Ergodicity and speed of convergence to equilibrium for diffusion processes.

Eva LÖCHERBACH*

November 14, 2015

Abstract

We discuss the long time behavior of diffusion processes (or more general Markov processes). We start by introducing the basic concept of Harris recurrence and establish the link with ergodic theory. We recall classical recurrence conditions from the theory of Markov chains (Doeblin condition, Dobrushin condition, local Doeblin condition). Then we turn to the study of one dimensional diffusions where hitting time moments determine the speed of convergence. We recall Kac's moments formula for hitting times and characterize the speed of convergence to equilibrium under Vertennikov's drift conditions both in the ergodic theorem and for the total variation distance. In higher dimensional models we show how to use coupling techniques in order to introduce regeneration times that allow to mimic the one dimensional case.

Key words : Harris recurrence, polynomial ergodicity, Nummelin splitting.

MSC 2000 : 60 J 55, 60 J 35, 60 F 10, 62 M 05

Contents

*CNRS UMR 8088, Département de Mathématiques, Université de Cergy-Pontoise, 95 000 CERGY-PONTOISE, France. E-mail: `eva.loecherbach@u-cergy.fr`

1 Harris-recurrence and ergodicity: Definitions, basic results and examples

1.1 Harris-recurrence

Consider a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, (P_x)_x)$, where $(\mathcal{F}_t)_{t\geq 0}$ is some right continuous filtration. Let $X = (X_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\geq 0}-$adapted process defined on $(\Omega, \mathcal{A})$. We suppose that X takes values in a locally compact Polish space $(E, \mathcal{E})$ and is strong Markov with respect to $(\mathcal{F}_t)_{t\geq 0}$ under P_x for any $x \in E$. Moreover we assume that X has càdlàg paths. The family $(P_x)_{x\in E}$ is a collection of probability measures on $(\Omega, \mathcal{A})$ such that $X_0 = x$ P_x-almost surely. We write $(P_t)_t$ for the transition semigroup of X. In particular we assume that the process is homogeneous in time.

We start by discussing the main concept of stochastic stability of Markov processes: the recurrence property. Basically, recurrence means that the process "comes back" almost surely and does not escape to ∞ as $t \to \infty$. In this statement, "coming back" and "not escaping to ∞" have to be defined in a precise manner. But before doing so, let us consider two very well known examples.

Example 1.1 *Consider standard Brownian motion $(B_t)_{t\geq 0}$ in dimension 1, $B_0 = 0$. This* *process is recurrent in an elementary sense, since*

$$\limsup_{t\to\infty} 1_{\{0\}}(B_t) = 1$$

almost surely; the process comes back to the point 0 *infinitely often and at arbitrary late times. The same is actually true for any other point* $a \in I\!R$ *: Let* $T_a = \inf\{t : B_t = a\}$, *then we have* $T_a < +\infty$ *almost surely and*

$$T_a \sim \frac{1}{\sqrt{2\pi}} e^{-a^2/t} \, |a| \, t^{-3/2} dt.$$

Proof. *Suppose w.l.o.g. that* $a > 0$. *Let* $\lambda > 0$, *then* $M_t = \exp\left(\lambda B_t - \frac{\lambda^2}{2} t\right)$ *is a martingale. Applying the stopping rule, we have for any* $N \geq 1$,

$$E_0 \left[e^{\lambda B_{T_a \wedge N} - \frac{\lambda^2}{2} T_a \wedge N} \right] = 1.$$

Since $e^{\lambda B_{T_a \wedge N}} \leq e^{\lambda a}$, *dominated convergence gives as* $N \to \infty$,

$$E_0 \left[e^{\lambda a - \frac{\lambda^2}{2} T_a} 1_{\{T_a < \infty\}} \right] = 1. \tag{1.1}$$

Letting $\lambda \to 0$ *implies that* $T_a < \infty$ *almost surely. (1.1) is the Laplace transform of* T_a *and gives easily the density. For* $a < 0$, *the result follows by symmetry.*

Let us consider a second example.

Example 1.2 $2-$*dimensional standard Brownian motion* $(B_t)_{t\geq 0}$ *is still recurrent, but does not come back to points any more. Actually, the following is true: For all sets* $A \in \mathcal{B}(I\!R^2)$ *having positive Lebesgue measure* $\lambda(A) > 0$, *we have* $\limsup_{t\to\infty} 1_A(B_t) = 1$ *almost surely. Thus, all Lebesgue positive sets are visited infinitely often by the process. The same is true for one-dimensional Brownian motion. For dimension* $d \geq 3$, *Brownian motion is not recurrent any more.*

These considerations lead to the following definition. Recall that a measure μ on $(E, \mathcal{E})$ is called invariant if $\mu P_t = \mu$ for all $t \geq 0$, i.e. $\int_E \mu(dx) P_t(x, dy) = \mu(dy)$.

Definition 1.3 X *is called recurrent in the sense of Harris if* X *possesses a* $\sigma-$*finite invariant measure* μ *such that for all* $A \in \mathcal{E}$,

$$\mu(A) > 0 \quad \textit{implies} \quad \limsup_{t\to\infty} 1_A(X_t) = 1 \; P_x- \textit{ almost surely for all } x \in E. \tag{1.2}$$

A deep result of Azéma, Duflo and Revuz (1969) implies that for a Harris recurrent process, the measure μ is necessarily unique, up to multiplication with a constant. Therefore we can introduce the following classification of Harris recurrent processes.

Definition 1.4 *Let* X *be Harris recurrent, with invariant measure* μ. *If* $\mu(E) < \infty$, *then* X *is called* **positive recurrent**, *else* **null recurrent**.

Hence, Brownian motion in dimension $d = 1$ and $d = 2$ is null recurrent, since Lebesgue measure is its invariant measure.

We consider another well known example of a one-dimensional diffusion process which is positive recurrent. It is the recurrent Ornstein Uhlenbeck process. This process is even recurrent in a much stronger sense, we will see that it is exponentially recurrent or exponentially ergodic in a sense that will be defined precisely later. The presentation of the next example is widely inspired by a course given by Reinhard Höpfner in Bonn in 1995/96, see also Höpfner (2008), *http://www.informatik.uni-mainz.de/ hoepfner/Material-MathStat.html*, Kap IX.

Example 1.5 *Let B be one-dimensional Brownian motion and*

$$dX_t = -aX_t dt + dB_t, X_0 = x_0, a > 0.$$

Apply the Itô-formula to $e^{at}X_t$ in order to obtain the explicit solution of the above equation

$$X_t = e^{-at}x_0 + \int_0^t e^{-a(t-s)} dB_s. \tag{1.3}$$

In particular,

$$X_t \sim \mathcal{N}(e^{-at}x_0, \frac{1}{2a}(1 - e^{-2at})).$$

This implies that for any bounded function f,

$$P_t f(x) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_t^2}(y-xe^{-at})^2} f(y)dy \to \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}y^2} f(y)dy,$$

as $t \to \infty$, where $\sigma_t^2 = \frac{1}{2a}(1 - e^{-2at})$ and $\sigma^2 = \frac{1}{2a}$. Therefore, $X_t \xrightarrow{\mathcal{L}} \mu := \mathcal{N}(0, \frac{1}{2a})$, as $t \to \infty$. It is immediate to see that μ is an invariant measure of X. [1]

We show that X is recurrent in the sense of Harris. We have to show that $\mu(A) > 0$ implies that X returns to A almost surely. But X is recurrent in an elementary sense: Suppose that $x_0 > 0$ and take Brownian motion starting from $x_0 : X_0 = B_0 = x_0$. Then $b(X_t) = -aX_t < 0$ on $[0, \tau[$, where $\tau = \inf\{t : X_t = 0\}$. Using the comparison theorem for diffusions, this implies $X_t \le B_t$ for all $t < \tau$. Therefore, $\tau \le \inf\{t : B_t = 0\} \stackrel{\mathcal{L}}{=} T_{x_0}$, where T_{x_0} is the hitting time of x_0 of a standard Brownian motion starting from 0 at time 0. Hence $\tau < \infty$ almost surely. The same argument applies to a starting point $x_0 < 0$.

As a consequence, we can cut the trajectory of X into i.i.d. excursions out of 0. Put

$$R_1 = \inf\{t : X_t = 0\}, R_{n+1} = \inf\{t > R_n + 1 : X_t = 0\},\ n \ge 1.$$

In the above definition, we take the infimum over all $t > R_n + 1$ (we could as well take $t > R_n + \varepsilon$ for any fixed $\varepsilon > 0$) in order to ensure that the sequence $(R_n)_n$ does not have accumulation points.

[1]First proof: We have to show that $\mu P_t = \mu$. But $\mu = \lim_{s\to\infty} P_s$, hence for any bounded measurable function f, $\mu P_t f = \lim_{s\to\infty} P_s P_t f = \lim_{s\to\infty} P_{s+t} f = \mu f$, where we have used the strong Feller property of the process. Second proof: Choose X_0 independently of $(B_t)_t$, $X_0 \sim \mu$. Then $e^{-at}X_0 \sim \mathcal{N}(0, e^{-2at}\frac{1}{2a})$. This implies by (1.3) that $X_t \sim \mathcal{N}(0, e^{-2at}\frac{1}{2a}) * \mathcal{N}(0, \frac{1}{2a}(1 - e^{-2at}) = \mu$.

Clearly, $R_n < \infty$ almost surely for all n, $R_n \to \infty$ almost surely as $n \to \infty$. Moreover, the trajectories $(X_{R_n+s})_{s \le R_{n+1}-R_n}, n \ge 1$, are i.i.d. The strong law of large numbers implies then that, with $\xi_k = \int_{R_k}^{R_{k+1}} 1_A(X_s)ds$,

$$\frac{1}{n}\int_0^{R_{n+1}} 1_A(X_s)ds = \frac{1}{n}\int_0^{R_1} 1_A(X_s)ds + \frac{1}{n}\sum_{k=1}^n \xi_k \to m(A),$$

as $n \to \infty$, where

$$m(A) = E\left[\int_{R_1}^{R_2} 1_A(X_s)ds\right]. \tag{1.4}$$

$m(A)$ is the mean occupation time of A during one excursion out of 0. This implies that for A, B such that $m(A) < \infty$,

$$\frac{\int_0^t 1_A(X_s)ds}{\int_0^t 1_B(X_s)ds} \to \frac{m(A)}{m(B)},$$

almost surely as $t \to \infty$. In particular, taking $B = I\!\!R$,

$$\frac{1}{t}\int_0^t 1_A(X_s)ds \to \frac{m(A)}{m(I\!\!R)},$$

where the limit is defined to be 0 if $m(I\!\!R) = \infty$. (We will see in a minute that this can not be the case.)

By dominated convergence, the above convergence holds true in $L^1(P_x)$, for all $x \in I\!\!R$. Hence

$$\frac{1}{t}\int_0^t P_s 1_A(x)ds \to \frac{m(A)}{m(I\!\!R)}.$$

On the other hand, we know that $P_s 1_A(x) \to \mu(A)$. As a consequence,

$$\frac{m(\cdot)}{m(I\!\!R)} = \mu = \mathcal{N}(0, \frac{1}{2a}), \tag{1.5}$$

and in particular, $m(I\!\!R) < \infty$ which is equivalent to $E(R_2 - R_1) < \infty$.

This finishes our proof: Indeed, $\mu(A) > 0$ being equivalent to $m(A) > 0$, by the strong law of large numbers, X visits A infinitely often almost surely. Thus, X is Harris with unique invariant probability measure μ.

Let us now collect some known results that hold for all Harris recurrent Markov processes. The proofs are postponed to the appendix.

Theorem 1.6 (Azéma, Duflo and Revuz (1969)) *For measurable, positive functions f and g with $0 < \mu(g) < \infty$,*

1. $\lim_{t\to\infty} \frac{E_x(\int_0^t f(X_s)ds)}{E_x(\int_0^t g(X_s)ds)} = \frac{\mu(f)}{\mu(g)}$ *$\mu-$a.s. (the exceptional set depends on f, g),*

2. $\lim_{t\to\infty} \frac{\int_0^t f(X_s)ds}{\int_0^t g(X_s)ds} = \frac{\mu(f)}{\mu(g)}$ *P_x-a.s. $\forall\, x$.*

3. In particular, if $\mu(E) < \infty$, we choose $\mu(E) = 1$. Then we have for any $x \in E$,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(X_s)ds = \mu(f)$$

P_x-almost surely and

$$\lim_{t\to\infty} \frac{1}{t} E_x \int_0^t f(X_s)ds = \mu(f). \tag{1.6}$$

Warning: (1.6) does not imply that $P_t f(x) \to \mu(f)$ as $t \to \infty$ – although this will be the case in most of the examples that we shall consider.

1.1.1 Poincaré inequality for the Ornstein-Uhlenbeck process

Let us come back to example 1.5. The invariant measure $\mu = \mathcal{N}(0, \frac{1}{2a})$ of the Ornstein-Uhlenbeck process has very good concentration properties which are related to the speed of convergence to equilibrium. Suppose for simplicity that $a = 1/2$, then $\mu = \mathcal{N}(0,1)$. The generator of the Ornstein-Uhlenbeck process is given by $Af(x) = -\frac{1}{2}xf'(x) + \frac{1}{2}f''(x)$ and the associated Dirichlet form by

$$\mathcal{E}(f,f) = -\int fAfd\mu.$$

Partial integration shows that for any function f such that $f' \in L^2(\mu)$,

$$\mathcal{E}(f,f) = \frac{1}{2}\int [f'(x)]^2 \mu(dx).$$

The measure μ satisfies the **Poincaré inequality**

$$\int f^2(x)\mu(dx) \leq 2\mathcal{E}(f,f), \ \int fd\mu = 0, \tag{1.7}$$

for any function $f \in L^2(\mu) \cap C^1(I\!R)$.

Proof of (1.7): Consider the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}, n \geq 0.$$

Then we have $H_{n+1}(x) = xH_n(x) - H_n'(x)$ and $H_n'(x) = nH_{n-1}(x)$. It is well known that $\{\frac{H_n}{\sqrt{n!}}, n \geq 0\}$ is an ONB of $L^2(\mu)$. We have, using that $H_n' = nH_{n-1}, H_n'' = nH_{n-1}'$,

$$AH_n(x) = \frac{1}{2}n\left[-xH_{n-1} + H_{n-1}'\right] = -\frac{n}{2}H_n.$$

We are now ready to prove (1.7). Let $f \in L^2(\mu)$, $\mu(f) = 0$. Hence $f = \sum_{n\geq 1} c_n H_n$, where $c_0 = 0$, since f is centered. Then

$$\int f^2 d\mu = \sum_{n=1}^{\infty} n!\, c_n^2.$$

On the other hand,

$$Af = \sum_{n=1}^{\infty} c_n A H_n = -\sum_{n=1}^{\infty} \frac{n}{2} c_n H_n,$$

thus

$$\begin{aligned} 2\mathcal{E}(f,f) = -2\int Af(x)f(x)\mu(dx) &= \sum_{n=1}^{\infty} n c_n^2 \int H_n^2 d\mu \\ &= \sum_{n=1}^{\infty} n(n!)\, c_n^2 \geq \sum_{n=1}^{\infty} (n!)\, c_n^2 = \int f^2(x)\mu(dx). \end{aligned}$$

This concludes the proof of the Poincaré inequality. •

Poincaré's inequality implies exponential speed of convergence to equilibrium. Write $\|f\|_2$ for $\mu(f^2)$. Then

Corollary 1.7 *For all $f \in L^2(\mu), \mu(f) = 0$,*

$$\|P_t f\|_2 \leq e^{-t/2}\|f\|_2. \tag{1.8}$$

Proof. There exists a direct proof, using the Hermite polynomials. We give a more general proof which works for any Markov process such that the Poincaré inequality is satisfied, and which is inspired by Liggett (1991).

Write $f_t = P_t f$. Suppose that $f \in \mathcal{D}(A) = \{f \in L^2(\mu) : \|\frac{P_t f - f}{t} - g\|_2 \to 0$ for some $g \in L^2(\mu)\}$, where $g =: Af$. Then we have $\frac{d}{dt}\int f_t^2 d\mu = 2\int f_t A f_t d\mu = -2\mathcal{E}(f_t, f_t)$. But the Poincaré inequality implies that $-2\mathcal{E}(f_t, f_t) \leq -\int f_t^2 d\mu$. So

$$\frac{d}{dt}\int f_t^2 d\mu \leq -\int f_t^2 d\mu,$$

which implies that

$$\int f_t^2 d\mu \leq e^{-t}\int f^2 d\mu \iff \|P_t f\|_2 \leq e^{-t/2}\|f\|_2.$$

For general $f \in L^2(\mu)$, the assertion follows, since $\mathcal{D}(A) \subset L^2(\mu)$ is dense. •

Remark 1.8 1. *The above results hold in $L^2(\mu)$. This means that we consider the strictly stationary process starting from its invariant measure $X_0 \sim \mu$. For the process starting from some fixed initial condition, in general, it is more difficult to obtain results in the above spirit.*

2. *The Poincaré inequality is also satisfied by the Lebesgue measure and (1.8) remains true for Brownian motion in dimension one. However there is a fundamental difference: Lebesgue measure is* **not** *a finite measure, and Brownian motion in dimension one is only null recurrent, i.e. "almost transient". It does not converge to its equilibrium in most of the common senses. For example we will see that for the recurrent Ornstein-Uhlenbeck process, the speed of convergence of $P_t(x, dy)$ to μ is exponential for the total variation distance. This is not the case for Brownian motion: $P_t(x, dy) = \mathcal{N}(x, t)$ does not converge at all as $t \to \infty$, so in particular not to λ.*

1.2 Harris-recurrence, ergodicity and strong mixing

By abuse of language, positive Harris recurrent processes are often called "ergodic". In this section, our aim is to recall briefly the basic concepts of ergodic theory and to establish links between "ergodicity" and "Harris recurrence". Our presentation is widely inspired by Delyon (2012) and Genon-Catalot, Jeantheau and Larédo (2000).

Throughout this section, we consider a process $(X_t)_{t \geq 0}$ defined on a fixed probability space $(\Omega, \mathcal{A}, P)$, having càdlàg paths, taking values in $(E, \mathcal{E})$. Write $D = D(\mathbb{R}_+, E)$ for the Skorokhod space of càdlàg functions $f : \mathbb{R}_+ \to E$, endowed with the Skorokhod topology, and let $\mathcal{D}$ be the associated Borel sigma field. We introduce the shift operator on D by $\vartheta_t f = f(t + \cdot)$ for any $f \in D, t \geq 0$. We write for short X for the whole process, i.e. for the $D-$valued random variable $X = (X_t)_{t \geq 0}$.

Definition 1.9 *X is called* **stationary** *if for all $t \geq 0$, X and $X_{t+\cdot}$ have the same law under P.*

Hence a Markov process X having invariant (probability) measure μ is a stationary process if $X_0 \sim \mu$. In other words, this means that we work under the probability measure $P_\mu := \int_E \mu(dx) P_x$.

Definition 1.10 *A set $A \in \mathcal{D}$ is called* **invariant** *if for all $t \geq 0$, we have $A = \vartheta_t^{-1}(A)$. We introduce the sigma field of invariant events*

$$\mathcal{J} = \{X^{-1}(A) : A \text{ is invariant}\}.$$

Now we are ready to introduce the notion of ergodicity.

Definition 1.11 *The process X is called* **ergodic** *if for all $B \in \mathcal{J}$, $P(B) \in \{0, 1\}$.*

Hence for ergodic processes, $\mathcal{J}-$measurable functions are $P-$almost surely constant.

Before making the link to Harris recurrence, let us state some of the classical results that are known to hold true for ergodic processes. The first is the famous Birkhoff ergodic theorem. We write $\mathbb{P}$ for the law of the stationary process X on the Skorokhod space $(D, \mathcal{D})$.

Theorem 1.12 *Let X be ergodic and $f : D \to \mathbb{R}, f \in L^1(\mathbb{P})$. Then*

$$\frac{1}{t} \int_0^t f(X_{s+\cdot}) ds \to E[f(X)|\mathcal{J}],$$

as $t \to \infty$. The above convergence holds $P-$almost surely and in $L^1(P)$.

Proof: We suppose w.l.o.g. that $\mathcal{J}$ is trivial and that f is bounded. Put $Y_n(f) = Y_n = \int_{n-1}^n f(X_{s+\cdot}) ds$ for all $n \geq 1$. Note that $Y_n = Y_1 \circ \vartheta_{n-1}(X)$. Finally, let

$$S_n = S_n(f) = Y_1 + \ldots + Y_n = \int_0^n f(X_{s+\cdot}) ds \text{ and } A = A(f) = \{\inf_n S_n(f) = -\infty\}.$$

Since f is bounded, we have

$$A(f) = \{\inf_t \int_0^t f(X_{s+\cdot})ds = -\infty\}.$$

The last equality shows that $A(f) \in \mathcal{J}$. We claim that

$$E(1_{A(f)} Y_1(f)) \leq 0. \tag{1.9}$$

Once (1.9) is obtained, we obtain the following implication: If $E(Y_1(f)) > 0$, then $P(A(f)) = 0$, since $A(f) \in \mathcal{J}$. We prove the claim (1.9) following Delyon (2012). Put $f_p = \inf(S_1, S_2, \ldots, S_p)$. Then

$$\begin{aligned} f_p &= Y_1 + \inf(0, Y_1 \circ \vartheta_1(X), \ldots, S_{p-1} \circ \vartheta_1(X)) \\ &= Y_1 + \inf(0, f_{p-1} \circ \vartheta_1(X)) \\ &\geq Y_1 + \inf(0, f_p \circ \vartheta_1(X)) \\ &= Y_1 + f_p^- \circ \vartheta_1(X), \end{aligned}$$

where $f_p^- = f_p \wedge 0$. Hence, by invariance of A and stationarity,

$$\begin{aligned} E(1_A Y_1) &\leq E(1_A f_p) - E(1_A f_p^- \circ \vartheta_1(X)) \\ &= E(1_A f_p) - E(1_A f_p^-) \\ &= E(1_A f_p^+). \end{aligned}$$

By definition of A, the last term tends to 0 as $p \to \infty$ (use dominated convergence and the fact that $f_p \leq Y_1$.) This concludes the proof of (1.9).

Now consider a fixed $\varepsilon > 0$ and $f_\varepsilon := f - E(f(X)) + \varepsilon$, $Y_1(f_\varepsilon) = \int_0^1 f_\varepsilon(X_{s+\cdot})ds$. Clearly, $E(Y_1(f_\varepsilon)) > 0$. Hence, applying (1.9) to f_ε,

$$\liminf_{t\to\infty} \frac{1}{t} \int_0^t f_\varepsilon(X_{s+\cdot})ds \geq 0$$

$P-$almost surely, which implies that

$$\liminf_{t\to\infty} \frac{1}{t} \int_0^t f(X_{s+\cdot})ds \geq E(f(X)) - \varepsilon$$

almost surely. Applying the same result to $-f$ and letting $\varepsilon \to 0$ implies the result. •

The next result is sometimes useful in order to establish whether a given process is ergodic.

Proposition 1.13 *Let $C \subset L^1(\mathbb{P})$ be dense. If for all $f \in C$*

$$\frac{1}{t} \int_0^t f(X_{s+\cdot})ds \to E(f(X)), \tag{1.10}$$

then the process is ergodic.

Proof: It is standard to show that (1.10) holds in fact for all $f \in L^1(\mathbb{P})$. Let $f \in L^1(\mathbb{P})$ be invariant. Thus $f(X_{s+\cdot}) = f(X)$ for all $s \geq 0$. In particular, the left hand side of (1.10) equals $f(X)$. Therefore, $f(X) = E(f(X))$ almost surely, which means that $f(X)$ is almost surely constant. This implies the ergodicity. •

If X is strong Markov with invariant probability measure μ, then the following theorem gives an important relation between ergodicity and the dimension of the eigenspace associated to the eigenvalue 0 of the generator A. We quote its proof from Genon-Catalot, Jeantheau and Larédo (2000). Recall that $\mathcal{D}(A) = \{f \in L^2(\mu) : \|\frac{P_t f - f}{t} - g\|_2 \to 0 \text{ for some } g \in L^2(\mu)\}$, where $g =: Af$.

Theorem 1.14 *For a strictly stationary strong Markov process X having invariant probability measure μ the following two statements are equivalent.*

1. *X is ergodic.*
2. *0 is a simple eigenvalue of $A : \{f \in \mathcal{D}(A) : Af = 0\}$ is the one-dimensional subspace of $L^2(\mu)$ spanned by the constants.*

For the convenience of the reader, we recall here the proof of the above theorem, see also the proof of Theorem 2.1 in Genon-Catalot, Jeantheau and Catalot (2000).

Proof: We show first that 2. implies 1. For that sake, let Z be a square integrable $\mathcal{J}-$measurable random variable. Put $h(x) = E(Z|X_0 = x)$. Then, using the Markov property and the invariance of Z,

$$h(X_t) = E(Z|X_t) = E(Z \circ \vartheta_t|X_t) = E(Z \circ \vartheta_t|\mathcal{F}_t) = E(Z|\mathcal{F}_t).$$

In the second and in the last equality, we have used that $Z \circ \vartheta_t = Z$. The third equality is the Markov property. As a consequence, $(h(X_t))_t$ is a square integrable martingale. Thus

$$P_t h(X_s) = E[h(X_{t+s})|\mathcal{F}_s] = h(X_s)$$

almost surely, for all $t, s \geq 0$. In particular, $P_t h = h$ in $L^2(\mu)$ which means that $h \in \mathcal{D}(A)$ and $Ah = 0$. By assumption, this implies that h is constant almost surely. Now, by the convergence theorem for martingales,

$$h(X_t) = E(Z|\mathcal{F}_t) \to E(Z|\mathcal{F}_\infty) = Z$$

almost surely as $t \to \infty$. (Note that being $\mathcal{J}-$measurable, the random variable Z is in particular $\mathcal{F}_\infty-$measurable.) Hence Z is constant almost surely, and this implies the ergodicity of the process.

We show now that 1. implies 2. Let $h \in \mathcal{D}(A)$ with $Ah = 0$. Hence, $\frac{d}{dt} P_t h = 0$, or $P_t h = P_0 h = h$ almost surely. This implies that $(h(X_t))_t$ is a square integrable martingale. Thus there exists a square integrable, $\mathcal{F}_\infty-$measurable random variable Z such that $h(X_t) = E(Z|\mathcal{F}_t)$ and $Z = \lim_{t\to\infty} h(X_t)$ almost surely. Moreover, for any $s > 0$, $Z = \lim_{t\to\infty} h(X_{t+s}) = Z \circ \vartheta_s(X)$, which means that Z is $\mathcal{J}-$measurable. Hence Z is constant $P-$almost surely. Since $h(X_t) = E(Z|\mathcal{F}_t)$, this implies that h is constant in $L^2(\mu)$. •

We are now ready to establish the relation between ergodicity and Harris recurrence.

Theorem 1.15 *Let X be a strictly stationary strong Markov process having invariant measure μ.*

1. *If X is strong Feller and supp$(\mu) = E$,*[2] *then X is positive Harris recurrent.*

2. *If X is positive Harris recurrent, then X is ergodic.*

Proof: We prove the first item. A strong Feller process having an invariant probability of full support is ergodic with unique invariant measure. Now let A be such that $\mu(A) > 0$ and put $C = \{\int_0^\infty 1_A(X_s)ds = \infty\}$. Clearly, $C \in \mathcal{J}$. Hence $P(C) \in \{0, 1\}$. We show that necessarily, $P(C) = 1$. Indeed, on C^c, we have that

$$\frac{1}{t}\int_0^t 1_A(X_s)ds \to 0.$$

On the other hand, the ergodic theorem implies that

$$\frac{1}{t}\int_0^t 1_A(X_s)ds \to E[1_A(X_0)|\mathcal{J}] = \mu(A) > 0,$$

hence $P(C^c) = 0$. Since $P = \int \mu(dx)P_x$, this implies

$$P_x(C) = 1 \text{ for } \mu\text{-almost all } x \in E.$$

Now by invariance of C,

$$P_x(C) = \int P_t(x, dy)P_y(C),$$

which implies that $x \mapsto P_x(C)$ is continuous, by the strong Feller property of the process. On the other hand, $P_x(C) = 1$ for μ-almost all x and *supp*$(\mu) = E$. This implies that $P_x(C) = 1$ for all $x \in E$, which is the Harris recurrence of X.

We prove the second item. Suppose that X is positive Harris recurrent and strictly stationary. Write $\mathbb{P}$ for its law on $(D, \mathcal{D})$ and $\mathbb{P}_x$ for the law of X under P_x, i.e. when $X_0 = x$. Then also the path valued process $(X_{t+\cdot})_{t\geq 0}$ is positive Harris recurrent, taking values in $D(\mathbb{R}_+, E)$, with invariant measure $\mathbb{P}$. This can be seen as follows. First, the transitions Q_t of the path valued process are given by $Q_t(f, dg) = \delta_{\vartheta_t(f)}(dg)$. From this follows that $\mathbb{P}$ is invariant. Now let $C \in D(\mathbb{R}_+, E)$ such that $\mathbb{P}(C) > 0$. Since $\mathbb{P} = \int_E \mu(dx)\mathbb{P}_x$, this means $\int_E \mu(dx)\mathbb{P}_x(C) > 0$, which implies that for some $\varepsilon > 0$,

$$\mu\left(\{x : \mathbb{P}_x(C) > \varepsilon\}\right) > 0.$$

By Harris recurrence of X, this implies that

$$\mathbb{P}_{X_t}(C) = P_x(X_{t+\cdot} \in C|\mathcal{F}_t) > \varepsilon \text{ infinitely often, } P_x\text{-a.s. for all } x.$$

Then the conditional Borel-Cantelli lemma implies that $X_{t+\cdot} \in C$ infinitely often, which is the Harris recurrence of the path valued process.

Now, using the ratio limit theorem 1.6, for any function $f \in L^1(\mathbb{P})$,

$$\frac{1}{t}\int_0^t f(X_{s+\cdot})ds \to E(f(X))$$

almost surely as $t \to \infty$. By Proposition 1.13, this implies the ergodicity of the process. •

[2] Recall that the support of a measure is defined as follows: *supp*$(\mu) = \{x : \text{ if } x \in B \text{ for some open set } B, \text{ then } \mu(B) > 0\}$.

We close this subsection with some remarks on strong mixing. The **strong mixing property** is a property which is stronger than ergodicity, implying the latter, and often easier to check. Introduce $\mathcal{T}_t = \sigma\{X_{t+s}, s \geq 0\}$ and the tail sigma field

$$\mathcal{T} = \bigcap_{t \geq 0} \sigma\{X_{t+s}, s \geq 0\}.$$

Clearly, $\mathcal{J} \subset \mathcal{T}$.

Definition 1.16 *A stationary process $(X_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{A}, P)$ is called* **strongly mixing** *if for all $A, B \in \mathcal{D}$,*

$$I\!P(A \cap \vartheta_t^{-1} B) - I\!P(A) I\!P(B) \to 0 \text{ as } t \to \infty. \tag{1.11}$$

Strong mixing implies ergodicity as shows the following proposition.

Proposition 1.17 *Let X be strongly mixing. Then X is ergodic.*

Proof: Take an invariant set A and apply (1.11) to $A = B$. By invariance, $\vartheta_t^{-1}(A) = A$. Hence $I\!P(A) = I\!P(A)^2$, which implies the result. •

Proposition 1.18 *If P is trivial on $\mathcal{T}$, then the process is strong mixing, and hence ergodic.*

Proof: The family $(\mathcal{T}_t)_{t\geq 0}$, $\mathcal{T}_t = \sigma\{X_{t+s}, s \geq 0\}$, is a decreasing family of sub-sigma fields of $\mathcal{F}_\infty$. Now take two bounded, $\mathcal{D}-$measurable functions f and g and consider

$$\begin{aligned} \int_D f \cdot (g \circ \vartheta_t)\, dI\!P &= E[f(X) g(X_{t+\cdot})] = E\left[g(X_{t+\cdot}) E[f(X)|\mathcal{T}_t]\right] \\ &= E\Big[g(X_{t+\cdot}) \Big(E[f(X)|\mathcal{T}_t] - E[f(X)|\mathcal{T}]\Big)\Big] + E\left[g(X_{t+\cdot}) E[f(X)|\mathcal{T}]\right]. \end{aligned}$$

Since P is trivial on $\mathcal{T}$, $E[f(X)|\mathcal{T}]$ is constant $P-$almost surely and hence equals $E[f(X)]$ almost surely. Hence, the last expression in the above formula equals $E\left[g(X)\right] E[f(X)]$. Moreover, $\Big(E[f(X)|\mathcal{T}_t] - E[f(X)|\mathcal{T}]\Big)_{t\geq 0}$ is a bounded backwards martingale which tends to 0 as $t \to \infty$. This implies the result. •

The next theorem is probably the most interesting result in the context of Markov processes. Recall that the total variation distance of two probability measures P and Q defined on $(\Omega, \mathcal{A})$ is defined as

$$\|P - Q\|_{TV} = \sup_{A \in \mathcal{A}} |P(A) - Q(A)|.$$

Actually, the total variation distance is closely related to coupling, and if we succeed to construct a good coupling of the measures P and Q, then this gives a good upper bound on $\|P - Q\|_{TV}$, see also Section 2 below.

Theorem 1.19 *Suppose that X is strong Markov, with invariant probability measure μ. If*

$$\|P_t(x, \cdot) - \mu\|_{TV} \to 0 \text{ for all } x \text{ as } t \to \infty,$$

then P_x is trivial on $\mathcal{T}$ for all $x \in E$, in particular the process is strong mixing and ergodic.

Proof: It can be shown that $\|P_t(x, \cdot) - \mu\|_{TV} \to 0$ as $t \to \infty$ is equivalent to P_ν trivial on $\mathcal{T}$ for all initial measures ν. This is a classical result proven in Orey (1971). •

2 Classical recurrence conditions

Let now $(X_t)_{t\geq 0}$ be a strong Markov process. Suppose for simplicity that it takes values in $\mathbb{R}^d$, for some $d \geq 1$, although this does not really matter. Is the process recurrent, and how can we tell it is?

In the above question, we encounter two types of difficulties. First, the process is in continuous time, most classical results are stated for chains, i.e. for processes indexed by discrete time parameter. Second, the infinite non countable state space is difficult to deal with, the situation of finite or countable state space is much easier to handle.

The most straightforward way of going back to chains, when considering processes, is the following: Fix a suitable time parameter $t_* > 0$ and consider the chain $(X_{nt_*})_{n\geq 1}$ which consists of observing the process along a time grid with step size t_*. The transition operator of the chain is $P_{t_*}(x, dy)$, and if we can show that the sampled chain $(X_{nt_*})_{n\geq 1}$ is Harris with invariant measure μ and if μ is also invariant for the process (this does not seem to be a direct consequence of $\mu P_{t_*} = \mu$ for the fixed t_*), then obviously the process is Harris as well. This leads us to the following question:

Let $(X_n)_{n\geq 1}$ be a Markov chain with transition operator P, taking values in $(E, \mathcal{E})$. What are criteria for recurrence of X?

2.1 The case of a finite or countable state space in discrete time

Firstly, suppose that E is finite. Then there exist always invariant measures of the chain, possibly infinitely many. But suppose that the chain is **irreducible**, i.e. for all $x, y \in E$, there exists $n = n(x, y)$ such that $P_n(x, y) > 0$. Now let μ be any invariant measure of the chain. Then $\mu(\{x\}) > 0$ for any $x \in E$. Therefore, Harris-recurrence is equivalent to point recurrence in this case. In other words, we have to show that $T_a < \infty$ P_x-almost surely for any a and any $x \in E$, where $T_a = \inf\{n \geq 1 : X_n = a\}$ is the first return time to a.

For irreducible chains, the well-known theorem of Kac shows:

Theorem 2.1 Theorem of Kac: *Any irreducible chain taking values in a finite state space possesses a unique invariant probability measure μ given by*

$$\mu(\{x\}) = \frac{1}{E_x(T_x)}, x \in E. \tag{2.12}$$

In particular, $E_x(T_x) < \infty$, and the chain is positive recurrent.

Proof: The proof of the above theorem is surprisingly simple in the case of a state space having two elements and has been communicated to the authors by Antonio Galves. The proof for general finite state space is more complicated, and we refer the reader to any classical textbook on Markov chains for the proof.

Take a stationary version of the chain and suppose that $E = \{0, 1\}$. Let $p = P(X_n = 1) = \mu(\{1\})$ and $B = \bigcup_{n<0}\{X_n = 1\}$. Then

$$B = \bigcup_{m<0} \{X_m = 1, X_n = 0 \text{ for all } n = m + 1, \ldots - 1\}.$$

Hence

$$P(B) = \sum_{m<0} P(X_m = 1, X_n = 0 \text{ for all } n = m+1, \ldots - 1),$$

but by stationarity,

$$P(X_m = 1, X_n = 0 \text{ for all } n = m+1, \ldots - 1) = P(T_1 > |m| - 1 \big| X_0 = 1) \cdot p.$$

Therefore,

$$P(B) = p \sum_{m \geq 0} P(T_1 > m | X_0 = 1) = pE_1(T_1),$$

which implies that

$$\mu(\{1\}) = p = \frac{P(B)}{E_1(T_1)}.$$

We will show that $P(B) = 1$. It is at this point that the restriction to a state space having two elements is essential. This is for the following reason: $B^c = \{X_n = 0 \text{ for all } n < 0\}$. But $P(0,0) < 1$. Then, stationarity and invariance of B^c imply that

$$P(B^c) = P(B^c)P(X_{-n} = 0, \ldots, X_{-1} = 0 | X_{-n-1} = 0) = P(B^c)P(0,0)^n \to 0$$

as $n \to \infty$. •

In the case of an infinite but countable state space, a simple sufficient condition for recurrence is the following condition. It states that a uniform coupling constant is strictly positive.

$$\beta(P) = \sum_{y \in E} \inf_{x \in E} P(x, y) > 0. \tag{2.13}$$

The constant $\beta(P)$ is often called **ergodicity coefficient**, see for instance Ferrari and Galves (2000).

Theorem 2.2 *Suppose that $\beta(P) > 0$. Then the following assertions hold true.*

1. *The chain possesses a unique invariant measure μ which is a finite measure (up to multiplication with a constant).*
2. *For any pair of starting values $x, y \in E$, we have*

$$\|P_n(x, \cdot) - P_n(y, \cdot)\|_{TV} \leq (1 - \beta(P))^n. \tag{2.14}$$

Hence we have ergodicity (and even strong mixing) with exponential (geometrical) speed of convergence.

The proof of the above theorem relies on the construction of a suitable coupling. Let us recall the definition of a coupling and the relation with the total variation distance.

Definition 2.3 *A coupling of two probability measures P and Q is any pair of random variables (X, Y) defined on the same probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ such that $X \sim P$ and $Y \sim Q$.*

The important link to the total variation distance is given by the following lemma:

Lemma 2.4

$$\|P - Q\|_{TV} = \inf\{\tilde{P}(X \neq Y) : (X, Y) \textit{ is a coupling of } P \textit{ and } Q\}.$$

A proof of this lemma can for example be found in the very instructive book by Levin, Peres and Wilmer (2009). In the following we only give a proof of the upper bound.

Proof We have for any coupling (X, Y) of P and Q and for any measurable set A,

$$\begin{aligned} P(A) - Q(A) &= \tilde{P}(X \in A) - \tilde{P}(Y \in A) \\ &\leq \tilde{P}(X \in A, Y \notin A) \\ &\leq \tilde{P}(X \neq Y). \end{aligned}$$

Hence we have that

$$\|P - Q\|_{TV} = \sup_{A \in \mathcal{A}} P(A) - Q(A) \leq \tilde{P}(X \neq Y) \tag{2.15}$$

for any coupling. •

As a consequence of the above proof, we see that it is sufficient to construct a **good** coupling of P and Q in order to control the total variation distance, this coupling does not need to be the optimal one that achieves the distance.

We now give the proof of Theorem 2.2. It follows ideas exposed in Ferrari and Galves (2000).

Proof of Theorem 2.2: We construct a coupling of the chains $(X_n^x)_{n \geq 1}$ and $(X_n^y)_{n \geq 1}$ starting from x and from y, respectively. Inspired by (2.15), the aim of this coupling is to force both chains to be as close as possible.

The coupling will be constructed with the help of a sequence of i.i.d. random variables $(U_n)_{n \geq 1}$ which are uniformly distributed on $[0, 1]$. Put

$$\nu(z) = \frac{\inf_x P(x, z)}{\beta(P)}, \; z \in E.$$

By definition, ν is a probability on $(E, \mathcal{E})$ and $P(x, z) \geq \beta(P)\nu(z)$ for all $x \in E$.

We put $X_0^x = x$ and $X_0^y = y$. Then we proceed recursively as follows. On $\{U_1 \leq \beta(P)\}$, we choose

$$X_1^x = X_1^y \sim \nu(dz).$$

Moreover, on $\{U_1 > \beta(P)\}$, we choose

$$X_1^x \sim \frac{P(x, dz) - \beta(P)\nu(dz)}{1 - \beta(P)},$$

and independently of this choice,

$$X_1^y \sim \frac{P(y, dz) - \beta(P)\nu(dz)}{1 - \beta(P)}.$$

It is easy to check, integrating over the possible values of U_1, that $\mathcal{L}(X_1^x) = P(x, dz)$ and $\mathcal{L}(X_1^y) = P(y, dz)$. Indeed,

$$P(X_1^x = z) = P(X_1^x = z | U_1 \leq \beta(P))\beta(P) + P(X_1^x = z | U_1 > \beta(P))(1 - \beta(P))$$
$$= \nu(z)\beta(P) + \left[\frac{P(x,z) - \beta(P)\nu(z)}{1 - \beta(P)}\right](1 - \beta(P)) = P(x, z).$$

If $X_1^x \neq X_1^y$, we continue in the same manner, using U_2 for the choice of X_2^x and X_2^y, and so on. We stop at the first time $T_{coup} = \inf\{n : X_n^x = X_n^y\}$. From that time on, we let both trajectories be sticked together, evolving according to the Markov transition kernel P without further coupling device (the above construction is only intended to make the two trajectories meet each other as soon as possible).

It is evident that $T_{coup} \leq \tau := \inf\{n : U_n \leq \beta(P)\}$, hence for any $A \in \mathcal{E}$,

$$\begin{aligned} |P_n(x, A) - P_n(y, A)| &= |E(1_A(X_n^x) - 1_A(X_n^y))| \\ &\leq \tilde{P}(X_n^x \neq X_n^y) \\ &\leq \tilde{P}(\tau > n) = (1 - \beta(P))^n. \end{aligned}$$

This shows (2.14).

The above construction proves moreover that for any $x \in E$, $\mu(\cdot) := \lim_n P_n(x, \cdot)$ exists and does not depend on the starting point. This is seen as follows. Fix n and let $T^{(n)} = \max\{k \leq n : U_k \leq \beta(P)\}$, where we put $\max \emptyset := 0$. On $\{T^{(n)} = n - k\}$, we have $P_n(x, \cdot) = \int \nu(dz) P_k(z, \cdot)$. Moreover, $\tilde{P}(T^{(n)} = n - k) = \beta(P)(1 - \beta(P))^k$. Therefore,

$$P_n(x, \cdot) = \sum_{k=0}^{n-1} \beta(P)(1 - \beta(P))^k \int \nu(dz) P_k(z, \cdot) + (1 - \beta(P))^n P_n(x, \cdot)$$
$$\to \sum_{k=0}^{\infty} \beta(P)(1 - \beta(P))^k \int \nu(dz) P_k(z, \cdot) =: \mu(\cdot). \quad (2.16)$$

Being the limit of $P_n(x, \cdot)$, the measure μ is an invariant probability of the chain. Integrating (2.14) against $\mu(dy)$, we obtain that

$$\|P_n(x, \cdot) - \mu(\cdot)\|_{TV} \leq (1 - \beta(P))^n,$$

which implies the uniqueness of the invariant measure. Hence the chain is mixing, thus ergodic and Harris recurrent.

It is also worthwhile to notice that we can prove the Harris recurrence directly, without using any results on mixing. For that sake, we use regeneration arguments similar to those in example 1.5. Let

$$R_1 = \inf\{k \geq 1 : U_k \leq \beta(P)\}, \ldots, R_{n+1} = \inf\{k > R_n : U_k \leq \beta(P)\}.$$

This is a sequence of regeneration times for the chain, and by its structure given in (2.16),

$$\mu(A) = \tilde{E} \sum_{k=R_1}^{R_2 - 1} 1_A(X_k^x).$$

Now we conclude exactly as in example 1.5. Using the strong law of large numbers, $\mu(A) > 0$ implies that the chain visits A i.o. P_x-almost surely for any x, and this concludes the proof. •

2.2 General state space, discrete time

The coupling presented in the proof of Theorem 2.2 is known as **Doeblin coupling**. It is widely used in order to control the speed of convergence to equilibrium of Markov chains and/or processes. In fact, the above ideas can be extended to general state space, by replacing the condition (2.13) by the

Doeblin condition: There exists $\alpha \in]0,1]$ and a probability measure ν on $(E, \mathcal{E})$ such that for all $x \in E$,

$$P(x, dy) \geq \alpha \nu(dy), \tag{2.17}$$

or equivalently, $P(x, A) \geq \alpha \nu(A)$ for all $A \in \mathcal{E}$.

Having a strictly positive ergodicity coefficient $\beta(P) > 0$ in case of countable state space is an example of a situation where the Doeblin condition holds: Take $\alpha = \beta(P)$ and $\nu(y) = \frac{\inf_x P(x,y)}{\beta(P)}$.

Corollary 2.5 *Under the Doeblin condition, the chain is geometrically ergodic, Harris-recurrent, and*

$$\|P_n(x, \cdot) - P_n(y, \cdot)\|_{TV} \leq (1-\alpha)^n,$$

in particular, $\|P_n(x, \cdot) - \mu(\cdot)\|_{TV} \leq (1-\alpha)^n$.

2.3 Coming back to continuous time

Let us come back to continuous time, by using the sampled chain $(X_{nt_*})_{n \geq 1}$. A direct translation of the Doeblin condition applied to the sampled chain gives the following sufficient condition for recurrence.

Doeblin condition in continuous time: There exists $t_* > 0$, $\alpha \in]0,1]$ and a probability measure ν on $(E, \mathcal{E})$ such that for all $x \in E$,

$$P_{t_*}(x, dy) \geq \alpha \nu(dy).$$

Remark 2.6 *Note that the Doeblin condition implies that for any* $x \in E$,

$$\nu \ll P_{t_*}(x, dy).$$

Corollary 2.7 *Suppose that the Doeblin condition in continuous time holds. Then the process is recurrent in the sense of Harris,* μ *is a probability measure, and*

$$\|P_t(x, \cdot) - \mu(\cdot)\|_{TV} \leq (1-\alpha)^{[t/t_*]}, \tag{2.18}$$

where $[\cdot]$ *denotes the integer part of the number.*

Proof: Take two different starting values x and y and construct the Doeblin coupling of the chains $(X^x_{nt_*})_n$ and $(X^y_{nt_*})_n$. Let $t > 0$, write $t = kt_* + r$, $0 \leq r < t_*$. Then for any measurable A, with μ the invariant probability measure of the chain,

$$P_t(x, A) = \int P_r(x, dy) P_{kt_*}(y, A); \quad \mu(A) = \int P_r(x, dy) \mu(A). \tag{2.19}$$

By Doeblin coupling of the chains, we have that

$$P_{kt_*}(y, A) - \mu(A) \leq (1-\alpha)^k$$

for any A, and integration $P_r(x, dy)$ yields

$$P_t(x, A) - \mu(A) \leq (1-\alpha)^k.$$

This gives the result since $k = [t/t_*]$. Finally, the Harris-recurrence of the process follows immediately from the Harris-recurrence of the chain. •

The exponential speed of convergence to equilibrium in (2.18) holds true uniformly with respect to all starting values x. This uniform behavior is actually extremely restrictive when dealing with general state space and can in general only be hoped for when looking at processes taking values in compact state space. In order to clarify this point let us consider uniformly elliptic diffusions in dimension d.

Example 2.8 *Consider a $d-$dimensional uniformly elliptic symmetric diffusion*

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

W $d-$dimensional Brownian motion, $a = \sigma\sigma^*$, $b^i = \sum_{j=1}^d \frac{\partial a_{ji}}{\partial x_j}$. Suppose that the coefficients a_{ij} are bounded and smooth and that there exists $\lambda_- > 0$ such that

$$< a(x)\xi, \xi > \geq \lambda_- \|\xi\|^2$$

for all $x, \xi \in I\!\!R^d$. This is the condition of **uniform ellipticity**.

Under these conditions, the transition operator $P_t(x, dy)$ admits a Lebesgue density $p_t(x, y)$ and Aronson's bounds are known to hold true, see e.g. Aronson (1968), Norris and Stroock (1990): we have upper and lower bounds

$$p_t(x, y) \leq c_1 t^{-d/2} \exp\left(-c_2 \frac{\|x-y\|^2}{t}\right),$$

$$p_t(x, y) \geq c_3 t^{-d/2} \exp\left(-c_4 \frac{\|x-y\|^2}{t}\right).$$

Checking Doeblin's condition in in this frame means that we have to find lower bounds of the transition density. In other words, we have to find suitable values of y such that

$$\inf_{x \in I\!\!R^d} p_{t_*}(x, y) > 0.$$

Using the upper bound of Aronson, we see that for any fixed y,

$$\inf_{x \in I\!\!R^d} p_{t_*}(x, y) \leq \inf_{x \in I\!\!R^d} c_1 (t_*)^{-d/2} \exp\left(-c_2 \frac{\|x-y\|^2}{t_*}\right) = 0,$$

which is simply seen by letting $\|x - y\| \to \infty$.

Let us resume the above discussion: Being in infinite state space such as $I\!R^d$ implies in general that the Doeblin condition can not be satisfied. The ingenious idea of Harris was now to replace the Doeblin condition by a localized one.

Localized Doeblin condition: There exists a measurable set C such that

$$P_{t_*}(x, dy) \geq \alpha \nu(dy) \quad \text{ for all } x \in C. \tag{2.20}$$

Such a set C is often called **petite set**, see the vast literature by Meyn and Tweedie and coauthors, we refer to the very complete monograph of Meyn and Tweedie (2009).

It is clear that in the case of uniformly elliptic symmetric diffusions as considered in the above example, **any compact set C does the job.** Actually this situation is quite typical, and compact sets will be often "petite" sets. The Doeblin coupling works whenever the process starts in C or comes back to the set C. But it is evident that we have to control excursions out of C, or equivalently, return times to C. To resume the situation, sufficient conditions implying Harris recurrence for Markov processes taking values in general state space are twofold:

1. First, find a good set C such that the local Doeblin condition (2.20) holds. This amounts of finding lower bounds for transition densities and sets where these densities are strictly positive. In order to establish the strict positivity, in case of diffusions, one can for example apply control arguments.
2. Find a way of controlling the return times to C, i.e. give precise estimations of $E_x(T_C)$, depending on the starting point, where $T_C = \inf\{t : X_t \in C\}$. This is often done using **Lyapunov functions**. We will come back to this point later.

It is easy to check the above points in dimension one. We will expose these ideas in the following section. Before doing so, let us close this section with a final remark.

Remark 2.9 *In the statement of the local Doeblin condition (2.20), the choice of the time parameter t_* is somewhat arbitrary. Probably the most natural way to make a link between chains and processes is to use the so called* **resolvant chain***: Let $(N_t)_{t\geq 0}$ be a standard Poisson process of rate 1, independent of the process X, and write $T_1, T_2, \ldots, T_n, \ldots$ for its jump times. Then $\bar{X} := (X_{T_n})_{n\geq 1}$ is a Markov chain, with transition operator*

$$U^1(x, dy) = \int_0^\infty e^{-t} P_t(x, dy) dt.$$

It is well known, see e.g. Theorem 1.4 of Höpfner and Löcherbach (2003), that $\bar{X}$ is Harris recurrent with invariant measure μ if and only if the same is true for the process X. Both $\bar{X}$ and X have the same invariant measure. So it is natural to state the localized Doeblin condition rather in terms of U^1 than in terms of P_{t_*} :

$$U^1(x, dy) \geq \alpha 1_C(x) \nu(dy). \tag{2.21}$$

To resume the situation: (2.21) together with a control of the return times to C implies Harris recurrence of the process. On the contrary, we know also that Harris recurrence implies the existence of a set C having $\mu(C) > 0$ such that (2.21) holds, see e.g. Proposition 6.7 of Höpfner and Löcherbach (2003).

We are now going to realize points 1. and 2. of the above program in the case of one-dimensional elliptic diffusions.

3 The case of one-dimensional diffusions

Let X be solution of the following stochastic differential equation in dimension one

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{3.22}$$

where W is one-dimensional Brownian motion. We suppose that $\sigma^2(x) > 0$ for all x and that b and σ are locally Lipschitz and of linear growth. Then there exists a unique strong non-exploding solution to (3.22).

Let us start by collecting some well known facts. We introduce

$$s(x) = \exp\left(-2\int_0^x \frac{b(u)}{\sigma^2(u)}du\right)$$

and define the **scale function**

$$S(x) = \int_0^x s(u)du.$$

The scale function satisfies

$$AS = bS' + \frac{1}{2}\sigma^2 S'' = 0$$

and provides a simple recurrence criterion.

Proposition 3.1 *The process X is recurrent in the sense of Harris if and only if S is a space transform, i.e. $S(+\infty) = +\infty$ and $S(-\infty) = -\infty$.*

The above result is classical, see e.g. Khasminskii (1980) and (2012), Revuz and Yor (2005), chapter VII. We give the proof for the convenience of the reader. The proof uses the speed measure density $m(x) = \frac{2}{\sigma^2(x)s(x)}$. The speed measure $m(x)dx$ equals the invariant measure of the diffusion (up to multiplication with a constant), as can be shown easily, cf. Revuz and Yor (2005).

Proof: We show first that Harris-recurrence implies that S is a space transform. We have for $a < x < b$,

$$P_x(T_a < T_b) = \frac{S(x) - S(b)}{S(a) - S(b)}, \tag{3.23}$$

where $T_a = \inf\{t : X_t = a\}$. This is easily deduced from $AS = 0$, applying Itô's formula to $S(X_t)$ and using the stopping rule with the stopping time $T_a \wedge T_b$.

Now let for instance $b \to \infty$. Since X is non-exploding, $T_b \to \infty$. Hence

$$P_x(T_a < \infty) = \frac{S(+\infty) - S(x)}{S(+\infty) - S(a)}.$$

By Harris-recurrence, $P_x(T_a < \infty) = 1$ for all x. This can be seen as follows. Take for instance $A = [a, +\infty[$. Since $\sigma^2(x) > 0$ for all x, $m \sim \lambda$ and $m(A) > 0$. This implies that $P_x(T_A < \infty) = 1$ for all x, where $T_A = \inf\{t : X_t \in A\}$. The continuity of the sample paths implies that for $x < a$, $P_x(T_A < \infty) = P_x(T_a < \infty)$. Hence for all $x < a$, $P_x(T_a < \infty) = 1$. The same argument applied to $]-\infty, a]$ shows that $P_x(T_a < \infty) = 1$ for all $x > a$.

Coming back to (3.23), this implies that $S(+\infty) = +\infty$. The other assertion, $S(-\infty) = -\infty$, follows from letting $a \to -\infty$.

Now we prove that S being a space transform implies that the process is recurrent. Put $Y_t = S(X_t)$. Applying Itô's formula, we obtain $dY_t = s(X_t)\sigma(X_t)dW_t$. Hence Y is a locally square integrable continuous martingale, and $< Y >_t = \int_0^t s^2(X_u)\sigma^2(X_u)du$. We show that $< Y >_\infty = +\infty$ almost surely. Indeed, using Lépingle (1978), on $\{< Y >_\infty < +\infty\}$, $\lim_{t\to\infty} Y_t$ exists and is finite almost surely. Now we use that S, as a space transform, is a bijection from $I\!R$ onto $I\!R$. Hence $X_t = S^{-1}(Y_t)$. As a consequence, on the set $\{< Y >_\infty < +\infty\}$, $\lim_{t\to\infty} X_t$ exists and is finite as well. This implies that on $\{< Y >_\infty < +\infty\}$, $\inf_{t\to\infty} s^2(X_t)\sigma^2(X_t) > 0$, hence $\int_0^\infty s^2(X_t)\sigma^2(X_t)dt = \infty$, which is a contradiction.

Being a locally square integrable martingale, Y can be represented as time-changed Brownian motion $Y_t = B_{<Y>_t}$, where B is some one-dimensional Brownian motion. B is recurrent and visits any fixed point infinitely often. Since $[0, < Y >_\infty [= [0, \infty[$, the process Y visits the same points as B. In particular, $\limsup_{t\to\infty} 1_{\{0\}}(Y_t) = 1$ almost surely. But $Y_t = 0$ if and only if $X_t = 0$, since $S(0) = 0$. This implies the result. •

To resume: If S is a space transform, then the diffusion is recurrent. This means that we have completely solved the two points stated in the last paragraph of the last section. The good set C is for example $C = \{0\}$. Being a singleton, it is evident that C is "petite", and by recurrence, we know that X comes back to $C = \{0\}$ almost surely in finite time. It is for this reason that the situation in dimension one is exceptionally nice, since we only have to control hitting times of points and since hitting times of points induce regeneration times.

Before proceeding, let us us show that the invariant measure of the process is indeed given by the speed measure $m(x)dx$.

Proposition 3.2 *Under the conditions of Proposition 3.1, the invariant measure of the process X is given by the speed measure $m(x)dx$, up to multiplication with a constant.*

Proof Fix two bounded, positive and measurable functions f and g and write $Y_t = S(X_t)$. Then Y is a local martingale, having diffusion coefficient $\tilde\sigma = (s \cdot \sigma) \circ S^{-1}$. Moreover,

$$\frac{\int_0^t f(X_s)ds}{\int_0^t s^2 \cdot \sigma^2(X_s)ds} = \frac{1}{< X >_t} \int_0^t f \circ S^{-1}(Y_t)dt,$$

which, as $t \to \infty$, behaves as

$$\frac{1}{t} \int_0^{\tau_t} f \circ S^{-1}(Y_s)ds,$$

where

$$\tau_t = \inf\{s \geq 0 : < X >_s = t\}.$$

Using that $Y_s = W_{<X>_s}$ is a time changed Brownian motion, we obtain that

$$\frac{1}{t} \int_0^{\tau_t} f \circ S^{-1}(Y_s)ds = \frac{1}{t} \int_0^t \frac{f \circ S^{-1}}{\sigma^2 \cdot s^2 \circ S^{-1}}(W_u)du,$$

and then, using the ratio limit theorem for one-dimensional Brownian motion and the change of variables $z = S^{-1}(x)$,

$$\lim_{t\to\infty} \frac{\int_0^t f(X_s)ds}{\int_0^t g(X_s)ds} = \lim_{t\to\infty} \frac{\int_0^t \frac{f\circ S^{-1}}{\sigma^2\cdot s^2\circ S^{-1}}(W_u)du}{\int_0^t \frac{g\circ S^{-1}}{\sigma^2\cdot s^2\circ S^{-1}}(W_u)du} = \frac{\int_{I\!R} \frac{f\circ S^{-1}}{\sigma^2\cdot s^2\circ S^{-1}}(x)dx}{\int_{I\!R} \frac{g\circ S^{-1}}{\sigma^2\cdot s^2\circ S^{-1}}(x)dx}$$
$$= \frac{\int_{I\!R} \frac{f}{\sigma^2\cdot s}(z)dz}{\int_{I\!R} \frac{g}{\sigma^2\cdot s}(z)dz} = \frac{\int_{I\!R} f(z)m(z)dz}{\int_{I\!R} g(z)m(z)dz}.$$

If m has finite total mass, then we may take $g \equiv 1$ in order to conclude that

$$\frac{1}{t}\int_0^t f(X_s)ds \to \frac{1}{m(I\!R)}\int f(x)m(x)dx \quad \text{as} \quad t \to \infty,$$

which implies the assertion. •

As in the preceding chapters, the letter μ will always be reserved for the invariant measure of the diffusion. Due to the preceeding proposition, up to multiplication with a constant, μ equals the speed measure $m(x)dx$. If μ is a finite measure, then we always renormalize it to be a probability. The main question is now: Is there any convergence to equilibrium in the sense that $P_t(x, \cdot) \to \mu$ as $t \to \infty$? And at which speed does the convergence to equilibrium hold true? (Do not forget that Brownian motion itself satisfies the above conditions: $S(x) = x$ is a space transform, however there is no convergence to equilibrium as we have discussed earlier, see Remark 1.8.)

We follow Löcherbach, Loukianova and Loukianov (2011) in order to answer these questions. Fix two points $a < b$ and introduce

$$S_1 = \inf\{t : X_t = b\}, R_1 = \inf\{t > S_1 : X_t = a\},$$

and

$$S_{n+1} = \inf\{t > R_n : X_t = b\}, R_{n+1} = \inf\{t > S_{n+1} : X_t = a\}, n \geq 1.$$

We call the sequence of times $R_1 \leq R_2 \leq \ldots$ **regeneration times** of X. They satisfy

1. $R_n \uparrow \infty$ as $n \to \infty$.
2. $R_{n+1} = R_n + R_1 \circ \vartheta_{R_n}$, for all $n \geq 1$.
3. $(R_{n+1} - R_n)_{n \geq 1}$ are i.i.d.
4. $X_{R_n + \cdot}$ is independent of $\mathcal{F}_{R_n -}$ and the path segments $(X_{R_n+s}, 0 \leq s \leq R_{n+1} - R_n)_{n \geq 1}$ are i.i.d. trajectories.

Definition 3.3 (Order of recurrence) *We say that the process is polynomially recurrent of order p if there exists some a such that $E_x T_a^p < \infty$ for all x. We say that the process is exponentially recurrent if there exist some a and some $\lambda > 0$ such that $E_x e^{\lambda T_a} < \infty$ for all x.*

In the sequel we will be mainly interested in the polynomially recurrent case. We will see that it is sufficient to suppose that there exists some x with $E_x T_a^p + E_a T_x^p < \infty$. This is the "all-or-none" property of Theorem 4.5 of Löcherbach, Loukianova and Loukianov (2011). For the exponentially recurrent case, we refer to Loukianova, Loukianov and Song (2011).

3.1 Kac's moment formula determines the speed of convergence to equilibrium

For $a < b$, let

$$T_{a,b} = \inf\{t \geq 0;\ X_t \notin [a, b]\}.$$

Consider X_t with $X_0 = x \in [a,b]$ and let $G_{[a,b]}$ be the Green function associated to the stopping time $T_{a,b}$, defined by

$$G_{[a,b]}(x,\xi) = \begin{cases} \frac{(S(x)-S(a))(S(b)-S(\xi))}{S(b)-S(a)}, & a \le x \le \xi \le b \\ \frac{(S(b)-S(x))(S(\xi)-S(a))}{S(b)-S(a)}, & a \le \xi \le x \le b \\ 0, & \text{otherwise.} \end{cases}$$

We have the following theorem.

Theorem 3.4 (Kac's moment formula) *Let $f : I\!R \to I\!R$ be such that the function $x \mapsto E_x f'(T_{a,b})$ is continuous on $[a,b]$. Then*

$$E_x f(T_{a,b}) = f(0) + \int G_{[a,b]}(x,\xi) E_\xi f'(T_{a,b}) m(\xi) d\xi. \tag{3.24}$$

Somehow this result means that in the one dimensional case, the question of moments of hitting times is explicitly solved since we have an explicit representation of the Green function. This is very specific of the one dimensional frame and does not hold any more in higher dimensions.

Theorem 3.4 is classical and can be found for example in Revuz and Yor (2005) or Fitzsimmons and Pitman (1999). The above cited version is Theorem 4.1 of Löcherbach, Loukianova and Loukianov (2011). We quote the proof from there:

Proof: Put $u(x) = \int G_{[a,b]}(x,\xi) E_\xi f'(T_{a,b}) m(\xi) d\xi$. Verify (this is easy and follows from $AS = 0$) that

$$Au(x) = -E_x f'(T_{a,b}), x \in]a,b[, \quad u(a) = u(b) = 0.$$

Applying Itô's formula we obtain for $x \in]a,b[$,

$$u(X_t) = u(x) - \int_0^t E_{X_s} f'(T_{a,b}) ds + M_t,$$

where $M_t = \int_0^t u'(X_s)\sigma(X_s) dW_s$ is a locally square integrable martingale such that $M^{T_{a,b}}$ is uniformly integrable. The stopping rule applied to $T_{a,b}$ gives

$$E_x(u(X_{T_{a,b}})) = u(x) - E_x \int_0^{T_{a,b}} E_{X_s} f'(T_{a,b}) ds,$$

thus

$$\begin{aligned} u(x) &= E_x \int_0^{T_{a,b}} E_{X_s} f'(T_{a,b}) ds \\ &= \int_0^\infty E_x \left(1_{\{s < T_{a,b}\}} E_x[f'(T_{a,b} \circ \vartheta_s) | \mathcal{F}_s] \right) ds \\ &= \int_0^\infty E_x \left(1_{\{s < T_{a,b}\}} f'(T_{a,b} - s) \right) ds \\ &= E_x \int_0^{T_{a,b}} f'(T_{a,b} - s) ds = E_x f(T_{a,b}) - f(0), \end{aligned}$$

and this concludes the proof. •

Letting $a \to -\infty$ and considering pairs (x,b) with $x < b$ or letting $b \to \infty$ and considering pairs (x,a) with $x > a$ leads to the following corollary.

Corollary 3.5 *Under the conditions of Theorem 3.4 the following holds.*

1. *For $x < b$, we have*

$$E_x f(T_b) = f(0) + (S(b) - S(x)) \int_{-\infty}^{x} E_\xi f'(T_b) m(\xi) d\xi + \int_x^b (S(b) - S(\xi)) E_\xi f'(T_b) m(\xi) d\xi. \quad (3.25)$$

In particular, $E_x f(T_b) < \infty$ if and only if $\xi \mapsto E_\xi f'(T_b) \in L^1(m(\xi) 1_{]-\infty,x]}(\xi) d\xi)$.

2. *For $x > a$, we have*

$$E_x f(T_a) = f(0) + (S(x) - S(a)) \int_x^{\infty} E_\xi f'(T_a) m(\xi) d\xi + \int_a^x (S(\xi) - S(a)) E_\xi f'(T_a) m(\xi) d\xi. \quad (3.26)$$

In particular, $E_x f(T_a) < \infty$ if and only if $\xi \mapsto E_\xi f'(T_a) \in L^1(m(\xi) 1_{[x,\infty[}(\xi) d\xi)$.

Remark 3.6 *Taking $f(x) = x^p$, we obtain that*

$$E_x T_b^p = p \left((S(b) - S(x)) \int_{-\infty}^{x} E_\xi T_b^{p-1} m(\xi) d\xi + \int_x^b (S(b) - S(\xi)) E_\xi T_b^{p-1} m(\xi) d\xi \right), \quad (3.27)$$

for $x < b$. Notice that this shows that $E_x T_b^p < \infty$ for one pair (x, b) with $x < b$ if and only if it holds for all such ordered pairs, see Theorem 4.5 of Löcherbach, Loukianova and Loukianov (2011).

Analogously, we have

$$E_x T_a^p = p \left((S(x) - S(a)) \int_x^{\infty} E_\xi T_a^{p-1} m(\xi) d\xi + \int_a^x (S(\xi) - S(a)) E_\xi T_a^{p-1} m(\xi) d\xi \right), \quad (3.28)$$

for $x > a$, which shows that $E_x T_a^p < \infty$ for one pair (x, a) with $x > a$ if and only if it holds for all such ordered pairs.

3.2 Veretennikov's drift conditions

We use Kac's moment formula in order to obtain an explicit control of $E_x T_a^p$ under suitable conditions on drift and diffusion coefficient of the process. These conditions are very much the same as those given in Veretennikov (1997). Veretennikov is mostly interested in obtaining polynomial ergodicity for diffusions. We will expose his ideas in Subsection 3.3 below. But before coming to polynomial ergodicity, let us first deal with the question of order of recurrence of the diffusion, under the following condition:

Drift condition: There exist $M_0 > 0$, $\lambda_- > 0$, $\gamma < 1$ and $r > 0$ such that

$$\sigma^2(x) \geq \lambda_- x^{2\gamma} \text{ and } -\frac{x b(x)}{\sigma^2(x)} \geq r, \text{ for all } |x| \geq M_0. \quad (3.29)$$

Under the above condition, the order of recurrence will be determined by $r + \gamma$. In most of the examples, $\gamma = 0$, thus r is the important parameter. (3.29) implies the following decay for the speed density

$$m(\xi) \leq C \xi^{-2r-2\gamma},$$

and if we put for some fixed $x > 0$,

$$p^* = \sup\{p : \int_x^\infty \xi^p m(\xi) d\xi + \int_{-\infty}^{-x} \xi^p m(\xi) d\xi < \infty\},$$

then clearly $p^* \geq 2r + 2\gamma - 1$.

We cite the following theorem from Löcherbach, Loukianova and Loukianov (2011).

Theorem 3.7 *Grant condition (3.29). Suppose that $2r+2\gamma > 1$. Then for any $m < \frac{1}{2}\frac{2r+1}{1-\gamma}$ and for all $x > a > M_0$ or $x < a < -M_0$,*

$$E_x T_a^m \leq C x^{2m(1-\gamma)}.$$

In particular, the process is recurrent of order m for any $m < \frac{1}{2}\frac{2r+1}{1-\gamma}$.

Proof: The proof follows from an iterated application of Kac's moment formula. Let $x > a$. Notice that for $\xi > x$, condition (3.29) yields

$$\frac{s(x)}{s(\xi)} \leq \exp\left(-2r \int_x^\xi \frac{1}{u} du\right) = \left(\frac{x}{\xi}\right)^{2r}.$$

This, together with (3.26), implies

$$\begin{aligned} \frac{d}{dx} E_x T_a &= \frac{d}{dx}\left((S(x) - S(a)) \int_x^\infty m(\xi) d\xi + \int_a^x (S(\xi) - S(a)) m(\xi) d\xi\right) \\ &= s(x) \int_x^\infty m(\xi) d\xi \leq C x^{2r} \int_x^\infty \xi^{-2r-2\gamma} d\xi \\ &\leq C x^{1-2\gamma}. \end{aligned}$$

This implies $E_x T_a \leq C x^{2-2\gamma}$. An inductive procedure, based on (3.27), allows then to deduce that for all $m \in I\!N$,

$$E_x T_a^m \leq C x^{2m(1-\gamma)},$$

as long as $\xi^{2(m-1)(1-\gamma)}$ is integrable with respect to $m(\xi)d\xi$, which is implied by

$$2(m-1)(1-\gamma) < 2r + 2\gamma - 1 \Leftrightarrow m < \frac{1}{2}\frac{2r+1}{1-\gamma}.$$

For non-integer m, let $\alpha = m - [m]$ and start with

$$E_x T_a^\alpha \leq (E_x T_a)^\alpha \leq C^\alpha x^{2\alpha(1-\gamma)},$$

and then iterate $[m]$ times in order to obtain the desired result. •

In order to give an application of the above results, let us show how the knowledge of the order of recurrence enables us to give a control of the speed of convergence in the ergodic theorem. For that sake, let $f \in L^1(\mu)$ be a bounded function (recall that $\mu = m(\xi)d\xi$ up to multiplication with a constant). Then

$$\frac{1}{t}\int_0^t f(X_s) ds \to \mu(f) \text{ as } t \to \infty,$$

and thus

$$P_x\left(|\frac{1}{t}\int_0^t f(X_s)ds - \mu(f)| \geq \varepsilon\right) \leq r(t)\Phi(f) \tag{3.30}$$

for all $x \in \mathbb{R}$, where $r(t) \to 0$ is a rate function and $\Phi(f)$ some power of $\|f\|_\infty$ that we define later. An inequality of the type (3.30) is called **deviation inequality**. One main question is to determine the speed of convergence at which $r(t)$ tends to 0 as t approaches ∞.

The main idea is as follows:

$$\frac{1}{t}\int_0^t f(X_s)ds \sim \frac{1}{t}\sum_{n=1}^{N_t} \xi_n,$$

where $\xi_n = \int_{R_n}^{R_{n+1}} f(X_s)ds$ and $N_t = \sup\{n : R_n \leq t\}$ is the number of regeneration events before time t. The $(\xi_n)_n$ are i.i.d. and have (at least) the same number of moments as a typical regeneration epoch $R_2 - R_1$, due to the boundedness of the function f. It can be shown that if the process is recurrent of order $p > 1$, then

$$r(t) = t^{-(p-1)} \text{ and } \Phi(f) = \begin{cases} \|f\|_\infty^{2(p-1)}, & \text{if } p \geq 2, \\ \|f\|_\infty^p, & \text{if } 1 < p < 2. \end{cases}$$

(See e.g. Corollary 5.9 of Löcherbach and Loukianova and Loukianov (2011) or Theorem 5.2 of Löcherbach and Loukianova (2012)). The idea of the proof is to use a maximal inequality for sums of i.i.d. centered random variables having a moment of order p. If $p \geq 2$, we use a Fuk-Nagaev inequality (see e.g. Rio (2000)), for $p < 2$, we use Burkholder-Davis-Gundy's inequality.

3.3 Polynomial ergodicity for one-dimensional diffusions à la Veretennikov

It is natural to ask whether under condition (3.29) also explicit bounds for the convergence rate to equilibrium with respect to the total variation distance can be given. More precisely, we are seeking for a control of $\|P_t(x,\cdot) - \mu\|_{TV}$ or of $\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV}$. In some exceptional situations, this control can be done explicitly, without any further work. Recall the recurrent Ornstein Uhlenbeck process of example 1.5. Then $P_t(x, dy) = \mathcal{N}(e^{-at}x, \frac{1}{2a}(1 - e^{-2at}))$ and $\mu = \mathcal{N}(0, \frac{1}{2a})$. We use Pinsker's inequality

$$2\|P - Q\|_{TV}^2 \leq D(P\|Q), \tag{3.31}$$

where $D(P\|Q)$ is the Kullback-Leibler distance. It is known that for $P = \mathcal{N}(\mu_1, \sigma_1^2)$ and $Q = \mathcal{N}(\mu_2, \sigma_2^2)$,

$$D(P\|Q) = \frac{(\mu_1 - \mu_2)^2}{2\sigma_2^2} + \frac{1}{2}\left(\frac{\sigma_1^2}{\sigma_2^2} - 1 - \ln\frac{\sigma_1^2}{\sigma_2^2}\right).$$

Applying this to our context, we obtain that

$$2\|P_t(x,\cdot) - \mu\|_{TV}^2 \leq ax^2e^{-2at} - \frac{1}{2}\left(e^{-2at} + \ln(1 - e^{-2at})\right)$$

whence

$$\|P_t(x,\cdot) - \mu\|_{TV}^2 \leq Cxe^{-at} + O(e^{-at}).$$

This shows once more that for the recurrent Ornstein Uhlenbeck process, speed of convergence to equilibrium is exponential.

For general diffusions, it is not possible to obtain such a precise control, and bounding the total variation distance $\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV}$ is mostly done by coupling two trajectories, one starting from x, the other starting from y (recall also the proof of Theorem 2.2). This is actually the precise context in which Veretennikov (1997) works. In what follows we sketch his ideas in dimension one (actually, they do not depend on the dimension), under the general conditions of this chapter and under condition (3.29) which is a slight improvement of his original condition.

From now on, for the simplicity of exposition, we suppose additionally to the general conditions of this section, that b and σ are twice continuously differentiable. Coupling two trajectories means that we have to make them meet as soon as possible. Even in dimension one, this is not quite evident, if the processes are taking values in the whole real line. To make meet two trajectories is in fact a two-dimensional problem, and it is no longer sufficient to consider return times to points. So instead of considering return times to points, we consider return times to sets - and the sets that we shall consider are petite sets. We choose $C = [-M, M]$ for some constant M that will be fixed later but that has to satisfy $M \geq M_0$. Fix some $t_* > 0$. Then, since b and $\sigma \in C^2$ and $\sigma^2(x) \neq 0$ for all x, we have the lower bound

$$P_{t_*}(x, dy) \geq 1_C(x) \alpha \nu(dy) \tag{3.32}$$

for $\alpha \in]0,1]$ and for some probability measure ν, both depending on t_* and M. Indeed, it suffices to note that under our conditions, $P_t(x, dy)$ possesses a transition density $p_t(x, y)$ which is continuous in x and y and strictly positive. Hence we may choose $\alpha = \int \inf_{x:|x| \leq M} p_{t_*}(x, y) dy$ and $\nu(dy) = \frac{1}{\alpha} \inf_{x:|x| \leq M} p_{t_*}(x, y) dy$.

Now we describe the construction of the coupling. Take two independent copies X_t^x and X_t^y of the process up to the first common hitting time

$$T_1 = \inf\{t : X_t^x \in C \text{ and } X_t^y \in C\}. \tag{3.33}$$

We let them evolve independently up to this time T_1. At time T_1, choose a uniform random variable U_1, independently of everything else. If $U_1 \leq \alpha$, then choose

$$X_{T_1 + t_*}^x = X_{T_1 + t_*}^y \sim \nu(dz).$$

Else, choose

$$X_{T_1 + t_*}^x \sim \frac{P_{t_*}(X_{T_1}^x, dz) - \alpha \nu(dz)}{1 - \alpha},$$

with a similar choice for $X_{T_1 + t_*}^y$. Integrating with respect to all possible values of U_1, we notice that the unconditional law of $X_{T_1+t_*}^x$ is $P_{t_*}(X_{T_1}^x, dz)$ as required, and the unconditional law of $X_{T_1+t_*}^y$ $P_{t_*}(X_{T_1}^y, dz)$, recall also the proof of Theorem 2.2 above.

Finally, fill in diffusion bridges between T_1 and $T_1 + t_*$, which is possible since condition (**??**) is satisfied. We refer to Fitzsimmons, Pitman and Yor (1992) for the details of the construction of the diffusion bridge.

Then we iterate the above procedure: For $n \geq 2$, let

$$T_n = \inf\{t > T_{n-1} + t_* : X_t^x \in C \text{ and } X_t^y \in C\}$$

and choose $X^x_{T_n+t_*}$ and $X^y_{T_n+t_*}$ analogously to the above described procedure. Once both trajectories meet, we merge them into a single trajectory which evolves from that time on freely, according to the underlying diffusion law.

Clearly $\inf\{t : X^x_t = X^y_t\} \le \tau := \inf\{T_n + t_* : U_n \le \alpha\}$, and thus

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} \le \tilde{P}_{x,y}(\tau > t) \le \frac{\tilde{E}_{x,y}(\tau^l)}{t^l}, \tag{3.34}$$

for any $l > 1$. It remains to establish up to which power l we can push the above bound. It is clear that the behavior of τ is related to that of T_1, and we quote the following result from Veretennikov (1997):

Proposition 3.8 *Grant (3.29) with $\gamma = 0$. Suppose moreover that $r > \frac{3}{2}$. Then there exists a constant $M_1 \ge M_0$, such that for all $M \ge M_1$,*

$$E_{x,y}T_1^{k+1} \le C(1 + |x|^p + |y|^p),$$

for all $k < r - \frac{3}{2}$ and for all $p \in]2k+2, 2r-1[$.

The proof of Proposition 3.8 is slightly lengthy, and we postpone it to the next section.

We close this section by showing that Proposition 3.8 indeed implies that the coupling time τ admits the same number of moments as T_1. Put $l = k + 1$. We have

$$\begin{aligned}
\tau^l &= \sum_{n=1}^{\infty} 1_{\{U_1>\alpha,\ldots,U_{n-1}>\alpha,U_n\le\alpha\}}(T_n + t_*)^l \\
&\le 2^{l-1}t_*^l + 2^{l-1}\sum_{n=1}^{\infty} 1_{\{U_1>\alpha,\ldots,U_{n-1}>\alpha,U_n\le\alpha\}}T_n^l \\
&\le 2^{l-1}t_*^l + 2^{l-1}\sum_{n=1}^{\infty} 1_{\{U_1>\alpha,\ldots,U_{n-1}>\alpha,U_n\le\alpha\}}n^{l-1}[T_1^l + (T_2 - T_1)^l + \ldots + (T_n - T_{n-1})^l].
\end{aligned}$$

Now we have

$$\tilde{E}_{x,y}1_{\{U_1>\alpha,\ldots,U_{n-1}>\alpha,U_n\le\alpha\}}T_1^l \le C(1-\alpha)^{n-1}(1 + |x|^p + |y|^p).$$

For any of the following terms, we use the Hölder inequality in order to decouple the choices of the uniform random variables and the evolution of the process. Choose $a, b > 1$ such that $a^{-1} + b^{-1} = 1$, $lb < r - \frac{1}{2}$. Then we obtain, using Markov's property with respect to T_1,

$$\begin{aligned}
\tilde{E}_{x,y}\left[1_{\{U_1>\alpha,\ldots,U_{n-1}>\alpha,U_n\le\alpha\}}(T_2 - T_1)^l\right] &\le (1-\alpha)^{(n-1)/a}\sup_{x,y\in C}\left(E_{x,y}(t_* + T_1\circ\vartheta_{t_*})^{lb}\right)^{1/b} \\
&\le (1-\alpha)^{(n-1)/a}\,2^l\left(t_*^{lb} + \sup_{x,y\in C}E_{x,y}(E_{X^x_{t_*},X^y_{t_*}}T_1^{lb})\right)^{1/b} \\
&\le C(1-\alpha)^{(n-1)/a}\,2^l\left(t_*^{lb} + \sup_{x,y\in C}[1 + E_x|X^x_{t_*}|^p + E_y|X^y_{t_*}|^p]\right)^{1/b} \le C(1-\alpha)^{(n-1)/a},
\end{aligned}$$

where we have used (3.42) below in order to control $E_x|X^x_{t_*}|^p + E_y|X^y_{t_*}|^p$. We conclude that

$$\tilde{E}_{x,y}\tau^l \le 2^{l-1}\left(t_*^l + C\sum_{n\ge 1} n^l(1-\alpha)^{(n-1)/a}(1+|x|^p+|y|^p)\right)$$
$$\le C(1+|x|^p+|y|^p), \quad (3.35)$$

as desired. Notice that the above upper bound can be integrated with respect to $m(y)dy$, since $p < 2r-1$. We obtain

Theorem 3.9 (Veretennikov (1997)) *Under the conditions of Proposition 3.8 we have*

$$\|P_t(x,\cdot)-\mu\|_{TV} \le Ct^{-(k+1)}(1+|x|^p),$$

for any $k < r - \frac{3}{2}$, $p \in]2k+2, 2r-1[$.

Remark 3.10 *Recall that* $l = k+1$. *Notice that in Theorem 3.7, we can go up to hitting time moments of any order* $l < r + \frac{1}{2}$, *whereas in the above Theorem, only moments* $l < r - \frac{1}{2}$ *can be considered. So we "loose one moment".*

3.4 Proof of Proposition 3.8

The following proof is taken from Veretennikov (1997).

Consider the two dimensional diffusion process $Z_t = (X^x_t, X^y_t)$ and apply Itô's formula to $f(t,x,y) = (1+t)^k(|x|^m+|y|^m)$, for some fixed $k < r - \frac{3}{2}$ and for some $m < p$ that will be chosen later. We obtain

$$\begin{aligned}
&(1+t)^k(|X^x_t|^m + |X^y_t|^m) = |x|^m + |y|^m + M_t\\
&+\int_0^t (1+s)^{k-1}|X^x_s|^{m-2}\left[k|X^x_s|^2 + m(1-s)X^x_s b(X^x_s) + \frac{1}{2}m(m-1)(1-s)\sigma^2(X^x_s)\right]ds\\
&+\int_0^t (1+s)^{k-1}|X^y_s|^{m-2}\left[k|X^y_s|^2 + m(1-s)X^y_s b(X^y_s) + \frac{1}{2}m(m-1)(1-s)\sigma^2(X^y_s)\right]ds.
\end{aligned} \quad (3.36)$$

We start by evaluating the second line in the above equality. For some $\varepsilon > 0$, we distinguish the events $\{|X^x_s|^2 \ge \varepsilon(1+s)\}$ and $\{|X^x_s|^2 < \varepsilon(1+s)\}$. This gives

$$\begin{aligned}
\int_0^t (1+s)^{k-1}|X^x_s|^{m-2}k|X^x_s|^2 ds &\le k\varepsilon\int_0^t (1+s)^k|X^x_s|^{m-2}ds\\
&\quad + \int_0^t k(1+s)^{k-1}|X^x_s|^m 1_{\{|X^x_s|^2\ge\varepsilon(1+s)\}}ds.
\end{aligned}$$

Plugging this into the second line of (3.36) we obtain

$$\begin{aligned}
&\int_0^t (1+s)^{k-1}|X^x_s|^{m-2}\left[k|X^x_s|^2 + m(1-s)X^x_s b(X^x_s) + \frac{1}{2}m(m-1)(1-s)\sigma^2(X^x_s)\right]ds\\
&\quad\le \int_0^t (1+s)^k|X^x_s|^{m-2}\left[k\varepsilon + mX^x_s b(X^x_s) + \frac{1}{2}m(m-1)\sigma^2(X^x_s)\right]ds\\
&\qquad + k\int_0^t (1+s)^{k-1}|X^x_s|^m 1_{\{|X^x_s|^2\ge\varepsilon(1+s)\}}ds.
\end{aligned}$$

We do not touch the last expression. But the second line will be treated considering successively the cases $|X^x_s| \le M_0$, $M_0 \le |X^x_s| < M$ and $|X^x_s| \ge M$. We choose ε sufficiently small such that

$$c_1 := m\lambda_- \left(r - \frac{1}{2}(m-1) \right) - k\varepsilon > 0$$

(recall that $r - \frac{1}{2}(m-1) > 0$ by choice of m). We use that for all $|x| \ge M_0$, since $\sigma^2(x) \ge \lambda_-$ (recall that $\gamma = 0$),

$$k\varepsilon + mxb(x) + \frac{1}{2}m(m-1)\sigma^2(x) \le -c_1,$$

for some positive constant $c_1 > 0$. This gives

$$\int_0^t (1+s)^k |X^x_s|^{m-2} \left[k\varepsilon + mX^x_s b(X^x_s) + \frac{1}{2}m(m-1)\sigma^2(X^x_s) \right] ds \\ \le -c_1 M^{m-2} \int_0^t (1+s)^k 1_{\{|X^x_s| \ge M\}} ds + c_2 M_0^{m-1} \int_0^t (1+s)^k ds, \quad (3.37)$$

where $c_2 = k\varepsilon + m \sup\{|b(x)| : |x| \le M_0\} + \frac{1}{2}m(m-1)\sup\{\sigma^2(x) : |x| \le M_0\}$. In the last line of (3.37), the last contribution comes from $|X^x_s| \le M_0$. The contributions coming from $M_0 \le |X^x_s| \le M$ are negative and we did not count them.

To resume we obtain the following upper bound for the r.h.s. of (3.36).

$$(1+t)^k(|X^x_t|^m + |X^y_t|^m) \le |x|^m + |y|^m + M_t \\ + k\int_0^t (1+s)^{k-1} \left[|X^x_s|^m 1_{\{|X^x_s| \ge \varepsilon(1+s)\}} + |X^y_s|^m 1_{\{|X^y_s| \ge \varepsilon(1+s)\}} \right] ds \\ - c_1 M^{m-2} \int_0^t (1+s)^k [1_{\{|X^x_s| \ge M\}} + 1_{\{|X^y_s| \ge M\}}] ds + c_2 M_0^{m-1} \int_0^t (1+s) ds, \quad (3.38)$$

where c_1 and c_2 are positive constants that do no depend on M. We will consider the above inequality up to the bounded stopping time $T_1 \wedge t$, recall also (3.33). Note that on $s \le T_1 \wedge t$, at least one of the two components X^x_s or X^y_s has absolute value bigger or equal than M, by definition of T_1. As a consequence, $1_{\{|X^x_s| \ge M\}} + 1_{\{|X^y_s| \ge M\}} \ge 1$.

We can now give the precise choice of m that we will consider in the above evaluations. Recall that $k < r - \frac{3}{2}$. Fix some $p \in]2k+2, 2r-1[$ and choose $m \in]2, p-2k[$. In Lemma 3.12 below we will see that for any $\varepsilon > 0$,

$$\int_0^\infty E_x \left((1+s)^{k-1} |X_s|^m 1_{\{X_s^2 \ge \varepsilon(1+s)\}} \right) ds \le C(1 + |x|^p). \quad (3.39)$$

We use this in order to control (3.38). Using the stopping rule with $t \wedge T_1$, we obtain

$$E_{x,y} \left[(1 + t\wedge T_1)^k (|X^x_{t\wedge T_1}|^m + |X^y_{t\wedge T_1}|^m) \right] \le C(1 + |x|^p + |y|^p) \\ - c_1 M^{m-2} E_{x,y} \int_0^{t\wedge T_1} (1+s)^k ds + c_2 M_0^{m-1} E\!\!\!E_{x,y} \int_0^{t\wedge T_1} (1+s)^k ds. \quad (3.40)$$

Now we choose M_1 such that

$$c_1 M_1^{m-2} > c_2 M_0^{m-1}$$

(recall that $m > 2$). Therefore, for all $M \geq M_1$, (3.40) implies, letting $t \to \infty$,

$$E_{x,y} T_1^{k+1} \leq C(1 + |x|^p + |y|^p), \tag{3.41}$$

and this finishes the proof. •

The main important point that made the above proof work is Lemma 3.12 below. In order to prove it, we first need the following statement.

Lemma 3.11 *Under the conditions of Proposition 3.8 we have for any* $2 < m < 2r - 1$,

$$E_x |X_t|^m \leq C(1 + |x|^m). \tag{3.42}$$

The above statement is in fact very strong. It implies that the process is $m-$ultimately bounded for all $m < 2r - 1$.

Proof: Suppose that $|x| \geq M_0$. Fix some $r' < r$ such that $m < 2r' - 1$, $r' > \frac{1}{2}$ and put $\tilde{b}(z) = -r' \frac{\sigma^2(z)}{z}$ for all z with $|z| \geq M_0$. Let $\tilde{X}$ be the diffusion process with drift $\tilde{b}$ and diffusion coefficient σ, starting from $\tilde{X}_0 = x$ and reflected at x and at $-x$ such that $|\tilde{X}_t| \geq |x|$ for all t :

$$d\tilde{X}_t = \tilde{b}(\tilde{X}_s)ds + \sigma(\tilde{X}_s)dW_s + d\tilde{L}_t^+ - d\tilde{L}_t^-, \tag{3.43}$$

where

$$\tilde{L}_t^+ = \int_0^t 1_{\{\tilde{X}_s = x\}} d\tilde{L}_s^+, \; \tilde{L}_t^- = \int_0^t 1_{\{\tilde{X}_s = -x\}} d\tilde{L}_s^-, \; E \int_0^\infty 1_{\{|\tilde{X}_s| = |x|\}} ds = 0.$$

Then the comparison theorem for diffusions shows that $|\tilde{X}_t| \geq |X_t|$ for all t almost surely. Hence it suffices to show (3.42) for $\tilde{X}$ instead of X. By construction, the speed density $\tilde{m}$ of $\tilde{X}$ satisfies (up to multiplication with a constant)

$$\tilde{m}(z) = \frac{1}{\sigma^2(z) z^{2r'}} \leq \frac{1}{\lambda_-} z^{-2r'} \text{ for all } |z| \geq |x|.$$

Write $\tilde{\mu}(dz) = \frac{1}{\tilde{m}(I\!R)} \tilde{m}(z)dz$. Finally let $\bar{X}_t$ be solution of the same equation (3.43), reflected at $|x|$, such that $\bar{X}_0 \sim \tilde{\mu}$. Then $|\bar{X}_t| \geq |\tilde{X}_t|$ for all t, almost surely. So it suffices to show that (3.42) holds in fact for $\bar{X}$. But

$$\begin{aligned} E|\bar{X}_t|^m = \frac{2}{\tilde{m}(I\!R)} \int_{|x|}^\infty z^m \frac{1}{\sigma^2(z) z^{2r'}} dz &\leq \frac{2}{\lambda_- \, \tilde{m}(I\!R)} \int_{|x|}^\infty z^{m-2r'} dz \\ &= \frac{2}{\lambda_- \, \tilde{m}(I\!R)(2r' - m - 1)} |x|^{m+1-2r'} \leq C|x|^m, \end{aligned}$$

since $r' > \frac{1}{2}$. Hence, by the comparison theorem,

$$E_x |X_t|^m \leq E|\tilde{X}_t|^m \leq E|\bar{X}_t| \leq C|x|^m,$$

for $|x| \geq M_0$. For $|x| < M_0$, the result follows analogously, by stopping at the first exit time of $[-M_0, M_0]$. •

The second lemma is as follows.

Lemma 3.12 *Under (3.29), for any $0 < m < p < 2r - 1$, for any $k < \frac{p-m}{2}$ and $\varepsilon > 0$, we have*

$$\int_0^\infty E_x\left((1+s)^{k-1}|X_s|^m 1_{\{X_s^2 \geq \varepsilon(1+s)\}}\right) ds \leq C(1+|x|^p). \tag{3.44}$$

We give the proof of Veretennikov (1997).

Proof: Write $p = m\tilde{p}$, $\tilde{p} > 1$, and let $\tilde{q} > 1$ such that $\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1$. Use Hölder's inequality

$$\begin{aligned} E_x|X_s|^m 1_{\{X_s^2 \geq \varepsilon(1+s)\}} &\leq [E_x|X_s|^p]^{1/\tilde{p}} \left[P_x(X_s^2 \geq \varepsilon(1+s))\right]^{1/\tilde{q}} \\ &\leq [E_x|X_s|^p]^{1/\tilde{p}} \left(\frac{E_x|X_s|^p}{[\varepsilon(1+s)]^{p/2}}\right)^{1/\tilde{q}} \\ &\leq C(1+s)^{-\frac{p}{2\tilde{q}}} E_x|X_s|^p \\ &\leq C(1+|x|^p)(1+s)^{-\frac{p}{2\tilde{q}}}. \end{aligned}$$

Here, we have used (3.42). Multiplying with $(1+s)^{k-1}$ and integrating with respect to ds, we obtain a finite expression if and only if $k - 1 - \frac{p}{2\tilde{q}} < -1$, which is equivalent to $2k < p - m$, since $p/\tilde{q} = p - m$, by definition of $\tilde{p}$ and $\tilde{q}$. •

4 Diffusions in higher dimensions

The aim of this section is to establish sufficient conditions for Harris-recurrence of diffusion processes in higher dimensions, by following the ideas presented at the end of Subsection 2.3. Let

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x \in I\!R^d, \tag{4.45}$$

W some $n-$dimensional Brownian motion. We impose the following conditions on the parameters:

$$b \text{ and } \sigma \text{ are in } C^2 \text{ and of linear growth and } \sigma \text{ is bounded.} \tag{4.46}$$

Moreover, we assume uniform ellipticity: Writing $a(x) = \sigma(x)\sigma^*(x)$ we have

$$< a(x)y, y > \geq \lambda_- \|y\|^2 \tag{4.47}$$

for all $y \in I\!R^d$, for some fixed $\lambda_- > 0$. Under the above conditions, we have pathwise uniqueness, the existence of a unique strong solution and the strong Markov property of solutions to the above equation (4.45), see Khasminskii (2012). Moreover, solutions have almost surely continuous sample paths. Classical results on lower bounds for transition densities of diffusions, see for instance Friedman (1964), imply that in this case the process admits a transition density for any time t, denoted $p_t(x, y)$, which is strictly positive and continuous in x and y. As a consequence, any compact set is petite.

4.1 Regenerative scheme and Harris-recurrence

We fix a compact set C and a time parameter $t_* > 0$ such that the local Doeblin condition (3.32)

$$P_{t_*}(x, dy) \geq \alpha 1_C(x)\nu(dy)$$

holds for a constant α and a probability measure ν. In what follows we show how to use (3.32) in order to introduce a sequence of regeneration times that allow to prove the Harris recurrence of X. These regeneration times do not exist for the original strong solution of the stochastic differential equation (4.45) directly, but they exist for a version of the process on an extended probability space, rich enough to support the driving Brownian motion of (4.45) and an independent i.i.d. sequence of uniform random variables $(U_n)_{n\geq 1}$.

More precisely, in what follows we construct a stochastic process $(Y_t)_{t\geq 0}$ on this richer probability space, such that $\mathcal{L}((Y_t)_{t\geq 0}) = \mathcal{L}((X_t)_{t\geq 0})$, where X_t denotes the original strong solution of (4.45). Our construction follows old ideas on "constructing Markov processes by piecing out" that have been exposed in Ikeda, Nagasawa and Watanabe (1966). It is done recursively as follows.

We start with $Y_t = X_t$ for all $0 \leq t \leq \tilde{S}_1$, where

$$\tilde{S}_1 = \inf\{t \geq t_* : X_t \in C\} \text{ and put } \tilde{R}_1 = \tilde{S}_1 + t_*.$$

At time $\tilde{S}_1$, we choose U_1, the first of the uniform random variables. If $U_1 \leq \alpha$, we choose

$$Y_{\tilde{R}_1} \sim \nu(dy). \tag{4.48}$$

Else, if $U_1 > \alpha$, given $Y_{\tilde{S}_1} = x$, we choose

$$Y_{\tilde{R}_1} \sim \frac{P_{t_*}(x, dy) - \alpha\nu(dy)}{1-\alpha}. \tag{4.49}$$

Finally, given $Y_{\tilde{R}_1} = y$, we fill in the missing trajectory between time $\tilde{S}_1$ and time $\tilde{R}_1$

$$(Y_t)_{t\in]\tilde{S}_1,\tilde{R}_1[} \text{ according to the diffusion bridge law } \frac{p_{t-\tilde{S}_1}(x,z)p_{\tilde{R}_1-t}(z,y)}{p_{t_*}(x,y)}dz, \tag{4.50}$$

compare also to Fitzsimmons, Pitman and Yor (1992). Notice that by construction, if we do not care about the exact choice of the auxiliary random variable U_1, then we have that $(Y_t)_{t\leq \tilde{R}_1} \stackrel{\mathcal{L}}{=} (X_t)_{t\leq \tilde{R}_1}$.

We continue this construction after time $\tilde{R}_1$: Choose Y_t equal to X_t for all $t \in]\tilde{R}_1, \tilde{S}_2]$, where

$$\tilde{S}_2 = \inf\{t > \tilde{R}_1 : X_t \in C\} \text{ and put } \tilde{R}_2 = \tilde{S}_2 + t_*.$$

At time $\tilde{S}_2$ we choose U_2 in order to realize the choice of $Y_{\tilde{R}_2}$ according to the splitting of the transition kernel P_{t_*}, as in (4.48) and (4.49). More generally, the construction is therefore achieved along the sequence of stopping times

$$\tilde{S}_{n+1} = \inf\{t > \tilde{R}_n : X_t \in C\}, \tilde{R}_{n+1} = \tilde{S}_{n+1} + t_*,\ n \geq 1,$$

where during each $]\tilde{R}_n, \tilde{S}_{n+1}]$, Y follows the evolution of the original strong solution of the SDE, whereas the intervals $[\tilde{S}_{n+1}, \tilde{R}_{n+1}]$ are used to construct the splitting. In particular, every time that we may choose a transition according to (4.48), we introduce a regeneration event for the process Y, and therefore the following two sequences of generalized stopping times will play a role. Firstly,

$$S_1 = \inf\{\tilde{S}_n : U_n \leq \alpha\}, \ldots, S_n = \inf\{\tilde{S}_m > S_{n-1} : S_m < \alpha\}, n \geq 1,$$

and secondly,

$$R_n = S_n + t_*, n \geq 1.$$

The above construction of the process Y implies in particular that we change the memory structure of the original process X, since at each time $\tilde{S}_n$, a projection into the future is made, by choosing $Y_{\tilde{R}_n}$. We refer the reader to Section 6.1 where this is made rigorous.

Let $\tilde{N}_t = \sup\{n : U_n \leq t\}$ and

$$\mathcal{F}_t^Y = \sigma\{Y_s, s \leq t, U_n, Y_{\tilde{R}_n}, n \leq \tilde{N}_t\}, t \geq 0,$$

be the canonical filtration of the process Y. The sequence of $(\mathcal{F}_t^Y)_{t\geq 0}-$stopping times $(R_n)_{n\geq 1}$ is a generalized sequence of **regeneration times** similar to that considered at the beginning of Section 3 and in Example 1.5.

Definition 4.1 *A sequence $(R_n)_{n\geq 1}$ is called generalized sequence of* **regeneration times**, *if*

1. *$R_n \uparrow \infty$ as $n \to \infty$.*
2. *$R_{n+1} = R_n + R_1 \circ \vartheta_{R_n}$.*
3. *$Y_{R_n+\cdot}$ is independent of $\mathcal{F}^Y_{S_n-}$.*
4. *At regeneration times, the process starts afresh from $Y_{R_n} \sim \nu(dy)$.*
5. *The trajectories $(Y_{R_n+s}, 0 \leq s \leq R_{n+1} - R_n)_n$ are $2-$independent, i.e. $(Y_{R_n+s}, 0 \leq s \leq R_{n+1} - R_n)$ and $(Y_{R_m+s}, 0 \leq s \leq R_{m+1} - R_m)$ are independent if and only if $|m - n| \geq 2$.*

Remark 4.2

1. *Let us make the following remark concerning item 3. of the above list of properties: In contrary to the one-dimensional case of Section 3, $Y_{R_n+\cdot}$ is not independent of $\mathcal{F}^Y_{R_n-}$. This is due to the fact that between S_n and R_n, we fill in the missing piece of trajectory by using a Markov bridge. But this bridge depends on the position Y_{R_n} where the process has to arrive.*
2. *As pointed out before, the trajectories of Y are not the same as those of the original strong solution X of the SDE (4.45). However, by definition, the Harris-recurrence is only a property in law. As a consequence, if, for a given set A, we succeed to show that almost surely, Y visits it infinitely often, the same is automatically true for X as well.*

As a consequence of the above discussion, the following theorem is now immediate.

Theorem 4.3 *If for all $x \in I\!R^d$ we have $P_x(R_1 < \infty) = 1$, then the process X is recurrent in the sense of Harris.*

Proof Define a measure π on $(I\!R^d, \mathcal{B}(I\!R^d))$ by

$$\pi(A) = E(\int_{R_1}^{R_2} 1_A(Y_s)ds), \; A \in \mathcal{B}(I\!R^d).$$

For any $n \geq 1$, put $\xi_n = \int_{R_{n-1}}^{R_n} 1_A(Y_s)ds$. By construction, the random variables $\xi_{2n}, n \geq 1$, are i.i.d. Moreover, the random variables $\xi_{2n+1}, n \geq 1$, are i.i.d., as well. Put

$$N_t = \sup\{n : R_n \leq t\}$$

and observe that $N_t \to \infty$ as $t \to \infty$. Hence, applying the strong law of large numbers separately to the sequence $(\xi_{2n})_{n\geq 1}$ and the sequence $(\xi_{2n+1})_{n\geq 1}$, we have that

$$\frac{\int_0^t 1_A(Y_s)ds}{N_t} \to \pi(A) \text{ as } t \to \infty,$$

P_x-almost surely, for any $x \in I\!R^d$. This implies that any set A such that $\pi(A) > 0$ is visited infinitely often by the process Y almost surely. Thus, we have the recurrence property also for the process X, for any set A such that $\pi(A) > 0$. Then by a deep theorem of Azéma, Duflo and Revuz (1969), see also Theorem 1.2. of Höpfner and Löcherbach (2003), the process is indeed Harris. •

The following definition is now a straightforward generalization of Definition 3.3.

Definition 4.4 (Order of recurrence) *We say that the process is recurrent of order p if $E_x R_1^p + E_x(R_2 - R_1)^p < \infty$ for all $x \in I\!R^d$.*

It is clear that in order to determine the order of recurrence, the following stopping time

$$T_C(t_*) = \inf\{t \geq t_* : X_t \in C\}$$

plays an important role. Indeed:

Remark 4.5 *The same proof as that leading to formula (3.35) above shows that due to the construction of the regeneration times (geometrical trials within successive visits to the set C), the process is recurrent of order p if*

$$E_x T_C(t_*)^p < \infty \text{ for all } x \in C^c \text{ and if } \sup_{x\in C} E_x T_C(t_*)^p < \infty, \tag{4.51}$$

see also the proof of Theorem 4.1 in Löcherbach and Loukianova (2011).

We use a condition which is reminiscent of Veretennikov's drift condition in order to establish sufficient conditions ensuring (4.51).

Drift condition: There exist $M_0 > 0$ and $r > 0$ such that

$$< b(x), x > + \frac{1}{2} tra(x) \ \leq -r < a(x)\frac{x}{\|x\|}, \frac{x}{\|x\|} >, \text{ for all } \|x\| \geq M_0. \tag{4.52}$$

The above drift condition will imply (4.51) for the set $C := \{x : \|x\| \leq M_0\}$, for a suitable choice of p. Since any compact set is petite, C is petite, and from now on, C will stand for this fixed choice of a petite set.

Remark 4.6 *Other drift conditions have been considered in the literature, see for instance example 5.1 in Douc, Fort and Guillin (2009). The above condition implies the weaker one*

$$< b(x), x > + \frac{1}{2} tra(x) \ \leq -\varepsilon$$

for some $\varepsilon > 0$, which just implies Harris recurrence, without any control on the speed of convergence, see Höpfner and Löcherbach (2010).

Condition (4.52) implies that we dispose of a Lyapunov function for our system. More precisely, let $f : I\!R^d \to [1, \infty[$ be a smooth function such that $f(x) = \|x\|^m$ for all x with $\|x\| \geq M_0$. Notice that we choose f to be lower bounded by 1. Then we have for $\|x\| \geq M_0$,

$$\begin{aligned} Af(x) &= m\|x\|^{m-2}\left[< x, b(x) > +\frac{1}{2}tra(x)\right] + \frac{1}{2}m(m-2)\|x\|^{m-2} < a(x)\frac{x}{\|x\|}, \frac{x}{\|x\|} > \\ &\leq m\|x\|^{m-2}\left(-r + \frac{1}{2}(m-2)\right) < a(x)\frac{x}{\|x\|}, \frac{x}{\|x\|} > . \end{aligned} \tag{4.53}$$

Recalling that $C = \{x : \|x\| \leq M_0\}$, we obtain therefore the following: If $m < 2r+2$, then

$$Af(x) \leq -c_1 f(x)^{(m-2)/m} + c_2 1_C(x), \tag{4.54}$$

where c_1 and c_2 are positive constants. In other words, the function f is a generalized Lyapunov function of the process in the following sense.

Definition 4.7 *A function $f : I\!R^d \to [1, \infty[$ satisfying*

1. *$f(x) \to \infty$ as $\|x\| \to \infty$;*
2. *$Af(x) \leq c_1 f(x)^{1-\alpha} + c_2 1_C(x)$; with C a petite set and $\alpha \in [0, 1[$,*

is called **generalized Lyapunov function** *of the process. If $\alpha = 0$, then we call f (classical) Lyapunov function.*

Remark 4.8 *The role of the parameter α in the above definition is important and determines the speed of convergence to equilibrium. The case $\alpha = 0$ is the most favorite one and gives rise to exponential speed of convergence. Notice that under our condition (4.52), the function $f(x) = \|x\|^m, \|x\| \geq M_0$, is a generalized Lyapunov function with $\alpha = \frac{2}{m}$.*

For any fixed $\delta > 0$, let $T_c(\delta) = \inf\{t \geq \delta : X_t \in C\}$. We show that under the above condition (4.52), the following holds (compare also to Theorem 3.7).

Theorem 4.9 *Grant condition (4.52) with $r > 0$. Then for any $2 < m < 2r + 2$ and for any δ, there exist two positive constants c_1 and c_2, depending on δ, such that*

$$E_x(T_C(\delta)^{m/2}) \leq c_1\|x\|^m + c_2.$$

Corollary 4.10 *Under condition (4.52) with $r > 0$, the process is recurrent of order p for any $p < r + 1$. In particular, the deviation inequality (3.30) holds, with $r(t) = t^{-(p-1)}$, for any $p < r + 1$.*

Veretennikov (1997) gives a proof of the above theorem which is analogous to the one dimensional proof that we presented in Section 3.4. His proof uses extensively the properties of diffusion processes: comparison of diffusions, local time, etc. Instead of presenting his proof in the multidimensional setting, we decided to present another approach which is entirely based on the use of Lyapunov functions. Probably the first paper in this context is Douc, Fort, Moulines and Soulier (2004). Our presentations follows closely Douc, Fort and Guillin (2009), proof of Theorem 4.1.

Proof of Theorem 4.9: We start by rewriting (4.54). Write

$$\Phi(t) = c_1 t^{1-\frac{2}{m}},\ t \geq 1,$$

then $Af(x) \leq -\Phi \circ f(x) + c_2 1_C(x)$. Put

$$H(u) = \int_1^u \frac{1}{\Phi(t)} dt = \frac{m}{2c_1}(u^{2/m} - 1)$$

and note that

$$H^{-1}(s) = (\frac{2c_1}{m} s + 1)^{m/2}.$$

Finally let

$$r(s) = \Phi \circ H^{-1}(s) = c_1 \left(\frac{2c_1}{m} s + 1 \right)^{\frac{m-2}{2}}$$

and note that $r(s) \geq c\, s^{\frac{m-2}{2}}$. The function r is log-concave which follows from r'/r non increasing (which is easy to check). This implies that $\frac{r(s+t)}{r(s)}$ is non increasing in s for any fixed $t > 0$. So $r(t+s) \leq \frac{1}{r(0)} r(t) r(s)$ for all $s, t \geq 0$.

Now, starting with (4.54), applying Itô's formula to $f(X_t)$, we obtain

$$E_x \left(\int_0^{T_C(\delta)} \Phi \circ f(X_s) ds \right) \leq f(x) - 1 + c_2 \delta \tag{4.55}$$

(recall that $f(\cdot) \geq 1$). This implies that $T_C(\delta) < \infty$ P_x-almost surely for all $x \in \mathbb{R}^d$.

Define $G(t,u) = H^{-1}(H(u) + t) - H^{-1}(t)$ and note that

$$\frac{\partial G(t,u)}{\partial u} = \frac{r(H(u)+t)}{\Phi(u)} = \frac{r(H(u)+t)}{r(H(u))},\ \frac{\partial G(t,u)}{\partial t} = r(H(u)+t) - r(t).$$

Since $\frac{r(h+t)}{r(h)}$ non increasing in h for any fixed $t > 0$ and H non decreasing, we see that $u \mapsto \frac{\partial G(t,u)}{\partial u}$ is non increasing. Hence, $u \mapsto G(t,u)$ is concave for any fixed t.

We define $\tau = T_C(\delta) \wedge \tau_K$, where $\tau_K = \inf\{t : \|X_t\| \geq K\}$. Moreover, for some fixed $\varepsilon > 0$, we put $t_k = \varepsilon k, k \geq 0$, and let $N_\varepsilon = \sup\{k \geq 1 : t_{k-1} < \tau\}$. Thus $\tau \leq \varepsilon N_\varepsilon$. Moreover, $\varepsilon N_\varepsilon$ is a stopping time, since $\{\varepsilon N_\varepsilon = k\varepsilon\} = \{(k-1)\varepsilon < \tau \leq k\varepsilon\} \in \mathcal{F}_{k\varepsilon}$. Hence we obtain, for any $M \geq \delta$, with $M_\varepsilon = [M/\varepsilon]$,

$$\begin{aligned} E_x \left(\int_0^{\tau \wedge M} r(s) ds \right) &\leq \limsup_{\varepsilon \to 0} E_x \left(\int_0^{(\varepsilon N_\varepsilon) \wedge M} r(s) ds \right) \\ &= \limsup_{\varepsilon \to 0} E_x \left(\int_0^{\varepsilon(N_\varepsilon \wedge M_\varepsilon)} r(s) ds \right) \leq \limsup_{\varepsilon \to 0} A(\varepsilon) + G(0, f(x)), \end{aligned}$$

where

$$A(\varepsilon) = E_x \left[G(\varepsilon(N_\varepsilon \wedge M_\varepsilon), f(X_{\varepsilon(N_\varepsilon \wedge M_\varepsilon)})) - G(0, f(x)) \right] + E_x \left(\int_0^{\varepsilon(N_\varepsilon \wedge M_\varepsilon)} r(s) ds \right).$$

This last inequality follows from $G(\varepsilon(N_\varepsilon \wedge M_\varepsilon), f(X_{\varepsilon(N_\varepsilon \wedge M_\varepsilon)})) \geq 0$.

Note that $G(0, f(x)) = f(x) - 1$, hence the proof is finished provided we succeed to control $\limsup_{\varepsilon\to 0} A(\varepsilon)$. This is done along the lines of the proof of Douc, Fort and Guillin (2009) as follows.

$$\begin{aligned}
&E_x\left[G(\varepsilon(N_\varepsilon \wedge M_\varepsilon), f(X_{\varepsilon(N_\varepsilon\wedge M_\varepsilon)})) - G(0, f(x))\right] \\
&= \sum_{k=1}^{M_\varepsilon} E_x\left[G(t_k, f(X_{t_k})) - G(t_{k-1}, f(X_{t_{k-1}}))1_{\{\tau > t_{k-1}\}}\right] \\
&= \sum_{k=1}^{M_\varepsilon} E_x\Big[E\left[G(t_k, f(X_{t_k})) - G(t_{k-1}, f(X_{t_{k-1}}))|\mathcal{F}_{t_{k-1}}\right] 1_{\{\tau > t_{k-1}\}}\Big].
\end{aligned}$$

Now, we use that by concavity of $u \mapsto G(t, u)$,

$$G(t_k, f(X_{t_k})) - G(t_k, f(X_{t_{k-1}})) \le \frac{\partial G}{\partial u}(t_k, f(X_{t_{k-1}}))(f(X_{t_k}) - f(X_{t_{k-1}}))$$

and obtain

$$\begin{aligned}
&E\left[G(t_k, f(X_{t_k})) - G(t_{k-1}, f(X_{t_{k-1}}))|\mathcal{F}_{t_{k-1}}\right] \\
&\quad \le \frac{\partial G}{\partial u}(t_k, f(X_{t_{k-1}}))E\left[f(X_{t_k}) - f(X_{t_{k-1}})|\mathcal{F}_{t_{k-1}}\right] + \int_{t_{k-1}}^{t_k} \frac{\partial G}{\partial t}(s, f(X_{t_{k-1}}))ds \\
&= \frac{\partial G}{\partial u}(t_k, f(X_{t_{k-1}}))E\left[f(X_{t_k}) - f(X_{t_{k-1}})|\mathcal{F}_{t_{k-1}}\right] + \int_{t_{k-1}}^{t_k} [r(H(f(X_{t_{k-1}})) + s) - r(s)]ds \\
&\le \frac{\partial G}{\partial u}(t_k, f(X_{t_{k-1}}))E\left[f(X_{t_k}) - f(X_{t_{k-1}})|\mathcal{F}_{t_{k-1}}\right] + r(H(f(X_{t_{k-1}})) + t_k)\,\varepsilon - \int_{t_{k-1}}^{t_k} r(s)ds.
\end{aligned}$$

Using (4.54), we have

$$\begin{aligned}
&\frac{\partial G}{\partial u}(t_k, f(X_{t_{k-1}}))E\left[f(X_{t_k}) - f(X_{t_{k-1}})|\mathcal{F}_{t_{k-1}}\right] \\
&\quad \le \frac{\partial G}{\partial u}(t_k, f(X_{t_{k-1}}))E\left[-\int_{t_{k-1}}^{t_k} \Phi\circ f(X_s)ds + c_2\int_{t_{k-1}}^{t_k} 1_C(X_s)ds \,\Big|\mathcal{F}_{t_{k-1}}\right] \\
&\quad = \frac{r(H(f(X_{t_{k-1}})) + t_k)}{\Phi(f(X_{t_{k-1}}))}E\left[-\int_{t_{k-1}}^{t_k} \Phi\circ f(X_s)ds + c_2\int_{t_{k-1}}^{t_k} 1_C(X_s)ds \,\Big|\mathcal{F}_{t_{k-1}}\right].
\end{aligned}$$

Moreover, we use that

$$\frac{r(H(f(X_{t_{k-1}})) + t_k)}{\Phi(f(X_{t_{k-1}}))} = \frac{r(H(f(X_{t_{k-1}})) + t_k)}{r(H(f(X_{t_{k-1}})))} \le \frac{r(t_k)}{r(0)},$$

in order to upper bound

$$\begin{aligned}
&\frac{r(H(f(X_{t_{k-1}})) + t_k)}{\Phi(f(X_{t_{k-1}}))}E\left[c_2\int_{t_{k-1}}^{t_k} 1_C(X_s)ds \,\Big|\mathcal{F}_{t_{k-1}}\right] \\
&\quad \le c\,r(t_k)E\left[\int_{t_{k-1}}^{t_k} 1_C(X_s)ds \,\Big|\mathcal{F}_{t_{k-1}}\right] \le cE\left[\int_{t_{k-1}}^{t_k} r(s+\varepsilon)1_C(X_s)ds \,\Big|\mathcal{F}_{t_{k-1}}\right],
\end{aligned}$$

since $r(t_k) \le r(s+\varepsilon)$ for all $s \in [t_{k-1}, t_k]$.

We conclude that

$$\begin{aligned}
E\left[G(t_k, f(X_{t_k})) - G(t_{k-1}, f(X_{t_{k-1}}))|\mathcal{F}_{t_{k-1}}\right] \\
\le E\left[r(H(f(X_{t_{k-1}})) + t_k)\left(\frac{-\int_{t_{k-1}}^{t_k} \Phi\circ f(X_s)ds}{\Phi(f(X_{t_{k-1}}))} + \varepsilon\right)\Big|\mathcal{F}_{t_{k-1}}\right] \\
+ cE\left[\int_{t_{k-1}}^{t_k} r(s+\varepsilon)1_C(X_s)ds\,\Big|\mathcal{F}_{t_{k-1}}\right] - \int_{t_{k-1}}^{t_k} r(s)ds,
\end{aligned}$$

and thus

$$\begin{aligned}
A(\varepsilon) \le E_x\left[\sum_{k=1}^{M_\varepsilon} r(H(f(X_{t_{k-1}})) + t_k)\left(\frac{-\int_{t_{k-1}}^{t_k} \Phi\circ f(X_s)ds}{\Phi(f(X_{t_{k-1}}))} + \varepsilon\right)1_{\{\tau > t_{k-1}\}}\right] \\
+ cE\left[\int_0^{\varepsilon(N_\varepsilon \wedge M_\varepsilon)} r(s+\varepsilon)1_C(X_s)ds\right].
\end{aligned}$$

We show that the first term in the above upper bound tends to 0 as $\varepsilon \to 0$. First, since $\|X_{t_{k-1}}\| \le K$ and $t_k \le M$, we have

$$r(H(f(X_{t_{k-1}})) + t_k) \le r(H(f(K)) + M) \text{ for all } k \le M_\varepsilon.$$

Hence

$$\begin{aligned}
\limsup_{\varepsilon\to 0} E_x\left[\sum_{k=1}^{M_\varepsilon} r(H(f(X_{t_{k-1}})) + t_k)\left(\frac{-\int_{t_{k-1}}^{t_k} \Phi\circ f(X_s)ds}{\Phi(f(X_{t_{k-1}}))} + \varepsilon\right)1_{\{\tau > t_{k-1}\}}\right] \\
\le \frac{r(H(f(K)) + M)}{\Phi(1)} \limsup_{\varepsilon\to 0} E_x\left[\sum_{k=1}^{M_\varepsilon}\left(\int_{t_{k-1}}^{t_k} |\Phi\circ f(X_s) - \Phi\circ f(X_{t_{k-1}})|ds\right)\right] = 0,
\end{aligned}$$

since X has continuous sample paths. As a consequence,

$$\limsup_{\varepsilon\to 0} A(\varepsilon) \le c \limsup_{\varepsilon\to 0} E\left[\int_0^{\varepsilon(N_\varepsilon\wedge M_\varepsilon)} r(s+\varepsilon)1_C(X_s)ds\right] = cE\left[\int_0^{\tau\wedge M} r(s)1_C(X_s)ds\right].$$

Finally, letting $K, M \to \infty$, we obtain

$$E_x\left(\int_0^{T_C(\delta)} r(s)ds\right) \le cE\left(\int_0^{T_C(\delta)} r(s)1_C(X_s)ds\right) + f(x) - 1 \le c\int_0^\delta r(s)ds + f(x) - 1.$$

Recall that $r(s) \ge cs^{(m-2)/2}$. Thus the above inequality implies that

$$E_x(T_C(\delta)^{m/2}) \le \tilde{c}_1 f(x) + \tilde{c}_2 \le c_1\|x\|^m + c_2,$$

where c_1 and c_2 are positive constants. This finishes the proof. •

We close this section with the following remark which is a simple consequence of (4.55),

$$E_x\left(\int_0^{T_C(\delta)} \Phi\circ f(X_s)ds\right) \le f(x) - 1 + c_2\delta :$$

Corollary 4.11 *Let μ be the unique invariant measure of the process X. Under condition (4.52) with $r > 0$, we have*

$$\mu(\|x\|^p) < \infty$$

for any $p < 2r$.

Proof: Instead of using the regeneration times R_n, we use the successful visits of the process Y to C, i.e. the times $S_1, S_2, \ldots$. The points 1., 2. and 5. of Definition 4.1 remain valid, in particular, $(Y_{S_n+s}, 0 \le s \le S_{n+1} - S_n)$ and $(Y_{S_m+s}, 0 \le s \le S_{m+1} - S_m)$ are independent for $|m - n| \ge 2$. Now, let g be a measurable positive function. Then we have by the law of large numbers applied to Y that

$$\frac{1}{t} \int_0^t g(Y_s) ds \to E \int_{S_n}^{S_{n+1}} g(Y_s) ds$$

almost surely. On the other hand, by the ergodic theorem for the process X, we have that

$$\frac{1}{t} \int_0^t g(X_s) ds \to \mu(g)$$

almost surely. Since both processes have the same law, the two limits have to be the same. Hence a possible explicit representation of the invariant measure of X is given by

$$\mu(g) = E \int_{S_n}^{S_{n+1}} g(Y_s) ds,$$

up to multiplication with a constant.

In the following, we show that

$$E \int_{S_n}^{S_{n+1}} g(Y_s) ds \le \frac{2}{\alpha} \sup_{x \in C} E_x \int_0^{T_C(t_*)} g(X_s) ds. \tag{4.56}$$

1. First of all, recall that $R_n = S_n + t_*$ and $Y_{R_n} \sim \nu(dy)$. Recall also (4.50). Let $T_1 = \inf\{\tilde{S}_m : \tilde{S}_m > S_n\} = S_n + T_C(t_*) \circ \vartheta_{S_n} = R_n + T_C \circ \vartheta_{R_n}$. Since $\nu \ll P_{t_*}(x, \cdot) \ll \lambda$, for any $x \in C$, ν is absolute continuous with respect to the Lebesgue measure. We write $\nu(y)$ for the Lebesgue density of ν and note that for any $x \in C$, $\alpha \nu(y) \le p_{t_*}(x, y)$. Hence,

$$\begin{aligned}
E \int_{S_n}^{T_1} g(Y_s) ds = \\
&E \left(\int_{\mathbb{R}^d} \nu(y) dy \int_0^{t_*} ds \int_{\mathbb{R}^d} dz \frac{p_s(Y_{S_n}, z) p_{t_*-s}(z, y)}{p_{t_*}(Y_{S_n}, y)} g(z) \right) \\
&\quad + \int_{\mathbb{R}^d} \nu(y) dy E_y \left(\int_0^{T_C} g(Y_s) ds \right) \\
&\le \frac{1}{\alpha} E \left(\int_{\mathbb{R}^d} dy \int_0^{t_*} ds \int_{\mathbb{R}^d} dz \, p_s(Y_{S_n}, z) p_{t_*-s}(z, y) g(z) \right) \\
&\quad + \frac{1}{\alpha} E \left(\int_{\mathbb{R}^d} dy \, p_{t_*}(Y_{S_n}, y) E_y \left(\int_0^{T_C} g(Y_s) ds \right) \right).
\end{aligned}$$

Here, we have used that
$$\frac{\nu(y)}{p_{t_*}(Y_{S_n}, y)} \leq \frac{1}{\alpha}.$$
Concerning the first term on the right hand side above, by integrating out dy, we obtain

$$E\left(\int_{I\!R^d} dy \int_0^{t_*} ds \int_{I\!R^d} dz\, p_s(Y_{S_n}, z) p_{t_*-s}(z, y) g(z)\right)$$
$$= E\left(\int_0^{t_*} ds \int_{I\!R^d} p_s(Y_{S_n}, z) g(z) dz\right),$$

which is nothing else than

$$E\left(\int_{S_n}^{S_n+t_*} g(Y_s) ds\right) = E\left(\int_{S_n}^{R_n} g(Y_s) ds\right).$$

As a consequence,

$$E\int_{S_n}^{T_1} g(Y_s) ds \leq \frac{1}{\alpha} \sup_{x\in C} E_x \int_0^{T_C(t_*)} g(X_s) ds.$$

2. Now, we iterate the above argument. Write $T_n = \inf\{\tilde{S}_m > T_{n-1}\}, n \geq 2$. On $\{S_n = \tilde{S}_m\}$, let $V_n = U_{m+n}$, $n \geq 1$, be the successive choices of uniform random variables. Then
$$\begin{aligned}\int_{S_n}^{S_{n+1}} g(Y_s) ds &= \int_{S_n}^{T_1} g(Y_s) ds + \sum_{n\geq 1} 1_{\{V_1>\alpha,\ldots,V_n>\alpha\}} \left(\int_{T_n}^{T_{n+1}} g(Y_s) ds\right) \\ &\leq \int_{S_n}^{T_1} g(Y_s) ds + \sum_{n\geq 1} 1_{\{V_1>\alpha,\ldots,V_{n-1}>\alpha\}} \left(\int_{T_n}^{T_{n+1}} g(Y_s) ds\right).\end{aligned}$$
But $\{V_1 > \alpha, \ldots, V_{n-1} > \alpha\} \in \mathcal{F}^Y_{T_{n-1}}$ and
$$E\left(\int_{T_n}^{T_{n+1}} g(Y_s) ds \Big| \mathcal{F}^Y_{T_{n-1}}\right) \leq \sup_{x\in C} E_x \int_0^{T_C(t_*)} g(X_s) ds.$$
Hence,
$$\int_{S_n}^{S_{n+1}} g(Y_s) ds \leq$$
$$\leq \frac{1}{\alpha} \sup_{x\in C} E_x \int_0^{T_C(t_*)} g(X_s) ds + \sum_{n\geq 1} (1-\alpha)^{n-1} \sup_{x\in C} E_x \int_0^{T_C(t_*)} g(X_s) ds$$
$$= \frac{2}{\alpha} \sup_{x\in C} E_x \int_0^{T_C(t_*)} g(X_s) ds,$$
and this concludes the proof of (4.56).

Taking $g(x) = \|x\|^p$ for any $p = m - 2$ and $m < 2r + 2$ – which means that $p < 2r$ – the claim now follows from (4.55). •

4.2 Polynomial ergodicity for multidimensional diffusions

Grant the general assumptions of Section 4 and condition (4.52). Then the same arguments as those exposed in Section 3.3 show that

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} \leq \tilde{P}_{x,y}(\tau > t) \leq \frac{\tilde{E}(\tau^l)}{t^l}, \tag{4.57}$$

where

$$\tau = \inf\{T_n + t_* : U_n \leq \alpha\}, \quad T_n = \inf\{t > T_{n-1} + t_* : X_t^x \in C \text{ and } X_t^y \in C\}, n \geq 1, T_0 := 0.$$

We have the following proposition.

Proposition 4.12 (Polynomial ergodicity) *Grant condition (4.52) with* $r > 2$. *Then for all* $k < r - 1$ *and for all* $p \in]4 + 2k, 2r + 2[$, *we have*

$$E_{x,y} T_1^{k+1} \leq C(1 + \|x\|^p + \|y\|^p).$$

As a consequence, integrating against the invariant measure and observing that $\|\cdot\|^p$ *is integrable for all* $p < 2r$, *we obtain that*

$$\|P_t(x,\cdot) - \mu\|_{TV} \leq Ct^{-(k+1)}(1 + \|x\|^p)$$

for all $k < r - 2$, $p \in]4 + 2k, 2r[$.

Proof The proof is the same as the proof of Proposition 3.8, using the Itô-formula for the function $f(t,x,y) = (1+t)^k(\|x\|^m + \|y\|^m)$. Here we choose $m \in]2, p-2-2k[$, which is possible by our choice of k and p. The only difference is then the control of the expression in (3.39). In what follows we show how to handle this term.

As in the proof of Lemma 3.12, we write $p = m\tilde{p}$, $\tilde{p} > 1$, and let $\tilde{q} > 1$ such that $\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1$. Use Hölder's inequality

$$\begin{aligned} E_x \|X_s\|^m 1_{\{\|X_s\|^2 \geq \varepsilon(1+s)\}} &\leq [E_x\|X_s\|^p]^{1/\tilde{p}} \left[P_x(\|X_s\|^2 \geq \varepsilon(1+s))\right]^{1/\tilde{q}} \\ &\leq [E_x\|X_s\|^p]^{1/\tilde{p}} \left(\frac{E_x\|X_s\|^p}{[\varepsilon(1+s)]^{p/2}}\right)^{1/\tilde{q}} \\ &\leq C(1+s)^{-\frac{p}{2\tilde{q}}} E_x\|X_s\|^p = C(1+s)^{-\frac{1}{2}(p-m)} E_x\|X_s\|^p. \end{aligned}$$

Here, we have used that $p/\tilde{q} = p - m$, by definition of $\tilde{p}$.

Apply now the Itô-formula to $\|X_t\|^p$ and use (4.54) with $f(x) = \|x\|^p$ for all $\|x\| \geq M_0$. This yields the following quite rough upper bound

$$E_x\|X_t\|^p \leq \|x\|^p + CE_x \int_0^t 1_C(X_s)ds \leq \|x\|^p + Ct \leq C(1 + \|x\|^p)(1+t),$$

where C is some constant. We use this upper bound and obtain that

$$E_x\|X_s\|^m 1_{\{\|X_s\|^2 \geq \varepsilon(1+s)\}} \leq C(1 + \|x\|^p)(1+s)^{-\frac{1}{2}(p-m)+1}.$$

Multiplying with $(1+s)^{k-1}$ and integrating with respect to ds, we obtain a finite expression if and only if $k - \frac{1}{2}(p-m) < -1$, which is equivalent to $m < p-2-2k$ which is guaranteed by our choice of m. As a consequence,

$$\int_0^\infty E_x \left((1+s)^{k-1} \|X_s\|^m 1_{\{\|X_s\|^2 \geq \varepsilon(1+s)\}} \right) ds \leq C(1 + \|x\|^p),$$

and then we may conclude our proof as in the proof of Proposition 3.8. •

5 Recurrence for degenerate diffusions

The results that we have presented in the two preceding sections are based on the essential assumption of ellipticity (even uniform ellipticity in the last section). Under ellipticity and convenient smoothness assumptions on drift and diffusion coefficient, the process possesses continuous transition densities $p_t(x, y)$ for any fixed time t which are strictly positive. As a consequence, any compact set is petite. If we drop these basic assumptions, things get considerably more complicated.

In this section we give an outline of what might work even in a complete degenerate situation, i.e. when $\sigma\sigma^*(x)$ is nowhere strictly positive. Consider the equation (4.45)

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x \in I\!R^d,$$

where W is some $m-$dimensional Brownian motion. Throughout this section we shall use the following notation. For $y_0 \in I\!R^d$ and $\delta > 0$ we denote by $B_\delta(y_0)$ the open ball of radius δ centered in y_0. For any open subset $A \subset I\!R^d$, $C_b^\infty(A)$ denotes the class of infinitely differentiable functions defined on A which are bounded together with all partial derivatives of any order.

The first difficulty in the degenerate case is to find sets where the local Doeblin condition (2.20) still holds. In the case of diffusions, it is natural to attack this problem by finding domains where the transition operator $P_t(x, dy)$ possesses strictly positive Lebesgue densities.

5.1 Local existence of strictly positive transition densities for degenerate diffusions having locally smooth coefficients

We impose the following conditions on b and σ.

1. σ is a measurable function from $I\!R^d$ to $I\!R^{d\times m}$ and b a continuous function from $I\!R^d$ to $I\!R^d$.

2. There exist strong solutions for (4.45) and we have pathwise uniqueness.

3. There exists a strictly growing sequence of compacts $K_n \subset I\!R^d$, $K_n \subset K_{n+1}$, such that the following holds. If the starting point x satisfies $x \in \bigcup_n K_n$, then we have

$$T_n := \inf\{t : X_t \notin K_n\} \to \infty \text{ almost surely as } n \to \infty.$$

   Therefore we can introduce

$$E = \bigcup_n K_n, \tag{5.58}$$

   which is the state space of the process.

4. We have local smoothness on each compact K_n : For all n,

$$\sigma \in C_b^\infty(K_n, I\!R^{d\times m}) \text{ and } b \in C_b^\infty(K_n, I\!R^d).$$

By convention, we write $\sigma = (\sigma_{ij})_{1\le i\le d, 1\le j\le m} = (\sigma^1, \ldots, \sigma^m)$, where σ^i is the $i-$th column of σ. Additionally to the above local smoothness conditions, we need a non-degeneracy condition which is weaker than ellipticity. It is classical to rely on the so-called Hörmander condition. In order to state the Hörmander condition, we have to rewrite equation (4.45) in the Stratonovitch sense. That means, we replace the drift function b by $\tilde b$ defined as

$$\tilde b^i(x) = b^i(x) - \frac{1}{2}\sum_{k=1}^d \sum_{j=1}^m \sigma_{kj}(x) \frac{\partial \sigma_{ij}}{\partial x^k}(x), \; x \in E, \; 1 \le i \le d,$$

where we denote elements $x \in I\!R^d$ by $x = (x^1, \ldots, x^d)$.

Now we can introduce the successive Lie brackets. The Lie bracket of $f, g \in C^1(I\!R^d, I\!R^d)$ is defined by

$$[f,g]^i = \sum_{j=1}^d \left(f^j \frac{\partial g^i}{\partial x^j} - g^j \frac{\partial f^i}{\partial x^j} \right).$$

Then we construct by recurrence the sets of functions $L_0 = \{\sigma^1, \ldots, \sigma^m\}$ and for any $k \ge 1$,

$$L_{k+1} = \{[b,\phi], [\sigma^1, \phi], \ldots, [\sigma^m, \phi], \phi \in L_k\}.$$

Finally, for any $x \in E$ and any $\eta \in I\!R^d$, we define

$$\mathcal{V}_k(x,\eta) = \sum_{\phi \in L_0 \cup \ldots \cup L_k} < \phi(x), \eta >^2$$

and

$$\mathcal{V}_k(x) = \inf_{\eta : \|\eta\| = 1} \mathcal{V}_L(x,\eta) \wedge 1. \tag{5.59}$$

We assume:

Assumption 5.1 *There exists y_0 with $B_{5R}(y_0) \subset E$ and some $k \ge 0$ such that the following local Hörmander condition holds:*

$$\textit{We have } \mathcal{V}_k(y) \ge c(y_0, R) > 0 \textit{ for all } y \in B_{3R}(y_0).$$

Remark 5.2 *Notice that if the local Hörmander condition holds with $k = 0$, then this means that $a = \sigma\sigma^*$ satisfies $a(y_0) > 0$ which is the ellipticity assumption at y_0. So Hörmander's condition is much weaker than the ellipticity assumption, since the drift coefficient and the successive Lie brackets are as well used in order to obtain a strictly positive quadratic form.*

We have the following result.

Theorem 5.3 *Grant the above conditions. Then for any initial condition $x \in E$ and for any t, the random variable X_t admits a Lebesgue density $p_t(x,y)$ on $B_R(y_0)$ which is continuous with respect to $y \in B_R(y_0)$. Moreover, for any fixed $y \in B_R(y_0)$, $E \ni x \to p_{0,t}(x,y)$ is lower semi-continuous.*

We shall not give the proof of this theorem here and refer the interested reader to Höpfner, Löcherbach and Thieullen (2012). The main ingredient of the proof is an estimate of the Fourier transform of the law of X_t which is based on ideas of Fournier (2008), Bally (2007) and De Marco(2011).

Corollary 5.4 *Suppose that in addition to the above conditions, for some $t_* > 0$,*

$$\int_{B_R(y_0)} \left(\inf_{x \in B_R(y_0)} p_{t_*}(x, y) \right) dy > 0. \tag{5.60}$$

Then the local Doeblin condition (2.20) holds with $C = B_R(y_0)$.

Remark 5.5 *In order to show (5.60) one uses classically the support theorem for diffusions and arguments from control theory.*

5.2 Control of the return times

Under condition (5.60), put $C = B_R(y_0)$. In order to establish Harris-recurrence of the process X, we have to find a control of the return times to the set C. This is done by means of Lypunov functions. Finding a good Lyapunov function is actually a quite difficult task and supposes that one has already a good understanding of the dynamics of the process. So in what follows we somehow cheat a little bit since we impose a theoretical condition – in practice, the Lyapunov function that should be used depends very much on the concrete model that one is studying.

We impose conditions implying polynomial ergodicity. Of course, depending on the concrete model one is interested in, other conditions might hold, and we refer the interested reader for example to Mattingly, Stuart and Higham (2002) for the excellent exposition of models having degenerate noise but exponential ergodicity due to the existence of a performant Lyapunov function.

Here are our conditions for polynomial recurrence, in spirit of those given in Veretennikov (1997). First of all we suppose that there exists some constant $r > 0$ such that

$$< b(x), x > \leq -r \text{ for all } x \in C^c. \tag{5.61}$$

Moreover, we need the following condition on the diffusion coefficient. There exist $\Lambda > 0$ and $\lambda_+ > 0$ such that

$$tra(x) \leq d\Lambda \text{ and } < a(x)x, x > \leq \lambda_+ \|x\|^2 \text{ for all } x \in C^c. \tag{5.62}$$

Note that on $C = B_R(y_0)$, by our local smoothness assumption, drift and diffusion coefficient are bounded. We then have the following control.

Corollary 5.6 *Grant all assumptions of Section 5.1, (5.61) and (5.62). If $2r > d\Lambda$, then for any $m > 2$ such that $(m-2)\lambda_+ < 2r - d\Lambda$, the function f defined by $f(x) = \|x\|^m$ for all $x \in C^c$, f something smooth on C, is a generalized Lyapunov function and*

$$Af(x) \leq -c_1 f(x)^{1-\frac{2}{m}} + c_2 1_C(x).$$

In particular, Theorem 4.9 remains valid, and the process is recurrent of order p for any $p < 1 + \frac{2r - d\Lambda}{2\lambda_+}$.

6 Appendix : Some basic facts on Harris recurrent Markov processes and Nummelin splitting

In the preceding sections, we were mainly interested in developing tools ensuring that a given (diffusion) process is Harris recurrent. Moreover we were interested in controlling the speed of convergence to equilibrium.

This appendix is devoted to a more theoretical study of the properties of general Harris recurrent processes - once we know already that they *are* recurrent. Two important points will be considered. The first important point is that for Harris recurrent processes, we can always introduce regeneration times which introduce an i.i.d. structure, following the well-known technique developed by Nummelin (1978) and Athreya and Ney (1978). The second point is that we can use this i.i.d. structure in order to prove one of the most important theorems that are known to hold true for Harris processes, the ergodic theorem or ratio limit theorem that we have stated in Theorem 1.6 above.

We recall the general setup we are working in: We consider a probability space $(\Omega, \mathcal{A}, (P_x)_x)$. $X = (X_t)_{t\geq 0}$ is defined on $(\Omega, \mathcal{A}, (P_x)_x)$. It is a strong Markov process, taking values in a locally compact Polish space $(E, \mathcal{E})$, with càdlàg paths. $(P_x)_{x\in E}$ is a collection of probability measures on $(\Omega, \mathcal{A})$ such that $X_0 = x$ P_x-almost surely. We write $(P_t)_t$ for the transition semigroup of X. Moreover, we shall write $(\mathcal{F}_t)_t$ for the filtration generated by the process.

We suppose throughout this section that X is positive Harris recurrent, with invariant probability measure μ. Moreover, we impose the following condition on the transition semigroup $(P_t)_t$ of X:

Assumption 6.1 *There exists a sigma-finite positive measure Λ on $(E, \mathcal{E})$ such that for every $t > 0$, $P_t(x, dy) = p_t(x, y)\Lambda(dy)$, where $(t, x, y) \mapsto p_t(x, y)$ is jointly measurable.*

Remark 6.2 *The above assumption is always satisfied with $\Lambda = \mu$ if the process is strong Feller.*

A trick which is quite often used in the theory of processes in continuous time is to observe the continuous time process after independent exponential times. This gives rise to the resolvent chain and allows to use known results in discrete time instead of working with the continuous time process.

Write $U^1(x, dy) := \int_0^\infty e^{-t} P_t(x, dy)dt$ for the resolvent kernel associated to the process. Introduce a sequence $(\sigma_n)_{n\geq 1}$ of i.i.d. $exp(1)$-waiting times, independent of the process X itself. Let $T_0 = 0$, $T_n = \sigma_1 + \ldots + \sigma_n$ and $\bar{X}_n = X_{T_n}$. Then the chain $\bar{X} = (\bar{X}_n)_n$ is recurrent in the sense of Harris, having the same invariant measure μ as the continuous time process, and its one-step transition kernel is given by $U^1(x, dy)$.

Since X is Harris, it can be shown, see Revuz (1984), see also Proposition 6.7 of Höpfner and Löcherbach (2003), that there exist $\alpha \in]0, 1]$ and a measurable set C such that

$$U^1(x, dy) \geq \alpha 1_C(x)\nu(dy), \tag{6.63}$$

where $\mu(C) > 0$ and ν a probability measure equivalent to $\mu(\cdot \cap C)$.

Remark 6.3 *The above lower bound is actually the most important tool in what follows. (6.63) holds* **for any Harris recurrent** *process, without any further assumption. Notice that since $\mu(C) > 0$, by recurrence, the process returns to C in finite time always surely. We will use this local Doeblin condition for the resolvant chain in order to introduce regeneration times for the process.*

6.1 Nummelin splitting and regeneration times

The main assertion of this section is the following:

Regeneration times can be introduced for any Harris recurrent strong Markov process under the Assumption 6.1 – without any further assumption.

We show how to do this: We construct regeneration times in continuous time by using the technique of Nummelin splitting which has been introduced for Harris recurrent Markov chains in discrete time by Nummelin (1978) and Athreya and Ney (1978). The idea is basically the same as what we have discussed in the beginning of Section 4, except that now we work with the resolvant kernel U^1 rather than with P_{t_*}. We recall the details of the construction from Löcherbach and Loukianova (2008).

We define on an extension of the original space $(\Omega, \mathcal{A}, (P_x))$ a Markov process $Z = (Z_t)_{t\geq 0} = (Z_t^1, Z_t^2, Z_t^3)_{t\geq 0}$, taking values in $E \times [0,1] \times E$ such that the times T_n are jump times of the process and such that $((Z_t^1)_t, (T_n)_n)$ has the same distribution as $((X_t)_t, (T_n)_n)$.

First of all we define the split kernel $Q((x,u), dy)$, a transition kernel $Q((x,u), dy)$ from $E \times [0,1]$ to E, which is defined by

$$Q((x,u),dy) = \begin{cases} \nu(dy) & \text{if } (x,u) \in C \times [0,\alpha] \\ \frac{1}{1-\alpha}\big(U^1(x,dy) - \alpha\nu(dy)\big) & \text{if } (x,u) \in C \times]\alpha, 1] \\ U^1(x,dy) & \text{if } x \notin C. \end{cases} \tag{6.64}$$

Remark 6.4 *This kernel is called split kernel since $\int_0^1 du Q((x,u),dy) = U^1(x,dy)$. Thus Q is a splitting of the resolvent kernel by means of the additional "color" u.*

Write $u^1(x,x') := \int_0^\infty e^{-t} p_t(x,x') dt$. We now show how to construct the process Z recursively over time intervals $[T_n, T_{n+1}[, n \geq 0$. We start with some initial condition $Z_0^1 = X_0 = x$, $Z_0^2 = u \in [0,1]$, $Z_0^3 = x' \in E$. Then inductively in $n \geq 0$, on $Z_{T_n} = (x,u,x')$:

1. Choose a new jump time σ_{n+1} according to

$$e^{-t} \frac{p_t(x,x')}{u^1(x,x')} dt \text{ on } I\!R_+,$$

where we define $0/0 := a/\infty := 1$, for any $a \geq 0$, and put $T_{n+1} := T_n + \sigma_{n+1}$.

2. On $\{\sigma_{n+1} = t\}$, put $Z_{T_n+s}^2 := u$, $Z_{T_n+s}^3 := x'$ for all $0 \leq s < t$.

3. For every $s < t$, choose

$$Z_{T_n+s}^1 \sim \frac{p_s(x,y)p_{t-s}(y,x')}{p_t(x,x')} \Lambda(dy).$$

Here, by convention we choose $Z^1_{T_n+s} := x_0$ for some fixed point $x_0 \in E$ on $\{p_t(x,x') = 0\}$. Moreover, given $Z^1_{T_n+s} = y$, on $s+u < t$, choose

$$Z^1_{T_n+s+u} \sim \frac{p_u(y,y')p_{t-s-u}(y',x')}{p_{t-s}(y,x')}\Lambda(dy').$$

Again, on $\{p_{t-s}(y,x') = 0\}$, choose $Z^1_{T_n+s+u} = x_0$.

4. At the jump time T_{n+1}, choose $Z^1_{T_{n+1}} := Z^3_{T_n} = x'$. Choose $Z^2_{T_{n+1}}$ independently of $Z_s, s < T_{n+1}$, uniformly on $[0,1]$. Finally, on $\{Z^2_{T_{n+1}} = u'\}$, choose $Z^3_{T_{n+1}} \sim Q((x',u'),dx'')$.

Note that by construction, given the initial value of Z at time T_n, the evolution of the process Z^1 during $[T_n, T_{n+1}[$ does not depend on the chosen value of $Z^2_{T_n}$.

We will write P_π for the measure related to X, under which X starts from the initial measure $\pi(dx)$, and $\mathbb{P}_\pi$ for the measure related to Z, under which Z starts from the initial measure $\pi(dx) \otimes U(du) \otimes Q((x,u),dy)$, where $U(du)$ is the uniform law on $[0,1]$. Hence, $\mathbb{P}_{x_0}$ denotes the measure related to Z under which Z starts from the initial measure $\delta_{x_0}(dx) \otimes U(du) \otimes Q((x,u),dy)$. In the same spirit we denote E_π the expectation with respect to P_π and $\mathbb{E}_\pi$ the expectation with respect to $\mathbb{P}_\pi$. Moreover, we shall write $\mathbb{F}$ for the filtration generated by Z, $\mathbb{G}$ for the filtration generated by the first two coordinates Z^1 and Z^2 of the process, and $\mathbb{F}^X$ for the sub-filtration generated by X interpreted as first coordinate of Z.

6.2 Basic properties of Z

The new process Z is a Markov process with respect to its filtration $\mathbb{F}$. For a proof of this result, the interested reader is referred to Theorem 2.7 of Löcherbach and Loukianova (2008). In general, Z will no longer be strong Markov. But for any $n \geq 0$, by construction, the strong Markov property holds with respect to T_n. Thus for any $f, g : E \times [0,1] \times E \to \mathbb{R}$ measurable and bounded, for any $s > 0$ fixed, for any initial measure π on $(E, \mathcal{E})$,

$$\mathbb{E}_\pi(g(Z_{T_n})f(Z_{T_n+s})) = \mathbb{E}_\pi(g(Z_{T_n})\mathbb{E}_{Z_{T_n}}(f(Z_s))).$$

Finally, an important point is that by construction,

$$\mathcal{L}((Z^1_t)_t|\mathbb{P}_x) = \mathcal{L}((X_t)_t|P_x)$$

for any $x \in E$, thus the first coordinate of the process Z is indeed a copy of the original Markov process X, when disregarding the additional colors (Z^2, Z^3).

However, adding the colors (Z^2, Z^3) allows to introduce regeneration times for the process Z (not for X itself). More precisely, write

$$A := C \times [0,\alpha] \times E$$

and put

$$S_0 := 0,\ R_0 := 0, S_{n+1} := \inf\{T_m > R_n : Z_{T_m} \in A\}, R_{n+1} := \inf\{T_m : T_m > S_{n+1}\}. \tag{6.65}$$

The sequence of $\mathbb{F}-$stopping times R_n generalizes the notion of life-cycle decomposition in the following sense.

Proposition 6.5 *[Proposition 2.6 and 2.13 of Löcherbach and Loukianova (2008)]*
a) Under $I\!P_x$, the sequence of jump times $(T_n)_n$ is independent of the first coordinate process $(Z^1_t)_t$ and $(T_n - T_{n-1})_n$ are i.i.d. exp(1)−variables.
b) At regeneration times, we start from a fixed initial distribution which does not depend on the past: $Z_{R_n} \sim \nu(dx)U(du)Q((x,u),dx')$ for all $n \geq 1$.
c) At regeneration times, we start afresh and have independence after a waiting time: $Z_{R_n+\cdot}$ is independent of $\mathcal{F}_{S_n-}$ for all $n \geq 1$.
d) The sequence of $(Z_{R_n})_{n\geq 1}$ is i.i.d.

Since the original process X is Harris with invariant measure μ, the new process Z will be Harris, too. (This seems to be obvious but has to be proved, see Proposition 6.8 below.) To begin with, the following proposition gives the exact formula for the invariant measure Π of Z.

Proposition 6.6 *Let $f : E \times [0,1] \times E \to I\!R_+$ be a measurable positive function. Then the invariant probability measure Π of Z is given by*

$$\Pi(f) = \int \mu(dx) \int_0^1 du \int Q((x,u),dz) \int \Lambda(dy) u^1(x,y) \frac{u^1(y,z)}{u^1(x,z)} f(y,u,z). \tag{6.66}$$

In other words,

$$\Pi(dy,du,dz) = \Lambda(dy)U(du) \left(\int \mu(dx)\; u^1(x,y) \frac{u^1(y,z)}{u^1(x,z)} Q((x,u),dz) \right).$$

The proof of (6.66) is given in Subsection 6.4 below.

Remark 6.7 *1. Suppose that $f = f(y,z)$ does not depend on u. Then, since we have $\int_0^1 Q((x,u),dz)du = U^1(x,dz)$ and since $\mu U^1 = \mu$,*

$$\begin{aligned}\Pi(f) = & \int \mu(dx) \int \Lambda(dz) u^1(x,z) \int \Lambda(dy) u^1(x,y) \frac{u^1(y,z)}{u^1(x,z)} f(y,z) \\ & = \int \mu(dy) \int U^1(y,dz) f(y,z).\end{aligned}$$

Hence the first and third marginal of Π is, as expected, $\int_0^1 \Pi(dy,du,dz) = \mu(dy)U^1(y,dz)$.
2. How can we interpret the above form of the invariant measure of the process Z? We have to understand what is the equilibrium behavior of Z_t. By the construction of Z over successive time intervals $[T_n, T_{n+1}[$, the law of Z_t under equilibrium will depend on the last jump time T_n before time t and the corresponding state of the process (x,u,z).

The delay time $t - T_{N_t}$ between the last jump before time t and time t is an exp (1) time truncated at time t if we are not already in equilibrium. Under equilibrium, however, the delay between last jump and time t is exactly an exponential time without truncation. So the process has started some exponential time ago in the past. At that time, it started in its initial equilibrium measure which is the equilibrium measure of the chain $(Z_{T_n})_n$, i.e. $\mu(dx)U(du)Q((x,u),dz)$. The second and third coordinate do not change during inter-jump intervals, hence they are still the same at time t. The first coordinate of the process, though, has evolved according to the transition operator of the original Markov process – this is the term $u^1(x,y)$ – conditioned on the arrival position z after one exponential period, which is the term $\frac{u^1(y,z)}{u^1(x,z)}$.

6.3 Regeneration, Chacon-Ornstein's ratio limit theorem and Harris recurrence of Z

One of the main results that hold for Harris recurrent processes is the ratio limit theorem, see Theorem 1.6. In what follows we give a proof of this theorem for the process Z. As we will see, this follows from the strong Markov property of Z with respect to the jump times T_n and from the regenerative structure of the times R_n.

Proposition 6.8 (Ratio limit theorem) *Let $f, g : E \times [0,1] \times I\!E \to I\!R_+$ be positive measurable functions such that $\Pi(g) > 0$. Then*

$$\frac{\int_0^t f(Z_s)ds}{\int_0^t g(X_s)ds} \to \frac{\Pi(f)}{\Pi(g)} \quad I\!P_x - \text{ almost surely, as } t \to \infty,$$

for any $x \in E$. Moreover, Z is recurrent in the sense of Harris and its unique invariant probability measure Π is given by

$$\Pi(f) = \ell \, I\!E_\pi \int_{R_1}^{R_2} f(Z_s)ds, \tag{6.67}$$

where $\ell = I\!E(R_2 - R_1)^{-1} > 0$.

In particular, applying the above Proposition to $\int_0^t f(Z_s)ds$ for functions $f(y,u,z) = f(y)$ depending only on the first variable gives a proof of Theorem 1.6, item 2., quoted in the first section.

Proof: Let $N_t = \sup\{n : R_n \le t\}$ the number of regeneration epochs before time t and $A_t = \int_0^t f(Z_s)ds$.

Firstly, we show that the following convergence holds place almost surely, for any starting point x :

$$\frac{A_t}{N_t} \to I\!E(A_{R_2} - A_{R_1}) \quad I\!P_x - \text{ almost surely, as } t \to \infty. \tag{6.68}$$

This is seen as follows. Put $\xi_n = \int_{R_{n-1}}^{R_n} f(Z_s)ds, n \ge 1$, which is the increment of the additive functional over one regeneration epoch. By construction, the increments over even regeneration epochs $\xi_{2n}, n \ge 1$, are i.i.d. The same is true for increments over odd regeneration epochs, if we exclude the first epoch between 0 and R_1. So the $\xi_{2n+1}, n \ge 1$, are i.i.d., as well. Hence, applying the strong law of large numbers separately to the even and the odd increments, we have that

$$\frac{A_{R_n}}{n} \Longrightarrow I\!E(A_{R_2} - A_{R_1}) \quad I\!P_x - \text{ almost surely, as } n \to \infty.$$

Hence

$$\frac{A_{R_{N_t}}}{N_t} \to I\!E(A_{R_2} - A_{R_1}) \quad I\!P_x - \text{ almost surely, as } t \to \infty.$$

Using that

$$\lim_{t\to\infty} \frac{A_{R_{N_t}}}{N_t} \le \lim_{t\to\infty} \frac{A_t}{N_t} \le \lim_{t\to\infty} \frac{A_{R_{N_t+1}}}{N_t},$$

by the positivity of A_t, we obtain (6.68). As a consequence, with $B_t = \int_0^t g(Z_s)ds$,

$$\frac{A_t}{B_t} \to \frac{\mathbb{E}(A_{R_2} - A_{R_1})}{\mathbb{E}(B_{R_2} - B_{R_1})} \quad \mathbb{P}_x - \text{ almost surely, as } t \to \infty.$$

Next, for any positive measurable function f such that $\mathbb{E} \int_{R_1}^{R_2} f(Z_s)ds$ is finite, put

$$\tilde{\Pi}(f) := \mathbb{E} \int_{R_1}^{R_2} f(Z_s)ds.$$

Note that we could also write

$$\tilde{\Pi}(f) = \mathbb{E}_\nu \int_0^{R_1} f(Z_s)ds.$$

We show that $\tilde{\Pi}$ is invariant for the process Z. For that sake let Q_t be the transition operator of Z. Let f be any positive measurable function. Then we have, by the simple Markov property,

$$\begin{aligned}
\tilde{\Pi}Q_t(f) &= \mathbb{E}_\nu \int_0^{R_1} Q_t f(Z_s)ds \\
&= \mathbb{E}_\nu \int_0^\infty 1_{\{s<R_1\}} \mathbb{E}_{Z_s} f(Z_t)ds \\
&= \int_0^\infty \mathbb{E}_\nu (1_{\{s<R_1\}} f(Z_{t+s}))ds \\
&= \mathbb{E}_\nu \int_t^{R_1} f(Z_s)ds + \mathbb{E}_\nu \int_{R_1}^{R_1+t} f(Z_s)ds.
\end{aligned}$$

Now we use the regeneration property with respect to R_1 that allows to conclude :

$$\mathbb{E}_\nu \int_{R_1}^{R_1+t} f(Z_s)ds = \mathbb{E}_\nu \int_0^t f(Z_s)ds,$$

and thus

$$\tilde{\Pi}Q_t f = \tilde{\Pi} f.$$

Thus $\tilde{\Pi}$ is invariant for Z.

Moreover, it is evident that $\tilde{\Pi}(A) > 0$ implies that $\int_0^\infty 1_A(Z_t)dt = \infty$ almost surely. Indeed, suppose that $\int_0^\infty 1_A(Z_t)dt < \infty$ on a set B of positive probability. Hence for any $\omega \in B$, $\lim 1_A(Z_t) = 0$, which means:

$$\exists\, t_0 = t_0(\omega) : \ \forall t \geq t_0,\ Z_t \notin A. \tag{6.69}$$

But $\tilde{\Pi}(A) > 0$ implies that for all n, $\int_{R_n}^{R_{n+1}} 1_A(Z_s)ds > 0$ with positive probability. Now using the Borel-Cantelli lemma separately along even regeneration epochs and along odd regeneration epochs, we obtain that $\int_{R_n}^{R_{n+1}} 1_A(Z_s)ds > 0$ for infinitely many n, almost surely. In particular, this implies: for infinitely many n, there exists $t \in [R_n, R_{n+1}[$, such that $Z_t \in A$. Since $R_n \to \infty$, this is in contradiction with (6.69). This means that the process Z is Harris with invariant measure $\tilde{\Pi}$. Then, by uniqueness of the invariant measure for Harris recurrent processes, $\tilde{\Pi} = c\,\Pi$. Since Π is a probability, $c = \mathbb{E}(R_2 - R_1)$. This concludes the proof. •

Since the process is positive recurrent, the expected length $\ell = I\!E(R_2 - R_1)^{-1}$ of one regeneration period is finite. Moreover, μ is the projection onto the first coordinate of Π. From this we deduce that the invariant probability measure μ of the original process X must be given by

$$\mu(f) = \ell \, I\!E_\pi \int_{R_1}^{R_2} f(X_s)ds. \tag{6.70}$$

In the above formula we interpret X as first coordinate of Z, under $I\!P_\pi$. Hence, Proposition 6.8 implies that

$$\frac{1}{t}\int_0^t f(X_s)ds \to \mu(f) \quad P_x - a.s. \, \forall \, x$$

as $t \to \infty$. This provides a proof of item 2. of Theorem 1.6.

We close this section with the proof of item 1. of Theorem 1.6.

Proof of Theorem 1.6 - continued We prove item 1. of Theorem 1.6 in the case when X is positive recurrent. It suffices to show that $\frac{1}{t}E_x \int_0^t f(X_s)ds \to \mu(f)$, for any positive function $f \in L^1(\mu)$.

We have, compare also to Remark 3.2 in Löcherbach and Loukianova (2012), since $f \geq 0$,

$$\int_0^t f(X_s)ds \leq \int_0^{R_1} f(X_s)ds + \sum_{n\geq 1} 1_{\{R_{n-1}\leq t\}} \int_{R_n}^{R_{n+1}} f(X_s)ds.$$

Taking conditional expectation with respect to $\mathcal{F}_{R_{n-1}}$ and using $E\int_{R_n}^{R_{n+1}} f(X_s)ds = \frac{1}{\ell}\mu(f)$ and $\sum_n P_x(R_{n-1} \leq t) = E_x(N_t + 1)$, $N_t = \sup\{n : R_n \leq t\}$, we obtain

$$E_x(\int_0^t f(X_s)ds) \leq E_x \int_0^{R_1} f(X_s)ds + \frac{1}{\ell}\mu(f)E_x(N_t + 1).$$

Since

$$\ell E \int_{R_n}^{R_{n+1}} f(X_s)ds = \ell E_\nu(\int_0^{R_1} f(X_s)ds) = \mu(f) < \infty,$$

we have that

$$E_x \int_0^{R_1} f(X_s)ds < \infty \quad \nu(dx) - a.s.,$$

hence $\mu(dx)-$almost surely, since $\nu \sim \mu(\cdot \cap C)$. Therefore,

$$\limsup_{t\to\infty} \frac{1}{t}E_x \int_0^t f(X_s)ds \leq \mu(f) \quad \mu(dx) - a.s.,$$

since

$$\lim_{t\to\infty} \frac{1}{t}E_x(N_t) = \ell.$$

Now, the assertion follows as an application of Fatou's lemma: Knowing that we have $\lim \frac{1}{t}\int_0^t f(X_s)ds = \mu(f)$ P_x-almost surely for any x, we deduce

$$\mu(f) \leq \liminf_{t\to\infty} \frac{1}{t}E_x \int_0^t f(X_s)ds \leq \limsup_{t\to\infty} \frac{1}{t}E_x \int_0^t f(X_s)ds \leq \mu(f),$$

$\mu(dx)-$almost surely. This implies the result. •

6.4 Proof of (6.66)

This section is devoted to the proof of (6.66). We start with the following formula: let $f : E\times[0,1]\times E \to I\!R_+$ be a measurable positive function. Then, writing $t_n := s_1+\ldots+s_n$, it can be shown that

$$
\begin{aligned}
& Q_t f(x,u,z) = I\!E(f(Z_t)|Z_0=(x,u,z)) \qquad (6.71)\\
& = e^{-t}\int p_t(x,y)\frac{u^1(y,z)}{u^1(x,z)}\Lambda(dy)f(y,u,z)\\
& +\sum_{n\geq 1}\int_0^t ds_1 e^{-s_1}\frac{p_{s_1}(x,z)}{u^1(x,z)}\int_0^t ds_2 e^{-s_2}\ldots\int_0^t ds_n e^{-s_n}1_{\{t_n\leq t\}}\int p_{s_2+\ldots+s_n}(z,x_n)\Lambda(dx_n)\\
& \int_0^1 du_n\int Q((x_n,u_n),dz_n)e^{-(t-t_n)}\int p_{t-t_n}(x_n,y)\frac{u^1(y,z_n)}{u^1(x_n,z_n)}\Lambda(dy)f(y,u_n,z_n).
\end{aligned}
$$

This follows directly from the definition of the process Z.

We now prove (6.66). For technical reasons, in the following we replace the variables x by x_0 and y by x. Since $\int_0^1 \Pi(dx,du,dz)=\mu(dx)U^1(x,dz)$, integrating (6.71) against Π yields

$$
\begin{aligned}
I\!E_\Pi f(Z_t) = & \int\mu(dx_0)\int_0^1 du\int Q((x_0,u),dz)\int\Lambda(dx)u^1(x_0,x)\frac{u^1(x,z)}{u^1(x_0,z)}\\
& e^{-t}\int p_t(x,y)\frac{u^1(y,z)}{u^1(x,z)}\Lambda(dy)f(y,u,z)\\
& +\sum_{n\geq 1}\int_0^t ds_1 e^{-s_1}\int_0^t ds_2 e^{-s_2}\ldots\int_0^t ds_n e^{-s_n}1_{\{t_n\leq t\}}e^{-(t-t_n)}\\
& \int\mu(dx_n)\int_0^1 du_n\int Q((x_n,u_n),dz_n)\int p_{t-t_n}(x_n,y)\frac{u^1(y,z_n)}{u^1(x_n,z_n)}\Lambda(dy)f(y,u_n,z_n)\\
=: & \, A+B.
\end{aligned}
$$

The first term can be rewritten as follows :

$$
\begin{aligned}
A = & \int\mu(dx_0)\int_0^1 du\int Q((x_0,u),dz)\int\Lambda(dx)u^1(x_0,x)\frac{u^1(x,z)}{u^1(x_0,z)}\\
& e^{-t}\int p_t(x,y)\frac{u^1(y,z)}{u^1(x,z)}\Lambda(dy)f(y,u,z) =\\
& \int\mu(dx_0)\int_0^1 du\int Q((x_0,u),dz)\int\Lambda(dy)[e^{-t}\int\Lambda(dx)u^1(x_0,x)p_t(x,y)]\\
& \frac{u^1(y,z)}{u^1(x_0,z)}f(y,u,z).
\end{aligned}
$$

But by Markov's property,

$$
\begin{aligned}
e^{-t}\int\Lambda(dx)u^1(x_0,x)p_t(x,y) &= e^{-t}\int_0^\infty e^{-s}ds\int\Lambda(dx)p_s(x_0,x)p_t(x,y)\\
&= e^{-t}\int_0^\infty e^{-s}ds\, p_{t+s}(x_0,y) = \int_t^\infty e^{-s}p_s(x_0,y)ds.
\end{aligned}
$$

Thus we get for the fist term in $I\!E_\Pi f(Z_t)$ the following expression

$$A = \int \mu(dx_0) \int_0^1 du \int Q((x_0,u),dz) \int \Lambda(dy) \int_t^\infty e^{-s} p_s(x_0,y) ds \frac{u^1(y,z)}{u^1(x_0,z)} f(y,u,z).$$

Moreover, for B we get, since

$$\tilde{\mathcal{L}}(t - \tilde{T}_{\tilde{N}_t}; \tilde{N}_t > 0) = e^{-s} 1_{[0,t]}(s) ds,$$

$$B = \int \mu(dx_0) \int_0^1 du \int Q((x_0,u),dz) \int \Lambda(dx) \int_0^t e^{-s} ds\, p_s(x_0,x) \frac{u^1(x,z)}{u^1(x_0,z)} f(x,u,z).$$

Finally,

$$A + B = \int \mu(dx_0) \int_0^1 du$$
$$\int Q((x_0,u),dz) \int \Lambda(dx) \int_0^\infty e^{-s} ds\, p_s(x_0,x) \frac{u^1(x,z)}{u^1(x_0,z)} f(x,u,z) = \Pi(f),$$

since $\int_0^\infty e^{-s} ds\, p_s(x_0,x) = u^1(x_0,x)$. This concludes the proof of (6.66).

Acknowledgments

The author thanks Arturo Kohatsu-Higa, Reinhard Höpfner and Vlad Bally for careful reading and stimulating discussions.

References

[1] D.G. Aronson. Non-negative solutions of linear parabolic equations. *Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser.*, 22:607–694, 1968.

[2] K.B. Athreya and P. Ney. A new approach to the limit theory of recurrent Markov chains. *Trans. Am. Math. Soc.*, 245:493–501, 1978.

[3] J. Azéma, M. Duflo, and D. Revuz. Mesure invariante des processus de Markov recurrents. 1969.

[4] Vlad Bally. Integration by parts formula for locally smooth laws and applications to equations with jumps I. Preprints Institut Mittag-Leffler, The Royal Swedish Academy of Sciences., 2007.

[5] Stefano De Marco. Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions. *Ann. Appl. Probab.*, 21(4):1282–1321, 2011.

[6] B. Delyon. Compléments sur les martingales et la théorie ergodique, 2012.

[7] Randal Douc, Gersende Fort, and Arnaud Guillin. Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Processes Appl.*, 119(3):897–923, 2009.

[8] Randal Douc, Gersende Fort, Eric Moulines, and Philippe Soulier. Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.*, 14(3):1353–1377, 2004.

[9] P.A. Ferrari and A. Galves. Coupling and regeneration for stochastic processes, notes for a minicourse presented in xiii escuela venezolana de matematicas, 2000.

[10] P.J. Fitzsimmons and Jim Pitman. Kac's moment formula and the Feynman-Kac formula for additive functionals of a Markov process. *Stochastic Processes Appl.*, 79(1):117–134, 1999.

[11] Nicolas Fournier. Smoothness of the law of some one-dimensional jumping S.D.E.s with non-constant rate of jump. *Electron. J. Probab.*, 13:135–156, 2008.

[12] A. Friedman. *Partial differential equations of parabolic type.* Englewood Cliffs, N.J.: Prentice-Hall, Inc. XIV, 347 p. , 1964.

[13] Valentine Genon-Catalot, Thierry Jeantheau, and Catherine Larédo. Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli*, 6(6):1051–1079, 2000.

[14] R. Hoepfner. Mathematische statistik, 2008.

[15] R. Hoepfner and E. Loecherbach. Limit theorems for null recurrent Markov processes. *Mem. Am. Math. Soc.*, 768:92 p., 2003.

[16] R. Hoepfner and E. Loecherbach. On some ergodicity properties for time inhomogeneous markov processes with t-periodic semigroup, 2010.

[17] Reinhard Hoepfner, Eva Loecherbach, and Michèle Thieullen. Transition densities for stochastic Hodgkin-Huxley models. arxiv.org/abs/1207.0195, 2012.

[18] N. Ikeda, M. Nagasawa, and S. Watanabe. A construction of Markov processes by piecing out. *Proc. Japan Acad.*, 42:370–375, 1966.

[19] Rafail Khasminskii. *Stochastic stability of differential equations. With contributions by G. N. Milstein and M. B. Nevelson. 2nd completely revised and enlarged ed.* Stochastic Modelling and Applied Probability 66. Berlin: Springer. xvii, 339 p., 2012.

[20] D. Lépingle. Sur le comportement asymptotique des martingales locales. Semin. Probab. XII, Univ. Strasbourg 1976/77, Lect. Notes Math. 649, 148-161, 1978.

[21] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times. With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson.* Providence, RI: American Mathematical Society (AMS). xvii, 371 p. , 2009.

[22] Thomas M. Liggett. L_2 rates of convergence for attractive reversible nearest particle systems: The critical case. *Ann. Probab.*, 19(3):935–959, 1991.

[23] Eva Loecherbach and Dasha Loukianova. On Nummelin splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multi-dimensional diffusions. *Stochastic Processes Appl.*, 118(8):1301–1321, 2008.

[24] Eva Loecherbach and Dasha Loukianova. Polynomial deviation bounds for recurrent harris processes having general state space. *ESAIM: Probability and Statistics*, 2011.

[25] Eva Loecherbach and Dasha Loukianova. Deviation inequalities for centered additive functionals of recurrent Harris processes having general state space. *J. Theor. Probab.*, 25(1):231–261, 2012.

[26] Eva Loecherbach, Dasha Loukianova, and Oleg Loukianov. Polynomial bounds in the Ergodic theorem for one-dimensional diffusions and integrability of hitting times. *Ann. I.H.P.*, 47:425–449, 2011.

[27] Oleg Loukianov, Dasha Loukianova, and Shiqi Song. Spectral gaps and exponential integrability of hitting times for linear diffusions. *Ann. I.H.P.*, 47:679–698, 2011.

[28] J.C. Mattingly, A.M. Stuart, and D.J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Processes Appl.*, 101(2):185–232, 2002.

[29] Sean Meyn and Richard L. Tweedie. *Markov chains and stochastic stability. Prologue by Peter W. Glynn. 2nd ed.* Cambridge Mathematical Library. Cambridge: Cambridge University Press. xviii, 594 p. , 2009.

[30] James R. Norris and Daniel W. Stroock. Estimates on the fundamental solution to heat flows with uniformly elliptic coefficients. *Proc. Lond. Math. Soc., III. Ser.*, 62(2):373–402, 1991.

[31] E. Nummelin. A splitting technique for Harris recurrent Markov chains. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 43:309–318, 1978.

[32] Steven Orey. *Lecture notes on limit theorems for Markov chain transition probabilities.* Van Nostrand Reinhold Mathematical Studies. 34. London etc.: Van Nostrand Reinhold Company. VIII, 108 p., 1971.

[33] D. Revuz. *Markov chains. 2nd rev. ed.* North-Holland Mathematical Library, Vol. 11. Amsterdam - New York - Oxford: North-Holland. XI, 374 p., 1984.

[34] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion. 3rd ed., 3rd. corrected printing.* Grundlehren der Mathematischen Wissenschaften 293. Berlin: Springer. xi, 606 p., 2005.

[35] Emmanuel Rio. *Asymptotic theory of weakly dependent stochastic processes. (Théorie asymptotique des processus aléatoires faiblement dépendants.).* Mathématiques & Applications (Paris). 31. Paris: Springer. xii, 170 p., 2000.

[36] A.Yu. Veretennikov. On polynomial mixing bounds for stochastic differential equations. *Stochastic Processes Appl.*, 70(1):115–127, 1997.