Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 118 (2008) 1301–1321
stochastic processes and their applications
www.elsevier.com/locate/spa

On Nummelin splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multi-dimensional diffusions

Eva Löcherbach[a], Dasha Loukianova[b,*]

[a] Centre de Mathématiques, Faculté de Sciences et Technologie, Université Paris XII, 61 avenue du Général de Gaulle, 94010 Créteil Cedex, France

[b] Département de Mathématiques, Université d'Evry-Val d'Essonne, Bd François Mitterrand, 91025 Evry Cedex, France

Received 9 August 2006; received in revised form 29 June 2007; accepted 13 September 2007
Available online 20 September 2007

Abstract

We introduce a sequence of stopping times that allow us to study an analogue of a life-cycle decomposition for a continuous time Markov process, which is an extension of the well-known splitting technique of Nummelin to the continuous time case. As a consequence, we are able to give deterministic equivalents of additive functionals of the process and to state a generalisation of Chen's inequality. We apply our results to the problem of non-parametric kernel estimation of the drift of multi-dimensional recurrent, but not necessarily ergodic, diffusion processes.

MSC: 60F99; 60J35; 60J55; 60J60; 62G99; 62M05

Keywords: Harris recurrence; Nummelin splitting; Continuous time Markov processes; Resolvents; Special functions; Additive functionals; Chacon–Ornstein theorem; Diffusion process; Nadaraya–Watson estimator

1. Introduction

Consider a Harris recurrent strong Markov process $X = (X_t)_{t\geq 0}$ with invariant measure μ. If such a process has a recurrent point x_0 (or more generally a recurrent atom), then it is possible to introduce a sequence of stopping times R_n – which is called *life-cycle decomposition* – such that

* Corresponding author. Tel.: +33 169470220; fax: +33 145171649.
E-mail addresses: locherbach@univ-paris12.fr (E. Löcherbach), dasha.loukianova@univ-evry.fr (D. Loukianova).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2007.09.003

1. For all n, $R_n < \infty$, $R_{n+1} = R_n + R_1 \circ \theta_{R_n}$. (Here, θ denotes the shift operator.)
2. $X_{R_n} = x_0$.
3. For all n, the process $(X_{R_n+t})_{t\geq 0}$ is independent of $\mathcal{F}_{R_n}$.

In this case, paths of the process can be decomposed into i.i.d. excursions $[R_i, R_{i+1}[$, $i = 1, 2, \ldots$, plus an initial segment $[0, R_1]$, and then limit theorems such as the ratio limit theorem for additive functionals of the process follow immediately as a direct application of the strong law of large numbers, in both the ergodic and the null recurrent case.

For general Harris processes, at least without further assumptions, recurrent atoms do not exist. However, for discrete time Harris chains, Athreya and Ney, see [1], and Nummelin, see [21], give a way of constructing a recurrent atom on an extended probability space, provided the transition operator of the chain satisfies a certain minorisation condition. This construction is called "splitting". A well-known idea, see for instance Meyn and Tweedie [19,20], is to consider the discrete chain $\bar{X} = (\bar{X}_n)_n$, called the resolvent chain or R-chain, instead of the process in continuous time X. This resolvent chain is obtained when observing the process at independent exponential times. We propose to apply the splitting technique to this resolvent chain which is always possible. Hence, we use splitting at random times when sampling the process after independent exponential times. We then fill in the original process in between two successive exponential times. In other words, we construct bridges of the process X between exponential times such that at the exponential times, the splitting is satisfied. This construction is not evident since we want to preserve the Markov property of the process. It is for that reason that we have to change the structure of the "history" of the process. Actually, we construct a process Z_t taking values in $E \times [0, 1] \times E$ together with a sequence of jump times T_n for this process such that at any time T_n, we know the present state of the process X_{T_n}, but also the future state $X_{T_{n+1}}$. Moreover, the following properties are fulfilled:

1. The first coordinate Z_t^1 of Z_t has the same dynamics as the original process X_t starting from $X_0 = x$ if we fix the initial condition $Z_0^1 = x$, but not Z_0^2 and Z_0^3.
2. On each interval $[T_n, T_{n+1}[$, Z_t^2 and Z_t^3 are constant, and the third coordinate $Z_{T_n}^3$ represents a choice of $X_{T_{n+1}}$ according to the splitting technique which has to be attained by the bridge (the process) between T_n and T_{n+1}.
3. The second coordinate is used only in order to model the splitting. It is not of further importance.

Then it is possible to define a sequence of stopping times (S_n, R_n) for this process which is a generalised life-cycle decomposition in the following sense:

1 For all n, $S_n < R_n < \infty$, $S_{n+1} = S_n + S_1 \circ \theta_{S_n}$, $R_n = \inf\{T_m : T_m > S_n\}$.
2 For every n, X_{R_n} is independent of $\sigma\{X_s : s \leq S_n\}$ and $\mathcal{L}(X_{R_n}) = \nu$ for some fixed probability measure ν.
3 For all μ-integrable functions f and for any initial measure π, $E_\pi(\int_{R_n}^{R_{n+1}} f(X_s)\mathrm{d}s) = \mu(f)$ (up to a multiplicative constant).

Note that it is not possible to divide the path of the process into real i.i.d. excursions. We obtain the independence only after the waiting time R_n after S_n. Some special attention has to be paid to the initial segment $\int_0^{R_1} f(X_s)\mathrm{d}s$ — and it is for that reason that we have to introduce and to investigate functions that are called *special functions*.

As a consequence, we establish a generalisation of Chen's inequality (compare also to [5], Lemma 1, (2.4)) for additive functionals of the Markov process (see Theorem 2.18) and get in