

Filling the gap between individual-based evolutionary models and Hamilton-Jacobi equations

Sylvie Méléard

Ecole Polytechnique, Centre de Mathématiques Appliquées

Conference on stochastic analysis and stochastic partial differential equations,
Barcelone, May 2022

Joint work with N. Champagnat (Nancy), S. Mirrahimi (Montpellier) and V.C. Tran (Marne-la-Vallée).



Evolutionary biology

The population has the capacity to generate as well to select the individual diversity.

The ability of an individual to survive and reproduce depends on phenotypic (or genetic) parameters called traits.

The evolution of the trait distribution results from the following mechanisms:

- **Heredity.** (Vertical) transmission of the ancestral trait to the offsprings.
- **Mutation.** Generates variability in the trait values.
- **Selection.** Individuals with traits increasing their survival probability or their reproduction ability will spread through the population over time.
The selection also results from the competition between individuals.

Asexual populations (cells, bacteria).

Usual biological assumptions:

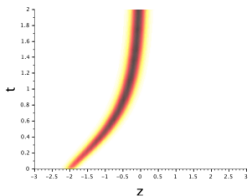
- large populations
- rare mutations
- small mutation steps
- long (evolutionary) time scale.

The main goal:

- predict the long term evolutionary dynamics.
- model and quantify the successive invasions of successful mutants: by mutation-selection, the population concentrates on advantageous mutants.

That is a multi-scale question : different mathematical approaches using different analytical tools.

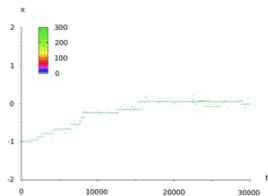
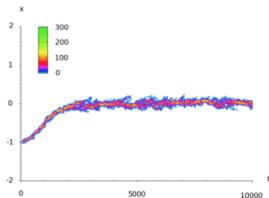
- **Game Theory - Dynamical Systems:**
Maynard-Smith 1974, Hofbauer-Sigmund 1990, Marrow-Law-Cannings 1992, Metz-Geritz-Meszéna et al. 1992, 1996, Dieckmann-Law 1996, Dieckmann 2004.
- **Partial or integro-differential and Hamilton-Jacobi equations (Hopf-Cole transformation):**
Perthame-Barles-Mirrahimi 07-10, Jabin, Desvillettes, Raoul, Mischler 08-10.
Concentration phenomenon on advantageous mutants but evolution seems too fast.



- **Stochastic individual-based processes** (birth and death processes with mutation and selection) :

(Bolker-Pacala 97, Kisdi 99, Dieckmann-Law 00, Fournier-M. 04, Ferrière-Champagnat-M. 06, Champagnat 06, Champagnat-M. 10).

Concentration phenomenon on advantageous mutants but evolution seems too slow (time scale separation between competition phases and mutation arrivals).



It is not clear how these models are related.

- Assume that the trait $x \in \mathcal{X}$. For an individual with trait x , the birth rate is $b(x)$, the mutation rate is $\rho(x)$ and the mutation kernel is $G(y - x)dy$, the death rate is $d(x)$ and the competition pressure is C .
- The stochastic population process is a point measure-valued process belonging to $\mathbb{D}(\mathbb{R}_+, M_F(\mathcal{X}))$ taking into account all birth and death events and renormalized by a parameter K , which is the order of the population size.
- When K tends to infinity, it converges in probability to the solution of the PDE

$$\begin{aligned} \partial_t n(t, x) = & \left(b(x) - d(x) - C \int n(t, y) dy \right) n(t, x) \\ & + \int_{\mathcal{X}} \rho(y) G(x - y) n(t, y) dy. \end{aligned}$$

To take into account the biological scales, the analysts (cf. Barles, Mirrahimi, Perthame 2009) assume that mutations are small and the evolutionary time scale is long. Therefore, they introduce $\varepsilon > 0$ and study

$$\begin{aligned} \varepsilon \partial_t n^\varepsilon(t, x) &= \left(b(x) - d(x) - C \int n^\varepsilon(t, y) dy \right) n^\varepsilon(t, x) \\ &+ \int_x p(y) \frac{1}{\varepsilon} G\left(\frac{x-y}{\varepsilon}\right) n^\varepsilon(t, y) dy. \end{aligned}$$

Writing

$$n^\varepsilon(t, x) = \exp\left(\frac{u^\varepsilon(t, x)}{\varepsilon}\right),$$

they prove (under technical assumptions) that when $\varepsilon \rightarrow 0$, the functions u^ε converge to the solution of a Hamilton-Jacobi equation.

But long term evolutionary dynamics may be strongly influenced by small subpopulations : emergence of new traits, local extinction.

How to keep track of small populations in large population models?

In particular, up to now, there were no results on
a good scaling relying directly the stochastic model to a Hamilton-Jacobi
equation.

Aim of the talk:

To introduce some scalings for which the stochastic population processes
converge to the viscosity solution of a Hamilton-Jacobi equation

The Population Model: a stochastic super-critical birth-death-mutation process

- **Large population:** the model is parametrized by a carrying capacity parameter K , $K \rightarrow +\infty$.
- Each individual is characterized by a trait $x \in \mathbb{T}$ modeling quantitative genetic or phenotypic information. (It could also model a location).
- The trait space is a discretization of the torus \mathbb{T} :

$$\mathcal{X}_K = \left\{ i\delta_K : i \in \{0, 1, \dots, \frac{1}{\delta_K} - 1\} \right\},$$

with δ_K such that $1/\delta_K \in \mathbb{N}$ and $\delta_K \rightarrow 0$ when $K \rightarrow \infty$.

- We define the stochastic processes:

$(N_i^K(t), t \geq 0)$: the number of individuals with trait $i\delta_K$ at time t .

The population process $((N_i^K(t), i \in \{0, 1, \dots, \frac{1}{\delta_K} - 1\}), t \geq 0)$ is a multitype birth-and-death process.

The transitions

- **Asexual reproduction:** an individual with trait $x \in \mathcal{X}_K$ gives birth at rate $b(x)$ to a new offspring with same trait x .
- **Natural death:** an individual with trait $x \in \mathcal{X}_K$ dies at rate $d(x)$.
- **Mutation:** For all $y \in \mathcal{X}_K$, an individual with trait $x \in \mathcal{X}_K$ gives birth to a mutant with trait y at rate

$$p(x)\delta_K \log K G(\log K(x - y)).$$

- b , d and p are Lipschitz continuous on \mathbb{T} and $\forall x \in \mathbb{T}$,

$$b(x) > d(x) \quad \text{and} \quad p(x) > 0.$$

- G is nonnegative, defined on \mathbb{R} , satisfies $\int_{\mathbb{R}} G(y) dy = 1$ and has exponential moments of any order.

For instance, $G(h) = \frac{1}{\sqrt{2\pi\sigma}} e^{-h^2/2\sigma^2}$.

The assumptions

- There exists $a_1 > 0$ such that for all K and $\forall i \in \{0, 1, \dots, \frac{1}{\delta_K} - 1\}$,

$$N_i^K(0) \geq K^{a_1}.$$

- There exists $a_2 < a_1$ such that as $K \rightarrow \infty$,

$$K^{-a_2/4} \ll \delta_K \ll \frac{1}{\log K}.$$

That implies

$$h^K := \delta_K \log K \ll 1.$$

Then for all $x \in \mathbb{T}$, the total mutation rate from an individual with trait $x_K = i_K \delta_K$ with $i_K = \lfloor x/\delta_K \rfloor$ to any trait, converges as $K \rightarrow +\infty$ to a positive value:

$$\lim_{K \rightarrow +\infty} p(x_K) \sum_{j=0}^{\lfloor \frac{1}{\delta_K} \rfloor} h_K G(h_K(i_K - j)) = p(x) \int_{\mathbb{R}} G(y) dy = p(x).$$

To summarize the different scalings:

- Population size of order $K \rightarrow +\infty$.
- The individual mutation rate $p(x)$ is of order 1.
- The mutations are small: scale $\frac{1}{\log K}$.
- Long time scale: $\log K$.
- Discretization mesh: scale $\ll \frac{1}{\log K}$.

We wish to capture the sub-populations of size K^α .

In the case where $p(x) = 0$ for all $x \in [0, 1]$ (no mutation), each process $N_i^K(t)$ is a branching process, hence

$$\mathbb{E}[N_i^K(t)] = \mathbb{E}(N_i^K(0))e^{(b(i\delta_K) - d(i\delta_K))t}.$$

Therefore, if $N_i^K(0) = K^{a_1}$,

$$\mathbb{E}[N_i^K(t \log K)] = K^{a_1 + (b(i\delta_K) - d(i\delta_K))t}.$$

This suggests to introduce

$$\beta_i^K(t) = \frac{\log(N_i^K(t \log K))}{\log(K)},$$

(with the convention that $\beta_i^K(t) = 0$ if $N_i^K(t \log K) = 0$).

We will study the convergence, as $K \rightarrow +\infty$, of the exponent processes $(\beta_i^K(t), i \in \{0, 1, \dots, \frac{1}{\delta_K} - 1\})_{t \geq 0}$.

For all $x \in \mathbb{T}$ and $K \geq 1$, let i be such that $x \in [i\delta_K, (i+1)\delta_K)$ and define

$$\tilde{\beta}^K(t, x) = \beta_i^K(t) \left(1 - \frac{x}{\delta_K} + i\right) + \beta_{i+1}^K(t) \left(\frac{x}{\delta_K} - i\right),$$

with the convention $\beta_{1/\delta_K}^K(t) = \beta_0^K(t)$.

The processes $\tilde{\beta}^K$ belong to $\mathbb{D}([0, T], C(\mathbb{T}, \mathbb{R}))$, endowed with the Skorokhod topology.

Assumption on the initial condition $\beta_i^K(0)$: there exists $A > 0$, such that

$$\lim_{K \rightarrow \infty} \mathbb{P} \left(\sup_{i \neq j} \frac{|\beta_i^K(0) - \beta_j^K(0)|}{\rho(j\delta_K, i\delta_K)} > A \right) = 0,$$

where ρ is the torus distance.

The main result

Theorem

Assume that $(\tilde{\beta}^K(0, \cdot))_K$ converges to a deterministic function $\beta_0(\cdot)$ and assumptions above. Then the processes $\tilde{\beta}^K$ converge in probability in $\mathbb{D}([0, T], C(\mathbb{T}, \mathbb{R}))$ to the unique Lipschitz viscosity solution of the Hamilton-Jacobi equation (HJ)

$$\begin{cases} \frac{\partial}{\partial t} \beta(t, x) = b(x) - d(x) + p(x) \int_{\mathbb{R}} G(h) e^{h\partial_x \beta(t, x)} dh, & (t, x) \in \mathbb{R}_+ \times \mathbb{T} \\ \beta(0, x) = \beta_0(x), & x \in \mathbb{T}. \end{cases}$$

That gives a constructive proof for the existence of a solution of (HJ) and could give a scheme for its simulation.

In our toy model, we have evacuated the main difficulty: a competition term which would create extinction events for the stochastic process.

Main steps of the proof

- Compactness-Uniqueness argument:

We prove that the sequence of laws of the processes $\tilde{\beta}^K$ is relatively compact in $\mathcal{P}(\mathbb{D}([0, T], C(\mathbb{T}, \mathbb{R})))$ and we identify any limiting value as the unique Lipschitz viscosity solution of (HJ).

- **Relative compactness:** criterion due to Jakubowski '86.

(i) $\forall \varepsilon > 0$, there exists a compact $C_\varepsilon \subset C(\mathbb{T}, \mathbb{R})$ such that

$$\forall K, \mathbb{P}(\tilde{\beta}^K \in \mathbb{D}([0, T], C_\varepsilon)) > 1 - \varepsilon.$$

(ii) $\forall f \in C(\mathbb{T}, \mathbb{R})$, the sequence of laws of the processes

$$\int_{\mathbb{T}} \tilde{\beta}^K(., x) f(x) dx$$

is relatively compact.

Using the characterization of C_ε by Ascoli theorem, we need to obtain equi-boundedness and equi-continuity estimates (uniformly in K and $t \leq T$) for

$$\tilde{\beta}^K(t, x) = x \left(\frac{\beta_{i+1}^K(t) - \beta_i^K(t)}{\delta_K} \right) + i(\beta_{i+1}^K(t) - \beta_i^K(t)) + \beta_i^K(t).$$

The main point is to estimate

$$\Delta_K \beta_i^K(t) = \frac{\beta_{i+1}^K(t) - \beta_i^K(t)}{\delta_K}.$$

We use the semimartingale decomposition of the processes $\beta_i^K, i \in \{0, \dots, 1/\delta_K - 1\}$.

We have that

$$\beta_i^K(t) = M_i^K(t) + A_i^K(t)$$

with M_i^K the martingale part and A_i^K a finite variation process .

The stochastic processes

$$\begin{aligned} A_i^K(t) &= \beta_i^K(0) \\ &+ \frac{1}{\log K} \int_0^{t \log K} N_i^K(s) \left(b(i\delta_K) \log \left(1 + \frac{1}{N_i^K(s)} \right) + d(i\delta_K) \log \left(1 - \frac{1}{N_i^K(s)} \right) \right) ds \\ &+ \frac{1}{\log K} \rho(i\delta_K) \sum_j h_K b(j\delta_K) G(h_K(j-i)) \int_0^{t \log K} N_j^K(s) \log \left(1 + \frac{1}{N_i^K(s)} \right) ds, \end{aligned}$$

M_i^K is a square integrable martingale such that

$$\begin{aligned} &\mathbb{E} \left(\sup_{t \leq T} (M_i^K(t))^2 \right) \\ &= \mathbb{E} \left(\frac{1}{\log^2 K} \int_0^{T \log K} N_i^K(s) \left(b(i\delta_K) \log^2 \left(1 + \frac{1}{N_i^K(s)} \right) + d(i\delta_K) \log^2 \left(1 - \frac{1}{N_i^K(s)} \right) \right) ds \right. \\ &\quad \left. + \frac{1}{\log^2 K} \rho(i\delta_K) \sum_j h_K b(j\delta_K) G(h_K(j-i)) \int_0^{T \log K} N_j^K(s) \log^2 \left(1 + \frac{1}{N_i^K(s)} \right) ds. \right) \end{aligned}$$

- Let $a \in (a_2, a_1)$ and define

$$\tau'_K = \inf\{t \geq 0, \exists i \in \{0, 1, \dots, \frac{1}{\delta_K} - 1\}; N_i^K(t \log K) < K^a\},$$

and for $L > 0$, we define

$$\tau_K = \tau_K(L) = \inf\{t \geq 0 : \exists i, j |\beta_j^K(t) - \beta_i^K(t)| > L \rho(i\delta_K, j\delta_K)\}.$$

and

$$\theta_K(L) = \tau_K(L) \wedge \tau'_K.$$

A key lemma: for all $T > 0$, there exists L_0 large enough such that

$$\lim_{K \rightarrow +\infty} \mathbb{P}(\theta_K(L_0) > T) = 1.$$

We will use repeatedly that, for all $t \leq \theta_K(L)$ and all $i, j \leq 1/\delta_K - 1$,

$$\frac{N_j^K(t \log K)}{N_i^K(t \log K)} = \exp(\log K (\beta_j(t) - \beta_i(t))) \leq e^{L \log K \rho(i\delta_K, j\delta_K)}.$$

Proof of tightness

Using

$$\sup_{K \geq 1} \sum_{j=0}^{\frac{1}{\delta_K} - 1} h_K G(h_K(i_k - j)) e^{L \log K \rho(j\delta_K, i_k\delta_K)} =: \bar{G}(L) < +\infty,$$

bounds on the quadratic variation and submartingale maximal lemma:

$$\mathbb{P}\left(\sup_{s \leq t \wedge \theta_K(L)} |\Delta_K M_i^K(s)| > \varepsilon\right) \leq \frac{1}{\varepsilon} \sqrt{\frac{2C(1 + \bar{G}(L))t}{\delta_K^2 K^a \log K}}.$$

Taking $\varepsilon_K = \delta_K^{-1} (K^a \log K)^{-1/4} \rightarrow 0$, we set

$$\Omega_K(L) = \left\{ \sup_{0 \leq i \leq 1/\delta_K - 1; t \leq T \wedge \theta_K(L)} |\Delta_K M_i^K(t)| \leq \varepsilon_K \right\}$$

Then there exists $C(L)$ such that

$$\mathbb{P}(\Omega_K^c(L)) \leq \sqrt{C(L)T} \varepsilon_K.$$

We will work on the probability subspace $\Omega_K(L)$.

Control of the variation part

One can prove that on the event $\Omega_K(L)$,

$$\begin{aligned} A_i^K(t \wedge \theta_K(L)) - A_i^K(s \wedge \theta_K(L)) &\leq C(\bar{b} + \bar{d})(t - s) \\ &+ \frac{1}{\log K} \sum_{\ell} h_K p((i + \ell)\delta_K) G(h_K \ell) \\ &\quad \int_{(s \wedge \theta_K) \log K}^{(t \wedge \theta_K) \log K} e^{\varepsilon_K \log K} \exp(\log K (A_{i+\ell}^K(u) - A_i^K(u))) du. \end{aligned}$$

We divide by $(t - s)$ and let s tend to $t \leq \theta_K(L)$.

Setting $\tilde{A}_i^K(t) = A_i^K(t) - 2C(\bar{b} + \bar{d})t - 2\bar{p}t$, we deduce that for any $t \leq \theta_K$,

$$\frac{d\tilde{A}_i^K(t)}{dt} < \bar{p} e^{\varepsilon_K \log K} \sum_{\ell} h_K G(h_K \ell) \exp(\log K (\tilde{A}_{i+\ell}^K(t) - \tilde{A}_i^K(t))) - 2\bar{p}. \quad (1)$$

Maximum principle (ω by ω).

Let us introduce

$$(i_K, t_K) = (i_K(\omega), t_K(\omega)) = \operatorname{argmax}_{i \in \{1, \dots, \frac{1}{\delta_K} - 1\}, t \in [0, \theta_K(L)]} \tilde{A}_i^K(t).$$

We have that

$$t_K = 0.$$

Indeed, if conversely we assume that $t_K > 0$, then the right term of (1) is non positive for $i = i_K$ and then the left term is negative, contradicting the fact that $\tilde{A}_{i_K}^K(t)$ is maximal for $t = t_K$.

Then, almost surely on the event $\Omega_K(L)$, for all $t \leq \theta_K(L)$ and $0 \leq i \leq 1/\delta_K - 1$,

$$A_i^K(t) \leq \max_{0 \leq j \leq 1/\delta_K - 1} \beta_j^K(0) + 2C(\bar{b} + \bar{d})t + 2\bar{p}t.$$

We deduce from the previous results that for all $T > 0$, there exists $C(T)$ such that,

$$\lim_{K \rightarrow +\infty} \mathbb{P} \left(\sup_{0 \leq i \leq 1/\delta_K - 1} \sup_{t \in [0, T \wedge \theta_K(L)]} \beta_i^K(t) \geq \|\beta_0\|_\infty + C(T) \right) = 0.$$

Let us now define

$$g_i^K(t) = \Delta_K A_i^K(t \wedge \theta_K) + \frac{\|\rho\|_{Lip}}{\underline{\rho}} A_{i+1}^K(t \wedge \theta_K).$$

We prove (tedious computation) that for $\omega \in \Omega_K(L)$ and $s \leq t$,

$$g_i^K(t) - g_i^K(s) \leq C(K, L)(t - s) + \bar{\rho} \log K \int_{(s \wedge \theta_K)}^{(t \wedge \theta_K)} \sum_{\ell} h_K G(h_K \ell) \left[g_{\ell+i}^K(v) - g_i^K(v) \right]_+ e^{h_K L |\ell|} dv.$$

Using again again the maximum principal, we prove that, almost surely on the event $\Omega_K(L)$, for all $t \leq \theta_K(L)$ and $0 \leq i \leq 1/\delta_K - 1$,

$$g_i^K(t) \leq \max_{0 \leq j \leq 1/\delta_K - 1} g_j^K(0) + 2C(K, L)t = A + 2C(K, L)t.$$

Finally, if $\tilde{\Omega}_K$ is the probability space allowing to control $\beta_i^K(0)$ and $\Delta_K \beta_i^K(0)$ when $K \rightarrow \infty$, then on $\Omega_K(L) \cap \tilde{\Omega}_K$, we have

$$|\Delta_K \beta_i^K(t \wedge \theta_K)| \leq C \left(A + \|\beta_0\|_\infty + T + 2 + \tilde{C}(K, L) \right),$$

with $\tilde{C}(K, L) < 1$ for K large enough (uniformly in L).

Thus the key lemma is proved ($L_0 = A + \|\beta_0\|_\infty + T + 3$).

The tightness of the sequence $(\tilde{\beta}^K)_K$ follows. More precisely we can prove:

Corollary:

The sequence of laws of $(\tilde{\beta}_t^K, t \in [0, T])_K$ is C -tight in $\mathcal{P}(\mathbb{D}([0, T], C(\mathbb{T}, \mathbb{R})))$.

In addition, for all $T > 0$, for any β distributed as a limiting value of the laws of $(\tilde{\beta}_t^K, t \in [0, T])_K$, we have almost surely

$$\sup_{t \in [0, T]} \sup_{x, y \in \mathbb{T} \text{ s.t. } x \neq y} \frac{|\beta(t, x) - \beta(t, y)|}{\rho(x, y)} \leq L_0.$$

Identification of the limit

Let $\beta \in C([0, +\infty) \times \mathbb{T}, \mathbb{R})$ be distributed as a limiting value of the laws of $(\tilde{\beta}_t^K, t \in [0, +\infty))_K$.

Using Skorokhod's representation theorem, there exist a new probability space and random variables again denoted by $\tilde{\beta}^{K_p}$ and β on this space, such that $(\tilde{\beta}^{K_p})$ converges almost surely to β on a set $\tilde{\Omega}_0$.

Then

$$\Omega_0 := \tilde{\Omega}_0 \cap \limsup_K \Omega_K(L_0) \cap \tilde{\Omega}_K$$

has probability 1.

To prove that β is a viscosity solution of (HJ), we work ω by ω in Ω_0 and proceed classically (cf. Mirrahimi-Barles-Perthame-Souganidis 2012).

Let $\omega \in \Omega_0$ and $T > 0$ be fixed and consider a smooth function $\varphi : [0, T] \times \mathbb{T}$ (depending on ω) such that $\beta(\omega) - \varphi$ attains a strict global maximum on $[0, T] \times \mathbb{T}$ at the point $(\bar{t}(\omega), \bar{x}(\omega))$ such that $\bar{t}(\omega) > 0$ and $\bar{x}(\omega) \in \mathbb{T}$.

We prove that

$$\frac{\partial}{\partial t} \varphi(\bar{t}, \bar{x}) \leq b(\bar{x}) - d(\bar{x}) + p(\bar{x}) \int_{\mathbb{R}} G(h) e^{h \partial_x \varphi(\bar{t}, \bar{x})} dh.$$

Then β is a viscosity sub-solution of (HJ) in $(0, T] \times \mathbb{T}$.

Following similar arguments, we also prove that β is a viscosity super-solution of (HJ) in $(0, T] \times \mathbb{T}$.

The theorem follows from the uniqueness of a Lipschitz viscosity solution of (HJ) (Barles, Mirrahimi, Perthame '09).

Perspectives

- To relax the periodic boundary conditions and to consider a changing sign growth rate
- To include a competition term
- To derive a modified Hamilton-Jacobi equation
- To study the properties of such an equation and to understand the population dynamics

Thank you for your attention!

