HAL
open science

QUASI-STATIONARY DISTRIBUTIONS FOR STOCHASTIC APPROXIMATION ALGORITHMS WITH CONSTANT STEP SIZE

Bastien Marmet

To cite this version:

Bastien Marmet. QUASI-STATIONARY DISTRIBUTIONS FOR STOCHASTIC APPROXIMATION ALGORITHMS WITH CONSTANT STEP SIZE. 2012. hal-00805201

HAL Id: hal-00805201

https://hal.science/hal-00805201v1

Preprint submitted on 28 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Submitted to the Annals of Applied Probability

QUASI-STATIONARY DISTRIBUTIONS FOR STOCHASTIC APPROXIMATION ALGORITHMS WITH CONSTANT STEP SIZE

By Bastien Marmet*

University of Neuchatel

Abstract: In this paper we investigate quasi-stationary distributions μ^N of stochastic approximation algorithms with constant step size which can be viewed as random perturbations of a time-continuous dynamical system. Inspired by ecological models these processes have a closed absorbing set corresponding to extinction. Under some large deviation assumptions and the existence of an interior attractor for the ODE, we show that the weak* limit points of the QSD μ^N are invariant measures for the ODE with support in the interior attractors.

CONTENTS

1 Introduction 1
2 Model, notations and hypotheses 3
3 Convergence of QSD and absorption time 7
3.1 Absorption time 9
3.2 Convergence of the QSD to an invariant measure 10
4 Support of the limiting measure 12
4.1 Absorption-preserving pseudo-orbit 18
4.2 Going back the hypotheses 4.1, 4.2 and 4.7 26
A Reminders 36
A.1 Quasi-stationary distributions 36
A.2 Stochastic approximation algorithms 38

1. Introduction.

One of the most considered issue in theoretical ecology is to find out under what kind of conditions one can expect a population of interacting species (animals, plants, microorganisms, ...) to survive on the long term with no extinctions. When these conditions are met the interacting populations are

*The author acknowledge financial support from the Swiss National Foundation Grant FN 200021-138242/1

AMS 2000 subject classifications: Primary 60J10, 34F05; Secondary 60F10, 92D25

Keywords and phrases: Stochastic Approximation Algorithm, Quasi-Stationary Distributions

said to persist or coexist. In the past, differential equations and nonlinear difference equations have been used to model these phenomena. Famous examples are Lotka [14] and Volterra [22] work on competitive and predator-prey interactions, Thompson [21], Nicholson and Bailey [16] on host-parasite interactions, and Kermack and McKendrick [12] on disease outbreaks. For these deterministic models, persistence definitions sometimes vary but most authors link persistence with the existence of an attractor bounded away from the extinction states, in which case persistence holds over an infinite time horizon, see e.g. [20]. In order to refine these models and allow for some "roughness" and/or influence of unpredictable outer events, randomness has been added to these models, leading to Markov processes models. However, extinctions being absorbent states and species dying out with positive probabilities, the underlying theory of Markov processes shows that, in finite time, extinction is inevitable. Yet, in the real world, with large sized pools of population, we don't observe that inevitable extinction. This finite extinction time may then be very large and the system may remain in some sort of "metastable state" bounded away from extinction for a long time. In [10], Faure and Schreiber studied this problem for randomly perturbed discrete time dynamical systems, showing that, under the appropriate assumptions about the random perturbations, there exists a positive attractor (i.e. an attractor which is bounded away from extinction states) for the unperturbed system, which implies two things as the number of individuals or particles gets large. First, when they exist, quasi-stationary distributions concentrate on the positive attractors of the unperturbed system. Second, the expected time to extinction for systems starting according to this quasi-stationary distribution grows exponentially with the system size. The aim of this paper is to extend their approach to a class of discrete time Markov process, that, up to a renormalization of time, can be seen as random perturbations of an ordinary differential equation.

In Section 2 we will introduce our setting and give some examples of systems that fall into it. Then, in Section 3 we will show that, under the hypothesis that the deterministic mean dynamic admits an interior attractor, the extinction time grows exponentially with the size of the system and that, when the system size goes to infinity, the limit set of the quasi-stationary distributions of the processes for the weak* convergence consists of invariant measures for the deterministic dynamic. Finally in Section 4 we will study the support of these invariant limiting measures and prove that, under some additional large deviations hypotheses, their support lies within attractors bounded away from the extinction states. To do that we will compare two

different notions of chain-recurrence, one given by the large deviations functional and the other a slight variation on Conley's δ-T chain-recurrence. Should the reader need some reminders about the objects used in this paper, he will find some basic properties and some references in the Appendix at the end.

2. Model, notations and hypotheses.

We denote by Δ the d-dimensional simplex.

$$\Delta = \{x \in \mathbb{R}^d \; ; \; \forall i = 1 \cdots d \;\; x_i \geqslant 0 \;\; \& \;\; \sum_{i=1}^{d} x_i = 1\}$$

We let $\mathring{\Delta}$ denote the relative interior of Δ and

$$\Delta_N = \Delta \cap \frac{1}{N}\mathbb{Z}^d$$

$$\Delta_N^* = \mathring{\Delta} \cap \frac{1}{N}\mathbb{Z}^d.$$

Let $F : \Delta \to \mathbb{R}^d$ be a locally Lipschitz vector field such that :

$$\forall x \in \Delta \quad \sum_{i=1}^{d} F_i(x) = 0$$

Unless specified otherwise, the topology considered will be the topology induced by the classical $\mathbb{R}^d$ metric topology on Δ. Throughout the paper, if A is a subset of a metric space (E, d), we will denote by $N^\varepsilon(A)$ its ε-neighborhood

$$N^\varepsilon(A) = \{x \in E \; ; \; d(x, A) < \varepsilon\}.$$

We consider a family of Markov chains $(X_n^N)_{n\in\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in the d-dimensional discrete simplex Δ_N.

We denote by $\mathcal{F}_n^N$ the σ-algebra generated by $\{X_i^N, i = 1, ..., n\}$. For $A \in \mathcal{F}$ we let $\mathbb{P}_x[A] = \mathbb{P}[A|X_0 = x]$.

Throughout the paper the following hypothesis will always be assumed to hold.

Standing Hypothesis 2.1 :
The Markov process X^N has the following properties :

(i) $X_{n+1}^N - X_n^N = \dfrac{1}{N}(F(X_n^N) + U_{n+1}^N)$

(ii) $\mathbb{E}[U_{n+1}^N|\mathcal{F}_n^N] = 0$

(iii) *There exists* $\Gamma \geqslant 0$ *such that* $\|U_n^N\| \leqslant \sqrt{\Gamma}$

(iv) *The boundary of the simplex is an absorbing set:*

(a) *for all* $x \in \partial\Delta$ $\mathbb{P}_x[X_1^N \in \partial\Delta] = 1$

(b) *for all* $x \in \Delta$ $\mathbb{P}_x[\exists n \,:\, X_n^N \in \partial\Delta] = 1$

(v) X^N *restricted to* Δ_N^* *is irreducible*

$$\forall x, y \in \Delta_N^* \quad \mathbb{P}_x[\exists n \,:\, X_n = y] > 0$$

and aperiodic

$$\forall x \in \Delta_N^* \quad gcd(\{n \,;\, \mathbb{P}_x[X_n = x] > 0\}) = 1$$

Example 1 (Guiding Thread)**:**
Let $(p_{i,j}(x))_{i,j\in\{1\dots d\}}$ be a family of real-valued continuous functions on Δ such that, for all $x \in \Delta$:

$$\forall i \neq j \quad p_{i,j}(x) = 0 \Leftrightarrow x_i x_j = 0, \tag{1}$$

$$p_{i,i}(x) = 0, \tag{2}$$

$$0 \leqslant p_{i,j}(x) \leqslant 1, \tag{3}$$

$$\sum_{i,j=1}^{d} p_{i,j}(x) \leqslant 1. \tag{4}$$

Let (X_k^N) be the random walk on Δ_N defined by:

$$\mathbb{P}\left[X_{k+1}^N = X_k^N + \frac{1}{N}(e_j - e_i)|X_k^n = x\right] = p_{i,j}(x)$$

where $(e_i)_{i=1\cdots d}$ is the canonical base of $\mathbb{R}^d$. This type of model often occurs in population games. In this setting N represents the size of the population. Each individual plays a pure strategy i and X^N represents then the vector of proportion of players of each strategy. The jump $X_{k+1}^N = X_k^N + \frac{1}{N}(e_j - e_i)$ means that an individual switches his strategy from i to j at time k. The conditions on the family $(p_{i,j})$ mean that :

- At each time k, it is always possible that a player switches from his strategy i to another strategy j that is currently in use in the population.
- No individual switches to an unused strategy. This makes sense for models based on strategy switching from imitations or models arising from ecology.

We define

$$p_i(x) = \sum_{j=1}^{d} p_{j,i}(x)$$
$$q_i(x) = \sum_{j=1}^{d} p_{i,j}(x)$$

$p(x)$ the vector of coordinates $p_i(x)$ and $q(x)$ the vector of coordinates $q_i(x)$.

In this case $F(x) = p(x) - q(x)$. Hypothesis 2.1(i) comes from the Markov property, 2.1(ii) and 2.1(iii) follow easily from the definition of the chain and 2.1(iv) and 2.1(v) follow from the fact that the functions $p_{i,j}$ are positive on the relative interior and vanish on the boundary.

A class of examples that falls in Example 1 setting is given by Imitative Protocols games, see e.g. [19]. Consider a population game with d pure strategies, we let $U(x) = (U_1(x), \cdots U_d(x))$ denote the vector of payoffs when the population is in state $x \in \Delta$ and $\overline{U}(x) = \sum_{i=1}^{d} x_i U_i(x)$ denotes the average payoff at population state x. Imitative protocols are of the form

$$p_{i,j}(x) = x_i x_j r_{ij}(U(x), x)$$

with some additional assumptions on r_{ij} to ensure that the problem is well-posed. Under such a protocol, at each time we pick an individual uniformly at random among the population and give him a revision opportunity. The opportunity unfolds as such

- The individual picks an opponent uniformly at random among the population (he/she can pick him/her-self) and observes his/her strategy.
- If the individual plays i and the opponent plays j then the individual switches from i to j with probability proportional to r_{ij}.

We now give some examples of Imitative Protocols

Example 2 (Pairwise Proportional Imitation)**:**
After selecting an opponent the agent imitates only if the opponent's payoff is higher than his own, doing so with probability proportional to the payoff difference.

$$p_{i,j}(x) = x_i x_j (U_j(x) - U_i(x))^+$$

where $(y)^+$ stands for $\max(y, 0)$.

The mean dynamic generated by this protocol is

$$\begin{aligned}
\dot{x} &= \sum_{j=1}^{d} p_{j,i}(x) - \sum_{j=1}^{d} p_{i,j}(x) \\
&= \sum_{j=1}^{d} x_j x_i \left(U_i(x) - U_j(x)\right)^+ - x_i \sum_{j=1}^{d} \left(U_j(x) - U_i(x)\right)^+ \\
&= x_i \sum_{j=1}^{d} x_j (U_i(x) - U_j(x)) \\
&= x_i (U_i(x) - \overline{U}(x))
\end{aligned}$$

We then get the well known replicator dynamic, a dynamic extensively studied in ecology and evolutionary game theory, see e.g. [11]

Example 3 (Aspiration and Random Imitation)**:**
A particular case of the former example is the aspiration and random imitation model, see e.g. [3], [4], [6] and [5]. At each time we pick an individual at random in the population and look at his/her "satisfaction", a payoff-like function $u_i(x)$ where i is the type of the drawn individual. If this satisfaction is lower than a certain aspiration level then switch to another type chosen at random in the population, otherwise stay at current type. The morality of this model is that, if your type isn't performant enough for your tastes then switch to another type. The aspiration levels are independent random variables uniformly distributed on intervals $[a_i(x), b_i(x)]$ with $a_i(x) \leqslant u_i(x) \leqslant b_i(x)$. This model gives us

$$p_{i,j}(x) = x_i x_j \frac{b_i(x) - u_i(x)}{b_i(x) - a_i(x)}$$

If we assume that the aspiration level bounds are not type-dependant, meaning $a_i(x) = a(x)$ and $b_i(x) = b(x)$ for all i then we get a mean field given by

$$\dot{x}_i = x_i \frac{u_i(x) - \sum_{j=1}^{d} x_j u_j(x)}{b(x) - a(x)}$$

which is a replicator dynamic with fitness functions $f_i(x) = \frac{u_i(x)}{b(x)-a(x)}$

Alternatively we can assume that the aspiration levels follow the type payoff by the relation $b_i(x) = \beta_i u_i(x)$ and $a_i(x) = \alpha_i u_i(x)$ with $\alpha_i < 1 < \beta_i$.

In this case we get a dynamic

$$\dot{x}_i = x_i \left(\sum_{j=1}^{d} x_j v_j - v_i \right)$$

where $v_j = \frac{\beta_j - 1}{\beta_j - \alpha_j}$.

Again this is a replicator dynamics with fitness function $f_i(x) = -v_i$.

Example 4 (Imitation Driven by Dissatisfaction)**:**
In this protocol, when a i player receives a revision opportunity, he opts to switch strategies with a probability that that is linearly decreasing in his current payoff. Should he decide to change, then he will imitate a randomly selected opponent. This protocol also falls under the former example of aspiration games with constant aspiration levels $a_i(x) = A$, $b_i(x) = B$. This gives the following dynamic

$$\dot{x}_i = \frac{1}{B - A}(F_i(x) - \overline{F}(x))$$

Again we recognize a replicator dynamic.

For other examples see e.g. [19].

3. Convergence of QSD and absorption time.

We denote by $\{\varphi_t\}$ the flow induced by F. In order to compare the trajectory of φ_t with those of (X_n^N) it's convenient to introduce the continuous process $\hat{X}^N : \mathbb{R} \to \mathbb{R}^m$ defined by

$$\hat{X}^N(k/N) = X_k^N \quad \forall k \in \mathbb{N}$$

and extended on every interval $[k/N, (k+1)/N]$ by piecewise linear interpolation.

Let

$$D^N(T) = \max_{0 \leqslant t \leqslant T} \|\hat{X}^N(t) - \varphi_t(X_0^N)\|$$

be the variable measuring the distance between the trajectories $t \mapsto \hat{X}^N(t)$ and $t \mapsto \varphi_t(X_0^N)$.

We recall this convergence theorem of Benaïm and Weibull.

Theorem 3.1 :
For every $T > 0$, there exists $c > 0$ (depending only on F,Γ and T) such that, for every $\varepsilon > 0$, and for N large enough :

$$\mathbb{P}[D^N(T) \geqslant \varepsilon] \leqslant 2de^{-\varepsilon^2 cN}$$

For a detailed proof of this result, see [3].

We define T_0^N to be the *absorption time*.

$$T_0^N = \inf\{n > 0; X_n^N \in \partial\Delta\}$$

Hypothesis 2.1(iv) implies that, whatever the initial state is, the process will almost surely be absorbed, i.e.

$$\forall x \in \Delta^N \qquad \mathbb{P}_x[T_0^N < \infty] = 1.$$

A probability measure μ on the discrete relative interior of the simplex Δ_N^* is said to be a *quasi-stationary distribution*, thereafter referred as QSD, if and only if, for every Borelian set $A \subset \Delta_N^*$ and every $n > 0$,

$$\mathbb{P}_\mu[X_n^N \in A | T_0^N > n] = \mu(A).$$

We remark that, in this case, μ is a fixed point for the conditional evolution

$$\nu \mapsto \mathbb{P}_\nu[X_n^N \in \bullet | T_0 > n]$$

The following proposition is a classic QSD result and follows easily from the Perron-Frobenius theorem.

Proposition 3.2 :
For every N, there exists an unique quasi-stationary distribution μ^N obtained as the only left eigenvector μ^N of the transition matrix of the Markov chain restricted to Δ_N^* verifying

$$\forall i \in \{1, \cdots d\} \quad \mu_i^N > 0 \quad ; \quad \sum_i \mu_i^N = 1$$

The corresponding eigenvalue $0 < \rho_N = e^{-\theta_N} < 1$ is such that

$$\mathbb{P}_{\mu^N}[T_0^N > n] = e^{-\theta_N n}$$

Hence, starting from μ^N, the expectation of T_0^N is

$$\mathbb{E}_{\mu^N}[T_0^N] = \frac{1}{1 - \rho_N}$$

<u>Proof :</u>
For a detailed proof see e.g. [15].

■

3.1. *Absorption time.*

A set $A \subset \Delta$ is called an *attractor* for the flow $\{\varphi_t\}$ if

(i) A is compact and invariant, i.e. for every $t \in \mathbb{R}$ $\varphi_t(A) = A$.
(ii) There exists a neighborhood U of A, called a fundamental neighborhood, such that
$$\lim_{t\to\infty} d(\varphi_t(x), A) = 0$$
uniformly in x in U.

Theorem 3.3 :
Starting from μ^N, the law of the absorption time and its expectation are given by Proposition 3.2. If we further assume that the flow $\{\varphi_t\}$ admits an attractor $A \subset \mathring{\Delta}$, then, there exists $\gamma > 0$ such that the following estimate holds :
$$0 \leqslant 1 - \rho_N \leqslant O\left(\frac{e^{-\gamma N}}{N}\right)$$
Thus, there exists a constant $C > 0$ such that
$$\mathbb{E}_{\mu^N}[T_0] \geqslant CNe^{\gamma N}$$

<u>Proof :</u>
Let $V \subset \mathring{\Delta}$ and let $k \in \mathbb{N}$. By the QSD property we have:
$$\begin{aligned}\rho_N^k \mu^N(V) &= \sum_{x\in\Delta_N} p_N^k(x,V)\mu^N(x)\\ &\geqslant \sum_{x\in V\cap\Delta_N} p_N^k(x,V)\mu^N(x)\\ &\geqslant \inf_{x\in V\cap\Delta_N} p_N^k(x,V)\mu^N(V).\end{aligned}$$

Thus
$$\rho_N^k \geqslant \inf_{x\in V\cap\Delta_N} p_N^k(x,V).$$

Let $U \subset \mathring{\Delta}$ be a fundamental neighborhood of the attractor A. We know that $d(\varphi_t(x), A)$ converges uniformly to 0 over U. Hence
$$\forall\varepsilon > 0 \quad \exists T(\varepsilon) > 0 \quad \forall t \geqslant T(\varepsilon) \quad \forall x \in U \quad d(\varphi_t(x), A) < \varepsilon.$$

Let $\alpha = d(A, U^c)$, $\varepsilon < \alpha$, $T = T(\varepsilon)$ and $\delta < \alpha - \varepsilon$.

For all $x \in U \cap \Delta_N$

$$\begin{aligned}
p_N^{[NT]}(x, U^c) &\leqslant \mathbb{P}_x[d(X_{[NT]}^N, A) > \alpha] \\
&\leqslant \mathbb{P}_x[d(X_{[NT]}^N, \varphi_T(x)) > \alpha - \varepsilon] \\
&\leqslant \mathbb{P}_x[d(X_{[NT]}^N, \hat{X}_T^N) + d(\hat{X}_T, \varphi_T(x)) > \alpha - \varepsilon] \\
&\leqslant \mathbb{P}_x\left[D_N(T) > \alpha - \varepsilon - (\|F\| + \sqrt{\Gamma})\frac{NT - [NT]}{N}\right] \\
&\leqslant 2de^{-\delta^2 cN} \text{ for } N \text{ large enough (see Theorem 3.1)}
\end{aligned}$$

Then

$$\begin{aligned}
\rho_N^{[NT]} &\geqslant \inf_{x \in U \cap \Delta_N} p_N^{[NT]}(x, U) \\
&\geqslant 1 - \max_{x \in U \cap \Delta_N} p_N^{[NT]}(x, U^c) \\
&\geqslant 1 - 2de^{-\delta^2 cN}
\end{aligned}$$

Therefore

$$1 - \rho_N \leqslant 1 - \left(1 - 2de^{-\delta^2 cN}\right)^{\frac{1}{[NT]}}$$

$$\begin{aligned}
\left(1 - 2de^{-\delta^2 cN}\right)^{\frac{1}{[NT]}} &= e^{\frac{1}{[NT]} \log\left(1 - 2de^{-\delta^2 cN}\right)} \\
&= e^{\frac{-1}{[NT]} 2de^{-\delta^2 cN}} + o(e^{-\delta^2 cN})) \\
&= 1 - \frac{2de^{-\delta^2 cN}}{[NT]} + o\left(\frac{2de^{-\delta^2 cN}}{[NT]}\right)
\end{aligned}$$

In conclusion we have

$$0 \leqslant 1 - \rho_N \leqslant O\left(\frac{e^{-\gamma N}}{N}\right)$$

■

3.2. *Convergence of the QSD to an invariant measure.*

A probability measure μ on Δ is called an *invariant measure* for the flow $\{\varphi_t\}$ if, for all $t \in \mathbb{R}$ and all borelian set $A \in \mathcal{B}(\Delta)$, $\mu(\varphi_t^{-1}(A)) = \mu(A)$.

Theorem 3.4 :
We suppose that the flow $\{\varphi_t\}$ admits an attractor $A \subset \mathring{\Delta}$. Then the set of limit points of $\{\mu^N\}$ for the weak topology is a subset of the set of invariant measures for the flow $\{\varphi_t\}$.*

Proof :

Let f be a Lipschitz function from Δ to $\mathbb{R}$ with constant L. We suppose that the sequence μ^N weakly converges to a measure μ. Let $t > 0$. We want to prove that

$$\lim_{N\to\infty} \int f(x)\mu^N(dx) - \int f(\varphi_t(x))\mu^N(dx) = 0$$

The QSD property gives us that, for all k

$$\mathbb{E}_\mu[f(X_n)] = \int f(x)\mu^N(dx) = \int \mathbb{E}_x\left[f(X_k^N)\middle| T_0^N > k\right]\mu^N(dx)$$

Let

$$I = \left|\int f(x)\mu^N(dx) - \int f(\varphi_t(x))\mu^N(dx)\right|$$

Then, for all k,

$$\begin{aligned} I &= \left|\int f(x)\mu^N(dx) - \int f(\varphi_t(x))\mu^N(dx)\right| \\ &= \left|\int \mathbb{E}_x\left[f(X_k^N)\middle| T_0^N > k\right]\mu^N(dx) - \int f(\varphi_t(x))\mu^N(dx)\right| \\ &= \left|\int \mathbb{E}_x\left[f(X_k^N) - f(\varphi_t(x))\middle| T_0^N > k\right]\mu^N(dx)\right| \end{aligned}$$

In particular, for $k = [Nt]$.

$$I = \left|\int \mathbb{E}_x\left[f(X_{[Nt]}^N) - f(\varphi_t(x))\middle| T_0^N > [Nt]\right]\mu^N(dx)\right|$$

By Theorem 3.1, we know that, for N large enough, we have

$$\mathbb{P}_x[D_N(t) > \delta] \leqslant 2de^{-\delta^2 cN}.$$

Thus

$$\mathbb{E}_x[D_N(t)] = \int_0^{+\infty} \mathbb{P}_x[D_N(t) > \delta]d\delta \leqslant \int_0^{+\infty} 2de^{-\delta^2 cN}d\delta = \frac{d\sqrt{\pi}}{\sqrt{cN}}$$

From Theorem 3.3, we can infer $1-\rho_N \leqslant O\left(\frac{e^{\gamma N}}{N}\right)$.

This implies $e^{\theta_N[Nt]}-1 \underset{N\to+\infty}{\longrightarrow} 0$, which gives us the boundedness of $e^{\theta_N[Nt]}$.

Hence

$$\begin{aligned}
I &= \left|\int \mathbb{E}_x\left[f(X^N_{[Nt]})-f(\varphi_t(x))\middle| T^N_0>[Nt]\right]\mu^N(dx)\right| \\
&\leqslant \left|\int \frac{\mathbb{E}_x\left[f(X^N_{[Nt]})-f(\varphi_t(x))\right]}{\mathbb{P}_x\left[T^N_0>[Nt]\right]}\mu^N(dx)\right| \\
&\leqslant \left|\int \frac{\mathbb{E}_x\left[L|X^N_{[Nt]}-\hat{X}_t+\hat{X}_t-\varphi_t(x)|\right]}{\mathbb{P}_x\left[T^N_0>[Nt]\right]}\mu^N(dx)\right| \\
&\leqslant \left|\int \frac{\mathbb{E}_x\left[L(D_N(t)+(Nt-[Nt])\frac{1}{N}(\|F\|+\Gamma))\right]}{\mathbb{P}_x\left[T^N_0>[Nt]\right]}\mu^N(dx)\right| \\
&\leqslant \left|L\left(\frac{d\sqrt{\pi}}{\sqrt{cN}}+\frac{\|F\|+\Gamma}{N}\right)e^{\theta_N[Nt]}\right| \underset{N\to+\infty}{\longrightarrow} 0
\end{aligned}$$

■

4. Support of the limiting measure.

Let $L=L(\{\mu_N\})$ denote the limit set of the sequence $(\mu_N)_{N\in\mathbb{N}}$ for the weak* topology. In view of Theorem 3.4, L consists of invariant measures. As the QSD have their support inside $\mathring{\Delta}$, it is natural to study whether the limiting measure also take their support in $\mathring{\Delta}$. However, by the Poincaré Recurrence Theorem, every $\mu\in L$ is supported by the Birkhoff center

$$BC(\varphi)=\overline{\{x\in\Delta\ ;\ x\in\omega(x)\}}$$

Since the boundary of Δ intersects the Birkhoff center (e.g the vertices of the simplex are equilibria and thus inside the Birkhoff center), knowing that the QSD converges to an invariant measure is not enough, we have to further study the support of the measure μ to ensure that it is strictly inside the interior of the simplex. For that we will need large deviation assumptions.

Hypothesis 4.1 :
For all $\alpha>0$, there exists a function

$$\begin{aligned}
S: V_\alpha\times\mathbb{R}\times\mathcal{C}_x([0,T],V_\alpha) &\longrightarrow \overline{\mathbb{R}}_+ \\
(x,T,\phi) &\longmapsto S(x,T,\phi)
\end{aligned}$$

with the following properties, where $V_\alpha = \Delta \setminus \overline{N^\alpha(\partial\Delta)}$, $\mathcal{C}([0,T],V_\alpha)$ *is the set of continuous functions* ψ *from* $[0,T]$ *to* V_α *and* $\mathcal{C}_x([0,T],V_\alpha)$ *is the set of continuous functions* ψ *from* $[0,T]$ *to* V_α *such that* $\psi(0)=x$, *both equipped with the topology of uniform convergence.*

- *For every* $s \in]0,\infty[$ *and* $T>0$, *the set*

$$\{\phi \in \mathcal{C}_x([0,T],V_\alpha) \text{ s.t. } S(x,T,\phi) \leqslant s\}$$

  is a compact set
- *For* $x \in \mathring{\Delta}$ *and* $T>0$, $S(x,T,\phi)=0 \Leftrightarrow \dot{\phi}_s = F(\phi_s)\ \forall s \in [0,T]$
- $\hat{X}$ *satisfies a large deviation principle with rate function* S *and speed* $1/N$ *uniformly in* x *on compact subsets of* V_α,
  i.e. for K *compact subset of* V_α, $T>0$ *and* $A \subset \mathcal{C}_x([0,T],V_\alpha)$

$$-\sup_{x\in K}\inf_{\phi\in\mathring{A}} S(x,T,\phi) \leqslant \liminf_{N\to\infty}\frac{1}{N}log\inf_{x\in K}\mathbb{P}_x[\hat{X}^N \in A]$$

$$\limsup_{N\to\infty}\frac{1}{N}log\sup_{x\in K}\mathbb{P}_x[\hat{X}^N \in A] \leqslant -\inf_{x\in K}\inf_{\phi\in\overline{A}} S(x,T,\phi)$$

  In particular when $K=\{x\}$ *we get the "classical" large deviation principle*

$$-\inf_{\phi\in\mathring{A}} S(x,T,\phi) \leqslant \liminf_{N\to\infty}\frac{1}{N}log\mathbb{P}_x[\hat{X}^N \in A]$$

$$\limsup_{N\to\infty}\frac{1}{N}log\mathbb{P}_x[\hat{X}^N \in A] \leqslant -\inf_{\phi\in\overline{A}} S(x,T,\phi)$$

- S *is linear with regards to the concatenation of functions, i.e. if* $\tilde{T}<T$

$$S(x,T,\phi) = S(x,\tilde{T},\phi|_{[0,\tilde{T}]}) + S(\phi(\tilde{T}),T-\tilde{T},\phi|_{[\tilde{T},T-\tilde{T}]})$$

- $\lim_{T\to 0} S(x,T,\phi)=0$ *uniformly in* $x \in V_\alpha$ *and* $\phi \in \mathcal{C}_x([0,T],V_\alpha)$.

Hypothesis 4.2 :
$\forall c>0\ \exists U_c$ *an open neighborhood of* $\partial\Delta$ *such that*

$$\lim_{N\to\infty}\inf_{x\in U_c}\frac{1}{N}log\mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] \geqslant -c$$

Example 5:
Before going further we will verify that Hypotheses 4.1 and 4.2 hold for the nearest neighbor random walk model introduced in Example 1.

To show that Hypothesis 4.1 holds we will use Theorem 6.3.3 in [9] with Conditions 6.2.1 and 6.3.1, which will gives us a Laplace principle for $\hat{X}^N\big|_{[0,T]}$ that holds uniformly on compacts, and Theorem 1.2.3 in [9] will, in turn, give us the desired uniform on compacts large deviation principle.

We only need to show that Conditions 6.2.1 and 6.3.1 holds

When transcribing our model in [9] setting, we get, for $x \in V_\alpha$

$$\mu(dy|x) = \sum_{i \neq j} p_{i,j}(x)\delta_{e_j - e_i}(dy) + \left(1 - \sum_{i \neq j} p_{i,j}(x)\right)\delta_0(dy)$$

We define

$$H_\mu(x,\alpha) = log \int_{\mathbb{R}^d} exp\langle \alpha, y\rangle \mu(dy|x)$$

and

$$L_\mu(x,\beta) = \sup_{\alpha \in \mathbb{R}^d} \{\langle \alpha, \beta\rangle - H_\mu(x,\alpha)\}$$

We refer to Chapter 6.2 in [9] for elementary properties of these functions.

Definition 4.3:
$\mu(dy|x)$ is said to verify Condition A (called 6.2.1 in [9]) if

(i) For each $\alpha \in \mathbb{R}^d$, $\sup_{x \in \mathbb{R}^d} H_\mu(x,\alpha) < \infty$
(ii) The function mapping $x \in \mathbb{R}^d \mapsto \mu(.|x)$ is continuous in the topology of weak convergence.

Definition 4.4:
$\mu(dy|x)$ is said to verify Condition B (called 6.3.1 in [9]) if

(i) The relative interior of the convex hull of the support of $\mu(dy|x)$, $Ri(Conv(Supp(\mu(.|x))))$, doesn't depend on x.
(ii) $0 \in Ri(Conv(Supp(\mu(.|x))))$

Let $\alpha > 0$, we'll now prove the L.D.P. on V_α The conditions, as they are written, demand that x may take values in all of $\mathbb{R}^d$ while, in our model, we only use Δ. To remedy to that we'll first embed our d-dimensional simplex in $\mathbb{R}^{d-1}$ and then extend μ to a kernel η defined on all of $\mathbb{R}^{d-1}$ by taking $\eta(dy|x) = \mu(dy|p_\alpha(x))$, where p_α is the convex projection on $\overline{V_\alpha}$.

This way we have a probability kernel η that is defined on all of $\mathbb{R}^{d-1}$, it is then easy to verify that Conditions A and B hold for η. We thus get a

LDP for η on all of $\mathbb{R}^{d-1}$ with speed $1/N$ and rate function

$$S_\eta = \begin{cases} \int_0^T L_\eta(\phi(t), \dot{\phi}(t))dt \text{ if } \phi \text{ is uniformly continuous} \\ 0 \text{ otherwise} \end{cases}$$

We now remark that, when $x \in V_\alpha$, $\eta(dy|x) = \mu(dy|x)$ and thus $H_\eta(x, v) = H_\mu(x, v)$ and $L_\eta(x, u) = L_\mu(x, u)$.

From that we deduce that, if $\phi \in \mathcal{C}([0, T], V_\alpha)$, then $L_\eta(\phi, \dot{\phi}) = L_\mu(\phi, \dot{\phi})$ and finally $S_\eta(x, T, \phi) = S_\mu(x, T\phi)$.

We finally get that Hypothesis 4.1 holds for our nearest neighbor random walk with the following rate function.

$$S(x, T, \phi) = \begin{cases} \int_0^T L_\mu(\phi(t), \dot{\phi}(t))dt \text{ if } \phi \text{ is uniformly continuous} \\ 0 \text{ otherwise} \end{cases}$$

We still have to verify that the wanted properties holds for this rate function:

- From Proposition 6.2.4 in [9] we get that, for every $s \in]0, \infty[$ and $T > 0$, the set

$$\{\phi \in \mathcal{C}_x([0, T], \Delta) \text{ s.t. } S(x, T, \phi) \leqslant s\}$$

  is a compact set.
- Let $x \in V_\alpha$, $T > 0$, it is already known that

$$S(x, T, \phi) = 0 \quad \Leftrightarrow \quad \dot{\phi}_s = F(\phi_s) \qquad \forall s \in [0, T]$$

- The LDP comes from Theorem 6.3.3 in [9]
- The linearity of $S(x, T, \phi)$ follows easily from it's definition as an integral.

Let's now prove that Hypothesis 4.2 holds too.

Proposition 4.5 :
Suppose that, for every couple $(i, j) \in \{1, \cdots d\}^2$ there exists $k \in \mathbb{N}$ such that $\dfrac{\partial^k p_{i_j}(x)}{\partial x_i^k} \neq 0$ on $\{x \in \Delta \; ; \; x_i = 0\}$. Then Hypothesis 4.2 holds

Proof :

We want to ensure that, for all $c > 0$ there exists an open neighborhood U_c of $\partial\Delta$ such that

$$\lim_{N \to \infty} \inf_{x \in U_c} \frac{1}{N} log \mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] \geqslant -c.$$

We know that $q_i(x)$ and $p_i(x)$ go to 0 as x_i goes to 0 and that there exists $k \in \mathbb{N}$ such that $\dfrac{\partial^k q_i(x)}{\partial x_i^k} > 0$ on $\{x \in \Delta \ ; \ x_i = 0\}$ and thus $\dfrac{\partial^k q_i(x)}{\partial x_i^k} > a$ on a sufficient small neighborhood of $\{x \in \Delta \ ; \ x_i = 0\}$ with $a > 0$.

Let $1 > b > 0$ and let $x \in U_c = \{x \ ; \ \exists i \ x_i < b\}$. We have

$$\begin{aligned}\mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] &\geqslant \mathbb{P}_x[X^N_{[Nb]} \in \partial\Delta] \\ &\geqslant \prod_{j=1}^{[Nb]} q_i(x^{(j)})\end{aligned}$$

where $(x^{(j)})_j$ is a sequence of points in Δ_N such that $x_i^{(j)} = \frac{j}{N}$.

If b is small enough we have, for $j \leqslant [Nb]$

$$q_i(x^{(j)}) \geqslant a \left(\frac{j}{N}\right)^k$$

Then

$$\prod_{j=1}^{[Nb]} q_i(x^{(j)}) \geqslant a^{[Nb]} \frac{[Nb]!}{N^{[Nb]}}$$

Thus

$$\begin{aligned}\inf_{x \in U_c} \frac{1}{N} log\mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] \geqslant & C \left(\frac{[Nb]}{N}(log(a) - 1) \right. \\ & \left. + \frac{[Nb]}{N} log \frac{[Nb]}{N} + \frac{1}{2N} log(2\pi[Nb]) \right)\end{aligned}$$

where C is a constant

As N goes to infinity, the right-hand term goes to $Cb(log(a)-1)+Cblog(b)$ which is greater than $-c$ for b small enough. Thus, for N large enough

$$\lim_{N \to \infty} \inf_{x \in U_c} \frac{1}{N} log\mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] \geqslant -c$$

Hence our random walk model satisfies Assumption 4.2.

■

Finally both hypotheses holds for our model.

Definition 4.6:
We define

$$\mathcal{L}(x, y) = \limsup_{T \to \infty} \inf_{\phi \in \mathcal{C}_x([0,T]), \, \phi(T) = y} S(x, T, \phi)$$

We will say that x $\mathcal{L}$-leads to y (denoted by $x \rightsquigarrow_{\mathcal{L}} y$) if, for every $\varepsilon > 0$, there exists a path of points $x = \xi_0, \xi_1, \cdots \xi_{n(\xi)} = y$ such that

$$A_{n(\xi)}(\xi) = \sum_{k=1}^{n(\xi)} \mathcal{L}(\xi_k, \xi_{k+1}) < \varepsilon$$

We define $B_{\mathcal{L}}(x, y) = \inf_{\xi \text{ linking } x \text{ to } y} A_{n(\xi)}(\xi)$.

Thus $x \rightsquigarrow_{\mathcal{L}} y$ iff $B_{\mathcal{L}}(x, y) = 0$

We will say that x is $\mathcal{L}$-chain recurrent if $x \rightsquigarrow_{\mathcal{L}} x$ and will denote by $\mathcal{R}_{\mathcal{L}}(\varphi)$ the set of all $\mathcal{L}$-chain recurrent points.

If x and y are two points of $\mathcal{R}_{\mathcal{L}}$ verifying $x \rightsquigarrow_{\mathcal{L}} y$ and $y \rightsquigarrow_{\mathcal{L}} x$ we will then denote $x \sim_{\mathcal{L}} y$. The equivalence classes for this relation will be called $\mathcal{L}$-basic classes. We define a partial order on these classes by $[x] \prec_{\mathcal{L}} [y]$ if $x \rightsquigarrow_{\mathcal{L}} y$. A maximal $\mathcal{L}$-basic class will be called a $\mathcal{L}$-quasi-attractor.

Hypothesis 4.7 :
There is only a finite number of $\mathcal{L}$-basic classes in $\mathring{\Delta}$ denoted by K_i, $i = 1 \cdots \nu$. We suppose that they are closed sets and indexed in such a way that the k first $\{K_i\}_{i=1\cdots k}$ are the $\mathcal{L}$-quasi-attractors and the $\nu - k$ others aren't.

Proposition 4.8 :
The function $\mathcal{L}$ has the following properties : for every sequence $(x_n)_{n\in\mathbb{N}} \in V_\alpha$ converging to an $x \in V_\alpha$ and every $y \in V_\alpha$ we have

$$\lim_{n\to\infty} \mathcal{L}(x_n, y) = \mathcal{L}(x, y) \qquad \lim_{n\to\infty} \mathcal{L}(y, x_n) = \mathcal{L}(y, x)$$

<u>Proof :</u>
This proposition follows easily from Hypothesis 4.1

■

The following theorem is the main result of this section, giving us more insight in the support of the limiting measure μ.

Theorem 4.9 :
We suppose that the flow $\{\varphi_t\}$ associated with the mean dynamic $\dot{x} = F(x)$ has an interior attractor. Under Hypotheses 4.1, 4.2 and 4.7, we have :

The limiting measure μ has its support in the union of the $\mathcal{L}$-quasi-attractor, i.e. in $\bigcup_{i=1}^{k} K_i$

We will prove this theorem under an intermediary set of hypotheses then we will prove that the announced hypotheses imply the intermediary hypotheses.

4.1. *Absorption-preserving pseudo-orbit.*
Here we introduce a different notion of chains for our dynamical system using absorption preserving δ, T pseudo-orbit, an analog to δ, T pseudo-orbits introduced by Conley [8], which have been extensively studied in the past.

Definition 4.10:
Let $\{\varphi_t\}_{t\in\mathbb{R}}$ be a flow given by an ordinary differential equation on (Δ, d) for which $\partial\Delta$ is an invariant set. We will call (δ, T) absorption preserving pseudo-orbit (δ, T-ap-pseudo orbit) from x to y a piecewise continuous path

$$\begin{aligned} x = x_0, &\{\varphi_t(x_1)\,;\, t \in [0, t_1]\}, \{\varphi_t(x_2)\,;\, t \in [0, t_2]\} \\ &\cdots \{\varphi_t(x_1)\,;\, t \in [0, t_k]\}, x_{k+1} = y \quad k \geqslant 1 \end{aligned}$$

which is uniquely defined by the sequences of points $x_0, \cdots x_{k+1}$ and times $t_1, \cdots t_k$ such that the following hypotheses hold:

$$\begin{cases} d(x, x_1) < \delta \\ d(\varphi_{t_j}(x_j), x_{j+1}) < \delta \quad \forall j = 1 \cdots k \\ t_i \geqslant T \quad \forall i = 1 \cdots k \\ x_j \in \partial\Delta \Rightarrow x_{j+1} \in \partial\Delta \quad \forall j = 0 \cdots k \end{cases}$$

We will denote then $x \rightsquigarrow_{ap,\delta,T} y$

If $x \rightsquigarrow_{ap,\delta,T} y$ for every $\delta > 0$ and every $T > 0$, we will denote $x \rightsquigarrow_{ap} y$

The point x will be said to be ap-chain recurrent if $x \rightsquigarrow_{ap} x$, we define $\mathcal{R}_{ap}(\varphi)$ the set of ap-chain recurrent points.

If x and y are two points of $\mathcal{R}_{ap}(\varphi)$ such that $x \rightsquigarrow_{ap} y$ and $y \rightsquigarrow_{ap} x$ we will write $x \sim_{ap} y$. The equivalence classes for this relation will be called ap-basic classes. We define a partial order on these classes by $[x] \prec_{ap} [y]$ if $x \rightsquigarrow_{ap} y$. A maximal ap-basic class will be called a ap-quasi-attractor.

The ap-chain recurrent points and the ap-basic classes have some interesting properties which will be of use when proving the result on the support of the limit measure μ. We enumerate and prove some of them.

Proposition 4.11 :
Let $x \in \mathcal{R}_{ap}$. Then $\overline{[x]}_{ap} \subset [x]_{ap} \cup \partial\Delta$. Moreover, for all $t > 0$ $[x]_{ap}$ is φ_t invariant.

Proof :

Let $y \in \overline{[x]}_{ap}$ and let $(y_k)_{k\in\mathbb{N}}$ be a sequence of elements of $[x]_{ap}$ converging to y. Suppose $y \notin \partial\Delta$.

Let $\delta > 0$ and $T > 0$ and let k such that $d(y_k, y) < \delta$.

There exists a δ, T pseudo-orbit $x, (x_1, t_1), \cdots, (x_n, t_n), y_k$ linking x to y_k. Thus $x, (x_1, t_1), \cdots, (x_n, t_n), y$ is a $2\delta, T$ pseudo-orbit linking x à y.

Hence $x \rightsquigarrow_{ap} y$.

A similar reasoning on δ, T-pseudo-orbit linking y_k to x gives us $y \rightsquigarrow_{ap} x$ and then $y \in [x]_{ap}$.

We show now that $[x]_{ap}$ is an invariant set.

Let $T, T', \varepsilon > 0$. $\varphi_{T'}$ is an uniformly continuous application. Let then $\delta < \varepsilon$ such that $d(x, y) < \delta$ implies $d(\varphi_{T'}(x), \varphi_{T'}(y)) < \varepsilon$.

We know that there exists a δ, T pseudo-orbit $x, (x_1, t_1), \cdots, (x_n, t_n), x$ linking x to x. Hence $x, (x_1, t_1), \cdots, (x_n, t_n + T'), \varphi_{T'}(x)$ is a ε, T pseudo-orbit linking x to $\varphi_{T'}(x)$.

Before proving the converse, let's just remark that, if $\delta_1 < \delta_2$ and $T_1 > T_2$ then every δ_1, T_1 pseudo-orbit is also a δ_2, T_2 pseudo-orbit.

Let's now suppose that $T > T'$

We consider again our δ, T pseudo-orbit $x, (x_1, t_1), \cdots, (x_n, t_n), x$ linking x to x. One can remark that $t_1 + t_2 - T' > T$.

Thus $\varphi_{T'}(x), (\varphi_{T'}(x_1), t_1 + t_2 - T), (x_3, t_3), \cdots, (x_n, t_n), x$ is a ε, T pseudo-orbit linking $\varphi_{T'}(x)$ to x if δ is chosen small enough.

Then $x \rightsquigarrow_{ap} \varphi_{T'}(x)$.

By composing the δ, T pseudo-orbits linking x and $\varphi_{T'}(x)$ by $\varphi_{-T'}$ we get that $x \sim_{ap} \varphi_T(x)$ for every T in $\mathbb{R}$.

■

Proposition 4.12 :

Let $x \in \Delta$. If $x \in \partial\Delta$ or $\omega(x) \subset \mathring{\Delta}$ then $\omega(x) \subset \mathcal{R}_{ap}$. From this we get that, for every x in Δ, $\omega(x) \cap \mathcal{R}_{ap} \neq \emptyset$.

Proof :

The first point is a well-known result for "classic" chain-recurrence and easily extended to ap-chain-recurrence. Furthermore, if $x \in \mathring{\Delta}$ and $\omega(x) \cap \partial\Delta \neq \emptyset$ then, by taking $y \in \omega(x) \cap \partial\Delta$, we get $\omega(y) \subset \omega(x)$ and $\omega(y) \subset \mathcal{R}_{ap}$.

■

Proposition 4.13 :

If $[x]_{ap}$ is maximal, then $x \rightsquigarrow_{ap} z$ if and only if $z \in [x]$. As a consequence we also get that every quasi-attractor is a closed set.

Proof :

This result is trivial as soon as $z \in \mathcal{R}_{ap}$, thus we only have to prove it for $z \notin \mathcal{R}_{ap}$. We know that $\omega(z) \cap \mathcal{R}_{ap} \neq \emptyset$ and that, if $u \in \omega(z) \cap \mathcal{R}_{ap}$ then $z \rightsquigarrow_{ap} u$ thus $x \rightsquigarrow_{ap} u$. Hence $u \in [x]_{ap}$, from that we get $x \rightsquigarrow_{ap} z \rightsquigarrow_{ap} u \rightsquigarrow_{ap} x$ which implies $z \in [x]_{ap} \subset \mathcal{R}_{ap}$.

■

The relation between being an attractor and being a quasi-attractor has been studied in the past, we recall this theorem of [1].

Proposition 4.14 :
Let C be a non-empty subset of Δ. The following assertions are equivalent :

(i) C is an irreducible attractor i.e. it doesn't contain any proper attractor.
(ii) C is an isolated quasi-attractor, i.e. there exists U, an open neighborhood of C, such that $U \cap \mathcal{R}_{ap} = C$.
(iii) C is an isolated connected component of $\mathcal{R}_{ap}$ and

$$C^+ = \{x \in M \,;\, \exists y \in \mathcal{R}_{ap} \setminus C \text{ s.t. } C \rightsquigarrow_{ap} y \rightsquigarrow_{ap} x\} = \emptyset$$

This result is proved in Part 5 of [1].

The following hypothesis is an analog of Hypothesis 4.1 adapted to the context of ap-pseudo-orbits.

Hypothesis 4.15 :
There is only a finite number of ap-basic classes in $\mathring{\Delta}$ denoted by K_i, $i = 1 \cdots \eta$. We suppose that they are closed sets and indexed in such a way that the k first $\{K_i\}_{i=1\cdots k}$ are the quasi-attractors and the $\eta - k$ others aren't.

Proposition 4.16 :
For every θ small enough, there exist $\delta(\theta) \in [0, \theta[$ and $T(\theta) \in]0, \infty]$ with $\delta(\theta), T(\theta) \neq (0, \infty)$ such that:

If there exists a δ, T pseudo-orbit $\xi_0 \cdots \xi_n$ verifying, $\delta < \delta(\theta)$ or $T > T(\theta)$ and, for a certain triplet $(i, i', j) \in \{1 \cdots \eta\}^3$,

$$d(\xi_0, K_i) < \delta \quad d(\xi_n, K_{i'}) < \delta \quad d(\xi_j, K_i) > \theta$$

Then $i \neq i'$ and $K_i \prec K_{i'}$.

This proposition means that, for δ small and T large, the δ, T pseudo-orbits respect the partial order.

Proof :

Suppose that, for every $\delta > 0$ and every $T > 0$, there exists a δ, T-pseudo-orbit $\xi_0 \cdots \xi_n$ such that

$$d(\xi_0, K_i) < \delta \quad d(\xi_n, K_{i'}) < \delta$$

then, we can construct δ, T-pseudo-orbits going from K_i to $K_{i'}$ which in turn implies $K_i \rightsquigarrow_{ap} K_{i'}$

Hence, if $K_i \not\rightsquigarrow_{ap} K_{i'}$, there exists $\tilde{\delta} \geqslant 0$ and $\tilde{T} \in]0, +\infty]$ such that a δ, T pseudo-orbit $\xi_0 \cdots \xi_n$ verifying

$$d(\xi_0, K_i) < \delta \quad d(\xi_n, K_{i'}) < \delta$$

may only exist if $\delta > \tilde{\delta}$ or $T < \tilde{T}$.

Suppose now that $i = i'$, we will show that there exists $\hat{\delta}$ and $\hat{T}$ such that every δ, T pseudo-orbit $\xi_0 \cdots \xi_n$ verifying either $\delta < \hat{\delta}$ or $T > \hat{T}$ and

$$d(\xi_0, K_i) < \delta \quad d(\xi_n, K_i) < \delta$$

doesn't contain any point at a distance greater than θ from K_i.

Suppose first that this assertion is false, thus we have real sequences $\delta_l \to 0$ and $T_l \to \infty$ and a sequence of δ_l, T_l pseudo-orbits verifying

$$d(\xi_0^l, K_i) < \delta \quad d(\xi_{n_l}^l, K_i) < \delta \quad d(\xi_{j_l}^l, K_i) > \theta.$$

Δ being a compact set, we may suppose that, up to an extraction, as l goes to infinity $x_0^l \to x \in K_i$, $x_{n_l}^l \to z \in K_i$ and $x_{j_l}^l \to y$ with $d(y, K_i) \geqslant \theta$. In that case we get $x \rightsquigarrow_{ap} y \rightsquigarrow_{ap} z$ and thus $y \in K_i$, which is absurd.

■

Proposition 4.17 :
For every $\delta > 0$, there exists $T_0 > 0$ such that every δ, T_0 pseudo-orbit intersects $N^\delta(\mathcal{R}_{ap})$.

<u>Proof :</u>

Let $x \in \Delta$ and $\gamma > 0$, we define $T^\gamma(x) = Inf\{t \geqslant 0\,;\, \varphi_t(x) \in N^\gamma(\mathcal{R}_{ap})\}$

As $\omega(x) \cap \mathcal{R}_{ap} \neq \emptyset$ we get $T^\gamma(x) < +\infty$.

For $\alpha > 0$ we will denote $N_\alpha = \{x \in \Delta\,;\, T^\gamma(x) \geqslant \alpha\}$ the level sets of T^γ

Let us show that N_α is closed. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of elements of N_α converging to y. By the continuity of φ_t we obtain then that, for every $t > 0$, $\lim\limits_{n\to\infty} \varphi_t(x_n) = \varphi_t(y)$. If $t < \alpha$ we get $\varphi_t(x_n) \in (N^\gamma(\mathcal{R}_{ap}))^c$ which is closed. Hence we have $\varphi_t(y) \notin N^\gamma(\mathcal{R}_{ap})$ for all $t < \alpha$ and thus $y \in N_\alpha$.

$T^\gamma(x)$ is then an upper semi-continuous function taking its values in $[0, +\infty[$, Δ being a compact set, we know then that $T^\gamma(x)$ attains its maximum on Δ which we will denote T^γ.

Taking $T_0 > T^\delta$ gives us the result.

■

The following corollary comes easily from the last two propositions.

Corollary 4.18 :
For every $\delta > 0$ and every $T > 0$ there exists a family V_i of open neighborhoods of the K_i and positive real numbers δ_1 and T_1 such that

- $\overline{N^{\delta_1}(K_i)} \subset V_i$ *for* $i = 1 \cdots n$
- *Every* δ_1, T_1 *pseudo-orbit starting from* V_i *stays in* V_i *for* $i = 1 \cdots k$
- *If there exists a* δ_1, T_1 *pseudo-orbit* $x, (\xi_1, t_1), \cdots, (\xi_p, t_p), y$ *with* $x \in N^{\delta_1}(K_i)$ *and* $y \in N^{\delta_1}(K_j)$ *such that*

$$\exists l \quad \exists \tilde{t} \in [0, t_l] \text{ such that } \varphi_{\tilde{t}}(\xi_l) \notin V_i$$

  Then $i \neq i'$ *and* $K_i \prec K_{i'}$.
- *Every* δ_1, T_1 *pseudo-orbit intersects* $N^{\delta}(\mathcal{R}_{ap})$

Definition 4.19:
For K compact subset of $\mathring{\Delta}$ we denote

$$\beta_{\delta,K}(N) = \sup_{x \in \Delta_N \cap K} \mathbb{P}_x[\hat{X}^N(1) \in \Delta \setminus N^{\delta}(\varphi_1(x))]$$

Proposition 4.20 :
If the flow $\{\varphi_t\}$ admits an attractor $A \subset \mathring{\Delta}$, then, for all K compact subset of $\mathring{\Delta}$ and neighborhood of A, there exists $\delta > 0$ such that $\rho_N^N \geqslant 1 - \beta_{\delta,K}(N)$. Moreover, if there exists V_K an open neighborhood of $\partial\Delta$ such that

$$\lim_{N \to \infty} \frac{\beta_{\delta,K}(N)}{\inf_{x \in V_K \cap \Delta_N} \mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta]} = 0$$

Then $\mu(V_K) = 0$.

<u>Proof :</u>

Let U be an open neighborhood of A and $\delta > 0$ such that $U \subset K$ and for every t, $\varphi_t(\overline{U}) \subset U$ and $N^{\delta}(\varphi_1(\overline{U})) \subset U$.

Then

$$\begin{aligned}
\rho_N^N \mu^N(U) &= \sum_{x\in\Delta_N} p_N^N(x,U)\mu^N(x) \\
&\geqslant \sum_{x\in U\cap\Delta_N} \inf_{x\in U\cap\Delta_N} p_N^N(x,U)\mu^N(x) \\
&\geqslant \mu^N(U)\left(1-\sup_{x\in U\cap\Delta_N} p_N^N(x,U^c)\right) \\
&\geqslant \mu^N(U)\left(1-\sup_{x\in U\cap\Delta_N} p_N^N(x,(N^\delta(\varphi_1(\overline{U})))^c)\right) \\
&\geqslant \mu^N(U)\left(1-\beta_{\delta,K}(N)\right)
\end{aligned}$$

We finally get $\rho_N^N \geqslant 1-\beta_\delta(N)$
From this we obtain

$$\begin{aligned}
1-\beta_{\delta,K}(N) &\leqslant \rho_N^N \mu^N(\mathring{\Delta}) \\
&\leqslant \sum_{x\in\Delta_N^*} \left(1-p_N^N(x,\partial\Delta)\right)\mu^N(x) \\
&\leqslant \mu^N(M\setminus V_K)+\mu^N(V_K)\left(1-\inf_{x\in V_0\cap\Delta_N} p_N^N(x,\partial\Delta)\right)
\end{aligned}$$

Hence

$$\mu^N(V_0) \leqslant \frac{\beta_{\delta,K}(N)}{\inf_{x\in V_0\cap\Delta_N} p_N^N(x,\partial\Delta)}$$

V_0 being an open set, the weak convergence of the measures μ^N gives us the desired result.

■

Hypothesis 4.21 :
We suppose that the flow φ_t admits an attractor $A\subset\mathring{\Delta}$ and that there exists K a compact neighborhood of A in $\mathring{\Delta}$ and V_K an open neighborhood of $\partial\Delta$ such that

$$\lim_{N\to\infty} \frac{\beta_{\delta,K}(N)}{\inf_{x\in V_K\cap\Delta_N} \mathbb{P}_x[\hat{X}^N(1)\in\partial\Delta]} = 0$$

The following assumption is a technical one but we will see later that it is true under the first set of hypotheses.

Hypothesis 4.22 :
Let $j \in \{k+1, \cdots, \nu\}$, i.e. such that K_j is not a quasi-attractor. Then

$$\exists \eta > 0 \quad \forall \gamma > 0 \quad \exists N_0 \quad \exists \chi : \mathbb{N} \to \mathbb{R} \quad \exists \zeta : \mathbb{N} \to \mathbb{R}$$

such that

$$\lim_{n\to\infty} \zeta(n) = \lim_{n\to\infty} \chi(n) = 0$$

and

$$\sup_{x \in N^{\eta}(K_j)} \mathbb{P}_x \left[\tau^N_{N^{\eta}(K_j)} > e^{\frac{\gamma}{\chi(N)}} \right] \leqslant \zeta(N)$$

We now arrive at the central theorem of this section.

Theorem 4.23 :
Under Hypotheses 4.15, 4.21 and 4.22, the limiting measure μ has its support inside the union of ap-quasi-attractors $\bigcup_{i=1}^{k} K_i$

<u>Proof :</u> Let $\delta > 0$ small enough (how small will be specified later)

K being an attractor, we know that $\rho_N \to 1$, that μ is φ-invariant and has its support inside $\mathcal{R}_{ap}$.

Hypothesis 4.21 and Proposition 4.20 gives us that $\mu(\partial\Delta) = 0$, hence $\mu(K) = 1$.

It only remains to be shown that, for every $j = k+1 \cdots \eta$, there exists an open neighborhood W_j of K_j such that $\mu(W_j) = 0$

We know that

$$\rho_N^{N^2} \mu^N(W_j) = \sum_{x \in M} \mathbb{P}_x[\hat{X}(N) \in W_j] = \int_M \mathbb{P}_x[\hat{X}^N_N \in W_j] d\mu^N(x)$$

We denote $t^i_N = \left[\frac{N}{i}\right]$ and define the following events :

$$E^{\delta}_N = \left\{ \exists \delta_1, T_1 \text{ pseudo-orbit closer than } \delta \text{ from } \hat{X}^N_t,\ t \in [0, N] \right\}$$

$$E'_N = \left\{ \forall i \in \{k+1, \cdots, \eta\}\ ,\ \forall q \geqslant N^2\ \hat{X}^N_p \in N^{\delta_1}(K_i) \Rightarrow \hat{X}^N_{p+q} \notin N^{\delta_1}(K_i) \right\}$$

E'_N is the event "after its first entry in $N^{\delta_1}(K_i)$ the Markov chain $\hat{X}$ will have left it after N^2 steps".

Let W_i be open neighborhoods of the K_i such that $N^{\delta}(W_i) \subset V_i$ for every $i = 1 \cdots \eta$. On $E^{\delta}_N \cap E'_N$ we get,

For $j \geqslant k$,

$$X_N^N \in W_j \Rightarrow \forall i \in \{2, \cdots b\} \quad \hat{X}_{t_N^i}^N \notin N^\delta(K)$$

Hence

$$\mathbb{P}[\hat{X}_N^N \in W_j | E_N^\delta \cap E'_N] \leqslant \mathbb{P}\left[\bigcap_{i=2}^{b} \{\hat{X}_{t_N^i}^N \notin N^{\delta_1}(K)\}\right]$$

On E_N^δ, starting from $N^{\delta_1-\delta}(K_i)$ the chain cannot enter $N^{\delta_1-\delta}(K_i)$ once it has left W_i^δ (and also V_i).

Thus

$$\begin{aligned}
\mathbb{P}[(E'_N)^c | E_N^\delta] &\leqslant \mathbb{P}[\exists i \text{ such that } \hat{X} \text{ doesn't leave } W_i \text{ before the time } N^2] \\
&\leqslant \sum_{i=k+1}^{\eta} \sup_{x \in N^{\delta_1-\delta}(K_i)} \mathbb{P}_x[\hat{\tau}_{V_i}^N \geqslant N^2] \\
&\leqslant \sum_{i=k+1}^{\eta} \sup_{x \in N^{\delta_1-\delta}(K_i)} \mathbb{P}_x[\hat{\tau}_{W_i}^N \geqslant N^2]
\end{aligned}$$

where $\hat{\tau}_U^N = Inf\{t \geqslant 0 \; ; \; \hat{X}_t^N \notin U\}$.

We now remark that, if A, B and C are three events, we get

$$\begin{aligned}
\mathbb{P}[C] &= \mathbb{P}[C \cap B^c \cap A] + \mathbb{P}[C \cap B \cap A] + \mathbb{P}[C \cap A^c] \\
&\leqslant \mathbb{P}[B^c \cap A] + \mathbb{P}[C | B \cap A] + \mathbb{P}[A^c] \\
&\leqslant \mathbb{P}[B^c | A] + \mathbb{P}[C | B \cap A] + \mathbb{P}[A^c]
\end{aligned}$$

From that, we obtain

$$\mathbb{P}_x[\hat{X}_N^N \in W_j] \leqslant \mathbb{P}_x[(E_N^\delta)^c] + \mathbb{P}_x[(E'_N)^c | E_N^\delta] + \mathbb{P}_x[\hat{X}_N^N \in W_j | E_N^\delta \cap E'_N]$$

It only remains to control $\mathbb{P}_x[(E_N^\delta)^c]$. In order to do that we will consider the pseudo-orbit $PO_N(t)$ defined by $x, (x, T_1), (\hat{X}_{T_1}^N, T_1) \cdots$

For N large enough, classic results on stochastic approximation algorithms (see e.g. [3]) give us the following estimate :

$$\begin{aligned}
\mathbb{P}_x[(E_N^\delta)^c] &\leqslant \mathbb{P}[PO_N(t) \text{ isn't a } \delta_1, T_1 \text{ pseudo-orbit}] \\
&\quad + \mathbb{P}_x[\sup_{t \in [0,N]} d(\hat{X}_t^N, PO_N(t)) > \delta] \\
&\leqslant O((1 - e^{-\varepsilon N})^N) = O(e^{N e^{-\varepsilon N}})
\end{aligned}$$

with $\varepsilon > 0$

In conclusion we get

$$\begin{aligned}
\mu^N(V_j) &= \frac{1}{\rho_N^N}\int_{\mathring{\Delta}} \mathbb{P}_x[\hat{X}_N^N \in W_j] d\mu^N(x) \\
&\leqslant \frac{1}{\rho_N^N}\int_{\mathring{\Delta}} \mathbb{P}_x[(E_N^\delta)^c] + \mathbb{P}_x[(E'_N)^c | E_N^\delta] d\mu^N(x) \\
&\qquad + \frac{1}{\rho_N^N}\int_{\mathring{\Delta}} \mathbb{P}_x[\hat{X}_N^N \in W_j | E_N^\delta \cap E'_N] d\mu^N(x) \\
&\leqslant \frac{1}{\rho_N^N}\left(O(e^{Ne^{-\varepsilon N}}) + \zeta(N) + \sum_{i=1}^{b} \int_{\mathring{\Delta}} \mathbb{P}_x[\hat{X}_{t_N^i} \notin N^\delta(K)] d\mu^N(x) \right) \\
&\leqslant \frac{1}{\rho_N^N}\left(O(e^{Ne^{-\varepsilon N}}) + \zeta(N) + \sum_{i=1}^{b} \rho_N^{Nt_N^i} \mu^N((N^\delta(K))^c) \right) \\
&\leqslant \frac{1}{\rho_N^N}\left(O(e^{Ne^{-\varepsilon N}}) + \zeta(N) + b\mu^N((N^\delta(K))^c) \right)
\end{aligned}$$

From Theorem 3.3, we can infer that $\frac{1}{\rho_N^N} \to 1$.

Hence we have $\lim_{N\to\infty} \mu^N(V_j) = 0$. As the sets V_j are open neighborhoods of the sets K_j we obtain $\mu(K_j) = 0$.

■

4.2. *Going back the hypotheses 4.1, 4.2 and 4.7.*
The aim of this section is to prove that our first set of hypotheses implies the second one, thus proving the announced Theorem 4.9. In particular we will show that $\mathcal{L}$-quasi-attractors and ap-quasi attractors are the same. In this section we will assume Hypotheses 4.1, 4.2 and 4.7 to be true.

Proposition 4.24 :
For every compact set $K \subset V_\alpha$, there exists a open neighborhood V_K of $\partial\Delta$ such that

$$\lim_{N\to\infty} \frac{\beta_{\delta,K}(N)}{\inf_{x\in V_K\cap\Delta_N} \mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta]} = 0$$

where $\beta_{\delta,K}(N) = \sup_{x\in K\cap\Delta_N} \mathbb{P}_x[\hat{X}^N(1) \notin N^\delta(\varphi_1(x))]$

Proof : Let K be a compact set in V_α and let $\delta_0 > 0$ such that

$$\delta_0 < \inf_{t\in[0,1], x\in K} d(\varphi_t(x), \partial\Delta)$$

Let $c(K) = \frac{1}{4} Inf\{S(x,1,\phi) | x \in K, d(\phi, \varphi(x)) \geqslant \delta_0\} > 0$.
For $x \in K$ we have

$$\begin{aligned}\limsup_{N\to\infty} \frac{1}{N} log\mathbb{P}_x \left[\hat{X}^N \in \{\phi \in \mathcal{C}_x([0,1]); d(\phi,\varphi(x)) > \delta_0\}\right] \\ \leqslant - \inf_{\phi; d(\phi,\varphi(x))>\delta_0} S(x,1,\phi) \\ \leqslant -4c(K)\end{aligned}$$

Thus, for N large enough and for every x in K.

$$\frac{1}{N} log\mathbb{P}_x \left[\hat{X}^N \in \{\phi \in \mathcal{C}_x([0,1]); d(\phi,\varphi(x)) > \delta_0\}\right] \leqslant -3c(K)$$

Moreover, there exists V_K, an open neighborhood of $\partial\Delta$ such that

$$\lim_{N\to\infty} \inf_{x\in V_K} \frac{1}{N} log\mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] \geqslant -2c(K)$$

Then, for N large enough, we have, for every x in V_K

$$\mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta] \geqslant e^{-\frac{2c(K)}{1/N}}$$

Finally we get

$$\lim_{N\to\infty} \frac{\beta_{\delta,K}(N)}{\inf_{x\in V_K} \mathbb{P}_x[\hat{X}^N(1) \in \partial\Delta]} \leqslant \lim_{N\to\infty} e^{-\frac{c(K)}{1/N}} = 0$$

■

Proposition 4.25 :
Let $T > 0$ and K be a compact subset of V_α. Then, for every $\delta > 0$, there exists $\varepsilon(K,T,\delta)$ such that

$$\forall x \in K \quad S(x,T,\phi) \leqslant \varepsilon \Rightarrow d(\phi, \varphi_.(x))_{[0,T]} \leqslant \delta$$

<u>Proof :</u> Let's suppose that this result is false, then

$$\exists\delta \quad \forall\varepsilon \quad \exists x \quad \exists\phi_\varepsilon \text{ s.t. } S(x,T,\phi_\varepsilon) < \varepsilon \text{ and } d(\varphi_.(x),\phi_\varepsilon) > \delta$$

We also know that $c = \inf\{S(x,T,\phi)\ ; x \in K\ ;\ d(\phi,\varphi_.(x)) > \delta\} > 0$
Thus, for every $\varepsilon > 0$

$$\varepsilon > S(X,T,\phi_\varepsilon) \geqslant c > 0$$

which is absurd.

■

Proposition 4.26 :
The function $B_{\mathcal{L}} : V_\alpha \times V_\alpha \to \mathbb{R}_+$ is upper semi-continuous.

Proof : Let x_n and y_n two sequences in V_α converging to x and y in V_α
For every $\delta > 0$ there exists a path $x = \xi_0, \cdots \xi_{n(\xi)} = y$ such that

$$B_{\mathcal{L}}(x, y) \leqslant A_{n(\xi)}(\xi) \leqslant B_{\mathcal{L}}(x, y) + \delta$$

Let's now consider the path ξ^n given by $x_n = \xi_0^N, \xi_1, \cdots \xi_{n(\xi)-1}, y_n = \xi_{n(\xi)}$
If we take n such that

$$|\mathcal{L}(x_n, \xi_1) - \mathcal{L}(x, \xi_1)| < \delta \text{ and } |\mathcal{L}(\xi_{n(\xi)-1}, y_n) - \mathcal{L}(\xi_{n(\xi)-1}, y)| < \delta$$

We obtain

$$B_{\mathcal{L}}(x_n, y_n) \leqslant A_{n(\xi)}(\xi^N) \leqslant B_{\mathcal{L}}(x, y) + 3\delta$$

Thus, for every $\delta > 0$, $\limsup_{n\to\infty} B_{\mathcal{L}}(x_n, y_n) \leqslant B_{\mathcal{L}}(x, y) + 3\delta$
Finally $\limsup_{n\to\infty} B_{\mathcal{L}}(x_n, y_n) \leqslant B_{\mathcal{L}}(x, y)$

■

Proposition 4.27 :
Let $x \in \mathcal{R}_{\mathcal{L}}$ such that $[x]_{\mathcal{L}}$ is a closed subset of V_α. Let $\theta > 0$ such that $N^\theta([x]_{\mathcal{L}}) \subset V_\alpha$.
Then,

$$\exists \delta > 0 \quad \exists T > 0 \quad \forall \Psi \text{ with } \Psi(0) \in [x]_{\mathcal{L}}, \quad \Psi(\tilde{T}) \in [x]_{\mathcal{L}}, \quad \tilde{T} > T,$$

$$\text{and } S(\Psi(0), \tilde{T}, \Psi) < \delta \quad \forall t \in [0, \tilde{T}] \qquad \Psi(t) \in N^\theta([x]_{\mathcal{L}}).$$

Proof : Let $x \in \mathcal{R}_{\mathcal{L}}$ such that $[x]_{\mathcal{L}}$ is a closed subset of V_α and let $\theta > 0$ such that $N^\theta([x]_{\mathcal{L}}) \subset V_\alpha$.

Suppose that the announced result is false, then, there exists a family of functions $\Psi_n \in \mathcal{C}([0, T_n], V_\alpha)$ such that

$$\Psi_n(0) \in [x]_{\mathcal{L}} \quad \Psi_n(T_n) \in [x]_{\mathcal{L}} \quad \lim_{n\to\infty} T_n = \infty \quad S(\Psi(0), T_n, \Psi) < 1/n$$

and

$$\forall n \quad \exists t_n > 0 \text{ such that } \Psi_n(t_n) \notin N^\theta([x]_{\mathcal{L}}).$$

$[x]_{\mathcal{L}}$ being a compact set, we can assume without loss of generality that $\Psi_n(0) \to u \in [x]_{\mathcal{L}}$ and $\Psi_n(T_n) \to v \in [x]_{\mathcal{L}}$.

Let $\tau_n = \inf\{t > 0\,;\, \Psi_n(t) \notin N^\theta([x]_{\mathcal{L}})\}$. Then we have

$$\Psi_n(\tau_n) \in \overline{V_\alpha} \setminus N^\theta([x]_{\mathcal{L}})$$

and $\overline{V_\alpha} \setminus N^\theta([x]_{\mathcal{L}})$ is a compact set.

Thus, without loss of generality, we can assume that

$$\Psi_n(\tau_n) \to w \in V_\alpha \setminus N^\theta([x]_{\mathcal{L}}).$$

Using the sequential continuity of $\mathcal{L}$ and the fact that u and v are $\mathcal{L}$-chain-recurrent points, we get that $u \rightsquigarrow_{\mathcal{L}} w$ and $w \rightsquigarrow_{\mathcal{L}} v$. Thus $w \in [x]_{\mathcal{L}}$ and we get a contradiction

■

Corollary 4.28 :
Suppose $\mathcal{R}_{\mathcal{L}}$ is a closed subset of V_α. Let $\theta > 0$ such that $N^\theta(\mathcal{R}_{\mathcal{L}}) \subset V_\alpha$ Then,

$$\exists \delta > 0 \quad \exists T > 0 \quad \forall \Psi \text{ with } \Psi(0) \in \mathcal{R}_{\mathcal{L}}, \quad \Psi(\tilde{T}) \in \mathcal{R}_{\mathcal{L}}, \quad \tilde{T} > T$$
$$\text{and } S(\Psi(0), \tilde{T}, \Psi) < \delta \quad \forall t \in [0, \tilde{T}] \qquad \Psi(t) \in N^\theta(\mathcal{R}_{\mathcal{L}}).$$

Proposition 4.29 :
Let $x \in \mathcal{R}_{\mathcal{L}}$ such that $[x]_{\mathcal{L}}$ is a closed subset of V_α. Then $x \in \mathcal{R}_{ap}$ and $[x]_{\mathcal{L}} \subset [x]_{ap}$.

Proof : Let $y \in [x]_{\mathcal{L}}$. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers and Ψ_n be a sequence of functions from $[0, T_n]$ to V_α such that $\Psi_n(0) = x$, $\Psi_n(T_n) = y$ and $S(x, T_n, \Psi_n) < 1/n$.

T_n is either a bounded sequence or it goes to infinity (up to a subsequence).

Let's suppose that $(T_n)_{n\in\mathbb{N}}$ is an increasing sequence of positive real numbers going to infinity and Ψ_n a sequence of functions such that $\Psi_n(0) = x$, $\Psi_n(T_n) = y$ and $S(x, T_n, \Psi_n) < 1/n$

The former proposition gives us a compact set $K \subset V_\alpha$, and positive numbers δ and T such that, for $T_n > T$ and $1/n < \delta$, Ψ_n lives in K

Let $\tilde{T} > T$, $\tilde{\delta} < \delta$ and $\varepsilon = \varepsilon(\tilde{\delta}, \tilde{T}, K)$ given by Proposition 4.25.

Then, from Proposition 4.25, we can infer that, for n large enough,

$$d(\Psi_n(\tilde{T}), \varphi_{\tilde{T}}(x)) \leqslant \delta$$

As $\Psi_n(\tilde{T}) \in K$, we also have

$$d(\varphi_{\tilde{T}}(\Psi_n(\tilde{T})), \Psi_n(2\tilde{T})) < \delta$$

By iterating this process we get that, for n large enough,

$$(x, \tilde{T}), (\Psi_n(\tilde{T}), \tilde{T}), \cdots$$

is a $\delta, \tilde{T}$ ap-pseudo orbit linking x to y.

Hence $x \rightsquigarrow_{ap} y$.

Let's suppose now that T_n is bounded by $\overline{T}$. By taking a sub-sequence we can assume that $T_n \to T$ as $n \to \infty$ for some T with $\overline{T} \geqslant T \geqslant 0$. we continue the function Ψ_n to the time $\overline{T}$ by concatenating the flow $\varphi_{\bullet}$. Then we know that $\Psi_n \to \varphi_{\bullet}(x)$ uniformly over $[0, \overline{T}]$ as $n \to \infty$. Hence $y = \lim_{n\to\infty} \Psi_n(T_n) = \varphi_T(x)$. Thus $T > 0$ and $y \in \gamma^+(x)$.

If $y = x$ we get then that, either $x \rightsquigarrow_{ap} x$ or x is periodic which implies $x \in \mathcal{R}_{ap}$.

Then, if $y \in [x]_{\mathcal{L}}$ we obtain that either $x \rightsquigarrow_{ap} y$ or $y \in \gamma^+(x) \subset [x]_{ap}$. Taking $y = x$ gives us $x \in \mathcal{R}_{ap}$, i.e. $\mathcal{R}_{\mathcal{L}} \subset \mathcal{R}_{ap}$. As the roles of x and y can be exchanged we also get $y \rightsquigarrow_{ap} x$ and in conclusion $[x]_{\mathcal{L}} \subset [x]_{ap}$.

■

Proposition 4.30 :
Closed $\mathcal{L}$-classes in V_α are positively invariant sets for the flow φ

<u>Proof :</u> Let $x \in \mathcal{R}_{\mathcal{L}}$ such that $[x]_{\mathcal{L}}$ is a closed $\mathcal{L}$ basic class in V_α. Let $T > 0$ and θ such that $N^\theta([x]_{\mathcal{L}}) \subset V_\alpha$.

If the paths linking x to itself have bounded length then x is periodic and in this case $\gamma^+(x) \subset [x]_{\mathcal{L}}$.

We now suppose that the paths have unbounded length.

Let $\varepsilon = \varepsilon([x]_{\mathcal{L}}, \tau, \theta)$ and let $\delta > 0$ and $\tau > 0$ given by Proposition 4.27, let Ψ be a path of length greater than τ and of cost smaller than $min(\varepsilon, \delta)$ linking x to itself.

Then, by the triangular inequality we get

$$d(\varphi_T(x), [x]_{\mathcal{L}}) \leqslant d(\varphi_T(x), \Psi(T)) + d(\Psi(T), [x]_{\mathcal{L}}) \leqslant 2\theta$$

Thus, by making θ go to zero, we obtain $\varphi_T(x) \in \overline{[x]_{\mathcal{L}}} = [x]_{\mathcal{L}}$;

■

Lemma 4.31 :
Let $\rightarrowtail$ be a binary, transitive relation on V_α, such that, for every $x \in V_\alpha$, every $y \in \alpha(x)$ and every $z \in \omega(x)$,

$$y \rightarrowtail x \quad x \rightarrowtail z \quad z \rightarrowtail z$$

Let $[x]_{\rightarrowtail}$ be a maximal, closed and isolated $\rightarrowtail$-basic class.

Then $[x]_{\rightarrowtail}$ is an attractor for the flow φ

Proof : Let W be an isolating open neighborhood of $[x]_{\rightarrowtail}$, i.e. W is an open neighborhood of $[x]_{\rightarrowtail}$ such that the only $\rightarrowtail$-recurrent points of W belong to $[x]_{\rightarrowtail}$.

If $[x]_{\rightarrowtail}$ isn't an attractor then there exists $p \in \partial\overline{W}$ such that

$$\gamma_-(p) \subset W \text{ and } \alpha(p) \subset [x]_{\rightarrowtail}$$

Let $y \in \omega(p)$, we have $p \rightarrowtail y$ and $y \rightarrowtail y$.

Similarly let $z \in \alpha(p)$, we have $z \rightarrowtail p$ and thus $z \rightarrowtail y$.

Let's define

$$[x]^+_{\rightarrowtail} = \{z \ , \exists y \in \mathcal{R}_{\rightarrowtail} \setminus [x]_{\rightarrowtail} \text{ and } x \rightarrowtail y \rightarrowtail z\}$$

As $[x]_{\rightarrowtail}$ is maximal, we have $[x]^+_{\rightarrowtail} = \emptyset$.

Yet we have $x \rightarrowtail y$ and $y \in \mathcal{R}_{\rightarrowtail}$.

Hence $y \in [x]_{\rightarrowtail}$, and, as $x \rightarrowtail p \rightarrowtail y$ we get $p \in [x]_{\rightarrowtail}$ which is absurd.

■

Proposition 4.32 :

Let $[x]_{\mathcal{L}}$ be a closed isolated quasi-attractor. Then $[x]_{\mathcal{L}}$ is an attractor.

Proposition 4.33 :

Let K be a compact subset of V_α.

For $\eta > 0$, $\delta > 0$, $\tilde{T} > 0$ there exists $N_0 > 0$ such that, for all $N > N_0$, all $x \in K$, all $T < \tilde{T}$ and all $\Psi \in \mathcal{C}_x([0,T],K)$

$$\mathbb{P}_x[d(\Psi, \hat{X}_N) < \eta] \geqslant exp\left(-\frac{S(x,T,\Psi)+\delta}{1/N}\right)$$

Proof : Let $\xi^K_{\gamma,T} = sup\{|S(x,T,\Psi) - S(x,T,\phi)| \, ; \, x,y \in K \, , \, d(\Psi,\phi) < \gamma\}$. The inferior semi-continuity of S gives us that $\lim\limits_{\gamma\to 0} \xi^K_{\gamma,T} = 0$

Let $\eta > 0$, $\delta > 0$, $\tilde{T} > 0$ and let $\gamma < \eta$ such that $\xi^K_{\gamma,\tilde{T}} < \delta$ and $N^\gamma(K) \subset V_\alpha$

The large deviation principle gives us that, for $\phi \in \mathcal{C}_x([0,T],K)$,

$$\liminf_{N\to\infty} \frac{1}{N} log \inf_{x\in K} \mathbb{P}_x[\hat{X} \in N^\gamma_x(\phi)]] \geqslant -\sup_{x\in K} \inf_{\Psi\in N^\gamma(\phi)} S(x,T,\Psi)$$

where $N^\gamma(\phi) = \{\Psi \in \mathcal{C}([0,T],K) \ ; \ \|\Psi - \phi\| < \gamma\}$. Thus, there exists $g_K(N)$ a function which goes to zero as N goes to infinity such that

$$\frac{1}{N} log \inf_{x\in K} \mathbb{P}_x[\hat{X} \in N_x^\gamma(\phi)]] \geqslant - \sup_{x\in K} \inf_{\Psi\in N^\gamma(\phi)} S(x, T\Psi) - g_K(N)$$

Hence

$$\begin{aligned}\mathbb{P}_x[d(\Psi, \hat{X}_N) < \eta] \geqslant \quad & \mathbb{P}_x[d(\Psi, \hat{X}_N) < \gamma] \geqslant \inf_{x\in K} \mathbb{P}_x[\hat{X} \in N^\gamma(\Psi)] \\ & \geqslant exp\left(\frac{-\sup_{x\in K} \inf_{\phi\in N^\gamma(\Psi)} S(x, T, \phi) - g_K(N)}{1/N}\right) \\ & \geqslant exp\left(\frac{-\xi_{\gamma,T}^K - S(x, T, \Psi) - g_K(N)}{1/N}\right)\end{aligned}$$

Taking N large enough such that $\xi_{\gamma,T}^K + g_K(N) < \delta$ gives us the result

■

Proposition 4.34 :
We denote by $BC(\varphi) = \overline{\{x \in \Delta\,;\, x \in \omega(x)\}}$ *the Birkhoff center of the flow* φ*. Then*

$$V_\alpha \cap BC(\varphi) \subset \mathcal{R}_\mathcal{L}$$

Proof : We know that $BC(\varphi) = \overline{Rec(\varphi)} = \overline{\{x \in \Delta\,;\, x \in \omega(x)\}}$. It's apparent that $Rec(\varphi) \subset \mathcal{R}_\mathcal{L}$. Hypothesis 4.7 allow us to conclude.

■

Corollary 4.35 :
Let μ *be an invariant measure for the flow* φ *whose support* S *lies within* V_α*. Then* $S \subset \mathcal{R}_\mathcal{L}$*.*

We arrive at the main theorem of this section, linking $\mathcal{L}$-chain recurrence with ap-chain recurrence.

Theorem 4.36 :
Under the hypotheses 4.1, 4.2 and 4.7 we have

$$\mathcal{R}_{ap} \cap V_\alpha = \mathcal{R}_\mathcal{L} \cap V_\alpha$$

$$\forall x \in \mathcal{R}_\mathcal{L} \cap V_\alpha \qquad [x]_\mathcal{L} = [x]_{ap}$$

Proof : By Proposition 4.29 we already know that, if $[x]_\mathcal{L}$ is a closed $\mathcal{L}$-basic class in $\mathring{\Delta}$, then $x \in \mathcal{R}_{ap}$ and $[x]_\mathcal{L} \subset [x]_{ap}$.

The function $B_\mathcal{L}$ is upper semi-continuous. Thus

$$\forall \gamma > 0 \quad \exists U_\gamma^i \text{ neighborhood of } K_i \text{ s.t. } \forall (a, b) \in (U_\gamma^i)^2 \quad B_\mathcal{L}(a, b) < \gamma$$

Lemma 4.37 :
Let U be a neighborhood of $\mathcal{R}_{\mathcal{L}} \cap \mathring{\Delta}$ such that $K = \overline{U}$ is a compact subset of $\mathring{\Delta}$. Thus there exists $T > 0$ such that, if $\gamma^+(x) \subset K$ then $\{\varphi_t(x)\,;\, t \in [0,T]\} \cap U \neq \emptyset$

Proof : Due to the fact that, if $\gamma_+(x) \subset K$, $\omega(x) \subset U$ we can't have $\gamma^+(x) \subset K \setminus U$. Let's define $\nu(x) = inf\{t\,;\, \varphi_t(x) \notin K \setminus U\} < +\infty$

We will show that $\nu(x)$ is upper semi-continuous.

Let $x \in K$ and $x_n \in K^{\mathbb{N}}$ such that $x_n \to x$.

The continuity of φ gives us $d(\varphi_{T_n}(x_n), \varphi_T(x)) \to 0$ with the convention that $\varphi_\infty(x) = \omega(x)$. Let $\tilde{T} = \nu(x)$, there exists a sequence $\varepsilon_m \to 0$ such that $\varphi_{\tilde{T}+\varepsilon_m}(x) \notin K \setminus U$.

Yet we have $\varphi_{\tilde{T}+\varepsilon_m}(x_n) \to \varphi_{\tilde{T}+\varepsilon_m}(x)$. Hence, for n large enough we get $\nu(x_n) \leqslant \tilde{T} + \varepsilon_m$

We now get $\limsup\limits_{n\to\infty} \nu(x_n) \leqslant \tilde{T} + \varepsilon_m$

Then $\limsup\limits_{n\to\infty} \nu(x_n) \leqslant \nu(x)$, i.e. ν is u.s.c.

Thus $T = \max\limits_{x \in K\setminus U} \nu(x)$ exists and we get, if $\gamma^+(x) \subset K$ then

$$\{\varphi_t(x)\,;\, t \in [0, T+\varepsilon]\} \cap U \neq \emptyset$$

■

Let $x \in \mathcal{R}_{ap} \cap \mathring{\Delta}$ and let $y \in [x]_{ap}$, there exists two sequences of positive real numbers $\delta_k \to 0$, $T_k \to \infty$ and a sequence of δ_k, T_k ap-pseudo-orbits linking x to y denoted Ψ^k.

Lemma 4.37 gives us $K \subset \mathring{\Delta}$ a compact neighborhood of $\mathcal{R}_{ap}$ such that, for k large enough, every δ_k, T_k ap-pseudo-orbit stays in K.

We know that $\mathcal{R}_{\mathcal{L}} \subset \mathcal{R}_{ap} \subset K$. Let γ such that $\bigcup\limits_i N^\gamma(K_i) \subset K$. Let then T given by the former lemma. For k large enough, we have $T_k > T$ and thus every continuous part of the δ_k, T_k ap-pseudo-orbit intersect U_γ.

Let $\varepsilon < \gamma$ and $V = \bigcup\limits_i N^\varepsilon(K_i)$, we assume ε to be small enough such that the ε-neighborhood of the K_i are disjoint.

We define two sequences of times $(\sigma_i(k))_{i=1\cdots p_k}$ and $(\tau_i(k))_{i=1\cdots q_k}$ by

$$\sigma_0(k) = 0 \qquad \tau_0(k) = min\{t\,;\, \Psi_t^k \notin V\}$$

$$\sigma_{i+1}(k) = min\{t > \tau_i(k)\,;\, \Psi_t^k \in V\} \quad \tau_{i+1}(k) = min\{t > \sigma_{i+1}(k)\,;\, \Psi_t^k \notin V\}$$

By Lemma 4.37 we know that $\sigma_{i+1}(k) - \tau_i(k) < 2T$.

If our pseudo-orbit ψ^k enters more than once the same $N^\varepsilon(K_i)$ then we truncate what happens between the first entry and the last exit and we will

keep the same name for the new path (which may no more be a pseudo-orbit). This way we get $q_k, p_k < \eta$.

If $y \in \bigcup_i K_i$ then $q_k = p_k + 1$, else $q_k = p_k$.

In the first case we get

$$B_{\mathcal{L}}(x,y) \leqslant \sum_{i=0}^{q_k-1} B_{\mathcal{L}}(\Psi^k_{\sigma_i(k)}, \Psi^k_{\tau_i(k)}) + \sum_{i=0}^{p_k} B_{\mathcal{L}}(\Psi^k_{\tau_i(k)}, \Psi^k_{\sigma_{i+1}(k)}) + B_{\mathcal{L}}(\Psi^k_{\tau_{q_k}}, y)$$

In the second case, we get

$$B_{\mathcal{L}}(x,y) \leqslant \sum_{i=0}^{q_k-1} B_{\mathcal{L}}(\Psi^k_{\sigma_i(k)}, \Psi^k_{\tau_i(k)}) + \sum_{i=0}^{p_k} B_{\mathcal{L}}(\Psi^k_{\tau_i(k)}, \Psi^k_{\sigma_{i+1}(k)}) + B_{\mathcal{L}}(\Psi^k_{\sigma_{p_k}}, y)$$

In either case we obtain

$$B_{\mathcal{L}}(x,y) \leqslant (1+\nu)\alpha + (1+\nu) Sup\left\{B_{\mathcal{L}}(a,b)\ ;\ d(b, \varphi_{[0,2T]}(a)) < \varepsilon_k\right\}$$

where

$$\varepsilon_k = \sup_{d(x,y)<\delta_k\ ,\ t\in[0,2T]} d(\varphi_t(x), \varphi_t(y))$$

From the continuity of φ, the positive invariance of $[x]_{ap}$ and Proposition 4.8 we infer that

$$\lim_{k\to 0} Sup\left\{B_{\mathcal{L}}(a,b)\ ;\ d(b, \varphi_{[0,2T]}(a)) < \varepsilon_k\right\} = 0$$

Hence we have, for every $\alpha > 0$, $B_{\mathcal{L}}(x,y) \leqslant (\nu+1)\alpha$.

In conclusion $x \rightsquigarrow_{\mathcal{L}} y$. Similarly $y \rightsquigarrow_{\mathcal{L}} x$, i.e. $y \in [x]_{\mathcal{L}}$

■

Remark : This proof gives us a little more, it proves that, if $y \in \mathcal{R}_{ap}$ and $x \rightsquigarrow_{ap} y$, then $x \rightsquigarrow_{\mathcal{L}} y$.

When we take α such that $\mathring{\Delta} \cap \mathcal{R}_{\mathcal{L}} = V_\alpha \cap \mathcal{R}_{\mathcal{L}}$, this proposition proves that the first set of hypotheses implies Hypothesis 4.22.

Proposition 4.38 :
Let $j \in \{k+1, \cdots, \eta\}$, i.e. such that K_j is not a ap-quasi-attractor. Then

$$\exists \lambda > 0 \quad \forall \gamma > 0 \quad \exists N_0 \quad \exists \zeta : \mathbb{N} \to \mathbb{R} \text{ with } \lim_{n\to\infty} \zeta(n) = 0 \text{ such that}$$

$$\sup_{x\in N^\lambda(K_j)} \mathbb{P}_x\left[\tau^N_{N^\lambda(K_j)} > e^{\frac{\gamma}{1/N}}\right] \leqslant \zeta(N)$$

Proof : K_j isn't a quasi-attractor, thus $\exists\lambda > 0$ such that $\overline{N^{2\lambda}(K_j)} \subset \mathring{\Delta}$ and, for all $\gamma > 0$ and all $x \in N^{\lambda}(K_j)$, there exists T^{γ} and Ψ^{γ} such that

$$\Psi^{\gamma}(0) = x, \quad \Psi^{\gamma}(T^{\gamma}) = y^{\gamma} \notin N^{2\lambda}(K_j), \quad S(x, T^{\gamma}, \Psi^{\gamma}) < \gamma$$

Let then $U = N^{2\lambda}(K_j)$, $V = \omega(\overline{U})$. Let $r > 0$ and let K be a compact subset of $\mathring{\Delta}$ such that $N^{r}(U) \subset K$.

As Ψ^{γ} starts in U and ends outside of U it must pass through $K \setminus U$. Without loss of generality we may suppose that Ψ^{γ} stays in K and $y^{\gamma} \in K \setminus U$.

Proposition 4.33 says that :

For $\lambda > 0$, $\delta > 0$, $\tilde{T} > 0$, there exists $N_0(K) > 0$ such that, for all $N > N_0$, all $T < \tilde{T}$, all $x \in K$ and all $\Psi \in \mathcal{C}_x([0,T], K)$

$$\mathbb{P}_x[d(\Psi, \hat{X}_N) < \lambda] \geqslant exp\left(-\frac{S(x,T,\Psi)+\delta}{1/N}\right)$$

Applying this to our case gives us N_0 such that, $\forall N > N_0$

$$\mathbb{P}_x[d(\Psi^{\gamma}, \hat{X}^N) < \lambda] \geqslant exp\left(-\frac{S(x,T^{\gamma},\Psi^{\gamma})+\delta}{1/N}\right) \geqslant exp\left(-\frac{\gamma+\delta}{1/N}\right)$$

Yet $\{d(\hat{X}^N, \Psi^{\gamma}) < \lambda\}$ implies that $\hat{X}^N$ leaves $N^{\lambda}(K_j)$ before the time T^{γ} Hence

$$\forall x \in K \quad \forall N \geqslant N_0(K) \quad \mathbb{P}_x[\tau_{N^{\lambda}(K_j)} > T^{\gamma}] < 1 - e^{-\frac{\gamma+\delta}{1/N}}$$

Then

$$\begin{aligned}
\mathbb{P}_x\left[\tau_{N^{\eta}(K_j)} > e^{\frac{\gamma}{1/N}}\right] &< \left(1 - e^{-\frac{\gamma+\delta}{1/N}}\right)^{\left[\frac{e^{\frac{2\gamma}{1/N}}}{T^{\gamma}}\right]} \\
&< e^{\left[\frac{e^{\frac{2\gamma}{1/N}}}{T^{\gamma}}\right] ln\left(1-e^{-\frac{\gamma+\delta}{1/N}}\right)} \\
&< e^{\left[\frac{e^{\frac{2\gamma}{1/N}}}{T^{\gamma}}\right] -e^{-\frac{\gamma+\delta}{1/N}}} \\
&< e^{-e^{-\frac{\gamma-\delta}{1/N}} \frac{1}{T^{\gamma}}}
\end{aligned}$$

Taking $\delta < \gamma$ and $\zeta(N) = e^{-e^{-\frac{\gamma-\delta}{1/N}}}$ allows us to conclude.

■

We linked ap-basic classes with $\mathcal{L}$-basic classes but, unless the partial orders on the class are similar, we might not have the same quasi-attractors for both partial orders. This proposition shows that the quasi-attractors are the same.

Proposition 4.39 :
Let $[x]$ be a basic-class. $[x]$ is a $\mathcal{L}$-quasi-attractor if and only if $[x]$ is an ap-quasi-attractor.

Proof :

Theorem 4.36 already gives us that, if $y \in \mathcal{R}_{ap}$ and $x \rightsquigarrow_{ap} y$, then $x \rightsquigarrow_{\mathcal{L}} y$. Hence, if $[x]$ is a $\mathcal{L}$-quasi-attractor, it is also an ap-quasi-attractor.

Suppose now that $[x]$ is an ap-quasi-attractor.

Let $y \in \mathcal{R}_{\mathcal{L}}$ such that $x \rightsquigarrow_{\mathcal{L}} y$. If the path linking x to y have bounded length we get $y \in \gamma^+(x)$. As $\mathcal{L}$-basic classes are positive invariant sets we get then $y \in [x]$.

Let us now suppose that the paths linking x to y have unbounded length. Propositions 4.28 and 4.8 give us a compact set containing the paths linking x to y for T large enough. Using the same technique as in the proof of the proposition 4.29 gives us $x \rightsquigarrow_{ap} y$, i.e. $y \in [x]$.

■

APPENDIX A: REMINDERS

The purpose of this appendix is to give small reminders about some basic properties of the objects used here. We won't give extensive proofs of the results or complete theory of these objects but will include some references should the reader want to know more.

A.1. Quasi-stationary distributions.

This subsection has been inspired by the notes of a course given by Sylvie Méléard at the VI Escuela de Probabilidad y Procesos Estocàsticos in Guanajuato in September 2009.

Here we will consider a Markov chain Z_t, either in discrete or continuous time, whose state space $E \subset \mathbb{R}^d$ admits an absorbing state, denoted by $\{0\}$. We will denote $E^* = E \setminus \{0\}$, $\mathcal{P}^*$ the set of probability measures whose support lies in E^*. We define T_0 to be the absorption time.

$$T_0 = \inf\{t > 0; Z_t = 0\}$$

We suppose that, whatever the initial state is, the process will almost surely be absorbed, i.e.

$$\forall z \in E \qquad \mathbb{P}_z[T_0 < \infty] = 1$$

Definition A.1:

1. A probability measure μ on E^* is said to be a quasi-stationary distribution (QSD) if and only if, for every Borelian set $A \subset E^*$ and every $t > 0$,
$$\mathbb{P}_\mu[Z_t \in A | T_0 > t] = \mu(A)$$
We remark that, in this case, μ is a fixed point for the conditional evolution
$$\nu \mapsto \mathbb{P}_\nu[Z_t \in \cdot | T_0 > t]$$
2. A quasi-limiting measure (QLD) for $\alpha \in \mathcal{P}^*$ is a probability measure ν on E^* such that, is the weak convergence sense,
$$\lim_{t\to\infty} \mathbb{P}_\alpha[Z_t \in \cdot | T_0 > t] = \nu$$
3. The Yaglom limit is the probability measure π on E^* defined by, for $A \in \mathcal{B}(E^*)$,
$$\pi(A) = \lim_{t\to\infty} \mathbb{P}_z[Z_t \in A | T_0 > t]$$
as soon as this limit exists and doesn't depend on $z \in E^*$

Proposition A.2 :
Suppose that μ is a QSD for this process Z_t. Then there exists a positive real number $\theta(\mu)$ such that

$$\mathbb{P}_\mu[T_0 > t] = e^{-\theta(\mu)t}$$

See [15] for a proof

The following theorem is true in a broader context (see e.g. Sylvie Méléard course notes or Bonsall paper [7]) but this version is sufficient for the problem at hand.

Theorem A.3 (Perron-Frobenius Theorem) **:**
Let E be a finite set and let X_n be a discrete time Markov chain on E with transition matrix Q. Up to merging some states, we suppose that X_n admits an unique absorbing state $\{0\}$. We denote by Q^* the matrix Q restricted to $E^* = E \setminus \{0\}$. We suppose Q^* to be irreducible. Then

1) There exists an unique quasi-stationary distribution (QSD) obtained as the only left eigenvector μ of Q^* verifying

$$\mu_i > 0 \quad ; \quad \sum_i \mu_i = 1$$

The corresponding eigenvalue $0 < \rho = e^{-\theta} < 1$ is such that

$$\mathbb{P}_\mu[T_0 > n] = e^{-\theta n}$$

2) The measure μ is the Yaglom limit, i.e.

$$\forall i, j \in E^* \lim_{n\to\infty} \mathbb{P}_i[X_n = j | T_0 > n] = \mu_j$$

3) For every couple (i, j) of points in E^* $\lim_{n\to\infty} e^{\theta n}\mathbb{P}_i[X_n = j] = \pi_i\mu_j$ where π is the unique right eigenvector of Q^* associated to ρ and such that $\pi_i > 0 \quad ; \quad \sum_i \mu_i\pi_i = 1$

4) For every couple (i, j) of E^* and every $n, m \in \mathbb{N}^*$

$$\lim_{n\to\infty} \frac{\mathbb{P}_i[T_0 > n + m]}{\mathbb{P}_i[T_0 > n]} = \frac{\pi_i}{\pi_j} e^{-\theta n}$$

The result concerning the Yaglom limit can be refined

Proposition A.4 :
We suppose Q to be $\mathbb{C}$ diagonalizable (this will be the case most of the time as the set of diagonalizable matrix of size n contains an open subset that is dense in $\mathcal{M}_n(\mathbb{C})$). We denote $\rho = \lambda_1 > |\lambda_2| > ... \geqslant |\lambda_n|$ the eigenvalues of Q. Let ν be a measure on E, then :

$$\|\mathbb{P}_\nu[X_k \in \cdot | T_0 > k] - \mu\|_2 \leqslant C \left(\frac{|\lambda_2|}{\rho}\right)^k$$

This result is an interpretation in our context of a classic result in numerical analysis on the power method, see e.g. [18].

For more information on the subject see [15] and [17].

A.2. Stochastic approximation algorithms.

This subsection has been inspired by [3] where you can find more information on the subject and complete proofs of the results.

Here we will consider a family of discrete processes $(X_n^N)_{n\in\mathbb{N}}$ taking values in K a compact subset of $\mathbb{R}^m$. The parameter N indexing the processes may take either real or integer values. We suppose that the X_n^N are defined on

a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and denote by $\mathcal{F}_n^N$ the σ-algebra generated by $\{X_i^N, i = 1, ..., n\}$. We suppose that

$$X_{n+1}^N - X_n^N = \frac{1}{N}(F(X_n^N) + U_{n+1}^N)$$

where

(i) $F : \mathbb{R}^m \to \mathbb{R}^m$ is a locally Lipschitz vector field
(ii) $\mathbb{E}[U_{n+1}^N | \mathcal{F}_n^N] = 0$
(iii) There exists $\Gamma \geqslant 0$ such that $\|U_n^N\| \leqslant \sqrt{\Gamma}$

F will be called the *mean field* associated to X^N. The following results may be proved under a broader set of hypotheses but these ones will be sufficient for our purposes.

We denote by $\{\varphi_t\}$ the flow induced by F. So as to compare the trajectory of $\{\varphi_t\}$ with those of (X_n^N) it's convenient to introduce the continuous in time process $\hat{X}^N : \mathbb{R} \to \mathbb{R}^m$ defined by

$$\hat{X}^N(k/N) = X_k^N \quad \forall k \in \mathbb{N}$$

and extended by linear interpolation on every $[k/N, (k+1)/N]$.

Let

$$D^N(T) = \max_{0 \leqslant t \leqslant T} \|\hat{X}^N(t) - \varphi_t(X_0^N)\|$$

be the variable measuring the distance between the trajectories $t \mapsto \hat{X}^N(t)$ and $t \mapsto \varphi_t(X_0^N)$.

Theorem A.5 :
For every $T > 0$, there exists $c > 0$ (depending only on F,Γ and T) such that, for every $\varepsilon > 0$, and for N large enough :

$$\mathbb{P}[D^N(T) \geqslant \varepsilon] \leqslant 2me^{-\varepsilon^2 cN}$$

See [3] for a proof.

We now suppose that $X^N = (X_n^N)_{n \in \mathbb{N}}$ is a Markov chain defined on a countable set K_N. We also suppose that X^N admits at least one invariant probability measure π_N.

Theorem A.6 :
The limit set of $\{\pi_N\}$ contains only probability measures that are invariant for the flow φ_t.

See [3] for a proof

For more information on the subject see e.g. [3] and [2].

ACKNOWLEDGEMENTS

The author wishes to thank Michel Benaïm and Mathieu Faure for their advices and proofreading on this work.

REFERENCES

[1] M. Benaïm. Recursive algorithms, urn processes, and the chaining number of chain recurrent sets. Ergodic Theory and Dynamical Systems, 18:53–87, 1998.
[2] M. Benaïm. Dynamics of stochastic approximation algorithms. In Le Séminaire de Probabilites, Lecture Notes in Math., To appear. Available at http://www-sv.cict.fr/lsp/Benaim/survey.html.
[3] M. Benaïm and J.W. Weibull. Deterministic approximation of stochastic evolution in games. Econometrica, 71:873–903, 2003.
[4] Ken Binmore, John Gale, and Larry Samuelson. Learning to be imperfect: The ultimatum game. Games and Economic Behavior, 8:56–90, 1995.
[5] Ken Binmore and Larry Samuelson. Muddling through: Noisy equilibrium selection. Journal of Economic Theory, 74:235–265, 1997.
[6] Jonas Björnerstedt and Jörgen W. Weibull. Nash equilibrium and evolution by imitation. In Kenneth J. Arrow et al., editors, The Rational Foundations of Economic Behavior, pages 155–181. St. Martin's Press, New York, 1996.
[7] F. F. Bonsall. Linear operators in complete positive cones. Proc. London Math. Soc., 1958.
[8] C.C. Conley. Isolated Invariant Sets and the Morse Index. American Mathematical Society, 1978.
[9] P. Dupuis and R. S. Ellis. A Weak Convergence Approach to the Theory of Large Deviations. Wiley, New York, NY, 1997.
[10] M. Faure and S. J. Schreiber. Quasi-stationary distributions for randomly perturbed dynamical systems. arXiv:1101.3420v1, Submitted, 2011.
[11] J. Hofbauer and K. Sigmund. Evolutionary Dynamics and Population Dynamics, 1998.
[12] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. Lond. A, 115:700–721, 1927.
[13] Y. Kifer. Random perturbations of dynamical systems. 1988.
[14] A. J. Lotka. Elements of Physical Biology. Williams and Witkins, Baltimore, 1925.
[15] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. arXiv:1112.4732v1, Submitted, 2011.
[16] A. J. Nicholson and V. A. Bailey. The balance of animal populations. Proc. Zool. Soc. Lond., pages 551–598, 1935.
[17] P. K. Pollett. Quasi-stationary distributions: A bibliography. 2012.
[18] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. Numerical mathematics, volume 37 of Texts in Applied Mathematics. Springer-Verlag, New York, 2000.
[19] William H. Sandholm. Population games and evolutionary dynamics. Economic Learning and Social Evolution. MIT Press, Cambridge, MA, 2010.
[20] S. J. Schreiber. Persistence despite perturbations for interacting populations. Journal of Theoretical Biology, 242:844–52, 2006.
[21] W. R. Thompson. La theorie mathematique de l'action des parasites entomophages et le facteur du hasard. Annales Faculte des Sciences de Marseille, 2:69–89, 1924.
[22] V. Volterra. Fluctuations in the abundance of a species considered mathematically. Nature, 118:558–560, 1926.

Institut de Mathématiques, Université de Neuchâtel,
Rue Emile-Argand 11. Neuchâtel. Switzerland.
E-mail: bastien.marmet@unine.ch