

L2 Mathématiques 2023-2024



2MA236

ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES

Rédacteurs : Ramona Anton, Valeria Banica et Pierre-Antoine Guihéneuf

Table des matières

1	Introduction et motivations	3
1.1	Définitions, historique et premiers exemples	3
1.2	Exemples de bilans entre t et $t + dt$	6
1.3	Difficultés et enjeux	9
2	Équations différentielles linéaires	11
2.1	Équations différentielles linéaires scalaires d'ordre un	11
2.2	Équations différentielles linéaires d'ordre 2	15
2.3	Systèmes d'équations d'ordre un à coefficients constants	19
2.3.1	Intermédiaire : normes d'opérateurs	19
2.3.2	Exponentielle d'une matrice	20
2.3.3	Le cas homogène	24
2.3.4	Le cas non homogène	26
2.3.5	Lien entre les deux méthodes	28
2.4	Tracé des trajectoires de solutions	30
2.4.1	Mécanique du point, portrait de phase	30
2.4.2	Quelques cas simples importants	33
2.4.3	Cas général	42
2.5	Stabilité des équations différentielles	43
3	Théorie générale	49
3.1	Théorème de Cauchy-Lipschitz	49
3.2	Solutions maximales	55
3.2.1	Existence et unicité	55
3.2.2	Dichotomie existence globale/explosion	57
3.3	Quelques conséquences de l'unicité	60
3.4	Critères d'existence globale	62
3.5	Deux exemples de résolution par des séries	65
3.5.1	Par série entière	65
3.5.2	Par série de Fourier	66
3.6	Tracé des trajectoires	67
3.6.1	Un peu de vocabulaire	67
3.6.2	Contraintes sur les solutions scalaires	68
4	Portraits de phase d'équations non-linéaires	73
4.1	Le système de Lotka-Volterra	73
4.2	Le pendule simple	79

Appendices

5	Deux types particuliers d'équations différentielles	87
5.1	Recollements	87
5.2	Équations à variables séparées	91
6	Calculs d'exponentielle de matrice	97
6.1	Quelques exemples	97
6.2	Cas général	99

Chapitre 1

Introduction et motivations

1.1 Définitions, historique et premiers exemples

Définition 1.1.1. Une **équation différentielle ordinaire** est une équation reliant une variable réelle t , une fonction inconnue x dépendant de la variable t et à valeurs dans¹ \mathbb{R}^m , ainsi que ses dérivées jusqu'à un ordre $n \geq 1$, c'est à dire $x'(t), x''(t), \dots, x^{(n)}(t)$. Le plus souvent, ces équations sont écrites sous la forme

$$x^{(n)}(t) = f(t, x(t), x'(t), \dots, x^{(n-1)}(t)), \quad (1.1)$$

où la fonction $f : \mathbb{R} \times \mathbb{R}^m \times \dots \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ est donnée. Autrement dit, la dérivée n -ième de x au point t est fonction de t et des valeurs des dérivées de x d'ordre strictement inférieur en ce même point t . L'entier n est appelé **ordre de l'équation**.

Ces équations peuvent être munies de certaines **conditions initiales** (ce qui correspond à interpréter la variable t comme variable temps), ou de **conditions au bord** (ce qui correspond à interpréter la variable t comme variable espace) : par exemple, les valeurs des dérivées de x jusqu'à l'ordre $n - 1$ sont prescrites en un point particulier t_0 :

$$x(t_0) = x_0, \quad x'(t_0) = x_1, \dots, \quad x^{(n-1)}(t_0) = x_{n-1}.$$

Plusieurs types de questions peuvent se poser :

- On peut chercher, pour x_0, \dots, x_{n-1} donnés, à déterminer l'ensemble des solutions de l'équation (1.1) ;
- On peut essayer de calculer explicitement ces solutions ;
- On peut chercher à comprendre le comportement qualitatif des solutions de (1.1) : ces solutions sont-elles définies sur tout \mathbb{R} ou sur un sous-intervalle ? Sont-elles périodiques ? Explosent-elles ? (c'est-à-dire : tendent-elles vers $\pm\infty$ au bord de leur intervalle de définition ?)
- Si on ne sait pas calculer explicitement les solutions, on peut essayer d'approcher à l'aide de l'ordinateur les solutions (attention, les fonctions ainsi obtenues ne sont pas elles-mêmes des solutions).

On reviendra sur ces différentes notions au cours du chapitre.

1. Dans ce cours on prendra $m = 1$ — EDO scalaires ou $m = 2$ — EDO vectorielles

Un peu d'histoire...

Le calcul différentiel naît à la fin du XVII^e siècle, à la suite des travaux de Leibniz et Newton (1671). Assez naturellement, la question de résoudre des équations différentielles se pose. À cette époque, les modèles étudiés viennent principalement de la mécanique (en particulier de la mécanique céleste) et de la géométrie. Des exemples multiples d'équations différentielles sont fournis par le principe fondamental de la dynamique, ou deuxième loi de Newton, qui s'écrit

$$F = mx''(t),$$

où x est le centre de masse d'un système, m sa masse, et F la résultante des forces s'appliquant sur le système. Notons que cette dernière équation est une équation différentielle pourvu que F puisse s'écrire comme une fonction de la position $x(t)$ et de la vitesse $x'(t)$.

À cette époque, deux méthodes de "résolution" étaient utilisées : soit on intégrait directement l'équation différentielle, c'est-à-dire qu'on disposait d'une formule explicite pour la solution (méthode appelée résolution par quadratures), soit on en calculait un développement en série entière. Un siècle plus tard, en 1769, Euler entreprend la classification de toutes les équations différentielles pour lesquelles on dispose de formules explicites.

Vers 1840 Liouville démontre que certaines équations ne peuvent être résolues analytiquement : la solution existe, mais ne peut être exprimée comme une combinaison de fonctions usuelles. C'est le cas par exemple de l'équation de type Riccati

$$x'(t) = t + x^2(t).$$

Dès lors, il apparaît crucial de disposer d'autres outils que le calcul explicite. Deux voies (complémentaires) sont possibles :

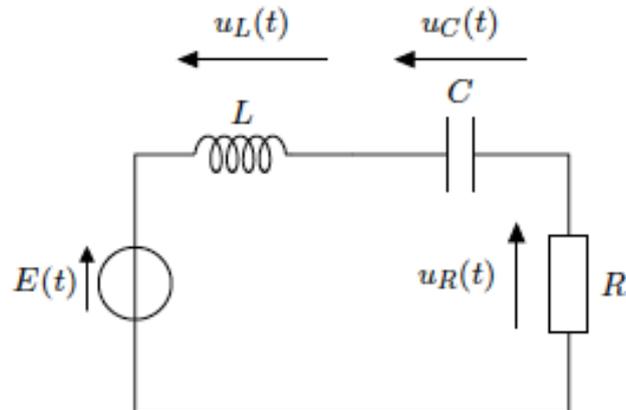
- l'analyse théorique, dont le but est de montrer l'existence et l'unicité des solutions d'équations différentielles à l'aide d'outils généraux ;
- l'analyse numérique, dont le but est de calculer des solutions approchées (fonctions approchant des solutions), avec la meilleure précision possible.

Vers 1870 Fuchs, puis Poincaré, cherchent ainsi à déduire de l'examen a priori de l'équation, les propriétés des solutions.

Parallèlement aux progrès mathématiques dans ce domaine, l'utilisation des équations différentielles se développe dans toutes les disciplines : la chimie (cinétique des réactions chimiques), les sciences du vivant (épidémies, modèles de populations)...

Quelques exemples connus

▷ *Circuit RLC en série :*



Soit $i(t)$ l'intensité qui circule dans le circuit. D'après la loi des mailles, on a

$$E(t) = u_R(t) + u_L(t) + u_C(t).$$

Par ailleurs, on a les relations suivantes pour les tensions électriques u_L, u_C, u_R :

— La charge aux bornes du condensateur est $Q(t) = Cu_C(t)$, et donc

$$i(t) = Q'(t) = Cu'_C(t);$$

— D'après la loi d'Ohm et l'équation ci-dessus, $u_R(t) = Ri(t) = RCu'_C(t)$;

— La tension aux bornes de la bobine est $u_L(t) = Li'(t) = LCu''_C(t)$.

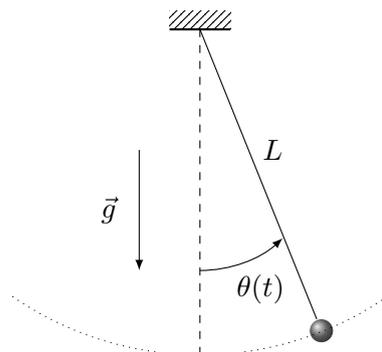
On obtient donc finalement

$$RCu'_C(t) + LCu''_C(t) + u_C(t) = E(t).$$

L'équation ci-dessus est une équation linéaire d'ordre deux, à coefficients constants. On verra au deuxième chapitre comment résoudre ce type d'équation.

Remarque 1.1.2. Dans ce cours, on notera systématiquement la dérivée d'une fonction x comme x' , et jamais comme \dot{x} .

▷ *Pendule simple :*



Le pendule, de masse m , est soumis à deux forces :

— Son poids $m\vec{g}$;

— La tension du fil, orientée selon l'axe du fil.

Ainsi si on projette le principe fondamental de la dynamique suivant la normale au fil, c-à-d la direction du mouvement, on obtient, puisque le mouvement est circulaire de rayon L :

$$mL\theta'' = -mg \sin \theta,$$

soit

$$\theta'' + \frac{g}{L} \sin \theta = 0. \quad (1.2)$$

Remarquons que cette équation est non linéaire : en général, si θ_1, θ_2 sont deux solutions de (1.2), alors $\theta_1 + \theta_2$ n'est pas solution. De façon générale, il est impossible de connaître explicitement les solutions de la plupart des équations non linéaires, mais on verra au deuxième chapitre que l'on peut dire des choses sur leur comportement qualitatif (et on tracera les trajectoires (θ, θ') dans l'Appendice 4.2).

On calcule souvent des solutions exactes d'une équation approchée de² (1.2) pour de petites oscillations en linéarisant le sinus autour de zéro : en effet, $\sin \theta \sim \theta$ pour $\theta \ll 1$. Pour la version linéarisée de l'équation, qui n'est valable que pour les petites oscillations, on obtient l'équation de l'oscillateur harmonique ou pendule élastique :

$$\theta'' + \frac{g}{L} \theta = 0. \quad (1.3)$$

Les solutions sont des combinaisons linéaires de $\sin(\omega_0 t)$ et de $\cos(\omega_0 t)$, où $\omega_0 = \sqrt{g/L}$.

1.2 Exemples de bilans entre t et $t + dt$

▷ *Équation logistique : Modèles de Malthus (1798) et de Verhulst (1836) :*

On considère une population d'individus dont l'effectif au cours du temps est représenté par une fonction à valeurs réelles $N : \mathbb{R}_+ \rightarrow \mathbb{R}_+$.

Dans un premier temps on suppose que le taux instantané de naissances (resp. de décès) dans la population est constant et égal à a (resp. b). On considère la variation de la population entre t et $t + dt$. On a

$$N(t + dt) - N(t) = \text{nombre de naissances} - \text{nombre de décès entre } t \text{ et } t + dt,$$

et donc

$$N(t + dt) - N(t) = aN(t)dt - bN(t)dt.$$

On arrive donc à l'équation différentielle

$$\frac{dN}{dt} = rN(t), \quad (1.4)$$

où $r = a - b$.

Les solutions de cette équation sont de la forme $N_0 e^{rt}$ (on rappellera au deuxième chapitre comment résoudre ce type d'équation). En particulier :

- Si $a > b$ ($r > 0$), autrement dit s'il y a plus de naissances que de décès, la population croît exponentiellement ;
- Si $a < b$ ($r < 0$), la population s'éteint avec une vitesse exponentielle ;
- Si $a = b$ ($r = 0$), on a une situation d'équilibre : la population est constante.

2. Ce qui n'est pas la même chose que de calculer des solutions approchées de l'équation exacte !

Le défaut principal de ce modèle est qu'il est peu réaliste qu'une population croisse indéfiniment à vitesse exponentielle. En effet, il faut tenir compte des ressources dont dispose cette population, qui vont rapidement s'épuiser si la population croît. Pour pallier ce problème, on peut modifier l'équation de Malthus (1.4) en ajoutant un terme quadratique proposé par Verhulst, censé réguler la population en fonction de sa taille :

$$\frac{dN}{dt} = rN(t) \left(1 - \frac{N(t)}{K} \right). \quad (1.5)$$

Ainsi, lorsque $N(t) > K$, $N'(t) < 0$: la population diminue lorsqu'elle dépasse la taille critique K .

Le paramètre K est appelé "charge utile de l'environnement". Il est fonction de la capacité du milieu en nourriture.

Lemme 1.2.1. *L'unique solution de l'équation (1.5) avec pour donnée initiale $N(0) = N_0$ est*

$$N(t) = N_0 \frac{K}{N_0 + (K - N_0)e^{-rt}}.$$

Démonstration. On vérifie que l'on a bien $N(0) = N_0$. Par ailleurs,

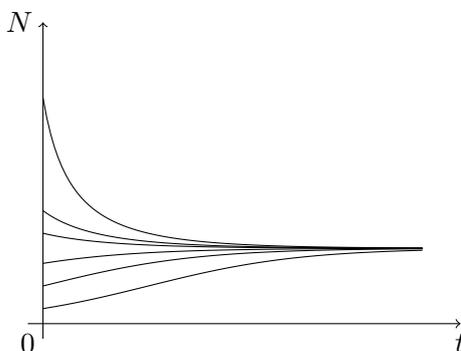
$$\begin{aligned} N'(t) &= -N_0 \frac{K}{(N_0 + (K - N_0)e^{-rt})^2} (K - N_0)(-r)e^{-rt} \\ &= rN(t) \frac{(K - N_0)e^{-rt}}{N_0 + (K - N_0)e^{-rt}} \\ &= rN(t) \left(1 - \frac{N_0}{N_0 + (K - N_0)e^{-rt}} \right) \\ &= rN(t) \left(1 - \frac{N(t)}{K} \right). \end{aligned}$$

Donc N est solution de l'équation (1.5) avec donnée initiale N_0 à $t = 0$.

On verra au Chapitre 3 un théorème d'unicité pour les solutions de l'équation (1.5) avec donnée initiale. On en déduit que $N(t)$ est l'unique solution de l'équation (1.5) avec donnée initiale N_0 à $t = 0$. \square

On vérifie aisément sur la formule que $N(t) \rightarrow K$ quand $t \rightarrow \infty$, quelle que soit la valeur de N_0 .

Regardons l'allure des solutions de (1.5) en fonction de la donnée initiale : en prenant comme paramètres $r = K = 1$, on a le **tracé de solutions** suivant :



Pour aller plus loin dans l'effet de l'environnement dans le cas de prédateurs, on peut inclure comme inconnue la population de prédateurs et considérer le système proie-prédateur de Lotka-Volterra comme suit.

▷ *Modèle de Lotka-Volterra :*

On considère une population de proies $N(t)$, et une population de prédateurs $P(t)$. Les hypothèses de modélisation sont les suivantes :

- Les proies sont supposées avoir une source illimitée de nourriture et se reproduire exponentiellement si elles ne sont soumises à aucune prédation ;
- Le taux de prédation p sur les proies est supposé proportionnel à la fréquence de rencontre entre les prédateurs et les proies. Cette fréquence de rencontre est elle-même proportionnelle au produit du nombre de prédateurs par le nombre de proies ;
- En l'absence de proies, la population de prédateurs s'éteint exponentiellement ;
- Le taux a d'accroissement de la population des prédateurs due à la prédation est proportionnelle au taux de prédation.

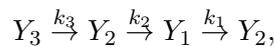
En effectuant un bilan entre t et $t + dt$ on obtient le système d'équations non-linéaire suivant :

$$\begin{cases} \frac{dN}{dt} = r_N N(t) - pN(t)P(t), \\ \frac{dP}{dt} = r_P P(t) + aN(t)P(t). \end{cases}$$

Nous allons tracer les solutions dans l'Appendice 4.1, dans certains cas. À noter qu'un tracé de solution $(N(t), P(t))$ sera une courbe du plan, pas nécessairement de la forme d'un graphe comme lors de solutions des équations à valeurs réelles. Une telle représentation graphique des solutions s'appelle un **portrait de phase**.

▷ *Cinétique chimique :*

Considérons une réaction chimique sur trois espèces :



où k_i sont les coefficients stœchiométriques. Notons c_i la concentration de l'espèce Y_i .

Entre t et $t + dt$, on a donc le bilan suivant :

$$\begin{aligned} c_3(t + dt) - c_3(t) &= -k_3 c_3(t) dt, \\ c_2(t + dt) - c_2(t) &= (k_3 c_3(t) - k_2 c_2(t) + k_1 c_1(t)) dt, \\ c_1(t + dt) - c_1(t) &= (k_2 c_2(t) - k_1 c_1(t)) dt. \end{aligned}$$

On est conduit au système d'équations différentielles

$$\begin{aligned} c_3'(t) &= -k_3 c_3(t), \\ c_2'(t) &= k_3 c_3(t) - k_2 c_2(t) + k_1 c_1(t), \\ c_1'(t) &= k_2 c_2(t) - k_1 c_1(t). \end{aligned}$$

Ce système est un système d'équations linéaires à coefficients constants d'ordre 1. On verra au deuxième chapitre comment obtenir une solution explicite pour un système d'EDO d'ordre 1.

▷ *Modélisation d'une épidémie :*

On pourra consulter les articles suivants :

<https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-1.html>

<https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-2.html>

1.3 Difficultés et enjeux

Comme on l'a déjà expliqué plusieurs fois, il est souvent difficile, voire impossible, d'obtenir des formules explicites pour les solutions d'une équation différentielle. En fait, les seuls cas pour lesquels on sait systématiquement donner une formule pour les solutions sont les équations linéaires à coefficients constants de n'importe quel ordre, ainsi que les équations linéaires d'ordre un. On connaît également des formules pour quelques équations non-linéaires remarquables (comme par exemple l'équation logistique (1.5), ou l'équation de Ricatti), mais ces cas sont l'exception plutôt que la règle.

Dès lors se pose la question de savoir comment calculer (à l'aide d'un ordinateur) les solutions d'équations différentielles. Comme on l'a vu, ces équations interviennent dans de nombreux domaines, allant de la physique et de la chimie aux sciences humaines, et comprendre comment calculer précisément les solutions est donc un enjeu crucial. L'idée clé est que l'on ne calcule quasiment jamais, numériquement, de solution exacte des équations, mais plutôt des solutions approchées (ou des suites de solutions approchées).

Une suite de solutions approchées est une suite $(x_n)_{n \in \mathbb{N}}$ de fonctions définies sur \mathbb{R} (ou sur le même intervalle I de \mathbb{R} que la solution x de l'équation (1.1)), et telle que, pour tout $t \in \mathbb{R}$,

$$x_n(t) \rightarrow x(t) \quad \text{quand } n \rightarrow \infty.$$

Ces fonctions doivent être calculables explicitement et relativement facilement. En revanche, on n'exige pas qu'elles soient des solutions de l'équation (1.1), et en général, elles ne le sont pas.

Un autre enjeu est de démontrer des comportements qualitatifs sans nécessairement savoir calculer la solution x de (1.1). Autrement dit, on cherchera à démontrer, sur la solution x de (1.1) dont on connaît l'existence et l'unicité grâce à un théorème abstrait, des propriétés du type³ :

- x est périodique ;
- x est croissante/décroissante ;
- l'intervalle de définition de x est \mathbb{R} (on dit dans ce cas que la solution est *globale*) ;
- il existe un temps $T^* \in \mathbb{R}$ tel que $x(t) \rightarrow \infty$ quand $t \rightarrow T^*$ (explosion) ;
- dans le cas où la solution est globale, $x(t)$ tend vers une constante quand $t \rightarrow \infty$...

3. Attention, certaines de ces propriétés sont incompatibles !

Chapitre 2

Équations différentielles linéaires et portraits de phases

Un **système d'équations différentielles linéaire** est une équation du type

$$X^{(n)}(t) = A_0(t)X(t) + A_1(t)X'(t) + \cdots + A_{n-1}(t)X^{(n-1)}(t) + B(t),$$

où $X : \mathbb{R} \rightarrow \mathbb{R}^m$ est l'inconnue, et $A_0, \dots, A_{n-1} : \mathbb{R} \rightarrow \mathcal{M}_m(\mathbb{R})$ et $B : \mathbb{R} \rightarrow \mathbb{R}^m$ sont données.

Nous commençons ce chapitre par la résolution du cas $m = 1$ et $n = 1$ (équation scalaire d'ordre 1). On se concentrera ensuite sur les cas $m = 1$ et $n = 2$ (équation scalaire d'ordre 2) et $m = 2$ et $n = 1$ (équations à valeurs vectorielle dans \mathbb{R}^2 d'ordre 1, c'est-à-dire systèmes de deux équations scalaires d'ordre 1), en gardant en tête que les cas m, n quelconques se traitent de façon analogue. On expliquera en particulier pourquoi ces deux problèmes sont équivalents, comment calculer à l'aide de la notion d'exponentielle de matrice les solutions dans le cas où A_0, \dots, A_{n-1} ne dépendent pas de t , et on présentera dans ce cas simple les tracés des solutions.

2.1 Équations différentielles linéaires scalaires d'ordre un

Soit I un intervalle de \mathbb{R} , et soit $a, b : I \rightarrow \mathbb{R}$ deux fonctions continues. On considère l'**équation non-homogène**

$$x'(t) = a(t)x(t) + b(t), \tag{2.1}$$

ainsi que sa version linéaire **homogène**

$$x'(t) = a(t)x(t). \tag{2.2}$$

Le terme "homogène" reflète ici le fait que la fonction nulle soit solution : on parle également d'**équation sans second membre**. En effet on écrit souvent (2.1) sous la forme

$$x'(t) - a(t)x(t) = b(t).$$

Résolution de l'équation homogène (2.2)

Le cas des équations différentielles linéaires d'ordre un homogènes est assez spécial : on connaît une forme générale pour les solutions. Elle est donnée par le théorème suivant.

Théorème 2.1.1. *Soit $t_0 \in I$, et soit $x_0 \in \mathbb{R}$ quelconque.*

L'équation homogène (2.2) munie de la condition initiale $x(t_0) = x_0$ admet une unique solution, donnée par

$$x(t) = x_0 e^{\int_{t_0}^t a(u) du}.$$

En particulier toute solution de (2.2) est de la forme

$$x(t) = C e^{\int_{t_0}^t a(u) du} = \tilde{C} e^{\int_{t_1}^t a(u) du}, C, \tilde{C}, t_1 \in \mathbb{R}.$$

Remarque 2.1.2. L'unicité n'est pas valable pour toutes des équations différentielles : par exemple $x'(t) = 2\sqrt{x(t)}$ avec $x(0) = 0$ admet deux solutions, celle nulle $x_1(t) = 0, \forall t \in \mathbb{R}$, et aussi $x_2(t) = 0$ si $t < 0$ et $x_2(t) = t^2$ pour $t \geq 0$.

Démonstration.

Existence : Posons

$$f(t) = x_0 e^{\int_{t_0}^t a(u) du}.$$

On vérifie immédiatement que $f(t_0) = x_0 e^0 = x_0$.

Montrons maintenant que f est bien solution de l'équation (2.2). On a, par le théorème fondamental de l'analyse,

$$f'(t) = x_0 a(t) e^{\int_{t_0}^t a(u) du} = a(t) f(t).$$

Donc f est solution de (2.2).

Unicité : Soit x une solution quelconque de (2.2). On remarque que x est nécessairement de classe \mathcal{C}^1 . On pose

$$C(t) = x(t) e^{-\int_{t_0}^t a(u) du}$$

Alors la fonction C est de classe \mathcal{C}^1 et

$$C'(t) = x'(t) e^{-\int_{t_0}^t a(u) du} - x(t) a(t) e^{-\int_{t_0}^t a(u) du} = (x'(t) - x(t) a(t)) e^{-\int_{t_0}^t a(u) du} = 0.$$

Donc la fonction C est constante sur l'intervalle I et vaut $C(t_0) = x(t_0) e^0 = x_0$. Donc

$$x(t) = C(t) e^{\int_{t_0}^t a(u) du} = x_0 e^{\int_{t_0}^t a(u) du}.$$

□

Exemple 2.1.3. — On prend $I = \mathbb{R}$, $a(t) = r \in \mathbb{R}$ pour tout $t \in \mathbb{R}$. Alors les solutions de l'équation (2.2) sont données par

$$x(t) = C e^{r(t-t_0)}.$$

— On prend $I = \mathbb{R}_+^*$, $t_0 \in \mathbb{R}_+^*$ et $a(t) = k/t$ pour $k \in \mathbb{R}^*$. Alors les solutions de l'équation (2.2) sont données par

$$x(t) = C e^{k \ln \frac{t}{t_0}} = C \left(\frac{t}{t_0} \right)^k.$$

On déduit en particulier du Théorème 2.1.1 que si $x_0 = 0$, alors l'unique solution de (2.2) est identiquement nulle. On obtient alors le corollaire suivant à propos de l'équation non-homogène (2.2) :

Corollaire 2.1.4. Soit $t_0 \in I$, et soit $x_0 \in \mathbb{R}$. Alors l'équation (2.1) munie de la condition initiale $x(t_0) = x_0$ admet au plus une solution.

Démonstration. Soient x et y deux solutions de (2.1) telles que $x(t_0) = y(t_0) = x_0$. Alors on vérifie immédiatement que $z := x - y$ est solution de (2.2) avec $z(t_0) = 0$. Par conséquent $z = 0$. \square

On cherche à présent à montrer l'existence des solutions de l'équation non-homogène (2.1), ainsi qu'une méthode pour calculer les solutions.

Résolution de l'équation non-homogène (2.1) par la méthode de variation de la constante

L'idée pour résoudre l'équation non-homogène (2.1) est d'utiliser le calcul fait dans le cas homogène pour changer de fonction inconnue et enlever ainsi le terme d'ordre zéro (c'est-à-dire $a(t)x(t)$) dans l'équation. Plus précisément, supposons que x est une solution de (2.1) et posons

$$C(t) = x(t)e^{-\int_{t_0}^t a(u)du},$$

ce qui revient à dire

$$x(t) = C(t)e^{\int_{t_0}^t a(u)du}.$$

On a alors

$$x'(t) = (C'(t) + a(t)C(t))e^{\int_{t_0}^t a(u)du},$$

et comme d'autre part

$$x'(t) = a(t)x(t) + b(t) = a(t)C(t)e^{\int_{t_0}^t a(u)du} + b(t),$$

on en déduit que

$$C'(t) = e^{-\int_{t_0}^t a(u)du}b(t).$$

Si on impose une condition initiale $x(t_0) = x_0$, on a alors $C(t_0) = x_0$ et en intégrant on obtient

$$C(t) = x_0 + \int_{t_0}^t b(s)e^{-\int_{t_0}^s a(u)du} ds.$$

On est donc conduits au théorème suivant :

Théorème 2.1.5. Soit $t_0 \in I$, $x_0 \in \mathbb{R}$. Alors l'unique solution de (2.1) telle que $x(t_0) = x_0$ est donnée par

$$\begin{aligned} x(t) &= \left(x_0 + \int_{t_0}^t e^{-\int_{t_0}^s a(u)du} b(s) ds \right) e^{\int_{t_0}^t a(u)du} \\ &= x_0 e^{\int_{t_0}^t a(u)du} + \int_{t_0}^t e^{\int_s^t a(u)du} b(s) ds. \end{aligned}$$

Remarque 2.1.6. On remarque que la solution de (2.1) est somme d'une solution de l'équation homogène (2.2) avec donnée initiale x_0 à t_0 et d'une solution de l'équation non-homogène (2.1) avec donnée initiale nulle à t_0 . Cette dernière est une "superposition" de solutions de l'équation homogène. Ce principe, qui consiste à exprimer les solutions d'une équation non homogène en fonction des solutions de l'équation homogène associée remonte à Euler et Lagrange, et dans le cas des équations aux dérivées partielles porte le nom de principe de Duhamel.

En termes plus savants, cela revient à dire que l'ensemble des solutions d'une équation linéaire scalaire non homogène d'ordre 1 est un espace affine de dimension 1, dont l'espace vectoriel associé est celui des solutions de l'équation linéaire scalaire homogène d'ordre 1 associée (ensemble qui est lui-même un espace vectoriel de dimension 1).

Démonstration. En reprenant le calcul ci-dessus, on vérifie facilement que x est solution de (2.1). L'unicité est assurée grâce au Corollaire 2.1.4. \square

On vient donc de montrer l'existence des solutions de (2.1), et de proposer une méthode pour calculer ces solutions. Cette méthode marche aussi en prenant une autre primitive et donnée initiale dans la définition de $C(t)$,

$$C(t) = x(t)e^{-\int_{t_1}^t a(u)du}, C(t_1) = C,$$

le choix de t_1 pouvant simplifier les calculs intermédiaires de la fonction $C(t)$; ensuite ce n'est qu'à la fin que l'on choisit C de façon à ce que la condition initiale en t_0 pour $x(t)$ soit vérifiée.

Exemple 2.1.7. On prend $I = \mathbb{R}_+^*$, et on considère l'équation

$$x'(t) - \frac{x(t)}{t} = t^2.$$

D'après la méthode de variation de la constante, on pose

$$x(t) = C(t)e^{\int_1^t \frac{1}{u}du} = tC(t),$$

et donc C est solution de

$$C'(t) = \frac{t^2}{t} = t,$$

d'où $C(t) = \frac{t^2}{2} + C$. On obtient finalement

$$x(t) = \frac{t^3}{2} + Ct.$$

Il faut ensuite choisir la constante C de façon à ce que la condition initiale en $t_0 > 0$ soit vérifiée.

Résolution de l'équation non-homogène (2.1) à l'aide d'une solution particulière

Dans les cas où le calcul explicite précédent n'est pas concluant, on dispose d'une autre méthode pour calculer les solutions de l'équation (2.1), qui est résumée dans le théorème suivant :

Théorème 2.1.8. Soit $t_0 \in I$, $x_0, x_1 \in \mathbb{R}$. Soit $x_p : I \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 telle que x_p est une solution de l'équation (2.1) et $x_p(t_0) = x_1$.

Alors l'unique solution de (2.1) telle que $x(t_0) = x_0$ est donnée par

$$x(t) = x_p(t) + (x_0 - x_1)e^{\int_{t_0}^t a(u)du}.$$

En particulier toute solution de (2.1) est de la forme

$$x(t) = x_p(t) + Ce^{\int_{t_0}^t a(u)du}, C \in \mathbb{R}.$$

Démonstration. L'unicité est assurée grâce au Corollaire 2.1.4. Il s'agit donc de montrer que la fonction $f(t) = x_p(t) + (x_0 - x_1)e^{\int_{t_0}^t a(u)du}$ est bien solution de l'équation. On a

$$f(t_0) = x_p(t_0) + (x_0 - x_1) = x_1 + x_0 - x_1 = x_0.$$

Par ailleurs,

$$\begin{aligned} f'(t) &= x_p'(t) + (x_0 - x_1)a(t)e^{\int_{t_0}^t a(u)du} \\ &= a(t)x_p(t) + b(t) + (x_0 - x_1)a(t)e^{\int_{t_0}^t a(u)du} \\ &= a(t)f(t) + b(t). \end{aligned}$$

Donc f est solution de (2.1) avec donnée initiale $f(t_0) = x_0$. \square

Exemple 2.1.9. On revient sur l'exemple précédent : $I = \mathbb{R}_+^*$, et on considère l'équation

$$x'(t) - \frac{x(t)}{t} = t^2.$$

On remarque que $x_p(t) = \frac{t^3}{2}$ est une solution particulière. Sachant que les solutions de l'équation homogène sont de la forme Ct , nous obtenons la forme générale des solutions :

$$x(t) = \frac{t^3}{2} + Ct.$$

Il faut ensuite choisir la constante C de façon à ce que la condition initiale en $t_0 > 0$ soit vérifiée. On remarque que l'on peut aussi résoudre l'équation sur $I = \mathbb{R}_-^*$. Pour les notions de recollements de solutions voir l'Appendice 5.1.

Exemple 2.1.10. On prend $I = \mathbb{R}$, et on considère l'équation

$$x'(t) + tx(t) = t^2 + 1.$$

On vérifie qu'une solution particulière de cette équation est $x_p(t) = t$. La solution générale de l'équation est donc donnée par

$$x(t) = t + Ce^{-t^2/2},$$

où la constante C est choisie de telle sorte que la condition initiale soit vérifiée.

On peut aussi résoudre cette équation avec la méthode de variation de la constante, comme suit. On pose $x(t) = C(t)e^{-t^2/2}$ et on obtient $C'(t) = e^{t^2/2}(t^2 + 1)$ donc

$$C(t) = C + \int_0^t e^{s^2/2}(s^2 + 1)ds = C + [se^{s^2/2}]_0^t = C + te^{t^2/2}.$$

2.2 Équations différentielles linéaires d'ordre 2

On considère dans ce paragraphe des équations de la forme

$$x''(t) + a(t)x'(t) + b(t)x(t) = h(t), \quad (2.3)$$

où $a, b, h : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions de classe \mathcal{C}^1 .

Passage à un système d'équations d'ordre un

On considère une solution $x : \mathbb{R} \rightarrow \mathbb{R}$ de l'équation (2.3). On pose, pour $t \in \mathbb{R}$,

$$X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} \in \mathbb{R}^2.$$

On vérifie alors que

$$X'(t) = \begin{pmatrix} x'(t) \\ x''(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} + \begin{pmatrix} 0 \\ h(t) \end{pmatrix}.$$

Posons

$$A(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix}, \quad F(t) = \begin{pmatrix} 0 \\ h(t) \end{pmatrix}.$$

La matrice A et le vecteur F sont des données du problème. La fonction $X : \mathbb{R} \rightarrow \mathbb{R}^2$ est solution de

$$X'(t) = A(t)X(t) + F(t). \quad (2.4)$$

On s'est donc ramené d'une équation différentielle scalaire (c'est-à-dire dont l'inconnue est à valeurs réelles) d'ordre 2 à un système de deux équations différentielles d'ordre un.

Réciproquement, soit X une solution de (2.4). On note (x_1, x_2) les coordonnées de X . On a alors

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 0 \\ h(t) \end{pmatrix}.$$

En particulier, $x_1'(t) = x_2(t)$ et

$$x_2'(t) = -b(t)x_1(t) - a(t)x_2(t) + h(t).$$

Donc x_1 est solution de (2.3). On a donc démontré le résultat suivant :

Proposition 2.2.1. *Soit $x : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 . Soit $X : t \in \mathbb{R} \mapsto \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} \in \mathbb{R}^2$. Alors*

$$x \text{ est solution de (2.3)} \Leftrightarrow X \text{ est solution de (2.4)}.$$

Remarque 2.2.2. De façon générale, et avec la même méthode, on peut montrer qu'une équation différentielle scalaire linéaire d'ordre n se ramène à un système linéaire de n équations différentielles d'ordre un, et qu'un système linéaire de m équations d'ordre n se ramène à un système linéaire de $m \times n$ équations différentielles d'ordre un.

Cette méthode permet, dans le cas linéaire, de restreindre l'étude au cas des systèmes d'ordre un. Aussi, dans la suite de cette partie, on étudiera uniquement des systèmes de deux équations linéaires d'ordre un (cas $m = 2, n = 1$).

Résolution dans le cas des coefficients constants

On considère tout d'abord des équations de la forme

$$x''(t) + ax'(t) + bx(t) = 0, \quad (2.5)$$

où $a, b \in \mathbb{R}$. L'enjeu est donc ici de trouver une expression explicite pour la solution de (2.5) munie de la condition (2.6) : en effet, si on arrive à trouver une solution particulière de (2.5) vérifiant les bonnes conditions initiales en $t = t_0$, le théorème ci-dessus assure que cette solution particulière est en réalité l'unique solution.

L'idée est de chercher des solutions comme des combinaisons linéaires de $\exp(\lambda t)$.

Théorème 2.2.3. Soit $a, b, t_0, x_0, x_1 \in \mathbb{R}$. On note x la solution de l'équation (2.5) munie des conditions initiales

$$x(t_0) = x_0, \quad x'(t_0) = x_1. \quad (2.6)$$

Soit $\Delta := a^2 - 4b$. On considère l'équation suivante, appelée **équation caractéristique** de (2.5) :

$$\lambda^2 + a\lambda + b = 0. \quad (2.7)$$

— *Premier cas : $\Delta \neq 0$: on note λ_1, λ_2 les deux racines distinctes de l'équation (2.7). Si les racines sont réelles alors il existe $A_1, A_2 \in \mathbb{R}$ tels que*

$$x(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}.$$

Si les racines sont complexes, alors il existe $A \in \mathbb{C}$ tel que

$$x(t) = A e^{\lambda_1 t} + \bar{A} e^{\lambda_2 t};$$

si on note $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta$, alors il existe $C_1, C_2 \in \mathbb{R}$ tels que

$$x(t) = e^{t\alpha} (C_1 \cos(\beta t) + C_2 \sin(\beta t)).$$

— *Deuxième cas : $\Delta = 0$: soit λ_0 l'unique racine réelle de (2.7) (il s'agit d'une racine double). Alors il existe $A, B \in \mathbb{R}$ tels que*

$$x(t) = (At + B)e^{\lambda_0 t}.$$

La preuve de ce théorème nécessite un résultat d'unicité qui sera vu par la suite (Théorème 2.3.8). Nous l'admettrons pour le moment, remettant sa preuve à un peu plus tard.

Démonstration. Première étape : les fonctions de la forme $A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}$ (si $\Delta \neq 0$) ou $(At + B)e^{\lambda_0 t}$ (si $\Delta = 0$) sont solutions de (2.5).

On se donne $A_1, A_2, A, B \in \mathbb{R}$ et pose

$$h(t) = \begin{cases} A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} & \text{si } \Delta \neq 0, \\ (At + B)e^{\lambda_0 t} & \text{si } \Delta = 0. \end{cases}$$

La fonction h ainsi définie est de classe C^∞ . Dans le premier cas ($\Delta \neq 0$), on vérifie que

$$\begin{aligned} & h''(t) + ah'(t) + bh(t) \\ &= \left(A_1 \lambda_1^2 e^{\lambda_1 t} + A_2 \lambda_2^2 e^{\lambda_2 t} \right) + a \left(A_1 \lambda_1 e^{\lambda_1 t} + A_2 \lambda_2 e^{\lambda_2 t} \right) + b \left(A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} \right) \\ &= A_1 e^{\lambda_1 t} (\lambda_1^2 + a\lambda_1 + b) + A_2 e^{\lambda_2 t} (\lambda_2^2 + a\lambda_2 + b) \\ &= 0 \end{aligned}$$

puisque λ_1, λ_2 sont solutions de l'équation caractéristique.

Dans le second cas ($\Delta = 0$),

$$\begin{aligned} h'(t) &= (\lambda_0 At + \lambda_0 B + A)e^{\lambda_0 t}, \\ h''(t) &= (\lambda_0^2 At + \lambda_0^2 B + 2\lambda_0 A)e^{\lambda_0 t}, \end{aligned}$$

et donc en utilisant le fait que $\lambda_0 = -a/2$,

$$\begin{aligned} h''(t) + ah'(t) + bh(t) &= e^{\lambda_0 t} [\lambda_0^2 At + \lambda_0^2 B + 2\lambda_0 A + a(\lambda_0 At + \lambda_0 B + A) + b(At + B)] \\ &= e^{\lambda_0 t} [At(\lambda_0^2 + a\lambda_0 + b) + B(\lambda_0^2 + a\lambda_0 + b) + A(2\lambda_0 + a)] \\ &= 0. \end{aligned}$$

Deuxième étape : on peut choisir les coefficients A_1, A_2 (cas $\Delta \neq 0$) ou A, B (cas $\Delta = 0$) pour que les conditions initiales soient vérifiées.

Il s'agit de montrer que les systèmes linéaires, d'inconnues A_1 et A_2

$$\begin{aligned} A_1 e^{\lambda_1 t_0} + A_2 e^{\lambda_2 t_0} &= x_0, \\ A_1 \lambda_1 e^{\lambda_1 t_0} + A_2 \lambda_2 e^{\lambda_2 t_0} &= x_1 \end{aligned} \tag{2.8}$$

dans le cas $\Delta \neq 0$, ou d'inconnues A et B

$$\begin{aligned} (At_0 + B)e^{\lambda_0 t_0} &= x_0, \\ (\lambda_0 At_0 + \lambda_0 B + A)e^{\lambda_0 t_0} &= x_1 \end{aligned} \tag{2.9}$$

dans le cas $\Delta \neq 0$ admettent une unique solution dans \mathbb{C}^2 .

Le déterminant du système (2.8) est $(\lambda_2 - \lambda_1)e^{(\lambda_1 + \lambda_2)t_0}$, qui est donc non nul puisque $\lambda_1 \neq \lambda_2$. Celui du système (2.9) (qui a pour inconnues A et B) est $-e^{2\lambda_0 t_0}$, qui est également non nul. Chacun des deux systèmes ci-dessus admet donc une unique solution dans \mathbb{C}^2 .

À ce stade, on a montré que l'on peut choisir des coefficients A_1, A_2, A, B de telle sorte que h soit une solution de (2.5) munie de la condition initiale (2.6). Il reste à montrer que la fonction h est à valeurs réelles.

Troisième étape : les coefficients A_1, A_2, A, B définis ci-dessus sont tels que la solution trouvée est à valeurs réelles.

Tout d'abord, si $\Delta > 0$ (resp. si $\Delta = 0$), les racines λ_1, λ_2 sont réelles (resp. $\lambda_0 \in \mathbb{R}$), donc la solution de (2.8) (resp. de (2.9)) est dans \mathbb{R}^2 . Dans ces cas il est donc clair que $h(t) \in \mathbb{R}$ pour tout t .

Il reste à examiner le cas $\Delta < 0$. Dans ce cas les deux racines λ_1, λ_2 appartiennent à $\mathbb{C} \setminus \mathbb{R}$ et sont des nombres complexes conjugués. En prenant le conjugué du système (2.8) et en utilisant le fait que

$$\bar{\lambda}_1 = \lambda_2 \quad \text{et} \quad \overline{e^{\lambda_1 t_0}} = e^{\lambda_2 t_0},$$

on en déduit que (\bar{A}_2, \bar{A}_1) est solution du même système que (A_1, A_2) . Par unicité des solutions du système (2.8), on obtient $A_2 = \bar{A}_1$, et donc

$$h(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} = 2\Re(A_1 e^{\lambda_1 t}).$$

En particulier on peut réécrire $h(t)$ en utilisant la notation $\lambda_1 = \alpha + i\beta$:

$$h(t) = e^{t\alpha}(C_1 \cos(\beta t) + C_2 \sin(\beta t)),$$

avec $C_1, C_2 \in \mathbb{R}$.

Quatrième étape : conclusion.

Ainsi $h : \mathbb{R} \rightarrow \mathbb{R}$ est une solution réelle de (2.5) qui vérifie (2.6). D'après le Théorème 2.3.8 qui assure l'unicité, on a $x = h$. \square

2.3 Systèmes d'équations d'ordre un à coefficients constants

Les paragraphes qui viennent sont dédiés à l'étude de systèmes d'équations différentielles du type

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

avec $X : \mathbb{R} \rightarrow \mathbb{R}^2$, $A \in \mathcal{M}_2(\mathbb{R})$, $B : \mathbb{R} \rightarrow \mathbb{R}^2$.

On parle de système à coefficients constants pour souligner le fait que la matrice A ne dépend pas de t . Si $B = 0$, le système est dit *homogène*, et son étude fait l'objet du paragraphe suivant. Si $B \neq 0$, le système est *non-homogène*, et son étude fera l'objet du paragraphe 2.3.4.

On rappelle que si $a \in \mathbb{R}$, les solutions de l'équation

$$y'(t) = ay(t)$$

sont de la forme $y(t) = y_0 e^{at}$. L'idée pour résoudre des systèmes d'équations est d'utiliser la généralisation aux matrices de $\mathcal{M}_2(\mathbb{R})$ de la définition de l'exponentielle d'un nombre réel $x \in \mathbb{R}$ comme série absolument convergente :

$$e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

2.3.1 Intermède : normes d'opérateurs

Dans ce paragraphe, on se place dans \mathbb{C}^n , $n \geq 1$. Des énoncés identiques restent vrais lorsqu'on remplace \mathbb{C}^n par \mathbb{R}^n .

Étant donnée une norme $\|\cdot\|$ sur \mathbb{C}^n , on définit une norme triple sur les matrices de $M_n(\mathbb{C})$:

$$\| \| A \| \| = \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\| AX \|}{\| X \|}.$$

Proposition 2.3.1. *L'application $\| \| \cdot \| \|$ définit une norme sur $M_n(\mathbb{C})$, qui est sous-multiplicative : pour $A, B \in M_n(\mathbb{C})$, on a $\| \| AB \| \| \leq \| \| A \| \| \| \| B \| \|$.*

Démonstration. Vérifions les trois propriétés de la norme.

- Par définition, $\| \| A \| \| = 0$ si et seulement si pour tout $X \in \mathbb{C}^n \setminus \{0\}$, on a $\| AX \| / \| X \| = 0$, ce qui est équivalent au fait que pour tout $X \in \mathbb{C}^n \setminus \{0\}$, on a $\| AX \| = 0$. Il revient au même de dire que $A = 0$. Ainsi, $\| \| A \| \| = 0 \iff A = 0$.
- Soit $t \geq 0$ et $A \in M_n(\mathbb{C})$. On a, par homogénéité de la norme $\|\cdot\|$,

$$\| \| tA \| \| = \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\| tAX \|}{\| X \|} = \sup_{X \in \mathbb{C}^n \setminus \{0\}} |t| \frac{\| AX \|}{\| X \|} = |t| \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\| AX \|}{\| X \|} = |t| \| \| A \| \|.$$

— Soient $A, B \in M_n(\mathbb{C})$. On a,

$$\begin{aligned} \|A + B\| &= \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|(A + B)X\|}{\|X\|} \\ &= \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|AX + BX\|}{\|X\|} \\ &\leq \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|AX\| + \|BX\|}{\|X\|} \quad (\text{par inégalité triangulaire sur } \|\cdot\|) \\ &\leq \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|AX\|}{\|X\|} + \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|BX\|}{\|X\|} \\ &\leq \|A\| + \|B\|. \end{aligned}$$

Il en reste plus qu'à démontrer la sous-multiplicativité. Soient $A, B \in M_n(\mathbb{C})$. Remarquons que si $B = 0$, alors $\|AB\| = 0 = \|A\| \|B\|$. On peut donc supposer $B \neq 0$. Ensuite, si $BX = 0$, alors $ABX = 0$. Ainsi,

$$\|AB\| = \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|ABX\|}{\|X\|} = \sup_{X \in \mathbb{C}^n \setminus \{0\}, BX \neq 0} \frac{\|ABX\|}{\|X\|}.$$

Par conséquent, (en utilisant le fait que le sup du produits de termes positifs est inférieur au produit des sups)

$$\|AB\| = \sup_{X \in \mathbb{C}^n \setminus \{0\}, BX \neq 0} \frac{\|ABX\|}{\|BX\|} \frac{\|BX\|}{\|X\|} \leq \sup_{Y \in \mathbb{C}^n \setminus \{0\}} \frac{\|AY\|}{\|Y\|} \sup_{X \in \mathbb{C}^n \setminus \{0\}} \frac{\|BX\|}{\|X\|} = \|A\| \|B\|.$$

□

2.3.2 Exponentielle d'une matrice

Théorème 2.3.2 (Exponentielle de matrice). *Soit $A \in \mathcal{M}_n(\mathbb{C})$. Pour $n \in \mathbb{N}$, on définit*

$$M_n = \sum_{k=0}^n \frac{A^k}{k!} \in \mathcal{M}_2(\mathbb{C}).$$

Alors la suite $(M_n)_{n \in \mathbb{N}}$ est convergente. Sa limite est appelée l'exponentielle de la matrice A , et notée $\exp(A)$ ou e^A .

Ce théorème utilise le fait que l'espace vectoriel normé $\mathcal{M}_n(\mathbb{C})$ (muni de la norme triple $\|\cdot\|$) est complet.

Démonstration. On montre que la suite (M_n) est de Cauchy, plus précisément que la série définissant (M_n) est absolument convergente. En effet, par sous-multiplicativité,

$$\sum_{k=0}^n \left\| \frac{A^k}{k!} \right\| \leq \sum_{k=0}^n \frac{\|A^k\|}{k!} \leq \sum_{k=0}^n \frac{\|A\|^k}{k!} = \exp(\|A\|).$$

Ainsi, la série définissant (M_n) est absolument convergente et donc convergente puisque $\mathcal{M}_n(\mathbb{C})$ est complet. □

Proposition 2.3.3 (Propriétés de l'exponentielle de matrice). *Soit $A \in \mathcal{M}_n(\mathbb{C})$, $P \in \mathcal{G}_n(\mathbb{C})$, et soit $\lambda, \mu \in \mathbb{C}$. On note $0_n \in \mathcal{M}_n(\mathbb{C})$ la matrice nulle et I_n la matrice identité. Nous avons les propriétés suivantes :*

- a) $e^{0n} = I_n$,
 b) $Ae^A = e^AA$,
 c) $e^{tA} = {}^t(e^A)$,
 d) $e^{\lambda A}e^{\mu A} = e^{(\lambda+\mu)A} = e^{\mu A}e^{\lambda A}$,
 e) $e^{PAP^{-1}} = Pe^AP^{-1}$.

Démonstration. Nous laissons cette preuve en exercice. □

Remarque 2.3.4. Attention, en général on n'a pas

$$e^Ae^B = e^{A+B}.$$

Cette propriété n'est vraie que si les matrices A et B commutent, c'est-à-dire si $AB = BA$. Voici un contre-exemple :

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

On vérifie alors aisément que $A^2 = B^2 = 0$. Par suite $A^k = B^k = 0$ pour tout $k \geq 2$. On obtient donc

$$e^A = I_2 + A, \quad e^B = I_2 + B.$$

D'autre part

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

On en déduit

$$e^Ae^B = I_2 + A + B + AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

$$e^Be^A = I_2 + A + B + BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

On a donc $e^Ae^B \neq e^Be^A$, ce qui implique que l'égalité $e^Ae^B = e^{A+B}$ ne peut être vraie.

Calculs d'exponentielle de matrice (pour plus de détails sur ces rappels consulter §6)

À partir d'ici nous traiterons uniquement le cas de la dimension 2. L'ensemble des résultats qui suivent peuvent se généraliser à des équations différentielles à coefficients constants d'ordre arbitrairement grand.

Pour une matrice quelconque de $\mathcal{M}_2(\mathbb{C})$, nous avons deux possibilités :

Type	Formule	Exponentielle
Matrice diagonalisable	$P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}$	$P \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} P^{-1}$
Matrice trigonalisable (mais pas diagonalisable)	$P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1}$	$P \begin{pmatrix} e^\lambda & \gamma e^\lambda \\ 0 & e^\lambda \end{pmatrix} P^{-1}$

Dans ce cours nous allons devoir souvent calculer des exponentielles de matrices réelles $A \in \mathcal{M}_2(\mathbb{R})$. Pour les exponentielles de matrices réelles on a les formules suivantes :

Type	Formule	Exponentielle
Matrice diagonalisable avec valeurs propres $\lambda_1, \lambda_2 \in \mathbb{R}$ distinctes ou confondues	$P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}, P \in \text{GL}_n(\mathbb{R})$	$P \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} P^{-1}$
Matrice diagonalisable avec valeurs propres non réelles $\alpha \pm i\beta, \beta \neq 0$	$P \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} P^{-1}, P \in \text{GL}_n(\mathbb{R})$	$P e^\alpha \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} P^{-1}$
Matrice non-diagonalisable avec une valeur propre double $\lambda \in \mathbb{R}$	$P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1}, P \in \text{GL}_n(\mathbb{R})$ $\gamma \in \mathbb{R}^*$	$P e^\lambda \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} P^{-1}$

Dans la pratique, si on veut calculer l'exponentielle d'une matrice réelle $A \in \mathcal{M}_2(\mathbb{R})$, d'abord on trouve ses valeurs propres, les deux racines λ_1 et λ_2 de l'équation $\det(A - \lambda I_2) = 0$. On note les sous-espaces de vecteurs propres correspondants $E_j = \ker(A - \lambda_j I_2)$, $j \in \{1, 2\}$. On distingue plusieurs cas, comme suit.

- Si $\lambda_1, \lambda_2 \in \mathbb{R}$ et sont distinctes (matrice diagonalisable) : On choisit $v_1 \in E_1, v_2 \in E_2$ deux vecteurs propres non-nuls réels, on pose $P \in \mathcal{M}_2(\mathbb{R})$ la matrice ayant v_1 en première colonne et v_2 en deuxième colonne, et on a la diagonalisation $A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}$ et par conséquent

$$e^A = P \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} P^{-1}.$$

Le même argument s'applique lorsque $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$ et $\dim E = 2$ (matrice diagonalisable).

- Si $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$ et $\dim E = 1$ (matrice non-diagonalisable) : On choisit $v_1 \in E$ un vecteur propre non nul réel, $v_2 \in \mathbb{R}^2 \setminus E$ et on détermine $\gamma \in \mathbb{R}^*$ tel que $Av_2 = \gamma v_1 + \lambda v_2$. On pose $P \in \mathcal{M}_2(\mathbb{R})$ la matrice ayant v_1 en première colonne et v_2 en deuxième colonne, et on a la trigonalisation $A = P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1}$ et par conséquent

$$e^A = P e^\lambda \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} P^{-1}.$$

- Si $\lambda_1 = \alpha + i\beta$ et $\lambda_2 = \alpha - i\beta$ avec $\beta \neq 0$ (matrice diagonalisable)

Lemme 2.3.5. Soit V un vecteur propre de A associé à la valeur propre $\lambda = \alpha + i\beta$, et $R = (\Re(V) \quad \Im(V))$. Alors

$$A = R \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} R^{-1}.$$

Démonstration. Remarquons tout d'abord que la matrice R est inversible. En effet, puisque V est un vecteur propre de A , associé à une valeur propre λ , on a $AV = \lambda V$, ce qui implique (en prenant le conjugué) que $A\bar{V} = \bar{\lambda}\bar{V}$. Cela implique que \bar{V} est un vecteur propre de A associé à la valeur propre $\bar{\lambda}$, et donc que la famille V, \bar{V} forme une base de \mathbb{C}^2 . Or

$$\frac{1}{2} \begin{pmatrix} V & \bar{V} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} V - \bar{V} & V + \bar{V} \end{pmatrix} = \begin{pmatrix} \Re(V) & \Im(V) \end{pmatrix},$$

donc (puisque la matrice $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ est inversible) la matrice $\begin{pmatrix} \Re(V) & \Im(V) \end{pmatrix}$ est inversible.

Montrons maintenant que

$$AR = R \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix},$$

en testant cette égalité sur une base de \mathbb{C}^2 . Tout d'abord, d'une part

$$AR \begin{pmatrix} 1 \\ i \end{pmatrix} = AV = \lambda V.$$

D'autre part

$$R \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = R \begin{pmatrix} \alpha + i\beta \\ -\beta + i\alpha \end{pmatrix} = \lambda R \begin{pmatrix} 1 \\ i \end{pmatrix} = \lambda V.$$

Ensuite, d'une part

$$AR \begin{pmatrix} 1 \\ -i \end{pmatrix} = A\bar{V} = \bar{\lambda}\bar{V}.$$

D'autre part

$$R \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = R \begin{pmatrix} \alpha - i\beta \\ -\beta - i\alpha \end{pmatrix} = \bar{\lambda} R \begin{pmatrix} 1 \\ -i \end{pmatrix} = \bar{\lambda}\bar{V}.$$

Ainsi, on a testé l'égalité du lemme sur la base

$$\left(\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix} \right)$$

de \mathbb{C}^2 , cela démontre le lemme. □

Lemme 2.3.6. *Pour $\alpha, \beta \in \mathbb{R}$, on a*

$$\exp \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = e^\alpha \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix}.$$

Démonstration. La propriété est triviale pour $\beta = 0$, supposons donc le contraire. Diagonalisons dans \mathbb{C} la matrice

$$B = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Le polynôme caractéristique de B est $\chi_B = X^2 - 2X\alpha + \alpha^2 + \beta^2$, on voit facilement que ses deux racines sont $\alpha \pm i\beta$. Trouvons un vecteur propre

$$V = \begin{pmatrix} x \\ y \end{pmatrix}$$

associé à la valeur propre $\lambda = \alpha + i\beta$. On a

$$\alpha x + \beta y = (\alpha + i\beta)x,$$

donc $\beta y = i\beta x$, et comme $\beta \neq 0$, on a $y = ix$. En prenant $x = 1$, on trouve

$$V = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

comme vecteur propre. Ainsi,

$$P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

est la matrice de passage d'une base propre de B . Notons que celle-ci est indépendante du choix de α et β définissant B . Ainsi,

$$P^{-1}BP = \begin{pmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{pmatrix},$$

donc

$$P^{-1}e^B P = \begin{pmatrix} e^{\alpha+i\beta} & 0 \\ 0 & e^{\alpha-i\beta} \end{pmatrix} = e^\alpha \begin{pmatrix} \cos \beta + i \sin \beta & 0 \\ 0 & \cos \beta - i \sin \beta \end{pmatrix}.$$

Puisque la matrice P est indépendante du choix de α et β , on peut appliquer le résultat au-dessus aux coefficients $\cos \beta$ et $\sin \beta$:

$$\begin{pmatrix} \cos \beta + i \sin \beta & 0 \\ 0 & \cos \beta - i \sin \beta \end{pmatrix} = P^{-1} \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} P.$$

Ceci démontre que

$$\exp \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = e^\alpha \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix}.$$

□

2.3.3 Le cas homogène

En utilisant les propriétés précédentes, nous arrivons finalement au résultat qui va permettre de trouver les solutions de $X'(t) = AX(t)$.

Lemme 2.3.7. *Soit $A \in \mathcal{M}_2(\mathbb{R})$. Alors l'application*

$$t \in \mathbb{R} \mapsto e^{tA}$$

est de classe¹ \mathcal{C}^1 , et

$$\frac{de^{tA}}{dt} = Ae^{tA} \quad \forall t \in \mathbb{R}.$$

Démonstration. On a

$$e^{tA} = \sum_{n \geq 0} \frac{(tA)^n}{n!} = \sum_{n \geq 0} \frac{A^n}{n!} t^n.$$

1. Cela signifie que toutes les applications $t \in \mathbb{R} \mapsto (e^{tA})_{ij}$, pour $1 \leq i, j \leq 2$, sont de classe \mathcal{C}^1 .

On va voir cette somme comme une série entière en t . Lorsqu'on prend le coefficient i, j (avec $1 \leq i, j \leq 2$) de cette égalité, on obtient une "vraie" série entière :

$$(e^{tA})_{i,j} = \sum_{n \geq 0} \frac{(A^n)_{i,j}}{n!} t^n.$$

Or, en appliquant A^n au vecteur $(0, 1)$ ou au vecteur $(1, 0)$ selon que $i = 1$ ou 2 , on voit que par définition de $\|\cdot\|$,

$$|(A^n)_{i,j}| \leq \|A^n\| \leq \|A\|^n$$

(la dernière inégalité est la sous-multiplicativité de $\|\cdot\|$). Ainsi le coefficient devant t^n dans la série entière $(e^{tA})_{i,j}$ est de module inférieur à $\|A\|^n/n!$. Par le critère de d'Alembert, cette série entière est de rayon de convergence infini. On peut donc dériver terme à terme et obtenir

$$\frac{d}{dt} \left((e^{tA})_{i,j} \right) = \sum_{n \geq 1} \frac{(A^n)_{i,j}}{n!} n t^{n-1} = \sum_{n \geq 1} \frac{(A^n)_{i,j}}{(n-1)!} t^{n-1} = \sum_{k \geq 0} \frac{(A^{k+1})_{i,j}}{k!} t^k$$

(la dernière égalité s'obtient par réindexation $k = n - 1$). Finalement,

$$\frac{d}{dt} (e^{tA}) = \sum_{k \geq 0} \frac{A^{k+1}}{k!} t^k = \sum_{k \geq 0} A \frac{(tA)^k}{k!} = A e^{tA}.$$

□

Nous arrivons donc finalement au résultat suivant :

Théorème 2.3.8. Soit $A \in \mathcal{M}_2(\mathbb{R})$, $t_0 \in \mathbb{R}$ et $X_0 \in \mathbb{R}^2$.

Le système différentiel muni de condition initiale :

$$X'(t) = AX(t) \quad \forall t \in \mathbb{R}, \quad X(t_0) = X_0 \quad (2.10)$$

admet une unique solution, donnée par

$$X(t) = e^{(t-t_0)A} X_0, \quad \forall t \in \mathbb{R}.$$

Démonstration. Le point remarquable ici est que l'exponentielle permet de montrer l'unicité des solutions sans avoir recours au théorème de Cauchy-Lipschitz général qui sera vu par la suite.

— Existence. On vérifie que la fonction \tilde{X} définie par

$$\tilde{X}(t) = e^{(t-t_0)A} X_0 \quad \forall t \in \mathbb{R}$$

est solution de (2.10). Cela découle immédiatement de la Proposition 2.3.3 (a) et (c), qui nous permet d'écrire $X(t) = e^{tA}(e^{-t_0A} X_0)$, et du Lemme 2.3.7.

— Unicité. Soit X une solution de (2.10). On pose

$$C(t) = e^{-tA} X(t).$$

Alors C est dérivable et

$$C'(t) = -Ae^{-tA} X(t) + e^{-tA} X'(t) = -e^{-tA} AX(t) + e^{-tA} AX(t),$$

puisque A et e^{-tA} commutent et que $X'(t) = AX(t)$. On obtient donc $C'(t) = 0$ pour tout $t \in \mathbb{R}$, donc C est une fonction constante : $C(t) = C(t_0) = e^{-t_0A}X_0$. En revenant à la définition de C , cela implique

$$e^{-tA}X(t) = e^{-t_0A}X_0 \quad \forall t \in \mathbb{R}.$$

On applique e^{tA} aux deux membres de l'équation et on obtient

$$X(t) = e^{(t-t_0)A}X_0.$$

□

Remarque 2.3.9. Attention : pour un système linéaire à coefficients dépendant du temps, c'est-à-dire dans le cas d'une équation du type

$$X'(t) = A(t)X(t),$$

avec $X : \mathbb{R} \mapsto \mathbb{R}^2$, $A : \mathbb{R} \mapsto \mathcal{M}_2(\mathbb{R})$, il est en général FAUX de dire

$$X(t) = e^{\int_{t_0}^t A(s)ds} X_0.$$

Cette égalité n'est vraie que si les matrices $A(s)$ et $A(t)$ commutent pour tout $s, t \in \mathbb{R}$.

2.3.4 Le cas non homogène

Lorsque le second membre n'est pas homogène, l'idée est la même que pour les équations différentielles scalaires du premier ordre (cas $m = n = 1$). On peut utiliser la méthode de variation des constantes pour démontrer l'existence de solutions, et éventuellement les calculer. On peut aussi utiliser une solution particulière de l'équation non-homogène, si on arrive à en trouver une.

Méthode de variation des constantes

Soit X la solution du problème

$$\begin{cases} X'(t) = AX(t) + B(t) & \forall t \in \mathbb{R}, \\ X(t_0) = X_0. \end{cases}$$

On définit, pour tout $t \in \mathbb{R}$,

$$C(t) = e^{-tA}X(t),$$

qui vérifie une équation de type plus simple, n'impliquant pas de terme $C(t)$:

$$C'(t) = -Ae^{-tA}X(t) + e^{-tA}X'(t) = e^{-tA}B(t).$$

Il "suffit" donc de trouver une primitive de $e^{-tA}B(t)$ pour connaître explicitement $C(t)$. On obtient alors X par la formule

$$X(t) = e^{tA}C(t).$$

Remarque 2.3.10. La présentation est faite ici pour des matrices carrées de taille 2, mais se généralise à des matrices carrées de taille quelconques, et donc à des systèmes linéaires à coefficients constants de m équations et d'ordre n avec n, m quelconques.

On a donc la proposition suivante :

Proposition 2.3.11. *Soit $A \in \mathcal{M}_2(\mathbb{R})$, I un intervalle de \mathbb{R} , $B : I \rightarrow \mathbb{R}^2$ une application continue. Soit $t_0 \in I$, $X_0 \in \mathbb{R}^2$.*

Alors le système différentiel

$$\begin{cases} X'(t) = AX(t) + B(t) & \forall t \in I, \\ X(t_0) = X_0 \end{cases} \quad (2.11)$$

admet une unique solution de classe \mathcal{C}^1 ,

$$X(t) = e^{(t-t_0)A} X_0 + \int_{t_0}^t e^{(t-s)A} B(s) ds.$$

Démonstration. Existence. On vérifie que l'application $t \rightarrow e^{(t-t_0)A} X_0 + \int_{t_0}^t e^{(t-s)A} B(s) ds$ est solution du problème de Cauchy (2.11).

Unicité. Comme dans le cas des équations scalaires, si X_1, X_2 sont deux solutions de classe \mathcal{C}^1 , alors $Y = X_1 - X_2$ est solution du système homogène et $Y(t_0) = 0$. Donc $Y = 0$. \square

Utilisation d'une solution particulière

Soit \tilde{X} une solution particulière de

$$\tilde{X}'(t) = A\tilde{X}(t) + B(t),$$

et soit $X_1 = \tilde{X}(t_0)$.

Soit X l'unique solution de (2.11). Alors $Y = X - \tilde{X}$ est solution de l'équation homogène

$$Y'(t) = AY(t), \quad Y(t_0) = X_0 - X_1.$$

Par conséquent, $Y(t) = e^{(t-t_0)A}(X_0 - X_1)$. On en déduit que

$$X(t) = \tilde{X}(t) + e^{(t-t_0)A}(X_0 - X_1).$$

Autrement dit, la solution générale de (2.11) est la somme d'une solution particulière et de la solution générale du système différentiel homogène associé.

Application. Cette méthode est aussi utile dans le cas des équations différentielles linéaires d'ordre deux à coefficients constants. En effet, soit $a, b \in \mathbb{R}$ et soit $f : \mathbb{R} \rightarrow \mathbb{R}$ continue. On considère l'équation différentielle

$$u''(t) + au'(t) + bu(t) = f(t). \quad (2.12)$$

On rappelle que cette équation peut être mise sous la forme d'un système d'ordre un en posant

$$X(t) = \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}, \quad B(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

Alors u est solution de (2.12) si et seulement si X est solution de (2.11), avec $X_0 = \begin{pmatrix} u(t_0) \\ u'(t_0) \end{pmatrix}$.

On peut donc appliquer ce qui précède pour trouver des solutions de (2.12). En particulier :

- On peut toujours mettre (2.12) sous la forme d'un système du type (2.11), et appliquer la méthode de variation des constantes. Néanmoins, les calculs peuvent rapidement s'avérer fastidieux.

- Supposons que l'on arrive à identifier une solution particulière \tilde{u} de (2.12). En particulier $\tilde{X}(t) = \begin{pmatrix} \tilde{u}(t) \\ \tilde{u}'(t) \end{pmatrix}$ est une solution particulière du système différentiel et nous avons vu que la solution générale de (2.11) est la somme de \tilde{X} et de la solution générale du système différentiel homogène associé, qui est facile à résoudre explicitement. On en déduit la solution générale de (2.12) comme somme de \tilde{u} et de la solution générale de l'équation différentielle homogène

$$v''(t) + av'(t) + bv(t) = 0,$$

c'est-à-dire $u(t) = \tilde{u}(t) + v(t)$.

Si on parvient à identifier une solution particulière, cette deuxième méthode est à privilégier car les calculs sont nettement plus faciles.

Remarque 2.3.12. Dans la pratique, pour chercher une solution particulière de (2.12), on pourra chercher une fonction \tilde{u} "du même type" que f , c'est-à-dire :

- Si f est un polynôme, on cherche \tilde{u} comme un polynôme ;
- Si f est une fonction trigonométrique ($f(t) = A \cos(\omega t) + B \sin(\omega t)$), avec $A, B, \omega \in \mathbb{R}$, on cherche \tilde{u} sous la forme

$$\tilde{u}(t) = A' \cos(\omega t) + B' \sin(\omega t);$$

- Si f est une exponentielle, on cherche u comme une exponentielle, etc.
- Plus généralement pour le cas $f = \Re(Ae^{\omega t})$ avec $A, \omega \in \mathbb{C}$, qui inclut les cas trigonométriques et exponentielle réelle, on cherche d'abord les solutions complexes de l'équation avec membre non-homogène $Ae^{\omega t}$ sous la forme :
 - $Be^{\omega t}$, avec $B \in \mathbb{C}$, si ω n'est pas racine de l'équation caractéristique $\lambda^2 + a\lambda + b = 0$;
 - $Bte^{\omega t}$, avec $B \in \mathbb{C}$, si ω est racine simple de l'équation caractéristique ;
 - $Bt^2e^{\omega t}$, avec $B \in \mathbb{C}$, si ω est racine double de l'équation caractéristique.

Ensuite on retrouve les solutions réelles cherchées initialement en prenant la partie réelle des solutions complexes.

2.3.5 Lien entre les deux méthodes

Pour calculer les solutions des équations linéaires à coefficients constants du type

$$x''(t) + ax'(t) + bx(t) = 0, \quad x(0) = x_0, x'(0) = x_1 \quad (2.13)$$

on a donc vu deux méthodes différentes :

1. Première méthode donnée par le Théorème 2.2.3 : recherche de solutions comme des combinaisons linéaires d'exponentielles en résolvant sur \mathbb{C} l'équation caractéristique $\lambda^2 + a\lambda + b = 0$. On note λ_1, λ_2 ses racines.
 - (i) Si $\lambda_1, \lambda_2 \in \mathbb{R}$ distinctes, la solution est de la forme

$$x(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}, \quad \text{avec } A_1, A_2 \in \mathbb{R}.$$

- (ii) Si les racines sont non-réelles $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta, \beta \neq 0$, alors la solution est de la forme

$$x(t) = e^{t\alpha}(C_1 \cos(\beta t) + C_2 \sin(\beta t)), \quad \text{avec } C_1, C_2 \in \mathbb{R}.$$

(iii) Si il y a une racine double $\lambda_1 = \lambda_2 = \lambda$, alors la solution est de la forme

$$x(t) = (A_1 t + A_2) e^{\lambda t}, \text{ avec } A_1, A_2 \in \mathbb{R}.$$

2. Deuxième méthode donnée par le Théorème 2.3.8 : on se ramène à l'étude d'un système d'équations d'ordre 1 en posant

$$X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}.$$

On a alors

$$X(t) = e^{tA} X_0,$$

$$\text{où } X_0 = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}.$$

En vue des Théorèmes 2.2.3 et 2.3.8 on sait que les deux formules donnent le même résultat, mais on va aussi le vérifier explicitement pour mieux comprendre le lien entre les deux méthodes. Tout d'abord, remarquons que le polynôme caractéristique de la matrice A est

$$\chi_A(\lambda) = -\lambda(-a - \lambda) + b = \lambda^2 + a\lambda + b.$$

Autrement dit, les racines de l'équation caractéristique de l'équation différentielle (2.13) sont exactement les valeurs propres de A . Notons comme au-dessus λ_1, λ_2 les racines complexes distinctes ou confondues de χ_A .

On va vérifier que la formule donnée par la deuxième méthode donne bien le même résultat que la première méthode, en utilisant les formules du tableau précédent.

(i) Si $\lambda_1, \lambda_2 \in \mathbb{R}$ distinctes il existe $P \in GL_2(\mathbb{R})$ tel que

$$A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1},$$

et on a

$$X(t) = P \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} P^{-1} X_0.$$

Notons

$$P^{-1} X_0 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad P = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}.$$

On en déduit que

$$x(t) = a_1 e^{\lambda_1 t} y_1 + a_2 e^{\lambda_2 t} y_2,$$

et on retrouve le résultat de la première méthode.

(ii) Si les racines sont non-réelles $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta, \beta \neq 0$ il existe $P \in GL_2(\mathbb{R})$ tel que

$$A = P \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} P^{-1},$$

et on a

$$X(t) = P e^{\alpha t} \begin{pmatrix} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{pmatrix} P^{-1} X_0.$$

On en déduit que

$$x(t) = a_1 e^{\alpha t} \cos(\beta t) y_1 - a_2 e^{\alpha t} \sin(\beta t) y_2,$$

et on retrouve le résultat de la première méthode.

(iii) S'il y a une racine double $\lambda_1 = \lambda_2 = \lambda$ alors la matrice A n'est pas diagonalisable (si elle l'était il existerait $P \in GL_2(\mathbb{R})$ tel que $A = P \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} P^{-1} = \lambda I_2$ en contradiction avec $A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}$), donc il existe $P \in GL_2(\mathbb{R})$, $\gamma \in \mathbb{R}^*$ tel que

$$A = P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1},$$

de sorte que

$$X(t) = P e^{\lambda t} \begin{pmatrix} 1 & \gamma t \\ 0 & 1 \end{pmatrix} P^{-1} X_0.$$

Avec les mêmes notations que ci-dessus, on obtient

$$x(t) = e^{\lambda t} [(a_1 y_1 + a_2 y_2) + a_1 y_2 \gamma t],$$

et ici encore on retrouve la formule de la première méthode.

2.4 Tracé des trajectoires de solutions d'équations différentielles linéaires à coefficients constants

2.4.1 Mécanique du point, portrait de phase

La mécanique du point est l'étude du mouvement des points matériels (ou des solides pouvant être décrits comme des points matériels). Le point de départ est le principe fondamental de la dynamique : le produit de la masse par l'accélération (c'est-à-dire la dérivée seconde de la position) est égal à la somme totale des forces s'exerçant sur le solide :

$$F = m x''(t).$$

Supposons que le mouvement se fasse sur une droite et qu'il ne dépende que d'un paramètre : autrement dit, $x(t) \in \mathbb{R}$, et que la force F puisse s'écrire comme une fonction de la position et de la vitesse : $F = \varphi(x(t), x'(t))$. On peut alors réécrire le principe fondamental de la dynamique sous la forme

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} = \begin{pmatrix} x'(t) \\ \frac{1}{m} \varphi(x(t), x'(t)) \end{pmatrix} = f(x(t), x'(t)), \quad (2.14)$$

où

$$f(x, v) = \begin{pmatrix} v \\ \frac{1}{m} \varphi(x, v) \end{pmatrix}.$$

On obtient donc un système différentiel du premier ordre dans \mathbb{R}^2 . On peut tracer les courbes représentatives des solutions dans le plan cartésien comme des courbes paramétrées, en mettant $x(t)$ en abscisse et $v = x'(t)$ en ordonnée. Une telle représentation graphique s'appelle un **portrait de phase**. À chaque paire de conditions initiales (x_0, v_0) correspond en général une courbe paramétrée (sous réserve qu'on ait existence et unicité des solutions de l'équation différentielle (2.14) pour chaque donnée initiale).

Exemple 2.4.1 (L'oscillateur harmonique, ou pendule élastique). L'équation du pendule élastique est donné par la loi de Newton

$$m x''(t) = -K x(t),$$

où $K > 0$ est la constante de raideur du ressort et $x(t)$ est l'élongation du ressort par rapport à la position d'équilibre qu'a le pendule au repos. En notant $\omega = \sqrt{\frac{K}{m}}$ on obtient l'équation différentielle

$$x''(t) + \omega_0^2 x(t) = 0, \quad (2.15)$$

que l'on munit des conditions initiales

$$x(0) = x_0, \quad x'(0) = v_0.$$

On sait alors que la forme générale des solutions de (2.15) est

$$x(t) = A \cos(\omega_0 t) + B \sin(\omega_0 t),$$

où A et B sont déterminés par les conditions initiales. Plus précisément, on a

$$x(0) = x_0 = A,$$

et comme

$$x'(t) = -\omega_0 A \sin(\omega_0 t) + \omega_0 B \cos(\omega_0 t),$$

on en déduit que

$$x'(0) = v_0 = \omega_0 B.$$

Pour tracer le portrait de phase associé à l'équation (2.15), il faut donc, pour chaque couple $(x_0, v_0) \in \mathbb{R}^2$, étudier la courbe paramétrée \mathcal{C}_{x_0, v_0} d'équation cartésienne

$$\begin{aligned} x(t) &= x_0 \cos(\omega_0 t) + \frac{v_0}{\omega_0} \sin(\omega_0 t), \\ y(t) &= v_0 \cos(\omega_0 t) - \omega_0 x_0 \sin(\omega_0 t), \end{aligned} \quad t \in \mathbb{R}.$$

Si $\omega_0 = 1$, on reconnaît facilement l'équation d'un cercle de centre zéro et de rayon $\sqrt{x_0^2 + v_0^2}$. Plus généralement, si ω_0 est quelconque, on vérifie facilement que

$$x(t)^2 + \frac{y(t)^2}{\omega_0^2} = x_0^2 + \frac{v_0^2}{\omega_0^2},$$

donc la courbe \mathcal{C}_{x_0, v_0} est une ellipse. On peut trouver son grand axe et son petit axe en mettant l'équation sous la forme canonique

$$\frac{x(t)^2}{a^2} + \frac{y(t)^2}{b^2} = 1,$$

où

$$a = \sqrt{x_0^2 + \frac{v_0^2}{\omega_0^2}}, \quad b = \omega_0 a = \sqrt{\omega_0^2 x_0^2 + v_0^2}.$$

En particulier, si $\omega_0 > 1$, on a $b > a$, et par conséquent le grand axe de l'ellipse est suivant (Oy) . Si $\omega_0 < 1$, alors $b < a$, et le grand axe de l'ellipse est suivant (Ox) .

Par exemple dans le cas $\omega_0 = 2, x_0 = -2, v_0 = 0$, c'est-à-dire départ du ressort comprimé à hauteur 2 par rapport à la position d'équilibre et sans vitesse initiale, on a

$$x(t) = -2 \cos(2t), \quad x'(t) = 4 \sin(2t).$$

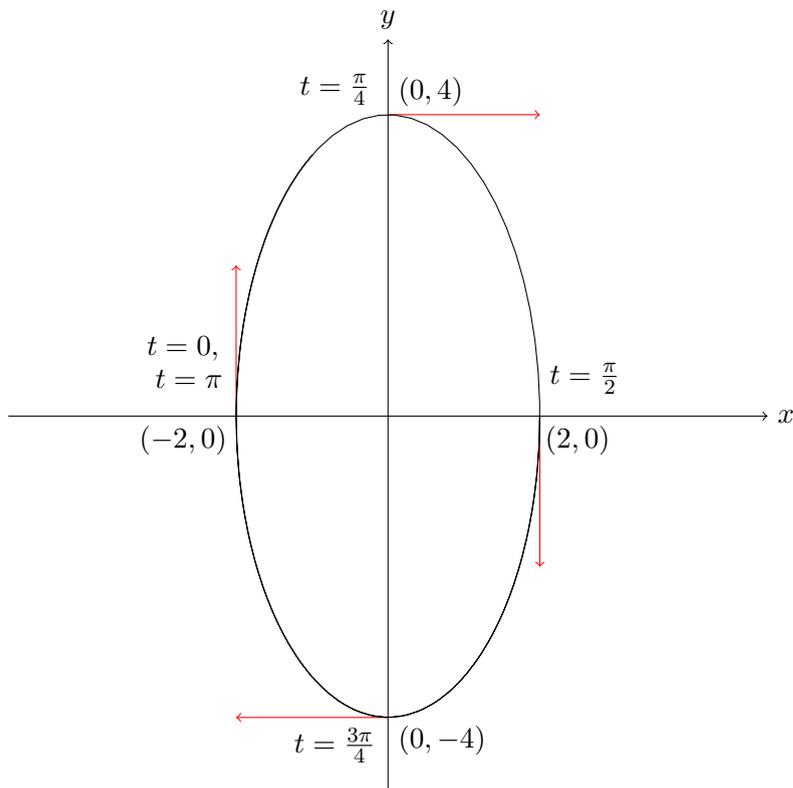
La trajectoire est π -périodique, il n'y a pas de points stationnaires, le vecteur tangent

$$T(t) = (x'(t), x''(t)) = (4 \sin(2t), 8 \cos(2t)),$$

vaut

$$T(0) = (0, 8), \quad T\left(\frac{\pi}{4}\right) = (4, 0), \quad T\left(\frac{\pi}{2}\right) = (0, -8), \quad T\left(\frac{3\pi}{4}\right) = (0, -8),$$

donc on a des tangentes verticales à $t = 0$ et $t = \frac{\pi}{2}$ et des tangentes horizontales à $t = \frac{\pi}{4}$ et $t = \frac{3\pi}{4}$. Le tracé de la solution $(x(t), x'(t))$ est le suivant, avec en rouge la direction des tangentes :



En particulier on peut lire sur le portrait de phase que de $t = 0$ jusqu'à $t = \frac{\pi}{4}$ la décompression du ressort fait que $x(t)$ s'approche la position d'équilibre, et sa vitesse croît. Passé le point d'équilibre à $t = \frac{\pi}{4}$ la résistance du ressort fait que la vitesse décroît jusqu'à s'annuler à $t = \frac{\pi}{2}$, temps auquel le ressort se retrouve allongé à distance 2 par rapport à la position d'équilibre. Ensuite jusqu'à $t = \pi$ on retrouve le même type de mouvement, dirigé vers le haut.

À la fin du poly on verra le portrait de phase du pendule simple, ainsi que celui de Lotka-Voterra. On pourra se faire une idée de comment est produit le portrait de phase du pendule simple en consultant l'animation de l'article suivant (par ailleurs fort intéressant !) : <http://images.math.cnrs.fr/Le-Rulpidon-5845.html>

L'objet de ce chapitre est de tracer les courbes paramétrées définies par $\{(x(t), x'(t)), t \in \mathbb{R}\}$, où x est solution de l'équation scalaire d'ordre 2 à coefficients constants

$$x''(t) + ax'(t) + bx(t) = 0, \quad x(0) = x_0, \quad x'(0) = x_1,$$

et plus généralement tracer les courbes paramétrées définies par $\{(x(t), y(t)), t \in \mathbb{R}\}$, où $X(t) = (x(t), y(t))$ est solution du système d'équations d'ordre 1 à coefficients constants

$$\begin{cases} X'(t) = AX(t) \\ X(0) = (x_0, y_0). \end{cases}$$

Il s'agit donc de tracer des courbes paramétrées définies par

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{At} X_0, \quad t \in \mathbb{R},$$

avec $A \in \mathcal{M}_2(\mathbb{R})$ et $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \mathbb{R}^2$.

2.4.2 Quelques cas simples importants

Avant de comprendre le cas général, on commence par des cas simples : matrice diagonale, un cas "modèle" correspondant à une matrice diagonalisable dans \mathbb{C} mais pas dans \mathbb{R} , et enfin le cas d'une matrice non-diagonalisable de type triangulaire supérieure. On donnera également un lemme permettant de comprendre le lien entre :

- d'une part la similitude de deux matrices A et \tilde{A} ;
- d'autre part le tracé des courbes paramétrées associées à $X(t) = e^{tA} X_0$ et à $Y(t) = e^{t\tilde{A}} Y_0$.

(a) Tracé de la courbe paramétrée d'équation $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{tA} X_0$, avec A diagonale :

On a

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

avec $\lambda_1, \lambda_2 \in \mathbb{R}$. On étudie donc la courbe paramétrée \mathcal{C}_{A, X_0} définie par

$$\begin{cases} x(t) = x_0 e^{\lambda_1 t}, \\ y(t) = y_0 e^{\lambda_2 t}, \end{cases} \quad t \in \mathbb{R}.$$

Le domaine d'étude est \mathbb{R} , et pour $X_0 \in \mathbb{R}^2$ fixé, les courbes n'ont pas de symétrie évidente. Cependant on peut remarquer que la courbe relative à la condition initiale $(-x_0, y_0)$ (resp. $(x_0, -y_0)$) se déduit de celle relative à la condition initiale (x_0, y_0) par une symétrie par rapport à l'axe des ordonnées (resp. par rapport à l'axe des abscisses). On peut donc se contenter d'étudier le cas où $x_0 \geq 0, y_0 \geq 0$ et en déduire les autres cas par symétrie.

On traite d'abord le cas $x_0 = 0$ (resp. $y_0 = 0$). On a alors $x(t) = e^{\lambda_1 t} x_0 = 0$ pour tout $t \in \mathbb{R}$ (resp. $y(t) = e^{\lambda_2 t} y_0 = 0$ pour tout t). Si $\lambda_2 = 0$ (resp. $\lambda_1 = 0$) la trajectoire est le point $(0, y_0)$ (resp. $(x_0, 0)$). Sinon on voit facilement que la trajectoire est le demi-axe des ordonnées positives si $y_0 > 0$ (resp. le demi-axe des abscisses positives si $x_0 > 0$), et le point $(0, 0)$ si $y_0 = 0$ (resp. $x_0 = 0$).

On se concentre donc à présent sur le cas où $x_0 > 0, y_0 > 0$. Si $\lambda_1 = 0$ et $\lambda_2 = 0$ la trajectoire est le point (x_0, y_0) . Si $\lambda_1 = 0$ et $\lambda_2 \neq 0$ la trajectoire est la demi-droite d'équation $x = x_0, y > 0$. On se place dans le dernier cas restant, $\lambda_1 \neq 0$. On remarque alors que $x(\mathbb{R}) = \mathbb{R}_+^*$. On fait alors le changement de paramétrage

$$x(t) = s \iff t = \frac{1}{\lambda_1} \log \frac{s}{x_0},$$

ce qui implique

$$y(t) = y_0 e^{\lambda_2 \left(\frac{1}{\lambda_1} \log \frac{s}{x_0} \right)} = \frac{y_0}{x_0^{\frac{\lambda_2}{\lambda_1}}} s^{\frac{\lambda_2}{\lambda_1}}.$$

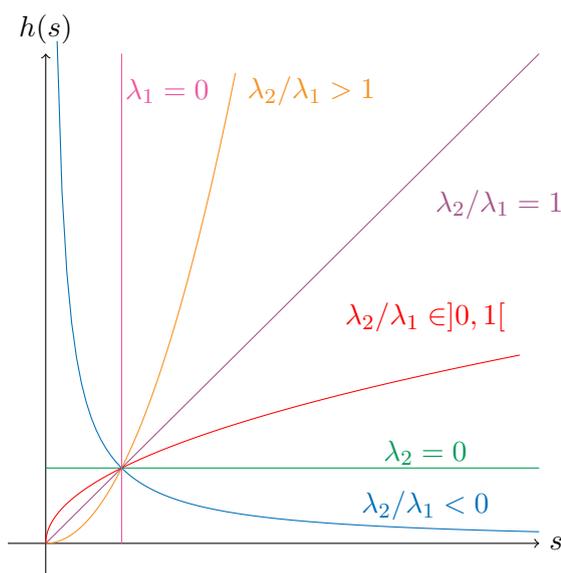
La courbe \mathcal{C}_{A,X_0} est donc la courbe \mathcal{C}_0 du graphe de la fonction

$$h(s) = \frac{y_0}{x_0^{\frac{\lambda_2}{\lambda_1}}} s^{\frac{\lambda_2}{\lambda_1}},$$

pour $s \in \mathbb{R}_+^*$. En vue du changement de paramétrage fait ci-dessus, si $\lambda_1 > 0$ le sens de parcours de la courbe \mathcal{C}_{A,X_0} lorsque t varie de $-\infty$ à $+\infty$ est le même que celui de \mathcal{C}_0 lorsque s varie de 0^+ à $+\infty$, et si $\lambda_1 < 0$ c'est le sens opposé. L'allure de la courbe \mathcal{C}_0 dépend de la valeur de $\frac{\lambda_2}{\lambda_1}$:

- (i) Si $\lambda_2/\lambda_1 < 0$, la fonction $s \in \mathbb{R}_+^* \mapsto s^{\frac{\lambda_2}{\lambda_1}}$ est strictement décroissante, tend vers $+\infty$ en 0^+ et vers 0 en $+\infty$. En particulier la courbe \mathcal{C}_0 admet en $s = 0^+$ une asymptote à l'axe des ordonnées, et en $s = +\infty$ une asymptote à l'axe des abscisses.
- (ii) Si $\lambda_2/\lambda_1 = 0$, la courbe \mathcal{C}_0 est la demi-droite d'équation $y = y_0$, $s > 0$.
- (iii) Si $\lambda_2/\lambda_1 \in]0, 1[$, la fonction $s \in \mathbb{R}_+^* \mapsto s^{\frac{\lambda_2}{\lambda_1}}$ est strictement croissante, tend vers 0 en 0^+ et vers $+\infty$ en $+\infty$. De plus la courbe \mathcal{C}_0 admet une tangente verticale au voisinage du point $(0,0)$ et une branche parabolique dans la direction de l'axe des abscisses quand $s \rightarrow +\infty$.
- (iv) Si $\lambda_2/\lambda_1 = 1$, la courbe \mathcal{C}_0 est la demi-droite d'équation $y = \frac{y_0}{x_0} s$, $s > 0$.
- (v) Si $\lambda_2/\lambda_1 > 1$, la fonction $s \in \mathbb{R}_+^* \mapsto s^{\frac{\lambda_2}{\lambda_1}}$ est strictement croissante, tend vers 0 en 0^+ et vers $+\infty$ en $+\infty$. De plus la courbe \mathcal{C}_0 admet une tangente horizontale au voisinage du point $(0,0)$ et une branche parabolique dans la direction de l'axe des ordonnées quand $s \rightarrow +\infty$.

L'allure des courbes est donc la suivante (avec ici $x_0 = y_0 = 1$) :



- (b) Tracé de la courbe paramétrée d'équation $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{tA} X_0$, avec $A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$, $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}^*$:

Il s'agit d'une matrice réelles ayant deux valeurs propres complexes conjuguées $\alpha \pm i\beta$, dont on a déjà vu la forme de l'exponentielle :

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{\alpha t} \begin{pmatrix} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

donc

$$\begin{cases} x(t) = e^{\alpha t}(\cos(\beta t)x_0 - \sin(\beta t)y_0), \\ y(t) = e^{\alpha t}(\cos(\beta t)y_0 + \sin(\beta t)x_0), \end{cases}$$

Nous allons privilégier un passage en coordonnées polaires. Si $(x_0, y_0) = (0, 0)$ et la trajectoire est réduite à un point. Sinon notons $(r_0, \theta_0) \in \mathbb{R}_+^* \times [0, 2\pi[$ l'unique couple tel que

$$x_0 = r_0 \cos \theta_0, \quad y_0 = r_0 \sin \theta_0.$$

Alors

$$\begin{cases} x(t) = r_0 e^{\alpha t}(\cos(\beta t) \cos(\theta_0) - \sin(\beta t) \sin(\theta_0)) = r_0 e^{\alpha t} \cos(\beta t + \theta_0), \\ y(t) = r_0 e^{\alpha t}(\cos(\beta t) \sin(\theta_0) + \sin(\beta t) \cos(\theta_0)) = r_0 e^{\alpha t} \sin(\beta t + \theta_0). \end{cases}$$

On remarque immédiatement que si $\alpha = 0$, la courbe est le cercle de rayon r_0 . On suppose donc dans toute la suite que $\alpha \neq 0$.

Nous allons tracer l'allure de la courbe \mathcal{C}_{A, X_0} pour $\theta_0 = 0$ et $\alpha, \beta > 0$ car le cas $\theta_0 \neq 0$ s'obtient en faisant une rotation d'angle θ_0 , le cas $\beta < 0$ s'obtient en faisant une symétrie par rapport à l'axe $(0x)$, et le cas $\alpha < 0$ s'obtient, en changeant la paramétrisation en $s = -t$, en changeant le sens de parcours de la courbe et en faisant une symétrie par rapport à l'axe $(0x)$.

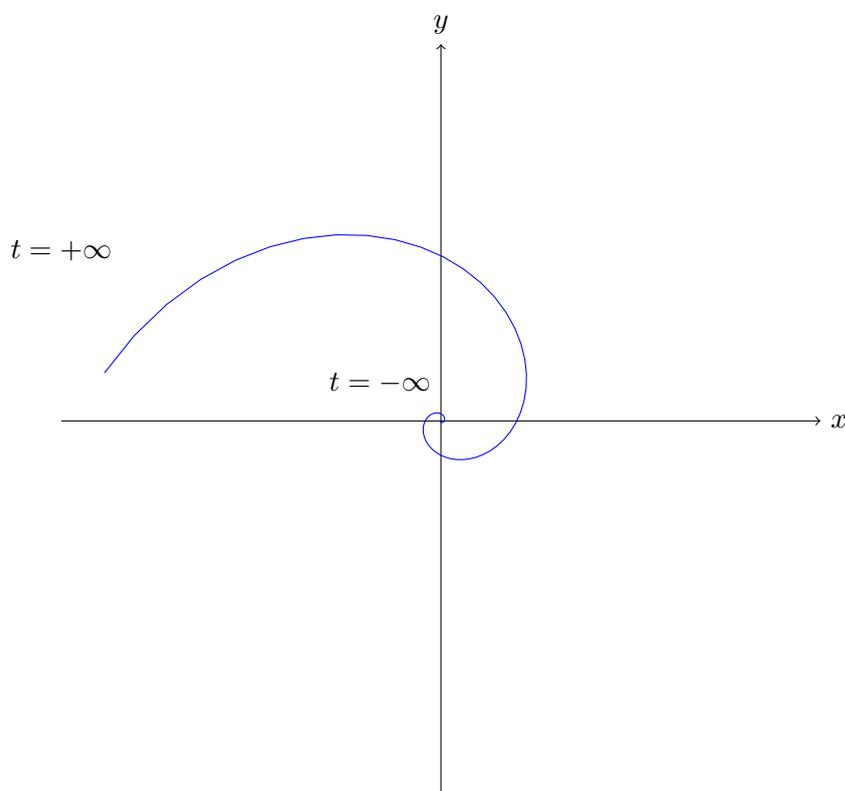
En faisant le changement de paramétrage $s = \beta t$ on obtient que \mathcal{C}_{A, X_0} est la courbe

$$\{(r_0 e^{\frac{\alpha}{\beta} s} \cos(s), r_0 e^{\frac{\alpha}{\beta} s} \sin(s)), s \in \mathbb{R}\},$$

d'équation polaire

$$r(s) = r_0 e^{\frac{\alpha}{\beta} s}.$$

On a $r'(s) = r_0 \frac{\lambda}{e} \frac{\alpha}{\beta} s > 0$, donc $r(s)$ croît de $\lim_{s \rightarrow -\infty} r(s) = 0$ à $\lim_{s \rightarrow +\infty} r(s) = +\infty$. On obtient une spirale qui s'enroule autour de l'origine quand $t \rightarrow -\infty$ et qui part vers l'infini quand $t \rightarrow +\infty$.



(c) **Tracé de la courbe paramétrée d'équation** $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{tA} X_0$, **avec A non-diagonalisable triangulaire supérieure :**

On prend à présent A de la forme

$$A = \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix},$$

avec $\lambda, \gamma \in \mathbb{R}$, $\gamma \neq 0$. On étudie donc la courbe paramétrée \mathcal{C}_{A, X_0} définie par

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{tA} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = e^{\lambda t} \begin{pmatrix} 1 & \gamma t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} (x_0 + \gamma t y_0) e^{\lambda t} \\ y_0 e^{\lambda t} \end{pmatrix}.$$

Si $\lambda = 0$, on a donc $g(t) = y_0$ pour tout t . On vérifie facilement que dans ce cas \mathcal{C}_{A, X_0} est la droite d'équation $y = y_0$ si $y_0 \neq 0$ et elle est réduite au point (x_0, y_0) si $y_0 = 0$. De même si $y_0 = 0$, $x_0 > 0$ (resp. $x_0 < 0$) \mathcal{C}_{A, X_0} est la demi-droite d'équation $y = 0$, $x > 0$ (resp. $x < 0$). Enfin si $y_0 = x_0 = 0$ la trajectoire est réduite au point $(0, 0)$.

On suppose donc à présent $\lambda \neq 0$, $y_0 \neq 0$. Comme les fonctions $x(t)$ et $y(t)$ n'ont pas de symétrie évidente, on passe à l'analyse de leurs variations.

Les fonctions $x(t)$ et $y(t)$ sont dérivables sur \mathbb{R} et on a

$$\begin{cases} x'(t) = e^{\lambda t} (\lambda x_0 + \lambda \gamma y_0 t + \gamma y_0), \\ y'(t) = \lambda y_0 e^{\lambda t}. \end{cases}$$

Ainsi $y'(t)$ ne s'annule pas sur \mathbb{R} et a le signe de λy_0 . La fonction $x'(t)$ s'annule en un unique point t_0 défini par

$$t_0 = -\frac{\lambda x_0 + \gamma y_0}{\lambda \gamma y_0} = -\frac{1}{\lambda} - \frac{x_0}{\gamma y_0}.$$

On a plusieurs cas de figure possibles suivant les signes de $\lambda, \lambda y_0, \lambda \gamma y_0$. Notons néanmoins que :

- On a toujours $x(t) \rightarrow 0, y(t) \rightarrow 0$ quand $\lambda t \rightarrow -\infty$ (c.à.d. $t \rightarrow -\text{sgn}(\lambda)\infty$), et de plus $y(t)/x(t) \rightarrow 0$ quand $t \rightarrow -\text{sgn}(\lambda)\infty$: il y a une tangente horizontale au voisinage du point $(0, 0)$.
- On a toujours $y(t) \rightarrow \text{sgn}(\lambda y_0)\infty, x(t) \rightarrow \text{sgn}(\gamma y_0)\infty$ quand $\lambda t \rightarrow +\infty$ (c.à.d. $t \rightarrow \text{sgn}(\lambda)\infty$). De plus $y(t)/x(t) \rightarrow 0$ quand $t \rightarrow \text{sgn}(\lambda)\infty$: il y a une branche parabolique dans la direction $(0x)$.
- $y(t)$ garde un signe constant (celui de λy_0).
- Il y a une tangente verticale au point de paramètre t_0 , qui a pour coordonnées $\exp(\lambda t_0)(-\frac{\gamma y_0}{\lambda}, y_0)$.
- La courbe croise l'axe des ordonnées en un unique point, qui a pour paramètre $t_1 = -x_0/(\alpha y_0)$. Notons que $t_1 = t_0 + \frac{1}{\lambda}$. Le point de paramètre t_1 est donc "le plus proche de la branche infinie".

L'allure de la courbe est donc toujours la même, modulo des symétries par rapport aux axes (Ox) et (Oy) . Plus précisément :

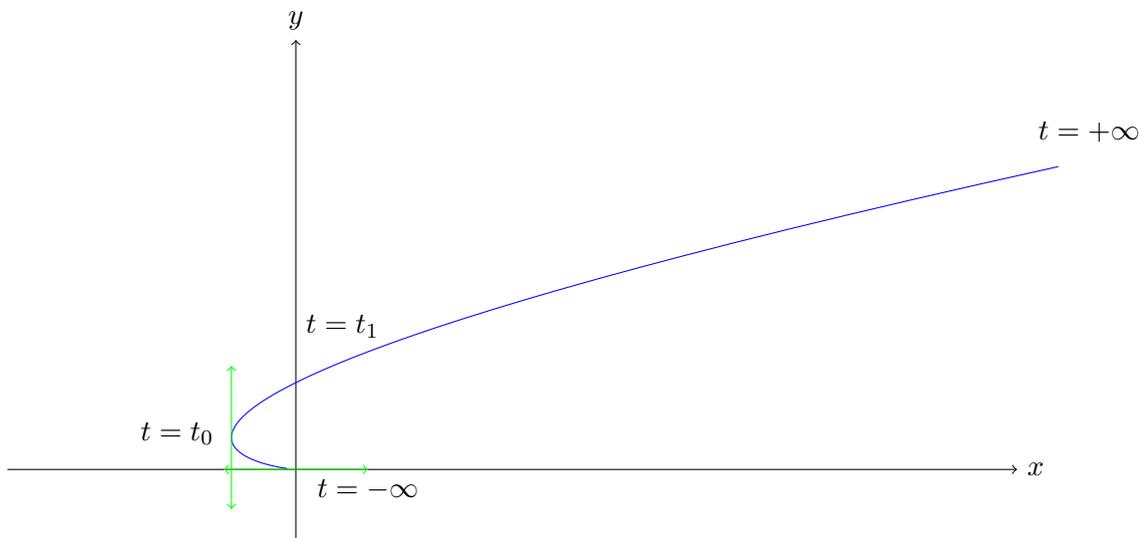
- Le signe de y_0 détermine le demi-plan dans lequel la courbe est incluse ($y > 0$ ou $y < 0$) ;
- Le signe de $\lambda \gamma y_0$ (ou de $\frac{\gamma y_0}{\lambda}$) détermine la position du point de tangente verticale par rapport à l'axe des ordonnées ;
- Le signe de λ détermine le sens dans lequel la courbe est parcourue (de 0 vers $+\infty$ ou le contraire).

On trace donc ci-dessous un exemple dans le cas $\gamma, \lambda, y_0 > 0$. Les autres cas sont laissés au lecteur et se traitent de façon rigoureusement analogue.

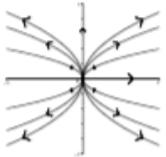
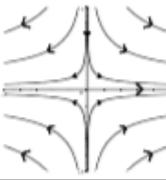
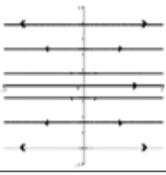
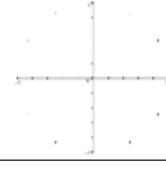
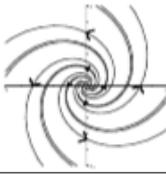
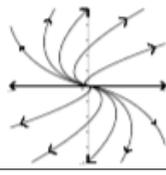
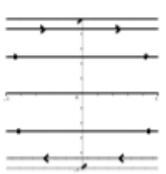
Dans ce cas le tableau de variations est le suivant :

t	$-\infty$		t_0		$+\infty$	
$x'(t)$	0	-	0	+		
$x(t)$	0	$-e^{\lambda t_0} \frac{\gamma y_0}{\lambda}$			$+\infty$	
$y(t)$	0	\longrightarrow				$+\infty$
$y'(t)$	0	+		+		

On obtient une courbe qui a l'allure suivante :



Nous résumons dans le tableau suivant les tracé type obtenues pour les trajectoires des solutions de $X'(t) = AX(t)$, et rappelons que le sens de parcours des trajectoires est inversé lorsque l'on change le signe de $\lambda_1, \alpha, \beta, \lambda, \gamma$.

Type	Sous-type	Trajectoires
Matrice (diagonalisable avec valeurs propres distinctes ou confondues $\lambda_1, \lambda_2 \in \mathbb{R}$) de forme $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ si $\frac{\lambda_2}{\lambda_1} < 1$; une rotation de $\frac{\pi}{2}$ de ce type de tracé si $\frac{\lambda_2}{\lambda_1} > 1$	valeurs propres distinctes de même signe, $\lambda_1 > 0$	
	valeurs propres de signes opposés, $\lambda_1 > 0$	
	valeur propre double, $\lambda_1 > 0$	
	$\lambda_2 = 0, \lambda_1 > 0$	
	$\lambda_2 = 0 = \lambda_1$	
Matrice (diagonalisable avec valeurs propres non-réelles $\alpha \pm i\beta, \beta \neq 0$), de forme $\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$	$\alpha > 0$; si $\beta > 0$; un symétrique de ce tracé par rapport à l'axe Ox si $\beta < 0$	
	$\alpha = 0, \beta > 0$	
Matrice (non diagonalisable), de forme $\begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix}$	valeur propre double $\lambda > 0$; si $\gamma > 0$; le symétrique de ce tracé par rapport à l'axe Oy si $\gamma < 0$	
	valeur propre double $\lambda = 0$	

Enfin, voici quelques remarques :

- en vue de la formule explicite des solutions, les trajectoires du système avec matrice $\begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}$, $\lambda_1, \lambda_2 \in \mathbb{R}$, s'obtiennent de celles du système avec matrice $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ par une symétrie par rapport à la droite $y = x$. Ceci s'explique aussi par le fait que la matrice de passage entre ces deux matrices est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, donc comme on le verra dans le prochain lemme, la trajectoire de la première en repère canonique Oxy est celle de deuxième en repère inversé,
- lorsque la matrice du système est $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ avec $\lambda_1, \lambda_2 \in \mathbb{R}$ distincts, la direction des vecteurs propres de la valeur propre λ_1 est celle de l'axe des abscisses, et la direction des vecteurs propres de la valeur propre λ_2 est celle de l'axe des ordonnées,
- les trajectoires issues d'un vecteur propre réel v d'une valeur propre λ restent sur la direction du vecteur propre car

$$e^{tA}v = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} v = \sum_{k=0}^{\infty} \frac{(t\lambda)^k}{k!} v = e^{t\lambda}v;$$

en effet, $e^{t\lambda}v$ est bien une solution car $(e^{t\lambda}v)' = e^{t\lambda}\lambda v = e^{t\lambda}Av = Ae^{t\lambda}v$. Si $\lambda = 0$ la solution est constante $X(t) = v$, et sa trajectoire est symbolisée sur le portrait de phase par une petite croix. Aussi, comme tout vecteur colinéaire à v est aussi vecteur propre de la même valeur propre λ , on retrouve sur le portrait de phase une droite entière de points représentant des solutions constantes.

Si $\lambda \neq 0$ alors la trajectoire $e^{t\lambda}v$ lorsque t varie de $-\infty$ à $+\infty$ est la demi-droite reliant l'origine (sans la contenir) et l'infini dans la direction v . Comme $-v$ est aussi un vecteur propre de la même valeur propre λ , on retrouve sur le portrait de phase les deux demi-droites données par la direction de v ,

- les directions des vecteurs propres sont donc des directions spéciales, le long desquelles on retrouve sur les portraits de phases ou des droites de points générant des solutions constantes, ou des trajectoires de type demi-droites. De plus, ces directions nous dictent l'allure de tout le portrait de phases. Par exemple, dans le premier cas du tableau, on a l'origine comme trajectoire de la solution constante symbolisée sur le portrait de phase par une petite croix, les trajectoires demi-droites au long des axes Ox et Oy , et tous les autres trajectoires courbes, confinées dans les quarts de plan déterminés par les axes Ox et Oy (directions des vecteurs propres de λ_1 et λ_2 respectivement), et construites comme suit : départ à côté de l'origine dans une direction tangente à l'axe Oy et ensuite branche parabolique à l'infini en direction de l'axe Ox , et enfin le sens des trajectoires est donné par le signe commun des valeurs propres - à travers les trajectoires on s'écarte de l'origine lorsque les valeurs propres sont strictement positives, et on s'approche de l'origine lorsque les valeurs propres sont strictement négatives,
- deux trajectoires non-confondues ne peuvent pas se croiser, car on a unicité des solutions,
- toujours grâce à l'unicité des solutions, la trajectoire issue de X_0 coïncide avec celle issues de $X(t_0)$ quel que soit $t_0 \in \mathbb{R}$.

2.4.3 Cas général

Les cas étudiés ci-dessus permettent en fait de comprendre le tracé des solutions dans le cas général d'une matrice $A \in \mathcal{M}_2(\mathbb{R})$. On va déduire du paragraphe précédent l'allure de la courbe paramétrée \mathcal{C}_{A,X_0} définie par

$$X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{tA} X_0, \quad t \in \mathbb{R},$$

avec $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \mathbb{R}^2$ et $A \in \mathcal{M}_2(\mathbb{R})$ quelconque. Il existe donc $P \in GL_2(\mathbb{R})$ tel que

$$A = P\tilde{A}P^{-1},$$

avec

- $\tilde{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ si A est diagonalisable avec valeurs propres $\lambda_1, \lambda_2 \in \mathbb{R}$ distinctes ou confondues,
- $\tilde{A} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$ si A est diagonalisable avec valeurs propres non-réelles $\alpha \pm i\beta, \beta \neq 0$,
- $\tilde{A} = \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix}$ si A est non-diagonalisable avec une valeur propre double $\lambda \in \mathbb{R}$.

Nous avons donc que la courbe \mathcal{C}_{A,X_0} est définie dans le repère canonique (Oxy) par les points de coordonnées

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P(e^{t\tilde{A}}P^{-1}X_0), \quad t \in \mathbb{R}.$$

Nous allons utiliser le résultat suivant :

Lemme 2.4.2. Soit $X : I \rightarrow \mathbb{R}^2, Y : I \rightarrow \mathbb{R}^2$ deux applications telles qu'il existe $P \in GL_2(\mathbb{R})$ telle que

$$X(t) = PY(t) \quad \forall t \in I.$$

On note $X = (x_1, x_2), Y = (y_1, y_2), v_1, v_2$ les vecteurs colonnes de la matrice $P, \mathcal{R}_P = (O, v_1, v_2)$ le repère dont les vecteurs de base sont v_1 et v_2 . Soit \mathcal{C} la courbe paramétrée définie dans le repère canonique (Oxy) par $\{(x_1(t), x_2(t)), t \in I\}$.

Alors la courbe \mathcal{C} est définie dans le repère \mathcal{R}_P par $\{(y_1(t), y_2(t)), t \in I\}$.

Démonstration. C'est une conséquence de la formule habituelle de changement de coordonnées : soit $M(t)$ le point de \mathcal{C} de paramètre t . Ce point a pour coordonnées $(x_1(t), x_2(t))$. Par hypothèse, on a

$$X(t) = PY(t) = (v_1 v_2) \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = y_1(t)v_1 + y_2(t)v_2.$$

On lit donc immédiatement sur la formule ci-dessus que les coordonnées de $M(t)$ dans le repère \mathcal{R}_P sont $y_1(t)$ et $y_2(t)$. \square

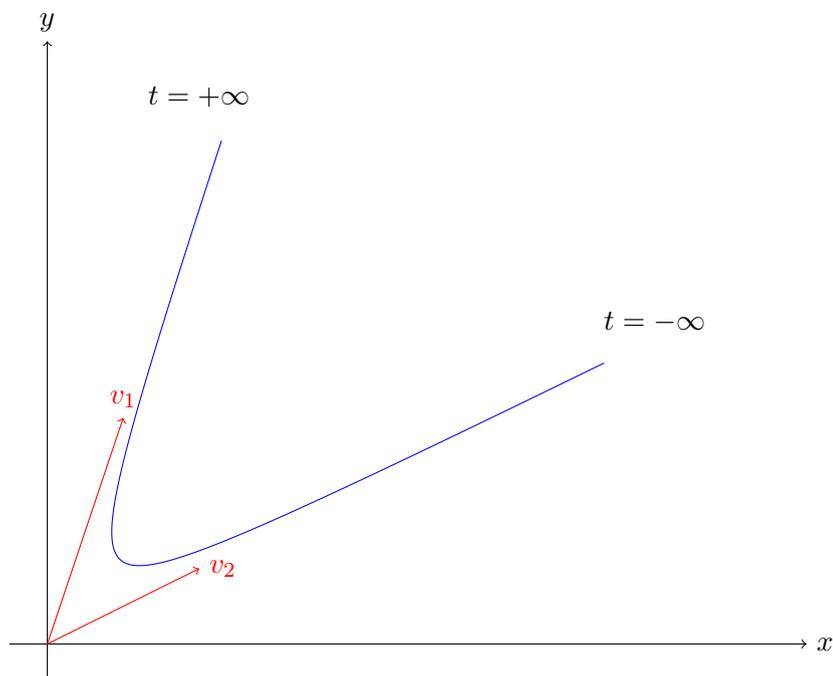
En vue du Lemme on obtient que la courbe \mathcal{C}_{A,X_0} est définie dans le repère \mathcal{R}_P par les points de coordonnées

$$e^{t\tilde{A}}(P^{-1}X_0), \quad t \in \mathbb{R}$$

Notons aussi que le point de coordonnées $P^{-1}X_0$ dans \mathcal{R}_P a pour coordonnées X_0 dans le repère (Oxy) .

On trace donc la courbe dans le repère \mathcal{R}_P en utilisant l'analyse faite au paragraphe précédent. Il faut cependant garder en tête qu'en général, le repère \mathcal{R}_P n'est pas ortho-normé, ni même orthogonal. La courbe a donc une allure déformée par ce changement de repère : la direction des abscisses (resp. ordonnées) devient la direction de v_1 , le premier vecteur colonne de la matrice de passage P (resp. la direction de v_2 , le deuxième vecteur colonne de la matrice de passage P). Par ailleurs nous remarquons que si $(1, 0)$ (resp. $(0, 1)$) est vecteur propre de \tilde{A} , alors v_1 (resp. v_2) est vecteur propre de A .

On trace ci-dessous un exemple de trajectoire de solution pour une matrice A diagonalisable avec valeurs propres réelles, lorsque $\lambda_2/\lambda_1 < 0$, $\lambda_1 > 0$:



Rapellons aussi que si $\lambda_1 < 0$ le sens de parcours de la courbe est inversé.

2.5 Que lit-on sur les portraits de phase ? Quelques notions de stabilité des équations différentielles

Les portraits de phase peuvent être utilisés pour comprendre les notions de **stabilité**. On s'intéresse à la partie des trajectoires des solutions pour laquelle $t \rightarrow +\infty$. De façon générale, les notions de stabilité s'intéressent au comportement d'un système au voisinage d'un point d'équilibre.

Définition 2.5.1. On appelle **point d'équilibre** d'une équation différentielle toute solution constante de cette équation différentielle.

Par la suite, on identifiera une solution constante avec la constante à laquelle est égale la solution en question.

Ici, comme le système est linéaire, les points d'équilibre X_0 sont caractérisés par la propriété $X_0 \in \ker A$. En effet, si X_0 est un point d'équilibre, on a $X(t) = X_0$ pour tout t , donc $AX_0 = AX(t) = X'(t) = 0$ pour tout t . Réciproquement, si $X_0 \in \ker A$, alors

$$e^{tA}X_0 = \left(\sum_{n \geq 0} \frac{A^n}{n!} \right) X_0 = X_0 + \sum_{n \geq 1} \frac{A^n X_0}{n!} = X_0,$$

si bien que X_0 est un point d'équilibre.

Ici $A \in \mathcal{M}_2(\mathbb{R})$, donc $\ker A$ est de dimension 0 (si 0 n'est pas valeur propre de A et dans ce cas $\ker A = \{0_{\mathbb{R}^2}\}$), 1 (si 0 est valeur propre simple de A , ou si 0 est valeur propre double de A mais A n'est pas diagonalisable), ou 2 (si 0 est valeur propre double de A et A est diagonalisable, donc si A est la matrice nulle).

On écarte le cas où A est la matrice nulle, qui n'a pas beaucoup d'intérêt (dans ce cas, les trajectoires sont toutes réduites à des points). On distingue au moins deux notions de stabilité :

— **Stabilité orbitale, ou stabilité au sens de Lyapunov :**

Définition 2.5.2. On dit que $X_0 \in \ker A$ est un *point d'équilibre stable au sens de Lyapunov* si la propriété suivante est vérifiée :

$$\forall \epsilon > 0, \exists \delta > 0, \forall Y_0 \in \mathbb{R}^2, \|Y_0 - X_0\| \leq \delta \Rightarrow \sup_{t \geq 0} \|e^{tA}Y_0 - X_0\| \leq \epsilon.$$

Autrement dit, un point d'équilibre est stable au sens de Lyapunov si la propriété suivante est vérifiée : si on part d'une position initiale proche de la position d'équilibre, alors on reste proche de cette position d'équilibre pour tout $t \geq 0$ (attention ici à la "direction" du temps : on ne repart pas en arrière, c'est-à-dire qu'on ne considère pas la portion de trajectoire qui correspond à $t < 0$).

— **Stabilité asymptotique :**

Définition 2.5.3. On dit que $X_0 \in \ker A$ est un *point d'équilibre asymptotiquement stable* si la propriété suivante est vérifiée :

$$\exists \epsilon_0 > 0, \forall Y_0 \in \mathbb{R}^2, \|Y_0 - X_0\| \leq \epsilon_0 \Rightarrow \lim_{t \rightarrow +\infty} \|e^{tA}Y_0 - X_0\| = 0.$$

Autrement dit, un point d'équilibre est asymptotiquement stable si la propriété suivante est vérifiée : si on part d'une position initiale suffisamment proche de la position d'équilibre, alors on converge en temps grand vers la position d'équilibre.

Interprétations graphiques, à comprendre sur le tableau page 46 :

- La stabilité orbitale signifie que les courbes du portrait de phase, si elles passent près de la position d'équilibre à un instant donné, restent ensuite toujours près de cette position d'équilibre. Intuitivement, cela veut dire qu'on ne s'éloigne pas trop de la position d'équilibre.
- La stabilité asymptotique signifie que toutes les courbes du portrait de phase qui passent près de la position d'équilibre convergent en temps grand vers cette position d'équilibre. Cela veut dire que l'on se rapproche de la position d'équilibre en temps grand.

Ainsi pour étudier la stabilité, il suffit d'examiner les trois cas décrits en "préliminaires", puisque le passage au cas général correspond simplement à une déformation des courbes par changement des axes de coordonnées.

Examinons à présent quelques cas décrits plus haut, en distinguant suivant la forme de la matrice A

1. Si $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, avec $\lambda_1, \lambda_2 \in \mathbb{R}$:

- Tout d'abord, si $\lambda_1 \neq 0$ et $\lambda_2 \neq 0$, seul $(0, 0)$ est point d'équilibre. On n'étudie donc la stabilité qu'autour du point $(0, 0)$. À partir des formules

$$\begin{aligned}x &= e^{\lambda_1 t} x_0, \\y &= e^{\lambda_2 t} y_0,\end{aligned}$$

on vérifie facilement que le point $(0, 0)$ est orbitalement ou asymptotiquement stable si et seulement si $\lambda_1 < 0, \lambda_2 < 0$ (par conséquent $\lambda_2/\lambda_1 > 0$). Sur le portrait de phase page 34, cela correspond aux courbes jaune, violette et rouge, parcourues dans le sens "infini vers zéro". La courbe bleue, qui correspond à $\lambda_2/\lambda_1 < 0$, est écartée : on vérifie bien que sur cette courbe, quel que soit le sens de parcours, on finit par s'éloigner infiniment de l'origine.

- Il reste à examiner le cas $\lambda_1 = 0, \lambda_2 \neq 0$ (le cas $\lambda_2 = 0$ se traite de façon similaire.) Dans ce cas les positions d'équilibre sont de la forme $(\alpha, 0)$, avec $\alpha \in \mathbb{R}$. Autrement dit l'ensemble des états d'équilibre est l'axe des abscisses tout entier.

D'un autre côté, les trajectoires du portrait de phase sont des demi-droites verticales (en rose sur le dessin page 34). Le signe de λ_2 détermine le sens de parcours de ces demi-droites. Si $\lambda_2 > 0$, on a

$$e^{\lambda_2 t} y_0 \rightarrow \operatorname{sgn}(y_0) \infty \quad \text{quand } t \rightarrow \infty,$$

et donc il n'y a pas de stabilité orbitale, ni de stabilité asymptotique.

On se concentre donc maintenant sur le cas $\lambda_2 < 0$. Dans ce cas les demi-droites sont parcourues dans le sens "haut vers bas" si $y_0 > 0$, "bas vers haut" si $y_0 < 0$. Autrement dit, si on part d'un point (x_0, y_0) , on converge vers la position d'équilibre $(x_0, 0)$.

On voit alors que les points d'équilibre sont tous orbitalement stables, mais pas asymptotiquement stables : en effet, si (x_0, y_0) est dans un voisinage de $(\alpha, 0)$ la trajectoire va rester dans un voisinage de $(\alpha, 0)$. En revanche, si $x_0 \neq \alpha$, aussi près soit-il de α , la trajectoire ne converge jamais vers $(\alpha, 0)$, et on n'a donc pas de stabilité asymptotique.

2. Si $A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$, avec $\alpha, \beta \in \mathbb{R}, \beta \neq 0$:

Seul $(0, 0)$ est un point d'équilibre.

Si $\alpha = 0$, les trajectoires sont des cercles centrés en 0. On a donc stabilité orbitale, mais pas stabilité asymptotique.

Si $\alpha \neq 0$, les trajectoires sont des spirales qui s'entourent autour de l'origine, et le signe de α détermine le sens de parcours de la spirale. On retrouve une fois de plus qu'on a stabilité orbitale et asymptotique si $\alpha < 0$, et ni stabilité orbitale ni stabilité asymptotique si $\alpha > 0$.

3. Si $A = \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix}$, avec $\lambda, \gamma \in \mathbb{R}, \gamma \neq 0$:

Si $\lambda \neq 0$, le portrait de phase a l'allure dessinée page 37. Le signe de λ détermine le sens de parcours de la courbe. On voit immédiatement qu'on a stabilité asymptotique (resp. stabilité orbitale) si et seulement si $\lambda < 0$.

Examinons maintenant le cas $\lambda = 0$. Dans ce cas les états d'équilibre sont de la forme $(r, 0)$ avec $r \in \mathbb{R}$. Par ailleurs

$$e^{tA} X_0 = X_0 + \gamma t \begin{pmatrix} y_0 \\ 0 \end{pmatrix}.$$

Ainsi si $y_0 \neq 0$, $\|e^{tA}X_0\| \rightarrow \infty$ quand $t \rightarrow \infty$. On ne peut donc pas avoir de stabilité asymptotique ni de stabilité orbitale.

On résume ces propriétés de stabilité des points d'équilibre, qui se transmettent aussi aux matrices semblables, dans le tableau suivant :

A diagonalisable, avec valeurs propres distinctes ou confondues $\lambda_1, \lambda_2 \in \mathbb{R}$	$\lambda_2 < 0$	$\lambda_1 < 0$ Stabilité orbitale et asymptotique	$\lambda_1 = 0$	$\lambda_1 > 0$
	$\lambda_2 = 0$	Stab. orbitale mais pas asymptotique		
	$\lambda_2 > 0$	Pas de stabilité orbitale, ni asymptotique		
A diagonalisable, avec valeurs propres $\alpha \pm i\beta$, $\beta \neq 0$		$\alpha < 0$ Stabilité orbitale et asymptotique	$\alpha = 0$ Stabilité orbitale, mais pas asymptotique	$\alpha > 0$ Pas de stabilité orbitale, ni asymptotique
A non diagonalisable, donc avec valeur propre double $\lambda \in \mathbb{R}$		$\lambda < 0$ Stabilité orbitale et asymptotique	$\lambda = 0$ Pas de stab. orbitale, ni asymptotique	$\lambda > 0$

Notons que si toutes les valeurs propres de A ont des parties réelles strictement négatives, alors les point d'équilibre sont asymptotiquement et orbitalement stables. Aussi, si l'une des valeurs propres de A a une partie réelle strictement positive, alors les points d'équilibre ne sont pas stables (ni asymptotiquement, ni orbitalement). Enfin, dans tous les cas de stabilité asymptotique on a aussi la stabilité orbitale. En général pour des équations plus compliquées la stabilité asymptotique n'implique automatiquement qu'une notion de stabilité orbitale plus faible que celle ci-dessus.

Remarque 2.5.4. Cette méthode est souvent utilisée en pratique pour déterminer la stabilité d'un état d'équilibre, même en dehors du cadre linéaire. En effet, supposons que X soit solution de l'équation autonome

$$X'(t) = b(X(t)),$$

et supposons qu'il existe $X_{eq} \in \mathbb{R}^2$ tel que $b(X_{eq}) = 0$. On part d'une position initiale dans un voisinage de X_{eq} . Alors tant que $X(t)$ est proche de X_{eq} , on a

$$b(X(t)) \approx Db(X_{eq})(X(t) - X_{eq}),$$

où Db est la matrice jacobienne

$$\begin{pmatrix} \frac{\partial b_1}{\partial x_1} & \frac{\partial b_1}{\partial x_2} \\ \frac{\partial b_2}{\partial x_1} & \frac{\partial b_2}{\partial x_2} \end{pmatrix}.$$

En notant $A = Db(X_{eq})$, $Y(t) = X(t) - X_{eq}$, on en déduit que tant que $X(t)$ est dans un voisinage de X_{eq} , et donc tant que $Y(t)$ est petit, l'évolution est gouvernée par l'équation

$$Y'(t) = AY(t).$$

En examinant les valeurs propres de la matrice A , on peut savoir si on a un espoir que le point d'équilibre X_{eq} soit stable. On peut démontrer (mais c'est difficile!) que si toutes les

valeurs propres de A ont des parties réelles strictement négatives, alors X_{eq} est un point d'équilibre asymptotiquement stable pour le système non linéaire de départ. Si l'une des valeurs propres de A a une partie réelle strictement positive, alors le point d'équilibre n'est pas stable (ni asymptotiquement, ni orbitalement). Dans les autres cas (valeurs propres avec parties réelles négatives ou nulles, mais pas toutes strictement négatives), on ne peut pas conclure en général.

Le raisonnement mené ci-dessus s'appelle une **linéarisation** de l'équation près du point d'équilibre. Il peut se généraliser à des cas plus complexes (étude d'équations aux dérivées partielles, qui peuvent s'apparenter à des systèmes d'équations différentielles en dimension infinie...). Le résultat cité précédemment, qui fait le lien entre les valeurs propres de $Df(X_{eq})$ et la stabilité du point d'équilibre correspondant, justifie *a posteriori* que l'on étudie les systèmes linéaires à coefficients constants de façon aussi approfondie.

Chapitre 3

Théorie générale pour les équations différentielles

3.1 Théorème de Cauchy-Lipschitz

On commence par un théorème d'existence et d'unicité assez puissant pour les équations différentielles d'ordre un en dimension quelconque (toutes les EDO d'ordre N peuvent se ramener à une EDO d'ordre 1 en prenant comme inconnue $(x, x', \dots, x^{(N-1)})$).

Pour préciser les idées, on définit tout d'abord ce qu'est la solution d'une équation différentielle munie d'une donnée initiale.

Définition 3.1.1. • Soit I un intervalle ouvert de \mathbb{R} , U un ouvert de \mathbb{R}^m , et soit $t_0 \in I$, $X_0 \in U$. Soit $f : I \times U \rightarrow \mathbb{R}^m$.

On considère l'équation différentielle avec donnée initiale

$$\begin{cases} X'(t) = f(t, X(t)) & \forall t \in I, \\ X(t_0) = X_0. \end{cases} \quad (3.1)$$

Une **solution** de (3.1) est une fonction X définie sur un intervalle I_X contenant t_0 , à valeurs dans $U \subset \mathbb{R}^m$, vérifiant l'équation sur I_X et la condition initiale.

• Le système (3.1) est aussi appelé **problème de Cauchy** (équation différentielle ET donnée initiale). Résoudre le problème de Cauchy consiste à trouver l'ensemble des solutions de (3.1) ainsi que l'intervalle sur lequel elles sont définies.

Exemple 3.1.2. — Une EDO scalaire :

$$\begin{cases} x'(t) = t + x^2(t), \\ x(0) = 2. \end{cases}$$

Dans le cas de ce problème de Cauchy $m = 1$ et $f(t, y) = t + y^2$. Si on considère comme équation $x'(t) = x^2(t) = f(t, x(t))$ alors il s'agit d'une équation **autonome** dans le sens que $f(t, y) = y^2$ ne dépend pas de t séparément et on peut écrire directement $x'(t) = f(x(t))$ avec $f(y) = y^2$.

— Système de Lotka-Volterra : les équations de Lotka-Volterra s'écrivent

$$\begin{cases} N'(t) = aN(t) - bN(t)P(t), \\ P'(t) = cN(t)P(t) - dP(t). \end{cases}$$

Dans ce cas, $m = 2$, $X = \begin{pmatrix} N \\ P \end{pmatrix}$, et $X'(t) = f(X(t))$ avec

$$f : \begin{pmatrix} \mathbb{R}^2 \\ x \\ y \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbb{R}^2 \\ ax - bxy \\ cxy - dy \end{pmatrix} .$$

C'est un système autonome : le champ de vecteurs f ne dépend pas de t .

— Pendule simple : on rappelle que l'équation régissant l'évolution du pendule simple s'écrit

$$\theta''(t) + \omega_0^2 \sin \theta(t) = 0,$$

que l'on munit des conditions initiales

$$\theta(0) = \theta_0, \quad \theta'(0) = v_0.$$

On met ce système sous forme vectorielle en posant $X = \begin{pmatrix} \theta \\ \theta' \end{pmatrix}$. On vérifie alors que l'équation du pendule simple peut se mettre sous la forme (3.1) en posant $X'(t) = f(X(t))$ avec

$$f : \begin{pmatrix} \mathbb{R}^2 \\ x \\ y \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbb{R}^2 \\ y \\ -\omega_0^2 \sin x \end{pmatrix},$$

et $X_0 = \begin{pmatrix} \theta_0 \\ v_0 \end{pmatrix}$. Il s'agit de nouveau d'un système autonome. Cet exemple montre comment on peut ramener l'étude d'une EDO scalaire d'ordre 2 à celui d'une EDO vectorielle d'ordre 1.

Remarque 3.1.3. Dans la définition ci-dessus, on ne suppose pas que t_0 appartient à l'intérieur de I_X (ni que l'intervalle I_X est ouvert) : ça peut être une extrémité de l'intervalle.

Proposition 3.1.4 (Forme intégrale des solutions). *Une fonction continue $X : I_X \rightarrow U$ est solution du problème de Cauchy (3.1) si et seulement si*

$$X(t) = X_0 + \int_{t_0}^t f(u, X(u)) du. \quad (3.2)$$

Démonstration. Si $X : I_X \rightarrow U$ est solution du problème de Cauchy (3.1) alors

$$Y(t) := X_0 + \int_{t_0}^t f(u, X(u)) du$$

satisfait, par le théorème fondamental de l'analyse,

$$Y'(t) = f(t, X(t)) = X'(t).$$

Dans la dérivation du terme intégral on a utilisé le fait que $f(u, X(u))$ est une fonction continue, grâce à l'hypothèse de continuité sur f et X . On en déduit que $Y(t) - X(t)$ est une constante. D'autre part $Y(t_0) = X_0 = X(t_0)$, donc Y et X coïncident sur I_X .

Réciproquement, si X satisfait (3.2) alors $X'(t) = f(t, X(t))$ et aussi $X(t_0) = X_0$, donc X est bien solution du problème de Cauchy (3.1). \square

Exemple 3.1.5. Considérons le problème de Cauchy suivant

$$\begin{cases} x'(t) = tx(t) & \forall t \in \mathbb{R}, \\ x(0) = 1. \end{cases}$$

Nous savons déjà trouver la solution explicitement, il s'agit de $e^{\frac{t^2}{2}}$. Nous allons la retrouver d'une autre façon, en ayant en tête la proposition précédente sur la forme intégrale des solutions.

Pour toute fonction continue $x : \mathbb{R} \rightarrow \mathbb{R}$ définissons une nouvelle fonction $Tx : \mathbb{R} \rightarrow \mathbb{R}$ par

$$(Tx)(t) := 1 + \int_0^t ux(u)du.$$

Posons x^0 la fonction constante donnée par la donnée initiale : $x^0(t) = 1$ pour tout $t \in \mathbb{R}$. Ensuite, pour tout $n \in \mathbb{N}^*$, on définit la fonction $x^n := Tx^{n-1}$, c'est-à-dire

$$x^n(t) = 1 + \int_0^t ux^{n-1}(u)du.$$

Nous pouvons calculer explicitement :

$$x^1(t) = 1 + \int_0^t udu = 1 + \frac{t^2}{2},$$

$$x^2(t) = 1 + \int_0^t u\left(1 + \frac{u^2}{2}\right)du = 1 + \frac{t^2}{2} + \frac{1}{2} \left(\frac{t^2}{2}\right)^2,$$

et par récurrence on trouve pour tout $n \in \mathbb{N}$

$$x^n(t) = 1 + \int_0^t u \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{u^2}{2}\right)^k du = \sum_{k=0}^n \frac{1}{k!} \left(\frac{t^2}{2}\right)^k.$$

Il s'ensuit donc que la suite de fonctions x^n vérifie

$$\lim_{n \rightarrow \infty} x^n(t) = e^{\frac{t^2}{2}},$$

donc on a retrouvé la solution du problème de Cauchy par itération de la fonction constante donnée par la donnée initiale à travers T . C'est cette approche que nous allons utiliser pour démontrer le résultat suivant dans le cas d'équations pas nécessairement résolubles explicitement.

Théorème 3.1.6 (Théorème de Cauchy-Lipschitz¹ – existence locale). *Soit I un intervalle ouvert de \mathbb{R} , U un ouvert de \mathbb{R}^m , et soit $t_0 \in I$, $X_0 \in U$. Soit $f : I \times U \rightarrow \mathbb{R}^m$ une fonction \mathcal{C}^1 (c'est-à-dire que les dérivées partielles de f par rapport à toutes les variables existent et sont continues.)*

Alors il existe $\tau > 0$ tel que le problème de Cauchy (3.1) admet une solution X de classe \mathcal{C}^1 sur l'intervalle $[t_0 - \tau, t_0 + \tau]$.

On a également la propriété d'unicité suivante : si Y est une autre solution de (3.1) définie sur un intervalle I_Y , X et Y coïncident sur $I_Y \cap [t_0 - \tau, t_0 + \tau]$. Plus généralement, si Y et \tilde{Y} sont deux solutions de (3.1) définies sur les intervalles I_Y et $I_{\tilde{Y}}$ respectivement, alors elles coïncident sur $I_Y \cap I_{\tilde{Y}}$.

1. Les théorèmes de Cauchy-Lipschitz sont aussi connus sous le nom de Picard-Lindelöf; ils datent de 1820-1900.

Remarque 3.1.7. Ce résultat assure que certaines équations différentielles (système de Lotka-Volterra, pendule simple...) admettent des solutions pour toute donnée initiale, au moins localement en temps (c'est-à-dire sur un petit intervalle autour de t_0). On verra plus tard des conditions pour que les solutions d'équations différentielles soient globales, c'est-à-dire définies pour tout temps.

Remarque 3.1.8. La condition $f \in \mathcal{C}^1$ peut être légèrement relaxée : il suffit que la fonction f soit continue et lipschitzienne sur tout compact en la 2ème variable. Mais en général, on peut pas faire "mieux" (dans le sens : "on ne peut pas supposer moins") que le caractère Lipschitz, comme le montre l'exemple suivant.

On considère le problème de Cauchy suivant :

$$\begin{cases} x'(t) = 2\sqrt{x(t)}, & t \in \mathbb{R}, \\ x(0) = 0. \end{cases} \quad (3.3)$$

On vérifie facilement que la fonction constante égale à zéro est solution de l'équation. On va montrer qu'il existe une infinité d'autres solutions : pour $T > 0$, on pose

$$x_T(t) = \begin{cases} 0 & \text{si } t < T, \\ (t - T)^2 & \text{si } t \geq T. \end{cases}$$

On vérifie facilement que x_T est également solution de l'équation différentielle (3.3) pour tout $T > 0$. Il n'y a donc pas unicité. Ce défaut d'unicité vient de ce que la fonction $x \mapsto \sqrt{x}$ n'est pas lipschitzienne au voisinage de zéro.

Démonstration. Comme expliqué dans l'Exemple 3.1.5, nous allons chercher la solution du problème de Cauchy par itérations de la fonction constante donnée par la donnée initiale.

Construction par itération d'une suite de fonctions :

Posons X^0 la fonction constante donnée par la donnée initiale : $X^0(t) = X_0$ pour tout $t \in I$. Ensuite, pour tout $n \in \mathbb{N}^*$, on définit par itérations successives la suite de fonctions

$$X^n(t) = X_0 + \int_{t_0}^t f(u, X^{n-1}(u)) du,$$

sur un intervalle de temps $[t_0 - \tau, t_0 + \tau]$ qui se prête à ce type d'itérations, et qu'on définit comme suit.

Choix de l'intervalle de définition :

Comme I est un intervalle ouvert de \mathbb{R} , U est un ouvert de \mathbb{R}^m , et $(t_0, X_0) \in I \times U$, il existe des nombres positifs a et b tels que $[t_0 - a, t_0 + a] \times \overline{B}(X_0, b) \subset I \times U$. Par exemple lorsque $m = 1$ ce cylindre fermé de \mathbb{R}^{m+1} se réduit au pavé $[t_0 - a, t_0 + a] \times [X_0 - b, X_0 + b]$ du plan \mathbb{R}^2 . S'agissant d'un ensemble compact et f étant continue, son image est aussi compacte. Soit $M > 0$ un majorant de $\|f\|$ sur $[t_0 - a, t_0 + a] \times \overline{B}(X_0, b)$. Aussi, comme f est de classe \mathcal{C}^1 , elle est différentiable et nous pouvons appliquer le théorème des accroissements finis pour obtenir une constante $K > 0$ telle que²

$$\|f(t, X) - f(t, Y)\| \leq K \|X - Y\|, \quad \forall (t, X), (t, Y) \in [t_0 - a, t_0 + a] \times \overline{B}(X_0, b). \quad (3.4)$$

Nous définissons alors

$$\tau = \min \left\{ a, \frac{b}{M}, \frac{1}{2K} \right\}.$$

2. Nous voyons ici qu'il suffit d'avoir f continue, localement Lipschitz en la 2ème variable.

Estimations sur la suite de fonctions :

D'abord nous allons démontrer par récurrence que pour tout $n \in \mathbb{N}$ nous avons

$$X^n(t) \in \overline{B}(X_0, b), \quad \forall t \in [t_0 - \tau, t_0 + \tau].$$

Ceci est vrai pour $n = 0$, et en le supposant pour n on a

$$\|X^{n+1}(t) - X_0\| = \left\| \int_{t_0}^t f(u, X^n(u)) du \right\| \leq M|t - t_0| \leq M\tau \leq b.$$

Ensuite nous avons tout $n \in \mathbb{N}^*$

$$\begin{aligned} \sup_{t \in [t_0 - \tau, t_0 + \tau]} \|X^{n+1}(t) - X^n(t)\| &= \left\| \int_{t_0}^t (f(u, X^n(u)) - f(u, X^{n-1}(u))) du \right\| \\ &\leq \int_{t_0}^t K \|X^n(t) - X^{n-1}(t)\| du \\ &\leq K\tau \sup_{t \in [t_0 - \tau, t_0 + \tau]} \|X^n(t) - X^{n-1}(t)\| \\ &\leq \frac{1}{2} \sup_{t \in [t_0 - \tau, t_0 + \tau]} \|X^n(t) - X^{n-1}(t)\|. \end{aligned}$$

En itérant cette estimations nous avons obtenu que pour tout $n \in \mathbb{N}^*$

$$\sup_{t \in [t_0 - \tau, t_0 + \tau]} \|X^{n+1}(t) - X^n(t)\| \leq \frac{1}{2^n} \sup_{t \in [t_0 - \tau, t_0 + \tau]} \|X^1(t) - X^0(t)\| \leq \frac{b}{2^n}. \quad (3.5)$$

Définition de la fonction limite :

De la majoration ci-dessus il s'ensuit que $\sum (X^n - X^{n-1})$ est une série de fonctions continues normalement convergente. En posant

$$X = X^0 + \sum_{n=1}^{\infty} (X^n - X^{n-1}),$$

on obtient donc une fonction $X : [t_0 - \tau, t_0 + \tau] \rightarrow \mathbb{R}^m$ continue, telle que

$$X(t) = \lim_{N \rightarrow \infty} (X^0(t) + \sum_{n=1}^N (X^n(t) - X^{n-1}(t))) = \lim_{N \rightarrow \infty} X^N(t).$$

Par ailleurs, il s'ensuit que

$$X(t) \in \overline{B}(X_0, b) \subset U, \quad \forall t \in [t_0 - \tau, t_0 + \tau].$$

Vérification du problème de Cauchy :

Pour montrer que X est bien solution du problème de Cauchy (3.1) nous remarquons que pour tout $t \in [t_0 - \tau, t_0 + \tau]$,

$$\begin{aligned} X(t) &= \lim_{N \rightarrow \infty} X^N(t) = X_0 + \lim_{N \rightarrow \infty} \int_{t_0}^t f(u, X^{N-1}(u)) du \\ &= X_0 + \int_{t_0}^t f(u, X(u)) du + \lim_{N \rightarrow \infty} \int_{t_0}^t (f(u, X^{N-1}(u)) - f(u, X(u))) du. \end{aligned}$$

Il nous reste donc à montrer que le dernier terme vaut zéro. Comme $X(u), X^{N-1}(u) \in \overline{B}(X_0, b) \subset U, \forall u \in [t_0 - \tau, t_0 + \tau]$ nous avons pour tout $t \in [t_0 - \tau, t_0 + \tau]$, par (3.5),

$$\begin{aligned} \left\| \int_{t_0}^t (f(u, X^{N-1}(u)) - f(u, X(u))) du \right\| &\leq \int_{t_0}^t K \|X^{N-1}(u) - X(u)\| du \\ &\leq K\tau \sup_{u \in [t_0 - \tau, t_0 + \tau]} \|X^{N-1}(u) - X(u)\| \\ &= K\tau \sup_{u \in [t_0 - \tau, t_0 + \tau]} \left\| \sum_{n \geq N} (X^{n+1}(u) - X^n(u)) \right\| \\ &\leq K\tau \sum_{n \geq N} \frac{b}{2^n} \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Unicité :

Nous allons démontrer directement le résultat d'unicité pour Y et \tilde{Y} deux solutions de (3.1) définies sur les intervalles I_Y et $I_{\tilde{Y}}$ respectivement. Si $I_Y \cap I_{\tilde{Y}} = \{t_0\}$ la conclusion découle du fait que Y et \tilde{Y} ont la même donnée initiale. On se place donc dans le cas où $I_Y \cap I_{\tilde{Y}}$ contient un intervalle centré en t_0 (le cas où $I_Y \cap I_{\tilde{Y}}$ contient un intervalle de type $[t_0, \beta[$ ou $]\alpha, t_0]$ se traite pareil).

Nous commençons par montrer qu'elles coïncident localement autour de t_0 . Pour $t \in I_Y \cap I_{\tilde{Y}}$ nous avons, en utilisant la forme intégrale des solutions,

$$Y(t) - \tilde{Y}(t) = \int_{t_0}^t (f(u, Y(u)) - f(u, \tilde{Y}(u))) du.$$

Les fonctions Y et \tilde{Y} étant continues et prenant la même valeur X_0 en t_0 , il existe $0 < \delta < \tau$ tel que $[t_0 - \delta, t_0 + \delta] \subset I_Y \cap I_{\tilde{Y}}$ et tel que

$$Y(u), \tilde{Y}(u) \in \overline{B}(X_0, b), \quad \forall u \in [t_0 - \delta, t_0 + \delta]$$

Il s'ensuit comme avant que pour tout $t \in [t_0 - \delta, t_0 + \delta]$

$$\|Y(t) - \tilde{Y}(t)\| \leq \int_{t_0}^t K \|Y(u) - \tilde{Y}(u)\| du,$$

et en particulier

$$\sup_{t \in [t_0 - \delta, t_0 + \delta]} \|Y(t) - \tilde{Y}(t)\| \leq K\delta \sup_{t \in [t_0 - \delta, t_0 + \delta]} \|Y(t) - \tilde{Y}(t)\| \leq \frac{1}{2} \sup_{t \in [t_0 - \delta, t_0 + \delta]} \|Y(t) - \tilde{Y}(t)\|.$$

On en déduit que $\sup_{t \in [t_0 - \delta, t_0 + \delta]} \|Y(t) - \tilde{Y}(t)\| = 0$, et donc que Y et \tilde{Y} coïncident sur $[t_0 - \delta, t_0 + \delta]$.

Nous allons maintenant démontrer que Y et \tilde{Y} coïncident sur $t \in I_Y \cap I_{\tilde{Y}} \cap \{t \geq t_0\}$ (montrer qu'ils coïncident sur $I_Y \cap I_{\tilde{Y}} \cap \{t \leq t_0\}$ se fait de façon similaire). Comme I_Y et $I_{\tilde{Y}}$ sont des intervalles, leur intersection l'est aussi. Notons t_1 l'extrémité supérieure de $I_Y \cap I_{\tilde{Y}}$. On considère l'ensemble

$$A = \{t \in I_Y \cap I_{\tilde{Y}} \mid t \geq t_0, Y(u) = \tilde{Y}(u) \forall u \in [t_0, t]\},$$

qui n'est pas vide car il contient t_0 , et c'est un ensemble majoré par t_1 , donc $t_+ = \sup A$ est un élément de \mathbb{R} , et $t_+ \leq t_1$. Si $t_+ = t_1$ cela implique que l'on a bien $Y = \tilde{Y}$ sur $t \in I_Y \cap I_{\tilde{Y}} \cap \{t \geq t_0\}$. Supposons maintenant par l'absurde que $t_+ < t_1$. D'après ce qui

précède sur l'unicité locale, on obtient aussi $t_0 < t_0 + \delta \leq t_+$. Pour t dans l'intervalle non-vide $[t_0, t_+[$ on a $Y(t) = \tilde{Y}(t)$, donc en passant à la limite quand $t \rightarrow t_+$, on obtient par continuité $Y(t_+) = \tilde{Y}(t_+)$. On pose $Y_+ = Y(t_+)$, et on considère le problème de Cauchy au point t_+ , avec la donnée initiale Y_+ , dont Y et \tilde{Y} sont solutions. Le résultat d'unicité locale démontré ci-dessus nous assure que Y et \tilde{Y} coïncident sur un petit intervalle $[t_+ - \delta_+, t_+ + \delta_+] \cap I_Y \cap I_{\tilde{Y}}$ avec $0 < \delta_+ < t_1 - t_+$, donc sur $[t_+ - \delta_+, t_+ + \delta_+] \subset I_Y \cap I_{\tilde{Y}}$. Nous déduisons que $Y(t) = \tilde{Y}(t)$ pour tout $t \in [t_0, t_+ + \delta_+]$, donc $t_+ + \delta_+ \in A$. Par conséquent,

$$t_+ + \delta_+ \leq \sup A = t_+,$$

ce qui est une contradiction car $\delta_+ > 0$, donc on ne peut pas avoir $t_+ < t_1$. \square

Nous remarquons que la partie unicité du théorème donne un résultat de non-croisement de solutions. Nous verrons plus tard (Proposition 3.3.2) cette propriété précisée dans le cas des EDO scalaires.

3.2 Solutions maximales

3.2.1 Existence et unicité

On va à présent définir la notion de solution maximale d'une équation différentielle. Pour cela, on introduit une relation d'ordre sur les solutions de l'équation différentielle (3.1) :

Définition 3.2.1. Soit X, Y deux solutions de (3.1), définies respectivement sur des intervalles I et J contenant t_0 . On dit que Y **est une extension de** X et l'on note $X \prec Y$ si et seulement si $I \subset J$ et $X(t) = Y(t)$ pour tout $t \in I$.

On dit que X_{max} est une **solution maximale** de (3.1) si c'est une solution de (3.1) et si pour toute solution X de (3.1),

$$X_{max} \prec X \Rightarrow X = X_{max}$$

Remarque 3.2.2. On vérifie facilement que \prec est une relation d'ordre (pas total) sur les solutions de (3.1), et que les solutions maximales sont précisément les éléments maximaux pour cette relation d'ordre.

Théorème 3.2.3 (Théorème de Cauchy-Lipschitz – solutions maximales). *Soit I un intervalle ouvert de \mathbb{R} , U un ouvert de \mathbb{R}^m , et soit $t_0 \in I$, $X_0 \in U$. Soit $f : I \times U \rightarrow \mathbb{R}^m$ une fonction \mathcal{C}^1 .*

Alors il existe une solution maximale X_{max} de l'équation (3.1), définie sur un intervalle ouvert $I_{max} \subset I$. Cette solution maximale est l'élément maximum (plus grand élément) pour la relation d'ordre \prec . Par conséquent, elle est unique.

Démonstration. On note \mathcal{S} l'ensemble des solutions de (3.1). D'après le théorème de Cauchy-Lipschitz, l'ensemble \mathcal{S} est non vide. On définit alors l'ensemble

$$I_{max} = \bigcup_{X \in \mathcal{S}} I_X,$$

où I_X est l'intervalle de définition d'une solution X de (3.1). On va montrer que I_{max} est un intervalle, et que l'on peut définir une solution X_{max} de (3.1) sur I_{max} . Il restera alors à démontrer que cette solution est une solution maximale.

Première étape : I_{max} est un intervalle. Il suffit de montrer que pour tout $t_1, t_2 \in I_{max}$, on a $[t_1, t_2] \subset I_{max}$ ³. Soit t_1, t_2 quelconques dans I_{max} . Il existe X_1, X_2 solutions de (3.1) tels que $t_1 \in I_{X_1}, t_2 \in I_{X_2}$. Or I_{X_1}, I_{X_2} sont des intervalles, et $t_0 \in I_X$ pour tout $X \in \mathcal{S}$. On en déduit que

$$[t_0, t_1] \subset I_{X_1}, \quad [t_0, t_2] \subset I_{X_2},$$

et par conséquent

$$[t_0, t_1] \cup [t_0, t_2] \subset I_{X_1} \cup I_{X_2} \subset I_{max}.$$

Or $[t_1, t_2] \subset [t_0, t_1] \cup [t_0, t_2]$. (Cette propriété se montre facilement en distinguant le cas où $t_0 \in [t_1, t_2]$ et le cas où $t_0 \notin [t_1, t_2]$.) Donc $[t_1, t_2] \subset I_{max}$.

Deuxième étape : Définition de X_{max} . On définit donc X_{max} de la façon suivante : soit $t \in I_{max}$ quelconque. Soit X une solution de (3.1) telle que $t \in I_X$. On pose

$$X_{max}(t) = X(t).$$

La définition de X_{max} ne dépend pas du choix de la solution X telle que $t \in I_X$. Effectivement, si $Y \in \mathcal{S}$ telle que $t \in I_Y$, alors d'après l'unicité du théorème de Cauchy-Lipschitz

$$X(u) = Y(u) \quad \forall u \in I_X \cap I_Y,$$

et en particulier $X(t) = Y(t)$. Donc X_{max} est définie de façon univoque.

Troisième étape : I_{max} est un intervalle ouvert et X_{max} est solution de (3.1). On va montrer que pour tout $\tilde{t} \in I_{max}$ quelconque il existe un voisinage de \tilde{t} inclut dans I_{max} sur lequel X_{max} vérifie l'équation. Comme de plus on a $X_{max}(t_0) = x_0$ par définition de X_{max} (toujours en utilisant le théorème de Cauchy-Lipschitz), on en déduit aussi que X_{max} est bien solution de (3.1).

Si $\tilde{t} = t_0$ on a la propriété ci-dessus grâce au théorème de Cauchy-Lipschitz.

Si $\tilde{t} \neq t_0$ on note $X_{\tilde{t}} = X_{max}(\tilde{t})$. Par définition de I_{max}, X_{max} , il existe une solution X de (3.1) telle que $\tilde{t} \in I_X$ et $X = X_{max}$ sur I_X . En particulier $[t_0, \tilde{t}] \subset I_X$ et $X(\tilde{t}) = X_{\tilde{t}}$.

D'autre part, d'après le théorème de Cauchy-Lipschitz (3.1.6), il existe $\tilde{\tau} > 0$ tel que le problème de Cauchy

$$X'(t) = f(t, X(t)), X(\tilde{t}) = X_{\tilde{t}}$$

possède une unique solution \tilde{X}_{CL} sur $[\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]$. Ainsi, on a nécessairement $X = \tilde{X}_{CL}$ sur $I_X \cap [\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]$, et en particulier sur $[t_0, \tilde{t}] \cap [\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]$. On définit à présent, sur $[t_0, \tilde{t}] \cup [\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]$,

$$\tilde{X}(t) = \begin{cases} X(t) & \text{si } t \in [t_0, \tilde{t}], \\ \tilde{X}_{CL}(t) & \text{si } t \in [\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]. \end{cases}$$

On vérifie aisément, à l'aide de ce qui précède, que la fonction \tilde{X} est bien définie et qu'elle est solution de (3.1). Donc $\tilde{X} \in \mathcal{S}$. En particulier,

$$[\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}] \subset [t_0, \tilde{t}] \cup [\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}] \subset I_{max}. \quad (3.6)$$

Aussi, on obtient $X_{max}(t) = \tilde{X}(t)$ pour tout $t \in [\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]$, et donc X_{max} est solution de l'équation

$$X'(t) = f(t, X(t))$$

sur l'intervalle $[\tilde{t} - \tilde{\tau}, \tilde{t} + \tilde{\tau}]$. On a donc bien obtenu l'existence d'un voisinage de \tilde{t} inclut dans I_{max} sur lequel X_{max} vérifie l'équation.

3. Les intervalles sont ici définis au sens large : si $a \geq b$, l'intervalle $[a, b]$ est égal à l'intervalle $[b, a]$.

Quatrième étape : X_{max} est une solution maximale, et c'est l'unique solution maximale. On observe que si $X \in \mathcal{S}$, alors $I_X \subset I_{max}$ par définition de I_{max} , et $X = X_{max}$ sur I_X par définition de X_{max} . Par conséquent, $X \prec X_{max}$ pour tout $X \in \mathcal{S}$. Donc X_{max} est le plus grand élément de \mathcal{S} pour la relation d'ordre \prec . Si $Y \in \mathcal{S}$ est aussi une solution maximale, alors puisque c'est une solution on a $Y \prec X_{max}$, et puisqu'elle est maximale, par définition on obtient $Y = X_{max}$. Donc X_{max} est l'unique solution maximale. \square

3.2.2 Dichotomie existence globale/explosion

On s'intéresse à présent à des solutions maximales uniquement, et à leur intervalle d'existence. On va montrer qu'il y a deux cas de figure possibles : soit les solutions sont *globales*, c'est-à-dire qu'elles sont définies sur \mathbb{R} tout entier, soit elles explosent quand elles approchent du bord de leur intervalle de définition.

On commence pour cela par un lemme très utile :

Lemme 3.2.4. *On suppose que les hypothèses du Théorème 3.1.6 sont vérifiées avec $I = \mathbb{R}$, $U = \mathbb{R}^m$, et on considère une solution maximale X_{max} de (3.1), définie sur un intervalle $I_{max} =]a, b[$. On suppose que X_{max} est bornée sur un voisinage à gauche de b . Alors $b = +\infty$.*

Démonstration. Raisonnons par l'absurde : supposons que $b < \infty$. Nous allons montrer dans un premier temps que X_{max} admet une limite à gauche en b .

Soit V un voisinage de b sur lequel X_{max} est bornée. Il existe donc un intervalle compact \bar{V} contenant V . Notons aussi $R = \sup_{t \in \bar{V}} \|X_{max}(t)\| < \infty$. On pose alors

$$C_0 = \sup_{t \in \bar{V}, \|Z\| \leq R} \|f(t, Z)\|.$$

Comme f est de classe \mathcal{C}^1 et \bar{V} est compact, il s'ensuit que C_0 est fini. Puisque X_{max} est solution de (3.1), on en déduit que

$$\|X'_{max}(t)\| \leq C_0 \quad \forall t \in V.$$

D'après l'inégalité des accroissements finis, pour tout $t, u \in V$, on a

$$\|X_{max}(t) - X_{max}(u)\| \leq C_0 |t - u|.$$

On en déduit que si $(t_n)_{n \in \mathbb{N}}$ est une suite quelconque de $]a, b[$ tendant vers b , alors la suite $(X_{max}(t_n))_{n \in \mathbb{N}}$ est de Cauchy dans \mathbb{R}^m , donc elle est convergente. De surcroît cette limite ne dépend pas de la suite $(t_n)_{n \in \mathbb{N}}$ choisie : en effet, si $(\tilde{t}_n)_{n \in \mathbb{N}}$ est une autre suite de $]a, b[$ tendant vers b , on pose, pour $n \in \mathbb{N}$,

$$\tau_n = \begin{cases} t_k & \text{si } n = 2k, k \in \mathbb{N}, \\ \tilde{t}_k & \text{si } n = 2k + 1, k \in \mathbb{N}. \end{cases}$$

On voit alors que $(\tau_n)_{n \in \mathbb{N}}$ est également une suite de $]a, b[$ tendant vers b . Par conséquent $(X(\tau_n))_{n \in \mathbb{N}}$ est une suite convergente, et toutes les suites extraites de $(X(\tau_n))_{n \in \mathbb{N}}$ convergent vers la même limite. Donc

$$\lim_{k \rightarrow \infty} X_{max}(t_k) = \lim_{k \rightarrow \infty} X_{max}(\tilde{t}_k).$$

Soit $X_b \in \mathbb{R}^m$ la limite commune de toutes les suites $(X_{max}(t_n))_{n \in \mathbb{N}}$ vérifiant les hypothèses précédentes :

$$X_b := \lim_{t \rightarrow b^-} X_{max}(t).$$

Cela permet de prolonger par continuité X_{max} en b .

Montrons maintenant que X_{max} est dérivable à gauche en b avec $X'_{max}(b) = f(b, X_{max}(b))$. Pour tout $t \in V$, on a $X'_{max}(t) = f(t, X_{max}(t))$. Par continuité de f , et par prolongement par continuité de X_{max} en b , cette fonction admet $f(b, X_b)$ comme limite pour $t \rightarrow b^-$. Par le théorème de prolongement des dérivées, on en déduit que le prolongement par continuité de X_{max} est dérivable à gauche en b , de dérivée $f(b, X_b)$.

Ainsi, X_{max} est solution du problème de Cauchy (3.1) sur $]a, b]$ ce qui est en contradiction avec le fait fait que X_{max} est solution maximale. \square

On déduit de ce lemme le résultat suivant :

Théorème 3.2.5 (Dichotomie existence globale/explosion). *On suppose que les hypothèses du Théorème 3.1.6 sont vérifiées avec $I = \mathbb{R}$, $U = \mathbb{R}^m$, et on considère une solution maximale X_{max} de (3.1), définie sur un intervalle $I_{max} =]a, b[$.*

Alors l'une et l'une seulement des deux propriétés suivantes est vérifiée :

- (i) soit $b = +\infty$;
- (ii) soit $b < +\infty$ et

$$\lim_{t \rightarrow b_-} \|X_{max}(t)\| = \infty.$$

La preuve de cette proposition reprend des éléments de la démonstration du lemme précédent.

Démonstration. L'une et l'une seulement des deux propriétés suivantes est vérifiée : soit $b = +\infty$, soit $b < +\infty$. Pour conclure il suffit de montrer que $b < +\infty$ implique que $\lim_{t \rightarrow b_-} \|X_{max}(t)\| = \infty$.

Raisonnons par l'absurde : supposons que $b < +\infty$ et que $\|X_{max}\|$ ne tend pas vers $+\infty$ en b_- . Par le Lemme 3.4.1, il suffit de montrer que X_{max} est bornée sur un voisinage à gauche de b : cela impliquera alors que $b = +\infty$, une contradiction.

Puisque $\|X_{max}\|$ ne tend pas vers $+\infty$ en b_- , par caractérisation séquentielle, il existe une suite de temps $(s_n)_{n \geq 0}$ tendant vers b par valeurs inférieures, ainsi que $R > 0$, tels que $\|X_{max}(s_n)\| \leq R$ pour tout $n \in \mathbb{N}$. Soit V un voisinage à gauche de b qui est borné et inclus dans $]a, b[$. On pose alors

$$C_0 = \sup_{t \in \bar{V}, \|Z\| \leq R+1} \|f(t, Z)\|.$$

Comme f est de classe \mathcal{C}^1 et \bar{V} est compact, il s'ensuit que C_0 est fini.

Pour tout $n \in \mathbb{N}$, on pose

$$\tilde{s}_n = \inf \{s \in [s_n, b[\mid \|X_{max}(s)\| > R + 1\}$$

Si un des \tilde{s}_n n'est pas bien défini (l'ensemble définissant l'inf est vide), on en déduit que X_{max} est bornée sur un voisinage à gauche de b , ce qu'on veut démontrer.

On suppose donc que tous les \tilde{s}_n sont bien définis, ils sont donc strictement inférieurs à b . Par continuité de X_{max} , on a également $\tilde{s}_n > s_n$ (sur un voisinage à droite de s_n , on a $\|X_{max}(t)\| \leq \|X_{max}(s_n)\| + 1 \leq R + 1$). De nouveau par continuité de X_{max} , on a que $\|X_{max}(\tilde{s}_n)\| = R + 1$.

Puisque X_{max} est solution de (3.1), on en déduit que

$$\|X'_{max}(t)\| \leq C_0 \quad \forall t \in [s_n, \tilde{s}_n].$$

Par l'inégalité des accroissements finis, cela implique que

$$\|X_{max}(s_n) - X_{max}(\tilde{s}_n)\| \leq (\tilde{s}_n - s_n)C_0.$$

Or $\|X_{max}(s_n)\| \leq R$ et $\|X_{max}(\tilde{s}_n)\| = R + 1$. Par l'inégalité triangulaire inversée, on en déduit que $\|X_{max}(s_n) - X_{max}(\tilde{s}_n)\| \geq 1$, si bien que $\tilde{s}_n - s_n \geq \frac{1}{C_0}$, donc $\tilde{s}_n \geq s_n + \frac{1}{C_0}$.

Or pour n assez grand on a $s_n \geq b - \frac{1}{2C_0}$, ce qui implique que $\tilde{s}_n > b$, ce qui est une contradiction avec la définition de \tilde{s}_n . \square

Exemple 3.2.6. On considère le problème de Cauchy

$$\begin{cases} x'(t) = x^2(t) \\ x(0) = 1. \end{cases}$$

C'est un problème de Cauchy de la forme (3.1) avec $m = 1$, $t_0 = 0$, $x_0 = 1$ et autonome avec $f(y) = y^2$. Comme f est une fonction \mathcal{C}^1 sur \mathbb{R} , on peut appliquer le théorème de Cauchy-Lipschitz et obtenir une solution maximale x_{max} du problème de Cauchy ci-dessus, définie sur un intervalle ouvert I_{max} contenant 0.

On remarque que la fonction constante égale à zéro sur \mathbb{R} est solution de $x'(t) = x^2(t)$. Par conséquent, la solution maximale x_{max} ne s'annule pas sur son intervalle de définition I_{max} (ce qui implique d'ailleurs que $x_{max}(t) > 0$ pour tout $t \in I_{max}$). Effectivement, s'il existait $t_1 \in I_{max}$ tel que $x_{max}(t_1) = 0$ alors d'après l'unicité du théorème de Cauchy-Lipschitz appliqué au problème de Cauchy avec donnée initiale égale à 0 en t_1 on obtiendrait que x_{max} coïncide avec la fonction nulle sur $I_{max} \cap \mathbb{R}$, donc en particulier en t_0 , ce qui n'est pas le cas puisque $x_{max}(0) = 1$.

Pour tout $t \in I_{max}$, on a donc

$$\frac{x'_{max}(t)}{x_{max}^2(t)} = 1,$$

et en intégrant entre 0 et t

$$1 - \frac{1}{x_{max}(t)} = t \quad \forall t \in I_{max},$$

ce qui nous donne déjà l'information $1 \notin I_{max}$ et on a

$$x_{max}(t) = \frac{1}{1-t} \quad \forall t \in I_{max}.$$

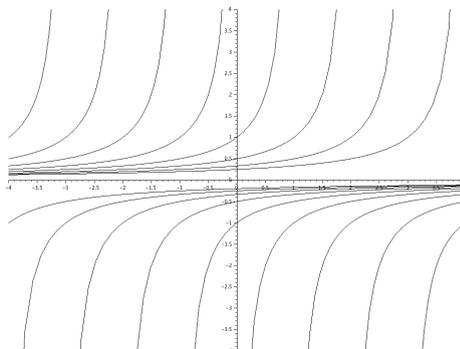
Comme I_{max} est un intervalle ouvert $]a, b[$ contenant 0 mais pas 1 donc $-\infty \leq a < 0$ et $0 < b \leq 1$. D'autre part, comme la fonction $\frac{1}{1-t}$ est bornée au voisinage de tout point de $] - \infty, 1[$ on conclut par la proposition précédente que $a = -\infty$ et que $b = 1$, donc

$$I_{max} =] - \infty, 1[.$$

Plus généralement, pour une donnée initiale $x(0) = x_0$ on trouve de la même façon que

$$x_{max}(t) = \begin{cases} \frac{1}{\frac{1}{x_0} - t}, & I_{max} =] - \infty, \frac{1}{x_0}[, & \text{si } x_0 > 0, \\ 0, & I_{max} = \mathbb{R}, & \text{si } x_0 = 0, \\ \frac{1}{\frac{1}{x_0} - t}, & I_{max} =]\frac{1}{x_0}, +\infty[, & \text{si } x_0 < 0, \end{cases}$$

En représentant dans le même repère plusieurs telles solutions, voir dessin ci-dessous, nous remarquons deux choses : par tout point du plan passe une solution (grâce à l'existence du théorème de Cauchy-Lipschitz), et les solutions ne se croisent pas (grâce à l'unicité du théorème de Cauchy-Lipschitz).



3.3 Quelques conséquences de l'unicité

Solutions à valeurs réelles

Pour des équations à valeurs réelles, l'unicité des solutions a plusieurs conséquences remarquables. La première est que les solutions d'équations autonomes sont soit constantes, soit strictement monotones.

Proposition 3.3.1 (Cas réel autonome). *Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . On considère une solution x de l'équation*

$$x'(t) = f(x(t)),$$

d'intervalle de définition I . Alors x est monotone, et plus précisément soit x est strictement monotone, soit x est constante.

Démonstration. Si $x'(t_0) = 0$, alors $f(x(t_0)) = 0$. Posons $y : \mathbb{R} \rightarrow \mathbb{R}$ la fonction constante égale à $x(t_0)$. Alors y est solution de $y' = f(y)$, avec $y(t_0) = x(t_0)$. Par unicité dans le théorème de Cauchy-Lipschitz, on en déduit que $x = y$.

Ainsi, si x' change de signe, alors par le théorème des valeurs intermédiaires il existe t tel que $x'(t) = 0$, ce qui implique par ce qu'on vient de voir que x est constante. Ainsi, x' est de signe constant (soit strictement négatif, soit nul, soit strictement positif). \square

Une autre conséquence est le fait que les solutions sont ordonnées.

Proposition 3.3.2 (Solutions ordonnées). *Soit $f : I \times \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . On considère deux solutions x, \tilde{x} de l'équation*

$$x'(t) = f(t, x(t)),$$

d'intervalles de définition I, \tilde{I} . On suppose qu'il existe $t_0 \in I \cap \tilde{I}$ tel que

$$x(t_0) < \tilde{x}(t_0).$$

Alors

$$x(t) < \tilde{x}(t) \quad \forall t \in I \cap \tilde{I}.$$

Démonstration. Raisonnons par l'absurde et supposons qu'il existe $t_1 \in I \cap \tilde{I}$ tel que

$$x(t_1) \geq \tilde{x}(t_1).$$

De deux choses l'une : soit $x(t_1) = \tilde{x}(t_1)$, soit $x(t_1) > \tilde{x}(t_1)$, et dans ce cas il existe $t_2 \in]t_0, t_1[$ tel que $x(t_2) = \tilde{x}(t_2)$, grâce au théorème des valeurs intermédiaires appliqué à la fonction $x(t) - \tilde{x}(t)$. Dans tous les cas, on obtient l'existence de $t_2 \in I \cap \tilde{I}$ tel que $x(t_2) = \tilde{x}(t_2) =: x_0$. Il s'ensuit que x et \tilde{x} sont deux solutions du problème de Cauchy avec donnée initiale x_0 à t_2 , donc par l'unicité du théorème de Cauchy-Lipschitz local on obtient que $x(t) = \tilde{x}(t)$ pour tout $t \in I \cap \tilde{I}$. Or par définition $t_0 \in I \cap \tilde{I}$. Donc $x(t_0) = \tilde{x}(t_0)$ ce qui contredit l'hypothèse. \square

Croisement de solutions

Dans le cas des équations autonomes, le théorème de Cauchy-Lipschitz a plusieurs conséquences frappantes. Le premier est un résultat de non croisement, qui implique que les trajectoires des solutions sont soit disjointes, soit confondues.

Proposition 3.3.3 (Non-croisement). *Si $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ est de classe \mathcal{C}^1 et si X_1, X_2 sont des solutions de l'équation*

$$X'(t) = f(t, X(t)),$$

d'intervalles de définition I_1, I_2 , telles qu'il existe $t_1 \in I_1$ et $t_2 \in I_2$ pour lesquels $X_1(t_1) = X_2(t_2)$ alors $X_1(t_1 + u) = X_2(t_2 + u)$ pour tout $u \in I_1 - t_1 \cap I_2 - t_2$.

Démonstration. Pour $i = 1, 2$, on pose

$$\begin{aligned} \tilde{X}_i : I_i - t_i &\longrightarrow \mathbb{R} \\ u &\longmapsto X_i(t_i + u). \end{aligned}$$

On a $\tilde{X}_1(0) = X_1(t_1) = X_2(t_2) = \tilde{X}_2(0)$, et de plus $\tilde{X}_i'(u) = X_i'(t_i + u) = f(X_i(t_i + u)) = f(\tilde{X}_i(u))$. Donc \tilde{X}_1 et \tilde{X}_2 sont tous les deux solutions du même problème de Cauchy ; par unicité dans le théorème de Cauchy-Lipschitz on a $\tilde{X}_1 = \tilde{X}_2$ sur l'intersection $I_1 - t_1 \cap I_2 - t_2$ de leurs intervalles de définition. \square

Un second exemple est une propriété de périodicité en temps, qui sera utilisée plus tard dans les exemples du pendule simple et du système de Lotka-Volterra.

Proposition 3.3.4 (Périodicité en temps). *Soit $f \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R}^m)$, $t_0 \in \mathbb{R}$, $X_0 \in \mathbb{R}^m$. On considère le problème de Cauchy différentiel*

$$\begin{cases} X'(t) = f(X(t)), \\ X(t_0) = X_0. \end{cases} \quad (3.7)$$

Soit X la solution maximale de (3.7), et soit I_{max} son intervalle (ouvert) de définition. On suppose qu'il existe $T \in \mathbb{R}^$ tel que $t_0 + T \in I_{max}$ et tel que*

$$X(t_0 + T) = X_0.$$

Alors $I_{max} = \mathbb{R}$ et X est périodique de période $|T|$.

Démonstration. Notons $I_{max} =]a, b[$. On définit

$$\tilde{X}(t) = X(t+T) \quad \forall t \in]a-T, b-T[.$$

Alors \tilde{X} est une fonction \mathcal{C}^1 , $\tilde{X}(t_0) = X(t_0+T) = X_0$, et

$$\tilde{X}'(t) = X'(t+T) = f(X(t+T)) = f(\tilde{X}(t)).$$

Donc \tilde{X} est solution de (3.7). Comme X est la solution maximale, on en déduit que $]a-T, b-T[\subset]a, b[$ et

$$\tilde{X}(t) = X(t) \quad \forall t \in]a-T, b-T[.$$

Pour fixer les idées, à ce stade, supposons $T > 0$ (le raisonnement est strictement analogue si $T < 0$). D'après la première propriété nécessairement $\inf I_{max} = -\infty$ et I est de la forme $] -\infty, b[$. On déduit de la deuxième propriété que

$$X(t) = X(t+T) \quad \forall t \in] -\infty, b-T[,$$

ce qui s'écrit encore

$$X(t-T) = X(t) \quad \forall t \in] -\infty, b[.$$

En vertu du lemme 3.2.4, on voit que nécessairement $b = +\infty$. En effet, si on a $b < +\infty$, alors $\exists t_n \in] -\infty, b[$, $\lim_{n \rightarrow \infty} t_n = b$ et telle que $\lim_{n \rightarrow \infty} \|X(t_n)\| = \infty$. Mais d'autre part,

$$\lim_{n \rightarrow \infty} X(t_n) = \lim_{n \rightarrow \infty} X(t_n - T) = X(b - T) \in \mathbb{R}^m,$$

contradiction.

Donc $I_{max} = \mathbb{R}$ et X est T -périodique. □

3.4 Critères d'existence globale

On va à présent donner quelques conditions pour que les solutions de l'équation (3.1) soient globales, c'est-à-dire définies sur \mathbb{R} . Pour cela, il convient de considérer des champs de vecteurs b tels que b est défini sur $\mathbb{R} \times U$, avec U ouvert de \mathbb{R}^m (autrement l'équation n'a pas de sens!). Remarquons que l'on sait déjà que si une solution est bornée, alors elle est globale (cf. Lemme 3.2.4). Ce lemme se généralise de la façon suivante :

Lemme 3.4.1. *Soit $f : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ une fonction de classe \mathcal{C}^1 , $t_0 \in \mathbb{R}$, et $X_0 \in \mathbb{R}^m$. On considère la solution maximale X_{max} de l'équation (3.1), définie sur un intervalle I_{max} .*

On suppose qu'il existe une fonction $h : \mathbb{R} \rightarrow \mathbb{R}_+$, continue, telle que

$$\|X_{max}(t)\| \leq h(t) \quad \forall t \in I_{max}.$$

Alors $I_{max} = \mathbb{R}$.

Démonstration. On raisonne par l'absurde et on suppose que $\sup(I_{max}) \in \mathbb{R}$. On pose

$$K = [t_0, \sup(I_{max})].$$

K est un intervalle compact et un voisinage à gauche de $\sup(I_{max})$. Comme h est continue sur \mathbb{R} , h est bornée sur K .

En utilisant l'hypothèse du lemme, on en déduit que X_{max} est bornée sur K . D'après le lemme 3.2.4, $\sup(I_{max}) = +\infty$: absurde.

De même, on montre que $\inf I_{max} = -\infty$. □

On va ensuite utiliser des propriétés structurelles de l'équation (ou plus spécifiquement, du champ de vecteurs f) qui assurent que l'on a une solution globale. Pour cela on utilise le résultat suivant :

Lemme 3.4.2 (Lemme de Gronwall – forme différentielle). *Soit I un intervalle de \mathbb{R} , $\varphi, \psi : I \rightarrow \mathbb{R}$ des fonctions telles que φ est de classe \mathcal{C}^1 et ψ est continue. On suppose qu'il existe une constante C telle que*

$$\varphi'(t) \leq \psi(t)(C + \varphi(t)) \quad \forall t \in I.$$

Alors pour tout $t \geq u$ dans I , on a

$$\varphi(t) \leq \varphi(u)e^{\int_u^t \psi(s)ds} + C \left(e^{\int_u^t \psi(s)ds} - 1 \right).$$

Démonstration. On remarque que pour tout $t \in I$, on a

$$\varphi'(t) - \psi(t)\varphi(t) \leq C\psi(t).$$

On multiplie cette inégalité par $e^{-\int_u^t \psi(s)ds}$ (on remarque que cette quantité est toujours positive). On obtient alors

$$\frac{d}{dt} \left(\varphi(t)e^{-\int_u^t \psi(s)ds} \right) \leq C \frac{d}{dt} \left(-e^{-\int_u^t \psi(s)ds} \right) \quad \forall t \in I. \quad (3.8)$$

On intègre cette inégalité entre u et t :

$$\varphi(t)e^{-\int_u^t \psi(s)ds} - \varphi(u) \leq C \left(-e^{-\int_u^t \psi(s)ds} + 1 \right).$$

En multipliant de nouveau par $\exp \left(\int_u^t \psi(s)ds \right)$, on obtient l'estimation annoncée.⁴ \square

Corollaire 3.4.3. *Soit I un intervalle de \mathbb{R} , $\varphi, \psi : I \rightarrow \mathbb{R}$ des fonctions telles que φ est de classe \mathcal{C}^1 et ψ est continue. On suppose qu'il existe une constante C telle que*

$$|\varphi'(t)| \leq \psi(t)(C + \varphi(t)) \quad \forall t \in I.$$

Alors pour tout t, u dans I , on a

$$\varphi(t) \leq \varphi(u)e^{\int_{\min\{t,u\}}^{\max\{t,u\}} \psi(s)ds} + C \left(e^{\int_{\min\{t,u\}}^{\max\{t,u\}} \psi(s)ds} - 1 \right).$$

Démonstration. Pour $t \geq u$ l'inégalité voulue est exactement celle de Gronwall. Pour $t \leq u$ on applique aussi Gronwall aux fonctions $-\varphi, -\psi$ avec la constante $-C$ car on a

$$-\varphi'(t) \leq \psi(t)(C + \varphi(t)) = -\psi(t)(-C - \varphi(t)) \quad \forall t \in I.$$

On obtient ainsi, pour $u \geq t$

$$-\varphi(u) \leq -\varphi(t)e^{\int_t^u (-\psi(s))ds} - C \left(e^{\int_t^u (-\psi(s))ds} - 1 \right),$$

ce qui nous donne, en multipliant par $e^{\int_t^u \psi(s)ds}$, l'inégalité de l'énoncé. \square

4. On peut aussi remarquer que la fonction $x(t) = \varphi(u)e^{\int_u^t \psi(s)ds} + C \left(e^{\int_u^t \psi(s)ds} - 1 \right)$ vérifie $x'(t) = \psi(t)(C + x(t))$ avec $x(u) = \varphi(u)$. Comme $(\varphi - x)' \leq \psi(\varphi - x)$ on a que la fonction $(\varphi(t) - x(t))e^{-\int_u^t \psi(s)ds}$ est décroissante et vaut 0 en u , donc on conclut que $\varphi(t) \leq x(t)$ pour tout $t \geq u$.

On arrive finalement au résultat suivant, dont les hypothèses ne portent que sur la forme de l'équation :

Théorème 3.4.4 (Théorème de Cauchy-Lipschitz global). *Soit $f : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^1 . On suppose qu'il existe une fonction continue $\psi : \mathbb{R} \rightarrow \mathbb{R}_+$ telle que*

$$\|f(t, Z)\| \leq \psi(t)(1 + \|Z\|) \quad \forall (t, Z) \in \mathbb{R} \times \mathbb{R}^m.$$

Soit X_{max} une solution maximale de l'équation (3.1). Alors elle est définie sur \mathbb{R} (on dit aussi qu'elle est globale).

Démonstration. Quitte à modifier la fonction ψ en la multipliant par une constante, on peut toujours supposer que $\|\cdot\|$ est la norme euclidienne. On pose alors

$$\varphi(t) = \|X_{max}(t)\|^2.$$

La fonction φ est de classe \mathcal{C}^1 sur l'intervalle de définition I_{max} de X_{max} , et pour tout $t \in I_{max}$, on a, en utilisant l'équation puis l'inégalité de Cauchy-Schwarz

$$\varphi'(t) = 2\langle X'_{max}(t), X_{max}(t) \rangle = 2\langle f(t, X_{max}(t)), X_{max}(t) \rangle \quad (3.9)$$

$$|\varphi'(t)| \leq 2\psi(t)(1 + \|X_{max}(t)\|)\|X_{max}(t)\| = 2\psi(t)(\sqrt{\varphi(t)} + \varphi(t)). \quad (3.10)$$

En utilisant l'inégalité

$$a \leq \frac{1}{2} + \frac{a^2}{2} \quad \forall a \in \mathbb{R},$$

avec $a = \sqrt{\varphi(t)}$ on obtient finalement, pour tout $t \in I_{max}$,

$$|\varphi'(t)| \leq 2\psi(t) \left(\frac{1}{2} + \frac{3}{2}\varphi(t) \right) = 3\psi(t) \left(\frac{1}{3} + \varphi(t) \right).$$

En utilisant le Corollaire 3.4.3 du Lemme de Gronwall, on en déduit que

$$\|X_{max}(t)\| = \sqrt{\varphi(t)} \leq g(t) \quad \forall t \in I_{max},$$

avec

$$g(t) = \left(\varphi(t_0) e^{\int_{\min\{t_0, t\}}^{\max\{t_0, t\}} 3\psi(s) ds} + \frac{1}{3} \left(e^{\int_{\min\{t_0, t\}}^{\max\{t_0, t\}} 3\psi(s) ds} - 1 \right) \right)^{\frac{1}{2}}.$$

Comme g est continue sur \mathbb{R} , en utilisant le Lemme 3.4.1, on en déduit que $I_{max} = \mathbb{R}$. \square

On a déjà vu que les solutions des équations différentielles linéaires à coefficients constants (avec second membre) sont globales. Le théorème de Cauchy-Lipschitz global permet de généraliser ce résultat au cas non constant.

Corollaire 3.4.5. *Les solutions des équations différentielles linéaires sont globales.*

Démonstration. Soit donc $A : \mathbb{R} \rightarrow M_m(\mathbb{R})$ et $B : \mathbb{R} \rightarrow \mathbb{R}^m$ deux applications continues. On se donne X_{max} une solution maximale de l'équation $X'(t) = A(t)X(t) + B(t)$, autrement dit on a ici $f(t, X) = A(t)X + B(t)$. Alors (rappelons que $\|A\|$ est la norme d'opérateur de A associée à la norme $\|\cdot\|$, voir le paragraphe 2.3.1).

$$\begin{aligned} \|f(t, X)\| &= \|A(t)X + B(t)\| \\ &\leq \|A(t)X\| + \|B(t)\| \\ &\leq \|A(t)\| \|X\| + \|B(t)\| \\ &\leq \psi(t)(1 + \|X\|) \end{aligned}$$

où on a posé $\psi(t) = \max(\|A(t)\| \|X\|, \|B(t)\|)$. Cette fonction ψ est continue, on peut donc appliquer le théorème de Cauchy-Lipschitz global qui assure que la solution X_{max} est définie sur \mathbb{R} . \square

3.5 Deux exemples de résolution par des séries

3.5.1 Par série entière

On veut résoudre l'équation différentielle suivante :

$$-y''(t) + ty'(t) + y(t) = 1. \quad (\text{EDSE})$$

Pour cela, on cherche une solution sous la forme d'une série entière

$$y(t) = \sum_{n \geq 0} a_n t^n$$

de rayon de convergence > 0 . On suppose donc que y est solution de (EDSE) à l'intérieur de son intervalle de convergence. Rappelons que pour t dans cet intervalle,

$$y'(t) = \sum_{n \geq 1} n a_n t^{n-1} \quad \text{et} \quad y''(t) = \sum_{n \geq 2} n(n-1) a_n t^{n-2}.$$

On injecte ces expressions dans (EDSE), ce qui donne :

$$-\sum_{n \geq 2} n(n-1) a_n t^{n-2} + t \sum_{n \geq 1} n a_n t^{n-1} + \sum_{n \geq 0} a_n t^n = 1,$$

et donc par réindexation de la première somme,

$$\sum_{n \geq 0} -(n+2)(n+1) a_{n+2} t^n + \sum_{n \geq 1} n a_n t^n + \sum_{n \geq 0} a_n t^n = 1,$$

d'où, en regroupant les sommes, et en isolant le terme pour $n = 0$,

$$\sum_{n \geq 1} (-(n+2)(n+1) a_{n+2} + n a_n + a_n) t^n + (2a_2 + a_0 - 1) = 0.$$

Or une série entière (à gauche) est nulle si et seulement si tous ses coefficients sont nuls, par conséquent d'une part $2a_2 = 1 - a_0$ et d'autre part, pour tout $n \geq 1$,

$$-(n+2)(n+1) a_{n+2} + n a_n + a_n = -(n+1)((n+2) a_{n+2} + a_n) = 0.$$

Ainsi,

$$a_2 = \frac{a_0 - 1}{2} \quad \text{et} \quad \forall n \geq 1, \quad a_{n+2} = \frac{a_n}{n+2}.$$

Par récurrence sur n , on en déduit que

$$a_{2n} = \frac{1}{(2n)(2n-2) \cdots 4} a_2 = \frac{1-a_0}{(2n)(2n-2) \cdots 4 \cdot 2} = \frac{a_0 - 1}{2^n n!},$$

d'autre part

$$\begin{aligned} a_{2n+1} &= \frac{1}{(2n+1)(2n-1) \cdots 3} a_1 \\ &= \frac{(2n)(2n-2) \cdots 2}{(2n+1)(2n)(2n-1)(2n-2) \cdots 3 \cdot 2} a_1 = \frac{2^n n!}{(2n+1)!} a_1. \end{aligned}$$

Remarquons que $|a_{2n}| \leq |1 - a_0|/n!$ et $|a_{2n+1}| \leq |a_1|/n!$ (la seconde majoration se voit par exemple sur la première ligne de la formule pour a_{2n+1}). Par conséquent, si on choisit

$a_0, a_1 \in \mathbb{R}$, et on pose $a_{2n} = \frac{a_0-1}{2^n n!}$ et $a_{2n+1} = \frac{2^n n!}{(2n+1)!} a_1$, alors le rayon de convergence de la série $\sum_{n \geq 0} a_n t^n$ est infini (les coefficients sont inférieurs en module à ceux de la série exponentielle). On peut tout à fait "remonter" les calculs ci-dessus, ce qui montre que la série entière ainsi définie est solution de (EDSE).

Conclusion : pour tout $a_0, a_1 \in \mathbb{R}$, il existe une solution de (EDSE) satisfaisant $y(0) = a_0$ et $y'(0) = a_1$; celle-ci est développable en série entière en 0, de rayon de convergence infini, et s'écrit $y(t) = \sum_{n \geq 0} a_n t^n$ avec, pour tout $n \geq 1$, les formules $a_{2n} = \frac{1-a_0}{2^n n!}$ et $a_{2n+1} = \frac{2^n n!}{(2n+1)!} a_1$.

De plus, cette solution est unique, par théorème de Cauchy-Lipschitz. En effet, si on pose $Y = (y, y')$, on voit que Y satisfait l'EDO

$$Y'(t) = \begin{pmatrix} 0 & 1 \\ -1 & -t \end{pmatrix} Y(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

et l'application

$$(t, Y) \mapsto \begin{pmatrix} 0 & 1 \\ -1 & -t \end{pmatrix} Y + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

est de classe C^1 , ce qui permet d'appliquer le théorème de Cauchy-Lipschitz local.

Notons que si on impose $y'(0) = 0$, on obtient

$$y(t) = a_0 + \sum_{n \geq 1} \frac{1-a_0}{2^n n!} x^{2n} = 2a_0 - 1 + (1-a_0) \sum_{n \geq 0} \frac{(x^2/2)^n}{n!} = 2a_0 - 1 + e^{x^2/2}.$$

3.5.2 Par série de Fourier

On veut résoudre l'équation différentielle suivante :

$$y''(t) + e^{it} y(t) = 0. \quad (\text{EDF})$$

On cherche une solution de classe C^2 de cette équation qui est 2π -périodique. Rappelons que

$$c_n(y) = \frac{1}{2\pi} \int_0^{2\pi} y(s) e^{-ins} ds.$$

Alors, par double intégration par parties, on obtient que

$$c_n(y'') = \frac{1}{2\pi} \int_0^{2\pi} y''(s) e^{-ins} ds = -n^2 c_n(y).$$

De plus, si y est solution de (EDF), alors $y''(t) = -e^{it} y(t)$, si bien que

$$c_n(y'') = \frac{1}{2\pi} \int_0^{2\pi} -e^{is} y(s) e^{-ins} ds = \frac{1}{2\pi} \int_0^{2\pi} -y(s) e^{-i(n-1)s} ds = -c_{n-1}(y),$$

et donc

$$-n^2 c_n(y) = -c_{n-1}(y),$$

autrement dit

$$c_n(y) = \frac{c_{n-1}(y)}{n^2}.$$

Ainsi, par une récurrence directe, on obtient pour tout $n \geq 1$:

$$c_n(y) = \frac{c_0(y)}{(n!)^2} \quad \text{et} \quad c_{-n}(y) = 0.$$

Rappelons que si y est une fonction 2π -périodique et de classe C^2 , alors sa série de Fourier converge (normalement). Par conséquent, si elle est de plus solution de (EDF), alors on peut écrire

$$y(t) = \sum_{n \geq 0} \frac{y(0)}{(n!)^2} e^{int}.$$

Réciproquement, si on pose

$$y(t) = \sum_{n \geq 0} \frac{y(0)}{(n!)^2} e^{int},$$

alors cela définit une série normalement convergente, et on vérifie facilement que cette série est solution de (EDF) de classe C^2 .

Conclusion : les seules solutions de (EDF) de classe C^2 et 2π -périodiques sont les fonctions s'écrivant

$$y(t) = C \sum_{n \geq 0} \frac{1}{(n!)^2} e^{int},$$

avec $C \in \mathbb{C}$

3.6 Tracé des trajectoires de solutions d'équations différentielles

3.6.1 Un peu de vocabulaire

Définition 3.6.1. On dit qu'une équation différentielle est **autonome** lorsqu'elle peut s'écrire sous la forme

$$x^{(n)}(t) = f(x(t), x'(t), \dots, x^{(n-1)}(t)),$$

avec $f : \mathbb{R}^m \times \dots \times \mathbb{R}^m \rightarrow \mathbb{R}^m$. Autrement dit, la fonction f intervenant dans l'équation différentielle ne dépend pas explicitement du temps.

Exemple 3.6.2. Donner ici des exemples dans lesquels on identifie la fonction f : cas autonomes et non autonomes.

Remarquons qu'on peut transformer une EDO non-autonome

$$X'(t) = f(t, X(t)),$$

en une EDO autonome, en considérant comme inconnue $Y(t) = (t, X(t)) \in \mathbb{R}^{m+1}$ qui satisfait l'équation

$$Y'(t) = g(Y(t)),$$

avec g définie sur \mathbb{R}^{m+1} par

$$g(Z) = (1, f(Z)).$$

Définition 3.6.3. Un **champ de vecteurs** de régularité \mathcal{C}^k est une application $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^k , définie par ses m composantes

$$f : \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1, \dots, x_m) \\ \vdots \\ f_m(x_1, \dots, x_m) \end{pmatrix}.$$

Si $m = 2$ (et éventuellement si $m = 3$), on peut représenter un champ de vecteurs en traçant, en un nombre fini mais élevé de points (x_1, x_2) (ou (x_1, x_2, x_3)), le vecteur de coordonnées $(f_1(x_1, x_2), f_2(x_1, x_2))$.

Par ailleurs, si $m = 2$, on peut aussi s'intéresser à l'équation différentielle autonome

$$X'(t) = f(X(t)).$$

Soit $X_0 \in \mathbb{R}^2$ tel que $f(X_0) = (0, 0)$. On vérifie alors que $X(t) = X_0$ est une solution globale de l'équation. Son tracé dans le plan est juste le point X_0 .

Soit $X_0 \in \mathbb{R}^2$ tel que $f(X_0) \neq (0, 0)$. Soit X une solution de l'équation différentielle précédente telle que $X(t_0) = X_0$ pour un certain $t_0 \in \mathbb{R}$. Il découle que le vecteur $f(X_0)$ est un vecteur tangent à la courbe paramétrée $\mathcal{C} = \{X(t) \mid t \in \mathbb{R}\}$. Ainsi la représentation du champ de vecteurs F évoquée ci-dessus donne une idée de l'allure des courbes représentant les solutions de l'équation différentielle : on trace les courbes en se laissant guider par le champ de vecteurs.

Définition 3.6.4. Les solutions de l'équation différentielle $X'(t) = f(X(t))$ sont appelées **courbes intégrales** du champ de vecteurs f .

Nous allons voir au chapitre suivant que nous pouvons tracer facilement dans le plan les trajectoires des solutions d'équations linéaires à coefficients constants d'ordre 1, vectorielles avec $m = 2$. Ceci est grâce au fait que nous saurons résoudre explicitement l'équation.

Dans le cas général une telle précision n'est pas attendue, mais nous pouvons, selon les cas, avoir une idée de l'allure du tracé de solutions. Dans la suite nous allons nous restreindre au cas d'équations différentielles scalaires.

3.6.2 Contraintes sur les solutions scalaires

Dans ce dernier sous-chapitre nous donnons quelques pistes pour comprendre l'allure du tracé de solutions d'équations différentielles scalaires

$$x'(t) = f(t, x(t)).$$

a) Étude du champ de directions.

Dans le cas d'équations différentielles scalaires les valeurs de la fonction f vont définir le **champ de directions** de l'équation. Plus précisément, la valeur $f(t_0, x_0)$ va être la pente de la tangente à t_0 au graphe de la solution du problème de Cauchy avec donnée initiale $x(t_0) = x_0$. Ainsi, en dessinant en plusieurs points du plan (t, y) des traits de pente $f(t, y)$ nous allons avoir une idée de l'allure du tracé des solutions de l'équation $x'(t) = f(t, x(t))$: on trace les courbes en se laissant guider par le champ de directions.

Par exemple, si $f(t_0, x_0) > 0$ alors la solution x de l'équation, telle que $x(t_0) = x_0$, est strictement croissante localement en t_0 . Il est utile donc de savoir quelles sont les endroits du plan où f prend des valeurs strictement positives, et celles où f prend des valeurs strictement négatives. Pour trouver ces zones il convient de déterminer les endroits où f s'annule.

Dans ce même esprit, il est utile de comprendre quels sont les endroits où le champs de directions a la même pente :

Définition 3.6.5. Soit $a \in \mathbb{R}$. On appelle **isocline de la pente a** de l'équation $x'(t) = f(t, x(t))$ l'ensemble des points (t, y) tels que $f(t, y) = a$.

Exemples : dessiner des tracés de solutions des ODE suivantes : $x'(t) = x(t)$, $x'(t) = 2x^2(t)$, $x'(t) = t(2t^2 - x(t))$.

b) Recherches de symétries.

- Si $f(t, -y) = -f(t, y)$ alors si $x(t)$ est solution avec $x(t_0) = x_0$, $\tilde{x}(t) = -x(t)$ est aussi une solution de l'équation satisfaisant $\tilde{x}(t_0) = -x_0$. Il s'ensuit qu'il suffit de tracer les solutions seulement dans le demi-plan supérieur, et faire après une symétrie par rapport à l'axe des abscisses (comme c'est le cas pour $x'(t) = x(t)$).
- Si $f(-t, y) = -f(t, y)$ alors si $x(t)$ est solution avec $x(t_0) = x_0$, $\tilde{x}(t) = x(-t)$ est aussi une solution de l'équation satisfaisant $\tilde{x}(-t_0) = x_0$. Il s'ensuit qu'il suffit de tracer les solutions seulement dans le demi-plan d'abscisses positives $t \geq 0$, et faire après une symétrie par rapport à l'axe des ordonnées (comme c'est le cas pour $x'(t) = t(2t^2 - x(t))$).
- Si $f(-t, -y) = f(t, y)$ alors si $x(t)$ est solution avec $x(t_0) = x_0$, $\tilde{x}(t) = -x(-t)$ est aussi une solution de l'équation satisfaisant $\tilde{x}(-t_0) = -x_0$. Il s'ensuit qu'il suffit de tracer les solutions seulement dans un demi-plan délimité par une droite passant par l'origine, et faire après une symétrie par rapport à l'origine (comme c'est le cas pour $x'(t) = 2x^2(t)$).

c) Sur-solutions et sous-solutions, pièges à solutions scalaires.

Définition 3.6.6. On appelle **sur-solution ou barrière supérieure** de l'équation

$$x'(t) = f(t, x(t))$$

toute fonction $u : I \rightarrow U \subset \mathbb{R}$ de classe \mathcal{C}^1 telle que

$$u'(t) \geq f(t, u(t)), \quad \forall t \in I.$$

De même, on appelle **sous-solution ou barrière inférieure** de l'équation toute fonction $v : I \rightarrow U \subset \mathbb{R}$ de classe \mathcal{C}^1 telle que

$$v'(t) \leq f(t, v(t)), \quad \forall t \in I.$$

Par exemple si f est la fonction nulle, toute fonction u croissante est un exemple de barrière supérieure et toute fonction v décroissante est un exemple de barrière inférieure. Par exemple dessiner $u(t) = t$, $v(t) = -t$, $(t_0, x_0) = (2, 1)$, $x(t) = 1$.

Une façon naturelle de trouver des sur-solutions est de trouver une fonction \tilde{f} telle que $\tilde{f}(t, y) \geq f(t, y)$ et telle qu'on sache résoudre explicitement l'équation $u'(t) = \tilde{f}(t, u(t))$ (et similairement pour les sous-solutions).

Nous allons montrer maintenant que toute solution qui se trouve sous une sur-solution à t_0 ne pourra jamais la dépasser pour $t > t_0$, et similairement pour les sous-solutions. Par ailleurs notons que le Lemme 3.4.2 de Gronwall en est une conséquence.

Proposition 3.6.7. Soit I un intervalle ouvert de \mathbb{R} , U un ouvert de \mathbb{R} , et soit $(t_0, x_0) \in I \times U$. Soit $f : I \times U \rightarrow \mathbb{R}$ une fonction \mathcal{C}^1 , et soit x une solution sur $] \alpha, \beta [$ ($\alpha, \beta \in \overline{\mathbb{R}}$) du problème de Cauchy avec donnée initiale $x(t_0) = x_0$. Alors :

Si $u :]\alpha, \beta[\rightarrow U$ est une sur-solution et $x(t_0) \leq u(t_0)$ alors $x(t) \leq u(t), \forall t \in [t_0, \beta[$.

Si $v :]\alpha, \beta[\rightarrow U$ est une sous-solution et $x(t_0) \geq v(t_0)$ alors $x(t) \geq v(t), \forall t \in [t_0, \beta[$.

En particulier, si x est piégée à t_0 entre une sous-solution et une sur-solution, elle reste piégée pour tout $t \in [t_0, \beta[$.

Démonstration. On considère l'ensemble non-vide

$$A = \{t \in [t_0, \beta[\mid \forall s \in [t_0, t], x(s) \leq u(s)\},$$

et on note

$$t^+ = \sup A.$$

On veut montrer que $t_+ = \beta$. On suppose par absurde que $t^+ < \beta$. En particulier ceci implique que $t_+ < +\infty$. D'après la définition de la borne supérieure, il existe une suite qui l'approche à valeurs dans l'ensemble : $\exists t_n \in A, t_n \xrightarrow{n \rightarrow \infty} t_+$. Ceci implique $x(t_n) \leq u(t_n)$ et par continuité $x(t_+) \leq u(t_+)$. Si on avait $x(t_+) < u(t_+)$ alors par continuité cette propriété resterait vraie dans un voisinage de t_+ ce qui contredirait le fait que $t_+ = \sup A$. On obtient donc que

$$x(t_+) = u(t_+).$$

Ceci implique en utilisant l'hypothèse :

$$u'(t_+) \geq f(t_+, u(t_+)) = f(t_+, x(t_+)) = x'(t_+) \implies (u - x)'(t_+) \geq 0.$$

Si $(u - x)'(t_+) > 0$, ce qui est le cas si $u'(t_+) > f(t_+, u(t_+))$, par continuité on obtient que $(u - x)'$ est strictement positive dans un voisinage de t_+ , donc localement croissante, donc en particulier $(u - x)(t) > (u - x)(t_+) = 0$ sur un voisinage à droite de t_+ , en contradiction avec le fait que $t_+ = \sup A$. Ainsi, nous obtenons $t_+ = \beta$ donc la conclusion de la proposition dans le cas de toute sur-solution stricte, c'est-à-dire $u'(t) > f(t, u(t)), \forall t \in]\alpha, \beta[$. La preuve est similaire pour le cas des sous-solutions strictes en considérant la borne supérieure de l'ensemble $\{t \in [t_0, \beta[, v(u) \leq x(u), \forall u \in [t_0, t]\}$.

Il nous reste à traiter le cas $(u - x)'(t_+) = 0$. Pour tout $\epsilon \in]0, 1]$ on considère le problème de Cauchy

$$\begin{cases} x'_\epsilon(t) = f(t, x_\epsilon(t)) - \epsilon, \\ x_\epsilon(t_+) = x(t_+). \end{cases} \quad (3.11)$$

La fonction $f_\epsilon(t, y) = f(t, y) - \epsilon$ étant encore de classe \mathcal{C}^1 , nous pouvons appliquer le théorème de Cauchy-Lipschitz pour obtenir l'existence d'une solution maximale x_ϵ définie sur un voisinage de t_+ . Soit $a, b > 0$ tels que $[t_+ - a, t_+ + a] \times [x(t_+) - b, x(t_+) + b] \subset I \times U$. Par continuité de x_ϵ , il existe $\tau(\epsilon) = \tau \in]0, a]$ tel que $[t_+ - \tau, t_+ + \tau]$ est inclus dans l'intervalle de définition de x_ϵ , et que

$$x_\epsilon(t) \in [x(t_+) - b, x(t_+) + b], \quad \forall t \in [t_+, t_+ + \tau].$$

On note également

$$K = \sup \{|f(t, x)| \mid (t, x) \in [t_+ - a, t_+ + a] \times [x(t_+) - b, x(t_+) + b]\}.$$

Il s'ensuit que pour tout $t \in [t_+, t_+ + \tau]$,

$$\begin{aligned} |x(t) - x_\epsilon(t)| &= \left| \int_{t_+}^t f(s, x(s)) - f(s, x_\epsilon(s)) + \epsilon ds \right| \\ &\leq \int_{t_+}^t K |x(s) - x_\epsilon(s)| ds + \epsilon \tau \\ &\leq K \tau \sup_{s \in [t_+, t_+ + \tau]} |x(s) - x_\epsilon(s)| + \epsilon \tau. \end{aligned}$$

En prenant $\tau \leq \frac{1}{2K}$ on a

$$\sup_{s \in [t_+, t_+ + \tau]} |x(s) - x_\epsilon(s)| \leq \frac{\epsilon\tau}{1 - K\tau} \leq 2\epsilon\tau \leq \frac{\epsilon}{K}.$$

Cela implique que

$$x_\epsilon(t) \leq x(t) + \frac{\epsilon}{K}, \quad \forall t \in [t_+, t_+ + \tau]. \quad (3.12)$$

D'une part, par continuité de x en t_+ , il existe $\tau_0 > 0$ tel que

$$x(t) \in [x(t_+) - b/2, x(t_+) + b/2], \quad \forall t \in [t_+ - \tau_0, t_+ + \tau_0],$$

ainsi, si $\epsilon < Kb/4$, et $\tau \leq \tau_0$,

$$x_\epsilon(t) \in]x(t_+) - b, x(t_+) + b[, \quad \forall t \in [t_+ - \tau, t_+ + \tau].$$

par le théorème de prolongement des solutions maximales, on en déduit qu'on peut choisir $\tau \leq \tau_1 := \min(a, 2/K, \tau_0)$: la taille de l'intervalle d'existence de x_ϵ est donc indépendante de ϵ .

D'autre part,

$$x(t) \leq x_\epsilon(t) + \frac{4\epsilon}{K}, \quad \forall t \in [t_+, t_+ + \tau_1], \forall \epsilon \in]0, Kb/4[.$$

D'autre part, u est une sur-solution stricte de l'équation du problème de Cauchy (3.11), avec $u(t_+) = x(t_+) = x_\epsilon(t_+)$, la première partie de la preuve assure alors que

$$x_\epsilon(t) \leq u(t), \quad \forall t \in [t_+, t_+ + \tau], \forall \epsilon \in]0, 1[.$$

À l'aide de (3.12), on obtient donc finalement

$$x(t) \leq u(t) + \frac{4\epsilon}{K}, \quad \forall t \in [t_+, t_+ + \tau_1], \forall \epsilon \in]0, Kb/4[.$$

En faisant tendre ϵ vers zéro on obtient

$$x(t) \leq u(t), \quad \forall t \in [t_+, t_+ + \tau_1].$$

Ceci est en contradiction avec la définition de t_+ , donc $t_+ = \beta$ et

$$x(t) \leq u(t), \quad \forall t \in [t_+, \beta[.$$

La preuve que toute solution qui se trouve au-dessus d'une sous-solution à t_0 ne pourra jamais la croiser pour $t > t_0$ est similaire. \square

Chapitre 4

Exemples d'études avec portraits de phase d'équations non-linéaires

4.1 Le système de Lotka-Volterra

Les équations de Lotka-Volterra, que l'on désigne aussi sous le terme de "modèle proie-prédateur", sont un couple d'équations différentielles non-linéaires du premier ordre, et sont couramment utilisées pour décrire la dynamique de systèmes biologiques dans lesquels un prédateur et sa proie interagissent. Elles ont été proposées indépendamment par Alfred James Lotka en 1925 et Vito Volterra en 1926. Le but de Volterra était de trouver un modèle pour la dynamique des populations de sardines et de requins en mer Adriatique, afin notamment d'expliquer pourquoi les quantités de sardines pêchées après l'interruption due à la guerre n'étaient plus aussi importantes que précédemment et pourquoi à la reprise de la pêche la proportion observée de requins avait augmenté.

Obtention des équations

On considère une population de proies $N(t)$, et une population de prédateurs $P(t)$. les hypothèses de modélisation sont les suivantes :

- Les proies sont supposées avoir une source illimitée de nourriture et se reproduire exponentiellement si elles ne sont soumises à aucune prédation ;
- Le taux de prédation sur les proies est supposé proportionnel à la fréquence de rencontre entre les prédateurs et les proies. Cette fréquence de rencontre est elle-même proportionnelle au produit du nombre de prédateurs par le nombre de proies ;
- En l'absence de proies, la population de prédateurs s'éteint exponentiellement ;
- La vitesse d'accroissement de la population des prédateurs due à la prédation est proportionnelle au taux de prédation.

En effectuant un bilan entre t et $t + dt$ comme au premier chapitre, on obtient les équations suivantes :

$$\begin{cases} \frac{dN}{dt} = aN(t) - bN(t)P(t), \\ \frac{dP}{dt} = cN(t)P(t) - dP(t), \end{cases} \quad (4.1)$$

où a, b, c, d sont des paramètres positifs.

La fonction

$$(x, y) \in \mathbb{R}^2 \mapsto (ax - bxy, cxy - dy) \in \mathbb{R}^2$$

est de classe \mathcal{C}^∞ (donc en particulier \mathcal{C}^1) car ses composantes sont polynomiales en x, y . Soit $(N_0, P_0) \in \mathbb{R}^2$ quelconque. On considère l'équation (4.1) munie des conditions initiales $N(t=0) = N_0, P(t=0) = P_0$. D'après le théorème de Cauchy-Lipschitz 3.1.6, ce problème de Cauchy admet une unique solution maximale (N, P) définie sur un intervalle ouvert I contenant 0.

Remarque 4.1.1. L'intervalle d'existence I est défini de part et d'autre de zéro. Cependant, comme ici on s'intéresse à l'évolution d'un système après avoir prescrit une donnée initiale, on ne considère ici que la famille $\{N(t), P(t)\}_{t \in \mathbb{R}_+ \cap I}$.

Positivité

- Soit $N_1 \in \mathbb{R}$ quelconque. On observe alors que le couple (\bar{N}, \bar{P}) défini par

$$\bar{N}(t) = N_1 e^{at}, \quad \bar{P}(t) = 0 \quad \forall t \in \mathbb{R}$$

est solution de (4.1).

De même, pour $P_1 \in \mathbb{R}$ quelconque, le couple $(\underline{N}, \underline{P})$ défini par

$$\underline{N}(t) = 0, \quad \underline{P}(t) = P_1 e^{-dt} \quad \forall t \in \mathbb{R}$$

est solution de (4.1).

- Soit $N_0 > 0, P_0 > 0$. Considérons à présent la solution maximale du système (4.1) muni de la condition initiale $N(t=0) = N_0, P(t=0) = P_0$. On cherche à montrer que $N(t) > 0$ et $P(t) > 0$ pour tout $t \in I$. On raisonne par l'absurde et on suppose par exemple qu'il existe $t_1 \in I$ tel que $N(t_1) = 0$. On pose $P_1 = P(t_1)e^{dt_1}$. On a alors $(N, P)(t_1) = (\underline{N}, \underline{P})(t_1)$, et donc d'après l'unicité du théorème de Cauchy-Lipschitz, $(N, P)(t) = (\underline{N}, \underline{P})(t)$ pour tout $t \in I$. En particulier, $N_0 = N(0) = \underline{N}(0) = 0$: absurde. Par conséquent, N ne s'annule pas sur I . Comme N est une fonction continue et que I est un intervalle, N garde un signe constant sur I . Donc $N(t) > 0$ pour tout $t \in I$. De même, $P(t) > 0$ pour tout $t \in I$.

“Énergie” du système

On a la propriété suivante, qui permet de mieux comprendre le comportement des solutions et de tracer les trajectoires :

Proposition 4.1.2. *On définit la fonction*

$$\begin{aligned} \mathbb{R}_+^{*2} &\rightarrow \mathbb{R} \\ H : (x, y) &\mapsto by + cx - a \ln y - d \ln x. \end{aligned}$$

Alors pour toute solution (N, P) de (4.1) telle que $N(0) = N_0 > 0, P(0) = P_0 > 0$, on a, pour tout $t \in I$

$$H(N(t), P(t)) = H(N_0, P_0).$$

Démonstration. On dérive par rapport à t la quantité $H(N(t), P(t))$: cette manipulation est autorisée puisque (N, P) est de classe \mathcal{C}^1 et que H est \mathcal{C}^∞ . Notons également que $H(N(t), P(t))$ est bien défini d'après ce qui précède puisque $(N(t), P(t)) \in \mathbb{R}_+^{*2}$.

En utilisant l'équation (4.1), on obtient

$$\begin{aligned} \frac{d}{dt}H(N(t), P(t)) &= b\frac{dP}{dt} + c\frac{dN}{dt} - a\frac{P'(t)}{P(t)} - d\frac{N'(t)}{N(t)} \\ &= bcP(t)N(t) - bdP(t) + caN(t) - bcP(t)N(t) - acN(t) + ad - da + dbP(t) \\ &= 0. \end{aligned}$$

On en déduit que la quantité $H(N(t), P(t))$ reste constante le long de chaque trajectoire. \square

Corollaire 4.1.3. *On suppose $N_0 > 0, P_0 > 0$. Alors il existe x_-, x_+, y_-, y_+ tels que*

$$\begin{aligned} 0 < x_- \leq \frac{d}{c} \leq x_+, \\ 0 < y_- \leq \frac{a}{b} \leq y_+ \end{aligned}$$

et tels que pour tout $t \in I$,

$$x_- \leq N(t) \leq x_+, \quad y_- \leq P(t) \leq y_+.$$

Par conséquent, $I = \mathbb{R}$.

Démonstration. On définit les fonctions $\varphi_{b,a}, \varphi_{c,d}$ par

$$\begin{aligned} \varphi_{b,a} : x \in \mathbb{R}_+^* &\mapsto bx - a \ln x, \\ \varphi_{c,d} : x \in \mathbb{R}_+^* &\mapsto cx - d \ln x, \end{aligned} \tag{4.2}$$

de sorte que

$$H(x, y) = \varphi_{b,a}(y) + \varphi_{c,d}(x) \quad \forall (x, y) \in \mathbb{R}_+^{*2}.$$

Le tableau de variations de $\varphi_{c,d}$ est

x	0	$\frac{d}{c}$	$+\infty$
$\varphi'_{c,d}(x)$	-	0	+
$\varphi_{c,d}(x)$	$+\infty$	$d(1 - \ln \frac{d}{c})$	$+\infty$

Celui de $\varphi_{b,a}$ est analogue.

On en déduit en particulier que

$$\begin{aligned} \varphi_{c,d}(x) &\geq d(1 - \ln \frac{d}{c}) \quad \forall x \in \mathbb{R}_+^*, \\ \varphi_{b,a}(y) &\geq a(1 - \ln \frac{a}{b}) \quad \forall y \in \mathbb{R}_+^*. \end{aligned}$$

Par conséquent, pour tout $t \in I$,

$$\varphi_{c,d}(N(t)) = H(N_0, P_0) - \varphi_{b,a}(P(t)) \leq H(N_0, P_0) - a(1 - \ln \frac{a}{b}).$$

Ainsi

$$\sup_{t \in I} \varphi_{c,d}(N(t)) \leq H(N_0, P_0) - a(1 - \ln \frac{a}{b}),$$

et de même

$$\sup_{t \in I} \varphi_{b,a}(P(t)) \leq H(N_0, P_0) - d(1 - \ln \frac{d}{c}).$$

On déduit des tableaux de variations de $\varphi_{c,d}$ et $\varphi_{b,a}$ qu'il existe x_{\pm}, y_{\pm} tels que

$$\begin{aligned} 0 < x_- \leq \frac{d}{c} \leq x_+, \\ 0 < y_- \leq \frac{a}{b} \leq y_+ \end{aligned}$$

et tels que

$$\varphi_{c,d}(x_{\pm}) = H(N_0, P_0) - a(1 - \ln \frac{a}{b}), \quad \varphi_{b,a}(y_{\pm}) = H(N_0, P_0) - d(1 - \ln \frac{d}{c}). \quad (4.3)$$

Par conséquent,

$$N(t) \in [x_-, x_+], \quad P(t) \in [y_-, y_+] \quad \forall t \in I.$$

Les solutions de (4.1) restent donc bornées sur leur intervalle maximal d'existence. La Proposition 3.2.5 implique alors que $I = \mathbb{R}$. \square

Tracé du portrait de phase

Traçons à présent quelques courbes paramétrées

$$\mathcal{C}_{N_0, P_0} = \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}_+, x = N(t), y = P(t)\},$$

où (N, P) est la solution de (4.1) de donnée initiale (N_0, P_0) .

D'après la proposition 4.1.2, ces courbes paramétrées sont incluses dans des lignes de niveau de la fonction H , c'est-à-dire des courbes d'équation $H(x, y) = \text{constante} = H(N_0, P_0)$. Par ailleurs, les fonctions N et P sont définies pour tout t et ne possèdent pas de symétrie évidente.

Notons \mathcal{D}_{H_0} la courbe d'équation

$$H(x, y) = H_0.$$

On a $\mathcal{C}_{N_0, P_0} \subset \mathcal{D}_{H(N_0, P_0)}$. Il faut donc d'une part tracer les courbes \mathcal{D}_{H_0} pour différentes valeurs de H_0 , et d'autre part montrer que $\mathcal{C}_{N_0, P_0} = \mathcal{D}_{H(N_0, P_0)}$, autrement dit que les solutions $(N(t), P(t))$ parcourent toute la courbe $\mathcal{D}_{H(N_0, P_0)}$.

▷ Tracé des courbes \mathcal{D}_{H_0} :

Tout d'abord, on observe que si

$$H_0 < \inf \varphi_{b,a} + \inf \varphi_{c,d},$$

la courbe \mathcal{D}_{H_0} est vide. Si

$$H_0 = \inf \varphi_{b,a} + \inf \varphi_{c,d} = a \left(1 - \ln \frac{a}{b}\right) + d \left(1 - \ln \frac{d}{c}\right),$$

la courbe est réduite au point $(a/b, d/c)$. Dans toute la suite, on étudie donc le cas où

$$H_0 > \inf \varphi_{b,a} + \inf \varphi_{c,d}.$$

On trace alors la courbe \mathcal{D}_{H_0} en séparant ses deux branches $y \in [y_-, a/b]$ et $y \in [a/b, y_+]$, où y_-, y_+ sont définis par (4.3).

On observe alors que $\varphi_{c,d}$ (resp. $\varphi_{b,a}$) est une bijection de $]0, d/c[$ sur $[d(1 - \ln \frac{d}{c}), +\infty[$ (resp. de $]0, a/b[$ sur $[a(1 - \ln \frac{a}{b}), +\infty[$), et également de $[d/c, +\infty[$ sur $[d(1 - \ln \frac{d}{c}), +\infty[$ (resp. de $[a/b, +\infty[$ sur $[a(1 - \ln \frac{a}{b}), +\infty[$). On définit les bijections réciproques ψ_-, ψ_+ de $\varphi_{b,a}$ restreinte aux intervalles $]0, a/b[$ et $[a/b, +\infty[$. Les fonctions ψ_-, ψ_+ sont toutes deux définies sur $[a(1 - \ln \frac{a}{b}), +\infty[$. Si $y \in [y_-, a/b]$, on a

$$(x, y) \in \mathcal{D}_{H_0} \Leftrightarrow y = \psi_-(H_0 - \varphi_{c,d}(x)),$$

et pour $y \in [a/b, y_+]$, on a

$$(x, y) \in \mathcal{D}_{H_0} \Leftrightarrow y = \psi_+(H_0 - \varphi_{c,d}(x)).$$

On est donc ramené à tracer les graphes des fonctions

$$f_{\pm} : x \mapsto \psi_{\pm}(H_0 - \varphi_{c,d}(x))$$

sur l'intervalle $[x_-, x_+]$. Notons que si $x \in [x_-, x_+]$, par définition de x_{\pm} , on a

$$\varphi_{c,d}(x) \leq H_0 - a \left(1 - \ln \frac{a}{b}\right)$$

et donc les fonctions f_+, f_- sont bien définies sur $[x_-, x_+]$.

La fonction ψ_- (resp. ψ_+) est décroissante (resp. croissante) sur $[a(1 - \ln a/b), +\infty[$, dérivable sur $]a(1 - \ln a/b), +\infty[$, et sa dérivée est

$$\psi'_{\pm} = \frac{1}{\varphi_{b,a} \circ \psi_{\pm}}.$$

Par conséquent les fonctions f_+ et f_- sont dérivables sur $]x_-, x_+[$, de dérivée

$$f'_{\pm}(x) = -\varphi'_{c,d}(x) \frac{1}{\varphi_{b,a} \circ f_{\pm}(x)}.$$

On observe que

$$\lim_{x \rightarrow x_-} f'_{\pm}(x) = \pm\infty, \quad \lim_{x \rightarrow x_+} f'_{\pm}(x) = \mp\infty.$$

On en déduit les tableaux de variations suivants :

x	x_-	$\frac{d}{c}$	x_+
$f'_-(x)$	-	0	+
$f_-(x)$	$\frac{a}{b}$	y_-	$\frac{a}{b}$
$f'_+(x)$	+	0	-
$f_+(x)$	$\frac{a}{b}$	y_+	$\frac{a}{b}$

La courbe possède donc des tangentes horizontales aux points $(d/c, y_-)$ et $(d/c, y_+)$ et des tangentes verticales aux points $(x_-, a/b)$ et $(x_+, a/b)$.

Enfin, on observe que y_- et x_- sont décroissants en H_0 , tandis que y_+ et x_+ sont croissants en H_0 . En considérant plusieurs valeurs de H_0 , on obtient donc une série de courbes qui ont l'allure représentée sur la figure 4.1. En particulier, on observe que les courbes d'équation $H(x, y) = H_0$ sont des **courbes fermées**.

▷ Description des trajectoires $(N(t), P(t))$ pour $t \geq 0$:

Remarquons tout d'abord que si $(N_0, P_0) = (d/c, a/b)$, alors on a

$$H(N_0, P_0) = \min_{\mathbb{R}_+^* \times \mathbb{R}_+^*} H = d(1 - \ln \frac{d}{c}) + a(1 - \ln \frac{a}{b}).$$

En particulier

$$H(N(t), P(t)) = \min_{\mathbb{R}_+^* \times \mathbb{R}_+^*} H \quad \forall t \in \mathbb{R},$$

et donc $(N(t), P(t)) = (N_0, P_0)$ pour tout t : le point $(d/c, a/b)$ est un point d'équilibre du système, et la courbe \mathcal{C}_{N_0, P_0} est un point.

Dans toute la suite, on suppose donc $N_0 > 0, P_0 > 0$, et $(N_0, P_0) \neq (d/c, a/b)$. On sait déjà que pour tout $t \in \mathbb{R}$, $(N(t), P(t)) \in \mathcal{D}_{H_0}$ avec $H_0 = H(N_0, P_0)$. La première question est de savoir dans quel sens la courbe va être parcourue. Or comme les fonctions N et P sont définies par une équation différentielle, il est assez facile de déterminer les signes de $N'(t)$ et $P'(t)$. On a

$$\begin{aligned} N'(t) > 0 &\iff P(t) < \frac{a}{b}, \\ P'(t) > 0 &\iff N(t) > \frac{d}{c}. \end{aligned}$$

On en déduit immédiatement que la courbe \mathcal{D}_{H_0} est parcourue dans le sens trigonométrique. En particulier, si on repère la position $(N(t), P(t))$ en fonction de son abscisse curviligne $s(t)$ sur la courbe \mathcal{D}_{H_0} , on voit que s est une fonction croissante du temps (en augmentant s de la longueur de la courbe à chaque fois que l'on fait un tour complet).

Il y a donc deux possibilités : soit s admet une limite finie en $+\infty$, et en ce cas il existe $(\bar{N}, \bar{P}) \in \mathcal{D}_{H_0}$ tel que

$$\lim_{t \rightarrow \infty} (N(t), P(t)) = (\bar{N}, \bar{P}),$$

soit $\lim_{t \rightarrow \infty} s(t) = +\infty$, et dans ce cas il existe $T > 0$ tel que $s(T) = s(0) + L$, c'est-à-dire $(N(T), P(T)) = (N_0, P_0)$.

Supposons que $(N_0, P_0) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \setminus \{(d/c, a/b)\}$. On va à présent montrer que la première possibilité ne peut pas se produire. Raisonnons par l'absurde et supposons que

$$\lim_{t \rightarrow \infty} (N(t), P(t)) = (\bar{N}, \bar{P}),$$

Alors nécessairement, on a aussi

$$\lim_{t \rightarrow \infty} (N'(t), P'(t)) = (a\bar{N} - b\bar{N}\bar{P}, c\bar{N}\bar{P} - d\bar{P}) = (c_1, c_2).$$

Si $(c_1, c_2) \neq (0, 0)$, alors on a nécessairement $\lim_{t \rightarrow \infty} |N(t)| = +\infty$ ou $\lim_{t \rightarrow \infty} |P(t)| = +\infty$: absurde. Donc $c_1 = c_2 = 0$, ce qui implique

$$(\bar{N}, \bar{P}) \in \left\{ (0, 0), \left(\frac{d}{c}, \frac{a}{b} \right) \right\}.$$

Or par conservation de l'énergie et continuité de H , on a aussi $H(\bar{N}, \bar{P}) = H(N_0, P_0) > \min H = H\left(\frac{d}{c}, \frac{a}{b}\right)$ et $H(\bar{N}, \bar{P}) < +\infty$: on aboutit à une contradiction.

On en déduit que si $(N_0, P_0) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \setminus \left\{\left(\frac{d}{c}, \frac{a}{b}\right)\right\}$, alors il existe $T > 0$ tel que $(N(T), P(T)) = (N_0, P_0)$. D'après la proposition 3.3.4, les trajectoires sont périodiques de période T .

Remarque 4.1.4. La propriété importante utilisée ici est que si les trajectoires sont portées par des courbes fermées, et que la vitesse ne s'annule pas, alors le mouvement est périodique.

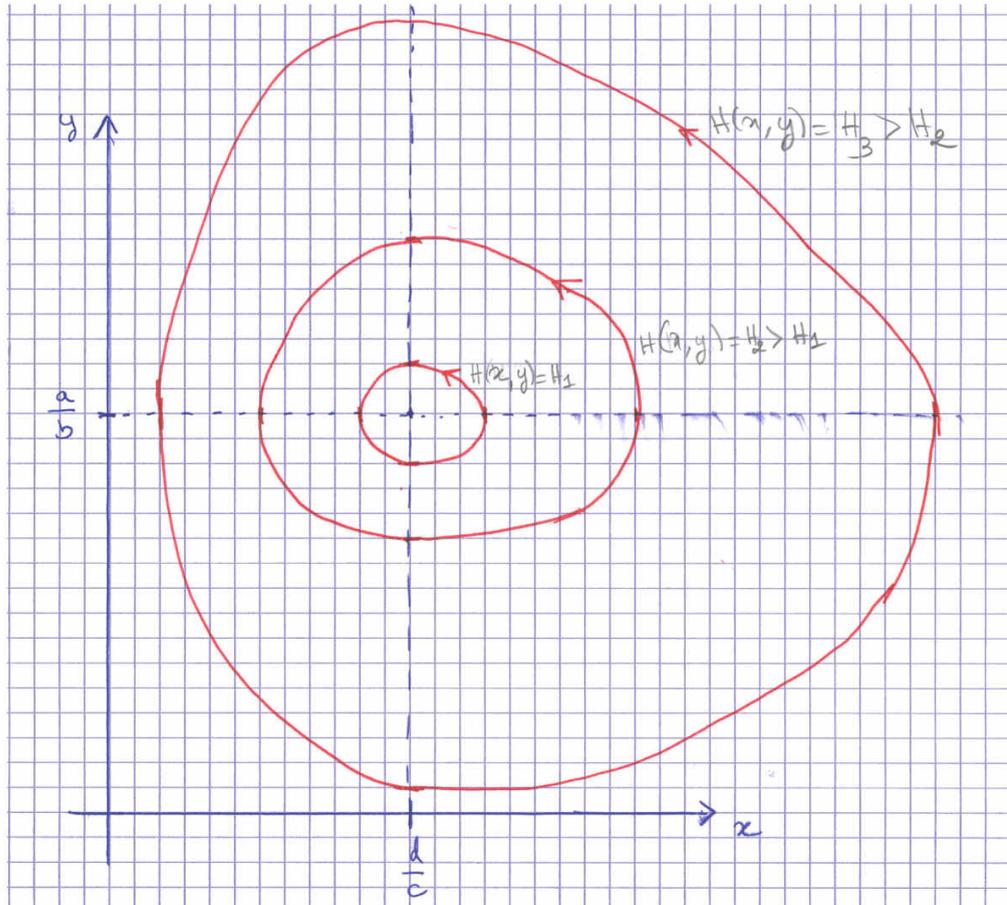


FIGURE 4.1 – Tracé de quelques trajectoires du système de Lotka-Volterra

4.2 Le pendule simple

On considère les solutions de l'équation (1.2)

$$\theta'' + \frac{g}{L} \sin \theta = 0. \quad (4.4)$$

Pour simplifier les calculs, on prendra dans tout ce paragraphe $\omega_0 = \sqrt{g/L} = 1$. On cherche à représenter, pour différentes valeurs des paramètres θ_0, v_0 , les courbes paramétrées

$$\mathcal{C}_{\theta_0, v_0} := \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}, x = \theta_{\theta_0, v_0}(t), y = \theta'_{\theta_0, v_0}(t)\},$$

où θ_{θ_0, v_0} est la solution de l'équation (1.2) telle que $\theta(0) = \theta_0, \theta'(0) = v_0$.

Dans la suite, pour alléger les notations, on omet les indices θ_0, v_0 ; il faut cependant bien garder en tête que les trajectoires $\{\theta(t)\}_{t \geq 0}$ dépendent des données initiales.

▷ *énergie du système* :

En multipliant (4.4) par $\theta'(t)$, on obtient

$$\theta'(t)\theta''(t) + \theta'(t) \sin \theta(t) = 0, \quad \forall t \in \mathbb{R},$$

soit

$$\frac{d}{dt} \left(\frac{(\theta'(t))^2}{2} - \cos \theta(t) \right) = 0.$$

On en déduit que l'énergie est conservée au cours du mouvement :

$$\frac{(\theta'(t))^2}{2} - \cos \theta(t) = \text{cste} = \frac{v_0^2}{2} - \cos \theta_0 \quad \forall t.$$

On pose

$$E_0 = \frac{v_0^2}{2} - \cos \theta_0.$$

Notons que $E_0 \geq -\cos \theta_0 \geq -1$. La courbe paramétrée $\mathcal{C}_{\theta_0, v_0}$ est portée par la courbe d'équation $y^2/2 - \cos x = E_0$. Chacune de ces courbes est symétrique par rapport à l'axe des abscisses et par rapport à celui des ordonnées. Il suffit donc de tracer les courbes sur le domaine $x \geq 0, y \geq 0$ et d'en déduire l'ensemble de la courbe par symétrie.

Ce cas est en fait un peu plus simple que celui du système de Lotka-Volterra car sur le domaine $x \geq 0, y \geq 0$, l'équation de la courbe relative à l'énergie E_0 est

$$y = \sqrt{2(E_0 + \cos x)}. \quad (4.5)$$

On a donc une équation cartésienne explicite pour chaque courbe. On va commencer par tracer les courbes d'équation (4.5) en distinguant les cas $E_0 < 1, E_0 = 1$ et $E_0 > 1$. On verra ensuite comment comprendre la trajectoire suivie par le pendule sur chacune des courbes.

Soit ϕ_{E_0} la fonction $\phi_{E_0} : x \in \mathbb{R} \mapsto \sqrt{2(E_0 + \cos x)}$.

La fonction ϕ_{E_0} est définie sur \mathbb{R} si $E_0 \geq 1$, et sur $[-\arccos(-E_0), \arccos(-E_0)]$ si $E_0 < 1$. Elle est paire, continue sur son ensemble de définition et dérivable sur \mathbb{R} si $E_0 > 1$, sur $\mathbb{R} \setminus \{\pi + 2k\pi, k \in \mathbb{Z}\}$ si $E_0 = 1$, et sur $] -\arccos(-E_0), \arccos(-E_0)[$ si $E_0 < 1$. Sur l'ensemble de dérivabilité de ϕ_{E_0} , on a

$$\phi'_{E_0}(x) = \frac{-\sin x}{\sqrt{2(E_0 + \cos x)}}.$$

On a donc deux types de tableaux de variations :

— Premier cas : $E_0 \geq 1$: dans ce cas la fonction ϕ_{E_0} est définie sur \mathbb{R} et périodique de période 2π , donc il suffit de tracer son tableau de variations sur l'intervalle $[-\pi, \pi]$:

x	$-\pi$	0	π
$\phi'_{E_0}(x)$	0	+	0
ϕ_{E_0}	$\sqrt{2(E_0 - 1)}$	$\sqrt{2(E_0 + 1)}$	$\sqrt{2(E_0 - 1)}$

Si $E_0 > 1$, la courbe admet des tangentes horizontales en $x = k\pi, k \in \mathbb{Z}$.

Si $E_0 = 1$, la courbe admet une tangente horizontale en $x = 2k\pi, k \in \mathbb{Z}$. Calculons la pente de la tangente en $x = \pi^\pm$: on pose $x = \pi + h$ avec $|h| \ll 1$ et $h \neq 0$. Alors

$$\begin{aligned}\sin x &= \sin(\pi + h) = -\sin h = -h + O(h^3), \\ 1 + \cos x &= 1 + \cos(\pi + h) = 1 - \cos h = \frac{h^2}{2} + O(h^4).\end{aligned}$$

On en déduit que

$$\phi_1'(x) \sim \frac{h}{|h|}.$$

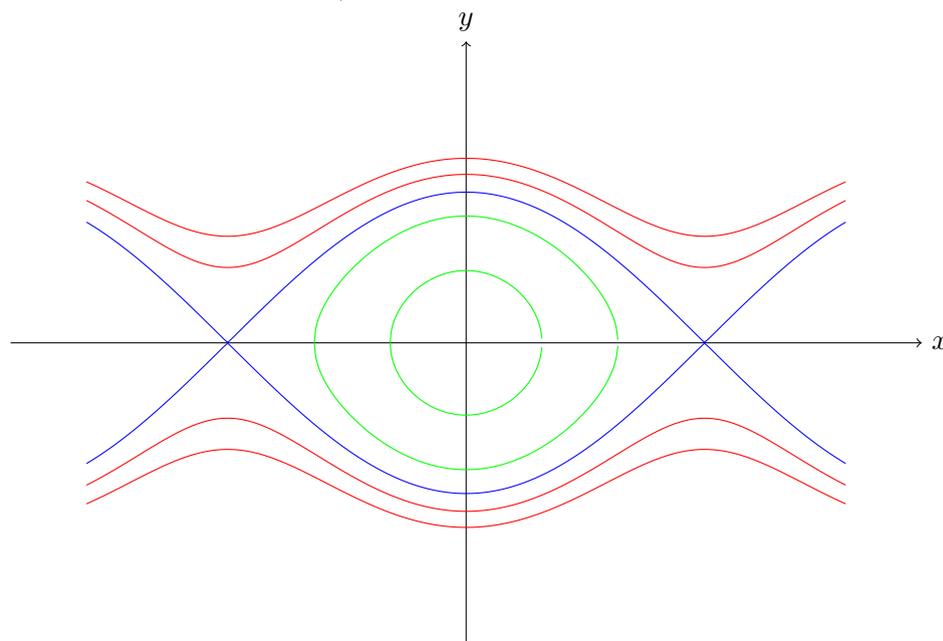
Ainsi la courbe représentative de ϕ_1 admet une tangente à gauche de pente -1 et une tangente à droite de pente 1 en $x = \pi$.

— Deuxième cas : $E_0 < 1$: on a alors le tableau de variations suivant, où l'on a posé $x_0 = \arccos(-E_0)$

x	$-x_0$	0	x_0
$\phi_{E_0}'(x)$	+	0	-
ϕ_{E_0}	0	$\sqrt{2(E_0 + 1)}$	0

La courbe admet une tangente horizontale en $x = k\pi$ et deux tangentes verticales en $x = \pm \arccos(-E_0)$.

Les courbes d'équation $y^2/2 - \cos x = E_0$ ont donc l'allure suivante :



En vert : deux courbes d'équation $y^2/2 - \cos x = E_0$ avec $E_0 < 1$. En bleu, la "séparatrice", avec $E_0 = 1$. En rouge, deux courbes avec $E_0 > 1$.

À présent que les courbes des trajectoires sont connues, voyons dans quel sens celles-ci sont parcourues :

▷ *Premier cas* : $E_0 < 1$:

Ce cas correspond à un pendule qui oscille (car lancé avec une énergie inférieure à l'énergie nécessaire pour faire un tour complet).

Si $E_0 < 1$, on a

$$-\cos \theta(t) = E_0 - \frac{(\theta'(t))^2}{2} \leq E_0,$$

et donc $\cos \theta(t) \geq -E_0$ pour tout t . Ainsi $\theta(t) \in [-\arccos(-E_0), \arccos(-E_0)] \pmod{2\pi}$ pour tout t . Comme par ailleurs la fonction θ est continue (car dérivable), et que $\theta_0 \in [-\pi, \pi]$, on en déduit que $\theta(t) \in [-\arccos(-E_0), \arccos(-E_0)]$ pour tout t .

On pose par ailleurs $v_{max} = \sqrt{2(E_0 + 1)}$. Notons qu'on a $|\theta'(t)| \leq v_{max}$ pour tout t .

On peut alors mener le même type d'analyse que pour le système de Lotka-Volterra. De nouveau, pour $E_0 = -1$, on a $\theta(t) = 0$ pour tout t : la position $\theta = 0$ est un point d'équilibre. Si $E_0 \in]-1, 1[$, la courbe d'équation

$$\frac{y^2}{2} - \cos x = E_0$$

est une courbe fermée. On peut définir l'abscisse curviligne s le long de la courbe, que l'on oriente par exemple dans le sens trigonométrique. On montre alors que s est strictement décroissante en t . En appliquant le même raisonnement que pour le système de Lotka-Volterra, on en déduit que les trajectoires sont périodiques.

Les courbes du portrait de phase correspondant à une énergie $E_0 < 1$ sont donc parcourues dans le sens des aiguilles d'une montre. On voit que pour chacune des trajectoires, le pendule parcourt toute la courbe d'équation $y^2/2 - \cos x = E_0$.

Les points $\theta = \pm \arccos(-E_0)$ correspondent aux extrémités des oscillations du pendule. Les points en lesquels $\theta = 0$ et $\theta' = \pm v_{max}$, correspondent aux passages du pendule par la position verticale. Une oscillation complète correspond au parcours de l'intégralité d'une courbe d'équation $y^2/2 - \cos x = E_0$.

▷ *Deuxième cas* : $E_0 > 1$

Dans ce cas θ' ne peut pas s'annuler (sinon on aurait $-\cos \theta = E_0 > 1$: impossible). Comme θ'' existe et est continue d'après l'équation (1.2), θ' est également continue. Donc θ' garde un signe constant, celui de v_0 . Ainsi

$$\theta'(t) = \operatorname{sgn}(v_0) \sqrt{2(E_0 + \cos \theta(t))}.$$

Par conséquent, pour une trajectoire donnée, on reste uniquement sur une branche de la courbe d'équation $y^2/2 - \cos x = E_0$. La branche de la courbe qui correspond au signe opposé de v_0 n'est jamais visitée (évidemment, elle le sera pour d'autres données initiales!)

De plus θ est strictement monotone (croissante si $v_0 < 0$, décroissante sinon). Le pendule parcourt uniquement une demi-branche de la courbe d'équation $y^2/2 - \cos x = E_0$.

Par ailleurs, on rappelle que les courbes d'équation

$$y = \pm \sqrt{2(E_0 + \cos x)}$$

sont périodiques de période 2π . Supposons que le pendule parte d'un angle θ_0 avec une vitesse $v_0 > 0$ telle que $E_0 > 1$. Alors θ est strictement croissante au cours du temps et on lit sur le portrait de phase qu'il existe un temps $T > 0$ tel que

$$\theta(T) = \theta_0 + 2\pi, \quad \theta'(T) = v_0.$$

Cela signifie que le pendule a fait un tour complet et est revenu à sa position initiale. Là encore, le mouvement est périodique, mais le pendule ne s'arrête jamais.

Les points de la courbe de coordonnées $x = 2k\pi, y = \sqrt{2(E_0 + 1)}$ correspondent aux passages du pendule par la position verticale en bas, et ceux de coordonnées $x = \pi + 2k\pi, y = \sqrt{2(E_0 - 1)}$ aux passages du pendule par la position verticale en haut.

▷ *Troisième cas* : $E_0 = 1$

On suppose dans tout ce paragraphe que $(\theta_0, v_0) \notin \{(\pi, 0), (-\pi, 0)\}$. En effet, si $(\theta_0, v_0) = (\pm\pi, 0)$, alors on a $\theta(t) = \pm\pi$ pour tout t : le point $\theta = \pi$ est une position d'équilibre. On se concentre donc sur les cas où $v_0^2/2 - \cos\theta_0 = 1$ avec $v_0 \neq 0$.

Ce dernier cas est un peu particulier : le pendule a tout juste assez d'énergie pour arriver au sommet. La question est de savoir s'il parcourt l'ensemble de la courbe, c'est à dire si la bille redescend après être passée par le maximum.

La réponse est non : le pendule ne parcourt qu'une branche de la courbe. En effet, on peut montrer qu'il met un temps infini à arriver à la position d'équilibre (instable) $\theta = \pi$, qui correspond au point d'altitude maximal de la bille. Pour voir cela, il y a deux preuves possibles :

- La première possibilité est de raisonner par l'absurde et de supposer qu'il existe un temps $T > 0$ tel que

$$\theta(T) = \pm\pi, \quad \theta'(T) = 0.$$

On vérifie alors aisément que $\tilde{\theta}(t) = \text{constante} = \pm\pi$ est solution de l'équation (1.2), avec $\tilde{\theta}(T) = \theta(T)$, $\tilde{\theta}'(T) = \theta'(T)$. D'après le théorème de Cauchy-Lipschitz, on en déduit que $\tilde{\theta} = \theta$, et donc en particulier $\theta(0) = \pm\pi$: absurde.

- L'autre possibilité est de calculer à la main le temps mis par le pendule pour parvenir au sommet. Pour cela, on peut utiliser le fait que θ' ne s'annule pas avant d'atteindre le sommet et donc garde un signe constant sur cet intervalle. Supposons pour fixer les idées que $v_0 > 0$. Soit $T_0 \in]0, +\infty[$ le temps auquel le sommet est atteint. On en déduit que $\theta|_{]0, T_0[} :]0, T_0[\rightarrow]\theta_0, \pi[$ est une bijection. D'après la formule de changement de variables, on a donc

$$\begin{aligned} T_0 &= \int_0^{T_0} dt \\ &= \int_{\theta_0}^{\pi} \frac{d\theta}{\theta'} \\ &= \int_{\theta_0}^{\pi} \frac{d\theta}{\sqrt{2(1 + \cos\theta)}}. \end{aligned}$$

Or pour $\theta = \pi - h$ avec $h > 0$, $h \ll 1$, on a

$$1 + \cos\theta = 1 + \cos(\pi - h) = \frac{h^2}{2} + O(h^4),$$

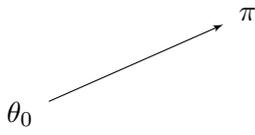
et donc

$$\frac{1}{\sqrt{2(1 + \cos\theta)}} \sim \frac{1}{\pi - \theta},$$

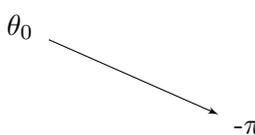
qui n'est pas intégrable au voisinage de $\theta = \pi$. Donc $T_0 = \infty$.

On en déduit que le pendule met un temps infini à atteindre la position d'équilibre $\theta = \pm\pi$. Par conséquent θ' ne s'annule pas sur \mathbb{R} et donc θ est strictement monotone (croissante si $v_0 > 0$, décroissante si $v_0 < 0$).

On obtient ainsi le tableau de variations suivant si $v_0 > 0$:

t	0	$+\infty$
$\theta'(t)$	+	
$\theta(t)$		

Si $v_0 < 0$, le tableau de variations est analogue :

t	0	$+\infty$
$\theta'(t)$	-	
$\theta(t)$		

Appendices

Chapitre 5

Deux types particuliers d'équations différentielles

5.1 Recollements

Toutes les équations différentielles étudiées jusqu'ici étaient de la forme

$$X'(t) = f(t, X(t)).$$

C'est le cadre pour lequel on dispose de théorèmes d'existence et d'unicité généraux.

Néanmoins, il se peut qu'on rencontre des équations différentielles qui ne puissent être mises sous cette forme de façon immédiate. L'exemple le plus simple est celui d'une équation différentielle de la forme

$$g(t)X'(t) = h(t, X(t)),$$

où la fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ est de classe \mathcal{C}^1 , et $h : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ est aussi de classe \mathcal{C}^1 , mais f peut s'annuler sur \mathbb{R} . Dans ce cas il n'y a *pas de théorie générale* assurant l'existence ou l'unicité de solutions. La stratégie est alors la suivante :

- On identifie les points d'annulation de g , et, par suite, les intervalles sur lesquels g ne s'annule pas ;
- Sur chacun des intervalles sur lequel g ne s'annule pas, on peut appliquer le théorème de Cauchy-Lipschitz. On cherche alors l'ensemble des solutions de l'équation différentielle sur cet intervalle (pour n'importe quelle valeur de la donnée initiale). Attention, pour pouvoir appliquer le point suivant, il faut que ces solutions soient définies (et bornées) dans un voisinage du point auquel on souhaite raccorder ;
- Il faut ensuite raccorder, ou recoller : par exemple, si f s'annule en $t_0 \in \mathbb{R}$, et que l'on a trouvé des solutions sur $] - \infty, t_0[$ et sur $]t_0, +\infty[$, il faut construire une fonction continue (et même \mathcal{C}^1 , si possible) X telle que

$$X(t) = \begin{cases} X_-(t) & \text{si } t < t_0, \\ X_+(t) & \text{si } t > t_0, \end{cases}$$

où X_- et X_+ sont des solutions sur $] - \infty, t_0[$ et sur $]t_0, +\infty[$ respectivement.

Suivant l'allure des fonctions X_- et X_+ , parfois il n'y aura aucune solution continue de l'équation, parfois il y en aura une unique, et parfois une infinité : tous les cas de figure sont possibles. Il convient également de remarquer que les problèmes de recollement ne sont pas associés à une condition initiale en général : en quelque sorte, le fait de devoir recoller en un point joue le rôle de condition initiale. On donne ci-dessous quelques exemples.

Exemple 1

On considère l'équation

$$(1 - t^2)u'(t) - 2tu(t) = 1.$$

La fonction $t \mapsto 1 - t^2$ s'annule en -1 et en 1 . On résout donc l'équation sur les intervalles $I_1 =] - \infty, -1[$, $I_2 =] - 1, 1[$, $I_3 =]1, +\infty[$.

Sur chacun des intervalles I_j , $j = \{1, 2, 3\}$, on a

$$u'(t) - \frac{2t}{1 - t^2}u(t) = \frac{1}{1 - t^2}.$$

Les solutions générales de l'équation homogène associée sont de la forme

$$\bar{u}(t) = C e^{\int \dots \frac{2s}{1-s^2} ds} = \frac{\tilde{C}}{|1 - t^2|}.$$

Pour trouver la solution générale de l'équation avec second membre, on applique la méthode de variation de la constante. On cherche des solutions sous la forme

$$u(t) = \frac{C(t)}{|1 - t^2|}.$$

Il vient alors

$$\frac{C'(t)}{|1 - t^2|} = \frac{1}{1 - t^2},$$

et donc $C(t) = t \operatorname{sgn}(1 - t^2) + C_j$, puisque $\operatorname{sgn}(1 - t^2)$ est constant sur les intervalles I_j .

Bilan : sur chacun des intervalles I_j , la solution est de la forme

$$u_j(t) = \frac{t \operatorname{sgn}(1 - t^2) + C_j}{|1 - t^2|} = \frac{t \operatorname{sgn}(1 - t^2) + C_j}{|(1 - t)(1 + t)|}.$$

On cherche maintenant à raccorder aux points -1 et 1 . Pour cela, il faut que les fonctions u_1 et u_2 admettent une limite en -1 , et que u_2 et u_3 admettent une limite en 1 . Or :

- u_1 admet une limite finie en -1^- si et seulement si $C_1 = -1$;
- u_2 admet une limite finie en -1^+ si et seulement si $C_2 = 1$;
- u_2 admet une limite finie en 1^- si et seulement si $C_2 = 1$;
- u_3 admet une limite finie en 1^+ si et seulement si $C_3 = -1$.

On voit donc que u_2 ne peut pas avoir simultanément des limites finies en -1^+ et en 1^- . Le problème n'a donc pas de solution globale, car **on ne peut pas recoller sur \mathbb{R} entier**. Par contre notons que l'on obtient un recollement continu sur $] - \infty, 1[$ (et similairement sur $] - 1, \infty[$), qui est, en prenant donc $C_1 = -1$ et $C_2 = 1$, la fonction continue $u(t) = \frac{1}{1-t}$. Cette fonction est aussi de classe \mathcal{C}^1 , donc solution de l'équation sur $] - \infty, 1[$.

Exemple 2

On considère une variante de l'équation "explosive",

$$tu'(t) = u(t)^2, \quad t \in \mathbb{R}.$$

Comme la fonction $t \mapsto t$ s'annule en zéro uniquement, on résout cette équation sur \mathbb{R}_\pm^* .

Tout d'abord, sur chacun de ces intervalles, on remarque que la fonction nulle est solution. On cherche à présent les solutions qui ne sont pas identiquement nulles. Soit u_{\pm} une solution de l'équation

$$u'(t) = \frac{1}{t}u(t)^2$$

sur \mathbb{R}_{\pm}^* , définie sur un intervalle $I_{\pm} \subset \mathbb{R}_{\pm}^*$, et soit $t_{\pm} \in I_{\pm}$ tel que $u(t_{\pm}) \neq 0$. On pose $a_{\pm} = u_{\pm}(t_{\pm})$. D'après la Proposition 3.3.2, u_{\pm} ne s'annule pas et garde un signe constant sur I_{\pm} . On a donc

$$\frac{u'_{\pm}(t)}{u_{\pm}(t)^2} = \frac{1}{t} \quad \forall t \in I_{\pm}.$$

On en déduit que

$$\frac{1}{a_{\pm}} - \frac{1}{u_{\pm}(t)} = \ln \left| \frac{t}{t_{\pm}} \right|,$$

et donc

$$u_{\pm}(t) = \frac{1}{\frac{1}{a_{\pm}} - \ln \left| \frac{t}{t_{\pm}} \right|}.$$

La fonction u_{\pm} est donc définie tant que la quantité

$$\frac{1}{a_{\pm}} - \ln \left| \frac{t}{t_{\pm}} \right|$$

reste non nulle et du même signe que a_{\pm} , autrement dit tant que

$$a_{\pm} \left(\frac{1}{a_{\pm}} - \ln \left| \frac{t}{t_{\pm}} \right| \right) = 1 - a_{\pm} \ln \left| \frac{t}{t_{\pm}} \right| > 0.$$

Par conséquent, si on veut que u_{\pm} soit définie au voisinage de zéro, il faut que $a_{\pm} > 0$ nécessairement. Dans ce cas l'intervalle de définition de u_{\pm} est

$$I_+ =]0, t_+ e^{\frac{1}{a_+}}[, \quad I_- =]t_- e^{\frac{1}{a_-}}, 0[.$$

On vérifie de surcroît que les fonctions u_{\pm} ainsi définies ont toutes pour limite zéro quand $t \rightarrow 0^{\pm}$. On peut donc raccorder par continuité quelles que soient les valeurs de t_{\pm}, a_{\pm} . Autrement dit, la fonction u définie par

$$u(t) = \begin{cases} \frac{1}{\frac{1}{a_+} - \ln \left(\frac{t}{t_+} \right)} & \text{si } t > 0, \\ \frac{1}{\frac{1}{a_-} - \ln \left(\frac{t}{t_-} \right)} & \text{si } t < 0, \end{cases}$$

avec $a_+, a_-, t_+ > 0, t_- < 0$ quelconques, se prolonge par continuité en zéro en posant $u(0) = 0$, et est solution de l'équation différentielle sur \mathbb{R}_+^* et sur \mathbb{R}_-^* .

Néanmoins le raccord ne sera jamais de classe \mathcal{C}^1 : en effet la fonction u_{\pm} a pour dérivée

$$u'_{\pm}(t) = \frac{1}{t \left(\frac{1}{a_{\pm}} - \ln \left| \frac{t}{t_{\pm}} \right| \right)^2} \quad \forall t \in I_{\pm},$$

qui a pour limite $\pm\infty$ quand $t \rightarrow 0^{\pm}$. Le raccord n'est donc pas \mathcal{C}^1 (on peut montrer qu'il n'est pas non plus dérivable en zéro).

Exemple 3

On considère l'équation

$$|t|u'(t) - u(t) = t^2.$$

Comme pour les exemples précédents, on commence par résoudre cette équation sur \mathbb{R}_-^* et sur \mathbb{R}_+^* .

Sur $I = \mathbb{R}_-^*$ ou \mathbb{R}_+^* , on a

$$u'(t) - \frac{1}{|t|}u(t) = |t|.$$

La solution générale de l'équation homogène est $t \mapsto C_+t$ sur \mathbb{R}_+^* , et $t \mapsto \frac{C_-}{t}$ sur \mathbb{R}_-^* . On applique ensuite la méthode de variation de la constante : on cherche u_+, u_- sous la forme

$$u_+(t) = C_+(t)t, \quad u_-(t) = \frac{C_-(t)}{t}.$$

On obtient

$$C'_+(t) = 1, \quad C'_-(t) = -t^2,$$

d'où

$$C_+(t) = t + c_+, \quad C_-(t) = -\frac{t^3}{3} + c_-,$$

avec $c_+, c_- \in \mathbb{R}$. Finalement, les solutions sur \mathbb{R}_-^* et \mathbb{R}_+^* sont de la forme

$$\begin{aligned} u_-(t) &= -\frac{t^2}{3} + \frac{c_-}{t}, & t < 0, \\ u_+(t) &= t^2 + c_+t, & t > 0. \end{aligned}$$

Par conséquent, u_- et u_+ ont des limites finies en zéro si et seulement si $c_- = 0$. Dans ce cas, la fonction définie par

$$u(t) = \begin{cases} -\frac{t^2}{3} & \text{si } t \leq 0, \\ t^2 + c_+t & \text{si } t > 0, \end{cases}$$

est continue sur \mathbb{R} et solution de l'équation différentielle sur \mathbb{R}_+^* et sur \mathbb{R}_-^* pour toutes les valeurs de $c_+ \in \mathbb{R}$.

Il reste à voir la question de la dérivabilité en zéro. On a

$$\begin{aligned} \lim_{t \rightarrow 0^-} u'(t) &= 0, \\ \lim_{t \rightarrow 0^+} u'(t) &= c_+. \end{aligned}$$

D'après le théorème de la limite de la dérivée, u admet des demi-dérivées à gauche et à droite en zéro. Elle est dérivable si et seulement si $c_+ = 0$. Dans ce cas, la fonction est de classe \mathcal{C}^1 sur \mathbb{R} .

Conclusion : l'unique solution de classe \mathcal{C}^1 sur \mathbb{R} de l'équation $|t|u'(t) - u(t) = t^2$ est

$$u(t) = \begin{cases} -\frac{t^2}{3} & \text{si } t \leq 0, \\ t^2 & \text{si } t > 0, \end{cases}$$

5.2 Équations à variables séparées

Le but est de résoudre des équations du type

$$y'(t) = g(y(t))f(t), \quad (5.1)$$

où $f, g : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions de classe \mathcal{C}^1 . On munit cette équation d'une donnée initiale $y(t_0) = y_0$.

On va commencer par faire un raisonnement manquant un peu de rigueur, mais qui donne l'intuition du résultat, puis on démontrera ce résultat rigoureusement à l'aide des théorèmes des paragraphes précédents.

On remarque immédiatement que si $g(y_0) = 0$, alors la fonction constante $t \mapsto y_0$ est solution de l'équation (5.1). D'après le théorème de Cauchy-Lipschitz et ses corollaires, c'est l'unique solution de donnée initiale $y(t_0) = y_0$.

On suppose donc à présent que $g(y_0) \neq 0$. Alors si y est une solution de (5.1), comme la fonction $g \circ y$ est continue (puisqu'elle est de classe \mathcal{C}^1), elle ne s'annule pas sur un voisinage de t_0 . Tant que $g \circ y$ ne s'annule pas, on a

$$\frac{y'(t)}{g(y(t))} = f(t).$$

Soit F une primitive de f , et soit H une primitive de $1/g$ sur un intervalle contenant y_0 et sur lequel g ne s'annule pas. On obtient

$$H(y(t)) = F(t) + C,$$

où C est une constante arbitraire. Encore une fois, cette formule reste vraie tant que $g \circ y$ ne s'annule pas. Si on arrive à inverser la fonction H , on a donc obtenu une formule permettant d'obtenir y en fonction de t .

Remarque 5.2.1. Ce type d'équation se rencontre souvent en physique. Les physiciens ont l'habitude d'écrire alors

$$y'(t) = \frac{dy}{dt} = g(y)f(t),$$

et par suite

$$\frac{dy}{g(y)} = f(t)dt.$$

Cette identité mène de nouveau à

$$H(y) = F(t) + C$$

On dit que l'on "sépare" les variables y et t , d'où le nom attribué à ces équations.

Remarque 5.2.2. Attention, la méthode ci-dessus n'est valable que pour un champ de vecteur $b(t, y)$ s'écrivant

$$b(t, y) = g(y)f(t).$$

Si on a $b(t, y) = g(y) + f(t)$, cette méthode échoue.

On peut justifier les calculs menés ci-dessus de façon rigoureuse, en montrant en particulier que la fonction $g \circ y$ ne s'annule jamais sur l'intervalle (maximal) de définition de y , et que la fonction H est toujours inversible. On obtient alors le résultat suivant :

Proposition 5.2.3. Soit I, U deux intervalles ouverts de \mathbb{R} , et soit $f : I \rightarrow \mathbb{R}$, $g : U \rightarrow \mathbb{R}$ deux fonctions de classe \mathcal{C}^1 .

Soit $E = \{z \in U, g(z) = 0\}$.

Soit $t_0 \in I, y_0 \in U$.

On considère la solution maximale y du problème de Cauchy

$$\begin{aligned} y'(t) &= f(t)g(y(t)), \\ y(t_0) &= y_0, \end{aligned}$$

et on note $J \subset I$ l'intervalle ouvert sur lequel elle est définie.

On a alors les propriétés suivantes :

- Si $y_0 \in E$, alors $J = I$ et $y(t) = y_0$ pour tout $t \in I$;
- Si $y_0 \notin E$: soit

$$U_0 = \bigcup_{\substack{O \text{ intervalle,} \\ y_0 \in O, \\ O \subset E^c}} O.$$

Alors $U_0 \neq \emptyset$, U_0 est ouvert, et U_0 est le plus grand intervalle contenant y_0 et inclus dans E^c . Soit H une primitive de $1/g$ sur U_0 , F une primitive de f sur I . On a :

- $y(t) \in U_0$ pour tout $t \in J$;
- Il existe une constante C telle que

$$H(y(t)) = F(t) + C \quad \forall t \in J;$$

- La fonction H est un \mathcal{C}^1 difféomorphisme de U_0 sur $H(U_0)$.

Démonstration. On a déjà justifié que si $g(y_0) = 0$ (i.e. $y_0 \in E$), alors $y(t) = y_0$ pour tout t . On s'intéresse donc au cas où $y_0 \notin E$.

Remarquons que $E = g^{-1}(\{0\})$, et g est continue; l'ensemble E est donc fermé (au sens de la "topologie induite" : $E = U \cap F$ où F est un fermé de \mathbb{R}), et E^c est un ouvert. Par conséquent, il existe $\delta > 0$ tel que $]y_0 - \delta, y_0 + \delta[\subset E^c$.

On remarque que

- U_0 est non vide (puisque $]y_0 - \delta, y_0 + \delta[\subset U_0$);
- U_0 est un intervalle contenant y_0 puisque c'est une union d'intervalles contenant y_0 ;
- $U_0 \subset E^c$;
- Si O est un intervalle contenant y_0 et inclus dans E^c , alors $O \subset U_0$ par définition de U_0 .

L'ensemble U_0 est donc le plus grand intervalle contenant y_0 et inclus dans E^c . De plus, si $z \in U_0$, d'une part il existe $O \subset E^c$ intervalle contenant y_0 et z , et d'autre part, puisque $z \in E^c$, il existe $\delta' > 0$ tel que $[z - \delta', z + \delta'] \subset E^c$. Alors $O \cup [z - \delta', z + \delta']$ est un intervalle contenant y_0 et inclus dans E^c , donc

$$[z - \delta', z + \delta'] \subset O \cup [z - \delta', z + \delta'] \subset U_0.$$

Par conséquent U_0 est ouvert.

Montrons à présent que $y(J) \subset U_0$: d'après le théorème des valeurs intermédiaires, $y(J)$ est un intervalle. De plus $y(J)$ contient y_0 . Montrons que $y(J) \subset E^c$: on raisonne par l'absurde et on suppose qu'il existe $y_1 \in E, t_1 \in J$ tel que $y(t_1) = y_1$. D'après l'unicité

pour le problème de Cauchy en $t = t_1$, on en déduit que $y(t) = y_1$ pour tout $t \in J$. En particulier $y_0 = y(t_0) = y_1 \in E$: absurde. Donc $y(J) \subset E^c$.

Par définition de U_0 , on en déduit que $y(J) \subset U_0$. Puisque $y(J) \subset E^c$, $g \circ y$ ne s'annule pas sur J et on peut écrire

$$\frac{y'(t)}{g(y(t))} = f(t) \quad \forall t \in J.$$

Cette égalité s'écrit encore

$$\frac{dH(y(t))}{dt} = f(t).$$

On en déduit qu'il existe une constante C telle que

$$H(y(t)) = F(t) + C \quad \forall t \in J.$$

Enfin, la fonction H est de classe \mathcal{C}^1 sur U_0 , et sa dérivée $1/g$ ne s'annule pas sur U_0 . Par conséquent H est un \mathcal{C}^1 difféomorphisme de U_0 sur $H(U_0)$. \square

Remarque 5.2.4. En pratique la constante C est déterminée par la relation

$$H(y_0) = F(t_0) + C.$$

Exemple 1

On considère l'équation différentielle

$$y'(t) = \frac{1}{y(t)},$$

avec $y(t_0) = y_0 \neq 0$.

On a ici $g : z \in \mathbb{R}^* \mapsto \frac{1}{z}$. On prend donc $U = \mathbb{R}_+^*$ si $y_0 > 0$, \mathbb{R}_-^* si $y_0 < 0$. On remarque alors que $E = \emptyset$, et donc $U_0 = U$. Soit y la solution maximale du problème de Cauchy, définie sur un intervalle ouvert J contenant t_0 . En particulier $y(t) \neq 0$ sur I_{max} donc $\text{sgn}(y(t)) = \text{sgn}(y_0)$.

Pour $t \in J$, on a

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

soit

$$\frac{1}{2} \frac{dy(t)^2}{dt} = 1.$$

On obtient donc

$$y(t)^2 - y_0^2 = 2t - 2t_0,$$

et

$$\forall t \in J, \quad y(t)^2 = y_0^2 + 2t - 2t_0.$$

On en déduit avec les arguments habituels que

$$J = \left] t_0 - \frac{y_0^2}{2}, +\infty \right[.$$

Puisque $y(t) \in U$ pour tout $t \in J$, on a

$$y(t) = \text{sgn}(y_0) \sqrt{y_0^2 + 2t - 2t_0} \quad \forall t \in J.$$

Exemple 2

On considère l'équation différentielle

$$y'(t) = y(t) - y(t)^2,$$

avec $y(t_0) = y_0 \in \mathbb{R}$.

La fonction $g : z \in \mathbb{R} \mapsto z - z^2$ s'annule en 0 et en 1. Ainsi, si $y_0 \in \{0, 1\}$, la fonction y est constante et égale à y_0 . On se concentre donc à présent sur les cas où $y_0 \notin \{0, 1\}$.

On a alors $U_0 = \mathbb{R}_*$ si $y_0 < 0$, $U_0 =]0, 1[$ si $y_0 \in]0, 1[$, $U_0 =]1, \infty[$ si $y_0 > 1$.

Avec les notations de la Proposition, pour tout $t \in J$, on a $y(t) \in U_0$, et

$$\frac{y'(t)}{y(t) - y(t)^2} = 1.$$

Pour appliquer la méthode, il faut calculer une primitive de $z \mapsto \frac{1}{z-z^2}$ sur U_0 . Pour cela, on décompose la fraction en éléments simples : on écrit que pour $z \in \mathbb{R} \setminus \{0, 1\}$,

$$\frac{1}{z - z^2} = \frac{1}{z(1 - z)} = \frac{1}{z} + \frac{1}{1 - z},$$

dont une primitive sur chacun des intervalles \mathbb{R}_* , $]0, 1[$, $]1, \infty[$ est

$$z \mapsto \ln |z| - \ln |1 - z| = \ln \left| \frac{z}{1 - z} \right|.$$

1. Premier cas : $y_0 < 0$: une primitive de $z \mapsto \frac{1}{z-z^2}$ sur \mathbb{R}_{*-} est

$$z \mapsto \ln \frac{-z}{1 - z} = \ln \left(1 - \frac{1}{1 - z} \right).$$

On obtient donc, pour tout $t \in J$,

$$\ln \left(1 - \frac{1}{1 - y(t)} \right) = t + C,$$

où

$$C = -t_0 + \ln \left(1 - \frac{1}{1 - y_0} \right) < -t_0.$$

En inversant la fonction du membre de gauche, on obtient finalement

$$y(t) = -\frac{e^{t+C}}{1 - e^{t+C}} \quad \forall t \in J,$$

d'où l'on déduit finalement que $J =] - \infty, -C[=] - \infty, t_0 - \ln \left(1 - \frac{1}{1 - y_0} \right) [$.

2. Deuxième cas : $y_0 \in]0, 1[$: une primitive de $z \mapsto \frac{1}{z-z^2}$ sur $]0, 1[$ est

$$z \mapsto \ln \frac{z}{1 - z} = \ln \left(\frac{1}{1 - z} - 1 \right).$$

On obtient donc, pour tout $t \in J$,

$$\ln \left(\frac{1}{1 - y(t)} - 1 \right) = t + C,$$

où

$$C = -t_0 + \ln\left(\frac{y_0}{1-y_0}\right) \in \mathbb{R}.$$

Dans ce cas on obtient

$$y(t) = \frac{e^{t+C}}{1+e^{t+C}} \quad \forall t \in J,$$

et donc finalement $J = \mathbb{R}$.

3. Troisième cas : $y_0 > 1$: une primitive de $z \mapsto \frac{1}{z-z^2}$ sur $]1, \infty[$ est

$$z \mapsto \ln \frac{z}{z-1} = \ln\left(1 + \frac{1}{z-1}\right).$$

On en déduit que pour tout $t \in J$,

$$\ln\left(1 + \frac{1}{y(t)-1}\right) = t + C, \quad \text{où } C = -t_0 + \ln\left(1 + \frac{1}{y_0-1}\right) > -t_0.$$

En inversant une nouvelle fois la formule, on arrive à

$$y(t) = \frac{e^{t+C}}{e^{t+C}-1} \quad \forall t \in J,$$

et donc $J =]-C, +\infty[=]t_0 - \ln\left(1 + \frac{1}{y_0-1}\right), +\infty[$.

Chapitre 6

Calculs d'exponentielle de matrice

6.1 Quelques exemples

Commençons par quelques exemples simples :

a) **Matrice diagonale :**

Soit $\lambda_1, \lambda_2 \in \mathbb{C}$, et soit $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Alors $e^A = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}$.

En effet : on montre facilement par récurrence que pour tout $k \in \mathbb{N}$,

$$A^k = \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix}.$$

On en déduit que

$$e^A = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}.$$

b) **Matrice nilpotente :**

Soit

$$N = \begin{pmatrix} 0 & \gamma \\ 0 & 0 \end{pmatrix}.$$

On vérifie aisément que $N^2 = 0$. Par suite, $N^k = 0$ pour tout $k \geq 2$. On obtient donc

$$e^N = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} N^k = I_2 + N = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix}.$$

c) **Matrices semblables :**

Soit $A, B \in \mathcal{M}_2(\mathbb{C})$, $P \in GL_2(\mathbb{C})$ telles que

$$B = PAP^{-1}.$$

Alors

$$e^B = Pe^AP^{-1}.$$

En effet, on vérifie facilement (par exemple par récurrence) que si $k \in \mathbb{N}$,

$$B^k = PA^kP^{-1}.$$

On en déduit alors que pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^n \frac{B^k}{k!} = P \left(\sum_{k=0}^n \frac{A^k}{k!} \right) P^{-1}.$$

La propriété s'ensuit.

d) **Matrices diagonalisables :**

Il découle des propriétés précédentes que l'exponentielle d'une matrice diagonalisable A se calcule facilement. En effet, si $A \in \mathcal{M}_2(\mathbb{C})$ est diagonalisable, alors il existe $\lambda_1, \lambda_2 \in \mathbb{C}$, $P \in GL_2(\mathbb{C})$ tels que

$$A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.$$

En ce cas,

$$e^A = P \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} P^{-1}.$$

e) **Un cas particulier de matrice réelle diagonalisable avec valeur propres non-réelles :** On considère la matrice

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix},$$

qui est une matrice réelle diagonalisable avec valeur propres non-réelles $\alpha \pm i\beta$ avec $\alpha, \beta \in \mathbb{R}, \beta \neq 0$ ¹. En calculant les vecteurs propres associées aux valeurs propres on conclut que

$$A = P \begin{pmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{pmatrix} P^{-1}, \text{ avec } P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}.$$

D'après l'expression de l'exponentielle pour les matrices semblables

$$\begin{aligned} e^A &= \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^{\alpha+i\beta} & 0 \\ 0 & e^{\alpha-i\beta} \end{pmatrix} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} \frac{1}{2i} \\ &= e^\alpha \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} \cos(\beta) + i \sin(\beta) & 0 \\ 0 & \cos(\beta) - i \sin(\beta) \end{pmatrix} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} \frac{1}{2i} \\ &= e^\alpha \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}. \end{aligned}$$

Résumons dans un tableau les propriétés obtenues jusqu'à présent :

Type	Formule	Exponentielle
Matrice diagonale	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$	$\begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}$
Un cas de matrice nilpotente	$\begin{pmatrix} 0 & \gamma \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix}$
Un cas de matrice réelle diagonalisable avec valeur propres non-réelles	$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$	$e^\alpha \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}$
Matrices semblables	$A = PBP^{-1}$	$e^A = Pe^BP^{-1}$

1. On peut aussi calculer $\exp(A)$ en remarquant que $A = \alpha I_2 + \beta N$. On utilise alors la formule $e^{M_1} e^{M_2} = e^{M_1+M_2}$ valable si les matrices M_1, M_2 commutent, et dont la démonstration est proche de celle de l'existence de l'exponentielle de matrice. On a donc $e^A = e^{\alpha I_2} e^{\beta N} = e^\alpha e^{\beta N}$ et on calcule $e^{\beta N} = \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}$ sachant que les puissances de N sont : $N^{2k} = (-1)^k I_2, N^{2k+1} = (-1)^k N$.

Pour traiter le cas général, on a besoin d'un résultat d'algèbre linéaire supplémentaire, comme suit.

6.2 Cas général

Théorème 6.2.1 (Décomposition de Dunford). *Soit $A \in \mathcal{M}_2(\mathbb{C})$ quelconque. Alors l'une des deux propriétés suivantes est vérifiée :*

1. *A est diagonalisable, i.e. il existe $\lambda_1, \lambda_2 \in \mathbb{C}$, $P \in GL_2(\mathbb{C})$ tels que*

$$A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.$$

2. *A n'est pas diagonalisable et dans ce cas A est trigonalisable : il existe $\lambda \in \mathbb{C}$, $\gamma \in \mathbb{C}^*$, $P \in GL_2(\mathbb{C})$ tels que*

$$A = P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1}.$$

Démonstration. On rappelle que les valeurs propres de A sont les racines de son polynôme caractéristique $\chi_A(X) = \det(A - XI_2) \in \mathbb{C}[X]$. De plus, d'après le théorème de d'Alembert-Gauss, χ_A est scindé et admet donc exactement deux racines complexes distinctes ou confondues. On distingue donc deux cas :

- Premier cas : χ_A admet deux racines distinctes $\lambda_1 \neq \lambda_2$:

En ce cas $\dim \ker(A - \lambda_1 I_2) \in \{1, 2\}$, $\dim \ker(A - \lambda_2 I_2) \in \{1, 2\}$, et comme

$$\dim \ker(A - \lambda_1 I_2) + \dim \ker(A - \lambda_2 I_2) \leq 2,$$

on en déduit que $\dim \ker(A - \lambda_1 I_2) = \dim \ker(A - \lambda_2 I_2) = 1$. Pour $i \in \{1, 2\}$, soit $v_i \in \ker(A - \lambda_i I_2) \setminus \{0\}$. Comme $\lambda_1 \neq \lambda_2$, v_1 et v_2 sont des vecteurs libres. Ils forment donc une base de \mathbb{C}^2 . Soit P la matrice de passage de la base canonique à la base (v_1, v_2) . On a alors

$$A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.$$

On est donc dans le premier cas énoncé dans le théorème.

- Deuxième cas : χ_A admet une seule racine $\lambda \in \mathbb{C}$:

Autrement dit, $\chi_A(X) = (X - \lambda)^2$.

On a de nouveau $\dim \ker(A - \lambda I_2) \in \{1, 2\}$. Si $\dim \ker(A - \lambda I_2) = 2 = \dim \mathbb{C}^2$, alors $A = \lambda I_2$: on est de nouveau dans le premier cas énoncé dans le théorème, avec $P = I_2$ et $\lambda_1 = \lambda_2 = \lambda$.

Il reste donc à examiner le cas où $\dim \ker(A - \lambda I_2) = 1$. Soit $v_1 \in \mathbb{C}^2$ tel que $\ker(A - \lambda I_2) = \mathbb{C}v_1$. On prend $v_2 \in \mathbb{C}^2 \setminus \ker(A - \lambda I_2)$. Alors les vecteurs v_1 et v_2 sont non nuls et non colinéaires. Ils sont donc libres et forment une base de \mathbb{C}^2 . En particulier il existe $\gamma, \mu \in \mathbb{C}$ tels que $Av_2 = \gamma v_1 + \mu v_2$.

Soit P la matrice de passage de la base canonique à la base (v_1, v_2) . Comme $Av_1 = \lambda v_1$, on note $Av_2 = \alpha v_1 + \mu v_2$ et on a :

$$A = P \begin{pmatrix} \lambda & \gamma \\ 0 & \mu \end{pmatrix} P^{-1},$$

où $\gamma, \mu \in \mathbb{C}$. Le polynôme caractéristique est invariant par similitude, donc

$$\chi_A(X) = (X - \lambda)^2 = (X - \lambda)(X - \mu).$$

On en déduit que $\lambda = \mu$. Aussi $\gamma \neq 0$ car sinon cela impliquerait $Av_2 = \lambda v_2$ donc $\dim \ker(A - \lambda I_2) = 2$. On est donc dans le deuxième cas énoncé dans le théorème. \square

Maintenant on peut calculer l'exponentielle d'une matrice quelconque $A \in \mathcal{M}_2(\mathbb{C})$ à l'aide du théorème précédent. Si la matrice A est diagonalisable alors on est dans le premier cas du théorème, et l'exponentielle d'une telle matrice a déjà été traité dans les remarques qui précèdent le théorème. Si la matrice A est non-diagonalisable nous sommes dans le deuxième cas du théorème et remarquons que

$$A = P(\lambda I_2 + N)P^{-1},$$

que $I_2 N = N I_2 = N$ et que $N^k = 0_2, \forall k \geq 2$. Par conséquent, pour tout $k \geq 0$,

$$(\lambda I_2 + N)^k = \sum_{j=0}^k C_k^j \lambda^j N^{k-j} = \lambda^k I_2 + k \lambda^{k-1} N = \begin{pmatrix} \lambda^k & \gamma k \lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}.$$

On obtient donc, pour tout $n \geq 0$

$$\sum_{k=0}^n \frac{A^k}{k!} = P \begin{pmatrix} \sum_{k=0}^n \frac{\lambda^k}{k!} & \gamma \sum_{k=1}^n \frac{\lambda^{k-1}}{(k-1)!} \\ 0 & \sum_{k=0}^n \frac{\lambda^k}{k!} \end{pmatrix} P^{-1}.$$

En passant à la limite quand $n \rightarrow \infty$, on arrive finalement à

$$e^A = P \begin{pmatrix} e^\lambda & \gamma e^\lambda \\ 0 & e^\lambda \end{pmatrix} P^{-1}.$$

Pour une matrice quelconque de $\mathcal{M}_2(\mathbb{C})$, on a donc deux possibilités :

Type	Formule	Exponentielle
Matrice diagonalisable	$P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}$	$P \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} P^{-1}$
Matrice trigonalisable (mais pas diagonalisable)	$P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1}$	$P \begin{pmatrix} e^\lambda & \gamma e^\lambda \\ 0 & e^\lambda \end{pmatrix} P^{-1}$

En pratique, nous allons devoir souvent calculer des exponentielles de matrices réelles $A \in \mathcal{M}_2(\mathbb{R})$. Pour calculer son exponentielle, on considère A comme une matrice complexe. Notons que e^A sera malgré tout une matrice réelle, comme on peut le voir sur la formule

$$e^A = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{A^k}{k!}.$$

De plus, si $A \in \mathcal{M}_2(\mathbb{R})$ a pour valeurs propres complexes λ_1, λ_2 , on a deux possibilités :

1. $\lambda_1 \in \mathbb{R}$ et $\lambda_2 \in \mathbb{R}$; dans ce cas les vecteurs propres peuvent être choisis réels et A est diagonalisable ou trigonalisable sur \mathbb{R} (autrement dit, dans le théorème 6.2.1, on a $P \in GL_2(\mathbb{R}), \gamma \in \mathbb{R}$).

2. On peut aussi calculer $\exp(A)$ en utilisant alors la formule $\exp(M_1) \exp(M_2) = \exp(M_1 + M_2)$ valable si les matrices M_1, M_2 commutent, et on obtient $e^A = P \begin{pmatrix} e^\lambda & 0 \\ 0 & e^\lambda \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} P^{-1} = P \begin{pmatrix} e^\lambda & \gamma e^\lambda \\ 0 & e^\lambda \end{pmatrix} P^{-1}$.

2. $\lambda_1 \in \mathbb{C} \setminus \mathbb{R}, \lambda_2 \in \mathbb{C} \setminus \mathbb{R}$; dans ce cas λ_1 et λ_2 sont des nombres complexes conjugués $\alpha \pm i\beta$. De plus, comme $\lambda_1 \notin \mathbb{R}$, on ne peut pas avoir $\lambda_1 = \lambda_2$ (sinon $\lambda_1 = \bar{\lambda}_1$). Donc les valeurs propres λ_1 et λ_2 sont nécessairement distinctes, et A est diagonalisable sur \mathbb{C} . Il existe donc $P \in \mathcal{M}_2(\mathbb{C})$ telle que

$$A = P \begin{pmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{pmatrix} P^{-1}.$$

On écrit alors

$$A = P(\alpha I_2 + \beta \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix})P^{-1} = \alpha I_2 + \beta P \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} P^{-1}.$$

Comme $A \in \mathcal{M}_2(\mathbb{R})$ ceci implique que $P \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} P^{-1} \in \mathcal{M}_2(\mathbb{R})$. D'autre part, cette matrice a comme valeurs propres $\pm i$, tout comme la matrice réelle $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Il s'agit donc de deux matrices réelles semblables dans \mathbb{C} . Or on a le lemme suivant, que l'on démontre à la fin de cette section :

Lemme 6.2.2. *Soit $n \in \mathbb{N}$, et soit $M, \tilde{M} \in \mathcal{M}_n(\mathbb{R})$. On suppose que M et \tilde{M} sont semblables dans \mathbb{C} : il existe $S \in \text{GL}_n(\mathbb{C})$ tel que $M = S\tilde{M}S^{-1}$.*

Alors M et \tilde{M} sont semblables dans \mathbb{R} : il existe $P \in \text{GL}_n(\mathbb{R})$ tel que $M = P\tilde{M}P^{-1}$.

Dans le cas présent on déduit donc l'existence de $R \in \text{GL}_2(\mathbb{R})$ telle que

$$P \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} P^{-1} = R \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} R^{-1}.$$

On obtient alors

$$A = \alpha I_2 + \beta R \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} R^{-1} = R(\alpha I_2 + \beta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix})R^{-1} = R \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} R^{-1},$$

donc

$$\exp(A) = R \exp(\alpha) \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} R^{-1}.$$

Résumons les formules obtenues pour les exponentielles de matrices réelles :

Type	Formule	Exponentielle
Matrice diagonalisable avec valeurs propres $\lambda_1, \lambda_2 \in \mathbb{R}$ distinctes ou confondues	$P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}, P \in \text{GL}_n(\mathbb{R})$	$P \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} P^{-1}$
Matrice diagonalisable avec valeurs propres non-réelles $\alpha \pm i\beta, \beta \neq 0$	$P \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} P^{-1}, P \in \text{GL}_n(\mathbb{R})$	$P e^\alpha \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} P^{-1}$
Matrice non-diagonalisable avec une valeur propre double $\lambda \in \mathbb{R}$	$P \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix} P^{-1}, P \in \text{GL}_n(\mathbb{R}), \gamma \in \mathbb{R}^*$	$P e^\lambda \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} P^{-1}$

Enfin, voici la preuve du Lemme 6.2.2 : On décompose S en sa partie réelle et sa partie imaginaire,

$$S = S_1 + iS_2, \quad S_1, S_2 \in \mathcal{M}_n(\mathbb{R}).$$

En prenant la partie réelle et la partie imaginaire de l'identité $MS = S\tilde{M}$, on obtient

$$MS_1 = S_1\tilde{M} \quad \text{et} \quad MS_2 = S_2\tilde{M}.$$

Attention : en général les matrices S_1 et S_2 ne sont pas inversibles !

Pour $x \in \mathbb{C}$, on considère la matrice $T(x) = S_1 + xS_2$ et son déterminant $p(x) = \det(T(x))$. On fait alors les observations suivantes :

- Pour tout $x \in \mathbb{R}$, $T(x) \in \mathcal{M}_n(\mathbb{R})$ et $MT(x) = T(x)\tilde{M}$;
- p est un polynôme,
- Si $x = i$, on a $T(i) = S_1 + iS_2 = S$, et donc $p(i) \neq 0$.

On déduit des deux dernières propriétés que p n'est pas le polynôme nul. Par conséquent, p s'annule au plus un nombre fini de fois sur \mathbb{R} . Il existe donc $x_0 \in \mathbb{R}$ tel que $p(x_0) \neq 0$. Ceci implique que la matrice $T(x_0)$ est inversible et à coefficients réels. En choisissant $P = T(x_0)$, on a bien $P \in \text{GL}_n(\mathbb{R})$ et $MP = P\tilde{M}$. Donc M et \tilde{M} sont semblables dans $\mathcal{M}_n(\mathbb{R})$.