

2M310 - Équations différentielles : éléments d'analyse et
approximation numérique

Anne-Laure Dalibard

13 janvier 2016

Chapitre 1

Introduction et motivation

1.1 Définition et premiers exemples

Une équation différentielle est une équation reliant une variable réelle t , une fonction inconnue x dépendant de la variable t et à valeurs dans \mathbb{R}^m (dans ce cours on prendra $m = 1$ ou $m = 2$), 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

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

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 .

Bien souvent, ces équations sont munies de “conditions initiales”, ou de “conditions au bord” : 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 de calculer des solutions approchées.

On reviendra sur ces différentes notions au cours du chapitre.

Un peu d'histoire...

Le calcul différentiel naît à la fin du XVIIe 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 $x'(t)$.

Un autre problème célèbre est celui de la chaînette, c'est-à-dire la question de savoir quelle forme prend une chaîne, suspendue aux deux extrémités, sous l'action de son propre poids. Ce problème a été résolu simultanément par Leibniz, Jean Bernoulli et Huygens (1691). 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, 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.

Au XIXe siècle, 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

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

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

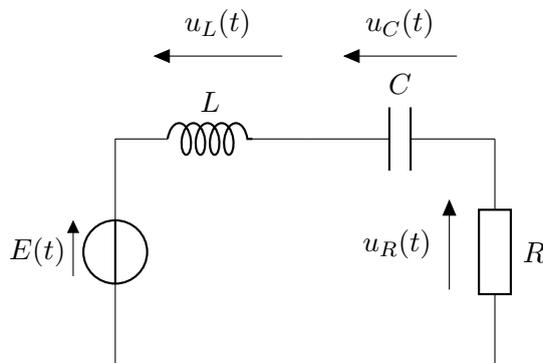
- l'analyse théorique, dont le but est de montrer l'existence et l'unicité des solutions d'équations différentielles à l'aide d'outils généraux ;
- l'analyse numérique, dont le but est de calculer des solutions approchées de ces solutions, avec la meilleure précision possible.

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

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

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

- D'après la loi d'Ohm, $u_R(t) = Ri(t)$;
- La charge aux bornes du condensateur est $Q(t) = Cu_C(t)$, et donc $i(t) = \frac{dQ(t)}{dt} = C \frac{du_C}{dt}$;
- La tension aux bornes de la bobine est $u_L(t) = L \frac{di}{dt}$.

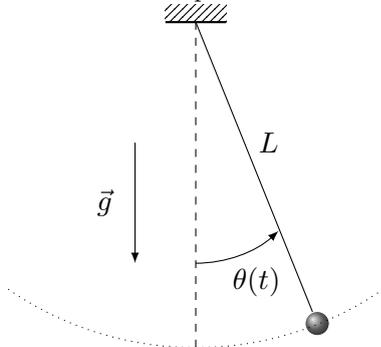
On obtient donc finalement

$$LCu_C''(t) + RCu_C'(t) + u_C(t) = E(t).$$

L'équation ci-dessus est une équation linéaire d'ordre deux, à coefficients constants. On verra plus tard comment la résoudre en général.

Remarque 1.1.1. Dans ce cours, on notera systématiquement la dérivée d'une fonction x comme x' , et jamais comme \dot{x} , par souci d'homogénéité.

▷ 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, 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 : si θ_1, θ_2 sont deux solutions de (1.2), en général $\theta_1 + \theta_2$ n'est pas solution. De façon générale, il est impossible de connaître explicitement les solutions d'équations non linéaires. On calcule donc souvent des solutions exactes d'une équation approchée de (1.2)¹ 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 a

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

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 $t + dt$

▷ 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 b (resp. d). On considère la variation de la population entre t et $t + dt$. On a

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

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

et donc

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

On arrive donc à l'équation différentielle

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

où $r = b - d$.

Les solutions de cette équation sont de la forme $N_0 \exp(rt)$. En particulier :

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

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 proposé par Verhulst :

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

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 $N(t = 0) = N_0$ est*

$$N(t) = N_0 \frac{Ke^{rt}}{K + N_0(e^{rt} - 1)}.$$

Démonstration. Posons

$$\bar{N}(t) = N_0 \frac{Ke^{rt}}{K + N_0(e^{rt} - 1)}.$$

On vérifie que l'on a bien $\bar{N}(0) = N(0) = N_0$. Par ailleurs,

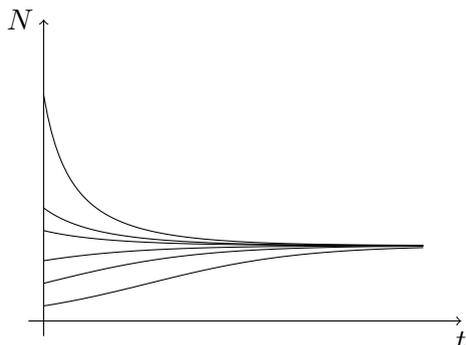
$$\begin{aligned} \bar{N}'(t) &= N_0 \frac{(K + N_0(e^{rt} - 1))Kre^{rt} - Ke^{rt}N_0re^{rt}}{(K + N_0(e^{rt} - 1))^2} \\ &= N_0 \frac{K^2re^{rt} - N_0Kre^{rt}}{(K + N_0(e^{rt} - 1))^2} \\ &= r\bar{N}(t) \frac{K - N_0}{K + N_0(e^{rt} - 1)} \\ &= r\bar{N}(t) \frac{K + N_0e^{rt} - N_0 - N_0e^{rt}}{K + N_0(e^{rt} - 1)} \\ &= r\bar{N}(t) \left(1 - \frac{\bar{N}(t)}{K} \right). \end{aligned}$$

Donc \bar{N} est solution de l'équation (1.5).

On verra au chapitre 3 un théorème d'unicité pour les solutions de l'équation (1.5). On en déduit que $N(t) = \bar{N}(t)$. □

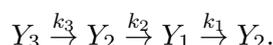
On vérifie aisément sur la formule que $N(t) \rightarrow K$ quand $t \rightarrow \infty$ quelle que soit la valeur de N_0 .

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



▷ *Cinétique chimique :*

Considérons une réaction chimique sur trois espèces :



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. On verra plus tard comment en obtenir une solution explicite.

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 à 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 type² :

- 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 $t \rightarrow \infty$;

1.4 Quelques notions sur les courbes paramétrées

Afin de représenter les trajectoires des solutions d'équations différentielles, on a besoin d'introduire la notion de courbe paramétrée. Le but de ce paragraphe est de présenter les idées majeures d'une telle représentation, en se concentrant sur les notions qui seront utilisées dans le reste du cours. D'autres résultats annexes, mais qui ne sont pas nécessaires pour la compréhension de ce cours, sont donnés en Appendice.

1.4.1 Généralités

Définition 1.4.1. Une courbe paramétrée \mathcal{C} est un ensemble de points de \mathbb{R}^2 de la forme

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2; \exists t \in I, x = f(t) \text{ et } y = g(t)\},$$

où I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$ et $g : I \rightarrow \mathbb{R}$ deux fonctions à valeurs réelles.

La variable t est appelée paramètre de la courbe.

Remarque 1.4.2. On parle également de courbes paramétrées en coordonnées cartésiennes, pour souligner le fait que les fonctions f et g donnent les coordonnées cartésiennes d'un point de la courbe.

Exemple 1.4.3. 1. Soit I un intervalle de \mathbb{R} , $h : I \rightarrow \mathbb{R}$ une fonction à valeurs réelles. Alors le graphe représentatif de h , c'est-à-dire l'ensemble des points

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2; y = h(x) \text{ et } x \in I\}$$

est une courbe paramétrée : il suffit de prendre dans la définition $f = \text{id}$ (i.e. $x = t$) et $g = h$.

2. Le cercle unité

$$\mathcal{C}_0 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\}$$

est une courbe paramétrée : en effet, on a aussi

$$\mathcal{C}_0 = \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}, x = \cos t \text{ et } y = \sin t\}.$$

2. attention, certaines de ces propriétés sont incompatibles !

Remarque 1.4.4. 1. Le paramétrage n'est pas unique : pour le cercle unité \mathcal{C}_0 , on peut aussi prendre

$$\mathcal{C}_0 = \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}, x = \sin t \text{ et } y = \cos t\}$$

ou encore

$$\mathcal{C}_0 = \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}, x = \cos(2t) \text{ et } y = \sin(2t)\}.$$

De façon générale, il suffit de dilater/translater l'intervalle de définition I (en modifiant les fonctions f et g de façon ad hoc) pour que cela modifie le paramétrage.

2. Physiquement, on peut interpréter le paramètre t comme un temps, et une courbe paramétrée comme la trajectoire d'un point en fonction du temps : à tout temps t correspond une position $(f(t), g(t))$.

Étude des branches infinies

Soit

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2; \exists t \in I, x = f(t) \text{ et } y = g(t)\},$$

une courbe paramétrée (avec les notations du paragraphe précédent), et soit $a \in \bar{I}$.

Définition 1.4.5. On dit que \mathcal{C} possède une branche infinie au voisinage de a si

$$\lim_{t \rightarrow a, t \in I} |f(t)| + |g(t)| = +\infty.$$

Plusieurs cas sont possibles :

1. Premier cas : seule l'une des deux limites $\lim_{t \rightarrow a} f(t)$ ou $\lim_{t \rightarrow a} g(t)$ est infinie (l'autre est finie).
2. Deuxième cas : les deux limites $\lim_{t \rightarrow a} f(t)$ et $\lim_{t \rightarrow a} g(t)$ sont infinies.
 1. Dans le premier cas, la courbe admet une asymptote horizontale ou verticale. Plus précisément :
 - (a) Si $\lim_{t \rightarrow a} f(t) = m \in \mathbb{R}$ et $\lim_{t \rightarrow a} |g(t)| = +\infty$, la droite d'équation $x = m$ est appelée asymptote de \mathcal{C} en a .
 - (b) Si $\lim_{t \rightarrow a} |f(t)| = +\infty$ et $\lim_{t \rightarrow a} g(t) = m \in \mathbb{R}$, la droite d'équation $y = m$ est appelée asymptote de \mathcal{C} en a .
 2. Dans le deuxième cas, il faut distinguer suivant la limite du rapport $g(t)/f(t)$ (si elle existe !)
 - (a) Si $\lim_{t \rightarrow a} \frac{g(t)}{f(t)} = 0$, on dit que \mathcal{C} possède une branche parabolique dans la direction (Ox) .
 - (b) Si $\lim_{t \rightarrow a} \frac{g(t)}{f(t)} = \pm\infty$, on dit que \mathcal{C} possède une branche parabolique dans la direction (Oy) .
 - (c) Si $\lim_{t \rightarrow a} \frac{g(t)}{f(t)} = m \in \mathbb{R}^*$, il faut regarder la limite de la quantité $g(t) - mf(t)$:
 - i. Si $\lim_{t \rightarrow a} g(t) - mf(t) = \pm\infty$, on dit que \mathcal{C} possède une branche parabolique dans la direction $y = mx$.
 - ii. Si $\lim_{t \rightarrow a} g(t) - mf(t) = p \in \mathbb{R}$, la droite d'équation $y = mx + p$ est appelée asymptote de \mathcal{C} en a .

Vecteur tangent

On reprend les notations de la définition 1.4.1.

Propriété 1.4.6. Soit $a \in I$. On suppose que les fonctions f et g sont de classe \mathcal{C}^1 sur I .

Si $(f'(a), g'(a)) \neq (0, 0)$, alors la tangente à la courbe \mathcal{C} au point de paramètre a est la droite qui passe par le point de coordonnées $(f(a), g(a))$ et dirigée par le vecteur de coordonnées $(f'(a), g'(a))$.

En particulier :

- Si $g'(a) = 0$ et $f'(a) \neq 0$, alors il y a une tangente horizontale à la courbe au point de coordonnées $(f(a), g(a))$.
- Si $g'(a) \neq 0$ et $f'(a) = 0$, alors il y a une tangente verticale à la courbe au point de coordonnées $(f(a), g(a))$.

Démonstration. On se place dans un voisinage du point $a \in I$.

- *Premier cas :* On suppose que $f'(a) \neq 0$.

On utilise alors le théorème d'inversion locale pour représenter \mathcal{C} comme le graphe d'une fonction \mathcal{C}^1 au voisinage de a . Plus précisément, comme f est de classe \mathcal{C}^1 et que $f'(a) \neq 0$, il existe un voisinage $V \subset I$ de a tel que f est un \mathcal{C}^1 difféomorphisme de V sur $f(V)$. On note $\varphi : f(V) \rightarrow V$ l'inverse de la fonction $f|_V$, c'est-à-dire que pour tout $t \in V$, pour tout $x \in f(V)$,

$$x = f(t) \iff t = \varphi(x).$$

Il s'ensuit que pour $t \in V$, $x \in f(V)$ on a

$$\begin{cases} x = f(t), \\ y = g(t) \end{cases} \iff y = g \circ \varphi(x),$$

et la fonction $g \circ \varphi$ est de classe \mathcal{C}^1 sur $f(V)$. La portion de \mathcal{C} correspondant au voisinage V de a est donc le graphe de $g \circ \varphi$ au voisinage du point $f(a)$. On sait alors que la tangente à \mathcal{C} au point de coordonnées $(f(a), g(a)) = (f(a), g \circ \varphi(f(a)))$ est la droite passant par $(f(a), g(a))$ et de coefficient directeur $(g \circ \varphi)'(f(a))$. Or

$$(g \circ \varphi)'(f(a)) = \varphi'(f(a))(g' \circ \varphi)(f(a)) = \frac{1}{f' \circ \varphi}(f(a))g'(a) = \frac{g'(a)}{f'(a)}.$$

On en déduit la propriété.

- *Deuxième cas :* On suppose que $f'(a) = 0$. Dans ce cas $g'(a) \neq 0$, et on peut de nouveau appliquer le théorème d'inversion locale. Il existe un voisinage $V \subset I$ de a tel que g est un \mathcal{C}^1 difféomorphisme de V sur $g(V)$. On note $\chi : g(V) \rightarrow V$ l'inverse de la fonction $g|_V$, c'est-à-dire que pour tout $t \in V$, pour tout $y \in g(V)$,

$$y = g(t) \iff t = \chi(y).$$

Il s'ensuit que pour $t \in V$, $y \in g(V)$ on a

$$\begin{cases} x = f(t), \\ y = g(t) \end{cases} \iff x = (f \circ \chi)(y),$$

et la fonction $f \circ \chi$ est de classe \mathcal{C}^1 sur $g(V)$. La portion de \mathcal{C} correspondant au voisinage V de a est donc l'image par la symétrie par rapport à la droite $y = x$ du graphe de $f \circ \chi$ au voisinage du point $g(a)$. On sait alors que la tangente à \mathcal{C} au point de coordonnées $(f(a), g(a)) = (f \circ \chi(g(a)), g(a))$ est l'image de la droite passant par $(g(a), f(a))$ et de coefficient directeur $\chi'(g(a))f'(a) = 0$. Il s'agit donc de la droite verticale d'équation $x = f(a)$.

□

Il reste à comprendre ce qui se passe lorsque $f'(a) = g'(a) = 0$. Dans ce cas, le point a est dit *stationnaire*. L'étude est alors plus compliquée. S'il existe deux entiers $m, n \geq 2$ avec $m \neq n$ tels que les vecteurs $(f^{(m)}(a), g^{(m)}(a))$ et $(f^{(n)}(a), g^{(n)}(a))$ sont non nuls et linéairement indépendants, alors l'analyse au voisinage du point de paramètre a est donnée dans la proposition 3.6.1 en Appendice. Néanmoins, cette situation ne sera pas rencontrée dans la suite du cours. En revanche, on rencontrera à plusieurs reprises le cas de figure suivant : pour un point $a \in \bar{I}$ tel que $(f(t), g(t))$ admet une limite quand $t \rightarrow a$ avec $t \in I$, on a

$$\lim_{t \rightarrow a} f^{(n)}(t) = \lim_{t \rightarrow a} g^{(n)}(t) = 0 \quad \forall n \in \mathbb{N}.$$

C'est par exemple le cas pour la courbe paramétrée définie par

$$f(t) = e^{-t}, \quad g(t) = te^{-2t}, \quad t \in \mathbb{R}, \quad (1.6)$$

avec $a = +\infty \in \bar{\mathbb{R}}$. On pourra alors utiliser les propriétés suivantes :

Propriété 1.4.7. *On pose $(x_0, y_0) = \lim_{t \rightarrow a}(f(t), g(t))$. Si $f(t) \neq x_0$ pour $t \in I$, on considère le taux d'accroissement*

$$\rho(t) := \frac{g(t) - y_0}{f(t) - x_0}.$$

On note \bar{C} la courbe définie par $\bar{C} := C \cup \{(x_0, y_0)\}$.

On a alors les résultats suivants :

1. Si $\rho(t)$ est bien défini dans un voisinage de a et si $\rho(t) \rightarrow \pm\infty$ quand $t \rightarrow a$, alors la courbe \bar{C} admet une tangente verticale en (x_0, y_0) .
2. Si $\rho(t)$ est bien défini dans un voisinage de a et si $\rho(t) \rightarrow 0$ quand $t \rightarrow a$, alors la courbe \bar{C} admet une tangente horizontale en (x_0, y_0) .
3. Si $\rho(t)$ est bien défini dans un voisinage de a et si $\rho(t) \rightarrow l \in \mathbb{R}^*$ quand $t \rightarrow a$, alors la courbe \bar{C} admet une tangente de pente l en (x_0, y_0) .

Démonstration. Les trois propriétés sont des conséquences du fait que la corde passant par les points $M(t)$ et (x_0, y_0) a pour pente $\rho(t)$. Par conséquent, si $\rho(t)$ a une limite l quand $t \rightarrow a$, cette valeur l est la pente de la tangente en (x_0, y_0) (avec la convention qu'une droite de pente infinie est verticale). □

Dans le cas de l'exemple (1.6), on a

$$\lim_{t \rightarrow +\infty} f(t) = 0, \quad \lim_{t \rightarrow +\infty} g(t) = 0,$$

donc $a = +\infty$ et $(x_0, y_0) = (0, 0)$. De plus

$$\rho(t) = \frac{g(t) - 0}{f(t) - 0} = t.$$

Donc $\lim_{t \rightarrow +\infty} \rho(t) = +\infty$: la courbe \bar{C} admet une tangente verticale en $(0, 0)$.

1.4.2 Exemple de tracé

Méthode générale

1. Déterminer l'ensemble de définition : c'est l'intersection des ensembles de définition de f et de g . Préciser aussi les dérivabilités de f et g ;
2. Rechercher les périodicités et les symétries éventuelles pour réduire le domaine d'étude :
 - Si $\begin{cases} f(-t) = -f(t) \\ g(-t) = g(t) \end{cases}$ alors le point $M(-t)$ peut se déduire de $M(t)$ par symétrie par rapport à l'axe des ordonnées.
 - Si $\begin{cases} f(-t) = f(t) \\ g(-t) = -g(t) \end{cases}$ alors le point $M(-t)$ peut se déduire de $M(t)$ par symétrie par rapport à l'axe des abscisses.
 - Si $\begin{cases} f(-t) = -f(t) \\ g(-t) = -g(t) \end{cases}$ alors le point $M(-t)$ peut se déduire de $M(t)$ par symétrie par rapport à l'origine.
 - Si $\begin{cases} f(-t) = g(t) \\ g(-t) = f(t) \end{cases}$ alors le point $M(-t)$ peut se déduire de $M(t)$ par symétrie par rapport à la droite d'équation $y = f$.
3. Établir le tableau de variation avec les lignes : t, f', f, g', g . Pour cela il faut calculer les dérivées et étudier leur signe ;
4. Placer sur la courbe les points à tangente verticale ($f' = 0$) ou horizontale ($g' = 0$), qui apparaissent sur le tableau ;
5. Étudier les branches infinies, tracer les asymptotes éventuelles ;
6. Tracer la courbe (et rétablir les symétries trouvées en 2) ;
7. Éventuellement, pour préciser le tracé, déterminer les points multiples (i.e. les points tels qu'il existe $t \neq t'$ avec $M(t) = M(t')$, les intersections de la courbe avec les axes de coordonnées et les points d'inflexion.

Exemple

On considère la courbe paramétrée \mathcal{C} définie par

$$\begin{cases} x = \frac{t^2}{t-1}, \\ y = \frac{t}{t^2-1} \end{cases}, \quad t \in \mathbb{R}.$$

1. Domaine de définition : $D = \mathbb{R} \setminus \{-1, 1\}$.
2. Pas de symétrie évidente ;
3. Tableau de variations :

Soit

$$\begin{aligned} f : t \in D &\mapsto \frac{t^2}{t-1}, \\ g : t \in D &\mapsto \frac{t}{t^2-1}. \end{aligned}$$

Les fonctions f et g sont dérivables sur D . Pour tout $t \in D$, on a

$$\begin{aligned} f'(t) &= \frac{t(t-2)}{(t-1)^2}, \\ g'(t) &= -\frac{t^2+1}{(t^2-1)^2}, \end{aligned}$$

On a donc le tableau de variations suivant :

t	$-\infty$	-1	0	1	2	$+\infty$
$f'(t)$	$+$	\vdots	$+$	0	$-$	$+$
$f(t)$	$-\infty$		0	$-\infty$	$+\infty$	4
$g'(t)$	$-$	\parallel	$-$	\vdots	$-$	$-$
$g(t)$	0	$-\infty$	$+\infty$	$-\infty$	$+\infty$	0

4. On voit sur le tableau de variations qu'il y a des tangentes verticales aux points de paramètres $t = 0$ et $t = 2$, qui ont respectivement pour coordonnées $(0, 0)$ et $(4, 2/3)$. Il n'y a pas de point à tangente horizontale.

5. Branches infinies :

- $t \rightarrow -\infty$: $\lim_{t \rightarrow -\infty} f(t) = -\infty$ et $\lim_{t \rightarrow -\infty} g(t) = 0^-$. Par conséquent la droite d'équation $y = 0$ est asymptote à la courbe quand $x \rightarrow -\infty$, et la courbe est en dessous de son asymptote.
- $t \rightarrow -1$: on a $f(-1) = -1/2$. Par conséquent la droite d'équation $x = -1/2$ est asymptote à la courbe. Pour $t \rightarrow -1^-$, on a $y = g(t) \rightarrow -\infty$ et $x = f(t) \rightarrow -1/2^-$: la courbe est à gauche de son asymptote quand $y \rightarrow -\infty$. Pour $t \rightarrow -1^+$, on a $y = g(t) \rightarrow +\infty$ et $x = f(t) \rightarrow -1/2^+$: la courbe est à droite de son asymptote quand $y \rightarrow +\infty$.
- $t \rightarrow 1$: on a $|f(t)| \rightarrow +\infty$ et $|g(t)| \rightarrow +\infty$. On étudie donc le rapport $g(t)/f(t)$. Pour $t \neq \pm 1$, on a

$$\frac{g(t)}{f(t)} = \frac{t}{t^2 - 1} \frac{t - 1}{t^2} = \frac{1}{t(t + 1)} \rightarrow \frac{1}{2} \quad \text{quand } t \rightarrow 1.$$

On étudie à présent le comportement de la quantité $g(t) - \frac{1}{2}f(t)$. On a

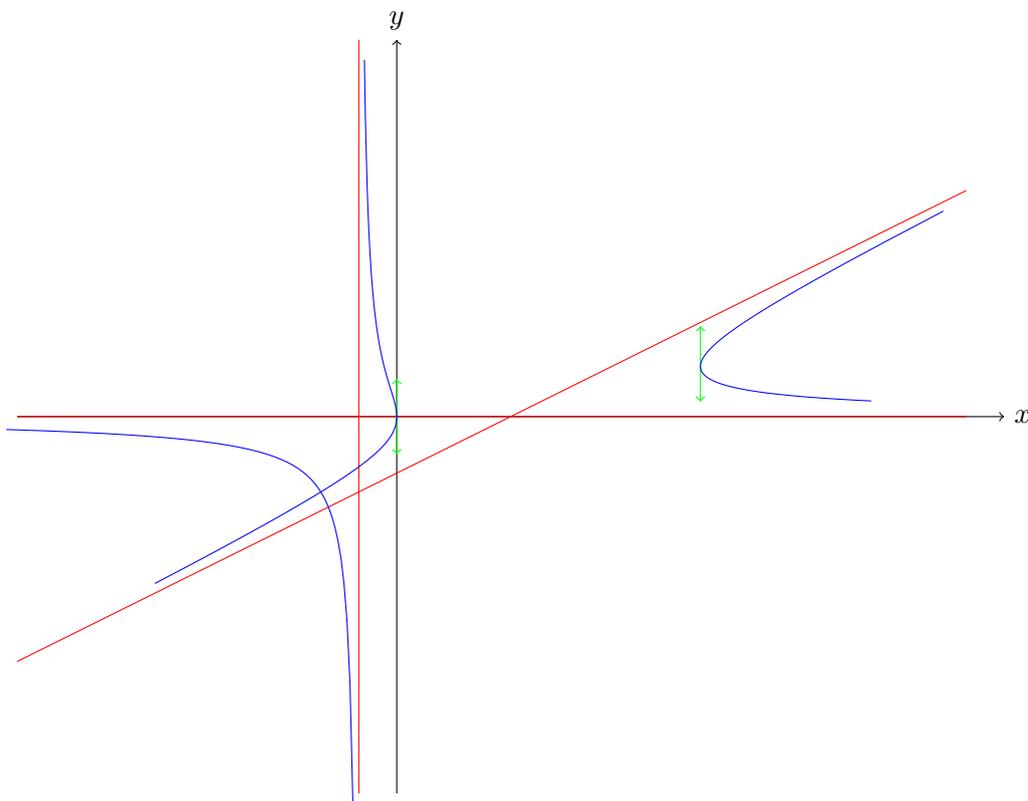
$$g(t) - \frac{1}{2}f(t) = \frac{-\frac{1}{2}t^3 - \frac{1}{2}t^2 + t}{t^2 - 1} = \frac{-\frac{1}{2}t(t - 1)(t + 2)}{t^2 - 1} = -\frac{t(t + 2)}{2(t + 1)} \rightarrow -\frac{3}{4} \quad \text{quand } t \rightarrow 1.$$

Par conséquent la droite d'équation $y = \frac{1}{2}x + \frac{3}{4}$ est asymptote à la courbe quand $x \rightarrow \pm\infty$.

- $t \rightarrow +\infty$: $f(t) \rightarrow +\infty$ et $g(t) \rightarrow 0^+$. Par conséquent la droite d'équation $y = 0$ est asymptote à la courbe quand $x \rightarrow +\infty$, et la courbe est au-dessus de son asymptote.

6. Intersections avec les axes des coordonnées : on vérifie facilement sur le tableau de variations que la courbe n'intersecte pas les axes de coordonnées quand le paramètre t appartient aux intervalles $] -\infty, -1[$ et $]1, +\infty[$. Sur l'intervalle $] -1, 1[$, la seule intersection se produit pour le point de paramètre $t = 0$, et la courbe passe alors par le point $(0, 0)$ où elle admet une tangente verticale.

7. Tracé :



1.4.3 Courbes en coordonnées polaires

Définition 1.4.8. Pour tout point $M = (x, y) \neq (0, 0)$ du plan \mathbb{R}^2 , il existe un unique couple $(r, \theta) \in \mathbb{R}_+^* \times [0, 2\pi[$ tel que

$$\begin{aligned}x &= r \cos \theta, \\y &= r \sin \theta.\end{aligned}$$

Le couple (r, θ) forme les “coordonnées polaires” du point M .

Remarque 1.4.9. On a, par définition,

$$r = \sqrt{x^2 + y^2}.$$

L'angle θ est donc uniquement déterminé (modulo 2π) en fonction des coordonnées cartésiennes par

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}.$$

Plus généralement, on peut associer à un couple (r, θ) de réels le point de coordonnées

$$\begin{aligned}x &= r \cos \theta, \\y &= r \sin \theta.\end{aligned}$$

On ne suppose alors plus que r est positif. On identifie alors le point associé au couple $(-r, \theta)$ avec $r > 0$ au point de coordonnées polaires $(r, \theta + \pi)$.

Exemples

1. Soit $\alpha_0 \in [0, 2\pi[$ fixé. L'ensemble des points

$$D := \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}, x = t \cos \alpha_0, y = t \sin \alpha_0\}$$

est la droite passant par l'origine et faisant un angle α_0 avec l'axe des abscisses.

2. Soit $r_0 > 0$. L'ensemble des points

$$C := \{(x, y) \in \mathbb{R}^2, \exists t \in [0, 2\pi[, \quad x = r_0 \cos t, \quad y = r_0 \sin t\}$$

est le cercle de centre 0 et de rayon r_0 .

Définition 1.4.10. Soit I un intervalle de \mathbb{R} , $\rho : I \rightarrow \mathbb{R}$ une fonction continue.

La courbe définie par

$$C := \{(x, y) \in \mathbb{R}^2, \exists \theta \in I, \quad x = \rho(\theta) \cos \theta, \quad y = \rho(\theta) \sin \theta\}$$

est une courbe paramétrée. On dit que l'équation polaire de cette courbe est

$$r = \rho(\theta).$$

Analyse des courbes paramétrées définies par une équation polaire

Comme dans le cas général, on commence par chercher le domaine d'étude de la fonction ρ et ses symétries éventuelles. On a en particulier les propriétés suivantes :

- Si ρ est paire, la courbe est symétrique par rapport à l'axe des abscisses ;
- Si ρ est impaire, la courbe est symétrique par rapport à l'axe des ordonnées ;
- Si ρ est définie sur \mathbb{R} et périodique de période 2π , on peut se contenter d'étudier la courbe sur l'intervalle $[0, 2\pi]$;
- S'il existe $\theta_0 \in \mathbb{R}$ tel que $\rho(\theta_0 - \theta) = \rho(\theta)$, la courbe est symétrique par rapport à la droite définie par $\theta = \theta_0/2 \pmod{\pi}$, et on peut restreindre le domaine d'étude à $[\theta_0/2, +\infty[$ puis appliquer la symétrie ;
- S'il existe $\theta_0 \in \mathbb{R}$ tel que $\rho(\theta_0 - \theta) = -\rho(\theta)$, la courbe est symétrique par rapport à la droite définie par $\theta = (\theta_0 + \pi)/2 \pmod{\pi}$, et on peut restreindre le domaine d'étude à $[\theta_0/2, +\infty[$ puis appliquer la symétrie.

Il faut également tracer le tableau de variations de ρ , même si celui-ci n'est pas aussi parlant que pour les courbes en coordonnées cartésiennes, dans la mesure où on ne peut rien déduire directement des annulations de ρ' en ce qui concerne les pentes des tangentes. En revanche, il est nécessaire, pour tracer la courbe, de connaître le signe de ρ partout.

▷ *Branches infinies :*

- Au voisinage de $\theta_0 \in \mathbb{R}$: si $\rho(\theta) \rightarrow +\infty$ quand $\theta \rightarrow \theta_0^\pm$, alors la courbe admet une branche asymptotique de direction la droite D_{θ_0} d'angle polaire $\theta = \theta_0$. On vérifie alors que la distance du point $M(\theta)$ à D_{θ_0} est $|\rho(\theta) \sin(\theta - \theta_0)|$. Notons $d(\theta)$ cette quantité.
 - Si $d(\theta)$ tend vers une limite finie \bar{d} quand $\theta \rightarrow \theta_0^\pm$, alors la droite parallèle à D_{θ_0} et à une distance $\pm \bar{d}$ de celle-ci est asymptote à la courbe.
 - Si $d(\theta) \rightarrow +\infty$, alors la courbe admet une branche parabolique dans la direction D_{θ_0} .
- Au voisinage de ∞ : on distingue plusieurs cas :
 - Si $\rho(\theta) \rightarrow \pm\infty$, alors la courbe présente une branche infinie en forme de spirale (qui s'écarte de plus en plus de l'origine).
 - Si $\rho(\theta) \rightarrow a \neq 0$, alors la courbe est une spirale qui s'enroule autour du cercle de centre O et de rayon $|a|$, appelé cercle asymptote de la courbe \mathcal{C} .
 - Si $\rho(\theta) \rightarrow 0$, alors la courbe se présente comme une spirale qui s'enroule autour du point O, qui est aussi appelé point asymptote.

▷ *Points stationnaires :*

Sur les formules $x = \rho(\theta) \cos \theta$, $y = \rho(\theta) \sin \theta$, on voit que si le point de paramètre θ est un point stationnaire, alors nécessairement

$$\begin{aligned} \rho'(\theta) \cos \theta - \rho(\theta) \sin \theta &= 0, \\ \rho'(\theta) \sin \theta + \rho(\theta) \cos \theta &= 0. \end{aligned}$$

En multipliant la première équation par $\sin \theta$, la seconde par $\cos \theta$ et en soustrayant, on en déduit que $\rho(\theta) = 0$, puis que $\rho'(\theta) = 0$.

▷ *Direction du vecteur tangent :*

Si le point de paramètre θ n'est pas un point stationnaire, alors un vecteur tangent à la courbe est

$$\vec{T}(\theta) = \begin{pmatrix} \rho'(\theta) \cos \theta - \rho(\theta) \sin \theta \\ \rho'(\theta) \sin \theta + \rho(\theta) \cos \theta \end{pmatrix}.$$

Si $\rho(\theta) = 0$, on en déduit que le vecteur $\begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ est un vecteur tangent à la courbe. Si $\rho(\theta) \neq 0$, alors

$$\vec{T}(\theta) = \frac{\rho'(\theta)}{\rho(\theta)} O\vec{M} + O\vec{M}^\perp,$$

où

$$O\vec{M}^\perp = \begin{pmatrix} -\rho(\theta) \sin \theta \\ \rho(\theta) \cos \theta \end{pmatrix}.$$

Comme $O\vec{M}$ et $O\vec{M}^\perp$ sont orthogonaux, on voit sur la formule précédente que $\vec{T}(\theta)$ fait un angle α avec le vecteur $O\vec{M}$ tel que

$$\begin{aligned} \tan \alpha &= \frac{\rho(\theta)}{\rho'(\theta)} \quad \text{si } \rho'(\theta) \neq 0, \\ \alpha &= \frac{\pi}{2} \quad \text{sinon.} \end{aligned}$$

Exemple : tracé de la courbe d'équation $r = \cos(2\theta)$

1. Réduction du domaine d'étude :

Soit $\rho : \theta \in \mathbb{R} \mapsto \cos(2\theta)$. La fonction ρ est de classe \mathcal{C}^∞ , elle est paire et elle vérifie

$$\rho(\pi - \theta) = \rho(\theta).$$

La courbe d'équation $r = \cos(2\theta)$ est donc symétrique par rapport à l'axe des abscisses et par rapport à l'axe des ordonnées. On peut donc se contenter de tracer la courbe sur l'intervalle $[0, \pi/2]$, puis déduire le reste de la courbe par symétrie.

2. Variations de la fonction ρ :

On a

$$\rho'(\theta) = -2 \sin(2\theta),$$

ce qui donne le tableau de variations suivant pour ρ :

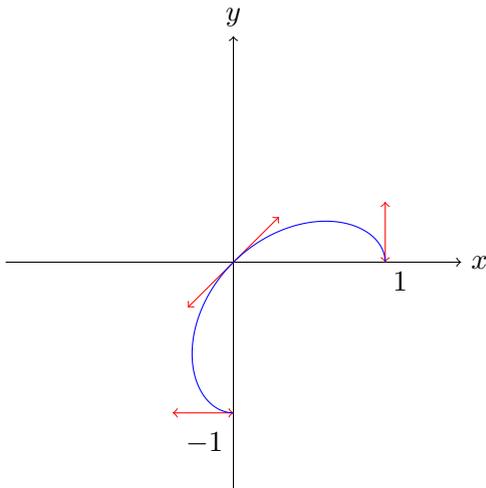
θ	0	$\pi/4$	$\pi/2$
$\rho'(\theta)$	0	—	0
ρ	1		

En particulier, sur l'intervalle $[0, \pi/2]$, on a $\rho(\theta) > 0$ si $\theta \in [0, \pi/4[$, $\rho(\theta) < 0$ si $\theta \in]\pi/4, \pi/2]$.

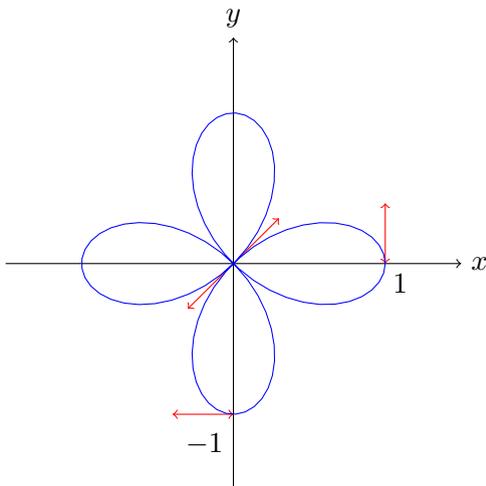
3. Il n'y a pas de branches infinies.

4. On n'a jamais d'annulation simultanée de $\rho(\theta)$ et de $\rho'(\theta)$, donc il n'y a pas de point stationnaire.
5. Pente de la tangente :
- En $\theta = 0$ et $\theta = \pi/2$, on a $\rho'(\theta) = 0$, donc la tangente fait un angle droit avec la droite $(OM(\theta))$. Ainsi la courbe admet une tangente verticale en $\theta = 0$ et une tangente horizontale en $\theta = \pi/2$.
 - En $\theta = \pi/4$, on a $r(\theta) = 0$, et donc le vecteur $\frac{\sqrt{2}}{2}(1, 1)$ est un vecteur tangent à la courbe.
 - Pour $\theta \in]0, \pi/2[$ et $\theta \neq \pi/4$, on a $\rho'(\theta) \neq 0$, $\rho(\theta) \neq 0$, donc la tangente fait un angle $\alpha = -\arctan \frac{\cos(2\theta)}{2\sin(2\theta)}$ avec la droite $(OM(\theta))$.

Sur l'intervalle $[0, \pi/2]$, on en déduit que la courbe a l'allure suivante :



Par symétrie, on en déduit l'allure de la courbe sur l'ensemble de l'intervalle $[0, \pi]$:



1.5 Lien avec les équations différentielles

Supposons que le couple (f, g) de fonctions intervenant dans la courbe paramétrée

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2, \exists t \in I, x = f(t), y = g(t)\}$$

est une solution de l'équation différentielle

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

où $X(t) = (f(t), g(t))$, $b : I \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$.

Les propositions vues au paragraphe précédent ont des conséquences immédiates :

Corollaire 1.5.1. *Soit $a \in I$ et $(x_0, y_0) \in \mathbb{R}^2$ tel que :*

- $M(a) = (x_0, y_0)$;
- $b(a, (x_0, y_0)) \neq 0$.

Alors $b(a, (x_0, y_0))$ est un vecteur tangent à la courbe en (x_0, y_0) .

En particulier, si b est indépendant du temps, c'est-à-dire s'il existe une fonction $\bar{b} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ telle que

$$b(t, X) = \bar{b}(X) \quad \forall t \in I, \forall X \in \mathbb{R}^2,$$

alors en tout point $M \in \mathcal{C}$ tel que $\bar{b}(M) \neq 0$, $\bar{b}(M)$ est un vecteur tangent à la courbe.

1.5.1 Un peu de vocabulaire

Définition 1.5.2. *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 1.5.3. *Donner ici des exemples dans lesquels on identifie la fonction f : cas autonomes et non autonomes.*

Définition 1.5.4. *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 $F(X_0) \neq 0$. 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 du paragraphe précédent que le vecteur $F(X_0)$ est un vecteur tangent à la courbe paramétrée $\mathcal{C} = \{(x, y) \in \mathbb{R}^2, \exists t \in \mathbb{R}, (x, y) = X(t)\}$. Ainsi la représentation du champ de vecteur F évoquée ci-dessus donne une idée de l'allure des courbes représentant les solutions de l'équation différentielle : il s'agit en réalité de tracer les tangentes à la courbe.

Définition 1.5.5. *Les solutions de l'équation différentielle $X'(t) = F(X(t))$ sont appelées courbes intégrales du champ de vecteurs.*

1.5.2 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 :

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

Supposons que le mouvement se fasse sur une droite (ou 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

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} = \begin{pmatrix} x'(t) \\ \frac{1}{m}\varphi(x(t), x'(t)) \end{pmatrix} = b(x(t), x'(t)), \quad (1.7)$$

où

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

On obtient donc un système différentiel du premier ordre dans \mathbb{R}^2 . Comme cela a été expliqué au paragraphe précédent, 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 (1.7) pour chaque donnée initiale).

Exemple 1.5.6 (L'oscillateur harmonique). Soit $\omega_0 > 0$. On considère l'équation différentielle

$$x''(t) + \omega_0^2 x(t) = 0, \quad (1.8)$$

muni des conditions initiales

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

On sait alors que la forme générale des solutions de (1.8) 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 (1.8), 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

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

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

$$(x, y) \in C_{x_0, v_0} \iff x^2 + \frac{y^2}{\omega_0^2} = x_0^2 + \frac{v_0^2}{\omega_0^2}.$$

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

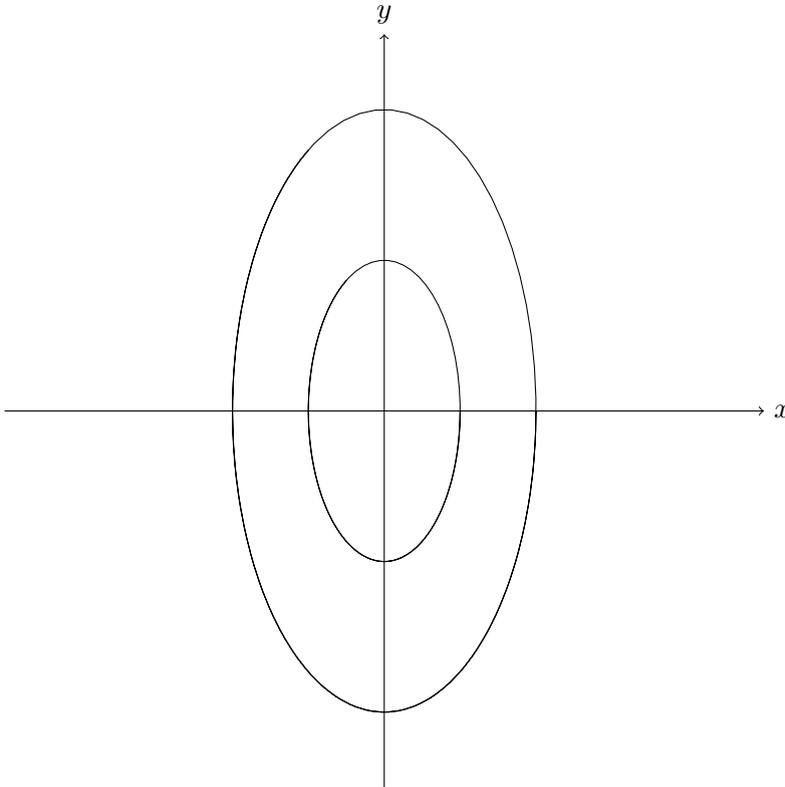
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

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) .

Dessin dans le cas $\omega_0 = 2$:



Chapitre 2

Systèmes d'équations différentielles linéaires

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

$$x^{(n)}(t) = A_0(t) + A_1(t)x'(t) + \dots + A_{n-1}(t)x^{(n-1)}(t),$$

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

Le cas où $m = n = 1$ a déjà été étudié en L1, et on commence ce chapitre par un rappel des principaux résultats dans ce cadre. On se concentre ensuite sur le cas $n = 1$ et $m = 2$, ou $n = 2$ et $m = 1$, en gardant en tête que les cas n, m quelconques se traitent de façon analogue. On expliquera en particulier comment calculer des solutions exactes dans le cas où A_0, \dots, A_{n-1} ne dépendent pas de t , et on présentera dans ce cas simple des exemples de calculs de solutions approchées.

2.1 Analyse théorique

2.1.1 Rappels sur les équations différentielles linéaires 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

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

ainsi que sa version 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

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

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

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

$$x(t) = x_0 \exp\left(\int_{t_0}^t a(t') dt'\right).$$

Démonstration. – Existence : Posons

$$\tilde{x}(t) = x_0 \exp\left(\int_{t_0}^t a(t') dt'\right).$$

On vérifie immédiatement que $\tilde{x}(t_0) = x_0 \exp(0) = x_0$.

Montrons maintenant que \tilde{x} est bien solution de l'équation (2.2). On a

$$\tilde{x}'(t) = x_0 a(t) \exp\left(\int_{t_0}^t a(t') dt'\right) = a(t) \tilde{x}(t).$$

Donc \tilde{x} 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

$$y(t) = x(t) \exp\left(-\int_{t_0}^t a(t') dt'\right)$$

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

$$y'(t) = x'(t) \exp\left(-\int_{t_0}^t a(t') dt'\right) - x(t) a(t) \exp\left(-\int_{t_0}^t a(t') dt'\right) = 0.$$

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

$$x(t) = y(t) \exp\left(\int_{t_0}^t a(t') dt'\right) = x_0 \exp\left(\int_{t_0}^t a(t') dt'\right).$$

□

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

Pour la solution de l'équation non-homogène (2.1), il n'y a pas de formule générale à retenir, mais plutôt une méthode. L'idée est d'utiliser le calcul fait dans le cas homogène pour changer de fonction inconnue et "tuer" 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

$$y(t) = x(t) \exp\left(-\int_{t_0}^t a(t') dt'\right),$$

ce qui revient à dire

$$x(t) = y(t) \exp\left(\int_{t_0}^t a(t') dt'\right).$$

On a alors

$$x'(t) = (y'(t) + a(t)y(t)) \exp\left(\int_{t_0}^t a(t') dt'\right),$$

et donc

$$x'(t) - a(t)x(t) = y'(t) \exp\left(\int_{t_0}^t a(t') dt'\right).$$

On en déduit que

$$y'(t) = b(t) \exp\left(-\int_{t_0}^t a(t') dt'\right).$$

Comme précédemment, on suppose que $x(t_0) = x_0 = y(t_0)$. On a alors

$$y(t) = x_0 + \int_{t_0}^t b(s) \exp\left(-\int_{t_0}^s a(t') dt'\right) ds.$$

Résolution de l'équation (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 2.1.2. Soit $t_0 \in I$, $x_0, x_1 \in \mathbb{R}$.

Soit $\bar{x} : I \rightarrow \mathbb{R}$ de classe C^1 telle que \bar{x} est une solution de l'équation (2.1) et $\bar{x}(t_0) = x_1$. Alors l'unique solution de (2.1) telle que $x(t_0) = x_0$ est donnée par

$$x(t) = \bar{x}(t) + (x_0 - x_1) \exp\left(\int_{t_0}^t a(t') dt'\right).$$

Démonstration. Pour l'unicité, on remarque que pour toute solution x de l'équation (2.1), la fonction $x - \bar{x}$ est solution de (2.2) avec $(x - \bar{x})(t_0) = x_0 - x_1$. D'après l'unicité dans le cas homogène, on a donc nécessairement

$$x(t) - \bar{x}(t) = (x_0 - x_1) \exp\left(\int_{t_0}^t a(t') dt'\right).$$

Il s'agit donc de montrer que la fonction $\tilde{x} : t \mapsto \bar{x}(t) + (x_0 - x_1) \exp\left(\int_{t_0}^t a(t') dt'\right)$ est bien solution de l'équation. On a

$$\tilde{x}(t_0) = \bar{x}(t_0) + (x_0 - x_1) = x_1 + x_0 - x_1 = x_0.$$

Par ailleurs,

$$\begin{aligned} \tilde{x}'(t) &= \bar{x}'(t) + (x_0 - x_1)a(t) \exp\left(\int_{t_0}^t a(t') dt'\right) \\ &= a(t)\bar{x}(t) + b(t) + (x_0 - x_1)a(t) \exp\left(\int_{t_0}^t a(t') dt'\right) \\ &= a(t)\tilde{x}(t) + b(t). \end{aligned}$$

□

Exemples

- On prend $I = \mathbb{R}$, $a(t) = r \in \mathbb{R}$ pour tout $t \in \mathbb{R}$. Alors les solutions de l'équation (2.1) sont données par

$$x(t) = x_0 \exp(r(t - t_0)).$$

- On prend $I = \mathbb{R}_+^*$, $a(t) = k/t$ pour $k \in \mathbb{R}^*$. Alors les solutions de l'équation (2.1) sont données par

$$x(t) = x_0 \exp\left(k \ln \frac{t}{t_0}\right) = x_0 \left(\frac{t}{t_0}\right)^k.$$

- 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 $\bar{x}(t) = t$. La solution générale de l'équation est donc donnée par

$$x(t) = t + C \exp(-t^2/2),$$

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

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

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

La solution générale de l'équation homogène est de la forme

$$h(t) = C \exp\left(\int_{t_0}^t \frac{1}{t'} dt'\right) = C_1 t.$$

On pose donc $y(t) = x(t)/t$. D'après la méthode de variation de la constante, y est solution de

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

d'où $y(t) = \frac{t^2}{2} + y_0$. On obtient finalement

$$x(t) = \frac{t^3}{2} + y_0 t.$$

Il faut ensuite choisir la constante y_0 de façon à ce que la condition "initiale" en $t_0 > 0$ soit vérifiée.

2.1.2 Équation différentielle linéaire 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 . Le résultat suivant découle d'un théorème qui sera vu au Chapitre 3 :

Théorème 2.1.3. *Soit $t_0 \in \mathbb{R}$, et soit $x_0, x_1 \in \mathbb{R}$. Alors il existe une unique solution de (2.5) telle que*

$$x(t_0) = x_0, \quad x'(t_0) = x_1. \quad (2.4)$$

Cette solution est définie sur \mathbb{R} .

Cas d'une équation à 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 (2.4) : 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 la bonne.

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

Théorème 2.1.4. *Soit $a, b, x_0, x_1 \in \mathbb{R}$. On note x la solution de l'équation (2.5) munie des conditions initiales (2.4).*

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

$$\lambda^2 + a\lambda + b = 0. \quad (2.6)$$

– Premier cas : $\Delta \neq 0$: on note λ_1, λ_2 les deux racines réelles ou complexes de l'équation (2.6). Alors il existe $A_1, A_2 \in \mathbb{C}$ tels que

$$x(t) = A_1 \exp(\lambda_1 t) + A_2 \exp(\lambda_2 t).$$

– Deuxième cas : $\Delta = 0$: soit λ_0 l'unique racine réelle de (2.6) (il s'agit d'une racine double). Alors il existe $A, B \in \mathbb{C}$ tels que

$$x(t) = (At + B) \exp(\lambda_0 t).$$

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

On pose

$$\bar{x}(t) = A_1 \exp(\lambda_1 t) + A_2 \exp(\lambda_2 t) \text{ si } \Delta \neq 0,$$

et

$$\bar{x}(t) = (At + B) \exp(\lambda_0 t) \text{ si } \Delta = 0.$$

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

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

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

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

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

et donc en utilisant le fait que $b = a^2/4$ et $\lambda_0 = -a/2$,

$$\begin{aligned} & \bar{x}''(t) + a\bar{x}'(t) + b\bar{x}(t) \\ &= \exp(\lambda_0 t) \left[\lambda_0^2 At + \lambda_0^2 B + 2\lambda_0 A + a(\lambda_0 At + \lambda_0 B + A) + b(At + B) \right] \\ &= \exp(\lambda_0 t) \left[At(\lambda_0^2 + a\lambda_0 + b) + B(\lambda_0^2 + a\lambda_0 + b) + A(2\lambda_0 + a) \right] \\ &= 0. \end{aligned}$$

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 en $t = 0$ soient vérifiées

Il s'agit de montrer que les systèmes linéaires

$$\begin{aligned} A_1 \exp(\lambda_1 t_0) + A_2 \exp(\lambda_2 t_0) &= x_0, \\ A_1 \lambda_1 \exp(\lambda_1 t_0) + A_2 \lambda_2 \exp(\lambda_2 t_0) &= x_1 \end{aligned} \tag{2.7}$$

dans le cas $\Delta \neq 0$, ou

$$\begin{aligned} (At_0 + B) \exp(\lambda_0 t_0) &= x_0, \\ (\lambda_0 At_0 + \lambda_0 B + A) \exp(\lambda_0 t_0) &= x_1 \end{aligned} \tag{2.8}$$

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

Le déterminant du système (2.7) 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 (2.8) (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 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 \bar{x} soit une solution de (2.5) munie de la condition initiale (2.4). Il reste à montrer que la fonction \bar{x} 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 (2.7) (resp. de (2.8)) est dans \mathbb{R}^2 . Dans ces cas il est donc clair que $\bar{x}(t) \in \mathbb{R}$ 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 (2.7) et en utilisant le fait que

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

on en déduit que (\bar{A}_2, \bar{A}_1) est solution du même système que (A_1, A_2) . Par unicité des solutions de (2.7), on obtient $A_2 = \bar{A}_1$, et donc

$$\bar{x}(t) = A_1 \exp(\lambda_1 t) + A_2 \exp(\lambda_2 t) = 2\Re(A_1 \exp(\lambda_1 t)).$$

Quatrième étape : conclusion.

Ainsi $\bar{x} : \mathbb{R} \rightarrow \mathbb{R}$ est une solution de (2.5) qui vérifie (2.4). D'après le théorème 2.1.3, $x = \bar{x}$. □

Passage à un système d'équations d'ordre un

On considère à présent 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

$$X'(t) = A(t)X(t) + F(t). \tag{2.9}$$

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 (2.9). 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 2.1.5. 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

$$x \text{ est solution de (2.3)} \Leftrightarrow X \text{ est solution de (2.9)}.$$

Remarque 2.1.6. 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. Ainsi, dans la suite de cette partie, on étudiera uniquement des systèmes de deux équations linéaires d'ordre un (cas $m = 2, n = 1$).

2.1.3 Systèmes homogènes d'ordre un à coefficients constants : exponentielle de matrice

Les deux 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 présent paragraphe. Si $B \neq 0$, le système est non-homogène, et son étude fera l'objet du paragraphe 2.1.4.

Dans la suite de ce paragraphe, on se concentre sur le cas $B = 0$.

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

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

sont de la forme $y(t) = y_0 \exp(at)$. L'idée est ici de généraliser la définition de l'exponentielle aux matrices de $\mathcal{M}_2(\mathbb{R})$.

Théorème 2.1.7 (Exponentielle de matrice). Soit $A \in \mathcal{M}_2(\mathbb{C})$. Pour $n \in \mathbb{N}$, on définit

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

On note, pour $n \in \mathbb{N}$,

$$M_n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}, \quad A^n = \begin{pmatrix} \alpha_n & \beta_n \\ \gamma_n & \delta_n \end{pmatrix}.$$

Alors :

1. On pose $C_0 = 2 \max(|\alpha_1|, |\beta_1|, |\gamma_1|, |\delta_1|)$. On a, pour tout $k \in \mathbb{N}$,

$$\max(|\alpha_k|, |\beta_k|, |\gamma_k|, |\delta_k|) \leq C_0^k. \quad (2.10)$$

2. Les suites $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$, $(c_n)_{n \in \mathbb{N}}$, $(d_n)_{n \in \mathbb{N}}$ sont de Cauchy dans \mathbb{C} , donc convergentes. On note $\bar{a}, \bar{b}, \bar{c}, \bar{d}$ leurs limites respectives. On définit l'exponentielle de la matrice A , et on note $\exp(A)$ ou e^A , la matrice définie par

$$\exp(A) = \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}.$$

Remarque 2.1.8. *Cé théorème, ainsi que l'ensemble des résultats qui suivent concernant l'exponentielle de matrice, peuvent se généraliser en dimension n quelconque (et donc à des équations différentielles à coefficients constants d'ordre arbitrairement grand).*

Démonstration. Montrons tout d'abord l'inégalité (2.10) par récurrence sur $k \in \mathbb{N}$:

- Initialisation : si $k = 0$ ou $k = 1$, l'inégalité est clairement vérifiée par définition de C_0 .
- Récurrence : soit $k \geq 1$ quelconque. Supposons l'inégalité vérifiée au rang $k - 1$. On a

$$\begin{aligned} A^k &= AA^{k-1} \\ &= \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \begin{pmatrix} \alpha_{k-1} & \beta_{k-1} \\ \gamma_{k-1} & \delta_{k-1} \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1\alpha_{k-1} + \beta_1\gamma_{k-1} & \alpha_1\beta_{k-1} + \beta_1\delta_{k-1} \\ \gamma_1\alpha_{k-1} + \delta_1\gamma_{k-1} & \gamma_1\beta_{k-1} + \delta_1\delta_{k-1} \end{pmatrix}. \end{aligned}$$

On a donc par exemple

$$|\alpha_k| \leq |\alpha_1| |\alpha_{k-1}| + |\beta_1| |\gamma_{k-1}|.$$

En utilisant l'hypothèse de récurrence, on obtient

$$|\alpha_k| \leq |\alpha_1| C_0^{k-1} + |\beta_1| C_0^{k-1} \leq 2 \max(|\alpha_1|, |\beta_1|) C_0^{k-1} \leq C_0^k.$$

L'inégalité est donc vérifiée au rang k .

Par récurrence, l'inégalité (2.10) est vérifiée pour tout $k \geq 0$.

Montrons à présent que la suite $(a_n)_{n \in \mathbb{N}}$ est de Cauchy : par définition, on a

$$a_n = \sum_{k=0}^n \frac{\alpha_k}{k!}.$$

Par conséquent, si $n \geq p \geq 0$,

$$|a_n - a_p| \leq \sum_{k=p+1}^n \frac{|\alpha_k|}{k!}$$

et d'après l'inégalité (2.10),

$$|a_n - a_p| \leq \sum_{k=p+1}^n \frac{C_0^k}{k!} \quad \forall n \geq p \geq 0.$$

La série $\sum \frac{C_0^k}{k!}$ est une série à termes positifs, convergente (sa somme est e^{C_0}). On en déduit que la suite $(\sum_{k=0}^n C_0^k/k!)_{n \in \mathbb{N}}$ est de Cauchy, et par conséquent la suite $(a_n)_{n \in \mathbb{N}}$ est de Cauchy.

On montre de même que les suites $(b_n)_{n \in \mathbb{N}}$, $(c_n)_{n \in \mathbb{N}}$, $(d_n)_{n \in \mathbb{N}}$ sont de Cauchy.

Comme \mathbb{R} est complet, les suites $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$, $(c_n)_{n \in \mathbb{N}}$, $(d_n)_{n \in \mathbb{N}}$ sont convergentes, ce qui conclut la preuve du théorème. □

Proposition 2.1.9 (Propriétés de l'exponentielle de matrice). *Soit $A \in \mathcal{M}_2(\mathbb{C})$, et soit $\lambda, \mu \in \mathbb{R}$. On note $0_2 \in \mathcal{M}_2(\mathbb{R})$ la matrice nulle, I_2 la matrice identité.*

1. $\exp(0_2) = I_2$;

2. $Ae^A = e^A A$;
3. $\exp(\lambda A) \exp(\mu A) = \exp((\lambda + \mu)A) = \exp(\mu A) \exp(\lambda A)$;
4. $\exp(A) \exp(-A) = I_2$;
5. $\exp({}^t A) = {}^t(\exp(A))$

Démonstration. 1. Pour tout $n \in \mathbb{N}$, on a

$$\sum_{k=0}^n \frac{0_2^k}{k!} = I_2.$$

En passant à la limite quand $n \rightarrow \infty$, on obtient la première propriété.

2. Notons que pour tout $k \in \mathbb{N}$,

$$A \cdot A^k = A^k \cdot A = A^{k+1}.$$

On en déduit que pour tout $n \in \mathbb{N}$,

$$A \sum_{k=0}^n \frac{A^k}{k!} = \sum_{k=0}^n \frac{A^{k+1}}{k!} = \sum_{k=0}^n \frac{A^k}{k!} A.$$

En passant à la limite quand $n \rightarrow \infty$, on en déduit que

$$A \exp(A) = \exp(A) A.$$

3. On a

$$\begin{aligned} \exp(\lambda A) \exp(\mu A) &= \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n \frac{(\lambda A)^k}{k!} \right) \left(\sum_{k=0}^n \frac{(\mu A)^k}{k!} \right) \\ &= \lim_{n \rightarrow \infty} \sum_{0 \leq l, m \leq n} \frac{\lambda^l \mu^m A^{l+m}}{l! m!} \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^{2n} A^k \sum_{\substack{l, m \geq 0, \\ l+m=k}} \frac{\lambda^l \mu^m}{l! m!}. \end{aligned}$$

Or pour tout $k \in \mathbb{N}$, d'après la formule du binôme de Newton,

$$\begin{aligned} \sum_{\substack{l, m \geq 0, \\ l+m=k}} \frac{\lambda^l \mu^m}{l! m!} &= \frac{1}{k!} \sum_{\substack{l, m \geq 0, \\ l+m=k}} \frac{k!}{l! m!} \lambda^l \mu^m \\ &= \frac{1}{k!} \sum_{l=0}^k \frac{k!}{l!(k-l)!} \lambda^l \mu^{k-l} \\ &= \frac{1}{k!} (\lambda + \mu)^k. \end{aligned}$$

En revenant à l'expression initiale, on en déduit

$$\begin{aligned} \exp(\lambda A) \exp(\mu A) &= \lim_{n \rightarrow \infty} \sum_{k=0}^{2n} \frac{(\lambda + \mu)^k A^k}{k!} \\ &= \exp((\lambda + \mu)A). \end{aligned}$$

Comme

$$\exp((\lambda + \mu)A) = \exp((\mu + \lambda)A),$$

on obtient finalement l'identité voulue.

4. La propriété découle de la première et de la troisième avec $\lambda = -\mu$.
 5. La propriété découle de la formule

$${}^t(A^k) = ({}^tA)^k$$

pour tout k .

□

Remarque 2.1.10. *Attention! En général on n'a PAS*

$$\exp(A)\exp(B) = \exp(A+B).$$

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

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

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

$$\exp(A) = I_2 + A, \quad \exp(B) = I_2 + B.$$

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

$$\begin{aligned} \exp(A)\exp(B) &= I_2 + A + B + AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \\ \exp(B)\exp(A) &= I_2 + A + B + BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}. \end{aligned}$$

On a donc déjà $e^A e^B \neq e^B e^A$, ce qui implique que l'égalité $\exp(A)\exp(B) = \exp(A+B)$ ne peut être vraie.

Calculons néanmoins $\exp(A+B)$ pour en avoir le cœur net : on a

$$A+B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (A+B)^2 = I_2.$$

On en déduit que

$$(A+B)^k = \begin{cases} A+B & \text{si } k \text{ est impair,} \\ I_2 & \text{si } k \text{ est pair.} \end{cases}$$

On obtient donc finalement

$$\exp(A+B) = \left(\sum_{k=0}^{\infty} \frac{1}{(2k)!} \right) I_2 + \left(\sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \right) (A+B).$$

La matrice $\exp(A+B)$ est donc de la forme

$$\begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix},$$

où $\alpha = \sum_{k=0}^{\infty} \frac{1}{(2k)!}$, $\beta = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}$. On a donc bien

$$\exp(A)\exp(B) \neq \exp(A+B).$$

En utilisant les propriétés précédentes, on arrive finalement au résultat qui va permettre de trouver les solutions de l'équation $y'(t) = Ay(t)$.

Lemme 2.1.11. *Soit $A \in \mathcal{M}_2(\mathbb{R})$. Alors l'application*

$$t \in \mathbb{R} \mapsto \exp(tA)$$

est de classe \mathcal{C}^1 , et

$$\frac{d \exp(tA)}{dt} = A \exp(tA) \quad \forall t \in \mathbb{R}.$$

Démonstration. Soit $t \in \mathbb{R}, h \in \mathbb{R} \setminus 0$. On écrit, à l'aide de la proposition précédente,

$$\frac{\exp((t+h)A) - \exp(tA)}{h} = \exp(tA) \frac{\exp(hA) - I_2}{h}.$$

Or

$$\frac{\exp(hA) - I_2}{h} = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{h^k A^{k+1}}{(k+1)!},$$

et donc

$$\frac{\exp(hA) - I_2}{h} = A + \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{h^k}{(k+1)!} \begin{pmatrix} \alpha_{k+1} & \beta_{k+1} \\ \gamma_{k+1} & \delta_{k+1} \end{pmatrix},$$

en reprenant les notations de la preuve du Théorème 2.1.7. D'après le théorème 2.1.7, on a

$$\begin{aligned} \left| \sum_{k=1}^n \frac{h^k}{(k+1)!} \alpha_{k+1} \right| &\leq \left| \sum_{k=1}^n \frac{|h|^k}{(k+1)!} C_0^{k+1} \right| \\ &\leq |h| C_0^2 \sum_{k=0}^{n-1} \frac{(|h| C_0)^k}{(k+2)!} \\ &\leq |h| C_0^2 \exp(|h| C_0). \end{aligned}$$

La même inégalité est vraie si on remplace α_k par $\beta_k, \gamma_k, \delta_k$. Si on suppose de surcroît que h appartient à un voisinage de zéro, par exemple $|h| \leq 1$, on obtient finalement

$$\frac{\exp((t+h)A) - \exp(tA)}{h} = \exp(tA)(A + O(h)).$$

On en déduit que l'application

$$t \mapsto \exp(tA)$$

est dérivable sur \mathbb{R} (donc continue), et que sa dérivée $t \mapsto A \exp(tA)$ est également continue. L'application considérée est donc de classe \mathcal{C}^1 . \square

On arrive donc finalement au résultat suivant :

Théorème 2.1.12. *Soit $X_0 \in \mathbb{R}^2, A \in \mathcal{M}_2(\mathbb{R}), I$ un intervalle de \mathbb{R} contenant 0.*

Soit $X : I \rightarrow \mathbb{R}$ une solution du système différentiel

$$X'(t) = AX(t) \quad \forall t \in \mathbb{R}, \quad X(0) = X_0. \quad (2.11)$$

Alors la définition de X peut être étendue à \mathbb{R} et

$$X(t) = \exp(At)X_0 \quad \forall t \in \mathbb{R}$$

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

Démonstration. Le point remarquable ici est que l'on va montrer l'unicité des solutions sans avoir recours au chapitre 3.

- Première étape : on vérifie que la fonction \bar{X} définie par

$$\bar{X}(t) = \exp(At)X_0 \quad \forall t \in \mathbb{R}$$

est solution du même système différentiel que X , avec la même condition initiale $\bar{X}(0) = X_0$. Cela découle immédiatement du Lemme 2.1.11.

- Deuxième étape : unicité : on définit à présent, pour $t \in I$,

$$Y(t) = \exp(-At)X(t).$$

Alors Y est dérivable sur I et

$$\frac{dY(t)}{dt} = -A \exp(-At)X(t) + \exp(-At)X'(t) = -\exp(-At)AX(t) + \exp(-At)AX(t),$$

puisque A et $\exp(-At)$ commutent et que $X'(t) = AX(t)$ pour tout t .

On obtient donc $Y'(t) = 0$ pour tout $t \in I$, donc Y est une fonction constante : $Y(t) = Y(0) = X_0$. En revenant à la définition de Y , cela implique

$$\exp(-At)X(t) = X_0 \quad \forall t \in I.$$

On applique $\exp(At)$ aux deux membres de l'équation et on obtient $X = \bar{X}$. □

Remarque 2.1.13. *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

$$X(t) = \exp\left(\int_0^t A(s)ds\right)y_0.$$

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

2.1.4 Systèmes non-homogènes d'ordre un à coefficients constants : méthode de variation des constantes

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 utilise la méthode de variation des constantes. Plus précisément, soit X la solution de l'équation

$$\begin{aligned} X'(t) &= AX(t) + B(t) \quad \forall t \in \mathbb{R}, \\ X(0) &= X_0. \end{aligned}$$

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

$$Z(t) = \exp(-At)X(t).$$

Alors

$$\forall t \in \mathbb{R} \quad Z'(t) = -A \exp(-At)X(t) + \exp(-At)X'(t) = \exp(-At)B(t).$$

Il "suffit" donc de trouver une primitive de $\exp(-At)B(t)$ pour connaître $Z(t)$. On obtient alors X par la formule

$$X(t) = \exp(At)Z(t).$$

Remarque 2.1.14. *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.*

2.1.5 Exemples de calculs d'exponentielle de matrice

Commençons par quelques exemples simples :

– **Matrice diagonale :**

Soit $\lambda_1, \lambda_2 \in \mathbb{C}$, et soit $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Alors $\exp(A) = \begin{pmatrix} \exp(\lambda_1) & 0 \\ 0 & \exp(\lambda_2) \end{pmatrix}$.

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

$$\exp(A) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} = \begin{pmatrix} \exp(\lambda_1) & 0 \\ 0 & \exp(\lambda_2) \end{pmatrix}.$$

– **Matrice nilpotente :**

Soit

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

On vérifie aisément que $N^2 = 0$. Par suite, $N^k = 0$ pour tout $k \geq 2$. On obtient donc

$$\exp(N) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} N^k = I_2 + N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

– **Matrices semblables :**

Soit $A, B \in \mathcal{M}_2(\mathbb{C})$, $P \in GL_2(\mathbb{C})$ telles que

$$B = PAP^{-1}.$$

Alors

$$\exp(B) = P \exp(A) P^{-1}.$$

En effet, on vérifie facilement (par exemple par récurrence) que si $k \in \mathbb{N}$,

$$B^k = PA^k P^{-1}.$$

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.

– **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

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

En ce cas,

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

Pour traiter le cas général, on a besoin d'un résultat d'algèbre linéaire supplémentaire.

Trigonalisation sur $\mathcal{M}_2(\mathbb{C})$: décomposition de Dunford

Théorème 2.1.15 (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 est trigonalisable (mais pas diagonalisable) : il existe $\lambda \in \mathbb{C}$, $\alpha \in \mathbb{C}$, $P \in GL_2(\mathbb{C})$ tels que*

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

Démonstration. On rappelle que les valeurs propres de A sont les racines de son polynôme caractéristique $\chi_A \in \mathbb{C}[X]$. 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

$$\dim \ker(A - \lambda_1 I_2) + \dim \ker(A - \lambda_2 I_2) = 2.$$

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 $\lambda_1 = \lambda_2$.

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 .

Soit P la matrice de passage de la base canonique à la base (v_1, v_2) . Comme $Av_1 = \lambda v_1$, on a

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

où $\alpha, \mu \in \mathbb{C}$. Le polynôme caractéristique est invariant par similitude, donc

$$\chi_A(X) = (X - \lambda)^2 = (X - \lambda)(X - \mu).$$

On en déduit que $\lambda = \mu$. On est donc dans le deuxième cas énoncé dans le théorème. \square

On peut calculer l'exponentielle de la matrice A à l'aide du théorème précédent. Le premier cas a déjà été traité dans les remarques qui précèdent le théorème. Pour le deuxième cas, remarquons que

$$A = P(\lambda I_2 + \alpha N)P^{-1},$$

et que $I_2 N = N I_2 = N$. Par conséquent, pour tout $k \geq 0$,

$$(\lambda I_2 + \alpha N)^k = \sum_{j=0}^k C_k^j \lambda^j \alpha^{k-j} N^{k-j} = \lambda^k I_2 + k \lambda^{k-1} \alpha N = \begin{pmatrix} \lambda^k & \alpha k \lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}.$$

On obtient donc, pour tout $n \geq 0$

$$\sum_{k=0}^n \frac{A^k}{k!} = P \begin{pmatrix} \sum_{k=0}^n \frac{\lambda^k}{k!} & \alpha \sum_{k=1}^n \frac{\lambda^{k-1}}{(k-1)!} \\ 0 & \sum_{k=0}^n \frac{\lambda^k}{k!} \end{pmatrix} P^{-1}.$$

En passant à la limite quand $n \rightarrow \infty$, on arrive finalement à

$$\exp(A) = P \begin{pmatrix} e^\lambda & \alpha e^\lambda \\ 0 & e^\lambda \end{pmatrix} P^{-1}.$$

Résumons les propriétés citées plus haut dans un tableau :

Type	Formule	Exponentielle
Matrice diagonale	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$	$\begin{pmatrix} \exp(\lambda_1) & 0 \\ 0 & \exp(\lambda_2) \end{pmatrix}$
Matrice nilpotente	$\begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$
Matrices semblables	$A = P B P^{-1}$	$\exp(A) = P \exp(B) P^{-1}$

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

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

En pratique, si $A \in \mathcal{M}_2(\mathbb{R})$, pour calculer son exponentielle, on considère A comme une matrice complexe. Notons que $\exp(A)$ sera malgré tout une matrice réelle, comme on peut le voir sur la formule

$$\exp(A) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{A^k}{k!}.$$

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 A est diagonalisable ou trigonalisable sur \mathbb{R} (autrement dit, dans le théorème 2.1.15, on a $P \in GL_2(\mathbb{R})$, $\alpha \in \mathbb{R}$).
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. 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} .

2.1.6 Lien entre les deux méthodes

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

$$f''(t) + af'(t) + bf(t) = 0, \quad f(0) = x_1, \quad f'(0) = y_2 \quad (2.12)$$

on a donc vu deux méthodes différentes :

1. Première méthode : recherche de solutions comme des combinaisons linéaires d'exponentielles :

On résout (sur \mathbb{C}) l'équation caractéristique $\lambda^2 + a\lambda + b = 0$, et on note λ_1, λ_2 ses racines complexes distinctes ou confondues. Si $\lambda_1 \neq \lambda_2$, la solution est de la forme

$$f(t) = A_1 \exp(\lambda_1 t) + A_2 \exp(\lambda_2 t), \quad \text{avec } A_1, A_2 \in \mathbb{C}.$$

Si $\lambda_1 = \lambda_2$, la solution est de la forme

$$f(t) = (A_1 t + A_2) \exp(\lambda_1 t), \quad \text{avec } A_1, A_2 \in \mathbb{C}.$$

2. Deuxième méthode : on pose

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

On a alors

$$X(t) = \exp(At)X_0,$$

où $X_0 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Vérifions que les deux formules donnent le même résultat. 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 (2.12) 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.

- Premier cas : $\lambda_1 \neq \lambda_2$:

Dans ce cas A est diagonalisable sur \mathbb{C} et il existe $P \in GL_2(\mathbb{C})$ tel que

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

On a alors

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

Notons

$$P^{-1} X_0 = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad P = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}.$$

On en déduit que

$$f(t) = a_1 \exp(\lambda_1 t) z_1 + a_2 \exp(\lambda_2 t) z_2,$$

et on retrouve le résultat de la première méthode dans le cas $\lambda_1 \neq \lambda_2$.

– Deuxième cas : $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$:

Dans ce cas A est trigonalisable sur \mathbb{R} (voir la fin du paragraphe précédent), et il existe donc $P \in GL_2(\mathbb{R})$, $\alpha \in \mathbb{R}$ tel que

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

de sorte que

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

Avec les mêmes notations que ci-dessus, on obtient

$$f(t) = \exp(\lambda t) [(a_1 z_1 + a_2 z_2) + a_1 z_2 \alpha t],$$

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

2.1.7 Portraits de phase dans le cas linéaire

L'objet de tout ce paragraphe est de tracer les courbes paramétrées définies par

$$\begin{cases} x = f(t), \\ y = f'(t), \end{cases} \quad t \in \mathbb{R},$$

où f est solution de (2.12). Dans tout ce paragraphe, on utilisera plutôt la deuxième méthode (avec l'exponentielle de matrice). Il s'agit donc de tracer des courbes paramétrées définies par

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

Préliminaires

Avant de comprendre le cas général, on commence par des cas simples : matrice diagonale, matrice triangulaire supérieure avec une seule valeur propre, ainsi qu'un cas "modèle" correspondant à une matrice diagonalisable dans \mathbb{C} mais pas dans \mathbb{R} . On donne également un lemme permettant de comprendre le lien entre :

- la similitude de deux matrices A et B ;
- et le tracé des courbes paramétrées associées à $X(t) = \exp(At)X_0$, $Y(t) = \exp(Bt)Y_0$.

1. *Tracé de la courbe paramétrée d'équation* $\begin{pmatrix} x \\ y \end{pmatrix} = \exp(At)X_0$, *avec* A *diagonale et* $X_0 \in \mathbb{R}^2$:

On a

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

avec $\lambda_1, \lambda_2 \in \mathbb{R}$. On pose également $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. On étudie donc la courbe paramétrée \mathcal{C}_{A, X_0} d'équation

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

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 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.

De plus, si $x_0 = 0$ (resp. $y_0 = 0$), alors on a $\exp(\lambda_1 t)x_0 = 0$ pour tout $t \in \mathbb{R}$ (resp. $\exp(\lambda_2 t)y_0 = 0$ pour tout t). On voit alors facilement que la trajectoire est le demi-axe (Oy) correspondant à $y > 0$ si $y_0 > 0$ (resp. le demi-axe (Ox) correspondant à $x > 0$ 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 $x_0 > 0, y_0 > 0$. On remarque alors que $f(t) > 0, g(t) > 0$ pour tout t . De plus, on vérifie que si $\lambda_1 \neq 0$,

$$g(t) = y_0 \left(\frac{f(t)}{x_0} \right)^{\frac{\lambda_2}{\lambda_1}}.$$

La courbe \mathcal{C}_{A, X_0} est donc portée par la courbe \mathcal{C}_0 d'équation cartésienne

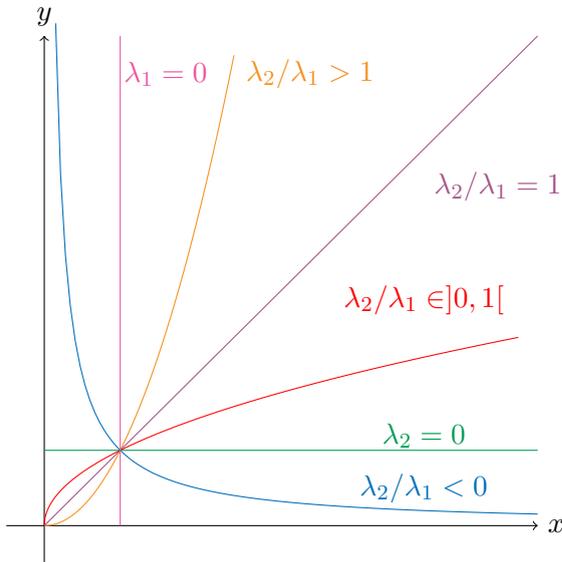
$$y = \frac{y_0}{x_0^{\frac{\lambda_2}{\lambda_1}}} x^{\frac{\lambda_2}{\lambda_1}}, \quad x > 0.$$

Comme $f(\mathbb{R}) = \mathbb{R}_+^*$, on parcourt en fait toute la courbe \mathcal{C}_0 . L'allure de la courbe \mathcal{C}_0 dépend de la valeur de $\frac{\lambda_2}{\lambda_1}$.

- (a) Si $\lambda_2/\lambda_1 < 0$, la fonction $x \in \mathbb{R}_+^* \mapsto x^{\frac{\lambda_2}{\lambda_1}}$ est strictement décroissante, tend vers $+\infty$ en 0^+ et vers 0 en $+\infty$.
- (b) Si $\lambda_2/\lambda_1 = 0$, la courbe \mathcal{C}_{A, X_0} est la demi-droite d'équation $y = y_0, x > 0$.
- (c) Si $\lambda_2/\lambda_1 \in]0, 1[$, la fonction $x \in \mathbb{R}_+^* \mapsto x^{\frac{\lambda_2}{\lambda_1}}$ est strictement croissante, tend vers 0 en 0^+ et vers $+\infty$ en $+\infty$. De plus la courbe \mathcal{C}_{A, X_0} admet une tangente verticale au voisinage du point $(0, 0)$ et une branche parabolique dans la direction $(0x)$ quand $x \rightarrow +\infty$.
- (d) Si $\lambda_2/\lambda_1 = 1$, la courbe \mathcal{C}_{A, X_0} est la demi-droite d'équation $y = \frac{y_0}{x_0}x, x > 0$.
- (e) Si $\lambda_2/\lambda_1 > 1$, la fonction $x \in \mathbb{R}_+^* \mapsto x^{\frac{\lambda_2}{\lambda_1}}$ est strictement croissante, tend vers 0 en 0^+ et vers $+\infty$ en $+\infty$. De plus la courbe \mathcal{C}_{A, X_0} admet une tangente horizontale au voisinage du point $(0, 0)$ et une branche parabolique dans la direction $(0y)$ quand $x \rightarrow +\infty$.

Il reste à étudier le cas $\lambda_1 = 0$. Dans ce cas \mathcal{C}_{A, X_0} est portée par la demi-droite d'équation $x = x_0, y > 0$.

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



2. Tracé de la courbe paramétrée d'équation $\begin{pmatrix} x \\ y \end{pmatrix} = \exp(At)X_0$, avec A triangulaire supérieure non diagonalisable et $X_0 \in \mathbb{R}^2$:

On prend à présent A de la forme

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

avec $\lambda, \alpha \in \mathbb{R}$, $\alpha \neq 0$ (sinon ce cas a déjà été traité dans le paragraphe précédent). On étudie de nouveau la courbe paramétrée \mathcal{C}_{A, X_0} d'équation

$$\begin{cases} x = f(t), \\ y = g(t), \end{cases} \quad t \in \mathbb{R}.$$

avec

$$\begin{pmatrix} f(t) \\ g(t) \end{pmatrix} = \exp(At) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

D'après le paragraphe précédent, on a

$$\exp(At) = \exp(\lambda t) \begin{pmatrix} 1 & \alpha t \\ 0 & 1 \end{pmatrix},$$

et donc

$$\begin{aligned} f(t) &= \exp(\lambda t)(x_0 + \alpha t y_0) \\ g(t) &= \exp(\lambda t)y_0. \end{aligned}$$

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 $y = y_0$. 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 à un point.

On suppose donc à présent $\lambda \neq 0, y_0 \neq 0$. Comme les fonctions f et g n'ont pas de symétrie évidente, on passe à l'analyse de leurs variations.

Les fonctions f et g sont dérivables sur \mathbb{R} et on a

$$\begin{aligned} f'(t) &= \exp(\lambda t)(\lambda x_0 + \lambda \alpha y_0 t + \alpha y_0), \\ g'(t) &= \lambda y_0 \exp(\lambda t). \end{aligned}$$

Ainsi g' ne s'annule pas sur \mathbb{R} et a le signe de λy_0 . f' s'annule en un unique point t_0 défini par

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

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

- On a toujours $f(t) \rightarrow 0, g(t) \rightarrow 0$ quand $t \rightarrow -\text{sgn}(\lambda)\infty$, et de plus $g(t)/f(t) \rightarrow 0$ quand $t \rightarrow -\text{sgn}(\lambda)\infty$: il y a une tangente horizontale au voisinage du point $(0, 0)$.
- On a toujours $g(t) \rightarrow \text{sgn}(y_0)\infty, f(t) \rightarrow \text{sgn}(\alpha y_0)\infty$ quand $t \rightarrow \text{sgn}(\lambda)\infty$. De plus $g(t)/f(t) \rightarrow 0$ quand $t \rightarrow \text{sgn}(\lambda)\infty$: il y a une branche parabolique dans la direction $(0x)$.
- g garde un signe constant (celui de y_0).
- Il y a une tangente verticale au point de paramètre t_0 , qui a pour coordonnées $\exp(\lambda t_0)(-\frac{\alpha y_0}{\lambda}, y_0)$.
- 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 :

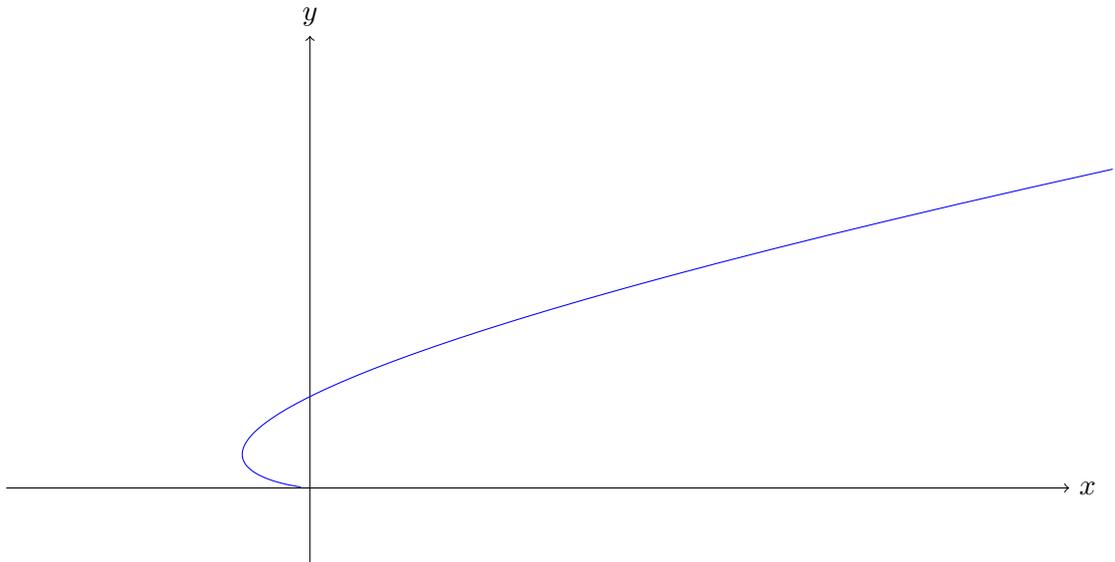
- Le signe de y_0 détermine le demi-plan dans lequel la courbe est incluse ($y > 0$ ou $y < 0$);
- Le signe de $\lambda \alpha y_0$ (ou de $\frac{\alpha y_0}{\lambda}$) détermine la position du point de tangente verticale par rapport à l'axe des ordonnées;
- Le signe de λ détermine le sens dans lequel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

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

Dans ce cas le tableau de variations est

t	$-\infty$	t_0	$+\infty$
$f'(t)$	0	-z	+
$f(t)$	0	$-e^{\lambda t_0} \frac{\alpha y_0}{\lambda}$	$+\infty$
$g'(t)$	0	+	
$g(t)$	0		$+\infty$

On obtient une courbe qui a l'allure suivante :



3. Tracé de la courbe paramétrée d'équation $x = f(t), y = g(t), t \in \mathbb{R}$, avec

$$\begin{pmatrix} f(t) \\ g(t) \end{pmatrix} = \exp(\lambda t) \left(\cos(\mu t) I_2 + \sin(\mu t) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right) X_0,$$

où $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \mathbb{R}^2$ et $\lambda, \mu \in \mathbb{R}, \mu \neq 0$.

On verra que ce cas correspond à (l'exponentielle d') une matrice réelle ayant deux valeurs propres complexes conjuguées (ici $\lambda \pm i\mu$).

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

$$\begin{aligned} f(t) &= e^{\lambda t} (\cos(\mu t) x_0 - \sin(\mu t) y_0), \\ g(t) &= e^{\lambda t} (\cos(\mu t) y_0 + \sin(\mu t) x_0). \end{aligned}$$

Si $X_0 = 0$, alors $f(t) = g(t) = 0$ pour tout t et la trajectoire est réduite à un point.

Si $X_0 \neq 0$, on peut écrire X_0 en coordonnées polaires : il existe un unique couple $(r_0, \theta_0) \in \mathbb{R}_+^* \times [0, 2\pi[$ tel que

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

En remplaçant dans les formules donnant f et g , on obtient

$$\begin{aligned} f(t) &= r_0 e^{\lambda t} (\cos(\mu t) \cos \theta_0 - \sin(\mu t) \sin \theta_0) = r_0 e^{\lambda t} \cos(\mu t + \theta_0), \\ g(t) &= r_0 e^{\lambda t} (\cos(\mu t) \sin \theta_0 + \sin(\mu t) \cos \theta_0) = r_0 e^{\lambda t} \sin(\mu t + \theta_0). \end{aligned} \tag{2.13}$$

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

Remarque 2.1.16. On vérifie facilement que la courbe a une équation polaire de la forme

$$r = r_0 \exp\left(\frac{\lambda}{\mu}(\theta - \theta_0)\right).$$

Ci-dessus on va mener l'étude sans utiliser cette remarque, mais le lecteur pourra, à titre d'exercice, reprendre le tracé en n'utilisant que l'équation polaire et les outils du chapitre 2.

Les fonctions f et g sont infiniment dérivables. On obtient, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} f'(t) &= r_0 \exp(\lambda t) (-\mu \sin(\mu t + \theta_0) + \lambda \cos(\mu t + \theta_0)), \\ g'(t) &= r_0 \exp(\lambda t) (\mu \cos(\mu t + \theta_0) + \lambda \sin(\mu t + \theta_0)). \end{aligned}$$

De nouveau, il existe un unique $\alpha \in [0, 2\pi[$ tel que

$$\lambda = \sqrt{\lambda^2 + \mu^2} \cos \alpha, \quad \mu = \sqrt{\lambda^2 + \mu^2} \sin \alpha.$$

Les formules donnant f' et g' deviennent alors, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} f'(t) &= r_0 \sqrt{\lambda^2 + \mu^2} \exp(\lambda t) \cos(\mu t + \theta_0 + \alpha), \\ g'(t) &= r_0 \sqrt{\lambda^2 + \mu^2} \exp(\lambda t) \sin(\mu t + \theta_0 + \alpha). \end{aligned}$$

Par conséquent, $f'(t)$ (resp. $g'(t)$) a le même signe que $\cos(\mu t + \theta_0 + \alpha)$ (resp. $\sin(\mu t + \theta_0 + \alpha)$), et on a

$$\begin{aligned} f'(t) = 0 &\iff \exists k \in \mathbb{Z}, \mu t + \theta_0 + \alpha = \frac{\pi}{2} + k\pi, \\ g'(t) = 0 &\iff \exists k \in \mathbb{Z}, \mu t + \theta_0 + \alpha = k\pi. \end{aligned}$$

On trace un tableau de variations sur un intervalle du type $[t_1 + \frac{2k\pi}{\mu}, t_1 + \frac{2(k+1)\pi}{\mu}]$, où

$$t_1 = -\frac{\theta_0 + \alpha}{\mu}.$$

(Attention : à cause du préfacteur $\exp(\lambda t)$, les fonctions f et g ne sont pas périodiques...) Pour simplifier, on ne traite que le cas $\mu > 0$ (le cas $\mu < 0$ peut être traité en changeant la paramétrisation $t \rightarrow -t$ et en changeant λ en $-\lambda$).

On remarque par exemple que

$$f\left(t_1 + \frac{2k\pi}{\mu}\right) = r_0 \exp\left(\lambda\left(t_1 + \frac{2k\pi}{\mu}\right)\right) \cos \alpha,$$

et $\cos \alpha = \frac{\lambda}{\sqrt{\lambda^2 + \mu^2}}$.

Pour alléger un peu les notations du tableau ci-dessous, on pose

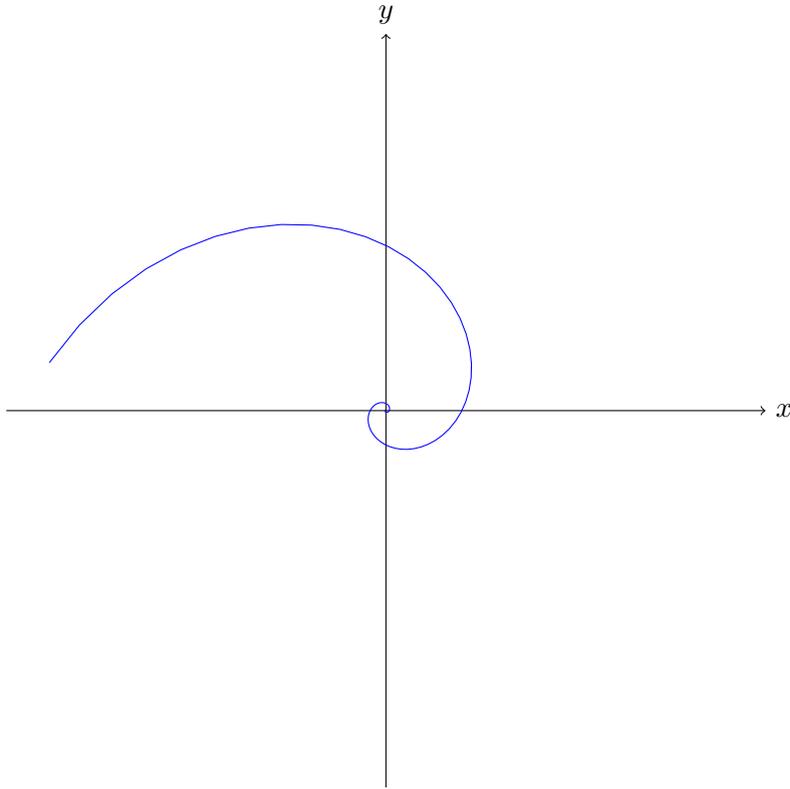
$$r_k = \frac{r_0}{\sqrt{\lambda^2 + \mu^2}} \exp\left(\lambda\left(t_1 + \frac{2k\pi}{\mu}\right)\right).$$

t	$t_1 + \frac{2k\pi}{\mu}$	$t_1 + \frac{\pi}{2\mu} + \frac{2k\pi}{\mu}$	$t_1 + \frac{(2k+1)\pi}{\mu}$	$t_1 + \frac{3\pi}{2\mu} + \frac{2k\pi}{\mu}$	$t_1 + \frac{2(k+1)\pi}{\mu}$		
$f'(t)$		+	0	-	-	0	+
$f(t)$	$r_1\lambda$		$r_1\mu e^{\frac{\lambda\pi}{2\mu}}$		$-r_1\mu e^{\frac{3\lambda\pi}{2\mu}}$		$r_1\lambda e^{\frac{2\lambda\pi}{\mu}}$
$g'(t)$	0	+	+	0	-	-	0
$g(t)$	$-r_1\mu$			$r_1\mu e^{\frac{\lambda\pi}{\mu}}$			$-r_1\mu e^{\frac{2\lambda\pi}{\mu}}$

Ainsi, sur un intervalle $[t_1 + \frac{2k\pi}{\mu}, t_1 + \frac{2(k+1)\pi}{\mu}]$, le point de coordonnées $(f(t), g(t))$ fait un tour complet autour de l'origine (remarquons que f et g prennent des valeurs positives et négatives, comme on peut le voir sur le tableau de variations ou sur la formule (2.13)). La courbe admet des tangentes horizontales aux points de paramètres $t_1 + k\pi/\mu$, des tangentes verticales aux points de paramètres $t_1 + \pi/(2\mu) + k\pi/\mu$. Enfin, on peut remarquer que deux points de paramètres t et $t + 2\pi/\mu$ ont le même argument mais des rayons différents : en effet, on a

$$\frac{\sqrt{f^2(t + 2\pi/\mu) + g^2(t + 2\pi/\mu)}}{\sqrt{f^2(t) + g^2(t)}} = \exp\left(\frac{2\pi\lambda}{\mu}\right)$$

En mettant bout à bout les intervalles $t_1 + 2k\pi/\mu, t_1 + 2(k + 1)\pi/\mu$, on obtient une spirale qui s'enroule autour de l'origine (quand $t \rightarrow -\text{sgn}(\lambda)\infty$) et qui part vers l'infini (quand $t \rightarrow \text{sgn}(\lambda)\infty$).



Les trois cas étudiés ci-dessus permettent en fait de comprendre le portrait de phase des solutions de l'équation (2.5) dans le cas général. Pour cela, on utilise le résultat suivant :

Lemme 2.1.17. 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 , et (x_P, y_P) les coordonnées dans le repère \mathcal{R}_P .

Soit \mathcal{C} la courbe paramétrée définie par l'équation

$$\begin{aligned} x &= X_1(t), \\ y &= X_2(t), \end{aligned} \quad t \in I.$$

Alors l'équation de \mathcal{C} dans le repère \mathcal{R}_P est

$$\begin{aligned} x_P &= Y_1(t), \\ y_P &= Y_2(t), \end{aligned} \quad t \in I.$$

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

Cas général

On en déduit immédiatement du paragraphe précédent l'allure des courbes paramétrées définies par

$$x = X_1(t), \quad y = X_2(t), \quad t \in \mathbb{R},$$

quand X est solution de l'équation différentielle

$$X'(t) = AX(t), \quad X(0) = X_0$$

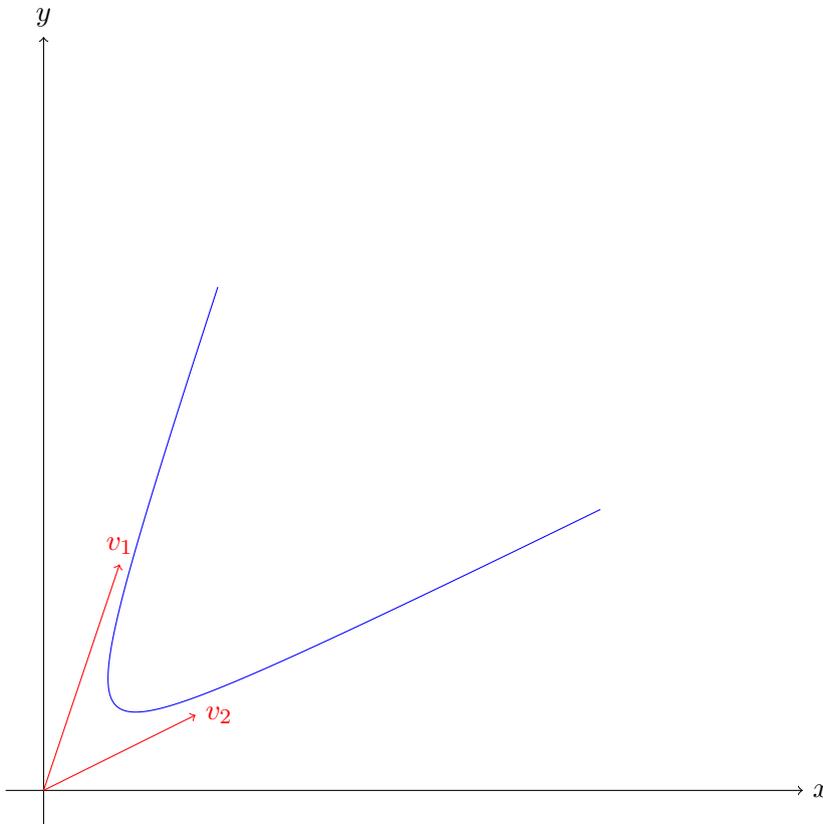
où $A \in \mathcal{M}_2(\mathbb{R})$ a deux valeurs propres réelles distinctes ou confondues λ_1 et λ_2 . En effet, si $\lambda_1 \neq \lambda_2$ il existe $P \in GL_2(\mathbb{R})$ tel que

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

L'équation de la courbe paramétrée dans le repère \mathcal{R}_P (avec les notations du Lemme 2.1.17) est donc

$$x_P = \exp(\lambda_1 t)y_P = x = \exp(\lambda_2 t)y_2, \quad t \in \mathbb{R},$$

où $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = P^{-1}X_0$. Notons que le point de coordonnées (y_1, y_2) 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 préliminaire faite au paragraphe précédent. Il faut cependant garder en tête qu'en général, le repère \mathcal{R}_P n'est pas orthonormé, ni même orthogonal. La courbe a donc une allure déformée par ce changement de repère : les tangentes horizontales (resp. verticales) deviennent des tangentes dans la direction v_1 (resp. dans la direction v_2). On trace ci-dessous un exemple de courbe avec $\lambda_2/\lambda_1 < 0$ et $y_1, y_2 > 0$:



De même, si $\lambda_1 = \lambda_2$, on obtient des courbes ayant le même aspect que dans le deuxième cas étudié au paragraphe précédent, après déformation par le changement de repère.

Il reste à étudier le cas où la matrice A possède deux valeurs propres $\lambda_1, \lambda_2 \in \mathbb{C} \setminus \mathbb{R}$, et à montrer que ce dernier cas peut se ramener au troisième cas étudié au paragraphe précédent.

On sait que dans ce cas $\lambda_1 = \bar{\lambda}_2$. Posons $\lambda = \Re(\lambda_1), \mu = \Im(\lambda_1) \neq 0$. Il existe une matrice $Q \in GL_2(\mathbb{C})$ telle que

$$A = Q \begin{pmatrix} \lambda + i\mu & 0 \\ 0 & \lambda - i\mu \end{pmatrix} Q^{-1}.$$

On a donc

$$\begin{aligned} \exp(At) &= Q \begin{pmatrix} \exp((\lambda + i\mu)t) & 0 \\ 0 & \exp((\lambda - i\mu)t) \end{pmatrix} Q^{-1} \\ &= Q \left(\exp(\lambda t) \begin{pmatrix} \cos(\mu t) + i \sin(\mu t) & 0 \\ 0 & \cos(\mu t) - i \sin(\mu t) \end{pmatrix} \right) Q^{-1} \\ &= Q \left(\exp(\lambda t) \cos(\mu t) I_2 + \exp(\lambda t) \sin \mu t \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \right) Q^{-1} \\ &= \exp(\lambda t) \cos(\mu t) I_2 + \exp(\lambda t) \sin(\mu t) Q \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} Q^{-1}. \end{aligned}$$

Posons à présent

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

Comme la matrice A est réelle, la matrice $\exp(At)$ est également réelle pour tout $t \in \mathbb{R}$. On en déduit que la matrice J est nécessairement réelle.

Par ailleurs, comme la matrice J est semblable à la matrice $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, ses valeurs propres sont i et $-i$. Considérons à présent la matrice

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

Son polynôme caractéristique est $\lambda^2 + 1$, donc ses valeurs propres dans \mathbb{C} sont i et $-i$. Elle est donc semblable dans \mathbb{C} à la matrice J . Comme les matrices J et J' sont toutes deux réelles, on en déduit qu'elles sont semblables dans \mathbb{R} . Il existe donc $P \in GL_2(\mathbb{R})$ telle que

$$J = PJ'P^{-1}.$$

On obtient alors

$$\exp(At) = P \left(\exp(\lambda t) \cos(\mu t) I_2 + \exp(\lambda t) \sin(\mu t) J' \right) P^{-1}.$$

L'équation de la courbe $\begin{pmatrix} x \\ y \end{pmatrix} = \exp(At)X_0$ dans le repère \mathcal{R}_P est du troisième type décrit dans le paragraphe précédent. La courbe paramétrée est donc une spirale qui s'entoure autour de l'origine.

2.1.8 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 du portrait de phase pour laquelle $t \in [0, \infty[$. De façon générale, les notions de stabilité s'intéressent au comportement d'un système au voisinage d'un point

d'équilibre. 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 $X'(t) = 0 = AX_0$).

Ici $A \in \mathcal{M}_2(\mathbb{R})$, donc $\ker A$ est de dimension 0 (si 0 n'est pas valeur propre de A), 1 (si 0 est valeur propre simple de A), ou 2 (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 2.1.18. *On dit que $X_0 \in \ker A$ est un point d'équilibre stable au sens de Lyapunov si la propriété suivante est vérifiée :*

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

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

$$\exists \epsilon_0 > 0, \forall Y_0 \in \mathbb{R}^2, \|Y_0 - X_0\| \leq \epsilon_0 \Rightarrow \lim_{t \rightarrow \infty} \|\exp(At)Y_0 - X_0\| = 0.$$

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 :

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

- 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

$$\begin{aligned} x &= \exp(\lambda_1 t)x_0, \\ y &= \exp(\lambda_2 t)y_0, \end{aligned}$$

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 36, 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 36). Le signe de λ_2 détermine le sens de parcours de ces demi-droites. Si $\lambda_2 > 0$, on a

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

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.

On résume ces propriétés dans le tableau suivant :

	$\lambda_1 < 0$	$\lambda_1 = 0$	$\lambda_1 > 0$
$\lambda_2 < 0$	Stabilité orbitale et asymptotique	Stabilité orbitale, mais pas de stabilité asymptotique	Pas de stabilité orbitale ni asymptotique
$\lambda_2 = 0$	Stabilité orbitale, mais pas de stabilité asymptotique	Stabilité orbitale, mais pas de stabilité asymptotique (Le système n'évolue pas!)	Pas de stabilité orbitale ni asymptotique
$\lambda_2 > 0$	Pas de stabilité orbitale ni asymptotique	Pas de stabilité orbitale ni asymptotique	Pas de stabilité orbitale ni asymptotique

2. Si $A = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}$:

Si $\lambda \neq 0$, le portrait de phase a l'allure dessinée page 38. 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 $(\beta, 0)$ avec $\beta \in \mathbb{R}$. Par ailleurs

$$\exp(At)X_0 = X_0 + \alpha t \begin{pmatrix} y_0 \\ 0 \end{pmatrix}.$$

Ainsi si $\alpha \neq 0$ (c'est à dire si A n'est pas la matrice nulle!) $\|\exp(At)X_0\| \rightarrow \infty$ quand $t \rightarrow \infty$. On ne peut donc pas avoir de stabilité asymptotique ni de stabilité orbitale.

3. Si les valeurs propres de A sont $\lambda \pm i\mu$ avec $\mu \neq 0$:
Seul $(0, 0)$ est état d'équilibre.

Si $\lambda = 0$, les trajectoires sont des cercles dans le repère ad hoc. On a donc stabilité orbitale, mais pas stabilité asymptotique.

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

Remarque 2.1.20. *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

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

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 pilotée 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 $Db(X_{eq})$ 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.

2.2 Analyse numérique

On a vu, dans ce Chapitre et dans le précédent, des façons de calculer explicitement des solutions d'équation différentielle en fonction du temps - lorsque c'était possible - et quelques cas où on peut donner des renseignements qualitatifs sur le comportement des solutions sans savoir les calculer explicitement.

On va aborder dans cette partie un point de vue complémentaire : comment faire pour calculer, avec une précision arbitrairement grande, des valeurs approchées des valeurs prises par les solutions d'équations différentielles ?

2.2.1 Notion de solution approchée

Dans toute cette partie, on utilisera la notion suivante

Définition 2.2.1. Soit I un intervalle de \mathbb{R} , et soit $X : I \rightarrow \mathbb{R}^m$ la solution d'une équation différentielle.

Une suite de solutions approchées est une suite de fonctions $(X^{(p)})_{p \in \mathbb{N}}$, avec $X^{(p)} : I \rightarrow \mathbb{R}^m$ pour tout p , et telle que

$$\lim_{p \rightarrow \infty} \sup_{t \in I} \|X^{(p)}(t) - X(t)\| = 0.$$

La démarche générale des schémas numériques qui seront vus dans ce cours est la suivante. On suppose que l'intervalle I est de la forme $I = [a, b]$. On commence par **discrétiser** l'intervalle : on découpe l'intervalle I en p sous-intervalles, et on ne va finalement s'intéresser qu'aux valeurs prises par la fonction u aux bornes de ces sous-intervalles.

Pour $p \in \mathbb{N}$, $p \geq 1$, on pose

$$h_p = \frac{b-a}{p}, \quad t_{k,p} = a + kh_p \text{ pour } 0 \leq k \leq p. \quad (2.14)$$

On remarque que l'on a $t_{0,p} = a$, $t_{p,p} = b$. Les points $t_{k,p}$ sont appelés **points de discrétisation**.

L'idée est de définir chaque fonction $X^{(p)}$ en prescrivant les valeurs prises en $t_{k,p}$ pour $k \in \{0, \dots, p\}$ et en ne travaillant qu'avec des fonctions affines par morceaux. Il suffit donc de connaître $X^{(p)}(t_{0,p})$, $X^{(p)}(t_{1,p})$, \dots , $X^{(p)}(t_{p,p})$ pour connaître la fonction $X^{(p)}$ sur l'ensemble de l'intervalle $[a, b]$. Par ailleurs, les valeurs de $X^{(p)}(t_{0,p})$, $X^{(p)}(t_{1,p})$, \dots , $X^{(p)}(t_{p,p})$ sont calculées par récurrence :

– on prend en général

$$X^{(p)}(a) = X^p(t_{0,p}) = X(a),$$

la quantité $X(a)$ étant par ailleurs donnée par la **condition initiale** prescrite avec l'équation différentielle étudiée ;

– $X^{(p)}(t_{k+1,p})$ est calculé en fonction de $X^{(p)}(t_{k,p})$. La formule reliant $X^{(p)}(t_{k+1,p})$ et $X^{(p)}(t_{k,p})$ détermine l'**algorithme** utilisé.

Récapitulons :

1. $X^{(p)}(a)$ est donné par la condition initiale de l'équation différentielle ;
2. $X^{(p)}(t_{k+1,p})$ et $X^{(p)}(t_{k,p})$ sont reliés par une formule de récurrence ;
3. pour $t \in]t_{k,p}, t_{k+1,p}[$, on prend

$$X^{(p)}(t) = \frac{t - t_{k,p}}{h_p} (X^{(p)}(t_{k+1,p}) - X^{(p)}(t_{k,p})) + X^{(p)}(t_{k,p})$$

Il y a plusieurs choses à vérifier :

- La première étape est de vérifier que $X^{(p)}$ est bien définie pour tout p . Compte tenu de la définition de $X^{(p)}$ ci-dessus, cela revient à montrer que la formule de récurrence reliant $X^{(p)}(t_{k+1,p})$ et $X^{(p)}(t_{k,p})$ permet bien de déterminer $X^{(p)}(t_{k+1,p})$ de façon unique.
- Ensuite il faut vérifier que la suite $(X^{(p)})_{p \geq 1}$ forme bien une suite de solutions approchées de X . Pour cela on peut commencer par montrer que

$$\lim_{p \rightarrow \infty} \sup_k \|X(t_{k,p}) - X^{(p)}(t_{k,p})\| = 0,$$

puis montrer par exemple que

$$\lim_{p \rightarrow \infty} \sup_k \sup_{t \in [t_{k,p}, t_{k+1,p}]} \left(\|X(t_{k,p}) - X(t)\| + \|X^{(p)}(t_{k,p}) - X^{(p)}(t)\| \right) = 0.$$

La première des deux conditions ci-dessus permet de dire qu'on a bien une convergence uniforme de $X^{(p)}$ vers X sur l'ensemble des points $t_{k,p}$. La seconde condition examine ce qui se passe entre $t_{k,p}$ et $t_{k+1,p}$, pour la fonction X et pour la fonction $X^{(p)}$. L'idée est que comme l'intervalle $[t_{k,p}, t_{k+1,p}]$ a une longueur h_p , avec $\lim_{p \rightarrow \infty} h_p \rightarrow 0$, si les fonctions X et $X^{(p)}$ sont régulières (par exemple si leur dérivée sur l'intervalle $]t_{k,p}, t_{k+1,p}[$ est uniformément bornée), la seconde condition sera réalisée.

Dans la pratique, dans l'étude des équations différentielles et pour les schémas qui y sont associés, la dernière condition est vérifiée de façon quasi-automatique. Par exemple, pour une équation du type

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

il suffit que la fonction b soit bornée.

Le schéma numérique en tant que tel ne s'intéresse donc qu'aux nombres $X^{(p)}(t_{k,p})$, pour $0 \leq k \leq p$. Pour alléger un peu les notations, on notera ces nombres X_k , $0 \leq k \leq p$, en omettant donc la dépendance par rapport à p . Il est cependant nécessaire que le lecteur la garde à l'esprit.

Remarque 2.2.2. La suite $(X_k)_{0 \leq k \leq p}$ est une **approximation** de la suite $(X(t_{k,p}))_{0 \leq k \leq p}$. Elle correspond aux valeurs prises par la suite de solutions approchées $X^{(p)}$ aux points de discrétisation $t_{k,p}$. Attention donc à **ne pas confondre** X_k et $X(t_{k,p})$!

Définition 2.2.3. Un schéma numérique est la donnée d'une **relation de récurrence** permettant de définir la suite $(X_k)_{0 \leq k \leq p}$.

2.2.2 Schémas d'Euler

On considère dans ce paragraphe une équation différentielle du type

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

avec $A : I \rightarrow \mathcal{M}_2(\mathbb{R})$, $B : I \rightarrow \mathbb{R}^2$, et $X : I \rightarrow \mathbb{R}^2$, où $I = [a, b]$ est un intervalle de \mathbb{R} .

On munit cette équation différentielle d'une condition initiale

$$X(a) = X_{ini},$$

et on cherche à calculer des solutions approchées de l'équation différentielle. Comme expliqué au paragraphe précédent, on choisit des points de discrétisation définis par (2.14). Pour définir le schéma numérique, on commence par considérer l'équation différentielle aux points de discrétisation :

$$X'(t_{k,p}) = A(t_{k,p})X(t_{k,p}) + B(t_{k,p}).$$

Il y a deux idées principales pour construire les schémas numériques qui seront vus dans le cadre de ce cours :

- On remplace $X'(t_{k,p})$ par un taux d'accroissement de la fonction X entre deux instants voisins de $t_{k,p}$. Par exemple, on peut écrire

$$X'(t_{k,p}) \simeq \frac{X(t_{k+1,p}) - X(t_{k,p})}{h_p} \quad (2.15)$$

ou encore

$$X'(t_{k,p}) \simeq \frac{X(t_{k,p}) - X(t_{k-1,p})}{h_p}. \quad (2.16)$$

– On remplace chaque occurrence de $X(t_{k,p})$ par X_k .

Les deux choix (2.15) et (2.16) donnent lieu à deux équations de récurrence différentes sur la suite $(X_k)_{0 \leq k \leq p}$:

1. **Choix (2.15) : schéma d'Euler explicite :**

On obtient l'équation de récurrence

$$\frac{X_{k+1} - X_k}{h_p} = A(t_{k,p})X_k + B(t_{k,p}),$$

ce qui donne, pour $0 \leq k \leq p-1$,

$$X_{k+1} = (I_2 + h_p A(t_{k,p}))X_k + h_p B(t_{k,p}). \quad (2.17)$$

L'initialisation de la récurrence est faite en prenant

$$X_0 = X_{ini}.$$

2. **Choix (2.16) : schéma d'Euler implicite :**

On obtient l'équation de récurrence

$$\frac{X_k - X_{k-1}}{h_p} = A(t_{k,p})X_k + B(t_{k,p}),$$

ce qui s'écrit encore

$$(I_2 - h_p A(t_{k,p}))X_k = X_{k-1} + h_p B(t_{k,p}).$$

Cette équation est dite *implicite* car la formule ne donne pas directement X_k en fonction de X_{k-1} : il faut d'abord inverser la matrice $I_2 - h_p A(t_{k,p})$ (après avoir montré que cette matrice est bien inversible !). En écrivant l'équation de récurrence au rang $k+1$, on obtient, si $I_2 - h_p A(t_{k+1,p})$ est inversible,

$$X_{k+1} = (I_2 - h_p A(t_{k+1,p}))^{-1} X_k + h_p (I_2 - h_p A(t_{k+1,p}))^{-1} B(t_{k+1,p}). \quad (2.18)$$

Lemme 2.2.4. *Pour tout $t \in I$, on écrit*

$$A(t) = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}$$

On suppose qu'il existe une constante $C_0 > 0$ telle que

$$\sup_{t \in I} \max(|a(t)|, |b(t)|, |c(t)|, |d(t)|) \leq C_0.$$

Alors il existe $p_0 \in \mathbb{N}$ tel que pour tout $p \geq p_0$, pour tout $t \in I$, la matrice $I_2 - h_p A(t)$ est inversible.

Par conséquent, pour tout $p \geq p_0$, le schéma d'Euler implicite donné par l'équation de récurrence (2.18) est bien défini.

Démonstration. Pour tout $t \in I$, $p \in \mathbb{N}^*$, la matrice $I_2 - h_p A(t)$ s'écrit

$$I_2 - h_p A(t) = \begin{pmatrix} 1 - h_p a(t) & -h_p b(t) \\ -h_p c(t) & 1 - h_p d(t) \end{pmatrix}.$$

On a donc

$$\begin{aligned}\det(I_2 - h_p A(t)) &= (1 - h_p a(t))(1 - h_p d(t)) - h_p^2 b(t)c(t) \\ &= 1 - h_p(a(t) + d(t)) + h_p^2(a(t)d(t) - b(t)c(t)).\end{aligned}$$

Ainsi, pour tout $t \in I$,

$$|\det(I_2 - h_p A(t)) - 1| \leq 2h_p C_0 + 2h_p^2 C_0^2.$$

On rappelle que $h_p = (b - a)/p$, de sorte que le membre de droite de l'inégalité ci-dessus tend vers zéro lorsque p tend vers l'infini. Soit $p_0 \in \mathbb{N}$ tel que

$$p \geq p_0 \Rightarrow 2h_p C_0 + 2h_p^2 C_0^2 \leq \frac{1}{2}.$$

On en déduit que si $p \geq p_0$, alors pour tout $t \in I$,

$$\frac{1}{2} \leq \det(I_2 - h_p A(t)) \leq \frac{3}{2}.$$

En particulier $\det(I_2 - h_p A(t)) \neq 0$ pour tout $t \in I$ et pour tout $p \geq p_0$, et la matrice $(I_2 - h_p A(t))$ est donc inversible. □

Remarque 2.2.5. Dans la preuve du lemme ci-dessus, on a en fait utilisé la continuité en zéro de l'application

$$h \mapsto \det(I - hA).$$

Plus spécifiquement, on utilise le fait que cette continuité est "uniforme" pour A dans une boule de $\mathcal{M}_2(\mathbb{R})$.

Remarque 2.2.6. On peut aussi faire d'autres choix que (2.15) et (2.16). Par exemple, on peut aussi écrire

$$X'(t_{k,p}) \simeq \frac{X(t_{k+1,p}) - X(t_{k,p})}{2h_p}.$$

Le lecteur pourra s'entraîner à écrire la relation de récurrence correspondante sur $(X_k)_{0 \leq k \leq p}$ et vérifier qu'il s'agit d'une suite récurrente d'ordre deux. On se concentre dans ce cours uniquement sur les schémas (2.17) et (2.18), mais il faut bien garder en tête qu'il en existe une multitude d'autres.

2.2.3 Notion de convergence d'un schéma numérique

Définition 2.2.7. On considère une équation différentielle du type

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

avec $X : [a, b] \rightarrow \mathbb{R}^m$, $b : [a, b] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$. On munit cette équation différentielle d'une donnée initiale $X(a) = X_{ini}$.

On considère un schéma numérique associé à cette équation différentielle; ce schéma permet de construire, pour tout p suffisamment grand, une suite $(X_k)_{0 \leq k \leq p}$ définie par $X_0 = X_{ini}$ et par une relation de récurrence.

On dit que ce schéma numérique est **convergent** si pour tout $X_{ini} \in \mathbb{R}^m$, on a

$$\lim_{p \rightarrow \infty} \sup_{0 \leq k \leq p} \|X(t_{k,p}) - X_k\| = \lim_{p \rightarrow \infty} \sup_{0 \leq k \leq p} \|X(t_{k,p}) - X^{(k)}(t_{k,p})\| = 0.$$

Examinons à présent cette notion de convergence pour les deux schémas d'Euler présentés plus haut, (2.17) et (2.18), dans le cas d'une équation linéaire à coefficients constants. Dans ce cas, on a

$$A(t) = A_0 \text{ et } B(t) = 0 \quad \forall t \in I.$$

1. Schéma d'Euler explicite : la relation de récurrence (2.17) devient

$$X_{k+1} = (I_2 + h_p A_0) X_k \quad \forall k \in \{0, \dots, p-1\}.$$

On en déduit que

$$X_k = (I_2 + h_p A_0)^k X_{ini}$$

pour tout $k \in \{0, \dots, p\}$. D'un autre côté on rappelle que

$$t_{k,p} = a + kh_p = a + \frac{k(b-a)}{p}.$$

Par conséquent

$$X_k = (I_2 + h_p A_0)^{\lfloor \frac{t_{k,p}-a}{b-a} p \rfloor} X_{ini}.$$

On a le résultat suivant (admis ici) :

Lemme 2.2.8. *Soit $R > 0$. Il existe une constante C_R telle que la propriété suivante est vérifiée :*

Pour toute matrice $A = (a_{i,j})_{1 \leq i,j \leq 2} \in \mathcal{M}_2(\mathbb{C})$ telle que

$$\max_{1 \leq i,j \leq 2} |a_{i,j}| \leq R,$$

et pour tout h dans un voisinage de zéro (dépendant de R), il existe une matrice $\tilde{A}_h = (\tilde{a}_{h,ij})_{1 \leq i,j \leq 2}$, telle que

$$\max_{1 \leq i,j \leq 2} |\tilde{a}_{h,ij}| \leq C_R h^2$$

et telle que

$$I_2 + hA = \exp(hA + \tilde{A}_h).$$

De surcroît, les matrices A et \tilde{A}_h commutent pour tout h dans un voisinage de zéro.

Par conséquent, pour tout $k \in \mathbb{N}$, pour tout $h > 0$,

$$(I_2 + hA)^k = \exp(khA + k\tilde{A}_h)$$

Remarque 2.2.9. *Si*

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

avec $\lambda_1, \lambda_2 \in \mathbb{R}$, alors le lemme est facile à démontrer. En effet, dans ce cas on a

$$(I_2 + hA)^k = P \begin{pmatrix} (1 + h\lambda_1)^k & 0 \\ 0 & (1 + h\lambda_2)^k \end{pmatrix} P^{-1},$$

et

$$1 + h\lambda = \exp(\ln(1 + h\lambda)) = \exp(h\lambda + O(h\lambda^2)).$$

On déduit du Lemme 2.2.8 la convergence du schéma d'Euler implicite. En effet,

$$X_k = (I_2 + h_p A_0)^{\lfloor \frac{t_{k,p} - a}{b-a} p \rfloor} X_{ini} = \exp \left(\left[\frac{t_{k,p} - a}{b-a} p \right] \left(\frac{b-a}{p} A_0 + \tilde{A}_{0,h_p} \right) \right) X_{ini}.$$

De plus, quand $p \rightarrow \infty$,

$$\left[\frac{t-a}{b-a} p \right] \frac{b-a}{p} = (t-a) + O\left(\frac{1}{p}\right).$$

En effet,

$$\frac{t-a}{b-a} p - 1 < \left[\frac{t-a}{b-a} p \right] \leq \frac{t-a}{b-a} p,$$

et donc

$$t-a - \frac{b-a}{p} < \left[\frac{t-a}{b-a} p \right] \frac{b-a}{p} \leq t-a.$$

Ainsi, pour tout $t \in [a, b]$ et pour tout $p \geq 0$,

$$\left\| \left[\frac{t_{k,p} - a}{b-a} p \right] \left(\frac{b-a}{p} A_0 + \tilde{A}_{0,h_p} \right) - (t-a) A_0 \right\| \leq \frac{b-a}{p} \|A_0\| + \left[\frac{t_{k,p} - a}{b-a} p \right] \|\tilde{A}_{0,h_p}\| \leq \frac{C'}{p},$$

où les constantes C et C' ne dépendent que de a, b , et A_0 (mais pas de p ni de t).

On en déduit que

$$X_k = \exp \left((t_{k,p} - a) A_0 + \left[\frac{t_{k,p} - a}{b-a} p \right] \frac{b-a}{p} - (t_{k,p} - a) \right) A_0 + \left[\frac{t_{k,p} - a}{b-a} p \right] \tilde{A}_{0,h_p} \right) X_{ini},$$

tandis que

$$X(t_{k,p}) = \exp((t_{k,p} - a) A) X_{ini}.$$

On a alors, pour tout $k \in \{0, \dots, p\}$,

$$\begin{aligned} & \|X_k - X(t_{k,p})\| \\ & \leq C \|\exp((t_{k,p} - a) A_0) X_{ini}\| \left\| I_2 - \exp \left(\left[\frac{t_{k,p} - a}{b-a} p \right] \frac{b-a}{p} - (t_{k,p} - a) \right) A_0 + \left[\frac{t_{k,p} - a}{b-a} p \right] \tilde{A}_{0,h_p} \right\| \\ & \leq C' \|\exp((t_{k,p} - a) A_0) X_{ini}\| \left\| \left[\frac{t_{k,p} - a}{b-a} p \right] \frac{b-a}{p} - (t_{k,p} - a) \right\| \left\| A_0 + \left[\frac{t_{k,p} - a}{b-a} p \right] \tilde{A}_{0,h_p} \right\| \\ & \leq C' \|\exp((t_{k,p} - a) A_0) X_{ini}\| \left(\frac{b-a}{p} \|A_0\| + p \|\tilde{A}_{0,h_p}\| \right) \\ & \leq C'' \left(\sup_{t \in [a,b]} \|\exp((t-a) A_0) X_{ini}\| \right) \frac{1}{p} \end{aligned}$$

où les constantes C, C', C'' ne dépendent ni de k ni de p . Pour finir,

$$\sup_{0 \leq k \leq p} \|X_k - X(t_{k,p})\| \leq C'' \left(\sup_{t \in [a,b]} \|\exp((t-a) A_0) X_{ini}\| \right) \frac{1}{p}.$$

Par conséquent, le schéma d'Euler explicite est convergent.

2. Schéma d'Euler implicite : le raisonnement est similaire. On prend p suffisamment grand pour que la matrice $I_2 - h_p A_0$ soit inversible. On a alors

$$X_k = (I_2 - h_p A_0)^{-\lfloor \frac{t_{k,p} - a}{b-a} p \rfloor} X_{ini}$$

où $B^{-k} = (B^{-1})^k$ pour $B \in GL_2(\mathbb{R})$, $k \in \mathbb{N}$.

En utilisant de nouveau le lemme 2.2.8, on a, pour tout h dans un voisinage de zéro,

$$I_2 - hA_0 = \exp(-hA_0 + \tilde{A}_{0,-h}),$$

et donc

$$(I_2 - hA_0)^{-1} = \left(\exp(-hA_0 + \tilde{A}_{0,-h}) \right)^{-1} = \exp(-(-hA_0 + \tilde{A}_{0,-h})) = \exp(hA_0 - \tilde{A}_{0,-h}).$$

La preuve de la convergence du schéma est ensuite identique à celle du schéma d'Euler explicite, en remplaçant $\tilde{A}_{0,h}$ par $-\tilde{A}_{0,-h}$.

On en déduit que le schéma d'Euler implicite est également convergent.

2.2.4 Perte de périodicité dans les schémas d'Euler

On se place dans le même cas qu'au paragraphe précédent, avec

$$A_0 = P \begin{pmatrix} i\mu & 0 \\ 0 & -i\mu \end{pmatrix} P^{-1}$$

pour $P \in GL_2(\mathbb{C})$. On a vu que dans ce cas,

$$X(t) = P \begin{pmatrix} \exp(i\mu(t-a)) & 0 \\ 0 & \exp(-i\mu(t-a)) \end{pmatrix} P^{-1} X_{ini}$$

Les trajectoires sont donc périodiques de période $2\pi/\mu$.

On peut maintenant voir ce qu'il en est pour les schémas d'Euler : prenons $b = a + \frac{2\pi}{\mu}$. A-t-on $X_p = X_{ini}$?

La réponse est non : en effet, pour tout $k \in \{0, \dots, p\}$, on a (dans le cas du schéma d'Euler explicite)

$$X_k = (I_2 + h_p A_0)^k X_{ini} = P \begin{pmatrix} (1 + ih_p \mu)^k & 0 \\ 0 & (1 - ih_p \mu)^k \end{pmatrix} P^{-1} X_{ini}.$$

Il s'agit donc de déterminer si on a

$$\begin{pmatrix} (1 + ih_p \mu)^k & 0 \\ 0 & (1 - ih_p \mu)^k \end{pmatrix} = I_2$$

pour un certain $k > 0$ (par exemple $k = p$), et donc si

$$(1 + i\mu)^k = 1$$

pour un certain $k > 0$. Or

$$|(1 + ih_p \mu)^k| = |1 + ih_p \mu|^k = (1 + h_p^2 \mu^2)^{k/2} > 1.$$

Les trajectoires du schéma d'Euler explicite ne sont donc pas périodiques. Elles forment des spirales qui s'éloignent de l'origine.

On peut également s'en convaincre en prenant un exemple explicite de matrice A_0 , par exemple

$$A_0 = \begin{pmatrix} 0 & -\mu & \mu & 0 \end{pmatrix}.$$

On vérifie alors que les trajectoires $\{X(t), t \in \mathbb{R}\}$ sont des cercles centrés en l'origine. Pour le schéma d'Euler, en posant $X_k = \begin{pmatrix} x_k \\ y_k \end{pmatrix}$, on obtient, puisque $X_{k+1} = (I_2 + h_p A_0)X_k$

$$\begin{aligned} x_{k+1} &= x_k - h_p \mu y_k, \\ y_{k+1} &= y_k + h_p \mu x_k, \end{aligned}$$

de sorte que

$$\|X_{k+1}\|^2 = x_{k+1}^2 + y_{k+1}^2 = (1 + h_p^2 \mu^2)(x_k^2 + y_k^2) = (1 + h_p^2 \mu^2)\|X_k\|^2 :$$

la distance à l'origine augmente à chaque itération, on ne reste pas sur le cercle centré en zéro et passant par X_{ini} .

Pour le schéma d'Euler implicite, le phénomène est similaire mais inversé : la distance à l'origine diminue à chaque itération. En reprenant les calculs de la seconde méthode, on a par exemple

$$X_{k+1} = (I_2 - h_p A_0)^{-1} X_k = \frac{1}{1 + h_p^2 \mu^2} \begin{pmatrix} 1 & -h_p \mu \\ h_p \mu & 1 \end{pmatrix},$$

et donc

$$\|X_{k+1}\|^2 = \frac{1}{1 + h_p^2 \mu^2} \|X_k\|^2.$$

Chapitre 3

Théorie générale pour les équations différentielles

3.1 Théorème de Cauchy-Lipschitz : énoncé et conséquences

On commence par un théorème d'existence et d'unicité assez puissant pour les équations différentielles d'ordre un en dimension quelconque (on a vu dans le chapitre précédent que toutes les équations différentielles peuvent se ramener à ce cas). On va en énoncer ici une version simplifiée, dont la preuve est admise.

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

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 $b : I \times U \rightarrow \mathbb{R}^m$

On considère l'équation différentielle

$$\begin{aligned} X'(t) &= b(t, X(t)), & t \in I, \\ X(t_0) &= X_0. \end{aligned} \tag{3.1}$$

Une solution de (3.1) est une fonction X définie sur un intervalle I_X contenant x_0 , à valeurs dans \mathbb{R}^m , dérivable sur I_X , et vérifiant les deux conditions de (3.1).

Remarque 3.1.2. Dans la définition ci-dessus, on ne suppose pas que x_0 appartient à l'intérieur de I_X (ni que l'intervalle I_X est ouvert) : ce peut être une extrémité de l'intervalle.

Théorème 3.1.3 (Théorème de Cauchy-Lipschitz). 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 $b : I \times U \rightarrow \mathbb{R}^m$ une fonction \mathcal{C}^1 (c'est-à-dire que les dérivées partielles de b par rapport à toutes les variables existent et sont continues.)

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

De plus, on a la propriété suivante : si Y est une autre solution de (3.1) définie sur un intervalle I_Y , X et Y coïncident sur $I_Y \cap [t_0 - \tau, t_0 + \tau]$.

Remarque 3.1.4. Ce résultat assure que les courbes paramétrées étudiées au chapitre 1, définies à partir des solutions d'équations différentielles (système de Lotka-Volterra, pendule simple...) sont bien définies, 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 3.1.5. La condition $b \in \mathcal{C}^1$ peut être légèrement relaxée : il suffit que la fonction b soit lipschitzienne sur tout compact. 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 l'équation

$$\begin{aligned} x'(t) &= 2\sqrt{x(t)}, \quad t \in \mathbb{R}, \\ x(0) &= 0. \end{aligned} \tag{3.2}$$

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 (3.2) 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.

Corollaire 3.1.6. Soit X et Y deux solutions de (3.1) définies sur des intervalles I_X et I_Y respectivement. Alors X et Y coïncident sur $I_X \cap I_Y$.

Démonstration. On note X_{CL} la solution de (3.1) donnée par le théorème de Cauchy-Lipschitz 3.1.3, définie sur un intervalle $[t_0 - \tau, t_0 + \tau]$, avec $\tau > 0$. Alors d'après le théorème 3.1.3, la fonction X (resp. Y) coïncide avec X_{CL} sur $[t_0 - \tau, t_0 + \tau] \cap I_X$ (resp. $[t_0 - \tau, t_0 + \tau] \cap I_Y$). Par conséquent X et Y coïncident sur $I_X \cap I_Y \cap [t_0 - \tau, t_0 + \tau]$.

On définit

$$\begin{aligned} \tau_+ &= \sup\{\tau' > 0, X \text{ et } Y \text{ coïncident sur } I_X \cap I_Y \cap [t_0 - \tau, t_0 + \tau']\} \in \mathbb{R}_+ \cup +\infty, \\ \tau_- &= \sup\{\tau' > 0, X \text{ et } Y \text{ coïncident sur } I_X \cap I_Y \cap [t_0 - \tau', t_0 + \tau]\} \in \mathbb{R}_+ \cup +\infty. \end{aligned}$$

Alors $\tau_+, \tau_- > \tau$. Montrons à présent que X et Y coïncident sur $I_X \cap I_Y \cap]t_0 - \tau_-, t_0 + \tau_+[$: il existe des suites $(\tau_n^\pm)_{n \in \mathbb{N}}$ telles que $\tau_n^\pm \rightarrow \tau_\pm$ quand $n \rightarrow \infty$, $\tau_n^\pm < \tau_\pm$ pour tout n , et X et Y coïncident sur $I_X \cap I_Y \cap [t_0 - \tau_n^-, t_0 + \tau_n^+]$. Donc X et Y coïncident sur $\cup_n I_X \cap I_Y \cap [t_0 - \tau_n^-, t_0 + \tau_n^+] = I_X \cap I_Y \cap]t_0 - \tau_-, t_0 + \tau_+[$.

Montrons que

$$t_0 + \tau_+ \geq \sup I_X \cap I_Y, \quad t_0 - \tau_- \leq \inf I_X \cap I_Y. \tag{3.3}$$

Raisonnons par l'absurde et supposons par exemple que $t_0 + \tau_+ < \sup I_X \cap I_Y$. On en déduit que $t_0 + \tau_+ \in I_X \cap I_Y$ (puisque $I_X \cap I_Y$ est un intervalle.) Posons

$$x_1 = X(t_0 + \tau_+), \quad y_1 = Y(t_0 + \tau_+).$$

Alors $]t_0, t_0 + \tau_+[\subset I_X \cap I_Y \cap]t_0, t_0 + \tau_+[$, et donc

$$\begin{aligned} x_1 &= X(t_0 + \tau_+) = \lim_{\substack{t \rightarrow t_0 + \tau_+, \\ t \in]t_0, t_0 + \tau_+[}} X(t) \\ &= \lim_{\substack{t \rightarrow t_0 + \tau_+, \\ t \in]t_0, t_0 + \tau_+[}} Y(t) \\ &= Y(t_0 + \tau_+) = y_1. \end{aligned}$$

On applique alors le théorème de Cauchy-Lipschitz au point $t_0 + \tau_+$, avec la donnée initiale x_1 . Le résultat d'unicité assure que X et Y coïncident sur un petit intervalle à droite de $t_0 + \tau_+$. Il existe donc $\bar{\tau}_+ > \tau_+$ tel que X et Y coïncident sur $[t_0, t_0 + \bar{\tau}_+] \cap I_X \cap I_Y$. Par définition de τ_+ , on a $\bar{\tau}_+ \leq \tau_+$: absurde.

On en déduit que la propriété (3.3) est vérifiée. On remarque alors les propriétés suivantes :

1. Si $t_0 + \tau_+ > \sup I_X \cap I_Y$, alors $t_0 + \tau_+ = +\infty$;
2. Si $\sup I_X \cap I_Y \in I_X \cap I_Y$, avec le même raisonnement que ci-dessus, $X(\sup I_X \cap I_Y) = Y(\sup I_X \cap I_Y)$. On en déduit que dans ce cas X et Y coïncident sur tous les intervalles du type $[t_0 - \tau, t_0 + \tau'] \cap I_X \cap I_Y$ avec $\tau' > \sup I_X \cap I_Y$. Donc dans ce cas encore $t_0 + \tau_+ = +\infty$;
3. Il reste à examiner le cas où $t_0 + \tau_+ = \sup I_X \cap I_Y$ et $\sup I_X \cap I_Y \notin I_X \cap I_Y$. On remarque que dans ce cas

$$[t_0, t_0 + \tau_+[\cap I_X \cap I_Y = [t_0, t_0 + \tau'] \cap I_X \cap I_Y$$

pour tout $\tau' > \tau$. Comme X et Y coïncident sur $[t_0, t_0 + \tau_+[\cap I_X \cap I_Y$, on en déduit encore que $t_0 + \tau_+ = \infty$.

Donc dans tous les cas, $t_0 + \tau_+ = +\infty$. De même, $t_0 - \tau_- = -\infty$. Donc X et Y coïncident sur $I_X \cap I_Y \cap]t_0 - \tau_-, t_0 + \tau_+[= I_X \cap I_Y \cap \mathbb{R} = I_X \cap I_Y$. □

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 3.1.7. Soit X, Y deux solutions de (3.1), définies respectivement sur des intervalles I et J contenant t_0 . On dit que $X \prec Y$ si et seulement si $I \subset J$ et $X(t) = Y(t)$ pour tout $t \in I$.

Soit X_{max} une solution de (3.1). On dit que X_{max} est une solution maximale de (3.1) si $X \prec X_{max}$ pour toute solution X de (3.1).

Remarque 3.1.8. On vérifie facilement que \prec est une relation d'ordre sur les solutions de (3.1), et que les solutions maximales sont précisément les éléments maximaux pour cette relation d'ordre. Ainsi la solution maximale, si elle existe, est unique.

Proposition 3.1.9. 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 $b : I \times U \rightarrow \mathbb{R}^m$ une fonction \mathcal{C}^1 .

Alors il existe une solution maximale de l'équation (3.1), définie sur un intervalle ouvert $I_{max} \subset I$. D'après la remarque qui précède, cette solution 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 $[t_1, t_2] \subset I_{max}$ ¹. 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

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

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

et par conséquent

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

Or $[t_1, t_2] \subset [t_1, t_0] \cup [t_2, t_0]$. (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} .

Soit $X, Y \in \mathcal{S}$. On a alors

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

d'après le corollaire 3.1.6. 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

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

D'après la remarque ci-dessus, la définition de X_{max} ne dépend pas du choix de la solution X telle que $t \in I_X$. X_{max} est donc définie de façon univoque.

Troisième étape : I_{max} est un intervalle ouvert et X_{max} est solution de (3.1).

Soit $\bar{t} \in I_{max}$ quelconque, et soit $\bar{x} = X_{max}(\bar{t})$. Par définition de I_{max}, X_{max} , il existe une solution \bar{X}_1 de (3.1) telle que $\bar{t} \in I_{\bar{X}_1}$ et $\bar{X}_1(\bar{t}) = \bar{x}$.

D'autre part, d'après le théorème de Cauchy-Lipschitz (3.1.3), il existe $\bar{\tau} > 0$ tel que l'équation

$$X'(t) = b(t, X(t)), X(\bar{t}) = \bar{x}$$

possède une unique solution \bar{X}_2 sur $[\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}]$. Ainsi, on a nécessairement $\bar{X}_1 = \bar{X}_2$ sur $[\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}] \cap I_{\bar{X}_1}$. Remarquons par ailleurs que l'intervalle $[\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}] \cap I_{\bar{X}_1}$ n'est pas réduit à \bar{t} : en effet, si $\bar{t} = t_0$, alors on peut toujours, grâce au Théorème de Cauchy-Lipschitz, choisir \bar{X}_1 de telle sorte que $I_{\bar{X}_1}$ contienne $[t_0 - \tau, t_0 + \tau]$. Et si $\bar{t} \neq t_0$, $I_{\bar{X}_1}$ contient $[t_0, \bar{t}]$, donc contient un voisinage à gauche ou à droite de \bar{t} .

Par conséquent $X_{max}(t) = \bar{X}_2(t)$ sur $[\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}] \cap I_{\bar{X}_1}$. On définit à présent, sur $I_{\bar{X}_1} \cup [\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}]$,

$$\tilde{X}(t) = \begin{cases} \bar{X}_1(t) & \text{si } t \in I_{\bar{X}_1}, \\ X_2(t) & \text{si } t \in [\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}]. \end{cases}$$

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,

$$[\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}] \subset I_{\bar{X}_1} \cup [\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}] \subset I_{max}. \quad (3.4)$$

De plus, $X_{max}(t) = \tilde{X}(t) = X_2(t)$ pour tout $t \in [\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}]$, et donc X_{max} est solution de l'équation

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

sur l'intervalle $[\bar{t} - \bar{\tau}, \bar{t} + \bar{\tau}]$.

Comme la propriété (3.4) est vraie pour tout $\bar{t} \in I_{max}$, on en déduit que I_{max} est un intervalle ouvert, et que l'équation différentielle ci-dessus est vérifiée en tout point de I_{max} . Comme de plus on a clairement $X_{max}(t_0) = x_0$ par définition de X_{max} (toujours en utilisant le théorème de Cauchy-Lipschitz), on en déduit que X_{max} est solution de (3.1).

Quatrième étape : X_{max} est une solution maximale.

Cette propriété est une conséquence immédiate des définitions de X_{max}, I_{max} et du fait que X_{max} est solution de (3.1). □

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 possible : 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.

Proposition 3.1.10 (Dichotomie existence globale/explosion). *On suppose que les hypothèses du théorème 3.1.3 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} .*

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

- soit $\sup I_{max} = +\infty$;
- soit $\sup I_{max} < \infty$, et dans ce cas il existe une suite $(t_n)_{n \in \mathbb{N}}$ de I_{max} telle que $\lim_{n \rightarrow \infty} t_n = \sup I_{max}$ et telle que

$$\lim_{n \rightarrow \infty} \|X_{max}(t_n)\| = \infty.$$

Démonstration. Commençons par un peu de logique. On rappelle que si P et Q sont deux assertions logiques, les assertions " P ou Q " et " $\text{non } P \Rightarrow Q$ " sont équivalentes. Ici, on va donc montrer

$$\text{non} \left(\begin{array}{l} \text{il existe une suite } (t_n)_{n \in \mathbb{N}} \text{ de } I_{max} \text{ telle que } \lim_{n \rightarrow \infty} t_n = \sup I_{max} \\ \text{et telle que } \lim_{n \rightarrow \infty} \|X_{max}(t_n)\| = \infty \end{array} \right) \Rightarrow \sup I_{max} = \infty.$$

Remarquons tout d'abord que la propriété

$$\text{" il existe une suite } (t_n)_{n \in \mathbb{N}} \text{ de } I_{max} \text{ telle que } \lim_{n \rightarrow \infty} t_n = \sup I_{max} \text{ et telle que } \lim_{n \rightarrow \infty} \|X_{max}(t_n)\| = \infty \text{"}$$

est équivalente à (cela est valable pour $\sup I_{max} \in \mathbb{R}$ ou $\sup I_{max} = \infty$)

$$\forall A > 0, \forall V \text{ voisinage de } \sup I_{max}, \exists t \in V \cap I_{max}, \|X(t)\| \geq A.$$

De plus, la négation de la propriété précédente est

$$\exists A > 0, \exists V \text{ voisinage de } \sup I_{max}, \forall t \in V \cap I_{max}, \|X(t)\| < A,$$

et cette dernière propriété signifie que X_{max} est bornée sur un voisinage de $\sup I_{max}$.

Pour démontrer la proposition, il suffit donc de montrer le lemme suivant (dont le résultat est intéressant en lui même) :

Lemme 3.1.11. *On suppose que les hypothèses de la Proposition 3.1.10 sont vérifiées. On suppose que X_{max} est bornée sur un voisinage de $\sup I_{max}$. Alors*

$$\sup I_{max} = +\infty.$$

Preuve du Lemme. Soit V un voisinage de $\sup I_{max}$ sur lequel X_{max} est bornée, et soit $R = \sup_{t \in V} \|X_{max}(t)\|$. Raisonnons par l'absurde et supposons que $\sup I_{max} < \infty$. Il existe donc un intervalle compact \bar{V} contenant V .

On pose alors

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

Comme b est de classe \mathcal{C}^1 , 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 $t, t' \in V$, on a

$$\|X_{max}(t) - X_{max}(t')\| \leq C_0|t - t'|.$$

On en déduit que si $(t_n)_{n \in \mathbb{N}}$ est une suite quelconque de I_{max} tendant vers $\sup I_{max}$, 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 $(t'_n)_{n \in \mathbb{N}}$ est une autre suite de I_{max} tendant vers $\sup I_{max}$, on pose, pour $n \in \mathbb{N}$,

$$\tau_n = \begin{cases} t_{\frac{n}{2}} & \text{si } n \text{ est pair,} \\ t'_{\frac{n-1}{2}} & \text{si } n \text{ est impair.} \end{cases}$$

On voit alors que $(\tau_n)_{n \in \mathbb{N}}$ est également une suite de I_{max} tendant vers $\sup I_{max}$. 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

$$\lim_{n \rightarrow \infty} X_{max}(t_n) = \lim_{n \rightarrow \infty} X_{max}(t'_n).$$

Soit $\bar{X} \in \mathbb{R}^m$ la limite commune de toutes les suites $(X_{max}(t_n))_{n \in \mathbb{N}}$ vérifiant les hypothèses précédentes. On voit aisément que

$$\lim_{t \rightarrow \sup I_{max}} X_{max}(t) = \bar{X}.$$

On considère alors le problème de Cauchy en $\sup I_{max}$, avec \bar{X} comme condition initiale. D'après le théorème de Cauchy-Lipschitz, il existe une solution de l'équation différentielle sur un voisinage ouvert de $\sup I_{max}$. Cette solution coïncide avec X_{max} sur un voisinage à gauche de $\sup I_{max}$. On peut donc prolonger X_{max} au delà de $\sup I_{max}$, ce qui contredit le fait que X_{max} est une solution maximale : absurde. □

□

Exemple 3.1.12. *On considère l'équation différentielle*

$$x' = x^2, \quad x(0) = 1.$$

On remarque la fonction constante égale à zéro est solution de $x' = x^2$. Par conséquent, d'après le corollaire (3.1.6), la solution maximale x_{max} du problème de Cauchy ci-dessus ne s'annule pas sur son intervalle de définition I_{max} .

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

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

et finalement

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

ce qui s'écrit encore

$$x(t) = \frac{1}{1-t} \quad \forall t \in I_{max} \cap]-\infty, 1[.$$

En particulier : x est bornée au voisinage de $\inf I_{max}$, donc $\inf I_{max} = -\infty$. De surcroît, $1 \notin I_{max}$ (sinon x aurait une limite finie en 1^- .) On en déduit finalement que

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

3.2 Quelques résultats plus avancés

Dans le cas des équations autonomes, le théorème de Cauchy-Lipschitz a plusieurs conséquences frappantes. Un premier 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 3.2.1. *Soit $b \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R}^m)$, $t_0 \in \mathbb{R}$, $X_0 \in \mathbb{R}^m$. On considère l'équation différentielle*

$$\begin{aligned} X'(t) &= b(X(t)), \\ X(t_0) &= X_0. \end{aligned} \tag{3.5}$$

Soit X la solution maximale de (3.5), et soit I son intervalle (ouvert) de définition. On suppose qu'il existe $T \in \mathbb{R}^$ tel que $t_0 + T \in I$ et tel que*

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

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

Démonstration. On pose $I_T = -T + I = \{t - T, t \in I\}$. On remarque que I_T est un intervalle ouvert et que comme $T + t_0 \in I$, $t_0 \in I_T$. Pour tout $t \in I_T$, on pose

$$\tilde{X}(t) = X(t + T).$$

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

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

Donc \tilde{X} est solution de (3.5). Comme X est la solution maximale, on en déduit que

$$\begin{aligned} I_T &\subset I, \\ \tilde{X}(t) &= X(t) \quad \forall t \in I_T. \end{aligned}$$

Pour fixer les idées, à ce stade, supposons $T > 0$ (le raisonnement est strictement analogue si $T < 0$). Comme $I_T \subset I$, nécessairement $\inf I = -\infty$ et I est de la forme $] -\infty, a[$ avec $a \in]t_0, +\infty[$. On déduit de la deuxième propriété que

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

ce qui s'écrit encore

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

En vertu du lemme 3.1.11, on voit que nécessairement $a = +\infty$: en effet, si on a $a < \infty$, alors

$$\lim_{t \rightarrow a^-} X(t) = X(a - T) \in \mathbb{R}^m,$$

donc X reste borné dans un voisinage à gauche de a : donc $a = +\infty$: absurde.

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

Pour des solutions à valeurs réelles, l'unicité des solutions a une autre conséquence remarquable : le non-croisement de solutions, ou encore le fait que les solutions sont ordonnées.

Proposition 3.2.2. Soit $b : I \times \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , $t_0 \in I$, et $x_0, x_1 \in \mathbb{R}$.
On considère les solutions maximales X_0, X_1 de l'équation différentielle

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

relatives aux conditions initiales $X_0(t_0) = x_0$, $X_1(t_0) = x_1$ respectivement. On note I_0, I_1 les intervalles de définition de X_0, X_1 .

On suppose que $x_0 < x_1$. Alors

$$X_0(t) < X_1(t) \quad \forall t \in I_0 \cap I_1.$$

Démonstration. Raisonnons par l'absurde et supposons qu'il existe $\tau \in I_1 \cap I_0$ tel que

$$X_0(\tau) \geq X_1(\tau).$$

De deux choses l'une : soit $X_0(\tau) = X_1(\tau)$, soit $X_0(\tau) > X_1(\tau)$, et dans ce cas il existe $\tau' \in]t_0, \tau[$ tel que $X_0(\tau') = X_0(\tau)$. Dans tous les cas, on obtient l'existence de $\tau_0 \in I_0 \cap I_1$ tel que $X_0(\tau_0) = X_1(\tau_0) =: x_2$.

On note \bar{X} la solution maximale de l'équation différentielle telle que $\bar{X}(\tau_0) = x_2$, et on note également \bar{I} son intervalle de définition. Comme \bar{X} est une solution maximale, nécessairement

$$I_0 \subset \bar{I}, \quad I_1 \subset \bar{I}, \\ X_0(t) = \bar{X}(t) \quad \forall t \in I_0, \quad X_1(t) = \bar{X}(t) \quad \forall t \in I_1.$$

Donc

$$X_0(t) = X_1(t) = \bar{X}(t) \quad \forall t \in I_0 \cap I_1.$$

Or par définition, $t_0 \in I_1 \cap I_0$. Donc $x_0 = x_1$: absurde. □

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 3.1.11). Ce lemme se généralise de la façon suivante :

Lemme 3.2.3. Soit $b : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ 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 $f : \mathbb{R} \rightarrow \mathbb{R}_+$, continue, telle que

$$\|X(t)\| \leq f(t) \quad \forall t \in I_{max}.$$

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 de $\sup I_{max}$. Comme f est continue sur \mathbb{R} , f est bornée sur K .

En utilisant l'hypothèse du lemme, on en déduit que X est bornée sur K . D'après le lemme 3.1.11, $\sup I_{max} = +\infty$: 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 b) qui assurent que l'on a une solution globale. Pour cela on utilise le résultat suivant :

Lemme 3.2.4 (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 $t \geq t'$ dans I , on a

$$\varphi(t) \leq \varphi(t') \exp\left(\int_{t'}^t \psi\right) + C \left(\exp\left(\int_{t'}^t \psi\right) - 1\right).$$

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 $\exp\left(-\int_{t'}^t \psi\right)$ (on remarque que cette quantité est toujours positive). On obtient alors

$$\frac{d}{dt} \left(\varphi(t) \exp\left(-\int_{t'}^t \psi\right) \right) \leq C\psi(t) \exp\left(-\int_{t'}^t \psi\right) \quad \forall t \in I. \quad (3.6)$$

Si $t \geq t'$: on intègre cette inégalité entre t' et t :

$$\varphi(t) \exp\left(-\int_{t'}^t \psi\right) - \varphi(t') \leq C \int_{t'}^t \psi(s) \exp\left(-\int_{t'}^s \psi\right) ds = C \left(1 - \exp\left(-\int_{t'}^t \psi\right)\right).$$

En multipliant de nouveau par $\exp\left(\int_{t_0}^t \psi\right)$, on obtient l'estimation annoncée. □

On arrive finalement au résultat suivant :

Proposition 3.2.5. *Soit $b : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^1 . On suppose qu'il existe une fonction continue $\psi : \mathbb{R} \rightarrow \mathbb{R}_+$ telle que*

$$\|b(t, X)\| \leq \psi(t)(1 + \|X\|) \quad \forall X \in \mathbb{R}^m.$$

Soit $t_0 \in \mathbb{R}$, $X_0 \in \mathbb{R}^m$, et soit X_{max} la solution maximale de l'équation (3.1), définie sur un intervalle I_{max} . Alors $I_{max} = \mathbb{R}$.

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 I_{max} , et pour tout $t \in I_{max}$, on a, en utilisant l'équation puis l'inégalité de Cauchy-Schwarz

$$\varphi'(t) = 2X'_{max}(t) \cdot X_{max}(t) = 2b(t, X_{max}(t)) \cdot X_{max}(t) \quad (3.7)$$

$$|\varphi'(t)| \leq 2\psi(t)(1 + \|X_{max}(t)\|)\|X_{max}(t)\|. \quad (3.8)$$

En utilisant l'inégalité

$$a \leq \frac{1}{2} + \frac{a^2}{2} \quad \forall a \in \mathbb{R},$$

on obtient finalement, pour tout $t \in I$,

$$\varphi'(t) \leq 2\psi(t) \left(\frac{1}{2} + \frac{3}{2}\varphi(t) \right) \leq 3\psi(t) \left(\frac{1}{3} + \varphi(t) \right),$$

et également

$$-\varphi'(t) \leq 2\psi(t) \left(\frac{1}{2} + \frac{3}{2}\varphi(t) \right) \leq -3\psi(t) \left(-\frac{1}{3} - \varphi(t) \right).$$

En utilisant le Lemme de Gronwall, on déduit de la première inégalité, pour tout $t \geq t_0$,

$$\varphi(t) \leq \varphi(t_0) \exp \left(\int_{t_0}^t 3\psi(s) ds \right) + \frac{1}{3} \left(\exp \left(\int_{t_0}^t 3\psi(s) ds \right) - 1 \right).$$

Pour majorer $\varphi(t)$ pour $t \leq t_0$, on utilise la deuxième inégalité :

$$-\varphi(t_0) \leq -\varphi(t) \exp \left(- \int_t^{t_0} 3\psi(s) ds \right) - \frac{1}{3} \left(\exp \left(- \int_t^{t_0} 3\psi(s) ds \right) - 1 \right),$$

ce qui donne finalement

$$\varphi(t) \leq \varphi(t_0) \exp \left(\int_t^{t_0} 3\psi(s) ds \right) + \frac{1}{3} \left(\exp \left(\int_t^{t_0} 3\psi(s) ds - 1 \right) \right).$$

Les fonctions φ, ψ sont à valeurs positives ou nulles, de sorte que les membres de droite des inégalités ci-dessus sont positifs ou nuls pour tout $t \in I$. Posons, pour $t \in \mathbb{R}$,

$$f(t) = \left[\varphi(t_0) \exp \left(\int_{t_0}^t 3\psi(s) ds \right) + \frac{1}{3} \left(\exp \left(\int_{t_0}^t 3\psi(s) ds \right) - 1 \right) \right]^{1/2} \quad \text{si } t \geq t_0,$$

et

$$f(t) = \left[\varphi(t_0) \exp \left(\int_t^{t_0} 3\psi(s) ds \right) + \frac{1}{3} \left(\exp \left(\int_t^{t_0} 3\psi(s) ds - 1 \right) \right) \right]^{1/2} \quad \text{si } t \leq t_0.$$

Alors f est continue sur \mathbb{R} et

$$\|X(t)\| \leq f(t) \quad \forall t \in I.$$

En utilisant le Lemme 3.2.3, on en déduit que $I = \mathbb{R}$.

□

3.3 Exemples

3.3.1 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 :

$$\begin{cases} \frac{dN}{dt} = aN(t) - bN(t)P(t), \\ \frac{dP}{dt} = cN(t)P(t) - dP(t), \end{cases} \quad (3.9)$$

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 (3.9) munie des conditions initiales $N(t = 0) = N_0, P(t = 0) = P_0$. D’après le théorème de Cauchy-Lipschitz 3.1.3 et ses conséquences, l’équation (3.9) admet, pour tout couple $(N_0, P_0) \in \mathbb{R}_+^{*2}$, une unique solution maximale (N, P) définie sur un intervalle ouvert I contenant 0.

Remarque 3.3.1. *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 $\{N(t), P(t)\}_{t \in \mathbb{R}_+ \cap I}$.*

“Énergie” du système

Remarquons que la solution (N, P) obtenue à l’aide du théorème de Cauchy-Lipschitz est de classe \mathcal{C}^1 sur I . Par conséquent, comme $N_0 > 0, P_0 > 0$, il existe un intervalle ouvert $I_0 \subset I$ tel que pour tout $t \in I_0$, $N(t) > 0, P(t) > 0$. On peut considérer l’intervalle I_0 maximal satisfaisant une telle propriété, défini par $I_0 =]T_-, T_+[$, où

$$\begin{aligned} T_+ &= \inf\{t \in I \cap \mathbb{R}_+, N(t) = 0 \text{ ou } P(t) = 0\}, \\ T_- &= \sup\{t \in I \cap \mathbb{R}_-, N(t) = 0 \text{ ou } P(t) = 0\}. \end{aligned}$$

On a alors la propriété suivante, qui permet de mieux comprendre le comportement des solutions et de tracer les trajectoires :

Proposition 3.3.2. *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 (3.9) telle que $N(0) = N_0, P(0) = P_0$, on a, pour tout $t \in I_0$

$$H(N(t), P(t)) = H(N_0, P_0).$$

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 la proposition ?? puisque $(N(t), P(t)) \in \mathbb{R}_+^{*2}$.

En utilisant l'équation (3.9), on obtient

$$\begin{aligned} \frac{d}{dt}H(N(t), P(t)) &= b\frac{dP}{dt} + c\frac{dN}{dt} - a\frac{dP}{dt}\frac{1}{P(t)} - d\frac{dN}{dt}\frac{1}{N(t)} \\ &= bcP(t)N(t) - bdP(t) + caN(t) - bcP(t)N(t) - acN(t) + ad - da + dbP(t) \\ &= 0. \end{aligned}$$

On en déduit que la quantité $H(N(t), P(t))$ reste constante le long de chaque trajectoire. \square

Corollaire 3.3.3. *Pour tout $(N_0, P_0) \in \mathbb{R}_+^{*2}$, on a $I = I_0 = \mathbb{R}$. De plus, il existe $m, M > 0$, dépendant uniquement de $H_0 = H(N_0, P_0)$, tel que pour tout $t \in \mathbb{R}$,*

$$m \leq N(t) \leq M, \quad m \leq P(t) \leq M.$$

Démonstration. On définit les fonctions $\varphi_{b,a}, \varphi_{c,d}$ par

$$\begin{aligned} \varphi_{b,a} : x \in \mathbb{R}_+^* &\mapsto bx - a \ln x, \\ \varphi_{c,d} : x \in \mathbb{R}_+^* &\mapsto cx - d \ln x, \end{aligned} \tag{3.10}$$

de sorte que

$$H(x, y) = \varphi_{b,a}(y) + \varphi_{c,d}(x) \quad \forall (x, y) \in \mathbb{R}_+^{*2}.$$

x	0	$\frac{d}{c}$	$+\infty$
$\varphi'_{c,d}(x)$		0	
$\varphi_{c,d}(x)$	$+\infty$	$d(1 - \ln \frac{d}{c})$	$+\infty$

Le tableau de variations de $\varphi_{c,d}$ est
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 $t \in I_0$,

$$\varphi_{c,d}(N(t)) \leq H(N_0, P_0) - \varphi_{b,a}(P(t)) \leq H(N_0, P_0) - a(1 - \ln \frac{a}{b}).$$

Ainsi

$$\sup_{t \in I_0} \varphi_{c,d}(N(t)) \leq H(N_0, P_0) - a(1 - \ln \frac{a}{b}),$$

et de même

$$\sup_{t \in I_0} \varphi_{b,a}(P(t)) \leq H(N_0, P_0) - d(1 - \ln \frac{d}{c}).$$

On déduit des tableaux de variations de $\varphi_{c,d}$ et $\varphi_{b,a}$ qu'il existe des constantes strictement positives m, M telles que

$$m \leq N(t) \leq M, \quad m \leq P(t) \leq M \quad \forall t \in I_0.$$

En revenant à la définition de I_0 , on en déduit que $I_0 = I$. Les solutions de (3.9) restent donc bornées sur leur intervalle maximal d'existence. La Proposition 3.1.10 implique alors que $I = \mathbb{R}$. □

Tracé des courbes

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 (3.9) de donnée initiale (N_0, P_0) .

D'après la proposition 3.3.2, ces courbes paramétrées sont 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 observe alors que $\varphi_{c,d}$ (resp. $\varphi_{d,a}$) 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 $[a(1 - \ln \frac{a}{b}), +\infty[$). En particulier, il existe

$$x_- < d \left(1 - \ln \frac{d}{c}\right) < x_+$$

tels que

$$\varphi_{c,d}(x_{\pm}) = H_0 - a \left(1 - \ln \frac{a}{b}\right) > d \left(1 - \ln \frac{d}{c}\right).$$

Soit $(x, y) \in \mathcal{D}_{H_0}$ quelconque. On a

$$\varphi_{c,d}(x) = H_0 - \varphi_{b,a}(y) \leq H_0 - a \left(1 - \ln \frac{a}{b}\right),$$

et d'après le tableau de variations de $\varphi_{c,d}$, $x \in [x_-, x_+]$. De même, on peut définir

$$y_- < a \left(1 - \ln \frac{a}{b}\right) < y_+$$

tels que

$$\varphi_{b,a}(y_{\pm}) = H_0 - d \left(1 - \ln \frac{d}{c}\right),$$

et on a

$$(x, y) \in \mathcal{D}_{H_0} \Rightarrow x_- \leq x \leq x_+ \text{ et } y_- \leq y \leq y_+.$$

On trace alors la courbe \mathcal{D}_{H_0} en séparant ses deux branches $y \leq a \left(1 - \ln \frac{a}{b}\right)$ et $y \geq a \left(1 - \ln \frac{a}{b}\right)$. 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 $[a \left(1 - \ln \frac{a}{b}\right), +\infty[$. Si $y \leq a \left(1 - \ln \frac{a}{b}\right)$, on a

$$(x, y) \in \mathcal{D}_{H_0} \Leftrightarrow y = \psi_-(H_0 - \varphi_{c,d}(x)),$$

et pour $y \geq a \left(1 - \ln \frac{a}{b}\right)$, 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 $[a \left(1 - \ln \frac{a}{b}\right), +\infty[$, dérivable sur $]a \left(1 - \ln \frac{a}{b}\right), +\infty[$, et sa dérivée est

$$\psi'_{\pm} = \frac{1}{\varphi_{b,a} \circ \psi_{\pm}}.$$

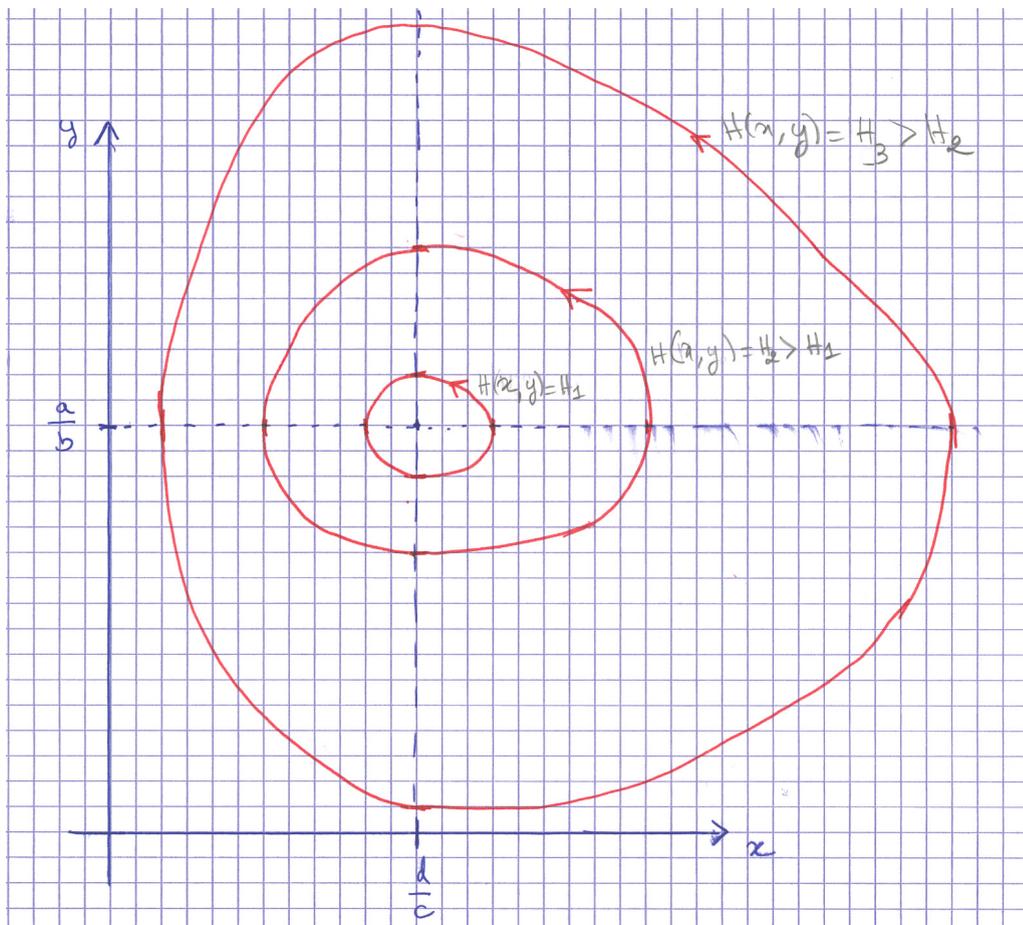
Par conséquent les fonctions f_+ et f_- sont dérivables sur $]x_-, x_+[$, de dérivée

$$f'_{\pm}(x) = -\varphi'_{c,d}(x) \frac{1}{\varphi_{b,a} \circ f_{\pm}(x)}.$$

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 l'allure des courbes suivantes : Finalement, on obtient une série de courbes qui ont l'allure représentée sur la figure 3.3.1. 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 $(N_0, P_0) = \left(\frac{d}{c}, \frac{a}{b}\right)$, alors on a

$$H(N_0, P_0) = \min_{\mathbb{R}_+^* \times \mathbb{R}_+^*} H = d\left(1 - \ln \frac{d}{c}\right) + a\left(1 - \ln \frac{a}{b}\right).$$

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 $\left(\frac{d}{c}, \frac{a}{b}\right)$ 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 $(N_0, P_0) \neq \left(\frac{d}{c}, \frac{a}{b}\right)$. 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 $(N_0, P_0) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \setminus \left\{\left(\frac{d}{c}, \frac{a}{b}\right)\right\}$. 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 3.2.1, les trajectoires sont périodiques de période T .

Remarque 3.3.4. *La propriété importante utilisée ici est que si les trajectoires sont portées par des courbes fermées, et qu'elles ne s'arrêtent pas, alors le mouvement est périodique.*

Portrait de phase du pendule simple

On considère les solutions de l'équation (1.2)

$$\theta'' + \frac{g}{L} \sin \theta = 0. \quad (3.11)$$

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 (3.11) 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

$$y = \sqrt{2(E_0 + \cos x)}. \quad (3.12)$$

On a donc une équation cartésienne explicite pour chaque courbe. On va commencer par tracer les courbes d'équation (3.12) 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

$$\phi'_{E_0} = \frac{-\sin x}{\sqrt{2(E_0 + \cos x)}}.$$

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
$\phi'_{E_0}(x)$	\emptyset	+	\emptyset	-	\emptyset
ϕ_{E_0}					

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

$$\sin x = \sin(\pi + h) = -\sin h = -h + O(h^3),$$

$$1 + \cos x = 1 + \cos(\pi + h) = 1 - \cos h = \frac{h^2}{2} + O(h^4).$$

On en déduit que

$$\phi'_1(x) \sim \frac{h}{|h|}.$$

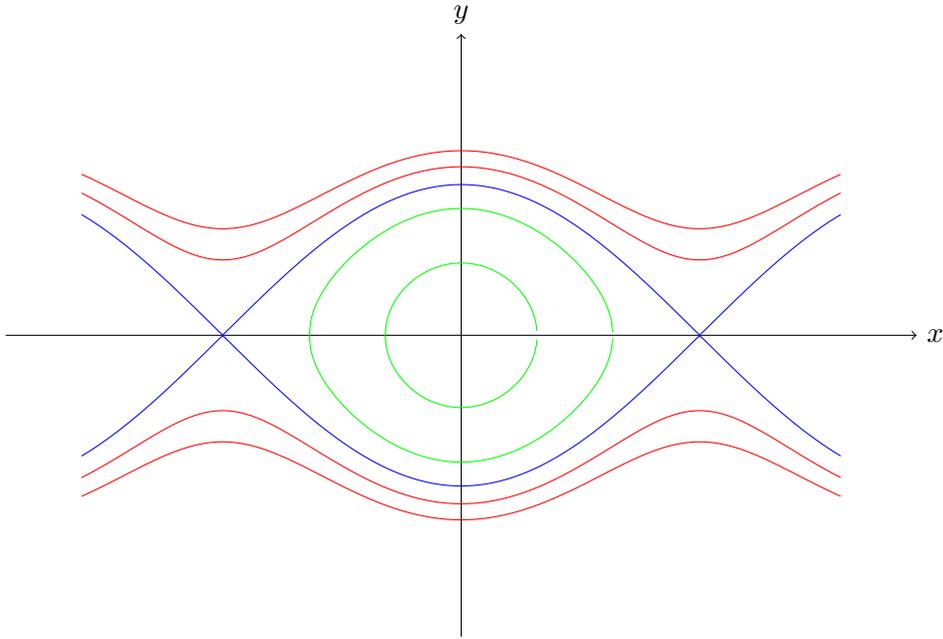
Ainsi la courbe représentative de ϕ_1 admet une tangente à gauche de pente -1 et une tangente à droite de pente 1 en $x = \pi$.

- 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
$\phi'_{E_0}(x)$		+	\emptyset	-	
ϕ_{E_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 croissante 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 trigonométrique. 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 $\tilde{\theta}(T) = \theta(T)$, $\tilde{\theta}'(T) = \theta'(T)$. En utilisant le résultat d'unicité que l'on verra au Chapitre 4, 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

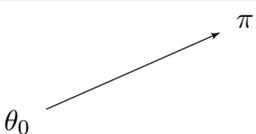
$$\frac{1}{\sqrt{2(1 + \cos \theta)}} \sim \frac{1}{\pi - \theta},$$

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 $v_0 < 0$).

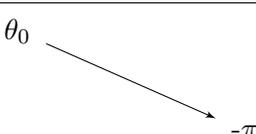
On obtient ainsi le tableau de variations suivant si $v_0 > 0$:

t	0	$+\infty$
$\theta'(t)$	+	
$\theta(t)$	θ_0	π



Si $v_0 < 0$, le tableau de variations est analogue :

t	0	$+\infty$
$\theta'(t)$	-	
$\theta(t)$	θ_0	$-\pi$



3.4 Recollements

Toutes les équations différentielles étudiées jusqu'ici étaient de la forme

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

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

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

où la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est de classe \mathcal{C}^1 , et $g : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ 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 f , et, par suite, les intervalles sur lesquels f ne s'annule pas ;
- Sur chacun des intervalles sur lequel f 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

$$\bar{u}(t) = C \exp\left(-\int^t \frac{2s}{1 - s^2} ds\right) = \frac{C}{1 - t^2}.$$

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 $C(t) = t + C_j$.

Bilan : sur chacun des intervalles I_j , la solution est de la forme

$$u_j(t) = \frac{t + C_j}{1 - t^2} = \frac{t + C_j}{(1 - t)(1 + t)}.$$

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. **On ne peut pas recoller.**

Exemple 2

On considère une variante de l'équation "explosive" déjà rencontrée précédemment, soit

$$tu'(t) = u(t)^2, \quad t \in \mathbb{R}.$$

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 (3.2.2), 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

$$1 - a_{\pm} - \ln \left| \frac{t}{t_{\pm}} \right| > 0.$$

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

$$I_+ =]0, t_+ \exp\left(\frac{1}{a_+}\right)[, \quad I_- =]t_- \exp\left(\frac{1}{a_-}\right), 0[.$$

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

$$u(t) = \begin{cases} -\frac{t^2}{3} & \text{si } t \leq 0, \\ t^2 & \text{si } t > 0, \end{cases}$$

3.5 Équations à variables séparées

Le but est de résoudre des équations du type

$$y'(t) = g(y(t))f(t), \quad (3.13)$$

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 (3.13). 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 (3.13), 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 3.5.1. *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 3.5.2. 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 3.5.3. 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 C^1 sur U_0 , et sa dérivée $1/g$ ne s'annule pas sur U_0 . Par conséquent H est un C^1 difféomorphisme de U_0 sur $H(U_0)$. □

Remarque 3.5.4. *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 $y(0) = y_0 \neq 0$.

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 zéro.

Pour $t \in J$, on a

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

soit

$$\frac{1}{2} \frac{dy(t)^2}{dt} = 1.$$

On obtient donc

$$y(t)^2 - y_0^2 = 2t,$$

et

$$\forall t \in J, \quad y(t)^2 = y_0^2 + 2t.$$

On en déduit avec les arguments habituels que

$$J = \left] -\frac{y_0^2}{2}, +\infty \right[.$$

Puisque $y(t) \in U$ pour tout $t \in J$, on a

$$y(t) = \operatorname{sgn}(y_0) \sqrt{y_0^2 + 2t} \quad \forall t \in J.$$

Exemple 2

On considère l'équation différentielle

$$y'(t) = y(t) - y(t)^2,$$

avec $y(0) = y_0 \in \mathbb{R}$.

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 $t \in J$,

$$\ln \left(1 - \frac{1}{1-y(t)} \right) = t + t_0,$$

où

$$t_0 = \ln \left(1 - \frac{1}{1-y_0} \right) < 0.$$

En inversant la fonction du membre de gauche, on obtient finalement

$$y(t) = -\frac{e^{t+t_0}}{1 - e^{t+t_0}} \quad \forall t \in J,$$

d'où l'on déduit finalement que $J =]-\infty, -t_0[$.

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 $t \in J$,

$$\ln \left(\frac{1}{1-y(t)} - 1 \right) = t + t_0,$$

où

$$t_0 = \ln \left(\frac{y_0}{1-y_0} \right) \in \mathbb{R}.$$

Dans ce cas on obtient

$$y(t) = \frac{e^{t+t_0}}{1+e^{t+t_0}} \quad \forall t \in J,$$

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 $t \in J$,

$$\ln \left(1 + \frac{1}{y(t)-1} \right) = t + t_0, \quad \text{où } t_0 = \ln \left(1 + \frac{1}{y_0-1} \right) > 0.$$

En inversant une nouvelle fois la formule, on arrive à

$$y(t) = \frac{e^{t+t_0}}{e^{t+t_0}-1} \quad \forall t \in J,$$

et donc $J =]-t_0, +\infty[$.

3.6 Analyse numérique

Les schémas d'Euler qui ont été définis au chapitre ?? peuvent évidemment être généralisés à un cadre non linéaire pour une équation de la forme

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

On définit ainsi le schéma d'Euler explicite par la formule de récurrence

$$X_{k+1} = X_k + hb(t_k, X_k),$$

et le schéma d'Euler implicite par la formule (implicite!)

$$X_{k+1} = X_k + hb(t_{k+1}, X_{k+1}).$$

L'analyse générale du caractère bien défini de ces schémas et de leur convergence sort du cadre de ce cours.

Appendice : vecteur tangent à une courbe paramétrée

Pour analyser le comportement de la courbe au voisinage du point a , on suppose que les fonctions f et g sont de classe \mathcal{C}^∞ , de sorte que l'on peut écrire un développement limité de f et g à tout ordre au voisinage de a . Notons $M(t)$ le point de \mathcal{C} de paramètre t , c'est-à-dire le point de coordonnées $(f(t), g(t))$. On a alors, pour tout $n \in \mathbb{N}$,

$$M(t) = M_0 + \sum_{k=1}^n M_k (t-a)^k + o((t-a)^n),$$

où M_k est le vecteur de coordonnées $(f^{(k)}, g^{(k)})(a)/k!$.

Proposition 3.6.1. *Soit $1 \leq m < n$ les plus petits entiers (strictement positifs) tels que les vecteurs M_m et M_n sont indépendants (s'ils existent!).*

La courbe admet alors en $M(a)$ une tangente de vecteur directeur M_m .

De plus, on a les cas suivants :

- *m impair, n pair (par exemple : $m = 1, n = 2$) : la courbe reste du même côté de sa tangente au voisinage de $M(a)$: il s'agit d'un point ordinaire ;*
- *m impair, n impair (par exemple : $m = 1, n = 3$) : la courbe admet un point d'inflexion en $M(a)$;*
- *m pair, n impair (par exemple : $m = 2, n = 3$) : la courbe admet un point de rebroussement de première espèce ;*
- *m pair, n pair (par exemple : $m = 2, n = 4$) : la courbe admet un point de rebroussement de deuxième espèce.*

Dans ce cours, on appellera points réguliers les points ordinaires et les points d'inflexion.

Démonstration. Tout d'abord, m est le plus petit entier strictement positif tel que M_m est non nul. On a donc

$$M(t) - M(a) = M(t) - M_0 = M_m (t-a)^m + o(|t-a|^m).$$

Par conséquent,

$$\|M(t) - M(a)\| \sim \|M_m\| |t-a|^m,$$

et

$$\frac{M(t) - M(a)}{\|M(t) - M(a)\|} = \frac{M_m}{\|M_m\|} \left(\frac{t-a}{|t-a|} \right)^m + o(1).$$

On en déduit que M_m est un vecteur tangent à la courbe en M_a .

Pour déterminer l'allure de la courbe au voisinage de $M(a)$, il faut étudier le signe de $(M(t) - M(a)) \cdot M_m$ d'une part, et de $(M(t) - M(a)) \cdot M_m^\perp$ d'autre part, où M_m^\perp est un vecteur orthogonal au vecteur M_m , autrement dit un vecteur normal à la courbe.

- Si $(M(t) - M(a)) \cdot M_m$ est de signe constant au voisinage de a , cela signifie que la courbe est entièrement incluse dans un demi-plan affine passant par $M(a)$ et de vecteur normal M_m . Par conséquent, la courbe admet un point de rebroussement en M_m .
- Si $(M(t) - M(a)) \cdot M_m$ change de signe au voisinage de a , alors la courbe traverse la droite passant par $M(a)$ et de vecteur normal M_m . On est donc dans un cas régulier (pas de point de rebroussement).
- Si $(M(t) - M(a)) \cdot M_m^\perp$ est de signe constant au voisinage de a , cela signifie que la courbe est entièrement incluse dans un demi-plan affine passant par $M(a)$ et de vecteur normal M_m^\perp . Autrement dit, la courbe ne traverse pas sa tangente en $M(a)$.
- Si $(M(t) - M(a)) \cdot M_m^\perp$ change de signe au voisinage de a , cela signifie que la courbe se situe de part et d'autre de sa tangente en a . S'il y a un point de rebroussement, c'est un point de rebroussement de première espèce. Si on est dans un cas régulier, la courbe admet un point d'inflexion en $M(a)$.

Enfin, pour étudier les signes de $(M(t) - M(a)) \cdot M_m$ et de $(M(t) - M(a)) \cdot M_m^\perp$, il suffit d'examiner les parités de m et n . Plus précisément, $(M(t) - M(a)) \cdot M_m \sim \|M_m\|^2(t-a)^m$, et par conséquent :

- Si m est impair, $(M(t) - M(a)) \cdot M_m$ change de signe : on a un point régulier.
- Si m est pair, $(M(t) - M(a)) \cdot M_m$ change de signe : on a un point de rebroussement.

Par ailleurs, on rappelle que les vecteurs M_k pour $m \leq k < n$ sont tous colinéaires à M_m , et donc

$$(M(t) - M(a)) \cdot M_m^\perp = M_n \cdot M_m^\perp (t-a)^n + o((t-a)^n).$$

Comme M_n et M_m sont linéairement indépendants, $M_n \cdot M_m^\perp \neq 0$ (en effet, si $M_n \cdot M_m^\perp$ était nul, M_n serait orthogonal à M_m^\perp , donc colinéaire à M_m , ce qui est exclu.)

Ainsi $(M(t) - M(a)) \cdot M_m^\perp$ change de signe si et seulement si n est impair. On en déduit donc le résultat de la proposition. □