ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

A statistical mechanics approach to cancer dynamics: a model for multiple myeloma bone disease

Relatore:
Prof. Armando Bazzani

Presentata da:
Maria Elena Corrias

Correlatori:
Prof. Amos Maritan
Dott. Marco Formentin
Prof.ssa Maria Giuliana Vannucchi

Anno Accademico 2017/2018

To whom is near, to whom is far,
and especially to whom cannot come back anymore

E tutto vidi nella creazione di Dio
Ma ad ogni cosa che esiste
A tutto quel che accade sotto il sole
Un senso l'uomo non riesce a dare
Lì sopra gli uomini si affaticano
Senza poter trovare
E il sapiente che dice di sapere
Neppure lui ha trovato

Qohèlet

Sommario

L'utilizzo di modelli matematici sta assumendo un ruolo sempre più centrale nella ricerca oncologica. La complessità del cancro ha stimolato gruppi di ricerca interdisciplinare nello sviluppo di modelli quantitativi per rispondere alle numerose domande aperte che riguardano l'insorgenza, la progressione, la diagnosi, la risposta al trattamento terapeutico e l'acquisizione della resistenza ai farmaci dei tumori. La varietà di approcci matematico-fisici ben si adatta allo studio di una materia così eterogenea. In questo lavoro presentiamo innanzitutto gli aspetti biologico-clinici che caratterizzano il cancro, per poi introdurre i modelli che sono stati utilizzati per comprenderli. Abbiamo preso in considerazione il caso del mieloma multiplo, una neoplasia che colpisce le plasmacellule. In particolare proponiamo un modello matematico per lo studio della patogenesi delle lesioni ossee causate dal mieloma. L'insorgere di questo tumore rompe l'equilibrio fisiologico del tessuto osseo, causando un aumento dell'attività degli osteoclasti ed una diminuzione dell'attività degli osteoblasti, fenomeni che, combinati, comportano le caratteristiche fratture. Abbiamo optato per un approccio di tipo ecologico, in cui i diversi tipi di cellule sono considerati come specie interagenti in meccanismi di cooperazione o sfruttamento. Questo fenomeno è stato modellizzato all'interno della classe degli Interacting Particle Systems, che sono sistemi di processi di Markov localmente interagenti. Abbiamo inizialmente studiato il caso dell'osso sano per poi passare a quello in cui sono presenti le cellule del mieloma. Infine, abbiamo svolto simulazioni per delineare l'evoluzione nel tempo delle specie cellulari. Abbiamo riservato una particolare attenzione alla definizione dei parametri del modello: non solo essi ci permettono di riprodurre diversi stadi e forme del mieloma, ma possono descrivere l'intervento terapeutico sul tumore, costituendo un nuovo strumento per la ricerca oncologica.

Abstract

Mathematical modelling has recently been gaining crucial importance in oncological research. The complexity of cancer has stimulated interdisciplinary research to address the many open questions about tumor initiation, progression, diagnosis, treatment response and drug resistance. The variety of mathematical and physical approaches is well-suited to the study of such a heterogeneous topic as cancer. In this thesis we firstly present the biological and clinical aspects that characterize cancer. Successively, we describe models that were developed to understand them. We decided to focus on the case of multiple myeloma, a neoplasia that affects plasmacells. In particular, we propose a mathematical model for the study of the pathogenesis of multiple myeloma bone disease. In fact, the onset of this tumor disrupts the physiological homeostasis of bone tissue. It enhances the osteoclasts activity and suppresses the osteoblast activity: these phenomena combined lead to bone lesions distinctive of myeloma. We opt for an ecological approach, where different kind of cells are considered as species, involved both in cooperative and exploitative interactions. This phenomenon has been modelled in the framework of the interacting particle systems, a class of systems composed by locally interacting Markov processes. Firstly, we study the healthy bone case, then the one in which myeloma cells are present. Lastly, we perform simulations to reproduce the time evolution of cellular species. We pay special attention to the model parameter setting: they allow us to simulate different stages and forms of myeloma and, moreover, they can describe the therapeutic intervention.

CONTENTS

INTRODUCTION

LEGGE DI WOLF
SULLA PIANIFICAZIONE
Un buon posto per cominciare
è dove sei.

During the past two decades, the study of a complex problem such as cancer has stimulated the interest of physicists, mathematicians, engineers and chemists. The encounter of different approaches is able to create a breeding ground ideal for new ideas for understanding and tackling cancer to blossom. This is especially crucial considering that, nowadays, cancer represents one of the most prominent causes of death.

Physics aims to provide quantitative models, in order to interpret and integrate clinical data. These have lately become more and more available, thanks to new techniques for DNA sequencing, epigenetic classification and so on.

Our thesis places itself in this framework, proposing a statistical mechanics model for the pathogenesis of multiple myeloma-induced bone disease. We opted for an ecological approach: simply put, we considered cancer as a result of interactions between cellular species. Successively, we

developed an interacting particle model and simulated it as a succession of points of a Poisson process over a lattice.

Scientists bring to interdisciplinary research their own knowledge, their methods but, unfortunately, also their language. Indeed, the major problem for experts from different fields is the ability to communicate. Because of this, a priority for us has been to maintain an impartial interdisciplinary approach towards the topic, discussing it both from a clinical-biological and from a mathematical-physical point of view.

In order to achieve this, we tried to produce a piece of work accessible to readers coming from various research fields. On the one hand, we did not go into in-depth mathematical details, that could prove disorientating to a biologist. On the other hand, we introduced two chapter of biomedical subject, alongside a glossary of biological terms, in order to facilitate the reading for physicists. We decided to alternate the chapters in order to highlight the process through which the toolbox of a physicist, who studies a real phenomenon, suits to the features of the phenomenon itself.

We commence Chapter 1 with an introduction, from a biological point of view, of the main features of cancer. Firstly, we are going to describe the hallmarks, i.e. the distinctive traits, of cancer, defined in 2000 in a well-known paper by D. Hananan and R. A. Weinberg [1]. We will deal with tumor progression, from the initial genome mutation to the diagnostic methods, working through the description of aspects such as angiogenesis and metastasis.

In Chapter 2, we are going to do a review of the approaches to cancer of mathematics and physics. First of all, we are going to try to explain how physics can be linked to oncology and especially how physics can be useful for both diagnosis and therapy. Later, following the structure of Chapter 1, we are going to introduce models and techniques applied by mathematical oncology. In particular, we will focus on some very useful tools such as branching processes.

In Chapter 3, we are going to focus on multiple myeloma, a neoplasia of plasmacells, often associated to lytic bone lesions. We will firstly introduce the physiology of healthy bone tissue and then we will describe the effect that myeloma has on it. Particularly, we are going to list the molecu-

lar pathways that lead to the pathogenesis of multiple myeloma-induced bone disease.

In Chapter 4, which is of a purely mathematical nature, we will dwell on the toolbox needed to develop our model. Having introduced the reader to the main concepts of Markov chains and Poisson processes, we will analyse in details two of examples interacting particle systems: the Curie-Weiss model and the Voter Model.

In Chapter 5 we are going to describe our model of pathogenesis of multiple myeloma-induced bone disease. Before doing this, we will briefly introduce the evolutionary game theory approach to the same problem. We will show step-by-step how we have created the model and how we designed the simulations.

In the following two sections, we will show the code we developed and the figures we obtained.

Two appendices will follow : the first one contains details about probability theory, that can be useful for understanding Chapter 4; the other one looks into the relation between evolutionary game theory and population dynamics.

To conclude, a glossary of the lesser-known biomedicals terms used was included.

CHAPTER 1

THE BIOLOGY OF CANCER

OSSERVAZIONE DI AVERY SUL TEATRO
Non importa se cadi per terra,
purché quando ti alzi prendi qualcosa
dal pavimento.

Cancer is a genetic disease of multicellular organism, consequence of the accumulation of somatic mutations. It is characterized by a breakdown of cooperation between individual cells [2] and by an abnormal cell growth [3]. Thus, all cancer types share a common pathogenesis [4]. It can be both benign, if it is localized *in situ*, or malignant, if it is invasive and causes metastasis. Cancer begins with a single genetically altered cell and then invades the adjacent tissue through clonal expansion. Cancer is the result of a process of Darwinian evolution among cell populations embedded in their environment. As a Darwinian process, its development is marked by two feature: firstly, the continuous acquisition of heritable genetic variations in individual cells by more-or-less random mutations and, secondly, the natural selection acting on the resultant phenotypic diversity.

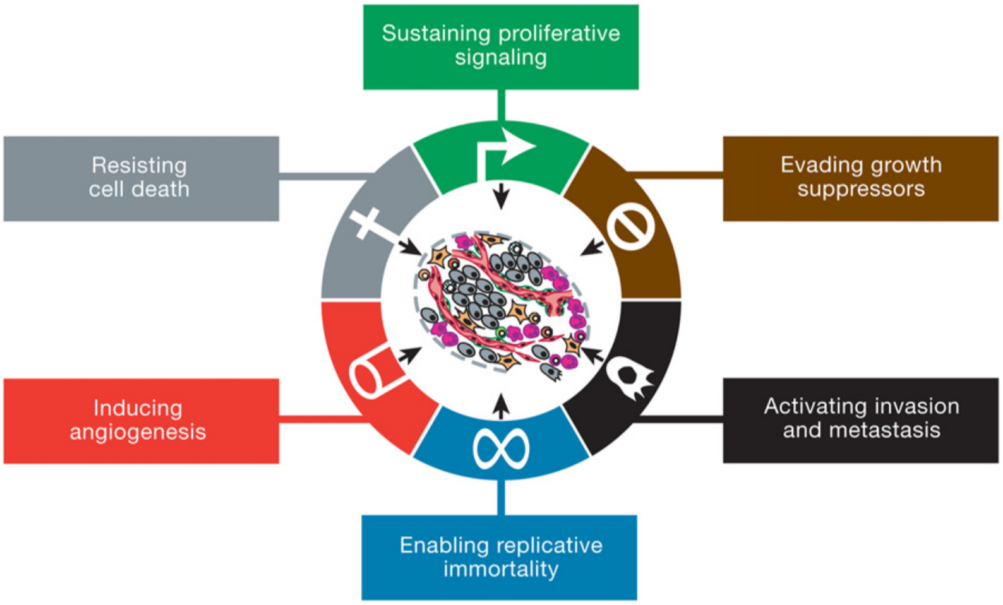


Figure 1.1: **Acquired capabilities of cancer.**

1.1 The hallmarks of cancer

D. Hanahan and R. A. Weinberg, at the beginning of this century, have tried [1] to find a logical framework for understanding cancer, identifying a small number of traits (molecular, biochemical and cellular) that all cancer types have in common. These functional capabilities are acquired during the multistep process that leads healthy cells to become malignant ones, called *tumorigenesis*.

Between all the unessential traits, they highlighted (Fig. 1.1):

- **Self-sufficiency in growth signals**
  In order to proliferate, a healthy cell requires signals that stimulate mitosis. Many oncongenes can simulate the effects of these growth signals: thus, the proliferation of malignant cells is independent from exogenous stimulation, but is self-sustained, other than chronic and unregulated.

- **Insensitivity to anti-growth signals**
  As for what concern the reproduction, also for cellular quiescence

are required antigrowth signals, that, blocking proliferation, guarantee tissue homeostasis. It has been discovered that cancer cells are insensitive to these growth-inhibitory signals: they disregard the programs that negatively regulate cell proliferation. When tumor occurs, many tumor suppressors, ordinarily responsible for limiting cell reproduction, become deactivated.

- **Tissue invasion and metastasis**
  Malignant cells become detached from the primary tumor and invade adjacent and also, flowing through blood vessels, distant tissues. New masses, called *metastasis*, are the cause of 90% of deaths related to cancer. Metastasis share with the primary tumor mass the same traits described here.

- **Sustained angiogenesis**
  Oxygen and nutrients are essential for the cell survival and they are carried throughout the body by circulatory system: the cell must reside within about 100 mμ of a capillary. Physiologically, after organogenesis is ended, the process of growth of a blood vessel, termed *angiogenesis*, is transitory and carefully regulated. By contrast, when cancer occurs, an angiogenic switch is almost always activated and remains on, causing normally quiescent vasculature to continually sprout new vessels.

- **Resisting cell death**
  Cancer cells acquire resistance toward apoptosis, that is programmed cell death.

- **Limitless replicative potential**
  The characteristics that we have already listed show that the growth program of a malignant cell is decoupled from signals in its environment. Nevertheless, disruption of cell-to-cell signaling is not enough in order to ensure expansive tumor growth: all types of mammalian cells are provided for a cell-autonomous program that limits their multiplication and initiates the process of senescence. In neoplastic

disease, some tumor suppressors are, as we have already said, deactivated, making these cells able to multiply. Cancer cells do not have a limited number of successive cell growth-and-division cycles: this trait is termed *immortalization*.

In 2011 the same authors have added [5] some other features (Fig. 1.2) to the hallmarks of cancer found in 2000.
First of all, they consider that the acquisition of the six hallmarks above described can be reached depending on two enabling characteristics:

- **Genome instability and mutation**
  Normal cells become malignant gaining, step by step, the capabilities that we are describing and this process happens thanks to a succession of random alterations in the genome. Some mutant genotypes confer selective advantage on subclones of cells, enabling their outgrowth and eventual dominance in local tissue environment. Also clonal expansion is usually triggered by the acquisition of a mutant genotype. It should be remembered that the stability of the DNA is ensured by an extremely efficient genome maintenance system. Nevertheless, cancer can damage the components of this machinery, making the mutability possible.

- **Promoting inflammation**
  Recent techniques have shown that every neoplastic lesion contains immune cells, demonstrating that tumor masses trigger the immune system response. Studies on this issue sustain that tumor can evade immune destruction, even if immune system, after it has recognized the cancer, tries to eradicate it. Tumor-associated inflammatory response, carried out especially by innate immune system cell, can, paradoxically, promote tumor progression. In fact, inflammation can contribute to multiple hallmark capabilities by supplying bioactive molecules to the tumor microenvironment, including growth factors that sustain proliferative signaling, survival factors that limit cell death, proangiogenic factors, extracellular matrix-modifying enzymes that facilitate angiogenesis, invasion, and metastasis.

In the second place, they defined two other hallmarks of cancer:

- **Avoiding immune destruction**
  An outstanding problem is the attempt of the immune system to resist or eradicate formation and progression of incipient neoplasia, already formed tumor masses and metastasis. In some way, in fact, tumors succeed in avoid immune system detection and thus they limit the extent of immunological killing.

- **Deregulating cellular energetics**
  The progression of cancer cells involves also an alteration of energy metabolism in order to fuel cell growth and division. In normal tissue, under aerobic conditions cells process glucose, first to pyruvate via glycolysis in the cytosol and thereafter to carbon dioxide in the mitochondria; under anaerobic conditions, glycolysis is favoured and relatively little pyruvate is dispatched to the oxygen-consuming mitochondria. In tumor masses, cells can modify their glucose metabolism, limiting it to glycolysis. This condition has been termed *aerobic glycolysis*.

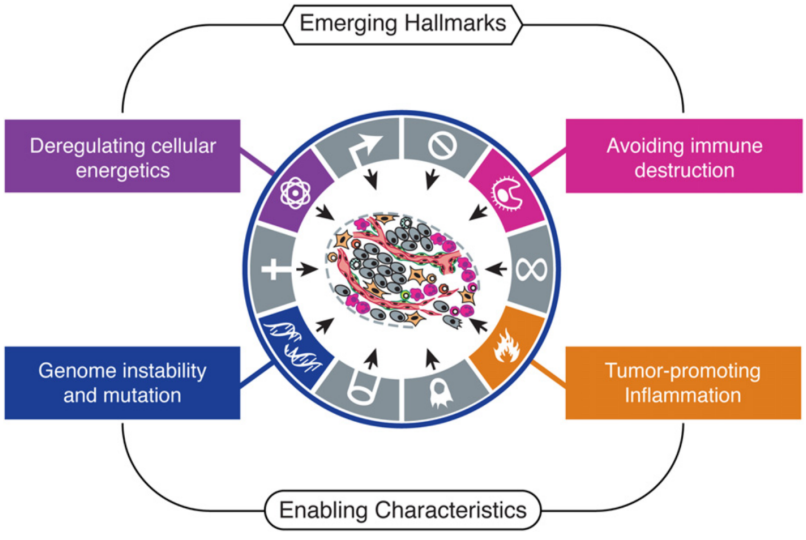


Figure 1.2: **Emerging hallmarks and enabling characteristics.**

In the paper of 2011, Hananhan and Weinberg also drew attention on the importance of what they called *tumor microenvironment* in tumorigenesis. They claimed that cancer is not, as it has been considered for long, an isolated mass of proliferating malignant cells. Rather, it is a complex tissue composed of multiple distinct cell types that interact with one another (see also [6]).

1.2 Mutations and cancer

As we have already said, cancer is initiated by a random mutation in a cell, followed by clonal expansion, during which the cell undergoes a series of mutation and selection events (a process that can be considered similar to the evolution of species) and would then expand its lineage. However, not all the somatic mutations that can occur are implicated in the development of cancer [4]. Some of them are deleterious and lead the clone straight to extinction. Others, called *passenger* mutation, are neutral, do not confer any advantage to the clone: they are somatic mutations without functional consequences. Only *driver* mutations are directly involved in oncogenesis: they confer to the clone a growth advantage that enables it to spread in the tissue. A recent study [7] suggests that passenger mutations can involve protein-coding genes and other functional elements that can potentially have deleterious effects on cancer progression. So, even if the individual effect of a passenger mutation is almost neutral, the collective burden of passengers can take action in the development of the cancer, leading to several oncological phenomena that are difficult to explain with traditional driver-centric view, suche as spontaneous regression.

1.3 Cancer risk

The aetiology of cancer is an outstanding problem. C. Tomasetti and B. Vogelstein [8], in 2015, proposed that the majority of cancer (about two third) does not originate from environmental factors or inherited predispositions, but is due to random mutations arising during DNA replica-

tion in normal, noncancerous stem cells. Simply put, cancer is due to "bad luck". This study has raised a significant debate. Some researchers [9], in fact, consider the report by Tomasetti and Vogelstein a pioneering work, that have highlighted the fact that stochastic accumulation of mutations during DNA replication is the major cause of variations in cancer incidence between tissues, opening new interesting questions. On the other hand, many scientists [10] [11] [12] [13] [14] [15] are extremely critical about this statement. They sustain that environmental and hereditary factors and lifestyle (such as sun exposure for melanoma, tobacco for lung cancer, viruses and obesity for hepatocellular carcinoma, and so on) play a fundamental role in carcinogenesis. Furthermore, they bring to light another delicate topic, the role of prevention: if cancer was due only to random mutation, prevention would be completely worthless.

As a matter of fact, many factors contribute to cancer initiation, and they are very difficult to be distinguished from each other. Studies that try to investigate this problem are complicated by the difficulty to collect data for all cancer types, and moreover by the fact that different populations (for example people from different countries) with different cancer patterns could provide different results.

1.4 Oncogenes and oncosuppressors

Mutations involved in tumor progression affect two types of genes:

- **Oncogenes**
  In the DNA there are some genes that normally help cell growth, the so-called *proto-oncogenes*. When these oncogenes are present in too many copies or when they undergo some mutation, they become malignant genes, the *oncogenes*. There are several mechanisms by which an oncogene can be activated: translocation, point mutation, deletion, insertional activation, amplification. When they are triggered, they become permanently responsible for increasing cell proliferation.

- **Oncosuppressors**
  Oncosuppressors are genes in charge of protecting the cell genome: when a cell expresses some mutated genes, they code for proteins that promote apoptosis, repress cell cycle regulation or repair DNA damage.

The most important oncosuppressor gene is p53, and its central role is reflected by its nickname: "guardian of the genome". This gene is evolutionary conserved in all animals and it plays a key role in many cellular metabolic pathways, where it acts as a regulator of the network. p53 reacts to stress signal and DNA damage by stimulating DNA repair or, in extreme cases, by inducing senescence or apoptosis. Unfortunately, in many cancer types, p53 does not function correctly. This happens for two reason: on the one hand, p53 could be itself mutated, or, on the other, it could be deactivated by another gene, MDM2, that fosters genetic degradation.

1.5 Cancer stem cells

Tumor masses are composed by a heterogeneous cell population, because of the accumulation of different driver and passenger mutations. According to the traditional view, all cancer cells have the same tumorigenic potential. Instead, following new suggestions, cancer cells can be considered hierarchically organized, with *cancer stem cells* (CSCs) at the top of the pyramid. CSCs are defined by two properties: they can re-grow the tumor from which they are isolated and their lineage differentiation is multipotent.

1.6 Angiogenesis and metastasis

Two of the most representative features of cancer are:

- **Angiogenesis**
  The tumor has the capability of inducing the growth of new blood vessels from the existing ones. This process is fundamental because the additional blood vessels are devoted both to bring oxygen and nutrients to cancer cells and to evacuate metabolic waste and carbon dioxide. Angiogenesis [16] is regulated by a balance between pro- and antiangiogenic molecules. These molecules can be secreted from cancer cells, endothelial cells, stromal cells, blood and connective cells. Their contribution, that control the angiogenesis, could change in respect of tumor type, site and developmental stage.

- **Formation of Metastasis**
  Cancer cells can spread across the body, flowing through blood vessels or lymphatic system (Fig. 1.3). Metastasis [17] is an orderly sequence of basic step, mediated by different classes of metastasis genes: local invasion, intravasation (that means the entry of tumor cells into the bloodstream), survival in the circulation, extravasation (the exit of tumor cells from capillary beds) and colonization. The temporal gap between organ infiltration and colonization produces a period of metastatic latency. This latency is considered an inefficient multistep process. In fact, malignant cells, during their journey through vessels and tissues, should overcome different obstacles, such as shear, compressive stresses and so on. In order to perform these tasks, cancer cells can modify their own structure and functionality. For example, they can change their cytoskeleton, or express new surface receptors. As well as stem cells, also metastatic cells define metastatic niches, made of peculiar locations, stromal cell types, diffusible signals and extracellular matrix proteins that sustain the survival and self renewal of malignant cells. It is interesting to note that different tumor types usually colonize some typical metastatic sites.

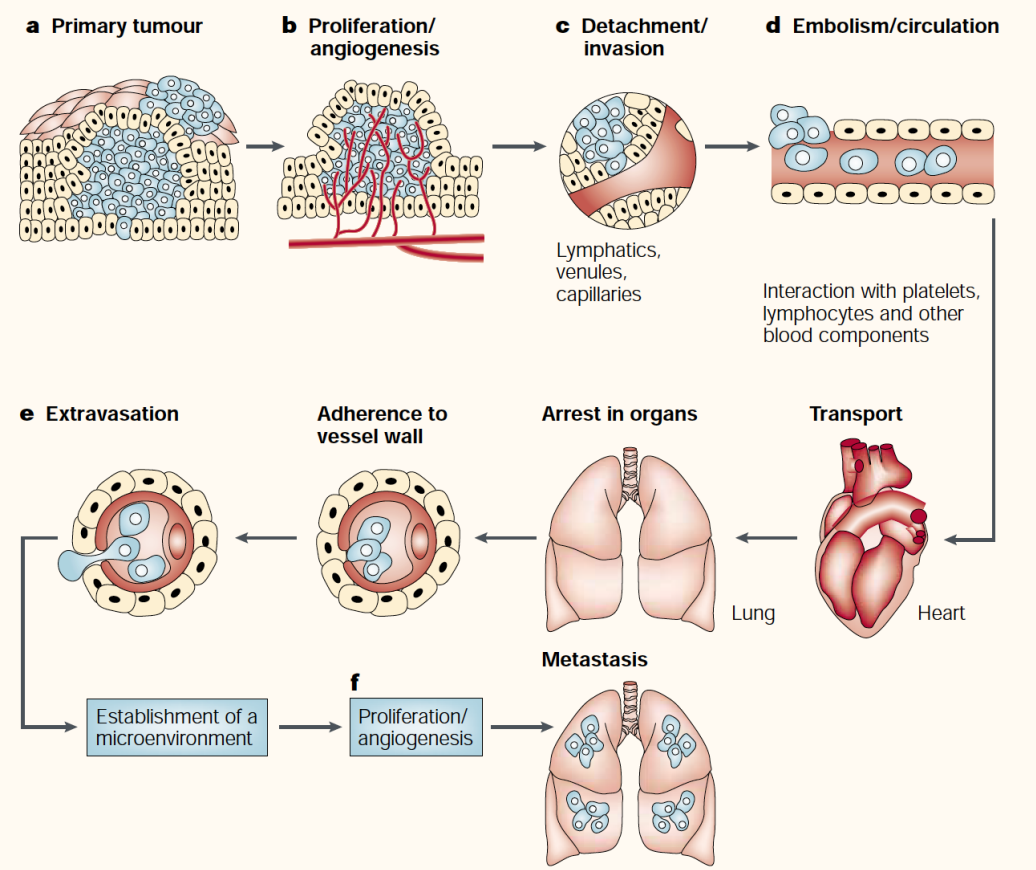


Figure 1.3: **Cellular transformation and tumour growth.**
a | Growth of neoplastic cells must be progressive, with nutrients for the expanding tumour mass initially supplied by simple diffusion. **b** | Extensive vascularization must occur if a tumour mass is to exceed 1–2 mm in diameter. The synthesis and secretion of angiogenic factors establish a capillary network from the surrounding host tissue. **c** | Local invasion of the host stroma by some tumour cells occurs by several parallel mechanisms. Thin-walled venules, such as lymphatic channels, offer very little resistance to penetration by tumour cells and provide the most common route for tumour-cell entry into the circulation. **d** | Detachment and embolization of single tumour cells or aggregates occurs next, most circulating tumour cells being rapidly destroyed. After the tumour cells have survived the circulation, they become trapped in the capillary beds of distant organs by adhering either to capillary endothelial cells or to subendothelial basement membrane that might be exposed. **e** | Extravasation occurs next, probably by mechanisms similar to those that operate during invasion. **f** | Proliferation within the organ parenchyma completes the metastatic process. To continue growing, the micrometastasis must develop a vascular network and evade destruction by host defences. The cells can then invade blood vessels, enter the circulation and produce additional metastases.

1.7 Big data in cancer

The last decades have seen the development of high-throughput or next-generation sequencing technologies, allowing researchers to obtain large amounts of data at relatively low cost for DNA sequence, transcriptome profiling, miRNA expressions, DNA-protein interactions and epigenetic classifications. These data, in order to be available for the entire scientific community, are usually collected in public databases. Some examples are: the *Cancer Genome Atlas* (TCGA: https://tcga-data.nci.nih.gov/tcga/), managed by the US National Cancer Institute; the UK-based *Catalogue of Somatic Mutations in Cancer* (COSMIC: http://cancer.sanger.ac.uk/cosmic/); the database of International Cancer Genome Consortium (https://icgc.org/) and the *Gene Expression Omnibus* (GEO: http://www.ncbi.nlm.nih.gov/geo/). Even if these data are available in enormous number, they are not so easy to be understood, because they are really heterogeneous. This happens for different reasons: first of all, mutation patterns are very different from patient to patient, even if they are affected by the same tumor. Another factor of variability is introduced by the experimental methods of data collecting and by the sequencing platforms used in each study.

1.8 Diagnostic methods

As it can be easily understood, the main goal of medicine is to detect the cancer as soon as possible, in order to intervene in the most appropriate way. Unfortunately, diagnosis often follows the illness of the patient; at that point, is likely that the cancer has already grown significantly. Moreover, when a tumor is discovered, it has already formed metastasis or have spread malignant cells that can remain in a latent state and come up later. After diagnosis, primary tumor mass can be removed by surgery and irradiation, whereas cancer cells spread throughout the organism can be treated by chemotherapy. If there are metastatic cells in latent state and they come up after a first treatment, they may have become chemoresistant and so insensitive to subsequent cycles of chemotherapy. This scenario is responsible for 90% of death due to neoplastic diseases. Regarding future

perspectives in this field, the purpose is to identify biomarkers that would allow the prevention of cancer cells spreading and also biomarkers that would help follow the treatment effect.

CHAPTER

2

THE PHYSICS OF CANCER

> Are there tangible examples of novel insights into cancer progression produced by mathematical modelling? Does the pay-off justify experimental and clinical cancer researchers "embracing the horror" of equations and calculus?
>
> A. Anderson and V. Quaranta [18]

2.1 What does physics have to do with cancer?

First of all we want to build a bridge to link biology and physics. Why physicists, mathematicians and engineers have started to worry about problems such as cancer?

As has been mentioned before, thanks to new sequencing techniques, the capacity to gather experimental or clinical cancer data has grown enormously. However, it is not an easy task to integrate and interpret these

large amount of data and to translate them into models and treatments. In a recent paper, F. Michor and her group noticed that physicists are well-trained to deal with problems in which large numbers of simple part interact, leading to the emergence of collective behaviours.

> Cancer is perhaps such a system. It has now become clear that cancer is not a strictly deterministic disease that progresses through a simple, fixed succession of specific mutations in two or three genes. Rather, there are many molecularly distinct routes to clinically identical cancers, and the final development of malignancy is influenced by a multitude of factors, encompassing the immune system, ageing, nutrition and microenvironmental details within particular tissues. Like other emergent phenomena, cancer cannot be readily understood by merely characterizing all its components. Developing a fundamental understanding of cancer that recognizes and embraces the great heterogeneity of tumors and their emergent properties may benefit from integrated teams of physicists, cancer biologists, mathematicians and engineers. [19]

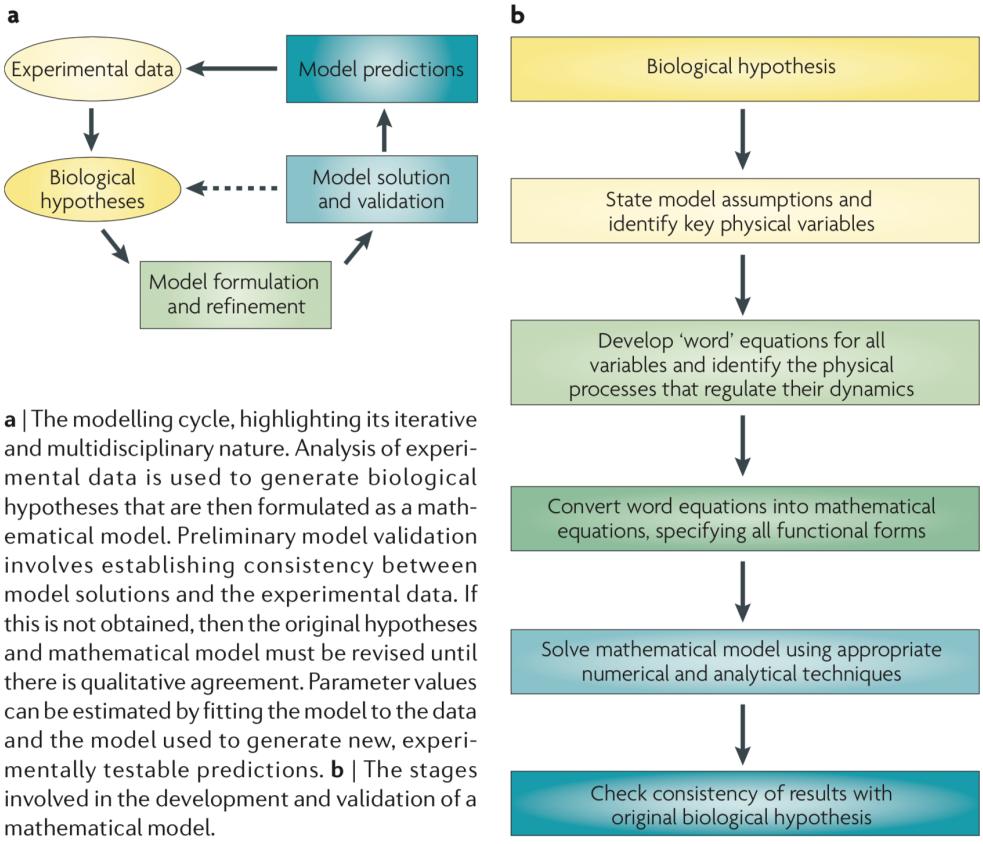
Thus, physics, mathematics, engineering and chemistry can make a decisive contribution to different aspects that concern oncology: they can define models, do data analysis, create new therapies, specifically tailor them to patient diseases, and so on. As concerns mathematical modelling, its greatest power lies in the capability to relate multiple components of a complex process to extract emergent properties that they themselves do not possess individually and also in the ability to reveal previously unknown and counterintuitive physical principles that might have been overlooked or missed by a qualitative approach to biology.

> When applied to experimental data, statistical techniques can reveal whether a particular intervention produces a significant response or whether a correlation exists between observable phenomena. Establishing why such correlations arise requires

> the statement of hypotheses postulating which physical processes are involved and how they interact. The biological experiments needed to test such hypotheses can be time-consuming, expensive and/or impossible with existing technology. In such cases, mathematical modelling can have an intermediate role, by providing an independent check of the consistency of the hypotheses: if a model is unable to reproduce the observed phenomena, then the original hypotheses should be revised before continuing. Mathematical models can also improve experimental design by highlighting which measurements are needed to test a particular theory and whether additional information can be gained by collecting supplementary data. [20]

Mathematical modelling of cancer is a multistage and iterative process: once that a first model is defined, it is necessary to check and validate its solutions and predictions through experimental data and biological and clinical observations. Then the model can be improved and its new predictions must be tested again (Fig. 2.1). Furthermore, according to the variety of mathematical tools that can be employed in studying cancer, it could be difficult to make a decision between different models, also considering that different approach can reproduce the same experimental results. Once that a model is chosen, mathematicians shall face another problem: the level of detail that should be reached.

> In such cases, it may be appropriate to appeal to Occam's razor to develop a model that includes sufficient detail to address the question of interest but not so much that it becomes obscured in detail. In practice, close collaboration between theoreticians and biomedical researchers is crucial to getting this balance right, because the models are only ever as good as the assumptions used to construct them and the data with which they are validated. Indeed, in many respects the form of the initial model is less important than starting the dialogue between experimentalists and modellers because the model is almost certain to be wrong. [20]



a | The modelling cycle, highlighting its iterative and multidisciplinary nature. Analysis of experimental data is used to generate biological hypotheses that are then formulated as a mathematical model. Preliminary model validation involves establishing consistency between model solutions and the experimental data. If this is not obtained, then the original hypotheses and mathematical model must be revised until there is qualitative agreement. Parameter values can be estimated by fitting the model to the data and the model used to generate new, experimentally testable predictions. **b** | The stages involved in the development and validation of a mathematical model.

Figure 2.1: **From biological hypothesis to testable prediction by mathematical modelling.**

In this frame, it would be desirable a more and more closer collaboration between mathematicians and biologists: everyone could take advantage from an effective interdisciplinary dialogue.

> Integrative mathematical oncology gradually creates a common language that on the one hand enables mathematicians to understand the biology, explain complex mathematics simply and build relevant realistic models, and on the other hand empowers cancer biologists with sufficient mathematical literacy to frame experiments in the context of quantitative models, transposing qualitative hypotheses into quantitative ones. [18]

A final remark is about the implementation of these theoretical models of cancer. They are usually complex and unfeasible to be solved by standard mathematical analysis and, consequently, they are nearly always sorted out through computational analysis. Such computational solutions, either numerical or simulation-based, require a great deal of computing power, which has only recently become widely available. It seems clear that we are now seeing the emergence of computational models as the dominant tool in mathematical models of cancer.

2.2 Mathematical oncology: integrating quantitative models

There are different way to classify mathematical models of cancer.

An initial approach [18] to this classification could be the distinction between descriptive and mechanistic models. The first ones try to reproduce the tumor cell population dynamics, leave out the cell biological details. The second ones, instead, focus on the biological aspects of tumor initiation and progression, in order to explain them. Obviously, there are many models that are between these distant perspectives.

Facing this issue from another point of view, models can be classified according to their mathematical and/or physical structure [21]. Single-cell-based models provide an appropriate description of cancer starting

from the properties of individual cells in order to predict collective behaviour of the whole tissue. They allow a portrayal of cell-cell and cell-microenvironment interactions. Several different individual-based models of tumor growth have been developed, including cellular automata models, Potts models, agent-based models or lattice-free models (for a review, see [22]). On the other hand, in continuous models molecules and cells are assumed to have a continuous distribution (for a review, see [23]). An example of this class are the deterministic reaction-diffusion models, that consider the tumor as a single continuous density varying both in space and time. Hybrid models, instead, combine the positive sides of discrete and continuous modelling techniques and portray, in a single model, chemical reactions and tissue environment. Between them, multiscale models play an important role: they are developed to describe interactions across different spatial and temporal scales and to encompass disparate components of a complex system, highlighting the emergence of collective behaviours.

2.2.1 The role of randomness

> All biological dynamical systems evolve under stochastic forces, if we define stochasticity as the parts of the dynamics that we either cannot predict or understand or that we choose not to include in the explicit modelling. To be realistic, models of biological systems should include random influences, since they are concerned with subsystems of the real world that cannot be sufficiently isolated from effects external to the model. The physiological justification to include erratic behaviours in a model can be found in the many factors that cannot be controlled, such as hormonal oscillations, blood pressure variations, respiration, variable neural control of muscle activity, enzymatic processes, energy requirements, cellular metabolism, sympathetic nerve activity, or individual characteristics like body mass index, genes, smoking, stress impacts, etc. Also to be considered are external influences, such as small differences in the ex-

> perimental procedure, temperature, differences in preparation and administration of drugs (if this is included in the experiment). In addition, experimental runs may be conducted by different experimentalists who inevitably will exhibit small differences in procedures within the protocols. Different sources of errors will require different modelling of the noise, and these factors should be considered as carefully as the modeling of the deterministic part, in order to make the model predictions and parameter values possible to interpret. [24]

Regarding models that deal with large groups of individuals, where deterministic models can be applied, the stochasticity that characterizes the individual behaviour is often considered negligible. However, it is clear that such approach is an approximation: if individuals evolve in a stochastic manner, so do finite populations. Thus, the best way to get in these phenomena is to define stochastic models and, possibly, to show that by some law-of-large-numbers effect it is well approximated by a deterministic simplification. In this scenario it is possible to estimate the magnitude of errors involved, that is, how much the "stochastic reality" differs from the "deterministic simplification".

> Population randomness through individual variability is called demographic stochasticity. Another source of randomness is environmental stochasticity, caused by spatial and/or temporal variation in environmental factors, that affects the population as a whole or their members individually. The environment in its turn can be influenced by the population. [...] Whereas the impact of demographic stochasticity can diminish for large population sizes, the effects of environmental stochasticity remain important for larger populations and should be included in model formulations, if relevant to the real biological system. Additional sources of randomness to be incorporated may be the effects of measurement errors or factors not explicitly included in a model, but lumped together into an unspecified random effect. [25]

Even though there are so many ways to make a distinction between mathematical models of cancer, we preferred a different approach compared to those described above. In the previous chapter we defined the main features that characterized cancer. Therefore, we introduce, section by section, the principal physical and mathematical tools and models, grouped according to the aspect of cancer they try to understand [26]. This can't be, for obvious reason, an exhaustive review of all the different mathematical models applied to oncology, but just a window over an area of research that is nowadays in the works.

2.3 Cancer initiation and progression

Historically, first modelling of tumor growth were based on the studies about population dynamics, starting from the ones about human population growth and demography made by Thomas Malthus between the 18th and the 19th century. Thus, the growing tumor has been considered as a deterministic dynamic system represented by ordinary differential equations.

> For tumor growth, the critical part is the period of growth (regression) and not so much the period of true stagnation which even may not be achieved before the death of the host. [...] The central concept of a dynamical system is the trajectory. For tumor growth the trajectory is the growth curve that describes the change in tumor size with time from the start of proliferation of initial tumor cells. The tumor size is expressed by mass or volume or cellularity, depending on what is measured. [27]

Two of the most widely used models for the growth of cancerous cell populations are the Gompertz and the logistic equations. The first one[1] is a sigmoid curve, described from the differential equation

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -gN \ln\left(\frac{N}{K}\right),$$

[1] Initially formulated by Benjamin Gompertz in 1825 to study the human mortality.

where N is the population and the constants g and K are respectively the growth rate and the carrying capacity (that is the asymptote for the population, when $t \to \infty$, established by the available resources). The solution of the above equation is

$$N(t) = Ke^{ae^{-gt}},$$

where the value of a is given by imposing the initial condition

$$N(0) = N_0 \quad \to \quad a = \left(\frac{N_0}{K}\right).$$

The logistic curve, or Verhulst[2] model, instead, is described, using the above notation, by

$$\frac{\mathrm{d}N}{\mathrm{d}t} = gN\left(1 - \frac{N}{K}\right).$$

Here a different assumption is made: the growth rate is proportional both to the population size and to available resources. The solution is

$$N(t) = \frac{KN_0\, e^{gt}}{K + N_0(\, e^{gt} - 1)}.$$

After these first, naive approaches, great deal of progress has been made.

2.3.1 Branching and Moran processes

Branching processes [25] are a class of simple stochastic models that have been used extensively to study the probabilistic growth of cell population. Many deterministic models are expectation versions of them. They can be defined in discrete time or continuous time and with evolution rules that may or may not depend on time.

A branching process is a Markov process (see Chapter 4 for further details) in which every, at time t, individual cell produces a random number of offspring at a later time $t + \Delta t$ (fig. 2.2).

[2]Introduced for the first time by Pierre F. Verhulst in 1838.

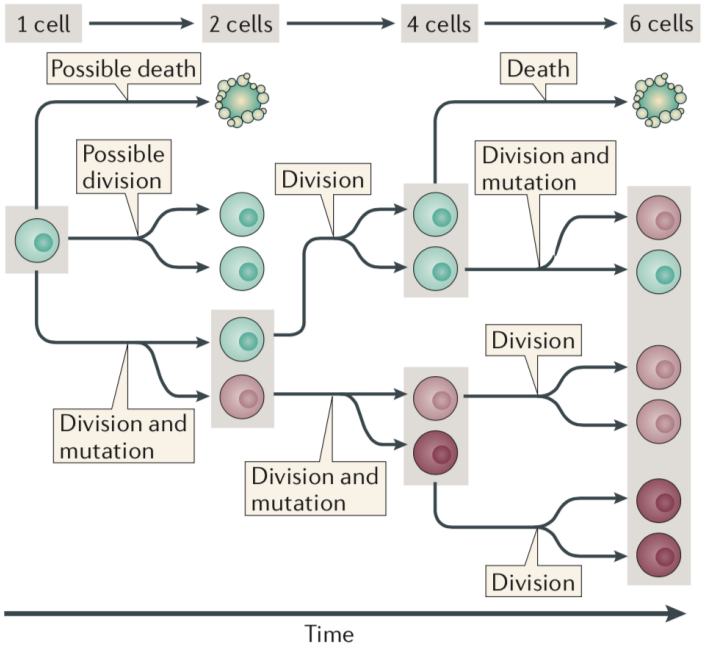


Figure 2.2: **Branching process: realization of three time steps.**

In discrete time, the Galton–Watson process is the oldest and best known branching process (for more details, see [28]). The underlying idea is the following (this is a slightly improved version of Galton-Watson model: the simplest one does not include mutations). An individual (a cell) goes through a number of events: proliferation, mutation or death. The basic assumption is that each event is characterized by a given rate which is independent of population size and composition. As mutations accumulate in the cell population, each new cell type that emerges may have a new set of rates. It is worth noting that independence in reproduction and survival among different individuals is assumed. The rationale for this is that the models are meant for small populations, in which can be presumed that resource limitations, for example, do not play an important role.

Approaches of this kind have been used, for instance, to investigate the onset of a driver mutation during tumorigenesis [29] or also to model the accumulation of passenger and driver mutations during cancer progression. In their paper [30], I. Bozic *et al.* modelled tumors as a discrete time branching process that starts with a single driver mutation and pro-

ceeds as each new driver mutation, reducing the death rate of mutated cells, leads to a slightly increased rate of clonal expansion. The authors proposed a formula that relates the number of driver mutations to the total number of mutations in the tumor, and applied this methodology to experimental data to infer the selective advantage conferred by typical somatic mutations. B. Bauer *et al.* [31], instead, used a branching process to study the impact of a large amount of passenger mutation on cancer progression, whereas C. Tomasetti and B. Vogelstein [32] (in a previous and preparatory paper of the one [8] we have discussed in the previous chapter) used the same mathematical tool to investigate the evolution of somatic mutations in which all relevant phases of a tissue history are considered: the tissue formation during development, the tissue homeostasis and self-renewal and the tumorigenesis, initiated by one driver mutation. They found that the number of somatic mutations in tumors of self-renewing tissues is positively correlated with the age of the patient at diagnosis, because older tissue have had more time to accumulate alterations, and they claimed that half or more of the somatic mutations in certain tumors of self-renewing tissues occurs before the onset of neoplasia.

If we introduce in the modelling, instead of a fixed fitness value, the concept of a fitness distribution such that a randomly drawn fitness value is assigned to each mutation, the best mathematical approach available is the fixed-size Moran process.
The Moran process (Fig. 2.3) is used to model the stochastic dynamics in a population of constant size. An individual is selected at random, but with probability proportional to its fitness. This individual produces an identical offspring, that replaces another individual randomly chosen to die. If there are n species, $i = 1, 2, \ldots, n$, the numbers of individuals for each type are N_1, N_2, ..., N_n, which sum to N (and N remains constant over time), and if the species can have different fitness value f_1, f_2, ..., f_n, then, the probability that the i-th species increases of one individual and

the i-th species decreases is

$$P(N_i \to N_i + 1, N_j \to N_j + 1) = \frac{N_i f_i}{N_i f_i + N_j f_j} \frac{N_j}{N}.$$

The Moran process can also include random mutations, nonrandom death proportional to "weakness" (or inverse fitness) or time-dependent fitness. Moran processes are used to model the dynamics of mutation accumulation. For example, J. Foo, K. Leder and F. Michor [33] used a Moran process to model the compartments of cells where cancer could arise. In a similar way, in another work [34], authors designed a mathematical model of the evolutionary processes of mutation accumulation both in healthy tissue during the phase prior to tumor initiation and during the clonal expansion phase of the tumor.

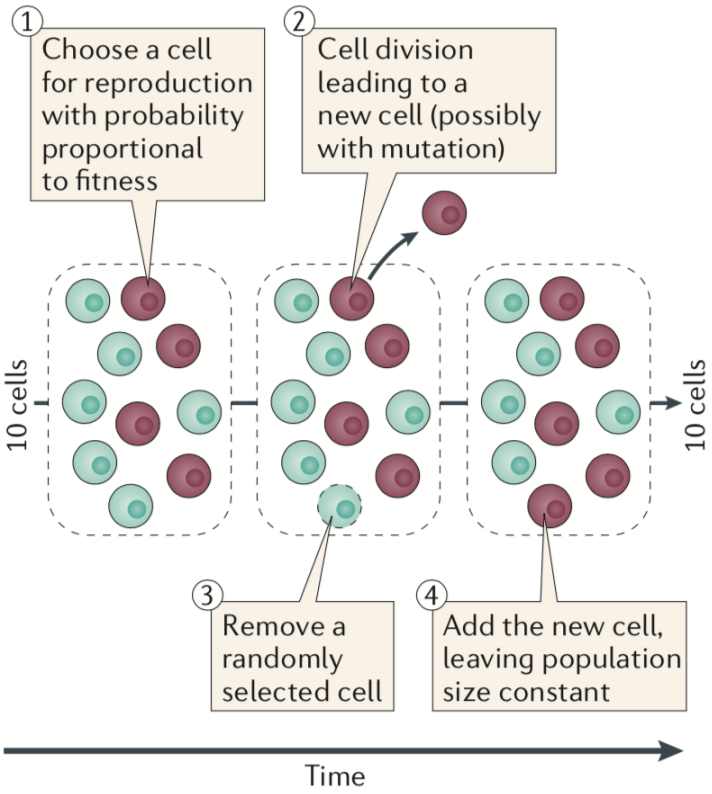


Figure 2.3: **Moran process: realization of one time step.**

In the study of the dynamics of mutation accumulation it is also important to investigate the cell type in which mutations arise. For this purpose, many models have been formulated. Some of them are deterministic, such as [35], where it is shown that the hierarchical organization strongly suppresses cells carrying multiple mutations and thus reduces the risk of cancer initiation. Nonetheless, many diseases are based on the accumulation of multiple mutations. Then, closed solutions for the deterministic clonal dynamics and the reproductive capacity of single clones are derived. Stochastic framework, instead, are useful to find out whether a stem cell, a progenitor cell or a terminally differentiated cell is more likely to undergoes a malignant mutation, becoming the origin of a tumor. The answer to this question depends on the type of tumor: for example, in hematopoietic neoplasia [36] progenitor cells seems to be the origin of tumorigenesis, meanwhile in brain cancer [37] stem cells are more likely to initiate the tumor.

The first attempts to model a cancer mass has been the uniform spherical models in one-dimension, where tumor growth is measured by the distance of a cell from the tumor center. Such models treat the tumor as a growing spherical mass: they consist of an ordinary differential equation, derived from mass conservation, coupled to one or more reaction-diffusion equations, that describe the distribution, within the tissue, of nutrients and growth factors. Usually the spheroid is composed by different layers: there is a central core of necrotic cells, surrounded by a rim of quiescent cells and by an outer annulus of proliferating cells [38] [39]. These models have been later extended to two dimension: in [40], S. Ferreira *et al.* analyzed the avascular cancer growth in a model including cell proliferation, motility and death, as well as competition for nutrients among normal and cancer cells, with particular interest in the role of capillary vessels. A further extension of these models has been the 3D simulation of tumors. X. Li *et al.* [41] developed an adaptive boundary integral method to simulate nonlinear tumor growth in 3D.

A cancerous mass is initially composed by cells that contain the same core of genetic mutation and the final heterogeneity observed is due to later alterations. B. Waclaw and his collaborators [42], instead, described, thanks to a 3D model, how short range dispersal and cell turnover can account for the expansion and mixing of these alterations inside the tumor. Following a different approach, reaction-diffusion modelling [43] predicted that glioma (a neoplasia that affects the glia cells of the brain and of the spine) growth can be modelled as a travelling wave with a diameter that increases linearly in time. Talking about brain tumor, it is worth emphasizing [44] the importance of emerging quantitative imaging methods that pave the way to a new generation of predictive methods. Indeed, magnetic resonance imaging (MRI) and positron emission tomography (PET) have matured to the point where they offer patient-specific measures of tumor status at the physiological, cellular, and molecular levels.

2.3.2 Temporal order of events in cancer progression

An outstanding question is about the temporal sequence in which alterations occur during human tumorigenesis. Some authors (for example [45], that identified different stages in colorectal carcinoma) assume that exists a single temporal sequence, a straight-like chain, and so events are always sequential, never simultaneous. But this idea, if has been proved successful for some cancer types, can be inadequate for others. So, further approaches have been introduced, especially in the light of recent improvement of cancer genome studies [46] [47].

One of the most thriving idea is the oncotree model, that is based on a probabilistic phylogenetic tree[3] approach [48] [49]. The temporal order of events is computed as a function of the distance of an event from the root node, that is, the time between the initiation and the event. Also oncotree models, however, show some limitation: for example they do not allow shared ancestors for multiple leaves. To overcome these aspects,

[3]That is, a branching, tree-structured graph that represents the evolutionary relationship among different mutational stages of tumor cell population, quantified by some measure of distance between individual cells or patient samples. [26]

other acyclic graphical models have been introduced [50]. These models determine the order of somatic alterations from cross-sectional data sets, but manage to do this at the cost of large computational burden owing to increased model complexity.

Another successful approach is borrowed from population dynamics, also coupled with optimization algorithms: it makes possible to study the evolutionary dynamics [51] of cell populations, with particular interest in the accumulation of malignant mutations. These models are often based on Moran processes, that are used to calculate transitions between different mutational states. They show that most cancer types are characterized by multiple evolutionary trajectories that lead to the fully transformed state, which suggests a large extent of heterogeneity in the temporal order of cancerous events.

Lastly, another tool employed to go into the temporal order of cancer event is the agent-based simulation[4]: in [52], researchers found that the utilisation of cross-sectional data to infer mutational order could be misleading, while phylogenetic methods based on sampling intratumor heterogeneity may reconstruct more accurately the evolutionary history of tumors.

2.3.3 Tumor microenvironment

Cancer progresses as a result of the collective dynamics that emerge from interactions between tumor cells and their microenvironment. However, for long time, the importance of the environment has been ignored[5], but recent studies have been more and more focused on the role of the tumor stroma, that includes the extracellular matrix, fibroblasts, immune and in-

[4]Agent-based simulation is a computational approach that models complex systems consisting of interacting discretized items or "agent". In cancer modelling, these agents often represent cells, which can mutate into other types, divide into two cells, die or move in space. These simulations can be implemented according to either probabilistic or deterministic laws. [26]

[5]An exception was the Paget's *seed and soil* hypothesis. The English surgeon Stephen Paget compared tumor cells with the seed of plants, in that they are both "carried in all directions; but they can only live and grow if they fall on congenial soil". Similarly, he argued that metastatic cells must thrive only where conditions are in some way favourable. [6]

flammatory cells and blood-vessels cell. The relative amount of stroma and its composition vary considerably from tumor to tumor and do not correlate with the degree of tumor malignancy. But the interactive signalling between tumor and stroma contributes to the formation of a complex multicellular organ, such as cancer [6].

As it can be easily understood, the mathematical modelling of microenvironmental interactions often requires complex model considerations, also because they must include biophysical properties and interrelated processes. The most useful and versatile mathematical tools for this purpose are ordinary and partial differential equations. Of course there is a huge mathematical literature about this topic, but we prefer to cite the few, intuitive words of P. Altrock, L. Liu and F. Michor:

> Systems that are deterministic (exactly or approximately) can be described by *ordinary differential equations* (ODEs). Their main characteristic is that they have one independent variable. For dynamic systems, the independent variable is time. Dependent variables can be the volume of a tumor, the fraction of a genetic alteration in a population or the chance of finding a receptor in a certain state at a certain time. ODEs can describe systems of few and many dimensions, and allow chaotic and complex behaviour. [...] For dynamic systems in which the quantities of interest — such as the concentration of oxygen — depend on more than one independent variable (for example, time and space), *partial differential equations* (PDEs) are used. This is beneficial especially when descriptions in higher dimensions are needed. For example, the concentration of oxygen in a tissue at time t in position x (for example, the distance to the centre of a blood vessel) can be denoted by $c(x, t)$. [...] Typically the oxygen concentration is only one component of a system of PDEs, on which the behaviour of tumor cell density depends. [26]

Talking about the interplay between malignant cells and tumor microenvironment, it emerges the key role of hybrid models. These models represent cells as individual discrete entities and use continuous concentration or density fields to model intracellular and extracellular environments. By their very nature, hybrid models are ideal for examining direct interactions between individual cells and between cells and their microenvironment, but they also allow us to analyze the emergent properties of complex multicellular systems.

Hybrid models can be sorted in two classes: on- and off-lattice [53], where this means that the positions of the cells are or are not imposed, but in both cases the underlying chemical or physical fields are typically defined on regular grids. The on-lattice model are easier to handle in computational implementation, but are less realistic: changes in the underlying chemical fields are modeled on the cell scale and therefore discontinuities in these values may not reflect the smooth changes in chemical gradients. These approaches on grids are often collected under the name of cellular automata (CA). Off-lattice models, instead, have a more realistic representation of cell spatial location, but they need special algorithms to handle, also from a computational point of view, the complexity of cells movement and of interactions with the underlying environment (Fig.2.4).

To take another example, in [54], the authors joined a continuous deterministic model of the extracellular matrix dynamics and a discrete cellular automata-like model of the cell migration and interaction in order to examine the effects of tumor cell heterogeneity upon the overall spatial structure of tumor and also to emphasize the importance of cell-cell and cell-environment interactions.

An hybrid cellular automata model [55] have been also developed to study the deregulation of glucose metabolism, that is the emergence of the glycolytic phenotype. The authors, through a neural network approach, analyzed the influence of the tissue oxygen concentration and extracellular matrix density on the dynamics of tumor growth: they observed that glycolytic phenotype is most likely to emerge in anaerobic condition and in tissue with a high matrix density.

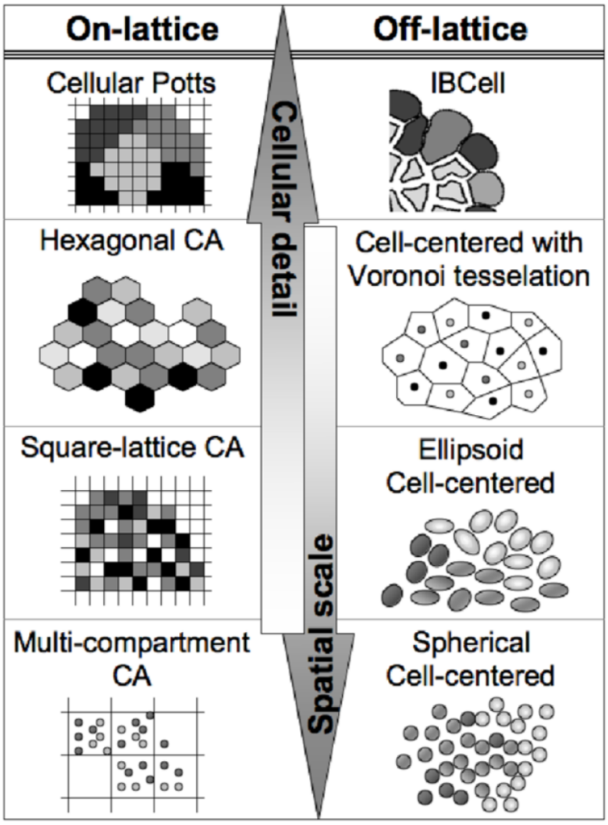


Figure 2.4: **Reciprocal relation between the numbers of cells handled by the models and the level of included cellular details.** For more details about "IBCell": `https://labpages.moffitt.org/rejniakk/IBCell.html`

2.4 Angiogenesis and metastasis

Angiogenesis is a well-orchestrated sequence of events involving endothelial cell migration, proliferation, degradation of tissue, new capillary vessel (sprout) formation, loop formation (*anastomosis*) and, crucially, blood flow through the network. Studies of tumor-induced angiogenesis have been almost always related to the ones about the interaction between tumor and environment. Some of them have been performed following a continuous, deterministic approach and they are capable of capturing some features of angiogenesis such as average sprout density and network expanding rate. Other ones, instead, used a discrete probabilistic framework, that makes easier to follow the motion of individual endothelial cells. However, as we have already made clear, many researchers eventually choose to use hybrid models [56] (for a review, see [57]).

In 1997, M. Orme and M. Chaplain [58] developed a two dimensional model of capillary-vessel formation. Three dimensional models have been designed later [59]. P. Macklin *et al.*, instead, formulated a multiscale mathematical model for tumor solid growth [60] which couples with a model of tumor-induced angiogenesis. They performed nonlinear simulations in order to show the importance of the coupling between the development and remodeling of the vascular network, the blood flow through the network and tumor progression.

The angiogenic process is both adaptive and dynamic: these feature have been studied [61] following the hypothesis that blood flow can affect the growth of the network in many ways: through the diffusion of angiogenic factors, through migratory cues via the extracellular matrix and through perfusion-related haemodynamic forces.

Historically, the first attempt to model cancer metastases has been the already mentioned seed and soil hypothesis by S. Paget, in 1889. This idea has been later resumed. In 2003, I. Fidler [62] has redefined it through three principles that comprehend the hetereogeneity and the clonal origin

of metastases masses and their specificity in colonizing organs. Ten years later, the seed and soil hypothesis has been renamed after "diaspora" [63], that is the scattering of people from an established homeland. According to this comparison, cancer, in this paper, is considered from an ecological point of view (we will spend a few words about this approach to oncology later). The diaspora paradigm takes into account several variables: the quality of the primary tumor microenvironment, the fitness of individual cancer cell migrants as well as migrant populations, the rate of bidirectional migration of cancer and host cells between cancer sites and the quality of the target microenvironments to establish metastatic sites.

However, many other approaches have been used to model metastasis features. Earliest descriptions [64] portrayed metastasis in the light of competition between healthy tissue and malignant cells, in the framework of population dynamics. More recently, cell-based model [65], with related agent-based computational simulation, has been used to understand how the interaction of cancer stem cells and their nonstem progeny can influence the formation of metastasis, highlighting the fact that tumor populations devoid of stem cells or developed from cancer stem cells could still persist as long-term dormant lesions.

Again, a Markov chain Monte Carlo mathematical approach [66] has been used to determine a pathway diagram that classifies metastatic tumors as "spreaders" or "sponges" and that orders the timescales of progression from site to site. The authors, thanks to this method, tried to quantify the stochastic, systemic and often multidirectional aspects of cancer progression. In particular, they take into account the self-seeding[6] of the primary tumor (primary seeding), the re-seeding of the primary tumor from a metastatic site (secondary seeding) and the re-seeding between metastatic tumors (metastatic seeding). Talking about different types of self-seeding, a study [67] has been conducted to elucidate the differences and relative probabilities between them, following the representation of the vascular system as a network throughout which the circulating tumor cells can spread. Researchers developed a model to test the relative likelihood of

[6]The phenomenon of tumor "self-seeding" occurs when circulating tumor cells (CTCs) repopulate the primary tumor and accelerate its growth.

primary and secondary seeding, showing that, in the end, secondary is far more likely than the other.

H. Heano and F. Michor [68], instead, designed a stochastic model of the evolution of tumor metastases in an expanding cancer cell population. They also calculated the probability of metastasis formation at a given time during tumor evolution, the expected number of metastatic sites and the total number of cancer cells as well as metastasized cells.

Talking about the already formed metastasis, many studies have investigated the heterogeneity of the cells of a single metastasis and also the differences between metastatic cells and primary tumors cells. V. Almendro *et al.* [69], by defining quantitative measures of intratumor cellular genetic and phenotypic heterogeneity in primary and metastatic breast tumors and by assessing tumor topology, found that distant metastatic tumors are the most diverse. This fact can explain the frequent therapy-resistance of advanced stage disease. Another reason of the heterogeneity of tumor metastasis has been found [70] to be the diversity in circulating tumor cells that mirrors the variety of the populations of competing phenotypes within the primary tumor.

A stochastic Markov chain model for metastatic progression have been developed by P. Netwon *et al.* [71]: they defined a network construction of metastatic sites with dynamics modeled as an ensemble of random walkers on the network. Successively, they calculated transition probabilities interpreted as random variables and used them to construct a circular bi-directional network of primary and metastatic locations.

A final mention has to be made about the models of metastasis that have tried to include the biophysical properties of the milieu through which circulating tumor cells spread. To cite just an example, A. Pathak e S. Kumar [72] tried to shed light, thanks to a 3D culture, over the role of interaction between migrating tumor cells and the extracellular matrix, emphasizing the importance of extracellular matrix stiffness and geometry.

2.5 Treatment and drug resistance

This section is going to be really heterogeneous. In fact, many aspects of clinical treatment and of drug response and resistance are related to physics or mathematical modelling: tumor forecasting[7], radiotherapy and its dosing strategies, immunotherapy, how the resistance to drugs evolves and so on.

As concerns radiotherapy, an earliest approach has been the attempt to find the growth curve of neoplasms treated with radiotherapy with respect to the one followed by untreated tumor (that is well represented by an exponential or, as we have seen before, by a Gompertzian growth curve). An important role has been played by linear-quadratic model, a prominent heuristic model used to describe cell survival under radiation. The number of surviving cells after a certain dose of radiation has been administered takes the form of an exponential function with a linear and a quadratic term in its argument. An example of this picture could be the paper by R. Dale [74], that investigated the application of a linear-quadratic dose-effect formalism in the context of radiotherapy.

More specifically, many studies have been conducted about the response to radiation therapy of glioblastoma multiforme (GBM), that is the most common and malignant form of glioma. R. Rockne *et al.* [75] applied a patient-specific, biologically-based mathematical model for GBM growth, that quantifies the response to radiotherapy: they used reaction-diffusion partial differential equations to describe tumor cell proliferation and invasion, along with the linear-quadratic model for the response to radiation therapy. Their model makes possible the creation of a virtual *in silico* tumor with the same growth kinetics as a particular patient and can not only predict treatment response in individual patients *in vivo* but also provide a basis for evaluation of response in each patient to any given therapy. Still talking about patient-specific therapies, a study by D. Corwin [76] exposes a biomathematical model of glioma proliferation, invasion and radiotherapy that tries to optimize the dose for treatment, using the pioneering metric of "Days Gained".

[7] That has been often compared to metereology. [73]

This last concept is a prognostic tool firstly introduced by M. Neal *et al.* in [77]: the metric of "Days Gained" is defined as the difference in time between the post-treatment MRI scan and the predicted time at which the tumor would have reached the same radius if the patient had not been treated; the latter is estimated relying on the pre-treatment MRI scan and on the subsequent tumor simulations.

Immunotherapy is considered one of the most promising cancer treatment approaches: the immune system is clearly capable of recognizing and eliminating tumor cells. However, tumors frequently interfere with the development and function of immune responses. Thus, the challenge for immunotherapy is to use advances in cellular and molecular immunology to develop strategies that effectively and safely augment antitumor responses [78], without harming healthy tissue. The phenomenon of competition between cancer and immune cells has been examined through deterministic models. *Inter alia*, a prominent role is played by prey-predator models[8], of ecological inspiration [27]. Later, multiscale mathematical models [79] have been developed: at a sub-cellular scale, the main activities are within the cells or at the cell membrane (genetic changes, expression of signals between cells,...); the cellular scale refers to the main interactive activities of the cells (activation and proliferation of tumor cells and competition with immune cells); lastly, at a macroscopic scale occur the phenomena that are typical of continuum systems: cell migration, diffusion and so on.

However, tumor response is often transient and therapy frequently fails due to emergence of resistant populations. That could happens because, although cancer is a highly dynamic systems, cancer therapy is typically administrated following a fixed protocol. A recent clinical approach has introduced the idea of "adaptive therapy" [80], that evolves in response to the temporal and spatial variability of tumor microenvironment and cellular phenotype as well as therapy-induced perturbations. The aim of

[8]These models, also known as Lotka-Volterra dynamics, are used to describe the dynamics of ecological species, or types, as a nonlinear deterministic process. They were originally used to describe population dynamics of predators and prey, taking into account abundance, interactions and population growth and diminution. They can also be used to describe mutualistic and competitive evolutionary dynamics. [26]

adaptive therapy is to enforce a stable tumor burden by permitting a significant population of chemo-sensitive cells to survive so that they, in turn, suppress proliferation of the less fit, but chemo-resistant, subpopulations.

Also methods from evolutionary dynamics have suggested therapeutic innovative strategies, such as combination therapy. Researchers [81] found, modeling cancer as a continuous time multitype branching process, that dual therapy results in long-term disease control for most patients, obviously if there are no single mutations that cause cross-resistance to both drugs; in patients with large disease burden, triple therapy is needed. Moreover, they discovered that simultaneous therapy with two drugs is more effective than sequential one.

Even considering the differences between techniques, the crucial, promising innovation in cancer treatment is targeted and patient-specific therapy. But, despite all, the biggest, already unsolved, problem is drug resistance.

One of the models more widely used to derive the probability of accumulating resistant cells in exponentially expanding populations is the Luria-Delbrück model[9].

> This approach assumes that both sensitive and resistant cells grow exponentially, and that sensitive cells can generate resistant cells during cell division. A two-type birth–death process can be used to calculate the probability of pre-existing resistance and the expected number of resistant cells before diagnosis. The latter was found to be independent of the mutation rate if mutations are rare, but to increase with the tumor size at detection. The probability of pre-existing resistance increases in proportion to both detection size and mutation rate. [26]

To investigate this phenomenon, other stochastic dynamical systems have been developed. Y. Iwasa, F. Michor and M. Nowak [83] tried to answer the following question: if a genetically diverse population of replicating

[9]The Luria-Delbrück experiment investigated whether mutations occur independently from, or owing to, selection. In its first formulation [82], data from growth experiments in which bacteria were challenged to a stochastic process model used to calculate the probability of having a certain number of resistant mutants. The findings suggested the mutations occurred randomly over time and were not a response to selection. [26]

organisms is challenged with a selection pressure that has the potential to eradicate it, what is the probability that this population will produce escape mutants? They used a multitype branching process to describe the accumulation of mutations in order to calculate evolutionary escape dynamics for a malignant population that has undergone a biomedical intervention with the potential aim to eradicate it. Depending on initial distribution and fitness landscape, they succeeded in quantifying how escape depends on the pre-existence or on the emergence of resistant mutants. A. Coldman and J. Goldie [84] [85] chosen a cell-based model for tumor growth coupled with a stochastic branching process to explain intrinsic and acquired resistance. N. Komarova and D. Wodarz [86] found that resistance arises mainly before the start of treatment and, for cancers with high turnover rates, combination therapy is less likely to yield an advantage over single-drug therapy.

2.6 An ecological point of view

Many of the models above introduced talk about *evolutionary* or *ecological* features of tumor onset and progression. Let us explain how it is possible to speak of an ecological perspective over a problem such as cancer.

> During the past decade, it has become increasingly recognized that a tumor is not genetically homogeneous but is rather composed of many genetically diverse cancer cells. If variability in the population is heritable and if it affects fitness, then the system is going to evolve, leading to competition for space and common resources and resulting in different clones being selected for or weeded out of the population due to natural selection. [...] From an ecological perspective, one can look at this process as an attempt of new species (cancer cells), which have different metabolic and reproductive strategies compared with the resident population (somatic cells) to invade a new habitat (tissue). Successful invasion will result in the formation of a primary solid tumor. [87]

B. Ujivari, B. Roche and F. Thomas [88] underlined five different "ecological" aspects of cancer:

- **Somatic selection shapes malignancies**
  As we have already explained in details, cancer is an evolutionary process, result of accumulation of somatic mutations. From an evolutionary game theory perspective[10] [87], individuals within a cell population in healthy tissue have been moving toward an evolutionarily stable strategy (ESS). Theoretically, when ESS is established, invasions are unlikely to happen. However, if one of the species escapes its ecological constraints and proliferates rapidly, the balance is destroyed, triggering extinctions of other species and ecosystem collapse [89].

- **Ecological principles explain how cancers interact with microenvironments**

  > There is no such thing as fitness for a gene or individual except in relation to a specific ecological environment. Fitness is a characteristic not of genotypes, but of phenotypes interacting with environments. The microenvironments inhabited by cancers influence their growth, as much as their genotypes. Furthermore, the growth of a tumor creates microenvironments that can speed or slow subsequent growth. [88]

  For example, hypoxia induces the angiogenic switch; clones gain an advantage if they have the capacity to adapt their metabolism to sudden changes in oxygen tension or nutrient availability without changes in their genomes; cancer cells may enter in a quiescent state if the nutrient supply is limited and so on. Once that cancerous cells begin to propagate in the tissue, they can construct their own niche. This last concept is purely ecological and, to be applied to

[10] We are going to show an example of evolutionary game theory applied to cancer in Chapter 5 and in Appendix B.

cancer, should include the definition of nutrients (glucose, phosphorus, iron, lipids and other materials necessary fro cell growth and reproduction), space (including extra cellular matrix, which is often destroyed by tumors) and predators (cells of the immune system). Invaders (cancer cells) can modify the niche to be better suited for them or they can exploit the niche in such a way to make it uninhabitable by anyone, inducing increased migration. According to this picture, isolated tumor and metastasis can be studied through the principles of island biogeography: cancer is similar to a geographic expansion of invasive species.

A last remark is about biodiversity: it makes the ecosystem more stable and resilient, and this hold true both for healthy and malignant ecosystems.

> Tumor heterogeneity can not only accelerate evolution but also directly select for more aggressive cancer. [...] Theoretical models also predict that the harsh microenvironments created by the tumor itself can create a selective pressure for more aggressive cancer lineages. This study also found that increasingly stressful microenvironments lead to a morphological transformation of the tumor from smooth and non-invasive margins to margins with finger-like protrusions that are associated with an aggressive and invasive phenotype. [89]

- **Behavioral ecological principles explain competition among cancer clones**
  Recent studies have been highlighted the possibility that different clones can provide resources that promote the survival and replication of other clones in a process that is something like cooperation. Even if the current knowledge of ecological dynamics within tumor is limited, till today many tools have been used to model cancer progression in this point of view: for example, the logistic growth curve or the Allee effect[11].

[11] A commonly observed deviation from logistic growth, with the per-capita growth

The Allee effect, in particular, could be exploited to describe many phenomena that frequently occur in cancer, such as the fact that large tumors are more difficult to be treated by drugs and radiotherapy. Or also the so called *minimal residual disease* (MRD), that is the possibility that some cancers recur after treatment, because a number, even very low, of malignant cell remains in the organism. Moreover, the Allee effect is also used to study cooperation among cells and clones, that can have a positive effect on population ecology, growth thresholds and resilience to perturbations.

- **Natural selection explains why cancer is rare**
  We have already discuss the ability of natural selection to shape mechanisms to prevent cancer, such as the activity of oncosuppressors. This is the first reason why cancer is so uncommon. From an ecological perspective, these mechanisms can be viewed as prey-predator systems.

- **Evolutionary medicine explains why cancer is common**
  This happens because evolutionary medicine tries to investigate why, despite the presence of the control mechanisms, some mutations lead to malignancies: it can depend on the environmental factors, on the stochastic probability that a mutation anyway happens, on the trade-offs that are intrinsic to all cancer prevention mechanisms and so on.

rate reaching a maximum at an intermediate population size. One often distinguishes between a strong Allee effect, when the growth rate is negative at small population sizes, and a weak Allee effect, when the growth rate at small population sizes small but positive. [89]

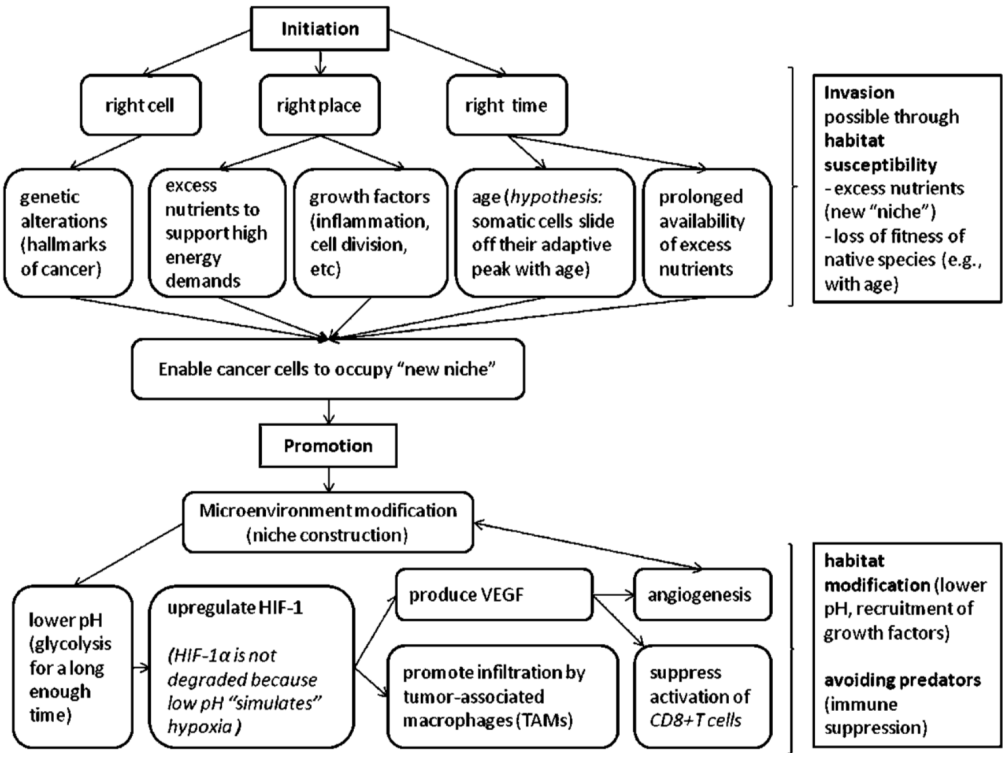


Figure 2.5: **Schematic representation of the possible mechanism of tumor initiation and progression from an ecological point of view.**
Tumor initiation corresponds to the mechanism of species invasion and is hypothesized to be possible when the environment is permissive, in particular, when there are excess nutrients (new niche) and when competitors (somatic cells) are less fit compared with the invaders. Tumor promotion corresponds to niche colonization and modification by the invading species through pH alteration, recruitment of growth factors, and others, as well as avoidance of predators (immune suppression).

CHAPTER 3

HEALTHY BONE TISSUE AND MULTIPLE MYELOMA

> Poi che null'altro che vacuo ci resta
> d'ogni cosa ch'esiste,
> Poi che difetto e sconfitta colgono al
> fine ogni cosa,
> Considera bene: ogni cosa che è, è in
> realtà nulla;
> Medita bene: ogni cosa ch'è nulla, è in
> realtà tutto.
>
> Omar Khayyâm

In the following sections we are going to introduce the physiology of the healthy bone tissue and the myeloma-related bone disease. This is not a complete description of these issues: we will focus just on the traits that are essential for understanding the problem and on the traits that can be modelled.

3.1 Bone tissue

Bone tissue is a kind of connective tissue composed by cells and extracellular matrix [90]. Its specificity is the high degree of mineralization, that gives it the distinctive hardness. The bone apparatus constitutes the scaffolding of the human body and protects the more delicate internal organs. More than its intuitive mechanical and protective functions, it has a very important role in the mineral turnover, since it takes part in the Calcium homeostasis. Even after the end of the somatic growth, it is able to modify its structure and architecture as a results of mechanical and hormonal stimuli (*skeletal homeostasis*). So, the bone tissue can be considered truly dynamic and plastic. Moreover, in its cavities is contained the bone marrow, that has haematopoietic functions. The bone is externally covered by the *periosteum*, a fibrous connective with intertwined bundles, excepted from the joints, that are covered by cartilage. In the shafts of long bones, the medullar canal is covered by the *endosteum*, a tiny, loose connective tissue.

The cells that can be found in the bone tissue are *osteoprogenitors*, *osteoblasts*, *osteocytes* and *osteoclasts*; they all derive from the *mesenchymal stem cell*, except from the osteoclasts, that derive from circulating cells (monocytes) originated from the the *granulocyte-monocyte colony-forming unit* (GM-CFU). The *organic component* of the extracellular matrix (both fibrillary and amorphous) is made of type I collagen fibers, glycoproteins, proteoglycans, and enzymes. The present minerals are mainly Calcium phosphates, Calcium carbonate, other salts and residues of Na, K, Sr, Mn, Zn, Cu.

3.1.1 Osteoblasts, ostecytes and osteoprogenitor cells

Osteoprogenitor cells, osteoblasts and osteocytes are three different stages of the same cellular type. The first ones derive from the mesenchymal stem cells, are localized in the periosteum and in the endosteum and are very active in mitosis. Osteoprogenitor are on the dormant surface of the bones. They probably play a role of mediation in the exchange between the blood and the interstitial liquid that flows through the bone lacunae.

Osteoblasts, that differentiate from the osteoprogenitors when it is necessary, are voluminous, mononuclear cells (Fig. 3.1). They have cubic shape and are aligned along the surface of the pre-osseous matrix in the process of deposit. Osteoblasts carry out a strong osteogenic activity: they produce the organic matrix and regulate the deposit of the inorganic one. Most osteoblasts remain walled in the extracellular matrix they have produced and turn into osteocytes. Otherwise, they return in the stage of osteoprogenitor cells.

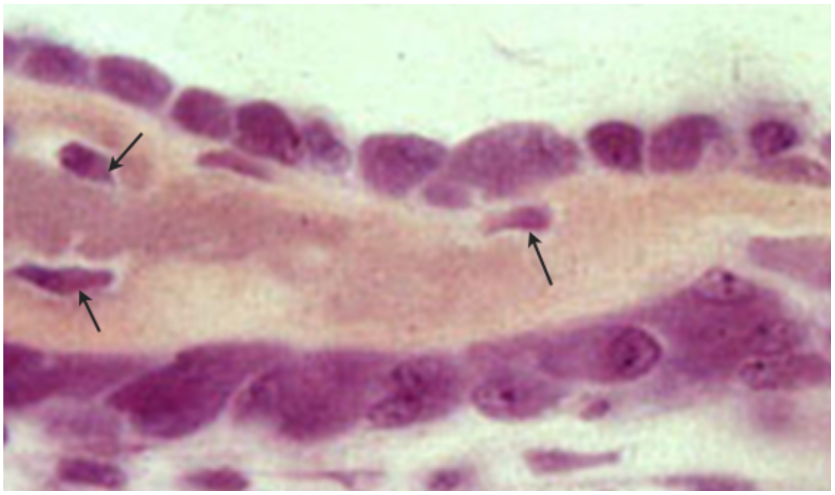


Figure 3.1: **Osteoblasts lining the osseous matrix they have produced.** The arrows indicate osteoblasts entrapped in the mineralized matrix that are becoming osteocytes.

Osteocytes are star-shaped cells that build up the mature bone. When an osteoblast secretes the extracellular matrix, it gets stuck in a lacuna, where it becomes quiescent and takes the name of osteocyte. The main body of the osteocyte remains in the lacuna, but the cell expands its cytoplasmatic extensions in some channels through the matrix, called *canaliculi*, that link the osteocyte to the *Haversian channels*, where the blood vessels flow. Osteocytes do not perform mitosis: the increase of the bone thickness can be realized just by apposition and is never of interstitial kind.

3.1.2 Osteoclasts

Osteoclasts are polynucleated large cells (Fig. 3.2). As already mentioned, osteoclasts result from mononucleated monocytes that merge together to form a syncytium. This event happens in reaction to receptor activator of nuclear factor (NF)-B ligand RANKL. Once formed, osteoclasts are in charge of the bone matrix degradation: they reabsorb the aged, damaged or immature matrix. Osteoclasts are found attached to the bone surface at sites of active bone resorption, where they form a cavity called *resorption bay* or *Howship lacuna*. These cells pursue this activity when, urged by local or systemic stimuli, they adhere to bone surface.

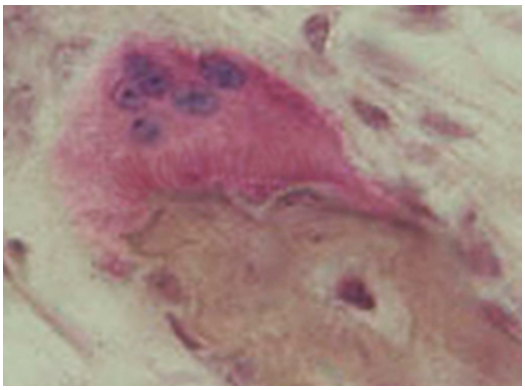


Figure 3.2: **An osteoclast adherent to the osseous matrix for reabsorbing it.** Note the numerous nuclei concentrated at the opposite side to that engaged in reabsorption.

3.1.3 Bone homeostasis

Over the course of the entire life, bones renew themselves, thanks both to resorption and to deposit and the rate of renewal depends on the age. Osteoblasts have receptors for the PTH (*parathyroid hormone*), that induces the secretion of *osteoprotogerin ligand* (OPGL) and of an osteoclasts stimulation factor, that are needed for the differentiation and for the activation and for the differentiation of the osteoclasts.

One of the fundamental regulatory system of bone remodeling is the RANK/RANKL signaling pathway. RANK (*receptor activator of nuclear factor* $(NF) - \kappa B$) is a transmembrane receptor that is expressed on the surface of osteoclast precursors. *RANK ligand* (RANKL) is a cytokine of the *tumour necrosis factor* (TNF) superfamily expressed by *bone marrow stromal cells* (BMSCs) of osteoblastic lineage. When RANKL binds to RANK, the osteoclast precursors fuse together into multinucleated cells that become mature osteoclasts. Once the mature osteoclasts attach to the bone surface, the bone resorption process is initiated. On the other hand, the function of bone resorption and the capability of differentiation are inhibited by the *osteoprotegerin* (OPG), a cytokine of the TNF superfamily produced by BMSCs and osteoblasts, that is a decoy for RANKL and inhibits osteoclastogenesis. The equilibrium between osteogenesis and osteolysis is therefore regulated by the osteoblasts, that adjust the osteoclasts activity, as a response to hormonal stimuli. The hormones involved are PTH and calcitonin, with antagonist effects.

3.2 Multiple myeloma disease

Multiple Myeloma disease (hereinafter simply MM) is the most important plasmacells neoplasia and constitutes the neoplastic proliferation of plasmacells resulted from a single clone [91] [92]. It is characterized by osteolytic bone lesions, hyper-calcaemia, anaemia, kidney failure, acquired immune anomalies and, more rarely, by infections and neurological symptoms. It usually shows up with cancerous masses spread throughout the skeletal system.

There are different kinds of myeloma, besides the MM as such, that is the more aggressive form of them: the *smouldering* myeloma, the *monoclonal gammopathy of undetermined significance* (MGUS), the *lymphoplasmacytic lymphoma* and the *solitary plasmacytoma*, all with different pathological features, staging and prognosis.

The MM aetiology has not yet been clarified. Unfortunately, malignant forms of MM are considered just treatable, but not curable, so they generally lead the patient to death.

3.2.1 Pathogenesis of bone disease in multiple myeloma

MM-induced bone disease is a hallmark of multiple myeloma [93] [94]. At the moment of diagnosis, up to 80% of the patients shows osteolytic bone lesions and has an increased risk of skeletal-related events associated with increased morbidity and mortality. MM cells interact with bone marrow microenvironment and activate molecular cascade. Other tumors, such as breast cancer, prostate cancer and lung cancer, can cause bone metastases, but there is a significant difference: both MM and other tumors show increased osteoclastic bone destruction, but only in MM osteoblast activity is either reduced or absent. Consequently, MM gives rise to the highest incidence of fracture (43%), compared to other types of cancer. Moreover, MM patients who experienced pathological fractures had at least a 20% increased risk of death compared to MM patients without them.

So, osteoclast activity boosts in MM. Besides, histological studies of bone biopsies from MM patients have shown that intensified osteoclast activity occurs adjacent to MM cells. This has led to the hypothesis that local cytokines produced or induced by MM cells are responsible both for the osteoclast formation and for the increased bone resorptive activity. In addition to this, osteoblast functionality is lowered: when MM occurs, bone formation becomes decreased or absent at all. Further, osteoblast apoptosis is increase due to high cytokine levels and physical interaction between osteoblasts and MM cells.

The main feature of MM-related bone disease is the deregulation of the RANK/RANKL/OPG pathway. MM cells produce themselves and stimulate the secretion of RANKL that promotes the process of osteolysis, and also reduce the OPG levels.

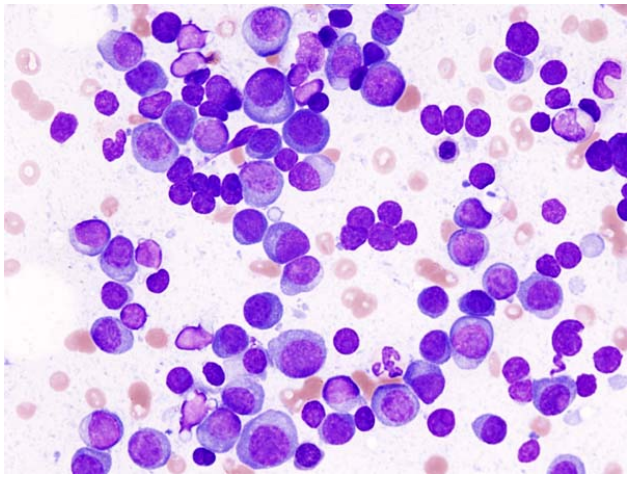


Figure 3.3: **Histopathological image of multiple myeloma cells.**

Below we list the main molecular pathways involved in the pathogenesis of MM-induced bone disease (Fig. 3.4).

Molecular pathways primarily implicated in increased osteoclast activity:

- **Notch pathway**
  The *Notch signaling pathway*, consisting of four trasnmembrane receptors (Notch 1-4) and of their ligands (Jagged 1, 2 and Delta-like 1, 3, 4), is actively implicated in MM-induced osteoclastogenesis: in fact it stimulates the production of RANKL by MM cells and by BMSCs. Moreover, it may facilitate the establishment of the pre-metastatic environment in the bone.

- **Osteopontin**
  Osteopontin is a bone matrix glycoprotein produced by osteoclasts. It is involved in osteoclast activation and in local angiogenesis. In MM patients high levels of this protein are usually associated with high osteolytic activity and advanced disease.

- **CCL-3 (MIP-1α)/CCL-20**
  Chemokine (C-C motif) ligand3 (CCL-3) and *chemokine (C-C motif) ligand 20* (CCL-20) are chemokines produced by MM cells and are implicated in the pathogenesis of MM induced bone disease.

CCL-3 and its receptors (CCR1 and CCR5), that are expressed on BM-SCs, osteoclasts, osteoblasts and MM cells, are involved in osteolytic bone disease and MM cells migration. CCL-20 and its receptor CCR6 are over expressed in the bone marrow and are also responsible for osteoclastogenesis.

- **Activin A**
  Activin A is a member of the TGFβ (*Transforming growth factor beta*) superfamily and may play different roles depending on the tumor microenvironment. Primarily, it induces RANK expression and osteoclasts differentiation.

- **Interleukins**
  Interleukin 3 (IL-3) is a bifunctional cytokine that stimulates osteoclast formation and inhibits osteoblast differentiation. Moreover it stimulates the production of activin A.
  Interleukin 6 (IL-6) is a multifunctional cytokine implicated in bone metabolism: it promotes the osteoclast differentiation and supports the MM cells survival. IL-6 production by marrow stromal cells is strongly induced by the myeloma plasmacells-derived *interleukin 1* (IL-1β).
  Interleukin 17 (IL-17) is a pro-inflammatory cytokine that is mainly secreted by T-helper cells (Th17). It enhances osteoclast activation and osteolytic lesions.

- **TNF superfamily**
  TNF-α is a signaling cytokine that is involved in the pathogenesis of MM bone disease. TNF-α operates together with RANKL in promoting osteoclatogenesis.

- **BTK and SDF-1α**
  Bruton's tyrosine kinase (BTK) is a nonreceptor tyrosine kinase that is involved in osteoclast differentiation. *Stromal cell-derived factor-1α* (SDF-1α) is a chemokine that supports the migration and homing of MM cells. Furthermore, it induces osteoclast activity.

- **Annexin II**
  Annexin II is a calcium-dependent phospolipid-binding member of the annexin family. It supports MM cells adhesion and growth, angiogenesis and osteoclastogenesis.

- **PU.1**
  Pu.1 is a transcriptional factor that has a fundamental role in osteoclast formation.

Molecular pathways primarily implicated in suppressed osteoblast activity:

- **WNT pathway**
  Wingless and integration-1 (WNT) is a signaling pathway that, when physiologically activated, is involved in gene expression responsible for promoting the proliferation, expansion and survival of osteoblastic cells. When, instead, it is aberrantly activated, it takes part in the proliferation of MM cells. Non-canonical WNT pathway is also involved in migration and invasion of myeloma plasmacells.

- **Sclerostin**
  Sclerostin is a protein secreted by osteocytes. In MM patients, sclerostin can be produced by MM cells in bone marrow. It can stimulate apoptosis of mature osteoblasts and prevent osteoblast-driven bone formation. Moreover, it inhibits the binding of *type I* and *type II bone morphogenetic proteins* (BMPs) to their receptor and, consequently, it reduces the BMP-mediated mineralization. Moreover, it increases the RANKL/OPG ratio and, thus, promotes osteoclastogenesis.

- **DKK-1**
  Dickkopf-1 (DKK1) is a member of the DKK family that acts against the WNT pathway. It gets involved in the BMSCs differentiation in mature osteoblasts, interrupting it. Furthermore, undifferentiated BMSCs produce IL-6, that induces the proliferation of MM plasmacells secreting DKK1. DKK1 and sclerostin work together to weaken osteoblast activity. DKK1, like sclerotin,also augments the RANKL/OPG ratio.

- **Periostin**
  Periostin is a disulfide-linked cell adhesion protein that is produced by BMSCs. It is involved in several forms of cancer and MM because it fosters tumor growth and metastasis formation.

- **RUNX2, GFI1 and IL-7**
  Runt-related transcription factor 2/core binding factor Runt domain subunit 1 (RUNX2/CBFA1) is part of the non-canonical WNT signaling pathway and acts as a regulator in osteoblastogenesis. MM cells stop RUNX2 activity in BMSCs and in osteoprogenitor cells and, thus, block osteoblast differentiation.
  Growth factor independence-1 (GFI1) is a transcriptional repressor that binds to RUNX2 and decreases its expression. High levels of GFI1 have been found in MM patients.
  Interleukin 7 (IL-7) downregulates RUNX2 transcriptional activity and so it prevents osteoclast differentiation. IL-7 also encourages the secretion of RANKL.

- **TGFβ and BMPs**
  Inactive form of TGFβ is secreted by osteocytes and osteoblasts in bone matrix and then activates by osteoclasts during bone resorption. When TGFβ pathway is not well regulated, it is involved in the MM bone disease, because it prevents BMSCs from differentiate in mature osteoblasts.
  Bone morphogenetic proteins (BMPs) are included in TGFβ superfamily. MM cells damage BMP-mediated osteoblast differentiation.

- **TNF superfamily**
  As we have already seen, TNF-α supports osteoclastogenesis. Besides that, it inhibits both osteoblast precursor recruitment from progenitor cells and RUNX2, jeopardising osteoblast differentiation.

- **EphB2/EphB4 signaling pathway**
  The Ephrin (Eph) receptors are tyrosine kinase receptors that are activated by ligands called ephrins and their dual function plays a fundamental role in bone homeostasis. On the one hand, EphB2

is expressed in osteoclasts and it is induced by PTH; on the other, EphB4 is expressed in osteoblasts and BMSCs. When EphB2 binds to EphB4, two different signaling cascades occur: the forward one supports osteoblast differentiation, whilst the reverse signaling inhibits osteoclast differentiation. In MM patients it has been found that both EphB2 ad EphB4 expression is reduced in BMSCs.

- **Adiponectin**
  Adiponectin is an adipocyte-derived hormone that is produced both by osteoblasts and BMSCs and that has an important role in bone remodeling; both osteoblasts and osteoclasts express adiponectin receptors. In MM patients, adiponectin deficiency is related to bone lesions.

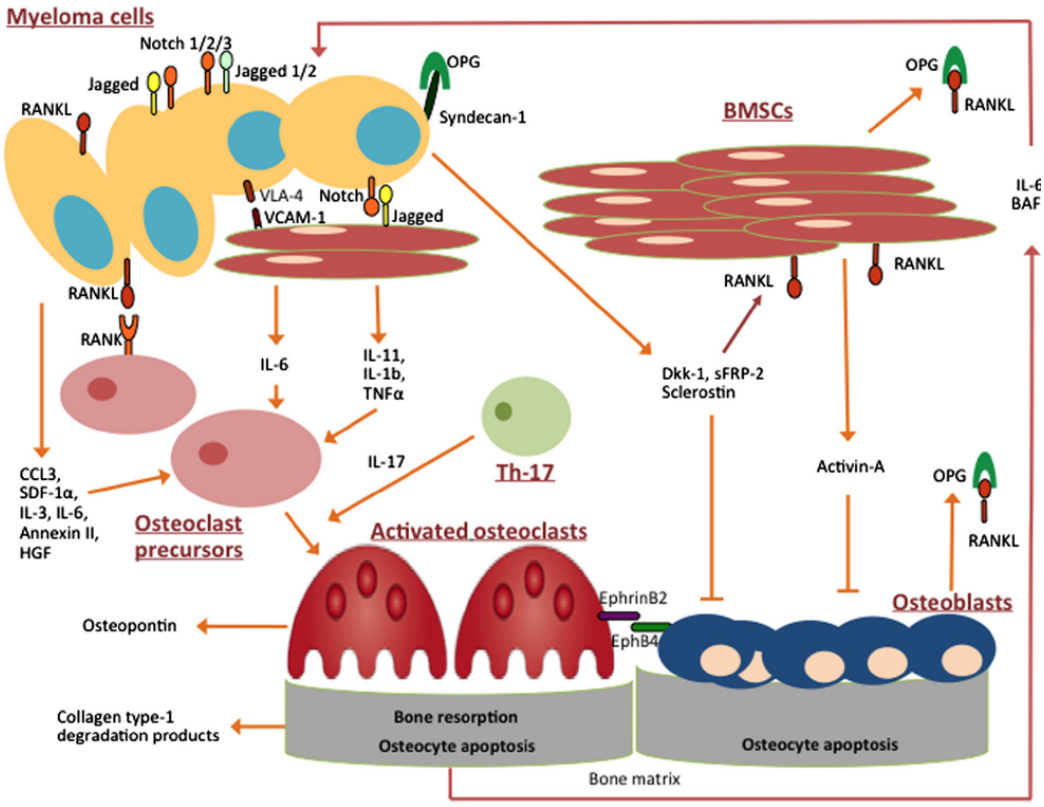


Figure 3.4: **Schematic overview of myeloma-related bone disease and of its characteristic molecular pathways.**

CHAPTER 4

MATHEMATICAL BACKGROUND

Milleseicentodieci, ai dieci di gennaio.
Galileo Galilei
vide che il cielo non c'era.

B. Brecht - Vita di Galileo

In this chapter we are going to shortly illustrate the theoretical background needed to develop our model and to implement the simulations. We are firstly going to define continuous-time Markov processes, that will be our modelling framework. Simply put, Markov chains are stochastic processes for which the future and the past states of the process are independent, once the present state is given. Then, we are going to give a representation of continuous-time Markov chains with finite state space in terms of competing Poisson point processes. Such a representation will be later used in the next chapter to define our model and, moreover, it will be very useful for simulations. In fact, we will portray cancer progression as a set of cells on a lattice, that undergoes a succession of Poisson point processes. In this theoretical framework, it is possible to derive rigorously mean field equations for the dynamics of two classical models from sta-

tistical mechanics: the mean field Ising model, also known as Curie-Weiss model and the Voter model.

For an introduction to the basics concepts of probability and to the Poisson distribution, we suggest to the reader the Appendix A.

4.1 Continuous-time Markov chains

Definition 4.1.1 *A stochastic process* $\{X(t)\}_{t\in\mathbb{R}^+}$ *with values in* $\mathcal{E}$ *is a continuous-time Markov chain if, for all* $i, j, i_1, \ldots, i_k \in \mathcal{E}$*, all* $t, s > 0$ *and all* $s_1, \ldots, s_k \geq 0$ *with* $s_l \leq s$*, for* $l = 1, \ldots, k$*, we have*

$$P(X(t+s) = j | X(t) = i, X(s_1) = i_1, \ldots, X(s_k) = i_k) = P(X(t+s) = j | X(s) = i).$$

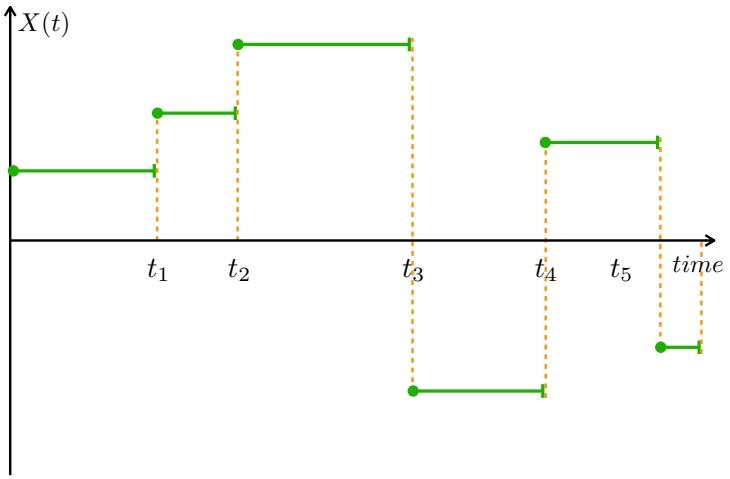


Figure 4.1: **Continuous-time Markov chain.**

Definition 4.1.2 *A continuous-time Markov chain* $\{X(t)\}_{t\in\mathbb{R}^+}$ *is homogeneous if*

$$P(X(t+s) = j | X(s) = i) = P(X(t) = j | X(0) = i).$$

4.1.1 Transition semigroup

Let then

$$\mathbf{P}(t) = \{p_{ij}(t)\}_{i,j\in\mathcal{E}}$$

where

$$p_{ij}(t) = P(X(t+s) = j \mid X(s) = i).$$

The family $\{\mathbf{P}(t)\}_{t\geq 0}$ is the *transition semigroup* of the continuous-time homogeneous Markov chain. Let us make some observations. First of all, it is possible to obtain the Chapman-Kolmogorov[1] equation

$$p_{ij}(t+s) = \sum_{k\in\mathcal{E}} p_{ik}(t)p_{kj}(s),$$

that, in compact form, is

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s).$$

Also

$$\mathbf{P}(0) = I,$$

where I is the identity matrix.

Let the distribution at time t of $X(t)$ be the vector $\mu(t) = \{\mu_i(t)\}_{i\in\mathcal{E}}$ where $\mu_i(t) = P(X(t) = i)$. Then $\mu(t)^T = \mu(0)^T\mathbf{P}(t)$.

The probability distribution of a continuous-time homogeneous Markov chain is entirely determined by its initial distribution and its transition semigroup. That is, for all $t_1, \ldots, t_k$ such that $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$, and for all states $i_0, i_1, \ldots, i_k$

$$P(X(t_k) = i_k, X(t_{k-1}) = i_{k-1}, \ldots, X(0) = i_0) = P(X(0) = i_0)\prod_{j=1}^{k} p_{i_{j-1}i_j}(t_j - t_{j-1}).$$

[1] For a proper derivation of the Chapman-Kolmogorov equation, we suggest the Chapter IV of [95].

4.1.2 Infinitesimal generator

Let $\{\mathbf{P}(t)\}_{t\geq 0}$ be a transition semigroup on $\mathcal{E}$, that is, for each $t, s \geq 0$,

i. $\mathbf{P}(t)$ is a stochastic matrix,

ii. $\mathbf{P}(0) = I$,

iii. $\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s)$.

Suppose, moreover, that the semigroup is continuous at the origin, that is,

$$\lim_{h\to 0} \mathbf{P}(h) = \mathbf{P}(0) = I,$$

where the convergence therein is pointwise and for each entry.

Theorem 4.1.1 (Local Characteristics) *Let $\{\mathbf{P}(t)\}_{t\geq 0}$ be a continuous transition semigroup on the countable state space $\mathcal{E}$. For any state i, there exists*

$$q_i = \lim_{h\to 0} \frac{1 - p_{ij}(h)}{h} \in [0, \infty]$$

and for any pair i, j of different states, there exists

$$q_{ij} = \lim_{h\to 0} \frac{p_{ij}(h)}{h} \in [0, \infty).$$

For a proof, see [96], Chapter VIII.

Definition 4.1.3 *The numbers q_{ij} are called the local characteristics of the semigroup, or of the corresponding continuous-time homogeneous Markov chain. The matrix*

$$\mathbf{L} = (q_{ij})_{i,j\in\mathcal{E}}$$

is called the infinitesimal generator of the semigroup, or of the continuous-time homogeneous Markov chain.

In compact notation,

$$\mathbf{L} = \lim_{h\to 0} \frac{\mathbf{P}(h) - \mathbf{P}(0)}{h},$$

where the meaning of this equation is given by the ones of Th.4.1.1. Thus, in this sense, the infinitesimal generator $\mathbf{L}$ is the derivative at 0 of the function $t \mapsto \mathbf{P}(t)$.

4.2 Finite state space

In view of the semigroup properties, for all $t \geq 0$ and all $h \geq 0$

$$\frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} = \mathbf{P}(t)\frac{\mathbf{P}(h) - I}{h} = \frac{\mathbf{P}(h) - I}{h}\mathbf{P}(t),$$

therefore, if the passage to the limit is allowed, which is the case when the state space $\mathcal{E}$ is finite, we obtain a differential system

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{L} = \mathbf{L}\mathbf{P}(t),$$

where $\mathbf{L}$ is the infinitesimal generator.
The equation

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{L}\mathbf{P}(t),$$

can be written explicitly. For all $i, j \in \mathcal{E}$,

$$\frac{d}{dt}p_{ij}(t) = -q_i p_{ij}(t) + \sum_{k \in E, k \neq i} q_{ik} p_{kj}(t).$$

Systems like these are Kolmogorov's *backward* differential systems.
The *forward* differential system is

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{L},$$

that is, for all $i, j \in \mathcal{E}$,

$$\frac{d}{dt}p_{ij}(t) = -p_{ij}(t) q_i + \sum_{k \in E, k \neq i} p_{kj}(t) q_{ik}.$$

For a continuous-time Markov chain with finite state space, for all states i, it holds

$$q_i = \sum_{j \in \mathcal{E}, j \neq i} q_{ij}.$$

This implies also that, for $f(i) : \mathcal{E} \to \mathbb{R}$

$$\mathbf{L}f = \mathbf{L}f(j) = \sum_{i \in \mathcal{E}} q_{ji} f(i) = \sum_{j \neq i} q_{ji}[f(i) - f(j)].$$

This means that the chain is determined by the generator, so these two concepts can be identified.

Theorem 4.2.1 *Let $\{X_k(t)\}_k$ be a sequence of Markov processes with values in $\mathcal{X}_k$ and denote by $\mathbf{L}_k$ the corresponding infinitesimal generators, defined on $\mathcal{D}(\mathbf{L}_k)$. Moreover, let $\mathbf{L}$ be the infinitesimal generator of another Markov process $X(t)$ with values on $\mathcal{X}$ and let $\mathcal{C}$ be the core for $\mathbf{L}$. Assume that, for every k, $\mathcal{X}_k \subseteq \mathcal{X}$. And each function in $\mathcal{C}$ is an element of $\mathcal{D}(\mathbf{L}_k)$, when restricted to $\mathcal{X}_k$. If the condition*

$$\lim_{k \to \infty} \sup_{x \in \mathcal{X}_k} |\mathbf{L}_k(f(x)) - \mathbf{L}(f(x))| = 0$$

holds for every $f \in \mathcal{C}$ and $X_k(0)$ converges to $X(0)$ in distribution, then the sequence of processes $\{X_k(t)\}_k$ converges to the process $X(t)$ in distribution as $k \to \infty$.

For a proof, see [97].

4.3 Point processes

A *random point process* is, roughly speaking, a countable random set of points on the real line [96]. A *point* of a point process is the time of occurrence of some event, and this is why points are also called *events*.

Definition 4.3.1 *A random point process on the positive half-line is a sequence $\{T_n\}_{n \geq 0}$ of nonnegative random variables such that*

i. $T_0 \equiv 0$,

ii. $0 < T_1 < T_2 < \dots$,

iii. $\lim_{n \to \infty} T_n = +\infty$.

Condition (*ii*) may be relaxed and thus multiple point allowed. When condition (*ii*) holds, one speaks of *simple point process*. Also, condition (*iii*) is not required in the more general definition, where it may occur that $P(\lim_{n\to\infty} T_n < +\infty)$: with positive probability there is an *explosion*, that is an accumulation of events in finite time.

For any interval $(a, b]$ in $\mathbb{R}^+$, we define

$$N((a, b]) = |T \cap (a, b]| = \sum_{n\geq 1} 1_{(a,b]}(T_n),$$

that is an integer-valued random variable counting the events occurring in the time interval (a,b]. For sake of simplicity, it will be usually denoted by $N(a, b]$. For $t \geq 0$, let us use the notation

$$N(t) = N(0, t].$$

In particular, $N(0) = 0$ and $N(a, b] = N(b) - N(a)$. The family of random variables $\{N(t)\}_{t\in\mathbb{R}^+}$ is called the *counting process* of the point process $\{T_n\}_{n\geq 1}$. Since the sequence of events can be recovered from $N(t)$, also the latter will be called point process.

4.4 Poisson processes

A Poisson process can be defined in different ways. We will introduce two of them: the first one clearly shows the link with random point processes, meanwhile the second, focusing on what happens in the infinitesimal time dt, offers a more intuitive view of this process, useful with the prospect of application to our model.

Definition 4.4.1 *Let $\{N(t)\}_{t\in\mathbb{R}^+}$ be a point process. It is a homogeneous Poisson process with intensity $\lambda > 0$ if*

i. for all times t_i, $i \in [1, k]$, such that $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$, the random variables $N(t_i, t_{i+1}]$, $i \in [1, k-1]$, are independent,

ii. for any interval $(a, b] \subset \mathbb{R}^+$, $N(a, b]$ is a Poisson random variable with mean $\lambda(b - a)$.

Thus, for all $k \geq 0$,

$$P(N(a,b] = k) = e^{-\lambda} \frac{[\lambda(b-a)]^k}{k!}$$

and in particular,

$$E[N(a,b]] = \lambda(b-a).$$

In this sense, λ *is the average density of points.*

Condition (*i*) is the property of *independence of increments* of Poisson processes. It implies in particular that for any interval $(a,b]$, the random variable $N(a,b]$ is independent of $(N(s), s \in (0,a])$. The increments of homogeneous Poisson processes have no memory of the past.

A second way to define Poisson processes is given by substituting the *(ii)* of Def. 4.4.1 with

$$P(|T \cap (t, t+dt]| = 1) = \lambda dt + o(dt)$$

where $o(dt)$ means that $\lim_{dt \to 0} \frac{o(dt)}{dt} = 0$.

4.4.1 Competing Poisson processes

Let $\{T_n^1\}_{n \geq 1}$ and $\{T_n^2\}_{n \geq 1}$ be two independent homogeneous Poisson processes on $\mathbb{R}^+$ with respective intensities $\lambda_1 > 0$ and $\lambda_2 > 0$. Their superposition is defined to be the sequence obtained by merging $\{T_n^1\}_{n \geq 1}$ and $\{T_n^2\}_{n \geq 1}$ (see Fig.4.2). It is noted that

- the point processes $\{T_n^1\}_{n \geq 1}$ and $\{T_n^2\}_{n \geq 1}$ have no points in common;
- the point process $\{T_n\}_{n \geq 1}$ is an homogeneous Poisson process with intensity $\lambda = \lambda_1 + \lambda_2$.

The above statement can be extended to several, possibly an infinity of, homogeneous processes as follows.

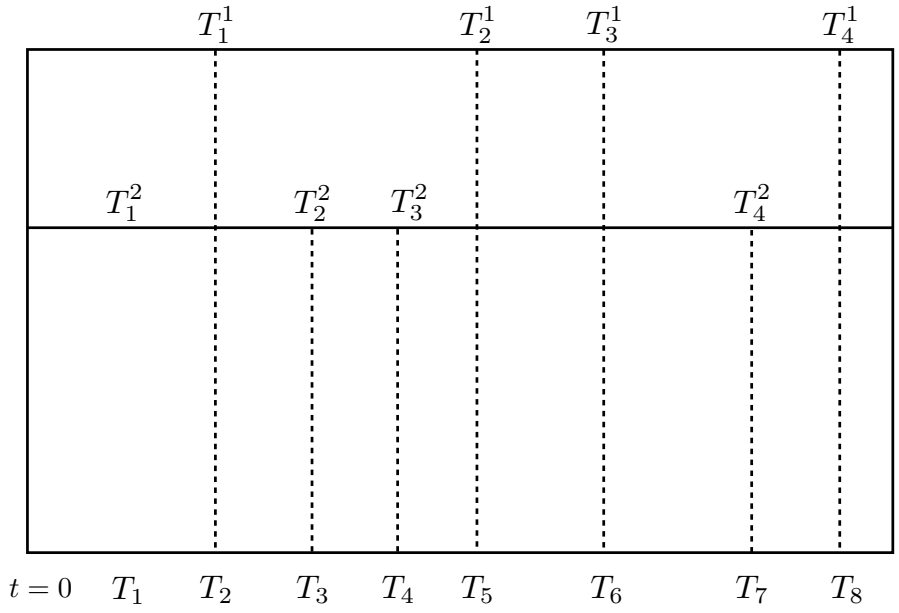


Figure 4.2: **Superposition, or sum, of two point processes.**

Theorem 4.4.1 *Let $\{N_i\}_{i\geq 1}$ be a family of independent homogeneous Poisson processes with respective positive intensities $\{\lambda_i\}_{i\geq 1}$. Then*

i. two distinct homogeneous Poisson processes of this family have no points in common,

ii. If

$$\sum_{i=1}^{\infty} \lambda_i = \lambda < \infty$$

then

$$N(t) = \sum_{i=1}^{\infty} N_i(t)$$

defines the counting process of a homogeneous Poisson process with intensity λ.

For a proof, see [96], Chapter VIII.

Let us now see a central result of the theory of Poisson system:

Theorem 4.4.2 (Competition Theorem) *Let $\{N_i\}_{i\geq 1}$ be a family of independent homogeneous Poisson processes with respective positive intensities $\{\lambda_i\}_{i\geq 1}$, where $\sum_{i=1}^{\infty}\lambda_i = \lambda < \infty$. Denote by Z the first event time of $N = \sum_{i=1}^{\infty} N_i$ and by J the index of the homogeneous Poisson process responsible for it; in particular, Z is the first event of N_J. Then*

$$P(J = i, Z \geq a) = P(J = i)P(Z \geq a) = \frac{\lambda_i}{\lambda}e^{-}.$$

In particular, J and Z are independent, $P(J = 1) = \lambda_i/\lambda$ and Z is exponential with mean λ^{-1}.

For a proof, see [96], Chapter VIII.

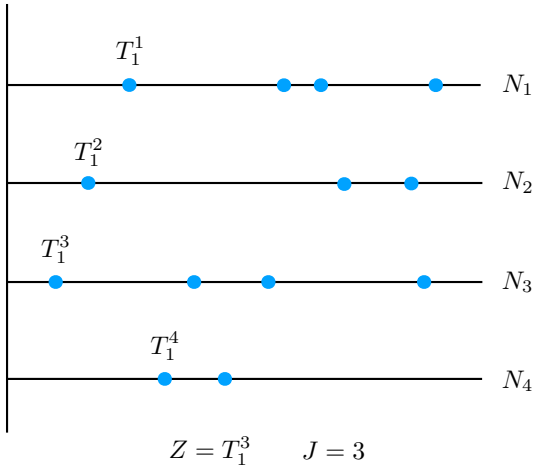


Figure 4.3: **Competition among four point processes.**

Furthermore, for a continuous-time Markov chain we have

$$P(X(t+h) = i \mid X(t) = i) = 1 - q_i h + o(h)$$

and, if $i \neq j$,

$$P(X(t+h) = j \mid X(t) = i) = q_{ij}h + o(h).$$

Once introduced the basic definitions above, we would like to give representation of Markov chain, related to the Competition Th.4.4.2 and hereafter we follow the notation set there), that could be particularly useful to understand the simulations of our model. Let $\{X(t)\}_{t\in\mathbb{R}^+}$ be a regular jump chain taking its values in $\mathcal{E} = \{1, 2, 3, \dots\}$. Suppose to start from $X(0) = k_0$ and indicate with $K_0^{\neq} = \{j_1, j_2, \dots\}$, the set of states that are different from k_0. Let's take N_{j_i} with intensity $q_{k_0 j_i}$. We note that, due to the fact that this is a regular jump process, the chain is stable and conservative, that is

$$\sum_{j\neq k_0} q_{k_0 j_i} = q_{k_0} < +\infty$$

Let set the dynamics of the chain as follows

- if $Z \in N_{j_i}$, then $X(t) = j_i$ till the next jump;
- there is competition between the Poisson processes with $\lambda_z = q_{j_i z}$, where $z \neq j_i$.

Let us calculate the transition semigroup of the chain

$$\begin{aligned} p_{ij}(h) &= P(z \in (0,h))P(J = j) \\ &= q_i \frac{q_{ij}}{q_i} + o(h) \\ &= q_{ij}. \end{aligned}$$

From here on out, we will refer to q_{ij} as the transition rate from state i to state j.

4.5 Interacting particle systems

Interacting particle systems are countable systems of locally interacting Markov processes. They are usually defined on a d-dimensional integer lattice: at each site of this lattice there is situated a continuous-time Markov process with a finite state space, whose transition rates depend on the states of the Markov processes on the near-by sites. For a deeper insight in this field, see [98] [99] [100].

Now we will briefly explain two paradigmatic examples of interacting particle systems: the Curie-Weiss model and the voter model.

4.6 Curie-Weiss model

The Curie-Weiss model [101] is one of the simplest classical systems of ferromagnetism: it corresponds to the mean field theory of the well known Ising model. It is also considered, in its dynamical version, as an example of interacting particle system.

Let us consider a set of N spins: we write $\{\sigma(t)\}_{t\in[0,T]}$ for the stochastic process defined as follows. If the state space of the system, at an arbitrary time t, is defined by the arbitrary configuration

$$\sigma = (\sigma_1, \dots \sigma_N) \quad \text{where} \quad \sigma_i = \{-1, +1\},$$

the dynamics is described by the probability of transition, in an infinitesimal time dt, from the previous state to

$$\sigma^i = (\sigma_1, \sigma_2, \dots, -\sigma_i, \dots, \sigma_N),$$

where the $i-$th spin is flipped. This probability is given by the product of the transition rate for dt

$$P(\sigma(t+dt) = \sigma^i \mid \sigma(t) = \sigma) = \omega_i \, dt = e^{-\sigma_i(t)\beta m_N(t)} dt,$$

where $\beta > 0$ is a parameter and $m_N = \frac{1}{N}\sum_{i=1}^{N} \sigma_i(t)$ is the magnetization.

The dynamics described above is ferromagnetic (or cooperative), in the sense that

- if σ_i and m_N have different sign (simply put, they are not aligned), the exponent of the transition rate is positive, $-\sigma_i(t)\beta m_N(t) > 0$. This means that the transition rate is high.
- if σ_i and m_N have the same sign (simply put, they are aligned), the exponent of the transition rate is negative, $-\sigma_i(t)\beta m_N(t) < 0$. This means that the transition rate is low.

Theorem 4.6.1 *As $N \to \infty$, the process $\{m_N(t)\}_{t\in[0,T]}$ converges in distribution to the solution of*

$$\begin{cases} \dot{m}(t) &= 2\sinh\left[\beta m(t)\right] - 2m(t)\cosh\left[\beta m(t)\right]; \\ m(0) &= m_0. \end{cases}$$

Using the theory developed in this chapter, we will briefly sketch the proof of this theorem. Let us consider the transition between the previous described states: from σ to σ^i

$$\sigma_i(t) \to -\sigma_i(t+dt).$$

Using the results of the section 4.4.2, we can write the infinitesimal generator for the markovian dynamics of m_N, that is

$$\mathbf{L}_N f(m_N) = \sum_{i=1}^{N} e^{-\sigma_i \beta m_N} \left[f\left(m_N - \frac{2\sigma_i}{N} \right) - f(m_N) \right].$$

Then, using the Th.4.2.1, we read the limiting process from the limit of $\mathbf{L}_N$, as N goes to infinity. We have

$$\begin{aligned}
\mathbf{L}_N f(m_N) =& \sum_{i=1}^{N} e^{-\sigma_i \beta m_N} \left[f\left(m_N - \frac{2\sigma_i}{N} \right) - f(m_N) \right] \\
=& \sum_{i=1}^{N} e^{-\sigma_i \beta m_N} \left[-\frac{2\sigma_i}{N} \frac{\partial f}{\partial x}\bigg|_{x=m_N} + o\left(\frac{1}{N}\right) \right] \\
=& \sum_{i:\sigma_i=1} e^{-\beta m_N} \left[-\frac{2}{N} \right] \frac{\partial f}{\partial x}\bigg|_{x=m_N} \\
&+ \sum_{i:\sigma_i=-1} e^{\beta m_N} \left[\frac{2}{N} \right] \frac{\partial f}{\partial x}\bigg|_{x=m_N} + o\left(\frac{1}{N}\right) \\
=& \sum_{i=1}^{N} \frac{1+\sigma_i}{2}\, e^{-\beta m_N} \left[-\frac{2}{N} \right] \frac{\partial f}{\partial x}\bigg|_{x=m_N} \\
&+ \sum_{i=1}^{N} \frac{1-\sigma_i}{2}\, e^{\beta m_N} \left[\frac{2}{N} \right] \frac{\partial f}{\partial x}\bigg|_{x=m_N} + o\left(\frac{1}{N}\right) \\
=& \left(e^{-\beta m_N} - m_N e^{-\beta m_N} + e^{\beta m_N} - m_N e^{\beta m_N} \right) \frac{\partial f}{\partial x}\bigg|_{x=m_N} + o\left(\frac{1}{N}\right) \\
=& \left(2\sinh\left[\beta m_N\right] - 2m_N \cosh\left[\beta m_N\right] \right) \frac{\partial f}{\partial x}\bigg|_{x=m_N} + o\left(\frac{1}{N}\right).
\end{aligned}$$

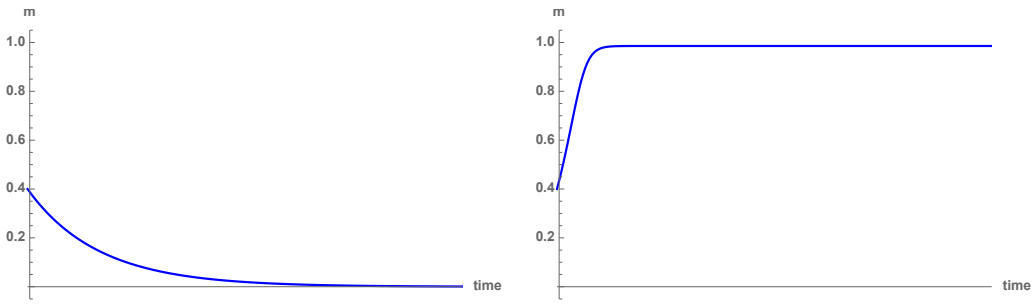
Hence, the limiting generator is

$$\mathbf{L} f(m_N) = \left(2\sinh\left[\beta m_N\right] - 2m_N \cosh\left[\beta m_N\right] \right) \frac{\partial f}{\partial x}\bigg|_{x=m_N},$$

that corresponds to the process

$$\dot{m}(t) = 2\sinh\left[\beta m(t)\right] - 2m(t)\cosh\left[\beta m(t)\right].$$

The time evolution of the magnetization m, that is called *order parameter* of the system, has different qualitative behaviours (Fig. 4.4), varying the values of the parameter β: these phenomenon is called *phase transition*. The threshold value for the phase transition is $\beta = 1$.



(a) $\beta < 1$ (b) $\beta > 1$

Figure 4.4: **Curie-Weiss model - Magnetization** m_N.

4.7 Voter model

The voter model [102] [103] is a toy spin system that has been firstly used to describe the evolution of opinions in a spineless population. Each individual resides on a site of an arbitrary network and can assume one of q equivalent opinion state. The basic picture occurs when each voter can be one of two state, for example, 1 or 0. A voter is selected at random and adopts the opinion of a randomly chosen of its neighbors. This update is repeated at unit rate until the finite population of N individuals reaches consensus. It is important to remark that each agent is influenced only by a fixed set of neighbours and that there are not other types of interactions.

Let us explain more in details the case of a voter model on a lattice with just two opinions. Thus, at each lattice site i, the opinion state of the voter, η_i, can be in one of two states, 0 or 1. Each spin flips at a rate that equals that fraction of its neighbours in the opposite opinion state (Fig. 4.5).

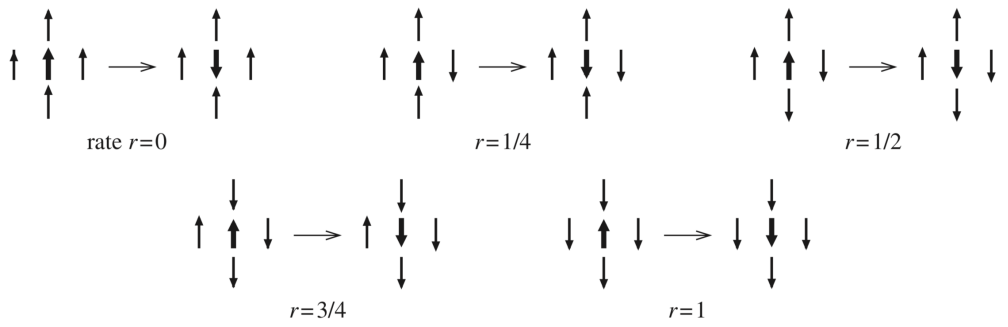


Figure 4.5: **The flip rates for the central voter as a function of the state of its local neighborhood on the square lattice.**

The flipping rule can be encoded by the rate w_i at which this voting in i state changes is

$$w_i(\eta(t)) = \frac{1}{2}\left(1 - \frac{\eta_i(t)}{z}\sum_{j\sim i}\eta_j(t)\right), \tag{4.1}$$

where z is that lattice coordination number and $j \sim i$ means that the sum is over the nearest neighbors of site i. Furthermore, the index i in the rate w_i points out that only the voter at site i changes opinion in an update event. A crucial feature of this transition rate is its linearity in the number of disagreeing neighbors, that underlies the solvability of the voter model.

Now we want to describe the mean field approximation of voter model. Let us consider a set of N spins: we write $\{\eta(t)\}_{t\in[0,T]}$ for the stochastic process defined as follows.

Let us assume that the state space of the system, at an arbitrary time t, is defined by the arbitrary configuration

$$\eta = (\eta_1, \ldots, \eta_N) \quad \text{where} \quad \eta_i = \{0, 1\}.$$

If we denote $\eta_i = 0$ with η_0 and $\eta_i = 1$ with η_1, the dynamics is described by

$$P(\eta(t+dt) = \eta_0 \mid \eta(t) = \eta) = \bar{\eta}_0 = \frac{1}{N}\sum_{i=1}^{N} \delta_{\eta_i}(0),$$
$$P(\eta(t+dt) = \eta_1 \mid \eta(t) = \eta) = \bar{\eta}_1 = \frac{1}{N}\sum_{i=1}^{N} \delta_{\eta_i}(1).$$

It is worthwhile to notice that, because in the voter model the probability of transition is determined by the number of individual of each type, the transition rates correspond to the mean values of individuals of each species

$$\omega_k(\eta) = \bar{\eta}_k = \frac{1}{N}\sum_{i=1}^{N} \delta_{\eta_i}(k).$$

Such as for the Curie-Weiss model, we can write the infinitesimal generator for the markovian dynamics of $\bar{\eta}_k$

$$\mathbf{L}_N f(\bar{\eta}_k) = \sum_{i=1}^{N} \Big\{ \bar{\eta}_0 \left[f(\eta_i = 0) - f(\eta) \right] + \bar{\eta}_1 \left[f(\eta_i = 1) - f(\eta) \right] \Big\}.$$

Then, using the Th.4.2.1, we obtain the limiting generator

$$\begin{aligned}
\mathbf{L}_N f(\bar{\eta}_0) &= \sum_{i=1}^{N} \Big\{ \bar{\eta}_0 \left[f(\eta_i = 0) - f(\eta) \right] + \bar{\eta}_1 \left[f(\eta_i = 1) - f(\eta) \right] \Big\} \\
&= \sum_{i:\eta_i=1} \bar{\eta}_0 \left(-\frac{1}{N} \right) \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_0} + \sum_{i:\eta_i=0} \bar{\eta}_1 \left(\frac{1}{N} \right) \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_0} \\
&= -\bar{\eta}_0 \left(\frac{1}{N}\sum_{i=1}^{N} \delta_{\eta_i}(1) \right) \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_0} \\
&\quad + \bar{\eta}_1 \left(\frac{1}{N}\sum_{i=1}^{N} \delta_{\eta_i}(0) \right) \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_0} \\
&= \left(-\bar{\eta}_0\, \bar{\eta}_1 + \bar{\eta}_1\, \bar{\eta}_0 \right) \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_0} \\
&= 0.
\end{aligned}$$

The same holds for $\bar{\eta}_1$, thus, summing up, the limiting generators are

$$\mathbf{L}_N f(\bar{\eta}_0) = 0,$$
$$\mathbf{L}_N f(\bar{\eta}_1) = 0,$$

that correspond to

$$\dot{\bar{\eta}}_0(t) = 0, \qquad \dot{\bar{\eta}}_1(t) = 0.$$

Thus, in average, fraction of each population remains constantly equal to its initial value. Nevertheless when N is finite, $\bar{\eta}_1 = 0$ and $\bar{\eta}_1 = 1$ are absorbing states for the microscopic dynamics and in the long run the chain will end in one of these states. In such a situation the quantity of interest is the mean time to absorption. For the simple voter model it scales polynomially with system size N while in more structured models, the same quantity may grow exponentially in N.

CHAPTER

5

OUR MODEL FOR THE PATHOGENESIS OF MM-INDUCED BONE DISEASE

> Stochastic evolutionary models of carcinogenesis are unique in the sense that they combine a large degree of reductionism (the simplicity of the model setting) with a high degree of analytical tractability, simultaneously providing insights into unknown aspects of carcinogenesis. [...] Models of this type have provided a window into the microscopic tumor dynamics.
>
> T. Jackson et al. [21]

Multiple myeloma is not a tumor of bone cells, thus, is not possible to study its rise up as a mutation of pre-existing cells during mitosis. A better way of describe the pathology initiation is in term of a new species that attempts to invade a resident species of normal cells. After that, the

development of cancer in the environment of healthy bone tissue can be modelled as the interaction between cell populations. For this reason, an ecological point of view over the problem is preferable to others.

5.1 Same problem, different approach

Our work was inspired by a paper [106] by J. Pacheco and his collaborators, in which they studied the pathogenesis of multiple myeloma bone disease from an evolutionary game theory (EGT) perspective. In this paragraph we sum up their work, without going through the details, in order to clarify the different method followed.

For a more complete understanding, we suggest to the reader the Appendix B.

> Understanding the dynamics of tumour growth and response to therapy is incomplete unless the interactions between the malignant cells and normal cells are investigated in the environment in which they take place.[...] Such processes impose costs and benefits to the participating cells that may be conveniently recast in the form of a game pay-off matrix. As a result, tumour progression and dynamics can be described in terms of evolutionary game theory, which provides a convenient framework in which to capture the frequency-dependent nature of ecosystem dynamics. [106]

First of all, the authors tried to model the phenomenon of normal bone remodeling. Due to the fact that osteoclast-mediated (OC) bone resorption and bone formation due to osteoblast (OB) activity are balanced, the best way to reproduce their interplay is a *co-existence game*, that can be realized by a pay-off matrix like this

$$\begin{array}{c} \\ \text{OC} \\ \text{OB} \end{array} \begin{array}{c} \begin{array}{cc} \text{OC} & \text{OB} \end{array} \\ \begin{bmatrix} 0 & a \\ e & 0 \end{bmatrix} \end{array}$$

where e and a are both positive reals, that means that the interactions between OB and OC are stronger than self-interactions.

Indeed, we have:

$$\phi_{OC}(x) = (1-x)\, a,$$
$$\phi_{OB}(x) = x\, e,$$

where x stands for the fraction of OC cells and $(1-x)$ for the fraction of OB cells. The replicator equation, then, reads

$$\dot{x} = x\,(1-x)\,(a\,(1-x) - e\,x).$$

This dynamics has three fixed points: $x^* = 0$, $x^* = 1$, both unstable, and $x^* = a/(a+e)$, that is stable and represents the balance of OB and OC[1]. When multiple myeloma (MM) occurs, as we have already explained, MM cells have a net benefit in the interaction, mediated by different molecular pathways, with OC and OB. This condition, illustrated by Fig. 5.1, is described by this pay-off matrix

$$\begin{array}{c} \\ \text{OC} \\ \text{OB} \\ \text{MM} \end{array} \begin{array}{c} \begin{array}{ccc} \text{OC} & \text{OB} & \text{MM} \end{array} \\ \begin{bmatrix} 0 & a & b \\ e & 0 & -d \\ c & 0 & 0 \end{bmatrix} \end{array}$$

where all parameters are non negative. The minimal pay-off matrix, that is the matrix with the smallest number of parameters compatible with this dynamics, is

$$\begin{array}{c} \\ \text{OC} \\ \text{OB} \\ \text{MM} \end{array} \begin{array}{c} \begin{array}{ccc} \text{OC} & \text{OB} & \text{MM} \end{array} \\ \begin{bmatrix} 0 & 1 & \beta \\ 1 & 0 & -\delta \\ \beta & 0 & 0 \end{bmatrix} \end{array}$$

The dynamics of two population, such as the OB-OC dynamics, could be represented along the segment $0 \leq x \leq 1$. The OB-OC-MM dynamics, instead, proceeds in a two-dimensional space, called simplex, and thus

[1] that is $\phi_{OB}(x^*) = \phi_{OC}(x^*)$.

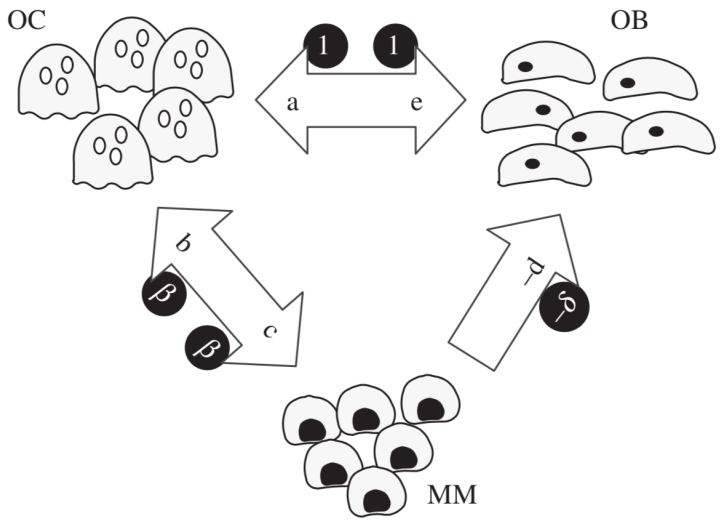


Figure 5.1: **Pathological bone turnover - EGT model**

could be depicted by an equilateral triangle, each vertex of which represents a monotypic population. Edges of the simplex represent population configurations in which at least one of the cell types is missing. The interior of the simplex, in turn, corresponds to configurations of population in which all cell types coexist, albeit with different fractions in different points. The number and nature of the fixed points in the simplex will naturally depend on the relative balance between β and δ in the pay-off matrix:

- if $\beta < 1$, MM cells can go extinct and OB and OC may again re-establish the stable dynamic equilibrium (Fig. 5.2 **a**).
- if $\beta < 1$ and $\beta + \delta > 1$ here appears a typical saddle point structure in the interior of the simplex, which still ensures that normal homoeostasis is a possible 'end-game' of the coevolutionary process. In this case, therapies that change β may provide important contributions to overall disease eradication. (Fig. 5.2 **b**)
- if $\beta > 1$, the OB population is drastically lowered, and the only stable equilibrium is the coexistence of OC and MM cells. (Fig. 5.2, **c**)

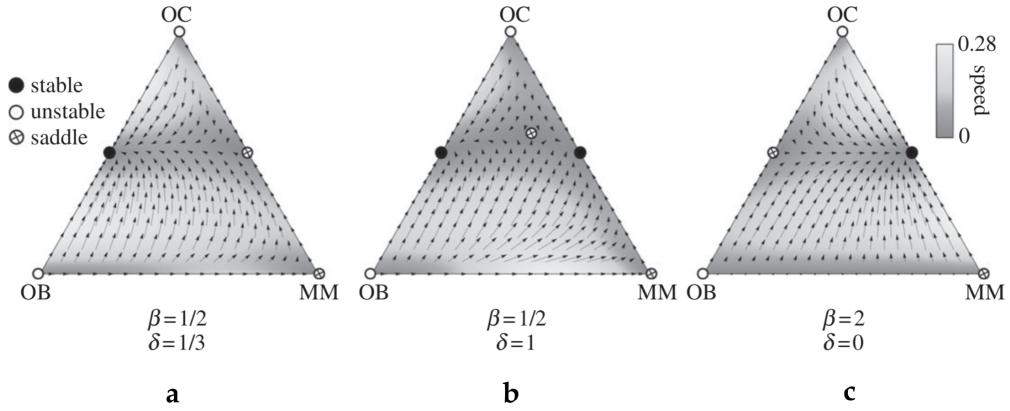


Figure 5.2: **Evolutionary dynamics of OC, OB and MM cell types - EGT model**

5.2 From a real phenomenon to a physical model

As we have shown in the third chapter, the pathogenesis of multiple myeloma bone disease is a very complex phenomenon. Of course, modelling every details is an impossible task, so we have made some reasonable simplifications. Many of them are already present in the paper of J. Pacheco [106], but we prefer to explain and underline them, in order to better understand our work.

We have considered just three cell populations: osteoblasts (OB), osteoclasts (OC) and multiple myeloma cells (MM). Thus, we have left out all the intermediate stages, such as osteoprogenitor cells or osteocytes.

To model the derivation of OB and OC, we have supposed the presence of two sources, initially with different rates of production for the two kind of cells.

We have pictured the interplay between OB and OC in the healthy bone tissue as cooperation. After the onset of myeloma, we have, on the one hand, summed up all the molecular pathways that enhance osteoclasts activity as cooperation between MM and OC. On the other, the fact that MM cells suppress osteoblasts activity is modelled as an increased death rate for OB when MM are present. In this initial model, each interaction could have different intensities depending on the direction.

Once the rate transition and the equations have been set, we made a further simplifications to make the simulations: we suppose that the sources of OB and OC are identical and that all the interactions of cooperation are specular.

5.3 Our model

We modelled the phenomenon of multiple myeloma bone disease as an interacting particle system. First of all, we described the healthy bone tissue, with two species only (OB and OC), to verify that these coexist and cooperate to maintain the tissue physiological equilibrium. Then we introduced in that framework the malignant cells of myeloma, modelling a system with three cellular species.

The system is defined over a two-dimensional square lattice, at each site of which there is a cell that undergoes a continuous-time Markov process.

An algorithmic description (Fig. 5.3) of this model update rule is the following

i. at time t, a random cell is picked with the same probability for each site (for the case with MM cells this point is slightly modified, as we will explain later);

ii. the selected cell is removed and substituted by another one, chosen depending on some transition rate;

iii. the time is incremented, $t \to t + dt$.

In other words, at each time step, in a site of the lattice can occur a point of a Poisson process, whose sequence determines the time evolution of cell populations.

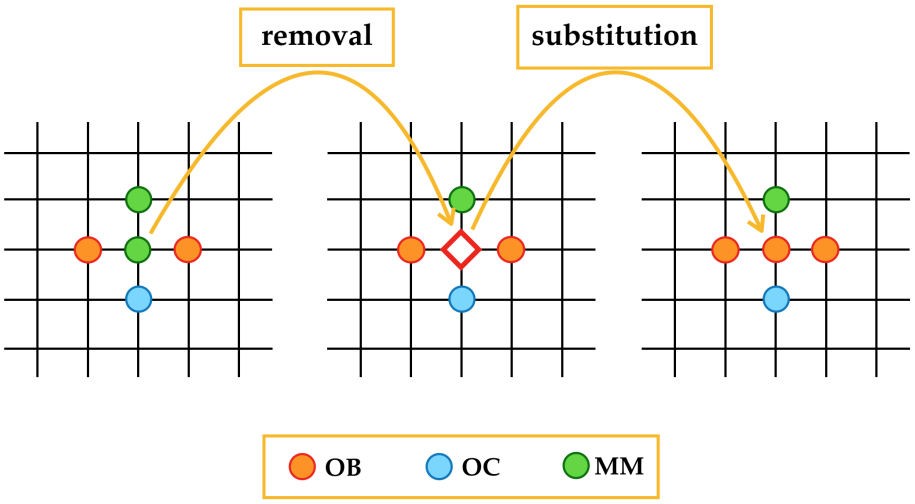


Figure 5.3: **Cell removal and substitution**

Let us introduce some notations.

The i-th empty site will be denoted with $\diamondsuit_i$.

For sake of simplicity, we will use B for indicate osteoblasts (OB), C for osteoclasts (OC) and M for multiple myeloma cells (MM) (e.g., n_B will stand for the number of OB).

Concentrations will be denoted as

$$\bar{\eta}_K(t) = \frac{n_K(t)}{n_{TOT}},$$

where $K = B$ (for OB), C (for OC) , M (for MM) and $n_{TOT} = n_B + n_C + n_M$.

The transition rate of $(K \rightarrow \diamondsuit_i \rightarrow L)$ will be ω_{KL}, where $K, L = B, C, M$. Of course there is no transition between, for example, OC and OC, so the terms with $K = L$ will be equal to zero.

As a first step, we study the time evolution of cell populations over the lattice in the mean field approximation, that is, when the rates are calculated over the whole lattice. In other words, each cell interacts with all the others and there is no concept of space. Then, we make a more realistic step: we develop a local model, supposing that each cell interacts only with its nearest neighbours.

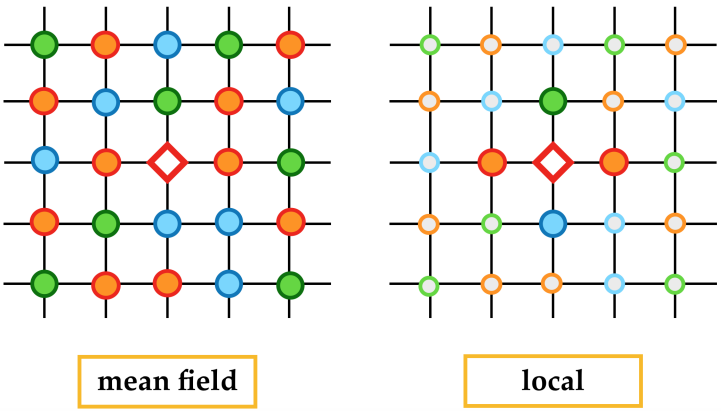


Figure 5.4: **Mean field and local configuration**

For the mean field case, in analogy with what we have done for the Curie-Weiss model and for the voter model, we can write the space state of the system as $\{\eta(t)\}_{t\in[0,T]}$. Let us assume that the space state of the system, at an arbitrary time t, is defined by the arbitrary configuration

$$\eta = (\eta_1, \ldots, \eta_i, \ldots, \eta_N) \quad \text{where} \quad \eta_i = \{OB,\ OC,\ MM\}$$

and that the dynamics is described by some transition rates ω_{KL}.

Then, the infinitesimal generator for the markovian dynamics of $\bar{\eta}_K$ is

$$\begin{aligned}\mathbf{L}_N f(\bar{\eta}_K) = \sum_{i=1}^{N} \Big\{ & \omega_{BC}\,[f(\eta_i = OC) - f(\eta_i = OB)] \\ & + \omega_{BM}\,[f(\eta_i = MM) - f(\eta_i = OB)] \\ & + \omega_{CB}\,[f(\eta_i = OB) - f(\eta_i = OC)] \\ & + \omega_{CM}\,[f(\eta_i = MM) - f(\eta_i = OC)] \\ & + \omega_{MB}\,[f(\eta_i = OB) - f(\eta_i = MM)] \\ & + \omega_{MC}\,[f(\eta_i = OC) - f(\eta_i = MM)] \Big\},\end{aligned}$$

where $f(\eta_i = L) - f(\eta_i = N)$ is the gradient of the configurations, considering the transition from an initial state L and a final state N.

All our results are shown in the appendix Supplementary Figures, but, to explain our model, we will exhibit some of them also in this chapter.

We have also reported, in the appendix Code, as examples, two commented codes of the simulations.

The discussion of the parameters, how they modify the behaviour of the system and their biological and clinical meaning will be the subject of the next section. Here we want just to clarify the theory behind our model and behind the simulations we performed.

5.3.1 Healthy bone homeostasis

First of all, we try to describe the homeostasis of the healthy bone tissue. We define the transition rates, that is the rate at which an empty site (indicated by $\diamondsuit_i$) is occupied by a cell of a specific type, taking into account the presence of sources and cooperative interactions.

$$\begin{aligned}(\diamondsuit_i \to OC) \quad &\text{with rate} \quad \omega_{BC} = a + b\, n_B, \\ (\diamondsuit_i \to OB) \quad &\text{with rate} \quad \omega_{CB} = c + d\, n_C,\end{aligned}$$

where (Fig. 5.5)

- a is the rate of formation of OC;
- b is the rate at which OB cells help OC cells, stimulating their activation;
- c is the rate of differentiation of OB by the osteoprogenitor cells;
- d is the rate at which OC cooperate with OB, in order to ensure homeostasis.

All the parameters are positive.

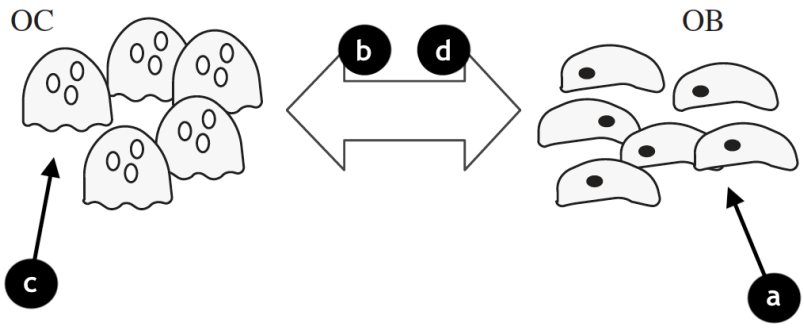


Figure 5.5: **Healthy bone tissue homeostasis**

It is important to remark that a and c represent two sources and that the production of OB and OC does not depend on the number of OB and OC that are already present.

- **Mean Field Approximation**

  We initially cope with the mean field approximation. We firstly translate the problems in terms of concentrations, so the total number of them sums to 1.

$$\bar{\eta}_B(t) = \frac{n_B(t)}{n_{TOT}}, \qquad \bar{\eta}_C(t) = \frac{n_C(t)}{n_{TOT}}.$$

Hence, we can write the infinitesimal generator for the dynamics of $\bar{\eta}_B$ and $\bar{\eta}_C$

$$\mathbf{L}_N f(\bar{\eta}_K) = \sum_{i=1}^{N} \Big\{ \omega_{BC} \left[f(\eta^i = OC) - f(\eta_i = OB)) \right] + \omega_{CB} \left[f(\eta^i = OB) - f(\eta_i = OC)) \right] \Big\}.$$

Using the Th.4.2.1, we read the limiting process from the limit of $\mathbf{L}_N$, as N goes to infinity. We can write

$$\begin{aligned}
\mathbf{L}_N f(\bar{\eta}_B) &= \sum_{i=1}^{N} \Big\{ \omega_{BC} \left[f(\eta^i = OC) - f(\eta_i = OB)) \right] \\
&\qquad + \omega_{CB} \left[f(\eta^i = OB) - f(\eta_i = OC)) \right] \Big\} \\
&= \sum_{i=1}^{N} \Big\{ (a + b\,\bar{\eta}_B) \left[f(\eta^i = OC) - f(\eta_i = OB)) \right] \\
&\qquad + (c + d\,\bar{\eta}_C) \left[f(\eta^i = OB) - f(\eta_i = OC)) \right] \Big\} \\
&= \sum_{i:\eta_i = OB} (a + b\,\bar{\eta}_B) \left(-\frac{1}{N} \right) \frac{\partial f}{\partial x}\Big|_{x=\bar{\eta}_B} \\
&\qquad + \sum_{i:\eta_i = OC} (c + d\,\bar{\eta}_C) \left(\frac{1}{N} \right) \frac{\partial f}{\partial x}\Big|_{x=\bar{\eta}_B} \\
&= -(a + b\,\bar{\eta}_B) \left(\sum_{i:\eta_i = OB} \frac{1}{N} \right) \frac{\partial f}{\partial x}\Big|_{x=\bar{\eta}_B} \\
&\qquad + (c + d\,\bar{\eta}_C) \left(\sum_{i:\eta_i = OC} \frac{1}{N} \right) \frac{\partial f}{\partial x}\Big|_{x=\bar{\eta}_B} \\
&= \left[\bar{\eta}_C (c + d\,\bar{\eta}_C) - \bar{\eta}_B (a + b\,\bar{\eta}_B) \right] \frac{\partial f}{\partial x}\Big|_{x=\bar{\eta}_B}.
\end{aligned}$$

The same holds for $\bar{\eta}_C$. Thus, the limiting generators are

$$\mathbf{L}_N f(\bar{\eta}_B) = \big[\bar{\eta}_C(c + d\,\bar{\eta}_C) - \bar{\eta}_B(a + b\,\bar{\eta}_B)\big] \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_B},$$
$$\mathbf{L}_N f(\bar{\eta}_C) = \big[\bar{\eta}_B(a + b\,\bar{\eta}_B) - \bar{\eta}_C(c + d\,\bar{\eta}_C)\big] \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_C},$$

that correspond to the mean field equations

$$\dot{\bar{\eta}}_B(t) = \bar{\eta}_C(t)(c + d\,\bar{\eta}_C(t)) - \bar{\eta}_B(t)(a + b\,\bar{\eta}_B(t)),$$
$$\dot{\bar{\eta}}_C(t) = \bar{\eta}_B(t)(a + b\,\bar{\eta}_B(t)) - \bar{\eta}_C(t)(c + d\,\bar{\eta}_C(t)).$$

It is also possible[2] to write the mean field equations, substituting $\bar{\eta}_B(t)$ with $x(t)$ and $\bar{\eta}_C(t)$ with $y(t)$, as

$$\dot{x}(t) = y(t)(c + d\,y(t)) - x(t)(a + b\,x(t)),$$
$$\dot{y}(t) = x(t)(a + b\,x(t)) - y(y)(c + d\,y(t)).$$

Considering that we are working with concentration, we have that $y(t) = 1 - x(t)$ for $\forall t$. Hence we need only one equation to describe the system in the mean field case

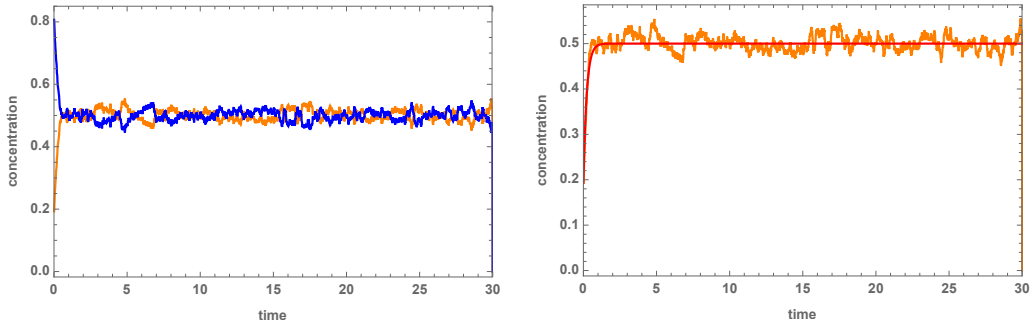
$$\dot{x}(t) = (1 - x(t))(c + d\,(1 - x(t))) - x(t)(a + b\,x(t)). \tag{5.1}$$

In Fig. 5.6 we show

- **(a)**: the time evolution of OB (in orange) and OC (in blue) populations simulated through the Poisson process over the lattice;
- **(b)**: the superposition of the simulated time evolution of OB cells and the mean field equation (in red) 5.1.

This comparison assures us that our model correctly reproduce the time evolution of the mean field differential equation derived from the theory.

[2]This substitution is useful to compare these equations with the code.



(a) (b)

Figure 5.6: **Healthy Bone Tissue - Mean Field:** $a = b = 1.2$; $c = d = 1$.

- **Local Model**

  Then we simulate the local model for the healthy tissue: as we have explained above, we compute the rate for each site just over the nearest neighbours on the lattice. In this way, for the healthy bone tissue we reproduce (Fig. 5.7) the same qualitative behaviour of the mean field equation.

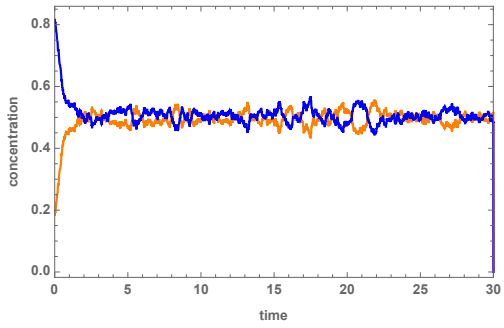


Figure 5.7: **Healthy Bone Tissue - Local Model:** $a = b = 1.2$; $c = d = 1$.

Both for the mean field and the local case, to perform the simulations, we do a further simplification. We reduce the number of parameters[3], supposing that the rate at which OB and OC are formed is the same, that is $a = c$.

[3] A parallel may be drawn here between some of our parameters and the ones of [106]: b(our model) $= a$(paper by Pacheco) and $d = e$.

5.3.2 Multiple myeloma progression

We want now to model the progression of multiple myeloma cells through the healthy tissue. As we have already underlined, multiple myeloma is a cancer of plasma cells, so there is no mutation of bone cells in malignant ones: we portray the system with MM cells already present. We have performed the same steps we made for the healthy case, but with some variations: so we are not going to repeat the arguments when they are analogous to the ones already explained, but we will dwell on the peculiar features of the myeloma case.

The first difference is that the cell are not picked at random with the same probability. To model the fact that MM cells reduce the OB activity and, consequently, their number, we have decided to make "more likely" OB cells to be chosen.

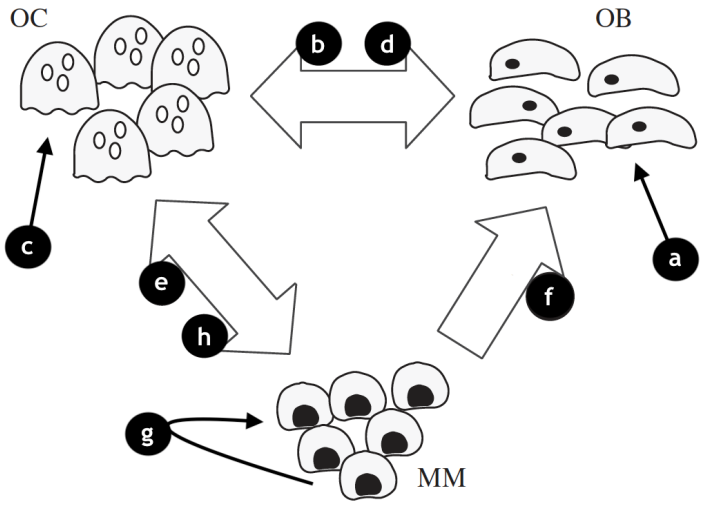


Figure 5.8: **Pathological bone turnover - Our model**

We define the model parameters as follows (Fig. 5.8).

- a, b, c, d are the same as in the healthy case;
- f is the rate at which MM cells have an harmful interaction with OB;
- R is the rate at which a random cell is picked;

- e is the rate at which MM cells enhance the OC activity;
- g is the reproduction rate of MM cells;
- h is the rate at which OC cells support MM cells.

The Heaviside step function is used to clarify that, if there is not MM cells, OC can not produce them. All the parameters are positive.

We want to clarify how the transition rates can include the fact that OB cells are more likely to be picked. Let us suppose that a cell is picked at random with rate R; we then assume that an OB cell is chosen with a rate $R + f\ n_M > R$.

If we take into account the transition $(OC \to \diamondsuit_i \to OB)$, we write

$$\begin{aligned} \omega_{CB} &= R\ \frac{c + d\ n_C}{R} \\ &= c + d\ n_C. \end{aligned}$$

Instead, if we consider the transition $(OB \to \diamondsuit_i \to OC)$, we write

$$\begin{aligned} \omega_{BC} &= (R + f\ n_M)\ \frac{a + b\ n_B + e\ n_M}{R} \\ &= R\ (1 + \frac{f}{R}\ n_M)\ \frac{a + b\ n_B + e\ n_M}{R} \\ &= (1 + \tilde{f}\ n_M)\ (\tilde{a} + \tilde{b}\ n_B + \tilde{e}\ n_M). \end{aligned}$$

Without losing generality, we can set $R = 1$, hence $\tilde{f} = f$, $\tilde{a} = a$ and so on.

Then, the transition rates are

$$
\begin{array}{rll}
(OB \to \diamondsuit_i \to OC) & \text{with rate} & \omega_{BC} = (1 + f\, n_M)(a + b\, n_B + e\, n_M), \\
(OB \to \diamondsuit_i \to MM) & \text{with rate} & \omega_{BM} = (1 + f\, n_M)(g\, n_M + h\, n_C\, \theta(n_M)), \\
(OC \to \diamondsuit_i \to OB) & \text{with rate} & \omega_{CB} = c + d\, n_C, \\
(OC \to \diamondsuit_i \to MM) & \text{with rate} & \omega_{CM} = g\, n_M + h\, n_C\, \theta(n_M), \\
(MM \to \diamondsuit_i \to OB) & \text{with rate} & \omega_{MB} = c + d\, n_C, \\
(MM \to \diamondsuit_i \to OC) & \text{with rate} & \omega_{MC} = a + b\, n_B + e\, n_M.
\end{array}
$$

- **Mean Field Approximation**

  In a similar way as in the healthy case, having translated the problem in terms of concentrations, we can derive the limiting generators

$$
\begin{aligned}
\mathbf{L}_N f(\bar{\eta}_B) = \Big[& \bar{\eta}_C(c + d\, \bar{\eta}_C) + \bar{\eta}_M((c + d\, \bar{\eta}_C)) \\
& - \bar{\eta}_B(1 + f\, \bar{\eta}_M)(a + b\, \bar{\eta}_C + e\, \bar{\eta}_M) \\
& - \bar{\eta}_B(1 + f\, \bar{\eta}_M)(g\, \bar{\eta}_M + h\, \bar{\eta}_C\, \theta(\bar{\eta}_M)) \Big] \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_B}, \\
\mathbf{L}_N f(\bar{\eta}_C) = \Big[& \bar{\eta}_B(1 + f\, \bar{\eta}_M)(a + b\, \bar{\eta}_C + e\, \bar{\eta}_M) \\
& + \bar{\eta}_M(a + b\, \bar{\eta}_C + e\, \bar{\eta}_M) \\
& - \bar{\eta}_C(c + d\, \bar{\eta}_C) - \bar{\eta}_C(g\, \bar{\eta}_M + h\, \bar{\eta}_C\, \theta(\bar{\eta}_M)) \Big] \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_C}, \\
\mathbf{L}_N f(\bar{\eta}_M) = \Big[& \bar{\eta}_B(1 + f\, \bar{\eta}_M)(g\, \bar{\eta}_M + h\, \bar{\eta}_C\, \theta(\bar{\eta}_M)) \\
& + \bar{\eta}_C(g\, \bar{\eta}_M + h\, \bar{\eta}_C\, \theta(\bar{\eta}_M)) \\
& - \bar{\eta}_M(c + d\, \bar{\eta}_C) - \bar{\eta}_M(a + b\, \bar{\eta}_C + e\, \bar{\eta}_M) \Big] \frac{\partial f}{\partial x}\bigg|_{x=\bar{\eta}_C}.
\end{aligned}
$$

Hence, the mean field equations are

$$\begin{aligned}
\dot{\bar{\eta}}_B(t) &= \bar{\eta}_C(t)(c + d\,\bar{\eta}_C(t)) + \bar{\eta}_M(t)(c + d\,\bar{\eta}_C(t)) \\
&\quad - \bar{\eta}_B(t)(1 + f\,\bar{\eta}_M(t))(a + b\,\bar{\eta}_C(t) + e\,\bar{\eta}_M(t)) \\
&\quad - \bar{\eta}_B(t)(1 + f\,\bar{\eta}_M(t))(g\,\bar{\eta}_M(t) + h\,\bar{\eta}_C(t)\,\theta(\bar{\eta}_M(t))), \\
\dot{\bar{\eta}}_C(t) &= \bar{\eta}_B(t)(1 + f\,\bar{\eta}_M(t))(a + b\,\bar{\eta}_C(t) + e\,\bar{\eta}_M(t)) \\
&\quad + \bar{\eta}_M(t)(a + b\,\bar{\eta}_C(t) + e\,\bar{\eta}_M(t)) \\
&\quad - \bar{\eta}_C(t)(c + d\,\bar{\eta}_C(t)) - \bar{\eta}_C(t)(g\,\bar{\eta}_M(t) + h\,\bar{\eta}_C(t)\,\theta(\bar{\eta}_M(t))), \\
\dot{\bar{\eta}}_M(t) &= \bar{\eta}_B(t)(1 + f\,\bar{\eta}_M(t))(g\,\bar{\eta}_M(t) + h\,\bar{\eta}_C(t)\,\theta(\bar{\eta}_M(t))) \\
&\quad + \bar{\eta}_C(t)(g\,\bar{\eta}_M(t) + h\,\bar{\eta}_C(t)\,\theta(\bar{\eta}_M(t))) \\
&\quad - \bar{\eta}_M(t)(c + d\,\bar{\eta}_C(t)) - \bar{\eta}_M(t)(a + b\,\bar{\eta}_C(t) + e\,\bar{\eta}_M(t)).
\end{aligned}$$

Substituting $\bar{\eta}_B(t)$ with $x(t)$, $\bar{\eta}_C(t)$ with $y(t)$ and $\bar{\eta}_M(t)$ with $z(t)$, we can write

$$\begin{aligned}
\dot{x}(t) &= y(t)(c + d\,y(t)) + z(t)(c + d\,y(t)) \\
&\quad - x(t)(1 + f\,z(t))(a + b\,y(t) + e\,z(t)) \\
&\quad - x(t)(1 + f\,z(t))(g\,z(t) + h\,y(t)\,\theta(z(t))), \\
\dot{y}(t) &= x(t)(1 + f\,z(t))(a + b\,y(t) + e\,z(t)) \\
&\quad + z(t)(a + b\,y(t) + e\,z(t)) \\
&\quad - y(t)(c + d\,y(t)) - y(t)(g\,z(t) + h\,y(t)\,\theta(z(t))), \\
\dot{z}(t) &= x(t)(1 + f\,z(t))(g\,z(t) + h\,y(t)\,\theta(z(t))) \\
&\quad + y(t)(g\,z(t) + h\,y(t)\,\theta(z(t))) \\
&\quad - z(t)(c + d\,y(t)) - z(t)(a + b\,y(t) + e\,z(t)).
\end{aligned}$$

Remembering that $x(t) + y(t) + z(t) = 1$ for $\forall t$, we can write $z(t) = 1 - x(t) - y(t)$ and reduce the number of equations

$$\begin{aligned}
\dot{x}(t) = & \, y(t)\,(c + d\,y(t)) + (1 - x(t) - y(t))(c + d\,y(t)) \\
& - x(t)(1 + f\,(1 - x(t) - y(t)))(a + b\,y(t) + e\,(1 - x(t) - y(t))) \\
& - x(t)(1 + f\,(1 - x(t) - y(t)))(g\,(1 - x(t) - y(t)) \\
& + h\,y(t)\,\theta((1 - x(t) - y(t))), \qquad (5.2) \\
\dot{y}(t) = & \, x(t)\,(1 + f\,(1 - x(t) - y(t)))(a + b\,y(t) + e\,(1 - x(t) - y(t))) \\
& + (1 - x(t) - y(t))(a + b\,y(t) + e\,(1 - x(t) - y(t))) \\
& - y(t)(c + d\,y(t)) - y(t)(g\,(1 - x(t) - y(t)) \\
& + h\,y(t)\,\theta((1 - x(t) - y(t))). \qquad (5.3)
\end{aligned}$$

Even in this case, both for the mean field and the local model, to perform the simulations we do a further simplification. We reduce the number of parameters, supposing that

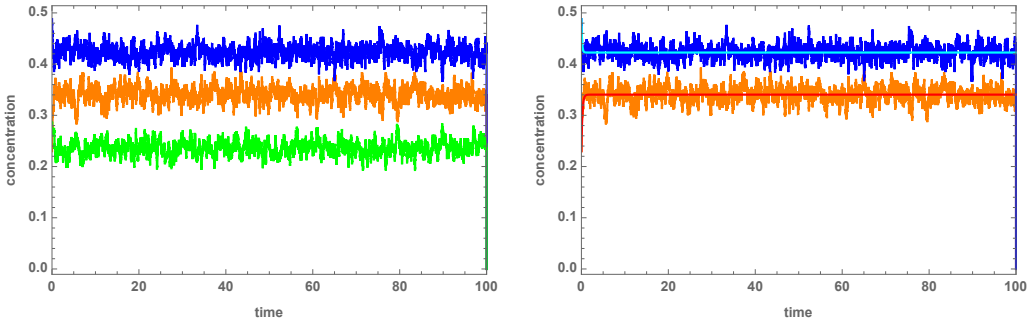
- the rate at which OB and OC are formed is the same: $a = c$;
- the cooperation between OB and OC has the same intensity in both directions: $b = d$;
- also the cooperation between OC and MM has the same intensity in both directions: $e = h$.

A parallel may be drawn here between some of our parameters and the ones of [106]; we will then renamed our reduced parameters after that used by J. Pacheco *et al.*: we will call $e = h \to \beta$ and $f \to \delta$.

In Fig. 5.9 we show

- **(a)**: the time evolution of OB (in orange), OC (in blue) and MM (in green) populations simulated through the Poisson process over the lattice;
- **(b)**: the superposition of the simulated time evolution of OB and OC cells and the mean field equations 5.2 (in red) and 5.3 (in light blue).

Also for the modelling of tumor progression, this comparison assures us that our model correctly reproduce the time evolution of the mean field differential equations derived from the theory.



(a) $g = 1.2$ (b) $g = 1.2$

Figure 5.9: **Multiple Myeloma - Mean Field**: $\beta = 2.0$ and $\delta = 0$.

- **Local Model**
  Simulating the dynamics of OB, OC and MM cells from a local point of view, we reproduce, for large time scale, that is after the equilibrium is reached, the qualitative behaviour (Fig. 5.10) of the mean field approximation.

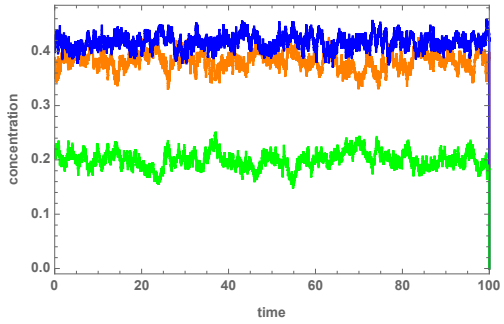


Figure 5.10: **Multiple Myeloma - Local Model**: $\beta = 2.0$; $\delta = 0$.

- **Initial concentration of MM cells**
  Lastly, we try to model the spatial spreading of MM cells. To do this, we drew at the center of the lattice a concentration of MM cells and fill the remaining of the lattice only with OB and OC cells. We then observe the time evolution of the system, thanks to our well-established simulation of Poisson processes. In this case, the qualitative behaviour of the system strongly depends on the parameter, as we explain in next section.

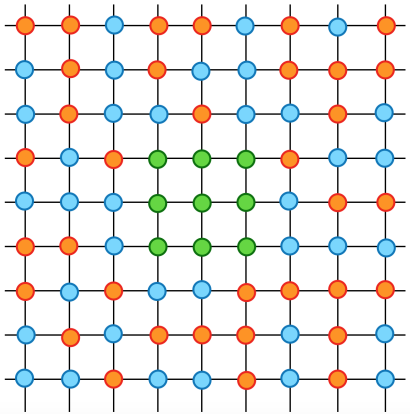


Figure 5.11: **An example of initial concentration of MM cells on the lattice.**

5.3.3 Parameters characterization

Our main results is the reproduction, varying parameters, of qualitative behaviours of time evolution of cell populations in presence of MM. Below we want to describe the meaning of the reduced parameters and how the system depends on them.

First of all, in our simulations we arbitrary set some parameters:

- $a = c = 1.2$, that is, the sources of OB and OC have the same rate of cells production;
- $b = d = 1.0$, that means that the cooperative interaction between OB and OC is symmetric.

The relevant parameter, for whose changing values we perform different simulations, are

- $e = h = \beta$, the intensity of symmetric cooperation between OC and MM cells. β has been linked [106] to the secretion of MIP-1α or of IL-1β. In general, it is related to the molecular pathways that enhance osteoclast activity. Actually, increasing β will lead to more bone destruction, higher tumor burden and faster tumor progression. From our simulations, we observe that, as long as $\beta < 1$, the MM cells population can go extinct and OB and OC may re-established the physiological bone homeostasis. On the contrary, if $\beta > 1$, at the equilibrium there is always the coexistence of the three cell types, that, unfortunately, means a negative prognosis for the patient.
- $f = \delta$, the intensity of harming action of MM on OB. This parameter is related with the molecular pathways that reduce osteoblast activity. In particular, [106] it has been associated to WNT pathway and to DKK-1. High values of δ portray a very diminished osteoblast activity.
- g, the rate of MM mitosis. We propose that, through this parameter, it may be possible to discriminate between different forms of myeloma [107] [91] [92].

We suggest, for example, that $g = 1.0$ can represent the *monoclonal gammopathy of undetermined significance* (MGUS), that is a plasma cells discrasia that may or not evolve into myeloma. Patients affected by MGUS show lower levels of antibodies, a lower number of plasma cells in the bone marrow, but they rarely have symptoms or major problems.

- The value $g = 1.2$ can be linked to the *lymphoplasmacytic lymphoma* or the *smouldering multiple myeloma* (SMM), that are precursor conditions of myeloma and that do not present bone destruction. Patients with SMM may have a stable disease state for many years (about the 75% of them move forward malignant myeloma in 15 years) before the progression to active myeloma.

Indolent multiple myeloma (IMM), instead, has a shorter pre-malignant state (10 months), thus can be represented by $g = 2.0$.

Lastly, $g = 3.5$ stands for the active multiple myeloma, the most aggressive form of plasma cells neoplasia.

The complete set of parameters, $\{a, b, c, d, e, f, g\}$, can describe a patient-specific case of tumor: they can reflect tumor-host interactions and the variability of disease due to differences in the host rather than the cancerous cell genotype.

Of course, different combinations of these parameters can lead to different equilibria for cell populations and some values of them can influence the model behaviour so much that the system will become non-sensitive to the other parameters variation. For example, if $g = 3.5$, the tumor will quickly develop and will invade the lattice (that is, the bone), regardless of the values of β and δ.

Considering the tumor progression, it can be possible to modify the parameters values through different kind of therapies, in order to offer to the patient a better prognosis. For example, both bone marrow transplantation and drug treatment can lower the values of β and δ, trying to avoid a MM cells invasion of the tissue.

Moreover this model could help to estimate the risk of progression of MGUS and other precursor forms into multiple myeloma.

CONCLUSIONS AND FUTURE PERSPECTIVES

PRIMA LEGGE DI FUDD
SULLA CREATIVITÀ
Per avere una buona idea,
fatti venire molte idee.

CONTROLEGGE DI FUDD
Più idee ti fai venire, più sarà difficile
riconoscere quella buona.

Having faced such a wide field as mathematical oncology, we now want to draw the conclusions of our work.

We tried to develop a model with a bottom-up approach, starting from the study of the biological problem. In such a manner, we tackled all the steps that lead to the creation of a model from a real phenomenon to its simulation. We also made the decisions required to render the problem manageable, both from a mathematical and a computational point of view. By doing so, we have offered a novel approach to the study of tumor progression through healthy tissue. We define the model for the specific case of multiple myeloma through bone tissue. However, it is applicable, via appropriate modifies of transition rates and parameters, to every kind of neoplasia.

In particular, we managed to formulate a realistic model of microscopic interactions between healthy and cancerous cells. Moreover, we were able to simulate the system from a local point of view, as well as in its mean field approximation. This factor (aspect, trait, feature..) opens up new interesting opportunities over the in-depth study of phenomena such as diffusion or segregation of malignant cells.

We hope that in the near future it will be possible to test models like ours on clinical data, in order to perform statistical inference over the free parameters. By doing this, the models may become powerful tools to make diagnoses and to analyse the feasibility of treatment, both surgically and pharmacologically. Moreover, the models lend themselves, by their very nature, to the application to patient-specific therapy, which is considered as the forthcoming landing of the clinical practice.

CODE

Below we report the code of the simulations for

- **the Healthy Mean Field Model**
  Firstly, we show the parameters and variables definition; then the core of the code, that is the Poisson process on the lattice for two species of cells (OB and OC) and, lastly, the simulation of the mean field differential equation;

- **the Myeloma Model with an initial concentration of cancerous cells**
  Firstly, we show the parameters and variables definition and how to make a disk of MM cells in the center of the square lattice. Successively, we show the core of the code, that is the Poisson process on the lattice for three species of cells (OB, OC and MM).

Healthy mean field model

```
(*set the number of iterations (Niter), the time interval (dt);
 the paramenters (a, b, c, d) and the dimensions of the lattice (L)*)
Niter = 3000;
dt = 0.01;
a = c = 1.2;
b = 1.;
d = 1.;
L = 30;
(*Ntot is the total number of sites of the lattice*)
Ntot = L * L;
(*define the initial arrays*)
(*rateOB and rateOC are the matrices of the rate of transitions,
calculated for each site of the lattice.
  They are calculated for each iteration of the simulation*)
rateOB = Table[0, {ix, 1, L}, {jx, 1, L}] ;
rateOC = Table[0, {ix, 1, L}, {jx, 1, L}];
(*Reticolo is the vector of matrices that describe the lattice.
  There are Niter matrices L*L.
  The matrices are initially filled with random number (0 or 1).
  0 stands for OB cells, 1 stands for OC cells.
   In this case the proportion of random number is OB=0.2, OC=0.8*)
Reticolo = Table[Table[RandomChoice[{0.2, 0.8} → {0, 1}],
    {ix, 1, L}, {jx, 1, L}] , {kx, 1, Niter}];
(*nOB and nOC are the vectors of the total number of OB or OC
 cells on the lattice for each iteration*)
nOB = Table[0, {k, 1, Niter}];
nOC = Table[0, {k, 1, Niter}];
(*popOB and popOC are the vectors of the fraction of the OB and OC
  population over the total number of cells.
  That is, for example, popOB=nOB/nTOT*)
popOB = Table[0, {k, 1, Niter}];
popOC = Table[0, {k, 1, Niter}];
```

```
For[k = 1, k ≤ Niter - 1, k++,
 (*for each iteration, nOC = number of 1 on the lattice,
 nOB = number of 0 on the lattice.
    popOB and popOC are the relative fractions
   over the total number of sites.
    These are the vectors that will be plot later*)
 nOC[[k]] = Total[Total[Reticolo[[k]]]];
 nOB[[k]] = Ntot - nOC[[k]];
 popOB[[k]] = N[nOB[[k]] / Ntot];
 popOC[[k]] = N[nOC[[k]] / Ntot];
 For[j = 1, j ≤ L, j++,
  For[i = 1, i ≤ L, i++,
    (*for each iteration, rateOC and rateOB
     are calculated for each site of the matrix*)
    (*N.B.: this is a mean field case,
    rates are computed over the entire lattice*)
    rateOC[[i, j]] = a + b * N[nOB[[k]] / Ntot];
    rateOB[[i, j]] = c + d * (1. - N[nOB[[k]] / Ntot]);
   ];
 ];
 For[jx = 1, jx ≤ L, jx++,
  For[ix = 1, ix ≤ L, ix++,
   (*for each site, a random number (z) is generated*)
   z = RandomReal[];
   (*rate is the sum, for each site of rateOB and rateOC*)
   rate = N[rateOC[[ix, jx]] + rateOB[[ix, jx]]];
   If[z < rate * dt,
    (*if z<rate*dt, there is a point of the Poisson process
     (that is, the cell on that site is removed)
     and another random number (w) is generated*)
```

```
w = RandomReal[];
          (*f is, for each site, rateOC/rate*)
          f = N[rateOC[[ix, jx]] / rate];
          If[w < f,
            (*if w<f, then the empty site is occupied by an OC,
            otherwise by an OB*)
            Reticolo[[k + 1]][[ix, jx]] = 1;,
            Reticolo[[k + 1]][[ix, jx]] = 0;
           ];,
          (*if z>rate*dt, nothing happens and
            the k-th matrix of Reticolo remains unchanged*)
          Reticolo[[k + 1]][[ix, jx]] = Reticolo[[k]][[ix, jx]];,
          Reticolo[[k + 1]][[ix, jx]] = Reticolo[[k]][[ix, jx]];
          ];
        ];
      ];
    ];

(*solve the mean field differential equation*)
s = NDSolve[{x'[t] == (1 - x[t]) (c + d * (1 - x[t])) - x[t] (a + b * x[t]),
                        x[0] == prop}, {x}, {t, (dt * Niter)}];
```

Myeloma model with the initial concentration of cancerous cells

```
(*set the number of iterations (Niter),
the time interval (dt); the paramenters (a, b, c, d, e, f)
 and the dimensions of the lattice (L)*)
Niter = 10 000;
dt = 0.01;
a = c = 1.2;
b = d = 1.;
e = h = 0.5;
f = 0.;
g = 1.0;
L = 30;
(*raggio is the radius of the circle of MM cells
 that we want to put at the center of the lattice*)
raggio = 2;
(*Ntot is the total number of sites of the lattice*)
Ntot = L * L;
(*define the initial arrays*)
(*rateOB, rateOC, rateMM  and deathOB are
 the matrices of the rate of transitions,
calculated for each site of the lattice.
  They are calculated for each iteration of the simulation*)
rateOB = Table[0, {ix, 1, L}, {jx, 1, L}] ;
rateOC = Table[0, {ix, 1, L}, {jx, 1, L}];
rateMM = Table[0, {ix, 1, L}, {jx, 1, L}] ;
deathOB = Table[0, {ix, 1, L}, {jx, 1, L}] ;
(*Reticolo is the vector of matrices that describe the lattice.
  There are Niter matrices L*L.
  The matrices are initially filled with random number (0 or 1).
  0 stands for OB cells, 1 stands for OC cells.
   In this case the proportion
  of random number is OB = 0.5, OC = 0.5*)
```

```
Reticolo = Table[Table[RandomChoice[{0.5, 0.5} → {1, 0}],
     {ix, 1, L}, {jx, 1, L}] , {kx, 1, Niter}];
(*disco is a L*L matrix of 0, with a disk of -1 of radius
  "raggio" at the center of the matrix*)
disco = Table[-(DiskMatrix[raggio, L]), {kx, 1, Niter}];
(*nOB nOC and nMM are the vectors of the total number
 of OB, OC or MM cells on the lattice for each iteration*)
nOB = Table[0, {k, 1, Niter}];
nOC = Table[0, {k, 1, Niter}];
nMM = Table[0, {k, 1, Niter}];
(*popOB, popOC and popMM are the vectors of the fraction of the OB,
OC and MM population over the total number of cells.
  That is, for example, popOB=nOB/nTOT*)
popOB = Table[0, {k, 1, Niter}];
popOC = Table[0, {k, 1, Niter}];
popMM = Table[0, {k, 1, Niter}];


(*here, for each iteration, the matrix disco
  is overlaid to the k-th matrix Reticolo.
   Thus, after this cycle, the matrix Reticolo
  (filled with 0 and 1) has a disk of -1 in the center*)
For[k = 1, k ≤ Niter, k++,
  For[i = 1, i ≤ L, i++,
    For[j = 1, j ≤ L, j++,
      If[disco[[k]][[i, j]] == -1,
        Reticolo[[k]][[i, j]] = disco[[k]][[i, j]],
        Reticolo[[k]][[i, j]] = Reticolo[[k]][[i, j]]
       ];
     ];
   ];
 ];
```

```
For[k = 1, k ≤ Niter - 1, k++,
 (*for each iteration, nOC = number of 1 on the lattice,
  nOB = number of 0 on the lattice,
 nMM = number of -1 on the lattice.
     popOB, popOC and popMM are the relative
  fractions over the total number of sites.
   These are the vectors that will be plot later*)
 nOC[[k]] = Count[Flatten[Reticolo[[k]]], 1];
 nOB[[k]] = Count[Flatten[Reticolo[[k]]], 0];
 nMM[[k]] = Count[Flatten[Reticolo[[k]]], -1];
 popOB[[k]] = N[nOB[[k]] / Ntot];
 popOC[[k]] = N[nOC[[k]] / Ntot];
 popMM[[k]] = N[nMM[[k]] / Ntot]; For[j = 1, j ≤ L, j++,
  For[i = 1, i ≤ L, i++,
   (*for each iteration, rateOC, rateOB, rateMM and deathOB
     are calculated for each site of the matrix*)
   (*N.B.: this is a local case, rates are
    computed just over the nearest neighbours*)
   rateOC[[i, j]] = a + (b / 4.) * Count[{Reticolo [[k ]]
          [[i, Mod[j - 1, L, 1]]], Reticolo [[k ]][[i, Mod[j + 1, L, 1]]],
        Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
        Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, 0] + (e / 4.) *
      Count[{Reticolo [[k ]][[i, Mod[j - 1, L, 1]]], Reticolo [[k ]]
          [[i, Mod[j + 1, L, 1]]], Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
        Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, -1];
   rateOB[[i, j]] = c + (d / 4.) *Count[{Reticolo [[k ]]
          [[i, Mod[j - 1, L, 1]]], Reticolo [[k ]][[i, Mod[j + 1, L, 1]]],
        Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
        Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, 1];
```

```
    rateMM[[i, j]] = (g / 4.) * Count[{Reticolo [[k ]]
            [[i, Mod[j - 1, L, 1]]], Reticolo [[k ]][[i, Mod[j + 1, L, 1]]],
          Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
          Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, -1] + (h / 4.) *
        Count[{Reticolo [[k ]][[i, Mod[j - 1, L, 1]]],
          Reticolo [[k ]][[i, Mod[j + 1, L, 1]]],
          Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
          Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, 1] *
        HeavisideTheta[Count[{Reticolo [[k ]]
            [[i, Mod[j - 1, L, 1]]], Reticolo [[k ]][[i, Mod[j + 1, L, 1]]],
          Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
          Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, -1]];
    deathOB[[i, j]] = (f / 4.) * Count[{Reticolo [[k ]]
          [[i, Mod[j - 1, L, 1]]], Reticolo [[k ]][[i, Mod[j + 1, L, 1]]],
        Reticolo [[k ]][[Mod[i - 1, L, 1], j]],
        Reticolo [[k ]][[Mod[i + 1, L, 1], j]]}, -1]
    ];
  ];
For[jx = 1, jx ≤ L, jx ++,
 For[ix = 1, ix ≤ L, ix ++,
  (*for each site, a random number (z) is generated*)
  z = RandomReal[];
  (*rate is the sum, for each site of rateOB, rateOC and rateMM*)
  rate = rateOC[[ix, jx]] + rateOB[[ix, jx]] + rateMM[[ix, jx]];
  (*rateplus is the sum, for each site of rateOB, rateOC,
  rateMM and deathOB*)
  rateplus = rate + deathOB[[ix, jx]];
  (*step is, for each site, rateOC/rate*)
  step = rateOC[[ix, jx]] / rate;
  (*step2 is, for each site, (rateOB+rateOC)/rate*)
  step2 = (rateOB[[ix, jx]] + rateOC[[ix, jx]]) / rate;
```

```
If[Reticolo[[k]][[ix, jx]] == 0,
 (*if the site is occupied by an OB,
  we have to consider rateplus*dt as a threshold*)
 If[z < rateplus * dt,
  (*if z<rateplus*dt, there is a point of the Poisson process
   (that is, the cell on that site is removed)
   and another random number (w) is generated*)
  w = RandomReal[];
  If[w < step,
   (*if w<step, then the empty site is occupied by an OC*)
   Reticolo[[k + 1]][[ix, jx]] = 1;
  ];

  If[step ≤ w ≤ step2,
   (*if step≤w≤step2, then the empty site is occupied by an OB*)
   Reticolo[[k + 1]][[ix, jx]] = 0;
  ];
  If[step2 ≤ w ≤ 1,
   (*if step2≤w≤1, then the empty site is occupied by an MM*)
   Reticolo[[k + 1]][[ix, jx]] = -1;
  ];,
  (*if z>rate*dt, nothing happens and
    the k-th matrix of Reticolo remains unchanged*)
  Reticolo[[k + 1]][[ix, jx]] = Reticolo[[k]][[ix, jx]];,
  Reticolo[[k + 1]][[ix, jx]] = Reticolo[[k]][[ix, jx]];
 ],
 (*if the site is occupied by and OC or by an MM,
 the threshold is rate*dt *)
```

```
        If[z < rate*dt,
          (*if z<rate*dt, there is a point of the Poisson process
           (that is, the cell on that site is removed)
            and another random number (w) is generated*)
          w = RandomReal[];
          If[w < step,
           (*if w<step, then the empty site is occupied by an OC*)
           Reticolo[[k + 1]][[ix, jx]] = 1;
          ];
          If[step ≤ w ≤ step2,
           (*if step≤w≤step2, then the empty site is occupied by an OB*)
           Reticolo[[k + 1]][[ix, jx]] = 0;
          ];
          If[step2 ≤ w ≤ 1,
           (*if step2≤w≤1, then the empty site is occupied by an MM*)
           Reticolo[[k + 1]][[ix, jx]] = -1;
          ];,
          (*if z>rate*dt, nothing happens and
            the k-th matrix of Reticolo remains unchanged*)
          Reticolo[[k + 1]][[ix, jx]] = Reticolo[[k]][[ix, jx]];,
          Reticolo[[k + 1]][[ix, jx]] = Reticolo[[k]][[ix, jx]];
         ];
        ];
      ];
    ];
  ];
```

SUPPLEMENTARY FIGURES

In this appendix we show the simulations we performed. Firstly, there are the ones about the healthy bone tissue, both mean field and local. Then, there are the plots about the progression of multiple myeloma: mean field, local and the simulations that start with an initial concentration of myeloma cells, with following dispersal or invasion.

For sake of simplicity we show here the color legend.

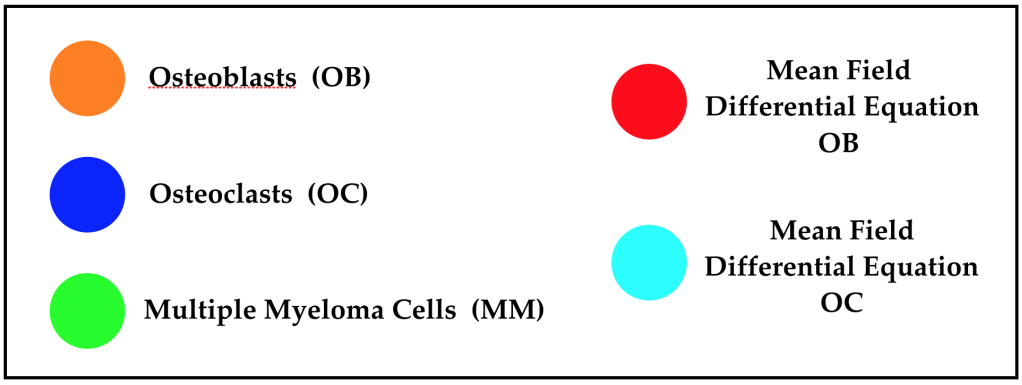


Figure 5.12: **Legend**

The initial conditions refer to the proportions according to which the OB, OC and MM cells are initially randomly spread over the lattice.

The time step is always 0.01.

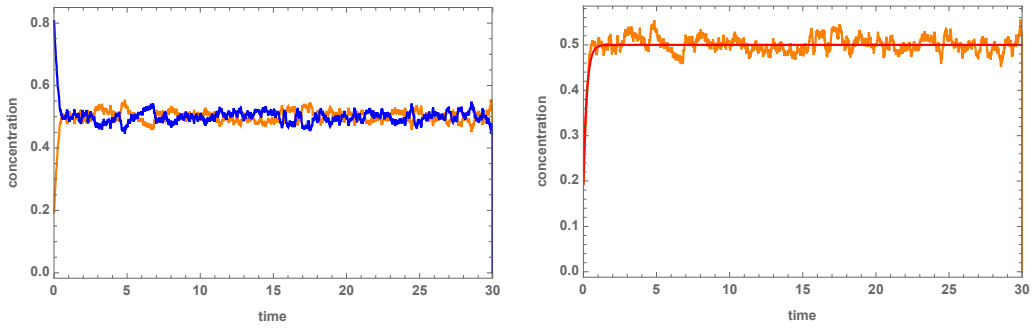
- **Healthy Bone Tissue - Mean Field**
  Iterations : 3000
  Lattice dimension : 30×30
  Initial conditions : **OB** = 0.2, **OC** = 0.8

- **Healthy Bone Tissue - Local Model**
  Iterations : 3000
  Lattice dimension : 30×30
  Initial conditions : **OB** = 0.2, **OC** = 0.8

- **Multiple Myeloma - Mean Field**
  Iterations : 10000
  Lattice dimension : 30×30
  Initial conditions : **OB** = 0.2, **OC** = 0.5, **MM** = 0.3

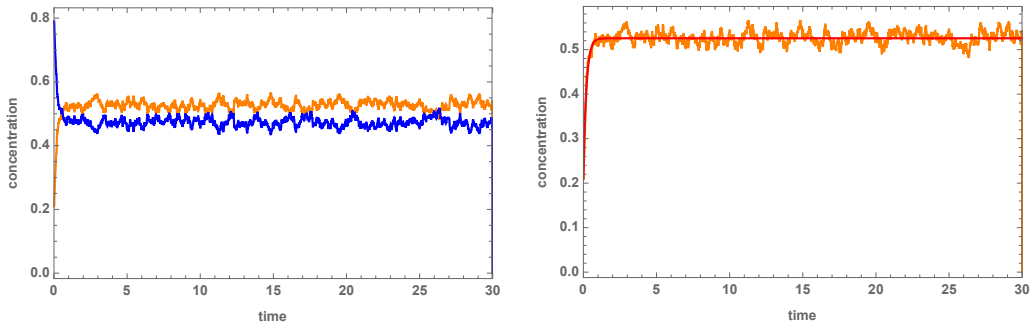
- **Multiple Myeloma - Local Model**
  Iterations : 10000
  Lattice dimension : 30×30
  Initial conditions : **OB** = 0.4, **OC** = 0.4, **MM** = 0.2

- **Multiple Myeloma with initial segregation**
  Iterations : 10000
  Lattice dimension : 30×30
  Initial conditions : **OB** = 0.5, **OC** = 0.5, with a disk of radius 2 of **MM**



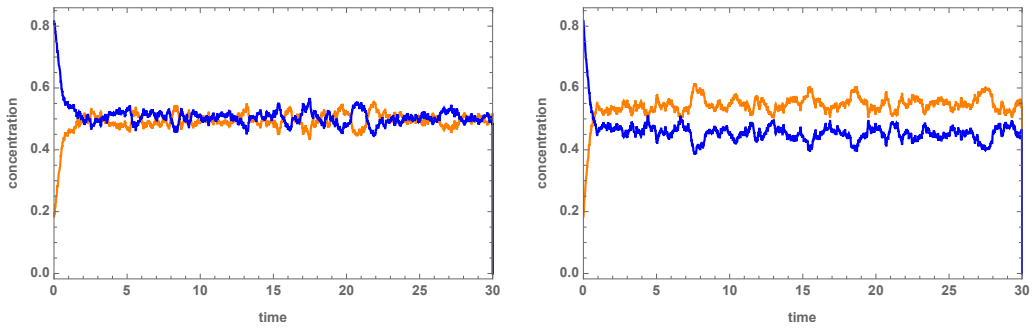
(a) (b)

Figure 5.13: **Healthy Bone Tissue - Mean Field:** $a = b = 1.2$; $c = d = 1$.



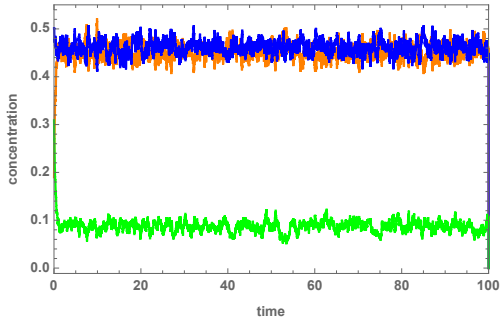
(a) (b)

Figure 5.14: **Healthy Bone Tissue - Mean Field:** $a = b = 1.2$; $c = 1.$; $d = 1.5$

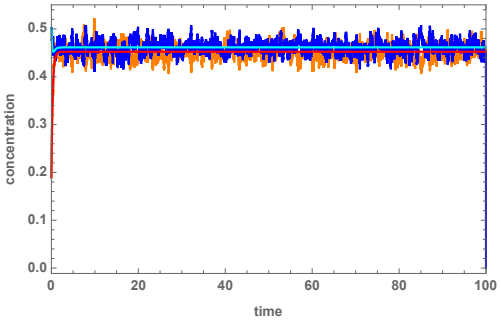


(a) $c = d = 1.$ (b) $c = 1.$; $d = 1.5$

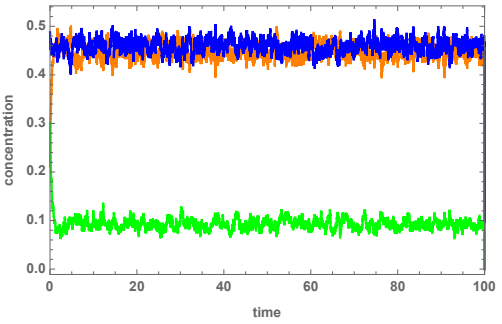
Figure 5.15: **Healthy Bone Tissue - Local Model:** $a = b = 1.2$



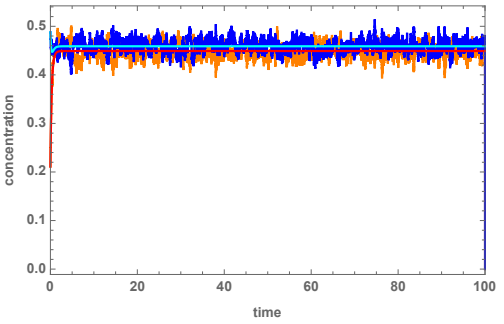
(a) $g = 1.0$



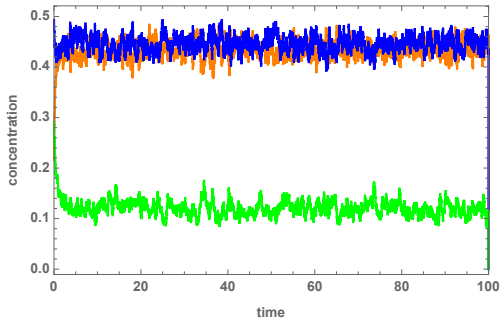
(b) $g = 1.0$



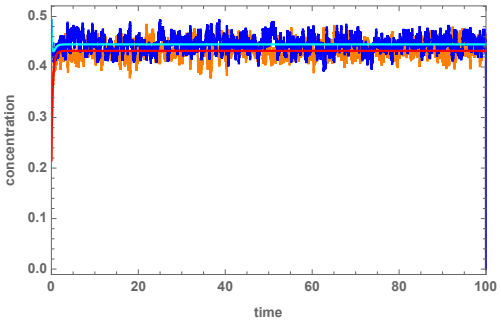
(c) $g = 1.2$



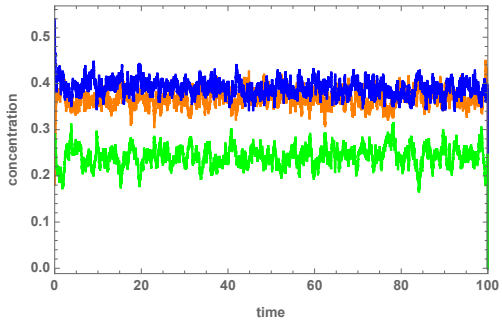
(d) $g = 1.2$



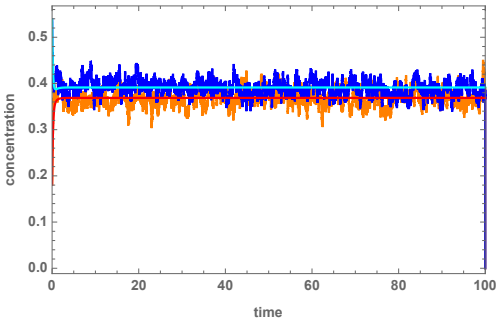
(e) $g = 2.0$



(f) $g = 2.0$

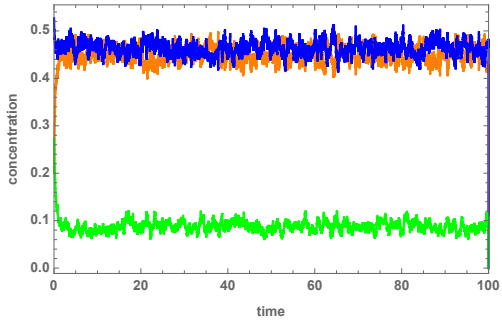


(g) $g = 3.5$

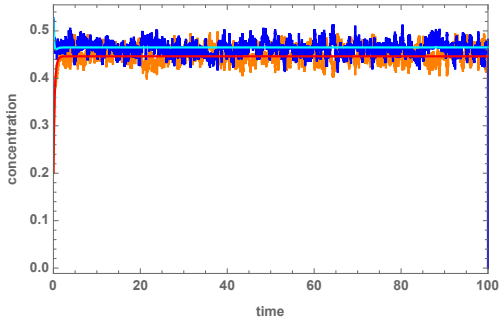


(h) $g = 3.5$

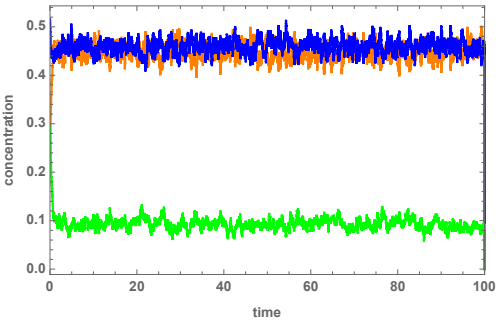
Figure 5.16: **Multiple Myeloma - Mean Field**: $\beta = 0.5$ and $\delta = 0$.



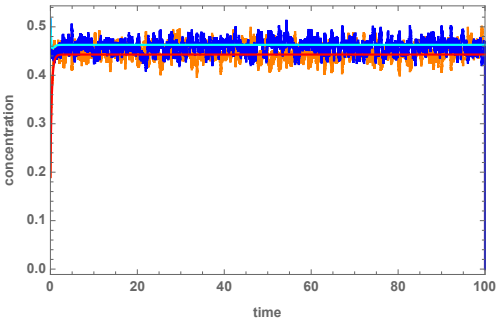
(a) $g = 1.0$



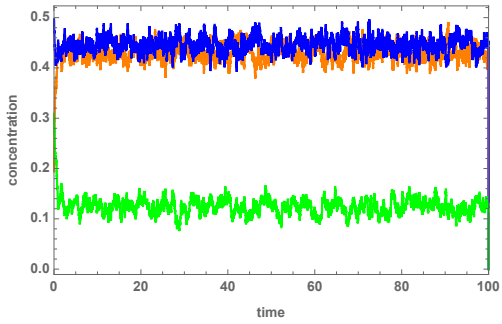
(b) $g = 1.0$



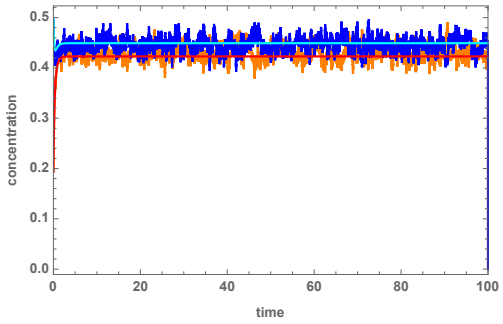
(c) $g = 1.2$



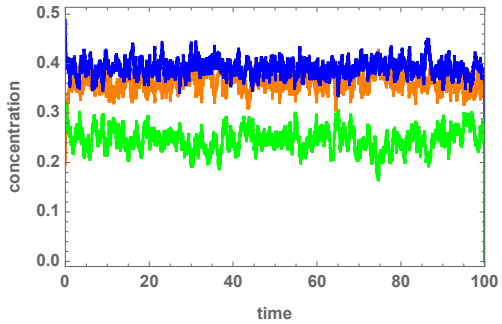
(d) $g = 1.2$



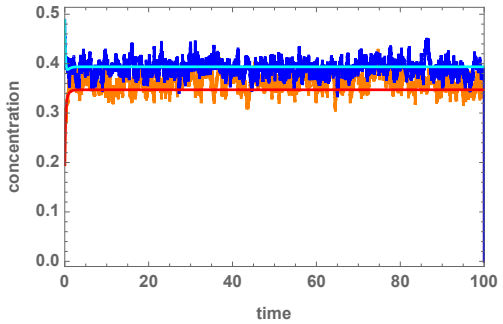
(e) $g = 2.0$



(f) $g = 2.0$

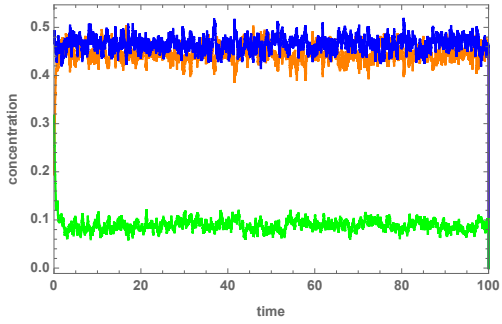


(g) $g = 3.5$

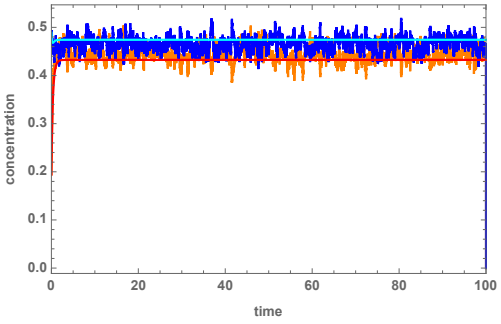


(h) $g = 3.5$

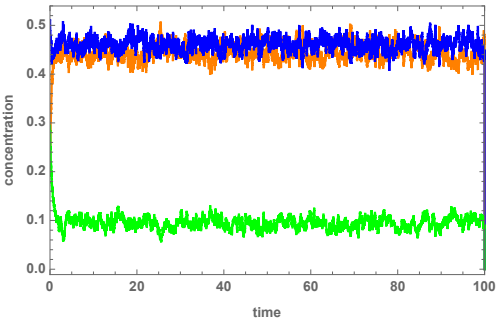
Figure 5.17: **Multiple Myeloma - Mean Field**: $\beta = 0.5$ and $\delta = 0.3$



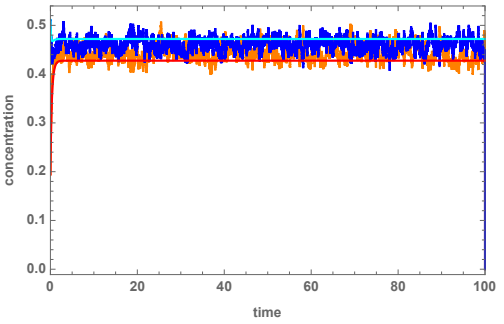
(a) $g = 1.0$



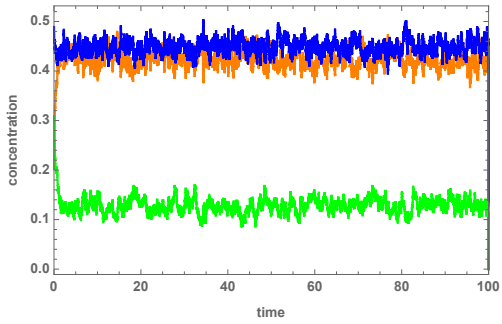
(b) $g = 1.0$



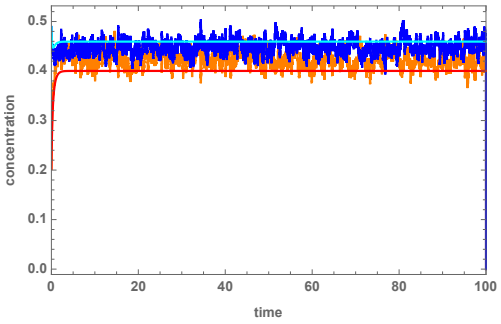
(c) $g = 1.2$



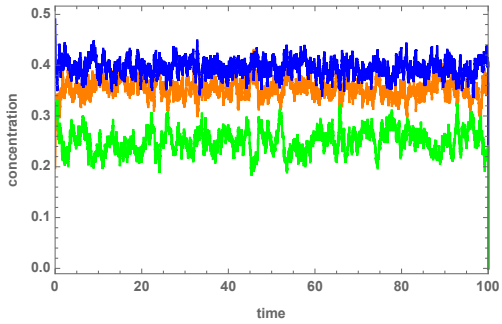
(d) $g = 1.2$



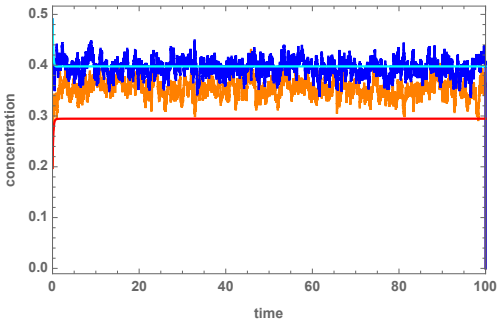
(e) $g = 2.0$



(f) $g = 2.0$

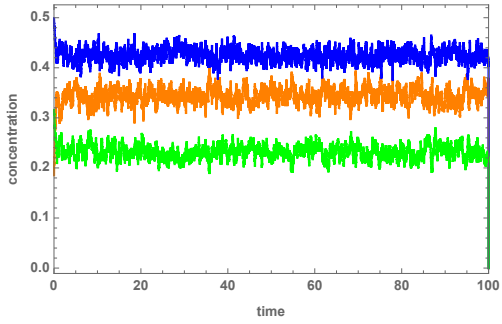


(g) $g = 3.5$

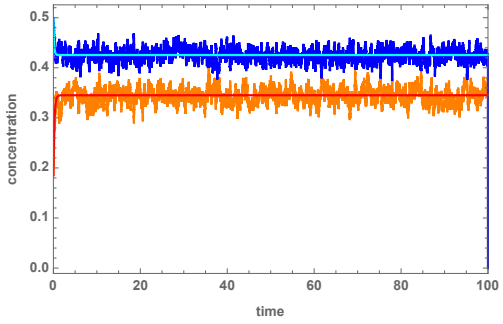


(h) $g = 3.5$

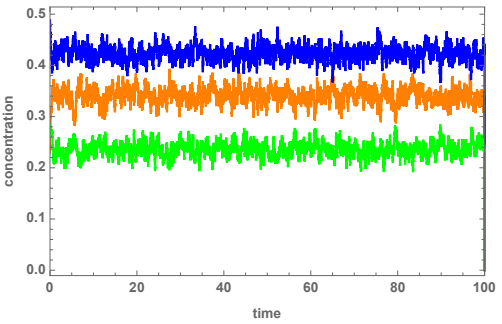
Figure 5.18: **Multiple Myeloma - Mean Field**: $\beta = 0.5$ and $\delta = 1.0$



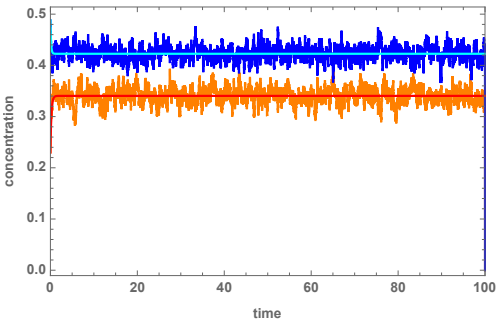
(a) $g = 1.0$



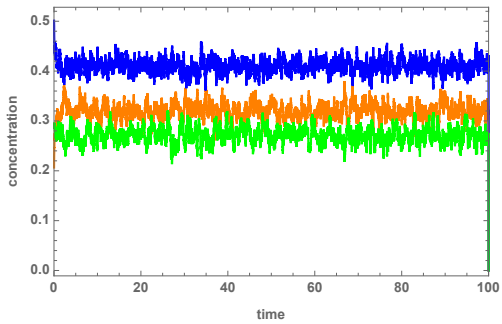
(b) $g = 1.0$



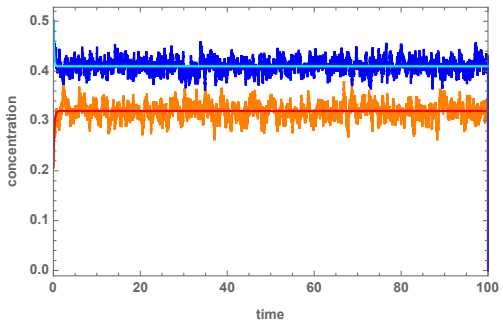
(c) $g = 1.2$



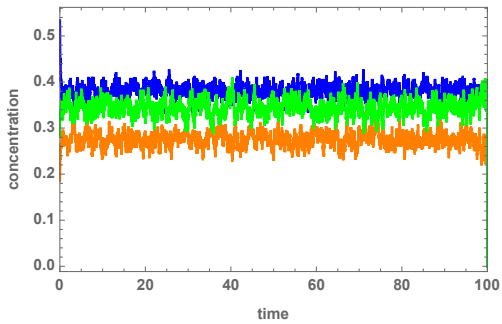
(d) $g = 1.2$



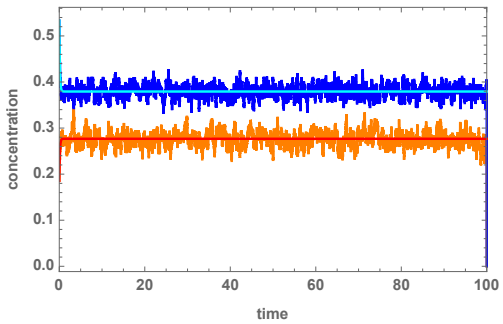
(e) $g = 2.0$



(f) $g = 2.0$

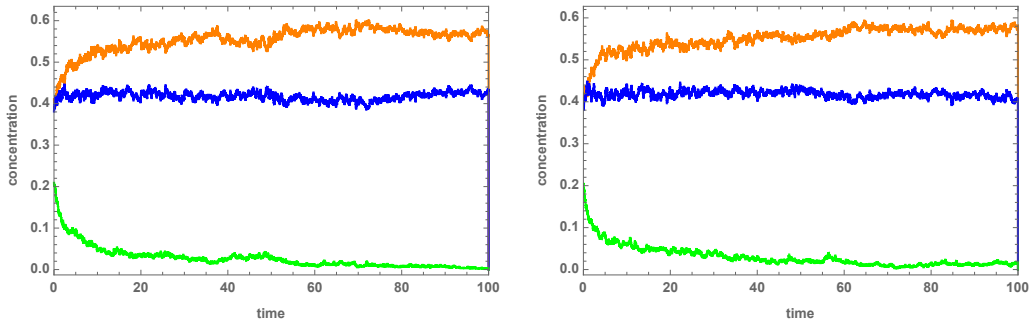


(g) $g = 3.5$



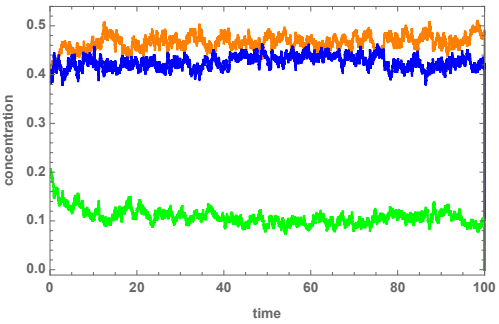
(h) $g = 3.5$

Figure 5.19: **Multiple Myeloma - Mean Field**: $\beta = 2.0$ and $\delta = 0$.

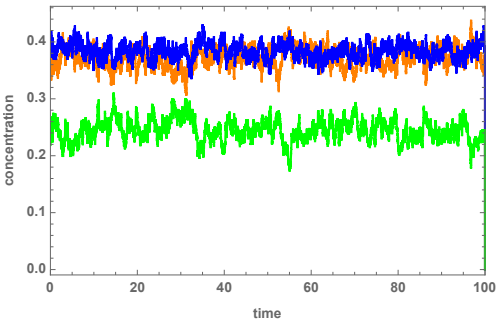


(a) $g = 1.0$

(b) $g = 1.2$

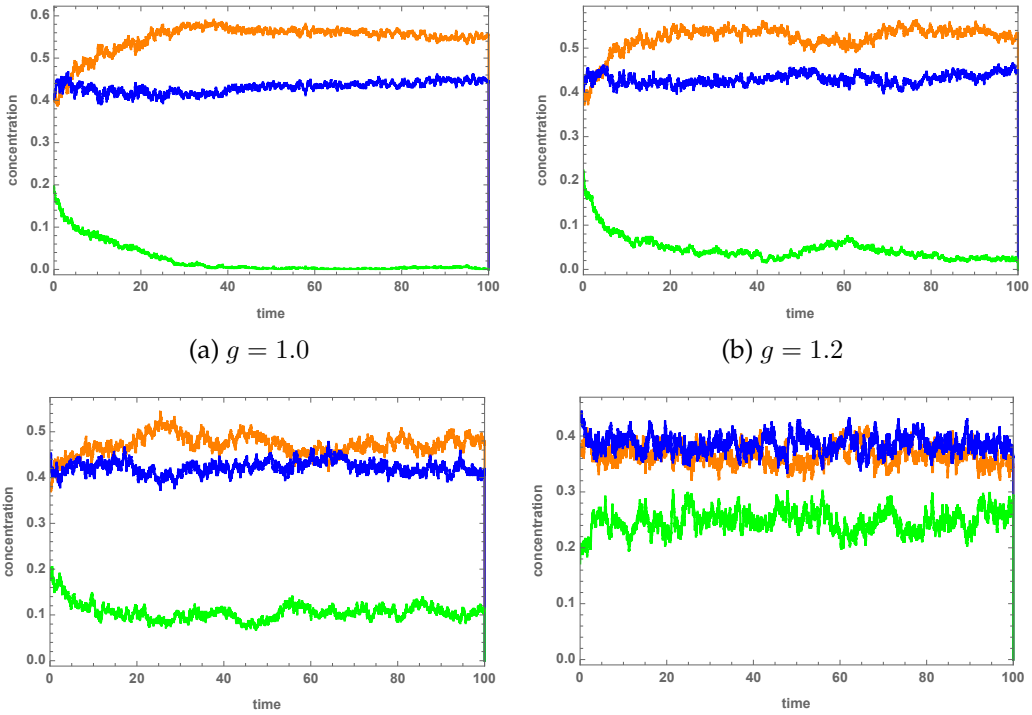


(c) $g = 2.0$



(d) $g = 3.5$

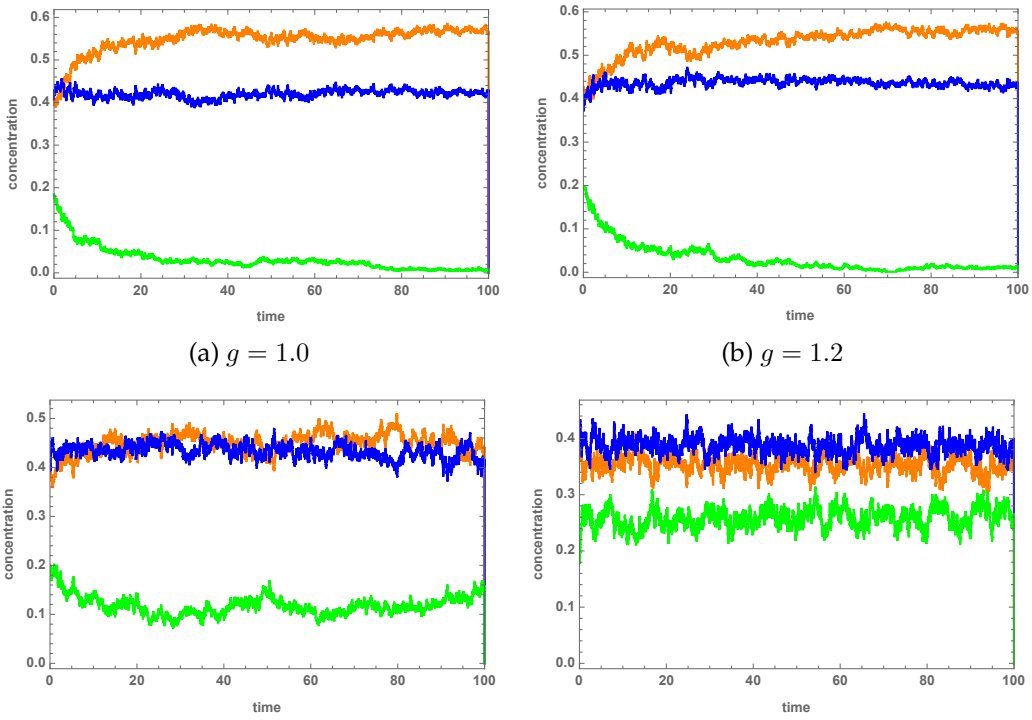
Figure 5.20: **Multiple Myeloma - Local Model**: $\beta = 0.5$; $\delta = 0$.



(a) $g = 1.0$ (b) $g = 1.2$

(c) $g = 2.0$ (d) $g = 3.5$

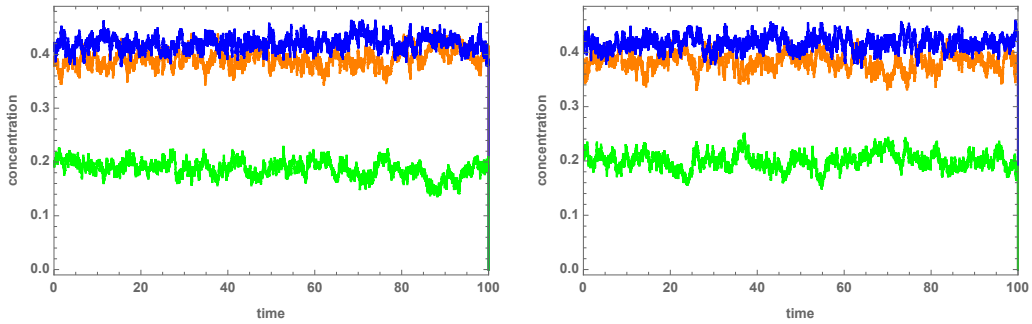
Figure 5.21: **Multiple Myeloma - Local Model**: $\beta = 0.5$; $\delta = 0.3$



(a) $g = 1.0$ (b) $g = 1.2$

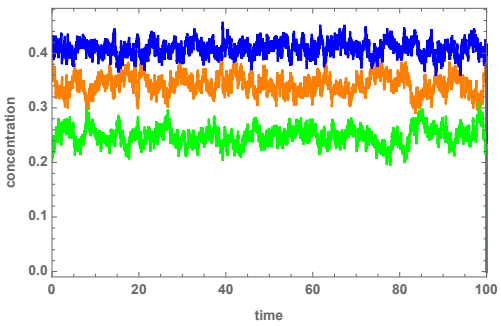
(c) $g = 2.0$ (d) $g = 3.5$

Figure 5.22: **Multiple Myeloma - Local Model**: $\beta = 0.5$; $\delta = 1.0$

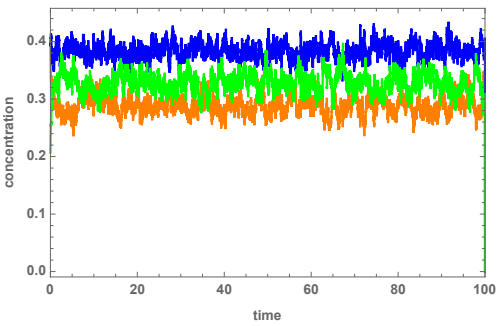


(a) $g = 1.0$

(b) $g = 1.2$

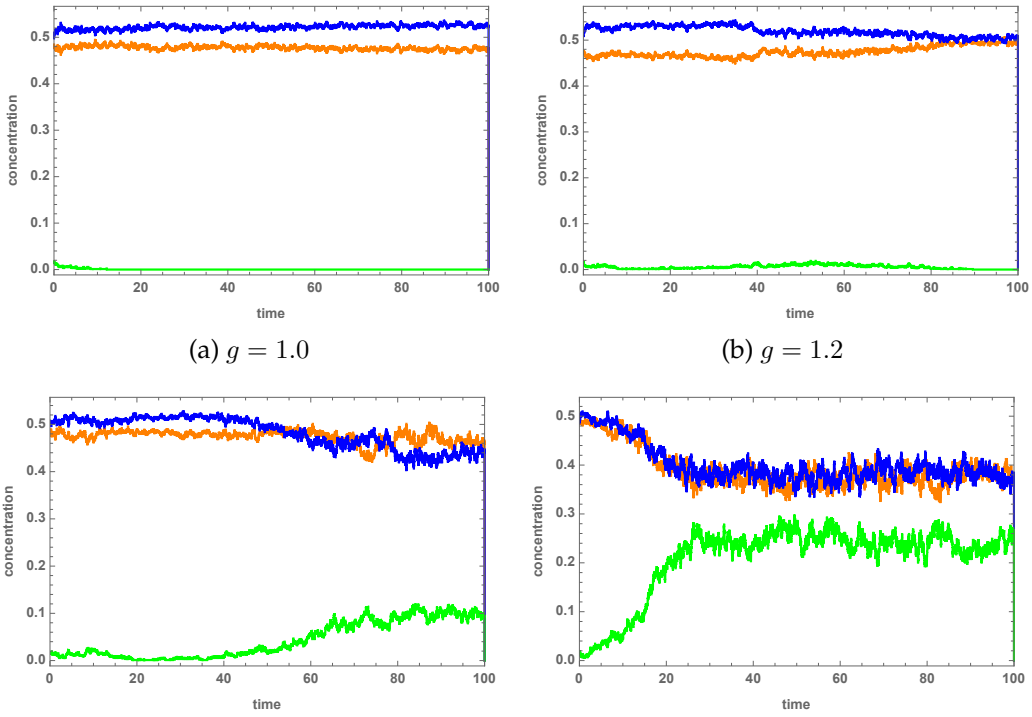


(c) $g = 2.0$



(d) $g = 3.5$

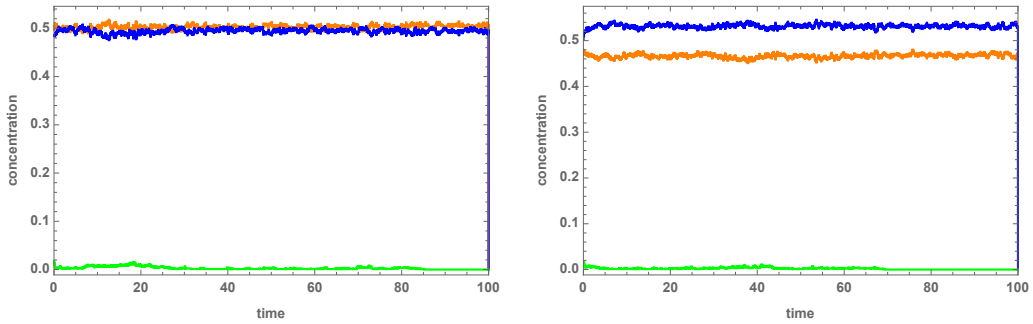
Figure 5.23: **Multiple Myeloma - Local Model**: $\beta = 2.0$; $\delta = 0$.



(a) $g = 1.0$ (b) $g = 1.2$

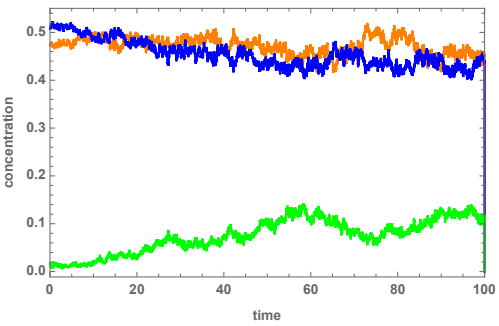
(c) $g = 2.0$ (d) $g = 3.5$

Figure 5.24: **Multiple Myeloma with initial segregation**: $\beta = 0.5$; $\delta = 0$.

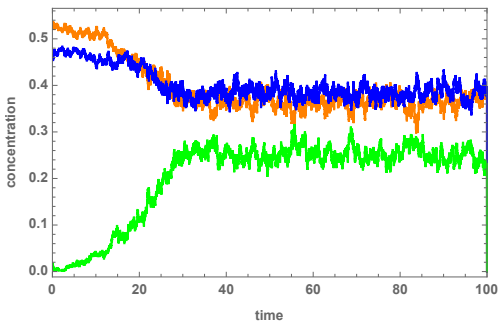


(a) $g = 1.0$

(b) $g = 1.2$

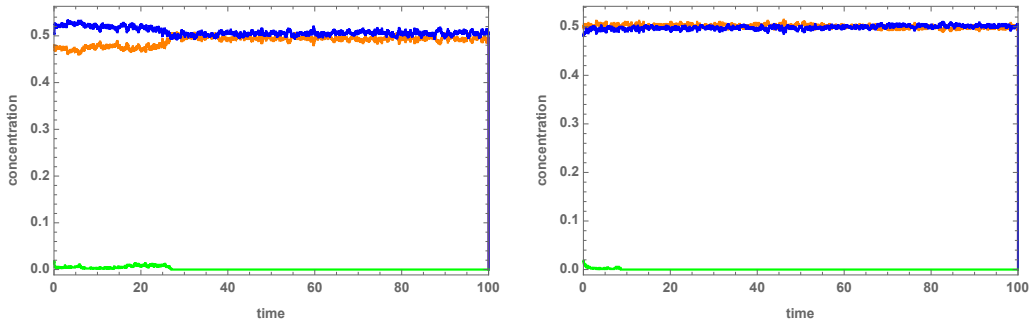


(c) $g = 2.0$



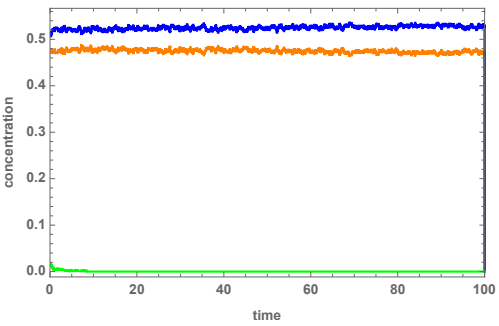
(d) $g = 3.5$

Figure 5.25: **Multiple Myeloma with initial segregation**: $\beta = 0.5$; $\delta = 0.3$

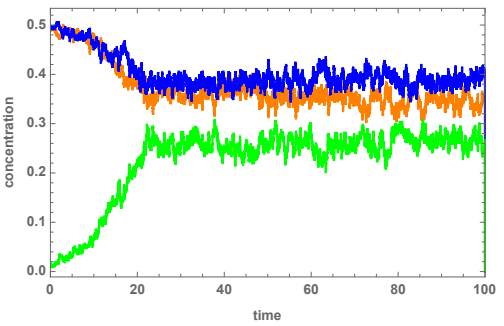


(a) $g = 1.0$

(b) $g = 1.2$

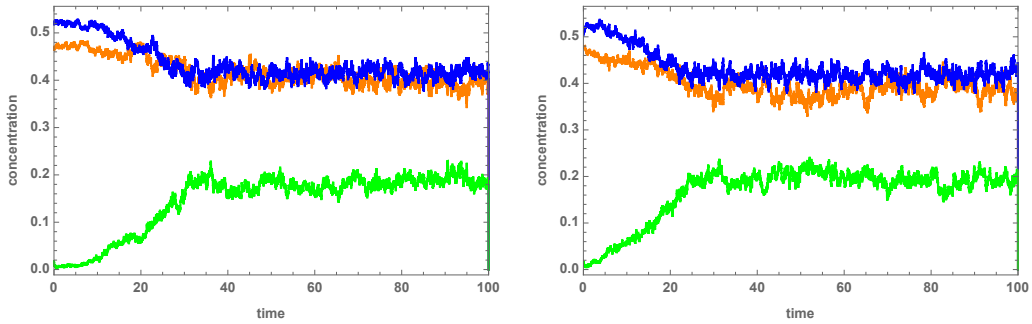


(c) $g = 2.0$



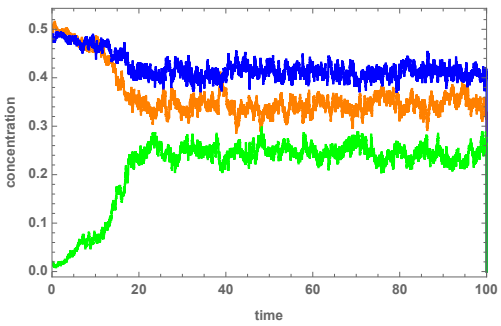
(d) $g = 3.5$

Figure 5.26: **Multiple Myeloma with initial segregation**: $\beta = 0.5$; $\delta = 1.0$

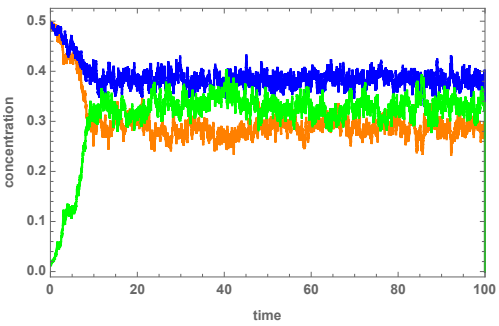


(a) $g = 1.0$

(b) $g = 1.2$



(c) $g = 2.0$



(d) $g = 3.5$

Figure 5.27: **Multiple Myeloma with initial segregation**: $\beta = 2.0$; $\delta = 0$.

APPENDIX A

Basic concepts of probability

Below we want to give a couple of definitions that could be useful to a reader that is not trained with probability theory. This appendix is not exhaustive, but collects just some concepts used in chapter VI.

Definition 5.3.1 (Independence) *Two events A and B are called independent if*

$$P(A \cap B) = P(A)P(B).$$

Two random variables X and Y are called independent if for all $a, b \in \mathbb{R}$,

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b).$$

Definition 5.3.2 (Conditional Probability) *The conditional probability of A given B is the number*

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

defined when $P(B) > 0$.

Definition 5.3.3 (Conditional Independence) *Two events A and B are conditionally independent given C if*

$$P(A \cap B|C) = P(A|C)P(B|C).$$

Let X, Y, Z be randonm variables taking their values in the denumerable sets $\mathcal{E}, \mathcal{F}, \mathcal{G}$, respectively. X and Y are conditionally independent given Z if, for all x, y, z in $\mathcal{E}, \mathcal{F}, \mathcal{G}$, respectively, events $\{X = x\}$ and $\{Y = y\}$ are conditionally independent given $\{Z = z\}$.

Theorem 5.3.1 (Markov property for events) *Let A_1, A_2, A_3 be three events of positive probability. Events A_1 and A_3 are conditionally independent given A_2 if and only if the "Markov property" holds, that is,*

$$P(A_3|A_1 \cap A_2) = P(A_3|A_2).$$

For a proof, see [96], Chapter I.

Theorem 5.3.2 (Markov Property for Random Variables) *Let X, Y, Z be three discrete random variables with values in $\mathcal{E}, \mathcal{F}$ and $\mathcal{G}$, respectively. If for some function $g : \mathcal{E} \times \mathcal{F} \to [0, 1]$, $P(X = x|Y = y, Z = z) = g(x, y)$ for all x, y, z, then $P(X = x|Y = y) = g(x, y)$ for all x, y and X and Y are conditionally independent given Z.*

For a proof, see [96], Chapter I.

Definition 5.3.4 (Stochastic Variables) *A stochastic, or random, variable is a function $X : \Omega \to \bar{\mathbb{R}}$ such that, for all $a \in \mathbb{R}$, the event $\{X \leq a\} = \{\omega; X(\omega) \leq a\}$ can be assigned a probability, that is,*

$$\{X \leq a\} \in \mathcal{F}.$$

A function $X : \Omega \to \mathcal{E}$, where $\mathcal{E}$ is a denumerable set is called a discrete random variable if for all $x \in \mathcal{E}$

$$\{X = x\} \in \mathcal{F}.$$

From a stochastic variable X derives and infinity of other stochastic variables, that is all quantities Y that are defined as functions of X by some mapping f.

Definition 5.3.5 (Stochastic Process) *These quantities Y may be any kind of mathematical object, in particular also functions of an additional variable t*

$$Y_X(t) = f(X, t).$$

Such a quantity $Y()t$ is called a random function or, since in most cases t stands for time, a stochastic process. Thus a stochastic process is a function of two variables, time t and a stochastic variable X.

Poisson Distribution

Suppose that you want to know how many times a binary event occur.

Just to make a naive example: if I flip a coin, how many tails (let us assume this means 1) can I, statistically, observe?

Consider a sequence $\{X_n\}_{n\geq 1}$ of random variables that can have two values, 0 or 1, with the same probability distribution

$$P(X_n = 1) = p,$$
$$P(X_n = 0) = 1 - p.$$

where $p \in (0, 1)$. Suppose, moreover, that the $\{X_n\}$ are independent. Since $P(X_j = a_j) = p$ or $1 - p$ depending upon $a_i = 1$ or 0, and since there are exactly $\sum_{j=1}^{n} a_j$ numbers among $a_1, \ldots, a_n$ that are equal to 1,

$$P(X_1 = a_1, \ldots, X_k = a_n) = p^{\sum_{j=1}^{n} a_j}(1 - p)^{n - \sum_{j=1}^{n} a_j}.$$

Let us define

$$S_N = X_i + \cdots + X_N$$

as a succession of N random binary events. For sake of simplicity, we assume that 1 corresponds to a success, while 0 to a failure. The probability of observing, after N repetitions, a succession of k successes and $N - k$ failures is

$$P(k) = P(S_N = k) = \binom{N}{k} p^k (1 - p)^{N-k},$$

where $\binom{N}{k} = \frac{N!}{k!(N-k)!}$ is the binomial coefficient. This is the well-known *binomial distribution*.

Now we ask ourselves a slightly different question: if I flip a coin, how many tails can I, statistically, observe, knowing the average number of times that this event occurs?

For a large number of turns N, the probability distribution is the following

$$P(X_n = 1) = \frac{\alpha}{N},$$

where α is the known average per time unit. Thus, the number S_N of tails is distributed according to

$$P(k) = P(S_N = k) = \binom{N}{k} \left(\frac{\alpha}{N}\right)^k \left(1 - \frac{\alpha}{N}\right)^{N-k}.$$

The limit of the distribution of S_N as $N \to \infty$ is

$$\lim_{N\to\infty} = e^{-\alpha} \frac{\alpha^k}{k!},$$

that is the *Poisson distribution*. By definition, a Poisson random variable with parameter $\theta > 0$ is a random variable such that, for all $k \geq 0$,

$$P(X = k) = e^{-\theta} \frac{\theta^k}{k!}.$$

Let X_1 and X_2 be two independent Poisson random variables with means $\theta_1 > 0$ and $\theta_2 > 0$ respectively. Then $X = X_1 + X_2$ is a Poisson random variable with mean $\theta = \theta_1 + \theta_2$.

Let $\{X_n\}_{n\geq 1}$ be independent random variables taking values 0 and 1 with probability $(1 - p)$ and p respectively, where $p \in (0, 1)$.
Let T be a Poisson random variable with mean $\theta > 0$, independent of $(X_n)_{n\geq 1}$. Define

$$S = X_1 + \ldots X_T.$$

Then S is a Poisson random variable with mean $p\,\theta$.

APPENDIX B

EGT and Cell Populations Dynamics

> Evolutionary game dynamics results from the transfer of economic ideas to biology. In economics, rational players try to find the best strategy to maximize their payoffs. In biology, those individuals who use the best strategy obtain the highest reproductive fitness and spread in the population. Traditionally, evolutionary game dynamics is considered in infinitely large, well–mixed populations. This typically leads to the replicator dynamics, a system of nonlinear differential equations governing the evolutionary dynamics. For any composition of the population, the replicator dynamics determines deterministically the direction and velocity of evolutionary dynamics. The replicator dynamics can be derived from microscopic models of strategy spreading, which are typically stochastic. [113]

First of all, it would be appropriate to introduce the main equation of EGT, that is, the *replicator equation* [114].

Let us assume that a population is divided into n different species, $M_1 ... M_n$ with corresponding frequencies $x_1, \ldots, x_n$. The fitness of the i-th population, f_i, is a function of the whole population, that is, of $\mathbf{x} = (x_1, ..., x_n)$. If the population is very large, and if the generations blend continuously into each other, we may assume that the state $x(t)$ evolves as a differentiable function of t. The rate of increase of a species, $\dot{x}_i/x_i$, can be considered as a measure of the evolutionary success of the i-th species. Thus, we may express it as the difference between the fitness $f_i(\mathbf{x})$ and the average fitness $\bar{f}(\mathbf{x}) = \sum x_i f_i(\mathbf{x})$

$$\frac{\dot{x}_i}{x_i} = f_i(\mathbf{x}) - \bar{f}(\mathbf{x}),$$

which yields to the replicator equation

$$\dot{x}_i = x_i \left(f_i(\mathbf{x}) - \bar{f}(\mathbf{x}) \right).$$

Let us now apply this concept to the case of cancer. We consider a population A of normal cells of initial population $N_A(0)$: every cell of this population replicate itself with a constant rate α. Assuming infinite resources, the equation that describe the time evolution of the population A is

$$\begin{aligned} \frac{\mathrm{d}N_A(t)}{\mathrm{d}t} &= \alpha N_A(t), \\ N_A(t) &= N_A(0)\, e^{\alpha t}. \end{aligned}$$

In the same way, for a population B composed by malignant mutated cells, with constant reproduction rate β

$$\begin{aligned} \frac{\mathrm{d}N_B(t)}{\mathrm{d}t} &= \beta N_B(t), \\ N_B(t) &= N_B(0)\, e^{\beta t}. \end{aligned}$$

In this unrealistic scenario, in which the overall population grows exponentially, there are two possible situations: if $\alpha/\beta > 1$, the normal cell population will overcome the mutant population; if $\alpha/\beta < 1$, the opposite happens.

To make the scenario a little bit more realistic, we can assume that the available resources impose a constant population size (*carrying capacity*). We can include this assumption in the above equations

$$\frac{\mathrm{d}N_A(t)}{\mathrm{d}t} = N_A(t)(\alpha - \omega),$$
$$\frac{\mathrm{d}N_B(t)}{\mathrm{d}t} = N_B(t)(\beta - \omega).$$

Imposing the conservation of total population size ($N_T \equiv const. = N_A(t) + N_B(t)$) leads to

$$N_T\, \omega = N_A(t)\, \alpha + N_B(t)\, \beta.$$

That is, ω stands for the average reproductive rate of population and all that matters is how α and β compare with ω: the population that reproduces with a rate bigger than the average, will outgrow the other. To describe the evolution of a two-population dynamics, due to the fact that we consider the total size of the population conserved, we need just one equation. Choosing the population of healthy cell, A, we have

$$\frac{\mathrm{d}N_A(t)}{\mathrm{d}t} = N_A(t)(N_T - N_A(t))(\alpha - \beta).$$

If population size is large enough, is possible to convert the cell number into frequencies x

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = x(t)(1 - x(t))(\alpha - \beta).$$

This nonlinear ordinary differential equation is a particular case of replicator equation. It describes the evolutionary dynamics of two species with constant rate of reproduction, assuming that the total population size is conserved. The rate of reproduction, at this stage, can be considered as

a measure of cell fitness. Indeed, this crude assumption is pretty rough for different reason: firstly, the cell population with higher fitness will certainly overgrow the other; secondly, the fitness of cells remains constant in time. Moreover, it doesn't matter if a population is composed by 10 or a million cells, the replication rate remains the same.

EGT provides one possible means to overcome this shortcomings: to make the fitness frequency-dependent. Replacing α with $\phi_A(x)$ and β with $\phi_B(x)$, it is possible to write down a general form of replicator equation for this scenario

$$\dot{x} = x(1-x)(\phi_A(x) - \phi_B(x)).$$

This simple form of frequency-dependence of the fitness derives from a mean field (or well-mixed) approximation, that is, every cell interacts with any other cell at the same time.

The interaction of two cells can be computed by inspection of the so-called *pay-off matrix*

$$\begin{bmatrix} p_{AA} & p_{AB} \\ p_{BA} & p_{BB} \end{bmatrix}$$

where p_{ij} stands for what a cell of type i gets from interacting with a cell of type j. Interaction here is understood in a very general sense, to include competition for space and nutrients as well as exchange of information through cytokines and growth factors. It can be assumed that this is how the fitness of a cell is, directly or indirectly, modified when this cell is in the presence of another cell (which can be of the same type or of another type). The average fitness is then

$$\phi_A(x) = x\, p_{AA} + (1-x)\, p_{AB},$$
$$\phi_B(x) = x\, p_{BA} + (1-x)\, p_{BB}.$$

Naturally, the previous situation, without frequency-dependent cell fitness, can be easily recovered setting $p_{AA} = p_{AB} = \alpha$ and $p_{BA} = p_{BB} = \beta$. It is obviously possible to generalize these results for the case of three cell types: there will be two replicator equations, with two independent frequencies and the pay-off matrix will be a 3×3 matrix.

GLOSSARY

Here we would try to give an explanation of the main terms, that we have used throughout the entire work, whose meaning is not so intuitive or well-known. This glossary does not pretend to be exhaustive, but just to be an helpful tool to understand also the clinical and the biological sides of the problem.

Definitions are taken from [116].

- **Chemokine**
  A diverse family of small secreted proteins that are chemotactic for leucocytes, a subset of the cytokines.

- **Clone**
  A population of cells or organisms derived from a single progenitor and therefore genetically identical.

- **Cytokine**
  A rather loose category of small proteins that are released by cells and that affect the behaviour of other cells. Normally taken to include interleukins, lymphokines, chemokines and several related signalling molecules such as tumour necrosis factor and interferons but not hormones or growth factors except perhaps transforming growth factor β (TGFβ).

- **Fibroblast**
  The mesodermally derived cells of connective tissue that are responsible for the synthesis and secretion of procollagen, fibronectin, collagenase, and extracellular matrix components.

- **Growth Factor**
  A diverse group of proteins that are important in the regulation of cell proliferation (growth) and differentiation. The distinction between growth factors and cytokines is blurred since some cytokines act as growth factors and some cytokines, originally described as important in the haematopoietic lineages, act on a broader range of cell types. Autonomous growth factor production or altered responsiveness to growth factors is a common characteristic of many neoplastic cells which thereby lose growth control.
  Trasforming Growth Factor
  A family of growth factors secreted by transformed cells that induce the phenotypic characteristics of cell transformation (e.g. the ability to grow in semisolid agar), but do not cause hereditable changes.

- **Ligand**
  Any molecule that binds to another; generally the smaller of the two, especially if there is no signal transduction by the receptor. Ligands are not necessarily small molecules such as hormones or neurotransmitters; sugar residues attached to proteins or lipids and incorporated in the membrane are often referred to as ligands for lectins.

- **Mesenchyme**
  Embryonic tissue that derives from mesoderm.

- **Mitosis**
  The process of nuclear division in the somatic cells of eukaryotes in which the genomic information, copied during S phase of the cell cycle, is distributed equally between two daughter cells so that each contains a diploid set of chromosomes identical to that of the parent cell. In prophase the nuclear envelope breaks down, the chromosomes condense, and the two chromatids become visible.

The centriole divides and the two daughter centrioles separate to define the two poles of the mitotic spindle. The spindle consists of interdigitated microtubules (spindle fibres) nucleated from the two pericentriolar organizing centres together with microtubules attached to kinetochores of chromosomes. In the next stage, metaphase, the chromosomes align on the equatorial plane of the spindle, the metaphase plate. Then in anaphase the paired chromatids separate, one toward each pole, a process that is completed during telophase. The chromatids, now considered to be chromosomes, lengthen and become diffuse, new nuclear envelopes form round the two sets of chromosomes, and the spindle disassembles. Mitosis is usually followed by cell division or cytokinesis in which the cytoplasm is also divided to give two daughter cells.

- **Neoplasia**
  Literally, a term meaning 'new growth' but referring to abnormal new growth that persists in the absence of the original stimulus. The term covers both tumours, where there is an actual swelling, and other proliferative disorders, such as leukaemias, all colloquially referred to as "cancer". The cells in benign tumours do not spread and such tumours are not life-threatening unless the growth pressure causes damage. Malignant neoplasia involves the loss of both growth control and positional control and the malignant cells invade territory normally occupied by other cells.

- **Parathyroid Hormone (PTH)**
  A peptide hormone (parathormone, parathyrin, 84 aa) secreted by the parathyroid glands. Parathyroid hormone acts on osteoclasts and causes an increase in blood calcium ion concentrations.

- **Parenchyma**
  The functional tissue of an organ, as opposed to the connective tissue (stroma).

- **Receptor**
  At a cellular level, an immobilized molecule, usually a membrane-

bound or membrane-enclosed protein, that binds to, or responds to, a smaller and more mobile ligand. Specificity of the interaction is implicit.

- **Repressor protein**
  A DNA-binding protein that binds to the operator region and blocks transcription. Originally described as a control mechanism for regulating bacterial gene expression (see lactose operon) but analogous mechanisms exist for eukaryotes, although they are more usually thought of as transcription factors.

- **T-cell or T lymphocyte**
  One of the two major lymphocyte classes, those that are of thymic origin. T-cells are involved in cell-mediated responses and regulated B-cells development. They have antigen receptors. Various subsets are now recognized T-helper cells, T-regulatory cells, etc.

- **T-helper cell**
  Classically two subclasses of T-helper cells, Th1 and Th2, have been recognized. Th1 cells are responsible for clearing intracellular pathogens and are involved in cell-mediated immunity. They produce IL-2, interferon-γ, and TNFα but not IL-4, IL-5, and IL-10. Selective activation of Th1 cells is promoted by interferon-γ and IL-12 and inhibited by IL-4 and IL-10, the products of Th2 cells. Th2 cells are involved with the humoral immune response, produce IL-4, IL-5, and IL-10, and promote antibody production; IL-4 is essential for growth and differentiation of Th2 cells. There is cross-inhibition between the two classes; if one subclass is activated it will inhibit the activity of the other so that the response is polarized. More recently a third class of Th cells that produce IL-17 in response to autoimmune tissue damage has been identified.

- **Transcription factor**
  A protein that binds to a specific DNA sequence upstream of a coding region and triggers the assembly of an RNA polymerase complex

and the production of mRNA or other RNA species. A range of different transcription factors are known.

- **Tyrosine kinase**
  A protein kinase that phosphorylates tyrosine residues. There are two subfamilies: receptor tyrosine kinases with an extracellular ligand-binding domain and an intracellular tyrosine kinase domain; and nonreceptor tyrosine kinases, which are soluble, cytoplasmic kinases. Both forms play important roles in signalling systems.

- **Stroma**
  Loose connective tissue that has few resident cells (stromal cells).
  Tumor stroma
  Compartment providing the connective-tissue framework of the tumour. It includes fibroblasts, immune and inflammatory cells, fat cells and blood-vessel cells.

L'originalità è l'arte
di nascondere le proprie fonti.

LIST OF FIGURES

BIBLIOGRAPHY

[1] D. Hanahan and R. A. Weinberg. "The hallmarks of cancer". In: *Cell* 100 (2000), pp. 57–70. DOI: 10.1016/S0092-8674(00)81683-9.

[2] M. R. Stratton, P. J. Campbell, and P. A. Futreal. "The cancer genome". In: *Nature* 458 (Apr. 2009), 719 EP -. DOI: https://doi.org/10.1038/nature07943.

[3] C. La Porta and S. Zapperi. *The Physics of Cancer*. Cambridge: Cambridge University Press, 2017. ISBN: 978-1-107-10959-9.

[4] N. Beerenwinkel et al. "Genetic progression and the waiting time to cancer". In: *PLOS Computational Biology* 3.11 (Nov. 2007), pp. 1–8. DOI: 10.1371/journal.pcbi.0030225.

[5] D. Hanahan and R. A. Weinberg. "Hallmarks of cancer: the next generation". In: *Cell* 5 (2011), pp. 646–674. DOI: https://doi.org/10.1016/j.cell.2011.02.013.

[6] M. M. Mueller and N. E. Fusenig. "Friends or foes—bipolar effects of the tumour stroma in cancer". In: *Nature Reviews Cancer* 4.11 (2004), p. 839.

[7] C. D. McFarland et al. “Impact of deleterious passenger mutations on cancer progression”. In: *Proceedings of the National Academy of Sciences* (2013). DOI: 10.1073/pnas.1213968110.

[8] C. Tomasetti and B. Vogelstein. “Variation in cancer risk among tissues can be explained by the number of stem cell divisions”. In: *Science* 347.6217 (2015), pp. 78–81. DOI: 10.1126/science.1260825.

[9] D. Wodarz and A. G. Zauber. “Cancer: Risk factors and random chances”. In: *Nature* 517.7536 (2015), p. 563.

[10] N. A. Ashford et al. “Cancer risk: role of environment”. In: *Science* 347.6223 (2015), pp. 727–727.

[11] J. D. Potter and R. L. Prentice. “Cancer risk: tumors excluded”. In: *Science* 347.6223 (2015), pp. 727–727.

[12] C. Wild et al. “Cancer risk: role of chance overstated”. In: *Science* 347.6223 (2015), pp. 728–728.

[13] C. Gotay, T. Dummer, and J. Spinelli. “Cancer risk: prevention is crucial”. In: *Science* 347.6223 (2015), pp. 728–728.

[14] M. Song and E. L. Giovannucci. “Cancer risk: many factors contribute”. In: *Science* 347.6223 (2015), pp. 728–729.

[15] M. O’Callaghan. “Cancer risk: accuracy of literature”. In: *Science* 347.6223 (2015), pp. 729–729.

[16] P. Carmeliet and R. K. Jain. “Angiogenesis in cancer and other diseases”. In: *nature* 407.6801 (2000), p. 249.

[17] D. X. Nguyen, P. D. Bos, and J. Massagué. “Metastasis: from dissemination to organ-specific colonization”. In: *Nature Reviews Cancer* 9.4 (2009), p. 274.

[18] A. R. A. Anderson and V. Quaranta. “Integrative mathematical oncology”. In: *Nature Reviews Cancer* 8 (Mar. 2008), 227 EP -. DOI: https://doi.org/10.1038/nrc2329.

[19] F. Michor et al. “What does physics have to do with cancer?” In: *Nature Reviews Cancer* 11 (Aug. 2011), 657 EP -. DOI: https://doi.org/10.1038/nrc3092.

[20] H. M. Byrne. “Dissecting cancer through mathematics: from the cell to the animal model”. In: *Nature Reviews Cancer* 10 (Mar. 2010), 221 EP -. DOI: https://doi.org/10.1038/nrc2808.

[21] T. Jackson, N. Komarova, and K. Swanson. “Mathematical oncology: using mathematics to enable cancer discoveries”. In: *The American Mathematical Monthly* 121.9 (2014), pp. 840–856. DOI: 10.4169/amer.math.monthly.121.09.840.

[22] A. Anderson and K. Rejniak. *Single-cell-based models in biology and medicine*. Springer Science & Business Media, 2007.

[23] R. Araujo and D. McElwain. “A history of the study of solid tumour growth: the contribution of mathematical modelling”. In: *Bulletin of Mathematical Biology* 66.5 (2004), pp. 1039–1091. DOI: https://doi.org/10.1016/j.bulm.2003.11.002.

[24] M. Bachar, J. Batzel, and S. Ditlevsen. *Stochastic biomathematical models with application to neuronal modeling*. Springer, 2013. ISBN: 978-3-642-32156-6.

[25] P. Haccou, P. Jagers, and V. A. Vatutin. *Branching rocesses - Variation, growth and extinction of populations*. Cambridge: Cambridge University Press, 2005. ISBN: 978-0-521-83220-5.

[26] P. M. Altrock, L. L. Liu, and F. Michor. “The mathematics of cancer: integrating quantitative models”. In: *Nature Reviews Cancer* 15.12 (2015), p. 730.

[27] J. A. Adam and N. Bellomo. *A survey of models for tumor-immune system dynamics*. Springer Science & Business Media, 2012.

[28] M. Kimmel and D. E. Axelrod. *Branching processes in biology*. New Tork: Springer, 2002. ISBN: 0-387-95340-X.

[29] T. Antal and P. Krapivsky. “Exact solution of a two-type branching process: models of tumor progression”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2011.08 (2011), P08018.

[30] I. Bozic et al. “Accumulation of driver and passenger mutations during tumor progression”. In: *Proceedings of the National Academy of Sciences* 107.43 (2010), pp. 18545–18550.

[31] B. Bauer, R. Siebert, and A. Traulsen. “Cancer initiation with epistatic interactions between driver and passenger mutations”. In: *Journal of theoretical biology* 358 (2014), pp. 52–60.

[32] C. Tomasetti, B. Vogelstein, and G. Parmigiani. “Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation”. In: *Proceedings of the National Academy of Sciences* 110.6 (2013), pp. 1999–2004.

[33] J. Foo, K. Leder, and F. Michor. “Stochastic dynamics of cancer initiation”. In: *Physical biology* 8.1 (2011), p. 015002.

[34] J. Foo et al. “An evolutionary approach for identifying driver mutations in colorectal cancer”. In: *PLoS computational biology* 11.9 (2015), e1004350.

[35] B. Werner, D. Dingli, and A. Traulsen. “A deterministic model for the occurrence and dynamics of multiple mutations in hierarchically organized tissues”. In: *Journal of The Royal Society Interface* 10.85 (2013), p. 20130349.

[36] H. Haeno et al. “A progenitor cell origin of myeloid malignancies”. In: *Proceedings of the National Academy of Sciences* 106.39 (2009), pp. 16616–16621.

[37] R. Ganguly and I. Puri. “Mathematical model for the cancer stem cell hypothesis”. In: *Cell proliferation* 39.1 (2006), pp. 3–14.

[38] H. M. Byrne and M. Chaplain. “Growth of necrotic tumors in the presence and absence of inhibitors”. In: *Mathematical biosciences* 135.2 (1996), pp. 187–216.

[39] J. A. Sherratt and M. A. Chaplain. “A new mathematical model for avascular tumour growth”. In: *Journal of mathematical biology* 43.4 (2001), pp. 291–312.

[40] S. Ferreira Jr, M. Martins, and M. Vilela. "Reaction-diffusion model for the growth of avascular tumor". In: *Physical Review E* 65.2 (2002), p. 021907.

[41] X. Li et al. "Nonlinear three-dimensional simulation of solid tumor growth". In: *Discrete and continuous dynamical systems* 7.3 (2007), p. 581.

[42] B. Waclaw et al. "A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity". In: *Nature* 525.7568 (2015), p. 261.

[43] K. R. Swanson et al. "Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion". In: *Journal of the neurological sciences* 216.1 (2003), pp. 1–10.

[44] T. E. Yankeelov et al. "Clinically relevant modeling of tumor growth and treatment response". In: *Science translational medicine* 5.187 (2013), 187ps9–187ps9.

[45] E. R. Fearon and B. Vogelstein. "A genetic model for colorectal tumorigenesis". In: *cell* 61.5 (1990), pp. 759–767.

[46] I. C. G. Consortium et al. "International network of cancer genome projects". In: *Nature* 464.7291 (2010), p. 993.

[47] S. A. Forbes et al. "COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer". In: *Nucleic acids research* 39.suppl_1 (2010), pp. D945–D950.

[48] R. Desper et al. "Inferring tree models for oncogenesis from comparative genome hybridization data". In: *Journal of computational biology* 6.1 (1999), pp. 37–51.

[49] R. Desper et al. "Distance-based reconstruction of tree models for oncogenesis". In: *Journal of Computational Biology* 7.6 (2000), pp. 789–803.

[50] M. Hjelm, M. Höglund, and J. Lagergren. "New probabilistic network models and algorithms for oncogenesis". In: *Journal of Computational Biology* 13.4 (2006), pp. 853–865.

[51] C. S. O. Attolini et al. "A mathematical framework to determine the temporal sequence of somatic genetic events in cancer". In: *Proceedings of the National Academy of Sciences* (2010).

[52] K. Sprouffske, J. W. Pepper, and C. C. Maley. "Accurate reconstruction of the temporal order of mutations in neoplastic progression". In: *Cancer prevention research* (2011), canprevres–0374.

[53] K. A. Rejniak and A. R. Anderson. "Hybrid models of tumor growth". In: *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 3.1 (2011), pp. 115–125.

[54] A. R. Anderson. "A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion". In: *Mathematical medicine and biology: a journal of the IMA* 22.2 (2005), pp. 163–186.

[55] P. Gerlee and A. R. Anderson. "A hybrid cellular automaton model of clonal evolution in cancer: the emergence of the glycolytic phenotype". In: *Journal of theoretical biology* 250.4 (2008), pp. 705–722.

[56] A. R. Anderson and M. Chaplain. "Continuous and discrete mathematical models of tumor-induced angiogenesis". In: *Bulletin of mathematical biology* 60.5 (1998), pp. 857–899.

[57] M. A. Chaplain, S. R. McDougall, and A. Anderson. "Mathematical modeling of tumor-induced angiogenesis". In: *Annu. Rev. Biomed. Eng.* 8 (2006), pp. 233–257.

[58] M. Orme and M. A. Chaplain. "Two-dimensional models of tumour angiogenesis and anti-angiogenesis strategies". In: *Mathematical Medicine and Biology* 14.3 (1997), pp. 189–205.

[59] H. B. Frieboes et al. "Three-dimensional multispecies nonlinear tumor growth—II: tumor invasion and angiogenesis". In: *Journal of theoretical biology* 264.4 (2010), pp. 1254–1278.

[60] P. Macklin et al. "Multiscale modelling and nonlinear simulation of vascular tumour growth". In: *Journal of mathematical biology* 58.4-5 (2009), pp. 765–798.

[61] S. R. McDougall, A. R. Anderson, and M. A. Chaplain. "Mathematical modelling of dynamic adaptive tumour-induced angiogenesis: clinical implications and therapeutic targeting strategies". In: *Journal of theoretical biology* 241.3 (2006), pp. 564–589.

[62] I. J. Fidler. "The pathogenesis of cancer metastasis: the'seed and soil'hypothesis revisited". In: *Nature Reviews Cancer* 3.6 (2003), p. 453.

[63] K. J. Pienta et al. "The Cancer Diaspora: Metastasis beyond the seed and soil hypothesis." In: *Clinical cancer research* (2013), clincanres–2158.

[64] G. M. Saidel, L. A. Liotta, and J. Kleinerman. "System dynamics of a metastatic process from an implanted tumor". In: *Journal of theoretical biology* 56.2 (1976), pp. 417–434.

[65] H. Enderling et al. "Paradoxical dependencies of tumor dormancy and progression on basic cell kinetics". In: *Cancer research* (2009), pp. 0008–5472.

[66] P. K. Newton et al. "Spreaders and sponges define metastasis in lung cancer: a Markov chain Monte Carlo mathematical model". In: *Cancer research* (2013).

[67] J. G. Scott et al. "A mathematical model of tumour self-seeding reveals secondary metastatic deposits as drivers of primary tumour growth". In: *Journal of The Royal Society Interface* 10.82 (2013), p. 20130011.

[68] H. Haeno and F. Michor. "The evolution of tumor metastases during clonal expansion". In: *Journal of theoretical biology* 263.1 (2010), pp. 30–44.

[69] V. Almendro et al. "Genetic and phenotypic diversity in breast tumor metastases". In: *Cancer research* (2014).

[70] J. Scott, P. Kuhn, and A. R. Anderson. "Unifying metastasis—integrating intravasation, circulation and end-organ colonization". In: *Nature Reviews Cancer* 12.7 (2012), p. 445.

[71] P. K. Newton et al. "A stochastic Markov chain model to describe lung cancer growth and metastasis". In: *PloS one* 7.4 (2012), e34637.

[72] A. Pathak and S. Kumar. “Independent regulation of tumor cell migration by matrix stiffness and confinement”. In: *Proceedings of the National Academy of Sciences* 109.26 (2012), pp. 10334–10339.

[73] T. E. Yankeelov et al. “Toward a science of tumor forecasting for clinical oncology”. In: *Cancer research* (2015).

[74] R. G. Dale. “The application of the linear-quadratic dose-effect equation to fractionated and protracted radiotherapy”. In: *The British journal of radiology* 58.690 (1985), pp. 515–528.

[75] R. Rockne et al. “Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: a mathematical modeling approach”. In: *Physics in Medicine & Biology* 55.12 (2010), p. 3271.

[76] D. Corwin et al. “Toward patient-specific, biologically optimized radiation therapy plans for the treatment of glioblastoma”. In: *PloS one* 8.11 (2013), e79115.

[77] M. L. Neal et al. “Response classification based on a minimal model of glioblastoma growth is prognostic for clinical outcomes and distinguishes progression from pseudoprogression”. In: *Cancer research* (2013).

[78] J. N. Blattman and P. D. Greenberg. “Cancer immunotherapy: a treatment for the masses”. In: *Science* 305.5681 (2004), pp. 200–205.

[79] N. Bellomo and L. Preziosi. “Modelling and mathematical problems related to tumor evolution and its interaction with the immune system”. In: *Mathematical and Computer Modelling* 32.3-4 (2000), pp. 413–452.

[80] R. A. Gatenby et al. “Adaptive therapy”. In: *Cancer research* 69.11 (2009), pp. 4894–4903.

[81] I. Bozic et al. “Evolutionary dynamics of cancer in response to targeted combination therapy”. In: *elife* 2 (2013), e00747.

[82] S. E. Luria and M. Delbrück. “Mutations of bacteria from virus sensitivity to virus resistance”. In: *Genetics* 28.6 (1943), p. 491.

[83] Y. Iwasa, F. Michor, and M. A. Nowak. "Evolutionary dynamics of escape from biomedical intervention". In: *Proceedings of the Royal Society of London B: Biological Sciences* 270.1533 (2003), pp. 2573–2578.

[84] A. Coldman and J. Goldie. "A stochastic model for the origin and treatment of tumors containing drug-resistant cells". In: *Bulletin of mathematical biology* 48.3-4 (1986), pp. 279–292.

[85] J. H. Goldie and A. J. Coldman. "The genetic origin of drug resistance in neoplasms: implications for systemic therapy". In: *Cancer research* 44.9 (1984), pp. 3643–3653.

[86] N. L. Komarova and D. Wodarz. "Drug resistance in cancer: principles of emergence and prevention". In: *Proceedings of the National Academy of Sciences* 102.27 (2005), pp. 9714–9719.

[87] I. Kareva. "What can ecology teach us about cancer?" In: *Translational oncology* 4.5 (2011), pp. 266–270.

[88] B. Ujvari, B. Roche, and F. Thomas. *Ecology and Evolution of Cancer*. London: Academic Press - Elsevier, 2017. ISBN: 978-0-12-804310-3.

[89] K. S. Korolev, J. B. Xavier, and J. Gore. "Turning ecology and evolution against cancer". In: *Nature Reviews Cancer* 14.5 (2014), p. 371.

[90] P. Rosati, R. Colombo, and N. Maraldi. *Istologia*. Milano: Edi Ermes, 2006. ISBN: 88-7051-294-0.

[91] D. L. Kasper et al. *Harrison. Principi di Medicina Interna - 16esima edizione*. Milano: McGraw-Hill, 2005. ISBN: 88-386-2999-4.

[92] V. Kumar, A. K. Abbas, and J. C. Aster. *Robbins e Cotran. Basi Patologiche delle Malattie*. Milano: Elsevier Italia, 2006. ISBN: 978-1-4557-2613-4.

[93] G. D. Roodman. "Pathogenesis of myeloma bone disease". In: *Leukemia* 23 (Nov. 2008), 435 EP -. DOI: https://doi.org/10.1038/leu.2008.336.

[94] E. Terpos et al. “Pathogenesis of bone disease in multiple myeloma: from bench to bedside”. In: *Blood Cancer Journal* 8.1 (2018), p. 7. DOI: 10.1038/s41408-017-0037-4.

[95] N. G. Van Kampen. *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier, 1992.

[96] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Vol. 31. Springer Science & Business Media, 2013.

[97] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*. Vol. 282. John Wiley & Sons, 2009.

[98] T. M. Liggett. *Interacting particle systems*. Vol. 276. Springer Science & Business Media, 2012.

[99] J. M. Swart. “A course in interacting particle systems”. In: *arXiv preprint arXiv:1703.10007* (2017).

[100] S. Grosskinsky. “Interacting particle systems”. In: *Notes* (2009).

[101] M. Kochmański, T. Paszkiewicz, and S. Wolski. “Curie-Weiss magnet—a simple model of phase transition”. In: *European Journal of Physics* 34 (Nov. 2013). DOI: 10.1088/0143-0807/34/6/1555.

[102] P. L. Krapivsky, S. Redner, and E. Ben-Naim. *A kinetic view of statistical physics*. Cambridge University Press, 2010.

[103] S. Redner. “Reality Inspired Voter Models: A Mini-review”. In: *arXiv e-prints* (Nov. 2018), arXiv:1811.11888.

[104] C. Gardiner. *Handbook of stochastic processes*. Springer-Verlag, New York, 1985.

[105] D. Tovazzi. “Self-sustained periodic behaviors in interacting systems: macroscopic limits and fluctuations”. Tesi di Dottorato di Ricerca in Scienze Matematiche. Università degli Studi di Padova.

[106] J. M. Pacheco, F. C. Santos, and D. Dingli. “The ecology of cancer from an evolutionary game theory perspective”. In: *Interface Focus* 4.4 (2014), p. 20140019. DOI: 10.1098/rsfs.2014.0019.

[107] J. A. Lust et al. “Induction of a chronic disease state in patients with smoldering or indolent multiple myeloma by targeting interleukin 1β-induced interleukin 6 production and the myeloma proliferative component”. In: 84.2 (2009), pp. 114–122.

[108] A. Bazzani. “Fisica dei sistemi complessi”. Dispense del corso.

[109] E. Ferrero. “Processi stocastici”. Dispense del corso.

[110] A. Celani. *Processi stocastici: istruzioni per l'uso*. 2010.

[111] C. A. Dinarello. “Targeting the pathogenic role of interleukin 1β in the progression of smoldering/indolent myeloma to active disease”. In: 84.2 (2009), pp. 105–107.

[112] C. Tu et al. “Reconciling cooperation, biodiversity and stability in complex ecological communities”. In: *arXiv preprint arXiv:1805.03527* (2018).

[113] P. M. Altrock and A. Traulsen. “Deterministic evolutionary game dynamics in finite populations”. In: *Physical Review E* 80.1 (2009), p. 011909.

[114] J. Hofbauer and K. Sigmund. *Evolutionary games and population dynamics*. Cambridge: Cambridge University Press, 1998.

[115] M. Agnolon. “Ecology of cancer: an evolutionary game theory approach to model cancer growth”. Tesi di Laurea Triennale in Matematica. Università degli Studi di Padova.

[116] J. Lackie. *A dictionary of biomedicine*. Oxford University Press, 2010.

Se cerchi la tua strada verso Itaca,
Spera in un viaggio lungo,
Avventuroso e pieno di scoperte.
I Lestrígoni e i Ciclopi non temerli,
Non temere l'ira di Poseidone.
Mai troverai tali mostri sulla tua via,
Se non li porti dentro, nel tuo cuore
Se l'anima non te li mette contro.

Devi augurarti che la strada sia lunga.
Che siano tanti i mattini d'estate,
Quando nei porti - finalmente e con che gioia -
Toccherai terra tu per la prima volta:
Negli empori fenici indugia e
Acquista madreperle coralli,
Ebano e ambre,
Tutta merce fina
E voluttuosi aromi d'ogni sorta;
Più profumi inebrianti che puoi.
Visita molte città egizie.
Impara quello che puoi dai sapienti.

Pensa ad Itaca, sempre
Il tuo destino ti ci porterà
Non hai bisogno di affrettare il corso
Fa' che il tuo viaggio duri anni, bellissimi,
E che tu giunga all'isola ormai vecchio,
Ricco di insegnamenti appresi per via,
Non sperare ti giungano ricchezze,
Il regalo di Itaca è il bel viaggio,
Senza di lei non l'avresti intrapreso,
Di più non ha da darti.

E se ti appare povera all'arrivo,
Non t'ha ingannato.
Carico di saggezza e di esperienza,
Avrai capito un'Itaca cos'è.

Kostantinos Kavafis