Mathematical Modeling

Vincent Bansaye and Thibaut Mastrolia

MAA 107 – Bachelor program, first year
Ecole Polytechnique – Spring 2021

General Introduction

The objective of this course is to introduce mathematical modeling to study systems evolving over time in a deterministic or random manner. The point of view adopted is to choose a few examples from different disciplinary fields (biology or ecology, physics or chemistry, economics or finance). We wish to introduce the reader to this modeling approach, which allows us to summarise and simplify a concrete problem in order to understand and analyse it mathematically. The objective is also to introduce the reader to new mathematical notions which will prove to be the fundamental tools for the study of the problem.

Mathematics

The course is divided into two parts. The first part focuses on deterministic dynamics. This means that knowing the state of a system at a given time, its state in the future will be predictable : it is uniquely determined by a mathematical equation. The question that will occupy us becomes: how to access the solution and the future state in question? We will either try to calculate this solution explicitly (obtaining a formula to give the state of the system at any given time); or describing it qualitatively (trajectory patterns) or quantitatively (numerical estimates). In the first course, we will consider time in a discrete way, that is, we will describe the state of the system step by step. The notion of recurrent sequence then appears to describe the evolution of this system. After recalling the formulas for calculating an arithmetic or geometric sequence, we will calculate the value of new sequences, using a simple transformation that allows us to return to a known sequence. In the second course, we will consider the dynamics $t \to x(t) \in \mathbb{R}$ in continuous time, which will be solutions of a differential equation of the form $x'(t) = f(x(t))$. We will see through an example both how an integral calculation method allows to obtain the solution explicitly but also how reasoning on the sign of the derivative allows to describe qualitatively its evolution in long time. In the upper dimensions, computable systems are essentially reduced to the special case of linear dynamical systems: $x'(t) = Ax(t)$, with x vector and A matrix. In the third course we will focus on a two-dimensional system $t \to (x(t), y(t)) \in \mathbb{R}$ in a non-linear case. We will use a preserved quantity over time to describe qualitatively the behaviour of the system. In the fourth course, we will no longer have access to a preserved quantity. We will have to introduce another mathematical technique. We will use an approximation method based on Taylor expansion to get back to a linear framework, and thus calculable via linear algebra. We will then be able to provide relevant information on the initial problem.

The second part of the course will focus on stochastic dynamics in discrete time of the form $X_1, X_2, X_3, ..., X_n, ...$ where all the X_n are random variables giving the state of the system at step n. The law of X_n will be defined as a function of the value of X_{n-1}: $\mathbb{P}(X_n = k|X_{n-1} = i)$ will be given explicitly by the modeling of the problem. In a first step, we will try to access the means and variances of these random variables to have a first idea and estimation on the behaviour of the system. If these quantities will give a first interesting idea of the system's trend, many behaviours and phenomena are in fact possible for the same evolution in mean and variance. The courses will then show different behaviours in long time, successively introducing different notions of convergence: almost sure convergence (course 6), convergence in law (course 7), convergence on average along the trajectory (course 8). The aim is not to study these notions of convergence in a general way, or to link them together, but rather to see how they come about in concrete problems.

Modeling

The word "modeling" refers to different notions depending on the field and the context. In this course we will give an overview of mathematical modeling of systems or processes evolving over time, in different disciplinary fields. Modeling here will initially mean simplifying, keeping what seems to be the essence of the problem of interest or a first relevant framework of study. We will be interested in biological or economic or physical systems, which may be very complicated. They may include many factors or an important typology of agents. We will not be able to integrate all these elements, nor even for many to describe precisely their effect. The first observation is therefore an admission of modesty: to begin by greatly simplifying the problem, to be able to reduce it to the study of a relevant mathematical object in order to start its study. Understanding the simplified problem in this way, criticizing the limits of this simplified approach, considering complexifications that seem relevant will be part of the course. We will therefore sometimes be in the discussion. The simplifications, the hypotheses made, the choice of certain parameters or effects that we want to integrate may seem difficult. The question of the robustness of the results, *i.e.* their dependency on modelling choices, will be legitimate. In any case, we will always try to explain and motivate this choice, to open the dialogue for alternatives or criticisms. The time will be also limited and the objective will be to learn techniques and discover mathematical notions...

This modeling step will take us from a concrete problem of interest to a mathematical object that can be studied to answer part of the problem, either through theory or through the use of machines (numerical approach and computer simulation); either qualitatively (what happens roughly or finely? who wins? who survives ? when does it increase?) or quantitative (can the number of remaining prey be counted? what is the optimal gain value? how fast does the average number of particles in a site decrease?). This first step will already allow to understand, with a simplified and chosen angle of attack, the problem. It will sometimes make it possible to identify effects that may be unexpected or fundamental difficulties. This step often leads to many others ...

In particular, we will address the following questions:

Session 1: How to optimize the reimbursement of a loan? a first approach to corporate finance and discrete systems.

Session 2: At what population size does a population competing for these resources saturate? Equilibrium, stability, and hysteresis.

Session 3: Why can a decrease in fishing have a negative effect on the size of the fished population? Modeling prey predators dynamics.

Session 4: How can a bioreactor be stable and what is its production speed? Interacting systems and linearization around equilibrium.

Session 5: What is the fair price of a financial product? A tale of European options.

Session 6 : Is a polymorphism doomed to disappear and after how long? Genetic drift in random transmission.

Session 7 : Is the evolution of the gas configuration in a room a reversible dynamic?

Contents

Part I

Dynamical deterministic systems

Reminders on sequences

Preliminary definitions

Definition. *Let $\mathbb{N}$ be the set of non-negative integers.*

- *A (real)* ***sequence*** *is a map from $\mathbb{N}$ into $\mathbb{R}$. We denote $\mathbb{R}^{\mathbb{N}}$ the space of real sequences. Let $u \in \mathbb{R}^{\mathbb{N}}$ then for any $n \in \mathbb{N}$, $u_n := u(n) \in \mathbb{R}$ is the nth term of the sequence u.*
- *A sequence $u = (u_n)_{n\in\mathbb{N}}$ is defined by induction if there exists a map f from $\mathbb{R}$ into $\mathbb{R}$ such that $u_{n+1} = f(u_n)$ for any $n \in \mathbb{N}$.*
- *We say that a sequence u converges to a finite real ℓ if $\lim_{n\to+\infty} u_n = \ell$, otherwise the sequence diverges (to $+\infty, -\infty$ or has no limit).*

Examples. *(a) The inductive sequence $u_{n+1} = u_n + d$ with $u_0 \in \mathbb{R} \setminus \{0\}$ where $d \in \mathbb{R} \setminus \{0\}$ diverges to $+\infty$ (resp. $-\infty$) if $d > 0$ (resp. $d < 0$).*

(b) The inductive sequence $u_{n+1} = \frac{1}{2}u_n$ with $u_0 > 0$ converges to 0.

(c) The sequence defined by general terms $u_n = (-1)^n$ diverges[1] since it has no limit.

We have in particular the fundamental properties

Properties. *Let $(u_n)_{n\in\mathbb{N}}$ be in $\mathbb{R}^{\mathbb{N}}$.*

(i) Any converging sequence is bounded.

(ii) If $(u_n)_{n\in\mathbb{N}}$ is nondecreasing (resp. nonincreasing), i.e. $u_{n+1} - u_n \geq 0$ (resp. $u_{n+1} - u_n \leq 0$) for any $n \in \mathbb{N}$, and bounded from above (resp. from below), then it converges.

(iii) Any increasing (resp. decreasing) sequence unbounded from above (resp. from below) diverges to $+\infty$ (resp. to $-\infty$).

We finally introduce the sommation of terms of a sequence.

Definition. *Let $(u_n)_{n\in\mathbb{N}}$ be in $\mathbb{R}^{\mathbb{N}}$. A* ***serie*** *with general term u_n is a sequence $(S_n)_{n\in\mathbb{N}}$ defined by*

$$S_n = \sum_{k=1}^{n} u_k.$$

We say that a serie converges is there exists a finite real L such that $\lim_{n\to+\infty} S_n = L$. Otherwise, the serie diverges.

[1] In fact, for all odd number the sequence takes the value -1 and for even numbers it takes the value 1. Although it does not converges, we say that it admits two clusters points 1 and -1.

In particular we have the following fundamental property

Proposition I.1. *Let $(S_n)_{n\in\mathbb{N}}$ be a serie associated with a sequence of general term u_n. If the serie $(S_n)_{n\in\mathbb{N}}$ converges then u_n converges to 0.*

Remark. *The previous result provides an easy way to prove that a serie diverges. For instance it says that the serie given by $S_n = \sum_{k=0}^n (-1)^k$ diverges when n goes to $+\infty$.*

Arithmetic progressions

We first study the properties associated with the sequence introduced in Example (a).

Definition. *An arithmetic progression $(u_n)_{n\geq 0}$ is a sequence defined by induction such that $u_{n+1} = u_n + d$, $n \geq 0$ with $u_0 \in \mathbb{R}$, where $d \in \mathbb{R}$ is a fixed constant called the common difference.*

The difference of two consecutive terms of an arithmetic sequence is constant and equal to d.

Proposition I.2. *Let $(u_n)_{n\geq 0}$ be an arithmetic sequence with common difference $d \in \mathbb{R}$. Then*

(i) $u_n = u_0 + nd$, $n \geq 0$,

(ii) for any $0 \leq m \leq n$, $\sum_{k=m}^n u_k = (n-m+1)u_0 + d\frac{(m+n)(n-m+1)}{2}$.

Proof. The proof of (i) is direct. We turn to (ii). Let $S_m^n = \sum_{k=m}^n u_k$ then,

$$S_m^n = \sum_{k=m}^n (u_0 + kd) = (n-m+1)u_0 + d\sum_{k=m}^n k.$$

By noting[2] that $\sum_{k=m}^n k = \frac{(m+n)(n-m+1)}{2}$ we get the result. □

Geometric progressions

We now study the properties associated with the sequence introduced in Example (b).

Definition. *A geometric progression $(u_n)_{n\geq 0}$ is a sequence defined by induction such that $u_{n+1} = qu_n$, $n \geq 0$ with $u_0 \in \mathbb{R}$, where $q \in \mathbb{R}$ is a fixed constant called the common ratio.*

The ratio of two consecutive terms of a geometric sequence is constant and equal to q.

Proposition I.3. *Let $(u_n)_{n\geq 0}$ be a geometric sequence with common ratio $q \in \mathbb{R}$. Then*

(i) $u_n = q^n u_0$, $n \geq 0$. *Hence, for $q = 1$ the sequence is constant and equals to u_0, and for $|q| < 1$, the sequence $(u_n)_n$ converges to 0 when n goes to $+\infty$. Otherwise it diverges.*

(ii) for any $0 \leq m \leq n$

$$\sum_{k=m}^n u_k = \begin{cases} u_0\frac{q^m - q^{n+1}}{1-q}, \ q \neq 1, \\ u_0(n-m+1), \ q = 1. \end{cases}$$

Proof. The proof of (i) is direct. The proof of (ii) is based on the same trick[3] than the proof of (ii) in Proposition I.2 by noting that $\sum_{k=m}^n u_k - q\sum_{k=m}^n u_k = u^m - u^{n+1}$. □

[2] Obviously it is useless to learn this formula! The trick consists in noting that $\sum_{k=m}^n k = m + (m+1) + \cdots + (n-1) + n$ and $\sum_{k=m}^n k = n + (n-1) + \cdots + (m+1) + m$. By adding the two equalities we see that $m+n$ is added $m-n+1$ times. We directly get $\sum_{k=m}^n k = \frac{(m+n)(n-m+1)}{2}$.

[3] Again, do not learn the formula! The Trick is very tractable to find the result quickly.

Affine progression

In this section, we mix the arithmetic and geometric properties seen before.

Definition. *An affine progression $(u_n)_{n\geq 0}$ with parameters $(d,q)\in\mathbb{R}^2$ is a sequence defined by induction such that*

$$u_{n+1} = qu_n + d,\ n\geq 0 \text{ with } u_0\in\mathbb{R}.$$

Of course, this notion generalizes the definitions of arithmetic and geometric progressions.

Proposition I.4. *Let $(u_n)_{n\geq 0}$ be an affine sequence with parameters $d,q\in\mathbb{R}$. Then*

$$u_n = \begin{cases} u_0 + nd, & \text{if } q=1 \\ \frac{d}{1-q} + (u_0 - \frac{d}{1-q})q^n, & \text{otherwise.} \end{cases}$$

Proof. We clearly have an arithmetic progression when $q=1$ with common difference d and from Proposition I.2 we deduce the result for $q=1$.
We now focus on the case $q\neq 1$.
Step 1[4]*: Let us find an equilibrium!* First, we aim at finding a fixed point for the sequence. More exactly, we are looking for a real numbers ℓ such that

$$\ell = q\ell + d, \text{ and we get } \ell = \frac{d}{1-q}.$$

Step 2: Yet another geometrical sequence. We now set $v_n = u_n - \frac{d}{1-q}$. Hence,

$$v_{n+1} = u_{n+1} - \frac{d}{1-q} = qu_n + d - \frac{d}{1-q} = q(v_n + \frac{d}{1-q}) + d - \frac{d}{1-q} = qv_n.$$

Hence, $(v_n)_n$ is a geometric progression with common ratio q. From Proposition I.3, $v_n = v_0 q^n$ with $v_0 = u_0 - \frac{d}{1-q}$. Thus, $u_n = \frac{d}{1-q} + (u_0 - \frac{d}{1-q})q^n$.

□

Remark. *The real $\ell = \frac{d}{1-q}$ is an equilibrium because if we start at $u_0 = \frac{d}{1-q}$ then $u_n = u_0 = \ell$. The system is thus stationary.*

[4] as a prequel of Lecture #2

Chapter 1

How to optimize the reimbursement of a loan? A first approach to discrete systems.

Reminders on arithmetic, geometrical and affine progressions. Applications to corporate finance: interest rate and payment schedule.

1.1 Introduction

Near 400 BC, Roman currency[1] was introduced in the Republic to organize more rigorously the cattle barter with bronze blocs (*aes rude*) improved later with a circle form by adding the face of Janus on one face. The birth of the monetary system in the Ancient Rome has induced new jobs related to money taking place at the *forum*. The *nummularius* tested the currency by ensuring its current value. He also helped at fluidifying the exchange of currencies or goods. The *argentarius* kept the money of his customers and ensured the most trading activity of this period: the *auctio* (auctions system). To echo an example of Jean Andreau[2], consider for instance an *equitis* wanting to exchange a small amount of money, he could go to the forum to deal with the *nummularius* or the *argentarius*. However, if the desired amount of money was too huge, he generally asked to his peers to borrow it by paying interests.
We thus see that the existence of financial markets has been motivated to make easier the exchange of cash flows between a financial agent lending money to an other one to finance a project, by paying interests for the service provided. We distinguish two types of financial agents: the borrowers and the lenders. For instance, a state borrows money to a bank to finance public services, people borrow money to a bank for mortgages. The aim of this first chapter is to understand what is the *real* value of a loan. In fact, the borrower not only reimburses his loan, he also pays interests to the lender for the service and for the management of this loan. We will see in the application to payment schedules that different types of reimbursements induce different credit costs. The main mathematical tool used to investigate this question is the theory of sequences.

[1] The example of roman currency is particularly interesting since the current monetary system has structurally inherited of the Ancient Rome system.

[2] *La vie financière dans le monde romain, les métiers de manieurs d'argent,* Bibliothèque des Écoles françaises d'Athènes et de Rome, 1987.

1.2 Interest rates

To come back to the example in the introduction, assume for instance that an *equitis* goes to the Roman forum and wants to invest $u_0 = 100$ *Sestertii*[3]. We develop two ways for the *equitis* to take a benefit of his initial investment.

Simple interest rates. The *equitis* earns $d > 0$ Sestertius each year. If u_n denotes the value of his capital at the year n, we have $u_{n+1} = u_n + d$ which is the example (a) of an arithmetic progression. In this example, if d is indexed as a percentage r of the initial capital u_0 invested by the *equitis*, we get $u_{n+1} = u_n + ru_0$.
An *annual simple interest rate* is a fixed number r such that an investment of u_0 euros in a bank account at time 0 with such interest rate yields ru_0 each year in addition to the investment u_0. Let u_n be the value of the bank account at time n, for any year n, we have

$$u_{n+1} = u_n + ru_0, \; n \geq 0.$$

$(u_n)_n$ is an arithmetic progression with common difference ru_0. Hence, after n years, the value of the bank account is $u_n = u_0 + nru_0 = (1 + nr)u_0$. The investor thus earns nru_0 euros.

Compound interest rates. Assume now that the income of the *equitis* at time $n + 1$ is indexed on the value of his investment at time n, so that the value of the bank account at time $n + 1$ and denoted by u_{n+1} pays a fixed rate r of the previous value of the investment u_n at time n. We get $u_{n+1} = u_n + ru_n = (1 + r)u_n$.
An *annual compound interest rate* is a fixed number r such that an investment of u_n euros in a bank account at time n lead to a wealth of $u_{n+1} = (1 + r)u_n$ the year $n + 1$. Hence, the value of the bank account is a geometric progression with common ratio $(1 + r)$. If we denote by u_0 the initial value of the bank account, the value of the bank account each year is $u_n = u_0(1 + r)^n$. The investor thus earns $u_0((1 + r)^n - 1)$ euros on the duration n years.

What is the best interest rate for a bank account? In practice, the rate r is very small (it is given in %) and positive. By using a limit order development, we get

$$((1 + r)^n - 1) \approx nr + \frac{n(n-1)}{2} r^2 + \dots$$

Thus, $((1 + r)^n - 1) > nr$. In other words, a financial agent benefits more of an investment in a bank account with compound interest rates compared to simple interest rates.

Exercise: compound interest rates with standing order. Assume now that an agent invests u_0 initially in a bank account and adds d euros each years. We assume that the bank account at time n with value u_n pays compound interests rates r, indexed on the previous value of the bank account u_{n-1} **before the standing order** d. Give the value of u_n with respect to u_0, r, d.

1.3 Application to corporate finance

In this section, we study a financial problem on the mechanism of specific financial *riskless* products called *loans*. The notion of risk in finance will be see in the second part of this course.

[3] The Sestertius was a Roman coin introduced at the end of the 3rd century BC in Roman Republic.

1.3.1 General definitions

Definition 1.3.1. *A **financial product** is a contract in which the involved parties exchange cash flows. It is said riskless if these flows are deterministic (i.e. a future cash flow is known at any moment). Otherwise, we use the terminology of risky financial product.*

Examples of financial products

- A *financial asset*. It is a (risky) contract ensuring to his owner incomes. For instance, it could be a part of the total value of a firm providing to his owner dividends or the right to give an advice in the management of the firm.
- A *bond*. A bond is a debt security so that the issuer owes money to the holder by paying coupons at fixed time during a period fixed in the contract with fixed interest rates.
- An *option*. It is a contract ensuring to the buyer (and so the owner) the **right** (and not the obligation) to buy/sell an underlying asset or other derivative products at a specified price at a specified time.

In this chapter we focus on some specific financial instruments called **loans**.

Definition 1.3.2. *A **loan** is a debt dues from one financial agent to another one at an interest rate fixed in the contract.*

A loan is characterized by four parameters: the amount of money borrowed, the interest rate, the duration of the loan, the debt repayment arrangement.
We assume that the borrower pays a cash flow to the lender every year. The total cash flow provided each year is called **an annuity.**

Remark 1.3.3. *An annuity is thus composed of two main parts: the part of the amount borrowed reimbursed at the considered year and the interest paid on the considered period.*

We distinguish three types of loans.

- **A bullet loan.** That is a loan such that the reimbursement of the initial cash borrowed occurs one times, at the end of the last period.
- **Straight line amortization.** The reimbursement of the cash initially borrowed is constant each years.
- **Mortgage-Style amortization.** The annuity paid is the same each year.

We insist on the fact the **Straight line amortization** is strictly different from the **Mortgage-Style amortization** in view of Remark 1.3.3. From now we denote by

- n the duration of the loan (number of years), r the annual (compound) interest rate,
- V_0 the amount borrowed intitally, V_k the outstanding capital at the end of the kth year,
- I_k the interest paid at the end of the kth year, indexed on V_k,
- D_k the part of capital reimbursed at the kth year,
- $A_k := I_k + D_k$ the annuity paid at the end of year k,

As a consequence of these definitions, we have for any $1 \leq k \leq n$

$$V_n = 0,\ V_k = V_{k-1} - D_k,\ \sum_{k=1}^{n} D_k = V_0,\ I_k = rV_{k-1}.$$

The credit cost is defined by $\sum_k I_k$. The aim of the challenge will be to compare the credit cost induced by the three reimbursement methods.

1.3.2 Paiement schedule for a bullet loan and extensions

We now investigate more deeply the structure of a bullet loan. By definition $D_k = 0$ for any $1 \leq k \leq n-1$ and $D_n = V_0$. Consequently, $V_k = V_0,\ 1 \leq k \leq n-1$ and $V_n = 0$ and $A_k = rV_0$ for any $1 \leq k \leq n-1$ and $A_n = V_0(1+r)$. The paiement schedule is thus given by

Period k	V_k	I_k	D_k	A_k
1	V_0	rV_0	0	rV_0
...	V_0	rV_0	0	rV_0
$n-1$	V_0	rV_0	0	rV_0
n	0	rV_0	V_0	$V_0(1+r)$
Total		nrV_0	V_0	$V_0(1+nr)$

Credit cost: nrV_0

In the **challenge #1**, you will have to implement (numerically or theoretically) the two others reimbursement methods and to compare the associated credit costs.

1.4 Pathological sequences and motivation to turn to continuous models

Basically, we can extend the existence of *nice* results for linear sequences of any order p, that is an inductive sequence defined for any $n \geq 0$ by

$$u_{n+p} = a_n u_n + \cdots + a_{n+p-1} u_{n+p-1},$$

where $a_n, \ldots, a_{n+p-1}$ are fixed constant and $u_0, \ldots, u_{p-1}$ are fixed initial terms. This induces to introduce the characteristic polynomial associated with it, find the zero and write a linear system. The linearity is the key to solve such system. As soon as we go outside the linear properties, it becomes more and more difficult.

Let us consider the famous logistic model introduced by R. May in population dynamics and defined by

$$u_{n+1} = \alpha u_n(1-u_n),\ \alpha \geq 0,\ 0 \leq u_0 \leq 1.$$

Here u_n can be seen as the ratio of existing population to the maximum possible population (that is the reason why we take $u_0 \in [0,1]$). Then,

- if $\alpha = 0$, the sequence $(u_n)_n$ is constant for $n \geq 1$,
- the system has two stationary points (that is real ℓ such that $u_0 = \ell$ leads to $u_n = u_0$) $u_0 = 0$ and $u_0 = \frac{\alpha-1}{\alpha}$.
- if $0 < \alpha \leq 1$ then u_n converges to 0 (exercise!) and if $1 < \alpha \leq 3$ then u_n converges to $\frac{\alpha-1}{\alpha}$ (more complicated).

- Otherwise, it is much more complicated... see the second lecture for preliminary results.

This motivates us to go to continuous time models in the following lectures.

1.5 Challenge #1: Payment schedules

In all these exercices we consider compound interest rates.

1.5.1 Preliminary questions

Let r be an annual interest rate for both a loan and an investment. What is the value today of 1 euros in 1 year? More generally, assume that you subscribe today to a contract ensuring at the kth year a cash flow of F_k for $1 \leq k \leq n$ with n the total duration of the contract, what is intuitively the actual value of this contract?

In view of the question above, provide intuitively a relation between V_k and the annuities A_i for $k+1 \leq i \leq n$. Prove rigorously this relation by using a backward induction.
Hint: We recall that $V_n = 0$.

1.5.2 Straight-line amortization–Mortgage-Style Amortization

Let V_0 the initial amount of euros borrowed with annual rate r during n year. What do you notice concerning the sequence of annuities $(A_k)_k$?

- Make the payment schedule for a Straight-line amortization. What do you notice concerning the sequence of anuities $(A_k)_k$?

- Make the payment schedule for Mortgage-Style Amortization. What do you notice concerning the sequence of reimbursements $(D_k)_k$?

- Discuss about the comparison of the credit costs between the three type of loans (bullet loan, straight-line amortization loan and Mortgage-Style Amortization).

1.5.3 Models with paiement delay.

We borrow $S = 1000$€ with annual interest rate $r = 5\%$ during 4 years.

1. The reimbursement is delayed of 2 years then the annuities are constant. Make the paiement schedule and compute the credit cost.

2. The annuities are delayed of 2 years, it means that the borrower pays no interest and reimburses nothing during the first 2 years. Then the annuities are constant. Make the paiement schedule and compute the credit cost.

3. Compare with a bullet loan with same characteristics. Interpretations?

1.5.4 Mortgage style amortization : applications

1. We borrow $S = 2000€$ with interest rate $r = 5\%$ during 4 year. We consider a loan of type "Mortgage-Style Amortization". Build the payment schedule and compute the credit cost.

2. We consider a loan of type "Mortgage-Style Amortization" with the following payment schedule rounded to the unit

Year	V_k	I_k	D_k	A_k
1		500	2320	
2				
3	2686			
4			2686	

   Complete the payment schedule.

   Hint: for the computation of the interest rate, you round decimals to the nearest hundredth, i.e. $r = x\%$ with x a positive integer.

1.6 Appendix: To go further...

Interest rate swaps: a way to manage credit risks

In this section, we study a particular derivative product called **interest rate swaps** (IRS for short). This type of contract allows his owner to modify the interest associated with a loan in order to improve the credit cost for instance before the end of the credit duration. In fact, one could distinguish different types of interest rates. Previously, we have considered fixed interest rated. We also may consider **floating interest rates** which move up and down according to a financial market. This type of rates is particularly used for mortgages.

The lasts days of IBOR rates and new interest rates.

We first give some particular interests rates in the class of IBOR (Interbank Offered Rate) models. These kind of rates are computed as an average value of the interest rates on a daily basis of leading banks answering to the question "which interest rate do you think other banks would ask to lend you money?".

- The LIBOR (London Inter-bank Offered Rate). It is a benchmark interest rate used on the market which offers to a bank to lend funds to one another in the international interbank market for short-term loans. It is one of the most floating interest rate used on the money market in the class of IBOR rates.

- Euribor: it is one of the most interest rate used in the Eurozone and published by the European Banking Federation. It acts until a fixed maturity in the future (for instance 3 months).

Another type of interest rate very common on the European market is the Eonia. It is calibrated as an average of all overnight unsecured lending transactions in the interbank market in Europe, and computed day by day by the European Central Bank, unlike IBOR models which are based on leading banks statements.

The great recession of 2007-2008 has led to an explosion of the standard deviation of the Libor rate. This rate being indexed on an average it has thus somehow lost its well-posedness. Beyond this technical issue, the fact that any IBOR model is based on bank statements has dramatically[4] led to interest conflicts (for instance traders were directly informed by leading banks before the setting of the interest rate, allowing advantages in the previsions of daily interest rates). In June 2012, criminal settlements by Barclays Bank revealed significant fraud and collusion by member banks connected to the rate submissions of the LIBOR.

The IBOR rates will disappear by 2021 and researches are still in progress to find successors of them. UK is developing[5] a new index named SONIA, US planed to launched[6] SOFTR and EU will introduced[7] ESTER in October 2019.

Interest Rate Swap

Definition 1.6.1. *An Interest Rate Swap (IRS for short) is a contract allowing to exchange at a fixed time different type of interest rates indexed on a fixed wealth.*

The most common IRS is a fixed for floating swap. An example of well-known IRS is the OIS[8] (Overnight Indexed Swap) in which the floating payment is indexed on daily compound interest investments.

A swap allows its participants to exchange the interests associated to a loan (and only the interests are exchanged, the initial wealth borrowed is never exchanged) with respect to their preferences on the types of interest rates. A swap is somehow the meeting of two financial agents having the same funding requirements with different preferences concerning the interests associated to this loan (one prefers fixed rate, the other prefers floating rate). This contract is interesting if it optimizes the credit costs of the two involved agents compared to the modalities of the same loan if it was made on the market directly. A bank can monitor the swap by taking banking margins specified in the contract. Let us give an example.
Two firms F_1 and F_2 want to borrow $1M€$ reimbursed totally in 10 years (bullet loan type). The credit rating of F_1 (resp. F_2) allows it to borrow this amount with a fixed rate of 6% (resp. 5%) or floating rate Euribor+0, 3% (resp. Euribor+1%).

- Firm F_1 (resp. F_2) prefers to borrow money at a fixed rate (resp. floating rate).
- (bank margin) The bank managing the operation quotes the swap 10 years on the Euribor: 5%-5,5%: it means that a firm wanting to pay the Euribor it receives 5% of interests and if the firm wants to receives the Euribor, it pays 5, 5% of interests to the bank.

Without swap: F_1 borrows at rate 6% and F_2 borrows at rate Euribor+1%. The bank does not earn anything.

[4] See for instance "Understanding the Libor Scandal", Council of Foreign Relations, J. McBride, Oct. 2016.

[5] "From LIBOR to SONIA - ensuring a smooth transition in the futures markets", London Stock Exchanges Group website, by A. Ross, CurveGlobal, August 2018.

[6] "What is SOFR? The new U.S. Libor alternative", Karen Brettell, Reuters, April 2018

[7] https://www.ecb.europa.eu/paym/initiatives/interest_rate_benchmarks/euro_short-term_rate/html/index.en.html

[8] A brief LIBOR-OIS spread skyrocketed was observed during the crisis and today the LIBOR-OIS spread is considered a key measure of credit risk within the banking sector.

With a swap:

(a) Assume that F_1 decides to borrow money on the market with floating interests and F_2 with fixed interests. F_1 pays Euribor+$0,3\%$ and F_2 pays 5%.

(b) According to the preferences of the firms, F_2 wants to pay the Euribor to the bank, it thus pays Euribor to the bank **(b1)** and receives 5% according to the bank margin in the contract **(b2)**.

(c) Conversely, F_1 wants to receive from the bank the Euribor (to cancel its floating interest). The bank thus gives the Euribor to F_1 **(c1)** and according to the bank margin, F_1 has to pay $5,5\%$ **(c2)**.

We sum up the operations in Figure 1.6

Operations	Interests paid by F_1	Interests paid by the bank	Interests paid by F_2
(a)	Euribor+$0,3\%$	0	5%
(b1)	0	-Euribor	Euribor
(b2)	0	5%	-5%
(c1)	-Euribor	+Euribor	0
(c2)	$5,5\%$	-$5,5\%$	0
Credit cost	$5,8\%$	$-0,5\%$	Euribor

Figure 1.1: An optimal swap between F_1 and F_2

We see that the swap has improved the payment schedule of both firm F_1 and firm F_2 and the bank has earned $0,5\%$ of interests so that the preferences of each firms are satisfied.

Exercise. Let two firms A and B wanted to borrow $1M€$ during 10 years. According to their credit ratings, we know that A can borrow at fixed rate 11% or floating rate Euribor+2% on the market and B can borrow at fixed rate 9% or floating rate Euribor+$0,25\%$ on the market. If A prefers fixed interests and B prefers floating interests, give a swap optimizing their loans and precise the bank margin. And if B prefers fixed rate and A floating rate?

References

- *Options, Futures, and Other Derivatives, 9th Edition,* John Hull, Pearson, 2015.
- *Derivatives and risk management,* S. Janakiramanan, 2011.

Chapter 2

Let us find the equilibrium !

This lecture is an introduction to one dimensional dynamical systems, their equilibria and their stability. We will see that computations can be achieved in some cases, while a systemic study is possible using monotonicity argument and the first derivative of the vecteur field.
The challenge provides a concrete example in classical mechanics and a first example of a non-reversible phenomenon : the hysteresis.

2.1 Introduction: first order one dimensional ODEs for population dynamics

2.1.1 Malthusian growth

Malthus (1766-1834) was interested in the links between increase of production and increase of population. He proposed a model of growth of population and was worried about the speed of growth of that latter. To derive this model, we make the following assumption : there are no interactions between individuals and each one die or give birth independently, with the same rate, constant in time. The numbers of deaths and births during the time interval $[t, t+\Delta t]$ are therefore proportional to the number of individuals $y(t)$ and the length of the time interval Δt:

$$\Delta y(t) = y(t+\Delta t) - y(t) = \alpha y(t)\Delta t$$

for some $\alpha \in \mathbb{R}$. Letting Δt be small and recalling that $y'(t) = \lim_{\Delta \to 0}(y(t+\Delta) - y(t))/\Delta$, we get:

$$y(t)' = \alpha y(t).$$

We can solve explicitly this ODE and obtain for any $t \geq 0$,

$$y(t) = y_0 e^{\alpha t}.$$

2.1.2 Logistic effect: compute or not compute ?

Malthus pointed out that an exponential growth raises a problem if resources available does not increase as fast as the population size. Verhulst (1804-1849) proposed then a model where the growth of the population is impacted by its size. Following this idea, let us consider an additional mortality term due to the competition between individuals. We assume that each individual's death rate is proportional to the size of the population: we get

$$\Delta y(t) = y(t + \Delta t) - y(t) = y(t)(\alpha - \beta y(t))\Delta t$$

for some $\alpha > 0$ (births-deaths without competition) and $\beta > 0$ (competition coefficient). It is worth nothing that the sequence obtained by iteration of this identity yields the discrete logistic model. It is indeed both non explicitly solvable nor simpler to study even at a rough qualitative level, see Section 2.4 for details.

Let us turn to the continuous framework. Again, if Δt is small, we get

$$y(t)' = y(t)(\alpha - \beta y(t)) = F(y(t)), \tag{2.1}$$

with

$$F(y) = y(\alpha - \beta y).$$

Explicit computation. We can also solve this equation via the separation of variables method. Setting G to be a primitive of $1/F, i.e$:

$$G'(y) = 1/F(y)$$

That means $G(y) = \int_a^y \frac{1}{F(x)} dx$, where a should be well chosen to avoid cancellation of F, we have

$$G(y(t))' = G'(y(t))y(t)' = 1$$

so that by integration

$$G(y(t)) - G(y_0) = t$$

and if we can invert G, we obtain

$$y(t) = G^{-1}(t + G(y_0))$$

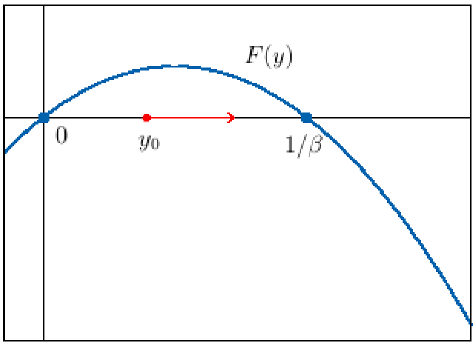
Exercice : find an explicit expression of G and $y(t)$ (in the case $\alpha = 1$ to simplify). Find

$$G(y) = \log\left(\frac{y(\beta - 1)}{\beta y - 1}\right), \quad G^{-1}(a) = \frac{1}{e^{-a}(1 - \beta) + \beta}, \quad y(t) = \frac{1}{e^{-t}(1/y_0 - \beta) + \beta}$$

for $t \geq 0$.

But in general you may not be able to find a primitive or invert it. Actually in dimension 2 or more, no explicit solution can be found, expect for very specific cases.

Qualitative study We keep $\alpha = 1$. We do not need to solve this equation (2.1) to describe its qualitative behavior when time runs. Let us determine the sign of $y \to F(y)$. It equals zero for $y = 0$ and $1 - \beta y = 0$, i.e. $y = 1/\beta$. F is positive on $(0, 1/\beta)$ and negative on $(1/\beta, \infty)$.



Thus, if a solution starts from 0, its derivative at time 0 is 0 and it stays at zero. Idem from $1/\beta$. The solutions $y(t) = 0$ and $y(t) = 1/\beta$ correspond to two particular solutions of (2.1) which are equilibriums (constant functions). The equilibrium 0 is rather expected : no people in the population, so no birth and no death. The equilibrium $1/\beta$ is more interesting : birth are compensated by deaths, the competition is just strong enough to maintain the size of the population at a constant size.

If a solution starts from any point in the interval $(0, 1/\beta)$, its derivative is positive and $t \to y(t)$ increases and goes to $1/\beta$.
Likewise, if a solution starts from any point in the interval $(1/\beta, \infty)$, it is decreasing and goes to $1/\beta$. As a conclusion,

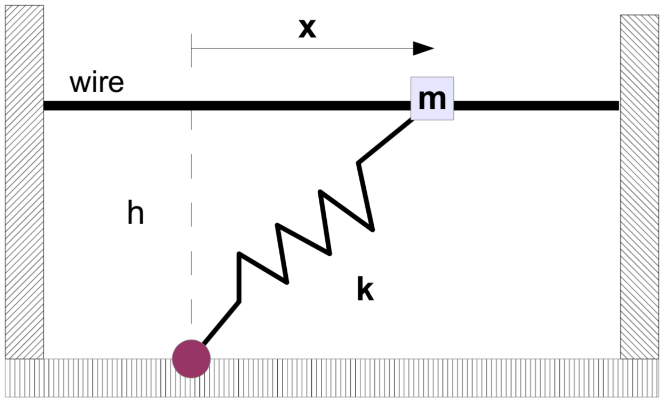
- 0 is an unstable equilibrium point : if we start close from it, the solution goes away from it.
- $1/\beta$ is a stable equilibrium : if we start close from it, the solution goes closer and closer.

Interpretation: population size tends to saturate at the equilibrium point $1/\beta$: if the size is larger, the competition is too strong and the population deacrases, and conversely. $1/\beta$ is seen as the carrying capacity of the system, i.e. the maximal size allowed when we let the process evolve.

Exercice: prove it! we admit that for each starting point there is a unique solution, and that it is differentiable. Uniqueness is clear in the example from the computations above, more generally it is a consequence of Cauchy Lipschitz theorem using the regularity of F.

2.2 The bead, the straight wire and the spring.

In this section, we focus on a more complex model coming from Newton's mechanics and discuss the equilibria.



A bead of mass m is constrained to slide along a straight horizontal wire. A spring of relaxed length L and spring constant k is attached to the mass and to a support point at distance h from the wire. We suppose that the motion of the bead is opposed by a viscous damping force proportional to the speed with a factor $b > 0$.
We denote by x the position of the bead. Newton's law says that the acceleration is proportional to the sum of the forces, i.e. writing $\phi(t)$ the angle between the wire and the spring

$$\begin{aligned} x''(t) &= k\left(L - \sqrt{x(t)^2 + h^2}\right)\cos(\phi(t)) - bx'(t) \\ &= k\left(L - \sqrt{x(t)^2 + h^2}\right)\frac{x(t)}{\sqrt{x(t)^2 + h^2}} - bx'(t). \end{aligned}$$

Let us focus on the case $L > h$, so that the spring may be compressed. An equilibrium corresponds to a solution x of (2.2) such that $x(t) = x_0$ for any $t \geq 0$, i.e. $x(t)' = x(t)'' = 0$ and x_0 solves

$$0 = x_0(L - \sqrt{x_0^2 + h^2})$$

i.e. the potential equilibria are

$$x_0 = 0, \quad \text{or} \quad x_0 = \sqrt{L^2 - h^2}, \quad x_0 = -\sqrt{L^2 - h^2}$$

One can check that $x(t) = x_0$ are indeed solutions of (2.2) for these three values.

Let us now deal with stability. Start from $x(0) = x_0 \in (0, \sqrt{L^2 - h^2})$, with speed $x'(0) = 0$, then x_0'' is positive and for small positive times the speed $x(t)'$ becomes positive and $x(t)$ increases. Thus 0 is not a stable equilibrium. Moreover the speed remains positive as long as $x(t) < \sqrt{L^2 - h^2}$.
Start now from $x_0 \in (\sqrt{L^2 - h^2}, +\infty)$, the acceleration is negative at the initial time, the speed becomes negative and again the bead comes closer to $\sqrt{L^2 - h^2}$. $\sqrt{L^2 - h^2}$ is a stable equilibrium.

Exercice (difficult). Prove in these two last cases that the position $x(t)$ of the bead converges to $\sqrt{L^2 - h^2}$ as t goes to infinity.

2.3 Challenge #2: The bead, the inclined wire and the spring

A bead of mass m is constrained to slide along a straight horizontal wire. A spring of relaxed length L and spring constant k is attached to the mass and to at support point at distance h from the wire. We suppose that the motion of the bead is opposed by a viscous damping force proportional to the speed with a factor $b > 0$.
We denote by x the position of the bead.
We keep the length L of the wire larger than h ($L > h$) and let us slightly incline the wire

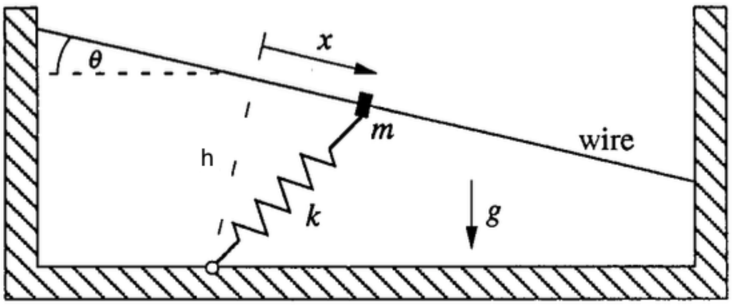


Figure 3.6.7

We write Newton's equation

$$x(t)'' = mg\sin(\theta) + k\left(L - \sqrt{x(t)^2 + h^2}\right)\frac{x(t)}{\sqrt{x(t)^2 + h^2}} - bx(t)'$$

Challenge number 2

- Find the number of equilibria as a function of the value of θ.
- Discuss their stability
- Start from $\theta = 0$ with the mass at the left equilibrium and fix h. Let θ increase very slowly (so the rleft point of the wire goes up slowly and the right point of the wire goes down slowly) and describe the equilibrium positions with time until θ is close to $\pi/2$.
- Make the converse operation and comment what happens !

We suggest you to make this analysis for the particular values of parameters given by

$$k = 10, \quad L = 10, \quad h = 5, \quad m = 5$$

and we recall that $g = 9.80665 (m.s^{-2})$.

2.4 Appendix: To go further ...

The discrete logistic model

The original paper of May, which introduced the model was

Simple mathematical models with very complicated dynamics

Indeed, as briefly mentioned in the introduced, the definition of the discrete model is more elementary since only a simple induction is involved, while it is much harder to study since some continuity property vanish and a much subtle dependence on the initial condition is observed...

A first example :

$$u_{n+1} = 4u_n(1-u_n), \qquad 0 \leq u_0 \leq 1.$$

a) Write $u_0 = \sin(2\pi\theta)^2$ ($\theta \in [0, 1[$) and show $u_n = \sin(2^n 2\pi\theta)^2$.
b) Prove then that if θ is rational, the sequence $(u_n)_n$ is periodic from some integer time n_0. For that purpose, you may decompose θ into base 2.

Chaotic behaviors. To know more, read and play with https://experiences.math.cnrs.fr/Iterations-de-l-application.html

Non-uniqueness and Cauchy Lipschitz theorem

First examples.
1) Consider the equation

$$y'(t) = \sqrt{y(t)}, \quad y(0) = 0.$$

a) Find two different solutions y_0 and y_1.
b) Show that there exists an infinite number of solutions to this equation.
2) Do the same for

$$y'(t) = y(t)^{1/3}, \quad y(0) = 0$$

Any difference?

General statement. Consider $y_0 \in \mathbb{R}$ and a function F. Existence and uniqueness of a solution on a maximal interval for $y' = F(y), y(0) = y_0$ are guaranteed by the Cauchy Lipschitz theorem. This latter requires a local Lipchitz property of the function F. The fact that F is C^1 provides a sufficient condition. More general conditions for uniqueness and existence can also be stated.

From discrete to continuous time : Euler approximation scheme.

The curious reader could see https://tutorial.math.lamar.edu/classes/de/eulersmethod.aspx. or https://www.khanacademy.org/math/ap-calculus-bc/bc-differential-equations-new/bc-7-5/v/eulers-method.

Chapter 3

Preys versus predators

We consider in this lecture a two dimensional dynamical system describing interactions between two populations. First, we explain the modeling leading to this dynamical system from the discrete setting. No explicit solution for this system can be obtained, neither for the discrete form nor for the continuous limit. But we can find a quantity which does vary along time, which will help us to determine precisely the behavior of the solutions.

Preys and predators systems have attracted lots of attention and arise in many different context of modeling, where one species attack another one and change/kill it. This includes of course prey predators interactions between plants or animals, but also epidemiology where infective people interact with sane and susceptible people, cell dynamics where virus attack cells, CLT attack cancer cells...

We focus in this lecture on a very simple model, relying on an interaction form which is probably the simplest one may think about at first glance : a bilinear form describing the number of interactions in terms of the sizes of each species. This famous model has been proposed by Volterra (1925) and after by Lotka (1926) to understand the evolution of population sizes of fishes after the first war.

Quite surprisingly maybe, this simple model is non explicitly solvable and its analysis is more involved than the one dimensional system we have seen in the previous lectures. It is not a matter of choice of model : interactions between two (or more species) lead to non linear dynamical systems for which explicit solutions are not known in general.

3.1 Introduction : prey predators in discrete time

We start with a simple modeling in discrete time. The time is indexed by $n \in \mathbb{N}$ and the successive integers correspond here to the successive generations. Let us write u_n the number of preys in generation $n = 1, 2, \ldots$ and v_n the number of predators in generation n.

Preys without predators. If the preys are alone, we assume that each prey has independently a mean number of offsprings equal to m and that each prey may die independently with probability p before the next generation. Therefore, the u_n preys of generation n give $u_n m$ offsprings in generation $n+1$ and $(1-p)u_n$ of the preys in generation n survive to the next generation. We obtain

$$u_{n+1} = (1-p)u_n + mu_n = u_n(1+\alpha), \tag{3.1}$$

where $\alpha = m - p$. We assume that if the preys are alone, then their population increases, i.e.

$$\alpha > 0.$$

Then $u_n = u_0(1+\alpha)^n$ and the size of the population of preys grows exponentially fast.

Predators without preys. Similarly, if the predators would live alone, we would get

$$v_{n+1} = (1-p')v_n + m'v_n = v_n(1-\beta) \tag{3.2}$$

where $\beta = p' - m'$. We assume that without preys, the population of predators can not survive and its size declines, i.e.

$$\beta > 0.$$

Then $v_n = v_0(1-\beta)^n$ and the size of the population of predators decreases exponentially fast.

Preys and predators. Now preys and predators live on the same place and are considered in the same model. The number of potential matchings between x preys and y predators is xy. We assume here that the number of matching that actually take place is a given deterministic fraction of this potential number, i.e. it is equal to γxy, with $\gamma \leq 1$, which is something we could (and will later) discuss for several reasons. Each of this effective matching leads to the death of the prey (we could add a probability that it indeed happens). Therefore the number of additional death of preys due to the presence of predators in generation n is $\gamma u_n v_n$ and taking into account this latter in (3.1), we obtain

$$u_{n+1} = u_n(1+\alpha) - \gamma u_n v_n$$

We now need to take into account the consequence of predations on preys. Here again, different modeling could be investigated and we make the following assumption : with probability p'', each predation yields a new predator. That amounts to make a transfer of mass of preys consumed into predators, with a coefficient $p'' \in [0,1]$, since the mass of prey is expected to be smaller than the mass of predator. Indeed, consumption of preys provide ressource to predators for survival of adults and breeding the new born. Such modeling also could be improved and criticized, but it give a first relevant and not too technical framework. Writing $\delta = \gamma p''$, the additional natality for the global population of predators is then $\delta u_n v_n$ in generation n and (3.2) becomes

$$v_{n+1} = v_n(1-\beta) + \delta u_n v_n.$$

To sum up, the dynamic is described by the following sequence :

$$\begin{cases} u_{n+1} = u_n(1+\alpha) - \gamma u_n v_n \\ \\ v_{n+1} = v_n(1-\beta) + \delta u_n v_n \end{cases} \tag{3.3}$$

3.2 Derivation of a continuous time model

We assume now that each step (generation) represents a short time $\Delta > 0$. As a consequence p and m (resp. p' and m') are small and we assume here that they are proportional to Δ. Thus we set

$$\alpha = \Delta a$$

Similarly

$$\beta = \Delta b, \quad \gamma = \Delta c, \quad \delta = \Delta d$$

For any $t = n\Delta$, we set

$$x(t) = u_n, \quad y(t) = v_n$$

the number of preys and predators at time t. We obtain

$$x(t+\Delta) = x(t)(1+\Delta a) - \Delta c x(t)y(t)$$

$$y(t+\Delta) = y(t)(1-\Delta b) + \Delta d x(t)y(t)$$

which becomes

$$\frac{x(t+\Delta) - x(t)}{\Delta} = ax(t) - cx(t)y(t)$$

$$\frac{y(t+\Delta) - y(t)}{\Delta} = -by(t) + dx(t)y(t)$$

Letting $\Delta \to 0$ and assuming that $t \to x(t)$ and $t \to y(t)$ are differentiable

$$\begin{cases} x'(t) = ax(t) - cx(t)y(t) \\ \\ y'(t) = -by(t) + dx(t)y(t) \end{cases} \tag{3.4}$$

which is the Lotka Volterra model.

3.3 Analysis

Let us start with the case where initially preys are alone :

$$(x(0), y(0)) = (x_0, y_0), \qquad x_0 > 0, \quad y_0 = 0.$$

We can then check that $(x(t), y(t))_{t \geq 0}$ defined by

$$x'(t) = ax(t), \qquad y(t) = 0$$

is a solution of the Lotka Volterra equation. It implies that

$$x(t) = x_0 e^{at}.$$

Actually, one can check that it is the unique solution of the Lotka Volterra with initial condition $(x_0, 0)$ equation using Cauchy Lipschitz theorem.
The result is expected from a biological point of view. No predators at the initial time implies no birth of predators and the population size remains equal to 0. The population of preys can then grow exponentially.

Similarly, let us consider the case when there is no prey (poor predators which have here nothing to eat !). Then $x_0 = 0$, ensures that $x(t) = 0$ for any $t \geq 0$ and

$$y'(t) = -by(t)$$

solves as $y(t) = y_0 e^{-bt}$. As expected, the population of predators decays, exponentially fast.

We deal now with the general case. Let us find the equilibriums, i.e. solve $x(t)' = y(t)' = 0$. We find two solutions

$$(x, y) = (0, 0), \qquad (x, y) = (b/d, a/c).$$

3.4 Challenge #3: preys versus predators

We study the Lotka Volterra model

$$\begin{cases} x'(t) = ax(t) - cx(t)y(t) \\ \\ y'(t) = -by(t) + dx(t)y(t) \end{cases} \tag{3.5}$$

with $a, b, c, d > 0$ and x_0 the initial quantity of preys, y_0 the initial quantity of predators. We have two equilibriums

$$(x_0, y_0) = (0, 0), \qquad (x_0, y_0) = (b/d, a/c).$$

- Determine the initial values (x_0, y_0) for which the population $x(t)$ decreases (resp increases) for small times. Do the same for the population of predators.
- Compute the variations of the function $t \to H(x(t), y(t))$ with $H(x, y) = dx - b\log(x) + cy - a\log(y)$
- Describe the behavior of the number of preys and predators when time goes. Draw their simultaneous evolution on a graph.
- Compute the mean number of preys and predators, i.e.

$$\frac{1}{T}\int_0^T x(t)dt, \qquad \frac{1}{T}\int_0^T y(t)dt$$

  where T has to be suitability chosen.
  Indication : one can consider $\int_0^T x'(t)/x(t)dt$.

- Determine the effect of fishing on these mean numbers, where the fishing is modeled by a parameter λ affecting identically preys and predators which is now described by the following ODE :

$$\begin{cases} x(t)' = (a-\lambda)x(t) - cx(t)y(t) \\ \\ y(t)' = -(b+\lambda)y(t) + dx(t)y(t) \end{cases} \tag{3.6}$$

3.5 Appendix: To go further...

Hamiltonian system

The study of the Lotka Volterra system above was carried out using the conservation of a functional over time

$$H(x(t), y(t)) = 0$$

for any $t \geq 0$. In physics, the conservation of a quantity such as energy is a fundamental property of many systems, which allows their mathematical study and analysis. More generally, a (p, q) system is Hamiltonian if there is a H function of two variables differentiable such as ‘

$$\begin{cases} p'(t) = \frac{\partial H}{\partial q}(p(t), q(t)) \\ q'(t) = -\frac{\partial H}{\partial p}(p(t), q(t)) \end{cases} \tag{3.7}$$

The function H then provides the functional conserved along the trajectory (first integral of the movement) :

$$H(p(t), q(t)) = H(p(0), q(0)) \text{pour tout } t \geq 0$$

Indeed

$$H(p(t), q(t))' = \frac{\partial H}{\partial p}(p(t), q(t))p'(t) + \frac{\partial H}{\partial q}(p(t), q(t))q'(t) = 0.$$

Exercise : consider again Lotka Volterra model (3.4) and

$$p(t) = \log(x(t)), \qquad q(t) = \log(y(t)).$$

Prove that it is an Hamiltonian system.
Is it also the case for the original one given by (3.4) ?
Propose a necessary and/or sufficient condition so that a two dimensional system $X' = F(X)$ is Hamiltonian.

Many systems with interactions

The interaction between two species is classical in modeling of epidemics, chemistry, viruses... As in this lecture, it is often quantified by a bilinear term $x.y$. Another famous model, which is also Hamiltonian, is the SIR model. It describes the propagation of an epidemics, when people who have been infected can not be infected again :

$$\begin{aligned} s'(t) &= -\beta s(t)i(t) \\ i(t)' &= \beta s(t)i(t) - \gamma i(t) \\ r(t)' &= \gamma i(t) \end{aligned}$$

where $s(t)$ is the number of susceptible individuals at time t, $i(t)$ the number of infected, $r(t)$ the number of recovered, β the contamination rate, γ the recovery rate. We assume also the the total population size is fixed and equals N: for any $t \geq 0$,

$$N = s(t) + i(t) + r(t).$$

Exercise. Find an equation to link the final size of the epidemics $r_\infty = \lim_{t\to\infty} r(t)$ to the

initial conditions $s(0), i(0)$ *and the size* N *of population.*
One can exploit the Hamiltonian structure or equivalently prove that $\gamma \log(s(t)) + \beta r(t)$ *is constant.*

Chapter 4

Speed of interactions and linearization

We consider more complex interactions in this lecture, focusing on motivations from chemical reactions. The interaction term won't be bilinear anylonger. As a consequence, we cannot exploit the same arguments (conservation of a some function of the coordinates) as in previous lecture to determine the evolution of the models. We introduce techniques of approximation, which is relevant in many other contexts. We first linearize the vector field around the equilibrium to compare the original problem with an explicitly solvable one.

4.1 Introduction : speed of interactions in enzyme kinetics

In biochemistry Michaelis-Menten kinetics is one of the best-known models of enzyme kinetics. In 1913, Michaelis and Menten proposed a mathematical model of reaction between enzyme and substrate. It involves an enzyme, E, binding to a substrate, S, to form a complex, ES, which in turn releases a product, P, regenerating the original enzyme. This may be represented schematically as

$$E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

where k_f (forward rate), k_r (reverse rate), and k_{cat} (catalytic rate) denote the rate constants, the double arrows between S (substrate) and ES (enzyme-substrate complex) represent the fact that enzyme-substrate binding is a reversible process, and the single forward arrow represents the formation of P (product).

We assume that the production (second reaction) is much slower than the binding (first reaction) :

$$\textit{Assumption :} \qquad k_{cat} = O(1) << k_f, k_r.$$

We write $p(t)$ the concentration of P and $s(t)$ the concentration of S at time t and $e(t)$ the total concentration of enzyme at time t (either alone or with substrate).
The speed of production of P is given by

$$p'(t) = k_{cat} e_0 \frac{s(t)}{c + s(t)}, \tag{4.1}$$

where $c = k_r/k_f$ and $e_0 = e(0)$. We observe that $v = k_{cat} e_0$ is the maximal speed of production.

Heuristic derivation of (4.1). Write $z(t)$ the concentration of ES and $x(t)$ the concentration of enzyme (alone) at time t. Then action mass law writes :

$$z'(t) = k_f x(t)s(t) - k_r z(t) - k_{cat} z(t).$$

Since $k_{cat} = O(1) << k_f, k_r$, the equilibrium of the first reaction (binding) is instantaneously reached and on a short time the systems reads

$$z'(t) \approx 0 \approx k_f x(t)s(t) - k_r z(t)$$

so that

$$k_f x(t)s(t) = k_r z(t).$$

We obtain that

$$c = \frac{k_r}{k_f} = \frac{x(t)s(t)}{z(t)}$$

is constant (and called Michaelis constant). Since $e(t) = x(t) + z(t)$ is constant in time, $e(t) = e_0$. Thus $x(t) = e_0 - z(t)$ and

$$c = \frac{(e_0 - z(t))s(t)}{z(t)}, \quad \text{i.e.} \quad z(t) = e_0 \frac{s(t)}{c + s(t)}.$$

We get

$$p'(t) = k_{cat} z(t) = k_{cat} e_0 \frac{s(t)}{c + s(t)}$$

and the expected result (4.1). □

4.2 Speed of production in a chemostat

4.2.1 Model

A chemostat (from chemical environment is static) is a bioreactor to which substrate (food) is continuously added to make mircroorganisms grow. The culture liquid containing left over nutriments, metabolic and products and microorganisms are continuously removed at the same rate to keep the culture volume constant.

Writing S the substrate, M the microorganisms, we first focus on the interactions between M and S. The substrate now allows the growth of the microorganism and we consider the following reactions

$$M + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} MS \xrightarrow{k} M + M$$

We assume that k_1 and k_2 are large compared to k. Following the first part, the speed of production of M is the speed of disparition of S and is given by

$$y'(t) = -s'(t) = ky(t)\frac{s(t)}{c + s(t)},$$

where $y(t)$ is the concentration of microorganism (either alone or in complex) at time t and $s(t)$ is the concentration of substrate.

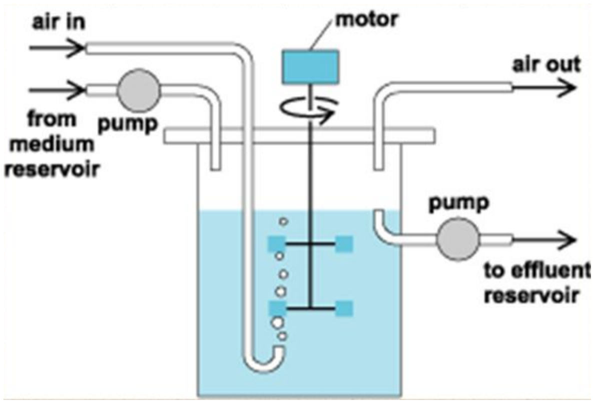


Figure 4.1: Chemostat extracted from https://encyclopedia2.thefreedictionary.com/chemostat

Let us now take into account that the substrate is added in a constant way and the volume is regulated and remains constant. Writing a the rate of arrival of substrate, the quantity of food brought increases linearly with speed a and

$$s'(t) = a, \quad \text{i.e} \quad s(t) = s_0 + at.$$

Moreover both food and microorganism are evacuated and each particle (M or S) leaves the cuve at a constant rate $d > 0$:

$$y'(t) = -dy(t), \qquad s'(t) = -ds(t)$$

Putting all that together, we finally obtain the system

$$\begin{cases} s'(t) = a - ds(t) - ky(t)\frac{s(t)}{c+s(t)} \\ \\ y'(t) = ky(t)\frac{s(t)}{c+s(t)} - dy(t) \end{cases} \tag{4.2}$$

with $c = k_2/k_1$.

4.2.2 Equilibria

First $(a/d, 0)$ is an equilibrium. It means that there is no microorganism and the quantity of substrate is a/d (arrival is compensated by evacuation). Is it stable? No it is not, since adding microorganisms, even very few, the population enjoys food (substrate) and grow ...

Is there an other equilibrium? Yes, we may have (s_*, y_*) given by

$$k\frac{s_*}{c+s_*} = d, \quad \text{i.e. } s_* = \frac{cd}{k-d}$$

and then

$$y_* = (a - ds_*)\frac{c+s_*}{ks_*}, \quad \text{i.e. } y_* = \frac{(k-d)a - d^2c}{d(k-d)}.$$

It is indeed well defined and non-negative if (and only if)

$$k > d \quad \text{and} \quad a(k-d) > cd^2 \qquad (H)$$

Exercice : prove that there is no other equilibrium. Explain what happens when (H) is not satisfied, i.e. what is the evolution of the system if we start with a non-zero quantity of microrganism and there is no other positive equilibrium than $(a/d, 0)$?

4.2.3 Stability of the non trivial equilibrium

One expect a stability of the production of microorganisms and thus of this equilibrium (s_*, y_*) (the other being unstable and without production of M). We thus assume (H) and consider an initial state (s_0, y_0) close to (s_*, y_*). Writing $S_t = s_t - s_*, Y_t = y_t - y_*$,

$$S'(t) = s'(t) = a - dS(t) - ds_* - k(Y(t) + y_*)\frac{s_* + S(t)}{c + s_* + S(t)}.$$

For small times, $(s(t), y(t))$ is close to (s_0, y_0) and $S(t)$ and $Y(t)$ are small, so that by Taylor expansion

$$\frac{s_* + S(t)}{c + s_* + S(t)} = \frac{s_*}{c + s_*} + \frac{c}{(c+s_*)^2}S(t) + o(S(t))$$

We obtain

$$S'(t) = -\left(d + k\frac{cy_*}{(c+s_*)^2}\right)S(t) - \frac{ks_*}{c+s_*}Y(t) + o(S(t)) + o(Y(t))$$

since

$$a - ds_* - ky_*\frac{1-s_*}{c+s_*} = 0.$$

Similarly

$$Y'(t) = y'(t) = k(Y(t) + y_*)\frac{S(t) + s_*}{c + S(t) + s_*} - d(Y(t) + y_*)$$

yields

$$Y'(t) = \left(k\frac{s_*}{c+s_*} - d\right)Y(t) + ky_*\frac{c}{(c+s_*)^2}S(t) + o(S(t)) + o(Y(t)) = ky_*\frac{c}{(c+s_*)^2}S(t) + o(S(t)) + o(Y(t)).$$

As a sum up, setting

$$X(t) = \begin{pmatrix} S(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} s(t) - s_* \\ y(t) - y_* \end{pmatrix},$$

we have obtained that

$$X'(t) = AX(t) + o(X(t)),$$

where

$$A = \begin{pmatrix} -\left(d + k\frac{cy^*}{(c+s_*)^2}\right) & -\frac{ks_*}{c+s_*} \\ ky_*\frac{c}{(c+s_*)^2} & 0 \end{pmatrix}.$$

As a consequence, the dynamics X is approximated by the linear differential system V

$$V'(t) = AV(t)$$

where $V(0) = X_0$, for small values of X_0 (i.e. initial conditions close to the equilibrium). A theorem (namely Hartman Grobman theorem, see Section "Going further") allows to give a rigorous sense to this approximation:

$$X \approx V \qquad \text{close to the equilibrium.}$$

We are now bound to describe V. Linear algebra can be invoked for V and reduction theory allows us to compute V. The result can be stated as follows. We write

$$\alpha = \left(d + k\frac{cy^*}{(c+s_*)^2}\right),\ \beta = \frac{ks_*}{c+s_*},\ \delta = ky_*\frac{c}{(c+s_*)^2}$$

and

$$\Delta = \alpha^2 - 4\beta\delta.$$

4.2.4 The case of a well : $\Delta > 0$

If $\Delta > 0$, we set

$$\lambda_+ = \frac{-\alpha + \sqrt{\Delta}}{2},\ \lambda_- = \frac{-\alpha - \sqrt{\Delta}}{2},$$

which are the two real solutions of $\lambda^2 + \alpha\lambda + \beta\delta = 0$.

Proposition 4.2.1. *We have*

$$A = PDP^{-1}$$

where

$$D = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}; \quad P = \begin{pmatrix} \lambda_+ & \lambda_- \\ \delta & \delta \end{pmatrix}.$$

Proof. The proof can be achieved by a simple computation : check that

$$AP = PD$$

using that both λ_+ and λ_- satisfy the equation $\lambda^2 + \alpha\lambda + \beta\delta = 0$. Multiply then this identity by P^{-1} on the right of both sides and conclude.

This proof requires the explicit expression of P and D; let us explain how one can find them (if they are not given like here) using linear algebra. First we search the eigenvectors, which means we solve the equation $AX = \lambda X$ with

$$A = \begin{pmatrix} -\alpha & -\beta \\ \delta & 0 \end{pmatrix}.$$

Writing $X = \binom{x}{y}$, this equation becomes

$$(\lambda + \alpha)x = -\beta y; \qquad \delta x = \lambda y.$$

We get, for $X \neq 0$,

$$\lambda^2 + \alpha\lambda + \beta\delta = 0.$$

Solving this second order polynomial equation using the discriminant Δ, we have two solutions for λ, namely λ_+ and $\lambda-$ given above. Then writting

$$X_1 = \begin{pmatrix} \lambda_+ \\ \delta \end{pmatrix}, \quad X_2 = \begin{pmatrix} \lambda_- \\ \delta \end{pmatrix},$$

we have

$$AX_1 = \lambda_+ X_1, \quad AX_2 = \lambda_- X_2.$$

The result follows by considering the matrix A in the new basis (X_1, X_2), which gives a diagonal form D. □

Setting

$$Z(t) = P^{-1}V(t),$$

we get

$$Z'(t) = P^{-1}V'(t) = P^{-1}AV(t) = P^{-1}APP^{-1}V(t) = DZ(t).$$

Since D is diagonale, Z can be computed explicitly and we get

Proposition 4.2.2. *For any $t \geq 0$,*

$$V(t) = P \begin{pmatrix} e^{\lambda_+ t} & 0 \\ 0 & e^{\lambda_- t} \end{pmatrix} P^{-1}X_0$$

and $V(t) \to 0$ as t tends to infinity.

Proof. Writing

$$Z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$

$Z'(t) = DZ(t)$ becomes

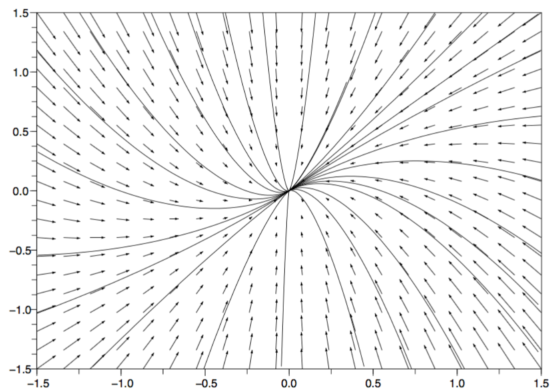
$$x'(t) = \lambda_+ x(t); \qquad y'(t) = \lambda_- y(t).$$

It can be solved :

$$x(t) = e^{\lambda_+} x_0; \quad y(t) = e^{\lambda_-} y_0$$

and obtain $V(t) = PZ(t)$ using $Z_0 = P^{-1}V_0 = P^{-1}X_0$. The fact that V tends to 0 is a consequence of the negativity of eigenvalues λ_+ and λ_-. The negativity of λ_+ is due to $0 \leq \Delta < \alpha^2$ and $\alpha > 0$. □

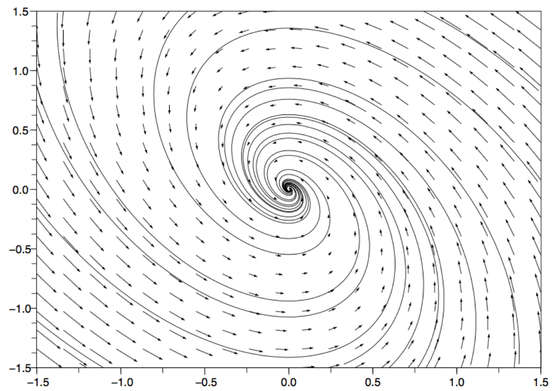
This ensures that the system is stable : solutions starting close to the equilibrium converge to equilibrium exponentially fast. It goes straight line :



,

4.2.5 The case of a spiral well : $\Delta < 0$

When $\Delta < 0$, the equation $\lambda^2 + \alpha\lambda + \beta\delta = 0$ has two complex solutions, which are conjugated. Then the trajectory turns (spiral) owing to the complex part:



This is left for the reader and follows the proof of the case above ($\Delta > 0$). Now D is of the form

$$D = \begin{pmatrix} \lambda_1 + i\lambda_2 & 0 \\ 0 & \lambda_1 - i\lambda_2 \end{pmatrix}.$$

The case $\Delta = 0$ is slightly different, it is left as an exercice.

4.3 Challenge #4: Rosenzweig MacArthur

Now we study a case where the substrate (or the prey) can duplicate with a saturation effect (logistic model). The microorganism (or the predator) needs substrate to duplicate. The evolution of the populations is given by

$$\begin{cases} x'(t) = x(t)(1 - x(t)/\gamma) - y(t)\frac{x(t)}{1+x(t)} \\ \\ y'(t) = \beta y(t)\left(\frac{x(t)}{1+x(t)} - \alpha\right) \end{cases} \tag{4.3}$$

with γ, α and β positive constants.

- Find the equilibria on $(\mathbb{R}^+)^2$ and interpret them in terms of survival/extinction/coexistence of the two terms.
- Determine the stability of these equilibriums.
- Explain what happens when all the equilibriums are non-stable.

Indicate : to solve these questions, you may use

http://experiences.math.cnrs.fr/Le-modele-de-Rosenzweig-MacArthur.html

for simulations and/or linearize the system for a theoretical approach.

4.4 Appendix: To go further...

The linearization technique consisted in linearizing the equation and compare the original system to a linear equation. First, we give the Taylor expansion, which give the general result which allows to linearize a function of several variables in a neighborhood of a point and extends the result with one variable you know. In the example above, this expansion has been achieved by a relevant use of one variable expansions but a direct derivation is possible using the expansion with two variables. Second, we provide the theoretical result which allows to reduce the problem to the its linearized version, at least locally : Hartman Grobman theorem.

Taylor expansion with several variables

Let U be an open set of $\mathbb{R}^n$ and $F : U \to \mathbb{R}$ be differentiable with respect to each variable at a point $a \in \mathbb{R}^n$. Then

$$F(a+h) = F(a) + \sum_{i=1}^{p} h_i \frac{\partial F}{\partial x_i}(a) + o(\| h \|_2),$$

where $\| h \|_2$ is the euclidian norm of $h \in \mathbb{R}^2$. It allows to approximate the function F by an affine function. The result can be generalized to higher orders.

Matrix exponential

When solving $x'(t) = ax(t)$ in $\mathbb{R}$, the solution is an exponential function : $x(t) = e^{at}x_0$. The exponential function appears in different way. For instance, it can be defined as :

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{k\to\infty} (1 + x/k)^k \quad (x \in \mathbb{R}).$$

An analogous result can be given in higher dimension, introducing the matrix exponential of a matrix A :

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \lim_{k\to\infty} (1 + A/k)^k.$$

The solution of $X'(t) = AX(t)$ where $X_0 \in \mathbb{R}^d$ and $A \in \mathcal{M}_d(\mathbb{R})$ is

$$X(t) = e^{At}X_0.$$

Hartman Grobman theorem

We want to reduce the analysis of a non-linear system to a linear one, which is computable as seen above. The idea is that if the linearized system is stable (resp. unstable), the same holds for the original one (which is non-linear). If the solution of the linearized system advances by turning in a spiral, the original one do the same. This comparison is valid when the linearized solution was attracted or repulsed, it is not valide otherwise. The linearized system indeed reduces the problem to the first order approximation, which is not enough to capture its behavior when this first order term is not significant for the stability.

Hartman Grobman theorem provides a general statement to deal with this issue and is stated as follows. Consider $X : \mathbb{R}^+ \to \mathbb{R}^d$ a solution of

$$X'(t) = F(X(t))$$

and an equilibrium x_0 :

$$F(x_0) = 0.$$

We assume that F is differentiable and write J_F the Jacobian matrix of F. We assume that all the real part of the eigenvalues of

$$A = J_F(x_0)$$

are non-zero. Then in a neighborhood of x_0, the flow of X is conjugated with the flow of Y in a neighborhood of 0 defined by

$$Y'(t) = AY(t), \qquad \text{i.e. } Y(t) = e^{At}X(0).$$

Let us explain what it means. The flow Φ of X is defined by

$$\Phi(x,t)' = F(\Phi(x,t)), \qquad \Phi(x,0) = x.$$

There exists a neighborhood $\mathcal{V}$ of x_0 and a diffeormorphism $B : \mathcal{V} \to B(\mathcal{V})$ such that for any $x \in \mathcal{V}$,

$$B(\Phi(x,t)) = e^{At}B(x).$$

Avoiding computations

Using tools of reduction theory, one could avoid computations in dimension two. One just need to consider the discriminant Δ of the characteristic polynomial and the trace Tr and determinant Det of the matrix A.

For instance, if Δ is positive, there are two real eigenvalues. If $Det > 0$, there have the same sign and if $Tr > 0$, there are both positive, which means that we have a wells. One could consider the other cases similarly.

Let us come back to the example

$$A = \begin{pmatrix} -\alpha & -\beta \\ \delta & 0 \end{pmatrix}.$$

The discriminant can be positive or negative, which means two real or complex conjugated eigenvalues. Adding that $Det = \beta\delta > 0$ and $Tr < 0$, the real values of the eigenvalues are both negative, so solutions go to zero.

About chemostat

The chemostat is a laboratory device that allows the culture and study of microorganism or plant cell species. The first introduction of the chemostat dates back to 1950 by their inventors Novick and Szilard and Monod. We find in the literature several works relating to the chemostat in journals of mathematics and biology as well as in journals of chemical engineering. In a chemostat, the nutrients required for cell growth continuously feed the culture container through a pump connected to the reservoir. The microorganisms inside the chemostat continuously grow on these nutrients. Residual nutrients and microorganisms are removed from the chemostat at the same rate, which allows the culture in the fermenter to be maintained at a constant volume. Laboratory chemostats generally contain 0.5 to 10 litres of culture, whereas industrial chemostat cultures can involve volumes up to 1300 m3 for the continuous production of microbial biomass. The chemostat is also used as a model for wastewater treatment processes. In fact, it can be used to grow microorganisms on highly toxic nutrients in order to reduce their concentrations. The chemostat is very useful in fields such as physiology, ecology and genetics of microorganisms. In its commercial form, it plays an important role in certain fermentation processes, in particular in the commercial production of products by genetically modified organisms (e. g. insulin production).

Part II

Dynamical random systems

Reminder of discrete probability theory

Expectation, variance, conditional expectation, law of large numbers.

Probability, independence and conditional probability.

To describe a random experiment, we can list the set of all possible events and associated to them a probability that each of them occurs. We now provide some definitions in the probability theory language. We recall that the set $\emptyset$ is the empty set (containing nothing).

- **Union of sets.** Let A and B be two abstract sets. Then an element is in $A \cup B$ if it is in A OR B.
- **Intersection of sets.** Let A and B be two abstract sets. Then an element is in $A \cap B$ if it is in A AND B. If $A \cap B = \emptyset$ we say that A and B are disjoints.

  For the sake of simplicity, we sometimes write (A, B) instead of $A \cap B$.

  Moreover, if A and B are disjoint, we write $A \dot{\cup} B$ instead of $A \cup B$ to specify that the set are disjoint.
- **Difference of sets.** Let A and B be two abstract sets. We denote by $A \setminus B$ the set of elements in A without being in B.
- **State space.** The state space denoted commonly by Ω is the state of all possible results induced by a random experiment. We denote by $\mathcal{P}(\Omega)$ the set of parts[1] of Ω.

[1] Of course, $\mathcal{P}(\Omega)$ contains Ω. In fact, at this step we should define the notion of $\sigma-$algebra fundamental to describe the possible combinations of events induced by the random experiment. We take the part to not speak about it for a first approach to the probability theory, it will be more deeply studied in the second year of the program.

$\hookrightarrow$ Examples:

- roll of a dice, the state space is composed by the union of the events

$$\text{" the result is } i\text{"}, \; 1 \leq i \leq 6.$$

  An element in $\mathcal{P}(\Omega)$ is for instance "the result is 1 or 2".
- Consider a market with a stock price with fixed initial value. The time after, it can increases or decrease. The state space is thus composed by two events "the stock price increases" and "the stock price decreases".

Each subset A of Ω is called an event.

- **Complementary set.** Let A be in $\mathcal{P}(\Omega)$. We denote by A^c or sometimes $\overline{A}$ the complementary set of A equals to $\Omega \setminus A$. That is the set containing all elements of Ω excepting those in A.

The fundamental idea in probability theory (and more generally in measure theory) is to associate to each event of the state space "a weight" representing the frequency of the realization of it. These weights are given through a function of the events of the considered random experiment named "probability measure".

Definition II.1. *A probability measure $\mathbb{P}$ is a map form $\mathcal{P}(\Omega)$ into $[0,1]$ such that*

- $\mathbb{P}(\Omega) = 1$,
- *if $A, B \subset \Omega$ with $A \cap B = \emptyset$ then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.*

Example: The uniform probability on a finite space Ω is defined by $\mathbb{P}(A) = \frac{\#A}{\#\Omega}$.

Remark II.1. *Be careful on the fact that the uniform probability is just an example of a probability measure. Of course, there exist infinitely other probability measures. For instance:*

- *Consider a rigged dice which never gives 1. Then, Ω is still composed by 6 events "I get i" for $1 \leq i \leq 6$. But the probability to get 1 is zero and not $1/6$.*
- *throw two (non-rigged) dices and study the sum of the numbers obtained. The event "I get 2" occurs with probability $1/36$ (only 1 and 1 on the 36 possibilities) but the event "I get 4" occurs with probability $3/36$ (1 and 3, 3 and 1, 2 and 2 on the 36 possibilities). We thus define a probability measure on the state space { "I get 2", "I get 3",..., "I get 12" } which is not the uniform law.*

Remark II.2. *Consequently, we say that two probabilities $\mathbb{P}$ and $\mathbb{Q}$ are equal on Ω if $\mathbb{P}(A) = \mathbb{Q}(A)$ for any A subset of Ω.*

Consider a probability measure $\mathbb{P}$ on a state space Ω, we have the following properties (the proof is given as an exercise)

i. $\mathbb{P}(\emptyset) = 0$,

ii. for any part A of Ω we have $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$,

iii. For any parts A, B of Ω and **not necessarily disjoints,** we have

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B).$$

iv. If $A \subset B$ are parts of Ω we have $\mathbb{P}(A) \leq \mathbb{P}(B)$.

v. if $(A_i)_{i\in I}$ is a family of disjoint sets then $\mathbb{P}(\bigcup_{i\in I} A_i) = \sum_{i\in I} P(A_i)$.

Finally, we say that a property holds $\mathbb{P}$-almost surely ($\mathbb{P}-a.s.$ for short) if it is satisfied outside a negligible set $\Omega' \subset \Omega$, (negligible: $\mathbb{P}(\Omega') = 0$).

Definition II.2. *Let Ω be the state space, $\mathcal{P}(\Omega)$ the parts of Ω and $\mathbb{P}$ a probability measure. Then $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ is a probability space.*

Conditioning and Independence.

Conditioning

The notion of conditioning is fundamental in probability theory. It is roughly speaking based on the fact that the realisation of an event can be impacted by the knowledge about the realization of an other event. We have the following mathematical definition.

Definition II.3. *Let $\mathbb{P}$ be a probability measure and A and B two elements in $\mathcal{P}(\Omega)$ with $\mathbb{P}(B) > 0$. Then we define the conditional probability of A knowing B has the number*

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

Note that $\mathbb{P}(A|B) \in [0, 1]$ as a consequence of the previous property v. and $A \cap B \subset B$.

As a consequence of this definition, we have the following result (exercise) if B is a part of a state space Ω and with $\mathbb{P}$ a probability measure, then the map from $\mathcal{P}(\Omega)$ into $[0, 1]$ defined for any element $A \in \mathcal{P}(\Omega)$ by $\mathbb{P}(A|B)$ is a probability measure in the sens of Definition II.1.

An important result that we will use later in the course is the following, named the total probability rule.

Proposition II.1. *Let I be a finite or countable set of indexes. Let $\mathbb{P}$ be a probability measure on a state space Ω and consider the family of disjoint[2] parts of Ω denoted by $(B_i)_{i\in I}$ such that $\mathbb{P}(B_i) > 0$ with $i \in I$ and $\dot{\bigcup}_{i\in I} B_i = \Omega$. Then, for any part A of Ω we have*

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

Sketch of the proof. Note that $\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A \cap B_i)$.

[2] We recall that it means $B_i \cap B_j = \emptyset$ for $i \neq j$.

Application. In a field, we have three types of plants denoted by A (30% of the field), B (20% of the field) and C (50% of the field). We know that

- 12% of A have a disease,
- 7% of B have a disease,
- 15% of C have a disease.

We pick randomly a plant. What is the probability that it has a desease?

Independence

Intuitively, we say that two events are independent if the realization of one of them does not provide information on the realization of the other. That is $\mathbb{P}(A|B) = \mathbb{P}(A)$ and $\mathbb{P}(B|A) = \mathbb{P}(B)$. From the definition of the conditioning, we deduce that the following definition is suitable for this notion.

Definition II.4. *We say that two parts of Ω denoted by A, B are independent if*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Remark II.3. *The notion of independence depends intrinsically on the definition of the probability measure $\mathbb{P}$ chosen. Do not make a confusion between this notion (which depend on the probability measure $\mathbb{P}$) and the notion of disjoint events (which is a property of set theory).*

Random variable

Finite valued discrete random variables

Definition. Let $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ be a probability set. Let $E \subset \mathbb{R}$ a **finite** list of reals, *i .e.* $E = \{x_1, \dots, x_n\}$ with $x_i \in \mathbb{R}$ for any index i. A **discrete**[3] random variable X on E is a map from Ω into E.

Remark II.4. *The notion of continuous random variables will be seen during the second year of the program.*

In other words, a random variable describes the random experiment by associating to each possible event a number. For instance if you throw a dice, X denotes the result so that for any event $\omega_i := \{ \text{ I get } i\}$ with $1 \leq i \leq 6$ we have $X(\omega_i) = i$ and here $E = \{1, 2, 3, 4, 5, 6\}$.

For any random variable we can associate a way to describe the probability of each value taken by defining the law of the random variable.

Definition II.5. *Let n be a positive integer and $X : \Omega \longrightarrow E$ be a discrete random variable with $E := \{x_1, \dots, x_n\}$ is finite with $x_i \in \mathbb{R}$ for any $1 \leq i \leq n$. The law of X denoted by P_X is the probability measure*

$$P_X : \mathcal{P}(E) \longrightarrow [0, 1]$$

such that $P_X(A) = \mathbb{P}(\{\omega \in \Omega,\ X(\omega) \in A\}) =: \mathbb{P}(X \in A)$ for any $A \subset \mathcal{P}(E)$.

[3]Such random variable should be named "discrete random variable $\mathbb{R}$-valued. Since we will only consider real valued random variable here, we make this abuse of definition. For more details see the course MAA 203 (Bachelor program, 2nd year).

The number $\mathbb{P}(X \in A)$ is an abuse of notations to write $\mathbb{P}(\{\omega \in \Omega,\ X(\omega) \in A\})$. In the following, for any $x_i \in E$, the notation $\mathbb{P}(X = x_i)$ has to be understood as $\mathbb{P}(\{\omega \in \Omega,\ X(\omega) = x_i\})$, which is the probability that the random experiment modelled by X returns x_i.

Remark II.5. *Notice that $\mathbb{P}(X \in A) = \sum_{x_i \in A} \mathbb{P}(X = x_i)$. Thus, the law of X is fully characterized by $\mathbb{P}(X = x_i)$ for any $x_i \in E$.*

Examples. *The probability that an asset increases or decreases. The probability of heads or tails.*

Extension to discrete countable valued random variables

We can extend the definition of random variables to countable subset E of $\mathbb{R}$. The definition are similar but the index n goes to $+\infty$ formally (so that we have to extend every notions inducing finite sommations to infinite sommation).

How to check that a function can be associated with a probability law of a (discrete) random variable taking values in $E = \mathbb{N}$? Let p be a function from $\mathbb{N}$ into $\mathbb{R}$ such that

(*a*) $p(n)$ is non-negative for all $n \in \mathbb{N}$,

(*b*) $\sum_{n \in \mathbb{N}} p(n) = 1$.

Hence, we can associate to p a probability measure[4] $\mathbb{P}$ describing the law of a random variable X given by $\mathbb{P}(X = n) = p(n) \in [0, 1]$ for all $n \in \mathbb{N}$.

Independence of random variables, identically distributed random variables and joint distribution

We can also define the independence of two random variables by the following

Independence of random variables. Let X and Y be two random variables on $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Assume that X takes values in $E_X := \{x_1, \ldots, x_n\}$ and Y takes values in $E_Y := \{y_1, \cdots, y_m\}$ with n, m positive integers and x_i, y_j are reals numbers for any $1 \leq i \leq n$ and $1 \leq j \leq m$. We say that X and Y are independent if for all $(x_i, y_j) \in \mathbb{E}_X \times E_Y$ we have

$$\mathbb{P}(X = x_i \cap Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j).$$

We set the notation $X \perp\!\!\!\perp Y$ when X and Y are independent.
A family of random variables $(X_i)_{i \in \mathbb{N}}$ such that each X_i takes its values in a finite set of reals E_{X_i}, is said to be a family of independent random variables if

$$\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \Pi_{i=1}^{n} \mathbb{P}(X_i = x_i),\ \forall x_i \in E_{X_i}.$$

[4] The fact that $p(n)$ is between $[0, 1]$ is a direct consequence of the two properties (a) and (b) satisfied by p.

Random variables identically distributed. We say that a family $(X_i)_{i\in\mathbb{N}}$ is a family of identically distributed random variables if the law P_{X_i} are all the same, that is $P_{X_i} = P_{X_0}$ for all $i \in \mathbb{N}$.

Definition II.6 (Joint distribution)**.** *Let X, Y two discrete random variables from Ω into $E_X := \{x_1, \dots, x_n\}$ and $E_Y := \{y_1, \cdots, y_m\}$ respectively with n, m positive integers and x_i, y_j are reals numbers for any $1 \leq i \leq n$ and $1 \leq j \leq m$. The joint distribution of X and Y is the law of the pair (X, Y), that is*

$$\mathbb{P}(X = x_i, Y = y_j),\ \forall (x_i, y_j) \in E_X \times E_Y.$$

We can determine from the joint law the marginal laws of X and Y, *i.e.*

$$\mathbb{P}(X = x_i) = \sum_{1\leq j\leq m} \mathbb{P}(X = x_i, Y = y_j), \quad 1 \leq i \leq n,$$

$$\mathbb{P}(Y = y_j) = \sum_{1\leq i\leq n} \mathbb{P}(X = x_i, Y = y_j), \quad 1 \leq j \leq m.$$

Expectation and variance

In this section, we aim at characterizing the notion of mean of a random experiment and the minimal distance expected when we make the experiment. These notions are called *expectation* and *variance* of a random experiment (or equivalently random variable).

Let $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ be a probability set and X a discrete random variable on this set E-valued with $E := \{x_1, \dots, x_n\}$, where n is a fixed positive integer and $x_i \in \mathbb{R}$ for any $1 \leq i \leq n$.

Expectation

Definition II.7 (Expectation)**.** *We denote by $\mathbb{E}^{\mathbb{P}}[X]$ the expectation of X under $\mathbb{P}$ and defined by*

$$\mathbb{E}^{\mathbb{P}}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i).$$

Without ambiguity on the underlying probability, we will simply write $\mathbb{E}[X]$**.** *If $\mathbb{E}[X] = 0$ we say that X is centered.*

Proposition II.2. *Let X, Y two random variables on Ω with values in E_X and E_Y respectively with $E_X := \{x_1, \dots, x_n\}$ and $E_Y := \{y_1, \dots, y_m\}$. Let $a \in \mathbb{R}$, then*

(i) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

(ii) $\mathbb{E}[aX] = a\mathbb{E}[X]$

(iii) If $X \perp\!\!\!\perp Y$ then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

(iv) If $X \leq Y$ almost surely, then $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

(v) Let f be a map from E_X into $\mathbb{R}$. Then $\mathbb{E}[f(X)] = \sum_{i=1}^{n} f(x_i)\mathbb{P}(X = x_i)$.

Proof. For (i).

$$\begin{aligned}\sum_{z=x_i+y_j} z\mathbb{P}(X+Y=z) &= \sum_{1\le i\le n, 1\le j\le m, x_i+y_j=z} (x_i+y_j)\mathbb{P}(X=x_i, Y=y_j)\\ &= \sum_{1\le i\le n, 1\le j\le m, x_i+y_j=z} x_i\mathbb{P}(X=x_i, Y=y_j)\\ &+ \sum_{1\le i\le n, 1\le j\le m, x_i+y_j=z} y_j\mathbb{P}(X=x_i, Y=y_j)\\ &= \sum_{1\le i\le n, 1\le j\le m} x_i\mathbb{P}(X=x_i, Y=y_j) + \sum_{1\le i\le n, 1\le j\le m} y_j\mathbb{P}(X=x_i, Y=y_j)\\ &= \sum_{1\le i\le n, 1\le j\le m} x_i\mathbb{P}(X=x_i) + \sum_{1\le i\le n, 1\le j\le m} y_j\mathbb{P}(Y=y_j).\end{aligned}$$

(ii) is trivial. (iii) and (iv) exercise! For property (v) note that $Z := f(X)$ is a random variable taking the values $z_i = f(x_i)$ with $x \in E_X$. It means that $Z = z_i$ is equivalent to $X \in f_{z_i}$ where f_{z_i} is the set of all x_i such that $f(x_i) = z_i$ (with possibly $z_i = z_j$ for some indexes $i \neq j$, since f may not be bijective). Hence

$$\mathbb{E}[Z] = \sum_{i=1}^n z_i\mathbb{P}(Z=z_i) = \sum_{i=1}^n f(x_i)\mathbb{P}(X=x_i).$$

□

BE CAREFUL. Property (iii) does not give a characterization of the independence of random variables. The converse part is of course wrong.

Variance

Definition II.8 (Variance). *We denote by $V[X]$ the variance of X defined by*

$$V[X] = \sum_{i=1}^n (x_i - \mathbb{E}[X])^2\mathbb{P}(X=x_i).$$

Proposition II.3. *Let X, Y two random variables on Ω with values in E_X and E_Y respectively with $E_X := \{x_1, \dots, x_n\}$ and $E_Y := \{y_1, \dots, y_m\}$. Let $a, b \in \mathbb{R}$, then*

(i) $V[X] = \sum_{i=1}^n |x_i|^2\mathbb{P}(X=x_i) - (\mathbb{E}[X])^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

(ii) $V[aX] = a^2V[X]$,

(iii) If $X \perp\!\!\!\perp Y$ *then* $V[aX+bY] = a^2V[X] + b^2V[Y]$.

BE CAREFUL Property (iii) is satisfied only if the random variable are independent. Otherwise there is an additional term (see the second year of the program).

Examples

X random variable from Ω into E which will be specified for each example.

Uniform law. $E := \{x_1, \dots, x_n\}$ with $x_i \in \mathbb{R}$. Then

$$\mathbb{P}(X = x_i) = \frac{1}{\#E} = \frac{1}{n}.$$

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i) = \frac{1}{n}\sum_{i=1}^{n} x_i, \text{ and } V[X] = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2$$

$\hookrightarrow$ Classical mean and variance in statistics.

Bernoulli law. $E = \{0, 1\}$ with $\mathbb{P}(X = 1) = p \in (0, 1)$. Then

$$\mathbb{E}[X] = p,\ V[X] = p - p^2 = p(1-p).$$

Rademacher law. $E = \{-1, 1\}$ with $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = \frac{1}{2}$. Then

$$\mathbb{E}[X] = 0,\ V[X] = 1.$$

Binomial law. $E = \{0, \dots, n\}$ with $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

$$\mathbb{E}[X] = \sum_{k=0}^{n} \binom{n}{k} k p^k (1-p)^{n-k} = np,\ V[X] = np(1-p).$$

Remark II.6. *In particular, if $(X_i)_{i\in\mathbb{N}}$ is a family of identically distributed random variables then they have the same mean (for instance $\mathbb{E}[X_0]$) and the same variance (for instance $V[X_0]$).*

Conditional expectations

Let X, Y two random variables on Ω with values in E_X and E_Y respectively with $E_X := \{x_1, \dots, x_n\}$ and $E_Y := \{y_1, \dots, y_m\}$. We define the conditional law of Y given a realization x_i of X with $\mathbb{P}(X = x_i) > 0$ by

$$P(Y|X = x_i)(y_j) := \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(X = x_i)},\ y_j \in E_Y.$$

It is a probability measure. In case of independence of the random variables involved, the conditioning disappears and we recover the law of Y.
Notice that $\mathbb{P}(Y = y_j|X = x_i) = P(Y|X = x_i)(y_j)$.

Definition II.9. *The conditional expectation of Y with respect to the event $\{X = x_i\}$ with $x_i \in E_X$ is the expectation associated with the probability measure $P(Y|X = x_i)$, i.e.*

$$\mathbb{E}[Y|X = x_i] = \sum_{j=1}^{m} y_j \mathbb{P}(Y = y_j|X = x_i).$$

Similar definitions hold for the conditional law of X given Y and its conditional expectation.

Law of large numbers

In this section[5] we will focus on a way to approach the expectation of a random experiment. This is mainly based on the idea that the empirical mean of identical and independent experiments should be relevant to characterize the mean of a random event when the number of the considered experiments is large. For instance, when we throw a non-rigged coin and we win 1€ if head occurs and we loose 1€ otherwise, we expect to win in mean 0€ when we repeat a lot of times the experiment. This induces to define the notion of convergence of a family of random experiments.

From now, $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ is a fixed probability space.

In this section we assume that E is not necessarily finite but countable part of $\mathbb{R}$. All the previous definitions can be extended to random variables taking values in E.

Mean and mean-square convergences toward a constant

Mean convergence Let $(X_n)_{n\in\mathbb{N}}$ be a family of random variables defined on $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ such that $\mathbb{E}[|X_n|] < +\infty$ for all $n \in \mathbb{N}$. We say that X_n converges in mean of order 1 (or simply in mean) to a constant $\mu \in \mathbb{R}$ if

$$\lim_{n\to+\infty} \mathbb{E}[|X_n - \mu|] = 0.$$

Mean-square convergence Similarly, assume that $\mathbb{E}[|X_n|^2] < +\infty$ for all $n \in \mathbb{N}$. We say that X_n converges in mean of order 2 (or also called mean-square convergence or L^2 convergence) to a constant $\mu \in \mathbb{R}$ if

$$\lim_{n\to+\infty} \mathbb{E}[|X_n - \mu|^2] = 0.$$

We now show that the mean square-convergence induces the mean-convergence.

Proposition II.4. *Let $(X_n)_{n\in\mathbb{N}}$ be a family of random variables defined on $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ with $x_i \in \mathbb{R}$ such that $\mathbb{E}[|X_n|^2] < +\infty$ for all $n \in \mathbb{N}$. Assume that X_n converges in mean-square to a constant μ. Then $\mathbb{E}[|X_n|] < +\infty$ and X_n converges in mean to μ.*

Proof. First note that $|x| \leq 1 + |x|^2$. Since $\mathbb{E}[|X_n|^2] < +\infty$, by using Proposition II.2 (iv), we deduce that $\mathbb{E}[|X_n|] < +\infty$. Assume now that

$$\lim_{n\to+\infty} \mathbb{E}[|X_n - \mu|^2] = 0.$$

[5] This section can be omitted, it will be more deeply studied during the second year.

By using the convexity[6] of $x \longmapsto (x-\mu)^2$ we deduce that[7]

$$\mathbb{E}[|X_n-\mu|^2] = \sum_{i=1}^{+\infty} |x_i-\mu|^2 \mathbb{P}(X_n = x_i) \geq \big(\sum_{i=1}^{+\infty} |x_i-\mu| \mathbb{P}(X_n = x_i)\big)^2 = \big(\mathbb{E}[|X_n-\mu|]\big)^2,$$

which gives the result. □

Law of large numbers

We can now give one particularly law of large number to approach the mean of a random experiment.

Theorem II.1. *Let $(X_i)_{i\in\mathbb{N}}$ be a family of random variables independent and identical distributed defined on $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ such that $\mathbb{E}[|X_i|^2] < +\infty$ and such that $\mathbb{E}[X_i] = \mu \in \mathbb{R}$ for all $i \in \mathbb{N}$. Then $Y_n := \frac{1}{n}\sum_{i=1}^n X_i$ converges in mean-square and in mean to μ when n goes to $+\infty$.*

Proof. Note that since all the X_i are independent and identical distributed, we have

$$\mathbb{E}[Y_n] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \mu, \text{ and } V[Y_n] = \frac{1}{n^2}\sum_{i=1}^n V[X_i] = \frac{V[X_0]}{n}.$$

Hence,

$$\begin{aligned}
\mathbb{E}[|Y_n-\mu|^2] &= \mathbb{E}[|Y_n|^2] + \mu^2 - 2\mu\mathbb{E}[Y_n] \\
&= V[Y_n] + (\mathbb{E}[Y_n])^2 + \mu^2 - 2\mu\mathbb{E}[Y_n] \\
&= \frac{V[Y_0]}{n} + \mu^2 + \mu^2 - 2\mu^2 \\
&= \frac{V[Y_0]}{n} \longrightarrow 0, \text{ when } n \to +\infty.
\end{aligned}$$

It proves that $Y_n := \frac{1}{n}\sum_{i=1}^n X_i$ converges in mean-square (and so in mean from Proposition II.4) to μ when n goes to $+\infty$. □

[6] We recall that if $f : x \longmapsto f(x)$ from F into $\mathbb{R}$ is convex then for any $\lambda_i \in [0,1]$, $1 \leq i \leq n$ such that $\sum_{i=1}^n \lambda_i = 1$ we have

$$f(\sum_{i=1}^n \lambda_i x_i) \leq \sum_{i=1}^n \lambda_i f(x_i), \ \forall x_i \in F.$$

We apply it to $f(x) = (x-\mu)^2$ by noting that $\sum_{x\in\mathcal{S}_{X_n}} \mathbb{P}(X_n = x) = 1$.

[7] In fact, this is a direct consequence of Jensen's Inequality that we directly prove here for discrete random variables. See the second year of the program for the generalization of this proof.

Chapter 5

What is the fair price of a financial product? A tale of arbitrage theory and European options

Dynamical portfolio, arbitrage, European options and Call-put parity.

5.1 Introduction to finance and first definitions

In the first lecture we have seen the mechanism of loans as a meeting between a financial agent wanting a certain amount of money and a lender without any uncertainty. In this lecture we will focus on the management of the randomness in risky financial products.

Financial markets are a way to exchange liquidity between different actors having different views on the price of some financial products. The aim of the exchange is to optimize the cash flow of each party in order to ensure huge liquidity on the market and to manage the different financial and economical risks by emitting sophisticated products. Here, we distinguish three types of financial products:

- A share. It is a small unit of ownership of a firm/a company ensuring to its owner dividends or the right to interfere in the management of the firm. The aggregation of all the shares is called the stock of a firm/company.
- A bond. It is a debt security for which the issuer has a debt (by paying coupons) to the owner at determined time in a contract. This financial product was studied a lot in the first lecture!
- The derivative products. This type of financial products is the more sophisticated since it groups all the possible combinations of simple financial products. Its value depends on an underlying financial product along a fixed period. In this lecture we will consider two kind of derivatives: futures contracts and options (see below for their respective definition).

Financial lingo. A short, or a short position, is created when a trader sells a security. Conversely, a long or a long position is the buying of a security.

Uncertainty on the price of a financial product results from its daily random variation on the market. The main challenge is thus to detemrined the fair price of a financial product for each party (the seller and the buyer of it), given that

- the buyer has a risk limited to this price,
- the seller is very sensitive to the variation on the market to compute her own risks.

As a consequence of the "law of one price" we expect that this price is determined as an equilibrium between the supply and the demand. All the point is to determined fairly such price given the randomness of the markets considered.

Louis Bachelier (1870-1946) is credited to be the precursor in the modeling of randomness on the markets in his thesis "Théorie de la spéculation". In his PhD thesis, Bachelier has used the Brownian motion (a very particular stochastic process, chaotic in its variations) to model the fluctuations of a market. Inspiring by the work of Bachelier, Peter Black and Myron Scholes (in 1973) have emphasized that the price of a financial product has to be exactly the price of the hedging strategy of the seller, by proving the famous Black-Scholes formula to compute explicitly the price of European options in a particular model. This work was then published in a mathematical framework by Merton. Determine the price of a financial products is still (and maybe even more these last decades) a very huge challenge. This lecture introduced the notion of *fair* price of a financial product and how to determine it in very simple models.

To quote Nicole El Karoui (M2 Probability and Finance lecture notes): "*The prices of financial products are not completely exogenous and strongly depends on underlying assets due to the "fundamental law of finance market":*

In a market with high liquidity, without neither transaction costs nor limit on the management (buy-sell) of basic asset, it is not possible to win money for sure without any investment. "

This notion of "wining money without any investment" is exactly at the heart of the notion of arbitrage in finance. We can thus provide a first informal definition

Definition 5.1.1 (Arbitrage (informal))**.** *We say that a market has an arbitrage opportunity if there exists a buying/selling strategy of different assets on the market without initial cost and which brings positive incomes for sure.*

5.2 Arbitrage opportunities

In this section we provide a rigorous definition of an arbitrage opportunity. From now, we assume that the following standing assumption holds

- Any assets can be infinitely divided (*i.e.* we can for instance buy/sell $\sqrt{2}$ assets on a market),
- The market is liquid (*i.e.* we can buy or sell at any moment),

- We can short sell[1] any asset,
- There is no transaction costs for any exchanged flow.

Although these assumptions are very restrictive and not realistic, as a first approach to the Arbitrage Theory, there are necessary to ensure the computations of the prices of derivatives together with hedging strategies in the models studied.

We consider a probability set $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, where Ω is the set of the possible states of the market and $\mathbb{P}$ is the probability of events' occurence (called the historical probability).

5.2.1 Assets on the market and general definitions

We consider a one period model for the market which evolves between an initial time $t = 0$ and some maturity $T > 0$. We consider three[2] types of assets on it:

- A risky asset, with initial price given by S_0, known at time 0, and with price S_T at time $T > 0$ which is a random variable on Ω with probability law $\mathbb{P}$. The term risky comes from the fact that the terminal value S_T is uncertain from the initial time.
- A non-risky asset with initial price 1, and with price $(1 + r)^T$ at time T where r is a deterministic positive constant (known at time $t = 0$).
- A general financial product proposed by a seller (denoted by Π) with initial price given by π_0 and random price π_T at time T as a function of the price of the risky asset and the non risky asset. The value π_T is called **the payoff** of such product at time T.

The general product Π as to be seen in this course as the financial asset that is traded between two financial agents: a seller and a buyer. More exactly, the seller proposes a contract to the buyer ensuring to him the value π_T at time T. The buyer has thus to pay the price π_0 to the seller if she accepts it. We will consider different financial product Π in the following: bonds (see Section 5.2.3), futures (see Section 5.3.1) and European options (see Section 5.3.2).

The notion of non-risky asset is set here for the sake of consistency concerning the objects considered. It has to be simply seen as a bank account in which a financial agent can invest at any moment. That is to say: buying (resp. selling) K non-risky assets is equivalent to invest (resp. borrow) $K \times 1$€ into a bank account with simple interest rate $r > 0$.

Of course, we also make the convention that buying (resp. selling) $K < 0$ assets remains to sell (resp. buy) $|K|$ assets.

Remark 5.2.1. *Be careful on the fact that π_T is a function of S and the non risky asset, that is to say $\pi_T = f(S_T, r)$, where f is known by all the financial agents. However, due to the randomness of S, this value π_T is itself random and thus uncertain, although the shape of f is known.*

[1] Short selling is the sale of a security that is not owned by the seller but which has to be paid at some fixed maturity.

[2] Again, here we make this assumption for the sake of simplicity to determine the non-arbitrage price of the product Π. More generally, we can consider more financial product but for this lecture it will not be necessary to do it to compute the non-arbitrage price. We will extend this model by considering four types of asset (risky assets, non-risky assets, call options and put options) to prove the call-put parity at the end of this lecture.

Definition 5.2.2 (Portfolio and strategy)**.** *Consider a market composed by the risky asset S, the non risky asset and the product Π.*

- *A financial portfolio or a strategy is a triplet (n_Π, n_S, n_M) of real numbers where n_Π, n_S, n_M denote respectively the number of general financial product Π, the number of risky asset S and the number of non-risky asset bought at time 0.*
- *The values at times $t = 0$ and $t = T$ of the portfolio associated with the triplet (n_Π, n_S, n_M) and denoted respectively by V_0 and V_T are given by*

$$V_0 = n_\Pi \pi_0 + n_S S_0 + n_M,$$

and

$$V_T = n_\Pi \pi_T + n_S S_T + n_M(1+r)^T, \ \mathbb{P} - a.s.$$

We can now state more rigorously the definition of an arbitrage

Definition 5.2.3 (Arbitrage)**.** *An arbitrage opportunity is a strategy (n_Π, n_S, n_M) associated to a financial product, the risky asset and the non risky asset such that*

$$V_0 = 0, \ V_T \geq 0, \mathbb{P} - a.s. \text{ and } \mathbb{P}(V_T > 0) > 0. \tag{5.1}$$

Remark 5.2.4. *Recall that V_T as a function of S is a random variable, that is why all the equalities/inequalities involving V_T are given $\mathbb{P}$ almost surely (i.e. outside a negligible set of possible event w.r.t. $\mathbb{P}$, see Lecture 5).*

Condition (5.1) simply means that a financial agent has made initially financial operations on the products considered, the risky asset and the non risky asset without paying anything ($V_0 = 0$), and at time $t = T$, his portfolio value is non-negative ($V_T \geq 0$, $\mathbb{P}$-a.s.) and positive in at least one possible state of the market ($\mathbb{P}(V_T > 0) > 0$). In other words, we say that there the "absence of arbitrage" opportunities holds if

$$\textbf{(AoA)}: V_0 = 0 \text{ and } V_T \geq 0, \mathbb{P} - a.s. \Longrightarrow \mathbb{P}(V_T > 0) = 0.$$

Pricing and Hedging problem: As said before and now stated formally, the seller proposed a contract to the buyer with initial price π_0 and ensuring to him the payoff π_T at time $t = T$.
The procedure is the following:

1. First, the buyer gives at time 0 her choice of desired payoff π_T, given at time T, to the seller.
2. The seller proposes a price π_0 paid by the buyer at time 0 in order to build a strategy[3] $(-1, n_S, n_M)$ ensuring to propose the payoff π_T for the buyer. This strategy is called **the hedging strategy** associated with the payoff π_T.
3. At time T, the seller gives the payoff to the buyer.

[3] Of course here $n_\Pi = -1$ since the seller gives the product to the buyer. The strategy remains to determine n_S and n_M.

The main difficulty here is to determined the price π_0 of the contract considered so that no arbitrage opportunity hold for both the seller and the buyer of such contract.

The problem of the seller is thus to find the arbitrage price π_0 and the hedging strategy (n_S, n_M) allowing her to replicate π_T.

5.2.2 Useful properties

In this section we derive from the definition of a portfolio a comparison principle under Assumption **(AoA)**.

Proposition 5.2.5 (Comparison Principle)**.** *Assume that **(AoA)** holds. If the value of a portfolio V is greater at time T than the value at time T of another $\tilde{V}$, the comparison relation is the same at time 0:*

$$V_T \geq \tilde{V}_T, \; \mathbb{P} - a.s. \Longrightarrow V_0 \geq \tilde{V}_0. \tag{5.2}$$

Proof. Let V and $\tilde{V}$ be two portfolios associated with the strategy (n_Π, n_S, n_M) and $(\tilde{n}_\Pi, \tilde{n}_S, \tilde{n}_M)$ respectively. Assume that $V_T \geq \tilde{V}_T$ and $V_0 < \tilde{V}_0$. Then, at time 0, we sell the position associated with $\tilde{V}$ and we buy the position associated with V. In other words, we make a short sale of $\tilde{n}_S$ risky asset, we buy n_S risky asset, we make a short sale of $\tilde{n}_M$ non risky asset (borrow at the bank), we buy n_M non risky asset (invest into the bank account), we sell $\tilde{n}_\Pi$ financial product Π and we buy n_Π of Π. The payoff at time 0 is

$$(\tilde{n}_\Pi - n_\Pi)\pi_0 + (\tilde{n}_S - n_S)S_0 + (\tilde{n}_M - n_M) = \tilde{V}_0 - V_0 > 0$$

that we put into the non-risky asset (bank account) to have 0. At time T, we have $-\tilde{n}_S S_T - \tilde{n}_M(1+r)^T$ (result of the short sale at time 0) and $-\tilde{n}_\Pi \pi_T$ (we have sold $\tilde{n}_\Pi$ products Π) $+n_S S_T + n_M(1+r)^T + n_\Pi \pi_T$ (we have bought portfolio V, at time 0) and $(\tilde{V}_0 - V_0)(1+r)^T$ (we have invested the difference into the non-risky asset). The terminal payoff is thus

$$V_T - \tilde{V}_T + (\tilde{V}_0 - V_0)(1+r)^T > 0, \; \mathbb{P} - a.s.$$

we have built an arbitrage. We deduce that the comparison principle (5.2) holds. □

As a direct consequence of this principle we have also the following result very useful to find the non arbitrage price of particular products (as soon as the dependancy of the payoff is for instance linear with respect to the underlying asset).

Corollary 5.2.6. *Assume that **(AoA)** holds. If two portfolio V and $\tilde{V}$ have the same values at time T, they have the same values at time 0:*

$$V_T = \tilde{V}_T, \; \mathbb{P} - a.s. \Longrightarrow V_0 = \tilde{V}_0.$$

5.2.3 Application to bonds

A bond[4] is a contract ensuring to his owner (the buyer) a (deterministic) coupon $F > 0$ at any future time T. It is quite intuitive (see the first lecture!) that the non-arbitrage price of such contract is

$$\pi_0 = \frac{F}{(1+r)^T}. \tag{5.3}$$

First method: by using (informal) Definition 5.1.1 and an arbitrage tabular We proceed by contradiction. Assume that $\pi_0 > \frac{F}{(1+r)^T}$. Then, the seller receives π_0 from the buyer and she invests π_0 into the non-risky asset. The total amount of money is thus 0 at time 0 in her portfolio. At time T, she must give to the buyer F and she wins the interest of the non-risky asset. Her portfolio value is thus

$$V_T = -F + \pi_0(1+r)^T > -F + \frac{F}{(1+r)^T}(1+r)^T = 0.$$

We can sum up this strategy with the following arbitrage tabular

Strat. time 0	Payoff time 0	Payoff time T
Sell the contract	$+\pi_0$	$-F$
Investment π_0 in the non-risky asset	$-\pi_0$	$\pi_0(1+r)^T$
Total	0	$-F+\pi_0(1+r)^T > 0$

Leading to an arbitrage, hence $\pi_0 \leq \frac{F}{(1+r)^T}$.

To prove the equality, we assume that $\pi_0 < \frac{F}{(1+r)^T}$ and we build an arbitrage strategy for the buyer similarly.

Strat. time 0	Payoff time 0	Payoff time T
Buy the contract	$-\pi_0$	$+F$
Sell π_0 non-risky asset	$+\pi_0$	$-\pi_0(1+r)^T$
Total	0	$F-\pi_0(1+r)^T > 0$

We have build an arbitrage for the buyer.

Hence, the non-arbitrage price is given by (5.3).

Second method: by using Corollary 5.2.6. Consider the portfolio P_1 with $n_M = \frac{F}{(1+r)^T}$. Consider now the portfolio P_2 with $n_\Pi = 1$ containing 1 bond. Hence, the values of P_1 and P_2 coincide at time T (since there are both equal to F) and we deduce from Proposition 5.2.6 that (5.3) holds.

[4] Again, for the sake of simplicity, we have consider a one-period model here. In reality, a bond provides a certain number N of deterministic coupons F_i at future times T_i with $T_1 < \cdots < T_N$. We can similarly prove by adapting our notations (exercise!) that the non-arbitrage price is given by

$$\pi_0 = \sum_{i=1}^{N} \frac{F_i}{(1+r)^{T_i}}.$$

5.3 Arbitrage and derivative products

In this section, we aim at finding the non-arbitrage price of two types of derivative products: forward/future contracts and European options. Before turning to the definition of these two contracts, let us recall the main characteristic of a derivative.

A derivative product involves two financial agents: the buyer and the seller of it. It is characterized by a contract between the two parties with characteristic parameters

- the maturity of the contract T,
- the price S of the underlying asset
- its payoff,
- the fixed price traded at time T, named the strike and denoted by K, which is a positive constant fixed by the contract.

The notion of strike is at the heart of the risk management. The aim of such contract it to hedge the buyer on possible important variations of the underlying asset.

5.3.1 Forward/futures contracts

Both forward and futures contracts involve the agreement between two parties to buy and sell an asset at a specified price at a certain time T. A forward contract is a private agreement that settles at the end of the agreement and is made *over the counter*. A futures contract is traded on an exchange and is settled on a daily basis until the end of the contract. The later is a standaridized product on the market, unlike the former.

Definition 5.3.1 (Forward/futures)**.** *A forward/futures sets at time* 0 *gives the its owner (the buyer)* ***the obligation*** *to buy an asset at a fixed price, named the strike* $K > 0$ *at the maturity* $T > 0$ *fixed in the contract* $T > 0$.

The seller has to propose **a non-arbitrage price** ensuring the underlying payoff associated with the futures.

Let S be the price of the underlying asset. If for instance the trader is long at time 0 in the futures, *i.e.* she has the obligation to buy the asset at time T at the price K, her payoff is given by $S_T - K$. In other words, she wins $S_T - K$ if $K < S_T$ (she buys an asset at a price cheaper than its market value). Conversely, she looses $S_T - K$ if $K > S_T$ since she has the obligation to buy an asset with market price S_T at a higher price K.

Remark 5.3.2. *For a short position with respect to this futures, the payoff is symmetrically given by* $-(S_T - K) = K - S_T$.

From now we consider only long positions in the futures, *i.e.* the buyer of the futures has the obligation to buy the underlying asset S at a fixed price K so that is payoff at time T is $S_T - K$. We want to determine the price of such contract.

As usual, in addition to the risky asset S, we consider a market with a fixed interest rate $r > 0$ in a bank account (non-risky asset).

Theorem 5.3.3. *The non-arbitrage price π_0^f at time 0 of a futures with associated payoff $S_T - K$ at the maturity T for its owner, with interest rate r and where K is the fixed strike of the contract, is given by*

$$\pi_0^f = S_0 - \frac{K}{(1+r)^T}.$$

We prove this theorem with the two methods introduced before.

First method: arbitrage table. Assume that $\pi_0^f > S_0 - \frac{K}{(1+r)^T}$ then

Strat. at time 0	Profit at time 0	Payoff at time T
Sell the future	$+\pi_0^f$	$-(S_T - K)$
Buy an asset	$-S_0$	$+S_T$
Borrow $\frac{K}{(1+r)^T}$	$+\frac{K}{(1+r)^T}$	$-K$
Diff. at the bank	$-(\pi_0^f - S_0 + \frac{K}{(1+r)^T})$	$(\pi_0^f - S_0 + \frac{K}{(1+r)^T})(1+r)^T$
Total	0	$(\pi_0^f - S_0 + \frac{K}{(1+r)^T})(1+r)^T > 0.$

This table and associated strategies lead to an arbitrage. Thus, $\pi_0^f \leq S_0 - \frac{K}{(1+r)^T}$.

Assume now that $\pi_0^f > S_0 - \frac{K}{(1+r)^T}$, we similarly (reverse the operations made in the table) prove that an arbitrage holds. Hence $\pi_0^f = S_0 - \frac{K}{(1+r)^T}$.

Second method: by using Corollary 5.2.6. Consider a first portfolio V with $n_\Pi = 0, n_M$[5] $= -\frac{K}{(1+r)^T}$ and $n_S = 1$. Its value at time 0 is given by $V_0 = S_0 - \frac{K}{(1+r)^T}$. Its value at time T is $S_T - K$ which is the payoff of a portfolio containing the futures. We deduce that $\pi_0^f = S_0 - \frac{K}{(1+r)^T}$.

5.3.2 European options

Note that a futures contract may induce very huge looses for its owner since the payoff is not bounded from below. To overcome this kind of risks, we aim at modifying the payoff of the futures so that its owner can leave the contract in case of negative payoff at time T.

Definition 5.3.4. *A European option is a contract between two parties giving to the buyer* ***the possibility*** *to buy (call option) or to sell (put option) an asset at fixed price K called the strike.*

As a consequence of this definition, if an agent is long in a call option (*i.e.* the agent is the buyer of a call option), her payoff[6] is : $(S_T - K)^+ = \max(S_T - K, 0)$. Conversely, if an agent is short in a call option (*i.e.* the agent is the seller of a call option), her payoff is : $-(S_T - K)^+$.

Hence, if at time T, $S_T > K$, the buyer (owner) of a call option wins the difference thanks to the contract, otherwise she does not exercise it so that her looses is limited to the price of the option. The seller loses $S_T - K$ in case of exercise.

[5] the negative sign corresponds to a short-sale of the non risky asset

[6] From now x^+ denotes the positive part of x that is x^+ values 0 if $x \leq 0$ and x if x is positive. In other words, $x^+ = \max(x, 0)$.

Symmetrically, if an agent is long in a put option (*i.e.* the agent is the buyer of a put option), her payoff is : $(K - S_T)^+ = \max(K - S_T, 0)$. Conversely, if an agent is short in a put option (*i.e.* the agent is the seller of a put option), her payoff is : $-(K - S_T)^+$.

Hence, if at time T, $S_T < K$, the buyer (owner) of a put option wins the difference thanks to the contract (she sells the asset at the price K greater that the market price of it), otherwise she does not exercise it. The seller loses $K - S_T$ in case of exercise.

Remark 5.3.5. *Unlike the futures, the payoff of a European option is not symmetric, that is to say, being long in a call option is not equivalent to be short in a put since* $x^+ \neq -(-x)^+$.

Some very famous properties on European options

As we will see in the section below, the price of European options is not easy to compute and will be studied in the Challenge in the famous binomial model. We give here three fundamental properties of Call/Put Options.

Proposition 5.3.6 (Monotonicity with respect to the strike.)**.** *The price of a call (resp. put) option is non-increasing (resp. non-decreasing) with respect to the strike.*

Proof. We denote by $\pi_0(a)$ the price of a call with strike $a > 0$. Let $0 \leq \tilde{K} \leq K$.

First proof: by using an arbitrage table. Assume that $\pi_0(K) > \pi_0(\tilde{K})$.

Strat. at time 0	Profit at time 0	Payoff at time T
Sell 1 call with strike K	$+\pi_0(K)$	$-(S_T - K)^+$
Buy 1 call with strike $\tilde{K}$	$-\pi_0(\tilde{K})$	$+(S_T - \tilde{K})^+$
Diff. at the bank	$-(\pi_0(K) - \tilde{\pi}_0(\tilde{K}))$	$(\pi_0(K) - \tilde{\pi}_0(\tilde{K}))(1+r)^T$
Total	0	$(\pi_0(K) - \tilde{\pi}_0(\tilde{K}))(1+r)^T + (S_T - \tilde{K})^+ - (S_T - K)^+$.

Since the function $k \longmapsto (S_T - k)^+$ is decreasing, we deduce that $(S_T - \tilde{K})^+ - (S_T - K)^+ \geq 0$ so that

$$(\pi_0(K) - \tilde{\pi}_0(\tilde{K}))(1+r)^T + (S_T - \tilde{K})^+ - (S_T - K)^+ > 0,$$

leading to an arbitrage. Hence, if $0 \leq \tilde{K} \leq K$ under **(AoA)** we have $\pi_0(K) \leq \pi_0(\tilde{K})$.

Second proof: by using the comparison principle. Consider a portfolio V long in a call option with strike K and a portfolio $\tilde{V}$ long in a call option with strike $\tilde{K}$. Assume that $\tilde{K} \leq K$. Hence, since $k \longmapsto (S_T - k)^+$ is decreasing, we deduce that $\tilde{V}_T \geq V_T$. From Proposition 5.2.5, we get $\pi_0(\tilde{K}) \geq \pi_0(K)$.

For put options: similar proofs by noting that $k \longmapsto (k - S_T)^+$ is increasing. ☐

Proposition 5.3.7 (Convexity with respect to the strike.). *The prices of a call and put options are convex*[7] *with respect to the strike.*

Proof. • **For call options.** We denote by $\pi_0(a)$ the price of a call with strike $a > 0$. Let $\lambda \in [0,1]$ and $K, \tilde{K}$ two positive strikes. We have to proved that

$$\pi_0(\lambda K + (1-\lambda)\tilde{K}) \leq \lambda \pi_0(K) + (1-\lambda)\pi_0(\tilde{K}). \tag{5.4}$$

First proof: by using an arbitrage table. Assume that $\Delta_{K,\tilde{K},\lambda} > 0$, with

$$\Delta_{K,\tilde{K},\lambda} := \pi_0(\lambda K + (1-\lambda)\tilde{K}) - \big(\lambda \pi_0(K) + (1-\lambda)\pi_0(\tilde{K})\big).$$

Strat. at time 0	Profit at time 0	Payoff at time T
Sell the call with strike $\lambda K + (1-\lambda)\tilde{K}$	$+\pi_0(\lambda K + (1-\lambda)\tilde{K})$	$-(S_T - (\lambda K + (1-\lambda)\tilde{K}))^+$
Buy λ call with strike K	$-\lambda \pi_0(K)$	$+\lambda(S_T - K)^+$
Buy $(1-\lambda)$ call with strike $\tilde{K}$	$-(1-\lambda)\pi_0(\tilde{K})$	$+(1-\lambda)(S_T - \tilde{K})^+$
Diff. at the bank	$-\Delta_{K,\tilde{K},\lambda}$	$+\Delta_{K,\tilde{K},\lambda}(1+r)^T$
Total	0	X

where

$$\begin{aligned} X &:= \Delta_{K,\tilde{K},\lambda}(1+r)^T + \lambda(S_T - K)^+ + (1-\lambda)(S_T - \tilde{K})^+ - \Big(S_T - (\lambda K + (1-\lambda)\tilde{K})\Big)^+ \\ &= \Delta_{K,\tilde{K},\lambda}(1+r)^T + \lambda(S_T - K)^+ + (1-\lambda)(S_T - \tilde{K})^+ - \Big(\lambda(S_T - K) + (1-\lambda)(S_T - \tilde{K})\Big)^+. \end{aligned}$$

Since the function $x \longmapsto x^+$ is convex we deduce that $X \geq \Delta_{K,\tilde{K},\lambda}(1+r)^T > 0$. This leads to an arbitrage. Hence, under **(AoA)** the relation (5.4) holds.

Second proof: by using the comparison principle. Consider a portfolio V containing a call with strike $\lambda K + (1-\lambda)\tilde{K}$ and another $\tilde{V}$ containing λ call with strike K and $(1-\lambda)$ call with strike $\tilde{K}$. The values at time T are given by

$$\begin{aligned} V_T &= \Big(S_T - (\lambda K + (1-\lambda)\tilde{K})\Big)^+ \\ &= \Big(\lambda(S_T - K) + (1-\lambda)(S_T - \tilde{K})\Big)^+ \\ &\leq \lambda(S_T - K)^+ + (1-\lambda)(S_T - \tilde{K})^+ \\ &= \tilde{V}_T, \end{aligned}$$

where the inequality holds by using the convexity of the function $x \longmapsto x^+$. From Proposition 5.2.5 we get (5.4).

[7] We recall that a function f is convex if for any $\lambda \in [0,1]$ and any y, z we have $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$.

- **For put option.** We denote by $\nu_0(a)$ the price of a put with strike $a > 0$. Let $\lambda \in [0, 1]$ and $K, \tilde{K}$ two positive strikes. We have to proved that

$$\nu_0(\lambda K + (1-\lambda)\tilde{K}) \leq \lambda \nu_0(K) + (1-\lambda)\nu_0(\tilde{K}). \tag{5.5}$$

First proof: by using an arbitrage table. Assume that $\Delta_{K,\tilde{K},\lambda} > 0$, with

$$\Delta_{K,\tilde{K},\lambda} := \nu_0(\lambda K + (1-\lambda)\tilde{K}) - \big(\lambda \nu_0(K) + (1-\lambda)\nu_0(\tilde{K})\big).$$

Strat. at time 0	Profit at time 0	Payoff at time T
Sell the put with strike $\lambda K + (1-\lambda)\tilde{K}$	$+\nu_0(\lambda K + (1-\lambda)\tilde{K})$	$-((\lambda K + (1-\lambda)\tilde{K}) - S_T)^+$
Buy λ put with strike K	$-\lambda\nu_0(K)$	$+\lambda(K - S_T)^+$
Buy $(1-\lambda)$ put with strike $\tilde{K}$	$-(1-\lambda)\nu_0(\tilde{K})$	$+(1-\lambda)(\tilde{K} - S_T)^+$
Diff. at the bank	$-\Delta_{K,\tilde{K},\lambda}$	$+\Delta_{K,\tilde{K},\lambda}(1+r)^T$
Total	0	X

where

$$\begin{aligned}
X &:= \lambda(K - S_T)^+ + (1-\lambda)(\tilde{K} - S_T)^+ - \Big((\lambda K + (1-\lambda)\tilde{K}) - S_T\Big)^+ + \Delta_{K,\tilde{K},\lambda}(1+r)^T \\
&= \Delta_{K,\tilde{K},\lambda}(1+r)^T + \lambda(K - S_T)^+ + (1-\lambda)(\tilde{K} - S_T)^+ - \Big(\lambda(K - S_T) + (1-\lambda)(\tilde{K} - S_T)\Big)^+.
\end{aligned}$$

Since the function $x \longmapsto x^+$ is convex we deduce that $X \geq +\Delta_{K,\tilde{K},\lambda}(1+r)^T > 0$. This leads to an arbitrage. Hence, under **(AoA)** the relation (5.5) holds.

Second proof: by using the comparison principle. Similarly, consider a portfolio V containing a put with strike $\lambda K + (1-\lambda)\tilde{K}$ and another $\tilde{V}$ containing λ put with strike K and $(1-\lambda)$ put with strike $\tilde{K}$. The values at time T are given by

$$\begin{aligned}
V_T &= \Big((\lambda K + (1-\lambda)\tilde{K}) - S_T\Big)^+ \\
&= \Big(\lambda(K - S_T) + (1-\lambda)(\tilde{K} - S_T)\Big)^+ \\
&\leq \lambda(K - S_T)^+ + (1-\lambda)(\tilde{K} - S_T)^+ \\
&= \tilde{V}_T,
\end{aligned}$$

where the inequality holds by using the convexity of the function $x \longmapsto x^+$. From Proposition 5.2.5 we get (5.5).

□

Proposition 5.3.8 (Call/Put parity.). *Let π_0 and ν_0 be respectively the prices of a call option and a put option at time 0 with same maturity T, same interest rate r and same strike K. Then, we have the following relation*

$$\pi_0 - \nu_0 = S_0 - \frac{K}{(1+r)^T}. \tag{5.6}$$

Proof. **First proof: by using an arbitrage table.** Assume that $R_{K,r} > 0$, with

$$R_{K,r} := \pi_0 - \nu_0 - S_0 + \frac{K}{(1+r)^T} > 0.$$

Strat. at time 0	Profit at time 0	Payoff at time T
Sell the call	$+\pi_0$	$-(S_T - K)^+$
Buy the put	$-\nu_0$	$+(K - S_T)^+$
Buy S	$-S_0$	$+S_T$
Borrow $\frac{K}{(1+r)^T}$ at the bank	$+\frac{K}{(1+r)^T}$	$-K$
Diff. at the bank	$-R_{K,r}$	$+R_{K,r}$
Total	0	$R_{K,r} + S_T - K + (K - S_T)^+ - (S_T - K)^+$

(*Recall that "Borrow $\frac{K}{(1+r)^T}$ at the bank" it is equivalent to short sale $\frac{K}{(1+r)^T}$ non risky asset with interest r.*)

We have[8] $S_T - K + (K - S_T)^+ - (S_T - K)^+ = 0$ and since $R_{K,r} > 0$ we get an arbitrage, hence we deduce that under **(AoA)**, we have $\pi_0 - \nu_0 - S_0 + \frac{K}{(1+r)^T} \leq 0$.

We prove that $\pi_0 - \nu_0 - S_0 + \frac{K}{(1+r)^T} < 0$ leads to an arbitrage similarly so that (5.6) holds.

Second proof: by using Corollary 5.2.6. Consider a portfolio V short in a put option and long in a call option. Consider a second portfolio $\tilde{V}$ with $n_S = 1$ and $n_M = -\frac{K}{(1+r)^T}$. At time T we have:

$$V_T = (S_T - K)^+ - (K - S_T)^+ = S_T - K = \tilde{V}_T.$$

Hence the values at time 0 coincides, by Corollary 5.2.6, so that Relation (5.6) holds. □

5.4 Challenge #5 : pricing in the binomial model

5.4.1 Additional results on the call

We consider a call option with strike K, underlying asset S, interest rate r and maturity T.

1. Prove that under the no-arbitrage assumption, the price π_0 of a call option satisfies

$$\left(S_0 - \frac{K}{(1+r)^T}\right)^+ \leq \pi_0 \leq S_0.$$

   Hint: we recall that $y^+ \leq x$ if $0 \leq x$ and $y \leq x$.

2. What can you say about the price of a call option and the price of a futures? Explain this result.

[8] Note that $x^+ - (-x)^+ = x$, by distinguishing the cases $x \geq 0$ and $x \leq 0$.

5.4.2 Arbitrage detection

A trainee presents her/his new program to compute European call and put prices. We assume that the considered asset does not provide dividends. Assume that the price of the asset is $S_0 = 100€$ today with $r = 1\%$, $T = 1$ (1 year). The program gives the following outputs for call and put prices with different strikes

Value Today	Call $K = 95€$	Call $K = 100€$	Call $K = 105€$	Put $K = 95€$	Put $K = 100€$	Put $K = 105€$
Price	11 €	9.5 €	7.27 €	6.95 €	10.51 €	9.83€

If these prices were market prices, detect some abnormalities in term of arbitrage theory in this tabular. Give some possible arbitrage strategies by using only call with different strikes, only put with different strikes and call/put with same strike.

5.4.3 Binomial model (fundamental exercise)

Pricing and hedging of a risky financial product in a one time period model.

In this section, we compute the price of a call option with strike K and maturity T in the binomial model or Cox-Ross-Rubinstein model (1979). In this model the risky asset can go up or go down on one period. In this problem, we aim at computing the non-arbitrage price of a general risky financial product together with the hedging strategy n_S used by the seller to cover the associated payoff.
Let S_0 be the price of the risky asset at time 0. At time 1, the price of this asset is denoted by S_1 and is either uS_0 with probability $p \in (0,1)$ or dS_0 with $u > d > 0$. We also assume that there exists a non-risky asset with compound interest rate $r > 0$ (meaning that 1 risky asset at time $t = 0$ gives $1 + r$ euros at time $t = 1$). We consider a contract giving to his/her owner the random payoff G at time $T = 1$. If $S_1 = uS_0$ the payoff G takes the value G_u, if $S_1 = dS_0$, G takes the value G_d.

We thus aim at finding the non-arbitrage price together with **the hedging strategy,** *i.e.* the number of risky asset S denoted by n_S together with the amount of cash n_M (number of non risky asset), used by the seller to replicated the payoff G at time $T = 1$.

1. Prove that there is no-arbitrage strategy if and only if $0 < d < 1 + r < u$.

2. We denote by $\widetilde{\pi}_0$ the initial amount of money needed at time 0 to replicate exactly the payoff G at time $T = 1$ and (n_S, n_M) the hedging strategy associated with it. Give the relation satisfied by $\tilde{\pi}_0, n_S$ and n_M and deduce that value of n_M with respect to $\tilde{\pi}_0, n_S$ and S_0.

3. Deduce the values of $\widetilde{\pi}_0$ and n_S as functions of r, u, d, S_0, G_u and G_d.

4. Prove that the non-arbitrage price π_0 is exactly $\widetilde{\pi}_0$.

   The no-arbitrage price of the contract considered coincides with its replication price.

   n_S is also called **the "Delta"** *(guess why...) of the derivative product.*

5. By setting $q := \frac{1+r-d}{u-d}$, deduce that there exists a probability $\mathbb{Q}$ such that
$$\pi_0 = \mathbb{E}^{\mathbb{Q}}\left[\frac{G}{1+r}\right],$$
where $\mathbb{E}^{\mathbb{Q}}$ is the expectation associated with the probability $\mathbb{Q}$.

   We call this probability measure the risk neutral probability. Note the historical probability given by p plays no role to get the price or the hedging strategy!

Application to European options

A European option is a contract with associated (random) payoff $G = (S_1 - K)^+$ at time $T = 1$ with fixed strike $K > 0$. By applying the previous question, gives the general formula for the price and the "Delta" (n_S in the hedging strategy) of such contract with respect to r, u, d, S_0, K only.

5.5 Appendix: To go further...

We provide in this section some fundamental macro-properties under no-arbitrage.

We aim at proving some easy parts of the first and second theorems of finance. The first theorem gives a link between the existence of a very particular probability (called the risk neutral probability) and a market without arbitrage opportunities. The second theorem gives a link between the uniqueness of such probability measure and the hedging of any asset on a market.

In all this section, we consider a market with a non-risky asset with value $S_0^0 = 1$ at time $t = 0$ and $S_0^0 = (1 + r)$ at time $t = 1$ where $r > 0$ is the interest rate. We also consider a risky asset with price $S_0 > 0$ at times 0 and S_1 at time $t = 1$ where S_1 is a (discrete) random variable defined on some probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ and taking countable values in $\mathbb{N}$.

Definition 5.5.1. *We say that $\mathbb{Q}$ is a risk neutral probability if $\mathbb{E}^{\mathbb{Q}}[\frac{S_1}{1+r}] = S_0$.*

In the following we consider V_0 the initial value of a portfolio and V_1 be the value of this portfolio at time $t = 1$.

Preliminary question. Recall the expression of V_0 and V_1 in terms of n_M the number of non-risky assets owned, n_S the number of risky assets owned, S_0, S_1 and r.

Arbitrage and existence of a risk neutral probability.

1. Show that $\mathbb{Q}$ is a risk neutral probability if and only if $\mathbb{E}^{\mathbb{Q}}[\frac{V_1}{1+r}] = V_0$.
2. Deduce that if there exists a risk neutral probability measure then the market is arbitrage-free (*i.e.* there does not exist arbitrage opportunity).

   The other part is also true (but much more difficult to prove) so that

   Theorem 5.5.2 (First theorem of finance (admitted))**.** *There exists a risk neutral probability if and only if the market is without any arbitrage opportunity.*
3. Recover the non-arbitrage condition of the binomial model with this theorem.

Hedging and uniqueness of a risk neutral probability and application to the trinomial model.

1. We now a market such that any random variable (payoff) G at time $t = 1$ can be replicated (we say that the market is **complete**). Prove that in this case and if the market is without any arbitrage, the risk neutral probability is unique. *Hint: consider the payoff* $\mathbf{1}_A$ *where* $A \subset \Omega$*, show that* $\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_A] = \mathbb{Q}(A)$ *for any probability measure.*

2. We consider a trinomial model which extended the binomial model by adding three possible values of S_1: $S_1 \in \{dS_0, mS_0, uS_0\}$ with $0 < d < m < u$. Gives a condition ensuring that this market is without any arbitrage opportunities but not complete under this condition.

Chapter 6

Almost sure convergence and population genetics

The objective is to introduce a dynamical random model coming from population genetics and to describe its asymptotic behavior. This course is the occasion to discover through an example a simple Markov chain in finite state space, a.s. convergence of a sequence of random variable and an interesting genetic phenomenom (the genetic drift !). TD : fixation probability and speed of convergence in Wright-Fisher model

6.1 Introduction : the Wright Fisher model

The Wright Fisher model, named after Sewall Wright and Ronald Fisher, is one of the most popular in population genetics. We first make a simplification by assuming that the population size is fixed and equal to N. Moreover the random transmission procedure is the following : each individual drops independently at random its allele from the previous generation by choosing uniformly at random its parent and getting the same allele. Finally, we consider only two alleles A, a and this model includes no selection, no mutation, no migration, non-overlapping generation times.

We let X_n be the random variable counting the number of alleles A in generation n and fix $X_0 = x_0 \in \{1, \dots N\}$ the initial number of alleles A. Thus there are $N - X_n$ alleles a in generation N.

Conditionally on $X_n = k$, the probability for each individual in generation $n+1$ to get allele A is equal to the proportion of allele A in generation n, i.e. k/N. Thus, we introduce $B_{i,n}^{(x)}$ independent and identically distributed Bernoulli random variables with parameter x. The random variable $B_{i,n}^{(k/N)}$ yields 1 if the ith individual of generation n gets allele A, which occurs with probability k/N. Otherwise, $B_i^{(k/N)}$ is equal to 0 and the term has no contribution to the amount of A. We count now the number of alleles A in the next generation. Conditionally on $X_n = k$, we have

$$X_{n+1} = \sum_{i=1}^{N} B_{i,n}^{(k/N)}$$

As a consequence,

$$X_{n+1} \stackrel{law}{=} \text{Binomial}(N, k/N),$$

and

$$\mathbb{P}(X_{n+1} = j | X_n = k) = \binom{N}{j} \left(\frac{k}{N}\right)^j \left(1 - \frac{k}{N}\right)^{N-j}.$$

6.2 First computations and properties

6.2.1 Mean behavior

We first compute the conditional expectation. Writing $B_i = B_{i,n}^{(k/N)}$ for convenience,

$$\mathbb{E}(X_{n+1}|X_n = k) = \mathbb{E}\left(\sum_{i=1}^{N} B_i\right) = \sum_{i=1}^{N} \mathbb{E}\left(B_i\right) = N\mathbb{E}(B_1) = N.k/N = k$$

One may also get this identity by using the definition of conditional expectation as sum of conditional probability :

$$\mathbb{E}(X_{n+1}|X_n = k) = \sum_{j=0}^{N} j\mathbb{P}(X_{n+1} = j|X_n = k) = \sum_{j=0}^{N} j\binom{N}{j} \left(\frac{k}{N}\right)^j \left(1 - \frac{k}{N}\right)^{N-j}.$$

Then

$$\begin{aligned}
\mathbb{E}(X_{n+1}) &= \sum_{i=0}^{N} i.\mathbb{P}(X_{n+1} = i) \\
&= \sum_{i=0}^{N} i. \sum_{k=0}^{N} \mathbb{P}(X_{n+1} = i \mid X_n = k)\mathbb{P}(X_n = k) \\
&= \sum_{k=0}^{N} \mathbb{P}(X_n = k) \sum_{i=0}^{N} i.\mathbb{P}(X_{n+1} = i \mid X_n = k) \\
&= \sum_{k=0}^{N} \mathbb{P}(X_n = k)\mathbb{E}(X_{n+1}|X_n = k) \\
&= \sum_{k=0}^{N} \mathbb{P}(X_n = k)k = \mathbb{E}(X_n)
\end{aligned}$$

Thus

$$\mathbb{E}(X_n) = \mathbb{E}(X_0) = x_0$$

and in average, the number of alleles A is conserved.

6.2.2 Variance

By expansion and writing $B_i = B_i^{(k/N)}$ for convenience,

$$\mathbb{E}(X_{n+1}^2|X_n = k) = \mathbb{E}\left(\left(\sum_{i=1}^{N} B_i\right)^2\right) = \sum_{i=1}^{N} \mathbb{E}(B_i^2) + \sum_{i \neq j} \mathbb{E}(B_i B_j).$$

and using that $\mathbb{E}(B_i B_j) = \mathbb{E}(B_i)\mathbb{E}(B_j)$ for $i \neq j$ by independence, we get

$$\mathbb{E}(X_{n+1}^2|X_n = k) = Nk/N + N(N-1)(k/N)^2 = k + k^2(N-1)/N.$$

Finally

$$\begin{aligned}
\mathbb{E}(X_{n+1}^2) &= \sum_{k=0}^{N} \mathbb{P}(X_n = k)\mathbb{E}(X_{n+1}^2|X_n = k) \\
&= \sum_{k=0}^{N} k\mathbb{P}(X_n = k) + \frac{N-1}{N}\sum_{k=0}^{N} k^2\mathbb{P}(X_n = k) \\
&= \mathbb{E}(X_n) + \mathbb{E}(X_n^2)(N-1)/N
\end{aligned}$$

We get an affine sequence !!! Setting

$$u_n = \mathbb{E}(X_n^2), \qquad \alpha = (N-1)/N = 1 - 1/N$$

and recalling $\mathbb{E}(X_n) = x_0$, we have

$$u_{n+1} = x_0 + \alpha u_n$$

Recall now the first course and set $z = x_0 + \alpha z$, so that $u_n - z$ is a geometric sequence with ratio α, $u_n = z + \alpha^n(x_0^2 - z)$, i.e. $z = Nx_0$ and

$$\mathbb{E}(X_n^2) = Nx_0 + (1-1/N)^n(x_0^2 - Nx_0), \; Var(X_n) = Nx_0 - x_0^2 + (1-1/N)^n(x_0^2 - Nx_0).$$

Thus the variance is not disappearing as n goes to infinity, it tends to $Nx_0 - x_0^2$ which is non zero (if x_0 is neither 0 nor N).

6.2.3 Absorption

The points 0 and N are absorbing, i.e. once X is equal to 0 or N, its keeps this value :

$$\mathbb{P}\left(\forall k \geq 0, X_{n+k} = N|X_n = N\right) = 1 = \mathbb{P}\left(\forall k \geq 0, X_{n+k} = 0|X_n = 0\right) = 1.$$

The other points are non absorbing : once we are there, we can move.

6.2.4 Law

Here again (see the other sessions), the fact that

$$\mathbb{P}(X_{n+1} = j \mid X_n = k) = P_{k,j} = \binom{N}{j}\left(\frac{k}{N}\right)^j\left(1 - \frac{k}{N}\right)^{N-j}$$

and we can prove inductively that

$$(\mathbb{P}(X_n = j \mid X_0 = x_0) : j = 0, \ldots, N) = \pi_{x_0}P^n$$

where $\pi_{x_0} = (0, \ldots, 0, 1, 0, \ldots 0)$ is the x_0th vector of the basis.
The problem is that we are not able to reduce the matrix P to capture the moments ...

6.3 Challenge #6: genetic drift, probability and speed of fixation

Recall that X_n be the random variable counting the number of alleles A in generation n and fix $X_0 = x_0 \in \{1, \dots N\}$ the initial number of alleles A. Conditionally on $X_n = k$, we have

$$X_{n+1} = \sum_{i=1}^{N} B_i^{(k/N)},$$

where $(B_i^{(x)} : i = 1, \dots, N)$ are independent and identically distributed Bernoulli random variable, with parameter x.

- Prove that there exists $\alpha > 0$ such that for any $k \in \{1, \dots, N\}$ and $n \geq 0$

$$\mathbb{P}(X_{n+1} = N | X_n = k) \geq \alpha.$$

- Introduce the absorption events

$$A_j = \{\exists\, n \text{ such that } \forall k \geq 0 : X_{n+k} = j\}$$

  and prove that there is almost sure absorption in 0 or N in the sense that

$$\mathbb{P}(A_0 \cup A_N) = 1.$$

- From the previous question, we can define

$$X_\infty = \lim_{n \to \infty} X_n$$

  with probability one. Thus, $X_\infty = 0$ on the event A_0 and $X_\infty = N$ on the event A_N. Compute the fixation probability of allele A

$$p_{fix} = \mathbb{P}(X_\infty = N).$$

- *Optional : obtain a geometric speed of convergence to fixation in probaility, i.e. prove that*

$$|\mathbb{P}(X_n = N) - p_{fix}| \leq C.\beta^n,$$

  for any $n \geq 0$, where $C \geq 0$ and $\beta \in [0, 1)$ have to be found.

- *Optional : propose now a model with selection, in the sense that we want to take into accound that allele A gives some advantage for survival and transmission to individuals, compared to allele a.*

Chapter 7

Reversibility in law and Ehrenfest model

The objective of this course is to study a stochastic model for gas molecules which exhibit an interesting behavior. Some reversibility property is holding for fixed number of molecules. But two variables and scales are intertwined and taking limits let different behavior appear depending on the relative sizes of variables. In this lecture, modeling is again leading us to study the long time behavior of a sequence of random variable. A new behavior will be observed...

7.1 Introduction : Ehrenfest model

In 1907, Paul and Tatyana Ehrenfest introduced the model that bears their name today. It is a simple probabilistic model, which makes it possible to describe the evolution of the pressure of a gas, an irreversible macroscopic evolution over time, by the reversible microscopic evolution of the molecules composing this gas. They propose this model in response to the strong criticism of Boltzmann's kinetic gas theory, which set out physical equations that were unchanged if we reverse the course of time, when they had to describe an irreversible phenomenon. This simple theoretical model shows how the laws of probability can produce an average trend towards equilibrium, while the model's behaviour is reversible over time and each of its states can be recurrent.

Let $N \in \mathbb{N}$. We consider $2N$ balls (gas molecules) spread in 2 boxes.
The model is dynamic and recursively defined as follows. At each step, one ball is chosen uniformly at random and is removed from its box and put in the other box.
We write

$$X_n \in \{0, \dots, 2N\}$$

the number of balls in the first box at step n. By construction of the model,

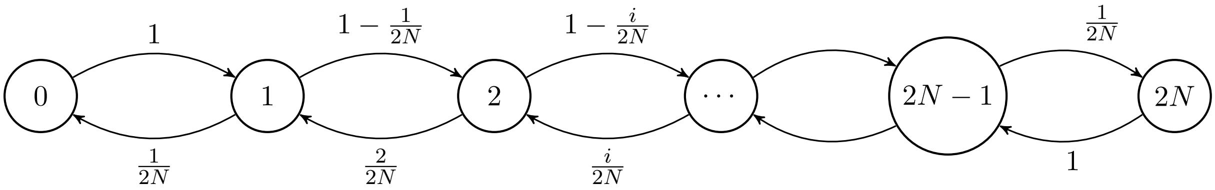
$$\mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i, \dots X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_1 = j | X_0 = i)$$

for $0 \leq i, j \leq 2N$ and $n \geq 0$. Thus defining $P = (P_{ij})_{0 \leq i,j \leq 2N}$ as

$$P_{ij} = \begin{cases} 1 - \frac{i}{2N} \ \text{ if } j = i+1 \\ \frac{i}{2N} \ \text{ if } j = i-1 \\ 0 \text{ otherwise} \end{cases}$$

we get

$$P_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i).$$



7.2 First properties : law, mean and variance

We can show inductively that

Proposition 7.2.1. *For any $n \in \mathbb{N}$ and $i, j \in \{0, \ldots, 2N\}$,*

$$\mathbb{P}(X_n = j \,|X_0 = i) = (P^n)_{ij}$$

In particular, for any i, j, there exists $n \in \mathbb{N}$ such that $\mathbb{P}(X_n = j \,|X_0 = i) > 0$. It is an irreducibility property of the model : for two given states, we can go from one to the other following a path and this path has a positive probability.

Proof. For $n = 0$ (initialization),

$$\mathbb{P}(X_0 = i \,|X_0 = j) = \delta_{i,j} = (Id)_{ij}$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise, and $Id = P^0$ is the identity matrix.
Let us proceed with the induction and assume that for a given n, $\mathbb{P}(X_n = j \,|X_0 = i) = (P^n)_{ij}$.
We compute

$$\begin{aligned} \mathbb{P}(X_{n+1} = j \mid X_0 = i) &= \sum_{k=0}^{2N} \mathbb{P}(X_{n+1} = j, X_n = k \mid X_0 = i) \\ &= \sum_{k=0}^{2N} \mathbb{P}(X_n = k \mid X_0 = i)\mathbb{P}(X_{n+1} = j \mid X_n = k, X_0 = i) \end{aligned}$$

Moreover by construction of the model $\mathbb{P}(X_{n+1} = j \mid X_n = k, X_0 = i) = \mathbb{P}(X_{n+1} = j \mid X_n = k) = P_{kj}$ and and by induction hypothesis $\mathbb{P}(X_n = k \mid X_0 = i) = (P^n)_{ik}$. We get

$$\mathbb{P}(X_{n+1} = j \mid X_0 = i) = \sum_{k=0}^{2N} (P^n)_{ik} P_{kj} = (P^{n+1})_{ij}$$

which ends the proof. $\square$

7.2.1 Mean behavior

Following the previous lecture, we try to compute the expectation of X_n and use that $\mathbb{E}(X_{n+1}|X_n = i) = (i+1)\mathbb{P}(X_{n+1} = i+1|X_n = i) + (i-1)\mathbb{P}(X_{n+1} = i-1|X_n = i)$. We get

$$\mathbb{E}(X_{n+1}|X_n = i) = (i+1)P_{i\,i+1} + (i-1)P_{i\,i-1} = (i+1)(1-i/2N) + (i-1)i/2N = \alpha i + 1$$

where $\alpha = 1 - 1/N$. Then

$$\mathbb{E}(X_{n+1}|X_n = i) = \alpha i + 1$$

We obtain (affine sequence)

$$\mathbb{E}(X_n) = N + (\mathbb{E}(X_0) - N)\alpha^n$$

and $(\mathbb{E}(X_n))_n$ converges to N as $n \to \infty$. This corresponds to a "monotone and irreversible mean behavior" and yields a balanced limiting value. This latter fact is coherent with the symmetry in the dynamic of the model.

7.2.2 Variance

Similarly

$$\mathbb{E}(X_{n+1}^2|X_n = i) = (i+1)^2 P_{i\,i+1} + (i-1)^2 P_{i\,i-1} = i^2(1-2/N) + 2i + 1$$

and writing $\beta = 1 - 2/N$,

$$\mathbb{E}(X_{n+1}^2) = \mathbb{E}(X_n^2)\beta + 2\mathbb{E}(X_n) + 1$$

Setting

$$u_n = \mathbb{E}(X_n), \qquad v_n = \mathbb{E}(X_n^2) - \mathbb{E}(X_n)^2 = Var(X_n)$$

and rescaling the value of u_n above,

$$v_{n+1} = \beta v_n + 1 - \alpha^{2n}(\mathbb{E}(X_0)/N - 1).$$

Inductively we can then show that

$$v_n = \frac{N}{2} + (v_0 - \frac{N}{2})\beta^n + (\mathbb{E}(X_0) - N)^2(\beta^n - \alpha^{2n}).$$

As a consequence, the variance v_n goes to $N/2$, which means that stochasticity does not vanish as time n goes to infinity.

7.3 Reversibility and stationary distribution

The following probability law is the binomial law with parameters $(2N, 1/2)$:

$$\pi_i = \frac{\binom{2N}{i}}{2^{2N}} \qquad (0 \le i \le 2N).$$

It satisfies the following "reversibility identity" :

$$\pi_i P_{ij} = \pi_j P_{ji}.$$

Indeed, one can distinguish the cases $j = i+1, j = i-1, j \notin \{i-1, i+1\}$ and for instance,

$$\pi_i P_{i\,i-1} = \frac{\binom{2N}{i}}{2^{2N}} \frac{i}{2N} = \frac{(2N)!i}{i!(2N-i)!2^{2N}2N}$$

$$\pi_{i-1} P_{i-1\,i} = \frac{\binom{2N}{i-1}}{2^{2N}} (1 - \frac{i-1}{2N}) = \frac{(2N)!(2N-i+1)}{(i-1)!(2N-i+1)!2^{2N}2N}.$$

This reversibility property ensures that

$$(\pi P)_i = \sum_j \pi_j P_{ji} = \pi_i \sum_j P_{ij}.$$

and thus π is an invariant law :

$$\pi P = \pi.$$

7.4 Challenge #7: Relaxation to equilbrium

We consider $2N$ balls and write X_n^N for the number of balls in the first box at the step n. We recall that

$$\mathbb{P}(X_{n+1}^N = j | X_n^N = i) = \mathbb{P}(X_1^N = j | X_0^N = i) = p_{ij}^N$$

where

$$p_{ij}^N = \begin{cases} 1 - \frac{i}{2N} \ \text{ if } j = i+1 \\ \frac{i}{2N} \ \text{ if } j = i-1 \\ 0 \text{ otherwise} \end{cases}$$

We start with all the balls in the first box ($X_0^N = 2N$) and define

$$Y_k^N = \frac{X_{2Nk}^N}{2N}$$

for every $k \geq 0$.

- Prove that for any $k > 0$,

$$\mathbb{E}(Y_k^N) \longrightarrow \frac{1}{2} + \frac{1}{2} e^{-2k}$$

  as $N \to \infty$

- Prove that for any $k > 0$,

$$Var(Y_k^N) \overset{N \to \infty}{\longrightarrow} 0,$$

  where $Var(.)$ is the variance.

- Prove that for any $\varepsilon > 0$ and any random variable Z such that $\mathbb{E}(Z^2) < \infty$,

$$\mathbb{P}(|Z| \geq \varepsilon) \leq \frac{\mathbb{E}(Z^2)}{\varepsilon^2}$$

- Prove for any $k \geq 0$ and $\varepsilon > 0$ the convergence

$$\mathbb{P}(\left|Y_k^N - \left(\frac{1}{2} + \frac{1}{2} e^{-2k}\right)\right| \geq \varepsilon) \longrightarrow 0$$

  as $N \to \infty$.

- Interpret the result. *You may plot a typical trajectory of the stochastic sequence $(Y_k^N : k \geq 0)$ for a fixed (large) N.*