Quasi stationary distributions.

P. Collet

Centre de Physique Théorique

CNRS UMR 7644

Ecole Polytechnique

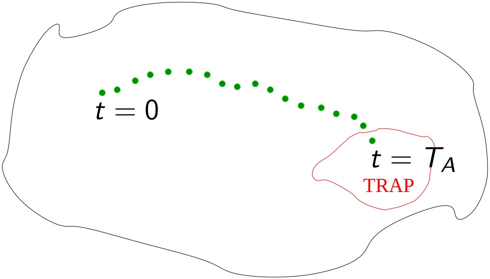
F-91128 Palaiseau Cedex (France)

e-mail: collet@cpht.polytechnique.fr

Content of Lecture 1.

- The setting.
- A simple example of dynamical system.
- A simple example of Markov chain.
- General definitions.
- General results.

We are interested in situations where in the phase space there is a subset which is absorbing for the (deterministic or stochastic) dynamics. This subset can be viewed as a trap, a hole, a cemetery etc. In other words, we will not consider the dynamics beyond the first time the state reaches the absorbing subset (the game is over). Let A be a subset of $\mathcal{X}$ which is imagined as a trap, namely if a trajectory arrives in A, it disappears (killed, stick etc).



For example in population dynamics, the number n of individuals of a specie is an integer, the phase space is $\mathcal{X} = \mathbb{N}^* = \{0, 1, 2, \ldots\}$. Individuals can die, reproduce, for example in a birth and death process. The number of individuals vary with time $n(t)$. However if there is no spontaneous generation, the state $n = 0$ is a trap, the specie disappears at this state. If the system has reached that state it stays there forever.

There are several natural questions associated to this setting.

Question I

Given an initial distribution μ on $\mathcal{X}$, what is the probability that a trajectory has survived up to time $t > 0$? (For example if μ is the Dirac measure on one point).

In other words, if we denote by T_A the first time the system enters in A (a function of the initial condition and of the randomness of the evolution), what is the behaviour of

$$\mathbb{P}_\mu\big(T_A > t\big) \ .$$

Often one can say something only for large t.

Question II

Assume a trajectory initially distributed with μ has survived up to time $t > 0$, what is the distribution of the state at time t?

In other words, can we say something about

$$\mathbb{P}_\mu\big(X_t \in B \big| T_A > t\big) = \frac{\mathbb{P}_\mu\big(X_t \in B\,,\ T_A > t\big)}{\mathbb{P}_\mu\big(T_A > t\big)}\ ,$$

B a measurable subset of $\mathcal{X}$.

Question III

Are there trajectories which never reach the trap A? ($T_A = \infty$).

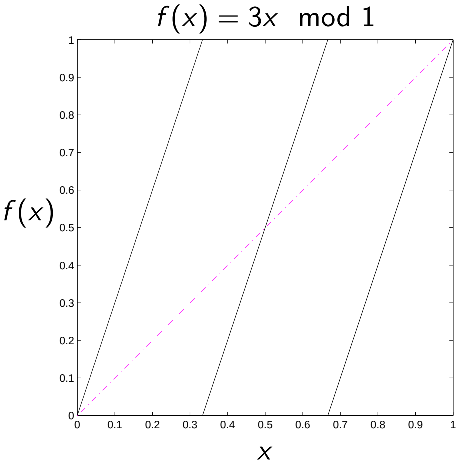
If so, how are they distributed?

How is this related to Question II?

Question II (distribution of survivors at large time) deals with trajectories which have survived for a large time but as we will see later, most of them are on the verge of falling in the trap. This will be the main subject of these lectures.
Question III deals with trajectories that will never see the trap, this is very different.
Question II (distribution of survivors at large time) and Question III (eternal life) have in general very different answers. This can be seen clearly in the case of dynamical systems (deterministic time evolution).

A simple example from dynamical systems.

Consider the map f of the unit interval $\mathcal{X} = [0, 1]$ given by $f(x) = 3x \pmod 1$.



It is easy to verify that the Lebesgue measure Leb is invariant, namely for any Borel set B

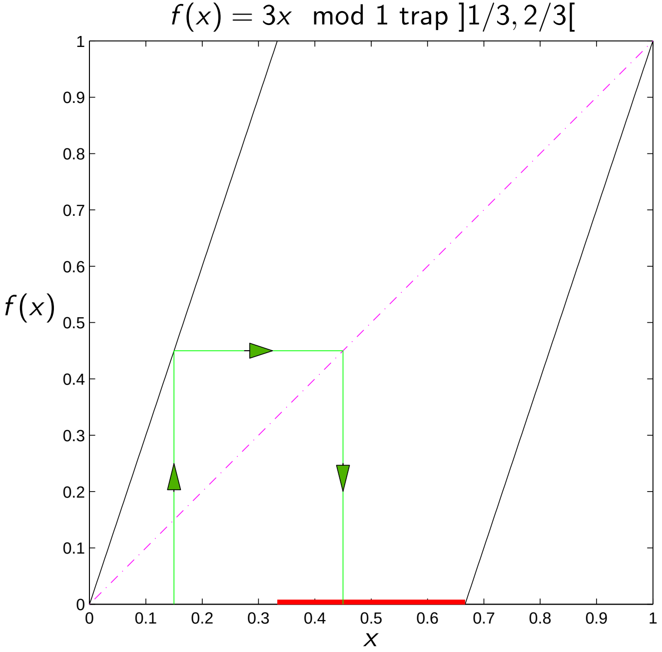
$$\mathrm{Leb}\big(f^{-1}(B)\big) = \mathrm{Leb}\big(B\big) \ ,$$

where $f^{-1}(B)$ is the preimage set of B, namely

$$f^{-1}(B) = \big\{x \in [0,1] \,\big|\, f(x) \in B\big\} \ .$$

Given an initial distribution μ_0 on $[0, 1]$, and a map f of $[0, 1]$, we can define a discrete time stochastic process on $[0, 1]$ as follows.
The probability space is $[0, 1]$ and the process is defined recursively by $X_{n+1} = f(X_n)$, X_0 being distributed according to μ_0.
The time evolution is deterministic, but the initial condition is chosen at random. This randomness is propagated (here "amplified") by the time evolution.
We have a measure on the set of trajectories, this is a stochastic process.

We use the interval $A =]1/3, 2/3[$ as a trap.

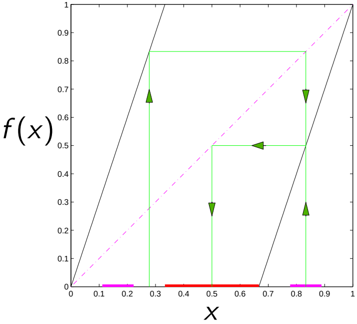


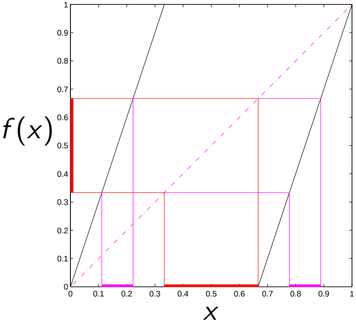
Consider the sequence of sets

$$\mathscr{S}_n = \{\, T_A > n \} \ .$$

It is easy to check that $\mathscr{S}_0 = [0, 1/3] \cup [1/3, 1]$,

$$\mathscr{S}_1 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$$





Consider the sequence of sets

$$\mathscr{S}_n = \{ T_A > n \} \ .$$

It is easy to check that $\mathscr{S}_0 = [0, 1/3] \cup [1/3, 1]$,

$$\mathscr{S}_1 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$$

and so on. Therefore

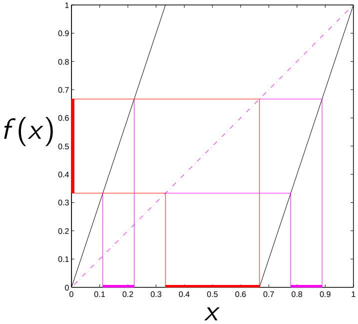
$$\text{Lebesgue}\big(\mathscr{S}_n\big) = \left(\frac{2}{3}\right)^{n+1} \ .$$

Namely we have the answer to question I: probability of surviving up to time n starting from the Lebesgue measure. This probability follows an exponential law.

The set of initial conditions which never die is

$$K = \bigcap_n \mathscr{S}_n \; .$$

This is the Cantor set.



The set of initial conditions which never die is

$$K = \bigcap_n \mathscr{S}_n \ .$$

This is the Cantor set.

K is of zero Lebesgue measure.

Moreover, during the recursive construction, if we start with the Lebesgue measure, all the intervals of $\mathscr{S}_n$ have the same length and hence the same weight. We get at the end the Cantor measure which is very singular (absolutely singular with respect to the Lebesgue measure).

Assume the initial conditions are distributed according to the Lebesgue measure on $A^c = [0, 1/3] \cup [1/3, 1]$. We can compute the distribution of the trajectories which have survived up to time one. For any $B \subset A^c$

$$\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in B \,\big|\, T_A > 1\big) = \frac{\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in B\,,\ T_A > 1\big)}{\mathbb{P}_{\mathrm{Leb}}\big(T_A > 1\big)} = \frac{\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in B\big)}{\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in A^c\big)}$$

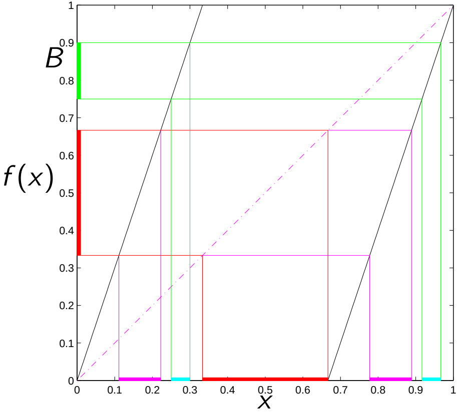
$$= \frac{\mathrm{Leb}\big(\big\{x \,\big|\, f(x) \in B\big\}\big)}{\mathrm{Leb}\big(\big\{x \,\big|\, f(x) \in A^c\big\}\big)} = \frac{\mathrm{Leb}\big(f^{-1}(B)\big)}{\mathrm{Leb}\big(f^{-1}(A^c)\big)}\ .$$

Using Thales theorem it is easy to verify that this is equal to $\mathrm{Leb}(B)/\mathrm{Leb}(A^c)$ when B is a finite union of intervals. The general case of a measurable B follows.

The graphical "proof"

$$\frac{\mathrm{Leb}\big(f^{-1}(B)\big)}{\mathrm{Leb}\big(f^{-1}(A^c)\big)} = \frac{\mathrm{Leb}(B)}{\mathrm{Leb}(A^c)}$$

$f(x) = 3x \mod 1$ trap $]1/3, 2/3[$



Assume the initial conditions are distributed according to the Lebesgue measure on $A^c = [0, 1/3] \cup [1/3, 1]$. We can compute the distribution of the trajectories which have survived up to time one. For any $B \subset A^c$

$$\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in B \,\big|\, T_A > 1\big) = \frac{\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in B\,,\ T_A > 1\big)}{\mathbb{P}_{\mathrm{Leb}}\big(T_A > 1\big)} = \frac{\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in B\big)}{\mathbb{P}_{\mathrm{Leb}}\big(X_1 \in A^c\big)}$$

$$= \frac{\mathrm{Leb}\big(\big\{x \,\big|\, f(x) \in B\big\}\big)}{\mathrm{Leb}\big(\big\{x \,\big|\, f(x) \in A^c\big\}\big)} = \frac{\mathrm{Leb}\big(f^{-1}(B)\big)}{\mathrm{Leb}\big(f^{-1}(A^c)\big)} \,.$$

Using Thales theorem it is easy to verify that this is equal to $\mathrm{Leb}(B)/\mathrm{Leb}(A^c)$ when B is a finite union of intervals. The general case of a measurable B follows.

By induction and using the Markov property, it follows that for any n

$$\mathbb{P}_{\mathrm{Leb}}\big(X_n \in B \,\big|\, T_A > n\big) = \mathrm{Leb}(B) \,.$$

We have solved Question II (distribution of survivors at large time) for the case of the (normalised) Lebesgue measure as initial condition: the distribution stays Lebesgue.
Question III (survive forever) deals with trajectories that will never see the trap. In the case of dynamical systems, these trajectories concentrate on a very small set (a Cantor set) which is invariant and disjoint from the trap.
Question II (distribution of survivors at large time) and Question III (eternal life) have very different answers.

Finite states Markov chains.

This is another simple interesting example. Consider the finite phase space $\mathcal{X} = \{0, 1, \ldots, k\}$, and let the trap be the state 0 ($A = \{0\}$). Let $(p_{i,j})$ be a Markov matrix, $\mathbb{P}(X_1 = j \big| X_0 = i) = p_{i,j}$. For simplicity we will assume that

$$\inf_{i,j} p_{i,j} > 0 \; .$$

We will denote by $\tilde{p}$ the sub-Markovian $k \times k$ matrix with entries $p_{i,j}$, $i, j \in \{1, \ldots, k\}$.

By the Perron-Frobenius theorem, there is a unique eigenvalue λ of $\tilde{p}$ with largest modulus. It is positive, smaller than one and a simple eigenvalue. There is an eigenvector h with all entries positive and an eigenvector α of the adjoint (a linear form) with the same property. We can assume $\alpha(h) = 1$, and $\sum_{i=1}^{k} \alpha(i) = 1$.

It is easy to verify that for $0 \leq m \leq n$, $x_0, x \in \{1, \ldots, k\}$

$$\mathbb{P}_{x_0}\big(X_m = x\,,\; T_0 > n\big) = \sum_{y \in \{1,\ldots,k\}} \tilde{p}^{m}_{x_0,x} \tilde{p}^{n-m}_{x,y}\ .$$

Therefore from spectral theory $(p^n_{x_0,x} = \lambda^n\, h(x_0)\, \alpha(x) + o\big(\lambda^n\big))$

$$\mathbb{P}_{x_0}\big(T_0 > n\big) = \lambda^n\ h(x_0) + o\big(\lambda^n\big)\ .$$

This is the answer to Question I: the survival probability decays asymptotically exponentially fast. Moreover

$$\lim_{n\to\infty} \mathbb{P}_{x_0}\big(X_n = x \,\big|\, T_0 > n\big) = \alpha(x)\ .$$

This is the answer to Question II: the trajectories which have survived up to time n (large) are asymptotically distributed accordingly to the probability α.

For a fixed $m \geq 0$

$$\lim_{n\to\infty} \mathbb{P}_{x_0}\big(X_m = x \,\big|\, T_0 > n\big) = \lambda^{-m} \, \frac{h(x)}{h(x_0)} \, \tilde{p}^m(x_0, x)$$

This is the answer to Question III: the trajectories which survive forever evolve according to a so called Q-process which is the h-transform of p. This is a Markov process with transition probabilities

$$\lambda^{-1} \, \frac{h(x)}{h(x_0)} \, \tilde{p}(x_0, x)$$

In particular, the invariant measure of the Q-process is the measure μ with $\mu(x) = \alpha(x)\, h(x)$, which in general is different from the measure α.

We see the same phenomenon as in the case of dynamical systems although less spectacular here.

One says that a probability measure ν is the Yaglom limit if for any initial point x, and any Borel set B

$$\lim_{t\to\infty} \frac{\mathbb{P}_x\big(X_t \in B\,,\ T_A > t\big)}{\mathbb{P}_x\big(T_A > t\big)} = \nu(B)\ .$$

A probability ν is a quasi limiting distribution (q.l.d.) if there is a probability π such that in distribution

$$\lim_{t\to\infty} \mathbb{P}_\pi\big(X_t \in \bullet \,\big|\, T_A > t\big) = \nu(\bullet)\ .$$

We say that ν (a probability on $\mathcal{X}$) is a quasi stationary distribution (q.s.d.) if for any $t \geq 0$

$$\mathbb{P}_\nu\big(X_t \in B \big| T_A > t\big) = \nu(B)$$

for any measurable set B. If there is no trap, this is a stationary measure.

It is easy to prove using the Markov property that any q.l.d. is a q.s.d. Any Yaglom limit is a q.l.d. $(\pi = \delta_x)$, hence a q.s.d.

There are few general results about q.s.d.

A first result concerns the entrance time T_A in the trap A for a Markov process (X_t).

Theorem 1

Let ν be a q.s.d. distribution for (X_t), then T_A has an exponential law, namely there is a number $\theta > 0$ such that

$$\mathbb{P}_\nu\big(T_A > t\big) = e^{-\theta t} \ .$$

This result follows from the Markov property and the definition of q.s.d. Note that this implies that for the q.s.d. most trajectories which have survived up to time t will die very soon.

The case $\theta = 0$ corresponds to invariant measure, and in the sequel we will only consider $\theta > 0$.

Proof of the exponential law.

One can verify by direct computations that this result holds in the two previous examples. In the discrete time case we have

$$\mathbb{P}_\nu\big(T_A > n\big) = \mathbb{P}_\nu\big(X_0 \notin A, \ldots, X_n \notin A\big)$$

$$= \mathbb{P}_\nu\big(X_n \notin A \big| X_0 \notin A, \ldots, X_{n-1} \notin A\big) \;\; \mathbb{P}_\nu\big(T_A > n-1\big)$$

and by the Markov property and stationarity

$$= \mathbb{P}_\nu\big(X_n \notin A \big| X_{n-1} \notin A\big) \;\; \mathbb{P}_\nu\big(T_A > n-1\big)$$

$$= \mathbb{P}_\nu\big(X_1 \notin A \big| X_0 \notin A\big) \;\; \mathbb{P}_\nu\big(T_A > n-1\big)$$

$$= \mathbb{P}_\nu\big(T_A > 1\big) \;\; \mathbb{P}_\nu\big(T_A > n-1\big)$$

and the result follows by iteration.

The problem of existence of q.s.d. is in a sense similar to the problem of existence of invariant measures.

There is however a supplementary difficulty: the rate of decay θ is unknown ($\theta = 0$ for invariant measures).

Up to now, no general existence theorem of q.s.d. is known, for example there is nothing similar to the Krylov-Bogoliubov theorem for stationary(invariant) measures.

We will see later on some particular results and techniques.

It is easy to construct examples where there is no q.s.d..

For example, if $\mathbb{P}(X_1 \in A \big| X_0 \in \mathcal{X} \backslash A) = 1$, there cannot be any q.s.d..

Indeed, here we have $T_A = 1$ (almost surely!), but we have seen before that if there is a q.s.d., this random variable should be exponential.

Another less trivial example is given by the Brownian motion (W_t) in dimension one with the trap A equal to the negative real line. There is no Yaglom limit, no q.s.d..

One can show easily that (for $x > 0$)

$$\mathbb{P}_x\big(W_t \in [y, y+dy]\,,\; T_0 > t\big) = \frac{1}{\sqrt{2\pi t}}\left(e^{-(x-y)^2/2t} - e^{-(x+y)^2/2t}\right) dy\;,$$

hence

$$\mathbb{P}_x\big(T > t\big) = x\,\sqrt{\frac{2}{\pi t}} + \mathcal{O}\left(\frac{1}{t^{3/2}}\right)\;,$$

no exponential decay. But there is renormalised Yaglom limit. It follows by a short computation that

$$\lim_{t\to\infty} \mathbb{P}_x\big(W_t/\sqrt{t} \in [y, y+dy] \,\big|\, T_0 > t\big) = ye^{-y^2/2}dy\;.$$

The strategy to survive is to escape to infinity.

Some existence theorems in particular cases.

For continuous time Markov processes on countable phase space ($\{0, 1, 2, \cdots\}$, with trap 0), Ferrari Kesten Martinez and Picco have proved the following result.

Theorem 2

Assume the process restricted to $\{1, 2, \cdots\}$ is irreducible, $\lim_{x\to\infty} \mathbb{P}_x(T < t) = 0$ for any $t > 0$ and $\mathbb{P}_x(T < \infty) = 1$ for one (and hence for all) $x \in \{1, 2, \cdots\}$. Then a necessary and sufficient condition for the existence of a q.s.d. is

$$\mathbb{E}_x\left(e^{\lambda T}\right) < \infty$$

for some $\lambda > 0$ and for one (hence for all) $x \in \{1, 2, \cdots\}$.

In a study of a birth and death process with mutation of the phenotype, with S.Méléard, S.Martínez and J.San Martín, we came to the following result.

Theorem 3

Let $\mathscr{X}$ be a polish space. Let S be a bounded positive linear operator on $C_b^0(\mathscr{X})$ the Banach space of bounded continuous functions on $\mathscr{X}$ satisfying

$$S1 > c > 0\ .$$

Assume there exists a continuous function $\varphi \geq 1$ such that for any $u > 1$, $\varphi^{-1}([1,u])$ is compact and there exists $D > 0$ and $\gamma \in]0, c[$ such that for any $\psi \in C_b^0(\mathscr{X})$ with $0 \leq \psi \leq \varphi$ we have

$$S\psi \leq \gamma\,\varphi + D\ .$$

Then there exists a probability measure ν on $\mathscr{X}$ satisfying $S^\dagger \nu = \beta\nu$ with $\beta = \nu(S1) > 0$, (and $\nu(\varphi) < \infty$).

For continuous time, this result is used to prove the existence of a q.s.d. by considering $S = P_1$ where P_t is the semi-group associated to the (killed) Makov process.

Example: birth and death processes. Let (a_n) and (b_n) be two sequences of strictly positive numbers, and consider the birth and death process $N(t)$ on $\mathbb{N}$ given by

$$\mathbb{P}\big(N(t+dt) = n+1 \,\big|\, N(t) = n\big) = n\, a_n\, dt \;,$$

$$\mathbb{P}\big(N(t+dt) = n-1 \,\big|\, N(t) = n\big) = n\, b_n\, dt \;,$$

$$\mathbb{P}\big(|N(t+dt) - n| > 1 \,\big|\, N(t) = n\big) = 0 \;.$$

We assume no spontaneous generation, namely $a_0 = b_0 = 0$.

Theorem 4

Assume that

$$\limsup_{n\to\infty} \frac{a_n}{b_n} = \beta < 1 \;, \quad \text{and} \quad \alpha = \liminf_{n\to\infty} b_n > 0 \;.$$

Then $(N(t)$ *has a q.s.d. (for the trap* $A = \{0\}$*).*

There are many ways to prove this theorem and one can get more information on the q.s.d.(s) (see for example the recent review by Van Doorn and Pollett). I will just illustrate how to use our abstract result in this case. The proof is almost the same when individuals carry phenotypic traits and birth can lead to mutations.

Preliminary observation: there exists $\epsilon > 0$ and $A > 0$ such that for any $t \in [0, 1]$

$$\epsilon + \alpha \left[\beta \left(e^{A+\epsilon t} - 1 \right) - \left(1 - e^{-A-\epsilon t} \right) \right] < 0 \, .$$

Let $f(t, n)$ be the function

$$f(t, n) = e^{n\,(A+\epsilon\, t)} \, .$$

It is easy to verify that there exists $C > 0$ such that for any $t \in [0, 1]$ and any n

$$\partial_t f(t, n) + Lf(t, n) < C \, ,$$

where L is the generator, namely

$$Lg(n) = n\, a_n \left(g(n+1) - g(n) \right) + n\, b_n \left(g(n-1) - g(n) \right) \, .$$

We now apply the Martingale representation

$$f(t,N(t)) = f(0,N(0))$$

$$+\int_0^t \big(\partial_s f(s,N(s)) + Lf(s,N(s))\big)\,ds + Martingale(t)\ .$$

Using the above estimate we get

$$\mathbb{E}_n\left(f(1,N(1))\right) \le C + f(0,n)\ .$$

This implies

$$\mathbb{E}_n\left(e^{(A+\epsilon)\,N(1)}\right) \le e^{A\,n} + C = e^{-\epsilon\,n}e^{(A+\epsilon)\,n} + C$$

$$\le \delta e^{(A+\epsilon)\,n} + C + \delta^{-1-A/\epsilon}$$

for any $\delta \in (0,1)$ and any $n \in \mathbb{N}$.

Let (P_t) be the semi-group associated to the process $N(t)$, since $L1 \geq -b_1$, we have

$$P_1 1 \geq e^{-b_1} \ .$$

If we denote by φ the function $\varphi(n) = \exp(n\,(A+\epsilon))$ we have for any $\delta \in (0,1)$

$$P_1\varphi \leq \delta\,\varphi + C + \delta^{-1-A/\epsilon} \ .$$

We now choose $\delta > 0$ such that $\delta\,\exp(-b_1) < 1$ and our previous result implies the existence of a q.s.d.

Remark. In the previous estimates I assumed the process $N(t)$ has exponential moments which is not known a priori. One can use a sequence of bounded functions to avoid this problem (see the formulation of the Theorem).

A short bibliography

E. Van Doorn, P.Pollett. Quasi-stationary distributions. Memorandum No. 1945, Department of Applied Mathematics, University of Twente (2011).

P.Collet, S.Martínez, J. San Martín. Quasi-Stationary Distributions, Stochastic Processes, Dynamical Systems, Applications. Prepared (2011).

P.Pollett. Quasi-stationary Distributions: A Bibliography. `www.maths.uq.edu.au/ pkp/papers/qsds/qsds.pdf`

P.Ferrari, H.Kesten, S.Martínez, P.Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. Ann. Probab. **23**, 501-521 (1995).

P.Collet, S.Méléard, S.Martínez and J.San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. Probab. Th. and Rel. fields. **151**, 191-231 (2011).

Q.S.D. for diffusions.

Content

- The setting.
- Bounded domain.
- Half line.
- The Ornstein Uhlenbeck process.
- Down from infinity.

We will consider diffusions which are solutions of a stochastic differential equation

$$dX = \alpha(X)\, dt + dW_t$$

where $X \in \mathbb{R}^n$, α is a regular vector field on $\mathbb{R}^n$. We will assume that α satisfies the standard hypothesis so that the process (X_t) is well defined for all (non negative) times. Many results below can be extended to processes satisfying $dX = \alpha(X)\, dt + \sigma(X)\, dW_t$. In order to simplify the exposition and for lack of time we will leave to the interested reader this more general case.

We will only consider two cases: the case of a bounded domain B in $\mathbb{R}^n$ with regular boundary (the trap is $A = B^c$), and the case $n = 1$ ($X_t \in \mathbb{R}$) with trap $A = (-\infty, 0]$.

Besides these cases, very few other situations have been studied (Cattiaux-Méléard, Villemonais).

Bounded domain.

Let B be a bounded (nonempty) open connected domain in $\mathbb{R}^n$ with regular boundary. The trap is $A = B^c$. The problem of q.s.d. and Q-process in this case was studied by Pinsky.

The process $(X_t)_{t \in \mathbb{R}^+}$ is well defined for $X_0 \in B$, and we will study its conditioning to the event $T > t$ where T is the first time the process hits the boundary of B.

Recall that the C_0 semi-group (P_t) is defined on $C^0(\bar{B})$ by

$$P_t f(x) = \mathbb{E}_x\big(f(X_t)\,\mathbb{1}_{T>t}\big)\ .$$

Fix $t > 0$, then P_t is a continuous linear map of $C^0(\bar{B})$. Moreover the cone of non-negative functions in $C^0(\bar{B})$ has nonempty interior. Therefore by a theorem of Krein, there exists a probability measure ν and a number $\lambda \geq 0$ such that

$$P_1^{\dagger}\nu = \lambda\,\nu\ .$$

By simple arguments it follows that there exists a q.s.d. (if $\lambda > 0$). This abstract argument does not provide precise information, although using that a q.s.d. is an eigenvector of the generator of the semi group we obtain that it is absolutely continuous with a regular density.

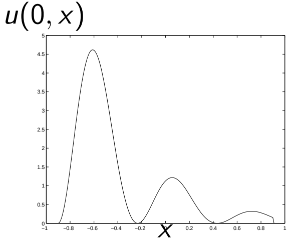
We will prove by other means (spectral theory)

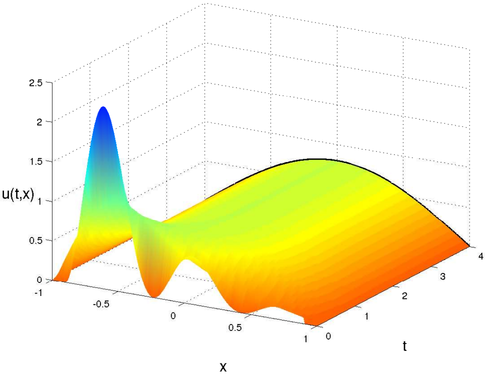
Theorem 5

There is a unique q.s.d. and it is absolutely continuous with respect to the Lebesgue measure.

Many more results follow from this approach: speed of convergence, central limit theorem etc.

Consider for example the ordinary diffusion in the interval $[-1, 1]$ killed at the boundary. Let $u(t, x)$ denote the density of $\mathbb{P}_{u(0,\cdot)}(X_t \in [x, x + dx])$

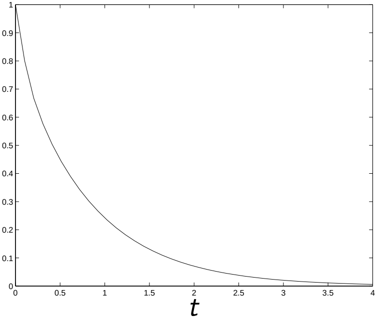




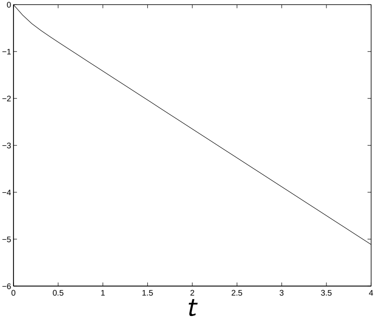
Left initial condition, right space time solution.

One can look at the probability of survival as a function of time, and to its logarithm to check the exponential decay

$\mathbb{P}(T > t)$



$\log \mathbb{P}(T > t)$



We will use spectral theory of the associated semi-group P_t to study q.s.d. and related properties.

Although the spectral theory of (P_t) can be studied in $C^0(\bar{B})$, it is slightly more convenient to work in $L^2(B)$. The relation is given by the following Lemma.

Theorem 6

For any $t > 0$, P_t is a continuous map from $L^2(B)$ to $C^0(\bar{B})$. (P_t) extends to a C_0 semi-group in $L^2(B)$.

Proof of the first part.

Let $\tilde{\alpha}$ be a smooth vector field with compact support, equal to α in $\bar{B}$, and $(\tilde{X}_t)$ the process satisfying $d\tilde{X} = \tilde{\alpha}(\tilde{X})dt + dW_t$. This process coincides with (X_t) as long as (X_t) has not left B. For $f \geq 0$ in $C^0(\bar{B})$, let $\tilde{f} \geq 0$ be a continuous function with compact support coinciding with f on $\bar{B}$ and satisfying $\|\tilde{f}\|_{L^2(\mathbb{R}^n)} \leq 2\|f\|_{L^2(B)}$. We have

$$P_t f(x) = \mathbb{E}_x\big(f(X_t)\,\mathbb{1}_{T>t}(X_.)\big) = \mathbb{E}_x\big(\tilde{f}(\tilde{X}_t)\,\mathbb{1}_{T>t}(\tilde{X}_.)\big) \leq \mathbb{E}_x\big(\tilde{f}(\tilde{X}_t)\big)$$

$$\overset{\text{Girsanov}}{=} \mathbb{E}_x\left(e^{\int_0^t \langle \tilde{\alpha}(w_s), dW_s\rangle - 1/2\int_0^t \tilde{\alpha}^2(s)ds}\tilde{f}(W_t)\right) \overset{Schwarz}{\leq}$$

$$\mathbb{E}_x\left(e^{2\int_0^t \langle \tilde{\alpha}(w_s), dW_s\rangle - (1/2)\int_0^t (2\tilde{\alpha})^2(s)ds}\right)^{1/2} \mathbb{E}_x\left(e^{\int_0^t \tilde{\alpha}^2(s)ds}\tilde{f}^2(W_t)\right)^{1/2}$$

$$\overset{\text{exp. martingale}}{\leq} C^{\text{te}}\left(\frac{1}{(2\pi t)^{d/2}}\int_{\mathbb{R}^n} e^{-(x-y)^2/(2t)}\tilde{f}^2(y)\,dy\right)^{1/2} \leq C^{\text{te}}\|f\|_{L^2(B)}$$

and the proof of the first part follows by density of $C^0(B)$ in $L^2(B)$.

Proof of the second part.

Note that in the above estimate, the constant is uniform on compact sets in t.

Since the domain B is bounded, we have from the previous inequality

$$\|P_t f\|_{L^2(B)} \leq C^{\text{te}} \big\|f\big\|_{L^2(B)} ,$$

with a constant uniform on compact sets in t. In other words, the semi-group (P_t) is uniformly bounded in $L^2(B)$.

The C_0 property in $L^2(B)$ follows from the density of $C^0(B)$ in $L^2(B)$ and a $3\,\epsilon$ argument.

Compactness in $L^2(B)$.

Theorem 7

The semi-group (P_t) is compact in $L^2(B)$.

There are several ways to prove this result.

One can apply for example Exercise VI.9.56 in Dunford-Schwartz volume 1.

One can also use Gaussian bounds on the kernel to prove that the semi-group (P_t) is Hilbert Schmidt.

We will see later on that the spectral radius is strictly positive. Hence the peripheral spectrum of any (P_t), is for $t > 0$ composed of a finite number of points which are finite dimensional eigenvalues. The rest of the spectrum lies in a closed disk of smaller radius. The following result will be used several times in the sequel. Let $B_r(y)$ denote the ball centered in y with radius r.

Lemma 8

For any $r > 0$, any $t > 0$, and any $x, y \in B$,

$$P_t \mathbb{1}_{B_r(y)}(x) > 0 \,.$$

It is enough to prove the result for r small enough. Assume assume first $x \in B_{r/2}(y)$. Then $P_t \mathbb{1}_{B_r(y)}(x) =$

$$\mathbb{E}_x\big(\mathbb{1}_{B_r(y)}(X_t)\,\mathbb{1}_{T>t}\big) \geq \mathbb{E}_x\big(\mathbb{1}_{B_r(y)}(X_t)\,\mathbb{1}_{T_{B_r(y)}>t}\big) = \mathbb{E}_x\big(\mathbb{1}_{T_{B_r(y)}>t}(X_.)\big)$$

$$\overset{\text{Girsanov}}{=} \mathbb{E}_x\left(e^{\int_0^t \langle \alpha(w_s), dW_s\rangle - 1/2\int_0^t \alpha^2(s)ds}\mathbb{1}_{T_{B_r(y)}>t}(W_.)\right) \overset{Jensen}{\geq}$$

$$C^{\text{te}}\,\mathbb{E}_x\left(\mathbb{1}_{T_{B_r(y)}>t}(W_.)\right) e^{\mathbb{E}_x\left(\mathbb{1}_{T_{B_r(y)}>t}(W_.)\,\int_0^t \langle \alpha(w_s), dW_s\rangle\right)/\mathbb{E}_x\left(\mathbb{1}_{T_{B_r(y)}>t}(W_.)\right)}$$

For a Brownian motion conditioned to stay in a ball there are explicit upper and lower bounds for $\mathbb{E}_x\left(\mathbb{1}_{T_{B_r(y)}>t}(W_.)\right)$. From the Schwarz inequality we have

$$\mathbb{E}_x\left(\mathbb{1}_{T_{B_r(y)}>t}(W_.)\,\int_0^t \langle \alpha(w_s), dW_s\rangle\right)$$

$$\geq -\mathbb{E}_x\left(\mathbb{1}_{T_{B_r(y)}>t}(W_.)\right)^{1/2}\,\mathbb{E}_x\left(\left(\int_0^t \langle \alpha(w_s), dW_s\rangle\right)^2\right)^{1/2}$$

and the result follows by standard estimates.

For x and y in B in general positions one can use a finite chain of overlapping balls contained in B, "joining" x to y (recall that we assumed B connected with regular boundary).

The following result is an important consequence.

Proposition 9

There is a function f_0 in $C^0(\bar{B})$ which is positive almost everywhere and a number $\lambda_0 > 0$ such that

$$P_t f_0 = e^{-\lambda_0 t} f_0 \ .$$

This result is used to prove uniqueness of the q.s.d.

We now sketch the proof.

We start by considering the operator P_1 which is compact and positivity preserving. There exists a (finite) set of $k \geq 1$ (different) eigenvalues $\rho_1, \ldots, \rho_k$, with equal modulus $\eta > 0$, and finite dimensional (bounded) projections $\pi_1, \ldots, \pi_k$ (in general not self-adjoint) in $L^2(B)$ such that for any integer n and for some $0 < \zeta < \eta$

$$P_1^n = \sum_{j=1}^{k} \rho_j^n \pi_j + \mathcal{O}(\zeta^n) \ .$$

We choose for ρ_1 the (an) eigenvalue nearest to η.

It easily follows that $\eta > 0$ from the above result $P_t \mathbb{1}_{B_r(y)}(x) > 0$.

The sequence

$$\frac{1}{N}\sum_{n=0}^{N}\eta^{-n}P_1^n 1$$

is non-negative and converges in $L^2(B)$ (use the spectral decomposition). If $\rho_1 \neq \eta$, this sequence converges to zero. If e_1 is an eigenvector of eigenvalue ρ_1 we have

$$|e_1| = \left|\frac{1}{N}\sum_{n=0}^{N}\rho_1^{-n}P_1^n e_1\right| \leq \frac{1}{N}\sum_{n=0}^{N}\eta^{-n}P_1^n|e_1|$$

$$\leq \|e_1\|_{C^0(B)}\ \frac{1}{N}\sum_{n=0}^{N}\eta^{-n}P_1^n 1\ .$$

and we conclude that $|e_1| = 0$ which is a contradiction (recall that $e_1 \in C^0(\bar{B})$ from a previous lemma).

We conclude that $\rho_1 = \eta$, and we can choose $f_0 = e_1$ non-negative. Positivity almost everywhere follows again from $P_t \mathbb{1}_{B_r(y)}(x) > 0$.

Existence of an a.c.q.s.d. (absolutely continuous q.s.d.)

We now consider the adjoint operator $P_1^\dagger$ in L^2. Since the positive functions generate L^2, there exists a positive function g_0 such that $\pi_1^\dagger g_0 \neq 0$. Since $P_t^\dagger g_0$ is positive for any t, we deduce that the function g given by

$$g = \pi_1^\dagger g_0 = \lim_{N\to\infty} \frac{1}{N-1} \sum_{n=0}^{N} \eta^{-n} {P_1^\dagger}^n g_0$$

is nonnegative (the limit exists by the same argument as before). We also have

$$0 \neq \langle g_0 \,,\, f_0 \rangle = \left\langle g_0 \,,\, \frac{1}{N} \sum_{n=0}^{N-1} \eta^{-n} P_1^n f_0 \right\rangle = \left\langle \frac{1}{N} \sum_{n=0}^{N-1} \eta^{-n} {P_1^\dagger}^n g_0 \,,\, f_0 \right\rangle .$$

Taking the limit $N \to \infty$ we get $\langle g \,,\, f_0 \rangle \neq 0$, hence $g \neq 0$. Since $P_1^\dagger g = \eta g$. it follows by the spectral mapping theorem that g is the density of a q.s.d.

We now prove uniqueness.

We first observe that if ν is a q.s.d., since $\nu \circ P_1 = \lambda\nu$, and P_1 maps L^2 to $\mathcal{C}^0$, ν is a continuous linear functional on L^2 and hence absolutely continuous. We denote by g its density (note that g also belongs to L^1). Hence we have to prove uniqueness in L^2.

Assume e_1 and e_1' are two independent non-negative eigenvectors of P_1 with eigenvalue η, which differ on a set of positive measure, and such that
$$\int_B g\, e_1\, dx = \int_B g\, e_1'\, dx = 1\ .$$
Then, from $P_1 \mathbb{1}_{B_r(y)}(x) > 0$ and the continuity of e_1 and e_1'
$$\eta \left|e_1 - e_1'\right| = \left|P_1 e_1 - P_1 e_1'\right| < P_1\left|e_1 - e_1'\right|\ .$$
Integrating against g we get
$$\eta \int_B g\left|e_1 - e_1'\right| dx < \int_B g\, P_1\left|e_1 - e_1'\right| dx = \eta \int_B g\left|e_1 - e_1'\right| dx$$
a contradiction, hence the eigenvalue η of P_1 is simple. The same result holds for the adjoint $P_1^\dagger$, proving uniqueness of the q.s.d.

One can show more, namely that there is no other spectral point on the circle of radius η.
This spectral gap implies an exponential rate of convergence in L^2 for q.l.d. and Yaglom limits.
One can also give results on the regularity of the density of the q.s.d., central limit theorem etc.

The Ornstein-Uhlenbeck process on $\mathbb{R}^+$.

This process provides an interesting example and was studied in details by Lladser and San Martín. Recall that

$$dX = -X\ dt + dW_t\ .$$

The trap is $(-\infty, 0]$.

Theorem 10 (Lladser San Martín)

For any $\theta \in (0,1]$, there is a q.s.d. ν_θ absolutely continuous with respect to the Lebesgue measure and such that

$$\mathbb{P}_{\nu_\theta}\big(T > t\big) = e^{-\theta\, t}\ .$$

In particular we see here that contrary to the case of bounded domains, there is a continuum of q.s.d.

The densities u_θ of these measures are related to special functions, they are given by

$$u_\theta(x) = e^{-x^2/2} y_\theta(x)$$

where y_θ is a parabolic cylinder function. In particular

$$u_1(x) = 2xe^{-x^2} \ .$$

Which one is the good one?

Lladser and San Martín give a criteria to answer this question.

They observe that for $0 < \theta < 1$

$$u_\theta(x) \overset{x\,\text{large}}{\simeq} x^{-1-\theta}$$

Recall that f (> 0) is regularly varying with exponent β if for all $c > 0$

$$\lim_{u \to +\infty} \frac{f(cu)}{f(u)} = c^\beta \, .$$

Theorem 11 (Lladser San Martin)

Assume $f > 0$ (integrable) is regularly varying of exponent $-(1+\theta)$ $(0 < \theta < 1)$. Then for any Borel set in $\mathbb{R}^+$

$$\lim_{t \to \infty} \mathbb{P}_{f\,dx}\big(X_t \in A \,|\, T > t\big) = \nu_\theta(A) = \int_A u_\theta(y) dy \, .$$

The case $\theta = 1$ was treated earlier by Mandl under different assumptions.

Down from infinity.

Why such a difference of behaviors between bounded and unbounded domains: a continuum of q.s.d. versus one? This seems to be related to the behavior of the drift near infinity.

Theorem 12

The process $dZ = -Z^2 dt + dW_t$ on $\mathbb{R}^+$ (trap $(-\infty, 0]$) has a unique q.s.d. which is absolutely continuous with respect to Lebesgue.

To prove this result we start by showing that this process comes down from infinity very fast. This is not the case for the Ornstein-Uhlenbeck process, which comes down from infinity but not fast enough (the transition kernel from y to x is $z(t)^{-1} \exp(-a(t)(x - b(t)y)^2)$ with $a(t) > 0$ and $0 < b(t) < 1$ for $t > 0$).

Theorem 13

There exists a constant $C > 0$ such that for any $x, y \in \mathbb{R}^+$

$$\mathbb{P}_x\big(Z_1 > y\ ,\ T > 1\big) \le C\ e^{-\sqrt{1+y}}$$

In order to prove this result, we will use the function

$$\varphi(t,x) = e^{t\sqrt{1+x}}\ .$$

It is easy to verify that for any $x \ge 0$ and $t \in [0,1]$

$$\partial_t \varphi(t,x) - x^2 \partial_x \varphi(t,x) + \frac{1}{2}\partial_x^2 \varphi(t,x) \le \left(\frac{2}{\sqrt{t}} + 1\right)\ \varphi\ .$$

From Ito's formula it follows that for some constant $C > 0$, for any $x > 0$ and $t \in [0,1]$

$$\mathbb{E}_x\big(\varphi(t,Z_t)\mathbb{1}_{T>t}\big) \le 1 + \int_0^t \left(1 + \frac{2}{\sqrt{s}}\right) \mathbb{E}_x\big(\varphi(s,Z_s)\mathbb{1}_{T>s}\big)\, ds\ .$$

It follows from Gronwall Lemma that $\mathbb{E}_x\big(\varphi(1,Z_1)\mathbb{1}_{T>1}\big) < \infty$ and the result follows by a Chebyshev inequality.

First consequence.

Theorem 14

There exists a q.s.d.

For the proof, apply Krein's Theorem in $C_b^0(\mathbb{R}^+)$ to the operator P_1. One gets a positive eigenvector ν in the dual space.

We have to prove the eigenvalue $\lambda \geq 0$ is not zero. It is easy to show that $\liminf_{x\to\infty} \mathbb{P}_x(T > 1) > 0$. Since $\nu(P_1 1) = \lambda\, \nu(1)$, if $\lambda = 0$ the functional is identically zero, a contradiction.

Finally the functional is a measure since tightness follows from the previous estimate and the eigenvalue equation $\nu = \lambda^{-1}\, \nu \circ P_1$.

Second consequence.

Theorem 15

Let $\mathcal{B} = e^{\sqrt{x}/2} C_b^0(\mathbb{R}^+)$ (this is a Banach space). The operator P_1 is compact in $\mathcal{B}$. The peripheral spectrum is finite and contains a positive point which is a simple eigenvalue with positive eigenvector.

For the proof, one first show that P_1 extends to $\mathcal{B}$ using the previous estimate. Better, P_1 maps continuously $\mathcal{B}$ to $C_b^0(\mathbb{R}^+)$. Using Girsanov's theorem one proves an estimate $\big|P_1 f(x)\big| \leq C^{\mathrm{te}}\, x \,\|f\|_{\mathcal{B}}$ for $x \in (0,1]$. On any compact interval (in x) one uses Harnack regularity and compactness follows. The rest of the theorem is proved as in the case of bounded domains. Another proof relies on Gaussian bounds for the kernel in some wheighted L^2 space.

Third consequence.

Theorem 16

The q.s.d. is unique, a.c. with a continuous density.

By the q.s.d. equation, $P_1^\dagger \nu = \lambda\, \nu$, a q.s.d. belongs to $\mathcal{B}^*$, and we know by the previous result that it is unique.

Moreover, in $\mathbb{R}^+$, the q.s.d. ν satisfies in the sense of distributions

$$\frac{1}{2}\,\frac{d^2}{dx^2}\nu + \frac{d}{dx}\big(x^2\,\nu\big) = \log(\lambda)\,\nu$$

and therefore is a function.

As in the case of bounded domains, one can prove the existence of a spectral gap etc.

Q.S.D. for Dynamical Systems.

I will briefly explain one of the results: the Pianigiani Yorke measure (for discrete time dynamical systems).

Theorem 17

Let T be a $C^{1+\alpha}$ map of $\mathbb{R}^n$, and assume that there is an open set Ω_0 with compact closure such that

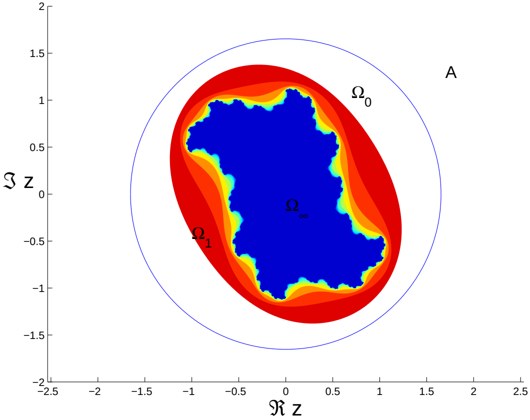
$$\overline{\Omega_0} \subset \mathrm{Int}\big(T(\Omega_0)\big) \ .$$

Assume there is a closed neighborhood V of $\overline{\Omega_0}$ such that

$$\sup_{x \in V} \big\|(DT_x)^{-1}\| < 1 \ .$$

Then there is an a.c.q.s.d. ν for the trap $A = \overline{\Omega_0}^c$.

The result can be generalized in many directions, and one has exponential convergence of q.l.d. (for adequate initial distributions).



An example of the setting for the map $z \mapsto T(z) = z^2 + .2 + .45\, i$ of the complex plane. Ω_0 is a disk, $\Omega_0 \subset T(\Omega_0)$ $\Omega_1 = \Omega_0 \cap T^{-1}(\Omega_0)$, etc. Ω_n is the set of initial conditions whose orbit stays in Ω_0 up to time n. The intersection Ω_∞ of this decreasing sequence is an invariant set (filled in Julia set). There is an a.c.q.s.d. with support in $\Omega_0 \backslash \Omega_\infty$.

The proof is rather similar to the above proofs for bounded domains and drifts coming down fast enough from infinity.
The density of a q.s.d. is an eigenvector of a Perron-Frobenius operator P given by

$$Pg(x) = \sum_{y\,,\,T(y)=x} \frac{g(y)}{J_T(y)}$$

where J_T is the Jacobian of T.
The main technical difference is that one shows that the operator P is quasi-compact in $C^{\alpha'}(\Omega_0)$ for some $0 < \alpha' < \alpha$. There is also a spectral gap.

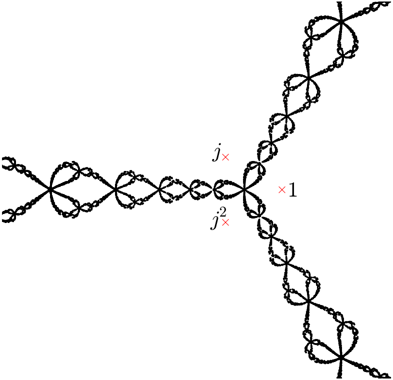
An example in numerical analysis: can you solve $z^3 = 1$?

Yes! but we will test Newton's method on this example. This amounts to iterating the map

$$T(z) = \frac{2\,z}{3} + \frac{1}{3\,z^2}\ .$$

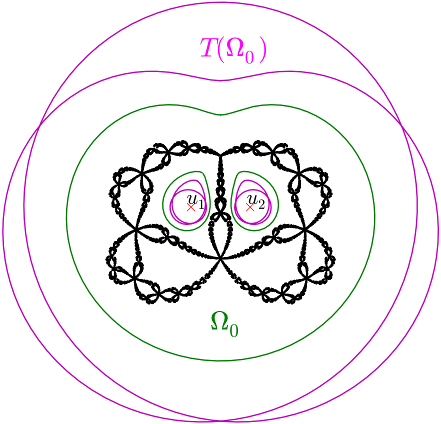
According to the choice of an initial condition, one converges to one of the three roots of $z^3 = 1$ which are the three (superstable) fixed points of T or not.

The set of initial conditions whose orbit does not converge to one of the roots of $z^3 = 1$.



The basins of attractions of 1, j, j^2 are three open sets with the same boundary.

In order to apply the Pianigiani-Yorke result, one can conjugate with a conformal mapping which maps 1 to infinity (and rotates).



This leads in the original coordinates to an absolutely continuous q.s.d. with density behaving like $|z|^{-4}$ near infinity which gives information about the convergence of Newton's method.

A short bibliography

P.Cattiaux, P.Collet, A.Lambert, S.Martínez, S.Méléard , J.San Martín. Quasi-stationary distributions and diffusions models in population dynamics. Ann. Probab. **37**, 1926-1969 (2009).

P.Collet, S.Martínez, J. San Martín. Quasi-Stationary Distributions, Stochastic Processes, Dynamical Systems, Applications. Prepared (2011).

M.Lladser, J.San Martín. Domain of Attraction of the Quasi-stationary distributions for the Ornstein-Uhlenbeck Process. Journal of Applied Probability **37**, 511-520 (2000).

T.Oikhberg, V.Troitsky. A theorem of Krein revisited. Rocky Mountain J. Math. **35**, 195-210 (2005).

R. Pinsky. On the convergence of diffusion processes conditioned to remain in bounded region for large time to limiting positive recurrent diffusion processes. Ann. Probab. **13**, 363-378 (1985).

P.Pollett. Quasi-stationary Distributions: A Bibliography.

`www.maths.uq.edu.au/ pkp/papers/qsds/qsds.pdf`