PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à MINES ParisTech

Diffusion dans les modèles d'agrégation et de cancer

Soutenue par
Frank Ernesto ALVAREZ BORGES
Le 07 décembre 2023

École doctorale n°543
ED SDOSE

Spécialité
Sciences

Composition du jury :

M. Luis ALMEIDA CNRS, Sorbonne Université	Examinateur
M. Jean CLAIRAMBAULT INRIA de Paris	Invité
Mme. Lucilla CORRIAS Université d'Evry	Rapporteure
M. Jean DOLBEAULT CNRS, Université Paris- Dauphine PSL	Examinateur
Mme. Raluca EFTIMIE Université de Franche-Comté	Rapporteure
Mme. Evelyne MIOT CNRS, Université Grenoble-Alpes	Examinatrice
M. Stéphane MISCHLER Université Paris- Dauphine PSL	Directeur de thèse
M. Mariano RODRIGUEZ Universidad de la Habana	Co-encadrant
M. Nicolas VAUCHELET Université Sorbonne Paris Nord	Examinateur

MINES ParisTech | PSL

Publications

Published

- Alvarez, F. E., Carrillo, J. A., & Clairambault, J. (2022). Evolution of a structured cell population endowed with plasticity of traits under constraints on and between the traits. *Journal of Mathematical Biology*, 85(6-7), 64. 10.1007/s00285-022-01820-5 | hal-04079919

Submitted

- Alvarez, F. E. & Viossat, Y. (2023). Tumor containment: a more general mathematical analysis. hal-03869114 | arXiv:2211.17041

- Alvarez, F. E. & Guilberteau, J. (2023). A particle method for non-local advection-selection-mutation equations. hal-04079919 | arXiv:2304.14210

- Alvarez, F. E. & Clairambault, J. (2023). Phenotype divergence and cooperation in isogenic multicellularity and in cancer. hal-04145070 | arXiv:2306.16880

In preparation

- Alvarez, F. E. & Rodriguez-Ricard M. (2023). Propagation of strong instabilities and twinkling patters in a tumour-immune system interaction system.

- Alvarez, F. E., Carrapatoso, K. & Mischler S. (2023). The parabolic-parabolic Keller-Segel equation with variable re-scaling parameter.

- Alvarez, F. E., Carrapatoso, K. & Mischler S. (2023). On the self-similar stability of the parabolic-parabolic Keller-Segel equation.

Résumé

Cette thèse est consacrée à l'étude de plusieurs problèmes issus de la modélisation mathématique des tumeurs. Plus spécifiquement, l'intérêt principal est orienté vers les interactions ayant lieu au sein de la tumeur et avec son environnement. Néanmoins, certains des modèles et méthodes présentés au coeur de la thèse ont une portée bien plus générale que l'étude du cancer. Les principaux résultats sont divisés en cinq chapitres. Dans le premier chapitre, par une nouvelle analyse mathématique comparant la taille des tumeurs entre traitements non pas en fonction du temps, mais en fonction de la taille de la population résistante, nous établissons une comparaison entre les résultats de différentes stratégies de traitement appliquées à une tumeur composée de deux sous-populations, une de cellules sensibles et une autre de cellules résistantes. Dans le deuxième chapitre, nous dérivons l'expression asymptotique d'un cycle limite apparaissant dans un modèle d'interaction tumeur-système immunitaire. Le troisième chapitre est consacré à la modélisation du bet-hedging, une stratégie évolutive d'intérêt pour la théorie atavique du cancer. L'existence et le caractère unique de la solution du modèle sont prouvés et deux phénomènes d'intérêt biologique sont mis en évidence par des simulations. Le chapitre quatre est un complément au troisième chapitre. On y développe une discussion philosophique sur la théorie atavique du cancer et on esquisse deux modèles différents pour l'émergence de la coopération. Le chapitre cinq concerne l'étude d'une méthode particulaire pour un modèle d'advection-réaction-diffusion non local d'une grande importance dans le domaine de les dynamiques adaptatives. La conservation du comportement asymptotique est analysée pour le schéma numérique proposé. Les chapitres six et sept sont consacrés à l'étude du système de Keller-Segel parabolique-parabolique où nous donnons respectivement quelques estimations de la solution et déterminons le comportement asymptotique pour le cas non radial

Mots clés: Populations structurées; modélisation du cancer; confinement des tumeurs; méthodes particulaires; théorie atavique; système de Keller-Segel parabolique-parabolique.

Abstract

This thesis is devoted to the study of several problems arising from the mathematical modelling of tumours. More specifically, the main interest is oriented towards the interactions taking place within the tumour and with its environment. Nevertheless, some of the models and methods presented at the core of the thesis have a much more general scope than the study of cancer. The main results are divided in five chapters. In the first chapter, by a novel mathematical analysis comparing tumor sizes across treatments not as a function of time, but as a function of the resistant population size, we establish a comparison between the outcomes of different treatment strategies applied to a tumour composed of two sub-populations, one of sensitive cells and another one of resistant cells. In the second chapter, we derive the asymptotic expression of a limit cycle arising in a tumour-immune system interaction model. The third chapter is devoted to the modeling of bet-hedging, an evolutionary strategy of interest for the atavistic theory of cancer. The existence and uniqueness of solution for the model is proved and two phenomena of biological interest are evidenced through simulations. Chapter four is a complement for the third chapter. On it, a philosophical discussion about the atavistic theory of cancer is developed and two different models for the emergence of cooperation are sketched. Chapter five is concerned with the study of a particle method for non-local advection-reaction-diffusion model of great importance in the area of adaptive dynamics. The conservation of asymptotic behaviour is analyzed for the proposed numerical scheme. Chapters six and seven are devoted to the study of the fully parabolic Keller-Segel system where we give some estimates over the solution and determine the asymptotic behaviour for the non-radial case, respectively

Keywords: Structured populations; cancer modelling; tumour containment; particle methods; atavistic theory; fully parabolic Keller-Segel system.

“Aut viam inveniam
aut faciam”

Contents

Introduction 8

1 Tumor containment: a more general mathematical analysis 34

1.1 Introduction 34
1.2 Model 36
1.3 Results 40
1.4 Proofs 43
1.4.1 Key lemmata 43
1.4.2 Proof of propositions 1.1 to 1.8 45
1.5 Discussion 49
1.6 Appendices 50
1.6.1 Mutations from sensitive to resistant cells 50
1.6.2 Comparison principles 53

2 Asymptotic expansion of a limit cycle arising from a tumour-immune system interaction model 55

2.1 Introduction 55
2.2 The model 56
2.3 Derivation of the expression for the limit cycle 58
2.3.1 Algorithm for obtaining the limit cycle 58
2.3.2 Expression for the limit cycle 59
2.4 Discussion and perspectives 64

3 Evolution of a structured cell population endowed with plasticity of traits under constraints on and between the traits 65

3.1 Introduction: biological background 65
3.2 The model 67
3.3 Existence of a weak solution and numerical approximation 71
3.3.1 Preliminaries on the finite volume method 71

3.3.2 Global existence, uniqueness, positivity and boundedness of the solution for the semi-discrete scheme . . . 74
3.3.3 Discrete gradient, L^2 norm estimate and compactness result . . . 75
3.3.4 Existence of weak solution . . . 85
3.3.5 A discrete implicit scheme . . . 91
3.4 Simulations . . . 94
3.4.1 Phenotypic Dimorphism . . . 96
3.5 Concluding remarks . . . 103

4 Phenotype divergence and cooperation in isogenic multicellularity and in cancer 106
4.1 Biological and evolutionary-developmental background . . . 106
4.1.1 Being or not teleological: the two settings considered . . . 106
4.1.2 The atavistic theory of cancer . . . 107
4.1.3 Why and how does multicellularity fail in cancer? . . . 108
4.1.4 A narrative of long-term evolution and cancer, freely exposed to the fire of philosophy of science . . . 108
4.2 Cell differentiation and phenotype divergence . . . 109
4.2.1 Heterogeneity and plasticity with respect to what? . . . 109
4.2.2 Long-term evolution as genetic adaptation of the body plan in animals . . . 109
4.2.3 A nonlocal phenotype-structured cell population model . . . 110
4.2.4 What this model tackles and what it leaves unexplained . . . 112
4.3 Cooperation . . . 113
4.3.1 Tinkered cooperation in the emergence of multicellularity vs. directed cooperation in constituted multicellular animals . . . 113
4.3.2 Prisoner's dilemma and reciprocity . . . 114
4.3.3 A continuously structured population model for the evolution of cooperation . . 118
4.4 Conclusion . . . 121

5 A particle method for non-local advection-selection-mutation equations 123
5.1 Introduction . . . 123
5.2 The problem . . . 127
5.2.1 Some bounds over the characteristics . . . 129
5.2.2 Existence of solution for smooth initial data . . . 131
5.2.3 Existence of solution for more general initial data . . . 135
5.3 Particle Method . . . 140
5.4 Convergence of the numerical solution towards a weak solution . . . 146

5.4.1 Convergence on a finite time interval 147
5.4.2 Asymptotic preserving properties 151
5.5 Simulations 163
5.6 Appendices 166
5.6.1 Proof of the results over the characteristics 166
5.6.2 Existence of solution for a system of ODEs with infinitely many unknowns and equations 170
5.6.3 A result from approximation theory 171
5.6.4 Proofs of convergence results from ODE theory 172

6 The parabolic-parabolic Keller-Segel equation with variable rescaling parameter 173

6.1 The parabolic-parabolic Keller-Segel equation and re-scaling parameters 173
6.1.1 The PPKS equation and the re-scaling parameters 173
6.1.2 Self-similar profiles 175
6.1.3 The linearized self-similar PPKS equation 175
6.1.4 Functional spaces and dissipativity estimate 177
6.1.5 Weyl's type theorem 178
6.2 Estimates on the self-similar profile 178
6.3 Some functional inequalities 183
6.4 Estimates on the first equation 192
6.4.1 Step 1. The principal term of classical Fokker-Planck type. 192
6.4.2 Step 2. The remainder term. 193
6.5 Estimates on the second equation 194
6.5.1 Step 1. The principal part in L^2. 194
6.5.2 The principal part for other norms. 195
6.5.3 Step 2. The remainder part in L^2 and proof of Theorem 6.1 196
6.6 Weyl's type result on the principal spectrum 198

7 On the self-similar stability of the parabolic-parabolic Keller-Segel equation 200

7.1 Introduction 200
7.2 Estimates over Q and P 203
7.3 Functional inequalities 204
7.4 Estimates for $\mathscr{L}_{1,1}$ 206
7.4.1 Dissipativity estimates related to $\mathscr{L}_{1,1}$. 206
7.4.2 Splitting of the operator $\mathscr{L}_{1,1}$ 207

7.4.3 Spectral analysis of $\mathscr{L}_{1,1}$ . . . 209
7.4.4 Semigroup decay estimates for $\mathscr{L}_{1,1}$ . . . 211
7.5 Estimates for $\mathscr{L}_{2,2}$ . . . 214
7.6 Semigroup estimates for the linearized system . . . 216
7.7 Stability for the nonlinear equation . . . 219

Introduction

This thesis is devoted to the study of several problems arising from the mathematical modelling of tumours. More specifically, the main interest is oriented towards the interactions taking place within the tumour and with its environment. Nevertheless, it is worth mentioning that some of the models and methods presented at the core of the thesis have a much more general scope than the study of cancer.
We first present some of the mathematical tools used in the modelling of population dynamics. We then expose different approaches for the study of tumour cells, and their evolution.
Finally, the main results obtained throughout the thesis will be summarized.

Some tools from differential equations theory for the mathematical modelling of population dynamics

Simple species growth model

The simplest model of population dynamics only takes one variable into account: the population size at time $t > 0$, usually denoted as $N(t)$. Its growth is described by means of an ordinary differential equation with an initial condition

$$\frac{dN(t)}{dt} = G(N(t)), \quad N(0) = N_0,$$

where $G(N)$ is a regular function and $N_0 \geqslant 0$. Some examples are the linear growth (if $G(N)$ is constant), exponential growth (if $G(N)$ is proportional to N), logistic growth $(G(N) := rN(K - N)))$ or gompertz growth $(G(N) := rN \ln(K/N))$, where r and K are the intrinsic growth rate of the population and the carrying capacity of the environment, respectively.
If the population is homogeneous, and surrounded by a rather stable environment (which are two very strong assumptions), simple species growth models are often enough to describe the population dynamics. However, once heterogeneity within the population and a variable environment come into play, the simplicity of these models does not allow for a sufficiently good representation of the population evolution.

Multiple species growth model

Heterogeneity inside the population often allows for a clear classification of its individuals in a finite amount of categories. Gender is the simplest example of such categories, but we could mention as

well resistance, or lack of it, towards a specific sickness or stimulation, or even dead or alive stages (for tumour cell populations, for example). Furthermore, entirely different species might be considered altogether.
These, and others, are all categories which directly affect the population dynamics. Therefore, is of interest to track the evolution of each sub-population independently. A system of ordinary differential equations is then very well suited for its mathematical representation. Consider $n \in \mathbb{N}$ sub-populations, and $N_i(t)$, $i \in \{1, \ldots, n\}$ their sizes at time $t > 0$. We denote $N(t) := (N_1(t), \cdots, N_n(t))$. The associated dynamic system then reads

$$\frac{dN(t)}{dt} = G(N(t)), \quad N(0) = N_0,$$

with $G : \mathbb{R}^n \to \mathbb{R}^n$ a vector valued regular function and $N_0 \in \mathbb{R}^n_+$.
The study of several sub-populations allows for the inclusion of various mechanisms such as competition, cooperation or mutations, that were not possible to describe with the simple species growth model.
Amid the most well known models we have Lotka-Volterra models, for predator-prey interactions, and SIR models for the spread of infectious diseases within a population. We present two models relying on a system of two differential equations, and derive basic results that will be further developed in this thesis.

Sensitive and resistant sub-populations on a tumour

Clinical and pre-clinical data suggests that treating some tumors at a mild, patient-specific dose might delay resistance to treatment and increase survival time. A recent mathematical model with sensitive and resistant tumor cells identified conditions under which a treatment aiming at tumor containment rather than eradication is indeed optimal. The equations associated to such model are

$$\begin{cases} \dfrac{dS}{dt}(t) = S(t)g_S(S(t), R(t), L(t)), \\ \dfrac{dR}{dt}(t) = R(t)g_R(S(t), R(t)), \\ \ S(0) = S_0 \geqslant 0, \ R(0) = R_0 \geqslant 0, \end{cases} \tag{1}$$

where $S(t)$ and $R(t)$ are the total number of sensitive and resistant cells at time t, $L(t)$ is the current dose or treatment level, and g_S and g_r are per-cell growth-rate functions. The total tumor population size is $N(t) = S(t) + R(t)$, with initial value $N_0 = S_0 + R_0$. We also fix the maximum tolerable level for the total tumor size N_{tol} and the critical size N_{crit} over which the patient cannot survive.
We set the following key assumptions

(a1) Resistant cells are fully resistant ($\partial_L g_r = 0$).

(a2) The function g_r is strictly decreasing in S ($\partial_S g_r < 0$).

(a3) The function g_s is non-increasing in R and strictly decreasing in L ($\partial_R g_s \leqslant 0$ and $\partial_L g_s < 0$).

(a4) As long as the patient is alive ($N < N_{crit}$), the size of an untreated or fully resistant tumor strictly increases ($g_r > 0$).

(a5) For any treatment level L, the function

$$R \to (N_{tol} - R)g_s(N_{tol} - R, R, L) + Rg_r(N_{tol} - R, R),$$

is increasing on $[0, N_{tol}]$.

Assumptions (a1) and (a2) imply that the only external factor influencing the evolution of the resistant cells is the competition with the sensitive cells. Conversely, assumption (a3) ensures that the sensitive population is affected by both the competition with resistant cells and the treatment. Assumption (a4) reflects the fact that the tumour is not curable and finally the last assumption ensures that if a tumor increases beyond the maximum tolerable level N_{tol} then it will never become smaller again. It also implies that the treatment level required to stabilize a tumor of size N_{tol} increases with the frequency of resistant cells.

A natural question is how outcomes related to different treatment levels compare. Consider two different treatment levels $L_1(t)$ and $L_2(t)$ and their respective associated solutions for problem (1), denoted as $(S_1(t), R_1(t))$ and $(S_2(t), R_2(t))$. In what follows, we assume that $S_1(0) \geqslant S_2(0)$ and $R_1(0) \leqslant R_2(0)$. If over a certain interval $[0, T]$ the relation $S_1(t) \geqslant S_2(t)$ holds true, the decreasing character of g_r with respect to S implies that

$$\frac{dR_1}{dt}(t) = R_1(t)g_r(S_1(t), R_1(t)) \leqslant R_1(t)g_r(S_2(t), R_1(t)),$$

which implies that $R_1(t)$ is a sub-solution of the equation satisfied by $R_2(t)$. The comparison principle allows us to conclude then that $R_1(t) \leqslant R_2(t)$ for all $t \in [0, T]$.

Similarly, under the relation $N_1(t) \geqslant N_2(t)$ we have that

$$\frac{dR_1}{dt}(t) = R_1(t)g_r(N_1(t) - R_1(t), R_1(t)) \leqslant R_1(t)g_r(N_2(t) - R_1(t), R_1(t)),$$

which again implies that $R_1(t) \leqslant R_2(t)$ for all $t \in [0, T]$ and additionally

$$S_1(t) = N_1(t) - R_1(t) \geqslant N_2(t) - R_2(t) = S_2(t).$$

In other words, as long as a treatment level keeps either the sensitive or the total population above the one related to a second treatment level, it will keep the resistant population below its counterpart.

Finally, if $L_1(t) \leqslant L_2(t)$ and the initial relations over $(S_1(0), R_1(0))$ and $(S_2(0), R_2(0))$ hold strictly, the previous argument ensures that $R_1(t) < R_2(t)$ and $S_1(t) > S_2(t)$ for all $t \in [0, T]$. Then, the continuity of the solutions with respect to the initial data guarantees that, when the initial relations are not strictly satisfied, then $R_1(t) \leqslant R_2(t)$ and $S_1(t) \geqslant S_2(t)$ for all $t \in [0, T]$, by using a perturbation argument.

The previous results, allow us to conclude that not treating will always minimize the amount of resistant cells while maximizing the sensitive ones. On the other hand, treating at the maximum tolerated dose will have the opposite effect, which is not desirable, as a fully resistant tumour does not leave many options for treatments. Furthermore, in [1] it was shown that containing the tumor at N_{tol} maximizes the survival time of the patient, as this strategy allows to keep the total population below any other outcome obtainable through any alternative treatment.

The literature regarding cancer containment includes Martin et al. (1992) [2], Monro-Gaffney (2009) [3], Gatenby et al. (2009) [4], Silva et al. (2012) [5], Carrère (2017) [6], Zhang et al. (2017) [7], Bacevic-Noble et al. (2017) [8], Hansen et al. (2017) [9], Gallaher et al. (2018) [10], Cunningham et al. (2018) [11], Pouchol (2018) [12], Carrère and Zidani (2020) [13], Strobl et al. (2020) [14], Cunningham et al. (2020) [15] and Viossat and Noble (2021) [1].

Periodic dynamics and Hopf bifurcation

Following the model presented in [16] and [17], let us now consider a system of two interacting and competing sub-populations with a different setting to the one presented on the previous section. The first one will be composed of tumour cells and the second one will be conformed by lymphocytes. The quantities $c(t)$ and $n(t)$ will represent the densities of tumour cell and lymphocytes at time t, respectively. The dynamics of these two populations are described by the following system

$$\begin{cases} \dfrac{d}{dt}c(t) = a_1 cF(c) - a_2\mu\phi(c)cn, \\ \dfrac{d}{dt}n(t) = -a_3 n\psi(c) + a_4 q(c), \end{cases} \tag{2}$$

where the function $\psi(c)$ describes the stimulatory effect of the tumour cells on the immune cells. It can be assumed that this function is positive (at least initially), $\psi(0) > 0$, and might be negative only in a finite interval. It is reasonable to assume $|\psi'(0)| \leqslant 1$, so that, at least initially, the death rate of lymphocytes is not greater than that in the linear model. The tumour growth rate $F(c)$ is a positive function which summarizes the carrying capacity (or malignancy) such that $F(0) > 0$, $F'(c) \leqslant 0$ and $\lim_{c\to\infty} cF(c) = 0$, with the additional assumption that initially it is $F'(0) = 0$. The loss of tumour cells, which depends on the competition with lymphocytes, is represented by the function $\phi(c)$ characterized by $\phi(c) > 0$, $\phi'(c) \leqslant 0$ and $\lim_{c\to\infty} c\phi(c) = l < \infty$. In other words, if the tumour growth tends to infinity the loss of tumour cells would tend to a constant rate. It can be also assumed that $\phi'(0) = 0$. Regarding the influx of immune cells $q(c)$ can be taken $q(0) = 1$, $|q'(0)| \leqslant 1$, so that, at least initially, the influx of effector cells is not greater than that in the linear model.
Through the change of variables

$$u = c, \qquad v = \frac{n}{a_4}, \qquad \tau = a_3 t,$$

and introducing the constants

$$a = \frac{a_1}{a_3}, \qquad b = \frac{1}{a_3}, \qquad \mu = \frac{a_2 a_4}{a_3},$$

we get the non-dimensional model

$$\begin{cases} \partial_t u = auF(u) - \mu\phi(u)uv, \\ \partial_t v = -v\psi(u) + bq(u). \end{cases}$$

Furthermore, the reaction term can be approximated using a second order Taylor expansion around the steady state $(0, b/\psi(0))$, as done in [16], where, after assuming $F'(0) = 0$, $\phi(0) = 1$ and $q''(0) = 0$, and grouping similar terms, we obtain the system

$$\begin{cases} \partial_t u = \alpha u - \mu uv, \\ \partial_t v = -\beta_1 uv - \beta_2 v + \beta_3 u + \beta_4 - \beta_5 vu^2, \end{cases} \tag{3}$$

where $\alpha = aF(0)$, $\beta_1 = \psi'(0)$, $\beta_2 = \psi(0)$, $\beta_3 = bq'(0)$, $\beta_4 = b$ and $\beta_5 = \frac{1}{2}\psi''(0)$. Hence, the following restrictions apply to the parameters:

$$\alpha > 0, \;\; \mu > 0, \;\; |\beta_1| \leqslant 1, \;\; \beta_2 > 0 \text{ and } |\beta_3| \leqslant \beta_4. \tag{4}$$

On the model around resistant and sensitive tumor cells, one of the sub-populations was assumed to be always increasing. This rules out the existence of a periodic solution for system (1). However, this is not the case for model (3), where the conditions for the appearance of such periodic solutions were already given in [16]. In fact, this is not a rare phenomenon in population dynamics. In particular, for predator-prey models (or homologous ones) one can expect that the growth of the prey population will imply an increase on the number of predators. In turns, this leads to the decay of the number of preys and consequently, the amount of predators also dwindles, allowing once again for the prey population to proliferate and triggering the whole process all over. This simple mechanism can be described through the following system of ordinary differential equations

$$\begin{cases} \dfrac{dx(t)}{dt} = a - y(t), \\ \dfrac{dy(t)}{dt} = x(t) - b, \end{cases}$$

where $x(t)$ and $y(t)$ represent the number of prey and predators respectively while a and b are the intrinsic growth factors of each sub-population in the absence of the other one. A simple computation shows that

$$\frac{d}{dt}\left((x(t)-b)^2 + (y(t)-a)^2\right) = 0,$$

for all values of t, which imply that the orbits for the dynamical system are non other than the circumferences

$$(x(t)-b)^2 + (y(t)-a)^2 = r, \quad r \in \mathbb{R}^+.$$

Determining the existence of periodic orbits (and their stability) is generally a difficult problem, which strongly depends on the parameters of the model. This type of studies usually falls under the theory of bifurcations.

A bifurcation occurs when a small smooth change made to the parameter values (the bifurcation parameters) of a system causes a sudden "qualitative" or topological change in its behavior. In particular, we are concerned with a Hopf bifurcation, which is a critical point where, as the parameter changes, the system's stability switches and a periodic solution arises. More accurately, it is a bifurcation in which a steady state of a dynamical system loses stability, as a pair of complex conjugate eigenvalues (of the linearization around the steady state) crosses the complex plane imaginary axis as the parameter crosses a threshold value.

We show with a simple example the conditions for the appearance of a Hopf bifurcation. Due to its simplicity and illustrative value, we consider the Liénard system, which is used while modelling oscillating circuits rather than population dynamics,

$$\begin{cases} \dfrac{du(t)}{dt} = f(u(t), v(t)) := v(t), \\ \dfrac{dv(t)}{dt} = g(u(t), v(t)) := -u(t) + (\mu - u^2(t))v(t). \end{cases}$$

Its only equilibrium point is the origin and the Jacobian matrix for the linearized system about the origin is

$$J_\mu = \begin{pmatrix} 0 & 1 \\ -1 & \mu \end{pmatrix}.$$

In order to identify the occurrence of Hopf bifurcation with respect to μ, the following conditions need to be fulfilled:

(b1) $tr(J_\mu) = 0$, $tr(J_\mu)$ is the trace of J_μ.

(b2) $\det(J_\mu) > 0$.

(b3) The derivative of $tr(J_\mu)$ with respect to μ, has to be different from 0.

(b4) The first Lyapunov coefficient l_1 has to be different from 0 where

$$\begin{aligned} l_1 = &\frac{1}{16}(f_{uuu} + g_{uuv} + f_{uvv} + g_{vvv}) \\ &+ \frac{1}{16\det(J_\mu)}(f_{xy}(f_{uu} + f_{vv})) - g_{uv}(g_{uu} + g_{vv}) - f_{uu}g_{uu} + f_{vv}g_{vv}), \end{aligned}$$

and

$$h_{u^i v^{k-i}} := \frac{\partial^k}{\partial u^i \partial v^{k-i}} h(0,0), \quad h \in \{f, g\}, k \in \{2, 3\}, 0 \leqslant i \leqslant k.$$

Condition (b1) and (b2) are necessary conditions related to the presence of imaginary eigenvalues which cross the real axis of the complex plane. Conditions (b3) and (b4) are the transversality and non-degeneracy conditions, respectively, which ensures the equivalence of the system to one of the normal forms of the Hopf bifurcation.
Regarding the Liénard system, $tr(J_\mu) = \mu$, which is 0 if $\mu = 0$. On the other hand

$$\det(J_\mu) = \frac{d}{d\mu} tr(J_\mu) = 1 > 0,$$

while $l_1 = -\frac{1}{8} \neq 0$. Hence, all the conditions of the Hopf Bifurcation Theorem are satisfied. Since the origin is stable for $\mu < 0$ and unstable for $\mu > 0$, the system has a supercritical Hopf bifurcation at $\mu = 0$.
Due to the various parameters present in (3), a similar study becomes harder to perform. In such cases, an alternative way of determining the appearance of a limit cycle is the Poincaré-Bendixon theorem. This theorem states that if an invariant subset of the phase portrait does not contain critical points, then there exists a periodic orbit. Indeed, this result was used in [16] to show that under certain conditions, a limit cycle appears for system (3). Amid the biological implications of this result, we have the fact that for a certain family of parameters, it is possible for the immune system to keep a tumor under control, given that a periodic solution, is always bounded. Of course, it remains to support these theoretical results with biological and medical evidence.
Amid the most recent literature regarding the study of limit cycles we have Mittal et al. (2020) [18], Li-Li (2023) [19], Bai et al. (2023) [20] and Zhang-Shateyi (2023) [21].

Continuous structure variables

A finite amount of classifications for the individuals is not always possible, or convenient. It is often useful to relate the elements of a population with a continuously quantifiable property, such as position, age, size or concentration of a certain chemical.
The quantity describing such property may be represented by $x \in \Omega \subset \mathbb{R}^d$ and is referred as a continuous structure variable.
Position in space is, perhaps, the most widely known among the continuous structure variables, and is of special interest in the study of cell populations. In such scenario, $n(t, x)$ denotes the density of

population at time $t > 0$ and point $x \in \mathbb{R}^d$, $d \in \{1, 2, 3\}$. The total population size $\rho(t)$ can be then computed as

$$\rho(t) := \int_\Omega n(t, x) dx,$$

and the dynamics of $n(t, x)$ can be described by the partial differential equation

$$\partial_t n(t, x) + \nabla \cdot (a(t, x, n(t, x))n(t, x)) + D(t, x, n(t, x)) = R(t, x, n(t, x)), \quad n(0, x) = n_0(x), \tag{5}$$

where $a(t, x, n)$ is a regular, vector valued function, $D(t, x, n(t, x))$ is a diffusion operator, usually an integral or a second order differential operator, and $R(t, x, n)$ is a real valued regular function, often called reaction term. In this context the divergence term usually represents transport in a given direction, the diffusion term represents invasion and the reaction term describes the selection mechanisms, which include those mentioned for the single and multiple species growth models. The inclusion of space as a variable allows to track not only the size, but the geometry and/or spatial distribution of the population, which sometimes hold important information.

Some of the non spatial structure variables are linked to a phenotype. In such cases, the interpretation of the terms in equation (5) is slightly different. The reaction term still describes the selection mechanisms, however the diffusion term stands for non-genetic epimutations, while the advection term models mutations that are being prompted by the environment.

An important aspect to consider is the fact that a, D and R could depend on $n(t, x)$ non-locally. This is, using global information of the function n over the domain Ω. The non-locality of such functions could be evidenced through a dependence on $\rho(t)$ or a more general integral operator

$$I_n(t, x) := \int_\Omega \varphi(t, x, y) n(t, y) dy,$$

for some known function $\varphi(t, x, y)$.

We present a simple example extracted from [22] where it is evidenced the concentration phenomena arising from these non-local problems. Consider the equation

$$\begin{cases} \partial_t n(t, x) = (r(x) - \rho(t))n(t, x), & x \in I \subset \mathbb{R}, t > 0, \\ n(0, x) = n_0(x), & x \in I, \end{cases} \tag{6}$$

where $r(x) > 0$ is the reproduction rate of the elements of the population with trait x, I is a bounded interval and $n_0(x)$ is a bounded and strictly positive function over I.

Notice that equation (6) has infinitely many steady states (in the sense of measures) given by the family of Dirac deltas $r(y)\delta(x - y)$, $y \in I$. However, a deeper analysis is needed to determine the stability for each of these solutions.

The function

$$N(t, x) := n(t, x) e^{\int_0^t \rho(s) ds}$$

satisfies the equation

$$\begin{cases} \partial_t N(t, x) = r(x) N(t, x), & x \in I \subset \mathbb{R}, t > 0, \\ N(0, x) = n_0(x), & x \in I, \end{cases}$$

hence, it is given explicitly by the expression

$$N(t, x) = n_0(x) e^{r(x)t}.$$

Therefore

$$\frac{d}{dt}e^{\int_0^t \rho(s)ds} = \rho(t)e^{\int_0^t \rho(s)ds} = \int_I N(t,x)dx = \int_I n_0(x)e^{r(x)t}dx.$$

Integrating in t we obtain

$$e^{\int_0^t \rho(s)ds} = \int_I \frac{n_0(x)}{r(x)}e^{r(x)t}dx + 1 - \int_I \frac{n_0(x)}{r(x)}dx,$$

which allows to give an explicit expression for $\rho(t)$

$$\rho(t) = \int_I n_0(x)e^{r(x)t}dx \left(\int_I \frac{n_0(x)}{r(x)}e^{r(x)t}dx + 1 - \int_I \frac{n_0(x)}{r(x)}dx\right)^{-1}.$$

Under the assumption that $r(x)$ attains its maximum at a single value $\bar{x}$, it is possible to compute the limit of $\rho(t)$ when t goes to infinity. Indeed, it is straightforward to obtain the bound

$$\rho(t) \leqslant r(\bar{x}) \int_I \frac{n_0(x)}{r(x)}e^{r(x)t}dx \left(\int_I \frac{n_0(x)}{r(x)}e^{r(x)t}dx + 1 - \int_I \frac{n_0(x)}{r(x)}dx\right)^{-1}.$$

The right-hand side of this relation converges towards $r(\bar{x})$, since $\int_I \frac{n_0(x)}{r(x)}e^{r(x)t}dx$ converges towards $+\infty$.
On the other hand, for any $\varepsilon > 0$, we define the set

$$I_\varepsilon := \{x \in I : r(x) \geqslant r(\bar{x}) - \varepsilon\}.$$

hence

$$\rho(t) \geqslant (r(\bar{x}) - \varepsilon) \int_{I_\varepsilon} \frac{n_0(x)}{r(x)}e^{r(x)t}dx \left(\int_I \frac{n_0(x)}{r(x)}e^{r(x)t}dx + 1 - \int_I \frac{n_0(x)}{r(x)}dx\right)^{-1}.$$

. The right-hand side of this relation converges towards $r(\bar{x}) - \varepsilon$, for all $\varepsilon > 0$, which implies that the limit of $\rho(t)$ is precisely $r(\bar{x})$.
Finally, recalling that

$$n(t,x) = n_0(x)e^{\int_0^t (r(x)-\rho(s))ds},$$

we see that for all $x \neq \bar{x}$, $n(t,x)$ converges to 0, which implies that the only stable steady state is $r(\bar{x})\delta(x-\bar{x})$.
The selection of the trait which maximizes the reproductive rate makes perfect sense from a biological perspective, and validates the modelling choice. It is important to remark the fact that, when second order differential operators are considered in order to model mutations, the regularizing effect of such operators does not allow for the formation of Dirac masses as a steady state. However, gaussian-like functions can still be observed, and it has been proved in [23] that in some cases, the steady states for problem (6) are the limit of similar problems where small mutations are considered.
The study of the dynamics of a structured population is a subject that has gained a lot of attention in recent years. Some of the references regarding this subject include Desvillettes et al. (2008) [24], Bouin et al. (2012) [25], Bouin-Calvez (2014) [26], Chisholm et al. (2016) [27], Pouchol et al. (2018) [12], Pouchol-Trélat (2018) [28], Guilberteau et al. (2023) [29] and Guilberteau et al. (2023) [30].

Angiogenesis and aggregation in tumours

Angiogenesis is a process whereby capillary sprouts are formed in response to a chemical stimuli which is externally supplied. This process occurs during embryogenesis, wound healing, arthritis and during the growth of solid tumors. Several mathematical models have been proposed and studied in order to represent angiogenesis, or more generally, the aggregation and chemo-attractant processes that governs it (see for example Anderson and Chaplain (1998) [31], Corrias et al. (2003) [32], Corrias et al. (2004) [33] and Perthame-Vasseur (2012) [34]).
One of the most well known models for chemotaxis is the Keller-Segel (or Patlak-Keller-Segel) system. A simplified version of this system is

$$\begin{cases} \partial_t f = \Delta f - \nabla \cdot (f \nabla u), \\ \varepsilon \partial_t u = \Delta u + f, & t > 0, x \in \mathbb{R}^d, \\ f(0,x) = f_0(x), \quad u(0,x) = u_0(x), \end{cases} \tag{7}$$

where $n(t,x)$ could represent the density of endothelial cells and $c(t,x)$ the concentration of chemo-attractant.
The global existence of solution for system (7), and its asymptotic behaviour is a question of interest that strongly depends on the parameters ε, d and $M := \int_{\mathbb{R}^d} f_0(x)dx$.
For example, consider the parabolic-elliptic Keller-Segel equation on the plane ($\varepsilon = 0$, $d = 2$). A formal computation shows that

$$\frac{d}{dt} \int_{\mathbb{R}^2} f(t,x)dx = 0,$$

hence, the mass M is conserved:

$$\int_{\mathbb{R}^2} f(t,x)dx = \int_{\mathbb{R}^2} f_0(x)dx = M, \quad \forall t > 0.$$

A similar computation shows that

$$\frac{d}{dt} \int_{\mathbb{R}^2} f(t,x)|x|^2 dx = 4M - \frac{M^2}{2\pi}.$$

This relation directly implies that, if $M > 8\pi$, then no global in time solutions can exist, because the second order moment of a positive solution would become negative after a finite amount of time, which is not possible. This is way the study of the parabolic-elliptic Keller-Segel equation on the plane is usually divided in three cases: the sub-critical case ($M < 8\pi$) where global solutions in time exist and converge towards a regular profile when t goes to $+\infty$ (see Blanchet et al. (2006) [35], Campos, Dolbeault (2014) [36] and Egaña-Mischler (2016) [37]), the critical case ($M = 8\pi$) where global solutions in time exist, but blow up when t goes to $+\infty$ (see Blanchet et al. (2008) [38], Ghoul-Masmoudi (2018) [39] and Davila et al. (2023) [40]) and the super-critical case ($M > 8\pi$) where solutions blow up in finite time (see Herrero-Velázquez (1997) [41] and Raphaël-Schweyer (2014) [42]).
The study of the fully parabolic Keller-Segel system is a much more complex problem, however, great advances have been accomplished as well. For example, important results regarding existence of solutions and their behaviour when $d > 2$ can be found in Corrias et al. (2004) [33], Corrias-Perthame (2006) [43], Corrias-Perthame (2008) [44] and Calvez et al. (2012) [45]. On the other hand, when $d = 2$, some advances have been made, (see, for example, Calvez-Corrias (2008) [46], Biler et al. (2011) [47] and Carrapatoso-Mischler (2017) [48]) but many questions remain open.

Two numerical methods for non-local partial differential equations

Numerical approximation plays a fundamental roll while understanding the solutions for partial differential equations. Not only serves as a way to illustrate qualitative and quantitative properties, but paired with functional analysis it provides a whole set of tools for the theoretical study of such solutions. We briefly describe the basis of two numerical methods which are well suited for the study of non-local integro-differential equations.

Finite volume method

The finite volume method is a discretization method which is well adapted for heat or mass transfer problems, among others. It may be used on arbitrary geometries and using structured or unstructured meshes. Furthermore, it locally conserves the numerical fluxes.
To apply this method, the equation is integrated on each discretization cell (which is often called "control volume"). Then, by the divergence formula, an integral formulation of the fluxes over the boundary of the control volume is given. Finally, the fluxes on the boundary are discretized with respect to the discrete unknowns.
Let us use the transport equation to showcase the method. Consider the equation

$$\begin{cases} \partial_t n(t,x) + \nabla \cdot (a(x)n(t,x)) = 0, & x \in \mathbb{R}^2, t > 0, \\ n(0,x) = n_0(x), & x \in \mathbb{R}^2. \end{cases}$$

where $a \in \mathscr{C}^1(\mathbb{R}^2, \mathbb{R}^2)$ and $n_0 \in L^\infty(\mathbb{R}^2)$. Let $\mathscr{M}$ be a mesh of $\mathbb{R}^2$ consisting of polygonal bounded convex subsets of $\mathbb{R}^2$ and let D_j be an element of the mesh $\mathscr{M}$. Integrating the transport equation over D_j yields

$$\rho_j'(t) + \int_{\partial D_j} a(x) \cdot \mathbf{n}_j(x) n(t,x) dS(x) = 0, \tag{8}$$

where

$$\rho_j(t) := \int_{D_j} n(t,x)dx,$$

is the discretized unknown, $\mathbf{n}_j(x)$ is the outward pointing normal vector to ∂D_j and dS denotes the one-dimensional Lebesgue measure on ∂D_j . The integral term in (8) may then be split as

$$\int_{\partial D_j} a(x) \cdot \mathbf{n}_j(x) n(t,x) dS(x) = \sum_{l \in N_j} \int_{\Gamma_{jl}} a(x) \cdot \mathbf{n}_{jl}(x) n(t,x) dS(x), \tag{9}$$

where N_j is the set of indexes such that D_l is a neighbor of D_j, Γ_{jl} is the boundary between D_j and D_l and $\mathbf{n}_{jl}(x)$ is the normal vector to Γ_{jl} pointing from D_j to D_l. Let

$$a_{jl} := |\Gamma_{jl}| a(\mathbf{x_{jl}}) \cdot \mathbf{n}_{jl}(\mathbf{x_{jl}}),$$

where $\mathbf{x_{jl}}$ is the center of mass of Γ_{jl}.By denoting as a_{jl}^+ and a_{jl}^- the positive and negative part of a_{jl} respectively, each term of the sum in the right-hand-side of (9) is then discretized as

$$F_{jl}(t) = a_{jl}^+ \rho_k(t) + a_{jl}^- \rho_l(t).$$

This "upwind" choice is classical for transport equations. It is crucial in the mathematical analysis; it ensures the stability properties of the finite volume scheme.
Therefore, we have derived the semi-discrete scheme

$$\begin{cases} \partial_t \rho_j(t) + \sum\limits_{l \in N_j} F_{jl}(t) = 0, & j \in \mathbb{N}, t > 0, \\ \\ \rho_j(0) = \int_{D_j} n_0(x) dx, & j \in \mathbb{N}, \end{cases}$$

which can be studied using classical ODE theory or further discretized in time.
This scheme is locally conservative in the sense that $F_{jl}(t) = -F_{lj}(t)$. This property makes the finite volume method a suitable choice in problems where the flux plays an important roll.

Weighted particle method

The weighted particle method is specially suited to treat non local diffusion terms. Consider, for example, the equation

$$\begin{cases} \partial_t n(t,x) + \nabla \cdot (a(x) n(t,x)) - D(n(t,x)) = 0, & x \in \mathbb{R}^2, t > 0, \\ \\ n(0,x) = n_0(x), & x \in \mathbb{R}^2. \end{cases}$$

where $a \in \mathscr{C}^1(\mathbb{R}^2, \mathbb{R}^2)$ and $D(n(t,x))$ is the diffusion operator

$$D(n(t,x)) := \int_{\mathbb{R}^2} \sigma(x,y) n(t,y) dy - \int_{\mathbb{R}^2} \sigma(y,x) n(t,x) dy,$$

for some regular function $\sigma(x,y)$.
Multiplying the differential equation by $\varphi \in \mathscr{C}_c^1(\mathbb{R}^2)$, integrating by parts the adjective term and changing the order of integration on the diffusion term shows that

$$\frac{d}{dt}\left(\int_{\mathbb{R}^2} n(t,x)\varphi(x)dx\right) - \int_{\mathbb{R}^2} n(t,x) a(x) \cdot \nabla\varphi(x) dx - \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \sigma(x,z) n(t,z)(\varphi(x) - \varphi(z)) dx dz = 0, \quad (10)$$

for all $\varphi \in \mathscr{C}_c^1(\mathbb{R}^2)$. On the other hand, for all $y \in \mathbb{R}^2$, the characteristic curves for this problem are given by the Cauchy problem

$$\begin{cases} \frac{d}{dt} X(t,y) = a(X(t,y)), & t > 0, \\ \\ X(0,y) = y. \end{cases}$$

Denoting as $J(t,y)$ the determinant of Jacobian matrix of $X(t,y)$, is a well established fact that

$$\frac{d}{dt} J(t,y) = J(t,y)(\nabla \cdot a)(X(t,y)).$$

The map $X(t,y)$ is a dipheomorphism of $\mathbb{R}^2$ into itself, for all values of t, hence, by making the change of variables $x = X(t,y)$, we get

$$\int_{\mathbb{R}^2} n(t,x)\varphi(x) dx = \int_{\mathbb{R}^2} n(t, X(t,y))\varphi(X(t,y)) J(t,y) dy.$$

We may approximate the integral on the right-hand side of this relation using the term

$$\sum_{k\in\mathscr{K}} \omega_k \nu_k(t)\varphi(X(t,y_k))J(t,y_k),$$

for some known set $(y_k,\omega_k)_{k\in\mathscr{K}}$ of points $y_k\in\mathbb{R}^2$ and weights $\omega_k\in\mathbb{R}^+$ and some unknown strength values $\nu_k(t)$. Injecting this approximation and homologous ones for the other integral terms in (10), we obtain the system to be solved:

$$\begin{cases} \dot{x}_k(t) = a(x_k), \\ \dot{w}_k(t) = \nabla\cdot a(x_k(t))w_k(t), \\ \dot{\nu}_k(t) = -\nabla\cdot a(x_i(t))\nu_i(t) + \sum_{j\in\mathscr{K}} w_j(t)\left(\sigma(x_k(t),x_j(t))\nu_j(t) - \sigma(x_j(t),x_k(t))\nu_k(t)\right), \\ x_k(0) = y_k,\ w_k(0) = \omega_k,\ \nu_k(0) = n_0(y_k), \end{cases} \tag{11}$$

where we have adopted the notations $x_k(t) := X(t,y_k)$ and $w_k(t) = J(t,y_k)$.
In other words, the particle method amounts to approximate the solution $n(t,x)$ by a sum of Dirac delta's centered at the points $x_k(t)$, with weights $w_k(t)$ and strengths $\nu_k(t)$.
As shown, this method is especially adapted for the linear advection equation, but has been generalised to many other kinds of equations which mostly come from physics [49], such as diffusion equations [50–54], advection-diffusion equations [55, 56], convection-diffusion equations [57], the Navier-Stokes equation [58, 59] or the Vlasov-Poisson equation [60, 61].

Main results of the thesis

The work being presented here deepens on the models and methods showcased so far and generalizes some of the already established results.

Cancer containment for Norton-Simons models

The way model (1) is stated neglects mutations from sensitive to resistant cells, and assumed that the growth-rate of sensitive cells is non-increasing in the size of the resistant population (first part of assumption (a3)). The latter is not true in standard models of chemotherapy.
In Chapter 1 we show how to dispense with this assumption and allow for mutations from sensitive to resistant cells. This is achieved by a novel mathematical analysis comparing tumor sizes across treatments not as a function of time, but as a function of the resistant population size.
As done previously, we consider a model with two types of tumors cells: sensitive to treatment, and fully resistant. Their growth is described by differential equations of the form:

$$\begin{cases} \dfrac{dS}{dt}(t) = \phi_S(S(t),R(t),L(t)), & S(0) = S_0 \geq 0 \\ \dfrac{dR}{dt}(t) = \phi_R(S(t),R(t)), & R(0) = R_0 > 0 \end{cases} \tag{12}$$

where ϕ_S and ϕ_r are continuously differentiable absolute growth-rate functions. The quantities $\phi_S(0, R, L)$ and $\phi_R(S, 0)$ are assumed non-negative to ensure that population sizes cannot become negative. Let $N(t) = S(t) + R(t)$ and $N_0 = S_0 + R_0$. We make the following assumptions:

- The patient dies when tumor size reaches a critical size $N_{crit} > N_0$.
- The size of an untreated tumor increases: $\phi_S(S, R, 0) + \phi_R(S, R) > 0$ if $N \leq N_{crit}$.
- The higher the treatment level, the lower the growth-rate of sensitive cells: ϕ_S is non-increasing in L.
- The resistant population keeps growing: $\phi_R(S, R) > 0$ whenever $R > 0$ and $N \leq N_{crit}$, so that the tumor is incurable if, as we assume, resistant cells are initially present.
- If $R \geq R_0$ and $N \leq N_{crit}$, for a given number of resistant cells, the larger the sensitive population, the lower the growth-rate of resistant cells: ϕ_R is non-increasing in S.

The difference with Viossat and Noble (2021) [1] (see assumptions (a1) through (a5)) is two-fold: first, the model is formulated in terms of absolute growth-rates, allowing for mutations from sensitive to resistant cells and back. Second, we make no assumption on how the growth-rate of sensitive cells depends on the number of resistant cells. In particular, ϕ_S is not assumed non-increasing in R.
More specifically, we show that, up to natural additional assumptions for comparisons of sensitive cell populations, all results of Viossat and Noble on (1) still hold on (12), in spite of our less restrictive assumptions.

The key point is that if treatment level is never larger than a given constant for treatment 1, and never smaller than the same constant for treatment 2, then the resistant population is never larger under treatment 1 than under treatment 2.

Proposition 0.1. *Consider solutions of 12 associated to two treatments $L_1(t)$ and $L_2(t)$. If there exists a constant $\bar{L}$ such that for all $t \geq 0$, $L_1(t) \leq \bar{L} \leq L_2(t)$, then $R_1(t) \leq R_2(t)$ for all $t \geq 0$.*

Proposition 0.1 is the key results which allow to compare the sizes of outcomes and survival times related to some of the more well known strategies. Amid these strategies we may find

• Constant dose treatments, including No treatment (noTreat): $L(t) = 0$, and Maximal Tolerated Dose (MTD): $L(t) = L_{max}$ throughout.

• Delayed MTD (del-MTD): do not treat until $N = N_{tol}$ for the first time, then treat at L_{max} for ever.

• Containment at N_{tol} (Cont): do not treat until $N = N_{tol}$ and then stabilize tumor size at N_{tol}, as long as possible with a treatment level $L(t) \leq L_{max}$. Finally, treat at L_{max} when $N > N_{tol}$. Formally, during the stabilization phase, the treatment level is chosen so that $dN/dt = 0$. Containment treatments are illustrated in Fig. 1

• Intermittent containment (Int), as in the prostate cancer clinical trial of Zhang et al. (2017): do not treat until $N = N_{tol}$, then treat at L_{max} until $N = N_{min} < N_{tol}$, then interrupt treatment until $N = N_{tol}$, and iterate as long as possible. Finally, treat at L_{max} when $N > N_{tol}$. This is illustrated by Fig. 2.

• An arbitrary treatment, called the alternative treatment (alt): we only assume that $0 \leq L(t) \leq L_{max}$ for all t and $L(t) = L_{max}$ if $N > N_{tol}$.

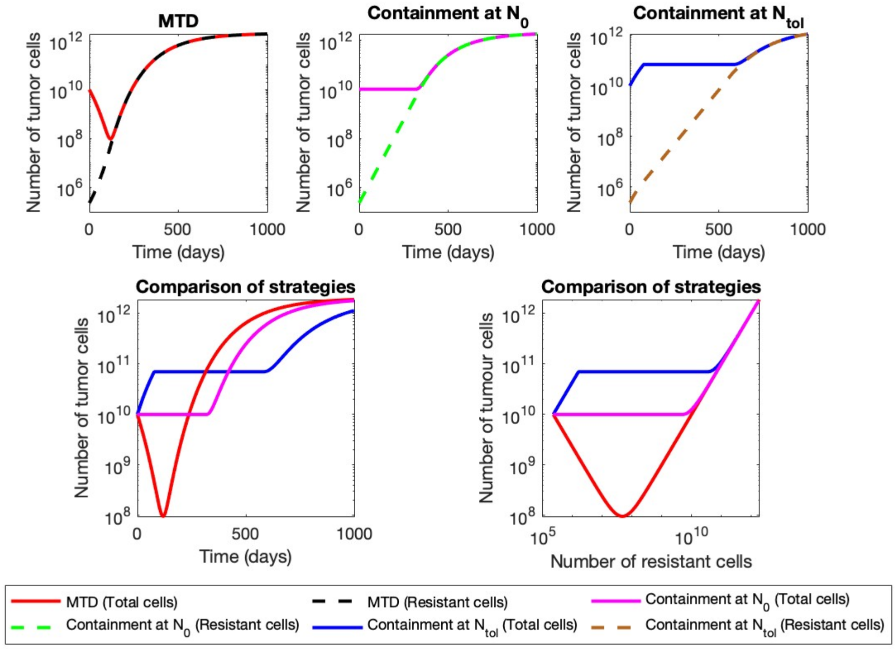


Figure 1: Number of resistant cells and number of tumor cells for different treatments. Top row: number of resistant cells and number of tumor cells as a function of time under MTD (left), Containment at the initial size N_0 (center), and Containment at the maximal tolerable size N_{tol} (right). Bottom-row: number of tumor cells under these three treatments as a function of time (left) or as a function of the number of resistant cells (right).

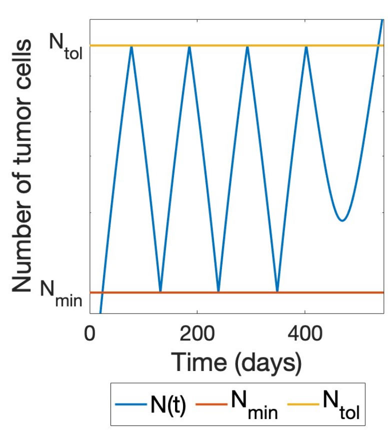


Figure 2: Total size under Intermittent treatment.

It follows that for constant dose treatments, lowering the dose or delaying treatment leads to a lower resistant population:

Proposition 0.2. *(constant dose treatments)*
a) Consider two constant dose treatments $L_1(t) = L_1$ and $L_2(t) = L_2$. If $L_2 \geq L_1$, then $R_1(t) \leq R_2(t)$ for all $t \geq 0$.

b) Assume that $L_1(t) = L > 0$ for all $t \geq 0$, while $L_2(t) = 0$ until $N = N_{start} \geq N_0$, and then $L_2(t) = L$. Then $R_1(t) \leq R_2(t)$ for all $t \geq 0$.

Proposition 0.1 also implies that not treating minimizes the resistant population while MTD maximizes it:

Proposition 0.3. *(MTD maximizes resistance)*
For all $t \geq 0$, $R_{noTreat}(t) \leq R_{alt}(t) \leq R_{MTD}(t)$.

Of course, not treating is typically not an option, as the number of sensitive cells would explode, but containment is. One of our main results is that containment minimizes the resistant population among all treatments treating at L_{max} after failing.

Proposition 0.4. *(containment minimizes resistance)*
For all $t \geq 0$, $R_{Cont}(t) \leq R_{alt}(t)$

More precise statements will be made for idealized versions of containment and MTD, and a comparison between all reference treatments will be established. The exhaustive list of comparisons can be found in section 1.3 while the proofs are in section 1.4.

Asymptotic expansion of the limit cycle for a tumour-immune system interaction model

In Chapter 2 we retake the issue of appearance of periodic orbits on a tumour-immune system interaction model. We provide an asymptotic expansion depending on the parameters of the problem for the limit cycle, paving the way for future studies of propagation of instabilities and appearance of twinkling patterns on an heterogeneous setting.

Bet-hedging and the atavistic theory of cancer

Confronted with the biological problem of managing plasticity in cell populations, which is in particular responsible for transient and reversible drug resistance in cancer, in Chapter 3 we propose a rationale consisting of an integro-differential reaction-advection-diffusion equation, the properties of which are studied theoretically and numerically. By using a constructive finite volume method, we show the existence and uniqueness of a weak solution and illustrate by numerical approximations and their simulations the capacity of the model to exhibit divergence of traits. This feature may be theoretically interpreted as describing a physiological step towards multicellularity in animal evolution and, closer to present-day clinical challenges in oncology, as a possible representation of bet hedging in cancer cell populations.
The link between the origin of multicellularity and cancer is justified by the atavistic theory of cancer,

states that the hallmark capabilities of cancer are based on latent functions already existing in the genome of normal human cells, and that cancer represents a reversion to a less differentiated and less cooperative cellular behavior. In the atavistic theory, accompanied or induced by blockade of differentiation or reverse differentiation of normally maturing cells, societies of cells in a multicellular organism (cancer is always a disease of multicellular organisms) somehow, in some location of the organism, escape the fine control under which they are normally placed and revert to a previous, coarse and disorganised state of multicellularity [62]. This may be understood as a process of "deDarwinisation", through which cancer cells gain a state of *plasticity* [63–66] representative of a former state in the evolution of multicellularity.

The passage from unicellular organisms to multicellular ones led to the regulation of capabilities, resulting in controlled proliferation and differentiation of cells leading to specialisation and cooperation between specialised cells. The role of environment-driven cellular stress in this process of specialisation has recently been stressed by various authors [67, 68]. The new genes responsible for these regulations became tumour suppressors. The atavistic theory states that if these new suppressors become damaged for some reason, then latent genes, associated with functions from unicellular organisms, will reappear and dominate the scenery, thus resulting in the unconstrained proliferation and the lack of cooperation with the other cells of the host organism, as actually found in tumours.

One can reasonably assume that those primitive organisms adopted bet hedging strategies, i.e., common risk-diversifying strategies in unpredictably changing and often aggressive environments, in order to maximise their phenotypic fitness [69, 70].

Among such commonly described strategies of living organisms (unicellular or multicellular) meant to ensure survival in changing environments have been classically described fright, fight and flight. Fright (or freeze) is not likely to induce phenotype evolution. Fight (establishing barriers, secreting poisons, gathering in colonies) and flight (motility to escape unbeatable predators) can. Differentiation between somatic and germinal cells is also a major step in evolution. Bet hedging strategies were not only present at elementary stages of evolution. They are a common adaptive tool that can still be found in nature at different levels of complexity, from prokaryotic organisms to vertebrate ones. In between, tumour cells, thanks to their high plasticity, in the presence of an aggressive environment provided by immune response of the host body or of any anti-cancer treatment, may adopt bet hedging as a strategy to guarantee a prolonged survival of their colony. The wide presence of bet hedging in nature as an evolutionary mechanism, and its many links to the development of cancer is what motivates us in the present attempt towards a mathematical model representing some of the factors that influence this phenomenon (natural selection, epimutations and environmental stress).

For the model, we consider a population (not necessarily of tumour cells) in which each individual has three defining traits: viability associated with the variable $x \in [0, 1]$ which reflects the potential to resist deadly insults, fecundity associated with the variable $y \in [0, 1]$ representing the potential to proliferate and plasticity associated with the variable $\theta \in [0, 1]$ which represents the potential to continue to differentiate within a differentiation tree. We assume furthermore that for a certain regular function $C : \mathbb{R}^2 \to \mathbb{R}$ and a positive constant K, $(x, y) \in \Omega := \{C(x, y) \leqslant K\}$, so $z = (x, y, \theta)$ ranges over the set $D := \{\Omega \times [0, 1]\}$. We then consider the evolution problem (13), (14), (15) on the density of population $n = n(t, z) \geqslant 0$.

$$\partial_t n + \nabla \cdot \Big(Vn - A(\theta)\nabla n\Big) = (r(z) - d(z)\rho(t))n, \tag{13}$$

$$\Big(Vn - A(\theta)\nabla n\Big) \cdot \mathbf{n} = 0, \text{ for all } z \in \partial D, \tag{14}$$

$$n(0, z) = n_0(z), \text{ for all } z \in D. \tag{15}$$

In the above equation, chosen for the sake of simplicity as diagonal, the matrix

$$A(\theta) = \begin{pmatrix} a_{11}(\theta) & 0 & 0 \\ 0 & a_{22}(\theta) & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

gives the speed at which non-genetic epimutations occur, otherwise said it is a minimally simple representation of how the internal plasticity trait θ affects the non-genetic instability of traits x and y, by tuning the diffusion term; the function

$$V(t,z) = (V_1(t,z), V_2(t,z), V_3(t,z))$$

represents the sensitivity of the population to abrupt changes in the environment;

$$\rho(t) = \int_D n(t,z)dz$$

stands for the total amount of individuals in the population at time t.
We introduce now the variational formulation of (13)-(15). Denote $H = L^2(D)$, with $(\cdot,\cdot)_H$ the usual scalar product in that space, and $\mathscr{V} = H^1(D)$ with $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{\mathscr{V}'\times\mathscr{V}}$ being the duality product in $\mathscr{V}$. For any given $n_0 \in H$, $T > 0$, we say that

$$n := n(t) \in X_T := C([0,T], H) \cap L^2((0,T), \mathscr{V}) \cap H^1([0,T], \mathscr{V}'),$$

is a variational solution of the problem (13)-(15) if it is a solution in the following weak sense

$$(n(t), \varphi(t))_H = (n_0, \varphi(0))_H + \int_0^t \Big(\langle Q[n](s), \varphi(s)\rangle + \langle \partial_s \varphi(s), n(s)\rangle \Big) ds, \tag{16}$$

where

$$\langle Q[n], \varphi\rangle = \int_D \Big(-A\nabla n \nabla\varphi + V n \nabla\varphi + (r(z) - \rho d(z)) n\varphi \Big) dz,$$

for any $\varphi \in X_T$. We say that n is a global solution if it is a solution on $[0,T]$ for any $T > 0$.

Theorem 0.1. *For all non-negative $n_0 \in L^p(D)$, $p > 2$, there exists a unique global non-negative weak solution for problem (13)-(15) in the sense of (16).*

We focus on giving a proof for this theorem using a discretized version of problem (13)-(15) after applying the Finite Volume Method to it. For this purpose, we define a set $D_h \supset D$, that can be covered by the union of N disjoint cubic cells, denoted as D_j, of side length h. After integrating the equation (13) over each of the cells D_j we derive the system of first-order differential equations

$$\frac{d}{dt}\nu_j(t) = M_j(t, \nu(t))\nu_j(t) + \sum_{l\in N_j} B_{jl}(t)\nu_l(t), \tag{17}$$

$$\nu_j(0) = \frac{1}{h^3}\int_{D_j} n_0(z)dz, \tag{18}$$

where ν_j is an approximation of the average of the solution $n(t,z)$ over D_j, N_j is the set of indexes corresponding to the neighbours of D_j and the coefficients M_j and B_{jl} are functions of $V(t,z)$, $A(\theta)$, $r(z)$ and $d(z)$. A full detailed derivation of the scheme is given in Section (3.3). We can then introduce the following result involving the solution for this system:

Theorem 0.2. *For all non-negative $n_0 \in L^p(D)$, $p > 2$, there exists a unique non-negative solution for problem (17)-(18). Furthermore, the function $\tilde{n}_h(t,z)$ defined by*

$$\tilde{n}_h(t,z) = \sum_j \nu_j(t) \mathbb{1}_{D_j \cap D},$$

converges in $L^2(D_T)$ to the unique non-negative weak solution of (13)-(15) as h goes to zero.

Figure (3) illustrates the convergence result from Theorem 0.2 by showing the dependence between the mesh size and the error with respect to a previously known solution.

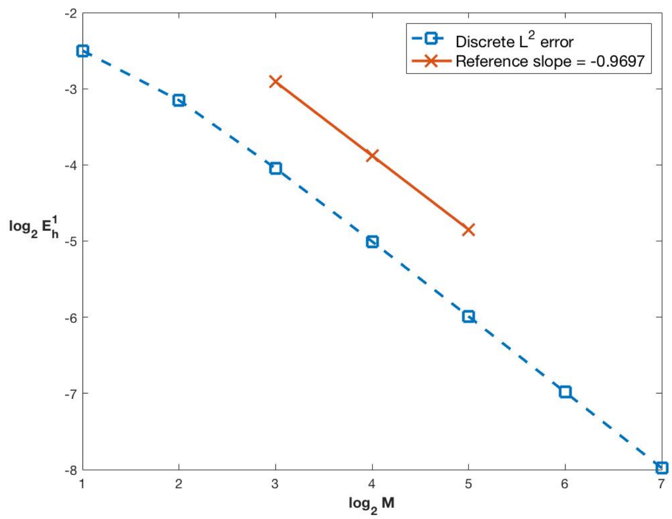


Figure 3: The discrete $L^2(D_T)$ error for the semi-discrete scheme, for $T = 10$ and M ranging between 2 and 128.

In addition to the theoretic results exposed in Theorems 0.1 and 0.2, some phenomena of biological interest are as well reproduced through simulations. The phenotype divergence due to the effect of the environment is evidenced in Figure 4 through an specific choice of the function $V(t,z)$ while the lost of plasticity in favor of specialization can be observed in Figures 5 and 6.

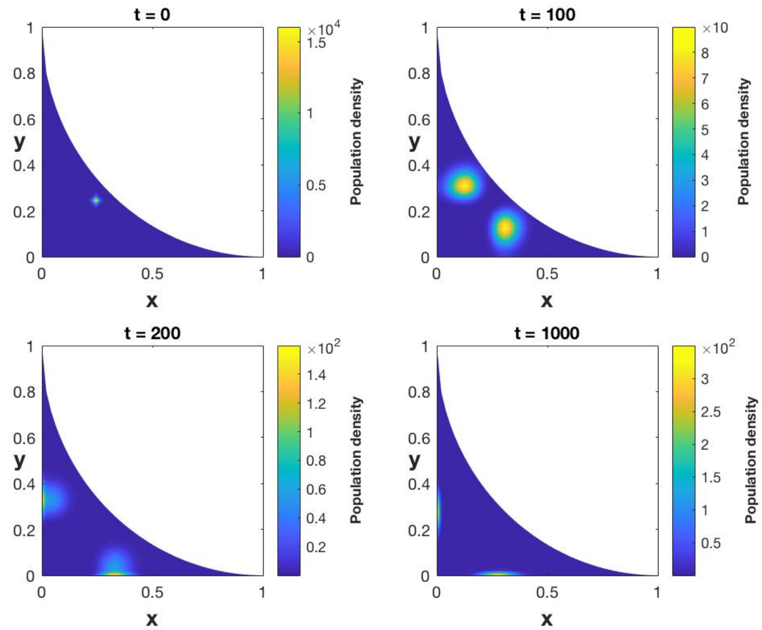


Figure 4: Evolution of a population under the effect of the environment.

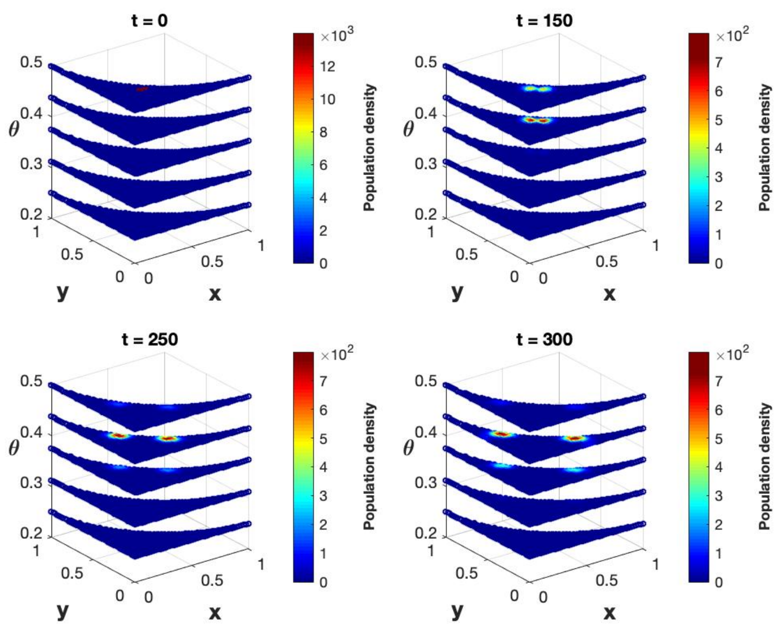


Figure 5: Initial stages of the population density for different values of θ: The differentiation process starts. At around $t = 250$ (bottom left) most of the population has already concentrated around the plasticity level $\theta = 0.4375$ and around $t = 300$ (bottom right) we observe that the migration towards a less plastic state continues.

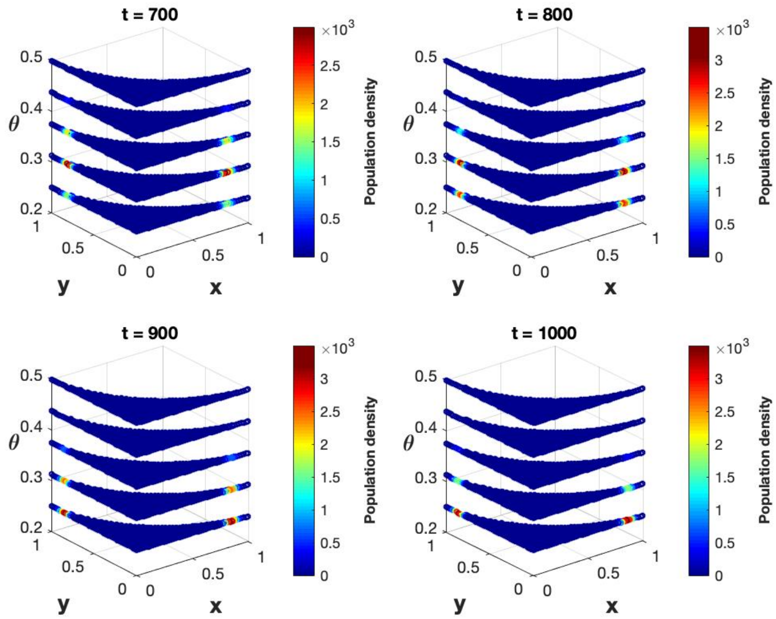


Figure 6: Final stages of the population density for different values of θ: Around $t = 900$ (bottom left) the differentiation process is over and most of the population has reached the plasticity level $\theta = 0.25$. At $t = 1000$ (bottom right) we observe that the population concentrated around any other level of plasticity is almost extinct, and only the one around $\theta = 0.25$ survives.

Chapter 4 serves as a complement and an extension to Chapter 3. On it, we discuss the mathematical modelling of two of the main mechanisms which pushed forward the emergence of multicellularity: phenotype divergence in cell differentiation, which was already treated on Chapter 3 and between-cell cooperation. In line with the atavistic theory of cancer, this disease being specific of multicellular animals, we set special emphasis on how both mechanisms appear to be reversed, however not totally impaired, rather hijacked, in tumour cell populations. Two settings are considered: the completely innovating, tinkering, situation of the emergence of multicellularity in the evolution of species, which we assume to be constrained by external pressure on the cell populations, and the completely planned - in the *body plan* - situation of the physiological construction of a developing multicellular animal from the zygote, or of bet hedging in tumours, assumed to be of clonal formation, although the body plan is largely - but not completely - lost in its constituting cells. We show how cancer impacts these two settings and we sketch mathematical models for them. We present there our contribution to the question at stake with a background from biology, from mathematics, and from philosophy of science.

A particle method for non-local models from adaptive dynamics

In Chapter 5 the well-posedness of a non-local advection-selection-mutation problem deriving from adaptive dynamics models is shown for a wide family of initial data. A particle method is then developed, in order to approximate the solution of such problem by a regularised sum of weighted Dirac masses whose characteristics solve a suitably defined ODE system. The convergence of the particle method over any finite interval is shown and an explicit rate of convergence is given. Furthermore, we investigate the asymptotic-preserving properties of the method in large times, providing sufficient conditions for it to hold true as well as examples and counter-examples. Finally, we illustrate the

method in two cases taken from the literature.
More specifically, the goal of this chapter is to develop a numerical method allowing to approximate the solutions of equations of the form

$$\begin{cases} \partial_t v(t,x) + \nabla_x \cdot (a(t,x,I_a v(t,x))v(t,x)) = R(t,x,I_g v(t,x))v(t,x) + \int\limits_{\mathbb{R}^d} m(t,x,y,I_d v(t,x))v(t,y)dy, \\ v \in \mathscr{C}([0,T], L^1(\mathbb{R}^d)), \\ v(0,\cdot) = v^0(\cdot) \in W^{1,1}(\mathbb{R}^d), \end{cases} \tag{19}$$

where

$$(I_l u)(t,x) = \int\limits_{\mathbb{R}^d} \psi_l(t,x,y)u(t,y)dy, \quad l = a, g, d$$

are non-local terms and a, R, m and ψ_l are smooth functions.
Upon establishing the well-posedness of (19), the chapter is concerned with the derivation of a particle method inspired by [57], the analysis of its convergence and asymptotic-preserving properties. However, we must emphasise the two main novelties with respect to that work: First, the use of non local terms, which as we will show, poses technical difficulties and affects the existence of smooth solutions in certain cases. Secondly, the study of the asymptotic preserving property, which guarantees that, under certain hypotheses, the long time behaviour of the solution is conserved.
The first main result we derive, is the existence and uniqueness of solution for problem(19) when the initial data is smooth

Theorem 0.3. *Consider $k \geqslant 1$ and $T > 0$. For all non-negative functions $v^0 \in \mathscr{C}_c^k(\mathbb{R}^d)$, there exists a unique non-negative classical solution $v \in \mathscr{C}^1([0,T], \mathscr{C}_c^k(\mathbb{R}^d))$ to problem (19). Furthermore, such solution satisfies*

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}, \tag{20}$$

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{W^{k,1}(\mathbb{R}^d)} \leqslant C_T \|v^0\|_{W^{k,1}(\mathbb{R}^d)}. \tag{21}$$

However, if more general initial data are taken into consideration, the regularity of the solution depends on the non-locality of the advective term. The following two theorems better explain this statement.

Theorem 0.4. *For all $k \geqslant 1$ and any non-negative functions $v^0 \in W^{k,\infty}(\mathbb{R}^d)$ with compact support, there exists a unique non-negative weak solution $v \in \mathscr{C}([0,T], \mathscr{C}_c^{k-1}(\mathbb{R}^d))$ to problem (19). Furthermore, such solution satisfies*

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}, \tag{22}$$

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{W^{k-1,1}(\mathbb{R}^d)} \leqslant C_T \|v^0\|_{W^{k-1,1}(\mathbb{R}^d)}, \tag{23}$$

and, for $k \geqslant 2$, $v \in \mathscr{C}^1([0,T], \mathscr{C}_c^{k-1}(\mathbb{R}^d))$.

Theorem 0.5. *If $\partial_I a = 0$, for all non-negative functions $v^0 \in W^{k,1}(\mathbb{R}^d)$, there exists a unique non-negative weak solution $v \in \mathscr{C}([0,T], W^{k,1}(\mathbb{R}^d))$ of problem (19). Furthermore, such a solution satisfies*

$$\sup_{t\in[0,T]} \|v\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}, \tag{24}$$

$$\sup_{t\in[0,T]} \|v\|_{W^{k,\infty}(\mathbb{R}^d)} \leqslant C_T \|v^0\|_{W^{k,1}(\mathbb{R}^d)}. \tag{25}$$

In other words, if the function a depends on a non-local term, the solution may experience a loss of regularity.
Once established the well posedness of the problem in question, the next step is to derive a semi-discrete scheme in order to approximate its solutions. Following [57], for a given set of indexes $\mathscr{J}_h$, the particle method associated to (19) consists in looking for a measure ν_h of the form

$$\nu_h(t) = \sum_{i\in\mathscr{J}_h} \nu_i(t) w_i(t) \delta_{x_i(t)},$$

where $(\nu := \{v_i(t)\}_{i\in\mathscr{J}_h}, w := \{w_i(t)\}_{i\in\mathscr{J}_h}, \overline{x} := \{x_i(t)\}_{i\in\mathscr{J}_h})$, is the solution of the following system

$$\begin{cases} \dot{x}_i(t) = A_{\nu,w}(t, x_i), \\ \dot{w}_i(t) = \mathrm{div}\ A_{\nu,w}(t, x_i(t)) w_i(t), \\ \dot{\nu}_i(t) = \big(-\mathrm{div}\ A_{\nu,w}(t, x_i(t)) + R(t, x_i(t), I_g(t, x_i(t), \nu, w))\big)\nu_i(t) \\ \qquad + \sum_{j\in\mathscr{J}_h} w_j(t)\nu_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t), \nu, w)), \\ x_i(0) = x_i^0,\ w_i(0) = w_i^0,\ \nu_i(0) = v^0(x_i^0), \end{cases} \tag{26}$$

where

$$A_{\nu,w}(t,x) = a(t, x, I_a(t, x, \nu, w)),$$

and

$$I_l(t, x, \nu, w) := \sum_{j\in\mathscr{J}_h} \nu_j(t) w_j(t) \psi_l(t, x, x_j(t)),$$

with $l \in \{a, g, d\}$.
Our next result ensures the existence of solution for this numerical scheme

Theorem 0.6. *For all $T > 0$ and all non-negative initial data $v^0 \in \ell^1(\mathscr{J}_h, \Omega^0)$ there exists a unique solution $x_i \in \mathscr{C}^1([0,T])$, for all $i \in \mathscr{J}_h$, $w := \{w_i(\cdot)\}_{i\in\mathscr{J}_h} \in \mathscr{C}([0,T], \ell^\infty(\mathscr{J}_h))$ and $0 \leqslant \nu := \{\nu_i(\cdot)\}_{i\in\mathscr{J}_h} \in \mathscr{C}([0,T], \ell^1(\mathscr{J}_h))$ of problem (26). Furthermore, there exist positive constants c_T and C_T such that the solution satisfies, for all $t \in [0,T]$*

$$c_T h \leqslant |x_i(t) - x_j(t)| \leqslant C_T h,\ \forall i, j \in \mathscr{J}_h,\ i \neq j, \tag{27}$$

$$c_T h^d \leqslant w_i(t) \leqslant C_T h^d,\ \forall i \in \mathscr{J}_h, \tag{28}$$

$$\|\nu w\|_{1,h} \leqslant \max\{\|v^0 h^d\|_{\ell^1}, \frac{I^*}{\underline{\psi_g}}\}. \tag{29}$$

The final sections in Chapter 5 study the convergence of the particle method solution towards a solution of (19). The first result states that up to any finite time T,

Theorem 0.7. *There exists $C > 0$ (which depends on T, a, R, m and $\overline{\rho}$), and positive values r, κ and μ such that*

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leqslant C\big(\varepsilon^r + \big(\frac{h}{\varepsilon}\big)^\kappa + h^\kappa\big)\|v^0\|_{W^{\mu,1}(\mathbb{R}^d)}, \quad \forall\, 0 \leqslant t \leqslant T,$$

where v_ε^h is a regularized version of the solution for (26).

In order to study the asymptotic behaviour of the solution for (19), convergence in finite time is not enough. It is necessary to check that the numerical scheme conserves the asymptotic behaviour. The last result in Chapter 5 identifies some conditions under which this property is satisfied.
Assume that there exists $\hat{x} \in \mathbb{R}^d$ an asymptotically stable equilibrium for the ODE '$\dot{x} = a(x)$' and that there exists $C, \delta > 0$ such that

$$\forall y \in \text{supp}(n^0), t \geq 0, \quad \|X(t,y) - \hat{x}\| \leq Ce^{-\delta t}. \tag{30}$$

Moreover, let us assume that there exist positive values D, I^m and I^M such that

$$R(x, I^m) \geqslant 0, \quad R(x, I^M) \leqslant 0 \text{ and } \partial_I R(x, I) \leqslant -D, \quad \forall x \in \text{supp}(v^0). \tag{31}$$

Theorem 0.8. *Let us assume that there exists $\hat{x} \in \mathbb{R}^d$ which is an asymptotically stable equilibrium for the ODE $\dot{x} = a(x)$ such that (30) holds. We assume as well that $m \equiv 0$ and that R satisfies (31). Then, v converges to $\hat{\rho}\, \delta_{\hat{x}}$ in the weak sense in the space of Radon measures, where $\hat{\rho}$ is the unique solution of*

$$R(\hat{x}, \psi_g(\hat{x}, \hat{x})\hat{\rho}) = 0.$$

Consequently, v_ε^h is an asymptotic preserving approximation of v.

The fully parabolic Keller-Segel system

Chapters 6 and 7 are devoted to the study of the re-scaled parabolic-parabolic Keller-Segel system

$$\begin{cases} \partial_t f &= \Delta f + \text{div}(\mu x f - f\nabla u) \\ \partial_t u &= \dfrac{1}{\varepsilon}(\Delta u + f) + \mu x \cdot \nabla u, \end{cases} \tag{32}$$

We first introduce the perturbation (g, v) defined by

$$f = Q + g, \quad u = P + v,$$

where $Q = Q_\varepsilon^\mu$ and $P = P_\varepsilon^\mu$ is a steady state of (32). If (f, u) is a solution to (32) then (g, v) satisfies the system

$$\begin{cases} \partial_t g = \Delta g + \text{div}(\mu x g - g\nabla P - Q\nabla v) - \text{div}(g\nabla v) \\ \partial_t v = \dfrac{1}{\varepsilon}(\Delta v + g) + \mu x \cdot \nabla v, \end{cases} \tag{33}$$

and reciprocally.

We are next interested on the linearized equation around a re-scaled self-similar profile

$$\left\{\begin{array}{rcl}\partial_t g &=& \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla v)\\ \partial_t v &=& \dfrac{1}{\varepsilon}(\Delta v + g) + \mu x\cdot\nabla v.\end{array}\right.$$

Let us define the Laplace kernel in the plane

$$\kappa(z) := -\frac{1}{2\pi}\log|z|, \quad \mathscr{K}(z) := \nabla\kappa(z) = -\frac{1}{2\pi}\frac{z}{|z|^2}, \tag{34}$$

so that $\omega := \kappa * \Omega$ is a solution to the Laplace equation

$$-\Delta\omega = \Omega \quad \text{in} \quad \mathbb{R}^2.$$

Next defining

$$w := v - \kappa * g,$$

the equation on w is

$$\partial_t w = \frac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + \mu x\cdot\nabla\kappa * g - \nabla\kappa * [\nabla g + \mu x g - g\nabla P - Q\nabla\kappa * g - Q\nabla w].$$

In fact, by using that

$$x\cdot\nabla\kappa * g - \nabla\kappa * (xg) \simeq \int \frac{(x-y)}{|x-y|^2}\big\{x\, g(y) - y\, g(y)\big\}\, dy \simeq \langle g\rangle$$

and $\langle g\rangle = 0$, the second equation simplifies. The system of equations on (g,w) becomes

$$\left\{\begin{array}{rcl}\partial_t g &=& \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla\kappa * g - Q\nabla w)\\ \partial_t w &=& \dfrac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g + Q\nabla w],\end{array}\right. \tag{35}$$

and we will focus on the dissipativity properties of the associated operator. More precisely, defining

$$\mathscr{L}(g,w) := (\mathscr{L}_1(g,v), \mathscr{L}_2(g,v))$$

with

$$\begin{array}{rcl}\mathscr{L}_1(g,w) &:=& \Delta g + \mathrm{div}(\mu x g - g\nabla P) - \mathrm{div}(Q\nabla\kappa * g + Q\nabla w)\\ \mathscr{L}_2(g,w) &:=& \dfrac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g + Q\nabla w],\end{array}$$

In Chapter 6, we will exhibit some scalar products $(\cdot,\cdot)_{\mathscr{H}}$ and associated norm $\|\cdot\|_{\mathscr{H}}$ such that

$$(\mathscr{L}(g,w),(g,w))_{\mathscr{H}} \le -\lambda\|(g,w)\|^2_{\mathscr{H}} + \dots,$$

with $\lambda > 0$ as large as possible and the remainder term "…" is essentially negative. We attempt to give an improved spectral analysis in the radially symmetric case in order to describe blowing up solutions in the critical 8π mass case in the spirit of what has been done for the parabolic-elliptic Keller-Segel equation in [39, 71, 72] (see also [40]). Although our project has been up to now unsuccessful, we are able to present the following moderate dissipative estimate

Theorem 0.9. *There exists a functional space $\mathscr{H}$, where there holds*

$$(\mathscr{L}(g,w),(g,w))_{\mathscr{H}} \le -\mu\|(g,w)\|_{\mathscr{H}}^2 - \frac{1}{2}\|\nabla g\|_{L^2_m}^2 - \frac{1}{2\varepsilon}\|\nabla w\|_{L^2}^2 + C\|g\|_{L^2(B_R)}^2,$$

for any $\varepsilon \in (0,\varepsilon_0)$ and $\mu \in (0,1]$, where $C, R > 0$ are some constants (independent of ε and μ) and $\varepsilon_0 > 0$ is small enough.

On the other hand, in Chapter 7 we give a description of the longtime self-similar behavior of solution in the sub-critical mass case without radially symmetric assumption in the spirit of what has been done in the radially symmetric case in [48]. More specifically, we are able to prove the following theorem

Theorem 0.10. *There exist functional spaces X and Y and there are $\varepsilon_0, \eta_0 > 0$ such that for any $\varepsilon \in (0,\varepsilon_0)$ and any initial data $(g_0,v_0) \in L^1_{m,0} \times (L^p \cap \dot{H}^1)$ with $|||(g_0,v_0)|||_X \le \eta_0$, there exists a unique global solution $(g,v) \in L^\infty_t(X) \cap L^2_t(Y)$ to (33), which verifies*

$$\|(g,v)\|_{L^\infty_t(X)} + \|(g,v)\|_{L^2_t(Y)} \lesssim |||(g_0,v_0)|||_X. \tag{36}$$

Moreover we have the decay estimate, for any $\lambda \in (0,\mu)$,

$$\|(g(t),v(t))\|_X \lesssim e^{-\lambda t}\|(g_0,v_0)\|_X. \tag{37}$$

Chapter 1

Tumor containment: a more general mathematical analysis

1.1 Introduction

The dominant paradigm in cancer therapy is to treat tumors aggressively. This makes sense if the tumor is curable, but might be counter-productive otherwise. Indeed, tumors contain a large number of cells, some of which may be resistant to treatment. By killing preferentially the most sensitive cells, an aggressive treatment could free resistant cells from competition with sensitive cells, allowing them to develop quickly: a phenomenon called competitive release in ecology (Gatenby 2009b [73], Enriquez-Navas et al., 2016 [74], Cunningham et al. 2019 [75]).

This led researchers to suggest that, at least for some tumors, treating at, or close to, the maximal tolerated dose should be replaced by treating at the minimal effective dose; that is, the minimal dose that stabilizes tumor size, subject to a sufficient quality of life of the patient, and not putting her life at risk in the short-run. The aim is to slow down the growth of resistant cells by maintaining competition with sensitive cells.

This idea, which is part of the broader framework of cancer adaptive therapy (Gatenby 2009 [4]), has been tested in vitro, in mice models and on human patients suffering from metastatic castrate-resistant prostate cancer (Gatenby et al. 2009 [4], Silva et al. 2012 [5], Enriquez-Navas et al. 2016 [74], Zhang et al. 2017 [7]: trial NCT02415621, Bacevic and Noble et al. 2017 [8], Smalley et al. 2019 [76], Strobl et al. 2020 [14], Bondarenko et al. 2021 [77], Wang et al. 2021a, 2021b [78, 79], Farrokhian et al. 2022 [80]). Other clinical trials are ongoing or starting in prostate cancer (NCT03511196, NCT05393791), melanoma (NCT03543969), rhabdomyosarcoma (NCT04388839) and ovarian cancer (ACTOv/NCT05080556)[1].

On the theoretical side, several mathematical models of tumor containment have been studied (e.g., Martin et al. 1992 [2], Monro and Gaffney 2009 [3], Gatenby et al. 2009 [4], Silva et al. 2012 [5], Carrère 2017 [6], Zhang et al. 2017 [7], Bacevic and Noble et al. 2017 [8], Hansen et al. 2017 [9], Gallaher et al. 2018 [10], Cunningham et al. 2018 [11], Pouchol 2018 [12], Carrère and Zidani 2020 [13], Strobl et al. 2020 [14], Cunningham et al. 2020 [15]) leading to the first workshop on Cancer Adaptive Therapy Models (CATMo; https://catmo2020.org/). However, many of these models make very specific assumptions, e.g., logistic tumor growth with a specific effect of intra-tumor competition and a specific

[1] The initial prostate cancer trial has been debated (Mistry 2021 [81], Zhang et al. 2021 [82])

treatment kill-rate (Zhang et al., 2017 [7], Cunningham et al. 2018 [11], Carrère 2017 [6], Strobl et al. 2020 [14]). This makes it difficult to generalize their conclusions.

Viossat and Noble (2021) [1] recently analysed a more general model with two types of tumor cells: sensitive and fully resistant to treatment. The model takes the form:

$$\begin{aligned}\frac{dS}{dt}(t) &= S(t)g_S(S(t), R(t), L(t)) \\ \frac{dR}{dt}(t) &= R(t)g_R(S(t), R(t))\end{aligned} \tag{Model 1}$$

where $S(t)$ and $R(t)$ are the total number of sensitive and resistant cells at time t, $L(t)$ is the current dose or treatment level, and g_S and g_r are per-cell growth-rate functions. They identified qualitative assumptions under which, among other results, containing the tumor at its initial size maximizes the time at which the tumor becomes larger than at the beginning of treatment (for an idealized form of containment) or is close to maximizing it (for a more realistic form). Similarly, an idealized form of containment at a larger threshold size maximizes the time at which tumor size becomes larger than this threshold. By contrast, eliminating all sensitive cells at treatment initiation - an idealized form of an aggressive treatment - leads to the quickest time to progression beyond any threshold size, among all treatments that eliminate sensitive cells before this threshold size is crossed.

Some of the assumptions of Viossat and Noble are however debatable. In particular, they assume that the higher the number of resistant cells, the lower the growth rate of sensitive cells. Formally, function g_S is non-increasing in R. This assumption helps to compare the size of sensitive populations across treatments. To see why, assume that sensitive cells hamper the growth of resistant cells (that is, g_R is non-increasing in S), and consider two constant dose treatments, with doses L_1 and $L_2 > L_1$, respectively, and the same initial conditions. Suppose that under the treatment with dose L_1 the solution $(S_1(t), R_1(t))$ is obtained, while $(S_2(t), R_2(t))$ follows from the treatment with dose L_2. Since treatment 2 is more aggressive, it initially leads to a smaller sensitive population, hence a larger resistant population than treatment 1: for $t > 0$ small enough, $S_2(t) < S_1(t)$ and $R_2(t) \geq R_1(t)$. If the growth-rate of sensitive cells g_S is non-increasing in R, the fact that treatment 2 is more aggressive and leads to a larger resistant population both negatively affect the sensitive population under treatment 2, ensuring that the sensitive population remains smaller under treatment 2 than under treatment 1: $S_2 < S_1$. This itself ensures that R_2 remains larger than R_1. The inequalities $S_2 < S_1$ and $R_2 \geq R_1$ thus propagate, and hold for all times $t > 0$. By contrast, if the growth-rate of sensitive cells g_S increases with R, the fact that $R_2 \geq R_1$ might boost the growth of sensitive cells under treatment 2, even though treatment 2 is more aggressive. But if the sensitive population becomes larger under treatment 2, the inequality $R_2 \geq R_1$ might also cease to hold, and the whole argument of Viossat and Noble seems to break.

Unfortunately, assuming g_S non-increasing in R, which may seem a natural consequence of competition between tumor cells, is actually problematic. Indeed, it is not satisfied in the Gompertzian model from Monro and Gaffney (2009) [3] that Viossat and Noble use for simulations:

$$\begin{aligned}\frac{dS}{dt}(t) &= \rho \ln(K/N(t))\,(1 - L(t))S(t), \\ \frac{dR}{dt}(t) &= \rho \ln(K/N(t))\,R(t),\end{aligned} \tag{Model 2}$$

where $N(t) = S(t)+R(t)$ is the total number of tumor cells. More precisely, in the absence of treatment ($L(t) = 0$), the growth-rate of sensitive cells is decreasing in N, hence in R; however, if the treatment level is high enough ($L(t) > 1$), the opposite happens, and a large resistant population slows down the regression of the sensitive population.

This reflects the fact that chemotherapy typically attacks cells that are actively dividing. For various reasons (e.g., boundary growth), a larger tumor size is thought to be associated with a lower growth-fraction, i.e., a lower proportion of cells actively dividing (Laird 1964 [83]; Norton and Simon 1977 [84]; Gerlee 2013 [85]). Thus the presence of additional resistant cells, by making the tumor larger, makes more sensitive cells quiescent, and shields them against the effect of treatment. As a result, the growth rate of sensitive cells is not always decreasing in R, and the assumptions of Viossat and Noble are not satisfied. The problem occurs for all Norton-Simon models (Norton and Simon 1977 [84]), where the growth of the sensitive population takes the form:

$$\frac{dS}{dt}(t) = S(t)g(N(t))(1 - L(t))$$

for some per-cell growth rate function g. It also occurs for birth-death models with a Norton-Simon treatment kill-rate (Strobl et al. 2020 [14]):

$$\frac{dS}{dt}(t) = S(t)\left[b(N(t))(1 - L(t)) - d(N(t))\right]$$

where $b(N)$ and $d(N)$ are birth- and death-rates in the absence of treatment.

Another issue is that Model 1 does not consider mutations from sensitive to resistant cells. This is problematic because one of the theoretical motivations for aggressive treatments is to decrease tumor size in order to limit the number of reproduction events, hence of possible appearance of resistant cells by mutation. Key-contributions to the tumor containment literature analyzed the trade-off between increasing competition (by allowing many sensitive cells to survive) and decreasing the number of mutations from sensitive to resistant cells (Martin et al. 1992 [2], Hansen et al. 2017 [9])

The purpose of our work is to generalize the results of Viossat and Noble to models that encompass Norton and Simon models, and, at least to a certain extent, allow for mutations from sensitive to resistant cells. Mathematically, this is achieved by formulating the model in terms of absolute growth-rates and, more importantly, by replacing a direct analysis of the evolution through time of the number of sensitive and resistant cells, $S(t)$ and $R(t)$, by an analysis of the induced trajectory in what we call the $R - N$ plane, where $N = S + R$ describes the total tumor size. These trajectories describe the evolution of the total size N of the tumor as a function of the size R of the resistant population. This turns out to be an efficient technique, allowing to generalize essentially all results of Viossat and Noble, including the optimality or near-optimality of containment treatments.

The remainder of this chapter is organized as follows: the model is described in the next section. Results are presented in Section 1.3, proved in Section 1.4 and discussed in Section 1.5. The Appendix elaborates on the extent to which our model allows for mutations from sensitive to resistant cells, and derives the comparison principle on which our results are based.

1.2 Model

We consider a model with two types of tumors cells: sensitive to treatment, and fully resistant. Their growth is described by differential equations of the form:

$$\begin{aligned} \frac{dS}{dt}(t) &= \phi_S(S(t), R(t), L(t)), \quad && S(0) = S_0 \geq 0 \\ \frac{dR}{dt}(t) &= \phi_R(S(t), R(t)), && R(0) = R_0 > 0 \end{aligned} \qquad \text{(Model 3)}$$

where ϕ_S and ϕ_r are continuously differentiable absolute growth-rate functions. The quantities $\phi_S(0, R, L)$ and $\phi_R(S, 0, L)$ are assumed non-negative to ensure that population sizes cannot become negative. Let $N(t) = S(t) + R(t)$ and $N_0 = S_0 + R_0$. We make the following assumptions:

- The patient dies when tumor size reaches a critical size $N_{crit} > N_0$. [2]
- The size of an untreated tumor increases: $\phi_S(S, R, 0) + \phi_R(S, R) > 0$ if $N \leq N_{crit}$.
- The higher the treatment level, the lower the growth-rate of sensitive cells: ϕ_S is non-increasing in L.
- The resistant population keeps growing: $\phi_R(S, R) > 0$ whenever $R > 0$ and $N \leq N_{crit}$, so that the tumor is incurable if, as we assume, resistant cells are initially present.
- If $R \geq R_0$ and $N \leq N_{crit}$, for a given number of resistant cells, the larger the sensitive population, the lower the growth-rate of resistant cells: ϕ_R is non-increasing in S.

The assumption that the resistant population keeps growing is technical: it ensures that the trajectory of the tumor in the $R - N$ plane (defined formally later on) is the graph of a function. It is no stronger than assuming that the tumor cannot be stabilized for ever. Indeed, if there exists a tumor state (S, R) with $S + R < N_{crit}$ and $\phi_R(S, R) \leq 0$, then from such a state, applying a dose that stabilizes the sensitive population size leads to a regression of the resistant population while the sensitive population stays constant, and so to permanent tumor control. In verbal descriptions, some of the early literature imagined that in an on-off treatment, the resistant population would decline during treatment holidays. This idea, however, was typically inconsistent with the mathematical models used (e.g. Zhang et al. 2017 [7]), and, as explained above, would imply that the tumor could be stabilized for ever. This is not the situation we study here.

The assumption that ϕ_R is non-increasing in S models competition for resources (space, glucose, oxygen) or some other form of inhibition of resistant cells by sensitive cells (Bondarenko et al, 2021 [77]). It neither forbids nor implies a cost of resistance, i.e., that in the absence of treatment, resistant cells grow slower than sensitive cells. In particular, we do not specify whether resistant cells compete more strongly with sensitive cells or with other resistant cells.

The difference with Viossat and Noble (2021) [1] is two-fold: first, the model is formulated in terms of absolute growth-rates, allowing for mutations from sensitive to resistant cells and back. Second, we make no assumption on how the growth-rate of sensitive cells depends on the number of resistant cells. In particular, ϕ_S is not assumed non-increasing in R. This model encompasses many previous models (Silva et al. 2012 [5], Carrère, 2017 [6], Bacevic and Noble et al. 2017 [8], Hansen et al. 2017 [9], Strobl et al. 2020 [14]), including Model 2, its original formulation with mutations (Monro and Gaffney, 2009 [3]), or explicit birth-death models with or without a Norton and Simon treatment effect (Strobl et al. 2020 [14]). Note that while one of our motivations is to include Norton and Simon models as Model 2, our assumptions are actually more general. In particular, sensitive cells may benefit from the presence of resistant cells even in the absence of treatment.

[2] This assumption is standard but debatable (Mistry, 2020). Technically, it ensures that properties of functions ϕ_S and ϕ_R when tumor size is unrealistically large are irrelevant. Conceptually, it could be replaced by the assumption that the treatment goal is to maintain tumor size below a given threshold as long as possible.

To analyse Model 3, it is useful to rewrite it in the equivalent form:

$$\begin{aligned}\frac{dN}{dt}(t) &= f_N(N(t), R(t), L(t)) \\ \frac{dR}{dt}(t) &= f_R(N(t), R(t))\end{aligned} \tag{Model 4}$$

where $f_N(N, R, L) = \phi_S(N - R, R, L) + \phi_R(N - R, R)$ and $f_R(N, R) = \phi_R(N - R, R)$. Our main assumptions are then that, on the domain $R_0 \leq R \leq N \leq N_{crit}$, f_N is non-increasing in L, positive if $L = 0$, and f_R is positive and non-increasing in N. We also assume that the treatment level cannot be larger than a constant L_{max} (the treatment level corresponding to the maximal tolerated dose). Other assumptions are technical:

- f_N and f_R are continuously differentiable (on a neighborhood of the relevant domain: $R_0 \leq R \leq N \leq N_{crit}$ and $0 \leq L \leq L_{max}$).
- $R(t)$ remains smaller than $N(t)$ (this must be biologically, and follows from our assumption on Model 3 that $\phi_S(0, R, L)$ is nonnegative).
- The treatment function $L(\cdot)$ is strongly piecewise continuously differentiable (our vocabulary) in the following sense: there exists a positive integer m and times $t_0 = 0 < t_1 < ... < t_m$ such that, on each interval $[t_k, t_{k+1})$, $k \in \{0, ..., m-1\}$, and on $[t_m, +\infty)$, L coincides with a continuously differentiable function defined on a neighborhood of this interval.

This ensures among other things that, for a given initial condition and treatment, there is a unique solution to Model 4. To fix ideas, we assume that the solutions $R(t)$ and $N(t)$ are defined for all times (though they have no clear interpretation once $N(t) > N_{crit}$), and that they remain bounded. Both properties can be ensured by modifying growth-rate functions f_N and f_R on the domain $N > N_{crit}$. This is without loss of generality since patients are then assumed already deceased.

Outcomes and treatments. We compare the effect of various treatments on the time at which tumor size becomes larger than a given threshold. Depending on this threshold, this may correspond to:

- time to progression, defined as the time at which tumor size progresses beyond its initial size N_0.[3]
- time to treatment failure: the time at which tumor size progresses beyond an hypothetical maximal tolerable tumor size $N_{tol} \geq N_0$, after which the life of the patient is considered at risk or side-effects of the disease are too strong.[4]
- survival time, defined as the time at which tumor size becomes larger than a critical size $N_{crit} \geq N_{tol}$.

[3]In the response evaluation criteria in solid tumors (RECIST), progressive disease is defined by a 20% increase in the sum of the largest diameters (LD) of target lesions, compared to the smallest LD sum recorded since the beginning of treatment. However, comparing to the smallest LD sum recorded would not be fair to aggressive treatments, and the 20% margin makes sense in medical practice, to take into account imperfect monitoring and imperfect forecast of treatment effect, but not for our deterministic mathematical model.

[4]The assumption $N_{tol} \geq N_0$ in without loss of generality in the following sense: if the initial size is larger than the maximal tolerable size, then all treatments we consider would treat at L_{max} until tumor size becomes tolerable ($N = N_{tol}$), and we could apply our analysis from that point on.

Mathematically, results on time to progression and survival time may be obtained through results on time to treatment failure by taking $N_{tol} = N_0$, or $N_{tol} = N_{crit}$, respectively. For this reason, we focus on time to treatment failure.

We consider the following treatments:

• Constant dose treatments, including No treatment (noTreat): $L(t) = 0$, and Maximal Tolerated Dose (MTD): $L(t) = L_{max}$ throughout.

• Delayed MTD (del-MTD): do not treat until $N = N_{tol}$ for the first time, then treat at L_{max} for ever.

• Containment at N_{tol} (Cont): do not treat until $N = N_{tol}$ and then stabilize tumor size at N_{tol}, as long as possible with a treatment level $L(t) \leq L_{max}$. Finally, treat at L_{max} when $N > N_{tol}$. Formally, during the stabilization phase, the treatment level is chosen so that $dN/dt = 0$ (e.g., $L(t) = N(t)/S(t)$ in Model 2). Containment treatments are illustrated in Fig. 1.1, see also Fig. 1 of Viossat and Noble.[5]

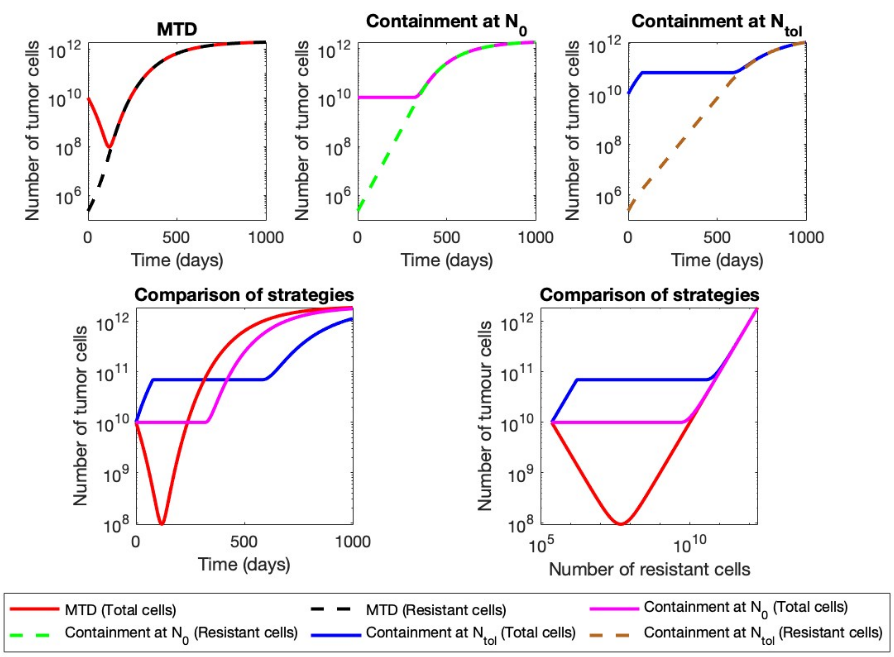


Figure 1.1: Number of resistant cells and number of tumor cells for different treatments. Top row: number of resistant cells and number of tumor cells as a function of time under MTD (left), Containment at the initial size N_0 (center), and Containment at the maximal tolerable size N_{tol} (right). Bottom-row: number of tumor cells under these three treatments as a function of time (left) or as a function of the number of resistant cells (right).

• Intermittent containment (Int), as in the prostate cancer clinical trial of Zhang et al. (2017): do not treat until $N = N_{tol}$, then treat at L_{max} until $N = N_{min} < N_{tol}$, then interrupt treatment until $N = N_{tol}$, and iterate as long as possible. Finally, treat at L_{max} when $N > N_{tol}$. This is illustrated by Fig. 1.2.[6]

[5]If, after crossing N_{tol}, tumor size comes back to N_{tol}, then the containment treatment stabilizes tumor size at N_{tol} again, as long as possible. Similar remarks apply to intermittent containment or other variants of containment.

[6]The above description is to fix ideas: our results are still valid for any other way of maintaining tumor size between N_{min} and N_{tol}, as long as this may be done with a dose $L(t) \leq L_{max}$.

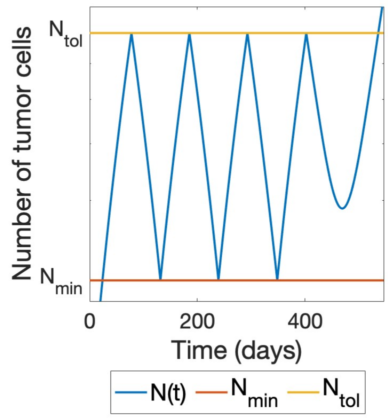


Figure 1.2: Total size under Intermittent treatment.

- An arbitrary treatment, called the alternative treatment (alt): we only assume that $0 \leq L(t) \leq L_{max}$ for all t and $L(t) = L_{max}$ if $N > N_{tol}$.

The times to treatment failure under these treatments will be denoted by $t_{noTreat}$, t_{MTD}, t_{delMTD}, t_{Cont}, t_{Int}, and t_{alt}, respectively.

Following Martin et al. (2012) [86], Hansen et al. (2017) [9], and Viossat and Noble (2021) [1], we also consider idealized treatments, which assume that the sensitive population may be reduced arbitrarily quickly. These treatments are not realistic but are useful theoretical references. Ideal MTD (idMTD) eliminates all sensitive cells instantaneously at the beginning of treatment. Delayed ideal MTD (del-idMTD) lets tumor grow to N_{tol} and then eliminates all sensitive cells. Ideal containment (idCont) lets tumor size grow to N_{tol}, and then stabilizes it as long as some sensitive cells remain. Finally, Ideal intermittent containment (idInt) lets tumor size grow to N_{tol} and then maintains it between $N_{min} \leq N_{tol}$ and N_{tol} as long as some sensitive cells remain.[7]

Under these idealized treatments, treatment fails (i.e. tumor size progresses beyond N_{tol}) when the resistant population reaches size N_{tol}. Sensitive cells have then been fully eliminated. Times to treatment failure are denoted by t_{idMTD}, $t_{del-idMTD}$, t_{idCont}, and t_{idInt}, respectively.[8]

To make our life easy, we assume that all treatments we consider may be implemented through a piecewise continuously differentiable treatment level function $L(t)$ (up to possible downward jumps in the sensitive population for idealized treatments), instead of deriving this result from the implicit function theorem and appropriate regularity assumptions.

1.3 Results

We show that, up to natural additional assumptions for comparisons of sensitive cell populations, all results of Viossat and Noble on Model 1 still hold on Model 3 (or equivalently Model 4), in spite of our less restrictive assumptions. The results are described below and proved in the next section.

[7] Viossat and Noble assumed to fixed ideas and for simulations that, each time tumor size reaches N_{tol}, it drops instantaneously to N_{min}, or to $R(t)$ if $R(t) > N_{min}$, but this is not needed.

[8] When comparing idealized treatments to an alternative treatment, for the comparison to be fair, we do not restrict treatment level under the alternative treatment either, and allow it to eliminate sensitive cells arbitrarily quickly.

The key point is that if treatment level is never larger than a given constant for treatment 1, and never smaller than the same constant for treatment 2, then the resistant population is never larger under treatment 1 than under treatment 2.

Proposition 1.1. *Consider solutions of Model 4 associated to two treatments $L_1(t)$ and $L_2(t)$.*[9] *If there exists a constant $\bar{L}$ such that for all $t \geq 0$, $L_1(t) \leq \bar{L} \leq L_2(t)$, then $R_1(t) \leq R_2(t)$ for all $t \geq 0$.*

It follows that for constant dose treatments, lowering the dose or delaying treatment leads to a lower resistant population:

Proposition 1.2. *(constant dose treatments)*
a) Consider two constant dose treatments $L_1(t) = L_1$ and $L_2(t) = L_2$. If $L_2 \geq L_1$, then $R_1(t) \leq R_2(t)$ for all $t \geq 0$.[10]

b) Assume that $L_1(t) = L > 0$ for all $t \geq 0$, while $L_2(t) = 0$ until $N = N_{start} \geq N_0$, and then $L_2(t) = L$. Then $R_1(t) \leq R_2(t)$ for all $t \geq 0$.

Proposition 1.1 also implies that not treating minimizes the resistant population while MTD maximizes it:

Proposition 1.3. *(MTD maximizes resistance)*
For all $t \geq 0$, $R_{noTreat}(t) \leq R_{alt}(t) \leq R_{MTD}(t)$.

Of course, not treating is typically not an option, as the number of sensitive cells would explode, but containment is. One of our main results is that containment minimizes the resistant population among all treatments treating at L_{max} after failing.

Proposition 1.4. *(containment minimizes resistance)*
For all $t \geq 0$, $R_{Cont}(t) \leq R_{alt}(t)$

It follows that $N_{Cont} \leq N_{alt} + (S_{Cont} - S_{alt})$. Thus, assuming that the tumor is eventually mostly resistant under the containment treatment, tumor size should eventually be smaller, or at least not substantially larger under the containment treatment than under any alternative one. This suggests that, under our assumptions, among treatments that treat at L_{max} when $N > N_{tol}$, containment should be close to maximizing survival time. Similarly, the fact that the resistant population is larger under MTD that under any alternative treatment suggests that most alternative treatments should eventually lead to a lower tumor size and a longer survival time than MTD.

More precise statements may be made for idealized versions of containment and MTD: ideal containment maximizes time to treatment failure, while ideal MTD minimizes it among all treatments eliminating sensitive cells before failing. Moreover, ideal containment eventually leads to a lower tumor size and ideal MTD to a larger tumor size than any such alternative treatment.

Proposition 1.5. *(comparison with ideal MTD and ideal containment)*
a) $t_{alt} \leq t_{idCont}$.

[9] Unless mentioned otherwise, when comparing two treatments, we assume the same initial conditions: $R_1(0) = R_2(0)$ and $N_1(0) = N_2(0)$.

[10] Exceptionally, we assume that even when $N_i > N_{tol}$, the dose stays equal to L_i and is not increased to L_{max}. A similar remark applies to b)

b) Consider an alternative treatment eliminating sensitive cells before failing, that is, such that $S_{alt}(t_{alt}) = 0$. Then:

b1) $t_{alt} \geq t_{idMTD}$;

b2) for all $t \geq 0$, $R_{idCont}(t) \leq R_{alt}(t) \leq R_{idMTD}(t)$;

b3) for all $t \geq t_{alt}$, $N_{idCont}(t) \leq N_{alt}(t) \leq N_{idMTD}(t)$.

In particular, survival time is larger with ideal containment and lower with ideal MTD than with any alternative treatment such that $S_{alt}(t_{alt}) = 0$.

The next result shows that intermittent containment between N_{min} and $N_{tol} > N_{min}$ leads to outcomes that are intermediate between those of containment at the larger threshold N_{tol} and those of containment at the lower threshold N_{min} (ContNmin). The latter lets tumor size grow until $N = N_{min}$ (or treats at L_{max} until $N = N_{min}$ if $N_0 > N_{min}$), and then stabilizes tumor size at N_{min} as long as possible with a treatment level $L(t) \leq L_{max}$. In the idealized form, ideal containment at N_{min}, tumor size is stabilized at N_{min} as long as some sensitive cells remain (and initially instantly reduced to the maximum of N_{min} and R_0, if $N_{min} > N_0$).

Proposition 1.6. *(dose modulation versus treatment vacation)*
a) For all $t \geq 0$, $R_{Cont}(t) \leq R_{Int}(t) \leq R_{ContNmin}(t)$, and similarly, $R_{idCont}(t) \leq R_{idInt}(t) \leq R_{idContNmin}(t)$.
b) $t_{idContNmin} \leq t_{idInt} \leq t_{idCont}$
c) For all $t \geq t_{idInt}$, $N_{idCont}(t) \leq N_{idInt}(t) \leq N_{idContNmin}(t)$.

This result suggests that, if the lower threshold N_{min} is close to the larger threshold N_{tol}, there should be little difference between outcomes of containment and intermittent containment, that is, between a continuous low dose treatment based on dose modulation and an intermittent high dose treatment based on treatment vacation. Of course, this disregards many possible differences between these two approaches. For instance, dose-modulation might lead to a more regular vascularization of the tumor, which might be key for an efficient drug delivery (Enriquez-Navas et al., 2016 [74]).

We now compare all reference treatments.

Proposition 1.7. *(comparison between all reference treatments)*
a) For all $t \geq 0$:
a1) $R_{noTreat}(t) \leq R_{Cont}(t) \leq R_{Int}(t) \leq R_{del-MTD}(t) \leq R_{MTD}(t) \leq R_{idMTD}(t)$
and a2) $R_{noTreat}(t) \leq R_{idCont}(t) \leq R_{idInt}(t) \leq R_{del-idMTD}(t) \leq R_{idMTD}(t)$
b) $t_{idMTD} \leq t_{del-idMTD} \leq t_{idInt} \leq t_{idCont}$
c) For all $t \geq t_{idCont}$, $N_{idCont}(t) \leq N_{idInt}(t) \leq N_{del-idMTD}(t) \leq N_{idMTD}(t)$.

Sensitive population sizes may also be compared under two mild additional assumptions:

(A1) Not treating maximizes the sensitive population.

(A2) The sensitive population decreases if the tumor is treated at L_{max}.

These assumptions hold for Model 2, assuming $L_{max} \geq 1$, and for most models we are aware of. They lead to the same comparison for sensitive population sizes as in Viossat and Noble, that is, the opposite as for resistant population sizes.[11]

[11] For Model 2, (A2) holds obviously if $L_{max} \geq 1$. To see that (A1) holds, note that $dN/dt = \rho \ln(K/N)N - \rho \ln(K/N)LS \geq \rho \ln(K/N)N$ with equality for $L = 0$. The comparison principle thus implies that not treating maximizes tumor size. Since not treating also minimizes the resistant population size (Proposition 1.3), it follows that it maximizes the sensitive population size.

Proposition 1.8. *(comparison of sensitive populations)*
Assume that (A1) and (A2) hold. Then for all $t \geq 0$:[12]
a) $S_{idMTD}(t) \leq S_{MTD}(t) \leq S_{alt}(t) \leq S_{Cont}(t) \leq S_{noTreat}(t)$
b) $S_{ContNmin}(t) \leq S_{Int}(t) \leq S_{Cont}(t)$ *and* $S_{del-MTD}(t) \leq S_{Int}(t)$
c) $S_{idContNmin}(t) \leq S_{idInt}(t) \leq S_{idCont}(t)$
d) $S_{idMTD}(t) \leq S_{del-idMTD}(t) \leq S_{idInt}(t) \leq S_{idCont}(t) \leq S_{noTreat}(t)$

1.4 Proofs

Viossat and Noble's proofs build on their Proposition 1, which gives conditions allowing them to compare the resistant populations or the sensitive populations under two different treatments. The following part of this result is still true in our framework, with the same proof:

Lemma 1.1. *Let $0 \leq t_0 \leq t_1$. Consider two solutions (S_1, R_1) and (S_2, R_2) of Model 3, associated to treatment functions L_1 and L_2, respectively. Assume that: i) $R_1(t_0) \leq R_2(t_0)$, and ii) $S_1(t_0) \geq S_2(t_0)$. If: iiia) $S_1(t) \geq S_2(t)$ on $[t_0, t_1]$, or: iiib) $N_1(t) \geq N_2(t)$ on $[t_0, t_1]$, then $R_1(t) \leq R_2(t)$ and $S_1(t) \geq S_2(t)$ on $[t_0, t_1]$.*

What is no longer true is that the same conclusions hold if iiia) or iiib) is replaced by iiic): $L_1(t) \leq L_2(t)$ for all t. For instance, in Model 2, if $L_1(t_0) = L_2(t_0) > 1$, $R_1(t_0) < R_2(t_0)$ and $S_1(t_0) = S_2(t_0)$, then S_2 becomes immediately larger than S_1. This will slow down the growth of the resistant population under treatment 2. Thus, conceivably, R_2 could later on become smaller than R_1.

We thus use a new proof technique. Instead of studying directly the evolution of the resistant population R, the sensitive population S, or the total tumor size N as a function of time, we first study, and compare across treatments, the evolution of tumor size N as a function of the number of resistant cells. In other words, we compare trajectories in the $R - N$ plane, that is, the sets of points $(R(t), N(t))$ for all $t \geq 0$.

To be more formal, fix a treatment L, and let $R^\infty = \lim_{t \to +\infty} R(t)$. Since the resistant population increases continuously, for any $r \in [R_0, R^\infty)$, there exists a unique time $t(r)$ at which the resistant population has size r, that it, $R(t(r)) = r$. Denote by $\tilde{S}(r)$, $\tilde{N}(r) = \tilde{S}(r) + r$, and $\tilde{L}(r)$, the number of sensitive cells, the total number of tumor cells, and the treatment level at time $t(r)$, that is, when the resistant population reaches size r. All these functions may be shown to be piecewise continuously differentiable, and $\tilde{S}$ and $\tilde{N}$ are also continuous. The graph of function $\tilde{N}$ coincides with the trajectory of the solution in the $R - N$ plane. It may be analyzed by noting that function $\tilde{N}$ satisfies the differential equation:

$$\frac{d\tilde{N}}{dr} = G(\tilde{N}, r) \text{ where } G(\tilde{N}, r) = \frac{f_N(\tilde{N}, r, \tilde{L}(r))}{f_R(\tilde{N}, r)}.$$

Trajectories in the $R - N$ plane, and their connections to the evolution of tumor size and of the resistant population as a function of time are illustrated in Figure 1.1.

1.4.1 Key lemmata

Our first result shows that if, for any resistant population level r, tumor size is larger under treatment 1 than under treatment 2, then at any time t, the resistant population is smaller under treatment 1

[12] The only inequality that uses (A1) is $S_{alt} \leq S_{Cont}$.

than under treatment 2. The intuition is the following: at the time $t_i(r)$ when the resistant population reaches size r under treatment i, the speed at which the resistant population increases is given by:

$$\frac{dR_i}{dt}(t_i(r)) = f_R(N_i(t_i(r)), R_i(t_i(r)) = f_R(\tilde{N}_i(r), r).$$

Since f_R is non-increasing in N, it follows that if $N_1(r) \geq N_2(r)$, the resistant population will increase quicker from r to $r + dr$ under treatment 2 than under treatment 1 (dr is a small positive increment). If this holds for all resistant population sizes r, then R_2 will remain no-smaller than R_1 at all times $t \geq 0$.

Lemma 1.2. *Let $L_1(t)$ and $L_2(t)$ be two different treatments. Consider solutions (N_1, R_1) and (N_2, R_2) of Model 4 associated to these treatments such that $R_1(0) = R_2(0) = R_0$. If $\tilde{N}_1(r) \geq \tilde{N}_2(r)$ for all r in $[R_0, \min\{R_1^\infty, R_2^\infty\})$, then $R_1^\infty \leq R_2^\infty$ and $R_1(t) \leq R_2(t)$ for all $t \geq 0$.*

Moreover, if on an interval $[r_1, r_2]$, $\tilde{S}_1(r)$ is non-increasing or $\tilde{S}_2(r)$ is non-increasing, then $S_1(t) \geq S_2(t)$ for all t in $[t_1(r_1), t_2(r_2)]$.

Proof. Consider a time $t \geq 0$ such that $R_1(t) < R_2^\infty$, so that $\tilde{N}_2(R_1(t))$ is well defined. Since $N(t) = \tilde{N}(R(t))$, $\tilde{N}_1 \geq \tilde{N}_2$ and f_N is non-increasing in N, we obtain:

$$\begin{aligned}\frac{dR_1}{dt}(t) = f_R(N_1(t), R_1(t)) = f_R(\tilde{N}_1(R_1(t)), R_1(t)) \leq & f_R(\tilde{N}_2(R_1(t)), R_1(t)) \\ =: & F(R_1(t))\end{aligned}$$

while

$$\frac{dR_2}{dt}(t) = f_R(\tilde{N}_2(R_2(t)), R_2(t)) = F(R_2(t)).$$

Since $R_1(0) = R_2(0)$, the comparison principle (Proposition 1.10, item b), in Appendix 1.6.2) implies that for all times $t \geq 0$ such that $R_1(t) < R_2^\infty$, we have $R_1(t) \leq R_2(t)$ (See Figure 1.3).

We now show that the inequality $R_1(t) < R_2^\infty$, hence the conclusion $R_1(t) \leq R_2(t)$, holds at all times $t \geq 0$. Indeed, otherwise there is a first time $t^* \geq 0$ such that $R_1(t^*) = R_2^\infty$, and $t^* > 0$. Since R_i is increasing, it follows that on $[0, t^*)$, $R_1(t) \leq R_2(t) \leq R_2(t^*) < R_2^\infty$. By continuity of R_1, this implies that $R_1(t^*) \leq R_2(t^*) < R_2^\infty$, a contradiction.

We now prove the result on sensitive cells. Assume that on $[r_1, r_2]$, $\tilde{S}_1(r)$ is non-increasing (which implicitly requires $r_2 < R_1^\infty$, so that $\tilde{S}_1(r)$ is well-defined on $[r_1, r_2]$). For all t in $[t_1(r_1), t_2(r_2)]$, $r_1 \leq R_1(t) \leq R_2(t) \leq r_2$. Moreover, the assumption $\tilde{N}_1(r) \geq \tilde{N}_2(r)$ is equivalent to $\tilde{S}_1(r) \geq \tilde{S}_2(r)$. Thus we obtain:

$$S_1(t) = \tilde{S}_1(R_1(t)) \geq \tilde{S}_1(R_2(t)) \geq \tilde{S}_2(R_2(t)) = S_2(t).$$

The first inequality follows from the fact that $\tilde{S}_1$ is non-increasing on $[r_1, r_2]$, the second from the fact that $\tilde{S}_1 \geq \tilde{S}_2$. If it is $\tilde{S}_2$ which is non-increasing (which implicitly requires $r_2 < R_2^\infty$, so that $\tilde{S}_2$ is well-defined on $[r_1, r_2]$), then:

$$S_1(t) = \tilde{S}_1(R_1(t)) \geq \tilde{S}_2(R_1(t)) \geq \tilde{S}_2(R_2(t)) = S_2(t).$$

The first inequality follows from the fact that $\tilde{S}_1 \geq \tilde{S}_2$, the second from the fact that $\tilde{S}_2$ is non-increasing on $[r_1, r_2]$. □

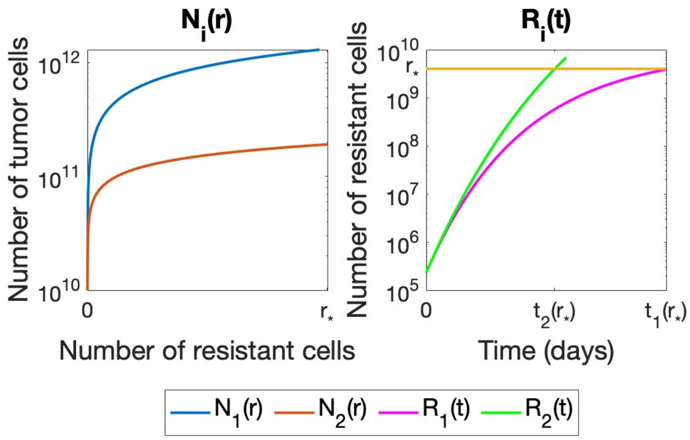


Figure 1.3: Using two different constant treatment levels on the Gompertzian model as an example, this figure illustrates that the relation $\tilde{N}_1 \geq \tilde{N}_2$ over the interval $[R_0, r_*)$, $r_* = 4 \times 10^9$ (Left panel), translates into $R_1(t) \leq R_2(t)$ over the interval $[0, \min\{t_1(r_*), t_2(r_*)\}]$ (Right panel). The same behavior will be observed for an arbitrary choice of r_*.

Assume now that for any resistant population level r, the treatment level when the resistant population reaches size r is lower for treatment 1 than for treatment 2. Our second result shows that the tumor size when the resistant population reaches size r is then always larger for treatment 1 than for treatment 2. By the previous lemma, this implies that the resistant population is always smaller under treatment 1 than under treatment 2.

Lemma 1.3. *Let $L_1(t)$ and $L_2(t)$ be two different treatments such that*

$$L_1(t) \leq \overline{L} \leq L_2(t),$$

for a certain positive number $\overline{L}$, or more generally such that $\tilde{L}_1(r) \leq \tilde{L}_2(r)$ for all r in $[R_0, R^*)$ where $R^* = \min\{R_1^\infty, R_2^\infty\}$. Consider solutions of Model 4 associated to these treatments such that $R_1(0) = R_2(0)$ and $N_1(0) \geq N_2(0)$. Then $\tilde{N}_1(r) \geq \tilde{N}_2(r)$ for all r in $[R_0, R^*)$. Therefore by Lemma 1.2, $R_1^\infty \leq R_2^\infty$ and $R_1(t) \leq R_2(t)$ for all $t \geq 0$.

Proof. Since f_N is non-increasing in L, and for all r in $[R_0, R^*)$, $\tilde{L}_1(r) \leq \tilde{L}_2(r)$, we get:

$$\frac{d\tilde{N}_1}{dr} = \frac{f_N(\tilde{N}_1, r, \tilde{L}_1(r))}{f_R(\tilde{N}_1, r)} \geq \frac{f_N(\tilde{N}_1, r, \tilde{L}_2(r))}{f_R(\tilde{N}_1, r)} =: G_2(\tilde{N}_1, r)$$

while $\dfrac{d\tilde{N}_2}{dr} = G_2(\tilde{N}_2, r)$. Moreover, $\tilde{N}_2(R_0) = N_2(0) \leq N_1(0) = \tilde{N}_1(R_0)$. Therefore, by the comparison principle (Proposition 1.10, item a), in Appendix 1.6.2), $\tilde{N}_1(r) \geq \tilde{N}_2(r)$ for all r in $[R_0, R^*)$. Then apply Lemma 1.2. □

1.4.2 Proof of propositions 1.1 to 1.8

Proposition 1.1 follows from Lemma 1.3, and Propositions 1.2 and 1.3 from Proposition 1.1.

Proof of Proposition 1.4. For later purposes, let us prove a more general result: For all $t \geq 0$,

$$R_{noTreat}(t) \leq R_{Cont}(t) \leq R_{alt}(t) \leq R_{MTD}(t) \leq R_{idMTD}(t). \quad (1.1)$$

This follows from Lemma 1.2 and the fact that, whenever these comparisons make sense:

$$\tilde{N}_{idMTD}(r) \leq \tilde{N}_{MTD}(r) \leq \tilde{N}_{alt}(r) \leq \tilde{N}_{Cont}(r) \leq \tilde{N}_{noTreat}(r) \quad (1.2)$$

To prove (1.2), note that for any alternative treatment, $\tilde{N}_{idMTD}(r) = r \leq \tilde{N}_{alt}(r)$, in particular, $\tilde{N}_{idMTD}(r) \leq \tilde{N}_{MTD}(r)$, and by Lemma 1.3 with $\bar{L} = 0$, $\tilde{N}_{alt}(r) \leq \tilde{N}_{noTreat}(r)$, in particular $\tilde{N}_{Cont}(r) \leq \tilde{N}_{noTreat}(r)$. Moreover, under the constraint $L_{alt}(t) \leq L_{max}$, it follows from Lemma 1.3 with $\bar{L} = L_{max}$ that $\tilde{N}_{MTD}(r) \leq \tilde{N}_{alt}(r)$.

It remains to prove that $\tilde{N}_{alt}(r) \leq \tilde{N}_{Cont}(r)$ for all $r \in [R_0, R^*)$, where $R^* = \min\{R^\infty_{alt}, R^\infty_{Cont}\}$. The notation we introduce is illustrated in Fig. 1.4. Let $r_1 = \min\{r \geq R_0, \tilde{N}_{Cont}(r) = N_{tol}\}$. When $r \leq r_1$, $\tilde{N}_{Cont}(r) = \tilde{N}_{noTreat}(r) \geq \tilde{N}_{alt}(r)$ as explained above. Moreover, for all $r \geq r_1$, $\tilde{N}_{Cont}(r) \geq N_{tol}$. Thus, assuming by contradiction that there exists $r_2 \geq r_1$ such that $\tilde{N}_{Cont}(r_2) < \tilde{N}_{alt}(r_2)$, it follows that $\tilde{N}_{alt}(r_2) > N_{tol}$. Let

$$r_{max} = \max\{r \leq r_2, \tilde{N}_{alt}(r) \leq N_{tol}\}.$$

Note that since $\tilde{N}_{alt}(r_1) < \tilde{N}_{Cont}(r_1) = N_{tol}$, we must have $r_{max} \geq r_1$. Therefore, $\tilde{N}_{alt}(r_{max}) = N_{tol} \leq \tilde{N}_{Cont}(r_{max})$. Moreover, on (r_{max}, r_2), $\tilde{N}_{alt}(r) > N_{tol}$, hence $\tilde{L}_{alt}(r) = L_{max} \geq \tilde{L}_{Cont}(r)$. By a variant of Lemma 1.3 (comparing treatments starting when the initial resistant population size is r_{max} rather than R_0), it follows that $\tilde{N}_{alt}(r_2) \leq \tilde{N}_{Cont}(r_2)$, a contradiction.

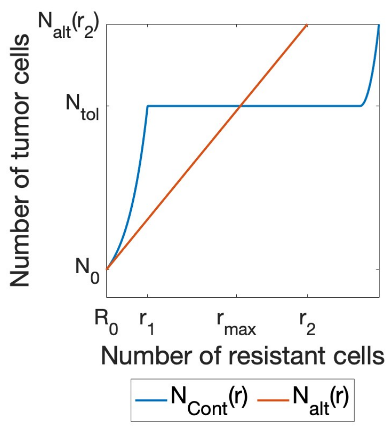


Figure 1.4: Comparison of $N_{Cont}(r)$ with an hypothetical curve which satisfies $N_{alt}(r_2) > N_{Cont}(r_2)$ for some point $r_2 \geq r_1$.

Proof of Proposition 1.5. Proof of a): Let $t_1 = \min\{t \geq 0, N_{Cont}(t) = N_{tol}\}$. For $t \leq t_1$, $R_{idCont}(t) = R_{noTreat}(t) \leq R_{alt}(t)$ (the inequality follows from Proposition 1.3). If $t_{alt} \geq t_1$ then, as in Viossat and Noble, on $[t_1, t_{alt}]$, $N_{idCont}(t) \geq N_{tol} \geq N_{alt}(t)$ so $R_{idCont}(t) \leq R_{alt}(t)$ by Lemma 1.1. Thus, $R_{idCont}(t_{alt}) \leq R_{alt}(t_{alt}) \leq N_{alt}(t_{alt}) = N_{tol}$. It follows that $t_{idCont} \geq t_{alt}$, since ideal containment fails when $R_{idCont} = N_{tol}$.

Proof of b): assume now that $S_{alt}(t_{alt}) = 0$ (which only makes sense for idealized alternative treatments).

Then $R_{alt}(t_{alt}) = N_{alt}(t_{alt}) = N_{tol}$. Thus, as in Viossat and Noble:

$$N_{tol} = N_{alt}(t_{alt}) = R_{alt}(t_{alt}) \leq R_{idMTD}(t_{alt}),$$

so $t_{idMTD} \leq t_{alt}$. This proves b1).

Let us prove the remaining results on ideal MTD. The inequality $R_{alt} \leq R_{idMTD}$ was shown in the proof of Proposition 1.4 (see Eq. 1.1). Moreover, on $[t_{idMTD}, t_{alt}]$, $N_{idMTD} = R_{idMTD} \geq N_{tol} \geq N_{alt}$, and for $t \geq t_{alt}$, $N_{idMTD} = R_{idMTD} \geq R_{alt} = N_{alt}$. This proves parts of b2) and b3).

We now prove the results on ideal containment. On $[t_{alt}, t_{idCont}]$,

$$R_{idCont}(t) \leq N_{idCont}(t) \leq N_{tol} = N_{alt}(t_{alt}) = R_{alt}(t_{alt}) \leq R_{alt}(t) = N_{alt}(t)$$

where the last inequality comes from the fact that for all $t \geq t_{alt}$, $S_{alt}(t) = 0$. Moreover, after treatment failure, R_{alt} and R_{idCont} both satisfy the autonomous equation $dR/dt = f_R(0, R)$. By invariance of solutions of autonomous equations through translation in time, this implies that for all $t \geq t_{idCont}$, $R_{idCont}(t) = R_{alt}(t - [t_{idCont} - t_{alt}]) \leq R_{alt}$. For $t \leq t_{idCont}$, the inequality $R_{idCont}(t) \leq R_{alt}(t)$ was derived in the proof of a). Therefore, $R_{idCont}(t) \leq R_{alt}(t)$ for all $t \geq 0$. Finally, for all t in $[t_{alt}, t_{idCont}]$, $N_{idCont}(t) \leq N_{tol} = R_{alt}(t_{alt}) \leq R_{alt}(t) = N_{alt}(t)$, while for all $t \geq t_{idCont}$, $N_{idCont}(t) = R_{idCont}(t) \leq R_{alt}(t) = N_{alt}(t)$. This completes the proof.

Proof of Proposition 1.6. Proof of a): The inequalities $R_{Cont} \leq R_{Int}$ and $R_{idCont} \leq R_{idInt}$ follow from the proof of Proposition 1.4 (see Eq. (1.1)) and from Proposition 1.5. The fact that $R_{Int} \leq R_{ContNmin}$ follows from Lemma 1.2 and the fact that, as shown below: for all r, $\tilde{N}_{ContNmin}(r) \leq \tilde{N}_{Int}(r)$. To prove this, note that for $r \leq r_{min} := \min\{r \geq R_0, \tilde{N}_{noTreat}(r) = N_{tol}\}$, both treatments coincide so $\tilde{N}_{ContNmin}(r) = \tilde{N}_{Int}(r)$. For $r \geq r_{min}$, the argument is as in the proof of $\tilde{N}_{Cont}(r) \geq \tilde{N}_{alt}(r)$ for $r \geq r_1$ in Proposition 1.4. Similarly, it is easily seen that $\tilde{N}_{idContNmin}(r) \leq \tilde{N}_{idInt}(r)$ for all r, so $R_{idInt} \leq R_{idContNmin}$ by Lemma 1.2.

Proof of b): $N_{tol} = N_{idInt}(t_{idInt}) = R_{idInt}(t_{idInt}) \leq R_{idContNmin}(t_{idInt})$ by a), hence $t_{idContNmin} \leq t_{idInt}$. The second inequality follows from item a) of Proposition 1.5.

Proof of c): Using a), for $t \geq t_{idInt}$, $N_{idInt} = R_{idInt} \leq R_{idContMin} = N_{idContMin}$, and the first inequality follows from item b3) of Proposition 1.5.

Proof of Proposition 1.7. Proof of a1): to see that $R_{Int}(t) \leq R_{del-MTD}(t)$, note that as long as tumor size is lower than N_{tol}, both treatments coincide, then apply Proposition 1.3 from that point on. The other inequalities have already been proved. The proof of a2) is similar.

Proofs of b) and c): in b), the inequality $t_{del-idMTD} \leq t_{idInt}$ follows from item b1) of Proposition 1.5, applied from the (common) time when tumor size reaches N_{tol} under both treatments, other inequalities were shown already. The proof of c) is as in Viossat and Noble.

Proof of Proposition 1.8. We first need a lemma.

Lemma 1.4. *a) Let $N^* \geq 0$. Consider a solution (N, R) of Model 4 under a treatment such that $L(t) = L_{max}$ whenever $N(t) > N^*$. Let $\bar{t} \geq 0$ be such that $N(\bar{t}) \geq N_{tol}$. If the sensitive population decreases when treated at L_{max}, then for all $t \geq \bar{t}$, $S(t) \leq S(\bar{t})$. If moreover $N(t) \geq N^*$ for all $t \geq \bar{t}$, then S is non-increasing on $[\bar{t}, +\infty)$.*

b) Under containment (respectively, containment at N_{min}), once tumor size reaches N_{tol} for the first time (respectively, N_{min}), the sensitive population is non-increasing.

Proof. a) The idea is that when $N > N^*$, S is non-increasing by assumption, and when $N(t) \leq N^*$ for $t > \bar{t}$, the sensitive population must have decreased since time $\bar{t}$ because the resistant population increased (by assumption) and total tumor size did not. Formally, let $t \geq \bar{t}$. If for all τ in $(\bar{t}, t)$, $N(\tau) > N^*$, hence $L(\tau) = L_{max}$, then S is non-increasing on $[\bar{t}, t]$ by assumption, therefore $S(t) \leq S(\bar{t})$. Otherwise, let $t_{max} = \max\{\tau \leq t, N(\tau) \leq N^*\}$. The previous argument implies that $S(t) \leq S(t_{max})$. Moreover, since R is increasing,

$$S(t_{max}) = N(t_{max}) - R(t_{max}) \leq N_{tol} - R(t_{max}) \leq N(\bar{t}) - R(\bar{t}) = S(\bar{t})$$

Therefore, $S(t) \leq S(\bar{t})$. Finally, if for any $t_1 \geq \bar{t}$, $N(t_1) \geq N^*$, then the previous result applied from t_1 on shows that for any $t_2 \geq t_1$, $S(t_2) \leq S(t_1)$, hence S is non-increasing on $[\bar{t}, \infty)$.

b) For containment, this follows from a) with $N^* = N_{tol}$ and the fact that once tumor size reaches N_{tol} under containment, it never becomes smaller. The proof for containment at N_{min} is the same with N_{min} replacing N_{tol}. □

We now prove Proposition 1.8. Proof of a): the first inequality is trivial since $S_{idMTD} = 0$ (we only mentioned it to show that all inequalities from Eq. 1.1 are reversed). The inequality $S_{MTD}(t) \leq S_{alt}(t)$ follows from Lemma 1.2, the fact that $\tilde{N}_{MTD}(r) \leq \tilde{N}_{alt}(r)$ (see Eq. (1.2)), and the fact that $\tilde{S}_{MTD}(r)$ is non-increasing by Assumption (A2). The last inequality follows from Assumption (A1), or, independently of (A1), from Lemma 1.2, the fact that $\tilde{N}_{Cont}(r) \leq \tilde{N}_{noTreat}(r)$ (see Eq. (1.2)), and that once Containment starts treating, S_{Cont} is non-increasing (Lemma 1.4, item b)).

Let us now prove that $S_{alt}(t) \leq S_{Cont}(t)$. Let $r_1 = \min\{r \geq R_0, \tilde{N}_{Cont}(r) = N_{tol}\}$. For $t \leq t_{Cont}(r_1)$, containment does not treat so $S_{Cont}(t) = S_{noTreat}(t) \geq S_{alt}(t)$ by Assumption (A1). Moreover, on $[r_1, R^\infty_{Cont})$, $\tilde{N}_{Cont}(r) \geq \tilde{N}_{alt}(r)$ and $\tilde{N}_{Cont}(r) \geq N_{tol}$, therefore $\tilde{S}_{Cont}(r)$ is non-increasing by Lemma 1.4. Thus, by Lemma 1.2, $S_{Cont}(t) \geq S_{alt}(t)$ for any t in $[t_{Cont}(r_1), t_{alt}(R^\infty_{Cont})\,)$.

Finally, let $t \geq \max(t_{Cont}(r_1), t_{alt}(R^\infty_{Cont}))$, that is, such that $R_{Cont}(t) \geq r_1$ and $R_{alt}(t) \geq R^\infty_{Cont}$. Since $N_{Cont}(t) \geq N_{tol}$ for all $t \geq t(r_1)$, it follows from Lemma 1.4 that S_{Cont} is non-increasing on $[t(r_1), +\infty[$, so $S_{Cont}(t) \geq S^\infty_{Cont}$. Thus, it suffices to show that $S_{alt}(t) \leq S^\infty_{Cont}$. There are two cases.

Case 1: If $\tilde{N}_{alt}(R^\infty_{Cont}) \geq N_{tol}$, then by Lemma 1.4, for all $t \geq t_{alt}(R^\infty_{Cont})$,

$$S_{alt}(t) \leq S_{alt}(t_{alt}(R^\infty_{Cont})) = \tilde{S}_{alt}(R^\infty_{Cont}) \leq S^\infty_{Cont}$$

where the last inequality follows from the fact that for $r < R^\infty_{Cont}$, $\tilde{S}_{alt}(r) \leq \tilde{S}_{Cont}(r)$ due to Eq. (1), so that

$$\tilde{S}_{alt}(R^\infty_{Cont}) = \lim_{r \to R^\infty_{Cont}} \tilde{S}_{alt}(r) \leq \lim_{r \to R^\infty_{Cont}} \tilde{S}_{Cont}(r) = \lim_{t \to +\infty} S_{Cont}(t) = S^\infty_{Cont}$$

Case 2: If $\tilde{N}_{alt}(R^\infty_{Cont}) < N_{tol}$, then as long as $N_{alt}(t) \leq N_{tol}$,

$$S_{alt}(t) \leq N_{tol} - R_{alt}(t) \leq N_{tol} - R^\infty_{Cont} \leq S^\infty_{Cont}$$

Moreover, if at some time $\bar{t}$, $N_{alt}(\bar{t}) = N_{tol}$ (which must indeed happen), then $S_{alt}(\bar{t}) \leq S^\infty_{Cont}$ by the previous argument, and for all $t \geq \bar{t}$, by Lemma 1.4, $S_{alt}(t) \leq S_{alt}(\bar{t}) \leq S^\infty_{Cont}$. This concludes the proof of a).

Proof of b): The inequality $S_{ContNmin}(t) \leq S_{Int}(t)$ follows from Lemma 1.2, the fact that $\tilde{N}_{contNmin}(r) \leq \tilde{N}_{Int}(r)$, and the fact that once tumor size reaches N_{min}, $S_{ContNmin}$ is non-increasing (Lemma 1.4, item b)). The proof of $S_{Int}(t) \leq S_{Cont}(t)$ is as the proof of $S_{alt}(t) \leq S_{Cont}(t)$ (except that Assumption (A1) is not needed). Finally, the inequality $S_{del-MTD} \leq S_{Int}$ follows from $S_{MTD} \leq S_{alt}$ applied from the time at which tumor size reaches N_{tol}.

Proof of c): we first prove $S_{idContNmin} \leq S_{idInt}$. Before tumor size reaches N_{min}, both treatments coincide, then until $t_{idContNmin}$, $N_{idContNmin} = N_{min} \leq N_{idInt}$ while $R_{idContNmin} \geq R_{idInt}$, so $S_{idContNmin} \leq S_{idInt}$. Finally, for $t \geq t_{idContNmin}$, $S_{idContNmin}(t) = 0 \leq S_{idInt}(t)$. The proof of $S_{idInt} \leq S_{idCont}$ is similar.

Proof of d): the first two inequalities are trivial, the third one was proved in c). The last inequality follows from (A1) but also, independently of (A1), from the following argument: for $t \leq t_{idCont}$, $N_{idCont} \leq N_{noTreat}$ while $R_{idCont} \geq R_{noTreat}$ by Proposition 1.7, so $S_{idCont} \leq S_{noTreat}$, and for $t \geq t_{idCont}$, $S_{idCont} = 0$.

1.5 Discussion

Viossat and Noble [1] provided qualitative conditions ensuring that a strategy aiming at containment, not elimination, minimizes resistance to treatment and is close to maximizing time to treatment failure. Some of these conditions were however debatable. In particular, their analysis did not allow for mutations from sensitive to resistant cells, a major concern of some key contributions to the field (Martin et al. 1992 [2], Hansen et al. 2017 [9]), and did not apply to Norton-Simon models [84], which are standard to model chemotherapy. We showed how a refined analysis allows to handle these two issues. This suggests that containment strategies are likely to perform well in more general situations than was previously known.

While Viossat and Noble compared across treatments the values of resistant and sensitive populations as a function of time, we first compare the induced trajectories in the R-N plane, that is, tumor sizes not at a given time, but when the resistant population reaches a given size. We made the additional assumption that the resistant population keeps increasing. This is no stronger than assuming that the tumor cannot be stabilized for ever, and is technically helpful (as the trajectory in the R-N plane is then the graph of a function), but we conjecture that our results hold without this assumption.

What is crucial is that, all else being equal, a larger sensitive population leads to a lower resistant population growth-rate. For this reason, our analysis only allows for mutations from sensitive to resistant cells if an increase in the sensitive population size is more detrimental to the growth of the resistant population (through competition, or some other form of inhibition of resistant cells by sensitive cells) than it is beneficial (through mutations from sensitive to resistant cells). We show in Appendix 1.6.1 that this is typically the case for Gompertzian growth, or power-law models, at least in the absence of a strong resistance cost.[13]

There are however many other concerns with containment. Mutations, or phenotypic switching, could be modeled in other ways, and the fact that maintaining a relatively large tumor burden may lead to an accumulation of driver mutations remains a concern. Modeling patient death as occurring when the tumor reaches a critical size favors containment, and models in which the probability of death increases continuously with tumor size may lead to the conclusion that the expected survival time is lower under containment strategies than under more aggressive treatments (Mistry 2020 [87]). Considering only two types of tumors cells is restrictive, and even with only two types, if resistant

[13]For logistic growth, our assumptions are likely to be valid if the variables N, R, S are interpreted as densities, as in Strobl et al. 2020 [14], but not necessarily if they are interpreted as numbers of cells in the whole tumor, see Appendix 1.6.1.

cells are only partially resistant, the logic changes, as the growth of resistant cells may be slowed down not only indirectly, through competition with sensitive cells, but also directly, through treatment effect. The impact of a containment strategy on the development of new metastases is also unclear. On the other hand, we did not consider additional benefits of containment, such as reduced treatment toxicity, less drug-induced mutations (Kuosmanen et al. 2021 [88]) or a possible stabilization of tumor vasculature that could increase the efficiency of drug delivery (Enriquez-Navas et al. 2016 [74]).

This chapter should not be seen as providing unambiguous support for containment strategies, but as part of a wider research program aiming at clarifying the conditions under which a strategy aiming at tumor stabilization is likely to perform better than a more aggressive treatment. Data allowing to fine-tune models is still scarce, but as new competition experiments are run, and new clinical trials open (NCT05393791, ACTOv/NCT05080556), more data should become available, allowing the community to reach more definite conclusions.

1.6 Appendices

1.6.1 Mutations from sensitive to resistant cells

The analysis in this appendix is related to the work of Martin et al. (1992) [2] and Hansen et al. (2017) [9]. Consider a basic Norton-Simon model with mutations (Norton and Simon 1977 [84], Goldie and Coldman 1979 [89], Monro and Gaffney 2009 [3]):

$$\begin{aligned} \frac{dS}{dt} &= g(N)\,(1-L)S - \tau_1 g(N)S + \tau_2 g(N)R, \\ \frac{dR}{dt} &= g(N)\,R + \tau_1 g(N)S - \tau_2 g(N)R, \end{aligned} \tag{Model 5}$$

where τ_1 and τ_2 are mutation and backmutation rates. Taking $g(N) = \rho \ln(K/N)$ leads to a version of Model 2 with mutations: the original Monro and Gaffney model (Monro and Gaffney, 2009 [3]).

If the growth-rate function g is decreasing in N, an increase in the size of the sensitive population leads to two opposite effects: it slows down the development of existing resistant cells (the competition effect), but usually increases the number of mutations from sensitive to resistant cells (the mutation effect). This trade-off has been studied by Martin et al. (1992 [2]) and Hansen et al. (2017 [9]). Here, we study whether such a model is compatible with our assumption that, during treatment, a larger sensitive population leads globally to a lower growth-rate of resistant cells. To do so, let ϕ_R denote the growth-rate function of resistant cells:

$$\phi_R(S,R) = g(N)R + \tau_1 g(N)S - \tau_2 g(N)R.$$

Denoting by $x_r = R/N$ the resistant fraction, it is easily checked that $\partial\phi_R/\partial S \leq 0$ if and only if:

$$\frac{x_r}{\tau_1}(1-\tau_1-\tau_2) + 1 \geq -\frac{g(N)}{Ng'(N)}. \tag{1.3}$$

Since the resistant fraction increases during treatment, this condition is bound to be hardest to satisfy at treatment initiation.

The resistant fraction obtained from Model 5 for the initial condition $S=1$, $R=0$ is then (Goldie and Coldman, 1979 [89]):

$$x_r = \frac{\tau_1}{\tau_1+\tau_2}(1 - N_0^{-\tau_1-\tau_2}) \simeq \tau_1 \ln N_0, \tag{1.4}$$

where we used the approximation $N^{-\tau} \simeq 1 - \tau \ln N$ for τ small. Injecting (1.4) into (1.3) and using that τ_1 and τ_2 are much smaller than 1 leads to:

$$\ln N_0 + 1 \geq -\frac{g(N_0)}{N_0 g'(N_0)}. \tag{1.5}$$

Let us now consider various growth-models.

Case 1 (power-law model): $g(N) = \rho N^{-\gamma}$ with $0 < \gamma < 1$. Eq. (1.5) becomes:

$$\ln N_0 + 1 \geq 1/\gamma.$$

Typical choices for γ are $\gamma = 1/3$ or $\gamma = 1/4$ (Gerlee, 2013 [85]; Benzekry et al., 2014 [90]; our γ corresponds to $1-\gamma$ in these references). The condition then holds by a huge margin for any detectable tumor size.

Case 2 (Gompertzian growth): $g(N) = \rho \ln(K/N)$. Eq (1.5) becomes:

$$\ln N_0 + 1 \geq \ln(K/N_0),$$

which is satisfied if $K \leq eN^2$. Standard values of the carrying capacity in Gompertzian models are in the range $10^{12} - 10^{13}$ (e.g., $K = 2 \times 10^{12}$ in Monro and Gaffney (2009) [3]). Eq. (1.5) is then satisfied for any detectable tumor size.

Case 3 (logistic growth) : $g(N) = \rho(1 - K/N)$. Eq (1.5) becomes:

$$\ln N_0 + 1 \geq \frac{K}{N_0} - 1.$$

This condition need not be satisfied, depending on the interpretation of the model and parameter choices. For instance, Monro and Gaffney (2009) [3] take $N_0 = 10^{10}$. Then $\ln N_0 \simeq 23$ and the condition is roughly $K \leq 2.5 \times 10^{11}$, which is not satisfied for standard values of the carrying capacity K.[14] The condition would however be satisfied for larger initial tumor sizes, modeling late-stage treatments. Actually, when logistic growth models are used in the adaptive therapy literature, the initial tumor size is often assumed to be a large fraction of the carrying capacity (e.g., Zhang et al. 2017 [7], Strobl et al. 2020 [14], Mistry 2020 [87]). This may be interpreted as modeling late-stage treatments, or as a model of local growth. In the latter case, the carrying capacity should be seen as the maximal number of cells <u>for the current tumor volume</u> (or equivalently, the variables N, S, R, K should be interpreted as densities). Assuming 10^{10} tumor cells at tumor initiation, the estimate $x_r/\tau_1 \simeq 23$ would still be valid, and Eq. (1.3) would become $K/N \leq 25$, which is bound to be satisfied in a model of local growth.

Let us now consider three variants of Model 5.

<u>Variant 1: birth-death model.</u> In Model 5, the number of mutations is assumed proportional to the net growth-rate of the tumor. It would be natural to assume that the number of mutations is proportional to the net birth-rate. This would increase the effective mutation rate (that is, the average number of mutations relative to a given increase of tumor size).[15] However, since the condition we found is insensitive to the precise mutation rates τ_1 and τ_2, this is unlikely to affect the previous analysis.

[14] The choice of carrying capacity may differ for a Gompertz or a logistic growth model. However, in Monro and Gaffney (2009) [3], the lethal tumor size is taken to be 5×10^{11} so in a logistic growth version of their model, K would have to be at least that large and Eq. (1.5) would not be satisfied.

[15] This is the exact effect when the turnover rate $b(N)/(b(N) - d(N))$ is constant, where $b(N)$ and $d(N)$ are the birth and death rates.

Variant 2: late-stage treatment. The previous analysis is better suited for a first line treatment than a second or third line treatment, especially if resistance to the first treatment may be associated with resistance to ulterior ones. However, in such a situation (late-stage treatment), the initial resistant population is likely to be larger than the one given by (1.4), and so condition (1.3) is more likely to be satisfied.

Variant 3: cost of resistance in the baseline growth-rate. Consider the following variant of Model 5, with a different growth-rate parameter for sensitive cells and for resistant cells:

$$\begin{aligned}\frac{dS}{dt} &= \rho_s g(N)\,(1-L)S - \tau_1\rho_s g(N)S + \tau_2\rho_r g(N)R,\\ \frac{dR}{dt} &= \rho_r g(N)\,R + \tau_1\rho_r g(N)S - \tau_2\rho_r g(N)R.\end{aligned} \tag{Model 6}$$

(the terms $g(N)$ in Model 5 correspond here to terms of the form $\rho g(N)$, with $\rho = \rho_s$ or $\rho = \rho_r$.) The absolute growth-rate of resistant cells is now

$$\phi_R(S,R) = \rho_r g(N)R + \tau_1\rho_s g(N)S - \tau_2\rho_r g(N)R.$$

The condition for $\partial\phi_R/\partial S$ to be nonpositive becomes:

$$\frac{x_r}{\tau_1}(\rho_r/\rho_s - \tau_2\rho_r/\rho_s - \tau_1) + 1 \geq -\frac{g(N)}{Ng'(N)}. \tag{1.6}$$

Moreover, if ρ_r is substantially smaller than ρ_s, then the resistant fraction at treatment initiation is no longer given by (1.4) but approximately by (see Viossat and Noble, 2020 [91], Section 7):

$$x_r = \frac{\tau_1}{1-\rho_r/\rho_s}. \tag{1.7}$$

Injecting (1.7) into (1.6) and using that τ_1 and τ_2 are much smaller than 1 leads to the condition:

$$\frac{1}{1-\rho_r/\rho_s} \geq -\frac{g(N)}{Ng'(N)}. \tag{1.8}$$

For a Power-law model, the condition becomes $\rho_r/\rho_s \geq 1-\gamma$. For $\gamma = 1/3$, this is satisfied if and only if $\rho_r/\rho_s \geq 2/3$. that is, if and only if the resistance cost is not too large.

For a Gompertzian model, the right-hand side is $\ln(K/N)$ and Eq. (1.8) may be written: $\rho_r/\rho_s \geq 1 - 1/\ln(K/N)$. With Monro and Gaffney's (2009) [3] values: $N_0 = 10^{10}$, $K = 2\times 10^{12}$, this is satisfied if $\rho_r/\rho_s \geq 0.81$.

With logistic growth, the right-hand side is $K/N - 1$ and the condition may be written as $\rho_r/\rho_s \geq (K-2N)/(K-N)$.[16] Assuming for instance $\rho_r/\rho_s = 4/5$, this boils down to $K \leq 6N$. This would not be satisfied at treatment initiation if N and K represent total numbers of cells in the whole tumor (except possibly for a late-stage treatment), but seems likely to be satisfied in a model of local growth.

We conclude that in the absence of resistance costs, our analysis applies to several standard models of tumor growth with mutations, such as Power-law models or Gompertzian growth, and possibly to logistic growth, at least when it models local growth. However, if the baseline growth-rate of resistant cells is substantially smaller than the baseline growth-rate of sensitive cells, our assumptions become

[16] This condition is approximately correct only if ρ_r/ρ_s is substantially different from 1, which explains that taking the limit $\rho_r/\rho_s \to 1$ does not lead to the condition obtained in the absence of a resistance cost.

more restrictive and might fail even for Gompertzian growth. This is in line with Hansen et al.'s (2017) [9] finding that, contrary to common wisdom, a resistance cost in the baseline growth rate may make it less likely that containment strategies outperform more aggressive treatments. Note however that the fact that we can no longer prove that containment outperforms MTD does not mean that it would not do so. Moreover, the analysis of Viossat and Noble (2020) [91], Section 7 of the supplementary material, suggests that the mutation effect could only make MTD marginally superior to containment.

1.6.2 Comparison principles

The following comparison principle is standard:

Proposition 1.9. *Let Ω be a nonempty open subset of $\mathbb{R}^2$. Let $f : \Omega \to \mathbb{R}$ be continuously differentiable. Consider the ordinary differential equation $x'(t) = f(t, x(t))$. Let $t_1 \geq t_0$. Let $u : [t_0, t_1] \to \mathbb{R}$ be solution of this ODE and let $v : [t_0, t_1] \to \mathbb{R}$ be a subsolution. That is, v is continuous, almost everywhere differentiable, $(t, v(t)) \in \Omega$ on $[t_0, t_1]$, and, almost everywhere, $v'(t) \leq f(t, v(t))$. If furthermore $v(t_0) \leq u(t_0)$, then $v(t) \leq u(t)$ for all $t \in [t_0, t_1]$.*

We want to apply a variant of this result to two equations: first,

$$\frac{d\tilde{N}}{dr} = G(r, \tilde{N}(r)) \text{ where } G(r, \tilde{N}) = \frac{f_N(\tilde{N}, r, \tilde{L}_2(r))}{f_R(\tilde{N}, r)},$$

(here $\tilde{N}(r)$ plays the role of $u(t)$, and G the role of f in Proposition 1.9). Second,

$$\frac{dR}{dt} = H(R(t)) \text{ where } H(R) = f_R(\tilde{N}_2(R), R),$$

where $\tilde{L}_2$ and $\tilde{N}_2$ are not continuously differentiable (otherwise Proposition 1.9 would directly apply), but piecewise C^1 in the strong sense we defined in Section 1.2. For instance, for $\tilde{L}_2$, which is defined on $[R_0, R_2^\infty)$, there exist values $r_0 = R_0 < r_1 < ... < r_n = R_2^\infty$ such that for each i in $\{1, .., n-1\}$, $\tilde{L}_2$ coincides on $[r_i, r_{i+1})$ with a continuously differentiable function $\tilde{L}_2^i$ defined on a neighborhood of $[r_i, r_{i+1})$.

We thus need variants of Proposition 1.9 where f is slightly less regular. The proof of these variants consists in repeated applications of Proposition 1.9.

Proposition 1.10. *The conclusion $u(t) \leq v(t)$ on $[t_0, t_1]$ of Proposition 1.9 still holds in the following cases:*

a) if f takes the form $f(t, x) = \psi(t, x, \phi(t))$, where ψ is continuously differentiable and ϕ is strongly piecewise C^1.

b) if f is a function of x only ($f : I \to \mathbb{R}$, where I is a nonempty interval) which is strongly piecewise C^1, and u and v are strictly increasing.

Proof. Proof of a): by assumption, there exists an integer n and real numbers $(\tau_k)_{0 \leq k \leq n}$ satisfying $t_0 = \tau_0 < \tau_1 < ... < \tau_n = t_1$ such that on $A_k = [t_k, t_{k+1}) \times \mathbb{R}$, f coincides with a continuously differentiable function f_k defined on a neighborhood of A_k. Assuming $u(\tau_k) \leq v(\tau_k)$, Proposition 1.9 implies that $u(t) \leq v(t)$ on $[t_k, t_{k+1})$ and this is also true at time t_{k+1} by continuity of u and v. An induction argument then gives the result.

Proof of b): By assumption, there exist a finite number of values x_0,...,x_n such that f coincides on $[x_k, x_{k+1})$ with a continuously differentiable function ϕ_k defined on a neighborhood of $[x_k, x_{k+1})$, and we may assume (up to adding an artificial point) that $x_0 \leq u(t_0)$. Since u and v are strictly increasing, there exists an integer $q \leq 2(n+1)$ and a sequence of times $\tau_0 = t_0 < \tau_1 < ... < \tau_q = t_1$ such that on (τ_j, τ_{j+1}), $u(t)$ and $v(t)$ are never equal to one of the x_k. Assuming $u(\tau_j) \leq v(\tau_j)$, then either: for all t in $[\tau_j, \tau_{j+1})$, $u(t)$ and $v(t)$ belong to the same interval $[x_k, x_{k+1})$ and Proposition 1.9 implies that $u(t) \leq v(t)$ on $[\tau_j, \tau_{j+1})$; or there exists k such that for all t in $[\tau_j, \tau_{j+1})$, $u(t) \leq x_k \leq v(t)$ and the same inequality holds trivially. In both cases, $u(\tau_{k+1}) \leq v(\tau_{k+1})$ by continuity of u and v. An induction argument then gives the result. □

Chapter 2

Asymptotic expansion of a limit cycle arising from a tumour-immune system interaction model

2.1 Introduction

The appearance of Hopf bifurcations is a relevant subject when studying tumor-immune system interaction models, as it can provide insights into the complex dynamics that are involved. A Hopf bifurcation indicates the emergence of oscillations or periodic behavior in the model. In the context of tumor-immune system interactions, this could represent cyclical fluctuations in tumor size and immune response strength. A deep understanding of this phenomena can be crucial for designing effective immunotherapy strategies. The timing and periodicity of immune system responses could impact treatment effectiveness (see [92, 93]). More specifically, the oscillatory behavior could lead to resonance phenomena. This means that interventions or treatments applied at specific frequencies may have enhanced effects on tumor control or immune system modulation. Another potential advantage is the identification of critical parameters within the model that drive the oscillatory behavior (tumour growth and death rates, influx of effector cells, among others).
It could be argued that the previously discussed reasons are motivation enough in order to further explore the results presented in [16]. On the first section of said reference, were given the conditions for the appearance of a periodic orbit on a system proposed in [94]. Said system describes, with great generality, the interactions between tumour cells and immune system cells. The work presented in [94] has sparked wide interest amid the researchers of the area. It has been referenced in several reviews of mathematical models such as [95, 96], where some positive points and shortcomings have been highlighted.
To perform their analysis, the authors from [16] first derive a second order approximation of the original system and then, by means of the Poincaré-Bendixon theorem, they conclude the existence of a limit cycle. Other references which perform qualitative studies of models directly based on the one proposed in [94] include [97, 98], but up to our knowledge, an expression for the limit cycle has not been derived yet. In this chapter, we provide an asymptotic expansion for the limit cycle by following the method described in [99]. We propose as well a way to extend our results for a model where spatial homogeneity is not taken into consideration.

2.2 The model

We briefly describe the model presented in [94] and [16]. The quantities $c(t)$ and $n(t)$ will represent the amount of tumour cell and lymphocytes at time t respectively. They evolve according to the system

$$\begin{cases} \dfrac{d}{dt}c(t) = a_1 c(t)F(c(t)) - a_2\mu\phi(c(t))c(t)n(t), \\ \dfrac{d}{dt}n(t) = -a_3 n(t)\psi(c(t)) + a_4 q(c(t)), \end{cases} \tag{2.1}$$

where the function $\psi(c)$ describes the stimulatory effect of the tumour cells on the immune cells. It can be assumed that this function is positive (at least initially), $\psi(0) > 0$, and might be negative only in a finite interval. It is reasonable to assume $|\psi'(0)| \leqslant 1$, so that, at least initially, the death rate of lymphocytes is not greater than that in the linear model. The tumour growth rate $F(c)$ is a positive function which summarizes the carrying capacity (or malignancy) such that $F(0) > 0$, $F'(c) \leqslant 0$ and $\lim_{c\to\infty} cF(c) = 0$, with the additional assumption that initially it is $F'(0) = 0$. The loss of tumour cells, which depends on the competition with lymphocytes, is represented by the function $\phi(c)$ characterized by $\phi(c) > 0$, $\phi'(c) \leqslant 0$ and $\lim_{c\to\infty} c\phi(c) = l < \infty$. In other words, if the tumour growth tends to infinity the loss of tumour cells would tend to a constant rate. It can be also assumed that $\phi'(0) = 0$. Regarding the influx of immune cells $q(c)$ can be taken $q(0) = 1$, $|q'(0)| \leqslant 1$, so that, at least initially, the influx of effector cells is not greater than that in the linear model.
By assuming

$$u = c, \qquad v = \frac{n}{a_4}, \qquad \tau = a_3 t,$$

and introducing

$$a = \frac{a_1}{a_3}, \qquad b = \frac{1}{a_3}, \qquad \mu = \frac{a_2 a_4}{a_3},$$

we get the non-dimensional model

$$\begin{cases} \dfrac{d}{dt}u(t) = au(t)F(u(t)) - \mu\phi(u(t))u(t)v(t), \\ \dfrac{d}{dt}v(t) = -v(t)\psi(u(t)) + bq(u(t)). \end{cases}$$

The reaction term can be approximated using a second order Taylor expansion around the steady state $(0, b/\psi(0))$, as done in [16], where, after assuming $F'(0) = 0$, $\phi(0) = 1$ and $q''(0) = 0$, and grouping similar terms, we obtain the system

$$\begin{cases} \dfrac{d}{dt}u(t) = \alpha u(t) - \mu u(t)v(t), \\ \dfrac{d}{dt}v(t) = -\beta_1 u(t)v(t) - \beta_2 v(t) + \beta_3 u(t) + \beta_4 - \beta_5 v(t)u^2(t), \end{cases} \tag{2.2}$$

where $\alpha = aF(0)$, $\beta_1 = \psi'(0)$, $\beta_2 = \psi(0)$, $\beta_3 = bq'(0)$, $\beta_4 = b$ and $\beta_5 = \frac{1}{2}\psi''(0)$. Hence, the following restrictions apply to the parameters:

$$\alpha > 0, \ \mu > 0, \ |\beta_1| \leqslant 1, \ \beta_2 > 0 \text{ and } |\beta_3| \leqslant \beta_4. \tag{2.3}$$

The system (2.2) presents 3 spatially homogeneous steady states. These were computed in [16] and they are

$$\begin{aligned} P_0 &= (0, \frac{\beta_4}{\beta_2}), \\ P_1 &= (p_1, q_1) = (\frac{1}{2\alpha\beta_5}\left[b_0 - \sqrt{D}\right], \frac{\alpha}{\mu}), \\ P_2 &= (p_2, q_2) = (\frac{1}{2\alpha\beta_5}\left[b_0 + \sqrt{D}\right], \frac{\alpha}{\mu}). \end{aligned}$$

where

$$\begin{aligned} b_0 &= -\alpha\beta_1 + \beta_3\mu, \\ D &= (\alpha\beta_1 - \beta_3\mu)^2 - 4\alpha\beta_5(\alpha\beta_2 - \beta_4\mu). \end{aligned}$$

It is withing our interest to study the existence of oscillating solutions around any of these steady states. However, an oscillating solution around P_0 would inevitably attain negative values, which lacks biological sense, and therefore we restrict our analysis to P_1 and P_2.
We start by imposing conditions which guarantee that such steady states are well defined and positive. The steady states P_1 and P_2 are well defined if and only if

$$\beta_5 \neq 0 \text{ and } D \geqslant 0. \tag{2.4}$$

Furthermore, the components of P_1 are positive if and only if either

$$b_0 \leqslant 0 \text{ and } \beta_5 < 0,$$

or

$$b_0 > 0 \text{ and } \alpha\beta_2 - \beta_4\mu > 0.$$

Similarly, the components of P_2 are non-negative if and only if either

$$b_0 \geqslant 0 \text{ and } \beta_5 > 0,$$

or

$$b_0 < 0 \text{ and } \alpha\beta_2 - \beta_4\mu < 0.$$

Knowing this, let us make a change of variables in such a way that $(0,0)$ becomes the steady state, this is

$$U = u - p_i, \quad V = v - \frac{\alpha}{\mu}, \tag{2.5}$$

with $i = 1, 2$. This way, system (2.2) can be written as

$$\begin{aligned} \partial_t U &= -\mu p_i U - \mu UV \\ \partial_t V &= k_1 U + k_2 V + k_3 UV + k_4 U^2 + k_5 V U^2 \end{aligned} \tag{2.6}$$

where

$$\begin{aligned} k_1 &= -\frac{\alpha\beta_1}{\mu} + \beta_3 - \frac{2\alpha p_i \beta_5}{\mu} = \frac{\pm\sqrt{D}}{\mu}, \\ k_2 &= -\beta_1 p_i - \beta_2 - \beta_5 p_i^2, \\ k_3 &= -\beta_1 - 2\beta_5 p_i, \\ k_4 &= -\beta_5 \frac{\alpha}{\mu}, \\ k_5 &= -\beta_5. \end{aligned}$$

2.3 Derivation of the expression for the limit cycle

In this section we briefly explain the algorithm for obtaining the limit cycle exposed in [99] and then proceed to apply it to system (2.6).

2.3.1 Algorithm for obtaining the limit cycle

Adopting the same notations as in [99], a non-linear system near the steady state such as (2.6) takes the form

$$\dot{X} = \mathscr{F}(X) = J_a X + \Psi(X), \tag{2.7}$$

where X is the two dimensional unknown, J_a is the parameter dependent Jacobian matrix at the origin and $\Psi(X)$ is the non linear term.
First, it is considered an invertible analytical transform of coordinates of a neighborhood of the origin onto another

$$Y = \mathscr{H}(X) = \Gamma X + \mathscr{G}(X), \tag{2.8}$$

driven by a non-singular matrix Γ, and the analytic vector function:

$$\mathscr{G}(X) = \begin{pmatrix} \overline{\varphi}(U,V) \\ \overline{\psi}(U,V) \end{pmatrix} = \begin{pmatrix} \sum\limits_{n=2}^{\infty} \sum\limits_{i+j=n} \overline{\varphi}_{ij} U^i V^j \\ \sum\limits_{n=2}^{\infty} \sum\limits_{i+j=n} \overline{\psi}_{ij} U^i V^j \end{pmatrix} \tag{2.9}$$

which is assumed to have a positive radius of convergence.
We shall denote the inverse of $\mathscr{H}(X)$ as

$$X = \mathscr{H}^{-1}(Y) = \Gamma^{-1} Y + \mathscr{K}(Y), \tag{2.10}$$

where

$$\mathscr{K}(Y) = \begin{pmatrix} \underline{\varphi}(y_1,y_2) \\ \underline{\psi}(y_1,y_2) \end{pmatrix} = \begin{pmatrix} \sum\limits_{n=2}^{\infty} \sum\limits_{i+j=n} \underline{\varphi}_{ij} y_1^i y_2^j \\ \sum\limits_{n=2}^{\infty} \sum\limits_{i+j=n} \underline{\psi}_{ij} y_1^i y_2^j \end{pmatrix}. \tag{2.11}$$

If $\mathscr{H}$ over the orbits of system (2.7) is such that

$$Y = \begin{pmatrix} z \\ \dot{z} \end{pmatrix} \tag{2.12}$$

being $z(t)$ an unknown function, the integration of the system can be reduced to the integration of a second order differential equation in the variable z.
The change of variables $\mathscr{H}$ verifies (2.12) if and only if

$$\frac{d}{dt}(\Pi_1 \mathscr{H}(X)) = \Pi_2 \mathscr{H}(X). \tag{2.13}$$

Equation (2.13) implies that the components γ_{ij} of Γ verify the following concordance condition with the Jacobian of the system at the steady state

$$J_a^T \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \end{pmatrix} = \begin{pmatrix} \gamma_{21} \\ \gamma_{22} \end{pmatrix}. \tag{2.14}$$

Developing the left hand side term in (2.13) we get

$$\begin{aligned}\frac{d}{dt}(\Pi_1\mathscr{H}(X)) &= \Pi_1\dot{\mathscr{H}(X)}\\ &= \Pi_1(\Gamma\dot{X} + \nabla\mathscr{G}(X)\cdot\dot{X})\\ &= \Pi_1(\Gamma(J_aX + \Psi(X))) + \nabla\overline{\varphi}(X)\cdot\dot{X}\\ &= \Pi_1(\Gamma J_aX) + \Pi_1(\Gamma\Psi(X)) + \nabla\overline{\varphi}\cdot\dot{X}.\end{aligned}$$

On the other hand, the right term on (2.13) is equal to

$$\Pi_2\mathscr{H}(X) = \Pi_2\Gamma X + \overline{\psi}(X). \tag{2.15}$$

Thanks to (2.14), we know that $\Pi_1(\Gamma J_aX) = \Pi_2\Gamma X$, therefore we get the relation

$$\overline{\psi}(X) = \Pi_1(\Gamma\Psi(X)) + \nabla\overline{\varphi}\cdot\dot{X} = \Pi_1(\Gamma\Psi(X)) + \nabla\overline{\varphi}\cdot\mathscr{F}(X). \tag{2.16}$$

Hence, determining the suitable change of variables $\mathscr{H}$ is reduced to appropriately fix the values of Γ and $\overline{\varphi}$. Deriving (2.15) with respect to t and using the expression for $\mathscr{H}^{-1}(Y)$, we obtain the aforementioned second order differential equation

$$\begin{aligned}\ddot{z}(t) =&\frac{d}{dt}\left(\Pi_2\mathscr{H}(X)\right)\\ =&\Pi_2\Gamma\dot{X} + \nabla\overline{\psi}(X)\cdot\mathscr{F}(X)\\ =&\Pi_2\Gamma(J_aX + \Psi(X)) + \nabla\overline{\psi}(X)\cdot\mathscr{F}(X)\\ =&\Pi_2\Gamma J_a\Gamma^{-1}Y + \Pi_2\Gamma\mathscr{K}(Y) + \Pi_2\Gamma J_a\Psi(\mathscr{H}^{-1}(Y)) + \nabla\overline{\psi}(\mathscr{H}^{-1}(Y))\cdot\mathscr{F}(\mathscr{H}^{-1}(Y)),\end{aligned}$$

or equivalently

$$\ddot{z}(t) - \Pi_2\Gamma J_a\Gamma^{-1}Y = G(Y), \tag{2.17}$$

where

$$G(Y) = \Pi_2\Gamma\mathscr{K}(Y) + \Pi_2\Gamma\Psi(\mathscr{H}^{-1}(Y)) + \nabla\overline{\psi}(\mathscr{H}^{-1}(Y))\cdot\mathscr{F}(\mathscr{H}^{-1}(Y)).$$

Once the system (2.7) is reduced to the second order differential equation (2.17), an approximation for the limit cycle can be obtained by applying the Krylov–Bogoliubov–Mitropolski averaging method (see [100]).

2.3.2 Expression for the limit cycle

Let us apply the previously described method to system (2.6). The Jacobian matrix at the steady state has coefficients

$$J = \begin{pmatrix} 0 & -\mu p_i \\ k_1 & k_2 \end{pmatrix}. \tag{2.18}$$

for $i = 1, 2$. Note that, for $i = 2$, $\det(J) = -\sqrt{D}p_2$, which is negative under the assumption that $p_2 > 0$, hence, no Hopf bifurcation occurs unless $p_2 < 0$. Again, due to the lack of biological sense for negative values of the solution, we discard the steady state P_2 and only perform our analysis over P_1. For all vectors $X = (x, y)^T$ we define the function

$$\mathscr{G}(X) = x^2\begin{pmatrix} g_1 \\ 1 \end{pmatrix}$$

for some g_1 to be fixed later and we consider the change of variables

$$Y = \mathscr{H}(X) = \Gamma X + \mathscr{G}(X), \tag{2.19}$$

where $\Gamma = a\Gamma_1 + b\Gamma_2$ is a linear combination of

$$\Gamma_1 = \begin{pmatrix} 1 & 0 \\ j_{11} & j_{12} \end{pmatrix} \text{ and } \Gamma_2 = \begin{pmatrix} 0 & 1 \\ j_{21} & j_{22} \end{pmatrix},$$

which will be determined later. This choice of Γ ensures that the concordance condition (2.14) is satisfied. We will fix the values of a, b and g_1 in such a way that, after denoting $X(t) = (U(t), V(t))^T$, we get $\mathscr{H}(X(t)) = (z(t), \dot{z}(t))^T$ for some function $z(t)$. It is simple to check that the appropriate values are those obtained after solving the linear system

$$\begin{aligned} -a\mu + bk_3 + 2g_1 j_{12} &= 0, \\ bk_5 - 2g_1\mu &= 0 \\ bk_4 + 2g_1 j_{11} &= 1. \end{aligned}$$

For the steady state P_1, we have $j_{11} = 0$, which implies that

$$\begin{aligned} a &= \frac{j_{12}k_5 + \mu k_3}{\mu^2 k_5} = \frac{\beta_1 + p_2\beta_5}{\alpha\beta_5}, \\ b &= 1/k_4 = -\frac{\mu}{\alpha\beta_5}, \\ g_1 &= -\frac{k_5}{2(-\mu))k_4} = \frac{1}{2\alpha}, \end{aligned}$$

and consequently

$$\Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} := \begin{pmatrix} \frac{\beta_1 + p_1\beta_5}{\alpha\beta_5} & -\frac{\mu}{\alpha\beta_5} \\ \frac{-\sqrt{D}}{\alpha\beta_5} & \frac{\mu\beta_2}{\alpha\beta_5} \end{pmatrix}.$$

The change of variables (2.19) accepts an explicit inverse which we proceed to derive. Given two pairs of vectors $X = (x_1, x_2)^T$ and $Y = (y_1, y_2)^T$ satisfying $Y = \mathscr{H}(X)$, we see that

$$\Gamma^{-1}Y = \Gamma^{-1}\mathscr{H}(X) = X + x_1^2\Gamma^{-1}\begin{pmatrix} g_1 \\ 1 \end{pmatrix}. \tag{2.20}$$

We define the quantities

$$L_1 = L_1(y_1, y_2) := |\Gamma|^{-1}(\Gamma_{22}y_1 - \Gamma_{12}y_2), \tag{2.21}$$
$$L_2 = L_2(y_1, y_2) := -|\Gamma|^{-1}(\Gamma_{21}y_1 - \Gamma_{11}y_2), \tag{2.22}$$

satisfying

$$\begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \Gamma^{-1}Y,$$

so the system of equations (2.20) can be written as

$$\begin{aligned} L_1 &= x_1 + |\Gamma|^{-1}(\Gamma_{22}g_1 - \Gamma_{12})x_1^2, \\ L_2 &= x_2 - |\Gamma|^{-1}(\Gamma_{21}g_1 - \Gamma_{11})x_1^2. \end{aligned} \tag{2.23}$$

The value of $\Gamma_{22}g_1 - \Gamma_{12} = \frac{\mu}{\alpha\beta_5}(1+\frac{\beta_2}{2\alpha})$ is always different from 0, and after solving the algebraic system (2.23) we obtain

$$X = \mathscr{H}^{-1}(Y) = \begin{pmatrix} L_1 + \chi(y_1,y_2) \\ L_2 - g^*\chi(y_1,y_2) \end{pmatrix} = \Gamma^{-1}Y + \chi\begin{pmatrix} 1 \\ -g^* \end{pmatrix}. \tag{2.24}$$

where

$$g^* = \frac{\Gamma_{11} - g_1\Gamma_{21}}{\Gamma_{12} - g_1\Gamma_{22}}, \tag{2.25}$$

$$\chi(y_1,y_2) = \frac{-1 - 2|\Gamma|^{-1}(g_1\Gamma_{22} - \Gamma_{12})L_1 + \sqrt{1 + 4|\Gamma|^{-1}(g_1\Gamma_{22} - \Gamma_{12})L_1}}{2|\Gamma|^{-1}(g_1\Gamma_{22} - \Gamma_{12})}. \tag{2.26}$$

With $\mathscr{H}(X)$ and its inverse being well defined, we have all of the necessary elements to state the main result of the current chapter.

Theorem 2.1. *Let τ_J and δ_J be the trace and the determinant of J respectively and define the quantities*

$$a_n = \begin{cases} \frac{(\Gamma_{12}-g_1\Gamma_{22})}{|\Gamma|}, & \text{if } n = 2, \\ \frac{(4n-6)}{n}\frac{(\Gamma_{12}-g_1\Gamma_{22})}{|\Gamma|}a_{n-1}, & \text{if } n > 2, \end{cases}$$

$$\begin{aligned}
\tilde{G}_1 &= |\Gamma|^{-3}\sqrt{\delta_J}(-\delta_J(\Gamma_{11} - g^*\Gamma_{12}) + \tau_J(\Gamma_{21} - g^*\Gamma_{22})),\\
\tilde{G}_2 &= |\Gamma|^{-3}\sqrt{\delta_J}(\Gamma_{22}k_4 + 2j_{11}),\\
\tilde{G}_3 &= |\Gamma|^{-3}\sqrt{\delta_J}(-\Gamma_{21}\mu + \Gamma_{22}k_3 + 2j_{12}),\\
\tilde{G}_4 &= |\Gamma|^{-3}\sqrt{\delta_J}(\Gamma_{22}k_5 - 2\mu)
\end{aligned}$$

and

$$\begin{aligned}
\pi_0 = &-\frac{3}{4}\left(a_3\tilde{G}_1 + 2a_2\tilde{G}_2 - a_2g^*\tilde{G}_3\right)\left(\Gamma_{22}^2\Gamma_{12} + \Gamma_{12}^3\delta_J\right)\\
&+\frac{1}{4}\left(a_2\tilde{G}_3 + \tilde{G}_4\right)\left(\Gamma_{22}^2\Gamma_{11} + 2\Gamma_{22}\Gamma_{12}\Gamma_{21} + 3\Gamma_{12}^2\Gamma_{11}\delta_J\right),\\
q_0 = &-\frac{3}{8}\left(a_3\tilde{G}_1 + 2a_2\tilde{G}_2 - a_2g^*\tilde{G}_3\right)\left(\Gamma_{22}^3 + \Gamma_{12}^2\Gamma_{22}\delta_J\right)\\
&+\frac{1}{8}\left(a_2\tilde{G}_3 + \tilde{G}_4\right)\left(3\Gamma_{22}^2\Gamma_{21} + (\Gamma_{12}^2\Gamma_{21} - 2\Gamma_{12}\Gamma_{11}\Gamma_{22})\delta_J\right).
\end{aligned}$$

System (2.6) admits a limit cycle if and only if $\tau_J\pi_0 > 0$ which admits the following asymptotic expansion with respect to τ_J

$$\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \Gamma^{-1}\begin{pmatrix} z(t) \\ \dot{z}(t) \end{pmatrix} + \chi(z(t),\dot{z}(t))\begin{pmatrix} 1 \\ g^* \end{pmatrix} + \mathscr{O}(\frac{|\tau_J|}{\sqrt{\delta_J}}), \tag{2.27}$$

where

$$z(t) = \sqrt{\frac{|\tau_J|}{|\pi_0|\sqrt{\delta_J}}}\sin((\sqrt{\delta_J} + \frac{q_0|\tau_J|}{|\pi_0|})t).$$

Proof. Noticing that

$$\Pi_2 \Gamma J \Gamma^{-1} Y = \tau_J \dot{z} - \delta_J z,$$

then, according to (2.17), the function $z(t) = \Pi_1 Y$ is the solution of the second order equation

$$\ddot{z} - \tau_J \dot{z} + \delta_J z = G(z, \dot{z}),$$

where $G(z, \dot{z})$ is defined as

$$G(z, \dot{z}) = \Pi_2 \{ \Gamma(J \mathscr{K}(Y) + \Psi(\mathscr{H}^{-1} Y)) + \langle \mathrm{grad}_X \mathscr{G}, \mathscr{F}) \rangle (\mathscr{H}^{-1} Y \},$$

Computing G we get

$$\begin{aligned} G(z, \dot{z}) =& (-\delta_J(\Gamma_{11} - g^* \Gamma_{12}) + \tau_J(\Gamma_{21} - g^* \Gamma_{22}))\chi \\ &+ (\Gamma_{22} k_4 + 2 j_{11})(L_1 + \chi)^2 \\ &+ (-\Gamma_{21} \mu + \Gamma_{22} k_3 + 2 j_{12})(L_1 + \chi)(L_2 - g^* \chi) \\ &+ (\Gamma_{22} k_5 - 2\mu)(L_1 + \chi)^2 (L_2 - g^* \chi). \end{aligned}$$

where we have omitted the dependence on z and $\dot{z}$ for the functions L_1, L_2 and χ.
We are looking for oscillations with small amplitude ε, so after making the change of variables

$$z(t) = \varepsilon \varsigma(\sqrt{\delta_J} t),$$

we get that ς satisfies the equation

$$\ddot{\varsigma} - \frac{\tau_J}{\sqrt{\delta_J}} \dot{\varsigma} + \varsigma = \varepsilon G_1(\varsigma, \dot{\varsigma}; \varepsilon).$$

where

$$G_1(\varsigma, \dot{\varsigma}; \varepsilon) = \delta_J^{-1} \varepsilon^{-2} G(\varepsilon \varsigma, \sqrt{\delta_J} \varepsilon \dot{\varsigma}).$$

Let us introduce the new variables $r(t)$ and $\theta(t)$ defined as

$$\begin{aligned} \varsigma &= r \sin(t + \theta), \\ \dot{\varsigma} &= r \cos(t + \theta). \end{aligned}$$

Aand after applying the Krylov–Bogoliubov–Mitropolski averaging method, we get the equations for r and θ

$$\dot{r} = \frac{r}{2} \Big(\frac{\tau_J}{\sqrt{\delta_J}} + p(r; \varepsilon) \Big), \tag{2.28}$$

$$\dot{\theta} = q(r; \varepsilon), \tag{2.29}$$

where

$$p(r; \varepsilon) = \frac{\varepsilon}{\pi r} \int_0^{2\pi} \cos\phi G_1(r \sin\phi, r\cos\phi; \varepsilon) d\phi,$$

$$q(r; \varepsilon) = -\frac{\varepsilon}{2\pi r} \int_0^{2\pi} \sin\phi G_1(r \sin\phi, r\cos\phi; \varepsilon) d\phi.$$

It is likely not possible to give an explicit expression for $p(r;\epsilon)$ and $q(r;\varepsilon)$, however, using the Taylor expansion of $\chi(z,\dot{z})$ given by

$$\chi(z,\dot{z}) = \sum_{n=2}^{\infty} a_n L_1^n(z,\dot{z}),$$

with

$$a_n = \begin{cases} \frac{(\Gamma_{12}-g_1\Gamma_{22})}{|\Gamma|}, & \text{if } n=2, \\ \frac{(4n-6)}{n}\frac{(\Gamma_{12}-g_1\Gamma_{22})}{|\Gamma|}a_{n-1}, & \text{if } n>2, \end{cases}$$

we can find the Taylor expansion in powers of r for both $p(r;\varepsilon)$ and $q(r;\varepsilon)$

$$p(r;\varepsilon) = \pi_0\varepsilon^2 r^2 + \mathcal{O}(\varepsilon^4 r^4),$$
$$q(r;\varepsilon) = q_0\varepsilon^2 r^2 + \mathcal{O}(\varepsilon^4 r^4),$$

where

$$\begin{aligned} \pi_0 = &-\frac{3}{4}\Big(a_3\tilde{G}_1 + 2a_2\tilde{G}_2 - a_2 g^*\tilde{G}_3\Big)\Big(\Gamma_{22}^2\Gamma_{12} + \Gamma_{12}^3\delta_J\Big) \\ &+\frac{1}{4}\Big(a_2\tilde{G}_3 + \tilde{G}_4\Big)\Big(\Gamma_{22}^2\Gamma_{11} + 2\Gamma_{22}\Gamma_{12}\Gamma_{21} + 3\Gamma_{12}^2\Gamma_{11}\delta_J\Big), \\ q_0 = &-\frac{3}{8}\Big(a_3\tilde{G}_1 + 2a_2\tilde{G}_2 - a_2 g^*\tilde{G}_3\Big)\Big(\Gamma_{22}^3 + \Gamma_{12}^2\Gamma_{22}\delta_J\Big) \\ &+\frac{1}{8}\Big(a_2\tilde{G}_3 + \tilde{G}_4\Big)\Big(3\Gamma_{22}^2\Gamma_{21} + (\Gamma_{12}^2\Gamma_{21} - 2\Gamma_{12}\Gamma_{11}\Gamma_{22})\delta_J\Big). \end{aligned}$$

and

$$\begin{aligned} \tilde{G}_1 &= |\Gamma|^{-3}\sqrt{\delta_J}(-\delta_J(\Gamma_{11} - g^*\Gamma_{12}) + \tau_J(\Gamma_{21} - g^*\Gamma_{22})), \\ \tilde{G}_2 &= |\Gamma|^{-3}\sqrt{\delta_J}(\Gamma_{22}k_4 + 2j_{11}), \\ \tilde{G}_3 &= |\Gamma|^{-3}\sqrt{\delta_J}(-\Gamma_{21}\mu + \Gamma_{22}k_3 + 2j_{12}), \\ \tilde{G}_4 &= |\Gamma|^{-3}\sqrt{\delta_J}(\Gamma_{22}k_5 - 2\mu). \end{aligned}$$

There will be a root for the right hand side of (2.28) if and only if

$$\tau_J\pi_0 > 0.$$

Assuming that this condition holds and choosing $\varepsilon^2 = \frac{\tau_J}{\sqrt{\delta_J}}$, the root for the right hand side of (2.28) is

$$r_0 = \frac{1}{\sqrt{|\pi_0|}}.$$

Substituting in the expression for $\varsigma(t)$ and then for $z(t)$, we get

$$z(t) = \sqrt{\frac{|\tau_J|}{|\pi_0|\sqrt{\delta_J}}}\sin((\sqrt{\delta_J} + \frac{q_0|\tau_J|}{|\pi_0|})t),$$

and finally using (2.19) we obtain the expression for an approximation of the limit cycle for system 2.6

$$\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \Gamma^{-1}\begin{pmatrix} z(t) \\ \dot{z}(t) \end{pmatrix} + \chi(z(t),\dot{z}(t))\begin{pmatrix} 1 \\ g^* \end{pmatrix} + \mathcal{O}(\frac{|\tau_J|}{\sqrt{\delta_J}}). \tag{2.30}$$

□

2.4 Discussion and perspectives

Knowing the expression for the limit cycle paves the way for two problems associated with a version of system (2.1) which is non homogeneous in space. Said problem is modeled through the system of partial differential equations

$$\begin{cases} \partial_t c - d_1 \Delta(c^{m_1}) = a_1 c F(c) - a_2 \mu \phi(c) c n, \\ \partial_t n - d_2 \Delta(n) = -a_3 n \psi(c) + a_4 q(c), \end{cases} \tag{2.31}$$

where the quantities $c(t, x)$ and $n(t, x)$ will represent the densities of tumour cell and lymphocytes at time t and at point x. It was assumed that cancer cell proliferation increases the local tissue pressure thus creating a velocity field, which justifies the diffusive term on the first equation while the lymphocytes will diffuse following its own concentration gradient.

The first question of interest related to system (2.31) is how do the strong instabilities propagate in the presence of of an spatially homogeneous limit cycle. The second problem to be tackled is the identification of conditions over the parameters that may lead to the appearance of twinkling patterns associated to Turing-Hopf instabilities. Similar questions for other various models have been approached in [101–105] and the references within.

Spatial heterogeneity and pattern formations are two phenomena which are present on the development of tumours, as observed in [106–108]. This is why, understanding the link between the parameters directing the evolution of cancer, and said heterogeneity, is vital when aiming for a stronger comprehension of this disease.

Chapter 3

Evolution of a structured cell population endowed with plasticity of traits under constraints on and between the traits

3.1 Introduction: biological background

One of the main theories explaining the origins of cancer states that the hallmark capabilities of cancer are based on latent functions already existing in the genome of normal human cells, and that cancer represents a reversion to a less differentiated and less cooperative cellular behavior. This theory is usually called the atavistic model of cancer [62]. It is opposed to the more commonly admitted somatic mutation theory [109], that states that cancer originated in a single cell, following a catastrophic sequence of stochastic tumorigenic mutations, and from which Darwinian selection produced by divisions from this unicellular basis a more or less organised society of cheating cells, more and more escaping controls by the host organism, i.e., a cancer. In the atavistic theory, accompanied or induced by blockade of differentiation or reverse differentiation of normally maturing cells, societies of cells in a multicellular organism (cancer is always a disease of multicellular organisms) somehow, in some location of the organism, escape the fine control under which they are normally placed and revert to a previous, coarse and disorganised state of multicellularity [62]. This may be understood as a process of "deDarwinisation", through which cancer cells gain a state of *plasticity* [63–66] representative of a former state in the evolution of multicellularity.

The atavistic theory has thus deep connections with the theory of evolution to multicellularity. The passage from unicellular organisms to multicellular ones led to the regulation of capabilities, resulting in controlled proliferation and differentiation of cells leading to specialisation and cooperation between specialised cells. The role of environment-driven cellular stress in this process of specialisation has recently been stressed by various authors [67, 68]. The new genes responsible for these regulations became tumour suppressors. The atavistic theory states that if these new suppressors become damaged for some reason, then latent genes, associated with functions from unicellular organisms, will reappear and dominate the scenery, thus resulting in the unconstrained proliferation and the lack of cooperation with the other cells of the host organism, as actually found in tumours.

Understanding how evolution led to the emergence of multicellularity then becomes a problem closely related to that of understanding cancer. Firstly because unravelling in detail what is lost (cohesion) and what is gained (plasticity) in this reverse evolutionary process may help understand the nature of cancer from a functional point of view. Secondly because it is reasonable to assume that Darwinian selection in tumours starts from a primitive state of multicellularity in which cells are very plastic with respect to their phenotypes, which sends us back to states in evolution close to the emergence of multicellularity. The division of work through cell differentiation achieved by coherent multicellularity (i.e., designing a stable, cohesive multicellular organism) [110] was of vital importance in order to evolve into more complex and more functionally efficient organisms. In a first stage, this differentiation was very likely reversible, due to the high plasticity that cells were endowed with in a primitive state of multicellularity. Under these conditions, one can reasonably assume that these primitive organisms adopted bet hedging strategies, i.e., common risk-diversifying strategies in unpredictably changing and often aggressive environments, in order to maximise their phenotypic fitness [69, 70].

In more detail, bet hedging in cell populations may be defined as a diversification of phenotypes in a cohesive (or at least bound by a common fate community, usually of genetic nature) cell population to optimise its fitness, in particular *a minima* to ensure its survival in a life-threatening environment, in other words to design a fail-safe strategy for the preservation of a propagating element; one cell may be enough to achieve this goal. Enhancing the ability to (quickly) diversify phenotypes by changing differentiation paths (by reversal, i.e., vertical de-differentiation, or by horizontal trandsdifferentiation in a differentiation tree) may in particular be achieved at the chromatin level by means of epigenetic enzymes, molecular instances of cell plasticity at work [64].

Among such commonly described strategies of living organisms (unicellular or multicellular) meant to ensure survival in changing environments have been classically described fright, fight and flight. Fright (or freeze) is not likely to induce phenotype evolution. Fight (establishing barriers, secreting poisons, gathering in colonies) and flight (motility to escape unbeatable predators) can. Differentiation between somatic and germinal cells is also a major step in evolution. Bet hedging strategies were not only present at elementary stages of evolution. They are a common adaptive tool that can still be found in nature at different levels of complexity, from prokaryotic organisms to vertebrate ones. In between, tumour cells, thanks to their high plasticity, in the presence of an aggressive environment provided by immune response of the host body or of any anti-cancer treatment, may adopt bet hedging as a strategy to guarantee a prolonged survival of their colony. The wide presence of bet hedging in nature as an evolutionary mechanism, and its many links to the development of cancer is what motivates us in the present attempt towards a mathematical model representing some of the factors that influence this phenomenon (natural selection, epimutations and environmental stress).

In our modelling choices, we have privileged, for the sake of biological interpretation, as in [111], two phenotypes that are often identified as such in theoretical ecology: viability (potential to resist a deadly insult: the elephant strategy) and fecundity (potential to proliferate, or in a way, surviving by numbers, even under hard environmental conditions: the rat strategy). They may influence cell behaviour in opposed directions, as two different choices, incompatible for the same cell, in the same way as fecundity and motility are notably incompatible (cells that divide do not move, and vice versa). To take into account the faculty of rapidly changing phenotypes in case of a life-threatening insult, a capacity reported about many cancer cells (e.g., epithelial to mesenchymal transition or drug-induced drug persistence) [64], here we have added plasticity as a complementary structuring phenotype (or trait) of cells. Among questions at stake we will in particular deal with is the optimisation of fitness strategy by concentration of phenotypes, either in two or more ranges surrounding fixed points (multimodality of traits, i.e., bet hedging) or around a central optimal point (unimodality of trait, i.e. no bet hedging when it is not favourable).

3.2 The model

We consider a population (not necessarily of tumour cells) in which each individual has three defining traits: viability associated with the variable $x \in [0,1]$ which reflects the potential to resist deadly insults, fecundity associated with the variable $y \in [0,1]$ representing the potential to proliferate and plasticity associated with the variable $\theta \in [0,1]$ which represents the potential to continue to differentiate (or de-differentiate, or transdifferentiate, as long as it has not been fixed at its lower bound θ=0) within a differentiation tree. We assume furthermore that for a certain regular function $C : \mathbb{R}^2 \to \mathbb{R}$ and a positive constant K, $(x, y) \in \Omega := \{C(x, y) \leqslant K\}$, so $z = (x, y, \theta)$ ranges over the set $D := \{\Omega \times [0, 1]\}$. We then consider the evolution problem (3.1), (3.2), (3.3) on the density of population $n = n(t, z) \geqslant 0$.

$$\partial_t n + \nabla \cdot \Big(Vn - A(\theta)\nabla n \Big) = (r(z) - d(z)\rho(t))n, \tag{3.1}$$

$$\Big(Vn - A(\theta)\nabla n \Big) \cdot \mathbf{n} = 0, \text{ for all } z \in \partial D, \tag{3.2}$$

$$n(0, z) = n_0(z), \text{ for all } z \in D. \tag{3.3}$$

In the above equation, chosen for the sake of simplicity as diagonal, the matrix

$$A(\theta) = \begin{pmatrix} a_{11}(\theta) & 0 & 0 \\ 0 & a_{22}(\theta) & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

gives the speed at which non-genetic epimutations occur, otherwise said it is a minimally simple representation of how the internal plasticity trait θ affects the non-genetic instability of traits x and y, by tuning the diffusion term; the function

$$V(t, z) = (V_1(t, z), V_2(t, z), V_3(t, z))$$

represents the sensitivity, meant here as the force of external evolutionary pressure, of the population to abrupt changes in the environment;

$$\rho(t) = \int_D n(t, z)dz$$

stands for the total amount of individuals in the population at time t.
From a biological point of view, the matrix $A(\theta)$ represents random epimutations leading to non genetic changes, increasing with plasticity θ, in the traits viability x and fecundity y, while V stands for an external factor (abrupt biophysical changes in the ecosystem, exposure to life-threatening drugs in the case of cancer cell populations) inducing cellular stress in the population. Throughout the chapter, we refer to $\nabla \cdot (Vn)$ as advection term, or drift term, indistinctly.
We chose the term $d(z)\rho(t)$ here as the simplest way to represent a death term in the evolution of the population. This multiplicative representation is classic, but most of the results we show in what follows can be extended to more general reaction terms.
Throughout our work we assume

(H1) *For some* $p > 2$*, the initial population density* $n_0(z)$ *belongs to* $L^p(D)$*.*

(H2) *The intrinsic growth rate in absence of competition* $r(z)$ *and the death rate* $d(z)$ *due to competition for individuals with trait* z *are positive bounded functions that satisfy* $0 < r^- \leqslant r(z) \leqslant r^+$ *and* $0 < d^- \leqslant d(z) \leqslant d^+$*.*

(H3) *The diffusion parameters* $a_{11}(\theta)$*,* $a_{22}(\theta)$ *and* a_{33} *are strictly positive, with* $a_{11}(\theta)$ *and* $a_{22}(\theta)$ *being non decreasing with respect to* θ*. Hence, the matrix* $A(\theta)$ *is elliptic for all values of* θ*.*

(H4) *The function* $V(t, z)$ *is continuously differentiable for all values of* $t > 0$ *and* $z \in D$*.*

Under these hypotheses we are able to prove the existence and uniqueness of a solution for the problem (3.1)-(3.3) using the Finite Volume method in order to obtain a convergent semi-discrete scheme.

The problem (3.1)-(3.3) underlying hypotheses *(H1)* to *(H4)* sets a structured population model of evolution that, as mentioned before, takes into account some of the factors that might lead to the occurrence of "bet hedging". Amongst the first works in this topic, we can find [112], where is studied the fraction of seeds that germinate and the fraction that remains dormant, in order to maximise the long term expectation of growth. In the same work, the similarities of this phenomenon with economic decision making under risk (so-called "fail-safe strategies") are noted. Other early works on the subject are [69] and [70]. The reaction term in (3.1) is a simple way of modelling the selection principle. This term and more general ones are used in [22] in order to study some basic properties of structured populations undergoing this type of behaviour. The same reaction term is also used in [28], where are analysed the global asymptotic stability properties for integro-differential systems of N species structured by different sets of traits. A similar competition term is used in [24] to provide results about the long time behaviour of such reaction models. The diffusion term here models non-genetic instabilities (also known as epimutations), which constitute the drift of phenotype without alteration of the genotype. In [22] an integral operator in order to model mutations arising during reproduction is used, and something similar could be done for the epimutations. The second order operator used in (3.1) can be obtained then after re-scaling the time variable. The effect of this phenomenon in cancer development is discussed in [113] from a biological point of view. Epimutations can also occur because of external stress, and this is represented in (3.1) by means of the advection terms. A biological example for a population changing phenotype due to external stress can be found in [114], where the effect of physical stress on the shape and the cell wall thickness of E.coli bacterias is discussed. Two different models are used in [111] to conclude that the three mechanisms described above might reversibly push an actively-proliferating, and drug-sensitive, cell population to transition into a weakly-proliferative and drug-tolerant state, which will eventually facilitate the emergence of more potent, proliferating and drug-tolerant cells. One of such models is an integro-differential model very similar to (3.1)-(3.3), but without including the effect of plasticity on the evolution of the population and without assuming the existence of a constraint between the traits x and y.

The main results of this chapter involve the variational formulation of (3.1)-(3.3), which we now introduce. Denote $H = L^2(D)$, with $(\cdot,\cdot)_H$ the usual scalar product in that space, and $\mathscr{V} = H^1(D)$ with $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{\mathscr{V}'\times\mathscr{V}}$ being the duality product in $\mathscr{V}$.
For any given $n_0 \in H$, $T > 0$, we say that

$$n := n(t) \in X_T := C([0,T], H) \cap L^2((0,T), \mathscr{V}) \cap H^1([0,T], \mathscr{V}'),$$

is a variational solution of the problem (3.1)-(3.3) if it is a solution in the following weak sense

$$(n(t),\varphi(t))_H =(n_0,\varphi(0))_H + \int_0^t \Big(\langle Q[n](s),\varphi(s)\rangle + \langle \partial_s\varphi(s), n(s)\rangle\Big)ds, \tag{3.4}$$

where

$$\langle Q[n],\varphi\rangle = \int\limits_D \Big(- A\nabla n\nabla\varphi + Vn\nabla\varphi + (r(z)-\rho d(z))n\varphi\Big)dz,$$

for any $\varphi\in X_T$. We say that n is a global solution if it is a solution on $[0,T]$ for any $T>0$.

Theorem 3.1. *For all non-negative $n_0\in L^p(D)$, $p>2$, there exists a unique global non-negative weak solution for problem (3.1)-(3.3) in the sense of (3.4).*

We focus on giving a proof for this theorem using a discretised version of problem (3.1)-(3.3) after applying the Finite Volume Method to it. For this purpose, we define a set $D_h\supset D$, that can be covered by the union of N disjoint cubic cells, denoted as D_j, of side length h. After integrating the equation (3.1) over each of the cells D_j we derive the system of first-order differential equations

$$\frac{d}{dt}\nu_j(t) =M_j(t,\nu(t))\nu_j(t) + \sum_{l\in N_j} B_{jl}(t)\nu_l(t), \tag{3.5}$$

$$\nu_j(0) =\frac{1}{h^3}\int\limits_{D_j} n_0(z)dz, \tag{3.6}$$

where ν_j is an approximation of the average of the solution $n(t,z)$ over D_j, N_j is the set of indexes corresponding to the neighbours of D_j and the coefficients M_j and B_{jl} are functions of $V(t,z)$, $A(\theta)$, $r(z)$ and $d(z)$. A full detailed derivation of the scheme is given in Section (3.3). We can then introduce the following result involving the solution for this system:

Theorem 3.2. *For all non-negative $n_0\in L^p(D)$, $p>2$, there exists a unique non-negative solution for problem (3.5)-(3.6). Furthermore, the function $\tilde n_h(t,z)$ defined by*

$$\tilde n_h(t,z)=\sum_j \nu_j(t)\mathbb{1}_{D_j\cap D},$$

converges in $L^2(D_T)$ to the unique non-negative weak solution of (3.1)-(3.3) as h goes to zero.

The existence and non-negativity of the solution for (3.5)-(3.6) results from the Cauchy-Lipschitz theorem, while the convergence of $\tilde n_h(t,z)$ is the consequence of the compactness of the sequence.

The existence result in Theorem (3.1) will be treated in Section (3.3), but the uniqueness can be directly obtained from the variational formulation with the help of some *a posteriori* estimates. Let us proceed then to prove the uniqueness before addressing the existence. Assume that there exists a non-negative weak solution n for (3.1)-(3.3). Assume as well that there exist $0\leqslant t_0<t_1$ such that $\rho(t)\leqslant \max\{\rho(0),\frac{r^+}{d^-}\}$ for $t\in[0,t_0]$, and $\rho(t)>\max\{\rho(0),\frac{r^+}{d^-}\}$ for $t\in(t_0,t_1)$. We may take on the variational formulation $\varphi(t,z)=\chi_\varepsilon(t)$, where $\chi_\varepsilon(t)$ is a sequence satisfying $\chi_\varepsilon(t)\to\mathbb{1}_{t>t_0}(t)$ and $\chi'_\varepsilon(t)\to\delta_{t_0}(t)$. For example, we could consider $\chi_\varepsilon(t)=\frac12(1-\cos(\frac{\pi(t-t_0+\varepsilon)}{2\varepsilon}))\mathbb{1}_{t\in[t_0,t_0+\varepsilon]}+\mathbb{1}_{t>t_0+\varepsilon}$. This leads to the equality

$$(n(t),\chi_\varepsilon(t))_H = \int_0^t \int\limits_D (r(z)-d(z)\rho(t))n(t,z)dz\chi_\varepsilon(s) + \langle\chi'_\varepsilon(s),n(s)\rangle\Big)ds.$$

Taking the limit when ε goes to 0, we obtain that, for all $t > t_0$

$$\rho(t) = \rho(t_0) + \int_{t_0}^{t} \int_{D} (r(z) - d(z)\rho(s))n(s,z)dzds,$$

where $\rho(t) = \int_D n(t,z)dz$ is a continuous function due to the fact that $n(t,z) \in \mathscr{C}([0,T],H)$. We can write the previous relation as

$$\rho(t) = \rho(t_0) + \int_{t_0}^{t} \int_{D} d(z) \left(\frac{r(z)}{d(z)} - \frac{r^+}{d^-} + \frac{r^+}{d^-} - \rho(s) \right) n(s,z)dzds,$$

which, thanks to the hypothesis on $\rho(t)$ over (t_0, t_1), leads to

$$\rho(t) \leqslant \rho(t_0) \leqslant \max \left\{ \rho(0), \frac{r^+}{d^-} \right\}$$

for all $t \in (t_0, t_1)$. This represents a contradiction, and implies that

$$\rho(t) \leqslant \max \left\{ \rho(0), \frac{r^+}{d^-} \right\} \tag{3.7}$$

for all values of t.
Taking now $\varphi = n$ on the variational formulation, using standard arguments to bound the linear terms from $Q[n]$ and the estimate (3.7) for the non-linear part together with the Gronwall Lemma, we obtain the relation

$$\frac{1}{2}\|n(t)\|_H^2 \leqslant \frac{1}{2}\|n(0)\|_H^2 + a \int_0^t \frac{1}{2}\|n(s)\|_H^2 ds,$$

for some real positive number a. Using Gronwall's lemma, this relation implies that

$$\|n(t)\|_H^2 \leqslant \|n_0\|_H^2 e^{2at},$$

for all values of t. Finally assume the existence of two non-negative weak solutions n_1 and n_2 for the same initial data n_0. Taking the difference between their respective variational formulations, choosing $\varphi = n_1 - n_2$, we get the equality

$$\frac{1}{2}\|n_1 - n_2\|_H^2 = \int_0^t \langle Q[n_1] - Q[n_2], n_1 - n_2 \rangle ds.$$

Once again, the linear part of $Q[n_1] - Q[n_2]$ can be easily bounded using standard methods, leading to

$$\begin{aligned}
\frac{1}{2}\|n_1 - n_2\|_H^2 \leqslant & a \int_0^t \|n_1 - n_2\|_H^2 ds - \int_0^t \rho_1(s) \int_D d(z)(n_1 - n_2)^2 dzds \\
& - \int_0^t (\rho_1(t) - \rho_2(t)) \int_D n_2 d(z)(n_1 - n_2)dzds \\
& \leqslant \int_0^t (a + d^+ \|n_0\|_H e^{as}) \|n_1 - n_2\|_H^2 ds
\end{aligned}$$

Consequently, thanks to Gronwall's lemma, $\|n_1 - n_2\|_H^2 = 0$ for all t, therefore, the solution is unique.

As stated before, the proof of existence of a weak solution will be carried on in Section (3.3). Following the ideas from [115], the semi-discrete numerical scheme (3.5)-(3.6) is developed and the convergence of its solution to the solution of (3.1)-(3.3) is demonstrated. A fully discrete numerical scheme is obtained starting from the semi-discrete one and its convergence is also proved. Section 4 is devoted to the numerical simulations, starting with some numerical computations of the approximation error by comparing the results with an exact solution. Finally, the solutions of some examples with biological meaning is presented.

3.3 Existence of a weak solution and numerical approximation

We aim to use the Finite Volume method in order to find a sequence of problems whose solutions converge to the solution of (3.4).

3.3.1 Preliminaries on the finite volume method

Consider $h = 1/M$ where M is a natural number and define the mesh

$$C_{ijk} = [\frac{i}{M}, \frac{i+1}{M}] \times [\frac{j}{M}, \frac{j+1}{M}] \times [\frac{k}{M}, \frac{k+1}{M}],$$

with $i, j, k = 0, \ldots, M-1$, such that $\bigcup_{i,j,k} C_{ijk} = [0,1]^3$.

Now introduce the sets

$$\mathscr{M} = \{C_{ijk} : C_{ijk} \cap D \neq \emptyset\}$$

and

$$D_h = \bigcup_{\mathscr{M}} C_{ijk}.$$

For simplicity, we define $N := |\mathscr{M}|$ as the amount of elements in $\mathscr{M}$, and denote each of its elements as D_j, for $j = 1, \ldots, N$. For each D_j, we denote its centre of mass as $\mathbf{z_j} := (x_j, y_j, \theta_j)$. For each j, define N_j as the set of indexes l such that D_l and D_j share a common boundary. Denote such common boundary as Γ_{jl}, its centre of mass as $\mathbf{z_{jl}}$ and $\mathbf{n}_{jl}$ the outer normal vector of D_j, in the direction of D_l. We remark that the distance between the centres of two neighbouring cells D_j and D_l will be equal to $|\mathbf{z_l} - \mathbf{z_j}|$. Having cubes as the mesh cells guarantees that the condition

$$\mathbf{n}_{jl} = \frac{\mathbf{z_l} - \mathbf{z_j}}{|\mathbf{z_l} - \mathbf{z_j}|}, \tag{3.8}$$

is fulfilled.

It is important to remark that if D has a regular enough boundary (for example: smooth or polygonal), then the area of $D_h \setminus D$, which we denote as $|D_h \setminus D|$, will converge to zero as h vanishes.

The approximated problem (3.1)-(3.3) is then given by

$$\partial_t \tilde{n}_h + \nabla \cdot \Big(V\tilde{n}_h - A(\theta)\nabla\tilde{n}_h \Big) = (r(z) - d(z)\tilde{\rho}_h(t))\tilde{n}_h, \text{ in } D_h \tag{3.9}$$

$$\Big(V\tilde{n}_h - A(\theta)\nabla\tilde{n}_h \Big) \cdot \mathbf{n} = 0, \text{ for all } z \in \partial D_h, \tag{3.10}$$

$$\tilde{n}_h(0,z) = n_0(z), \text{ for all } z \in D_h, \tag{3.11}$$

where $\tilde{\rho}_h(t) = \int\limits_{D_h} \tilde{n}_h(t,z)dz$. We propose a classical finite volume method based on local averages of the unknown density over cell grids defined by

$$\nu_j(t) := \frac{1}{h^3}\int\limits_{D_j} \tilde{n}_h(t,z)dz = n(t,\mathbf{z_j}) + \mathcal{O}(h^2). \tag{3.12}$$

Assume that the coefficients and the solution from equation (3.9) are smooth. Then, integrating it over a cell D_j yields the equality

$$\frac{d}{dt}\int\limits_{D_j} \tilde{n}_h dz = -\int\limits_{D_j} \nabla \cdot \Big(V\tilde{n}_h - A\nabla\tilde{n}_h \Big) dz + \int\limits_{D_j} r(z)\tilde{n}_h dz - \tilde{\rho}_h(t)\int\limits_{D_j} d(z)\tilde{n}_h dz.$$

After integrating by parts and using the boundary conditions (3.10), we get

$$-\int\limits_{D_j} \nabla \cdot \Big(V\tilde{n}_h - A\nabla\tilde{n}_h \Big) dz = -\sum_{l\in N_j}\int\limits_{\Gamma_{jl}} \Big(V\tilde{n}_h - A\nabla\tilde{n}_h \Big) \cdot \mathbf{n}_{jl} dS.$$

For a real function $f(t)$, define the positive and negative part of f as

$$f^+(t) = \begin{cases} f(t), & \text{if} \quad f(t) \geqslant 0, \\ 0, & \text{if} \quad f(t) < 0, \end{cases}$$

and

$$f^-(t) = \begin{cases} 0, & \text{if} \quad f(t) \geqslant 0, \\ f(t), & \text{if} \quad f(t) < 0, \end{cases}$$

respectively. Using an upwind approximation technique for the advection term, we conclude

$$\int\limits_{\Gamma_{jl}} V\tilde{n}_h \cdot \mathbf{n}_{jl} dS = |\Gamma_{jl}|\Big(\nu_j(t)u_{jl}^+(t) + \nu_l(t)u_{jl}^-(t) \Big) + \mathcal{O}(h^2), \tag{3.13}$$

where $u_{jl}(t) = V(t,\mathbf{z_{jl}}) \cdot \vec{\mathbf{n}}_{jl}$. On the other hand, we have

$$-\int\limits_{\Gamma_{jl}} \Big(A\nabla\tilde{n}_h \Big) \cdot \mathbf{n}_{jl} dS = -|\Gamma_{jl}|\Big(A(\theta_{jl})\nabla\tilde{n}_h(\mathbf{z_{jl}}) \Big) \cdot \mathbf{n}_{jl} = -|\Gamma_{jl}|\nabla\tilde{n}_h(\mathbf{z_{jl}}) \cdot A(\theta_{jl})\mathbf{n}_{jl}.$$

As $A(\theta)$ is a diagonal matrix and $\mathbf{n}_{jl}$ is either one of the vectors from the euclidean canonical base, or one of their opposites, we have the relation

$$A(\theta_{jl})\mathbf{n}_{jl} = A_{jl}\mathbf{n}_{jl},$$

where $A_{jl} = (A(\theta_{jl})\mathbf{n}_{jl} \cdot \mathbf{n}_{jl})$. Notice that, thanks to hypothesis *(H3)*, there exists $\alpha > 0$ such that $A_{jl} \geqslant \alpha$, for all j and l. So that, together with the expression of the normal vectors, (3.8) this implies

$$-\int_{\Gamma_{jl}} \Big(A\nabla\tilde{n}_h\Big) \cdot \mathbf{n}_{jl} dS = -\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|}\Big(\nabla\tilde{n}_h(\mathbf{z_{jl}})\Big) \cdot \Big(\mathbf{z_l} - \mathbf{z_j}\Big).$$

Due to the approximation of the gradient

$$\Big(\nabla\tilde{n}_h(\mathbf{z_{jl}})\Big) \cdot \Big(\mathbf{z_j} - \mathbf{z_l}\Big) = \tilde{n}_h(\mathbf{z_l}) - \tilde{n}_h(\mathbf{z_j}) + \mathcal{O}(h^2),$$

we can finally write

$$\begin{aligned} -\int_{\Gamma_{jl}} \Big(A\nabla\tilde{n}_h\Big) \cdot \mathbf{n}_{jl} dS &= -\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|}\Big(\tilde{n}_h(\mathbf{z_j}) - \tilde{n}_h(\mathbf{z_l})\Big) + \mathcal{O}(h^2) \\ &= -\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|}\Big(\nu_l(t) - \nu_j(t)\Big) + \mathcal{O}(h^2). \end{aligned} \tag{3.14}$$

Taking into account that

$$\rho_h(t) = \int_{D_h} \tilde{n}_h(t,z)dz = \sum_{l=1}^{N} \int_{D_l} \tilde{n}_h(t,z)dz = \sum_{l=1}^{N} h^3 \nu_l(t),$$

the reaction terms can be easily approximated

$$\begin{aligned} \int_{D_j} r(z)\tilde{n}_h dz - \tilde{\rho}_h(t)\int_{D_j} d(z)\tilde{n}_h dz &= h^3\Big(r(\mathbf{z_j}) - \rho_h(t)d(\mathbf{z_j})\Big)\tilde{n}_h(t,\mathbf{z_j}) + \mathcal{O}(h^5) \\ &= h^3\Big(r_j - \tilde{\rho}_h(t)d_j\Big)\nu_j(t) + \mathcal{O}(h^5) \\ &= h^3\Big(r_j - d_j\sum_{l=1}^{N} h^3\nu_l(t)\Big)\nu_j(t) + \mathcal{O}(h^5), \end{aligned} \tag{3.15}$$

where we have adopted the notation $r_j := r(\mathbf{z_j})$ and $d_j := d(\mathbf{z_j})$. Finally, using again (3.12), we get

$$\frac{d}{dt}\int_{D_j} \tilde{n}_h(z)dz = h^3\nu_j'(t).$$

Consequently, collecting (3.13), (3.14) and (3.15), and getting rid of the approximation orders we obtain the semi-discrete scheme

$$\frac{d}{dt}\nu_j(t) = M_j(t,\tilde{\rho}_h(t))\nu_j(t) + \sum_{l\in N_j} B_{jl}(t)\nu_l(t), \tag{3.16}$$

where

$$\begin{aligned} M_j(t,\tilde{\rho}_h(t)) &= -\sum_{l\in N_j} \frac{|\Gamma_{jl}|}{h^3}\Big(u_{jl}^+(t) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|}\Big) + \Big(r_j - d_j\tilde{\rho}_h(t)\Big), \\ B_{jl} &= \frac{|\Gamma_{jl}|}{h^3}\Big(-u_{jl}^-(t) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|}\Big). \end{aligned}$$

This system of equations can be complemented with the set of initial data

$$\nu_j(0) = \nu_j^0 := \frac{1}{h^3} \int_{D_j} \overline{n}_0(z) dz, \tag{3.17}$$

where $\overline{n}_0(z)$ is the extension by 0 of $n_0(z)$ to all of $\mathbb{R}^3$.

3.3.2 Global existence, uniqueness, positivity and boundedness of the solution for the semi-discrete scheme

We prove the local existence and uniqueness of the solution for the problem (3.16)-(3.17), by using the Cauchy-Lipschitz theorem. Then, such solution is proved to be non-negative and, as a consequence, bounded independently of t. Finally the boundedness property is used to prove the global existence of the solution.

Proposition 3.1 (Local existence of solution). *For all sets of initial data $\{\nu_j^0\}$, there exists $0 < T^* < \infty$ such that the problem (3.16)-(3.17) has an unique solution over $[0, T^*)$. Furthermore, if $\nu_j^0 \geqslant 0$ for all j, then $\nu_j(t) \geqslant 0$ for all time $t \in [0, T^*)$ and all j.*

Proof. The RHS term in (3.16) is Lipschitz continuous for all values of t and ν_j, therefore the existence and uniqueness of solution over a certain interval $[0, T^*)$ is a direct consequence of the Cauchy-Lipschitz theorem.
On the other hand, consider a strictly positive set of initial values ν_j^0 and define the continuous function $f(t) = \min_j \nu_j(t)$. If $f(t) \geqslant 0$ for all $t < T^*$, then the solution remains positive at all times. If $f(t) < 0$ for some $t \in (0, T^*)$, then there exists $t_0 > 0$ such that $f(t_0) = 0$ and $f(t) \geqslant 0$ for $t < t_0$. This implies the existence of j_0 such that $\nu_{j_0}(t_0) = 0$ with $\nu'_{j_0}(t_0) \leqslant 0$. If $\nu_l(t_0) > 0$ for some $l \in N_{j_0}$, then, thanks to (3.16) we have

$$\begin{aligned} \nu'_{j_0}(t_0) &= \sum_{l \in N_j} B_{jl}(t_0) \nu_l(t_0) \\ &= \sum_{l \in N_j} \frac{|\Gamma_{jl}|}{h^3} \Big(-u_{jl}^-(t_0) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|} \Big) \nu_l(t_0) > 0, \end{aligned}$$

which is a contradiction with the previously established fact that $\nu'_{j_0}(t_0) \leqslant 0$. Consequently $\nu_l(t_0) = 0$ for all $l \in N_{j_0}$. Furthermore, from the definitions of $f(t)$ and t_0, we also have that $\nu'_l(t_0) \leqslant 0$ for all $l \in N_{j_0}$. We can iterate the previous argument in order to obtain that $\nu_j(t_0) = 0$ for all j and consequently, thanks to the uniqueness of the solution, $\nu_j(t) \equiv 0$ for all j and all t. This is a contradiction with the assumption that $f(t)$ is negative for some value of t and consequently $f(t) \geqslant 0$ for all t.
Finally, for non-negative initial values ν_j^0 and ε small enough, we define ν_j^ε as

$$\nu_j^\varepsilon = \begin{cases} \nu_j^0 & \text{if } \nu_j^0 > 0, \\ \varepsilon & \text{if } \nu_j^0 = 0. \end{cases}$$

Thanks to the previous step, the solution of (3.16) associated to ν_j^ε remains non-negative for all t and all ε. Hence, thanks to the continuous dependence of the solution of a system with respect to its initial data, we conclude that $\nu_j^0 \geqslant 0$ implies $\nu_j(t) \geqslant 0$ for all $t \in [0, T^*)$. □

Proposition 3.2 (Discrete L^1 bound). *The L^1 norm $\tilde{\rho}_h(t) := \sum_l h^3 \nu_l(t)$ satisfies the bounds*

$$\underline{\rho} := \min\{\tilde{\rho}_h(0), \frac{r^-}{d^+}\} \leqslant \tilde{\rho}_h(t) \leqslant \max\{\tilde{\rho}_h(0), \frac{r^+}{d^-}\} =: \overline{\rho}, \ \forall t \geqslant 0, \tag{3.18}$$

where r^-, r^+, d^- and d^+ are the bounds given in (H1) for $r(z)$ and $d(z)$ respectively.

Proof. Multiplying (3.16) by h^3 for each j, adding up all the equations, recalling that $|\Gamma_{jl}| = |\Gamma_{lj}|$, $u^+_{jl} = -u^-_{lj}$, $A_{jl} = A_{lj}$ and $|\mathbf{z_l} - \mathbf{z_j}| = |\mathbf{z_j} - \mathbf{z_l}|$ we obtain that

$$\tilde{\rho}'_h(t) = \sum_{j=1}^{N} \Big(r_j - d_j \tilde{\rho}_h(t) \Big) h^3 \nu_j(t).$$

The non-negativity of the solution implies then

$$\Big(r^- - d^+ \tilde{\rho}_h(t) \Big) \tilde{\rho}_h(t) \leqslant \tilde{\rho}'_h(t) \leqslant \Big(r^+ - d^- \tilde{\rho}_h(t) \Big) \tilde{\rho}_h(t).$$

These differential inequalities directly imply the bounds over $\tilde{\rho}_h(t)$. □

Notice that the L^1 bounds are independent of t. Additionally, the upper bound also implies that

$$h^3 \nu_j(t) \leqslant \tilde{\rho}_h(t) \leqslant \overline{\rho}, \text{ for all } t \in [0, T^*), \text{ for all } j,$$

So that in general $\nu_j(t) \leqslant \frac{\overline{\rho}}{h^3}$, which is the key estimate in order to prove global existence of solution for (3.16)-(3.17).

Proposition 3.3 (Global existence of solution). *For all sets of initial data $\{\nu_j^0\}$, there exists a unique solution of problem* (3.16)-(3.17) *for all $t > 0$. Such solution is non-negative and satisfies the estimate* (3.18).

Proof. For each h, assume that there exists a finite maximal time of existence T_h. However, the estimate (3.18) on $\tilde{\rho}_h(t)$ implies that for all j, $\nu_j(T_h) \leqslant \frac{\overline{\rho}}{h^3} < \infty$, which, thanks to the Lipschitz continuity of the right hand side of (3.16) allows to extend to solution to a certain interval $[T_h, T_h^*)$, contradicting this way the maximality of T_h. □

3.3.3 Discrete gradient, L^2 norm estimate and compactness result

In this section we introduce some piecewise constant functions depending on the solution of (3.16)-(3.17) together with some estimates related to such functions. Then, some compactness properties will be proved in order to ensure that such functions converge to some function that will be proved to be a weak solution of (3.1)-(3.2), and their derivatives, respectively.

We first introduce $n_h(t,z)$ defined as

$$n_h(t,z) = \sum_{j=1}^{N} \nu_j(t) \mathbb{1}_{D_j}(z).$$

Notice that $\|n_h\|_{L^1} = \tilde{\rho}_h(t)$. Now, for each $l \in N_j$, define the polygonal subsets of D_j, denoted D_{jl}, having Γ_{jl} as the common side and $\mathbf{z_j}$ as a vertex. The subsets D_{jl} are pyramids of area $s_{jl} = \frac{|\Gamma_{jl}| d(\mathbf{z_j}, \Gamma_{jl})}{3} = \frac{h^3}{6}$. Let us define the piecewise constant function

$$v_h(t,z) = \sum_{j=1}^{N} \sum_{l \in N_j} \frac{|\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|} \Big((\mathbf{z_{jl}} - \mathbf{z_j}) \frac{\nu_l(t) - \nu_j(t)}{s_{jl}} \mathbb{1}_{D_{jl}}(z) \Big).$$

This function can be regarded as a discrete gradient for $n_h(t)$.

Proposition 3.4 (L^2 bound)**.** *For each value of h, define the space $H_h := L^2(D_h)$. Then, there exists positive constants a and b, independent of h, such that the functions $n_h(t,z)$ and $v_h(t,z)$ satisfy the following estimate*

$$\|n_h\|^2_{H_h} + a \int_0^T \|v_h\|^2_{H_h} \leqslant e^{bT} \|n_0\|^2_{H_h}, \text{ for all } T > 0. \tag{3.19}$$

Proof. Multiplying equation (3.16) by $h^3 \nu_j(t)$ for each j, and adding them up, we obtain the relation

$$\sum_{j=1}^{N} h^3 \nu_j(t) \nu_j'(t) = A(t) + D(t) + R(t), \tag{3.20}$$

where

$$A(t) = -\sum_{j=1}^{N} \nu_j(t) \sum_{l \in N_j} |\Gamma_{jl}| \Big(\nu_j(t) u^+_{jl}(t) + \nu_l(t) u^-_{jl}(t) \Big),$$

$$D(t) = \sum_{j=1}^{N} \nu_j(t) \sum_{l \in N_j} \frac{A_{jl} |\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|} \Big(\nu_l(t) - \nu_j(t) \Big),$$

$$R(t) = \sum_{j=1}^{N} h^3 \Big(r_j - d_j \sum_{l} h^3 \nu_l(t) \Big) \nu_j^2(t).$$

Let us proceed to estimate each of these terms. In order to simplify the notation, we define

$$\mu_{jl}(t) = |\Gamma_{jl}| (\nu_j(t) u^+_{jl}(t) + \nu_l(t) u^-_{jl}(t)),$$

and write

$$A(t) = -\frac{1}{2} \sum_{j=1}^{N} \sum_{l \in N_j} \Big(\nu_j(t) \mu_{jl}(t) + \nu_l(t) \mu_{lj}(t) \Big).$$

Consequently, knowing that $-u_{jl}^{+} = u_{lj}^{-} \leqslant 0$, we have

$$
\begin{aligned}
A(t) &= \frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nu_l(t)-\nu_j(t)\Big)\Big(u_{jl}^{+}(t)\nu_j(t)+u_{jl}^{-}(t)\nu_l(t)\Big)\\
&= \frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nu_l(t)-\nu_j(t)\Big)\Big(u_{jl}(t)\nu_j(t)+u_{jl}^{-}(t)(\nu_l(t)-\nu_j(t))\Big)\\
&\leqslant \frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nu_l(t)-\nu_j(t)\Big)\Big(u_{jl}(t)\nu_j(t)\Big)\\
&= \sum_{j=1}^{N}\sum_{l\in N_j}\frac{|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}|\mathbf{z_{jl}}-\mathbf{z_j}|\Big(\nu_l(t)-\nu_j(t)\Big)\Big(u_{jl}(t)\nu_j(t)\Big),
\end{aligned}
$$

where we used that $|\mathbf{z_l}-\mathbf{z_j}| = 2|\mathbf{z_{jl}}-\mathbf{z_j}|$. From the definition of u_{jl}, we conclude that

$$|u_{jl}| \leqslant |V(t,\mathbf{z_{jl}})\cdot\vec{\mathbf{n}}_{jl}| \leqslant \overline{V},$$

where $\overline{V} := \max_{z,t}|V(t,z)|$. This implies

$$
\begin{aligned}
|A(t)| &\leqslant \overline{V}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}|\mathbf{z_{jl}}-\mathbf{z_j}|\Big(|\nu_l(t)-\nu_j(t)|\Big)\nu_j(t)\\
&= \overline{V}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}|\mathbf{z_{jl}}-\mathbf{z_j}|\frac{\Big(|\nu_l(t)-\nu_j(t)|\Big)}{s_{jl}}\nu_j(t)s_{jl}\\
&= \overline{V}\int_D |v_h(t,z)||n_h(t,z)|dz.
\end{aligned}
$$

Then, for all $\varepsilon > 0$, Young's inequality implies

$$|A(t)| \leqslant \overline{V}\Big(\frac{\varepsilon}{2}\|v_h\|_{H_h}^2 + 2\varepsilon^{-1}\|n_h\|_{H_h}^2\Big). \tag{3.21}$$

On the other hand

$$
\begin{aligned}
D(t) &= \frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_j(t)\Big(\nu_l(t)-\nu_j(t)\Big)+\nu_l(t)\Big(\nu_j(t)-\nu_l(t)\Big)\Big)\\
&= -\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_l(t)-\nu_j(t)\Big)^2\\
&= -\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}s_{jl}|\mathbf{z_l}-\mathbf{z_j}|}{|\Gamma_{jl}||\mathbf{z_{jl}}-\mathbf{z_j}|^2}\frac{|\Gamma_{jl}|^2|\mathbf{z_{jl}}-\mathbf{z_j}|^2}{d_{jl}^2}\frac{\Big(\nu_l(t)-\nu_j(t)\Big)^2}{s_{jl}^2}s_{jl}.
\end{aligned}
$$

The ellipticity and boundedness of the matrix $A(\theta)$ imply that there exist positive constants $\underline{\alpha}$, $\overline{\alpha}$ such that $\underline{\alpha} \leqslant A_{jl} \leqslant \overline{\alpha}$ for all j and l. From this and the value of s_{jl} we deduce that

$$\frac{A_{jl}s_{jl}|\mathbf{z_l}-\mathbf{z_j}|}{|\Gamma_{jl}||\mathbf{z_{jl}}-\mathbf{z_j}|^2} = \frac{4A_{jl}h^4}{6h^4} \geqslant \frac{2\underline{\alpha}}{3},$$

and consequently

$$D(t) \leqslant -\frac{\alpha}{3}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{|\Gamma_{jl}|^2|\mathbf{z_{jl}}-\mathbf{z_j}|^2}{d_{jl}^2}\frac{\Big(\nu_l(t)-\nu_j(t)\Big)^2}{s_{jl}^2}s_{jl}$$

$$= -\frac{\alpha}{3}\|v_h\|_{H_h}^2, \tag{3.22}$$

notice that the last identity is true since v_h is a piecewise constant function. Using the bounds over $r(z)$ and the positiveness of $d(z)$ and $\tilde{\rho}_h(t)$ we see that

$$R(t) := \sum_{j=1}^{N} h^3\Big(r_j - d_j\sum_{l} h^3\nu_l(t)\Big)\nu_j^2(t) = \sum_{j=1}^{N} h^3\Big(r_j - d_j\tilde{\rho}_h(t)\Big)\nu_j^2(t)$$

$$\leqslant r^+\sum_{j=1}^{N} h^3\nu_j^2(t) = r^+\|n_h\|_{H_h}^2. \tag{3.23}$$

Using (3.21), (3.22) and (3.23) in (3.20), with the relation

$$\sum_{j=1}^{N} h^3\nu_j(t)\nu_j'(t) = \frac{1}{2}\Big(\sum_{j=1}^{N} h^3\nu_j^2(t)\Big)' = \frac{1}{2}\Big(\|n_h\|_{H_h}^2\Big)',$$

yields the differential inequality

$$\frac{1}{2}\Big(\|n_h\|_{H_h}^2\Big)' + \Big(\frac{\alpha}{3} - \frac{\varepsilon\overline{V}}{2}\Big)\|v_h\|_{H_h}^2 \leqslant \Big(\frac{\varepsilon^{-1}\overline{V}}{2} + r^+\Big)\|n_h\|_{H_h}^2, \tag{3.24}$$

which, after taking $\varepsilon = \frac{\alpha}{3\overline{V}}$ and using Gronwall's lemma, leads to the estimate (3.19). □

The result of Proposition (3.4) can be easily generalised if instead of multiplying each equation (3.16) by $h^3\nu_j(t)$, we multiply by $h^3\nu_j^{p-1}(t)$, for any $p>1$, which would lead to the following uniform L^p bound:

Proposition 3.5 (L^p bound)**.** *There exists positive constants a_p and b_p, independents of h, such that the functions $n_h^p(t,z)$ and $v_h^p(t,z)$ defined as*

$$n_h^p(t,z) = \sum_{j=1}^{N}\nu_j^{p/2}(t)\mathbb{1}_{D_j}(z)$$

$$v_h^p(t,z) = \sum_{j=1}^{N}\sum_{l\in N_j}\frac{|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big((\mathbf{z_{jl}}-\mathbf{z_j})\frac{\nu_l^{p/2}(t)-\nu_j^{p/2}(t)}{s_{jl}}\mathbb{1}_{D_{jl}}(z)\Big).$$

satisfy the following estimate

$$\|n_h^p\|_{H_h}^2 + a_p\int_0^T\|v_h^p\|_{H_h}^2 \leqslant e^{b_pT}\|n_0\|_{L^p(D)}^2, \text{ for all } T>0. \tag{3.25}$$

Before stating the compactness result that will allow us to extract a convergent subsequence from n_h we give two important results that are a consequence from estimate (3.19).

Proposition 3.6. *For all vectors $\eta \in \mathbb{R}^3$ define the translation operator $\pi_\eta : L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ as $\pi_\eta(u)(z) = u(z+\eta)$. Then*

$$\lim_{\eta \to 0} \|\pi_\eta \overline{n}_h - \overline{n}_h\|_{L^2(\mathbb{R}^3 \times (0,T))} = 0,$$

uniformly in h, where $\overline{n}_h$ represents the extension by 0 of n_h to all of $\mathbb{R}^3$.

Proof. The proof of this proposition follows the same steps as Lemma 18.3 and Remark 18.8 in [116]. In particular, thanks to the discrete trace inequality given in the Lemma 10.5 within the same reference, we can prove the existence of a constant C, independent of h and η, such that

$$\|\pi_\eta \overline{n}_h - \overline{n}_h\|^2_{L^2(\mathbb{R}^3 \times (0,T))} \leqslant |\eta| C \left(\int_0^T \left(\|v_h\|^2_{H_h} + \|n_h\|^2_{H_h} \right) dt \right),$$

and we get the result from Proposition (3.6) thanks to (3.19). □

From now on we introduce the notation $D_T := (0,T) \times D$.

Proposition 3.7. *There exists a constant C_1, independent of h, such that for all $\psi \in \mathscr{D}(D_T)$,*

$$\left| \int_0^T \langle \frac{d\overline{n}_h}{dt}, \psi \rangle_{H^{-3} \times H^3} dt \right| \leqslant C_1 \sum_{|k| \leqslant 3} \|\partial_z^k \psi\|_{L^2(D_T)}.$$

Proof. Using the scheme (3.16), we have that, for all $\psi \in \mathscr{D}(D)$

$$\begin{aligned} \langle \frac{d\overline{n}_h}{dt}, \psi \rangle_{H^{-3} \times H^3} &= \sum_j \frac{d\nu_j(t)}{dt} \int_{D_j} \psi dz \\ &= \sum_j h^3 \Big(M_j(t, \tilde{\rho}_h(t)) \nu_j(t) + \sum_{l \in N_j} B_{jl}(t) \nu_l(t) \Big) \psi_j, \end{aligned}$$

where $\psi_j = h^{-3} \int_{D_j} \psi dz$ is the mean value of ψ over D_j. Reordering the terms in this equality we have

$$\begin{aligned} \langle \frac{d\overline{n}_h}{dt}, \psi \rangle_{H^{-3} \times H^3} =& \frac{1}{2} \sum_{j=1}^N \sum_{l \in N_j} |\Gamma_{jl}| \Big(\psi_l - \psi_j \Big) \Big(u^+_{jl}(t) \nu_j(t) + u^-_{jl}(t) \nu_l(t) \Big) \\ & - \frac{1}{2} \sum_{j=1}^N \sum_{l \in N_j} \frac{A_{jl} |\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|} \Big(\nu_l(t) - \nu_j(t) \Big) \Big(\psi_l - \psi_j \Big) \\ & + \sum_{j=1}^N h^3 \Big(r_j - d_j \tilde{\rho}_h(t) \Big) \nu_j \psi_j \end{aligned}$$

We can bound each term on this equality as follows: From the definition of $u_{jl}(t)$ we get that $|u^{\pm}_{jl}| \leqslant \|V\|_{L^\infty(D_T)}$ and consequently

$$\begin{aligned}
&\Big|\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\Big(\psi_l-\psi_j\Big)\Big(u^{+}_{jl}(t)\nu_j(t)+u^{-}_{jl}(t)\nu_l(t)\Big)\Big|\\
=&\Big|\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}h^3\left(\frac{\psi_l-\psi_j}{h}\right)\Big(u^{+}_{jl}(t)\nu_j(t)+u^{-}_{jl}(t)\nu_l(t)\Big)\Big|\\
\leqslant&\frac{1}{2}\|\psi\|_{\mathscr{C}^1(D)}\|V\|_{L^\infty(D_T)}\sum_{j=1}^{N}\sum_{l\in N_j}h^3\Big(\nu_j(t)+\nu_l(t)\Big)\\
\leqslant&3\|\psi\|_{\mathscr{C}^1(D)}\|V\|_{L^\infty(D_T)}\|n\|_{H_h}\,.
\end{aligned}$$

Using the previously established relation $A_{jl}\leqslant\overline{\alpha}$ we get

$$\begin{aligned}
&\Big|\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_l(t)-\nu_j(t)\Big)\Big(\psi_l-\psi_j\Big)\Big|\\
=&\Big|\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}A_{jl}h^3\Big(\frac{\nu_l(t)-\nu_j(t)}{h}\Big)\Big(\frac{\psi_l-\psi_j}{h}\Big)\Big|\\
\leqslant&\overline{\alpha}\|\psi\|_{\mathscr{C}^1(D)}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{h^3}{2}\Big|\frac{\nu_l(t)-\nu_j(t)}{h}\Big|\\
\leqslant&\overline{\alpha}\|\psi\|_{\mathscr{C}^1(D)}\sum_{j=1}^{N}\|v_h\|_{H_h}\,.
\end{aligned}$$

Finally, using the boundedness of $r(z)$, $d(z)$ and $\tilde{\rho}_h$ we see that

$$\begin{aligned}
\Big|\sum_{j=1}^{N}h^3\Big(r_j-d_j\tilde{\rho}_h(t)\Big)\nu_j\psi_j\Big| &\leqslant \sum_{j=1}^{N}h^3|r_j-d_j\tilde{\rho}_h(t)||\psi_j|\nu_j\\
&\leqslant (r^+ + d^+\overline{\rho})\|\psi\|_{\mathscr{C}^1(D)}\sum_{j=1}^{N}h^3\nu_j\\
&\leqslant (r^+ + d^+\overline{\rho})\|\psi\|_{\mathscr{C}^1(D)}\|n\|_{H_h}\,.
\end{aligned}$$

Putting everything together, we obtain the existence of a constant C independent of h, such that

$$|\langle\frac{d\overline{n}}{dt},\psi\rangle_{H^{-3}\times H^3}| \leqslant C\|\psi\|_{\mathscr{C}^1(D)}\Big(\|n\|_{H_h}+\|v_h\|_{H_h}\Big),\ \text{for all } t\in(0,T).$$

Finally, using the inclusions $H^3(D)\subset\mathscr{C}^{1,\frac{1}{2}}(D)\subset\mathscr{C}^1(D)$ that hold true in any open subset of $\mathbb{R}^3$ with smooth enough boundary thanks to the Sobolev inequalities, integrating over $(0,T)$, using the Cauchy-Schwartz inequality and estimate (3.19), we get to the result stated in the proposition. □

Proposition 3.8. *If $n_0\in L^p(D)$ for some $p>2$, then there exists a function $n\in L^2(0,T;\mathscr{V})$ such that, up to the extraction of a sub-sequence, the sequence of functions n_h strongly converges to n in $L^2(D_T)$ and v_h weakly converges to ∇n in $L^2(D_T)$.*

Proof. Propositions (3.6) and (3.7) give all the necessary tools in order to ensure the relative compactness of the set of functions $\{n_h\}_h$ in $L^2(D_T)$. In order to prove that such set is indeed compact, we will use a variant adapted to our purposes of the proof of [117, Theorem 3] to show that $\{n_h\}_h$ is a Cauchy sequence in $L^2(D_T)$. We consider a sequence of mollifiers $\Phi_\varepsilon(z) = \varepsilon^{-3}\Phi(\varepsilon^{-1}z)$ for a positive, symmetric function $\Phi \in \mathscr{D}(B(0,1))$ satisfying $\int_{\mathbb{R}^3}\Phi(z)dz = 1$.

Step 1: We claim that

$$\lim_{\varepsilon\to 0}\|\Phi_\varepsilon * \overline{n}_h - \overline{n}_h\|_{L^2((0,T)\times\mathbb{R}^3)} = 0,$$

uniformly on h.
We have

$$\begin{aligned}|\Phi_\varepsilon * \overline{n}_h(z) - \overline{n}_h(z)| &\leqslant \int_{\mathbb{R}^3} |\overline{n}_h(z-y) - \overline{n}_h(z)|\Phi_\varepsilon(y)dy \\ &\leqslant \left(\int_{\mathbb{R}^3} |\overline{n}_h(z-y) - \overline{n}_h(z)|^2\Phi_\varepsilon(y)dy\right)^{1/2},\end{aligned}$$

thanks to the Cauchy-Schwarz inequality and the value of the integral of $\Phi_\varepsilon(y)$. Consequently

$$\begin{aligned}\|\Phi_\varepsilon * \overline{n}_h - \overline{n}_h\|_{L^2((0,T)\times\mathbb{R}^3)} &\leqslant \int_0^T\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |\overline{n}_h(z-y) - \overline{n}_h(z)|^2\Phi_\varepsilon(y)dydzdt \\ &= \int_{B(0,\varepsilon)} \Phi_\varepsilon(y)\|\pi_{-y}\overline{n}_h - \overline{n}_h\|^2_{L^2((0,T)\times\mathbb{R}^3)}dy,\end{aligned}$$

and due to Proposition (3.6), we obtain the strong convergence of $\Phi_\varepsilon * \overline{n}_h$ to $\overline{n}_h$ in $L^2((0,T)\times\mathbb{R}^3)$, uniformly in h.

Step 2: We prove that for every fixed ε and any compact $\omega \subset D$, the sequence $\Phi_\varepsilon * \overline{n}_h$ is uniformly bounded in $H^1((0,T)\times\omega)$.
Thanks to Young's inequality, for a fixed ε,

$$\|\nabla(\Phi_\varepsilon * \overline{n}_h)\|_{L^2(D_T)} \leqslant \|\nabla\Phi_\varepsilon\|_{L^2(D)}\|\overline{n}_h\|_{L^1(D_T)} \leqslant C_\varepsilon\overline{\rho}T.$$

Furthermore, for any compact $\omega \subset D$, $\psi \in \mathscr{D}((0,T)\times\omega)$ and ε small enough, we have $\Phi_\varepsilon * \psi \in \mathscr{D}(D_T)$ and consequently, using Proposition (3.7) and Young's inequality for the convolution product, we get

$$\begin{aligned}\left|\int_0^T \langle \frac{d}{dt}\Phi_\varepsilon * \overline{n}_h, \psi\rangle dt\right| &\leqslant \int_0^T \left|\langle\frac{d\overline{n}_h}{dt}, \Phi_\varepsilon * \psi\rangle\right| dt \\ &\leqslant C\sum_{|k|\leqslant 3}\|\partial_z^k\Phi_\varepsilon * \psi\|_{L^2(D_T)} \\ &\leqslant C_\varepsilon\|\psi\|_{L^1(D_T)}.\end{aligned}$$

Then, by duality, for a fixed ε, $\frac{d}{dt}\Phi_\varepsilon * \overline{n}_h$ is uniformly bounded in $L^\infty((0,T)\times\omega)$ and we conclude that for each fixed ε, $\Phi_\varepsilon * \overline{n}_h$ is uniformly bounded in $H^1((0,T)\times\omega)$.

Step 3: We claim that, for all compacts $\omega \subset D$, the sequence $\overline{n}_h$ is a Cauchy sequence in $L^2((0,T)\times\omega)$. Write

$$\overline{n}_{h_1} - \overline{n}_{h_2} = (\overline{n}_{h_1} - \Phi_\varepsilon * \overline{n}_{h_1}) + (\Phi_\varepsilon * \overline{n}_{h_1} - \Phi_\varepsilon * \overline{n}_{h_2}) + (\Phi_\varepsilon * \overline{n}_{h_2} - \overline{n}_{h_2}).$$

Thanks to *Step 1*, for all $\eta > 0$, we can fix ε in such a way that

$$\|\overline{n}_{h_1} - \Phi_\varepsilon * \overline{n}_{h_1}\|_{L^2((0,T)\times\omega)} < \eta/3,$$
$$\|\Phi_\varepsilon * \overline{n}_{h_2} - \overline{n}_{h_2}\|_{L^2((0,T)\times\omega)} < \eta/3.$$

Thanks to *Step 2* and Rellich-Kondrachov's theorem, we know that for a fixed ε, $\Phi_\varepsilon * \overline{n}_h$ is a Cauchy sequence in $L^2((0,T)\times\omega)$, hence, for h_1 and h_2 close enough

$$\|\Phi_\varepsilon * \overline{n}_{h_1} - \Phi_\varepsilon * \overline{n}_{h_2}\|_{L^2((0,T)\times\omega)} < \eta/3$$

Consequently, for all positive $\eta > 0$, there exists h_1 and h_2 such that

$$\|\overline{n}_{h_1} - \overline{n}_{h_2}\|_{L^2((0,T)\times\omega)} < \eta,$$

which proves that $\overline{n}_h$ is a Cauchy sequence.

Step 4: The sequence $\overline{n}_h$ is a Cauchy sequence in $L^2(D_T)$.
Fix a natural number m, define $\omega_m := \{z \in D : d(z,\partial D) \geqslant 1/m\}$, $\omega_m^T := (0,T)\times\omega_m$ and write

$$\begin{aligned}\|\overline{n}_{h_1} - \overline{n}_{h_2}\|_{L^2(D_T)} &= \|\overline{n}_{h_1} - \overline{n}_{h_2}\|_{L^2(\omega_m^T)} + \|\overline{n}_{h_1} - \overline{n}_{h_2}\|_{L^2((0,T)\times(D\setminus\omega_m))}\\ &\leqslant \|\overline{n}_{h_1} - \overline{n}_{h_2}\|_{L^2(\omega_m^T)} + 2\sup_h \|n_h\|_{L^2((0,T)\times(D\setminus\omega_m))}\end{aligned} \tag{3.26}$$

Taking $q > 1$ and using Hölder's inequality we have

$$\begin{aligned}\|n_h\|^2_{L^2((0,T)\times(D\setminus\omega_m))} &= \int_0^T \int_{D\setminus\omega_m} n_h^2 dz dt\\ &\leqslant \int_0^T \|(n_h)^q\|_{L^2(D)} dt |D\setminus\omega_m|^{1/q^*}.\end{aligned}$$

Hence, for $q = p/2$, thanks to (3.25), we can conclude that $\sup_h \|n_h\|_{L^2((0,T)\times(D\setminus\omega_m))}$ goes to 0 when m goes to infinity. Consequently, for all $\eta > 0$, we can fix m big enough so that the second term in (3.26) is smaller than $\eta/2$ and then, thanks to *Step 3*, we can choose h_1 and h_2 in such a way that the first term is also smaller than $\eta/2$, which proves that $\{n_h\}_h$ is a Cauchy sequence in $L^2(D_T)$.

Step 5: Let n be the limit of a suitable subsequence of n_h and $v := (v^1, v^2, v^3)$ the weak limit of v_h (such limit exists due to the fact that v_h is bounded in $(L^2(D_T))^3$). We claim that $n \in L^2(0,T;\mathscr{V})$ and $v = \nabla n$.
For all $\phi(t,z) \in \mathscr{C}([0,T];\mathscr{D}(D))$, the convergence of n_h to n implies that

$$\int_0^T \int_D n(t,z)\partial_i\phi(t,z) dz dt = \lim_{h\to 0} \int_0^T \int_D n_h(t,z)\partial_i\phi(t,z) dz dt, \tag{3.27}$$

where $\partial_1\phi(z) = \partial_x\phi(z)$, $\partial_2\phi(z) = \partial_y\phi(z)$ and $\partial_3\phi(z) = \partial_\theta\phi(z)$. Writing $\partial_i\phi(z) = \nabla\cdot(\phi(t,z)\mathbf{e_i})$, where $\{\mathbf{e}_i\}$ is the eucledian canonical basis, we have that

$$\begin{aligned}\int_D n_h(t,z)\partial_i\phi(t,z)dz &= \sum_{j=1}^N \nu_j(t)\int_{D_j}\nabla\cdot(\phi(t,z)\mathbf{e_i})dz\\ &= \sum_{j=1}^N \nu_j(t)\sum_{l\in N_j}\int_{\Gamma_{jl}}\phi(t,z)\mathbf{e_i}\cdot\mathbf{n_{jl}}dS\\ &= \sum_{j=1}^N\sum_{l\in N_j}\nu_j(t)\int_{\Gamma_{jl}}\phi(t,z)n_{jl}^i dS.\end{aligned}$$

Therefore, we obtain

$$\int_D n_h(t,z)\nabla\phi(t,z)dz = \sum_{j=1}^N\sum_{l\in N_j}\nu_j(t)\int_{\Gamma_{jl}}\phi(t,z)\mathbf{n_{jl}}dS.$$

Recalling that $\mathbf{z_{jl}}$ is the center of mass of Γ_{jl}, we may estimate the integral over Γ_{jl} as

$$\int_{\Gamma_{jl}}\phi(t,z)\mathbf{n_{jl}}dS = \phi(t,\mathbf{z_{jl}})\mathbf{n_{jl}} + \mathcal{O}(|\Gamma_{jl}|),$$

as done in (3.12) with the integral over D_j. Using the fact that $|\Gamma_{jl}| = h^2$ and the amount of cells on the mesh is $\mathcal{O}(h^{-3})$, we have

$$\int_D n_h(t,z)\nabla\phi(t,z)dz = \sum_{j=1}^N\sum_{l\in N_j}\nu_j(t)|\Gamma_{jl}|\phi(t,\mathbf{z_{jl}})\mathbf{n_{jl}} + \mathcal{O}(h). \tag{3.28}$$

Reordering the terms of this last sum, we have

$$\begin{aligned}\sum_{j=1}^N\sum_{l\in N_j}\nu_j(t)|\Gamma_{jl}|\phi(t,\mathbf{z_{jl}})\mathbf{n_{jl}} &= \frac{1}{2}\sum_{j=1}^N\sum_{l\in N_j}\Big(\nu_j(t)|\Gamma_{jl}|\phi(t,\mathbf{z_{jl}})\mathbf{n_{jl}} + \nu_l(t)|\Gamma_{lj}|\phi(t,\mathbf{z_{lj}})\mathbf{n_{lj}}\Big)\\ &= -\frac{1}{2}\sum_{j=1}^N\sum_{l\in N_j}|\Gamma_{jl}|\mathbf{n_{jl}}\Big(\nu_l(t)-\nu_j(t)\Big)\phi(t,\mathbf{z_{jl}})\\ &= -\sum_{j=1}^N\sum_{l\in N_j}\frac{|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}(\mathbf{z_{jl}}-\mathbf{z_j})\frac{\Big(\nu_l(t)-\nu_j(t)\Big)}{s_{jl}}\phi(t,\mathbf{z_{jl}})s_{jl}\\ &= -\int_D v_h(z)\phi_h(t,z)dz, \end{aligned}\tag{3.29}$$

where

$$\phi_h(t,z) = \sum_{j=1}^N\sum_{l\in N_j}\phi(t,\mathbf{z_{jl}})\mathbb{1}_{D_{jl}}(z).$$

This function strongly converges in $L^2(D_T)$ to ϕ. Substituting (3.29) in (3.28) and then in (3.27), and using the weak convergence of v_h to v, we obtain that, for all $\phi \in \mathscr{C}([0,T];\mathscr{D}(D))$

$$\int_0^T \int_D n(t,z)\partial_i \phi(t,z)dzdt = -\int_0^T \int_D v^i(t,z)\phi(t,z)dzdt.$$

Taking $\phi(t,z) = \varphi(z)\chi(t)$, with $\varphi \in \mathscr{D}(D)$ and $\chi(t) \in \mathscr{C}([0,T])$, this last equality becomes

$$\int_0^T \int_D n(t,z)\partial_i \varphi(z)dz\chi(t)dt = -\int_0^T \int_D v^i(t,z)\varphi(z)dz\chi(t)dt,$$

for all $\varphi \in \mathscr{D}(D)$ and $\chi(t) \in \mathscr{C}([0,T])$, which implies that for each $\varphi \in \mathscr{D}(D)$

$$\int_D n(t,z)\partial_i \varphi(z)dz = -\int_D v^i(t,z)\varphi(z)dz, \text{ a.e. } [0,T].$$

The separability of $\mathscr{D}(D)$ finally implies that

$$\int_D n(t,z)\partial_i \varphi(z)dz = -\int_D v^i(t,z)\varphi(z)dz, \ \forall \varphi \in \mathscr{D}(D), \text{ a.e. } [0,T].$$

As $v(t,z) \in L^2(D)$ for almost all $t \in [0,T]$, $n(t,z)$ belongs to $\mathscr{V}$ for almost all $t \in [0,T]$, which proves the statements of the Proposition. □

An immediate consequence of Proposition (3.8), together with estimate (3.19) is that

$$\|n\|_H^2 + a\int_0^T \|v\|_H^2 \leqslant e^{bT}\|n_0\|_H^2, \text{ for all } T > 0. \tag{3.30}$$

Noticing that

$$\begin{aligned}
\int_0^T (\rho_h(t) - \rho(t))^2 dt &= \int_0^T \left(\int_{D_h} n_h(t,z)dz - \int_D n(t,z)dz \right)^2 dt \\
&= \int_0^T \left(\int_{D_h \setminus D} n_h(t,z)dz + \int_D (n_h(t,z) - n(t,z)dz \right)^2 dt \\
&\leqslant \int_0^T (|D_h \setminus D|^{1/2}\|n_h\|_{H_h} + |D|^{1/2}\|n_h - n\|_{L^2(D)})^2 dt \\
&\leqslant 2(|D_h \setminus D|\|n_h\|^2_{L^2((0,T)\times D_h} + |D|\|n_h - n\|^2_{L^2(D_T)}),
\end{aligned}$$

the definition of D_h and Proposition (3.8) implies that the sequence of functions $\rho_h(t)$ strongly converges to $\rho(t) := \int_D n(t,z)dz$ in $L^2((0,T))$.

3.3.4 Existence of weak solution

This section is devoted to prove that the function n is a weak solution of problem (3.1)-(3.2).

Proposition 3.9. *The function n satisfies*

$$-\langle n_0, \varphi(0)\rangle = \int_0^T \langle Q[n], \varphi\rangle + \langle \partial_t \varphi(t), n\rangle dt, \tag{3.31}$$

for all $\varphi \in \mathscr{C}_c^1([0,T), \mathscr{V})$.

Proof. First consider $\varphi \in \mathscr{C}_c^1([0,T), \mathscr{C}_c^\infty(\mathbb{R}^3))$, and for each j, multiply equation (3.16) by $h^3\varphi_j(t) := h^3\varphi(t, \mathbf{z_j})$, and add them up for all j, obtaining the relation

$$\sum_{j=1}^N h^3 \nu_j'(t)\varphi_j(t) = A_\varphi(t) + D_\varphi(t) + R_\varphi(t),$$

where

$$A_\varphi(t) = -\sum_{j=1}^N \varphi_j(t) \sum_{l\in N_j} |\Gamma_{jl}| \Big(\nu_j(t) u_{jl}^+(t) + \nu_l(t) u_{jl}^-(t)\Big),$$

$$D_\varphi(t) = \sum_{j=1}^N \varphi_j(t) \sum_{l\in N_j} \frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l} - \mathbf{z_j}|}\Big(\nu_l(t) - \nu_j(t)\Big),$$

$$R_\varphi(t) = \sum_{j=1}^N h^3 \Big(r_j - d_j \sum_l h^3 \nu_l(t)\Big)\nu_j(t)\varphi_j(t).$$

Reordering the terms from $A_\varphi(t)$, we get

$$\begin{aligned}
A_\varphi(t) &= \frac{1}{2}\sum_{j=1}^N \sum_{l\in N_j} |\Gamma_{jl}| \Big(\varphi_l(t) - \varphi_j(t)\Big)\Big(u_{jl}^+(t)\nu_j(t) + u_{jl}^-(t)\nu_l(t)\Big) \\
&= \frac{1}{2}\sum_{j=1}^N \sum_{l\in N_j} |\Gamma_{jl}| \Big(\varphi_l(t) - \varphi_j(t)\Big)\Big(u_{jl}(t)\nu_j(t) + u_{jl}^-(t)(\nu_l(t) - \nu_j(t))\Big) \\
&= \frac{1}{2}\sum_{j=1}^N \sum_{l\in N_j} |\Gamma_{jl}| \Big(\varphi_l(t) - \varphi_j(t)\Big) u_{jl}(t)\nu_j(t) + A_\varphi^1(t),
\end{aligned} \tag{3.32}$$

where

$$A_\varphi^1(t) = \frac{1}{2}\sum_{j=1}^N \sum_{l\in N_j} u_{jl}^-(t)|\Gamma_{jl}| \Big(\varphi_l(t) - \varphi_j(t)\Big)\Big(\nu_l(t) - \nu_j(t)\Big).$$

Thanks to the boundedness of u_{jl} and the regularity of φ, this term satisfies

$$\begin{aligned}
|A_\varphi^1(t)| &\leqslant \frac{\overline{V}}{2}\|\nabla\varphi\|_{L^\infty(\mathbb{R}^3)} h \sum_{j=1}^N \sum_{l\in N_j} |\Gamma_{jl}||\nu_l(t) - \nu_j(t)| \\
&= \frac{h}{6}\overline{V}\|\nabla\varphi\|_{L^\infty(\mathbb{R}^3)}\|v_h\|_{L^1(D_h)} \\
&\leqslant \frac{h}{6}|D_h|^{1/2}\overline{V}\|\nabla\varphi\|_{L^\infty(\mathbb{R}^3)}\|v_h\|_{L^2(D_h)}.
\end{aligned} \tag{3.33}$$

Recalling the definition of u_{jl} and the property (3.8), we get from (3.32)

$$A_\varphi(t) = \frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\frac{\Big(\varphi_l(t)-\varphi_j(t)\Big)}{|\mathbf{z_l}-\mathbf{z_j}|}V_{jl}(t)\cdot(\mathbf{z_l}-\mathbf{z_j})\nu_j(t) + A^1_\varphi(t),$$

$$= \sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\frac{\Big(\varphi_l(t)-\varphi_j(t)\Big)}{|\mathbf{z_l}-\mathbf{z_j}|}V_{jl}(t)\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\nu_j(t) + A^1_\varphi(t),$$

Defining $V_j(t) := V(t,\mathbf{z_j})$ allows us to write

$$A_\varphi(t) = \sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\frac{\Big(\varphi_l(t)-\varphi_j(t)\Big)}{|\mathbf{z_l}-\mathbf{z_j}|}V_j(t)\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\nu_j(t) + A^1_\varphi + A^2_\varphi(t),$$

with

$$A^2_\varphi(t) = \sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\frac{\Big(\varphi_l(t)-\varphi_j(t)\Big)}{|\mathbf{z_l}-\mathbf{z_j}|}(V_{jl}(t)-V_j(t))\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\nu_j(t).$$

Thanks to the regularity of φ and V, we have

$$|A^2_\varphi(t)| \leqslant \frac{h}{4}\|\nabla\varphi\|_{L^\infty(\mathbb{R}^3)}\|\nabla V\|_{L^\infty(\mathbb{R}^3)}\sum_{j=1}^{N}\sum_{l\in N_j}h^3\nu_j(t)$$

$$\leqslant \frac{h}{4}|D_h|^{1/2}\|\nabla\varphi\|_{L^\infty(\mathbb{R}^3)}\|\nabla V\|_{L^\infty(\mathbb{R}^3)}\|n_h\|_{L^2(D_h)}. \tag{3.34}$$

Finally, we write

$$A_\varphi(t) = \sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nabla\varphi(\mathbf{z_j})\cdot\mathbf{n_{jl}}\Big)V_j(t)\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\nu_j(t)$$
$$+ A^1_\varphi(t) + A^2_\varphi(t) + A^3_\varphi(t)$$
$$:= A^0_\varphi(t) + A^1_\varphi(t) + A^2_\varphi(t) + A^3_\varphi(t),$$

where

$$A^3_\varphi(t) = \sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\left(\frac{\Big(\varphi_l(t)-\varphi_j(t)\Big)}{|\mathbf{z_l}-\mathbf{z_j}|} - \nabla\varphi(\mathbf{z_j})\cdot\mathbf{n_{jl}}\right)V_j(t)\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\nu_j(t).$$

Again, the regularity of φ and the boundedness of V imply

$$|A^3_\varphi| \leqslant \frac{h}{4}\overline{V}|D_h|^{1/2}\|\nabla^2\varphi\|_{L^\infty(\mathbb{R}^3)}\|n_h\|_{L^2(D_h)}.$$

And this together with (3.33), (3.34) and estimate (3.19) ensures that there exists a constant C_A independent of h such that

$$\int_0^T |A^1_\varphi(t) + A^2_\varphi(t) + A^3_\varphi(t)|dt \leqslant C_A h.$$

On the other hand, we can rewrite the remaining term as

$$
\begin{aligned}
A^0_\varphi(t) &= \sum_{j=1}^{N}\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nabla\varphi(\mathbf{z_j})\cdot\mathbf{n_{jl}}\Big)V_j(t)\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\nu_j(t)\\
&= \sum_{j=1}^{N}\nu_j(t)V_j(t)\cdot\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nabla\varphi(\mathbf{z_j})\cdot\mathbf{n_{jl}}\Big)(\mathbf{z_{jl}}-\mathbf{z_j}).
\end{aligned}
$$

In [118], it was proven that

$$
\sum_{l\in N_j}|\Gamma_{jl}|\Big(\nabla\varphi(\mathbf{z_j})\cdot\mathbf{n_{jl}}\Big)(\mathbf{z_{jl}}-\mathbf{z_j}) = h^3\nabla\varphi(\mathbf{z_j}),
$$

consequently

$$
A^0_\varphi(t) = \sum_{j=1}^{N}\nu_j(t)V_j(t)\cdot\nabla\varphi(\mathbf{z_j})h^3 = \int_D n_h(t)\Big(V\cdot\nabla\varphi\Big)_h(t),
$$

where

$$
\Big(V\cdot\nabla\varphi\Big)_h(t) := \sum_{j=1}^{N}V_j(t)\cdot\nabla\varphi(\mathbf{z_j})\mathbb{1}_{D_j}(z).
$$

Moreover, the sequence $\Big(V\cdot\nabla\varphi\Big)_h(t)$ strongly converges to $V\cdot\nabla\varphi$ in $L^2(D_T)$, which implies that

$$
\int_0^T A_\varphi(t)dt \longrightarrow \int_0^T\int_D nV\cdot\nabla\varphi dzdt, \text{ for all } \varphi\in\mathscr{C}^1_c([0,T),\mathscr{C}^\infty_c(\mathbb{R}^3)). \tag{3.35}
$$

Reordering as well the terms from D_φ, we get

$$
\begin{aligned}
D_\varphi(t) &= -\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_l(t)-\nu_j(t)\Big)\Big(\varphi_l(t)-\varphi_j(t)\Big)\\
&= -\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_l(t)-\nu_j(t)\Big)\Big(\nabla\varphi(\mathbf{z_{jl}})\cdot(\mathbf{z_l}-\mathbf{z_j})\Big)+D^1_\varphi(t)\\
&= -\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_l(t)-\nu_j(t)\Big)\Big(\nabla\varphi(\mathbf{z_{jl}})\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\Big)+D^1_\varphi(t)
\end{aligned}
$$

where

$$
D^1_\varphi(t) = -\frac{1}{2}\sum_{j=1}^{N}\sum_{l\in N_j}\frac{A_{jl}|\Gamma_{jl}|}{|\mathbf{z_l}-\mathbf{z_j}|}\Big(\nu_l(t)-\nu_j(t)\Big)\Big(\varphi_l(t)-\varphi_j(t)-\nabla\varphi(\mathbf{z_{jl}})\cdot(\mathbf{z_l}-\mathbf{z_j})\Big).
$$

Thanks to the boundedness of the coefficients A_{jl} and the regularity of φ we get

$$
|D^1_\varphi| \leqslant h|D_h|^{1/2}\frac{\overline{\alpha}}{4}\|\nabla^2\varphi\|_{L^\infty(\mathbb{R}^3)}\|v_h\|_{L^2(D_h)}. \tag{3.36}
$$

Recalling the definition of A_{jl}, the fact that $A(\theta)$ is a diagonal matrix and the normal vectors to the boundary of D_j are elements of the canonical euclidean basis or their opposites, we have that, for all j and l

$$A_{jl}\Big(\nabla\varphi(\mathbf{z_{jl}})\cdot(\mathbf{z_{jl}}-\mathbf{z_j})\Big)=\Big(A(\theta_{jl})\nabla\varphi(\mathbf{z_{jl}})\Big)\cdot(\mathbf{z_{jl}}-\mathbf{z_j}).$$

Consequently, we infer that

$$D_\varphi(t)=-\int\limits_D v_h\cdot\Big(A(\theta)\nabla\varphi(z)\Big)_h dz+D_\varphi^1(t),$$

where

$$\Big(A(\theta)\nabla\varphi(z)\Big)_h(t):=\sum_{j=1}^N\sum_{l\in N_j}A(\theta_{jl})\nabla\varphi(\mathbf{z_{jl}})\mathbb{1}_{D_{jl}}(z)$$

strongly converges to $A(\theta)\nabla\varphi(z)$ in $\Big(L^2(D_T)\Big)^3$. The regularity of φ, the boundedness of v_h in $L^2(D_T)$ and (3.36) guarantee the existence of C_D such that

$$\int\limits_0^T|D_\varphi^1(t)|dt\leqslant C_Dh,$$

so that, consequently,

$$\int\limits_0^T D(t)dt\longrightarrow\int\limits_0^T\int\limits_D A(\theta)\nabla n\nabla\varphi dzdt,\ \text{for all } \varphi\in\mathscr{C}_c^1([0,T),\mathscr{C}_c^\infty(\mathbb{R}^3)). \tag{3.37}$$

The sequence of functions

$$R_h(t,z)=\sum_{j=1}^N\Big(r_j-d_j\rho_h(t)\Big)\mathbb{1}_{D_j}(z),$$

belongs to $L^\infty(D_T)$ and strongly converges in $L^2(D_T)$ to $r(z)-d(z)\rho(t)$, which implies that

$$\int_0^T R_\varphi(t)dt\longrightarrow\int\limits_0^T\int\limits_D\Big(r(z)-d(z)\rho(t)\Big)n\varphi dt,\ \text{for all } \varphi\in\mathscr{C}_c^1([0,T),\mathscr{C}_c^\infty(\mathbb{R}^3)). \tag{3.38}$$

Finally, we conclude that

$$\begin{aligned}\int\limits_0^T\sum_{j=1}^N h^3\nu_j'(t)\varphi_j(t)dt&=\sum_{j=1}^N h^3\int\limits_0^T\nu_j'(t)\varphi_j(t)dt\\&=-\sum_{j=1}^N h^3\left(\nu_j(0)\varphi_j(0)+\int\limits_0^T\nu_j(t)\varphi_j'(t)dt\right),\end{aligned}$$

which converges to

$$-\int\limits_D n_0\varphi(0)dz-\int\limits_0^T\int\limits_D n(t,z)\partial_t\varphi(t,z)dzdt.$$

Putting this result together with (3.35), (3.37) and (3.38), we get that, for all $\varphi \in \mathscr{C}_c^1([0,T),\mathscr{C}_c^\infty(\mathbb{R}^3))$

$$-\langle n_0,\varphi(0)\rangle = \int_0^T \langle Q[n],\varphi\rangle + \langle\partial_t\varphi(t),n\rangle dt. \tag{3.39}$$

As $\mathscr{C}_c^1([0,T),\mathscr{C}_c^\infty(\mathbb{R}^3))$ is dense in $\mathscr{C}_c^1([0,T),H^1(\omega))$ for all compacts $\omega\subset D$, and the functions n and ∇n belong to $L^2(D_T)$, (3.39) also holds for all $\varphi\in\mathscr{C}_c^1([0,T),\mathscr{V})$. □

Proposition 3.10. *The function n belongs to $H^1((0,T);\mathscr{V}')$.*

Proof: Taking $\varphi := \chi(t)\psi(z)$ with $\chi\in\mathscr{D}((0,T))$ and $\psi\in\mathscr{V}$ in equation (3.31), we get

$$\Big\langle \int_0^T n\chi' dt,\psi\Big\rangle = \int_0^T \langle\psi\chi',n\rangle dt = \int_0^T\langle\partial_t\varphi,n\rangle dt$$
$$= -\int_0^T\langle Q[n],\varphi\rangle dt = -\langle\int_0^T Q[n]\chi dt,\psi\rangle.$$

As this holds true for all $\psi\in\mathscr{V}$, this equation is equivalent with

$$\int_0^T n\chi' dt = -\int_0^T Q[n]\chi dt \text{ in } \mathscr{V}' \text{ for any } \chi\in\mathscr{D}((0,T)).$$

or in other words

$$\partial_t n = Q[n] \text{ in the sense of distributions in } \mathscr{V}'.$$

The estimate (3.30) implies that $Q[n]\in L^2((0,T);\mathscr{V}')$, consequently, n belongs to $H^1((0,T);\mathscr{V}')$. □

Proposition 3.11. *The function n belongs to $\mathscr{C}([0,T],L^2(D))$.*

Proof. From Proposition (3.6) we have that $n\in L^2((0,T),\mathscr{V})$. Define $\overline{n}(t,z) = n(t,z)\mathbb{1}_{[0,T]}(t)$ and the approximation to the identity sequence

$$\Phi_\varepsilon(t) := \varepsilon^{-1}\Phi(\varepsilon^{-1}t),$$

where $\Phi(t)$ is a mollifier with compact support included in $(-1,-1/2)$. The sequence $n_\varepsilon(t) := n *_t \Phi_\varepsilon$ belongs to $\mathscr{C}^1(\mathbb{R},\mathscr{V})$, $n_\varepsilon\to n$ a.e. on $[0,T]$ and in $L^2((0,T),\mathscr{V})$. For a fixed $\tau\in(0,T)$ and for any $t\in(0,\tau)$ and any $0<\varepsilon<T-\tau$, we have $s\to\Phi_\varepsilon(t-s)\in\mathscr{D}(0,T)$, since $\mathrm{supp}\Phi_\varepsilon(t-\cdot)\subset[t+\varepsilon/2,t+\varepsilon]\subset[\varepsilon/2,\tau+\varepsilon]$. Therefore, we get

$$n'_\varepsilon = \int_{\mathbb{R}} \partial_t\Phi_\varepsilon(t-s)\overline{n}(s)ds$$
$$= -\int_0^T \partial_s\Phi_\varepsilon(t-s)n(s)ds = \int_0^T \Phi_\varepsilon(t-s)n'(s)ds = \Phi_\varepsilon *_t \overline{n'}.$$

As a consequence $n'_\varepsilon \to n'$ a.e. and in $L^2((0,\tau),\mathscr{V}')$. Now fix $\tau \in (0,T)$ and $\varepsilon, \varepsilon' \in (0, T-\tau)$, and compute

$$\frac{d}{dt}\|n_\varepsilon(t) - n_{\varepsilon'}(t)\|_H^2 = 2\langle n'_\varepsilon - n'_{\varepsilon'}, n_\varepsilon - n_{\varepsilon'}\rangle_{\mathscr{V}'\times\mathscr{V}},$$

so that for any $t_1, t_2 \in [0,\tau]$

$$\|n_\varepsilon(t_2) - n_{\varepsilon'}(t_1)\|_H^2 = \|n_\varepsilon(t_1) - n_{\varepsilon'}(t_1)\|_H^2 + 2\int_{t_1}^{t_2} \langle n'_\varepsilon - n'_{\varepsilon'}, n_\varepsilon - n_{\varepsilon'}\rangle dt.$$

Since $n_\varepsilon \to n$ a.e. in $[0,\tau]$ in V, fix t_1 such that $n_\varepsilon(t_1) \to n(t_1)$ in V, so as a consequence of $n_\varepsilon \to n$ in $L^2((0,T),\mathscr{V})$ and $n'_\varepsilon \to n'$ in $L^2((0,\tau),\mathscr{V}')$ we have

$$\limsup_{\varepsilon,\varepsilon'\to 0} \sup_{[0,\tau]} \|n_\varepsilon - n_{\varepsilon'}\|_H^2 \leqslant \lim_{\varepsilon,\varepsilon'\to 0} \int_0^\tau \|n'_\varepsilon - n'_{\varepsilon'}\|_{\mathscr{V}'}\|n_\varepsilon - n_{\varepsilon'}\|_{\mathscr{V}} ddt = 0.$$

So that n_ε is a Cauchy sequence in $\mathscr{C}([0,\tau], L^2(D))$, and then n_ε converges in $\mathscr{C}([0,\tau], L^2(D))$ to a limit $\tilde{n} \in \mathscr{C}([0,\tau], L^2(D))$. That proves $n = \tilde{n}$ a.e. and $n \in \mathscr{C}([0,\tau], L^2(D))$. By taking $\Phi(-t)$ as the mollifier function in the previous proof, and choosing $\tau \in (0,T)$ and $\varepsilon \in (0,\tau)$ can be proven that $n \in \mathscr{C}([\tau, T], L^2(D))$, and consequently $n \in \mathscr{C}([0,T], L^2(D))$. □

Proposition 3.12. *The function n is a weak solution of problem (3.1)-(3.2).*

Proof. Assume first $\varphi \in \mathscr{C}_c([0,T), H) \cap L^2((0,T),\mathscr{V}) \cap H^1([0,T],\mathscr{V}')$. We define $\varphi_\varepsilon(t) = \varphi *_t \Phi_\varepsilon$ for a mollifier Φ_ε with compact support included in $(0,\infty)$ so that $\varphi_\varepsilon \in \mathscr{C}^1_c([0,T);\mathscr{V})$ for any $\varepsilon > 0$ small enough and

$$\varphi_\varepsilon \to \varphi \text{ in } \mathscr{C}([0,T], H) \cap L^2((0,T),\mathscr{V}) \cap H^1([0,T],\mathscr{V}').$$

Writing the equation (3.31) for φ_ε and passing to the limit $\varepsilon \to 0$ we get that (3.31) also holds true for φ.
Assume now $\varphi \in X_T = C([0,T], H) \cap L^2((0,T),\mathscr{V}) \cap H^1([0,T],\mathscr{V}')$. Fix $\chi \in \mathscr{C}^1(\mathbb{R})$, such that $\operatorname{supp}\chi \subset (-\infty, 0)$, $\chi' \leqslant 0$, $\chi' \in \mathscr{C}_c((-1,0))$, and such that the integral of χ' is -1. For example, fix $\delta < 1/2$ and define

$$\chi(s) = \begin{cases} 1 & \text{if} \quad s \leqslant -1+\delta \\ \frac{1}{2}(1+\cos(\frac{\pi(t+1-\delta)}{1-2\delta})) & \text{if} \quad -1+\delta \leqslant s \leqslant -\delta \\ 0 & \text{if} \quad s \geqslant -\delta \end{cases}$$

Now define $\chi^t_\varepsilon = \chi(\frac{s-t}{\varepsilon})$, so that $\varphi_\varepsilon := \varphi\chi^t_\varepsilon \in \mathscr{C}_c([0,T);H)$ and $\chi^t_\varepsilon \to \mathbf{1}_{[0,t]}$, $(\chi^t_\varepsilon)' \to -\delta_t$ as $\varepsilon \to 0$. Equation (3.31) for the function φ_ε writes

$$-(n_0, \varphi(0)) - \int_0^T (n, \varphi(s))(\chi^t_\varepsilon)' ds = \int_0^T \chi^t_\varepsilon \Big(\langle Q[n](s), \varphi(s)\rangle + \langle \varphi'(s), n\rangle\Big) ds.$$

Passing to the limit when ε goes to 0 we obtain that n is a solution for the variational formulation. □

3.3.5 A discrete implicit scheme

Once we have established the proof of existence of a solution for problem (3.16)-(3.17), together with the obtention of a semi-discrete scheme in order to approximate such solution, we proceed to derive an implicit discrete scheme starting from problem (3.16)-(3.17), and to prove its convergence.
Consider a natural number K and define $\Delta t = \frac{T}{K}$ and $t_k = k\Delta t$, $\nu_j^k := \nu_j(t_k)$, $k = 1, \ldots, K$. Using a forward difference approximation for the time derivative in (3.16) we get the implicit scheme

$$\frac{\nu_j^{k+1} - \nu_j^k}{\Delta t} = M_j^{k+1}\nu_j^{k+1} + \sum_{l\in N_j} B_{jl}^{k+1}\nu_l^{k+1}, \tag{3.40}$$

where

$$M_j^{k+1} := M_j(t_{k+1}) = -\frac{|\Gamma_{jl}|}{h^3}\sum_{l\in N_j}\Big(u_{jl}^+(t_{k+1}) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|}\Big) + \Big(r_j - d_j\sum_{l=1}^{N} h^3\nu_l^{k+1}\Big),$$
$$B_{jl}^{k+1} := B_{jl}(t_{k+1}) = \frac{|\Gamma_{jl}|}{h^3}\Big(-u_{jl}^-(t_{k+1}) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|}\Big).$$

Theorem 3.3. *Let ν_j^0 be non-negative initial data with mass $\rho_0 = \sum_{j=1}^{N} h^3\nu_j^0$ and assume that*

$$\Delta t < \frac{1}{\left(\sqrt{r^+ + d^+\overline{\rho}} + \sqrt{d^+\overline{\rho}}\right)^2}, \tag{3.41}$$

then there exists a unique non-negative solution ν_j^k, $k = 1, \ldots, N$ to scheme (3.40). Furthermore, for each h, the sequence of piecewise constant functions

$$\nu_{\Delta t}^j(t) = \sum_{k=0}^{K}\nu_j^k\mathbb{1}_{(t_k,t_{k+1})},$$

strongly converges to the solution of (3.16)-(3.17) in $(L^2((0,T)))^N$ when Δt goes to 0.

Proof. For all Δt satisfying (3.41), there exists $\lambda < 1$ such that

$$\Delta t(r^+ + d^+\overline{\rho}_\lambda) < \lambda, \tag{3.42}$$

where $\overline{\rho}_\lambda = \frac{\overline{\rho}}{1-\lambda}$. Consider the set

$$\mathscr{X} = \{\eta \in \mathbb{R}^M : \eta_j \geqslant 0\,\forall j,\ \|\eta\|_1 = \sum_j h^3\eta_j \leqslant \overline{\rho}_\lambda\},$$

and assume $\nu^k \in \mathbb{R}^M$ to be the solution of (3.40) for a previous iteration, having all non-negative components and satisfying $\sum_j h^3\nu_j^n \leqslant \overline{\rho}$. Define the operator $\nu = F(\eta) : \mathscr{X} \to \mathbb{R}^M$ as the solution of the linear system

$$P(\eta)\nu = \nu^k, \tag{3.43}$$

where the components of matrix $M(\eta)$ are defined as

$$P_{jl}(\eta) = \begin{cases} -\Delta t B_{jl}^{n+1}, & \text{if } l \in N_j, \\ 1 + \Delta t\Big(\frac{|\Gamma_{jl}|}{h^3} \sum_{l \in N_j} \Big(u_{jl}^+(t_{n+1}) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|}\Big) - \Big(r_j - d_j\|\eta\|_1\Big)\Big) & \text{if } j = l. \end{cases}$$

From the definition of $\mathscr{X}$ and the choice of Δt we have that $P_{jl}(\eta)$ is positive for all η if $j = l$, and non-positive if $j \neq l$. Furthermore $P(\eta)$ is a column-dominant matrix, so that we may conclude that $P(\eta)$ is an M-matrix, which implies that its inverse exists and has only non-negative entries. As ν^k has all non-negative components, then $\nu = F(\eta)$ also has non-negative components for all $\eta \in \mathscr{X}$. Multiplying system (3.43) by h^3 and adding up all equations, thanks to (3.42) we obtain

$$\begin{aligned} \sum_{j=1}^N h^3 \nu_j =& \sum_{j=1}^N h^3 \nu_j^k + \Delta t \sum_{j=1}^N h^3 \Big(r_j - d_j\|\eta\|_1\Big)\nu_j \\ &\leqslant \overline{\rho} + \Delta t(r^+ + d^+\overline{\rho}_\lambda) \sum_{j=1}^N h^3 \nu_j \\ &< \overline{\rho} + \lambda \sum_{j=1}^N h^3 \nu_j. \end{aligned}$$

This implies that $\sum_{j=1}^N h^3 \nu_j < \overline{\rho}_\lambda$ and consequently, $F(\eta)$ is a continuous application going from $\mathscr{X}$ to itself, and thanks to Brouwer's fixed point theorem, $F(\eta)$ has at least a fixed point on $\mathscr{X}$. Furthermore, a fixed point of $F(\eta)$ will satisfy

$$\sum_{j=1}^N h^3 \nu_j = \sum_{j=1}^N h^3 \nu_j^k + \Delta t \sum_{j=1}^N h^3 \Big(r_j - d_j\|\nu\|_1\Big)\nu_j$$

which in turn implies $\|\nu\|_1 \leqslant \overline{\rho}$. As a consequence, the implicit Euler scheme satisfies

$$\|\nu^k\|_1 \leqslant \overline{\rho} \qquad \text{for all } k. \tag{3.44}$$

For the uniqueness, assume there exists two different solutions ν and μ to scheme (3.40). Let us denote the sign of $\beta \in \mathbb{R}$ as $sign(\beta)$. Taking the difference for each equation, multiplying by $h^3 sign(\nu_j - \mu_j)$, adding up all the equations and recalling that

$$\sum_{l \in N_j} B_{jl}^{n+1} = \frac{|\Gamma_{jl}|}{h^3} \sum_{l \in N_j} \Big(u_{jl}^+(t_{n+1}) + \frac{A_{jl}}{|\mathbf{z_l} - \mathbf{z_j}|}\Big),$$

we obtain that

$$\begin{aligned} \sum_j h^3 |\nu_j - \mu_j| =& \Delta t \sum_j h^3 \Big((r_j - d_j\|\nu\|_1)|\nu_j - \mu_j| \\ &+ d_j sign(\nu_j - \mu_j)(\|\nu\|_1 - \|\mu\|_1)\mu_j\Big) \\ &\leqslant \Delta t(r^+ + d^+\overline{\rho}) \sum_j h^3 |\nu_j - \mu_j|. \end{aligned}$$

The condition over Δt then implies that $\|\nu-\mu\|_1 = 0$ and consequently $\nu = \mu$.
To prove the convergence of $\nu^j_{\Delta t}$ to ν_j, we define the sequences of continuous functions

$$\mu^j_{\Delta t}(t) = \sum_{k=1}^{K} \frac{(t-t_k)\nu_j^{k+1} + (t_{k+1}-t)\nu_j^k}{\Delta t} \mathbb{1}_{(t_k,t_{k+1})}.$$

These sequences are in $\mathscr{C}([0,T])$ and are uniformly bounded because for all values of j and k, $\nu_j^k \leqslant \frac{\bar{\rho}}{h^3}$ due to (3.44). Furthermore, thanks to (3.40)

$$|\mu^j_{\Delta t} - \nu^j_{\Delta t}| \leqslant \max_k |\nu_j^{k+1} - \nu_j^k| = \Delta t \max_k |M_j^{k+1}\nu_j^{k+1} + \sum_{l\in N_j} B_{jl}^{k+1}\nu_l^{k+1}| \leqslant C\Delta t,$$

where C is independent of Δt. Consequently, for each j, when Δt goes to 0 both sequences $\nu^j_{\Delta t}$ and $\mu^j_{\Delta t}$ strongly converge in $L^2((0,T))$ to a certain continuous functions $\nu_j^*(t)$.

We consider now a function $\varphi \in \mathscr{C}^1_0((0,T))$ and we define $\varphi_k := \varphi(t_k)$. For all k, multiply (3.40) by $\Delta t \varphi_k$ and add over k in order to obtain

$$\sum_{k=1}^{K} (\nu_j^k - \nu_j^{k-1})\varphi_k = \sum_{k=1}^{K} \Delta t \Big(M_j^k \nu_j^k + \sum_{l\in N_j} B_{jl}^k \nu_l^k \Big)\varphi_k,$$

or, after reordering the sum on the left side

$$\sum_{k=1}^{K} \Delta t \nu_j^k \frac{\varphi_k - \varphi_{k-1}}{\Delta t} = \sum_{k=1}^{K} \Delta t \Big(M_j^k \nu_j^k + \sum_{l\in N_j} B_{jl}^k \nu_l^k \Big)\varphi_k. \tag{3.45}$$

For all φ in $\mathscr{C}^1_0((0,T))$, the sequence

$$\sum_{k=1}^{K} \frac{\varphi_k - \varphi_{k-1}}{\Delta t} \mathbb{1}_{(t_k,t_{k+1})},$$

strongly converges in $L^2((0,T))$ to φ'. The boundedness of the coefficients M_j^k and B_{jl}^k together with strong convergence of $\nu^j_{\Delta t}$ imply that

$$\sum_{k=1}^{K} \Big(M_j^k \nu_j^k + \sum_{l\in N_j} B_{jl}^k \nu_l^k \Big) \mathbb{1}_{(t_k,t_{k+1})},$$

strongly converges to $M_j(t,\rho_h(t))\nu_j^* + \sum\limits_{l\in N_j} B_{jl}(t)\nu_l^*$, so, taking the limit in (3.45), we get

$$\int_0^T \nu_j^* \varphi'(t) dt = \int_0^T \Big(M_j(t,\nu_j^*)\nu_j^* + \sum_{l\in N_j} B_{jl}(t)\nu_l^* \Big)\varphi(t) dt,$$

which implies that ν_j^* is in $\mathscr{C}^1$ and is a solution for (3.16). Furthermore, as ν_j^* is the point-wise limit of $\mu^j_{\Delta t}$, it also satisfies the initial conditions (3.17). □

3.4 Simulations

The first part of this section is devoted to the numerical analysis of the approximation error. For certain values of the coefficients of problem (3.1)-(3.3), it is possible to obtain an analytical solution, which we will use in order to compare with our numerical approximation.
Assume that $r(x, y, \theta)$ and $d(x, y, \theta)$ are constants, that $V(t, x, y, \theta)$ is independent of t and that exists $W(x, y, \theta)$ such that

$$V(z) = A(\theta)\nabla W(z).$$

Assume as well that

$$n_0(z) = C\frac{e^{W(z)}}{\int_D e^{W(z)}dz}.$$

Then, the solution for problem (3.1)-(3.3) is

$$n(t, z) = \frac{e^{W(z)}}{\int_D e^{W(z)}dz}\frac{re^{rt}}{d(K + e^{rt})},$$

where

$$K = \frac{r - Cd}{Cd}.$$

The existence of an analytic solution allows us to compare its values with those obtained from solving (3.16) for different values of h, and this way, numerically establish the error order of the method. Choosing

$$D := \{(x, y, \theta) \in [0, 1]^3 : x^2 + y^2 \leqslant 1\},$$

$$V(t, z) = \begin{pmatrix} -(\theta+1)x \\ -(\theta+1)y \\ 1 \end{pmatrix}, \ A(\theta) = \begin{pmatrix} \theta+1 & 0 & 0 \\ 0 & \theta+1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$r = d = 1$ and $n_0 = \Big(\pi(1 - e^{-1/2})(e - 1)\Big)^{-1} e^{-\frac{x^2}{2} - \frac{y^2}{2} + \theta}$, the exact solution for (3.1)-(3.3) is

$$n(t, z) = \Big(\frac{\pi}{2}(1 - e^{-1/2})(e - 1)\Big)^{-1} e^{-\frac{x^2}{2} - \frac{y^2}{2} + \theta}\frac{e^t}{1 + e^t}.$$

For a grid of points $\{(t_k, z_j)\}$ with $t_k = k\Delta t$, $k = 1, \dots, K$, $\Delta t > 0$ and $j = 1, \dots, N$, we define the discrete $L^2(D_T)$ error for the semi-discrete scheme (3.16) as

$$E^1(\Delta t, h) = \left(\sum_{k=1}^{K}\sum_{j=1}^{N}(n(t_k, z_j) - \nu_j(t_k))^2 h^3 \Delta t_k\right)^{1/2},$$

where $\nu(t)$ is the solution of the scheme for the functions introduced above. We set $\Delta t = 0.01$ and in Figure (3.1) we show the dependence in log-log scale of $E_h = E(0.01, h)$ with respect to the inverse of the cell size $M = 1/h$.

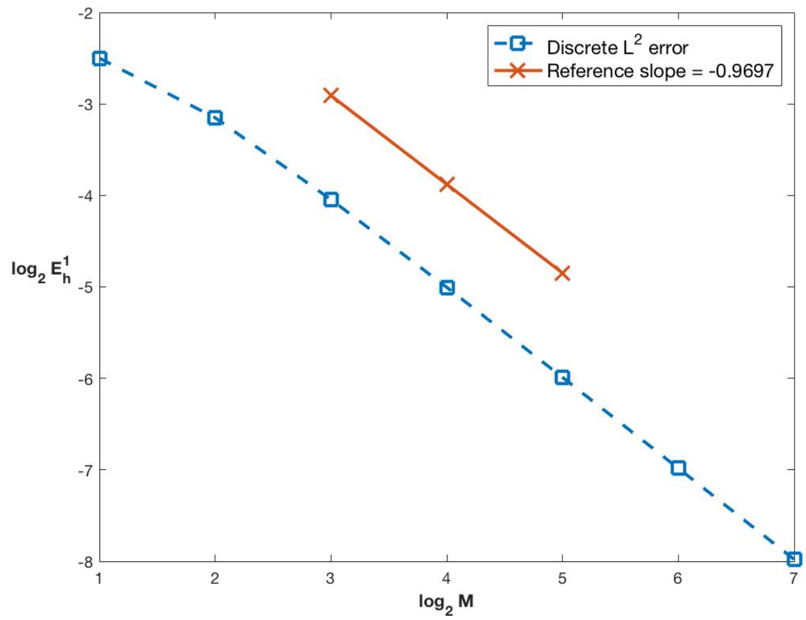


Figure 3.1: The discrete $L^2(D_T)$ error for the semi-discrete scheme, for $T = 10$ and M ranging between 2 and 128.

In the same way, we define the discrete $L^2(D_T)$ error for the discrete scheme (3.40) as

$$E^2(\Delta t, h) = \left(\sum_{k=1}^{K} \sum_{j=1}^{N} (n(t_k, z_j) - \nu_j^k)^2 h^3 \Delta t_k \right)^{1/2},$$

where ν_j^k is the solution of (3.40). We set $h = 1/50$ and plot the dependence of $E^2_{\Delta t} := E^2(\Delta t, 0.02)$ with respect to $\Delta t = 1/M_1$, again in log-log scale.

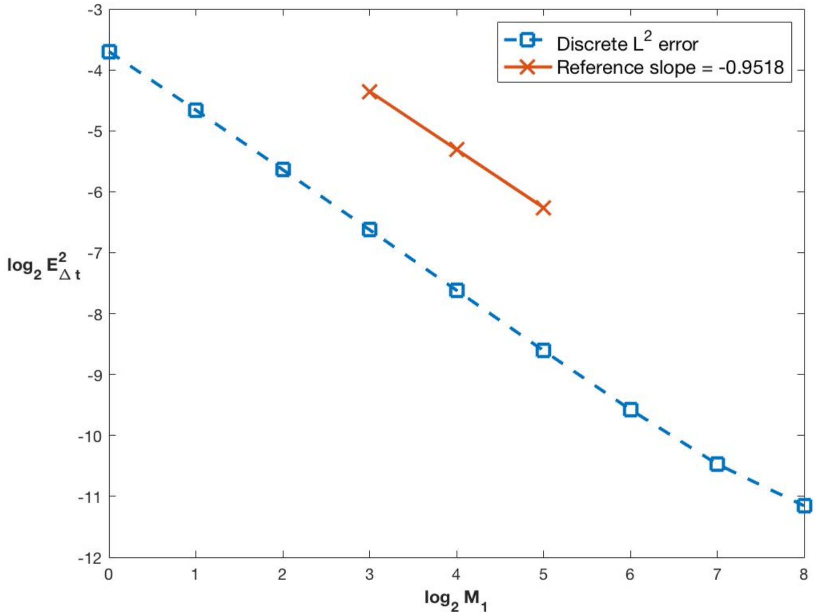


Figure 3.2: The discrete $L^2(D_T)$ error for the discrete scheme, for $T = 10$ and M_1 ranging between 2 and 256.

3.4.1 Phenotypic Dimorphism

We present now several examples illustrating the effect of the environment on a population and how the plasticity trait plays a role in surviving effects.

Monomorphic Population

On the first place, we show the evolution of a population which is under the effects of natural selection and non-genetic epimuations, but without considering plasticity as a trait nor accounting for the environmental pressure. Specifically, we solve the problem (3.1)-(3.2) over the domain

$$\Omega = \{(x, y) \in [0, 1]^2 : (x - 1)^2 + (y - 1)^2 > 1\}.$$

This choice of domain represents the existence of a trade-off between traits and it is evidenced here by noticing that the individuals of the population which are close to the maximal value of one of the traits ($x = 1$ or $y = 1$), must forcibly be close as well to the minimal value of the other trait ($y = 0$ or $x = 0$). We take the growth rate as $r(x, y) = e^{-(x-0.1)^2-(y-0.1)^2}$, the death rate as $d(x, y) = 0.5$, the diffusion parameters $a_{11}(\theta) = a_{22}(\theta) = 10^{-6}$, and the drift terms $V(t, x, y) = (0, 0)$. We take an initial condition (3.3) given by the expression

$$n_0(x, y) = a\mathbb{1}_{\{f(x,y)<1\}}e^{-\frac{1}{1-f(x,y)}},$$

with $f(x, y) = \frac{(x-0.25)^2+(y-0.25)^2}{(0.025)^2}$. We choose the value of a in such a way that $\rho_0 = \int_\Omega n_0(x, y)dxdy = 1$. With this choice of n_0 we intend to represent a "strongly" monomorphic population. This is, a population where most of the individuals are concentrated around a single set of traits.

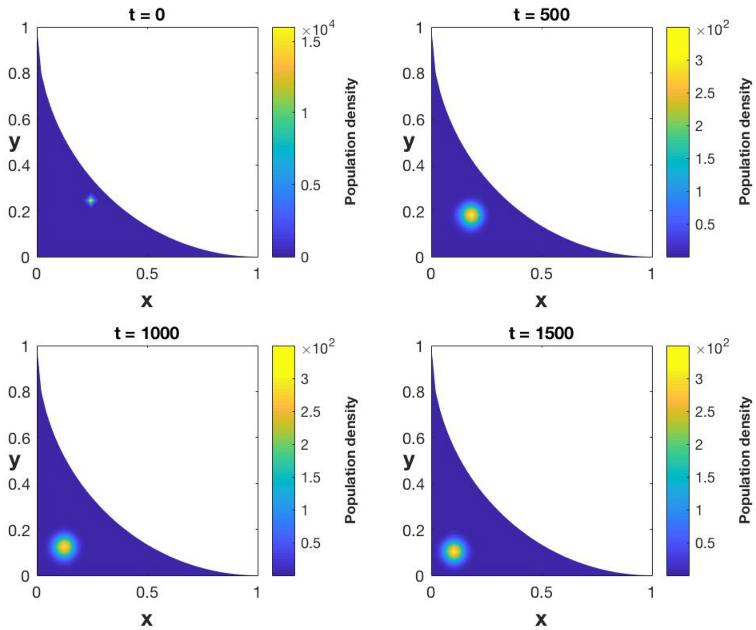


Figure 3.3: Evolution of a population only subjected to natural selection and non-genetic epimutations.

We observe in figure (3.3) that the dominant phenotype slowly converges to the point which maximises the fitness, in this case, the point $(0.1, 0.1)$, which maximises the growth rate $r(x, y)$.

Dimorphism due to the effect of the environment

For a second example we keep the same parameters, but add a drift term accounting for the effect of the environment (biologically, a “cellular stress”). Specifically we choose

$$V(t, x, y) = 10^{-3} \left(\mathbb{1}_{(y>x)} \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \mathbb{1}_{(y<x)} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right).$$

Notice that we still are not including plasticity as a trait in our analysis. Also notice that the function $V(t, x, y)$ is not continuous, while the results presented in previous sections needed $V(t, z)$ to be smooth in order to ensure the existence and uniqueness of solution for the problem, as well as the convergence of the finite volume method. This is not a big issue, because the conditions of our problem allow us to use a density argument in order to extend our results to any $V(t, z) \in \mathscr{C}([0, T), L^2(D))$, and in any case all numerical approximations are smoothing approximations of the drift V.
The choice of V can be seen (and experimentally replicated) as a certain type of “training”: all the individuals of the population are “pushed” in the direction where they show their largest potential.

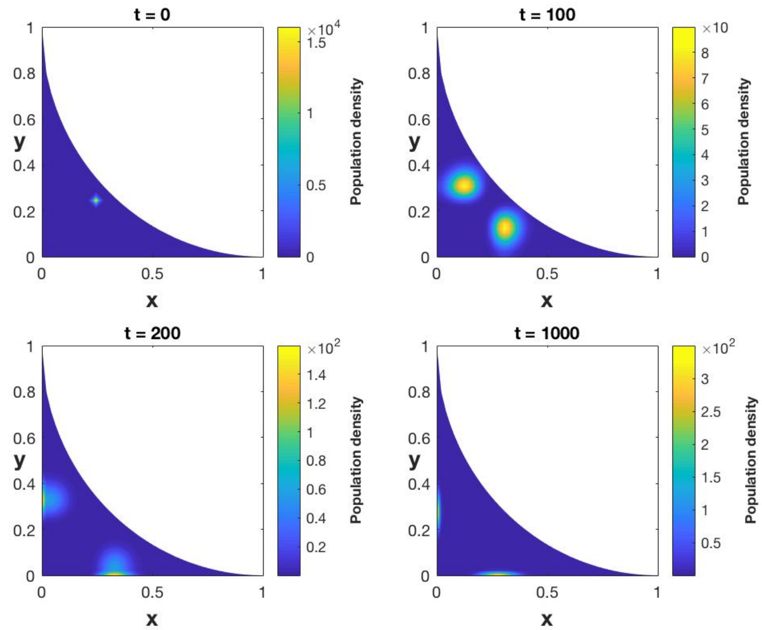


Figure 3.4: Evolution of a population under the effect of the environment.

We can appreciate in figure (3.4) that adding a drift term resulted in the appearance of dimorphism: the final population is concentrated around two different trait configurations. This evolution into dimorphism due the action of environment, contrasts with the results obtained in [23], where it was proved that dimorphism can occur in the absence of a drift term, given that the growth rate $r(x, y)$ has several maximum point satisfying certain conditions.

Plasticity, environmental effect and dimorphism

We will now consider plasticity as a trait and modify the parameters from the previous examples accordingly. We first consider the growth rate as

$$r(x, y, \theta) = e^{-(x-0.1)^2-(y-0.1)^2} + e^{-(z-0.8)^2},$$

and keep the constant death rate $d(x, y, \theta) = 0.5$. We consider the diffusion matrix

$$A(\theta) = \begin{pmatrix} (\theta+1)10^{-6} & 0 & 0 \\ 0 & (\theta+1)10^{-6} & 0 \\ 0 & 0 & 10^{-6} \end{pmatrix},$$

and the drift term

$$V(t, z) = 10^{-3}\theta \left(\mathbb{1}_{(y>x)} \begin{pmatrix} -1 \\ 1 \\ -x^2 - y^2 \end{pmatrix} + \mathbb{1}_{(y<x)} \begin{pmatrix} 1 \\ -1 \\ -x^2 - y^2 \end{pmatrix} \right).$$

This choice of V is similar to the one shown before, only that now differentiation imposes a cost on adaptability: the more specialised you are, the harder it gets to adapt to new situations. Notice that a higher plasticity increases the effect of non-genetic epimutations (given by the diffusion term) and stress induced mutations (given by the drift term). For the initial data (3.3) we take a function of the form

$$n_0(z) = a\left(\mathbb{1}_{\{f_1(z)<1\}}e^{-\frac{1}{1-f_1(z)}} + \mathbb{1}_{\{f_2(z)<1\}}e^{-\frac{1}{1-f_2(z)}}\right),$$

with $f_i(z) = \frac{\|z-z_i\|^2}{(0.025)^2}$, $i = 1, 2$, $z_1 = (0.25, 0.25, 0.25)$ and $z_1 = (0.25, 0.25, 0.75)$. The value of a is again chosen in a way that the total population size over $D = \Omega \times [0, 1]$ is equal 1.

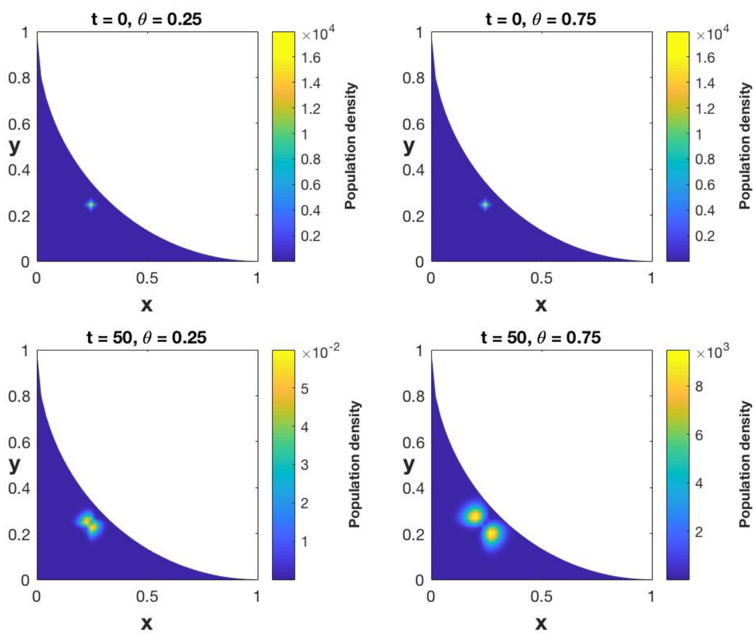


Figure 3.5: Evolution of two sub-populations with different levels of plasticity: Initial stages.

We observe in figure (3.5) that for this set of parameters, the sub-population with the lowest plasticity quickly gets extinct (notice the scale of the density values), while the emergence of dimorphism can be appreciated for the one with higher plasticity.

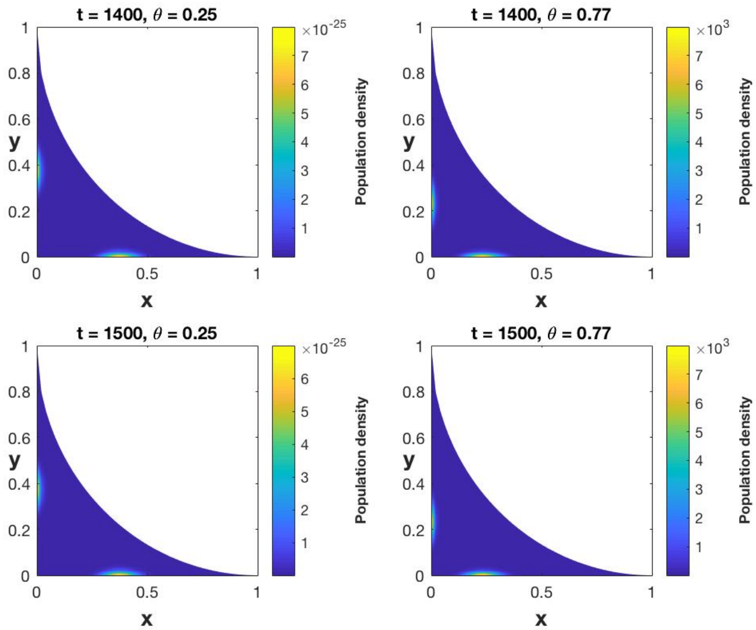


Figure 3.6: Evolution of two sub-populations with different levels of plasticity: Final stages.

This kind of behaviour persists up to the final stages of the evolution, when we can observe in Figure (3.6) that the low plasticity sub-population is completely extinct while the high plasticity one has completed the differentiation process.

Another example of emergence of dimorphism

We present now a different example of emergence of dimorphism, but this time, not as a result of a response to the effect of the environment, but as a consequence of the existence of two maximum points for the growth rate. We will observe how a population initially concentrated around a single phenotype configuration, will evolve with time into a dimorphic population, in which each upcoming sub-population is more specialised and less plastic that in the initial configuration. For this purpose, over the domain $D = \Omega \times [0, 1]$ we consider an initial density given by the expression

$$n_0(z) = a\mathbb{1}_{\{f(z)<1\}} e^{-\frac{1}{1-f(z)}},$$

with $f(z) = \frac{\|z-z_0\|^2}{(0.025)^2}$, where $z_0 = (0.25, 0.25, 0.5)$ and $\|\cdot\|$ is the euclidean norm. We choose the value of a in such a way that $\rho_0 = \int_D n_0(z) = 1$.
We set the growth rate and the death rate as

$$r(x, y, \theta) = \mathbb{1}_{\{y>x\}} e^{-(0.1-x)^2-(0.9-y)^2} + \mathbb{1}_{\{x\geqslant y\}} e^{-(0.1-y)^2-(0.9-x)^2},$$
$$d(x, y, \theta) = \frac{1}{2}.$$

We choose the diffusion matrix

$$A(\theta) = \begin{pmatrix} (\theta+1)10^{-6} & 0 & 0 \\ 0 & (\theta+1)10^{-6} & 0 \\ 0 & 0 & 10^{-6} \end{pmatrix},$$

and finally the drift term

$$V(t,z) = 10^{-3}\theta \begin{pmatrix} -y \\ -x \\ -(x+y) \end{pmatrix}.$$

This time, the "push" towards specialisation imposed by V is inversely proportional to the current set of traits (individuals with traits (x, y) are specialising with a rate proportional to $(-y, -x)$. The growth rate was chosen in such a way that it satisfies the sufficient conditions given in [23] in order to guarantee the appearance of phenotypic polymorphism.

We show in Figure (3.7) that initially the population is concentrated around the phenotype $z_0 = (0.25, 0.25, 0.5)$, and gradually starts differentiating while loosing plasticity.

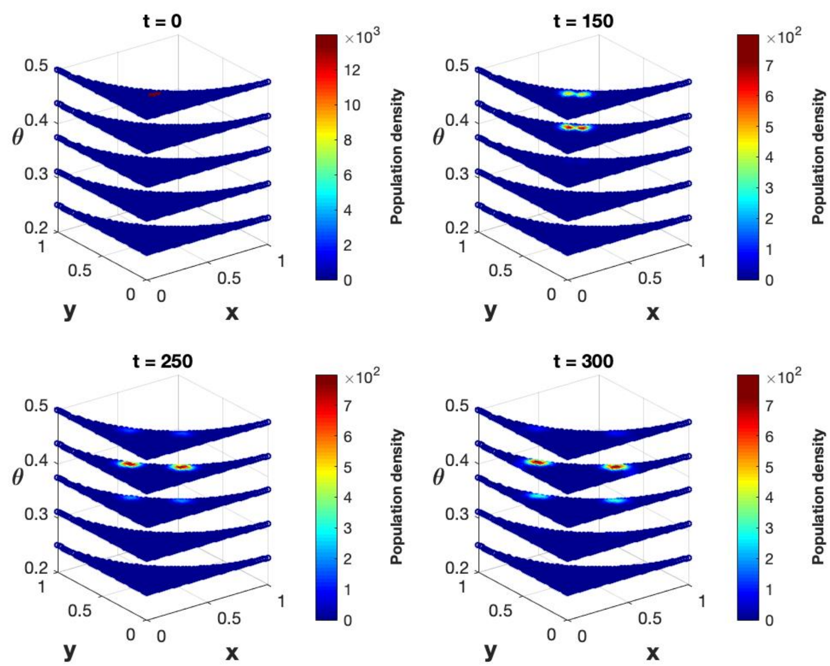


Figure 3.7: Initial stages of the population density for different values of θ: The differentiation process starts. At around $t = 250$ (bottom left) most of the population has already concentrated around the plasticity level $\theta = 0.4375$ and around $t = 300$ (bottom right) we observe that the migration towards a less plastic state continues.

As the two new sub-populations become more and more differentiated, the loss in plasticity becomes more evident, and we see in Figure (3.8) that most of the mass is migrating towards less plastic states, while the differentiation process continues.

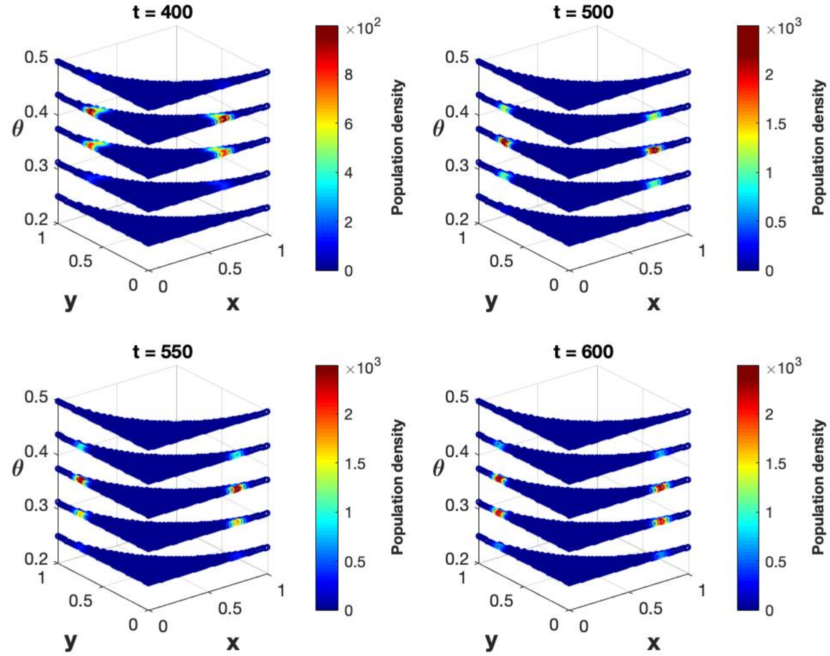


Figure 3.8: Intermediate stages of the population density for different values of θ: While the differentiation process continues, we observe further loss in plasticity. Around $t = 500$ (top right) most of the population has reached $\theta = 0.375$ and at subsequent times the migration continues.

Finally we observe in Figure (3.9) that once the sub-populations are fully specialised, the concentration process continues and at the final stage $t = 1000$ we have a dimorphic population which is more specialised but less plastic that the initial one.

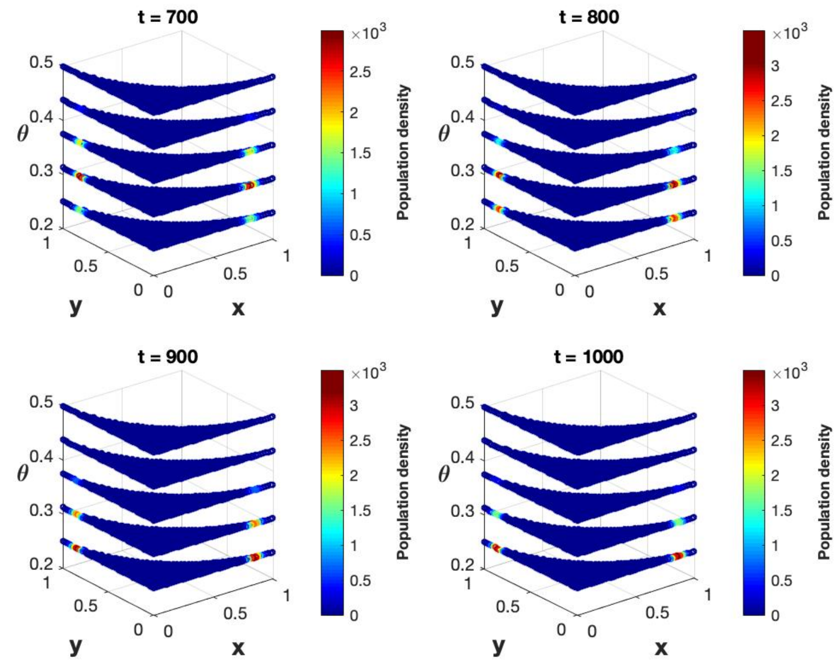


Figure 3.9: Final stages of the population density for different values of θ: Around $t = 900$ (bottom left) the differentiation process is over and most of the population has reached the plasticity level $\theta = 0.25$. At $t = 1000$ (bottom right) we observe that the population concentrated around any other level of plasticity is almost extinct, and only the one around $\theta = 0.25$ survives.

Overall, in the absence of a drift term, the effects of the growth and death rate, with small (or zero) diffusion has been widely studied, for example, see [22, 23, 28]. In general, the points where the population concentrates are entirely determined by the reaction part, while the diffusion coefficients determine how concentrated the population is.
The shape of the domain Ω directly affects where the fitness function (which depends on the birth and death rates) attains its maxima, therefore, affecting as well where the concentration phenomena occurs.
The choice of initial state does not appear to affect the final configuration, but rather the dynamics of the population during the evolution in time. We have shown that if we start with an already concentrated population, we witness the continuous "migration" of the the population towards the fittest state.
Finally, if we consider the presence of the advection term, we observe that it may affect not only the amount of selected sets of traits, but their positions and the dynamics of convergence towards them as well.

3.5 Concluding remarks

The validity of the model we constructed is strengthened by the different evolutionary mechanisms described in [68], where the authors focus their attention on Stress-Induced Evolutionary Innovation, and compare it to plasticity-based models, in particular the Plasticity First Hypothesis. Quoting the authors: "SIEI and PFH are not competing models but explain different kinds of evolutionary processes that are sometimes distinct and sometimes combined over evolutionary time". Similar mechanisms were taken into consideration in the construction of our model, environmental stress (aka environmental pressure, biologically, at the single-cell level, "cellular stress") in the form of an advection term and and mutations thanks to plasticity in the form of a diffusion term, both accompanied by natural selection in the form of reaction term.

It is important to highlight the novelty that represents the inclusion of plasticity as a trait, which has not been considered before (with the exception of [25, 26] and the subsequent works on the cane toad spreading rate, where a similar structural variable is used, but to denote spatial diffusivity, and not adaptability potential, as in our case). Another novelty is the inclusion of constraints between traits modelled through a certain relation between the structural variables. Most of the previous work, and in particular [111], have considered the variable space as the entire unit square, or all of $\mathbb{R}^d$, $d \in \mathbb{N}$, disregarding the existence of constraints between different traits, and the possible effects on the dynamics of the population this might have.

The finite volume method offers a powerful tool in order to numerically approximate the solution of integro-differential or reaction-diffusion equations, such as the one treated in the present chapter. The preservation of the structure of the original problem at a semi-discrete level and the excellent approximation for the non-local terms are just two of the reasons why we chose this method. This way, we were able to obtain two numerical schemes in order to approximate the solution for an evolution problem modelling bet hedging strategies.

We proved the existence and uniqueness of solutions for such schemes, and constructed sequences of functions converging to the solution of the original problem. We approximated the convergence error by establishing a comparison with an exact solution. It is worth mentioning that the constructive character of the proofs may provide new and interesting tools in order to obtain further theoretical results.

After simulating various situations, we observed different ways in which a population can respond to external stress, depending on the plasticity levels of its individuals. A highly plastic sub-population can quickly adapt to its surrounding environment, guaranteeing this way its survival, while a less plastic sub-population might go extinct under the same external factor. Another strategy consists in "trading" some of the plasticity by a higher differentiation level.

Furthermore, the emergence of dimorphism as a consequence of external stress, not only is an interesting alternative to the previously established results from [23], but also shows that bet-hedging strategies are a suitable response to (abrupt) external changes in the environment, and, at the same time, a possible way to survive them. It is fair mentioning that, throughout all the simulations, the symmetry hypothesis required in the reference [23] in order to observe dimorphism were respected. It remains to establish what are the essential conditions that will lead to the appearance of dimorphism when an advection term is present.

We thus provided here a rigorous model for the study of the emergence of dimorphism, an event that is likely to have been at the evolutionary origin of multicellularity by divergence of phenotypes and may thus provide a rationale for a renewed conception of animal evolution towards multicellular organisms, and, more pragmatically and consistently with the atavistic theory of cancer, for a possible origin of phenotype bet hedging in cancer cell populations.

Bet hedging in cancer cell populations is indeed a strategy susceptible to yield maximal probabilities of survival to a plastic cell population exposed to life-threatening insults such as by drugs or other deadly therapies. The modelling setting presented here may thus help in the future to test and optimise combined anticancer therapies involving chemotherapies, targeted therapies, and - what is likely still ahead of us for the present time - possible control of cell plasticity by epigenetic drugs.

Chapter 4

Phenotype divergence and cooperation in isogenic multicellularity and in cancer

4.1 Biological and evolutionary-developmental background

4.1.1 Being or not teleological: the two settings considered

Although this may seem completely trivial to state, let us emphasise that for us there is no such thing as teleology, i.e., orientation in a given direction or towards a given goal, in the general evolution of multicellular animals, which is constituted of a succession of haphazard strategic choices of adaptation to changing environments in existing evolutionary units, at one stage of evolution towards an identified next one. Such adaptations, often resulting in branchings of clades, as solutions to existential problems, imposed by external constraints as stresses [119, 120] induced by changes in the environment, are by no means unique, admitting that evolution proceeds by trials and errors, and by tinkering [121] from available material to solve such problems. We proposed in Chapter 3 a mathematical scheme to model the phenotypic divergence that may be a basis for such environmental stress-induced evolutionary steps.

Conversely, teleology is of course present in the embryonic development of multicellular animals, which, according to Haeckel's formula "Ontogeny recapitulates phylogeny" [122, 123], follows in each species the evolutionary choices made at each branching step of the evolution of species, leading from the fecundated egg (most frequent form of elementary material evolutionary unit in multicellular animals [124], those who are subject to cancer [125, 126]) to adult animals with their completely differentiated cell types, following the *body plan* [127, 128] characteristic of the species. From this holistic point of view, evolution of species is nothing but evolution of the body plan, evolution of genes and of gene regulatory networks being completely dependent upon this master regulator. We suggest here that understanding the cooperation principles that have been optimised (noting that an optimisation problem may have diverse solutions) at each developmental step may benefit from a close look at the mechanisms of the evolutionary steps that have determined the species body plan, and we sketch mathematical ways to achieve this task.

One of the main difficulties in understanding and representing the design of the body plan is how to introduce mechanisms of coherence (for signals) and cohesion (for tissues) that make a multicellular organism stable and functional, with compatibility and cooperation between tissues and organs, and we are aware of the fact that such complete understanding still lies ahead of us. However, localised absence of coherence between tissues of an organism by lack of control on differentiations is precisely the main characteristic of cancer, the second and in our opinion resulting from the first one, being absence of control on proliferation [129]. We propose that evolution of cooperation between cells, that has been identified in tumours [130–132], is a reactivation of mechanisms present in the body plan that are still present, although chaotic, uncontrolled and doomed to fail at the level of the organism, in tumour cells, may rely on elementary evolutionary mechanisms that have been designed in the evolutionary past of their body plan, so that this point should be better understood to efficiently represent cooperation in tumours.

4.1.2 The atavistic theory of cancer

Recently popularised by physicists Paul Davies and Charles Lineweaver, together with oncologist Mark Vincent, the atavistic theory of cancer [62, 133–136], had in fact been envisioned already in 1996 by oncologist Lucien Israel [137], and likely as early as 1914 by biologist Theodor Boveri [138], although none of these scientists seem to have been initially aware of the works of their predecessors. It helps us understand tumour progression and intratumoral organisation from a long-term evolutionary viewpoint. Briefly, it relies on the ideas that 1) all cancer cells are multicellular animal cells, results of a billion year-old evolution from unicellular organisms, and as such keep in their genomes powerful remnants of the organismic defence and construction mechanisms borne in their body plans (even if this term is not used by Davies and Lineweaver, they only mention their genomes); 2) tumours are results of a regression in the development of the organism, corresponding to early, incoherent versions of *"an ancient genetic toolkit of pre-programmed behaviors"*, which we may freely identify as an unachieved evolutionary version of the species body plan, and which they name "Metazoa 1.0". The atavistic theory thus clearly states that a tumour is not just the result of some aberrant stochastic mutation in somatic cells (the somatic mutation theory, SMT, recently reviewed and compared to the atavistic theory in [134]), but that it rather follows predictable paths in such regression towards a poorly organised, incoherent population of cells, nevertheless constituted of animal cells that are highly plastic (and thus resistant to external therapeutic pressure by anticancer drugs), as they have the power to differentiate and de-differentiate, and also to loosely cooperate between them in tumours. The works of David Goode and colleagues [139–142] have evidenced in cancer samples silencing of genes of multicellularity and compatibility between expression of genes of multicellularity and of unicellularity, resulting in escaping organismic control on cell differentiation (in other words, developing cell plasticity) and on proliferation, tending to a widely autonomic behaviour which is a characteristic of cells in tumour tissues.

The atavistic theory of cancer is little by little, as more evidence in the study of ancient genes becomes known and published [139–142], gaining recognition among theorists of cancer biology, however still quite limited in the field of oncology, where people question its amenability to produce innovations in the therapeutics of cancer. Innovating theories may take a long time to reverse the argument of "authority of tradition" [143]. The present situation may remind us, *mutatis mutandis*, of the way geographers received in 1912 with much skepticism Alfred Wegener's theory of continental drift [144], until it was completely justified fifty years later by the theory of plate tectonics and progressively admitted by all geophysicists. A limitation to a wider acceptance of the atavistic theory is the present lack of sufficient evidence susceptible to convince biologists and philosophers of cancer, who prefer to keep on the "safe" side of science under development and, at least temporarily, reject it as not sufficiently relying on facts. Indeed, when it is mentioned in recent texts of philosophy of science - by authors who nevertheless must be commended for at least mentioning it -, the atavistic theory of cancer is not always correctly summed up, sometimes even presented in an off-hand way with arguments against it that show but partial understanding, as in [145]. A mere hypothesis, really? At least a uniting one in understanding cancer, fully compatible with the holistic point of view on evolution that we have mentioned above.

4.1.3 Why and how does multicellularity fail in cancer?

Cancer is thus, taking the atavistic theory of cancer for granted - although it tells us nothing about the very origin of the disease -, the progressive result of a failed maintenance of the teleological (or teleonomical, if one wants to explicitly exclude any intentionality, which is our position) construction of an animal. It may be described as essentially "a deunification of the individual" [146]. In the perspective of evolved multicellularity, it is tempting to describe - an epistemological position we assume - such material construction at the level of genes and gene regulatory networks, initially not from the zygote, but from nonclonal colonies of cells (i.e., before the invention of the egg [124] and of the *body plan* contained in it) in three successive steps.

At the first step, the colony level, exist only genes of the cell division cycle and cell death, likely by quorum sensing. At the second step are introduced genes coding for transcription factors and (unregulated) differentiation. At the third step appear genes coding for epigenetic regulations, the top level of fine local regulations, that are themselves subject to central regulations in higher-level animals such as bilaterians. Such hierarchy is remarkably found, in a reverse order, in the evolution in malignancy found in fresh blood samples of patients with acute myelogenous leukaemia [147], which induces us to propose a scenario for cancer progression as relying firstly on epigenetic gene alterations (which includes differentiation control), secondly on alterations in differentiation, and only very late on alterations in cell cycle regulations, which are the strongest basis of proliferation. Unfortunately so far, with the remarkable and recent exception of the successes of immunotherapy, cancer therapies target mainly this strength [133].

4.1.4 A narrative of long-term evolution and cancer, freely exposed to the fire of philosophy of science

We need not justify any given evolutionary path that led to such and such animal, and rather see paths followed in evolution as diverse evolutionary strategies adapted to external constraints that imposed changes on the behaviour of the actors of the evolutionary paths at stake. Let us mention here that we hold, from our point of view, which resorts to functional, physiological and anatomical evolution, these actors, or evolutionary units, to be the *body plans* [127, 128] of multicellular animals, and not the individual genes, nor the gene regulatory networks that are mere effectors of evolutionary strategies, not determinants, and are only secondarily affected by them, as reflected in observations. A paleoanthropological analogy in evolution, *mutatis mutandis*, of such strategies at the level of divergence from a common ancestor in the Hominin lineage between Paranthropus and early Homo, relying on different dietary choices, may be found in [148]. Such haphazard strategical choices in long-term, Darwinian, evolution, that have become fixed in the body plan of animal species by genetic mutations and success in species fitness, may fail in cancer, as described in the previous section.

These firstly non determined (tinkered [121]) strategies led to epigenetic modifications (aka epimutations), later to fixed mutations of the genes coding for the epigenetic enzymes that determine these epigenetically defined strategies yielding functional body plans, that are the bases of physiology and anatomy construction in multicellular animals. Cancer cannot change the body plan of an animal in that of another animal, and it is certainly not a new form of life. However, by loss of organismic control on differentiations, it can reverse a cohesive body plan in a given species to some intermediate, poorly defined, unachieved form of the body plan of this species, yielding a collection of still very plastic cells, in other words a tumour, or a Metazoan 1.0 in the words of the atavistic theory of cancer [62]. The causes of such loss of control on differentiations are unknown, and the atavistic theory tells us nothing about them. However they may consist of an abrupt change in the environmental pressure on the tissue at stake, but also may be identified as due to a mutation in the genes responsible for epigenetic control [147].

4.2 Cell differentiation and phenotype divergence

4.2.1 Heterogeneity and plasticity with respect to what?

Cell populations, healthy and cancer, are heterogeneous w.r.t. various *continuous* traits under study, that are used to describe their biological variability, such as cell size, age in the cell division cycle, expression of genes of drug resistance, or more functional and abstract traits determining cell population fate such as viability, fecundity, motility, plasticity, according to the biological question at stake. Plasticity [149, 150] in a given trait is its capacity to change under the pressure of external constraints, such as drugs, and it has long been recognised as as relying on epigenetic factors [151]. Plasticity may be considered as a speed of evolution from one trait distribution to another one when the surrounding environment of the cell population changes, slowly or abruptly. Such evolution may be accelerated in equations by terms of advection (especially when abrupt changes in the environment force the cell population to adapt quickly) and diffusion (representing uncertainty in phenotype determination).

Differentiation in cell lineages, such as the ones constituting the paths of haematopoiesis, may consist either of simple maturation, following the same line towards a terminally differentiated cell type, such as the different granulocytes (neutrophils, eosinophils, and basophils) among white blood cells, or of branching, e.g., in haematopoiesis from pluripotent haematopoietic stem cells to myeloid versus lymphoid progenitors. Phenotype divergence is the biological phenomenon by which branching occurs between precursors of terminally differentiated cell types. The first identified phenomenon relying on phenotype divergence in evolution from unicellularity towards multicellularity was likely the separation between germinal cells (the germen) and germen-supporting somatic cells (the soma), proposed in 1892 by August Weismann [152] and later mentioned by John Maynard Smith and Eörs Szathmáry as the first step from unicellularity towards multicellularity, one of the major transitions in evolution [153]. Basis of heterogeneity in cell populations within a cohesive multicellular individual, or within a tumour, phenotype divergence necessarily relies on phenotype plasticity, and it is the phenomenon we here tackle to represent in phenotype-structured equations.

4.2.2 Long-term evolution as genetic adaptation of the body plan in animals

As mentioned in the introduction, we consider that the fundamental evolutionary unit in the great Darwinian evolution of animals is the body plan [127, 128], which is virtually (as it is abstract, indeed as a plan, self-developing, written as a self-extracting archive in genetic code, its dynamic extraction occurring continuously during the process of animal development) present in every physiologically complete nucleated animal cell, starting from the zygote, i.e., the initial fecundated egg. The genes and gene regulatory networks that materially proceed from it and serve to design and cohesively maintain the construction of the animal when it is achieved, are its observable materialisation.

Anatomically in 3D observations, physiologically by the observation of the great functions of the organism, and genetically by investigation the genes that have been identified (e.g., by KO experiments) in different species to correspond to anatomic structures and physiological functions, and their expression, we may have access to material reflections of the body plan, and thus partially reconstitute its evolution across species. This is precisely what has been investigated about the genes at the origin of multicellularity and their correspondence with the genes that are altered in cancer by Domazet-Lošo and Tautz [154,155], and later by Trigos *et al.* [139–142] in David Goode's team, giving rise and genetic arguments to the atavistic theory of cancer [62,133,134].

4.2.3 A nonlocal phenotype-structured cell population model

The reaction-diffusion-advection model proposed in Chapter 3 to exemplify *bet hedging* as a 'tumour strategy' to diversify its phenotypes in response to deadly stress (e.g., by cytotoxic drugs), but also to represent phenotypic divergence in evolution towards multicellularity, runs as follows.

Let $D = \Omega \times [0,1]$, where $\Omega := \{C(x,y) \leqslant K\}$ (a constraint between competing traits x and y) and $\theta \in [0,1]$. The evolution with time t of a plastic cell population of density $n(z,t)$ structured in a 3D phenotype $z = (x, y, \theta)$, where x=viability, y=fecundity, θ=plasticity, with $r(z)$ and $d(z)$ growth and death rates, is given by

$$\partial_t n + \nabla \cdot \Big(Vn - A(\theta)\nabla n \Big) = (r(z) - d(z)\rho(t))n, \tag{4.1}$$

with $\Big(Vn - A(\theta)\nabla n\Big)\cdot \mathbf{n} = 0$ for all $z \in \partial D$ ($\mathbf{n}$ is a normal vector to ∂D), $n(0,z) = n_0(z)$ for all $z \in D$, where $\Omega = \{(x,y) \in [0,1]^2 : (x-1)^2 + (y-1)^2 > 1\}$, and the diffusion matrix is

$A(\theta) = \begin{pmatrix} a_{11}(\theta) & 0 & 0 \\ 0 & a_{22}(\theta) & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$, with a_{11} and a_{22} non-decreasing functions of θ, influencing the speed at which non-genetic epimutations occur, otherwise said, it is a representation of how the internal plasticity trait θ affects the non-genetic instability of traits x and y, by tuning the diffusion term $\nabla \cdot \{A(\theta)\nabla n\}$; the advection term

$$\nabla \cdot \{V(t,z)n\} = \nabla \cdot \{(V_1(t,z), V_2(t,z), V_3(t,z))n\}$$

represents the cellular stress exerted on the population by external evolutionary pressure, i.e., by changes in the cell population environment, here chosen as tearing apart the cell population between competing traits x (viability) and y (fecundity); and $\rho(t) = \int_D n(t,z)dz$ stands for the total mass of individuals in the cell population at time t.

The existence and uniqueness of solutions is obtained in finite time in a constructive way by using the compactness of a sequence of numerical solutions, which are the result of the algorithms used to discretise the model. Simulations may be obtained with instances of the functions used in the equations. For instance, to obtain phenotypic divergence (which we take as the basis of both bet hedging in cancer and of emergence of multicellularity in evolution), we consider over the domain $D = \Omega \times [0,1]$ an initial density given by

$$n_0(z) = a\mathbb{1}_{\{f(z)<1\}} e^{-\frac{1}{1-f(z)}},$$

with $f(z) = \frac{\|z-z_0\|^2}{(0.025)^2}$, where $z_0 = (0.25, 0.25, 0.5)$ and $\|\cdot\|$ is the euclidean norm. We choose the value of a in such a way that $\rho_0 = \int_D n_0(z) = 1$.
We set the growth rate and the death rate as

$$r(x,y,\theta) = \mathbb{1}_{\{y>x\}} e^{-(0.1-x)^2-(0.9-y)^2} + \mathbb{1}_{\{x\geqslant y\}} e^{-(0.1-y)^2-(0.9-x)^2}, \quad d(x,y,\theta) = \frac{1}{2}.$$

We choose the diffusion matrix

$$A(\theta) = \begin{pmatrix} (\theta+1)10^{-6} & 0 & 0 \\ 0 & (\theta+1)10^{-6} & 0 \\ 0 & 0 & 10^{-6} \end{pmatrix},$$

and the advection term, tearing apart traits x and y, is chosen as $V(t,z) = 10^{-3}(-y, -x, -(x+y))$, or $10^{-3}\theta(-y, -x, -(x+y))$ if we want plasticity θ to impinge also on the advection term, representing in all cases the influence of the tumour ecosystem on the tumour cell population.

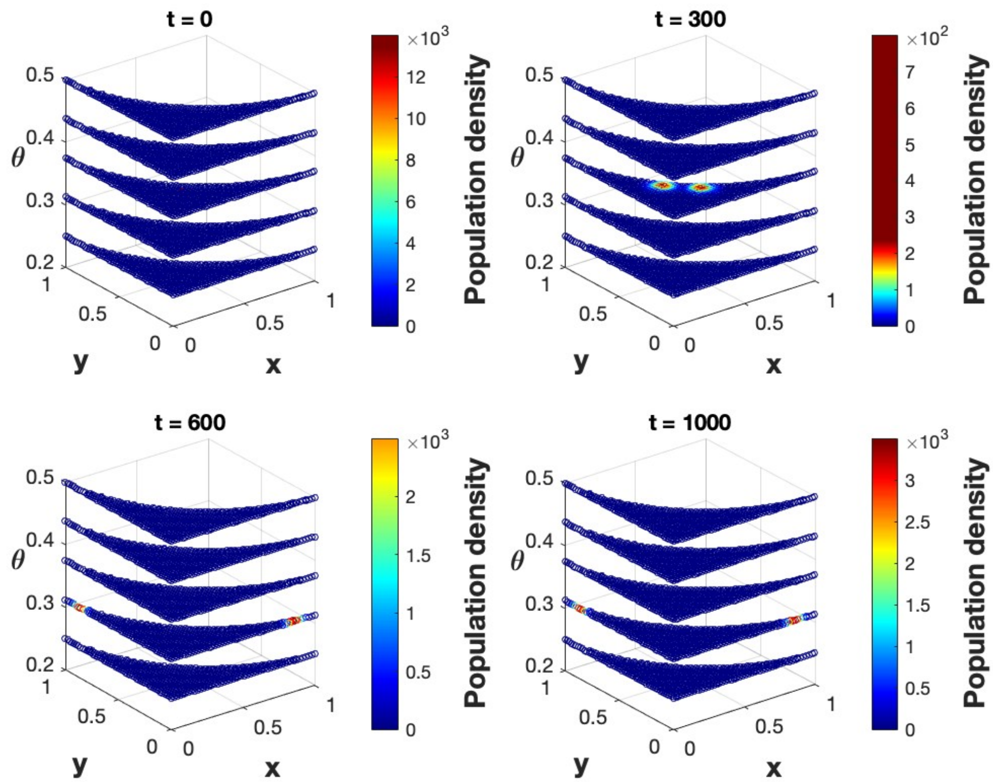


Figure 4.1: Phenotype divergence and loss of plasticity. On these cartoon-like figures, one can follow the progressive distancing of an initial cell population arbitrarily set at $z = (0.25, 0.25, 0.5)$, submitted to an advection gradient that tends to split the cell population into two subpopulations migrating towards the two extreme points $(0, 1)$ and $(1, 0)$ of the domain Ω, while the plasticity variable θ decreases towards 0.

The reader is sent to Chapter 3 for more detailed explanations and illustrations.

4.2.4 What this model tackles and what it leaves unexplained

Our reaction-diffusion-advection equations give the most important part in modelling phenotype divergence to the drift (advection) term representing environmental pressure from the ecosystem towards separation of phenotypes. Plasticity is naturally already present in the reaction term of this continuous phenotype-structured cell population model of adaptive dynamics, and the diffusion term adds to phenotype adaptability by uncertainty in its determination. Nevertheless, the sensitivity of phenotype adaptation and the trade-off we set between the supposed contradictory 1D phenotypes is mainly represented by the advection term and the bounded region within which the phenotypes evolve, that together represent constraints and offer possibilities of trade-offs between the phenotypes.

This model is clearly a mathematical abstraction that may be applied as such to every possible branching situation in the physiological development of multicellular animals or in bet hedging of phenotypes in tumours. For instance, one could model more precisely in glioblastoma cells such branching situations as the "go-or-grow" alternative between enhancing a proliferation potential (fecundity) and a motion potential (motility) [156], which would need to represent in the same kind of model the biological mechanisms that account for them, and about the constraints (likely of energetic nature) between them. This would help us design more precisely the advection term and the domain in phenotype space within which phenotypes evolve. It would imply efficient transdisciplinary collaboration on this subject between mathematicians and biologists of cancer, which we hope to develop in the future.

4.3 Cooperation

4.3.1 Tinkered cooperation in the emergence of multicellularity vs. directed cooperation in constituted multicellular animals

Noting that the question of cooperation and of division of labour has been considered by many authors at different stages of associations between individuals, including animal societies [153]. To follow again the metaphor of the separation in evolution between Paranthropus and early Homo, the situation with respect to phenotype divergence between body plans of animals is as if, *mutatis mutandis*, in evolution from their common hominin ancestor, Paranthropus and early Homo, after their genetic separation starting by fixation of initial epigenetic haphazard strategic adaptive choices (since evolution under changes in environmental pressure proceeds by tinkering [121]), had found interest in developing mutualistic interactions, living in symbiosis, less and less independently of one another. However, since the Paranthropus species eventually became extinct, likely due to climate changes incompatible with his too specialised vegetalian diet, whereas Homo survived, having adapted his diet to meat eating, this was actually not the case, or not in a permanent way, in the evolution of hominins.

We are aware of the fact that this metaphor is by no means perfect, and that reversible development, of epigenetic nature, within an isogenic individual (or a tumour) is not the same process as evolution of species, which is based on fixed, irreversible, genetic separations by branchings. Nevertheless, hypothesising that genetic specialisation is likely to begin with reversible epigenetic phenotype divergence before being fixed by gene mutations, we hope that it sheds some light on the processes that are at work in elementary steps in the evolution towards multicellularity and in bet hedging in tumours.

Cooperation between populations of cells resulting from such phenotype divergence may be considered as the glue that holds together all cell subpopulations in an isogenic multicellular organism. It may occur when mutualistic interactions are beneficial for all the interacting cell populations, provided that none of them becomes extinct. And it may also not occur, in which case no trace of such missed mutualism is found in the evolution of body plans. It is indeed, in our representation, the body plan that has kept memory, in each species, in constitutive intercellular gene regulatory networks, of the proper strategic choices w.r.t. phenotype divergences that lead to the design of an anatomically and physiologically cohesive animal. No tinkering is present anymore in these programmed choices designed in the body plan, and this is what we would like to represent now.

We will present two different possible approaches to the study of evolution of cooperation. The first one takes the prisoner's dilemma as a starting point, and considers reciprocity as a factor influencing the strategies of both players. The possible outcomes for a long running game are studied, and finally, a way to model a scenario with n players is described. The second modelling choice is through an integro-differential system structured according to the probability of cooperation. In this case, reciprocity is represented by an advective term. For a simple set of hypotheses we show that cooperation might mark the difference between extinction or proliferation for two interacting populations.

4.3.2 Prisoner's dilemma and reciprocity

According to [157], an initial intention for cooperation and the existence of reciprocity are crucial for the evolution of cooperation, even in an environment composed of egoistic individuals. However, one may wonder what are the conditions that guarantee this to be true; after all, it can be expected that, if reciprocity is stronger in the absence of cooperation, then cooperation becomes less usual. In other words: when is reciprocity a catalyst for cooperation? The following (very simple) model tackles this question.
Consider two players (that can range from cells to entire groups of individuals, such as governments) involved in the repeated prisoner's dilemma game. Player A will initially cooperate with probability $p_0 > 0$ while player B will do so with probability $q_0 > 0$. We assume both values to be strictly positive to account for the initial intention of cooperation described in [157]. Both players will modify their probabilities of cooperation at turn $k+1$ (denoted as p_{k+1} and q_{k+1} respectively) by following the rule:

$$p_{k+1} = \begin{cases} p_k + \varepsilon_{11}(1-p_k), & \text{if player B cooperated in turn k,} \\ \\ p_k(1-\varepsilon_{12}), & \text{if not,} \end{cases}$$

and

$$q_{k+1} = \begin{cases} q_k + \varepsilon_{21}(1-q_k), & \text{if player A cooperated in turn k,} \\ \\ q_k(1-\varepsilon_{22}), & \text{if not,} \end{cases}$$

where $0 < \varepsilon_{ij} < 1$ for $i, j \in \{1, 2\}$. According to this model, both players modify their strategy by "learning" from each other. A different strategy was already studied in [158], where players could modify their strategy by imitation.
We recall that the payoff matrix of the prisoner's dilemma game is given by

$$\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix},$$

where b is the benefit and c is the cost of cooperation ($b > c$). Hence, the expected gain for players A and B at turn k are given by

$$E_A^k = (b-c)p_k q_k + b(1-p_k)q_k - cp_k(1-q_k) = bq_k - cp_k \text{ and } E_B^k = bp_k - cq_k,$$

respectively. Therefore, the average expected gain at turn k is given by the relation

$$E_k = \frac{(b-c)}{2}(p_k + q_k).$$

Given that the probability of both players cooperating at turn k is equal to $p_k q_k$, our interest falls then on the question: What are the conditions over the values ε_{ij}, $i, j \in \{1, 2\}$, such that the sequence (p_k, q_k) converges towards a non trivial limit ? In such cases, when does the average expected gain can be expected to increase ?

In order to answer these questions we first explicitly give the values of p_{k+1} and q_{k+1} as functions of p_k and q_k. Thanks to the law of total probability, we get the relations

$$\begin{aligned} p_{k+1} &= q_k(p_k + \varepsilon_{11}(1-p_k)) + (1-q_k)p_k(1-\varepsilon_{12}) \\ &= (1-\varepsilon_{12})p_k + \varepsilon_{11}q_k + (\varepsilon_{12} - \varepsilon_{11})p_k q_k =: f_1(p_k, q_k), \\ q_{k+1} &= p_k(q_k + \varepsilon_{21}(1-q_k)) + (1-p_k)q_k(1-\varepsilon_{22}) \\ &= (1-\varepsilon_{22})q_k + \varepsilon_{21}p_k + (\varepsilon_{22} - \varepsilon_{21})p_k q_k =: f_2(p_k, q_k). \end{aligned}$$

If this sequence has a limit (p^*, q^*), it must satisfy the relation

$$\begin{cases} p^* &= f_1(p^*, q^*), \\ \\ q^* &= f_2(p^*, q^*). \end{cases} \tag{4.2}$$

In the following proposition we will identify the possible values for (p^*, q^*) and determine their stability.

Proposition 4.1. *Consider a couple (p_0, q_0) and the value $e = \varepsilon_{11}\varepsilon_{21} - \varepsilon_{12}\varepsilon_{22}$.*

i) If $e < 0$, then the only possible values for (p^*, q^*) are $(0,0)$ and $(1,1)$. The first one is a stable fixed point and the second one is an unstable fixed point.

ii) If $e > 0$, then the only possible values for (p^*, q^*) are $(0,0)$ and $(1,1)$. The first one is an unstable fixed point and the second one is an stable fixed point.

iii) If $e = 0$ then (p^*, q^*) is the unique solution of

$$\varepsilon_{22}p_0 + \varepsilon_{11}q_0 = \varepsilon_{22}p^* + \varepsilon_{11}q^*,$$
$$q^* = \frac{\varepsilon_{12}p^*}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11})p^*},$$

and it is a stable fixed point.

Proof. Notice that the values $(0,0)$ and $(1,1)$ are always a solution of (4.2). The stability of said fixed points (and others we wil determine) can be study by means of the eigenvalues of the Jacobian matrix.

First case: The unbalanced scenario: $(\varepsilon_{11}\varepsilon_{21} \neq \varepsilon_{12}\varepsilon_{22})$ If this condition is satisfied, a simple computation shows that there are not non-trivial solutions for (4.2). Hence, if a limit exists, it has to be either the $(0,0)$ or the $(1,1)$. The Jacobian matrix of the system at each of these points is equal to

$$J_0 := J(0,0) = \begin{pmatrix} 1 - \varepsilon_{12} & \varepsilon_{11} \\ \varepsilon_{21} & 1 - \varepsilon_{22} \end{pmatrix} \text{ and } J_1 := J(1,1) = \begin{pmatrix} 1 - \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{22} & 1 - \varepsilon_{21} \end{pmatrix}.$$

The eigenvalues of J_0 are then

$$\lambda_1^0 = \frac{2 - (\varepsilon_{12} + \varepsilon_{22}) - \sqrt{(\varepsilon_{12} + \varepsilon_{22})^2 + 4e}}{2} \text{ and } \lambda_2^0 = \frac{2 - (\varepsilon_{12} + \varepsilon_{22}) + \sqrt{(\varepsilon_{12} + \varepsilon_{22})^2 + 4e}}{2},$$

while those of J_1 are

$$\lambda_1^1 = \frac{2 - (\varepsilon_{11} + \varepsilon_{21}) - \sqrt{(\varepsilon_{11} + \varepsilon_{21})^2 - 4e}}{2} \text{ and } \lambda_2^1 = \frac{2 - (\varepsilon_{11} + \varepsilon_{21}) + \sqrt{(\varepsilon_{11} + \varepsilon_{21})^2 - 4e}}{2}.$$

If $e < 0$, then $-1 < 1 - (\varepsilon_{12} + \varepsilon_{22}) < \lambda_1^0 < \lambda_2^0 < 1$ and $\lambda_2^1 > 1$, hence $(0,0)$ is stable and $(1,1)$ is unstable. On the other hand, if $e > 0$, then $\lambda_2^0 > 1$ and $-1 < 1 - (\varepsilon_{11} + \varepsilon_{21}) < \lambda_1^1 < \lambda_2^1 < 1$, hence $(0,0)$ is unstable and $(1,1)$ is stable.

Second case: The balanced scenario ($\varepsilon_{11}\varepsilon_{21} = \varepsilon_{12}\varepsilon_{22}$) Under this condition, it is straightforward to notice the relation

$$\varepsilon_{22}p_{k+1} + \varepsilon_{11}q_{k+1} = \varepsilon_{22}p_k + \varepsilon_{11}q_k, \text{ for all } k \in \mathbb{N},$$

hence, if a limit (p^*, q^*) exists, it satisfies

$$r^* := \varepsilon_{22}p^* + \varepsilon_{11}q^* = \varepsilon_{22}p_0 + \varepsilon_{11}q_0 =: r_0.$$

Furthermore, directly from the relation $f_1(p^*, q^*) = p^*$ we get the equality

$$q^* = \frac{\varepsilon_{12}p^*}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11})p^*}.$$

Hence, the value of (p^*, q^*) is given by the unique solution of the system

$$\begin{cases} r_0 & = \quad \varepsilon_{22}p^* + \varepsilon_{11}q^*, \\ q^* & = \quad \dfrac{\varepsilon_{12}p^*}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11})p^*}. \end{cases} \tag{4.3}$$

Computing the Jacobian matrix at (p^*, q^*) gives

$$J_* := J(p^*, q^*) = \begin{pmatrix} 1 - \varepsilon_{11}\frac{q^*}{p^*} & \varepsilon_{12}\frac{p^*}{q^*} \\ \varepsilon_{22}\frac{q^*}{p^*} & 1 - \varepsilon_{21}\frac{p^*}{q^*} \end{pmatrix}.$$

The eigenvalues of J_* are

$$\lambda_1^* = 1 - (\varepsilon_{11}\frac{q^*}{p^*} + \varepsilon_{21}\frac{p^*}{q^*}) \text{ and } \lambda_2^* = 1.$$

Given that the second eigenvalue is equal to 1, we cannot immediately give a conclusion to the stability of (p^*, q^*). However, we can proceed as follows: for a fixed (p_0, q_0), p^* is solution of the equation

$$\begin{aligned} p^* =& f_1(p^*, \frac{r^* - \varepsilon_{22}p^*}{\varepsilon_{11}}) \\ =& (1 - \varepsilon_{12})p^* + (r^* - \varepsilon_{22}p^*) + (\varepsilon_{12} - \varepsilon_{11})p^*\frac{r^* - \varepsilon_{22}p^*}{\varepsilon_{11}} \\ =& r^* + (1 - (\varepsilon_{12} + \varepsilon_{22}) + (\varepsilon_{12} - \varepsilon_{11})\frac{r^0}{\varepsilon_{11}})p^* + (\varepsilon_{22} - \varepsilon_{21})(p^*)^2 \\ =:& f(p^*). \end{aligned}$$

This is, p^* is a fixed point of $f(p)$. Therefore, in order to determine the stability of (p^*, q^*), it suffices to study the value of

$$\begin{aligned} f'(p^*) &= (1 - (\varepsilon_{12} + \varepsilon_{22}) + (\varepsilon_{12} - \varepsilon_{11})\frac{r^0}{\varepsilon_{11}}) + 2(\varepsilon_{22} - \varepsilon_{21})p^* \\ &= 1 - (\varepsilon_{11}\frac{q^*}{p^*} + \varepsilon_{21}\frac{p^*}{q^*}), \end{aligned}$$

which is precisely the first eigenvalue of J^*. Since $\lambda_1^* < 1$, (p^*, q^*) will be a stable fixed point if and only if $\lambda_1^* > -1$, or equivalently, if and only if

$$g(p^*) := \frac{\varepsilon_{11}\varepsilon_{12}}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11})p^*} + \frac{\varepsilon_{21}}{\varepsilon_{12}}(\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11})p^*) = \varepsilon_{11}\frac{q^*}{p^*} + \varepsilon_{21}\frac{p^*}{q^*} < 2.$$

Since $g(p)$ is a convex function over $[0, 1]$, which satisfies $g(0) = \varepsilon_{12} + \varepsilon_{22} < 2$ and $g(1) = \varepsilon_{11} + \varepsilon_{21} < 2$, we conclude $g(p^*) < 2$ for all possible values of p^*. Therefore (p^*, q^*) is a stable fixed point. □

Let us discuss the results from Proposition 4.1. There are two scenarios for the unbalanced case. If the players reaction to the lack of cooperation is stronger than the reaction to the presence of it ($e < 0$), then both players will eventually adopt the no cooperation strategy, making the average expected gain equal to 0. On the other hand, two players that are highly responsive to cooperation, and not to the lack of it ($e > 0$), will eventually always cooperate, maximising this way the average expected gain. We observe a far more complicated outcome when the responses of both players are balanced ($e = 0$). Given that (p^*, q^*) satisfies system (4.3), then the average expected gain will increase if $\varepsilon_{11} < \varepsilon_{22}$ and the initial values p_0 and q_0 satisfy

$$q_0 > \frac{\varepsilon_{12} p_0}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11}) p_0},$$

or if $\varepsilon_{11} > \varepsilon_{22}$ and

$$q_0 < \frac{\varepsilon_{12} p_0}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11}) p_0}.$$

Thanks to the balance condition, $\varepsilon_{11} < \varepsilon_{22}$ implies that $\varepsilon_{12} < \varepsilon_{21}$. This is, in a way, player A has more shy responses than player B. According to the previously established conditions, interactions between these two players will lead to an increase in the average expected gain only if the initial probability of cooperation for player B is sufficiently big. An analogous interpretation can be given when $\varepsilon_{11} > \varepsilon_{22}$. Figure 4.2 shows several initial configurations for (p_0, q_0) and their respective limiting values satisfying the relation.

$$q^* = \frac{\varepsilon_{12} p^*}{\varepsilon_{11} + (\varepsilon_{12} - \varepsilon_{11}) p^*}.$$

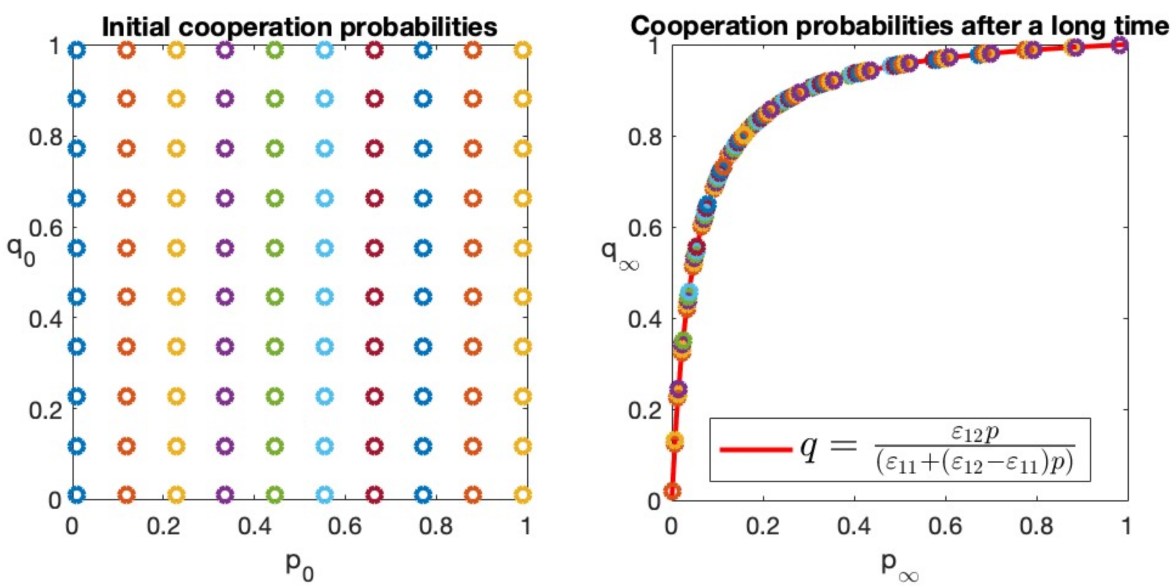


Figure 4.2: Left panel: Several initial configurations of cooperation probabilities. Right panel: Limiting values of the sequences (p_k, q_k) associated to initial values showcased on the previous figure.

Assume now the presence of n players, each one with an initial probability of cooperation p_0^i and reciprocity constants $(\varepsilon_{i1}, \varepsilon_{i2})$, for $i \in \{1, \ldots, n\}$. The previous model can be adapted in such a way that each player modifies its strategy by taking into account the global cooperation level. This is, p_k^i satisfies the relation

$$\begin{aligned} p_{k+1}^i &= q_k^i(p_k + \varepsilon_{i1}(1 - p_k^i)) + (1 - q_k^i)p_k^i(1 - \varepsilon_{i2}) \\ &= (1 - \varepsilon_{i2})p_k^i + \varepsilon_{i1}q_k^i + (\varepsilon_{i2} - \varepsilon_{i1})p_k^i q_k^i, \end{aligned}$$

with

$$q_k^i := \frac{1}{n-1} \sum_{j \neq i} p_k^j,$$

being the average probability of cooperation from the co-players of player i. As in the previous case, it can be expected that the amount of fixed points for this recurrence, and its stability will depend on a family of conditions over the values of $(\varepsilon_{i1}, \varepsilon_{i2})$, however, for the moment being, we will not study this case any further.

An element that was not considered in these models was the effect of the average expected gain on the relation between (p_k, q_k) and (p_{k+1}, q_{k+1}). For example, considering variable reciprocity coefficients which directly depend on the average expected gain would create a mutual feedback between the cooperation probabilities and the gain, resulting this way in a far more complex, interesting and realistic model.

4.3.3 A continuously structured population model for the evolution of cooperation

Take $p \in [0, 1]$ to be a continuous structure variable representing a probability of cooperation. Consider two populations A and B, each one composed by individuals with different probabilities of cooperation with the elements on the other population. Let $n_A(t, p)$ and $n_B(t, p)$ be their respective population densities of individuals with probability of cooperation equal to p at time t. The total populations at time t are given by

$$\rho_A(t) := \int_0^1 n_A(t, p)dp \text{ and } \rho_B(t) := \int_0^1 n_B(t, p)dp,$$

and the mean cooperation probabilities by

$$\tilde{p}_A(t) := \frac{\int_0^1 p n_A(t, p)dp}{\rho_A(t)} \text{ and } \tilde{p}_B(t) := \frac{\int_0^1 p n_B(t, p)dp}{\rho_B(t)}.$$

These choices allow to define the global expected gain for each population. For the first population its global expected gain is defined then as

$$E_A(t) := (b - c)\tilde{p}_A(t)\tilde{p}_B(t) + b(1 - \tilde{p}_A(t))\tilde{p}_B(t) - c\tilde{p}_A(t)(1 - \tilde{p}_B(t)) = b\tilde{p}_B(t) - c\tilde{p}_A(t),$$

where b and c are the benefit and cost, respectively, of cooperation in the prisoner's dilemma setting[1]. Similarly, the expected gain for population B is given by

$$E_B(t) := b\tilde{p}_A(t) - c\tilde{p}_B(t).$$

[1]For a more general model,the values of b and c could be dependent on p, this is, the cost and benefit of cooperation might depend on the probability of cooperation itself

This way, we may consider that the population densities evolve following the system of equations

$$\begin{cases} \partial_t n_A(t,p) + \varepsilon_A \partial_p \left((\tilde{p}_B(t) - p) n_A(t,p)\right) = & g_A(p, E_A(t)) n_A(t,p), \\ \partial_t n_B(t,p) + \varepsilon_B \partial_p \left((\tilde{p}_A(t) - p) n_B(t,p)\right) = & g_B(p, E_B(t)) n_B(t,p), \\ n_A(0,p) = n_A^0(p), \quad n_B(0,p) = n_B^0(p), \end{cases} \tag{4.4}$$

where ε_A and ε_B are reciprocity coefficients and g_A, g_B are continuous and increasing functions of E_A and E_B respectively, while the elements of both populations modify their probabilities of cooperation, depending on the global probability of cooperation of their counterpart.

Cooperation or extinction, an easy choice From model (4.4), we will illustrate, for an specific choice of g_A and g_B, how cooperation may make a difference between extinction or persistence. Consider

$$\begin{aligned} g_A(p, E_A(t)) &:= r_A(p) + \gamma_A(p) E_A(t) = r_A(p) + \gamma_A(p)(b\tilde{p}_B(t) - c\tilde{p}_A(t)), \\ g_B(p, E_B(t)) &:= r_B(p) + \gamma_B(p) E_B(t) = r_B(p) + \gamma_B(p)(b\tilde{p}_A(t) - c\tilde{p}_B(t)), \end{aligned} \tag{4.5}$$

where $r_A(p), r_B(p)$ are the respective intrinsic growth rates of populations A and B and the non-negative functions $\gamma_A(p), \gamma_B(p)$ represent the effect of the expected gain on the growth rate of each population. This choice of g_A and g_B makes system (4.4) bear a striking resemblance to the model studied in [28], where conditions under which there is persistence of all species are given. Nevertheless, there are several differences: In our case the non local terms are given by the mean cooperation probabilities, the functions γ_A and γ_B are non-negative and we consider no restrictions over the signs of $r_A(p)$ and $r_B(p)$. Despite these differences, we do not rule out the fact that the tools and techniques used within the cited reference may be useful for the study of problem (4.4) as well. For specific choices of ε_A, ε_B, γ_A and γ_B it is possible to identify the conditions over r_A, r_B, b and c which guarantee that one or both populations will either go extinct or proliferate. Such conditions are stated on the following proposition:

Proposition 4.2. *Consider $\varepsilon_A = \varepsilon_B = 0$, $\gamma_A(p) \equiv \gamma_A$ and $\gamma_B(p) \equiv \gamma_B$, with γ_A, γ_B non negative constants. Suppose $r_A(p)$, $r_B(p)$, $n_A^0(p)$ and $n_B^0(p)$ to be continuous functions such that the maximum value of $r_A(p)$ over the support of $n_A^0(p)$ is attained at a single point p_A^*, and the maximum value of $r_B(p)$ over the support of $n_B^0(p)$ is attained at a single point p_A^*. Then*

i) If $r_A(p_A^*) + \gamma_A(bp_B^* - cp_A^*) < 0$, population A will go extinct.

ii) If $r_A(p_A^*) + \gamma_A(bp_B^* - cp_A^*) > 0$, there exists and interval I satisfying $p_A^* \in I \subset [0,1]$ such that population A will blow up for all $p \in I$.

iii) The same is true for population B, depending on the sign of $r_B(p_B^*) + \gamma_B(bp_A^* - cp_B^*)$.

Proof. Under these hypotheses, the expression for $n_A(p)$ and $n_B(p)$ are implicitly given by the expressions

$$n_A(t,p) = n_0^A(p) e^{r_A(p)t + \gamma_A \int_0^t E_A(s)ds} \text{ and } n_B(t,p) = n_0^B(p) e^{r_B(p)t + \gamma_B \int_0^t E_B(s)ds},$$

respectively. This allows to explicitly compute the values of $\tilde{p}_A(t)$ and $\tilde{p}_B(t)$:

$$\tilde{p}_A(t) = \frac{\int_0^1 p n_A^0(p) e^{r_A(p)t} dp}{\int_0^1 n_A^0(p) e^{r_A(p)t} dp} \text{ and } \tilde{p}_B(t) = \frac{\int_0^1 p n_A^0(p) e^{r_B(p)t} dp}{\int_0^1 n_A^0(p) e^{r_B(p)t} dp}.$$

From here, it is not hard to prove that, under the hypotheses of Proposition 4.2, $\tilde{p}_A(t)$ and $\tilde{p}_B(t)$ converge towards p_A^* and p_B^* respectively. This implies that, for all positive ε there exists $T > 0$ such that $r(p) + \gamma_A(b\tilde{p}_B(t) - c\tilde{p}_A(t)) \leqslant r_A(p_A^*) + \gamma_A(bp_B^* - cp_A^*) + \varepsilon$ for all $t > T$. If ε is chosen small enough, then $r(p) + \gamma_A(b\tilde{p}_B^*(t) - c\tilde{p}_A^*(t)) < 0$ for all $t > T$ which gives the convergence to 0 of the population. Conversely, if $r_A(p_A^*) + \gamma_A(bp_B^* - cp_A^*) > 0$, we set $\delta := \frac{r_A(p_A^*)+\gamma_A(bp_B^*-cp_A^*)}{2}$, and define

$$I = \{p \in [0,1] : r_A(p) > r_A(p_A^*) - \delta\}.$$

Hence, for all $p \in I$ there exists $T > 0$ such that

$$r(p) + \gamma_A(b\tilde{p}_B(t) - c\tilde{p}_A(t)) \geqslant r_A(p_A^*) - \delta + \gamma_A(b\tilde{p}_B(t) - c\tilde{p}_A(t)) - \varepsilon = \delta - \varepsilon,$$

for all $t > T$. Once again, by choosing ε small enough we obtain the strictly positive growth rate for all values of $p \in I$, which implies the blow up of the population for all such values of p.
The proof for population B is analogous. □

Let us illustrate the result of Proposition 4.2 with an example. Consider

$$r_A(p) = r_B(p) = p(1-p) - \frac{1}{2} < 0.$$

It is straightforward to conclude that, if there is no cooperation ($\gamma_A(p) = \gamma_B(p) = 0$ or $n_0^A(p) = n_0^B(p) = \rho_0\delta_0(p)$) then both populations will go extinct, at an exponential rate. On the other hand, consider $\gamma_A(p) = \gamma_B(p) = 1$, $n_0^A(p) \equiv n_0^A$ and $n_0^B(p) \equiv n_0^B$. Under these assumptions, we have

$$n_A(t,p) = n_0^A e^{r_A(p)t + \int_0^t E_A(s)ds} \text{ and } n_B(t,p) = n_0^B e^{r_B(p)t + \int_0^t E_B(s)ds},$$

and consequently we get

$$\tilde{p}_A(t) = \frac{\int_0^1 pe^{r_A(p)t}dp}{\int_0^1 e^{r_A(p)t}dp} = \frac{1}{2} \text{ and } \tilde{p}_B(t) = \frac{\int_0^1 pe^{r_B(p)t}dp}{\int_0^1 e^{r_B(p)t}dp} = \frac{1}{2},$$

after integrating by means of a substitution. This way, the equations for $n_A(t,p)$ and $n_B(t,p)$ are reduced to

$$\begin{cases} \partial_t n_A(t,p) = (r_A(p) + \frac{(b-c)}{2}) n_A(t,p), \\ \\ \partial_t n_B(t,p) = (r_B(p) + \frac{(b-c)}{2}) n_B(t,p), \\ \\ n_A(0,p) = n_A^0, \quad n_B(0,p) = n_B^0. \end{cases}$$

It is then evident that, as long as $(b-c) > 1$ there will be values of p for which $r_A(p) + \frac{(b-c)}{2} > 0$ and $r_B(p) + \frac{(b-c)}{2} > 0$, hence, the population densities $n_A(t,p)$ and $n_B(t,p)$ will be proliferating exponentially. An interesting question left unanswered is the case $r_A(p_A^*) + \gamma_A(bp_B^* - cp_A^*) = 0$. In this scenario, additional conditions over the parameters of the problem might be needed in the general case in order to determine the behaviour of the solution. For the previous illustrative example, this condition is equivalent to choosing $b - c = \frac{1}{2}$, which leads to a solution which decreases for all $p \neq 1/2$ and that remains constant for $p = 1/2$. It is also of interest to identify whether the effect of cooperation on the populations dynamics proved in Proposition 4.2 can be observed for a more general family of conditions, and choices of g_A and g_B. As mentioned before, the tools presented in [28] might be of use in order to better understand the long time dynamics of both populations, and identify concentration phenomena,

stable steady states and rates of convergence or explosion.
The study of the effect of the advection term, representing reciprocity adds a layer of complexity to the study of the problem. We refer to [30] and the references therein, where similar non local advection-reaction problems have been studied, but for a single population. A diffusion term can be considered as well in both equations of system (4.4) in order to model random instabilities of the probabilities of cooperation. This term may be a second order differential operator and suitable boundary conditions, or an integral term with a mutation kernel. A similar model, for only one population, excluding the advection term and depending on the population size as the non-local term was already studied in [23]. Finally, if all three terms are considered, the resulting model will follow the same principles as in model (4.1), where the diffusion term represents the non-genetic instability of trait p, the advection term represents the external stress exerted over each population as in Chapter 3 or the existence of a bias in the direction of epimutations, as in [27] (in our case such stress or bias is prompted by the global cooperation probability of the other population) and the reaction term accounts for selection mechanisms.

4.4 Conclusion

We have sketched in this short essay, relying on concepts of philosophy of science and on mathematical models under development, the two settings of evolution in which phenotype divergence and cooperation between phenotypes in the constitution of animal multicellularity should be considered from our point of view. They are the billion-year Darwinian evolution of species - which we assimilate with the evolution of body plans - and the short-term construction, in embryogenesis and development, of an isogenic animal from the zygote to the constituted, terminally differentiated multicellular organism.

In the first case, phenotype divergence is considered to be determined by changes in the environment, and it is represented by an advection term in a PDE, yielding different optimal adaptive strategies that are chosen randomly in the initial body plan and resulting in (at least) two different body plans, that in the first place should be reversible, before being fixed by stabilising mutations.

In the second case, the body plan of a given coherent multicellular animal, that has been established in Darwinian evolution in a deterministic machinery of embryogenesis and organism maintenance, governs the process of development from the zygote of the animal individual on principles of compatibility and cooperativity between physiological functions, organs and tissues, that relies on cell differentiations. Of note, cellular stress-induced genes might evolve into developmental organisers, according to a mechanism proposed in the Chlamydomonas/Volvox lineage [159]. Such differentiations are by nature theoretically reversible, relying on epigenetic enzyme activities which graft methyl or acetyl radicals on the DNA or on the histones that constitute the genome on animal, and dedifferentiations indeed have been shown to be experimentally possible in 2006 by Takahashi and Yamanaka [160]. However, they are physiologically excluded, except in particular situations such as wound healing, by a strict control of the expression of these epigenetic enzymes. Plasticity in cancer cells alters such normal organismic control.

In cancer, which is a disease characteristic of multicellular animals, differentiations are (locally, in the tissue from which it originates) out of organismic control, so that tumours, as poorly organised cell colonies that nevertheless are made of cells bearing in each one of them the body plan of a multicellular organism, can reactivate a process of phenotype divergence in response to a deadly insult (such as a chemotherapy at high doses), resulting in cancer bet hedging, i.e., developing diverse transient (reversible) phenotypes without organised control, with the goal to preserve the proliferation potential of their cells.

We are aware that the mathematical models presented here are sketches that need refinement, and that in particular the cooperativity part should be oriented towards defining a compulsory common gain (likely represented by, again, an advection term in a PDE) that determines the precise construction of an individual animal organism designed by its body plan. Much still remains to be done towards this goal, and in particular the body plan - whose effects are patent in embryogenesis and development, but is still not properly defined as a programme - needs to be better defined in a mathematical representation. It is likely made of an organised ensemble of gene regulatory networks, as evidenced in the works of Eric Davidson [127] and his colleagues, and systematically described in the diversity of its functions in hypothetical *Urmetazoa* by W.E.G. Muller and his colleagues [128]. A mathematical representation of the body plan, as a programme of construction of the individual and as the evolutionary unit on which relies Darwinian evolution and the design of animal anatomy and physiology, is a challenge that awaits philosophers, evolutionary biologists, and mathematical modellers and analysts, a challenge we have merely sketched in this short essay.

Chapter 5

A particle method for non-local advection-selection-mutation equations

5.1 Introduction

Presentation of the model

The goal of this chapter is to develop a numerical method allowing to approximate the solutions of equations of the form

$$\begin{cases} \partial_t v(t,x) + \nabla_x \cdot (a(t,x,I_a v(t,x))v(t,x)) = R(t,x,I_g v(t,x))v(t,x) + \displaystyle\int_{\mathbb{R}^d} m(t,x,y,I_d v(t,x))v(t,y)dy, \\ v \in \mathscr{C}([0,T], L^1(\mathbb{R}^d)), \\ v(0,\cdot) = v^0(\cdot) \in W^{1,1}(\mathbb{R}^d), \end{cases} \tag{5.1}$$

where

$$(I_l u)(t,x) = \int_{\mathbb{R}^d} \psi_l(t,x,y)u(t,y)dy, \quad l = a, g, d$$

are non-local terms and a, R, m and ψ_l are smooth functions.

This general formulation aims to bring together a wide family of PDE models typically used in the field of *adaptive dynamics*. In this context, x represents a phenotypic trait (usually simply called 'phenotype' or 'trait') which is a characteristic inherent to individuals, and $x \mapsto v(t,x)$ represents the density of the studied population at time $t \geq 0$. One purpose of adaptive dynamics is to understand the combined effect of selection and mutations (which are usually assumed to be rare and small [161–163]) on living populations [164]. The literature concerning phenotype-structured equations is abundant [22,165–170]. The model proposed in this chapter (which includes, among others, the equations studied in [24,30,171–174]) takes into account

- Selection and growth, via the term '$R(t,x,I_g v(t,x))v(t,x)$', where R can be interpreted as the instantaneous growth rate, which depends on the trait x and the whole population.

- Mutation, via the term '$\int_{\mathbb{R}^d} m(t,x,y,I_d v(t,x))v(t,y)dy$', where the function m can be seen as the probability density for a cell of trait y to mutate into a cell of trait x.

- Advection, via the term '$\nabla_x \cdot (a(t,x,I_a v(t,x))v(t,x))$'. This term models how the environment drives the individuals towards specific regions, as opposed to random mutations. Among others, this term can be used in order to model a cell differentiation phenomenon.

Mutations can also be modelled through a second order differential operator such as in [111] and Chapter 3. Laplacian-like terms can be approximated by integral operators, as shown in [57], which means that, after choosing an appropriate integral approximation, our analysis could be extended to deal with second order equations. The other two non local-terms (I_a and I_g) allow to take into account the influence of the environment, created by the whole population, over the behaviour of the individuals [174, 175], or competition between individuals [22].

The long time behaviour of models considering only one phenomenon among selection, advection and mutation is well-known: broadly speaking, it has been shown that considering selection alone, or advection alone, leads to concentration phenomena (towards a finite number of traits) [22, 171, 176], meaning that the density converges to a sum of Dirac masses, while mutations by themselves have a smoothing effect [177]. Nevertheless, the combined effects of these terms remains unclear, and may lead to different and non-intuitive behaviours. As an example, considering both selection and advection can lead to convergence either to a Dirac mass or to a continuous function [30,178], and considering mutation and selection leads either to convergence to a non-smooth measure or to a continuous function [173]. Note that this model also includes the equation studied in [179], which was also approximated with a particle method.

Upon establishing the well-posedness of (5.1), this chapter is concerned with the derivation of a particle method inspired by [57], the analysis of its convergence and asymptotic-preserving properties. However, we must emphasise the two main novelties with respect to that work: First, the use of non local terms, which as we will show, poses technical difficulties and affects the existence of smooth solutions in certain cases. Secondly, the study of the asymptotic preserving property, which guarantees that, under certain hypotheses, the long time behaviour of the solution is conserved. As will be seen, these equations are naturally posed in the space of Radon measures, making particle methods a natural tool to approximate then. Compared to finite volume or finite element methods, they are more easily implemented. Furthermore, a change in model leads to very few changes in the corresponding code, a clear advantage over other methods.

Particle method

Particle methods use ODE resolution in order to approximate the solution of PDEs. This makes them particularly easy to implement, as they only require a classical ODE solver. The main idea is to seek a sum of weighted Dirac masses, called particle solution, which is denoted

$$v^N(t) = \sum_{i=1}^{N} \alpha_i(t)\delta_{x_i(t)}, \tag{5.2}$$

where the weights α_i and the points x_i are solutions of a suitable ODE system.

In order to recover a smooth function close to the solution of the studied PDE, the particle solution needs to be regularised: this is usually done by means of a convolution with a so-called 'cut-off function'

φ_ε which must satisfy some specific properties. We denote this regular solution

$$v_\varepsilon^N(t,x) = \sum_{i=1}^{N} \alpha_i(t)\varphi_\varepsilon(x - x_i(t)), \tag{5.3}$$

where the scaling parameter ε is a function of N.

This method is especially adapted for the linear advection equation '$\partial_t v(t,x) + \nabla \cdot (a(x)v(t,x)) = 0$', but has been generalised to many other kinds of equations which mostly come from physics [49], such as diffuson equations [50–54], advection-diffusion equations [55, 56], convection-diffusion equations [57], the Navier-Stokes equation [58, 59] or the Vlasov-Poisson equation [60, 61].

We apply the particle method by following its main three steps, as described in [180]:

1. **Particle approximation of the initial data.** This first step consists of approaching the initial condition of v^0 with a sum of weighted Dirac masses, *i.e.* choosing $N \in \mathbb{N}$, $x_1^0, ..., x_N^0 \in \mathbb{R}^d$, $\alpha_1^0, ..., \alpha_N^0 \in \mathbb{R}$ such that

$$v_0^N := \sum_{i=1}^{N} \alpha_i^0 \delta_{x_i^0} \sim v^0,$$

   in the sense of Radon measures, which means that, for any $\phi \in \mathscr{C}_c^0(\mathbb{R}^d)$,

$$\sum_{i=1}^{N} \alpha_i^0 \phi(x_i^0) \underset{N\to+\infty}{\longrightarrow} \int_{\mathbb{R}^d} \phi(x)v^0(x)dx.$$

   Assuming that v^0 has a compact support, a canonical way of choosing these values is to choose a finite collection of subsets $\Omega_i^0 \subset \text{supp}\,(v^0)$ satisfying

$$\Omega_i^0 \cap \Omega_j^0 = \emptyset, \text{ if } i \neq j, \text{ and } \bigcup_{i\in\{1,...,N\}} \Omega_i^0 = \text{supp}\left(v^0\right),$$

   and to take, for any $i \in \{1, ..., N\}$

$$x_i^0 \in \Omega_i^0, \quad w_i^0 = |\Omega_i^0|, \quad \nu_i^0 = v^0(x_i^0) \quad \alpha_i = \nu_i w_i.$$

2. **Time-evolution of the particles.** By using a weak formulation of the PDE, we determine the ODE satisfied by the positions (denoted x_i), the volumes (denoted w_i) and the weights (denoted ν_i) associated to each of the N particles, with initial conditions (x_i^0, w_i^0, ν_i^0) given at the previous step. The exact ODE, and the way the particles are correlated with each other depends on the complexity of the PDE. In the case where the advection term is local ($a(t,x,I) = a(t,x)$), the positions and the volumes satisfy

$$\dot{x}_i(t) = a(t, x_i(t)), \quad \dot{w}_i(t) = \nabla_x \cdot (a(t, x_i(t)))w_i(t) \tag{5.4}$$

   and the formula for the ν_i depends on the selection and the mutation terms. The method described in the core of this chapter also allows to use this method in the case of a non-local advection, which modifies the ODE satisfied by the positions of the particles. Full formulas are given in Section 5.3.

Rewriting $\alpha_i = \nu_i w_i$, with the volumes w_i which satisfy (5.4) is required for the approximation of the different integral terms. Indeed, by Liouville's formula [181], for any f smooth enough,

$$\int_{\mathbb{R}^d} f(x)dx \sim \sum_{i=1}^{N} f(x_i(t))w_i(t).$$

Formally, v and v^N (as defined by (5.2)) are both solution of (5.1) in the weak sense, with $v^N(0) \sim v^0$, which implies that, assuming that the parameters of the PDE are smooth enough, for any given time $T > 0$, $v^N(T) \sim v(T, \cdot)$.

3. **Regularisation.** In order to transform the discrete measure v^N into a smooth function, we use a regularisation process based on convolution, which writes as a sum, shown in (5.3), since the convoluted measure is a sum of Dirac masses. The function

$$\varphi_\varepsilon := \frac{1}{\varepsilon}\varphi\left(\frac{\cdot}{\varepsilon}\right),$$

depending on a parameter $\varepsilon > 0$, is a scaling of the so-called *cut-off function* φ, which must satisfy some regularity and symmetry properties (which we specify in Section 5.4.1). The choice of ε, which intimately depends on the choice of N, is intricate: if ε is too large, then the solution is 'over-regularised', and the scheme loses its accuracy. Conversely, if ε is to small, then some of the particles will be neglected, and the scheme does not converge towards the solution. Choosing the optimal ε as a function of N is thus not a trivial question, and it is possible in some cases to optimise the convergence rate by improving this regularisation step [182, 183].

Main results

Well-posedness. We first prove that problem (5.1) is well-posed, *i.e.* that for any family of parameters satisfying some regularity properties, defined in $\mathscr{C}([0,+\infty), L^1(\mathbb{R}^d))$. The proof heavily relies on the use of the characteristic curves $X_u(t,y)$, solution to the equation

$$\begin{cases} \dot{X}_u(t,y) = a(t, X_u(t,y), (I_a u)(t, X_u(t,y))) =: \mathscr{A}_u(t, X_u(t,y)), \quad t \in [0,T], \\ X_u(0,y) = y, \end{cases}$$

for $y \in \mathbb{R}^d$ and $u \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$. The main difference with respect to the approach taken for similar problems, as in [57], is the need for continuity results for $X_u(t,y)$, not only with respect to the trait variable y, but with respect to u as well. The required results are stated in Section 5.2.1 and proved in Appendix 5.6.1.

The well posedness of the problem for smooth initial data is proved in Section 5.2.2 using a standard fixed point argument. Moreover, in Section 5.2.3, we consider a more general family of initial conditions and we prove that the regularity of $v(t,\cdot)$ is linked to that of v^0: more precisely, if $v^0 \in W^{k,\infty}(\mathbb{R}^d)$ with compact support, then $v \in \mathscr{C}([0,+\infty), W^{k-1,1}(\mathbb{R}^d))$. This result can be improved if the advection is local, *i.e.* $a(t,x,I) = a(t,x)$: In this case, for any initial condition in $W^{k,1}(\mathbb{R}^d)$, the solution v is in $\mathscr{C}([0,+\infty), W^{k,1}(\mathbb{R}^d))$.

Particle method: definition and well-posedness. In Section 5.3, we define the particle method corresponding to this PDE, by deriving the ODE system satisfied by the particles (x_i, w_i and ν_i). For the non-local case, the equation we obtain is a coupled system with infinitely many equations and unknowns. We generalise some classical results from the Cauchy-Lipschitz theory in order to deal with this problem, and prove that the ODE system is well posed.

Convergence of the particle method. Section 5.4 is structured as follows: in Section 5.4.1 we prove the following estimate, detailed in Theorem 5.5:

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leqslant C\big(\varepsilon^r + \big(\frac{h}{\varepsilon}\big)^\kappa + h^\kappa\big)\|v^0\|_{W^{\mu,1}(\mathbb{R}^d)}, \quad \text{for all } \ 0 \leqslant t \leqslant T,$$

))Nevertheless, in general this scheme is not asymptotic preserving. Therefore, in subsection 5.4.2, we show examples for which the scheme is asymptotic preserving, and others for which it is not. In general, the asymptotic behaviour of a solution is preserved when it converges to a sum of Dirac masses, and is not when it converges to a smooth solution.

Perspectives and open problems

Although a loss of regularity appears to be taking place when advection is non local, we do not know whether such a loss of regularity does happen in certain cases. The construction of such an example or, on the contrary, the improvement of our results in order to prove that, in fact, no regularity is lost could be a first way to extend our work.
Another open problem is the optimisation of the order of convergence for the numerical solution, improving upon the order $\frac{\kappa r}{\kappa + r}$ obtained in the present work. The approach from [182], where the local averages are viewed as point values of an approximation of the solution, and the regularisation of the solution at time $t > 0$ is performed by interpolation rather than convolution, could be a suitable choice. Lastly, as mentioned before, another direction could be the extension of our results in order to deal with second order equations, as done in [57], where the Laplacian operator is approximated by an appropriate sequence of mutation kernels.

5.2 The problem

For $T > 0$ and $k \in \mathbb{N}$ we consider the functions

$$(t, x, I) \mapsto a(t, x, I) \in W^{1,\infty}\left([0, T], (W^{k+1,\infty}(\mathbb{R}^{d+1}))^d\right), \tag{5.5}$$

$$(t, x, I) \mapsto R(t, x, I) \in \mathscr{C}\left([0, T] \times \mathbb{R}_x^d, W_{loc}^{k+1,\infty}(\mathbb{R}_I)\right) \cap \mathscr{C}\left([0, T] \times \mathbb{R}_I, W^{k+1,\infty}(\mathbb{R}_x^d)\right). \tag{5.6}$$

We consider as well $(t, x, y, I) \mapsto m(t, x, y, I)$ such that

$$0 \leqslant m \in \mathscr{C}\left([0, T] \times \mathbb{R}_x^d \times \mathbb{R}_y^d, W^{k+1,\infty}(\mathbb{R}_I)\right) \cap \mathscr{C}\left([0, T] \times \mathbb{R}_x^d \times \mathbb{R}_I, L^\infty(\mathbb{R}_y^d)\right) \cap \mathscr{C}\left([0, T] \times \mathbb{R}_y^d \times \mathbb{R}_I, \mathscr{C}_c^k(\mathbb{R}_x^d)\right),$$

which is globally Lipschitz with respect to the non local variables and uniformly compactly supported with respect to the x variable. That is, we suppose m to satisfy the following hypotheses:

- There exists $\mu > 0$ such that

$$\sup_{t,x,y} \sum_{i=1}^{k} \sum_{|\alpha|=i} \sum_{j=1}^{k} |\partial_x^\alpha \partial_I^j m(t,x,y,I) - \partial_x^\alpha \partial_I^j m(t,x,y,J)| \leqslant \mu |I - J|. \tag{5.7}$$

- There exists a compact set $\mathscr{K}$ such that the function

$$M(x) := \sup_{t,y,I} \sum_{i=1}^{k} \sum_{|\alpha|=i} \sum_{j=1}^{k} |\partial_x^\alpha \partial_I^j m(t,x,y,I)|, \tag{5.8}$$

  satisfies

$$\text{supp } M(x) \subset \mathscr{K}. \tag{5.9}$$

  Furthermore, we assume that

$$\|M\|_{L^\infty(\mathbb{R}^d)} \leqslant \overline{M} < \infty. \tag{5.10}$$

  From a modelling point of view, assuming m to be uniformly compactly supported reflects the fact that some traits are realistically out of reach for a given population, and that, in general, mutations are rare and small.

We remark that hypotheses (5.9) and (5.10) imply that

$$\|M\|_{L^1(\mathbb{R}^d)} \leqslant |\mathscr{K}|\overline{M} =: K < \infty. \tag{5.11}$$

For all functions $u \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$ we consider the linear mappings I_a, I_g and I_d which satisfy, for all $t \in [0,T]$, $x \in \mathbb{R}^d$,

$$(I_a u)(t,x) := \int_{\mathbb{R}^d} \psi_a(t,x,y) u(t,y) dy, \tag{5.12}$$

$$(I_g u)(t,x) := \int_{\mathbb{R}^d} \psi_g(t,x,y) u(t,y) dy, \tag{5.13}$$

$$(I_d u)(t,x) := \int_{\mathbb{R}^d} \psi_d(t,x,y) u(t,y) dy, \tag{5.14}$$

where

$$\psi_a \in W^{1,\infty}\left([0,T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d)\right) \cap \mathscr{C}\left([0,T] \times \mathbb{R}_y^d, W^{k+1,\infty}(\mathbb{R}_x^d)\right), \tag{5.15}$$

$$0 < \underline{\psi_g} \leqslant \psi_g \in \mathscr{C}\left([0,T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d)\right) \cap \mathscr{C}\left([0,T] \times \mathbb{R}_y^d, W^{k+1,\infty}(\mathbb{R}^d)\right), \tag{5.16}$$

$$\psi_d \in \mathscr{C}\left([0,T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d)\right) \cap \mathscr{C}\left([0,T] \times \mathbb{R}_y^d, W^{k+1,\infty}(\mathbb{R}^d)\right), \tag{5.17}$$

for a certain $\underline{\psi_g} > 0$. We remark that $I_a u, I_g u, I_d u \in \mathscr{C}([0,T], L^\infty(\mathbb{R}^d))$. The functions ψ_a and ψ_d do not need to be positive, reflecting this way how different traits have different impacts (which are not always beneficial) on the environment, and ultimately on the population itself. On the other hand, ψ_g has to be bounded away from zero. This hypothesis reflects the fact that, at least for the growth term, all interactions between individuals are of the same type. These interactions may be interpreted as

either strictly competitive or strictly cooperative. In particular, this means that only “very” non-local dependence with respect to u are allowed. This excludes partial densities, for instance on part of the traits.
Lastly, we assume that there exist non-negative constants I^* and r^* such that, for all $t \in [0,T]$, $x \in \mathbb{R}^d$, and $I \geqslant I^*$,

$$R(t,x,I) + K < -r^*, \tag{5.18}$$

uniformly on t and x. It is somewhat natural to assume that R is negative for a large population: for example, if a carrying capacity is assumed to exist, and the population size is approaching such value, then the growth rate will inevitably drop to levels where no amount of mutations will be able to compensate for it.
For a given function $v^0 \in L^1(\mathbb{R}^d)$, we will study the existence and uniqueness of solution for the problem

$$\begin{cases} \partial_t v(t,x) + \nabla_x \cdot (a(t,x,I_a v(t,x))v(t,x)) = R(t,x,I_g v(t,x))v(t,x) + \int\limits_{\mathbb{R}^d} m(t,x,y,I_d v(t,x))v(t,y)dy, \\ v \in \mathscr{C}([0,T), L^1(\mathbb{R}^d)), \\ v(0,\cdot) = v^0(\cdot). \end{cases} \tag{5.19}$$

We will show that, under additional hypotheses, either over a or v^0, we can guarantee the well-posedness of this problem. In particular, we provide the results regarding the cases of local advection ($\partial_I a = 0$) and non-local advection ($\partial_I a \neq 0$). We will see that this distinction directly affects the set of initial data v^0 for which the existence of solutions is guaranteed.

5.2.1 Some bounds over the characteristics

Consider a satisfying (5.5) and ψ_a satisfying (5.15) for some $k \geqslant 1$. For all $y \in \mathbb{R}^d$ and $u \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$ we define the characteristic curve $t \mapsto X_u(t,y)$ as the unique solution the following ODE

$$\begin{cases} \dot{X}_u(t,y) = a(t, X_u(t,y), (I_a u)(t, X_u(t,y))) =: \mathscr{A}_u(t, X_u(t,y)), \quad t \in [0,T], \\ X_u(0,y) = y, \end{cases} \tag{5.20}$$

where $(I_a u)(t,x)$ is defined in (5.12). Since the function ψ_a belongs to $W^{1,\infty}\left([0,T] \times \mathbb{R}^d_x, L^\infty(\mathbb{R}^d_y)\right)$, then $(I_a u)(t,x)$ belongs to $W^{1,\infty}\left([0,T] \times \mathbb{R}^d_x\right) \subset \mathscr{C}^{0,1}([0,T] \times \mathbb{R}^d_x)$. The regularity of a then implies that $\mathscr{A}_u(t,x)$ is a Lipschitz function with respect to the x variable, uniformly with respect to t, guaranteeing this way the global existence of solution for (5.20).

For all $u \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$, we define the norms

$$\|u\|_1 := \|u\|_{L^1([0,T]\times\mathbb{R}^d)} = \int\limits_0^T \int\limits_{\mathbb{R}^d} |u(t,x)| dx dt \text{ and } \|u\| := \sup_{t \in [0,T]} \|u(t,\cdot)\|_{L^1(\mathbb{R}^d)}.$$

We present some results involving the characteristics. The proofs for such results are given in Appendix 5.6.1.

The first property we describe is the continuity of the family of characteristics with respect to the spatial variables and the function u.

Lemma 5.1. *Let ψ_a satisfy (5.15) and a satisfy (5.5) for $k = 0$. Consider $y_1, y_2 \in \mathbb{R}^d$ and $u_1, u_2 \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$. Then there exists a positive constant $C(T, \|u_1\|, \|u_2\|)$, such that the solutions X_{u_1} and X_{u_2} of (5.20) satisfy for any $t \in [0,T]$,*

$$\sum_{j=1}^{d} |X^j_{u_1}(t, y_1) - X^j_{u_2}(t, y_2)| \leqslant C(T, \|u_1\|, \|u_2\|) \left(|y_1 - y_2| + \|u_1 - u_2\|_1\right).$$

Secondly, we claim that the spatial derivatives of the characteristics remain bounded by a constant only depending on T and $\|u\|$. We also claim that the spatial derivatives are continuous with respect to the spatial variables and the function u.

Lemma 5.2. *Let a satisfy (5.5) and ψ_a satisfy (5.15) for some $k \geqslant 1$. Consider $u \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$. Then there exists a positive constant $C(T, \|u\|)$, such that the solution $X_u(t, y)$ of (5.20) satisfies, for all $t \in [0,T]$ and $y \in \mathbb{R}^d$,*

$$\sum_{i=1}^{k} \sum_{|\alpha| \leqslant i} \sum_{j=1}^{d} |\partial_y^\alpha X_u^j(t, y)| \leqslant C(T, \|u\|). \tag{5.21}$$

Furthermore, for any two points $y_1, y_2 \in \mathbb{R}^d$, and any two functions $u_1, u_2 \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$, there exists a positive constant $C_2(T, \|u_1\|, \|u_2\|)$, such that the solutions X_{u_1} and X_{u_2} of (5.20) satisfy for any $t \in [0,T]$

$$\sum_{i=0}^{k} \sum_{|\alpha| \leqslant i} \sum_{j=1}^{d} |\partial_y^\alpha X^j_{u_1}(t, y_1) - \partial_y^\alpha X^j_{u_2}(t, y_2)| \leqslant C(T, \|u_1\|, \|u_2\|)(|y_1 - y_2| + \|u_1 - u_2\|_1). \tag{5.22}$$

Since for all $t \in [0,T]$, $y \mapsto X_u(t, y)$ is a $\mathscr{C}^1$-diffeomorphism from $\mathbb{R}^d$ onto itself, we may define its inverse as the function satisfying $X_u(t, X_u^{-1}(t, x)) = x$ for all $(t, x) \in [0,T] \times \mathbb{R}^d$. We have the following results for X_u^{-1}.

Lemma 5.3. *Let a satisfy (5.5) and ψ_a satisfy (5.15) for some $k \geqslant 1$. Consider $u \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$. Then there exists a positive constant $\widetilde{C}(T, \|u\|)$, such that the inverse of the solution $X_u(t, y)$ of (5.20) satisfies, for all $t \in [0,T]$, $x \in \mathbb{R}^d$*

$$\sum_{i=1}^{k} \sum_{|\alpha| \leqslant i} \sum_{j=1}^{d} |\partial_x^\alpha \left(X_u^{-1}\right)^j (t, x)| \leqslant \widetilde{C}(T, \|u\|). \tag{5.23}$$

Lemma 5.4. *Let a satisfy (5.5) and ψ_a satisfy (5.15) for some $k \geqslant 1$. Consider any two functions $u_1, u_2 \in \mathscr{C}([0,T], L^1(\mathbb{R}^d))$. Then there exists a positive constant $\widetilde{\widetilde{C}}(T, \|u_1\|, \|u_2\|)$, which satisfies $\lim_{T \to 0} \widetilde{\widetilde{C}}(T, \|u_1\|, \|u_2\|) = 0$ and such that the inverses of the solutions X_{u_1} and X_{u_2} of (5.20) satisfy, for all $t \in [0,T]$ and $x \in \mathbb{R}^d$*

$$\sum_{i=0}^{k-1} \sum_{|\alpha| \leqslant i} \sum_{j=1}^{d} | \left(\partial_x^\alpha X_{u_1}^{-1}\right)^j (t, x) - \partial_x^\alpha \left(X_{u_2}^{-1}\right)^j (t, x)| \leqslant \widetilde{\widetilde{C}}(T, \|u_1\|, \|u_2\|) \|u_1 - u_2\|_1. \tag{5.24}$$

We remark that thanks to the relation $\|u\|_1 \leqslant T\|u\|$, the relations (5.22) and (5.24) also hold true when replacing $\|u_1 - u_2\|_1$ by $\|u_1 - u_2\|$. Lastly, we give a result regarding the regularity of $X_u(t, x)$ with respect to t.

Lemma 5.5. *Let a satisfy (5.5) and ψ_a satisfy (5.15) for some $k \geqslant 1$. Consider $u \in \mathscr{C}^1([0,T_1)\times\mathbb{R}^d)$ such that $\sup_{t\in[0,T_1)} (\|u(t,\cdot)\|_{L^1(R^d)} + \|\partial_t u(t,\cdot)\|_{L^1(R^d)}) < +\infty$. Then $X_u(t,y) \in \mathscr{C}^1([0,T_1],\mathscr{C}^k(\mathbb{R}^d))$. As a consequence, $X_u^{-1}(t,x) \in \mathscr{C}^1([0,T_1],\mathscr{C}^k(\mathbb{R}^d))$.*

Proof. Thanks to Lemma 5.2, we know that under these hypotheses, for all $T < T_1$, $X_u(t,y)$ exists and belongs to $\mathscr{C}^1([0,T],\mathscr{C}^k(\mathbb{R}^d))$. Consider $0 < t_1, t_2 < T_1$, then

$$\begin{aligned}\sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\partial_y^\alpha X_u^j(t_1,y)-\partial_y^\alpha X_u^j(t_2,y)| &= \sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\int_{t_1}^{t_2}\partial_y^\alpha \dot{X}_u^j(s,y)ds| \\ &= \sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\int_{t_1}^{t_2}\partial_y^\alpha a_j(s,X_u(s,y),(I_a u)(s,X_u(s,y)))ds|.\end{aligned}$$

Thanks to the regularity of a and ψ_a, the bounds given in Lemma 5.2 for the derivatives of $X_u(t,y)$ and the uniform bound for $\|u\|_{L^1(\mathbb{R}^d)}$ we conclude that there exists a positive constant such that

$$\sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\partial_y^\alpha X_u^j(t_1,y)-\partial_y^\alpha X_u^j(t_2,y)| \leqslant C|t_1-t_2|.$$

Similarly, we have

$$\sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\partial_y^\alpha \dot{X}_u^j(t_1,y)-\partial_y^\alpha \dot{X}_u^j(t_2,y)| = \sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\partial_y^\alpha(\mathscr{A}_u)_j(t_1,y)-\partial_y^\alpha(\mathscr{A}_u)_j(t_2,y)|,$$

where

$$(\mathscr{A}_u)_j(t,y) := a_j(t,X_u(t,y),(I_a u)(t_1,X_u(t,y))).$$

Again, the regularity up to order $k+1$ of the involved coefficients allow us to conclude that

$$\sum_{i=0}^{k}\sum_{|\alpha|\leqslant i}\sum_{j=1}^{d}|\partial_y^\alpha \dot{X}_u^j(t_1,y)-\dot{X}_u^j(t_1,y)| \leqslant C|t_1-t_2|.$$

We have shown that $X_u(t,\cdot)$ is a Cauchy sequence in $\mathscr{C}^k(\mathbb{R}^d)$ when t goes to T_M, therefore, there exists $X^*(x) \in \mathscr{C}^k(\mathbb{R}^d)$ and $Y^*(x) \in \mathscr{C}^k(\mathbb{R}^d)$ such that

$$\lim_{t\to T_M}\|X_u(t,\cdot)-X^*\|_{\mathscr{C}^k(\mathbb{R}^d)} + \|\dot{X}_u(t,\cdot)-Y^*\|_{\mathscr{C}^k(\mathbb{R}^d)},$$

which is the desired result. □

5.2.2 Existence of solution for smooth initial data

We first provide the proof of existence and uniqueness of solution for problem (5.19) when the initial condition v^0 is a smooth enough function. We still assume hypotheses (5.5) through (5.18) hold. For a smooth initial condition v^0 we denote by classical solution any function $v \in \mathscr{C}^1([0,T]\times\mathbb{R}^d)$ which satisfies problem (5.19).
The following a priori estimate will allow us to guarantee the global existence of a classical solution given that such solution exists over a certain interval $[0,T_1]$.

Lemma 5.6. *Let $v^0 \in \mathscr{C}^1_c(\mathbb{R}^d)$ and $T_1 > 0$ be such that a classical solution $v \in \mathscr{C}^1([0,T]\times\mathbb{R}^d)$ exists for problem (5.19) for all $T < T_1$, which is positive and has compact support with respect to the x variable. Then, such solution satisfies the estimate*

$$\sup_{t\in[0,T_1]} \|v(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}. \tag{5.25}$$

Proof. Let v be the aforementioned positive solution. Denoting $\rho(t) := \|v(t,\cdot)\|_{L^1(\mathbb{R}^d)}$, we see directly from equation (5.19) that

$$\dot\rho(t) = \int\limits_{\mathbb{R}^d} \left(R(t,y,(I_g v)(t,y)) + \int_{\mathbb{R}^d} m(t,x,y,(I_d v)(t,x))dx \right) v(t,y)dy.$$

If there exists t such that $\rho(t) > \frac{I^*}{\underline{\psi_g}}$, then $(I_g v)(t,y) \geqslant \underline{\psi_g}\rho(t) > I^*$, which allows us to use hypothesis (5.18) in order to conclude

$$\dot\rho(t) < -r^*\rho(t) < 0.$$

This way, we see that either $\rho(t)$ is smaller than $\frac{I^*}{\underline{\psi_g}}$ or $\rho(t)$ is decreasing, which in turn implies the bound (5.25). □

A fixed point argument together with estimate (5.25) will allow us to conclude the existence of solution for problem (5.19) for smooth initial conditions.

Theorem 5.1. *Consider $k \geqslant 1$ and $T > 0$. Under hypotheses (5.15) through (5.18), for all non-negative functions $v^0 \in \mathscr{C}^k_c(\mathbb{R}^d)$, there exists a unique non-negative classical solution $v \in \mathscr{C}^1([0,T],\mathscr{C}^k_c(\mathbb{R}^d))$ to problem (5.19). Furthermore, such solution satisfies*

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}, \tag{5.26}$$

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{W^{k,1}(\mathbb{R}^d)} \leqslant C_T\|v^0\|_{W^{k,1}(\mathbb{R}^d)}. \tag{5.27}$$

Proof. Consider $v^0 \in \mathscr{C}^k_c(\mathbb{R}^d)$. For $t \geq 0$ and the function M introduced in (5.8) we define $r_t := 2\|a\|_{L^\infty} t$, B_{r_t} the open ball centred at 0 and of radius r_t,

$$O_t := \text{supp }(v^0) \cup \text{supp }(M) + B_{r_t},$$

and for $\alpha > 1$ we define $\overline{\rho}_\alpha := \max\{\alpha\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}$. We consider

$$u \in D^T_\alpha := \left\{ u \in \mathscr{C}([0,T]\times\mathbb{R}^d) : u \geqslant 0, \text{supp }(u(t,\cdot)) \subset O_t, \int_{\mathbb{R}^d} u(t,x)dx \leqslant \overline{\rho}_\alpha,\ \forall t \in [0,T] \right\}.$$

We denote as Φu the mapping defined by $v = \Phi u$, where v is the solution of

$$\begin{cases} \partial_t v(t,x) + \nabla_x \cdot (a(t,x,(I_a u)(t,x))v(t,x)) - R(t,x,(I_g u)(t,x))v(t,x) = \int_{\mathbb{R}^d} m(t,x,y,(I_d u)(t,x))u(t,y)dy \\ v(0,\cdot) = v^0(\cdot). \end{cases}$$

Let us denote, for any $y \in \mathbb{R}^d$, as $X_u(\cdot, y)$ the unique solution of

$$\begin{cases} \dot{X}_u(t,y) = a(t, X_u(t,y), (I_a u)(t, X_u(t,y))) =: \mathscr{A}_u(t, X_u(t,y)) \quad t \geq 0, \\ X(0,y) = y. \end{cases}$$

As stated before, for any $t \in [0,T]$, $x \mapsto X_u(t,x)$ is a $\mathscr{C}^1$-diffeomorphism, therefore, for all $t \geq 0$ and all $x \in \mathbb{R}^d$ there exists a unique $y \in \mathbb{R}^d$ such that $x = X_u(t,y)$, which we denote $y = X_u^{-1}(t,x)$.
We see that,

$$\begin{aligned} \frac{d}{dt} v(t, X_u(t,y)) =& \big[R(t, X_u(t,y), (I_g u)(t, X_u(t,y))) - \operatorname{div} \mathscr{A}_u(t, X_u(t,y))\big] v(t, X_u(t,y)) \\ &+ \int_{\mathbb{R}^d} m(t, X_u(t,y), z, (I_d u)(t, X_u(t,y))) u(t,z) dz, \end{aligned}$$

and thus, denoting

$$\mathscr{G}_u(t,y) := R(t, X_u(t,y), (I_g u)(t, X_u(t,y))) - \operatorname{div} \mathscr{A}_u(t, X_u(t,y)),$$

we get

$$\begin{aligned} \Phi u(t, X_u(t,y)) =& v(t, X_u(t,y)) \\ =& v^0(y) \exp\left(\int_0^t \mathscr{G}_u(s,y) ds\right) \\ &+ \int_0^t \int_{\mathbb{R}^d} m(s, X_u(s,y), z, (I_d u)(s, X_u(s,y))) u(s,z) dz \exp\left(\int_s^t \mathscr{G}_u(\tau, y) d\tau\right) ds \end{aligned}$$

or, equivalently

$$\begin{aligned} \Phi u(t,x) =& v(t,x) \\ =& v^0(X_u^{-1}(t,x)) \exp\left(\int_0^t \mathscr{G}_u(s, X_u^{-1}(t,x)) ds\right) \\ &+ \int_0^t \int_{\mathbb{R}^d} m(s, X_u(s, X_u^{-1}(t,x)), z, (I_d u)(s, X_u(s, X_u^{-1}(t,x)))) u(s,z) dz \exp\left(\int_s^t \mathscr{G}_u(\tau, X_u^{-1}(t,x)) d\tau\right) ds \end{aligned}$$

The solution v is thus non-negative, according to the non-negativity of v^0, u and m. Thanks to Lemma 5.3, v belongs to $\mathscr{C}^1([0,T], \mathscr{C}^\kappa(\mathbb{R}^d)) \subset \mathscr{C}([0,T] \times \mathbb{R}^d)$. Furthermore, the fact that for all $t \in [0,T]$, $|X(t,y) - y| \leqslant \|a\|_{L^\infty} t$ implies that supp $(v(t,\cdot)) \subset O_t$.
Additionally, directly from its definition, we see that

$$\mathscr{G}_u(s, X_u^{-1}(t,x)) \leqslant \gamma := \|R\|_{L^\infty_{t,x,I}} + \|a\|_{W^{1,\infty}_x L^\infty_{t,I}} + \|a\|_{W^{1,\infty}_I L^\infty_{t,x}} \|\psi_a\|_{W^{1,\infty}_x L^\infty_y} \overline{\rho}_\alpha,$$

and consequently, for all $u \in D_\alpha^T$, we have

$$\begin{aligned} \|\Phi u(t,\cdot))\|_{L^1(\mathbb{R}^d)} \leqslant& e^{\gamma T} \left(\int_{\mathbb{R}^d} v^0(X_u^{-1}(t,x)) dx \right. \\ &\left. + \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} m(s, X_u(s, X_u^{-1}(t,x)), z, (I_d u)(s, X_u(s, X_u^{-1}(t,x)))) u(s,z) dz dx ds \right) \\ \leqslant& e^{\gamma T} \left(\int_{\mathbb{R}^d} v^0(X_u^{-1}(t,x)) dx + \int_0^t \int_{\mathbb{R}^d} M(s, X_u(s, X_u^{-1}(t,x))) dx \int_{\mathbb{R}^d} u(s,z) dz ds \right). \end{aligned}$$

Making the changes of variables $y = X_u^{-1}(t,x)$ and $y = X_u(s, X_u^{-1}(t,x))$ respectively on each of the integrals on the last expression, recalling that, according to Liouville's formula

$$|J_{X_u(t,y)}| = e^{\int_0^t \operatorname{div} \mathscr{A}_u(s,X_u(t,y))ds},$$
$$|J_{X_u(t,X_u^{-1}(s,y))}| = e^{\int_0^t \operatorname{div} \mathscr{A}_u(\tau,X_u(t,X_u^{-1}(s,y)))d\tau - \int_0^s \operatorname{div} \mathscr{A}_u(\tau,y)d\tau},$$

and using the hypotheses over a and m we obtain that for all $t \in [0,T]$,

$$\|\Phi u(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant e^{(\gamma+2\tilde{a})T} \left(\|v^0\|_{L^1(\mathbb{R}^d)} + KT\|u\|_{L^1(\mathbb{R}^d)}\right),$$

where $\tilde{a} := \|a\|_{W_x^{1,\infty}L_{t,I}^\infty} + \|a\|_{W_I^{1,\infty}L_{t,x}^\infty}\|\psi_a\|_{W_x^{1,\infty}L_y^\infty}\overline{\rho}_\alpha$. Finally, using the hypothesis over v^0 and u we see that

$$\|\Phi u(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant e^{(\gamma+2\tilde{a})T} \left(\frac{1}{\alpha} + KT\right) \overline{\rho}_\alpha.$$

Thanks to the condition $\alpha > 1$ there exists T_α (only depending on α and on the coefficients of the problem) such that $\|\Phi u\|_{L^1(\mathbb{R}^d)} \leqslant \overline{\rho}_\alpha$ for all $t \in [0,T_\alpha]$. In other words, $\Phi : D_\alpha^{T_\alpha} \to D_\alpha^{T_\alpha}$.

We now claim that the mapping Φu is a contraction on some $D_\alpha^{T_1}$, $0 < T_1 \leqslant T_\alpha$, with respect to the usual norm in $\mathscr{C}([0,T]\times\mathbb{R}^d)$. For two functions $u_1, u_2 \in D_\alpha^{T_\alpha}$ and any $t \in [0,T_\alpha]$, thanks to Lemma 5.1 and Lemma 5.4, we have

$$\begin{aligned}|\Phi u_1 - \Phi u_2| \leqslant& Ce^{\gamma t}|X_{u_1}^{-1}(t,x) - X_{u_2}^{-1}(t,x)| \\ \leqslant& Ce^{\gamma t}\|u_1 - u_2\|_1 \\ \leqslant& Ce^{\gamma t}|O_t||t\|u_1 - u_2\|_{\mathscr{C}([0,t]\times\mathbb{R}^d)}.\end{aligned}$$

Clearly, for $t = T_1$ small enough, Φu is a contraction, and therefore, thanks to the Banach fixed point theorem, there exists a unique $v \in D_\alpha^{T_1}$ such that $\Phi v = v$. Such v is a solution of (5.19) over $[0,T_1]$. Furthermore, directly from the relation $v = \Phi v$ we see that $v \in \mathscr{C}^1([0,T],\mathscr{C}_c^\kappa(\mathbb{R}^d))$.

Let us now assume that there exists T_M, a finite maximal time such that a solution exists in B_α^T for all $T < T_M$, and let v be such solution. Directly from Lemma 5.6, the solution v satisfies estimate (5.25) over $[0,T_M)$. Furthermore, thanks to the relation $v = \Phi v$, we are able to show that

$$\sup_{[0,T_M)} \left(\|v(t,\cdot)\|_{L^1(\mathbb{R}^d)} + \|\dot{v}(t,\cdot)\|_{L^1(\mathbb{R}^d)}\right) < +\infty.$$

From Lemma 5.5 we get then that $X_v^{-1}(t,x) \in \mathscr{C}^1([0,T_M],\mathscr{C}^k(\mathbb{R}^d))$, and by composition of functions, so is $v = \Phi v$. We can then iterate the previous ideas using $v(T_M) \in \mathscr{C}^k(\mathbb{R}^d)$ as a starting point in order to obtain the existence of solution over a certain interval $[T_M, T_M+\delta)$, contradicting this way the maximal character of T_M. Hence, there exists a classical solution of (5.19) for all $t > 0$.

The uniqueness on $\mathscr{C}([0,T],\mathscr{C}_c^\kappa(\mathbb{R}^d))$ comes from the fact that every other solution on D_α^∞ will coincide with v at least over a small interval $(0,t_0)$ and then, by continuity, the same would hold for all t.

To obtain the $W^{k,1}(\mathbb{R}^d)$ estimates, we differentiate the relation $v = \Phi v$ and notice that for all multi-index β such that $|\beta| \leqslant k$

$$\partial_x^\beta v = \exp\left(\int_0^t \mathscr{G}_v(s, X_v^{-1}(t,x))ds\right) \sum_{|\gamma|\leqslant|\beta|} \partial_x^\gamma v^0(X_v^{-1}(t,x))F_1^\gamma(t,x) + \int_0^t \int_{\mathbb{R}^d} F_2(s,x,y)v(s,y)dy \exp\left(\int_s^t \mathscr{G}_v(\tau,x)d\tau\right) ds$$

where the functions F_1^γ and F_2 are combinations of sums and multiplications of the derivatives up to order $|\beta|$ of X_v^{-1}, R, m, I_g and I_d. Taking absolute values, integrating over $\mathbb{R}^d$ and using the boundedness of all the involved coefficients we arrive at

$$\|v(t,\cdot)\|_{W^{k,1}(\mathbb{R}^d)} \leqslant C_T^1 \|v^0\|_{W^{k,1}(\mathbb{R}^d)} + C_T^2 \int_0^t \|v(s,\cdot)\|_{W^{k,1}(\mathbb{R}^d)} ds$$

and thanks to Grönwall's lemma we get (5.27). □

5.2.3 Existence of solution for more general initial data

Depending on whether $\partial_I a = 0$ or $\partial_I a \neq 0$, we will have a different class of initial data for which we are able to guarantee existence of solution for problem (5.19). Furthermore, the regularity of such solution might also be affected.
We first prove that, when $\partial_I a \neq 0$, a solution exists (in a sense that will be defined below) for any initial condition $v^0 \in W^{k,\infty}(\mathbb{R}^d)$, with compact support. However, we do not prove that the regularity of the solution is preserved over time, even if we did not manage to highlight the existence of cases where a loss of regularity is observed. Secondly, we will show that, when $\partial_I a = 0$, not only the set of initial conditions for which we can claim existence of solution is more general ($v^0 \in W^{k,1}(\mathbb{R}^d)$), but the regularity of such solution is preserved for all $t > 0$.
We introduce the definition of weak solution for problem (5.19). We say that v is a weak solution of problem (5.19) associated to $v^0 \in L^p(\mathbb{R}^d)$ if

$$v \in L^\infty([0,T], L^p(\mathbb{R}^d)),$$

and it satisfies the equation in the following weak sense

$$\int_0^T \int_{\mathbb{R}^d} v L_v^* \varphi dx dt = \int_{\mathbb{R}^d} v^0 \varphi dx,$$

for any $\varphi \in \mathscr{C}_c^1([0,T) \times \mathbb{R}^d)$, where we define the operator L_v^* by

$$L_v^*\varphi(t,x) = -\partial_t\varphi(t,x) - a(t,x,(I_a v)(t,x))\cdot\nabla\varphi(t,x) - R(t,x,(I_g v)(t,x))\varphi(t,x) - \int_{\mathbb{R}^d} m(t,y,x,(I_d v)(t,y))\varphi(t,y)dy,$$

for all $t \in [0,T]$, $y \in \mathbb{R}^d$. We remark that a classical solution is always a weak solution.

Theorem 5.2. *Under hypotheses (5.5) through (5.18), for all $k \geqslant 1$ and any non-negative functions $v^0 \in W^{k,\infty}(\mathbb{R}^d)$ with compact support, there exists a unique non-negative weak solution $v \in \mathscr{C}([0,T], \mathscr{C}_c^{k-1}(\mathbb{R}^d))$ to problem (5.19). Furthermore, such solution satisfies*

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\psi_g}\}, \tag{5.28}$$

$$\sup_{t\in[0,T]} \|v(t,\cdot)\|_{W^{k-1,1}(\mathbb{R}^d)} \leqslant C_T \|v^0\|_{W^{k-1,1}(\mathbb{R}^d)}, \tag{5.29}$$

and, for $k \geqslant 2$, $v \in \mathscr{C}^1([0,T], \mathscr{C}_c^{k-1}(\mathbb{R}^d))$.

Proof. Directly from Morrey's inequality, we get the relation $v^0 \in W^{k,\infty}(\mathbb{R}^d) \subset \mathscr{C}^{k-1,1}(\mathbb{R}^d)$. If $k \geqslant 2$, we are able to apply Theorem 5.1 in order to get the desired result.
Consider now $k = 1$. The compact support of v^0 and the $W^{1,\infty}(\mathbb{R}^d)$ regularity imply that $v^0 \in W^{k,1}(\mathbb{R}^d)$. This means that, there exists a sequence of compactly supported functions $v^0_\varepsilon \in \mathscr{C}^1_c(\mathbb{R}^d)$ such that

$$\begin{gathered}\text{supp }(v^0)_\varepsilon \subset \text{ supp }(v^0),\\ \|v^0_\varepsilon\|_{W^{1,\infty}(\mathbb{R}^d)} \leqslant \|v^0\|_{W^{1,\infty}(\mathbb{R}^d)},\\ \lim_{\varepsilon\to 0}\|v^0 - v^0_\varepsilon\|_{W^{1,1}(\mathbb{R}^d)} = 0.\end{gathered}$$

We denote as v_ε the solution of problem (5.19) associated to v^0_ε, and we claim that v_ε is a Cauchy sequence in $\mathscr{C}([0,T], L^1(\mathbb{R}^d))$.
We recall that for all ε,

$$\sup_{t\in[0,T]} \|v_\varepsilon\|_{L^1(\mathbb{R}^d)} \leqslant \overline{\rho}_\varepsilon := \max\{\|v^0_\varepsilon\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}.$$

Furthermore, the equality $v_\varepsilon = \Phi v_\varepsilon$ holds true, where Φ was defined on the proof of Theorem 5.1. Consequently, for $\varepsilon_1, \varepsilon_2 > 0$ we have the relation

$$\begin{aligned}\Delta V_{\varepsilon_1\varepsilon_2} :=& v_{\varepsilon_1} - v_{\varepsilon_2}\\ =& \Phi v_{\varepsilon_1} - \Phi v_{\varepsilon_2}\\ =& \Big(v^0_{\varepsilon_1}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_2}}(t,x))\Big)\exp\Big(\int_0^t \mathscr{G}_1(\tau,t,x)d\tau\Big) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_2}}(t,x))\Big(\Delta E(0,t,x)\Big)\\ &+ \int_0^t\int_{\mathbb{R}^d}\Big(\mathscr{M}_1(s,t,x,z) - \mathscr{M}_2(s,t,x,z)\Big)v_{\varepsilon_1}(s,z)dz\exp\Big(\int_s^t \mathscr{G}_1(\tau,t,x)d\tau\Big)ds\\ &+ \int_0^t\int_{\mathbb{R}^d}\mathscr{M}_2(s,t,x,z)\Big(\Delta V_{\varepsilon_1\varepsilon_2}(s,z)\Big)dz\exp\Big(\int_s^t \mathscr{G}_1(\tau,t,x)d\tau\Big)ds\\ &+ \int_0^t\int_{\mathbb{R}^d}\mathscr{M}_2(s,t,x,z)v_{\varepsilon_j}(s,z)dz\Big(\Delta E(s,t,x)\Big)ds,\end{aligned}$$

where

$$\begin{aligned}\mathscr{G}_i(\tau,t,x) &= \mathscr{G}_{v_{\varepsilon_i}}(\tau, X^{-1}_{v_{\varepsilon_i}}(t,x)),\\ \Delta E(s,t,x) &:= \exp\left(\int_s^t \mathscr{G}_1(\tau,t,x)d\tau\right) - \exp\left(\int_s^t \mathscr{G}_2(\tau,t,x)d\tau\right),\\ \mathscr{M}_i(s,t,x,z) &:= m(s, X_{v_{\varepsilon_i}}(s, X^{-1}_{v_{\varepsilon_i}}(t,x)), z, (I_d v_{\varepsilon_i})(s, X_{v_{\varepsilon_i}}(s, X^{-1}_{v_{\varepsilon_i}}(t,x)))).\end{aligned}$$

We write

$$\begin{aligned}v^0_{\varepsilon_1}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_2}}(t,x)) \quad = \quad & v^0_{\varepsilon_1}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_1}}(t,x))\\ &+ v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_2}}(t,x)).\end{aligned} \tag{5.30}$$

Thanks to the change of variables $y = X^{-1}_{v_{\varepsilon_1}}(t,x)$ and the relations

$$\begin{aligned}|J_{X_{v_{\varepsilon_1}}(t,y)}| =& e^{\int_0^t \text{div } \mathscr{A}_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}(t,y))ds},\\ \mathscr{G}(s, X^{-1}_{v_{\varepsilon_1}}(t,x)) \leqslant& \gamma := \|R\|_{L^\infty_{t,x,I}} + \|a\|_{W^{1,\infty}_x L^\infty_{t,I}} + \|a\|_{W^{1,\infty}_I L^\infty_{t,x}}\|\psi_a\|_{W^{1,\infty}_x L^\infty_y}\overline{\rho},\\ \text{div } \mathscr{A}_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}(t,y)) \leqslant& \tilde{a} := \|a\|_{W^{1,\infty}_x L^\infty_{t,I}} + \|a\|_{W^{1,\infty}_I L^\infty_{t,x}}\|\psi_a\|_{W^{1,\infty}_x L^\infty_y}\overline{\rho},\end{aligned}$$

we conclude that

$$\begin{aligned}
&\int\limits_{\mathbb{R}^d} \left(v^0_{\varepsilon_1}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_1}}(t,x)) \right) \exp\left(\int_0^t \mathscr{G}_1(\tau,t,x) d\tau \right) dx \\
=&\int\limits_{\mathbb{R}^d} \left(v^0_{\varepsilon_1}(y) - v^0_{\varepsilon_2}(y) \right) \exp\left(\int_0^t \mathscr{G}_{v_{\varepsilon_i}}(\tau,y) d\tau \right) |J^{-1}_{X_{v_{\varepsilon_1}}(t,y)}| dy \\
\leqslant & e^{(\gamma+\tilde{\alpha})T} \|v^0_{\varepsilon_1} - v^0_{\varepsilon_2}\|_{L^1(\mathbb{R})}.
\end{aligned} \tag{5.31}$$

On the other hand, the compactness of the support of $v^0_{\varepsilon_2}$, together with the relation

$$|X_{v_\varepsilon}(t,y) - y| \leqslant \|a\|_{L^\infty} t,$$

implies that

$$v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_2}}(t,x)) = 0, \text{ for all } x \notin O_t := \text{supp } v^0 + B_{r_t},$$

where B_{r_t} is the ball of radius $\|a\|_{L^\infty} t$. Hence, we obtain

$$\begin{aligned}
&\int\limits_{\mathbb{R}^d} \left(v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_1}}(t,x)) - v^0_{\varepsilon_2}(X^{-1}_{v_{\varepsilon_2}}(t,x)) \right) \exp\left(\int_0^t \mathscr{G}_1(\tau,t,x) d\tau \right) dx \\
\leqslant & |O_t| \|v^0_{\varepsilon_2}\|_{W^{1,\infty}(\mathbb{R}^d)} e^{\gamma T} |X^{-1}_{v_{\varepsilon_1}}(t,x) - X^{-1}_{v_{\varepsilon_2}}(t,x)| \\
\leqslant & \widetilde{C}_T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1,
\end{aligned} \tag{5.32}$$

where we have used (5.24) on the second line[1].
From the definition of $\mathscr{G}_u(t,x)$, we observe that

$$\begin{aligned}
|\Delta E(s,t,x)| \leqslant & e^{\gamma T} \int_s^t |\mathscr{G}_1(\tau,t,x) - \mathscr{G}_2(\tau,t,x)| d\tau \\
\leqslant & \widetilde{C}_T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1,
\end{aligned} \tag{5.33}$$

where $\widetilde{C}_T$ depends on $\overline{\rho}$, T, the derivatives of R, a and ψ_a, and on the constant appearing in (5.24). Using the change of variables $y = X_{v_{\varepsilon_2}}(s, X^{-1}_{v_{\varepsilon_2}}(t,x))$ and recalling that

$$|J_{X_{v_{\varepsilon_2}}(t, X^{-1}_{v_{\varepsilon_2}}(s,x))}| = e^{\int_0^t \text{div } \mathscr{A}_u(\tau, X_{v_{\varepsilon_2}}(t, X^{-1}_{v_{\varepsilon_2}}(s,x)))d\tau - \int_0^s \text{div } \mathscr{A}_{v_{\varepsilon_2}}(\tau,x)d\tau},$$

we see that

$$\begin{aligned}
\int\limits_{\mathbb{R}^d} \mathscr{M}_2(s,t,x,z) dx =& \int\limits_{\mathbb{R}^d} m(s,y,z,(I_d v_{\varepsilon_2})(s,y)) |J_{X_{v_{\varepsilon_2}}(t, X^{-1}_{v_{\varepsilon_2}}(s,y))}|^{-1} dy \\
\leqslant & e^{2\tilde{\alpha} T} \int\limits_{\mathbb{R}^d} \sup_{s,z,I} m_{\varepsilon_i}(s,y,z,I) dy \leqslant e^{2\tilde{\alpha} T} |\mathscr{K}| \overline{M}.
\end{aligned}$$

[1]This term is responsible for the possible loss of regularity for $t > 0$: In order to prove that v_ε is a Cauchy sequence in $\mathscr{C}([0,T], W^{1,1}(\mathbb{R}^d))$, we would need a $W^{2,\infty}(\mathbb{R}^d)$ estimate over v^0_ε, which we do not have.

Therefore, we have the bounds

$$\int_{\mathbb{R}^d} v^0(X^{-1}_{v_{\varepsilon_2}}(t,x))|\Delta E(0,t,x)|dx \leqslant e^{\tilde{\alpha}T}\|v^0\|_{L^1(\mathbb{R}^d)}\widetilde{C}_T\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_1, \tag{5.34}$$

$$\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathscr{M}_2(s,t,x,z)\Big(\Delta V_{\varepsilon_1\varepsilon_2}(s,z)\Big)dz\exp\left(\int_s^t \mathscr{G}_1(\tau,t,x)d\tau\right)dxds \leqslant e^{(\gamma+2\tilde{a})T}|\mathscr{K}|\overline{M}\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_1, \tag{5.35}$$

$$\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathscr{M}_2(t,x,z)v_{\varepsilon_j}(s,z)dz\Big(|\Delta E(s,t,x)|\Big)dxds \leqslant e^{2\tilde{a}T}|\mathscr{K}|\overline{M}\overline{\rho}T\widetilde{C}_T\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_1. \tag{5.36}$$

The function $m(t,x,z,I)$ having a compact support on the x variable, leads to

$$\mathscr{M}_1(s,t,x,z)-\mathscr{M}_2(s,t,x,z)=0, \text{ for all } x\notin\mathscr{K}+B_{2r_t}.$$

On the other hand, the function m being differentiable and Lipschitz, implies that, for all $x\in\mathscr{K}+B_{2r_t}$

$$\begin{aligned}|\mathscr{M}_1(s,t,x,z)-\mathscr{M}_2(s,t,x,z)| \leqslant&\|m\|_{W^{1,\infty}_x}|X_{v_{\varepsilon_1}}(s,X^{-1}_{v_{\varepsilon_1}}(t,x))-X_{v_{\varepsilon_2}}(s,X^{-1}_{v_{\varepsilon_2}}(t,x))|\\ &+\mu|(I_d v_{\varepsilon_1})(s,X_{v_{\varepsilon_1}}(s,X^{-1}_{v_{\varepsilon_1}}(t,x)))-(I_d v_{\varepsilon_2})(s,X_{v_{\varepsilon_2}}(s,X^{-1}_{v_{\varepsilon_2}}(t,x)))|\\ \leqslant&(\|m\|_{W^{1,\infty}_x}+\|\psi_d\|_{W^{1,\infty}_x}\overline{\rho})|X_{v_{\varepsilon_1}}(s,X^{-1}_{v_{\varepsilon_1}}(t,x))-X_{v_{\varepsilon_2}}(s,X^{-1}_{v_{\varepsilon_2}}(t,x))|\\ &+\|\psi_d\|_{L^\infty}\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Using first Lemma 5.1 and then Lemma 5.4, we conclude that there exists a constant $\widetilde{C}_T$ such that

$$\begin{aligned}|\mathscr{M}_1(s,t,x,z)-\mathscr{M}_2(s,t,x,z)| \leqslant&\widetilde{C}_T\Big(|X^{-1}_{v_{\varepsilon_1}}(t,x)-X^{-1}_{v_{\varepsilon_2}}(t,x)|+\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_1+\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)}\Big)\\ \leqslant&\widetilde{C}_T\Big(\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_1+\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)}\Big).\end{aligned}$$

Therefore, we have the bound

$$\begin{aligned}&\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big(\mathscr{M}_1(s,t,x,z)-\mathscr{M}_2(s,t,x,z)\Big)v_{\varepsilon_1}(s,z)dz\exp\left(\int_s^t\mathscr{G}_1(\tau,t,x)d\tau\right)dxds\\ \leqslant&|\mathscr{K}+B_{2r_T}|\overline{\rho}\widetilde{C}_Te^{\gamma T}\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_1.\end{aligned} \tag{5.37}$$

Putting together the bounds (5.31) through (5.37), we get

$$\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} \leqslant \widetilde{C}_T\left(\|v^0_{\varepsilon_1}-v^0_{\varepsilon_2}\|_{L^1(\mathbb{R})}+\int_0^T\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)}ds\right).$$

Thanks to Grönwall's lemma, we have then the relation

$$\sup_{t\in[0,T]}\|v_{\varepsilon_1}-v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} \leqslant \widetilde{C}_T\|v^0_{\varepsilon_1}-v^0_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)}$$

for some $\widetilde{C}_T$ independent of ε_1 and ε_2, which proves that, up to the extraction of a sub-sequence, v_ε is a Cauchy sequence in $\mathscr{C}([0,T],L^1(\mathbb{R}^d))$. Therefore, there exists $v \in \mathscr{C}([0,T],L^1(\mathbb{R}^d))$ such that

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,T]} \|v_\varepsilon - v\|_{L^1(\mathbb{R}^d)} = 0.$$

Furthermore, such function satisfies the bounds (5.25) and (5.27).
We claim now that the sequence $L^*_{v_\varepsilon}\varphi$ converges to $L^*_v\varphi$ in $L^\infty([0,T]\times\mathbb{R}^d)$ for all $\varphi\in\mathscr{C}^1_c([0,T)\times\mathbb{R}^d)$. This is a direct consequence of the relation

$$|L^*v_\varepsilon\varphi - L^*v_\varepsilon\varphi| \leqslant (L_r\|\psi_g\|_{L^\infty} + \mu\|\psi_d\|_{L^\infty})\|\varphi\|_{L^\infty(\mathbb{R}^d)}\|v^\varepsilon - v\|_{L^1(\mathbb{R}^d)}.$$

In order to conclude, we recall that all classical solutions are weak solutions, and therefore, for all $\varepsilon > 0$

$$\int_0^T\int_{\mathbb{R}^d} v_\varepsilon L^*_{v_\varepsilon}\varphi dxdt = \int_{\mathbb{R}^d} v^0_\varepsilon\varphi dx,$$

and taking the limit when ε goes to 0 we see that v is a weak solution of problem (5.19). □

Considering initial data with compact support might be enough in order to model most of the biological scenarios found in nature. However, the hypothesis $v^0 \in W^{k,\infty}(\mathbb{R}^d)$ might be too restrictive for some real life scenarios. Furthermore, the study of the problem when more general initial conditions are present, is of theoretical interest. We show below that, when $\partial_I a = 0$, a solution exists for any initial data $v^0\in W^{k,1}(\mathbb{R}^d)$, $k\geqslant 1$.

Theorem 5.3. *Under hypothesis (5.15) through (5.18), if $\partial_I a = 0$, for all non-negative functions $v^0 \in W^{k,1}(\mathbb{R}^d)$, there exists a unique non-negative weak solution $v\in\mathscr{C}([0,T],W^{k,1}(\mathbb{R}^d))$ of problem (5.19). Furthermore, such a solution satisfies*

$$\sup_{t\in[0,T]}\|v\|_{L^1(\mathbb{R}^d)} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi_g}}\}, \tag{5.38}$$

$$\sup_{t\in[0,T]}\|v\|_{W^{k,\infty}(\mathbb{R}^d)} \leqslant C_T\|v^0\|_{W^{k,1}(\mathbb{R}^d)}. \tag{5.39}$$

Proof. As in the proof for $k=1$ when $\partial_I a\neq 0$, we can approximate any function $v^0\in W^{k,1}$ by a smooth, compactly supported sequence of functions v^0_ε. The same arguments as in the previous proof will show that v_ε, the sequence of solutions associated to v^0_ε, is a Cauchy sequence in $\mathscr{C}([0,T],L^1(\mathbb{R}^d))$. Furthermore, given that the second term in (5.30), which is responsible for the possible loss of regularity in the previous case, is equal 0 when $\partial_I a = 0$, we show that v_ε is a Cauchy sequence in $\mathscr{C}([0,T],W^{k,1}(\mathbb{R}^d))$ as well. We prove as in Theorem 5.2 that the limit of v_ε is the required weak solution. □

Given that the regularity of the solution varies depending on whether $\partial_I a = 0$ or $\partial_I a\neq 0$, and that such regularity will be of importance in the upcoming sections, we define the parameter

$$\kappa := \begin{cases} k-1, & \text{if } \partial_I a \neq 0,\\ k, & \text{if } \partial_I a = 0,\end{cases}$$

which encompasses the information over said regularity.

We remark that if we had $\partial_I m = 0$, we might obtain existence for a larger class of mutation functions m. For conciseness, we will however not consider such cases in the present work.

5.3 Particle Method

The particle method basically consists in searching for an approximate solution of problem (5.19) which is a sum of weighted Dirac masses.
Throughout the following section we suppose

$$\psi_a \in \mathscr{C}\left([0,T]\times\mathbb{R}^d_x, W^{1,\infty}(\mathbb{R}^d)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}_y, \mathscr{C}^2(\mathbb{R}^d_x)\cap W^{2,\infty}(\mathbb{R}^d_x)\right), \tag{5.40}$$

$$0<\underline{\psi_g}\leqslant\psi_g \in \mathscr{C}\left([0,T]\times\mathbb{R}^d_x, W^{1,\infty}(\mathbb{R}^d_y)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}^d_y, \mathscr{C}^1(\mathbb{R}^d_x)\cap W^{1,\infty}(\mathbb{R}^d_x)\right) \tag{5.41}$$

$$\psi_d \in \mathscr{C}\left([0,T]\times\mathbb{R}^d_x, W^{1,\infty}(\mathbb{R}^d_y)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}^d_y, \mathscr{C}^1(\mathbb{R}^d_x)\cap W^{1,\infty}(\mathbb{R}^d_x).\right) \tag{5.42}$$

Notice that, unlike the set of hypotheses (5.15)-(5.17), we have imposed $W^{1,\infty}(\mathbb{R}^d)$ regularity for the y variable, which is needed in order to approximate the integral terms by sums over a countable set.
Consider as well

$$0\leqslant m\in\mathscr{C}\left([0,T]\times\mathbb{R}^d_x\times\mathbb{R}^d_y, W^{1,\infty}(\mathbb{R}_I)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}^d_x\times\mathbb{R}_I, W^{1,\infty}(\mathbb{R}^d_y)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}^d_y\times\mathbb{R}_I, \mathscr{C}^1_c(\mathbb{R}^d_x)\right) \tag{5.43}$$

satisfying hypotheses (5.7) through (5.18).
For $h>0$, consider a countable set of indices $\mathscr{J}_h\in\mathbb{Z}^d$, points $x_i^0\in\mathbb{R}^d$ and weights w_i^0 for $i\in\mathscr{J}_h$. The weights w_i^0 can be regarded as the respective masses of a collection of subsets $\Omega_i^0\subset\mathbb{R}^d$ satisfying

$$\Omega_i^0\cap\Omega_j^0=\emptyset,\ \text{if } i\neq j,\ \text{and}\ \bigcup_{i\in\mathscr{J}_h}\Omega_i^0=\mathbb{R}^d. \tag{5.44}$$

For example, we may choose the Ω_i^0 as the set of all non intersecting cubes of side length equal h having the points hi as centers , $i\in\mathbb{Z}^d$. This way, $w_i^0=h^d$, with each of the x_i^0 being a point in Ω_j^0. In general we assume that there exist positive constants c and C such that

$$ch\leqslant|x_i^0-x_j^0|\leqslant Ch,\ \forall i\neq j, \tag{5.45}$$

$$ch^d\leqslant w_i^0\leqslant Ch^d,\ \forall i\in\mathscr{J}_h. \tag{5.46}$$

Following [57], the particle method then consists in looking for a measure ν_h of the form

$$\nu_h(t)=\sum_{i\in\mathscr{J}_h}\nu_i(t)w_i(t)\delta_{x_i(t)},$$

where $(\nu:=\{v_i(t)\}_{i\in\mathscr{J}_h}, w:=\{w_i(t)\}_{i\in\mathscr{J}_h}, \overline{x}:=\{x_i(t)\}_{i\in\mathscr{J}_h})$, is the solution of the following system

$$\begin{cases} \dot{x}_i(t)=A_{\nu,w}(t,x_i),\\ \dot{w}_i(t)=\text{div }A_{\nu,w}(t,x_i(t))w_i(t),\\ \dot{\nu}_i(t)=\big(-\text{div }A_{\nu,w}(t,x_i(t))+R(t,x_i(t),I_g(t,x_i(t),\nu,w))\big)\nu_i(t)\\ \qquad+\sum_{j\in\mathscr{J}_h}w_j(t)\nu_j(t)m(t,x_i(t),x_j(t),I_d(t,x_i(t),\nu,w)),\\ x_i(0)=x_i^0,\ w_i(0)=w_i^0,\ \nu_i(0)=v^0(x_i^0), \end{cases} \tag{5.47}$$

where

$$A_{\nu,w}(t,x) = a(t,x,I_a(t,x,\nu,w)),$$

and

$$I_l(t,x,\nu,w) := \sum_{j\in\mathscr{J}_h} \nu_j(t)w_j(t)\psi_l(t,x,x_j(t)),$$

with $l \in \{a,g,d\}$.
In what follows we assume that h and x_k^0 are chosen in such a way that

$$\|v^0\|_{1,h} := \sum_{i\in\mathscr{J}_h} v^0(x_i^0)w_i^0 < \infty. \tag{5.48}$$

We define the subset of indices

$$\mathscr{J}_h^m := \{i \in \mathscr{J}_h : x_i^0 \in \text{supp } m + B_{\|a\|_{L^\infty}T}\},$$

where $B_{\|a\|_{L^\infty}T}$ is the ball of radius $\|a\|_{L^\infty}T$. The compact support of m implies that $|\mathscr{J}_h^m| < \infty$.

For a positive value of h and a set of indexes $\mathscr{J}_h$ we define the functional spaces

$$\ell^1(\mathscr{J}_h) := \big\{u = \{u_i\}_{i\in\mathscr{J}_h} : \sum_{i\in\mathscr{J}_h}|u_i| < +\infty\big\},$$
$$\ell^\infty(\mathscr{J}_h) := \big\{w = \{w_i\}_{i\in\mathscr{J}_h} : \sup_{i\in\mathscr{J}_h}|w_i| < +\infty\big\}.$$

We equip these spaces with the norms

$$\|u\|_{\ell^1} := \sum_{i\in\mathscr{J}_h}|u_i| \text{ and } \|w\|_{\ell^\infty} := \sup_{i\in\mathscr{J}_h}|w_i|$$

respectively. It is clear that for all $u \in \ell^1(\mathscr{J}_h)$ and $v \in \ell^\infty(\mathscr{J}_h)$, then $uv \in \ell^1(\mathscr{J}_h)$. For $T > 0$ we define as well the spaces

$$X_h^T := \mathscr{C}([0,T],\ell^1(\mathscr{J}_h)) \text{ and } Y_h^T := \mathscr{C}([0,T],\ell^\infty(\mathscr{J}_h)),$$

equipped with the norms

$$\|u\|_{1,h} := \sup_{t\in[0,T]}\|u(t)\|_{\ell^1} \text{ and } \|w\|_{\infty,h} := \sup_{t\in[0,T]}\|w(t)\|_{\ell^\infty}.$$

Problem (5.47) is a strongly coupled system of ODEs, with an infinite number of unknowns and equations. In some cases the system becomes uncoupled (for example if $\partial_I a = 0$) or with a finite number of equations and unknowns (for example if v^0, a and m have compact support), however, for the sake of generality, we present below the proof of existence of solution in the general case, and later discuss briefly these particular scenarios.
We start by giving two results that will be of great use for the proof of existence for problem (5.47). First, we deal with the existence of solution for a simpler system of infinite equations with infinitely many unknowns:

Lemma 5.7. *Consider $a \in \mathscr{C}([0,T],(W^{1,\infty}(\mathbb{R}^{d+1})))^d$, $u \in X_h^T$, $w \in Y_h^T$ and ψ_a satisfying hypothesis (5.40). Then there exists a unique family of functions $x := \{x_i\}_{i\in\mathscr{J}_h}$, $x_i \in \mathscr{C}^1([0,T])$ for all $i \in \mathscr{J}_h$ which is solution of the system of equations*

$$\dot{x}_i(t) = A_{u,w}(t,x_i), \quad t\in[0,T],\ x_i(0) = x_i^0. \tag{5.49}$$

When $\partial_I a = 0$, system (5.49) becomes uncoupled, each individual equation has a solution, thanks to the classic Cauchy-Lipschitz theory. The proof of the general case is given in Appendix 5.6.2.
The second auxiliary result comes from approximation theory, and it will also be of great use in Section 5.4:

Lemma 5.8.

$$\forall \varphi \in W^{k,1}(\mathbb{R}^d), \quad \left| \int_{\mathbb{R}^d} \varphi(x)dx - \sum_{i \in \mathscr{J}_h} w_i(t)\varphi(x_i(t)) \right| \leqslant Ch^k \|\varphi\|_{k,1},$$

where C is a constant which depends on a, ψ_a, $\|vw\|_{1,h}$ and T.

This result is a direct corollary of Lemma 8 in [184]. More details regarding its proof are given in Appendix 5.6.3.
From now on, we suppose h to be small enough so that for any $t \in [0, T]$,

$$\sum_{i \in \mathscr{J}_h} w_i(t) m(t, x_i(t), y, I_d(t, x_i(t), \nu, w)) < K + \frac{r^*}{2}, \tag{5.50}$$

where the values of K and r^* are given in (5.11) and (5.18) respectively. Such a choice is always possible thanks to Lemma 5.8.

Theorem 5.4. *Under hypothesis (5.5) through (5.18) and (5.40) through (5.46), for all $T > 0$ and all non-negative initial data $v^0 \in \ell^1(\mathscr{J}_h, \Omega^0)$ there exists a unique solution $x_i \in \mathscr{C}^1([0, T])$, for all $i \in \mathscr{J}_h$, $w := \{w_i(\cdot)\}_{i \in \mathscr{J}_h} \in \mathscr{C}([0, T], \ell^\infty(\mathscr{J}_h))$ and $0 \leqslant \nu := \{\nu_i(\cdot)\}_{i \in \mathscr{J}_h} \in \mathscr{C}([0, T], \ell^1(\mathscr{J}_h))$ of problem (5.47). Furthermore, there exist positive constants c_T and C_T such that the solution satisfies, for all $t \in [0, T]$*

$$c_T h \leqslant |x_i(t) - x_j(t)| \leqslant C_T h, \ \forall i, j \in \mathscr{J}_h, \ i \neq j, \tag{5.51}$$

$$c_T h^d \leqslant w_i(t) \leqslant C_T h^d, \ \forall i \in \mathscr{J}_h, \tag{5.52}$$

$$\|\nu w\|_{1,h} \leqslant \max\{\|v^0 h^d\|_{\ell^1}, \frac{I^*}{\underline{\psi_g}}\}. \tag{5.53}$$

Proof. Consider $v^0 \in \ell^1(\mathscr{J}_h, \Omega^0)$, satisfying $v^0 \geqslant 0$. Consider as well $\alpha > 1$, and define

$$\overline{\rho}_\alpha := \max\{\alpha h^d \|v^0\|_{\ell^1}, \frac{I^*}{\underline{\psi_g}}\},$$

$$\tilde{a} := \|a\|_{W_x^{1,\infty} L_{I,t}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty}} C_\alpha.$$

For $T > 0$ we define the set

$$D_\alpha^T := \{(u, w) \in X_h^T \times Y_h^T : \|uw\|_{1,h} \leqslant \overline{\rho}_\alpha, \forall t \in [0, T], u(t) \geqslant 0, w(t) \geqslant 0, h^d e^{-\tilde{a}t} \leqslant w_k(t) \leqslant h^d e^{\tilde{a}t}\}.$$

For any $(u,w) \in D_\alpha^T$ we introduce the problem, for $t \in [0,T]$,

$$\begin{cases} \dot{x}_i(t) = A_{u,w}(t,x_i), \\ \dot{\omega}_i(t) = \text{div } A_{u,w}(t,x_i(t))\omega_i(t), \\ \dot{\nu}_i(t) = \big(-\text{div } A_{u,w}(t,x_i(t)) + R(t,x_i(t),I_g(t,x_i(t),u,w))\big)\nu_i(t) \\ \qquad\qquad + \sum_{j\in\mathscr{J}_h} \omega_j(t)u_j(t)m(t,x_i(t),x_j(t),I_d(t,x_i(t),u,w)), \\ x_i(0) = x_i^0,\ \omega_i(0) = w_i^0,\ \nu_i(0) = v^0(x_i^0). \end{cases} \tag{5.54}$$

We denote $(\nu,\omega) = \Phi(u,w)$.

For each pair (u,w), the existence and uniqueness of x_i is immediate from Lemma 5.7. Furthermore, for all values of i, we have the following explicit expression for ω_i

$$\omega_i(t) = w_i^0 e^{\int_0^t \text{div } A_{u,w}(s,x_i(s))ds},$$

which satisfies, for any $t \in [0,T]$

$$h^d e^{-\tilde{a}t} \leqslant \omega_i(t) \leqslant h^d e^{\tilde{a}t}.$$

On the other hand, for all $(u,w) \in D_\alpha^T$ and all values of i, the right-hand side of the differential equation in (5.54) is well defined, as we have for all $t \in [0,T]$

$$\sum_{j\in\mathscr{J}_h} w_j(t)u_j(t)m(t,x_i(t),x_j(t),I_d(t,x_i(t),u,w)) \leqslant \overline{M} \sum_{j\in\mathscr{J}_h} w_j(t)u_j(t) \leqslant \overline{M}\overline{\rho}_\alpha.$$

Therefore, the expression for ν_i is given by

$$\nu_i(t) = v^0(x_i^0)e^{\int_0^t \mathscr{G}_i(s)ds} + \int_0^t \sum_{j\in\mathscr{J}_h} w_j(s)u_j(s)m(s,x_i(s),x_j(s),I_d(s,x_i(s),u,w))e^{\int_s^t \mathscr{G}_i(\tau)d\tau}ds, \tag{5.55}$$

where

$$\mathscr{G}_i(t) := -\text{div } A_{u,w}(s,x_i(s)) + R(s,x_i(s),I_g(s,x_i(s),u(s),w(s))),$$

satisfies

$$\sup_{t\in[0,T]} |\mathscr{G}_i(t)| \leqslant \gamma := \|R\|_{L^\infty_{t,x,I}} + \|a\|_{W^{1,\infty}_x L^\infty_{t,I}} + \|a\|_{W^{1,\infty}_I L^\infty_{t,x}} \|\psi_a\|_{W^{1,\infty}_x L^\infty_y} \overline{\rho}_\alpha.$$

The positiveness of ν is immediate from the positiveness of v^0 and m.

Furthermore, given that $|x_i(t) - x_i^0| \leqslant \|a\|_{L^\infty}T$, for all $k \in \mathscr{J}_h$ and $t \in [0,T]$, we have

$$m(t,x_i(s),y,I) = 0,$$

for all $i \notin \mathcal{J}_h^m$, $t \in [0,T]$, $y \in \mathbb{R}^d$ and $I \in \mathbb{R}$. As a result, multiplying (5.55) by $\omega_i(t)$ for each i and adding for all values of i, we obtain

$$\begin{aligned}
\sum_{i\in\mathcal{J}_h} \nu_i(t)\omega_i(t) &\leqslant e^{(\gamma+\tilde{\alpha})T}\left(\sum_{i\in\mathcal{J}_h} v^0(x_i^0)h^d + h^d\int_0^t \sum_{i\in\mathcal{J}_h}\sum_{j\in\mathcal{J}_h} w_j(s)u_j(s)m(s,x_i(s),x_j(s),I_d(s,x_i(s),u,w))ds\right)\\
&= e^{(\gamma+\tilde{\alpha})T}\left(\sum_{i\in\mathcal{J}_h} v^0(x_i^0)h^d + h^d\int_0^t \sum_{i\in\mathcal{J}_h^m}\sum_{j\in\mathcal{J}_h} w_j(s)u_j(s)m(s,x_i(s),x_j(s),I_d(s,x_i(s),u,w))ds\right)\\
&\leqslant e^{(\gamma+\tilde{\alpha})T}\left(\sum_{i\in\mathcal{J}_h} v^0(x_i^0)h^d + h^d|\mathcal{J}_h^m|\overline{M}\int_0^t \sum_{j\in\mathcal{J}_h} w_j(s)u_j(s)ds\right)\\
&\leqslant e^{(\gamma+\tilde{\alpha})T}(\frac{1}{\alpha}+TK_h)\overline{\rho}_\alpha,
\end{aligned}$$

where $K_h := h^d|\mathcal{J}_h^m|\overline{M}$[2]. Thanks to the condition $\alpha > 1$ there exists T_α (only depending on α and on the coefficients of the problem) such that $\Phi : D_\alpha^T \to D_\alpha^T$, for all $T \leqslant T_\alpha$.
We now prove that there exists $T \in (0, T_\alpha)$ such that Φ is a contraction over D_α^T.
Step 1: Bounds over $x = \{x_i\}_{i\in\mathcal{J}_h}$
Let (u^1,w^1) and (u^2,w^2) be two pairs in D_α^T, and let x^1, x^2 be the respective solutions of

$$\dot{x}_i^j = A_{u^j,w^j}(t,x_i^j).$$

By following the same ideas as in the proof of Lemma 5.1 (see Appendix 5.6.1), we obtain that for all $t \in [0,T]$,

$$\|x^1(t)-x^2(t)\|_{\infty,h} \leqslant C(T,h)\left(\|w^1-w^2\|_{\infty,h} + \|u^1-u^2\|_{1,h}\right),$$

where the constant $C(T,h)$ satisfies $\lim_{T\to 0} C(T,h) = 0$.
Step 2: Bounds over $\omega = \{\omega_i\}_{i\in\mathcal{J}_h}$
From the expression for ω, we get, for all $t \in [0,T]$

$$\begin{aligned}
|\omega_i^1(t)-\omega_i^2(t)| \leqslant & h^dTe^{\tilde{a}T}\Big(\|a\|_{W^{2,\infty}_{x,I}}(1+|\partial_x I_a(t,x_i^2,u^2,w^2)|)\|x^1-x^2\|_{\infty,h}\\
&+\|a\|_{W^{2,\infty}_{x,I}}(1+|\partial_x I_a(t,x_i^2,u^2,w^2)|)|I_a(t,x_i^1,u^1,w^1)-I_a(t,x_i^2,u^2,w^2)|\\
&+\|a\|_{W^{1,\infty}_I}|\partial_x I_a(t,x_i^1,u^1,w^1)-\partial_x I_a(t,x_i^2,u^2,w^2)|\Big).
\end{aligned}$$

On the other hand we have

$$\begin{aligned}
|\partial_x I_a(t,x_i^2,u^2,w^2)| \leqslant & \|\psi_a\|_{W^{1,\infty}_x}\overline{\rho}_\alpha,\\
|I_a(t,x_i^1,u^1,w^1)-I_a(t,x_i^2,u^2,w^2)| \leqslant & \overline{\rho}_\alpha\|\psi_a\|_{W^{1,\infty}_{x,y}}\|x^1-x^2\|_{\infty,h},\\
&+\|\psi_a\|_{L^\infty_{x,y}}e^{\tilde{a}T}(h^d\|u^1-u^2\|_{1,\infty}+\frac{\overline{\rho}_\alpha}{h^d}\|w^1-w^2\|_{\infty,h}),
\end{aligned}$$

and

$$\begin{aligned}
|\partial_x I_a(t,x_i^1,u^1,w^1)-\partial_x I_a(t,x_i^2,u^2,w^2)| \leqslant & \overline{\rho}_\alpha\|\psi_a\|_{W^{2,\infty}_{x,y}}\|x^1-x^2\|_{\infty,h}\\
&+\|\psi_a\|_{W^{1,\infty}_x}e^{\tilde{a}T}(h^d\|u^1-u^2\|_{1,\infty}+\frac{\overline{\rho}_\alpha}{h^d}\|w^1-w^2\|_{\infty,h}).
\end{aligned}$$

[2]Notice that $h^d|\mathcal{J}_h^m| \approx |\text{supp } m + B_{\|a\|_{L^\infty}T}|$

In conclusion, there exists a constant $C(T,h)$, satisfying $\lim_{T\to 0} C(T,h)=0$ such that

$$\|\omega^1-\omega^2\|_{\infty,h}\leqslant C(T,h)\left(\|w^1-w^2\|_{\infty,h}+\|u^1-u^2\|_{1,h}\right).$$

Step 3: Bounds over $\nu=\{\nu_i\}_{i\in\mathscr{J}_h}$

Using the expression for ν, the regularity of m and bounds similar to those used for ω, we see that there exists a constant $C(T,h)$ satisfying $\lim_{T\to 0} C(T,h)=0$ such that

$$\|\nu^1-\nu^2\|_{1,h}\leqslant C(T,h)\left(\|w^1-w^2\|_{\infty,h}+\|u^1-u^2\|_{1,h}\right).$$

Consequently, there exists a constant $C(T,h)$ satisfying $\lim_{T\to 0} C(T,h)=0$, such that

$$\|\omega^1-\omega^2\|_{\infty,h}+\|\nu^1-\nu^2\|_{1,h}\leqslant C(T,h)\left(\|w^1-w^2\|_{\infty,h}+\|u^1-u^2\|_{1,h}\right),$$

which implies that, for $0<T_1\leqslant T_\alpha$ small enough, Φ is a contraction over $D_\alpha^{T_1}$, and therefore it has a unique fixed point. Such fixed point is a solution of problem (5.47) over $[0,T_1)$.

We now claim that the solution exists for T arbitrary, and furthermore, it satisfies the relation (5.53). Let T_f be the maximal time of existence of solution. Suppose that there exists $t_0\in(0,T_f]$ such that $\|\nu w\|_{1,h}>\overline{\rho}_\alpha$. This implies that there exist $\delta\geqslant 0$ and $t^*>0$ such that for a certain finite subset of $\mathscr{J}_h$, that we denote as $\mathscr{K}_h$, the following statements are true:

$$\sum_{i\in\mathscr{K}_h}\nu_i(t)w_i(t)\leqslant\overline{\rho}_\alpha,\ \forall\ t\in[t^*-\delta,t^*],$$
$$\sum_{i\in\mathscr{K}_h}\nu_i(t)w_i(t)>\overline{\rho}_\alpha,\ \forall\ t\in(t^*,t^*+\delta].$$

This implies the existence of $t_1\in[t^*,t^*+\delta]$ such that the following properties are satisfied simultaneously

$$\sum_{i\in\mathscr{K}_h}\nu_i(t_1)w_i(t_1)\geqslant\overline{\rho}_\alpha \text{ and } \left(\sum_{i\in\mathscr{K}_h}\nu_i w_i\right)'(t_1)\geqslant 0. \tag{5.56}$$

Multiplying the equation satisfied by $\nu_i(t)$ by $w_i(t)$ we get the relation

$$\begin{aligned}\dot\nu_i(t)w_i(t)=&\big(-\operatorname{div} A_{\nu,w}(t,x_i(t))+R(t,x_i(t),I_g(t,x_i(t),\nu,w))\big)\nu_i(t)w_i(t)\\&+w_i(t)\sum_{j\in\mathscr{J}_h}w_j(t)\nu_j(t)m(t,x_i(t),x_j(t),I_g(t,x_i(t),\nu,w)),\end{aligned}$$

while, directly from the equation for $w_i(t)$ we deduce

$$\nu_i(t)\dot w_i(t)=\operatorname{div} A_{\nu,w}(t,x_i(t))\nu_i(t)w_i(t).$$

Therefore, adding both relations for $i\in\mathscr{K}_h$ and using (5.50), we get

$$\begin{aligned}\left(\sum_{i\in\mathscr{K}_h}\nu_i w_i\right)'(t)=&\left(\sum_{i\in\mathscr{K}_h}R(t,x_i(t),I_g(t,x_i(t),\nu,w))\nu_i(t)w_i(t)\right)\\&+\sum_{j\in\mathscr{J}}w_j(t)\nu_j(t)\sum_{i\in\mathscr{K}_h\cap\mathscr{J}_h^m}m(t,x_i(t),x_j(t),I_d(t,x_i(t),\nu,w))w_i(t)\\\leqslant&\left(\sum_{i\in\mathscr{K}_h}R(t,x_i(t),I_g(t,x_i(t)))\nu_i(t)w_i(t)\right)+(K+\frac{r^*}{2})\sum_{j\in\mathscr{K}_h}w_j(t)\nu_j(t).\end{aligned}$$

Given that $\|\nu(t_1)w(t_1)\|_{\ell^1} \geqslant \overline{\rho}_\alpha \geqslant \frac{I^*}{\underline{\psi}_g}$, then $I_g(t_1, x_i(t_1), \nu(t_1), w(t_1)) \geqslant I^*$, and consequently

$$R(t_1, x_i(t_1), I_g(t_1, x_i(t_1), \nu(t_1), w(t_1))) < -r^* - K,$$

for all values of i, which in turn implies that

$$\left(\sum_{i\in\mathscr{K}_h} \nu_i w_i\right)'(t_1) \leqslant -\frac{r^2}{2}\sum_{i\in\mathscr{K}_h} \nu_i(t_1)w_i(t_1) < 0,$$

which contradicts (5.56). Therefore, $\|\nu w\|_{1,h} \leqslant \overline{\rho}_\alpha$ for all values of $\alpha > 1$ and for all $t \in (0, T_f)$. We can then iterate the arguments used to prove existence of a solution, and conclude that the solution can be extended to any interval $[0, T]$. As $\|\nu w\|_{1,h} \leqslant \overline{\rho}_\alpha$ for all t, independently of α, taking the limit when α goes to 1, we obtain (5.53). □

5.4 Convergence of the numerical solution towards a weak solution

We study now the conditions under which a solution of problem (5.54) converges towards a solution of problem (5.19), in a certain sense that will be defined later. We split our analysis in two cases: first, the study of convergence on a finite interval of time $[0, T]$. We will see that for any $T > 0$, the solution obtained through the particle method converges towards the solution of the PDE (5.19). However, the speed of convergence might be affected by the value of T. The second case we study is the asymptotic proximity of both solutions when t goes to ∞. This is a far more complex and interesting issue, and we show different examples exposing some of the behaviours that can be observed.

Directly from the study of existence of solutions for each problem, we notice that the sets of hypotheses we have used, do not coincide. We give a set of hypotheses which simultaneously guarantees the existence of solution for both problems, while taking into account the distinction of cases involved in the definition of κ.

For a certain $T > 0$ and $k > 0$, we consider the functions

$$\psi_a \in \mathscr{C}([0,T]\times\mathbb{R}^d_x, W^{1,\infty}(\mathbb{R}^d_y)) \cap \mathscr{C}\left([0,T]\times\mathbb{R}^d_y, \mathscr{C}^{k+1}(\mathbb{R}^d_x)\cap W^{k+1,\infty}(\mathbb{R}^d_x)\right), \tag{5.57}$$

$$0 < \underline{\psi_g} \leqslant \psi_g \in \mathscr{C}\left([0,T]\times\mathbb{R}_x, W^{1,\infty}(\mathbb{R}^d_y)\right) \cap \mathscr{C}\left([0,T]\times\mathbb{R}_y, \mathscr{C}^{\kappa}(\mathbb{R}^d_x)\cap W^{\kappa,\infty}(\mathbb{R}^d_x)\right), \tag{5.58}$$

$$\psi_d \in \mathscr{C}\left([0,T]\times\mathbb{R}^d_x, W^{1,\infty}(\mathbb{R}^d_y)\right) \cap \mathscr{C}\left([0,T]\times\mathbb{R}^d_y, \mathscr{C}^{\kappa}(\mathbb{R}^d_x)\cap W^{\kappa,\infty}(\mathbb{R}^d_x)\right). \tag{5.59}$$

As in Section 5.2 we introduce

$$a \in \mathscr{C}([0,T], W^{k+1,\infty}(\mathbb{R}^{d+1})), \tag{5.60}$$

$$R \in \mathscr{C}\left([0,T]\times\mathbb{R}_x, W^{\kappa,\infty}_{loc}(\mathbb{R}^d_y)\right) \cap \mathscr{C}\left([0,T]\times\mathbb{R}_y, \mathscr{C}^{\kappa}(\mathbb{R}^d_x)\cap W^{\kappa,\infty}(\mathbb{R}^d_x))\right) \tag{5.61}$$

We consider as well

$$0 \leqslant m \in \mathscr{C}\left([0,T]\times\mathbb{R}^d_x\times\mathbb{R}^d_y, W^{\kappa,\infty}(\mathbb{R}_I)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}^d_x\times\mathbb{R}^d_I, W^{\kappa,\infty}(\mathbb{R}^d_y)\right)\cap\mathscr{C}\left([0,T]\times\mathbb{R}^d_y\times\mathbb{R}^d_I, \mathscr{C}^{\kappa}_c(\mathbb{R}^d_x)\right). \tag{5.62}$$

satisfying hypothesis (5.7) through (5.18).

Finally, we consider $v^0 \in W^{k,1}(\mathbb{R}^d)$ if $\partial_I a = 0$ and $v^0 \in W^{k,1}(\mathbb{R}^d)\cap W^{k,\infty}(\mathbb{R}^d)$ with compact support otherwise.

5.4.1 Convergence on a finite time interval

We recall that the function v represents the solution of problem (5.19) while x_i, w_i and ν_i, $i \in \mathscr{J}_h$ represents that of problem (5.54). We recall as well that

$$\max\{\|v\|, \|\nu w\|_{1,h}\} \leqslant \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \|v^0 h^d\|_{\ell^1}, \frac{I^*}{\underline{\psi_g}}\} =: \overline{\rho}.$$

Let $\varepsilon > 0$, $r \in \mathbb{R}$ and $\varphi \in \mathscr{C}_c(\mathbb{R}^d)$ satisfy the following conditions

$$\int_{\mathbb{R}^d} \varphi(x)dx = 1, \tag{5.63}$$

$$\int_{\mathbb{R}^d} x^\alpha \varphi(x)dx = 0, \quad \forall\, \alpha \in \mathbb{N}^n, |\alpha| \leqslant r-1. \tag{5.64}$$

. We define, for all $t \in [0,T]$, $x \in \mathbb{R}^d$

i)

$$v^h(t,x) = \sum_{i\in\mathscr{J}_h} \nu_i(t) w_i(t) \delta(x - x_i(t)), \tag{5.65}$$

a time dependent measure obtained as a sum of weighted Dirac deltas at $x_i(t)$,

ii)

$$v^h_\varepsilon(t,x) = (v^h(t) * \varphi_\varepsilon)(x) = \sum_{i\in\mathscr{J}_h} \nu_i(t) w_i(t) \varphi_\varepsilon(x - x_i(t)), \tag{5.66}$$

a regular function obtained as the space convolution of $v^h(t,x)$ and $\varphi_\varepsilon(x)$, where

$$\varphi_\varepsilon := \frac{1}{\varepsilon^d}\varphi(\frac{\cdot}{\varepsilon}).$$

We also introduce the following operator, for any function $v \in L^\infty(0,T;L^1(\mathbb{R}^d))$:

$$\Big(\Pi^h_\varepsilon(t) v\Big)(x) = \sum_{i\in\mathscr{J}_h} w_i(t) v(t, x_i(t)) \varphi_\varepsilon(x - x_i(t)).$$

We recall a direct corollary of the **Theorem 3** in [184]:

Proposition 5.1. *Let k, r be two integers, with $k > d$, and let us assume that $a \in L^\infty\big(0,T;W^{k+1,\infty}(\mathbb{R}^d)^d\big)$, and that $\varphi \in \mathscr{C}^1_c(\mathbb{R}^d) \cap W^{k+1,1}(\mathbb{R}^d)$ satisfies conditions (5.63) and (5.64). Then, for any $p \in [1,+\infty]$, there exists $C = C(T) > 0$ such that, for any $u \in W^{\mu,1}(\mathbb{R}^d)$ ($\mu = \max(r,k)$),*

$$\|u - \Pi^h_\varepsilon(t) u\|_{L^p(\mathbb{R}^d)} \leqslant C\big(\varepsilon^r \|u\|_{W^{r,p}(\mathbb{R}^d)} + \big(\frac{\varepsilon}{h}\big)^k \|u\|_{W^{k,p}(\mathbb{R}^d)}\big).$$

We seek to prove the following approximation result between v^h_ε and v, the solution of problem (5.19).

Theorem 5.5. *Assume that hypotheses (5.57) through (5.62) are satisfied, and that $\varphi \in \mathscr{C}_c^1(\mathbb{R}^d) \cap W^{k+1,1}(\mathbb{R}^d)$ satisfies (5.63) and (5.64). Then, there exists $C = C(T, a, R, m, \overline{\rho}) > 0$, a positive constant which depends on T, a, R, m and $\overline{\rho}$ such that*

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leqslant C\big(\varepsilon^r + \big(\frac{h}{\varepsilon}\big)^\kappa + h^\kappa\big)\|v^0\|_{W^{\mu,1}(\mathbb{R}^d)}, \quad \forall\, 0 \leqslant t \leqslant T,$$

where $\mu = \max(r, \kappa)$.

The proof of Theorem 5.5 strongly relies on Proposition 5.1 and the following result:

Proposition 5.2. *Under hypotheses (5.57) through (5.62), there exists a constant $C_T > 0$, depending only on T and on the parameters of problems (5.19) and (5.54), such that their respective solutions satisfy, for all $t \in [0, T]$,*

$$\sum_{i \in \mathscr{J}_h} |v(t, x_i(t)) - \nu_i(t)| w_i(t) \leqslant C_T h^{k-1} \|v^0\|_{W^{\mu,1}(\mathbb{R}^d)}. \tag{5.67}$$

Proof. Consider β_ε as in (5.88). We define $e = \{e_i(\cdot)\}_{i \in \mathscr{J}}$ where for all $i \in \mathscr{J}$ and $t \in [0, T]$,

$$\begin{aligned} e_i(t) &:= v(t, x_i(t)) - \nu_i(t), \\ e_{\varepsilon,i}(t) &:= \beta_\varepsilon(e_i(t)) w_i(t), \end{aligned}$$

and compute

$$\dot{e}_{\varepsilon,i}(t) = \beta_\varepsilon'(e_i(t))\dot{e}_i(t) w_i(t) + \beta_\varepsilon(e_i(t))\dot{w}_i(t). \tag{5.68}$$

We recall that

$$\begin{aligned} \dot{e}_i(t) =& \Big(G_{\nu,w}(t, x_i(t)) - \mathscr{G}_v(t, x_i(t))\Big)\nu_i(t) + \Big(A_{\nu,w}(t, x_i(t)) - \mathscr{A}_v(t, x_i(t))\Big)\nabla v(t, x_i(t)) \\ & - \mathscr{G}_v(t, x_i(t)) e_i(t) + \Delta\mathscr{M}(t, x_i(t)), \end{aligned}$$

where

$$\begin{aligned} G_{\nu,w}(t, x_i(t)) &:= \operatorname{div} A_{\nu,w}(t, x_i(t)) - R(t, x_i(t), I_g(t, x_i(t), \nu, w)), \\ \mathscr{G}_v(t, x_i(t)) &:= \operatorname{div} \mathscr{A}_v(t, x_i(t)) + R(t, x_i(t), (I_g v)(t, x_i(t))), \\ \Delta\mathscr{M}(t, x_i(t)) &:= \int_{\mathbb{R}^d} m(t, x_i(t), y, (I_d v)(t, x_i)) v(t, y) dy - \sum_{j \in \mathscr{J}_h} w_j(t)\nu_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t), \nu, w)). \end{aligned}$$

The functions a and R being Lipschitz, there exists a constant C depending on the parameters of the problem and the value $\overline{\rho}$, such that

$$\begin{aligned} \Big|G_{\nu,w}(t, x_i(t)) - \mathscr{G}_v(t, x_i(t))\Big| \leqslant& C\Big(|(I_a v)(t, x_i(t)) - I_a(t, x_i(t), \nu, w)| \\ & + |\partial_x (I_a v)(t, x_i(t)) - \partial_x I_a(t, x_i(t), \nu, w)| \\ & + |(I_g v)(t, x_i(t)) - I_g(t, x_i(t), \nu, w)|\Big). \end{aligned}$$

Notice that

$$\begin{aligned}
|(I_a v)(t,x_i(t)) - I_a(t,x_i(t),\nu,w)| =& \left| \int\limits_{\mathbb{R}^d} \psi_a(t,x_i(t),y)v(y)dy - \sum_{j\in\mathscr{J}_h} \psi_a(t,x_i(t),x_j(t))\nu_j(t)w_j(t) \right| \\
\leqslant& \left| \int\limits_{\mathbb{R}^d} \psi_a(t,x_i(t),y)v(y)dy - \sum_{j\in\mathscr{J}_h} \psi_a(t,x_i(t),x_j(t))v(t,x_j(t))w_j(t) \right| \\
&+ \left| \sum_{j\in\mathscr{J}_h} \psi_a(t,x_i(t),x_j(t))\Big(v(t,x_j(t)-\nu_j(t))\Big)w_j(t) \right| \\
\leqslant& C\Bigg(h^\kappa \|v\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j\in\mathscr{J}_h} |e_j(t)|w_j(t)\Bigg),
\end{aligned}$$

where in the last line we have used Lemma 5.8 and the $W^{\kappa,1}(\mathbb{R}^d)$ regularity of v. Similar results are true for $|\partial_x(I_a v)(t,x_i(t)) - \partial_x I_a(t,x_i(t),\nu,w)|$ and $|(I_g v)(t,x_i(t)) - I_g(t,x_i(t),\nu,w)|$. In conclusion, thanks to (5.27), there exists a constant C_T, only depending on T, the parameters of the problem and the value $\overline{\rho}$, such that

$$\Big|G_{\nu,w}(t,x_i(t)) - \mathscr{G}_v(t,x_i(t))\Big| \leqslant C_T\Bigg(h^\kappa\|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j\in\mathscr{J}_h}|e_j(t)|w_j(t)\Bigg). \tag{5.69}$$

Again, using the Lipschitz regularity of a, we see that

$$\begin{aligned}
\Big|A_{\nu,w}(t,x_i(t)) - \mathscr{A}_v(t,x_i(t))\Big| \leqslant& C|(I_a v)(t,x_i(t)) - I_a(t,x_i(t),\nu,w)| \\
\leqslant& C_T\Bigg(h^\kappa\|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j\in\mathscr{J}_h}|e_j(t)|w_j(t)\Bigg).
\end{aligned} \tag{5.70}$$

The boundedness of a and R implies the existence of a constant $\overline{G}$ such that for all $i \in \mathscr{J}$ and $t \in [0,T]$,

$$|\mathscr{G}_v(t,x_i(t))| \leqslant \overline{G}. \tag{5.71}$$

We recall from the previous section that

$$\Delta\mathscr{M}(t,x_i(t)) = 0,\ \forall i \notin \mathscr{J}_h^m,$$

where the set of indexes $\mathscr{J}_h^m$ has a finite number of elements, which depends on T. On the other hand, for those $i \in \mathscr{J}_h^m$, we have

$$\begin{aligned}
&|\Delta\mathscr{M}(t,x_i(t))| \\
\leqslant& \left| \int\limits_{\mathbb{R}^d} m(t,x_i(t),y,(I_d v)(t,x_i(t)))v(t,y)dy - \sum_{j\in\mathscr{J}_h} w_j(t)v(t,x_j(t))m(t,x_i(t),x_j(t),(I_d v)(t,x_i(t))) \right| \\
&+ \sum_{j\in\mathscr{J}_h} w_j(t)v(t,x_j(t))\Big|m(t,x_i(t),x_j(t),(I_d v)(t,x_i(t))) - m(t,x_i(t),x_j(t),I_d(t,x_i(t),\nu,w))\Big| \\
&+ \sum_{j\in\mathscr{J}_h} w_j(t)|e_j(t)|m(t,x_i(t),x_j(t),(I_d v)(t,x_i(t))) \\
\leqslant& C\Bigg(h^\kappa\|v\|_{W^{\kappa,1}(\mathbb{R}^d)} + \mu|(I_d v)(t,x_i(t)) - I_d(t,x_i(t),\nu,w)|\sum_{j\in\mathscr{J}_h} w_j(t)v(t,x_j(t)) + \overline{M}\sum_{j\in\mathscr{J}_h}|e_j(t)|w_j(t)\Bigg),
\end{aligned}$$

where we have used again Lemma 5.8, the $W^{\kappa,1}(\mathbb{R}^d)$ regularity of $m(t,x,y,I)v(t,y)$ with respect to the y variable, and the Lipschitz regularity of m. Furthermore, Lemma 5.8 gives us the bound

$$\sum_{j\in\mathscr{J}_h} w_j(t)v(t,x_j(t)) \leqslant \|v\|_{L^1(\mathbb{R}^d)} + Ch^\kappa\|v\|_{W^{\kappa,1}(\mathbb{R}^d)},$$

which together with manipulations similar to those made for $|(I_a v)(t,x_i(t)) - I_a(t,x_i(t),\nu,w)|$, and the bound (5.27), gives

$$|\Delta\mathscr{M}(t,x_i(t))| \leqslant C_T\left(h^\kappa\|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j\in\mathscr{J}_h}|e_j(t)|w_j(t)\right). \tag{5.72}$$

We denote as $\mathscr{K}$ an arbitrary finite subset of $\mathscr{J}_h$. If we add (5.68) for all values of $i\in\mathscr{K}$, and use bounds (5.69) through (5.72), together with the equation for w_i, we get

$$\left(\sum_{i\in\mathscr{K}} e_{\varepsilon,i}(t)\right)' \leqslant C_T B(t)\left(h^\kappa\|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j\in\mathscr{J}_h}|e_j(t)|w_j(t)\right),$$

where

$$B(t) := \sum_{i\in\mathscr{K}}(\nu_i(t) + |\nabla v(t,x_i(t))|)w_i(t) + \overline{G} + |\mathscr{K}\cap\mathscr{J}_h^m|h^d e^{\tilde{a}T} + \tilde{a}.$$

Given that

$$\sum_{i\in\mathscr{K}}\nu_i(t)w_i(t) \leqslant \overline{\rho},$$

and

$$\sum_{i\in\mathscr{K}} w_j(t)|\nabla v(t,x_j(t))| \leqslant \|\nabla v\|_{L^1(\mathbb{R}^d)} + Ch^{\kappa-1}\|\nabla v\|_{W^{\kappa-1,1}(\mathbb{R}^d)},$$

thanks to (5.27), we conclude that there exists a constant B_T, independent of the choice of $\mathscr{K}$ and h, such that $B(t)\leqslant B_T$. Consequently, for all values of $t\in[0,T]$, h small enough and any finite subset of $\mathscr{J}_h$, we have the relation

$$\left(\sum_{i\in\mathscr{K}} e_{\varepsilon,i}(t)\right)' \leqslant C_T\left(h^\kappa + \sum_{j\in\mathscr{J}_h}|e_j(t)|w_j(t)\right).$$

Integrating between 0 and t, taking the limit when ε goes to zero, and using Grönwall's lemma, we obtain that there exists a constant C_T, independent of $\mathscr{K}$ such that

$$\sum_{i\in\mathscr{K}}|e_i(t)|w_i(t) \leqslant C_T h^\kappa\|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)}.$$

Being C_T independent of $\mathscr{K}$ and h, (5.67) is immediate. □

In other words, we proved in Proposition 5.2 that the piece-wise constant functions that take values $v(t,x_k(t))$ and $v_k(t)$ respectively over the intervals $\Omega_k(t)$ are close in $L^1(\mathbb{R}^d)$.

Proof of Theorem 5.5. According to the triangle inequality,

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leqslant \|v - \Pi_\varepsilon^h(t)v\|_{L^1(\mathbb{R}^d)} + \|\Pi_\varepsilon^h(t)v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)}, \tag{5.73}$$

it only remains to bound both terms on the right hand side.

i) According to Proposition 5.1 with $p=1$, and bound (5.27)

$$\|v-\Pi_\varepsilon^h(t)v\|_{L^1(\mathbb{R}^d)} \leqslant C(\varepsilon^r+(\frac{\varepsilon}{h})^\kappa)\|v\|_{\mu,p} \leqslant C_T(\varepsilon^r+(\frac{\varepsilon}{h})^\kappa)\|v^0\|_{\mu,p}.$$

ii) On the other hand, one computes

$$\begin{aligned}\|\Pi_\varepsilon^h(t)v-v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d}\Big|\sum_{i\in\mathscr{J}_h} w_i(t)\varphi_\varepsilon(x-x_i(t))\big(v(t,x_i(t))-\nu_i(t)\big)\Big|dx \\ &\leqslant \sum_{i\in\mathscr{J}_h}\left(w_i(t)|v(t,x_i(t))-\nu_i(t)|\int_{\mathbb{R}^d}|\varphi_\varepsilon(x-x_i(t))|dx\right).\end{aligned}$$

According to the definition of φ_ε, with the change of variable $x'=\frac{x-x_k(t)}{\varepsilon}$, we note that

$$\int_{\mathbb{R}^d}|\varphi_\varepsilon(x-x_k(t))|dx=\int_{\mathbb{R}^d}|\varphi(x)|dx<+\infty,$$

by hypothesis on φ. We have then, according to Proposition 5.2, that

$$\|\Pi_\varepsilon^h(t)v-v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leqslant C_T h^\kappa\|v^0\|_{\kappa,1} \leqslant Ch^\kappa\|v^0\|_{\mu,1},$$

which concludes the proof of Theorem 5.5.

□

5.4.2 Asymptotic preserving properties

The study of the asymptotic behaviour of the solution for adaptive dynamics models, such as (5.19), is one of the main interests often treated in the literature (see [24, 30, 165, 166, 170–174, 178]). For this reason, the design of numerical methods which preserve the asymptotic behaviour, or at least, the identification of the problems for which the asymptotics are preserved under a certain numerical scheme, is a priority. In other words, given that $v(t,\cdot)$ converges to a measure μ when t goes to infinity, we expect to identify the conditions under which $\lim_{h\to 0}\lim_{t\to+\infty} v^h_{\varepsilon(h)}(t,\cdot)=\mu$, that is, ensuring the commutativity of diagram (5.74):

$$\begin{array}{ccc} v(t,\cdot) & \xrightarrow[t\to\infty]{} & \mu \\ {\scriptstyle h\to 0}\uparrow & & \uparrow{\scriptstyle h\to 0} \\ v^h_{\varepsilon(h)(t)} & \xrightarrow[t\to\infty]{} & \mu_h \end{array} \tag{5.74}$$

In what follows, we formally define the concept of an asymptotic preserving approximation, and give examples and counter-examples of this concept.

We recall that, according to the Riesz representation theorem, the space of finite Radon measures can be identified with the topological dual space of $\mathscr{C}_c(\mathbb{R}^d)$. Hence, we say that a sequence of finite

Radon measures $\{\mu_n\}_{n\in\mathbb{N}}$ converges weakly to a finite Radon measure μ (denoted $\mu_n \underset{n\to+\infty}{\rightharpoonup} \mu$) if for all $\phi \in \mathscr{C}_c(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \phi(x) d\mu_n(x) \underset{n\to+\infty}{\longrightarrow} \int_{\mathbb{R}^d} \phi(x) d\mu(x).$$

This leads us to introduce the following definition:

Definition 5.1. *We say that the particle solution v_ε^h defined in (5.66) is an asymptotic preserving approximation of v, the solution to (5.19), if for all $\varepsilon : (0,1] \to \mathbb{R}_+^*$ which converges to 0 when h goes to 0, and all $\phi \in \mathscr{C}_c(\mathbb{R}^d)$,*

$$\limsup_{t\to+\infty} \left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t,x) dx - \int_{\mathbb{R}^d} \phi(x) v(t,x) dx \right| \underset{h\to 0}{\longrightarrow} 0.$$

The following lemma ensures that, in the previous definition, $v_{\varepsilon(h)}^h$, introduced in (5.66), can be replaced by v^h, introduced in (5.65).

Lemma 5.9. *The function v_ε^h is an asymptotic preserving approximation of v if and only if*

$$\limsup_{t\to+\infty} \left| \int_{\mathbb{R}^d} \phi(x) dv^h(t,x) - \int_{\mathbb{R}^d} \phi(x) v(t,x) dx \right| \underset{h\to 0}{\longrightarrow} 0.$$

Proof. Let us prove that for all $\varepsilon : (0,1] \to \mathbb{R}_+^*$ which converges to 0 as h goes to 0, and all $\phi \in \mathscr{C}_c(\mathbb{R}^d)$,

$$\limsup_{t\to+\infty} \left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t,x) dx - \int_{\mathbb{R}^d} \phi(x) dv^h(t,x) \right| \underset{h\to 0}{\longrightarrow} 0.$$

Let $h > 0$. According to the definitions of v^h and v_ε^h, and since $\int_{\mathbb{R}^d} \varphi_{\varepsilon(h)}(x - x_i(t)) dx = 1$ for all $i \in \mathscr{J}_h$ we get, for all $t \geqslant 0$,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t,x) dx - \int_{\mathbb{R}^d} \phi(x) dv^h(t,x) \right| &= \left| \int_{\mathbb{R}^d} \sum_{i\in\mathscr{J}_h} \nu_i(t) w_i(t) \big(\phi(x) - \phi(x_i(t))\big) \varphi_{\varepsilon(h)}(x - x_i(t)) dx \right| \\ &\leqslant \sum_{i\in\mathscr{J}_h} |\nu_i(t) w_i(t)| \int_{\mathbb{R}^d} \left| \big(\phi(x) - \phi(x_i(t))\big) \varphi_{\varepsilon(h)}(x - x_i(t)) \right| dx. \end{aligned}$$

With the change of variable '$y = \frac{x - x_i(t)}{\varepsilon(h)}$', we get, for all $i \in \mathscr{J}_h$,

$$\int_{\mathbb{R}^d} \left| \big(\phi(x) - \phi(x_i(t))\big) \varphi_{\varepsilon(h)}(x - x_i(t)) \right| dx = \int_K \left| \big(\phi(\varepsilon(h) y + x_i(t)) - \phi(x_i(t))\big) \varphi(y) \right| dy,$$

where K is the support of φ. Let $\eta > 0$. Since ϕ is continuous with a compact support, and thus uniformly continuous, then $|\phi(\varepsilon(h)x + x_i(t)) - \phi(x_i(t))| \leqslant \eta$ for all $i \in \mathscr{J}_h$, $x \in K$, $t \geqslant 0$ and any h small enough. Therefore, for any h small enough,

$$\left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t,x) dx - \int_{\mathbb{R}^d} \phi(x) dv^h(t,x) \right| \leqslant \eta \sum_{i\in\mathscr{J}_h} \nu_i(t) w_i(t),$$

which concludes the proof, since there exists $\overline{\rho} > 0$ such that $0 \leqslant \sum_{i\in\mathscr{J}_h} \nu_i(t) w_i(t) \leqslant \overline{\rho}$ for all $h > 0$, $t \geqslant 0$, as proved in Theorem 5.4. ∎

The problem of determining if v_ε^h is an asymptotic preserving approximation of v is generally a difficult question. In what follows, we deal with cases where we are able to determine the asymptotic behaviour of both v and v^h, and we check if the necessary and sufficient condition from Lemma 5.9 holds. From now on, we assume that a is local and not time dependent, *i.e.* $a(t,x,I) = a(x)$ and that that the functions m, R, ψ_g and ψ_d are not time-dependent. We assume as well that the function $(x,y,I) \mapsto m(x,y,I)$ is not only uniformly compactly supported as a function of the x variable, but relative to the y variable as well. That is, there exist two compact sets K_x and K_y such that $\sup_{y,I} m(x,y,I) = 0$ for all y outside of K_x and $\sup_{x,I} m(x,y,I) = 0$ for all x outside of K_y. We denote $K_{xy} := K_x \cup K_y$. Finally, we assume ψ_g and ψ_d to be compactly supported as functions of the y variable.

Necessary conditions of convergence towards a Radon measure

With the help of necessary conditions, we would be able to rule out those cases where v does not converge towards certain types of Radon measures, which are the object of our interest. We start by giving a general result, involving the necessary conditions of convergence towards any Radon measure.

Lemma 5.10. *Let us assume that $v(t,\cdot) \underset{t\to+\infty}{\rightharpoonup} \mu$ in the weak sense in the space of finite Radon measures. Then, for $\gamma \in \{g,d\}$,*

$$I_\alpha(t,x) \underset{t\to+\infty}{\longrightarrow} \int_{\mathbb{R}^d} \psi_\alpha(x,y)d\mu(y) =: \overline{I_\alpha}(x) \quad \forall x \in \mathbb{R}^d,$$

and for all $\phi \in \mathscr{C}_0^1(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \Big(a(x)\cdot\nabla\phi(x) + R\left(x,\overline{I_g}(x)\right)\phi(x)\Big) d\mu(x) + \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} m\left(x,y,\overline{I_d}(x)\right)\phi(x)dx\right) d\mu(y) = 0.$$

Proof. Let us assume that $v(t,\cdot) \underset{t\to+\infty}{\rightharpoonup} \mu$. By definition of the weak convergence in the space of finite Radon measure, for any $\phi \in \mathscr{C}_c^1(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \phi(x)v(t,x)dx \underset{t\to+\infty}{\to} \int_{\mathbb{R}^d} \phi(x)d\mu(x).$$

The first identity is thus a direct consequence of the definition of the weak convergence, applied with $\phi = \psi_\gamma(x,\cdot)$, $\gamma \in \{g,d\}$.

Let $\phi \in \mathscr{C}_c(\mathbb{R}^d)$. One computes

$$
\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^d}\phi(x)v(t,x)dx = &-\int_{\mathbb{R}^d}\phi(x)\nabla\cdot(a(x)v(t,x)) + \int_{\mathbb{R}^d}R(x,I_g(t,x))\phi(x)v(t,x)dx\\
&+\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}m(x,y,I_d(t,x))v(t,y)\right)\phi(x)dx\\
=&\int_{\mathbb{R}^d}\Big(a(x)\cdot\nabla\phi(x)+R(x,I_g(t,x))\phi(x)\Big)v(t,x)dx\\
&+\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}m(x,y,I_d(t,x))\phi(x)dx\right)v(t,y)dy\\
\underset{t\to+\infty}{\longrightarrow}&\int_{\mathbb{R}^d}\Big(a(x)\cdot\nabla\phi(x)+R\left(x,\overline{I_g}(x)\right)\phi(x)\Big)d\mu(x)\\
&+\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}m\left(x,y,\overline{I_d}(x)\right)\phi(x)dx\right)d\mu(y).
\end{aligned}
$$

Thus, $t\mapsto\int_{\mathbb{R}^d}\phi(x)v(t,x)dx$ is a convergent function with a convergent derivative, which ensures that the limit of its derivative is zero, which concludes the proof. □

The following proposition provides a necessary condition for the convergence to a sum of Dirac masses.

Proposition 5.3. *Let us assume that* $v(t,\cdot)\underset{t\to+\infty}{\rightharpoonup}\sum_{i=1}^N C_i\delta_{x_i}$*, with* $x_1,\dots,x_N\in\mathbb{R}^d$*,* $C_1,\dots,C_N>0$*. Then, for* $\alpha=g,d$*,* $I_\alpha(t,x)\underset{t\to+\infty}{\to}\sum_{i=1}^N C_i\psi_\alpha(x,x_i)=:\overline{I_\alpha}(x)$*. Moreover, for all* $i\in\{1,\dots,N\}$*,* $a(x_i)=0$*,* $R\left(x_i,\overline{I_g}(x_i)\right)=0$*, and for all* $x\in\mathbb{R}^d$*,* $m(x,x_i,\overline{I_d}(x))=0$*.*

Proof. According to the previous lemma, for all $\phi\in\mathscr{C}_c^1(\mathbb{R}^d)$,

$$
\sum_{i=1}^N C_i a(x_i)\cdot\nabla\phi(x_i)+\sum_{i=1}^N C_iR\left(x_i,\overline{I_g}(x_i)\right)\phi(x_i)+\sum_{i=1}^N C_i\int_{\mathbb{R}^d}m\left(x,x_i,\overline{I_d}(x)\right)\phi(x)dx=0.
$$

Let $\varepsilon>0$. For any non-negative function $\phi_\varepsilon\in\mathscr{C}_c^1(\mathbb{R}^d)$ with a support on $\mathbb{R}^d\backslash\bigcup_{i=1}^N B(x_i,\varepsilon)$, we have that

$$
\int_{\mathbb{R}^d}\sum_{i=1}^N C_i m\left(x,x_i,\overline{I_d}(x)\right)\phi_\varepsilon(x)dx=0,
$$

which proves that $x\mapsto\sum_{i=1}^N C_i m\left(x,x_i,\overline{I_d}(x)\right)$ is 0 on $\mathbb{R}^d\backslash\bigcup_{i=1}^N B(x_i,\varepsilon)$. Since m is a non-negative function, for all $i\in\{1,...N\}$, $x\mapsto m\left(x,x_i,\overline{I_d}(x)\right)$ is 0 as well on $\mathbb{R}^d\backslash\bigcup_{i=1}^N B(x_i,\varepsilon)$, and therefore on $\mathbb{R}^d$, since the result holds for any $\varepsilon>0$. Hence,

$$
\sum_{i=1}^N C_i a(x_i)\cdot\nabla\phi(x_i)+\sum_{i=1}^N C_iR\left(x_i,\overline{I_g}(x_i)\right)\phi(x_i)=0,
$$

for any $\phi\in\mathscr{C}_c^1(\mathbb{R}^d)$. By choosing ϕ_j^1 such that

$$
\phi_j^1(x_i)=\delta_{ij}\text{ and }\nabla\phi_j^1(x_i)=0,\ i,j\in\{1,...N\},
$$

where δ_{ij} represents the Kronecker delta, one proves that $R\left(x_i, \overline{I_g}(x_i)\right) = 0$ for all $i \in \{1, ...N\}$. Finally, by choosing ϕ^2_{ij} such that

$$\nabla \phi^2_{jl}(x_i) = \delta_{ij} e_l,\ i, j \in \{1, ...N\},\ l \in \{1, ...d\},$$

where $\{e_l\}_{l=1}^d$ represents the euclidean basis in $\mathbb{R}^d$, one proves that $a(x_i) = 0$, for any $i \in \{1, \ldots, N\}$. $\square$

Limit identification and asymptotic preserving approximations

In some cases, it is possible to guarantee the existence of a limit for v^h and identify it.
Assume that there exists $\hat{x} \in \mathbb{R}^d$ an asymptotically stable equilibrium for the ODE '$\dot{x} = a(x)$' and that there exists $C, \delta > 0$ such that

$$\forall y \in \mathrm{supp}(n^0), t \geq 0, \quad \|X(t, y) - \hat{x}\| \leq Ce^{-\delta t}. \tag{5.75}$$

Moreover, let us assume that there exist positive values D, I^m and I^M such that

$$R(x, I^m) \geqslant 0, \quad R(x, I^M) \leqslant 0 \text{ and } \partial_I R(x, I) \leqslant -D, \quad \forall x \in \mathrm{supp}(v^0). \tag{5.76}$$

Then, we can compute the limit of v^h when t goes to $+\infty$, whatever the value of m, as stated in the following proposition.

Proposition 5.4. *Let us assume that supp(v^0) is a compact set such that hypotheses (5.75) and (5.76) hold. Then, v^h converges to $\hat{\rho}_h\, \delta_{\hat{x}}$ in the weak sense in the space of Radon measures, where $\hat{\rho}_h$ is the unique solution of*

$$R(\hat{x}, \psi_g(\hat{x}, \hat{x})\hat{\rho}_h) = 0.$$

The following lemma, proved in Appendix 5.6.4, is required in the proof of this result.

Lemma 5.11. *Let $u \in \mathscr{C}^2(\mathbb{R}_+, \mathbb{R})$ be a bounded function, and let us assume that there exist $p_0 > 0$, $p : \mathbb{R}_+ \to \mathbb{R}_+$ a function which satisfies $p \geqslant p_0$ and $B \in L^1(\mathbb{R}_+)$ an integrable function such that*

$$\ddot{u}(t) \geqslant -p(t)\dot{u}(t) + B(t).$$

Then, there exists $u_\infty \in \mathbb{R}$ such that $\lim\limits_{t\to+\infty} u(t) = u_\infty$.

Proof of Proposition 5.4. Given that supp$(v^0) \cup K_{xy}$ is a compact set which is strictly contained in the basin of attraction of $\hat{x}$ we will have the existence of $\mathscr{J}_h^0 \subset \mathscr{J}_h$, with $|\mathscr{J}_h^0| < +\infty$ such that $\nu^h(t) \not\equiv 0$ only for $i \in \mathscr{J}_h^0$.
Let us denote, for all $i \in \mathscr{J}_h^0$, $\alpha_i(t) := \nu_i(t) w_i(t)$, and

$$\rho_h(t) := \sum_{i \in \mathscr{J}_h} \alpha_i(t) = \sum_{i \in \mathscr{J}_h^0} \alpha_i(t).$$

Let us note that, according to the hypotheses on a, for all $i \in \mathscr{J}_h^0$, $x_i(t)$ converges to $\hat{x}$. Thus,

$$v^h(t) - \rho_h(t)\delta_{\hat{x}} \underset{t\to+\infty}{\rightharpoonup} 0.$$

Hence, it only remains to prove that ρ_h converges to the expected limit. According to the definition of ρ_h,

$$\dot\rho_h(t) = \sum_{i\in\mathscr{J}_h^0} R(x_i(t), I_g(t,x_i(t)))\alpha_i(t) + \underbrace{\sum_{i,j\in\mathscr{J}_h^0} w_i(t) m(x_i(t), x_j(t), I_d(t,x_i(t)))\alpha_j(t)}_{:=\varepsilon(t)}. \tag{5.77}$$

According to hypothesis 5.75, there exist $C, \delta > 0$ such that

$$\|x_i(t) - \hat x\| \leqslant Ce^{-\delta t}, \quad \forall i \in \mathscr{J}_h^0, \quad \forall t \geqslant 0.$$

Thus, for all $i \in \mathscr{J}_h^0$,

$$w_i(t) = e^{\int_0^t \mathrm{div}\ a(x_i(s))ds} w_i(0) \leqslant \underbrace{e^{\int_0^t \mathrm{div}\ a(x_i(s)) - \mathrm{div}\ a(\hat x)ds}}_{\leqslant \tilde C} e^{\mathrm{div}\ a(\hat x)t} w_i(0),$$

which proves, since div $a(\hat x) < 0$, that there exist C', δ' such that for all $t \geqslant 0$,

$$0 \leqslant \max_{i\in\mathscr{J}_h^0} w_i(t) \leqslant C' e^{-\delta' t}. \tag{5.78}$$

In particular, it proves that $t \mapsto \varepsilon(t)$ defined in (5.77), converges to zero with an exponential speed, since

$$|\varepsilon(t)| \leqslant \max_{i\in\mathscr{J}_h} w_i(t) |\mathscr{J}_h^0| \|m\|_{L^\infty} \rho_h(t),$$

and ρ_h is bounded, according to Theorem 5.4.

Moreover, according to the hypothesis on ψ_g, $\underline{\psi_g}\,\rho(t) \leqslant I_g(t, x_i(t)) \leqslant \|\psi_g\|_\infty\, \rho(t)$. Thus, according to the hypotheses (5.76) on R, the relation

$$\dot\rho_h(t) \geqslant \left(\min_{i\in\mathscr{J}_h^0} \left(R(x_i(t), I_g(t, x_i(t)))\right)\right) \rho_h(t)$$

implies that, as soon as ρ_h becomes small, and so does I_g, ρ_h becomes increasing, which proves that ρ_h is lower bounded by a positive constant. Moreover, since ρ_h and ε are bounded, and

$$\begin{aligned}\|\dot\alpha\|_{\ell_1} = \sum_{i\in\mathscr{J}_h^0} |\dot\alpha(t)| &\leqslant \sum_{i\in\mathscr{J}_h^0} \left(2|\mathrm{div}\ a(x_i(t))| + |R(x_i(t), I_g(t, x_i(t)))|\right)\alpha_i(t) + \varepsilon(t)\\ &\leqslant \left(2\|a\|_{W^{1,\infty}(\mathbb{R}^d)} + \overline{R}\right)\rho_h(t) + \varepsilon(t),\end{aligned}$$

where

$$\overline{R} := \max_{t\in\mathbb{R}, i\in\mathscr{J}_h^0} |R(x_i(t), I_g(t, x_i(t)))|,$$

then $\|\dot\alpha\|_{\ell_1}$ is also bounded.

Now, let us prove that ρ_h satisfies the equality of Lemma 5.11. First, for $\gamma \in \{g, d\}$, we compute

$$\begin{aligned}\left|\frac{d}{dt} I_\gamma(t, x_i(t)) - \psi_\gamma(\hat x, \hat x)\dot\rho_h(t)\right| \leqslant &\left|\sum_{j\in\mathscr{J}_h^0}\Big(a(x_i(t))\partial_x\psi_\gamma(x_i(t), x_j(t)) + a(x_j(t))\partial_y\psi_\gamma(x_i(t), x_j(t))\Big)\alpha_j(t)\right|\\ &+ \left|\sum_{j\in\mathscr{J}_h^0}\left(\psi_\gamma(x_i(t), x_j(t)) - \psi_\gamma(\hat x, \hat x)\right)\dot\alpha_j(t)\right|\\ \leqslant &2\max_{i\in\mathscr{J}_h^0} |a(x_i(t))| \|\psi_\gamma\|_{W^{1,\infty}(\mathbb{R}^d)}\rho_h(t) + \max_{i,j\in\mathscr{J}_h^0} |\psi_\gamma(x_i(t), x_j(t)) - \psi_\gamma(\hat x, \hat x)| \|\dot\alpha\|_{\ell_1}.\end{aligned}$$

Since hypothesis (5.75), is satisfied, the functions $t \mapsto \max_{i \in \mathcal{J}_h^0} |a(x_i(t))|$ and $t \mapsto \max_{i,j \in \mathcal{J}_h^0} |\psi_\gamma(x_i(t), x_j(t)) - \psi_\gamma(\hat{x}, \hat{x})|$ converge to zero with an exponential speed. Since $|\mathcal{J}_h^0| < +\infty$, this proves that, for $\gamma \in \{g, d\}$,

$$\sum_{i \in \mathcal{J}_h^0} \left| \frac{d}{dt} I_\gamma(t, x_i(t)) - \psi_\gamma(\hat{x}, \hat{x}) \dot{\rho}_h(t) \right| = O(e^{-\delta t}), \tag{5.79}$$

for a certain $\delta > 0$.

Thus, by differentiating (5.77), we get

$$\begin{aligned} \ddot{\rho}_h(t) = &\sum_{i \in \mathcal{J}_h^0} a(x_i(t)) \partial_x R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) + \sum_{i \in \mathcal{J}_h^0} \left(\frac{d}{dt} I_g(t, x_i(t)) \right) \partial_I R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) \\ &+ \sum_{i \in \mathcal{J}_h^0} \dot{\alpha}_i(t) R(x_i(t), I_g(t, x_i(t))) + \sum_{i,j \in \mathcal{J}_h^0} w_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \dot{\alpha}_j(t) \\ &+ \sum_{i,j \in \mathcal{J}_h^0} \dot{w}_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t) + \sum_{i,j \in \mathcal{J}_h^0} w_i(t) \frac{d}{dt} m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t), \end{aligned}$$

where

i) $\sum_{i \in \mathcal{J}_h^0} a(x_i(t)) \partial_x R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) = O(e^{-\delta t})$, since $\max_{i \in \mathcal{J}_h} |a(x_i(t))|$ converges to zero with exponential speed, and $\partial_x R$ and ρ_h are bounded,

ii) according to (5.79), $\sum_{i \in \mathcal{J}_h^0} \left(\frac{d}{dt} I_g(t, x_i(t)) \right) \partial_I R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) = -p(t) \dot{\rho}_h(t) + O(e^{-\delta t})$, where

$$p(t) = -\psi_g(\hat{x}, \hat{x}) \sum_{i \in \mathcal{J}_h^0} \partial_I R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) \geqslant D \underline{\psi_g} \min_{t \geqslant 0} \rho_h(t) > 0,$$

iii)

$$\begin{aligned} \sum_{i \in \mathcal{J}_h^0} \dot{\alpha}_i(t) R(x_i(t), I_g(t, x_i(t))) = &\underbrace{\sum_{i \in \mathcal{J}_h^0} R(x_i(t), I_g(t, x_i(t)))^2 \alpha_i(t)}_{:= P(t) \geqslant 0} \\ &+ \underbrace{\sum_{i,j \in \mathcal{J}_h^0} w_i(t) R(x_i(t), I_g(t, x_i(t))) \alpha_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t)))}_{= O(e^{-\delta t})}, \end{aligned}$$

where the relation for the second term was proved thanks to the bound

$$\left| \sum_{i,j \in \mathcal{J}_h^0} w_i(t) R(x_i(t), I_g(t, x_i(t))) \alpha_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t))) \right| \leqslant \overline{M}\, \overline{R} \rho_h(t) |\mathcal{J}_h^0| \max_{i \in \mathcal{J}_h} w_i(t),$$

the boundedness of ρ_h and inequality (5.78), iv)

$$\sum_{i,j\in\mathscr{I}_h^0} w_i(t)m(x_i(t),x_j(t),I_d(t,x_i(t)))\dot\alpha_j(t)+\sum_{i,j\in\mathscr{I}_h^0}\dot w_i(t)m(x_i(t),x_j(t),I_d(t,x_i(t)))\alpha_j(t)$$
$$+\sum_{i,j\in\mathscr{I}_h^0} w_i(t)\frac{d}{dt}m(x_i(t),x_j(t),I_d(t,x_i(t)))\alpha_j(t)=O(e^{-\delta t}),$$

since for all $t\geq 0$,

$$\left|\sum_{i,j\in\mathscr{I}_h^0} w_i(t)m(x_i(t),x_j(t),I_d(t,x_i(t)))\dot\alpha_j(t)\right|\leqslant \overline{M}|\mathscr{I}_h^0|\|\dot\alpha\|_{\ell^1}\max_{i\in\mathscr{I}_h^0} w_i(t),$$

$$\left|\sum_{i,j\in\mathscr{I}_h}\dot w_i(t)m(x_i(t),x_j(t),I_d(t,x_i(t)))\alpha_j(t)\right|\leqslant \overline{M}\rho_h(t)|\mathscr{I}_h^0|\|a\|_{W^{1,\infty}(\mathbb{R}^d)}\max_{i\in\mathscr{I}_h^0} w_i(t),$$

and

$$\left|\sum_{i,j\in\mathscr{I}_h} w_i(t)\frac{d}{dt}m(x_i(t),x_j(t),I_d(t,x_i(t)))\alpha_j(t)\right|\leqslant \overline{M}\left(\|a\|_{L^\infty(\mathbb{R})}+\left|\frac{d}{dt}I_g(t,x_i(t))\right|\right)\rho_h(t)|\mathscr{I}_h^0|\max_{i\in\mathscr{I}_h^0} w_i(t),$$

where $\frac{d}{dt}I_d(t,x_i(t))$ is bounded thanks to (5.79). For these three inequalities, we conclude with (5.78).

Hence, $\ddot\rho_h(t)\geqslant -p(t)\dot\rho_h(t)+O(e^{-\delta t})$. According to Lemma 5.11, ρ_h has a limit when t goes to ∞, which we denote $\hat\rho_h$. Since $\dot\rho_h(t)=\sum_{i\in\mathscr{I}_h^0} R(x_i(t),I_g(t,x_i(t)))\alpha_i(t)+\varepsilon(t)$, which converges to $R(\hat x,\psi_g(\hat x,\hat x)\hat\rho_h)\hat\rho_h$, we deduce that $R(\hat x,\psi_g(\hat x,\hat x)\hat\rho_h)=0$, since $\hat\rho_h>0$. □

When $m\equiv 0$ and under the same assumptions for the remaining coefficients of the problem as in Proposition 5.4, we are able to identify the limit of v and prove that it coincides with the limit of v^h. According to Lemma 5.9, this ensures that v_ε^h is an asymptotic preserving approximation of v.

Theorem 5.6. *Let us assume that there exists $\hat x\in\mathbb{R}^d$ which is an asymptotically stable equilibrium for the ODE $\dot x=a(x)$ such that hypothesis (5.75) holds. We assume as well that $m\equiv 0$ and that hypotheses (5.76) hold. Then, v converges to $\hat\rho\,\delta_{\hat x}$ in the weak sense in the space of Radon measures, where $\hat\rho$ is the unique solution of*

$$R(\hat x,\psi_g(\hat x,\hat x)\hat\rho)=0.$$

Consequently, v_ε^h is an asymptotic preserving approximation of v.

Proof. Let us recall that

$$\rho(t)=\int_{\mathbb{R}^d} v(t,x)dx.$$

We recall as well that, the function a being only dependent of x, the characteristic lines $X_t(x):=X(t,x)$ satisfy $X_t^{-1}(x):=X^{-1}(t,x)=X_{-t}(x)$ and $X_t(X_s(x))=X_{t+s}(x)$. By using the fact that, for any $x\in\mathbb{R}^d$,

$$v(t,x)=v^0(X_{-t}(x))e^{\mathscr{G}(0,t,x)},$$

where
$$\mathscr{G}(s,t,x) := \int_s^t R(X_{\tau-t}(x), (I_g v)(\tau, X_{\tau-t}(x))) - \operatorname{div}\, a(X_{\tau-t}(x)) d\tau,$$
and that, by hypothesis, $K := \operatorname{supp}(v^0)$ is a compact set included in the basin of attraction of $\hat{x}$, one proves that $\operatorname{supp}(v(t,\cdot))$ is the image of $\operatorname{supp}(v^0)$ by $X_t(\cdot)$. Since $X_t(y)$ converges to $\hat{x}$ for all $y \in \operatorname{supp}(v^0)$, we prove that $K_t := \operatorname{supp}(v(t,\cdot)) = X_t(\operatorname{supp}(v^0)) = X_t(K)$ is a compact set included in the basin of attraction of $\hat{x}$, for all $t \geqslant 0$. By (5.75), there exist $C > 0$ and $\delta > 0$ such that
$$\|X_t(y) - \hat{x}\| \leqslant Ce^{-\delta t}, \quad \forall y \in K, \quad \forall t \geqslant 0.$$

Let $\phi \in \mathscr{C}_c(\mathbb{R}^d)$. By definition of ρ and by using the change of variable '$x = X_t(y)$' we get
$$\begin{aligned}\left| \int_{\mathbb{R}^d} \phi(x) v(t,x) - \rho(t)\phi(\hat{x}) \right| &\leqslant \int_{K_t} |\phi(x) - \phi(\hat{x})| v(t,x) dx \\ &= \int_K |\phi(X_t(y)) - \phi(\hat{x})| v(t, X_t(y)) e^{\int_0^t \operatorname{div}\, a(X_t(y))} dy \\ &\leqslant \max_{y\in K} |\phi(X_t(y)) - \phi(\hat{x})| \int_K v(t, X_t(y)) e^{\int_0^t \operatorname{div}\, a(X_t(y))} dy \\ &= \max_{y\in K} |\phi(X_t(y)) - \phi(\hat{x})| \rho(t).\end{aligned}$$

Since ρ is bounded, this proves that
$$v(t,\cdot) - \rho(t)\delta_{\hat{x}} \underset{t\to+\infty}{\rightharpoonup} 0. \tag{5.80}$$

Hence, it only remains to prove that ρ converges to the expected limit. We see that
$$\dot{\rho}(t) = \int_{\mathbb{R}^d} R(x, (I_g v)(t,x)) v(t,x) dx. \tag{5.81}$$

First, let us note that ρ has a positive lower bound. Indeed, according to the hypothesis on ψ_g, for all $x \in \mathbb{R}^d$, $\underline{\psi_g}\, \rho(t) \leqslant (I_g v)(t,x) \leqslant \|\psi_g\|_{L^\infty}\, \rho(t)$. Thus,
$$\dot{\rho}(t) \geqslant \left(\min_{x\in K_t} (R(x, (I_g v)(t,x)) \right) \rho(t).$$

Hence, as soon as $\rho(t)$ becomes small, and so does $I_g(t,x)$, ρ becomes increasing, which proves that ρ is lower bounded by a positive constant. We get an upper bound for $|\dot{\rho}(t)|$ from the relation
$$|\dot{\rho}(t)| \leqslant \left(\max_{x\in K_t} (R(x, (I_g v)(t,x)) \right) \rho(t),$$

and the boundedness of $\rho(t)$.
We introduce now the function
$$\tilde{v}(t,y) = v(t, X_t(y)) e^{\int_0^t \operatorname{div}\, a(X_s(y)) ds},$$

which satisfies
$$\int_K \tilde{v}(t,y) dy = \int_{K_t} v(t,x) dx = \rho(t).$$

Moreover,

$$\partial_t \tilde{v}(t,y) = R(X_t(y),(I_g v)(t,X_t(y)))\tilde{v}(t,y). \tag{5.82}$$

Before proving that ρ satisfies the equality of Lemma 5.11, we observe that the following relation holds for all $y \in K$:

$$\begin{aligned}\left|\frac{d}{dt}(I_g v)(t,X_t(y)) - \psi_g(\hat{x},\hat{x})\dot{\rho}(t)\right| \leqslant & \left|\int_{K_t} a(X_t(y))\cdot\nabla_x\psi_g(X_t(y),z)v(t,z)dz\right| \\ & + \left|\int_{K_t}\big(\psi_g(X_t(y),z)-\psi_g(\hat{x},\hat{x})\big)\partial_t v(t,z)dz\right|. \end{aligned}\tag{5.83}$$

From the equation satisfied by v, we see that

$$\begin{aligned} &\int_{K_t}\big(\psi_g(X_t(y),z)-\psi_g(\hat{x},\hat{x})\big)\partial_t v(t,z)dz \\ =&\int_{K_t}\nabla_z\psi_g(X_t(y),z)a(z)v(t,z)dz \\ &+\int_{K_t}\big(\psi_g(X_t(y),z)-\psi_g(\hat{x},\hat{x})\big)R(z,(I_g v)(t,z))v(t,z)dz \\ =&\int_K \nabla_z\psi_g(X_t(y),X_t(\bar{z}))a(X_t(\bar{z}))\tilde{v}(t,\bar{z})d\bar{z} \\ &+\int_K\big(\psi_g(X_t(y),X_t(\bar{z}))-\psi_g(\hat{x},\hat{x})\big)R(X_t(\bar{z}),(I_g v)(t,X_t(\bar{z})))\tilde{v}(t,\bar{z})d\bar{z}\end{aligned}$$

which allows us to conclude that

$$\begin{aligned}\left|\int_{K_t}\big(\psi_g(X_t(y),z)-\psi_g(\hat{x},\hat{x})\big)\partial_t v(t,z)dz\right| \leqslant & \max_{\bar{z}\in K}\|a(X_t(\bar{z}))\|\|\psi_g\|_{W^{1,\infty}(\mathbb{R}^d)}\rho(t) \\ & + \max_{y,\bar{z}\in K}|\psi_g(X_t(y),X_t(\bar{z}))-\psi_g(\hat{x},\hat{x})|\,\overline{R}\rho(t).\end{aligned}$$

Using this relation in (5.83) gives

$$\begin{aligned}\left|\frac{d}{dt}(I_\gamma v)(t,X_t(y)) - \psi_\gamma(\hat{x},\hat{x})\dot{\rho}(t)\right| \leqslant & 2\max_{\bar{z}\in K}\|a(X_t(\bar{z}))\|\|\psi_\gamma\|_{W^{1,\infty}(\mathbb{R}^d)}\rho(t) \\ & + \max_{y,\bar{z}\in K}|\psi_\gamma(X_t(y),X_t(\bar{z}))-\psi_\gamma(\hat{x},\hat{x})|\,\overline{R}\rho(t),\end{aligned}$$

which proves, according to hypothesis (5.75), that for all $y \in K$

$$\frac{d}{dt}I_g(t,X(t,y)) = \psi_g(\hat{x},\hat{x})\dot{\rho}(t) + O(e^{-\delta t})), \tag{5.84}$$

for a certain $\delta > 0$.

Differentiating (5.81) we get

$$\begin{aligned}\ddot{\rho}(t) =&\frac{d}{dt}\left(\int_{K_t} R(x,(I_g v)(t,x))v(t,x)dx\right)\\ =&\frac{d}{dt}\left(\int_{K} R(X_t(y),(I_g v)(t,X_t(y)))\tilde{v}(t,y)dy\right)\\ =&\int_K a(X_t(y))\cdot\nabla_x R(X_t(y),(I_g v)(t,X_t(y)))\tilde{v}(t,y)dy\\ &+\int_K \frac{d}{dt}(I_g v)(t,X_t(y))\partial_I R(X_t(y),(I_g v)(t,X_t(y)))\tilde{v}(t,y)dy\\ &+\int_K R(X_t(y),(I_g v)(t,X_t(y)))\partial_t\tilde{v}(t,y)dy.\end{aligned}$$

Let us note that

i) Since $a(X_t(y))$ converges uniformly to zero with an exponential speed, then

$$\int_{\mathbb{R}^d} a(X_t(y))\cdot\nabla_x R\left(X_t(y),(I_g v)(t,X_t(y))\right)\tilde{v}(t,y))dy = O(e^{-\delta t}),$$

thanks to the boundedness of $\nabla_x R$ and $\rho(t)$.

ii) According to (5.84),

$$\int_K \frac{d}{dt}(I_g v)(t,X_t(y))\partial_I R(X_t(y),(I_g v)(t,X_t(y)))\tilde{v}(t,y)dy = -p(t)\dot{\rho}(t)+O(e^{-\delta t}),$$

with

$$p(t) := -\psi_g(\hat{x},\hat{x})\int_K \partial_I R(X_t(y),(I_g v)(t,X_t(y)))\tilde{v}(t,y)dy \geqslant \psi_g(\hat{x},\hat{x})D\min_{t\geqslant 0}\rho(t) > 0.$$

iii) Directly from (5.82),

$$\begin{aligned}\int_K R(X_t(y),(I_g v)(t,X_t(y)))\partial_t\tilde{v}(t,y)dy =& \int_K\Big(R(X_t(y),(I_g v)(t,X_t(y)))\Big)^2\tilde{v}(t,y)dy\\ =&: P(t)\geqslant 0.\end{aligned}$$

Thus, $\ddot{\rho}(t)\geqslant -p(t)\dot{\rho}(t)+O(e^{-\delta t})$, hence, ρ converges, thanks to Lemma 5.11.

Recalling (5.80), $v(t,\cdot)$ thus converges to $\hat{\rho}\delta_{\hat{x}}$, where $\hat{\rho}$ is the limit of ρ. We conclude, according to Proposition 5.3, that $\hat{\rho}$ satisfies the expected equality.
Having proved that v and v^h share the same limit is enough then to conclude, thanks to Lemma 5.9, that v_ε^h is an asymptotic preserving approximation. □

If m is not 0, under very specific hypotheses over its support, we can extend the result of Theorem 5.6. The explanation behind this is simple: as long as the population is composed of traits that are not prone to mutatations, it will evolve as in the case where mutations are not possible at all.

Theorem 5.7. *Let us assume that there exists $\hat{x} \in \mathbb{R}^d$ which is an asymptotically stable equilibrium for the ODE $\dot{x} = a(x)$ such that (5.75) holds. Moreover, let us assume that $supp\,(v^0) \cup K_x$ is a compact set such that there exist $C', \delta' > 0$ such that*

$$\forall y \in supp\left(v^0\right) \cup K_x, \quad t \geq 0, \quad \|X(t,y) - \hat{x}\| \leq C' e^{-\delta' t},$$

that $\bigcup_{s \geqslant 0} (X_s(supp(v^0) \cup K_x)) \cap K_y = \emptyset$ and that hypothesis (5.76) holds. Then, v converges to $\hat{\rho}\,\delta_{\hat{x}}$ in the weak sense in the space of Radon measures, where $\hat{\rho}$ is the unique solution of

$$R(\hat{x}, \psi_g(\hat{x}, \hat{x})\hat{\rho}) = 0.$$

Consequently, v_ε^h is an asymptotic preserving approximation of v.

Proof. By using the fact that, for any $x \in \mathbb{R}^d$,

$$v(t,x) = v^0(X_{-t}(x))e^{\mathscr{G}(0,t,x)} + \int_0^t \int_{\mathbb{R}^d} m(X_{s-t}(x), z, (I_d v)(s, X_{s-t}(x)))v(s,z)dz e^{\mathscr{G}(s,t,x)} ds,$$

where

$$\mathscr{G}(s,t,x) := \int_s^t R(X_{\tau-t}(x), (I_g v)(\tau, X_{\tau-t}(x))) - \text{div}\, a(X_{\tau-t}(x)) d\tau,$$

we observe that $\text{supp}(v(t,x)) \subset X_t(\text{supp}(v^0)) \cup \bigcup_{0 \leqslant s \leqslant t} X_s(K_x) \subset \bigcup_{s \geqslant 0} (X_s(\text{supp}(v^0) \cup K_x))$, therefore

$$\begin{aligned} v(t,x) =& v^0(X_{-t}(x))e^{\mathscr{G}(0,t,x)} + \int_0^t \int_{\mathbb{R}^d} m(X_{s-t}(x), z, (I_d v)(s, X_{s-t}(x)))v(s,z)dz e^{\mathscr{G}(s,t,x)} ds \\ =& v^0(X_{-t}(x))e^{\mathscr{G}(0,t,x)} + \int_0^t \int_{\text{supp}(v(t,x)) \cap K_y} m(X_{s-t}(x), z, (I_d v)(s, X_{s-t}(x)))v(s,z)dz e^{\mathscr{G}(s,t,x)} ds \\ =& v^0(X_{-t}(x))e^{\mathscr{G}(0,t,x)}. \end{aligned}$$

Therefore, we can replicate the proof for the case $m \equiv 0$. □

The result of Theorem 5.6 does not generalize when $\text{supp}\,(v^0)$ is not strictly contained in the basin of attraction of x_s, as shown in the following result:

Proposition 5.5. *Let us consider the one-dimensional PDE*

$$\begin{cases} \partial_t v(t,x) + \nabla_x \cdot (a(x)v(t,x)) = (r(x) - \rho(t))v(t,x), \\ \rho(t) = \int_{\mathbb{R}} v(t,x)dx, \\ n(0,\cdot) = n^0(\cdot), \end{cases} \tag{5.85}$$

which is a particular case of (5.19), and let us assume that there exist $x_u < x_s$ such that $a(x_u) = a(x_s) = 0$, $a'(x_u) > 0$, $a'(x_s) < 0$, $\text{supp}(n^0) \subset [x_u, x_s]$, $n^0(x_u) = 0$, and that there exists $\alpha > 0$ such that ${n^0}'(x) = \underset{x \to x_u^+}{O}((x - x_u)^\alpha)$ and that $r(x_u) - (1+\alpha)f'(x_u) > r(x_s)$. Then, v_ε^h is not an asymptotic preserving approximation of v.

Proof. The long-time behaviour of the solution of (5.85) has been studied in detail in [30], and it has been proved, under the hypotheses of Proposition 5.5, that v converges to a function in L^1. Let us now compute the limit of v^h. Since $n^0(x_u) = 0$, we can assume, without loss of generality, that for all $i \in \mathscr{J}_h$, $x_i^0 \in (0,1]$. Thus, since $a > 0$ on (x_u, x_s), for all $t \geqslant 0$, $x_i(t)$ converges to x_s. As seen in the proof of Proposition 5.4, v^h therefore converges to $r(x_s)\delta_{x_s}$, and v_ε^h is therefore not a asymptotic preserving approximation of v. □

5.5 Simulations

In this section, we present some simulations obtained with the particle method developed throughout the chapter. In Figure 5.1, we deal with the non-local advection equation presented in [174], which writes

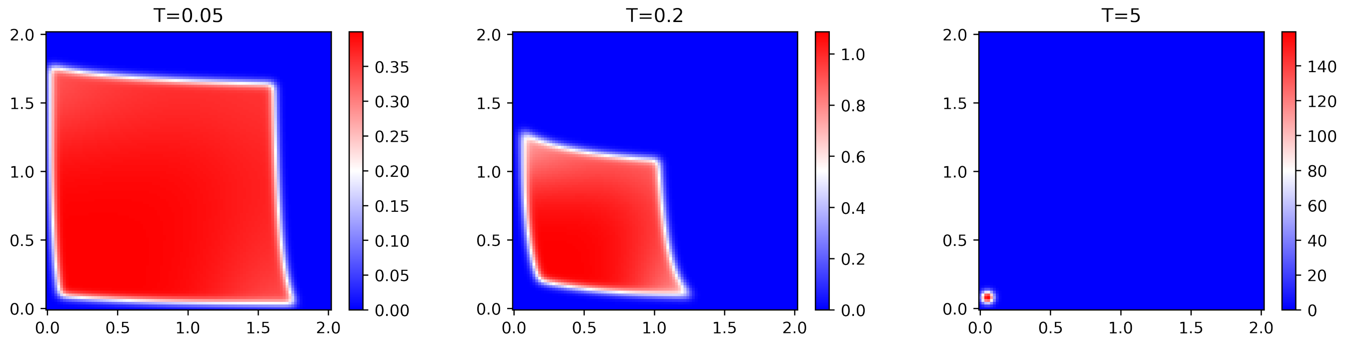
$$\partial_t v(t,x) + \nabla_x \left(a(t, I_1 v(t,x), I_2 v(t,x)\right) = 0, \quad x \in \mathbb{R}^d, t \geq 0, \tag{5.86}$$

with $I_j v(t,x) = \int_{\mathbb{R}^2} x_j v(t,x)dx$, for $j \in \{1,2\}$. Note that this equation is not exactly a particular case of (5.1), since there are two non-local terms involved for advection, but the particle method can straightforwardly be adapted to this case. As in this chapter, we show that, depending on the parameters, the solution of this PDE can converge to a single Dirac mass, to a sum of two Dirac masses, or to the sum of four Dirac masses. The parameters used for the simulations are the same as the one detailed in Figures 9, 10 and 11 of [174].

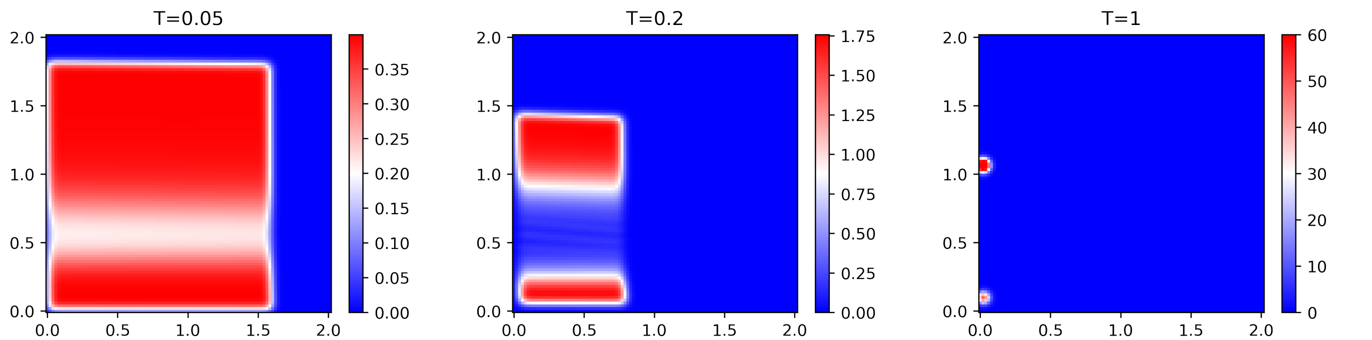
Figure 5.2 illustrates different scenarios for the equation

$$\begin{cases} \partial_t v(t,x) + \nabla_x \left(a(x)v(t,x)\right) = \left(r(x) - \rho(t)\right) n(t,x), \\ \rho(t) = \int_{\mathbb{R}} v(t,x)dx, \\ n(0,x) = n^0(x), \end{cases} \tag{5.87}$$

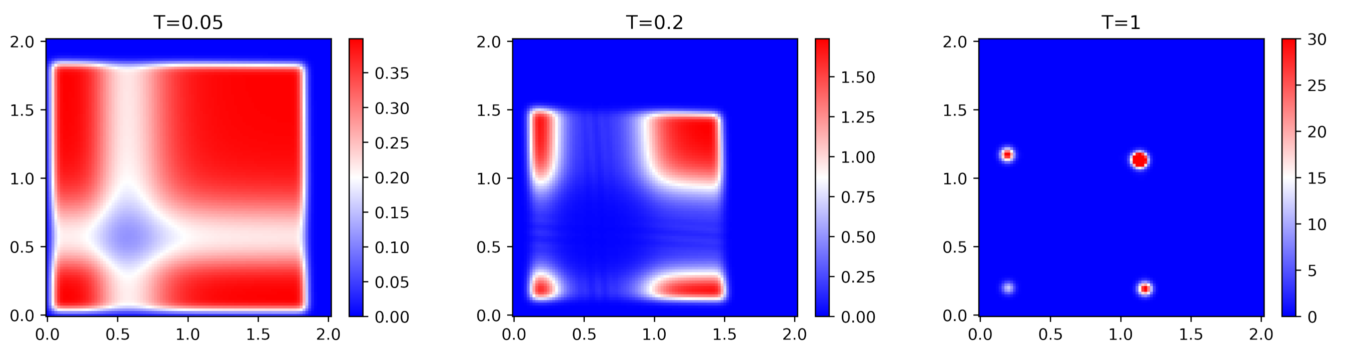
with $a(x) = x(1-x)$, and an initial solution supported in $[0,1]$. This equation has been studied in [30] where it has been proved that its solution can either converge to a function in L^1, (which depends on the initial condition), or to a Dirac mass on 1, depending on the functions r and n^0.



(a) Monostability

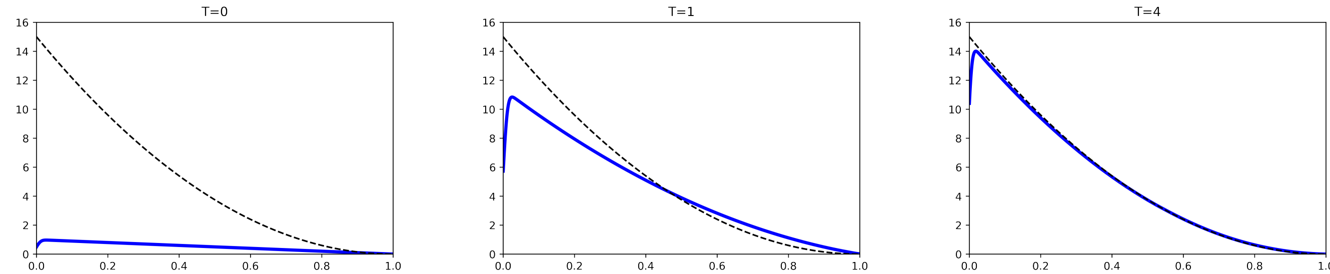


(b) Bistability

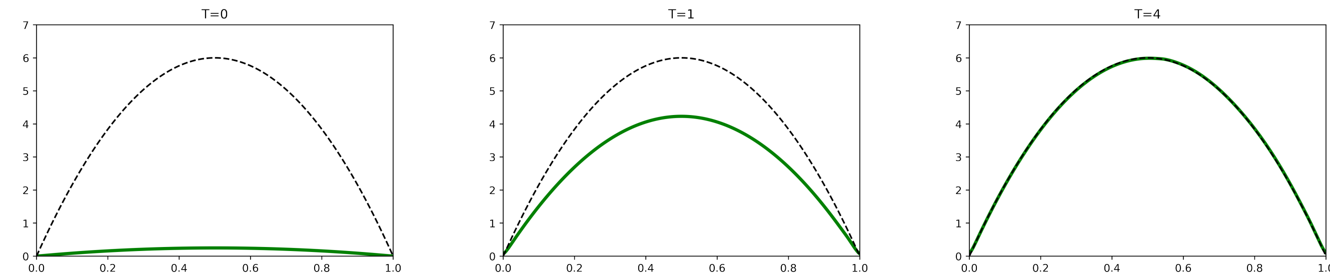


(c) Quadstability

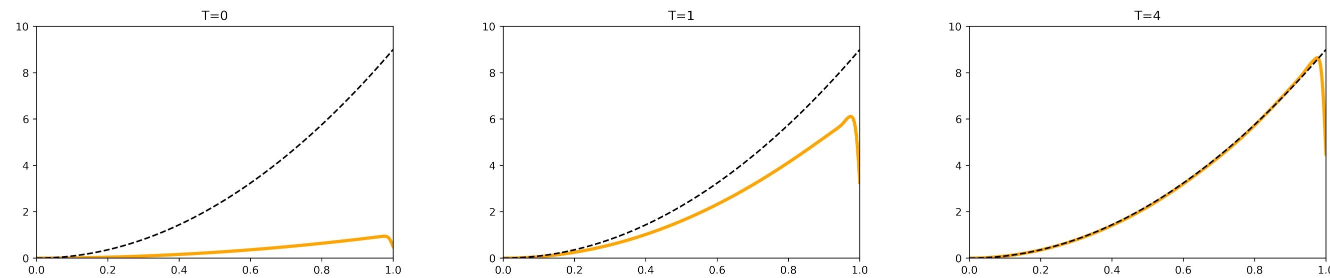
Figure 5.1: The three possible regimes of convergence for equation (5.86), obtained with the particle method. The lines (a), (b) and (c) respectively show the convergence to a single Dirac mass, two Dirac masses and four Dirac masses, and have been obtained by choosing the parameters of Figures 9, 10 and 11 of [174]. In the three cases, we have chosen $N = 100$, $h = 2./100$, $\varepsilon = h^{0.8}$, and the cut off function φ is a Gaussian.



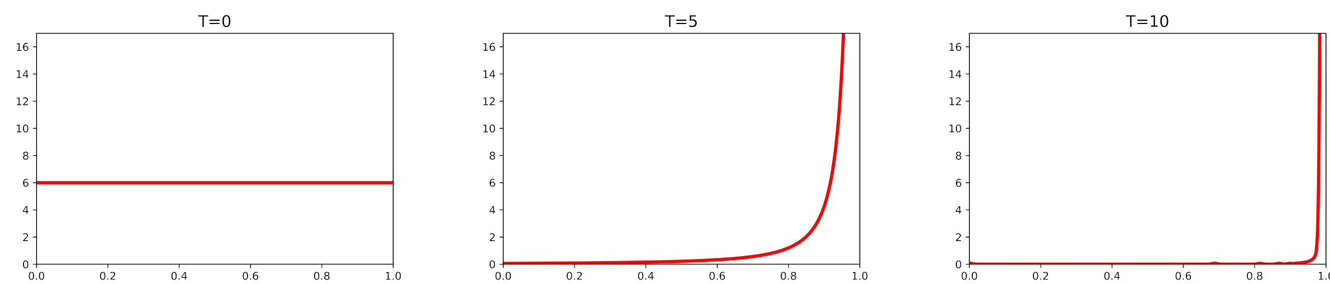
(a) Solution of (5.87) at different time steps, with $n^0(x) = 1 - x, r(x) = 6 - 4x$.



(b) Solution of (5.87) at different time steps, with $n^0(x) = x(1-x), r(x) = 6 - 4x$.



(c) Solution of (5.87) at different time steps, with $n^0(x) = x^2, r(x) = 6 - 4x$.



(d) Solution of (5.87) at different time steps, with $n^0(x) = 6, r(x) = 6 - 0.5x$.

Figure 5.2: Different possible regimes of convergence for the solution of (5.87). The first three lines (green, blue and orange curves), show the convergence to a function in L^1, which can be explicitly computed (see [30]), and is represented by a black dashed line. Note that the limit function is different when the initial condition changes. The last line (red curves) shows the convergence to a Dirac mass in 1. In the four cases, we have chosen $a(x) = x(1-x)$, $N = 5000$, $h = \frac{1}{N}$, $\varepsilon = \sqrt{h}$ and the cut-off function φ is a Gaussian.

5.6 Appendices

5.6.1 Proof of the results over the characteristics

In order to prove results which involve the use of absolute values, we introduce a smooth re-normalizing sequence of functions. Consider a sequence of smooth positive functions β_ε satisfying $\beta_\varepsilon(0) = 0$, $\beta_\varepsilon(s) > 0$ for all $s \neq 0$, $\beta_\varepsilon(s) \leqslant |s|$, $\beta_\varepsilon(s) \to |s|$ almost everywhere, $|\dot\beta_\varepsilon(s)| \leqslant 1$ and $s\dot\beta_\varepsilon(s) \to |s|$ almost everywhere. For example we may choose

$$\beta_\varepsilon(s) = \begin{cases} -s - \varepsilon(1 - \frac{2}{\pi}) & \text{if} \quad s \leqslant -\varepsilon, \\ \frac{2\varepsilon}{\pi}\left(1 - \cos(\frac{\pi}{2\varepsilon}s)\right) & \text{if} \quad -\varepsilon < s < \varepsilon, \\ s - \varepsilon(1 - \frac{2}{\pi}) & \text{if} \quad s \geqslant \varepsilon. \end{cases} \tag{5.88}$$

Proof of Lemma 5.1. We introduce the notation

$$\Delta X_j(t) := X^j_{u_1}(t, y_1) - X^j_{u_2}(t, y_2).$$

For all $t \in [0, T]$, the function

$$U_\varepsilon(t) := \sum_{j=1}^d \beta_\varepsilon(\Delta X_j(t))$$

satisfies then the relation

$$\begin{aligned}
\dot U_\varepsilon(t) =& \sum_{j=1}^d \dot\beta_\varepsilon(\Delta X_j(t))\left(a_j(t, X_{u_1}(t,y_1), (I_a u_1)(t, X_{u_1}(t,y_1))) - a_j(t, X_{u_2}(t,y_2), (I_a u_2)(t, X_{u_2}(t,y_2)))\right) \\
\leqslant& \sum_{j=1}^d \left(\|a_j\|_{W^{1,\infty}_x} \sum_{i=1}^d |\Delta X_i(t)| + \|a_j\|_{W^{1,\infty}_I} |(I_a u_1)(t, X_{u_1}(t,y_1)) - (I_a u_2)(t, X_{u_2}(t,y_2))| \right) \\
\leqslant& d\|a\|_{W^{1,\infty}_{x,I}} \left(\sum_{i=1}^d |\Delta X_i(t)| + |(I_a u_1)(t, X_{u_1}(t,y_1)) - (I_a u_2)(t, X_{u_2}(t,y_2))| \right) \\
\leqslant& d\|a\|_{W^{1,\infty}_{x,I}} \left((1 + \|\psi_a\|_{W^{1,\infty}_x L^\infty_y} \|u_1\|) \sum_{i=1}^d |\Delta X_i(t)| + \|\psi_a\|_{L^\infty} \|u_1 - u_2\|_{L^1(\mathbb{R}^d)} \right).
\end{aligned}$$

Integrating between 0 and t we get

$$\begin{aligned}
&U_\varepsilon(t,y) - U_\varepsilon(0,y) \\
&\leqslant d\|a\|_{L^\infty_t W^{1,\infty}_{x,I}} \left((1 + \|\psi_a\|_{L^\infty_{t,y} W^{1,\infty}_x} \|u_1\|) \int_0^t \sum_{i=1}^d |\Delta X_i(t)| ds + \|\psi_a\|_{L^\infty} \int_0^t \|u_1 - u_2\|_{L^1(\mathbb{R}^d)} ds \right).
\end{aligned}$$

Taking the limit when ε goes to 0 and applying Grönwall's lemma we get the desired result. □

Proof of Lemma 5.2. We explicitly give the proof for $k = 1$. The proof for higher values of k follows the same ideas.

Thanks to the hypothesis over a, the function X_u is one time differentiable with respect to y, and directly from (5.20) we get the system of equations

$$\begin{cases} \partial_{y_i} \dot{X_u}(t,y) = J_a(t, X_u(t,y)) \partial_{y_i} X_u(t,y), \quad t \in [0,T], \\ \partial_{y_i} X_u(0,y) = e_i, \end{cases} \tag{5.89}$$

for all values of $i \in \{1, \dots, d\}$, where

$$[J_a(t,x)]_{ij} := \partial_{x_i} a_j(t,x,(I_a u)(t,x)) + \partial_I a_j(t,x,(I_a u)(t,x)) \int_{\mathbb{R}^d} \partial_{x_i} \psi_a(t,x,y) u(t,y) dy, \tag{5.90}$$

is the Jacobian matrix of the function $a(t,x,(I_a u)(t,x))$ and the e_i represent the canonical basis of $\mathbb{R}^d$. The function

$$V_\varepsilon(t,y) := \sum_{i,j=1}^{n} \beta_\varepsilon(\partial_{y_i} X_u^j(t,y))$$

satisfies then

$$\dot{V_\varepsilon}(t,y) = \sum_{i,j=1}^{n} \dot{\beta_\varepsilon}(\partial_{y_i} X_u^j(t,y)) \sum_{k=1}^{d} [J_a(t, X_u(t,y))]_{kj} \, \partial_{y_i} X_u^k(t,y)$$

and consequently

$$\dot{V_\varepsilon}(t,y) \leqslant d \|a\|_{W^{1,\infty}_{x,I}} (1 + \|\psi_a\|_{W^{1,\infty}_x L^\infty_y} \|u\|) \sum_{i,j=1}^{n} |\partial_{y_i} X_u^j(t,y)|.$$

Integrating between 0 and t, we obtain the relation

$$V_\varepsilon(t,y) - V_\varepsilon(0,y) \leqslant d\, \tilde{\alpha}_1 \int_0^t \sum_{i,j=1}^{n} |\partial_{y_i} X_u^j(s,y)| ds$$

with $\tilde{\alpha}_1 := \sup_{t\in[0,T]} \|a\|_{W^{1,\infty}_{x,I}} (1 + \|\psi_a\|_{W^{1,\infty}_x L^\infty_y} \|u\|)$, which after taking the limit when ε goes to 0 leads to

$$\sum_{i,j=1}^{n} |\partial_{y_i} X_u^j(t,y)| \leqslant d + d\tilde{\alpha}_1 \int_0^t \sum_{i,j=1}^{n} |\partial_{y_i} X_u^j(s,y)| ds.$$

We obtain (5.21) thanks to Grönwall's lemma.
In order to prove (5.22) we adopt the notation

$$\Delta \partial_{y_k} X^j(t) := \partial_{y_k} X_{u_1}^j(t,y_1) - \partial_{y_k} X_{u_2}^j(t,y_2)$$

and define

$$[J_{X_u}(t,y)]_{jk} := \partial_{y_k} X_u^j(t,y),$$

which satisfies the relation

$$\dot{J_{X_u}}(t,y) = J_a(t, X_u(t,y)) J_{X_u}(t,y), \; J_{X_u}(0,y) = I_d,$$

with $J_a(t,x)$ as defined on (5.90). Consequently, we have that

$$D_\varepsilon(t) := \sum_{j=1}^{d} \sum_{k=1}^{d} \beta_\varepsilon \left(\Delta \partial_{y_k} X^j(t) \right)$$

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

$$\begin{aligned}
\dot{D}_\varepsilon(t) &= \sum_{j=1}^{d}\sum_{k=1}^{d} \dot{\beta}_\varepsilon\left(\Delta\partial_{y_k}X^j(t)\right)\sum_{i=1}^{d}\left([J_a(t,X_{u_1}(t,y_1))]_{ij}\partial_{y_k}X^i_{u_1}(t,y_1) - [J_a(t,X_{u_2}(t,y_2))]_{ij}\partial_{y_k}X^i_{u_2}(t,y_2)\right)\\
&\leqslant \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{i=1}^{d}\left|[J_a(t,X_{u_1}(t,y_1))]_{ij}\partial_{y_k}X^i_{u_1}(t,y_1) - [J_a(t,X_{u_2}(t,y_2))]_{ij}\partial_{y_k}X^i_{u_2}(t,y_2)\right|\\
&\leqslant \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{i=1}^{d}\left|[J_a(t,X_{u_1}(t,y_1))]_{ij}\right|\left|\Delta\partial_{y_k}X^i(t)\right|\\
&\quad + \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{i=1}^{d}\left|[J_a(t,X_{u_1}(t,y_1))]_{ij} - [J_a(t,X_{u_2}(t,y_2))]_{ij}\right|\left|\partial_{y_k}X^i_{u_2}(t,y_2)\right|.
\end{aligned}$$

From the hypothesis over a and ψ_a we see that

$$|[J_a(t,X_{u_1}(t,y_1))]_{ij}| \leqslant \|a\|_{W^{1,\infty}_{x,I}}(1+\|\psi_a\|_{W^{1,\infty}_x L^\infty_y}\|u_1\|).$$

Furthermore, from the definition of $J_a(t,x)$ we conclude that there exists a constant C, depending only on $\|a\|_{W^{2,\infty}_{x,I}}$, $\|\psi_a\|_{W^{2,\infty}_x L^\infty_y}$ and $\|u_i\|$ such that

$$\begin{aligned}
|[J_a(t,X_{u_1}(t,y_1))]_{ij} - [J_a(t,X_{u_2}(t,y_2))]_{ij}| &\leqslant C\Big(\sum_{j=1}^{d}|X^j_{u_1}(t,y_1) - X^j_{u_2}(t,y_2)| + \|u_1-u_2\|_{L^1(\mathbb{R}^d)}\Big)\\
&\leqslant C(|y_1-y_2| + \|u_1-u_2\|_1 + \|u_1-u_2\|_{L^1(\mathbb{R}^d)}),
\end{aligned}$$

where we have used the results from Lemma 5.1 on the second line.
Putting all estimates together, we conclude that there exist constants C_1 and C_2 only depending on $\|a\|_{W^{2,\infty}_{x,I}}$, $\|\psi_a\|_{W^{2,\infty}_x L^\infty_y}$ and $\|u_i\|$, such that

$$\dot{D}_\varepsilon(t) \leqslant C_1\sum_{j=1}^{d}\sum_{k=1}^{d}|\partial_{y_k}X^j_{u_1}(t,y_1) - \partial_{y_k}X^j_{u_2}(t,y_2)| + C_2(|y_1-y_2| + \|u_1-u_2\|_1 + \|u_1-u_2\|_{L^1(\mathbb{R}^d)}).$$

Integrating in time, using Grönwall's lemma and taking the limit when ε goes to zero, we obtain (5.22). □

Proof of Lemma 5.3. We explicitly give the proof for $k=1$. The proof for higher values of k follows the same ideas.
Differentiating once each component of the equality $X_u(t,X_u^{-1}(t,x)) = x$ with respect to each of the variables x_k, we obtain the family of relations

$$\sum_{i=1}^{d}\partial_{y_i}X^j_u(t,X_u^{-1}(t,x))\partial_{x_k}\left(X_u^{-1}\right)^i(t,x) = \delta_{jk},\ \ j,k=1,\dots,d,$$

where δ_{jk} represents the Kronecker's delta. Written in matrix form, this equality reads

$$J_{X_u}(t,X_u^{-1}(t,x))J_{X_u^{-1}}(t,x) = I_d.$$

It is known that the matrix $J_{X_u}(t,y)$ is invertible for all values of x, furthermore, its determinant is given by the expression

$$det(J_{X_u}(t,y)) = e^{\int_0^t \nabla_x \cdot a(s,y,(I_a u)(s,y)) + \partial_I a(s,y,(I_a u)(s,y)) \cdot \int\limits_{\mathbb{R}^d} \nabla_x \psi_a(s,y,z) u(s,z) dz ds} \geqslant c_T > 0$$

for all values of $t \in [0,T]$ and $y \in \mathbb{R}^d$.
We conclude by writing

$$J_{X_u^{-1}}(t,x) = J_{X_u}^{-1}(t, X_u^{-1}(t,x)),$$

and noticing that all of the components of $J_{X_u}^{-1}(t, X_u^{-1}(t,x))$ are a combination of sums and multiplications of the components of $J_{X_u}(t, X_u^{-1}(t,x))$, divided by $det(J_{X_u}(t,y))$. The bound (5.21) from Lemma 5.2, together with the lower bound for the determinant of $J_{X_u}(t,y)$ gives the bound (5.23) over the components of $J_{X_u^{-1}}(t,x)$. $\square$

Proof of Lemma 5.4. We explicitly give the proof for $k=1$. The proof for higher values of k follows the same ideas.
Differentiating with respect to t the relation $X_{u_i}(t, X_{u_i}^{-1}(t,x)) = x$, for $i = 1,2$, we see that

$$a(t,x,(I_a u_i)(t,x)) + J_{X_{u_i}}(t, X_{u_i}^{-1}(t,x)) \dot{X}_{u_i}^{-1}(t,x) = 0,$$

which gives

$$\dot{X}_{u_i}^{-1}(t,x) = -J_{X_{u_i}}^{-1}(t, X_{u_i}^{-1}(t,x)) a(t,x,(I_a u_i)(t,x)). \tag{5.91}$$

From now on we adopt the notations

$$\begin{aligned} A^i(t,x) &:= a(t,x,(I_a u_i)(t,x)), \\ K^i(t,x) &:= -J_{X_{u_i}}^{-1}(t, X_{u_i}^{-1}(t,x)), \\ \Delta X_j^{-1}(t,x) &:= \left(X_{u_1}^{-1}\right)^j(t,x) - \left(X_{u_2}^{-1}\right)^j(t,x). \end{aligned}$$

The function

$$W_\varepsilon(t,x) := \sum_{j=1}^d \beta_\varepsilon(\Delta X_j^{-1}(t,x))$$

satisfies the relation

$$\begin{aligned} \dot{W}_\varepsilon(t,x) &= \sum_{j=1}^d \sum_{k=1}^d \dot{\beta}_\varepsilon(\Delta X_j^{-1}(t,x)) \left(K_{jk}^1(t,x) A_k^1(t,x) - K_{jk}^2(t,x) A_k^2(t,x)\right) \\ &\leqslant \sum_{j=1}^d \sum_{k=1}^d |K_{jk}^1(t,x) A_k^1(t,x) - K_{jk}^2(t,x) A_k^2(t,x)|. \end{aligned}$$

From Lemma 5.3 we know that all components of K^i are uniformly bounded by a constant only depending on T and $\|u_i\|$. We deduce from the hypothesis over a that the components of A^i are uniformly bounded by $\tilde{a} := \|a\|_{L^\infty}$. Therefore

$$\dot{W}_\varepsilon(t,x) \leqslant d\tilde{C}(T, \|u_1\|) \sum_{k=1}^d |A_k^1(t,x) - A_k^2(t,x)| + \tilde{a} \sum_{j=1}^d \sum_{k=1}^d |K_{jk}^1(t,x) - K_{jk}^2(t,x)|.$$

The function a being L-Lipschitz with respect to the I variable, we have that, for all values of k

$$|A_k^1(t,x) - A_k^2(t,x)| \leqslant L\|\psi_a\|_{L^\infty}\|u_1 - u_2\|_{L^1(\mathbb{R}^d)}.$$

On the other hand, from the definition of K^i and Lemma 5.2 we conclude that there exists C, depending on T, $\|a\|_{W^{2,\infty}_{x,I}}$, $\|\psi_a\|_{W^{2,\infty}_x L^\infty_y}$ and $\|u_i\|$ such that

$$|K_{jk}^1(t,x) - K_{jk}^2(t,x)| \leqslant \begin{cases} 0, & \text{if} \quad \partial_I a = 0, \\ C\left(\sum_{j=1}^d |\left(X_{u_1}^{-1}\right)^j(t,x) - \left(X_{u_2}^{-1}\right)^j(t,x)| + \|u_1 - u_2\|_1\right), & \text{if} \quad \partial_I a \neq 0. \end{cases}$$

Putting everything together, integrating between 0 and t, taking the limit when ε goes to 0 and applying Grönwall's lemma we get (5.24). □

5.6.2 Existence of solution for a system of ODEs with infinitely many unknowns and equations

Proof of Lemma 5.7. For all $u \in X_h^T$, there exists a sequence of elements $u^\delta \in X_h^T$ such that:

1) $\mathscr{K}_h^\delta := \{k \in \mathscr{J}_h : u_k^\delta \neq 0\}$ has a finite number of elements.

2)
$$\lim_{\delta \to 0} \|u - u^\delta\|_{1,h} = 0.$$

We denote $K^\delta := |\mathscr{K}_h^\delta|$ and notice that the system

$$\dot{x_k^\delta}(t) = A_{u^\delta,w}(t, x_k^\delta), \;\; x_k^\delta(0) = x_k^0, \tag{5.92}$$

is composed of a coupled system of K^δ equations and unknowns (corresponding to those $k \in \mathscr{K}_h^\delta$), and an uncoupled infinite number of equations, corresponding to those $k \notin \mathscr{K}_h^\delta$. Therefore, thanks to the classic Cauchy-Lipschitz theory, the system (5.92) has a unique solution $x_k^\delta \in \mathscr{C}^1([0,T])$, $k \in \mathscr{J}_h$.
We claim that for all values of k, the sequence $x_k^{\delta_1} - x_k^{\delta_2}$ is a Cauchy sequence in $\mathscr{C}^1([0,T])$, therefore is has a limit that we will call $x_k(t)$, which is solution to (5.49).
We first remark that $x^{\delta_1} - x^{\delta_2} \in Y_h^T$ due to the fact that $|x_k^\delta(t) - x_k^0| \leqslant \|a\|_{L^\infty} T$ for all values of k and δ.
Consider now β_ε as defined in (5.88), then

$$\begin{aligned} \dot{\beta_\varepsilon}(x_k^{\delta_1} - x_k^{\delta_2}) &\leqslant |A_{u^{\delta_1},w}(t, x_k^{\delta_1}) - A_{u^{\delta_2},w}(t, x_k^{\delta_2})| \\ &\qquad \|a\|_{W^{1,\infty}_x}|x_k^{\delta_1} - x_k^{\delta_2}| + \|a\|_{W^{1,\infty}_I}|I_a(t, x_k^{\delta_1}, u^{\delta_1}, w) - I_a(t, x_k^{\delta_1}, u^{\delta_2}, w)|. \end{aligned}$$

Noticing that

$$\begin{aligned} |I_a(t, x_k^{\delta_1}, u^{\delta_1}, w) - I_a(t, x_k^{\delta_1}, u^{\delta_2}, w)| &= |\sum_{j \in \mathscr{J}_h}\left(u_j^{\delta_1}(t)\psi_a(t, x_k^{\delta_1}, x_j^{\delta_1}(t)) - u_j^{\delta_2}(t)\psi_a(t, x_k^{\delta_2}, x_j^{\delta_2}(t))\right)w_j(t)| \\ &\leqslant (\|\psi_a\|_{L^\infty}\|u^{\delta_1} - u^{\delta_2}\|_{1,h} + \|u^{\delta_2}\|_{1,h}\|\psi_a\|_{W^{1,\infty}_{x,y}}\|x^{\delta_1} - x^{\delta_2}\|_{\infty,h})\|w\|_{\infty,h}, \end{aligned}$$

we deduce the existence of two constants[3], C_1 and C_2, only depending on a, ψ_a, u and w, such that

$$\dot{\beta}_\varepsilon(x_k^{\delta_1} - x_k^{\delta_2}) \leqslant C_1\|x^{\delta_1} - x^{\delta_2}\|_{\infty,h} + C_2\|u^{\delta_1} - u^{\delta_2}\|_{1,h}.$$

Integrating between 0 and t, taking the maximum over k and t and using Grönwall's lemma, we conclude that there exists a constant C_T, only depending on T and the coefficients of the problem, such that

$$\|x^{\delta_1} - x^{\delta_2}\|_{\infty,h} \leqslant C_T\|u^{\delta_1} - u^{\delta_2}\|_{1,h}.$$

Proceeding in a similar way with the absolute value of $\dot{x}^{\delta_1} - \dot{x}^{\delta_2}$ we obtain that

$$\|\dot{x}^{\delta_1} - \dot{x}^{\delta_2}\|_{\infty,h} \leqslant C_T\|u^{\delta_1} - u^{\delta_2}\|_{1,h}.$$

Recalling that u^δ is a Cauchy sequence on X_h^T, then so it is x_k^δ on $\mathscr{C}^1([0,T])$, for each k.
Let $x := \{x_k\}_{k\in\mathscr{I}_h}$ be the limit of x^δ when δ goes to 0. With a simple continuity argument we conclude that x is a solution of (5.49) over $[0,T]$. The uniqueness can be obtained by assuming the existence of two solutions, deriving the equation satisfied by the difference and using Grönwall's lemma to conclude that they have to be equal. □

5.6.3 A result from approximation theory

As mentioned before, Lemma 5.8 is a direct corollary of Lemma 8 in [184], that we recall here

Lemma 5.12. *Let $k > d$ an integer. Assume that*

$$a \in (L^\infty(0,T;W^{k+1,\infty}(\mathbb{R}^d)))^d.$$

Then, there exists a constant $C > 0$ such that for all functions $\varphi \in W^{k,p}(\mathbb{R}^d)$, $1 \leqslant p \leqslant +\infty$, and $t \in [0,T]$,

$$\|\varphi - \sum_{i\in\mathscr{I}_h} w_i(t)\varphi(x_i(t))\delta(\cdot - x_i(t))\|_{W^{-k,p}(\mathbb{R}^d)} \leqslant Ch^k\|\varphi\|_{W^{k,p}(\mathbb{R}^d)}.$$

Given that for fixed functions $\nu \in X_h^T$ and $w \in Y_h^T$ we have the inclusion

$$A_{\nu,w} : (t,x) \mapsto a(t,x,I_a(t,x,\nu,w)) \in (L^\infty(0,T;W^{k+1,\infty}(\mathbb{R}^d)))^d,$$

then Lemma 5.12 holds true as well for the values of x_i obtained in Section 5.3.

Proof of Lemma 5.8. We recall that $W^{-k,1}(\mathbb{R}^d)$ is the dual space of $W^{k,\infty}(\mathbb{R}^d)$. Thus, for any $\psi \in W^{-k,1}(\mathbb{R}^d)$ we have

$$\|\psi\|_{-k,1} = \sup_{f\in W^{k,\infty}(\mathbb{R}^d)} \frac{|\langle\psi, f\rangle|}{\|f\|_{k,\infty}}.$$

[3]Notice that in order to obtain the estimate over I_a, we used the hypothesis of differentiability over both variables on ψ_a

Since the function $f \equiv 1$ belongs to $W^{k,\infty}(\mathbb{R}^d)$, and has norm equal to 1 in this space, we get for all $\varphi \in W^{k,p}(\mathbb{R}^d)$

$$\left|\int_{\mathbb{R}^d} \varphi(x)dx - \sum_{i\in\mathcal{J}_h} w_i(t)\varphi(x_i(t))\right| = \left|\langle \varphi - \sum_{i\in\mathcal{J}_h} w_i(t)\varphi(x_i(t))\delta(\cdot - x_i(t)), 1\rangle\right|$$
$$\leqslant \|\varphi - \sum_{i\in\mathcal{J}_h} w_i(t)\varphi(x_i(t))\delta(\cdot - x_i(t))\|_{-k,1}.$$

We conclude by applying Lemma 5.12 with $p = 1$. □

5.6.4 Proofs of convergence results from ODE theory

This appendix is dedicated to the proofs of lemma 5.11, used in subsection 5.4.2. In order to prove this lemma, we use the following result:

Lemma 5.13. *Let $\alpha > 0$ and $B \in L^1(\mathbb{R}_+)$. Then, all the solutions of the ODE*

$$\dot{u}(t) = -\alpha u(t) + B(t)$$

are in $L^1(\mathbb{R}^+)$.

Proof of lemma 5.13. The solution of this ODE is explicitly given by

$$u(t) = u(0)e^{-\alpha t} + \int_0^t e^{-\alpha(t-s)} B(s)ds.$$

Hence,

$$\int_0^{+\infty} |u(t)|dt \leqslant |u(0)| \int_0^{+\infty} e^{-\alpha t}dt + \iint_{\mathbb{R}_+^2} e^{-\alpha(t-s)}|B(s)|\mathbb{1}_{\{s\leqslant t\}}dsdt.$$

With the change of variables $y = s, z = t - s$, we get

$$\iint_{\mathbb{R}_+^2} e^{-\alpha(t-s)}|B(s)|\mathbb{1}_{\{s\leqslant t\}}dsdt \leqslant \int_0^{+\infty} |B(y)|dy \int_0^{+\infty} e^{-\alpha z}dz,$$

which concludes the proof.

□

Proof of Lemma 5.11. First, let us note that if $\dot{u}$ is a BV function, *i.e.* if $\int_0^{+\infty} |\dot{u}(t)|dt < +\infty$, then u is a Cauchy function, and thus converges. Let us denote $v := \dot{u}$. Since u is assumed to be bounded, and $|v| = v + 2v^-$, where v^- denotes the negative part of v, it is enough to prove that $v^- \in L^1(\mathbb{R}^d)$. By hypothesis,

$$\dot{v}(t) = -p(t)v(t) + P(t) + B(t),$$

which implies that

$$\dot{\overline{v^-}}(t) \leqslant -p_0\, v^-(t) + B(t).$$

We conclude, according to lemma 5.13, that $v^- \in L^1(\mathbb{R}^d)$, which implies that u converges. □

Chapter 6

The parabolic-parabolic Keller-Segel equation with variable rescaling parameter

In this chapter we present some estimates on the parabolic-parabolic Keller-Segel equation which has been developed in order to tackle two problems:

(1) an improved spectral analysis in the radially symmetric case in order to describe blowing up solutions in the critical 8π mass case in the spirit of what has been done for the parabolic-elliptic Keller-Segel equation in [39, 71, 72] (see also [40]);

(2) a description of the longtime self-similar behavior of solution in the sub-critical mass case without radially symmetric assumption in the spirit of what has been done in the radially symmetric case in [48].

Although the first project has been up to now unsuccessful, we believe that the material we have developed can be useful for future works. A series of obtained intermediate estimates are thus presented below. The solution to the second problem is presented in details in the next chapter.

6.1 The parabolic-parabolic Keller-Segel equation and re-scaling parameters

6.1.1 The PPKS equation and the re-scaling parameters

The original parabolic-parabolic Keller-Segel (PPKS in short) equation in the plane writes

$$\left\{ \begin{array}{lll} \partial_t F & = & \Delta F + \mathrm{div}(-F \nabla U) \quad \text{in} \quad (0,\infty) \times \mathbb{R}^2 \\ \varepsilon\, \partial_t U & = & \Delta U + F \quad \text{in} \quad (0,\infty) \times \mathbb{R}^2, \end{array} \right. \tag{6.1}$$

complemented with an initial condition

$$F(0,\cdot) = F_0 \geq 0 \quad \text{and} \quad U(0,\cdot) = U_0 \geq 0 \quad \text{in} \quad \mathbb{R}^2. \tag{6.2}$$

Here $t \geq 0$ is the time variable, $x \in \mathbb{R}^2$ is the space variable, $F = F(t,x) \geq 0$ stands for the *mass density of cells* while $U = U(t,x) \geq 0$ is the *chemo-attractant concentration* and $\varepsilon > 0$ is a constant.

We refer to the work [46] as well as to the reviews [185, 186] and the references quoted therein for biological motivation and mathematical introduction.

In short, the KS equation models a cells population which is subject to two inverse mechanisms:

- a brownian motion (responsible to the diffusion term ΔF in the first equation of (6.1)) modeling the fact that any cell change of direction and move in a completely erratic way and which global effect is to spread out the population all over the plane $\mathbb{R}^2$;

- an aggregation mechanism (responsible to the drift term $\nabla(-F\nabla U)$ in the first equation of (6.1)) modeling the fact that cells have a tendency to follow the gradient lines of the chemo-attractant, which is itself produced and diffused according to the second equation in (6.1). That mechanism has a concentration effect, which is quite strong due to the fact that the associated *interaction kernel* is singular.

From a mathematical point of view, both mechanisms are almost at the same order, and that makes the rigorous analysis of the model particularly difficult and interesting.

When assuming the initial mass sub-critical, namely

$$\varrho := \langle F_0 \rangle = \int_{\mathbb{R}^2} F_0(x)dx < 8\pi,$$

one can show that the solution (F, U) to (6.1) is global in time and mass preserving

$$\langle F(t,\cdot)\rangle \equiv \varrho < 8\pi, \quad \forall\, t \geq 0.$$

We introduce the self similar variables (f, u) through the change of variables

$$F(t,x) = R(t)^{-2}\mathscr{F}(\log R(t), R(t)^{-1}x) \text{ and } U(t,x) = \mathscr{U}(\log R(t), R(t)^{-1}x),$$

where $R(t) = \sqrt{1+2t}$. This re-scaled functions solve the system

$$\left\{ \begin{array}{rcl} \partial_t \mathscr{F} &=& \Delta\mathscr{F} + \operatorname{div}(x\mathscr{F} - \mathscr{F}\nabla\mathscr{U}) \\ \partial_t \mathscr{U} &=& \dfrac{1}{\varepsilon}(\Delta\mathscr{U} + \mathscr{F}) + x\cdot\nabla\mathscr{U}, \end{array} \right. \tag{6.3}$$

with

$$\langle F(t,\cdot)\rangle = \langle F_0\rangle = \varrho < 8\pi.$$

It is shown in [47,187,188] the existence, uniqueness (up to a constant for the second unknown), radially symmetric property and smoothness of an associated steady state $(\mathscr{G}_{\varepsilon,\varrho}, \mathscr{V}_{\varepsilon,\varrho})$ for a given sub-critical mass, so that

$$\left\{ \begin{array}{rcl} 0 &=& \Delta\mathscr{G}_{\varepsilon,\varrho} + \operatorname{div}(x\mathscr{G}_{\varepsilon,\varrho} - \mathscr{G}_{\varepsilon,\varrho}\nabla\mathscr{V}_{\varepsilon,\varrho}), \quad \langle \mathscr{G}_{\varepsilon,\varrho}\rangle = \varrho, \\ 0 &=& \dfrac{1}{\varepsilon}(\Delta\mathscr{V}_{\varepsilon,\varrho} + \mathscr{G}_{\varepsilon,\varrho}) + x\cdot\nabla\mathscr{V}_{\varepsilon,\varrho}, \end{array} \right. \tag{6.4}$$

for any $\varrho \in (0, 8\pi)$ and $\varepsilon > 0$. Such steady states are parameterized according to the mass ϱ. This implies that there exists a one to one relation between ϱ and, for example, the value $\mathscr{G}_{\varepsilon,\varrho}(0)$, which allows us to redefine the aforementioned parameterization. Let us set

$$\mu = \frac{8}{\mathscr{G}_{\varepsilon,\varrho}(0)}, \quad f(t,x) = \mu\mathscr{F}(\mu t, \sqrt{\mu}x), \quad v(t,x) := \mathscr{V}(\mu t, \sqrt{\mu}x). \tag{6.5}$$

An easy computation shows that (f, v) is solution to the modified parabolic-parabolic Keller-Segel equation in self-similar variables

$$\begin{cases} \partial_t f &= \Delta f + \mathrm{div}(\mu x f - f \nabla u) \\ \partial_t u &= \dfrac{1}{\varepsilon}(\Delta u + f) + \mu x \cdot \nabla u, \end{cases} \tag{6.6}$$

and in the sequel we will focus on this problem. In fact, regardless of the value of μ, such a re-scaling will always yield a solution for (6.6). The motivation behind this specific choice for the parameter can be found in [39] where the same choice allows for a refined study of the steady states for the sub-critical parabolic-elliptic case, see also [40, 42, 71, 72].

6.1.2 Self-similar profiles

We denote by $Q = Q^\mu_\varepsilon$ and $P = P^\mu_\varepsilon$ the associated self-similar profiles defined as the stationary solutions to the Keller-Segel system (6.6), that is

$$\begin{cases} 0 = \Delta Q + \mathrm{div}(\mu x Q - Q \nabla P) \\ 0 = \Delta P + Q + \varepsilon \mu x \cdot \nabla P. \end{cases} \tag{6.7}$$

The functions Q and P should be equivalently defined by $Q(x) := \mu \mathscr{G}_{\varepsilon,\varrho}(\sqrt{\mu} x)$ and $P(x) := \mathscr{U}_{\varepsilon,\varrho}(\sqrt{\mu} x)$, where μ is still defined by the first equation in (6.5). It is worth noticing here that whatever is the value of $\varepsilon \in [0, 1\varepsilon_0)$, we may establish that

$$Q^\mu_\varepsilon \to Q_0, \quad P^\mu_\varepsilon \to P_0, \quad \text{as} \quad \mu \to 0,$$

where (Q_0, P_0) is defined by

$$Q_0(x) := \frac{8}{(1 + |x|^2)^2}, \quad \Delta P_0 = Q_0$$

and it is a stationary solution to parabolic-elliptic Keller-Segel system (6.1) (corresponding to the case $\varepsilon = 0$) and mass $\langle Q_0 \rangle = 8\pi$ (and with infinite second moment). The relevancy of the choice of μ in (6.5) comes from the fact that $\langle Q^\mu_\varepsilon \rangle \to 8\pi$ the critical mass as $\mu \to 0$. In some sense the system (6.6) with sub-critical mass $\langle Q^\mu_\varepsilon \rangle$ is asymptotically close to the system (6.1) with parameter $\varepsilon = 0$ and critical mass 8π. It turns out that an accurate analysis of the first sub-critical problem may be helpful for the analysis of the second critical one and such an idea is developed in [39] when $\varepsilon = 0$. That is also the strategy we try to follow here in order to understand the possible infinite time blow up behaviour of solutions of the Parabolic-Parabolic Keller-Segel system (6.6) (when thus $\varepsilon > 0$) for the critical mass. We have not been able to perform the complete analogical work when $\varepsilon > 0$ as what is known when $\varepsilon = 0$. Nevertheless, we present in the sequel of this note the very first steps of such an analysis which we believe can be interesting for their own and for the realization of the full program. In particular in section 6.2 and as a first step, we establish families of estimates on the profiles functions $(Q^\mu_\varepsilon, P^\mu_\varepsilon)$ with quantitative dependence on the parameters $\mu, \varepsilon \geq 0$.

6.1.3 The linearized self-similar PPKS equation

We introduce the perturbation (g, v) defined by

$$f = Q + g, \quad u = P + v,$$

where $Q = Q^\mu_\varepsilon$ and $P = P^\mu_\varepsilon$ verify (6.7). If (f,u) is a solution to (6.6) then (g,v) satisfies the system

$$\begin{cases} \partial_t g = \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla v) - \mathrm{div}(g\nabla v) \\ \partial_t v = \dfrac{1}{\varepsilon}(\Delta v + g) + \mu x \cdot \nabla v, \end{cases} \tag{6.8}$$

and reciprocally.

We are next interested on the linearized equation around a re-scaled self-similar profile

$$\left\{ \begin{array}{rcl} \partial_t g &=& \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla v) \\ \partial_t v &=& \dfrac{1}{\varepsilon}(\Delta v + g) + \mu x \cdot \nabla v. \end{array} \right.$$

Let us define the Laplace kernel in the plane

$$\kappa(z) := -\frac{1}{2\pi}\log|z|, \quad \mathscr{K}(z) := \nabla\kappa(z) = -\frac{1}{2\pi}\frac{z}{|z|^2}, \tag{6.9}$$

so that $\omega := \kappa * \Omega$ is a solution to the Laplace equation

$$-\Delta\omega = \Omega \quad \text{in} \quad \mathbb{R}^2.$$

Next defining

$$w := v - \kappa * g,$$

the equation on w is

$$\partial_t w = \frac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + \mu x \cdot \nabla\kappa * g - \nabla\kappa * [\nabla g + \mu x g - g\nabla P - Q\nabla\kappa * g - Q\nabla w].$$

In fact, by using that

$$x\cdot\nabla\kappa * g - \nabla\kappa * (xg) \simeq \int \frac{(x-y)}{|x-y|^2}\big\{x\,g(y) - y\,g(y)\big\}\,dy \simeq \langle g\rangle$$

and $\langle g\rangle = 0$, the second equation simplifies. The system of equations on (g,w) becomes

$$\left\{ \begin{array}{rcl} \partial_t g &=& \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla\kappa * g - Q\nabla w) \\ \partial_t w &=& \dfrac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g + Q\nabla w], \end{array} \right. \tag{6.10}$$

and we will focus on the dissipativity properties of the associated operator. More precisely, defining

$$\mathscr{L}(g,w) := (\mathscr{L}_1(g,v), \mathscr{L}_2(g,v))$$

with

$$\begin{array}{rcl} \mathscr{L}_1(g,w) &:=& \Delta g + \mathrm{div}(\mu x g - g\nabla P) - \mathrm{div}(Q\nabla\kappa * g + Q\nabla w) \\ \mathscr{L}_2(g,w) &:=& \dfrac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g + Q\nabla w], \end{array}$$

we want to exhibit some scalar products $(\cdot,\cdot)_{\mathscr{H}}$ and associated norm $\|\cdot\|_{\mathscr{H}}$ such that

$$(\mathscr{L}(g,w),(g,w))_{\mathscr{H}} \le -\lambda\|(g,w)\|^2_{\mathscr{H}} + \dots,$$

with $\lambda > 0$ as large as possible and the remainder term "..." is essentially negative.

6.1.4 Functional spaces and dissipativity estimate

In order to obtain the dissipativity property of the operator, we will introduce several Hilbert spaces which more or less conveniently reveal the dissipativity mechanism.

For the first component, we will work in weighted Lebesgue spaces. For a weight function $m : \mathbb{R}^2 \to [1, \infty)$, the weighted Lebesgue space $L^p_m(\mathbb{R}^2)$, for $1 \le p \le \infty$, is defined by

$$L^p_m(\mathbb{R}^2) := \{f \in L^1_{loc}(\mathbb{R}^2);\ \|f\|_{L^p_m} := \|f\, m\|_{L^p} < \infty\},$$

and the norm of the higher-order Sobolev spaces $W^{\ell,p}_m(\mathbb{R}^2)$ is defined by

$$\|f\|^p_{W^{\ell,p}_m} := \sum_{|\alpha|\le\ell} \|m\, \partial^\alpha f\|^p_{L^p},$$

we the usual shorthand $H^\ell_m := W^{\ell,2}_m$. For the weight function m, we take

$$m^2 := e^{\vartheta \frac{\mu}{2}|x|^2 + 4\log\langle x\rangle}, \quad \vartheta \in (0,1), \tag{6.11}$$

because it will conveniently behaves as $\mu \to 0$.

For the second component, we will work in several work in several different spaces. The simplest one is to consider the Lebesgue space L^2, so that $\mathscr{H} = L^2_m \times L^2$ and more precisely

$$\|(g,w)\|^2_{\mathscr{H}} := \|g\|^2_{L^2_m} + \eta\|w\|^2_{L^2}, \tag{6.12}$$

but this choice leads to a moderate dissipative estimate.

Theorem 6.1. *With the choice (6.12) for the norm of $\mathscr{H}$, there holds*

$$(\mathscr{L}(g,w),(g,w))_{\mathscr{H}} \le -\mu\|(g,w)\|^2_{\mathscr{H}} - \frac{1}{2}\|\nabla g\|^2_{L^2_m} - \frac{1}{2\varepsilon}\|\nabla w\|^2_{L^2} + C\|g\|^2_{L^2(B_R)},$$

for any $\varepsilon \in (0, \varepsilon_0)$ and $\mu \in (0,1]$, where $C, R > 0$ are some constants (independent of ε and μ) and $\varepsilon_0 > 0$ is small enough.

In a radially symmetric framework, we define the cumulant function W of the radially symmetric function w by

$$W(s) = (\mathscr{C}w)(s) := \int_{\sqrt{s}}^\infty w(r) r dr, \quad s > 0, \tag{6.13}$$

where we abuse notation by writing $w(|x|) = w(x)$ for any $x \in \mathbb{R}^2$, so that

$$W(s) = \int_0^{\sqrt{s}} w(r) r dr \ \text{ if } \ w \in L^1,\ \langle w\rangle = 0. \tag{6.14}$$

In that situation, we may choose $\mathscr{H} = (L^2_{\text{rad}} \cap L^2_m) \times L^2_{\text{rad}}$ and more precisely

$$\|(g,w)\|^2_{\mathscr{H}} := \|g\|^2_{L^2_m} + \eta\,\|w\|^2_{L^2} + \eta\|s^{-\alpha}W\|^2_{L^2}, \quad \eta > 0,\ \alpha \in (0,1). \tag{6.15}$$

That choice leads to a stronger dissipative estimate at least for the *principal part* $\widetilde{\mathscr{L}}_{2,2}$ of the operator.

Remark 1. *With the choice (6.15) for the norm of $\mathscr{H}$, there holds*

$$(\mathscr{L}(g,w),(g,w))_{\mathscr{H}} \le -3(1-\alpha)\mu\|(g,w)\|^2_{\mathscr{H}} - \frac{1}{2}\|\nabla g\|^2_{L^2_m} - \frac{1}{2\varepsilon}D_\alpha(w) + C\|g\|^2_{L^2(B_R)},$$

for any $\varepsilon \in (0,\varepsilon_0)$, $\mu \in (0,1]$, $(g,w) \in \mathscr{H}$, $w \in L^1$ such that $\langle w\rangle = 0$, where $C, R > 0$ are some constants (independent of ε and μ) and $\varepsilon_0 > 0$ is small enough, and where we define

$$D(w) := \|\nabla w\|^2_{L^2} + 4\int s^{1-\alpha}(\partial_s W)^2 + 2\alpha(1-\alpha)\int W^2 s^{-1-\alpha}.$$

It is however worth emphasizing that $s^{-\alpha}W \in L^2$ does not ensure that $w \in L^1$ and thus the property (6.14) is not continuous for the topology associated to that norm. That can be an issue when performing an abstract spectral analysis of $\mathscr{L}$.

A last possibility will be briefly considered. For the second component, one can take the space $\mathscr{H}_{\varepsilon\mu}$ associated to the Lebesgue norm

$$\|w\|^2_{\mathscr{H}_{\varepsilon\mu}} := \int w^2 e^{\varepsilon\mu\frac{|x|^2}{2}},$$

in which the principal part of the operator $\mathscr{L}_2$ has a similar dissipativity property as in the space considered in the previous example. The issue is the strong dependency on the parameters $\varepsilon, \mu > 0$.

These dissipativity estimates will be discussed in Sections 6.4 and 6.5.

6.1.5 Weyl's type theorem

From the previous analysis and because we are able to exhibit some explicit eigenfunctions, the following information on principal part of the spectrum of $\mathscr{L}$ can be established.

Theorem 6.2. *In the space $\mathscr{H}$ defined by the norm (6.12), there holds*

$$\Sigma(\mathscr{L}) \cap \Delta_{-\mu} = \Sigma_d(\mathscr{L}) \cap \Delta_{-\mu}$$
$$\Sigma_P(\mathscr{L}) \supset \{-2\mu, -\mu, 0\}.$$

The first main issue is the possibility to improve the above result by choosing adequately the space $\mathscr{H}$ and then to establish

$$\Sigma(\mathscr{L}) \cap \Delta_{-\lambda} = \Sigma_d(\mathscr{L}) \cap \Delta_{-\lambda} = \{-2\mu, -\mu, 0\},$$

for some $\lambda \in (2\mu, 3\mu)$.

6.2 Estimates on the self-similar profile

In this section, we present an accurate estimate on the self-similar profile $(P^\mu_\varepsilon, Q^\mu_\varepsilon) =: (P,Q)$.

Proposition 6.1. *There exists $\varepsilon^* > 0$, such that, for all $\varepsilon \in (0,\varepsilon^*)$, $Q^\mu_\varepsilon(x)$ converges uniformly to $Q_0(x)$ in $\mathbb{R}^2$. Furthermore, there exists $\alpha^* \in (0,1)$ independent of μ and ε, such that for all $\varepsilon \in (0,\varepsilon^*)$ there exists $\alpha := \alpha(\varepsilon) \in (\alpha^*, 1)$ such that the following inequalities hold true:*

i) (Bounds over P^μ_ε) For all $x \in \mathbb{R}^2$ there holds

$$P_0(x) - \frac{\mu\alpha|x|^2}{2} < P(x) - \frac{\mu\alpha|x|^2}{2} < P_0(x) < P(x) < 0, \tag{6.16}$$

ii) (Bounds over $x \cdot \nabla P$) For all $x \in \mathbb{R}^2$ there holds

$$x \cdot \nabla P(x) - \mu\alpha|x|^2 < x \cdot \nabla P_0(x) < x \cdot \nabla P(x) < 0. \tag{6.17}$$

iii) (Bounds over Q) For all $x \in \mathbb{R}^2$, there holds

$$Q_0(x)e^{-\mu\frac{|x|^2}{2}} < Q(x) < Q_0(x)e^{-\mu(1-\alpha)\frac{|x|^2}{2}}. \tag{6.18}$$

Proof. In order to prove (6.16) and (6.17), we follow the same ideas as the proof for Proposition 4.1 in [39], but including the necessary modifications to handle the terms depending on ε that appear for this new problem.
First notice that $P_0(x)$ and $P(x)$ are radial functions solving the equations

$$\begin{aligned}
&\Delta P_0(x) = -Q_0(x) = -8e^{P_0(x)}\\
&\Delta P + \mu\varepsilon x \cdot \nabla P = -Q(x) = -8e^{P(x)-\mu\frac{|x|^2}{2}},
\end{aligned}$$

respectively, with $P_0(0) = P(0) = 0$. When translated to polar variables, these equations read as

$$\begin{aligned}
&P_0''(r) + \frac{1}{r}P_0'(r) = -8e^{P_0(r)}\\
&(P)''(r) + \left(\frac{1}{r} + \mu\varepsilon r\right)(P)'(r) = -8e^{P(r)-\mu\frac{|r|^2}{2}},
\end{aligned} \tag{6.19}$$

with $P_0(0) = P_0'(0) = 0 = P(0) = P'(0)$. After solving we get

$$P_0(r) = -8\int_0^r \frac{1}{\rho}\int_0^\rho \tau e^{P_0(\tau)} d\tau d\rho, \tag{6.20}$$

$$P(r) = -8\int_0^r \frac{e^{-\mu\varepsilon\frac{\rho^2}{2}}}{\rho}\int_0^\rho \tau e^{P(\tau)-\mu(1-\varepsilon)\frac{\tau^2}{2}} d\tau d\rho. \tag{6.21}$$

Plugging an expansion in powers of r up to order 4 for $P^\mu_\varepsilon(r)$ in (6.19), the coefficients of such expansion can be computed, which gives

$$P(r) - \mu\alpha\frac{r^2}{2} = -(2 + \frac{\mu\alpha}{2})r^2 + (1 + \frac{\mu}{4}(1+\varepsilon))r^4 + o(r^4), \tag{6.22}$$

$$P_0(r) = -2r^2 + r^4 + o(r^4), \tag{6.23}$$

$$P(r) = -2r^2 + (1 + \frac{\mu}{4}(1+\varepsilon))r^4 + o(r^4), \tag{6.24}$$

for all $\alpha \in (0,1)$. This implies that there exists a certain $r_0 > 0$ (which depends on α) such that for all $r \in (0, r_0)$ the relation

$$P_0(r) - \frac{\mu\alpha|r|^2}{2} < P(r) - \frac{\mu\alpha|r|^2}{2} < P_0(r) < P(r) < 0$$

holds true.
Set $\alpha = 1-\varepsilon$ and assume now that there exists $r_1 > 0$ such that $P_0(r_1) = P(r_1)$ and

$$P(r) - \frac{\mu(1-\varepsilon)|r|^2}{2} < P_0(r) < P(r)$$

for all $r \in (0, r_1)$. Using (6.20) and (6.21) we get,

$$\begin{aligned} 0 &= P_0(r_1) - P(r_1) \\ &= -8\int_0^{r_1} \frac{1}{\rho}\int_0^{\rho} \tau \left(e^{P_0(\tau)} - e^{P(\tau)-\mu(1-\varepsilon)\frac{\tau^2}{2}-\mu\varepsilon\frac{\rho^2}{2}} \right) d\tau d\rho \\ &\leqslant -8\int_0^{r_1} \frac{1}{\rho}\int_0^{\rho} \tau \left(e^{P_0(\tau)} - e^{P(\tau)-\mu(1-\varepsilon)\frac{\tau^2}{2}} \right) d\tau d\rho \\ &< 0 \end{aligned}$$

due to the fact that $P(r) - \frac{\mu(1-\varepsilon)|r|^2}{2} < P_0(r)$ for all $r \in (0, r_1)$. This is a contradiction and therefore $P_0(r) < P(r)$ for all $r > 0$.
We consider again $\alpha \in (0,1)$. Suppose now that there exists r_α such that

$$\Psi_\varepsilon^\mu(\alpha, r_\alpha) := P(r_\alpha) - \frac{\mu\alpha|r_\alpha|^2}{2} = P_0(r_\alpha)$$

and

$$P(r) - \frac{\mu\alpha|r|^2}{2} < P_0(r) < P(r)$$

for all $r \in (0, r_\alpha)$. Using again (6.21) we have

$$\begin{aligned} \Psi_\varepsilon^\mu(\alpha, r_\alpha) &= -8\int_0^{r_\alpha} \frac{1}{\rho}\int_0^{\rho} \tau e^{P(\tau)-\mu(1-\varepsilon)\frac{\tau^2}{2}-\mu\varepsilon\frac{\rho^2}{2}} d\tau d\rho - \frac{\mu\alpha|r_\alpha|^2}{2} \\ &< -8\int_0^{r_\alpha} \frac{1}{\rho}\int_0^{\rho} \tau e^{P_0(\tau)-\mu(1-\varepsilon)\frac{\tau^2}{2}-\mu\varepsilon\frac{\rho^2}{2}} d\tau d\rho - \frac{\mu\alpha|r_\alpha|^2}{2} =: \tilde{\Psi}(\alpha, r_\alpha, \mu). \end{aligned}$$

Notice that $\tilde{\Psi}(\alpha, r_\alpha, 0) = P_0(r_\alpha)$. If we prove that there exist values of α such that $\partial_\mu \tilde{\Psi}(\alpha, r_\alpha, 0) < 0$ then, in a neighbourhood of $\mu = 0$ we have $\Psi_\varepsilon^\mu(\alpha, r_\alpha) < \tilde{\Psi}(\alpha, r_\alpha, \mu) < \tilde{\Psi}(\alpha, r_\alpha, 0) = P(r_\alpha)$ which is a contradiction. Since

$$\begin{aligned} \partial_r\partial_\mu \tilde{\Psi}(\alpha, r, 0) &= \frac{4(1-\varepsilon)}{r}\int_0^r \tau^3 e^{P_0(\tau)} d\tau + 4\varepsilon r \int_0^r \tau e^{P_0(\tau)} d\tau - \alpha r \\ &= \frac{2(1-\varepsilon)}{r}\left(\ln(1+r^2) + \frac{1}{1+r^2} - 1 \right) + 4\varepsilon r \left(1 - \frac{1}{1+r^2} \right) - \alpha r \end{aligned}$$

is smaller than 0 for all $r > 0$ if $(\varepsilon, \alpha) \in [0, \varepsilon^*] \times [\alpha^*(\varepsilon), 1]$, with $\varepsilon^* = 0,206$ and $\alpha^*(\varepsilon)$ and increasing function such that $\alpha^*(0) = 0,45$ and $\alpha^*(\varepsilon^*) = 1$, we have that

$$\partial_\mu \tilde{\Psi}(\alpha, r_\alpha, 0) = \int_0^{r_\alpha} \partial_r\partial_\mu \tilde{\Psi}(\alpha, r, 0) dr < 0,$$

which leads to the desired contradiction. This finishes the proof of (6.16). We arrive to (6.17) in a similar way, but using the expressions for $x \cdot \nabla P_0$ and $x \cdot \nabla P$.
To prove (6.18), it suffices to take the exponential on the relation

$$P_0(x) - \frac{\mu|x|^2}{2} < P(x) - \frac{\mu\alpha|x|^2}{2} < P_0(x)$$

which leads to

$$Q_0(x)e^{-\frac{\mu|x|^2}{2}} < Q(x) < Q_0(x)e^{-\frac{\mu(1-\alpha)|x|^2}{2}}$$

for all $\alpha \in (\alpha^*(\varepsilon), 1)$. □

As an immediate consequence of Proposition 6.1 we have

Lemma 6.1. *There exist constants $C > 0$ and $\vartheta \in (0,1)$ independent of μ and ε such that for any $\varepsilon \in (0, \varepsilon^*/2]$*

$$0 \leq Q(x) \leq C\, e^{-\mu\vartheta|x|^2/2}\langle x\rangle^{-4}, \tag{6.25}$$

$$\sup_{x\in\mathbb{R}^2}\big(\frac{1}{|x|} + \langle x\rangle\big)\,|\nabla P(x)| \leq C, \tag{6.26}$$

and

$$\sup_{x\in\mathbb{R}^2} |\Delta P(x)| \leq C. \tag{6.27}$$

Proof. Fix $\bar\alpha := \sup_{\varepsilon\in(0,\varepsilon^*/2)} \alpha^*(\varepsilon)$. We know from the proof of Proposition 6.1 that $\vartheta := 1-\bar\alpha$ is independent from μ and ε. Furthermore, choosing $\alpha = \bar\alpha$ in (6.18) directly gives (6.25).
Computing the explicit expression for ∇P gives

$$\nabla P = -e^{-\mu\varepsilon\frac{|x|^2}{2}}\frac{x}{|x|^2}\int_0^{|x|} Q(r)e^{\mu\varepsilon\frac{r^2}{2}}rdr,$$

hence

$$|\nabla P| \leqslant \frac{1}{|x|}\int_0^{|x|} Q(r)e^{\mu\varepsilon\frac{r^2}{2}}dr \leqslant \frac{1}{|x|}\int_0^{|x|}\langle r\rangle^{-4}e^{-\mu(\vartheta-\varepsilon)\frac{r^2}{2}}rdr.$$

Given that ϑ is independent from ε, then, at least for ε small enough, we get

$$|\nabla P| \leqslant \frac{1}{|x|}\int_0^{|x|}\langle r\rangle^{-4}rdr \leqslant |x|\langle x\rangle^{-2},$$

which directly implies (6.26).
Finally

$$\Delta P = -\mu\varepsilon x\cdot\nabla P - Q,$$

and we conclude (6.27) thanks to (6.17) and (6.18). □

We recall the definition (6.11) for the value of ϑ given in Lemma 6.1, this is

$$m^2 := e^{\frac{\vartheta\mu}{2}|x|^2}\langle x\rangle^4, \tag{6.28}$$

Corollary 6.3. *There exists $\vartheta > 0$, $\varepsilon^* > 0$ and C_i such that*

$$0 \leq Qm \leq C_1, \quad Q|\nabla m|\langle x\rangle \leq C_2, \tag{6.29}$$

for any $\varepsilon \in (0,\varepsilon^*)$ and $\mu \in (0,1)$.

Proof. For the first relation in (6.29) we use

$$Qm = \frac{Qm^2}{m} \leqslant CQm^2,$$

followed by (6.25).
Next, given that $\nabla m = x(\frac{\vartheta\mu}{2} + 4\langle x\rangle^{-2})m$, we have that

$$Q|\nabla m|\langle x\rangle \leqslant (Qm^2)(m^{-1}|x|\langle x\rangle)(\frac{\vartheta\mu}{2} + 4\langle x\rangle^{-2}) \leqslant Cm^{-1}|x|\langle x\rangle.$$

We conclude thanks to the fact that $m^{-1} \leqslant C\langle x\rangle^{-2}$, where all constant are independent from ε and μ. □

We recall now the functions Q_μ and P_μ studied in [39], which with our notations can be defined as $Q_\mu := Q_0^\mu$ and $P_\mu = P_0^\mu$. We have that (Q_μ, P_μ) is the solution of the system

$$\Delta Q_\mu - \nabla \cdot (Q_\mu \nabla P_\mu - \mu x Q_\mu) = 0, \tag{6.30}$$
$$\Delta P_\mu + Q_\mu = 0, \tag{6.31}$$

Proposition 6.2. *There exist constants $\vartheta \in (0,1)$, $C_i > 0$, $i = 1, \dots, 4$, independent of μ and ε such that for all $x \in \mathbb{R}^d$*

$$|P - P_\mu| \leqslant \mu\varepsilon C_1 |x|^2, \tag{6.32}$$
$$|\nabla P - \nabla P_\mu| \leqslant \mu\varepsilon C_2 |x|, \tag{6.33}$$
$$|Q - Q_\mu| \leqslant \mu\varepsilon C_3 |x|^2 Q(x) e^{-\frac{\vartheta\mu|x|^2}{2}}, \tag{6.34}$$
$$|\nabla Q - \nabla Q_\mu| \leqslant \mu\varepsilon C_4 |x|^3 Q(x) e^{-\frac{\vartheta\mu|x|^2}{2}}. \tag{6.35}$$

Proof. We recall that in radial variables, we have the expressions

$$P_\mu(r) = -8\int_0^r \frac{1}{\rho}\int_0^\rho e^{P_\mu - \mu\frac{|\tau|^2}{2}} \tau d\tau d\rho,$$
$$P(r) = -8\int_0^r \frac{e^{-\mu\varepsilon\frac{\rho^2}{2}}}{\rho}\int_0^\rho e^{P - \mu(1-\varepsilon)\frac{|\tau|^2}{2}} \tau d\tau d\rho$$

which imply

$$\begin{aligned} P - P_\mu &= \left(P - \int_0^r P' e^{\mu\varepsilon\frac{|\tau|^2}{2}} d\tau\right) + \left(\int_0^r P' e^{\mu\varepsilon\frac{|\tau|^2}{2}} d\tau - P_\mu\right) \\ &= -\int_0^r \frac{e^{-\mu\varepsilon\frac{\rho^2}{2}} - 1}{\rho}\int_0^\rho Q(\tau) e^{\mu\varepsilon\frac{\tau^2}{2}} \tau d\tau d\rho \\ &\quad - 8\int_0^r \frac{1}{\rho}\int_0^\rho (e^{P - \mu(1-\varepsilon)\frac{|\tau|^2}{2}} - e^{P_\mu - \mu\frac{|\tau|^2}{2}}) \tau d\tau. \end{aligned}$$

Directly from Proposition 6.1 we know that $\int_0^\rho Q(\tau)e^{\mu\varepsilon\frac{\tau^2}{2}}\tau d\tau \leqslant \int_0^\rho Q_0(\tau)\tau d\tau \leqslant 8\pi$, and the mean value theorem gives us the relation

$$e^{P - \mu(1-\varepsilon)\frac{|\tau|^2}{2}} - e^{P_\mu - \mu\frac{|\tau|^2}{2}} = (P - P_\mu + \mu\varepsilon\frac{\tau^2}{2}) e^{h(\tau)}$$

with $h(\tau)$ satisfying

$$h(\tau) \leqslant \max\{P - \mu(1-\varepsilon)\frac{|\tau|^2}{2}, P_\mu - \mu\frac{|\tau|^2}{2}\} \leqslant P_0(x) - \frac{\vartheta\mu|x|^2}{2},$$

which thanks to Proposition 6.1 gives

$$8e^{h(\tau)} \leqslant Q_0(\tau)e^{-\frac{\vartheta\mu|x|^2}{2}}.$$

Putting everything together we get

$$\begin{aligned}|P - P_\mu| \leqslant& 8\pi\int_0^r \frac{1-e^{-\mu\varepsilon\rho^2}}{\rho}d\rho + \int_0^r \frac{1}{\rho}\int_0^\rho |P - P_\mu + \mu\varepsilon\frac{\tau^2}{2}|Q_0(\tau)\tau d\tau\\ \leqslant& \mu\varepsilon K r^2 + \int_0^r\int_0^\rho |P - P_\mu|Q_0(\tau)d\tau,\end{aligned}$$

where K is a constant independent of μ and ε. Integrating by parts the integral term we get

$$|P - P_\mu| \leqslant \mu\varepsilon K r^2 + r\int_0^r |P - P_\mu|Q_0(\tau)d\tau,$$

which thanks to Gronwall's Lemma gives (6.32). A similar manipulation with the gradients of P and P_μ gives (6.33).
On the other hand, we have

$$\begin{aligned}|Q - Q_\mu| =& 8|e^{P_\varepsilon^\mu - \mu(1-\varepsilon)\frac{|\tau|^2}{2}} - e^{P_\mu - \mu\frac{|\tau|^2}{2}}|\\ =& |P - P_\mu + \mu\varepsilon\frac{\tau^2}{2})|8e^{h(\tau)}\\ \leqslant& \mu\varepsilon C_2|r|^2 Q_0(r)e^{-\frac{\vartheta\mu r^2}{2}}.\end{aligned}$$

Repeating the process for $\nabla(Q - Q_\mu)$ gives (6.35). □

Corollary 6.4. *There exists $C > 0$ such that*

$$|\Delta(P - P_\mu)| \leq \mu\varepsilon C, \tag{6.36}$$

for any $\varepsilon \in (0, \varepsilon^*)$, $\mu \in (0,1)$ and $x \in \mathbb{R}^2$.

Proof. Using the equations for P and P_μ, we have

$$\Delta(P - P_\mu) = -(Q - Q_\mu) - \mu\varepsilon x \cdot \nabla P.$$

We conclude thanks to (6.34) for the first term and thanks to (6.17) for the second one. □

6.3 Some functional inequalities

We present in this section a series of estimates for functions defined in the plane and which will be useful in the sequel. Thanks to [48, Lemma B.2], which uses Fourier arguments, we have

$$\|\nabla\kappa * f\|_{L^2} = \|\mathscr{K} * f\|_{L^2} \leq C\|f\|_{L^2_\ell}, \quad \forall\, \ell > 2,\ \forall\, f \in L^2_{\ell,0} \tag{6.37}$$

and

$$\|\nabla\kappa * f\|_{L^2_1} = \|\mathscr{K} * f\|_{L^2_1} \le C\|f\|_{L^2_\ell}, \quad \forall\, \ell > 3,\ \forall\, f \in L^2_{\ell,1}$$

where we recall that κ and $\mathscr{K} = \nabla\kappa$ are defined in (6.9). We sometime use the shorthand $L^r_{\ell+0} = L^r_{\ell+}$ for saying that an estimate is true in $L^r_{\ell+\varepsilon}$ whatever is $\varepsilon > 0$ (small enough).

The above estimates are quite sharp as we can figure out from the following series of estimates.

Lemma 6.2. *For any $\varepsilon > 0$ small enough, we have*

$$\|\nabla\kappa * g\|_{L^2_{-\varepsilon}} \lesssim \|g\|_{L^2_{1+\varepsilon}}, \quad \forall\, g \in L^2_{1+\varepsilon}. \tag{6.38}$$

In the radially symmetric case, we may slightly improve (6.37) by

$$\|\nabla\kappa * g\|_{L^2} \lesssim \|g\|_{L^2_{1+0}}, \quad \forall\, g \in L^2_{\rm rad} \cap L^2_{1+0,0}. \tag{6.39}$$

In general, the estimate

$$\|\nabla\kappa * g\|_{L^2} \lesssim \|g\|_{L^2_\ell}, \quad \forall\, g \in L^2_\ell, \tag{6.40}$$

is not true whatever is $\ell > 0$ and even in the radially symmetric case.

Proof of Lemma 6.2. Step 1. We prove (6.38). We split $|\mathscr{K}| := K_1 + K_2$, with

$$K_1 := \frac{1}{|x|}\mathbf{1}_{|x|\le 1} \in L^r,\ \forall\, r < 2, \quad K_1 := \frac{1}{|x|}\mathbf{1}_{|x|\ge 1} \in L^r,\ \forall\, r > 2.$$

We have

$$\begin{aligned} \|\mathscr{K} * f\|_{L^2_{-\varepsilon}} &\le \|K_1 * f\|_{L^2} + \|K_2 * f\|_{L^r} \\ &\le \|K_1\|_{L^1}\|f\|_{L^2} + \|K_2\|_{L^r}\|f\|_{L^1} \\ &\lesssim \|f\|_{L^2_{1+0}} \end{aligned}$$

where for $\varepsilon > 0$ fixed, we have chosen $r > 2$, $r - 2 > 0$ small enough (take $2 < r < 2/(1-\varepsilon)$), and we have used the Holder inequality in the first line, where we have used twice the convolution embedding $L^1 * L^q \subset L^q$ in the second line and where finally we have used the Cauchy-Schwarz inequality in the last line in order to prove $L^2_{1+0} \subset L^1$.

Step 2. We prove (6.39). In the radially symmetric case, we define $u := \kappa * g$ in the sense

$$u'' + \frac{1}{r}u' = \Delta u = g, \quad u'(0) = 0, \tag{6.41}$$

so that $u' = \nabla\kappa * g$. We deduce $(ru')' = rg$ and thus

$$u' = \frac{1}{r}\int_0^r g(s)s\,ds.$$

Assuming that $g \in L^1(\mathbb{R}^2)$ and $\langle g\rangle = 0$, we deduce that u' satisfies (6.41) associated to the limit conditions $u'(0) = u'(\infty) = 0$. We finally observe that

$$\begin{aligned} \int (u')^2 r\,dr &= \int_0^1 \Big(\frac{1}{r}\int_0^r g(s)s\,ds\Big)^2 r\,dr + \int_1^\infty \Big(\frac{1}{r}\int_r^\infty g(s)s\,ds\Big)^2 r\,dr \\ &\le \int_0^1 \frac{1}{r}\int_0^r g(s)^2 s\,ds \int_0^r s\,ds\,dr + \int_1^\infty \frac{1}{r}\int_r^\infty g(s)^2 s^{3+\varepsilon} ds \int_r^\infty s^{-1-\varepsilon} ds\,dr \\ &\le \int_0^1 g(s)^2 s \int_s^1 \frac{r}{2}\,dr\,ds + \frac{1}{\varepsilon}\int_1^\infty \frac{dr}{r^{1+\varepsilon}}\int_0^\infty g(s)^2 s^{3+\varepsilon} ds \\ &\le \Big(\frac{1}{2} + \frac{1}{\varepsilon^2}\Big)\int_0^\infty g(s)^2(1+s^{2+\varepsilon})s\,ds, \end{aligned}$$

where we have used the condition $\langle g\rangle = 0$ at the first line, the Cauchy-Schwarz inequality at the second line and next elementary arguments. In other words, we have established (6.39).

Step 3. We prove that (6.40) *is not true in general.* We may indeed fix $g := e^{-r^2/2} \in \bigcap_\ell L^2_\ell$ so that

$$u' = \frac{1}{r}\int_0^r e^{-s^2/2} s ds \underset{r\to\infty}{\sim} \frac{1}{r} \in L^2_{-\varepsilon}(\mathbb{R}^2)\backslash L^2(\mathbb{R}^2).$$

The condition $\langle g\rangle = 0$ is necessary in (6.39) and we cannot improve (6.38) in general. $\square$

A variant of Lemma 6.2 in L^p spaces.

Lemma 6.3. *For any $p > 2$ and $q \ge 2$ (with $q > 2$ if $p = \infty$), there exists $k \in (1, 2+0)$ and $C = C(p, q, k)$ such that*

$$\|\nabla\kappa * g\|_{L^p} \le C\|g\|_{L^q_k}, \quad \forall\, g \in L^q_k. \tag{6.42}$$

Proof of Lemma 6.3. We only consider the case $p < \infty$, the case $p = \infty$ may be handled in a very similar way. With the same notations as in Lemma 6.2, we have

$$\begin{aligned} \|\mathscr{K} * f\|_{L^p} &\le \|K_1 * f\|_{L^p} + \|K_2 * f\|_{L^p} \\ &\lesssim \|K_1\|_{L^r}\|f\|_{L^2} + \|K_2\|_{L^p}\|f\|_{L^1} \\ &\lesssim \|f\|_{L^q_k}, \end{aligned}$$

where we have used the convolution embedding $L^1 * L^q \subset L^q$ and the Young convolution embedding $L^r * L^2 \subset L^p$ with $r := (1/2 + 1/p)^{-1} \in (1, 2)$ in the second line and where finally we have used the Holder inequality in the last line in order to prove $L^q_k \subset L^1 \cap L^2$.

We establish now several elementary estimates on the norm on the cumulant as defined in (6.13).

Lemma 6.4. *We have*

$$\|W\|_{L^2_\ell} \lesssim \|w\|_{L^2_{2+2\ell}}, \tag{6.43}$$

for any $w \in L^2_{\rm rad} \cap L^2_{2\ell+2,0}$, $\ell \ge 0$. We also have

$$\|W\|_{L^2_{-\ell}} \lesssim \|w\|_{L^2_{2-2\ell}}, \tag{6.44}$$

for any $w \in L^2_{\rm rad} \cap L^2_{2-2\ell}$, $\ell > 1/2$. By interpolation, we deduce

$$\|W\|_{L^2_{-\ell}} \lesssim \|w\|_{L^2_{2-2\ell}}, \tag{6.45}$$

for any $w \in L^2_{\rm rad} \cap L^2_{2-2\ell,0}$, $\theta \in (0, 1/2)$.

Proof of Lemma 6.4. Step 1. Because $\langle w\rangle = 0$, we may write

$$\begin{aligned} W(s)^2 &= \Big(\int_0^{\sqrt{s}} wrdr\Big)^2 = \Big(\int_{\sqrt{s}}^\infty wrdr\Big)^2 \\ &\le \int_{\sqrt{s}}^\infty r^{2\alpha+1}w^2 dr \int_{\sqrt{s}}^\infty r^{1-2\alpha} dr \\ &\lesssim \int_{\sqrt{s}}^\infty r^{2\alpha+1} w^2 dr\, s^{1-\alpha} \end{aligned}$$

for any $s > 0$, thanks to the Cauchy-Schwarz inequality by choosing $\alpha > 1$. We deduce

$$\begin{aligned}
\|W\|^2_{L^2_\ell} &:= \int_0^\infty W^2 \langle s\rangle^{2\ell}\, ds \\
&\lesssim \int_0^\infty \int_{\sqrt{s}}^\infty r^{2\alpha+1} w^2 dr\, \langle s\rangle^{2\ell} s^{1-\alpha} ds \\
&\lesssim \int_0^\infty r^{2\alpha+1} w^2 \int_0^{r^2} s^{1-\alpha} \langle s\rangle^{2\ell} ds dr \\
&\lesssim \int_0^\infty r^5 \langle r\rangle^{4\ell} w^2 \\
&\lesssim \|w\|^2_{L^2_{2\ell+2}}
\end{aligned}$$

by choosing $\alpha \in (1,2)$.

Step 2. We now write

$$\begin{aligned}
W(s)^2 &= \Big(\int_0^{\sqrt{s}} w r dr\Big)^2 \\
&\le \int_0^{\sqrt{s}} w^2 r^{1+2\alpha} dr \int_0^{\sqrt{s}} r^{1-2\alpha} dr \\
&\lesssim s^{1-\alpha} \int_0^{\sqrt{s}} w^2 r^{1+2\alpha} dr,
\end{aligned}$$

thanks to the Cauchy-Schwarz inequality when $\alpha < 1$. We deduce

$$\begin{aligned}
\|W\|^2_{L^2_\ell} &:= \int_0^\infty W^2 \langle s\rangle^{-2\ell}\, ds \\
&\lesssim \int_0^\infty s^{1-\alpha} \langle s\rangle^{-2\ell} \int_0^{\sqrt{s}} w^2 r^{1+2\alpha} dr\, ds \\
&\lesssim \int_0^\infty r^{1+2\alpha} w^2 \int_{r^2}^\infty s^{1-\alpha} \langle s\rangle^{-2\ell} ds dr \\
&\lesssim \int_0^\infty r \langle r\rangle^{4-4\ell} w^2\, dr \\
&\lesssim \|w\|^2_{L^2_{2-2\ell}}
\end{aligned}$$

by choosing $\alpha \ge 0$ such that $\alpha > 2-2\ell$ so that $1-\alpha-2\ell < -1$ (remind that $\ell > 1/2$ so that $2-2\ell < 1$).

Step 2. The estimate (6.45) with $\ell := 1-\theta$, $\theta \in (1/2,1)$, follows the interpolation between the estimate (6.43) with $\ell = 0$ and the estimate (6.44) with $\ell = -1$. □

The cumulant function may be useful for controlling the original one as we establish now.

Lemma 6.5. *We have*

$$\|w\|_{L^2_1} \lesssim \|\nabla w\|_{L^2} + \|s^{1/2} W'\|_{L^2}, \tag{6.46}$$

for any $w \in L^2_{\rm rad}$ such that $\nabla w \in L^2$ and $s^{1/2} W' \in L^2$.

Proof of Lemma 6.5. From the very definition, we have

$$w(\sqrt{s}) = 2W'(s),$$

so that (6.46) is equivalent to

$$\int_0^\infty w^2(1+r^2)rdr \lesssim \int_0^\infty (w')^2 rdr + \int_0^\infty w^2 r^3 dr. \tag{6.47}$$

On the one hand, we write

$$w(r) = \int_1^2 w(\rho)d\rho + \int_1^2 \int_r^\rho w'(\sigma)d\sigma d\rho, \quad \text{for} \quad 0 < r < 1,$$

and thanks to the Cauchy-Schwarz inequality, we deduce

$$\int_0^1 w^2 rdr \lesssim \int_1^2 w^2 d\rho + \int_0^2 (w')^2 \sigma d\sigma.$$

The other part of the domain of integration is trivially handled. □

We are now interested in controlling the norm of a function by a norm of its gradient. A first possible result writes as follows.

Lemma 6.6. *We have*

$$\|f\|_{L^2(\mathbb{R}^2)} \lesssim \|\nabla f\|_{L^2_1(\mathbb{R}^2)}, \tag{6.48}$$

for any radially symmetric measurable function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $f(x) \to 0$ as $|x| \to \infty$ and $\nabla f \in L^2_1$. Similarly, we have

$$\|G\|_{L^2} \lesssim \|G'\|_{L^2_1}, \tag{6.49}$$

for any measurable function $G : (0,\infty) \to \mathbb{R}$ such that $G(s) \to 0$ as $s \to \infty$ and $G' \in L^2_1$.

Proof of Lemma 6.6. Writing

$$f(r) = -\int_r^\infty f'(s)ds, \quad \forall\, r > 0,$$

we have

$$\begin{aligned}
\int_{\mathbb{R}^2} f^2 dx &\simeq \int_0^\infty \Big(\int_r^\infty f'(s)ds \Big)^2 rdr \\
&\lesssim \int_0^\infty \int_r^\infty (f'(s))^2 (1+s)^{1+\varepsilon} ds\, r^{-\varepsilon}\, rdr \\
&\lesssim \int_0^\infty (f'(s))^2 (1+s)^{1+\varepsilon} s^{2-\varepsilon} ds \\
&\lesssim \int_0^\infty |\nabla f|^2 \langle x \rangle^2 \, dx
\end{aligned}$$

by using the Cauchy-Schwarz inequality and the Fubini theorem. □

Remark 2. *We may observe that the function*

$$\mathfrak{f}(x) := \log\log\langle r\rangle, \quad \langle r\rangle := e + r,$$

satisfies

$$\begin{aligned}
\|\nabla \mathfrak{f}\|^2_{L^2(\mathbb{R}^2)} &\simeq \int \frac{1}{(\langle r\rangle \log\langle r\rangle)^2} r dr < \infty \\
\|\mathfrak{f}\|^2_{L^2_{-k}(\mathbb{R}^2)} &\simeq \int \frac{(\log\log\langle r\rangle)^2}{\langle r\rangle^{2k-1}} dr < \infty, \qquad \text{if} \quad k > 1, \\
\|\mathfrak{f}\|^2_{L^2_{-1}(\mathbb{R}^2)} &\simeq \int \frac{(\log\log\langle r\rangle)^2}{\langle r\rangle} dr = \infty.
\end{aligned}$$

As a consequence, we cannot hope for the estimate

$$\|f\|_{L^2_{-1}} \lesssim \|\nabla f\|_{L^2}$$

because it fails on $\mathfrak{f}$, nor for the estimate

$$\|f\|_{L^2_{-k}} \lesssim \|\nabla f\|_{L^2}, \quad k > 1, \tag{6.50}$$

because it fails on the constants.

We give a second estimate of the norm of a function by a norm of its gradient where the weighted Lebesgue norm concerns the function (instead of its gradient).

Lemma 6.7. *We have*

$$\int gw \lesssim \|\nabla w\|_{L^2} \|g\|_{L^2_k} \tag{6.51}$$

for any $g \in L^2_{\mathrm{rad}} \cap L^2_{k,0}$, $k > 1$, and any radially symmetric measurable function $w : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla w \in L^2$, we note $w \in \dot H^1_{\mathrm{rad}}(\mathbb{R}^2)$. As a consequence, we have

$$\|w\|_{L^2_{-k}/\mathbb{R}} \lesssim \|\nabla w\|_{L^2}, \quad \forall w \in \dot H^1_{\mathrm{rad}}(\mathbb{R}^2), \ \forall\, k > 1, \tag{6.52}$$

which is a convenient alternative to (6.50), and

$$\|g\|_{\dot H^1} \lesssim \|g\|_{L^2_k}, \quad \forall\, g \in L^2_{\mathrm{rad}} \cap L^2_{k,0}, \ \forall\, k > 1. \tag{6.53}$$

Proof of Lemma 6.7. We start writing

$$\begin{aligned}
\int gw &= \int_0^\infty g(s) w(s) s ds \\
&= -\int_0^\infty g(s) \int_s^\infty w'(\sigma) d\sigma s ds \\
&= -\int_0^\infty w'(\sigma) \int_0^\sigma g(s) s ds d\sigma \\
&\leq \Big(\int_0^\infty w'(\sigma)^2 \sigma d\sigma\Big)^{1/2} I^{1/2},
\end{aligned}$$

with

$$I := \int_0^\infty \frac{1}{\sigma}\Big(\int_0^\sigma g(s) s ds\Big)^2 d\sigma,$$

where we have used Fubini and Cauchy-Schwarz inequality. On the one hand, we have

$$\begin{aligned} I_1 &:= \int_0^1 \frac{1}{\sigma}\Big(\int_0^\sigma g(s) s ds\Big)^2 d\sigma \\ &\le \int_0^1 \frac{1}{\sigma}\int_0^\sigma g(s)^2 s ds \frac{\sigma^2}{2} d\sigma \\ &\le \frac{1}{4}\int_0^1 g(s)^2 s ds, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality in the second line. On the other hand, we have

$$\begin{aligned} I_2 &:= \int_1^\infty \frac{1}{\sigma}\Big(\int_0^\sigma g(s) s ds\Big)^2 d\sigma \\ &\lesssim \int_1^\infty \frac{1}{\sigma}\int_1^\infty g(s)^2 \langle s\rangle^{2k} s ds \langle\sigma\rangle^{2-2k} d\sigma \\ &\lesssim \int_1^\infty g(s)^2 \langle s\rangle^{2k} s ds, \end{aligned}$$

where we have used the zero mean condition on g, the Cauchy-Schwarz inequality and twice the fact that $2k-1>1$. Observing that

$$I = I_1 + I_2 \lesssim \|g\|^2_{L^2_k},$$

we immediately conclude. □

We now two last estimates of the same kind where the radial symmetric property is removed.

Lemma 6.8. *We have*

$$\Big|\int_{\mathbb{R}^2} gw\Big| \lesssim \|g\|_{L^2_k}\|\nabla w\|_{L^2} \tag{6.54}$$

for any $g \in L^2_{k,0}(\mathbb{R}^2)$, $k>2$, and any measurable function $w:\mathbb{R}^2\to\mathbb{R}$ such that $\nabla w \in L^2$, we note $w \in \dot H^1(\mathbb{R}^2)$. As a consequence, we have

$$\|w\|_{L^2_{-k}/\mathbb{R}} \lesssim \|\nabla w\|_{L^2}, \quad \forall w \in \dot H^1(\mathbb{R}^2),\ \forall k>2, \tag{6.55}$$

$$\|g\|_{\dot H^1} \lesssim \|g\|_{L^2_k}, \quad \forall g \in L^2_{k,0},\ \forall k>2. \tag{6.56}$$

Proof of Lemma 6.8. We only have to prove (6.54) what we do by adapting a standard proof of the Poincaré inequality. Observing that

$$\int \frac{c_k}{\langle x\rangle^{2k}} dx = 1, \quad k>2, c_k>0,$$

for $g \in L^2_{k,0}$ and $w \in C^1_c(\mathbb{R}^d)$, we may write

$$\int gw = \int g(w - M[w]) \le \|g\|_{L^2_k} I_w$$

with

$$I_w := \|w - M[w]\|_{L^2_{-k}}, \quad M[w] := \int w(y)\frac{c_k}{\langle y\rangle^{2k}}dy.$$

We next write

$$\begin{aligned} w(x) - M[w] &= c_k \int (w(x) - w(y))\frac{dy}{\langle y\rangle^{2k}} \\ &= c_k \int_{\mathbb{R}^2}\int_0^1 (\nabla w)(z_t)\cdot(x-y)\frac{dy}{\langle y\rangle^{2k}}dt, \quad z_t := (1-t)x + ty. \end{aligned}$$

We then split

$$I_w = I_0 + I_1,$$

with

$$I_0^2 := c_k^2 \int_{\mathbb{R}^2}\Big(\int_{\mathbb{R}^2}\int_0^{1/2}(\nabla w)(z_t)\cdot(x-y)\frac{dy}{\langle y\rangle^{2k}}dt\Big)^2\frac{dx}{\langle x\rangle^{2k}}$$

and

$$I_1^2 := c_k^2 \int_{\mathbb{R}^2}\Big(\int_{\mathbb{R}^2}\int_{1/2}^{1}(\nabla w)(z_t)\cdot(x-y)\frac{dy}{\langle y\rangle^{2k}}dt\Big)^2\frac{dx}{\langle x\rangle^{2k}}.$$

We compute for instance

$$\begin{aligned} I_0^2 &\le \frac{c_k}{2}\int_0^{1/2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\nabla w(z_t)|^2|x-y|^2\frac{dy}{\langle y\rangle^{2k}}\frac{dx}{\langle x\rangle^{2k}}dt \\ &\le \frac{c_k}{2}\int_0^{1/2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\nabla w(z_t)|^2\frac{dx}{\langle x\rangle^{2k-2}}\frac{dy}{\langle y\rangle^{2k-2}}dt \\ &\le \frac{c_k}{2}\int_0^{1/2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\nabla w(z)|^2\frac{dz}{(1-t)^2}\frac{dy}{\langle y\rangle^{2k-2}}dt \\ &= \frac{c_k}{2c_{k-1}}\int_{\mathbb{R}^2}|\nabla w(z)|^2dz, \end{aligned}$$

where we have used twice the Cauchy-Schwarz inequality in the first line, the elementary estimate $|x-y| \le \langle x\rangle\langle y\rangle$ in the second line and the change of variable $x \mapsto z$ in the third line. With the same computation for the I_1^2 term, we obtain

$$I_w^2 \le \frac{c_k}{c_{k-1}}\int_{\mathbb{R}^2}|\nabla w(z)|^2dz,$$

and the conclusion. □

We give a variant of the last estimate.

Lemma 6.9. *For any $u \in L^2_{loc}(\mathbb{R}^2)$, $\nabla u \in H^1(\mathbb{R}^2)$, we have $u \in L^\infty_{-1}(\mathbb{R}^2)$ and more precisely*

$$\sup \frac{|u(x)|}{\langle x\rangle} \lesssim \|u\|_{L^2(B_1(0))} + \|\nabla u\|_{H^1(\mathbb{R}^2)}.$$

Proof of Lemma 6.8. Step 1. We may classically write

$$u(x) - N[u] = \int_{B_1(0)}\int_0^1 (\nabla w)(z_t)\cdot(x-y)dydt, \quad z_t := (1-t)x + ty,$$

with

$$N[u] := \int_{B_\varrho(0)} u dx, \quad |B_\varrho| = \pi\varrho^2 = 1.$$

For a given $a \in \mathbb{R}^2 \backslash B_{3\varrho}(0)$, we aim to estimate

$$I = \int_{B_\varrho(a)} (u - N[u])^2 dx = I_0 + I_1,$$

where we define

$$I_0 := \int_{B_\varrho(a)} \Big(\int_{B_\varrho(0)} \int_0^{1/2} (\nabla u)(z_t) \cdot (x-y) dy dt \Big)^2 dx$$

and

$$I_1 := \int_{B_\varrho(a)} \Big(\int_{B_\varrho(0)} \int_{1/2}^1 (\nabla u)(z_t) \cdot (x-y) dy dt \Big)^2 dx.$$

We compute first

$$\begin{aligned} I_0 &\leq \frac{1}{2} \int_0^{1/2} \int_{B_\varrho(a)} \int_{B_\varrho(0)} |\nabla u(z_t)|^2 |x-y|^2 dy dx dt \\ &\lesssim |a|^2 \int_0^{1/2} \int_{B_\varrho(0)} \int_{B_\varrho(a)} |\nabla u(z_t)|^2 dy dx dt \\ &\lesssim |a|^2 \int_0^{1/2} \int_{B_\varrho(0)} \int_{B_\varrho((1-t)a)} |\nabla u(z)|^2 \frac{dz}{(1-t)^2} dy dt \\ &\lesssim |a|^2 \int_{\mathbb{R}^2} |\nabla u(z)|^2 dz, \end{aligned}$$

where we have used twice the Cauchy-Schwarz inequality in the first line, the elementary estimate $|x-y| \lesssim |a|$ in the second line and the change of variable $x \mapsto z$ in the third line. Similarly, we compute

$$\begin{aligned} I_1^2 &\leq \frac{1}{2} \int_{1/2}^1 \int_{B_\varrho(a)} \int_{B_\varrho(0)} |\nabla u(z_t)|^2 |x-y|^2 dy dx dt \\ &\lesssim |a|^2 \int_{1/2}^1 \int_{B_\varrho(a)} \int_{B_\varrho(0)} |\nabla u(z_t)|^2 dy dx dt \\ &\lesssim |a|^2 \int_{1/2}^1 \int_{B_\varrho(a)} \int_{B_\varrho((1-t)a)} |\nabla u(z)|^2 \frac{dz}{t^2} dx dt \\ &\lesssim |a|^2 \int_{\mathbb{R}^2} |\nabla u(z)|^2 dz, \end{aligned}$$

where we have rather used the change of variable $y \mapsto z$ in the third line. Both estimates together, implies

$$\int_{B_\varrho(a)} (u - N[u])^2 dx \lesssim |a|^2 \int_{\mathbb{R}^2} |\nabla u(z)|^2 dz,$$

from what we deduce

$$\int_{B_\varrho(a)} |u|^2 dx \lesssim |a|^2 \int_{\mathbb{R}^2} |\nabla u(z)|^2 dz + \int_{B_\varrho(0)} |u|^2 dx.$$

Step 2. We know from Sobolev and Morrey inequality that

$$\|u\|_{L^\infty(B)} \lesssim \|u\|_{H^2(B)},$$

so that

$$\begin{aligned} |u(x)| &\leq \|u\|_{L^\infty(B_\varrho(x))} \\ &\lesssim \|u\|_{L^2(B_\varrho(x))} + \|Du\|_{L^2(B_\varrho(x))} + \|D^2 u\|_{L^2(B_\varrho(x))} \\ &\lesssim \|u\|_{L^2(B_\varrho(0))} + \langle x\rangle \|Du\|_{L^2} + \|D^2 u\|_{L^2}, \end{aligned}$$

and the conclusion. □

6.4 Estimates on the first equation

In this section, we set

$$\begin{aligned} &\mathscr{L}_1(g,w) := \mathscr{L}_{1,1} g + \mathscr{L}_{1,2}(g,w), \\ &\mathscr{L}_{1,1} g := \Delta g + \operatorname{div}(\mu x g - g\nabla P), \quad \mathscr{L}_{1,2}(g,w) := -\operatorname{div}(Q\nabla\kappa * g + Q\nabla w). \end{aligned}$$

Defining m by (6.28) and setting $m_0 := m/\langle x\rangle$, we aim to establish the following result. We use here the shorthand $Q = Q^\mu_\varepsilon$ and $P = P^\mu_\varepsilon$.

Proposition 6.3. *There holds*

$$(\mathscr{L}_1(g,w), g)_{L^2_m} \leq -\frac{1}{2}\|\nabla g\|^2_{L^2_m} - 3\mu\|g\|^2_{L^2_m} - \frac{C_0}{2}\|g\|^2_{L^2_{m_0}} + \|\nabla w\|^2_{L^2} + C_1\|g\|^2_{L^2(B_R)}, \tag{6.57}$$

for some constants $C_i, R > 0$ and uniformly in $\mu \in (0,\mu_*)$, $\varepsilon \in (0,1)$.

6.4.1 Step 1. The principal term of classical Fokker-Planck type.

We recall that for

$$\mathscr{L} f := \Delta f + \operatorname{div}(Ff),$$

there holds

$$\begin{aligned} \int (\mathscr{L} f) f m^2 &= -\int |\nabla f|^2 m^2 + \int |f|^2 m^2 \psi_1 \\ &= -\int |\nabla (fm)|^2 + \int |f|^2 m^2 \psi_2 \\ &= -\int |\nabla (fm^2)|^2 m^{-2} + \int f^2 m^2 \psi_3, \end{aligned}$$

with

$$\psi_1 := \frac{|\nabla m|^2}{m^2} + \frac{\Delta m}{m} + \frac{1}{2}\operatorname{div} F - F\cdot\frac{\nabla m}{m}, \tag{6.58}$$

$$\psi_2 := \frac{|\nabla m|^2}{m^2} + \frac{1}{2}\operatorname{div} F - F\cdot\frac{\nabla m}{m} \tag{6.59}$$

and

$$\psi_3 := 3\frac{|\nabla m|^2}{m^2} - \frac{\Delta m}{m} + \frac{1}{2}\mathrm{div}\, F - F\cdot\frac{\nabla m}{m}. \tag{6.60}$$

We take

$$F = \nabla W, \quad W = \mu\frac{|x|^2}{2} + V, \quad V = \log\langle x\rangle^4, \quad m = Q^{-1/2}, \quad Q = e^{-W},$$

so that first

$$\begin{aligned}
&\frac{\nabla m}{m} = \tfrac{1}{2}\nabla W, \qquad \frac{\Delta m}{m} = \tfrac{1}{2}\Delta W + \tfrac{1}{4}|\nabla W|^2,\\
&\nabla W = \mu x + \nabla V = \mu x + 4\frac{x}{\langle x\rangle^2},\\
&\Delta W = 2\mu + \Delta V = 2\mu + \frac{8}{\langle x\rangle^4}
\end{aligned}$$

and then

$$\begin{aligned}
\psi_2 &= \tfrac{1}{2}\Delta W - \tfrac{1}{4}|\nabla W|^2\\
&= \tfrac{1}{2}\big(2\mu + \frac{8}{\langle x\rangle^4}\big) - \tfrac{1}{4}\big|\mu x + 4\frac{x}{\langle x\rangle^2}\big|^2\\
&= \mu + \frac{4}{\langle x\rangle^4} - \frac{1}{4}\mu^2|x|^2 - 2\mu\frac{|x|^2}{\langle x\rangle^2} - 4\frac{|x|^2}{\langle x\rangle^4}\\
&= \frac{\mu}{\langle x\rangle^2}(1-|x|^2) + \frac{4}{\langle x\rangle^4}(1-|x|^2) - \frac{1}{4}\mu^2|x|^2
\end{aligned}$$

We observe that

$$\liminf_{x\in\mathbb{R}^2} \tfrac{1}{4}\big|\mu x + 4\frac{x}{\langle x\rangle^2}\big|^2 = \min_{y>0}\tfrac{1}{4}\big(\mu y + \frac{4}{y}\big)^2 = 4\mu!!$$

We thus probably have $\exists\, M, R > 1$ large enough such that

$$\int (\mathscr{L} f - M\chi_R f) f m^2 \le -3\mu\int f^2 m^2.$$

We accept that the term

$$T_1 := \int (\Delta g + \mathrm{div}(\mu x g - g\nabla P)) g m^2$$

is fine.

6.4.2 Step 2. The remainder term.

We consider the remainder term

$$\begin{aligned}
T_2 &:= \int \mathrm{div}(-Q\nabla\kappa_g - Q\nabla w) g m^2\\
&= \int Q(\nabla\kappa_g + \nabla w)\cdot\nabla(g m^2)\\
&= \int (\nabla\kappa_g + \nabla w)\cdot[(m\nabla g)Qm + (Q\langle x\rangle\nabla m)(m_0 g)].
\end{aligned}$$

We have then

$$|T_2| \lesssim \|\nabla\kappa_g + \nabla w\|_{L^2}(\|m\nabla g\|_{L^2} + \|m_0 g\|_{L^2})$$

by using the Cauchy-Scwharz inequality and Corollary 6.3. We deduce

$$|T_2| \lesssim (\|g\|_{L^2_{1+0}} + \|\nabla w\|_{L^2})(\|\nabla g\|_{L^2_m} + \|g\|_{L^2_{m_0}})$$

by using Lemma 6.2. Thanks to the Young inequality we have the following estimate on the remainder term

$$|T_2| \lesssim (\delta\|g\|_{L^2_{m_0}} + C_\delta\|g\|_{L^2(B(0,\delta^{-1}))} + \|\nabla w\|_{L^2})(\|\nabla g\|_{L^2_m} + \|g\|_{L^2_{m_0}}).$$

Gathering this estimate with the one obtained on T_1, we conclude to (6.57).

6.5 Estimates on the second equation

In this part we are concerned with the second equation

$$\partial_t w = \mathscr{L}_2(g,w) := \frac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g + Q\nabla w]. \tag{6.61}$$

We introduce the splitting

$$\mathscr{L}_2(g,w) := \mathscr{L}_{2,1}g + \mathscr{L}_{2,2}w,$$
$$\mathscr{L}_{2,1}g := g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g], \quad \mathscr{L}_{2,2}w := \frac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + \nabla\kappa * [Q\nabla w],$$

and we first investigate the dissipativity property of a part of $\mathscr{L}_{2,2}$ in several spaces. In a second time, we investigate the dissipativity property of $\mathscr{L}_2$ by adding the contribution of the remainder terms.

6.5.1 Step 1. The principal part in L^2.

We consider here the first part of the equation, namely

$$\partial_t w = \widetilde{\mathscr{L}}_{2,2}w = \frac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w, \tag{6.62}$$

for which we investigate the dissipativity for several Hilbert norms.

- For the L^2 norm, we observe that

$$\frac{1}{2}\frac{d}{dt}\int w^2 = -\frac{1}{\varepsilon}\int|\nabla w|^2 - \mu\int w^2,$$

or in other words

$$\langle\mathscr{L}_{2,2}w, w\rangle_{L^2} = -\frac{1}{\varepsilon}\|\nabla w\|^2_{L^2} - \mu\|w\|^2_{L^2}. \tag{6.63}$$

Also observe that a solution w to equation (6.62) satisfies the conservation

$$\frac{d}{dt}\int w = -2\mu\int w,$$

what is a sharper estimate than for the L^2 norm and what suggests that we may improve the dissipativity property by working on the cumulant.

6.5.2 The principal part for other norms.

We investigate the dissipativity property of the operator $\widetilde{\mathscr{L}}_{2,2}$ in other spaces.

- We estimate the cumulant. In a radial symmetric framework, equation (6.62) writes

$$\partial_t w = \frac{1}{\varepsilon}\frac{1}{r}(rw')' + \mu r w'.$$

We deduce

$$\partial_t w r = \frac{1}{\varepsilon}(rw')' + \mu r^2 w'.$$

We assume $w_0 \in L^1 \cap L^2$ and $\langle w_0 \rangle = 0$ so that the cumulant W of w satisfies

$$W(t,s) := \int_0^{\sqrt{s}} r w(t,r) dr = -\int_{\sqrt{s}}^{\infty} r w(t,r) dr$$

because of (6.13), (6.14) and the mass conservation $\langle w \rangle = 0$. Observing that

$$w(\sqrt{s}) = 2W'(s), \quad w'(\sqrt{s}) = 4\sqrt{s}W''(s),$$

we get

$$\begin{aligned}
\partial_t W &= \frac{1}{\varepsilon}(r\partial_r w)(\sqrt{s}) + \mu \int_0^{\sqrt{s}} r^2 \partial_r w dr \\
&= \frac{1}{\varepsilon} 4s\partial_{ss} W + \mu [r^2 w]_0^{\sqrt{s}} - 2\mu \int_0^{\sqrt{s}} r w dr.
\end{aligned}$$

That gives the following equation on the cumulant

$$\partial_t W = \frac{1}{\varepsilon} 4s\partial_{ss} W + 2\mu s \partial_s W - 2\mu W. \tag{6.64}$$

We may improve the previous L^2 estimate by considering (6.64) and writing

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\int W^2 &= -\frac{4}{\varepsilon}\int \{s(\partial_s W)^2 + W\partial_s W\} + \mu \int s\partial_s W^2 - 2\mu \int W^2 \\
&= -\frac{4}{\varepsilon}\int s(\partial_s W)^2 + \frac{2}{\varepsilon} W(0)^2 - 3\mu \int W^2 \\
&= -\frac{4}{\varepsilon}\int s(\partial_s W)^2 - 3\mu \int W^2,
\end{aligned}$$

because $W(0) = \langle w \rangle = 0$.

Defining now $|||\cdot|||$ by

$$|||w|||^2 := \|w\|_{L^2}^2 + \|W\|_{L^2}^2,$$

and gathering the two previous identities, we compute

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}|||w|||^2 &= -\frac{1}{\varepsilon}\int_{\mathbb{R}^2} |\nabla w|^2 - \mu \int_{\mathbb{R}^2} w^2 - \frac{4}{\varepsilon}\int_0^{\infty} s(\partial_s W)^2 - 3\mu \int_0^{\infty} W^2 \\
&\leq -\frac{1}{2\varepsilon}\int_{\mathbb{R}^2} |\nabla w|^2 - \frac{2}{\varepsilon}\int_0^{\infty} s(\partial_s W)^2 - 3\mu |||w|||^2,
\end{aligned}$$

for $\varepsilon > 0$ small enough, thanks to Lemma 6.5.

More generally, for $\alpha \in (0,1)$, we alternatively compute

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\int W^2 s^{-\alpha} &= -\frac{4}{\varepsilon}\int \{s^{1-\alpha}(\partial_s W)^2 + (1-\alpha)s^{-\alpha}W\partial_s W\} + \mu\int s^{1-\alpha}\partial_s W^2 - 2\mu\int W^2 s^{-\alpha} \\ &= -\frac{4}{\varepsilon}\int s^{1-\alpha}(\partial_s W)^2 - \frac{2}{\varepsilon}\alpha(1-\alpha)\int W^2 s^{-1-\alpha} - (3-\alpha)\mu\int W^2 s^{-\alpha}\end{aligned}$$

because $W(s) = o(s^{1/2})$ when $s \sim 0$. An issue is the fact that $\langle w\rangle = 0$ in order to justify the computation on the cumulant, so that we need $w \in L^1$ and the L^1 does not enjoy sharper dissipativity property than -2μ.

- Because the solution w to (6.62) also satisfies

$$\partial_t w = \frac{1}{\varepsilon}\mathrm{div}(\nabla w + \varepsilon\mu x w) - 2\mu w,$$

the classical computation leads to

$$\frac{1}{2}\frac{d}{dt}\int w^2 e^{\varepsilon\mu\frac{|x|^2}{2}} = -\frac{1}{\varepsilon}\int |\nabla(w e^{\varepsilon\mu\frac{|x|^2}{2}})|^2 e^{-\varepsilon\mu\frac{|x|^2}{2}} - 2\mu\int w^2 e^{\varepsilon\mu\frac{|x|^2}{2}}.$$

From the classical Poincaré inequality

$$\int |\nabla h|^2 e^{\frac{|x|^2}{2}} \ge \int h^2 e^{\frac{|x|^2}{2}}, \quad \text{if } \int h e^{\frac{-|x|^2}{2}} = 0,$$

applied to $h(x) := w(x/\sqrt{\mu\varepsilon})e^{\frac{-|x|^2}{2}}$, we get

$$\int |\nabla(w e^{\varepsilon\mu\frac{|x|^2}{2}})|^2 e^{-\varepsilon\mu\frac{|x|^2}{2}} \ge \varepsilon\mu\int w^2 e^{\varepsilon\mu\frac{|x|^2}{2}}, \quad \text{if } \langle w\rangle = 0.$$

We thus conclude to

$$\frac{1}{2}\frac{d}{dt}\int w^2 e^{\varepsilon\mu\frac{|x|^2}{2}} \le -3\mu\int w^2 e^{\varepsilon\mu\frac{|x|^2}{2}},$$

if $\langle w_0\rangle = 0$. Denoting by $\mathscr{H}_\varepsilon$ the associated Hilbert space, we have equivalently

$$\frac{1}{2}\frac{d}{dt}\|w\|^2_{\mathscr{H}_\varepsilon} \le -3\mu\|w\|^2_{\mathscr{H}_\varepsilon}.$$

6.5.3 Step 2. The remainder part in L^2 and proof of Theorem 6.1

We are concerned with the second part of the operator in (6.61), namely

$$\widetilde{\mathscr{L}}_{2,1}(g,w) := g + \nabla\kappa * [g\nabla P + Q\nabla\kappa_g + Q\nabla w].$$

We recall that we assume here that both g and w are radially symmetric functions.

- We start with

$$\mathscr{I}_1 := \int wg \lesssim \|\nabla w\|_{L^2}\|g\|_{L^2_{1+}} \lesssim \|\nabla w\|_{L^2}\|g\|_{L^2_{m_0}},$$

thanks to Lemma 6.7.

- Next, we have

$$\begin{aligned}\mathscr{I}_2 &:= \int w(\nabla\kappa * (g\nabla P)) \\ &\leq \|w\|_{L^2}\|\nabla\kappa * (g\nabla P)\|_{L^2} \\ &\lesssim \|w\|_{L^2}\|g\nabla P\|_{L^2_{2+0}} \\ &\lesssim \|w\|_{L^2}\|g\|_{L^2_{m_0}},\end{aligned}$$

where we have used the Cauchy-Schwarz inequality in the second line, the fact that $g\nabla P \in L^2_{2+0,0}$ and (6.37) in the third line and the estimate (6.26) on ∇P in the last line.

- Similarly, we have

$$\begin{aligned}\mathscr{I}_3 &:= \int w(\nabla\kappa * (Q\nabla\kappa * g)) \\ &\leq \|w\|_{L^2}\|\nabla\kappa * (Q\nabla\kappa * g)\|_{L^2} \\ &\lesssim \|w\|_{L^2}\|Q\nabla\kappa * g\|_{L^2_{2+0}} \\ &\lesssim \|w\|_{L^2}\|\nabla\kappa * g\|_{L^3} \\ &\lesssim \|\nabla w\|_{L^2}\|g\|_{L^2_{m_0}},\end{aligned}$$

where we have used the Cauchy-Schwarz inequality in the second line, the fact that $Q\nabla\kappa * g \in L^2_{2+0,0}$ and (6.37) in the third line, the Holder inequality and the estimate (6.18) on Q in the fourth line and Lemma 6.3 in the last line.

- For the last term, we have

$$\begin{aligned}\mathscr{I}_4 &:= \int w(\nabla\kappa * (Q\nabla w)) \\ &\leq \|w\|_{L^2}\|\nabla\kappa * (Q\nabla w)\|_{L^2} \\ &\lesssim \|w\|_{L^2}\|Q\nabla w\|_{L^2_{2+0}} \\ &\lesssim \|w\|_{L^2}\|\nabla w\|_{L^2},\end{aligned}$$

where we have used the Cauchy-Schwarz inequality in the second line, the fact that $Q\nabla w \in L^2_{2+0,0}$ and (6.37) in the third line and the estimate (6.18) on Q again in the last line.

We are now in position to end the

Proof of Theorem 6.1. Summing the four above contributions $\mathscr{I}_i$, we find

$$\int w\widetilde{\mathscr{L}}_{2,2}(w,g) \lesssim \|\nabla w\|_{L^2}(\|g\|_{L^2_{m_0}} + \|w\|_{L^2})$$

Together with (6.63) and using the Young inequality, we have

$$\langle \mathscr{L}_2(g,w), w\rangle_{L^2} \leq (1 + \frac{1}{\eta} - \frac{1}{\varepsilon})\|\nabla w\|^2_{L^2} + \frac{\eta}{4}\|g\|^2_{L^2_{m_0}} - \mu\|w\|^2_{L^2}.$$

Combining that last estimate with (6.57) and recalling the notation $\mathscr{H} := L^2_m \times L^2$, we have

$$\begin{aligned}\langle \mathscr{L}(g,w),(g,w)\rangle_{\mathscr{H}} \ &\le \ -\mu\|(g,w)\|^2_{\mathscr{H}} - \frac{1}{2}\|\nabla g\|^2_{L^2_m} + \big(2+\frac{1}{\eta}-\frac{1}{\varepsilon}\big)\|\nabla w\|^2_{L^2} \\ &\quad +\big(\frac{\eta}{4}-\frac{C_0}{2}\big)\|g\|^2_{L^2_{m_0}} + C_1\|g\|^2_{L^2(B_R)},\end{aligned}$$

and we conclude by choosing first $\eta > 0$ small and next $\varepsilon_0 > 0$ small enough. □

6.6 Weyl's type result on the principal spectrum

In the space L^2_k we have

$$\Sigma(\Lambda_\varepsilon) \cap \Delta_{-3\mu} \supset \{0, -\mu, -2\mu\}.$$

We recall that

$$\begin{cases} 0 = \Delta Q + \mathrm{div}(\mu x Q - Q\nabla P), \\ 0 = \Delta P + Q + \varepsilon\mu x\cdot\nabla P. \end{cases}$$

We set

$$\begin{aligned} f_1 &:= \partial_\mu Q - \frac{\mathrm{div}(xQ)}{2\mu}, \quad v_1 := \partial_\mu P - \frac{x\cdot\nabla P}{2\mu}, \\ f_2^i &:= \partial_{x_i} Q, \quad v_2^i := \partial_{x_i} P, \quad i\in\{1,2\}, \\ f_3 &:= \mathrm{div}(xQ), \quad v_3 := x\cdot P. \end{aligned}$$

We define $D := x\cdot\nabla$ and we observe that

$$D(\nabla\cdot a) = \nabla\cdot(Da) - \nabla\cdot a, \quad D(\nabla b) = \nabla(Db) - \nabla b, \quad D(xb) = xDb + xb,$$

for a vector field a and a function b, so that

$$D\Delta b = \nabla[D\nabla b - \Delta b] = \Delta D b - 2\Delta b.$$

We compute

$$\begin{aligned} 0 &= D\nabla\cdot(\nabla Q + \mu x Q - Q\nabla P) \\ &= \nabla\cdot D(\nabla Q + \mu x Q - Q\nabla P) \\ &= \Delta(DQ - Q) + \mathrm{div}(\mu x DQ + \mu x Q) - \mathrm{div}(DQ\nabla P + Q[\nabla(DP) - \nabla P]), \end{aligned}$$

where we have used the first identity and once the stationary state equation in the first line and we have used the second identity and the third identity in the second line. Adding 3 times the first stationary state equation, we get

$$0 = \Delta\mathrm{div}(xQ) + \mathrm{div}(\mu x\mathrm{div}(xQ)) - \mathrm{div}(\mathrm{div}(xQ)\nabla P + Q\nabla DP) + 2\mu\mathrm{div}(xQ),$$

or equivalently

$$\mathscr{L}_1(\mathrm{div}(xQ), DP) = -2\mu\mathrm{div}(xQ).$$

On the other hand, we compute

$$\begin{aligned} 0 &= D(\Delta P + Q + \varepsilon\mu DP) \\ &= \Delta DP - 2\Delta P + DQ + \varepsilon\mu x \cdot \nabla DP \end{aligned}$$

and adding twice the second equation, we find

$$0 = \Delta DP + \varepsilon\mu x \cdot \nabla DP + \mathrm{div}(xQ) + 2\varepsilon\mu x \cdot \nabla P,$$

or equivalently

$$\widetilde{\mathscr{L}}_2(\mathrm{div}(xQ), DP) = -2\mu DP.$$

We define

$$G_3 := f_3, \quad W_3 := v_3 - \kappa * (f_3) = x \cdot \nabla P - \nabla\kappa * (xQ),$$

and we observe that

$$\mathscr{L}(G_3, W_3) = -2\mu(G_3, W_3).$$

Deriving the stationary state equations with respect to μ yields

$$\begin{aligned} \mathscr{L}_1(\partial_\mu Q, \partial_\mu P) &= -\mathrm{div}(xQ) = \frac{\mathscr{L}_1(f_3, v_3)}{2\mu}, \\ \widetilde{\mathscr{L}}_2(\partial_\mu Q, \partial_\mu P) &= -x \cdot \nabla P = \frac{\widetilde{\mathscr{L}}_2(f_3, v_3)}{2\mu} \end{aligned}$$

which thanks to the linearity of $\mathscr{L}_1$ and $\widetilde{\mathscr{L}}_2$, gives

$$\mathscr{L}_1(f_1, v_1) = 0 \text{ and } \widetilde{\mathscr{L}}_2(f_1, v_1) = 0.$$

Hence, defining

$$G_1 := f_1, \quad W_1 := v_1 - \kappa * (f_1),$$

gives

$$\mathscr{L}(G_3, W_3) = 0.$$

Finally, deriving with respect to x_i the steady state equations directly gives,

$$\mathscr{L}_1(f_2^i, v_2^i) = -\mu f_2^i \text{ and } \widetilde{\mathscr{L}}_2(f_2^i, v_2^i) = -\mu v_2^i.$$

which implies that

$$G_2^i := f_2^i, \quad W_2^i := v_2^i - \kappa * (f_2^i),$$

satisfies

$$\mathscr{L}(G_3^i, W_3^i) = -2\mu(G_3^i, W_3^i).$$

Chapter 7

On the self-similar stability of the parabolic-parabolic Keller-Segel equation

7.1 Introduction

In this Chapter, we are concerned with the parabolic-parabolic Keller-Segel equation in self-similar variables:

$$\begin{cases} \partial_t f = \Delta f + \operatorname{div}(\mu x f - f \nabla u) \\ \partial_t u = \dfrac{1}{\varepsilon}(\Delta u + f) + \mu x \cdot \nabla u, \end{cases} \tag{7.1}$$

with fixed drift parameter $\mu > 0$ and $\varepsilon > 0$. We establish a polynomial weighted $H^1 \times H^2$ exponential stability of the self-similar profile with sub-critical mass in the quasi parabolic-elliptic regime, that is for small values of the time scale $\varepsilon > 0$, without assuming any radial symmetry property on the initial datum. This extends a previous analysis performed in Carrapatoso-Mischler [48] in a radially symmetric framework. As in this last reference, the proof of the stability is based on a perturbation argument which takes advantage of the exponential stability of the self-similar profile for the parabolic-elliptic Keller-Segel equation established by Campos-Dolbeault [36] and Egaña-Mischler [37]. The proof however differs from [48] because it uses among other things (1) a different, and somehow more standard, perturbation argument performed at the level of the main part of the first component of the linearized operator instead of at the level of the whole linearized system and (2) a purely semigroup analysis of the linear and nonlinear stability of the system.

Our result implies that in a quasi-parabolic-elliptic regime and for some class of initial data with sub-critical mass but without assuming any radial symmetry property, the associated solution to the parabolic-parabolic KS system in standard variables (corresponding to $\mu = 0$) has a self-similar longtime behaviour, which in particular means that no concentration occurs in large time and thus the diffusion mechanism is really the dominant phenomenon all along the time evolution. It is worth mentioning that, although we work in a strongly perturbative regime, where existence of solutions is guaranteed only for a very restricted set of initial data, which allows us to determine their behaviour for all times, an alternative approach has been developed where weak solutions are guaranteed to exist for very general initial values (see for example Blanchet et al. (2006) [35] and Calvez-Corrias (2008) [46]).

We denote by $Q = Q^\mu_\varepsilon$ and $P = P^\mu_\varepsilon$ the associated self-similar profiles defined as the stationary solutions to the Keller-Segel system (7.1), that is

$$\begin{cases} 0 = \Delta Q + \operatorname{div}(\mu x Q - Q\nabla P) \\ 0 = \Delta P + Q + \varepsilon \mu x \cdot \nabla P, \end{cases} \tag{7.2}$$

which existence, uniqueness, radially symmetric property and smoothness have been established in [47, 187, 188]. We introduce the perturbation (g, v) defined by

$$f = Q + g, \quad u = P + v,$$

where $Q = Q^\mu_\varepsilon$ and $P = P^\mu_\varepsilon$ verify (7.2). If (f, u) is a solution to (7.1) then (g, v) satisfies the system

$$\begin{cases} \partial_t g = \Delta g + \operatorname{div}(\mu x g - g\nabla P - Q\nabla v) - \operatorname{div}(g\nabla v) \\ \partial_t v = \dfrac{1}{\varepsilon}(\Delta v + g) + \mu x \cdot \nabla v, \end{cases} \tag{7.3}$$

and reciprocally.
For the spaces $X := L^2_m \times (L^p \cap \dot H^1)$ and $Y = H^1_m \times (L^p \cap \dot H^2)$ we define the norms

$$\begin{aligned} |||(g, v)|||_X &:= \|g\|_{L^2_k} + \|v - \kappa * g\|_{L^p} + \|v - \kappa * g\|_{\dot H^1}, \\ |||(g, v)|||_Y &:= \|g\|_{H^1_k} + \|v - \kappa * g\|_{L^p} + \|v - \kappa * g\|_{\dot H^2}, \end{aligned} \quad k > 3,$$

where $\kappa(x) := -\frac{1}{2\pi} \ln |x|$ is the Laplace kernel, the weighted Lebesgue space $L^p_k(\mathbb{R}^2)$, for $1 \le p \le \infty$, $k \ge 0$, with weight

$$\langle x \rangle := (1 + |x|^2)^{1/2}$$

is defined by

$$L^p_k(\mathbb{R}^2) := \{f \in L^1_{loc}(\mathbb{R}^2);\ \|f\|_{L^p_k} := \|f \langle x\rangle^k\|_{L^p} < \infty\},$$

and the norm of the higher-order Sobolev spaces $W^{\ell,p}_k(\mathbb{R}^2)$ is defined by

$$\|f\|^p_{W^{\ell,p}_k} := \sum_{|\alpha| \le \ell} \|\langle x\rangle^k \partial^\alpha f\|^p_{L^p}.$$

We also denote by H^{-1}_m the duality space of H^1_m for the scalar product $\langle \cdot, \cdot\rangle_{L^2_m}$, namely

$$\|\phi\|_{H^{-1}_m} = \sup_{\|f\|_{H^1_m} \le 1} \langle \phi, f\rangle_{L^2_m} = \sup_{\|g\|_{H^1} \le 1} \langle m\phi, g\rangle_{L^2} = \|m\phi\|_{H^{-1}},$$

so that we may identify

$$H^{-1}_m = \Big\{ F_0 + \operatorname{div} F_1;\ F_i \in L^2_m \Big\},$$

and for $k > 1$ so that $L^2_k \subset L^1$, we denote

$$L^2_{k,0} := \Big\{ f \in L^2_k;\ \int_{\mathbb{R}^2} f \, dx = 0 \Big\}.$$

Theorem 7.1. *There are $\varepsilon_0, \eta_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$ and any initial data $(g_0, v_0) \in L^1_{m,0} \times (L^p \cap \dot H^1)$ with $|||(g_0, v_0)|||_X \le \eta_0$, there exists a unique global solution $(g, v) \in L^\infty_t(X) \cap L^2_t(Y)$ to (7.3), which verifies*

$$\|(g, v)\|_{L^\infty_t(X)} + \|(g, v)\|_{L^2_t(Y)} \lesssim |||(g_0, v_0)|||_X. \tag{7.4}$$

Moreover we have the decay estimate, for any $\lambda \in (0, \mu)$,

$$\|(g(t), v(t))\|_X \lesssim e^{-\lambda t} \|(g_0, v_0)\|_X. \tag{7.5}$$

This result improves [48, Theorem 1.4] where similar estimates are established with the restriction that the initial datum is radially symmetric.
In what follows, we denote $w := v - \kappa * f$ so that (g, w) satisfies the modified system

$$\begin{cases} \partial_t g = \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla\kappa * g - Q\nabla w) - \mathrm{div}(g\nabla\kappa * g) - \mathrm{div}(g\nabla w) \\ \partial_t w = \dfrac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g + Q\nabla w] + \nabla\kappa * [g\nabla w + g\nabla\kappa * g]. \end{cases} \tag{7.6}$$

Equivalently, defining the operator

$$\mathscr{L}(g, w) = (\mathscr{L}_1(g, w), \mathscr{L}_2(g, w))$$

by

$$\mathscr{L}_i(g, w) = \mathscr{L}_{i,1} g + \mathscr{L}_{i,2} w, \quad i = 1, 2,$$

with

$$\begin{cases} \mathscr{L}_{1,1} g = \Delta g + \mathrm{div}(\mu x g - g\nabla P - Q\nabla\kappa * g), \quad \mathscr{L}_{1,2} w = -\mathrm{div}(Q\nabla w), \\ \mathscr{L}_{2,1} g = g + \nabla\kappa * [g\nabla P + Q\nabla\kappa * g], \quad \mathscr{L}_{2,2} w = \dfrac{1}{\varepsilon}\Delta w + \mu x\cdot\nabla w + \nabla\kappa * [Q\nabla w], \end{cases} \tag{7.7}$$

the system (7.6) rewrites as

$$\begin{cases} \partial_t (g, w) = \mathscr{L}(g, w) + (-\mathrm{div}(g\nabla\kappa * g) - \mathrm{div}(g\nabla w), \nabla\kappa * [g\nabla w + g\nabla\kappa * g]) \\ (g, w)_{|t=0} = (g_0, w_0), \end{cases} \tag{7.8}$$

with $w_0 = v_0 - \kappa * g_0$.
In the initial sections, we present some estimates over the steady state (Q, P), in addition to some functional inequalities that will be of use throughout the chapter. Then, we establish the dissipativity of the operator $\mathscr{L}_{1,1}$ for small enough values of $\varepsilon > 0$ thanks to a perturbation argument and by taking advantage of the dissipativity of the limit operator for $\varepsilon = 0$ corresponding to the usual linearized parabolic-elliptic Keller-Segel operator. Afterwards, we prove in a more direct way the dissipativity of the operator $\mathscr{L}_{2,2}$. We deduce then the decay of the semigroup $S_{\mathscr{L}}$ associated to $\mathscr{L}$ by writing in a proper accurate enough semigroup way the two decay estimates of $S_{\mathscr{L}_{ii}}$ and by showing that both out of the diagonal contribution $\mathscr{L}_{i,j}$, $i \neq j$, are small enough. The above two arguments significantly differ from those used in the proof of [48, Theorem 1.4]. We conclude the proof of Theorem 7.1 by a classical nonlinear stability trick.

In the sequel, for two functions S and T defined on $\mathbb{R}_+$, we define the convolution $S * T$ by

$$(S * T)(t) = \int_0^t S(t-s)T(s)\,\mathrm{d}s, \quad \text{for all } t \geq 0,$$

so that in particular the Duhamel formula associated to an evolution equation

$$\partial_t g = \Lambda g + G, \quad g(0) = g_0,$$

writes

$$g = S_\Lambda g_0 + S_\Lambda * G.$$

Moreover, for $\lambda \in \mathbb{R}$, we denote $\mathbf{e}_\lambda : t \mapsto e^{\lambda t}$.

7.2 Estimates over Q and P

We denote by $Q = Q_\varepsilon^\mu$ and $P = P_\varepsilon^\mu$ the associated self-similar profiles defined as the stationary solutions to the Keller-Segel system (7.1), that is

$$\begin{cases} 0 = \Delta Q + \mathrm{div}(\mu x Q - Q\nabla P) \\ 0 = \Delta P + Q + \varepsilon\mu x \cdot \nabla P. \end{cases} \tag{7.9}$$

It is worth noticing here that whatever is the value of $\varepsilon \in [0, \varepsilon_0)$, we may establish that

$$Q_\varepsilon^\mu \to Q^0, \quad P_\varepsilon^\mu \to P^0, \quad \text{as} \quad \mu \to 0,$$

where (Q^0, P^0) is defined by

$$Q^0(x) := \frac{8}{(1+|x|^2)^2}, \quad \Delta P^0 = Q^0$$

Proposition 7.1. *There exists $\varepsilon_0 > 0$, such that, for all $\varepsilon \in (0, \varepsilon_0)$, $Q(x)$ converges uniformly to $Q^0(x)$ in $\mathbb{R}^2$. Furthermore, there exists $\alpha^* \in (0,1)$ independent of μ and ε, such that for all $\varepsilon \in (0, \varepsilon^*)$ there exists $\alpha := \alpha(\varepsilon) \in (\alpha^*, 1)$ such that the following inequalities hold true:*

i) (Bounds over P) For all $x \in \mathbb{R}^2$ there holds

$$P^0(x) - \frac{\mu\alpha|x|^2}{2} < P(x) - \frac{\mu\alpha|x|^2}{2} < P^0(x) < P(x) < 0, \tag{7.10}$$

ii) (Bounds over $x \cdot \nabla P$) For all $x \in \mathbb{R}^2$ there holds

$$x \cdot \nabla P(x) - \mu\alpha|x|^2 < x \cdot \nabla P^0(x) < x \cdot \nabla P(x) < 0. \tag{7.11}$$

iii) (Bounds over Q) For all $x \in \mathbb{R}^2$, there holds

$$Q^0(x) e^{-\mu\frac{|x|^2}{2}} < Q(x) < Q^0(x) e^{-\mu(1-\alpha)\frac{|x|^2}{2}}. \tag{7.12}$$

Some immediate consequence of Proposition 7.1 are established on the following Lemmas.

Lemma 7.1. *There exist constants $C > 0$ and $\vartheta \in (0,1)$ independent of μ and ε such that for any $\varepsilon \in (0, \varepsilon_0]$*

$$0 \le Q(x) \le C\, e^{-\mu\vartheta|x|^2/2} \langle x \rangle^{-4}, \tag{7.13}$$

$$\sup_{x \in \mathbb{R}^2} \Big(\frac{1}{|x|} + \langle x \rangle\Big) |\nabla P(x)| \le C, \tag{7.14}$$

and

$$\sup_{x \in \mathbb{R}^2} |\Delta P(x)| \le C. \tag{7.15}$$

Lemma 7.2. *For all $k > 3$ there exists C depending on μ and k such that*

$$\|Q(x)\langle x \rangle^k\|_{L^\infty(\mathbb{R}^2)} \leqslant C(\mu, k). \tag{7.16}$$

Furthermore, $C(\mu, k) \leqslant C\mu^{-\max\{0, (\frac{k}{2}-2)\}}$ for some C independent of μ and ε.

Lemma 7.3. *There exist constants $C_1 > 0$ and $C_2 > 0$ independent of μ and ε such that for any $\varepsilon \in (0, \varepsilon^*/2]$*

$$\|\nabla^2 P\|_{L^\infty(\mathbb{R}^2)} \leqslant C_1 \tag{7.17}$$

and

$$\|\nabla Q m^2\|_{L^\infty(\mathbb{R}^2)} \le C_2. \tag{7.18}$$

Finally, we include a convergence result between (P, Q) and (Q_0, P_0) (a solution to (7.2) with $\varepsilon = 0$).

Lemma 7.4. *There exist constants $\vartheta \in (0,1)$, $C_i > 0$, $i = 1, \dots, 4$, independent of μ and ε such that for all $x \in \mathbb{R}^d$*

$$|P - P_\mu| \leqslant \mu\varepsilon C_1 |x|^2, \tag{7.19}$$

$$|\nabla P - \nabla P_\mu| \leqslant \mu\varepsilon C_2 |x|, \tag{7.20}$$

$$|Q - Q_\mu| \leqslant \mu\varepsilon C_3 |x|^2 Q^0(x) e^{-\frac{\vartheta\mu|x|^2}{2}}. \tag{7.21}$$

$$|\nabla Q - \nabla Q_\mu| \leqslant \mu\varepsilon C_4 |x|^3 Q^0(x) e^{-\frac{\vartheta\mu|x|^2}{2}}. \tag{7.22}$$

Proofs for Proposition 7.1 and Lemmas 7.1, 7.2, 7.3 and 7.4 were given in Chapter 6

7.3 Functional inequalities

We gather in this section some functional inequalities that we shall use through the paper. First, we provide some estimates over the solution for the Poisson problem.

Lemma 7.5. *For any $g \in L^2_{k,0}$ with $k > 3$ there holds*

$$\|\nabla\kappa * g\|_{L^2} \lesssim \|g\|_{L^2_k}. \tag{7.23}$$

Proof. We observe that

$$\|\nabla\kappa * g\|_{L^2} \approx \|\xi(\widehat{\kappa * g})\|_{L^2}$$

and $\widehat{\kappa * g}(\xi) = |\xi|^{-2}\widehat{g}(\xi)$. Therefore

$$\|\nabla\kappa * g\|^2_{L^2} \approx \int \frac{|\widehat{g}(\xi)|^2}{|\xi|^2}\,\mathrm{d}\xi = \int \mathbf{1}_{|\xi|\le 1}\frac{|\widehat{g}(\xi)|^2}{|\xi|^2}\,\mathrm{d}\xi + \int \mathbf{1}_{|\xi|>1}\frac{|\widehat{g}(\xi)|^2}{|\xi|^2}\,\mathrm{d}\xi =: I_1 + I_2.$$

For the second term we have

$$I_2 \lesssim \int |\widehat{g}(\xi)|^2\,\mathrm{d}\xi \lesssim \|g\|_{L^2}.$$

For the first term we use that $\widehat{g}(0) = 0$, since $\langle g\rangle = 0$, to write

$$\widehat{g}(\xi) = \xi \cdot \int_0^1 D_\xi \widehat{g}(\theta\xi)\,\mathrm{d}\theta,$$

and thus we obtain

$$I_1 \lesssim \left(\sup_{|\xi|\le 1} |D_\xi \widehat{g}(\xi)|\right)\int \mathbf{1}_{|\xi|\le 1}\,\mathrm{d}\xi \lesssim \sup_{|\xi|\le 1}|D_\xi\widehat{g}(\xi)| \lesssim \|xg\|_{L^1} \lesssim \|g\|_{L^2_k}.$$

□

Lemma 7.6. *For $k > 1$ and $p > 2$, $2 \le q \le p$, we have*

$$\|\nabla\kappa * g\|_{L^p} \lesssim \|g\|_{L^q_k}, \quad \forall\, g \in L^q_k. \tag{7.24}$$

Proof. We split $|\mathscr{K}| := K_1 + K_2$, with

$$K_1 := \frac{1}{|x|}\mathbf{1}_{|x|\le 1} \in L^r,\, \forall\, r < 2, \quad K_2 := \frac{1}{|x|}\mathbf{1}_{|x|\ge 1} \in L^r,\, \forall\, r > 2.$$

We have

$$\begin{aligned}\|\mathscr{K} * f\|_{L^p} &\le \|K_1 * f\|_{L^p} + \|K_2 * f\|_{L^p}\\ &\le \|K_1\|_{L^{\frac{1}{1+\frac{1}{p}-\frac{1}{q}}}}\|f\|_{L^q} + \|K_2\|_{L^p}\|f\|_{L^1}\\ &\lesssim \|f\|_{L^q_k},\end{aligned}$$

where we have used the convolution embeddings $L^{\frac{1}{1+\frac{1}{p}-\frac{1}{q}}} * L^q \subset L^p$ and $L^1 * L^p \subset L^p$ in the second line, and the Cauchy-Schwarz inequality in the last line in order to prove $L^q_k \subset L^1$ with $q \ge 2$. □

We recall the following two particular cases of the Gagliardo-Nirenberg interpolation Theorem in dimension 2:

Lemma 7.7. *1. For any $p > 2$ we have*

$$\|\nabla f\|_{L^p} \lesssim \|f\|_{L^p}^{1-\theta}\|\nabla^2 f\|_{L^2}^{\theta}, \quad \theta = \frac{p}{2+p}. \tag{7.25}$$

2. We have the Ladyzhenskaya inequality

$$\|f\|_{L^4} \lesssim \|f\|_{L^2}^{1/2}\|\nabla f\|_{L^2}^{1/2}. \tag{7.26}$$

We also have:

Lemma 7.8. *Let $p \in (2,\infty)$. For any $\beta > 0$ there is $C_\beta > 0$ such that*

$$\|\nabla f\|_{L^2}^2 \le \beta\|f\|_{L^p}^2 + C_\beta\|\nabla^2 f\|_{L^2}^2. \tag{7.27}$$

Proof. We write, using Hölder's inequality,

$$\begin{aligned}\|\nabla f\|_{L^2}^2 &\lesssim \int_{|\xi|\le 1}|\xi|^2|\widehat{f}|^2 + \int_{|\xi|>1}|\xi|^2|\widehat{f}|^2\\ &\lesssim \int_{|\xi|\le 1}(\langle\xi\rangle^{-1}|\widehat{f}|)(|\xi|^2|\widehat{f}|) + \int_{|\xi|>1}(\langle\xi\rangle^{-2/3}|\widehat{f}|^{2/3})(|\xi|^{8/3}|\widehat{f}|^{4/3})\\ &\lesssim \|\langle\xi\rangle^{-1}\widehat{f}\|_{L^2}\||\xi|^2\widehat{f}\|_{L^2} + \|\langle\xi\rangle^{-1}\widehat{f}\|_{L^2}^{2/3}\||\xi|^2\widehat{f}\|_{L^2}^{4/3}\\ &\lesssim \|f\|_{H^{-1}}\|\nabla^2 f\|_{L^2} + \|f\|_{H^{-1}}^{2/3}\|\nabla^2 f\|_{L^2}^{4/3}\\ &\lesssim \|f\|_{L^p}\|\nabla^2 f\|_{L^2} + \|f\|_{L^p}^{2/3}\|\nabla^2 f\|_{L^2}^{4/3},\end{aligned}$$

where we haved the continuous embedding $L^p(\mathbb{R}^2) \subset H^{-1}(\mathbb{R}^2)$ (consequence of the embedding $H^1(\mathbb{R}^2) \subset L^{p'}(\mathbb{R}^2)$). We then conclude by applying Young's inequality. □

7.4 Estimates for $\mathscr{L}_{1,1}$

In this section, we establish some dissipativity estimates and related semigroup decay estimates successively on the operators $\mathscr{L}_{1,1}$ and related operators.

7.4.1 Dissipativity estimates related to $\mathscr{L}_{1,1}$.

In order to keep track of the $\varepsilon \geq 0$ dependence, let us denote

$$\Lambda_\varepsilon := \mathscr{L}_{1,1}.$$

We start with a first fundamental dissipativity estimate.

Lemma 7.9. *There is $\varepsilon_0 > 0$ small enough such that for any $\varepsilon \in (0, \varepsilon_0)$ the following holds: For any $m = \langle x \rangle^k$ with $k > 3$ there are constants $C_0, \varrho_0 > 0$ such that*

$$\langle \Lambda_\varepsilon g, g \rangle_{L^2_m} \leq -\mu(k-2)\|g\|^2_{L^2_m} - \frac{1}{2}\|\nabla g\|^2_{L^2_m} + C_0\|g\|^2_{L^2(B_{\varrho_0})}. \tag{7.28}$$

Proof. We briefly repeat the proof of [48, Lemma 4.4].

We compute

$$\begin{aligned}\langle \Lambda_\varepsilon g, g \rangle_{L^2_m} &= \int \Delta g g m^2 + \int \operatorname{div}(\mu x g) g m^2 - \int \operatorname{div}(g \nabla P) g m^2 - \int \operatorname{div}(Q \nabla \kappa * g) g m^2 \\ &= I_1 + I_2 + I_3 + I_4,\end{aligned}$$

and estimate each term separately.

For the two first terms we have

$$I_1 + I_2 = -\int |\nabla g|^2 m^2 + \int \psi_1 g^2 m^2$$

where

$$\begin{aligned}\psi_1 &= \frac{|\nabla m|^2}{m^2} + \frac{\Delta m}{m} + \mu - \mu x \cdot \frac{\nabla m}{m} \\ &= k(2k+\mu)\langle x \rangle^{-2} - \mu(k-1) - 2k\langle x \rangle^{-4}.\end{aligned}$$

Moreover, for the third term we compute

$$\begin{aligned}I_3 &= \int g \nabla P \cdot \nabla g m^2 + 2\int g^2 m^2 \nabla P \cdot \frac{\nabla m}{m} \\ &= -\frac{1}{2}\int \Delta P g^2 m^2 + \int \nabla P \cdot \frac{\nabla m}{m} g^2 m^2.\end{aligned}$$

Thanks to the properties of P established in Proposition 7.1, we observe that

$$\left| \nabla P \cdot \frac{\nabla m}{m} \right| \leq \|x \cdot \nabla P\|_{L^\infty} \langle x \rangle^{-1} \leq C_1 \langle x \rangle^{-1}$$

for some constant $C_1 > 0$, and also that

$$\Delta P = -Q - \varepsilon \mu x \cdot \nabla P$$

with

$$|x\cdot\nabla P|\le C_2,\quad Q\le C_3\langle x\rangle^{-1},$$

for constants $C_2, C_3>0$, which implies

$$I_3\le \frac{\varepsilon\mu C_2}{2}\|g\|^2_{L^2_m}+\left(C_1+\frac{C_3}{2}\right)\|\langle x\rangle^{-\frac12}g\|^2_{L^2_m}.$$

For the last term we write

$$I_4=\int Q(\nabla\kappa * g)\nabla g m^2+2\int Q(\nabla\kappa * g)\frac{\nabla m}{m}\,gm^2.$$

Since $\|Qm^2\|_{L^\infty}\le C_4$ and $\|Qm\nabla m\|_{L^\infty}\le C_4$ thanks to Lemma 7.2, we obtain

$$\begin{aligned}
I_4 &\le C_4\|\nabla\kappa * g\|_{L^2}\|\nabla g\|_{L^2}+2C_4\|\nabla\kappa * g\|_{L^2}\|g\|_{L^2}\\
&\le C_4'\|\nabla\kappa * g\|^2_{L^2}+\frac12\|\nabla g\|^2_{L^2}+C_4'\|g\|^2_{L^2}\\
&\le C_4''\|\langle x\rangle^{-1}g\|^2_{L^2_m}+\frac12\|\nabla g\|^2_{L^2},
\end{aligned}$$

where we have used Lemma 7.5 and Young's inequality.
Gathering previous estimates yields

$$\langle\Lambda_\varepsilon g,g\rangle_{L^2_m}\le -\frac12\int|\nabla g|^2m^2+\int\bar\psi_1 g^2m^2$$

with

$$\begin{aligned}
\bar\psi_1 &= -\mu\left(k-1-\frac{\varepsilon C_2}{2}\right)+\left(C_4''+C_1+\frac{C_3}{2}\right)\langle x\rangle^{-1}+k(2k+\mu)\langle x\rangle^{-2}-k^2\langle x\rangle^{-4}\\
&\le -\mu\left(k-1-\frac{\varepsilon C_2}{2}\right)+C_5\langle x\rangle^{-1}.
\end{aligned}$$

We remark that, for any $\varrho_0\ge1$, we have

$$\langle x\rangle^{-1}m^2\le \varrho_0^{2k-1}\mathbf{1}_{\langle x\rangle\le\varrho_0}+\frac{1}{\varrho_0}m^2,$$

thus we obtain

$$\langle\Lambda_\varepsilon g,g\rangle_{L^2_m}\le -\frac12\|\nabla g\|^2_{L^2_m}-\mu\left(k-1-\frac{\varepsilon C_2}{2}-\frac{C_5}{\mu\varrho_0}\right)\|g\|^2_{L^2_m}+C_0\|g\|^2_{L^2(B_{\varrho_0})}\tag{7.29}$$

where $C_0=C_5\varrho_0^{2k-1}$.

We therefore choose $\varepsilon_0>0$ small enough such that $\varepsilon_0C_2\le1$ and $\varrho_0\ge1$ large enough such that $\frac{C_5}{\mu\varrho_0}\le 1/2$, which concludes the proof. □

7.4.2 Splitting of the operator $\mathscr{L}_{1,1}$

We introduce the splitting

$$\Lambda_\varepsilon=\mathscr{A}+\mathscr{B}_\varepsilon,\quad \mathscr{A}:=M\chi_\varrho,\quad \mathscr{B}_\varepsilon:=\Lambda_\varepsilon-\mathscr{A},$$

with $\chi_\varrho(x):=\chi(x/\varrho)$, $\chi\in\mathscr{D}(\mathbb{R}^2)$, $\mathbf{1}_{B(0,1)}\le\chi\le\mathbf{1}_{B(0,2)}$, and constants $M,\varrho>0$. We immediately deduce from Lemma 7.9 that $\mathscr{B}_\varepsilon$ is dissipative, more precisely:

Corollary 7.2. *For any $\varepsilon \in (0, \varepsilon_0)$, any $m = \langle x \rangle^k$ with $k > 3$, any constants $M \geq C_0$ and $\varrho \geq \varrho_0$ there holds*

$$\langle \mathscr{B}_\varepsilon g, g \rangle_{L^2_m} \leq -\mu(k-2)\|g\|^2_{L^2_m} - \frac{1}{2}\|\nabla g\|^2_{L^2_m} \leq -\lambda\|g\|^2_{L^2_m} - \sigma\|g\|^2_{H^1_m} \tag{7.30}$$

for any $0 \leq \lambda < \mu(k-2)$ with $\sigma = \min(1/2, \mu - \lambda)$, and where $\varepsilon_0, C_0, \varrho_0 > 0$ are from Lemma 7.9.

Remark 3. *We shall fix hereafter the parameters $M \geq C_0$ and $\varrho \geq \varrho_0$ in the definition of $\mathscr{B}_\varepsilon$ such that Lemma 7.9 holds.*

In order to work at the level of the semigroup, we reformulate (7.30) in the following way.

Lemma 7.10. *For any $\varepsilon \in (0, \varepsilon_0)$, any $m = \langle x \rangle^k$ with $k > 3$, any constants $M \geq C_0$ and $\varrho \geq \varrho_0$ there holds*

1. *For all $0 \leq \lambda < \mu(k-2)$ and all $g \in L^2_m$, we have*

$$\|\mathbf{e}_\lambda S_{\mathscr{B}_\varepsilon}(\cdot) g\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda S_{\mathscr{B}_\varepsilon}(\cdot) g\|_{L^2_t H^1_m} \lesssim \|g\|_{L^2_m}.$$

2. *For all $0 \leq \lambda < \mu(k-2)$ and all $\mathbf{e}_\lambda R \in L^2_t H^{-1}_m$, we have*

$$\|\mathbf{e}_\lambda (S_{\mathscr{B}_\varepsilon} * R)\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda (S_{\mathscr{B}_\varepsilon} * R)\|_{L^2_t H^1_m} \lesssim \|\mathbf{e}_\lambda R\|_{L^2_t H^{-1}_m}.$$

Proof. Let $0 \leq \lambda < \mu(k-2)$. We first consider $f := \mathbf{e}_\lambda S_{\mathscr{B}_\varepsilon}(\cdot) g$ which is a solution to the evolution equation

$$\partial_t f = \mathscr{B}_\varepsilon f + \lambda f, \quad f(0) = g.$$

Because of (7.30), we have

$$\frac{1}{2}\frac{d}{dt}\|f\|^2_{L^2_m} = \langle \mathscr{B}_\varepsilon f, f \rangle_{L^2_m} + \lambda\|f\|^2_{L^2_m} \leq -\sigma\|f\|^2_{H^1_m}$$

from which we deduce (1) thanks to the Grönwall's lemma.

We next consider $f := \mathbf{e}_\lambda (S_{\mathscr{B}_\varepsilon} * R)$ which is a solution to the evolution equation

$$\partial_t f = \mathscr{B}_\varepsilon f + \lambda f + \mathbf{e}_\lambda R, \quad f(0) = 0.$$

Because of (7.30) and the Young inequality, we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|f\|^2_{L^2_m} &= \langle \mathscr{B}_\varepsilon f, f \rangle_{L^2_m} + \lambda\|f\|^2_{L^2_m} + \langle \mathbf{e}_\lambda R, f \rangle_{L^2_m} \\ &\leq -\sigma\|f\|^2_{H^1_m} + \|\mathbf{e}_\lambda R\|_{H^{-1}_m}\|f\|_{H^1_m} \\ &\leq -\frac{\sigma}{2}\|f\|^2_{H^1_m} + C\|\mathbf{e}_\lambda R\|^2_{H^{-1}_m}, \end{aligned}$$

for some constant $C = C(\mu, \lambda) > 0$. We deduce (2) thanks to the Grönwall's lemma again. □

7.4.3 Spectral analysis of $\mathscr{L}_{1,1}$

We deduce a nice localization of the spectrum of $\mathscr{L}_{1,1}$ from the previous estimates and a perturbation argument. Let us denote by Λ_0 the linearized operator of the parabolic-elliptic Keller-Segel equation which given by

$$\Lambda_0 g = \Delta g + \operatorname{div}(\mu x g - g\nabla P_0 - Q_0 \nabla\kappa * g)$$

where (Q_0, P_0) is a solution to (7.2) with $\varepsilon = 0$. From [36, 37], we know that for any weight function $m = \langle x\rangle^k$ with $k > 3$ there holds: For all $0 < \lambda < \mu$ there exists a constant $C = C(\lambda, \mu, k) \geq 1$ such that

$$\|S_{\Lambda_0}(t) f\|_{L^2_m} \lesssim e^{-\lambda t}\|f\|_{L^2_m}, \quad \forall\, f \in L^2_{m,0},$$

and the spectrum verifies

$$\Sigma(\Lambda_0) \cap \Delta_{-\mu} = \{0\} \tag{7.31}$$

where $\Delta_{-\mu} := \{z \in \mathbb{C} : \Re e z > -\mu\}$.

By a perturbation argument similar to the one used in [189] (see also [190, 191]) we are able to obtain a similar picture for the operator $\mathscr{L}_{1,1} = \Lambda_\varepsilon$.

Proposition 7.2. *Let $m = \langle x\rangle^k$ with $k > 3$. For any $0 < \lambda < \mu$, there is $\varepsilon_* > 0$ small enough, such that*

$$\Sigma(\Lambda_\varepsilon) \cap \Delta_{-\mu} = \{0\}, \quad \forall\, \varepsilon \in (0, \varepsilon^*).$$

Proof. We split the proof into several steps.

Step 1. We claim that

$$\mathscr{U}_\varepsilon(z) := \mathscr{R}_{\mathscr{B}_\varepsilon}(z) - \mathscr{R}_{\Lambda_0}(z)\mathscr{A}\mathscr{R}_{\mathscr{B}_\varepsilon}(z)$$

is uniformly bounded in $\mathscr{B}(L^2_m)$ and $\mathscr{B}(H^{-1}_m, H^1_m)$ for any $z \in \Omega := \Delta_{-\lambda}\backslash B(0, r/2)$ any $\varepsilon \geq 0$ and $0 < r < \lambda < \mu$. On the one hand, $\mathscr{R}_{\mathscr{B}_\varepsilon}(z) \in \mathscr{B}(L^2_m)$ is just an immediate consequence of the growth estimate on $S_{\mathscr{B}_\varepsilon}$ established in Lemma 7.10-(1). For proving $\mathscr{R}_{\mathscr{B}_\varepsilon}(z) \in \mathscr{B}(H^{-1}_m, H^1_m)$, we consider first $g \in L^2_m$, $z \in \Delta_{-\mu}$ and we define $f := \mathscr{R}_{\mathscr{B}_\varepsilon}(z) g$, so that $(z - \mathscr{B}_\varepsilon) f = g$. Using (7.30) and the fact that $\mu(k-2) \geq \mu$, we deduce

$$\frac{1}{2}\|\nabla f\|^2_{L^2_m} + (\Re e z + \mu)\|f\|^2_{L^2_m} \leq \langle (z - \mathscr{B}_\varepsilon) f, f\rangle_{L^2_m} = \langle f, g\rangle_{L^2_m} \leq \|f\|_{H^1_m}\|g\|_{H^{-1}_m}$$

and thus

$$\|\nabla f\|_{L^2_m} \leq \max\big(2, \mu^{-1}\big)\|g\|_{H^{-1}_m}. \tag{7.32}$$

By a density argument, the same holds for any $g \in H^{-1}_m$. From (7.31), we also have $\mathscr{R}_{\Lambda_0}(z) \in \mathscr{B}(L^2_m)$ uniformly bounded in $\mathscr{B}(L^2_m)$ for any $z \in \Omega := \Delta_{-\lambda}\backslash B(0, r/2)$. Moreover, the proof of the bound in $\mathscr{B}(L^2_m, H^1_m)$ is exactly the same as for $\mathscr{R}_{\mathscr{B}_\varepsilon}(z)$: Indeed arguing as in Lemma 7.9 we first obtain

$$\langle \Lambda_0 f, f\rangle_{L^2_m} \leq -\mu\|f\|^2_{L^2_m} - \frac{1}{2}\|\nabla f\|^2_{L^2_m} + C_0\|f\|^2_{L^2(B_{\varrho_0})},$$

thus defining $f := \mathscr{R}_{\Lambda_0}(z) g$ we deduce

$$(\Re e z + \mu)\|f\|^2_{L^2_m} + \frac{1}{2}\|\nabla f\|^2_{L^2_m} - C_0\|f\|^2_{L^2(B_{\varrho_0})} \leq \langle (z - \Lambda_0) f, f\rangle_{L^2_m} = \langle g, f\rangle_{L^2_m} \leq C\|g\|^2_{L^2_m}.$$

Step 2. We claim that the operators converge in the sense

$$\|\Lambda_\varepsilon - \Lambda_0\|_{\mathscr{B}(H^1_k, L^2_k)} \le \eta_1(\varepsilon) \to 0.$$

We indeed write

$$\begin{aligned}(\Lambda_\varepsilon - \Lambda_0) g &= (\mathscr{B}_\varepsilon - \mathscr{B}_0) g \\ &= -\mathrm{div}(g \nabla (P_\varepsilon - P_0)) - \mathrm{div}((Q_\varepsilon - Q_0)\nabla\kappa * g) \\ &= -\nabla g \cdot \nabla (P_\varepsilon - P_0) + g\Delta(P_\varepsilon - P_0) \\ &\quad + \nabla(Q_\varepsilon - Q_0)\cdot \nabla\kappa * g - (Q_\varepsilon - Q_0) g,\end{aligned}$$

so that

$$\begin{aligned}\|(\Lambda_\varepsilon - \Lambda_0) g\|_{L^2_m} \le\ & \|\nabla(P_\varepsilon - P_0)\|_{L^\infty} \|\nabla g\|_{L^2_m} + \|\Delta(P_\varepsilon - P_0)\|_{L^\infty}\|g\|_{L^2_m} \\ & + \|m\nabla(Q_\varepsilon - Q_0)\|_{L^\infty}\|g\|_{L^2_{1+0}} + \|Q_\varepsilon - Q_0\|_{L^\infty}\|g\|_{L^2_m}.\end{aligned}$$

We immediately conclude since we are able to prove (see Lemma 7.4)

$$\nabla(P_\varepsilon - P_0) \to 0, \quad \Delta(P_\varepsilon - P_0) \to 0, \quad m\nabla(Q_\varepsilon - Q_0) \to 0, \quad Q_\varepsilon - Q_0 \to 0$$

uniformly in $L^\infty(\mathbb{R}^2)$ and we take $m = \langle x\rangle^k$ with $k > 1$.

Step 3. We claim that $\Sigma(\Lambda_\varepsilon) \cap \Delta_a \subset B(0, r/2)$ for any $\varepsilon \in (0, \varepsilon_0)$, choosing $\varepsilon_0 > 0$ small enough. On the one hand, we write the two resolvent equations

$$\begin{aligned}\mathscr{R}_{\Lambda_\varepsilon} &= \mathscr{R}_{\mathscr{B}_\varepsilon} - \mathscr{R}_{\Lambda_\varepsilon}\mathscr{A}\mathscr{R}_{\mathscr{B}_\varepsilon}, \\ \mathscr{R}_{\Lambda_\varepsilon} &= \mathscr{R}_{\Lambda_0} - \mathscr{R}_{\Lambda_\varepsilon}(\Lambda_\varepsilon - \Lambda_0)\mathscr{R}_{\Lambda_0},\end{aligned}$$

from what we deduce

$$\mathscr{R}_{\Lambda_\varepsilon} = \mathscr{R}_{\mathscr{B}_\varepsilon} - \mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_\varepsilon} + \mathscr{R}_{\Lambda_\varepsilon}(\Lambda_\varepsilon - \Lambda_0)\mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_\varepsilon},$$

or equivalently

$$\mathscr{R}_{\Lambda_\varepsilon}(I + \mathscr{K}_\varepsilon) = \mathscr{U}_\varepsilon,$$

with

$$\mathscr{K}_\varepsilon := (\Lambda_0 - \Lambda_\varepsilon)\mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_\varepsilon}.$$

On the other hand, from *Step 1*, we have $\mathscr{R}_{\Lambda_0}(z)\mathscr{A}\mathscr{R}_{\mathscr{B}_\varepsilon}(z)$ is bounded in $\mathscr{B}(L^2_m, H^1_m)$ unifomly in $z \in \Omega := \Delta_{-\lambda}\backslash B(0, r/2)$ and $\Lambda_0 - \Lambda_\varepsilon$ is small in $\mathscr{B}(H^1_m, L^2_m)$ for $\varepsilon > 0$ small, so that both estimates together imply

$$\sup_{z\in\Delta_{-\lambda}\backslash B(0,r/2)} \|\mathscr{K}_\varepsilon(z)\|_{\mathscr{B}(L^2)} < 1$$

for any $0 < r < \lambda < \mu$ and $\varepsilon \in (0, \varepsilon_0)$, with $\varepsilon_0 = \varepsilon_0(r, \lambda) > 0$ small enough. This implies that $I + \mathscr{K}_\varepsilon$ is inversible on $\Omega := \Delta_{-\lambda}\backslash B(0, r/2)$ so that

$$\mathscr{R}_{\Lambda_\varepsilon} = \mathscr{U}_\varepsilon(I + \mathscr{K}_\varepsilon)^{-1}$$

is bounded on Ω, which ends the proof.

Step 4. We define now

$$\Pi_\varepsilon := \frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\Lambda_\varepsilon}(z)\,dz, \quad \Gamma := \{z\in\mathbb{C}; |z| = r\},$$

the Dunford projector on the eigenspace associated to eigenvalues of Λ_ε which belong to the ball $B(0,r)$. We write

$$\begin{aligned}\Pi_\varepsilon &= \frac{i}{2\pi}\int_\Gamma \mathscr{U}_\varepsilon\,(I+\mathscr{K}_\varepsilon)^{-1}\,dz\\ &= \frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\mathscr{B}_\varepsilon}\left\{I-\mathscr{K}_\varepsilon\,(I+\mathscr{K}_\varepsilon)^{-1}\right\}dz - \frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\Lambda_0}\,\mathscr{A}\,\mathscr{R}_{\mathscr{B}_\varepsilon}\,(I+\mathscr{K}_\varepsilon)^{-1}\,dz\\ &= -\frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\mathscr{B}_\varepsilon}\mathscr{K}_\varepsilon\,(I+\mathscr{K}_\varepsilon)^{-1}\,dz - \frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\Lambda_0}\,\mathscr{A}\,\mathscr{R}_{\mathscr{B}_\varepsilon}\,(I+\mathscr{K}_\varepsilon)^{-1}\,dz,\end{aligned}$$

and

$$\begin{aligned}\Pi_0 &= \frac{i}{2\pi}\int_\Gamma \{\mathscr{R}_{\mathscr{B}_0} - \mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_0}\}\,dz\\ &= -\frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_0}\{(I+\mathscr{K}_\varepsilon)^{-1} + \mathscr{K}_\varepsilon(I+\mathscr{K}_\varepsilon)^{-1}\}\,dz.\end{aligned}$$

We deduce

$$\begin{aligned}\Pi_\varepsilon - \Pi_0 &= \frac{i}{2\pi}\int_\Gamma (\mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_0} - \mathscr{R}_{\mathscr{B}_\varepsilon})\mathscr{K}_\varepsilon\,(I+\mathscr{K}_\varepsilon)^{-1}\,dz\\ &\quad + \frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\Lambda_0}\mathscr{A}\{\mathscr{R}_{\mathscr{B}_0} - \mathscr{R}_{\mathscr{B}_\varepsilon}\}(I+\mathscr{K}_\varepsilon)^{-1}\,dz\\ &= -\frac{i}{2\pi}\int_\Gamma \mathscr{U}_\varepsilon\mathscr{K}_\varepsilon\,(I+\mathscr{K}_\varepsilon)^{-1}\,dz\\ &\quad + \frac{i}{2\pi}\int_\Gamma \mathscr{R}_{\Lambda_0}\mathscr{A}\mathscr{R}_{\mathscr{B}_0}\{\mathscr{B}_0 - \mathscr{B}_\varepsilon\}\mathscr{R}_{\mathscr{B}_\varepsilon}(I+\mathscr{K}_\varepsilon)^{-1}\,dz.\end{aligned}$$

From what we get $\|\Pi_\varepsilon - \Pi_0\|_{\mathscr{B}(L^2)} = \mathscr{O}(\varepsilon) < 1$ for $\varepsilon > 0$ small enough by taking adavantage of the estimates established in *Step 1* and *Step 2*. By classical operator theory (see for instance the arguments presented in [191] in order to prove [191, Chap 1, (4.43)]) one deduces that $\dim \Pi_\varepsilon = \dim \Pi_0 = 1$. On the other hand, at first glance we have $\Lambda_\varepsilon^* 1 = 0$ and $1 \in (L^2_m)'$ so that $0 \in \Sigma(\Lambda_\varepsilon^*) = \Sigma(\Lambda_\varepsilon)$, and 0 is the only spectral value of Λ_ε in the ball $B(0,r)$. □

7.4.4 Semigroup decay estimates for $\mathscr{L}_{1,1}$

We are now able to deduce a nice semigroup decay estimate on $S_{\mathscr{L}_{1,1}}$ from the previous estimates on the resolvent.

Proposition 7.3. *With the notation of Proposition 7.2, for all $0 < \lambda < \mu$ and all $\varepsilon \in (0, \varepsilon_*)$ there holds, for any $g \in L^2_{m,0}$,*

$$\|S_{\Lambda_\varepsilon}(t)g\|_{L^2_m} \lesssim e^{-\lambda t}\|g\|_{L^2_m}.$$

Proof. It is a consequence of Proposition 7.2 and of the splitting structure of the operator Λ_ε. More precisely, we may for instance apply the quantitative mapping theorem [192, Theorem 2.1], where it is worth emphasizing that $\mathscr{R}_{\mathscr{B}_\varepsilon}(z) : L^2_m \to H^1_m \subset D(\Lambda_\varepsilon^{1/2})$ with uniformly bound in $z \in \Delta_{-\lambda}$, which is a strong enough information in order to establish [192, (2.23)] without checking [192, **(H2)**]. Alternatively, one can use the Gearhart-Prüss-Greiner theorem [193–195] in order to get the same conclusion. $\square$

Thanks to the previous estimate for S_{Λ_ε} and the estimates for $S_{\mathscr{B}_\varepsilon}$ in Lemma 7.10, we are able to deduce semigroup estimates for S_{Λ_ε} (in Propositions 7.4 below) similar to those satisfied by $S_{\mathscr{B}_\varepsilon}$.

We start by observing that, thanks to Duhamel's formula, we have

$$S_{\Lambda_\varepsilon} = S_{\mathscr{B}_\varepsilon} + S_{\mathscr{B}_\varepsilon}\mathscr{A} * S_{\Lambda_\varepsilon} \quad \text{and} \quad S_{\Lambda_\varepsilon} = S_{\mathscr{B}_\varepsilon} + S_{\Lambda_\varepsilon} * \mathscr{A} S_{\mathscr{B}_\varepsilon}.$$

Denoting $\Pi^\perp g = g - \Pi g$ where Π is the projection onto $\mathrm{Ker}(\Lambda_\varepsilon)$, we obtain

$$S_{\Lambda_\varepsilon}\Pi^\perp = S_{\mathscr{B}_\varepsilon}\Pi^\perp + (S_{\mathscr{B}_\varepsilon}\mathscr{A} * S_{\Lambda_\varepsilon}\Pi^\perp) \quad \text{and} \quad S_{\Lambda_\varepsilon}\Pi^\perp = \Pi^\perp S_{\mathscr{B}_\varepsilon} + (S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A} S_{\mathscr{B}_\varepsilon}),$$

using that $S_{\Lambda_\varepsilon}\Pi^\perp = \Pi^\perp S_{\Lambda_\varepsilon}$, and iterating this formula also yields

$$S_{\Lambda_\varepsilon}\Pi^\perp = S_{\mathscr{B}_\varepsilon}\Pi^\perp + S_{\mathscr{B}_\varepsilon}\mathscr{A} * \Pi^\perp S_{\mathscr{B}_\varepsilon} + S_{\mathscr{B}_\varepsilon}\mathscr{A} * S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A} S_{\mathscr{B}_\varepsilon}$$

and

$$S_{\Lambda_\varepsilon}\Pi^\perp = \Pi^\perp S_{\mathscr{B}_\varepsilon} + S_{\mathscr{B}_\varepsilon}\Pi^\perp * \mathscr{A} S_{\mathscr{B}_\varepsilon} + S_{\mathscr{B}_\varepsilon}\mathscr{A} * S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A} S_{\mathscr{B}_\varepsilon}$$

Proposition 7.4. *Let $0 \le \lambda < \mu$. There is $\varepsilon_* > 0$ small enough such that for any $\varepsilon \in (0, \varepsilon_*)$ and any $m = \langle x \rangle^k$ with $k > 3$ the following holds:*

1. *For all $g \in L^2_{m,0}$ we have*

$$\|\mathbf{e}_\lambda S_{\Lambda_\varepsilon}(\cdot) g\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda S_{\Lambda_\varepsilon}(\cdot) g\|_{L^2_t H^1_m} \lesssim \|g\|_{L^2_m}.$$

2. *For all $\mathbf{e}_\lambda R \in L^2_t H^{-1}_m$ with $\Pi R = 0$, we have*

$$\|\mathbf{e}_\lambda (S_{\Lambda_\varepsilon} * R)\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda (S_{\Lambda_\varepsilon} * R)\|_{L^2_t H^1_m} \lesssim \|\mathbf{e}_\lambda R\|_{L^2_t H^{-1}_m}.$$

Proof. Step 1: Proof of (1). Remark that $g = \Pi^\perp g$ since $g \in L^2_{m,0}$. The first estimate is nothing but Proposition 7.3. We in particular deduce from this one that

$$\|\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp\|_{L^1_t \mathscr{B}(L^2_m)} \lesssim \|\mathbf{e}_{\lambda'} S_{\Lambda_\varepsilon}\Pi^\perp\|_{L^\infty_t \mathscr{B}(L^2_m)} \lesssim 1, \tag{7.33}$$

by choosing $\lambda < \lambda' < 1$. For the second one, we write

$$S_{\Lambda_\varepsilon}(\cdot) g = S_{\mathscr{B}_\varepsilon}(\cdot) g + (S_{\mathscr{B}_\varepsilon} * \mathscr{A} S_{\Lambda_\varepsilon} g)(\cdot).$$

Thanks to Lemma 7.10 and the first estimate, we have

$$\|\mathbf{e}_\lambda S_{\mathscr{B}_\varepsilon}(\cdot) g\|_{L^2_t H^1_m} \lesssim \|g\|_{L^2_m}$$

and

$$\begin{aligned} \|\mathbf{e}_\lambda (S_{\mathscr{B}_\varepsilon} * \mathscr{A} S_{\Lambda_\varepsilon})(\cdot) g\|_{L^2_t H^1_m} &\lesssim \|\mathbf{e}_\lambda \mathscr{A} S_{\Lambda_\varepsilon}(\cdot) g\|_{L^2_t H^{-1}_m} \\ &\lesssim \|\mathbf{e}_\lambda S_{\Lambda_\varepsilon}(\cdot) g\|_{L^2_t L^2_m} \\ &\lesssim \|g\|_{L^2_m}, \end{aligned}$$

from what we immediately obtain the second estimate.

Step 2: Proof of (2). We remark that $R(t) = \Pi^\perp R(t)$ for all $t \geq 0$. For the first term we write

$$S_{\Lambda_\varepsilon} * R = S_{\Lambda_\varepsilon}\Pi^\perp * R = \Pi^\perp(S_{\mathscr{B}_\varepsilon} * R) + S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A}S_{\mathscr{B}_\varepsilon} * R,$$

and thus

$$\mathbf{e}_\lambda(S_{\Lambda_\varepsilon} * R) = \mathbf{e}_\lambda\Pi^\perp(S_{\mathscr{B}_\varepsilon} * R) + (\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp) * \mathscr{A}[\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)].$$

Therefore we get

$$\begin{aligned}
\|\mathbf{e}_\lambda(S_{\Lambda_\varepsilon} * R)\|_{L^\infty_t L^2_m} \lesssim & \|\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)\|_{L^\infty_t L^2_m} + \|(\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp) * \mathscr{A}[\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)]\|_{L^\infty_t L^2_m} \\
\lesssim & \|\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)\|_{L^\infty_t L^2_m} \\
& + \|\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp\|_{L^1_t(\mathscr{B}(L^2_m))}\|\mathscr{A}\|_{\mathscr{B}(L^2_m)}\|\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)\|_{L^\infty_t L^2_m} \\
\lesssim & \|\mathbf{e}_\lambda R\|_{L^2_t H^{-1}_m},
\end{aligned}$$

where we have used Lemma 7.10 and (7.33) in last line.

For the second term we write

$$S_{\Lambda_\varepsilon} * R = S_{\Lambda_\varepsilon}\Pi^\perp * R = \Pi^\perp(S_{\mathscr{B}_\varepsilon} * R) + S_{\mathscr{B}_\varepsilon}\Pi^\perp * \mathscr{A}S_{\mathscr{B}_\varepsilon} * R + S_{\mathscr{B}_\varepsilon}\mathscr{A} * S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A}S_{\mathscr{B}_\varepsilon} * R$$

thus

$$\mathbf{e}_\lambda(S_{\Lambda_\varepsilon} * R) = \Pi^\perp\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R) + \mathbf{e}_\lambda[S_{\mathscr{B}_\varepsilon} * (\Pi^\perp\mathscr{A}S_{\mathscr{B}_\varepsilon} * R)] + \mathbf{e}_\lambda[(S_{\mathscr{B}_\varepsilon}\mathscr{A}) * (S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A}S_{\mathscr{B}_\varepsilon} * R)].$$

From Lemma 7.10 we have

$$\|\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)\|_{L^2_t H^1_m} \lesssim \|\mathbf{e}_\lambda R\|_{L^2_t H^{-1}_m},$$

and

$$\|\mathbf{e}_\lambda[S_{\mathscr{B}_\varepsilon} * (\Pi^\perp\mathscr{A}S_{\mathscr{B}_\varepsilon} * R)]\|_{L^2_t H^1_m} \lesssim \|\mathbf{e}_\lambda(\Pi^\perp\mathscr{A}S_{\mathscr{B}_\varepsilon} * R)\|_{L^2_t H^{-1}_m},$$

thus we deduce

$$\begin{aligned}
\|\mathbf{e}_\lambda[S_{\mathscr{B}_\varepsilon} * (\Pi^\perp\mathscr{A}S_{\mathscr{B}_\varepsilon} * R)]\|_{L^2_t H^1_m} &\lesssim \|\mathbf{e}_\lambda(\Pi^\perp\mathscr{A}S_{\mathscr{B}_\varepsilon} * R)\|_{L^2_t L^2_m} \\
&\lesssim \|\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)\|_{L^2_t L^2_m} \\
&\lesssim \|\mathbf{e}_\lambda R\|_{L^2_t H^{-1}_m}.
\end{aligned}$$

Moreover, we also have

$$\begin{aligned}
&\|\mathbf{e}_\lambda[(S_{\mathscr{B}_\varepsilon}\mathscr{A}) * (S_{\Lambda_\varepsilon}\Pi^\perp * \mathscr{A}S_{\mathscr{B}_\varepsilon} * R)]\|_{L^2_t H^1_m} \\
\lesssim & \|(\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp) * [\mathscr{A}\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)]\|_{L^2_t H^{-1}_m} \\
\lesssim & \|(\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp) * [\mathscr{A}\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)]\|_{L^2_t L^2_m} \\
\lesssim & \|\mathbf{e}_\lambda S_{\Lambda_\varepsilon}\Pi^\perp\|_{L^1_t(\mathscr{B}(L^2_m))}\|\mathscr{A}\|_{\mathscr{B}(L^2_m)}\|[\mathbf{e}_\lambda(S_{\mathscr{B}_\varepsilon} * R)]\|_{L^2_t L^2_m} \\
\lesssim & \|\mathbf{e}_\lambda R\|_{L^2_t H^{-1}_m}.
\end{aligned}$$

□

7.5 Estimates for $\mathscr{L}_{2,2}$

Recall that

$$\mathscr{L}_{2,2}w = \frac{1}{\varepsilon}\Delta w + \mu x \cdot \nabla w + \nabla\kappa * [Q\nabla w].$$

Lemma 7.11. *Let $p > 2$, then there holds*

$$\int (\mathscr{L}_{2,2}w)w^{p-1} \leq -\frac{c_p}{\varepsilon}\|\nabla w^{p/2}\|_{L^2}^2 - \frac{2\mu}{p}\|w\|_{L^p}^p + C\|w\|_{L^p}^{p/p'}\|\nabla w\|_{L^p}.$$

Proof. We compute

$$\int (\mathscr{L}_{2,2}w)w^{p-1} = -\frac{c_p}{\varepsilon}\int |\nabla w^{p/2}|^2 - 2\frac{\mu}{p}\int w^p + \int w^{p-1}\nabla\kappa * (Q\nabla w),$$

then we remark that we have

$$\begin{aligned}\int w^{p-1}\nabla\kappa * (Q\nabla w) &\leq \|w\|_{L^p}^{p/p'}\|\nabla\kappa * (Q\nabla w)\|_{L^p}\\ &\lesssim \|w\|_{L^p}^{p/p'}\|Q\nabla w\|_{L^p_2}\\ &\lesssim \|w\|_{L^p}^{p/p'}\|\nabla w\|_{L^p},\end{aligned}$$

where we have used Hölder's inequality in the first line, Lemma 7.6 in the third line, and Lemma 7.1 in the fourth one. □

We observe that we have

$$\nabla\mathscr{L}_{2,2}w = \frac{1}{\varepsilon}\Delta(\nabla w) + \mu x \cdot \nabla(\nabla w) + \mu\nabla w + \nabla^2\kappa * [Q\nabla w],$$

where we denote

$$(x \cdot \nabla\Phi)_i = x_\ell\partial_\ell\Phi_i, \quad (\nabla^2\kappa * \Phi)_i = \partial_{i\ell}\kappa * \Phi_\ell$$

for any vector Φ.

Lemma 7.12. *There holds*

$$\langle\mathscr{L}_{2,2}w, w\rangle_{\dot{H}^1} = -\frac{1}{\varepsilon}\|\nabla^2 w\|_{L^2}^2 - \|Q^{1/2}\nabla w\|_{L^2}^2.$$

Proof. A straightforward computation gives

$$\begin{aligned}\langle\nabla\mathscr{L}_{2,2}w, \nabla w\rangle_{L^2} &= \int (\frac{1}{\varepsilon}\Delta\nabla w + \mu\nabla w + \mu x \cdot \nabla^2 w + \nabla^2\kappa * [Q\nabla w])\nabla w\\ &= -\frac{1}{\varepsilon}\int |\nabla^2 w|^2 - \int Q|\nabla w|^2,\end{aligned}$$

where we have performed two integrations by parts for the last term and we have used the identity $\Delta\kappa = -\delta$. □

As a consequence of previous estimates, we obtain decay and regularization estimates for the semigroup $S_{\mathscr{L}_{2,2}}$ in the following result.

Lemma 7.13. *Let $p \in (2, \infty)$ and $0 \le \vartheta < \frac{2\mu}{p}$. There is $\varepsilon_2 > 0$ small enough such that for any $\varepsilon \in (0, \varepsilon_2)$ the following holds:*

1. *For all $w \in L^p \cap \dot{H}^1$ we have*

$$\|\mathbf{e}_\vartheta S_{\mathscr{L}_{2,2}}(\cdot)w\|_{L_t^\infty(L^p\cap\dot{H}^1)} + \|\mathbf{e}_\vartheta S_{\mathscr{L}_{2,2}}(\cdot)w\|_{L_t^2L^p} + \frac{1}{\sqrt{\varepsilon}}\|\mathbf{e}_\vartheta S_{\mathscr{L}_{2,2}}(\cdot)w\|_{L_t^2\dot{H}^2} \lesssim \|w\|_{L^p\cap\dot{H}^1}.$$

2. *For all $\mathbf{e}_\vartheta S \in L_t^2(L^p \cap \dot{H}^1)$ we have*

$$\left\|\mathbf{e}_\vartheta(S_{\mathscr{L}_{2,2}} * S)\right\|_{L_t^\infty(L^p\cap\dot{H}^1)} + \left\|\mathbf{e}_\vartheta(S_{\mathscr{L}_{2,2}} * S)\right\|_{L_t^2L^p} + \frac{1}{\sqrt{\varepsilon}}\left\|\mathbf{e}_\vartheta(S_{\mathscr{L}_{2,2}} * S)\right\|_{L_t^2\dot{H}^2} \lesssim \|\mathbf{e}_\vartheta S\|_{L_t^2L^p} + \|\mathbf{e}_\vartheta S\|_{L_t^2\dot{H}^1}.$$

Proof. Consider $\phi = \mathbf{e}_\vartheta S_{\mathscr{L}_{2,2}}(\cdot)w$ which is a solution to the evolution equation

$$\partial_t\phi = \mathscr{L}_{2,2}\phi + \vartheta\phi, \quad \phi(0) = w.$$

Thanks to Lemma 7.11 we have

$$\frac{1}{p}\frac{d}{dt}\|\phi\|_{L^p}^p \le \left(\vartheta - \frac{2\mu}{p}\right)\|\phi\|_{L^p}^p + C\|\phi\|_{L^p}^{p/p'}\|\nabla\phi\|_{L^p},$$

therefore, using also that

$$\frac{1}{2}\frac{d}{dt}\|\phi\|_{L^p}^2 = \|\phi\|_{L^p}^{2-p}\left(\frac{1}{p}\frac{d}{dt}\|\phi\|_{L^p}^p\right),$$

we get

$$\frac{1}{2}\frac{d}{dt}\|\phi\|_{L^p}^2 \le \left(\vartheta - \frac{2\mu}{p}\right)\|\phi\|_{L^p}^2 + C\|\phi\|_{L^p}\|\nabla\phi\|_{L^p}.$$

Combining this with Lemma 7.12 yields

$$\frac{1}{2}\frac{d}{dt}\left\{\|\phi\|_{L^p}^2 + \|\nabla\phi\|_{L^2}^2\right\} \le \left(\vartheta - \frac{2\mu}{p}\right)\|\phi\|_{L^p}^2 + \vartheta\|\nabla\phi\|_{L^2}^2 - \frac{1}{\varepsilon}\|\nabla^2\phi\|_{L^2}^2 + C\|\phi\|_{L^p}\|\nabla\phi\|_{L^p}.$$

Using (7.25) and Young's inequality we get

$$C\|\phi\|_{L^p}\|\nabla\phi\|_{L^p} \le C\|\phi\|_{L^p}^{2-\theta}\|\nabla^2\phi\|_{L^2}^\theta \le C\varepsilon^{\frac{\theta}{2-\theta}}\|\phi\|_{L^p}^2 + \frac{1}{2\varepsilon}\|\nabla^2\phi\|_{L^2}^2$$

with $\theta = p/(2+p)$, thus we obtain

$$\frac{1}{2}\frac{d}{dt}\left\{\|\phi\|_{L^p}^2 + \|\nabla\phi\|_{L^2}^2\right\} \le -\left(\frac{2\mu}{p} - \vartheta - C\varepsilon^{\frac{\theta}{2-\theta}}\right)\|\phi\|_{L^p}^2 + \vartheta\|\nabla\phi\|_{L^2}^2 - \frac{1}{2\varepsilon}\|\nabla^2\phi\|_{L^2}^2.$$

Finally, taking $\varepsilon > 0$ small enough and using Lemma 7.8 we hence deduce

$$\frac{1}{2}\frac{d}{dt}\left\{\|\phi\|_{L^p}^2 + \|\nabla\phi\|_{L^2}^2\right\} \le -\sigma\|\phi\|_{L^p}^2 - \frac{1}{4\varepsilon}\|\nabla^2\phi\|_{L^2}^2$$

for some $\sigma > 0$, from which (1) follows by Grönwall's lemma.

We now consider $\phi = \mathbf{e}_\vartheta(S_{\mathscr{L}_{2,2}} * S)$ which is a solution to the evolution equation

$$\partial_t \phi = \mathscr{L}_{2,2}\phi + \vartheta\phi + \mathbf{e}_\vartheta S, \quad \phi(0) = 0.$$

Arguing as above we have

$$\frac{1}{2}\frac{d}{dt}\left\{\|\phi\|_{L^p}^2 + \|\nabla\phi\|_{L^2}^2\right\} \le -\left(\frac{2\mu}{p} - \vartheta - C\varepsilon^{\frac{\theta}{2-\theta}}\right)\|\phi\|_{L^p}^2 + \vartheta\|\nabla\phi\|_{L^2}^2 - \frac{1}{2\varepsilon}\|\nabla^2\phi\|_{L^2}^2 + C\|\phi\|_{L^p}\|\mathbf{e}_\vartheta S\|_{L^p} + C\|\nabla\phi\|_{L^2}\|\mathbf{e}_\vartheta \nabla S\|_{L^2}.$$

By Young's inequality we get, for any $\beta > 0$ and some $C_\beta > 0$,

$$\frac{1}{2}\frac{d}{dt}\left\{\|\phi\|_{L^p}^2 + \|\nabla\phi\|_{L^2}^2\right\} \le -\left(\frac{2\mu}{p} - \vartheta - C\varepsilon^{\frac{\theta}{2-\theta}} - \beta\right)\|\phi\|_{L^p}^2 + (\vartheta+\beta)\|\nabla\phi\|_{L^2}^2 - \frac{1}{2\varepsilon}\|\nabla^2\phi\|_{L^2}^2 + C_\beta\|\mathbf{e}_\vartheta S\|_{L^p}^2 + C_\beta\|\mathbf{e}_\vartheta \nabla S\|_{L^2}^2.$$

We then conclude to (2) arguing as before by taking $\varepsilon, \beta > 0$ small enough, using Lemma 7.8 and applying Grönwall's lemma again. □

7.6 Semigroup estimates for the linearized system

We start with some estimates on the out of the diagonal operators $\mathscr{L}_{1,2}$ and $\mathscr{L}_{2,1}$.

Lemma 7.14. *Let $m = \langle x\rangle^k$ with $k > 3$ and $p > 2$. For any $w \in L^p \cap \dot H^2$ and $g \in H^1_m$, there holds:*

$$\|\mathscr{L}_{1,2} w\|_{H^{-1}_m} \lesssim \|w\|_{L^p}^{1-\theta}\|\nabla^2 w\|_{L^2}^\theta \quad \textit{with} \quad \theta = \frac{p}{2+p},$$

and

$$\|\mathscr{L}_{2,1} g\|_{L^p} + \|\mathscr{L}_{2,1} g\|_{\dot H^1} \lesssim \|g\|_{H^1_m}.$$

Proof. For the first estimate we obtain

$$\begin{aligned}\|\mathrm{div}(\nabla Q\cdot\nabla w)\|_{H^{-1}_m} &\lesssim \|\nabla Q\cdot\nabla w\|_{L^2_m}\\ &\lesssim \|\nabla w\|_{L^p}\\ &\lesssim \|w\|_{L^p}^{1-\theta}\|\nabla^2 w\|_{L^2}^\theta\end{aligned}$$

where we haved used the exponential decay of Q (see Lemma 7.1) in the third line together with Hölder's inequality, and also (7.25) in the last one.

For the second estimate we first compute

$$\|\mathscr{L}_{2,1} g\|_{L^p} \le \|g\|_{L^p} + \|\nabla\kappa * [g\nabla P]\|_{L^p} + \|\nabla\kappa * [Q\nabla\kappa * g]\|_{L^p}.$$

On the one hand we have

$$\begin{aligned}\|\nabla\kappa * [g\nabla P]\|_{L^p} &\lesssim \|g\nabla P\|_{L^p_2}\\ &\lesssim \|g\|_{L^p_1}\end{aligned}$$

where we have used Lemma 7.6 in the second line and Lemma 7.1 in the third one. Similarly, we also obtain

$$\begin{aligned}\|\nabla\kappa * [Q\nabla\kappa * g]\|_{L^p} &\lesssim \|Q\nabla\kappa * g\|_{L^p_2}\\ &\lesssim \|\nabla\kappa * g\|_{L^p}\\ &\lesssim \|g\|_{L^p_2}\end{aligned}$$

by using sucessively Lemma 7.6, Lemma 7.1 and Lemma 7.6 again. Therefore we get

$$\|\mathscr{L}_{2,1}g\|_{L^p} \lesssim \|g\|_{L^p_2} \lesssim \|g\|_{H^1_m}$$

thanks to the Sobolev embedding $H^2(\mathbb{R}^2) \subset L^p(\mathbb{R}^2)$.

For the third estimate we have

$$\begin{aligned}\|\mathscr{L}_{2,1}g\|_{\dot H^1} &\le \|\nabla g\|_{L^2} + \|\nabla^2\kappa * [g\nabla P]\|_{L^2} + \|\nabla^2\kappa * [Q\nabla\kappa * g]\|_{L^2}\\ &\lesssim \|\nabla g\|_{L^2} + \|g\nabla P\|_{L^2} + \|Q\nabla\kappa * g\|_{L^2}.\end{aligned}$$

Thanks to Lemma 7.1 we get

$$\|g\nabla P\|_{L^2} \lesssim \|g\|_{L^2}$$

and also

$$\|Q\nabla\kappa * g\|_{L^2} \lesssim \|\kappa * g\|_{L^p} \lesssim \|g\|_{L^2_2}$$

where we have used Hölder's inequality and then Lemma 7.6. This implies

$$\|\mathscr{L}_{2,1}g\|_{\dot H^1} \le \|g\|_{H^1_m}.$$

□

As a consequence of Proposition 7.4 and Lemmas 7.13 and 7.14, and recalling the definition of the spaces $X := L^2_m \times (L^p \cap \dot H^1)$ and $Y = H^1_m \times (L^p \cap \dot H^2)$, we obtain:

Proposition 7.5. *Let $0 \le \lambda < \mu$. There is $\varepsilon_* > 0$ small enough such that for any $\varepsilon \in (0, \varepsilon_*)$ there holds:*

1. *For any $(g_0, w_0) \in L^2_{m,0} \times (L^p \cap \dot H^1)$ with $2 < p < \frac{2\mu}{\lambda}$ we have*

$$\|\mathbf{e}_\lambda S_{\mathscr{L}}(\cdot)(g_0,w_0)\|_{L^\infty_t(X)} + \|\mathbf{e}_\lambda S_{\mathscr{L}}(\cdot)(g_0,w_0)\|_{L^2_t(Y)} \lesssim \|(g_0,w_0)\|_X.$$

2. *For any $\mathbf{e}_\lambda \mathscr{R} = \mathbf{e}_\lambda(\mathscr{R}_1, \mathscr{R}_2) \in L^2_t(H^{-1}_m \times (L^p \cap \dot H^1))$ with $\Pi\mathscr{R}_1 = 0$ we have*

$$\|\mathbf{e}_\lambda(S_{\mathscr{L}} * \mathscr{R})\|_{L^\infty_t(X)} + \|\mathbf{e}_\lambda(S_{\mathscr{L}} * \mathscr{R})\|_{L^2_t(Y)} \lesssim \|\mathbf{e}_\lambda \mathscr{R}\|_{L^2_t(H^{-1}_m \times (L^p\cap \dot H^1))}.$$

Proof. We split the proof into two steps.

Step 1. We write

$$(g(t), w(t)) = S_{\mathscr{L}}(t)(g_0, w_0)$$

so that

$$g(t) = S_{\mathscr{L}_{1,1}}(t)g_0 + (S_{\mathscr{L}_{1,1}} * \mathscr{L}_{1,2}w)(t)$$

and

$$w(t) = S_{\mathscr{L}_{2,2}}(t)w_0 + (S_{\mathscr{L}_{2,2}} * \mathscr{L}_{2,1}g)(t).$$

We observe that $\langle \mathscr{L}_{1,2} w \rangle = 0$ so that $\Pi(\mathscr{L}_{1,2} w) = 0$ and then we can hereafter apply the results of Proposition 7.4 to $S_{\mathscr{L}_{1,1}} * \mathscr{L}_{1,2} w$.

From Proposition 7.4, we have

$$\|\mathbf{e}_\lambda S_{\mathscr{L}_{1,1}}(\cdot) g_0\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda S_{\mathscr{L}_{1,1}}(\cdot) g_0\|_{L^2_t H^1_m} \lesssim \|g_0\|_{L^2_m},$$

and using also Lemma 7.14, we have

$$\begin{aligned}\|\mathbf{e}_\lambda (S_{\mathscr{L}_{1,1}} * \mathscr{L}_{1,2} w)\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda (S_{\mathscr{L}_{1,1}} * \mathscr{L}_{1,2} w)\|_{L^2_t H^1_m} &\lesssim \|\mathbf{e}_\lambda \mathscr{L}_{1,2} w\|_{L^2_t H^{-1}_m} \\ &\lesssim \|\mathbf{e}_\lambda w\|^{1-\theta}_{L^2_t L^p} \|\mathbf{e}_\lambda w\|^{\theta}_{L^2_t \dot{H}^2}\end{aligned}$$

with $\theta = p/(2+p)$. This implies

$$\|\mathbf{e}_\lambda g\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m} \le C_1 \|g_0\|_{L^2_m} + C_2 \|\mathbf{e}_\lambda w\|^{1-\theta}_{L^2_t L^p} \|\mathbf{e}_\lambda w\|^{\theta}_{L^2_t \dot{H}^2}, \tag{7.34}$$

for some constant $C_1, C_2 > 0$.

Furthermore from Lemma 7.13, we have

$$\|\mathbf{e}_\lambda S_{\mathscr{L}_{2,2}}(\cdot) w_0\|_{L^\infty_t (L^p \cap \dot{H}^1)} + \|\mathbf{e}_\lambda S_{\mathscr{L}_{2,2}}(\cdot) w_0\|_{L^2_t L^p} + \frac{1}{\sqrt{\varepsilon}} \|\mathbf{e}_\lambda S_{\mathscr{L}_{2,2}}(\cdot) w_0\|_{L^2_t \dot{H}^2} \lesssim \|w_0\|_{L^p \cap \dot{H}^1},$$

and using also Lemma 7.14, we have

$$\begin{aligned}\|\mathbf{e}_\lambda (S_{\mathscr{L}_{2,2}} * \mathscr{L}_{2,1} g)\|_{L^\infty_t (L^p \cap \dot{H}^1)} + \|\mathbf{e}_\lambda (S_{\mathscr{L}_{2,2}} * \mathscr{L}_{2,1} g)\|_{L^2_t L^p} + \frac{1}{\sqrt{\varepsilon}} \|\mathbf{e}_\lambda (S_{\mathscr{L}_{2,2}} * \mathscr{L}_{2,1} g)\|_{L^2_t \dot{H}^2} \\ \lesssim \|\mathbf{e}_\lambda \mathscr{L}_{2,1} g\|_{L^2_t L^p} + \|\mathbf{e}_\lambda \mathscr{L}_{2,1} g\|_{L^2_t \dot{H}^1} \\ \lesssim \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m}\end{aligned}$$

which gives,

$$\|\mathbf{e}_\lambda w\|_{L^\infty_t (L^p \cap \dot{H}^1)} + \|\mathbf{e}_\lambda w\|_{L^2_t L^p} + \frac{1}{\sqrt{\varepsilon}} \|\mathbf{e}_\lambda w\|_{L^2_t \dot{H}^2} \le C_3 \|w_0\|_{L^2_m} + C_4 \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m}, \tag{7.35}$$

for some constants $C_3, C_4 > 0$.

Thanks to Young's inequality, we deduce from (7.34) that for any $\beta > 0$ there is some $C_\beta > 0$ such that

$$\|\mathbf{e}_\lambda g\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m} \le C_1 \|g_0\|_{L^2_m} + \beta \|\mathbf{e}_\lambda w\|_{L^2_t L^p} + C_\beta \|\mathbf{e}_\lambda w\|_{L^2_t \dot{H}^2},$$

which combining with (7.35) yields

$$\begin{aligned}\|\mathbf{e}_\lambda g\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m} \le C_1 \|g_0\|_{L^2_m} + \beta C_3 \|w_0\|_{L^2_m} + \beta C_4 \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m} \\ + \sqrt{\varepsilon} C_\beta C_3 \|w_0\|_{L^2_m} + \sqrt{\varepsilon} C_\beta C_4 \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m}.\end{aligned}$$

Therefore choosing first $\beta > 0$ small enough and then $\varepsilon > 0$ small enough gives

$$\|\mathbf{e}_\lambda g\|_{L^\infty_t L^2_m} + \|\mathbf{e}_\lambda g\|_{L^2_t H^1_m} \le C_5 \|g_0\|_{L^2_m} + C_6 \|w_0\|_{L^2_m}$$

for some constants $C_5, C_6 > 0$. We then conclude part (1) by gathering this last estimate with (7.35).

Step 2. We argue similarly as for part (1) by using the estimates of Proposition 7.4–(2) and Lemma 7.13–(2). □

7.7 Stability for the nonlinear equation

We come back to the nonlinear system (7.6) given by

$$\begin{cases} \partial_t g = \mathscr{L}_1(g,w) - \operatorname{div}(g\nabla w) - \operatorname{div}(g\nabla\kappa * g) \\ \partial_t w = \mathscr{L}_2(g,w) + \nabla\kappa * [g\nabla w + g\nabla\kappa * g] \\ (g,w)_{|t=0} = (g_0,w_0), \end{cases} \tag{7.36}$$

that we rewrite

$$\begin{cases} \partial_t(g,w) = \mathscr{L}(g,w) + (-\operatorname{div}(g\nabla w) - \operatorname{div}(g\nabla\kappa * g), \nabla\kappa * [g\nabla w + g\nabla\kappa * g]) \\ (g,w)_{|t=0} = (g_0,w_0). \end{cases} \tag{7.37}$$

Theorem 7.3. *There are $\varepsilon_0, \eta_0 > 0$ such that for any $\varepsilon \in (0,\varepsilon_0)$ and any initial data $(g_0,w_0) \in L^1_{m,0} \times (L^p \cap \dot H^1)$ with $\|(g_0,w_0)\|_{L^2_m \times (L^p\cap \dot H^1)} \le \eta_0$, there exists a unique global solution $(g,w) \in L^\infty_t(L^2_m \times (L^p \cap \dot H^1)) \cap L^2_t(H^1_m \times (L^p \cap \dot H^2))$ to (7.37), which verifies*

$$\|(g,w)\|_{L^\infty_t(L^2_m \times (L^p\cap\dot H^1))} + \|(g,w)\|_{L^2_t(H^1_m \times (L^p\cap \dot H^2))} \lesssim \|(g_0,w_0)\|_{L^2_m \times (L^p\cap\dot H^1)}. \tag{7.38}$$

Moreover we have the decay estimate, for any $\lambda \in (0,\mu)$,

$$\|\mathbf{e}_\lambda(g,w)\|_{L^\infty_t(L^2_m \times (L^p\cap\dot H^1))} + \|\mathbf{e}_\lambda(g,w)\|_{L^2_t(H^1_m \times (L^p\cap \dot H^2))} \lesssim \|(g_0,w_0)\|_{L^2_m \times (L^p\cap\dot H^1)}. \tag{7.39}$$

Proof. We fix $0 < \lambda < \mu$.

Step 1: Existence. Consider the space

$$\mathscr{X} = \left\{ (g,w) \in L^\infty_t(L^1_{m,0} \times (L^p \cap \dot H^1)) \cap L^2_t(H^1_m \times (L^p \cap \dot H^2)) \,\middle|\, \|(g,w)\|_{\mathscr{X}} < \infty \right\}$$

where

$$\|(g,w)\|_{\mathscr{X}} = \|\mathbf{e}_\lambda(g,w)\|_{L^\infty_t(L^2_m \times (L^p\cap\dot H^1))} + \|\mathbf{e}_\lambda(g,w)\|_{L^2_t(H^1_m \times (L^p\cap \dot H^2))}.$$

Define the map $\Phi : \mathscr{X} \to \mathscr{X}$, $(g,w) \mapsto \Phi[g,w]$ given by, for all $t \ge 0$,

$$\Phi[g,w](t) = S_{\mathscr{L}}(t)(g_0,w_0) + (S_{\mathscr{L}} * \mathscr{R}[(g,w),(g,w)])(t),$$

where

$$\mathscr{R}[(g,w),(g,w)] = (R_1[(g,w),(g,w)] + S_1[(g,w),(g,w)], R_2[(g,w),(g,w)] + S_2[(g,w),(g,w)]),$$

with

$$\begin{aligned} R_1[(g,w),(\bar g,\bar w)] &= -\operatorname{div}(g\nabla \bar w) \\ S_1[(g,w),(\bar g,\bar w)] &= -\operatorname{div}(g\nabla\kappa * \bar g) \\ R_2[(g,w),(\bar g,\bar w)] &= \nabla\kappa * [g\nabla\bar w] \\ S_2[(g,w),(\bar g,\bar w)] &= \nabla\kappa * [g\nabla\kappa * \bar g]. \end{aligned}$$

We observe here that the first component of $\Phi[g,w](t)$ belongs to $L^2_{m,0}$ since

$$\Pi R_1[(g,w),(g,w)] = \Pi S_1[(g,w),(g,w)] = 0,$$

thus in the sequel we can apply the results of Proposition 7.5.

Thanks to Proposition 7.5 we have on the one hand

$$\|S_{\mathscr{L}}(\cdot)(g_0, w_0)\|_{\mathcal{X}} \lesssim \|(g_0, w_0)\|_{L^2_m \times (L^p \cap \dot{H}^1)},$$

and on the other hand

$$\begin{aligned}
\|S_{\mathscr{L}} * \mathscr{R}\|_{\mathcal{X}} \lesssim & \|\mathbf{e}_\lambda R_1[(g,w),(g,w)]\|_{L^2_t H^{-1}_m} + \|\mathbf{e}_\lambda S_1[(g,w),(g,w)]\|_{L^2_t H^{-1}_m} \\
& + \|\mathbf{e}_\lambda R_2[(g,w),(g,w)]\|_{L^2_t L^p} + \|\mathbf{e}_\lambda R_2[(g,w),(g,w)]\|_{L^2_t \dot{H}^1} \\
& + \|\mathbf{e}_\lambda S_2[(g,w),(g,w)]\|_{L^2_t L^p} + \|\mathbf{e}_\lambda S_2[(g,w),(g,w)]\|_{L^2_t \dot{H}^1},
\end{aligned}$$

and we now estimate each term separately.

For the term associated to R_1, we first write

$$\begin{aligned}
\|\mathrm{div}(g\nabla w)\|_{H^{-1}_m} &\lesssim \|g\nabla w\|_{L^2_m} \\
&\lesssim \|g\|_{L^4_m} \|\nabla w\|_{L^4} \\
&\lesssim \|g\|^{1/2}_{L^2_m} \|g\|^{1/2}_{H^1_m} \|\nabla w\|^{1/2}_{L^2} \|\nabla^2 w\|^{1/2}_{L^2},
\end{aligned}$$

where we have used Hölder's inequality in the second line, and Ladyzhenskaya inequality (7.26) in the last one. Therefore we obtain

$$\begin{aligned}
\|\mathbf{e}_\lambda R_1[(g,w),(g,w)]\|_{L^2_t H^{-1}_m} &\lesssim \|\mathbf{e}_\lambda g\|^{1/2}_{L^\infty_t L^2_m} \|\nabla w\|^{1/2}_{L^\infty_t L^2} \|\mathbf{e}_\lambda g\|^{1/2}_{L^2_t H^1_m} \|\nabla^2 w\|^{1/2}_{L^2_t L^2} \\
&\lesssim \|(g,w)\|^2_{\mathcal{X}}.
\end{aligned} \tag{7.40}$$

For the term associated to S_1, arguing similarly as above with Hölder's inequality and (7.26) we get

$$\begin{aligned}
\|\mathrm{div}(g\nabla\kappa * g)\|_{H^{-1}_m} &\lesssim \|g\nabla\kappa * g\|_{L^2_m} \\
&\lesssim \|g\|_{L^4_m} \|\nabla\kappa * g\|_{L^4} \\
&\lesssim \||g\|^{1/2}_{L^2_m} \|g\|^{1/2}_{H^1_m} \|\nabla\kappa * g\|^{1/2}_{L^2} \|\nabla^2\kappa * g\|^{1/2}_{L^2} \\
&\lesssim \|g\|_{L^2_m} \|g\|_{H^1_m},
\end{aligned}$$

where in the last line we have also used Lemma 7.5. We hence obtain

$$\begin{aligned}
\|\mathbf{e}_\lambda S_1[(g,w),(g,w)]\|_{L^2_t H^{-1}_m} &\lesssim \|\mathbf{e}_\lambda g\|_{L^\infty_t L^2_m} \|g\|_{L^2_t H^1_m} \\
&\lesssim \|(g,w)\|^2_{\mathcal{X}}.
\end{aligned} \tag{7.41}$$

For the term associated to R_2 we have thanks to Lemma 7.6

$$\|\nabla\kappa * (g\nabla w)\|_{L^p} \lesssim \|g\nabla w\|_{L^2_m}$$

and also

$$\|\nabla\kappa * (g\nabla w)\|_{\dot{H}^1} \lesssim \|g\nabla w\|_{L^2},$$

thus we can argue as above for obtaining (7.40) to deduce

$$\|\mathbf{e}_\lambda R_2[(g,w),(g,w)]\|_{L^2_t L^p} + \|\mathbf{e}_\lambda R_2[(g,w),(g,w)]\|_{L^2_t \dot{H}^1} \lesssim \|(g,w)\|^2_{\mathcal{X}}.$$

Finally, for the term associated to S_2, we write thanks to Lemma 7.6

$$\|\nabla\kappa * (g\nabla\kappa * g)\|_{L^p} \lesssim \|g\nabla\kappa * g\|_{L^2_m}$$

and also

$$\|\nabla\kappa * (g\nabla\kappa * g)\|_{\dot H^1} \lesssim \|g\nabla\kappa * g\|_{L^2},$$

therefore arguing as for obtaining (7.41) yields

$$\|\mathbf{e}_\lambda S_2[(g,w),(g,w)]\|_{L^2_t L^p} + \|\mathbf{e}_\lambda S_2[(g,w),(g,w)]\|_{L^2_t \dot H^1} \lesssim \|(g,w)\|_{\mathscr{X}}^2.$$

We have hence obtained a first estimate

$$\|\Phi[g,w]\|_{\mathscr{X}} \le C_0\|(g_0,w_0)\|_{H^1_m\times H^2} + C_1\|(g,w)\|_{\mathscr{X}}^2. \tag{7.42}$$

For $(g,w),(\bar g,\bar w)\in\mathscr{X}$, we remark that

$$\Phi[g,w]-\Phi[\bar g,\bar w] = S_{\mathscr{L}} * (R_1^*+S_1^*, R_2^*+S_2^*)$$

with

$$\begin{aligned}
R_1^* &= R_1[(g,w),(g,w)] - R_1[(\bar g,\bar w),(\bar g,\bar w)]\\
&= R_1[(g,w),(g,w)-(\bar g,\bar w)] + R_1[(g,w)-(\bar g,\bar w),(\bar g,\bar w)]\\
S_1^* &= S_1[(g,w),(g,w)] - S_1[(\bar g,\bar w),(\bar g,\bar w)]\\
&= S_1[(g,w),(g,w)-(\bar g,\bar w)] + S_1[(g,w)-(\bar g,\bar w),(\bar g,\bar w)]\\
R_2^* &= R_2[(g,w),(g,w)] - R_2[(\bar g,\bar w),(\bar g,\bar w)]\\
&= R_2[(g,w),(g,w)-(\bar g,\bar w)] + R_2[(g,w)-(\bar g,\bar w),(\bar g,\bar w)]\\
S_2^* &= S_2[(g,w),(g,w)] - S_2[(\bar g,\bar w),(\bar g,\bar w)]\\
&= S_2[(g,w),(g,w)-(\bar g,\bar w)] + S_2[(g,w)-(\bar g,\bar w),(\bar g,\bar w)].
\end{aligned}$$

Arguing as above, we also deduce a second estimate

$$\|\Phi(g,w)-\Phi(\bar g,\bar w)\|_{\mathscr{X}} \le C_2\left(\|(g,w)\|_{\mathscr{X}} + \|(\bar g,\bar w)\|_{\mathscr{X}}\right)\|(g,w)-(\bar g,\bar w)\|_{\mathscr{X}}. \tag{7.43}$$

As a consequence of the estimates (7.42) and (7.43), one can construct a global mild solution $(g,w)\in X$ to (7.37) by a standard fixed-point argument, which verifies moreover estimates (7.38) and (7.39).

Step 2: Uniqueness. Let (g,w) and $(\bar g,\bar w)$ be two solutions to (7.37) in $L^\infty_t(X)\cap L^2_t(Y)$ associated to the same initial data (g_0,w_0) such that $\|(g_0,w_0)\|_{L^2_m\times(L^p\cap\dot H^1)}\le\eta_0$ and verifying estimate (7.38).

Arguing as in the previous step, we obtain that

$$\begin{aligned}
\|(g,w)-(\bar g,\bar w)\|_{L^\infty_t(X)\cap L^2_t(Y)} \lesssim &\left(\|(g,w)\|_{L^\infty_t(X)\cap L^2_t(Y)} + \|(\bar g,\bar w)\|_{L^\infty_t(X)\cap L^2_t(Y)}\right)\\
&\times\|(g,w)-(\bar g,\bar w)\|_{L^\infty_t(X)\cap L^2_t(Y)},
\end{aligned}$$

and we conclude using (7.38) and the fact that $\eta_0>0$ is small enough. □

Bibliography

[1] Y. Viossat and R. Noble, "A theoretical analysis of tumour containment," *Nature Ecology & Evolution*, vol. 5, pp. 826–835, jun 2021.

[2] R. B. Martin, M. E. Fisher, R. F. Minchin, and K. L. Teo, "Optimal control of tumor size used to maximize survival time when cells are resistant to chemotherapy," *Mathematical Biosciences*, vol. 110, pp. 201–219, jul 1992.

[3] H. C. Monro and E. A. Gaffney, "Modelling chemotherapy resistance in palliation and failed cure," *Journal of Theoretical Biology*, vol. 257, pp. 292–302, mar 2009.

[4] R. A. Gatenby, A. S. Silva, R. J. Gillies, and B. R. Frieden, "Adaptive Therapy," *Cancer Research*, vol. 69, pp. 4894–4903, jun 2009.

[5] A. S. Silva, Y. Kam, Z. P. Khin, S. E. Minton, R. J. Gillies, and R. A. Gatenby, "Evolutionary Approaches to Prolong Progression-Free Survival in Breast Cancer," *Cancer Research*, vol. 72, pp. 6362–6370, dec 2012.

[6] C. Carrère, "Optimization of an in vitro chemotherapy to avoid resistant tumours," *Journal of Theoretical Biology*, vol. 413, pp. 24–33, jan 2017.

[7] J. Zhang, J. J. Cunningham, J. S. Brown, and R. A. Gatenby, "Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer," *Nature Communications*, vol. 8, p. 1816, dec 2017.

[8] K. Bacevic, R. Noble, A. Soffar, O. Wael Ammar, B. Boszonyik, S. Prieto, C. Vincent, M. E. Hochberg, L. Krasinska, and D. Fisher, "Spatial competition constrains resistance to targeted cancer therapy," *Nature Communications*, vol. 8, p. 1995, dec 2017.

[9] E. Hansen, R. J. Woods, and A. F. Read, "How to Use a Chemotherapeutic Agent When Resistance to It Threatens the Patient," *Plos Biology*, vol. 15, p. e2001110, feb 2017.

[10] J. A. Gallaher, P. M. Enriquez-Navas, K. A. Luddy, R. A. Gatenby, and A. R. Anderson, "Spatial Heterogeneity and Evolutionary Dynamics Modulate Time to Recurrence in Continuous and Adaptive Cancer Therapies," *Cancer Research*, vol. 78, pp. 2127–2139, apr 2018.

[11] J. J. Cunningham, J. S. Brown, R. A. Gatenby, and K. Staňková, "Optimal control to develop therapeutic strategies for metastatic castrate resistant prostate cancer," *Journal of Theoretical Biology*, vol. 459, pp. 67–78, dec 2018.

[12] C. Pouchol, J. Clairambault, A. Lorz, and E. Trélat, "Asymptotic analysis and optimal control of an integro-differential system modelling healthy and cancer cells exposed to chemotherapy," *Journal de Mathématiques Pures et Appliquées*, vol. 116, pp. 268–308, aug 2018.

[13] C. Carrère and H. Zidani, "Stability and reachability analysis for a controlled heterogeneous population of cells," *Optimal Control Applications and Methods*, vol. 41, pp. 1678–1704, sep 2020.

[14] M. A. Strobl, J. West, Y. Viossat, M. Damaghi, M. Robertson-Tessi, J. S. Brown, R. A. Gatenby, P. K. Maini, and A. R. Anderson, "Turnover modulates the need for a cost of resistance in adaptive therapy," *Cancer Research*, nov 2020.

[15] J. Cunningham, F. Thuijsman, R. Peeters, Y. Viossat, J. Brown, R. Gatenby, and K. Staňková, "Optimal control to reach eco-evolutionary stability in metastatic castrate-resistant prostate cancer," *Plos one*, vol. 15, p. e0243386, dec 2020.

[16] C. Cattani and A. Ciancio, "Qualitative analysis of second-order models of tumor–immune system competition," *Mathematical and Computer Modelling*, vol. 47, no. 11, pp. 1339 – 1355, 2008.

[17] C. Cattani and A. Ciancio, "Early stage of tumor-immune competition with time dependent parameter and external source," *Scientific Bulletin. Series A: Applied Mathematics and Physics. Politehnica University of Bucharest*, vol. 72, 01 2010.

[18] S. Mittal, A. Mukhopadhyay, and S. Chakraborty, "Evolutionary dynamics of the delayed replicator-mutator equation: Limit cycle and cooperation," *Physical Review E*, vol. 101, no. 4, p. 042410, 2020.

[19] K. Li and S. Li, "On the poincaré–bendixson index theorem for a class of piecewise linear differential systems," *Qualitative Theory of Dynamical Systems*, vol. 23, no. 1, p. 9, 2023.

[20] D. Bai, J. Zheng, and Y. Kang, "Global dynamics of a predator-prey model with a smith growth function and the additive predation in prey," *Discrete and Continuous Dynamical Systems-B*, pp. 0–0, 2023.

[21] E. Zhang and S. Shateyi, "Exploring limit cycle bifurcations in the presence of a generalized heteroclinic loop," *Mathematics*, vol. 11, no. 18, p. 3944, 2023.

[22] B. Perthame, "Transport equations in biology," *Springer Science & Business Media*, 2006.

[23] T. Lorenzi and C. Pouchol, "Asymptotic analysis of selection-mutation models in the presence of multiple fitness peaks," *Nonlinearity*, vol. 33, p. 5791–5816, Oct 2020.

[24] L. Desvillettes, P. E. Jabin, S. Mischler, and G. Raoul, "On selection dynamics for continuous structured populations," *Commun. Math. Sci.*, vol. 6, pp. 729–747, 09 2008.

[25] E. Bouin, V. Calvez, N. Meunier, S. Mirrahimi, B. Perthame, G. Raoul, and R. Voituriez, "Invasion fronts with variable motility: phenotype selection, spatial sorting and wave acceleration," *C. R. Acad. Sci. Paris*, vol. 350, no. 15-16, pp. 761–766, 2012.

[26] E. Bouin and V. Calvez, "Travelling waves for the cane toads equation with bounded traits," *Nonlinearity*, vol. 27, pp. 2233–2253, aug 2014.

[27] R. H. Chisholm, T. Lorenzi, L. Desvillettes, and B. D. Hughes, "Evolutionary dynamics of phenotype-structured populations: from individual-level mechanisms to population-level consequences," *Zeitschrift für angewandte Mathematik und Physik*, vol. 67, no. 4, p. 100, 2016.

[28] C. Pouchol and E. Trélat, "Global stability with selection in integro-differential Lotka-Volterra systems modelling trait-structured populations," *Journal of Biological Dynamics*, vol. 12, no. 1, pp. 872–893, 2018.

[29] J. Guilberteau, P. Jain, M. K. Jolly, N. P. Duteil, and C. Pouchol, "An integrative phenotype-structured partial differential equation model for the population dynamics of epithelial-mesenchymal transition," *arXiv preprint arXiv:2309.09569*, 2023.

[30] J. Guilberteau, C. Pouchol, and N. P. Duteil, "Long-time behaviour of an advection-selection equation," *arXiv preprint arXiv:2301.02470*, 2023.

[31] A. R. Anderson and M. A. Chaplain, "Continuous and discrete mathematical models of tumor-induced angiogenesis," *Bulletin of mathematical biology*, vol. 60, no. 5, pp. 857–899, 1998.

[32] L. Corrias, B. Perthame, and H. Zaag, "A chemotaxis model motivated by angiogenesis," *Comptes Rendus Mathematique*, vol. 336, no. 2, pp. 141–146, 2003.

[33] L. Corrias, B. Perthame, and H. Zaag, "Global solutions of some chemotaxis and angiogenesis systems in high space dimensions," *Milan Journal of Mathematics*, vol. 72, pp. 1–28, 2004.

[34] B. Perthame and A. F. Vasseur, "Regularization in keller-segel type systems and the de giorgi method," *Communications in Mathematical Sciences*, vol. 10, no. 2, pp. 463–476, 2012.

[35] A. Blanchet, J. Dolbeault, and B. Perthame, "Two-dimensional keller-segel model: Optimal critical mass and qualitative properties of the solutions.," *Electronic Journal of Differential Equations (EJDE)[electronic only]*, vol. 2006, pp. Paper–No, 2006.

[36] J. F. Campos and J. Dolbeault, "Asymptotic estimates for the parabolic-elliptic keller-segel model in the plane," *Communications in Partial Differential Equations*, vol. 39, no. 5, pp. 806–841, 2014.

[37] G. Egaña Fernández and S. Mischler, "Uniqueness and long time asymptotic for the keller–segel equation: the parabolic–elliptic case," *Archive for Rational Mechanics and Analysis*, vol. 220, pp. 1159–1194, 2016.

[38] A. Blanchet, J. A. Carrillo, and N. Masmoudi, "Infinite time aggregation for the critical patlak-keller-segel model in $\mathbb{R}^2$," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 61, no. 10, pp. 1449–1481, 2008.

[39] T.-E. Ghoul and N. Masmoudi, "Minimal mass blowup solutions for the patlak-keller-segel equation," *Communications on Pure and Applied Mathematics*, vol. 71, no. 10, pp. 1957–2015, 2018.

[40] J. Davila, M. del Pino, J. Dolbeault, M. Musso, and J. Wei, "Existence and stability of infinite time blow-up in the keller-segel system," *arXiv preprint arXiv:1911.12417*, 2019.

[41] M. A. Herrero and J. J. Velázquez, "A blow-up mechanism for a chemotaxis model," *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, vol. 24, no. 4, pp. 633–683, 1997.

[42] P. Raphaël and R. Schweyer, "On the stability of critical chemotactic aggregation," *Mathematische Annalen*, vol. 359, pp. 267–377, 2014.

[43] L. Corrias and B. Perthame, "Critical space for the parabolic-parabolic keller–segel model in $\mathbb{R}^d$," *Comptes rendus. Mathématique*, vol. 342, no. 10, pp. 745–750, 2006.

[44] L. Corrias and B. Perthame, "Asymptotic decay for the solutions of the parabolic–parabolic keller–segel chemotaxis system in critical spaces," *Mathematical and Computer Modelling*, vol. 47, no. 7-8, pp. 755–764, 2008.

[45] V. Calvez, L. Corrias, and M. A. Ebde, "Blow-up, concentration phenomenon and global existence for the keller–segel model in high dimension," *Communications in Partial Differential Equations*, vol. 37, no. 4, pp. 561–584, 2012.

[46] V. Calvez and L. Corrias, "The parabolic-parabolic keller-segel model in $\mathbb{R}^2$," *Communications in Mathematical Sciences*, vol. 6, no. 2, pp. 417–447, 2008.

[47] P. Biler, L. Corrias, and J. Dolbeault, "Large mass self-similar solutions of the parabolic–parabolic keller–segel model of chemotaxis," *Journal of mathematical biology*, vol. 63, pp. 1–32, 2011.

[48] K. Carrapatoso and S. Mischler, "Uniqueness and long time asymptotics for the parabolic–parabolic keller–segel equation," *Communications in Partial Differential Equations*, vol. 42, no. 2, pp. 291–345, 2017.

[49] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles*. crc Press, 2021.

[50] P. Degond and F.-J. Mustieles, "A deterministic approximation of diffusion equations using particles," *SIAM Journal on Scientific and Statistical Computing*, vol. 11, no. 2, pp. 293–310, 1990.

[51] G. Russo, "Deterministic diffusion of particles," *Communications on Pure and Applied Mathematics*, vol. 43, no. 6, pp. 697–733, 1990.

[52] J. A. Carrillo, K. Craig, and F. S. Patacchini, "A blob method for diffusion," *Calculus of Variations and Partial Differential Equations*, vol. 58, pp. 1–53, 2019.

[53] A. Leonard, "Vortex methods for flow simulation," *Journal of Computational Physics*, vol. 37, no. 3, pp. 289–335, 1980.

[54] A. Leonard, "Computing three-dimensional incompressible flows with vortex elements," *Annual Review of Fluid Mechanics*, vol. 17, no. 1, pp. 523–559, 1985.

[55] B. Lucquin-Desreux and S. Mas-Gallic, "A conservative particle approximation for a boundary advection-diffusion problem," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 26, no. 6, pp. 757–791, 1992.

[56] F. Hermeline, "A deterministic particle method for transport diffusion equations: Application to the fokker-planck equation," *Journal of Computational Physics*, vol. 82, no. 1, pp. 122–146, 1989.

[57] P. Degond and S. Mas-Gallic, "The weighted particle method for convection-diffusion equations. i. the case of an isotropic viscosity," *Mathematics of computation*, vol. 53, no. 188, pp. 485–507, 1989.

[58] G. Cottet and S. Mas-Gallic, "A particle method to solve the navier-stokes system," *Numerische Mathematik*, vol. 57, no. 1, pp. 805–827, 1990.

[59] J. Choquin and S. Huberson, "Particles simulation of viscous flow," *Computers & fluids*, vol. 17, no. 2, pp. 397–410, 1989.

[60] P.-A. Raviart, "An analysis of particle methods," in *Numerical methods in fluid dynamics*, pp. 243–324, Springer, 1985.

[61] G.-H. Cottet and P.-A. Raviart, "On particle-in-cell methods for the vlasov-poisson equations," *Transport Theory and Statistical Physics*, vol. 15, no. 1-2, pp. 1–31, 1986.

[62] P. C. W. Davies and C. H. Lineweaver, "Cancer tumors as metazoa 1.0: tapping genes of ancient ancestors," *Physical Biology*, vol. 8, p. 015001, feb 2011.

[63] J. Clairambault, "Stepping from modeling cancer plasticity to the philosophy of cancer," *Frontiers in Genetics*, vol. 11, p. 1160, 2020.

[64] S. Shen and J. Clairambault, "Cell plasticity in cancer cell populations," *F1000Research*, p. 635, 2020.

[65] D. Hanahan, "Hallmarks of Cancer: New Dimensions," *Cancer Discovery*, vol. 12, pp. 31–46, 01 2022.

[66] S. Yuan, R. J. Norgard, and B. Z. Stanger, "Cellular plasticity in cancer.," *Cancer Discov.*, vol. 9, no. 7, pp. 837–851, 2019.

[67] A. M. Nedelcu and R. E. Michod, "Stress responses co-opted for specialized cell types during the early evolution of multicellularity," *BioEssays*, vol. 42, no. 5, p. 2000029, 2020.

[68] G. P. Wagner, E. M. Erkenbrack, and A. C. Love, "Stress-induced evolutionary innovation: A mechanism for the origin of cell types," *BioEssays*, vol. 41, no. 4, pp. 180–188, 2019.

[69] T. Philippi and J. Seger, "Hedging one's evolutionary bets, revisited," *Trends in Ecology & Evolution*, vol. 4, no. 2, pp. 41 – 44, 1989.

[70] J. Seger, "What is bet-hedging?," *Oxford surveys in evolutionary biology*, vol. 4, pp. 182–211, 1987.

[71] C. Collot, T.-E. Ghoul, N. Masmoudi, and V. T. Nguyen, "Refined description and stability for singular solutions of the 2d keller-segel system," *Communications on Pure and Applied Mathematics*, vol. 75, no. 7, pp. 1419–1516, 2022.

[72] C. Collot, T.-E. Ghoul, N. Masmoudi, and V. T. Nguyen, "Spectral analysis for singularity formation of the two dimensional keller–segel system," *Annals of PDE*, vol. 8, no. 1, p. 5, 2022.

[73] R. A. Gatenby, "A change of strategy in the war on cancer," *Nature*, vol. 459, pp. 508–509, may 2009.

[74] P. M. Enriquez-Navas, Y. Kam, T. Das, S. Hassan, A. Silva, P. Foroutan, E. Ruiz, G. Martinez, S. Minton, R. J. Gillies, and R. A. Gatenby, "Exploiting evolutionary principles to prolong tumor control in preclinical models of breast cancer," *Science Translational Medicine*, vol. 8, pp. 327ra24–327ra24, feb 2016.

[75] J. J. Cunningham, "A call for integrated metastatic management," *Nature Ecology and Evolution*, vol. 3, pp. 996–998, jun 2019.

[76] I. Smalley, E. Kim, J. Li, P. Spence, C. J. Wyatt, Z. Eroglu, V. K. Sondak, J. L. Messina, N. A. Babacan, S. S. Maria-Engler, L. De Armas, S. L. Williams, R. A. Gatenby, Y. A. Chen, A. R. Anderson, and K. S. Smalley, "Leveraging transcriptional dynamics to improve braf inhibitor responses in melanoma," *EBioMedicine*, vol. 48, pp. 178–190, oct 2019.

[77] M. Bondarenko, M. Le Grand, Y. Shaked, Z. Raviv, G. Chapuisat, C. Carrère, M.-P. Montero, M. Rossi, E. Pasquier, M. Carré, *et al.*, "Metronomic chemotherapy modulates clonal interactions to prevent drug resistance in non-small cell lung cancer," *Cancers*, vol. 13, p. 2239, may 2021.

[78] J. Wang, Y. Zhang, X. Liu, and H. Liu, "Is the fixed periodic treatment effective for the tumor system without complete information?," *Cancer Management and Research*, vol. 13, p. 8915, nov 2021.

[79] J. Wang, Y. Zhang, X. Liu, and H. Liu, "Optimizing adaptive therapy based on the reachability to tumor resistant subpopulation," *Cancers*, vol. 13, oct 2021.

[80] N. Farrokhian, J. Maltas, M. Dinh, A. Durmaz, P. Ellsworth, M. Hitomi, E. McClure, A. Marusyk, A. Kaznatcheev, and J. G. Scott, "Measuring competitive exclusion in non-small cell lung cancer," *bioRxiv*, 2022.

[81] H. B. Mistry, "On the reporting and analysis of a cancer evolutionary adaptive dosing trial," *Nature Communications*, vol. 12, jan 2021.

[82] J. Zhang, J. J. Cunningham, J. S. Brown, and R. A. Gatenby, "Response to mistry," *Nature Communications*, vol. 12, jan 2021.

[83] A. K. Laird, "Dynamics of tumor growth," *British Journal of Cancer*, vol. 18, pp. 490–502, sep 1964.

[84] L. W. Norton and R. Simon, "Tumor size, sensitivity to therapy, and design of treatment schedules.," *Cancer treatment reports*, vol. 61, pp. 1307–1317, oct 1977.

[85] P. Gerlee, "The model muddle: in search of tumor growth laws.," *Cancer research*, vol. 73, pp. 2407–11, apr 2013.

[86] R. B. Martin, M. E. Fisher, R. F. Minchin, and K. L. Teo, "Low-intensity combination chemotherapy maximizes host survival time for tumors containing drug-resistant cells," *Mathematical Biosciences*, vol. 110, pp. 221–252, jul 1992.

[87] H. B. Mistry, "Evolutionary Based Adaptive Dosing Algorithms: Beware the Cost of Cumulative Risk," *bioRxiv*, 2020.

[88] T. Kuosmanen, J. Cairns, R. Noble, N. Beerenwinkel, T. Mononen, and V. Mustonen, "Drug-induced resistance evolution necessitates less aggressive treatment," *PLoS computational biology*, vol. 17, p. e1009418, sep 2021.

[89] J. H. Goldie and A. J. Coldman, "A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate," *Cancer treatment reports*, vol. 63, pp. 1727–1733, nov-dec 1979.

[90] S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J. M. L. Ebos, L. Hlatky, and P. Hahnfeldt, "Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth," *PLoS Computational Biology*, vol. 10, p. e1003800, aug 2014.

[91] Y. Viossat and R. Noble, "The logic of containing tumors," *bioRxiv*, 2020.

[92] M. Ankathatti Munegowda, Y. Deng, S. J. Mulligan, and J. Xiang, "Th17 and th17-stimulated cd8+ t cells play a distinct role in th17-induced preventive and therapeutic antitumor immunity," *Cancer Immunology, Immunotherapy*, vol. 60, pp. 1473–1484, 2011.

[93] J. L. Markman and S. L. Shiao, "Impact of the immune system and immunotherapy in colorectal cancer," *Journal of Gastrointestinal Oncology*, vol. 6, no. 2, 2014.

[94] A. d'Onofrio, "A general framework for modeling tumor-immune system competition and immunotherapy: Mathematical analysis and biomedical inferences," *Physica D: Nonlinear Phenomena*, vol. 208, no. 3-4, pp. 220–235, 2005.

[95] R. Eftimie, J. L. Bramson, and D. J. Earn, "Interactions between the immune system and cancer: a brief review of non-spatial mathematical models," *Bulletin of mathematical biology*, vol. 73, pp. 2–32, 2011.

[96] K. P. Wilkie, "A review of mathematical models of cancer–immune interactions in the context of tumor dormancy," *Systems biology of tumor dormancy*, pp. 201–234, 2013.

[97] M. Itik and S. P. Banks, "Chaos in a three-dimensional cancer model," *International Journal of Bifurcation and Chaos*, vol. 20, no. 01, pp. 71–79, 2010.

[98] P. Bi and S. Ruan, "Bifurcations in delay differential equations and applications to tumor and immune system interaction models," *SIAM Journal on Applied Dynamical Systems*, vol. 12, no. 4, pp. 1847–1888, 2013.

[99] M. Ricard and S. Mischler, "Turing instabilities at hopf bifurcation," *Journal of Nonlinear Science*, vol. 19, pp. 467–496, 10 2009.

[100] N. M. Krylov and N. N. Bogoliubov, *Introduction to non-linear mechanics.* No. 11 in Annals of Mathematics Studies, Princeton university press, 1950.

[101] W. Abid, R. Yafia, M. Aziz Alaoui, H. Bouhafa, and A. Abichou, "Instability and pattern formation in three-species food chain model via holling type ii functional response on a circular domain," *International Journal of Bifurcation and Chaos*, vol. 25, no. 06, p. 1550092, 2015.

[102] W. Abid, R. Yafia, M. Aziz-Alaoui, and A. Aghriche, "Turing instability and hopf bifurcation in a modified leslie–gower predator–prey model with cross-diffusion," *International Journal of Bifurcation and Chaos*, vol. 28, no. 07, p. 1850089, 2018.

[103] J. Sarria-Gonzalez, I. Sgura, and M. Ricard, "Bifurcations in twinkling patterns for the lengyel–epstein reaction–diffusion model," *International Journal of Bifurcation and Chaos*, vol. 31, no. 11, p. 2150164, 2021.

[104] M. Chen, R. Wu, B. Liu, and L. Chen, "Turing-turing and turing-hopf bifurcations in a general diffusive brusselator model," *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, vol. n/a, no. n/a, p. e201900111, 2023.

[105] M. Chen and T. Wang, "Qualitative analysis and hopf bifurcation of a generalized lengyel–epstein model," *Journal of Mathematical Chemistry*, vol. 61, no. 1, pp. 166–192, 2023.

[106] X. Fu, Y. Zhao, J. I. Lopez, A. Rowan, L. Au, A. Fendler, S. Hazell, H. Xu, S. Horswell, S. T. Shepherd, *et al.*, "Spatial patterns of tumour growth impact clonal diversification in a computational model and the tracerx renal study," *Nature ecology & evolution*, vol. 6, no. 1, pp. 88–102, 2022.

[107] C. Carmona-Fontaine, V. Bucci, L. Akkari, M. Deforet, J. A. Joyce, and J. B. Xavier, "Emergence of spatial structure in the tumor microenvironment due to the warburg effect," *Proceedings of the National Academy of Sciences*, vol. 110, no. 48, pp. 19402–19407, 2013.

[108] Y. Yuan, "Spatial heterogeneity in the tumor microenvironment," *Cold Spring Harbor perspectives in medicine*, vol. 6, no. 8, 2016.

[109] C. H. Lineweaver and P. C. W. Davies, "Comparison of the atavistic model of cancer to somatic mutation theory: Phylostratigraphic analyses support the atavistic model," in *The Physics of Cancer. Research Advances* (B. S. Gerstman, ed.), ch. 12, pp. 243–261, Singapore: World Scientific, 2020.

[110] W. E. Muller, "Review: How was metazoan threshold crossed? The hypothetical Urmetazoa," *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, vol. 129, no. 2, pp. 433–460, 2001.

[111] R. H. Chisholm, T. Lorenzi, A. Lorz, A. K. Larsen, L. N. d. Almeida, A. Escargueil, and J. Clairambault, "Emergence of drug tolerance in cancer cell populations: An evolutionary outcome of selection, nongenetic instability, and stress-induced adaptation," *Cancer Research*, vol. 75, no. 6, pp. 930–939, 2015.

[112] D. Cohen, "Optimizing reproduction in a randomly varying environment," *Journal of Theoretical Biology*, vol. 12, no. 1, pp. 119 – 129, 1966.

[113] S. Huang, "Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells," *Cancer metastasis reviews*, vol. 32, p. 423—448, December 2013.

[114] K. Matuła, Ł. Richter, M. Janczuk-Richter, W. Nogala, M. Grzeszkowiak, B. Peplińska, S. Jurga, E. Wyroba, S. Suski, H. Bilski, *et al.*, "Phenotypic plasticity of Escherichia coli upon exposure to physical stress induced by ZnO nanorods," *Scientific reports*, vol. 9, 2019.

[115] J. A. Carrillo, F. Filbet, and M. Schmidtchen, "Convergence of a finite volume scheme for a system of interacting species with cross-diffusion," *Numerische Mathematik*, vol. 145, no. 3, 2020.

[116] R. Eymard, T. Gallouët, and R. Herbin, "Finite volume methods," *Handbook of Numerical Analysis*, vol. 7, pp. 713 – 1018, 2000.

[117] A. Moussa, "Some variants of the classical Aubin–Lions lemma," *Journal of Evolution Equations*, vol. 16, no. 1, pp. 65–93, 2016.

[118] R. Eymard, T. Gallouët, and R. Herbin, "A cell-centred finite volume approximation for second order partial derivative operators with full matrix on unstructured meshes in any space dimension," *IMA Journal of Numerical Analysis*, vol. 26, pp. 326–353, 2006.

[119] A. M. Nedelcu and R. E. Michod, “Stress responses co-opted for specialized cell types during the early evolution of multicellularity,” *BioEssays*, vol. 42, no. 5, p. 2000029, 2020.

[120] G. P. Wagner, E. M. Erkenbrack, and A. C. Love, “Stress-induced evolutionary innovation: A mechanism for the origin of cell types,” *BioEssays : news and reviews in molecular, cellular and developmental biology*, vol. 41, no. 4, 2019.

[121] F. Jacob, “Evolution and tinkering,” *Science*, vol. 196, no. 4295, pp. 1161–1166, 1977.

[122] S. Gould, *Ontogeny and Phylogeny*. Harvard University Press, 1977.

[123] E. Haeckel, *Generelle morphologie der organismen*. G. Reimer Verlag, 1866.

[124] L. Wolpert and E. Szathmáry, “Evolution and the egg,” *Nature*, vol. 420, 2002.

[125] C. A. Aktipis, A. M. Boddy, G. Jansen, U. Hibner, M. E. Hochberg, C. C. Maley, and G. S. Wilkinson, “Cancer across the tree of life: cooperation and cheating in multicellularity,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, no. 1673, p. 20140219, 2015.

[126] N. A. M., “The evolution of multicellularity and cancer: views and paradigms,” *Biochemical Society transactions*, vol. 48, no. 4, pp. 1505–1518, 2020.

[127] E. Davidson, K. Peterson, and R. Cameron, “Origin of bilaterian body plans: evolution of developmental regulatory mechanisms.,” *Science (New York, N.Y.)*, vol. 270, no. 0036-8075, p. 1319–1325, 1995.

[128] W. Müller, M. Wiens, T. Adell, V. Gamulin, H. C. Schröder, and I. M. Müller, “Bauplan of urmetazoa: basis for genetic complexity of metazoa.,” *International Review of Cytology*, vol. 235, pp. 53–92, 2004.

[129] M. Bertolaso, *Philosophy of Cancer. A Dynamic and Relational View*. Springer Science+Business Media B.V., 2016.

[130] A. S. Cleary, T. L. Leonard, and E. J. Gunther, “Tumour cell heterogeneity maintained by cooperating subclones in wnt-driven mammary cancers,” *Nature*, vol. 508, no. 7494, pp. 113–117, 2014.

[131] K. Polyak and A. Marusyk, “Clonal cooperation,” *Nature*, vol. 508, no. 7494, pp. 52–53, 2014.

[132] D. P. Tabassum and K. Polyak, “Tumorigenesis: it takes a village,” *Nature reviews. Cancer*, vol. 15, no. 8, pp. 473–483, 2015.

[133] L. C.H., D. P.C., and V. M.D., “Targeting cancer’s weaknesses (not its strengths): Therapeutic strategies suggested by the atavistic model,” *Bioessays*, vol. 36, no. 9, pp. 827–35, 2014.

[134] C. H. Lineweaver and P. C. Davies, *Comparison of the Atavistic Model of Cancer to Somatic Mutation Theory: Phylostratigraphic Analyses Support the Atavistic Model*, ch. Chapter 12, pp. 243–261. World Scientific, 2020.

[135] C. H. Lineweaver, K. J. Bussey, A. C. Blackburn, and P. C. W. Davies, “Cancer progression as a sequence of atavistic reversions,” *BioEssays : news and reviews in molecular, cellular and developmental biology*, vol. 43, no. 7, 2021.

[136] V. M., “Cancer: a de-repression of a default survival program common to all cells?: a life-history perspective on the nature of cancer,” *BioEssays : news and reviews in molecular, cellular and developmental biology*, vol. 34, no. 1, pp. 72–82, 2012.

[137] I. L., “Tumour progression: random mutations or an integrated survival response to cellular stress conserved from unicellular organisms?,” *Journal of theoretical biology*, vol. 178, no. 4, pp. 375—-380, 1996.

[138] T. Boveri, “Concerning the origin of malignant tumours by theodor boveri. translated and annotated by henry harris.,” *Journal of cell science*, vol. 121 Suppl 1, pp. 1–84, 2008.

[139] A. S. Trigos, R. B. Pearson, A. T. Papenfuss, and D. L. Goode, “Altered interactions between unicellular and multicellular genes drive hallmarks of transformation in a diverse range of solid tumors,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 24, p. 6406–6411, 2017.

[140] A. S. Trigos, R. B. Pearson, A. T. Papenfuss, and D. L. Goode, “How the evolution of multicellularity set the stage for cancer,” *British Journal of Cancer*, vol. 118, no. 2, pp. 145–152, 2018.

[141] A. S. Trigos, R. B. Pearson, A. T. Papenfuss, and D. L. Goode, “Somatic mutations in early metazoan genes disrupt regulatory links between unicellular and multicellular genes in cancer,” *eLife*, vol. 8, p. e40947, feb 2019.

[142] A. S. Trigos, F. Bongiovanni, M. Zethoven, R. Tothill, R. Pearson, T. Papenfuss, and D. L. Goode, “Disruption of metazoan gene regulatory networks in cancer alters the balance of co-expression between genes of unicellular and multicellular origins,” *bioRxiv*, 2023.

[143] P. Bayle, “Pensées diverses sur la comète,” *CH VII, Paris*, 1682.

[144] A. Wegener, “Die entstehung der kontinente,” *Geol Rundsch*, vol. 3, p. 276–292, 1912.

[145] T. *et al.*. Pradeu, “Reuniting philosophy and science to advance cancer research,” *Biological reviews of the Cambridge Philosophical Society*, vol. 98, no. 5, pp. 1668–1686, 2023.

[146] T. Pradeu, *Philosophy of Immunology.* Elements in the Philosophy of Biology, Cambridge University Press, 2020.

[147] P. Hirsch, Y. Zhang, R. Tang, V. Joulin, H. Boutroux, E. Pronier, H. Moatti, P. Flandrin, C. Marzac, D. Bories, F. Fava, H. Mokrani, A. Betems, F. Lorre, R. Favier, F. Féger, M. Mohty, L. Douay, O. Legrand, C. Bilhou-Nabera, F. Louache, and F. Delhommeau, “Genetic hierarchy and temporal variegation in the clonal history of acute myeloid leukaemia,” *Nature Communications*, vol. 7, no. 1, p. 12475, 2016.

[148] V. Balter, J. Braga, P. Télouk, and J. Thackeray, “Evidence for dietary change but not landscape use in south african early hominins,” *Nature*, vol. 489, no. 7417, pp. 558–560, 2012.

[149] J. Clairambault, “Plasticity in cancer cell populations: Biology, mathematics and philosophy of cancer,” in *Mathematical and Computational Oncology* (G. Bebis, M. Alekseyev, H. Cho, J. Gevertz, and M. Rodriguez Martinez, eds.), (Cham), pp. 3–9, Springer International Publishing, 2020.

[150] J. Clairambault, "Mathematical modelling of cancer growth and drug treatments: Taking into account cell population heterogeneity and plasticity," in *2023 European Control Conference (ECC)*, pp. 1–6, 2023.

[151] K. McCullough, W. Coleman, S. Ricketts, W. J.W., S. G.J., and J. Grisham, "Plasticity of the neoplastic phenotype in vivo is regulated by epigenetic factors," *Proc Natl Acad Sci U S A*, vol. 95, no. 26, pp. 15333–8, 1998.

[152] F. Weismann, "Das keimplasma eine theorie der vererbung," *Nature*, vol. 47, p. 265–266, 1893.

[153] J. Maynard Smith and E. Szathmary, *The Major Transitions in Evolution.* Oxford University Press, 10 1997.

[154] T. Domazet-Loso and D. Tautz, "An ancient evolutionary origin of genes associated with human genetic diseases," *Molecular biology and evolution*, vol. 25, no. 12, pp. 2699–2707, 2008.

[155] T. Domazet-Loso and D. Tautz, "Phylostratigraphic tracking of cancer genes suggests a link to the emergence of multicellularity in metazoa," *BMC Biology*, vol. 8, no. 1, p. 66, 2010.

[156] H. Hatzikirou, D. Basanta, M. Simon, K. Schaller, and A. Deutsch, "'go or grow': The key to the emergence of invasion in tumour progression?," *Mathematical medicine and biology: a journal of the IMA*, vol. 29, pp. 49–65, Mar. 2012.

[157] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[158] Y. Murase, "Evolution of direct reciprocity in group-structured populations," *Scientific Reports*, vol. 12, no. 1, p. 18645, 2022.

[159] K. S. G. and N. A. M., "The genetic basis for the evolution of soma: mechanistic evidence for the co-option of a stress-induced gene into a developmental master regulator," *Proc. R. Soc. B.*, vol. 287, 202O.

[160] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *Cell*, vol. 126, no. 4, pp. 636–676, 2006.

[161] U. Dieckmann and R. Law, "The dynamical theory of coevolution: a derivation from stochastic ecological processes," *Journal of mathematical biology*, vol. 34, no. 5, pp. 579–612, 1996.

[162] S. A. Geritz, E. Kisdi, J. A. Metz, *et al.*, "Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree," *Evolutionary ecology*, vol. 12, no. 1, pp. 35–57, 1998.

[163] J. A. Metz, S. A. Geritz, G. Meszéna, F. J. Jacobs, and J. S. Van Heerwaarden, "Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction," *WP-95-099*, 1995.

[164] R. Bürger, *The mathematical theory of selection, recombination, and mutation.* John Wiley & Sons, 2000.

[165] G. Barles, S. Mirrahimi, and B. Perthame, "Concentration in lotka-volterra parabolic or integral equations: a general convergence result," *Methods and Applications of Analysis*, vol. 16, no. 3, pp. 321–340, 2009.

[166] J. Coville, "Convergence to equilibrium for positive solutions of some mutation-selection model," *arXiv preprint arXiv:1308.6471*, 2013.

[167] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame, "The dynamics of adaptation: an illuminating example and a hamilton–jacobi approach," *Theoretical population biology*, vol. 67, no. 4, pp. 257–271, 2005.

[168] A. Lorz, T. Lorenzi, M. E. Hochberg, J. Clairambault, and B. Perthame, "Populational adaptive evolution, chemotherapeutic resistance and multiple anti-cancer therapies," *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, vol. 47, no. 2, pp. 377–399, 2013.

[169] B. Perthame and G. Barles, "Dirac concentrations in lotka-volterra parabolic pdes," *Indiana University Mathematics Journal*, pp. 3275–3301, 2008.

[170] P.-E. Jabin and G. Raoul, "On selection dynamics for competitive interactions," *Journal of mathematical biology*, vol. 63, no. 3, pp. 493–517, 2011.

[171] T. Lorenzi and C. Pouchol, "Asymptotic analysis of selection-mutation models in the presence of multiple fitness peaks," *Nonlinearity*, vol. 33, no. 11, p. 5791, 2020.

[172] À. Calsina, S. Cuadrado, L. Desvillettes, and G. Raoul, "Asymptotics of steady states of a selection–mutation equation for small mutation rate," *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, vol. 143, no. 6, pp. 1123–1146, 2013.

[173] O. Bonnefon, J. Coville, and G. Legendre, "Concentration phenomenon in some non-local equation," *arXiv preprint arXiv:1510.01971*, 2015.

[174] A. Friedman, C.-Y. Kao, and C.-W. Shih, "Asymptotic phases in a cell differentiation model," *Journal of Differential Equations*, vol. 247, no. 3, pp. 736–769, 2009.

[175] A. Yates, R. Callard, and J. Stark, "Combining cytokine signalling with t-bet and gata-3 regulation in th1 and th2 differentiation: a model for cellular decision-making," *Journal of theoretical biology*, vol. 231, no. 2, pp. 181–196, 2004.

[176] R. J. DiPerna and P.-L. Lions, "Ordinary differential equations, transport theory and sobolev spaces," *Inventiones mathematicae*, vol. 98, no. 3, pp. 511–547, 1989.

[177] M. Gyllenberg and G. Meszéna, "On the impossibility of coexistence of infinitely many strategies," *Journal of mathematical biology*, vol. 50, no. 2, pp. 133–160, 2005.

[178] D. B. Cooney and Y. Mori, "Long-time behavior of a pde replicator equation for multilevel selection in group-structured populations," *Journal of Mathematical Biology*, vol. 85, no. 2, pp. 1–67, 2022.

[179] L. Barbarroux, P. Michel, M. Adimy, and F. Crauste, "Multi-scale modeling of the cd8 immune response," in *AIP Conference Proceedings*, vol. 1738, p. 320002, AIP Publishing LLC, 2016.

[180] A. Chertock, "A practical guide to deterministic particle methods," in *Handbook of numerical analysis*, vol. 18, pp. 177–202, Elsevier, 2017.

[181] P. Hartman, *Ordinary differential equations*. SIAM, 2002.

[182] A. Cohen and B. Perthame, "Optimal approximations of transport equations by particle and pseudoparticle methods," *SIAM Journal on Mathematical Analysis*, vol. 32, no. 3, pp. 616–636, 2000.

[183] T. Y. Hou, "Convergence of a variable blob vortex method for the euler and navier–stokes equations," *SIAM Journal on Numerical Analysis*, vol. 27, no. 6, pp. 1387–1404, 1990.

[184] S. Mas-Gallic and P. Raviart, "A particle method for first-order symmetric systems," *Numerische Mathematik*, vol. 51, no. 3, pp. 323–352, 1987.

[185] T. Hillen and K. J. Painter, "A user's guide to pde models for chemotaxis," *Journal of mathematical biology*, vol. 58, no. 1-2, pp. 183–217, 2009.

[186] M. J. Tindall, P. K. Maini, S. L. Porter, and J. P. Armitage, "Overview of mathematical approaches used to model bacterial chemotaxis ii: bacterial populations," *Bulletin of mathematical biology*, vol. 70, pp. 1570–1607, 2008.

[187] Y. Naito, T. Suzuki, and K. Yoshida, "Self-similar solutions to a parabolic system modeling chemotaxis," *Journal of Differential Equations*, vol. 184, no. 2, pp. 386–421, 2002.

[188] L. Corrias, M. Escobedo, and J. Matos, "Existence, uniqueness and asymptotic behavior of the solutions to the fully parabolic keller–segel system in the plane," *Journal of Differential Equations*, vol. 257, no. 6, pp. 1840–1878, 2014.

[189] S. Mischler and C. Mouhot, "Stability, convergence to self-similarity and elastic limit for the boltzmann equation for inelastic hard spheres," *Communications in Mathematical Physics*, vol. 288, no. 2, pp. 431–502, 2009.

[190] I. Tristani, "Boltzmann equation for granular media with thermal force in a weakly inhomogeneous setting," *Journal of Functional Analysis*, vol. 270, no. 5, pp. 1922–1970, 2016.

[191] T. Kato, *Perturbation theory for linear operators*, vol. 132. Springer Science & Business Media, 2013.

[192] S. Mischler and J. Scher, "Spectral analysis of semigroups and growth-fragmentation equations," in *Annales de l'IHP Analyse non linéaire*, vol. 33, pp. 849–898, 2016.

[193] L. Gearhart, "Spectral theory for contraction semigroups on hilbert space," *Transactions of the American Mathematical Society*, vol. 236, pp. 385–394, 1978.

[194] J. Prüss, "On the spectrum of c^0-semigroups," *Transactions of the American Mathematical Society*, vol. 284, no. 2, pp. 847–857, 1984.

[195] W. Arendt, A. Grabosch, G. Greiner, U. Moustakas, R. Nagel, U. Schlotterbeck, U. Groh, H. P. Lotz, and F. Neubrander, *One-parameter semigroups of positive operators*, vol. 1184. Springer, 1986.

RÉSUMÉ

Cette thèse est consacrée à l'étude de plusieurs problèmes issus de la modélisation mathématique des tumeurs. Plus spécifiquement, l'intérêt principal est orienté vers les interactions ayant lieu au sein de la tumeur et avec son environnement. Néanmoins, certains des modèles et méthodes présentés au coeur de la thèse ont une portée bien plus générale que l'étude du cancer. Les principaux résultats sont divisés en cinq chapitres. Dans le premier chapitre, par une nouvelle analyse mathématique comparant la taille des tumeurs entre traitements non pas en fonction du temps, mais en fonction de la taille de la population résistante, nous établissons une comparaison entre les résultats de différentes stratégies de traitement appliquées à une tumeur composée de deux sous-populations, une de cellules sensibles et une autre de cellules résistantes. Dans le deuxième chapitre, nous dérivons l'expression asymptotique d'un cycle limite apparaissant dans un modèle d'interaction tumeur-système immunitaire. Le troisième chapitre est consacré à la modélisation du bet-hedging, une stratégie évolutive d'intérêt pour la théorie atavique du cancer. L'existence et le caractère unique de la solution du modèle sont prouvés et deux phénomènes d'intérêt biologique sont mis en évidence par des simulations. Le chapitre quatre est un complément au troisième chapitre. On y développe une discussion philosophique sur la théorie atavique du cancer et on esquisse deux modèles différents pour l'émergence de la coopération. Le chapitre cinq concerne l'étude d'une méthode particulaire pour un modèle d'advection-réaction-diffusion non local d'une grande importance dans le domaine de les dynamiques adaptatives. La conservation du comportement asymptotique est analysée pour le schéma numérique proposé. Les chapitres six et sept sont consacrés à l'étude du système de Keller-Segel parabolique-parabolique où nous donnons respectivement quelques estimations de la solution et déterminons le comportement asymptotique pour le cas non radial.

MOTS CLÉS

Populations structurées; modélisation du cancer; confinement des tumeurs; méthodes particulaires; théorie atavique; système de Keller-Segel parabolique-parabolique.

ABSTRACT

This thesis is devoted to the study of several problems arising from the mathematical modelling of tumours. More specifically, the main interest is oriented towards the interactions taking place within the tumour and with its environment. Nevertheless, some of the models and methods presented at the core of the thesis have a much more general scope than the study of cancer. The main results are divided in five chapters. In the first chapter, by a novel mathematical analysis comparing tumor sizes across treatments not as a function of time, but as a function of the resistant population size, we establish a comparison between the outcomes of different treatment strategies applied to a tumour composed of two sub-populations, one of sensitive cells and another one of resistant cells. In the second chapter, we derive the asymptotic expression of a limit cycle arising in a tumour-immune system interaction model. The third chapter is devoted to the modeling of bet-hedging, an evolutionary strategy of interest for the atavistic theory of cancer. The existence and uniqueness of solution for the model is proved and two phenomena of biological interest are evidenced through simulations. Chapter four is a complement for the third chapter. On it, a philosophical discussion about the atavistic theory of cancer is developed and two different models for the emergence of cooperation are sketched. Chapter five is concerned with the study of a particle method for non-local advection-reaction-diffusion model of great importance in the area of adaptive dynamics. The conservation of asymptotic behaviour is analyzed for the proposed numerical scheme. Chapters six and seven are devoted to the study of the fully parabolic Keller-Segel system where we give some estimates over the solution and determine the asymptotic behaviour for the non-radial case, respectively.

KEYWORDS

Structured populations; cancer modelling; tumour containment; particle methods; atavistic theory; fully parabolic Keller-Segel system.