \mathcal{F}_t$ *ist ein 0-1-Ereignis. D.h. die Sprünge eines Lévy-Prozesses erzeugen entweder mit Wahrscheinlichkeit* 1 *unendliche Variation oder sie besitzen mit Wahrscheinlichkeit* 1 *endliche Variation. Dies besagt schon Proposition 4.27. Man kann dies aber auch einfacher sehen. Definiere dazu für festes* $t \in \mathbb{R}_+$ *die Folge von* σ*-Algebren*

$$\sigma_n := \sigma\left((\Delta X_s)1_{\{1/(n+1)<|\Delta X_s|\leq 1/n\}}\ 0\leq s\leq t\right), \quad n\in\mathbb{N}.$$

σ_n *enthält alle Informationen über Sprünge von* X *auf* $[0,t]$*, die betragsmäßig größer als* $1/(n+1)$ *und kleiner als* $1/n$ *sind. Das Ereignis* $\left\{\int_{0\leq s\leq t,\ x\in\mathbb{R}^d} |x|J_X(ds\times dx) = \infty\right\}$ *ist bzgl. der Folge* $(\sigma_n)_{n\in\mathbb{N}}$ *ein* **terminales Ereignis***, d.h.*

$$\left\{\int_{0\leq s\leq t,\ x\in\mathbb{R}^d} |x|J_X(ds\times dx) = \infty\right\} \in \bigcap_{n\in\mathbb{N}} \sigma\left(\bigcup_{m\geq n}\sigma_m\right).$$

Sei $n\in\mathbb{N}$*. Da es auf jedem Pfad nur endlich viele Sprünge größer als* $1/n$ *gibt, sind diese Sprünge für die Frage, ob der Pfad endliche oder unendliche Variation besitzt, nicht relevant. Hat man also für irgendein* $n\in\mathbb{N}$ *die Information über alle* σ_m *mit* $m\geq n$*, dann weiß man, ob der Pfad endliche Variation besitzt oder nicht. Also ist das Ereignis, dass der Pfad unendliche Variation besitzt, für alle* $n\in\mathbb{N}$ *in der* σ*-Algebra* $\sigma\left(\bigcup_{m\geq n}\sigma_m\right)$ *enthalten und damit definitionsgemäß auch in* $\bigcap_{n\in\mathbb{N}}\sigma\left(\bigcup_{m\geq n}\sigma_m\right)$*.*

Zudem ist das Ereignis wegen Lemma 4.22 von allen σ_n *stochastisch unabhängig. Damit folgt mit Kolmogorovs 0-1-Gesetz, dass*

$$P\left(\int_{0\leq s\leq t,\ x\in\mathbb{R}^d} |x|J_X(ds\times dx) = \infty\right) \in \{0,1\}.$$

Proposition 4.30. *Sei* $n\in\mathbb{N}$ *und* X *ein reellwertiger Lévy-Prozess mit charakteristischem Tripel* (γ, A, ν)*. Das* n*-te absolute Moment von* X_t*, also* $E(|X_t|^n)$*, ist genau dann endlich für ein* $t>0$ *oder äquivalent für alle* $t>0$*, wenn*

$$\int_{|x|\geq 1} |x|^n\nu(dx) < \infty$$

Beweis: siehe Beweis von Theorem 25.3 in Sato [10].

4.2 Modellierung von Aktienpreisen mit Lévy-Prozessen

Modelliere Aktienpreisprozess $S = (S_t)_{t\geq 0}$ durch

$$S_t = S_0 \exp(L_t), \quad t \geq 0,$$

wobei L ein Lévy-Prozess ist.

Theorem 4.31. *Sei L ein Lévy-Prozess. Es gilt*

$$(\exp(L_t))_{t\geq 0} \text{ ist ein lokales Martingal} \quad \Leftrightarrow \quad (\exp(L_t))_{t\geq 0} \text{ ist ein Martingal.}$$

Proof. Da jedes Martingal auch ein lokales Martingal ist, muss nur "$\Rightarrow$" gezeigt werden.

Sei also L ein Lévy-Prozess und $\exp(L)$ ein lokales Martingal. Es reicht zu zeigen, dass für alle $T \in \mathbb{R}_+$ der bei T abgestoppte Prozess ein Martingal ist. O.B.d.A. $T = 1$. Da $\exp(L)$ nichtnegativ ist, folgt, dass $\exp(L)$ ein echtes Supermartingal ist, insbesondere sind alle $\exp(L_t)$ integrierbar mit

$$E\left(\exp(L_t)\right) \leq 1.$$

Da $\exp(L_1) > 0$ folgt

$$a := \ln\left(\frac{1}{E(\exp(L_1))}\right) \in \mathbb{R}_+$$

und damit

$$E\left(\exp(L_1 + a)\right) = 1.$$

Wegen obiger Integrierbarkeitsbedingung gilt (ohne Beweis, siehe Theorem 25.17 in Sato [10])

$$\begin{aligned} E\left(\exp(L_t + at)\right) &\overset{\text{ohne Beweis}}{=} \exp\left(t\left(\gamma + a + \sigma^2 + \int_{\mathbb{R}} (e^x - 1 - x1_{\{|x|\leq 1\}})\nu(dx)\right)\right) \\ &= \left(\exp\left(\gamma + a + \sigma^2 + \int_{\mathbb{R}} (e^x - 1 - x1_{\{|x|\leq 1\}})\nu(dx)\right)\right)^t \\ &= \left(E\left(\exp(L_1 + a)\right)\right)^t \\ &= 1, \quad \forall t \in [0,1]. \end{aligned}$$

Damit folgt

$$\begin{aligned} E\left(\exp(L_1 + a) \mid \mathcal{F}_t\right) &= \exp(L_t + at) E\left(\exp(L_1 - L_t + a(1-t)) \mid \mathcal{F}_t\right) \\ &= \exp(L_t + at) E\left(\exp(L_1 - L_t + a(1-t))\right) = \exp(L_t + at). \end{aligned}$$

Also ist der Prozess $(\exp(L_t + at))_{t\in[0,1]}$ ein Martingal und damit von Klasse (D). Wegen

$$\exp(L_t) \leq \exp(L_t + at), \quad \forall t \in [0,1],$$

ist auch $(\exp(L_t))_{t\in[0,1]}$ von Klasse (D). Da jedes lokale Martingal von Klasse (D) ein Martingal ist, folgt die Behauptung. $\square$

Beispiel 4.32 (**Merton Modell**).

$$L_t = \mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$$

wobei $Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$. *Die charakteristische Funktion von* L_1 *ist*

$$\Psi_{L_1}(u) = \exp\left(i\gamma u - \frac{1}{2}\sigma^2 u^2 + \lambda\left(e^{i\mu_Y u - \sigma_Y^2 u^2/2} - 1\right)\right).$$

Es gilt

$$E(L_1) = \mu + \lambda \mu_Y$$

und

$$\mathrm{Var}(L_1) = \sigma^2 + \lambda \mu_Y^2 + \lambda \sigma_Y^2.$$

Beispiel 4.33 (**Kou Modell**).

$$L_t = \mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i,$$

wobei Y_i *„doppelt exponentialverteilt" mit Parametern* $p \in [0,1]$, $\theta_1, \theta_2 > 0$ *ist, also die Dichtefunktion*

$$f_Y(x) = \begin{cases} p\theta_1 e^{-\theta_1 x} & : \textit{für } x \geq 0 \\ (1-p)\theta_2 e^{\theta_2 x} & : \textit{für } x < 0. \end{cases}$$

besitzt. Die charakteristische Funktion von L_1 *ist*

$$\Psi_{L_1}(u) = \exp\left(i\gamma u - \frac{1}{2}\sigma^2 u^2 + \lambda\left(\frac{p}{\theta_1 - iu} - \frac{1-p}{\theta_2 + iu}\right)\right).$$

Es gilt

$$E(L_1) = \mu + \frac{\lambda p}{\theta_1} - \frac{\lambda(1-p)}{\theta_2}$$

und

$$\mathrm{Var}(L_1) = \sigma^2 + \frac{\lambda p}{\theta_1^2} - \frac{\lambda(1-p)}{\theta_2^2}.$$

To be continued

Literatur

[1] Cont, R. and Tankov, P. (2004) Financial Modelling with Jump Processes. *Chapman & Hall/CRC.*

[2] Fleming, W.H., Soner, H.M. (2006) Controlled Markov Processes and Viscosity Solutions. *Springer-Verlag*, 2. Auflage.

[3] Kallenberg, O. (1997) Foundations of Modern Probability. *Springer-Verlag.*

[4] Klenke, A. (2008) Wahrscheinlichkeitstheorie. *Springer-Verlag*, 2. Auflage.

[5] Korn, R. (1997) Optimal Portfolios – Stochastic Models for Optimal Investment and Risk Management in Continuous Time. *World Scientific.*

[6] Kühn, C. Vorlesungsskript "Stochastische Analysis mit Finanzmathematik". *http://ismi.math.uni-frankfurt.de/kuehn/.*

[7] Kyprianou, A.E. (2006) Introductory Lectures on Fluctuations of Lévy Processes with Applications. *Springer-Verlag.*

[8] Øksendal, B. and Sulem, A. (2007) Applied Stochastic Control of Jump Diffusions. *Springer-Verlag*, 2. Auflage.

[9] Protter, P. (2004) Stochastic Integration and Differential Equations. *Springer-Verlag*, 2. Auflage.

[10] Sato, K.-I. (1999) Lévy Processes and Infinitely Divisible Distributions. *Cambridge studies in advanced mathematics.*

[11] Touzi, N. (2002) Stochastic Control Problems, Viscosity Solutions, and Applications to Finance. *Special Research Semester, Lecture Notes, Pisa.*

[12] Pham, H. (2009) Continuous-time Stochastic Control and Optimzation with Financial Applications. *Springer-Verlag.*

[13] Yong, J. and Zhou, X.Y. (1999) Stochastic Controls – Hamiltonian Systems and HJB Equations. *Springer-Verlag.*