Über die Asymptotik des Bayesrisikos bei Diffusionsmodellen

Jochen Voß

Dritte, verbesserte Auflage der Diplomarbeit von Jochen Voß.
Diese wurde unter demselben Titel im Sommer 1997 fertiggestellt.

http://www.mathematik.uni-kl.de/~wwwstoch/voss/

Inhalt

Inhalt

.

Einleitung

In dieser Arbeit untersuche ich Eigenschaften von Diffusionsprozessen. Unter einem Diffusionsprozeß verstehe ich hierbei ein mathematisches Modell für die Bewegung eines sehr kleinen Teilchens in einem Gas. Ein solches Teilchen wechselt durch Stöße mit Gasmolekülen ständig seine Geschwindigkeit; zwischen den Stößen fliegt es nahezu geradlinig. Im mathematischen Modell wird auf die Behandlung der stoßenden Gasmoleküle verzichtet und stattdessen angenommen, das Teilchen folge einer zufälligen Bahn. Die Bewegung ist dadurch idealisiert, daß hier ständige Richtungsänderungen angenommen werden, die Phasen geradlinigen Fluges kommen nicht mehr vor. Das Modell wird also durch eine spezielle Wahrscheinlichkeitsverteilung auf der Menge der möglichen Bahnkurven charakterisiert. Die ständigen Richtungsänderungen schlagen sich dann beispielsweise darin nieder, daß die Wahrscheinlichkeit für irgendwo differenzierbare Bahnkurven Null beträgt.

Strömt nun das Gas in eine Richtung, so wird sich das Partikel trotz seiner unregelmäßigen Zitterbewegung im Laufe der Zeit in dieselbe Richtung bewegen. Durch Beobachtung dieser Drift kann man von der Bahn des Partikels auf die Strömung des Gases zurückschließen. In dem von mir betrachteten Fall stehen nur zwei mögliche Strömungsrichtungen zur Auswahl. Das Problem ist dann folgendes: man beobachtet das Partikel über eine bestimmte Zeit und versucht dann aufgrund der beobachteten Bahn zu „erraten“, welche der beiden möglichen Strömungsrichtungen vorlag. Dieses „Raten“ erfolgt durch einen statistischen Test. Nimmt man die Auswahl der Strömungsrichtung vor der Beobachtung als zufällig an, so kann man die Wahrscheinlichkeit bestimmen, daß der Test das falsche Ergebnis liefert. Diese Wahrscheinlichkeit heißt Bayesrisiko.

Es ist plausibel, daß man um so mehr Information über die Strömung gewinnt, je länger man das Partikel beobachtet. Es stellt sich heraus, daß auch das Bayesrisiko gegen Null geht. Aus der Theorie der großen Abweichungen ergibt sich in dieser Arbeit, daß hier das Bayesrisiko mit der Zeit t sogar exponentiell schnell gegen Null geht, daß es sich also schließlich wie die Exponentialfunktion $e^{-c \cdot t}$ für eine positive Konstante c verhält. Das Negative dieser Konstante heißt dann Konvergenzrate des Bayesrisikos. Mein Ziel ist es nun, für bestimmte Situationen diese Konvergenzrate zu bestimmen.

Dabei gehe ich folgendermaßen vor.

Im ersten Kapitel präzisiere ich den Begriff „Diffusionsprozeß“ und liste einige Eigenschaften solcher Prozesse auf, die für das Folgende nützlich sind. Dieser Teil ist im wesentlichen dem Buch von v. Weizsäcker und Winkler (1990) entnommen. Aus den Lösungen der stochastischen Differentialgleichung

$$dX = b(t, X_t) \bullet dt + \sigma(t, X_t) \bullet dB$$

erhält man Diffusionsprozesse. In der eingangs beschriebenen Situation entspricht hier X_t dem Ort des Partikels zur Zeit t, der Term $b(X_t, t)$ entspricht der Drift, der das Teilchen unterworfen ist, und die Diffusionsmatrix $\sigma(X_t, t)$ beschreibt die Stärke der Fluktuation (die z.B. von der Temperatur des umgebenden Mediums abhängen könnte).

Indem man die Situation auf zeitunabhängige Drift und Diffusionsmatrix einschränkt, kann man zu dem Prozeß als zusätzliches Hilfsmittel eine Operatorhalbgruppe auf dem Banachraum $C_b(\mathbb{R}^d)$ konstruieren.

Im letzten Teil des Kapitels zeige ich, wie man die Girsanov-Formel verwenden kann, um die Dichte der Verteilung der Pfade eines Diffusionsprozesses bezüglich des Wienermaßes auf $C\big([0;\infty), \mathbb{R}^d\big)$ zu bestimmen.

Im zweiten Kapitel beweise ich einen Satz aus der Theorie der großen Abweichungen. Satz und Beweis stammen aus dem Buch von Varadhan (1980). Der Satz ermöglicht es, für gewisse Familien von Ereignissen, deren Wahrscheinlichkeit gegen Null konvergiert, die Geschwindigkeit dieser Konvergenz zu beschreiben. Man erhält ein Ergebnis der Form

$$\begin{aligned} -\inf_{x\in A^\circ} \mathrm{I}(x) \;&\le\; \liminf_{\varepsilon\downarrow 0} \varepsilon\cdot\log P(\frac{A}{\sqrt{\varepsilon}}) \\ &\le\; \limsup_{\varepsilon\downarrow 0} \varepsilon\cdot\log P(\frac{A}{\sqrt{\varepsilon}}) \;\le\; -\inf_{x\in\overline{A}} \mathrm{I}(x), \end{aligned}$$

wobei P ein Wahrscheinlichkeitsmaß und A ein Ereignis ist. Die Ratenfunktion I ist dort klein, wo das Maß P konzentriert ist, und wächst „nach außen". Im Satz wird als Wahrscheinlichkeitsmaß das Wienermaß auf $C\big([0;1],\mathbb{R}^d\big)$ verwendet.

Anschließend wird noch das sogenannte Varadhansche Integrallemma bewiesen, das man immer dann anwenden kann, wenn man für die Konvergenzraten eine Abschätzung wie die obige hat. Es bringt das Ergebnis in eine Form, in der ich es später verwenden werde.

Im dritten Kapitel löse ich das eingangs beschriebene Problem für Prozesse, die der stochastischen Differentialgleichung

$$dX_t \;=\; b\bullet dt + dB_t$$

genügen. Dies sind Diffusionsprozesse mit orts- und zeitunabhängiger Drift b. Wie bereits angedeutet stellt sich heraus, daß das Bayesrisiko für die Unterscheidung zweier solcher Prozesse exponentiell schnell fällt. Die Konvergenzrate ist proportional zum Quadrat der Differenz beider Driftrichtungen. Insbesondere bestätigt das die naheliegende Vermutung, daß sich — durch Beobachtung der zugehörigen Prozesse — zwei Driftrichtungen um so besser unterscheiden lassen, je unterschiedlicher sie sind.

Ich gebe zwei Methoden zur Bestimmung der Konvergenzrate an. Die erste Methode führt das Problem auf eine Aussage über die Normalverteilung auf dem $\mathbb{R}^d$ zurück. Die Behauptung ergibt sich dann z.B. als einfache Folgerung aus dem Satz von Varadhan, wie er im zweiten Kapitel steht. Die andere Methode verwendet das Varadhansche Integrallemma zur Bestimmung der Konvergenzrate.

Im zweiten Teil des Textes stelle ich Begriffe und Methoden vor, um dieses Problem für andere Klassen von Diffusionsprozessen anzugehen. Zunächst werden dazu, im vierten Kapitel, Diffusionsprozesse betrachtet, die eine stationäre Verteilung μ besitzen, für die also

$$\mathcal{L}(X_t) \;=\; \mu \qquad \text{für alle } t\ge 0$$

gilt. Solche Prozesse lassen sich durch die zugehörige Operatorhalbgruppe und deren Generator, charakterisieren. Die Operatorhalbgruppe, wie sie im ersten Kapitel eingeführt wurde, besitzt für Prozesse mit stationärer Verteilung eine Fortsetzung zu einer Halbgruppe auf dem Hilbert-Raum $L^2(\mathbb{R}^d,\mu)$.

Weiter beweise ich, daß aus der Existenz einer stationären Verteilung unter gewissen Voraussetzungen auch Stationarität folgt. Das Verhalten des Prozesses ist dann invariant unter Zeitverschiebungen. Im letzten Teil des Kapitels zeige ich, daß man einen stationären Prozeß, der zur Zeit Null startet, zu einem Prozeß mit Indexmenge $\mathbb{R}$ fortsetzen kann.

Das fünfte Kapitel beschreibt Eigenschaften von reversiblen Diffusionsprozessen. Dies sind Prozesse, deren Verhalten sich unter Zeitumkehr nicht ändert; insbesondere sind solche Prozesse stationär. Ich gebe mehrere unterschiedliche Beschreibungen für Reversibilität an.

Der letzte Teil dieses Kapitels zitiert einen Satz über große Abweichungen der empirischen Verteilung solcher Prozesse.

Im sechsten Kapitel stelle ich einen Zusammenhang zwischen stationären Lösungen der Schrödingergleichung und reversiblen Diffusionsprozessen her. Die Rechnungen diese Kapitels stammen großteils aus einem Artikel von S. K. Mitter (1986). Die Schrödingergleichung hat ihren Ursprung in der Quantenmechanik und beschreibt, wie sich die Aufenthaltswahrscheinlichkeiten eines Partikels mit der Zeit verändern. Stationäre Lösungen sind solche, bei denen diese Wahrscheinlichkeiten zeitlich konstant sind. Einer solchen Lösung ordnet man einen reversiblen Diffusionsprozeß X zu, der dieselben Aufenthaltswahrscheinlichkeiten besitzt; es gilt dann

$$\int_A \psi_t^2(x)\,dx \;=\; P(X_t \in A),$$

wobei ψ_t^2 die Dichte der Aufenthaltswahrscheinlichkeit aus der Schrödingergleichung ist.

Anschließend zeige ich, wie man, diesmal von einem Diffusionsprozeß ausgehend, wieder das quantenmechanische System zurückgewinnt.

Die Ergebnisse dieses Kapitels erweisen sich im auch im Folgenden als nützlich: Man erhält einen Zusammenhang zwischen dem Hamiltonoperator aus der Schrödingergleichung und dem Generator der Halbgruppe des Diffusionsprozesses. Diese Beziehung erleichtert es, die Ratenfunktion für den Satz über große Abweichungen aus dem fünften Kapitel auszuwerten.

Das siebte Kapitel illustriert die bisherigen Ergebnisse am Beispiel des Ornstein-Uhlenbeck-Prozesses. Dieser löst die stochastische Differentialgleichung

$$dX \;=\; -\alpha X \bullet dt + dB$$

für eine Zahl $\alpha > 0$. Da die Drift $-\alpha X$ vom jeweiligen Ort des Prozesses aus stets zum Ursprung hin gerichtet ist, wird der Prozeß immer wieder zur Nullage hin gezogen; der Fluktuationsterm dB verhindert allerdings, daß er dort zur Ruhe kommt. Es stellt sich heraus, daß dieser Prozeß reversibel ist und als stationäre Verteilung eine Normalverteilung um den Ursprung besitzt.

Die Ergebnisse aus Kapitel 6 liefern einen Zusammenhang zum harmonischen Oszillator, wie er aus der Physik bekannt ist. Den Abschluß dieser Arbeit bildet eine Teillösung des Problems, die Konvergenzrate des Bayesrisikos für diesen Prozeß zu bestimmen.

.

Zur Notation

Im Folgenden fasse ich kurz die Bedeutungen einiger Symbole zusammen, die in der Arbeit verwendet werden. Die meisten Symbole werden auch im eigentlichen Text definiert.

$\mathcal{B}(X)$	bezeichnet die Borel-σ-Algebra auf einem topologischen Raum X.
$C_b(\mathbb{R}^d)$	ist eine Abkürzung für $C_b(\mathbb{R}^d, \mathbb{R})$ und bezeichnet die beschränkten, stetigen Abbildungen von $\mathbb{R}^d$ nach $\mathbb{R}$.
$C_c(\mathbb{R}^d)$	bezeichnet die stetigen Abbildungen von $\mathbb{R}^d$ nach $\mathbb{R}$, die einen kompakten Träger besitzen.
$C_c^n(\mathbb{R}^d)$	für $n \in \mathbb{N}$ bezeichnet die stetigen Abbildungen mit kompaktem Träger von $\mathbb{R}^d$ nach $\mathbb{R}$, die darüberhinaus stetige n-te Ableitungen besitzen.
$C_0^\infty(\mathbb{R}^d)$	bezeichnet die beliebig oft differenzierbaren, stetigen Abbildungen von $\mathbb{R}^d$ nach $\mathbb{R}$, die die zusammen mit allen ihren Ableitungen für $\lvert x\rvert \to \infty$ gegen Null konvergieren.
$\mathcal{D}(L)$	bezeichnet den Definitionsbereich des linearen Operators L.
$L^p_{\mathrm{loc}}(\mathbb{R}^d)$	ist die Menge aller meßbaren Abbildungen von $\mathbb{R}^d$ nach R, die lokal in $L^p(\mathbb{R}^d)$ liegen, d.h. genau dann ist $f \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$, wenn $1_K \cdot f \in L^p(\mathbb{R}^d)$ für jede kompakte Menge $K \subseteq \mathbb{R}^d$ gilt.
$L^2(\mu)$	ist eine Abkürzung für $L^2(\mathbb{R}^d, \mu)$, wobei μ ein Wahrscheinlichkeitsmaß ist, und wird vor Satz 4.1 eingeführt.
$(\,\cdot\,,\,\cdot\,)_\mu$	bezeichnet das Skalarprodukt auf $L^2(\mu)$.
$M^\complement$	bezeichnet gelegentlich das Komplement einer Menge M. Die gemeinsame Obermenge geht dann aus dem Kontext hervor.
$\mathcal{M}_1(X)$	bezeichnet die Menge alle Wahrscheinlichkeitsmaße auf X.
$\mathbb{R}_+$	steht für $[0;\infty)$ und enthält bei mir also die Null.
$S(x,\varepsilon)$	bezeichnet die offene ε-Kugel um x.
W	bezeichnet den Wienerprozeß, d.h. für alle $\omega \in C\big([0;\infty), \mathbb{R}^d\big)$ und alle $t \geq 0$ gilt $W_t(\omega) = \omega_t$.
$\mathbb{W}$	ist das Wienermaß auf $C\big([0;\infty), \mathbb{R}^d\big)$. Bezüglich $\mathbb{W}$ ist der Wienerprozeß W eine Brownsche Bewegung.

1. Diffusionsprozesse

In diesem Kapitel führe ich den Begriff des Diffusionsprozesses ein und liste diejenigen Eigenschaften solcher Prozesse auf, die ich im Folgenden verwenden werde. Die Darstellung ist an Kapitel 12 des Buches von v. Weizsäcker und Winkler (1990) angelehnt.

Definition. Es bezeichne S_+^d die Menge der symmetrischen, nicht-negativ definiten $d \times d$-Matrizen. Seien $b: \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ und $a: \mathbb{R}_+ \times \mathbb{R}^d \to S_+^d$ meßbar und auf kompakten Teilmengen von $\mathbb{R}_+ \times \mathbb{R}^d$ beschränkt. Dann definiere den Differentialoperator L_t durch

$$L_t\varphi(x) \;=\; \sum_{i=1}^{d} b_i(t,x)\frac{\partial\varphi}{\partial x_i}(x) + {}^1\!/_2 \sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial\varphi^2}{\partial x_i \partial x_j}(x).$$

für alle $\varphi \in T := C_c^\infty(\mathbb{R}^d)$.

Sei nun $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$. Dann heißt $X: [0;\infty) \times \Omega \to \mathbb{R}^d$ ein **Diffusionsprozeß**, falls folgende Bedingungen erfüllt sind:

(j) X ist ein $(\mathcal{F}_t)$-adaptierter, f.s. stetiger stochastischer Prozeß.

(ij) Es gilt

$$\lim_{h\downarrow 0} \frac{1}{h} E\big(\varphi(X_{t+h}) - \varphi(X_t) \;\big|\; \mathcal{F}_t\big) \;=\; L\varphi(X_t) \qquad \text{f.s.}$$

für alle $t \geq 0$ und alle $\varphi \in T$.

(iij) X besitzt die starke Markoveigenschaft, d.h. für jede endliche Stoppzeit S und jede beschränkte, $\sigma(X_{S+h} \mid h \geq 0)$-meßbare Funktion $\Phi: \Omega \to \mathbb{R}$ gilt

$$E(\Phi|\mathcal{F}_S^+) = E(\Phi|X_S).$$

b heißt **Drift** und a **Diffusionsmatrix** von X. Die Elemente des Raumes T heißen **Testfunktionen**.

Der nächste Satz (v. Weizsäcker und Winkler (1990), Theorem 12.1.8) besagt, daß sich Diffusionsprozesse zum Beispiel als Lösungen stochastischer Differentialgleichungen konstruieren lassen.

Satz 1.1 *Seien $b: \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ und $\sigma: \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n}$ stetige Abbildungen, die folgende Lipschitz-Bedingung erfüllen: Für jedes $t^* < \infty$ gibt es ein $c \in \mathbb{R}$, so daß*

$$\left.\begin{aligned} |b(t,x) - b(t,x')| &\leq c \cdot |x - x'| \\ |\sigma(t,x) - \sigma(t,x')| &\leq c \cdot |x - x'| \end{aligned}\right. \qquad \text{für alle } t \in [0;t^*]$$

gilt. $B = (B^1, \ldots, B^n)$ sei eine n-dimensionale Brownsche Bewegung.

Dann hat das Anfangswertproblem zu der stochastischen Differentialgleichung

$$dX \;=\; b(t,X_t) \bullet dt + \sigma(t,X_t) \bullet dB$$

für jede Anfangsbedingung X_0 eine (pfadweise, f.s.) eindeutig bestimme Lösung $X: \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$. Falls X_0 zweite Momente besitzt, ist dieser Prozeß ein Diffusionsprozeß mit Drift b und Diffusionsmatrix $\sigma\sigma^T$.

Aus der pfadweisen Eindeutigkeit der Lösung der stochastischen Differentialgleichung im letzten Satz folgt auch die schwache Eindeutigkeit (a.a.O., Theorem 12.4.4):

Satz (Yamada, Watanabe). *Sind die Drift $b: \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ und die Diffusionsmatrix $\sigma: \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n}$ der stochastischen Differentialgleichung*

$$dX \;=\; b(t,X_t) \bullet dt + \sigma(t,X_t) \bullet dB$$

so gewählt, daß es für jede Anfangsbedingung X_0 und jede $(\mathcal{F}_t)$-Brownsche Bewegung eine (pfadweise) eindeutig bestimmte Lösung gibt, so ist die Lösung **schwach eindeutig**, *d.h. je zwei Lösungen mit derselben Anfangsverteilung $\mathcal{L}(X_0)$ haben auf $C\big([0;\infty), \mathbb{R}^d\big)$ dieselbe Verteilung.*

Das folgende Lemma (a.a.O., Theorem 11.2.3) ermöglicht es, für $p \geq 2$ die p-ten Momente der Lösung einer stochastischen Differentialgleichung abzuschätzen.

Notation. $\mathcal{H}^d$ bezeichne den Raum der $\mathbb{R}^d$-wertigen, adaptierten, f.s. regulären und f.s. linksstetigen Prozesse.

Lemma 1.1 *Seien S, T Stoppzeiten mit $S \leq T$ und $F: \mathcal{H}^d \to \mathcal{H}^{d \times n}$ eine Abbildung. Weiter sei Y ein Semimartingal mit Werten in $\mathbb{R}^n$ und $Y = M + A$ für ein rechtsstetiges lokales L^2-Martingal M und einen rechtsstetigen, adaptierten Prozeß A von lokal endlicher Variation mit $A_0 = 0$. Es gelte*

$$\sum_{i=1}^{n} \int_S^T \left(d\langle M^i \rangle + |dA^i| \right) \leq \tau .$$

Zuletzt sei $p \geq 2$. Dann gilt folgendes:

(j) *Existiert ein $\kappa \in \mathbb{R}$, so daß F auf $]S;T]$ die Wachstumsbedingung*

$$\left| F(H)_t(\omega) \right| \leq \kappa \cdot \left(1 + \sup_{s \leq t} |H_s(\omega)| \right) \quad \textit{für alle } H \in \mathcal{H}^d$$

erfüllt, so gibt es ein $\gamma \in \mathbb{R}$ mit

$$E\left(\sup_{t \leq T} |X_t|^p \right) \leq \gamma \left(1 + E\left(\sup_{s \leq S} |X_s|^p \right) \right)$$

für jeden Prozeß X, der auf dem Intervall $]S;T]$ die stochastische Differentialgleichung $dX = F(X) \bullet dY$ löst.

(ij) *Existiert ein $\kappa \in \mathbb{R}$, so daß F auf $]S;T]$ die Bedingung*

$$\left| F(H)_t(\omega) - F(H')_t(\omega) \right| \leq \kappa \cdot \sup_{s \leq t} \left| H_s(\omega) - H'_s(\omega) \right| \quad \textit{für alle } H, H' \in \mathcal{H}^d$$

erfüllt, so gibt es ein $\gamma' \in \mathbb{R}$ mit

$$E\left(\sup_{t \leq T} |X - X'|_t^p \right) \leq \gamma' E\left(\sup_{s \leq S} |X - X'|_s^p \right)$$

für alle Prozesse X, X', die auf dem Intervall $]S;T]$ die stochastische Differentialgleichung $dX = F(X) \bullet dY$ lösen.

Ich betrachte ab jetzt Lösungen der zeithomogenen Gleichung

$$dX \;=\; b(X_t) \bullet dt + \sigma(X_t) \bullet dB. \tag{1.1}$$

Diese stochastische Differentialgleichung kann man mit

$$(FH)_t \;=\; \begin{pmatrix} b_1(H_t) & \sigma_{1,1}(H_t) & \cdots & \sigma_{1,n}(H_t) \\ \vdots & \vdots & \ddots & \vdots \\ b_d(H_t) & \sigma_{d,1}(H_t) & \cdots & \sigma_{d,n}(H_t) \end{pmatrix}$$

und

$$Y_t \;=\; \begin{pmatrix} t \\ B_t^1 \\ \vdots \\ B_t^n \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ B_t^1 \\ \vdots \\ B_t^n \end{pmatrix} + \begin{pmatrix} t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \;=:\; M + A$$

als $dX = F(X) \bullet dY$ schreiben. Sei $S = 0$ und $T = t \in \mathbb{R}_+$. Da hier $d\langle M^j \rangle = |dA^0| = dt$ gilt, erfüllen dann M und A für $\tau = (n+1) \cdot t$ die Bedingung aus Lemma 1.1.

Sind b und σ lipschitzstetig, so existiert ein $\kappa \in \mathbb{R}$ mit

$$\big|b_i(x)\big| \leq \big|b_i(x) - b_i(0)\big| + \big|b_i(0)\big| \leq \kappa \cdot \big(|x| + 1\big)$$

und genauso

$$\big|\sigma_{ij}(x)\big| \leq \kappa \cdot \big(|x| + 1\big)$$

für $i \in \{1, \ldots, d\}$ und $j \in \{1, \ldots, n\}$. Verwendet man auf $\mathbb{R}^{d \cdot (n+1)}$ die Maximumsnorm, so erhält man

$$\big|F(H)_t(\omega)\big| \;\leq\; \max_{\substack{i:=1,\ldots,d \\ j:=1,\ldots,n}} \big\{b_i(H_t(\omega)), \sigma_{ij}(H_t(\omega))\big\} \;\leq\; \kappa \cdot \big(|H_t(\omega)| + 1\big).$$

Also erfüllt F die Wachstumsbedingung aus Teil (j) des Lemmas und man erhält

Satz 1.2 *Seien $b\colon \mathbb{R}^d \to \mathbb{R}^d$ und $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d \times n}$ lipschitzstetig, $t \in \mathbb{R}_+$ und $p \geq 2$. Dann gibt es ein $\gamma \in \mathbb{R}$ mit*

$$E\big(\sup_{s \leq t} |X_s|^p\big) \leq \gamma\big(1 + E(|X_0|^p)\big)$$

für jeden Prozeß X, der Gleichung (1.1) *löst.*

Setze $a(x) := \sigma(x)\sigma(x)^T$. Durch direktes Nachrechnen (a.a.O., Lemma 12.1.1) erhält man

$$dX^i \bullet dX^j \;=\; a_{ij}(X) \bullet dt.$$

Ist X eine Lösung von (1.1), so ist X nach Satz 1.1 unter geeigneten Voraussetzungen ein Diffusionsprozeß. Der zugehörige Differentialoperator L hat die Form

$$L\varphi(x) \;=\; \sum_{i=1}^{d} b_i(x)\partial_i\varphi(x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial_i\partial_j\varphi(x). \tag{1.2}$$

Für beliebige Lösungen X von (1.1) und $\varphi \in C^2(\mathbb{R}^d)$ erhält man aus der Itô-Formel durch einfaches Nachrechnen die Beziehung

$$d\varphi(X) - L\varphi(X) \bullet dt \;=\; \sum_{i=1}^{d}\sum_{j=1}^{n} \partial_i\varphi(X)\sigma_{ij}(X) \bullet dB^i.$$

Als Summe von stochastischen Integralen ist die rechte Seite in dieser Gleichung ein f.s. stetiges lokales Martingal. Unter geeigneten Bedingungen ist sie — und damit auch die linke Seite der Gleichung — sogar ein Martingal. Der nächste Satz (a.a.O., Theorem 12.1.4) gibt dafür ein hinreichendes Kriterium an.

Satz 1.3 *Sei X eine Lösung der stochastischen Differentialgleichung* (1.1) *und $\varphi \in C^2(\mathbb{R}^d)$. Weiter gelte*

$$E\Big(\int_0^t \big(\partial_i\varphi(X_s)\sigma_{ij}(X_s)\big)^2\, ds\Big) \;<\; \infty$$

für alle $i \in \{1, \ldots, d\}$, $j \in \{1, \ldots, n\}$ und $t \geq 0$. Dann ist der Prozeß

$$\varphi(X_\cdot) - \varphi(X_0) - \int_0^\cdot L\varphi(X_s)\, ds$$

ein Martingal. Insbesondere gilt in diesem Fall

$$E\big(\varphi(X_{t+h}) - \varphi(X_t) \bigm| \mathcal{F}_t\big) \;=\; E\Big(\int_t^{t+h} L\varphi(X_s)\, ds \bigm| \mathcal{F}_t\Big).$$

Mit der letzten Gleichung kann man für spezielle φ beispielsweise Eigenschaften, wie in Bedingung (ij) aus der Definition von „Diffusionsprozeß“ nachweisen. Stellt man geeignete Forderungen an b, σ und X_0, so gelingt dies — wie im nächsten Satz — auch für Funktionen φ, die keinen kompakten Träger besitzen. Zunächst beweise ich aber noch ein Lemma.

Lemma 1.2 *Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $\mathcal{G} \subseteq \mathcal{F}$ eine Sub-σ-algebra von $\mathcal{F}$ und $f: \Omega \to \mathbb{R}$ eine $\mathcal{G}$-meßbare Funktion. $f_h: \Omega \to \mathbb{R}$ für $h > 0$ seien Funktionen mit $f_h \to f$ in $L^2(\Omega)$ für $h \downarrow 0$. Dann gilt auch*

$$E(f_h \mid \mathcal{G}) \xrightarrow{h \downarrow 0} f \qquad \text{in } L^2(\Omega).$$

Beweis. Es gilt

$$\begin{aligned} E\Big(\big| E(f_h \mid \mathcal{G}) - f \big|^2\Big) &= E\Big(\big| E(f_h - f \mid \mathcal{G}) \big|^2\Big) \\ &\overset{\text{Jensen}}{\leq} E\Big(E\big(|f_h - f|^2 \mid \mathcal{G}\big)\Big) \\ &= E\big(|f_h - f|^2\big) \quad \xrightarrow{h\downarrow 0} \quad 0. \end{aligned}$$

(q.e.d.)

Satz 1.4 *Sei X eine Lösung von Gleichung (1.1) mit $X_0 \in L^2(\Omega)$. Die Drift b und der Diffusionskoeffizient σ seien lipschitzstetig. Für alle $t \geq 0$ gilt dann*

$$\lim_{h \downarrow 0} \frac{1}{h} E(X_{t+h} - X_t \mid \mathcal{F}_t) \;=\; b(X_t) \qquad \text{in } L^2(\Omega).$$

Weiter gilt

$$\lim_{h \downarrow 0} \frac{1}{h} E\big((X_{t+h} - X_t) f(X_t)\big) \;=\; E\big(b(X_t) f(X_t)\big) \qquad \text{für alle } f \in C_b(\mathbb{R}^d).$$

Beweis. Sei $k \in \{1, \ldots, d\}$ und $\varphi(x) := x_k$ für alle $x \in \mathbb{R}^d$. Nach Voraussetzung existiert ein $\kappa \geq 0$ mit

$$\sigma_{ij}(x) \leq \kappa \cdot \big(|x| + 1\big) \quad \text{und} \quad b_i(x) \leq \kappa \cdot \big(|x| + 1\big)$$

für alle i, j. Wegen $X_0 \in L^2(\Omega)$ liefert Satz 1.2, daß $E\big(\sup_{s \leq t} |X_s|^2\big) < \infty$ ist. Aus $\partial_i \varphi \equiv \delta_{ik}$ folgt

$$\int_0^t \big(\partial_i \varphi(X_s) \sigma_{ij}(X_s)\big)^2 \, ds \;\leq\; \int_0^t \big(\sigma_{ij}(X_s)\big)^2 \, ds \;\leq\; 4t \cdot \kappa^2 \big(\sup_{s \leq t} |X_s|^2 + 1\big),$$

also

$$E\Big(\int_0^t \big(\partial_i \varphi(X_s) \sigma_{ij}(X_s)\big)^2 \, ds\Big) \;<\; \infty.$$

Dies ist aber die Voraussetzung des Satzes 1.3, mit dem man nun

$$\frac{1}{h} E\big(X^k_{t+h} - X^k_t \mid \mathcal{F}_t\big) \;=\; E\Big(\frac{1}{h} \int_t^{t+h} b_k(X_s) \, ds \mid \mathcal{F}_t\Big)$$

erhält.

Da $s \mapsto b_k(X_s)$ stetig ist, gilt

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} b_k(X_s) \, ds \;=\; b_k(X_t) \quad \text{f.s.}$$

Für $\varepsilon > 0$ und $h < \varepsilon$ folgt mit Satz 1.2 weiter

$$\begin{aligned} E\Big(\Big|\frac{1}{h} \int_t^{t+h} b_k(X_s) \, ds\Big|^2\Big) &\leq E\big(\sup_{s \leq t+\varepsilon} |b_k(X_s)|^2\big) \\ &\leq E\big(4\kappa^2 (\sup_{s \leq t+\varepsilon} |X_s|^2 + 1)\big) \;<\; \infty \end{aligned}$$

und mit dem Satz von der dominierten Konvergenz

$$\lim_{h\downarrow 0} \frac{1}{h} \int_t^{t+h} b_k(X_s)\,ds \;=\; b_k(X_t) \qquad \text{in } L^2(\Omega).$$

Nun läßt sich Lemma 1.2 anwenden, um

$$\begin{aligned} \lim_{h\downarrow 0} \frac{1}{h} E\big(X_{t+h}^k - X_t^k \mid \mathcal{F}_t\big) \;&=\; \lim_{h\downarrow 0} E\big(\frac{1}{h}\int_t^{t+h} b_k(X_s)\,ds \mid \mathcal{F}_t\big) \\ &=\; b_k(X_t) \qquad \text{in } L^2(\Omega) \end{aligned}$$

zu erhalten. Dies ist die erste Behauptung.

Sei nun noch $f \in C_b(\mathbb{R}^d)$. Dann gilt

$$\begin{aligned} & E\Big(\big|f(X_t)\frac{1}{h}\int_t^{t+h} b_k(X_s)\,ds \;-\; f(X_t)b_k(X_t)\big|\Big) \\ &\qquad \le\; \|f\|_\infty \cdot E\Big(\big|\frac{1}{h}\int_t^{t+h} b_k(X_s)\,ds \;-\; b_k(X_t)\big|\Big) \quad \longrightarrow \quad 0 \end{aligned}$$

und damit

$$\begin{aligned} E\Big(\frac{1}{h}(X_{t+h} - X_t)f(X_t)\Big) \;&=\; E\Big(\frac{1}{h}E\big(X_{t+h} - X_t \mid \mathcal{F}_t\big)\cdot f(X_t)\Big) \\ &=\; E\Big(\frac{1}{h}E\big(\int_t^{t+h} b(X_s)\,ds \mid \mathcal{F}_t\big)\cdot f(X_t)\Big) \\ &=\; E\Big(\frac{1}{h}\int_t^{t+h} b(X_s)\,ds \cdot f(X_t)\Big) \\ &\longrightarrow \quad E\Big(b(X_t)f(X_t)\Big) \qquad \text{für } h \downarrow 0. \end{aligned}$$

Dies ist aber die zweite Behauptung des Satzes. (q.e.d.)

Nun folgt noch ein Lemma, das ich später zum Beweis des Satzes 5.2 verwenden werde.

Lemma 1.3 *Sei $f \in C_c^\infty(\mathbb{R}^d)$ und X eine Lösung der stochastischen Differentialgleichung*

$$dX \;=\; b(X) \bullet dt + dB$$

mit Anfangsbedingung $X_0 \in L^2(\Omega)$. Die Drift $b\colon \mathbb{R}^d \to \mathbb{R}^d$ sei lipschitzstetig. Dann gilt

$$\lim_{h\downarrow 0} E\Big(\frac{1}{h}\big(X_{t+h} - X_t\big)\big(f(X_{t+h}) - f(X_t)\big)\Big) \;=\; E\big(\operatorname{grad} f(X_t)\big).$$

Beweis. Nach Satz 1.1 ist X ein Diffusionsprozeß; der zugehörige Differentialoperator L ist durch

$$L\varphi(x) \;=\; b(x)\cdot \operatorname{grad}\varphi(x) + \frac{1}{2}\Delta\varphi(x) \qquad \text{für alle } x \in \mathbb{R}^d$$

definiert. Sei $k \in \{1, \ldots, d\}$. Dann ist

$$\begin{aligned} \big(X_{t+h}^k - X_t^k\big)\big(f(X_{t+h}) - f(X_t)\big) \;&=\; \big(X_{t+h}^k f(X_{t+h}) - X_t^k f(X_t)\big) \\ &\quad - X_t^k\big(f(X_{t+h}) - f(X_t)\big) \\ &\quad - \big(X_{t+h}^k - X_t^k\big) f(X_t). \end{aligned} \tag{1.3}$$

Ich betrachte zunächst die drei Hauptterme der rechten Seite von (1.3) einzeln. Zum ersten Term: Definiere $\varphi \in C_c^\infty(\mathbb{R}^d)$ durch

$$\varphi(x) \;:=\; x_k \cdot f(x) \qquad \text{für alle } x \in \mathbb{R}^d.$$

Dann ist

$$\partial_i \varphi(x) \;=\; \delta_{ik} f(x) + x_k \partial_i f(x)$$

und

$$\partial_i^2 \varphi(x) \;=\; 2\delta_{ik}\partial_i f(x) + x_k \partial_i^2 f(x),$$

also

$$\begin{aligned} L\varphi(x) \;&=\; \sum_{i=1}^{d} b_i(x)\partial_i\varphi(x) + \frac{1}{2}\partial_i^2\varphi(x) \\ &=\; b_k(x)f(x) + x_k b(x)\cdot \operatorname{grad} f(x) + \partial_k f(x) + \frac{1}{2}x_k \Delta f(x) \end{aligned}$$

für alle $x \in \mathbb{R}^d$. Nach Eigenschaft (ij) aus der Definition von „Diffusionsprozeß" gilt

$$\lim_{h\downarrow 0} \frac{1}{h} E\big(\varphi(X_{t+h}) - \varphi(X_t) \bigm| \mathcal{F}_t\big) \;=\; L\varphi(X_t) \qquad \text{f.s.}$$

und, da φ beschränkt ist, auch

$$\begin{aligned} &\lim_{h\downarrow 0} \frac{1}{h} E\big(X_{t+h}^k f(X_{t+h}) - X_t^k f(X_t)\big) \;=\; E\big(L\varphi(X_t)\big) \\ &\quad=\; E\big(b_k(X_t)f(X_t) + X_t^k b(X_t)\cdot \operatorname{grad} f(X_t) + \partial_k f(X_t) + \frac{1}{2}X_t^k \Delta f(X_t)\big). \end{aligned}$$

Für den zweiten Term aus der Rechten Seite von (1.3) erhält man folgendes: Es ist $f \in C_c^\infty(\mathbb{R}^d)$. Da b lipschitzstetig ist, folgt mit Satz 1.3, daß

$$E\big(f(X_{t+h}) - f(X_t) \bigm| \mathcal{F}_t\big) \;=\; E\Big(\int_t^{t+h} Lf(X_s)\,ds \bigm| \mathcal{F}_t\Big)$$

gilt. Weiter ist $s \mapsto X_t^k Lf(X_s)$ stetig, also

$$\lim_{h\downarrow 0} X_t^k \frac{1}{h}\int_t^{t+h} Lf(X_s)\,ds \;=\; X_t^k Lf(X_t) \qquad \text{f.s.}$$

Darüberhinaus ist Lf beschränkt und nach Satz 1.2 $E\big(|X_t^k|^2\big) < \infty$. Also gilt

$$E\Big(\,\big|X_t^k \tfrac{1}{h}\int_t^{t+h} Lf(X_s)\,ds\big|^2\Big) \;\le\; \|Lf\|_\infty^2 E\big(|X_t^k|^2\big) \;<\; \infty$$

und mit dem Satz von der majorisierten Konvergenz folgt

$$\begin{aligned} \lim_{h\downarrow 0} X_t^k \frac{1}{h}\int_t^{t+h} Lf(X_s)\,ds \;&=\; X_t^k Lf(X_t) \\ &=\; X_t^k b(X_t)\cdot \operatorname{grad} f(X_t) + \frac{1}{2}X_t^k \Delta f(X_t) \qquad \text{in } L^2(\Omega). \end{aligned}$$

Man erhält

$$\begin{aligned} \lim_{h\downarrow 0} \frac{1}{h} E\Big(X_t^k\big(f(X_{t+h}) - f(X_t)\big) \Bigm| \mathcal{F}_t\Big) \;&=\; \lim_{h\downarrow 0} \frac{1}{h} E\Big(X_t^k \int_t^{t+h} Lf(X_s)\,ds \Bigm| \mathcal{F}_t\Big) \\ &=\; E\big(X_t^k b(X_t)\cdot \operatorname{grad} f(X_t) + \frac{1}{2}X_t^k \Delta f(X_t)\big). \end{aligned}$$

Da f beschränkt ist, liefert Satz 1.4 für den dritten Term noch

$$\lim_{h\downarrow 0} \frac{1}{h} E\big((X_{t+h}^k - X_t^k) f(X_t)\big) \;=\; E\big(b_k(X_t)f(X_t)\big).$$

Insgesamt ergibt das

$$
\begin{aligned}
&\lim_{h\downarrow 0} \frac{1}{h} E\Big(\big(X_{t+h}^k - X_t^k\big)\big(f(X_{t+h}) - f(X_t)\big)\Big) \\
&\quad = \lim_{h\downarrow 0} \frac{1}{h} E\Big(\big(X_{t+h}^k f(X_{t+h}) - X_t^k f(X_t)\big)\Big) \\
&\qquad - \lim_{h\downarrow 0} \frac{1}{h} E\Big(X_t^k \big(f(X_{t+h}) - f(X_t)\big)\Big) \;-\; \lim_{h\downarrow 0} \frac{1}{h} E\Big(\big(X_{t+h}^k - X_t^k\big) f(X_t)\Big) \\
&\quad = E\Big(b_k(X_t) f(X_t) + X_t^k b(X_t) \cdot \operatorname{grad} f(X_t) + \partial_k f(X_t) + \frac{1}{2} X_t^k \Delta f(X_t)\Big) \\
&\qquad - E\Big(X_t^k b(X_t) \cdot \operatorname{grad} f(X_t) + \frac{1}{2} X_t^k \Delta f(X_t)\Big) \;-\; E\Big(b_k(X_t) f(X_t)\Big) \\
&\quad = E\big(\partial_k f(X_t)\big).
\end{aligned}
$$

(q.e.d.)

Seien wieder b und σ lipschitzstetig. Gleichung (1.1) besitzt dann für jede Brownsche Bewegung eine eindeutig bestimmt Lösung mit Anfangsbedingung $X_0 = x \in \mathbb{R}^d$. Nach dem Satz von Yamada und Watanabe ist die Verteilung der Lösung auf $C\big([0;\infty), \mathbb{R}^d\big)$ eindeutig bestimmt (Die Stetigkeit des Prozesses erreicht man, indem man den zugrundeliegenden Wahrscheinlichkeitsraum einschränkt: statt Ω betrachtet man $\Omega \setminus N$ mit der induzierten σ-Algebra. Dabei ist N die Nullmenge, für die die Pfade unstetig sind.) Im Folgenden schreibe ich für diese Verteilung P^x und bezeichne den Erwartungswert bezüglich P^x mit E^x. Man erhält nun eine alternative Form der starken Markoveigenschaft (ebd., Theorem 12.5.1).

Satz 1.5 *Sei* $(\Omega, \mathcal{B}(\Omega), P)$ *ein polnischer Wahrscheinlichkeitsraum mit Filtration* $(\mathcal{F}_t)_{t\geq 0}$*,* X *eine stetige Lösung von Gleichung* (1.1) *und* S *eine Stoppzeit. Weiter sei* $f\colon C\big([0;\infty), \mathbb{R}^d\big) \to \mathbb{R}$ *Borel-meßbar. Dann gilt*

$$
E\big(f(X_{S+\,\cdot}) \,\big|\, \mathcal{F}_S^+\big) \;=\; E^{X_S}(f) \qquad \text{f.s.}
$$

Wenn ich im Folgenden die Verteilungen P^x oder die zugehörigen Erwartungen E^x verwende, so setze ich stets voraus, daß Ω ein polnischer Raum ist, damit man obigen Satz anwenden kann.

Aus diesem Satz erhält man einen Zusammenhang zwischen den Maßen E^x auf dem Pfadraum und den bedingten Erwartungen $E(\,\cdot \mid \mathcal{F}_0^+)$ auf Ω. Ist $f\colon C\big([0;\infty), \mathbb{R}^d\big) \to \mathbb{R}$ meßbar, so gilt

$$
E^{X_0}(f) \;=\; E\big(f(X) \,\big|\, \mathcal{F}_0^+\big) \qquad \text{f.s.} \tag{1.4}
$$

und damit

$$
E\big(E^{X_0}(f)\big) \;=\; E\big(f(X)\big).
$$

Mittels der Erwartungswerte $(E^x)_{x\in\mathbb{R}^d}$, die der Drift b und der Diffusionsmatrix σ zugeordnet sind, kann man eine Operatorhalbgruppe definieren. Dies geschieht im nächsten Satz (ebd., Theorem 12.5.3).

Satz 1.6 *Seien* $b\colon \mathbb{R}^d \to \mathbb{R}^d$ *und* $\sigma\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n}$ *lipschitzstetig. Dann wird durch*

$$
P_t\varphi(x) := E^x\big(\varphi(W_t)\big) \qquad \textit{für alle } x \in \mathbb{R}^d
$$

eine Operatorhalbgruppe auf dem Banachraum $C_b(\mathbb{R}^d, \mathbb{R})$ *definiert. Der Raum* $C_c^2(\mathbb{R}^d)$ *ist im Definitionsbereich des infinitesimalen Generators der Halbgruppe enthalten und auf diesem Raum stimmt der Generator mit dem* L *aus Gleichung* (1.2) *überein.*

Im Folgenden betrachte ich meist Prozesse mit lipschitzstetiger Drift und Diffusionsmatrix um die gerade konstruierte Halbgruppe verwenden zu können.

Die Verteilung von Diffusionsprozessen auf $C\big([0;\infty),\mathbb{R}^d\big)$ hat oft eine Dichte bezüglich des Wienermaßes. In manchen Situationen kann man diese mittels der Girsanov-Formel (ebd., Theorem 10.2.1) bestimmen:

Notation. $\mathcal{B}$ bezeichne die Klasse aller Prozesse, die f.s. gleich einem lokal beschränkten, vorhersagbaren (engl.: predictible) Prozeß sind.

Nach Proposition 7.1.4 bei v. Weizsäcker und Winkler enthält $\mathcal{B}$ alle adaptierten Prozesse, die f.s. regulär und f.s. linksstetig sind.

Satz 1.7 (Girsanov). *Sei $(\Omega,\mathcal{F},P)$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$, B eine $(\mathcal{F}_t)$-Brownsche Bewegung und $H\in\mathcal{B}$. Setze*

$$Z_t \ := \ \exp\Big(\int_0^t H_s\,dB_s - \frac{1}{2}\int_0^t {H_s}^2\,ds\Big) \qquad \textit{für alle } t\geq 0. \tag{1.5}$$

Zuletzt sei T eine Stoppzeit, so daß $Z_{\cdot\wedge T}$ ein durch Z_T abgeschlossenes Martingal ist, und Q das durch $dQ = Z_T dP$ definierte Maß. Dann ist der Prozeß B^* mit

$$B_t^* \ := \ B_{t\wedge T} - \int_0^{t\wedge T} H_s\,ds \qquad \textit{für alle } t\geq 0$$

eine in T gestoppte $(\mathcal{F}_t)$-Brownsche Bewegung auf dem Raum $(\Omega,\mathcal{F},Q)$.

In diesem Satz startet B als $(\mathcal{F}_t)$-Brownsche Bewegung im Ursprung. Dies stellt aber keine Einschränkung dar: Sei nämlich B eine $(\mathcal{F}_t)$-Brownsche Bewegung mit beliebiger Startverteilung $\mathcal{L}(B_0)$. Dann ist $\tilde{B} := B_\cdot - B_0$ eine Brownsche Bewegung, die in Null startet, und man kann obigen Satz auf $\tilde{B}$ anwenden. Da der Wert eines stochastischen Integrals nur von den Zuwächsen des Integrators abhängt, ist $\int H\,d\tilde{B} = \int H\,dB$ und man kann in Z weiterhin B statt $\tilde{B}$ schreiben. Ist Z ein Martingal, so liefert der Satz, daß $\tilde{B}^*$ mit

$$\tilde{B}_t^* \ := \ B_{t\wedge T} - B_0 - \int_0^{t\wedge T} H_s\,ds \qquad \text{für alle } t\geq 0$$

eine $(\mathcal{F}_t)$-Brownsche Bewegung bezüglich Q ist, und man erhält, daß

$$B^* \ := \ \tilde{B}^* + B_0 \ = \ B_{\cdot\wedge T} - \int_0^{\cdot\wedge T} H_s\,ds$$

eine in T gestoppte $(\mathcal{F}_t)$-Brownsche Bewegung bezüglich Q mit Startverteilung $\mathcal{L}(B_0)$ ist. Also gilt der Satz unverändert auch für $(\mathcal{F}_t)$-Brownsche Bewegungen mit beliebiger Startverteilung.

Korollar 1.1 *Sei $(\Omega,\mathcal{F},P)$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$, sowie B eine $(\mathcal{F}_t)$-Brownsche Bewegung und X eine Lösung der stochastischen Differentialgleichung*

$$dX \ = \ b(X)\bullet dt + dB.$$

Weiter sei Z der durch

$$Z_t \ := \ \exp\Big(\int_0^t -b(X_s)\,dB_s - \frac{1}{2}\int_0^t b^2(X_s)\,ds\Big) \qquad \textit{für alle } t\geq 0$$

definierte Prozeß und $T\geq 0$ so, daß $Z_{\cdot\wedge T}$ ein durch Z_T abgeschlossenes Martingal ist. Ist dann $\varphi_T\colon C\big([0;\infty),\mathbb{R}^d\big)\to\mathbb{R}$ eine $\mathcal{F}_T$-meßbare Abbildung mit

$$\varphi_T(X) = \exp\Big(\int_0^T b(X_s)\,dX_s - \frac{1}{2}\int_0^T b^2(X_s)\,ds\Big) \qquad P\textit{-f.s.},$$

so gilt

$$P(X \in A) \;=\; \int_A \varphi_T \, dW \qquad \textit{für alle } A \in \mathcal{F}_T.$$

Beweis. Es gilt

$$X_t \;=\; B_t - \int_0^t -b(X_s)\, ds \qquad \text{für alle } t \geq 0.$$

Wendet man die Girsanov-Formel für $H_s = -b(X_s)$ an, so ergibt sich, daß X eine in T gestoppte $(\mathcal{F}_t)$-Brownsche Bewegung auf dem Raum $(\Omega, \mathcal{F}, Q)$ ist, wobei Q das durch $dQ = Z_T \, dP$ definierte Maß ist. Weiter gilt

$$b(X) \bullet dB \;=\; b(X) \bullet \big(dX - b(X) \bullet dt\big) \;=\; b(X) \bullet dX - b^2(X) \bullet dt,$$

also

$$dQ \;=\; \exp\big(-\int_0^T b(X_s)\, dX_s + \frac{1}{2}\int_0^T b^2(X_s)\, ds\big)\, dP$$

und man erhält

$$\begin{aligned} P(X \in A) \;&=\; \int 1_{\{X \in A\}}\, dP \\ &=\; \int 1_{\{X \in A\}} \exp\big(\int_0^T b(X_s)\, dX_s - \frac{1}{2}\int_0^T b^2(X_s)\, ds\big)\, dQ \\ &=\; \int 1_A(X)\varphi_T(X)\, dQ \\ &=\; \int 1_A(\omega)\varphi_T(\omega)\, dW(\omega) \qquad \text{für alle } A \in \mathcal{F}_T. \end{aligned}$$

(q.e.d.)

Um die Girsanov-Formel anwenden zu können, benötigt man Kriterien, um zu entscheiden, ob $Z_{\cdot \wedge T}$ ein Martingal ist. Eine äquivalente Bedingung ist, daß $E(Z_T) = 1$ gilt (ebd., Theorem 10.2.3). Eine hinreichende Bedingung erhält man aus dem Novikov-Kriterium (Liptser und Shiryayev (1977), Theorem 6.1).

Satz (Novikov). *Sei B eine d-dimensionale $(\mathcal{F}_t)$-Brownsche Bewegung und H sei ein adaptierter, f.s. regulärer, f.s. linksstetiger Prozeß mit*

$$E\Big(\exp\big(\frac{1}{2}\int_0^T H_s^2\, ds\big)\Big) < \infty \qquad \textit{für alle } T \geq 0.$$

Dann ist der durch (1.5) *definierte Prozeß Z ein Martingal.*

Bemerkung. Liptser und Shiryayev setzen die sogenannten „usual conditions“ voraus, die ich (dem Buch von v. Weizsäcker und Winkler folgend) in dieser Arbeit nicht verwende. Dies sind folgende Bedingungen an die Filtration

(j) Die Filtration $(\mathcal{F}_t)_{t \geq 0}$ ist rechtsstetig.

(ij) Jede P-Nullmenge ist in allen $\mathcal{F}_t$ enthalten.

(iij) Alle Teilmengen von P-Nullmengen sind $\mathcal{F}_t$-meßbar.

Die Aussage des Novikov-Kriteriums läßt sich dennoch folgendermaßen verwenden: Man erweitert die Filtration $(\mathcal{F}_t)_{t \geq 0}$ zu einer Filtration $(\tilde{\mathcal{F}}_t^+)_{t \geq 0}$, die die „usual conditions“ erfüllt, indem man die geforderten Nullmengen hinzunimmt und die Filtration rechtsstetig macht. Dann ist die $(\mathcal{F}_t)$-Brownsche Bewegung B auch eine $(\tilde{\mathcal{F}}_t^+)$-Brownsche Bewegung. Nun kann man den Satz anwenden.

Ist dann T eine $(\mathcal{F}_t)$-Stoppzeit, so ist T auch eine $(\tilde{\mathcal{F}}_t)$-Stoppzeit, und der Stoppsatz liefert $E(Z_T) = 1$. Dieses Ergebnis ist aber unabhängig von der verwendeten Filtration, und hinreichend dafür, daß $Z_{\cdot \wedge T}$ bezüglich der ursprünglichen Filtration ein Martingal ist.

Aus dem Satz von Novikov erhält man auch folgendes Kriterium, das als Korollar 3.5.16 in dem Buch von Karatzas und Shreve (1991) steht.

Satz (Beneš). *Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$ und B eine d-dimensionale $(\mathcal{F}_t)$-Brownsche Bewegung. Weiter sei H ein adaptierter, f.s. regulärer, f.s. linksstetiger Prozeß mit Werten in $\mathbb{R}^d$ und für alle $T \geq 0$ gebe es eine Konstante K_T, so daß P-fast sicher*

$$|H_t| \;\leq\; K_T \cdot \big(\sup_{0\leq s\leq T} |B_s| + 1\big) \qquad \textit{für alle } t \in [0;T]$$

gilt. Dann ist der Prozeß Z mit

$$Z_t \;=\; \exp\Big(\int_0^t H_s\, dB_s - \frac{1}{2}\int_0^t H_s^2\, ds\Big) \qquad \textit{für alle } t \geq 0$$

ein Martingal.

Beweis. Sei $T > 0$. Wähle $0 = t_0 < \ldots < t_N = T$ so, daß $t_n - t_{n-1} < D := 1/(8TK_T^2)$ für alle $n \in \{1, \ldots, N\}$ gilt.

Die Brownsche Bewegung B ist ein Martingal und die Abbildung

$$f\colon x \longmapsto \exp\big(DK_T^2(|x|+1)^2/4\big)$$

ist konvex. Weiter ist

$$\begin{aligned}
E\big|f(B_t)\big| \;&\leq\; \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{DK_T^2(x^2+1)} \cdot \mathrm{e}^{-x^2/2t}\, dx \\
&=\; \frac{\mathrm{e}^{DK_T^2}}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{(DK_T^2 - 1/2t)x^2}\, dx \\
&<\; \infty \qquad \text{für alle } t \in [0;T],
\end{aligned}$$

da $DK_T^2 = 1/8T < 1/2t$ gilt. Nach der mehrdimensionalen Jensenschen Ungleichung für bedingte Wahrscheinlichkeiten (Dudley (1989), Satz 10.2.7) ist der Prozeß $\big(f(B_t)\big)_{t\in[0;T]}$ dann ein Submartingal. Indem man die Doobsche L^2-Ungleichung (Bauer (1991), Satz 46.4) auf dieses Submartingal anwendet erhält man

$$\begin{aligned}
E\Big(\exp\big(\frac{1}{2}\int_{t_{n-1}}^{t_n} |H_t|^2\, dt\big)\Big) \;&\leq\; E\Big(\exp\big(\frac{1}{2}DK_T^2\big(\sup_{0\leq s\leq T}|B_s|+1)^2\big)\Big) \\
&=\; E\Big(\big(\sup_{0\leq s\leq T} f(B_s)\big)^2\Big) \\
&\overset{\text{Doob}}{\leq}\; 4E\big(f^2(B_T)\big) \\
&\leq\; 4\frac{\mathrm{e}^{2DK_T^2}}{(2\pi T)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{(2DK_T^2 - 1/2T)x^2}\, dx \\
&<\; \infty.
\end{aligned}$$

Damit läßt sich das Novikov-Kriterium auf $\big(H_s \cdot 1_{(t_{n-1};t_n]}(s)\big)_{s\geq 0}$ anwenden, und man erhält

$$E\Big(\exp\big(\int_{t_{n-1}}^{t_n} H_s\, dB_s - \frac{1}{2}\int_{t_{n-1}}^{t_n} H_s^2\, ds\big) \;\Big|\; \mathcal{F}_{t_{n-1}}\Big) = 1.$$

Dies liefert

$$\begin{aligned}
E(Z_{t_n}) \;&=\; E\Big(E\Big(\exp\big(\int_0^{t_n} H_s\, dB_s - \frac{1}{2}\int_0^{t_n} H_s^2\, ds\big) \;\Big|\; \mathcal{F}_{t_{n-1}}\Big)\Big) \\
&=\; E\Big(\exp\big(\int_0^{t_{n-1}} H_s\, dB_s - \frac{1}{2}\int_0^{t_{n-1}} H_s^2\, ds\big) \\
&\qquad\qquad \cdot E\Big(\exp\big(\int_{t_{n-1}}^{t_n} H_s\, dB_s - \frac{1}{2}\int_{t_{n-1}}^{t_n} H_s^2\, ds\big) \;\Big|\; \mathcal{F}_{t_{n-1}}\Big)\Big) \\
&=\; E(Z_{t_{n-1}} \cdot 1) \qquad \text{für } n := 1, \ldots, N
\end{aligned}$$

und mit Induktion folgt $E(Z_T) = E(Z_{t_0}) = 1$. Nach Proposition 10.2.3 bei v. Weizsäcker und Winkler ist $Z_{\cdot \wedge T}$ dann ein Martingal. Da dies für alle $T \geq 0$ gilt, ist Z ein Martingal. (q.e.d.)

Korollar 1.2 *Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$ und B eine $(\mathcal{F}_t)$-Brownsche Bewegung. Weiter sei $z \in \mathbb{R}^d$ sowie X eine Lösung der stochastischen Differentialgleichung*

$$dX = b(X) \bullet dt + dB$$

mit $X_0 = z$ f.s. und lipschitzstetiger Drift $b\colon \mathbb{R}^d \to \mathbb{R}^d$. Dann ist der durch

$$Z_t \;:=\; \exp\Big(-\int_0^t b(X_s)\,dB_s - \frac{1}{2}\int_0^t b^2(X_s)\,ds\Big) \qquad \text{für alle } t \geq 0$$

definierte Prozeß Z ein Martingal.

Beweis. Sei $z \in \mathbb{R}^d$ und $w \in C\big([0;\infty), \mathbb{R}^d\big)$. Weiter sei x die Lösung der (gewöhnlichen, deterministischen) Differentialgleichung

$$x_t \;=\; z + w_t + \int_0^t b(x_s)\,ds \qquad \text{für alle } t \geq 0. \tag{1.6}$$

Nach Voraussetzung ist b lipschitzstetig, so daß es ein $\kappa > 0$ mit $\big|b(x)\big| \leq \kappa \cdot \big(|x| + 1\big)$ gibt. Folglich gilt

$$\begin{aligned}
\big|b(x_t)\big| \;&\leq\; \kappa \cdot \big(|x_t| + 1\big) \\
&\leq\; \kappa \cdot \Big(|z| + |w_t| + \int_0^t \big|b(x_s)\big|\,ds + 1\Big) \\
&=\; \kappa \cdot \big(|w_t| + |z| + 1\big) + \int_0^t \kappa \cdot \big|b(x_s)\big|\,ds
\end{aligned}$$

und mit dem Gronwall-Lemma (Hilfssatz 3.1.2 bei Werner und Arndt) erhält man die Abschätzung

$$\begin{aligned}
\big|b(x_t)\big| \;&\leq\; \kappa \cdot \big(|w_t| + |z| + 1\big) + \int_0^t \kappa \cdot \kappa \cdot \big(|w_s| + |z| + 1\big) \cdot \exp\Big(\int_s^t \kappa\,dr\Big)\,ds \\
&\leq\; \kappa \cdot \big(\sup_{0\leq s\leq T} |w_s| + |z| + 1\big) + \kappa \cdot \big(\sup_{0\leq s\leq T} |w_s| + |z| + 1\big) \cdot (\mathrm{e}^{\kappa\cdot t} - 1) \\
&\leq\; \kappa \cdot \big(|z| + 1\big)\mathrm{e}^{\kappa\cdot T}\big(\sup_{0\leq s\leq T} |w_s| + 1\big) \qquad \text{für alle } t \in [0;T].
\end{aligned}$$

Da $t \mapsto t$ von lokal endlicher Variation ist, ist $b(X) \bullet dt$ eine Version von $\int b\big(X_s(\omega)\big)\,ds$ und damit genügt $t \mapsto X_t(\omega)$ für P-fast alle ω der Gleichung (1.6) mit $w = B(\omega)$. Also gilt fast sicher $\big|-b(X_t)\big| \leq \kappa \cdot \big(|z| + 1\big)\mathrm{e}^{\kappa\cdot T}\big(\sup_{0\leq s\leq T} |B_s| + 1\big)$. Damit kann man den Satz von Beneš anwenden und erhält, daß Z ein Martingal ist. (q.e.d.)

Korollar 1.3 *Sei μ ein Wahrscheinlichkeitsmaß auf $\mathbb{R}^d$, W der Wienerprozeß mit Startverteilung μ und $b\colon \mathbb{R}^d \to \mathbb{R}^d$ eine lipschitzstetige Abbildung. Dann ist der Prozeß*

$$Z \;=\; \exp\Big(\int_0^{\cdot} b(W_s)\,dW_s \;-\; \frac{1}{2}\int_0^{\cdot} b^2(W_s)\,ds\Big)$$

ein Martingal.

Beweis. Es bezeichne E^x und E^μ die Erwartungswerte bezüglich des Wienermaßes mit Startpunkt x bzw. Startverteilung μ. Weiter sei $T > 0$ und $x \in \mathbb{R}^d$. Da b lipschitzstetig ist, gibt es ein $K \geq 0$ mit

$$\begin{aligned}
\big|b(W_t + x)\big| \;&\leq\; \big|b(W_t + x) - b(x)\big| + \big|b(x)\big| \\
&\leq\; K\big|W_t + x - x\big| + K \\
&\leq\; K \cdot \big(\sup_{0\leq s\leq T} |W_s| + 1\big) \qquad \text{für alle } t \in [0;T].
\end{aligned}$$

Nach dem Satz von Beneš ist also

$$\begin{aligned} E^x(Z_T) &= E^0\Big(\exp\big(\int_0^T b(W_s+x)\,dW_s - \frac{1}{2}\int_0^T b^2(W_s+x)\,ds\big)\Big) \\ &= 1 \qquad \text{für alle } x \in \mathbb{R}^d, \end{aligned}$$

und damit

$$E^\mu(Z_T) = \int_{\mathbb{R}^d} E^x(Z_T)\,d\mu(x) = 1 \qquad \text{für alle } T \geq 0.$$

Dies zeigt, daß Z ein Martingal ist. (q.e.d.)

Eine weitere Anwendung der Girsanov-Formel besteht darin, die schwache Lösbarkeit von stochastischen Differentialgleichungen zu zeigen. Zu diesem Zweck muß man einen Wahrscheinlichkeitsraum und einen Prozeß auf diesem Raum konstruieren, so daß der Prozeß die vorgegebene Gleichung löst.

Satz 10.2.5 bei v. Weizsäcker und Winkler liefert eine Version der Girsanov-Formel, bei der man eine Brownsche Bewegung anstatt einer gestoppten Brownschen Bewegung erhält.

Satz 1.8 *Sei $\Omega = C\big([0;\infty), \mathbb{R}^d\big)$, W der Wienerprozeß auf $(\Omega, \mathcal{B}(\Omega), \mathbb{W})$ und $(\mathcal{F}_t)_{t\geq 0}$ die natürliche Filtration von W. Weiter sei $H \in \mathcal{B}$ so, daß der Prozeß*

$$Z := \exp\big(\int_0^\cdot H_s\,dW_s - \frac{1}{2}\int_0^\cdot {H_s}^2\,ds\big)$$

ein Martingal ist. Dann gibt es eine eindeutig bestimmtes Wahrscheinlichkeitsmaß Q auf $\mathcal{F}$, so daß $dQ = Z_t\,d\mathbb{W}$ auf $\mathcal{F}_t$ für alle $t \geq 0$ gilt, und auf dem Raum $(\Omega, \mathcal{F}, Q)$ ist der Prozeß

$$B = W - \int_0^\cdot H_s\,ds$$

eine $(\mathcal{F}_t)$-Brownsche Bewegung.

Sei nun W der Wienerprozeß und Z das exponentielle Martingal von $\int_0^\cdot b(W)\,dW$, also

$$Z_t = \exp\big(\int_0^t b(W_s)\,dW_s - \frac{1}{2}\int_0^t b^2(W_s)\,ds\big) \qquad \text{für alle } t \geq 0.$$

Weiter sei b so, daß Z ein Martingal ist. Dann ist nach der Girsanov-Formel, in der Form von Satz 1.8, der Prozeß B mit

$$B_t = W_t - \int_0^t b(W_s)\,ds \qquad \text{für alle } t \geq 0$$

eine Brownsche-Bewegung auf $\big(C(\mathbb{R}_+), \mathcal{F}, Q\big)$ und nach Definition von B ist W auf diesem Raum eine Lösung der stochastischen Differentialgleichung

$$dW = b(W) \bullet dt + dB.$$

2. Große Abweichungen

In diesem Kapitel beweise ich zwei Sätze über große Abweichungen für das Wienermaß. Die Darstellung stammt aus dem Buch von Varadhan (1980). Diese Sätze werde ich im nächsten Kapitel verwenden, um die Konvergenzrate des Bayesrisikos für Diffusionsprozesse mit konstanter Drift zu berechnen. Dabei beobachtet man die Pfade des Prozesses bis zur Zeit $T > 0$ und läßt später T gegen unendlich gehen. Damit alle diese Beobachtungen vergleichbar werden, kann man die Pfade des Prozesses so stauchen, daß das Intervall $[0;T]$ auf $[0;1]$ abgebildet wird. Dadurch erhält man einen Prozeß auf $[0;1]$, dessen Verteilung von T abhängt. Das Transformationsverhalten desjenigen Anteils des Prozesses, der von der Brownschen Bewegung her kommt, wird durch folgendes Lemma beschrieben.

Lemma 2.1 *Sei $(B_t)_{t\geq 0}$ eine d-dimensionale Brownsche Bewegung und $T > 0$. Dann ist auch der Prozeß $(\frac{1}{\sqrt{T}}B_{t\cdot T})_{t\geq 0}$ eine Brownsche Bewegung.*

Beweis. Setze $X_t := \frac{1}{\sqrt{T}}B_{t\cdot T}$ für alle $t \geq 0$. Ist ein Pfad $t \mapsto B_t(\omega)$ stetig, so ist auch $t \mapsto X_t(\omega)$ stetig, d.h. X hat f.s. stetige Pfade.

Weiter gilt $\mathcal{L}\big(B_{t\cdot T} - B_{s\cdot T}\big) = \mathcal{N}^d\big(0,(t-s)T\big)$, also

$$\mathcal{L}\big(X_t - X_s\big) \;=\; \mathcal{L}\big(\frac{B_{t\cdot T} - B_{s\cdot T}}{\sqrt{T}}\big) \;=\; \mathcal{N}^d\big(0,\frac{(t-s)T}{T}\big) \;=\; \mathcal{N}^d\big(0,t-s\big),$$

und auch die Unabhängigkeit der Zuwächse überträgt sich von B auf X. (q.e.d.)

Das Lemma zeigt, daß die Verteilung der Brownschen Bewegung erhalten bleibt, wenn die Transformation den Ort um den Faktor $\sqrt{T}$ staucht. Ich verwende aber eine Stauchung des Raums um den Faktor T, so daß der Einfluß der Brownschen Bewegung mit steigendem T abnimmt. Sei also $\mathbb{W}$ das Wienermaß und W der Wienerprozeß auf $C\big([0;\infty),\mathbb{R}^d\big)$. Definiere für $T > 0$ die Abbildung $\tau\colon C\big([0;\infty),\mathbb{R}^d\big) \to C\big([0;1],\mathbb{R}^d\big)$ durch

$$\big(\tau\omega\big)(t) \;:=\; \frac{1}{T}\,\omega(t\cdot T) \qquad \text{für alle } t\in[0;1],\, \omega \in C\big([0;\infty),\mathbb{R}^d\big)$$

und setze

$$\mathbb{W}_{1/T} \;:=\; \mathbb{W}\circ\tau^{-1} \qquad \text{für alle } T > 0.$$

Für $\varepsilon > 0$ ist $\mathbb{W}_\varepsilon$ dann ein Maß auf $C\big([0;1],\mathbb{R}^d\big)$.

Sei $A \subseteq C\big([0;1],\mathbb{R}^d\big)$ meßbar. Dann gilt

$$\begin{aligned}\mathbb{W}_\varepsilon(A) \;&=\; \mathbb{W}\big((\varepsilon W_{t/\varepsilon})_{t\in[0;1]} \in A\big) \;=\; \mathbb{W}\big((\sqrt{\varepsilon}W_{t/\varepsilon})_{t\in[0;1]} \in \frac{A}{\sqrt{\varepsilon}}\big)\\ &\overset{\text{Lemma 2.1}}{=}\; \mathbb{W}\big((W_t)_{t\in[0;1]} \in \frac{A}{\sqrt{\varepsilon}}\big) \;=\; \mathbb{W}\big(\frac{A}{\sqrt{\varepsilon}}\big),\end{aligned}$$

wobei das $\mathbb{W}$ im letzten Term die Einschränkung des Wienermaßes auf $C[0;1]$ bezeichnet. Dies läßt sich auch als

$$\mathbb{W}_\varepsilon\big(\frac{W}{\sqrt{\varepsilon}} \in A\big) \quad=\quad \mathbb{W}_\varepsilon\big(W \in \sqrt{\varepsilon}\cdot A\big) \quad=\quad \mathbb{W}\big(W \in A\big)$$

schreiben, d.h. $\frac{W}{\sqrt{\varepsilon}}$ ist eine (in 1 gestoppte) Brownsche Bewegung bzgl. $\mathbb{W}_\varepsilon$.

Weiter definiere $\mathrm{I}: C\big([0;1],\mathbb{R}^d\big) \to \mathbb{R}$ durch

$$\mathrm{I}(\omega) \;:=\; \begin{cases} \frac{1}{2}\int_0^1 |\dot\omega_s|^2 ds, & \text{falls } \omega \text{ absolut stetig mit } \omega_0 = 0 \text{ ist, und} \\ +\infty & \text{sonst.} \end{cases}$$

Dieses I ist so etwas wie eine Kurvenlänge: Für $\mathrm{I}(\omega) < +\infty$ gilt

$$\frac{1}{2}|\omega_1|^2 \;=\; \frac{1}{2}\Big|\int_0^1 1 \cdot \dot\omega_s \, ds\Big|^2 \overset{\text{Hölder}}{\leq} \frac{1}{2}\Big(\int_0^1 |1|^2 ds\Big) \cdot \Big(\int_0^1 |\dot\omega_s|^2 ds\Big) \;=\; \mathrm{I}(\omega).$$

Gleichheit gilt in der Hölderschen Ungleichung genau dann, wenn die Faktoren linear abhängig sind, also hier wenn $\dot\omega \equiv x \cdot 1$ und damit $\omega_t = x \cdot t$ für ein $x \in \mathbb{R}^d$ ist. Insbesondere gilt

$$\inf\big\{\mathrm{I}(\omega) \;\big|\; \omega \in C[0;1], \omega_1 = x\big\} \;=\; \mathrm{I}\big((x \cdot t)_{t\in[0;1]}\big) \;=\; \frac{1}{2}|x|^2. \tag{2.1}$$

Lemma 2.2 *Sei $\ell \geq 0$. Dann ist die Menge $K_\ell := \big\{\omega \in C[0;1] \;\big|\; \mathrm{I}(\omega) \leq \ell\big\}$ kompakt. Insbesondere ist* I *nach unten halbstetig, d.h. für alle ℓ ist K_ℓ abgeschlossen.*

Manchmal führt man im Zusammenhang mit Sätzen über große Abweichungen folgende Begriffe ein. Das Lemma besagt dann, daß I eine gute Ratenfunktion ist.

Definition. Sei X ein topologischer Raum, $I: X \to [0;\infty]$ nach unten halbstetig. Dann heißt I eine **Ratenfunktion**. Gilt darüberhinaus, daß alle Mengen der Form $\big\{x \in X \;\big|\; \mathrm{I}(x) \leq \ell\big\}$ mit $\ell \geq 0$ kompakt sind, so heißt I eine **gute Ratenfunktion**.

Beweis des Lemmas. Ich zeige zuerst, daß I halbstetig ist. Sei daher $(\omega_m)_{m\in\mathbb{N}}$ eine Folge in K_ℓ mit $\omega_m \to \omega$ gleichmäßig für ein $\omega \in C\big([0;1],\mathbb{R}^d\big)$. Weiter sei $\big([s_i,t_i] \;\big|\; i := 1,\dots,n\big)$ eine Familie paarweise disjunkter Intervalle in $[0;1]$. Da die ω_m absolut stetig sind, erhält man

$$\begin{aligned} \sum_{i=1}^n \big|\omega_m(t_i) - \omega_m(s_i)\big| \;&=\; \sum_{i=1}^n \Big|\int_{s_i}^{t_i} \dot\omega_m(t)\, dt\Big| \\ &\overset{\text{Hölder}}{\leq} \sum_{i=1}^n \big|t_i - s_i\big|^{1/2} \cdot \Big(\int_{s_i}^{t_i} |\dot\omega_m|^2 \, dt\Big)^{1/2} \\ &\overset{\text{Hölder}}{\leq} \Big(\sum_{i=1}^n |t_i - s_i|\Big)^{1/2} \cdot \Big(\sum_{i=1}^n \int_{s_i}^{t_i} |\dot\omega_m|^2 \, dt\Big)^{1/2} \\ &\leq\; \Big(\sum_{i=1}^n |t_i - s_i|\Big)^{1/2} \cdot \sqrt{2\ell} \qquad \text{für alle } m \in \mathbb{N}, \end{aligned}$$

und für $m \to \infty$ auch

$$\sum_{i=1}^n \big|\omega(t_i) - \omega(s_i)\big| \;\leq\; \Big(\sum_{i=1}^n |t_i - s_i|\Big)^{1/2} \cdot \sqrt{2\ell}.$$

Also ist ω absolut stetig.

Weiter gilt

$$\left|\frac{\omega_m(t+h) - \omega_m(t)}{h}\right|^2 \;=\; \frac{1}{h^2}\left|\int_t^{t+h} \dot\omega_m(s) ds\right|^2 \overset{\text{Hölder}}{\leq} \frac{1}{h}\int_t^{t+h} |\dot\omega_m(s)|^2 ds,$$

also

$$\begin{aligned} \int_0^{1-h} \left|\frac{\omega_m(t+h) - \omega_m(t)}{h}\right|^2 dt \;&\leq\; \frac{1}{h}\int_0^{1-h}\int_0^h |\dot\omega_m(t+s)|^2 ds\, dt \\ &\overset{\text{Tonelli}}{=} \frac{1}{h}\int_0^h \int_0^{1-h} |\dot\omega_m(t+s)|^2 dt\, ds \\ &\leq\; \frac{1}{h} \cdot 2\ell \int_0^h ds \;=\; 2\ell \end{aligned}$$

für alle $m \in \mathbb{N}$ und damit auch

$$\int_0^{1-h} \left|\frac{\omega(t+h)-\omega(t)}{h}\right|^2 dt \quad \leq \quad 2\ell.$$

Läßt man nun h gegen 0 gehen, so erhält man

$$\begin{aligned} I(\omega) \quad &= \quad \frac{1}{2}\int_0^1 \liminf_{h\downarrow 0} 1_{[0;1-h]}(t)\cdot\left|\frac{\omega(t+h)-\omega(t)}{h}\right|^2 dt \\ &\overset{\text{Fatou}}{\leq} \quad \liminf_{h\downarrow 0} \frac{1}{2}\int_0^1 1_{[0;1-h]}(t)\cdot\left|\frac{\omega(t+h)-\omega(t)}{h}\right|^2 dt \quad \leq \quad \ell, \end{aligned}$$

d.h. $\omega \in K_\ell$. Also ist K_ℓ abgeschlossen.

Außerdem gilt

$$\begin{aligned} \big|\omega(t_2)-\omega(t_1)\big| \quad &= \quad \Big|\int_{t_1}^{t_2} 1\cdot\dot\omega(s)\,ds\Big| \\ &\overset{\text{Hölder}}{\leq} \quad \sqrt{|t_2-t_1|}\cdot\Big(\int_{t_1}^{t_2}|\dot\omega(s)|^2\,ds\Big)^{1/2} \\ &\leq \quad \sqrt{|t_2-t_1|}\cdot\sqrt{2\ell} \end{aligned}$$

für alle $\omega \in K_\ell$, d.h. K_ℓ ist gleichgradig stetig und wegen $\omega_0 = 0$ für alle $\omega \in K_\ell$ auch beschränkt. Nach dem Satz von Arzelà-Ascoli ist $K_\ell = \overline{K_\ell}$ dann kompakt. (q.e.d.)

Lemma 2.3 *Sei* $C \subseteq C\big([0;1],\mathbb{R}^d\big)$ *abgeschlossen,* $C^\delta := \bigcup_{\omega\in C} S(\omega,\delta)$ *für alle* $\delta > 0$. *Dann gilt*

$$\lim_{\delta\to 0}\big(\inf_{\omega\in C^\delta} \mathrm{I}(\omega)\big) \quad = \quad \inf_{\omega\in C}\mathrm{I}(\omega).$$

Beweis. Für $\delta \downarrow 0$ wächst $\delta \mapsto \inf_{\omega\in C^\delta}\mathrm{I}(\omega)$ monoton und es gilt

$$\inf_{\omega\in C^\delta}\mathrm{I}(\omega) \quad \leq \quad \inf_{\omega\in C}\mathrm{I}(\omega) \qquad \text{für alle } \delta > 0.$$

Also existiert $\ell := \lim_{\delta\to 0}\big(\inf_{\omega\in C^\delta}\mathrm{I}(\omega)\big)$ und es gilt $\ell \leq \inf_{\omega\in C}\mathrm{I}(\omega)$.

Setze nun $\Delta := \{\,1/n \mid n\in\mathbb{N}\,\}$ und wähle für jedes $\delta\in\Delta$ ein $\omega_\delta \in C^\delta$, so daß die Menge $\{\,\mathrm{I}(\omega_\delta) \mid \delta\in\Delta\,\}$ durch ein $\alpha\in\mathbb{R}$ beschränkt ist und $\mathrm{I}(\omega_{1/n}) \longrightarrow \ell$ für $n\to\infty$ gilt. Dann ist

$$\begin{aligned} \big|\omega_\delta(t_2)-\omega_\delta(t_1)\big| \;&=\; \Big|\int_{t_1}^{t_2}\dot\omega(t)\,dt\Big| \\ &\overset{\text{Hölder}}{\leq} \; \sqrt{|t_2-t_1|}\cdot\Big(\int_{t_1}^{t_2}|\dot\omega_\delta(t)|^2\,dt\Big)^{\frac{1}{2}} \;\leq\; \sqrt{|t_2-t_1|\cdot 2\alpha}, \end{aligned}$$

d.h. $\{\,\omega_\delta \mid \delta\in\Delta\,\}$ ist gleichgradig stetig und wegen $\omega_\delta(0) = 0$ für alle δ auch beschränkt in $\big(C\big([0;1],\mathbb{R}^d\big),\|\cdot\|_\infty\big)$. Nach dem Satz von Arzelà-Ascoli ist damit $\{\omega_\delta\}$ relativ kompakt, d.h. es existiert eine Folge (δ_n) in Δ und eine Funktion $\omega \in C\big([0;1],\mathbb{R}^d\big)$ mit $\delta_n \downarrow 0$ und $\omega_{\delta_n}\to\omega$ gleichmäßig. Offensichtlich ist $\omega \in \bigcap_{\delta>0} C^\delta = C$.

Es gilt $\mathrm{I}(\omega_{\delta_n}) \to \ell$. Für jedes $\varepsilon > 0$ gibt es also ein $N\in\mathbb{N}$ mit $\mathrm{I}(\omega_{\delta_n}) \leq \ell+\varepsilon$ für alle $n \geq N$ und mit Lemma 2.2 folgt $\mathrm{I}(\omega) \leq \ell+\varepsilon$. Damit ist auch $\inf_{\omega\in C}\mathrm{I}(\omega) \leq \ell$ gezeigt.

(q.e.d.)

Satz (Varadhan). *Sei* $(\mathbb{W}_\varepsilon)_{\varepsilon>0}$ *die oben konstruierte Familie von Maßen auf dem Raum* $C\big([0;1],\mathbb{R}^d\big)$ *und* I *die eingangs definierte Ratenfunktion. Dann gilt*

(j) *für jede abgeschlossene Teilmenge* $C \subseteq C\big([0;1],\mathbb{R}^d\big)$ *ist*

$$\limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\mathbb{W}_\varepsilon(C) \leq -\inf_{\omega\in C}\mathrm{I}(\omega)$$

und

(ij) *für jede offene Teilmenge* $G \subseteq C\big([0;1],\mathbb{R}^d\big)$ *ist*

$$\liminf_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\mathbb{W}_\varepsilon(G) \geq -\inf_{\omega\in G}\mathrm{I}(\omega).$$

Beweis (von Teil (j) des Satzes). Sei $C \subseteq C[0;1]$ eine abgeschlossene Menge, $W = (W_1, \dots, W_d)$ der d-dimensionale Wienerprozeß auf $[0;1]$ und $W^n = (W_1^n, \dots, W_d^n)$ die Linearisierung von W durch die Stützstellen (t_j, W_{t_j}) mit $t_j := j/n$ und $j \in \{0, \dots, n\}$. Auch für $W \in C$ ist im allgemeinen nicht $W^n \in C$. Der Beweis liefert zunächst Abschätzungen für die Wahrscheinlichkeiten $\mathbb{W}_\varepsilon(W^n \notin C^\delta)$ und $\mathbb{W}_\varepsilon(W^n \in C^\delta)$. Diese Zahlen hängen von ε, δ und n ab. Dann werden nacheinander die Grenzwerte ε gegen Null, danach n gegen Unendlich und zuletzt δ gegen Null betrachtet.

Für $t \in [t_j; t_{j+1}]$ gilt

$$\begin{aligned} \big|W_r^n(t) - W_r(t)\big| &\leq \big|W_r^n(t) - W_r(t_j)\big| + \big|W_r(t_j) - W_r(t)\big| \\ &\leq \big|W_r^n(t_{j+1}) - W_r(t_j)\big| + \big|W_r(t) - W_r(t_j)\big| \\ &\leq 2 \sup_{s \in [t_j; t_{j+1}]} \big|W_r(s) - W_r(t_j)\big| \end{aligned}$$

und folglich

$$\frac{1}{2}\big|W_r^n(t) - W_r(t)\big| \leq \max_{0 \leq j < n} \sup_{s \in [t_j; t_{j+1}]} \big|W_r(s) - W_r(t_j)\big| \qquad \text{für alle } t \in [0;1].$$

Ist nun $\big|W^n(t) - W(t)\big| \geq \delta$, so gibt es ein $r \in \{1, \dots, d\}$ mit

$$\big|W_r^n(t) - W_r(t)\big| \geq \frac{\delta}{\sqrt{d}}$$

und mit dem gerade Gezeigten folgt

$$\max_{0 \leq j < n} \sup_{s \in [t_j; t_{j+1}]} \big|W_r(s) - W_r(t_j)\big| \geq \frac{\delta}{2\sqrt{d}}.$$

Daher gilt

$$\mathbb{W}_\varepsilon\Big(\|W^n - W\|_\infty \geq \delta \Big) \leq \mathbb{W}_\varepsilon\Big(\bigcup_{r=1}^{d} \Big\{ \max_{0 \leq j < n} \sup_{s \in [t_j; t_{j+1}]} \big|W_r(s) - W_r(t_j)\big| \geq \frac{\delta}{2\sqrt{d}} \Big\} \Big).$$

Da W eine d-dimensionale Brownsche Bewegung mit $W(0) = 0$ f.s. ist, sind auch die Komponenten W_r Brownsche Bewegungen mit $W_r(0) = 0$ f.s. Aufgrund der starken Markoveigenschaft gilt

$$\mathbb{W}_\varepsilon\Big(\sup_{s \in [t_j; t_{j+1}]} \big|W_r(s) - W_r(t_j)\big| \geq \frac{\delta}{2\sqrt{d}} \Big) = \mathbb{W}_\varepsilon\Big(\sup_{0 \leq t \leq \frac{1}{n}} |W_r(t) - W_r(0)| \geq \frac{\delta}{2\sqrt{d}} \Big)$$

und damit

$$\begin{aligned} \mathbb{W}_\varepsilon\Big(\|W^n - W\|_\infty \geq \delta \Big) &\leq \sum_{r=1}^{d} \sum_{j=1}^{n} \mathbb{W}_\varepsilon\Big(\sup_{0 \leq t \leq \frac{1}{n}} |W_r(t) - W_r(0)| \geq \frac{\delta}{2\sqrt{d}} \Big) \\ &= d \cdot n \cdot \mathbb{W}_\varepsilon\Big(\sup_{0 \leq t \leq \frac{1}{n}} |W_1(t)| \geq \frac{\delta}{2\sqrt{d}} \Big) \\ &\leq 2d \cdot n \cdot \mathbb{W}_\varepsilon\Big(\sup_{0 \leq t \leq \frac{1}{n}} W_1(t) \geq \frac{\delta}{2\sqrt{d}} \Big) \\ &= 2d \cdot n \cdot \mathbb{W}\Big(\sup_{0 \leq t \leq \frac{1}{n}} W_1(t) \geq \frac{\delta}{2\sqrt{d}\sqrt{\varepsilon}} \Big) \\ &\overset{\substack{\text{Spiegelungs-}\\ \text{prinzip}}}{=} 4dn \cdot \mathbb{W}\Big(W_1(\frac{1}{n}) \geq \frac{\delta}{2\sqrt{d\varepsilon}} \Big) \\ &= 4dn \int_{\frac{\delta}{2\sqrt{d\varepsilon}}}^{\infty} \frac{1}{\sqrt{2\pi \frac{1}{n}}} \exp\Big(-\frac{y^2}{2 \cdot \frac{1}{n}}\Big)\, dy \\ &\overset{x := \sqrt{n} y}{=} 4dn \int_{\frac{\delta\sqrt{n}}{2\sqrt{d\varepsilon}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)\, dx \end{aligned}$$

Für $a > 0$ gilt

$$a \cdot \int_a^\infty e^{-\frac{x^2}{2}} dx \quad \leq \quad \int_a^\infty x \cdot e^{-\frac{x^2}{2}} dx \quad = \quad e^{-\frac{a^2}{2}}.$$

Mit dieser Abschätzung erhält man

$$\begin{aligned} \mathbb{W}_\varepsilon\Big(\|W^n - W\|_\infty \geq \delta\Big) \;&\leq\; \frac{4dn}{\sqrt{2\pi}} \cdot \frac{2\sqrt{d\varepsilon}}{\delta\sqrt{n}} \exp\left(-\frac{n\delta^2}{8\varepsilon d}\right) \\ &=\; \underbrace{\frac{8d\sqrt{dn}}{\sqrt{2\pi}}}_{=:\, C_1(n)} \cdot \frac{\sqrt{\varepsilon}}{\delta} \exp\left(-\frac{n\delta^2}{8\varepsilon d}\right). \end{aligned}$$

Setze nun $\ell_\delta := \inf\{\, \mathrm{I}(\omega) \mid \omega \in C^\delta \,\}$ für das oben definierte C^δ. Dann gilt

$$\begin{aligned} \mathbb{W}_\varepsilon\big(W^n \in C^\delta\big) \;&\leq\; \mathbb{W}_\varepsilon\big(\mathrm{I}(W^n) \geq \ell_\delta\big) \\ &=\; \mathbb{W}_\varepsilon\Big(\frac{1}{2}\sum_{j=0}^{n-1} \frac{|W(\frac{j+1}{n}) - W(\frac{j}{n})|^2}{1/n} \;\geq\; \ell_\delta\Big) \\ &=\; \mathbb{W}\Big(Y_1^2 + \cdots + Y_{n\cdot d}^2 \geq \frac{2\ell_\delta}{\varepsilon}\Big), \end{aligned}$$

wobei $Y_1 := \sqrt{n}\big(W_1(\frac{1}{n}) - W_1(0)\big), \ldots, Y_{n\cdot d} := \sqrt{n}\big(W_d(1) - W_d(\frac{n-1}{n})\big)$ unabhängige $\mathcal{N}(0,1)$-verteilte Zufallsvariablen sind.

Es gilt also

$$\begin{aligned} &\mathbb{W}\big(Y_1^2 + \cdots + Y_{n\cdot d}^2 \geq \frac{2\ell_\delta}{\varepsilon}\big) \\ &\quad=\; \left(\frac{1}{2\pi}\right)^{\frac{n\cdot d}{2}} \int_{y_1^2+\cdots+y_{nd}^2 \geq \frac{2\ell_\delta}{\varepsilon}} \exp\big(-\frac{y_1^2 + \cdots + y_{nd}^2}{2}\big)\, dy_1 \cdots dy_{nd} \\ &\quad=\; C(n) \int_{\sqrt{\frac{2\ell_\delta}{\varepsilon}}}^\infty \exp(-\frac{r^2}{2}) \cdot r^{nd-1}\, dr \\ &\quad\overset{s := \frac{r^2}{2}}{=}\; C'(n) \int_{\ell_\delta/\varepsilon}^\infty \exp(-s) \cdot s^{\frac{nd-2}{2}}\, ds. \end{aligned}$$

Partielle Integration ergibt

$$\begin{aligned} \int_\alpha^\infty e^{-s} s^k\, ds \quad &= \quad -e^{-s} s^k \Big|_\alpha^\infty + \int_\alpha^\infty e^{-s} k s^{k-1}\, ds \\ &\overset{\text{Induktion}}{=} \; e^{-\alpha}\alpha^k + k \cdot e^{-\alpha}\alpha^{k-1} + \ldots + k! e^{-\alpha}\alpha^0 \\ &= \; e^{-\alpha}(\alpha^k + k \cdot \alpha^{k-1} + \ldots + k! \cdot 1). \end{aligned}$$

Für gerade n und hinreichend kleine ε erhält man durch Einsetzen

$$\mathbb{W}_\varepsilon(W^n \in C^\delta) \;\leq\; C'(n) \cdot \int_{\ell_\delta/\varepsilon}^\infty e^{-s} \cdot s^{\frac{nd}{2}-1}\, ds \;\leq\; C_2(n) \cdot e^{-\ell_\delta/\varepsilon} \left(\frac{\ell_\delta}{\varepsilon}\right)^{\frac{nd}{2}-1}.$$

Sei nun also $\delta > 0$, $n \in \mathbb{N}$ gerade und ε hinreichend klein für obige Abschätzung. Da für $W \in C$ entweder $\|W^n - W\|_\infty \geq \delta$ oder $W^n \in C^\delta$ folgt, erhält man dann

$$\begin{aligned} \mathbb{W}_\varepsilon(C) \;&\leq\; \mathbb{W}_\varepsilon\big(\|W^n - W\|_\infty \geq \delta\big) \;+\; \mathbb{W}_\varepsilon\big(W^n \in C^\delta\big) \\ &\leq\; C_1(n) \cdot \frac{\sqrt{\varepsilon}}{\delta} \exp\Big(-\frac{n\delta^2}{8\varepsilon d}\Big) \;+\; C_2(n) \cdot \Big(\frac{\ell_\delta}{\varepsilon}\Big)^{\frac{nd}{2}-1} \exp\Big(-\frac{\ell_\delta}{\varepsilon}\Big) \\ &\leq\; 2 \cdot \max\left(C_1(n) \cdot \frac{\sqrt{\varepsilon}}{\delta} \exp\Big(-\frac{n\delta^2}{8\varepsilon d}\Big), C_2(n) \cdot \Big(\frac{\ell_\delta}{\varepsilon}\Big)^{\frac{nd}{2}-1} \exp\Big(-\frac{\ell_\delta}{\varepsilon}\Big)\right), \end{aligned}$$

also

$$\varepsilon \cdot \log \mathbb{W}_\varepsilon(C) \;\le\; \varepsilon \log 2 \,+\, \varepsilon \cdot \max\Big(\log\Big(C_1(n) \cdot \frac{\sqrt{\varepsilon}}{\delta} \exp\big(-\frac{n\delta^2}{8\varepsilon d}\big)\Big), \\ \log\Big(C_2(n) \cdot \big(\frac{\ell_\delta}{\varepsilon}\big)^{\frac{nd}{2}-1} \exp\big(-\frac{\ell_\delta}{\varepsilon}\big)\Big)\Big).$$

Indem man ε gegen 0 gehen läßt, erhält man

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \mathbb{W}_\varepsilon(C) \;\le\; \max\big(-\frac{n\delta^2}{8d}, -\frac{\ell_\delta}{1}\big)$$

für alle geraden $n \in \mathbb{N}$ und alle $\delta > 0$. Also gilt auch

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \mathbb{W}_\varepsilon(C) \;\le\; -\ell_\delta \qquad \text{für alle } \delta > 0$$

und mit dem zuletzt bewiesenen Lemma wird dies zu

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \mathbb{W}_\varepsilon(C) \;\le\; -\inf_{\omega \in C} \mathrm{I}(\omega).$$

(q.e.d.)

Zum Beweis des zweiten Teils des Satzes von Varadhan zeige ich zunächst zwei Lemmata.

Lemma 2.4 *Die Abbildung* $\omega \mapsto \big(\int_0^1 f(s)\, dW_s\big)(\omega)$ *ist f.s. antisymmetrisch.*

Beweis. Sei $\sigma(\omega) := -\omega$ für alle ω. Dann ist $W \circ \sigma = -W$ und $\mathbb{W} \circ \sigma^{-1} = \mathbb{W}$. Mit l.u.p.-lim bezeichne ich den Grenzwert bezüglich der lokal-gleichmäßigen Konvergenz in Wahrscheinlichkeit (engl.: locally uniform in probability). Nach Proposition 5.5.6 bei v. Weizsäcker und Winkler gilt

$$\begin{aligned} -\int_0^{\cdot} f(s)\, dW_s &= \underset{n\to\infty}{\text{l.u.p.-lim}} \sum_{i=1}^n -f(\tfrac{i-1}{n})\big(W_{\frac{i}{n}\wedge\,\cdot} - W_{\frac{i-1}{n}\wedge\,\cdot}\big) \\ &= \underset{n\to\infty}{\text{l.u.p.-lim}} \sum_{i=1}^n f(\tfrac{i-1}{n})\big(W_{\frac{i}{n}\wedge\,\cdot} \circ \sigma - W_{\frac{i-1}{n}\wedge\,\cdot} \circ \sigma\big) \\ &= \underset{n\to\infty}{\text{l.u.p.-lim}} \Big(\sum_{i=1}^n f(\tfrac{i-1}{n})\big(W_{\frac{i}{n}\wedge\,\cdot} - W_{\frac{i-1}{n}\wedge\,\cdot}\big)\Big) \circ \sigma. \end{aligned}$$

Andererseits gilt auch

$$\Big(\int_0^{\cdot} f(s)\, dW_s\Big) \circ \sigma \;=\; \underset{n\to\infty}{\text{l.u.p.-lim}} \Big(\sum_{i=1}^n f(\tfrac{i-1}{n})\big(W_{\frac{i}{n}\wedge\,\cdot} - W_{\frac{i-1}{n}\wedge\,\cdot}\big)\Big) \circ \sigma$$

und es folgt

$$\Big(\int_0^1 f(s)\, dW_s\Big) \circ \sigma \;=\; -\int_0^1 f(s)\, dW_s \quad \text{f.s.}$$

(q.e.d.)

Die Aussage des folgenden Lemmas bezeichnet man manchmal als „lokales Prinzip der großen Abweichungen“ (engl.: local large deviations priciple). Wenn man das Lemma bewiesen hat, folgt die globale Aussage ohne Mühe.

Lemma 2.5 *Sei $\omega_0 \in C\big([0;1],\mathbb{R}^d\big)$ und $S(\omega_0,\delta)$ die offene Kugel um ω_0 mit Radius δ. Dann gilt*

$$\liminf_{\varepsilon\downarrow 0} \varepsilon \cdot \log \mathbb{W}_\varepsilon\big(S(\omega_0,\delta)\big) \geq -\mathrm{I}(\omega_0).$$

Beweis. Für $\mathrm{I}(\omega_0) = +\infty$ ist die Behauptung klar. Sei also $\mathrm{I}(\omega_0) < \infty$. Wie oben gesehen ist $\frac{W}{\sqrt{\varepsilon}}$ eine Brownsche Bewegung bezüglich $\mathbb{W}_\varepsilon$. Definiert man den (deterministischen) Prozeß H durch

$$H_s \;:=\; -\frac{\dot{\omega}_0(s)}{\sqrt{\varepsilon}} \qquad \text{für alle } s \in [0;1],$$

so ist

$$\int_0^1 H_s^2\,ds \;=\; \frac{1}{\varepsilon}\int_0^1 |\dot{\omega}_0(s)|^2\,ds \;=\; \frac{2\mathrm{I}(\omega_0)}{\varepsilon} \;<\; \infty.$$

Dann erfüllt H nach Bemerkung 10.2.2 bei v. Weizsäcker und Winkler (bzw. nach dem Novikov-Kriterium) die Voraussetzungen aus der Girsanov-Formel (Satz 1.7). Diese liefert dann, daß der Prozeß

$$\left(\frac{W_{t\wedge 1}}{\sqrt{\varepsilon}} + \frac{\omega_0(t\wedge 1)}{\sqrt{\varepsilon}}\right)_{t\geq 0}$$

eine in 1 gestoppte Brownsche Bewegung bezüglich des Maßes Q mit

$$\begin{aligned} dQ \;&=\; \exp\Big(\int_0^1 H_s\,d\tfrac{W_s}{\sqrt{\varepsilon}} - \frac{1}{2}\int_0^1 H_s^2\,ds\Big)\,d\mathbb{W}_\varepsilon \\ &=\; \exp\Big(-\int_0^1 \frac{\dot{\omega}_0(s)}{\varepsilon}\,dW_s - \frac{1}{\varepsilon}\mathrm{I}(\omega_0)\Big)\,d\mathbb{W}_\varepsilon \end{aligned}$$

ist. Daraus folgt

$$\begin{aligned} \mathbb{W}_\varepsilon\,(S(\omega_0,\delta)) \;&=\; \int 1_{S(0,\delta)}(\omega-\omega_0)\,d\mathbb{W}_\varepsilon(\omega) \\ &=\; \int 1_{S(0,\delta)}(\omega+\omega_0-\omega_0)\,dQ(\omega) \\ &=\; \int 1_{S(0,\delta)}(\omega)\exp\Big(-\int_0^1 \frac{\dot{\omega}_0(s)}{\varepsilon}\,dW_s - \frac{1}{\varepsilon}\mathrm{I}(\omega_0)\Big)\,d\mathbb{W}_\varepsilon(\omega) \\ &=\; e^{-\mathrm{I}(\omega_0)/\varepsilon}\,\mathbb{W}_\varepsilon\big(S(0,\delta)\big)\,\frac{\int_{S(0,\delta)}\exp\big(-\int_0^1 \frac{\dot{\omega}_0(s)}{\varepsilon}\,dW_s\big)\,d\mathbb{W}_\varepsilon}{\mathbb{W}_\varepsilon\big(S(0,\delta)\big)} \\ &\overset{\text{Jensen}}{\geq}\; e^{-\mathrm{I}(\omega_0)/\varepsilon}\,\mathbb{W}_\varepsilon\big(S(0,\delta)\big)\,\exp\left(\frac{\int_{S(0,\delta)} - \int_0^1 \frac{\dot{\omega}_0(s)}{\varepsilon}\,dW_s\,d\mathbb{W}_\varepsilon}{\mathbb{W}_\varepsilon(S(0,\delta))}\right) \\ &\overset{\text{Lemma 2.4}}{=}\; e^{-\mathrm{I}(\omega_0)/\varepsilon}\,\mathbb{W}_\varepsilon\big(S(0,\delta)\big)\exp(0). \end{aligned}$$

Man erhält

$$\varepsilon\cdot\log\mathbb{W}_\varepsilon\big(S(\omega_0,\delta)\big) \;\geq\; -\mathrm{I}(\omega_0) + \varepsilon\cdot\log\mathbb{W}\big(S(0,\tfrac{\delta}{\sqrt{\varepsilon}})\big) \qquad \text{für alle } \varepsilon > 0,$$

und mit $\mathbb{W}\big(S(0,\frac{\delta}{\sqrt{\varepsilon}})\big) \to \mathbb{W}(\mathbb{R}^d) = 1$ für $\varepsilon \downarrow 0$ folgt

$$\liminf_{\varepsilon\downarrow 0} \varepsilon\cdot\log\mathbb{W}_\varepsilon\big(S(\omega_0,\delta)\big) \;\geq\; -\mathrm{I}(\omega_0).$$

(q.e.d.)

Beweis (von Teil (ij) des Satzes). Sei $\omega \in G$. Dann existiert ein $\delta > 0$ mit $S(\omega,\delta) \subseteq G$, und nach Lemma 2.5 gilt

$$\liminf_{\varepsilon\downarrow 0} \varepsilon \cdot \log W_\varepsilon(G) \;\geq\; \liminf_{\varepsilon\downarrow 0} \varepsilon \cdot \log W_\varepsilon(S(\omega,\delta)) \;\geq\; -\mathrm{I}(\omega).$$

Bildet man nun das Supremum über alle $\omega \in G$, so erhält man

$$\liminf_{\varepsilon\downarrow 0} \varepsilon \cdot \log W_\varepsilon(G) \quad\geq\quad \sup_{\omega\in G}(-\mathrm{I}(\omega)) \quad=\quad -\inf_{\omega\in G} \mathrm{I}(\omega).$$

(q.e.d.)

Ich beweise noch ein Lemma, das ich auch im Folgenden mehrfach verwenden werde. Dieses ermöglicht es, die Konvergenzrate der Summe zweier Funktionen aus den einzelnen Raten zu bestimmen.

Lemma 2.6 *Seien $f_1, f_2\colon \mathbb{R}_+ \to \mathbb{R}_+$ zwei Funktionen. Dann gilt*

$$\begin{aligned}\liminf_{t\to\infty} \frac{1}{t}\log\big(f_1(t)+f_2(t)\big) \quad&\geq\quad \max_{k:=1,2}\big(\liminf_{t\to\infty}\frac{1}{t}\log f_k(t)\big) \quad \textit{und}\\ \limsup_{t\to\infty} \frac{1}{t}\log\big(f_1(t)+f_2(t)\big) \quad&=\quad \max_{k:=1,2}\big(\limsup_{t\to\infty}\frac{1}{t}\log f_k(t)\big)\end{aligned}$$

Beweis. Es gilt $\log\big(f_1(t)+f_2(t)\big) \geq \log f_k(t)$ für $k := 1,2$ und alle $t \geq 0$, also im Grenzwert auch

$$\liminf_{t\to\infty} \frac{1}{t}\log\big(f_1(t)+f_2(t)\big) \quad\geq\quad \liminf_{t\to\infty}\frac{1}{t}\log f_k(t) \qquad \text{für } k := 1,2$$

und damit

$$\liminf_{t\to\infty} \frac{1}{t}\log\big(f_1(t)+f_2(t)\big) \quad\geq\quad \max_{k:=1,2}\big(\liminf_{t\to\infty}\frac{1}{t}\log f_k(t)\big).$$

Genauso folgt

$$\limsup_{t\to\infty} \frac{1}{t}\log\big(f_1(t)+f_2(t)\big) \quad\geq\quad \max_{k:=1,2}\big(\limsup_{t\to\infty}\frac{1}{t}\log f_k(t)\big).$$

Andererseits gilt auch

$$\begin{aligned}\limsup_{t\to\infty} \frac{1}{t}\log\big(f_1(t)+f_2(t)\big) \quad&\leq\quad \limsup_{t\to\infty}\frac{1}{t}\log\Big(2\cdot\max_{k:=1,2}\big(f_k(t)\big)\Big)\\ &=\quad \limsup_{t\to\infty}\max_{k:=1,2}\Big(\frac{1}{t}\log\big(2\cdot f_k(t)\big)\Big)\\ &=\quad \max_{k:=1,2}\limsup_{t\to\infty}\Big(\frac{1}{t}\log f_k(t)+\frac{\log 2}{t}\Big)\\ &=\quad \max_{k:=1,2}\Big(\limsup_{t\to\infty}\frac{1}{t}\log f_k(t)\Big).\end{aligned}$$

(q.e.d.)

Ist $(W_t)_{t\geq 0}$ der Wienerprozeß, so ist W_1 normalverteilt. Dadurch kann man aus obigem Satz auch einfach Aussagen über die Normalverteilung auf $\mathbb{R}^d$ gewinnen.

Korollar 2.1 *Sei $A \subseteq \mathbb{R}^d$ meßbar, $\mathcal{N}$ bezeichne die d-dimensionale Standardnormalverteilung. Dann gilt*

$$-\inf_{x\in A^\circ} \frac{1}{2}|x|^2 \;\le\; \liminf_{\varepsilon\downarrow 0} \varepsilon\cdot\log\mathcal{N}\Big(\frac{A}{\sqrt{\varepsilon}}\Big) \le \limsup_{\varepsilon\downarrow 0} \varepsilon\cdot\log\mathcal{N}\Big(\frac{A}{\sqrt{\varepsilon}}\Big) \;\le\; -\inf_{x\in\overline{A}} \frac{1}{2}|x|^2.$$

Beweis. Es ist

$$\mathcal{N}\Big(\frac{A}{\sqrt{\varepsilon}}\Big) \;=\; \mathbb{W}\big(W_1 \in \frac{A}{\sqrt{\varepsilon}}\big) \;=\; \mathbb{W}_\varepsilon\big(\{\,\omega \mid \omega_1 \in A\,\}\big).$$

Also kann man den Satz von Varadhan anwenden, um obige Grenzwerte zu berechnen. Wie bereits gesehen, gilt

$$\inf_{\omega_1 = x} \mathrm{I}(\omega) \;=\; \mathrm{I}\big((x\cdot t)_{t\in[0;1]}\big) \;=\; \frac{1}{2}|x|^2 \qquad \text{für alle } x\in\mathbb{R}^d$$

und damit

$$\inf_{\omega_1\in M} \mathrm{I}(\omega) \;=\; \inf_{x\in M}\inf_{\omega_1 = x} \mathrm{I}(\omega) \;=\; \inf_{x\in M}\frac{1}{2}|x|^2 \qquad \text{für alle } M\subseteq\mathbb{R}^d.$$

Wendet man nun Teil (j) des Satzes von Varadhan auf $\{\omega_1 \in \overline{A}\}$ und Teil (ij) auf $\{\omega_1 \in A^\circ\}$ an, so folgt die Behauptung. (q.e.d.)

Der Satz von Varadhan besagt, daß die oben definierten Maße $(\mathbb{W}_\varepsilon)_{\varepsilon>0}$ mit der Funktion I folgende Annahmen erfüllen.

Annahmen. Sei Ω ein metrischer Raum und $\mathcal{F}$ die Borel-σ-Algebra auf Ω. Weiter sei $(\mathbb{W}_\varepsilon)_{\varepsilon>0}$ eine Familie von Wahrscheinlichkeitsmaßen auf Ω und $\mathrm{I}\colon\Omega\to\mathbb{R}$ eine Funktion, die folgende Eigenschaften aufweist:

(j) Für alle $\ell\in[0;\infty)$ ist $\{\,\omega \mid \mathrm{I}(\omega)\le\ell\,\}$ kompakt (d.h. I ist eine gute Ratenfunktion).

(ij) Für jede abgeschlossene Menge $C\subseteq\Omega$ gilt

$$\limsup_{\varepsilon\downarrow 0}\; \varepsilon\cdot\log\mathbb{W}_\varepsilon(C) \le -\inf_{\omega\in C}\mathrm{I}(\omega).$$

(iij) Für jede offene Menge $G\subseteq\Omega$ gilt

$$\liminf_{\varepsilon\downarrow 0}\; \varepsilon\cdot\log\mathbb{W}_\varepsilon(G) \ge -\inf_{\omega\in G}\mathrm{I}(\omega).$$

Satz 2.1 („Varadhan-Lemma"). *Sei $F\colon\Omega\to\mathbb{R}$ beschränkt und stetig. Unter den obigen Annahmen gilt dann folgendes.*

(j) *Für jede abgeschlossene Menge $C\subseteq\Omega$ ist*

$$\limsup_{\varepsilon\downarrow 0}\; \varepsilon\cdot\log\int_C \exp\Big(\frac{F(\omega)}{\varepsilon}\Big)\,d\mathbb{W}_\varepsilon(\omega) \quad\le\quad \sup_{\omega\in C}\big(F(\omega)-\mathrm{I}(\omega)\big) \qquad \textit{und}$$

(ij) *für jede offene Teilmenge $G\subseteq\Omega$ ist*

$$\liminf_{\varepsilon\downarrow 0}\; \varepsilon\cdot\log\int_G \exp\Big(\frac{F(\omega)}{\varepsilon}\Big)\,d\mathbb{W}_\varepsilon(\omega) \quad\ge\quad \sup_{\omega\in G}\big(F(\omega)-\mathrm{I}(\omega)\big).$$

Beweis. Sei $C \subseteq \Omega$ abgeschlossen und $L := \sup_{\omega\in C}\big(F(\omega) - \mathrm{I}(\omega)\big)$. Weiter sei $M \in \mathbb{R}$ eine obere Schranke für F.

Ist $L = -\infty$, so gilt

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\int_C \exp\frac{F(\omega)}{\varepsilon}\,dW_\varepsilon(\omega) \ &\le\ \limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\big(e^{M/\varepsilon}\cdot W_\varepsilon(C)\big)\\ &=\ M\ +\ \limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log W_\varepsilon(C)\\ &\le\ M\ -\ \inf_{\omega\in C}\mathrm{I}(\omega)\ =\ -\infty,\end{aligned}$$

d.h. in diesem Falle ist die erste Behauptung klar. Sei also $L \in \mathbb{R}$.

Weiter sei $\delta > 0$. Für $\omega_0 \in C$ gibt es dann eine Umgebung N von ω_0, so daß

$$F(\omega)\ \le\ F(\omega_0) + \delta \qquad \text{für alle } \omega \in N$$

erfüllt ist. Gäbe es Elemente ω mit $\mathrm{I}(\omega) < \mathrm{I}(\omega_0) - \delta$, die beliebig nahe an ω_0 lägen, so wäre wegen der Halbstetigkeit von I auch $\mathrm{I}(\omega_0)\ \le \mathrm{I}(\omega_0) - \delta$. Da dies ein Widerspruch ist, kann man N so klein wählen, daß auch

$$\mathrm{I}(\omega)\ \ge\ \mathrm{I}(\omega_0) - \delta \qquad \text{für alle } \omega \in N$$

gilt.

Da Ω metrischer Raum ist, gibt es eine offene Menge $G(\omega_0)$ mit

$$\omega_0 \in G(\omega_0) \subseteq \overline{G(\omega_0)} \subseteq N.$$

Es gilt

$$\int_{\overline{G(\omega_0)}} \exp\Big(\frac{F(\omega)}{\varepsilon}\Big)\,dW_\varepsilon(\omega)\ \le\ \exp\Big(\frac{F(\omega_0)+\delta}{\varepsilon}\Big)\cdot W_\varepsilon\left(\overline{G(\omega_0)}\right),$$

also

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\int_{\overline{G(\omega_0)}} \exp\frac{F(\omega)}{\varepsilon}\,dW_\varepsilon(\omega)\ &\le\ F(\omega_0)+\delta+\limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log W_\varepsilon(N)\\ &\le\ F(\omega_0)+\delta-\inf_{\omega\in N}\mathrm{I}(\omega)\\ &\le\ L\ +\ 2\delta.\end{aligned}$$

Ist C kompakt, so gibt es eine endliche Menge $\{\omega_1,\dots,\omega_n\} \subseteq C$ mit

$$C\ \subseteq\ \bigcup_{i=1}^n G(\omega_i)\ =:\ G_\delta.$$

Daher gilt

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\int_C \exp\Big(\frac{F}{\varepsilon}\Big)\,dW_\varepsilon \quad &\le\quad \limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\sum_{i=1}^n\int_{G(\omega_i)}\exp\Big(\frac{F}{\varepsilon}\Big)\,dW_\varepsilon\\ &\overset{\text{Lemma 2.6}}{=}\ \max_{i:=1,\dots,n}\ \limsup_{\varepsilon\downarrow 0}\ \varepsilon\log\int_{G(\omega_i)}\exp\big(\frac{F}{\varepsilon}\big)\,dW_\varepsilon\\ &\le\quad L\ +\ 2\delta\end{aligned}$$

und, da δ beliebig war, auch

$$\limsup_{\varepsilon\downarrow 0}\ \varepsilon\cdot\log\int_C \exp\frac{F(\omega)}{\varepsilon}\,dW_\varepsilon \quad\le\quad L\ =\ \sup_{\omega\in C}\big(F(\omega)-\mathrm{I}(\omega)\big).$$

Diese zeigt die erste Behauptung für kompakte Mengen.

Sei nun C eine beliebige abgeschlossene Menge und $\ell > 0$ so groß, daß $M - \ell \leq L$ gilt. Nach Voraussetzung ist $K_\ell := \{\,\omega \mid \mathrm{I}(\omega) \leq \ell\,\}$ und damit auch $K_\ell \cap C$ kompakt. Wie oben gesehen, gibt es dann für jedes $\delta > 0$ eine offene Menge G_δ, die $K_\ell \cap C$ enthält, und für die

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_{G_\delta} \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\leq\; L + 2\delta$$

gilt. Damit ist auch

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_{G_\delta \cap C} \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\leq\; L + 2\delta.$$

Andererseits gilt

$$\int_{G_\delta^\complement \cap C} \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\leq\; \mathrm{e}^{M/\varepsilon} \cdot \mathbb{W}_\varepsilon(G_\delta^\complement \cap C)$$

und folglich

$$\begin{aligned} \limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_{G_\delta^\complement \cap C} \exp\Big(\frac{F}{\varepsilon}\Big)\, d\mathbb{W}_\varepsilon \;&\leq\; M \;+\; \limsup_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{W}_\varepsilon\big(G_\delta^\complement \cap C\big) \\ &\leq\; M \;-\; \inf_{\omega \in G_\delta^\complement \cap C} \mathrm{I}(\omega). \end{aligned}$$

Wegen $G_\delta \supseteq K_\ell \cap C$ ist $G_\delta^\complement \subseteq K_\ell^\complement \cup C^\complement$, also $G_\delta^\complement \cap C \subseteq K_\ell^\complement \cap C$. Für $\omega \in G_\delta^\complement \cap C$ gilt also $\omega \notin K_\ell$ und damit $\mathrm{I}(\omega) > \ell$. Man erhält

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_{G_\delta^\complement \cap C} \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\leq\; M - \ell \;\leq\; L + 2\delta.$$

Mit Lemma 2.6 folgt aus den beiden letzten Abschnitten zusammen die Aussage

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_C \exp\Big(\frac{F(\omega)}{\varepsilon}\Big)\, d\mathbb{W}_\varepsilon(\omega) \;\leq\; L + 2\delta \qquad \text{für alle } \delta > 0$$

und da δ beliebig war gilt auch hier

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_C \exp\Big(\frac{F(\omega)}{\varepsilon}\Big)\, d\mathbb{W}_\varepsilon(\omega) \;\leq\; \sup_{\omega \in G}\big(F(\omega) - \mathrm{I}(\omega)\big).$$

Damit ist Teil (j) des Satzes bewiesen.

Sei jetzt also $G \subseteq \Omega$ offen, $\omega_0 \in G$ und $\delta > 0$. Dann gibt es eine Umgebung $N \subseteq G$ von ω_0 mit $F(\omega) \geq F(\omega_0) - \delta$ für alle $\omega \in N$. Für diese Umgebung gilt

$$\int_G \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\geq\; \int_N \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\geq\; \exp\Big(\frac{F(\omega_0) - \delta}{\varepsilon}\Big) \cdot \mathbb{W}_\varepsilon(N)$$

und damit auch

$$\begin{aligned} \liminf_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_G \exp\Big(\frac{F}{\varepsilon}\Big)\, d\mathbb{W}_\varepsilon \;&\geq\; F(\omega_0) - \delta + \liminf_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{W}_\varepsilon(N) \\ &\geq\; F(\omega_0) - \delta - \inf_{\omega \in N} \mathrm{I}(\omega) \\ &\geq\; F(\omega_0) - \mathrm{I}(\omega_0) - \delta \qquad \text{für alle } \delta > 0, \end{aligned}$$

also

$$\liminf_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_G \exp\Big(\frac{F}{\varepsilon}\Big)\, d\mathbb{W}_\varepsilon \;\geq\; F(\omega_0) - \mathrm{I}(\omega_0) \qquad \text{für alle } \omega_0 \in G,$$

und indem man das Supremum über alle $\omega_0 \in G$ bildet, folgt die Behauptung

$$\liminf_{\varepsilon \downarrow 0} \varepsilon \cdot \log \int_G \exp \frac{F(\omega)}{\varepsilon}\, d\mathbb{W}_\varepsilon(\omega) \;\geq\; \sup_{\omega \in G}\big(F(\omega) - \mathrm{I}(\omega)\big).$$

(q.e.d.)

3. Diffusionsprozesse mit konstanter Drift

In diesem Kapitel bestimme ich explizit die Konvergenzrate des Bayesrisikos für die Unterscheidung zweier Diffusionsprozesse mit konstanter Drift. Ich gebe zwei unterschiedliche Methoden zur Bestimmung dieser Rate an. Das erste Verfahren verwendet dazu Korollar 2.1. Das zweite Verfahren bestimmt den Wert mit Hilfe des Varadhan-Lemmas.

Setze $\Theta = \{1,2\}$. Für $\vartheta \in \Theta$ sei $b^\vartheta \in \mathbb{R}^d$ und X^ϑ eine Lösung der stochastischen Differentialgleichung

$$dX^\vartheta = b^\vartheta \bullet dt + dB$$

mit $X_0 = 0$ f.s. Dann ist X^ϑ ein Diffusionsprozeß mit konstanter Drift b^ϑ. Für $T \geq 0$ setze $P_T^\vartheta := \mathcal{L}(X^\vartheta)\big|_{\mathcal{F}_T}$. Mit Korollar 1.1 kann man die Dichte von P_T^ϑ bezüglich des Wienermaßes bestimmen: Da die Drift konstant ist, ist $Z_{\cdot\wedge T}$ mit dem dortigen Z ein Martingal. Definiert man weiter

$$\varphi_T^\vartheta(\omega) := \exp\big(b^\vartheta \cdot \omega_T - \frac{1}{2}T|b^\vartheta|^2\big) \qquad \text{für alle } \omega \in C\big([0;\infty),\mathbb{R}^d\big),$$

so ist

$$\varphi_T^\vartheta(X) = \exp\Big(\int_0^T b^\vartheta\, dX^\vartheta - \frac{1}{2}\int_0^T |b^\vartheta|^2\, ds\Big) \qquad P\text{-fast sicher}$$

und das Korollar liefert

$$dP_T^\vartheta = \varphi_T^\vartheta\, dW \qquad \text{für alle } T \geq 0.$$

Nun versucht man folgendes Problem zu lösen: Man beobachtet einen Prozeß X, von dem man weiß, daß er sich wie X^ϑ für ein zufällig gewähltes $\vartheta \in \Theta$ verhält, und versucht aus dem beobachteten Pfad $t \mapsto X_t(\omega)$ Informationen über das ϑ zu gewinnen. Dazu verwendet man einen „maximum-likelyhood Test". Sei zum Beispiel $\mathrm{Ws}(b = b^1) = \lambda_1$ und $\mathrm{Ws}(b = b^2) = \lambda_2$. Hat man dann bis zur Zeit T den Pfad $\xi\colon [0;T] \to \mathbb{R}^d$ beobachtet, so nimmt man $\vartheta = 1$ an, falls $\lambda_1\varphi_T^1(\xi) \geq \lambda_2\varphi_T^2(\xi)$ ist, und $\vartheta = 2$ sonst.

Definition. Für $\lambda \in \mathcal{M}_1(\Theta)$ und $T \geq 0$ setze $P_T := \lambda_1 P_T^1 + \lambda_2 P_T^2$. Dann ist das **Bayesrisiko** $B(\lambda,T)$ durch

$$B(\lambda,T) := \int \min_{\vartheta\in\Theta}\Big(\frac{d\lambda_\vartheta P_T^\vartheta}{dP_T}(\omega)\Big)\, dP_T(\omega)$$

definiert.

P_T ist die Verteilung der Pfade, die man bei zufälliger Drift b bis zur Zeit T beobachtet. Es ist

$$\frac{d\lambda_\vartheta P_T^\vartheta}{dP_T} = \frac{\lambda_\vartheta \varphi_T^\vartheta}{\lambda_1\varphi_T^1 + \lambda_2\varphi_T^2}. \tag{3.1}$$

Der Test liefert das falsche Ergebnis, falls $b = b^1$ und $X \in \{\lambda_1\varphi_T^1 < \lambda_2\varphi_T^2\}$ oder falls $b = b^2$ und $X \in \{\lambda_1\varphi_T^1 \geq \lambda_2\varphi_T^2\}$ ist. Dies geschieht mit der Wahrscheinlichkeit

$$\begin{aligned}
\mathrm{Ws}(b = {} & b^1)\cdot P_T^1(\lambda_1\varphi_T^1 < \lambda_2\varphi_T^2) + \mathrm{Ws}(b = b^2)\cdot P_T^2(\lambda_1\varphi_T^1 \geq \lambda_2\varphi_T^2) \\
&= \lambda_1 \cdot \int_{\{\lambda_1\varphi_T^1 < \lambda_2\varphi_T^2\}} \frac{dP_T^1}{dP_T}\, dP_T + \lambda_2 \cdot \int_{\{\lambda_1\varphi_T^1 \geq \lambda_2\varphi_T^2\}} \frac{dP_T^2}{dP_T}\, dP_T \\
&\overset{(3.1)}{=} \int_{\{\frac{d\lambda_1 P_T^1}{dP_T} < \frac{d\lambda_2 P_T^2}{dP_T}\}} \frac{d\,\lambda_1 P_T^1}{dP_T}\, dP_T + \int_{\{\frac{d\lambda_1 P_T^1}{dP_T} \geq \frac{d\lambda_2 P_T^2}{dP_T}\}} \frac{d\,\lambda_2 P_T^2}{dP_T}\, dP_T \\
&= B(\lambda,T).
\end{aligned} \tag{3.2}$$

Also ist das Bayesrisiko tatsächlich die Fehlerwahrscheinlichkeit des oben beschriebenen Tests. Im Folgenden zeige ich, daß diese Wahrscheinlichkeit für $b^1 \neq b^2$ exponentiell schnell gegen Null konvergiert, und daß die Konvergenzrate nicht von λ abhängt.

In der speziellen Situation konstanter Drift erhält man für das Bayesrisiko folgende Darstellung.

Lemma 3.1 *Sei* $\lambda \in \mathcal{M}_1(\Theta)$ *mit* $\lambda_1, \lambda_2 \neq 0$ *und* $T \in [0;\infty)$. *Setze* $\eta_T := \frac{\log \lambda_1/\lambda_2}{T}$, *sowie*

$$\begin{aligned} A_T^1 &:= \big\{ x \in \mathbb{R}^d \mid x \cdot (b^2 - b^1) > \frac{b^2+b^1}{2} \cdot (b^2 - b^1) + \eta_T \big\} \quad \textit{und} \\ A_T^2 &:= \big\{ x \in \mathbb{R}^d \mid x \cdot (b^2 - b^1) \leq \frac{b^2+b^1}{2} \cdot (b^2 - b^1) + \eta_T \big\}. \end{aligned}$$

Dann gilt

$$B(\lambda, T) \quad = \quad \lambda_1 P_T^1\big(\frac{W_T}{T} \in A_T^1\big) + \lambda_2 P_T^2\big(\frac{W_T}{T} \in A_T^2\big).$$

Beweis. Es gilt

$$\begin{aligned} \lambda_2 \varphi_T^2(W) > \lambda_1 \varphi_T^1(W) \quad &\Longleftrightarrow \quad \lambda_2 \exp\big(b^2 \cdot W_T - \frac{1}{2}|b^2|^2 T \big) \\ &\qquad\qquad > \lambda_1 \exp\big(b^1 \cdot W_T - \frac{1}{2}|b^1|^2 T \big) \\ &\Longleftrightarrow \quad \exp\Big((b^2 - b^1) \cdot W_T - \frac{1}{2}\big(|b^2|^2 - |b^1|^2\big)T \Big) > \frac{\lambda_1}{\lambda_2} \\ &\Longleftrightarrow \quad (b^2 - b^1) \cdot W_T - \frac{1}{2}\big(|b^2|^2 - |b^1|^2\big)T > \log \frac{\lambda_1}{\lambda_2} \\ &\Longleftrightarrow \quad \frac{W_T}{T} \cdot (b^2 - b^1) > \frac{b^2+b^1}{2} \cdot (b^2 - b^1) + \frac{log \lambda_1/\lambda_2}{T} \\ &\Longleftrightarrow \quad \frac{W_T}{T} \in A_T^1. \end{aligned}$$

Aus (3.2) erhält man daher

$$B(\lambda, T) = \lambda_1 P_T^1\big(\frac{W_T}{T} \in A_T^1\big) + \lambda_2 P_T^2\big(\frac{W_T}{T} \in A_T^2\big).$$

(q.e.d.)

Definiere $\sigma: C\big([0;\infty), \mathbb{R}^d\big) \to C\big([0;\infty), \mathbb{R}^d\big)$ durch

$$\big(\sigma(\omega)\big)_t \quad := \quad \frac{1}{\sqrt{T}} \omega_{t \cdot T}.$$

Dann ist $\omega_T = \sqrt{T}\sigma(\omega)_1$ und für meßbare Mengen $A \subseteq \mathbb{R}^d$ gilt

$$\sigma(\omega)_1 \in \sqrt{T} \cdot A \iff \frac{\omega_T}{T} \in A.$$

Mittels dieser Transformation kann man Aussagen über die P_T^ϑ in Aussagen über die Normalverteilung auf $\mathbb{R}^d$ umformen. Dadurch läßt sich dann Korollar 2.1 anwenden. Lemma 2.1 liefert $\mathbb{W} \circ (\sigma^{-1}) = \mathbb{W}$ und durch Substitution erhält man

$$\begin{aligned} P_T^\vartheta\big(\frac{W_T}{T} \in A\big) \quad &= \quad \int_{\{\, \omega \mid \frac{\omega_T}{T} \in A \,\}} \exp\big(b^\vartheta \cdot \omega_T - \frac{1}{2} T\, |b^\vartheta|^2 \big)\, d\mathbb{W}(\omega) \\ &\overset{\text{Subst.}}{=} \quad \int_{\{\omega_1 \in \sqrt{T}A\}} \exp\big(\sqrt{T}\omega_1 \cdot b^\vartheta - \frac{1}{2} T |b^\vartheta|^2 \big)\, d\mathbb{W}(\omega) \\ &= \quad \frac{1}{(2\pi)^{d/2}} \int_{\sqrt{T} \cdot A} \exp\big(\sqrt{T}x \cdot b^\vartheta - \frac{1}{2} T |b^\vartheta|^2 \big) exp\big(-\frac{x^2}{2}\big)\, dx \\ &\overset{y = x - \sqrt{T}b^\vartheta}{=} \quad \frac{1}{(2\pi)^{d/2}} \int_{\sqrt{T} \cdot (A - b^\vartheta)} exp\big(-\frac{y^2}{2}\big)\, dy \\ &= \quad \mathcal{N}\big(\sqrt{T}(A - b^\vartheta)\big), \end{aligned}$$

wobei $\mathcal{N}$ die Standardnormalverteilung auf dem $\mathbb{R}^d$ bezeichnet.

Zur Abkürzung bezeichne ich ab hier den Halbraum $\{ x \in \mathbb{R}^d \mid n \cdot x > \alpha \}$ für $n \in \mathbb{R}^d \setminus \{0\}$ und $\alpha \in \mathbb{R}$ durch $H(n,\alpha)$ und durch $\overline{H}(n,\alpha)$ den Abschluß von $H(n,\alpha)$. Ist dann $\alpha < \beta$, so gilt $H(n,\beta) \subset H(n,\alpha)$ und für $y \in \mathbb{R}^d$ ist

$$\begin{aligned} H(n,\alpha) - b &= \{ x - b \mid n \cdot x > \alpha \} \\ &= \{ x \mid n \cdot (x+b) > \alpha \} \\ &= H(n, \alpha - n \cdot b). \end{aligned}$$

Außerdem gilt $H(\lambda n, \lambda\alpha) = H(n,\alpha)$ für alle $\lambda > 0$. Ist $|n| = 1$ und $\alpha \geq 0$, so gilt $\inf_{x \in H(n,\alpha)} |x| = \inf_{x \in \overline{H}(n,\alpha)} |x| = \alpha$, und es folgt

$$-\frac{1}{2} \inf_{x \in H(n,\alpha)} |x|^2 = -\frac{1}{2} \inf_{x \in \overline{H}(n,\alpha)} |x|^2 = -\frac{1}{2}\alpha^2.$$

Mit den Bezeichnungen aus Lemma 3.1 ist weiter

$$A_T^1 = H\big(b^2 - b^1, \frac{1}{2}(b^2 + b^1)(b^2 - b^1) + \eta_T\big).$$

Für $\eta > 0$ definiere

$$\begin{aligned} H_\eta^+ &:= H\big(b^2 - b^1, \frac{1}{2}(b^2 + b^1)(b^2 - b^1) + |b^2 - b^1| \cdot \eta\big) \quad \text{und} \\ H_\eta^- &:= H\big(b^2 - b^1, \frac{1}{2}(b^2 + b^1)(b^2 - b^1) - |b^2 - b^1| \cdot \eta\big). \end{aligned}$$

Wegen $\eta_T \to 0$ für $T \to \infty$ ist schließlich $\eta_T \in \big[-|b^2 - b^1| \cdot \eta \ ; \ +|b^2 - b^1| \cdot \eta\big]$. Es gilt also $H_\eta^+ \subseteq A_T^1 \subseteq H_\eta^-$ und daher

$$\begin{aligned} \mathcal{N}\big(\sqrt{T} \cdot (H_\eta^+ - b^1)\big) &= P_T^1\big(\frac{W_T}{T} \in H_\eta^+\big) \\ &\leq P_T^1\big(\frac{W_T}{T} \in A_T^1\big) \\ &\leq P_T^1\big(\frac{W_T}{T} \in H_\eta^-\big) = \mathcal{N}\big(\sqrt{T} \cdot (H_\eta^- - b^1)\big), \end{aligned}$$

wobei

$$\begin{aligned} H_\eta^+ - b^1 &= H\big(b^2 - b^1, \frac{1}{2}|b^2 - b^1|^2 + |b^2 - b^1| \cdot \eta\big) \\ &= H\Big(\frac{b^2 - b^1}{|b^2 - b^1|}, \frac{1}{2}|b^2 - b^1| + \eta\Big) \end{aligned}$$

und

$$H_\eta^- - b^1 = H\Big(\frac{b^2 - b^1}{|b^2 - b^1|}, \frac{1}{2}|b^2 - b^1| - \eta\Big)$$

ist. Aus Korollar 2.1 ergibt sich nun die gesuchte Rate: für alle hinreichend kleinen η gilt

$$\begin{aligned} -\frac{1}{2}\big(\frac{|b^2 - b^1|}{2} + \eta\big)^2 &\leq \liminf_{T \to \infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot (H_\eta^+ - b^1)\big) \\ &\leq \liminf_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in A_T^1\big) \\ &\leq \limsup_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in A_T^1\big) \\ &\leq \limsup_{T \to \infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot (H_\eta^- - b^1)\big) \\ &\leq -\frac{1}{2}\big(\frac{|b^2 - b^1|}{2} - \eta\big)^2 \end{aligned}$$

und daher

$$\lim_{T\to\infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in A_T^1\big) \;=\; -\frac{1}{2}\big(\frac{b^2-b^1}{2}\big)^2.$$

Analog folgt wegen

$$A_T^2 \;=\; \overline{H}\big(b^1-b^2, \frac{1}{2}(b^1+b^2)(b^1-b^2) - \eta_T\big)$$

auch

$$A_T^2 - b^2 \;=\; \overline{H}\big(b^1-b^2, \frac{1}{2}(b^2-b^1)^2 - \eta_T\big),$$

und für P_T^2 erhält man ebenfalls

$$\lim_{T\to\infty} \frac{1}{T} \log P_T^2\big(\frac{W_T}{T} \in A_T^2\big) \;=\; -\frac{1}{2}\big(\frac{b^2-b^1}{2}\big)^2.$$

Lemma 2.6 liefert

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{T} \log B(\lambda, T) &= \max_{\vartheta:=1,2} \lim_{T\to\infty} \frac{1}{T} \log \lambda_\vartheta P_T^\vartheta(A^\vartheta) \\
&= \max_{\vartheta:=1,2} \lim_{T\to\infty} \Big(\frac{\log \lambda_\vartheta}{T} + \frac{\log P_T^\vartheta(A^\vartheta)}{T}\Big) \\
&= \max_{\vartheta:=1,2} \lim_{T\to\infty} \frac{1}{T} \log P_T^\vartheta(A^\vartheta) \\
&= -\frac{\big(b^2-b^1\big)^2}{8}.
\end{aligned}$$

Dies ist die Konvergenzrate des Bayesrisikos für Diffusionsprozesse mit konstanter Drift.

Wie oben bereits angekündigt gebe ich noch eine zweite Methode zur Bestimmung dieser Rate an: Anstelle der Transformation σ verwende ich hier die Abbildung τ aus dem letzten Kapitel. Dort war $W \circ \tau^{-1} = W_{1/T}$ und $\tau(\omega)_1 = \frac{\omega_T}{T}$. Für meßbare $A \subseteq \mathbb{R}^d$ gilt dann

$$\tau(\omega)_1 \in A \iff \frac{\omega_T}{T} \in A,$$

und man erhält

$$\begin{aligned}
P_T^\vartheta\big(\frac{W_T}{T} \in A\big) &= \int_{\{\,\omega \,\mid\, \frac{\omega_T}{T} \in A\,\}} \exp\big(b^\vartheta \cdot \omega_T - \frac{1}{2} T\,|b^\vartheta|^2\big)\, dW(\omega) \\
&\overset{\text{Subst.}}{=} \int_{\{\omega_1 \in A\}} \exp\big(\frac{\omega_1 b^\vartheta - \frac{1}{2}\,|b^\vartheta|^2}{1/T}\big)\, dW_{1/T}(\omega).
\end{aligned} \tag{3.3}$$

Zur Bestimmung der Konvergenzrate von $P_T^\vartheta\big(\frac{W_T}{T} \in A\big)$ möchte ich nun das Varadhan-Lemma (Satz 2.1) auf $F\colon C[0;1] \to \mathbb{R}$ mit

$$F(\omega) \;:=\; \omega_1 \cdot b^\vartheta - \frac{1}{2}|b^\vartheta|^2 \qquad \text{für alle } \omega \in C[0;1]$$

anwenden. Dies ist aber nicht direkt möglich, da diese Funktion nicht beschränkt ist. Daher benötigt man ein zusätzliches Abschneideargument.

Lemma. *Sei $M \subseteq \mathbb{R}^d$ meßbar mit $M \neq \emptyset$ und $\overline{M^\circ} = \overline{M}$. Weiter sei $r > 0$ so, daß $M \cap S(0,r) \neq \emptyset$ ist. Dann gilt*

$$\lim_{T\to\infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot M\big) \;=\; \lim_{T\to\infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot (M \cap S(0,r))\big).$$

Beweis. Für $\overline{M^\circ} = \overline{M}$ wird Korollar 2.1 zu

$$\lim_{T\to\infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot M\big) \;=\; -\inf_{x\in M} \frac{1}{2}|x|^2.$$

Sei nun $x \in M \cap S(0,r)$, also $x \in M \subseteq \overline{M^\circ}$ und $|x| < r$. Dann gibt es für alle $n \in \mathbb{N}$ ein $x_n \in M^\circ$ mit $|x_n - x| < \max(1/n, r - |x|)$ und wegen $x_n \in M^\circ \cap S(0,r) \subseteq \big(M \cap S(0,r)\big)^\circ$, sowie $x = \lim_{n\to\infty} x_n$ liegt x im Abschluß von $\big(M \cap S(0,r)\big)^\circ$. Daher gilt

$$\overline{M \cap S(0,r)} \;=\; \overline{\big(M \cap S(0,r)\big)^\circ}$$

und damit

$$\lim_{T\to\infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot (M \cap S(0,r))\big) \;=\; -\inf_{x\in M\cap S(0,r)} \frac{1}{2}|x|^2.$$

Wegen $M \cap S(0,r) \neq \emptyset$ ist $\inf_{x\in M} \frac{1}{2}|x|^2 < \frac{1}{2}r^2$. Für $y \in M \setminus S(0,r)$ gilt aber $\frac{1}{2}|y|^2 \geq \frac{1}{2}r^2$, d.h. diese y spielen bei der Bestimmung des Infimums keine Rolle. Also ist

$$\begin{aligned}\lim_{T\to\infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot M\big) \;&=\; -\inf_{x\in M} \frac{1}{2}|x|^2 \\ &=\; -\inf_{x\in M\cap S(0,r)} \frac{1}{2}|x|^2 \;=\; \lim_{T\to\infty} \frac{1}{T} \log \mathcal{N}\big(\sqrt{T} \cdot (M \cap S(0,r))\big).\end{aligned}$$

(q.e.d.)

Sei also $A \subseteq \mathbb{R}^d$ meßbar mit $A \neq \emptyset$ und $\overline{A} = \overline{A^\circ}$. Weiter sei $r > 0$ so groß, daß $A \cap S(b^\vartheta, r) \neq \emptyset$ gilt. Dann ist

$$P_T^\vartheta\big(\frac{W_T}{T} \in A\big) \;=\; \mathcal{N}\big(\sqrt{T}(A - b^\vartheta)\big)$$

und wegen $\big(A \cap S(b^\vartheta, r)\big) - b^\vartheta = (A - b^\vartheta) \cap S(0,r)$ auch

$$P_T^\vartheta\big(\frac{W_T}{T} \in A \cap S(b^\vartheta, r)\big) \;=\; \mathcal{N}\Big(\sqrt{T}\big((A - b^\vartheta) \cap S(0,r)\big)\Big).$$

Das Lemma zeigt, daß beide Wahrscheinlichkeiten dieselbe Konvergenzrate besitzen, daß also

$$\begin{aligned}\lim_{T\to\infty} \frac{1}{T} \log P_T^\vartheta\big(\frac{W_T}{T} \in A\big) \;&=\; \lim_{T\to\infty} \frac{1}{T} \log P_T^\vartheta\big(\frac{W_T}{T} \in A \cap S(b^\vartheta, r)\big) \\ &\overset{(3.3)}{=}\; \frac{1}{T} \log \int_{\{\omega_1 \in A\cap S(b^\vartheta, r)\}} \exp\big(\frac{F(\omega)}{1/T}\big)\, dW_{1/T}(\omega)\end{aligned}$$

gilt, falls einer dieser Grenzwerte existiert.

Da F auf der Menge $\{\omega_1 \in A \cap S(b^\vartheta, r)\}$ beschränkt ist, kann man F durch eine „abgeschnittene“ Funktion F' ersetzen. Sei also

$$F'(\omega) \;=\; \begin{cases} K & \text{falls } F(\omega) > K, \\ -K & \text{falls } F(\omega) < -K \text{ und} \\ F(\omega) & \text{sonst} \end{cases}$$

für ein hinreichend großes $K \in \mathbb{R}$. Dann ist F' beschränkt und stetig, und auf der Menge $\{\omega_1 \in A \cap S(b^\vartheta, r)\}$ gilt $F' = F$. Daher kann man das Varadhan-Lemma auf F' anwenden.

Ich gebe zunächst eine Vorüberlegung an, mit deren Hilfe man die Rate in eine günstigere Form bringen kann: Für $x \in \mathbb{R}^d$ setze $\Omega_x := \{\, \omega \in C[0;1] \bigm| \omega_1 = x \,\}$. Für $\omega \in \Omega_x$ ist dann $F(\omega) = x \cdot b^\vartheta - \frac{1}{2}|b^\vartheta|^2$, und es folgt

$$\begin{aligned}
\sup_{\{\,\omega \mid \omega_1 \in A \cap S(b^\vartheta, r)\,\}} \big(F'(\omega) - I(\omega)\big) &= \sup_{x \in A \cap S(b^\vartheta, r)} \sup_{\omega \in \Omega_x} \big(F(\omega) - I(\omega)\big) \\
&= \sup_{x \in A \cap S(b^\vartheta, r)} x \cdot b^\vartheta - \frac{1}{2}|b^\vartheta|^2 - \inf_{\omega \in \Omega_x} I(\omega) \\
&\overset{(2.1)}{=} \sup_{x \in A \cap S(b^\vartheta, r)} x \cdot b^\vartheta - \frac{1}{2}|b^\vartheta|^2 - \frac{1}{2}|x|^2 \\
&= -\frac{1}{2} \inf_{x \in A \cap S(b^\vartheta, r)} |x - b^\vartheta|^2 \\
&= -\frac{1}{2} \inf_{x \in A} |x - b^\vartheta|^2 \\
&= -\frac{1}{2} \inf_{x \in A - b^\vartheta} |x|^2 .
\end{aligned}$$

Die Abbildung $x \mapsto \frac{1}{2}|x|^2$ ist stetig. Daher sind für $\overline{A} = \overline{A^\circ}$ die obere und die untere Rate aus dem Varadhan-Lemma gleich. Es gilt dann

$$\begin{aligned}
\lim_{T \to \infty} \frac{1}{T} \log P_T^\vartheta\big(\frac{W_T}{T} \in A\big) &= \lim_{T \to \infty} \frac{1}{T} \log \int_{\{\omega_1 \in A \cap S(b^\vartheta, r)\}} \exp\big(\frac{F'(\omega)}{1/T}\big)\, dW_{1/T} \\
&= -\frac{1}{2} \inf_{x \in A - b^\vartheta} |x|^2 .
\end{aligned}$$

Nun fährt man ähnlich fort, wie im ersten Teil des Kapitels. Ich verwende wieder die dortigen Bezeichnungen. Für hinreichend kleine $\eta > 0$ folgt dann

$$\begin{aligned}
-\frac{1}{2}\big(\frac{|b^2 - b^1|}{2} + \eta\big)^2 &= -\frac{1}{2} \inf_{x \in H_\eta^+ - b^\vartheta} |x|^2 \\
&= \lim_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in H_\eta^+\big) \\
&\le \liminf_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in A_T^1\big)
\end{aligned}$$

und

$$\begin{aligned}
\limsup_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in A_T^1\big) &\le \lim_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in H_\eta^-\big) \\
&= -\frac{1}{2} \inf_{x \in H_\eta^- - b^\vartheta} |x|^2 \\
&= -\frac{1}{2}\big(\frac{|b^2 - b^1|}{2} - \eta\big)^2 .
\end{aligned}$$

Also erhält man hier ebenfalls

$$\lim_{T \to \infty} \frac{1}{T} \log P_T^1\big(\frac{W_T}{T} \in A_T^1\big) = -\frac{1}{2}\big(\frac{b^2 - b^1}{2}\big)^2$$

und durch eine analoge Rechnung auch

$$\lim_{T \to \infty} \frac{1}{T} \log P_T^2\big(\frac{W_T}{T} \in A_T^2\big) = -\frac{1}{2}\big(\frac{b^2 - b^1}{2}\big)^2 .$$

Lemma 2.6 liefert wieder die Konvergenzrate

$$\lim_{T \to \infty} \frac{1}{T} \log B(\lambda, T) = -\frac{\big(b^2 - b^1\big)^2}{8} .$$

4. Prozesse mit stationärer Verteilung

Dieses Kapitel beschreibt Eigenschaften von Diffusionsprozessen, die eine stationäre Verteilung besitzen. Ich gebe mehrere unterschiedliche Charakterisierungen stationärer Verteilungen an. Um die Operatorhalbgruppe aus Satz 1.6 verwenden zu können, verlange ich hier meist, daß Drift und Diffusionsmatrix lipschitzstetig sind.

Definition. Sei X ein Diffusionsprozeß mit Werten in $\mathbb{R}^d$ und $\mathcal{L}(X_0) = \mu$. Ist dann auch $\mathcal{L}(X_t) = \mu$ für alle $t > 0$, so heißt μ **stationäre Verteilung** oder **stationäres Maß** von X.

Das folgende Lemma charakterisiert die Existenz stationärer Verteilungen mit Hilfe der Halbgruppe aus dem ersten Kapitel.

Lemma 4.1 *Sei $\mu \in \mathcal{M}_1(\mathbb{R}^d)$, X eine Lösung von* (1.1) *mit $\mathcal{L}(X_0) = \mu$ und $(P_t)_{t\geq 0}$ die zugehörige Operatorhalbgruppe. Die Drift $b\colon \mathbb{R}^d \to \mathbb{R}^d$ und der Diffusionskoeffizient $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times n}$ seien lipschitzstetig. Dann sind folgende Aussagen äquivalent.*

(j) *μ ist stationäre Verteilung von X.*

(ij) *Es gilt*

$$\int \varphi\, d\mu \;=\; \int P_t\varphi\, d\mu \qquad \text{für alle } t \geq 0,\ \varphi \in C_b(\mathbb{R}^d).$$

(iij) *Es gilt*

$$\int \varphi\, d\mu \;=\; \int P_t\varphi\, d\mu \qquad \text{für alle } t \geq 0,\ \varphi \in C_c^2(\mathbb{R}^d).$$

Beweis. Sei $t \geq 0$ und $\varphi \in C_b(\mathbb{R}^d)$. Dann gilt

$$E\big(\varphi(X_t)\big) = E\big(E(\varphi(X_t)|\mathcal{F}_0^+)\big) = E\big(E^{X_0}(\varphi(W_t))\big) = \int P_t\varphi\, d\mu.$$

Besitzt X eine stationäre Verteilung, so ist $E\big(\varphi(X_t)\big) = E\big(\varphi(X_0)\big)$, und wegen $P_0\varphi = \varphi$ folgt die Bedingung (ij).

Aus (ij) folgt trivialerweise die Bedingung (iij).

Gilt nun (iij), so erhält man aus obiger Gleichungskette $E\big(\varphi(X_t)\big) = E\big(\varphi(X_0)\big)$ für alle $\varphi \in C_c^2(\mathbb{R}^d)$ und es folgt $\mathcal{L}(X_t) = \mathcal{L}(X_0) = \mu$ für alle $t \geq 0$. Also ist Eigenschaft (j) erfüllt.

(q.e.d.)

Ist μ ein Wahrscheinlichkeitsmaß, so ist die Abbildung

$$f \;\longmapsto\; \int_{\mathbb{R}^d} f\, d\mu \qquad \text{für alle } f \in C_b(\mathbb{R}^d)$$

linear und wegen $|\int f\, d\mu| \leq \int |f|\, d\mu \leq \|f\|_\infty$ auch stetig. Dadurch kann man μ als Element des Dualraums $(C_b(\mathbb{R}^d))^*$ auffassen. Auf diesem Raum ist der zum Generator L adjungierte Operator L^* durch

$$\begin{array}{l} \mu \in \mathcal{D}(L^*) \\ L^*\mu = \mu^* \end{array} \qquad \Longleftrightarrow \qquad \begin{array}{l} \mu, \mu^* \in \big(C_b(\mathbb{R}^d)\big)^* \\ \mu(Lf) = \mu^*(f) \text{ für alle } f \in \mathcal{D}(L) \end{array}$$

definiert.

Satz. *Sei X eine Lösung der stochastischen Differentialgleichung* (1.1). *Die Drift $b\colon \mathbb{R}^d \to \mathbb{R}^d$ und der Diffusionskoeffizient $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times n}$ seien lipschitzstetig. Weiter sei L der Generator der zugehörigen Operatorhalbgruppe und $\mu \in \mathcal{M}_1(\mathbb{R}^d)$. Dann sind folgende Bedingungen äquivalent:*

(j) *μ ist stationäre Verteilung von X.*

(ij) *Es ist $L^*\mu = 0$.*

Beweis. Es gilt

$$\Big|\frac{1}{h}\Big(\int P_h\varphi\,d\mu - \int \varphi\,d\mu\Big) - \int L\varphi\,d\mu\Big| \;\le\; \int \Big|\frac{P_h\varphi-\varphi}{h} - L\varphi\Big|\,d\mu \\ \le\; \Big\|\frac{P_h\varphi-\varphi}{h} - L\varphi\Big\|_\infty$$

und nach Satz 1.6 geht die rechte Seite für $h \downarrow 0$ gegen Null. Also konvergiert auch die linke Seite dieser Gleichung gegen Null und es folgt

$$\frac{d}{dt}\int P_t\varphi\,d\mu \;=\; \lim_{h\downarrow 0}\frac{1}{h}\Big(\int P_{t+h}\varphi\,d\mu - \int P_t\varphi\,d\mu\Big) \;=\; \int LP_t\varphi\,d\mu$$

für alle $\varphi \in \mathcal{D}(L)$.

Sei nun μ stationäre Verteilung von X. Dann gilt nach dem voranstehenden Lemma

$$\int P_t\varphi\,d\mu \;=\; \int \varphi\,d\mu \qquad \text{für alle } t\ge 0,\ \varphi\in C_b(\mathbb{R}^d)$$

und es folgt

$$\int L\varphi\,d\mu \;=\; \frac{d}{dt}\int P_t\varphi\,d\mu\Big|_{t=0} \;=\; 0 \qquad \text{für alle } \varphi\in\mathcal{D}(L).$$

Damit ist $\mu \in \mathcal{D}(L^*)$ und $L^*\mu = 0$.

Sei umgekehrt $L^*\mu = 0$, also

$$\int L\varphi\,d\mu = 0 \qquad \text{für alle } \varphi\in\mathcal{D}(L).$$

Dann ist

$$\frac{d}{dt}\int P_t\varphi\,d\mu \;=\; \int L(P_t\varphi)\,d\mu \;=\; 0$$

und folglich

$$\int P_t\varphi\,d\mu \;=\; \int P_0\varphi\,d\mu \;=\; \int\varphi\,d\mu \qquad \text{für alle } \varphi\in\mathcal{D}(L).$$

Nach Satz 1.6 ist $C_c^2(\mathbb{R}^d) \subseteq \mathcal{D}(L)$ und mit dem Lemma folgt, daß μ stationäre Verteilung von X ist. (q.e.d.)

Für Prozesse mit stationärer Verteilung $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ kann man die Halbgruppe $(P_t)_{t\ge 0}$ aus Satz 1.6 zu einer Halbgruppe $(\bar{P}_t)_{t\ge 0}$ auf dem Hilbertraum $L^2(\mathbb{R}^d,\mu)$ fortsetzen. Diese Fortsetzung wird in folgendem Satz konstruiert.

Notation. Ist μ ein Wahrscheinlichkeitsmaß auf $\mathbb{R}^d$, so schreibe ich zur Abkürzung $L^2(\mu)$ für $L^2(\mathbb{R}^d,\mu)$.

Satz 4.1 *Seien $b\colon \mathbb{R}^d \to \mathbb{R}^d$ und $\sigma\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n}$ lipschitzstetig und so geartet, daß Lösungen der stochastischen Differentialgleichung* (1.1) *eine stationäre Verteilung μ besitzen. Dann wird durch*

$$\big(\bar{P}_t f\big)(x) \;:=\; E^x\big(f(W_t)\big) \qquad \textit{für alle } x\in\mathbb{R}^d,\ f\in L^2(\mu)$$

eine stark stetige Kontraktionshalbgruppe auf $L^2(\mu)$ definiert.

Beweis. Sei X eine Lösung von (1.1) mit $\mathcal{L}(X_0) = \mu$ und $f \in L^2(\mu)$. Dann ist

$$\begin{aligned} E\Big(E^{X_0}\big(|f(W_t)|\big)\Big) &= E\big(|f(X_t)|\big) \\ &\overset{\text{Jensen}}{\leq} \sqrt{E\big(|f(X_t)|^2\big)} \quad = \quad \|f\|_{\mu,2} \quad < \quad \infty, \end{aligned}$$

also $E^x\big(|f(W_t)|\big) < \infty$ für μ-fast alle $x \in \mathbb{R}^d$, d.h. für μ-fast alle $x \in \mathbb{R}^d$ ist $f(W_t)$ bezüglich P^x integrierbar.

Sei nun $t \geq 0$. Dann gilt

$$\begin{aligned} \|\bar{P}_t f\|^2_{\mu,2} &= E\Big(\big|E^{X_0}\big(\,f(W_t)\,\big)\big|^2\Big) \\ &\overset{\text{Jensen}}{\leq} E\Big(E^{X_0}\big(\,|f(W_t)|^2\,\big)\Big) \\ &= E\big(\,|f(X_t)|^2\,\big) \quad = \quad \|f\|^2_{\mu,2}. \end{aligned}$$

Also ist $\bar{P}_t f \in L^2(\mu)$ und, da sich die Linearität des Erwartungswertes auf $\bar{P}_t$ überträgt, ist $\bar{P}_t$ ein linearer Operator auf $L^2(\mu)$ mit $\|\bar{P}_t\| \leq 1$.

Für $s, t \geq 0$ gilt weiter

$$\begin{aligned} E^{X_0}\Big(\big(\bar{P}_t f\big)(W_s)\Big) &= E\Big(\big(\bar{P}_t f\big)(X_s)\;\Big|\;\mathcal{F}_0^+\Big) \\ &= E\Big(E^{X_s}\big(f(W_t)\big)\;\Big|\;\mathcal{F}_0^+\Big) \\ &\overset{\text{Markov}}{=} E\Big(E\big(f(X_{s+t})\;\big|\;\mathcal{F}_t^+\big)\;\Big|\;\mathcal{F}_0^+\Big) \\ &= E^{X_0}\big(f(W_{s+t})\big), \end{aligned}$$

also

$$\begin{aligned} \big(\bar{P}_s(\bar{P}_t f)\big)(x) &= E^x\Big(\big(\bar{P}_t f\big)(W_s)\Big) \\ &= E^x\big(f(W_{s+t})\big) \quad = \quad \big(\bar{P}_{s+t} f\big)(x) \end{aligned}$$

für μ-fast alle $x \in \mathbb{R}^d$, also $\bar{P}_s\bar{P}_t = \bar{P}_{s+t}$ in $L^2(\mu)$.

Jetzt ist nur noch die Stetigkeit zu zeigen. Sei dafür zuerst $f \in C_b(\mathbb{R}^d, \mathbb{R})$. Da X f.s. stetig ist, gilt

$$f(X_h) \xrightarrow{h\downarrow 0} f(X_0) \quad \text{f.s.},$$

und mit dem Satz von der majorisierten Konvergenz folgt

$$f(X_h) \xrightarrow{h\downarrow 0} f(X_0) \quad \text{in } L^2(\Omega).$$

Lemma 1.2 liefert nun $P_h f(X_0) = E\big(f(X_h)\;\big|\;\mathcal{F}_0\big) \to f(X_0)$ in $L^2(\Omega)$, d.h. $P_h f \to f$ in $L^2(\mu)$. Da $C_b(\mathbb{R}^d, \mathbb{R})$ in $L^2(\mu)$ dicht liegt, folgt mit dem Satz von Banach-Steinhaus

$$P_h f \xrightarrow{h\downarrow 0} f \qquad \text{für alle } f \in L^2(\mu).$$

(q.e.d.)

Offensichtlich ist $\bar{P}_t\big|_{C_b(\mathbb{R}^d)} = P_t$ für alle $t \geq 0$. Mittels Satz 1.3 kann man den Generator dieser Halbgruppe bestimmen.

Satz 4.2 *Sei $(\bar{P}_t)_{t\geq 0}$ die Operatorhalbgruppe aus Satz* 4.1 *und $\bar{L}$ der infinitesimale Generator dieser Halbgruppe. Weiter sei $(L, \mathcal{D}(L))$ der Generator der Halbgruppe aus Satz* 1.6*. Dann gilt $\mathcal{D}(L) \subseteq \mathcal{D}(\bar{L})$ und auf diesem Raum ist $\bar{L} = L$.*

Beweis. Es gilt

$$\|f\|_{\mu,2} \;=\; \big(E_\mu(|f|^2)\big)^{1/2} \;\le\; \big(\|f\|_\infty^2\big)^{1/2} \;=\; \|f\|_\infty$$

für alle $f \in C_b(\mathbb{R}^d)$. Sei nun $\varphi \in \mathcal{D}(L) \subseteq C_b(\mathbb{R}^d)$. Dann gilt

$$\Big\|\frac{\bar P_t\varphi - \varphi}{t} - L\varphi\Big\|_{\mu,2} \;\le\; \Big\|\frac{P_t\varphi - \varphi}{t} - L\varphi\Big\|_\infty \longrightarrow 0 \qquad \text{für } t \downarrow 0,$$

d.h. es ist auch $\varphi \in \mathcal{D}(\bar L)$ und es gilt $\bar L\varphi = L\varphi$. (q.e.d.)

Nach Satz 1.6 stimmt der Generator L und damit auch $\bar L$ auf $C_c^2(\mathbb{R}^d)$ mit dem Diffusionsoperator aus Formel (1.2) überein. Davon, daß $\bar L$ auf $C_c^2(\mathbb{R}^d)$ diese Gestalt besitzt, kann man sich auch direkt durch folgende Rechnung überzeugen: Sei X eine Lösung von (1.1) mit $\mathcal{L}(X_0) = \mu$ und $f \in C_c^2(\mathbb{R}^d)$. Dann hat $\partial_i f$ kompakten Träger und da σ_{ij} lipschitzstetig ist, gilt $\|\partial_i f \cdot \sigma_{ij}\|_\infty < \infty$. Damit folgt

$$\int_0^t \big(\partial_i f(X_s)\sigma_{ij}(X_s)\big)^2\,ds \;\le\; t \cdot \|\partial_i f \cdot \sigma_{ij}\|_\infty^2 \;<\; \infty \qquad \text{für alle } t \ge 0,$$

und nach Satz 1.3 ist

$$E\big(f(X_h) - f(X_0) \;\big|\; \mathcal{F}_0\big) \;=\; E\Big(\int_0^h Lf(X_s)\,ds \;\Big|\; \mathcal{F}_0\Big).$$

Da $s \mapsto b_i(X_s)\cdot\partial_i f(X_s) + \frac{1}{2}a_{ij}(X_s)\partial_i\partial_j f(X_s)$, und damit auch $s \mapsto Lf(X_s)$, stetig ist, gilt

$$\lim_{h\downarrow 0}\frac{1}{h}\int_0^h Lf(X_s)\,ds \;=\; Lf(X_0) \qquad \text{f.s.} \tag{$*$}$$

Außerdem ist

$$\Big|\frac{1}{h}\int_0^h Lf(X_s)\,ds\Big| \;\le\; \|b\cdot\nabla f\|_\infty + \frac{1}{2}\max_{ij}\|a_{ij}\partial_i\partial_j f\|_\infty \;<\; \infty,$$

und mit dem Satz von der majorisierten Konvergenz erhält man den Grenzwert aus $(*)$ auch in $L^2(\Omega)$. Mittels Lemma 1.2 folgt

$$\begin{aligned}\frac{1}{h}\big(\bar P_h f(X_0) - f(X_0)\big) \;&=\; \frac{1}{h}E\big(f(X_h) - f(X_0) \;\big|\; \mathcal{F}_0^+\big)\\ &=\; E\Big(\frac{1}{h}\int_0^h Lf(X_s)\,ds \;\Big|\; \mathcal{F}_0^+\Big)\\ &\overset{h\downarrow 0}{\longrightarrow}\; Lf(X_0) \qquad \text{in } L^2(\Omega),\end{aligned}$$

also

$$\frac{\bar P_h f - f}{h} \;\overset{h\downarrow 0}{\longrightarrow}\; Lf \qquad \text{in } L^2(\mu).$$

Für Lösungen der stochastischen Differentialgleichung (1.1) ist die Existenz einer stationären Verteilung äquivalent dazu, daß der Prozeß stationär ist: Daß stationäre Prozesse eine stationäre Verteilung besitzen, ist klar. Die Umkehrung folgt aus der Markoveigenschaft.

Satz 4.3 *Sei X eine Lösung von Gleichung* (1.1) *mit stationärer Verteilung. Die Drift b und die Diffusionsmatrix σ seien lipschitzstetig. Dann ist X stationär.*

Beweis. Seien $0 \leq t_1 < t_2 < \ldots < t_n$ und $A_1, \ldots, A_n \subseteq \mathbb{R}^d$ meßbar. Setze

$$f := 1_{\{\, \omega \in C(\mathbb{R}_+) \mid \omega_{t_1} \in A_1, \, \ldots, \, \omega_{t_n} \in A_n \,\}}$$

Dann gilt

$$\begin{aligned} P\big(X_{t_1+h} \in A_1, \ldots, X_{t_n+h} \in A_n\big) &= E\big(f(X_{\,\cdot\,+h})\big) \\ &= E\Big(E\big(f(X_{\,\cdot\,+h}) \mid \mathcal{F}_h^+\big)\Big) \\ &\overset{\text{Satz 1.5}}{=} E\big(E^{X_t}(f)\big) \\ &= E\big(E^{X_0}(f)\big), \end{aligned}$$

also

$$P\big(X_{t_1+h} \in A_1, \ldots, X_{t_n+h} \in A_n\big) = P\big(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\big)$$

für alle $h \geq 0$. (q.e.d.)

Nun zeige ich noch, daß man einen stationären Prozeß mit Indexmenge $\mathbb{R}_+$ zu einem Prozeß mit Indexmenge $\mathbb{R}$ fortsetzen kann. Dazu verwende ich einen Satz von Kolmogorov, wie er als Theorem 12.1.2 bei Dudley (1989) steht.

Notation. Ist I eine Menge, so bezeichne $\mathcal{E}(I)$ die Menge aller endlichen Teilmengen von I.

Definition. Seien E und I Mengen, sowie $\mathcal{B}$ eine σ-Algebra auf E. Für jedes $J \in \mathcal{E}(I)$ sei P_J ein Maß auf $(E^J, \mathcal{B}^J)$. Dann heißt die Familie $(P_J)_{J \in \mathcal{E}(I)}$ **konsistent** oder **projektiv**, falls für $J, H \in \mathcal{E}(I)$ und $\emptyset \subset J \subseteq H$ die Projektionseigenschaft

$$P_J = P_H \circ (p_J^H)^{-1}$$

mit

$$p_J^H \colon \begin{array}{ccc} E^H & \to & E^J \\ (H \to E) & \mapsto & (H \to E)\big|_J \end{array}$$

erfüllt ist.

Definition. Sei Σ eine σ-Algebra auf der Menge M. Dann heißt der Meßraum (M, Σ) **standard**, falls es eine Topologie $\mathcal{O}$ gibt, so daß $(M, \mathcal{O})$ ein kompakter metrischer Raum ist und daß $\Sigma = \mathcal{B}(\mathcal{O})$ gilt.

Lemma. *Ist $\mathcal{B}$ die Borel-σ-Algebra auf $\mathbb{R}$, so ist $(\mathbb{R}, \mathcal{B})$ standard.*

Satz (Kolmogorov). *Sei I eine Menge, $(S_t, \mathcal{B}_t)_{t \in I}$ eine Familie von standard-Räumen und $(P_J)_{J \in \mathcal{E}(I)}$ eine projektive Familie über $(S^J, \mathcal{B}^J)_{J \in \mathcal{E}(I)}$. Dann hat $(P_J)_{J \in \mathcal{E}(I)}$ einen eindeutig bestimmten projektiven Limes P_I auf $(S^I, \mathcal{B}^I)$.*

Sei nun $(X_t)_{t \geq 0}$ stationär. Für $T \in \mathcal{E}(\mathbb{R})$, also $T = \{t_1, \ldots, t_n\}$ mit $n \in \mathbb{N}$ und $t_1 < \cdots < t_n$, setze

$$P_T := \mathcal{L}\big(\mathrm{proj}_T \circ (X_{t-t_1})_{t \geq t_1}\big) \in \mathcal{M}_1\big((\mathbb{R}^d)^T\big),$$

wobei proj_T die kanonische Projektion des Pfadraums auf $(\mathbb{R}^d)^T$ ist. Da X stationär ist, erhält man dadurch eine konsistente Familie von Wahrscheinlichkeitsmaßen.

Nach obigem Satz hat (P_T) einen eindeutig bestimmten projektiven Limes $\hat{P}$ auf $\{\, \omega \mid \omega \colon \mathbb{R} \to \mathbb{R}^d \,\}$. Bezüglich $\hat{P}$ ist der Koordinatenprozeß Y auf diesem Raum ein stationärer Prozeß, und es gilt

$$\hat{P}(Y_{t_1} \in A_1, \ldots, Y_{t_n} \in A_n) = P(X_{t_1} - t_1 \in A_1, \ldots, X_{t_n} - t_1 \in A_n)$$

für alle meßbaren Mengen $A_1, \ldots, A_n$ und Zeiten $t_1 < \cdots < t_n$.

5. Reversible Prozesse

In diesem Kapitel beschreibe ich Eigenschaften von reversiblen Diffusionsprozessen.

Definition. Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum. Ein stationärer Diffusionsprozeß X auf diesem Raum mit Werten in $\mathbb{R}^d$ heißt **reversibel** oder **symmetrisch** (engl. symmetric), falls

$$P(X_0 \in A, X_t \in B) \;=\; P(X_t \in A, X_0 \in B) \qquad \text{für alle } A, B \in \mathcal{B}(\mathbb{R}^d) \tag{5.1}$$

gilt.

Bemerkung. Ist X eine Lösung von (1.1) mit lipschitzstetiger Drift und Diffusionsmatrix, so folgt die Stationarität von X bereits aus Bedingung (5.1). Es gilt nämlich

$$\begin{aligned} P(X_0 \in A) \;&=\; P(X_0 \in A, X_t \in \mathbb{R}^d) \\ &=\; P(X_t \in A, X_0 \in \mathbb{R}^d) \;=\; P(X_t \in A) \end{aligned}$$

für alle $A \in \mathcal{B}(\mathbb{R}^d)$ und damit $\mathcal{L}(X_t) = \mathcal{L}(X_0)$ für alle $t \geq 0$. Daher besitzt X eine stationäre Verteilung und nach Satz 4.3 ist der Prozeß dann stationär.

Sei nun X reversibel. Dann folgt aus der Stationarität zum Beispiel

$$\begin{aligned} P(X_t \in A, X_{t+h} \in B) \;&=\; P(X_0 \in A, X_h \in B) \\ &=\; P(X_h \in A, X_0 \in B) \;=\; P(X_{t+h} \in A, X_t \in B) \end{aligned} \tag{5.2}$$

für $t, h \geq 0$ und meßbare Mengen $A, B \subseteq \mathbb{R}^d$.

Ich zeige, daß die Verteilung von X invariant unter Zeitumkehr ist. Die Überlegung am Ende des vorangehenden Kapitels zeigt, daß man aus $(X_t)_{t\geq 0}$ einen Prozeß $(Y_t)_{t\in\mathbb{R}}$ auf einem Raum $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ konstruieren kann, der in folgendem Sinne eine Fortsetzung von X ist: Y ist stationär und es gilt

$$\hat{P}(Y_{t_1} \in A_1, \ldots, Y_{t_n} \in A_n) \;=\; P(X_{t_1 - t_0} \in A_1, \ldots, X_{t_n - t_0} \in A_n)$$

für alle meßbaren Mengen $A_1, \ldots, A_n \subseteq \mathbb{R}^d$ und alle Zeiten $t_0 \leq t_1 < \cdots < t_n$. Setzt man nun $\hat{X}_t := Y_{-t}$ für alle $t \geq 0$, so ist $\hat{X}$ ein Prozeß, der dem Prozeß X mit rückwärts laufender Zeit entspricht.

Satz. *Sei X reversibel und $\hat{X}$ der oben konstruierte, rückwärts laufende Prozeß. Weiter seien $A_1, \ldots, A_n \in \mathbb{R}^d$ meßbar und $0 \leq t_1 < \cdots < t_n$. Dann gilt*

$$\hat{P}(\hat{X}_{t_1} \in A_1, \ldots, \hat{X}_{t_n} \in A_n) \;=\; P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n).$$

Beweis. Es ist

$$\begin{aligned} \hat{P}(\hat{X}_{t_1} \in A_1, \ldots, \hat{X}_{t_n} \in A_n) \;&=\; \hat{P}(Y_{-t_1} \in A_1, \ldots, Y_{-t_n} \in A_n) \\ &=\; P(X_{t_n - t_1} \in A_1, \ldots, X_{t_n - t_n} \in A_n). \end{aligned} \tag{5.3}$$

Ist $P(X_t \in A_k) = 0$ für ein k, so folgt

$$\hat{P}(\hat{X}_{t_1} \in A_1, \ldots, \hat{X}_{t_n} \in A_n) \;=\; 0 \;=\; P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n).$$

Also kann man $P(X_t \in A_k) > 0$ für alle k annehmen. X ist reversibel und damit auch stationär. Mittels einer einfachen Rechnung erhält man

$$P(X_{t_n - t_k} \in A_k \mid X_{t_n - t_{k-1}} \in A_{k-1}) \;=\; P(X_{t_k} \in A_k \mid X_{t_{k-1}} \in A_{k-1}).$$

Also gilt

$$
\begin{aligned}
&\hat{P}(\hat{X}_{t_1} \in A_1, \ldots, \hat{X}_{t_n} \in A_n) \\
&\qquad \overset{(5.3)}{=} P(X_{t_n - t_1} \in A_1) \cdot \prod_{k=2}^{n} P(X_{t_n - t_k} \in A_k \mid X_{t_n - t_{k-1}} \in A_{k-1}) \\
&\qquad = P(X_{t_1} \in A_1) \cdot \prod_{k=2}^{n} P(X_{t_k} \in A_k \mid X_{t_{k-1}} \in A_{k-1}) \\
&\qquad = P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n).
\end{aligned}
$$

(q.e.d.)

Der nächste Satz charakterisiert die Reversibilität eines Prozesses durch die zugehörige Halbgruppe.

Satz 5.1 *Sei X eine Lösung der stochastischen Differentialgleichung* (1.1) *mit* $\mathcal{L}(X_0) = \mu$. *Drift und Diffusionsmatrix seien lipschitzstetig. Weiter sei $(P_t)_{t\geq 0}$ die zugehörige Halbgruppe aus Satz* 1.6 *mit Generator L. Dann sind folgende Bedingungen äquivalent:*

(i) *X ist reversibel.*

(ii) *Für jedes $t \geq 0$ gilt*

$$\int f\, P_t g\, d\mu = \int g\, P_t f\, d\mu \qquad \text{für alle } f, g \in C_b(\mathbb{R}^d).$$

(iii) *Es gilt*

$$\int f\, Lg\, d\mu = \int g\, Lf\, d\mu \qquad \text{für alle } f, g \in C_c^2(\mathbb{R}^d).$$

(iv) *Es gilt*

$$\int f\, Lg\, d\mu = \int g\, Lf\, d\mu \qquad \text{für alle } f, g \in C_c^\infty(\mathbb{R}^d).$$

Beweis. Es gelte zuerst (ii). Dann folgt mit $g \equiv 1$ insbesondere $\int f\, d\mu = \int P_t f\, d\mu$ für alle $t \geq 0$, und nach Lemma 4.1 ist μ stationäre Verteilung von X. Nach Satz 4.1 kann man die Halbgruppe dann zu einer Halbgruppe $(\bar{P}_t)_{t\geq 0}$ auf $L^2(\mu)$ fortsetzen. Auf $C_b(\mathbb{R}^d)$ stimmen P_t und $\bar{P}_t$ überein.

Sei nun $t \geq 0$ und $A, B \subseteq \mathbb{R}^d$ meßbar. Dann gilt

$$(1_A, \bar{P}_t 1_B)_\mu = E\Big(1_A(X_0) E^{X_0}\big(1_B(W_t)\big)\Big) = P(X_0 \in A, X_t \in B).$$

Sind (a_n) und (b_n) Folgen in $C_b(\mathbb{R}^d)$ mit $a_n \to 1_A$, $b_n \to 1_B$ in $L^2(\mu)$, so ist

$$
\begin{aligned}
P(X_0 \in A, X_t \in B) &= \lim_{n\to\infty} (a_n, \bar{P}_t b_n)_\mu \\
&= \lim_{n\to\infty} (\bar{P}_t a_n, b_n)_\mu = P(X_0 \in B, X_t \in A).
\end{aligned}
$$

Also folgt, daß X reversibel ist.

Ist andererseits X reversibel, so gilt

$$
\begin{aligned}
(1_A, \bar{P}_t 1_B)_\mu &= P(X_0 \in A, X_t \in B) \\
&= P(X_t \in A, X_0 \in B) = (\bar{P}_t 1_A, 1_B)_\mu.
\end{aligned}
$$

Da $\bar{P}_t$ und das Skalarprodukt linear sind, erhält man diese Beziehung auch für Treppenfunktionen. Sind dann $f, g \in C_b(\mathbb{R}^d)$ und (f_n) sowie (g_n) Folgen von Treppenfunktionen mit $f_n \to f$, $g_n \to g$ in $L^2(\mu)$, so folgt

$$
\begin{aligned}
(f, \bar{P}_t g)_\mu &= \lim_{n\to\infty} (f_n, \bar{P}_t g_n)_\mu \\
&= \lim_{n\to\infty} (\bar{P}_t f_n, g_n)_\mu = (\bar{P}_t f, g)_\mu.
\end{aligned}
$$

Also gilt die Bedingung (ii).

Nun zeige ich noch die Äquivalenz von (ij), (iij) und (iv). Seien $f, g \in C_c^2(\mathbb{R}^d)$. Da L der Generator von $(P_t)_{t\geq 0}$ ist, gilt

$$\frac{1}{h}(P_h f - f) \to Lf$$

bezüglich $\|\cdot\|_\infty$ und damit auch in $L^2(\mu)$. Es folgt

$$\frac{1}{h}(P_h f - f, g)_\mu \to (Lf, g)_\mu.$$

Es gelte zuerst (ij). Dann folgt

$$\begin{aligned}(Lf, g)_\mu &= \lim_{h\downarrow 0} \frac{1}{h}(P_h f - f, g)_\mu \\ &= \lim_{h\downarrow 0} \frac{1}{h}(f, P_h g - g)_\mu \;=\; (f, Lg)_\mu,\end{aligned}$$

d.h. es gilt (iij).

Bedingung (iij) impliziert banalerweise (iv).

Nimmt man nun (iv) an, so erhält man

$$\begin{aligned}\frac{d}{dt}\big((P_t f, g) - (f, P_t g)\big) &= \lim_{h\downarrow 0} \frac{1}{h}(P_h P_t f - P_t f, g)_\mu - \frac{1}{h}(f, P_h P_t g - P_t g)_\mu \\ &= (LP_t f, g)_\mu - (f, LP_t g)_\mu \;=\; 0 \qquad \text{für } f, g \in C_c^\infty(\mathbb{R}^d).\end{aligned}$$

Daher ist $(P_t f, g) - (f, P_t g) = (P_0 f, g) - (f, P_0 g) = 0$ für alle $t \geq 0$ und da $C_c^\infty(\mathbb{R}^d) \subseteq C_b(\mathbb{R}^d)$ bezüglich $\|\cdot\|_{\mu,2}$ dicht liegt, ist (ij) gezeigt. (q.e.d.)

Für Prozesse, die von vornherein eine stationäre Verteilung besitzen, erhält man aus dem voranstehenden Satz das

Korollar 5.1 *Sei X eine Lösung von* (1.1) *mit stationärer Verteilung μ. Drift und Diffusionsmatrix seien lipschitzstetig. Weiter sei $\bar{L}$ der Generator der zugehörigen Operatorhalbgruppe auf $L^2(\mu)$. Dann sind folgende Bedingungen äquivalent:*

(j) *X ist reversibel.*

(ij) *$\bar{L}$ ist selbstadjungiert.*

Beweis. Wie oben gezeigt ist $C_c^2(\mathbb{R}^d) \subseteq \mathcal{D}(\bar{L})$. Ist $\bar{L}$ selbstadjungiert, so ist insbesondere die dritte Bedingung aus dem Satz erfüllt und es folgt, daß X reversibel ist.

Ist andererseits X reversibel, so folgt mit dem Satz, daß $\bar{P}$ auf $C_b(\mathbb{R}^d)$ symmetrisch ist. Da $C_b(\mathbb{R}^d) \subseteq L^2(\mu)$ dicht ist, ist $\bar{P}_t$ auch auf dem ganzen Raum symmetrisch und damit selbstadjungiert. Der verbleibende Teil des Beweises stammt aus dem Buch von Goldstein (1985), Satz 4.3, wo gezeigt wird, daß eine Operatorhalbgruppe genau dann selbstadjungiert ist, wenn ihr Generator diese Eigenschaft besitzt.

Seien $f, g \in \mathcal{D}(\bar{L})$. Dann ist

$$\begin{aligned}(Lf, g)_\mu &= \lim_{h\downarrow 0} \frac{1}{h}(P_h f - f, g)_\mu \\ &= \lim_{h\downarrow 0} \frac{1}{h}(f, P_h g - g)_\mu \;=\; (f, Lg)_\mu,\end{aligned}$$

d.h. $\bar{L}$ ist symmetrisch und damit $\mathcal{D}(\bar{L}) \subseteq \mathcal{D}(\bar{L}^*)$.

Ist $g \in \mathcal{D}(\bar{L}^*)$, so folgt

$$\begin{aligned}(\bar{P}_t g - g, f)_\mu &= (g, \bar{P}_t f - f)_\mu \\ &= (g, \textstyle\int_0^t \bar{L}\bar{P}_s f\, ds)_\mu \\ &= \int_0^t (g, \bar{L}\bar{P}_s f)_\mu\, ds \\ &= \int_0^t (\bar{P}_s \bar{L}^* g, f)_\mu\, ds \\ &= (\textstyle\int_0^t \bar{P}_s \bar{L}^* g\, ds, f)_\mu \qquad \text{für alle } f \in \mathcal{D}(\bar{L}),\end{aligned}$$

also $\bar{P}_t g - g = \int_0^t \bar{P}_s \bar{L}^* g \, ds$ und damit

$$\lim_{t \downarrow 0} \frac{\bar{P}_t g - g}{t} \;=\; \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \bar{P}_s \bar{L}^* g \, ds \;=\; \bar{L}^* g \qquad \text{in } L^2(\mu).$$

Dies zeigt, daß auch $g \in \mathcal{D}(\bar{L})$ und $\bar{L} g = \bar{L}^* g$ gilt. (q.e.d.)

Ein Prozeß, der durch die Drift gezwungen wird „im Kreis zu laufen“ aber auf dem ganzen „Kreis“ dieselbe Dichte besitzt, könnte eine stationäre Verteilung besitzen. Da sein Verhalten sich aber unter Zeitumkehr ändert, ist er nicht reversibel. So kann man vermuten, daß reversible Prozesse eine rotationsfreie Drift besitzen müssen. Wie der nächste Satz zeigt, ist diese Vermutung zumindest für Lösungen der stochastischen Differentialgleichung

$$dX = b(X) \bullet dt + dB \tag{5.4}$$

korrekt.

Satz 5.2 *Sei X eine Lösung der stochastischen Differentialgleichung* (5.4) *mit lipschitzstetiger Drift $b: \mathbb{R}^d \to \mathbb{R}^d$ und $\mathcal{L}(X_0) = \mu$. Es sei $X_0 \in L^2(\Omega)$. Dann sind folgende Bedingungen äquivalent.*

(j) *X ist reversibel.*

(ij) *Es gibt ein Potential $\Phi: \mathbb{R}^d \to \mathbb{R}$ mit $b = -\operatorname{grad} \Phi$ (im verallgemeinerten Sinne) und $d\mu = \exp(-2\Phi) dx$.*

Bemerkung. Der Gradient in Bedingung (ij) des Satzes ist in folgendem Sinne zu verstehen: es gilt

$$\int_{\mathbb{R}^d} e^{-2\Phi} \partial_j f \, dx \;=\; -\int_{\mathbb{R}^d} 2 b_j e^{-2\Phi} f \, dx \qquad \text{für alle } j \in \{1, \ldots, d\}, f \in C_c^\infty(\mathbb{R}^d). \tag{5.5}$$

Ist Φ im gewöhnlichen Sinne differenzierbar, so kann man die linke Seite in (5.5) umformen:

$$\int_{\mathbb{R}^d} e^{-2\Phi(x)} \partial_j f(x) \, dx \overset{\text{part. Int.}}{=} -\int_{\mathbb{R}^d} (-2) \partial_j \Phi(x) e^{-2\Phi(x)} f(x) \, dx.$$

Daher ist (5.5) für differenzierbare Φ äquivalent zu der Bedingung $b = -\operatorname{grad} \Phi$.

Ich zeige zunächst ein Lemma.

Lemma 5.1 *Sei X eine Lösung der stochastischen Differentialgleichung* (5.4) *mit lipschitzstetiger Drift $b: \mathbb{R}^d \to \mathbb{R}^d$ und $\mathcal{L}(X_0) = \mu$. Es gelte Bedingung* (ij) *aus dem Satz. Zuletzt sei L der Generator der Halbgruppe $(P_t)_{t \geq 0}$. Dann ist*

$$\int_{\mathbb{R}^d} L f \, g \, d\mu \;=\; -\frac{1}{2} \int_{\mathbb{R}^d} \nabla f \cdot \nabla g \, d\mu \qquad \textit{für alle } f, g \in C_c^\infty(\mathbb{R}^d).$$

Beweis des Lemmas. Seien $f, g \in C_c^\infty(\mathbb{R}^d)$. Dann ist

$$L f \;=\; b \cdot \nabla f + \frac{1}{2} \Delta f.$$

Nach Voraussetzung existiert ein $\Phi: \mathbb{R}^d \to \mathbb{R}$ mit $d\mu = \exp(-2\Phi) dx$ und $b = -\operatorname{grad} \Phi$ im verallgemeinerten Sinne von Gleichung (5.5). Damit erhält man

$$\begin{aligned}
\int_{\mathbb{R}^d} Lf\, g\, d\mu \;&=\; \frac{1}{2}\sum_{j=1}^{d}\int_{\mathbb{R}^d} 2b_j \cdot \partial_j f\, g\, e^{-2\Phi(x)} dx \;+\; \frac{1}{2}\int_{\mathbb{R}^d} \Delta f\, g\, e^{-2\Phi(x)} dx \\
&\overset{(5.5)}{=}\; -\frac{1}{2}\sum_{j=1}^{d}\int_{\mathbb{R}^d} e^{-2\Phi(x)}\, \partial_j(g\partial_j f)\, dx \;+\; \frac{1}{2}\int_{\mathbb{R}^d} \Delta f\, g\, e^{-2\Phi(x)} dx \\
&=\; -\frac{1}{2}\sum_{j=1}^{d}\int_{\mathbb{R}^d} e^{-2\Phi(x)}\, (\partial_j g \cdot \partial_j f + g \cdot \partial_j^2 f)\, dx \;+\; \frac{1}{2}\int_{\mathbb{R}^d} \Delta f\, g\, e^{-2\Phi(x)} dx \\
&=\; -\frac{1}{2}\sum_{j=1}^{d}\int_{\mathbb{R}^d} e^{-2\Phi(x)}\, \partial_j g \cdot \partial_j f\, dx \\
&=\; -\frac{1}{2}\int_{\mathbb{R}^d} \nabla g \cdot \nabla f\, d\mu.
\end{aligned}$$

(q.e.d.)

Den nun folgenden Beweis habe ich einer Vorlesung über Diffusionen entnommen, die Prof. Dr. H. v. Weizsäcker im Wintersemester 1995/96 in Kaiserslautern hielt.

Beweis des Satzes. Sei zuerst X reversibel. Weiter sei W der Wienerprozeß auf $\mathbb{R}^d$ mit Startverteilung μ, $(\mathcal{F}_t)_{t\geq 0}$ die natürliche Filtration von W und

$$Z \;:=\; \exp\Big(\int_0^{\cdot} b(W_s)\, dW_s \;-\; \frac{1}{2}\int_0^{\cdot} b^2(W_s)\, ds\Big).$$

Korollar 1.3 besagt, daß dieses Z ein Martingal ist. Nach Satz 1.8 gibt ein eindeutig bestimmtes Wahrscheinlichkeitsmaß Q, so daß $dQ \;=\; Z_t\, d\mathbb{W}^\mu$ auf $\mathcal{F}_t$ für alle $t \;\geq\; 0$ gilt, und bezüglich Q ist

$$B^* \;=\; W - \int_0^{\cdot} b(W_s)\, ds$$

eine Brownsche Bewegung. Nach Definition dieser Brownschen Bewegung löst W die stochastische Differentialgleichung

$$dW \;=\; b(W) \bullet dt + dB^*,$$

und nach Satz 1.1 und dem Satz von Yamada-Watanabe ist $Q \circ W^{-1} = P \circ X^{-1}$, also

$$d\mathcal{L}(X) \;=\; dQ \;=\; Z\, d\mathbb{W}^\mu.$$

Es bezeichne nun $\sim$ die Äquivalenz von Maßen, also $\mu \;\sim\; \nu$ genau dann, wenn jede μ-Nullmenge auch eine ν-Nullmenge ist, und umgekehrt. Da Z echt positiv ist, ist dann $\mathcal{L}(X) \sim \mathbb{W}^\mu$ und es folgt

$$\mu \;=\; \mathcal{L}(X_t) \;=\; Q \circ W_t^{-1} \;\sim\; \mathbb{W}^\mu \circ W_t^{-1} \;=\; \mu * \mathcal{N}(0,t) \;\sim\; \lambda^d,$$

wobei $*$ die Faltung von Maßen, $\mathcal{N}(0,t)$ die Normalverteilung auf $\mathbb{R}^d$ mit Kovarianzmatrix $t \cdot I$ und λ^d das Lebesguemaß auf $\mathbb{R}^d$ bezeichnet. Also existiert die Radon-Nikodym-Dichte $q := d\mu/d\lambda^d$ und ist fast sicher positiv. Setze

$$\Phi \;:=\; -\frac{1}{2}\log q.$$

Sei nun $f \in C_c^\infty(\mathbb{R}^d)$. Nach Satz 1.4 gilt dann

$$\lim_{h\downarrow 0} \frac{1}{h} E\big((X_{t+h} - X_t) f(X_t)\big) \;=\; E\big(b(X_t) f(X_t)\big)$$

und Lemma 1.3 liefert

$$\lim_{h\downarrow 0} E\Big(\frac{1}{h}(X_{t+h}-X_t)\big(f(X_{t+h})-f(X_t)\big)\Big) \;=\; E\big(\operatorname{grad} f(X_t)\big).$$

Zusammengenommen ergibt dies Folgendes.

$$\begin{aligned}
E\big(b(X_t)f(X_t)\big) &= \lim_{h\downarrow 0} E\Big(\frac{1}{h}(X_{t+h}-X_t)f(X_t)\Big)\\
&\overset{(5.2)}{=} \lim_{h\downarrow 0} E\Big(\frac{1}{h}(X_t-X_{t+h})f(X_{t+h})\Big)\\
&= -\lim_{h\downarrow 0} E\Big(\frac{1}{h}(X_{t+h}-X_t)f(X_t)\Big)\\
&\qquad\qquad -\lim_{h\downarrow 0} E\Big(\frac{1}{h}(X_{t+h}-X_t)\big(f(X_{t+h})-f(X_t)\big)\Big)\\
&= -E\big(b(X_t)f(X_t)\big)-E\big(\operatorname{grad} f(X_t)\big),
\end{aligned}$$

also

$$-2\cdot E\big(b(X_t)f(X_t)\big) \;=\; E\big(\operatorname{grad} f(X_t)\big).$$

Damit folgt

$$\begin{aligned}
\int_{\mathbb{R}^d} q(x)\cdot\partial_j f(x)\,dx &= E\big(\partial_j f(X_t)\big)\\
&= E\big(-2b_j(X_t)f(X_t)\big)\\
&= -\int_{\mathbb{R}^d} 2b_j(x)f(x)q(x)\,dx
\end{aligned}$$

für alle $j\in\{1,\dots,d\}$, $f\in C_c^2(\mathbb{R}^d)$. Dies ist aber gerade Formel (5.5), so daß Bedingung (ij) gezeigt ist.

Es gelte nun Bedingung (ij). L sei wieder der Generator der Halbgruppe $(P_t)_{t\geq 0}$. Dann liefert das Lemma

$$\int Lf\,g\,d\mu \;=\; -\frac{1}{2}\int_{\mathbb{R}^d}\nabla f\cdot\nabla g\,d\mu \;=\; \int f\,Lg\,d\mu \qquad \text{für alle } f,g\in C_c^\infty(\mathbb{R}^d).$$

Dies ist aber die Bedingung (*iv*) aus dem letzten Satz, und es folgt, daß X reversibel ist.

(q.e.d.)

Aus dem Satz liest man einige Bedingungen für die Drift b von reversiblen Lösungen der Gleichung (5.4) ab: Zunächst muß b Gradientenfeld, also rotationsfrei, sein. Da μ ein Wahrscheinlichkeitsmaß mit Dichte $\exp(-2\Phi)$ ist, muß Φ außen im Raum groß werden; $b=-\operatorname{grad}\Phi$ ist dort nach innen gerichtet. Auch diese Bedingung ist plausibel, da X eine stationäre Verteilung besitzt und daher nicht immer weiter nach außen laufen kann. Außerdem besitzt die stationäre Verteilung μ eine Dichte bezüglich des Lebesguemaßes, die echt positiv ist.

Für reversible Prozesse hat die Dichte der Verteilung bezüglich des Wienermaßes eine einfache Form: Sei X eine Lösung der stochastischen Differentialgleichung (5.4). Die Drift $b\colon\mathbb{R}^d\to\mathbb{R}^d$ sei lipschitzstetig und so geartet, daß es eine Abbildung $\Phi\in C^2(\mathbb{R}^d,\mathbb{R})$ mit $b=-\operatorname{grad}\Phi$ und

$$\int_{\mathbb{R}^d}\exp\big(-2\Phi(x)\big)\,dx \;=\; 1$$

gibt. Setze $\psi:=\exp(-\Phi)$. Nach Satz 5.2 bedeutet dies, daß X für die Anfangsverteilung $\mathcal{L}(X_0)$ mit $d\mathcal{L}(X_0)=\psi^2\,dx$ reversibel ist.

Weiter setze $v := \Delta\psi/(2\psi)$. Dann ist

$$\Delta\psi \;=\; \operatorname{div}\bigl(-\operatorname{grad}(\Phi)\mathrm{e}^{-\Phi}\bigr) \;=\; \bigl((\operatorname{grad}\Phi)^2 - \Delta\Phi\bigr)\mathrm{e}^{-\Phi},$$

also

$$v \;=\; \frac{1}{2}\bigl((\operatorname{grad}\Phi)^2 - \Delta\Phi\bigr) \;=\; \frac{1}{2}(b^2 + \operatorname{div} b). \tag{5.6}$$

Sei nun $\varphi_t : C\bigl([0;\infty), \mathbb{R}^d\bigr) \to \mathbb{R}$ durch

$$\varphi_t(\omega) \;:=\; \frac{\psi(\omega_t)}{\psi(\omega_0)} \exp\bigl(-\int_0^t v(\omega_s)\,ds\bigr) \qquad \text{für alle } \omega \in C\bigl([0;\infty), \mathbb{R}^d\bigr)$$

definiert. Die Itô-Formel liefert hier

$$\begin{aligned} -d\Phi(X) \;&=\; -\operatorname{grad}\Phi(X) \bullet dX - \frac{1}{2}\Delta\Phi(X) \bullet dt \\ &=\; b(X) \bullet dX - \frac{1}{2}\Delta\Phi(X) \bullet dt, \end{aligned}$$

und wegen

$$\begin{aligned} b(X) \bullet dX - \frac{1}{2}\bigl|b(X)\bigr|^2 \bullet dt \;&=\; -d\Phi(X) - \frac{1}{2}\bigl(|b(X)|^2 - \Delta\Phi(X)\bigr) \bullet dt \\ &=\; -d\Phi(X) - v(X) \bullet dt \end{aligned}$$

gilt dann

$$\begin{aligned} \varphi_t(X) \;&=\; \exp\bigl(-\Phi(X_t)\bigr) \cdot \exp\bigl(\Phi(X_0)\bigr) \cdot \exp\bigl(-\int_0^t v(X_s)\,ds\bigr) \\ &=\; \exp\bigl(\int_0^t b(X_s)\,dX_s - \frac{1}{2}\int_0^t b^2(X_s)\,ds\bigr). \end{aligned}$$

Sei nun $X_0 = 0$ fast sicher. Da b lipschitzstetig ist, kann man Korollar 1.2 anwenden, um zu zeigen, daß das exponentielle lokale Martingal

$$\exp\bigl(-\int_0^\cdot b(X_s)\,dB_s - \frac{1}{2}\int_0^\cdot b^2(X_s)\,ds\bigr)$$

ein Martingal ist. Für $t \geq 0$ liefert Korollar 1.1 dann

$$P(X \in A) \;=\; \int_A \varphi_t\,dW \qquad \text{für alle } A \in \mathcal{F}_t.$$

Ich gebe nun noch einen Satz über große Abweichungen der empirischen Verteilung eines reversiblen Prozesses an. Das Ergebnis stammt aus Kapitel 6.2 des Buches von Deuschel und Stroock (1989) und findet bei mir im siebten Kapitel Verwendung. Sei X wieder eine Lösung der stochastischen Differentialgleichung (5.4). Weiter sei $\Phi \in C^2(\mathbb{R}^d, \mathbb{R})$ eine Abbildung mit $b = -\operatorname{grad}\Phi$ und

$$\int_{\mathbb{R}^d} \exp\bigl(-2\Phi(x)\bigr)\,dx \;=\; 1.$$

Zuletzt sei wieder μ das Wahrscheinlichkeitsmaß auf $\mathbb{R}^d$ mit $d\mu = \exp(-2\Phi)\,dx$. Dann definiere die quadratische Form $\mathcal{E}_0$ durch

$$\mathcal{E}_0(f,g) \;:=\; \frac{1}{2}\int_{\mathbb{R}^d} \nabla f \cdot \nabla g\,d\mu \qquad \text{für alle } f, g \in C_c^\infty(\mathbb{R}^d, \mathbb{R}).$$

Lemma 5.2 *Sei X eine reversible Lösung der stochastischen Differentialgleichung* (5.4) *mit lipschitzstetiger Drift $b: \mathbb{R}^d \to \mathbb{R}^d$ und $\mathcal{L}(X_0) = \mu$. Es sei $X_0 \in L^2(\Omega)$. Weiter sei $\mathcal{E}_0$ die oben definierte, quadratische Form und $\bar{L}$ der Generator der Halbgruppe aus Satz* 4.2. *Dann gilt*

$$\mathcal{E}_0(f,g) \;=\; (-\bar{L}f, g)_\mu \qquad \textit{für alle } f, g \in C_c^\infty(\mathbb{R}^d).$$

Beweis. Dies ist eine direkte Folgerung aus Lemma 5.1. (q.e.d.)

Nach Theorem X.23 bei Reed und Simon bzw. Satz 6.2.9 bei Deuschel und Stroock ist $\mathcal{E}_0$ zu einer quadratischen Form $\big(\mathcal{E},\mathcal{D}(\mathcal{E})\big)$ abschließbar; für $f \in L^2(\mu)$ gilt also

$$f \in \mathcal{D}(\mathcal{E}) \quad \Longleftrightarrow \quad \begin{array}{l} \text{es existieren } f_n \in C_c^\infty(\mathbb{R}^d) \text{ und } g_1,\dots,g_d \in L^2(\mu) \\ \text{mit } f_n \to f \text{ und } \partial_j f_n \to g_j \text{ in } L^2(\mu) \text{ für } j := 1,\dots,d \end{array}$$

und in diesem Falle ist

$$\mathcal{E}(f,f) \;=\; \lim_{n\to\infty} \mathcal{E}_0(f_n,f_n) \;=\; \frac{1}{2}\int_{\mathbb{R}^d} \sum_{j=1}^{d} g_j^2 \, d\mu. \tag{5.7}$$

Diese Funktionen g_j sind die verallgemeinerten partiellen Ableitungen von f; im Folgenden schreibe ich gelegentlich $\partial_j f$ statt g_j. Der Ableitungsoperator $\big(\partial_j,\mathcal{D}(\mathcal{E})\big)$ ist dann ein abgeschlossener Operator auf $L^2(\mu)$.

Lemma. *Sei $f \in \mathcal{D}(\mathcal{E})$ und $\psi \in C^\infty(\mathbb{R};\mathbb{R})$ lipschitzstetig. Dann ist auch $\psi \circ f \in \mathcal{D}(\mathcal{E})$ und es gilt die Kettenregel $\partial_j(\psi\circ f) = \psi'(f)\cdot\partial_j f$.*

Beweis. Nach Definition von $\mathcal{D}(\mathcal{E})$ gibt es eine Folge $(f_n)_{n\in\mathbb{N}}$ in $C_c^\infty(\mathbb{R}^d)$ und Funktionen $g_1,\dots,g_d \in L^2(\mu)$ mit $f_n \to f$ und $\partial_j f_n \to g_j$ in $L^2(\mu)$. Da ψ lipschitzstetig ist, existiert ein $k \geq 0$ mit $\big|\psi(x)-\psi(y)\big| \leq k\cdot|x-y|$. Insbesondere ist $\big|\psi(f_n)-\psi(f)\big| \leq k\cdot|f_n - f|$ und es folgt

$$\psi(f_n) \;\longrightarrow\; \psi(f) \qquad \text{in } L^2(\mu).$$

Weiter gilt

$$\partial_j\psi(f_n) \;=\; \psi'(f_n)\cdot\partial_j f_n \;\longrightarrow\; \psi'(f)\cdot g_j \qquad \mu\text{-fast sicher.}$$

Ich zeige nun, daß diese Konvergenz auch in $L^2(\mu)$ erfolgt: Sei dazu $\varepsilon > 0$. Da $\partial_j f_n \to g_j$ in $L^2(\mu)$ gilt, ist $|\partial_j f_n|^2$ gleichmäßig integrierbar. Also existiert ein $L > 0$, so daß

$$\int_{\{|\partial_j f_n|^2 \geq L/k^2\}} |\partial_j f_n|^2 \, d\mu \;\leq\; \varepsilon/k^2$$

ist. Es folgt

$$\begin{aligned} \int_{\{|\psi'(f_n)\partial_j f_n|^2 \geq L\}} |\psi'(f_n)\partial_j f_n|^2\, d\mu \quad &\leq \quad k^2 \int_{\{k^2|\partial_j f_n|^2\geq L\}} |\partial_j f_n|^2\, d\mu \\ &\leq \quad k^2\cdot\varepsilon/k^2 \;=\; \varepsilon, \end{aligned}$$

d.h. auch $\big|\psi'(f_n)\partial_j f_n\big|^2$ ist gleichmäßig integrierbar, und man erhält

$$\partial_j\psi(f_n) \;\longrightarrow\; \psi'(f)\cdot g_j \qquad \text{in } L^2(\mu)$$

für alle $j \in \{1,\dots,d\}$. Damit ist $\psi(f) \in \mathcal{D}(\mathcal{E})$ und es gilt die behauptete Kettenregel.

(q.e.d.)

Lemma 5.3 *Sei $f \in \mathcal{D}(\mathcal{E})$. Dann ist auch $|f| \in \mathcal{D}(\mathcal{E})$ und es existiert eine monoton fallende Folge $(f_n)_{n\in\mathbb{N}}$ von positiven Funktionen aus $C_c^\infty(\mathbb{R}^d)$ mit $f_n \to |f|$ in $L^2(\mu)$ und $\mathcal{E}_0(f_n,f_n) \to \mathcal{E}\big(|f|,|f|\big)$ für $n\to\infty$.*

Beweis. Zum Beweis approximiert man die Betragsfunktion auf $\mathbb{R}$ durch geeignete glatte Funktionen. Sei also $(\psi_\varepsilon)_{\varepsilon>0}$ eine Familie von Funktionen, die folgende Eigenschaften aufweisen:

(j) Für alle $\varepsilon > 0$ ist $\psi_\varepsilon \in C_c^\infty(\mathbb{R};\mathbb{R})$.
(ij) Für alle $x \in \mathbb{R}$ mit $|x| \geq \varepsilon$ gilt $\psi_\varepsilon(x) = |x|$.
(iij) Alle ψ_ε sind lipschitzstetig mit Lipschitzkonstante 1. Insbesondere gilt also $\|\psi_\varepsilon'\|_\infty = 1$.
(iv) Es gilt $\big\|\psi_\varepsilon - |\cdot|\big\|_\infty \to 0$ für $\varepsilon \to 0$.

Nach dem letzten Lemma ist $\psi_\varepsilon(f) \in \mathcal{D}(\mathcal{E})$ für alle $\varepsilon > 0$ und mit dem Satz von der dominierten Konvergenz erhält man

$$\psi_\varepsilon(f) \longrightarrow |f| \qquad \text{in } L^2(\mu)$$

für $\varepsilon \to 0$. Weiter gilt punktweise

$$\big|\partial_j(\psi_\varepsilon(f))\big| \;=\; \big|\psi_\varepsilon'(f))\big| \cdot \big|\partial_j f\big| \;\le\; \big|\partial_j f\big|,$$

also

$$\big\|\partial_j(\psi_\varepsilon(f))\big\|_{\mu,2} \;\le\; \big\|\partial_j f\big\|_{\mu,2} \qquad \text{für alle } j \in \{1,\dots,d\}.$$

Da die Operatoren $\big(\partial_j, \mathcal{D}(\mathcal{E})\big)$ abgeschlossen sind, folgt auch $|f| \in \mathcal{D}(\mathcal{E})$. (q.e.d.)

Lemma 5.4 *Sei $f \in C^\infty(\mathbb{R}^d) \cap L^2(\mu)$ mit $\partial_j f \in L^2(\mu)$ für alle $j \in \{1,\dots,d\}$. Dann ist $f \in \mathcal{D}(\mathcal{E})$, und es gilt*

$$\mathcal{E}(f,f) \;=\; \frac{1}{2}\int_{\mathbb{R}^d} (\operatorname{grad} f)^2 \, d\mu.$$

Beweis. Sei $(\chi_n)_{n\in\mathbb{N}}$ eine Folge in $C_c^\infty(\mathbb{R}^d)$, so daß $\chi_n(x) = 1$ für alle $x \in \mathbb{R}^d$ mit $|x| \le n$, sowie $\|\chi_n\|_\infty = 1$ und $\|\partial_j \chi_n\|_\infty \le 1$ für alle $n \in \mathbb{N}$, $j \in \{1,\dots,d\}$ gilt. Dann ist $\chi_n \cdot f \in C_c^\infty(\mathbb{R}^d)$, es gilt $|\chi_n \cdot f| \le |f|$, und mit dem Satz von der dominierten Konvergenz folgt

$$\chi_n \cdot f \longrightarrow f \qquad \text{in } L^2(\mu).$$

Weiter ist

$$\partial_j(\chi_n \cdot f) \;=\; (\partial_j \chi_n) \cdot f + \chi_n \cdot \partial_j f,$$

also

$$|\partial_j(\chi_n \cdot f)| \;=\; |(\partial_j \chi_n) \cdot f + \chi_n \cdot \partial_j f| \;\le\; |f| + |\partial_j f|.$$

Für diese Schranke gilt

$$\big\|\, |f| + |\partial_j f| \,\big\|_{\mu,2}^2 \;\le\; \|f\|_{\mu,2}^2 + 2\int |f \cdot \partial_j f| \, d\mu + \|\partial_j f\|_{\mu,2}^2 \;<\; \infty$$

und wieder mit dem Satz von der dominierten Konvergenz folgt

$$\partial_j(\chi_n \cdot f) \longrightarrow 0 \cdot f + 1 \cdot \partial_j f \;=\; \partial_j f \qquad \text{in } L^2(\mu).$$

Dies zeigt, daß $f \in \mathcal{D}(\mathcal{E})$ ist, und daß

$$\mathcal{E}(f,f) \;=\; \frac{1}{2}\int \sum_{j=1}^d (\partial_j f)^2 \, d\mu \;=\; \frac{1}{2}\int (\operatorname{grad} f)^2 \, d\mu$$

gilt. (q.e.d.)

Definiere jetzt noch $J_\mathcal{E}\colon \mathcal{M}_1(\mathbb{R}^d) \to [0;\infty]$ durch

$$J_\mathcal{E}(\nu) \;:=\; \begin{cases} \mathcal{E}(f,f), & \text{falls } \nu \ll \mu \text{ mit } d\nu = f^2 d\mu \text{ für ein } f \in \mathcal{D}(\mathcal{E}), \text{ und} \\ +\infty & \text{sonst.} \end{cases}$$

Weiter sei $\lambda\big|_{[0;t]}$ das Lebesguemaß auf dem Intervall $[0;t]$ und die empirische Verteilung $L_t^\omega \in \mathcal{M}_1(\mathbb{R}^d)$ für $\omega \in C\big([0;\infty), \mathbb{R}^d\big)$ durch

$$L_t^\omega \;:=\; \frac{1}{t}\,\lambda\big|_{[0;t]} \circ \big(\omega\big|_{[0;t]}\big)^{-1}$$

definiert. Dann ist L_t eine Abbildung von $C\big([0;\infty), \mathbb{R}^d\big)$ nach $\mathcal{M}_1(\mathbb{R}^d)$. Zuletzt sei P_ν die Verteilung von X auf $C\big([0;\infty), \mathbb{R}^d\big)$ mit der Anfangsbedingung $\mathcal{L}(X_0) = \nu \in \mathcal{M}_1(\mathbb{R}^d)$. Dann kann man durch das Wahrscheinlichkeitsmaß P_ν auf $C\big([0;\infty), \mathbb{R}^d\big)$ die empirische Verteilung als ein zufälliges Maß in $\mathcal{M}_1(\mathbb{R}^d)$ auffassen.

$\mathcal{M}_1(\mathbb{R}^d)$ sei mit der schwachen Topologie und der Borel-σ-Algebra versehen.

In dieser Situation kann man Satz 6.2.21 bei Deuschel und Stroock (1989) anwenden. (Die Notation dort weicht etwas von der hier gebrauchten ab: was dort V heißt, ist hier $2v$ und dem dortigen U entspricht 2Φ hier.) Der Satz liefert Folgendes.

Satz 5.3 *Sei v so, daß $\{ x \in \mathbb{R}^d \mid v(x) \leq R \}$ für jedes $R \geq 0$ kompakt ist. Dann ist $J_{\mathcal{E}}$ eine gute Ratenfunktion und es gilt*

$$
\begin{aligned}
-\inf_{\Gamma^\circ} J_{\mathcal{E}} \;&\leq \inf_{\nu \in \mathcal{M}_1(\mathbb{R}^d)} \liminf_{t\to\infty} \frac{1}{t} \log P_\nu\big(\{ \omega \mid L_t^\omega \in \Gamma \}\big) \\
&\leq \sup_{\nu \in \mathcal{M}_1(\mathbb{R}^d)} \limsup_{t\to\infty} \frac{1}{t} \log P_\nu\big(\{ \omega \mid L_t^\omega \in \Gamma \}\big) \leq -\inf_{\overline{\Gamma}} J_{\mathcal{E}}
\end{aligned}
$$

für alle meßbaren $\Gamma \subseteq \mathcal{M}_1(\mathbb{R}^d)$.

6. Die Schrödingergleichung

In diesem Kapitel beschreibe ich einen Zusammenhang zwischen reversiblen Lösungen der stochastischen Differentialgleichung (5.4) und dem Grundzustand des (mathematischen) Hamiltonoperators. Diesem Grundzustand entspricht eine stationäre Lösung der Schrödingergleichung. Die Rechnungen stammen größtenteils aus dem Artikel von Mitter (1986). Ich beginne die Darstellung damit, daß ich den mathematischen Hamiltonoperator definiere und kurz auf den Zusammenhang zum quantenmechanischen Hamiltonoperator eingehe.

Sei $v \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$. Dann definiere den Operator H_0 durch

$$\big(H_0\varphi\big)(x) \;:=\; -\frac{1}{2}\Delta\varphi(x) \,+\, v(x)\cdot\varphi(x) \qquad \text{für alle } x\in\mathbb{R}^d,\ \varphi\in C_c^2(\mathbb{R}^d).$$

Für $\varphi \in C_c^2(\mathbb{R}^d)$ sind $\Delta\varphi$ und $v\cdot\varphi$ in $L^2(\mathbb{R}^d)$. Also gilt dies auch für $H_0\varphi$, und man kann H_0 als linearen Operator auf $L^2(\mathbb{R}^d)$ auffassen.

Falls H_0 eine Fortsetzung zu einem selbstadjungierten Operator auf $L^2(\mathbb{R}^d)$ besitzt, so heißt diese Fortsetzung **Hamiltonoperator** zum Potential v. Hinreichend für die Existenz eines Hamiltonoperators ist zum Beispiel die Bedingung

$$\operatorname*{ess\,inf}_{x\in\mathbb{R}^d} v(x) > -\infty.$$

Dies sieht man folgendermaßen: Für $\varphi,\psi \in C_c^2(\mathbb{R}^d)$ liefert partielle Integration

$$\begin{aligned}
(H_0\varphi,\psi) \;&=\; \int \big(-\frac{1}{2}\Delta\varphi(x) + v(x)\cdot\varphi(x)\big)\cdot\psi(x)\,dx \\
&=\; \int \frac{1}{2}\nabla\varphi(x)\cdot\nabla\psi(x)\,dx \;+\; \int \varphi(x)\cdot\big(v\psi\big)(x)\,dx \\
&=\; \int \varphi(x)\cdot\big(-\frac{1}{2}\Delta\psi(x) + v(x)\cdot\psi(x)\big)\,dx \\
&=\; (\varphi, H_0\psi),
\end{aligned}$$

d.h. H_0 ist symmetrisch. Weiter ist $\underline{v} := \operatorname{ess\,inf} v > -\infty$, und wegen

$$(H_0\psi,\psi) \;=\; \int \frac{1}{2}\big(\nabla\psi(x)\big)^2\,dx + \int v(x)\psi^2(x)\,dx \;\geq\; 0 + \underline{v}\cdot(\psi,\psi)$$

ist H_0 nach unten halbbeschränkt. Nun ist folgender Satz (Riesz und Nagy (1973), Nr. 124) anwendbar.

Satz (Friedrichs-Fortsetzung). *Sei A ein linearer Operator auf einem Hilbertraum. A sei symmetrisch und nach unten halbbeschränkt. Dann besitzt A eine selbstadjungierte Fortsetzung mit derselben unteren Schranke.*

Also kann man H_0 zu einem selbstadjungierten Operator $\big(H,\mathcal{D}(H)\big)$ auf dem Hilbertraum $L^2(\mathbb{R}^d)$ fortsetzen. Dann ist H der Hamiltonoperator zum Potential v. Das folgende Lemma zeigt noch, daß die untere Schranke eines Operators im obigen Sinne auch eine untere Schranke des Spektrums ist.

Lemma. *Sei A ein selbstadjungierter, linearer Operator auf einem Hilbertraum $\mathcal{H}$. Der Operator A sei nach unten halbbeschränkt, d.h. es gebe ein $c\in\mathbb{R}$ mit $(Af,f)\geq c\cdot(f,f)$ für alle $f\in\mathcal{H}$. Dann ist c auch eine untere Schranke für das Spektrum von A.*

Beweis. Sei zuerst λ ein Eigenwert von A. Dann gibt es ein $f \in \mathcal{H}$ mit $f \neq 0$ und $Af = \lambda f$, also

$$\lambda \cdot \|f\|^2 \;=\; (\lambda f, f) \;=\; (Af, f) \;\geq\; c \cdot \|f\|^2$$

und damit $\lambda \geq c$.

Nun sei λ ein Element des kontinuierlichen Spektrums von A. Dann ist $A - \lambda I$ injektiv und $(A - \lambda I)^{-1}$ unstetig. Für alle $n \in \mathbb{N}$ gibt es also ein $f_n \in \mathcal{D}\big((A - \lambda I)^{-1}\big)$ mit

$$\big\|(A - \lambda I)^{-1} f_n\big\| \;\geq\; n \cdot \|f_n\|.$$

Setze

$$g_n \;:=\; \frac{(A - \lambda I)^{-1} f_n}{\|(A - \lambda I)^{-1} f_n\|}\,.$$

Dann ist $\|g_n\| = 1$ für alle $n \in \mathbb{N}$, und es gilt

$$\big\|(A - \lambda I) g_n\big\| \;=\; \frac{\|f_n\|}{\big\|(A - \lambda I)^{-1} f_n\big\|} \;\leq\; \frac{1}{n} \;\longrightarrow\; 0.$$

Damit erhält man auch

$$((A - \lambda I) g_n, g_n) \;\leq\; \big\|(A - \lambda I) g_n\big\| \cdot \|g_n\| \;\longrightarrow\; 0.$$

Angenommen es gelte $g_n \to g \in \mathcal{H}$, also

$$A g_n \;=\; (A - \lambda I) g_n + \lambda g_n \;\longrightarrow\; 0 + \lambda g \qquad \text{für } n \to \infty.$$

Da A abgeschlossen ist, wäre dann $g \in \mathcal{D}(A)$ mit $Ag = \lambda g$ und

$$\|g\| \;=\; \lim_{n \to \infty} \|g_n\| \;=\; 1.$$

Da dies im Widerspruch zur Injektivität von $A - \lambda I$ stände, kann (g_n) nicht konvergent sein.

Also existiert ein $\varepsilon > 0$ und eine Teilfolge $(g_{n_k})_{k \in \mathbb{N}}$ mit $\|g_{n_k}\|^2 \geq \varepsilon$ für alle $k \in \mathbb{N}$. Für diese Teilfolge gilt

$$(c - \lambda)\varepsilon \;\leq\; (c - \lambda)(g_{n_k}, g_{n_k}) \;\leq\; ((A - \lambda I) g_{n_k}, g_{n_k}) \;\longrightarrow\; 0,$$

also auch in diesem Fall $\lambda \geq c$. (q.e.d.)

Bemerkung. Die selbstadjungierte Erweiterung von H ist nicht notwendigerweise eindeutig bestimmt (vgl. das Beispiel nach Theorem VIII.3 bei Reed und Simon. Dort wird ein selbstadjungierter Operator mit überabzählbar vielen unterschiedlichen selbstadjungierten Erweiterungen konstruiert). Außerdem ist folgendes zu beachten: selbstadjungierte Operatoren sind zwar abgeschlossen, und symmetrische Operatoren A sind stets abschließbar (z.B. ist A^{**} abgeschlossen), aber keine der abgeschlossenen Fortsetzungen von A muß dann selbstadjungiert sein (vgl. die Bemerkung nach Theorem VIII.15 bei Reed und Simon).

Sei nun H ein Hamiltonoperator. Eine Familie $(\psi_t)_{t \geq 0}$ von Funktionen aus $\mathcal{D}(H)$ erfüllt die (mathematische) **Schrödingergleichung**, falls

$$H\psi_t = \frac{\partial}{\partial t}\psi_t \qquad \text{für alle } t \geq 0 \tag{6.1}$$

gilt. Die Lösung heißt **stationär**, falls $\psi_t = \psi_0$ für alle $t \geq 0$ gilt. Da H selbstadjungiert ist, kann man mittels des Spektralsatzes den Operator $\exp(t \cdot H)$ definieren. Man erhält zu jeder Anfangsbedingung $\psi \in \mathcal{D}\big(\exp(tH)\big)$ eine Lösung von (6.1): Es ist nämlich

$$\frac{\partial}{\partial t}\exp(tH)\psi = H\exp(tH)\psi,$$

d.h. $\big(\exp(tH)\psi\big)_{t \geq 0}$ löst die Schrödingergleichung.

Ist das Infimum des Spektrums von H ein Eigenwert, so heißen die zugehörigen Eigenfunktionen **Grundzustand** von H. Ist die Dimension des Eigenraums größer als eins, so sagt man, der Grundzustand sei „entartet“. Der folgende Satz ist Theorem XIII.47 aus dem Buch von Reed und Simon. Er gibt ein hinreichendes Kriterium für die Existenz eines Grundzustandes an.

Satz. *Sei $v \in L^2_{\mathrm{loc}}$, $v \geq 0$ mit $v(x) \to \infty$ für $|x| \to \infty$. Dann hat $-\Delta + v$ einen nicht entarteten Grundzustand, der echt positiv ist.*

Mit v erfüllt auch $2v$ die Voraussetzungen des Satzes. Daher gibt es ein ψ mit $-\Delta\psi + 2v\psi = \lambda\psi$, und wegen $-\frac{1}{2}\Delta\psi + v\psi = \frac{\lambda}{2}\psi$ ist ψ Grundzustand des oben definierten Hamiltonoperators. Also kann man im Satz auch $-\frac{1}{2}\Delta$ statt $-\Delta$ schreiben. Ist H der Hamiltonoperator zum Potential $v \geq 0$, und ist $H\psi = \lambda\psi$, so gilt

$$-\frac{1}{2}\Delta\psi + (v-c)\psi \;=\; (-\frac{1}{2}\Delta + v)\psi - c\psi \;=\; (\lambda - c)\psi. \tag{6.2}$$

Daher kann man die Bedingung $v \geq 0$ des Satzes durch $v \geq c$ für $c \in \mathbb{R}$ ersetzen und erhält folgendes

Korollar. *Sei $v \in L^2_{\mathrm{loc}}$ nach unten beschränkt mit $v(x) \to \infty$ für $|x| \to \infty$. Dann existiert der Hamiltonoperator zum Potential v und besitzt einen nicht entarteten Grundzustand, der echt positiv ist.*

Der oben definierte Hamiltonoperator steht in engem Zusammenhang zum Hamiltonoperator aus der Quantenmechanik. Bei der quantenmechanischen Beschreibung eines Partikels im $\mathbb{R}^d$ bestimmt man nicht — wie in der klassischen Mechanik — die Bahnkurve des Teilchens. Sein Verhalten wird vielmehr als zufällig angenommen, und es sind nur noch Aussagen darüber möglich, mit welcher Wahrscheinlichkeit man das Teilchen zu einer bestimmten Zeit in einem bestimmten Gebiet des Raumes finden kann. Als Zustandsraum wählt man den Hilbertraum $L^2_{\mathbb{C}}(\mathbb{R}^d)$. Die physikalisch möglichen Zustände des Systems werden dann durch Vektoren ψ dieses Raums mit $\|\psi\|_2 = 1$ dargestellt. Ist ψ_t der Zustand des Systems zur Zeit t, so ist $\int_A |\psi_t|^2\,dx$ die Wahrscheinlichkeit, daß sich das Teilchen zur Zeit t in der Menge $A \subseteq \mathbb{R}^d$ aufhält. Der Hamiltonoperator H_{QM} wird in der Quantenmechanik durch

$$\big(H_{\mathrm{QM}}\psi\big)(x) \;:=\; -\frac{\hbar^2}{2m}\Delta\psi(x) \;+\; v(x)\cdot\psi(x) \qquad \text{für alle } x \in \mathbb{R}^d$$

definiert, wobei $\hbar \approx 1{,}055 \cdot 10^{-34} Js$ die Plancksche Konstante und m die Masse des Teilchens ist. Dieser Operator spielt unter anderem in folgenden beiden Beziehungen eine Rolle: Erstens wird die Zeitevolution des Systems durch die physikalische Schrödingergleichung

$$i\hbar\frac{\partial}{\partial t}\psi_t \quad = \quad H_{\mathrm{QM}}\psi_t$$

beschrieben. Zweitens hat die Energie eines Teilchens im Zustand ψ den Erwartungswert $(H_{\mathrm{QM}}\psi, \psi)$.

Physikalische und mathematische Schrödingergleichung unterscheiden sich also durch den Faktor $i\hbar$ vor der Zeitableitung. Hier ist dieser Unterschied irrelevant, da ich nur stationäre Prozesse betrachte und somit stets $i\hbar\partial_t\psi_t = 0 = \partial_t\psi_t$ gilt. Im allgemeinen spielt dieser Unterschied jedoch eine große Rolle; dies wird bei Nagasawa (1993) behandelt.

Das physikalische Verhalten des Systems ändert sich nicht, wenn man zu v eine Konstante $v_0 \in \mathbb{R}$ addiert: Ist $(\psi_t)_{t\geq 0}$ eine Lösung der Schrödingergleichung mit Potential v, so ist wegen

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\big(\mathrm{e}^{-iv_0t/\hbar}\psi_t\big) \quad &= \quad i\hbar\big(-\frac{iv_0}{\hbar}\mathrm{e}^{-iv_0t/\hbar}\psi_t \;+\; \mathrm{e}^{-iv_0t/\hbar}\frac{\partial}{\partial t}\psi_t\big) \\
&= \quad v_0\cdot\mathrm{e}^{-iv_0t/\hbar}\psi_t \;+\; \mathrm{e}^{-iv_0t/\hbar}H_{\mathrm{QM}}\psi_t \\
&= \quad \big(H_{\mathrm{QM}} + v_0\big)\big(\mathrm{e}^{-iv_0t/\hbar}\psi_t\big)
\end{aligned}$$

die Familie $(e^{-iv_0 t/\hbar}\psi_t)_{t\geq 0}$ ebenfalls Lösung der Schrödingergleichung mit Potential $v+v_0$; da $|e^{-iv_0 t/\hbar}\psi_t|^2 = |\psi_t|^2$ gilt, bleiben z.B. die Aufenthaltswahrscheinlichkeiten des Teilchens unverändert.

Ich gebe nun eine Zuordnung zwischen reversiblen Lösungen X von Gleichung (5.4) und stationären Lösungen ψ der Schrödingergleichung an. Der Zusammenhang zwischen diesen Objekten besteht in der Beziehung

$$\int_A \psi^2(x)\,dx \;=\; P(X_t \in A)$$

für meßbare Mengen $A \subseteq \mathbb{R}^d$, d.h. die Aufenthaltswahrscheinlichkeiten des Partikels aus dem quantenmechanischen System und des Diffusionsprozesses stimmen überein. ψ muß der Grundzustand des Hamiltonoperators sein. Ich benötige für beide Richtungen der Äquivalenz, daß ψ zweite Ableitungen besitzt. Zuerst konstruiere ich aus dem Grundzustand den zugehörigen Diffusionsprozeß.

Sei also $v \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$ nach unten beschränkt mit $v(x) \to \infty$ für $|x| \to \infty$ und H der Hamiltonoperator zum Potential v. Dann ist nach obigem Korollar das Infimum λ des Spektrums vom H ein Eigenwert, und es gibt einen zugehörigen Eigenvektor $\psi \in L^2(\mathbb{R}^d)$ mit $\|\psi\|_2 = 1$ und $\psi(x) > 0$ für alle $x \in \mathbb{R}^d$. Gleichung (6.2) zeigt, daß man durch Addition einer Konstanten zu v auch $\lambda = 0$ erreichen kann; wie oben gezeigt beschreibt dieser modifizierte Hamiltonoperator dasselbe physikalische System. Dann ist ψ eine stationäre Lösung der Schrödingergleichung (6.1)[1].

Die Abbildung ψ sei zweimal stetig differenzierbar und $\partial_j\psi$ sowie $\partial_j^2\psi$ seien für $j \in \{1,\ldots,d\}$ in $L^2(\mathbb{R}^d)$. Dann gilt $-\frac{1}{2}\Delta\psi + v\cdot\psi = H\psi = 0$, also $v = \Delta\psi/(2\psi)$. Definiere $\Phi := -\log\psi$. Dann ist $\psi = \exp(-\Phi)$ und aus Gleichung (5.6) folgt

$$v \;=\; \frac{1}{2}\big((\operatorname{grad}\Phi)^2 - \Delta\Phi\big).$$

Sei nun W der d-dimensionale Wienerprozeß, $(\mathcal{F}_t)_{t\geq 0}$ die natürliche Filtration von W und Z das exponentielle lokale Martingal zu $-\operatorname{grad}\Phi(W)\bullet dW$, also

$$Z_t \;=\; \exp\Big(-\int_0^t \operatorname{grad}\Phi(W_s)\,dW_s - \frac{1}{2}\int_0^t \big(\operatorname{grad}\Phi\big)^2(W_s)\,ds\Big).$$

Indem man die gerade gewonnene Form für v einsetzt, kann man diesen Ausdruck umformen: Die Itô-Formel liefert

$$d\,\Phi(W) \;=\; \operatorname{grad}\Phi(W)\bullet dW \;+\; \frac{1}{2}\Delta\Phi(W)\bullet dt,$$

folglich

$$\begin{aligned}
-\operatorname{grad}\Phi(W)\bullet dW \;&-\; \frac{1}{2}\big(\operatorname{grad}\Phi\big)^2(W)\bullet dt \\
&= -d\,\Phi(W) \;-\; \frac{1}{2}\Big(\big(\operatorname{grad}\Phi\big)^2(W) - \Delta\Phi(W)\Big)\bullet dt \\
&= -d\,\Phi(W) \;-\; v(W)\bullet dt
\end{aligned}$$

und damit

$$\begin{aligned}
Z_t \;&=\; \exp\Big(-\Phi(W_t) + \Phi(W_0) \;-\; \int_0^t v(W_s)\,ds\Big) \\
&=\; \frac{\psi(W_t)}{\psi(W_0)}\cdot\exp\Big(-\int_0^t v(W_s)\,ds\Big).
\end{aligned}$$

Ich zeige nun, daß Z ein Martingal ist.

[1] Daß ich diese Voraussetzungen an v benötige, stellt eine Lücke dar: Erstens erfüllt diese Richtung der Rechnung damit nicht das oben versprochene. Zweitens wird bei der Rückrichtung nicht gezeigt, daß das dort konstruierte v diese Eigenschaften besitzt.

Lemma. *Sei $g \in L^1(\mathbb{R}^d)$ auf einer Nullumgebung beschränkt. Dann ist*

$$E\Big(\int_0^t |g(W_s)|\,ds\Big) \;<\; \infty.$$

Beweis. Nach Voraussetzung existieren $\varepsilon > 0$ und $K > 0$, so daß $\big|g(x)\big| \leq K$ für alle $x \in \mathbb{R}^d$ mit $|x| < \varepsilon$ gilt. Definiere $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ durch

$$\Phi(x) \;:=\; \int_0^t \frac{1}{(2\pi s)^{d/2}} \exp(-\frac{|x|^2}{2s})\,ds.$$

Indem man Integration in Polarkoordinaten verwendet, findet man dann

$$\begin{aligned}\int_{\{|x|<\varepsilon\}} \Phi(x)\,dx \;&=\; \int_{\{|x|<\varepsilon\}} \int_0^t \frac{1}{(2\pi s)^{d/2}} \exp(-\frac{|x|^2}{2s})\,ds\,dx \\ &=\; C \cdot \int_0^\varepsilon \int_0^t \frac{1}{s^{d/2}} \exp(-\frac{r^2}{2s}) r^{d-1}\,ds\,dr\end{aligned}$$

für eine geeignete Konstante C. Nun läßt sich der Transformationssatz für das Lebesgueintegral anwenden: Definiere $\varphi\colon (0;\varepsilon) \times (0;t) \to \mathbb{R}^2$ durch

$$\varphi(r,s) \;:=\; (r^2, r^2/s) \qquad \text{für alle } (r,s) \in (0;\varepsilon) \times (0;t).$$

Dann ist

$$\det D\varphi(r,s) \;=\; \det \begin{pmatrix} 2r & 0 \\ 2r/s & -r^2/s^2 \end{pmatrix} \;=\; -\frac{2r^3}{s^2}.$$

Definiere nun noch $f(u,v) := \exp(-v/2) v^{d/2-2}/2$. Dann ist

$$\begin{aligned} f \circ \varphi(r,s) \cdot \big|\det D\varphi(r,s)\big| \;&=\; \frac{1}{2} \exp\big(-\frac{r^2}{2s}\big) \frac{r^{d-4}}{s^{d/2-2}} \cdot \frac{2r^3}{s^2} \\ &=\; \exp\big(-\frac{r^2}{2s}\big) \frac{r^{d-1}}{s^{d/2}}.\end{aligned}$$

Der Transformationssatz liefert also

$$\begin{aligned}\int_{\{|x|<\varepsilon\}} \Phi(x)\,dx \;&=\; C \cdot \int_{(0;\varepsilon)\times(0;t)} \exp\big(-\frac{r^2}{2s}\big) \frac{r^{d-1}}{s^{d/2}}\,d\lambda^2(r,s) \\ &=\; C \cdot \int_{(0;\varepsilon)\times(0;t)} f \circ \varphi(r,s) \cdot \big|\det D\varphi(r,s)\big|\,d\lambda^2(r,s) \\ &=\; C \cdot \int_{\{\,(u,v)\,|\,0<u<\varepsilon^2,\; v>u/t\,\}} f(u,v)\,d\lambda^2(u,v) \\ &=\; C \cdot \int_0^{\varepsilon^2} \int_{u/t}^\infty \frac{1}{2} \exp(-v/2) v^{d/2-2}\,dv\,du.\end{aligned}$$

Setze jetzt $a := \varepsilon^2/t$. Dann ist auf dem Integrationsbereich stets $u/t \leq a$ und man kann den Ausdruck auf der rechten Seite folgendermaßen zerlegen:

$$\begin{aligned}\int_{\{|x|<\varepsilon\}} \Phi(x)\,dx \;&=\; \frac{C}{2} \cdot \int_0^{\varepsilon^2} \int_a^\infty \exp(-v/2) v^{d/2-2}\,dv\,du \\ &\qquad + \frac{C}{2} \cdot \int_0^{\varepsilon^2} \int_{u/t}^a \exp(-v/2) v^{d/2-2}\,dv\,du \\ &\leq\; \frac{C\varepsilon^2}{2} \cdot \int_a^\infty \exp(-v/2) v^{d/2-2}\,dv\,du \\ &\qquad + \frac{C}{2} \cdot \int_0^{\varepsilon^2} \int_{u/t}^a v^{d/2-2}\,dv\,du \\ &=\; \frac{C\varepsilon^2}{2} \cdot \int_a^\infty \exp(-v/2) v^{d/2-2}\,dv\,du \\ &\qquad + \begin{cases} \frac{C}{2} \cdot \int_0^{\varepsilon^2} \ln(a) - \ln(u/t)\,du, & \text{falls } d=2, \text{ und} \\ \frac{C}{d-2} \cdot \int_0^{\varepsilon^2} a^{d/2-1} - (u/t)^{d/2-1}\,du & \text{sonst.} \end{cases}\end{aligned}$$

Der erste Summand ist endlich, da für große v der Exponentialterm schnell abfällt. Der zweite Summand ist ebenfalls endlich, da $\log(u)$ für $u \to 0$ genügend langsam gegen $-\infty$ fällt und $u^{d/2-1}$ für $d > 0$ auf $(0;\varepsilon^2)$ integrierbar ist. Also gilt

$$\int_{\{|x|<\varepsilon\}} \Phi(x)\, dx \;<\; \infty.$$

Für $|x| \geq \varepsilon$ gilt andererseits

$$\big|\Phi(x)\big| \;\leq\; \int_0^t \frac{1}{(2\pi s)^{d/2}} \exp\big(-\frac{\varepsilon^2}{2s}\big)\, ds \;\leq\; \underbrace{t \cdot \big\|\frac{1}{(2\pi s)^{d/2}} \exp\big(-\frac{\varepsilon^2}{2s}\big)\big\|_\infty}_{=:\, K'} \;<\; \infty.$$

Damit folgt

$$\begin{aligned} E\Big(\int_0^t |g(W_s)|\, ds\Big) &\overset{\text{Tonelli}}{=} \int_{\mathbb{R}^d} \int_0^t \frac{1}{(2\pi s)^{d/2}} \exp\big(-\frac{x^2}{2s}\big)\big|g(x)\big|\, ds\, dx \\ &= \int_{\{|x|<\varepsilon\}} \Phi(x)\big|g(x)\big|\, dx + \int_{\{|x|\geq\varepsilon\}} \Phi(x)\big|g(x)\big|\, dx \\ &\leq K \cdot \int_{\{|x|<\varepsilon\}} \Phi(x)\, dx + K' \cdot \int_{\{|x|\geq\varepsilon\}} \big|g(x)\big|\, dx \\ &< \infty. \end{aligned}$$

(q.e.d.)

Sei nun $K < 0$ eine untere Schranke für v und $t \geq 0$. Mit Hilfe des Lemmas kann man zeigen, daß die Menge $\mathcal{M} := \{\, Z_T \mid T \leq t \text{ ist eine Stoppzeit} \,\}$ beschränkt in $L^2(\mathbb{R}^d)$ ist: Es gilt

$$\big|Z_T\big|^2 \;=\; \frac{|\psi(W_T)|^2}{|\psi(W_0)|^2} \cdot \exp\big(-2\int_0^T v(W_s)\, ds\big) \;\leq\; \frac{|\psi(W_T)|^2}{|\psi(W_0)|^2} \cdot \mathrm{e}^{-2tK}.$$

Also genügt es, eine Schranke für $E\big(\psi^2(W_T)\big)$ zu finden, die unabhängig von T ist. Nach der Itô-Formel gilt

$$d\psi^2(W) \;=\; 2\psi(W)\psi'(W) \bullet dW + \big(\psi(W)\Delta\psi(W) + (\operatorname{grad}\psi(W))^2\big) \bullet dt.$$

Sei nun $(T_n)_{n\in\mathbb{N}}$ eine lokalisierende Folge für das lokale Martingal

$$2\int_0^{\cdot} \psi(W_s)\psi'(W_s)\, dW_s.$$

Dann ist $(2\int_0^{t\wedge T_n} \psi(W_s)\psi'(W_s)\, dW_s)_{t\geq 0}$ ein Martingal, und nach dem Stoppsatz (Satz 3.2.7 bei v. Weizsäcker und Winkler) gilt

$$E\Big(2\int_0^{T\wedge T_n} \psi(W_s)\psi'(W_s)\, dW_s\Big) \;=\; 0.$$

Damit folgt

$$\begin{aligned} E\Big(\psi^2(W_{T\wedge T_n})\Big) &= E\Big(\psi^2(W_0)\Big) + E\Big(2\int_0^{T\wedge T_n} \psi(W_s)\psi'(W_s)\, dW_s\Big) \\ &\quad + E\Big(\int_0^{T\wedge T_n} \psi(W_s)\Delta\psi(W_s) + (\operatorname{grad}\psi(W_s))^2\, ds\Big) \\ &\leq E\Big(\psi^2(W_0)\Big) + 0 \\ &\quad + E\Big(\int_0^t \big|\psi(W_s)\Delta\psi(W_s) + (\operatorname{grad}\psi(W_s))^2\big|\, ds\Big) \\ &=: C, \end{aligned}$$

	v	ψ	Φ	b
v	—	$v = \frac{1}{2}\frac{\Delta\psi}{\psi}$	$v = \frac{1}{2}((\nabla\Phi)^2 - \Delta\Phi)$	$v = \frac{1}{2}(b^2 + \operatorname{div} b)$
ψ	ψ ist der Grundzustand des Hamiltonop. $H = -\frac{1}{2}\Delta + v$	—	$\psi = e^{-\Phi}$	Sei X Lösung von $dX = b(X_\cdot) \bullet dt + dW$ und μ stationäre Verteilung von X. Dann ist $\psi = \sqrt{\frac{d\mu}{dx}}$.
Φ		$\Phi = -\log\psi$	—	
b		$b = \nabla\log\psi = \frac{\nabla\psi}{\psi}$	$b = -\nabla\Phi$	—

Abbildung 6.1: Tabellarische Darstellung der im Text genannten Zusammenhänge zwischen dem Grundzustand ψ, der Drift b und den Potentialen v und Φ.

und nach dem Lemma ist C endlich. Da (T_n) eine lokalisierende Folge ist, gilt fast sicher $T_n \to \infty$, also $\psi^2(W_{T\wedge T_n}) \to \psi^2(W_T)$ fast sicher für $n \to \infty$. Das Lemma von Fatou liefert schließlich

$$E\big(\psi^2(W_T)\big) \;\le\; \liminf E\big(\psi^2(W_{T\wedge T_n})\big) \;\le\; C \;<\; \infty.$$

Nach dem Voranstehenden ist damit gezeigt, daß $\mathcal{M}$ in $L^2(\mathbb{R}^d)$ beschränkt ist. Lemma 4.2.3 bei v. Weizsäcker und Winkler besagt dann, daß $Z_{\cdot\wedge t}$ für alle $t \ge 0$ und damit auch Z ein Martingal ist.

Setze $b := -\operatorname{grad}\Phi$. Dann ist nach der Girsanov-Formel (in der Form, wie sie in Satz 1.8 steht) der Prozeß B mit

$$B_t \;:=\; W_t - \int_0^{\cdot} b(W_s)\,ds \qquad \text{für alle } t \ge 0$$

eine $(\mathcal{F}_t)$-Brownsche Bewegung auf $(\Omega, \mathcal{F}, Q)$, wobei Q das eindeutig bestimmte Maß mit $dQ = Z_t\,dW$ auf $\mathcal{F}_t$ für alle $t \ge 0$ ist. Also löst $X := W$ auf diesem Raum die stochastische Differentialgleichung (5.4). Ist μ das Wahrscheinlichkeitsmaß auf $\mathbb{R}^d$ mit $d\mu = \psi^2\,dx$, so gilt $b = -\operatorname{grad}\Phi$ sowie $d\mu/dx = \exp(-2\Phi)$, und mit Satz 5.2 folgt, daß X für $\mathcal{L}(X_0) = \mu$ reversibel ist.

Sei umgekehrt $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, B eine Brownsche Bewegung auf diesem Raum, $b\colon \mathbb{R}^d \to \mathbb{R}^d$ lipschitzstetig und X eine reversible Lösung der stochastischen Differentialgleichung

$$dX \;=\; b(X) \bullet dt + dB.$$

Dann hat X eine stationäre Verteilung μ und nach Satz 5.2 hat μ eine echt positive Dichte. Sei ψ die (positive) Wurzel aus dieser Dichte, also $\psi > 0$ und $d\mu = \psi^2 dx$.

Zu diesem Prozeß erhält man wieder eine zugehörige Operatorhalbgruppe $(\bar{P}_t)_{t\ge0}$ auf $L^2(\mu)$ mit Generator $\bar{L}$. Wie bereits gesehen gilt

$$\bar{L}f \;=\; b\cdot\nabla f + \frac{1}{2}\Delta f \qquad \text{für alle } f \in C_c^2(\mathbb{R}^d).$$

Nach Korollar 5.1 ist $\bar{L}$, und damit auch $-\bar{L}$, selbstadjungiert.

Definiere die Abbildung $U\colon L^2(\mathbb{R}^d) \to L^2(\mu)$ durch

$$\big(U\varphi\big)(x) \;:=\; \frac{\varphi(x)}{\psi(x)} \qquad \text{für alle } x \in \mathbb{R}^d,\ \varphi \in L^2(\mathbb{R}^d).$$

Es gilt

$$\|U\varphi\|_{\mu,2}^2 \;=\; \int_{\mathbb{R}^d} \Big(\frac{\varphi}{\psi}\Big)^2 \psi^2 dx \;=\; \int_{\mathbb{R}^d} \varphi^2\, dx \;=\; \|\varphi\|_2^2,$$

d.h. für $\varphi \in L^2(\mathbb{R}^d)$ ist tatsächlich $U\varphi \in L^2(\mu)$. Genauso erhält man

$$\big(U\varphi, U\gamma\big)_\mu \;=\; \int_{\mathbb{R}^d} \frac{\varphi \cdot \gamma}{\psi^2}\, \psi^2 dx \;=\; \int_{\mathbb{R}^d} \varphi \cdot \gamma\; dx \;=\; (\varphi, \gamma)$$

für alle $\varphi, \gamma \in L^2(\mathbb{R}^d)$, d.h. U ist eine Isometrie. Weil genau dann $f \cdot \psi \in L^2(\mathbb{R}^d)$ gilt, wenn $f \in L^2(\mu)$ ist, ist U surjektiv. Daher ist U auch unitär und es gilt $U^* = U^{-1}$.

Sei nun H durch $H = -U^* \bar{L} U$ definiert, also

$$\varphi \in \mathcal{D}(H) \;\Leftrightarrow\; U\varphi \in \mathcal{D}(\bar{L}) \quad \text{und} \quad H\varphi \;=\; -\psi \cdot \bar{L}(\varphi/\psi) \quad \text{für alle } \varphi \in \mathcal{D}(H)$$

beziehungsweise

$$\psi f \in \mathcal{D}(H) \;\Leftrightarrow\; f \in \mathcal{D}(\bar{L}) \quad \text{und} \quad H(\psi f) \;=\; -\psi \cdot \bar{L} f \quad \text{für alle } f \in \mathcal{D}(\bar{L}).$$

Aus dieser Definition folgt insbesondere

$$(H(\psi f), \psi g) \;=\; (-\bar{L} f, g)_\mu \qquad \text{für alle } f, g \in \mathcal{D}(\bar{L}). \tag{6.3}$$

Außerdem ist folgendes Diagramm kommutativ.

$$\begin{array}{ccc} \mathcal{D}(H) \subseteq L^2(\mathbb{R}^d) & \xrightarrow{\;U\;} & \mathcal{D}(\bar{L}) \subseteq L^2(\mu) \\ \big\downarrow H & & \big\downarrow -\bar{L} \\ L^2(\mathbb{R}^d) & \xrightarrow{\;U\;} & L^2(\mu) \end{array}$$

Das nächste Lemma zeigt, daß mit $-\bar{L}$ auch H selbstadjungiert ist.

Lemma. *Seien $\mathcal{H}_1$ und $\mathcal{H}_2$ Hilberträume, $U\colon \mathcal{H}_1 \to \mathcal{H}_2$ unitär und B ein selbstadjungierter linearer Operator auf $\mathcal{H}_2$. Der Operator A auf $\mathcal{H}_1$ sei durch $A = U^* B U$ definiert. Dann ist A selbstadjungiert.*

Beweis. Seien $\varphi, \gamma \in \mathcal{D}(A)$. Dann sind $U\varphi, U\gamma \in \mathcal{D}(B)$, und es gilt

$$\begin{aligned} (A\varphi, \gamma)_1 \;&=\; (U^* B U \varphi, \gamma)_1 \;=\; (BU\varphi, U\gamma)_2 \\ &=\; (U\varphi, BU\gamma)_2 \;=\; (\varphi, U^* BU\gamma)_1 \;=\; (\varphi, A\gamma)_1, \end{aligned}$$

d.h. A ist symmetrisch und es gilt $\mathcal{D}(A) \subseteq \mathcal{D}(A^*)$.

Sei nun $\varphi \in \mathcal{D}(A^*)$. Dann gibt es ein $\varphi^* \in \mathcal{H}_1$ mit

$$(A\gamma, \varphi)_1 \;=\; (\gamma, \varphi^*)_1 \qquad \text{für alle } \gamma \in \mathcal{D}(A).$$

Sei weiter $g \in \mathcal{D}(B)$. Dann ist $U^* g \in \mathcal{D}(A)$ und daher

$$\begin{aligned} (Bg, U\varphi)_2 \;&=\; (UAU^* g, U\varphi)_2 \;=\; (AU^* g, \varphi)_1 \\ &=\; (U^* g, \varphi^*)_1 \;=\; (g, U\varphi^*)_2. \end{aligned}$$

Also ist $U\varphi \in \mathcal{D}(B^*) = \mathcal{D}(B)$, folglich $\varphi \in \mathcal{D}(A)$, und damit auch $\mathcal{D}(A^*) \subseteq \mathcal{D}(A)$. (q.e.d.)

Bezeichnet man mit 1 die konstante Eins-Funktion in $L^2(\mu)$, so ist $\big(\bar{P}_t 1\big)(x) = E^x(1) = 1$, also $\bar{P}_t 1 - 1 = 0$ für alle $t \geq 0$ und folglich $1 \in \mathcal{D}(\bar{L})$ mit $\bar{L}1 = 0$. Nach Definition von H ist damit $\psi = \psi \cdot 1 \in \mathcal{D}(H)$ und $H\psi = 0$.

Sei nun ψ zweimal stetig differenzierbar. Aus Satz 5.2 folgt dann $b = (\nabla\psi)/\psi$, und für $\varphi \in C_c^2(\mathbb{R}^d)$ gilt

$$\begin{aligned}
\frac{1}{2}\Delta\frac{\varphi}{\psi} &= \frac{1}{2}\operatorname{div}\frac{\psi\cdot\nabla\varphi - \varphi\cdot\nabla\psi}{\psi^2} \\
&= \frac{1}{2}\operatorname{div}\frac{\nabla\varphi}{\psi} - \frac{1}{2}\operatorname{div}\frac{\varphi\cdot\nabla\psi}{\psi^2} \\
&= \frac{1}{2}\frac{\psi\cdot\Delta\varphi - \nabla\psi\nabla\varphi}{\psi^2} - \frac{1}{2}\frac{\psi^2(\nabla\varphi\nabla\psi + \varphi\Delta\psi) - 2\psi\nabla\psi\cdot\varphi\nabla\psi}{\psi^4} \\
&= \big(\frac{1}{2}\Delta\varphi - \frac{\Delta\psi}{2\psi}\varphi\big)/\psi - \big(\nabla\varphi\frac{\nabla\psi}{\psi} - \varphi(\frac{\nabla\psi}{\psi})^2\big)/\psi,
\end{aligned}$$

also

$$\begin{aligned}
H\varphi &= -\psi\big(b\cdot\nabla\frac{\varphi}{\psi} + \frac{1}{2}\Delta\frac{\varphi}{\psi}\big) \\
&= -\psi\big(\frac{\nabla\psi}{\psi}\cdot\frac{\psi\nabla\varphi - \varphi\nabla\psi}{\psi^2}\big) - \frac{1}{2}\Delta\varphi + \frac{\Delta\psi}{2\psi}\varphi + \big(\nabla\varphi\frac{\nabla\psi}{\psi} - \varphi\cdot(\frac{\nabla\psi}{\psi})^2\big) \\
&= -\frac{1}{2}\Delta\varphi + \frac{\Delta\psi}{2\psi}\varphi \qquad \text{für alle } \varphi \in C_c^2(\mathbb{R}^d).
\end{aligned}$$

Damit ist H Hamiltonoperator zum Potential $\Delta\psi/(2\psi)$.

Bemerkung. Durch den obigen Zusammenhang zwischen dem Hamiltonoperator H und dem Generator $\bar{L}$ kann man aus der Feynman-Kac-Formel auch Aussagen über zeitabhängige Lösungen einer gestörten Schrödingergleichung gewinnen: Sei wieder wie oben $\psi \in L^2(\mathbb{R}^d)$ mit $H\psi = 0$. Dann gilt

$$f \in L^2(\mu) \quad \Longleftrightarrow \quad \psi\cdot f \in L^2(\mathbb{R}^d)$$

und weiter

$$\begin{aligned}
\partial_t(\psi\cdot f_t) = \big(H+v\big)(\psi\cdot f_t) \;&\Leftrightarrow\; \psi\cdot\partial_t f_t = H(\psi\cdot f_t) + \psi\cdot v f_t \\
&\Leftrightarrow\; \partial_t f_t = -\bar{L}f_t + v\cdot f_t.
\end{aligned}$$

Letzteres ist aber genau die partielle Differentialgleichung aus der Feynman-Kac-Formel. Diese liefert, daß

$$\Big(\exp\big(-\int_0^t v(X_s)\,ds\big)\cdot f_t(X_t)\Big)_{t\geq 0}$$

ein lokales Martingal ist. Unter geeigneten Integrabilitätsbedingungen erhält man, daß dieser

---

Satz („Feynman-Kac-Formel" nach v. Weizsäcker und Winkler, 1990). *Sei $G \subseteq \mathbb{R}^d$ offen, X eine Lösung der stochastischen Differentialgleichung*

$$dX = b(t, X_t)\bullet dt + \sigma(t, X_t)\bullet dB$$

und T eine Stoppzeit mit $X_{\cdot\wedge T} \in G$ fast sicher. Weiter sei $v\colon G \to \mathbb{R}$ lokal beschränkt und borelmeßbar, sowie $f \in C^{1,2}(\mathbb{R}_+ \times G)$ eine Lösung der Gleichung

$$\partial_t f_t = -L_t f_t + v\cdot f_t.$$

Dann ist der Prozeß

$$\Big(\exp\big(-\int_0^{t\wedge T} v(X_s)\,ds\big)\cdot f_{t\wedge T}(X_{t\wedge T})\Big)_{t\geq 0}$$

ein lokales Martingal [...].

Prozeß auch ein Martingal ist. In diesem Falle gilt also

$$f_0(x) \;=\; E^x\Big(\exp\big(-\int_0^t v(X_s)\,ds\big)\cdot f_t(X_t)\Big) \qquad \text{für alle } t \geq 0.$$

Falls v auch negative Werte annehmen kann, ist es allerdings mühsam nachzuweisen, daß der Prozeß ein Martingal ist.

7. Der Ornstein-Uhlenbeck-Prozeß

Zum Abschluß illustriere ich in diesem Kapitel die bisherigen Ergebnisse anhand eines ausführlichen Beispiels. Als Diffusionsprozeß dient hier der Ornstein-Uhlenbeck-Prozeß. Mit den Methoden aus Kapitel 6 stelle ich einen Zusammenhang zum harmonischen Oszillator aus der Quantenmechanik her. Anschließend gebe ich zwei Abschätzungen an, die das Problem, die Konvergenzrate des Bayesrisikos bei der Unterscheidung zweier Ornstein-Uhlenbeck Prozesse zu bestimmen, teilweise lösen.

Sei X eine Lösung der stochastischen Differentialgleichung

$$dX \;=\; -\alpha X \bullet dt + dB.$$

Dann heißt X der **Ornstein-Uhlenbeck-(Geschwindigkeits-)Prozeß** mit Konstante α (siehe z.B. v. Weizsäcker und Winkler (1990), S. 10ff).

Zunächst zeige ich, daß X (unter einer geeigneten Startverteilung) reversibel ist: Die stochastische Differentialgleichung ist die Spezialisierung von (5.4) auf die Drift b mit $b(x) = -\alpha x$ für alle $x \in \mathbb{R}^d$. Definiere $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ durch

$$\Phi(x) \;:=\; \frac{\alpha}{2}x^2 - \frac{d}{4}\log\frac{\alpha}{\pi} \qquad \text{für alle } x \in \mathbb{R}^d.$$

Dann ist $b = -\operatorname{grad}\Phi$ und für das Maß μ mit $d\mu = \exp(-2\Phi)\,dx$ gilt

$$\mu(A) \;=\; \int_A e^{-\alpha x^2}(\alpha/\pi)^{d/2}\,dx \;=\; \left(\frac{1}{2\pi\frac{1}{2\alpha}}\right)^{d/2}\int_A \exp\Big(-\frac{x^2}{2\frac{1}{2\alpha}}\Big)\,dx,$$

d.h. μ ist die d-dimensionale Normalverteilung mit Kovarianzmatrix $\frac{1}{2\alpha}I$. Nach Satz 5.2 ist X dann für $\mathcal{L}(X_0) = \mu$ reversibel.

Wie in Kapitel 5 kann man die Verteilung von X bestimmen: Setze $\psi := \exp(-\Phi)$ und $v := \Delta\psi/(2\psi)$. Das folgende Lemma hilft, v explizit darzustellen.

Lemma 7.1 *Sei $\alpha > 0$ und ψ_α durch $\psi_\alpha(x) := (\alpha/\pi)^{d/4}e^{-\frac{1}{2}\alpha x^2}$ für alle $x \in \mathbb{R}^d$ definiert. Dann ist $\psi_\alpha \in L^2(\mathbb{R}^d) \cap C^2(\mathbb{R}^d)$ mit $\|\psi_\alpha\|_2 = 1$, und es gilt*

$$\partial_j\psi_\alpha \;=\; -\alpha x_j\psi_\alpha \quad \textit{und} \quad \Delta\psi_\alpha \;=\; (\alpha^2x^2 - \alpha\cdot d)\psi_\alpha,$$

sowie

$$(-\Delta\psi_\alpha, \psi_\alpha) \;=\; \frac{d}{2}\cdot\alpha \quad \textit{und} \quad (x^2\psi_\alpha, \psi_\alpha) \;=\; \frac{d}{2}\cdot\frac{1}{\alpha}.$$

Beweis. Es ist

$$\psi_\alpha^2 \;=\; \Big(\frac{1}{2\pi\frac{1}{2\alpha}}\Big)^{d/2}\exp\Big(-\frac{x^2}{2\frac{1}{2\alpha}}\Big),$$

d.h. ψ_α^2 ist die Dichte der Normalverteilung auf $\mathbb{R}^d$ mit Kovarianzmatrix $\frac{1}{2\alpha}I$. Daraus folgt natürlich

$$\|\psi_\alpha\|_2^2 \;=\; \int_{\mathbb{R}^d}\psi_\alpha^2(x)\,dx \;=\; 1.$$

Zunächst erhält man nun

$$\partial_j\psi_\alpha(x) \;=\; \Big(\frac{\alpha}{\pi}\Big)^{d/4}(-\frac{1}{2}\alpha)2x_je^{-\frac{1}{2}\alpha x^2} \;=\; -\alpha x_j\psi_\alpha(x)$$

und

$$\partial_j^2\psi_\alpha(x) \;=\; \Big(\frac{\alpha}{\pi}\Big)^{d/4}(\alpha^2x_j^2 - \alpha)e^{-\frac{1}{2}\alpha x^2} \;=\; (\alpha^2x_j^2 - \alpha)\psi_\alpha(x)$$

für alle $x \in \mathbb{R}^d$ und folglich

$$\Delta\psi_\alpha \;=\; \sum_{j=1}^d \partial_j^2 \psi_\alpha \;=\; \sum_{j=1}^d (\alpha^2 x_j^2 - \alpha)\psi_\alpha \;=\; (\alpha^2 x^2 - \alpha \cdot d)\psi_\alpha .$$

Weiter gilt

$$\begin{aligned}
\int_{\mathbb{R}^d} x_j^2 \psi_\alpha^2(x)\,dx \;&=\; \Big(\frac{\alpha}{\pi}\Big)^{d/2} \int_{\mathbb{R}} x_j^2 \mathrm{e}^{-\alpha x_j^2}\,dx_j \;\cdot\; \prod_{k \neq j} \int_{\mathbb{R}} \mathrm{e}^{-\alpha x_k^2}\,dx_k \\
&=\; \Big(\frac{\alpha}{\pi}\Big)^{d/2} \frac{\sqrt{\pi}}{2\sqrt{\alpha}^3} \cdot \Big(\frac{\pi}{\alpha}\Big)^{(d-1)/2} \\
&=\; \frac{1}{2\alpha}
\end{aligned}$$

und damit

$$(x^2\psi_\alpha, \psi_\alpha) \;=\; \sum_{j=1}^d (x_j^2 \psi_\alpha, \psi_\alpha) \;=\; \frac{d}{2\alpha}.$$

Die verbleibende Behauptung läßt sich auf das bisher Gezeigte zurückführen: Aus der Formel für $\Delta\psi_\alpha$ ergibt sich

$$\begin{aligned}
(-\Delta\psi_\alpha, \psi_\alpha) \;&=\; -\alpha^2 (x^2\psi_\alpha, \psi_\alpha) + \alpha d(\psi_\alpha, \psi_\alpha) \\
&=\; -\alpha^2 \frac{d}{2\alpha} + \alpha d \;=\; \frac{\alpha d}{2}.
\end{aligned}$$

(q.e.d.)

Mit den Bezeichnungen des Lemmas ist $\psi = \psi_\alpha$, und es folgt

$$v(x) \;=\; \frac{\Delta\psi}{2\psi} \;=\; \frac{1}{2}(\alpha^2 x^2 - \alpha d) \qquad \text{für alle } x \in \mathbb{R}^d.$$

Die Bemerkung nach dem Beweis von Satz 5.2 zeigt, daß der Prozeß (für $X_0 = 0$) die Beziehung

$$P(X \in A) \;=\; \int_A \varphi_t\,dW \qquad \text{für alle } A \in \mathcal{F}_t, t \geq 0$$

erfüllt, wobei

$$\begin{aligned}
\varphi_t \;&=\; \frac{\psi(W_t)}{\psi(W_0)} \exp\big(-\int_0^t v(W_s)\,ds\big) \\
&=\; \exp\Big(-\Phi(W_t) + \Phi(W_0) - \frac{1}{2}\int_0^t \alpha^2 W_s^2 - \alpha d\,ds\Big) \\
&=\; \exp\Big(-\frac{\alpha}{2}W_t^2 + \frac{\alpha}{2}W_0^2 + t \cdot \frac{\alpha d}{2} - \frac{\alpha^2}{2}\int_0^t W_s^2\,ds\Big) \\
&=\; \exp\Big(-\frac{\alpha}{2}(W_t^2 - t \cdot d) - \frac{\alpha^2}{2}\int_0^t W_s^2\,ds\Big) \qquad W\text{-fast sicher}
\end{aligned} \tag{7.1}$$

ist.

Mit den Methoden aus Kapitel 6 findet man ein zugehöriges physikalisches System. Dort wird gezeigt, daß der Hamiltonoperator H zum Potential v existiert, und daß

$$H\psi \;=\; 0$$

gilt. Mit dem hier gefundenen, speziellen v ist

$$H\varphi(x) \;=\; -\frac{1}{2}\Delta\varphi(x) + v(x)\cdot\varphi(x) \;=\; -\frac{1}{2}\Delta\varphi(x) + \frac{1}{2}\big(\alpha^2 x^2 - \alpha d\big)\cdot\varphi(x) \tag{7.2}$$

für alle $x \in \mathbb{R}^d$ und $\varphi \in C_c^2(\mathbb{R}^d)$.

Das zugrundeliegende System könnte z.B. Folgendes sein: Ein sehr kleines Teilchen wird von einer Kraft F zu einem festen Punkt (der „Ruhelage“) hin gezogen; es ist

$$F(x) \;=\; -\operatorname{grad} v(x) \;=\; -\alpha^2 x,$$

d.h. die Stärke der Kraft ist proportional zur Auslenkung des Teilchens aus der Ruhelage. Außer dieser Kraft wirken keine äußeren Einflüsse auf das Teilchen. Dieses System ist in der Mechanik als **harmonischer Oszillator** bekannt.

Im verbleibenden Teil des Kapitels gebe ich noch eine Teillösung für die Bestimmung des Bayesrisiko bei der Unterscheidung zweier Ornstein-Uhlenbeck-Prozesse an. Dazu betrachte ich zwei Ornstein-Uhlenbeck-Prozesse mit unterschiedlichen Driftkonstanten und gehe zunächst wie in Kapitel 3 vor. Sei also wieder $\Theta = \{1,2\}$, und für $\vartheta \in \Theta$ weiter $\alpha_\vartheta > 0$ und X^ϑ die Lösung der stochastischen Differentialgleichung

$$dX^\vartheta \;=\; -\alpha_\vartheta X^\vartheta \bullet dt + dB$$

mit $X_0^\vartheta = 0$. Ohne Beschränkung der Allgemeinheit kann man $\alpha_1 \le \alpha_2$ annehmen.

Setze $P_T^\vartheta \;:=\; \mathcal{L}\big(X^\vartheta\big)\big|_{\mathcal{F}_T}$ für alle $T \;\ge\; 0$. Zuletzt sei φ_T^ϑ die Dichte von P_T^ϑ aus Gleichung (7.1). Für $\lambda \in \mathcal{M}_1(\Theta)$ und $T \ge 0$ ist das Bayesrisiko dann

$$B(\lambda,T) \;=\; \lambda_1 P_T^1(\lambda_1\varphi_T^1 < \lambda_2\varphi_T^2) \;+\; \lambda_2 P_T^2(\lambda_1\varphi_T^1 \ge \lambda_2\varphi_T^2). \tag{7.3}$$

Für $\alpha_1 = \alpha_2$ gilt nach obiger Gleichung trivialerweise $B(\lambda,T) = \min\{\lambda_1,\lambda_2\}$. Im Folgenden sei daher $\alpha_1 < \alpha_2$. Dann gilt

$$\begin{aligned}
\lambda_1\varphi_T^1 \;\le\; \lambda_2\varphi_T^2 \quad &\Longleftrightarrow \quad -\frac{\alpha_1}{2}(W_T^2 - T\cdot d) \;-\; \frac{\alpha_1^2}{2}\int_0^T W_s^2\,ds \\
&\qquad\qquad \le\; -\frac{\alpha_2}{2}(W_T^2 - T\cdot d) \;-\; \frac{\alpha_2^2}{2}\int_0^T W_s^2\,ds + \log\frac{\lambda_2}{\lambda_1} \\
&\Longleftrightarrow \quad \frac{\alpha_2-\alpha_1}{2}(W_T^2 - T\cdot d) \;+\; \frac{\alpha_2^2-\alpha_1^2}{2}\int_0^T W_s^2\,ds \;\le\; \log\frac{\lambda_2}{\lambda_1} \\
&\Longleftrightarrow \quad \frac{W_T^2}{T} \;+\; (\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2\,ds \;\le\; d + \eta_T
\end{aligned}$$

mit $\eta_T := \frac{2}{(\alpha_2-\alpha_1)T}\log\frac{\lambda_2}{\lambda_1}$, also

$$P_T^1\big(\lambda_1\varphi_T^1 \le \lambda_2\varphi_T^2\big) \;=\; P_T^1\Big(\frac{W_T^2}{T} + (\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2\,ds \;\le\; d+\eta_T\Big). \tag{7.4}$$

Ich betrachte zunächst den Ausdruck $(\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2\,ds$.

Bezeichnet $\lambda\big|_{[0;t]}$ das Lebesguemaß auf dem Intervall $[0;t]$ und ist das Wahrscheinlichkeitsmaß $L_t^\omega \in \mathcal{M}_1(\mathbb{R}^d)$ für $\omega \in C\big([0;\infty), \mathbb{R}^d\big)$ wieder durch

$$L_t^\omega \;:=\; \frac{1}{t}\,\lambda\big|_{[0;t]} \circ \big(\omega\big|_{[0;t]}\big)^{-1}$$

definiert, so ist

$$\frac{1}{t}\int_0^t \omega_s^2\,ds \;=\; \frac{1}{t}\int_{\mathbb{R}^d} x^2\,d\big(\lambda \circ (\omega|_{[0;t]})^{-1}\big)(x) \;=\; \int_{\mathbb{R}^d} x^2\,dL_t^\omega(x)$$

und für alle $C \in \mathbb{R}$ gilt

$$P\Big(\frac{1}{t}\int_0^t (X_s^\vartheta)^2\,ds \le C\Big) \;=\; P_t^\vartheta\Big(L_t \in \{\,\nu \in \mathcal{M}_1(\mathbb{R}^d) \mid \textstyle\int x^2\,d\nu \le C\,\}\Big).$$

Aus Satz 5.3 kann man Aussagen über diesen Ausdruck gewinnen.

Die Menge $M := \big\{\,\nu \in \mathcal{M}_1(\mathbb{R}^d) \;\big|\; \int x^2\,d\nu \le C\,\big\}$ ist bezüglich der schwachen Topologie auf $\mathcal{M}_1(\mathbb{R}^d)$ abgeschlossen: Sei $\nu \in \mathcal{M}_1(\mathbb{R}^d) \setminus M$. Für $n \to \infty$ konvergiert $\min(x^2, n) \to x^2$ punktweise monoton von unten. Mit dem Satz von der monotonen Konvergenz folgt

$$\lim_{n\to\infty} \int \min(x^2, n)\,d\nu \;=\; \int x^2\,d\nu \;>\; C.$$

Sei nun $\varepsilon > 0$ so klein, daß

$$\int x^2\,d\nu - \varepsilon \;>\; C + \varepsilon$$

gilt. Dann existiert ein $n \in \mathbb{N}$ mit

$$\int_{\mathbb{R}^d} \min(x^2, n)\,d\nu(x) \;\ge\; \int_{\mathbb{R}^d} x^2\,d\nu(x) - \varepsilon \;>\; C + \varepsilon.$$

In der schwachen Topologie ist

$$U \;:=\; \Big\{\,\alpha \in \mathcal{M}_1(\mathbb{R}^d) \;\Big|\; \Big|\int \min(x^2, n)\,d\alpha - \int \min(x^2, n)\,d\nu\Big| < \varepsilon\,\Big\}$$

eine Umgebung von ν. Für $\alpha \in U$ gilt aber

$$\int x^2\,d\alpha \;\ge\; \int \min(x^2, n)\,d\alpha \;=\; \int \min(x^2, n)\,d\nu - \varepsilon \;>\; C,$$

also $\alpha \in M^{\mathsf{C}}$. Damit ist $U \subseteq M^{\mathsf{C}}$, und es folgt, daß M^{C} offen und damit M abgeschlossen ist.

Nun bestimme ich das Infimum der Ratenfunktion aus Satz 5.3 auf der Menge M. Dazu verwende ich den folgenden Satz, der eine Entsprechung zu der allgemeinen Unschärferelation aus der Quantenmechanik formuliert.

Definition. Für lineare Operatoren A und B definiere den **Kommutator** von A und B als

$$[A, B] \;:=\; AB - BA \qquad \text{mit} \qquad \mathcal{D}\big([A,B]\big) := \mathcal{D}(AB) \cap \mathcal{D}(BA).$$

Satz (Allgemeine Unschärferelation). *Seien A_j und B_j für $j \in \{1,\dots,d\}$ lineare Operatoren auf einem Hilbertraum. Die Operatoren A_j seien symmetrisch und B_j antisymmetrisch, also*

$$(A_j f, g) \;=\; (f, A_j g) \qquad \text{für alle } f,g \in \mathcal{D}(A)$$

und

$$(B_j f, g) \;=\; -(f, B_j g) \qquad \text{für alle } f,g \in \mathcal{D}(B).$$

Setze $A^2 := \sum_{j=1}^d A_j^2$ und $B^2 := \sum_{j=1}^d B_j^2$. Ist dann $f \in \mathcal{D}(A^2) \cap \mathcal{D}(B^2) \cap \mathcal{D}\big(\sum_{j=1}^d [A_j, B_j]\big)$, so sind $(A^2 f, f)$ und $(-B^2 f, f)$ positiv, und es gilt

$$\sqrt{(A^2 f, f)} \cdot \sqrt{(-B^2 f, f)} \;\geq\; \frac{1}{2}\Big|\sum_{j=1}^d \big([A_j, B_j] f, f\big)\Big|.$$

Beweis. Zunächst ist

$$\begin{aligned} (A^2 f, f) \;&=\; \sum_{j=1}^d (A_j^2 f, f) \;=\; \sum_{j=1}^d (A_j f, A_j f) \;\geq\; 0 \qquad \text{und} \\ (-B^2 f, f) \;&=\; \sum_{j=1}^d (-B_j^2 f, f) \;=\; \sum_{j=1}^d (B_j f, B_j f) \;\geq\; 0. \end{aligned}$$

Für $\lambda \in \mathbb{R}$ gilt weiter

$$\begin{aligned} 0 \;&\leq\; \sum_{j=1}^d \big((A_j + \lambda B_j) f, (A_j + \lambda B_j) f\big) \\ &=\; (A^2 f, f) - \lambda^2 (B^2 f, f) + \lambda \sum_{j=1}^d (A_j B_j f - B_j A_j f, f), \end{aligned}$$

also

$$(A^2 f, f) - \lambda^2 (B^2 f, f) \;\geq\; -\lambda \sum_{j=1}^d \big([A_j, B_j] f, f\big)$$

und mit $-\lambda$ statt λ auch

$$(A^2 f, f) - \lambda^2 (B^2 f, f) \;\geq\; \lambda \sum_{j=1}^d \big([A_j, B_j] f, f\big).$$

Zusammen ergibt das

$$(A^2 f, f) + \lambda^2 (-B^2 f, f) \;\geq\; \Big|\lambda \sum_{j=1}^d \big([A_j, B_j] f, f\big)\Big| \qquad \text{für alle } \lambda \in \mathbb{R}$$

und insbesondere

$$\begin{aligned} \sqrt{(A^2 f, f)} \cdot \sqrt{(-B^2 f, f)} \;&=\; \frac{1}{2}\sqrt{\tfrac{(-B^2 f, f)}{(A^2 f, f)}}\left((A^2 f, f) \;+\; \sqrt{\tfrac{(A^2 f, f)}{(-B^2 f, f)}}^{\,2} (-B^2 f, f)\right) \\ &\geq\; \frac{1}{2}\sqrt{\tfrac{(-B^2 f, f)}{(A^2 f, f)}}\;\left|\sqrt{\tfrac{(A^2 f, f)}{(-B^2 f, f)}}\;\sum_{j=1}^d \big([A_j, B_j] f, f\big)\right| \\ &=\; \frac{1}{2}\Big|\sum_{j=1}^d \big([A_j, B_j] f, f\big)\Big|. \end{aligned}$$

(q.e.d.)

Da ich hier den Hilbertraum $L^2(\mathbb{R}^d)$ statt $L^2_{\mathbb{C}}(\mathbb{R}^d)$ verwende, kann ich im Beweis nicht, wie sonst üblich, die Operatoren iA und iB betrachten. Daher erscheint das Ergebnis in einer etwas ungewohnten Form.

Als Spezialfall erhält man das Analogon zur Heisenbergschen Unschärferelation: Für $j \in \{1, \dots, d\}$ definiere die Operatoren A_j und B_j auf $L^2(\mathbb{R}^d)$ durch

$$A_j\varphi(x) \;:=\; x_j\varphi(x) \qquad \text{für alle } x \in \mathbb{R}^d, \varphi \in L^2(\mathbb{R}^d) \cap C_0^2(\mathbb{R}^d)$$

und

$$B_j\varphi(x) \;:=\; \partial_j\varphi(x) \qquad \text{für alle } x \in \mathbb{R}^d, \varphi \in L^2(\mathbb{R}^d) \cap C_0^2(\mathbb{R}^d).$$

Dann folgt $(A_j\varphi, \psi) = (\varphi, A_j\psi)$ und mit partieller Integration auch $(B_j\varphi, \psi) = -(\varphi, B_j\psi)$ für alle $\varphi, \psi \in C_0^2(\mathbb{R}^d)$. Weiter ist

$$\begin{aligned}[A_j, B_j]\varphi(x) &= A_jB_j\varphi(x) - B_jA_j\varphi(x) \\ &= x_j\partial_j\varphi(x) - \partial_j\big(x_j\varphi(x)\big) \\ &= x_j\partial_j\varphi(x) - \varphi(x) - x_j\partial_j\varphi(x) \\ &= -\varphi(x).\end{aligned}$$

Sei nun $\varphi \in C_0^2(\mathbb{R}^d)$. Dann liefert die allgemeine Unschärferelation

$$\sqrt{(A^2\varphi, \varphi)} \cdot \sqrt{(-B^2\varphi, \varphi)} \;\geq\; \frac{1}{2}\Big|-\sum_{j=1}^{d}(\varphi, \varphi)\Big|,$$

also

$$\sqrt{(x^2\varphi, \varphi)} \cdot \sqrt{(-\Delta\varphi, \varphi)} \;\geq\; \frac{d}{2}\|\varphi\|_2^2. \tag{7.5}$$

Bemerkung. Lemma 7.1 zeigt, daß es für alle $(p, q) \in \mathbb{R}_+ \times \mathbb{R}_+$ mit $p \cdot q = d^2/4$ eine Funktion $\varphi \in L^2(\mathbb{R}^d) \cap C_0^2(\mathbb{R}^d)$ mit $\|\varphi\|_2 = 1$ gibt, die

$$(-\Delta\varphi, \varphi) \;=\; p \quad \text{und} \quad (x^2\varphi, \varphi) \;=\; q$$

erfüllt (nämlich $\varphi = \psi_{d/(2a)}$ mit den dortigen Bezeichnungen). Also ist Ungleichung (7.5) scharf.

Lemma 7.2 *Seien $b, c > 0$. Dann gilt*

$$\inf\{\, p + q \mid p, q \geq 0 \textit{ mit } p \cdot q \geq c^2, q \leq b \,\} \;=\; \begin{cases} b + c^2/b, & \textit{falls } b < c, \textit{ und} \\ 2c & \textit{sonst,} \end{cases}$$

sowie

$$\inf\{\, p + q \mid p, q \geq 0 \textit{ mit } p \cdot q \geq c^2, q \geq b \,\} \;=\; \begin{cases} b + c^2/b, & \textit{falls } b > c, \textit{ und} \\ 2c & \textit{sonst.} \end{cases}$$

Beweis (vgl. Abb. 7.1). Definiere $f\colon (0; \infty) \to \mathbb{R}$ durch

$$f(t) \;:=\; t + c^2/t \qquad \text{für alle } t > 0.$$

Dann ist $f'(t) = 1 - c^2 \cdot t^{-2}$, d.h. f ist auf dem Intervall $(0; c]$ monoton fallend, auf $[c; \infty)$ monoton steigend und nimmt in c mit $f(c) = 2c$ sein globales Minimum an. Für $(p, q) \in \mathbb{R}_+ \times \mathbb{R}_+$ mit $p \cdot q \geq c^2$ gilt weiter

$$p + q \;\geq\; c^2/q + q \;=\; f(q),$$

also

$$\inf\{\, p + q \mid p \cdot q \geq c^2, q \in A \,\} \;=\; \inf\{\, f(q) \mid q \in A \,\}$$

für alle $A \subseteq (0; \infty)$.

Ist $c \in A$, so folgt

$$\inf\{\, p + q \mid p \cdot q \geq c^2, q \in A \,\} \;=\; f(c) \;=\; 2c.$$

Ist $c \notin A$ und $A = (0; b]$ bzw. $A = [b; \infty)$, so folgt

$$\inf\{\, p + q \mid p \cdot q \geq c^2, q \in A \,\} \;=\; f(b).$$

(q.e.d.)

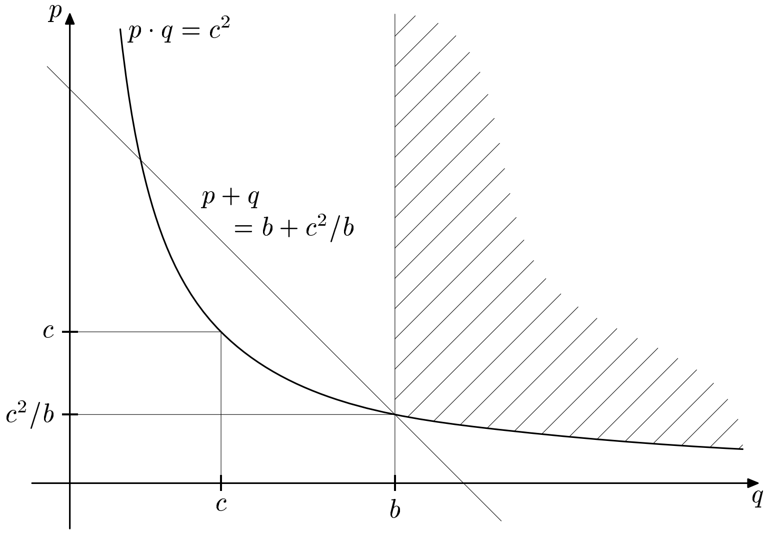

Abbildung 7.1

Satz 7.1 *Sei $r > 0$. Dann gilt*

$$\inf\{\, J_{\mathcal{E}}(\nu) \bigm| \nu \in \mathcal{M}_1(\mathbb{R}^d) \text{ mit } \textstyle\int x^2\, d\nu \le r \,\} = \begin{cases} \dfrac{\alpha^2}{2r}\Big(\dfrac{d}{2\alpha} - r\Big)^2, & \text{falls } r < \dfrac{d}{2\alpha}, \text{ und} \\ 0 & \text{sonst.} \end{cases}$$

Beweis. Sei zuerst $0 < r < d/(2\alpha)$. Weiter sei $\nu \in \mathcal{M}_1(\mathbb{R}^d)$ mit $\int x^2\, d\nu \le r$ und $J_{\mathcal{E}}(\nu) < \infty$. Dann ist $\nu \ll \mu$ mit $d\nu = f^2\, d\mu$ für ein $f \in \mathcal{D}(\mathcal{E})$. Also existiert eine Folge $(f_n)_{n\in\mathbb{N}}$ in $C_c^\infty(\mathbb{R}^d)$ mit $f_n \to f$ in $L^2(\mu)$ und

$$J_{\mathcal{E}}(\nu) \;=\; \mathcal{E}(f,f) \;=\; \lim_{n\to\infty} \mathcal{E}_0(f_n, f_n).$$

Wie Lemma 5.3 zeigt, kann man $f \ge 0$ und $f_n \ge 0$ für alle $n \in \mathbb{N}$ annehmen und die f_n monoton fallend wählen. Die L^2-Konvergenz von f_n gegen f impliziert auch μ-fast sichere Konvergenz. Mit dem Satz von der monotonen Konvergenz erhält man dann

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} x^2 f_n^2(x)\, d\mu(x) \;=\; \int_{\mathbb{R}^d} x^2 f^2(x)\, d\mu(x) \;=\; \int_{\mathbb{R}^d} x^2\, d\nu(x) \;\le\; r.$$

ψ bezeichne wieder den Grundzustand von H. Setze $\varphi_n := \psi f_n$ und $c_n := \|\varphi_n\|_2^2 \alpha d/4$. Wegen $\|f\|_{\mu,2} = 1$ gilt dann $c_n \to \alpha d/4$ für $n \to \infty$. Sei ε so klein, daß $r + \varepsilon < d/(2\alpha) - \varepsilon$ gilt. Daher existiert ein $N \in \mathbb{N}$ mit

$$(\frac{1}{2}\alpha^2 x^2 \varphi_n, \varphi_n) \;=\; \frac{\alpha^2}{2} \int x^2 f_n^2(x)\, d\mu(x) \;<\; \frac{\alpha^2}{2}(r+\varepsilon) \qquad \text{für alle } n \ge N$$

und

$$c_n \;\ge\; \frac{\alpha d}{4} - \frac{\alpha^2}{2}\varepsilon \qquad \text{für alle } n \ge N.$$

Die Unschärferelation aus Formel (7.5) liefert

$$(-\frac{1}{2}\Delta\varphi_n, \varphi_n) \cdot (\frac{1}{2}\alpha^2 x^2 \varphi_n, \varphi_n) \quad\ge\quad \frac{1}{2}\frac{\alpha^2}{2}\frac{d^2}{4}\|\varphi_n\|_2^4 \quad=\quad c_n^2.$$

Jetzt kann man Lemma 7.2 anwenden. Wegen

$$\frac{\alpha^2}{2}(r+\varepsilon) \quad<\quad \frac{\alpha^2}{2}(\frac{d}{2\alpha} - \varepsilon) \quad=\quad \frac{\alpha d}{4} - \frac{\alpha^2}{2}\varepsilon \quad\le\quad c_n$$

liefert es

$$(-\frac{1}{2}\Delta\varphi_n, \varphi_n) + (\frac{1}{2}\alpha^2 x^2 \varphi_n, \varphi_n) \quad \geq \quad \frac{\alpha^2}{2}(r+\varepsilon) + \frac{2c_n^2}{\alpha^2(r+\varepsilon)}, \tag{7.6}$$

also

$$\begin{aligned}
\mathcal{E}_0(f_n, f_n) &\stackrel{\text{Lemma 5.2}}{=} (-\bar{L}f_n, f_n)_\mu \\
&\stackrel{(6.3)}{=} (H\varphi_n, \varphi_n) \\
&\stackrel{(7.2)}{=} (-\frac{1}{2}\Delta\varphi_n, \varphi_n) + (\frac{1}{2}\alpha^2 x^2\varphi_n, \varphi_n) - (\frac{\alpha d}{2}\varphi_n, \varphi_n) \\
&\stackrel{(7.6)}{\geq} \frac{\alpha^2}{2}(r+\varepsilon) + \frac{2c_n^2}{\alpha^2(r+\varepsilon)} - \frac{\alpha d}{2}\|\varphi_n\|_2^2
\end{aligned}$$

für alle $n \geq N$. Damit erhält man

$$\begin{aligned}
J_{\mathcal{E}}(\nu) &= \lim_{n\to\infty} \mathcal{E}_0(f_n, f_n) \\
&\geq \frac{\alpha^2}{2}(r+\varepsilon) + \frac{2\alpha^2 d^2}{16\alpha^2(r+\varepsilon)} - \frac{\alpha d}{2}
\end{aligned}$$

für alle $\varepsilon > 0$ und folglich auch

$$J_{\mathcal{E}}(\nu) \quad \geq \quad \frac{\alpha^2 r}{2} + \frac{d^2}{8r} - \frac{\alpha d}{2} \quad = \quad \frac{\alpha^2}{2r}\Big(\frac{d}{2\alpha} - r\Big)^2.$$

Ich zeige nun, daß diese Ungleichung scharf ist: Mit den Bezeichnungen aus Lemma 7.1 sei $\varphi := \psi_{d/(2r)}$ und $f := \varphi/\psi$. Weiter sei ν das Maß auf $\mathbb{R}^d$ mit $d\nu = f^2\,d\mu$. Wegen

$$\int f^2\,d\mu \ = \ \int \varphi^2(x)\,dx \stackrel{\text{Lemma 7.1}}{=} 1$$

ist ν ein Wahrscheinlichkeitsmaß auf $\mathbb{R}^d$, und es gilt

$$\begin{aligned}
\int_{\mathbb{R}^d} x^2\,d\nu(x) &= \int_{\mathbb{R}^d} x^2 f^2(x)\,d\mu(x) \\
&= \int_{\mathbb{R}^d} x^2 \psi_{d/(2r)}^2(x)\,d\mu(x) \\
&\stackrel{\text{Lemma 7.1}}{=} \frac{d}{2}\cdot\frac{2r}{d} \ = \ r.
\end{aligned}$$

Das Maß ν ist also ein zulässiger Kandidat für die Bildung des Infimums.

Nach Lemma 5.4 ist $f \in \mathcal{D}(\mathcal{E})$, und es gilt

$$\mathcal{E}(f,f) \ = \ \frac{1}{2}\int (\text{grad}\, f)^2\,d\mu.$$

Weiter ist

$$\psi \cdot \partial_j f \ = \ \frac{\partial_j\varphi\cdot\psi - \partial_j\psi\cdot\varphi}{\psi} \stackrel{\text{Lemma 7.1}}{=} -\frac{d}{2r}x_j\varphi + \alpha x_j\varphi,$$

also

$$\begin{aligned}
J_{\mathcal{E}}(\nu) \ = \ \mathcal{E}(f,f) &= \frac{1}{2}\sum_{j=1}^d \int (\partial_j f)^2\,d\mu \\
&= \frac{1}{2}\sum_{j=1}^d \int \big(\alpha - \frac{d}{2r}\big)^2 x_j^2\varphi^2\,dx \\
&\stackrel{\text{Lemma 7.1}}{=} \frac{1}{2}\big(\alpha - \frac{d}{2r}\big)^2\cdot\frac{d}{2}\cdot\frac{2r}{d} \\
&= \frac{\alpha^2}{2r}\Big(\frac{d}{2\alpha} - r\Big)^2.
\end{aligned}$$

Damit ist für den Fall $r < d/(2\alpha)$ alles gezeigt.

Sei nun noch $r \geq d/(2\alpha)$. Aus Formel (5.7) entnimmt man, daß stets $J_\mathcal{E} \geq 0$ gilt. Also ist nur noch zu zeigen, daß diese Schranke auch angenommen wird. Wieder aus Lemma 7.1 erhält man

$$\int x^2 \, d\mu \;=\; \int x^2 \psi_\alpha^2(x) \, dx \;=\; \frac{d}{2\alpha} \;\leq\; r.$$

Also ist hier die stationäre Verteilung μ ein Kandidat für die Bildung des Infimums. Es gilt aber $d\mu = 1 \, d\mu$ und mit Lemma 5.4 weiter $1 \in \mathcal{D}(\mathcal{E})$ sowie

$$J_\mathcal{E}(\mu) \;=\; \mathcal{E}(1,1) \;=\; \int_{\mathbb{R}^d} (\operatorname{grad} 1)^2 \, d\mu \;=\; 0.$$

(q.e.d.)

Bemerkung. Indem man die zweite Aussage aus Lemma 7.2 verwendet, kann man genauso

$$\begin{aligned} &\inf\{\, J_\mathcal{E}(\nu) \mid \nu \in \mathcal{M}_1(\mathbb{R}^d) \text{ mit } \textstyle\int x^2 \, d\nu > r \,\} \\ &\qquad = \begin{cases} \dfrac{\alpha^2}{2r}\Big(r - \dfrac{d}{2\alpha}\Big)^2, & \text{falls } r > \dfrac{d}{2\alpha}, \text{ und} \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

beweisen.

Zusammengenommen ergeben die Rechnungen auf den letzten Seiten

$$\begin{aligned} &\limsup_{t\to\infty} \frac{1}{t} \log P_t^1\Big((\alpha_1+\alpha_2)\frac{1}{t}\int_0^t W_s^2 \, ds \leq C\Big) \\ &\quad = \limsup_{t\to\infty} \frac{1}{t} \log P_t^1\Big(L_t \in \Big\{\, \nu \in \mathcal{M}_1(\mathbb{R}^d) \;\Big|\; \int x^2 \, d\nu \leq \frac{C}{\alpha_1+\alpha_2} \,\Big\}\Big) \\ &\quad \overset{\text{Satz 5.3}}{\leq} -\inf\{\, J_\mathcal{E}(\nu) \mid \nu \in \mathcal{M}_1(\mathbb{R}^d) \text{ mit } \textstyle\int x^2 \, d\nu \leq \frac{C}{\alpha_1+\alpha_2} \,\} \\ &\quad \overset{\text{Satz 7.1}}{=} \begin{cases} -\dfrac{\alpha_1^2(\alpha_1+\alpha_2)}{2C}\Big(\dfrac{d}{2\alpha_1} - \dfrac{C}{\alpha_1+\alpha_2}\Big)^2, & \text{falls } C \leq d(\alpha_1+\alpha_2)/2\alpha_1, \text{ und} \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Da diese Rate stetig von C abhängt, kann man, wie im dritten Kapitel ausgeführt, den Term η_T in Formel (7.4) vernachlässigen und erhält

$$\begin{aligned} &\limsup_{T\to\infty} \frac{1}{T} \log P_T^1(\lambda_1\varphi_T^1 \leq \lambda_2\varphi_T^2) \\ &\qquad = \limsup_{T\to\infty} \frac{1}{T} \log P_T^1\Big(\frac{W_T^2}{T} + (\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2 \, ds \;\leq\; d + \eta_T\Big) \\ &\qquad \leq \limsup_{T\to\infty} \frac{1}{T} \log P_T^1\Big((\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2 \, ds \;\leq\; d + \eta_T\Big) \\ &\qquad \leq -\frac{\alpha_1^2(\alpha_1+\alpha_2)}{2d}\Big(\frac{d}{2\alpha_1} - \frac{d}{\alpha_1+\alpha_2}\Big)^2 \\ &\qquad = -\frac{d(\alpha_2-\alpha_1)^2}{8(\alpha_1+\alpha_2)}. \end{aligned} \tag{7.7}$$

Dies ist eine Abschätzung für die obere Rate der Wahrscheinlichkeit, irrtümlich die Konstante α_2 anzunehmen, obwohl man einen Prozeß mit Konstante $\alpha_1 < \alpha_2$ beobachtete. Da die Menge $\{\, \nu \in \mathcal{M}_1(\mathbb{R}^d) \mid \int x^2 \, d\nu \leq C \,\}$ bezüglich der schwachen Topologie auf $\mathcal{M}_1(\mathbb{R}^d)$ leeres Inneres hat, erhält man aus Satz 5.3 keine sinnvolle Abschätzung für die untere Rate dieser Wahrscheinlichkeit.

Für die zweite Fehlerwahrscheinlichkeit aus der Darstellung des Bayesrisikos in Formel (7.3) verhält es sich genau umgekehrt: man erhält eine Abschätzung für die untere Rate, aber keine sinnvolle Aussage über die obere Rate. Zunächst kann man aus der Bemerkung nach Satz 7.1 folgende Abschätzung gewinnen:

$$\begin{aligned}
&\liminf_{t\to\infty}\frac{1}{t}\log P_t^2\Big((\alpha_1+\alpha_2)\frac{1}{t}\int_0^t W_s^2\,ds > C\Big)\\
&\quad= \liminf_{t\to\infty}\frac{1}{t}\log P_t^2\Big(L_t\in\Big\{\,\nu\in\mathcal{M}_1(\mathbb{R}^d)\;\Big|\;\int x^2\,d\nu > \frac{C}{\alpha_1+\alpha_2}\,\Big\}\Big)\\
&\quad\overset{\text{Satz 5.3}}{\geq} -\inf\big\{\,J_{\mathcal{E}}(\nu)\;\big|\;\nu\in\mathcal{M}_1(\mathbb{R}^d)\text{ mit }\textstyle\int x^2\,d\nu > \frac{C}{\alpha_1+\alpha_2}\,\big\}\\
&\quad= \begin{cases} -\dfrac{\alpha_2^2(\alpha_1+\alpha_2)}{2C}\Big(\dfrac{C}{\alpha_1+\alpha_2}-\dfrac{d}{2\alpha_2}\Big)^2, & \text{falls } C > d(\alpha_1+\alpha_2)/2\alpha_2,\text{ und}\\ 0 & \text{sonst.}\end{cases}
\end{aligned}$$

Daraus erhält man dann

$$\begin{aligned}
&\liminf_{T\to\infty}\frac{1}{T}\log P_T^2\big(\lambda_1\varphi_T^1 > \lambda_2\varphi_T^2\big)\\
&\qquad= \liminf_{T\to\infty}\frac{1}{T}\log P_T^2\Big(\frac{W_T^2}{T}+(\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2\,ds \;>\; d+\eta_T\Big)\\
&\qquad\geq \liminf_{T\to\infty}\frac{1}{T}\log P_T^2\Big((\alpha_1+\alpha_2)\frac{1}{T}\int_0^T W_s^2\,ds \;>\; d+\eta_T\Big)\\
&\qquad\geq -\frac{\alpha_2^2(\alpha_1+\alpha_2)}{2d}\Big(\frac{d}{\alpha_1+\alpha_2}-\frac{d}{2\alpha_2}\Big)^2\\
&\qquad= -\frac{d(\alpha_2-\alpha_1)^2}{8(\alpha_1+\alpha_2)}.
\end{aligned}\tag{7.8}$$

Dies ist eine Abschätzung für die untere Rate der Wahrscheinlichkeit, irrtümlich die Konstante α_1 anzunehmen, obwohl man einen Prozeß mit Konstante $\alpha_2 > \alpha_1$ beobachtete. Da die Menge $\{\,\nu\in\mathcal{M}_1(\mathbb{R}^d)\mid\int x^2\,d\nu > C\,\}$ bezüglich der schwachen Topologie in $\mathcal{M}_1(\mathbb{R}^d)$ dicht liegt, erhält man aus Satz 5.3 für diesen Fall keine sinnvolle Abschätzung für die obere Rate dieser Wahrscheinlichkeit.

Die Formeln (7.7) und (7.8) schätzen die Summanden aus Gleichung (7.3) in unterschiedliche Richtungen ab. Daher kann man hier nicht, wie im dritten Kapitel, eine Aussage über das Bayesrisiko ableiten. Da aber die Grenzen in beiden Abschätzungen übereinstimmen, ist

$$\lim_{T\to\infty}\frac{1}{T}\log B(\lambda,T) \;=\; -\frac{d(\alpha_2-\alpha_1)^2}{8(\alpha_1+\alpha_2)}$$

eine naheliegende Vermutung.

Nachwort

Diese Arbeit beschäftigt sich mit dem asymptotischen Verhalten von Diffusionsprozessen, und dabei insbesondere mit der Bestimmung der Konvergenzrate des Bayesrisikos. Für eine spezielle Klasse von Prozessen, nämlich solcher, deren Drift konstant ist, wird das Problem hier gelöst.

Der darauf folgende Teil des Textes stellt Hilfsmittel zur Verfügung, um weitere Klassen von Prozessen zu untersuchen. Speziell beschäftigt er sich mit reversiblen Prozessen. Das abschließende Kapitel löst das Problem der Bestimmung der Konvergenzrate des Bayesrisikos für einen speziellen Prozeß, den Ornstein-Uhlenbeck-Prozeß, teilweise. Dazu wird hier ein Satz über große Abweichungen der empirischen Verteilung (bezüglich der schwachen Topologie auf der Menge aller Wahrscheinlichkeitsmaße) verwendet.

Um weitergehende Ergebnisse zu erzielen, könnte man versuchen, ein Prinzip der großen Abweichungen bezüglich einer stärkeren Topologie zu verwenden. Damit könnte man eine Abschätzung für die gesuchte Konvergenzrate gewinnen. Um den genauen Wert der Rate zu finden, könnte man die gemeinsame Verteilung der Terme X_t^2/t und $\frac{1}{t}\int_0^t X_s^2\,ds$, wie sie in der Darstellung des Bayesrisikos im siebten Kapitel auftauchen, betrachten.

Eine naheliegende Verallgemeinerung dieser Problemstellung besteht darin, beliebige reversible Diffusionsprozesse zu betrachten.

Literatur

Folgende Liste faßt die Literatur, die ich zur Erstellung dieser Arbeit verwendet habe, zusammen. Die Hauptquellen waren das Buch von v. Weizsäcker und Winkler für das erste Kapitel, das Buch von Varadhan für das zweite Kapitel und der Artikel von Mitter für das sechste Kapitel. Die Liste ist alphabetisch nach Name des (ersten) Autors sortiert.

Heinz Bauer, *Wahrscheinlichkeitstheorie.* 4. Auflage. Walter de Gruyter; Berlin, 1991.

Jean-Dominique Deuschel, Daniel W. Stroock, *Large Deviations.* Academic Press, 1989.

Richard M. Dudley, *Real Analysis and Probability.* Wadsworth; Belmont (California), 1989.

Jerome A. Goldstein, *Semigroups of Linear Operators and Applications.* Oxford University Press, 1985.

Ioannis Karatzas, Steven E. Shreve, *Brownian Motion and Stochastic Calculus.* Second Edition. Springer; New York, 1991.

Robert Shevilevich Liptser, Al'bert Nikolaevich Shiryayev, *Statistics of Random Processes.* Zwei Bände. Springer, 1977.

S. K. Mitter, *Estimation Theory and Statistical Physics.* In: Itô Hida (Hrsg.), *Stochastic Processes and their Applications, Proceedings of the International Conference held in Nagoya, July 2–6, 1985.* Lecture Notes in Mathematics, Vol. 1203. Springer; Berlin, Heidelberg, New York, 1986.

Masao Nagasawa, *Schrödinger Equations and Diffusion Theory.* Birkhäuser; Basel, Boston, Berlin, 1993.

Michael Reed, Barry Simon, *Methods of Modern Mathematical Physics.* Vier Bände. Academic Press; New York, 1972 (Bd. I).

Friedrich Riesz, Béla Sz.-Nagy, *Vorlesungen über Funktionalanalysis.* Dritte Auflage, VEB deutscher Verlag der Wissenschaften; Berlin, 1973.

S. R. S. Varadhan, *Lectures on Diffusion Problems and Partial Differential Equations.* Springer, 1980.

Heinrich v. Weizsäcker, Gerhard Winkler, *Stochastic Integrals.* Vieweg, 1990.

Helmut Werner, Herbert Arndt, *Gewöhnliche Differentialgleichungen. Eine Einführung in Theorie und Praxis.* Springer, 1986.

.

Index

Verweise auf Definitionen sind durch fettgedruckte Seitennummern gekennzeichnet.