Electron. Commun. Probab. **0** (2016), no. 0, 1–10.
DOI: 10.1214/ECP.vVOL-PID
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

On the f-Norm Ergodicity of Markov Processes in Continuous Time

Ioannis Kontoyiannis[*] Sean P. Meyn[†]

Abstract

Consider a Markov process $\mathbf{\Phi} = \{\Phi(t) : t \geq 0\}$ evolving on a Polish space X. A version of the f-Norm Ergodic Theorem is obtained: Suppose that the process is ψ-irreducible and aperiodic. For a given function $f \colon \mathsf{X} \to [1, \infty)$, under suitable conditions on the process the following are equivalent:

(i) There is a unique invariant probability measure π satisfying $\int f \, d\pi < \infty$.

(ii) There is a closed set C satisfying $\psi(C) > 0$ that is "self f-regular."

(iii) There is a function $V \colon \mathsf{X} \to (0, \infty]$ that is finite on at least one point in X, for which the following Lyapunov drift condition is satisfied,

$$\mathcal{D}V \leq -f + b\mathbb{I}_C\,, \tag{V3}$$

where C is a closed small set and $\mathcal{D}$ is the extended generator of the process.

For discrete-time chains the result is well-known. Moreover, in that case, the ergodicity of $\mathbf{\Phi}$ under a suitable norm is also obtained: For each initial condition $x \in \mathsf{X}$ satisfying $V(x) < \infty$, and any function $g \colon \mathsf{X} \to \mathbb{R}$ for which $|g|$ is bounded by f,

$$\lim_{t\to\infty} \mathsf{E}_x[g(\Phi(t))] = \int g \, d\pi.$$

Possible approaches are explored for establishing appropriate versions of corresponding results in continuous time, under appropriate assumptions on the process $\mathbf{\Phi}$ or on the function g.

Keywords: Markov process ; continuous time ; generator ; stochastic Lyapunov function ; ergodicity.
AMS MSC 2010: 60J25 ; 37A30 : 47H99.
Submitted to ECP on December 2, 2015, final version accepted on November 14, 2016.
Supersedes arXiv:1512.00523.

[*]Department of Informatics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece. E-mail: yiannis@aueb.gr, http://pages.cs.aueb.gr/users/yiannisk/. I.K. was supported by the European Union and Greek National Funds through the Operational Program Education and Lifelong Learning of the National Strategic Reference Framework through the Research Funding Program Thales-Investing in Knowledge Society through the European Social Fund.

[†]Department of Electrical and Computer Engineering, University of Florida, Gainesville, USA. E-mail: meyn@ece.ufl.edu, http://www.meyn.ece.ufl.edu/. S.P.M. was supported in part by the National Science Foundation ECS-0523620, and AFOSR grant FA9550-09-1-0190. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or AFOSR.

1 Introduction

Consider a Markov process $\mathbf{\Phi} = \{\Phi(t) : t \geq 0\}$ in continuous time, evolving on a Polish space X, equipped with its Borel σ-field $\mathcal{B}$. Assume it is a nonexplosive Borel right process: It satisfies the strong Markov property and has right-continuous sample paths [1, 8].

The distribution of the process $\mathbf{\Phi}$ is described by the initial condition $\Phi(0) = x \in \mathsf{X}$ and the transition semigroup: For any $t \geq 0$, $x \in \mathsf{X}$, $A \in \mathcal{B}$,

$$P^t(x, A) := \mathsf{P}_x\{\Phi(t) \in A\} := \Pr\{\Phi(t) \in A \mid \Phi(0) = x\}.$$

A set C is called *small* if there is probability measure ν on $(\mathsf{X}, \mathcal{B})$, a time $T > 0$, and a constant $\varepsilon > 0$ such that,

$$P^T(x, A) \geq \varepsilon\nu(A), \qquad \textit{for every } A \in \mathcal{B}.$$

It is assumed that the process is ψ-irreducible and aperiodic, where ψ is a probability measure on $(\mathsf{X}, \mathcal{B})$. This means that for each set $A \in \mathcal{B}$ satisfying $\psi(A) > 0$, and each $x \in \mathsf{X}$,

$$P^t(x, A) > 0, \qquad \textit{for all } t \textit{ sufficiently large.}$$

It follows that there is a countable covering of the state space by small sets [7, Prop. 3.4].

The Lyapunov theory considered in this paper and in our previous work [4, 8] is based on the *extended generator* of $\mathbf{\Phi}$, denoted $\mathcal{D}$. A function $h \colon \mathsf{X} \to \mathbb{R}$ is in the domain of $\mathcal{D}$ if there exists a function $g \colon \mathsf{X} \to \mathbb{R}$ such that the stochastic process defined by,

$$M(t) = h(\Phi(t)) - \int_0^t g(\Phi(s))\, ds, \qquad t \geq 0, \tag{1.1}$$

is a *local martingale*, for each initial condition $\Phi(0)$ [1, 12]. We then write $g = \mathcal{D}h$.

For example, consider a diffusion on $\mathsf{X} = \mathbb{R}^d$, namely, the solution of the stochastic differential equation,

$$d\Phi(t) = u(\Phi(t))dt + M(\Phi(t))dB(t), \quad t \geq 0,\ \Phi(0) = x, \tag{1.2}$$

where $u = (u_1, u_2, \dots, u_d)^{\intercal} : \mathsf{X} \to \mathbb{R}^d$ and $M : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^k$ are Lipschitz, and $\boldsymbol{B} = \{B(t) : t \geq 0\}$ is k-dimensional standard Brownian motion. If the function $h \colon \mathsf{X} \to \mathbb{R}$ is C^2 then we can write [12],

$$\mathcal{D}h\,(x) = \sum_i u_i(x) \frac{d}{dx_i} h\,(x) + \frac{1}{2} \sum_{ij} \Sigma_{ij}(x) \frac{d^2}{dx_i\, dx_j} h\,(x), \qquad x \in \mathsf{X}.$$

The Lyapunov condition considered in this paper is Condition (V3) of [8]: For a function $V \colon \mathsf{X} \to (0, \infty]$ which is finite for at least one $x \in \mathsf{X}$, a function $f \colon \mathsf{X} \to [1, \infty)$, a constant $b < \infty$, and a closed, small set $C \in \mathcal{B}$,

$$\mathcal{D}V \leq -f + b\mathbb{I}_C\,. \tag{V3}$$

It is entirely analogous to its discrete-time counterpart [10], in which the extended generator is replaced by a difference operator $\mathcal{D} = P - I$, where P is the transition kernel of the discrete-time chain and I is the identity operator.

The lower bound $f \geq 1$ is imposed in (V3) because this function is used to define two norms: One on measurable functions $g \colon \mathsf{X} \to \mathbb{R}$ via,

$$\|g\|_f := \sup_{x \in \mathsf{X}} \frac{|g(x)|}{f(x)},$$

and a second norm on signed measures μ on $(\mathsf{X}, \mathcal{B})$:

$$\|\mu\|_f = \sup_{g:|g|\le f} |\mu(g)|.$$

Our main goal is to establish the erodicity of $\mathbf{\Phi}$ in terms of this norm: There is an invariant measure π for the semi-group $\{P^t\}$ satisfying,

$$\lim_{t\to\infty} \|P^t(x,\,\cdot\,) - \pi(\,\cdot\,)\|_f = 0\,. \tag{1.3}$$

The following result is a partial extension of the f-Norm Ergodic Theorem of [10] to the continuous time setting.

Theorem 1.1. *Suppose that the Markov process $\mathbf{\Phi}$ is ψ-irreducible and aperiodic, and let $f \ge 1$ be a function on X. Then the following conditions are equivalent:*

(i) *The semi-group admits an invariant probability measure π satisfying:*

$$\pi(f) := \int \pi(dx) f(x) < \infty.$$

(ii) *There exists a closed, small set $C \in \mathcal{B}$ such that,*

$$\sup_{x\in C} \mathsf{E}_x\Big[\int_0^{\tau_C(1)} f(\Phi(t))\,dt\Big] < \infty, \tag{1.4}$$

where $\tau_C(1) := \inf\{t \ge 1 : \Phi(t) \in C\}$ and E_x denotes the expectation operator under $X_0 = x$.

(iii) *There exists a closed, small set C and an extended-valued non-negative function V satisfying $V(x_0) < \infty$ for some $x_0 \in \mathsf{X}$, such that Condition (V3) holds.*

Moreover, if (iii) holds then there exists a constant b_f such that,

$$\mathsf{E}_x\Big[\int_0^{\tau_C(1)} f(\Phi(t))\,dt\Big] \le b_f(V(x)+1), \quad x \in \mathsf{X} \tag{1.5}$$

where V and C satisfy the conditions of (iii). The set $S_V = \{x : V(x) < \infty\}$ is absorbing ($P^t(x, S_V) = 1$ for each $x \in S_V$ and all $t \ge 0$), and also full ($\pi(S_V) = 1$).

Proof. Theorem 1.2 (b) of [9] gives the equivalence of (i) and (ii). Theorem 4.3 of [9] gives the implication (iii) $\Rightarrow$ (ii), along with the bound (1.5).

Conversely, if (ii) holds then we can define,

$$V(x) = \int_0^\infty \mathsf{E}_x\Big[f(\Phi(t))\exp\Big(-\int_0^t \mathbb{I}\{\Phi(s)\in C\}\,ds\Big)\Big]\,dt. \tag{1.6}$$

We show in Proposition 2.2 that this is a solution to (V3) and that it is uniformly bounded on C. ☐

The function V in (1.6) has the following interpretation. Let $\widetilde{T}$ denote an exponential random variable that is independent of $\mathbf{\Phi}$, and denote,

$$\tilde{\tau}_C = \min\Big\{t : \int_0^t \mathbb{I}\{\Phi(s)\in C\}\,ds = \widetilde{T}\Big\}.$$

We then have,

$$V(x) = \mathsf{E}_x\Big[\int_0^{\tilde{\tau}_C} f(\Phi(t))\,dt\Big], \tag{1.7}$$

where now the expectation is over both $\mathbf{\Phi}$ and $\widetilde{T}$. Consequently, this construction is similar to the converse theorems found in [10] for discrete-time models.

Theorem 1.1 is almost identical to the f-Norm Ergodic Theorem of [10], except that it leaves out the implications to ergodicity of the process. This brings us to two open problems: Under the conditions of Theorem 1.1:

Q1 *Can we conclude that* (1.3) *holds for any initial condition* $x \in S_V$*?*

Q2 Assume in addition that $\pi(V) < \infty$. *Can we conclude that there exists a finite constant* B_f *such that, for all* $x \in S_V$*,*

$$\int_0^\infty \|P^t(x,\,\cdot\,) - \pi\|_f \, dt \le B_f(V(x) + 1). \tag{1.8}$$

In discrete time, questions Q1 and Q2 are answered in the affirmative by the f-Norm Ergodic Theorem of [10], with the integral replaced by a sum in (1.8).

Q2 is resolved in the affirmative in this paper by an application of the discrete-time counterpart:

Theorem 1.2. *Suppose that the Markov process* $\mathbf{\Phi}$ *is* ψ*-irreducible and aperiodic, and that there is a solution to (V3) with* V *everywhere finite. Then there is a constant* B_f^0 *such that for each* $x, y \in \mathsf{X}$*,*

$$\int_0^\infty \|P^t(x,\,\cdot\,) - P^t(y,\,\cdot\,)\|_f \, dt \le B_f^0(V(x) + V(y) + 1) \tag{1.9}$$

If in addition $\pi(V) < \infty$*, then* (1.8) *also holds for some constant* B_f *and all* x*.*

Although the full resolution of Q1 remains open, in Section 3 we discuss how (1.3) can be established under additional conditions on the process $\mathbf{\Phi}$.

We begin, in the following section, with the proof of the implication (ii) $\Rightarrow$ (iii), which is based on theory of generalized resolvents and f-regularity [7]. Following this result, it is shown in Proposition 2.3 that f-regularity of the process is equivalent to f_Δ-regularity for the sampled process, where Δ is the sampling interval, and,

$$f_\Delta(x) = \int_0^\Delta \mathsf{E}_x[f(\Phi(t))]\, dt, \quad x \in \mathsf{X}. \tag{1.10}$$

This is the basis of the proof of Theorem 1.2 that is contained in Section s:ergodic.

Acknowledgment. The work reported in this note was prompted by a question of Yuanyuan Liu who, in a private communication, pointed out to us that some results in our earlier work [3] were stated inaccurately. Specifically: (1.) The implication (ii) $\Rightarrow$ (iii) in Theorem 2.2 of [3], which is the same as the corresponding result in our present Theorem 1.1, was stated there without proof; and (2.) The convergence in (1.3) was stated as a consequence of any of the three equivalent conditions (i)—(iii), again without proof. This note attempts to address and correct these omissions, although the relevant statements in [3] were only discussed as background material and do not affect any of the subsequent results in that paper.

2 f-Regularity

Following [7], we denote for each $r \ge 0$ and $B \in \mathcal{B}$,

$$G_B(x, f; r) := \mathsf{E}_x\Big[\int_0^{\tau_B(r)} f(\Phi(t))\, dt\Big], \tag{2.1}$$

where $\tau_B(r) = \inf\{t \ge r : \Phi(t) \in B\}$, and we write $G_B(x, f) = G_B(x, f; 0)$. The Markov process is called *f-regular* if there exists $r_0 > 0$ such that $G_B(x, f; r_0) < \infty$ for every x and every $B \in \mathcal{B}$ satisfying $\psi(B) > 0$.

The following result, given here without proof, is a simple consequence of Lemma 4.1 and Prop. 4.3 of [7]:

Proposition 2.1. *Suppose that the set C is closed and small, and that the following self-regularity property holds: There exists $r_0 > 0$ such that $\sup_{x\in C} G_C(x, f; r_0) < \infty$. Then:*

(i) *There is $b_C < \infty$ such that $G_C(x, f; r) < G_C(x, f; r_0) + b_C r$ for each x and r.*

(ii) *For each $B \in \mathcal{B}$ satisfying $\psi(B) > 0$, for each $r \ge 0$, and for each $x \in \mathsf{X}$,*

$$G_C(x, f; r) < \infty \Rightarrow G_B(x, f; r) < \infty.$$

Consequently, the process is f-regular if $G_C(x, f; r_0) < \infty$ for each x.

We next show that the function V in (1.6) is finite-valued on $\{x \in \mathsf{X} : G_C(x, f; r_0) < \infty\}$. We show that V is in the domain of the extended generator, and obtain an expression for $\mathcal{D}V$.

Consider the generalized resolvent developed in [7, 11]: For a function $h\colon \mathsf{X} \to \mathbb{R}_+$, $A \in \mathcal{B}$, and $x \in \mathsf{X}$, denote,

$$R_h(x, A) = \int_0^\infty \mathsf{E}_x\Big[\mathbb{I}_A(\Phi(t)) \exp\Big(-\int_0^t h(\Phi(s))\, ds\Big)\Big]\, dt.$$

With the usual interpretation of P^t, or any kernel $Q(x, dy)$, as a lineal operator, $g \mapsto Qg = \int g(y)Q(\cdot, dy)$, it is shown in [11] that the following resolvent equation holds: For any functions $g \ge h \ge 0$,

$$R_h = R_g + R_g I_{g-h} R_h, \tag{2.2}$$

where, for any function g, I_g denotes the (operator induced by the) kernel $I_g(x, dy) = g(x)\delta_x(dy)$.

When $h \equiv \alpha$ is constant, we obtain the usual resolvent,

$$R_\alpha := \int_0^\infty e^{-\alpha t} P^t\, dt, \qquad \alpha > 0, \tag{2.3}$$

In the case $\alpha = 1$ we write $R := R_1 = \int_0^\infty e^{-t} P^t\, dt$, and call R “the” resolvent kernel. For any non-negative function $g\colon \mathsf{X} \to \mathbb{R}_+$ for which Rg is finite valued, the function $\gamma = Rg$ is in the domain of the extended generator, with,

$$\mathcal{D}\gamma = Rg - g. \tag{2.4}$$

Proposition 2.2. *Suppose that the assumptions of Theorem 1.1 (ii) hold: There is a closed, small set $C \in \mathcal{B}$ such that, $\sup_{x\in C} G_C(x, f; r_0) < \infty$ with $r_0 = 1$. Then the function V defined in* (1.7) *is finite on the full set $S_V \subset \mathsf{X}$ and (V3) holds with this function V and this closed set C.*

Proof. Proposition 4.3 (ii) of [7] implies that the set of x for which $G_C(x, f; 1) < \infty$ is a full set. This result combined with Proposition 4.4 (ii) of [7] implies that V is bounded on C.

For arbitrary x we have $\tilde\tau_C > \tau_C = \min\{t \ge 0 : \Phi(t) \in C\}$. Consequently, by the strong Markov property and the representation (1.7),

$$\begin{aligned} V(x) &= \mathsf{E}_x\Big[\int_0^{\tau_C} f(\Phi(t))\, dt\Big] + \mathsf{E}_x\Big[\mathsf{E}_{\Phi(\tau_C)}\Big[\int_0^{\tilde\tau_C} f(\Phi(t))\, dt\Big]\Big] \\ &\le G_C(x, f; 1) + \sup_{x'\in C} V(x'). \end{aligned}$$

Hence $V(x)$ is finite whenever $G_C(x, f; 1)$ is finite.

To establish (V3), first observe that the function V in (1.7) can be expressed,

$$V = R_h f, \qquad \text{with } h = \mathbb{I}_C.$$

Taking $g \equiv 1$, the resolvent equation gives,

$$R_h = R + R I_{1-h} R_h = R[I + I_{C^c} R_h],$$

where, for any set B and kernel Q, $I_B Q$ denotes the kernel $\mathbb{I}_B(x) Q(x, dy)$. Combining the representation of V above with (2.4) we obtain,

$$\begin{aligned} V &= R[I + I_{C^c} R_h] f \\ \text{and} \qquad \mathcal{D}V &= (R - I)[I + I_{C^c} R_h] f. \end{aligned}$$

The second equation can be decomposed as follows,

$$\mathcal{D}V = D_1 - D_2 - f,$$

with $D_1 = R[I + I_{C^c} R_h] f = V$ and $D_2 = I_{C^c} R_h f = I_{C^c} V$. Substitution then gives,

$$\mathcal{D}V = -f + \mathbb{I}_C V.$$

This establishes (V3) with $b = \sup_{x \in C} V(x)$. □

The final results in this section concern the Δ-*skeleton chain*. This is the discrete-time Markov chain with transition kernel P^Δ, where $\Delta \geq 1$ is given. It can be realized by sampling the Markov process with sampling interval Δ. The sampled process is denoted,

$$X(i) = \Phi(i\Delta), \qquad i \geq 0. \tag{2.5}$$

In prior work, the skeleton chain is used to translate ergodicity results for discrete-time Markov chains to the continuous time setting. For example, Theorem 6.1 of [8] implies that a weak version of the ergodic convergence (1.3) holds for an f-regular Markov process:

$$\lim_{t \to \infty} \|P^t(x, \cdot\,) - \pi(\,\cdot\,)\|_1 = 0\,. \tag{2.6}$$

The proof consists of two ingredients: (i) The corresponding ergodicity result holds for the Δ-skeleton chain, and (ii) the error $\|P^t(x, \cdot\,) - \pi(\,\cdot\,)\|_1$ is non-increasing in t.

In the next section we use a similar approach to address question Q2. The f_Δ norm is considered, where the function f_Δ is defined in (1.10). Denote,

$$\sigma_C^\Delta = \min\{i \geq 0 : X(i) \in C\}, \qquad \tau_C^\Delta = \min\{i \geq 1 : X(i) \in C\}\,.$$

The Δ-skeleton is called f_Δ-*regular* if,

$$G_B^\Delta(x, f_\Delta) := \mathsf{E}_x\Big[\sum_{i=0}^{\tau_B^\Delta} f_\Delta(X(i))\Big] < \infty,$$

for every $x \in \mathsf{X}$ and every $B \in \mathcal{B}$ satisfying $\psi(B) > 0$.

Proposition 2.3. *If the process* $\boldsymbol{\Phi}$ *is* f*-regular, then each* Δ*-skeleton is* f_Δ*-regular. Moreover, there is a closed* f*-regular set* C *such that:*

(i) *For a finite-valued function* $V_\Delta \colon \mathsf{X} \to (0, \infty]$ *and a finite constant* b*,*

$$P^\Delta V_\Delta \leq V_\Delta - f_\Delta + b\mathbb{I}_C\,, \tag{2.7}$$

and $\sup_x |V_\Delta(x) - G_C(x, f)| < \infty$.

(ii) *For every* $x \in \mathsf{X}$ *and every* $B \in \mathcal{B}$ *satisfying* $\psi(B) > 0$*, there is a constant* $c_B < \infty$ *such that,*

$$G_B^{\Delta}(x, f_{\Delta}) \leq G_C(x, f) + c_B. \tag{2.8}$$

Proof. It is enough to establish (i). Theorem 14.2.3 of [10] then implies that for every $B \in \mathcal{B}$ satisfying $\psi(B) > 0$, there is a constant $c_B^{\Delta} < \infty$ satisfying $G_B^{\Delta}(x, f_{\Delta}) \leq V_{\Delta}(x) + c_B^{\Delta}$.

Let C denote any closed f-regular set for the process, satisfying $\psi(C) > 0$. For $V_0(x) = G_C(x, f)$ we obtain a bound similar to (2.7) through the following steps. First write,

$$P^{\Delta} V_0\,(x) = \mathsf{E}_x\Big[\int_{\Delta}^{\tau_C(\Delta)} f(\Phi(t))\,dt\Big]\,.$$

The integral can be expressed as a sum,

$$\begin{aligned}\int_{\Delta}^{\tau_C(\Delta)} f(\Phi(t))\,dt = &\int_{\Delta}^{\tau_C(\Delta)} f(\Phi(t))\,dt \mathbb{I}\{\tau_C \leq \Delta\} \\ &+ \int_{\Delta}^{\tau_C} f(\Phi(t))\,dt \mathbb{I}\{\tau_C > \Delta\}.\end{aligned}$$

By the strong Markov property,

$$\begin{aligned}\mathsf{E}_x\Big[\mathbb{I}\{\tau_C \leq \Delta\} \int_{\Delta}^{\tau_C(\Delta)} f(\Phi(t))\,dt\Big] &\leq \mathsf{E}_x\Big[\mathbb{I}\{\tau_C \leq \Delta\} \int_{\tau_C}^{\tau_C(\Delta)} f(\Phi(t))\,dt\Big] \\ &\leq \mathsf{P}_x\{\tau_C \leq \Delta\} \sup_y G_C(y, f; \Delta).\end{aligned}$$

Consequently,

$$P^{\Delta} V_0\,(x) \leq \mathsf{E}_x\Big[\int_{\Delta}^{\tau_C} f(\Phi(t))\,dt\Big] + b_0 s(x) = V_0(x) - f_{\Delta}(x) + b_0 s(x), \tag{2.9}$$

where $b_0 = \sup_y G_C(y, f; \Delta) < \infty$, and $s(x) = \mathsf{P}_x\{\tau_C \leq \Delta\}$.

To eliminate the function s in (2.9) we establish the following bound: For some $\varepsilon_0 > 0$ and $k_0 \geq 1$,

$$P^{k_0\Delta}(x, C) \geq \varepsilon_0 s(x), \qquad x \in \mathsf{X}. \tag{2.10}$$

The proof is again by the strong Markov property:

$$\begin{aligned}P^{k_0\Delta}(x, C) &\geq \mathsf{E}_x[\mathbb{I}\{\tau_C \leq \Delta\}\mathbb{I}\{\Phi(k_0\Delta) \in C\}] \\ &= \int_{r=0}^{\Delta} \int_y \mathsf{P}_x\{\tau_C \in dr,\, \Phi(r) \in dy\} P^{k_0\Delta - r}(y, C) \\ &\geq \varepsilon(k) s(x),\end{aligned}$$

where $\varepsilon(k) = \inf\{P^{k_0\Delta - r}(y, C) : y \in C,\ 0 \leq r \leq \Delta\}$. This is strictly positive for sufficiently large k because (2.6) holds. This establishes (2.10).

The Lyapunov function can now be specified as,

$$V_{\Delta}(x) = V_0(x) + b_0 G_C^{\Delta}(x, s),$$

where b_0 is defined in (2.9). The required bound $\sup_x |V_{\Delta}(x) - G_C(x, f)| < \infty$ holds because $V_0(x) = G_C(x, f)$, and the second term is uniformly bounded:

$$\begin{aligned} G_C^\Delta(x,s) &= \mathsf{E}_x\Big[\sum_{i=0}^{\tau_B^\Delta} s(X(i))\Big] \\ &\le \varepsilon_0^{-1}\mathsf{E}_x\Big[\sum_{i=0}^{\tau_B^\Delta} P^{k_0\Delta}(\Phi(i\Delta),C)\Big] \\ &= \varepsilon_0^{-1}\mathsf{E}_x\Big[\sum_{i=0}^{\tau_B^\Delta} \mathbb{I}\{X(i+k_0)\in C\}\Big] \le \varepsilon_0^{-1}(k_0+1). \end{aligned}$$

Consequently, from familiar arguments,

$$\begin{aligned} PV_\Delta(x) - V_\Delta(x) &\le -f_\Delta(x) + b_0 s(x) \\ &\qquad + b_0\Big\{G_C^\Delta(x,s) - s(x) + \mathbb{I}_C(x)\varepsilon_0^{-1}(k_0+1)\Big\}. \end{aligned}$$

This establishes (2.7) with $b = b_0\varepsilon_0^{-1}(k_0+1)$. □

3 f-Norm Ergodicity

In this section we consider the implications to the ergodicity of the process. We assume that (V3) holds for a finite-valued function $V\colon \mathsf{X}\to(0,\infty)$, so that the process is f-regular.

Q1. f-norm ergodicity. The ergodicity of $\mathbf{\Phi}$ in terms of the f-norm as in (1.3) has only been established under special conditions. Theorem 5.3 of [9] implies that (1.3) will hold if f is subject to this additional bound: For some $\beta \ge 0$,

$$P^t f \le \beta e^{\beta t} f, \qquad t \ge 0.$$

This holds for example if $f \equiv 1$ and $\beta = 1$.

It is likely that the application of coupling bounds will lead to a more general theory. Under stronger conditions on the process, such a coupling time was obtained in [5], and it was used in [6] to obtain rates of convergence in the law of large numbers. However, to construct the coupling time, it is assumed in this prior work that the semi-group $\{P^t\}$ admits a density for each t. No such assumptions are required in the discrete-time setting, so the full answer to Q1 remains open.

Q2. Proof of Theorem 1.2. The complete resolution of Q2 is possible by applying Proposition 2.3, which implies that the skeleton chain $\{X(i) = \Phi(i\Delta) : i \ge 0\}$ is f_Δ-regular. The bound (2.8) is the main ingredient in the proof of Theorem 1.2, but we also require the following relationship between a norm for the process and a norm for the sampled chain.

Lemma 3.1. *For any signed measure μ,*

$$\|\mu\|_{f_\Delta} \ge \int_0^\Delta \|\mu P^t\|_f \, dt,$$

where, for any measure ν and kerner Q, νQ denotes the measure $\nu Q(\cdot) = \int \nu(dx)Q(x,\cdot)$.

Proof. We first consider the right-hand side. Consider the signed measure Γ on $[0,\Delta]\times\mathsf{X}$ defined by:

$$\Gamma(dt,dy) = \mu P^t(dy)dt.$$

Define $f^{\Delta}\colon [0,\Delta]\times\mathsf{X}\to[1,\infty)$ via $f(t,y)=f(y)$ for each pair t,y, and the associated norm,

$$\|\Gamma\|_{f^{\Delta}}=\sup\iint g(t,y)\Gamma(dt,dy),$$

where the supremum is over all g satisfying $|g(t,y)|\le f^{\Delta}(t,y)$ for all t,y. It is shown next that the norm can be expressed,

$$\|\Gamma\|_{f^{\Delta}}=\int_0^{\Delta}\|\mu P^t\|_f\,dt. \tag{3.1}$$

The Jordan decomposition theorem [2] implies that there is a minimal decomposition, $\Gamma=\Gamma_+-\Gamma_-$, in which the two measures on the right-hand side are non-negative, with disjoint supports denoted S_+,S_-, resoectively. Hence $|\Gamma|:=\Gamma_++\Gamma_-$ is a non-negative measure. In this notation the norm is expressed,

$$\begin{aligned}\|\Gamma\|_{f^{\Delta}}&=\iint f^{\Delta}(t,y)|\Gamma|(dt,dy)\\&=\iint f(y)\big(\mathbb{I}_{S_+}(t,y)-\mathbb{I}_{S_-}(t,y)\big)\Gamma(dt,dy)\\&=\int_0^{\Delta}\Big[\int_{y\in\mathsf{X}}f(y)\big(\mathbb{I}_{S_+}(t,y)-\mathbb{I}_{S_-}(t,y)\big)\mu P^t(dy)\Big]\,dt.\end{aligned}$$

For each t, the measure on $(\mathsf{X},\mathcal{B})$ defined by $\big(\mathbb{I}_{S_+}(t,y)-\mathbb{I}_{S_-}(t,y)\big)\mu P^t(dy)$ is the marginal of $|\Gamma|$, and is hence a non-negative measure for a.e. t. It follows that for such t,

$$\int_{y\in\mathsf{X}}f(y)\big(\mathbb{I}_{S_+}(t,y)-\mathbb{I}_{S_-}(t,y)\big)\mu P^t(dy)=\|\mu P^t\|_f,$$

which gives (3.1).

Consider next the left-hand side of the inequality in the lemma. Letting $\mu=\mu_+-\mu_-$ denote the Jordan decomposition for the signed measure μ, and $|\mu|=\mu_++\mu_-$, we have,

$$\|\mu\|_{f_\Delta}=\int f^{\Delta}(x)|\mu|(dx)=\int_{t=0}^{\Delta}\int_{x\in\mathsf{X}}|\mu|(dx)P^t(x,dy)f(y).$$

The right-hand side can be expressed as,

$$\int_0^{\Delta}\int|\mu|(dx)P^t(x,dy)f(y)=\iint f^{\Delta}(t,y)\Lambda_+(dt,dy)+\iint f^{\Delta}(t,y)\Lambda_-(dt,dy),$$

where $\Lambda_\pm(dt,dy)=\mu_\pm P^t(dy)dt$ defines a decomposition:

$$\Gamma=\Lambda_+-\Lambda_-\,.$$

It follows that $\|\mu\|_{f_\Delta}\ge\|\Gamma\|_{f^{\Delta}}$, by the minimality of the Jordan decomposition. This bound combined with (3.1) completes the proof. $\square$

Proof of Theorem 1.2. Theorem 1.1 combined with the result of Proposition 2.3 establishes f_Δ-regularity of the skeleton chain under (V3): The skeleton chain satisfies (V3) with Lyapunov function V_Δ that satisfies $\sup_x|V_\Delta(x)-G_C(x,f)|<\infty$. The bound (1.5) in Theorem 1.1 implies that $V_\Delta(x)\le b_f^{\Delta}(V(x)+1)$ for some constant b_f^{Δ} and all x.

Theorem 14.3.4 of [10] then gives the bound, for some finite constant $M_f^0<\infty$,

$$\sum_{k=0}^{\infty}\|P^{\Delta k}(x,\,\cdot\,)-P^{\Delta k}(y,\,\cdot\,)\|_{f_\Delta}\le M_f^0(V(x)+V(y)+1). \tag{3.2}$$

Next apply Lemma 3.1 with $\mu(\,\cdot\,) = P^{\Delta k}(x,\,\cdot\,) - P^{\Delta k}(y,\,\cdot\,)$ to obtain,

$$\|P^{\Delta k}(x,\,\cdot\,) - P^{\Delta k}(y,\,\cdot\,)\|_{f_\Delta} \geq \int_0^\Delta \|\mu P^t\|_f \, dt, \tag{3.3}$$

and recognize that $\mu P^t(\,\cdot\,) = P^{\Delta k+t}(x,\,\cdot\,) - P^{\Delta k+t}(y,\,\cdot\,)$. Substituting the resulting bound into (3.2) establishes (1.9).

The proof of (1.8) is similar: If in addition $\pi(V) < \infty$, then Theorem 14.3.5 of [10] gives, for some constant $M_f < \infty$,

$$\sum_{k=0}^{\infty} \|P^{\Delta k}(x,\,\cdot\,) - \pi(\,\cdot\,)\|_{f_\Delta} \leq M_f(V(x)+1). \tag{3.4}$$

This combined with (3.3) completes the proof. □

References

[1] Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York. Characterization and convergence. MR-0838085

[2] Halmos, P. R. (1950). *Measure Theory*. D. Van Nostrand Company, Inc., New York, N. Y. MR-0033869

[3] Kontoyiannis, I. and Meyn, S. P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.* **13**, 1, 304–362. MR-1952001

[4] Kontoyiannis, I. and Meyn, S. P. (2005). Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab. 10*, no. 3, 61–123 (electronic). MR-2120240

[5] Löcherbach, E. and Loukianova, D. (2008). On Nummelin splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multi-dimensional diffusions. *Stochastic Process. Appl.* **118**, 8, 1301–1321. MR-2427041

[6] Löcherbach, E., Loukianova, D., and Loukianov, O. (2011). Polynomial bounds in the ergodic theorem for one-dimensional diffusions and integrability of hitting times. *Ann. Inst. Henri Poincaré Probab. Stat.* **47**, 2, 425–449. MR-2814417

[7] Meyn, S. P. and Tweedie, R. L. (1993). Generalized resolvents and Harris recurrence of Markov processes. In *Doeblin and modern probability (Blaubeuren, 1991)*. Contemp. Math., Vol. **149**. Amer. Math. Soc., Providence, RI, 227–250. MR-1229967

[8] Meyn, S. P. and Tweedie, R. L. (1993). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.* **25**, 3, 487–517. MR-1234294

[9] Meyn, S. P. and Tweedie, R. L. (1993). Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25**, 3, 518–548. MR-1234295

[10] Meyn, S. P. and Tweedie, R. L. (1993). *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London, Ltd., London. MR-1287609

[11] Neveu, J. (1972). Potentiel Markovien récurrent des chaînes de Harris. *Ann. Inst. Fourier (Grenoble)* **22**, 2, 85–130. MR-0380992

[12] Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge. Itô calculus, Reprint of the second (1994) edition. MR-1780932