Quasi-invariant measures for a model in evolutionary biology

Mauro Mariani
Jointly with E. Pardoux

Dipartimento di Matematica, Università di Roma La Sapienza

Project partly supported by A*Midex Initiative d'excellence Aix-Marseille

Lyon, June 18th 2015

Outline

1. Muller's ratchet
   - Motivations
   - The model
2. Quasi-invariant measures
   - Definition
   - The 1d case
   - A possible approach in $>$ 1d
3. QIM for the Muller's ratchet
4. Open problems

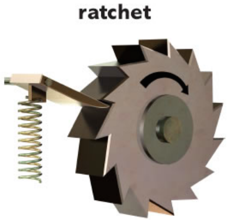
Outline

Muller's ratchet

- The Muller's ratchet is a basic paradigm in evolutionary biology. The concept was introduced by Muller in the 30s'. The point: genetic recombination provides evolution advantages when compared to asexual reproduction. The idea: in an asexual population, the accumulation of random deleterious mutations over many generations will drastically decrease the fitness of the species.

Muller's ratchet

- The Muller's ratchet is a basic paradigm in evolutionary biology. The concept was introduced by Muller in the 30s'. The point: genetic recombination provides evolution advantages when compared to asexual reproduction. The idea: in an asexual population, the accumulation of random deleterious mutations over many generations will drastically decrease the fitness of the species.
- The ratchet metaphor has been introduced in the 60s, to summarise the irreversibility of harmful genetic mutations.

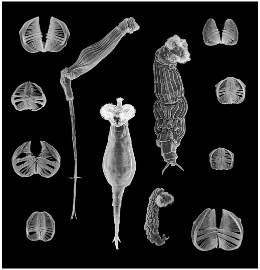


Muller's ratchet

- This has been a successful paradigm to explain the appearance of sexual reproduction or more generally genetic recombination. Nevertheless, there exist 'ancient' species featuring basic asexual reproduction (e.g. bdelloideas).

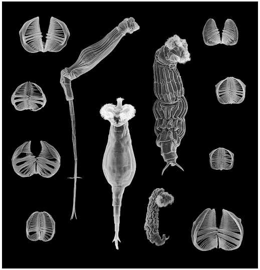
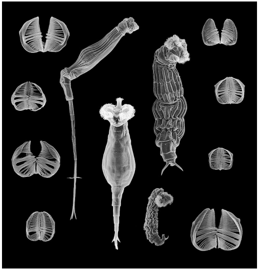
Muller's ratchet

- This has been a successful paradigm to explain the appearance of sexual reproduction or more generally genetic recombination. Nevertheless, there exist 'ancient' species featuring basic asexual reproduction (e.g. bdelloideas).
- Easy guess: superiority of evolution via genetic recombination.



Muller's ratchet

- This has been a successful paradigm to explain the appearance of sexual reproduction or more generally genetic recombination. Nevertheless, there exist 'ancient' species featuring basic asexual reproduction (e.g. bdelloideas).
- Easy guess: superiority of evolution via genetic recombination.

Stochastic diffusion model

- Mathematical modelling started in 70s'. We consider a Fleming-Viot model by Etheridge et al. 2008. The population size N is fixed, the species has d genes ($d = \infty$ is relevant), and $X_k(t)$ denotes the percentage of individuals carrying k mutations.

Stochastic diffusion model

- Mathematical modelling started in 70s'. We consider a Fleming-Viot model by Etheridge et al. 2008. The population size N is fixed, the species has d genes ($d = \infty$ is relevant), and $X_k(t)$ denotes the percentage of individuals carrying k mutations.
- $\{X_k(t),\ t \geq 0,\ k = 0,\ 1, \ldots\}$ solves the finite or infinite SDEs

$$\dot{X}_k(t) = \alpha(M(X) - k)\,X_k + \sum_h \lambda_{h,k} X_h + \sqrt{\frac{X_k\,X_h}{N}}\,\dot{B}_{k,h}$$

  $\{B_{k,h}\}$ are independent BMs but $B_{k,h} = -B_{h,k}$ (so $B_{k,k} = 0$). α= positive parameter accounting for fitness of healthy individuals. $\lambda_{h,k}$=rate from h mutations to k mutations. $\lambda_{k,k} = -\sum_{h \neq k} \lambda_{k,h}$ and $\lambda_{h,k} = 0$ if $h > k$.
- $1 = \sum_k X_k(t)$ and $M(X) = \sum_k k\,X_k$.

Remarks on the model

- If $N = \infty$ (no noise), then there is an attractive fixed point (if $d = \infty$ and mutation are nearest neighbours, then the fixed point is Poisson(α/λ)).
- Notice that $X_0 = 0$ is absorbing. If τ is the absorption time, then $\mathbb{P}(\tau < +\infty) = 1$. This is immediate if d is finite. Audiffren and Pardoux 2013 for the infinite system.
- Once $X_0 = 0$, then $X_1 = 0$ is absorbing and so on. That's the key feature of the ratchet model.

Several questions

One may consider many asymptotic problems.

- Large N (small noise).
- Large d.
- Many clicks.
- Long times provided no click happened.

Several questions

One may consider many asymptotic problems.

- Large N (small noise).
- Large d.
- Many clicks.
- Long times provided no click happened.

We focus on the last issue, and possibly large d. We fix $N = 1$ hereafter. To simplify notation:

$$\dot{X}_k(t) = \alpha(M(X) - k)\, X_k + \lambda\big(\varepsilon X_0 + (1-\varepsilon)(X_{k-1} - X_k)\big)$$
$$+ \sum_h \sqrt{X_k\, X_h}\dot{B}_{k,h}$$

$$a = (a_{h,k})_{h,k} = X_h\, \delta_{h,k} - X_h X_k$$

Outline

Quasi-invariant measures

- Let S_t be a sub-Markov semigroup (one can always realise it as a Markov semigroup, but that's a bad idea). A *probability* measure α is quasi-invariant if

$$\alpha(S_t f) = \alpha(f)\alpha(S_t \mathbf{1}) \qquad \text{for all bounded measurable } f$$

Otherwise stated

$$\mathbb{E}_\alpha(f(X_t)|\tau > t) = \alpha(f)$$

Quasi-invariant measures

- Let S_t be a sub-Markov semigroup (one can always realise it as a Markov semigroup, but that's a bad idea). A *probability* measure α is quasi-invariant if

$$\alpha(S_t f) = \alpha(f)\alpha(S_t \mathbf{1}) \qquad \text{for all bounded measurable } f$$

  Otherwise stated

$$\mathbb{E}_\alpha(f(X_t)|\tau > t) = \alpha(f)$$

- Necessarily $\alpha(S_t\mathbf{1}) = \mathbb{P}_\alpha(\tau > t) = \exp(-\theta\, t)$ for some $\theta \geq 0$. If $\theta = 0$ then the measure is actually *invariant*.

Quasi-invariant measures

- Let S_t be a sub-Markov semigroup (one can always realise it as a Markov semigroup, but that's a bad idea). A *probability* measure α is quasi-invariant if

$$\alpha(S_t f) = \alpha(f)\alpha(S_t \mathbf{1}) \qquad \text{for all bounded measurable } f$$

  Otherwise stated
$$\mathbb{E}_\alpha(f(X_t)|\tau > t) = \alpha(f)$$

- Necessarily $\alpha(S_t \mathbf{1}) = \mathbb{P}_\alpha(\tau > t) = \exp(-\theta\, t)$ for some $\theta \geq 0$. If $\theta = 0$ then the measure is actually *invariant*.
- It's a relevant notion in population dynamics. For the ratchet, τ is the first time when $X_0 = 0$.

Quasi-invariant measures

- Let S_t be a sub-Markov semigroup (one can always realise it as a Markov semigroup, but that's a bad idea). A *probability* measure α is quasi-invariant if

$$\alpha(S_t f) = \alpha(f)\alpha(S_t \mathbf{1}) \qquad \text{for all bounded measurable } f$$

  Otherwise stated

$$\mathbb{E}_\alpha(f(X_t)|\tau > t) = \alpha(f)$$

- Necessarily $\alpha(S_t \mathbf{1}) = \mathbb{P}_\alpha(\tau > t) = \exp(-\theta\, t)$ for some $\theta \geq 0$. If $\theta = 0$ then the measure is actually *invariant*.
- It's a relevant notion in population dynamics. For the ratchet, τ is the first time when $X_0 = 0$.
- If $d = \infty$, it appears hard to prove that the ratchet model is Markov in this sense.

Quasi-invariant measures

For a one-dimensional diffusion on the half line (and absorption at 0), the problem is well understood. Cattiaux, Collet, Lambert, Martinez, Méléard, San Martin 2009 consider

$$\dot{X} = -V'(X) + \dot{W}$$

V is confining at ∞ and $\exp(-2V(x))$ is not integrable at 0 (that is, there is absorption).

- Consider the invariant σ-finite measure $m(dx) = e^{-2V(x)}$.
- The generator is self-adjoint and admits a spectral gap on $L_2(dm)$. Largest (negative) eigenvalue has positive eigenfunction η.
- Under mild assumptions $\alpha(dx) = \eta(x)\, m(dx)$.

What's bad about the ratchet

The Muller's ratchet is a good test-bed for quasi-invariant measures techniques. Even in finite dimension

- Irreversible process. Unknown σ-finite invariant measure. The generator asymmetric part is singular, no matter the (non-invariant) reference measure considered.
- Non-locally flat manifold.
- Complex topological nature of the boundary (points at infinity, attainable, absorbing).

Spectral approach

We use a non-invariant reference measure m. The following general fact has an elementary proof. Suppose (informal statement)

- S_t extends to a bounded semigroup in $L_2(dm)$.
- There exists $\bar{t}$ such that $S_{\bar{t}}$ acts boundedly from $L_\infty(dm)$ to $L_2(dm)$.
- There exists $\theta \geq 0$, $\eta \in L_2(m)$, $\eta \geq 0$, such that $S_t^\dagger \eta = e^{-\theta\, t}\eta$.

Then $\eta \in L_1(m)$ and $\eta\, m$ is a quasi-invariant probability.

Spectral approach

A sufficient condition for the previous approach: assume

- Technicalities on the generator.
- Dirichlet form(f) + const. $\times \|f\|^2_{L_2(dm)} \geq 0$.
- Resolvent operator is (quasi)-compact in $L_2(dm)$.
- $\int_0^1 dt \int dm(x)\, \mathbb{P}_x(\tau > t)^2 < +\infty$.

Outline

What's good about the ratchet

- Ergodic diffusion on an Einstein manifold with positive curvature.

What's good about the ratchet

- Ergodic diffusion on an Einstein manifold with positive curvature.
- Ansatz for a reference measure that supposedly catches the singular behaviour of the invariant measure

$$V(x) = \tfrac{1}{2}\log\det(a) + \alpha M(x) + \lambda\Big(\frac{\varepsilon}{d}x_0\sum_{i=1}^{d}\log x_i + (1-\varepsilon)\big(x_d + \sum_{i=1}^{d} x_{i-1}\log x_i\big)\Big)$$

What's good about the ratchet

- Ergodic diffusion on an Einstein manifold with positive curvature.
- Ansatz for a reference measure that supposedly catches the singular behaviour of the invariant measure

$$V(x) = \frac{1}{2}\log\det(a) + \alpha M(x)$$
$$+ \lambda\Big(\frac{\varepsilon}{d}x_0\sum_{i=1}^{d}\log x_i + (1-\varepsilon)\big(x_d + \sum_{i=1}^{d}x_{i-1}\log x_i\big)\Big)$$

- Close to the boundaries, the diffusion matrix features a 1d-degeneracy. This gives spectral gap.

What's good about the ratchet

- Ergodic diffusion on an Einstein manifold with positive curvature.
- Ansatz for a reference measure that supposedly catches the singular behaviour of the invariant measure

$$V(x) = \frac{1}{2}\log\det(a) + \alpha M(x) + \lambda\Big(\frac{\varepsilon}{d}x_0\sum_{i=1}^{d}\log x_i + (1-\varepsilon)\big(x_d + \sum_{i=1}^{d} x_{i-1}\log x_i\big)\Big)$$

- Close to the boundaries, the diffusion matrix features a 1d-degeneracy. This gives spectral gap.
- Estimating of the probability of absorption is easy. (Change of variables and comparison with a 1d process).

What's good about the ratchet

- Ergodic diffusion on an Einstein manifold with positive curvature.
- Ansatz for a reference measure that supposedly catches the singular behaviour of the invariant measure

$$V(x) = \frac{1}{2}\log\det(a) + \alpha M(x) + \lambda\Big(\frac{\varepsilon}{d} x_0 \sum_{i=1}^{d} \log x_i + (1-\varepsilon)\big(x_d + \sum_{i=1}^{d} x_{i-1}\log x_i\big)\Big)$$

- Close to the boundaries, the diffusion matrix features a 1d-degeneracy. This gives spectral gap.
- Estimating of the probability of absorption is easy. (Change of variables and comparison with a 1d process).

Theorem

The Muller's ratchet admits a QIM measure $\alpha_d = \eta_d m_d$ for each d.

Outline

Open problems

- We have uniform estimates in d and t (up to infinity). But no compactness of α_d!

Open problems

- We have uniform estimates in d and t (up to infinity). But no compactness of α_d!
- (For finite or infinite d) Recall $1/N$ as a coefficient for the noise. Prove that, when $N \to \infty$, α concentrates on the fixed point of the deterministic dynamics. Is the rate of large deviations of α in N given by the Freidlin-Wentzell quasi-potential for the $\mathbb{Q}$-process?

Open problems

- We have uniform estimates in d and t (up to infinity). But no compactness of α_d!
- (For finite or infinite d) Recall $1/N$ as a coefficient for the noise. Prove that, when $N \to \infty$, α concentrates on the fixed point of the deterministic dynamics. Is the rate of large deviations of α in N given by the Freidlin-Wentzell quasi-potential for the $\mathbb{Q}$-process?
- (For finite or infinite d) Is the rate of large deviations of the empirical measure of the conditioned process given by the Donsker-Varadhan rate of the $\mathbb{Q}$-process?