Stability of linear dynamical systems and application to Lotka-Volterra equations and epidemiology

Benoît Müller

Bachelor project in mathematics
École Polytechnique Fédérale de Lausanne

Supervisor: Joachim Krieger
Co-supervisor: Shengquan Xiang

Spring 2021

Abstract

The goal of this project is to study the stability of some linear and nonlinear autonomous ordinary differential systems including the "predator–prey" based Lotka-Volterra models, and to further find applications to epidemiology such as the Covid-19 evolution.

The first chapter is devoted to the investigation of the stability of linear autonomous ODE systems, where a complete classification is given thanks to spectrum information. We refine these results to give a stability decomposition of the solution space, and show that in general, some stability properties are preserved to related nonlinear systems thanks to linearization arguments.

Then, we turn in the second chapter to a more specific form of ODE systems, say Lotka-Volterra models. Instead of considering classical Lotka-Volterra systems, we look at a more general form. Because at some special critical points the linearization stability theory is not always applicable, we introduced the Lyapunov approach, which can even lead to global stability results.

Finally, the third chapter is an application of the preceding chapters, where we try to model Covid-19 systems concerning patients and vaccinations, and to predict the evolution of the pandemic via stability investigation.

Contents

Introduction

In the following paragraphs, we detail the composition of this project and develop the ideas of the main results.

In the first chapter we introduce linear dynamical systems. They are autonomous ordinary differential equations, where the derivative of a solution is linear with respect to the solution itself. We write it as a matrix product. Solutions always exist on the all real space and are unique for a same initialisation.

We describe the space of solutions by showing it has the same dimension as the image space because the linear dependence of solutions is the same as their linear dependence at any fixed time. We can compute the solution starting at any point by multiply the initial position by a matrix called fundamental matrix. Its value can be computed as an infinite sum. We allow our self complex solutions, seeing that their real and imaginary parts are real solutions too, and we find the general form of the solutions, by treating particular solutions that start on the generalised eigenvectors of the system matrix. They are called eigensolutions and are the product of two factors: an exponential and a polynomial. The eigenvalue appears in the argument exponential, and the coefficients of the polynomial are generalised eigenvectors of the same eigenspace. A basis of generalised eigenvectors form a basis of eigensolutions for the solution space that we can compute and understand.

Then we introduce the stability of a fixed point. We say it is Lyapunov-stable when all solutions that start near enough the fixed point never leave an arbitrary neighbourhood. It is asymptotically stable if in addition to that, all solutions that start near enough the fixed point, converge to it. From the understanding of the solutions, we deduce necessary and sufficient conditions on the matrix that define the equation, to have stability on the origin. All conditions relate to the sign of the real part of its eigenvalues, which appear in the argument of the exponential factor of the eigensolutions. The linear system is Lyapunov-stable if and only if all eigenvalues have a non-positive real part, and a negative one when the eigenvalue is defective. It is asymptotically stable if and only if all eigenvalues have a negative part, and we call such eigenvalues stable. We give a classification of linear dynamical systems, that describe the kind of stability they have. They take into account the range of stability of the eigenvalues.

We present a direct sum that decompose the space of solutions in three subspaces: the generalised eigenspaces associated to eigenvalues of real parts that are respectively negative, null, and positive. These three spaces, called respectively stable, center and unstable eigenspace have a second description that only take the asymptotic comportment of solutions when t tends to plus or minus infinity. Two systems, whose center eigenspace are trivial, are actually linked by a homeomorphism between their solutions. We show a direct application of our understanding of the linear case, the linearization theorem. It says that if the linearization of the system in a fixed point is asymptotically stable, then the nonlinear system is asymptotically stable too.

We motivate the study of a particular nonlinear dynamical system on the plane, called Lotka-Volterra equations, by showing how it is linked to population dynamics. The functions that are set in an equation are the size of two population of prey and predators. It is based on the concept of logistic growth and the understanding of the growth rate, which is the derivative divided by the function itself. It represent the mean of change an individual add to the population by unit of time. This rate of growth is linear in the Lotka-Volterra equation. We present a modification that change the affect between the species. The idea is to consider in the equation not just the effective size of population, but a function of it. This is useful to have a better description of the affect each species have on the others. We consider two situations. In the first system, a function does not appears in its own growth rate, and in the second it does.

We investigate the nature of the solutions and their stability. For the first system, the linearization in a particular fixed point gives an elliptic linear system, a system with cyclic solutions. We search then for a constant quantity that is conserved along time, and this gives a implicit equation for the trajectory of the solutions. From this, we conclude that the solutions are indeed cyclic, they orbit around the fixed point. In the second system, the linearization in a particular fixed point gives an stable linear system, so we know that the fixed point is asymptotically stable. Motivated by this result, and inspired by the conserved quantity of the first system, we develop the theory of Lyapunov functions. They are function that represent the energy of a position, and such that all solutions loose energy along time. When such a function exist and has a minimum in a fixed point, we can obtain asymptotic stability without condition on the starting point of the solutions, they all converge to the fixed point. Taking a similar expression of the constant quantity of the first system, we find a Lyapunov function. From this, we conclude that all the solutions turn in a spiral around the fixed point and converge to it, we have global asymptotic stability.

From the study of this problem, and the basic strategic idea of this modelisation, we present adaptations to epidemiology dynamics. We investigate the modelisation of a population that is confronted to a virus. We use the Lotka-Volterra system to describe the population of vulnerable people that have not been infected, the population of infected people, and the population of people that have recovered after infection. We see that this model has no isolated fixed point so fixed points are not asymptotically stable. We derive again constant quantities along time, that give implicit formulas for the trajectories of solutions and the limits of them. If the initial number of vulnerable people is bigger than a certain threshold, the number of infected people increase to a maximum located in this threshold, and eventually tends to zero. If the initial number of vulnerable people is smaller than this threshold, the number of infected people directly tends to zero. We give some description of the asymptotic comportment and the reminded size of groups. In a second modelisation, we introduce the vaccine to the population and separate each group in two, regarding if they are vaccinated or not. The analysis of the problem can be done with the same tools as the precedent, with a bit more work. We obtain a pair of equations that gives some explicit relations between the functions. In particular, we have a bijective relation between the vulnerable groups and show the affect of the vaccine. Both models aren't asymptotically stable but are L-stables in their domain.

Chapter 1

Linear dynamical systems and stability

In this chapter, we study explicit autonomous ordinary differential equations of first order that are linear, namely equations of the form

$$\dot{\mathbf{x}} = A\mathbf{x}$$

where A is a $n \times n$ matrix and $\mathbf{x}$ is a differentiable function from a real interval to $\mathbb{R}^n$.

Our main goal is to investigate stability. For this, we present a brief description of the solution space and explicit forms of its elements. This first section will help us for the second, in the study of the stability.

1.1 Solution space

First of all, using results by Cauchy and Picard, and the particular smoothness of $\mathbf{x} \to A\mathbf{x}$, we know that the solutions locally exist, and that they are unique for a same initialisation. Since

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\|^2 = 2\langle \mathbf{x}, \dot{\mathbf{x}} \rangle = 2\langle \mathbf{x}, A\mathbf{x} \rangle \leq 2\|A\|\|\mathbf{x}\|^2,$$

we have

$$\|\mathbf{x}(t)\|^2 \leq \|\mathbf{x}(0)\|^2 e^{2\|A\|t}$$

by Grönwall and solutions cannot explode, they are defined on all $\mathbb{R}$. We note that since the derivative of a solution is just itself multiplied by a matrix, it is derivable too, and by induction $\mathbf{x}$ is C^∞ with $\mathbf{x}^{(k)} = A^k\mathbf{x}$.

The equation is autonomous, so any solution can be reparametrised by a translation of the time: If $\dot{\mathbf{x}}(t) = A\mathbf{x}(t)$, then $\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{x}(t+\tau)) = A\mathbf{x}(t+\tau)$. That is why we will generally initialise them at $t = 0$ without loss of generality. Solutions are described by a *flow* $\phi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ where $t \mapsto \phi(\mathbf{x}_0, t)$ is the solution starting in $\mathbf{x}_0$ at $t = 0$. However, existence theorems does not tell us much about their actual form. We present then their existence in this particular case. First, we observe that the set of solution is a vector subspace of C^∞. Indeed for all solutions $\mathbf{x}$, $\mathbf{y}$ and a scalar α,

$$(\alpha\mathbf{x} + \mathbf{y})' = \alpha\dot{\mathbf{x}} + \dot{\mathbf{y}} = \alpha A\mathbf{x} + A\mathbf{y} = A(\alpha\mathbf{x} + \mathbf{y})$$

and the identical null function is trivially in the space. This motivate us to find a basis of this subspace and understand how to construct it.

Definition 1. A collection of k solutions $\mathbf{x}_1, \dots, \mathbf{x}_k$ are said to be *linearly independent* or *independent*, if $\mathbf{x}_1(t), \dots, \mathbf{x}_k(t)$ are linearly independents for all t.

Lemma 1. *For any scalar τ, a collection $\mathbf{x}_1, \dots, \mathbf{x}_k$ of solutions are linearly independents if and only if $\mathbf{x}_1(\tau), \dots, \mathbf{x}_k(\tau)$ are linearly independents.*

Proof. If the solutions are independent we have the result by definition. In the other way, by uniqueness of the solutions up to initialisation, we have that if for time τ, the positions are not independent, then we must have non all null scalars $\alpha_1, \dots, \alpha_k$ such that $\alpha_1\mathbf{x}_1(\tau) + \dots + \alpha_k\mathbf{x}_k(\tau) = 0$. But then 0 is a trivial solution starting there and actually $\alpha_1\mathbf{x}_1(t) + \dots + \alpha_k\mathbf{x}_k(t) = 0$ for all t. □

This lemma prove that the space of solution has the same dimension as the space $\mathbb{R}^n$, *i.e.* n. In order to manipulate the solutions, we put them in a matrix, like $(\mathbf{x}_1, \dots, \mathbf{x}_k)$. In this form we see that we can actually extend the equation to matrix entries:

$$\dot{X} = AX$$

and since

$$\dot{X} = (\dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_k) \text{ and } AX = (A\mathbf{x}_1, \dots, A\mathbf{x}_k),$$

a matrix X is a solution if and only if its columns are vector solutions. We see that for a vector $v \in \mathbb{R}^k$, Xv is a vector solution $(Xv)' = \dot{X}v = AXv$. This solution ca be written as $Xv = v_1\mathbf{x}_1 + \dots + v_k\mathbf{x}_k$ and is then a linear combination of the column solutions of X. Thus, a matrix solution allow us to easily write new solutions as linear combinaison of a collection of solutions. As the dimension of the space of solutions is n, exactly n linear independents solutions will be enough to construct all solutions as a product of the matrix solution constructed with them.

Definition 2. If a matrix solution $M = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ is square (k=n) and of full rank, then it is called a *fundamental matrix solution* and $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ a *fundamental set of solutions.*

With such a M, $M\mathbb{R}^n$ is the all space of solutions. At time $t = 0$, $M(0)v = v_1\mathbf{x}_1(0) + \dots + v_k\mathbf{x}_k(0)$ meaning that with the condition $\mathbf{x}(0) = x_0$, v must be chosen to be the vector of coordinates of x_0 in the basis of the columns of M.

We can now look at the forms the solutions can take. In one dimension, the problem is $\dot{x} = ax$ and if a is non null, non trivial solutions satisfy $a = (\log|x|)'$ which give $\log|x(t)| = at + c$ and $x(t) = \pm e^{at+c} = x_0e^{at}$. We check easily that e^{at} is indeed a solution. The n-dimensional case is more complicated. We try to put the equation in its integral form and derive a property by recurrence (supposing $X_0 = I$ for simplicity):

$$\begin{aligned} X(t) &= I + \int_0^t AX(s)\mathrm{d}s = I + \int_0^t A\Big(I + \int_0^{s_1} AX(s_2)\mathrm{d}s_2\Big)\mathrm{d}s_1 = I + tA + \int_0^t \int_0^{s_1} A^2X(s_2)\mathrm{d}s_2\mathrm{d}s_1 \\ &= I + tA + \frac{1}{2}(tA)^2 + \int_0^t \int_0^{s_1} \int_0^{s_3} A^3X(s_3)\mathrm{d}s_2\mathrm{d}s_1. \end{aligned}$$

We see the Taylor expansion of exp appear, as a generalisation to matrices, where we define $A^0 = I$ for any matrix A. This motivate the following definition:

Definition 3. If it converges, we define the *exponential* e^A of a square matrix A and the underlying function exp as the infinite sum

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!}A^n.$$

Lemma 2. *The exponential of a matrix always converges.*

Proof. We use $\|.\|$ for the operator norm on matrices, keeping in mind that all matricies norms are equivalent. Now we have by basic properties of this norm that

$$\sum_{n=0}^{N} \|\frac{1}{n!}A^n\| \leq \sum_{n=0}^{N} \frac{1}{n!}\|A^n\| \leq \sum_{n=0}^{N} \frac{1}{n!}\|A\|^n,$$

and this is the Taylor finite expansion of $\|A\|$, which converge for all value of $\|A\|$. As a result the sum is absolutely convergent for the operator norm, and then must converge. □

Theorem 1. *For a matrix A and a scalar t, the quantity e^{At} is differentiable with respect to t and has derivative* $\mathrm{d}/\mathrm{d}t(e^{At}) = Ae^{At}$.

Proof. Each coordinate of the series is actually a convergent Taylor series in t and is then analytic. The theory of analytic functions tell us that they are C^∞ and that we can derive term by term. We get

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{n=0}^{\infty}\frac{1}{n!}t^n A^n = \sum_{n=1}^{\infty}\frac{1}{(n-1)!}t^{n-1}A^n = A\sum_{n=0}^{\infty}\frac{1}{n!}(tA)^n = Ae^{At}$$

□

All of this tell us that as we expected, e^{tA} is a matrix solution to the linear differential equation.

Corollary 1. *The matrix e^{tA} is a fundamental solution and each solution write* $\mathbf{x}(t) = e^{tA}\mathbf{x}(0)$

Proof. We evaluate it in $t = 0$ by direct calculation since the sum become finite:

$$e^{0A} = \sum_{n=0}^{\infty}\frac{1}{n!}(0A)^n = I.$$

So the initial matrix is non singular, meaning that it will stay along the time non singular and e^{tA} is a fundamental matrix solution with formula $x(t) = e^{tA}x_0$. □

For matrices that are not in a special form, we cannot directly compute the series of the exponential and then then neither the solutions. Specials forms of matrices whose exponential are commutable are for example the ones where the power of the matrix has a general formula, like diagonal matrices. We know that diagonal matrices are directly related to eigenvalues and eigenvectors. The eigenvector are the directions where the matrix act like in the one dimensional case and as often, it is a way to find similar results in the higher dimensional problems.

Let us investigate what happens in theses directions, and choose a candidate that look like a one dimensional solution,$e^{\lambda t}v$, for a real eigenvalue λ of the matrix A and the corresponding eigenvector $\mathbf{v}$. We compute its derivative and obtain

$$\mathrm{d}/\mathrm{d}t(e^{\lambda t}v) = e^{\lambda t}\lambda v = e^{\lambda t}Av = A(e^{\lambda t}v).$$

This give us indeed a non trivial solution since the eigenvector is non null. Theses kind of solutions can be seen in the computation of the exponential when we have a basis of eigenvectors: the matrix A is diagonalisable like $A = PDP^{-1}$ where D is the diagonal matrix of the eigenvalues and P is the non singular matrix formed by the eigenvectors. Then

$$A^n = (PDP^{-1})^n = PDP^{-1}\cdots PDP^{-1} = PD^nP^{-1}$$

and

$$e^{tA} = Pe^{tD}P^{-1} = P\mathrm{diag}(e^{\lambda_j t})P^{-1} = (v_1e^{\lambda_1 t}, \dots, v_ne^{\lambda_n t})P^{-1}$$

which is the fundamental matrix of the kind of solutions we find before, up to a reparametrization.

This was the simple case. More generally the matrix can have complex eigenvalues with complex eigenvectors, but the following lemma show us that complex solutions are made of real solutions :

Lemma 3. *If $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ is a complex solution of real and imaginary parts $\mathbf{x}$ and $\mathbf{y}$, then $\mathbf{x}$ and $\mathbf{y}$ are (real) solutions.*

Proof. This can be shown by the simple fact that

$$\dot{\mathbf{x}} + i\dot{\mathbf{y}} = \dot{\mathbf{z}} = A\mathbf{z} = A\mathbf{x} + iA\mathbf{y}.$$

As A is real, we can directly read the real and imaginary parts as being $A\mathbf{x}$ and $A\mathbf{y}$. The real and respectively imaginary parts being the same along equalities implies $\dot{\mathbf{x}} = A\mathbf{x}$ and $\dot{\mathbf{y}} = A\mathbf{y}$ which is the result. □

We should then take into account complex solutions obtained by complex eigenvalues because they actually give us new reals solutions. We sum up the results together in the following theorem.

Theorem 2. *For a complex eigenvalue $\sigma = \alpha + i\beta$ of A with eigenvector $w = u + iv$, we have the real solutions*

$$e^{\alpha t}(\cos(\beta t)u - \sin(\beta t)v)$$
$$e^{\alpha t}(\cos(\beta t)v + \sin(\beta t)u).$$

They are independent if ad only if u and v are independent. If σ is real, this give only one solution $e^{\alpha t}u$.

Proof. First we have that $e^{\sigma t}w$ is a complex solution, since

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{\sigma t}w) = e^{\sigma t}\sigma w = e^{\sigma t}Aw = A(e^{\sigma t}w).$$

Then by Lemma 3, the real and the imaginary parts are real solutions, we compute them by rewriting the complex solution:

$$e^{\sigma t}w = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t))(u + iv) = e^{\alpha t}(\cos(\beta t)u - \sin(\beta t)v) + ie^{\alpha t}(\cos(\beta t)v + \sin(\beta t)u).$$

We recognise the resulting solutions in the real and imaginary part. If we evaluate them in t=0, we get respectively u and v so the independence follow by Lemma 1. By setting $\beta = 0$, the solutions are $e^{\alpha t}(\cos(0)u - \sin(0)v) = e^{\alpha t}u$ and $e^{\alpha t}v$. But since the eigenvalue is real, the eigenvector is real too and $v = 0$ letting only one non-trivial solution. □

Remark. Note that complex values come by pairs of conjugates, as well as the eigenvectors: $Aw = \sigma w$ implies

$$A\overline{w} = \overline{A}\overline{w} = \overline{Aw} = \overline{\sigma w} = \bar{\sigma}\bar{w}.$$

However, this will not give us new real solutions by the method of Theorem 2 because

$$e^{\overline{\sigma}t}\overline{w} = \overline{e^{\sigma t}}\overline{w} = \overline{e^{\sigma t}w}$$

has the same real and imaginary part as $e^{\sigma t}w$ up to the sign. As a result, a complex eigenvalue give two solutions but together with its conjugate, they give us still two solutions, so we can maybe find as many independent solutions as independent eigenvector we find, real or not.

Now we have to deal with the case when we do not have a basis of eigenvectors. In this case, some eigenvalue λ with algebraic multiplicity μ_a and a geometric multiplicity

$$\mu_g = \dim\ker(A - \lambda I) < \mu_a.$$

Such an eigenvalue is called *defective*, and the matrix is said *defective* too when it has at least one defective eigenvalue. We use here the concept of generalized eigenvector that come from the result about the Jordan form:

Definition 4. A vector $\mathbf{w}$ is a *generalized eigenvector* of rank m of a matrix A and corresponding to an eigenvalue λ, if it is a vector (possibly complex if λ is complex) that satisfy

$$(A-\lambda I)^m\mathbf{w}=0 \quad \text{and} \quad (A-\lambda I)^{m-1}\mathbf{w}\neq 0$$

for $m \in \mathbb{N}^*$.

A *canonical* basis of generalised eigenvectors is a basis of generalised eigenvector such that for all generalized eigenvector $\mathbf{w}$ of rank m that is in the basis, for all $0<j<m$, $(A-\lambda I)^j\mathbf{w}$ are generalised eigenvectors of rank $m-j$ with respect to λ, and are in the basis too.

Remark. As before, a square matrix powered by 0 is the identity matrix. Note that $(A-\lambda I)^{m-1}\mathbf{w}\neq 0$ implies that in particular $\mathbf{w}\neq 0$. As a consequence, a generalised eigenvector of rank 1 is a usual eigenvector.

Theorem 3. *A matrix always have a canonical basis of generalised eigenvectors.*

Proof. Linear algebra result, see the appendix of [3]. □

In term of computation, we start from $(A-\lambda I)\mathbf{v}=0$ to find the eigenvectors. Then we search for some $\mathbf{w}$ such that $(A-\lambda I)\mathbf{w}=\mathbf{v}$, assuring that $(A-\lambda I)^2\mathbf{w}=(A-\lambda I)\mathbf{v}=0$ and hence that $\mathbf{w}$ is a generalised vector of rank 2. We continue like this making a chain of generalized eigenvectors, even if they are complex.

We show that this basis is useful if we use its elements as initialisation points.

Theorem 4 (General form of the eigensolutions)**.** *For a canonical basis B of generalised eigenvectors, we have a basis of complex solutions with form $e^{t\lambda}p_{\mathbf{w}}(t)$, where λ is the eigenvalue associated to a generalised eigenvector $\mathbf{w}$ of rank m in the basis, and $p_{\mathbf{w}}$ is a polynomial of degree $m-1$ with coefficients that are in the eigenspace of w .*

Proof. We see easily that the basic property of the real exponential that changes sum into product is true for the matrices with the condition that they commute. The proof is similar to the real case. We write $A=\lambda I+(A-\lambda I)$ and λI is diagonal hence commutative with all matrices. Taking $\mathbf{w}$ from a canonical basis, we get

$$e^{tA}\mathbf{w}=e^{t\lambda I+t(A-\lambda I)}\mathbf{w}=e^{t\lambda I}e^{t(A-\lambda I)}\mathbf{w}=e^{t\lambda}\sum_{n=0}^{\infty}\frac{1}{n!}t^n(A-\lambda I)^n\mathbf{w}=e^{t\lambda}\sum_{n=0}^{m-1}\frac{1}{n!}t^n\mathbf{w}_n \tag{1.1}$$

where the $\mathbf{w}_n=(A-\lambda I)^n\mathbf{w}$ are other generalised vectors by definition of the canonical basis. So have indeed a polynomial in t with generalised eigenvectors coefficients. These are surely independent solutions for different choices of $\mathbf{w}$ since the generalised eigenvectors are supposed independent and they are the initial values of these solutions, the proof is complete. □

We have now a complex basis but we know that eigenvalues and eigenvectors come by pairs of conjugates. It is the same for generalised eigenvectors and eigenvalues:

$$(A-\lambda I)^m\mathbf{w}=0 \quad \text{and} \quad (A-\lambda I)^{m-1}\mathbf{w}\neq 0$$

implies

$$(A-\overline{\lambda}I)^m\overline{\mathbf{w}}=0 \quad \text{and} \quad (A-\overline{\lambda}I)^{m-1}\overline{\mathbf{w}}\neq 0$$

So we will get solutions like (1.1) by pairs of conjugates $e^{tA}\mathbf{w}$ and $e^{tA}\overline{\mathbf{w}}$. by subtracting them or summing them, we get two real new solutions that we will call *degenerated*:

$$e^{tA}\mathbf{w}+e^{tA}\overline{\mathbf{w}}=e^{tA}(\mathbf{w}+\overline{\mathbf{w}})=2e^{tA}\Re(\mathbf{w}),$$

$$e^{tA}\mathbf{w}-e^{tA}\overline{\mathbf{w}}=e^{tA}(\mathbf{w}-\overline{\mathbf{w}})=2e^{tA}\Im(\mathbf{w}).$$

However, we can see that the formulas for the real and the imaginary part are not very convenient:

$$
\begin{aligned}
& e^{t\lambda}\sum_{n=0}^{m-1}\frac{1}{n!}t^n\mathbf{w}_n \\
&= e^{t\alpha}(\cos(\beta t)+i\sin(\beta t))\sum_{n=0}^{m-1}\frac{1}{n!}t^n(\Re\mathbf{w}_n+i\Im\mathbf{w}_n) \\
&= e^{t\alpha}\Big(\cos(\beta t)\sum_{n=0}^{m-1}\frac{1}{n!}t^n\Re\mathbf{w}_n-\sin(\beta t)\sum_{n=0}^{m-1}\frac{1}{n!}t^n\Im\mathbf{w}_n\Big)+ie^{t\alpha}\Big(\cos(\beta t)\sum_{n=0}^{m-1}\frac{1}{n!}t^n\Im\mathbf{w}_n+\sin(\beta t)\sum_{n=0}^{m-1}\frac{1}{n!}t^n\Re\mathbf{w}_n\Big).
\end{aligned}
$$

We know that we have a (complex) basis of the complex space solution. The dimension of the $\mathbb{C}$-vector space $\mathbb{C}^n$ is n, and since $\mathbb{R}\subset\mathbb{C}$, all real solutions are included in this description. However, the method of taking the real part of a basis does not give us a priori a independent set of real solutions. For this reason, we will stay in the the complex space because it present a good and simple description thanks to its algebraic closure, and we will be able too deduce from the complex case results about what we need for the real case. We state a corollary of the precedent theorem, giving all possible forms that can take the reals solutions, but without giving a explicit formula.

Corollary 2. *All solutions to linear system have coordinates that are linear combinations of the following functions :*

- $e^{\lambda t}$
- $e^{\alpha t}\cos(\beta t)$, $e^{\alpha t}\sin(\beta t)$
- $t^j e^{\lambda t}$, $t^j e^{\alpha t}\cos(\beta t)$, $t^j e^{\alpha t}\sin(\beta t)$

where λ is a real eigenvalue, $\sigma=\alpha+i\beta$ is an complex eigenvalue, $0\le j<m_a$ is a natural number with m_a the algebraic multiplicity of σ. Note that each point is a generalisation of the precedent.

Remark. This actually describe the complex space too, since we can rebuild the complex solutions we used to derive this corollary by summing the real part and i times imaginary one.

1.2 Stability

In other terms, Theorem 2 tell us that each non null real eigenvalue gives direction(s) where the trajectories are straight and of exponential velocity. Null eigenvalues give direction(s) where trajectories are fixed. Non real eigenvalue gives direction(s) where the trajectories are like ellipses that can change of size exponentially. The Theorem 4 add other sorts of solutions as polynomials resized by a exponential and correspond to the case of degenerated eigenvalues.

All these considerations are only on the special directions. Depending on the sign of the real part of the eigenvalues, theses specials solutions go very fast to 0, stay in an orbit, diverge very fast to infinity values, or maybe follow a polynomial. This motivate us to see how these special solutions act together, and what is the asymptotic comportment of trajectories, how stable is the origin, and make a classification about all theses factors. First of all we define concepts about stability, and asymptotic convergence. For this we put our self in a more general cadre which will be useful later, with function $\mathbf{F}$ and the equation

$$\dot{\mathbf{x}}=\mathbf{F}(\mathbf{x})$$

We suppose $\mathbf{F}$ regular enough to have a flow $\phi(\mathbf{x}_0,t)$ which encapsulate all solutions, which are unique and defined in $\mathbb{R}$:

$$
\begin{aligned}
&\dot{\phi}(\mathbf{x}_0,t)=\mathbf{F}(\phi(\mathbf{x}_0,t)) \\
&\phi(\mathbf{x}_0,0)=\mathbf{x}_0
\end{aligned}
$$

Definition 5. A solution x is *Lyapunov stable* or *L-stable* if the continuity of the flow with respect to the initial condition $\mathbf{x}_0$ is uniform with respect to positive times. Namely, if for all $\epsilon > 0$ there exist a $\delta > 0$ such that $\|\mathbf{z} - \mathbf{x}(0)\| < \delta$ implies that for all $t \geq 0$, $\|\phi(\mathbf{z}, t) - \mathbf{x}(t)\| < \epsilon$.

Definition 6. Two solutions $\mathbf{x}$ and $\mathbf{y}$ are *ω-attracted* to each other if $\lim_{t\to\infty} \|\mathbf{y}(t) - \mathbf{x}(t)\| = 0$. The resulting equivalence class is called the *basin of attraction.* A solution $\mathbf{x}$ is said *ω-attracting* or *attracting* if there is a $\delta > 0$ such that $\|\mathbf{z}_0 - \mathbf{x}_0\| < \delta$ implies that $\phi(\mathbf{z}_0, t)$ is attracted by $\mathbf{x}$. A solution is said *globally ω-attracting* (or *globally attracting*) on a set if this set is in the basin of attraction. We do not need to specify the set if it is the all solution space.

Remark.

The basin of attraction can be see as a space of solutions, as well as a space of points, identifying the solutions by their initial positions.

Definition 7. A solution is *asymptotically stable* if it is L-stable and attracting. It is said *globally asymptotically stable* if it is L-stable and globally attracting

Remark. Note that these two notions are different. For examples of non-attracting point which is L-stable, we can take simply $F(x) = 0$ or for a non trivial case, we take the origin of $\dot{x} = -y$, $\dot{y} = x$ who describe the circle trajectories $x(t) = \cos(t)$, $y(t) = \sin(t)$.

In the other way, there exist non L-stable points which are attracting. Such a point is the limit of all near trajectories but they always go a bit far before converging, like a detour. For this we place our-self in polar coordinates. We want the trajectories to follow the circle and finish in $(1, 0)$ for this we make the θ always go and stop to 2π, and r go and stop to 1. For this we can write $(\dot{r}, \dot{\theta}) = (1-r, 2\pi-\theta) = G(r, \theta)$. But if we want $(G_1 \cos(G_2), G_1 \sin(G_2))$ to be continuous on $\mathbb{R}_+ \times 0$, we should write $\dot{r} = r(1-r)$, $\dot{\theta} = \theta(2\pi-\theta)$. And to obtain the continuity of the derivative, in purpose to have a flow, we could write $\dot{r} = r(1-r^2)$, $\dot{\theta} = \theta(2\pi-\theta^2)$. This gives us what we need but to be able to explicitly change the coordinates into cartesian, we prefer $\dot{\theta} = \sin(\theta/2)^2 = (1-\cos(\theta))/2$. The fact that we started as a polar coordinate and that they are actually computable in the polar equation, actually gives a proof, by construction, that all solutions tends indeed to $(1, 0)$ but solutions beginning near $(1, 0)$ and in the upper plane, pass behind the origin before returning to the fixed point $(1, 0)$. This shows that it is attracting but not L-stable. We compute however the formula of the equation:

$$\begin{aligned}\dot{x} = (r\cos(\theta))' &= r'\cos(\theta) - r\sin(\theta)\theta' = r(1-r^2)\cos(\theta) - r\sin(\theta)\frac{1-\cos(\theta)}{2} \\ &= (1-x^2-y^2)x - y(1 - \frac{x}{\sqrt{x^2+y^2}}) = x - y + x(x^2+y^2) + \frac{x^2}{\sqrt{x^2+y^2}}.\end{aligned}$$

Similarly, we find

$$\dot{y} = x + y - y(x^2+y^2)\frac{xy}{\sqrt{x^2+y^2}}$$

The following graphic shows the behave of solutions and how indeed they all make a detour:

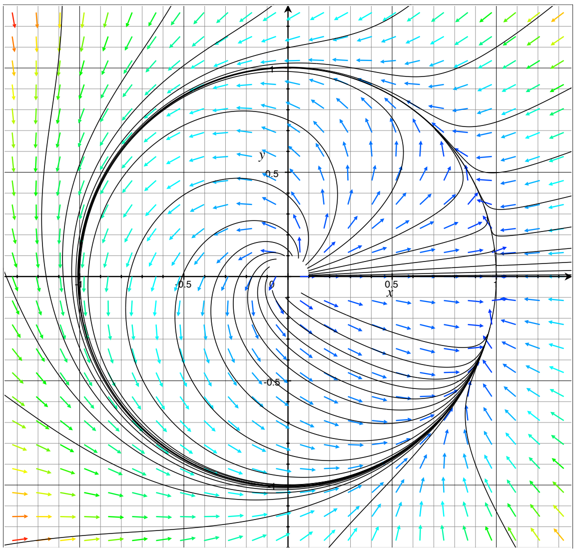

Figure 1.1: An attracting, but not Lyapunov stable fixed point $(1, 0)$ of a non-linear dynamical system

Lemma 4. *When* $\mathbf{F}$ *is linear, i.e.* $\dot{\mathbf{x}} = A\mathbf{x}$, *all solutions have the same L-stability and attractivity.*

Proof. In the linear case, ϕ is linear in the first variable, indeed $\phi(\mathbf{x}_0, t) = e^{tA}\mathbf{x}_0 =: X(t)\mathbf{x}_0$. Now for any solution $\mathbf{x}$ and all initial point $\mathbf{z}$,

$$\|\phi(\mathbf{z}, t) - \mathbf{x}(t)\| = \|X(t)\mathbf{z} - X(t)\mathbf{x}(0)\| = \|X(t)(\mathbf{z} - \mathbf{x}_0)\| = \|\phi((\mathbf{z} - \mathbf{x}_0), t) - \phi(0, t)\|,$$

meaning that L-stability and attractivity is entirely determined by the stability of the trivial solution $\phi(0, t) = 0$. □

Remark. Therefore, we can speak of the L-stability and the attractivity of a linear system, meaning that it applies to all solutions, or none of them, and doing it by looking at the trivial solution 0. In this case the system is L-stable if and only if there exists a $\delta > 0$ such that $\|\mathbf{x}_0\| < \delta$ implies that for all $t \geq 0$, $\|\phi(\mathbf{x}_0, t)\| < \epsilon$. The system is attracting if and only if there exists a $\delta > 0$ such that $\|\mathbf{x}_0\| < \delta$ implies that $\phi(\mathbf{x}_0, t) \to 0$.

Theorem 5. *The linear system is L-stable if and only if each of its solutions is bounded for positive times.*

Proof. Suppose the system is L-stable, and for contradiction that a solution $\mathbf{x}$ is not bounded. Let $\delta > 0$ be the distance given by the stability, such that all solutions that start with a norm smaller than δ do not go away the unit ball. We can then define an other solution $\mathbf{y} = \delta\mathbf{x}/\|\mathbf{x}(0)\|$ and $\|\mathbf{y}(0)\| = \delta$, so $\mathbf{y}$ wont go away the unit ball and is bounded. This contradict the fact that $\mathbf{x}$ and $\mathbf{y}$ are proportionals and all solutions must be bounded.

Suppose now that all solutions are bounded. Then the columns of $X = e^{tA}$ the fundamental system are bounded, implying that actually the norm of X is bounded (all norm on finite dimensional spaces are equivalent). Now we get

$$\|\phi(x_0, t)\| = \|X(t)x_0\| \leq \|X(t)\|\|\mathbf{x}_0\| \leq \max_{t\geq 0} \|X(t)\|\|\mathbf{x}_0\|$$

which is smaller than any positive ϵ as soon as

$$\|\mathbf{x}_0\| < \frac{\epsilon}{\max_{t\geq 0} \|X(t)\|}$$

and give us the stability of the system. □

Theorem 6. *The linear system is globally asymptotically stable if and only if it is attracting.*

Proof. By definition, global asymptotic stability implies global attractivity and so in particular attractivity with any radius condition. In the other direction, if it is attracting with radius condition δ, such that when $\|\mathbf{x}_0\| < \delta$, $\|\phi(\mathbf{x}_0, t)\| \to 0$. Then any solution $\mathbf{x}$ can be written $\mathbf{x} = \|\mathbf{x}(0)\|\mathbf{y}/\delta$ where $\mathbf{y} = \delta\mathbf{x}/\|\mathbf{x}(0)\|$ is a proportional solution that start with norm $\delta\|\mathbf{x}(0)\|/\|\mathbf{x}(0)\| = \delta$ and is small enough to converge to zero, implying that $\mathbf{x} = \mathbf{x}(0)\mathbf{y}/\delta$ converge to zero too. The system is globally attracting by arbitrarity of $\mathbf{x}$. Now that all solutions are converging to zero, they are all bounded and by Theorem 5, we know that the system is actually L-stable. Both condition of stability and global attractivity are reunited, the system is globally asymptotically stable. $\square$

Example 1 (saddle)**.** Consider the system

$$\begin{cases} \dot{x} = x + 3y \\ \dot{y} = 3x + ay \end{cases} \quad \text{rewritten with } \mathbf{z} = (x, y)^\top : \quad \dot{\mathbf{z}} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \mathbf{z} = A\mathbf{z}$$

The characteristic polynomial is $p_A(\lambda) = \det(A - \lambda I) = (1 - \lambda)^2 - 9 = (-2 - \lambda)(4 - \lambda)$ and has roots $\lambda_1 = -2$, $\lambda_2 = 4$. For the eigenvalue -2,

$$A - \lambda_1 I = A + 2I = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}$$

has $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ in the kernel and this is the first eigenvector. The stable associated eigensolution is

$$e^{-2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xrightarrow[t \to \infty]{} 0.$$

Similary,

$$A - \lambda_2 I = A - 4I = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$$

give the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and the unstable eigensolution $e^{4t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow[t \to \infty]{\|\cdot\|} \infty$. We plot the results:

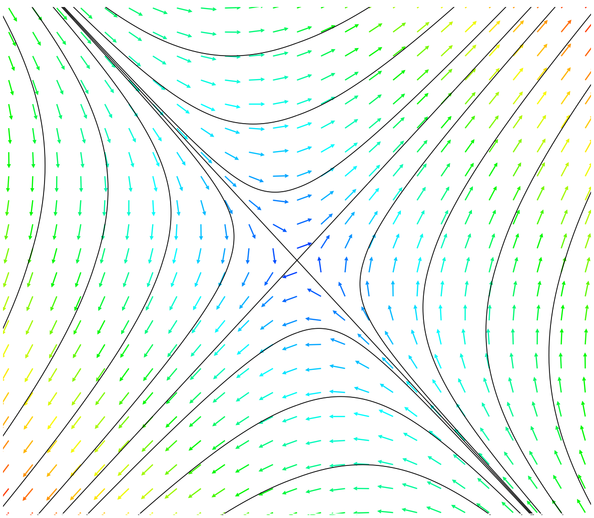

Figure 1.2: Saddle phase plane

We see that we have only two straight solutions, the eigensolutions. Their norm evolve exponentially but they are time inverted from each other. All the other solutions are combinations of those two and are like hyperbolas, they come near the origin and then goes back away. The only one that goes to zero is the eigensolution with eigenvalue -2. We can be as near to zero as we want for the start, the solution eventually goes to infinity if it is not this eigensolution. We does not see any stability, a such system is called a saddle and we will define a general definition for it later.

Example 2 (focus)**.** Consider the system

$$\dot{\mathbf{z}} = \begin{pmatrix} -4 & -2 \\ 3 & -11 \end{pmatrix} \mathbf{z} = A\mathbf{z}$$

The characteristic polynomial is $(-4-\lambda)(11-\lambda)+6 = (-5-\lambda)(-10-\lambda)$ and has roots $\lambda_1 = -10$, $\lambda_2 = -5$. For the eigenvalue -10,

$$(A - \lambda_1 I)\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 & -2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = 0.$$

The stable associated eigensolution is $e^{-10t}\begin{pmatrix} 1 \\ 3 \end{pmatrix} \xrightarrow[t\to\infty]{} 0$. Similary,

$$(A - \lambda_2 I)\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 3 & -6 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 0$$

give the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and the second stable eigensolution $e^{-5t}\begin{pmatrix} 2 \\ 1 \end{pmatrix} \xrightarrow[t\to\infty]{} 0$. We plot the results:

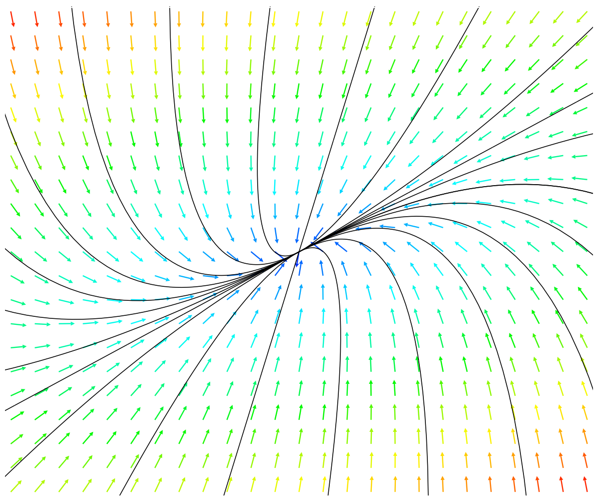

Figure 1.3: Focus phase plane

Here the two eigensolutions goes to zero exponentially and thus, all linear combinations do the same. The system is attracting and so asymptotically stable. We will call this situation later a focus, and more generally a stable node.

Example 3 (null eigenvalue)**.** Consider the system

$$\dot{\mathbf{z}} = \begin{pmatrix} -1 & -3 \\ -1 & -3 \end{pmatrix} \mathbf{z} = A\mathbf{z}$$

The characteristic polynomial is $(-1-\lambda)(-3-\lambda)-3=\lambda(-4-\lambda)$ and has roots $\lambda_1=-4$, $\lambda_2=0$. The eigenvectors are respectively $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}3\\-1\end{pmatrix}$. They give the solutions $e^{-4t}\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}3\\-1\end{pmatrix}$ (constant solution).

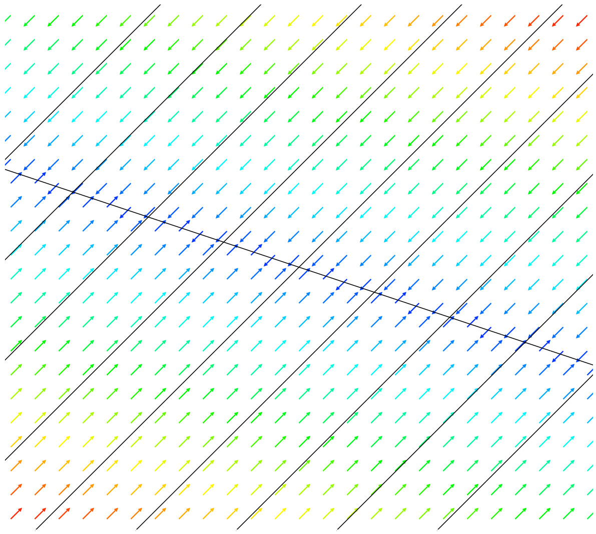

Figure 1.4: Null eigenvalue phase plane

We see that we have a line where the solutions are fixed and a other where the solution goes exponentially to the origin. As a result the system is L-stable because solutions do not go away, but it is not attracting or asymptotically stable because this fixed line contain solutions that never move to the origin.

These are simple examples where the matrix is diagonalisable. When it is not the case, the situation is more complicated because of the generalised eigenvectors:

Example 4 (defective eigenvalue)**.** Consider the system

$$\dot{\mathbf{z}}=\begin{pmatrix}-2 & 1\\-1 & 0\end{pmatrix}\mathbf{z}=A\mathbf{z}.$$

The characteristic polynomial is $(1+\lambda)^2$ and has one root $\lambda=-1$, $\lambda_2=0$; $(A+I)=\begin{pmatrix}-1 & 1\\-1 & 1\end{pmatrix}$ give the eigenvector $\mathbf{v}=\begin{pmatrix}1\\1\end{pmatrix}$. It is the only one because $\det\begin{pmatrix}-1 & 1\\-1 & 1\end{pmatrix}=0$ and it must be of rank one. This give us a solutions $e^{-t}\begin{pmatrix}1\\1\end{pmatrix}$. We search for a generalised eigenvector:

$$(A+I)\mathbf{w}=\mathbf{v}\ ,\ i.e.\ \begin{pmatrix}-1 & 1\\-1 & 1\end{pmatrix}\mathbf{w}=\begin{pmatrix}1\\1\end{pmatrix}.$$

We find $\mathbf{w}=\begin{pmatrix}0\\1\end{pmatrix}$ and the second solution is $e^{\lambda t}(\mathbf{w}+t\mathbf{v})=e^{-t}(\begin{pmatrix}0\\1\end{pmatrix}+t\begin{pmatrix}1\\1\end{pmatrix})=e^{-t}\begin{pmatrix}t\\1+t\end{pmatrix}$.

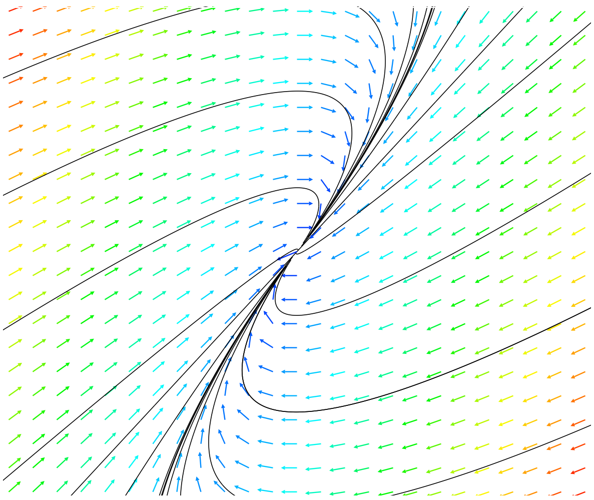

Figure 1.5: Defective eigenvalue phase plane

Remark. Interested readers might want to have a look online to this interactive phase plane [1], to see the diversity of possibilities on the plane.

Theorem 7. *The linear system is*

1. *L-stable if and only if all the eigenvalues of the matrix have non positive real parts, and all the one that are defective have negative real part.*
2. *globally asymptotically stable if and only if all the eigenvalues of the matrix have negative real parts.*

Proof.

1. By Theorem 5, we just have to show that the conditions on the eigenvalues are equivalent to the fact that the solutions are bounded. The Corollary 2 told us the possible forms of all solutions. The functions $t^j e^{\alpha t}\cos(\beta t)$, $t^j e^{\alpha t}\sin(\beta t)$ describe all the possibilities of linear combinations for the coordinates. Since a linear combination is a finite sum, the solution is bounded if all these possible terms are bounded.

   For both, if the eigenvalue $\sigma = \alpha + i\beta$ is non defective, then $j = 0$ only and they are bounded if α is non positive. If σ is defective, $j > 0$ and they are bounded if α is negative because the exponential is (t^j) for all j's. The sytem is now L-stable.

   Alternatively, if there exists an eigenvalue $\sigma = \alpha + i\beta$ with eigenvector $w = u + iv$ that have a positive real part, then by Theorem 2, there exist a solution $e^{\alpha t}(cos(\beta t)u - \sin(\beta t)v)$ and its norm $e^{\alpha t}\|cos(\beta t)u - \sin(\beta t)v\|$ is not bounded since $cos(\beta t)u - \sin(\beta t)v$ does not converge to zero. If σ is defective and is purely imaginary, then Theorem 4 tell us that there exist a solution $p_w(t)$, polynomial of non null degree, and $p_w(t) \to \infty$ when $t \to \infty$ because of the dominant term. There exists solutions that are unbounded , and the system is not L-stable.

2. By Theorem 5, we just have to show that the conditions on the eigenvalues are equivalent to the fact that the system is attracting.

   But following the considerations of the first part, if all eigenvalues have negative real parts, there is always a $e^{\alpha t}$ with $\alpha < 0$ term in front of the bases solutions and they all converge to zero. The system is attracting, and globally asymptotically stable.

[1]Copyright © 2009–2015 H. Miller | Powered by WordPress, https://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/

Alternatively if it is not the case and that $\sigma = \alpha + i\beta$ has a non negative part, we have a solution $e^{\alpha t} p_w(t)$ that satisfy $\|e^{\alpha t} p_w(t)\| \geq \|p_w(t)\|$ which does not go to zero. The system is not attracting and then not globally asymptotically stable.

□

Remark. The computation of eigenvalues can be complicated and a bit cumbersome in high dimensions; there exists some criteria on the matrix to have only eigenvalues with negative real part, that are more easy to compute. We can cite for instance the Hurwitz criterion given in [1] page 85, that reduces the problem to a computation of principal minors of a nn matrix.

Example 5. The second condition of the first point seem to point out a very specific situation. Would the theorem work without this condition, meaning that this situation is algebraically impossible? Indeed, we usually do not often see real matrices with a defective and purely imaginary eigenvalue. This is maybe because the following remarks: it cannot be in one dimension because eigenvalues are real; it cannot be in two dimension because imaginary eigenvalues come by pair of conjugates and then they would be opposed, hence different and non-defective. To have a purely imaginary eigenvalue of multiplicity two we must have its conjugate of multiplicity two too. So we try in four dimension. We search for a matrix with characteristic polynomial $(\lambda^2 + 1)^2$. The basic bloc that introduce i eigenvalue is

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

We could try with $\text{diag}(J, J)$, but this gives a non-defective eigenspace. With a little modification on it, we find a matrix that is defective:

$$\begin{pmatrix} J & I_2 \\ 0 & J \end{pmatrix}.$$

It has the only two independent eigenvectors $(\pm i, 1, 0, 0)^\top$. Hence in this pathological situation, the defective eigensolutions are of the form $e^{it}p(t) = (\cos(t) + i\sin(t))p(t)$, which is neither constant, neither exponential, neither cyclic, neither a polynomial, neither bounded.

The last theorem is one of the main results of the section, it presents a complete understanding of the stability. It motivates the categorisation of linear systems, taking into account the nature of their stability.

Definition 8. For a sytem $\dot{\mathbf{x}} = A\mathbf{x}$, we set the following denomination of the matrix A.

- *Stable* : all eigenvalues have negative real part. $(i\mathbb{R} + \mathbb{R}_-^*)$
- *unstable* : $-A$ is stable. i.e. all eigenvalues have positive real part. $(i\mathbb{R} + \mathbb{R}_+^*)$
- *saddle* : all eigenvalues have a non null real part and they do not all have the same sign.
- *hyperbolic* : stable, unstable or saddle; *i.e.* all eigenvalues have non null real part. $(i\mathbb{R} + \mathbb{R}^*)$
- *center* : all eigenvalues are purely imaginary. $(i\mathbb{R}^*)$
- *elliptic* : all eigenvalues are non null, purely imaginary, and non defective.
- *focus* : all eigenvalues are real, negative, and non defective.
- *source* : all eigenvalues are real, positive, and non defective.
- *degenerated* : there exists a defective eigenvalue.

More generally, we use these terms to speak about a specific eigenvalue, seeing it as a one-dimensional matrix, meaning that it must respect the generic condition (this does not apply to saddle).

Remark. The matrix A is unstable (resp. a focus) if $-A$ is stable (resp. a source).

We point out the similarity between the words "L-stabe" and "stable", which speak respectively about Lyapunov stability and eigenvalues, do not confuse them. A stable linear system is L-stable but not not in the other way because of the elliptic case.

For example of the classification, we present a focus resulting of the matrix

$$\begin{pmatrix} -4 & 5 \\ -5 & 2 \end{pmatrix},$$

which has no real eigenvalues but two complex and stable ones. We do not present the computations, they are the same as before. In this case the solutions are in a convergent spiral because of the trigonometric functions that appear in the formula and the negativity of the real part.

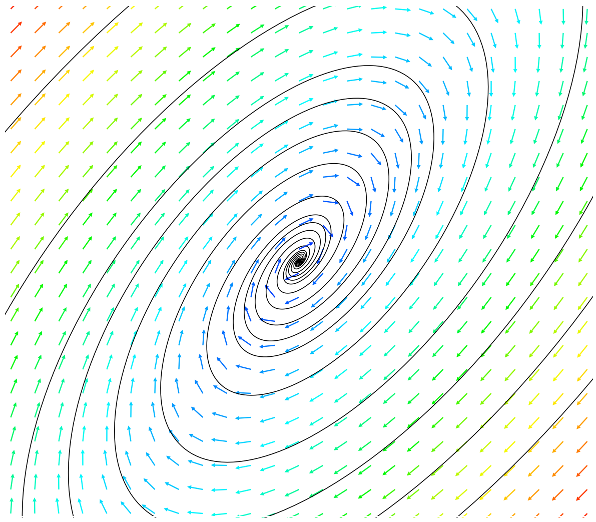

Figure 1.6: Saddle phase plane

Corollary 3.

- *Stable(and focus in particular) and elliptic linear systems are L-stable.*
- *Sources and saddles linear systems are not L-stable, so neither asymptotically stable.*
- *Stable (and focus in particular) linear systems are globally asymptotically stable.*

All this classification and considerations on the eigenvalues, show us that there exists directions where the trajectories act kind of exponentially, and these are the directions that are linear combinations of some generalised eigenvectors. We want now to find a structure on the space, to describe the case where we don't have the stability, but some directions act like in a stable case, and this is the starting point of the next section.

1.3 Structure of stability and links to the nonlinear case

Inspired by the theorems about stability that we established, and the classification we made, we present a separation of the space of solution, which tell us where the solutions act like in a stable, unstable, or center system.

Definition 9. For a linear system $\dot{\mathbf{x}} = A\mathbf{x}$ of dimension n, we define the stable eigenspace, the unstable eigenspace, and the center eigenspace as follows respectively :

$$\mathbb{E}^s = \text{Span}\{v \in \mathbb{C}^n |\, v \text{ is a generalised eigenvector of the system with eigenvalue } \lambda \text{ such that } \Re(\lambda) < 0\}$$
$$\mathbb{E}^u = \text{Span}\{v \in \mathbb{C}^n |\, v \text{ is a generalised eigenvector of the system with eigenvalue } \lambda \text{ such that } \Re(\lambda) > 0\}$$
$$\mathbb{E}^c = \text{Span}\{v \in \mathbb{C}^n |\, v \text{ is a generalised eigenvector of the system with eigenvalue } \lambda \text{ such that } \Re(\lambda) = 0\}$$

They are the generalised eigenspaces of the stable, unstable, and center generalised eigenvectors. In the same idea we want to compare them to the directions where the solutions act exponentially or not:

$$V^s = \{v \in \mathbb{C}^n |\, \exists a, b > 0 : \forall t \geq 0, \|e^{At}v\| \leq be^{-at}\}$$
$$V^u = \{v \in \mathbb{C}^n |\, \exists a, b > 0 : \forall t \leq 0, \|e^{At}v\| \leq be^{at}\}$$
$$V^c = \{v \in \mathbb{C}^n |\, \forall a > 0,\, e^{At}ve^{-a|t|} \to 0 \text{ when } |t| \to \infty\}.$$

Namely, the spaces where the solutions act exponentially or not.

Remark. We can precise the value of the matrix as an argument when dealing with multiples systems and does not use argument when the context is clear. Note that they are all subspaces and $\mathbb{E}^u(A) = \mathbb{E}^s(-A)$ as well as $V^u(A) = V^s(-A)$. Taking $-A$ instead of A correspond actually to inverse time.

For the center eigenspace, the reason why it is not enough to watch if a solution is, and why we compare it to an exponential, is only because of the pathological case described in Example 5, where the eigenvalue is of center type, but is defective and the eigensolution is polynomial, with a trigonometric factor in front of it. It is neither bounded, neither exponential.

Lemma 5. *For $\sigma \in \{s, u, c\}$, we have the equalities $\mathbb{E}^\sigma = V^\sigma$, the direct sum $\mathbb{C}^n = \mathbb{E}^s \oplus \mathbb{E}^u \oplus \mathbb{E}^c$, and the $\mathbb{E}^\sigma$ are invariants in the sense that $\phi(\mathbb{E}^\sigma \times \mathbb{R}) = \mathbb{E}^\sigma$,* i.e. *all the solutions stay in the space.*

Proof. The invariance of the sets comes from the general form of the eigensolutions presented in Theorem 2. Such a solution looks like $e^{\lambda t}p(t)$, with an eigenvalue λ and a polynomial $p(t)$ whose coefficients are in the generalised eigenspace of λ. It is a linear combination of vectors in the same $\mathbb{E}^\sigma$ for all t. The eigensolutions stay then in the space, as well as a span of them, $\mathbb{E}^\sigma$.

The direct sum come from the linear algebra fact that eigenspaces associated to different eigenvalues are independents. The three spaces are just a separation of them in three sums. Namely $\mathbb{E}^\sigma$ is just the direct sum of the generalised eigenspaces associated to eigenvalues of type σ.

Now we prove the equalities. First, we show $\mathbb{E}^\sigma \subset V^\sigma$. Each hyperbolic generalised eigensolution associated to a generalised eigenvector $\mathbf{w}$, can be written $e^{\delta t}p(t)$ for a polynomial p with complex generalised eigenvector coefficients associated to the same eigenvalue $\delta = a + ib$ (a,b scalars).

$\mathbb{E}^s$: When $a < 0$, we can just bound the norm as

$$\|e^{\delta t}p(t)\| = e^{at}\|p(t)\| \leq e^{ta/2} \max_{t\geq 0} \||p(t)|e^{ta/2}\|,$$

The max exists because the continuous expression $p(t)e^{ta/2} \to 0$ when $t \to \infty$ if $a < 0$ (by Taylor expansion or iterating Hospital).

$\mathbb{E}^u$: When $a > 0$, we directly obtain the $\mathbb{E}^u(A) = \mathbb{E}^s(-A) \subset V^s(-A) = V^u(A)$ by the remark made above.

$\mathbb{E}^c$: When $a = 0$, we have

$$\|e^{-\alpha|t|}e^{\delta t}p(t)\| = e^{-\alpha|t|}\|p(t))\| \to 0 \text{ when } |t| \to \infty.$$

for any $\alpha > 0$.

We get then $\mathbf{w} \in \mathbb{E}^\sigma$ for the three σ. We clearly see that the V^σ's are subspaces, and since $\mathbf{w}$ was chosen as an arbitrary generalised eigenvector, actually the whole span satisfy $\mathbb{E}^\sigma \subset V^\sigma$.

Secondly, we prove the inverse inclusion. Each vector v of V^σ can be written as a unique sum of vectors of the three $\mathbb{E}^\sigma$: $v = v_s + v_u + v_c$. We define a new norm on $\mathbb{C}^n$ using this unique writing.

$$N(v) = \left\|(\|v_s\|, \|v_u\|, \|v_c\|)\right\|_1 = \|v_s\| + \|v_u\| + \|v_c\|.$$

Indeed, absolute homogeneity and positive definiteness are clear, while subadditivity come from the unique decomposition: for vectors v and w, we have $v_s + w_s + v_u + w_u + v_c + w_c = v + u$. This implies necessarily that $(u + v)_\sigma = v_\sigma + w_\sigma$, and then

$$N(v+w) = \|v_s+w_s\|+\|v_u+w_u\|+\|v_c+w_c\| \leq \|v_s\|+\|w_s\|+\|v_u\|+\|w_u\|+\|v_c\|+\|w_c\| = N(v)+N(w).$$

Now, N is equivalent to any other norm, *i.e.* we have a constant c such that $c^{-1}\|v\| \leq N(v) \leq c\|v\|$. If $\sigma = s$, $N(e^{At}v) \to 0$ exponentially when $t \to \infty$. By looking at the norm N, we see that it must be the same for the directions v_s, v_u and v_c. From the general form of eigensolutions we see that unstable and center eigensolutions does not even goes to zero:

$$\|e^{\delta t}p(t)\| = \|e^{\Re(\delta)t}p(t)\| \geq e^{\Re(\delta)t} \min \|p\| \geq \min \|p\| > 0$$

Where the min exist and is positive because the polynomial is not trivial. Then actually $v_u = v_c = 0$, and $v = v_s \in \mathbb{E}^s$, so $V^s \subset \mathbb{E}^s$. As before we directly conclude that

$$V^u(A) = V^s(-A) \subset \mathbb{E}^s(-A) = \mathbb{E}^u(A).$$

Same kind of argument for a $v = v_s + v_u + v_c \in V^c$:

$$e^{-\alpha|t|}\|e^{\delta t}p(t)\| = e^{-\alpha|t|}e^{at}\|p(t)\| \geq e^{(at/|t|-\alpha)|t|} \min \|p\| \geq \min \|p\| > 0$$

for a small $0 < \alpha < |a|$ and alternatively for positive time if δ unstable ($a > 0$), or for negative time if δ stable ($a < 0$). We get that v is in $\mathbb{E}^c$. We have proven the both inclusions and obtain the result. □

This result gives more precision about the velocity of convergence in a context of stability, it is exponential. It gives indication in the other way too, the velocity of the convergence of a solution can tell us about its composition with respect to the eigen-basis. It must be a somewhat "purely" stable composed solution to vanish exponentially, all other (composed) solutions are not exponential.

Now that we understand well the complex eigenspaces of solutions, we can easily deal with the real case:

Corollary 4. *We have the equalities* $\Re(\mathbb{E}^\sigma) = \Re(V^\sigma)$ *and* $\dim(\Re(\mathbb{E}^\sigma)) = \dim(\mathbb{E}^\sigma)$*, the direct sum* $\mathbb{R}^n = \Re(\mathbb{E}^s) \oplus \Re(\mathbb{E}^u) \oplus \Re(\mathbb{E}^c)$*, and the sets* $\mathbb{E}^s$ *are invariant:* $\phi(\Re(\mathbb{E}^\sigma) \times \mathbb{R}) = \Re(\mathbb{E}^\sigma)$.

Proof. The equality $\Re(\mathbb{E}^\sigma) = \Re(V^\sigma)$ is trivial. We have the simple sum

$$\mathbb{R}^n = \Re(\mathbb{C}^n) = \Re(\mathbb{E}^s \oplus \mathbb{E}^u \oplus \mathbb{E}^c) = \Re(\mathbb{E}^s) + \Re(\mathbb{E}^u) + \Re(\mathbb{E}^c).$$

This is actually a direct sum because the conjugate invariance of the $\mathbb{E}^\sigma$'s implies that $\Re(\mathbb{E}^\sigma) \subset \mathbb{E}^\sigma$. This last inclusion implies that $\dim(\Re(\mathbb{E}^\sigma)) \leq \dim(\mathbb{E}^\sigma)$ and we have equality thanks to the fact that the dimensions must sum up to n. The invariance of the spaces along solutions come from the fact that a real solution stay real in the time and $\mathbb{E}^\sigma$ is invariant. □

With this descrption, we understand better what could look like a $\mathbb{R}^3$ example (or even in more dimensions). There will be some planes or lines where the solutions act perfectly like in the $\mathbb{R}^2$ case. For graphics in $\mathbb{R}^3$, see [7] page 34 for instance.

Remark. By taking the real part of the spaces (just like the imaginary part) and summing them, we can construct all the real space and the induced real solutions. However, note that taking the real part of a basis will not a priory give a real basis. Indeed the complex basis can use complex coefficient to build all real solutions while a real basis can only use real coefficients and the real part of a product is not just the product of the real parts.

We see that the eigenspaces seem to act independently from each other and the general comportment of solutions in them, seem to be somewhat uniform. A change of an eigenvalue can change totally the stability of a system; the flow $e^{At}x$ is continuous on A, but how does the uniformity continuity change along t and x? We can show that actually, in the hyperbolic case, the dimensions of the eigenspaces do not change when we chose a second matrix near enough the first one, in some sense the general stability structure depends continuously on A:

Theorem 8. *For $\sigma = s, u$, the map $A \mapsto$ dim $\mathbb{E}^\sigma(A)$ is locally continuous on hyperbolic matrices, and so locally constant.*

Proof. We treat only the stable eigenspace since $\dim \mathbb{E}^u(A) = \dim \mathbb{E}^s(-A)$. It is a classical but non trivial analysis result that the roots of a polynomial are continuous, so we prove that if we can find a neighbourhood where the roots stay stable, the structure of the generalised eigenspaces stay the same. Since no eigenvalue is in the imaginary line, we can surround all the stable eigenvalues by a positively oriented simple closed rectifiable curve γ, that does not cross any eigenvalue, the imaginary line neither. Since the characteristic polynomial p_A vanishes exactly on the eigenvalues, and is holomorphic, we have by the residues that

$$\frac{1}{2\pi i}\int_\gamma \frac{p'_A}{p_A} = \dim \mathbb{E}^s(A).$$

Indeed, if p_A has a root λ of order k, the residue of p'_A/p_a in λ is just k. Since p_A is a polynomial, the order of a zero is just the its multiplicity, and it correspond then to the dimension of the generalised eigenspace of this eigenvalue. The space $\mathbb{E}^\sigma$ is a direct sum of the generalised eigenspaces of stable eigenvalues, whose dimensions sum up to the dimension of the stable eigenspace itself. Since the coefficients of p_A are polynomials of the coefficients of A, they change continuously with A. We can suppose that A varies in a small neighbourhood so p_A still do not vanish in the curves γ (that does not change with A). As a result, the formula is well defined around A and is continuous with respect to p_A and hence to A. ☐

As we see, the stability structure of the system is the same for matrices in a neighbourhood. But more generally, if two matrices A and B are far from each other, but still have the same dimensions of eigenspaces, do them have similarities? We introduce the following tool of comparison, that match all the trajectories from both system in a continuous way.

Definition 10. Two flows ϕ and ψ on the space $\mathbb{R}^n$, are said *topologically conjugate*, if there exist a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ such that $h \circ \phi_t = \psi_t \circ h$, i.e. the image of a trajectory is a trajectory.

Lemma 6. *If a matrix is stable with maximum real part of eigenvalues $-a < 0$, there exists a scalar product $\langle ., . \rangle_a$ such that $\langle x, Ax\rangle_a < -a\|x\|_a^2$.*

Proof. Linear algebra result, see [3] at page 145 for a proof. ☐

Theorem 9 (Hartman and Grobman)**.** *Two stable systems on the same space are topologically conjugate.*

Proof. To find a homeomorphism h, we first search one on a subset of the space that content points that are part of all non trivial solutions and then we will extend it along the trajectories. The idea is to find a kind of sphere that trajectories cross only one time. The Lemma 6 exhibit a special scalar product that seem to be related to the problem. We consider two matrices A, B and their associated linear systems. We set the spheres S_A and S_B associated to the specials products $\langle,\rangle_A$, $\langle,\rangle_B$ of the matrices A and B. We show that trajectories cross them exactly one time by showing that the norm chosen is monotone along the solutions:

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\|_A = \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{x},\mathbf{x}\rangle_A = 2\langle \mathbf{x},\dot{\mathbf{x}}\rangle_A = 2\langle \mathbf{x}, A\mathbf{x}\rangle_A \leq -a\|\mathbf{x}\|_A^2$$

For a scalar $a > 0$. So before applying Grönwal and conclude the exponential convergence to zero, we see that this convergence is strictly monotone definite in the sense of the norm $\|.\|_A$. Thus, $\mathbf{x}$ only passes in S_A one time. The same hold for B by symetry of the situation. We can set now an homeomorphism

$$\begin{aligned} h_0 : S_A &\to S_B \\ \mathbf{x} &\mapsto \mathbf{x}/\|\mathbf{x}\|_B \end{aligned}$$

It is well defined because the image is obviously in S_B, $\mathbf{x}$ is never 0 and the denominator never vanishes. The continuity come from continuity of the norm. By symmetry of the situation its inverse should be $g_0 : \mathbf{x} \mapsto \mathbf{x}/\|\mathbf{x}\|_A$. Indeed for a $\mathbf{x} \in S_B$,

$$(h_0 \circ g_0)(\mathbf{x}) = h_0(\frac{\mathbf{x}}{\|\mathbf{x}\|_A}) = \frac{\frac{\mathbf{x}}{\|\mathbf{x}\|_A}}{\|\frac{\mathbf{x}}{\|\mathbf{x}\|_A}\|_B} = \frac{\mathbf{x}}{\|\mathbf{x}\|_B} = \mathbf{x}$$

so $h_0 \circ g_0 = \mathrm{Id}_{S_B}$ and by the symetry of the situation, $g_0 \circ h_0 = \mathrm{Id}_{S_A}$. The function h_0 is a continuous bijection, an homeomorphism.

We construct now the final h. For this we want to take a initial point, find the intersection of its trajectory with the sphere, go to the other sphere, and return go back in the trajectory as much as we came in the first sphere. For this we need to know for any point $\mathbf{x}$, how many time $\tau(x)$ does it take to go to the sphere, *i.e.* $e^{A\tau(x)}x \in S_A$. The value $\tau(x)$ is uniquely defined by the choice of the sphere. We set h by

$$h(\mathbf{x}) = \begin{cases} 0 \text{ if } \mathbf{x} = 0 \\ e^{-B\tau(\mathbf{x})}h_0(e^{A\tau(\mathbf{x})}\mathbf{x}) \text{ else} \end{cases}$$

We prove first that it has indeed the relation $h(e^{At}\mathbf{x}) = e^{Bt}h(\mathbf{x})$. We note that $\tau(e^{At}x) = \tau(x) - t$ because

$$\phi(e^{At}x, \tau(x) - t) = e^{A(\tau(x)-t)}e^{At}x = e^{A\tau(\mathbf{x})}x \in S_A.$$

When $\mathbf{x}$ is not the origin we compute,

$$\begin{aligned} h(e^{At}\mathbf{x}) &= e^{-B\tau(e^{At}\mathbf{x})}h_0(e^{A\tau(e^{At}\mathbf{x})}e^{At}\mathbf{x}) \\ &= e^{-B(\tau(\mathbf{x})-t)}h_0(e^{A(\tau(\mathbf{x})-t)}e^{At}\mathbf{x}) \\ &= e^{Bt}e^{-B\tau(\mathbf{x})}h_0(e^{A\tau(\mathbf{x})}\mathbf{x}) \\ &= e^{Bt}h(\mathbf{x}) \end{aligned}$$

If we write it $h = h_{AB}$, we show that it has an inverse $g = h_{BA}$. By looking at the definition of $h(\mathbf{x})$, we see that

$$e^{B\tau(\mathbf{x})}h(\mathbf{x}) = h_0(e^{A\tau(\mathbf{x})}\mathbf{x}) \in S_B$$

so $\tau(h(\mathbf{x})) = \tau(\mathbf{x})$. We compute that for all $\mathbf{x} \in S_A$

$$\begin{aligned} g \circ h(\mathbf{x}) &= e^{-A\tau(h(\mathbf{x}))}g_0(e^{B\tau(h(\mathbf{x}))}h(\mathbf{x})) \\ &= e^{-A\tau(\mathbf{x})}g_0(e^{B\tau(\mathbf{x})}h(\mathbf{x})) \\ &= e^{-A\tau(\mathbf{x})}g_0(e^{B\tau(\mathbf{x})}e^{-B\tau(\mathbf{x})}h_0(e^{A\tau(\mathbf{x})}\mathbf{x})) \\ &= e^{-A\tau(\mathbf{x})}g_0(h_0(e^{A\tau(\mathbf{x})}\mathbf{x})) \\ &= e^{-A\tau(\mathbf{x})}e^{A\tau(\mathbf{x})}\mathbf{x} \\ &= x \end{aligned}$$

and $g \circ h = \mathrm{Id}_{S_A}$. By symmetry of the situation, $h \circ g = \mathrm{Id}_{S_B}$ too and h is a bijection with inverse g.

It remind to show the continuity of h, which is the same for g. Apart of the origin, everything is continuous if we can prove it for τ. We put the case orgine away and will do it separately and for now we can use the implicit function theorem on $(\mathbf{x}, t) \mapsto \|e^{At}\mathbf{x}\|_A - 1$, which is continuous. The partial derivative with respect to t and in a point $(\mathbf{x}_0, t_0)$ is

$$\partial_t(\|e^{At_0}\mathbf{x}_0\|_A - 1) = 2\langle e^{At_0}\mathbf{x}_0, Ae^{At_0}\mathbf{x}_0\rangle < -a\|e^{At_0}\| < 0$$

and is inversible because monotone. We must have then a continuous map $\tilde{\tau}$ in a neighbourhood of $\mathbf{x}_0$ such that $\|e^{A\tilde{\tau}(\mathbf{x})}\mathbf{x}\|_A - 1 = 0$. By unicity of τ (and $\tilde{\tau}$ actually), we know that in this neighbourhood, $\tau = \tilde{\tau}$ is continuous. Since $\mathbf{x}_0$ was chosen arbitrarily, τ is locally continuous and so continuous. For $\mathbf{x} \neq 0$, $h(\mathbf{x})$ is a composition of continuous function and is itself continuous.

For the continuity in $\mathbf{x} = 0$, we take a sequence $(\mathbf{x}_k)_k \to 0$ and we have necessarily that $\tau(\mathbf{x}_k)$ goes to $-\infty$. If not, then we must have a bounded sub-sequence $\mathbf{x}_{k_j}$ since τ is necessarily negative in the unit ball. The result is that we can suppose this sub-sequence converging to a time τ^*, and get by the continuity of the flow

$$S_A \ni \phi(\mathbf{x}_{k_j}, \tau(\mathbf{x}_{k_j})) \xrightarrow[j\to\infty]{} \phi(0, \tau^*) = 0,$$

which is a contradiction with the closure of the sphere. We use $\tau(\mathbf{x}_k) \to -\infty$ to get

$$\|h(\mathbf{x}_k)\|_B = \|e^{-B\tau(\mathbf{x}_k)}h_0(e^{A\tau(\mathbf{x}_k)}\mathbf{x}_k)\|_B \leq \|e^{-B\tau(\mathbf{x}_k)}\|_B \to 0$$

because B is stable. Finally, h is continuous in the origin and everywhere and it is a homeomorphism. □

This theorem give us a total correspondence between two flows, and explain how the eigenvalues are decisive in the structure of the stability. More than that, it show that the stability itself, which is a local property, is determinant for the all flow. This exhibit the importance of the study of stability from a topological point of view, which is somewhat the qualitative way to see the comportment of solutions.

We can actually show that the unstable and saddle cases have a similar result, built on last theorem. We present a partial proof to see what is going on, see [8] page 113 for a detailled proof.

Corollary 5. *Two hyperbolic linear systems with stable eigenspace of same dimension, are topologically conjugates.*

Proof. Let A and B be unstable matrices with induced linear flows ϕ and ψ. The linear systems of matrices $-A$ and $-B$ are stable, they are topologically conjugate by a homeomorphism h. Now simply

$$h \circ \phi_t(\mathbf{x}) = h(e^{At}\mathbf{x}) = h(e^{-A(-t)}\mathbf{x}) = e^{-B(-t)}h(\mathbf{x}) = e^{Bt}h(\mathbf{x}) = \phi_t \circ h(\mathbf{x}),$$

so the same h works for the unstable case. We do not prove here the saddle case, but we present a sketch of proof: When the two matrices are saddle, but with stable eigenspaces of the same dimension k and m, we can decompose the space like

$$\mathbb{R}^n = \mathbb{E}^s(A) + \mathbb{E}^u(A) = \mathbb{E}^s(B) + \mathbb{E}^u(B)$$

Finite and same dimensional spaces are homeomorphic by a bijective linear map so we can use the theorem in each eigenspace independently as it was $\mathbb{R}^d$ with d the dimension of the stable or unstable eigenspace. We would get two homeomorphism $h_s : \mathbb{E}^s(A) \to \mathbb{E}^s(B)$ and $h_u : \mathbb{E}^u(A) \to \mathbb{E}^u(B)$. To define a h in the all space we can use the projections π_s and π_u of the decomposition $\mathbb{R}^n = \mathbb{E}^s(A) + \mathbb{E}^u(A)$ and we would get something like $h = h_s \circ \pi_s + h_u \circ \pi_u$ that will give us a homeomorphism. □

We have studied the stability of linear dynamical system under multiples point of view. One can think that Linear systems are a very specific situation and less intricate. It is indeed in all generality, but we present here a first link to the non-linear case, whose stability appears to be determined in some way by a linear associated system. This motivate the study of linear system and their possible applications to non-linear dynamics.

Theorem 10 (Linearization, Hartman–Grobman)**.** *Consider a system* $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ *such that* $\mathbf{F}$ *is* C^1*, and* $\mathbf{x}^*$ *a fixed point. If the linearized system* $\dot{\mathbf{x}} = D\mathbf{F}(\mathbf{x}^*)\mathbf{x}$ *is stable, then* $\mathbf{x}^*$ *is asymptotically stable for the non linear system.*

Proof. If the fixed point is not the origin, we set $\mathbf{y} = \mathbf{x} - \mathbf{x}^*$ which satisfy $\dot{\mathbf{y}} = \dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{y} + \mathbf{x}^*)$. Its flow is just a translation of the flow of the initial problem, is therefore isometric to it, and has the corresponding fixed point in the origin. Since $\mathbf{F}$ is C^1, we write its Taylor expression around the origin:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(0) + D\mathbf{F}(\mathbf{x})\mathbf{x} + h(\mathbf{x})\mathbf{x} =: A\mathbf{x} + h(\mathbf{x})\mathbf{x}.$$

for some function $h(\mathbf{x}) \to 0$ when $\mathbf{x} \to 0$. We use now the special scalar product $\langle ., . \rangle_a$ related to the stable matrix A of Lemma 6 and the metric induced by it, which is equivalent to the euclidean metric. We get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{x}, \mathbf{x}\rangle_a &= 2\langle \mathbf{x}, \dot{\mathbf{x}}\rangle_a = 2\langle \mathbf{x}, \mathbf{F}(\mathbf{x})\rangle_a = 2\langle \mathbf{x}, A\mathbf{x} + h(\mathbf{x})\mathbf{x}\rangle_a = 2\langle \mathbf{x}, A\mathbf{x}\rangle_a + 2\langle \mathbf{x}, h(\mathbf{x})\mathbf{x}\rangle_a \\
&\leq -a\|\mathbf{x}\|_a^2 + 2\|\mathbf{x}\|_a^2 h(\mathbf{x}) \\
&= (2h(x) - a)\|\mathbf{x}\|_a^2. \\
&\leq -c\|\mathbf{x}\|_a^2
\end{aligned}$$

We can take $\|\mathbf{x}\|$ small enough such that $2h(x) - a$ is negative, let us say for any $a > \epsilon > 0$, $\|\mathbf{x}\| < \delta(\epsilon)$ implies $\alpha = 2h(x) - a < -\epsilon$. Then by Grönwall, $\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\|_a^2 < -\alpha\|\mathbf{x}\|_a^2$ implies

$$\|\mathbf{x}\|_a^2 < \|\mathbf{x}(0)\|_a^2 e^{-\alpha t} \to 0$$

exponentially when $t \to \infty$ and the system is attracting since it happen in a neighbourhood $B_{\|.\|_a}(0, \delta(\epsilon))$ of the origin. Note that taking a neighbourhood small enough, we can attain a exponential convergence of any rate $\alpha < a$. The L-stability come just from the fact that the convergence of the attractivity is direct and does not go out its chosen neighbourhood, namely a sphere for the metric $\|.\|_a$. This give us the asymptotic stability. □

The idea behind this proof is that we have defined a quantity, $\|\mathbf{x}\|_a$, and that we have showed that this quantity was always decreasing along the trajectories, like a potential energy. This kind of function is called a Lyapunov function, we will define it in the next chapter and see that it has nice general applications.

Chapter 2

Lotka-Volterra equations and modifications

In this chapter, we present a model for the evolution of the population between preys and predators. For predator and prey populations, we define a function that quantify the size of the population with respect to time. The goal is to motivate a choice of differential equations that will describe the interaction and give a possible evolution of the two populations. Lotka-Volterra equations are well known equation in mathematical biology. To get straight to the point, the equations are

$$\begin{cases} \dot{x} = x(\alpha - \beta y), \\ \dot{y} = y(-\gamma + \delta x). \end{cases}$$

where x and y represent the size of the population of preys, and predators respectively. The parameters $\alpha, \beta, \gamma, \delta$ are positive scalars. In the first chapter we presented a pragmatic and mathematical analysis of the linear systems. Here, we propose first to speak about the qualitative understanding of the equation. We will progressively motivate the choice behind this modelisation by showing equations related to it, and then assert a modification on the equations that will give us more possible equations but for which the same study still works as in the classic system.

2.1 Motivation

We go back to the basics and remind that for a function x of one variable t, the notations

$$x'(t) = \frac{\mathrm{d}}{\mathrm{d}t}x(t) = \dot{x}(t) = \lim_{h \to 0} \frac{x(t+h) - x(t)}{h}$$

are the derivative when the limit exist. Qualitatively it indicates the amount of change of the function with respect to the variable. This amount is absolute with respect to the size of x and doesn't depend relatively on it. In consequence, when x is non null which is an assumption we will always make for initial value, $\frac{\dot{x}}{x}$ quantify the relative rate of change of the function. This could be in our situation the mean number of descendant of an individual. For example if a specie reproduce itself always with the same speed no matter the population or the environment, we can say that their grow rate is a constant c:

$$\dot{x} = cx,$$

If the population is positive,

$$\frac{\dot{x}}{x} = \frac{d}{dt}(\log x) = c$$

and then by integrating,

$$\log x(t) = ct + \log x(0),$$

giving us $x(t) = x(0)e^{ct}$, the exponential growth. This is quite intuitive, if a population double each step time the general formula is of powers of two, the population will just grow indefinitely. Alternatively if c is taken negative, it mean we have loss of individuals, and the population decrease.

Obviously, this is doesn't encapsulate the reality as the function grow very fast forever. The growth rate must decrease as the population increase. This come from multiple complex reasons such as competition due to environment capacities in food, space etc. For now we suppose it is from the simplest form, a linear growth rate $\alpha - \beta x$:

$$\dot{x} = x(\alpha - \beta x).$$

That gives us the logistic equation, where α represent the initial growth rate, and β how fast the growth rate slow down as the population size increase. Here we have one non trivial equilibrium when $\dot{x}(t) = 0$ i.e. when $x(t) = \alpha/\beta = x_*$. If not, we remark that $\dot{x} > 0$ when $x < x_*$, and $\dot{x} < 0$ when $x < x_*$, meaning that x_* seems stable. Indeed we do the computations and

$$\alpha = \dot{x}\frac{x_*}{x(x_* - x)} = \frac{\dot{x}}{x} + \frac{\dot{x}}{x_* - x} = (\log x + \log|x_* - x|)'.$$

By integrating from 0 to t we get

$$\alpha t = \log(x(t)) - \log|x_* - x(t)| - (\log(x_0) - \log|x_* - x_0|) = \log\frac{x(t)(x_* - x_0)}{x_0(x_* - x(t))}$$

and rearranging terms

$$x(t) = \frac{x_0 x_* e^{\alpha t}}{x_* + x_0(e^{\alpha t} - 1)}$$

which is well defined for $t \in \mathbb{R}$ if $0 < x_0 < x_*$, and for $t \in [1/\alpha \log(1 - x_*/x_0), \infty]$ if $0 < x_* < x_0$. We see that in both cases $x(t) \to x_*$ as $t \to \infty$.

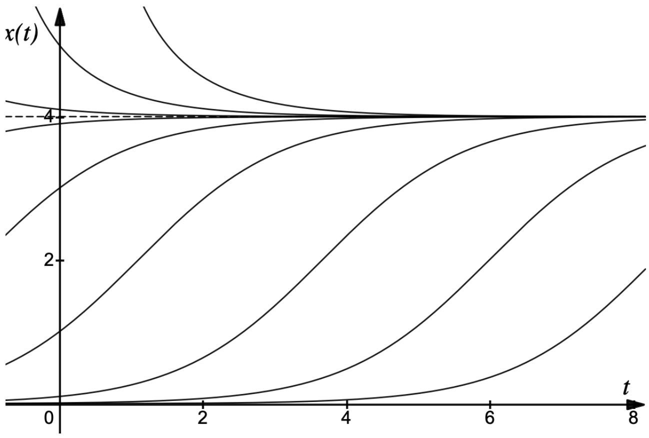

Figure 2.1: Logistic equation for some initialisations, with $x_* = 4$ and $\alpha = 1$.

In conclusion for this logistic growth, the population will always stabilise in the direction of a unique non trivial equilibrium.

Now we want to introduce a second specie, the predator which alter the growth of the prey population, the same way we made the size of the population itself affect the rate of change. That mean that the growth rate will decrease in function of the prey, let's say linearly:

$$\dot{x} = x(\alpha - \beta y).$$

In the opposite, the growth rate of the predator increase together with the population. But it decrease without them:

$$\dot{y} = x(-\gamma + \delta x).$$

This is the Lotka-Volterra equations for a prey-predator system. When we added the affect of the other species, with didn't keep the affect of the population on itself. We can consider both of these affects and add a term in each growth rate:

$$\begin{cases} \dot{x} &= x(\alpha - \beta y - \mu x) \\ \dot{y} &= y(-\gamma + \delta x - \nu y). \end{cases}$$

We obtained the equations we wanted to investigate. Let us talk deeper about them and about the possible modifications:

The idea is that we want to be able to modify the way the size of the species affect the rate. Traditional Lotka-Volterra equation consider the rate of growth as linear with respect to the populations, let's change that. We use functions F and G that are strictly monotonic, hence bijective, and increase from 0 to ∞ to replace the linearity:

$$\begin{cases} \dot{B} = F(B)(\alpha - \beta G(\Phi)), \\ \dot{\Phi} = G(\Phi)(-\gamma + \delta F(B)). \end{cases} \tag{2.1}$$

Here again, $\alpha, \beta, \gamma, \delta$ are positive scalars.

Note that $F(B)$ and other similar forms are an abuse of notation and mean $F \circ B$. When using the dot notation we always mean derivation with respect to time. So if a function has not time as an argument, for example F, then $\dot{F}(B) = (F \circ B)'$ and $\dot{F}(B(t)) = (F \circ B)'(t) = \frac{\mathrm{d}}{\mathrm{d}t}(F(B(t)))$.

Let us consider a real case where these considerations are meaningful. The paper [2] proposes marine phages and bacteria. In this case bacteria doesn't seem to be limited by the environment and their size. We can suppose the affect of itself as negligible against the affect of the phages. With the modification, the effective and the physical size of the phage population are not the same. It seem that one of them quantify as the power p of the other. The reason behind this choice is that in the traditional Lotka-Volterra equations, we assume one predator meets one prey at a time. Here laboratory tests tend to say that the important meetings are when two (or maybe three) phages meet a bacteria.

This consideration make the power of $p = 2$ appear on the value of the page population in the equation. Indeed we first wanted to wrote $\dot{\Phi} = \Phi(-\gamma + \delta B)$ meaning that the amount of change is the interactions between phages and bacteria, something like phages$\times$bacteria. But we need two phages, so something like phages$\times$phages$\times$bacteria which gives $\dot{\Phi} = \Phi^2(-\gamma + \delta B)$. A similar thinking motivate the writing of $\dot{B} = B(\alpha - \beta\Phi^2)$. In other words, we can say that they are hunting teams of p phages. For the mathematical study, we will keep general functions F and G that we can change in function of these kind of considerations on the affect.

2.2 Mathematical study

We have a non trivial equilibrium where $\dot{B} = \dot{\Phi} = 0$:

$$B_* = F^{-1}(\gamma/\delta), \quad \Phi_* = G^{-1}(\alpha/\beta)$$

Note that if one of the two populations is constant, then the other must be constant too and they must be in this total equilibrium. Let's rewrite (2.1) using this notation:

$$\begin{cases} \dot{B} = \beta F(B)(G(\Phi_*) - G(\Phi)), \\ \dot{\Phi} = \delta G(\Phi)(-F(B_*) + F(B)). \end{cases} \tag{2.2}$$

Now we can study the sign of the derivatives $\dot{B}$ and $\dot{\Phi}$ and draw a phase plane. The positive values B_* and Φ_* divide the positive plane $\mathbb{R}^2_+$ in four regions as follow:

$$\begin{array}{ll} \dot{B} < 0 \text{ and } \dot{\Phi} < 0 & \dot{B} < 0 \text{ and } \dot{\Phi} > 0 \\ \dot{B} > 0 \text{ and } \dot{\Phi} < 0 & \dot{B} > 0 \text{ and } \dot{\Phi} > 0. \end{array}$$

Trajectories seem to turn around the center of equilibrium but we cannot see if the solutions are converging to the fixed point, are cyclic, or even diverging. To test this, we compute the linearlized system in purpose to use the Theorem 10 and possibly obtain asymptotic stability. If not we will not be able to conclude for the moment, but this will give a hint to search more.

$$\begin{aligned} & D_{B,\Phi} \begin{pmatrix} F(B)(\alpha - \beta G(\Phi)) \\ G(\Phi)(-\gamma + \delta F(B)) \end{pmatrix} \\ = & \begin{pmatrix} F'(B)(\alpha - \beta G(\Phi)) & -F(B)\beta G'(\Phi) \\ G(\Phi)\delta F'(B) & G'(\Phi)(-\gamma + \delta F(B)) \end{pmatrix}. \end{aligned}$$

We evaluate in the equilibrium and obtain

$$\begin{pmatrix} 0 & -F(B_*)\beta G'(\Phi_*) \\ G(\Phi_*)\delta F'(B_*) & 0 \end{pmatrix}.$$

This matrix has obviously imaginary eigenvalues because the characteristic polynomial is just

$$\lambda^2 + F(B_*)\beta G'(\Phi_*)G(\Phi_*)\delta F'(B_*)$$

and vanishes in $\pm i\sqrt{F(B_*)\beta G'(\Phi_*)G(\Phi_*)\delta F'(B_*)}$. This means that the linearized system is elliptic, and we cannot conclude with the theorem we have. However, this gives a hint, the solutions might be cyclic, just like in the linearized version.

For this we search a first integral by taking the cross product of (2.2) and dividing by $F(B)G(\Phi)$:

$$\begin{aligned} 0 &= \big(\dot{B}\delta G(\Phi)(-F(B_*) + F(B)) - \dot{\Phi}\beta F(B)(G(\Phi_*) - G(\Phi))\big)\frac{1}{F(B)G(\Phi)} \qquad (2.3) \\ &= \delta(-\frac{\dot{B}}{F(B)}F(B_*) + \dot{B}) - \beta(\frac{\dot{\Phi}}{G(\Phi)}G(\Phi_*) - \dot{\Phi}) \\ &= \Big(\delta(-P(B)F(B_*) + B) - \beta(Q(\Phi)G(\Phi_*)) - \Phi)\Big)' \end{aligned}$$

where P and Q are primitives of $1/F$ and $1/G$, and exist because F and G are continuous. This give us a conserved quantity

$$\begin{aligned} V(B,\Phi) &= \delta\big(-P(B)F(B_*) + B\big) - \beta\big(Q(\Phi)G(\Phi_*) - \Phi\big) \\ &= V(B(0),\Phi(0)) \\ &=: V_0 \end{aligned}$$

Now we know then that $(B,\Phi) \in V^{-1}(\{V_0\})$, a closed set as V is continuous. To understand the level set, we compute the gradient:

$$\nabla V(B,\Phi) = \big(\delta(-\frac{F(B_*)}{F(B)} + 1), -\beta(\frac{G(\Phi_*)}{G(\Phi)} - 1)\big)^\top.$$

It never vanishes, except in the equilibrium and is continuous. As a result, the level set doesn't have interior otherwise the function would be constant and the gradient null. Since we have the local unicity of the solution, the level set is actually a closed line, *i.e.* the solutions are in an closed orbit. The following graph shows the trajectory of some solutions, and the direction in each part of the first quadrant:

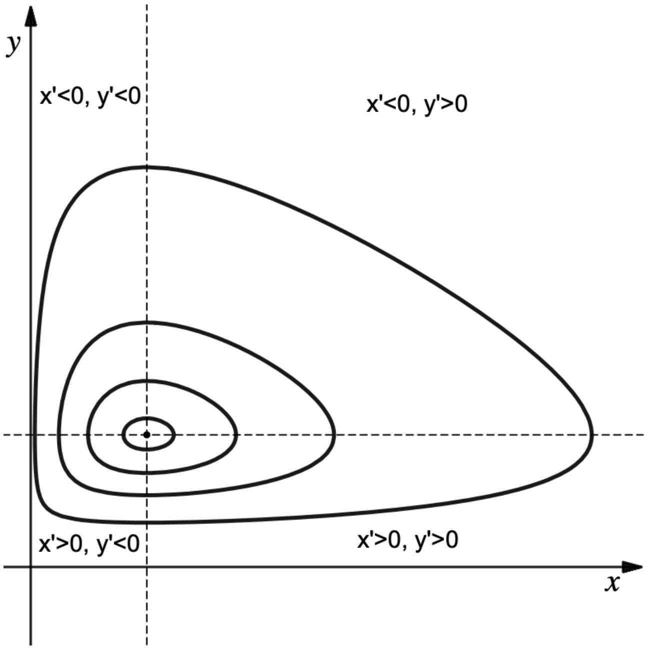

Figure 2.2: Some trajectories of the modified Lotka-Volterra system.[1]

Where we take $F = \text{Id}$, $G(\Phi) = \Phi^2$, and with a fixed point $(3.04, 3.4)$. We see that as the formula of the gradient shows and the system itself, the extrema of Φ and B are when the other function is on the equilibrium. Now that we know that it's cyclic with a period τ, we derive a modified Volterra principle by integrating these qualities obtained from (2.2):

$$\dot{B}/F(B) = \beta(G(\Phi_*) - G(\Phi)).$$

the left side gives

$$\int_0^\tau \frac{\dot{B}}{F(B)} = \int_0^\tau (P(B))' = P(B(\tau)) - P(B(0)) = P(B(0)) - P(B(0)) = 0,$$

and then the rigth side

$$0 = \int_0^\tau \beta(G(\Phi_*) - G(\Phi)) = \tau\beta G(\Phi_*) - \int_0^\tau \beta G(\Phi)$$

which implies that

$$G(\Phi_*) = \frac{1}{\tau}\int_0^\tau G(\Phi).$$

Note that if $G = \text{Id}$, this says that the mean of the population Φ during time is $F(B_*)$. By the same argument with the second equation, we obtain that similarly

$$F(B_*) = \frac{1}{\tau}\int_0^\tau F(B).$$

[1]Interested readers can find here an online interactive graphic that we made for the Lotka-Volterra system, and here a version for the modified system.
(The urls are *https://www.desmos.com/calculator/xsoo8fqwth* and *https://www.desmos.com/calculator/go5ata2oee* .)

In Section 2.1, we explained how the own size of a population can affect its growth. We presented the logistic growth consequence of this equation:

$$\dot{x} = x(\alpha - \beta x).$$

We add then a term in the rates of (2.1) to modelise the fact that the growth rate of a population decrease with the size of population due to environment capabilities or competition. Again we use F and G to quantify their importance:

$$\begin{cases} \dot{B} &= F(B)(\alpha - \beta G(\Phi) - \mu F(B)), \\ \dot{\Phi} &= G(\Phi)(-\gamma + \delta F(B) - \nu G(\Phi)). \end{cases} \tag{2.4}$$

with new positive scalars μ and ν. We search for non trivial equilibrium where $\dot{B}_{**} = \dot{\Phi}_{**} = 0$ and obtain a linear system in $F(B_{**})$ and $G(\Phi_{**})$:

$$\begin{cases} -\alpha = -\mu F(B_{**}) - \beta G(\Phi_{**}), \\ \gamma = \delta F(B_{**}) - \nu G(\Phi_{**}), \end{cases}$$

which give

$$\begin{aligned} \mu\gamma - \delta\alpha &= \mu\delta F(B_{**}) - \mu\nu G(\Phi_{**}) - \delta\mu F(B_{**}) - \delta\beta G(\Phi_{**}) \\ &= (-\mu\nu - \delta\beta) G(\Phi_{**}) \end{aligned}$$

and then as $\mu\nu + \delta\beta > 0$ and supposing $\alpha\delta > \gamma\mu$

$$G(\Phi_{**}) = \frac{\alpha\delta - \gamma\mu}{\beta\delta + \nu\mu}, \quad \Phi_{**} = P^{-1}\Big(\frac{\alpha\delta - \gamma\mu}{\beta\delta + \nu\mu}\Big). \tag{2.5}$$

Similarly,

$$F(B_{**}) = \frac{\beta\gamma + \nu\alpha}{\beta\delta + \nu\mu}, \quad B_{**} = F^{-1}\Big(\frac{\beta\gamma + \nu\alpha}{\beta\delta + \nu\mu}\Big). \tag{2.6}$$

Here we cannot derive a first integral like we did in (2.3) by separating variables B and Φ. Instead, we want to test the stability of (B_{**}, Φ_{**}). First we can try to use the theorem of linearization to check the possible asymptotic stability of the system:

$$\begin{aligned} & D_{B,\Phi} \begin{pmatrix} F(B)(\alpha - \beta G(\Phi) - \mu F(B)) \\ G(\Phi)(-\gamma + \delta F(B) - \nu G(\Phi)) \end{pmatrix} \\ = & \begin{pmatrix} F'(B)(\alpha - \beta G(\Phi) - \mu F(B)) - F(B)\mu F'(B) & -F(B)\beta G'(\Phi) \\ G(\Phi)\delta F'(B) & G'(\Phi)(-\gamma + \delta F(B) - \nu G(\Phi)) - G(\Phi)\nu G'(\Phi) \end{pmatrix}. \end{aligned}$$

We evaluate in the equilibrium and obtain

$$\begin{pmatrix} -F(B_{**})\mu F'(B_{**}) & -F(B_{**})\beta G'(\Phi_{**}) \\ G(\Phi_{**})\delta F'(B_{**}) & -G(\Phi_{**})\nu G'(\Phi_{**}) \end{pmatrix}.$$

Now we could replace $F(B_{**})$ and $G(\Phi_{**})$ by their expression (2.6) and (2.5), and compute the characteristic polynomial. But we would still have the the expressions $F'(B_{**})$ and $G'(\Phi_{**})$, that we know to be positive but nothing else. The polynomial isn't in a simple form and the computation of the eigenvalues become excessively cumbersome, because of the number of parameters and the hypothesis on them. More simply we show that a 2×2 matrix with the same signs of this one is always stable. Indeed all the factors are positives if we suppose that $F(B_{**})\beta$ and $G'(\Phi_{**})$ aren't null, and we can write the matrix like

$$\begin{pmatrix} -a & -b \\ c & -d \end{pmatrix}$$

with characteristic polynomial $\lambda^2 + (a+d)\lambda + ad + bc$ of discriminant $(a+d)^2 - 4(ad+bc)$. If the discriminant is negative, then the real part of the root will be $-(a+d)/2 < 0$, and else

$$\frac{-(a+d) \pm \sqrt{(a+d)^2 - 4(ad+bc)}}{2} < \frac{-(a+d) \pm (a+d)}{2} \leq 0$$

So the eigenvalues are stable and the linearized system too. By the theorem of linearization, the non-linear system is asymptotically stable in the equilibrium. In other words all solutions near enough this point tends to exponentially to it. However, we don't know how big is the basin of attraction.

For this we develop the theory of stability of Lyapunov. Consider a differential equation $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, such that there exist a unique solution for each initial point and for all $t \geq 0$. Such solutions are denoted by the flow ϕ, such that $t \mapsto \phi(\mathbf{x}_0, t)$ is the solution initialised at $\mathbf{x}_0$. We recall the following definitions

Definition 11. A fixed point $\mathbf{x}_*$ of is said *Lyapunov stable* or *L-stable* if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $\mathbf{x_0}$ and for all $t \geq 0$, $\|\mathbf{x_0} - \mathbf{x}_*\| < \delta$ implies $\|\phi(\mathbf{x_0}, t) - \mathbf{x}_*\| < \epsilon$.

Definition 12. A fixed point $\mathbf{x}_*$ of is said *attracting* if there exists a $\delta > 0$ s.t. for all $\mathbf{x_0}$, $\|\mathbf{x_0} - \mathbf{x}_*\| < \delta$ implies that $\phi(\mathbf{x_0}, t) \to \mathbf{x}_*$ as $t \to \infty$.

Definition 13. A fixed point which is L-stable and attracting is said *asymptoticaly stable*

Remark. We remind that have shown that these two notions are different in Remark 1.2.

Now we can introduce a tool that will be useful to understand the limit comportment of the trajectories and will a tool to proove asymptotic stability.

Definition 14. Assume $\mathbf{x}_*$ is a fixed point of a equation $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, and let a function $L : U \to \mathbb{R}$ defined in a neighbourhood of $\mathbf{x}_*$.
The function L is called a *weak Lyapunov function* if

$$L(\mathbf{x}) > L(\mathbf{x}_*)$$

and

$$\dot{L}(\mathbf{x}) := \frac{\mathrm{d}}{\mathrm{d}t} L(\phi(\mathbf{x}, t))|_{t=0} \leq 0$$

for all $\mathbf{x}$ in U.
The function L is called a *Lyapunov function* (or a *strict Lyapunov function*) provided that we have the strict inequality $\dot{L}(\mathbf{x}) < 0$ when $\mathbf{x} \neq \mathbf{x}_*$.

Theorem 11. *Let $\mathbf{x}_*$ a fixed point of the differential equation $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ and L a weak Lyapunov function on $U \ni \mathbf{x}_*$. Then $\mathbf{x}_*$ is L-stable. If L is a strict Lyapunov function, $\mathbf{x}_*$ is attracting too and thus L is asymptotically stable.*

Proof. L-stable: Let's suppose first that L is a weak Lyapunov function. For any solution $\mathbf{x}$, $\dot{L}(\mathbf{x}) \leq 0$ and $L(\mathbf{x})$ is decreasing. Let be $\epsilon > 0$. Up to taking it smaller, we suppose $B(\mathbf{x}_*, \epsilon) \subset U$. But L is continuous and $\partial B(\mathbf{x}_*, \epsilon)$ is compact, we can then define

$$m = \min_{\partial B(\mathbf{x}_*, \epsilon)} L \; > \; L(\mathbf{x}_*) = L_*,$$

and since L is continuous, there exists a $\epsilon > \delta > 0$ such that $L_* < L(\mathbf{x}) < m$ when $0 < \|\mathbf{x} - \mathbf{x}_*\| < \delta$. Now, for all solution starting in $B(\mathbf{x}_*, \delta) \subset B(\mathbf{x}_*, \epsilon)$, $L(\mathbf{x}) < m$ mean that $\mathbf{x}$ cannot go out of $B(\mathbf{x}_*, \epsilon)$, otherwise it would cross $\partial B(\mathbf{x}_*, \epsilon)$ in a certain $t > 0$, and we would have the contradiction $m > L(\mathbf{x}(0)) \geq L(\mathbf{x}(t)) = m$. Whave the L-stability.

Attracting when strict Lyapunov: Since $L(\mathbf{x})$ is decreasing and bounded over $L(\mathbf{x}_*)$, $L(\mathbf{x}(t))$ must have a limit when $t \to \infty$, let's say L_∞. This implies that $\dot{L}(\mathbf{x}(t)) \to 0$. Using first part, let be $\delta_2 > 0$

such that for all solution $\mathbf{x}$ starting in $B(\mathbf{x}_*, \delta_2)$, and for all $t > 0$, $\mathbf{x}(t) \in B(\mathbf{x}_*, \epsilon) \subset U$. Now since $\mathbf{x}(t)$ stay in a compact, there exists an accumulation point $\mathbf{z}$ and a sequence $(t_n)_n$ growing to infinity such that $\lim_{n\to\infty} \mathbf{x}(t_n) = \mathbf{z}$. Because $\dot{L}$ is continuous, the limit $\lim_{n\to\infty} \dot{L}(\mathbf{x}(t_n)) = 0$ is actually $\dot{L}(\mathbf{z}) = 0$ and by hypothesis, $\mathbf{z}$ must be $\mathbf{x}_*$. Now because of the first part, as $\mathbf{x}(t_n)$ approaches $\mathbf{x}_* = \mathbf{z}$, when n is big enough $\mathbf{x}(t_n)$ will be arbitrarily near from $\mathbf{x}_*$ for all $t > t_n$, assuring the convergence for all $t \to \infty$, and $\mathbf{x}_*$ is attracting.

From the two last parts we conclude that $\mathbf{x}_*$ is asymptotically stable. □

We now return to our original problem: test the stability of (B_{**}, Φ_{**}), the fixed point of

$$\begin{cases} \dot{B} &= F(B)(\alpha - \beta G(\Phi) - \mu F(B)), \\ \dot{\Phi} &= G(\Phi)(-\gamma + \delta F(B) - \nu G(\Phi)). \end{cases} \tag{2.7}$$

Generally speaking, there is no particular general method to find a Lyapunov function for generic equations. We note that concerning (2.1), we used a conserved quantity V to prove periodicity of the trajectory. Let us try to use a similar version of this V to obtain a quantity that would not be constant and for our goal, decreasing:

$$W(B, \Phi) = \delta\big(- P(B)F(B_{**}) + B\big) - \beta\big(Q(\Phi)G(\Phi_{**}) - \Phi\big).$$

We compute its derivative along a solution (B, Φ):

$$\begin{aligned}
\dot{W}(B, \Phi) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\delta\big(- P(B)F(B_{**}) + B\big) - \beta\big(Q(\Phi)G(\Phi_{**}) - \Phi\big)\Big) \\
&= \delta\big(- \frac{\dot{B}}{F(B)}F(B_{**}) + \dot{B}\big) - \beta\big(\frac{\dot{\Phi}}{G(\Phi)}G(\Phi_{**}) - \dot{\Phi}\big) \\
&= \delta\frac{\dot{B}}{F(B)}\big(- F(B_{**}) + F(B)\big) - \beta\frac{\dot{\Phi}}{G(\Phi)}\big(G(\Phi_{**}) - G(\Phi)\big) \\
&= \delta(\alpha - \beta G(\Phi) - \mu F(B))\Delta_F + \beta(-\gamma + \delta F(B) - \nu G(\Phi))\Delta_G \\
&= \delta(\mu F(B_{**}) + \beta G(\Phi_{**}) - \beta G(\Phi) - \mu F(B))\Delta_F + \beta(-\delta F(B_{**}) + \nu G(\Phi_{**}) + \delta F(B) - \nu G(\Phi))\Delta_G \\
&= \delta(-\mu\Delta_F - \beta\Delta_G)\Delta_F + \beta(\delta\Delta_F - \nu\Delta_G)\Delta_G \\
&= -\delta\mu\Delta_F^2 - \beta\nu\Delta_G^2,
\end{aligned}$$

with $\Delta_F = F(B) - F(B_{**})$ and $\Delta_G = G(\Phi) - G(\Phi_{**})$. As a result, $\dot{W}(B, \Phi)$ is null only on (B_{**}, Φ_{**}), otherwise it is strictly negative on $U = (\mathbb{R}_+^*)^2 \backslash \{(B_{**}, \Phi_{**})\}$. This is a Lyapunov function if we can prove that (B_{**}, Φ_{**}) is a strict minimum. For this we simply compute the gradient and the Hessian:

$$\nabla W(B, \Phi) = (\delta\big(- \frac{F(B_{**})}{F(B)} + 1\big), -\beta\big(\frac{G(\Phi_{**})}{G(\Phi)} - 1\big))^\top$$

vanishes only in the fixed point, and

$$H_W(B, \Phi) = \mathrm{diag}(\delta\frac{F(B_{**})F'(B)}{(F(B))^2}, \beta\frac{G(\Phi_{**})G'(\Phi)}{(G(\Phi))^2})$$

is a diagonal matrix and has its eigenvalues in the diagonal. Because F and G are supposed positive and strictly increasing, the diagonal is strictly since its diagonal is strictly positive and we have supposed not being on the axes. As a result, the hessian is positive definite and W is strictly convex, with a unique minimum on the only stationary point of W, the fixed point of the system. The function W is Lyapunov and the fixed point is asymptotically stable. For now, we don't know much about the size of the basin of attraction. Actually, in Theorem 11, we have used the neighbourhood of the L-stability only because solutions would be bounded. We would like to prove a global convergence result. First we recall the notion:

Definition 15. A fixed point $\mathbf{x}_*$ of $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ is said *globally attractive* on a set U, if for all $\mathbf{x_0} \in U$, $\phi(\mathbf{x_0}, t) \to \mathbf{x}_*$ as $t \to \infty$. In other words it is attractive without condition on the proximity of the initial point.

Definition 16. A fixed point $\mathbf{x}_*$ of $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ is said *globally asymptotically stable* on a set U, if it is globally attractive and L-stable. In other words it is asymptotically stable without condition on the proximity of the initial point for the convergence.

Now, if we look at our proof of Theorem 11 on Lyapunov functions, we see that the proximity that we needed to obtain convergence, was just in purpose to have a compact and invariant set. Without loss of generality we can take the δ_2 as big as we want, as long as $\phi(B(\mathbf{x}_*, \delta_2) \times \mathbb{R}_+)$ is bounded in U. A criteria on the boundedness of $L^{-1}([L(\mathbf{x}_*), L_0])$ will help us to obtain global attractivity:

Corollary 6. *Under the suppositions of Theorem 11, if we have in addition that for a scalar $L_0 > L(\mathbf{x}_*)$,*

$$U_{L_0} = \{\mathbf{x} \in U | L(\mathbf{x}) \leq L_0\} = L^{-1}\big([L(\mathbf{x}_*); L_0]\big)$$

is bounded and whose closure is contained in U, then $\mathbf{x}_*$ is globally asymptotically stable on U_{L_0}.

Proof. First of all, U_{L_0} is closed in the topology of U because of the continuity of L. But since its euclidean closure is in U, it's actually closed in the all space. In addition, it is bounded, and thus, compact. Finally, all $\mathbf{x}$ starting in U_{L_0} will stay in U_{L_0}, because of the monotonicity of $L(\mathbf{x})$ and we have the convergence by the same argument of Theorem 11, i.e. $L(\mathbf{x}(t)) \to L_\infty$, $\dot{L}(\mathbf{x}(t)) \to 0$, $\mathbf{x}(t_n) \to \mathbf{z}$ (because of the compacity of U_{L_0}), $\mathbf{z} = \mathbf{x}_*$, and $\mathbf{x}(t) \to \mathbf{x}_*$. This assure the global attractivity of $\mathbf{x}_*$ on U_{L_0}, and hence, the global asymptotic stability, since the L-stability was assured yet. □

Remark. If the hypothesis is true for all $L_0 > L(\mathbf{x}_*)$ then it is actually globally asymptotically stable everywhere on U. More generally, if a closed set $U_0 \subset U$ (possibly U itself), is invariant and each trajectory is bounded, we have global asymptotic stability on it.

We go back to our problem and recall that our Lyapunov function is

$$W(B, \Phi) = \delta\big(- P(B)F(B_{**}) + B\big) - \beta\big(Q(\Phi)G(\Phi_{**}) - \Phi\big)$$

defined in $(\mathbb{R}_+^*)^2$. For all $W_0 > W(B_{**}, \Phi_{**})$, under the condition that W tends to infinity when B or Φ tends to infinity, $U_{W_0} = W^{-1}([W(B_*, \Phi_*); W_0])$ must be bounded and closed in the interior of the positive quadrant. If not, we would have a sequence inside U_{W_0} converging to infinity or to the axes whose image by W is bounded which is impossible by supposition. If $F = G =$Id for instance, it is obviously the case because $P = G =$log which explode in zero and is dominated by Id at infinity. In all cases, the convexity of W and the strict min in the fixed point assures that at infinity the function indeed explodes. For the axes, we see that the form of $P(B)$ (similar for $Q(\Phi)$) can be written

$$P(B) = \int_1^B \frac{1}{F(B)} + c$$

and as $F(B) > F(0) = 0$, it explodes to $-\infty$ if $F(B)$ become small enough when B tends to zero. A sufficient condition for $F(B)$ is to be $O(B)$ for example. Then looking at the formula of W, it explodes to infinity too. The fixed point is globally asymptotically stable on the positive quadrant if F and G are $O(x)$ in zero.

Chapter 3

Modelisation about Covid-19

In the last chapter, we gained a good intuition how and why Lotka-Volterra equations model population interactions. We have seen what are the effects of adding terms in the growth rate of a equation. We recall that the general understanding of growth rate of a scalar function x_j is $\dot{x}_j/x_j$ when $x_j \neq 0$. This is mostly interesting when we have to deal with a equation of the form $\dot{x}_j = x_j f(x)$, because the growth rate is just $f(x)$. We would like to use these ideas of the second chapter to elaborate an equation that would simulate a pandemic. We do not have the audacity to say that the model will be a perfect modelisation of the problem and that it will even predict the future. It's actually not the goal, our main goal is to learn how to create a differential equation so that it fit the kind of comportment we need. The biological context is a motivation, the beginning of some real world application, the definition of a particular situation, which determine the qualitative considerations that our final solution will satisfy.

3.1 Quantities and considerations

First we present the quantities we will have to deal with and link together with differential equations. We suppose we have a population of a size P that will be confronted to a virus. The general idea is that the virus is the predator and the people are the prey. We have a number V of vulnerable people that have never been infected. We could have set a variable for virus but there is no such idea of quantity of virus, so we consider that infected people become part of the predators, and we set a number I of people that are infected. As it is the case most of the time for virus, there is a phenomena of immunity (at least partial) that follow an infection. As a result, the number R of resistant people is the number of people that have been infected in the past. We choose the case where infected people are almost immediately infectious after infection, so V is the number of infectious people too.

In conclusion we have a population P where each person can successively belong to each of the three parts V, I, and R. In other words we have $V + I + R = P$ and the scheme is $V \to I \to R$. As the number of people is supposed to be big, we can use real variable for all these quantities even if the populations are integers, and we make them vary along the free variable of time.

3.2 A Lotka-Volterra approach

We want to use the Lotka-Volterra model, just like if infected people where the predators of the vulnerable ones. The general form of this system is

$$\begin{aligned} \dot{V} &= V(\alpha + \beta I) \\ \dot{I} &= I(\gamma + \delta V). \end{aligned}$$

But what are the sign of the scalars? If there is no (-more) infected people, *i.e.* $I = 0$, nothing sould move so $0 = \dot{V} = \alpha V$ implies that $\alpha = 0$. The parameter β represent how the infected affect the vulnerable people. It's then negative because more infected implies less vulnerable. More precisely to explain the nature of the term in β, if the mean number of contacts of a person is μ then the mean of infection by unit of time is μ times the proportion of infected, I/P. This give indeed a term $-\mu/PI = \beta I$. For the second equality, the rate of infected people contains a term for the people that add from the vulnerable group, the same term as before but with opposite sign, $-\beta I$, and $\delta = \beta$. The other term models the people that are leaving the infected group for the resistant one. We can suppose that the period of infectivity is significantly smaller than the velocity of the pandemic, namely the change of size of the groups. So the number of people that leave the infected group at a time t is the same as the number of people that was going into it a bit sooner. We can suppose that in average, a fraction of the infected group leave it and so it give a term γI with $\gamma < 0$. With all these considerations we rename the paramters and take them all positive so we can write

$$\begin{aligned} \dot{V} &= -\beta V I \\ \dot{I} &= I(-\alpha + \beta V). \end{aligned}$$

We would like the resistant group to have the same oppostite rate as the term $-\alpha I$ that represent people leaving the infected group. So we have $\dot{R} = \alpha I$. We have finally

$$\begin{cases} \dot{V} = -\beta V I, \\ \dot{I} = I(-\alpha + \beta V) \\ \dot{R} = \alpha I. \end{cases} \tag{3.1}$$

We see that all of this seem coherent since

$$(V + I + R)' = \dot{V} + \dot{I} + \dot{R} = -\beta V I + I(-\alpha + \beta V) + \alpha R = 0$$

and the population is always conserved (we did not include deaths and demographic considerations). The last function R does not appear in the equations of the two first, so we can treat it aside: we study first V and I, and will deduce R as a consequence.

There is no isolated fixed point of the system (3.1), but all the axis where $I = 0$ is fixed. This show that once the virus is eradicate there is no more change. Here, the fixed axis $I = 0$ will introduce a zero eigenvalue in the linearization, and we cannot use the theorem of linearization because we need stable eigenvalues. More than that, because the fixed points are not isolated, it implies that they are not asymptotically stable because there exists non convergent solutions that start as near as we want. Instead of this method, we will use first integrals, *i.e.* constant quantities along time. We consider the system terminated if I or V reach zero. The function V is decreasing, and I change of direction when V passes α/β. We try to obtain a first integral like in the second chapter by separating the variables. Supposing $I \neq 0 \neq V$, we take the cross product of

$$\begin{aligned} \frac{\dot{V}}{V} &= -\beta I \\ \frac{\dot{I}}{I} &= -\alpha + \beta V. \end{aligned}$$

and obtain

$$0 = \frac{\dot{V}}{V}(-\alpha + \beta V) + \beta \dot{I} = \dot{V}(-\frac{\alpha}{V} + \beta) + \beta \dot{I} = (-\alpha \log(V) + \beta V + \beta I)',$$

so we have the constant quantity

$$L(V) := -\alpha \log(V) + \beta V + \beta I = -\alpha \log(V_0) + \beta V_0 + \beta I_0$$

$$0 = -\alpha \log(\frac{V}{V_0}) + \beta(V - V_0) + \beta(I - I_0)$$

and an explicit form for the curve :

$$I = \frac{\alpha}{\beta} \log(\frac{V}{V_0}) - (V - V_0) + I_0 = \frac{\alpha}{\beta} \log(\frac{V}{V_0}) - V + P. \quad (3.2)$$

Since its derivative is negative, V is decreasing and positive so it has a limit V_∞. Taking the limit of the expression we just get when time goes to infinity, we see that as I is bounded by P, V cannot goes to zero, and $V_\infty > 0$. For I, we see that $(I + V)' = -\alpha I \leq 0$, so similarly, the positive quantity $I + V$ must converge, and its derivative must converge to zero, so $I = (I + V)'/\alpha \to 0$. As a result, we can now take the limit of (3.2) and get

$$0 = \frac{\alpha}{\beta} \log(\frac{V_\infty}{V_0}) - V_\infty + P.$$

We show that we always have a point V_∞ such that this is satisfied. Indeed $L(V) - L(0) \to -\infty$ when $V \to 0$ and

$$L(P) - L(0) = \frac{\alpha}{\beta} \log(\frac{P}{V_0}) \geq \frac{\alpha}{\beta} \log(1) = 0.$$

Moreover,

$$\frac{d}{dV}(L(V) - L(0)) = \frac{\alpha}{\beta}\frac{1}{V} - 1$$

changes sign only one time in $V = \frac{\alpha}{\beta}$ where there is a max, so if there is two zeros, P must be between them and there is exactly one zero such that $V_\infty \leq P$. The max of I is

$$I_{max} = L(\alpha/\beta) - L(0) = \frac{\alpha}{\beta}(\log(\frac{\alpha}{\beta V_0}) - 1) + P$$

All in all, we have a pandemic (the number of infected increase) if and only if $V_0 > \alpha/\beta$, and in this case the infected group increase to I_{max}, before to go down to zero, and there will be $V_\infty > 0$ people that will never have been infected.

3.3 Inclusion of the vaccine in the equation

In this section, we modify our first model such that it include a group that has a better immunity against the virus than vulnerable people, but has never had been infected. For this we need to make two version of V and I depending if the people are vaccinated or not. For this we index them by v (vaccinated) and u (unvaccinated): V_u, V_v, I_u, I_v. In the first section, we have use parameters α and β to quantify the affect of the virus. Now, the vaccine have an effect on that, and reduce the susceptibility to infection, let's say it reduces by a factor β the capacity of the vulnerable group to be infected by a infectious; it reduces by a factor δ the capacity of the infected group to infect a vulnerable one (they may be differents parameters); it increase by a factor σ the capacity of the infected group to recover and become resistant. Using our understanding of each term of the equation in section 1, we establish a new equation and detail it after for clarity :

$$\begin{aligned}
\dot{V}_u &= -\beta V_u(I_u + \epsilon I_v) \\
\dot{V}_v &= -\delta\beta V_v(I_u + \epsilon I_v) \\
\dot{I}_u &= \beta V_u(I_u + \epsilon I_v) - \alpha I_u \\
\dot{I}_v &= \delta\beta V_v(I_u + \epsilon I_v) - \sigma\alpha I_v.
\end{aligned}$$

Here again, we use the concept of growth rate to justify the equation. The growth rate of the two vulnerable groups can be separated in two terms that represent the interactions with the vaccinated and unvaccinated infectious group. We have to take into account the fact that the vaccinated infectious group has an effect reduced by a factor ϵ and that the vaccinated vulnerable group has a vulnerability reduced by a factor δ. These two terms sum up and combine the two new factors, which justify the first two lines of the system. For the two last equations, we first have to add the term of people who leave the respective vulnerable group depending if they have vaccine or not. Then we have to add the term that represent the people that leave the infected group to become resistant. It is of the same form as before, but with a factor augmentation for the vaccinated ones. The general equation for the evolution of R is just the sum of the two terms we just described, it is $\dot{R} = -\alpha I_u - \sigma\alpha I_v$. As wanted, the total population $P = V_u + V_v + I_u + I_v + R$ is constant because as before, everything vanishes in $(V_u + I_u + I_v + V_v + R)' = 0$.

Now we try to do the same analysis as before. We have again for the set of fixed points, the 2 dimensional subspace where $I_u = I_v = 0$, and we don't have asymptotic stability on them. The constant quantity of motion doesn't come easily as before, but we see similarities. The V's are decreasing, and the I's may change direction during the process. We want to integrate the first two equations because we can separate at least the V's from the other functions. Using $\alpha_v = \sigma\alpha$, $\alpha_u = \alpha$, $\beta_v = \delta\beta$ ($\beta_u = \beta$), and a index w that go through $\{u, v\}$, we see that again, $(V_w + I_w)' = \alpha_w I_w$ and we get

$$(\log V_w)' = \frac{\dot{V}_w}{V_w} = -\beta_w(I_u + \epsilon I_v) = -\beta_w\Big(\frac{(V_u + I_u)'}{\alpha_u} + \epsilon\frac{(V_v + I_v)'}{\alpha_v}\Big)$$
$$= \Big(-\beta_w\Big(\frac{1}{\alpha_u}(V_u + I_u) + \frac{\epsilon}{\alpha_v}(V_v + I_v)\Big)\Big)'$$

This gives us a pair of constant quantities

$$L_w(V_u, V_v, I_u, I_v) = \log V_w + \beta_w\Big(\frac{1}{\alpha_u}(V_u + I_u) + \frac{\epsilon}{\alpha_v}(V_v + I_v)\Big). \tag{3.3}$$

We take a linear combination of them to make the second term disapear and obtain that $\delta \log V_u - \log V_v$ is constant too, so $V_v = \frac{V_v(0)}{V_u(0)^\delta} V_u^\delta$. This show how the vaccine affect the vulnerable group depending on if they are vaccinated or not. Like in the first model in Section 3.2, $(V_w + I_w)' = \alpha_w I_w$ implies that I_w tends to zero and V_w tends to a positive value $V_w(\infty)$. We take the limit of the pair of constant quantity (3.3) and obtain

$$L_0 = \log V_w(\infty) + \beta_w\Big(\frac{1}{\alpha_u}V_u(\infty) + \frac{\epsilon}{\alpha_v}V_v(\infty)\Big)$$

We can substitute the relation $V_v = \frac{V_v(0)}{V_u(0)^\delta} V_u^\delta$ into it to obtain a implicit formula for these limit values. It has is again a linear combination of log and the identity, it as then the same form as in the first model and a similar computation.

In the last two sections, we have not been able to investigate the stability as before. The linearization theorem could not be applied ans we did not find asymptotically stable fixed points. However, we showed the asymptotic value of the quantities, and we have seen they tends to a fixed axis, and with a continuous formula for them. This show that the solutions near a position on the fixed axis will tend near to it, and we have Lyanupov stability, restricted to the first quadrant. Without doing an mathematical analysis, we can intuitively see that the scheme $V \to I \to R$ we choose, cannot give a equilibrium, because V is monotone. People only leave the vulnerable group, but nobody enter it. A possible model to investigate, which would allow to have a balance between input and output, and a possible equilibrium, is a demographic model where births and deaths are included in the modelisation as well as immigration and emigration.

Bibliography

[1] L. Y. Adrianova. *Introduction to Linear Systems of Differential Equations.* eng. American Mathematical Society. ISBN: 1-4704-4563-8.

[2] C. Gavin et al. "Dynamics of a Lotka-Volterra type model with applications to marine phage population dynamics". In: *Journal of Physics: Conference Series* 55 (Dec. 2006), pp. 80–93. DOI: 10.1088/1742-6596/55/1/008. URL: https://doi.org/10.1088/1742-6596/55/1/008.

[3] M. W. Hirsch and S. Smalle. *Differential equations, dynamical systems, and linear algebra.* eng. Pure and applied mathematics 60. New York: Academic Press, 1974. ISBN: 0123495504.

[4] J. Hofbauer. *Evolutionary games and population dynamics.* eng. Cambridge [etc: Cambridge University Press, 1998. ISBN: 0521623650.

[5] J. Hofbauer and K. Sigmund. "The Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection". eng. In: *The Quarterly review of biology* 64.4 (1989). ISSN: 0033-5770.

[6] M. C. Irwin. *Smooth dynamical systems.* eng. Advanced series in nonlinear dynamics vol. 17. Singapore: World Scientific, 2001. ISBN: 9810245998.

[7] R. C. Robinson. *An Introduction to Dynamical Systems: Continuous and Discrete, Second Edition.* eng. Providence: American Mathematical Society, 2012. ISBN: 0821891359.

[8] R. C. Robinson. *Dynamical systems : stability, symbolic dynamics, and chaos.* eng. Second ed. Studies in advanced mathematics. Boca Raton [etc: CRC Press, 1999. ISBN: 9780849384950.

[9] Y. Takeuchi. *Global dynamical properties of Lotka-Volterra systems.* eng. Singapore ; World Scientific, 1996. ISBN: 981-283-054-5.