Università degli studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Magistrale in Matematica

Mathematical modelling of cancer evolution

Relatore:
Prof. Fabio Marcuzzi
Correlatore:
Prof. Andrei Korobeinikov

Laureando:
Stefano Pedarra, 1150708

20 Luglio 2018

Anno Accademico 2017-2018

When you think you are so over it,
shout out and fight till the very end.

Contents

Introduction

The term *cancer* refers to a group of diseases which can affect almost any tissue and organ and are characterised by the uncontrolled growth of abnormal cells, whose cell cycle is much faster than the one of a healthy cell. The carcinogenesis, i.e. the production of a cancer, is due to mutations that occur physiologically during the DNA replication process. Such mutations are usually repaired by some enzymes; however, their functions can be inhibited for genetical reasons or affected by some environmental factor, such as exposition to X rays, pollution and some chemical substances. In this case, it may happen that the cell does not repair mutations and the accumulation of a certain amount of them can bring the healthy cell to be a cancer cell.

From a mathematical point of view, many aspects can be studied. For instance, one can model how cancer developes, how a treatment can affect its evolution or how angiogenesis vascularises it. In this thesis, we focus on tumour initial developement, modelling the oncogenesis through a system of ordinary differential equations (ODEs), taking into account the possibility that a mutation occurs and the competition among healthy cells and cancer cells for the available resource, such as physical space to develop, nutrients and oxygen. The model is later analysed: it is done analitically when we are differentiating between healthy cells and cancer ones; otherwise, if we want to consider n different genotypes, the study is numerical.

The thesis is structured as following: *Chapter 1* contains a short review of classical models of mathematical biology, namely Malthusian and logistic models and a competition model, all based on ODEs. In *Chapter 2*, starting from this last model, we model the cancer evolution adding the possibility of mutation and, then, the so-obtained model is non-dimensionalised to make this analysis simpler. *Chapter 3* contains the analysis of the model for two different populations of cells, namely healthy and cancer cells; the number of fixed points is studied depending on the values of the physical parameters of the model and some sufficient and sometimes also necessary conditions for their stability are obtained; all these conditions are interpreted biologically. Finally, in *Chapter 4* the outcomes of some numerical simulations are presented and compared, in the case of a big variety of different genotypes; some biological interpretation is given to the numerical results.

Chapter 1

Background on population dynamics models

The mathematical models constructed to study population dynamics are called population models. If a population model takes into account competition factors among populations, it is called competition model. Both kinds of model let us understand better the quantitative or, at least, qualitative consequences of complex interactions among the populations and the environment. In this chapter, we will present some classical population and competition models.

1.1 Malthusian growth

The first and the simplest population model is the Malthusian growth model. It was proposed by the British economist Thomas Robert Malthus (1766 - 1834) in *An essay on the principle of population*, published in 1798. Malthus thought that a population grows exponentially with a constant rate. To model this situation, it is assumed that the rate of increase of the population is proportional to the size of the population. Denoting by $P(t)$ the size of the population at time t and by a the population growth rate, the differential equation which describes the model is

$$\frac{dP(t)}{dt} = aP(t)\,. \tag{1.1}$$

Consequently, if the inital condition is $P(0) = P_0$, the evolution of the size of the population is described by the function

$$P(t) = P_0 e^{at}\,.$$

However, this model is as simple as unrealistic. In the very short term, it gives a good prediction, but in the long term the approximation is not accurate at all. For this reason,

it is necessary to improve it.

1.2 Logistic growth model

The logistic growth model represents a refinement of the Malthusian one. It takes into account the limited carrying capacity of the environment where the population lives, due, for instance, to the limited resources. The simplest form of this model, presented in [7] is

$$\frac{dP(t)}{dt} = aP(t)\left(1 - \frac{P(t)}{K}\right), \tag{1.2}$$

where parameters a and K denote, respectively, the growth rate and the carrying capacity of the environment. We can notice that, if $P(t) < K$, then $\dot{P}(t) > 0$, that is the population increases; conversely, if $P(t) > K$, then $\dot{P}(t) < 0$, that is the population decreases. Otherwise, if $P(t) = K$, then $\dot{P}(t) = 0$, which means that $P(t) = K$ is an equilibrium state of (1.2), as is $P(t) = 0$.

Apart from this remark, we see that the differential equation (1.2) is separable. Hence, it is possible to integrate it, obtaining the solution

$$P(t) = \frac{KP_0}{P_0 + (K - P_0)e^{-rt}} \xrightarrow{t\to\infty} K\,,$$

if $P(0) = P_0$.

1.3 Competition model

Until now we have considered models describing the growth of a single population. However, in a real-life context, populations interact among them and compete for resources, e.g. food or territory. Let us suppose that there are two populations of sizes, respectively, $A(t)$ and $C(t)$. In the absence of the other population, each of these governed by a logistic growth. If there are individuals of the second population, the death rate of the first population is proportional to both $A(t)$ and $C(t)$ and vice versa. For this reason, the model, proposed, for instance, in [7] can be written as following:

$$\begin{aligned} \frac{dA}{dt} &= pA\left(1 - \frac{A}{K_0} - \frac{b_{01}C}{K_0}\right), \\ \frac{dC}{dt} &= rC\left(1 - \frac{C}{K_1} - \frac{b_{10}A}{K_1}\right), \end{aligned} \tag{1.3}$$

where $p,\ r,\ b_{01},\ b_{10},\ K_0,\ K_1$ are all positive constants. In particular, p and r are the growth rates, K_0 and K_1 are the carrying capacities and b_{01} and b_{10} measure the competitive effects.

In order to simplify the analysis of the model, let us non-dimensionalize it by setting

$$x = \frac{A}{K_0}\,,\ y = \frac{C}{K_1}\,,\ \tau = pt\,,\ \rho = \frac{r}{p}\,,\ a_{01} = b_{01}\frac{K_1}{K_0}\,,\ a_{10} = b_{10}\frac{K_0}{K_1}\,.$$

Then, the system becomes

$$\begin{aligned} \frac{dx}{d\tau} &= \dot{x} = x(1 - x - a_{01}y)\,, \\ \frac{dy}{d\tau} &= \dot{y} = \rho y(1 - y - a_{10}x)\,. \end{aligned} \tag{1.4}$$

The steady states of the system (1.4) are the solutions of $\dot{x} = \dot{y} = 0$. Namely, they are:

- $P_{00} = (0, 0)$, which occurs when both populations do not exist;
- $P_{10} = (1, 0)$, when only the population A exists;
- $P_{01} = (0, 1)$, when only the population C exists;
- $P_* = (x_*, y_*) = \left(\frac{1-a_{01}}{1-a_{01}a_{10}}, \frac{1-a_{10}}{1-a_{01}a_{10}}\right)$, when the populations coexist.

Since both variables represent population sizes, the last steady state exists if and only if $1 - a_{01}a_{10} \neq 0$ and $x_*, y_* \geq 0$. Hence, it existes if and only if either $1 - a_{01}$, $1 - a_{10}$ and $1 - a_{01}a_{10}$ are positive at the same time or they are negative at the same time.

Model analysis

We can study the stability of the steady states by linearizing the system near the four steady states and applying the following theorem, presented, for instance, in [10]:

Theorem 1.3.1 (Grobman-Hartman)**.** *Let $F \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^n)$ be a vector field. Let us consider the dynamical system associated to the equation*

$$\dot{x} = F(x)\,.$$

Let x^* be an hyperbolic fixed point, i.e. $F(x^*) = 0$ and $J(x^*)$ has no eigenvalue with real part equal to zero, where $J(x^*)$ is the Jacobian matrix of F in x^*. Then, the local phase portrait near x^* is topologically equivalent to the phase portrait of the linearization, that is there exists a neighborhood N of x^* and a homeomorphism $h : N \to \mathbb{R}^n$ such that $h(x^*) = 0$ and such that in the neighbourhood N the flow of $\dot{x} = F(x)$ is topologically conjugate by the continuous map $y = h(x)$ to the flow of its linearization $\dot{y} = J(x^*)y$.

In particular, it means that, if (x^*, y^*) is a fixed point of the dynamical system

$$\dot{x} = P(x, y)\,,$$
$$\dot{y} = Q(x, y)\,,$$

then:

- if (x^*, y^*) is a saddle point for the linearized system, then it is a local saddle point of the system;
- if (x^*, y^*) is a stable/unstable node for the linearized system, then it is a local stable/unstable node of the system;
- if (x^*, y^*) is a stable/unstable focus for the linearized system, then it is a local stable/unstable focus of the system.

The Jacobian matrix associated to the system (1.4) is

$$J(x, y) = \begin{pmatrix} \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial y} \\ \frac{\partial \dot{y}}{\partial x} & \frac{\partial \dot{y}}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 - 2x - a_{01}y & -a_{01}x \\ -\rho a_{10}y & \rho(1 - 2y - a_{10}x) \end{pmatrix}.$$

Let us now evaluate at in the four steady states to analyze their stability.

(i) Trivial steady state P_{00}.

The Jacobian matrix in the steady state P_{00} is

$$J(0, 0) = J_{00} = \begin{pmatrix} 1 & 0 \\ 0 & \rho \end{pmatrix}.$$

Its eigenvalues are evidently $\lambda_1 = 1$ and $\lambda_2 = \rho$, which are both positive. Therefore, the fixed point P_{00} is always a repulsore node and, hence, unstable.

(ii) One population steady states P_{10} and P_{01}.

As regards the steady state $P_{10} = (1, 0)$, the Jacobian matrix evalutated in this point is

$$J(1, 0) = J_{10} = \begin{pmatrix} -1 & -a_{01} \\ 0 & \rho(1 - a_{10}) \end{pmatrix},$$

whose eigenvalues are are $\lambda_1 = -1 < 0$ and $\lambda_2 = \rho(1 - a_{10})$. Then, since $\rho > 0$, the stability of P_{10} depends only on the value of a_{10}:

- if $a_{10} \in [0, 1[$, the eigenvalues have opposite signs and, hence, P_{10} is a saddle point, which is unstable;
- if $a_{10} \in]1, +\infty$, both eigenvalues are negative and, hence, P_{10} is an attractive node, which is stable.

In an analogous way, it can be observed that the stability of P_{01} depends only on the value of a_{01}. In particular:

- if $a_{01} \in [0, 1[$, the eigenvalues have opposite signs and, hence, P_{01} is a saddle point, which is unstable;
- if $a_{01} \in]1, +\infty$, both eigenvalues are negative and, hence, P_{01} is an attractive node, which is stable.

(iii) Coexistence steady state P_*.

For the last steady state P_*, the Jacobian matrix is

$$J\left(\frac{1-a_{01}}{1-a_{01}a_{10}}, \frac{1-a_{01}}{1-a_{01}a_{10}}\right) = J_* = \frac{1}{1-a_{01}a_{10}} \begin{pmatrix} -1+a_{01} & -a_{01}(1-a_{01}) \\ -\rho a_{10}(1-a_{10}) & \rho(-1+a_{10}) \end{pmatrix}.$$

To study the stability of the steady state, we compute the trace and the determinant of J_*, obtaining:

$$tr(J_*) = \frac{-(1-a_{01}) - \rho(1-a_{10})}{1-a_{01}a_{10}}, \qquad det(J_*) = \frac{\rho(1-a_{01})(1-a_{10})}{1-a_{01}a_{10}}.$$

Let us observe that neither $a_{01} < 1 < a_{10}$ nor $a_{10} < 1 < a_{01}$ is possible, as otherwise P_* would not belong to the first quadrant. Hence, there are only two cases to considerate:

- if $a_{01}, a_{10} < 1$, then $det(J_*) > 0$ and $tr(J_*) < 0$, which means that both eigenvalues of J_* are real and negative. In other words, P_* is an attractive node and it is stable;
- if $a_{01}, a_{10} > 1$, then $det(J_*) < 0$, which means that the eigenvalues of J_* are real and have opposite signs. In other words, P_* is a saddle point and it is unstable.

Cases of dynamics

First of all, let us notice that the nullclines of the system are the straight lines described by

$$x = 0\,, \qquad y = 0\,, \qquad x + a_{01}y = 1\,, \qquad a_{10}x + y = 1\,.$$

At this point, each possible case can be studied. As obtained before, $P_0 = (0, 0)$ is always an unstable equilibrium, independently of the value of the parameters.

1. $a_{01}\,,\ a_{10} < 1$

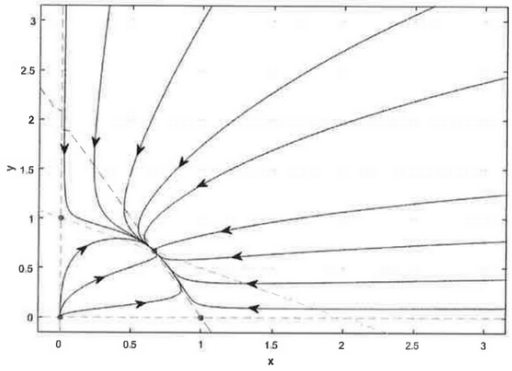


Figure 1.1: Phase plane of the system (1.4) for $a_{01}, a_{10} < 0$ [2].

As we can see in the figure, both P_{01} and P_{10} are saddle points, whilst P_* is an attractive node. Therefore, P_* is stable, although P_0, P_{01} and P_{10} are unstable. Since it is evident that P_* is a global attractor, it means that, in the long term, the two populations tend to coexist in a stable way.

2. $a_{01}\,,\ a_{10} > 1$

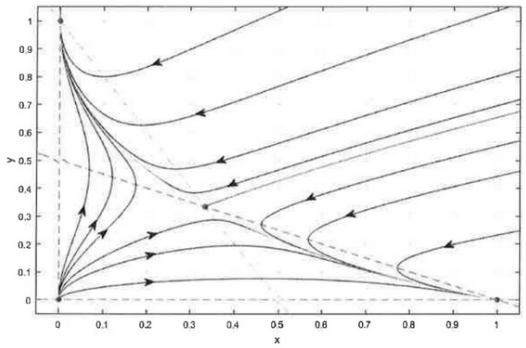


Figure 1.2: Phase plane of the system (1.4) for a_{01}, $a_{10} > 0$ [2].

In this case, both P_{01} and P_{10} are attractive nodes, whereas P_0 is a repulsor node and P_* is a saddle point. Therefore, P_0 and P_* are unstable, although P_{01} and P_{10} are stable. It is evident in the figure that just one population survive in the long term and it depends on the initial condition.

3. $a_{01} < 1\,,\ a_{10} > 1$

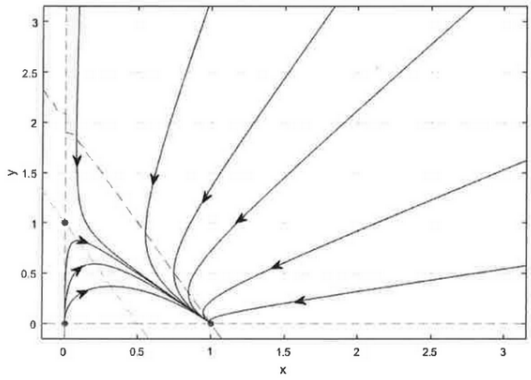


Figure 1.3: Phase plane of the system (1.4) for $a_{01} < 0\,, a_{10} < 0$ [2].

For these values of the parameters, there is no possibility of coexistence of the two populations. P_{10} is an attractive node, although P_0 is a repulsive node and P_{01} is a saddle point. In this case, it is evident that P_{10} is a global attractor, therefore the population x survives, while the population y extincts.

4. $a_{01} > 1\,,\ a_{10} < 1$

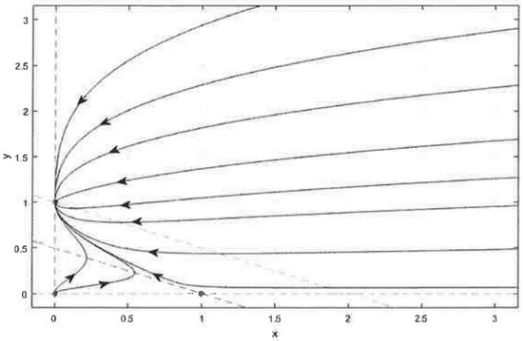


Figure 1.4: Phase plane of the system (1.4) for $a_{01} > 0\,,\ a_{10} < 0$ [2].

In this case, there is no possibility of coexistence of the two populations. P_{01} is an attractive node, although P_0 is a repulsive node and P_{10} is a saddle point. In this case, it is evident that P_{01} is a global attractor, therefore the population y survives, while the population x extincts.

5. $a_{01} = a_{10} = 1$

   In this case, the differential system is

$$\frac{dx}{d\tau} = x(1 - x - y)\,,$$
$$\frac{dy}{d\tau} = \rho y(1 - x - y)\,.$$

   and the nullclines are

$$x = 0\,,\ y = 0\,,\ x + y = 1\,.$$

   All the points of the last nullclines, which divides the positive quadrant in two region, are fixed points. Let us study their stability through the sign of partial derivatives in the two regions.

   Let (x_1, y_1) be a point of the positive quadrant such that $x_1 + y_1 < 1$. Then:

$$\frac{dx}{d\tau}|_{(x_1,y_1)} = x_1(1 - x_1 - y_1) = x_1(1 - (x_1 + y_1)) > 0\,,$$

$$\frac{dy}{d\tau}|_{(x_1,y_1)} = \rho y_1(1 - x_1 - y_1) = \rho y_1(1 - (x_1 + y_1)) > 0\,.$$

   Although, if (x_2, y_2) is a point of the positive quadrant such that $x_2 + y_2 > 1$, then:

$$\frac{dx}{d\tau}|_{(x_2,y_2)} = x_2(1 - x_2 - y_2) = x_2(1 - (x_2 + y_2)) < 0\,,$$

$$\frac{dy}{d\tau}|_{(x_1,y_1)} = \rho y_2(1 - x_2 - y_2) = \rho y_2(1 - (x_2 + y_2)) < 0\,.$$

   Due to the positivity of $\dot{x}$ and $\dot{y}$, trajectories of the bounded region $\mathcal{R}_1 = \{(x, y)\,|\,x > 0\,, y > 0\,, x +$ have positive slopes; whereas, trajectories of the region $\mathcal{R}_2 = \{(x, y)\,|\,x > 0\,, y > 0\,, x + y > 1\}$ have negative slopes, since $\dot{x}$, $\dot{y} < 0$. It means that each positive point of the nullcline

$$\mathcal{L}_2 = \{(x, y)\,|x + y = 1\}$$

   is an attractive point; as a consequence, the nullcline is an attractive line.

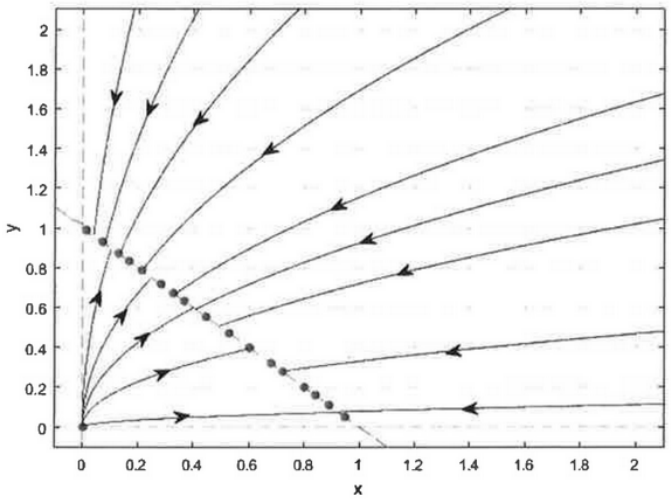


Figure 1.5: Phase plane of the system (1.4) for $a_{01} = a_{10} = 1$ [2].

Chapter 2

Model

In this chapter, a new model for cancer evolution is developed. After the construction of the model, it will be non-dimensionalized in order to obtain a more compact expression of the model and to make it easier to study.

2.1 Construction of the model

Let us assume that there are n cancer cell genotypes. We will denote the size of the population of cells with the i-th genotype at time t by $C_i(t)$ and the size of the normal cells population at the same time by $C_0(t)$. The phenomena that the model must consider are the following:

- cells reproduce and die;
- there are limited resources, hence the growth cannot be unlimited;
- there is a competitive interaction among cells with different genotypes for the limited resources;
- in the process of mitosis, with some probability p_{ij}, a cell of type i can produce a mutant daughter cell of type j, that goes to the j-th population.

The models presented in *Chapter 1* are not sufficient to describe such a system. However, they can be used as a basis in order to build a more appropriate model. Considering the competition model (1.3), one can observe that there is one factor missing: namely, the model disregards a possibility of mutation. To understand the behaviour of the model, let us write it for $n = 1$ (that is for a system "healthy cells-cancer cells") as following:

$$\begin{aligned}\dot{C}_0(t) =& r_0 C_0 \left(1 - \frac{C_0 + b_{01} C_1}{K_0}\right), \\ \dot{C}_1(t) =& r_1 C_0 \left(1 - \frac{b_{10} C_0 + C_1}{K_1}\right).\end{aligned} \tag{2.1}$$

By multiplying the numerator and the denominator of the fractions involving K_i $(i = 0, 1)$ for suitable constants, one can homogenize the values of these parameters, obtaining an unique carrying capacity K of the system. This lead to get the following system of ODEs:

$$\begin{aligned}\dot{C}_0(t) =& r_0 C_0 \left(1 - \frac{\bar{b}_{00} C_0 + \bar{b}_{01} C_1}{K}\right), \\ \dot{C}_1(t) =& r_1 C_0 \left(1 - \frac{\bar{b}_{10} C_0 + \bar{b}_{11} C_1}{K}\right).\end{aligned}$$

For simplicity of notation, let us rename the parameters $\bar{b}_{ij}$ as b_{ij}. In this way, we can generalize the model to n tumoral genotypes as:

$$\dot{C}_i(t) = r_i C_i \left(1 - \frac{1}{K} \sum_{j=0}^{n} b_{ij} C_j\right). \tag{2.2}$$

In these equations, $r = (r_i)_{i=0}^n$ represents the vector of the growth rates and the elements of matrix $B = (b_{ij})_{i,j=0}^n$ represent the competitive abilities of the genotypes.

In order to introduce a possibility of mutation into the model, it is necessary to separate the birth term and the death one: in fact, mutations can appear only when cells reproduce, thus in the birth term.
In the simplest case of a single population we have

$$\text{population size growth} = \text{birth term - death term.}$$

In formulas, if C is the size of the population, this leads to the following ODE:

$$\dot{C} = \underbrace{aC\left(1 - h\frac{C}{K}\right)}_{\text{birth}} - \underbrace{dC\left(1 + g\frac{C}{K}\right)}_{\text{death}}, \tag{2.3}$$

where the new parameters a and d are respectively the birth and death rate, while h and g are to fine-tune the effects of the lack of resources on the births and deaths (usually, such effects on the birth term is very small, i.e. $h \approx 0$).

Since we know that the population growth follows the logistic growth, rearranging the terms of (2.3) one must find the logistic law

$$\dot{C} = rC\left(1 - \frac{C}{K}\right), \tag{2.4}$$

where r is the growth rate and K the carrying capacity of the system. In order to have the same ODE, the conditions $r = a - d$ and $r = ah + dg$ must hold. Since d represents the inverse of the average life span, then, if r is known, a can be simply computed. Likewise, g can be calculated if we neglect h, since, as said before, its value is usually very small.

Going back to the principal model (2.2), following the same procedure, we can separate the birth term and the death term, which are the following:

$$\text{birth term for the } i\text{-th population} = a_i C_i \left(1 - h_i \frac{\sum_{k=0}^{n} b_{ik} C_k}{K}\right), \tag{2.5}$$

$$\text{death term for the } i\text{-th population} = -d_i C_i \left(1 - g_i \frac{\sum_{j=0}^{n} b_{ik} C_k}{K}\right), \tag{2.6}$$

where a_i and d_i denote the birth and death rates of the cells of genotype i and h_i and g_i are the fine-tuning parameters. These parameters are related to r_i as following:

$$r_i = a_i - d_i\,; \qquad a_i h_i + d_i g_i = r_i\,.$$

It is possible now to modify the birth term (2.5) introducing the possibility of mutation, as suggested in [4], obtaining, for the i-th genotype,

$$\sum_{j=0}^{n} \left(p_{ji} a_j C_j \left(1 - h_j \frac{\sum_{k=0}^{n} b_{jk} C_k}{K}\right)\right) \tag{2.7}$$

where p_{ij} denote the probability that a cell of genotype i produces a cell of genotype j.

By the combination of (2.7) and (2.6), we get the following form of the system of ODEs that we were looking for:

$$\dot{C}_i = \sum_{j=0}^{n} \left(p_{ji} a_j C_j \left(1 - h_j \frac{\sum_{k=0}^{n} b_{jk} C_k}{K}\right)\right) - d_i C_i \left(1 + g_i \frac{\sum_{k=0}^{n} b_{ik} C_k}{K}\right). \tag{2.8}$$

Remark 2.1.1. *All the physical parameters of the model are positive real numbers, except the elements of matrix P, which can be zeros.*

2.2 Non-dimensionalization of the model

In order to reduce the number of parameters of the model, we non-dimensionalize the system of ODEs (2.8). For $n = 1$, the system is

$$\begin{aligned}
\dot{C}_0 =& p_{00}a_0C_0\left(1 - h_0\frac{b_{00}C_0 + b_{01}C_1}{K}\right) + p_{10}a_1C_1\left(1 - h_1\frac{b_{10}C_0 + b_{11}C_1}{K}\right) + \\
&- d_0C_0\left(1 + g_0\frac{b_{00}C_0 + b_{01}C_1}{K}\right), \\
\dot{C}_1 =& p_{01}a_0C_0\left(1 - h_0\frac{b_{00}C_0 + b_{01}C_1}{K}\right) + p_{11}a_1C_1\left(1 - h_1\frac{b_{10}C_0 + b_{11}C_1}{K}\right) + \\
&- d_1C_1\left(1 + g_1\frac{b_{10}C_0 + b_{11}C_1}{K}\right).
\end{aligned}$$

Rearranging its terms, one can obtain:

$$\begin{aligned}
\frac{dC_0}{dt} &= (p_{00}a_0 - d_0)C_0 - (p_{00}a_0h_0 + d_0g_0)C_0\frac{b_{00}C_0 + b_{01}C_1}{K} + p_{10}a_1C_1\left(1 - h_1\frac{b_{10}C_0 + b_{11}C_1}{K}\right), \\
\frac{dC_1}{dt} &= (p_{11}a_1 - d_1)C_1 - (p_{11}a_1h_1 + d_1g_1)C_1\frac{b_{10}C_0 + b_{11}C_1}{K} + p_{01}a_0C_0\left(1 - h_0\frac{b_{00}C_0 + b_{01}C_1}{K}\right).
\end{aligned}$$

Let us denote

$$x_i = \frac{b_{ii}}{K}C_i\,, \qquad \tau = (p_{00}a_0h_0 + d_0g_0)t\,. \tag{2.9}$$

Then, the system becomes:

$$\begin{aligned}
\frac{dx_0}{d\tau} =& x_0 + p_{10}\frac{a_1}{\sigma_0}e_{01}x_1 - x_0\left(x_0 + f_{01}x_1\right) - p_{10}\frac{a_1h_1}{\sigma_0}e_{01}x_1\left(f_{10}x_0 + x_1\right), \\
\frac{dx_1}{d\tau} =& p_{01}\frac{a_0}{\sigma_0}e_{10}x_0 + \frac{l_1}{\sigma_0}x_1 - \frac{\sigma_1}{\sigma_0}x_1\left(f_{10}x_0 + x_1\right) - p_{01}\frac{a_0h_0}{\sigma_0}e_{10}x_0\left(x_0 + f_{01}x_1\right),
\end{aligned}$$

where new parameters are

$$l_i = p_{ii}a_i - d_i\,, \qquad \sigma_i = p_{ii}a_ih_i + d_ig_i\,, \qquad e_{ij} = \frac{b_{ii}}{b_{jj}}\,, \qquad f_{ij} = \frac{b_{ij}}{b_{jj}}\,. \tag{2.10}$$

Generalizing this process to n cancer genotypes, by the change of variables and parameters given by (2.9) and (2.10), we get the non-dimensionalized system of ODEs

$$\frac{dx_i}{d\tau} = x_i \left(\frac{l_i}{\sigma_0} - \frac{\sigma_i}{\sigma_0} \left(\sum_{k=0}^{k=n} f_{ik} x_k \right) \right) + \sum_{\substack{j=0 \\ j \neq i}}^{j=n} p_{ji} \frac{a_j}{\sigma_0} e_{ij} x_j \left(1 - h_j \left(x_j + \sum_{k=0}^{k=n} f_{jk} x_k \right) \right) , \quad (2.11)$$

for $i = 0, \ldots, n$.

Remark 2.2.1. *Let us notice that σ_0 is always a positive quantity, since $d_0, g_0 > 0$.*

It is possible to write the system in an easier way, observing that it is possible to separate the linear and the non-linear parts as following:

$$\frac{dx_i}{d\tau} = \frac{1}{\sigma_0} \left(\sum_{j=0}^{n} u_{ij} x_j - \sum_{j=0}^{n} v_{ij} \sum_{k=0}^{n} x_j f_{jk} x_k \right) , \; i = 0, \ldots, n , \quad (2.12)$$

where the new parameters are

$$u_{ij} = \begin{cases} l_j & \text{if } j = i , \\ a_j p_{ji} e_{ij} & \text{if } j \neq i , \end{cases} \quad (2.13)$$

and

$$v_{ij} = \begin{cases} \sigma_j & \text{if } j = i , \\ a_j h_j p_{ji} e_{ij} & \text{if } j \neq i . \end{cases} \quad (2.14)$$

Chapter 3

Analysis of model properties: case of a planar system

In this chapter, the non-dimensional model is studied for $n = 1$, that is for two genotypes, one normal and one tumoural. First of all, the nullclines are analyzed in order to establish the number of non-negative fixed points of the model and their location. Then, some sufficient (and, in some cases, also necessary) conditions for the stability of these points are stated. Lastly, a biological interpretation of the analytical results is provided.

From now on, we introduce a reasonable hypothesis on the system in order to preserve the positivity of the model: namely, we suppose that $h_i = 0$ for all $i = 0, \dots, n$. Mathematically, this condition is necessary to preserve the positive-invariance of the first quadrant, otherwise the model would not be realistic. Biologically, this means that competion influences the death but not the proliferation. In terms of equation, the system (2.12) can be re-written as following:

$$\frac{dx_i}{d\tau} = \frac{1}{\sigma_0}\left(\sum_{j=0}^{n} u_{ij}x_j - d_i g_i \sum_{j=0}^{n} x_i f_{ij} x_j\right) \,, \; i = 0, \dots, n \,, \tag{3.1}$$

3.1 Equilibrium states

For simplicity, we start with $n = 1$, that is with two genotypes, normal and malignant. Writing (3.1) for this case, one obtains

$$\begin{aligned}
\dot{x_0} &= \frac{1}{\sigma_0}(u_{00}x_0 + u_{01}x_1 - d_0 g_0 f_{00} x_0^2 - d_0 g_0 f_{01} x_0 x_1) \,, \\
\dot{x_1} &= \frac{1}{\sigma_0}(u_{10}x_0 + u_{11}x_1 - d_1 g_1 f_{10} x_0 x_1 - d_1 g_1 f_{11} x_1^2) \,.
\end{aligned}$$

To simplify the notation, we set

$$x = x_0\,, \qquad y = x_1$$

and

$$A = u_{00} = l_0\,, \qquad B = u_{01} = a_1 p_{10}\frac{b_{00}}{b_{11}}\,, \qquad C = u_{10} = a_0 p_{01}\frac{b_{11}}{b_{00}}\,, \qquad D = u_{11} = l_1\,,$$

$$\alpha = d_0 g_0 f_{00} = d_0 g_0\,, \qquad \beta = d_0 g_0 f_{01} = d_0 g_0 \frac{b_{01}}{b_{11}}\,,$$

$$\gamma = d_1 g_1 f_{10} = d_1 g_1 \frac{b_{10}}{b_{00}}\,, \qquad \delta = d_1 g_1 f_{11} = d_1 g_1\,.$$

By these notation, the system can be written as following:

$$\begin{aligned} \dot{x} &= \frac{1}{\sigma_0}(Ax + By - \alpha x^2 - \beta xy)\,, \\ \dot{y} &= \frac{1}{\sigma_0}(Cx + Dy - \gamma xy - \delta y^2)\,. \end{aligned} \tag{3.2}$$

Equilibria of the model satisfy to the system of algebraic equations

$$\begin{aligned} Ax + By - \alpha x^2 - \beta xy &= 0\,, \\ Cx + Dy - \gamma xy - \delta y^2 &= 0\,. \end{aligned} \tag{3.3}$$

Then, in this case, the equilibrium states are represented by the intersections of two conic curves defined by (3.3), which could degenerate in a couple of straight lines, depending on the value of the parameters of the algebraic system. Let us start with some trivial observations expressed in the following lemmas.

Lemma 3.1.1. *The origin $P_0 = (0, 0)$ is always a fixed point.*

Proof. The origin always belongs to both curves because the equations which define them have constant terms equal to 0. Hence, they cross in the origin, which means that the origin is a steady state. □

This fact is biologically meaningful, because it means that cells cannot appear out of nowhere.

Lemma 3.1.2. *The dynamical system in analysis can have from one to four fixed points.*

Proof. By the previous lemma, we have that there is always a fixed point, namely the origin. Moreover, the steady states are exactly the solutions of (3.3), which is equivalent to a fourth-degree equation; therefore, they can be at most four. □

Another limit to the number of fixed point is given by the positivity condition: the variables x and y represent sizes of populations and, hence, they cannot assume negative values: this means that the interesting intersections are the ones which are in the first quadrant, i.e. the ones which have both coordinates non-negative.

Remark 3.1.3. *In order to find an upper bound on the number of non-negative equilibrium points, one could use an algebraic criterion, called Descartes' rule of signs, fully explained in [3].*

To start, it is possible to state the following result:

Lemma 3.1.4. *Apart from the origin, the model* (3.2) *has no fixed point on the coordinate axes.*

Proof. Let us suppose, by contradiction, that the model admits a steady state on the y-axis, which has the form $(0, y^*)$, $y^* \neq 0$. It means that $(0, y^*)$ is a solution of the algebraic system (3.3); hence, substituting $x = 0$ in (3.3), one obtains

$$y = 0\,, \qquad y\left(1 - \frac{\delta}{D}y\right) = 0\,.$$

It implies that $y^* = 0$, which is a contradiction.

Likewise, it can be proved that, $(x^*, 0)$ cannot be a solution unless $x^* = 0$.

□

To get the objective, first of all let us study the properties of the two curves.

First curve: $Ax + By - \alpha x^2 - \beta xy = 0$

It is convenient to isolate the variable y in it, obtaining

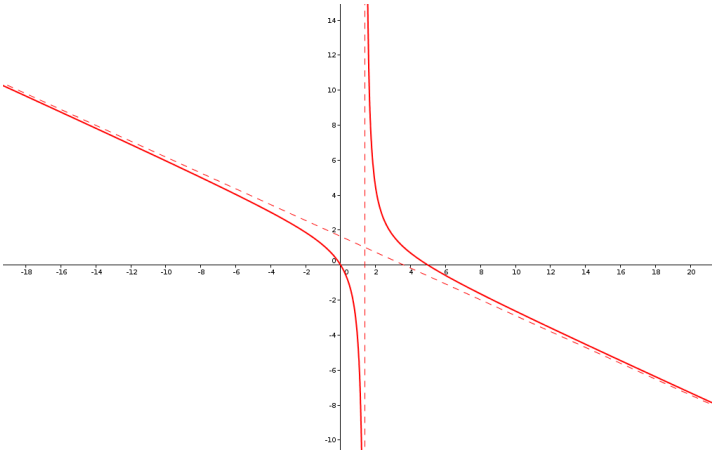
$$y = f(x) = \frac{\alpha x^2 - Ax}{B - \beta x}\,.$$

The first conic curve is exactly the graph of the function f.

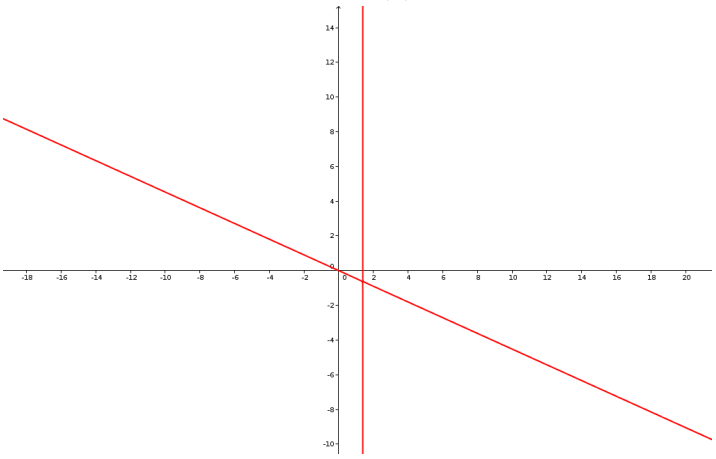
It is convenient to study which kind of curve it is. It is evident that it has a vertical asymptote described by the equation

$$x = \frac{B}{\beta}\,,$$

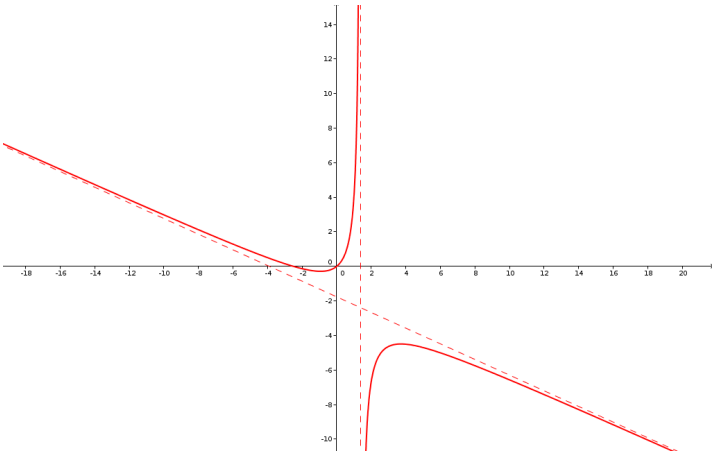
where $B/\beta \geq 0$.



(a)



(b)



(c)

Figure 3.1: The first hyperbola for: (a) $A\beta - B\alpha > 0$, (b) $A\beta - B\alpha = 0$, (c) $A\beta - B\alpha < 0$.

Moreover, we can see if it has another asymptote, which, evidently, cannot be horitzontal. Then, if it exists, it must be oblique and its slope would be given by

$$m_1 = \lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \frac{\alpha x^2 - Ax}{Bx - \beta x^2} = -\frac{\alpha}{\beta}\,,$$

where $m_1 < 0$. Since the previous limit is finite, the oblique asymptote exists and its constant term is

$$q_1 = \lim_{x\to\infty} (f(x) - m_1 x) = \lim_{x\to\infty} \left(\frac{\alpha x^2 - Ax}{B - \beta x} + \frac{\alpha}{\beta} x \right) = \frac{A\beta - B\alpha}{\beta^2}\,.$$

We have just proved the following

Lemma 3.1.5. *The curve whose equation is $Ax + By - \alpha x^2 - \beta xy = 0$ is a hyperbola. Its asymptotes are given by equations*

$$x = \frac{B}{\beta} \qquad \textit{and} \qquad y = -\frac{\alpha}{\beta} x + \frac{A\beta - B\alpha}{\beta^2}\,.$$

The hyperbola degenerates into a couple of straight lines if and only if $A\beta - B\alpha = 0$; in this case, the straight lines correspond to the asymptotes.

Depending on the sign of the constant term of the equation of the oblique asymptote, the position of the two branches changes as it can be seen in the figures 1a-1c. Anyway, it is possible to state the following result:

Lemma 3.1.6. *One of the two branches of the hyperbola has no points in the first quadrant.*

Proof. Let us distinguish three possible cases, that cover all the possibilities:

1. $A\beta - B\alpha > 0$ (*Figure 1(a)*)

   Let us compute the first derivative of f, which is

   $$f'(x) = \frac{2B\alpha x - \alpha\beta x^2 - AB}{(B - \beta x)^2}\,.$$

   The sign of $f'(x)$ depends only on the numerator of the fraction, whose terms can be re-arranged in the following way:

   $$2B\alpha x - \alpha\beta x^2 - AB = -\frac{\alpha}{\beta} \left(x - \frac{B}{\beta} \right)^2 - \frac{B}{\beta}(A\beta - B\alpha) < 0$$

In other words, f is a decreasing function for all $x \neq B/\beta$. In particular, since $f(0) = 0$ (because the origin belongs to the hyperbola), $f(x) > 0$ for $x < 0$ and $f(x) < 0$ for $0 < x < B/\beta$. It means that one of the two branches has no point in the first quadrant.

2. $A\beta - B\alpha = 0$ (*Figure 1(b)*)

   The hyperbola degenerates to its asymptotes which cross in the point

   $$(B/\beta, -B\alpha/\beta^2).$$

   Since one of these two straigth lines is vertical and the other one has negative slope and pass through the origin $(0, 0)$, a branch of the hyperbola lays entirely in the fourth quadrant.

3. $A\beta - B\alpha < 0$ (*Figure 1(c)*)

   Since the origin belongs to the graph of f, the branch of the hyperbola which contains it lays in the region delimited on the right by the vertical asymptote and downwards by the oblique one. It means that the other branch lays in the opposite region, which belongs entirely to the fourth quadrant.

□

Second curve: $Cx + Dy - \gamma xy - \delta y^2 = 0$

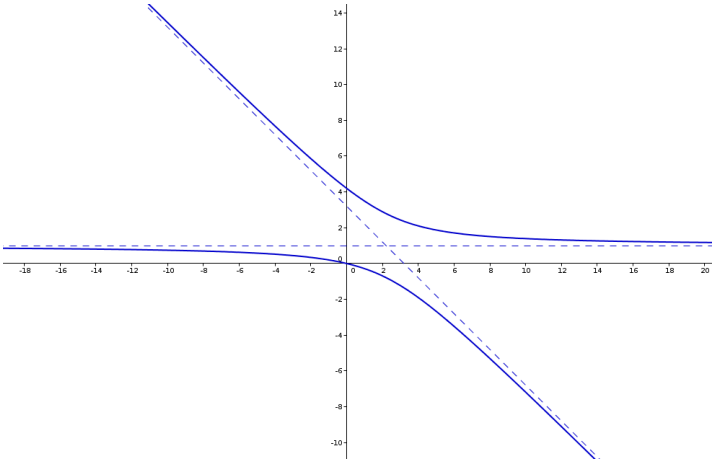
Likewise, it is convenient to isolate the variable x in the equation, obtaining

$$x = g(y) = \frac{\delta y^2 - Dy}{C - \gamma y}\,.$$

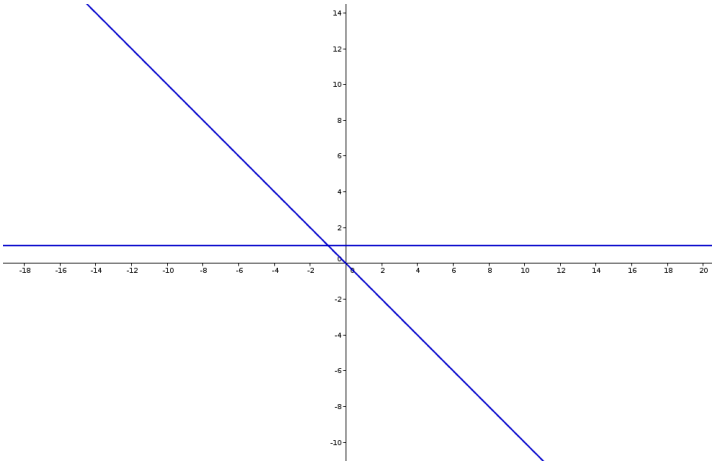
The second curve, then, is the graph of the function g. Repeating similar computations to the previous section, it results that the two asymptotes are described respectively by the equations

$$y = \frac{C}{\gamma} \geq 0 \qquad \text{and} \qquad x = -\frac{\delta}{\gamma}y + \frac{D\gamma - C\delta}{\gamma^2}\,.$$

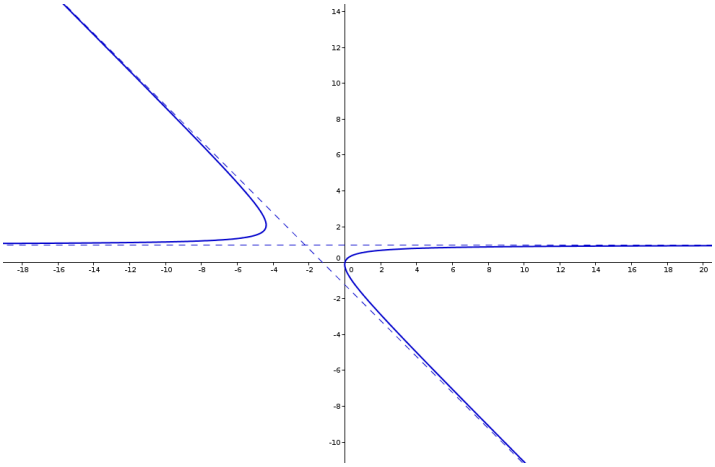
and, as a consequence, the conic curve is a hyperbola. It degenerates into a couple of straight lines if and only if $D\gamma - C\delta = 0$; in this case, the straight lines correspond to the asymptotes. Depending on the sign of the constant term of the equation of the oblique asymptote, the position of the two branches changes as it can be seen in the figures 2a-2c.



(a)



(b)



(c)

Figure 3.2: The second hyperbola for: (a) $D\gamma - C\delta > 0$, (b) $D\gamma - C\delta = 0$, (c) $D\gamma - C\delta < 0$.

As for the first hyperbola, just a portion of one of the two branches of the hyperbola belongs to the first quadrant. As a consequence of this fact, it is possible to state the following result:

Lemma 3.1.7. *The model* (3.2) *has either exactly one or exactly two non-negative steady states.*

Proof. As stated in a previous lemma, the origin is always a steady state. Then, we can restrict to prove that there exists, at most, one steady state different from it. Let us consider the following six cases, that cover all the possibilities:

1. $A\beta - B\alpha > 0,\ D\gamma - C\delta > 0$

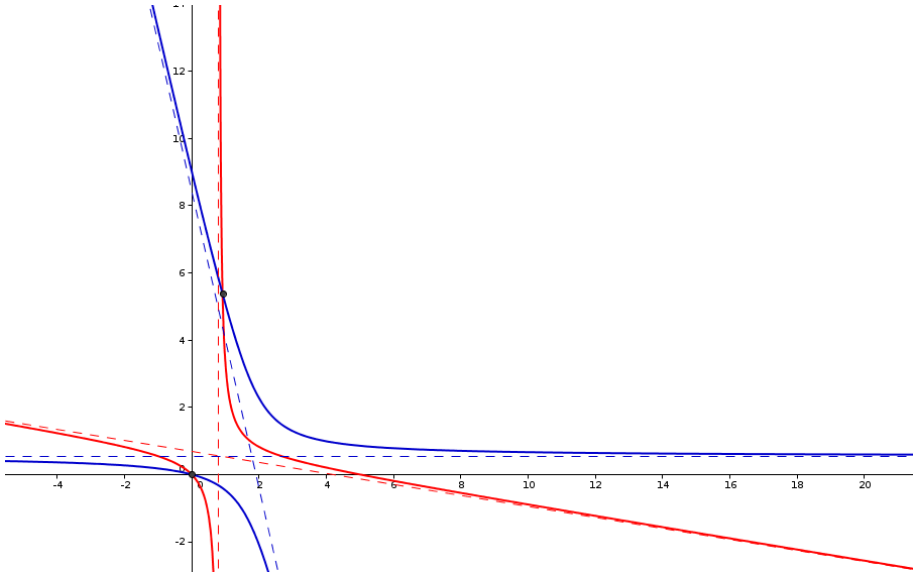


Figure 3.3(a): Intersections for $A\beta - B\alpha > 0$ and $D\gamma - C\delta > 0$.

For all $x \in]B/\beta, +\infty[$ the function f is strictly decreasing, as proved before; $f(x) \to +\infty$ as $x \to (B/\beta)^+$ and $f(x) \to -\infty$ as $x \to +\infty$ since the hyperbola tends to the oblique asymptote, that has negative slope. The same holds for g: it is strictly decreasing for all $y \in]C/\gamma, +\infty[$, $g(y) \to +\infty$ as $y \to (C/\gamma)^+$ and $g(y) \to -\infty$ as $y \to +\infty$. It means that, in this case, there is exactly one intersection different from the origin and, hence, there are exactly two non-negative steady states.

2. $A\beta - B\alpha > 0,\ D\gamma - C\delta < 0$

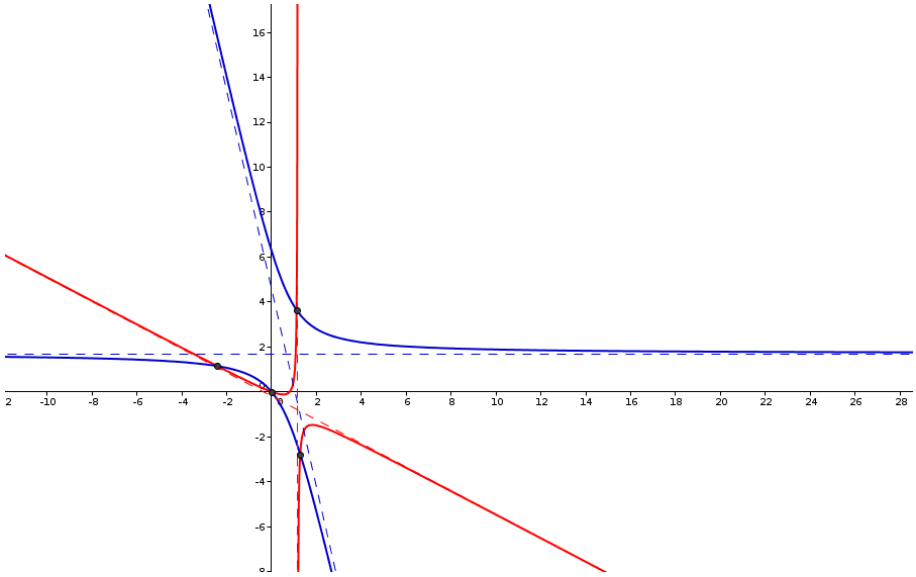


Figure 3.3(b): Intersections for $A\beta - B\alpha > 0$ and $D\gamma - C\delta < 0$.

As in the previous case, for all $x \in]B/\beta, +\infty[$ the function f is strictly decreasing, as proved earlier, $f(x) \to +\infty$ as $x \to (B/\beta)^+$ and $f(x) \to -\infty$ as $x \to +\infty$. Regarding to the second hyperbola, let us remind that

$$g'(y) = \frac{2C\delta y - \gamma\delta y^2 - CD}{(C - \gamma y)^2}\,.$$

Therfore, the sign of g' depends only on the numerator, which is positive if and only if $y \in]y_1, y_2[\setminus \{C/\gamma\}$, where

$$y_{1,2} = \frac{C}{\gamma} \pm \frac{\sqrt{C\delta(C\delta - D\gamma)}}{\gamma\delta}$$

exist because the discriminant is non-negative by hypothesis. In particular g is strictly increasing for all $y \in]y_1, C/\gamma[$. Since $g(y) \to +\infty$ as $y \to -\infty$ and $g(y) \to +\infty$ as $y \to (C/\gamma)^-$ and y_1 is the unique stationary point in the interval $]-\infty, C/\gamma[$, then it must be an absolute minimum point of this interval and, therefore, $g(y_1) < g(0) = 0$. Hence, independently of the sign of y_1, there exists a left neighbourhood $I = [y^*, C/\gamma[$ of C/γ , with $\max\{0, y_1\} \leq y^* < C/\gamma$ and $g(y^*) = 0$, where the function g is strictly increasing and, then, positive. It is simple to prove that either $y^* = 0$, if $D \leq 0$, or $y^* = D/\delta$, if $D > 0$ (in other words, $y^* = \max\{0, D/\delta\}$).

Therefore, in this case there is exactly one non-negative steady state different from the origin.

3. $A\beta - B\alpha < 0,\ D\gamma - C\delta > 0$

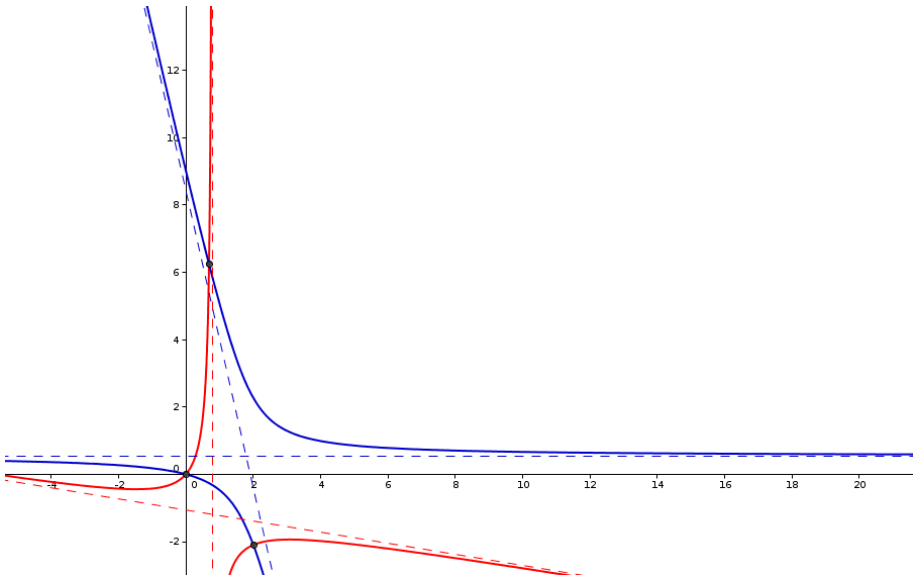


Figure 3.3(c): Intersections for $A\beta - B\alpha < 0$ and $D\gamma - C\delta > 0$.

Simmetrical of the previous case.

4. $A\beta - B\alpha < 0,\ D\gamma - C\delta < 0$

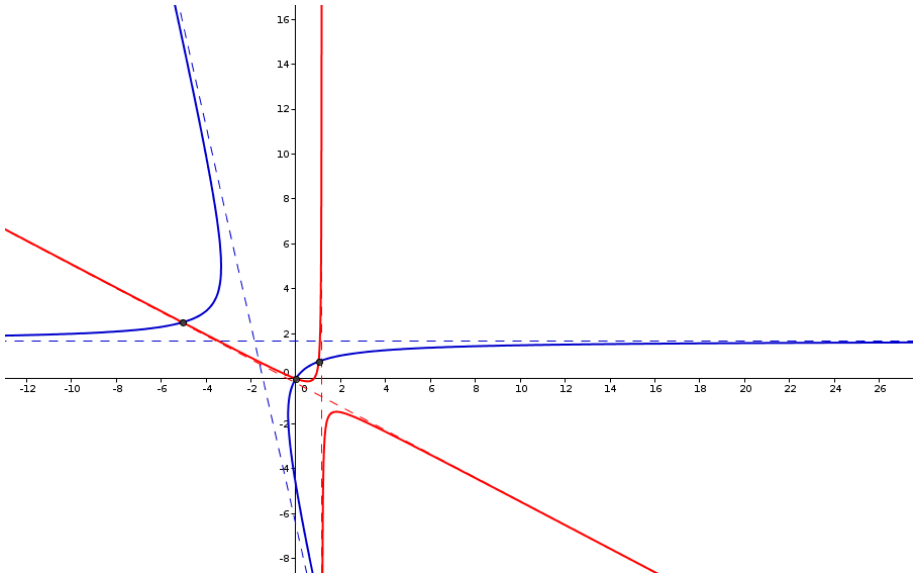


Figure 3.3(d): Intersections for $A\beta - B\alpha < 0$ and $D\gamma - C\delta < 0$.

As seen in the previous two cases, the portion of hyperbolas in the first quadrant are such that they can cross at most in one non-negative point different from the origin.

5. $A\beta - B\alpha = 0$

In this case, the first hyperbola degenerates into a couple of straigth lines. As observed before, the oblique one has a negative slope and crosses the origin, hence it lays in the second and in the fourth quadrant. As a consequence, just an intersection of the second hyperbola with the vertical line of equation $x = B/\beta$ can be a non-negative fixed point. Let us study the number of intersections in the first quadrant; first of all, the first coordinate of such a point B/β and its second coordinate must satisfy the following equation, obtained by substituting B/β to x in the $Cx + Dy - \gamma xy - \delta y^2 = 0$:

$$\beta\delta y^2 + (B\gamma - D\beta)y - BC = 0\,.$$

The solutions of this equation are

$$y_\pm = \frac{D\beta - B\delta \pm \sqrt{(B\gamma - D\beta)^2 + 4BC\beta\delta}}{2\beta\delta}.$$

Both of them are real since the discriminant $\Delta = (B\gamma - D\beta)^2 + 4BC\beta\delta$ is non negative thanks to the non-negativity of the parameters B, C, β, δ. It is evident from the geometrical observations done in the previous cases that only y_+ can be positive. Hence, there is only one fixed point strictly in the first quadrant, which is $(B/\beta, y_+)$.

6. $D\gamma - C\delta = 0$

In this case, the second hyperbola degenerates into a couple of straigth lines. By the same scheme of the previous point, we obtain that, if it does not coincide with the origin, the second steady state is $(x_+, C/\gamma)$, where

$$x_+ = \frac{C\beta - A\gamma + \sqrt{(C\beta - A\gamma)^2 + BC\alpha\gamma}}{2\alpha\gamma}\,.$$

□

3.2 Stability of P_0

In order to study the stability of the origin, we can proceed by linearizing the system (3.2) around it. The Jacobian matrix is

$$J(0,0) = J_0 = \frac{1}{\sigma_0}\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3.4}$$

and its determinant and trace are, respectively

$$det(J_0) = \frac{1}{\sigma_0^2}(AD - BC) \qquad \text{and} \qquad tr(J_0) = \frac{1}{\sigma_0}(A + D)\,.$$

Another interesting quantity is

$$tr(J_0)^2 - 4det(J_0) = \frac{1}{\sigma_0^2}[(A - D)^2 + 4BC]$$

which is always non-negative. It means that, in any case, the eigenvalues of J_0 are real and, hence, $(0,0)$ is not a focus. Furthermore, as a consequence, no Hopf bifurcation is possible in the origin. Then, there are only three possibilities:

- if $AD - BC < 0$, the origin is a saddle point, hence it is unstable;
- if $AD - BC > 0$ and $(A + D) > 0$, the origin is a repulsive node, hence it is unstable;
- if $AD - BC > 0$ and $(A + D) < 0$, the origin is an attractive node, hence it is stable.

3.3 Existence and location of the positive fixed point

Let us considerate some possible cases in order to study the existence and the location of the possible positive steady state, that will be denoted by $P_* = (x_*, y_*)$.

1. $A\beta - B\alpha = 0$

   As seen in Lemma 3.1.7, in this case

$$P_* = \left(\frac{B}{\beta}, \frac{D\beta - B\delta + \sqrt{(B\gamma - D\beta)^2 + 4BC\beta\delta}}{2\beta\delta}\right).$$

2. $D\gamma - C\delta = 0$

   As seen in Lemma 3.1.7, in this case

$$P_* = \left(\frac{C\beta - A\gamma + \sqrt{(C\beta - A\gamma)^2 + BC\alpha\gamma}}{2\alpha\gamma}, \frac{C}{\gamma}\right).$$

3. $A\beta - B\alpha > 0,\ D\gamma - C\delta > 0$

   In this case, P_* always exists. The part of the first hyperbola in the first quadrant is the graph of f for $x \in]B/\beta, A/\alpha[$. At the same way, the part of the second hyperbola in the first quadrant is the graph of g for $y \in]C/\gamma, D/\delta[$. This means that

   $$P_* \in]B/\beta, A/\alpha[\times]C/\gamma, D/\delta[.$$

4. $A\beta - B\alpha > 0,\ D\gamma - C\delta < 0$

   In this case, P_* always exists. Observing the interval of x and y where the graph of f and g, respectively, are in the first quadrant, it results that:

   - if $D \leq 0,\ P_* \in]B/\beta, A/\alpha[\times]0, C/\gamma[$;
   - if $D > 0,\ P_* \in]B/\beta, A/\alpha[\times]D/\delta, C/\gamma[$.

5. $A\beta - B\alpha < 0,\ D\gamma - C\delta > 0$

   In this case, P_* always exists. Simmetrically to the previous case,

   - if $A \leq 0,\ P_* \in]0, B/\beta[\times]C/\gamma, D/\delta[$;
   - if $A > 0,\ P_* \in]A/\alpha, B/\beta[\times]C/\gamma, D/\delta[$.

6. $A\beta - B\alpha < 0,\ D\gamma - C\delta < 0$

   This is the only case for which P_* may both exist and not exist. If, by changing the parameters of the system, P_* appears, it means that a bifurcation occurs. A saddle-node bifurcation is not possible because it would means that there is a third possible steady state, in contraddiction with Lemma 3.1.7, and for the same reason neither is a pitchfork bifurcation. Hence, the unique possibility is that the bifurcation is transcritical. Therefore, P_* is generated when the two hyperbolas are tangent in the origin. This occurs when the raws of the matrix (3.4) are proportional, that is there exists k such that

   $$A = kC\,, \qquad B = kD\,.$$

   It means that a transcritical bifurcation occurs when $AD - BC = 0$. The interesting fact, at this point, is if P_* exists when $AD - BC < 0$ or when $AD - BC > 0$. We will study this problem in a further section.

3.4 Global analysis

It would be interesting to make a global analysis of the system, in order to make some previsions about the large-time behaviour of the system. To do it, let us remind a very important theorem, presented, for instance, in [10].

Theorem 3.4.1 (Poincaré-Bendixson)**.** *Let D be a closed bounded region of the plane and*

$$\begin{aligned} \dot{x} &= P(x,y) \\ \dot{y} &= Q(x,y) \end{aligned}$$

be a dynamical system in which P and Q are continuously differentiable. If a trajectory of the dynamical system is such that it remains in D for all $t \geq 0$, then the trajectory must

1. *be a closed orbit or*
2. *approach a closed orbit or*
3. *approach an equilibrium point as $t \to +\infty$.*

Let us notice now that the model (3.2) can be written in the form

$$\begin{aligned} \dot{x} &= \frac{1}{\sigma_0}((A-\alpha x)x + (B-\beta x)y) = P(x,y)\,, \\ \dot{y} &= \frac{1}{\sigma_0}((C-\gamma y)x + (D-\delta y)y) = Q(x,y)\,. \end{aligned} \tag{3.5}$$

Let us study the direction of the vector field (P,Q) in a point (x,y) of the first quadrant. It is simple to observe that, if $x > \max\{A/\alpha, B/\beta\}$, then $P(x,y) < 0$; analogously, if $y > \max\{C/\gamma, D/\delta\}$, then $Q(x,y) < 0$. Moreover, on the positive semi-axes, the vector field points at the first quadrant. It means that the compact square $S = [0, \max\{A/\alpha, B/\beta\}] \times [0, \max\{C/\gamma, D/\delta\}]$ is an attractive region, therefore it contains at least one stable limit cycle or at least one stable fixed point. For this system, it is possible to exclude the existence of a limit cycle in some cases due to the following result, proved, for instance, in [10]:

Theorem 3.4.2 (Bendixson-Dulac)**.** *If there exists a $\mathcal{C}^1$ function $\varphi(x,y)$ such that the quantity $div(\varphi P, \varphi Q)$ has the same sign almost everywhere in a simply connected region of the plane, then the planar system*

$$\begin{aligned} \dot{x} &= P(x,y) \\ \dot{y} &= Q(x,y) \end{aligned}$$

has no non-constant periodic solutions lying entirely within the region. In particular, no limit cycle is entirely contained within the region.

Choosing $\varphi(x, y) = \sigma_0$, one immediately obtains for the system (3.2)

$$div(\varphi P, \varphi Q) = A + D - (2\alpha + \gamma)x - (\beta + 2\delta)y\,,$$

which is always negative if $A + D < 0$. Therefore, since S is a simply connected region, according to Bendixson-Dulac theorem no limit cycle lays within it. It means that all orbits tend to a stable fixed point. Recall that, for $A + D < 0$, the origin is stable if and only if $AD - BC > 0$. Hence, for $AD - BC < 0$, P_* must exist and be stable. This means that the condition for the existence of P_* in the sixth case of *Section 3.3* is that $AD - BC < 0$ holds.

3.5 Stability of P_*

Apart from the global analysis done before, it is possible to do a local analysis in order to obtain more information about the stability of the positive fixed point P_*. The linearized system around P_* has Jacobian matrix

$$J(x_*, y_*) = J_* = \frac{1}{\sigma_0}\begin{pmatrix} A - 2\alpha x_* - \beta y_* & B - \beta x^* \\ C - \gamma y_* & D - 2\delta y_* - \gamma x_* \end{pmatrix}. \tag{3.6}$$

Let us remind that (x_*, y_*) are coordinates of the point P_* and are solution of (3.2); therefore,

$$A - \alpha x_* - \beta y_* = -\alpha x_* - B\frac{y_*}{x_*} \qquad \text{and} \qquad D - \delta y_* - \gamma = -\delta y_* - C\frac{x_*}{y_*}\,.$$

This allows us to re-write J_* as following:

$$J(x_*, y_*) = J_* = \frac{1}{\sigma_0}\begin{pmatrix} -\alpha x_* - By_*/x_* & B - \beta x^* \\ C - \gamma y_* & -\delta y_* - Cx_*/y_* \end{pmatrix}. \tag{3.7}$$

Then, the trace of J_* is

$$tr(J_*) = \frac{1}{\sigma_0}\left(-\alpha x_* - B\frac{y_*}{x_*} - \delta y_* - C\frac{x_*}{y_*}\right)$$

and its determinant is

$$
\begin{aligned}
det(J_*) &= \frac{1}{\sigma_0^2}\left[(\alpha\delta-\beta\gamma)x_*y_* + B\delta\frac{y_*^2}{x_*} + C\alpha\frac{x_*^2}{y_*} + B\gamma y_* + C\beta x_*\right] \\
&= \frac{1}{\sigma_0^2}\left[(\alpha\delta-\beta\gamma)x_*y_* + B\frac{\delta y_*^2+\gamma x_*y_*}{x_*} + C\frac{\alpha x_*^2+\beta x_*y_*}{y_*}\right] \\
&= \frac{1}{\sigma_0^2}\left[(\alpha\delta-\beta\gamma)x_*y_* + B\frac{Cx_*+Dy_*}{x_*} + C\frac{Ax_*+By_*}{y_*}\right] \\
&= \frac{1}{\sigma_0^2}\frac{1}{x^*y^*}\left[(\alpha\delta-\beta\gamma)x_*^2y_*^2 + ACx_*^2 + 2BCx_*y_* + BDy_*^2\right] .
\end{aligned}
$$

At this point, it is possible to state the following

Lemma 3.5.1. *No Hopf bifurcation is possible at P_*.*

Proof. Let us remind that a Hopf bifurcation occurs around at a fixed point when the trace of the Jacobian matrix is zero and its determinant is positive. In this case, $tr(J_*)$ is always non-zero; in particular, it is always negative. In fact

$$
\frac{1}{\sigma_0} > 0 \qquad \text{and} \qquad -\alpha x_* - B\frac{y_*}{x_*} - \delta y_* - C\frac{x_*}{y_*} < 0
$$

always because $B, C \geq 0$ and $\alpha x_*, \delta y_* > 0$. □

Therefore, reminding that no Hopf bifurcation is possible at $(0,0)$, one can deduce the following

Corollary 3.5.2. *The model* (3.2) *admits no Hopf bifurcation.*

Since $tr(J_*)$ is always negative, P_* can be stable: in particular, a sufficient (but not necessary) condition for its stability is that $\alpha\delta - \beta\gamma > 0$, which makes the determinant positive. Depending on the values of the parameters, in this case, P_* can be a stable node or a stable focus.

3.6 Biological interpretation

Let us compute the relevant parameters we have found so far as a function of the physical parameters:

$$
A\beta - B\alpha = \frac{d_0g_0}{b_{11}}(a_0p_{00}b_{01} - a_1p_{10}b_{00} - d_0b_{01})\,;
$$

$$
D\gamma - C\delta = \frac{d_1g_1}{b_{00}}(a_1p_{11}b_{10} - a_0p_{01}b_{11} - d_1b_{10})\,;
$$

$$AD - BC = (a_0 p_{00} - d_0)(a_1 p_{11} - d_1) - a_0 a_1 p_{01} p_{10}\,;$$

$$A + D = a_0 p_{00} + a_1 p_{11} - (d_0 + d_1)\,;$$

$$\alpha\delta - \beta\gamma = \frac{d_0 d_1 g_0 g_1}{b_{00} b_{11}}(b_{00} b_{11} - b_{01} b_{10})\,.$$

When we talk about *pure growth rate*, we refer to the growth rate of the cells with the i-th genotype proceeding from cells with the same genotype. At the same way, the expression *net growth rate* of the i-th genotype refers to the quantity $l_i = a_i p_{ii} - d_i$.

About the stability of P_0

We have seen that P_0 is unstable if and only if

- $A + D > 0$ or
- $AD - BC < 0$.

Re-writing it in terms of the physical parameters, the first condition is equivalent to

$$a_0 p_{00} + a_1 p_{11} - (d_0 + d_1) > 0\,,$$

that is $a_0 p_{00} + a_1 p_{11} > d_0 + d_1$: biologically, it can be interpretated as the fact that, the new cells produced by the proliferation of cells with the same genotype are more that the cells that die. This is exactly what one could expect from a system of this kind, since it is exactly a generalization of a competition model where the populations grow in a logistic way.

With regard to the second condition, we see that it is equivalent to $(a_0 p_{00} - d_0)(a_1 p_{11} - d_1) - a_0 a_1 p_{01} p_{10} < 0$, that is the product of the net growth rates is smaller than the product of the mutation rates $a_0 p_{01}$ and $a_1 p_{10}$. Even if this result is not exactly a generalization of the result we obtain for the competition model (1.4), we notice that it is meaningful, because it means that cells do not disappear if mutation probabilities are big enough.

About the nonexistence of limit cycles

We have seen that a sufficient condition for the nonexistence of limit cycles is that $A + D < 0$. In terms of physical parameters, it means that $l_0 + l_1 < 0$, that is that the sum of the net growth rates is negative, then either one of them is negative or both. It means that there are more cells of genotype i that are dying than that are produced by cells with the same genotype, at least for one value of i.

About the existence of P_*

Regarding the existence of the positive steady state, we have proved that it always exists except when $A\beta - B\alpha < 0$, $D\gamma - C\delta < 0$ and $AD - BC > 0$ at the same time. Equivalently,

1. $(a_0 p_{00} - d_0)b_{01} < a_1 p_{10} b_{00}$;
2. $(a_1 p_{11} - d_1)b_{10} < a_0 p_{01} b_{11}$;
3. $(a_0 p_{00} - d_0)(a_1 p_{11} - d_1) > a_0 a_1 p_{01} p_{10}$.

At this point, for the first time the effects of competition appear, in particular in the first two conditions. We can see that competition factors fine-tune the relation between the net growth rate of the genotype i and the mutation rate of the other genotype j to the genotype i, for $i, j = 0, 1$.

About the stability of P_*

We have found two sufficient condition for the stability of P_*

- $\alpha\delta - \beta\gamma > 0$ or
- $AD - BC < 0$ and $A + D < 0$.

We have already analyzed the second condition talking about the stability of P_0, therefore let us focus on the first one. In terms of the physical parameters, the condition is equivalent to $b_{00}b_{11} - b_{01}b_{10} > 0$, that is $detB > 0$. Biologically, it means that the system achieves an equilibrium if the competition effects among cells with the same genotype are bigger than the ones between the two genotypes.

Chapter 4

Numerical simulations

In this chapter, some numerical simulations are performed for several values of physical parameters. Moreover, some statistical parameters are computed and compared. A biological interpretation of the simulations is provided.

In order to explore the behaviour of the system for $n > 1$, we run numerical simulations. In order to do it, we have written a Matlab code to solve the given system of ODEs. In the code, presented in *Appendix* A, an explicit Runge-Kutta method of order 2 is used. This method is classical and its derivation can be found in [8].

4.1 First simulation

In the first simulation, the parameters have the following characteristics:

- the time is measured in days;
- the carrying capacity is $K = 100000$ cells;
- the non-dimensional competition factors b_{ij} depend only on i and decrease linearly from 2 to 1, that is

$$b_{ij} = 2 - \frac{i}{n}\,, \qquad i, j = 0, \ldots, n\,; \tag{4.1}$$

- the mutation probability matrix P is tridiagonal: for $i = 2, \ldots, n-1$,

$$p_{ij} = \begin{cases} 0.8\,, & \text{if } j = i\,, \\ 0.1\,, & \text{if } j = i-1, i+1\,, \\ 0\,, & \text{otherwise;} \end{cases}$$

for $i = 0$,

$$p_{ij} = \begin{cases} 0.9\,, & \text{if } j = 0\,, \\ 0.1\,, & \text{if } j = 1\,, \\ 0\,, & \text{otherwise;} \end{cases}$$

for $i = n$,

$$p_{ij} = \begin{cases} 0.9\,, & \text{if } j = n\,, \\ 0.1\,, & \text{if } j = n-1\,, \\ 0\,, & \text{otherwise;} \end{cases}$$

In fact, one can think that the indeces i and j represent not specifical genotypes, but the numbers of mutations that they have respect to healthy cells. This choice of matrix P reflects a scenario where it is possible that, during every cell cycle, the replication can produce at most one mutation, which can make the genotype more or less different from the healthy genotype different.

The used initial conditions are

$$C_i(0) = \begin{cases} K-1\,, & \text{if } j = 0\,, \\ 1\,, & \text{if } j = 1\,, \\ 0\,, & \text{otherwise.} \end{cases}$$

It means that we are supposing that, at the beginning, there are only healthy cells, except one mutated cell with the first genotype.

To run the first simulations, we have used:

- growth rates a_i are equal for all genotypes and equal to 2;
- death rates d_i are equal for all genotypes and equal to 0.2;
- $h_i = 0$ for all $i = 0, \ldots, n$ and, as a consequence, $g_i = (a_i - d_i)/d_i = 9$ for all $i = 0, \ldots, n$.

Next up, the graphical results of some simulation are shown.

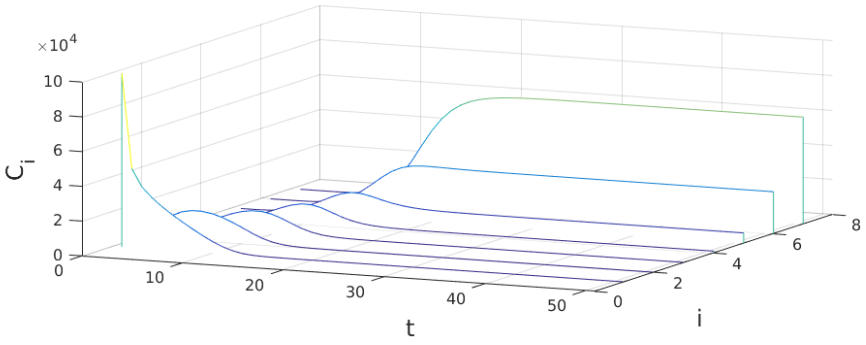


Figure 4.1: Simulation for $n = 6$.

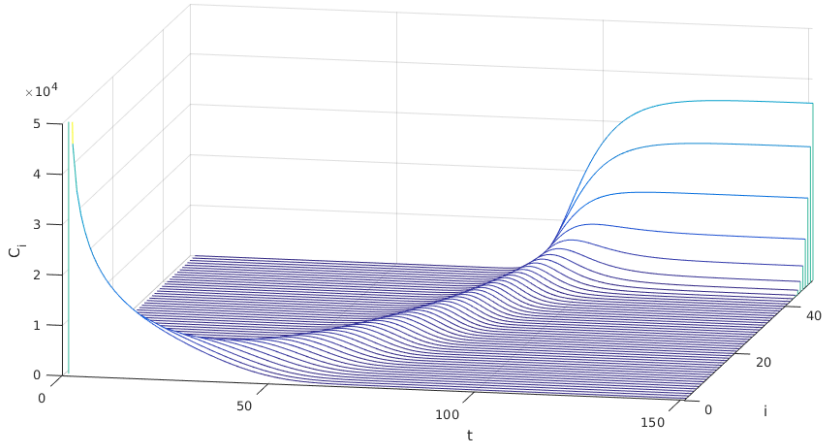


Figure 4.2: Simulation for $n = 50$.

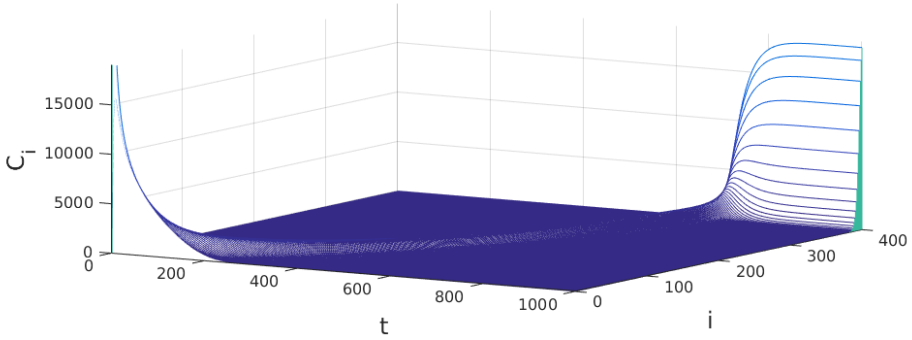


Figure 4.3: Simulation for $n = 400$.

As one can see, the dynamics in all cases is the same: independently of the number of possible genotypes, the number of healthy cells decreases and the formation of a travelling wave takes place, until arriving to the n-th genotype; at this point, it seems that the system is getting closer and closer to a steady state, since, apparently, there are no relevant changes in the number of cells for each genotype. The accumulation of cells close to the last genotype can be considered a numerical artefact, due to considering a small number of possible genotypes. In fact, in a cell there are so many possible mutations which can occur that it is impossible to arrive to a "last" genotype in a lifetime.

4.2 Second simulation

Let us keep the same values for all the parameters, except for B and a_i. Just for a computational reason, we set $a_i = 10$. Moreover, we set $b_{ij} = 1$ for all $i, j = 0, \dots, n$, that is all the genotypes compete equally for the resources. By this choice of the competition matrix, we obtain the following result:

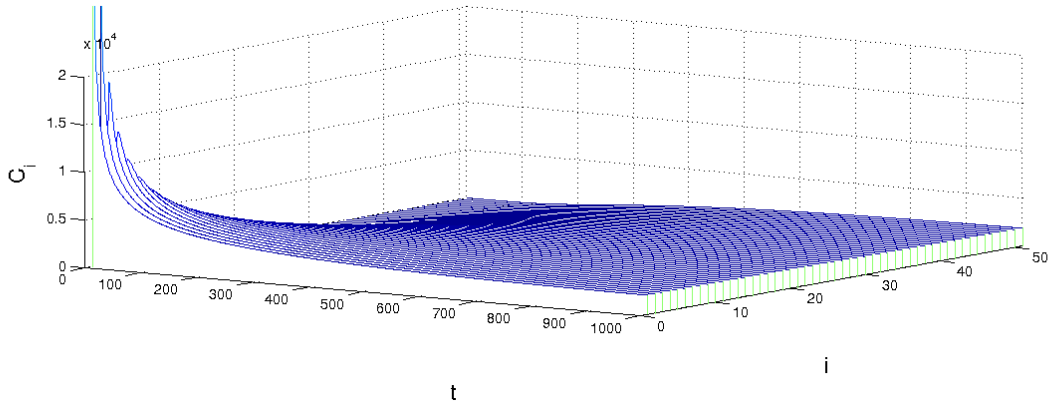


Figure 4.4: Second simulation for $n = 50$.

As we can see, in this case, the system tends to an equilibrium where all the genotypes are represented by the same number of cells. One could expect this behaviour since proliferation and death rates are equal for all the genotypes, P is symmetrical and they compete on equal terms for the resources.

4.3 Third simulation

Let us change again B, keeping $a_i = 10$ for all $i = 0, \dots, n$. We set

$$b_{ij} = 1 + \frac{i}{n} \qquad \text{for all } i, j = 0, \dots, n\,;$$

that is the competition factors increase linearly from 1 to 2 as i increases. By this choice of the competition matrix, we obtain the following result:

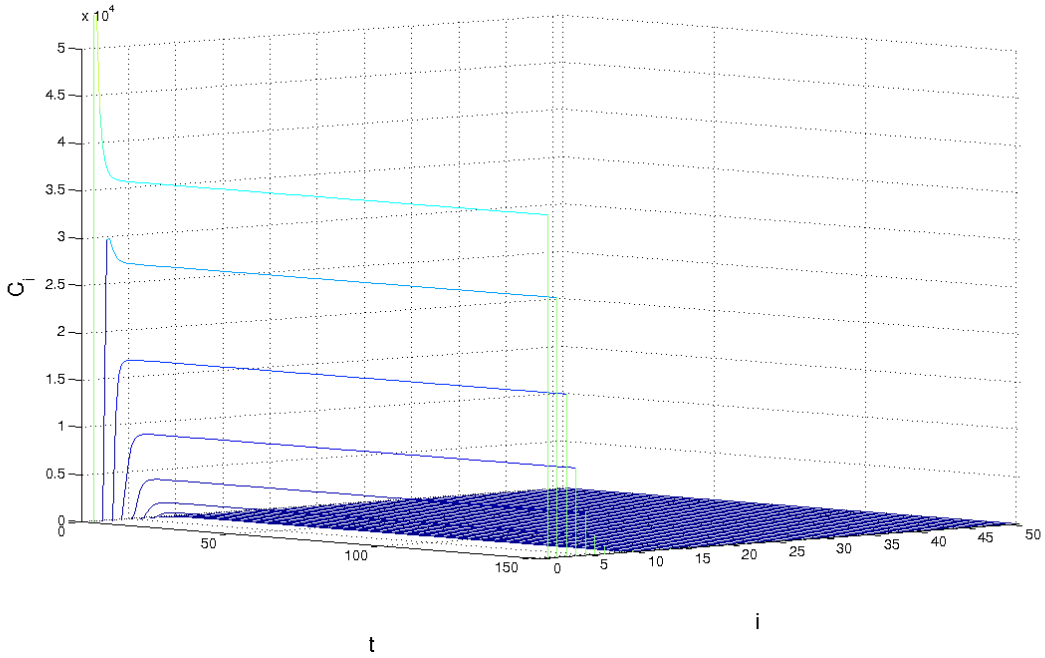


Figure 4.5: Third simulation for $n = 50$.

As the plot shows, in this case, the healthy genotype is predominant and the more mutated a genotype is, fewer are the cells having that genotype.

4.4 Fourth simulation

Going back to the competition matrix defined by (4.1), it is possible now to run a more realistic simulation: in fact, biologists observed that cancer cells proliferate faster than healthy cells and the proliferation rate depends, above all, on the degree of differentiation. For exemple, to suppose that

$$a_i = 4 - 2e^{-i} \qquad \text{and} \qquad d_i = 0.2 - 0.1e^{-i}\,. \tag{4.2}$$

For this values of a_i and d_i, we obtain that

$$g_i = \frac{a_i - d_i}{d_i} = 19\,.$$

In this case, using the same initial condition as before, the numerical simulations produce the following results:

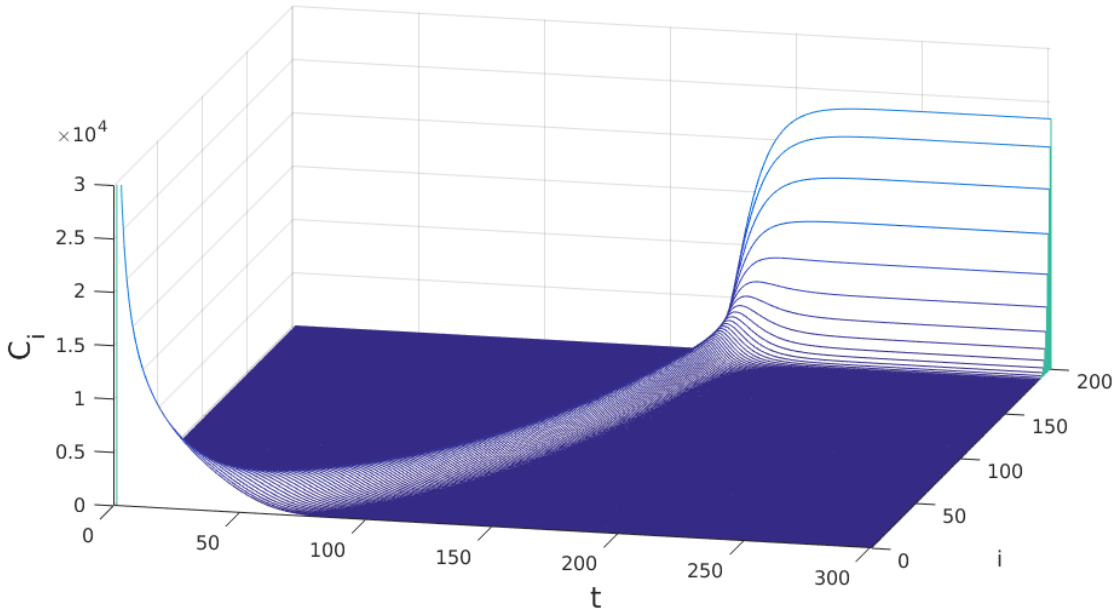


Figure 4.6: Fourth simulation for $n = 200$.

As we can see, in this case a travelling wave is forming again but it travels faster than in the previous case, when the proliferation and death rates were constant. One could expect this behaviour, since these parameters now are increasing as i increases.

4.5 Fifth simulation

Just to verify that the model is valid, we can run a last simulation, where we are assuming that:

- growth rates a_i are equal for all genotypes and equal to 1;
- death rates d_i are equal for all genotypes and equal to 1.5;
- $h_i = 0$ for all $i = 0, \dots, n$ and, as a consequence, $g_i = (a_i - d_i)/d_i = -0.25$ for all $i = 0, \dots, n$;
- B is defined as in (4.1).

One can expect that, in this case, at least for some initial condition, the solution tends to zero as time grows. This is exactly what happens, for instance, for the following initial condition:

$$C_i(0) = \begin{cases} K/2\,, & \text{if } j = 0\,, \\ 1\,, & \text{if } j = 1\,, \\ 0\,, & \text{otherwise.} \end{cases}$$

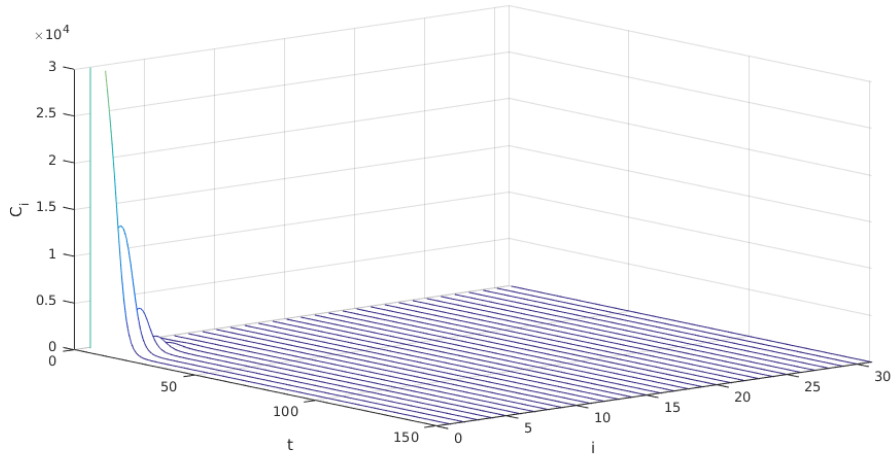


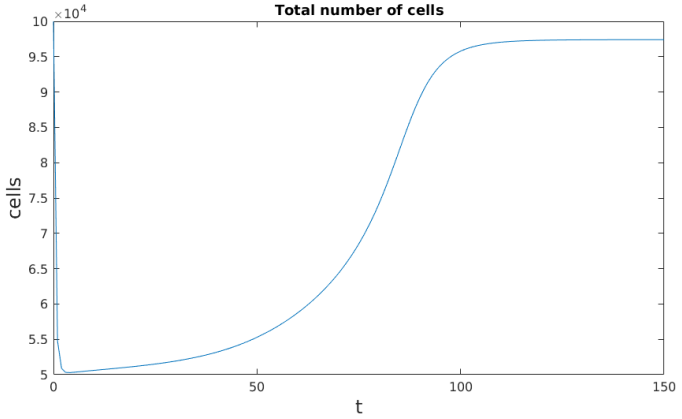
Figure 4.7: Fifth simulation for $n = 30$.

4.6 Comparison

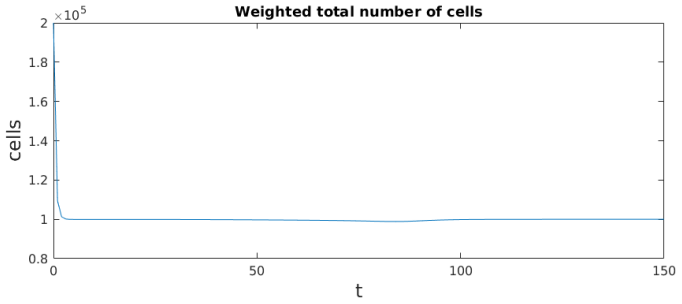
Let us compare now some statistical parameters in all the mentioned cases, that is:

- the total number of cells $T(t) = \sum_{i=0}^{n} C_i(t)$;
- the weighted total number of cells $W(t) = \sum_{i=0}^{n} b_{i0} C_i(t)$;
- the mean genotype $M(t) = \sum_{i=0}^{n} iC_i(t) / \sum_{i=0}^{n} C_i(t)$;
- the variance of the genotype $V(t) = \sum_{i=0}^{n} i^2 C_i(t) / \sum_{i=0}^{n} C_i(t) - M(t)^2$.

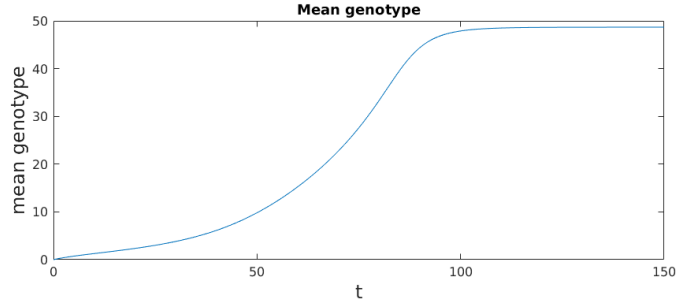
It is done, in this case, for $n = 50$.



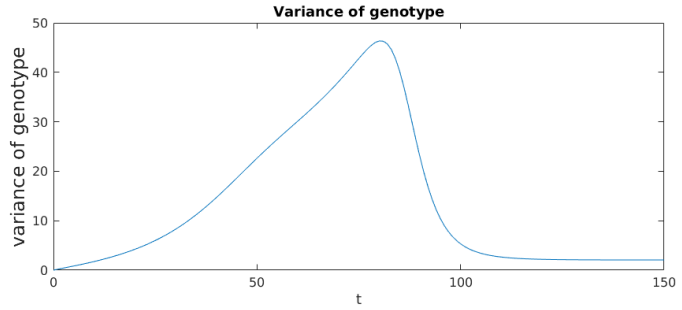
(a)



(b)

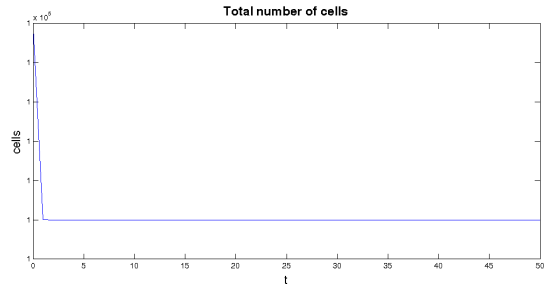


(c)

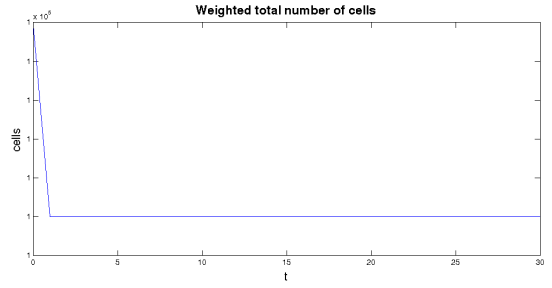


(d)

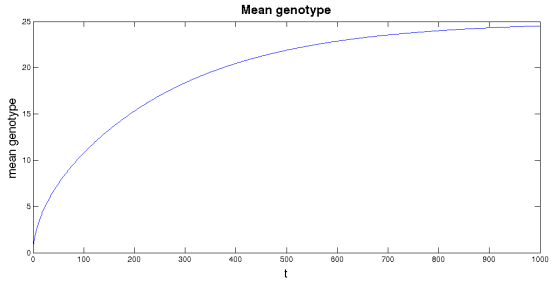
Figure 4.8: Statistical parameters for $n = 50$, physical parameters as in *Section* 4.1.



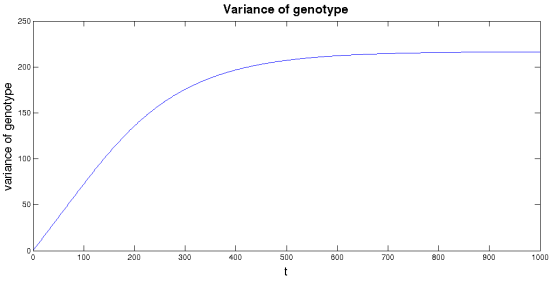
(a)



(b)

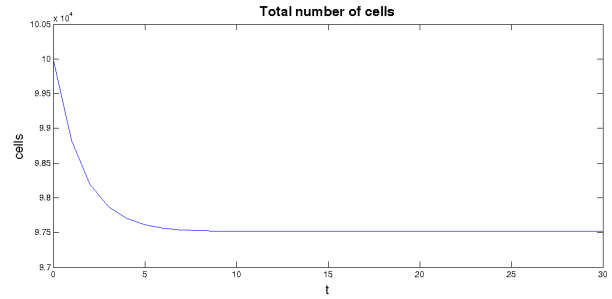


(c)

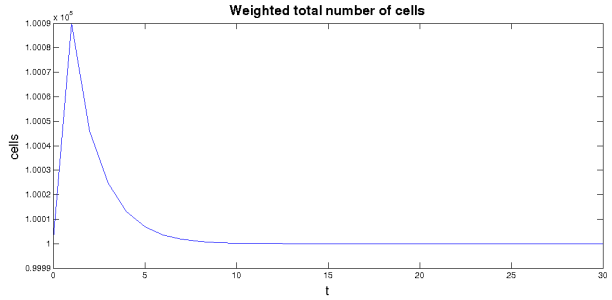


(d)

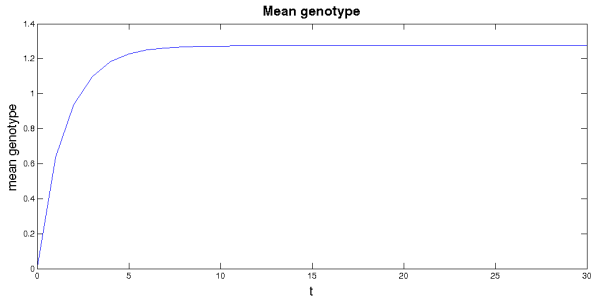
Figure 4.9: Statistical parameters for $n = 50$, physical parameters as in *Section* 4.2.



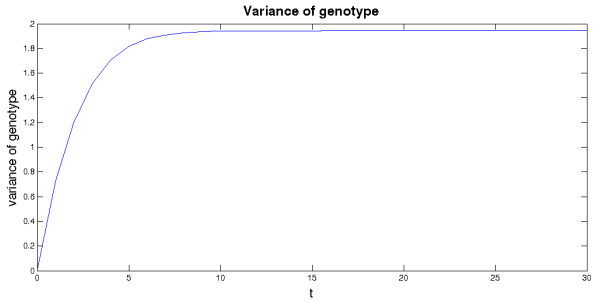
(a)



(b)

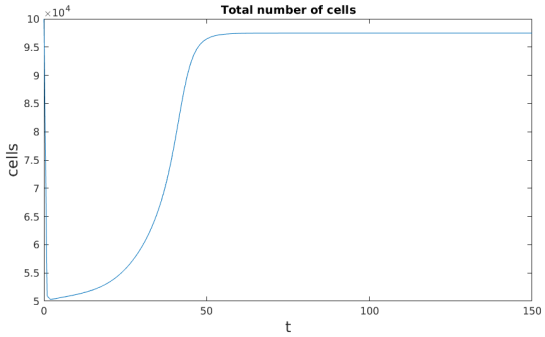


(c)

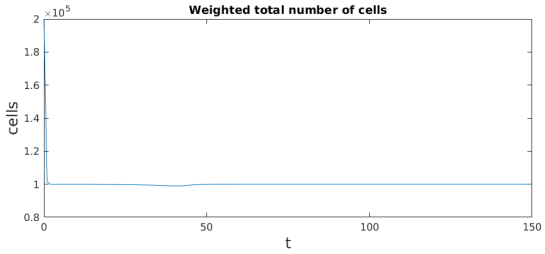


(d)

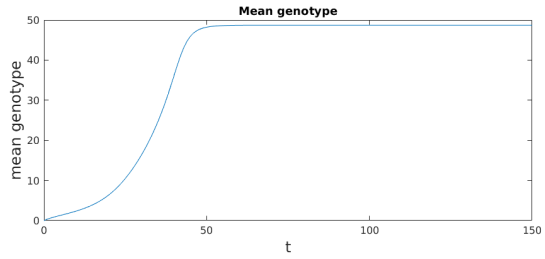
Figure 4.10: Statistical parameters for $n = 50$, physical parameters as in *Section* 4.3.



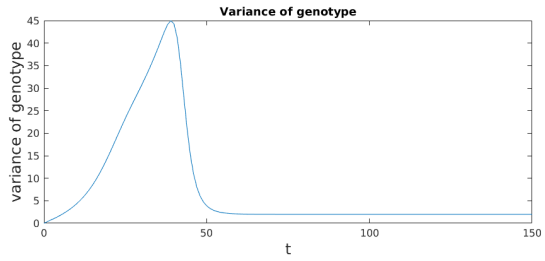
(a)



(b)

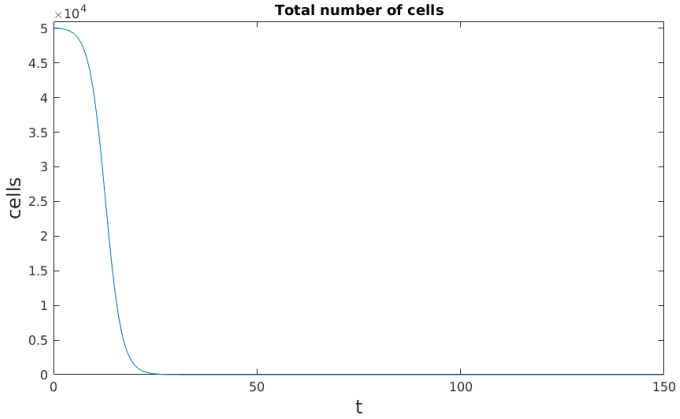


(c)

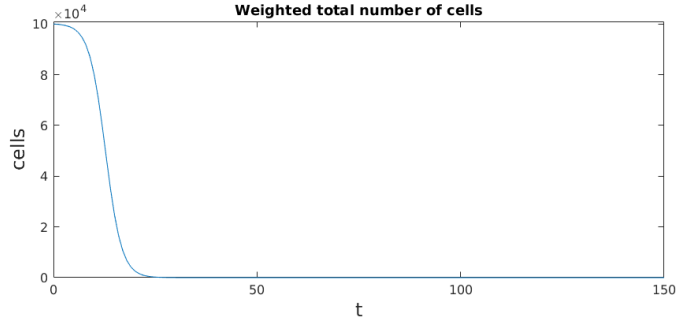


(d)

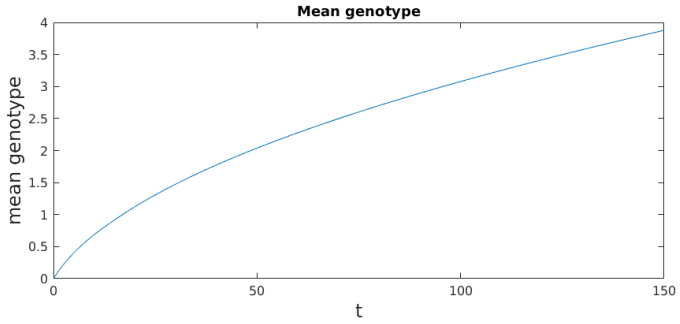
Figure 4.11: Statistical parameters for $n = 50$, physical parameters as in *Section* 4.4.



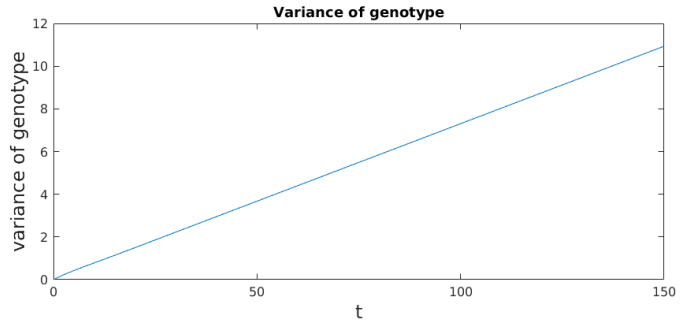
(a)



(b)



(c)



(d)

Figure 4.12: Statistical parameters for $n = 50$, physical parameters as in *Section* 4.5.

As we can see, in the first and in the fourth cases (*Figures* 4.8-4.11) the system has similar behaviours: the only difference is that in the fourth case the system gets the equilibrium faster, that is equivalent to the fact that the travelling wave travels faster. Anyway, in both cases the mean genotype grows and gets closer to the last genotype. The variance, instead, grows until reaching a maximum for a certain time t^*, which represents the instant where there is a big variety of cells in the environment, and then it decreases until getting a value close to 0.

In the fifth case (*Figure* 4.12) it is evident that the total number of cells $T(t)$ tends to 0 as $t \to \infty$, that is all genotypes are disappearing. Instead, the value of the variance increases and we guess that it is because, from a certain time on, all genotypes disappear uniformely.

Comparing then the first simulation with the second and the third ones (*Figures* 4.8-4.9-4.10), we can observe that a change in the values of the competition factors b_{ij} change qualitatively the behaviour of the system:

- if the competition factors are equal, the system tend to an homogenuous distribution of genotypes, at least when proliferation and death rates are the same for all the genotypes;

- although, depending on the relation among these factors, the system tends to the predominance of a certain genotype, which seems to be the genotype j with the smallest value of b_{ij}, in the case of b_{ij} independent of i.

  Moreover, as we can observe, in all the cases, either $W(t) \to 0$ or $W(t) \to K$ as $t \to K$, which is exactly what occurs in a classical competition model.

Conclusions

In this thesis, we have studied the dynamics of a discrete finite genotype space model of cancer competitive evolution. The model allows us to predict the qualitative behaviour of a system consisting on a variety of genotypes co-existing in the same environment. We have assumed that:

- there are $n+1$ possible genotypes, from 0 to n, where the index 0 is associated to the normal genotype;
- for all genotypes, the growth rates follow the logistic growth;
- the genotypes compete among them for limited resources (food, oxygen, ...);
- random mutations can occur during cell replication process.

Denoting the number of cells with the i-th genotype at time $t \in [0, \infty[$ by $C_i(t)$, starting from the well-known competition model (1.3), we have obtained the following system of ordinary differential equations that model the biological system:

$$\dot{C}_i = \sum_{j=0}^{n} \left(p_{ji} a_j C_j \left(1 - h_j \frac{\sum_{k=0}^{n} b_{jk} C_k}{K} \right) \right) - d_i C_i \left(1 + g_i \frac{\sum_{k=0}^{n} b_{ik} C_k}{K} \right),$$

for $i = 0, \ldots, n$. In the previous equation, a_i represents the proliferation rate of i-th genotype, d_i its death rate, p_{ij} the probability that a cell with genotype i produces a daugther cell with genotype j, b_{ij} the competition factor between the genotypes i and j and h_i and g_i fine-tune the effects of competition respectively on proliferation and on death. All the parameters are positive, except p_{ij} which can be zero if $i \neq j$.

In order to study the model analitically and to run simulations, we have assumed that $h_i = 0$ for all i, because it garantees that the positivity of the model is preserved: it means that, if an orbit starts from a point $(C_i(0))_i$ with $C_i(0) > 0$ for all $i = 0, \ldots, n$, then $C_i(t) > 0$ for all $t > 0$ and for all $i = 0, \ldots, n$.

The analytical study of the model for $n = 1$, that is for just two population (healthy cells and cancer cells), shows that there are at most two possible non-negative fixed

points for the model: one of them is the origin P_0 and the other one P_* is strictly positive. Depending on the value of the parameters, the second fixed point exists or it does not. Since all the orbits, from a certain time on, are trapped in a compact region of the first quadrant, due to Poincaré-Bendixson theorem, they can tend to one of the two equilibria or to a limit cycle, if it exists (we have not been able to exclude its existence in all cases). A biological interpretation of the system have led us to the conclusion that the new model represents actually a generalization of the competition model (1.3) taken as starting point.

After running simulations for $n > 1$ we can conclude that, in the general case:

- performing reasonable changes in the values of the proliferation and death rates does not affect the evolution of the system qualitatively: in any case, we assist to the formation of a travelling wave, whose speed is affected by the aforesaid rates;
- if the proliferation rates are smaller than the death ones, the total population tends to zero as $t \to \infty$;
- changing the values of b_{ij}, the dynamics is affected in different ways: if the competition matrix is constant, the final distribution of the genotypes is homogenous; otherwise, some genotypes are more predominant than other ones;
- for all simulations where the system tends to a positive fixed point, whatever the values of the different parameters are, the total weighted population tends to carrying capacity K.

Therefore, we have observed that, for real life situations with a predominance of healthy cells at the beginning, it is possible to model different scenarios where cancer is preponderant or not (just changing the competition matrix B). However, some improvements can be done to simulate some situation that are more realistic. For example:

- the parameters h_i can be considered strictly positive. This would mean that the competion has effects on the proliferation as well. Anyway, these parameters should be very small in order to preserve the positivity of the model at least in a compact region of the positive region of the phase space;
- one can suppose that cells can mutate into any other genotype, not just into the previous and the next ones. For example, it can be done supposing that healthy cells have low probability of mutating into cancer cells and cancer cells have a higher probability to produce other types of cancer cells;
- to implement random mutations, instead of using a deterministic model, one could use a stochastic model, as proposed, for instance, in [1]. However, this model would be harder to simulate and would give probably the same results: in fact, we are working with a huge number of cells and, hence, averaging over a probabilistic model would have the same outcomes as our deterministic model;

- since a same genotype can have different phenotypical expressions, there are so many phenotypes that this spectrum can be considered continuous. This leads to a continuous formulation of the model (2.8), using integro-partial differential equations, analyzed, for instance, in [5] and [6]. The outcomes of these works can be read as generalizations of our model, and, vice-versa, our model can be seen as a discretization of the continuous ones.

Appendix A

The code

In this appendix, the code used to simulate numerically the model is presented.

A.1 Implementing the vector field

First of all, we need to define the vector field of the dynamical system described by (2.8).

```
function f=vecfield(x)

%INPUT: an n+1 dimensional vector x.
%OUTPUT: the vector field evaluated at t.

%It is necessary to define the parameters of the system as global
%in order to  introduce them in a following script.

global a
global P
global d
global g
global B
global k

n=length(x)-1;

%The vector field is calculated.
f=zeros(n+1,1);
for i=0:n
    t=0;
```

```
    m=0;
    for l=0:n
        t=t+B(i+1,l+1)*x(l+1);
        m=m+P(l+1,i+1)*a(l+1)*x(l+1);
    end
    t=(t*g(i+1)/k + 1)*d(i+1)*x(i+1);
    f(i+1)=m-t;
end
```

A.2 Implementing the Runge-Kutta method

The explicit Runge-Kutta method used to run numerical simulations has order 2.

```
function Y=RK2(tfin,y0,N)
%INPUT: final time tgin
%               initial condition y0
%               number of iterations N

n=length(y0)-1;

h=tfin/N; %time step

%The matrix of the numerical solution is computed
Y=zeros(n+1,N+1);
Y(:,1)=y0;
x=y0;
for j=1:N
        k1=vecfield(x);
        k2=vecfield(x+h/2*k1);
        x=x+h*k2;
        Y(:,j+1)=x;
end
```

A.3 The script

In this last script, after the inizialization of the parameters, the numerical solution is computed and plotted and so are the statistical parameters (total population, total weighted population, mean genotype and variance of the genotype).

```
n=50;       %number of mutated genotypes
```

```
tfin=150;  %final time
N=tfin*10; %number of iterations

%The parameters are set as global
global a
global P
global d
global g
global B
global k

%carrying capacity
k=100000;

%mutation probability matrix
P=zeros(n+1,n+1);
P(1,1)=0.9;
P(1,2)=0.1;
P(n+1,n)=0.1;
P(n+1,n+1)=0.9;
for i=2:n
    for j=1:(n+1)
        if j==i-1
           P(i,j)=0.1;
        elseif j==i+1
            P(i,j)=0.1;
        elseif j==i
            P(i,j)=0.8;
        end
    end
end

%proliferation rates
a=2*ones(n+1,1); %a constant
%  for i=1:(n+1) %a exponential
%       a(i)=4-2*exp(-n+1);
%  end

%death rates
d=0.2*ones(n+1,1); %d constant
```

```
% for i=1:(n+1) %d exponential
%     d(i)=0.2-0.1*exp(-n+1);
% end

%competition effects on death
g=zeros(n+1,1);
for i=1:(n+1)
    g(i)=a(i)/d(i) - 1;
end

%competition parameters
B=ones(n+1,n+1);
for i=1:(n+1)
    for j=1:(n+1)
        B(i,j)=2-(1/n)*(i-1);
    end
end

%initial condition
c0=zeros(n+1,1);
c0(1)=k-1;
c0(2)=1;

%The numerical solution is computed
Y=RK2(tfin,c0,N);

%Plot of the numerical solution
Z=Y(:,1:10:N+1);
figure
waterfall(Z)

%Computation of the statistical parameters
time=0:tfin;
tot=zeros(tfin+1,1);
tot_w=zeros(tfin+1,1);
mean=zeros(tfin+1,1);
variance=zeros(tfin+1,1);
for i=1:tfin+1
    tot(i)=sum(Z(:,i));
    tot_w(i)=B(:,1)'*Z(:,i);
```

```
    mean(i)=(0:n)*Z(:,i)/tot(i);
    variance(i)=((0:n).^2)*Z(:,i)/tot(i) - mean(i)^2;
end
figure
plot(time,tot)

title('Total number of cells')

figure
plot(time,tot_w)

title('Weighted total number of cells')

figure
plot(time,mean)

title('Mean genotype')

figure
plot(time,variance)

title('Variance of genotype')
```

Bibliography

[1] N. Bellomo, M. Delital, *From the mathematical kinetic, and stochastic game theory to modelling mutations, onset, prograssion and immune competition of cancer cells*, Physics of Life Reviews, Vol. 5, No. 4, 2008, pp. 183-206.

[2] A. C. Buira, S. Gómez, *Mathematical modelling viral and cancer evolution*, Bachelor Thesis, 2017.

[3] D. R. Curtiss, *Recent extensions of Descartes' rule of signs*, Annals of Mathematics, Vol. 19, No. 4, 1918, pp. 251–278.

[4] A. Korobeinikov, K. E. Starkov, P. A. Valle, *Modeling cancer evolution: evolutionary escape under immune system control*, IOP Conf. Series: Journal of Physics, Vol. 811, 2017.

[5] D. Masip, A. Korobeinikov, *A continuous phenotype space model of cancer evolution*, IOP Conf. Series: Journal of Physics, Vol. 811, 2017.

[6] D. Moreno, *Mathematical modelling cancer evolution*, Bachelor Thesis, 2018.

[7] J. D. Murray, *Mathematical Biology*, Third edition, Springer, 2001.

[8] A. Quarteroni, R. Sacco, F. Saleri, *Numerical mathematics*, Springer, 2000.

[9] R. Ruddon, *Cancer biology*, Oxford University Press, 2007.

[10] S. H. Strogatz, *Nonlinear dynamics and chaos*, First edition, Perseus Books Publishing, 1994.

Acknowledgements

First of all, I am grateful to my supervisor, prof. Andrei Korobeinikov, for all the help he has provided me during the last year and for all the lessons he taught me. I am also grateful to the *Centre de Recerca Matemàtica* (CRM) for providing me with a place to work at and for giving me the opportunity to do a research experience.

Secondly, I want to thank my institutional supervisor, prof. Fabio Marcuzzi, for letting me study this particular subject in a foreign country under the supervision of a professor he did not know and for trusting me. I also want to thank prof. Àngel Calsina Ballesta for the mathematical support and for helping me when I could not see the light.

Then, thanks to those people who contributed to my education and, indirectly, to my personal growth, to Carolina Dutti, to Teresa Macolino, to Letizia Saponaro and, in particular, to Rosa Miranda, whose memory will always be in my heart and in my mind.

Finally, I would like to express my gratitude to all those people who trust me, to Bianca, Marco, Elisabetta, Francesca and all my closest friends. But my biggest thanks goes to my parents Rosario and Pina and to my sister Simona, who always stay by my side, for better or for worse, and have always believed in me, also when I could not do it.