Kapitel IX

LAN, LAMN und LAQ in statistischen Modellen für Diffusionsprozesse

Reinhard Höpfner

Vorlesung Mathematische Statistik im Winter 2004/2005 und im Sommer 2005

Institut für Mathematik, Fachbereich 08, Johannes Gutenberg Universität Mainz

06.10.05

Übersicht zu Kapitel IX :

A. Statistische Modelle für Diffusionsprozesse

B. Ein Beispiel für lokale Asymptotik: der Ornstein-Uhlenbeck-Prozeß mit unbekanntem Parameter unter langer Beobachtungsdauer

Anhang A9 (Harris-Rekurrenz)

Kapitel IX erfordert eine gewisse Vertrautheit mit Grundlagen über stochastische Prozesse in stetiger Zeit, Martingale und Semimartingale, und mit stochastischen Differentialgleichungen. Von Lesern, denen diese Begriffe nicht vertraut sind, sollte es übersprungen werden; Beweise werden im Rahmen dieses Textes nicht gegeben.

Ein schön geschriebene Einführung in die Theorie der stochastischen Prozesse bietet das Buch von Métivier (1981); eine schnelle Übersicht über einige wichtige Grundtatsachen findet man im Anhang des Buches von Brémaud (1981); eine gute Referenz zum Einarbeiten insbesondere in stochastische Differentialgleichungen ist Karatzas und Shreve (1991). Die tiefergehenden Resultate – insbesondere für lokale Absolutstetigkeit und Likelihoodratioprozesse – findet man in den Büchern von Liptser und Shiryaev (1981, 2001) sowie Jacod und Shiryaev (1987, 2003). Die detaillierteste und sorgfältigste Darstellung der Theorie der stochastischen Prozesse bieten die Bände von Dellacherie und Meyer (insbesondere: Ch. I–IV: 1975, Ch. V–VIII: 1980); eine ähnlich detaillierte und sorgfältige Abhandlung über stochastische Differentialgleichungen findet man im Buch von Ikeda und Watanabe (1989). Einen reichen Überblick über Schätzverfahren in eindimensionalen ergodischen Diffusionen und ihre Effizienzeigenschaften gibt Kutoyants (2003).

A. Statistische Modelle für Diffusionsprozesse

Wir erläutern einige der wesentlichen Fakten über statistische Modelle für Diffusionsprozesse. Die folgenden Notationen werden durch das ganze Kapitel hindurch benutzt werden.

9.1 Bezeichnungen und Voraussetzungen: a) Wir betrachten wie in 5.11 oder in 6.9 den kanonischen Pfadraum $(C, \mathcal{C})$ für $\mathbb{R}^d$-wertige stochastische Prozesse mit stetigen Pfaden. Auf $(C, \mathcal{C})$ sei $X = (X_t)_{t\geq 0}$ der kanonische Prozeß und $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ die von X erzeugte (rechtsstetige) Filtration, d.h. $\mathcal{G}_t := \bigcap_{r>t} \sigma\left(X_t : 0 \leq t \leq r\right)$ ist die σ-Algebra aller Ereignisse bis zur Zeit t.
b) Wir betrachten zwei stochastische Differentialgleichungen (SDE) in $\mathbb{R}^d$

$$\text{(I)} \qquad d\xi_t \;=\; b(t,\xi_t)\,dt \;+\; \sigma(t,\xi_t)\,dW_t\,, \quad t \geq 0\,, \quad \xi_0 \equiv x_0$$

$$\text{(II)} \qquad d\xi'_t \;=\; b'(t,\xi'_t)\,dt \;+\; \sigma(t,\xi'_t)\,dW'_t\,, \quad t \geq 0\,, \quad \xi_0 \equiv x_0$$

mit treibenden m-dimensionalen Brownschen Bewegungen W, W', mit Diffusionskoeffizienten

$$\sigma \;:\; [0,\infty)\times\mathbb{R}^d \;\longrightarrow\; \mathbb{R}^{d\times m}$$

(beide Gleichungen (I) und (II) haben dasselbe $\sigma(\cdot,\cdot)$!) und mit Driftkoeffizienten

$$b\,,\,b' \;:\quad [0,\infty)\times\mathbb{R}^d \;\longrightarrow\; \mathbb{R}^d\;,$$

für die wir Lipschitzstetigkeit

$$|b(t,x)-b(t,y)| + |b'(t,x)-b'(t,y)| + \|\sigma(t,x)-\sigma(t,y)\| \;\le\; K\,|x-y|$$

(für $t\ge 0$ und $x,y\in\mathbb{R}^d$) und eine lineare Wachstumsbedingung

$$|b(t,x)|^2 + |b'(t,x)|^2 + \|\sigma(t,x)\|^2 \;\le\; K\,(1+|x|^2)$$

mit globalen Konstanten K voraussetzen. Wir schreiben

$$c(t,x) \;:=\; \sigma(t,x)\,\sigma^\top(t,x)\;.$$

c) Unter den Voraussetzungen aus b) gibt es für die Gleichungen (I) und (II) starke Lösungen (siehe Karatzas und Shreve, 1991, Ch. 5.2.B) jeweils auf dem Wahrscheinlichkeitsraum, der die treibenden Brownschen Bewegungen W bzw. W' trägt. Uns interessieren nur die *Verteilungen* dieser Lösungen, d.h. die Bildmaße

$$P \;:=\; \mathcal{L}(\,\xi\,) \quad \text{und} \quad P' \;:=\; \mathcal{L}(\,\xi'\,) \quad \text{auf dem kanonischen Pfadraum} \quad (C,\mathcal{C},\mathbb{G})\;.$$

Wir setzen stets voraus, daß in beiden Gleichungen (I) und (II) derselbe Startwert x_0 vorliegt.

Das folgende Ergebnis beweist man mit den Argumenten, die zu Liptser und Shiryaev (2001, Thm. 7.7 und Thm 7.18) führen. Eine Variante unter etwas anders gearteten Voraussetzungen findet man bei Kutoyants (2003, S.34, S.27).

9.2 Fakt: Es gebe zu b, b' aus 9.1 eine meßbare Funktion $\gamma:[0,\infty)\times\mathbb{R}^d\longrightarrow\mathbb{R}^d$, so daß die folgenden Bedingungen (+) und (++) erfüllt sind:

$$(+)\qquad b'(t,x) \;=\; b(t,x) \;+\; c(t,x)\,\gamma(t,x)\;,\quad t\ge 0\;,\; x\in\mathbb{R}^d\;,$$

$$(++)\qquad \int_0^t \left(\gamma^\top c\,\gamma\right)(s,X_s)\,ds \;<\; \infty \quad \text{für jedes } t<\infty,\;\; P\text{- und } P'\text{-fast sicher.}$$

Dann sind die Wahrscheinlichkeitsmaße P und P' *lokal äquivalent auf* $(C,\mathcal{C},\mathbb{G})$, d.h. es gilt

$$P^{\,t} \;\sim\; P'^{\,t} \quad \text{für jedes }\; 0\le t<\infty \quad (\text{Schreibweise: } P \overset{\text{loc}}{\sim} P'\;)$$

wobei wie in 5.11 Q^t die Restriktion eines Wahrscheinlichkeitsmaßes Q auf $(C,\mathcal{C})$ auf die Sub-σ-Algebra $\mathcal{G}_t$ bezeichnet. Der Dichteprozeß von P' zu P relativ zu $\mathbb{G}$ – beachte die Bemerkungen in 5.11 – ist das $(P,\mathbb{G})$-Martingal

$$L = (L_t)_{t\geq 0}\ , \quad L_t \;=\; \exp\left\{ \int_0^t \gamma^\top(s,X_s)\, dm_s^P \;-\; \frac{1}{2}\int_0^t \left(\gamma^\top c\, \gamma\right)(s,X_s)\, ds \right\}$$

wobei m^P den Martingalteil des kanonischen Prozesses X auf $(C,\mathcal{C},\mathbb{G})$ unter P bezeichnet

$$m^P \;=\; (m_t^P)_{t\geq 0}\ , \quad m_t^P \;=\; X_t - X_0 - \int_0^t b(s,X_s)\, ds$$

und das stochastische Integral

$$\int \gamma^\top(s,X_s)\, dm_s^P \;=:\; M$$

unter P gebildet ist. M ist ein lokal quadratintegrables lokales P-Martingal; P-fast alle Pfade sind stetig. Mit der üblichen Schreibweise $\langle M\rangle$ für den Kompensator von M^2 hat L die Struktur

$$L_t \;=\; \exp\left\{ M_t \;-\; \frac{1}{2}\langle M\rangle_t \right\}\ , \quad t\geq 0\ ;$$

damit ist der Dichteprozeß L das *Exponential von* M im Sinne einer Lösung der Gleichung

$$L_t \;=\; 1 \;+\; \int_0^t L_t\, dM_t\ , \quad t\geq 0\ .$$

Eine übliche Schreibweise für das Exponential von M ist $\mathcal{E}(M)$.

9.3 Spezialfall: Sei für alle $t\geq 0$ und $x\in \mathbb{R}^d$ die $d\times d$-Matrix $c(t,x)$ invertierbar. Ausgehend vom Driftkoeffizienten $b(\cdot,\cdot)$ in SDE (I) definieren wir ein statistisches Modell

$$\{\, P_h : h\in \mathbb{R}^d \,\} \quad \text{auf} \quad (C,\mathcal{C},\mathbb{G})\ , \quad \text{mit} \quad P_0 \;:=\; P\ ,$$

indem wir in SDE (II) anstelle von $b'(\cdot,\cdot)$ betrachten

$$b_h(t,x) \;:=\; b(t,x) \;+\; c(t,x)\, h\ , \quad h\in \mathbb{R}^d\ .$$

Dann sind alle Voraussetzungen von 9.2 (mit konstantem $\gamma(\cdot,\cdot)\equiv h$) erfüllt, es gilt

$$P_h \overset{\text{loc}}{\sim} P_0 \quad \text{relativ zu} \quad \mathbb{G}\ , \quad \text{mit} \quad P_0 \;=\; P\ ,$$

und der Dichteprozeß von P_h zu P_0 relativ zu $\mathbb{G}$ schreibt sich als

$$L_t^{h/0} \;=\; \exp\left\{ h^\top m_t^0 - \frac{1}{2} h^\top \langle m^0\rangle_t\, h \right\} \;=\; \mathcal{E}_t\left(h^\top m^0 \right)$$

mit Schreibweise m^0 für den Martingalteil des kanonischen Prozesses X unter $P_0 = P$. Wegen Stetigkeit und Invertierbarkeit von $c(t,x)$ für alle t,x ist die zufällige $d{\times}d$-Matrix

$$\langle m^0 \rangle_t \;=\; \int_0^t c(s,X_s)\,ds$$

invertierbar auf ganz C, für jedes $0<t<\infty$, folglich ist für festes $0<t<\infty$ das Modell

$$(\circ) \qquad \{\, P_h^t : \; h \in I\!R^d \,\} \;=\; \mathcal{E}(S,J) \;: \quad S \;:=\; m_t^0 \;, \quad J \;:=\; \langle m^0 \rangle_t$$

ein quadratisches statistisches Modell im Sinne von 6.1. Die Aussage bleibt gültig, wenn die deterministische Zeit $0<t<\infty$ in $(\circ)$ durch eine $I\!\!G$-Stopzeit τ mit der Eigenschaft $0<\tau<\infty$ auf C ersetzt wird (vgl. 5.11).

Beachte, daß auf diese Weise um die Verteilung der Lösung P jeder SDE (I), die den Voraussetzungen aus 9.1 genügt, in kanonischer Weise ein (im 'filtrierten' Sinne) quadratisches statistisches Modell gelegt werden kann.
Auf genau diese Weise war in 5.11 a), ausgehend von der durch $\mathcal{J}$ skalierten Brownschen Bewegung ($P = \mathcal{L}(\mathcal{J}^{1/2}B)$ als Spezialfall $b \equiv 0$ und $\sigma \equiv \mathcal{J}^{1/2}$ von SDE (I)) das Gauß-Shift-Modell 5.12 – 'Brownsche Bewegung mit unbekannter Drift, skaliert durch $\mathcal{J}$' – erklärt worden.

Selbstverständlich ist es auf vielfältige Weise möglich, durch ein Wahrscheinlichkeitsmaß P auf $(C,\mathcal{C},I\!\!G)$ parametrische Pfade (in einem 'filtrierten' Sinn) zu legen. Als Beispiel illustrieren wir dies in Dimension $d=1$ am Ornstein-Uhlenbeck Modell mit Startwert 0, das wir als einparametrigen Pfad durch das Wienermaß in $(C,\mathcal{C},I\!\!G)$ auffassen.

9.4 Beispiel: Sei $d=1$, sei $\sigma(\cdot,\cdot) \equiv 1$ unabhängig von $t \ge 0$ und $x \in I\!R$, sei W eine eindimensionale Standard-Brownsche Bewegung. Für $h \in I\!R$ betrachte die Ornstein-Uhlenbeck-SDE

$$d\xi_t \;=\; h\,\xi_t\,dt \;+\; dW_t \;, \quad t \ge 0 \;, \quad \xi_0 \equiv 0$$

mit Startwert 0; die explizite Lösung ist

$$\xi_t \;=\; e^{ht} \int_0^t e^{-hs}\,dW_s \;, \quad t \ge 0 \;.$$

Schreibt man $P_h = \mathcal{L}(\xi)$ für die Verteilung der Lösung auf dem Pfadraum $(C,\mathcal{C},I\!\!G)$, in Dimension $d=1$, so ist P_0 das Wienermaß und es gilt für alle $h \in I\!R$ (mit $\gamma(t,x) := h\,x$ in 9.2)

$$P_h \overset{\mathrm{loc}}{\sim} P_0 \;, \quad L_t^{h/0} \;=\; \exp\left\{ h \int_0^t X_s\,dm_s^0 \;-\; \frac{1}{2}\,h^2 \int_0^t X_s^2\,ds \right\} \;.$$

Unter dem Wienermaß P_0 ist der kanonische Prozeß X auf $(C,\mathcal{C},\mathbb{G})$ als eindimensionale Standard-Brownsche Bewegung selbst ein P_0-Martingal. Also gilt $m^0 = X$ (bis auf P_0-Ununterscheidbarkeit), und die Ito-Formel zeigt

$$\int_0^t X_s \, dX_s \;=\; \frac{1}{2}\left(X_t^2 - t\right) \;,\quad t \geq 0 \;,\quad P_0\text{-fast sicher} \;.$$

Folglich schreibt sich der Dichteprozeß $L^{h/0}$ für beliebiges $h \in I\!R^d$ explizit als

$$L_t^{h/0} \;=\; \exp\left\{ h \, \frac{1}{2}\left(X_t^2 - t\right) \;-\; \frac{1}{2} h^2 \int_0^t X_s^2 \, ds \right\} \;,\quad t \geq 0 \;.$$

Bei deterministischer Beobachtungsdauer $0 < t < \infty$ ist das statistische Modell

$$(\circ) \qquad \{ P_h^t : \, h \in I\!R \} \;=\; \mathcal{E}(S,J) \;:\quad S := \frac{1}{2}\left(X_t^2 - t\right) \;,\quad J := \int_0^t X_s^2 \, ds$$

ein quadratisches, aber kein gemischt-normales statistisches Modell:
wegen $X_t^2 \geq 0$ ist die Verteilung des Score an der Stelle 0 in $(\circ)$

$$\mathcal{L}\left(\frac{1}{2}\left(X_t^2 - t\right) \mid P_0 \right) \qquad \text{konzentriert auf die Halbachse } [-\tfrac{t}{2}, \infty)$$

nicht symmetrisch; unter gemischter Normalität nach 6.3 wären aber für P_0^J-fast alle j die bedingten Verteilungen $P_0^{S|J=j} = \mathcal{N}(0,j)$ symmetrisch, folglich auch $\mathcal{L}(S \mid P_0)$ als Mischung über symmetrische Verteilungen. Genauso hätte man auch mit der Information J argumentieren können: für den kanonischen Prozeß X hängt

$$\mathcal{L}\,(\,J \mid P_h\,) \;=\; \mathcal{L}\left(\int_0^t X_s^2 \, ds \mid P_h \right)$$

ganz offenkundig vom Wert des Parameters $h \in I\!R^d$ ab. □

Ein Schlüsselbegriff, der unter schwachen Zusatzvoraussetzungen zur lokalasymptotische Normalität statistischer Modelle führt, ist *Ergodizität*; dieser Begriff wird in einer Azéma, Duflo und Revuz (1969) folgenden Darstellung in A9.2 im Anhang A9 zu diesem Kapitel präzisiert.

B. Ein Beispiel für lokale Asymptotik: Ornstein-Uhlenbeck-Prozeß mit unbekanntem Parameter unter langer Beobachtungsdauer

Wir beginnen mit einem Überblick über die asymptotischen Eigenschaften des Ornstein-Uhlenbeck-Prozesses in Abhängigkeit vom Wert des Driftparameters.

9.5 Bemerkungen: a) Sei $d = 1$. Die Ornstein-Uhlenbeck SDE mit Parameter $\vartheta \in I\!R$

$$(*) \qquad d\xi_t \;=\; \vartheta\, \xi_t\, dt \;+\; dW_t\;, \quad t \geq 0\;, \quad \xi_0 \equiv x_0$$

und festgelegtem Startwert $x_0 \in I\!R$ besitzt die explizite Lösung

$$\xi_t \;=\; e^{\vartheta\, t} \left(x_0 \;+\; \int_0^t e^{-\vartheta\, s}\, dW_s \right) \;, \quad t \geq 0\;.$$

Indem man z.B. die Funktion $s \to e^{-\vartheta\, s}$ durch Stufenfunktionen approximiert, sieht man

$$\mathcal{L}\left(e^{\vartheta\, t} \int_0^t e^{-\vartheta\, s}\, dW_s \right) \;=\; \left\{ \begin{array}{cl} \mathcal{N}\left(0\,,\, \frac{1}{2\,\vartheta}\left(e^{2\vartheta\, t} - 1\right)\right) & \text{falls } \vartheta \neq 0 \\ \mathcal{N}\left(0\,,\, t\right) & \text{falls } \vartheta = 0 \end{array} \right\}$$

und hat damit auch die Halbgruppe der Übergangswahrscheinlichkeiten

$$(+) \qquad P_t(x, dy) \;=\; \left\{ \begin{array}{cl} \mathcal{N}\left(e^{\vartheta\, t}\, x_0\,,\, \frac{1}{2\,\vartheta}\left(e^{2\vartheta\, t} - 1\right)\right) & \text{falls } \vartheta \neq 0 \\ \mathcal{N}\left(x_0\,,\, t\right) & \text{falls } \vartheta = 0 \end{array} \right\}\,.$$

b) *Nullrekurrenz im Fall* $\vartheta = 0$*:* Unter $\vartheta = 0$ ist ξ als Brownsche Bewegung mit Start in x_0 nullrekurrent – siehe A9.5 – mit invariantem Maß $\lambda\!\!\lambda$.

c) *Ergodizität im Fall* $\vartheta < 0$*:* Unter $\vartheta < 0$ wirkt die Drift $b_\vartheta(x) = \vartheta x$ der SDE $(*)$ als 'hinreichend starke' Rückstellkraft, die Trajektorien in Richtung auf den Punkt $0 \in I\!R$ zurücktreibt. Diese anschauliche Formulierung wird in der folgenden Weise rigoros gemacht:
Man verifiziert, daß die Bedingung (A9.9) in A9.7 erfüllt ist (mit $s(y) = e^{|\vartheta| y^2}$, $y \in I\!R$). Damit ist der Prozeß X unter $\vartheta < 0$ positiv rekurrent im Sinn von Harris, mit invariantem Maß

$$m^\vartheta \;:=\; \mathcal{N}\left(0, \frac{1}{2|\vartheta|}\right)$$

(benutze (A9.8) oder explizit die Gestalt der Halbgruppe). Im langen Lauf wird X Mengen von positivem invariantem Maß immer wieder besuchen, vgl. A9.1–A9.4, für jede Wahl eines Startpunktes. Nach dem Ratio Limit Theorem (A9.3) gilt für $f \in L^1(m^\vartheta)$

$$(9.6) \qquad \lim_{t\to\infty} \frac{1}{t} \int_0^t f(\xi_s)\, ds \;=\; m^\vartheta(f) \quad \text{fast sicher für } t \to \infty\;.$$

Damit ist ein Grenzwertsatz für eine große Klasse additiver Funktionale von ξ zur Hand.

d) *Transienz im Fall* $\vartheta > 0$*:* Die Drift der SDE $(*)$ wirkt für $\vartheta > 0$ zentrifugal und treibt Trajektorien von ξ schließlich exponentiell schnell gegen $+\infty$ oder $-\infty$:

Auf $(\Omega, \mathcal{A}, I\!F, P)$, wobei $I\!F$ die von der Brownschen Bewegung W in $(*)$ erzeugte Filtration ist, betrachte das $(P, I\!F)$-Martingal Y:

$$Y = (Y_t)_{t \geq 0} \ , \quad Y_t := e^{-\vartheta t} \xi_t \ = \ x_0 + \int_0^t e^{-\vartheta s} \, dW_s \ , \ t \geq 0 \ .$$

Wegen $E\left((Y_t - x_0)^2\right) = \int_0^t e^{-2\vartheta s} \, ds$ hat man

$$\sup_{t \geq 0} E\left(Y_t^2\right) \ < \ \infty \ ;$$

folglich konvergiert Y_t für $t \to \infty$ P-fast sicher und in $L^2(\Omega, \mathcal{A}, P)$ gegen eine Limesvariable $Y_\infty \in L^2(\Omega, \mathcal{F}_\infty, P)$. Die explizite Gestalt von Y_t liefert

$$(9.6') \qquad Y_\infty \ = \ x_0 + \int_0^\infty e^{-\vartheta s} \, dW_s \quad \sim \quad \mathcal{N}\left(x_0 \, , \, \frac{1}{2\vartheta} \right) \ , \quad \vartheta > 0 \ .$$

Aus der fast sicheren Konvergenz von Y_t gegen Y_∞ ergibt sich für die Trajektorien von ξ

$$(9.7) \qquad \xi_t(\omega) \ \sim \ Y_\infty(\omega) \, e^{\vartheta t} \ , \quad t \to \infty \ , \quad \text{für } P\text{-fast alle } \omega \in \Omega \ .$$

Dies liefert für einige (wenige!) additive Funktionale von ξ einen Grenzwertsatz, insbesondere

$$(9.8) \qquad \int_0^t \xi_s^2(\omega) \, ds \ \sim \ Y_\infty^2(\omega) \, \frac{1}{2\vartheta} \, e^{2\vartheta t} \ , \quad t \to \infty \ , \quad \text{für } P\text{-fast alle } \omega \in \Omega \ .$$

Wegen der Asymptotik (9.7) der Pfade von ξ gilt für jedes Kompaktum K in $I\!R$

$$\int_0^\infty 1_K(\xi_s) \, ds \ < \ \infty \quad P\text{-fast sicher} \ ;$$

damit ist der Prozeß ξ unter $\vartheta > 0$ transient. □

9.9 Das statistische Experiment: Setze $\Theta := I\!R$, schreibe P_ϑ für die Verteilung der Lösung der SDE $(*)$ unter $\vartheta \in \Theta$, und X für den kanonischen Prozeß auf $(C, \mathcal{C}, I\!\!G)$, dem kanonischen Pfadraum für reellwertige Prozesse mit stetigen Pfaden. Wir fixieren den Startwert x_0 in $(*)$ unabhängig von $\vartheta \in \Theta$.

a) Nach 9.4 gilt für beliebiges $\vartheta \in \Theta$

$$P_\vartheta \overset{\text{loc}}{\sim} P_0 \ , \quad L_t^{\vartheta/0} \ = \ \exp\left\{ \vartheta \int_0^t X_s \, dm_s^0 \ - \ \frac{1}{2} \vartheta^2 \int_0^t X_s^2 \, ds \right\} \ .$$

Halten wir eine Festlegung Y des stochastischen Integrals $\int_0^t X_s \, dm_s^0$ unter P_0 fest, so hat der Maximum-Likelihood-Schätzer für den unbekannten Parameter die Form

$$(\circ) \qquad \widehat{\vartheta}_t \ = \ \frac{Y_t}{\int_0^t X_s^2 \, ds} \ , \quad t \geq 0 \ .$$

Y ist aber auch eine Festlegung des stochastischen Integrals $\int_0^t X_s\, dm_s^\vartheta + \vartheta \int X_s^2\, ds$ unter P_ϑ. Folglich liefert

$$\exp\left\{ (\xi - \vartheta)\, Y_t \;-\; \frac{1}{2}\,(\xi^2 - \vartheta^2) \int_0^t X_s^2\, ds \right\} \;,\quad t \geq 0$$

zugleich eine Festlegung des Quotienten $L^{\xi/0}/L^{\vartheta/0}$ unter P_0 und eine Festlegung des Dichteprozesses $L^{\xi/\vartheta}$ unter P_ϑ gemäß 9.2:

$$(+) \qquad L_t^{\xi/\vartheta} \;=\; \exp\left\{ (\xi - \vartheta) \int_0^t X_s\, dm_s^\vartheta \;-\; \frac{1}{2}\,(\xi - \vartheta)^2 \int_0^t X_s^2\, ds \right\} \;,\quad t \geq 0\;.$$

Mit demselben Argument erhält man aus ($\circ$) die Martingaldarstellung der Maximum-Likelihood-Schätzfehler unter P_ϑ

$$(++) \qquad \widehat{\vartheta}_t \;-\; \vartheta \;=\; \frac{\int_0^t X_s\, dm_s^\vartheta}{\int_0^t X_s^2\, ds} \;,\quad t \geq 0\;.$$

b) Für beliebiges $\vartheta \in \Theta$ und festes $0 < t < \infty$ entsteht in (+) ein quadratisches Modell an der Stelle ϑ, wobei Score und Information an der Stelle ϑ in stochastische Prozesse eingebettet sind: wir nennen

$$\left(\int_0^t X_s\, dm_s^\vartheta \right)_{t\geq 0} \quad \text{und} \quad \left(\int_0^t X_s^2\, ds \right)_{t\geq 0} \;=\; \left\langle \int X\, dm^\vartheta \right\rangle$$

Score-Martingal und *Informationsprozeß* unter P_ϑ. Beachte, daß im hier betrachteten Modell der Informationsprozeß nicht vom Parameter abhängt.

c) Für zeitkontinuierliche Beobachtung des kanonischen Prozesses X unter unbekanntem Parameter über ein langes Zeitintervall $[0, n]$ betrachten wir nun – bei geeigneter (und noch zu präzisierender) Wahl eines lokalen Maßstabs $\delta_n(\vartheta)$ an der Stelle $\vartheta \in \Theta$ – Folgen lokaler Modelle

$$(9.9') \qquad \left(C\,,\, \mathcal{G}_n\,,\, \left\{ P^n_{\vartheta+\delta_n(\vartheta)\,h} :\; h \in I\!R \right\} \right) \;,\quad n \geq 1$$

an der Stelle ϑ. In diesen schreiben sich wegen (+) die log-Likelihoods in Form

$$(\diamond) \qquad \Lambda^{h/0}_{\vartheta,n}(t) \;=\; \log\, L_t^{(\vartheta+\delta_n(\vartheta)h)/\vartheta} \;=\; h\left(\delta_n(\vartheta) \int_0^n X_s\, dm_s^\vartheta \right) - \frac{1}{2} h^2 \left(\delta_n^2(\vartheta) \int_0^n X_s^2\, ds \right)$$

und wegen (++) die reskalierten Maximum-Likelihood-Schätzfehler als

$$(\diamond\diamond) \qquad \delta_n^{-1}(\vartheta) \left(\widehat{\vartheta}_n \;-\; \vartheta \right) \;=\; \left(\delta_n^2(\vartheta) \int_0^n X_s^2\, ds \right)^{-1} \left(\delta_n(\vartheta) \int_0^n X_s\, dm_s^\vartheta \right)$$

Folglich ist – bei Beobachtung des kanonischen Prozesses bis zur Zeit n – der 'richtige' lokale Maßstab an der Stelle ϑ nichts anderes als die 'richtige' Folge von Normierungskonstanten für das

Paar aus Score-Martingal und Informationsprozeß an der Stelle ϑ: in 7.1 braucht man schwache Konvergenz der Paare

$$\left(\delta_n(\vartheta) \int_0^n X_s \, dm_s^\vartheta \ , \ \delta_n^2(\vartheta) \int_0^n X_s^2 \, ds \right) \quad \text{unter } P_\vartheta, \text{ für } n \to \infty$$

gegen ein Paar $(S(\vartheta), J(\vartheta))$, das wie in 6.1 als Score und Information in einem Limesmodell $\mathcal{E}(S(\vartheta), J(\vartheta))$ interpretierbar ist.

9.10 LAN im ergodischen Fall $\vartheta < 0$: Unter positiver Rekurrenz hat man nach (9.6) mit der invarianten Wahrscheinlichkeit m^ϑ aus 9.5 c) den Grenzwertsatz

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t X_s^2 \, ds \ = \ \int x^2 \, m^\vartheta(dx) \ = \ \frac{1}{2|\vartheta|} \ =: \ \mathcal{J} \quad P_\vartheta\text{-fast sicher} .$$

Im lokalen Modell an der Stelle ϑ muß folglich der lokale Maßstab $\delta_n(\vartheta)$ das (aus der Welt klassischer iid-Modelle mit Fisher-Information) gewohnte

$$(9.10') \qquad \delta_n(\vartheta) \ = \ n^{-1/2} \quad \text{im Fall } \vartheta < 0$$

sein; entsprechend skalieren wir das Score-Martingal an der Stelle ϑ in Raum und Zeit

$$M_\vartheta^n \ = \ (M_\vartheta^n(t))_{t\geq 0} \ , \quad M_\vartheta^n(t) \ := \ n^{-1/2} \int_0^{t\,n} X_s \, dm_s^\vartheta \ , \quad t \geq 0 \ , \quad n \geq 1 \ .$$

Dies liefert eine Familie M_ϑ^n stetiger P_ϑ-Martingale bezüglich $\mathbb{G}^n := (\mathcal{G}_{tn})_{t\geq 0}$, $n \geq 1$, deren vorhersehbare quadratische Variation die Bedingung

$$(*) \qquad \forall\, t \geq 0 \ : \quad \langle M_\vartheta^n \rangle_t \ = \ \frac{1}{n} \int_0^{t\,n} X_s^2 \, ds \ \longrightarrow \ t \, \frac{1}{2|\vartheta|} \ = \ t \, \mathcal{J} \quad P_\vartheta\text{-fast sicher}$$

erfüllt. Wegen $(*)$ liefert der Martingalkonvergenzsatz (siehe Jacod und Shiryaev 1987, Cor. VIII.3.24) schwache Konvergenz im Skorohodraum D aller càdlàg-Funktionen $[0, \infty) \to \mathbb{R}$ gegen die mit $\mathcal{J}$ skalierte Standard-Brownschen Bewegung B:

$$M_\vartheta^n \ \longrightarrow \ \mathcal{J}^{1/2} B \quad (\text{schwache Konvergenz in } D \text{ unter } P_\vartheta, \text{ für } n \to \infty) \ .$$

Nun sind aber Auswertungsabbildungen $D \ni \alpha \to \alpha(t) \in \mathbb{R}$ stetig an allen Stellen $\alpha \in C \subset D$ (siehe Jacod und Shiryaev 1987, VI.2.1, Aussage b5), und C ist eine Menge von vollem Maß unter $\mathcal{L}\left(\mathcal{J}^{1/2} B\right)$; der Satz über stetige Abbildungen zeigt also

$$M_\vartheta^n(1) \ \longrightarrow \ \mathcal{J}^{1/2} B_1 \quad (\text{schwache Konvergenz in } \mathbb{R} \text{ unter } P_\vartheta, \text{ für } n \to \infty)$$

insbesondere für $t = 1$. Dies kann durch Hinzufügen einer deterministischen Komponente zu

$$(M_\vartheta^n(1) \, , \mathcal{J}) \;\longrightarrow\; \left(\mathcal{J}^{1/2} B_1 \, , \mathcal{J} \right) \quad \text{schwach in } I\!R^2 \text{ unter } P_\vartheta,\ n \to \infty$$

umgeschrieben werden, was nach nochmaliger Ausnutzung von $(*)$ zu

$$(\diamond \diamond \diamond) \qquad \left(M_\vartheta^n(1) \, , \langle M_\vartheta^n \rangle_1 \right) \;\longrightarrow\; \left(\mathcal{J}^{1/2} B_1 \, , \mathcal{J} \right) \quad (\text{schwach in } I\!R^2 \text{ unter } P_\vartheta, \text{ für } n \to \infty)$$

führt. Indem man nun $(\diamond)$–$(\diamond\diamond)$ aus 9.9 mit der Konvergenzaussage $(\diamond \diamond \diamond)$ kombiniert, sind für die lokalen Modelle (9.9')+(9.10') an Stellen $\vartheta < 0$ folgende Aussagen gezeigt:

- an jeder Stelle $\vartheta < 0$ gilt lokalasymptotische Normalität mit lokalem Maßstab $n^{-1/2}$ bei Beobachtung der Trajektorie des kanonischen Prozesses bis zur Zeit n, $n \to \infty$ (vgl. 7.1);
- das Limesexperiment an der Stelle $\vartheta < 0$ ist die durch $\mathcal{J} = \frac{1}{2|\vartheta|}$ skalierte Brownsche Bewegung mit unbekannter Drift, beobachtet über das Einheits-Zeitintervall $[0,1]$, wie in (5.13);
- die Maximum-Likelihood-Schätzfolge $(\widehat{\vartheta}_n)_n$ für den unbekannten Parameter ist regulär und effizient an der Stelle $\vartheta < 0$ (vgl. 7.12);
- die Maximum-Likelihood-Schätzfolge schöpft an der Stelle $\vartheta < 0$ die lokalasymptotische Minimaxschranke aus (vgl. 7.13). □

9.11 LAMN im transienten Fall $\vartheta > 0$: a) Für $\vartheta > 0$ hat man nach (9.8) und (9.6')

$$e^{-2\,\vartheta\, t} \int_0^t X_s^2 \, ds \;\longrightarrow\; Y_\infty^2 \, \frac{1}{2\,\vartheta} \quad P_\vartheta\text{-fast sicher für } t \to \infty \, , \quad Y_\infty = Y_\infty(\vartheta) \;\sim\; \mathcal{N}\left(x_0, \frac{1}{2\vartheta} \right) \, .$$

Folglich muß man für $\vartheta > 0$ den lokalen Maßstab im lokalen Modell (9.9') parameterabhängig als

$$(9.11') \qquad \delta_n(\vartheta) \;=\; e^{-\vartheta\, n} \quad \text{im Fall } \vartheta > 0$$

festlegen. Das Score-Martingal skaliert man in Raum und Zeit entsprechend durch

$$M_\vartheta^n \;=\; \left(M_\vartheta^n(t) \right)_{t \geq 0} \; , \quad M_\vartheta^n(t) \;:=\; e^{-\vartheta\, n} \int_0^{(n+\log(t))^+} X_s \, dm_s^\vartheta \; , \quad t \geq 0 \; , \quad n \geq 1$$

mit $f^+ = f \vee 0$, und erhält eine Familie M_ϑ^n stetiger P_ϑ-Martingale bezüglich

$$I\!\!G^n \;:=\; \left(\mathcal{G}_{(n+\log(t))^+} \right)_{t \geq 0} \; , \quad n \geq 1 \; ,$$

deren vorhersehbare quadratische Variation für jedes feste t und $n \to \infty$ die Bedingung

$$\langle M_\vartheta^n \rangle_t \;=\; e^{-2\,\vartheta\, n} \int_0^{(n+\log(t))^+} X_s^2 \, ds \;\longrightarrow\; Y_\infty^2(\vartheta) \, \frac{1}{2\,\vartheta} \, t^{2\,\vartheta} \quad P_\vartheta\text{-fast sicher}$$

erfüllt. Die Limesvariable $Y^2_\infty(\vartheta)$ folgt wegen (9.6') einer nichtzentralen Gammaverteilung, siehe z.B. Barra (1981, Ch. VII.1), und kann nach Abänderung auf einer P_ϑ-Nullmenge als strikt positiv aufgefaßt werden. Damit ergibt sich auf der rechten Seite in Abhängigkeit von $t \geq 0$ ein stochastischer Prozeß $\varphi_\vartheta = (\varphi_\vartheta(t))_{t\geq 0}$ mit den Eigenschaften

$$(*) \qquad \varphi_\vartheta(0) \equiv 0\ , \quad \varphi_\vartheta(t) > 0 \ \text{ für } \ t > 0\ , \quad \lim_{t\to\infty} \varphi_\vartheta(t) = +\infty$$

$$(**) \qquad \text{für } t \geq 0 \text{ fest } : \quad \langle M^n_\vartheta \rangle_t \ \longrightarrow \ \varphi_\vartheta(t) \quad P_\vartheta\text{-fast sicher für } n \to \infty\ .$$

Eine Erweiterung des Martingalkonvergenzsatzes (siehe Jacod und Shiryaev 1987, Thm. VIII.5.7, und S. 419) erlaubt es, aus $(**)$ die schwache Konvergenz der raum-zeit-skalierten Score-Martingale M^n_ϑ zu folgern: mit einer von φ_ϑ unabhängigen Standard-Brownschen Bewegung B gilt

$$M^n_\vartheta \ \longrightarrow \ B \circ \varphi_\vartheta \quad (\text{schwache Konvergenz in } D, \text{ unter } P_\vartheta, \text{ für } n \to \infty)\ .$$

Wegen der Stetigkeit aller hier betrachteten lokalen Martingale darf man in dieser Konvergenz die vorhersehbaren quadratischen Variationen mitnehmen (siehe Jacod und Shiryaev 1987, Thm. 6.6.1) und erhält schwache Konvergenz

$$(\, M^n_\vartheta\, ,\, \langle M^n_\vartheta \rangle) \ \longrightarrow \ (\, B \circ \varphi_\vartheta\, ,\, \varphi_\vartheta\,)$$

im Skorohod-Raum der càdlàg-Funktionen $[0,\infty) \to I\!R^2$, unter P_ϑ, für $n \to \infty$. Mit Stetigkeit der Auswertungsabbildungen auf der Teilmenge der stetigen Funktionen $[0,\infty) \to I\!R^2$ folgt

$$(\diamond\diamond\diamond) \quad \big(M^n_\vartheta(1)\, ,\, \langle M^n_\vartheta \rangle_1 \big) \ \longrightarrow \ (\, B(\varphi_\vartheta(1))\, ,\, \varphi_\vartheta(1)\,) \quad (\text{schwach in } I\!R^2, \text{ unter } P_\vartheta, \text{ für } n \to \infty)\ .$$

b) Die rechte Seite von $(\diamond\diamond\diamond)$ erzeugt ein gemischt normales Limesmodell:
Unabhängigkeit der Brownschen Bewegung B von φ_ϑ bedeutet

$$\mathcal{L}\,(\, B(\varphi_\vartheta(1))\, ,\, \varphi_\vartheta(1) \mid P_\vartheta\,) \ = \ \int \mathcal{L}\,(\varphi_\vartheta(1) \mid P_\vartheta)\,(dj) \ \big(\mathcal{N}(0,j) \otimes \epsilon_j\big)\ ,$$

also definiert man auf $\big(I\!R^2, \mathcal{B}(I\!R^2)\big)$ für $h \in I\!R$ Wahrscheinlichkeitsmaße

$$P'_h(\vartheta) \ := \ \mathcal{L}\,(\, B(\varphi_\vartheta(1)) + h\varphi_\vartheta(1)\, ,\, \varphi_\vartheta(1) \mid P_\vartheta\,) \ = \ \int \mathcal{L}\,(\varphi_\vartheta(1) \mid P_\vartheta)\,(dj) \ \big(\mathcal{N}(hj,j) \otimes \epsilon_j\big)$$

und hat damit wie in 6.2–6.3 ein Experiment

$$\big(I\!R^2, \mathcal{B}(I\!R^2), \{P'_h(\vartheta) : h \in I\!R\}\big)\ ,$$

in dem die Information unabhängig von $h \in I\!\!R$ wegen

$$\mathcal{L}\left(\varphi_1(\vartheta) \mid P_\vartheta\right) \;=\; \mathcal{L}\left(Y_\infty^2(\vartheta)\,\frac{1}{2\vartheta} \mid P_\vartheta\right) \;=\; \mathcal{L}\left(\frac{1}{2\vartheta}\left(x_0 + \int_0^\infty e^{-\vartheta s}\,dW_s\right)^2\right)$$

und (9.6') dezentral Gamma-verteilt ist.

c) Schritte a) und b) zusammen zeigen lokalasymptotisch gemischte Normalität an der Stelle $\vartheta > 0$. Indem wir ($\diamond$) und ($\diamond\diamond$) aus 9.9 mit ($\diamond \diamond \diamond$) aus a) kombinieren, erhalten wir für lokale Modelle (9.9')+(9.11') im 'transienten' Fall $\vartheta > 0$ die folgenden Aussagen:

- an jeder Stelle $\vartheta > 0$ gilt LAMN mit (parameterabhängigem) lokalem Maßstab $e^{-\vartheta\, n}$ für $n \to \infty$ (vgl. 7.1 und 7.2);
- die Maximum-Likelihood-Schätzfolge $(\widehat{\vartheta}_n)_n$ für den unbekannten Parameter ist regulär und effizient an der Stelle $\vartheta > 0$ (vgl. 7.12);
- die Maximum-Likelihood-Schätzfolge schöpft an der Stelle $\vartheta > 0$ die lokalasymptotische Minimaxschranke aus (vgl. 7.13).

Wir bemerken (ohne Details) noch folgendes: das in b) angegebene Limesexperiment ist nichts anderes als (6.14) im filtrierten Experiment 'Brownsche Bewegung mit unbekannter Drift unter unabhängiger Zeittransformation' aus 6.12. Also kann noch eine vierte griffige Formulierung angefügt werden:

- das Limesexperiment an der Stelle $\vartheta > 0$ ist Brownsche Bewegung mit unbekannter Drift, unabhängig zeittransformiert durch

$$t \;\longrightarrow\; Y_\infty^2(\vartheta)\,\frac{1}{2\vartheta}\,t^{2\vartheta}\;, \quad Y_\infty(\vartheta) \;=\; \lim_{t\to\infty} e^{-\vartheta\, t}\,X_t \quad \text{unter } P_\vartheta$$

und beobachtet über das Einheits-Zeitintervall $[0,1]$. □

9.12 LAQ an der Stelle $\vartheta = 0$: Im Fall $\vartheta = 0$ ist P_0 auf $(C, \mathcal{C})$ die Verteilung einer Brownschen Bewegung mit Start in x_0, und für das Paar aus Score und Information an der Stelle $\vartheta = 0$ gilt

$$(*) \qquad \mathcal{L}\left(\left(\frac{1}{n}\int_0^n X_s\,dm_s^0\,,\;\frac{1}{n^2}\int_0^n X_s^2\,ds\right) \mid P_0\right) \;\longrightarrow\; \mathcal{L}\left(\frac{1}{2}\left(B_1^2 - 1\right)\;,\;\int_0^1 B_s^2\,ds\right)\;,\; n \to \infty$$

mit einer Standard-Brownschen Bewegung B; dies sieht man so.

Mit $B = (B_t)_{t\geq 0}$ ist für jedes feste $t \in (0,\infty)$ auch $\left(\sqrt{t}\,B_{s/t}\right)_{s\geq 0}$ eine Standard-Brownsche Bewegung. Über die Ito-Formel erhält man Verteilungsgleichheit

$$\left(\int_0^t B_s\,dB_s\,,\;\int_0^t B_s\,ds\,,\;\int_0^t B_s^2\,ds\right) \quad = \quad \left(\frac{1}{2}\left(B_t^2 - t\right)\;,\;\int_0^t B_s\,ds\,,\;\int_0^t B_s^2\,ds\right)$$

$$\stackrel{d}{=} \quad \left(\frac{1}{2} \left([\sqrt{t}\, B_1]^2 - t \right) \, , \, \int_0^t \left(\sqrt{t}\, B_{\frac{s}{t}} \right) ds \, , \, \int_0^t \left(\sqrt{t}\, B_{\frac{s}{t}} \right)^2 ds \right)$$

$$\stackrel{d}{=} \quad \left(t\, \frac{1}{2} \left(B_1^2 - 1 \right) \, , \, t^{3/2} \int_0^1 B_s \, ds \, , \, t^2 \int_0^1 B_s^2 \, ds \right) .$$

In den Notationen aus 9.9 gilt damit für den kanonischen Prozeß X unter $P_0 = \mathcal{L}(x_0 + B)$

$$\frac{1}{n} \int_0^n X_s \, dm_s^0 \;=\; \frac{1}{n} \int_0^n (X - X_0)_s \, dm_s^0 \;+\; O_{P_0}\left(\frac{x_0}{\sqrt{n}} \right)$$

$$\frac{1}{n^2} \int_0^n X_s^2 \, ds \;=\; \frac{1}{n^2} \int_0^n (X - X_0)_s^2 \, ds \;+\; \frac{x_0^2}{n} \;+\; O_{P_0}\left(\frac{x_0}{\sqrt{n}} \right) .$$

woraus $(*)$ folgt. Im nullrekurrenten Fall $\vartheta = 0$ übernimmt also die Skalierungseigenschaft der Brownschen Bewegung die Rolle, die Grenzwertsätze (vgl. 9.10 a) oder 9.11 a)) für das Paar aus Score und Information

$$\left(\delta_n(\vartheta) \int_0^n X_s \, dm_s^\vartheta \, , \, \delta_n^2(\vartheta) \int_0^n X_s^2 \, ds \right) \, , \quad n \to \infty \, , \quad \vartheta = 0 \, ,$$

in den anderen Fällen spielten. Als richtige Wahl des lokalen Maßstabs ergibt sich wegen $(*)$

(9.12′) $$\delta_n(\vartheta) \;:=\; \frac{1}{n} \quad \text{falls } \vartheta = 0 \, ,$$

und als Limesmodell für die lokalen Modelle an der Stelle $\vartheta = 0$

$$\left(C \, , \, \mathcal{G}_n \, , \, \left\{ P_{0+h/n}^n : h \in I\!R \right\} \right) \, , \quad n \to \infty$$

findet man $\mathcal{E}(S, J)$ aus 9.4 mit $t = 1$ wieder:

$$\mathcal{E}(S, J) \, , \quad S \;=\; \frac{1}{2} \left(B_1^2 - 1 \right) \;=\; \int_0^1 B_s \, dB_s \, , \quad J \;=\; \int_0^1 B_s^2 \, ds \, .$$

Dieses war (vgl. 9.4) quadratisch, aber nicht gemischt normal. Als statistische Konsequenz folgt:

- an der Stelle $\vartheta = 0$ gilt LAQ mit lokalem Maßstab $\frac{1}{n}$;
- reskalierte Maximum-Likelihood-Schätzfehler in kleinen Umgebungen der Stelle $\vartheta = 0$ haben unabhängig von $n \geq 1$ die Verteilung

$$\mathcal{L}\left(n \left(\widehat{\vartheta}_n - (0 + h/n) \right) \mid P_{0+h/n} \right) \;=\; \mathcal{L}\left(\frac{\int_0^1 B_s \, dB_s}{\int_0^1 B_s^2 \, ds} \;-\; h \right) \, , \quad h \in I\!R \, ,$$

B eine Standard-Brownsche Bewegung, als Spezialfall von 7.10 und 7.11;

- an der Stelle $\vartheta = 0$ hat man Effizienzkriterien weder im Sinne des lokalasymptotischen Faltungssatzes 7.7+7.9 noch im Sinne des lokalasymptotischen Minimaxsatzes 7.13. □

Anhang A9 (Harris-Rekurrenz)

In diesem Anhang sei $(P_x)_{x \in \mathbb{R}^d}$ eine Markov-Familie auf dem kanonischen Pfadraum $(C, \mathcal{C}, \mathbb{G})$ für stetige $\mathbb{R}^d$-wertige Prozesse, etwa die Familie der Verteilungen der Lösung der SDE (I) aus 9.1 in Abhängigkeit vom Startwert x, und X der kanonische Prozeß auf $(C, \mathcal{C}, \mathbb{G})$. Die Markov-Halbgruppe von X wird bezeichnet mit $P_t(x, dy)$, $t \geq 0$, $x, y \in \mathbb{R}^d$; ein invariantes Maß ist ein σ-endliches Maß m auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit der Eigenschaft

$$\int m(dx)\, P_t(x, A) \;=\; m(A) \;, \quad A \in \mathcal{B}(\mathbb{R}^d) \;,\; t \geq 0 \;.$$

A9.1 Definition: X heißt *Harris-rekurrent* (oder kurz: *Harris*) falls es ein σ-endliches Maß ν auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ gibt mit der folgenden Eigenschaft (H):

$$(H) \quad A \in \mathcal{B}(\mathbb{R}^d) \;,\; \nu(A) > 0 \quad \Longrightarrow \quad \int_0^\infty 1_A(X_s)\, ds \;=\; +\infty \quad P_x \text{ fast sicher für jedes } x \in \mathbb{R}^d \;.$$

(H) besagt, daß die Aufenthaltszeit des Prozesses X in Mengen von positivem ν-Maß im langen Lauf jede noch so große Schranke übersteigen wird, unabhängig von der Wahl des Startpunkts x.

A9.2 Fakt: 1) Unter *(H)* existiert ein (bis auf Multiplikation mit einer Konstanten) eindeutig bestimmtes invariantes Maß m für den Prozeß X, und die Aussage *(H)* gilt mit m anstelle von ν. Siehe Azéma, Duflo und Revuz (1969, Theorem 2.5 mit Bemerkung).
Man nennt X *positiv rekurrent* oder *ergodisch* falls das invariante Maß m ein *endliches* Maß auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ist (damit o.E. ein Wahrscheinlichkeitsmaß), und *nullrekurrent* sonst.
2) Unter *(H)* gilt das *Ratio Limit Theorem*

$$\text{für jedes } x \text{ gilt } P_x\text{-fast sicher} \;:\; \frac{\int_0^t f(X_s)\, ds}{\int_0^t g(X_s)\, ds} \;\longrightarrow\; \frac{m(f)}{m(g)} \;, \quad t \to \infty$$

für Funktionen $g \in L^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m)$ mit $m(g) \neq 0$ und $f \in L^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m)$ (oder nur: $f \geq 0$ meßbar); dabei schreiben wir $m(g) = \int m(dy) g(y)$. Siehe Azéma, Duflo und Revuz (1969, Theorem 3.1). Im Fall positiver Rekurrenz und $m(\mathbb{R}^d) = 1$ vereinfacht sich die Formulierung zu

$$(A9.3) \qquad \text{für jedes } x \text{ gilt } P_x\text{-fast sicher} \;:\; \frac{1}{t} \int_0^t f(X_s)\, ds \;\longrightarrow\; m(f) \;, \quad t \to \infty \;.$$

Der Grenzwertsatz (A9.3) unter positiver Rekurrenz ist für den Statistiker ein wichtiges Werkzeug zum Nachweis von LAN. Unter Nullrekurrenz gibt es andere Typen von Grenzwertsätzen, siehe

Bingham, Goldie und Teugels (1987) und Khasminskii (1980); siehe auch Höpfner und Löcherbach (2003).

A9.4 Bemerkung: Der Nachweis der Harris-Rekurrenz wird einfacher, wenn man bereits ein invariantes Maß m für X kennt: in diesem Fall folgt *(H)* mit $\nu := m$ aus der Bedingung *(H')*:

$$(H') \quad A \in \mathcal{B}(\mathbb{R}^d)\ ,\ m(A) > 0 \quad \Longrightarrow \quad \limsup_{t\to\infty} 1_A(X_t) = 1 \quad P_x\text{-fast sicher für jedes } x \in \mathbb{R}^d\ ;$$

siehe Revuz und Yor (1991, S. 395–396). *(H')* bedeutet, daß X in Mengen von positivem invariantem Maß im langen Lauf stets weitere Besuche abstatten wird, für jede Wahl eines Startpunkts x.

Wir geben einige Beispiele in Dimension $d = 1$.

A9.5 Beispiel: Die eindimensionale Brownsche Bewegung ist nullrekurrent im Sinn von Harris:

Schreibe P_x für die Verteilung von $B + x$ auf $(C, \mathcal{C}, \mathbb{G})$, B eine Standard-Brownsche Bewegung. Das Lebesguemaß $\lambda\!\!\lambda$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ist ein invariantes Maß; dies sieht man sofort aus

$$\int dy\, \mathcal{N}(y,t)(A) \;=\; \lambda\!\!\lambda(A)\ , \quad A \in \mathcal{B}(\mathbb{R})\ .$$

Sei $A \in \mathcal{B}(\mathbb{R})$ mit $\lambda\!\!\lambda(A) > 0$; dabei reicht es, $A \subset [m, m+1[$ für $m \in \mathbb{Z}$ zu betrachten. Das Gesetz vom iterierten Logarithmus angewandt auf $B + x$ zeigt zusammen mit der Stetigkeit der Pfade, daß der Prozeß X im langen Lauf jeden Streifen $[m, m+1[$ P_x-fast sicher unendlich oft überqueren wird. Damit gilt auch

$$\limsup_{t\to\infty} 1_A(X_t) = 1 \quad P_x\text{-fast sicher für jedes } x \in \mathbb{R}^d\ ,$$

und Bedingung *(H')* aus A9.4 ist nachgewiesen; $\lambda\!\!\lambda(\mathbb{R}) = +\infty$ bedeutet Nullrekurrenz.
(Alternativ kann man auch mit Zyklenzerlegungen der Trajektorie von X argumentieren: man kennt die Verteilung der Level-Crossing-Zeiten der Standard-Brownschen Bewegung (dies sind stabile Verteilungen mit Parameter $\frac{1}{2}$, als Folgerung aus dem Spiegelungsprinzip); definiert man für beliebige Paare $-\infty < K_1 < K_2 < \infty$ $\mathbb{G}$-Stopzeiten

$$R_i \;:=\; \inf\{t > S_i : X_t = K_2\} \quad \text{mit} \quad S_i \;:=\; \inf\{t > R_{i-1} : X_t = K_1\}\ , \quad i \geq 1\ , \quad R_0 \equiv 0\ ,$$

so gilt mit $a := 2(K_2 - K_1)$ und $c := |x - K_1|$ für $i \geq 1$

$$P_x\,(S_1 \in dt) \;=\; \frac{c}{\sqrt{2\pi}}\, t^{-\frac{3}{2}}\, e^{-\frac{c^2}{2t}}\ , \quad P_\bullet\,(R_{i+1} - R_i \in dt) \;=\; \frac{a}{\sqrt{2\pi}}\, t^{-\frac{3}{2}}\, e^{-\frac{a^2}{2t}}\ , \quad t \geq 0\ .$$

Keine dieser Wahrscheinlichkeitsverteilungen besitzt endliche Erwartungswerte). □

A9.6 Beispiel: In Dimension $d = 1$: Ist $\sigma(\cdot)$ strikt positiv, so ist die Lösung der SDE (I) mit $b(\cdot) \equiv 0$ rekurrent im Sinn von Harris, mit invariantem Maß $m(dx) = \frac{1}{\sigma^2(x)}\,dx$:

Sei P_x die Verteilung der Lösung der SDE $d\xi_t = \sigma(\xi_t)\,dW_t$ zum Startwert x; wir brauchen hier – schwächer als in 9.1 – nur die Voraussetzung, daß $\sigma(\cdot)$ stetig und strikt positiv auf $I\!R$ ist; siehe Ikeda und Watanabe (1989, p. 173) oder Karatzas und Shreve (1991, p. 332). Die Behauptung wird sich aus einer Darstellung von X als zeittransformierte Brownsche Bewegung ergeben.

1) Auf dem kanonischen Pfadraum $(C, \mathcal{C}, \mathbb{G})$ sind alle Pfade des Prozesses

$$(+) \qquad t \longrightarrow \Phi_t := \int_0^t \sigma^2(X_s)\,ds\,, \quad t \geq 0$$

stetig und streng monoton wachsend; wir zeigen:

$$(++) \qquad \Phi_\infty := \int_0^\infty \sigma^2(X_s)\,ds \;=\; \infty \quad P_x\text{-fast sicher, für jedes } x \in I\!R\,.$$

Angenommen, das Ereignis $\{\int_0^\infty \sigma^2(X_s)\,ds \;<\; \infty\}$ hätte positive P_x-Wahrscheinlichkeit für ein x. Der Prozeß (+) ist die spitze Klammer des lokalen P_x-Martingals $X - X_0$, also gilt

$$\lim_{t\to\infty} X_t \quad \text{existiert in } I\!R \;\; P_x\text{-fast sicher auf} \quad \{\int_0^\infty \sigma^2(X_s)\,ds \;<\; \infty\}$$

(Lepingle 1978). Offensichtlich gilt stets (wegen $\sigma(\cdot)$ stetig und strikt positiv)

$$\int_0^\infty \sigma^2(X_s)\,ds \;=\; \infty \quad \text{auf dem Ereignis} \quad \{\lim_{t\to\infty} X_t \quad \text{existiert in } I\!R\}\,,$$

was unter der gemachten Annahme auf einen Widerspruch führt: also war diese absurd. Wegen (++) liefert (+) eine zufällige Zeittransformation $[0,\infty] \leftrightarrow [0,\infty]$, P_x-fast sicher für jedes x, und

$$B \;:=\; (X - X_0) \circ \Phi^{-1}$$

ist eine Standard Brownsche Bewegung auf $\left(C, \mathcal{C}, \left(\mathcal{G}_{\Phi_t^{-1}}\right)_{t\geq 0}\right)$, unter jedem P_x. Dabei bezeichnet Φ^{-1} den zu Φ inversen Prozeß, d.h. $\Phi_v^{-1} = \inf\{t \geq 0 : \Phi_t > v\}\,,\, v \geq 0$. Wegen

$$\int_0^{\Phi_v^{-1}} f(x + B_{\Phi_s})\,ds \;=\; \int_0^v f(x + B_r)\,d(\Phi_r^{-1}) \;=\; \int_0^v f(x + B_r)\,\frac{dr}{\sigma^2(X_{\Phi_r^{-1}})}$$

($f(\cdot)$ meßbar und nichtnegativ) gilt für beliebiges festes x

$$(*) \qquad \int_0^\infty f(X_s)\,ds \;=\; \int_0^\infty \left(\frac{f}{\sigma^2}\right)(x + B_r)\,dr \quad P_x\text{-fast sicher}$$

und speziell für $x = 0$

$$(**) \qquad \int_0^t f(X_s)\, ds \;=\; \int_0^{\Phi_t} \left(\frac{f}{\sigma^2}\right)(B_r)\, dr\ , \quad t \geq 0\ , \quad P_0\text{-fast sicher}\ .$$

2) Wir zeigen, daß der Prozeß X Harris-rekurrent ist.

Sei $A \in \mathcal{B}(\mathbb{R})$ mit $\lambda\!\!\lambda(A) > 0$; es reicht zu betrachten $A \subset [m, m+1[$ für ein $m \in \mathbb{Z}$. Nach Voraussetzung ist $\sigma(\cdot)$ (stetig und strikt positiv) auf $[m, m+1[$ beschränkt weg von 0 und ∞. Dann aber zeigt $(*)$ mit $f := 1_A$ wegen der Harris-Eigenschaft der Brownschen Bewegung

$$\int_0^\infty 1_A(X_s)\, ds \;=\; +\infty \quad P_x\text{-fast sicher, für jedes } x \in \mathbb{R}\ .$$

Damit ist für den Prozeß X die Bedingung *(H)* mit $\nu := \lambda\!\!\lambda$ erfüllt, und X ist Harris.

3) Wir zeigen, daß das invariante Maß m von X durch $\frac{1}{\sigma^2(x)}\, dx$ gegeben ist.

i) Es gibt $g \geq 0$ meßbar mit den Eigenschaften $0 < m(g) < \infty$ und $0 < \lambda\!\!\lambda(\frac{1}{\sigma^2} g) < \infty$:

Wähle $g \in L^1(m)$ mit $m(g) > 0$; o.E. sei dabei g nichtnegativ, beschränkt, und $\{g > 0\}$ sei enthalten in $[m, m+1[$ für ein $m \in \mathbb{Z}$. Wegen $\sigma(\cdot)$ stetig und strikt positiv liegt dann $\widetilde{g} := \frac{1}{\sigma^2} g$ in $L^1(\lambda\!\!\lambda)$; wir zeigen $\lambda\!\!\lambda(\widetilde{g}) > 0$. Wäre $\lambda\!\!\lambda(\widetilde{g}) = 0$, hätte man für die Standard-Brownsche Bewegung B aus Schritt 1) $E(\widetilde{g}(B_s)) = 0$ für jedes s, damit f.s. $\int_0^v \widetilde{g}(B_s) ds = 0$ für jedes $v \geq 0$, und damit nach $(**)$ den Widerspruch

$$\int_0^\infty \widetilde{g}(B_s) ds \;=\; \int_0^\infty g(X_s) dx \;=\; 0 \quad P_0\text{-fast sicher}$$

zu $m(g) > 0$ und zur Harris-Eigenschaft *(H)* (mit $\nu = m$) von X .

ii) Sei g wie in i) fest. Wir betrachten die Klasse aller nichtnegativen stetigen Funktionen f mit kompaktem Träger. Mit dem Ratio Limit Theorem für die Brownsche Bewegung, mit $(**)$ und dem Ratio Limit Theorem für X folgt P_0-fast sicher

$$(0,\infty) \ni \frac{\lambda\!\!\lambda\left(\frac{f}{\sigma^2}\right)}{\lambda\!\!\lambda\left(\frac{g}{\sigma^2}\right)} \;=\; \lim_{v\to\infty} \frac{\int_0^v \left(\frac{f}{\sigma^2}\right)(B_r)\, dr}{\int_0^v \left(\frac{g}{\sigma^2}\right)(B_r)\, dr} \;=\; \lim_{t\to\infty} \frac{\int_0^t f(X_s)\, ds}{\int_0^t g(X_s)\, ds} \;=\; \frac{m(f)}{m(g)}\ .$$

Als Konsequenz der Wahl von g ergibt sich

$$\lambda\!\!\lambda\left(\frac{f}{\sigma^2}\right) \;=\; c(g) \cdot m(f) \quad \text{für alle } f \geq 0 \text{ stetig mit kompaktem Träger}\ .$$

mit $c(g) := \lambda\!\!\lambda\left(\frac{g}{\sigma^2}\right)/m(g)$; damit hat m eine Lebesgue-Dichte der Form $x \to \frac{1}{c(g)} \frac{1}{\sigma^2(x)}$. Auf multiplikative Konstanten kommt es bei der Festlegung des invarianten Maßes für X nicht an. $\square$

A9.7 Beispiel: (siehe Has'minskii 1980, Section III.8, Ex. 2) In Dimension $d = 1$: Lösungen der SDE (I) sind rekurrent im Sinn von Harris, mit invariantem Maß

$$(A9.8) \qquad m(dx) \ := \ \frac{1}{\sigma^2(x)} \, exp\left(\int_0^x \frac{2b}{\sigma^2}(v)\, dv \right)$$

auf $(I\!R, \mathcal{B}(I\!R))$, falls die folgende Bedingung erfüllt ist:

> $\sigma(\cdot)$ ist strikt positiv, und die $\mathcal{C}^2$-Abbildung $S : I\!R \to I\!R$
>
> $$(A9.9) \qquad S(x) \ := \ \int_0^x s(y)\, dy \quad , \quad s(y) \ := \ \exp\left(- \int_0^y \frac{2b}{\sigma^2}(v)\, dv \right)$$
>
> ist eine Raumtransformation, d.h. eine Bijektion auf $I\!R$.

Dies sieht man so: Für S definiert durch (A9.9) zeigt die Ito-Formel unter P_x, daß der Prozeß

$$\widetilde{X} \ := \ S \circ X \ = \ (S(X_t))_{t \geq 0}$$

unter P_x eine Diffusion ohne Drift ist, mit Diffusionskoeffizient

$$\widetilde{\sigma} \ := \ (s \cdot \sigma) \circ S^{-1}$$

und Startwert $\widetilde{x} := S(x)$. Dies gilt für jedes $\widetilde{x} \in I\!R$. Folglich ist nach A9.5 der Prozeß $\widetilde{X}$ Harris mit invariantem Maß

$$\widetilde{m}(dx) \ := \ \frac{dx}{\widetilde{\sigma}^2(x)} \ = \ \frac{1}{(s \cdot \sigma)^2(S^{-1}(x))}\, dx \ .$$

Die Harris-Eigenschaft *(H)* für $\widetilde{X}$

$$\widetilde{m}(A) > 0 \quad \Longrightarrow \quad \int_0^\infty 1_A(S(X_t))\, dt \ = \ +\infty \quad P_x\text{-fast sicher für jedes } x$$

übersetzt sich aber sofort in eine Bedingung *(H)* (mit $\nu := \widetilde{m}$) für den Prozeß X: damit ist auch X Harris. Bezeichne m das invariante Maß von X. Das Ratio Limit Theorem für X kombiniert mit dem für $\widetilde{X}$ zeigt ähnlich wie in A9.6

$$m(f) \ = \ c \cdot \widetilde{m}(f \circ S^{-1}) \quad \text{für alle } f \geq 0 \text{ stetig mit kompaktem Träger}$$

für eine Konstante c, auf die es bei der Festlegung von m nicht ankommt. Also ist das invariante Maß m von X das Bild von $\widetilde{m}$ unter S^{-1}; nach Transformationsformel gilt (A9.8). □

Einen eher analytischen Zugang zum invarianten Maß einer Diffusion findet man in Büchern wie z.B. Bass (1998, p. 56).