Some epidemic systems are long range interacting particle systems

CHRISTIAN LÉONARD

1. INTRODUCTION

We present some recent results about dynamical interacting particle systems in the setting of epidemiology. The individuals are particles whose states (of health) depend on their relative positions. These individuals interact since they fall ill more often when their neighbours are infectious. We begin with a description of the interaction between two individuals at the microscopic level. Then we study the behaviour of the whole system at the macroscopic level, when the number of individuals tends to infinity.

Let $U_i^N(\omega)$ be the random state of the individual i in the system of N individuals: $\{j; 1 \leq j \leq N\}$. For instance U_i^N may specify its position in a geographical space, its state of health and its deterministic type. Let us denote $\mathbf{X}$ the *fiber*, that is the set of the possible values of each $U_i^N(\omega)$. We suppose that all individuals of the same type are similar, therefore the *configuration* $(U_i^N(\omega); 1 \leq i \leq N) \in \mathbf{X}^N$ is completely described by the empirical probability measure

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{U_i^N(\omega)} \in \Pi(\mathbf{X})$$

Here δ_x is the Dirac measure at point x. In all what follows, if M is a topological space equipped with its Borel σ-algebra, $\Pi(M)$ stands for the set of all probability measures on M and it is endowed with its natural topology $\sigma(\Pi(M), C_b(M))$.

When studying the behaviour of the system as N tends to infinity, it is worth representing it in terms of its empirical measure. Indeed, it allows us to imbed the sequence of sets $(\mathbf{X}^N, N \geq 1)$ in the unique set $\Pi(\mathbf{X})$. Therefore, $\Pi(\mathbf{X})$ is the natural set of all configurations. One could consider

$$\mathbf{X} = \{\text{position at time } t\} \times \{\text{state of health}\} \times \{\text{type}\} \subset \mathbb{R}^k$$

and study the evolution of the measure-valued stochastic process

$$(\omega, t) \Longrightarrow \frac{1}{N}\sum_{i=1}^{N} \delta_{U_i^N(\omega,t)} \in \Pi(\mathbb{R}^k)$$

Instead of this, we shall look at the random empirical measure

$$\omega \Longrightarrow \frac{1}{N}\sum_{i=1}^{N} \delta_{U_i^N(\omega,\cdot)} \in \Pi((\mathbb{R}^k)^{[0,T]})$$

Hence, the fiber $\mathbf{X}$ is the set $(\mathbb{R}^k)^{[0,T]}$ of all the paths from [0,T] into $\mathbb{R}^k$. Other measure-valued stochastic processes have already been used to modelize spatial branching processes arising in biology. One should have a look to the survey paper, written

J.-P. Gabriel et al. (eds.), *Stochastic Processes in Epidemic Theory*

by *D.A. Dawson (1984)*, on this subject. The first result we shall present, is a law of large numbers. Under some regularity assumptions, it states the existence of a unique probability measure μ on $\mathbf{X} = (\mathbb{R}^k)^{[0,T]}$ of all the paths from [0,T] into $\mathbb{R}^k$ such that:

$$(1.1) \qquad \text{almost surely in } \omega \ , \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{U_i^N(\omega,\cdot)} = \mu \qquad \text{in } \Pi(\mathbf{X})$$

Then, we shall give an estimate for the probability of the large deviations from this almost sure event. This result is of the type:

$$(1.2) \qquad \lim_{N\to\infty} \frac{1}{N} \log P(\omega; \frac{1}{N} \sum_{i=1}^{N} \delta_{U_i^N(\omega,\cdot)} \in A) = -\inf\{I(v); v \in A\}$$

for "regular enough" subsets A of $\Pi(\mathbf{X})$.

The function $I(\cdot)$ on $\Pi(\mathbf{X})$ takes its values in $\mathbb{R}_+ \cup \{+\infty\}$ and is such that $I(v) = 0$ if and only if $v = \mu$ (which is given by (1.1)). It is called the *rate function* for the large deviations, since the probability of the event $\{\omega; \frac{1}{N}\sum_{i=1}^{N} \delta_{U_i^N(\cdot,\omega)} \in A\}$ is approximately $\exp(-N \inf\{I(v);\ v \in A\})$ when N is large. If A is such that $\inf\{I(v);\ v \in A\}$ is positive, this event becomes rarer and rarer as N increases. Notice that in this case, μ does not belong to A and the event "contradicts" the law of large numbers (1.1). At section 5, we shall explain why solving a variational problem involving the rate function $I(\cdot)$ can provide an answer to a question such as:

"What is the most probable path leading to a given final state from a given initial state?" .

2. The epidemic system

Let us describe our model at the microscopic level. Let N individuals: $1 \le i \le N$, each of type a or b (let us say that a=male and b=female). All these individuals move randomly in $\mathbb{R}^2$ and are likely to be struck down by a contagious disease. The individual i at time t $(0 \le t \le T)$, is described by:

$$\begin{cases} s_i^N \in \{a,b\} & : \text{its deterministic and constant type} \\ y_i^N(\omega,t) \in \mathbb{R}^2 & : \text{its random position} \\ e_i^N(\omega,t) \in \{S,I,D\} & : \text{its random state of health.} \end{cases}$$

Here S stands for "susceptible", I for "infectious" and D for "dead". The possible transitions for the state of health are : $S \rightleftharpoons I \rightarrow D$. A susceptible individual may fall ill after getting in contact with infectious ones of the alternate type (heterosexual transmission of a s.t.d.). An infectious individual can recover or die after random waiting times. Once it is dead, it stops moving and cannot infect anyone anylonger. We shall only consider Markov dynamics, so that the random waiting times have exponential distributions and the individuals move accordingly to a Markov process $(Y_i^N)_{1\le i\le N} :\ \ \Omega \times [0,T] \implies (\mathbb{R}^2)^N$. With the notations of the previous section, we take

$$\mathbf{X} = D\left([0,T], \mathbb{R}^2 \times \{S,I,D\}\right) \times \{a,b\}$$

where $D([0,T], \mathbb{R}^2 \times \{S,I,D\})$ is the space of the right continuous paths from $[0,T]$ to $\mathbb{R}^2 \times \{S,I,D\}$ which admit left limit everywhere.

Each individual moves independently from the others according to a jump Markov process in $\mathbb{R}^2$ with generator

$$g(\cdot) \implies \int_{\mathbb{R}^2_*} \{g(\cdot + x) - g(x)\} \, \mathcal{L}_Y(\cdot, dx)$$

Here $\mathcal{L}_Y$ is a *Lévy kernel*, that is:

for every $y \in \mathbb{R}^2$, $\mathcal{L}_Y(y,\cdot)$ is a nonnegative measure on the set of jumps $\mathbb{R}^2_* = \mathbb{R}^2 \setminus \{0\}$. If for every $y \in \mathbb{R}^2_*$, $\mathcal{L}_Y(y,\cdot)$ is a bounded measure, then one can interpret this generator as follows. Conditionally on $y_t = z \in \mathbb{R}^2$, the probability of a jump during the time interval $]t, t+dt]$ is $(\int_{\mathbb{R}^2_*} \mathcal{L}_Y(z,dx))dt \; + \; o(dt)$ and conditionally on the occurence of this jump, the law of this jump is the probability measure $\mathcal{L}_Y(z,\cdot)/\int_{\mathbb{R}^2_*} \mathcal{L}_Y(z,dx)$ on $\mathbb{R}^2_*$.

On the other hand, knowing the configuration of the whole system at time t : $(y_j^N(t), e_j^N(t))_{1 \leq j l N}$, the state of health of the individual i will perform a jump during the time interval $]t, t+dt]$

$$\begin{cases} \text{from the state } e_i(t) = I \text{ to the state } D \text{ with probability:} & C_{ID}dt + o(dt) \\ \text{from the state } e_i(t) = I \text{ to the state } S \text{ with probability:} & C_{IS}dt + o(dt) \quad \text{and} \\ \text{from the state } e_i(t) = S \text{ to the state } I \text{ with probability:} & \end{cases}$$

$$C_{SI}\left(y_i^N(t), s_i^N, \frac{1}{N}\sum_{j=1}^{N} \delta_{(y_j^N(t), e_j^N(t), s_j^N)}\right) dt + o(dt) \qquad \text{where } C_{ID}, C_{IS}, C_{SI}(\cdot) \geq 0.$$

The instantaneous rate of transition $C_{SI}(i)$ from S to I of the individual i depends on its type s_i^N and its position $y_i^N(t)$ but also on the configuration of the whole system by means of its empirical measure $\frac{1}{N}\sum_{j=1}^N \delta_{(y_j^N(t), e_j^N(t), s_j^N)}$ which belongs to $\Pi(\mathbb{R}^2 \times \{S,I,D\} \times \{a,b\})$. We shall choose for $C_{SI}(i)$ an increasing function of the ratio of infectious individuals in the disk of centre $y_i^N(t)$ and radius R whose types are different from i's one.

Denoting $u_t = \frac{1}{N}\sum_{j=1}^N \delta_{(y_j^N(t), e_j^N(t), s_j^N)}$ and $J(y)$ the indicator function of the set $\{\|y\| \leq R\}$, we take:

(2.1)

$$C_{SI}(y_i^N(t), a, u_t) = c_a \left(\int_{\mathbb{R}^2} J(y - y_i^N(t)) u_t(dy, I, b)\right) \Big/ \left(\varepsilon + \int_{\mathbb{R}^2} J(y - y_i^N(t)) u_t(dy, \{I,S\}, b)\right)$$

$$C_{SI}(y_i^N(t), b, u_t) = c_b \left(\int_{\mathbb{R}^2} J(y - y_i^N(t)) u_t(dy, I, a)\right) \Big/ \left(\varepsilon + \int_{\mathbb{R}^2} J(y - y_i^N(t)) u_t(dy, \{I,S\}, a)\right)$$

with $c_a, c_b \geq 0$ and $\varepsilon > 0$ (to prevent dividing by zero).

Finally, the model is described by the sequence $(s_i^N; 1 \le i \le N)$ in $\{a,b\}$ and the Markov process $\{(Y_i^N, E_i^N); 1 \le i \le N\} : \Omega \times [0,T] \Longrightarrow (\mathbb{R}^2 \times \{S,I,D\})^N$ whose generator applied to the real function $(y_1, e_1; \dots; y_N, e_N) \to f(y_1, e_1; \dots; y_N, e_N)$ gives:

$$\sum_{i=1}^{N} \int_{\mathbb{R}^2_*} \{f(\dots; y_{i-1}, e_{i-1}; y_i + x, e_i; y_{i+1}, e_{i+1}; \dots) - f(\dots; y_i, e_i; \dots)\} \mathbb{1}_{e_i \neq D} \mathcal{L}_Y(y_i; dx)$$

$$+ \sum_{i=1}^{N} \int_{\{S,I,D\} \to \{S,I,D\}} \{f(\dots; y_{i-1}, e_{i-1}; y_i, \overbrace{e_i + (e_i \to e)}^{=e}; y_{i+1}, e_{i+1}; \dots) - f(\dots; y_i, e_i; \dots)\}$$
$$\mathcal{L}_E\left(y_i, e_i, s_i^N, \frac{1}{N}\sum_{j=1}^{N} \delta_{(y_j(t), e_j(t), s_j^N)}; d(e_i \to e)\right)$$

Here $(e_i \to e)$ is the jump from e_i to e and the nonnegative measure $\mathcal{L}_E(y, e, s, u; \cdot)$ on the set of jumps $\{S,I,D\} \to \{S,I,D\}$ is defined for any $y \in \mathbb{R}^2, e \in \{S,I,D\}, s \in \{a,b\}$ and any probability measure u on $\mathbb{R}^2 \times \{S,I,D\} \times \{a,b\}$ by :

$$\begin{cases} \mathcal{L}_E(y, D, s, u; \cdot) = & 0 \\ \mathcal{L}_E(y, I, s, u; \cdot) = & C_{IS}\delta_{(I \to S)}(\cdot) + C_{ID}\delta_{(I \to D)}(\cdot) \\ \mathcal{L}_E(y, S, s, u; \cdot) = & C_{SI}(y, s, u)\delta_{(S \to I)}(\cdot) \end{cases} .$$

3. A general long range interacting particle system

The above epidemic system is a particular case of a long range interacting (or mean field) particle system, which we describe below. Let us give ourself a triangular array $\{(s_i^N)_{1 \le i \le N}; N \ge 1\}$ of a subset $\mathcal{S}$ of $\mathbb{R}^k$ and let us suppose that there exists $m \in \Pi(\mathcal{S})$ such that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{s_i^N} = m(ds) \quad \text{in } \Pi(\mathcal{S}) \tag{3.1}$$

Let us consider the Markov process $(X_i^N)_{1 \le i \le N}$ with values in $\mathcal{Z}^N$, where $\mathcal{Z}$ is a subset of $\mathbb{R}^d$ and whose generator applied to a function f on $\mathcal{Z}^N$ gives :

$$\sum_{i=1}^{N} \int_{(\mathcal{Z}-\mathcal{Z}) \setminus \{0\}} \{f(\dots, z_{i-1}, z_i + \Delta, z_{i+1}, \dots) \tag{3.2}$$
$$- f(\dots, z_{i-1}, z_i, z_{i+1}, \dots)\} \mathcal{L}(z_i, s_i^N, \frac{1}{N} \sum_{j=1}^{N} \delta_{(z_j, s_j^N)})(d\Delta)$$

Here, $\{\mathcal{L}(z, s, u)(\cdot); z \in \mathcal{Z}, s \in \mathcal{S}, u \in \Pi(\mathcal{Z} \times \mathcal{S})\}$ is a Lévy kernel on the set of jumps $(\mathcal{Z} - \mathcal{Z}) \setminus \{0\}$.

The epidemic system fits this model with

$$\mathcal{S} = \{a,b\}, \mathcal{Z} = \mathbb{R}^2 \times \{S,I,D\}, z = (y,e) \qquad \text{and} \tag{3.3}$$

$$\mathcal{L}\left((y,e),s,u;d(\Delta_y,\Delta_e)\right) = 1\!\!1_{\{e\neq D\}}\mathcal{L}_Y\left(y,d\Delta_y\right)\delta_0(d\Delta_e) + \delta_0(d\Delta_y)\mathcal{L}_E\left((y,e),s,u;d\Delta_e\right)$$

The Lévy kernel $\mathcal{L}(z_i, s_j^N, \frac{1}{N}\sum_{i=1}^N \delta_{(z_j, s_j^N)})$ drives the evolution of the particle i. As it depends on the whole system $(z_j, s_j^N)_{1\leq j\leq N}$, there is an interaction. Since this dependence occurs only via the empirical measure $\frac{1}{N}\sum_{j=1}^N \delta_{(z_j, s_j^N)})$, the interaction is said to be a long range (or a mean field) one.

We want the limit result as N tends to infinity for the empirical measure

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N, s_i^N)} \in \Pi\left(D([0,T],\mathcal{Z})\times\mathcal{S}\right)$$

The space $D = D([0,T],\mathcal{Z})$ is endowed with its usual Skorokhod topology (see *P. Billingsley (1968)*). Let us assume the following hypotheses.

Hypotheses (H)

$\mathbf{H_1}$ There is a bounded nonnegative measure Λ on $(\mathcal{Z}-\mathcal{Z})\setminus\{0\}$ such that

a) for all $z\in\mathcal{Z}, s\in\mathcal{S}$ and $u\in\Pi(\mathcal{Z}\times\mathcal{S})$, $\mathcal{L}(z,s,u)(\cdot)$ is absolutely continuous with respect to $\Lambda(\cdot)$.

b) there is a unique Markov process with generator

$$g(\cdot) \longrightarrow \int_{\mathcal{Z}-\mathcal{Z}}\{g(\cdot+\Delta)-g(\cdot)\}\Lambda(d\Delta)$$

$\mathbf{H_2}$ There is a version of $\dfrac{d\mathcal{L}}{d\Lambda}$ such that

a) $(z,s,u,\Delta)\rightarrow \dfrac{d\mathcal{L}(z,s,u)}{d\Lambda}(\Delta)$ is bounded and continuous on $\mathcal{Z}\times\mathcal{S}\times\Pi(\mathcal{Z}\times\mathcal{S})\times((\mathcal{Z}-\mathcal{Z})\setminus\{0\})$

b) $\{\dfrac{d\mathcal{L}(\cdot,\cdot,\cdot)}{d\Lambda}(\Delta);\Delta\in(\mathcal{Z}-\mathcal{Z})\setminus\{0\}\}$ is an equi-uniformly continuous family of functions on $\mathcal{Z}\times\mathcal{S}\times\Pi(\mathcal{Z}\times\mathcal{S})$.

4. A LAW OF LARGE NUMBERS FOR THE GENERAL LONG RANGE INTERACTING PARTICLE SYSTEM

A. The law of large numbers. In this section we give a result looking like (1.1), for the system which is described by (3.1) and (3.2). We shall add to the hypotheses (H), the following assumption on the initial condition :

(4.1)

for m-almost every s in $\mathcal{S}$, there exists μ_0^s in $\Pi(\mathcal{Z})$, such that almost surely in ω:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N(\omega,t=0), s_i^N)}(dz,ds) = \mu_0^s(dz)m(ds) \quad \text{in } \Pi(\mathcal{Z}\times\mathcal{S}) \quad .$$

Notice that m is given by (3.1) and one can interpret μ_0^s as a conditional version of the probability measure $\mu_0(dz,ds) = \mu_0^s(dz)m(ds)$ on $\mathcal{Z}\times\mathcal{S}$, knowing s.

In what follows, viewing a probability measure v on $D \times \mathcal{S}$ as the law of a $D \times \mathcal{S}$-valued random vector (X, S),

- $v_{\mathcal{S}}$ stands for the marginal law of S on $\mathcal{S}$
- v_t is the law of (X_t, S) on $\mathcal{Z} \times \mathcal{S}$, $\quad (0 \leq t \leq T)$
- v_t^s denotes a regular version of the conditional law on $\mathcal{Z}$ of X_t knowing that $S = s$, $\quad (0 \leq t \leq T, s \in \mathcal{S})$
- v^s denotes a regular version of the conditional law on D of (X, S), knowing that $S = s, (s \in \mathcal{S})$.

PROPOSITION 4.1. *Let us assume (H) and* (4.1).*Then, almost surely in ω, the sequence $\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N(\omega,\cdot),s_i^N)}\right)_{N\geq 1}$ is relatively compact in $\Pi(D \times \mathcal{S})$.*
Let $\mu \in \Pi(D \times \mathcal{S})$ be any limit point of this sequence, then its marginal on $\mathcal{S}$ is $\mu_{\mathcal{S}}(ds) = m(ds)$ and for m-almost every s in $\mathcal{S}$, its conditional version on D knowing s is the law of a nonhomogeneous Markov process on $\mathcal{Z}$, with the initial condition μ_0^s and the family of generators $(G^s_{\mu_t}; 0 \leq t \leq T)$, defined by

$$\left(G^s_{\mu_t}(g)\right)(z) = \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(z+\Delta) - g(z)\}\, \mathcal{L}(z, s, \mu_t)(d\Delta)$$

for any continuous function g with compact support in $\mathcal{Z}$ and any $0 \leq t \leq T$.
In the above formula, μ_t is the marginal at time t, on $\mathcal{Z} \times \mathcal{S}$, of the limit point μ.

If $\mathcal{S}$ has a unique element (then it is no need to consider this set) and if $(X_t)_{0\leq t\leq T}$ stands for the Markov process on $\mathcal{Z}$ whose law is μ, the transition kernel $P(X_{t+dt} \in dy \mid X_t = x)$ depends on t through the law μ_t of X_t. The flow $(t \to \text{Law}(X_t); 0 \leq t \leq T)$ is a solution of the nonlinear Kolmogorov equation

$$\begin{aligned}&\frac{d}{dt}\int_{\mathcal{Z}} g(z)\text{Law}(X_t)(dz)\\ (4.2)\qquad &= \int_{\mathcal{Z}}\left(\int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(z+\Delta) - g(z)\}\, \mathcal{L}\left(z, \text{Law}(X_t)\right)(d\Delta)\right) \text{Law}(X_t)\,(dz)\end{aligned}$$

for any continuous function g with compact support in $\mathcal{Z}$.

To be more precise, the law μ of X is such that,

$$(4.3)\qquad g(X_t) - g(X_0) - \int_0^t du \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(X_u+\Delta) - g(X_u)\}\, \mathcal{L}\left(X_u, \mu_u\right)(d\Delta)$$

is a μ-martingale for any continuous function g with compact support in $\mathcal{Z}$.

Now, if $\mathcal{S}$ has several elements, denoting $(X_t^s)_{0\leq t\leq T}$ the Markov process on $\mathcal{Z}$ whose law is μ^s, for m-almost every s in $\mathcal{S}$, $P(X^s_{t+dt} \in dy \mid X_t = s)$ depends on t via the set of laws $\mu_t^{s'} = \mathcal{L}\text{aw}(X_t^{s'})$ as s' describes $\mathcal{S}$. As in (4.3), the law μ of X is such that
(4.4)

$$g(X_t^s) - g(X_0^s) - \int_0^t dt \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(X_u^s+\Delta) - g(X_u^s)\}\mathcal{L}\left(X_u^s, s, \mu_u^{s'}(dz')m(ds')\right)(d\Delta)$$

is a μ^s-martingale, for any function g with a compact support in $\mathcal{Z}$.

In *H.P.Mc Kean (1967)* introduced a class of Markov processes whose evolution equations are similar to (4.3). He called them nonlinear Markov processes. Later, A.S.Sznitman studied in [Szn] the corresponding nonlinear martingale problem, which is (4.3) for our model. Clearly, if the nonlinear martingale problem (4.4) admits a unique solution in $\Pi(D \times \mathcal{S})$, then the sequence $\left(\frac{1}{N}\sum_{i=1}^N \delta_{(X_i^N, s_i^N)}\right)_{N\geq 1}$ converges to this solution in $\Pi(D \times \mathcal{S})$. We shall need additional regularity assumptions to get this uniqueness result. In order to obtain it, let us assume that:

(4.5)
for any $s \in \mathcal{S}, u \in \Pi(\mathcal{Z}\times\mathcal{S})$ and for any continuous bounded function f on $(\mathcal{Z}-\mathcal{Z})\setminus\{0\}$,

$$\text{the function : } \quad z \longrightarrow \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} f(\Delta)\,\mathcal{L}(z,s,u)(d\Delta) \quad \text{is Lipschitz on } \mathcal{Z}$$

(4.6)
for any $z \in \mathcal{Z}, s \in \mathcal{S}$ and for any continuous bounded function f on $(\mathcal{Z} - \mathcal{Z}) \setminus \{0\}$,

$$\text{the function : } \quad u \longrightarrow \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} f(\Delta)\,\mathcal{L}(z,s,u)(d\Delta) \quad \text{is Lipschitz on } \Pi(\mathcal{Z} \times \mathcal{S}).$$

In (4.6), $\Pi(\mathcal{Z} \times \mathcal{S})$ is endowed with the Wasserstein metric. For the definition of this metric, see *A.S. Szitman (1984)* for instance. Let us just recall that this metric yields the usual $\sigma(\Pi, C_b)$-topology.

PROPOSITION 4.2 (STRONG LAW OF LARGE NUMBERS). *Let us assume the hypotheses (H), (4.1), (4.5) and (4.6). If μ_0 satisfies $\int_{\mathcal{Z}\times\mathcal{S}} |z|^2 \mu_0^s(dz)m(ds) < +\infty$, then there exists a unique probability measure μ on $D \times \mathcal{S}$ satisfying the properties of proposition 4.1. More,*

$$(4.7) \qquad \textit{almost surely in } \omega, \qquad \lim_{N\to\infty} \sum_{i=1}^N \delta_{(X_i^N(\omega,\cdot), s_i^N)} = \mu \qquad \textit{in } \Pi(D \times \mathcal{S})$$

The limit of proposition 4.2 is often called Mc Kean - Vlasov limit. Several authors have already studied it in different contexts. Some of them are: *H.P. Mc Kean (1967), D.A. Dawson (1984), A.S. Sznitman (1984), K. Oelschläger (1984)* or *C. Léonard (1986)*. In all these papers, the law of large numbers is stated in terms of convergence in law (weak law of large numbers). It is well-known that the almost sure convergence result (strong law of large numbers) is a natural consequence of the large deviations (theorem 5.1 below) and of Borel-Cantelli lemma. For a proof of this improvement, see for instance *C. Léonard (1987)*.

B. The propagation of chaos. In a different context, *M. Kac (1967)* introduced the propagation of chaos. Let us illustrate this notion. Suppose that the initial condition is an independtly distributed sequence, whose law is

$$(4.8) \qquad \text{Law}\left((X_i^N(t=0))_{1\leq i\leq N}\right) = \otimes_{i=1}^N \mu^{s_i^N} \in \Pi(\mathcal{Z}^N).$$

Because of the interaction, for any $t > 0$, $(X_i^N(t))_{1\leq i\leq N}$ becomes correlated. Nevertheless, in the limit $N \to \infty$, the independence propagates at any positive times. This property, whose rigorous statement is given in the next proposition, is called the propagation of chaos.

PROPOSITION 4.3 (PROPAGATION OF CHAOS). *Let us assume* (4.8) *and the hypotheses of proposition 4.2. For any $k \geq 1$, let us choose $(s_1, \ldots, s_k)$ in $\mathcal{S}^k$ and let us assume that*

$$s_j^N = s_j, \quad \text{for any } 1 \leq j \leq k \text{ and any } N \geq k. \tag{4.9}$$

Then,

$$\lim_{N\to\infty} Law(X_1^N, \ldots, X_k^N) = \otimes_{j=1}^k \mu^{s_j} \qquad \text{in } \Pi(D^k).$$

Notice that (4.9) does not contradict (3.1). Proposition 4.3 means that if k tagged particles are independently distributed at the initial time, then they will move "almost" independently from each others provided that they live in a "large enough" system. It implies, in particular, that

for any $t > 0$ and any $f_1, \ldots, f_k$ continuous bounded functions on $\mathcal{Z}$, then

$$\lim_{N\to\infty} E\left(\prod_{j=1}^k f_j(X_j^N(t))\right) = \prod_{j=1}^k Ef_j(X_j^\infty(t))$$

where the law of $X_j^\infty(t)$ is $\mu_t^{s_j}$.

C. **Back to the epidemic system.** Let us consider again the epidemic system which was described in section 2. It fits the general interacting system of section 3 with $\mathcal{S}, \mathcal{Z}$ and $\mathcal{L}$ given by (3.3). In order that the assumptions of our previous results are satisfied, one should assume that:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N 1\!\!I_{s_i^N=a} = m_a \quad , \quad \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N 1\!\!I_{s_i^N=b} = m_b$$

(hence $m = m_a\delta_a + m_b\delta_b \in \Pi(\mathcal{S})$ with $m_a + m_b = 1$) and that the function $J(y)$ appearing in (2.1) is a continuous bounded function on $\mathbb{R}^2$; for instance, a smooth approximation of $1\!\!I_{\{\|y\|<R\}}$.

If all the individuals are independently distributed at time $t = 0$ with the common law μ_0^a (resp. μ_0^b) on $\mathbb{R}^2 \times \{S, I, D\}$ for the individuals of type a (resp. type b), then by proposition 4.2 :

almost surely in ω,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \delta_{(X_i^N(\omega,\cdot),s_i^N)} = \mu \text{ in } \Pi\left(D([0,T], \mathbb{R}^2 \times \{S,I,D\}) \times \{a,b\}\right)$$

Here, $\mu(dx, ds) = m_a\,\mu^a(dx)\delta_a(ds) + m_b\,\mu^b(dx)\delta_b(ds)$, μ_a and μ_b being the unique Markov laws on $D([0,T], \mathbb{R}^2 \times \{S, I, D\})$ such that $t \to (\mu_t^a, \mu_t^b)$ is a solution of the <u>nonlinear</u> differential equations

$$
\begin{aligned}
&\frac{d}{dt}\sum_{e\in\{S,I,D\}}\int_{\mathbb{R}^2} g(y,e)\,\mu_t^a(dy\times\{e\}) = \\
&\qquad \int_{\mathbb{R}^2}\{g(y,I)-g(y,S)\}\,C_{SI}(y,a,m_a\mu_t^a\otimes\delta_a+m_b\mu_t^b\otimes\delta_b)m_a\,\mu_t^a(dy\times\{S\}) \\
&\qquad +\int_{\mathbb{R}^2}\{g(y,S)-g(y,I)\}\,C_{IS}\,m_a\,\mu_t^a(dy\times\{I\}) \\
&\qquad +\int_{\mathbb{R}^2}\{g(y,D)-g(y,I)\}\,C_{ID}\,m_a\,\mu_t^a(dy\times\{I\}) \\
&\qquad +\sum_{e\in\{S,I\}}\int_{\mathbb{R}^2}\left(\int_{\mathbb{R}^2_*}\{g(y+\Delta_y,e)-g(y,e)\}\mathcal{L}_Y(y,d\Delta_y)\right) m_a\,\mu_t^a(dy\times\{e\})
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{d}{dt}\sum_{e\in\{S,I,D\}}\int_{\mathbb{R}^2} h(y,e)\,\mu_t^b(dy\times\{e\}) = \\
&\qquad \int_{\mathbb{R}^2}\{h(y,I)-h(y,S)\}\,C_{SI}(y,a,m_a\mu_t^a\otimes\delta_a+m_b\mu_t^b\otimes\delta_b)m_b\,\mu_t^b(dy\times\{S\}) \\
&\qquad +\int_{\mathbb{R}^2}\{h(y,S)-h(y,I)\}\,C_{IS}\,m_b\,\mu_t^b(dy\times\{I\}) \\
&\qquad +\int_{\mathbb{R}^2}\{h(y,D)-h(y,I)\}\,C_{ID}\,m_b\,\mu_t^b(dy\times\{I\}) \\
&\qquad +\sum_{e\in\{S,I\}}\int_{\mathbb{R}^2}\left(\int_{\mathbb{R}^2_*}\{h(y+\Delta_y,e)-h(y,e)\}\mathcal{L}_Y(y,d\Delta_y)\right) m_b\,\mu_t^b(dy\times\{e\})
\end{aligned}
$$

for any continuous bounded functions g and h on $\mathbb{R}^2\times\{S,I,D\}$.

5. Large deviations for the general long range interacting particle system

All the results of this section have been proved in *C. Léonard (1989)* Analogous results have been obtained by *F. Comets (1987)* for the long range Ising model, and by *D.A. Dawson and J. Gärtner (1987)* for weakly interacting diffusion systems. The proofs of *F. Comets (1987)*, *D.A. Dawson and J. Gärtner (1987)* and *C. Léonard (1989)* are quite different from each others. One should have a look at the paper written by *D.A. Dawson and J. Gärtner (1989)*, on this subject.

For general features about large deviations, one should look at the courses of *R. Azencott (1980)* and *S.R.S. Varadhan (1989)*.

In theorem 5.1, one gives a rigorous statement of (1.2) , then one expresses the function $I(\cdot)$ at theorems 5.2 and 5.3.

A. Natural questions involving large deviations. It has been seen in the previous section that, whenever N is large and the initial distribution satisfies (4.1), the empirical measure $\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N(\omega,\cdot),s_i^N)}$ tends, almost surely in ω, to the probability measure μ on $D\times\mathcal{S}$ which is given by propositions 4.1 and 4.2 . The speed of convergence is made precise by a result of type (1.2).

Let us suppose that the limiting measure μ on $D\times\mathcal{S}$ is such that its family of marginals $\mu_t\in\Pi(\mathcal{Z}\times\mathcal{S})$ for all times t, tends to μ_∞ in $\Pi(\mathcal{Z}\times\mathcal{S})$, as t tends to infinity. One would like to get an estimate,for large N and large T, for the probability of the event

" $\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N(T),s_i^N)}$ belongs to the neighbourhood $\mathcal{N}(\nu)$ of a probability measure ν on $\mathcal{Z}\times\mathcal{S}$ " (ν may differ from the asymptotic state μ_∞). A rough answer to this question is given by (1.2), which can be written as follows:

$$P\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N,s_i^N)}\in A\right)\asymp\exp\left(-N\inf\{I(v);v\in A\}\right)\qquad\text{, as } N \text{ tends to infinity}$$

with

$$A=\{v\in\Pi(D\times\mathcal{S})\text{ such that }v_0\in\mathcal{N}(\mu_0)\text{ and }v_T\in\mathcal{N}(\nu)\}\,. \tag{5.1}$$

Among all the paths leading from $\mathcal{N}(\mu_0)$ to $\mathcal{N}(\nu)$ during the time interval [0,T], the most probable ones , as N tends to infinity, are those $\hat{v}$ in $\Pi(D\times\mathcal{S})$ such that:

$$I(\hat{v})=\inf\{I(v);v\in A\}\quad\text{and}\quad\hat{v}\in A \tag{5.2}$$

where A is given by (5.1). Indeed, if $\tilde{v}\in A$ and $I(\tilde{v})=I(\hat{v})+\varepsilon$ $(\varepsilon>0)$, then:

$$\frac{P\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N,s_i^N)}\in\mathcal{N}(\hat{v})\right)}{P\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N,s_i^N)}\in\mathcal{N}(\tilde{v})\right)}\asymp\exp\left(-N\left(I(\hat{v})-I(\tilde{v})\right)\right),\qquad\text{as } N \text{ tends to infinity}$$
$$=\exp(N\varepsilon),\qquad\text{which tends to infinity with } N.$$

This leads us to the optimization problem (5.1) & (5.2) .

B. The large deviation principle. In this subsection, we give a rigorous statement of (1.2) in the following theorem.

THEOREM 5.1. *Let us assume hypotheses (H) and (4.1). The sequence of random variables* $\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N,s_i^N)}\right)_{N\geq1}$ *obeys a large deviation principle in the space* $\Pi(D\times\mathcal{S})$ *endowed with its natural topology.*
This means that one can find a function $I(\cdot)$ *on* $\Pi(D\times\mathcal{S})$ *with values in* $\mathbb{R}_+\cup\{+\infty\}$*, such that:*
for any closed subset C *of* $\Pi(D\times\mathcal{S})$*,*

$$\limsup_{N\to\infty}\frac{1}{N}\log P\left(\omega;\frac{1}{N}\sum_{i=1}^{N}\delta_{(X_i^N(\omega,\cdot),s_i^N)}\in C\right)\leq-\inf\{I(v);v\in C\}$$

and for any open subset G of $\Pi(D \times \mathcal{S})$,

$$\liminf_{N\to\infty} \frac{1}{N} \log P\left(\omega; \frac{1}{N}\sum_{i=1}^{N} \delta_{(X_i^N(\omega,\cdot),s_i^N)} \in G\right) \geq -\inf\{I(v); v \in G\}.$$

If the subset A of $\Pi(D \times \mathcal{S})$ is "regular for $I(\cdot)$" , in the sense that:

$$\inf\{I(v); v \text{ in the interior of } A\} = \inf\{I(v); v \text{ in the closure of } A\}$$

then:
$$\lim_{N\to\infty} \frac{1}{N} \log P\left(\omega; \frac{1}{N}\sum_{i=1}^{N} \delta_{(X_i^N(\omega,\cdot),s_i^N)} \in A\right) = -\inf\{I(v); v \in A\},$$

which is (1.2) .

C. The rate function $I(\cdot)$**.** The following result states that if $I(v)$ is finite, then $t \to v_t$ is an "absolutely continuous" path in the sense of (5.4) below.

THEOREM 5.2. *Let us assume the hypotheses (H) and* (4.1). *Let v be a probability measure on $D \times \mathcal{S}$ such that*

$$I(v) < +\infty$$

then, it satisfies the following properties.

- (5.3) *The marginal law $v_{\mathcal{S}}$ on $\mathcal{S}$ is $v_{\mathcal{S}} = m$ (see* (4.1)*)* .
- (5.4) *One can find a Lévy kernel $\{\mathcal{L}_v(z,s,t)(\cdot); z \in \mathcal{Z}, s \in \mathcal{S}, 0 \leq t \leq T\}$ such that for m-almost every s in $\mathcal{S}$ and for any continuous function g with compact support in $\mathcal{Z}$, the function $t \to \int_{\mathcal{Z}} g(z)\, v_t^s(dz)$ is absolutely continuous and*

$$\frac{d}{dt}\int_{\mathcal{Z}} g(z)\, v_t^s(dz) = \int_{\mathcal{Z}}\left(\int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(z+\Delta)-g(z)\}\mathcal{L}_v(z,s,t)(d\Delta)\right) v_t^s(dz)$$

The evolution equation (5.4) *is a Kolmogorov equation, but this does not imply that v^s is a Markov law. Nevertheless,*

- *for m-almost every s in $\mathcal{S}$, one can find a unique probability measure v^{*s} on D which is a Markov law that satisfies* (5.4) *and which is absolutely continuous with respect to the law of the Markov process with generator : $g(\cdot) \to \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(\cdot+\Delta)-g(\cdot)\}\, \Lambda(d\Delta)$. ($\Lambda$ is defined in hypothesis H_1).*

<u>The markovian projection</u>. It is natural to define the "markovian projection" v^* of v, for any $v \in \Pi(D \times \mathcal{S})$ such that $I(v) < +\infty$, by:

$$v^*(dx,ds) = v^{*s}(dx)\, m(ds) \qquad ; v^* \in \Pi(D \times \mathcal{S})$$

where v^{*s} is given by theorem 5.2 .

In order to express $I(v)$ in theorem 5.3 below, we need a generalization of the definition of Kullback information.

Kullback information between two nonnegative measures. Let α and β be two nonnegative measures on a measurable space Y. The Kullback information $K(\beta,\alpha)$ of the measure β with respect to the measure α is defined by:

$$K(\beta,\alpha)=\begin{cases}\int_Y\{\frac{d\beta}{d\alpha}(y)\log\frac{d\beta}{d\alpha}(y)-\frac{d\beta}{d\alpha}(y)+1\}\,\alpha(dy) & \text{if } \beta\ll\alpha\\ +\infty & \text{otherwise}\end{cases} .$$

It is easy to chek that:

$$\begin{aligned}&K(\beta,\alpha)\geq 0\\ &K(\beta,\alpha)=0 \iff \beta=\alpha\\ &\beta\longrightarrow K(\beta,\alpha) \text{ is a convex function.}\end{aligned}$$

Notice that if both α and β are probability measures, the above definition of $K(\beta,\alpha)$ matches with the classical one:

$$K(\beta,\alpha)=\begin{cases}\int_Y\log\frac{d\beta}{d\alpha}(y)\,\beta(dy) & \text{if } \beta\ll\alpha\\ +\infty & \text{otherwise}\end{cases} .$$

We are now ready to give an expression for $I(v)$.

THEOREM 5.3. *Let us assume the hypotheses (H),* (4.1) *and that the initial condition* $\left\{(X_i^N(t=0),s_i^N)_{1\leq i\leq N}\,;N\geq 1\right\}$ *obeys a large deviation principle in the topological space* $\Pi(\mathcal{Z}\times\mathcal{S})$ *with a rate function* $I_0\colon \Pi(\mathcal{Z}\times\mathcal{S})\to \mathbb{R}_+\cup\{+\infty\}$. *Then, for any probability measure* v *on* $D\times\mathcal{S}$,

$$I(v)=\begin{cases}K(v,v^*)+I(v^*) & \text{if } v \text{ satisfies } (5.3)\ \&\ (5.4)\\ +\infty & \text{otherwise ,}\end{cases} \tag{5.5}$$

with

$$I(v^*)=I_0(v_0)+\int_0^T\left(\int_{\mathcal{Z}\times\mathcal{S}}K\left(\mathcal{L}_v(z,s,t),\mathcal{L}(z,s,v_t)\right)\,v_t^s(dz)\,m(ds)\right)dt \tag{5.6}$$

Let us recall that v^* is the markovian projection of v, $K(\cdot,\cdot)$ is the Kullback information, $\mathcal{L}_v$ comes from (5.4), $\mathcal{L}$ from (3.2), m from (3.1) & (4.1) and I_0 from the above assumption on the initial condition.

D. The variational principle (5.1) & (5.2). It can be proved that, whenever $I(v)<+\infty$,

$$v_t=v_t^*,\qquad \text{for any } 0\leq t\leq T. \tag{5.7}$$

As an indication, $t\to v_t$ and $t\to v_t^*$ both satisfy the same evolution equation (5.4). But (5.4) does not necessarly admit a unique solution and (5.7) is far from being easy to prove. *H. Föllmer (1989)* proved such a result for interacting diffusions. A different proof of (5.7) is given in *C. Léonard (1989)*.

A straightforward consequence of (5.5) and (5.7) is the following corollary.

COROLLARY 5.4. *Under the assumptions of theorem 5.3, for any probability measures ν and λ on $\mathcal{Z} \times \mathcal{S}$:*

$$\begin{aligned} &\inf \{I(v); v \text{ such that } v_0 = \nu \text{ and } v_T = \lambda\} \\ = &\inf \{I(v^*); v \text{ such that } v_0 = \nu \text{ and } v_T = \lambda\} \\ = &\inf \{I(v); v \text{ such that } v_0 = \nu, v_T = \lambda \text{ and } v \text{ is a Markov law. }\}. \end{aligned}$$

Such a result is worth when solving the variational problem (5.1) & (5.2). Indeed, the minimizing solutions are necessarly Markov laws and

$$I(v) = I(v^*) \qquad \text{for any Markov law } v$$

whose explicit form is given by (5.6).

Let us denote $\alpha \in \Pi(D)$, the law of the Markov process with generator: $g(\cdot) \to \int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{g(\cdot + \Delta) - g(\cdot)\}\, \Lambda(d\Delta)$. By theorem 5.2 and Girsanov's formula, one can find a nonnegative measurable function $l(\cdot)$ on $\mathcal{Z} \times \mathcal{S} \times [0,T] \times ((\mathcal{Z} - \mathcal{Z}) \setminus \{0\})$, such that $v^* = \psi(l)$, where

$$v^*(dx, ds) = v^{*s}(dx)m(ds) = \psi(l)^s m(ds) = \psi(l)(dx, ds) \tag{6.1}$$

with

$$\begin{aligned} &\frac{d\psi(l)^s}{d\alpha}(x) = \\ &\exp\left[\sum_{0 \le t \le T x_t \ne x_{t-}} \log l(x_{t-}, s, t^-, x_t - x_{t-}) - \int_0^T \left(\int_{\mathcal{Z}-\mathcal{Z}} \log l(x_t, s, t, \Delta)\, \Lambda(d\Delta)\right) dt\right] \\ &\times \frac{dv_0^s}{d\alpha_0}(x_0) \exp\left[-\int_0^T \left(\int_{\mathcal{Z}-\mathcal{Z}} \{l(x_t, s, t, \Delta) - \log l(x_t, s, t, \Delta) - 1\} \Lambda(d\Delta)\right) dt\right] \end{aligned}$$

This means that v^* is the Markov law with Lévy kernel $\mathcal{L}_v(z,s,t)(d\Delta) = l(z,s,t,\Delta)\, \Lambda(d\Delta)$. Denoting

$$\begin{aligned} \phi(l)(z,s,t,\Delta) &= \frac{d\mathcal{L}_v(z,s,t)}{d\mathcal{L}(z,s,\psi(l)_t)}(\Delta) \\ &= l(z,s,t,\Delta)\left(\frac{d\mathcal{L}(z,s,t,\psi(l)_t)}{d\Lambda}(d\Delta)\right)^{-1} \end{aligned}$$

one can write $I(v^*) = I(\psi(l))$ with

(6.3)

$$I(\psi(l)) = \int_0^T dt \int_{\mathcal{Z}\times\mathcal{S}} \left(\int_{(\mathcal{Z}-\mathcal{Z})\setminus\{0\}} \{\phi(l)\log(\phi(l)) - \phi(l) + 1\}(z,s,t,\Delta)\, \mathcal{L}(z,s,\psi(l)_t)(d\Delta)\right) \psi(l)_t^s\,(dz)\, m(ds) dt$$

Hence, to solve (5.1) & (5.2), one should consider the above function l on $\mathcal{Z} \times \mathcal{S} \times [0,T] \times ((\mathcal{Z} - \mathcal{Z}) \setminus \{0\})$ as a new parameter instead of v^*, with the transformation (6.1) & (6.2). The problem is now to minimize $l(\cdot) \longrightarrow I(\psi(l(\cdot)))$ given by (6.3). Bon courage.

References

R. Azencott, *Grandes déviations et applications, Saint-Flour VII, 1978*, Lecture Notes in Mathematics **774** (1980), Springer-Verlag.

P. Billingsley, "Convergence of Probability Measures," J. Wiley and Sons, New-York, 1968.

F. Comets, *Nucleation for a long range magnetic model*, Ann. Inst. H. Poincaré **23, N.2** (1987), 135–178.

D.A. Dawson, *Measure-valued processes. Construction, qualitative behavior and stochastic geometry.*, Lecture Notes in Mathematics **1212** (1984), Springer-Verlag.

D.A. Dawson, *Critical dynamics and fluctuations for a mean-field model of cooperative behavior.*, Journal of Statistical Physics **31, N.1** (1983), 29–85.

D.A. Dawson and J. Gärtner, *Large deviations from the Mc Kean-Vlasov limit for weakly interacting diffusions.*, Stochastics **20** (1987), 247–308.

D.A. Dawson and J. Gärtner, *Large deviations, free energy funtional and quasi-potential for a mean field model of interating diffusions.*, Memoirs of the A.M.S. **78, N.398** (1989).

H. Föllmer, *Random fields and diffusion processes. Saint-Flour XV-XVII, 1985-1987*, Lecture Notes in Mathematics **1362** (1989), Springer-Verlag.

M. Kac, "Probability and related topics in physical sciences.," Interscience, New-York, 1958.

C. Léonard, *Une loi des grands nombres pour des systèmes de diffusions avec interaction et à coefficients non bornés.*, Ann. Inst. H. Poincaré **22, N.2** (1986), 237–262.

C. Léonard, *Large deviations and law of large numbers for a mean field type interacting particle system.*, Stochastic Processes and their Appl. **25, N.2** (1987), 215–235.

C. Léonard, *Large deviations and Orlicz spaces. Long range interacting Markov processes with jumps.*, Prébublications (1989), Université d' Orsay.

H.P. Mc Kean, Jr, *Propagation of chaos for a class of nonlinear parabolic equations*, Lecture Series in Differential Equations **t. 7** (1967), Catholic University (Washington D. C.).

K. Oelschläger, *A martingale approach to the law of large numbers for weakly interacting stochastic processes.*, The Annals of Proba. **t. 12, N.22** (1984), 458–479.

A.S. Sznitman, *Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated.*, Journal of Functional Analysis **t. 56** (1984), 311–336.

S.R.S. Varadhan, *Large deviations. Saint-Flour XV-XVII, 1985-1987*, Lecture Notes in Mathematics **1362** (1989), Springer-Verlag.

Bâtiment 425. Université de Paris-Sud . 91405 Orsay Cedex. France.
Laboratoire de mathématiques de l'Université de Paris-Sud.
Equipe de Statistiques. U.A. 743.