Elements of Mathematical Oncology

Franco Flandoli

Draft under construction

Contents

Introduction

The material of these notes is manifestly incomplete with respect to the present knowledge of Mathematical Oncology, naïve from the biomedical viewpoint, often not rigorous from the Mathematical side. We apologize for the mistakes and, with several experts and colleagues, for the lack of references. The only reason to write these notes in the present poor form is to provide notes for lectures and the beginning of discussion with groups of students. I thank several master and Ph.D. students in Pisa (Manuela Benedetti, Michele Coghi, Valeria De Mattei, Dario Domingo, Mario Maurelli, Giovanni Zanco and others) for sharing with me the first steps in this activity and the Ph.D. program and students in Padova for giving me the opportunity to prepare these notes and give the lectures.

Part I

Macroscopic models

Chapter 1

Introduction to macroscopic and microscopic models

1.1 Introduction to a few elements of biology of cancer

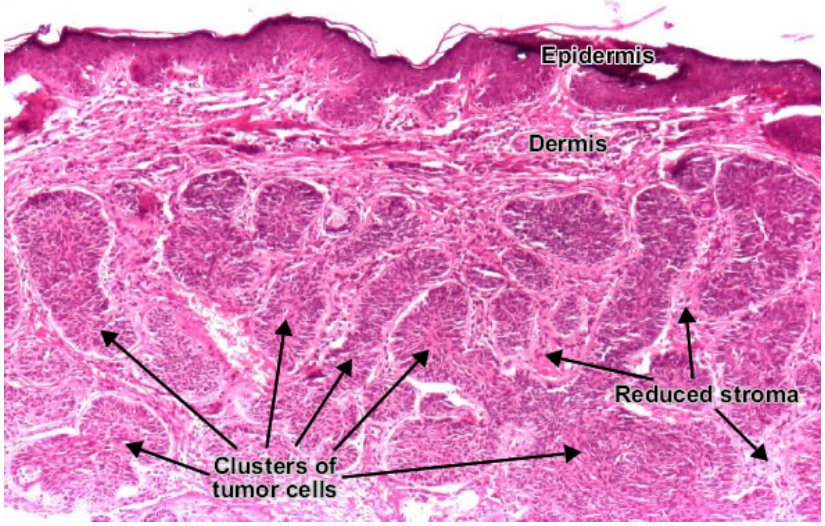
Cancer starts from a single normal cell which, due to external factors or just random mutations, undergoes a number of genetic modification which give her the characteristics of a cancel cell. The first one of these characteristics (phenotypes) is that it duplicates without need of special inputs and ignoring all alarm systems which would stop it or kill it (5-6 special genetic mutations are necessary, hence tumors arise either by chance after a considerable time of life, or due to strong external factors like radiations, smoke etc.); and the generated cells are able to duplicate again. The full cell cycle (duplication) takes an average of 16 hours. Plus some rest, we may roughly think that with a frequency higher than one day each cancer cell duplicates. Thus, from the single initial cell, in a relatively short period a very large number of new cancer cells appear.

These cells occupy a portion of some tissue (skin, liver, etc.). The tumor mass starts to be detectable when there are around 10^9 cancer cells, a portion of tissue of diameter of few millimeters. The diameter of a single (normal or cancer) cell is of the order of 10 μn (1/100 millimeters). Usually, before, it is invisible and innocuous. But 10^{12} cancer cells is the usual order of magnitude leading to death, so in principle the distance separating the two scales is not so large.

Let us stop to "admire" a light micrograph (photomicrograph) of tumor cells in the epithelial region. These pictures (visit the indicated sites) give us a rather realistic idea of the complexity of the medium in which a tumor takes place.

From: http://www.pathologyatlas.ro/index.php

see http://www.pathologyatlas.ro/squamous-cell-carcinoma-skin.php



See also pictures at the sites:
http://histol.narod.ru/atlas-en/content-en.htm
http://histology-world.com/photoalbum

A tumor can be in different *phases*. The initial phase is called *in situ*. Cancer cells of such tumor have the proliferating phenotype but not worse. They duplicate and, due to pressure caused by the volume they have to occupy (each duplication doubles the occupied volume), the tumor mass spreads, increases. But cells keep close one to the other, they still are bounded by adhesion forces.

Later on, let us say at a random time, some cancer cell acquires the invasive phenotype and the tumor enter a new phase, the *invasive phase*: the cell may move, adhesion to other cancer cells is no more imposed, so the cell starts to move (very slowly!) in the surrounding space, called stroma, filled in with Extracellular Matrix (ECM), a compound of several objects and ingredients. The cell performs some sort of erratic motion, with a tendency to move in the direction of gradients of some quantity, that can be Oxygen or density of ECM. When an invasive cell reaches a blood vessel, it tries to enter it. Entering the blood stream requires special skills, phenotypes. If it succeeds, it moves in the flow and tries to attach somewhere, to colonize, namely to exit the stream and restart proliferating (other phenotypes). This is a very complex procedure, where only few cells over millions survive, but it is the metastatic phase, the most dangerous of course.

A tumor, both in the in situ or invasive phase, may start the so called angiogenic cascade, namely it enters the *angiogenic phase*. Primarily, this is due to the scarcity of Oxygen inside the tumor mass. Thus cells inside the tumor send messengers to blood vessels which sprout, send filaments of microvessels in the direction of the tumor. When the tumor is reached by these new vessels, oxygen supply is restored. This has the bad

consequence that the tumor becomes stronger, but even worse the consequence that a huge amount of cancer cells come in contact with blood vessels (not only the invasive cells which move). Moreover, the microvessels created by this procedure are less strong than the already existing vessels, and thus it is easier for the tumor cells to enter the blood stream. Thus angiogenesis is often associated with a higher risk of metastasis.

Tumors are attached either by means of surgery, or by chemotherapy, or both. In both cases there is a possibility that some cells have not been taken away or killed. In the case of surgery, the main problem is the amount of tissue, the boarder: it should be a little more extended than the mass observed by ultrasonography, or other means, but how much? Of course, more extend, more invasive is the surgery. Similarly, in the case of chemotherapy typically only a percentage of cells is killed, not all. Chemotherapy has the advantage to kill cells everywhere, hence also those which moved away from the main mass; but it has the disadvantage to kill also sane cells, typically those which proliferate for physiological reasons.

It is obvious that one should say much more to introduce problems of tumor biology and medicine, but let us stop and say more in subsequent sections.

1.2 Macroscopic and microscopic models

Cancer is a phenomenon at *cell level*, a few micron, which become visible and dangerous at the scale of a few millimeters or centimeters, the *tissue level*. These two levels, cell and tissue ones, will be called *microscopic* and *macroscopic* levels.

This is different from statistical physics, where microscopic refers to molecular level. Another possible misunderstanding is that macroscopic refers to the tissue level, or even just to a portion of tissue, not the full human body.

A macroscopic tumor, as we said above, has $10^8 - 10^{12}$ cells. It is reasonable to apply ideas similar to statistical physics, although gap of scales is much smaller.

At the microscopic level, stochasticity is essential, even more than in the common descriptions of molecular dynamics in statistical physics. At the macroscopic level, it is reasonable to use deterministic models.

To exemplify, we have in mind to use stochastic particle systems to describe cell motion and interaction; deterministic partial differential equations to describe a tumor from the macroscopic viewpoint.

The concept of *macroscopic limit* is essential for the sequel: it is the link between the two levels. When we can, we try to understand which are the PDEs arising from a microscopic model of interacting particles, in the limit when the number of particles tends to infinity.

1.2.1 Mechanical perspective

We adopt a mechanical perspective: we are interested in the phenomena that can be described and explained by mechanical ideas: motion of cells (considered like particles), change in time of the density of cells. In the biomedical study of tumors there are a lot of other aspects, like those related to genetic mutations, protein networks which we do not touch.

1.2.2 Special pedagogical path

If the mathematical theory of tumors would be more mature, it could be convenient to proceed step by step, from the most elementary ideas and models to the more difficult ones.

It is not so, the literature is rich but still disorganized. Thus I prefer to start in a more provocative way: I will present a very advanced PDE, macroscopic, model of invasive tumor with angiogenesis.

Motivated by it, we shall make steps in a backward direction, to understand its pieces, identifying also some drawbacks of the big model, asking questions which lead us in other directions. At the end I hope that some coherent picture arises, but not "given from above", as in well structured mathematical theories.

Thus:

- the remaining part of Chapter 1 is devoted to formulate the big complex model, its biological motivation, and also start thinking to a corresponding microscopic stochastic model of cells.
- Chapter 2 extracts a small part of the big model, a sub-model, called Fisher-KPP (it existed long time before the big one) and start analyzing it in some detail, rigorous and numerical. Again we see some element of microscopic description, but just a few ones.
- Chapter 3 goes back to the big model and identifies a number of difficulties, critical points, which to a large extent I do not know how to solve. Numerics of the big model will be shown. Chapters 1-2-3 have been devoted mainly to macroscopic aspects, with few initial remarks on particle systems.
- Chapters 4-5-6 are more intensively devoted to microscopic stochastic models, and their macroscopic limits. Thus we restart from cells, learn more and describe them from scratch, directly (not as a by-product of the PDE description). Chapter 4 is also devoted to the in situ phase and to the mathematics of proliferation.
- Chapter 5 describes, in rather informal way, different regimes of interaction (long range, short range, intermediate one) and different expected PDEs in the macroscopic limit.

- Chapter 6 is devoted to some mathematical details about these interacting particle systems and their large scale limit. We present in particular some elements of the mean field theory and of the intermediate regime case. At the end we reach a picture of micro and macroscopic models, with different level of complexity.

1.3 Role of Mathematics

Nowadays the mathematical models are still too rough to give really useful informations for cancer medicine. The question is what we may hope to get. It is better to discuss this topic at the end of the lectures but at least one sentence is useful now.

Assume that, thanks to a few observations (like three radiographies, once a week, for three weeks) we may compare different models and decide whether it is more plausible that a tumor is in the in situ phase instead of the beginning of the invasive phase or the angiogenic phase. To understand this by visual inspection (by ultrasonography etc.) is not easy. The mathematical model captures the differences in evolution and may reveal which model fits better the recorded evolution.

Then, the doctor has more informations to decide about surgery/chemotherapy and their amount.

For this reason, we shall pay particular attention, in these lectures, to the mathematical differences in the description of different phases.

Let us also mention that another, more mainstream, declared purpose is the possibility to run simulations under different therapies, schedules of chemotherapy, combination of surgery and chemotherapy, looking for optimal choices.

1.4 An advanced model of invasive tumor with angiogenesis

Let us describe a model introduced by [22]. This model is made of 7 coupled PDE-ODE. Let us immediately emphasize that this is not "the model" of cancer growth. It is one model, with its own degree of sophistication, and describing a tumor in a particular phase. It is a model of mechanical type, dealing with aspects like random motion, motion along gradients, proliferation, change of type.

1.4.1 Normoxic, hypoxic and apoptotic cells

Cancer cells are mainly characterized by their tendency to duplicate. However, when they receive an insufficient amount of oxygen, or when they do not have space enough, their proliferation is inhibited.

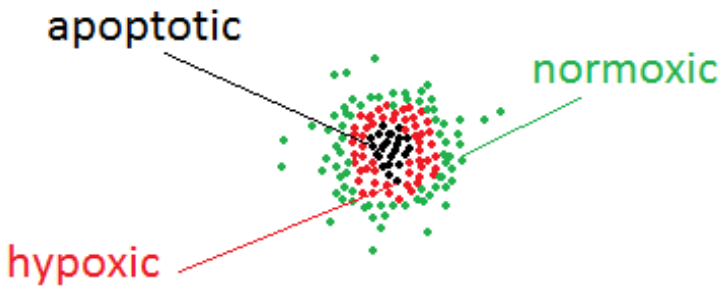
The model of [22] splits the category of cancer cells in three classes:

1. *normoxic* cells: healthy, proliferating tumor cells, with *normal oxygen* supply

2. *hypoxic* cells: quiescent tumor cells, with *poor oxygen* supply

3. *apoptotic* cells: death or *programmed to death* tumor cells

The figure below shows schematically the obvious fact that in a three-dimensional tumor mass the hypoxic cells are those inside, with an even smaller core of apoptotic cells.



PDE for normoxic cells

The model prescribes the following PDE for normoxic cells:

$$\begin{aligned}\frac{\partial \mathcal{N}}{\partial t} = & \underbrace{k_1 \Delta \mathcal{N}}_{\text{background diffusion}} \qquad (\mathcal{N}(t,x) = \text{normoxic cell density}) \\ & \underbrace{\operatorname{div}(\sigma(\mathcal{N})\nabla\mathcal{N})}_{\text{crowding-driven diffusion}} + \underbrace{c_1 \mathcal{N}(\mathcal{V}_{\max} - \mathcal{V})}_{\text{proliferation}} \\ & - \underbrace{\chi_1 \operatorname{div}(\mathcal{N}\nabla m)}_{\text{transport along ECM gradient}} \\ & - \underbrace{\alpha_{\mathcal{N}\to\mathcal{H}} 1_{o\le o_{\mathcal{H}}} \mathcal{N}}_{\text{normoxic}\to\text{hypoxic}} + \underbrace{\alpha_{\mathcal{H}\to\mathcal{N}} 1_{o>o_{\mathcal{H}}} \mathcal{H}}_{\text{hypoxic}\to\text{normoxic}}\end{aligned}$$

- First, a background diffusion is admitted; and no adhesion constraint is imposed; this is the *invasive phase*. The value of the constant k_1 is, however, extremely small, the diffusion is extremely slow.

- Second, a crowding-driven diffusion is introduced. This is a very interesting term, which will occupy our effort quite often, but not now, it is too early. Just as a first idea, this term enforces the diffusion when the density of cells is larger.

- Third, proliferation. If we neglect space, the time evolution of a complete proliferation is given by the differential equation $x'(t) = \lambda x(t)$, where λ is the proliferation rate.

One can consider a model with time-dependent rate. In the PDE above the time-dependent rate is $c_1(\mathcal{V}_{\max} - \mathcal{V})$: it decreases to zero when the total density of cell (plus ECM)

$$\mathcal{V} = \mathcal{N} + \mathcal{H} + \mathcal{A} + \mathcal{E} + m = \text{total density of cells plus ECM}$$

approaches the threshold $\mathcal{V}_{\max}$.

- Then, transport: normoxic cells, beyond the random motion described by $k_1 \Delta \mathcal{N}$, have a tendency to move along the gradient of m, the ECM density.
- Finally, some normoxic cells become hypoxic (when $o \leq o_{\mathcal{H}}$, namely the oxygen is too low) and some hypoxic cells are restored to the normoxic state (when $o > o_{\mathcal{H}}$).

ODE for hypoxic cells

Hypoxic cells do not move and do not proliferate. Their number increases when some normoxic cell deteriorate to the hypoxic state (for $o \leq o_{\mathcal{H}}$). And decreases either when they are restored to the normoxic state (for $o > o_{\mathcal{H}}$) or when they degenerate to apoptotic cells (for $o \leq o_{\mathcal{A}}$):

$$\begin{aligned} \frac{d\mathcal{H}}{dt} = & \underbrace{\alpha_{\mathcal{N}\to\mathcal{H}} 1_{o \leq o_{\mathcal{H}}} \mathcal{N}}_{\text{normoxic} \to \text{hypoxic}} - \underbrace{\alpha_{\mathcal{H}\to\mathcal{N}} 1_{o \geq o_{\mathcal{H}}} \mathcal{H}}_{\text{hypoxic} \to \text{normoxic}} \\ & - \underbrace{\alpha_{\mathcal{H}\to\mathcal{A}} 1_{o \leq o_{\mathcal{A}}} \mathcal{H}}_{\text{hypoxic} \to \text{apoptotic}} \qquad (\mathcal{H}(t,x) = \text{hypoxic cell density}) \end{aligned}$$

ODE for apoptotic cells

These cells are programmed to death. It means that they dissolve in a regulated way, not by necrosis and causing infections. The number of apoptotic cells can only increase, due to the hypoxic cells that deteriorate:

$$\frac{d\mathcal{A}}{dt} = \underbrace{\alpha_{\mathcal{H}\to\mathcal{A}} 1_{o \leq o_{\mathcal{A}}} \mathcal{H}}_{\text{hypoxic} \to \text{apoptotic}} \qquad (\mathcal{A}(t,x) = \text{apoptotic cell density})$$

ODE for ECM

In this model, it is assumed that the Extracellular Matrix can only deteriorate, due to the invasion of normoxic cells:

$$\frac{dm}{dt} = - \underbrace{\beta m \mathcal{N}}_{\text{degradation by normoxic cells}} \qquad (m(t,x) = \text{ExtraCellular Matrix})$$

Crowding-driven diffusion

In the equation for normoxic cells we have the term

$$\operatorname{div}\left(\sigma\left(\mathcal{N}\right)\nabla\mathcal{N}\right)$$

which we have called crowding-driven diffusion. The prescription for $\sigma\left(\mathcal{N}\right)$ in [22] is (up to constants)

$$\sigma\left(\mathcal{N}\right)=\max\left(\mathcal{N}-\mathcal{N}_0,0\right).$$

In other words, when the density of normoxic cells passes the threshold $\mathcal{N}_0$, an additional diffusion starts. This term is very intriguing and typical of [22]. It is similar to (up to constants)

$$\sigma\left(\mathcal{N}\right)=\mathcal{N}$$

which corresponds to

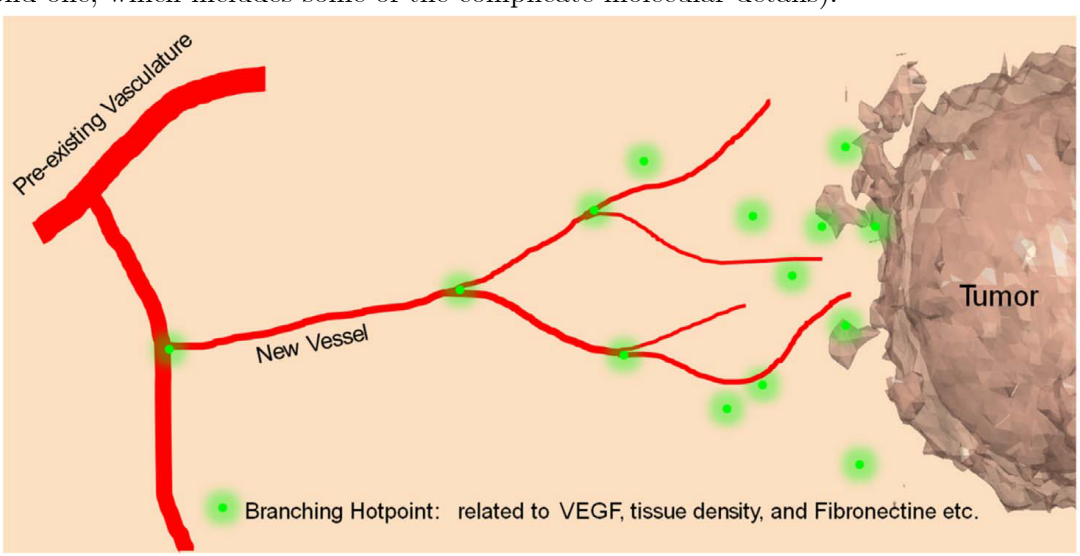
$$\frac{1}{2}\Delta\mathcal{N}^2$$

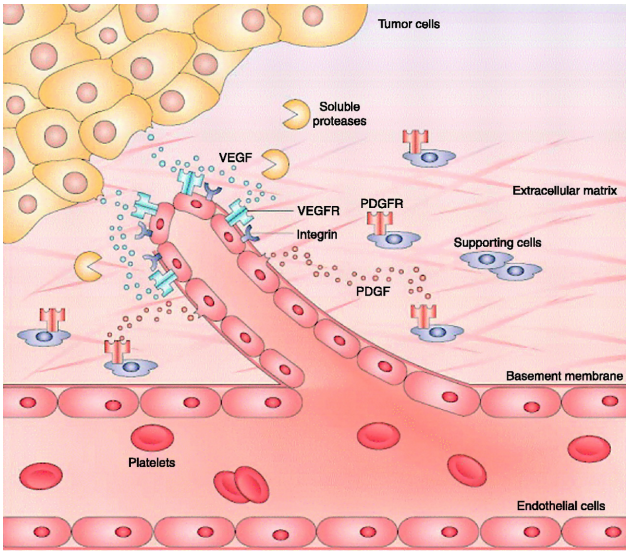
also called porous media diffusion.

1.4.2 The endothelial cascade

Hypoxic cells need more oxygen to survive. Thus they initiate a cascade of cellular interactions. The result is angiogenesis: new vascularization is developed to supply the tumor (microvessels branching from main vessels in the direction of the tumor).

A messenger from hypoxic cells is sent to endothelial cells: it is called VEGF (Vascular Endothelial Growth Factor). Here are two pictures taken from the web (see [30] for the second one, which includes some of the complicate molecular details):





See also the movies at the web sites [28], [29].

PDEs for the VEGF concentration

VEGF is a density of object at molecular level. These proteins have always a diffusion (opposite to cells, which diffuse only under special circumstances):

$$\frac{\partial g}{\partial t} = \underbrace{k_4 \Delta g}_{\text{diffusion}} \qquad (g(t,x) = \text{VEGF concentration})$$
$$+ \underbrace{\alpha_{\mathcal{H}\to g}\mathcal{H}}_{\text{production by hypoxic cells}} - \underbrace{\alpha_{g\to\mathcal{E}}\mathcal{E}g}_{\text{uptake by endothelial cells}}$$

and the constant k_4 is much bigger than k_1.

Moreover, the concentration of VEGF is produced by the hypoxic cells, hence increases due to their presence, and VEGF is absorbed by endothelial cells (those forming the boundary of blood vessels).

PDEs for endothelial ramification

As the pictures above show, endothelial cells do not diffuse as isolated individuals but they propagate as microvessels. However, keeping track of their topological structure is difficult and perhaps not so important (in the opinion of the authors of this model), hence the concept of density of endothelial cells, or maybe more precisely density of endothelial

ramification, is introduced. The PDE is:

$$\frac{\partial \mathcal{E}}{\partial t} = \underbrace{k_2 \Delta \mathcal{E}}_{\text{diffusion}} \qquad (\mathcal{E}(t,x) = \text{density of endothelial ramification})$$
$$- \underbrace{\chi_2 \operatorname{div}(\mathcal{E} \nabla g)}_{\text{transport along VEGF gradient}} + \underbrace{c_2 \mathcal{E} g (\mathcal{V}_{\max} - \mathcal{V})}_{\text{proliferation under VEGF presence}}$$

where

- random motion is considered (however, as for normoxic cells, the value of the constant k_1 is extremely small)
- endothelial cells move along the VEGF gradient, in order to reach the area occupied by hypoxic cells
- in order to build new vessels, they need to proliferate; the rate of proliferation is $c_2 g (\mathcal{V}_{\max} - \mathcal{V})$, namely it is proportional to VEGF concentration, and is inhibited by the same volume constraint of normoxic proliferation.

PDEs for oxygen concentration

Oxygen is also molecular-level hence it diffuses, with k_3 is much bigger than k_1 (also bigger than k_4)

$$\frac{\partial o}{\partial t} = \underbrace{k_3 \Delta o}_{\text{diffusion}} \qquad (o(t,x) = \text{oxygen concentration})$$
$$+ \underbrace{c_3 \mathcal{E} (o_{\max} - o)}_{\text{production by endothelial cells}} - \underbrace{\alpha_{o \to \mathcal{N},\mathcal{H},\mathcal{E}} (\mathcal{N} + \mathcal{H} + \mathcal{E}) o}_{\text{uptake by all living cells}} - \underbrace{\gamma o}_{\text{oxygen decay}}$$

and

- oxygen concentration increases proportionally to the density of endothelial ramification, but only up to $o_{\max}$
- it decreases due to absorption by various cells (not the apoptotic ones)
- it decays (slowly).

Summary of variables

It may be useful to summarize the list of variables:

$\mathcal{N}(t,x)$ = density of normoxic cells
$\mathcal{H}(t,x)$ = density of hypoxic cells
$\mathcal{A}(t,x)$ = density of apoptotic cells
$\mathcal{E}(t,x)$ = density of endothelial cells (or density of vasculature)
$o(t,x)$ = oxygen concentration
$g(t,x)$ = angiogenic growth factor (VEGF) concentration
$m(t,x)$ = ECM (ExtraCellular Matrix)

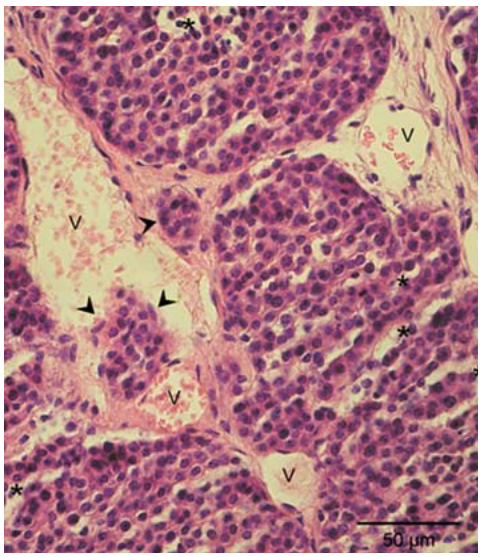
Summary of constants. Difficulties

The next list of constants is given mostly to emphasize a main difficulty with this kind of complex models: some of these constants are poorly known and parameter fit is excluded by the impossibility of real experiment.

k_1 = background random motility coefficient of normoxic cells
k_2 = random motility coefficient of endothelial cells
k_3 = diffusion coefficient of oxygen
k_4 = diffusion coefficient of angiogenic factor
χ_1 = transport coefficient of normoxic cells along ECM gradient
χ_2 = transport coefficient of endothelial cells along VEGF gradient
$\mathcal{V}_{cr}$ = threshold for crowding-driven diffusion
$\mathcal{V}_{\max}$ = limit to total volume of cells and ECM
c_1 = proliferation rate of normoxic cells
c_2 = proliferation rate of endothelial cells
c_3 = production rate of oxygen
$\alpha_{\mathcal{N}\to\mathcal{H}}$ = decay rate from normoxic to hypoxic cells
$\alpha_{\mathcal{H}\to\mathcal{N}}$ = restoration rate from hypoxic to normoxic cells
$\alpha_{\mathcal{H}\to\mathcal{A}}$ = decay rate from hypoxic to apoptotic cells
$\alpha_{\mathcal{H}\to g}$ = production rate of VEGF from hypoxic cells
$\alpha_{o\to\mathcal{N},\mathcal{H},\mathcal{E}}$ = uptake rate of oxygen from all living cells
$\alpha_{g\to\mathcal{E}}$ = uptake rate of VEGF from endothelial cells
$o_{\max}$ = maximum oxygen concentration
$o_{\mathcal{H}}$ = oxygen threshold for transition normoxic $\leftrightarrow$ hypoxic
$o_{\mathcal{A}}$ = oxygen threshold for transition hypoxic $\leftrightarrow$ apoptotic
β = rate of ECM degradation
γ = oxygen decay rate

To stress the difficulties, let us also mention the fact that a tissue is a highly complex environment, possibly highly heterogeneous. The first picture of Chapter one, like

for instance the next one, are examples of the geometrical complexity of a tissue. This complexity is not considered by the model above, at least is this idealized form.

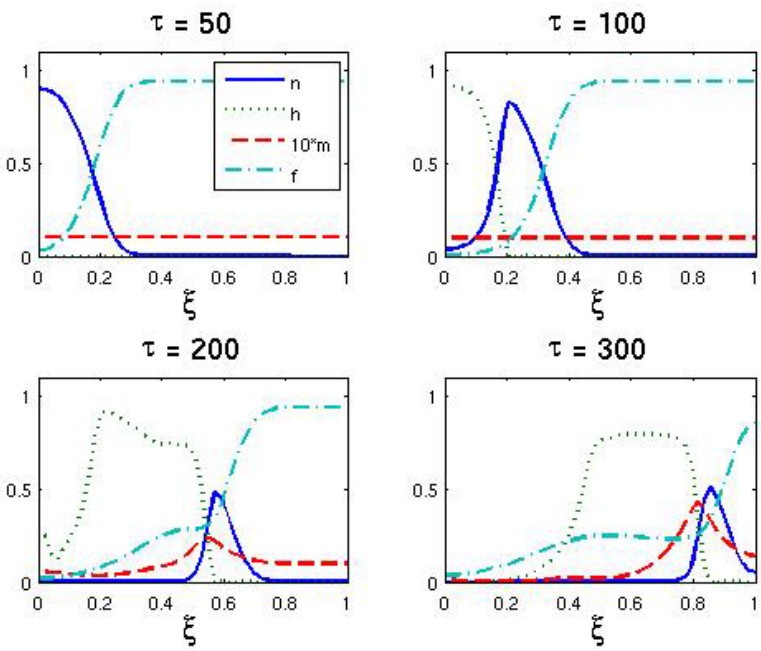


Simulations

In Chapter 3 we shall devote some time to discuss numerical simulation of the full system and a number of its reductions. However, just to give a first impression, let us see some of the pictures reported by the paper [22].

Blue: density of normoxic cells; light blue (green): extracellular matrix

Dotted black: density of hypoxic cells; red: density of endothelial ramification.



We see a front of normoxic cells propagating in the medium (they degradate ECM), followed by a front of hypoxic cells (it also apparently disappears due to the decay to apoptotic cells, which are not shown). The endothelial ramification is silent for a while, then it is triggered.

1.5 Stochastic dynamics associated to some element of the previous model

Some elements of the previous model have a straightforward translation into a microscopic, cell level, random dynamics.

1.5.1 SDE and Fokker-Planck equation

Let us recall, for this purpose, the relation between Fokker-Planck equations and SDEs. Let $(W_t)_{t\geq 0}$ be a Brownian motion in $\mathbb{R}^k$, defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $\left(\mathcal{F}_t^B\right)_{t\geq 0}$ be the associated completed filtration. Let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times k}$ be continuous functions such that, for some constants $L_b, L_\sigma > 0$,

$$|b(t,x) - b(t,y)| \leq L_b |x-y|$$
$$|\sigma(t,x) - \sigma(t,y)| \leq L_\sigma |x-y|$$

for all $x, y \in \mathbb{R}^d$ and all $t \in [0,T]$. Let X_0 be an $\mathcal{F}_0$-measurable r.v. with values in $\mathbb{R}^d$. Consider the SDE in $\mathbb{R}^d$

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t \tag{1.1}$$

and say that a stochastic process $(X_t)_{t\in[0,T]}$ is a strong solution if it is continuous, adapted to $\left(\mathcal{F}_t^B\right)_{t\in[0,T]}$ and a.s. satisfies

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s$$

for all $t \in [0,T]$. Two solutions coincide if they are indistinguishable processes. Let us recall the following classical result:

Theorem 1 *There exists a unique strong solution.*

Consider now the following parabolic PDE on $[0,T] \times \mathbb{R}^d$

$$\frac{\partial p}{\partial t} = \frac{1}{2} \sum \partial_i \partial_j (a_{ij} p) - \operatorname{div}(pb), \qquad p|_{t=0} = p_0 \tag{1.2}$$

called Fokker-Planck equation. Here

$$a = \sigma\sigma^T.$$

Although in many cases it has regular solutions, in order to minimize the theory it is convenient to introduce the concept of measure-valued solution. We restrict to the case of probability measures. We loosely write

$$\frac{\partial \mu_t}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{d} \partial_i \partial_j (a_{ij} \mu_t) - \operatorname{div}(\mu_t b), \qquad \mu|_{t=0} = \mu_0 \tag{1.3}$$

but we mean the following concept. By $\langle \mu_t, \phi \rangle$ we mean $\int_{\mathbb{R}^d} \phi(x)\, \mu_t(dx)$. By $C_c^\infty\left(\mathbb{R}^d\right)$ we denote the space of smooth compact support functions $\phi : \mathbb{R}^d \to \mathbb{R}$, and by $C_c^0\left([0,T] \times \mathbb{R}^d\right)$ the space of continuous compact support functions $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$,

Definition 2 *A measure-valued solution of the Fokker-Planck equation (1.3) is a family of Borel probability measures $(\mu_t)_{t\in[0,T]}$ on $\mathbb{R}^d$ such that $t \mapsto \langle \mu_t, \varphi(t,.) \rangle$ is measurable for all $\varphi \in C_c^0\left([0,T] \times \mathbb{R}^d\right)$ and*

$$\langle \mu_t, \phi \rangle - \langle \mu_0, \phi \rangle = \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \langle \mu_s, a_{ij}(s,.)\, \partial_i \partial_j \phi \rangle\, ds + \int_0^t \langle \mu_s, b(s,.) \cdot \nabla \phi \rangle\, ds$$

for every $\phi \in C_c^\infty\left(\mathbb{R}^d\right)$.

Denote now by μ_t the law of X_t, the solution of equation (1.1). To simplify the proof we need a little extra assumption. The simplest is to assume that σ is bounded (otherwise see Remark 4 below).

Theorem 3 *The law μ_t of X_t is a measure-valued solution of the the Fokker-Planck equation (1.3).*

Proof. By Itô formula for $\phi(X_t)$:

$$d\phi(X_t) = \nabla\phi(X_t)\cdot dX_t + \frac{1}{2}\sum_{i,j=1}^{d}\partial_i\partial_j\phi(X_t)\,a_{ij}(t,X_t)\,dt$$

$$= \nabla\phi(X_t)\cdot b(t,X_t)\,dt + \nabla\phi(X_t)\cdot\sigma(t,X_t)\,dW_t + \frac{1}{2}\sum_{i,j=1}^{d}\partial_i\partial_j\phi(X_t)\,a_{ij}(t,X_t)\,dt.$$

We have $E\int_0^T|\nabla\phi(X_t)\cdot\sigma(t,X_t)|^2\,dt<\infty$ (we use here that σ is bounded; and also $\nabla\phi$ is bounded), hence $E\int_0^t\nabla\phi(X_s)\cdot\sigma(s,X_s)\,dW_s=0$ and thus

$$E[\phi(X_t)]-E[\phi(X_0)] = E\int_0^t\nabla\phi(X_s)\cdot b(t,X_s)\,ds + \frac{1}{2}\sum_{i,j=1}^{d}E\int_0^t\partial_i\partial_j\phi(X_s)\,a_{ij}(s,X_s)\,ds.$$

Since $E[\phi(X_t)]=\int_{\mathbb{R}^d}\phi(x)\,\mu_t(dx)$ (and similarly for the other terms) we get the weak formulation of equation (1.3). The preliminary property that $t\mapsto\langle\mu_t,\varphi(t,.)\rangle=E[\varphi(t,X_t)]$ is measurable for all $\varphi\in C_c^0\left([0,T]\times\mathbb{R}^d\right)$ is easy. ■

Remark 4 *Instead of assuming σ bounded, it is sufficient to assume that $E\left[|X_0|^2\right]<\infty$. This implies $E\left[\sup_{t\in[0,T]}|X_t|^2\right]<\infty$, which, along with the linear growth of σ, will give us the estimate $E\int_0^T|\nabla\phi(X_t)\cdot\sigma(t,X_t)|^2\,dt<\infty$ needed in the proof above.*

Remark 5 *Under suitable assumptions, like the simple case when a_{ij} is the identity matrix, if μ_0 has a density p_0 then also μ_t has a density $p(t,\cdot)$, often with some regularity gained by the parabolic structures, and thus the Fokker-Planck equation in the differential form (1.2) holds. We do not insist rigorously in this direction since our purpose here is only to understand the modeling aspects, more precisely the shape of the microscopic model behind the PDEs of the previous sections.*

1.5.2 Non-interacting particles and Fokker-Planck as macroscopic limit

We may reformulate Theorem 3 as a macroscopic limit of a system of non-interacting particles.

Let W_t^n, $n \in \mathbb{N}$, be a sequence of independent Brownian motions in $\mathbb{R}^k$, defined on a probability space $(\Omega, \mathcal{F}, P)$. Let b, σ as above. Consider the sequence of SDEs in $\mathbb{R}^d$

$$dX_t^n = b\left(t, X_t^n\right) dt + \sigma\left(t, X_t^n\right) dW_t^n$$

with X_0^n given independent $\mathbb{R}^d$-r.v.'s, $\mathcal{F}_0$-measurable, with the same law μ_0. Then, since weak uniqueness holds for the SDE, the processes X_t^n have the same law; in particular, with the notations above, the marginal at time t of X_t^n is μ_t. Moreover, the processes X_t^n are independent, since each X_t^n is adapted to the corresponding Brownian motion W_t^n, which are independent.

Consider, for each $N \in \mathbb{N}$, the random probability measure, called *empirical measure*,

$$S_t^N := \frac{1}{N} \sum_{n=1}^{N} \delta_{X_t^n}$$

namely, for $\phi \in C_c^0\left(\mathbb{R}^d\right)$,

$$\left\langle S_t^N, \phi \right\rangle = \frac{1}{N} \sum_{n=1}^{N} \phi\left(X_t^n\right).$$

It is a sort of discrete density of particles. The following simple theorem is our first example of macroscopic limit of a system of microscopic particles.

Theorem 6 *For every $t \in [0, T]$ and $\phi \in C_c^0\left(\mathbb{R}^d\right)$, a.s.*

$$\lim_{N \to \infty} \left\langle S_t^N, \phi \right\rangle = \langle \mu_t, \phi \rangle.$$

In other words, S_t^N converges weakly, a.s., to a measure-valued solution of the Fokker-Planck equation (1.3).

Proof. Since, at each time t, the r.v. $\phi\left(X_t^n\right)$ are bounded i.i.d., by the strong law of large numbers we have, a.s.,

$$\lim_{N \to \infty} \left\langle S_t^N, \phi \right\rangle = E\left[\phi\left(X_t^1\right)\right] = \langle \mu_t, \phi \rangle.$$

And by Theorem 3 above, μ_t is a measure-valued solution of the the Fokker-Planck equation (1.3). The only technical detail that we could discuss more deeply is the precise meaning of the sentence "converges weakly, a.s.". The simplest meaning is the a.s. convergence $\lim_{N \to \infty} \left\langle S_t^N, \phi \right\rangle = \langle \mu_t, \phi \rangle$, for every a priori given $\phi \in C_c^0\left(\mathbb{R}^d\right)$. But this implies the stronger concept that, chosen a priori $\omega \in \Omega$ a.s., we have $\lim_{N \to \infty} \left\langle S_t^N, \phi \right\rangle = \langle \mu_t, \phi \rangle$ for every $\phi \in C_c^0\left(\mathbb{R}^d\right)$ (namely the null set of ω's where the convergence could fail is independent of ϕ). To reach this result it is sufficient to notice first that it holds for any given countable set $\{\phi_n\}$. Then, it is possible to chose such set $\{\phi_n\} \subset C_c^0\left(\mathbb{R}^d\right)$ in a way

that it is dense in $C_c^0\left(\mathbb{R}^d\right)$, in the topology of uniform convergence on the full space $\mathbb{R}^d$. Therefore, taken $\phi \in C_c^0\left(\mathbb{R}^d\right)$, taken a subsequence $\left\{\phi_{n_k}\right\}$ which converges uniformly to ϕ, one has

$$\begin{aligned}\left|\left\langle S_t^N, \phi\right\rangle - \langle \mu_t, \phi\rangle\right| &\leq \left|\left\langle S_t^N, \phi\right\rangle - \left\langle S_t^N, \phi_{n_k}\right\rangle\right| + \left|\left\langle S_t^N, \phi_{n_k}\right\rangle - \left\langle \mu_t, \phi_{n_k}\right\rangle\right| + \left|\left\langle \mu_t, \phi_{n_k}\right\rangle - \langle \mu_t, \phi\rangle\right| \\ &\leq 2\left\|\phi_{n_k} - \phi\right\|_0 + \left|\left\langle S_t^N, \phi_{n_k}\right\rangle - \left\langle \mu_t, \phi_{n_k}\right\rangle\right|.\end{aligned}$$

Choose $\omega \in \Omega$ such that $\lim_{N\to\infty}\left\langle S_t^N(\omega), \phi_{n_k}\right\rangle = \left\langle \mu_t, \phi_{n_k}\right\rangle$ for all $k \in \mathbb{N}$. Given $\varepsilon > 0$, we first take k such that $\left\|\phi_{n_k} - \phi\right\|_0 < \frac{\varepsilon}{4}$; then we take N_0 such that $\left|\left\langle S_t^N(\omega), \phi_{n_k}\right\rangle - \left\langle \mu_t, \phi_{n_k}\right\rangle\right| < \frac{\varepsilon}{2}$ for all $N > N_0$. We get $\left|\left\langle S_t^N(\omega), \phi\right\rangle - \langle \mu_t, \phi\rangle\right| < \varepsilon$ for all $N > N_0$. Hence $\lim_{N\to\infty}\left\langle S_t^N, \phi\right\rangle = \langle \mu_t, \phi\rangle$ on the same set of ω's where the convergence was true for all ϕ_{n_k}, $k \in \mathbb{N}$. The proof is complete. ■

1.5.3 Fragments of microscopic model associated to the PDE system

The full system of 7 equations above cannot be split into exact sub-systems. But we can consider simplified versions just to start understanding the translation from Macroscopic to Microscopic. Consider for instance the closed system between normoxic cell density and ECM, neglecting all other terms and equations:

$$\frac{\partial \mathcal{N}}{\partial t} = \underbrace{k_1 \Delta \mathcal{N}}_{\text{background diffusion}} - \underbrace{\chi_1 \operatorname{div}(\mathcal{N}\nabla m)}_{\text{transport along ECM gradient}}$$

$$\frac{dm}{dt} = -\underbrace{\beta m \mathcal{N}}_{\text{degradation by normoxic cells}}$$

If m would be a priori given, we could say that $\mathcal{N}$ is the weak limit of

$$S_n^{\mathcal{N}}(t) := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i^{\mathcal{N}}(t)}$$

where $X_i^{\mathcal{N}}(t)$, $i = 1, 2, ...$ are the positions of independent normoxic cells subject to the microscopic dynamics

$$dX_i^{\mathcal{N}}(t) = \chi_1 \nabla m\left(t, X_i^{\mathcal{N}}(t)\right) dt + \sqrt{2k_1}\, dW_i^{\mathcal{N}}(t).$$

But $m(t,x)$ is not given. To simplify, we could decide that the dynamics acts in two steps, like in the numerical methods called slitting up or decomposition: in one step, m is considered as a given function, and the scheme above applies; in the next step, $\mathcal{N}(t,x)$ is considered as a given function, and m is computed from equation $\frac{dm}{dt} = -\beta m \mathcal{N}$; to be realistic, the two steps should be short.

Otherwise, if we do not want to use a two-step mechanism, we have to couple the two equations. In this case we have to solve the system

$$\begin{aligned}\frac{dm}{dt} &= -\beta m\mathcal{N}\\ \mathcal{N}(t,\cdot) &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i^{\mathcal{N}}(t)}\\ dX_i^{\mathcal{N}}(t) &= \chi_1\nabla m\left(t,X_i^{\mathcal{N}}(t)\right)dt+\sqrt{2k_1}dW_i^{\mathcal{N}}(t).\end{aligned}$$

Even more realistically, we should not pass to the limit inside the system, but only a posteriori. We mane that we should define some relatively smooth function $\mathcal{N}_n(t,x)$ based on $\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i^{\mathcal{N}}(t)}$ and consider the system

$$\begin{aligned}\frac{dm_n}{dt} &= -\beta m_n\mathcal{N}_n\\ dX_i^{\mathcal{N}}(t) &= \chi_1\nabla m_n\left(t,X_i^{\mathcal{N}}(t)\right)dt+\sqrt{2k_1}dW_i^{\mathcal{N}}(t).\end{aligned}$$

In absence of better ideas, to define $\mathcal{N}_n(t,x)$ one could take the convolution of the measure $\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i^{\mathcal{N}}(t)}$ with a smooth mollifier $\theta_n(x)=\epsilon_n^{-d}\theta\left(\epsilon_n^{-1}x\right)$ (under usual assumptions on θ), namely

$$\mathcal{N}_n(t,x)=\frac{1}{n}\sum_{i=1}^{n}\theta_n\left(x-X_i^{\mathcal{N}}(t)\right).$$

Now the limit as $n\to\infty$ of the solution of the system

$$\begin{aligned}\frac{dm_n(t,x)}{dt} &= -\frac{\beta}{n}\sum_{i=1}^{n}\theta_n\left(x-X_i^{\mathcal{N}}(t)\right)m_n(t,x)\\ dX_i^{\mathcal{N}}(t) &= \chi_1\nabla m_n\left(t,X_i^{\mathcal{N}}(t)\right)dt+\sqrt{2k_1}dW_i^{\mathcal{N}}(t)\end{aligned}$$

has to be investigated ex novo, it is not a simple consequence of the facts seen above on Fokker-Planck equations.

Similar arguments can be applied to the equation for endothelial density,

$$\frac{\partial\mathcal{E}}{\partial t}=\underbrace{k_2\Delta\mathcal{E}}_{\text{diffusion}}-\underbrace{\chi_2\operatorname{div}(\mathcal{E}\nabla g)}_{\text{transport along VEGF gradient}}$$

by introducing the positions of single endothelial cells $X_i^{\mathcal{E}}(t)$, $i=1,2,...$, subject to the equations

$$dX_i^{\mathcal{E}}(t)=\chi_2\nabla g\left(t,X_i^{\mathcal{E}}(t)\right)dt+\sqrt{2k_2}dW_i^{\mathcal{E}}(t).$$

Clearly, to reach a full microscopic description we should be able to describe other difficult terms, like nonlinear diffusions

$$\underbrace{\operatorname{div}\left(\sigma\left(\mathcal{N}\right)\nabla\mathcal{N}\right)}_{\text{crowding-driven diffusion}}$$

change of type

$$\underbrace{\alpha_{\mathcal{H}\to\mathcal{N}}1_{o\geq o_{\mathcal{H}}}\mathcal{H}}_{\text{hypoxic}\,\to\,\text{normoxic}}$$

proliferation

$$\underbrace{c_1\mathcal{N}\left(\mathcal{V}_{\max}-\mathcal{V}\right)}_{\text{proliferation}}$$

and so on. The message of this section is that linear diffusion terms and transport ones have a simple microscopic counterpart, suggested by the theory of Fokker-Planck equations.

Chapter 2

Fisher-Kolmogorov-Petrovskii-Piskunov model

2.1 Introduction

The complexity of the model of Chapter 1, invasive with angiogenesis, does not allow a straightforward mathematical analysis and thus it is convenient to start by understanding an easier model, considered in the earlier literature on oncology and other applications in biology.

It deals with the space-time evolution of a single quantity, the density of tumor cells. It is assumed subject to diffusion and proliferation:

$$\frac{\partial u}{\partial t} = D\Delta u + \rho u \left(u_{\max} - u\right), \qquad u|_{t=0} = u_0. \tag{2.1}$$

The diffusion term is the usual one, based on Fick law, with constant diffusion coefficient D. Proliferation happens at each point x. Proliferation rate is $\rho\left(u_{\max} - u\right)$, ρ a positive constant, hence damped by the factor $(u_{\max} - u)$ which reduces proliferation to zero when u reaches the threshold $u_{\max}$. Where the density is higher, proliferation is inhibited.

The dynamical mechanism can be intuitively described as follows: if only the term $\rho u\left(u_{\max} - u\right)$ was present, the density u would increase up to $u_{\max}$ at each point x; diffusion redistributes the density from higher to lower zones. Progressively, the density will occupy new and new space (due to diffusion) and at the same time will increase towards $u_{\max}$.

2.1.1 Scaling transformations

Setting $U = u/u_{\max}$, dividing (2.1) by $u_{\max}$ we get

$$\frac{\partial U}{\partial t} = D\Delta U + \widetilde{\rho} U (1 - U)$$

with $\widetilde{\rho} = \rho u_{\max}$, hence we may replace $u_{\max}$ by 1 with a simple transformation (modifying ρ). Setting $V(t,x) = U(\lambda t, \mu x)$, with $\lambda, \mu > 0$, we have

$$\begin{aligned}\frac{\partial V}{\partial t}(t,x) &= \lambda \frac{\partial U}{\partial t}(\lambda t, \mu x) = \lambda D (\Delta U)(\lambda t, \mu x) + \lambda \widetilde{\rho} U(\lambda t, \mu x)(1 - U(\lambda t, \mu x)) \\ &= \widehat{D} \Delta V(t,x) + \widehat{\rho} V(t,x)(1 - V(t,x))\end{aligned}$$

with $\widehat{D} = \frac{\lambda D}{\mu^2}$, $\widehat{\rho} = \lambda \widetilde{\rho}$. Thus, by a simple transformation, we may arbitrarily change both diffusion and proliferation constants. For these reasons, it is sufficient to restrict to the *canonical* model

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + u(1 - u). \tag{2.2}$$

Moreover, the function $v = 1 - u$ satisfies

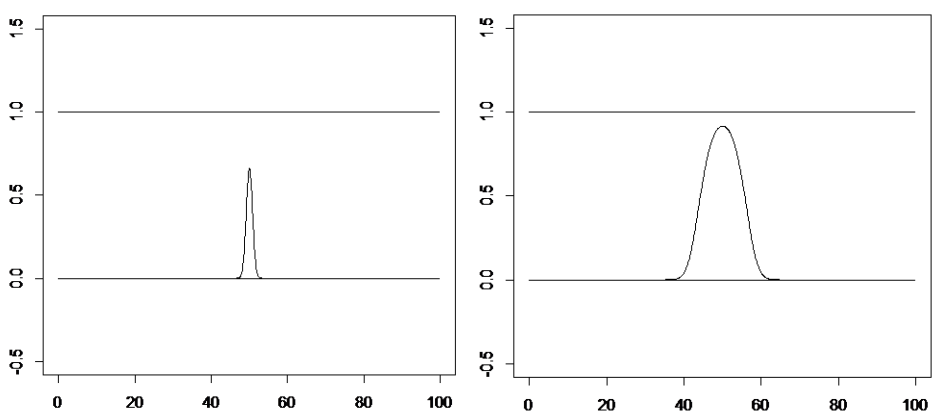
$$\begin{aligned}\frac{\partial v}{\partial t} &= -\frac{\partial u}{\partial t} = -\frac{1}{2}\Delta u - u(1-u) \\ &= \frac{1}{2}\Delta v - uv = \frac{1}{2}\Delta v - (1-v)v\end{aligned}$$

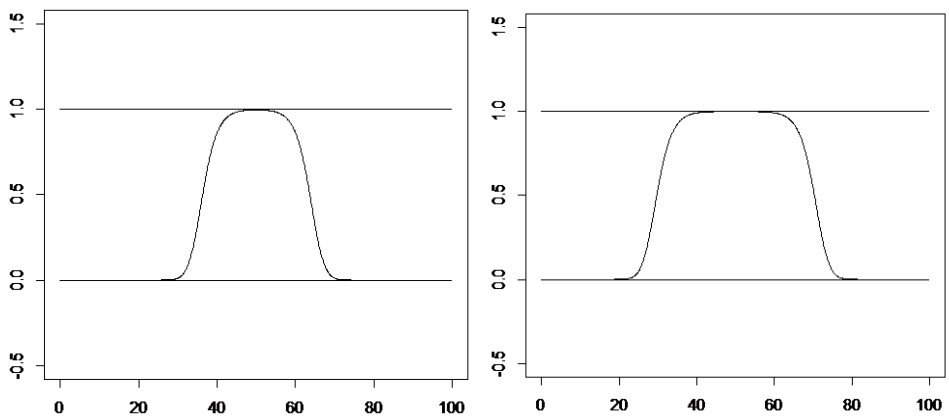
hence we may either study (2.2), or

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v + v^2 - v.$$

2.1.2 Simulations

The R codes for the numerical simulations shown here are given in the Appendix (Chapter 7).





2.2 Traveling waves

Simulations with $0 \leq u \leq 1$ clearly reveal the existence of a front, a wave, which moves. Assume there is a solution of the form

$$u(t,x) = w(x - ct).$$

Substituting into equation (2.2) we get

$$-cw' = \frac{1}{2}w'' + w - w^2. \tag{2.3}$$

Assume we start with the initial condition $u(0,x) = 1_{x<0}(x)$. If it happens that it becomes closer and closer to one of these traveling waves (for a suitable value of c), then the function w must fulfill

$$\lim_{x\to-\infty} w(x) = 1, \qquad \lim_{x\to+\infty} w(x) = 0. \tag{2.4}$$

Let us also ask that w is decreasing.

Theorem 7 *If $0 \leq c < \sqrt{2}$, there are no solutions of (2.3)-(2.4). If $c \geq \sqrt{2}$, there exists one and only one solution, denoted in the sequel by $w_c(x)$.*

Proof. Let us sketch the proof of [7]. Equation (2.3) is equivalent to the system

$$\begin{aligned} w' &= z \\ z' &= -2cz - 2w + 2w^2. \end{aligned}$$

To have a better dynamical intuition, let us write t for x (time) and (x,y) for (w,w'). Hence we study the system

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= -2cy(t) - 2x(t) + 2x(t)^2. \end{aligned}$$

We are looking for solutions $(x(t), y(t))$ defined on the whole $\mathbb{R}$, such that

$$\lim_{t\to-\infty} x(t) = 1, \qquad \lim_{t\to+\infty} x(t) = 0.$$

The fixed points (x, y) of the system satisfy

$$\begin{aligned} y &= 0 \\ -2cy - 2x + 2x^2 &= 0 \end{aligned}$$

hence $-2x + 2x^2 = 0$, $x = 0$ or $x = 1$, namely

$$A = (0, 0), \qquad B = (1, 0).$$

Denoting by $f(x, y)$ the vector field on the right of the system, we have

$$Df(x, y) = \begin{pmatrix} 0 & 1 \\ -2 + 4x & -2c \end{pmatrix}$$

hence

$$Df(A) = \begin{pmatrix} 0 & 1 \\ -2 & -2c \end{pmatrix}, \qquad Df(B) = \begin{pmatrix} 0 & 1 \\ 2 & -2c \end{pmatrix}.$$

With some computations (here done by Maple) one discovers that $Df(A)$ has eigenvectors

$$e_1^A = \begin{pmatrix} \frac{1}{2}\sqrt{c^2-2} - \frac{1}{2}c \\ 1 \end{pmatrix}, \qquad e_2^A = \begin{pmatrix} -\frac{1}{2}c - \frac{1}{2}\sqrt{c^2-2} \\ 1 \end{pmatrix}$$

with eigenvalues

$$-c - \sqrt{c^2-2}, \qquad \sqrt{c^2-2} - c$$

while $Df(B)$ has eigenvectors

$$e_1^B = \begin{pmatrix} \frac{1}{2}c - \frac{1}{2}\sqrt{c^2+2} \\ 1 \end{pmatrix}, \qquad e_2^B = \begin{pmatrix} \frac{1}{2}c + \frac{1}{2}\sqrt{c^2+2} \\ 1 \end{pmatrix}$$

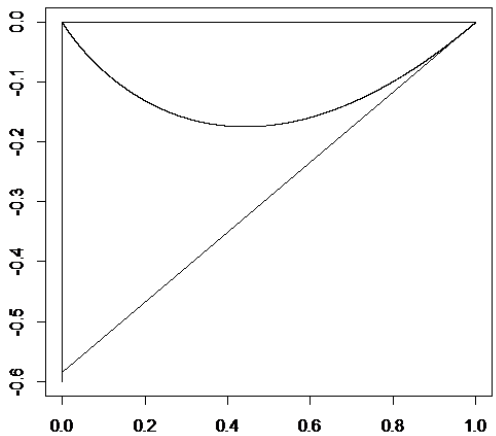
with eigenvalues

$$-c - \sqrt{c^2+2}, \qquad \sqrt{c^2+2} - c.$$

We always have $c \geq 0$, hence both eigenvalues of $Df(A)$ are negative, therefore A is a locally attractive stationary point; while $Df(B)$ has eigenvalues of opposite sign, namely it is an hyperbolic point with unstable manifold tangent to e_2^B, which spans a line through the first and third quadrant.

The intuition is the following one: if we have a solution, it arises at time $-\infty$ from B (since $\lim_{t\to-\infty} x(t) = 1$) and converges to A as $t \to +\infty$ (because $\lim_{t\to+\infty} x(t) = 0$); from B, being forced to exit along the unstable manifold, and being constrained to have

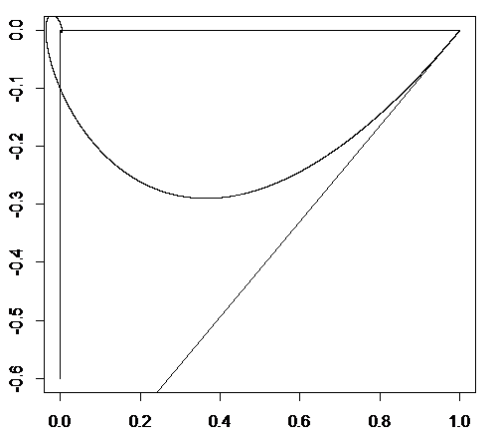
$x \in [0,1]$, it must exit in the third quadrant (centered at B), along $-e_2^B$. The next picture shows a numerical simulation of the system, for $c = \sqrt{2}$, with an initial condition of the form $-\epsilon e_2^B$ with very small ϵ. It goes precisely to A.



The R codes for the numerical simulations shown here are given in the Appendix (Chapter 7).

A rigorous proof that, for $c \geq \sqrt{2}$, a solution exists, requires additional arguments which we omit. One should for instance consider the triangle shown in the figure, given by the axes and the line through $-e_2^B$, and prove that the vector field f is directed inside the triangle at every non critical boundary point.

The non existence for $c < \sqrt{2}$ can be understood, on the contrary, from the fact that, for $c < \sqrt{2}$, point A attracts spiraling, as shown in the figure ($c = 0.8$):



which violates the constraint $x \in [0,1]$. Spiraling can be analytically seen looking for the solutions of the linearized system $x' = y$, $y' = -2cy - 2x$, namely $x'' = -2cx' - 2x$, which have the form $C_1 \exp\left(-t\left(c + \sqrt{c^2-2}\right)\right) + C_2 \exp\left(-t\left(c - \sqrt{c^2-2}\right)\right)$, understanding the exponential also in the complex case; they are decreasing exponentials for $c \geq \sqrt{2}$, but

oscillate for $c < \sqrt{2}$. ■

Theorem 8 *If $u(0,x) = 1_{x<0}(x)$, the solution converges to the traveling profile with $c = \sqrt{2}$. Precisely*

$$\lim_{t\to+\infty} u(t, m_t + x) = w_{\sqrt{2}}(x)$$

for every $x \in \mathbb{R}$, where m_t is the unique point such that $u(t, m_t) = 1/2$.

A probabilistic proof can be found in [10].

Remark 9 *One should not think that every initial condition converges to such traveling wave. One can prove, for instance, that if $u_0 \in [0,1]$ is such that for some $b \in (0, \sqrt{2}]$ the limit $\lim_{x\to+\infty} e^{bx}(1 - f(x))$ exists finite and different from zero, then $\lim_{t\to+\infty} u(t, ct + x) = w_c(x)$, with $c = 1/b + b/2$.*

2.3 A probabilistic representation

Consider the equation

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v + \lambda\left(v^2 - v\right).$$

McKean proved the following probabilistic representation

$$u(t,x) = 1 - E\left[\prod_{i=1}^{N_t} u_0\left(x + X_t^i\right)\right]$$

where the processes X_t^i, N_t are defined as follows. First we give an informal definition, then we formalize it. At time $t = 0$, from $x = 0$ a Brownian motion X_t^1 starts. At the random time $T_0 \sim Exp(\lambda)$, independent of X^1, the process X^1 ends its existence and, at position $X_{T_0}^1$, two new and independent Brownian motions start, X_t^1 and X_t^2. Each one lives an exponential time $Exp(\lambda)$ (independent of the previous objects) then dies and generates two new Brownian motions; and so on. At any time t there are N_t points alive, that we call X_t^i, $i = 1, ..., N_t$. They are the random points which appear in the probabilistic representation formula above.

Let us write a rigorous scheme. On a probability space $(\Omega, \mathcal{F}, P)$ assume we have a countable family of independent exponential times, $Exp(\lambda)$, indexed by finite sequences of 1,2: $T_{i_1,...,i_n}$, $i_k \in \{1,2\}$, $k = 1, ..., n$, $n \in \mathbb{N}$; for $n = 0$ let us write T_0. This family identifies a binary tree with branches of random length, but the tree structure in itself can be maintained behind, hidden, in or description. Set $\tau_{i_1,...,i_n} = T_0 + T_{i_1} + T_{i_1,i_2} + ... + T_{i_1,...,i_n}$, $\tau_0 = T_0$. Assume moreover that we have a countable family of independent Brownian motions (independent among themselves and with respect to the random times above) $B_{i_1,...,i_n}(t)$, $i_k \in \{1,2\}$, $k = 1, ..., n$, $n \in \mathbb{N}$; for $n = 0$ let us write $B_0(t)$. Now define a

countable family of processes $X_{i_1,...,i_n}(t)$, $i_k \in \{1,2\}$, $k = 1,...,n$, $n \in \mathbb{N}$, taking values in $\mathbb{R} \cup \delta$, where δ is an auxiliary point outside $\mathbb{R}$; for $n = 0$ let us write $X_0(t)$. All processes X take the values δ except on a random time interval that we are going to define. For $t \in [0, \tau_0)$ set $X_0(t) = B_0(t)$. For $t \in [\tau_0, \tau_{i_1})$, $i_1 = 1,2$, set $X_{i_1}(t) = X_0(\tau_0) + B_{i_1}(t - \tau_0)$. For $t \in [\tau_{i_1}, \tau_{i_1+i_2})$, $i_1, i_2 \in \{1,2\}$, set $X_{i_1+i_2}(t) = X_{i_1}(\tau_{i_1}) + B_{i_1+i_2}(t - \tau_{i_1})$. By induction, for $t \in [\tau_{i_1,...,i_n}, \tau_{i_1,...,i_n,i_{n+1}})$ set $X_{i_1,...,i_n,i_{n+1}}(t) = X_{i_1,...,i_n}(\tau_{i_1,...,i_n}) + B_{i_1,...,i_n,i_{n+1}}(t - \tau_{i_1,...,i_n})$. Denote by Λ_t the set of multi-indices $(i_1,...,i_n)$ such that $X_{i_1,...,i_n}(t) \neq \delta$. The precise meaning of the formula above is now

$$u(t,x) = 1 - E\left[\prod_{a \in \Lambda_t} u_0(x + X_a(t))\right].$$

From this formula one can immediately see a few facts, like the property $u(t,x) \in [0,1]$ when $u_0 \in [0,1]$, or the fact that the function $x \mapsto u(t,x)$ is decreasing if u_0 is decreasing. In [10] it is used to prove several facts stated above on traveling waves.

Let us finally see the proof of the formula. We disintegrate the expected value with respect to T_0:

$$v(t,x) :=$$

$$\begin{aligned}
E\left[\prod_{a \in \Lambda_t} u_0(x + X_a(t))\right] &= P(T_0 > t)\, E\left[\prod_{a \in \Lambda_t} u_0(x + X_a(t)) \,|T_0 > t\right] \\
&\quad + \int_0^t \lambda e^{-\lambda s} E\left[\prod_{a \in \Lambda_t} u_0(x + X_a(t)) \,|T_0 = s\right] ds \\
&= e^{-\lambda t} E\left[u_0(x + B_0(t))\right] \\
&\quad + \int_0^t \lambda e^{-\lambda s} E\left[\prod_{a \in \Lambda_t^1} u_0(x + X_a(t)) \prod_{a \in \Lambda_t^2} u_0(x + X_a(t)) \,|T_0 = s\right] ds
\end{aligned}$$

where Λ_t^k is the set of elements $(i_1,...,i_n) \in \Lambda_t$ such that $i_1 = k$, $k = 1,2$

$$\begin{aligned}
&= e^{-\lambda t}\left(e^{\frac{1}{2}\Delta t} u_0\right)(x) + \int_0^t \lambda e^{-\lambda s} E\left[E\left[\prod_{a \in \Lambda_{t-s}} u_0\left(x' + X_a(t-s)\right)\right]^2_{x' = x + B_0(s)}\right] ds \\
&= \left(e^{\left(\frac{1}{2}\Delta - \lambda\right)t} u_0\right)(x) + \int_0^t \left(e^{\left(\frac{1}{2}\Delta - \lambda\right)s} \lambda v^2(t-s,\cdot)\right)(x)\, ds
\end{aligned}$$

and this is the PDE in the mild sense.

2.4 Is FKPP the macroscopic equation of the tree structure?

The previous probabilistic formula is based on a particle system which seems a reasonable microscopic model of cells: each cell moves like a Brownian motion and at a random time duplicates. However, FKPP is *not* the macroscopic limit of this particle system. To see this, let us discover the true macroscopic limit of this particle system.

In Chapter 1 we have shown (in greater generality) that given a family of independent Brownian motions B_t^i, w.r.t. a filtration $\mathcal{F}_t$, and i.i.d. $\mathcal{F}_0$-adapted initial conditions X_0^i, the empirical measure $S_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ of the particle system

$$X_t^i = X_0^i + B_t^i$$

weakly converges to a measure-valued solution of the heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u$$

with initial condition given by the law of X_0^1. In this particular case we also have the representation formula

$$u(t,x) = E\left[u_0\left(X_t^1\right)\right]$$

but this coincidence should not be translated to more non-linear mechanisms.

Consider now the process defined in the previous section and let us investigate its macroscopic limit. What do we send to infinity? As in the case just recalled of independent Brownian motions B_t^i, let us consider an family of N independent processes of that form, with i.i.d. initial conditions X_0^i, independent of the Brownian motions and exponential times of the process. Consider the empirical measure

$$S_t^N := \frac{1}{N}\sum_{i=1}^N \sum_{a\in\Lambda_t^i} \delta_{X_a^i(t)}.$$

Theorem 10 *S_t^N weakly converges to a measure-valued solution of the equation*

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \lambda u. \tag{2.5}$$

Proof. Let us apply also here the LLN:

$$\left\langle S_t^N, \phi\right\rangle = \frac{1}{N}\sum_{i=1}^N \sum_{a\in\Lambda_t^i} \phi\left(X_a^i(t)\right) \to E\left[\sum_{a\in\Lambda_t} \phi\left(X_a(t)\right)\right] =: \left\langle u_t, \phi\right\rangle.$$

Now we have to find the equation satisfied by the time-dependent probability measure u_t. Introduce $\rho_\phi(t,x)$:

$$\begin{aligned}
\rho_\phi(t,x) &:= E\left[\sum_{a\in\Lambda_t}\phi(x+X_a(t))\right] \\
&= P(T_0>t)\,E\left[\sum_{a\in\Lambda_t}\phi(x+X_a(t))\,|T_0>t\right]+\int_0^t \lambda e^{-\lambda s}E\left[\sum_{a\in\Lambda_t}\phi(x+X_a(t))\,|T_0=s\right]ds \\
&= e^{-\lambda t}E\left[\phi(x+B_0(t))\right]+\int_0^t \lambda e^{-\lambda s}E\left[\sum_{a\in\Lambda_t^1}\phi(x+X_a(t))+\sum_{a\in\Lambda_t^2}\phi(x+X_a(t))\,|T_0=s\right]ds \\
&= e^{-\lambda t}\left(e^{\frac{1}{2}\Delta t}\phi\right)(x)+\int_0^t \lambda e^{-\lambda s}E\left[2E\left[\sum_{a\in\Lambda_{t-s}}\phi\left(x'+X_a(t-s)\right)\right]_{x'=x+B_0(s)}\right]ds \\
&= \left(e^{\left(\frac{1}{2}\Delta-\lambda\right)t}\phi\right)(x)+2\int_0^t\left(e^{\left(\frac{1}{2}\Delta-\lambda\right)s}\lambda\rho_\phi(t-s,\cdot)\right)(x)\,ds.
\end{aligned}$$

This implies

$$\begin{aligned}
\frac{\partial}{\partial t}\rho_\phi(t,x) &= \frac{1}{2}\Delta\rho_\phi(t,x)-\lambda\rho_\phi(t,x)+2\lambda\rho_\phi(t,x) \\
&= \frac{1}{2}\Delta\rho_\phi(t,x)+\lambda\rho_\phi(t,x) \\
\rho_\phi(0,x) &= \phi(x).
\end{aligned}$$

We aim to prove that u_t satisfies

$$\frac{\partial u_t}{\partial t}=\frac{1}{2}\Delta u_t+\lambda u_t$$

knowing that $\rho_\phi(t,0)=\langle u_t,\phi\rangle$ satisfies the equation written above. We have $\rho_\phi(t,x)=\langle u_t,\phi(x+\cdot)\rangle$ hence

$$\frac{\partial}{\partial t}\langle u_t,\phi(x+\cdot)\rangle=\frac{1}{2}\Delta\langle u_t,\phi(x+\cdot)\rangle+\lambda\langle u_t,\phi(x+\cdot)\rangle$$

but $\Delta\langle u_t,\phi(x+\cdot)\rangle=\langle u_t,(\Delta\phi)(x+\cdot)\rangle$ hence

$$\frac{\partial}{\partial t}\langle u_t,\phi(x+\cdot)\rangle=\frac{1}{2}\langle u_t,(\Delta\phi)(x+\cdot)\rangle+\lambda\langle u_t,\phi(x+\cdot)\rangle$$

hence $(x=0)$

$$\frac{\partial}{\partial t}\langle u_t,\phi\rangle=\frac{1}{2}\langle u_t,\Delta\phi\rangle+\lambda\langle u_t,\phi\rangle$$

that is the PDE we wanted, in weak form. ■

Remark 11 *Equation (2.5) is not FKPP.*

Remark 12 *The solutions equation (2.5) increase exponentially. The solutions of FKPP cannot overcome a threshold.*

Remark 13 *In order to find a microscopic model of FKPP we have to modulate the birth rate by the factor* $1-u(t,x)$*. But* $u(t,x)$ *is not known at the microscopic level. One could then use* $1-\left(\theta_{\epsilon_N} * S_t^N\right)(x)$*, where* θ_{ϵ_N} *is a mollifier which goes to the delta Dirac at zero. It allows one to talk of "*S_t^N *at point* x*". We shall deal with this problem in Chapter 4.*

2.5 Existence, uniqueness, invariant regions

Consider the equation (we drop 1/2 since we do not use probabilistic arguments in this section)

$$\frac{\partial u}{\partial t} = \Delta u + u(1-u) \tag{2.6}$$

for $t \in [0,T]$ and $x \in \mathbb{R}^d$. We are interested in solutions which are bounded, precisely $u \in [0,1]$, and of class $C^{1,2}$ ($u \in C^{1,2}$ if the partial derivatives $\frac{\partial u}{\partial t}$, $\partial_{x_i} u$, $\partial_{x_j}\partial_{x_i} u$ exist and are continuous); we shall call them *classical solutions*.

About the regularity of the initial condition u_0, maybe we ask a little bit too much. Let $UC_b\left(\mathbb{R}^d\right)$ be the space of uniformly continuous and bounded functions $f : \mathbb{R}^d \to \mathbb{R}$. We assume that $u_0 \in C^2\left(\mathbb{R}^d\right)$ with u and its first and second derivatives in $UC_b\left(\mathbb{R}^d\right)$; we write $u_0 \in UC_b^2\left(\mathbb{R}^d\right)$. We address [19], Chapter 14, Section A for a discussion of the advantages of $UC_b\left(\mathbb{R}^d\right)$ in this framework.

When we write expressions like $u_0 \in [0,1]$, $u \in [0,1]$, we mean that the property holds for all values of the independent variables x, (t,x) etc.

Theorem 14 *Equation (2.6) with* $UC_b^2\left(\mathbb{R}^d\right)$*-initial condition* $u_0 \in [0,1]$*, has one and only one classical solution* $u \in [0,1]$ *(uniqueness holds in the larger class of mild solutions, see the proof; the solution is also* $UC_b\left(\mathbb{R}^d\right)$ *at every time, plus other regularity properties).*

Proof. Let us only sketch the main steps, some of which will be expanded in Section 2.6.

Step 1. We prove the result for the auxiliary equation

$$\frac{\partial u}{\partial t} = \Delta u + u(1-u) + h(u) \tag{2.7}$$

where $h(\cdot)$ is a smooth function, strictly negative for $u=1$, strictly positive for $u=0$. This step is divided in three sub-steps.

Notice that the introduction of the auxiliary term $h(u)$ is not always needed. It depends on the proof we make of the invariance of the region $[0,1]$ (Step 1c). We present two proofs; the first one does not need $h(u)$, the second (deeper) one requires $h(u)$.

Step 1a. For a general equation of the form

$$\frac{\partial u}{\partial t} = \Delta u + f(u) \tag{2.8}$$

with smooth f, one can prove local existence and uniqueness of *mild solutions* (bounded continuous functions satisfying the mild form of the equation). If we then prove an a priori bound on solutions in uniform norm, this will lead to global existence (of mild solutions).

Step 1b. One can prove that mild solutions are also classical ones, when u_0 is regular.

Step 1c. For the particular equation (2.7) an a priori bound in uniform norm is proved (this is the main conceptual step). Precisely, we prove that, if u is a classical solution with $u_0(x) \in [0,1]$, then $u(t,x) \in [0,1]$. Putting together the three sub-steps, one gets the theorem for equation (2.7).

Step 2. (This step is needed only if we had to introduce $h(u)$ above.) Consider now the problem

$$\frac{\partial u^\delta}{\partial t} = \Delta u^\delta + b(t,x) \cdot \nabla u^\delta + u^\delta \left(1 - u^\delta\right) - \delta h\left(u^\delta\right)$$

where $h(u)$ is like above. We have $u^\delta \in [0,1]$ by Step 1. One can prove that u^δ converges uniformly on compact sets (first locally in time, then globally), as $\delta \to 0$. Then one can prove that the limit u is a solution of the original equation. Therefore there exists a solution $u \in [0,1]$.

Step 3. Uniqueness holds locally by Step 1a. This completes the proof.

See Section 2.6 for additional details. ■

2.6 Some proofs about Theorem 14

2.6.1 Homogeneous and non-homogeneous heat equation

Consider equation

$$\frac{\partial u}{\partial t} = \Delta u + g, \qquad u|_{t=0} = u_0 \tag{2.9}$$

where $u_0 \in UC_b^2\left(\mathbb{R}^d\right)$, $g : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is of class $C\left([0,T]; UC_b^2\left(\mathbb{R}^d\right)\right)$. We say that $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a *classical solution* if it is of class $C^{1,2}$ and satisfies the equation pointwise, on $[0,T] \times \mathbb{R}^d$. One can check that, for $g = 0$ and $u_0 \in UC_b^2\left(\mathbb{R}^d\right)$, there exists one and only one classical solution, given by

$$u(t,x) = \int_{\mathbb{R}^d} G(t, x-y)\, u_0(y)\, dy$$

$$G(t,x) = (4\pi k t)^{-d/2} \exp\left(-\frac{|x|^2}{4kt}\right).$$

That this expression defines a function of class $C^{1,2}$ and that this function satisfies the equation can be checked by several but elementary computations, that we do not repeat. Conversely, if u is a classical solution, we multiply the equation by $G(t, x' - x)$, integrate on $\mathbb{R}^d$, integrate by parts and get the formula.

Introduce the operator $P_t : C_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ given by

$$(P_t v)(x) = \int_{\mathbb{R}^d} G(t, x - y)\, v(y)\, dy, \qquad v \in C_b\left(\mathbb{R}^d\right).$$

One can check that, for $g \in C\left([0,T]; UC_b^2\left(\mathbb{R}^d\right)\right)$, there exists one and only one classical solution of equation (2.9), given by

$$u(t,x) = (P_t u_0)(x) + \int_0^t (P_{t-s} g(s))(x)\, ds.$$

2.6.2 Mild and regular solutions (inside Step 1b)

Definition 15 *We say that $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a mild solution of equation (2.8) if it is of class $C_b(\mathbb{R}^d)$ (bounded continuous) and satisfies pointwise the identity*

$$u(t,x) = (P_t u_0)(x) + \int_0^t (P_{t-s} f(u(s)))(x)\, ds.$$

Proposition 16 *If u is a classical solution, then it is a mild solution. If u is a mild solution and $u_0 \in UC_b^2(\mathbb{R}^d)$, then it is a classical solution.*

2.6.3 Local existence and uniqueness

Assume that f, in equation (2.8), is a locally Lipschitz continuous function.

Theorem 17 *Given $u_0 \in UC_b(\mathbb{R}^d)$, there exists one and only one mild solution of equation (2.8) on some interval $[0,\tau]$. The size τ depends only on the uniform norm of u_0.*

Proof. Consider the Banach space $C\left([0,\tau]; UC_b\left(\mathbb{R}^d\right)\right)$ and let $X_{\tau,\theta} \subset C\left([0,\tau]; UC_b\left(\mathbb{R}^d\right)\right)$ be the set of all functions v such that

$$\|v\|_0 := \sup_{(t,x)\in[0,\tau]\times\mathbb{R}^d} |v(t,x)| \le \theta.$$

Given $v \in X_{\tau,\theta}$, set

$$(\Gamma v)(t,x) = (P_t u_0)(x) + \int_0^t (P_{t-s} f(v(s)))(x)\, ds$$

for $t \in [0,\tau]$. The function Γv is well defined and of class $C\left([0,\tau]\,;UC_b\left(\mathbb{R}^d\right)\right)$. We have

$$\|\Gamma v\|_0 \leq \sup_{(t,x)\in[0,\tau]\times\mathbb{R}^d} \left| \int_0^t \left(P_{t-s} f\left(v\left(s\right)\right)\right)(x)\, ds \right| \leq C_{f,\theta}\tau$$

where

$$C_{f,\theta} = \max_{|u|\leq\theta} |f\left(u\right)|\,.$$

Moreover we have

$$\|\Gamma v - \Gamma w\|_0 \leq \sup_{(t,x)\in[0,\tau]\times\mathbb{R}^d} \left| \int_0^t \left(P_{t-s}\left(f\left(v\left(s\right)\right) - f\left(w\left(s\right)\right)\right)\right)(x)\, ds \right|$$

$$\leq \int_0^\tau L_{f,\theta} \|v-w\|_0\, ds = \tau L_{f,\theta} \|v-w\|_0$$

where

$$L_{f,\theta} = \max_{|u|\leq\theta} |\nabla f\left(u\right)|\,.$$

It is now easy to verify that, for τ small enough, the set $X_{\tau,\theta}$ is invariant under Γ and on such set Γ is a contraction. This gives us local existence and uniqueness, by the contraction principle. ■

2.6.4 Invariant regions (Step 1c and completion of Step 1)

Approach by linear equations

Let u be a local classical solution. Just for notational simplicity, we assume it is define on $[0,T]$. Setting $\lambda\left(t,x\right) = 1 - u\left(t,x\right)$, we have

$$\frac{\partial u}{\partial t} = \Delta u + \lambda u.$$

For this kind of equations it is well known that:

Lemma 18 *If u is a classical solution on $[0,T]$ with $u_0 \geq 0$, then $u \geq 0$ on $[0,T]$.*

Proof. Among different proof let us choose a probabilistic one: by the well-known Feynman-Kac formula

$$u\left(t,x\right) = E\left[e^{\int_0^t \lambda\left(t-s,x+\sqrt{2}B_s\right)ds} u_0\left(x+\sqrt{2}B_t\right)\right]$$

we immediately deduce the result. Here B_t is a Brownian motion in $\mathbb{R}^d$. Let us recall the proof of Feynman-Kac formula in this simple case. Given $T_0 > 0$, by Itô formula for $t \in [0, T_0]$ we have

$$\begin{aligned} du\left(T_0 - t, x + \sqrt{2}B_t\right) &= -\frac{\partial u}{\partial t}\left(T_0 - t, x + \sqrt{2}B_t\right) dt + \nabla u\left(T_0 - t, x + \sqrt{2}B_t\right) \cdot \sqrt{2} dB_t \\ &\quad + \Delta u\left(T_0 - t, x + \sqrt{2}B_t\right) dt \\ &= -\lambda u\left(T_0 - t, x + \sqrt{2}B_t\right) dt + \nabla u\left(T_0 - t, x + \sqrt{2}B_t\right) \cdot \sqrt{2} dB_t \end{aligned}$$

hence the process $\xi_t := u\left(T_0 - t, x + \sqrt{2}B_t\right)$ satisfies $d\xi_t = -\lambda_t \xi_t dt + dM_t$, where $\lambda_t = \lambda\left(T_0 - t, x + \sqrt{2}B_t\right)$, $M_t = \int_0^t \nabla u\left(T_0 - s, x + \sqrt{2}B_s\right) \cdot \sqrt{2} dB_s$, whence

$$d\left(e^{\int_0^t \lambda_s ds}\xi_t\right) = \lambda_t\left(e^{\int_0^t \lambda_s ds}\xi_t\right) dt + e^{\int_0^t \lambda_s ds}\left(-\lambda_t \xi_t dt + dM_t\right) = e^{\int_0^t \lambda_s ds} dM_t.$$

This implies

$$E\left[e^{\int_0^{T_0} \lambda_s ds} u_0\left(x + \sqrt{2}B_{T_0}\right)\right] = u\left(T_0, x\right)$$

which is the desired formula. ■

Now let us set, as we have done above, $v = 1 - u$. We have

$$\frac{\partial v}{\partial t} = \Delta v - v\left(1 - v\right)$$

hence

$$\frac{\partial v}{\partial t} = \Delta v + \widetilde{\lambda} v$$

where $\widetilde{\lambda} = v - 1 = -u$. By the lemma we deduce

Corollary 19 *If u is a classical solution on $[0, T]$ with $u_0 \leq 1$, then $u \leq 1$ on $[0, T]$.*

Proof. The lemma applied to the equation for v gives us $v \geq 0$ on $[0, T]$. Hence $u \leq 1$ on $[0, T]$. ■

Approach by the theory of invariant regions

The previous approach is very fast but it has the disadvantage to be applicable only when the algebraic structure of the terms is particularly simple. In order to deal with the full system described in Chapter 1, we think it is convenient to know the following somewhat deeper approach. We address to [19], Chapter 14 for more details on this approach.

Our aim is to prove that the interval $[0, 1]$ is an invariant region for equation (2.7) (it will be also for equation (2.6)).

Let u be a classical solution of

$$\frac{\partial u}{\partial t} = \Delta u + u(1-u) + h(u)$$

on some interval $[0,\tau]$. We want to prove that, if $u_0 \in [0,1]$, then $u \in [0,1]$ on $[0,\tau]$. For simplicity of notation we take $\tau = T$.

Let us argue by contradiction; hence, assume that $u_0 \in [0,1]$ but $[0,1]$ is not invariant (there is some (t',x') with $u(t',x') \notin [0,1]$).

Assume that, due to the lack of invariance, we can prove that there is $t_0 \in [0,T]$ with the following properties:

i) $t_0 > 0$

ii) $u(t,x) \in [0,1]$ for every $x \in \mathbb{R}^d$ and every $t \in [0,t_0]$

iii) $u(t_0,x_0) \in \{0,1\}$ for some $x_0 \in \mathbb{R}^d$.

Let us analyze the case $u(t_0,x_0) = 1$. We deduce:

a) $\nabla u(t_0,x_0) = 0$ and $\Delta u(t_0,x_0) \leq 0$ (because $u(t_0,x) \leq 1$ for all x and $u(t_0,x_0) = 1$, and $u(t_0,\cdot)$ is C^2)

b) $\frac{\partial u}{\partial t}(t_0,x_0) < 0$ (from the equation, because $\Delta u(t_0,x_0) \leq 0$, $u(t_0,x_0)(1-u(t_0,x_0)) = 0$, $h(u(t_0,x_0)) < 0$)

c) $\frac{\partial u}{\partial t}(t_0,x_0) \geq 0$ (because $u(t,x_0) \leq 1$ for $t \in [0,t_0]$ and $u(t_0,x_0) = 1$).

Thus there is a contradiction. In the case $u(t_0,x_0) = 0$ we have

a') $\nabla u(t_0,x_0) = 0$ and $\Delta u(t_0,x_0) \geq 0$ (because $u(t_0,x) \geq 0$ for all x and $u(t_0,x_0) = 0$, and $u(t_0,\cdot)$ is C^2)

b') $\frac{\partial u}{\partial t}(t_0,x_0) > 0$ (from the equation, because $\Delta u(t_0,x_0) \geq 0$, $u(t_0,x_0)(1-u(t_0,x_0)) = 0$, $h(u(t_0,x_0)) > 0$)

c') $\frac{\partial u}{\partial t}(t_0,x_0) \leq 0$ (because $u(t,x_0) \geq 0$ for $t \in [0,t_0]$ and $u(t_0,x_0) = 0$)

hence again a contradiction.

Thus it remains to prove that $t_0 \in [0,T]$ exists with properties (i)-(iii). Intuitively this is reasonable, so we may split the proof in two parts: the main conceptual part is the one just given; a second part deals with the proof of existence of such t_0. This second part is complicated by two potential difficulties.

Let us start with the definition of t_0:

$$t_0 = \inf\left\{t' \in [0,T] : \exists x' \text{ such that } u\left(t',x'\right) \notin [0,1]\right\}$$

or equivalently

$$t_0 = \sup\left\{t \in [0,T] : \forall x \text{ we have } u(t,x) \in [0,1]\right\}. \tag{2.10}$$

The first difficulty comes from the fact that t_0 may be equal to zero, violating property (i). To avoid this, one could argue as follows: we first assume $u_0 \in [\epsilon, 1-\epsilon]$ for some $\epsilon > 0$. In this case, since we know from the existence theorem that $t \mapsto u(t)$ is continuous in the space $UC_b\left(\mathbb{R}^d\right)$, there is some small time interval $[0,\delta]$ such that that $u(t,x) \in [\epsilon/2, 1-\epsilon/2]$ for $t \in [0,\delta]$; therefore t_0 defined above by (2.10) is strictly positive. Later, when the

proof of invariance is done for initial conditions satisfying $u_0 \in [\epsilon, 1-\epsilon]$, given a general $u_0 \in [0,1]$ it is sufficient to take $u_0^n = \left(1 - \frac{1}{n}\right) u_0 + \frac{1}{2n}$ and prove that the solutions u^n with initial conditions u_0^n converge uniformly on compact sets to u. The continuous dependence on initial conditions can be established similarly to the proof of existence by contraction principle (or the proof of continuous dependence on δ in step 2 above).

The second difficulty is related to lack of compactness. Given the definition above of t_0, with properties (i)-(ii), can we say that (iii) holds? We know that for every $\epsilon > 0$ there is $t'_\epsilon \in (t_0, t_0 + \epsilon)$ and $x'_\epsilon \in \mathbb{R}^d$ such that $u(t'_\epsilon, x'_\epsilon) \notin [0,1]$, so by continuity there is $t''_\epsilon \in [t_0, t_0 + \epsilon)$ and $x''_\epsilon \in \mathbb{R}^d$ such that $u(t''_\epsilon, x''_\epsilon) \in \{0,1\}$. Obviously $\lim_{\epsilon \to 0} t''_\epsilon = t_0$, but what about the family $\{x''_\epsilon\}$?

One way to solve this problem is to prove that there is a compacts set $K \subset \mathbb{R}^d$ such that $u(t,x) \in (0,1)$ for all $t \in [0,T]$ and $x \notin K$. When this is proved, we know that $\{x''_\epsilon\} \subset K$, hence we may extract a sequence which converges to a point x_0, so that $u(t_0, x_0) \in \{0,1\}$.

The existence of K requires some technical work but, conceptually, it is based on simple ingredients: the property of decay to zero at infinity is preserved by the heat semigroup, the image of the heat semigroup is made of functions which are strictly positive everywhere (infinite speed of diffusion). Notice that this way of solving the problem restricts the set of initial conditions: they have to decay to zero at infinity. Using other arguments one can treat more general $u_0 \in [0,1]$.

2.6.5 Continuous dependence on h (Step 2)

This part can be done in two ways. One is by estimating the distance between u and u^δ, by triangle inequality and suitable estimates. The other by proving that Ascoli-Arzelà theorem can be applied, and then passing to the limit.

Chapter 3

Remarks on the full invasive model with angiogenesis

3.1 Invariant regions for the full system?

3.1.1 The total volume constraint is not fulfilled

The full system of Chapter 1 seems apparently based on the same ingredients of the FKPP equation and contains terms, like the proliferation term $\mathcal{N}(\mathcal{V}_{\max} - \mathcal{V})$, which have been devised having in mind that the total volume $\mathcal{V}$ keeps below a maximum value $\mathcal{V}_{\max}$.

However, there are troubles. To understand a first one, let us consider the reduced system

$$\begin{aligned}
\frac{\partial \mathcal{N}}{\partial t} &= k_1 \Delta \mathcal{N} + c_1 \mathcal{N}(\mathcal{V}_{\max} - \mathcal{V}) - \alpha_{\mathcal{N}\to\mathcal{A}} \mathcal{N} \\
\frac{d\mathcal{A}}{dt} &= \alpha_{\mathcal{N}\to\mathcal{A}} \mathcal{N}
\end{aligned}$$

with $\mathcal{V} = \mathcal{N} + \mathcal{A}$.

The natural invariant region is now

$$\Sigma = \{(\mathcal{N}, \mathcal{A}) : \mathcal{N} \geq 0, \mathcal{A} \geq 0, \mathcal{N} + \mathcal{A} \leq \mathcal{V}_{\max}\}.$$

For the biological meaning of the equation, it is necessary to prove that Σ is invariant: if the constraints $\mathcal{N} \geq 0, \mathcal{A} \geq 0$ are lost, the meaning of $\mathcal{N}$, $\mathcal{A}$ is lost; if the constraint $\mathcal{N} + \mathcal{A} \leq \mathcal{V}_{\max}$ is violated, the proliferation term $c_1 \mathcal{N}(\mathcal{V}_{\max} - \mathcal{V})$ changes sign, it becomes a killing term, without meaning!

But Σ is not invariant. The properties $\mathcal{N} \geq 0, \mathcal{A} \geq 0$ persist but the constraint $\mathcal{N} + \mathcal{A} \leq \mathcal{V}_{\max}$ is violated. The best way to understand it is to repeat the proof of

invariance and see where it fails. So, assume we are dealing with the equation modified by $h(u)$. Assume there is $t_0 \in [0,T]$ with the following properties:

i) $t_0 > 0$

ii) $(\mathcal{N},\mathcal{A})(t,x) \in \Sigma$ for every $x \in \mathbb{R}^d$ and every $t \in [0,t_0]$

iii) $(\mathcal{N}+\mathcal{A})(t_0,x_0) = \mathcal{V}_{\max}$ for some $x_0 \in \mathbb{R}^d$.

Can we deduce a contradiction? Following the intuition and also the prescriptions of [19], Chapter 14, we should consider the function $\mathcal{V} = \mathcal{N}+\mathcal{A}$ and compute $\frac{\partial \mathcal{V}}{\partial t}$, looking for a contradiction of signs. We have

$$\frac{\partial \mathcal{V}}{\partial t} = k_1 \Delta \mathcal{N} + c_1 \mathcal{N}(\mathcal{V}_{\max} - \mathcal{V}) + h$$

because

$$\begin{aligned}\frac{\partial \mathcal{V}}{\partial t} &= \frac{\partial \mathcal{N}}{\partial t} + \frac{\partial \mathcal{A}}{\partial t} \\ &= k_1 \Delta \mathcal{N} + c_1 \mathcal{N}(\mathcal{V}_{\max} - \mathcal{V}) - \alpha_{\mathcal{N}\to\mathcal{A}}\mathcal{N} + \alpha_{\mathcal{N}\to\mathcal{A}}\mathcal{N} + h \\ &= k_1 \Delta \mathcal{N} + c_1 \mathcal{N}(\mathcal{V}_{\max} - \mathcal{V}) + h.\end{aligned}$$

From (i)-(iii) above we deduce

$$\frac{\partial \mathcal{V}}{\partial t}(t_0,x_0) \geq 0$$

$$\nabla \mathcal{V}(t_0,x_0) = 0, \qquad \Delta \mathcal{V}(t_0,x_0) \leq 0$$

and from the equation (recall h)

$$\frac{\partial \mathcal{V}}{\partial t}(t_0,x_0) < k_1 \Delta \mathcal{N}(t_0,x_0)$$

but we have no reason to claim that $\Delta \mathcal{N}(t_0,x_0) \leq 0$ and deduce a contradiction.

This argument is not only logical: it indicates what goes wrong: when and where $\mathcal{V}$ approaches the threshold, $\Delta \mathcal{N}$ may remain > 0, with the effect that some portion of $\mathcal{N}$ diffuses there instead of diffusing away. $\mathcal{N}$ may continue to increase when and where $\mathcal{V} = \mathcal{V}_{\max}$.

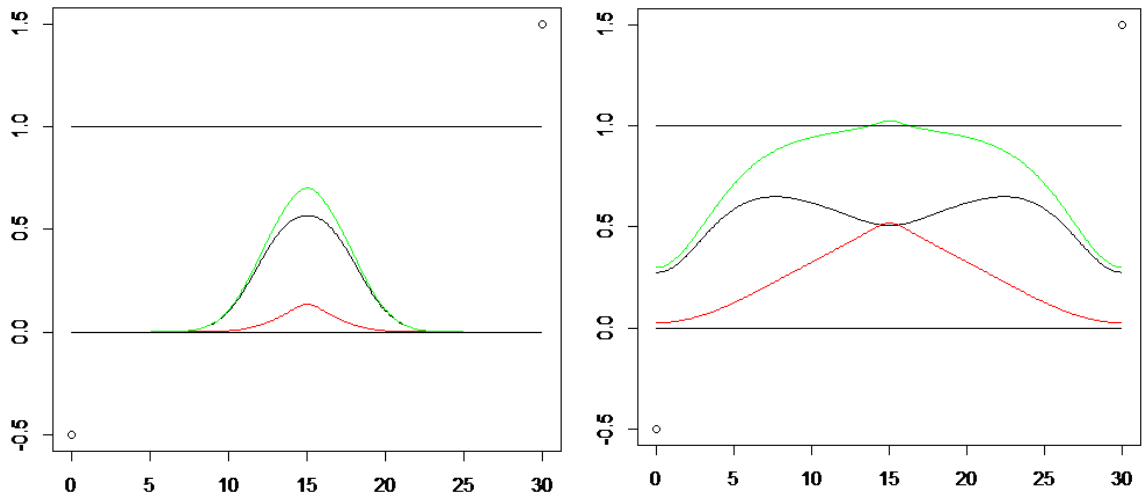
Are we wrong? Is there another argument which proves that Σ is invariant? Unfortunately not. I am sure that one can construct an analytic proof based on the following idea, but mostly the numerical simulation below is unambiguous. The analytical idea, in dimension 1, is to take at time $t = 0$

$$\begin{aligned}\mathcal{N}(0,x) &= \mathcal{V}_{\max} \cdot 1_{[5,10]}(x) + \mathcal{V}_{\max} \cdot 1_{[-10,-5]}(x) \\ \mathcal{A}(0,x) &= \mathcal{V}_{\max} \cdot 1_{[-5,5]}(x).\end{aligned}$$

The constraint $\mathcal{N}+\mathcal{A} \leq \mathcal{V}_{\max}$ is fulfilled. But a second later part of the density $\mathcal{N}$ will be diffused in the region $[-5,5]$, because of the term $k_1\Delta\mathcal{N}$, while $\mathcal{A}$ cannot go down in that region. Thus we shall have points in the interval $[-5,5]$ where $\mathcal{N}+\mathcal{A} > \mathcal{V}_{\max}$.

It remains the doubt that this problem is caused by the simplification of this reduced system, with respect to the full one, but I do not think so. A more rigorous or precise investigation should be done.

The figure shows normoxic in black, apoptotic in red, total volume in green. The R codes for the numerical simulations shown here are given in the Appendix (Chapter 7).



Problem 20 *How could we modify the system to have Σ invariant?*

3.1.2 Invariant regions for different diffusion operators

Consider, as a model problem, the equation

$$\frac{\partial u}{\partial t} = \mathcal{D}u + u\,(1-u)$$

where $\mathcal{D}u$ is one of the following operators, all (except the first one) having something to do with the idea that the diffusion is damped when we approach the threshold $u = 1$:

$$\begin{aligned}
\mathcal{D}_1 u &= \Delta u \\
\mathcal{D}_2 u &= (1-u)\,\Delta u \\
\mathcal{D}_3 u &= \operatorname{div}\left((1-u)\,\nabla u\right) \\
\mathcal{D}_4 u &= \Delta\left((1-u)\,u\right).
\end{aligned}$$

We state the next result as a Proposition but it is only the verification of a key ingredient of the result, following the proof of invariant regions given in Chapter 2.

Proposition 21 *The region $[0,1]$ is invariant in cases 1,2,3. In case 4 solutions remain positive but the constraint $u \leq 1$ is not necessarily preserved.*

Proof. Assume we are dealing with the equation modified by $h(u)$. Assume there is $t_0 \in [0,T]$ with the following properties:

i) $t_0 > 0$

ii) $u(t,x) \in [0,1]$ for every $x \in \mathbb{R}^d$ and every $t \in [0,t_0]$

iii) $u(t_0,x_0) \in \{0,1\}$ for some $x_0 \in \mathbb{R}^d$.

Let us analyze the case $u(t_0,x_0) = 1$. We deduce:

a) $\nabla u(t_0,x_0) = 0$ and $\Delta u(t_0,x_0) \leq 0$

b) $\frac{\partial u}{\partial t}(t_0,x_0) \geq 0$ (because $u(t,x) \leq 1$ for $t \in [0,t_0]$ and $u(t_0,x_0) = 1$).

These facts were already known. In addition we now have:

$$\begin{aligned} \mathcal{D}_2 u(t_0,x_0) &= 0 \\ \mathcal{D}_3 u(t_0,x_0) &= 0 \\ \mathcal{D}_4 u &= -u(t_0,x_0)\,\Delta u(t_0,x_0) \geq 0 \end{aligned}$$

because

$$\begin{aligned} \mathcal{D}_3 u &= \nabla(1-u)\cdot\nabla u + \mathcal{D}_2 u = -|\nabla u|^2 + \mathcal{D}_2 u \\ \mathcal{D}_4 u &= \mathcal{D}_2 u - u\Delta u - 2|\nabla u|^2 . \end{aligned}$$

Hence, from the equation:

c) $\frac{\partial u}{\partial t}(t_0,x_0) < 0$ in cases 1, 2, 3, but not in case 4.

Then we get a contradiction in the first three cases only.

Let us finally analyze the case $u(t_0,x_0) = 0$. We deduce:

a) $\nabla u(t_0,x_0) = 0$ and $\Delta u(t_0,x_0) \geq 0$

b) $\frac{\partial u}{\partial t}(t_0,x_0) \leq 0$

These facts were already known. In addition we now have:

$$\begin{aligned} \mathcal{D}_2 u(t_0,x_0) &\geq 0 \text{ (we know that } 1-u \geq 0) \\ \mathcal{D}_3 u(t_0,x_0) &\geq 0 \\ \mathcal{D}_4 u &\geq 0. \end{aligned}$$

Hence, from the equation:

c) $\frac{\partial u}{\partial t}(t_0,x_0) > 0$ in all cases.

Then we get a contradiction in all cases. ■

Consider now the following operators, all (except the first one) having something to do with the idea that the diffusion increases when the density is larger (like for pressure-driven diffusion):

$$\begin{aligned} \mathcal{D}_5 u &= u\Delta u \\ \mathcal{D}_6 u &= \operatorname{div}(u\nabla u) = \frac{1}{2}\Delta u^2 \end{aligned}$$

Proposition 22 *The region* $[0,1]$ *is invariant in both cases 5,6.*

Proof. Assume we are dealing with the equation modified by $h(u)$. Assume there is $t_0 \in [0,T]$ with the following properties:

i) $t_0 > 0$

ii) $u(t,x) \in [0,1]$ for every $x \in \mathbb{R}^d$ and every $t \in [0,t_0]$

iii) $u(t_0,x_0) \in \{0,1\}$ for some $x_0 \in \mathbb{R}^d$.

Let us analyze the case $u(t_0,x_0) = 1$. We deduce:

a) $\nabla u(t_0,x_0) = 0$ and $\Delta u(t_0,x_0) \leq 0$

b) $\frac{\partial u}{\partial t}(t_0,x_0) \geq 0$ (because $u(t,x_0) \leq 1$ for $t \in [0,t_0]$ and $u(t_0,x_0) = 1$).

These facts were already known. In addition we now have:

$$\mathcal{D}_5 u(t_0,x_0) \leq 0$$
$$\mathcal{D}_6 u(t_0,x_0) \leq 0$$

because

$$\mathcal{D}_6 u = \operatorname{div}(u\nabla u) = |\nabla u|^2 + \mathcal{D}_5 u.$$

Hence, from the equation:

c) $\frac{\partial u}{\partial t}(t_0,x_0) < 0$ in both cases 5, 6.

Then we get a contradiction in both cases.

Let us finally analyze the case $u(t_0,x_0) = 0$. We deduce:

a) $\nabla u(t_0,x_0) = 0$ and $\Delta u(t_0,x_0) \geq 0$

b) $\frac{\partial u}{\partial t}(t_0,x_0) \leq 0$

These facts were already known. In addition we now have:

$$\mathcal{D}_5 u(t_0,x_0) = 0$$
$$\mathcal{D}_6 u(t_0,x_0) = 0.$$

Hence, from the equation:

c) $\frac{\partial u}{\partial t}(t_0,x_0) > 0$ in both cases.

Then we get a contradiction in both cases. ■

Remark 23 *The intuition (concerning the constraint $u \leq 1$) in the group 2-3 is opposite to the case of group 5-6 but the result is the same. The factor $(1-u)$ damps the diffusion when we approach $u = 1$, so the diffusion term does not contribute to the motion of mass; what happens depends on the other terms of the equation. The factor u increases the diffusion when u is larger and thus helps to damp the profile, when it is closer to $u = 1$.*

Remark 24 *If the other terms of the equation preserve the constraint $u \leq 1$, the factor $(1-u)$ is just innocuous, it is neutral; but it cannot help if someone of the other terms violates the constraint. On the contrary, the factor u has a damping effect and thus it may help.*

Remark 25 *The simpler method of invariance described in Chapter 2, based on the linear equation, does not apply here, at least not so easily.*

Remark 26 *If the region* $[0,1]$ *is preserved by two of the previous operators* $\mathcal{D}_i$ *and* $\mathcal{D}_j$*, then it is preserved by the sum* $\mathcal{D}_i + \mathcal{D}_j$*. In particular, the example*

$$\frac{\partial u}{\partial t} = k_1 \Delta u + k_2 \Delta u^2 + u(1-u)$$

preserves $[0,1]$*.*

Exercise 27 *Investigate the invariance of* $[0,1]$ *for the crowding-driven diffusion operator*

$$\mathcal{D}_7 u = \operatorname{div}(\sigma(u) \nabla u)$$

with

$$\sigma(u) = \max(u - u_0, 0)$$

for some $u_0 \in (0,1)$*.*

3.1.3 Back to the system of Section 3.1.1

The previous section investigates nonlinear operators for single equations but the ideas can be extended to systems. Let us solve Problem 20 using these ideas. One possibility is to consider the system

$$\begin{aligned}
\frac{\partial \mathcal{N}}{\partial t} &= k_1 (\mathcal{V}_{\max} - \mathcal{V}) \Delta \mathcal{N} + c_1 \mathcal{N} (\mathcal{V}_{\max} - \mathcal{V}) - \alpha_{\mathcal{N} \to \mathcal{A}} \mathcal{N} \\
\frac{d\mathcal{A}}{dt} &= \alpha_{\mathcal{N} \to \mathcal{A}} \mathcal{N}
\end{aligned}$$

namely to replace $\Delta \mathcal{N}$ by $(\mathcal{V}_{\max} - \mathcal{V}) \Delta \mathcal{N}$. Let us show that the method of invariant regions work for this case.

Proposition 28 *The region*

$$\Sigma = \{(\mathcal{N}, \mathcal{A}) : \mathcal{N} \geq 0, \mathcal{A} \geq 0, \mathcal{N} + \mathcal{A} \leq \mathcal{V}_{\max}\}$$

is invariant.

Proof. Assume we are dealing with the equation modified by $h(u)$. Assume there is $t_0 \in [0,T]$ with the following properties: ■

i) $t_0 > 0$

ii) $(\mathcal{N}, \mathcal{A})(t,x) \in \Sigma$ for every $x \in \mathbb{R}^d$ and every $t \in [0, t_0]$

iii) $(\mathcal{N}+\mathcal{A})(t_0,x_0)=\mathcal{V}_{\max}$ for some $x_0\in\mathbb{R}^d$.
Consider the function $\mathcal{V}=\mathcal{N}+\mathcal{A}$. We have

$$\begin{aligned}\frac{\partial\mathcal{V}}{\partial t}&=\frac{\partial\mathcal{N}}{\partial t}+\frac{\partial\mathcal{A}}{\partial t}\\&=k_1\left(\mathcal{V}_{\max}-\mathcal{V}\right)\Delta\mathcal{N}+c_1\mathcal{N}\left(\mathcal{V}_{\max}-\mathcal{V}\right)+h\end{aligned}$$

From (i)-(iii) above we deduce

$$\frac{\partial\mathcal{V}}{\partial t}(t_0,x_0)\geq 0$$

$$\nabla\mathcal{V}(t_0,x_0)=0,\qquad \Delta\mathcal{V}(t_0,x_0)\leq 0$$

and from the equation (recall h)

$$\frac{\partial\mathcal{V}}{\partial t}(t_0,x_0)<0$$

hence we have a contradiction. Concerning the constraint $\mathcal{N}\geq 0$, at the critical point, using $\mathcal{V}\leq\mathcal{V}_{\max}$, we get

$$\begin{aligned}\frac{\partial\mathcal{N}}{\partial t}(t_0,x_0)&=k_1\left(\mathcal{V}_{\max}-\mathcal{V}(t_0,x_0)\right)\Delta\mathcal{N}(t_0,x_0)+c_1\mathcal{N}(t_0,x_0)\left(\mathcal{V}_{\max}-\mathcal{V}(t_0,x_0)\right)-\alpha_{\mathcal{N}\to\mathcal{A}}\mathcal{N}(t_0,x_0)\\&=k_1\left(\mathcal{V}_{\max}-\mathcal{V}(t_0,x_0)\right)\Delta\mathcal{N}(t_0,x_0)\geq 0\end{aligned}$$

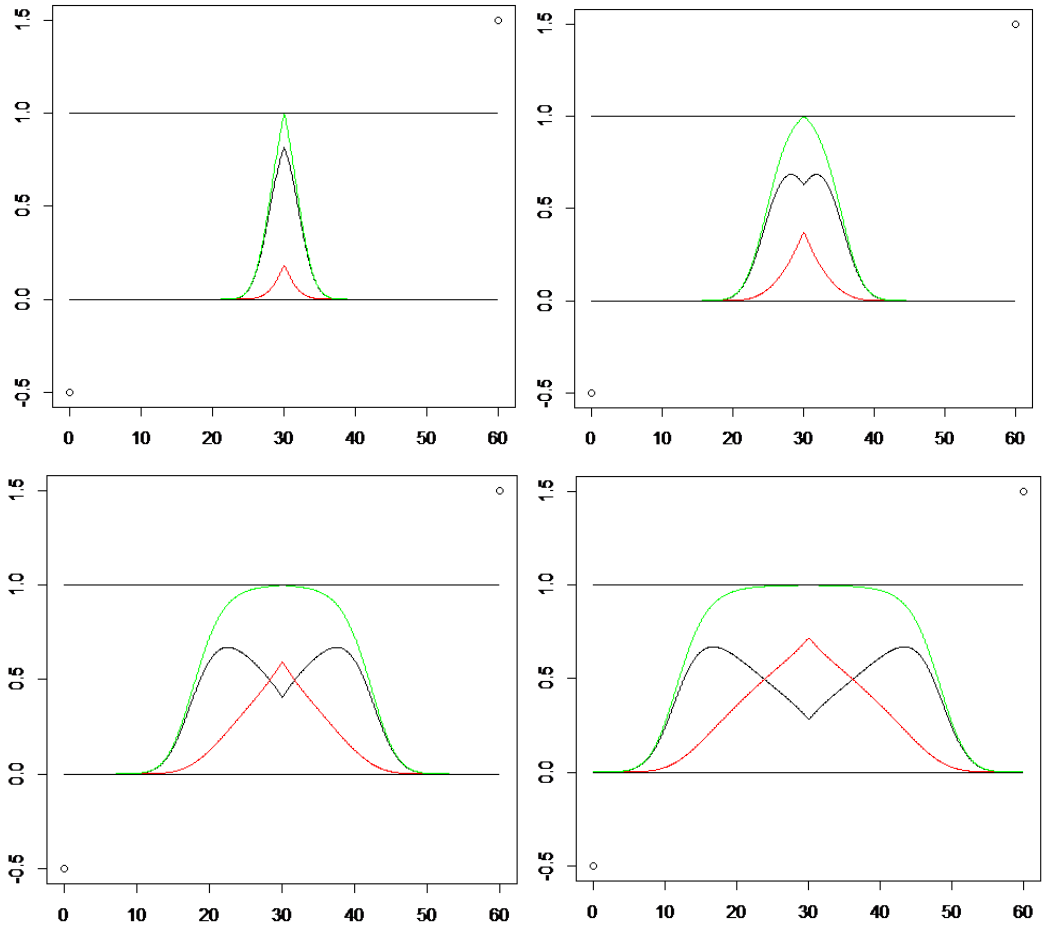
which yields the contradiction. The constraint $\mathcal{A}\geq 0$ is then obvious, since $\mathcal{A}$ is non decreasing.

Exercise 29 *Investigate the invariance of Σ when the diffusion operator is of the form*

$$\begin{aligned}&\operatorname{div}\left(\left(\mathcal{V}_{\max}-\mathcal{V}\right)\nabla\mathcal{N}\right)\\&\Delta\left(\mathcal{V}_{\max}-\mathcal{V}\right)\mathcal{N}\\&\mathcal{W}\Delta\mathcal{N}\\&\operatorname{div}\left(\mathcal{W}\nabla\mathcal{N}\right)\\&\Delta\left(\mathcal{W}\mathcal{N}\right).\end{aligned}$$

with $\mathcal{W}=\mathcal{V}$ or $\mathcal{W}=\mathcal{N}$.

The following numerical simulations confirm the prediction of the theorem. The R codes for the numerical simulations shown here are given in the Appendix (Chapter 7).



3.1.4 Brainstorming on the problem of bounds on cell density

1. The bound on the total cell density $V \leq V_{\max}$ is something new, not common when we deal with fluids or gases of molecules. For cells, it has a meaning.

2. But is necessary to impose it strictly, or a mild form is sufficient? Living tissues may deform, stretch, may accommodate higher density a little bit. How to describe mathematically this mild accommodation possibility?

3. For a single equation like
$$\frac{\partial u}{\partial t} = \Delta u + b \cdot \nabla u + cu$$
the constraint
$$u \geq 0$$

holds (is maintained). On the contrary, the constraint

$$u \leq u_0$$

holds (is maintained) only when $c \leq 0$. Otherwise, we only have

$$u(t,x) \leq \sup u_0 \cdot e^{\sup[c]^+ t}.$$

4. For a single equation of Fokker-Planck type

$$\frac{\partial u}{\partial t} = \Delta u - \operatorname{div}(bu)$$

we have

$$u \geq 0$$
$$\int u dx = const$$

but only

$$u(t,x) \leq \sup u_0 \cdot e^{\sup[-\operatorname{div} b]^+ t}.$$

In other words, if the vector field b has positive or null divergence, all right, otherwise there is concentration and a threshold on the density may be easily overcome.

5. For a single equation, preservation of bounds like

$$0 \leq u \leq 1$$

hold also for certain diffusion operators different from Δu. See above.

6. But for systems, if we want to preserve $V(t,x) \leq V_{\max}$, diffusion terms *have* to be modified, Δu is not sufficient, as we have showed above.

7. Even worse, for systems, terms transport terms of the form

$$-\operatorname{div}(u \nabla \theta)$$

which couple different variables, may be very critical and difficult, they may even lead to blow-up. One should have and estimate on the quantity $[-\operatorname{div} \nabla \theta]^+ = [-\Delta \theta]^+$ which is not under control. Every place where θ is more concentrated becomes an attracting place.

8. Should we expect this phenomenon in our full model? If instead of having a smooth distribution of endothelial cells we more realistically assume they are concentrated in vessels, this could be more dangerous (but actually we do not know what precisely could happen). Presumably, with the present model, θ does not concentrate too much and thus, when V is not to close to $V_{\max}$, problems should not occur.

9. Even if we discover how to handle these complicate couplings, it is hardly plausible that invariant regions exist. One should then try to include special mechanisms into the model, to avoid concentration.

3.1.5 Transport terms and possibility of blow-up

It is well known that terms of the form $b \cdot \nabla u$ transport the profile u along characteristic lines; hence u cannot increase or decrease beyond its maximum and minimum values, due to the term $b \cdot \nabla u$.

But all our terms are in divergence form

$$\operatorname{div}\left(b\left(t,x\right)u\left(t,x\right)\right). \tag{3.1}$$

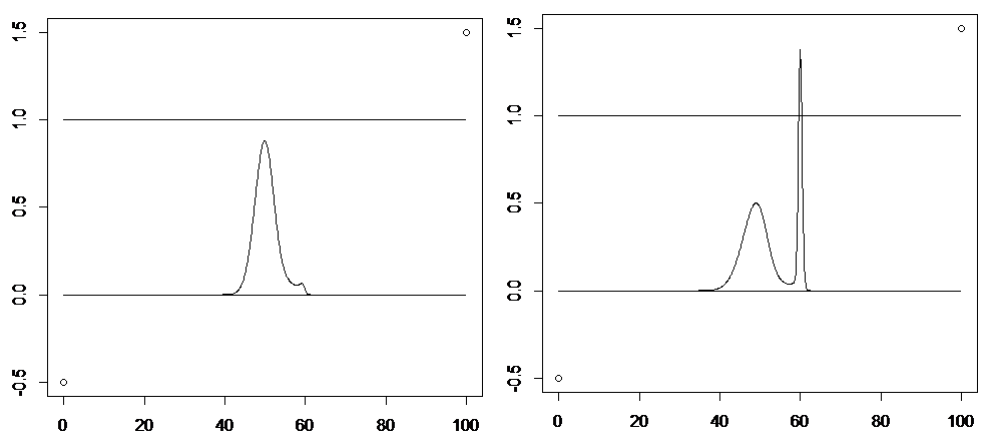
This reduces to $b \cdot \nabla u$ only when b is divergence free. Under (3.1), the profile u may increase beyond the original maximum, when the flux lines concentrate. The next pictures show the case of the simple heat equation with transport ($d = 1$)

$$\frac{\partial u}{\partial t} = \Delta u - \operatorname{div}\left(bu\right)$$

along the vector field

$$b\left(x\right) = -5\left(x - x_0\right)e^{-0.08\left|x-x_0\right|^2}$$

which is very concentrated around x_0 and has flux lines which move in the direction of x_0 from both sides:



The divergence of b around x_0 is roughly $\operatorname{div} b\left(x\right) \sim -5e^{-0.08\left|x-x_0\right|^2}$. See the code in the appendix.

On the contrary, let us show that positivity is maintained. We do it for instance for

$$\begin{aligned}
\frac{\partial \mathcal{N}}{\partial t} &= k_1\left(\mathcal{V}_{\max} - \mathcal{V}\right)\Delta\mathcal{N} + \operatorname{div}\left(b\mathcal{N}\right) + c_1\mathcal{N}\left(\mathcal{V}_{\max} - \mathcal{V}\right) - \alpha_{\mathcal{N}\to\mathcal{A}}\mathcal{N} \\
\frac{d\mathcal{A}}{dt} &= \alpha_{\mathcal{N}\to\mathcal{A}}\mathcal{N}.
\end{aligned}$$

If (t_0, x_0) is a usual critical point, where $\mathcal{N}(t_0, x_0) = 0$ and before $\mathcal{N}(t, x) \geq 0$, we get

$$\begin{aligned}\frac{\partial \mathcal{N}}{\partial t}(t_0, x_0) &= k_1 (\mathcal{V}_{\max} - \mathcal{V}(t_0, x_0)) \Delta \mathcal{N}(t_0, x_0) + \operatorname{div}(b(t_0, x_0) \mathcal{N}(t_0, x_0)) + c_1 \mathcal{N}(t_0, x_0)(\mathcal{V}_{\max} - \mathcal{V}(t_0, x_0)) - \alpha_{\mathcal{N}} \\ &= k_1 (\mathcal{V}_{\max} - \mathcal{V}(t_0, x_0)) \Delta \mathcal{N}(t_0, x_0) + b(t_0, x_0) \cdot \nabla \mathcal{N}(t_0, x_0) + \mathcal{N}(t_0, x_0) \operatorname{div} b(t_0, x_0) \\ &= k_1 (\mathcal{V}_{\max} - \mathcal{V}(t_0, x_0)) \Delta \mathcal{N}(t_0, x_0) \geq 0\end{aligned}$$

which yields a contradiction.

The proof that $\mathcal{V} \leq \mathcal{V}_{\max}$ is maintained, on the contrary, does not work, since

$$\begin{aligned}\frac{\partial \mathcal{V}}{\partial t}(t_0, x_0) &= k_1 (\mathcal{V}_{\max} - \mathcal{V}(t_0, x_0)) \Delta \mathcal{N}(t_0, x_0) + b(t_0, x_0) \cdot \nabla \mathcal{N}(t_0, x_0) + \mathcal{N}(t_0, x_0) \operatorname{div} b(t_0, x_0) + c_1 \mathcal{N}(t_0, x_0)(\mathcal{V}_{\mathrm{n}} \\ &= b(t_0, x_0) \cdot \nabla \mathcal{N}(t_0, x_0) + \mathcal{N}(t_0, x_0) \operatorname{div} b(t_0, x_0)\end{aligned}$$

and nothing is known of these two terms.

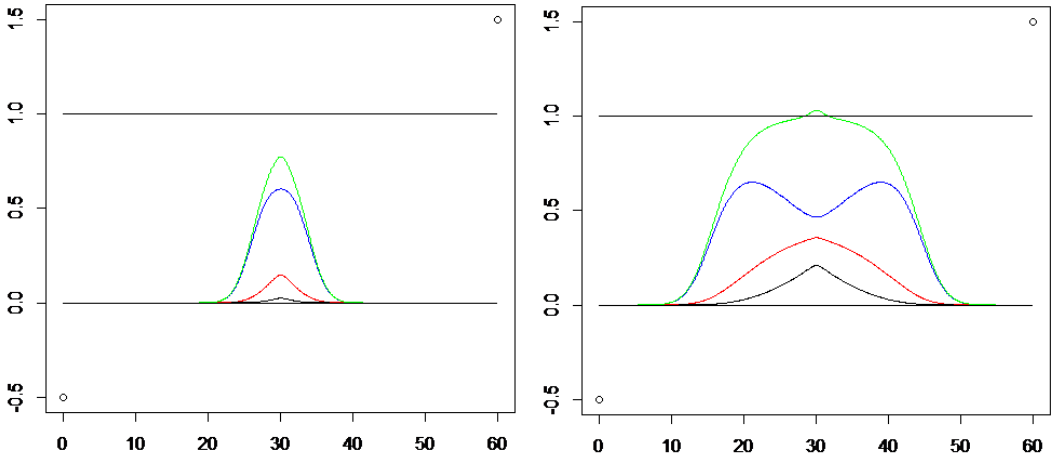
Not only the profile could increase beyond the required limits but, even worse, it may blow-up in special cases. It may happen when the drift b depends on the solution, maybe the solution of an associated PDE (it is the case of our example). Let us mention an example which received a lot of attention in the literature: the chemotaxis equations, in particular the Keller-Siegel model (see also models under the name "aggregation models"). An example is the system, in dimension ≥ 2

$$\begin{aligned}\frac{\partial u}{\partial t} &= \Delta u - \chi \operatorname{div}(u \nabla v) \\ \Delta v &= 1 - u.\end{aligned}$$

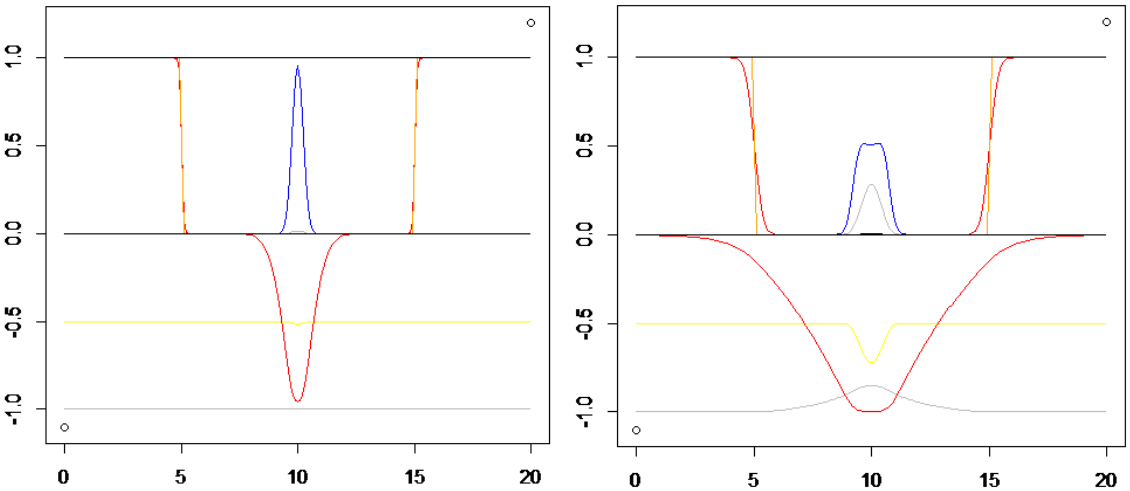
It is known that there is a value $\chi_* > 0$ such that, for all $\chi > \chi_*$, radially symmetric positive solutions can be constructed which blow-up in finite time. However, blow-up does not happen for small values of χ and in dimension 1.

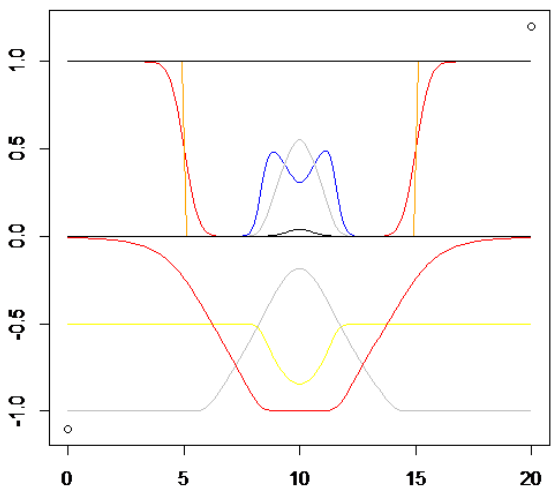
3.2 Simulations about the full system

First we show a reduced system with only normoxic (blue), hypoxic (red) and apoptotic (black) (total volume in green, again over the threshold) (no angiogenic cascade):



Then we show (a simplified version of) the full system. Each picture presents two figures, above the profiles of normoxic (blue), hypoxic (grey), apoptotic (black), endothelial (red) with the initial profile of endothelial (orange) for comparison (since they move very slowly). Below, we see the profile of oxygen (red) ECM (yellow), VEGF (grey).





The R codes for the numerical simulations shown here are given in the Appendix (Chapter 7).

3.3 Fick or Fokker-Planck?

Another unclear issue is concerned with the terms generically speaking of the form

$$\operatorname{div}\left(a\left(t,x\right)\nabla u\left(t,x\right)\right) \tag{3.2}$$

$$\Delta\left(a\left(t,x\right)u\left(t,x\right)\right). \tag{3.3}$$

Assume we want to include a function $a\left(t,x\right)$ in the diffusion term (usually $a\left(t,x\right)$ is not given a priori but it depends on the unknown densities), either because of inhomogenuities or, more important, to modulate the diffusion with some logic as described above. From the modelling viewpoint is it more correct (3.2) or (3.3)?

We mention these two forms for the following reason. Form (3.2) is the celebrated inhomogeneous Fick law, so often used in the biological literature. On the contrary, (3.3) is the Fokker-Planck type of diffusion, compare with the Fokker-Planck equation in Chapter 1.

Usually the literature based on Fick law does not even discuss the other possibility. There are a few papers in the physical literature which make a comparison between the two forms and have some tendency to prefer Fokker-Planck but one should carefully understand whether the objects considered there correspond to the applications to cell dynamics; this is not so clear. From my side the problem is unsolved, except that above we have seen a good algebraic reason (Proposition 21) to prefer Fick law (3.2), when $a\left(t,x\right)=1-u\left(t,x\right)$. Moreover, we shall see in Chapters 5-6 that the macroscopic limit of certain microscopic dynamics lead to Fick law.

A natural question for probabilists is: which SDE corresponds to Fick law? Consider the model problem, in Fick form

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial}{\partial x}\left(\sigma^2 \frac{\partial}{\partial x} p\right).$$

Obviously

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2 p\right) - \frac{\partial}{\partial x}\left(\sigma' \sigma p\right).$$

Hence, recalling the theorem on Fokker-Planck equation in Chapter 1, we may associate the SDE

$$dX_t = \sigma'\left(t, X_t\right)\sigma\left(t, X_t\right)dt + \sigma\left(t, X_t\right)dB_t$$

to the Fick diffusion. However, is this a natural equation, from the viewpoint of microscopic dynamics, in some suitable sense, or does it just look like an artificial association? Is there any physical justification of the drift term $\sigma'\left(t, X_t\right)\sigma\left(t, X_t\right)dt$? I cannot answer this question. Let me only remark that this is *not* Stratonovich equation

$$dX_t = \sqrt{2}\sigma\left(t, X_t\right) \circ dB_t.$$

If it were, since Stratonovich equations are for good reasons more physical than Itô equations, we would have a wonderful reason to prefer Fick diffusion. But it is not so. The rewriting of our Stratonovich equation in Itô form is

$$dX_t = \frac{1}{2}\sigma'\left(t, X_t\right)\sigma\left(t, X_t\right)dt + \sigma\left(t, X_t\right)dB_t.$$

Just by the factor $\frac{1}{2}$, we miss this very interesting interpretation.

3.4 Modelling the crowding-driven diffusion

3.4.1 Different nonlinear diffusion terms

We have found the nonlinear diffusion term

$$\operatorname{div}\left(\mathcal{N}\nabla\mathcal{N}\right).$$

The model at the beginning prescribed

$$\operatorname{div}\left(\max\left(\left(\mathcal{N} - \mathcal{N}_{\min}\right) \wedge 0\right)\nabla\mathcal{N}\right)$$

which is similar. Even if not equal, it is of Fick type!

Which one is more correct is not clear but I tend to vote for $\max\left(\left(\mathcal{N} - \mathcal{N}_{\min}\right) \wedge 0\right)$. If so, we have found a microscopic interacting model which is very similar to the expected one but not exactly.

We are working to find microscopic models which are more flexible, to produce

$$\operatorname{div}\left(g\left(\mathcal{N}\right)\nabla\mathcal{N}\right)$$

for suitable functions g.

Remark. It is known that hydrodynamic limits of discrete exclusion-type models also give rise to such Fick terms but perhaps not so flexible again.

3.5 Conclusions

The first three Chapters have been devoted mainly to the macroscopic viewpoint, based on nonlinear versions of Fokker-Planck equations, possibly systems. We have also discussed a few topics about the microscopic description by stochastic processes and SDEs, but the viewpoint has been to associate the correct processes and SDEs to the PDEs; we didn't analyze in detail the microscopic view. Thus, a first problem that we aim to treat in the next Chapters is:

- restart from the biology of cells and give a direct microscopic description; then develop the inverse path to the one above, namely try to go from micro to macro.

Concerning tools at the microscopic level, the simple diffusion of Brownian motion and transport terms are clear, in clear correspondence with the macroscopic ones. But we have understood very little of:

- the tools to describe proliferation
- how to get different nonlinear diffusion operators, by means of microscopic particle systems.

All these questions are interesting in themselves but may also help to throw light into the problems and difficulties which emerged in the previous macroscopic analysis:

- which models preserve, at least approximately, the global density of cells?
- which diffusion operators are more natural?

Part II

Microscopic models

Chapter 4

The tumor from the microscopic viewpoint

4.1 Cell models

Let us start this second part of the lectures from a microscopic viewpoint. We want to build microscopic models, based on stochastic systems, and investigate their macroscopic limits. Tumors are cell anomalies, opposite to other diseases which involve the function of tissues and organs. Hence a cell-level description is very appropriate, although until now we have given a tissue-level description to capture global features.

Let us distinguish between continuous and discrete space models, the first ones based on stochastic differential equations, the second ones on the so called cellular automata, where the position is constrained on some discrete structure.

A cell occupies a volume, a portion of space. It is not pointwise, as we usually assume to be molecules when we treat a gas. A cell can be widely deformed and a little bit compressed, but we cannot ignore the fact that it is not pointwise. Cellular automata take this fact into account intrinsically, when they ask that a cell may occupy only a position of a grid and each position may be occupied only by a cell (exclusion type models). If we consider only one class of cells, it is sufficient to associate 0 (empty) or 1 (occupied) to each site of the grid. If we consider cells of different types (like normoxic, hypoxic etc.), it is sufficient to associate colors or numbers 0, 1, ..., N to each site (the so called Potts models).

On the contrary, when we study continuous space models, and we associate a point, a position, to each cell, the volume constraint must be "simulated" by means of an interaction term between the pointwise cells, interaction that forbids (maybe in a mild sense, to accommodate some degree of compression and deformation) two point to become too close. One has to introduce a repulsive potential, ideally with compact support (opposite to most potentials of physics) because the volume constraint acts only when cells are close

each other. Sometimes also a damping term is introduced; combined with a non-compact support potential it has similar performances to compact support ones.

Normal cells, with the exception of a few types like those of the blood, have adhesion constraint, usually with cells of the same class. This adhesion guarantees solidity to tissues. Some tumor cells maintain this feature but others, after additional genetic changes, loose it. Usually, the cells of the initial phase of the tumor are still attached each other (tumor *in situ*); later, after several mutations, they loose this property and may move (invasive tumor), very slowly (on the scale of days) through the surrounding ECM and then at long distance if they succeed to penetrate blood vessels or lymphatic system. Sometimes they move in groups, like that adhesion is only partially lost.

In principle, we may choose between a Newtonian dynamics (second order differential equations) where we have to specify the set of forces acting on particles, and a first order dynamics, more phenomenological, where we have to prescribe the sources of variation of position. We follow the second approach.

To simplify the notations in this initial section, assume we have cells of only one type, just tumor cells. Denote their positions by $X^1(t)$, $X^2(t)$, ... If we take into account only volume and adhesion constraints, and we impose a Brownian behavior when the adhesion is lost, we have the following equations

$$\frac{d}{dt}X^i(t) = \sum_{j\neq i} K_{vol}\left(X^i(t) - X^j(t)\right) + \underbrace{\sum_{j\neq i} K_{adh}\left(X^i(t) - X^j(t)\right)}_{\text{in the adhesive phase}}$$

$$\frac{d}{dt}X^i(t) = \sum_{j\neq i} K_{vol}\left(X^i(t) - X^j(t)\right) + \underbrace{\frac{d}{dt}W^i(t)}_{\text{in the invasive phase}}$$

The volume interaction $K_{vol}(\cdot)$ may be assumed for simplicity of gradient form

$$K_{vol}(x) = -\nabla U_{vol}(x)$$

and in addition we may assume that the potential $U_{vol}(x)$ depends only on the distance, $U_{vol}(x) = U_{vol}(|x|)$ (with little abuse of notations) where $U_{vol}(r)$ is a function defined for $r > 0$. Hence

$$K_{vol}(x) = -U'_{vol}(|x|)\frac{x}{|x|}.$$

The volume constraint is repulsive. This means that the contribution $K_{vol}\left(X^i(t) - X^j(t)\right)$ to the displacement of $X^i(t)$ is oriented as $X^i(t) - X^j(t)$. This translates in the condition

$$U'_{vol}(r) < 0, \qquad r > 0$$

Moreover, the function $U_{vol}(\cdot)$ should be assumed compact support, with support related to the typical radius of a cell; with $K_{vol}(x)$ very strong, possibly diverging, when $x \to 0$, because some compression of the cell is acceptable but not to a point.

Similarly we assume that the adhesion interaction $K_{adh}(\cdot)$ is of gradient form, $K_{adh}(x) = -\nabla U_{adh}(x)$ and that $U_{adh}(x) = U_{adh}(|x|)$, hence

$$K_{adh}(x) = -U'_{adh}(|x|)\frac{x}{|x|}.$$

Opposite to the previous case, it is an attractive force, $K_{adh}\left(X^i(t) - X^j(t)\right)$ is oriented as $-\left(X^i(t) - X^j(t)\right)$. This translates into

$$U'_{had}(r) > 0, \qquad r > 0.$$

Moreover, $U_{adh}(x)$ acts only when the distance between two cells is close to the typical one: when the cells are far away, they do not attract each other. So also $U_{adh}(\cdot)$ is compact support, with a little larger support than $U_{vol}(\cdot)$.

The free motion when adhesion forces are lost has been described above by a simple Brownian motion, but this is clearly a first simplification. For instance one can include systematic drifts related to chemical gradients.

Sometimes a damping term of the form

$$-\alpha X^i(t)$$

is added, which generically describes the viscosity of the medium in which the motion takes place (the ECM). It is useful in simulations and it may allow one to use non-compact support potentials.

What has been described until now is just the motion, the part of the evolution of the system that is described by differential equations. In addition, we have to incorporate proliferation and change of type (from normoxic to hypoxic etc.) and cell death in some model (we mean that cells may disappear, natural after apoptosis). These aspects of the evolution are not incorporated into the differential equations but must be super-imposed by means of discrete-event techniques, like Poisson processes; we shall see below. For the time being, let us remark that proliferation of tumor cells is not unconditional but it may be slowed down or inhibited by excessive cell density or lack of nutrients.

We may now state the main purposes of our future investigation. We want to consider a large family of cells described by the microscopic dynamics introduced here and send the number of cells to infinity, namely investigate the macroscopic limit. We already guess that in the invasive phase the presence of Brownian motion will lead to a diffusion term of the form Δu. We already know that, if we incorporate a drift along a chemical gradient, it will converge to a transport term of the form $\operatorname{div}(u\nabla g)$. We guess that a suitable Poissonian mechanism of duplication, modulated by the density of cells, may converge to the proliferation terms of the form $u(1-u)$ (this fact is still among those we have to clarify at the technical level, recall the exponential growth model of Chapter 2). But what is more obscure is what happens to the local interactions given by the volume constraint (always working) and the adhesion constraint (working in the in situ phase). Maybe they

produce additional diffusion operators, presumably nonlinear, but of which shape? Can we explain the crowding-driven diffusion operator of the previous chapters in this way, or other interesting nonlinear operators investigated in Chapter 3, having the property of maintaining regions invariant? Could it be that we find operators which control the transport terms, also in presence of negative divergence of a drift?

To summarize, two specific aims are:

1. describe proliferation at microscopic scale and find its macroscopic limit

2. find the macroscopic limit of volume and adhesion constraint.

In the rest of this chapter we investigate 1 and also the in situ phase from some particular viewpoints, not yet those based on 2. In the next two chapters we shall then investigate more deeply the macroscopic limit of interacting particle systems and try to approach a partial answer to question 2.

4.2 The non invasive (*in situ*) phase

4.2.1 Prediction of the exponential model

Let us forget for a moment about the motion of cells and let us assume that each cell has an exponential clock such that, when it rings, the cell duplicates. And let us assume that the two new cells which replace the original cell will undergo the same mechanism, independently of each other and of the past. The spatial aspect will only be superimposed to the duplication mechanism.

Since exponential clocks are memory free, if we know the number N_t of cells at some time t, we may restart all the clocks and make future predictions independently of the past. Hence N_t can be taken as the "state" of the system.

It is a jump Markov process, on positive integers (a pure birth process, in the language of birth-and-death processes). The transition rate from n to $n+1$ is $n\lambda$

$$P\left(N_{t+h} = n+1 | N_t = n\right) = n\lambda h + o\left(h\right)$$

(the minimum between exponential r.v.'s $T_1, ..., T_n$ with parameters $\lambda_1, ..., \lambda_n$ is an exponential r.v. with parameter $\lambda_1 + ... + \lambda_n$) and it is equal to zero for all other transitions.

Approximately we thus have

$$\begin{aligned}
E\left[N_{t+h}\right] &= \sum_{n=0}^{\infty} E\left[N_{t+h}|N_t = n\right] P\left(N_t = n\right) \\
&\sim \sum_{n=0}^{\infty} \left(nP\left(N_{t+h} = n|N_t = n\right) + (n+1) P\left(N_{t+h} = n+1|N_t = n\right)\right) P\left(N_t = n\right) \\
&\sim \sum_{n=0}^{\infty} \left(n\left(1 - n\lambda h\right) + (n+1) n\lambda h\right) P\left(N_t = n\right) \\
&= E\left[N_t\right] + \lambda h E\left[N_t\right]
\end{aligned}$$

namely

$$\frac{d}{dt} E\left[N_t\right] = \lambda E\left[N_t\right]$$

which yields (with $E\left[N_0\right] = 1$, a single initial cell)

$$E\left[N_t\right] = e^{\lambda t}.$$

This result is simple but wrong: experimentally this exponential increase is roughly true only in very first stages of development of the tumor; very soon the number of cells increases like t^3. See [6].

4.2.2 Damping proliferation in absence of space and nutrients

Without spatial structure we cannot develop a precise argument but let us make a rough computation. Assume that, if the cells alive at time t, in number N_t, are grouped in roughly spherical fashion, just the external part of the sphere, of thickness δ_t, proliferates. Let us recall a few formulae about the sphere:

i) the surface of a sphere or radius R is $4\pi R^2$

ii) the volume of a sphere or radius R is $\frac{4}{3}\pi R^3$ (it comes from (i) integrating $4\pi r^2$ in r, from $r = 0$ to $r = R$)

iii) for a sphere or radius R, the volume of an external layer of thickness δ, namely of the portion of sphere which has distance less than δ from the boundary, is $\frac{4}{3}\pi R^3 - \frac{4}{3}\pi\left(R - \delta\right)^3$ (by difference between the volumes of the original sphere and the complementary part to the external layer, again a sphere).

Hence, assuming cells of equal volume V_{cell} and no empty space, the radius R_t at time t satisfies the relation

$$\frac{4}{3}\pi R_t^3 = V_{cell} N_t.$$

The external layer of thickness δ_t has volume $\frac{4}{3}\pi R_t^3 - \frac{4}{3}\pi\left(R_t - \delta_t\right)^3$ and thus the number M_t of cells which belong to it satisfies the relation

$$\frac{4}{3}\pi R_t^3 - \frac{4}{3}\pi\left(R_t - \delta_t\right)^3 = V_{cell} M_t.$$

If we assume that only the cells of the external layer proliferate, we get

$$N_{t+h} - N_t \sim \left(e^{\lambda h} - 1\right) M_t \sim \lambda h \left(\frac{4}{3V_{cell}} \pi R_t^3 - \frac{4}{3V_{cell}} \pi \left(R_t - \delta_t\right)^3\right)$$

namely

$$\frac{d}{dt} N_t \sim \lambda \left(\frac{4}{3V_{cell}} \pi R_t^3 - \frac{4}{3V_{cell}} \pi \left(R_t - \delta_t\right)^3\right)$$

namely

$$\frac{d}{dt} R_t^3 = \lambda \left(R_t^3 - \left(R_t - \delta_t\right)^3\right)$$

$$3R_t^2 \frac{d}{dt} R_t = \lambda \left(R_t^3 - \left(R_t - \delta_t\right)^3\right)$$

$$\frac{d}{dt} R_t = \lambda \frac{R_t^3 - \left(R_t - \delta_t\right)^3}{3R_t^2}.$$

Since

$$R_t^3 - \left(R_t - \delta_t\right)^3 = 3R_t^2 \delta_t - 3R_t \delta_t^2 + \delta_t^3$$

we get

$$\frac{d}{dt} R_t = \lambda \frac{3R_t^2 \delta_t - 3R_t \delta_t^2 + \delta_t^3}{3R_t^2} = \lambda \delta_t \left(-\frac{\delta_t}{R_t} + \frac{\delta_t^2}{3R_t^2}\right).$$

If δ_t is small with respect to R_t, essentially we obtain a linear growth, hence $N_t \sim t^3$.

4.2.3 Fisher-Kolmogorov

The model

$$\frac{\partial u}{\partial t} = k \Delta u + \chi u \left(u_{\max} - u\right)$$

contains some dynamical features that resembles those just described. First of all, let us clarify the meaning of $u(x)$. If a cell occupies a volume V_{cell}, in a (much larger) volume V we have at most $N_{\max} = \frac{V}{V_{cell}}$ cells. The density of cells, at this level of approximation, should be understood as the ratio between the number of cells and the volume under consideration; hence it is smaller than $\frac{N_{\max}}{V} = \frac{1}{V_{cell}}$. Hence

$$u_{\max} = \frac{1}{V_{cell}}.$$

Proliferation will be very week where the cells saturate the volume at disposal, intense where there are rarefied. Hence the term $\chi u \left(u_{\max} - u\right)$ corresponds to the ideas of the previous section.

The term $k \Delta u$ is necessary to increase the volume occupied by the cells. This mechanism of increase is not entirely correct, since it should be due just to pressure, not to

random Brownian diffusion. In this sense the term $k\Delta u$ is not realistic for *in situ* tumors. However, it is a good mathematical simplification. An interesting problem is to devise other operators, different from $k\Delta u$, which are more realistic.

Assume we start, at time $t = 0$, from a small cluster of cancer cells, roughly spherical, around the origin

$$u_0(x) = u_0 1_{|x|\leq\rho_0}(x).$$

How does the solution $u(t,x)$ develop? Let us discuss only a 1D model for shortness. From the theory of Chapter 2 we know that $u(t,x)$ remains between 0 and $u_{\max} = \frac{1}{V_{cell}}$.

Let us think to the evolution in time of the solution by means of a splitting method (we use it in our simulations, see Chapter 7). Because of the term $k\Delta u$, the density spreads; if ρ_0 is small, the solution is qualitatively of the form $C_t e^{-\mu_t x^2}$. Because of the term $\chi u(u_{\max} - u)$, the profile u increases everywhere, percentually more where it is small, infinitesimally when it is very close to $u_{\max}$. The combination of these two elements creates, asymptotically, a traveling profile of the form

$$u(t,x) = \phi(x - vt).$$

The profile moves linearly in t, in accordance with the rule discovered in the previous section: if we think to the portion of space where u is very high as a sphere, its radius increases linearly, as in the previous section.

4.3 FKPP as macroscopic limit. The mathematics of proliferation

This section has two aims. Inside this Chapter, devoted to microscopic models and specifically to the in situ phase, it serves as an example of microscopic model giving rise to a PDE which is a first approximate model of in situ tumors, the FKPP equation. More generally, this section presents the elements needed to treat proliferation, a building block of all models of tumor growth, in all regimes and degree of complexity.

The result we are going to present is a particular case of [13]; see also [21]. We present some detail in a different manner.

4.3.1 Time-inhomogeneous Poisson processes

Let $\lambda(t)$ be a non-negative locally integrable function. A non-homogeneous Poisson process N_t with intensity $\lambda(t)$ on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathcal{G}_t$ is an adapted process, null at zero, taking values in the non negative integer numbers, with càdlàg trajectories, with independent increments and $N_t - N_s$ distributed as $\mathcal{P}\left(\int_s^t \lambda(u)\,du\right)$, namely

$$P(N_t - N_s = k) = e^{-\Lambda(s,t)} \frac{\Lambda(s,t)^k}{k!}$$

having set $\Lambda(t) = \int_0^t \lambda(u)\,du$, $\Lambda(s,t) = \Lambda(t) - \Lambda(s)$. When $\lambda(t) = 1$, we say it is a standard Poisson process. If N_t^0 is a standard Poisson process w.r.t. a filtration $\mathcal{F}_t$, then

$$N_t := N^0_{\Lambda(t)}$$

is a non-homogeneous Poisson process N_t with intensity $\lambda(t)$, w.r.t. the filtration $\mathcal{G}_t = \mathcal{F}_{\Lambda(t)}$. Indeed it has the same properties of adaptedness and independence, and of trajectories, and

$$N_t - N_s = N^0_{\Lambda(t)} - N^0_{\Lambda(s)} \sim \mathcal{P}(\Lambda(s,t)).$$

The process $N_t - \Lambda(t)$ is a martingale w.r.t. $\mathcal{G}_t$:

$$E[N_t - N_s - \Lambda(s,t)\,|\mathcal{G}_s] = E[N_t - N_s] - \Lambda(s,t) = 0.$$

Because of this, the function $\Lambda(t)$ is called the compensator of N_t.

Under suitable assumptions, the same result remains true when $\lambda(t)$ is a stochastic process. This is a delicate theory which we address to specialized books. It is the concept of random time change. The a priori exceptional feature of this theory is that $\lambda(t)$ may depend on N^0 itself.

Mix of SDE and Poisson jumps

SDEs driven by jump processes (instead of Brownian motion) is a delicate and specialized subject and presumably it is possible to restate what we are going to do in those terms. But here, precisely in the next section, we need only a special combination of SDEs and jumps. To prepare it, let us consider a model problem.

Assume to have two SDEs

$$\begin{aligned} dX_t^1 &= b_1\left(t, X_t^1\right) dt + dW_t^1 \\ dX_t^2 &= b_2\left(t, X_t^2\right) dt + dW_t^2 \end{aligned}$$

and to have a standard Poisson process $N_0(t)$, with the three processes $W_t^1, W_t^2, N_0(t)$ that are independent. Assume that we start with an $\mathcal{F}_0$-measurable initial condition X_0 by solving the equation for X_t^1. And that, at a random time T which we are going to describe, we "jump" to the equation for X_t^2, namely we start the evolution of such equation at time T from the initial condition X_T^1 (the precise technical meaning of starting an SDE at a random time requires some attention but it is intuitive, it can be done and thus it is left to the interested reader). Overall, we consider the process X_t defined by X_t^1 on the random interval $[0,T)$ and by X_t^2 on $[T,\infty)$.

If T is a priori given, stopping time w.r.t. $\mathcal{F}_t$, the previous description works. But we are interested in the case when T depends on X_t itself. This makes things less obvious.

We assume that the time T is the first jump time of the non-homogeneous Poisson process N_t with intensity $\lambda(t)$, explicitly given by a deterministic function $g(t,x) > 0$ (say continuous, to simplify) composed with the solution X_t:

$$\lambda(t) = g(t, X_t).$$

This is a *random intensity function.* We then consider the process N_t defined as

$$N_t = N_0\left(\int_0^t g(s, X_s)\, ds\right).$$

The first jump of this process is the random time T we want to define.

It can be described also without explicit use of Poisson processes. Let T_0 be an exponential r.v. of rate 1 independent of W_t^1, W_t^2, X_0; it can be the first random jumps of $N_0(t)$. Define T by the equation

$$\int_0^T g(s, X_s)\, ds = T_0.$$

This is a stopping time and it is precisely the first jump time of the process N_t above.

There is a recursion in the previous definition of the dynamics: we need to know the final process X_t, which includes informations about T, in the definition of the intensity $\lambda(t)$, and thus in the definition of T, which is needed to define X_t. This recursion must be solve by a fixed point argument. We suggest the reader to think about this delicate issue and accept it is solvable, with a little degree of intuition. One possibility to understand the solution is to consider a retarded approximation scheme, which mimics a discrete-time scheme (indeed in the discrete-time case we could easily avoid the implicit recursion, making only an explicit recursion).

4.3.2 Proliferating Brownian particles modulated by the density

In Chapter 2 we have seen a rigorous definition of a family of proliferating Brownian particles. However, there the rate of proliferation was constant and equal for all particles. The consequence was the exponential increase of the number of particles. Here we want to modulate proliferation in a way that it is smaller when the density is closer to the threshold 1.

We generically write a for a multi-index $(k, i_1, ..., i_n)$ with $i_1, ..., i_n \in \{1, 2\}$ and $k = 1, ..., N$, where $N \in \mathbb{N}$ is given. We include $a = (k)$ in the class of multi-indices. Denote by Λ the set of all multi-indices. If $a = (k, i_1, ..., i_{n-1}, i_n)$, we write (a, j) for $(k, i_1, ..., i_n, j)$ and $(a, -)$ for $(i_0, i_1, ..., i_{n-1})$.

We are going to describe a dynamic of particles denoted by X_t^a, $a \in \Lambda$, not all alive at time t (to formalize this concept it is sufficient to introduce a point δ outside $\mathbb{R}^d$ and say that the particle is not alive when $X_t^a = \delta$). The life period of particle X_t^a is the interval

$$I^a := [T_0^a, T_1^a)$$

with the stopping times T_0^a, T_1^a have to be specified. Denote by Λ_t the set of all $a \in \Lambda$ such that X_t^a is alive at time t (namely $t \in I^a$) and set

$$S_t^N = \frac{1}{N} \sum_{a \in \Lambda_t} \delta_{X_t^a} = \frac{1}{N} \sum_{a \in \Lambda} \delta_{X_t^a} 1_{t \in I^a}$$

which is the empirical measure at time t.

Let $N_t^{a,0}$, B_t^a, $a \in \Lambda$, X_0^k, $k \in \mathbb{N}$, be independent processes (adapted to a filtration $\mathcal{F}_t$) and r.v.'s ($\mathcal{F}_0$-measurable), where $N_t^{a,0}$ are standard Poisson processes, B_t^i Brownian motions, X_0^i r.v.'s with the same given pdf u_0. When T_0^a and $X_{T_0^a}^a$ will be defined for all a, we set

$$X_t^a = X_{T_0^a}^a + B_t^a - B_{T_0^a}^a, \qquad t \geq T_0^a$$

namely we solve the SDE

$$dX_t^a = dB_t^a$$

on $[T_0^a, \infty)$ with initial condition $X_{T_0^a}^a$. It remains to specify $X_{T_0^a}^a$, T_0^a, T_1^a. We set

$$T_0^a = T_1^{(a,-)}, \qquad X_{T_0^a}^a = X_{T_1^{(a,-)}}^{(a,-)}.$$

It remains to specify the family T_1^a, $a \in \Lambda$, and the first step of the "iteration", namely $T_0^{(k)}$ and $X_{T_0^{(k)}}^{(k)}$. The latter is simply

$$T_0^{(k)} = 0, \qquad X_0^{(k)} = X_0^k.$$

On the contrary, the definition of T_1^a, $a \in \Lambda$, is the difficult part of the story (until now it was just book-keeping).

Let θ_ϵ be a family of classical mollifiers and $\epsilon_N \to 0$ to be specified later on. Set

$$g_t^N(x) = \left(\theta_{\epsilon_N} * S_t^N\right)(x) = \int \theta_{\epsilon_N}(x-y)\, S_t^N(dy).$$

Then introduce the intensities

$$\begin{aligned} \lambda^a(t) &= \left[1 - g_t^N(X_t^a)\right]^+, \qquad t \geq T_0^a \\ \Lambda^a(t) &= \int_{T_0^a}^t \lambda^a(s)\, ds, \qquad t \geq T_0^a. \end{aligned}$$

Finally, let us introduce the process

$$N_t^a = N_{\Lambda^a(t)}^{a,0}, \qquad t \geq T_0^a$$

and denote by T_1^a the first jump time of N_t^a.

We omit the proof that this set of self-referring definitions is well given (the intensities, necessary to define the stopping times, are based on the empirical process which is based on the stopping times). One possibility is to discretize time and avoid implicit references in the recursion; another is to delay some times; see the previous section.

Instead of using a full process N_t^a, it is equivalent to have just the family $T^{a,0}$ of the first random jumps of $N_t^{a,0}$, which are independent exponential r.v. of rate 1 and define T_1^a by the equation

$$\Lambda^a\left(T_1^a\right) = T^{a,0}.$$

The process X_t^a is defined on $[T_0^a, \infty)$, but it will be considered alive only on $[T_0^a, T_1^a)$. The intuition is that the particle a at position $X_{T_1^a}^a$ is replaced by two particles which move like

$$X_t^{(a,j)} := X_{T_1^a}^a + B_t^{(a,j)} - B_{T_1^a}^{(a,j)}, \qquad t \geq T_1^a, \qquad j = 1, 2.$$

Then process $X_t^{(a,j)}$ is stopped at time $T_1^{(a,j)}$, when it splits in two new independent processes, and so on.

The proof of macroscopic limit is made in a few steps:

1. Proof that the family of laws of $S_\cdot^N$ is tight.
2. Proof that, if $S_t^{N_k}$ converges weakly to μ_t, then μ_t is a weak solution of the PDE.
3. Proof that μ_t is the unique solution and has a density $u(t,x)$, so that the full sequence S_t^N converges to $u(t,x)\,dx$.

Let us only sketch point 2, since the other two ones are technical; but point 2 reveals why the particle system should converge to the right PDE, if it converges (namely if 1 holds true).

Let $\varphi \in C_c^2\left(\mathbb{R}^d\right)$ be given. With a little interpretation of $d\varphi\left(X_t^a\right)$ as increment on infinitesimal interval (the problem is with jumps) we can write

$$d\varphi\left(X_t^a\right) = 1_{t\in I_a}\nabla\varphi\left(X_t^a\right)\cdot dB_t^a + 1_{t\in I_a}\Delta\varphi\left(X_t^a\right)dt + 1_{t=T_0^a}\varphi\left(X_{T_0^a}^a\right) - 1_{t=T_1^a}\varphi\left(X_{T_1^a}^a\right)$$

namely

$$\begin{aligned}\varphi\left(X_t^a\right) - \varphi\left(X_0^a\right) &= \int_0^t 1_{s\in I_a}\nabla\varphi\left(X_s^a\right)\cdot dB_s^a + \int_0^t 1_{s\in I_a}\Delta\varphi\left(X_s^a\right)ds \\ &\quad + 1_{T_0^a\in[0,t]}\varphi\left(X_{T_0^a}^a\right) - 1_{T_1^a\in[0,t]}\varphi\left(X_{T_1^a}^a\right)\end{aligned}$$

Hence

$$\begin{aligned}\left\langle S_t^N, \varphi\right\rangle - \left\langle S_0^N, \varphi\right\rangle &= \frac{1}{N}\sum_{a\in\Lambda}\int_0^t 1_{s\in I_a}\nabla\varphi\left(X_s^a\right)\cdot dB_s^a + \frac{1}{N}\sum_{a\in\Lambda}\int_0^t 1_{s\in I_a}\Delta\varphi\left(X_s^a\right)ds \\ &\quad + \frac{1}{N}\sum_{a\in\Lambda} 1_{T_0^a\in[0,t]}\varphi\left(X_{T_0^a}^a\right) - \frac{1}{N}\sum_{a\in\Lambda} 1_{T_1^a\in[0,t]}\varphi\left(X_{T_1^a}^a\right)\end{aligned}$$

namely

$$\left\langle S_t^N, \varphi \right\rangle - \left\langle S_0^N, \varphi \right\rangle = M_t^{1,N} + \int_0^t \left\langle S_s^N, \Delta\varphi \right\rangle ds + \frac{1}{N} \sum_{a \in \Lambda} 1_{T_1^a \in [0,t]} \varphi \left(X_{T_1^a}^a \right)$$

where $M_t^{1,N} = \frac{1}{N} \sum_{a \in \Lambda} \int_0^t 1_{s \in I_a} \nabla\varphi \left(X_s^a \right) \cdot dB_s^a$ satisfies

$$E \left[\left| M_t^{1,N} \right|^2 \right] = \frac{1}{N^2} \sum_{a \in \Lambda} E \int_0^t 1_{s \in I_a} \left| \nabla\varphi \left(X_s^a \right) \right|^2 ds = \frac{1}{N} \int_0^t E \left[\left\langle S_s^N, |\nabla\varphi|^2 \right\rangle \right] ds \leq \frac{T \left\| \nabla\varphi \right\|_0^2}{N} \to 0$$

as $N \to \infty$. We have simplified

$$\frac{1}{N} \sum_{a \in \Lambda} 1_{T_0^a \in [0,t]} \varphi \left(X_{T_0^a}^a \right) - \frac{1}{N} \sum_{a \in \Lambda} 1_{T_1^a \in [0,t]} \varphi \left(X_{T_1^a}^a \right) = \frac{1}{N} \sum_{a \in \Lambda} 1_{T_1^a \in [0,t]} \varphi \left(X_{T_1^a}^a \right).$$

because two particles appear in the same position whenever a particle disappears. The key result is now:

Theorem 30 *The process*

$$M_t^{2,N} = \frac{1}{N} \sum_{a \in \Lambda} 1_{T_1^a \in [0,t]} \varphi \left(X_{T_1^a}^a \right) - \frac{1}{N} \sum_{a \in \Lambda} \int_0^t \varphi \left(X_s^a \right) 1_{s \in I^a} \lambda_s^a ds$$

is a martingale, and

$$\lim_{N \to \infty} E \left[\left| M_t^{2,N} \right|^2 \right] = 0.$$

Due to this result we rewrite the equation above in the form

$$\left\langle S_t^N, \varphi \right\rangle - \left\langle S_0^N, \varphi \right\rangle = M_t^{1,N} + M_t^{2,N} + \int_0^t \left\langle S_s^N, \Delta\varphi \right\rangle ds + \int_0^t \left\langle S_s^N, \varphi \left[1 - g_s^N \right]^+ \right\rangle ds$$

because

$$\frac{1}{N} \sum_{a \in \Lambda} \int_0^t \varphi \left(X_s^a \right) 1_{s \in I^a} \lambda_s^a ds = \int_0^t \left\langle S_s^N, \varphi \left[1 - g_s^N \right]^+ \right\rangle ds.$$

The theorem heuristically gives us

$$\left\langle \rho_t, \varphi \right\rangle - \left\langle \rho_0, \varphi \right\rangle = \int_0^t \left\langle \rho_s, \Delta\varphi \right\rangle ds + \int_0^t \left\langle \rho_s \left(1 - \rho_s \right)^+, \varphi \right\rangle ds$$

in the limit. For this equation one can prove preservation of the domain $[0,1]$, and thus $(1 - \rho_s)^+ = (1 - \rho_s)$. This is the weak form of FKPP equation.

Concerning the proof of Theorem 30, it requires a good part of the theory of Poisson processes and the corresponding stochastic integration. Let us give only some hint. Consider for simplicity the case $a = (k)$. Take the standard Poisson process $N_t^{a,0}$ and time-change it by $\Lambda^a(t)$:

$$\begin{aligned}
\lambda^a(t) &= \left[1 - g_t^N(X_t^a)\right]^+, \qquad t \geq 0 \\
\Lambda^a(t) &= \int_0^t \lambda^a(s)\, ds, \qquad t \geq 0 \\
N_t^a &= N_{\Lambda^a(t)}^{a,0}, \qquad t \geq 0.
\end{aligned}$$

The process $N_t^a - \Lambda^a(t)$ is a martingale, w.r.t. a suitable filtration and the stopped process $N_{t\wedge T_1^a}^a - \Lambda^a(t \wedge T_1^a)$ is also a martingale. But

$$N_{t\wedge T_1^a}^a = 1_{T_1^a \in [0,t]}$$

thus $\Lambda^a(t \wedge T_1^a)$ is the compensator of the random time $1_{T_1^a \in [0,t]}$. We may interpret $1_{T_1^a \in [0,t]} \varphi\left(X_{T_1^a}^a\right)$ as

$$1_{T_1^a \in [0,t]} \varphi\left(X_{T_1^a}^a\right) = \int_0^t \varphi(X_s^a)\, 1_{s \in I^a} dN_s^a.$$

The process

$$\int_0^t \varphi(X_s^a)\, 1_{s \in I^a} d\left(N_s^a - \Lambda^a(s)\right)$$

is again a martingale and its quadratic variation is equal to

$$\int_0^t \varphi^2(X_s^a)\, 1_{s \in I^a} dN_s^a$$

so

$$E\left[\left|\int_0^t \varphi(X_s^a)\, 1_{s \in I^a} d\left(N_s^a - \Lambda^a(s)\right)\right|^2\right] = E\left[\int_0^t \varphi^2(X_s^a)\, 1_{s \in I^a} dN_s^a\right] = E\left[\int_0^t \varphi^2(X_s^a)\, 1_{s \in I^a} \lambda^a(s)\, ds\right].$$

One can also show that the joint quadratic variation between different a's is zero since jumps cannot occur at the same time. It follows

$$\begin{aligned}
E\left[\left|M_t^{2,N}\right|^2\right] &= \frac{1}{N^2} \sum_{a \in \Lambda} E\left[\left|\int_0^t \varphi(X_s^a)\, 1_{s \in I^a} d\left(N_s^a - \Lambda^a(s)\right)\right|^2\right] \\
&= \frac{1}{N^2} \sum_{a \in \Lambda} E\left[\int_0^t \varphi^2(X_s^a)\, 1_{s \in I^a} \lambda^a(s)\, ds\right] \\
&= \frac{1}{N} \int_0^t \left\langle S_s^N, \varphi^2 \left[1 - g_s^N\right]^+ \right\rangle ds \\
&\leq \frac{1}{N} \|\varphi\|_0^2 \to 0.
\end{aligned}$$

Chapter 5

Different interactions and macroscopic limits

5.1 Preliminaries and summary

5.1.1 Number of particles in a volume.

Our interacting system will always have either N particles, in the case without proliferation, or a number which is larger than N but still comparable, at least on short times.

This family of particles, up to a few of them, will occupy a finite portion of space, which to some extent we may idealize to be a compact set. Up to finite multiples, it is like the cube of size one (in length).

If we take a cube with side of length ϵ, or a ball of diameter ϵ, its volume is proportional to ϵ^d.

If the particles are fairly distributed, still with differences in local concentration but of finite size, there are

$$\epsilon^d N \text{ particles in a cube of diameter } \epsilon$$

up to constants.

The global empirical measure at time t is

$$S_t^N = \frac{1}{N} \sum_{\mathbb{R}^d} \delta_{X_t^i}$$

where N is the initial number of particles and the symbol $\sum_{\mathbb{R}^d}$ means the sum extended to all particles $X_t^i \in \mathbb{R}^d$, all the particles which are alive at time t.

If we take a set $\mathcal{U}_\epsilon$, cube or ball or similar, of linear size ϵ, the empirical measure in $\mathcal{U}_\epsilon$ at time t will be defined as

$$S_t^{N,\mathcal{U}_\epsilon} = \frac{1}{N\epsilon^d} \sum_{\mathcal{U}_\epsilon} \delta_{X_t^i}$$

where the symbol $\sum_{\mathcal{U}_\epsilon}$ means the sum extended to all particles $X_t^i \in \mathcal{U}_\epsilon$, all the particles which are alive and belong to $\mathcal{U}_\epsilon$ at time t.

5.1.2 Our model

Both volume constraint and cell adhesion work between cells in contact, hence points at a distance of the order $\frac{1}{N^{1/d}}$. Hence we deal with a system

$$dX_t^i = \sum_j K_N \left(X_t^i - X_t^j \right) dt + \sigma dW_t^i \tag{5.1}$$

where

$$K_N(x) = \frac{1}{\alpha_N} K \left(N^{1/d} x \right), \qquad K \text{ compact support.}$$

The question is: which α_N (and which K)? The shape of K can be decided on the basis of the two interactions. But the intensity, α_N, how large should it be? More precisely, when N increases to infinity, how large should α_N be? Notice that the sum $\sum_j$ has a finite number of addends.

Deducing from the physics the precise asymptotics of α_N is obscure to me. For sure, $\frac{1}{\alpha_N}$ must be very large, compared to σ, since the speed of the local displacement produced by volume constraint and adhesion is enormous, on the scale of seconds, compared to other sources of motions, which are on the scale of days.

Thus, in absence of a more strict motivation, our choice is that we are in the case of short range interactions with $\frac{1}{\alpha_N} = N$.

5.2 Different interactions

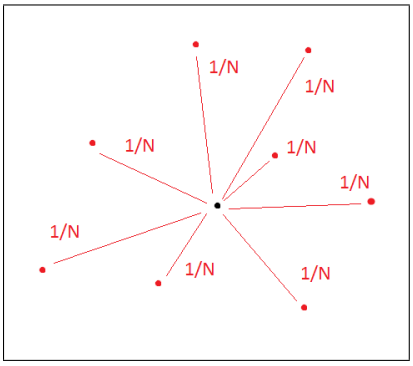
Different classes of interactions, for systems of type (5.1), are considered in the literature. A general source of inspiration is [20], although mainly devoted either to deterministic ODEs or cellular automata. More specific references are given below.

5.2.1 Long range (mean field)

It is the case

$$K_N(x) = \frac{1}{N} K(x)$$

without special restrictions on K (just bounded and Lipschitz). It means that each particle interacts with each other, with a weak intensity $\sim \frac{1}{N}$, hence each particle feels an average of inputs; it is not surprising that in the limit each particle feels an average input ρ, the so called mean field of the system.



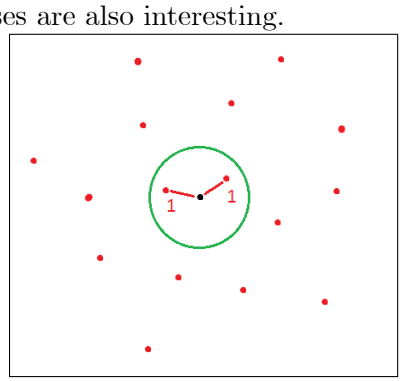
5.2.2 Short range

It is the case (like contact interaction):

$$K_N(x) = a_N K\left(N^{1/d}x\right), \qquad K \text{ compact support.}$$

It means that particles interact only if they are at distance $\frac{1}{N^{1/d}}$ each other, which is the typical distance between nearest neighbor particles in a set of N particles in a finite volume (when no accumulation occurs), hence only nearest neighbor particles (or close to it) interact. The amplitude a_N is not prescribed a priori, we may have in mind different regimes. If we set $\epsilon_N = \frac{1}{N^{1/d}}$, we may think to the ball (or cube) $\mathcal{U}_{\epsilon_N}$ or diameter ϵ_N centered at X_t^i and we have that X_t^i interacts with particles in $\mathcal{U}_{\epsilon_N}$ only, which are a finite number.

In the figure we have denoted by "1" the strength of each single interaction but other cases are also interesting.

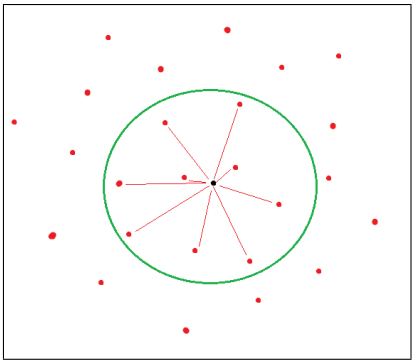


5.2.3 Intermediate range

Intermediate between long and short range is the case (called *moderate interactions* in part of the literature)

$$K_N(x) = a_N K\left(N^{\beta/d}x\right), \qquad \beta \in (0,1), \qquad K \text{ compact support.}$$

It is an intermediate regime between the previous two ones. If we set $\epsilon_N = \frac{1}{N^{\beta/d}}$, we may think to the ball (or cube) $\mathcal{U}_{\epsilon_N}$ or diameter ϵ_N centered at X_t^i and we have that X_t^i interacts with particles in $\mathcal{U}_{\epsilon_N}$ only; the number of such particles is of the order $\epsilon_N^d N = \frac{N}{N^\beta} = N^{1-\beta}$.



5.2.4 Common notations

To unify notations, we may always write

$$K_N(x) = a_N K\left(N^{\beta/d} x\right)$$

and distinguish:

1. long range: $\beta = 0$, $a_N = \frac{1}{N}$
2. short range: $\beta = 1$, K compact support
3. intermediate range: $\beta \in (0,1)$, K compact support.

5.2.5 Mean field amplitude

The literature focuses on two choices of the amplitude a_N. The first one will be called *mean field amplitude* and is given by

$$a_N = \frac{1}{N} N^\beta = \begin{cases} \frac{1}{N} & \text{long range} \\ \frac{1}{N^{1-\beta}} & \text{intermediate range} \\ 1 & \text{short range.} \end{cases}$$

Since $N^{1-\beta} = N\epsilon_N^d$ is proportional to the number of particles in $\mathcal{U}_{\epsilon_N}$ (when no accumulation occurs), the input on each particle is bounded, like in the long range case; hence tightness is "easy". If $K(x) = \theta(x) v$ where θ is a classical mollifier and v is a given vector, then

$$\begin{aligned} K_N(x) &= a_N \theta\left(N^{\beta/d} x\right) v \\ &= \frac{1}{N} N^\beta \theta\left(N^{\beta/d} x\right) v \end{aligned}$$

looks like a mean field problem with kernel converging to δ_0. However, only in the intermediate range case we still expect a law of large numbers, while in the short range case not, since it is an average over a finite number of values.

5.2.6 Large amplitude

The second typical example of amplitude in the literature, that here will be called with *large amplitude*, is inspired to the choice

$$K(x) = \nabla\theta(x)$$

where θ is a classical mollifier. It is

$$a_N = \frac{1}{N} N^{\beta\left(1+\frac{1}{d}\right)} = \begin{cases} \frac{1}{N} & \text{long range} \\ \frac{1}{N^{1-\beta\left(1+\frac{1}{d}\right)}} & \text{intermediate range} \\ N^{1/d} & \text{short range} \end{cases}$$

(hence in the long range case it reduces again to the mean field situation). It is natural because

$$\begin{aligned} K_N(x) &= a_N(\nabla\theta)\left(N^{\beta/d}x\right) \\ &= \frac{1}{N} N^{\beta} N^{\beta/d}(\nabla\theta)\left(N^{\beta/d}x\right) \\ &= \frac{1}{N}\nabla\theta_N(x) \end{aligned}$$

where $\theta_N(x) = N^{\beta}\theta\left(N^{\beta/d}x\right) \to \delta_0(x)$. It looks like a mean field problem with kernel converging to $\nabla\delta_0$. This case is singular: think to the short range case to understand why (the other case is similar). We have

$$\sum_j K_N\left(X_t^i - X_t^j\right) = N^{1/d}\sum_j (\nabla\theta)\left(N^{1/d}\left(X_t^i - X_t^j\right)\right).$$

The sum is typically finite, hence this input to the motion of X_t^i diverges with N, opposite to the mean field cases above. This divergence must be compensated by a special symmetry of K, a symmetry which typically makes $\sum_j K\left(N^{1/d}\left(X_t^i - X_t^j\right)\right)$ very small.

5.3 Different PDE limits

In all cases, given

$$S_t^N = \frac{1}{N}\sum_i \delta_{X_t^i}$$

we have to understand the limit of the identity

$$\begin{aligned}\left\langle S_t^N, \varphi\right\rangle - \left\langle S_0^N, \varphi\right\rangle &= \frac{1}{N}\sum_i \int_0^t \nabla\varphi\left(X_s^i\right)\cdot\sum_j K_N\left(X_s^i - X_s^j\right) ds \\ &\quad + \frac{\sigma^2}{2}\int_0^t \left\langle S_s^N, \Delta\varphi\right\rangle ds + M_t^N\end{aligned}$$

where

$$M_t^N = \sigma\frac{1}{N}\sum_i \int_0^t \nabla\varphi\left(X_s^i\right)\cdot dW_s^i.$$

Assume we already know that $S_t^N \rightharpoonup \rho_t$. We have to identify the equation satisfied by ρ_t.

The only difficult issue is to understand the limit of

$$\frac{1}{N}\sum_i \int_0^t \nabla\varphi\left(X_s^i\right)\cdot\sum_j K_N\left(X_s^i - X_s^j\right) ds.$$

This term is equal to

$$\int_0^t \left\langle S_s^N, \nabla\varphi\sum_j K_N\left(\cdot - X_s^j\right)\right\rangle ds$$

and to

$$N\int_0^t \int\int \nabla\varphi\left(x\right) K_N\left(x-y\right) S_s^N\left(dx\right) S_s^N\left(dy\right) ds.$$

If we use the first formula, the problem is to understand the function

$$x \mapsto \sum_j K_N\left(x - X_s^j\right).$$

5.3.1 Long range, mean field, case

Exercise 31 *If $\mu_N \rightharpoonup \mu$ then $\mu_N \otimes \mu_N \rightharpoonup \mu\otimes\mu$ (use Skorohod representation theorem and two independent copies of the space).*

If

$$K_N\left(x\right) = \frac{1}{N}K\left(x\right)$$

then

$$\begin{aligned}&N\int_0^t \int\int \nabla\varphi\left(x\right) K_N\left(x-y\right) S_s^N\left(dx\right) S_s^N\left(dy\right) ds \\ =\ &\int_0^t \int\int \nabla\varphi\left(x\right) K\left(x-y\right) S_s^N\left(dx\right) S_s^N\left(dy\right) ds\end{aligned}$$

which converges to

$$\int_0^t \int \int \nabla\varphi(x) K(x-y) \rho_s(dx) \rho_s(dy) ds$$

by the exercise and dominated convergence theorem in ds. The result is

$$\int_0^t \langle \rho_s, \nabla\varphi \cdot (\rho_s * K) \rangle ds = -\int_0^t \langle \varphi, \operatorname{div}(\rho_s(\rho_s * K)) \rangle ds.$$

The final PDE is

$$\frac{\partial \rho}{\partial t} = \frac{\sigma^2}{2}\Delta\rho - \operatorname{div}(\rho(\rho * K)).$$

It is a non-local non-linear PDE. A general main reference is [23].

5.3.2 Intermediate range with mean field amplitude

Consider the case

$$K_N(x) = \frac{1}{N} N^\beta K\left(N^{\beta/d} x\right), \qquad K \text{ compact support}$$

We have

$$\begin{aligned} & N \int_0^t \int \int \nabla\varphi(x) K_N(x-y) S_s^N(dx) S_s^N(dy) ds \\ = & \int_0^t \int \int \nabla\varphi(x) N^\beta K\left(N^{\beta/d}(x-y)\right) S_s^N(dx) S_s^N(dy) ds. \end{aligned}$$

Assume for simplicity $K(x) = v\theta(x)$ where θ is a classical mollifier. Hence the natural limit is

$$\begin{aligned} & \int_0^t \int \int \nabla\varphi(x) \cdot v\delta_0(x-y) \rho_s(x) \rho_s(y) dxdyds \\ = & \int_0^t \int \int \nabla\varphi(x) \cdot v\rho_s^2(x) dxds = -\int_0^t \int \int \varphi(x) \operatorname{div}\left(v\rho_s^2(x)\right) dxds. \end{aligned}$$

The final PDE is

$$\frac{\partial \rho}{\partial t} = \frac{\sigma^2}{2}\Delta\rho - \operatorname{div}\left(v\rho^2\right).$$

It is a local semi-linear PDE. A result in this direction can be found in [12].

Is such limit so natural? Let us mention the following simple fact.

Lemma 32 *Let $\mu_n, \mu, \rho_n := \theta_n * \mu_n$. Then $\langle \mu_n, \phi \rangle \to \langle \mu, \phi \rangle$ for all $\phi \in UC_b$ if and only if $\langle \rho_n, \phi \rangle \to \langle \mu, \phi \rangle$ for all $\phi \in UC_b$.*

Proof. $\langle \mu_n, \phi \rangle \to \langle \mu, \phi \rangle$ and $\langle \rho_n, \phi \rangle \to \langle \mu, \phi \rangle$ differ by

$$\langle \mu_n, \phi \rangle - \langle \rho_n, \phi \rangle$$

and thus it is sufficient to show that this difference goes to zero for all $\phi \in UC_b$. We have

$$\begin{aligned} \langle \rho_n, \phi \rangle &= \int \left(\int \theta_n (x-y) \mu_n (dy) \right) \phi (x) dx \\ &= \int \left(\int \theta_n (x-y) \phi (x) dx \right) \mu_n (dy) = \left\langle \mu_n, \widetilde{\theta}_n * \phi \right\rangle \end{aligned}$$

hence we have to show that

$$\left\langle \mu_n, \widetilde{\theta}_n * \phi - \phi \right\rangle \to 0.$$

We have

$$\left| \left\langle \mu_n, \widetilde{\theta}_n * \phi - \phi \right\rangle \right| \leq \left\langle \mu_n, \left| \widetilde{\theta}_n * \phi - \phi \right| \right\rangle \leq \left\| \widetilde{\theta}_n * \phi - \phi \right\|_0$$

and it is a simple exercise to show that $\left\| \widetilde{\theta}_n * \phi - \phi \right\|_0 \to 0$ as $n \to \infty$, when $\phi \in UC_b$. ■

Based on this lemma, $\int N^\beta K \left(N^{\beta/d} (x-y) \right) S_s^N (dy)$ "weakly" converges to $\rho_s (x)$ (by "weakly" we mean against UC_b functions, sufficient if we reformulate the problem on a torus; to reach C_b functions one needs ad hoc control on several objects at infinity). But then we have to compute

$$\left\langle \int N^\beta K \left(N^{\beta/d} (\cdot - y) \right) S_s^N (dy), \nabla \varphi (x) S_s^N \right\rangle$$

so we have a classical problem of convergence of product of two weakly convergent objects. The result is not unique, it depends on details of the sequences.

In the intermediate range case the intuition is that it works, the limit is the product of the weak limits: this is due to the fact that the family of mollifiers $N^\beta K \left(N^{\beta/d} (x-y) \right)$ "observe" a large part of points of S_s^N.

Completely different would be the short range case, when the interaction is $NK \left(N^{1/d} (x-y) \right)$. In this case, if particles do not aggregate too much, only very few, a finite number of particles are "observed" by the mollifiers. Think to the extreme case when the support of θ is too small compared to the typical distance between neighbor particles. The mollifiers do not match any pair of particles, or just very few due to some fluctuation, hence

$$\int_0^t \int \int \nabla \varphi (x) N^1 K \left(N^{1/d} (x-y) \right) S_s^N (dx) S_s^N (dy) ds$$

is zero or almost zero, opposite to the conjectured limit $\int_0^t \int \int \nabla \varphi (x) \cdot v \rho_s^2 (x) dx ds$.

5.3.3 Intermediate range with large amplitude

Consider the case

$$K_N(x) = \frac{1}{N} N^{\beta\left(1+\frac{1}{d}\right)} (\nabla\theta)\left(N^{\beta/d}x\right) = \frac{1}{N}\nabla\theta_N(x), \qquad K \text{ compact support}$$

where $\theta_N(x) = N^{\beta}\theta\left(N^{\beta/d}x\right)$. We have

$$N\int_0^t \int\int \nabla\varphi(x) K_N(x-y) S_s^N(dx) S_s^N(dy)\, ds = \int_0^t \int\int \nabla\varphi(x) \nabla\theta_N((x-y)) S_s^N(dx) S_s^N(dy)\, ds.$$

The intuition is that it converges to

$$\begin{aligned} & \int_0^t \int\int \nabla\varphi(x) \nabla\delta_0(x-y) \rho_s(x) \rho_s(y)\, dxdyds \\ = & -\int_0^t \int \nabla\varphi(x) \rho_s(x) \nabla\rho_s(x)\, dxds \\ = & \int_0^t \int \varphi(x) \operatorname{div}(\rho_s(x) \nabla\rho_s(x))\, dxds. \end{aligned}$$

The final PDE is

$$\frac{\partial\rho}{\partial t} = \frac{\sigma^2}{2}\Delta\rho + \operatorname{div}(\rho\nabla\rho).$$

It is a local quasi-linear PDE. When $\sigma = 0$ it is known as porous media equation and it has the property of maintaining compact support of solutions. The basic reference for us will be [12]. See also [17].

Obviously the same remarks apply here as in the mean field case: only in the true intermediate range case we may expect the result to be true; it cannot be extended to short range.

5.3.4 Short range and other regimes

Under appropriate conditions, the limit PDE has the form

$$\frac{\partial\rho}{\partial t} = \Delta P(\rho).$$

Unfortunately, the function P is not known explicitly; it is the pressure of a certain equilibrium ensemble. The basic reference on this problem is [27]. See also [25] and [14].

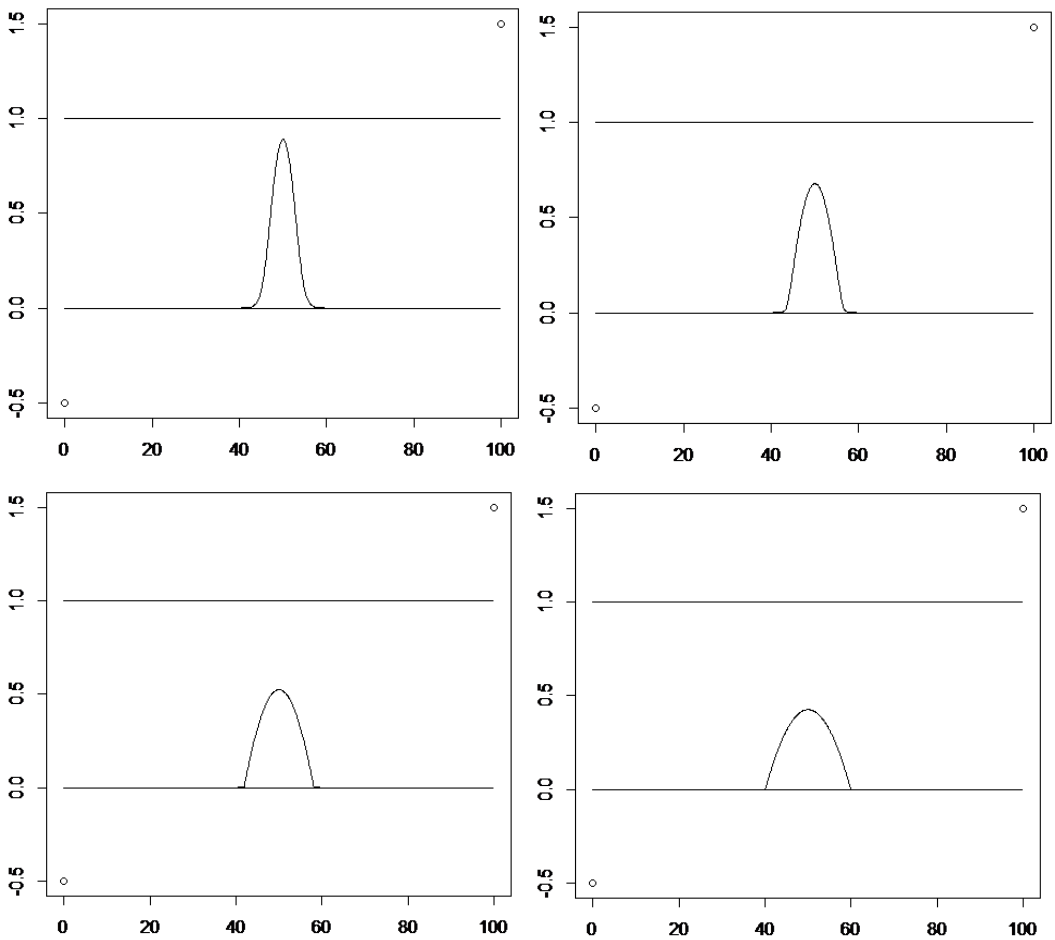
Let us also mention that other regimes, not explicitly discussed here, can be found in [2].

5.4 Simulations

We have seen in Chapter 2 and 3 several simulations of FKPP and other examples. We add here the simulation of two equations with nonlinear diffusion term, of the form found above. The first one is the so called porous media equation,

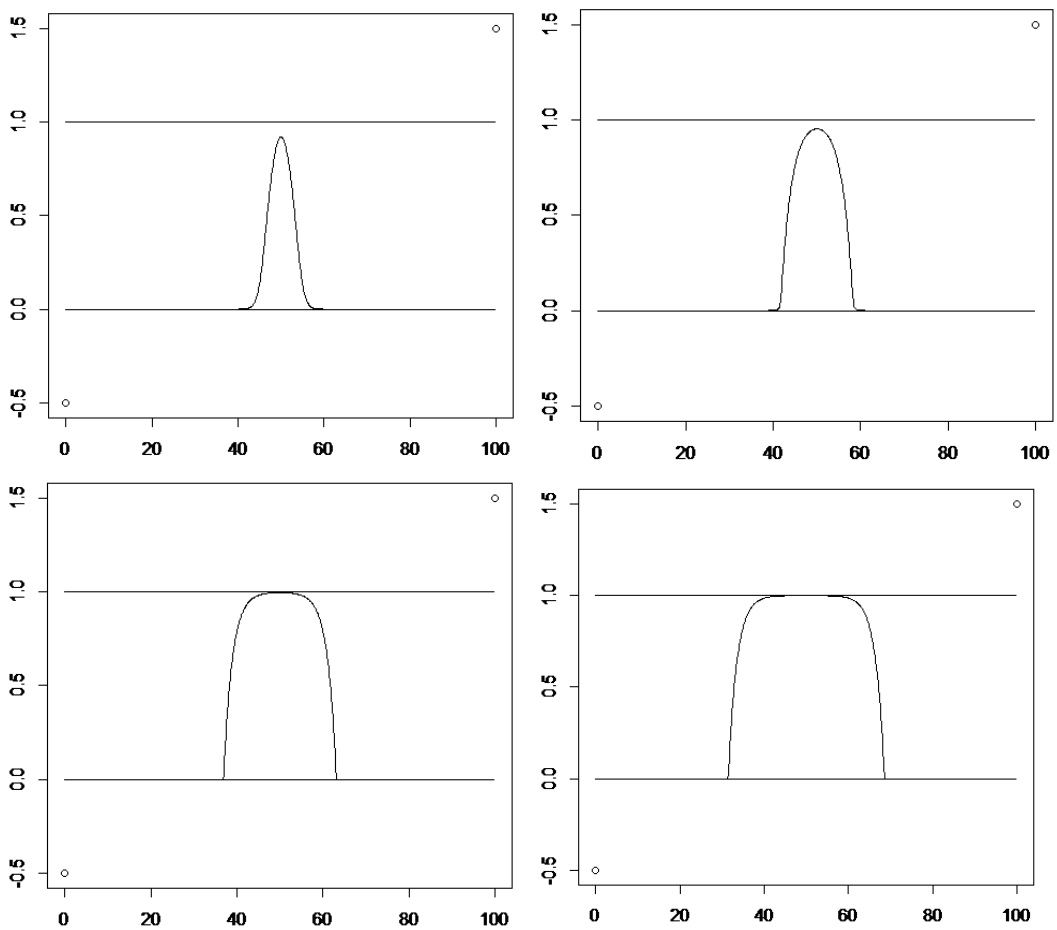
$$\frac{\partial \rho}{\partial t} = \Delta \rho^2$$

(notice that $\Delta \rho^2 = 2 \operatorname{div}(\rho \nabla \rho)$). The solution, which initially is an exponential, tends to become "compact support". It is known, for this equation, that a compact support initial condition remains compact support.



The previous pictures are just nonlinear diffusion. The next ones add proliferation:

$$\frac{\partial \rho}{\partial t} = \Delta \rho^2 + \rho (1 - \rho)$$



This looks very promising to describe accurately the growth of a tumor in situ. We presumably have a sort of traveling wave here too, like for FKPP. Let us sketch the analysis of this problem, with factor $\frac{1}{2}$ for comparison with FKPP.

Consider the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u^2 + u\left(1-u\right).$$

Assuming we have a solution of the form

$$u\left(t,x\right) = w\left(x-ct\right)$$

we find, by substitution

$$-cw' = w'w' + ww'' + w - w^2. \tag{5.2}$$

We impose again

$$\lim_{x\to-\infty} w\left(x\right) = 1, \qquad \lim_{x\to+\infty} w\left(x\right) = 0 \tag{5.3}$$

and that w is decreasing.

Equation (2.3), with the notations $x(t) := w$, $y(t) = w'$, is equivalent to the system

$$\begin{aligned} x' &= y \\ y' &= \frac{-cy - y^2 - x + x^2}{x} = -c\frac{y}{x} - \frac{y^2}{x} - 1 + x. \end{aligned}$$

We look for solutions $(x(t), y(t))$ defined on the whole $\mathbb{R}$ such that

$$\lim_{t\to-\infty} x(t) = 1, \qquad \lim_{t\to+\infty} x(t) = 0.$$

The fixed points are the pairs (x, y) such that

$$\begin{aligned} y &= 0 \\ -cy - y^2 - x + x^2 &= 0 \end{aligned}$$

hence $-x + x^2 = 0$, $x = 0$ or $x = 1$, therefore

$$A = (0, 0), \qquad B = (1, 0).$$

Denoting the vector field by $f(x, y)$, we have

$$Df(x, y) = \begin{pmatrix} 0 & 1 \\ c\frac{y}{x^2} + \frac{y^2}{x^2} + 1 & -c\frac{1}{x} - 2\frac{y}{x} \end{pmatrix} = \frac{1}{x^2}\begin{pmatrix} 0 & x^2 \\ cy + y^2 + x^2 & -cx - 2xy \end{pmatrix}$$

By Maple we have found that $Df(B) = \begin{pmatrix} 0 & 1 \\ 1 & -c \end{pmatrix}$ has eigenvectors

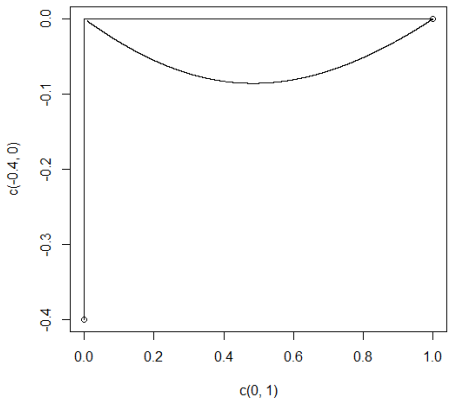
$$e_1^B = \begin{pmatrix} \frac{1}{2}c - \frac{1}{2}\sqrt{c^2+4} \\ 1 \end{pmatrix}, \qquad e_2^B = \begin{pmatrix} \frac{1}{2}c + \frac{1}{2}\sqrt{c^2+4} \\ 1 \end{pmatrix}$$

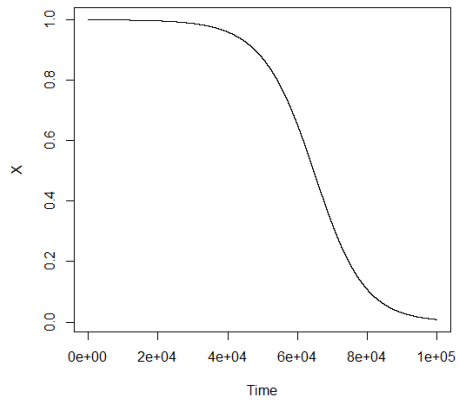
with eigenvalues

$$-\frac{1}{2}c - \frac{1}{2}\sqrt{c^2+4}, \qquad \frac{1}{2}\sqrt{c^2+4} - \frac{1}{2}c.$$

Hence $Df(B)$ has eigenvalues of opposite signs, an hyperbolic point with unstable manifold tangent to e_2^B, pointing in the first and third quadrant. As for FKPP, a solution exists from B along $-e_2^B$. It moves in the forth quadrant initially in south-west direction, then turns towards the origin in the north-west direction.

The analysis of point A is more difficult because $Df(w, z)$ diverges as $(w, z) \to A$, a fact that could be a signature of a faster approach and a compact support profile. Let us see a simulation for $c = 3$. The figure on the left is the heterocline. The figure on the right is the corresponding profile.


c(-0.4, 0)
0.0
-0.1
-0.2
-0.3
-0.4
0.0
0.2
0.4
0.6
0.8
1.0
c(0, 1)


X
1.0
0.8
0.6
0.4
0.2
0.0
0e+00
2e+04
4e+04
6e+04
8e+04
1e+05
Time

Chapter 6

Some mathematical details on the macroscopic limits

6.1 Mean field limit

The next discussion is partially based on [23].

Consider the equation

$$dX_t^i = \frac{1}{N}\sum_j K\left(X_t^i - X_t^j\right)dt + \sigma dW_t^i, \qquad i = 1, ..., N \tag{6.1}$$

with

$$K \text{ bounded Lipschitz continuous.}$$

We shall also write it in the form

$$dX_t^i = b_t^N\left(X_t^i\right)dt + \sigma dW_t^i \tag{6.2}$$

where

$$b_t^N(x) := \frac{1}{N}\sum_j K\left(x - X_t^j\right) = \int K(x-y)\, S_t^N(dy).$$

To simplify, assume that the initial conditions X_0^i, $\mathcal{F}_0$-measurable, are i.i.d. with law $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$ (defined in the next lines). There are two main approaches, described in the next subsections.

As a general preliminary, denote by $\mathrm{Pr}_1\left(\mathbb{R}^d\right)$ the space of probability measures on $\mathbb{R}^d$ with finite first moment, endowed with the 1-Wasserstein metric $\mathcal{W}_1$. Recall that

$$\mathcal{W}_1(\mu,\nu) = \inf E\left[|X - Y|\right]$$

where the infimum is taken over all the pairs (X, Y) of r.v. having laws $\mathcal{L}(X) = \mu$, $\mathcal{L}(Y) = \nu$. And one also has

$$\mathcal{W}_1(\mu, \nu) = \sup \left| \int \phi(x) (\mu - \nu)(dx) \right|$$

over all $\phi : \mathbb{R}^d \to \mathbb{R}$ with $[\phi]_{Lip} \leq 1$, $\|\phi\|_0 \leq 1$. Consider also the space $C\left([0,T]; \mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$ of continuous families $(\mu_t)_{t\in[0,T]}$, with $\mu_t \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$, endowed with the metric

$$d(\mu, \nu) = \sup_{t\in[0,T]} \mathcal{W}_1(\mu_t, \nu_t).$$

A known tightness criterion for the family of laws of a sequence of measure-valued processes $(\mu^n_\cdot)_{n\in\mathbb{N}}$ with paths in $C\left([0,T]; \mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$ (see [8]) requires that:

i) for every $\epsilon > 0$ there is compact set $K_\epsilon \subset \mathrm{Pr}_1\left(\mathbb{R}^d\right)$ such that

$$P(\mu^n_t \in K_\epsilon \text{ for all } t \in [0,T]) > 1 - \epsilon \tag{6.3}$$

for all $n \in \mathbb{N}$; and that

ii)

$$E\left[\mathcal{W}_1(\mu^n_t, \mu^n_s)^p\right] \leq C|t-s|^{1+\alpha} \tag{6.4}$$

for some $p, \alpha > 0$.

6.1.1 Compactness approach

The first one could be summarized under the sentence "by compactness". It consists in three steps. First one has to prove:

Theorem 33 *The family of laws of the empirical measure processes*

$$S^N_t := \frac{1}{N} \sum_j \delta_{X^j_t}$$

is tight in $C\left([0,T]; \mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$.

Proof. We do not discuss the proof of (6.3) which is a little technical; the reader who wants to have a complete result could restrict the framework to diffusions on the torus (as in [27]). Let us prove (6.4), by the following elegant computation: for every $\phi : \mathbb{R}^d \to \mathbb{R}$ with $[\phi]_{Lip} \leq 1$, $\|\phi\|_0 \leq 1$,

$$\mathcal{W}_1\left(S^N_t, S^N_s\right) \leq \left| \int \phi(x) \left(S^N_t - S^N_s\right)(dx) \right|$$

$$\begin{aligned}
&\leq \frac{1}{N}\sum_i \left|\phi\left(X_t^i\right)-\phi\left(X_s^i\right)\right| \leq \frac{1}{N}\sum_i \left|X_t^i - X_s^i\right| \\
&\leq \frac{1}{N}\sum_i \int_s^t \left|b_r^N\left(X_r^i\right)\right| dr + \sigma\frac{1}{N}\sum_i \left|W_t^i - W_s^i\right| \\
&\leq \|K\|_0 |t-s| + \sigma\frac{1}{N}\sum_i \left|W_t^i - W_s^i\right|
\end{aligned}$$

from (6.2) and the boundedness of K. Moreover, $\left(\frac{1}{N}\sum\left|W_t^i - W_s^i\right|\right)^4 \leq \frac{1}{N}\sum\left|W_t^i - W_s^i\right|^4$, hence

$$E\left[\left(\frac{1}{N}\sum_i \left|W_t^i - W_s^i\right|\right)^4\right] \leq 3\left(t-s\right)^2.$$

Thus we get

$$E\left[\mathcal{W}_1\left(S_t^N, S_s^N\right)^4\right] \leq 4\|K\|_0^4 |t-s|^4 + 12\sigma^4 |t-s|^2$$

which implies (6.4). ■

The second step consists in proving that all weak limit points of S_t^N satisfy the PDE. Let us take a weakly converging subsequence, still denoted by S_t^N for notational simplicity (at the end the question of uniqueness of the limit is important). Denote by μ_t its limit, which a priori is a measure-valued stochastic process. We have to identify the equation satisfied, in weak sense, by μ_t. We have already outlined this step in Chapter 5: from Itô formula, for every $\phi \in C_c^2\left(\mathbb{R}^d\right)$ we have

$$\begin{aligned}
d\left\langle S_t^N, \phi\right\rangle &= \frac{1}{N}\sum_i d\phi\left(X_t^i\right) \\
&= \frac{1}{N}\sum_i \nabla\phi\left(X_t^i\right)\left\langle S_t^N, K\left(X_t^i - \cdot\right)\right\rangle dt + \frac{1}{N}\sum_i \nabla\phi\left(X_t^i\right)\sigma dW_t^i + \frac{1}{N}\sum_i \frac{\sigma^2}{2}\Delta\phi\left(X_t^i\right) dt.
\end{aligned}$$

Think to the integrated version in t: on the left we have

$$\left\langle S_t^N, \phi\right\rangle - \left\langle S_0^N, \phi\right\rangle$$

which converges to

$$\left\langle \mu_t, \phi\right\rangle - \left\langle \mu_0, \phi\right\rangle.$$

Similarly the term

$$\int_0^t \frac{1}{N}\sum_i \frac{\sigma^2}{2}\Delta\phi\left(X_s^i\right) ds = \int_0^t \frac{\sigma^2}{2}\left\langle S_s^N, \Delta\phi\right\rangle ds$$

converges to

$$\int_0^t \frac{\sigma^2}{2}\left\langle \mu_s, \Delta\phi\right\rangle ds.$$

The term $\frac{1}{N}\sum_i \int_0^t \nabla\phi\left(X_s^i\right)\sigma dW^i(s)$ converges to zero in mean square:

$$E\left[\left|\frac{1}{N}\sum_i \int_0^t \nabla\phi\left(X_s^i\right)\sigma dW^i(s)\right|^2\right] = \frac{\sigma^2}{N^2}\sum_i \int_0^t E\left[\left|\nabla\phi\left(X_s^i\right)\right|^2\right] ds$$
$$\leq \frac{\sigma^2}{N^2}\sum_i \|\nabla\phi\|_\infty^2 t = \frac{\sigma^2}{N}\|\nabla\phi\|_\infty^2 t \to 0.$$

Finally,

$$\int_0^t \frac{1}{N}\sum_i \nabla\phi\left(X_s^i\right)\left\langle S_t^N, K\left(X_s^i - \cdot\right)\right\rangle ds = \int_0^t \int\int \nabla\phi(x) K(x-y) S_s^N(dx) S_s^N(dy) ds$$

converges to

$$\int_0^t \int\int \nabla\phi(x) K(x-y) \mu_s(dx) \mu_s(dy) ds$$

as explained rigorously in Chapter 5. Thus μ_t satisfies the identity

$$\langle\mu_t,\phi\rangle - \langle\mu_0,\phi\rangle = \int_0^t \int\int \nabla\phi(x) K(x-y) \mu_s(dx) \mu_s(dy) ds + \int_0^t \frac{\sigma^2}{2}\langle\mu_s,\Delta\phi\rangle ds \quad (6.5)$$

which is the weak form of

$$\frac{\partial\mu_t}{\partial t} = \frac{\sigma^2}{2}\Delta\mu_t - \operatorname{div}\left(\mu_t \int K(x-y)\mu_t(dy)\right). \quad (6.6)$$

The third and last step requires to prove that this PDE has a unique weak solution. This can be done as in Theorem 37 below. Since S_0^N convergence weakly, a.s., to the deterministic measure μ_0, it follows that μ_t is deterministic. Moreover, it is the same for every converging subsequence of S_t^N, hence the full sequence S_t^N converges to μ_t. We have proved:

Theorem 34 *S_t^N converges, in the topology of $C\left([0,T];\mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$, to the unique deterministic measure-valued solution $(\mu_t)_{t\in[0,T]}$ of equation (6.6) with initial condition $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$.*

6.1.2 By comparison with the mean field SDE

The second approach (following e.g. [23]) is potentially less general than the compactness one but it gives more informations. It starts with a preliminary analysis of the limit SDE

$$dX_t = b_t(X_t)dt + \sigma dW_t \quad (6.7)$$

$$b_t(x) = \int K(x-y)\,\mu_t(dy) \tag{6.8}$$

$$\mu_t = \text{Law of } X_t \tag{6.9}$$

with X_0 a given $\mathcal{F}_0$-measurable r.v.; assume for simplicity $E[|X_0|] < \infty$. One can prove that:

Theorem 35 *There is a unique strong solution (X, μ) of problem (6.7)-(6.9).*

Proof. We give only the scheme of the proof (the details are not difficult). Denote by $\mathrm{Pr}_1(\mathbb{R}^d)$ the space of probability measures on $\mathbb{R}^d$ with finite first moment, endowed with the 1-Wasserstein metric $\mathcal{W}_1$. Consider the space $C([0,T]; \mathrm{Pr}_1(\mathbb{R}^d))$ of continuous families $(\mu_t)_{t\in[0,T]}$, with $\mu_t \in \mathrm{Pr}_1(\mathbb{R}^d)$, endowed with the metric

$$d(\mu,\nu) = \sup_{t\in[0,T]} \mathcal{W}_1(\mu_t, \nu_t).$$

First, one consider the map Γ in $C([0,T']; \mathrm{Pr}_1(\mathbb{R}^d))$, T' to be chosen, defined as

$$\begin{aligned}
(\Gamma\mu)_t &= Law(Y_t^\mu) \\
dY_t^\mu &= b_t^\mu(Y_t^\mu)\,dt + \sigma dW_t, \qquad Y_0^\mu = X_0 \\
b_t^\mu(x) &= \int K(x-y)\,\mu_t(dy).
\end{aligned}$$

It is a contraction (for small T'). Let us show only this computation. By definition/characterization of 1-Wasserstein metric, we have

$$\mathcal{W}_1((\Gamma\mu)_t, (\Gamma\nu)_t) \le E[|Y_t^\mu - Y_t^\nu|].$$

By the equation,

$$\begin{aligned}
|Y_t^\mu - Y_t^\nu| &\le \int_0^t |b_s^\mu(Y_s^\mu) - b_s^\nu(Y_s^\nu)|\,ds \\
&\le \int_0^t |b_s^\mu(Y_s^\mu) - b_s^\mu(Y_s^\nu)|\,ds + \int_0^t |b_s^\mu(Y_s^\nu) - b_s^\nu(Y_s^\nu)|\,ds \\
&\le L_K \int_0^t |Y_s^\mu - Y_s^\nu|\,ds + L_K \int_0^t \mathcal{W}_1(\mu_s, \nu_s)\,ds
\end{aligned}$$

because

$$\left|b_t^\mu(x) - b_t^\mu(x')\right| \le \int \left|K(x-y) - K(x'-y)\right| \mu_t(dy) \le L_K\left|x - x'\right|$$

$$|b_t^\mu(x) - b_t^\nu(x)| = L_K \left|\int L_K^{-1} K(x-y)(\mu_t - \nu_t)(dy)\right| \le L_K \mathcal{W}_1(\mu_t, \nu_t)$$

by the other definition/characterization of 1-Wasserstein metric. Thus, by Gronwall lemma,

$$|Y_t^\mu - Y_t^\nu| \leq e^{L_K} L_K \int_0^t \mathcal{W}_1(\mu_s, \nu_s)\, ds.$$

We conclude

$$\mathcal{W}_1((\Gamma\mu)_t, (\Gamma\nu)_t) \leq e^{L_K} L_K \int_0^{T'} \mathcal{W}_1(\mu_s, \nu_s)\, ds.$$

Then Γ is a contraction in $C\left([0,T'];\mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$ for T' small enough, namely such that $e^{L_K} L_K T' < 1$.

This gives us local existence and uniqueness of μ, fixed point of Γ, from which one deduces existence and uniqueness of a strong solution on $[0,T']$ by a classical result on SDE's and identification of μ as the law of this solution, due to the definition of Γ. The argument can then be repeated on intervals of equal length, because the choice of T' depends only on L_K. ■

Definition 36 *A measure-valued solution $(\mu_t)_{t\in[0,T]}$ of equation (6.6) with initial condition $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$ is a family in $C\left([0,T];\mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$ such that (6.5) holds for every $\phi \in C_c^2\left(\mathbb{R}^d\right)$.*

Theorem 37 *Given $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$, there exists a unique measure-valued solution $(\mu_t)_{t\in[0,T]}$ of equation (6.6) with initial condition $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$.*

Proof. We give only the scheme of the proof. Existence readily follows from the previous theorem: it is sufficient to apply the result on Fokker-Planck equations given in Chapter 1. The proof of uniqueness is more involved.

First (this is the more technical step) one can prove uniqueness of measure-valued solutions $(\mu_t)_{t\in[0,T]}$ of the linear (Fokker-Planck) equation

$$\frac{\partial \mu_t}{\partial t} = \frac{\sigma^2}{2}\Delta\mu_t - \mathrm{div}\,(\mu_t b_t)$$

with initial condition $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$. The definition is

$$\langle \mu_t, \phi\rangle - \langle \mu_0, \phi\rangle = \int_0^t \langle b_s\mu_s, \nabla\phi\rangle\, ds + \int_0^t \frac{\sigma^2}{2}\langle \mu_s, \Delta\phi\rangle\, ds$$

for all $\phi \in C_c^2\left(\mathbb{R}^d\right)$. The assumption on b here is that it is a given uniformly-in-time Lipschitz continuous vector field $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$. One proof is by duality: one first prove that, given an arbitrary pair $T' \in [0,T]$ and $g \in C_c^2\left(\mathbb{R}^d\right)$, the backward Kolmogorov equation on $[0,T']$

$$\frac{\partial u_t}{\partial t} + \frac{\sigma^2}{2}\Delta u_t + b_t\cdot\nabla u_t = 0 \qquad u_{T'} = g$$

has a sufficiently regular solution to apply the rules of calculus required in the next argument. Second, one extends the definition of measure-valued solution from every $\phi \in C_c^2\left(\mathbb{R}^d\right)$ to every time-dependent test function u_t in the call of regularity of solutions of the previous Kolmogorov equation. At this point we know that

$$\langle \mu_t, u_t \rangle - \langle \mu_{t'}, u_{t'} \rangle = \int_{t'}^{t} \left\langle \mu_s, \frac{\partial u_s}{\partial s} \right\rangle ds + \int_{t'}^{t} \langle b_s \mu_s, \nabla u_s \rangle \, ds + \int_0^t \frac{\sigma^2}{2} \langle \mu_s, \Delta u_s \rangle \, ds$$

for every $0 \leq t' \leq t \leq T'$. By the equation satisfied by u we get

$$\langle \mu_t, u_t \rangle = \langle \mu_{t'}, u_{t'} \rangle$$

and in particular

$$\langle \mu_t, g \rangle = \langle \mu_0, u_0 \rangle .$$

Since $g \in C_c^2\left(\mathbb{R}^d\right)$ is arbitrary, this identifies μ_t. Hence the linear equation has a unique solution.

Now, let ν_t be a measure-valued solution of equation (6.6) with initial condition $\mu_0 \in \mathrm{Pr}_1\left(\mathbb{R}^d\right)$. It is solution of

$$\frac{\partial \nu_t}{\partial t} = \frac{\sigma^2}{2} \Delta \nu_t - \mathrm{div}\left(\nu_t b_t\right)$$

where $b_t = K * \nu_t$. It is unique (by the argument just explained); but also the marginal of the law ν of a process Y solving

$$dY_t = b_t\left(Y_t\right) dt + \sigma dB_t$$

is a solution. Hence (Y_t, ν_t) is another solution of problem (6.7)-(6.9), thus it is equal to (X_t, μ_t); namely ν_t is equal to μ_t. ■

Now, for every $i \in \mathbb{N}$, consider

$$\begin{aligned} d\overline{X}^i(t) &= v\left(\overline{X}^i(t), t\right) dt + \sigma dW^i(t) \\ v(x,t) &= \int K(x-y)\, \mu_t(dy) \\ \mu_t &= \text{Law of } \overline{X}^i(t) . \end{aligned}$$

The next theorem is the key step in the proof of the macroscopic limit but it is also a very interesting result in itself: it describes the asymptotic (in N) dynamics of each particle X_t^i. It immediately has an interesting corollary, called propagation of chaos.

Theorem 38 *We have*

$$E\left[\sup_{t \in [0,T]} \left| X_t^i - \overline{X}_t^i \right|\right] \leq \frac{C}{\sqrt{N}}.$$

Proof. We have

$$\begin{aligned}
X_t^i - \overline{X}_t^i &= \int_0^t \frac{1}{N} \sum_j K\left(X_s^i - X_s^j\right) ds - \int_0^t v\left(\overline{X}_s^i, s\right) ds \\
&= \int_0^t \frac{1}{N} \sum_j \left(K\left(X_s^i - X_s^j\right) - K\left(\overline{X}_s^i - X_s^j\right)\right) ds \\
&+ \int_0^t \frac{1}{N} \sum_j \left(K\left(\overline{X}_s^i - X_s^j\right) - K\left(\overline{X}_s^i - \overline{X}_s^j\right)\right) ds \\
&+ \int_0^t \left(\frac{1}{N} \sum_j K\left(\overline{X}_s^i - \overline{X}_s^j\right) - v\left(\overline{X}_s^i, s\right)\right) ds
\end{aligned}$$

hence

$$\begin{aligned}
E\left[\left|X_t^i - \overline{X}_t^i\right|\right] &\leq L_K \int_0^t E\left[\left|X_s^i - \overline{X}_s^i\right|\right] ds + L_K \frac{1}{N} \sum_j \int_0^t E\left[\left|X_s^j - \overline{X}_s^j\right|\right] ds \\
&+ \int_0^t E\left[\left|\frac{1}{N} \sum_j K\left(\overline{X}_s^i - \overline{X}_s^j\right) - v\left(\overline{X}_s^i, s\right)\right|\right] ds.
\end{aligned}$$

By exchangeability

$$E\left[\left|X_t^i - \overline{X}_t^i\right|\right] \leq 2L_K \int_0^t E\left[\left|X_s^i - \overline{X}_s^i\right|\right] ds + \int_0^t E\left[\left|\frac{1}{N} \sum_j K\left(\overline{X}_s^i - \overline{X}_s^j\right) - v\left(\overline{X}_s^i, s\right)\right|\right] ds$$

which implies, by Gronwall lemma,

$$E\left[\left|X_t^i - \overline{X}_t^i\right|\right] \leq e^{2L_K T} \int_0^t E\left[\left|\frac{1}{N} \sum_j K\left(\overline{X}_s^i - \overline{X}_s^j\right) - v\left(\overline{X}_s^i, s\right)\right|\right] ds.$$

This is a great estimate: it controls the difference between the solution of a coupled system and the solution of the independent one, by an expression which depends only on the independent one.

We have

$$E\left[\left|\frac{1}{N} \sum_j K\left(\overline{X}_s^i - \overline{X}_s^j\right) - v\left(\overline{X}_s^i, s\right)\right|\right] = E\left[E\left[\left|\frac{1}{N} \sum_j K\left(x - \overline{X}_s^j\right) - v\left(x, s\right)\right|\right]_{x=\overline{X}_s^i}\right]$$

$$\begin{aligned} E\left[\left|\frac{1}{N}\sum_j K\left(x-\overline{X}_s^j\right)-v\left(x,s\right)\right|\right] &\leq E\left[\left|\frac{1}{N}\sum_j K\left(x-\overline{X}_s^j\right)-v\left(x,s\right)\right|^2\right]^{1/2} \\ &\leq \left(\frac{Var\left[K\left(x-X_s\right)\right]}{N}\right)^{1/2} \leq \frac{C}{\sqrt{N}} \end{aligned}$$

hence

$$E\left[\left|\frac{1}{N}\sum_j K\left(\overline{X}_s^i-\overline{X}_s^j\right)-v\left(\overline{X}_s^i,s\right)\right|\right] \leq \frac{C}{\sqrt{N}}$$

$$E\left[\left|X_t^i-\overline{X}_t^i\right|\right] \leq e^{2L_K T}\frac{CT}{\sqrt{N}}.$$

■

Before we complete the main program about the convergence of S_t^N, we may deduce from the previous theorem the so called property of *propagation of chaos*: the independence of initial conditions propagates to an approximate independence of positions at any time t, in the limit as $N \to \infty$. And, the previous theorem also tells us, in the limit each independent particle satisfies an SDE (the equation for $\overline{X}_t^i$) in interaction with a mean field (the density of $\overline{X}_t^i$ itself), no more with single particles.

Corollary 39 *For every given k and every $\varphi_1, ..., \varphi_k \in C_b\left(\mathbb{R}^d\right) \cap Lip\left(\mathbb{R}^d\right)$, we have*

$$\lim_{N\to\infty} E\left[\prod_{i=1}^k \varphi_i\left(X_t^i\right)\right] = \prod_{i=1}^k E\left[\varphi_i\left(X_t\right)\right] = \prod_{i=1}^k \left\langle \varphi_i, \mu_t\right\rangle.$$

Proof. It is sufficient to explain the proof for $k=2$. We have

$$\begin{aligned} & E\left[\varphi_1\left(X_t^1\right)\varphi_2\left(X_t^2\right)\right] - E\left[\varphi_1\left(X_t\right)\right]E\left[\varphi_2\left(X_t\right)\right] \\ = \; & E\left[\varphi_1\left(X_t^1\right)\varphi_2\left(X_t^2\right) - \varphi_1\left(\overline{X}_t^1\right)\varphi_2\left(\overline{X}_t^2\right)\right] \end{aligned}$$

and in absolute value it is bounded above by

$$\begin{aligned} &\leq E\left[\left|\varphi_1\left(X_t^1\right)\varphi_2\left(X_t^2\right) - \varphi_1\left(\overline{X}_t^1\right)\varphi_2\left(\overline{X}_t^2\right)\right|\right] \\ &\leq \|\varphi_1\|_0 E\left[\left|\varphi_2\left(X_t^2\right) - \varphi_2\left(\overline{X}_t^2\right)\right|\right] + \|\varphi_2\|_0 E\left[\left|\varphi_2\left(X_t^2\right) - \varphi_2\left(\overline{X}_t^2\right)\right|\right] \\ &\leq 2\|\varphi_1\|_{Lip}\|\varphi_2\|_{Lip} E\left[\left|X_t^1 - \overline{X}_t^1\right|\right] \end{aligned}$$

which converges to zero as $N \to \infty$. ■

Remark 40 *The proof also shows that*

$$\left|E\left[\varphi_1\left(X_t^1\right)\varphi_2\left(X_t^2\right)\right]-E\left[\varphi_1\left(X_t\right)\right]E\left[\varphi_2\left(X_t\right)\right]\right|\leq 2C\left\|\varphi_1\right\|_{Lip}\left\|\varphi_2\right\|_{Lip}/\sqrt{N}. \tag{6.10}$$

We complete the section with the following result, which solves the problem of the macroscopic limit, since we already know that μ_t is the solution of the PDE:

Theorem 41 *If μ_t denotes the law of X_t, we have*

$$E\left[\left|\left\langle S_t^N,\phi\right\rangle-\left\langle\mu_t,\phi\right\rangle\right|^2\right]\leq\frac{C}{\sqrt{N}}$$

for all $\phi\in C_b\left(\mathbb{R}^d\right)\cap Lip\left(\mathbb{R}^d\right)$, for a suitable constant $C>0$.

Proof. We give only an idea of proof. One has

$$E\left[\left|\left\langle S_t^N,\phi\right\rangle-\left\langle\mu_t,\phi\right\rangle\right|^2\right]=E\left[\left\langle S_t^N,\phi\right\rangle^2\right]+\left\langle\mu_t,\phi\right\rangle^2-2E\left[\left\langle S_t^N,\phi\right\rangle\right]\left\langle\mu_t,\phi\right\rangle.$$

By exchangeability,

$$\begin{aligned}E\left[\left\langle S_t^N,\phi\right\rangle^2\right] &= \frac{1}{N^2}\sum_{ij}E\left[\phi\left(X_t^i\right)\phi\left(X_t^j\right)\right]=\frac{1}{N^2}\sum_{i}E\left[\phi\left(X_t^1\right)^2\right]+\frac{1}{N^2}\sum_{i\neq j}E\left[\phi\left(X_t^1\right)\phi\left(X_t^2\right)\right]\\ &= \frac{1}{N}E\left[\phi\left(X_t^1\right)^2\right]+\frac{N-1}{N}E\left[\phi\left(X_t^1\right)\phi\left(X_t^2\right)\right]\end{aligned}$$

$$\left\langle\mu_t,\phi\right\rangle^2=E\left[\phi\left(X_t\right)\right]^2$$

$$E\left[\left\langle S_t^N,\phi\right\rangle\right]\left\langle\mu_t,\phi\right\rangle=\frac{1}{N}\sum_{i}E\left[\phi\left(X_t^i\right)\right]E\left[\phi\left(X_t\right)\right]=E\left[\phi\left(X_t^1\right)\right]E\left[\phi\left(X_t\right)\right].$$

Hence

$$\begin{aligned}E\left[\left\langle S_t^N,\phi\right\rangle^2\right] &= \frac{1}{N}E\left[\phi\left(X_t^1\right)^2\right]+\frac{N-1}{N}E\left[\phi\left(X_t^1\right)\phi\left(X_t^2\right)\right]+E\left[\phi\left(X_t\right)\right]^2-2E\left[\phi\left(X_t^1\right)\right]E\left[\phi\left(X_t\right)\right]\\ &= \frac{1}{N}E\left[\phi\left(X_t^1\right)^2\right]-\frac{1}{N}E\left[\phi\left(X_t^1\right)\phi\left(X_t^2\right)\right]\\ &\quad +E\left[\phi\left(X_t^1\right)\phi\left(X_t^2\right)\right]-E\left[\phi\left(X_t\right)\right]^2\\ &\quad +2E\left[\phi\left(X_t\right)\right]^2-2E\left[\phi\left(X_t^1\right)\right]E\left[\phi\left(X_t\right)\right].\end{aligned}$$

The term in the first line is bounded by $2\left\|\phi\right\|_0^2/N$. The term in the second line is bounded by $2C\left\|\phi\right\|_{Lip}^2/\sqrt{N}$ (see the remark above) and similarly the term in the third line. ■

6.2 Intermediate regime with large amplitudes

We consider now the case

$$dX_t^i = -\frac{1}{N}\sum_j \nabla V_N\left(X_t^i - X_t^j\right)dt + \sigma dW_t^i$$

in the intermediate regime with large amplitudes, namely when V_N are standard mollifiers converging to δ_0 and the guess is that the limit PDE is

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2}\Delta u + \operatorname{div}\left(u\nabla u\right).$$

The approach above based on the comparison with the mean-field SDE, Section 6.1.2, has been extended to a problem of this form (under some special assumptions on the smallness of parameters, and also with $\sigma = \sigma_N \to 0$) by [17]. We limit ourselves here to a short description of the other approach, by compactness, following [12], see also [14].

We want to prove tightness of the empirical measure S_t^N. The size of the input $\frac{1}{N}\sum_j \nabla V_N\left(X_t^i - X_t^j\right)$ is too large to make easy a priori estimates as in Theorem 33 above. Specifically, we cannot estimate so easily $\frac{1}{N}\sum\left|X_t^i - X_s^i\right|$. The idea then comes from the limit PDE. For this equation, the standard energy estimate reads

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}\int u^2 dx &= \int u\frac{\partial u}{\partial t}dx = \int u\left(\frac{\sigma^2}{2}\Delta u + \operatorname{div}\left(u\nabla u\right)\right)dx \\ &= -\frac{\sigma^2}{2}\int \left|\nabla u\right|^2 dx - \int u\left|\nabla u\right|^2 dx\end{aligned}$$

hence

$$\frac{1}{2}\int u_t^2 dx + \int_0^t\left(\frac{\sigma^2}{2}\int\left|\nabla u\right|^2 dx + \int u\left|\nabla u\right|^2 dx\right)ds = \frac{1}{2}\int u_0^2 dx.$$

This provides estimates on ∇u. Potentially, in a very intuitive sense, they are what we need. indeed,

$$-\frac{1}{N}\sum_j \nabla V_N\left(X_t^i - X_t^j\right)dt = -\nabla u_N\left(t, X_t^i\right)$$

where

$$u_N\left(t,x\right) = \left(V_N * S_t^N\right)\left(x\right) = \frac{1}{N}\sum_j V_N\left(x - X_t^j\right)$$

and, due to the mollifier property of V_N, we expect (intuitively) $u_N\left(t,x\right) \to u\left(t,x\right)$ (in some sense). Thus we could expect a bound on ∇u_N, similarly to the bound on ∇u. And a bound on ∇u_N may provide tightness of S_t^N since

$$dX_t^i = -\nabla u_N\left(t, X_t^i\right)dt + \sigma dW_t^i. \tag{6.11}$$

Thus the idea is to look for an "energy estimate" for u_N, in place of u. But u_N does not satisfy a PDE. Is it true? Does it satisfy some modified PDE? Let us first argue heuristically.

From Itô formula we have

$$\begin{aligned} d\left\langle S_t^N, \varphi\right\rangle &= \frac{1}{N}\sum_i \nabla\varphi\left(X_t^i\right)\cdot\left(-\frac{1}{N}\sum_j \nabla V_N\left(X_t^i - X_t^j\right)dt + \sigma dW_t^i\right) + \frac{\sigma^2}{2N}\sum_i \Delta\varphi\left(X_t^i\right)dt \\ &= -\left\langle S_t^N, \nabla\varphi\cdot\nabla V_N * S_t^N\right\rangle dt + dM_t^N + \frac{\sigma^2}{2}\left\langle S_t^N, \Delta\varphi\right\rangle dt \end{aligned}$$

hence, if we neglect dM_t^N, we formally have the equation

$$\frac{\partial S_t^N}{\partial t} = \operatorname{div}\left(S_t^N \nabla V_N * S_t^N\right) + \frac{\sigma^2}{2}\Delta S_t^N.$$

In order to write an energy estimate we have to mollify S_t^N, we need a function and not a pointwise measure. Apply the convolution with a mollifier W_N (in fact the next equation can be obtained rigorously from the formula above for $d\left\langle S_t^N, \varphi\right\rangle$, by a suitable choice of φ)

$$\frac{\partial\left(W_N * S_t^N\right)}{\partial t} = \operatorname{div}\left(W_N * \left(S_t^N \nabla V_N * S_t^N\right)\right) + \frac{\sigma^2}{2}\Delta\left(W_N * S_t^N\right)$$

where the rigorous meaning of $W_N * \left(S_t^N \nabla V_N * S_t^N\right)$ is

$$W_N * \left(S_t^N \nabla V_N * S_t^N\right)(x) := \int W_N(x-y)\left(\nabla V_N * S_t^N\right)(y)\, S_t^N(dy).$$

Now we have

$$\frac{d}{dt}\frac{1}{2}\int\left(W_N * S_t^N\right)^2 dx = -\frac{\sigma^2}{2}\int\left|\nabla\left(W_N * S_t^N\right)\right|^2 dx - \int \nabla W_N * S_t^N \cdot W_N * \left(S_t^N \nabla V_N * S_t^N\right) dx.$$

We cannot say anything of the sign of the last term. But is we make the very special choice

$$\begin{aligned} V_N &= W_N * W_N^- \\ \nabla V_N &= \nabla W_N * W_N^- = W_N^- * \nabla W_N \end{aligned}$$

the last term becomes

$$\begin{aligned} &= -\int W_N^- * \nabla W_N * S_t^N \cdot \nabla V_N * S_t^N S_t^N(dx) \\ &= -\int \nabla V_N * S_t^N \cdot \nabla V_N * S_t^N S_t^N(dx) \\ &= -\int\left|\nabla V_N * S_t^N\right|^2 S_t^N(dx). \end{aligned}$$

This incredible trick provides the estimate

$$\frac{d}{dt}\frac{1}{2}\int\left(W_N * S_t^N\right)^2 dx + \frac{\sigma^2}{2}\int\left|\nabla\left(W_N * S_t^N\right)\right|^2 dx + \int\left|\nabla V_N * S_t^N\right|^2 S_t^N(dx) = 0$$

in the case $\sigma = 0$ (which is not our case, but it has been treated in the literature).

We have used the following simple facts.

Lemma 42 *i)* $f * g = g * f$

ii)

$$(f * g) * h = f * (g * h)$$

hence we may write

$$f * g * h$$

without ambiguities

iii)

$$\int (f * g)\, h = \int g\left(f^- * h\right)$$

where $f^-(x) = f(-x)$.

Proof.

$$\begin{aligned}
((f * g) * h)(x) &= \int (f * g)(x - y)\, h(y)\, dy \\
&= \int\int f(x - y - z)\, g(z)\, h(y)\, dz dy \\
&= \int\int f\left(x - y'\right) g(z)\, h\left(y' - z\right) dz dy' \\
&= \int f\left(x - y'\right) dy' \int g(z)\, h\left(y' - z\right) dz.
\end{aligned}$$

$$\begin{aligned}
\int (f * g)(x)\, h(x)\, dx &= \int\int f(x - y)\, g(y)\, dy h(x)\, dx \\
&= \int\int f(x - y)\, h(x)\, dx g(y)\, dy.
\end{aligned}$$

■

Let us formalize a few details outlined above, also to show how to deal with the stochastic case, $\sigma \neq 0$..

Theorem 43 *The empirical measure process*

$$S_t^N := \frac{1}{N}\sum_j \delta_{X_t^j}$$

is tight in $C\left([0, T]; \mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$.

Proof. From

$$\begin{aligned} d\left\langle S_t^N, \varphi\right\rangle &= \frac{1}{N}\sum_i \nabla\varphi\left(X_t^i\right)\cdot\left(-\frac{1}{N}\sum_j \nabla V_N\left(X_t^i - X_t^j\right)dt + \sigma dW_t^i\right) + \frac{\sigma^2}{2N}\sum_i \Delta\varphi\left(X_t^i\right)dt \\ &= -\left\langle S_t^N, \nabla\varphi\cdot\nabla V_N * S_t^N\right\rangle dt + dM_t^{N,\varphi} + \frac{\sigma^2}{2}\left\langle S_t^N, \Delta\varphi\right\rangle dt \end{aligned}$$

with $\varphi(y) = \varphi_x(y) := W_N(x-y)$ and $\langle .,. \rangle$ interpreted as $\int ...dy$, where

$$dM_t^{N,\varphi} = \frac{\sigma}{N}\sum_i \nabla\varphi\left(X_t^i\right)\cdot dW_t^i$$

we get

$$\begin{aligned} d\left(W_N * S_t^N\right)(x) &= \frac{1}{N}\sum_i \nabla\varphi\left(X_t^i\right)\cdot\left(-\frac{1}{N}\sum_j \nabla V_N\left(X_t^i - X_t^j\right)dt + \sigma dW_t^i\right) + \frac{\sigma^2}{2N}\sum_i \Delta\varphi\left(X_t^i\right)dt \\ &= \operatorname{div}\left(W_N * \left(S_t^N \nabla V_N * S_t^N\right)\right)dt + \frac{\sigma^2}{2}\Delta\left(W_N * S_t^N\right)(x)\,dt + dM_t^{N,\varphi_x} \end{aligned}$$

because

$$\left\langle S_t^N, \Delta\varphi_x\right\rangle = \Delta\left(W_N * S_t^N\right)(x)$$

$$\begin{aligned} \left\langle S_t^N, \nabla\varphi_x\cdot\nabla V_N * S_t^N\right\rangle &= \int \nabla_y W_N(x-y)\cdot\left(\nabla V_N * S_t^N\right)(y)\,S_t^N(dy) \\ &= -\operatorname{div}\int W_N(x-y)\left(\nabla V_N * S_t^N\right)(y)\,S_t^N(dy) \\ &= -\operatorname{div}\left(W_N * \left(S_t^N \nabla V_N * S_t^N\right)\right). \end{aligned}$$

Therefore, by Itô formula,

$$\begin{aligned} d\frac{1}{2}\left|\left(W_N * S_t^N\right)(x)\right|^2 &= \left(W_N * S_t^N\right)(x)\left(\operatorname{div}\left(W_N * \left(S_t^N \nabla V_N * S_t^N\right)\right)dt + \frac{\sigma^2}{2}\Delta\left(W_N * S_t^N\right)(x)\,dt\right) \\ &\quad + \left(W_N * S_t^N\right)(x)\,dM_t^{N,\varphi_x} \\ &\quad + d\left[M^{N,\varphi_x}\right]_t \end{aligned}$$

and now we use the brilliant computation done above on the term $\int \left(W_N * S_t^N\right)\operatorname{div}\left(W_N * \left(S_t^N \nabla V_N * S_t^N\right)\right)$

$$\begin{aligned} & d\frac{1}{2}\int\left|\left(W_N * S_t^N\right)\right|^2 dx + \frac{\sigma^2}{2}\int\left|\nabla\left(W_N * S_t^N\right)\right|^2 dxdt + \int\left|\nabla V_N * S_t^N\right|^2 S_t^N(dx)\,dt \\ = & \int\left(W_N * S_t^N\right)(x)\,dM_t^{N,\varphi_x}dx + \frac{\sigma^2}{N^2}\sum_i\left|\nabla\varphi\left(X_t^i\right)\right|^2 dt. \end{aligned}$$

The last term goes to zero and the previous one ha zero expectation. Hence at least we get

$$\begin{aligned} & E\left[\frac{1}{2}\int\left|\left(W_N * S_t^N\right)\right|^2 dx\right] + \frac{\sigma^2}{2}\int_0^t\int E\left[\left|\nabla\left(W_N * S_t^N\right)\right|^2\right] dxdt + \int_0^t\int E\left[\left|\nabla V_N * S_t^N\right|^2 S_t^N(dx)\right] dt \\ \leq\ & E\left[\frac{1}{2}\int\left|\left(W_N * S_0^N\right)\right|^2 dx\right] + \frac{C\sigma^2}{N} =: c_N. \end{aligned}$$

Under appropriate assumptions on initial conditions one has

$$C : \sup_N c_N < \infty.$$

This is the surrogate of the estimate on ∇u guessed above. Now we have to use it rigorously to prove tightness. As in the proof of Theorem 33, we initially have

$$\begin{aligned} \mathcal{W}_1\left(S_t^N, S_s^N\right) &\leq \frac{1}{N}\sum_i\left|X_t^i - X_s^i\right| \\ &\leq \frac{1}{N}\sum_i\int_s^t\left|\nabla u_N\left(r, X_r^i\right)\right| dr + \sigma\frac{1}{N}\sum_i\left|W_t^i - W_s^i\right| \end{aligned}$$

and

$$E\left[\left(\frac{1}{N}\sum\left|W_t^i - W_s^i\right|\right)^{2p}\right] \leq c_p\,(t-s)^p$$

for every $p \geq 1$. Here we have used (6.11). Moreover,

$$\int_s^t\frac{1}{N}\sum\left|\nabla u_N\left(r, X_r^i\right)\right| dr = \int_s^t\left\langle\left|\nabla u_N(r,\cdot)\right|, S_r^N\right\rangle dr \leq (t-s)^{1/2}\left(\int_s^t\left\langle\left|\nabla u_N(r,\cdot)\right|^2, S_r^N\right\rangle dr\right)^{1/2}$$

and recall $u_N(t,x) = \left(V_N * S_t^N\right)(x)$. Hence

$$\begin{aligned} E\left[\left(\int_s^t\frac{1}{N}\sum\left|\nabla u_N\left(r, X_r^i\right)\right| dr\right)^2\right] &\leq (t-s)\int_0^t\int E\left[\left|\nabla V_N * S_t^N\right|^2 S_t^N(dx)\right] dt \\ &\leq C\,(t-s). \end{aligned}$$

Hence we get the bound

$$E\left[\left\|S_t^N - S_s^N\right\|_W^2\right] \leq (C+1)\,(t-s).$$

To reach $(t-s)^2$ on the right-hand-side (we need (6.4)), it requires additional work with stopping times, that we omit, see [12]; similarly for the proof of (6.3). ■

With due little effort, the bound

$$\int_0^t\int E\left[\left|\nabla\left(W_N * S_t^N\right)\right|^2\right] dxdt \leq C$$

gives us the information that the limit measure μ_t of S_t^N has a density ρ_t w.r.t. Lebesgue measure and

$$\int_0^t \int E\left[\left|\nabla \rho_t(x)\right|^2\right] dxdt \leq C.$$

We omit the non trivial remaining parts of the proof (and also the precise statement of the theorem), namely the fact that $\rho_t(x)$ satisfies in a weak sense the PDE and that it is deterministic and unique. In a certain philosophy of PDE theory, we pretend that the a priori bounds are the hard core of the matter, the part without which the result cannot be true, hence we felt the urgency to show them, in contrast to other (still very tricky) details. See [12].

Chapter 7

Appendix on R codes

We copy here some of the R codes used to produce the pictures above, that the reader may cut and paste on R. These are highly non-professional codes, that every scientist with a minimum of experience may improve.

7.1 Codes of Chapter 2

7.1.1 FKPP equation

Next code simulates FKPP equation. It allows us to see the beginning of formation of a traveling wave. Moreover, in the first phases, one can see first the diffusive depletion followed by the proliferating expansion, up to the threshold.

The simulation is done in such a way that we can see pictures every `Temp` time steps. The numerical method is based on the so called splitting method. The parabolic step is done by finite differences. The ODE nonlinear step is done by explicit Euler method. At the boundary we impose no flux conditions, to avoid depletion.

```
Temp=20; NT0 = 500
NT= NT0*Temp; Nx=1000; dx=0.1; K=1; npas=4
L=dx*Nx
dt=(dx^2)/(4*K)
Nx.virt=Nx+1
u = matrix(nrow=Nx.virt, ncol=NT)
v = matrix(nrow=Nx.virt, ncol=npas+1)
w = matrix(nrow=Nx.virt, ncol=npas+1)
X=seq(0,L,dx)
u[,1]=exp(-2*abs((X-L/2)^2))
for (t0 in 1:(NT0-1)) {
for (t1 in 1:Temp) {
t=(t0-1)*Temp + t1
```

```
v[,1]=u[,t]
for (k in 1:npas) {
for (i in 2:Nx) {
v[i,k+1] = v[i,k] + K * dt * ((v[i+1,k]-2*v[i,k]+v[i-1,k])) / (dx^2)
}
v[1,k+1]=v[2,k+1]; v[Nx.virt,k+1]=v[Nx,k+1]
}
w[,1]=v[,npas+1]
for (k in 1:npas) {
for (i in 1:Nx.virt) {
w[i,k+1] = w[i,k] + dt * w[i,k]*(1-w[i,k])
}
}
u[,t+1]=w[,npas+1]
}
plot(c(0,L),c(-0.5,1.5),type="n")
lines(X,u[,t+1])
lines(c(0,L),c(1,1))
lines(c(0,L),c(0,0))
}
```

7.1.2 ODEs for traveling waves

We have "proved" the existence and no existence of traveling waves by solving an ODE, which is simulated here.

The first code shows the heteroclinic orbit of the ODE which proves existence of a traveling wave, for c=sqrt(2):

```
N=30000; X=1:N; Y=1:N; dt=0.001; c=sqrt(2); eps=0.0001
ex=0.5*c+0.5*sqrt(c^2+2); ey=1
X[1]=1-eps*ex; Y[1]=-eps*ey
for (i in 1:(N-1)) {
X[i+1]=X[i]+dt*Y[i]
Y[i+1]=Y[i]+dt*(-2*c*Y[i]-2*X[i]+2*X[i]^2)
}
plot(c(0,1),c(-0.6,0),type="n")
lines(X,Y)
lines(c(0,1),c(0,0))
lines(c(0,0),c(-0.6,0))
lines(c(0,1),c(-1/ex,0))
```

The second code shows the heteroclinic orbit for c=1 which turns around the origin and thus proves the non-existence of traveling wave:

```
N=30000; X=1:N; Y=1:N; dt=0.001; c=0.8; eps=0.0001
ex=0.5*c+0.5*sqrt(c^2+2); ey=1
X[1]=1-eps*ex; Y[1]=-eps*ey
for (i in 1:(N-1)) {
X[i+1]=X[i]+dt*Y[i]
Y[i+1]=Y[i]+dt*(-2*c*Y[i]-2*X[i]+2*X[i]^2)
}
plot(c(0,1),c(-0.6,0),type="n")
lines(X,Y)
lines(c(0,1),c(0,0))
lines(c(0,0),c(-0.6,0))
lines(c(0,1),c(-1/ex,0))
```

7.2 Chapter 3

7.2.1 Normoxic + apoptotic cells only

The following code simulates a system composed only by normoxic (black) + apoptotic (red) cells. It is useful to show that the total volume (in green) may exceed the threshold.

```
Temp=10; NT0 = 100
NT= NT0*Temp; Nx=600; dx=0.1; K=1; chi=0.1; npas=4; int=0
L=dx*Nx
dt=(dx^2)/(4*K)
Nx.virt=Nx+1
u = matrix(nrow=Nx.virt, ncol=NT)
a = matrix(nrow=Nx.virt, ncol=NT)
v = matrix(nrow=Nx.virt, ncol=npas+1)
w = matrix(nrow=Nx.virt, ncol=npas+1)
z = matrix(nrow=Nx.virt, ncol=npas+1)
X=seq(0,L,dx)
u[,1]=exp(-2*abs((X-L/2)^2))
a[,1]=exp(-2*abs((X-L/3)^2))*int
for (t0 in 1:(NT0-1)) {
for (t1 in 1:Temp) {
t=(t0-1)*Temp + t1
v[,1]=u[,t]
for (k in 1:npas) {
for (i in 2:Nx) {
v[i,k+1] = v[i,k] +K* dt * ((v[i+1,k]-2*v[i,k]+v[i-1,k])) / (dx^2)
}
v[1,k+1]=v[2,k+1]; v[Nx.virt,k+1]=v[Nx,k+1]
```

```
}
w[,1]=v[,npas+1]
z[,1]=a[,t]
for (k in 1:npas) {
for (i in 1:Nx.virt) {
w[i,k+1] = w[i,k] + dt * w[i,k]*(1-w[i,k]-z[i,k]) - dt *chi*w[i,k]
z[i,k+1] = z[i,k] + dt * chi*w[i,k]
}
}
u[,t+1]=w[,npas+1]
a[,t+1]=z[,npas+1]
}
plot(c(0,L),c(-0.5,1.5))
lines(X,u[,t+1],col="black")
lines(X,a[,t+1],col="red")
lines(X,u[,t+1]+a[,t+1],col="green")
lines(c(0,L),c(1,1))
lines(c(0,L),c(0,0))
}
```

7.2.2 Normoxic + apoptotic but with damped diffusion

The previous example is modified by pre-multiplying the diffusion operator by *(1-v[i,k]-a[i,t])

The threshold is not overcome, as the theory of invariant regions prove.

```
Temp=20; NT0 = 500
NT= NT0*Temp; Nx=600; dx=0.1; K=1; chi=0.1; npas=4; int=0
L=dx*Nx
dt=(dx^2)/(4*K)
Nx.virt=Nx+1
u = matrix(nrow=Nx.virt, ncol=NT)
a = matrix(nrow=Nx.virt, ncol=NT)
v = matrix(nrow=Nx.virt, ncol=npas+1)
w = matrix(nrow=Nx.virt, ncol=npas+1)
z = matrix(nrow=Nx.virt, ncol=npas+1)
X=seq(0,L,dx)
u[,1]=exp(-2*abs((X-L/2)^2))
a[,1]=exp(-2*abs((X-L/3)^2))*int
for (t0 in 1:(NT0-1)) {
for (t1 in 1:Temp) {
t=(t0-1)*Temp + t1
```

```
v[,1]=u[,t]
for (k in 1:npas) {
for (i in 2:Nx) {
v[i,k+1] = v[i,k] +K* dt * ((v[i+1,k]-2*v[i,k]+v[i-1,k])) / (dx^2) *(1-v[i,k]-a[i,t])
}
v[1,k+1]=v[2,k+1]; v[Nx.virt,k+1]=v[Nx,k+1]
}
w[,1]=v[,npas+1]
z[,1]=a[,t]
for (k in 1:npas) {
for (i in 1:Nx.virt) {
w[i,k+1] = w[i,k] + dt * w[i,k]*(1-w[i,k]-z[i,k]) - dt *chi*w[i,k]
z[i,k+1] = z[i,k] + dt * chi*w[i,k]
}
}
u[,t+1]=w[,npas+1]
a[,t+1]=z[,npas+1]
}
plot(c(0,L),c(-0.5,1.5))
lines(X,u[,t+1],col="black")
lines(X,a[,t+1],col="red")
lines(X,u[,t+1]+a[,t+1],col="green")
lines(c(0,L),c(1,1))
lines(c(0,L),c(0,0))
}
```

7.2.3 Trasport terms may overcome given thresholds

The next code simulate the heat equation with a drift (transport) which concentrates mass around a point.

```
Temp=20; NT0 = 500; npas=4
NT= NT0*Temp; Nx=1000; dx=0.1; K=1; chi=0.1; int=0
L=dx*Nx
dt=(dx^2)/(4*K)
Nx.virt=Nx+1
u = matrix(nrow=Nx.virt, ncol=NT)
v = matrix(nrow=Nx.virt, ncol=npas+1)
w = matrix(nrow=Nx.virt, ncol=npas+1)
z = matrix(nrow=Nx.virt, ncol=npas+1)
X=seq(0,L,dx)
b = -5*(X-L/2-10)*exp(-0.08*abs((X-L/2-10)^2))
```

```
   u[,1]=exp(-0.1*abs((X-L/2)^2))
   for (t0 in 1:(NT0-1)) {
   for (t1 in 1:Temp) {
   t=(t0-1)*Temp + t1
   v[,1]=u[,t]
   for (k in 1:npas) {
   for (i in 2:Nx) {
   v[i,k+1] = v[i,k] +K*dt*((v[i+1,k]-2*v[i,k]+v[i-1,k]))/(dx^2) - dt*(b[i+1]*v[i+1,k]-b
   }
   v[1,k+1]=v[2,k+1]; v[Nx.virt,k+1]=v[Nx,k+1]
   }
   u[,t+1]=v[,npas+1]
   }
   plot(c(0,L),c(-0.5,1.5))
   lines(X,u[,t+1],col="black")
   lines(c(0,L),c(1,1))
   lines(c(0,L),c(0,0))
   }
```

7.2.4 Normoxic + hypoxic + apoptotic

```
Temp=10; NT0 = 100; npas=4
   NT= NT0*Temp; Nx=600; dx=0.1; K=1; chi1=0.1; chi2=0.1 ; int=0
   L=dx*Nx
   dt=(dx^2)/(4*K)
   Nx.virt=Nx+1
   n = matrix(nrow=Nx.virt, ncol=NT)
   h = matrix(nrow=Nx.virt, ncol=NT)
   a = matrix(nrow=Nx.virt, ncol=NT)
   nd = matrix(nrow=Nx.virt, ncol=npas+1)
   nn = matrix(nrow=Nx.virt, ncol=npas+1)
   hh = matrix(nrow=Nx.virt, ncol=npas+1)
   aa = matrix(nrow=Nx.virt, ncol=npas+1)
   X=seq(0,L,dx)
   n[,1]=exp(-2*abs((X-L/2)^2))
   h[,1]=exp(-2*abs((X-L/3)^2))*int
   a[,1]=exp(-2*abs((X-L/3)^2))*int
   for (t0 in 1:(NT0-1)) {
   for (t1 in 1:Temp) {
   t=(t0-1)*Temp + t1
   nd[,1]=n[,t]
```

```
for (k in 1:npas) {
for (i in 2:Nx) {
nd[i,k+1] = nd[i,k] +K* dt * ((nd[i+1,k]-2*nd[i,k]+nd[i-1,k])) / (dx^2)
}
nd[1,k+1]=nd[2,k+1]; nd[Nx.virt,k+1]=nd[Nx,k+1]
}
nn[,1]=nd[,npas+1]
hh[,1]=h[,t]
aa[,1]=a[,t]
for (k in 1:npas) {
for (i in 1:Nx.virt) {
nn[i,k+1] = nn[i,k] + dt * nn[i,k]*(1-nn[i,k]-hh[i,k]-aa[i,k]) - dt *chi1*nn[i,k]
hh[i,k+1] = hh[i,k] + dt *chi1*nn[i,k] - dt *chi2*hh[i,k]
aa[i,k+1] = aa[i,k] + dt * chi2*hh[i,k]
}
}
n[,t+1]=nn[,npas+1]
h[,t+1]=hh[,npas+1]
a[,t+1]=aa[,npas+1]
}
plot(c(0,L),c(-0.5,1.5))
lines(X,n[,t+1],col="blue")
lines(X,h[,t+1],col="red")
lines(X,a[,t+1],col="black")
lines(X,n[,t+1]+h[,t+1]+a[,t+1],col="green")
lines(c(0,L),c(1,1))
lines(c(0,L),c(0,0))
}
```

7.2.5 Full system

```
Temp=5; NT0 = 400; npas=4; NT= NT0*Temp
   Nx=200; dx=0.1; L=dx*Nx; Nx.virt=Nx+1; X=seq(0,L,dx)
   Kn=0.1; xhi =1; Kg=15; rho=5; Ke=0.1; Ko=15; OH=0.5; OA=0.2; Vmax=1; rho2=5;
Omax=1; rho3=10
   Cnh=1; Cha=0.1 ; Cmn=2; Cgh=10; Ceg=1; Chn=1; Cocell=0.1; gam=0.000
   int=0; m0=0.5
   dt=(dx^2)/(4*max(Kn,Kg,Ke,Ko))
   n = matrix(nrow=Nx.virt, ncol=NT); h = matrix(nrow=Nx.virt, ncol=NT); a =
matrix(nrow=Nx.virt, ncol=NT)
   m = matrix(nrow=Nx.virt, ncol=NT); g = matrix(nrow=Nx.virt, ncol=NT)
```

```
   e = matrix(nrow=Nx.virt, ncol=NT); o = matrix(nrow=Nx.virt, ncol=NT)
   nd = matrix(nrow=Nx.virt, ncol=npas+1); gd = matrix(nrow=Nx.virt, ncol=npas+1)
   ed = matrix(nrow=Nx.virt, ncol=npas+1); od = matrix(nrow=Nx.virt, ncol=npas+1)
   nn = matrix(nrow=Nx.virt, ncol=npas+1); hh = matrix(nrow=Nx.virt, ncol=npas+1)
   aa = matrix(nrow=Nx.virt, ncol=npas+1); b = 1:Nx.virt; mm = matrix(nrow=Nx.virt,
ncol=npas+1)
   gg = matrix(nrow=Nx.virt, ncol=npas+1); ee = matrix(nrow=Nx.virt, ncol=npas+1)
   oo = matrix(nrow=Nx.virt, ncol=npas+1)
   n[,1]=exp(-10*abs((X-L/2)^2)); h[,1]=0; a[,1]=0; m[,1]=m0; g[,1]=0
   e[,1]=(sign(X-3*L/4)+1)/2 + (sign(L/4-X)+1)/2; o[,1]=1
   for (t0 in 1:(NT0-1)) {
   for (t1 in 1:Temp) {
   t=(t0-1)*Temp + t1; nd[,1]=n[,t]; gd[,1]=g[,t]; ed[,1]=e[,t]; od[,1]=o[,t]
   for (i in 2:Nx.virt) {
   b[i] = (m[i,t]-m[i-1,t])/dx
   }
   for (k in 1:npas) {
   for (i in 2:Nx) {
   nd[i,k+1] = nd[i,k] +Kn* dt * ((nd[i+1,k]-2*nd[i,k]+nd[i-1,k])) / (dx^2)-
dt*(b[i+1]*nd[i+1,k]-b[i]*nd[i,k])/dx
   gd[i,k+1] = gd[i,k] +Kg* dt * ((gd[i+1,k]-2*gd[i,k]+gd[i-1,k])) / (dx^2)
   ed[i,k+1] = ed[i,k] +Ke* dt * ((ed[i+1,k]-2*ed[i,k]+ed[i-1,k])) / (dx^2)
   od[i,k+1] = od[i,k] +Ko* dt * ((od[i+1,k]-2*od[i,k]+od[i-1,k])) / (dx^2)
   }
   nd[1,k+1]=nd[2,k+1]; nd[Nx.virt,k+1]=nd[Nx,k+1]
   gd[1,k+1]=gd[2,k+1]; gd[Nx.virt,k+1]=gd[Nx,k+1]
   ed[1,k+1]=ed[2,k+1]; ed[Nx.virt,k+1]=ed[Nx,k+1]
   od[1,k+1]=od[2,k+1]; od[Nx.virt,k+1]=od[Nx,k+1]
   }
   nn[,1]=nd[,npas+1]; hh[,1]=h[,t]; aa[,1]=a[,t]; mm[,1]=m[,t]; gg[,1]=gd[,npas+1];
ee[,1]=ed[,npas+1]; oo[,1]=od[,npas+1]
   for (k in 1:npas) {
   for (i in 1:Nx.virt) {
   nn[i,k+1] = nn[i,k] + dt * rho* nn[i,k]*(Vmax-nn[i,k]-hh[i,k]-aa[i,k]-ee[i,k])
- dt *Cnh*nn[i,k]*(sign(OH-oo[i,k])+1)/2 + dt *Chn*hh[i,k]*(sign(oo[i,k]-OH)+1)/2
   hh[i,k+1] = hh[i,k] + dt *Cnh*nn[i,k]*(sign(OH-oo[i,k])+1)/2 - dt *Chn*hh[i,k]*(sign(
- dt *Cha*hh[i,k]*(sign(OA-oo[i,k])+1)/2
   aa[i,k+1] = aa[i,k] + dt * Cha*hh[i,k]*(sign(OA-oo[i,k])+1)/2
   mm[i,k+1] = mm[i,k] - dt * Cmn*nn[i,k]*mm[i,k]
   gg[i,k+1] = gg[i,k] + dt * Cgh*hh[i,k] - Ceg*ee[i,k]*gg[i,k]
   ee[i,k+1] = ee[i,k] + dt * rho2* ee[i,k]*gg[i,k]*(Vmax-nn[i,k]-hh[i,k]-aa[i,k]-ee[i,k
```

```
   oo[i,k+1] = oo[i,k] + dt * rho3* ee[i,k]*(Omax-oo[i,k])- Cocell*oo[i,k]*(nn[i,k]+hh[i,k])
- gam*oo[i,k]
   }
   }
   n[,t+1]=nn[,npas+1]; h[,t+1]=hh[,npas+1]; a[,t+1]=aa[,npas+1]; m[,t+1]=mm[,npas+1];
g[,t+1]=gg[,npas+1]
   e[,t+1]=ee[,npas+1]; o[,t+1]=oo[,npas+1]
   }
   plot(c(0,L),c(-1.1,1.2))
   lines(X,n[,t+1],col="blue"); lines(X,h[,t+1],col="grey"); lines(X,a[,t+1],col="black");
lines(X,e[,t+1],col="red")
   lines(X,m[,t+1]-1,col="yellow"); lines(X,g[,t+1]-1,col="grey"); lines(X,o[,t+1]-1,col="red"
   lines(X,e[,1],col="orange")
   lines(c(0,L),c(1,1)); lines(c(0,L),c(0,0))
   }
```

7.3 Chapter 5

7.3.1 Porous media equation

The following code simulates the porous media equation, a nonlinear diffusion:

```
   Temp=20; NT0 = 500; npas=4
   NT= NT0*Temp; Nx=1000; dx=0.1; K=1; chi=0.1; int=0
   L=dx*Nx
   dt=(dx^2)/(4*K)
   Nx.virt=Nx+1
   u = matrix(nrow=Nx.virt, ncol=NT)
   v = matrix(nrow=Nx.virt, ncol=npas+1)
   w = matrix(nrow=Nx.virt, ncol=npas+1)
   z = matrix(nrow=Nx.virt, ncol=npas+1)
   X=seq(0,L,dx)
   u[,1]=exp(-0.1*abs((X-L/2)^2))
   for (t0 in 1:(NT0-1)) {
   for (t1 in 1:Temp) {
   t=(t0-1)*Temp + t1
   v[,1]=u[,t]
   for (k in 1:npas) {
   for (i in 2:Nx) {
   v[i,k+1] = v[i,k] +K*dt*((v[i+1,k]^2-2*v[i,k]^2+v[i-1,k]^2))/(dx^2)
   }
   v[1,k+1]=v[2,k+1]; v[Nx.virt,k+1]=v[Nx,k+1]
```

```
}
u[,t+1]=v[,npas+1]
}
plot(c(0,L),c(-0.5,1.5))
lines(X,u[,t+1],col="black")
lines(c(0,L),c(1,1))
lines(c(0,L),c(0,0))
}
```

7.3.2 Porous media equation with proliferation

The next code simulates the porous media equation with proliferation; we presumably see a traveling wave.

```
Temp=20; NT0 = 500; npas=4
NT= NT0*Temp; Nx=1000; dx=0.1; K=1; chi=0.1; int=0
L=dx*Nx
dt=(dx^2)/(4*K)
Nx.virt=Nx+1
u = matrix(nrow=Nx.virt, ncol=NT)
v = matrix(nrow=Nx.virt, ncol=npas+1)
w = matrix(nrow=Nx.virt, ncol=npas+1)
z = matrix(nrow=Nx.virt, ncol=npas+1)
X=seq(0,L,dx)
u[,1]=exp(-0.1*abs((X-L/2)^2))
for (t0 in 1:(NT0-1)) {
for (t1 in 1:Temp) {
t=(t0-1)*Temp + t1
v[,1]=u[,t]
for (k in 1:npas) {
for (i in 2:Nx) {
v[i,k+1] = v[i,k] +K*dt*((v[i+1,k]^2-2*v[i,k]^2+v[i-1,k]^2))/(dx^2)
}
v[1,k+1]=v[2,k+1]; v[Nx.virt,k+1]=v[Nx,k+1]
}
w[,1]=v[,npas+1]
for (k in 1:npas) {
for (i in 1:Nx.virt) {
w[i,k+1] = w[i,k] + dt * w[i,k]*(1-w[i,k])
}
}
u[,t+1]=w[,npas+1]
```

```
}
plot(c(0,L),c(-0.5,1.5))
lines(X,u[,t+1],col="black")
lines(c(0,L),c(1,1))
lines(c(0,L),c(0,0))
}
```

7.3.3 Traveling wave for porous media with proliferation

```
N=100000; X=1:N; Y=1:N; dt=0.0004; c=3; eps=0.0001
   ex=0.5*c+0.5*sqrt(c^2+4); ey=1
   X[1]=1-eps*ex; Y[1]=-eps*ey
   for (i in 1:(N-1)) {
   X[i+1]=X[i]+dt*Y[i]
   Y[i+1]=Y[i]+dt*(-c*Y[i]/X[i]-Y[i]^2/X[i]-1+X[i])
   }
   plot(c(0,1),c(-0.4,0))
   lines(X,Y)
   lines(c(0,1),c(0,0))
   lines(c(0,0),c(-0.4,0))
```

Asking `ts.plot(X)` one gets, in addition, the front.

Bibliography

[1] A. R. A. Anderson, A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion, Mathematical Medicine and Biology (2005) 22, 163–186.

[2] M. Bodnar, J. J. L. Velazquez, Derivation of macroscopic equations for individual cell-based models: a formal approach, *Math. Methods Appl. Sci.* **28** (2005), no. 15, 1757–1779.

[3] H. M. Byrne, Mathematical Biomedicine and modeling avascular tumor growth, OCCAM Preprint Number 12/96.

[4] H. Byrne, D. Drasdo, Individual-based and continuum models of growing cell populations: a comparison, J. Math. Biol. (2009) 58:657–687.

[5] V. Capasso, D. Morale, Stochastic modelling of tumour-induced angiogenesis, *J. Math. Biol.* **58** (2009), no. 1-2, 219–233.

[6] D. Drasdo, S. Hoehme, M. Block, On the role of Physics in the growth and pattern formation of multi-cellular systems: what can we learn from individual-cell based models? *J. Stat. Physics* **128** (2007), n.1/2, 287-345.

[7] Jean-Pierre Françoise, Oscillations en biologie: Analyse qualitative et modèles, Springer 2006.

[8] I.I. Gikhman, A.V. Skorohod, The Theory of Stochastic Processes I, Springer, Berlin 1974.

[9] A. Lorz, T. Lorenzi, J. Clairambault, A. Escargueil, B. Perthame, Effects of space sctructure and combination therapies on phenotypic heterogeneity and drug resistance in solid tumors, arxhiv 2013.

[10] H. P. McKean, Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov, Comm. Pure Appl. Math. XXVIII (1975), 323-331

[11] D. Morale, V. Capasso, Vincenzo, K. Oelschläger, An interacting particle system modelling aggregation behavior: from individuals to populations, *J. Math. Biol.* **50** (2005), no. 1, 49–66.

[12] K. Oelschläger, A law of large numbers for moderately interacting diffusion processes, *Z. Wahrsch. Verw. Gebiete* **69** (1985), no. 2, 279–322.

[13] K. Oelschläger, On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes, *Probab. Theory Related Fields* **82** (1989), no. 4, 565–586.

[14] K. Oelschläger, Large systems of interacting particles and the porous medium equation, *J. Differential Equations* **88** (1990), no. 2, 294–346.

[15] S. Olla, S. R. S.Varadhan, H.-T. Yau, Hydrodynamical limit for a Hamiltonian system with weak noise, *Comm. Math. Phys.* **155** (1993), no. 3, 523–560.

[16] R. Philipowski, Interacting diffusions approximating the porous medium equation and propagation of chaos, *Stochastic Process. Appl.* **117** (2007), no. 4, 526–538.

[17] R. Philipowski, Stochastic interacting particle systems and nonlinear partial differential equations from fluid mechanics, RIMS Kôkyûroku Bessatsu B6 (2008), 201–211.

[18] M. Scianna, L. Preziosi, A Hybrid Model Describing Different Morphologies of Tumor Invasion Fronts, Math. Model. Nat. Phenom. Vol. 7 (2012), No. 1, pp. 78-104.

[19] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Second Ed., Springer, New-York, 1994.

[20] H. Spohn, Large Scale Dynamics of Interacting Particles, Springer, Berlin 1991.

[21] A. Stevens, The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems, *SIAM J. Appl. Math.* **61** (2000), no. 1, 183–212.

[22] P. Hinow, P. Gerlee, L. J. McCawley, V. Quaranta, M. Ciobanu, S. Wang, J. M. Graham, B. P. Ayati, J. Claridge, K. R. Swanson, M. Loveless, A. R. A. Anderson, A spatial model of tumor-host interaction: application of chemoterapy, *Math. Biosc. Engin.* 2009.

[23] A.-S. Sznitman, Topics in propagation of chaos, Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989, Lecture Notes in Mathematics Volume 1464, 1991, pp 165-251.

[24] Youshan Tao, Jianjun Paul Tian, Mathematical Oncology, Springer, Berlin 2010.

[25] C. Tremoulet, Hydrodynamic limit for interacting Ornstein-Uhlenbeck particles, *Stochastic Process. Appl.* **102** (2002), no. 1, 139–158.

[26] S.E. Wang, P. Hinow, N. Bryce, A. M. Weaver, L. Estrada, C. L. Arteaga, G. F. Webb, A mathematical model quantifies proliferation and motility effects of TGF-β on cancer cells, *Comput. Math. Methods in Medicine*, 2013.

[27] S. R. S. Varadhan, Scaling limits for interacting diffusions, *Comm. Math. Phys.* **135** (1991), no. 2, 313–353.

[28] http://www.cancerresearchuk.org/cancer-info/cancerandresearch/all-about-cancer/what-is-cancer/grow-and-spread/growing-a-blood-supply/growing-a-blood-supply

[29] http://www.cancerresearchuk.org/about-cancer/cancers-in-general/what-is-cancer/grow/how-a-cancer-gets-its-blood-supply

[30] http://galleryhip.com/tumor-angiogenesis.html