Università degli Studi di Trento
Dipartimento di Matematica

Dottorato di Ricerca in Matematica

Ph.D. Thesis

Deterministic Approximation of Stochastic Metapopulation Models

Francesca Arrigoni

Supervisor

Prof. A. Pugliese

Gennaio 2001

ADDRESS: Via Sommarive, 14 38050 POVO - Trento
TEL:+39-461-881513.
FAX: +39-461-881624.
E-MAIL: arrigoni@science.unitn.it

Acknowledgments

I wish to thank my advisor, Professor A. Pugliese, who introduced me to Biomathematics, Professor M. Iannelli, and Professor L.Tubaro, for many fruitful discussions about the interplay between the probabilistic and analytical approaches to study models arising from biological problems.

I am very grateful to Professor A.D. Barbour, for his suggestions and for careful reading of the thesis.

I wish to thank Lorenza, who shared with me the Ph.D. experience and the 'adventures' at 'via Fiume 16' , my friends and colleagues at the Mathematical Department in Povo (Beppe, Claudio, Davide, Elisa, Gippo, Roberto, Sonia, Vincenzo and all the others), and my friends Paola, Annamaria, Chiara, Elena, Mara, Silvia, Cristina, who have been close to me, in spite of the distance between Trento and Milano.

I would also like to express my thanks to Amerigo, Miriam and the administrative and technical staff of the Department.

Finally, a very special thank to my sister Serena and my parents, who have always encouraged, supported and loved me.

Contents

Introduction

Most population models usually assume that all individuals of the population live in the same habitat and interact homogeneously with each other, but natural populations often have a more complex structure : in this thesis, we deal with metapopulation models.
The term *metapopulation* was introduced in the works of Levins in 1969 and 1970 ([Lev69], [Lev70]) to denote a population, which is divided into local subpopulations (in the literature, these local entities are called patches).
In table 1, we show some characteristics of mathematical models for metapopulation dynamics (see also [met97]).

Model	**Number of habitat patches**	**Variables**	**Mathematical tools**
Levins' model	infinite no localized interactions	p fraction of occupied sites	ODE
Deterministic structured models	infinite no localized interactions	p_i fraction of sites with i-individuals or p(x) density of size x	system of ODEs or PDE
Spatially explicit Approaches	infinite interaction only among nearby cells	a vector of presence and absence	lattice model cellular automata
Stochastic models	finite	x_i fraction of sites with i individuals or vector of presence and absence	Markov chains

Table 1: Metapopulation model

We are actually interested in the relationship between deterministic and stochastic models, focusing especially on the structured ones: before introducing our models, we give a brief description of the models that appear in Table 2

	Deterministic	Stochastic
Unstructured	Levins	Gyllenberg & Silvestrov
Structured	Val et al. Gyllenberg & Hanski	Metz Ch.2, Ch3

Table 2: Deterministic versus stochastic

Levins' model

Levins' metapopulation model makes the following assumptions:

A) the number of patches is very large (effectively infinite);

B) all patches are either occupied or empty and have the same size and quality;

C) the exchange rate of individuals among local populations is so low that migration has no real effect on local dynamics;

D) spatial arrangement of patches is ignored and all patches communicate equally with all others.

This approach leads to an ordinary differential equation

$$\frac{dP(t)}{dt} = mP(1-P) - eP \tag{1}$$

for the scalar variable $P(t)$, which is the fraction of patches occupied at time t, with m and e colonization and extinction parameters respectively. The equation (1) is formally the same as an $S-I-S$ epidemic model, if we identify the occupied patches of the metapopulation with the infective individuals and the empty patches with the susceptibles. Also, in an epidemiological framework, it is often assumed that there is homogeneous mixing of a large number of individuals, and that all infective individuals are equally infectious. These assumptions can be satisfactory for a disease spread in a population of freely moving individuals, but are not completely satisfactory for local populations with a fixed spatial location.

Deterministic structured models

To relax assumption B), structured metapopulation models have been suggested: in Levin and Paine [LP74],[LP] the structure variables are age and size, in Hastings and Wolin [HW89] the variable is age, but neither model takes into account the effect of migration on local dynamics (they keep assumption C)).
The model proposed in Gyllenberg- Hanski [GH92] gives up B) and C) and describes explicitly the within-patch consequences of migration, introducing the dispersers. To build their model, they implicitly assume that a minimum k of inhabitants per patch is required for it to be considered occupied, and they take $\varepsilon = \frac{1}{k}$ as a parameter in a heuristic model, where the local population size is the structuring variable and it is labelled by x (we can interpret x as the number of individuals, rescaled by k); the population in a patch is considered large enough that its growth can be described by a deterministic model

$$\frac{dx(t)}{dt} = g(x),$$

where g is a known function that governs the local dynamics. The migration is modeled by introducing a migration rate $\gamma_\varepsilon(x)$ and the number $D_\varepsilon(t)$ of dispersers per patch at time t; the local population size can change from

$$\begin{array}{lll} x \to x - \varepsilon & \text{at rate } \gamma_\varepsilon(x) & \text{due to emigration} \\ x \to x + \varepsilon & \text{at rate } \alpha D_\varepsilon(t) & \text{due to immigration of dispersers;} \end{array}$$

the dispersers that arrive in an empty patch colonize it with probability ρ_ε, and the patch population instantaneously reaches the critical size $1-\varepsilon$.
Let $P_\varepsilon(t)$ be the proportion of occupied patches, expressed by $\int_{1-\varepsilon}^{\infty} p_\varepsilon(t,x)dx$, and let $E_\varepsilon(t) = 1 - P_\varepsilon(t)$ be the proportion of empty patches.
Their eventual PDE model describes the evolution of the normalized patch-size distribution $p(t,x)$, which is obtained as a limit of $p_\varepsilon(t,x)$ as $\varepsilon \downarrow 0$. Assuming the

existence of the following limits

$$\begin{aligned}\gamma(x) :& = \lim_{\varepsilon\to 0} \varepsilon\gamma_\varepsilon\\ \rho :& = \lim_{\varepsilon\to 0} \frac{\rho_\varepsilon}{\varepsilon}\\ D(t) :& = \lim_{\varepsilon\to 0} \varepsilon D_\varepsilon(t)\\ p(t,x) :& = \lim_{\varepsilon\to 0} p_\varepsilon(t,x)\\ E(t) :& = \lim_{\varepsilon\to 0} E_\varepsilon(t),\end{aligned}$$

they can write down the following set of differential equations

$$\begin{aligned}\partial_t p(t,x) + \partial_x(v(t,x)p(t,x)) &= -\mu(x)p(t,x)\\ \frac{dD(t)}{dt} &= -(\alpha+\nu)D(t) + \int_1^\infty \gamma(x)p(t,x)dx\\ v(t,1)p(t,1) &= \rho\alpha D(t)\left(1 - \int_1^\infty p(t,x)\right)\\ p(0,x) = p_0(x) & \qquad D(0) = D_0\end{aligned} \tag{2}$$

where $\int_1^\infty p(t,x)dx$ represents the proportion of occupied patches, $\mu(x)$ is the extinction rate of a local population of size x, ν is the death rate of dispersers and α is the rate at which the dispersers arrive at a patch.

The model proposed in Val *et al.* [VVM96] for a structured metapopulation is very similar to the previous one: it is a deterministic model with the size of a patch as a structuring variable and with migration that influences local dynamics. Here no critical size for the patch is assumed (in [GH92] the critical size is 1), so that the proportion of occupied patches $P(t)$ can be written as $\int_0^\infty p(t,x)dx$; moreover, partial disasters are allowed, causing so that local size can change from $y \to x > 0$ with probability $\varphi_1(x,y)$ and from $y \to 0$ with probability $\varphi_0(y)$ under the condition $\varphi_0(y) + \int_0^y \varphi_1(x,y)dx = 1$.

The evolution equations for this model are

$$\begin{aligned}\partial_t p(t,x) \quad + \quad \partial_x(v(t,x)p(t,x)) &= -\mu(x)p(t,x) + \int_0^\infty \varphi_1(x,z)\mu(z)p(t,z)dz\\ \frac{dD(t)}{dt} &= -(\alpha+\nu)D(t) + \int_0^\infty \gamma(x)p(t,x)dx\\ v(t,0)p(t,0) &= \theta\left(1 - \int_0^\infty p(t,x)\right) + (1-u)\int_0^\infty \varphi_0(x)\mu(x)p(t,x)dx,\\ p(0,x) &= p_0(x) \quad D(0) = D_0\end{aligned} \tag{3}$$

where u is the probability that, after a total disaster, a patch becomes unsuitable for habitation and $\frac{1}{\theta}$ is the mean time to regain habitat quality after a disaster.

The model of [GH92] is essentially different from this one in the boundary condition (see (2) in comparison with (3)): if $u = 0$, there are no empty patches, because, after a total disaster, patches get suddenly colonized; even if we set $u = 1$ and $\varphi_0(x) \equiv 1$, (3) covers (2)-case, only when the size of the instream of dispersers affects the recovery of habitat, i.e. θ is a function of D.

In those articles, the authors are interested in equilibrium solutions and in their stability: under some assumptions on the function $g - \gamma$ (which show the existence of a carrying capacity of the patch, as a function of the number of dispersers), the models allow alternative stable equilibria, not found in the models that ignore the effect of dispersal on local dynamics. From those structured models, it is also possible to come back to a simplified one, in the spirit of Levins', when the local dynamics is supposed to take place on a much faster time scale than the dynamics at the metapopulation level.

Stochastic models

In reality, we have a finite, though possibly large, number of patches with different characteristics and different local population sizes, which are coupled in a complicated manner. When the number of patches is finite, the model should be stochastic (the dynamic processes are actually discrete and random in their development). A mathematical tool to describe this kind of modelis a stochastic Markov process, as used throughout this thesis.

A stochastic metapopulation model with finite number of patches is proposed by Gyllenberg and Silvestrov [GS94] (see this article for further references about stochastic models) and [Pol99]: it takes the spatial arrangements of patches into account and keeps track of which patches are occupied and which are empty. The time-evolution of the system is described by a time-discrete Markov chain $\bar{\eta}(t) = (\eta_1(t), \ldots, \eta_M(t))$ with finite state space $\bar{X} = \left\{\bar{x} = (x_1, \ldots, x_M) : x_i \in \{0, 1\}\right\}$ (the random variable $\eta_i(t)$ takes the value 1 if the i-th patch is occupied and 0 if the i-th patch is empty at time t). The dynamics is assigned by an interaction matrix $Q = (q_{ji})$: q_{ii} is the probability that, in the absence of migration, the population in patch i will go extinct in one time-step, q_{ji} is the probability that patch i will not be colonized in one time-step by a migrant originating from patch j. Under the assumption that local extinction processes and colonization attempts from different local populations are all independent, the conditional probabilities $q_i(\bar{x})$ for patch i to be empty at time $t + 1$, given the configuration $\bar{x}$ at time t is

$$q_i(\bar{x}) = \prod_{j=1}^{M} q_{ji}^{x_j}$$

Hence the transition matrix for this homogeneous Markov chain is

$$P(\bar{x},\bar{y}) = \prod_{j=1}^{M} q_i(\bar{x})^{1-y_i}(1-q_i(\bar{x}))^{y_i} \qquad \bar{x},\bar{y} \in \bar{X}$$

They are interested in the long term behaviour of the metapopulation : if there is no mainland ($q_{ii} > 0$ for all i), the metapopulation will certainly go extinct, so they study in particular the case with $q_{11} = 0$ and $q_{ii} > 0$ for $i > 1$ (from the metapopulation point of view, the patch labelled by 1 plays the role of the mainland), introducing a perturbation that depends on a small parameter ε which is of the same order as the extinction probability of patch 1 (the 'quasi-mainland'). A perturbed chain $\bar{\eta}^{(\varepsilon)}(t)$ is built up from a perturbed interaction matrix $q_{ji}^{(\varepsilon)} = q_{ji} + \varepsilon\hat{q}_{ji} + o(\varepsilon)$ for $\varepsilon \to 0$ with $\hat{q}_{ji} \in [0,+\infty)$. The limiting procedure is performed as $\varepsilon \to 0$ and $t \to \infty$ simultaneously, assuming that

$$t \to \infty \quad \text{as} \;\; \varepsilon \to 0 \text{ and } \varepsilon t \to s \text{ as } \varepsilon \to 0, \text{ where } 0 \leq s \leq \infty.$$

The authors point out that their model is adequate for a moderate number of patches: when the number of patches is very large, a different limit procedure, where the number of patches tends to infinity, seems to be more satisfactory.

Deterministic approximation of stochastic models

The main purpose of this thesis is to look into the relationship between some deterministic and stochastic models that can describe the evolution in time of a metapopulation : is it possible to obtain deterministic models as an approximation of a family of stochastic ones?

This kind of problem has been suggested by H. Metz [Met]: his starting point is a model with M equal finite patches each of size N: the population size in each patch follows a stochastic birth and death process, with the possibility of catastrophes, which set a patch population to zero; the patches are coupled by migration in a fully symmetric manner, through a common disperser pool; both the catastrophe rates, the emigration and immigration rates may depend on the local population size. He also adds two parameters measuring the overall migration speed and catastrophe rate. The aim is to let either or both the parameters M and N go to infinity and either of the two rate parameters go to zero, after some appropriate rescaling of local population sizes, patch numbers or time.

Deterministic models as approximations of stochastic ones have been studied in a more general setting for a long time, especially in works by T.G. Kurtz ([Kur70], [Kur71],[EK86]) and A.D. Barbour [Bar74], [Bar80]: for a large class of differential equations, the deterministic solutions can be viewed as limits in probability of pure jump Markov processes. The results in [Kur70],[Kur71],[Bar80] and in [Pol90] concern the so-called density -dependent or asymptotically density-dependent one-parameter family of Markov chains; they are useful to describe epidemic models, population growth, chemical reactions [EK86] [Bal99], parasite models [Pol90], [BK93]. In particular, the density dependent models for macro-parasites [BK93], [KA93] show some similarities with our metapopulation models (the hosts can be considered as patches and the parasites per host represents the local size); in [KA93] the model is deterministic, while in [BK93] the deterministic system is the limit of a stochastic one, when the number of hosts is large, and it describes the evolution over time of the proportions of the population with different parasite loads.

The model presented in Chapter 2 of this thesis describes a metapopulation with M patches and size of patches N, similar to that described above by Metz; we analyse the model behaviour as M increases to infinity. From the mathematical point of view, we consider a pure jump M-dimensional Markov process, whose components are the number of individuals in the i-th patch: the local dynamics includes births, deaths, migrations and catastrophes. To study the limit, we choose as variables of our process the proportion of patches with j individuals; thus we are dealing with a Markov process with values in $(\frac{\mathbb{Z}}{M} \cap [0,1])^{\mathbb{N}}$. The result is that this process tends in distribution to a deterministic one, under some assumptions concerning the asymptotic behaviour of the initial distribution. The limit satisfies a infinite system of differential equations (there are analogies to the model in [JM99], where the number of differential equations is finite, because the number of individuals per patches cannot exceed a fixed k). The mathematical tools used in this chapter are martingale theorems and convergence of processes on the space $\mathcal{D}[0,T]$ (with references to [Bal99], [BK93], [Luc99]) to get the deterministic approximation, and semigroup theory to prove the existence and uniqueness of the solution of the deterministic system ([Paz83],[BIT91]).

In the model of Chapter 3, which describes the spread of an epidemic of S-I-S type in a population structured in subpopulations (as has been introduced in [Bal99] and in [BMST97]), the number of individuals per patch is finite, with the population at each site fixed at a constant N. Also for this model, we have two parameters: M, the number of patches, and N, the number of individuals per patch: if we keep one parameter fixed and let the other tend to infinity, we can directly apply the theorems of Kurtz [EK86] and Barbour [BK93] to get a deterministic limit, which satisfies a finite system of ordinary differential equations: when N is going to infinity, the variables are the proportion of infectives per patch, and when M is going to infinity, the variables are the proportion of patches with j infectives.
The main results in this chapter concern the case when both the parameters go to infinity, taking the proportion of patches with a percentage of infectives in a assigned range as the variable that describes the spread of the epidemic: from a mathematical point of view, this is a stochastic process with values in $\mathcal{M}([0,1])$, the space of probability measures on $[0,1]$. The two iterated limits both lead to the

same non-linear differential equation for an unknown measure $\mu(t)$: if we assume that this measure is absolutely continuous for all t in the interval $[0, T]$, that is $\mu(t) = h(t, x)dx$, the differential equation turns into a first order PDE for $h(t, x)$. We show the existence and uniqueness of the solution to the equation in the general case, and we study the asymptotic behaviour in time.

Among the problems that still remain open, we mention in particular the proof of a central limit theorem studying the deviations of the process from its deterministic approximation, under a proper rescaling (as is shown for general density population processes in [EK86] and recently applied to an epidemic model in [AB00]), the analysis of the asymptotic behaviour in time for the deterministic metapopulation model and the possibility of performing the double limit also in the metapopulation model of Chapter 2. As both M and N increase to infinity, the idea is to introduce a carrying capacity K for the local population and to perform the limit also as $K \to \infty$, in order to get a partial differential equation for the distribution of the structuring variable x (heuristically, $x = \frac{j}{K}$, with j the number of individuals in a patch); it may be possible that such limit procedures give rise to deterministic models similar to those in [GH92] and in [VVM96], under proper assumptions on the asymptotic behaviour of the other parameters involved.

Chapter 1

Preliminaries

In this preliminary chapter we introduce some basic notation and concepts, that will be used to study the models from the mathematical point of view: we give some basic definition about stochastic processes, Markov chains and martingales.
We suggest some works from the reference that cover this material:[And91], [BR97], [EK86],[Fel68],[Fel66], [Nor98], [Reu57],[RY99].

Let E a metric space, $\mathcal{B}$ the σ-algebra of the Borel sets in E and $I = [0, +\infty)$; let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space.

Definition 1.1 *A stochastic process X with state space (E) defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a function defined on $I \times \Omega$ with values in E such that, for all t, $X(t, \cdot) : \Omega \to E$ is a E-valued random variable.*

We say that a process X is measurable if $X : I \times \omega \to E$ is $\mathcal{B}(I) \times \mathcal{F}$-measurable.
The function $X(\cdot, \omega)$ is called the sample path of the process at ω; if for almost $\omega \in \Omega$ the sample path $X(\cdot, \omega)$ is (right, left)continuous , then the process is said (right, left) continuous.
The information obtained by observing the process X up to time t is 'contained' in the natural filtration associated to to X

$$\mathcal{F}_t^X = \sigma(X(s) : s \leq t)$$

A process is adapted to a given filtration $\mathcal{F}_t$ if $\mathcal{F}_t^X \subset \mathcal{F}_t$.
We notice that every right (left) continuous $\mathcal{F}_t$-adapted process is measurable.

The processes that will appear in the following chapters are pure jump processes with values in a countable state space and their sample paths are constant except for isolated jumps and right continuous .

A key concept related to the evolution in time of stochastic processes is the following

Definition 1.2 *A real valued process X with $\mathbb{E}(|X(t)|) < \infty$ for all $t \in I$ and adapted to a filtration $\mathcal{F}_t$ is an $\{\mathcal{F}_t\}$-martingale if*

$$\mathbb{E}[X(t+s)|\mathcal{F}_t] = X(t) \quad t, s \geq 0$$

The processes involved in the description of our model are Markov processes.

Definition 1.3 *X is a Markov process if*

$$\mathbb{P}[X(t+s) \in \Gamma | \mathcal{F}_t^X] = \mathbb{P}[X(t+s) \in \Gamma | X(t)] \tag{1.1}$$

for all $s, t \geq 0$ and $\Gamma \in \mathcal{B}(E)$.
Note that (1.1) implies

$$\mathbb{E}[f(X(t+s)) | \mathcal{F}_t^X] = \mathbb{E}[f(X(t+s)) | X(t)]$$

for all $s, t \geq 0$ and $f \in B(E)$(f is a bounded function defined in E).

Definition 1.4 *A function $P(s, x; t, \Gamma)$ defined on $\mathbb{R}_+ \times E \times \mathbb{R}_+ \times \mathcal{B}(E)$ for $s \leq t$ is a transition function if*

1. *$P(s, x; t, \cdot)$ is a probability measure on E*
2. *$P(s, \cdot; t, \Gamma)$ is $\mathcal{B}(E)$-measurable*
3. *$P(s, x; s, \Gamma) = \delta_x(\Gamma)$*
4. $$P(s, x; t, \Gamma) = \int_E P(u, y; t, \Gamma) P(s, x; u, dy) \qquad s < u < t$$

The transition function is called homogeneous if $P(s, x; t, \Gamma) = P(s+r, x; t+r, \Gamma)$ so that we define $P(t, x, \Gamma) = P(0, x; t, \Gamma)$.

A transition function is a transition function for a Markov process X if

$$\mathbb{P}[X(t) \in \Gamma | X(s)] = P(s, X(s); t, \Gamma) \qquad a.e.$$

Taking a version of the conditional probability, the last property can be written as

$$\mathbb{P}[X(t) \in \Gamma | X(s) = x] = P(s, x; t, \Gamma)$$

The simplest Markov process to describe is a *Markov jump process* with countable state space, that we denote by $Z_t(\cdot) = X(t, \cdot)$; we labelled its countable states with the integers $i \in \mathbb{Z}^+$ and introduce two new functions

1. *the intensity function* $q_i(t)$, which has the following probabilistic interpretation:
   $q_i(t)h + o(h)$ is the probability that Z_t will undergo a random change in the interval $(t, t+h)$, conditioned by $Z_t = i$;

2. the *relative transition probability function* $\Pi_{ij}(t)$, which is the conditional probability of Z assuming the value j at time $t+h$, given $Z_t = i$ and a change has happened in the interval $(t, t+h)$ with the obvious properties $\Pi_{ii}(t) = 0$ and $\sum_{j \neq i} \Pi_{ij} = 1$.

For small values of h, we assume that

$$\mathbb{P}[Z(t+h) = j | Z(t) = i] = (1 - q_i(t)h)\delta_{ij} + q_i(t)\Pi_{ij}(t)h + o(h) \tag{1.2}$$

where $\lim_{h \to 0} \frac{o(h)}{h} = 0$ uniformly in i, j. Put

$$\begin{aligned} q_{ii}(t) &= -q_i(t) \\ q_{ij}(t) &= q_i(t)\Pi_{ij}(t), \qquad i \neq j. \end{aligned}$$

The elements $q_{ij}(t)$ have the following properties

$$\begin{aligned} q_{ij}(t) &\geq 0 \qquad && i \neq j \\ q_{ii}(t) &\leq 0 && \\ \sum_{j=0}^{+\infty} q_{ij}(t) &\leq 0 && \textit{for all } i. \end{aligned}$$

The quantities $q_{ij}(t)h + o(h)$ are called the *infinitesimal transition probabilities* of the Markov process Z_t. We will adopt the following notation throughout the thesis

$$i \to j \quad \text{at rate} \quad q_{ij}(t)$$

If the process is homogeneous in time, $q_{ij}(t) = q_{ij}$.

For a homogeneous Markov pure jump process, one can define the times of jumps $\{\tau_n\}$

$$\tau_0 = 0, \quad \tau_{n+1} = \inf\{t > \tau_n \, : Z_t - Z_{t^-} \neq 0\},$$

with the convention that $\inf \emptyset = \infty$.
By the Markov property, conditionally on $\{Z_{\tau_n} = i\}$ for $\tau_{n+1} < \infty$ *a.e.* we have

a) $\tau_{n+1} - \tau_n$ is exponentially distributed with a parameter that depens on i and that we denote by $\varrho(z)$

b) $Z_{\tau_{n+1}}$ has the probability distribution $\pi(i, \cdot) = \Pi_{i\cdot}$

If $\lim_{n \to \infty} \tau_n = +\infty$, the process is called regular or non-explosive and it can be can be written as

$$Z_t = Z_0 + \sum_{n \geq 1} J_n \mathbb{1}_{[\tau_n \leq t]}$$

where $\{J_n\}$ are its jumps, defined by

$$J_0 = 0, \quad J_{n+1} = Z_{\tau_{n+1}} - Z_{\tau_n}.$$

If $\lim_{n\to\infty} \tau_n = \tau_\infty$, explosions are possible; the process is not determined by its infinitesimal transitions, then to completely describe the process, it is necessary to specify its evolution after the time τ_∞ of the first explosion. In this thesis, we only deal with regular processes. The process

We mention in particular the Poisson process , with parameter $\lambda > 0$ where, for $i \geq 0$

$$q_i = \lambda, \quad \Pi_{ij} = \begin{cases} 1 & \text{if} j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

and the birth-and-death process , with parameters $\lambda_i \geq 0, \mu_i \geq 0$ where for $i \geq 0$

$$q_i = \lambda_i + \mu_i, \quad \Pi_{ij} = \begin{cases} \dfrac{\lambda_i}{\lambda_i + \mu_i} & \text{if} \quad j = i+1 \\ \dfrac{\mu_i}{\lambda_i + \mu_i} & \text{if} \quad j = i-1 \\ 0 & \text{if} \quad |i-j| > 1 \end{cases}$$

with $\mu_0 = \lambda_{-1} = 0$.

A homogeneous Markov process generates a semigroup on the space of bounded functions $B(E)$ (not necessarily a $\mathcal{C}_0$-semigroup on the whole space): we say that an E-valued Markov process X corresponds to the semigroup $T(t)$ defined on a closed subspace $G \subset B(E)$ if

$$(T(s)f)(X(t)) = \mathbb{E}(f(X(t+s))|\mathcal{F}_t^X] \qquad f \in G, \ s, t \geq 0.$$

Given a homogeneous transition function, we can define a semigroup on $B(E)$ by

$$\int f(y)P(t,x,dy) = T(t)f(x) \tag{1.3}$$

If P is the transition function of the homogeneous Markov process X, then the semigroup defined by (1.3) corresponds to X. For example, if we take as state space $\mathbb{Z}^+$, the space $B(E) = \ell^\infty$: a transition function generates a $\mathcal{C}_0$-semigroup on $G = \ell^1$. (See for details [And91]).
We have the following proposition from [Kur81]

Proposition 1.5 *Let X be an E-valued Markov process with initial distribution ν. Let $T(t)$ be a semigroup on $G \subset B(E)$ corresponding to X. If G is bp-dense in $B(E)$ (that is, G is dense in $B(E)$ in the topology of bounded pointwise convergence), then $T(t)$ and ν determine the finite dimensional distribution of X. In particular*

$$\mathbb{E}[f(X(t))] = \int_E (T(t)f(x))\,\nu(dx).$$

Let $\|\cdot\|$ denote the sup norm on $B(E)$ and G be a Banach subspace of $B(E)$. Let $T(t)$ be a contraction semigroup defined on G such that $T(t)f(x)$ is $\mathcal{B}(I)\times\mathcal{B}(E)$ measurable for all $f \in G$. We can define

Definition 1.6 *The full generator* $\tilde{L}$ *of* $T(t)$ *is the (in general multivalued) operator defined by*

$$\tilde{L} = \left\{(f,h) \in G \times G : \, T(t)f - f = \int_0^t T(s)h \, ds\right\}$$

The (strong)infinitesimal generator L *of* $T(t)$ *is the (single-valued) operator defined by*

$$L = \left\{(f,h) \in G \times G : \, \lim_{t\to 0} \frac{T(t)f - f}{t} = h\right\}$$

For any multivalued operators

$$\begin{aligned} \mathcal{D}(\tilde{L}) \quad &= \{f : (f,h) \in \tilde{L}\} \\ \mathcal{R}(\tilde{L}) \quad &= \{g : (f,h) \in \tilde{L}\} \end{aligned}$$

For a Markov jump processes, the infinitesimal generator is

$$Lf(z) = \varrho(z) \int (f(y) - f(z)) \, \pi(z, dy)$$

If X is a Markov process with respect to $\mathcal{F}_t$ corresponding to a semigroup $T(t)$ with full generator $\tilde{L}$, then $(f,g) \in \tilde{L}$ implies that

$$f(X(t)) - \int_0^t g(X(s)) \, ds$$

is a $\mathcal{F}_t$-martingale

It is possible to use the martingale property to characterize the Markov process corresponding to a given generator. Let $\tilde{L} \in \mathcal{B}(E) \times \mathcal{B}(E)$ and ν a probability measure on E: by a solution of the martingale problem for $(\tilde{L}, \nu)$ we mean a measurable stochastic process X with values in E and initial distribution ν defined on some probability space such that

$$f(X(t)) - \int_0^t h(X(s))ds$$

is a $\mathcal{F}_t^X$-martingale for every $(f,h) \in \tilde{L}$. We have the following theorem (see [Kur81])

Theorem 1.7 *Let* $\tilde{L} \in \mathcal{B}(E) \times \mathcal{B}(E)$ *be a closed, linear, dissipative operator such that* $\mathcal{R}(\lambda_0 - A) \supset \mathcal{D}(\tilde{L})$ *for some* $\lambda_0 > 0$*, and suppose that* $\mathcal{D}(\tilde{L})$ *is bp-dense in* $B(E)$*. Let* ν *be a probability measure on* E *and let* X *a solution of the martingale problem for* $(\tilde{L}, \nu)$*. Then* X *is a Markov process corresponding to the semigroup generated by* $\tilde{A}$ *and the martingale problem is well posed (that is, there exists a solution and every solution has the same finite dimensional distributions).*

Remark Also for some kinds of non-homogeneous Markov jump processes, it is possible to define a family of operators $L(t)$ in an analogous way

$$L(t)f(z) = \varrho(t,z) \int (f(y) - f(z))\, \pi(t,z,dy)$$

If there exists a measurable locally integrable function r on $[0,+\infty)$, such that $\varrho(t,z) \leq r(t)$ for all t and z, we define

$$\tilde{L} = \{(f, \varrho(,) \int (f(y+\cdot) - f(\cdot))\, \pi(\cdot,\cdot,dy) : f \in B(E)\} \subset B(E) \times B(E \times I)$$

then the martingale problem for $\tilde{L}$ is well-posed (see [EK86]).

Chapter 2

A metapopulation model

2.1 Introduction

In this chapter, we present a stochastic structured metapopulation model, where the structuring variable is the number of inhabitants per patch: we consider a population composed of M subpopulations, distributed in patches, labelled with i, for $i = 1, \dots, M$.
The metapopulation dynamics can be described by a Markov jump process, which takes into account local dynamics (births, deaths and catastrophes) and also migration from each patch to the others: in the first section, we define the continuous Markov chains that govern the evolution of the metapopulation .
The main purpose of this chapter is to study the behaviour of the process, when the number of patches M goes to infinity:in this context, the more suitable variable seems to be $\underline{x}^M(t)$, whose components $x_j^M(t)$ represents the fraction of patches with j-inhabitants, for $j \geq 0$.

We notice that for every fixed t, $x_j^M(t) \geq 0$, $\displaystyle\sum_{j=0}^{+\infty} x_j^M(t) = 1$, so that $\{x_j^M(t)\}_{j\geq 0}$ can be considered as a random probability measure on $\mathbb{N}$.
The quantity

$$E^M(t) = \sum_{j=0}^{+\infty} j x_j^M(t)$$

is the mean of this measure and, if it is finite, it represents the average number of inhabitants per patch; the quantity

$$E_2^M(t) = \sum_{j=0}^{+\infty} j^2 x_j^M(t)$$

is the second moment of this probability measure.

Our goal is to prove that the random process $\underline{x}^M$ converges in distribution to a deterministic process $\underline{p}$ under some natural assumptions on the convergence of the initial value of the process: we assume, in fact, that the initial distribution of inhabitants $\underline{x}^M(0)$ has finite second moment and that

$$s_2 := \sup_M \sum_{j=0}^{+\infty} j^2 x_j^M(0) < +\infty \qquad a.e.; \tag{2.1}$$

from this condition, we have existence and boundedness of the first moments, and we define

$$s_1 = \sup_M \sum_{j=0}^{+\infty} j x_j^M(0). \tag{2.2}$$

The proof consists of many steps. In section [2.4], we will show that the sequence of probability measures defined on $(\mathcal{D}([0,T]))^\infty$ and associated to the stochastic process $\underline{x}^M$ is tight (here we use only the fact that $s_1 < +\infty$, (2.4)). Because of tightness, we can extract a weakly convergent subsequence, and in section [2.5] we prove that every limit $\underline{p}$ of such a convergent subsequence satisfies a set of differential equations (here we use the the fact that $s_2 < \infty$ (2.1)). The existence and uniqueness of the solution for the deterministic system is studied in section (2.6.1) by semigroups methods: the Banach space on which our sequences live is

$$m^1 = \{\underline{x} \in \ell^1, \sum_j j|x_j| < \infty\}.$$

In particular, the uniqueness of the solution implies that the limit for every weakly convergent subsequence of $\underline{x}^M$ is always the same; this fact together with the tightness is enough to establish that the whole sequence $\underline{x}^M$ is weakly convergent to $\underline{p}$ as $M \to \infty$.

2.2 The model

We denote by Y_i the number of individuals in the $i-th$ patch and by $\underline{e}_i$ the $i-th$ coordinate vector in $\mathbb{N}^M$; $\underline{Y} = (Y_1, \ldots, Y_M)$ is a Markov process with state space $\mathbb{N}^M$ and transition rates

$$\begin{array}{llll} \underline{Y} & \to \underline{Y} + \underline{e}_i & \text{at rate} & Y_i b_{Y_i} \\ \underline{Y} & \to \underline{Y} - \underline{e}_i & \text{at rate} & Y_i d_{Y_i} + \gamma Y_i (1-\rho) \\ \underline{Y} & \to \underline{Y} - \underline{e}_i + \underline{e}_l & \text{at rate} & Y_i \gamma \rho \frac{1}{M} \\ \underline{Y} & \to \underline{Y} - Y_i \underline{e}_i & \text{at rate} & \nu \end{array}$$

for $i \geq 1$. The parameters in the model are

- $b_j > 0,\ j \geq 1$ the *per capita* birth rate. Let us assume that jb_j is concave and increasing in j (in particular, b_j is decreasing in j) and set $b = \max\{j \in \mathbb{N} : b_j\}$;
- $d_j > 0,\ j \geq 1$ the *per capita* death rate. Let us assume that jd_j is convex and increasing in j (in particular, d_j is increasing in j)and set $d = \min\{j \in \mathbb{N} : d_j\}$;
- γ is the emigration rate;
- $0 \leq \rho \leq 1$ is the fraction of emigrants from one patch that succeed in reaching another patch;
- ν is the catastrophe rate.

In order the study the behaviour of the stochastic process when the number of patches increases to infinity, we consider the process

$$\underline{x}^M = (x_0^M, \ldots, x_r^M, \ldots) \qquad x_r^M = \frac{\sum_{i=1}^{M} 1\!\!1_{[Y_i = r]}}{M} \qquad r = 0, 1, 2, \ldots \tag{2.3}$$

This is a Markov jump process with state-space $\left(\frac{\mathbb{Z}}{M} \cap [0,1]\right)^{\mathbb{N}}$ and transition rates

$$\begin{array}{llllll} \underline{x} & \to \underline{x} & + \frac{1}{M}(\underline{e}_{i+1} - \underline{e}_i) & \text{at rate} & M x_i b_i i & i \geq 1 \\ \underline{x} & \to \underline{x} & + \frac{1}{M}(\underline{e}_{i-1} - \underline{e}_i) & \text{at rate} & M x_i (d_i i + \gamma(1-\rho) i) & i > 1 \\ \underline{x} & \to \underline{x} & + \frac{1}{M}(\underline{e}_{i+1} - \underline{e}_i) & \text{at rate} & M \rho \gamma k x_k\, x_i & i \geq 0, \\ & & + \frac{1}{M}(\underline{e}_{k-1} - \underline{e}_k) & & & k \geq 1 \\ \underline{x} & \to \underline{x} & + \frac{1}{M}(\underline{e}_0 - \underline{e}_i) & \text{at rate} & M x_i \nu & i \geq 2 \\ \underline{x} & \to \underline{x} & + \frac{1}{M}(\underline{e}_0 - \underline{e}_1) & \text{at rate} & M x_1 (\nu + d_1 + (1-\rho)\gamma) & \end{array}$$

Let $E^M(t) = \frac{1}{M}\sum_{i=1}^{M} Y_i(t)$: it represents the empirical mean number of individuals per patch. Using the assumption that

$$E^M(0) = \frac{1}{M}\sum_{i=1}^{M} Y_i(0) = \mathbb{E}(\sum_{j=0}^{+\infty} j x_j^M(0)) \leq s_1 \tag{2.4}$$

and a coupling argument (for some basic details about coupling see [Lin92]), we will prove

Proposition 2.2.1 *The stochastic process $\tilde{E}^M = ME^M = \sum_{i=1}^{M} Y_i$ is stochastically smaller than a pure birth process F^M with per capita birth rate b*

PROOF. Let $(\underline{Z}, \underline{\tilde{Z}})$ be a pure jump Markov process on a probability space Ω with state space $\mathbb{N}^{2M}$ and transition rates given by

$$\begin{array}{lll}
(\underline{Z}, \underline{\tilde{Z}}) \to (\underline{Z} + \underline{e}_i, \underline{\tilde{Z}}) & \text{at rate} & Z_i b_{Z_i} \\
(\underline{Z}, \underline{\tilde{Z}}) \to (\underline{Z}, \underline{\tilde{Z}} + \underline{e}_i) & \text{at rate} & \tilde{Z}_i b + Z_i(b - b_{Z_i}) \\
(\underline{Z}, \underline{\tilde{Z}}) \to (\underline{Z} - \underline{e}_i, \underline{\tilde{Z}} + \underline{e}_i) & \text{at rate} & Z_i d_{Z_i} + \gamma Z_i(1-\rho) \\
(\underline{Z}, \underline{\tilde{Z}}) \to (\underline{Z} - Z_i\underline{e}_i, \underline{\tilde{Z}} + Z_i\underline{e}_i) & \text{at rate} & \nu \\
(\underline{Z}, \underline{\tilde{Z}}) \to (\underline{Z} - \underline{e}_i + \underline{e}_l, \underline{\tilde{Z}}) & \text{at rate} & \rho\gamma\frac{1}{M}Z_i
\end{array}$$

for all $1 \leq i, l \leq M$. The process $\underline{Z}$ is Markovian having the same generator as $\underline{Y}$; if $\underline{\tilde{Z}}(0) = 0$ the process $\underline{\tilde{Z}}$ is positive (that is $\tilde{Z}_i \geq 0$ for $i = 1, \ldots M$) for all $t \in [0, T]$. If $\underline{Z}(0)$ has the same distribution as $\underline{Y}(0)$, then the process $F^M = \sum_{i=1}^{M}(Z_i + \tilde{Z}_i)$ is a pure birth process with *per capita* rate b; $\tilde{E}^M \leq F^M$, and $\tilde{E}^M(0)$ has the same distribution of $F^M(0)$.
Because of Proposition (2.2.1), we have

Corollary 2.2.2

$$\mathbb{E}[\sup_{0\leq u\leq s} \tilde{E}^M(u)] \leq \mathbb{E}[\sup_{0\leq u\leq s} F^M(u)]$$

and then (under our hypothesis the empirical mean $E^M(0)$ has finite expectation)

$$\mathbb{E}[\sup_{0\leq u\leq s} E^M(u)] \leq \frac{1}{M}e^{bs}\mathbb{E}(F^M(0)) = \mathbb{E}(E^M(0))e^{bs}$$

Corollary 2.2.3 *If C is a positive constant, then*

$$\mathbb{P}[E^M(t) > C] \leq \frac{1}{C}\mathbb{E}(E^M(0))e^{bt}$$

PROOF. Because of 2.2.1, $\mathbb{P}[E^M(t) > C] \leq \mathbb{P}[F^M(t) > CM]$.
For a fixed t, let us apply the Markov inequality to the positive random variable $F^M(t)$

$$\begin{aligned}\mathbb{P}[F^M(t) > CM] &\leq \frac{1}{CM}\mathbb{E}(F^M(t)) = \frac{1}{CM}\mathbb{E}(F^M(0))e^{bt} = \frac{1}{C}\mathbb{E}(E^M(0))e^{bt} \\ &\leq \frac{1}{C}\mathbb{E}(E^M(0))e^{bT}\end{aligned}$$

for all $t \in [0, T]$.

The main purpose of this chapter is the proof of

Theorem 2.2.4 *If* $\underline{x}_0^M \Rightarrow \underline{p}^0$ *in distribution, with* $\sup_M \sum_{j\leq 1} j^2 x_j^M(0) < \infty$ *a.e., then* $\underline{x}^M \Rightarrow \underline{p}$ *in* $(\mathcal{D}[0,T])^\infty$ *in distribution, where* $\underline{p}$ *is the unique positive solution of the Cauchy problem*

$$\begin{cases} p_i'(t) = & -\left[(b_i + d_i + \gamma)\, i + \nu + \rho\gamma \sum_{j=0}^{\infty} j p_j(t)\right] p_i(t) \qquad i \geq 1 \\ & + \left[b_{i-1}(i-1) + \rho\gamma \sum_{j=0}^{\infty} j p_j(t)\right] p_{i-1}(t) \\ & + [d_{i+1} + \gamma]\,(i+1)\, p_{i+1}(t) \\ p_0'(t) = & \nu(1 - p_0(t)) + (d_1 + \gamma) p_1(t) - \rho\gamma \left(\sum_{j=0}^{\infty} j p_j(t)\right) p_0(t) \\ \underline{p}(0) = & \underline{p}^0 \end{cases}$$

2.3 Martingale tools

Our aim in this section is to show that some processes, functions of our process $\underline{x}^M$, are martingales with respect to the natural filtration $\mathcal{F}_t^M = \sigma\{\underline{x}^M(s), 0 \leq s \leq t\}$.

We recall some results about Markov processes and martingales. First we have the following expression for the generator L of a Markov jump process Z_t:

$$Lf(z) = \varrho(z) \int [f(y) - f(z)]\pi(z, dy). \tag{2.5}$$

where f belongs to the domain of the generator.
In the following table, we mention the positive values of the purely atomic measure $\varrho\pi$ related to our process $\underline{x}^M$

y	$\varrho(z)\pi(z, dy)$
$\frac{1}{M}(\underline{e}_{i+1} - \underline{e}_i)$	$Mx_ib_ii \quad i \geq 1$
$\frac{1}{M}(\underline{e}_{i-1} - \underline{e}_i)$	$Mx_i(d_ii + \gamma(1-\rho)i) \quad i \geq 2$
$\frac{1}{M}(\underline{e}_{i+1} - \underline{e}_i) + \frac{1}{M}(\underline{e}_{k-1} - \underline{e}_k)$	$M\rho\gamma k x_k\, x_i \quad i, k \geq 1$
$\frac{1}{M}(\underline{e}_0 - \underline{e}_i)$	$Mx_i\nu \quad i \geq 2$
$\frac{1}{M}(\underline{e}_0 - \underline{e}_1)$	$Mx_1(\nu + d_1 + (1-\rho)\gamma)$

It is known that, if f is a bounded function, the process

$$N_t^f = f(Z_t) - f(Z_0) - \int_0^t Lf(Z_s)ds \tag{2.6}$$

is a martingale. In Hamza-Klebaner [HK95] and Luchsinger[Luc99], the following sufficient conditons are given for the integrability of the process $f(Z_t)$ and the martingale property of N_t^f, for unbounded functions f.

Theorem 2.3.1 *Let Z_t be a regular process with values in a general space S; we consider a function $f : S \to \mathbb{R}$ and assume one of the following conditions:*

a) There exists a positive constant c for which

$$|L|f(z) = a(z) \int |f(y) - f(z)|\pi(z, dy) \leq c(1 \vee |f(z)|)$$

b) There exists a function F such that

$$|L|f(z) \leq F(z), \qquad \mathbb{E}\left[\sup_{u\in[0,s]} F(Z_u)\right] < \infty$$

If $f(Z_0)$ is integrable, then $f(Z_t)$ is integrable and N_t^f is a martingale.

For our purpose, the space S will be the space

$$m_+^1 = \{x \in \ell^1, x_j \geq 0, \sum_j x_j \leq 1, \sum_j jx_j < \infty\}; \tag{2.7}$$

the function $f : m_+^1 \to \mathbb{R}$ will be the j-th projection $j \geq 0$

$$f_j(z) = z_j, \qquad z \in m_+^1$$

$j \geq 0$ and the process Z_t will be $\underline{x}^M(t)$.

From now on, we adopt the following notation for $j \geq 1$

$$\begin{aligned} q_{j-1,j} &= b_{j-1}(j-1) & \tilde{q}_{j-1,j}(\underline{x}) &= q_{j-1,j} + \rho\gamma \textstyle\sum_k kx_k \\ q_{j+1,j} &= (d_{j+1} + \gamma)(j+1) & \tilde{q}_{j+1,j}(\underline{x}) &= q_{j+1,j} \\ q_{j,j} &= -(q_{j,j-1} + q_{j,j+1}) & \tilde{q}_{j,j}(\underline{x}) &= -(\tilde{q}_{j,j-1}(\underline{x}) + \tilde{q}_{j,j+1}(\underline{x})) - \nu \end{aligned}$$

$$\begin{aligned} \alpha_j(\underline{x}) &= \tilde{q}_{j-1,j}(\underline{x})x_{j-1} + \tilde{q}_{j,j}(\underline{x})x_j + \tilde{q}_{j+1,j}(\underline{x})x_{j+1} \\ \alpha_0(\underline{x}) &= -\left(\nu + \rho\gamma \textstyle\sum_k kx_k\right) x_0 + (d_1 + \gamma)x_1 + \nu \\ \beta_j(\underline{x}) &= \tilde{q}_{j-1,j}(\underline{x})x_{j-1} - \tilde{q}_{j,j}(\underline{x})x_j + \tilde{q}_{j+1,j}(\underline{x})x_{j+1} \\ \beta_0(\underline{x}) &= \left(\nu + \rho\gamma \textstyle\sum_k kx_k\right) x_0 + (d_1 + \gamma)x_1 + \nu \end{aligned}$$

Lemma 2.3.2 *Let us assume that $\mathbb{E}(x_j^M(0)) < \infty$, $j \geq 0$; we define*

$$U_j^M(t) = x_j^M(t) - x_j^M(0) - \int_0^t \alpha_j(x^M(u))\, du$$

Then U_j^M is a martingale with respect to the natural filtration $\mathcal{F}_t^M$.

PROOF. We notice that

$$U_j^M(t) = f_j(\underline{x}^M(t)) - f_j(\underline{x}^M(0)) - \int_0^t (Lf_j)(\underline{x}^M(s))\, ds.$$

Since $|L|(f_j)(z) \leq F(z)$ where

$$\begin{aligned} F(z) &= 2\left(\max(b_{j-1}, b_j)j + \max(d_{j+1}, d_j)(j+1) + \gamma(j+1) + \nu + \rho\gamma \sum_k kz_k\right) \\ &= H_j + 2\rho\gamma \sum_k kz_k; \end{aligned}$$

and, by Corollary (2.2.2)

$$\begin{aligned} \mathbb{E}\left[\sup_{0\leq u\leq s} F(x^M(u))\right] &= H_j + 2\rho\gamma\mathbb{E}\left[\sup_{0\leq u\leq s} \sum_k kx_k^M(u)\right] \\ &= H_j + 2\rho\gamma\mathbb{E}\left[\sup_{0\leq u\leq s} E^M(u)\right] < \infty, \end{aligned}$$

we can apply Theorem 2.3.1 b).

Before stating next lemma, we present the results in Corollaries A5, A6 of [Luc99], which involve the higher moments of the distribution $\pi(z, dy)$ for a general process Z_t ; we define

$$\mathbf{m}_k(z) = \int (f_j(y))^k \pi(z, dy) \qquad |\mathbf{m}_k|(z) = \int |f_j(y)|^k \pi(z, dy)$$

for $k = 1, 2, \ldots$.

Proposition 2.3.3 *Let us assume that*

- $\{f_j(Z_0)\}^k$ *is integrable*
- *there exists a function* $F : S \to \mathbb{R}$ *such that*

$$\varrho(z)|\boldsymbol{m}_i(z)| \leq F(z), \qquad i = 1, \ldots, k$$

with

$$\mathbb{E}\left(\sup_{0\leq u\leq s} F(Z_u)\right) < \infty$$

Then $(f_j(Z_t))^k$ *is integrable and*

$$(f_j(Z_t))^k - (f_j(Z_0))^K - \sum_{i=0}^{k-1} \binom{k}{i} \int_0^t \varrho(Z_s)\, (f_j(Z_s))^i\, \boldsymbol{m}_{k-i}(Z_s)\, ds$$

is a martingale.

We write down the expressions for these martingales in the cases $k=1,2$.

$$\begin{aligned} k=1 \quad \mathbf{N}_j^1(Z_t) &= (Z_t)_j - (Z_0)_j - \int \varrho(Z_s)\mathbf{m}_1(Z_s)ds \\ k=2 \quad \mathbf{N}_j^2(Z_t) &= (Z_t)_j^2 - (Z_0)_j^2 - \int \varrho(Z_s)\mathbf{m}_2(Z_s)ds - 2\int \varrho(Z_s))f_j(Z_s)\mathbf{m}_1(Z_s)ds \end{aligned} \tag{2.8}$$

Remark : Let $\left[\mathbf{N}_j^1, \mathbf{N}_j^1\right]_t$ be the quadratic variation process, then

$$(Z_t)_j^2 - (Z_0)_j^2 - 2\int \varrho(Z_s)f_j(Z_s)\mathbf{m}_1(Z_s)ds - \left[\mathbf{N}_j^1, \mathbf{N}_j^1\right]_t \tag{2.9}$$

is a martingale.

Also

$$\begin{aligned} &\left[\mathbf{N}_j^1, \mathbf{N}_j^1\right]_t - \int \varrho(Z_s))\mathbf{m}_2(Z_s)ds \\ &\qquad = \mathbf{N}_j^2(t) - \left\{(Z_t)_j^2 - (Z_0)_j^2 - 2\int \varrho(Z_s)f_j(Z_s)\mathbf{m}_1(Z_s)ds - \left[\mathbf{N}_j^1, \mathbf{N}_j^1\right]_t\right\} \end{aligned} \tag{2.10}$$

is a martingale.

Lemma 2.3.4 *Let us assume that* $\mathbb{E}((x_j^M)^2(0)) < \infty,\ j \geq 0$*; we define*

$$V_j^M(t) = \left(U_j^M(t)\right)^2 - \frac{1}{M}\int_0^t \beta_j(\underline{x}^M(s))\,ds$$

Then $V_j^M(t)$ *is a martingale with respect to the natural filtration* $\mathcal{F}_t^M$.

PROOF. We notice that

$$V_j^M(t) = \left(U_j^M(t)\right)^2 - \int_0^t \varrho(\underline{x}^M(s))\mathbf{m}_2(\underline{x}^M(s))\,ds,$$

and, according to the previous notation,

$$U_j^M(t) = N_j^1(\underline{x}^M(t)), \qquad \text{and} \qquad \varrho(z)|\mathbf{m}_2|(z) \leq F(z)$$

with the same function F as in Lemma 2.3.2, so that

$$\left[\mathbf{N}_j^1, \mathbf{N}_j^1\right]_t - \int \varrho(\underline{x}^M(s))\mathbf{m}_2(x^M(s))ds = \left[U_j^M, U_j^M\right]_t - \int \varrho(\underline{x}^M(s))\mathbf{m}_2(x^M(s))ds$$

is a martingale. Since $U_j^M(t)$ is a martingale, the process

$$\left(U_j^M(t)\right)^2 - \left[U_j^M, U_j^M\right]_t$$

is a martingale, and by differencing we obtain that

$$\left(U_j^M(t)\right)^2 - \int_0^t \varrho(\underline{x}^M(s))\mathbf{m}_2(\underline{x}^M(s))\,ds = V_j^M(t)$$

is a martingale.

2.4 Tightness

Proposition 2.4.1 *For each $j \geq 0$, the sequence $\{x_j^M\}_{M\geq 0}$ is tight in $\mathcal{D}[0,T]$ and the limit of any convergent subsequence belongs to $C([0,T])$*

PROOF. In order to show that the sequence $\{x_j^M\}_{M\geq 0}$ is tight in $\mathcal{D}[0,T]$, we have to prove that $\forall \varepsilon$, $\eta > 0$, there exist $\delta \in (0,1)$ and $M_0 \geq 0$ such that

$$\mathbb{P}\left[\sup_{0\leq s\leq t\leq T,\,(t-s)<\delta} |x_j^M(t) - x_j^M(s)| \geq \varepsilon\right] \leq \eta, \quad M \geq M_0$$

We define

$$w(x_j^M;\delta) = \sup_{0\leq s\leq t\leq T,\,t-s<\delta} |x_j^M(t) - x_j^M(s)|,$$

and

$$A_s^M = \{\sup_{t\in(s,s+\delta)} |x_j^M(t) - x_j^M(s)| \geq \frac{\varepsilon}{3}\}$$

First of all, we notice that, setting $\delta = \dfrac{T}{n}$

$$\{w(x_j^M;\delta) \geq \varepsilon\} \subset \bigcup_{i=0}^{n-1} A_{i\delta}^M \tag{2.11}$$

Let $\delta = \dfrac{T}{n}$ for $n > 0$.

$$\mathbb{P}[\,w(x_j^M;\delta) \geq \varepsilon] \leq \mathbb{P}[\,w(x_j^M;\delta) \geq \varepsilon,\ \sum_{j\geq 0} jx_j^M < C] + \mathbb{P}[\sum_{j\geq 0} jx_j^M \geq C]. \tag{2.12}$$

Since $\mathbb{P}[\sum_{j\geq 0} jx_j^M \geq C] \leq \dfrac{1}{C}e^{bT}s_1$ (Corollary 2.2.3), let us choose C such that $\dfrac{1}{C}e^{bT}s_1 \leq \dfrac{\eta}{2}$.
Because of (2.11), we have

$$\mathbb{P}[\,w(x_j^M;\delta) \geq \varepsilon,\ \sum_{j\geq 0} jx_j^M < C] \leq \sum_{i=0}^{n-1} \mathbb{P}[A_{i\delta}^M\,,\ \sum_{j\geq 0} jx_j^M < C] \tag{2.13}$$

Let us denote

$$\mathbb{P}[A_s^M\,,\ \sum_{j\geq 0} jx_j^M < C] = P_C(A_s^M)$$

Since

$$|x_j^M(t) - x_j^M(s)| \leq |U_j^M(t) - U_j^M(s)| + \int_s^t |\alpha_j(x^M(u))|\,du$$

we can estimate

$$
\begin{aligned}
P_C(A_s^M) \;&\le P_C\left(\sup_{0\le s\le t\le s+\delta}|U_j^M(t)-U_j^M(s)|>\frac{\varepsilon}{6}\right)+\\
&+P_C\left(\sup_{0\le s\le t\le s+\delta}\left|\int_s^t|\alpha_j(x^M(u))|\,du>\frac{\varepsilon}{6}\right)
\end{aligned}
\tag{2.14}
$$

For the second term on the right side of (2.14), we have

$$
P_C\left(\int_s^{s+\delta}|\alpha_j(x^M(u))|\,du>\frac{\varepsilon}{6}\right)\le P_C\left((H_j+2\rho\gamma C)\delta>\frac{\varepsilon}{6}\right).
$$

Let us choose n such that $\delta<\dfrac{\frac{\varepsilon}{6}}{H_j+2\rho\gamma C}$; then

$$
P_C\left((H_j+2\rho\gamma C)\delta>\frac{\varepsilon}{6}\right)\le\mathbb{P}\left((H_j+2\rho\gamma C)\delta>\frac{\varepsilon}{6}\right)=0.
\tag{2.15}
$$

For the first term on the right side of (2.14), we use the martingale property

$$
\begin{aligned}
\mathbb{P}\left(\sup_{0\le s\le t\le s+\delta}|U_j^M(t)-U_j^M(s)|>\frac{\varepsilon}{6}\right)&\le\frac{36}{\varepsilon^2}\mathbb{E}\left[(U_j^M(s+\delta)-U_j^M(s))^2\right]\\
=\frac{36}{\varepsilon^2}\{\mathbb{E}[(U_j^M(s+\delta))^2]-\mathbb{E}[U_j^M(s))^2]\}&=\frac{36}{\varepsilon^2}\mathbb{E}[\frac{1}{M}\int_s^{s+\delta}\beta_j(x^M(u))\,du]\\
&\le\frac{36}{\varepsilon^2M}\delta(H_j+2\rho\gamma e^{bT}s_1).
\end{aligned}
$$

We can conclude that

$$
\sum_{i=0}^{n-1}P_C(A_{i\delta}^M)\le n\delta\frac{36}{\varepsilon^2M}(H_j+2\rho\gamma e^{bT}s_1)=T\frac{36}{\varepsilon^2M}(H_j+2\rho\gamma e^{bT}s_1).
\tag{2.16}
$$

Let us choose M such that $T\dfrac{36}{\varepsilon^2M}(H_j+2\rho\gamma e^{bT}s_1)<\dfrac{\eta}{2}$. Then

$$
\begin{aligned}
\mathbb{P}[w(x_j^M;\delta)\ge\varepsilon]\;&\le\mathbb{P}[\,w(x_j^M;\delta)\ge\varepsilon,\ \textstyle\sum_{j\ge0}jx_j^M<C]+\mathbb{P}[\textstyle\sum_{j\ge0}jx_j^M\ge C]\\
&\le\sum_{i=0}^{n-1}\mathbb{P}[A_{i\delta}^M,\sum_{j\ge0}jx_j^M<C]+\mathbb{P}[\sum_{j\ge0}jx_j^M\ge C]\\
&\le\frac{\eta}{2}+\frac{\eta}{2}=\eta
\end{aligned}
$$

Lemma 2.4.2 $\underline{x}^M(t)$ *is tight in* $(\mathcal{D}[0,T])^\infty$

PROOF. Given $\epsilon>0$ let K_j be a compact set in $\mathcal{D}[0,T]$ such that $\mathbb{P}[x_j^M\in K_j]>1-2^{-(j+1)}\epsilon$. Then $K=\prod_{j\ge0}K_j$ is compact in $(\mathcal{D}[0,T])^\infty$ and $\mathbb{P}[x^M\in K]\ge1-\epsilon$

2.5 Weak convergence

From the previous section, we know there exists a weakly convergent subsequence $\{\underline{x}^{M_k}\}$ (for sake of simplicity, we denote it by $\{\underline{x}^M\}$); let $\underline{p}(t) = (p_0(t), p_1(t), \dots)$ be its weak limit. The aim of this section is to prove that

$$\mathbb{P}\left[p_j(t) = p_j^0(0) + \int_0^t \alpha_j(\underline{p}(u))\, du \quad \text{for all } 0 \le t \le T\right] = 1 \tag{2.17}$$

First of all, we have

Lemma 2.5.1 *For each $j \ge 0$ and $\epsilon > 0$*

$$\lim_{M\to\infty} \mathbb{P}\left[\sup_{t\in[0,T]} |U_j^M(t)| > \epsilon\right] = 0$$

PROOF. Applying the Doob-Kolmogorov inequality to U_j^M and using the martingale property of V_j^M, we get

$$\begin{aligned}
\mathbb{P}\left[\sup_{t\in[0,T]} |U_j^M(t)| > \epsilon\right] &\le \frac{1}{\epsilon^2}\mathbb{E}\left(U_j^M(T)^2\right) \le \frac{1}{\epsilon^2 M}\int_0^T \mathbb{E}(\beta_j(\underline{x}^M(u)))du \\
&\le \frac{1}{\epsilon^2 M}\int_0^T \left(H_j + \mathbb{E}(\textstyle\sum_k k x_k^M(u))\right) du \\
&\le \frac{1}{\epsilon^2 M}(H_j + e^b T s_1)T
\end{aligned}$$

which converges to zero as $M \to \infty$.

Lemma 2.5.2 *If $\underline{x}^M \Rightarrow \underline{p}$, then $U_j^M \Rightarrow U_j$ for any $j \ge 0$ where*

$$U_j(t) = p_j(t) - p_j^0 - \int_0^t \alpha_j(\underline{p}(u))du$$

PROOF. For $\underline{x} \in (\mathcal{D}[0,T])^\infty$ and $j \ge 0$ fixed, we define

$$h(\underline{x})(t) = x_j(t) - x_j(0) - \int_0^t \left((B\underline{x})_j + \rho\gamma\left(\sum_{k=1}^{\infty} k x_k(u)\right)(x_j(u) - x_{j-1}(u))\right) du$$

and

$$h_N(\underline{x})(t) = x_j(t) - x_j(0) - \int_0^t \left((B\underline{x})_j + \rho\gamma\left(\sum_{k=1}^{N} k x_k(u)\right)(x_j(u) - x_{j-1}(u))\right) du$$

where

$$(B\underline{x})_j = \sum_k q_{kj} x_k - \nu x_j - \nu\delta_0^j \qquad x_{-1}(u) \equiv 0$$

According to that notation, $U_j^M = h(\underline{x}^M)$ and $U_j = h(\underline{p})$; we must show that

$$\lim_{M\to\infty} \mathbb{E}(f(h(\underline{x}^M))) = \mathbb{E}(f(h(\underline{p}))) \qquad \forall f \in UC_b(\mathcal{D}[0,T]).$$

Since

$$\begin{aligned}\left|\mathbb{E}(f(h(\underline{x}^M))) - \mathbb{E}(f(h(\underline{p})))\right| &\le \left|\mathbb{E}(f(h(\underline{x}^M))) - \mathbb{E}(f(h_N(\underline{x}^M)))\right| \\ &+ \left|\mathbb{E}(f(h_N(\underline{x}^M))) - \mathbb{E}(f(h_N(\underline{p})))\right| \\ &+ \left|\mathbb{E}(f(h_N(\underline{p}))) - \mathbb{E}(f(h(\underline{p})))\right|,\end{aligned}$$

we are going to estimate every term on the right side.
For the first term, we use the uniform continuity of f and the following Lemmas . For every positive ϵ, we can choose a $\delta > 0$ such that if $z, z' \in \mathcal{D}[0,T]$ with $d(z,z') < \delta$ then $|f(z) - f(z')| < \frac{\epsilon}{6}$ (remembering that if $z, z' \in \mathcal{D}[0,T]$, $d(z,z') \le \sup_{0\le t\le T} |z(t) - z'(t)|$).
$h(\underline{x}^M)$ and $h_N(\underline{x}^M)$ are elements of $\mathcal{D}[0,T]$ and

$$h(\underline{x}^M) - h_N(\underline{x}^M)(t) = \int_0^t \rho\gamma \left(\sum_{k=N+1}^{\infty} k x_k^M(u)\right)(x_j^M(u) - x_{j-1}^M(u))du.$$

Therefore, if $\sup_{0\le t\le T} \left|h(\underline{x}^M)(t) - h_N(\underline{x}^M)(t)\right| < \delta$ then $|f(h(\underline{x}^M)) - f(h_N(\underline{x}^M))| < \dfrac{\epsilon}{6}$.
We use some lemmas to get an estimate of the probability of the set

$$B_N^M = \left\{ \sup_{0\le t\le T} \left| \rho\gamma \int_0^t \left(\sum_{k=N+1}^{\infty} k x_k^M\right)(x_{j-1}^M(u) - x_j^M(u))\,du \right| \ge \delta \right\}.$$

Lemma 2.5.3 *Let $\underline{Z}^M$ be a Markov process with transition rates*

$$\begin{array}{llll} \underline{Z} & \to \underline{Z} + \underline{e}_i & \textit{at rate} & Z_i b \\ \underline{Z} & \to \underline{Z} - \underline{e}_i & \textit{at rate} & Z_i d + \gamma Z_i(1-\rho) \\ \underline{Z} & \to \underline{Z} - \underline{e}_i + \underline{e}_l & \textit{at rate} & Z_i \gamma\rho \frac{1}{M} \end{array}$$

then $\underline{Y}^M \prec^s \underline{Z}^M$.

PROOF. Let $(\underline{W},\underline{\tilde{W}})$ be a $2M$-dimensional pure jump Markov process on a probability space Ω with transition rates given by

$$
\begin{array}{llll}
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W}+\underline{e}_i,\underline{\tilde{W}}) & \text{at rate} & W_i b_{W_i} \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W}-\underline{e}_i,\underline{\tilde{W}}) & \text{at rate} & W_i \tilde{d}_{W_i} \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W},\underline{\tilde{W}}+\underline{e}_i) & \text{at rate} & (W_i+\tilde{W}_i)b - W_i b_{W_i} \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W}-\underline{e}_i,\underline{\tilde{W}}+\underline{e}_i) & \text{at rate} & W_i(d_{W_i}-d) \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W},\underline{\tilde{W}}-\underline{e}_i) & \text{at rate} & \tilde{W}_i \tilde{d} \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W}-\underline{e}_i+\underline{e}_l,\underline{\tilde{W}}) & \text{at rate} & W_i \frac{1}{M}\rho\gamma \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W},\underline{\tilde{W}}-\underline{e}_i+\underline{e}_l) & \text{at rate} & \tilde{W}_i \frac{1}{M}\rho\gamma \\
(\underline{W},\underline{\tilde{W}}) & \to (\underline{W}-W_i\underline{e}_i,\underline{\tilde{W}}+W_i\underline{e}_i) & \text{at rate} & \nu
\end{array}
$$

with $\tilde{d}_i = d_i + (1-\rho)\gamma$ and $\tilde{d} = d + (1-\rho)\gamma$. All the transition are non-negative, because of the assumptions on the death and birth rates.
The process $\underline{W}$ is Markovian with the same generator as the birth and death process $\underline{Y}^M$, the process $\underline{W}+\underline{\tilde{W}}$ is Markovian with the same generator as the process $\underline{Z}^M$. Thus we can couple the realization $\underline{W}$ of the process $\underline{Y}^M$ with the same initial distribution and the realization $\underline{W}+\underline{\tilde{W}}$ of the process $\underline{Z}^M$ with $\underline{\tilde{W}}(0)=0$; it is plain that $\underline{\tilde{W}}(t)\geq 0$ for all t and then $\underline{W}(t)+\underline{\tilde{W}}(t)\geq \underline{W}(t)$ for all t.

We define

$$
\underline{z}^M = (z_0^M,\dots,z_r^M,\dots) \qquad z_r^M = \frac{\sum_{i=1}^M \mathbb{1}_{[Z_i=r]}}{M} \qquad r=0,1,2,\dots,
$$

so that then, by stochastic comparison, $\mathbb{E}[\sum_{j\geq N} jx_j^M(t)] \leq \mathbb{E}[\sum_{j\geq N} jz_j^M(t)],\ t\in[0,T]$.
We use a martingale argument to estimate the right side of the previous inequality. Let

$$
m_2^M(t) = \sum_{j=1}^{\infty} j^2 z_j^M(t); \tag{2.18}
$$

according to the notation we have introduced in Section 2.3, we define

$$
f(\underline{z}) = \sum_{j=1}^{\infty} j^2 z_j \quad \text{and} \quad c_2(\underline{z}) = Lf(\underline{z});
$$

after some calculations, we get

$$
\begin{aligned}
c_2(\underline{z}) &= \sum_{j=1}^{\infty}\left[(j+1)^2-j^2\right] jz_j + \sum_{j=1}^{\infty}\left[(j-1)^2-j^2\right] djz_j + (1-\rho)\gamma\sum_{j=1}^{\infty}\left[(j-1)^2-j^2\right] jz_j + \\
&+\rho\gamma\sum\sum\nolimits_{j,k}\left[(j+1)^2-j^2+(k-1)^2-k^2\right] kz_kz_j
\end{aligned}
\tag{2.19}
$$

■

Using Theorem 2.3.1, we can prove

Proposition 2.5.4 *If* $\mathbb{E}\big(f(\underline{z}^M(0))\big) < \infty$, *then* $\mathbb{E}\big(f(\underline{z}^M(t))\big) < \infty,\ t \in [0,T]$ *and*

$$W_2^M(t) = m_2^M(t) - m_2^M(0) - \int_0^t c_2\big(\underline{z}^M(u)\big)du \tag{2.20}$$

is a martingale.

PROOF. Since

$$\begin{aligned} |L|f(\underline{z}) &= \sum_{j=1}^{\infty} \big|(j+1)^2 - j^2\big|\, bjz_j + \sum_{j=1}^{\infty} \big|(j-1)^2 - j^2\big|\, djz_j \\ &\quad + (1-\rho)\gamma \sum_{j=1}^{\infty} \big|(j-1)^2 - j^2\big|\, jz_j \\ &\quad + \rho\gamma \sum\sum_{j,k}^{\infty} \big|(j+1)^2 - j^2 + (k-1)^2 - k^2\big|\, kz_kz_j \\ &= b\sum_{j=1}^{\infty} |2j+1| jz_j + (d + (1-\rho)\gamma) \sum_{j=1}^{\infty} |2j-1| jz_j \\ &\quad + 2\rho\gamma \sum_{j,k} |j-k+|kz_jz_k, \end{aligned} \tag{2.21}$$

it follows that

$$|L|f(\underline{z}) \leq 3bf(z) + 3(d + (1-\rho)\gamma)f(z) + 2\rho\gamma\left[\left(\sum_k kz_k\right)^2 + \left(\sum_k k^2 z_k\right)(1-z_0)\right].$$

For $\underline{z} \in m_+^1$, with $\sum_{j=0}^{\infty} z_j = 1$, we have

$$\left(\sum_k kz_k\right)^2 \leq \left(\sum_k k^2 z_k\right)$$

so that

$$|L|f(\underline{z}) \leq 3(b + d + (1-\rho)\gamma + 2\rho\gamma)|f(z)| \tag{2.22}$$

Thus, we can apply Hamza-Klebaner theorem (2.3.1).
Hypothesis (2.1) gives us an estimate for $\mathbb{E}(m_2^M(t))$ which is uniform in M: let us denote $C_2 = 3(b + d + \gamma(1-\rho)) + 6\rho\gamma$; taking the expectation in (2.20) we get

$$\mathbb{E}(m_2^M(t)) = \mathbb{E}(m_2^M(0)) + \int_0^t \mathbb{E}(c_2\big(\underline{z}^M(u)\big))du$$

but

$$|c_2(\underline{z}(t))| \leq C_2 m_2^M(t)$$

so that

$$\mathbb{E}(m_2^M(t)) \leq \mathbb{E}(m_2^M(0)) + C_2 \int_0^t \mathbb{E}(m_2^M(u))du$$

From the Gronwall inequality, it follows that

$$\mathbb{E}(m_2^M(t)) \leq \mathbb{E}(m_2^M(0))e^{C_2 t}.$$

Finally

$$\begin{aligned} \mathbb{E}\left(\sum_{j\geq N} jx_j^M(t)\right) \leq \mathbb{E}\left(\sum_{j\geq N} jz_j^M(t)\right) &\leq \frac{1}{N}\mathbb{E}\left[\sum_{j\geq N} j^2 z_j^M(t)\right] \leq \frac{1}{N}\mathbb{E}(m_2^M(t)) \\ &\leq \frac{1}{N}\mathbb{E}(m_2^M(0))e_2^C t \end{aligned} \tag{2.23}$$

Applying Lemma 2.5.3, we can give an estimate of the probability of the set B_N^M; indeed,

$$\begin{aligned} B_N^M \subset & \left\{ \sup_{0\leq t\leq T} \int_0^t 2\rho\gamma \left(\sum_{k=N+1}^{\infty} kx_k^M(u) \right) du \geq \delta \right\} \\ = & \left\{ \int_0^T \left(\sum_{k=N+1}^{\infty} kx_k^M(u) \right) du \geq \frac{\delta}{\rho\gamma 2} \right\}, \end{aligned}$$

and thus

$$\begin{aligned} \mathbb{P}[B_N^M] &\leq \mathbb{P}\left[\int_0^T \left(\sum_{k=N+1}^{\infty} kx_k^M(u)\right) du \geq \frac{\delta}{2\rho\gamma}\right] \\ &\leq \frac{2\rho\gamma}{\delta}\mathbb{E}\left(\int_0^T \left(\sum_{k=N+1}^{\infty} kx_k^M(u)\right) du\right) = \frac{2\rho\gamma}{\delta}\int_0^T \mathbb{E}\left(\sum_{k=N+1}^{\infty} kx_k^M(u)\right) du \\ &\leq \frac{2\rho\gamma}{\delta}\int_0^T \frac{1}{N}\mathbb{E}\big(m_2^M(0)\big)e^{C_2 u}\, du = \frac{1}{N}\frac{2\rho\gamma}{\delta}\mathbb{E}\big(m_2^M(0)\big)\frac{1}{C_2}\left(e^{C_2 t}-1\right) \\ &\leq \frac{1}{N}\left(\frac{2\rho\gamma}{\delta}s_2\frac{1}{C_2}\left(e^{C_2 T}-1\right)\right) \end{aligned} \tag{2.24}$$

It thus follows from (2.24) that

$$\begin{aligned} \left|\mathbb{E}(f(h(\underline{x}^M))) - \mathbb{E}(f(h_N(\underline{x}^M)))\right| &\leq \int_\Omega |f(h(\underline{x}^M)) - f(h_N(\underline{x}^M))|d\mathbb{P} \\ &= \int_{B_N^M} |f(h(\underline{x}^M)) - f(h_N(\underline{x}^M))|d\mathbb{P} + \\ &+ \int_{B_N^{M\,c}} |f(h(\underline{x}^M)) - f(h_N(\underline{x}^M))|d\mathbb{P} \\ &\leq 2\|f\|_\infty \mathbb{P}[B_N^M] + \tfrac{\epsilon}{6} \end{aligned}$$

If we choose N such that

$$\frac{1}{N}2\|f\|_\infty\left(\frac{2\rho\gamma}{\delta}s_2\frac{1}{C_2}\left(e^{C_2T}-1\right)\right)\le\frac{\epsilon}{6},$$

then

$$\left|\mathbb{E}(f(h(\underline{x}^M)))-\mathbb{E}(f(h_N(\underline{x}^M)))\right|<\frac{\epsilon}{3}$$

A similar argument can be used for $\left|\mathbb{E}(f(h_N(\underline{p})))-\mathbb{E}(f(h(\underline{p})))\right|$ because

$$\begin{aligned}\mathbb{E}\left(\sum_{k=N+1}^{\infty}kp_k(t)\right) &= \lim_{s\to\infty}\mathbb{E}\left(\sum_{k=N+1}^{s}kp_k(t)\right)\le\lim_{s\to\infty}\liminf_{M\to\infty}\mathbb{E}\left(\sum_{k=N+1}^{s}kx_k^M(t)\right)\\ &\le\liminf_{M\to\infty}\mathbb{E}\left(\sum_{k=N+1}^{\infty}kx_k^M(t)\right)\le\frac{1}{N}\mathbb{E}(m_2^M(0))e_2^Ct\end{aligned}$$

Finally, since $f\circ h_N\in C_b(\mathcal{D}[0,T])$ for every N, we have

$$\lim_{M\to\infty}\left|\mathbb{E}(f(h_N(\underline{x}^M)))-\mathbb{E}(f(h_N(\underline{p})))\right|=0$$

and thus we can choose M such that

$$\left|\mathbb{E}(f(h_N(\underline{x}^M)))-\mathbb{E}(f(h_N(\underline{p})))\right|\le\frac{\epsilon}{3}.$$

As a consequence of Lemma 2.5.1 , Lemma 2.5.2 and because $\underline{x}^m\Rightarrow\underline{p}^0$ we therefore have (3.14) for any weak limit $\underline{p}$.

2.6 Abstract setting for the limit equation

In this section, we will prove the existence and uniqueness of the following Cauchy problem

$$\begin{cases} p_i'(t) = & -\left[(b_i + d_i + \gamma)\, i + \nu + \rho\gamma \sum_{j=0}^{\infty} j p_j(t)\right] p_i(t) \qquad i \geq 1 \\ & + \left[b_{i-1}(i-1) + \rho\gamma \sum_{j=0}^{\infty} j p_j(t)\right] p_{i-1}(t) \\ & + [d_{i+1} + \gamma]\,(i+1)\, p_{i+1}(t) \\ p_0'(t) = & \nu(1 - p_0(t)) + (d_1 + \gamma) p_1(t) - \rho\gamma \left(\sum_{j=0}^{\infty} j p_j(t)\right) p_0(t) \\ \underline{p}(0) = & \underline{p}^0 \end{cases} \tag{2.25}$$

It follows from the previous section (2.5) that the deterministic limit $\underline{p} = (p_0, p_1, p_2, \dots)$ of the stochastic process $\underline{x}^M(t)$ must satisfies (2.25). We can write down the system for $i \geq 0$ in a more compact way as

$$\begin{aligned} p_i' &= -(\lambda_i + \mu_i + \nu) p_i + \lambda_{i-1} p_{i-1} + \mu_{i+1} p_{i+1} \\ &\quad + \rho\gamma \left(\textstyle\sum_{j=0}^{\infty} j p_j\right) (p_{i-1} - p_i) + \nu \delta_0^i \end{aligned} \tag{2.26}$$

where

$$p_{i-1} \equiv 0 \qquad \mu_0 = 0, \quad \lambda_{-1} = 0, \quad \lambda_i = b_i i, \quad \mu_i = (d_i + \gamma) i = \bar{d}_i i$$

We identify the problem as an abstract evolution equation in a Banach space X.

2.6.1 Abstract Cauchy Problem

Let X be a Banach space and let A be a linear operator from $\mathcal{D}(A) \subset X$ into X. Given $u_0 \in X$, the abstract Cauchy problem for A with initial data $u_0 \in X$ consists of finding a solution $u(t)$ to the initial value problem

$$\begin{cases} \dfrac{du(t)}{dt} &= Au(t), \qquad t > 0 \\ u(0) &= u_0 \end{cases} \tag{2.27}$$

where by a solution we mean an X valued function $u(t)$ such that u(t) is continuous for $t \geq 0$, continuously differentiable and $u(t) \in \mathcal{D}(A)$ for $t > 0$ and (2.27) is satisfied.

Our problem (2.25) actually gives rise to a semilinear initial value problem

$$\begin{cases} \dfrac{du(t)}{dt} &= Au(t) + F(u(t)), \qquad t > 0 \\ u(0) &= u_0 \end{cases} \tag{2.28}$$

where A is the generator of a $\mathcal{C}_0$ semigroup (we denote the semigroup by e^{tA}) and $F : X \to X$ satisfies a Lipschitz condition in u. If there exists a solution of (2.28), it satisfies the integral equation

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds \tag{2.29}$$

We call a continuous solution u of the *mild equation* (2.29) a *mild solution* of the initial value problem (2.28). A classical result assures us of the existence and uniqueness of mild solution of (2.28) for Lipschitz continuous functions F.

Theorem 2.6.1 *([Paz83], VI 1.4) Let $F : X \to X$ be locally Lipschitz continuous . If A is the infinitesimal generator of a $\mathcal{C}_0$ semigroup e^{tA} on X, then for every $u_0 \in X$ there is a $t_{max} \leq \infty$ such that the initial value problem (2.28) has a unique mild solution u on $[0, t_{max})$. Moreover, if $t_{max} < \infty$, then $\lim\limits_{t \uparrow t_{max}} \|u(t)\| = \infty$.*

If we take a function F that is regular enough, then the mild solution is a classical solution of (2.28); we have the following theorem of regularity

Theorem 2.6.2 *([Paz83], VI Th. 1.5) Let A be the infinitesimal generator of a $\mathcal{C}_0$-semigroup; if F is continuously differentiable, then the mild solution of (2.28) with $u_0 \in \mathcal{D}(A)$ is a classical solution of the initial value problem.*

In the next section, we transform the 'concrete' Cauchy problem (2.25) into an abstract one.

2.6.2 The space m^1

For our purpose, the appropriate Banach space X is

$$m^1 = \{\underline{x} \in \ell^1, \sum_j j|x_j| < \infty\}.$$

m^1 is a subspace of the Banach space ℓ^1 and it is a Banach space itself, equipped with the norm

$$|||\underline{x}||| = |x_0| + \sum_{i=0}^{\infty} i|x_i|. \tag{2.30}$$

It is actually an L^1- space of the integrable functions defined in $\mathbb{N}$, when we consider the measurable space $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ with the measure $dm = \delta_0 + \sum_{i \geq 1} i\delta_i$. Therefore its dual space consists of the essentially bounded sequences, that coincides with the space of bounded sequences, with the supremum norm. The duality map $< \cdot, \cdot >$ between m^1 and ℓ^∞ can be explicitly written as

$$< \phi, u >= \int \phi\, u dm = \phi_0 u_0 + \sum_{i=1}^{+\infty} i\phi_i u_i \tag{2.31}$$

We construct the operator A that consists of the linear part of the left part of (2.25); this operator takes into account the linear birth and death part of the dynamics of our metapopulation . Let $\mathbf{Q}$ be the q-matrix

$$(\mathbf{Q})_{ij} = q_{ij} = \begin{cases} b_{j-1}(j-1), & \text{if } i = j-1 \\ -(b_j + (d_j + \gamma))j & \text{if } i = j \\ (d_{j+1} + \gamma)(j+1) & \text{if } i = j+1 \end{cases}$$

Let us denote $q_j = -q_{jj}$, $\bar{d}_i = d_i + \gamma$ and consider the following linear operator on m^1

$$Au = u\mathbf{Q}, \qquad (Au)_i = \sum_k q_{ki} u_k \qquad \mathcal{D}(A) = \{u \in m^1 \,:\, \sum_k k q_k |u_k| < \infty\}. \tag{2.32}$$

The operator A is well-define, that is Au belongs to m^1 for $u \in \mathcal{D}(A)$

$$\begin{aligned} \textstyle\sum_{i=1}^{+\infty} i|(Au)_i| &= \textstyle\sum_{i=1}^{+\infty} i\,|q_{i-1,i}u_{i-1} + q_{ii}u_i + q_{i+1,i}u_{i+1}| \\ &= \textstyle\sum_{i=1}^{+\infty} i\,\left|b_{i-1}(i-1)u_{i-1} - q_i u_i + \bar{d}_{i+1}(i+1)u_{i+1}\right| \\ &\le 2\textstyle\sum_{i=1}^{+\infty} iq_i|u_i| + \sum_{i=1}^{+\infty} iq_i|u_i| + 2\sum_{i=1}^{+\infty} iq_i|u_i| \le 5\sum_{i=1}^{+\infty} iq_i|u_i| \end{aligned}$$

Let us define the map $F : m^1 \to m^1$

$$F(\underline{p}) = \rho\gamma(\sum_{j=0}^{\infty} jp_j)\Big(T_{-1}(\underline{p}) - I(\underline{p})\Big) - \nu I + \nu\delta_0 \tag{2.33}$$

where $(T_{-1}(\underline{p}))_i = p_{i-1}, \quad (I(\underline{p}))_i = p_i, \quad (\delta_0)_i = \delta_0^i$.
According to that notation, the system (2.26) can be written as

$$\begin{cases} \underline{p}'(t) &= A\underline{p}(t) + F(\underline{p}(t)) \\ \underline{p}(0) &= \underline{p}^0 \end{cases} \tag{2.34}$$

2.6.3 A is a $\mathcal{C}_0$-semigroup generator

To prove that A is the infinitesimal generator of $\mathcal{C}_0$-semigroup on m^1, we first study the operator $\mathcal{A} = A - bI$ (we recall that $b = \max\{j \in \mathbb{N} \,:\, b_j\}$), where $I : m^1 \to m^1$ is the identity map; our aim is to apply the Lumer-Phillips theorem ([Paz83], 1.4) to the closure of $\mathcal{A}$: the operator $\bar{\mathcal{A}}$ will be the generator of a $\mathcal{C}_0$-semigroup of contractions on m^1, so the closure $\bar{A}$ of our operator, by a perturbation argument, generates a semigroup on our Banach space.

We recall the definition of the duality set of u, $D(u)$, with u belonging to a general Banach space X

Definition 2.6.3 $D(u) = \{\varphi \in X^\star : <\varphi, u> = \|u\|^2 = \|\varphi\|\}$

Definition 2.6.4 *A linear operator B is dissipative if for every $u \in \mathcal{D}(B)$ there is a $\varphi \in D(u)$ such that Re $< Bu, \varphi > \leq 0$*

Proposition 2.6.5 *The operator $A - bI$ is dissipative.*

Proof. Our Banach space X is the space m^1; if we denote by $\partial\|u\|$ the sub-differential of u defined by

$$\partial\|u\| = \{\phi \in X^\star : <\phi, u> = \|u\|, \ \|\phi\| = 1\}$$

one can easily see that the operator $\mathcal{A}$ is dissipative if for every $u \in \mathcal{D}(A)$ there exists a $\phi \in \partial\|u\|$ such that

$$< (A - b)u, \varphi > \leq 0$$

(indeed, if $\phi \in \partial\|u\|$, then $\varphi = \|u\|\phi \in D(u)$).
The sequence $\phi \in \ell^\infty$ defined as

$$(\phi)_i = \begin{cases} 1, & \text{if } u_i > 0 \\ -1, & \text{if } u_i < 0 \\ c_i, & \text{if } u_i = 0 \end{cases}$$

with $-1 \leq c_i \leq 1$ is an element of $\partial\|\underline{u}\|$ and

$$< (A - b)u, \phi > = \phi_0((A - bI)u)_0 + \sum_{i\geq 1} i\phi_i \left((A - bI)u\right)_i$$

Now

$$\begin{aligned}
\phi_0((A-bI)u)_0 &= -\phi_0 b u_0 + \phi_0 \bar{d}_1 u_1, \\
\text{and} & \\
\sum_{i\geq 1} i\phi_i\left((A-bI)u\right)_i &= \sum_{i\geq 1} i\phi_i(q_{i-1,i}u_{i-1} + q_{i,i}u_i + q_{i+1,i}u_{i+1} - bu_i) \\
&= \sum_{i\geq 1} i\phi_i(b_{i-1}(i-1)u_{i-1} - b_i i u_i - bu_i) \\
&\quad + \sum_{i\geq 1} i\phi_i(\bar{d}_{i+1}(i+1)u_{i+1} - \bar{d}_i i u_i) \\
&= \sum_{i\geq 1}(i+1)\phi_{i+1}b_i i u_i - \sum_{i\geq 1} i^2\phi_i b_i u_i - \sum_{i\geq 1} bi\phi_i u_i \\
&\quad + \sum_{i\geq 2}(i-1)\phi_{i-1}\bar{d}_i i u_i - \sum_{i\geq 1} i^2\phi_i\bar{d}_i u_i \\
&= \sum_{i\geq 1}(\phi_{i+1}u_i - \phi_i u_i)i^2 b_i + \sum_{i\geq 1} i(\phi_{i+1}u_i b_i - \phi_i u_i b)i \\
&\quad + \sum_{i\geq 2}(\phi_{i-1}u_i(i-1) - i\phi_i u_i)\bar{d}_i i - \phi_1\bar{d}_1 u_1.
\end{aligned}$$

Since $0 \leq b_i \leq b$, because of the assumptions on the birth rates, we have

$$\begin{array}{ll}
\phi_{i-1}u_i \leq \phi_i u_i, & \phi_{i+1}u_i \leq \phi_i u_i \\
\phi_{i-1}u_i(i-1) \leq \phi_i u_i(i-1) \leq \phi_i u_i i & \phi_{i+1}u_i b_i \leq \phi_i u_i b_i \leq \phi_i u_i b.
\end{array}$$

Hence

$$\begin{aligned}
< (A-b)u, \phi > &= -\phi_0 b u_0 + \phi_0\bar{d}_1 u_1 - \phi_1\bar{d}_1 u_1 + \\
&\quad + \sum_{i\geq 1}(\phi_{i+1}u_i - \phi_i u_i)i^2 b_i + \sum_{i\geq 1} i(\phi_{i+1}u_i b_i - \phi_i u_i b)i \\
&\quad \sum_{i\geq 2}(\phi_{i-1}u_i(i-1) - i\phi_i u_i)\bar{d}_i i \leq 0.
\end{aligned}$$

Proposition 2.6.6 *If $\mu > b$, the range of the operator $\mu I - \mathcal{A}$ is dense in m^1.*

PROOF. Let $\lambda = \mu + b$; first of all, let us take $v \in m^1$, $v \geq 0$; we will construct a solution u of

$$(\lambda - A)u = v$$

or componentwise

$$\begin{aligned}
\lambda u_j - (Au)_j &= v_j \qquad j \geq 0 \\
\lambda u_j - \textstyle\sum_k q_{kj}u_k &= v_j.
\end{aligned} \tag{2.35}$$

Let us take the finite approximation of order m of the infinite system of equations (2.35)

$$\lambda u_j^{(m)} - \sum_{k \leq m-1} u_k^{(m)} q_{kj} = v_j^{(m)} \qquad j = 0, \ldots, m-1 \tag{2.36}$$

where $v^{(m)}$ is an m-dimensional vector defined by $v_j^{(m)} = v_j$, for $j = 0, \ldots, m-1$; we can write down (2.36) in a matrix form

$$(\lambda I^{(m)} - A^{(m)}) u^{(m)} = v^{(m)}. \tag{2.37}$$

Now $A_{jk}^{(m)} = q_{kj}$, $\quad I^{(m)} = \delta_{jk}$, $\quad B_\lambda^{(m)} := \lambda I^{(m)} - A^{(m)}$;
$j = 0, \ldots, m-1$, $k = 0, \ldots, m-1$.

$$B_\lambda^{(m)} = \begin{bmatrix} \lambda - q_{00} & -q_{10} & 0 & \ldots & \ldots & 0 \\ -q_{01} & \lambda - q_{11} & -q_{21} & 0 & \ldots & 0 \\ 0 & -q_{12} & \lambda - q_{22} & -q_{32} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & 0 & -q_{m-2,m-1} & \lambda - q_{m-1,m-1} \end{bmatrix}$$

$$v^{(m)} = \begin{bmatrix} v_0 \\ v_1 \\ \ldots \\ \ldots \\ v_{m-1} \end{bmatrix}, \qquad u^{(m)} = \begin{bmatrix} u_0^{(m)} \\ u_1^{(m)} \\ \ldots \\ \ldots \\ u_{m-1}^{(m)} \end{bmatrix}$$

For every m, the linear system (2.37) has a positive solution $u^{(m)}$, because of the special form of the matrix $B_\lambda^{(m)} = \lambda I_m - A^{(m)}$; we have the following lemma (relevant definitions and proof can be found in the appendix (A).

Lemma 2.6.7 *The $m \times m$ matrix $B_\lambda^{(m)}$ is a non-singular M-matrix, so that $[B_\lambda^{(m)}]^{-1}$ is strictly positive.*

The components u_j of a solution of (2.35) can be built as limit of the components $u_j^{(m)}$ for $m \to \infty$: for every fixed j, the sequence $u_j^{(m)}$ is positive, non-decreasing in m and bounded.
To see this, note that, for $j = 0, \ldots, m-1$

$$\begin{aligned} \lambda u_j^{(m+1)} - \textstyle\sum_{k \leq m} u_k^{(m+1)} q_{kj} &= v_j^{(m+1)} \\ \lambda u_j^{(m+1)} - \textstyle\sum_{k \leq m-1} u_k^{(m+1)} q_{kj} &= v_j^{(m+1)} + u_m^{(m+1)} q_{mj} \end{aligned}$$

Let us set

$$\tilde{v}^{(m)} = \begin{bmatrix} v_0 \\ v_1 \\ \dots \\ \dots \\ v_{m-1} \end{bmatrix}, \qquad \tilde{u}^{(m)} = \begin{bmatrix} u_0^{(m+1)} \\ u_1^{(m+1)} \\ \dots \\ \dots \\ u_{m-1}^{(m+1)} \end{bmatrix}, \qquad h^{(m)} = \begin{bmatrix} u_m^{(m+1)} q_{m0} \\ u_m^{(m+1)} q_{m1} \\ \dots \\ \dots \\ u_m^{(m+1)} q_{m,m-1} \end{bmatrix}$$

The vector $h^{(m)}$ is non-negative and $\tilde{v}^{(m)} + h^{(m)} \geq v^{(m)}$, so we have

$$\tilde{u}^{(m)} = \left(B_\lambda^{(m)}\right)^{-1} \left(\tilde{v}^{(m)} + h^{(m)}\right) \geq \left(B_\lambda^{(m)}\right)^{-1} \left(\tilde{v}^{(m)}\right) = \left(B_\lambda^{(m)}\right)^{-1} \left(v^{(m)}\right) = u^{(m)}$$

which implies $u_j^{(m+1)} \geq u_j^{(m)}$ for $j = 0, \dots, m-1$. To prove the that the solutions are bounded, let us estimate $\sum_{j=0}^{m-1} u_j^{(m)}$; summing up in (2.36) for $j = 0. \dots, m-1$ and remembering that for $k \leq m-1$, $\sum_{j=0}^{m-1} q_{kj} \leq 0$, we have

$$\begin{aligned} \lambda \sum_{j=0}^{m-1} u_j^{(m)} &= \sum_{j=0}^{m-1} \sum_{k \leq m-1} u_k^{(m)} q_{kj} + \sum_{j=0}^{m-1} v_j^{(m)} \\ &= \sum_{k \leq m-1} u_k^{(m)} \sum_{j=0}^{m-1} q_{kj} + \sum_{j=0}^{m-1} v_j^{(m)} \leq \sum_{j=0}^{m-1} v_j^{(m)} \end{aligned} \tag{2.38}$$

From (2.38), we get, for each $l < m$,

$$\lambda u_l^{(m)} \leq \sum_{j=0}^{m-1} u_j^{(m)} \leq \sum_{j=0}^{m-1} v_j^{(m)} \leq \sum_{j=0}^{\infty} v_j.$$

Hence there exists $u_j = \lim_{m \to \infty} u_j^{(m)}$ and it satisfies (2.35).

If $\underline{v}$ is a generic element of m^1 (not necessarily positive) we can take v_+ and v_- the positive and negative part of v and solve

$$(\lambda - A)u_+ = v_+, \qquad (\lambda - A)u_- = v_-$$

The solution we are asking for is $u = u^+ - u^-$.

Let us give a sufficient condition on v to obtain that our solution u actually belongs to $\mathcal{D}(A)$. The proof needs two lemmas.

Lemma 2.6.8 *If $v \geq 0$ and $\sum_i iv_i < \infty$ and $\lambda > b$, then $\sum_i id_i u_i < \infty$ (where u is the solution of (2.35).* Proof. *We multiply by i and sum over i for $i = 0, \dots, m-1$*

in (2.36) with $u_i^{(m)} = 0$ *for* $i \geq m$, *obtaining*

$$\begin{aligned}
\sum_{i=1}^{m-1} iv_i^{(m)} &= \lambda \sum_{i=1}^{m-1} iu_i^{(m)} + \sum_{i=1}^{m-1} i^2 d_i u_i^{(m)} - \sum_{i=1}^{m-1} i(i+1)u_{i+1}^{(m)} d_{i+1} \\
&\quad + \sum_{i=1}^{m-1} i^2 b_i u_i^{(m)} - \sum_{i=1}^{m-1} i(i-1)u_{i-1}^{(m)} b_{i-1} \\
&= \lambda \sum_{i=1}^{m-1} iu_i^{(m)} + \sum_{i=1}^{m-1} i^2 (d_i + b_i) u_i^{(m)} \\
&\quad - \sum_{i=1}^{m-1} (i-1) i d_i u_i^{(m)} - \sum_{i=1}^{m-2} i(i+1) b_i u_i^{(m)} \\
&= \lambda \sum_{i=1}^{m-1} iu_i^{(m)} + \sum_{i=1}^{m-1} i d_i u_i^{(m)} + (m-1)^2 b_{m-1} u_{m-1}^{(m)} - \sum_{i=1}^{m-2} i b_i u_i^{(m)} \\
&\geq \lambda \sum_{i=1}^{m-1} iu_i^{(m)} + \sum_{i=1}^{m-1} i d_i u_i^{(m)} - b \sum_{i=1}^{m-1} iu_i^{(m)}
\end{aligned}$$

so that

$$(\lambda - b) \sum_{i=1}^{m-1} iu_i^{(m)} + \sum_{i=1}^{m-1} i d_i u_i^{(m)} \leq \sum_{i=1}^{m-1} iv_i^{(m)}$$

If $\lambda > b$, *then* $\sum_i i d_i u_i = \lim_{m\to\infty} \sum_{i=1}^{m-1} i d_i u_i^{(m)} \leq \lim_{m\to\infty} \sum_{i=1}^{m-1} iv_i^{(m)} = \sum_i iv_i < \infty.$

Lemma 2.6.9 *If* $v \geq 0$ $\sum_i i^2 v_i < \infty$ *and* $\lambda > 2b$, *then* $\sum_i i^2 d_i u_i < \infty$, *where* u *is the solution of (2.35).*

PROOF. We multiply by i^2 and sum over i for $i = 0, \ldots, m-1$ in (2.36) with

$u_i^{(m)} = 0$ for $i \geq m$,obtaining

$$\begin{aligned}
\sum_{i=1}^{m-1} i^2 v_i^{(m)} &= \lambda \sum_{i=1}^{m-1} i^2 u_i^{(m)} + \sum_{i=1}^{m-1} i^3 d_i u_i^{(m)} - \sum_{i=1}^{m-1} i^2 (i+1) u_{i+1}^{(m)} d_{i+1} \\
&\quad + \sum_{i=1}^{m-1} i^3 b_i u_i^{(m)} - \sum_{i=1}^{m-2} i^2 (i-1) u_{i-1}^{(m)} b_{i-1} \\
&= \lambda \sum_{i=1}^{m-1} i^2 u_i^{(m)} + \sum_{i=1}^{m-1} i^3 (d_i + b_i) u_i^{(m)} \\
&\quad - \sum_{i=1}^{m-1} (i-1)^2 i d_i u_i^{(m)} - \sum_{i=1}^{m-1} i (i+1)^2 b_i u_i^{(m)} \\
&= \lambda \sum_{i=1}^{m-1} i^2 u_i^{(m)} + 2 \sum_{i=1}^{m-1} i^2 d_i u_i^{(m)} + (m-1)^3 b_{m-1} u_{m-1}^{(m)} \\
&\quad -2 \sum_{i=1}^{m-2} i^2 b_i u_i^{(m)} - \sum_{i=1}^{m-1} i d_i u_i^{(m)} - \sum_{i=1}^{m-2} i b_i u_i^{(m)} \geq \\
&\geq \lambda \sum_{i=1}^{m-1} i^2 u_i^{(m)} + 2 \sum_{i=1}^{m-1} i^2 d_i u_i^{(m)} - 2 \sum_{i=1}^{m-1} i^2 b_i u_i^{(m)} - \sum_{i=1}^{m-1} i (b_i + d_i) u_i^{(m)} \\
&\geq (\lambda - 2b) \sum_{i=1}^{m-1} i^2 u_i^{(m)} + 2 \sum_{i=1}^{m-1} i^2 d_i u_i^{(m)} - \sum_{i=1}^{m-1} i (b_i + d_i) u_i^{(m)}
\end{aligned}$$

Thus

$$(\lambda - 2b) \sum_{i=1}^{m-1} i^2 u_i^{(m)} + 2 \sum_{i=1}^{m-1} i^2 d_i u_i^{(m)} \leq \sum_{i=1}^{m-1} i^2 v_i + \sum_{i=1}^{m-1} i (b_i + d_i) u_i^{(m)}$$

Since there exists a positive constant C such that $b_i \leq b \leq Cd \leq Cd_i$, we have

$$\sum_{i=1}^{m-1} i (b_i + d_i) u_i^{(m)} \leq (1 + C) \sum_{i=1}^{m-1} i d_i u_i^{(m)},$$

and hence, using the previous lemma,

$$\begin{aligned}
\sum_i i^2 d_i u_i = \lim_{m \to \infty} \sum_{i=1}^{m-1} i^2 d_i u_i^{(m)} &\leq \frac{1}{2} \left(\lim_{m \to \infty} \sum_{i=1}^{m-1} i^2 v_i^{(m)} + (1 + C) \sum_{i=1}^{m-1} i d_i u_i^{(m)} \right) \\
&= \frac{1}{2} \left(\sum_i i^2 v_i + (1 + C) \sum_i i v_i \right)
\end{aligned}$$

for $\lambda > 2b$.

If we take $v \in m^2 = \{\underline{v} \in m^1 \ : \sum_i i^2|v_i| < \infty\}$, then both v^+ and v^- belong to m^2, so applying the previous lemmas, we get

$$\begin{aligned} \sum_i iq_i|u_i| = \sum_i i^2(b_i + d_i)|u_i| &\le (C+1)\sum_i i^2 d_i|u_i| = (C+1)\sum_i i^2 d_i(u_i^+ + u_i^-) \\ &\le (C+1)\left[\sum_i i^2 d_i u_i^+ + \sum_i i^2 d_i u_i^-\right] \le \\ &\le \left[\sum_i i^2 v_i^+ + \sum_i i^2 v_i^-\right] \le (C+1)\sum_i i^2|v_i| \end{aligned}$$

Since m^2 is a dense subspace of m^1, for $\mu > 2b$ the range of $(\mu - \mathcal{A})$ is dense in m^1.

Theorem 2.6.10 *The closure $\bar{\mathcal{A}}$ of the operator $\mathcal{A}$ generates a continuous semigroup on m^1*

PROOF. Since $\mathcal{A}$ is dissipative, it is closable and its closure is dissipative too; besides, there exists a μ such that the range of $\mu - \mathcal{A}$ covers m^1. We can apply Lumer-Phillips theorem [Paz83] to show that $\bar{A} - bI$ is the infinitesimal generator of a C_0 semigroup of contraction of m^1.
The closure of our operator A is a perturbation of the operator $\bar{\mathcal{A}}$ via a bounded operator bI, so then $\bar{A}$ generates a C_0 semigroup . The next proposition shows that it is a positive semigroup.

Proposition 2.6.11 *$\bar{A}$ generates a positive C_0 semigroup.*

PROOF. If $\mu > b$, the resolvent operator $R(\mu, \bar{A}) = \overline{R(\mu, A)}$ is a positive operator: let $v \ge 0$ and $u = R(\mu, A)$, then

$$(\mu I - A)u = v$$

Reproducing the steps in the proof of Proposition 2.6.6, one can show that u is positive.

2.6.4 The map F

The non-linear part F (2.33) is regular enough to guarantee the local existence of the mild solution of (2.28).

Proposition 2.6.12 *The map F is locally Lipschitz continuous .*

PROOF. Let $\underline{p}, \underline{r} \in m^1$, $\||\underline{p}\|| \le R, \||\underline{r}\|| \le R$ and let us denote $s(\underline{p}) = \sum_k kp_k$.
First, clearly,

$$|s(\underline{p}) - s(\underline{r})| \le \||\underline{p} - \underline{r}\||.$$

Then we have

$$
\begin{aligned}
|||F(\underline{p}) - F(\underline{r})||| &= \rho\gamma(|s(\underline{p})p_0 - s(\underline{r})r_0| + \sum_k k|s(\underline{p})(p_{k-1} - p_k) - s(\underline{r})(r_{k-1} - r_k)|) \\
&\quad + \nu|||\underline{p} - \underline{r}||| \\
&\leq \rho\gamma \big(|s(\underline{p})p_0 - s(\underline{p})r_0| + |s(\underline{p}) - s(\underline{r})|\,|r_0| \\
&\quad + |s(\underline{p})| \left|\sum_k (k+1)(p_k - r_k)\right| + |s(\underline{p}) - s(\underline{r})| \left|\sum_k (k+1) r_k\right| \\
&\quad + |s(\underline{r})| \sum_k k|(p_k - r_k)| + |s(\underline{p}) - s(\underline{r})| \left|\sum_k k p_k\right| + \big)\, \nu|||\underline{p} - \underline{r}||| \\
&\leq \big(3\rho\gamma \left(|||\underline{p}||| + |||\underline{r}|||\right) + \nu\big)\, |||\underline{p} - \underline{r}|||
\end{aligned}
$$

We can actually prove that the map F is continuously differentiable.

Proposition 2.6.13 *The map $F : m^1 \to m^1$ is differentiable and the Fréchet differential $F'(\underline{p})$ applied to $\underline{z} \in m^1$ is*

$$
(F'(\underline{p}) \cdot \underline{z})_i = \rho\gamma \left[(\sum_{j=1}^{\infty} j p_j)(z_{i-1} - z_i) + (\sum_{j=1}^{\infty} j z_j)(p_{i-1} - p_i) \right] - \nu z_i, \qquad i \geq 0,
$$

where $z_{-1} = p_{i-1} = 0$

PROOF.

$$
\left(F(\underline{p} + \underline{z}) - F(\underline{p}) - F'(\underline{p}) \cdot \underline{z}\right)_i = \rho\gamma(\sum_{j=1}^{\infty} j z_j)(z_{i-1} - z_i)
$$

and

$$
|||(\sum_{j=1}^{\infty} j z_j)(z_{i-1} - z_i)||| \leq 3|||\underline{z}|||^2
$$

Proposition 2.6.14 *The Cauchy problem (2.34) has a unique mild solution $u(t) = \underline{p}(t)$ a priori defined on a maximal interval $[0, t_{\max}]$).(Ch 6, Th 1.4 of [Paz83])*

We have actually proved that the map F is continuously differentiable, so that the solution $u(t)$ is differentiable.

We are now interested in the positive solutions of (2.34), when we start with a positive initial datum.

2.6.5 Positive and bounded solution

First of all, we notice that for any $\alpha > 0$, the solution of the problem (2.29) is equivalent to the solution of

$$
u(t) = e^{-\frac{1}{\alpha}t} e^{tA} u_0 + \frac{1}{\alpha} \int_0^t e^{-\frac{1}{\alpha}(t-s)} e^{(t-s)A} [u(s) + \alpha F(u(s))]\, ds. \tag{2.39}
$$

where e^{tA} is the semigroup generated by the closure of A.

Lemma 2.6.15 *Let* $0 < \alpha < \dfrac{1}{R\rho\gamma+\nu}$. *If* $\underline{p} \geq 0$ $|||\underline{p}||| \leq R$, *then* $\underline{p} + \alpha F(\underline{p}) \geq 0$.

PROOF.

$$(\underline{p}+\alpha F(\underline{p}))_i = (1-\nu\alpha)p_i + \alpha\rho\gamma(\sum_j jp_j)(p_{i-1}-p_i)$$
$$= \alpha\rho\gamma(\sum_j jp_j)p_{i-1} + (1-\alpha\nu-\rho\gamma\alpha(\sum_j jp_j))p_i$$

$$\alpha \leq \frac{1}{R\rho\gamma+\nu} \Longrightarrow \alpha \leq \frac{1}{(\sum_j jp_j)\rho\gamma+\nu} \Longrightarrow (1-\alpha(\nu+\rho\gamma\sum_j jp_j)) \geq 0$$

Proposition 2.6.16 *If* $\underline{p}^0 \geq 0$, *then the solution of (2.39)* $\underline{p}(t)$ *is non-negative on* $t \in [0, t_{\max})$.

PROOF. Let $K = [0, t_k] \subset [0, t_{\max})$ be a compact interval and let us take $R = \max_{t\in K} \|u(t)\|$.
If $u_0 \geq 0$, then from (2.39) one can easily see that $u(t) \geq 0$ for $t \in K$.

Proposition 2.6.17 *If* $\underline{p}_0 \geq 0$ *then there is a global solution of (2.29), that is* $t_{\max} = +\infty$ *(The solution is bounded on bounded intervals).*

PROOF. We have proved that, if the initial datum $\underline{p}^0 \in m^1$ is nonnegative, then the solution $u(t)$ is non negative too on $[0, t_{\max})$: for such a solution , we shall give an estimate of the quantities

$$\sum_{i=0}^{\infty} p_i(t) \quad \text{and} \quad p_0 + \sum_{i=1}^{\infty} ip_i(t).$$

Let $L \in (m^1)^\star$ be a linear continuous functional; let $\underline{p}^0 \in \mathcal{D}(\bar{A})$, so that the solution $u(t)$ is differentiable, we can take the derivative with respect to t of the function $\mathcal{L}(t) = L(u(t))$ and

$$\frac{d\mathcal{L}(t)}{dt} = L(Au(t) + F(u(t))).$$

For our purpose, we deal with two different functionals on m^1:

$$L_0(\underline{p}) = \sum_{i=0}^{\infty} p_i \qquad \text{and} \qquad L_1(\underline{p}) = p_0 + \sum_{i=1}^{\infty} ip_i.$$

Since

$$\begin{aligned}
\frac{d}{dt}\left(\sum_{i=0}^{\infty} p_i\right) &= \frac{d}{dt}L_0(u(t)) \\
&= \textstyle\sum_{i=0}^{\infty}\left(A\underline{p}+F(\underline{p})\right)_i \\
&= \nu(1-p_0)+\bar{d}_1 p_1-\rho\gamma\left(\textstyle\sum_{j=1}^{\infty} j p_j\right) \\
&\quad+\textstyle\sum_{i=1}^{\infty} -\left[\left(b_i+d_i+\gamma\right) i+\nu+\rho\gamma\displaystyle\sum_{j=0}^{\infty} j p_j(t)\right] p_i(t) \qquad i \geq 1 \\
&\quad+\textstyle\sum_{i=1}^{\infty}\left[b_{i-1}(i-1)+\rho\gamma\displaystyle\sum_{j=0}^{\infty} j p_j(t)\right] \\
&\quad+\textstyle\sum_{i=1}^{\infty}\left[d_{i+1}+\gamma\right](i+1)\, p_{i+1}(t) \\
&= \nu-\nu\sum_{i=0}^{\infty} p_i(t),
\end{aligned}$$

it follows that, if $\displaystyle\sum_{i=0}^{\infty} p_i^0 \leq 1$ then $\displaystyle\sum_{i=0}^{\infty} p_i(t) \leq 1$.

To estimate the norm of the positive solution on the interval $[0,t]$, we use the functional L_1, obtaining

$$\begin{aligned}
\frac{d}{dt}L_1(u(t)) &= \frac{d}{dt}\left(p_0(t)+\sum_{i=1}^{\infty} i p_i(t)\right) = \sum_{i=1}^{\infty} i\left(A\underline{p}+F(\underline{p})\right)_i \\
&= -\sum_{i=1}^{\infty} i^2 b_i p_i-\sum_{i=1}^{\infty} i^2 d_i p_i-\gamma\sum_{i=1}^{\infty} i^2 p_i-\nu\sum_{i=1}^{\infty} i p_i-\rho\gamma\left(\sum_{i=1}^{\infty} i p_i\right)^2+ \\
&\quad+\sum_{i=1}^{\infty} i(i-1) b_{i-1} p_{i-1}+\rho\gamma\left(\sum_{i=1}^{\infty} i p_i\right)\sum_{i=1}^{\infty} i p_{i-1}+ \\
&\quad+\sum_{i=1}^{\infty} i(i+1) d_{i+1} p_{i+1}+\gamma\sum_{i=1}^{\infty} i(i+1) p_{i+1}+ \\
&\quad+\nu(1-p_0)+(d_1+\gamma) p_1-\rho\gamma\left(\sum_{i=1}^{\infty} i p_i\right) p_0
\end{aligned}$$

$$\begin{aligned}\frac{d}{dt}L_1(u(t)) = &-\sum_{i=1}^{\infty} i^2 b_i p_i - \sum_{i=1}^{\infty} i^2 d_i p_i - \gamma \sum_{i=1}^{\infty} i^2 p_i - \nu \sum_{i=1}^{\infty} i p_i - \rho\gamma \left(\sum_{i=1}^{\infty} i p_i\right)^2 + \\ &+ \sum_{i=1}^{\infty} i(i+1) b_i p_i + \rho\gamma \left(\sum_{i=1}^{\infty} i p_i\right) \sum_{i=0}^{\infty} (i+1) p_i + \\ &+ \sum_{i=1}^{\infty} i(i-1) d_i p_i + \gamma \sum_{i=1}^{\infty} i(i-1) p_i + \\ &+ \nu(1-p_0) - \rho\gamma \left(\sum_{i=1}^{\infty} i p_i\right) p_0 \\ = &\sum_{i=1}^{\infty} i(b_i - d_i) p_i - \nu\|\underline{p}\| + \nu + \left(\sum_{i=1}^{\infty} i p_i\right)\left(\rho\gamma(\sum_{i=1}^{\infty} p_i) - \gamma\right)\end{aligned}$$

and thus

$$\begin{aligned}\frac{d}{dt}L_1(u(t)) &\leq (b-d)\sum_{i=1}^{\infty} i p_i - \nu\|\underline{p}\| + \nu + \rho\gamma \sum_{i=1}^{\infty} i p_i \\ &\leq (b-d)\sum_{i=1}^{\infty} i p_i - \nu\|\underline{p}\| + \nu + \rho\gamma\|\underline{p}\| \\ &\leq |b-d|\sum_{i=1}^{\infty} i p_i - \nu\|\underline{p}\| + \nu + \rho\gamma\|\underline{p}\| \\ &\leq (|b-d| + \nu + \rho\gamma)\,\|\underline{p}\| + \nu\end{aligned}$$

We denote $(|b-d| + \nu + \rho\gamma)$ by K and from Gronwall inequality, we get

$$\|\underline{p}(t)\| \leq \frac{\nu}{K}(e^{Kt} - 1) + \|\underline{p}(0)\| e^{Kt}$$

Corollary 2.6.18 *If the initial datum $\underline{p}^0 \geq 0$, the unique mild solution of (2.34) is defined on $[0, +\infty)$.*

PROOF. If $t_{\max} < \infty$, $\sup_{t\in[0,t_{\max}]} \|u(t)\|$ would be finite and this is in contradiction with Theorem 2.6.1.

2.7 Conclusion

We recall the following theorem from [Bil68], that states a necessary and sufficient condition of weak convergence for probability measures

Theorem 2.7.1 *We have $P_M \Rightarrow P$ if and only if each subsequence $\{P_{M'}\}$ contains a further subsequence $\{P_{M''}\}$ such that $P_{M''} \Rightarrow P$*

Now we can write down the proof of theorem (2.2.4)
PROOF. The sequence of stochastic processes $\underline{x}^M$ is tight; because of Prokhorov theorem [Bil68], it is weakly compact, so that every subsequence $\underline{x}^{M'}$ contains a convergent subsequence, whose limit is $\underline{p}$, depending a priori on the given subsequence; but this limit satisfies (2.25) and the mild equation (2.29); because of the uniqueness of the mild solution, the limit $\underline{p}$ is the same for all the subsequence; Theorem 2.7.1 assures that the whole sequence $\underline{x}^M$ converges to $\underline{p}$ as $M \to \infty$.

Chapter 3

An SIS model in a structured population

3.1 Introduction

The stochastic SIS model under discussion describes the spread of an epidemic among a population partitioned into M households, each of N individuals, as introduced by F.Ball [Bal99]. From a mathematical point of view, this process can be modeled by a continuous time Markov chain, which involves the two size-parameters M and N and the infections rates. Our aim in this chapter is to study a deterministic approximation to the stochastic model, when both the parameters M and N increase to infinity.
In section 3.2,we give the definition of the Markov chain that describes the evolution over time of the epidemics, and we suggest two different deterministic approximations ; the first one is suitable for not so many households with large population, the second one instead describes a situation of a population of a large number of small households. Both these stochastic processes are density dependent Markov processes and we can apply a "law of large numbers", as Kurtz [EK86] has shown.
In section 3.3, we introduce a new variable, that gives a global description of the system at time t. It is a random probability measure; roughly speaking, it represents the fraction of households with a percentage of infectives in an assigned range. We discuss the existence and uniqueness of the solution of a non-linear differential equation for the deterministic measure, that can be obtained in a formal way as the limit of the random probability measures.
Sections 3.4 and 3.5 are devoted to the study of the deterministic approximation of the random measure as the two parameters M and N go to infinity; we can apply a slight modification of Kurtz [EK86] and Pollett [Pol90] results for asymptotically density dependent processes. In section 3.6, the time evolution of the approximating measure is analysed (when the epidemic threshold $\frac{c+d}{\gamma} < 1$, every subpopulation goes to the same endemic equilibrium as $t \to \infty$).

3.2 The stochastic model

The epidemic spreads among a population consisting of M households , labelled $1, \ldots, M$, each containing N individuals. We describe this process by means of a continuous time Markov chain with values in the lattice $\mathbb{Z}^M$:

$$\underline{Y} = (Y_1, Y_2, \ldots, Y_M)$$

where $Y_i(t)$ denotes the number of infectives at time t in the i-th household; the constants involved are c (the local infection rate), d (the global infection rate) and γ (the removal rate). Let $\underline{k} = (k_1, \ldots, k_m)$ be an M-dimensional vector with $k_j \in \{0, \ldots, N\}$ and let $\underline{e}_j$ be the j-th coordinate vector; $\underline{Y}$ is a jump Markov process with transition rates

$$\begin{cases} \underline{Y} & \to \underline{Y} + \underline{e}_j \text{ at rate } \quad (\mathrm{N} - \mathrm{Y_j})(\mathrm{c}\frac{\mathrm{Y_j}}{\mathrm{N}} + \frac{\mathrm{d}}{\mathrm{M}} \sum_{\mathrm{i}=1}^{\mathrm{M}} \frac{\mathrm{Y_i}}{\mathrm{N}}) \\ \underline{Y} & \to \underline{Y} - \underline{e}_j \text{ at rate } \quad \gamma \mathrm{Y_j} \end{cases}$$

To start with, we recall two results of approximation of the stochastic model with a deterministic one: in the first case, let the number of individuals in each site go to infinity, keeping the number of sites fixed; in the second formulation, let the number of sites go to infinity, with fixed population for each household.

- Let $X_N(t)$ be the Markov process, whose components represent the fraction of infectives in each household

$$\underline{X}_N(t) = \left(\frac{Y_1(t)}{N}, \ldots, \frac{Y_M(t)}{N} \right).$$

$\underline{X}_N$ is the density process associated to $\underline{Y}(t)$. A result of [EK86] for density-dependent Markov processes shows that if $\underline{X}_N(0) \to \underline{X}^0$ a.s. for $N \to \infty$, then

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \|\underline{X}_N(t) - \underline{X}(t)\| = 0 \qquad a.s.$$

where the vector function $\underline{X} = (X_1, \ldots X_M)$ satisfies the following ODE system

$$\begin{cases} \dot{X}_i(t) = (1 - X_i(t))(cX_i(t) + \frac{d}{M} \sum_{j=1}^{M} X_j(t)) - \gamma X_i(t) \\ X_i(0) = X_i^0 \end{cases} \tag{3.1}$$

for $i = 1, \ldots M$

- As $M \to \infty$, it is useful to consider the fraction of households with j infectives:

$$\begin{aligned} \underline{x}^M(t) &= (x_0, \dots, x_N) \\ x_j^M(t) &= \frac{\sum_{j=1}^M \chi_{[Y_i=j]}(t)}{M} \end{aligned}$$

$\{\underline{x}^M(t), t \in [0,\infty)\}$ is a continuous Markov chain with countable state-space

$$\left([0,1] \cap \frac{\mathbb{Z}}{M}\right)^{N+1}$$

and transition rates

$$\begin{cases} \underline{x}^M & \to \underline{x}^M + \frac{1}{M}(\underline{e}_{j+1} - \underline{e}_j) \quad \text{at rate} \quad \mathrm{M(N-j)x_j(c\frac{j}{N} + \frac{d\sum_{l=1}^{N} lx_l}{N})} \\ \\ \underline{x}^M & \to \underline{x}^M + \frac{1}{M}(\underline{e}_{j-1} - \underline{e}_j) \quad \text{at rate} \quad \mathrm{M\gamma j x_j} \\ \end{cases}$$

The results of *Barbour* [BK93] and *Kurtz* [EK86] show that assuming that $\underline{x}^M(0)$ converges a.s., as M goes to infinity, to $y \in [0,1]^{N+1}$ with $\sum_{j=1}^{N} y_j = 1$, the stochastic process $\underline{x}^M(t)$ converges a.s. to $\xi = (\xi_0, \dots, \xi_N)$, the solution of the following deterministic ODE system:

$$\begin{cases} \dot{\xi}_j(t) &= \gamma[(j+1)\xi_{j+1} - j\xi_j] + \\ & \quad +c\left[(N-j+1)\frac{j-1}{N}\xi_{j-1} - (N-j)\frac{j}{N}\xi_j\right] + \\ & \quad +d\frac{\sum_{l=0}^N \xi_l l}{N}\left[(N-j+1)\xi_{j-1} - (N-j)\xi_j\right] \\ \\ \dot{\xi}_0(t) &= \gamma\xi_1 - \xi_0 d\sum_{l=0}^N \xi_l l \\ \xi_j(0) &= y_j \end{cases} \tag{3.2}$$

where $0 \le \xi_j(t) \le 1, \quad \xi_j \in C[0,T], \quad j = 0, \dots, N$, and $\xi_{N+1}(t) \equiv 0$.

3.3 Weak limit equation

Our aim is to study the behaviour of the process as both parameters go to infinity. In order to do that, we construct a family of probability measures on the interval

$[0,1]$ depending on the time parameter t; for every measurable set $I \subset [0,1]$, let us define

$$Y^{N,M}(t)(I) := \frac{\sum_{i=1}^{M} \xi_{[\frac{Y_i(t)}{N} \in I]}}{M} = \frac{1}{M} \sum_{i=1}^{M} \delta_{\frac{Y_i}{N}}(I). \tag{3.3}$$

If $I = (a, b]$, $Y^{N,M}(t)(I)$ is the percentage of households with a fraction of infectives between a and b at time t; if $a = \frac{k-1}{N}$and $b = \frac{k}{N}$, we have $Y^{N,M}(t)(I) = x_k^m(t)$.
$Y^{N,M}(t)$ is a stochastic process indexed by two parameters and takes values in the space of probability measures on the unit interval; with the notation of the previous section, we get

$$Y^{N,M}(t)(I) = \sum_{j=0}^{N} x_j^M(t) \delta_{\frac{j}{N}}(I).$$

The following sections deal with two limits:

A) $\lim_{M\to\infty} \lim_{N\to\infty} Y^{N,M}(t)$

B) $\lim_{N\to\infty} \lim_{M\to\infty} Y^{N,M}(t)$

They could a 'priori' be different, but it will be shown that they both lead to the same measure on the interval $[0,1]$, whose evolution is described by the following non linear 'weak' equation

$$< f, \mu(t) >=< f, \mu_0 > + \int_0^t < H_{\mu(s)} f, \mu(s) > ds \qquad \forall f \in C^1([0,1]) \tag{3.4}$$

where we denote the action of the generic probability measure μ on a continuous function $f : [0,1] \to R$ with $< \cdot, \cdot >$

$$< f, \mu >= \int_0^1 f(x) \mu(dx)$$

and

$$\begin{aligned} Rf(x) &= \gamma x f'(x) \\ Sf(x) &= cx(1-x)f'(x) \\ Tf(x) &= d(1-x)f'(x) \\ E(\mu) &= \textstyle\int_0^1 x\mu(dx) \\ H_\mu f &= Rf + Sf + E(\mu)Tf. \end{aligned}$$

In order to show the existence and uniqueness of a solution of the equation (3.4), let us first consider the linear equation

$$< f, \mu(t) >=< f, \mu(0) > + \int_0^t < r(s,\cdot)f'(\cdot), \mu(s) > ds \qquad \forall f \in C^1([0,1]) \tag{3.5}$$

where $r(t,x)$ is a continuous function defined on $[0,T]\times[0,1]$. Formally, we can write $\partial_t\mu = A^\star(t)\mu$ where $A^\star$ denotes the adjoint of the time dependent operator $(A(t)f)(x) = r(t,x)f'(x)$.
Let us denote by $\Phi(t;t_0,x)$ the solution to the Cauchy problem

$$\begin{cases} \dot{y}(t) & = r(t,y(t)) \\ y(t_0) & = x \end{cases} \tag{3.6}$$

and assume that this solution is defined for $t\in[t_0,+\infty)$ and $x\in[0,1]$.

Lemma 3.3.1 *Let $E\in C([0,1])$ with $0\le E(t)\le 1$, $x\in[0,1]$ and let $\phi^E(t,x)$ be the solution of (3.6) with $t_0=0$ and $r(t,y)=cy(1-y)-\gamma y+E(t)d(1-y)$.
Then $0\le\phi^E(t,x)\le 1 \quad ,x\in[0,1]$.*

PROOF. The equation $r(t,y)=0$ defines a function $\alpha(t)$ such that $0\le\alpha(t)<1$ and

$$\begin{aligned} A_+ &:= \{(t,y) : t\in\mathbb{R}_+, y<\alpha(t)\}, \quad r_{|A_+}>0 \\ A_- &:= \{(t,y) : t\in\mathbb{R}_+, y>\alpha(t)\}, \quad r_{|A_+}<0. \end{aligned}$$

Now $\phi^E(t,x)$ is increasing in t if $(t,\phi^E(t,x))\in A_+$ and is decreasing if $(t,\phi^E(t,x))\in A_-$. Thus $0\le\phi^E(t,x)\le 1$
As in the case of an absolutely continuous measure, we can construct the measure solution of (3.5) carrying the initial μ_0 along the flow: for every measurable set A in $\Omega=[0,1]$, let

$$\begin{aligned} \phi_t(x) &= \Phi(t;0,x) \\ \psi_t(x) &= \Phi(0;t,x) \\ \mu_t(A) &= \mu_0\left[\phi_t^{-1}(A)\cap\Omega\right] \end{aligned} \tag{3.7}$$

Since the solution of the (3.6) is unique , we have

$$\Phi(t,s,\Phi(s,t,x)) = x \tag{3.8}$$

We take the derivative with respect to t and to x of both sides of (3.8) ($\partial_i,\ i=1,2,3$ denotes the partial derivative of Φ with respect to the $i-th$ variable)

$$\begin{aligned} (\partial_1\Phi)(t,s,\Phi(s,t,x)) + (\partial_3\Phi)(t,s,\Phi(s,t,x))(\partial_2\Phi(s,t,x)) &= 0 \\ (\partial_3\Phi(t,s,\Phi(s,t,x))(\partial_3\Phi(s,t,x)) &= 1 \end{aligned}$$

Multiplying the first equality by $\partial_3\Phi(s,t,x)$ and recalling that

$$\partial_1\Phi(t,s,\Phi(s,t,x)) = r(t,\Phi(t,s,\Phi(s,t,x)) = r(t,x),$$

we have

$$\partial_2\Phi(s,t,x) + r(t,x)\partial_3\Phi(s,t,x) = 0 \tag{3.9}$$

and for $s = 0$

$$\partial_x \psi_t(x) r(t,x) + \partial_t \psi_t(x) = 0 \tag{3.10}$$

Lemma 3.3.2 *The equation (3.5) has a unique solution.*

PROOF.

Note that the one-parameter family satisfies (3.5): for every measurable function f defined on Ω

$$\int_0^1 f(x)\mu_t(dx) = \int_0^1 f(\phi_t(y))\mu_0(dy)$$

(see for example [DS88],Th.III.10.8); hence

$$\begin{aligned}
\tfrac{d}{dt} < f, \mu_t > &= \tfrac{d}{dt}\left[\int_0^1 f(\phi_t(y))\mu_0(dy)\right] \\
&= \int_0^1 \tfrac{\partial}{\partial t} f(\phi_t(y))\mu_0(dy) = \\
&= \int_0^1 f'(\phi_t(y))\tfrac{\partial \phi_t(y)}{\partial t}\mu_0(dy) \\
&= \int_0^1 f'(\phi_t(y)) r(t, \phi_t(y))\mu_0(dy) \\
&= \int_0^1 f'(x) r(t,x)\mu_t(dx).
\end{aligned}$$

Let ν_t be a solution of the equation (3.5). First note that, $\forall g \in C^1([0,T] \times [0,1])$

$$\frac{d}{dt} < g(t,\cdot), \nu_t >= \int_0^1 \frac{\partial}{\partial t} g(t,y)\nu_t(dy) + \int_0^1 \frac{\partial}{\partial y} g(t,y) r(t,y)\nu_t(dy). \tag{3.11}$$

To show that the family of measures (3.7) is the only solution of equation (3.5), let us construct the family of measures on Ω

$$\lambda_t(B) = \nu_t\left[\psi_t^{-1}(B) \cap \Omega\right]$$

For all $f \in C^1([0,1])$

$$\begin{aligned}
\tfrac{d}{dt} < f, \lambda_t > &= \tfrac{d}{dt}\int_0^1 f(\psi_t(y))\nu_t(dy) = \\
&= \int_0^1 \left[\tfrac{\partial}{\partial t} f(\psi_t(y))\right]\nu_t(dy) + \int_0^1 \left[\tfrac{\partial}{\partial y}\left(f(\psi_t(y))\right)\right] r(t,y)\nu_t(dy) = \\
&= \int_0^1 f'(\psi_t(y))\partial_t \psi_t(y)\nu_t(dy) + \int_0^1 f'(\psi_t(y))\partial_y(\psi_t(y)) r(t,y)\,\nu_t(dy) \\
&= \int_0^1 f'(\psi_t(y))\left[\partial_y \psi_t(y) r(t,y) + \partial_t \psi_t(y)\right]\nu_t(dy) = 0
\end{aligned}$$

because of (3.10). Hence, $\lambda_t = \lambda_0 \quad \forall t \in [0,T]$; for all measurable sets B ,

$$\begin{aligned}
\nu_t(B) &= \nu_t\big(\psi_t^{-1}(\phi_t^{-1}(B)) \cap \Omega\big) = \lambda_t\big(\phi_t^{-1}(B)\big) \\
&= \lambda_0\big(\phi_t^{-1}(B)\big) = \nu_0\big(\psi_0^{-1}(\phi_t^{-1}(B) \cap \Omega)\big) \\
&= \nu_0\big(\phi_t^{-1}(B) \cap \Omega\big) = \mu_0\big(\phi_t^{-1}(B) \cap \Omega\big) = \mu_t(B)
\end{aligned}$$

In our original weak equation, the function r depends on the measure μ but we can apply the previous argument to construct a contraction map with the usual sup-norm $\|\cdot\|$ on $C([0,T])$: a fixed point of this map is the solution of our original problem.

Theorem 3.3.3 *Let us consider $\mathcal{M} = \{E \in C([0,T]), 0 \le E(t) \le 1\}$ equipped with the supremum norm , and let μ_t^E be the solution of (3.5) where*

$$r(t,y) = cy(1-y) - \gamma y + d(1-y)E(t).$$

If $T < T_0$, with T_0 such that $T_0 \exp((|c-\gamma|+1)T_0) = 1$, the map $\mathbf{T} : \mathcal{M} \to \mathcal{M}$

$$\mathbf{T}(E)(t) = \int_0^1 x\mu_t^E(dx).$$

is a contraction.

PROOF. Let $E, F \in \mathcal{M}$ and denote $Q_T = [0,1] \times [0,T]$

$$\begin{aligned}
|\phi^E(t,x) - \phi^F(t,x)| &= \left| \int_0^t r^E(s,\phi^E(s,x)) - r^F(s,\phi^F(s,x))\, ds \right| \\
&\le \int_0^t \left| r^E(s,\phi^E(s,x)) - r^F(s,\phi^E(s,x)) \right| ds \\
&\quad + \int_0^t \left| r^F(s,\phi^E(s,x)) - r^F(s,\phi^F(s,x)) \right| ds \\
&\le \left(\sup_{Q_T} |r^E - r^F| \right) t + \left(\sup_{Q_T} \left| \frac{\partial r^F}{\partial x} \right| \right) \int_0^t |\phi^E(s,x) - \phi^F(s,x)|\, ds \\
&\le d \sup_{[0,T]} |E(t) - F(t)| t + (|c-\gamma|+1) \int_0^t |\phi^E(s,x) - \phi^F(s,x)|\, ds
\end{aligned}$$

By Gronwall inequality

$$|\phi^E(t,x) - \phi^F(t,x)| \le dt\|E-F\| \exp[(|c-\gamma|+1)t]$$

Hence

$$\left| (\mathbf{T}(E) - \mathbf{T}(F))(t) \right| = \left| \int_0^1 [\phi^E(t,x) - \phi^F(t,x)]\, \mu_0(dx) \right| \le dt\|E-F\| \exp(|c-\gamma|+1)t$$

for all $t \in [0,T]$

By iterating the previous procedure, it follows easily that the solution of the equation (3.4) exists on any compact time interval.

3.4 First limit

If $\lim_{N\to\infty} \underline{X}_N(0) = \underline{X}^0$ *a.s.* then weakly

$$\lim_{N\to\infty} Y^{N,M}(t) = \lim_{N\to\infty} \frac{1}{M}\sum_{i=1}^{M} \delta_{\frac{Y_i}{N}(t)} = \frac{1}{M}\sum_{i=1}^{M} \delta_{X_i^M(t)}$$

where $\{X_j^M\}_{j=1}^M$ is the solution of the ODE system (3.1) Hence, we work with the following probability measures for $t \in [0,T]$

$$\mu_t^M(I) = \frac{1}{M}\sum_{i=1}^{M} \delta_{X_i^M(t)} \tag{3.12}$$

Since the trajectories X_i^M belong to the space of continuous functions, we can consider the sequence of probability measures μ_M on the space $C([0,T])$, each with a finite support consisting of the M functions X_i^M: for every Borel set A of the space $C([0,T])$

$$\mu_M(A) = \frac{1}{M}\sum_{i=0}^{M} \delta_{X_i^M}(A).$$

Our aim is to prove that the sequence μ_M is tight; tightness implies the relative compactness of the sequence via Prokhorov theorem [Bal99].
We remind ecall that a sequence of probability measures P_N on $C([0,T])$ is tight iff the next two conditions hold

i) For each positive η, there exists an a such that

$$P_N\{x : |x(0)| > a\} \le \eta \qquad N \ge 1$$

ii) For each positive ϵ and η, there exist a δ and an integer N_0 such that

$$P_N\{x : \sup_{|s-t|<\delta} |x(s) - x(t)| \ge \epsilon\} \le \eta, \quad N \ge N_0$$

Lemma 3.4.1 *The sequence of probability measures $\{\mu_M\}$ is tight*

PROOF. The first condition, which is equivalent to the tightness of the initial finite -dimensional measure is automatically fulfilled because we deal with measures on a compact space. Since the functions X_i^M are Lipschitz with a uniform Lipschitz constant

$$\begin{aligned} |\dot{X}_i(t)| &\le c|1 - X_i(t)|\,|X_i(t)| + |1 - X_i(t)|\tfrac{d}{M}\textstyle\sum_{i=0}^{M} |X_i^M| + \gamma|X_i(t)| \\ &\le c + d + \gamma, \end{aligned}$$

the second condition is satisfied.
For every M, $\mu_M(t)$ satisfies the weak equation :

$$
\begin{aligned}
< f,\mu_M(t) > - < f,\mu_M(0) >= & \int_0^t \frac{d}{ds} < f,\mu_M(s) > ds \\
= & \int_0^t \sum_{j=1}^{M} \frac{1}{M}\frac{d}{ds} f(X_i^M(s))\, ds \\
= & \int_0^t \sum_{j=1}^{M} \frac{1}{M} f\prime(X_i^M(s))[(1-X_i(s))(cX_i(s)+\frac{d}{M}\sum_{j=1}^{M} X_j(s)) - \gamma X_i(s)]\, ds \\
= & \int_0^t (c < Sf,\mu_M(s) > +dE(\mu_M) < Tf,\mu_M(s) > -\gamma < Rf,\mu_M(s) >)\, ds
\end{aligned}
$$

for all $f \in C^1([0,1])$.
It is easy to see that if $\mu_M(0) \to \mu(0) \text{as} M \to \infty$, the limit of every convergent subsequence of measures satisfies the equation (3.4); because of the uniqueness of the solution of (3.4), we have that the whole sequence μ_M has μ as a weak limit.

3.5 Second limit

First of all, according to the result of section (3.2), if $\lim_{M\to\infty} \underline{x}^M(0) = \underline{y}$ we have that

$$\lim_{M\to\infty} \sum_{j=0}^{N} x_j^M(t)\delta_{\frac{j}{N}} = \sum_{j=0}^{N} \xi_j^N(t)\delta_{\frac{j}{N}}$$

where $\{\xi_j^N(t)\}_{j=0}^N$ satisfies (3.2). Let $E^N(t) = \sum_{j=0}^{N} \xi_j^N(t)\frac{j}{N}$, and define the sequence of probability measures

$$\mu_N(t) = \sum_{j=0}^{N} \xi_j^N(t)\delta_{\frac{j}{N}}.$$

In order to prove the convergence for $N \to \infty$ of this family of measures, we give a probabilistic interpretation of such a measure: we construct a time continuous Markov chain $Z^N(t)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with inhomogeneous transitions and state-space $S_N = \{0, 1, ...N\}$: $\mu_N(t)$ will be the one-dimensional distribution of that process.
First of all, for every $k \in S_N$ we can construct a Markov process $J_k^N(t)$ on a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ characterized by

$$
\begin{aligned}
\mathbb{P}[J_k^N(t+h) = j+l | J_k^N(t) = j] &= q^N(j, j+l, t)\, h + o(h) \qquad l \neq 0 \\
\mathbb{P}[J_k^N(0) = k] &= 1 \qquad k \in S_N
\end{aligned}
\tag{3.13}
$$

with transition rates

$$\begin{aligned} q^N(j,j+1,t) &= c(N-j)\tfrac{j}{N} + d(N-j)E^N(t) \\ &= N\left[c(1-\tfrac{j}{N})\tfrac{j}{N} + d(1-\tfrac{j}{N})E^N(t)\right] \\ q^N(j,j-1,t) &= \gamma j = \gamma N(\tfrac{j}{N}) \\ q^N(j,j+l,t) &= 0 \qquad |l|>1 \end{aligned}$$

The functions $q^N(j,j+l,t)$ satisfies the hypothesis of Theorem B.2 with

$$\begin{aligned} \beta^N(x,1,t) &= c(1-x)x + d(1-x)E^N(t) \\ \beta^N(x,-1,t) &= \gamma x \\ \beta^N(x,l,t) &= 0 \qquad |l|>1 \\ F^N(x,t) &= c(1-x)x + d(1-x)E^N(t) - \gamma x \end{aligned}$$

One can easily show that

Lemma 3.5.1 *The family of function $E^N(t)$ is equicontinuous and uniformly bounded on the compact time interval $[0,T]$.*

If we restrict ourselves to a subsequence, we have that $E(t) = \lim_{N\to\infty} E^N(t)$, hence setting $F(x,t) = \lim_{N\to\infty} F^N(x,t) = c(1-x)x + d(1-x)E(t) - \gamma x$, we have $\lim_{N\to\infty} F^N(x,t) = F(x,t)$ and we can apply Theorem B.2 of the Appendix, obtaining

Corollary 3.5.2 *Let $x_N = \frac{J_k^N(0)}{N}$ and $Z_{x_N}^N(t) = \dfrac{J_{Nx_N}^N(t)}{N}$. If $\lim\limits_{N\to\infty} x_N = x$, then*

$$\lim_{N\to\infty} \sup_{t\in[0,T]} |Z_{x_N}^N - Z_x| = 0 \qquad a.e.$$

where Z_x satisfies

$$Z_x(t) = x + \int_0^t F\left(Z_x(s), s\right)\, ds \tag{3.14}$$

Let us suppose that the sequence of measures μ_0^N converges in distribution to a probability measure μ_0 on $[0,1]$. First we have the result ([Bil95], Theorem 25.6, page 333)

Lemma 3.5.3 *If $\lim\limits_{N\to\infty} \mu_0^N = \mu_0$ then there exist random variables Z_0^N, Z_0 on a common probability space Ω'', such that Z_0^N has distribution μ_0^N, Z_0 has distribution μ_0, and $\lim\limits_{N\to\infty} Z_0^N = Z_0$ for all $\omega'' \in \Omega''$.*

We can consider now the processes

$$Z(t,\omega'') := Z_{Z_0(\omega'')}(t) \text{ and } Z_N(t,\omega',\omega'') = Z_{Z_0^N(\omega'')}(t,\omega')$$

on the probability space $\Omega = \Omega' \otimes \Omega''$ with the product measure. From Corollary 3.5.2 and Lemma 3.5.3, it follows

Lemma 3.5.4 *The measure $\mu_N(t)$ is the one-dimensional distribution associated to the process $Z_N(t)$ and for almost every $\omega' \in \Omega'$ $\lim_{N\to\infty} Z^N(t,\omega',\omega'') = Z(t,\omega'')$ solution of*

$$Z(t,\omega'') = Z_0(\omega'') + \int_0^t F(Z(s,\omega''),s)\,ds \qquad \text{for all } \omega'' \in \Omega''. \tag{3.15}$$

The one-dimensional distribution $\mu(t)$ of $Z(t)$ is the requested limit measure. From the a.s. convergence, we have the weak convergence of the sequence of measures $\mu_N(t)$, therefore for every $f \in C([0,1])$

$$\lim_{n\to\infty} \int_0^1 f(x)\mu_t^N(dx) = \int_0^1 f(x)\mu_t(dx)$$

If we choose the function $f(x) = x$, we have that

$$E(t) = \lim_{N\to\infty} E^N(t) = \lim_{N\to\infty} \int_0^1 x\mu_t^N(dx) = \int_0^1 x\mu_t(dx)$$

The family of measures $\mu_N(t)$ satisfies the following equation

$$< f, \mu_N(t) > = < f, \mu_N(0) > + \int_0^t \sum_{j=0}^N \dot{\xi}_j(s) f(\frac{j}{N})\,ds \qquad \forall f \in C([0,1])$$

and

$$\begin{aligned}
\textstyle\sum_{j=0}^N \dot{\xi}_j(t) f(\frac{j}{N}) &= \gamma \textstyle\sum_{j=0}^N f(\frac{j}{N})[(j+1)\xi_{j+1} - j\xi_j] + \\
&\quad + c\textstyle\sum_{j=0}^N f(\frac{j}{N})\left[(1-\frac{j-1}{N})(j-1)\xi_{j-1} - (1-\frac{j}{N})j\xi_j\right] + \\
&\quad + d\left[\textstyle\sum_{l=0}^N(\frac{l}{N})\xi_l \sum_{j=0}^N f(\frac{j}{N})(1-\frac{j-1}{N})\xi_{j-1} - (1-\frac{j}{N})\xi_j\right] \\
&= -\gamma\textstyle\sum_{j=0}^N \nabla_-^N f(\frac{j}{N})\frac{j}{N}\xi_j + c\sum_{j=0}^N(1-\frac{j}{N})\frac{j}{N}\xi_j\nabla_+^N f(\frac{j}{N}) + \\
&\quad + E^N(t)\textstyle\sum_{j=0}^N(1-\frac{j}{N})\nabla_+^N f(\frac{j}{N})
\end{aligned}$$

where

$$\nabla_+^N f(x) = \frac{f(x+\frac{1}{N}) - f(x)}{\frac{1}{N}} \qquad \nabla_-^N f(x) = \frac{f(x) - f(x-\frac{1}{N})}{\frac{1}{N}}$$

Hence we have

$$\begin{aligned} < f, \mu_N(t) > - < f, \mu_N(0) >= \int_0^t [& - \gamma < \nabla_-^N f(\cdot)\cdot, \mu_N(s) > \\ & + c < (1-\cdot)\cdot \nabla_+^N f(x), \mu_N(s) > \\ & + E^N(s) < (1-\cdot)\nabla_+^N f(x), \mu_N(s) >] \, ds \end{aligned} \tag{3.16}$$

If we choose a function $f \in C^2([0,1])$ (dense subspace of $C([0,1])$), we can pass to the limit in the equation (3.16), using

$$\begin{aligned} \sup_{x\in[0,1]} \left|(f'(x) - \nabla_-^N f(x))x\right| & \leq \sup_{x\in[0,1]} |f''(x)| \tfrac{1}{N} \\ \sup_{x\in[0,1]} \left|(f'(x) - \nabla_+^N f(x))(1-x)\right| & \leq \sup_{x\in[0,1]} |f''(x)| \tfrac{1}{N}; \end{aligned}$$

this gives

$$\begin{aligned} \big|< \nabla_-^N f(\cdot)\cdot, \mu_N > & + < f'(\cdot)\cdot, \mu >| \leq \\ & \leq \big|< \nabla_-^N f(\cdot)\cdot + f'(\cdot)\cdot, \mu_N > - < f'(\cdot)\cdot, \mu_N - \mu >\big| \\ & \leq \|\mu_n\| \, \|f''\| \tfrac{1}{N} + | < f'(\cdot)\cdot, \mu_n - \mu > | \end{aligned}$$

and

$$\big|< \nabla_+^N f(\cdot)\cdot, \mu_N > - < f'(\cdot)\cdot, \mu >\big| \leq \|\mu_n\| \, \|f''\| \frac{1}{N} + | < f'(\cdot)(1-\cdot), \mu_n - \mu > |.$$

The equation for the measure $\mu(t)$ is again the equation (3.4) with

$$\mu_0 = \lim_{N\to\infty} (\sum_{j=0}^{N} \xi_j(0) \delta_{\frac{j}{N}}).$$

Since there exists a unique solution of (3.4), any subsequence of $\mu_N(t)$ converges to $\mu(t)$ and we have that the whole sequence $\mu^N(t)$ has $\mu(t)$ as a weak limit.

3.6 Asymptotic behaviour of the limit measure

We search for an equilibrium measure for the equation (3.4) as a delta measure $\mu^\star = \delta_{x^\star}$ with support $x^\star \in (0,1)$. Since $E(\mu^\star) = x^\star$, $< H_{\mu^\star} f, \mu^\star >= 0$ for all $f \in C^1([0,1])$ if and only if we have

$$G(x^\star) := c(1-x^\star) + d(1-x^\star) - \gamma = 0 \tag{3.17}$$

The previous equation has a unique solution $x^\star \in (0,1)$ if and only if $c + d > \gamma$. We notice that, in this case, $G(z) > 0$ if and only if $z < x^\star$.
In order to study the asymptotic behaviour of the measure $\mu(t)$ solution of (3.4),

we can take advantage of its representation as the one-dimensional measure of the stochastic process $Z(t)$, which is the solution of the following Cauchy problem (as one can easily see from (3.15))

$$\begin{cases} Z'(t) &= cZ(t)(1-Z(t)) + d(1-Z(t))E(t) - \gamma Z(t) := F(Z(t),t) \\ Z(0) &= Z_0 \end{cases} \tag{3.18}$$

where Z_0 is a random variable with distribution μ_0 and $E(t) = \int_0^1 x\mu_t(dx)$.
We consider the solutions $Z_m(t)$ and $Z_M(t)$ of the Cauchy problems

$$\begin{cases} Z'_m(t) &= F(Z_m(t),t) \\ Z_m(0) &= 0 \end{cases} \qquad \begin{cases} Z'_M(t) &= F(Z_M(t),t) \\ Z_M(0) &= 1 \end{cases}$$

They bound the solution $Z(t)$, that is $0 \le Z_m(t) \le Z(t,\omega'') \le Z_M(t) \le 1$; hence $Z_m(t) \le E(t) \le Z_M(t)$

Proposition 3.6.1 *Let $x^\star$ be the unique positive solution of (3.17) if $c+d > \gamma$ and $x^\star = 0$ if $c+d \le \gamma$; both $Z_m(t)$ and $Z_M(t)$ go to $x^\star$ for $t \to +\infty$.*

PROOF. If $x^\star = 0$, $Z'_M(t) \le Z_M(t)G(Z_M(t))$ and the conclusion is trivial.
Let $x^\star > 0$. If there exists t_0 such that $Z_m(t) < x^\star$, then, using $E(t) \ge Z_m(t)$,

$$Z'_m(t) = F(Z_m(t),t) \ge Z_m(t)\, G(Z_m(t)) \ge 0$$

$Z_m(t)$ is not decreasing in t, so $\lim\limits_{t\to\infty} Z_m(t) = m^\star \le x^\star$. If $m^\star < x^\star$ we have $Z'_m(t) \ge m^\star G(m^\star) > 0$, in contrast with the existence of a finite limit; so $m^\star = x^\star$.
Analogously, we can prove that, if $Z_M \ge x^\star$, $\quad t \in [t_0,+\infty)$, then $\lim\limits_{t\to\infty} Z_M(t) = x^\star$.
Assume now that there exists τ such that $x^\star = Z_m(\tau) < Z_M(\tau)$, $E(\tau) > x^\star$, then

$$Z'_m(\tau) = F(\mathbb{Z}_m(\tau),\tau) > x^\star G(x^\star) = 0. \tag{3.19}$$

Hence $Z_m(t) > x^\star$ for $t \in (\tau, \tau+\epsilon)$; if there were $\tau' > \tau$ such that $Z_m(\tau') = x^\star$ and $Z_m(t) > x^\star$, $\quad t \in (\tau,\tau')$ we could again obtain (3.19) reading a contradiction; we will then have $Z_m(t) > x^\star$ for $t > \tau$. Since $x^\star < Z_m(t) \le Z_M(t)$, $\quad t \in (\tau,+\infty)$, we can conclude as before that $\lim_{t\to\infty} Z_M(t) = x^\star$ and from $x^\star < Z_m(t) < Z_M(t)$, also that $\lim\limits_{t\to\infty} Z_m(t) = x^\star$. Analogously the case where there exists τ such that $E(\tau) < x^\star$, $Z_m(\tau) < Z_M(\tau) = x^\star$.

Since the support of the measure $\mu(t)$ is contained in the interval $(Z_m(t), Z_M(t))$, we can easily conclude

Theorem 3.6.2 *As $t \to \infty$, the measure $\mu(t)$ tends weakly to the Dirac measure $\delta_{x^\star}$.*

Appendix A

M - matrices

In this appendix, we recall some definitions and propositions about the so-called M-matrices (for detailed references, see [BP94] and [Win89]).

We assume that all matrices under consideration are real and square of order $n \geq 2$.

Definition A.1 *A matrix $A = (a_{ij})$ is said to be strongly row diagonally dominant if*

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \qquad \forall i.$$

If the transpose A^T is strongly row diagonally dominant, then A is said to be strongly column diagonally dominant.

Proposition A.2 *If A is strongly row or column diagonally dominant matrix, then $detA \neq 0$*

Let

$$Z^{n \times n} = \{A = (a_{ij}) \in \mathbb{R}^{n \times n} \ : \ a_{ij} \leq 0, i \neq j\}$$

The M matrices are a special subclass of matrices in $Z^{n \times n}$.

Definition A.3 *Any matrix of the form*

$$A = sI - B, \quad s > 0, \quad B \geq 0 \tag{A.1}$$

for which $s \geq r(B)$, the spectral radius of B, is called an M-matrix.

The non-singular M-matrices are those of the form (A.1) for which $s > r(B)$).

The matrices $B_\lambda^{(m)}$ involved in Proposition 2.6.6, equation (2.37) are L-matrices

Definition A.4 *A matrix $A = (a_{ij})$ is called an L-matrix if*

$$a_{ii} > 0, \forall i, \qquad and \qquad a_{ij} \leq 0, \ i \neq j.$$

Since any non-singular M-matrix have all positive diagonal entries, such an M-matrix is an L-matrix; the converse is in general not true, but

Theorem A.5 *Let A be an L-matrix which is strongly row or column diagonally dominant. Then A is a non-singular M-matrix.*

Theorem A.6 *Let $A \in Z^{n\times n}$; A is a non-singular M-matrix if and only if A is inverse-positive, that is A^{-1} exists and*

$$A^{-1} \geq 0$$

Appendix B

Random time change

Our aim is to give a representation of a pure jump Markov process as a sum of independent Poisson processes. We apply some results due to Kurtz [EK86], which can be used also for some kinds of non-homogeneous transition rates.
Let $B(t)$ be the operator defined on $\mathcal{B}_0(\mathbb{Z})$, the space of bounded functions on the integers , that vanish off a compact set

$$B(t)f(x) = \sum_{l\in\mathbb{Z}} \beta_l(x,t)\left(f(x+l) - f(x)\right) \tag{B.1}$$

where $\beta_j : \mathbb{Z} \times [0,T]$ are functions such that the martingale problem for $B(t)$ is well posed (see for instance [EK86], Theorem 7.3, page 223).
We can state a slight modification of [EK86], Theorem 4.1, page 236

Theorem B.1 *Let $E = \mathbb{Z} \cup \Delta$, the one-point compactification of $\mathbb{Z}$. If $J(t)$ is a solution of the martingale problem for $B(t)$ with sample paths in $D_E([0,+\infty))$, $J(0) = x$, $\tau_\infty = \{\inf t : J(t) = \Delta\}$ and Y_j are independent Poisson processes then*

$$\begin{aligned} J(t) &= x + \sum_{j\in\mathbb{Z}} jY_j\left(\int_0^t \beta_j(J(s),s)\,ds\right) && t < \tau_\infty \\ J(t) &= \Delta && t \geq \tau_\infty \end{aligned}$$

This representation is useful for investigating the asymptotic behaviour of this kind of processes. We deal with a family of continuous-time Markov chains $\{J^N(t)\}_{N\in\mathbb{N}}$ with values in $S_N = \{0,1,2,\ldots,N\} \subset \mathbb{Z}$ and transition rates $q^N(j,k,t),\quad j,k \in S_N, t \in [0,T]$. According to definition 3.1 of [Pol90], the family $\{J^N(t)\}_{N\in\mathbb{N}}$ is asymptotically density dependent if there exists an open set $E \in \mathbb{R}$ and a family of continuous functions $\{\beta^N\} : E \times \mathbb{Z} \times [0,T]$ such that

$$q^N(k,k+l,t) = N\beta^N(\frac{k}{N},l,t) \qquad l \neq 0,$$

$\sum_l l\beta^N(x,l,t)$ converges for all $(x,t) \in E \times [0,T]$ and there exists a function F such that $F^N(x,t) := \sum_l l\beta^N(x,l,t)$ converges to $F(x,t)$ on $E \times [0,T]$. We recall the theorem 3.1 of [Pol90]

Theorem B.2 *If* $|F(x,t) - F(y,t)| < M|x-y|$ *and for all* N

$$\sup_{(x,t)\in E\times[0,T]} \sum |l|\beta_N(x,t,l) < \infty$$
$$\lim_{\delta\to\infty} \sup_{(x,t)\in E\times[0,T]} \sum_{l:|l|>\delta} |l|\beta_N(x,t,l) = 0$$

$$\lim_{N\to\infty} \sup_{E\times[0,T]} |F^N(x,t) - F(x,t)| = 0$$

and $\lim_{N\to\infty} \frac{J^N(t)}{N}(0) = x$, *then*

$$\lim_{N\to\infty} \sup_{[0,T]} \left| \frac{J^N(t)}{N}(t) - Z_x(t) \right| = 0 \qquad a.e.$$

where

$$Z_x(t) = x + \int_0^t F(Z_x(s), s)\, ds$$

Bibliography

[AB00] Håkan Andersson and Tom Britton. Stochastic epidemics in dynamic populations:quasi-stationarity and extinction. *J. Math. Biol.*, 41:559–580, 2000.

[And91] William J. Anderson. *Continuous-time Markov chains*. Springer-Verlag, New York, 1991. An applications-oriented approach.

[Bal99] Frank Ball. Stochastic and deterministic models for SIS epidemics among a population partitioned into households. *Math. Biosci.*, 156(1-2):41–67, 1999. Epidemiology, cellular automata, and evolution (Sofia, 1997).

[Bar74] Andrew D. Barbour. On a functional central limit theorem for Markov population processes. *Advances in Appl. Probability*, 6:21–39, 1974.

[Bar80] A. D. Barbour. Density dependent Markov population processes. In *Biological growth and spread (Proc. Conf., Heidelberg, 1979)*, pages 36–49. Springer, Berlin, 1980.

[Bil68] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[Bil95] Patrick Billingsley. *Probability and measure*. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[BIT91] Stavros N. Busenberg, Mimmo Iannelli, and Horst R. Thieme. Global behavior of an age-structured epidemic model. *SIAM J. Math. Anal.*, 22(4):1065–1080, 1991.

[BK93] A. D. Barbour and Maria Kafetzaki. A host-parasite model yielding heterogeneous parasite loads. *J. Math. Biol.*, 31(2):157–176, 1993.

[BMST97] Frank Ball, Denis Mollison, and Gianpaolo Scalia-Tomba. Epidemics with two levels of mixing. *Ann. Appl. Probab.*, 7(1):46–89, 1997.

[BP94] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original.

[BR97] A. T. Bharucha-Reid. *Elements of the theory of Markov processes and their applications.* Dover Publications Inc., Mineola, NY, 1997. Reprint of the 1960 original.

[DS88] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part I.* John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes.* John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[Fel66] William Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York, 1966.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* John Wiley & Sons Inc., New York, 1968.

[GH92] Mats Gyllenberg and Ilkka Hanski. Single-species metapopulation dynamics: a structured model. *Theoret. Population Biol.*, 42(1):35–61, 1992.

[GS94] Mats Gyllenberg and Dmitrii S. Silvestrov. Quasi-stationary distributions of a stochastic metapopulation model. *J. Math. Biol.*, 33(1):35–70, 1994.

[HK95] K. Hamza and F. C. Klebaner. Conditions for integrability of Markov chains. *J. Appl. Probab.*, 32(2):541–547, 1995.

[HW89] A. Hastings and C.L. Wolin. Within-patch dynamics in a metapopulation. *Ecology*, 70, 1989.

[JM99] Metz J.A.J. and Gyllenberg M. How should we define fitness in structured metapopulations? including an application to the calculation of evolutionarily stable dispersal strategies. Technical report, University of Turku, Institute of Applied Mathematics, 1999.

[KA93] M. Kretzschmar and F.R. Adler. Aggregated distributions in models for patchy populations. *Theoret. Population Biol.*, 1993.

[Kur70] Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probability*, 7:49–58, 1970.

[Kur71] T. G. Kurtz. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probability*, 8:344–356, 1971.

[Kur81] Thomas G. Kurtz. *Approximation of population processes.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1981.

[Lev69] R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.*, 1969.

[Lev70] R. Levins. Extinction. In M. Gesternhaber, editor, *Some mathematical problems in biology*, pages 75–107. American Mathematical Society, Providence, RI., 1970.

[Lin92] Torgny Lindvall. *Lectures on the coupling method.* John Wiley & Sons Inc., New York, 1992. A Wiley-Interscience Publication.

[LP] S.A. Levin and R.T. Paine. The role of disturbance in models of community structure. In *Ecosystem Analysis and Prediction.*

[LP74] S.A. Levin and R.T. Paine. Disturbance, patch formation, and community structure. *Proc.Natl. Acad. Sci. U.S.A.*, 1974.

[Luc99] C.J. Luchsinger. *Mathematical models of a parasitic disease.* PhD thesis, 1999.

[Met] J. A. J. Metz. Limit relations between metapopulation models. Manuscript.

[met97] *Metapopulation biology. Ecology, Genetics,and Evolution.* Academic Press, 1997.

[Nor98] J. R. Norris. *Markov chains.* Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.

[Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations.* Springer-Verlag, New York, 1983.

[Pol90] P. K. Pollett. On a model for interference between searching insect parasites. *J. Austral. Math. Soc. Ser. B*, 32(2):133–150, 1990.

[Pol99] P.K. Pollett. Modelling quasi-stationary behaviour in metapopulations. *Mathematics and Computers in Simulation*, 1999.

[Reu57] G. E. H. Reuter. Denumerable Markov processes and the associated contraction semigroups on l. *Acta Math.*, 97:1–46, 1957.

[RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion.* Springer-Verlag, Berlin, third edition, 1999.

[VVM96] J. Val, J. Verboom, and J.A.J. Metz. A deterministic size-structured metapopulation model. Manuscript, 1996.

[Win89] Günther Windisch. *M-matrices in numerical analysis*. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1989. With German, French and Russian summaries.