

**Notes de cours**

**Équations différentielles ordinaires**

**Licence de mathématiques**

**Université d'Aix-Marseille**

Version du 17 novembre 2022

# Cours 1. Équations linéaires du premier ordre

Ce premier cours a pour objet les équations différentielles ordinaires (EDO) linéaires du premier ordre. Les premières notions seront illustrées par des exemples issus de la biologie, de la physique nucléaire ainsi que d'autres applications.

**Définition 1** (Équation différentielle ordinaire du premier ordre). *Une équation différentielle ordinaire (EDO) du premier ordre est une équation qui a pour inconnue une fonction  $y$  d'une variable réelle  $t$ , à valeurs dans  $\mathbb{R}$ , qui s'écrit sous la forme suivante :*

$$y'(t) = f(t, y(t)), \quad t \in I \quad (1.1)$$

où  $I$  est un intervalle ouvert de  $\mathbb{R}$  (par exemple  $I = \mathbb{R}$ ),  $f$  est une fonction continue de  $I \times \mathbb{R}$  dans  $\mathbb{R}$  qui est connue. On cherche alors les fonctions  $y$  de classe  $C^1$  qui vérifient (1.1).

En pratique, on est souvent intéressé par une équation différentielle avec une condition initiale. Le problème obtenu s'appelle alors "problème de Cauchy<sup>1</sup>". Il s'écrit :

$$y'(t) = f(t, y(t)), \quad t \in ]t_0, T[ = I, \quad (1.2a)$$

$$y(t_0) = y_0, \quad (1.2b)$$

où  $t_0, T, f$  et  $y_0$  sont donnés,  $t_0 < T \leq +\infty$ ,  $f$  est continue de  $]t_0, T[ \times \mathbb{R}$  à valeurs dans  $\mathbb{R}$ ,  $y_0 \in \mathbb{R}$ . On cherche alors  $y \in C^1([t_0, T[, \mathbb{R})$  et de classe  $C^1$  sur  $]t_0, T[$  qui vérifie le problème (1.2)<sup>2</sup>. Se posent alors les questions naturelles suivantes.

- 1) Existence d'une solution : locale (c'est-à-dire sur  $]t_0, t_0 + \alpha[$  avec un nombre  $\alpha > 0$  tel que  $t_0 + \alpha < T$ ) ou globale (c'est-à-dire sur  $]t_0, T[$ ) ?
- 2) Unicité de la solution ?
- 3) Stabilité de la solution (en un ou plusieurs sens à déterminer) ?
- 4) Comportement de la solution globale quand  $t \rightarrow T$  ?
- 5) Calcul de la solution ?

On s'intéresse ici au cas linéaire : on choisit  $f(t, y(t)) = a(t)y(t) + b(t)$  où les fonctions  $a$  et  $b$  sont des fonctions continues sur un intervalle  $I \subset \mathbb{R}$  dans  $\mathbb{R}$ ; à  $t$  donné, la fonction  $y \mapsto f(t, y)$  est donc

1. Augustin Louis, baron Cauchy, (1789–1857, mathématicien français, membre de l'Académie des sciences et professeur à l'École polytechnique.

2. Une telle fonction est alors aussi dérivable à droite en  $t_0$ , et vérifie l'équation (1.2a) en  $t_0$ .

affine (mais on parle toutefois incorrectement d'équation linéaire). En se limitant, pour simplifier, au cas  $I = \mathbb{R}$ , l'équation (1.1) devient :

$$y'(t) = a(t)y(t) + b(t), \quad t \in \mathbb{R} \quad (1.3)$$

En supposant qu'une solution à cette équation existe (ce qu'on démontrera un peu plus loin), on cherche tout d'abord à trouver l'expression de  $y$  en fonction de  $t$ .

On résout dans un premier temps l'équation homogène associée à l'équation (1.3), c.à.d. l'équation (1.3) avec  $b(t) = 0$  pour tout  $t$ .

## 1.1 Équation linéaire homogène

Soit l'équation homogène associée à (1.3)

$$y'(t) = a(t)y(t), \quad t \in \mathbb{R}. \quad (1.4)$$

La fonction  $a$  est continue de  $\mathbb{R}$  dans  $\mathbb{R}$ . On note  $A$  une primitive de  $a$  et  $C$  un nombre réel. La fonction  $y$  définie par, pour  $t \in \mathbb{R}$ ,

$$y(t) = Ce^{A(t)}, \quad (1.5)$$

est solution de (1.4). En effet, si  $y$  est de la forme (1.5) alors

$$y'(t) = (Ce^{A(t)})' = CA'(t)e^{A(t)} = a(t)y(t),$$

et donc  $y$  vérifie bien (1.4).

On peut maintenant remarquer que toutes les solutions de (1.4) sont de la forme (1.5). En effet, soit  $y$  une solution de (1.4). On définit alors la fonction  $z \in C^1(\mathbb{R}, \mathbb{R})$  par

$$z(t) = y(t)e^{-A(t)}, \quad t \in \mathbb{R},$$

de sorte que, pour tout  $t \in \mathbb{R}$ ,  $y(t) = z(t)e^{A(t)}$ . On a donc, pour tout  $t \in \mathbb{R}$ ,

$$y'(t) = z'(t)e^{A(t)} + a(t)z(t)e^{A(t)} = z'(t)e^{A(t)} + a(t)y(t).$$

Comme  $y$  est solution de (1.4), on a donc  $z'(t) = 0$  pour tout  $t \in \mathbb{R}$ . La fonction  $z$  est donc constante et  $y$  est bien sous la forme (1.5).

On peut noter qu'une démonstration analogue donnerait toutes les solutions de (1.4) si on travaillait sur  $I$  intervalle ouvert de  $\mathbb{R}$  au lieu de  $\mathbb{R}$ . Toutes les solutions seraient bien sous la forme (1.5) (avec  $t \in I$ ).

Le problème de Cauchy associé à l'équation (1.4) sur l'intervalle  $[t_0, +\infty[$ , avec  $a \in C([t_0, +\infty[)$ , s'écrit :

$$y'(t) = a(t)y(t), \quad t \in ]t_0, +\infty[, \quad (1.6a)$$

$$y(t_0) = y_0. \quad (1.6b)$$

La solution de ce problème est obtenue en écrivant que la solution générale  $y$  de (1.6a), définie par (1.5), satisfait (1.6b) :

$$y(t_0) = Ce^{A(t_0)} = y_0 \text{ et donc } C = y_0 e^{-A(t_0)}$$

ce qui donne finalement

$$y(t) = y_0 e^{A(t) - A(t_0)}, \quad (1.7)$$

**Définition 2** (Équation autonome). On dit que l'équation (1.1) est autonome si la fonction  $f$  ne dépend pas de  $t$ .

L'équation (1.4) est autonome si  $a$  ne dépend pas de  $t$ . La solution générale de (1.4) est alors  $y(t) = Ce^{at}$ . Voyons maintenant quelques exemples :

**Exemple 1** (Modèles de population). Le modèle de Malthus<sup>3</sup> pour la croissance d'une population<sup>4</sup> dont le nombre initial est  $y_0$  au temps  $t_0$  s'écrit :

$$y'(t) = ay(t) \quad t \in ]t_0, +\infty[ \quad (1.8)$$

$$y(t_0) = y_0 \quad (1.9)$$

avec  $a > 0$  donné. Le paramètre  $a$  est souvent appelé taux d'accroissement de la population, et son unité est l'inverse de celle du temps (car  $y$  est un nombre, et  $y'$  un nombre par unité de temps). L'unité de temps doit donc être correctement spécifiée lors de la modélisation. La solution de ce problème de Cauchy est donnée par (1.7) avec  $A(t) = at$ , elle s'écrit donc  $y(t) = y_0 e^{a(t-t_0)}$ . Ce modèle a été utilisé pour modéliser l'évolution de la population américaine durant les années 1960-1970; il s'est avéré réaliste jusqu'à l'année 2000. Le paramètre utilisé est  $a = \frac{2}{100}$  (ici l'unité de temps est l'année). Il est intéressant de calculer le temps qu'il faut pour que la population ait doublé, i.e. le temps  $T_2$  tel tel que  $y(t_0 + T_2) = y_0 e^{aT_2} = 2y_0$  soit encore  $e^{aT_2} = 2$ . On obtient  $T_2 = \frac{\ln 2}{a} = 100 \frac{0,7}{2}$  c.à.d. à peu près 35 ans. Ce modèle de croissance exponentielle n'est valide que si les ressources ne sont pas limitées, et en l'absence d'épidémies, guerres etc... Si l'on veut prendre en compte la limitation des ressources, il faut introduire un modèle non linéaire. On peut utiliser par exemple le modèle de Verhulst<sup>5</sup>, dit aussi "modèle logistique", s'écrit, avec  $a > 0$  et  $b > 0$  donnés :

$$\begin{cases} y'(t) = ay(t) - by^2(t), \\ y(t_0) = y_0. \end{cases} \quad (1.10)$$

On montrera en TD qu'en  $+\infty$  la solution du problème (1.10), pour  $y_0 > 0$ , tend vers  $\frac{a}{b}$ .

L'allure des solutions des deux modèles est illustrée sur la figure suivante :

Comme le suggère la figure 1.1, la solution du modèle de Malthus tend vers  $+\infty$  lorsque  $t$  tend vers  $+\infty$ , alors que celle du modèle du Verhulst converge vers une asymptote.

Revenons à l'équation différentielle linéaire du premier ordre (1.3) avec la fonctions  $a$  constante et la fonction  $b$  nulle (c'est-à-dire une équation différentielle linéaire du premier ordre autonome et homogène). On cherche donc les solutions de l'EDO suivante

$$y'(t) = ay(t), \quad t \in \mathbb{R}. \quad (1.11)$$

On a déjà vu que la solution est  $y(t) = Ce^{at}$  (avec  $C \in \mathbb{R}$  quelconque). Si on a oublié la forme de la solution générale de (1.11), on peut la retrouver en supposant, par exemple, que  $y(t) > 0$  pour tout  $t$ ,

3. Thomas Robert Malthus est un économiste britannique qui a vécu entre 1761 et 1834, surtout connu pour ses travaux sur les rapports entre les dynamiques de croissance de la population et la production. Son nom est associé à une doctrine, le malthusianisme, qui inclut une politique active de contrôle de la natalité.

4. Dans ce type de modèle, on modélise l'évolution du nombre (entier) d'individus d'une population, par une équation différentielle, dont la solution est une fonction continue (et même dérivable) à valeurs réelles, ce qui peut paraître surprenant. Toutefois pour une population de grande taille, et avec une échelle de temps appropriée, les variations de la population pourront effectivement être considérées comme continues.

5. Pierre-François Verhulst, 1804-1849, mathématicien belge. Inspiré par l'« Essai sur le principe de population » de Thomas Malthus, il proposa en 1838 le modèle de Verhulst, décrivant l'évolution des populations animales grâce à un modèle qui ne soit pas exponentiel. C'est dans la publication de 1845 qu'il nomme cette courbe « logistique » sans donner l'explication de ce terme.

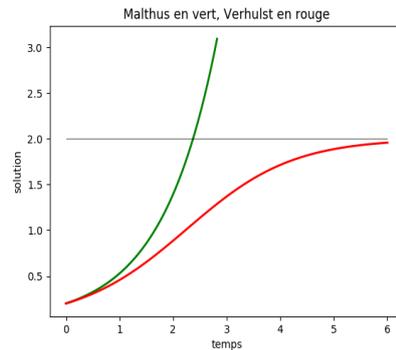


Fig. 1.1: Solution des deux modèles avec  $y_0 = 0.2$ ,  $a = 1$  et  $b = 1/2$

on a alors  $\frac{y'(t)}{y(t)} = a$  pour tout  $t$ , d'où l'on déduit qu'il existe  $C_1 \in \mathbb{R}$  tel que  $\ln(y(t)) = at + C_1$ , et donc  $y(t) = Ce^{at}$  (pour tout  $t$ ), avec  $C = e^{C_1}$ . Remarquons que cette solution vérifie bien l'hypothèse de départ  $y(t) > 0$  si  $C > 0$ . Ce calcul permet de retrouver la forme générale la solution ; on peut vérifier ensuite que toutes les solutions sont de cette forme : reprenons pour cela le raisonnement vu au paragraphe 1.1 lorsque  $a$  n'est pas nécessairement une fonction constante. Soit  $y$  une solution de l'équation (1.11). Posons  $z(t) = y(t)e^{-at}$  ; on a alors

$$\begin{aligned} z'(t) &= y'(t)e^{-at} - y(t)ae^{-at} \\ &= \underbrace{(y'(t) - ay(t))}_{=0} e^{-at} = 0. \end{aligned}$$

On en déduit que  $z$  est constante. Les solutions de (1.11) sont donc toutes de la forme  $y(t) = Ce^{at}$ . On a ainsi démontré le lemme suivant.

**Lemme 1** (Solutions d'une EDO linéaire homogène autonome). *L'ensemble des solutions de l'EDO linéaire homogène autonome (1.11) est un espace vectoriel de dimension 1. C'est l'ensemble des fonctions de la forme  $t \mapsto Ce^{at}$ ,  $C \in \mathbb{R}$ . Cet espace vectoriel est engendré par la fonction  $t \mapsto e^{at}$ .*

Le problème de Cauchy associé à l'équation (1.11) s'écrit

$$y'(t) = ay(t), \quad t \in [t_0, +\infty[, \quad (1.12a)$$

$$y(t_0) = y_0. \quad (1.12b)$$

La condition initiale permet de fixer la constante  $C$  de la solution  $y(t) = Ce^{at}$ , qui s'écrit au temps initial :  $y(t_0) = y_0 = Ce^{at_0}$ . On obtient  $C = y_0 e^{-at_0}$ . La solution du problème de Cauchy (1.12) est donc  $y(t) = y_0 e^{a(t-t_0)}$ . On a ainsi démontré le lemme suivant :

**Lemme 2** (Solution d'une EDO linéaire homogène autonome avec condition initiale). *La solution du problème de Cauchy (1.12) existe, elle est unique, et donnée par  $y(t) = y_0 e^{a(t-t_0)}$ .*

**Exemple 2** (Datation au carbone 14.). *Le carbone 14 est un isotope radioactif du carbone<sup>6</sup> noté  $^{14}_6\text{C}$ . Son évolution suit une loi exponentielle, et le problème de Cauchy associé s'écrit :*

$$y'(t) = -\lambda y(t), \quad t > t_0, \quad (1.13a)$$

$$y(t_0) = y_0. \quad (1.13b)$$

6. Le carbone (C) possède 15 isotopes connus, de nombre de masse variant de 8 à 22, dont deux stables,  $^{12}\text{C}$  et  $^{13}\text{C}$ , le nombre devant  $C$  désignant le nombre de protons.

où  $\lambda > 0$  est connu. Dans le cas de la datation d'un os préhistorique, par exemple, la valeur de  $y_0$  est la concentration<sup>7</sup> en  $^{14}\text{C}$  de l'os au moment de la mort de son propriétaire. La solution du problème (1.13) est  $y(t) = y_0 e^{-\lambda(t-t_0)}$ ; on en déduit que  $e^{-\lambda(t-t_0)} = \frac{y(t)}{y_0}$  et  $-\lambda(t-t_0) = \ln\left(\frac{y(t)}{y_0}\right)$  et donc que

$$t - t_0 = -\frac{1}{\lambda} \ln\left(\frac{y(t)}{y_0}\right) > 0.$$

Dans le cas du modèle de Malthus de croissance d'une population, nous avons calculé plus haut le temps de doublement de la population. Pour la mesure de la décroissance de la radioactivité, une mesure très utilisée en radioactivité est demi-vie<sup>8</sup>, qui est le temps au bout duquel la moitié des noyaux radioactifs d'une source se sont désintégrés<sup>9</sup>. La demi-vie  $T_{\frac{1}{2}}$  du carbone 14 est donc définie par  $e^{-\lambda T_{\frac{1}{2}}} = \frac{1}{2}$ ; on a donc  $T_{\frac{1}{2}} = \frac{\ln 2}{\lambda}$ . Connaissant la demi-vie de  $^{14}\text{C}$  qui est de 5568 ans, on en déduit la valeur de  $\lambda$ ,  $\lambda = \ln 2 / 5568$  c'est-à-dire environ  $1,26 \cdot 10^{-4}$ .

**Lemme 3** (Demi-vie). On considère l'équation (1.11) dont la solution est  $y(t) = Ce^{at}$ ,  $C \in \mathbb{R}$ .

Si  $a < 0$ , cette solution est décroissante et on appelle demi-vie le temps  $T_{\frac{1}{2}}$  tel que  $y(t + T_{\frac{1}{2}}) = \frac{1}{2}y(t)$  pour tout  $t > 0$ .

Si  $a > 0$ , cette solution est croissante et on appelle temps de doublement le temps  $T_2$  tel que  $y(t + T_2) = 2y(t)$  pour tout  $t > 0$ .

Plus généralement, pour  $\alpha > 0$ , le temps  $T_\alpha$  tel que  $y(t + T_\alpha) = \alpha y(t)$  est indépendant du temps  $t$ , et il donné par la formule

$$T_\alpha = \frac{\ln \alpha}{a}. \quad (1.14)$$

*Démonstration.* Soit  $\alpha > 0$ ,  $t > 0$  et  $T_\alpha$  tel que  $y(t + T_\alpha) = \alpha y(t)$ . Comme  $y(t) = Ce^{at}$ , on a donc

$$y(t + T_\alpha) = Ce^{at} e^{aT_\alpha} = \alpha y(t) = \alpha Ce^{at}.$$

On en déduit que  $e^{aT_\alpha} = \alpha$ , ce qui donne (1.14). Le temps  $T_\alpha$  est donc bien indépendant de  $t$ .  $\square$

7. Cette concentration est aussi la concentration en carbone 14 de l'atmosphère à la même époque, en raison des échanges respiratoires des êtres vivants. On admet en général que cette concentration a très peu varié au cours des siècles, cette hypothèse est cruciale pour la datation au carbone 14. Cette hypothèse est en fait contestable. Par exemple, la concentration en  $^{14}\text{C}$  atmosphérique a significativement varié après les essais nucléaires entre les années 1950 et 1970.

8. La valeur de la demi-vie est très utile pour estimation de la durée de vie de l'élément radioactif et donc pour l'évaluation de sa dangerosité. Par exemple, dans le cas d'un accident nucléaire, un grand nombre de radionucléides s'échappent dans l'environnement, parmi lesquels :

- l'uranium 238  $^{238}_{92}\text{U}$  a une demi-vie de  $4,5 \cdot 10^9$  années, mais qui n'est que faiblement radioactif. L'activité d'un échantillon radioactif (et donc sa dangerosité) est inversement proportionnel à sa demi-vie, voir par exemple <https://courses.lumenlearning.com/physics/chapter/31-5-half-life-and-activity/>. La mesure de la demi-vie d'un radionucléide faiblement radioactif tel que l'uranium 238 n'est d'ailleurs pas chose aisée, voir par exemple <https://hps.org/publicinformation/ate/q8270.html>.
- l'iode 131  $^{131}_{53}\text{I}$  qui a une demi-vie de 8 jours. L'iode  $^{131}_{53}\text{I}$  est dangereux car il se fixe sur la thyroïde des personnes exposées. Une protection consiste à administrer pendant quelques jours des comprimés d'iode 127 stable empêchant ainsi l'iode  $^{131}_{53}\text{I}$  de se fixer.

9. Attention, deux demi-vies ne correspondent pas à la vie complète !

## 1.2 Équation linéaire non homogène

Il s'agit ici de l'équation  $y'(t) = a(t)y(t) + b(t)$  avec une fonction  $b$  non nulle. Le problème de Cauchy associé s'écrit

$$\begin{aligned} y(t) &= a(t)y(t) + b(t), & t \in ]t_0, T[, \\ y(t_0) &= y_0. \end{aligned} \tag{1.15}$$

Les fonctions  $a$  et  $b$  sont des fonctions continues de  $]t_0, T[$  dans  $\mathbb{R}$  et on cherche  $y \in C^1(]t_0, T[, \mathbb{R}) \cap C([t_0, T[, \mathbb{R})$  solution de (1.15).

Nous verrons dans la section 1.2.2 qu'il existe une unique solution au problème (1.15) (c'est aussi une conséquence du théorème de Cauchy-Lipschitz<sup>10</sup> que nous verrons dans un cours ultérieur). Nous allons décrire plusieurs méthodes pour trouver la solution générale de l'équation  $y(t) = a(t)y(t) + b(t)$ . Cette solution dépend d'un paramètre dont la valeur est choisie pour satisfaire la condition initiale  $y(t_0) = y_0$ , ce qui permet d'obtenir l'unique solution de (1.15).

### 1.2.1 Recherche d'une solution particulière

Pour la résolution d'une EDO linéaire, la méthode à privilégier est la recherche d'une solution particulière, car elle nécessite le moins de calculs.

On suppose avoir trouvé une solution particulière  $y_p$  de l'équation  $y(t) = a(t)y(t) + b(t)$ . On a donc  $y'_p(t) = a(t)y_p(t) + b(t)$  pour tout  $t$ . Soit  $y$  une autre solution de l'équation  $y(t) = a(t)y(t) + b(t)$ , on pose  $z = y - y_p$  et on a donc, grâce à la linéarité de l'équation,

$$\begin{aligned} z' &= y' - y'_p \\ &= ay + b - ay_p - b \\ &= az. \end{aligned}$$

On a donc  $z' = az$ , ce qui donne  $z(t) = Ce^{A(t)}$ , où  $A$  est une primitive de  $a$  et  $C$  est un réel arbitraire. On en déduit que  $y = y_p + Ce^{A(t)}$ . Autrement dit, la solution générale de l'équation  $y(t) = a(t)y(t) + b(t)$  s'écrit comme la somme d'une solution particulière (notée ici  $y_p$ ) et de la solution générale de l'équation homogène associée (équation  $y'(t) = a(t)y(t)$ ).

Si une solution particulière  $y_p$  est facile à trouver, le problème est alors résolu : la solution générale de l'équation (1.3) est  $y(t) = Ce^{A(t)} + y_p(t)$  où  $A$  est une primitive de  $a$ . Sinon, il faut passer par une des méthodes proposées ci-après (section 1.2.2).

**Exemple 3** (Solution particulière évidente). *Si les fonctions  $a$  et  $b$  sont constantes,  $a(t) = a \in \mathbb{R}^*$  et  $b(t) = b \in \mathbb{R}$  pour tout  $t$ , alors  $y_p(t) = -\frac{b}{a}$  est une solution particulière car  $y'_p = 0$  et  $ay_p + b = 0$ . La solution générale de l'équation  $y' = ay + b$  est donc  $y(t) = Ce^{at} - \frac{b}{a}$ .*

**Remarque 1** (Principe de superposition). *Dans le cas où le second membre est une somme de plusieurs termes, grâce au caractère linéaire de l'équation, on peut calculer les solutions particulières pour chacun des termes et les ajouter à la solution de l'équation homogène. Etudions par exemple le cas  $b(t) = b_1(t) + b_2(t)$ , où  $b_1$  et  $b_2$  sont des fonctions continues d'un intervalle de  $]t_0, T[$  dans  $\mathbb{R}$  (le cas où  $b$  est la somme de  $n$  fonctions,  $n \geq 2$  s'en déduit facilement). Reprenons l'étude du problème (1.15) avec cette fonction  $b$ . Le problème de Cauchy s'écrit donc*

$$y'(t) = a(t)y(t) + b_1(t) + b_2(t), \quad t \in ]t_0, T[, \tag{1.16}$$

$$y(t_0) = y_0. \tag{1.17}$$

10. Rudolph Otto Sigismund Lipschitz (1832–1903), mathématicien allemand.

Montrons qu'en raison de la linéarité de l'équation, la solution de ce problème s'écrit  $y = y_h + y_{p,1} + y_{p,2}$  où  $y_{p,1}$  (resp.  $y_{p,2}$ ) est une solution particulière de l'équation  $y'(t) = a(t)y(t) + b_1(t)$  (resp.  $y'(t) = a(t)y(t) + b_2(t)$ ). En effet, soit  $y_{p,i}$  une solution particulière de l'équation  $y'(t) = ay(t) + b_i(t)$ , pour  $i = 1, 2$ . Si  $y$  est solution de (1.16), on pose  $z = y - y_{p,1} - y_{p,2}$ . La fonction  $z$  est alors solution de l'équation homogène associée (équation  $y'(t) = a(t)y(t)$ ). On a donc  $z(t) = C \exp A(t)$  où  $A$  est une primitive de  $a$  et  $C$  une constante arbitraire. La solution générale de l'équation (1.16) est donc  $y(t) = C \exp A(t) + y_{p,1}(t) + y_{p,2}(t)$  (où  $A$  est une primitive de  $a$  et  $C$  une constante arbitraire). On calcule alors la valeur de  $C$  avec (1.17).

## 1.2.2 Variation de la constante, Facteur intégrant, Réduction d'ordre

Lorsqu'on ne sait pas trouver facilement de solution particulière à l'équation (1.15), on utilise une méthode de calcul à partir de la solution de l'équation homogène. Cette méthode peut être présentée de plusieurs façons, et a donc plusieurs dénominations possibles : méthode de *la variation de la constante*, la méthode du *facteur intégrant* ou la méthode de *réduction d'ordre*. Cette méthode (que l'on a déjà utilisée précédemment lors de l'étude des solutions de l'équation homogène, voir paragraphe 1.1) utilise la forme des solutions de l'équation homogène  $y'(t) = a(t)y(t)$ , c'est-à-dire  $y(t) = Ce^{A(t)}$  (avec  $A$  primitive de  $a$  et  $C \in \mathbb{R}$ ).

Soit  $y$  solution de l'équation différentielle donnée dans (1.15) (c'est-à-dire  $y'(t) = a(t)y(t) + b(t)$ ). On définit la fonction  $z$  de  $[t_0, T[$  dans  $\mathbb{R}$  par

$$z(t) = y(t)e^{-A(t)},$$

ce qui donne

$$y(t) = z(t)e^{A(t)}.$$

La fonction  $t \mapsto e^{-A(t)}$  est le *facteur intégrant* et la fonction  $z$  remplace  $C$  dans l'expression de la solution de l'équation homogène  $y(t) = Ce^{A(t)}$ , d'où l'expression, un peu contradictoire, de *variation de la constante*.

On obtient alors, pour tout  $t \in ]t_0, T[$ ,

$$y'(t) = z(t)a(t)e^{A(t)} + z'(t)e^{A(t)} = a(t)y(t) + z'(t)e^{A(t)}.$$

Comme  $y$  est solution de  $y'(t) = a(t)y(t) + b(t)$ , on a donc nécessairement, pour tout  $t \in ]t_0, T[$ ,

$$z'(t) = b(t)e^{-A(t)}.$$

L'équation sur la fonction  $z'$  est d'ordre 0 car ses dérivées n'interviennent pas dans l'équation, ce qui explique l'expression "réduction d'ordre".)

Comme  $z$  est continue sur  $[t_0, T[$ , on a donc, pour tout  $t \in [t_0, T[$ ,

$$z(t) = z(t_0) + \int_{t_0}^t b(s)e^{-A(s)} ds,$$

c'est-à-dire en utilisant la condition initiale  $y(t_0) = y_0$ ,

$$z(t) = y(t)e^{-A(t)} = y_0e^{-A(t_0)} + \int_{t_0}^t b(s)e^{-A(s)} ds,$$

ou encore, en choisissant pour  $A$  la primitive de  $a$  s'annulant en  $t_0$ , c'est-à-dire  $A(t) = \int_{t_0}^t a(s) ds$ ,

$$y(t) = y_0e^{A(t)} + \int_{t_0}^t b(s)e^{A(t)-A(s)} ds, \text{ pour tout } [t_0, T[. \quad (1.18)$$

On a obtenu ainsi l'unique solution de (1.15). (La méthode donne aussi toutes les solutions de l'EDO  $y'(t) = a(t)y(t) + b(t)$ .)

Dans le cas où la fonction  $a$  est constante (et on note  $a$  sa valeur) et  $t_0 = 0, T = +\infty$ , on obtient, pour tout  $t \geq 0$ ,

$$y(t) = y_0 e^{at} + \int_{t_0}^t b(s) e^{a(t-s)} ds. \quad (1.19)$$

Cette formule s'appelle formule de Duhamel<sup>11</sup>. Nous verrons dans la suite qu'elle se généralise pour les systèmes différentiels (elle se généralise d'ailleurs aussi pour des équations différentielles en dimension infinie).

**Exemple 4.** On cherche à résoudre le problème de Cauchy

$$y' + y = \frac{1}{t^2 + 1} \text{ sur } I = [1, \infty[, \quad (1.20a)$$

$$y(1) = y_0. \quad (1.20b)$$

Ici  $a = -1$ , la primitive s'annulant en 1 de  $a$  est donc  $A(t) = 1 - t$  et la formule (1.18) s'écrit donc

$$y(t) = y_0 e^{1-t} + \int_1^t \frac{e^{s-t}}{s^2 + 1} ds, \quad t \geq 1.$$

### 1.2.3 Application : l'expertise de tableaux d'art

L'analyse de la teneur de certains composants radioactifs contenus dans les peintures utilisées par les artistes permet d'expertiser les tableaux<sup>12</sup>. C'est par exemple ainsi qu'on a pu déterminer que le tableau "La Cène", attribué pendant longtemps à Vermeer<sup>13</sup> était en fait un faux, peint par van Meegeren<sup>14</sup>.

Les scientifiques de trois laboratoires différents ont prouvé que le tableau était récent grâce à sa teneur en plomb 210, radioactif, qui a une demi-vie de 22 ans. Il serait trop long d'expliquer ici le modèle en détails car il s'appuie sur la connaissance des processus de désintégration des matériaux radioactifs et sur l'étude des minerais d'où proviennent les pigments utilisés lors de la fabrication des peintures.



La Cène, du peintre et faussaire Van Meegeren

On renvoie au livre de Braun<sup>15</sup> pour une explication complète du modèle. Toujours est-il qu'après simplification, le problème de Cauchy à étudier s'écrit de la manière suivante :

$$y'(t) = -\lambda y(t) + b, \quad t \in [t_0, +\infty[, \quad (1.21a)$$

$$y(t_0) = y_0, \quad (1.21b)$$

11. Jean-Marie Duhamel (1797-1872), mathématicien et physicien français. Travaux sur les équations aux dérivées partielles, avec des applications en acoustique et propagation de la chaleur

12. Voir aussi l'expertise du faux "saint suaire" [http://www.unice.fr/zetetique/articles/HB\\_suaire\\_C14.html](http://www.unice.fr/zetetique/articles/HB_suaire_C14.html), qui continue à faire couler beaucoup d'encre.

13. Johannes ou Jan Van der Meer, dit Vermeer, peintre baroque néerlandais, 1632-1675.

14. Henricus Antonius van Meegeren, peintre, restaurateur d'œuvres d'art et faussaire hollandais, 1889-1947

15. Braun, Differential equations, Texts in Applied Mathematics book series (TAM, volume 11), Springer Verlag, 1994

avec  $b \in \mathbb{R}_+$ . Soit  $y$  une solution de (1.21); en posant  $z(t) = y(t) - \frac{b}{\lambda}$ , on a immédiatement que  $z$  est solution de

$$\begin{aligned} z'(t) &= -\lambda z(t) \\ z(t_0) &= z_0 = y_0 - \frac{b}{\lambda}, \end{aligned}$$

dont la solution est  $z(t) = z_0 e^{-\lambda(t-t_0)}$ . On en déduit que

$$y(t) = z(t) + \frac{b}{\lambda} = \left(y_0 - \frac{b}{\lambda}\right) e^{-\lambda(t-t_0)} + \frac{b}{\lambda}.$$

A partir de la résolution de cette équation, et de la modélisation décrite dans le livre de Braun cité plus haut, il est alors possible de déterminer si un tableau a été peint il y a 30 ans ou il y a 300 ans.

### 1.3 Stabilité par rapport à la donnée initiale

Le terme “stabilité” regroupe un ensemble de notions assez différentes. Ici on se pose la question de la stabilité par rapport à la condition initiale (ou donnée initiale, les termes sont synonymes) : on ajoute une perturbation à la condition initiale, c’est-à-dire un petit terme  $\varepsilon$ , et on se demande quel est le comportement de la solution avec cette nouvelle donnée initiale lorsque  $\varepsilon$  tend vers 0. La solution sera dite stable par rapport à la donnée initiale, si elle tend (en un sens à définir) vers la solution du problème sans perturbation.

**Définition 3** (Stabilité par rapport à la donnée initiale). *On considère le problème de Cauchy (1.2), et on suppose qu’il admet une unique solution  $y$ . Soit  $y_\varepsilon$  la solution du problème de Cauchy où la condition initiale est remplacée par*

$$y(0) = y_0 + \varepsilon.$$

*On dit que la solution  $y$  est stable à l’instant  $t$  par rapport à la donnée initiale si*

$$\lim_{\varepsilon \rightarrow 0} (y_\varepsilon(t) - y(t)) = 0.$$

*On dit que la solution  $y$  est uniformément stable (on dira souvent simplement “stable”) par rapport à la donnée initiale si*

$$\lim_{\varepsilon \rightarrow 0} \sup_{t > t_0} |y_\varepsilon(t) - y(t)| = 0.$$

Etudions ce comportement dans le cas d’une EDO linéaire, homogène et à coefficient constant  $a \in \mathbb{R}$ . Soit  $t_0 \in \mathbb{R}$ , le problème de Cauchy s’écrit alors

$$\begin{cases} y'(t) = ay(t), & t \in [t_0, +\infty[ \\ y(t_0) = y_0. \end{cases} \quad (1.22)$$

Le problème de Cauchy avec donnée initiale perturbée s’écrit, pour  $\varepsilon \neq 0$ ,

$$\begin{cases} y'(t) = ay(t), & t \in [t_0, +\infty[ \\ y(t_0) = y_0 + \varepsilon. \end{cases} \quad (1.23)$$

Les solutions respectives des problèmes (1.22) et (1.23) sont

$$y(t) = y_0 e^{a(t-t_0)} \text{ et } y_\varepsilon(t) = (y_0 + \varepsilon) e^{a(t-t_0)} = y_0 e^{a(t-t_0)} + \varepsilon e^{a(t-t_0)}.$$

On a donc

$$y_\varepsilon(t) - y(t) = \varepsilon e^{a(t-t_0)}.$$

La solution est donc stable à l'instant  $t$  par rapport à la donnée initiale car

$$\forall t \in [t_0, +\infty[, \quad \lim_{\varepsilon \rightarrow 0} (y_\varepsilon(t) - y(t)) = 0. \quad (1.24)$$

Si  $a > 0$ , la solution n'est pas uniformément stable car

$$\forall \varepsilon \neq 0, \quad \sup_{t > t_0} |y_\varepsilon(t) - y(t)| = +\infty.$$

Par contre, si  $a \leq 0$ , la solution est uniformément stable car on a bien

$$\lim_{\varepsilon \rightarrow 0} \sup_{t > t_0} |y_\varepsilon(t) - y(t)| = 0.$$

# Cours 2. Équations non linéaires du premier ordre

On considère le problème de Cauchy suivant, avec  $T \in ]0, +\infty]$  (noter que  $T$  peut donc être fini ou infini),  $y_0 \in \mathbb{R}$  et  $f \in C([0, T[ \times \mathbb{R}, \mathbb{R})$  donnés :

$$y'(t) = f(t, y(t)), \quad t \in ]0, T[, \quad (2.1a)$$

$$y(0) = y_0. \quad (2.1b)$$

On cherche alors  $y \in \mathcal{C}([0, T[) \cap \mathcal{C}^1(]0, T[)$  vérifiant (2.1). (La relation (2.1a) est alors aussi vraie en  $t = 0$  en considérant seulement la dérivée à droite de  $y$  en 0.)

Les questions que nous allons aborder ici sont fondamentales :

- 1) existence locale, c.à.d. existe-t-il  $\alpha > 0$  (avec  $\alpha \leq T$ ) et une fonction  $y : [0, \alpha[ \rightarrow \mathbb{R}$  telle que  $y$  satisfasse (2.1) avec  $\alpha$  au lieu de  $T$  dans (2.1a) ?
- 2) existence globale, c.à.d. peut-on prendre  $\alpha = T$  dans l'item précédent ?
- 3) unicité ?

## 2.1 Exemples d'existence locale, globale et de non unicité

Commençons par des exemples pour illustrer ces notions.

**Exemple 5** (Un cas d'existence locale et non globale). *On considère le problème de Cauchy (2.1) avec  $f(t, x) = x^2$  et  $T = +\infty$ , c.à.d.*

$$y'(t) = y^2(t), \quad t > 0, \quad (2.2a)$$

$$y(0) = 1. \quad (2.2b)$$

Si une solution  $y$  à ce problème existe, on aura alors  $y(t) > 0$  pour tout  $t > 0$  (car  $y(0) > 0$  et  $y'(t) = y^2(t) \geq 0$  pour tout  $t > 0$  et donc  $y(t) \geq y(0) > 0$  pour tout  $t \geq 0$ ). On a donc  $\frac{y'}{y^2} = 1$ , ce

qui s'écrit aussi  $\left(-\frac{1}{y}\right)' = 1$ . Il existe donc  $C \in \mathbb{R}$  tel que  $-\frac{1}{y(t)} = t + C$ . La condition initiale (2.2b) donne  $C = -1$  et donc

$$y(t) = \frac{1}{1-t}, \quad (2.3)$$

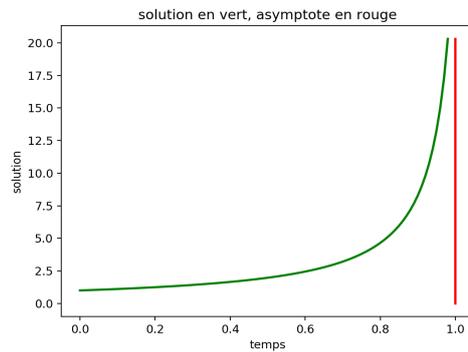


Fig. 2.1: La fonction  $t \mapsto \frac{1}{1-t}$ ; asymptote en  $t = 1$ .

dont le graphe est donnée sur la figure 2.1. Or cette fonction n'est pas définie en 1. Ceci montre que  $y$  existe sur  $[0, 1[$  et qu'il n'y a pas de solution sur  $[0, \alpha[$  si  $\alpha > 1$ ; en effet, une solution éventuelle sur  $[0, \alpha[$  doit être continue sur  $[0, \alpha[$  et on vient de montrer que ceci est impossible si  $\alpha > 1$  car la solution est nécessairement donnée par (2.3) pour  $0 \leq t < 1$ . Il y a existence de  $y$  sur  $]0, 1[$  et donc existence locale d'une solution au problème (2.2), mais il n'y a pas d'existence globale.

**Remarque 2** (Existence globale pour une autre condition initiale). Si on remplace la condition initiale (2.2b) dans le problème 2.2 de l'exemple 5 par  $y_0 = -1$ , la solution devient  $y(t) = -\frac{1}{1+t}$  qui est définie sur tout  $\mathbb{R}_+$ , comme le montre la figure 2.2. Il y a donc existence globale dans ce cas.

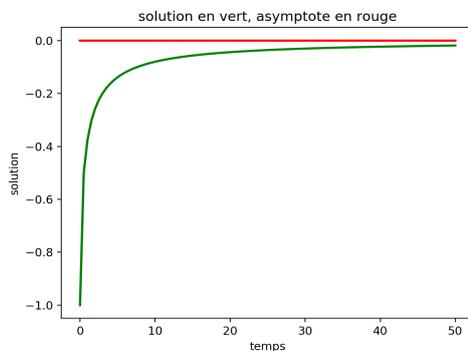


Fig. 2.2: Existence globale pour  $y_0 = -1$

**Remarque 3** (Méthode des variables séparables). Dans certains cas, on peut résoudre les équations différentielles en utilisant la méthode dite des variables séparables. Cette méthode consiste à mettre l'équation sous la forme  $h(y)y'(t) = g(t)$ . En prenant une primitive de  $h$ , notée  $H$ , cette équation est équivalente à  $(H(y))'(t) = g(t)$ . En notant  $G$  une primitive de  $g$ , ceci donne l'existence de  $C \in \mathbb{R}$  tel que  $H(y(t)) = G(t) + C$  pour tout  $t$ . C'est ainsi que nous avons raisonné pour trouver une solution de (2.2a).

**Exemple 6** (Un cas de non unicité). On considère maintenant le problème de Cauchy (2.1) avec  $f(t, x) = \sqrt{x}$ ,  $y_0 = 0$  et  $T = +\infty$ , c.à.d.

$$y'(t) = \sqrt{y(t)}, \quad t > 0, \quad (2.4a)$$

$$y(0) = 0. \quad (2.4b)$$

L'équation (2.4a) n'a un sens que si  $y(t) \geq 0$  pour tout  $t$  (on rappelle qu'on cherche des solutions à valeurs réelles). Cette équation admet alors plusieurs solutions dont l'une solution dite "triviale" est :

$$y_1(t) = 0, \quad \forall t \geq 0,$$

une autre solution se calcule facilement par la méthode des variables séparables :

$$y_2(t) = \frac{t^2}{4}, \quad \forall t \geq 0.$$

En fait on peut construire une infinité de solutions du problème de Cauchy 2.4. En effet, pour tout  $t_0 > 0$ , la fonction définie par

$$y(t) = 0 \text{ si } 0 \leq t \leq t_0,$$

$$y(t) = \frac{(t - t_0)^2}{4} \text{ si } t > t_0,$$

est solution de 2.4. Ceci est illustré dans la figure 2.3.

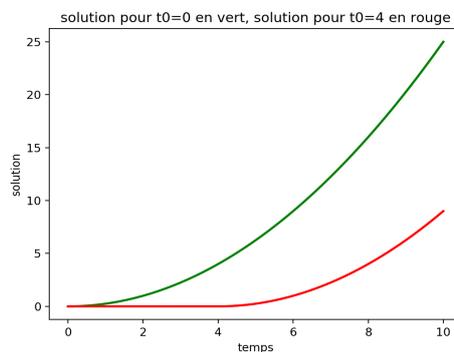


Fig. 2.3: Infinité de solutions pour le problème 2.4

## 2.2 Résultats d'existence et d'unicité

Des résultats d'existence et d'unicité pour le problème (2.1) s'obtiennent grâce au caractère localement lipschitzien de la fonction  $f$  qui apparaît dans l'EDO (2.1a). Rappelons d'abord la notion de fonction lipschitzienne.

**Définition 4** (Fonction lipschitzienne). On dit qu'une fonction  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  est lipschitzienne s'il existe  $C \in \mathbb{R}$ , tel que :

$$|f(x) - f(y)| \leq C|x - y|, \quad \forall x, y \in \mathbb{R}.$$

Notons qu'une fonction  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$ , de classe  $C^1$  et dont la dérivée est bornée, est une fonction lipschitzienne. En effet, en posant  $C = \sup_{s \in \mathbb{R}} |f'(s)|$ , on a  $C < +\infty$  et, pour tout  $(x, y) \in \mathbb{R}^2$ ,  $y \leq x$ ,

$$|f(x) - f(y)| \leq \int_y^x |f'(s)| ds \leq C|x - y|.$$

**Définition 5** (Fonction localement lipschitzienne). Soient  $f \in \mathcal{C}([0, T[ \times \mathbb{R}, \mathbb{R})$  et  $0 < T \leq +\infty$ . On dit que  $f$  est localement lipschitzienne par rapport à son deuxième argument si :

$\forall \alpha \in ]0, T[, \forall R > 0, \exists C_{\alpha, R} > 0 ; |f(t, x) - f(t, y)| \leq C_{\alpha, R}|x - y|, \forall t \in [0, \alpha], \forall (x, y) \in \mathcal{B}_R^2$ ,  
où  $\mathcal{B}_R = \{z \in \mathbb{R}, |z| \leq R\} = [-R, R]$ .

**Proposition 1** (Condition suffisante pour le caractère localement lipschitzien).

Soient  $f \in \mathcal{C}([0, T[ \times \mathbb{R}, \mathbb{R})$  et  $0 < T \leq +\infty$ . On suppose que  $f$  est dérivable par rapport à son deuxième argument et que

$$\frac{\partial f}{\partial y} \in \mathcal{C}([0, T[ \times \mathbb{R}, \mathbb{R})$$

où  $\frac{\partial f}{\partial y}$  désigne la dérivée de  $f$  par rapport à son deuxième argument. Alors,  $f$  est localement lipschitzienne par rapport à son deuxième argument.

*Démonstration.* Pour  $t \in [0, T[$  et  $x, y \in \mathbb{R}$ ,  $x < y$ ,

$$f(t, y) - f(t, x) = \int_x^y \frac{\partial f}{\partial y}(t, z) dz,$$

et donc

$$|f(t, y) - f(t, x)| \leq \int_x^y \left| \frac{\partial f}{\partial y}(t, z) \right| dz. \quad (2.5)$$

Pour  $0 < \alpha < T$  et  $R > 0$ ,  $\frac{\partial f}{\partial y} \in \mathcal{C}([0, \alpha] \times \mathcal{B}_R)$ , on a donc

$$\sup_{\substack{t \in [0, \alpha] \\ y \in \mathcal{B}_R}} \left| \frac{\partial f}{\partial y}(t, y) \right| = C_{\alpha, R} < +\infty.$$

On a utilisé ici le fait qu'une fonction continue sur un compact est bornée.

L'inégalité (2.5) donne alors, pour tout  $t \in [0, \alpha]$  et tout  $x, y \in [-R, R]$ ,

$$|f(t, y) - f(t, x)| \leq C_{\alpha, R}|x - y|.$$

Ceci prouve bien que  $f$  est localement lipschitzienne par rapport à son deuxième argument.  $\square$

**Théorème 1** (Existence locale : Cauchy-Lipschitz). Soient  $f \in \mathcal{C}([0, T[ \times \mathbb{R}, \mathbb{R})$ ,  $0 < T \leq +\infty$  et  $y_0 \in \mathbb{R}$ . On suppose que  $f$  est localement lipschitzienne par rapport à son deuxième argument (voir Définition 5). Alors il existe  $\alpha$ ,  $0 < \alpha \leq T$  et  $y \in \mathcal{C}^1([0, \alpha[, \mathbb{R})$  solution du problème de Cauchy (2.1). Il y a donc existence locale de la solution de ce problème.

La démonstration de ce théorème fondamental est assez longue et fera l'objet du cours suivant dans son entier.

On s'intéresse maintenant au problème de l'unicité. On commence par un lemme très utile pour la suite.

**Lemme 4** (Gronwall). Soient  $\alpha > 0$ ,  $\psi \in C([0, \alpha], \mathbb{R})$  et  $a, b \geq 0$ . On suppose que, pour tout  $t \in [0, \alpha]$ ,

$$0 \leq \psi(t) \leq a + b \int_0^t \psi(s) ds. \quad (2.6)$$

On a alors, pour tout  $t \in [0, \alpha]$ ,  $\psi(t) \leq ae^{bt}$ .

*Démonstration.* Pour  $t \in [0, \alpha]$ , on pose

$$\varphi(t) = a + b \int_0^t \psi(s) ds,$$

de sorte que  $\varphi(0) = a$ ,  $\varphi$  est dérivable et  $\varphi'(t) = b\psi(t)$  pour tout  $0 < t < \alpha$ .

D'après (2.6),  $\psi \leq \varphi$  et donc  $\varphi'(t) \leq b\varphi(t)$  (pour tout  $t$ ). En utilisant la méthode du facteur intégrant,

$$(\varphi(t)e^{-bt})' = \varphi'(t)e^{-bt} - b\varphi(t)e^{-bt} = (\varphi'(t) - b\varphi(t))e^{-bt} \leq 0.$$

La fonction  $t \mapsto \varphi(t)e^{-bt}$  est donc décroissante, et donc

$$\forall t \in [0, \alpha] \quad \varphi(t)e^{-bt} \leq \varphi(0)e^0 = \varphi(0) = a.$$

On en déduit  $0 \leq \psi(t) \leq \varphi(t) \leq ae^{bt}$ , pour tout  $t \in [0, \alpha]$ . □

**Théorème 2** (Unicité). Soient  $f \in C([0, T] \times \mathbb{R}, \mathbb{R})$ ,  $0 < T \leq +\infty$ , et  $y_0 \in \mathbb{R}$ . On suppose que  $f$  est localement lipschitzienne par rapport à son deuxième argument (voir définition 5). Soient  $y, z$  deux solutions du problème de Cauchy (2.1), définies et continues sur  $[0, \alpha]$  avec  $0 < \alpha < T$ , alors  $z(t) = y(t)$  pour tout  $t \in [0, \alpha]$ .

*Démonstration.* Les fonctions  $y$  et  $z$  étant continues sur l'intervalle compact  $[0, \alpha]$ , elles sont bornées sur cet intervalle. Il existe donc  $R > 0$  tel que

$$\forall t \in [0, \alpha], \quad |y(t)| \leq R \quad \text{et} \quad |z(t)| \leq R.$$

Comme  $y'(t) = f(t, y(t))$  et  $z'(t) = f(t, z(t))$ , on a

$$y'(t) - z'(t) = f(t, y(t)) - f(t, z(t)).$$

En intégrant cette équation entre 0 et  $t$  et en remarquant que  $y(0) = z(0) = y_0$ , on a donc

$$y(t) - z(t) = \int_0^t (y'(s) - z'(s)) ds = \int_0^t (f(s, y(s)) - f(s, z(s))) ds, \quad \forall t \in [0, \alpha].$$

En prenant la valeur absolue, pour tout  $t \in [0, \alpha]$ ,

$$\begin{aligned} |y(t) - z(t)| &\leq \int_0^t |f(s, y(s)) - f(s, z(s))| ds \\ &\leq \int_0^t C_{\alpha, R} |y(s) - z(s)| ds, \end{aligned}$$

où  $C_{\alpha,R}$  est la constante de Lipschitz de  $f$ , voir définition 5. On a donc :

$$\forall t \in [0, \alpha] \quad |y(t) - z(t)| \leq \int_0^t C |y(s) - z(s)| ds \quad (2.7)$$

Pour  $t \in [0, \alpha]$ , on pose  $\psi(t) = |y(t) - z(t)|$ , de sorte que (2.7) donne, pour tout  $t \in [0, \alpha]$ ,

$$0 \leq \psi(t) \leq C \int_0^t \psi(s) ds.$$

Le lemme 4 de Gronwall avec  $a = 0$  et  $b = C$  donne alors, pour tout  $t \in [0, \alpha]$ ,  $0 \leq \psi(t) \leq 0 e^{Ct} = 0$  et donc  $\psi(t) = |y(s) - z(s)| = 0$  pour tout  $t \in [0, \alpha]$ , ce qui termine la preuve du théorème 2.  $\square$

On généralise dans le théorème 3 le résultat d'unicité (théorème 2) en montrant que les trajectoires des solutions de (2.1a) ne se rencontrent jamais. Par trajectoire de la solution  $y$ , on entend ici le graphe de la fonction  $t \mapsto y(t)$ .

**Théorème 3** (Les trajectoires ne se rencontrent pas). Soient  $f \in \mathcal{C}([0, T[ \times \mathbb{R}, \mathbb{R}), 0 < T \leq +\infty$ . On suppose que  $f$  est localement lipschitzienne par rapport à son deuxième argument (voir définition 5). Soient  $y, z$  deux solutions de (2.1a), définies et continues sur  $[0, \alpha]$  avec  $0 < \alpha < T$ . On suppose qu'il existe  $t_0 \in [0, \alpha]$  tel que  $y(t_0) = z(t_0)$ , alors  $z(t) = y(t)$  pour tout  $t \in [0, \alpha]$ .

*Démonstration.* Comme dans la preuve du théorème 2, il existe  $R > 0$  tel que

$$\forall t \in [0, \alpha], \quad |y(t)| \leq R \quad \text{et} \quad |z(t)| \leq R.$$

et donc, pour tout  $s \in [0, \alpha]$ ,

$$|f(s, y(s)) - f(s, z(s))| \leq C |y(s) - z(s)|,$$

où  $C = C_{\alpha,R}$  est défini dans la définition 5.

On va montrer avec le lemme 4 que  $y = z$  sur  $[0, \alpha]$  en distinguant les intervalles  $[0, t_0[$  et  $]t_0, \alpha]$ .

**Intervalle**  $]t_0, \alpha]$

Pour  $0 < t < \alpha - t_0$  on a (comme  $y(t_0) = z(t_0)$ )

$$y(t_0+t) - z(t_0+t) = \int_0^t (y'(t_0+s) - z'(t_0+s)) ds = \int_0^t (f(t_0+s, y(t_0+s)) - f(t_0+s, z(t_0+s))) ds,$$

et donc

$$|y(t_0+t) - z(t_0+t)| \leq \int_0^t |f(t_0+s, y(t_0+s)) - f(t_0+s, z(t_0+s))| ds \leq C \int_0^t |y(t_0+s) - z(t_0+s)| ds.$$

En posant  $\psi(t) = |y(t_0+t) - z(t_0+t)|$ , le lemme de Gronwall donne  $\psi(t) = 0$  pour tout  $t \in [0, \alpha - t_0]$ , c'est-à-dire  $y = z$  sur  $]t_0, \alpha]$ .

**Intervalle**  $[0, t_0[$

Pour  $0 < t < t_0$  on a (comme  $y(t_0) = z(t_0)$ )

$$y(t_0-t) - z(t_0-t) = \int_0^t (-y'(t_0-s) + z'(t_0-s)) ds = - \int_0^t (f(t_0-s, y(t_0-s)) - f(t_0-s, z(t_0-s))) ds,$$

et donc

$$|y(t_0-t) - z(t_0-t)| \leq \int_0^t |f(t_0-s, y(t_0-s)) - f(t_0-s, z(t_0-s))| ds \leq C \int_0^t |y(t_0-s) - z(t_0-s)| ds.$$

En posant  $\psi(t) = |y(t_0-t) - z(t_0-t)|$ , le lemme de Gronwall donne  $\psi(t) = 0$  pour tout  $t \in [0, t_0]$ , c'est-à-dire  $y = z$  sur  $[0, t_0]$ .

□

Le théorème 3 permet d'effectuer une *étude qualitative* des solutions du problème de Cauchy (2.1), grâce aux notions de *solution d'équilibre*, de *temps maximal d'existence* et d'*explosion en temps fini*, que nous allons maintenant expliquer et illustrer par le modèle de Verhulst (1.10) vu au chapitre précédent, en choisissant comme paramètres  $a = 1$  et  $b = 1$ , et qui s'écrit dans ce cas :

$$y'(t) = y(t)(1 - y(t)), t \in ]0, +\infty[ \quad (2.8a)$$

$$y(t_0) = y_0 \in ]0, 1[. \quad (2.8b)$$

Il y a deux solutions d'équilibre à l'équation (2.8a) qui sont  $y_1 \equiv 0$  et  $y_2 \equiv 1$ , au sens de la définition suivante.

**Définition 6** (Solution d'équilibre). *On appelle solution d'équilibre, ou solution stationnaire, de l'EDO (2.1a) toute fonction  $y$  vérifiant (2.1a) et indépendante de temps. Une solution stationnaire est donc telle que  $y'(t) = 0$  pour tout  $t \in ]0, T[$  ce qui est équivalent à dire que  $f(t, y(t)) = 0$  pour tout  $t \in ]0, T[$ .*

Les fonctions constantes  $y_1 = 1$  et  $y_2 = 0$  sont donc bien des solutions stationnaires de l'équation (2.8a), et donc aussi solutions du problème de Cauchy (2.8) avec données initiales respectives  $y_0 = 1$  et  $y_0 = 0$ . Par le théorème 3 (qui généralise le théorème d'unicité), on en déduit que la solution de  $y'(t) = y(t)(1 - y(t))$ , avec  $y(0) = y_0 \in ]0, 1[$  prend toujours ses valeurs entre 0 et 1 strictement, puisque les trajectoires des solutions ne peuvent pas se croiser.

C'est une propriété très importante, en particulier en raison des résultats sur le temps maximal  $T_m$  d'existence de la solution et sur le comportement de cette dernière lorsque  $T_m < +\infty$ ; ces résultats sont énoncés dans le théorème suivant :

**Théorème 4** (Solution maximale, explosion en temps fini). *Soit  $0 < T \leq +\infty$ ,  $y_0 \in \mathbb{R}$  et  $f \in \mathcal{C}([0, T[ \times \mathbb{R}, \mathbb{R})$ . On suppose que  $f$  est localement lipschitzienne par rapport à son deuxième argument et on considère le problème de Cauchy (2.1). Soit  $T_m = \sup\{\alpha \in ]0, T[ \mid \text{il existe } y \text{ solution de 2.1 sur } [0, \alpha[ \}$ . Alors, il existe  $y$  solution de (2.1) sur  $[0, T_m[$ , et on dit que  $T_m$  est le temps maximal d'existence. Cette solution  $y$  est l'unique solution sur  $[0, T_m[$ , elle s'appelle solution maximale du problème (2.1).*

*De plus, si  $T_m < T$ , alors  $\lim_{t \rightarrow T_m} |y(t)| = +\infty$ .*

*Démonstration.* Remarquons d'abord que théorème 1 d'existence locale montre que  $T_m$  est bien défini et que  $0 < T_m \leq T$  : en effet, par ce théorème, l'ensemble des réels  $\alpha > 0$  tels qu'il existe  $y$  solution de (2.1) sur  $[0, \alpha[$  est non vide et admet donc une borne supérieure.

Par la définition de  $T_m$ , il existe une suite  $(\alpha_n)_{n \in \mathbb{N}} \subset ]0, T_m[$  telle que

$$\alpha_{n+1} \geq \alpha_n, \text{ pour tout } n \in \mathbb{N}, \text{ et } \lim_{n \rightarrow +\infty} \alpha_n = T_m.$$

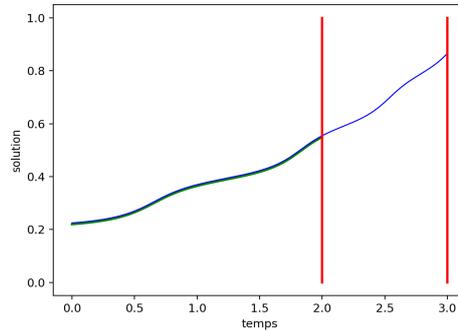


Fig. 2.4: Solution (vert) avec  $\alpha_n = 2$  et (bleu) avec  $\alpha_{n+1} = 3$

Soit  $y_n$  l'unique solution de 2.1 sur  $[0, \alpha_n[$ .

Par le théorème d'unicité (théorème 2),  $y_n = y_{n+1}$  sur  $[0, \alpha_n[$ . Ceci permet de définir  $y$  sur  $[0, T_m[$  en posant, pour tout  $n$ ,  $y = y_n$  sur l'intervalle  $[0, \alpha_n[$ . La fonction  $y$  est ainsi solution sur  $[0, T_m[$ . Ainsi,

- Si  $T_m < T$ , il n'y a pas de solution sur  $[0, \beta[$  dès que  $\beta > T_m$ .
- Si  $T_m = T$ , alors  $y$  est solution globale.

La preuve de l'explosion en temps fini, *i.e.* du fait que  $T_m < T$  implique  $\lim_{t \rightarrow T_m} y(t) = +\infty$  est effectuée dans le chapitre suivant.  $\square$

Le lemme suivant permet de montrer que si la solution d'une équation différentielle autonome admet une limite lorsque  $t \rightarrow +\infty$ , la fonction constante et égale à cette limite est une solution stationnaire. Sa démonstration peut se faire avec le théorème des accroissements finis.

**Lemme 5** (Limite des solutions globales). *Soient  $y_0, t_0 \in \mathbb{R}$  et  $f \in C(\mathbb{R}, \mathbb{R})$ ; on considère le problème de Cauchy*

$$y'(t) = f(y(t)), \quad t > t_0, \quad (2.9)$$

$$y(t_0) = y_0. \quad (2.10)$$

*On suppose qu'il existe une solution globale  $y$  de (2.9)-(2.10), et que  $\lim_{t \rightarrow +\infty} y(t) = \ell \in \mathbb{R}$ . Alors la fonction  $t \mapsto \ell$  est une solution stationnaire (c'est-à-dire que  $\ell$  est un point d'équilibre).*

*Démonstration.* En utilisant le théorème des accroissements finis et l'équation (2.9), on obtient que, pour tout  $t \geq t_0$ , il existe  $\theta_t \in ]t, t+1[$  tel que

$$y(t+1) - y(t) = y'(\theta_t) = f(y(\theta_t)).$$

Quand  $t \rightarrow +\infty$ , on en déduit que  $f(\ell) = 0$  (car  $\lim_{t \rightarrow +\infty} y(\theta_t) = \ell$  et  $f$  est continue).  $\square$

Le résultat suivant permet, grâce à une comparaison, de déterminer le caractère "explosif" d'une solution.

**Lemme 6** (Explosion par comparaison). *Soient  $y_0, t_0 \in \mathbb{R}$  et  $f \in C(\mathbb{R}, \mathbb{R})$ ; on considère le problème de Cauchy (2.10)-(2.9). On suppose qu'il existe une solution de (2.9)-(2.10) sur  $]t_0, T[$  avec  $y(t) \geq y_0 > 0$  pour tout  $t \in [t_0, T[$  et  $f(y(t)) \geq y(t)^2$ . Alors  $T < +\infty$ .*

*Démonstration.* Si  $f(y(t)) \geq y(t)^2$ , alors comme  $y(t) > 0$  pour tout  $t \in [t_0, T[$ ,

$$\frac{y'(t)}{y^2(t)} \geq 1 \text{ pour tout } t \in [t_0, T[$$

et donc en intégrant entre  $t_0$  et  $t < T$ ,

$$\frac{1}{y_0} - \frac{1}{y(t)} \geq t - t_0 \text{ pour tout } t \in [t_0, T[.$$

Donc  $\frac{1}{y(t)} \leq \frac{1}{y_0} - t + t_0$ , et comme  $y(t) > 0$ , on en déduit que  $t \leq t_0 + \frac{1}{y_0}$  pour tout  $t \in [t_0, T[$ , ce qui montre que  $T < +\infty$ ; il n'y a pas existence globale.  $\square$

**Exemple 7** (Comportement qualitatif des solutions du modèle de Verhulst). *Reprenons le problème de Cauchy (2.8). Comme on a montré que la solution du problème reste toujours entre 0 et 1, elle ne peut pas tendre vers l'infini et donc, grâce au théorème 4,  $T_m = +\infty$ . La solution existe sur tout l'intervalle  $[0, +\infty[$ .*

*Une fois qu'on a prouvé l'existence, il est alors assez facile de trouver la limite en  $+\infty$  de cette solution en appliquant le lemme 5. En effet, la fonction  $y$  est croissante et prend donc ses valeurs entre  $y_0$  et 1. Il existe donc  $\ell \in ]0, 1]$  tel que  $\lim_{t \rightarrow +\infty} y(t) = \ell$  où  $0 < \ell \leq 1$ , et  $\ell$  est donc un point d'équilibre. Comme  $\ell \neq 0$ , on a donc forcément  $\ell = 1$ . En fait, pour cet exemple, il est même possible de calculer la solution exacte. Ceci sera fait en TD. La figure 2.5 donne cette solution pour deux valeurs différentes de  $y_0$ .*

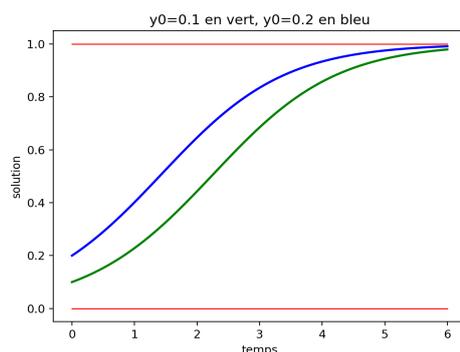


Fig. 2.5: Solution (vert) pour  $y_0 = 0.1$  et (bleu) pour  $y_0 = 0.2$

Sous les hypothèses du théorème 4, il est intéressant de noter que  $\lim_{t \rightarrow T_m} y(t) = +\infty$  n'implique pas que  $T_m < T$ , comme le montre l'exemple suivant.

**Exemple 8** ( $\lim_{t \rightarrow T_m} y(t) = +\infty \not\Rightarrow T_m < T$ ). *Pour  $T = +\infty$ ,  $t_0 = 0$  et  $y_0 \in \mathbb{R}$ , on considère le problème de Cauchy suivant :*

$$\begin{aligned} y'(t) &= 1, & t > 0, \\ y(0) &= y_0. \end{aligned}$$

*Pour ce problème de Cauchy, on a existence globale, c'est-à-dire  $T_m = T = +\infty$ . La solution est  $y(t) = t + y_0$  pour tout  $t \geq 0$ . Pour cet exemple  $\lim_{t \rightarrow T_m} |y(t)| = +\infty$  et pourtant  $T_m = T$ . La réciproque à la dernière assertion du théorème 4 est donc évidemment fausse.*

Enfin, on termine ce cours par la définition de la stabilité pour les points d'équilibre pour une équation autonome, c'est-à-dire l'équation

$$y'(t) = f(y(t)), \quad t > 0. \quad (2.11)$$

**Définition 7** (Stabilité des points d'équilibre). *Soit  $a \in \mathbb{R}$  tel que  $f(a) = 0$ . La fonction constante et égale à  $a$  est donc une solution de (2.11), le point  $a$  est un point d'équilibre. Ce point d'équilibre est stable (plus précisément uniformément stable, définition 3) si pour toute donnée initiale suffisamment proche de  $a$ , le problème de Cauchy associé à (2.11) a une solution globale et si :*

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tel que } (y \text{ solution de (2.11)}, |y(0) - a| \leq \delta) \Rightarrow \sup_{t > 0} |y(t) - a| \leq \varepsilon.$$

On montrera dans un chapitre ultérieur que, si  $f$  est de classe  $C^1$ ,  $a$  est un point d'équilibre stable de (2.11) si  $f'(a) < 0$ , et que  $a$  est instable (c'est-à-dire non stable) si  $f'(a) > 0$ . Le cas où  $f'(a) = 0$  demande une étude plus précise de la fonction  $f$ .

## Cours 3. Les preuves : Cauchy Lipschitz et explosion en temps fini

Soient  $f \in \mathcal{C}([t_0, T[ \times \mathbb{R}, \mathbb{R})$ ,  $t_0 < T \leq +\infty$  et  $y_0 \in \mathbb{R}$ . On cherche  $y \in \mathcal{C}([t_0, T[, \mathbb{R}) \cap \mathcal{C}^1(]t_0, T[, \mathbb{R})$ , solution du problème Cauchy suivant :

$$\begin{cases} y'(t) = f(t, y(t)), & t \in ]t_0, T[, \\ y(t_0) = y_0. \end{cases} \quad (3.1)$$

Le cours 2 abordait les questions suivantes :

- **Question 1** : A-t-on *existence locale*, c'est-à-dire, existe-t-il  $\alpha > 0$  et  $y$  définie sur  $[t_0, t_0 + \alpha]$  ( $t_0 + \alpha < T$ ) solution de (3.1) avec  $t_0 + \alpha$  au lieu de  $T$  ?
- **Question 2** : A-t-on *existence globale*, c'est-à-dire, existe-t-il  $y$  définie sur  $[t_0, T[$  solution de (3.1).

Nous allons supposer, comme dans le cours précédent, que la fonction  $f$  est localement lipschitzienne par rapport à son deuxième argument, c'est-à-dire que pour tout  $a < T$  et tout  $A > 0$ , il existe  $C_{a,A}$  tel que :

$$\forall t \in [t_0, a], \quad \forall x, y \in [-A, A] = \mathcal{B}_A, \quad |f(t, x) - f(t, y)| \leq C_{a,A} |x - y|. \quad (3.2)$$

Les théorèmes du cours précédent ont été énoncés avec  $t_0 = 0$ , pour simplifier les notations, mais il n'y a aucune difficulté à remplacer dans ces théorèmes  $[0, T[$  par  $[t_0, T[$  avec  $t_0 < T \leq +\infty[$ .

Le théorème de Cauchy-Lipschitz (théorème 1 page 15, énoncé au cours précédent) répond à la question 1. Sa preuve fait l'objet de la section suivante. Nous abordons dans une deuxième partie la preuve de la deuxième partie du théorème d'existence d'une solution maximale, théorème 4 (cette deuxième partie montre "l'explosion en temps fini" lorsque la solution maximale n'est pas globale).

### 3.1 Preuve du théorème 1

#### 3.1.1 Forme intégrale du problème de Cauchy

Pour démontrer le théorème 1, nous allons utiliser une forme intégrale du problème (3.1) que nous donnons dans la proposition suivante.

**Proposition 2** (Forme intégrale du problème de Cauchy). *Soient  $y_0$  et  $f \in \mathcal{C}([t_0, T[ \times \mathbb{R}, \mathbb{R})$ . Soit  $\alpha > 0$ ,  $t_0 + \alpha < T$ . Alors  $y \in \mathcal{C}([t_0, t_0 + \alpha], \mathbb{R}) \cap \mathcal{C}^1(]t_0, t_0 + \alpha[, \mathbb{R})$ , est solution de (3.1) avec  $t_0 + \alpha$  au lieu de  $T$  si et seulement si*

$$\begin{cases} y \in \mathcal{C}([t_0, t_0 + \alpha], \mathbb{R}), \\ y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds, \quad \forall t \in [t_0, t_0 + \alpha]. \end{cases} \quad (3.3)$$

*Démonstration.* On suppose tout d'abord que  $y \in \mathcal{C}([t_0, t_0 + \alpha], \mathbb{R}) \cap \mathcal{C}^1(]t_0, t_0 + \alpha[, \mathbb{R})$  est solution de (3.1) avec  $t_0 + \alpha$  au lieu de  $T$ . On va montrer que  $y$  est solution de (3.3). Comme  $y$  est de classe  $\mathcal{C}^1$  sur l'intervalle  $]t_0, t_0 + \alpha[$ , on a pour  $\varepsilon$  et  $t$  tels que  $t_0 < t_0 + \varepsilon < t < t_0 + \alpha$

$$y(t) - y(t_0 + \varepsilon) = \int_{t_0 + \varepsilon}^t y'(s) ds$$

et donc, avec l'équation donnée par (3.1),

$$y(t) - y(t_0 + \varepsilon) = \int_{t_0 + \varepsilon}^t f(s, y(s)) ds.$$

Lorsque  $\varepsilon \rightarrow 0$ , obtient, pour tout  $t_0 < t < t_0 + \alpha$ ,

$$y(t) - y_0 = \int_{t_0}^t f(s, y(s)) ds.$$

(Noter que cette égalité est aussi vraie pour  $t = t_0$  et aussi, par continuité pour  $t = t_0 + \alpha$ .) Ceci montre que  $y$  est bien solution de (3.3).

On montre maintenant la réciproque. On suppose donc que  $y \in \mathcal{C}([t_0, t_0 + \alpha], \mathbb{R})$  et que  $y$  est solution de l'équation intégrale donnée en (3.3). Comme  $f$  est continue, la fonction  $t \mapsto f(t, y(t))$  est continue. L'équation (3.3) nous montre alors que  $y$  est de classe  $\mathcal{C}^1$  sur l'intervalle  $]t_0, t_0 + \alpha[$  et que  $y'(t) = f(t, y(t))$  pour tout  $t \in ]t_0, t_0 + \alpha[$ .  $\square$

### 3.1.2 Décomposition de la preuve du théorème 1

La démonstration du théorème d'existence locale (théorème 1) est constructive : elle consiste à construire une suite  $(y_n)_{n \geq 1}$  de fonctions continues de  $[t_0, t_0 + \alpha]$  dans  $\mathbb{R}$  de manière à ce que cette suite converge vers une solution de la forme intégrale (3.3) du problème de Cauchy.

Pour  $\alpha > 0$  (avec  $t_0 + \alpha < T$ ), cette suite est construite ainsi :

**Initialisation :**  $y_1(t) = y_0$  pour tout  $t \in [t_0, t_0 + \alpha]$ .

**Itérations :** Pour  $n \geq 1$ ,  $y_{n+1}(t) = y_0 + \int_{t_0}^t f(s, y_n(s)) ds$  pour tout  $t \in [t_0, t_0 + \alpha]$ .

On se donne  $A > 0$  et  $a > 0$  tels que  $y_0 \in [-A, A]$  et  $t_0 < a < T$ . On va montrer que l'on peut choisir  $0 < \alpha$  tel que  $t_0 + \alpha \leq a$  et :

- (P1)  $y_n \in \mathcal{C}([t_0, t_0 + \alpha], [-A - 1, A + 1])$  pour tout  $n \in \mathbb{N}^*$ ,
- (P2)  $y_n$  converge uniformément sur  $[t_0, t_0 + \alpha]$  vers une fonction  $y$ , qui est continue.
- (P3) Cette fonction continue  $y$  est solution de (3.3) (et donc solution de (3.1) sur  $]t_0, t_0 + \alpha[$  grâce à la proposition 2).

Il est intéressant de noter que la choix de  $\alpha$  va dépendre de  $A$  et  $a$  (et, bien sûr, de la fonction  $f$ ) mais ne va pas dépendre de  $t_0$  et  $y_0$  pourvu que  $y_0 \in [-A, A]$  et que  $t_0 + \alpha \leq a$  (c'est-à-dire que l'on peut changer le temps initial et la donnée initiale sans changer  $\alpha$ ). Ceci sera très utile pour montrer la propriété 3.8 (explosion en temps fini, qui constitue la dernière partie du théorème 4, voir section 3.2).

### 3.1.3 Preuve de (P1)

On commence par montrer que la propriété (P1) est vraie pour  $\alpha$  suffisamment petit. On pose

$$M = \max_{\substack{s \in [t_0, a] \\ z \in [-A-1, A+1]}} |f(s, z)| \text{ et } \alpha_1 = 1/M.$$

Comme  $f$  est continue et que  $[-A-1, A+1] \times [t_0, a]$  est compact, on a  $M < +\infty$ . On a alors  $\alpha_1 > 0$ . On va montrer par récurrence sur  $n$  que, si  $\alpha \leq \alpha_1$  et  $t_0 + \alpha \leq a$ ,

$$y_n \in \mathcal{C}([t_0, t_0 + \alpha], [-A-1, A+1]) \text{ pour tout } n \geq 1.$$

On a bien  $y_1 \in \mathcal{C}([t_0, t_0 + \alpha], [-A-1, A+1])$  car  $y_1(t) = y_0$  pour tout  $t$ .

Soit  $n \geq 1$ . On suppose que  $y_n \in \mathcal{C}([t_0, t_0 + \alpha], [-A-1, A+1])$ . La définition de  $y_{n+1}$  à partir de  $y_n$  donne alors si  $\alpha \leq \alpha_1$  et  $t_0 + \alpha \leq a$ ,

$$\begin{aligned} \forall t \in [t_0, t_0 + \alpha], |y_{n+1}(t)| &\leq |y_0| + \int_{t_0}^t \underbrace{|f(s, y_n(s))|}_{\leq M} ds \\ &\leq A + \alpha M \leq A + \alpha_1 M \leq A + 1. \end{aligned}$$

On a donc  $y_{n+1} \in \mathcal{C}([t_0, t_0 + \alpha], [-A-1, A+1])$ . On a bien montré, par récurrence sur  $n$ , que la condition ( $\alpha \leq \alpha_1$  et  $t_0 + \alpha \leq a$ ) donne  $y_n \in \mathcal{C}([t_0, t_0 + \alpha], [-A-1, A+1])$ .

La propriété (P1) est donc vraie dès que  $\alpha \leq \alpha_1$  et  $t_0 + \alpha \leq a$ .

Noter que  $\alpha_1$  ne dépend que de  $a$  et  $A$  (et, bien sûr,  $f$ ).

### 3.1.4 Preuve de (P2)

On suppose maintenant que  $0 < \alpha \leq \alpha_1$  et  $t_0 + \alpha \leq a$  et on cherche à vérifier la propriété (P2). Pour cela, on va devoir réduire  $\alpha$ .

Pour montrer la propriété (P2), on utilise le fait que  $f$  est localement lipschitzienne par rapport à son deuxième argument. Plus précisément on utilise l'inégalité (3.2) pour  $a$  et  $A+1$ . Il existe  $C$  (correspondant à  $C_{a, A+1}$  dans l'inégalité (3.2)) tel que

$$|f(t, x) - f(t, z)| \leq C|x - z|, \quad \forall t \in [t_0, a] \quad \forall x, z \in [-A-1, A+1]. \quad (3.4)$$

(Noter qu'on peut supposer  $C > 0$ .)

Comme  $y_n(s), y_{n-1}(s) \in [-A-1, A+1]$  (grâce à (P1)) l'inégalité (3.4) donne pour tout  $n \geq 2$  et tout  $t \in [t_0, t_0 + \alpha]$ ,

$$|y_{n+1}(t) - y_n(t)| = \left| \int_{t_0}^t f(s, y_n(s)) ds - \int_{t_0}^t f(s, y_{n-1}(s)) ds \right| \quad (3.5)$$

$$\leq \int_{t_0}^t |f(s, y_n(s)) - f(s, y_{n-1}(s))| ds \quad (3.6)$$

$$\leq \int_{t_0}^t C |y_n(s) - y_{n-1}(s)| ds \quad (3.7)$$

Sur l'espace vectoriel  $\mathcal{C}([t_0, t_0 + \alpha], \mathbb{R})$ , on peut définir une norme par

$$\|z\| = \max_{t \in [t_0, t_0 + \alpha]} |z(t)|.$$

Avec cette norme, l'inégalité (3.7) donne pour tout  $n \geq 2$  et tout  $t \in [t_0, t_0 + \alpha]$ ,

$$\begin{aligned} |y_{n+1}(t) - y_n(t)| &\leq \int_{t_0}^t C \|y_n - y_{n-1}\| ds \\ &\leq C\alpha \|y_n - y_{n-1}\| \\ \underbrace{\|y_{n+1} - y_n\|}_{\max_{t \in [t_0, t_0 + \alpha]} |y_{n+1}(t) - y_n(t)|} &\leq C\alpha \|y_n - y_{n-1}\| \end{aligned}$$

On pose  $\alpha_2 = \frac{1}{2C}$  de sorte que  $C\alpha_2 = \frac{1}{2}$  et on choisit  $\alpha$  tel que  $\alpha \leq \alpha_1$ ,  $\alpha \leq \alpha_2$  et  $t_0 + \alpha \leq a$  (par exemple  $\alpha = \min\{\alpha_1, \alpha_2, a - t_0\}$ ).

(Noter encore que  $\alpha_2$  ne dépend que de  $a$ ,  $A$  et  $f$ .) Avec ce choix de  $\alpha$ ,

$$\|y_{n+1} - y_n\| \leq \frac{1}{2} \|y_n - y_{n-1}\| \text{ pour tout } n \geq 2.$$

Ce qui donne pour  $n = 2$

$$\|y_3 - y_2\| \leq \frac{1}{2} \|y_2 - y_1\|,$$

pour  $n = 3$

$$\|y_4 - y_3\| \leq \frac{1}{2} \|y_3 - y_2\| \leq \left(\frac{1}{2}\right)^2 \|y_2 - y_1\|,$$

et donc, par récurrence,

$$\|y_{n+1} - y_n\| \leq \left(\frac{1}{2}\right)^{n-1} \|y_2 - y_1\| \text{ pour tout } n \geq 1.$$

En effectuant une inégalité triangulaire pour la norme  $\|\cdot\|$ , on obtient

$$\begin{aligned} \|y_{n+p} - y_n\| &\leq \|y_{n+p} - y_{n+p-1}\| + \|y_{n+p-1} - y_{n+p-2}\| + \dots + \|y_{n+1} - y_n\| \\ &\leq \left[ \left(\frac{1}{2}\right)^{n+p-2} + \left(\frac{1}{2}\right)^{n+p-3} + \dots + \left(\frac{1}{2}\right)^{n-1} \right] \|y_2 - y_1\| \\ &\leq \left(\frac{1}{2}\right)^{n-1} \left[ 1 + \frac{1}{2} + \dots + \left(\frac{1}{2}\right)^{p-1} \right] \|y_2 - y_1\| \\ &\leq \left(\frac{1}{2}\right)^{n-1} 2 \|y_2 - y_1\|, \quad n \geq 1, \quad p \geq 1 \end{aligned}$$

Dans la dernière ligne de ce calcul, nous avons utilisé que, pour  $0 < b < 1$ ,

$$\sum_{p=0}^{+\infty} b^p = \frac{1}{1-b}.$$

Nous venons donc de montrer que  $(y_n)_{n \in \mathbb{N}^*}$  est une suite de Cauchy de l'espace vectoriel  $\mathcal{C}([t_0, t_0 + \alpha], \mathbb{R})$  pour la norme  $\|\cdot\|$ , c'est-à-dire que

$$\forall \varepsilon > 0, \quad \exists n_0 \quad | \quad n \geq n_0, p \geq 1 \Rightarrow \|y_{n+p} - y_n\| \leq \varepsilon$$

En particulier pour tout  $t \in [t_0, t_0 + \alpha]$  la suite  $(y_n(t))_{n \geq 1}$  est une suite de Cauchy dans  $\mathbb{R}$ , car

$$|y_{n+p}(t) - y_n(t)| \leq \|y_{n+p} - y_n\|.$$

Donc, la suite  $(y_n(t))_{n \geq 1}$  converge dans  $\mathbb{R}$ . Pour tout  $t \in [t_0, t_0 + \alpha]$ , on pose  $\lim_{n \rightarrow +\infty} y_n(t) = y(t)$ . On a ainsi défini une fonction

$$y : [t_0, t_0 + \alpha] \longrightarrow \mathbb{R}.$$

On montre maintenant que  $y \in \mathcal{C}([t_0, t_0 + \alpha], [-A - 1, A + 1])$ .

On remarque tout d'abord que  $y_n(t) \in [-A - 1, A + 1]$  pour tout  $n \geq 1$  et tout  $t \in [t_0, t_0 + \alpha]$  et donc

$$y(t) = \lim_{n \rightarrow +\infty} y_n(t) \in [-A - 1, A + 1].$$

Pour montrer la continuité de  $y$ , on remarque que

$$|y_{n+p}(t) - y_n(t)| \leq \|y_{n+p} - y_n\| \leq \underbrace{\left(\frac{1}{2}\right)^{n-1} \times 2\|y_2 - y_1\|}_{\text{ne dépend pas de } p}$$

Dans cette inégalité, on fixe  $n$  et  $t$ . Lorsque  $p$  tend vers  $+\infty$  on obtient

$$|y(t) - y_n(t)| \leq \left(\frac{1}{2}\right)^{n-1} \times 2\|y_2 - y_1\|.$$

On déduit que

$$y_n \longrightarrow y, \quad \text{uniformément sur } [t_0, t_0 + \alpha]$$

Pour conclure, on utilise le lemme suivant, dont la preuve est laissée à titre d'exercice.

**Lemme 7** (Continuité d'une limite uniforme de fonctions continues).

Soient  $I$  un intervalle de  $\mathbb{R}$  et  $(y_n)_{n \geq 1}$  une suite de fonctions continues de  $I$  dans  $\mathbb{R}$ . On suppose que cette suite converge uniformément sur  $I$  vers une fonction  $y$ . La fonction  $y$  est alors continue.

On obtient donc que  $y$  est continue sur  $[t_0, t_0 + \alpha]$ . Si on récapitule ce que nous venons de démontrer,

$$\alpha = \min\{\alpha_1, \alpha_2, a - t_0\} > 0, \quad y_n \xrightarrow{\text{uniformément}} y, \quad \text{et} \quad y \in \mathcal{C}([t_0, t_0 + \alpha], [-A - 1, A + 1]).$$

### 3.1.5 Preuve de (P3)

Il reste à montrer que  $y$  est solution de (3.3). Pour tout  $n \geq 1$ ,

$$\forall t \in [t_0, t_0 + \alpha], \quad y_{n+1}(t) = y_0 + \int_{t_0}^t f(s, y_n(s)) ds$$

Dans cette égalité, on fixe  $t$  et on fait tendre  $n$  vers  $+\infty$ , alors  $y_{n+1}(t) \longrightarrow y(t)$ . Puis pour le deuxième membre de cette égalité, on utilise les résultats suivants, laissés à titre d'exercice.

- 1) On pose  $z_n(s) = f(s, y_n(s))$ . La fonction  $f$  est uniformément continue sur le compact  $[t_0, t_0 + \alpha] \times [-A - 1, A + 1]$  et  $y_n$  converge uniformément vers  $y$  sur  $[t_0, t_0 + \alpha]$ . Alors, on peut montrer que

$$z_n \longrightarrow f(\cdot, y) \quad \text{uniformément sur } [t_0, t_0 + \alpha].$$

2) Grâce à la convergence uniforme de la suite  $(z_n)_{n \geq 1}$ , on peut montrer que pour tout  $t \in [t_0, t_0 + \alpha]$ ,

$$\int_{t_0}^t z_n(s) ds \longrightarrow \int_{t_0}^t f(s, y(s)) ds.$$

Avec ces deux résultats, on conclut que  $y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds$  pour tout  $t \in [t_0, t_0 + \alpha]$ .

Ceci termine la démonstration du théorème 1 (théorème d'existence locale). On rappelle enfin que cette existence locale est montrée avec  $\alpha = \min\{\alpha_1, \alpha_2, a - t_0\}$  et que  $\alpha_1$  et  $\alpha_2$  ne dépendent que de  $a$ ,  $A$  et  $f$  (sous la condition que  $y_0 \in [-A, A]$ ).

**Remarque 4.** Pour ceux qui connaissent le théorème du point fixe de contraction, le théorème 1 est une conséquence de ce théorème du point fixe en prenant (avec les notations de la démonstration ci-dessus) l'espace  $E = C([t_0, t_0 + \alpha], [-A - 1, A + 1])$  (qui est métrique complet avec sa distance naturelle) et l'application  $T$  de  $E$  dans  $E$  définie par

$$T(y)(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds, \text{ pour } t \in [t_0, t_0 + \alpha].$$

La démonstration faite ci dessus correspond à celle du théorème du point fixe.

## 3.2 Preuve du résultat d'explosion en temps fini

### 3.2.1 Rappel du résultat et preuve

On rappelle que sous l'hypothèse que  $f$  est localement lipschitzienne par rapport à son deuxième argument, on a toujours unicité (théorème 2) et qu'il existe donc une solution maximale (théorème 4), c'est-à-dire qu'il existe  $T_m$  et  $y$  tels que  $t_0 < T_m \leq T$  et  $y$  solution de (3.1) sur  $[t_0, T_m[$  et il n'y a pas de solution sur  $[t_0, t_0 + \alpha]$  avec  $T_m < t_0 + \alpha \leq T$ . On montre maintenant la dernière partie du théorème 4, c.à.d. que sous les hypothèses du théorème 4, si  $y$  est la solution maximale définie sur  $[t_0, T_m[$ , avec  $T_m \leq T$ , alors

$$T_m < T \quad \Rightarrow \quad \lim_{t \rightarrow T_m} |y(t)| = +\infty. \quad (3.8)$$

On raisonne par l'absurde. Soit  $y$  la solution maximale. Elle est définie sur l'intervalle  $[0, T_m[$ .

On suppose que  $T_m < T$  et que  $|y(t)|$  ne tend pas vers  $+\infty$  lorsque  $t \rightarrow T_m$ , donc

$$\begin{cases} \exists A > 0 \text{ et une suite } (t_n)_{n \in \mathbb{N}} \text{ tels que } \lim_{n \rightarrow +\infty} t_n = T_m \\ \text{et } |y(t_n)| \leq A \text{ pour tout } n. \end{cases}$$

On choisit maintenant  $a$  tel que  $T_m < a < T$ . et, pour  $n \geq 3$ , on considère le problème

$$\begin{aligned} z'_n(t) &= f(t, z_n(t)) \text{ pour } t \in [t_n, T[, \\ z_n(t_n) &= y(t_n). \end{aligned}$$

La démonstration du théorème 1 (que nous venons de faire) nous montre que ce problème a une solution sur l'intervalle  $[t_n, t_n + \alpha_n]$  avec  $\alpha_n = \min\{\alpha_1, \alpha_2, a - t_n\}$ . Les nombres  $\alpha_1, \alpha_2$  ne dépendent pas de  $n$  (ils ne dépendent que de  $a$  et  $A$ , le nombre  $A$  étant un majorant de  $|y(t_n)|$  pour tout  $n$ ).

On définit maintenant la fonction  $u_n$  (voir figure 3.1) par

$$\begin{aligned} u_n(t) &= y(t) \text{ pour } t \in [t_0, t_n], \\ u_n(t) &= z_n(t) \text{ pour } t \in [t_n, t_n + \alpha_n]. \end{aligned}$$

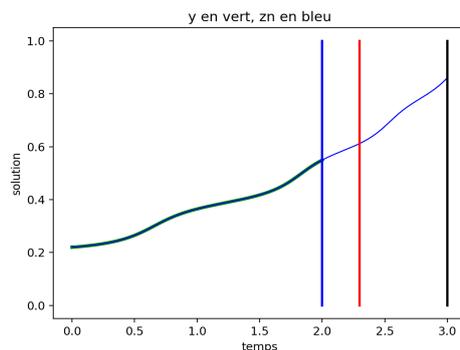


Fig. 3.1:  $t_0 = 0$ ,  $t_n = 2$ ,  $\alpha_n = 1$ ,  $T_m = 2.3$

La fonction  $u_n$  est alors solution de (3.1) sur  $[t_0, t_n + \alpha_n]$  (on remarque par exemple qu'elle est solution de (3.3) sur  $[t_0, t_n + \alpha_n]$ ). Or  $\lim_{n \rightarrow +\infty} (t_n + \alpha_n) = T_m + \min\{\alpha_1, \alpha_2, a - T_m\} > T_m$ . Pour  $n$  assez grand on a donc  $t_n + \alpha_n > T_m$ , ce qui est en contradiction avec la définition de  $T_m$ .

### 3.2.2 Commentaires et exemples

On rappelle qu'il est possible d'avoir  $T_m = T$  et  $\lim_{t \rightarrow T_m} |y(t)| = +\infty$ , voir exemple 8 du cours précédent.

La principale conséquence du théorème 2 est que, sous les hypothèses du théorème 4, pour montrer que  $T_m = T$  (et donc l'existence globale de la solution) il suffit de montrer que

$$\sup_{t \in [t_0, t_0 + \alpha]} |y(t)| < +\infty, \quad \text{si } t_0 + \alpha < T$$

On illustre cette conséquence sur l'exemple suivant, vu en TD.

**Exemple 9** (Etude qualitative d'une EDO). *On considère le problème de Cauchy (vu en TD)*

$$y'(t) = ay(t) - by(t)^2, \quad t > 0, \tag{3.9}$$

$$y(0) = y_0 \text{ avec } 0 < y_0 < \frac{a}{b}. \tag{3.10}$$

Les nombres  $a, b$  sont donnés,  $a > 0, b > 0$ . L'équation (3.9) admet 0 et  $\frac{a}{b}$  comme solutions stationnaires. Il y a existence locale au problème (3.9)-(3.10). Par le théorème 3 sur les trajectoires, la solution est "coincée" entre 0 et  $\frac{a}{b}$ . Par le théorème 4, si  $T_m < +\infty$  alors la solution explose, mais ceci est impossible car  $0 < y(t) < \frac{a}{b}$ . On en déduit que  $T_m = +\infty$ . Il y a donc existence globale.

# Cours 4. Schémas numériques

## 4.1 Introduction

Nous allons introduire dans ce chapitre des schémas numériques permettant un calcul approché de la solution des équations différentielles. Ces schémas peuvent être utilisés pour des équations dont on connaît la solution exacte (avec des fonctions “usuelles”), et dans ce cas on peut évaluer facilement leur précision grâce à la connaissance de cette solution exacte, mais ils sont surtout utiles pour des équations dont on ne connaît pas la solution exacte et pour lesquelles c’est donc un moyen d’obtenir une solution approchée.

Nous rappelons ici le problème de Cauchy :

$$\begin{aligned}y'(t) &= f(t, y(t)), \quad t \in ]t_0, T[ = I, \\y(t_0) &= y_0,\end{aligned}\tag{4.1}$$

où  $t_0, T, f$  et  $y_0$  sont données,  $t_0, y_0 \in \mathbb{R}, t_0 < T \leq +\infty$ . On suppose que la fonction  $f$  appartient à  $C([t_0, T[ \times \mathbb{R}, \mathbb{R})$  et est localement lipschitzienne par rapport à son deuxième argument. On note  $y$  la solution maximale de (4.1). On rappelle qu’elle est définie sur l’intervalle  $[t_0, T_m[$  avec  $t_0 < T_m \leq T$ . On cherche à calculer  $y$ , tout en sachant qu’elle existe.

Dans les chapitres précédents nous avons pu parfois exprimer  $y$  en fonction des fonctions usuelles (trigonométriques, exponentielle, polynomiales...) mais ceci n’est pas toujours possible, comme le montre l’exemple suivant.

**Exemple 10** (Problèmes de Cauchy avec et sans solution explicite). *Soit le problème suivant :*

$$\begin{aligned}y'(t) &= f(t, y(t)) = y(t)(1 - y(t)), \quad t > 0, \\y_0 &= \frac{1}{2}.\end{aligned}\tag{4.2}$$

*Nous pouvons calculer aisément la solution de (4.2), nous obtenons :*

$$y(t) = \frac{1}{1 + e^{-t}}\tag{4.3}$$

*Considérons maintenant le problème suivant :*

$$\begin{aligned}y'(t) &= f(t, y(t)) = y(t)(1 + e^{-y(t)}) + e^{2t}, \quad t > 0, \\y_0 &= 0.\end{aligned}\tag{4.4}$$

*On ne sait pas exprimer la solution du problème (4.4) à partir de fonctions “usuelles”. Mais, grâce aux schémas numériques qui sont l’objet de ce chapitre, on peut calculer des approximations aussi précises que l’on veut de cette solution exacte (voir le tp 1 pour le cas du problème (4.4)).*

## 4.2 Schémas numériques explicites

### 4.2.1 Euler explicite

On se place sous les hypothèses rappelées en introduction. Les données sont  $t_0, T, f$  et  $y_0$ . La fonction  $f$  appartient à  $C([t_0, T] \times \mathbb{R}, \mathbb{R})$  et est localement lipschitzienne par rapport à son deuxième argument. On note  $y$  la solution maximale de (4.1), elle est définie sur l'intervalle  $[t_0, T_m[$  où  $t_0 < T_m \leq T$ . Pour simplifier les notations, on suppose  $t_0 = 0$ .

On se donne  $\alpha$ , avec  $0 < \alpha < T_m$ . La fonction  $y$  est continue sur le compact  $[0, \alpha]$ , elle est donc bornée. On note  $A = \max\{|y(t)|, t \in [0, \alpha]\}$ . Pour présenter le schéma d'Euler explicite et montrer sa convergence sur  $[0, \alpha]$  (et donc aussi sur  $[0, T_m[$  car  $\alpha$  est arbitrairement choisi entre 0 et  $T_m$ ), on suppose que  $y$  est de classe  $C^2$  sur  $]0, \alpha[$  et que  $M_2 = \max\{|y''(s)|, s \in ]0, \alpha[\} < +\infty$ . (Cette hypothèse est vérifiée si la fonction  $f$  est suffisamment régulière, ce qui est le cas le plus fréquent.)

Notre objectif est de calculer  $y$  sur l'intervalle  $[0, \alpha]$ . Pour cela, on divise l'intervalle  $[0, \alpha]$  en  $N$  sous-intervalles, où  $N \in \mathbb{N}^*$ . On pose  $h = \alpha/N$ , le nombre  $h$  est le pas de temps du schéma numérique.

On pose  $t_n = nh$  pour  $n \in \llbracket 0, N \rrbracket$  et on note  $y_n$  l'approximation recherchée de  $y(nh)$ , pour  $n \in \llbracket 0, N \rrbracket$ . Comme  $y$  est de classe  $C^2$  et que  $y'(t) = f(t, y(t))$ , un développement de Taylor nous donne pour tout  $n \in \mathbb{N}$ ,

$$\frac{y(t_{n+1}) - y(t_n)}{h} = \underbrace{f(t_n, y(t_n))}_{y'(t_n)} + \varepsilon_n$$

où  $|\varepsilon_n| \leq hM_2$ . Le terme  $\varepsilon_n$  est appelé *erreur de consistance*.

Ce développement de Taylor suggère alors une méthode pour calculer une valeur approchée  $y_n$  de  $y(t_n)$  pour tout  $n$  :

$$\forall n \in \llbracket 0, N-1 \rrbracket, \quad y_{n+1} = y_n + hf(t_n, y_n). \quad (4.5)$$

À partir de  $y_0$ , qui est connu, on calcule ainsi tous les  $y_n$ ,  $n \in \llbracket 1, N \rrbracket$ .

#### Questions

- 1) Bien sûr, en général,  $y_n \neq y(t_n)$ ; la question naturelle à se poser est la convergence (quand  $h \rightarrow 0$ ) de la solution approchée définie à partir des valeurs  $y_n$ , vers la solution exacte  $y$ .
- 2) Dans l'affirmative, à quelle vitesse converge-t-elle, et plus précisément, avec quel ordre par rapport au pas de discrétisation  $h$  ?

Pour répondre à ces questions on va montrer que :

$$\max_{0 \leq n \leq N} |y_n - y(t_n)| \leq Ch, \quad (4.6)$$

où  $C \in \mathbb{R}$  ne dépend que de  $y_0, f, \alpha$ . On posera par la suite  $e_n = |y_n - y(t_n)|$ . Cette erreur est appelée *erreur de discrétisation*. L'inégalité (4.6) entraîne que l'erreur est au moins d'ordre 1. Plus généralement, la définition de l'ordre d'un schéma est la suivante.

**Définition 8** (Ordre d'un schéma numérique). *Un schéma est dit au moins d'ordre  $k$  (avec  $k \geq 1$ ), s'il existe  $C \in \mathbb{R}_+$  tel que*

$$\max_{0 \leq n \leq N} |y_n - y(t_n)| \leq Ch^k, \quad (4.7)$$

*Il est d'ordre  $k$  (avec  $k \geq 1$ ), s'il existe  $C_1, C_2 \in \mathbb{R}_+$  tels que*

$$C_1 h^k \max_{0 \leq n \leq N} |y_n - y(t_n)| \leq C_2 h^k, \quad (4.8)$$

**Proposition 3** (Ordre du schéma d'Euler explicite). *Sous les hypothèses précédentes, le schéma d'Euler explicite est d'ordre 1.*

*Démonstration.* Pour démontrer cette proposition, nous allons montrer que la solution approchée donnée par le schéma d'Euler explicite satisfait (4.7) avec  $k = 1$ , c'est-à-dire (4.6). Pour  $0 \leq \alpha \leq T$  et  $h = \alpha/N$ , avec  $N \in \mathbb{N}^*$  fixé, l'objectif est donc de démontrer qu'il existe  $C \in \mathbb{R}$ , ne dépendant que de  $y_0, f, \alpha$ , tel que

$$e_n = |y_n - y(t_n)| \leq Ch \text{ pour tout } n \in \llbracket 0, N \rrbracket. \quad (4.9)$$

Noter que pour  $n = 0$ , l'inégalité (4.9) est vraie quelque soit  $C$  car  $y_0 = y(0)$ . On rappelle le schéma numérique et la majoration de l'erreur de consistance :

$$y_{n+1} = y_n + hf(t_n, y_n), \quad \forall n \in \llbracket 0, N-1 \rrbracket, \quad (4.10a)$$

$$y(t_{n+1}) = y(t_n) + hf(t_n, y(t_n)) + \varepsilon_n, \quad |\varepsilon_n| < h^2 M_2. \quad (4.10b)$$

En soustrayant l'équation (4.10b) à l'équation (4.10a) on obtient :

$$y_{n+1} - y(t_{n+1}) = y_n - y(t_n) + h(f(t_n, y_n) - f(t_n, y(t_n))) - \varepsilon_n \quad (4.11)$$

En passant en valeur absolue on a donc :

$$|y_{n+1} - y(t_{n+1})| \leq |y_n - y(t_n)| + h|f(t_n, y_n) - f(t_n, y(t_n))| + h^2 M_2,$$

c.à.d.

$$e_{n+1} \leq e_n + h|f(t_n, y_n) - f(t_n, y(t_n))| + h^2 M_2. \quad (4.12)$$

On a envie alors d'en déduire une relation de la forme :

$$e_{n+1} \leq e_n + hL|y_n - y(t_n)| + h^2 M_2, \quad (4.13)$$

en utilisant le caractère localement lipschitzien de  $f$  par rapport à son deuxième argument ( $L$  serait la constante de Lipschitz de  $f$ ). Mais  $f$  n'est que localement lipschitzienne, et donc pour déduire (4.13) de (4.12) il faut que  $y_n$  et  $y(t_n)$  restent dans un intervalle borné fixe (c'est-à-dire indépendant de  $n$  et  $h$ , cet intervalle borné ne dépendant que de  $y_0, f, \alpha$ ). Comme  $y$  est continue sur  $[0, \alpha]$ ,  $\alpha < +\infty$ , on a

$$y(t_n) \in [-A, A], \quad 0 \leq t_n \leq \alpha,$$

avec  $A = \max\{|y(t)|, t \in [0, \alpha]\}$ . Mais il n'y a pas de raison que  $y_n \in [-A, A]$  pour tout  $1 \leq n \leq N$ . Par contre si on montre qu'il existe  $B > A$  tel que  $y_n \in [-B, B]$  pour tout  $1 \leq n \leq N$ , on aura (4.13) avec  $L = L_{\alpha, B}$ , où  $L_{\alpha, B}$  est donné par le caractère localement lipschitzien de  $f$  (voir Définition 5) et on pourra en déduire (4.6). En fait, on va montrer qu'il existe  $B > A$  et  $h_0 > 0$  tel que  $y_n \in [-B, B]$  pour tout  $1 \leq n \leq N$  et pour tout  $h \leq h_0$ . La relation (4.6) sera ainsi vérifiée pour tout  $h \leq h_0$ .

Notons que le caractère localement lipschitzien de  $f$  est donc suffisant pour montrer (4.6), ce qui est un résultat intéressant, car la plupart des fonctions rencontrées ne sont pas globalement lipschitziennes.

On se donne maintenant  $B > A$  on va montrer qu'il existe  $h_0 > 0$  tel que, si  $h \leq h_0$ ,  $y_n \in [-B, B]$  pour tout  $n$  et (4.6) est vrai.

On commence par un calcul liminaire. Supposons que  $y_n \in [-B, B]$  pour tout  $n$ . On note alors  $L_f$  la constante de Lipschitz (par rapport à son deuxième argument) de  $f$  sur  $[0, \alpha] \times [-B, B]$ . Comme on a aussi  $y(t_n) \in [-B, B]$  (car  $B > A$ ) l'inégalité (4.13) est vraie, c'est-à-dire

$$\begin{aligned} \forall n \in \llbracket 0, N-1 \rrbracket, \quad e_{n+1} &\leq e_n + hL_f e_n + h^2 M_2, \\ &\leq e_n(1 + hL_f) + h^2 M_2. \end{aligned}$$

On pose alors :

- 1)  $a = 1 + hL_f$ ,
- 2)  $b = h^2 M_2$ ,
- 3)  $F_n = \frac{e_n}{a^n}$  de sorte que  $F_{n+1} \leq F_n + \frac{b}{a^{n+1}}$ ,  $F_0 = 0$ .
- 4)  $G_{n+1} = G_n + \frac{b}{a^{n+1}}$ ,  $G_0 = F_0$ .

On montre par récurrence que  $F_n \leq G_n$  pour tout  $n$ .

On a bien  $F_0 \leq G_0$ , puis si  $F_n \leq G_n$ , on a

$$F_{n+1} \leq F_n + \frac{b}{a^{n+1}} \leq G_n + \frac{b}{a^{n+1}} = G_{n+1}.$$

On calcule maintenant  $G_n$ ,

$$G_n = G_{n-1} + \frac{b}{a^n} = G_{n-2} + \frac{b}{a^{n-1}} + \frac{b}{a^n} = G_0 + \frac{b}{a} + \dots + \frac{b}{a^n}.$$

On a alors :

$$G_n = \sum_{p=1}^n \frac{b}{a^p} = \frac{b}{a} \left( \sum_{p=0}^{n-1} \frac{1}{a^p} \right) = \frac{b}{a} \cdot \frac{1 - \frac{1}{a^n}}{1 - \frac{1}{a}},$$

et donc

$$F_n \leq G_n \leq \frac{b}{a-1} \frac{a^n - 1}{a^n}.$$

Nous pouvons en déduire que

$$e_n = a^n F_n \leq \frac{b}{a-1} (a^n - 1).$$

Nous avons finalement

$$|y_n - y(t_n)| = e_n \leq \frac{M_2 h^2}{hL_f} ((1 + hL_f)^n - 1).$$

Or :  $1 + hL_f \leq e^{hL_f}$ , d'où pour  $0 \leq n \leq N$  et comme  $Nh = \alpha$ , on a

$$\begin{aligned} e_n &\leq \frac{M_2}{L_f} h (e^{nhL_f} - 1) \\ &\leq \frac{M_2}{L_f} (e^{\alpha L_f} - 1) h. \end{aligned}$$

On obtient donc que

$$e_n = |y_n - y(t_n)| \leq Ch, \quad 0 \leq n \leq N, \quad (4.14)$$

avec  $C = \frac{M_2}{L_f} (e^{\alpha L_f} - 1)$ . Pour l'obtenir, nous avons supposé que  $y_n \in [-B, B]$  pour tout  $n$ .

Pour montrer  $y_n \in [-B, B]$  pour tout  $n$  (au moins pour  $h$  assez petit), on choisit  $B > A$  (on peut prendre par exemple  $B = A + 1$ ) et on définit  $h_0$  par la formule (avec  $C = C_{\alpha, B}$ )

$$\frac{M_2}{C} (e^{\alpha C} - 1) h_0 = B - A.$$

On remarque alors que si  $h \leq h_0$  la formule (4.14) donne

$$|y_n| \leq |y(t_n)| + (B - A) \leq B.$$

Une récurrence simple sur  $n$  permet alors de montrer que  $y_n \in [-B, B]$  pour tout  $n$ . Plus précisément,  $y_0 \in [-B, B]$  et si  $y_p \in [-B, B]$  pour  $p \leq n$ , les calculs précédents, valables jusqu'à l'ordre  $n$ , donnent la formule (4.14) et donc  $y_{n+1} \in [-B, B]$ . □

## 4.2.2 Schéma de Heun : RK2, Runge-Kutta 2

Le défaut du schéma d'Euler est qu'il n'est pas très précis : l'erreur se comporte comme  $h$  et il faut donc prendre des pas de temps très petits pour bien approcher la solution. L'objectif est ici d'obtenir un schéma plus précis que le précédent : pour un pas de discrétisation égal, nous voulons avoir une erreur qui soit plus petite que la précédente, de la forme  $e_n \leq \tilde{D}h^2$  (ce qui est plus petit que  $Dh$  pour  $h \leq D/\tilde{D}$ ). Le schéma est alors dit d'ordre au moins 2. Pour construire ce schéma plus précis que le précédent, nous allons d'abord nous intéresser à la forme intégrale du problème. Comme  $y'(t) = f(t, y(t))$ ,

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(s, y(s)) ds. \quad (4.15)$$

À partir de cette forme intégrale nous pouvons approcher  $\int_{t_n}^{t_{n+1}} f(s, y(s)) ds$  par  $hf(t_n, y(t_n))$  avec une erreur  $\varepsilon_n$  :

$$y(t_{n+1}) = y(t_n) + hf(t_n, y(t_n)) + \varepsilon_n,$$

où  $|\varepsilon_n| \leq h^2 M_2$ , ce qui permet d'obtenir le schéma suivant :

$$y_{n+1} = y_n + hf(t_n, y_n). \quad (4.16)$$

On reconnaît le schéma d'Euler explicite.

Comment peut-on à partir de (4.15) obtenir un schéma numérique encore plus précis ? Une idée est d'améliorer l'approximation de l'intégrale dans (4.15) en utilisant la formule des trapèzes :

$$y(t_{n+1}) = y(t_n) + \frac{h}{2} \left[ \underbrace{f(t_n, y(t_n))}_{y'(t_n)} + \underbrace{f(t_{n+1}, y(t_{n+1}))}_{y'(t_{n+1})} \right] + \tilde{\varepsilon}_n \quad (4.17)$$

où l'erreur  $|\tilde{\varepsilon}_n| \leq M_3 h^3$  avec  $M_3 = \max\{|y'''(s)|, s \in ]0, \alpha[ \}$  (nous supposons  $y$  assez régulière pour avoir  $M_3 < +\infty$ ).

L'équation (4.17) nous suggère alors le schéma suivant

$$y_{n+1} = y_n + \frac{h}{2} (f(t_n, y_n) + f(t_{n+1}, y_{n+1})). \quad (4.18)$$

L'erreur de consistance  $\tilde{\varepsilon}_n$  est plus petite que l'erreur  $\varepsilon_n$  (quand  $h \rightarrow 0$ ), nous pouvons alors espérer que ce nouveau schéma donne une solution plus proche des valeurs exactes donc plus précis que le précédent. Le problème de ce schéma est qu'il est *implicite* : en effet  $y_{n+1}$  apparaît dans le deuxième membre de (4.18), et on n'a donc pas en général pas d'expression explicite de  $y_{n+1}$  en fonction de  $y_n, t_n$  et  $t_{n+1}$ . Pour remédier à cet inconvénient, on remplace dans (4.18)  $y_{n+1}$  par l'approximation obtenue par le schéma d'Euler explicite à partir de  $y_n$ , c.à.d.  $y_n + hf(t_n, y_n)$  et on obtient le schéma de Heun :

$$y_{n+1} = y_n + \frac{h}{2} (f(t_n, y_n) + f(t_{n+1}, y_n + hf(t_n, y_n))). \quad (4.19)$$

Ce schéma a l'avantage d'être explicite et d'être similaire au schéma (4.18) du point de vue de l'erreur de consistance, c'est-à-dire que

$$y(t_{n+1}) = y(t_n) + \frac{h}{2} [f(t_n, y(t_n)) + f(t_{n+1}, y(t_n) + hf(t_n, y_n))] + \tilde{\tilde{\varepsilon}}_n, \quad (4.20)$$

où l'erreur  $|\tilde{\tilde{\varepsilon}}_n| \leq \tilde{M}_3 h^3$  avec  $\tilde{M}_3$  ne dépendant que des dérivées de  $y$  jusqu'à l'ordre 3.

La démonstration de  $|\tilde{\varepsilon}_n| \leq M_3 h^3$  et  $|\tilde{\varepsilon}_n| \leq \tilde{M}_3 h^3$  est fastidieuse (mais conceptuellement assez simple). Elle consiste seulement à utiliser des développements de Taylor. Elle n'est pas faite ici. Pour obtenir une majoration de l'erreur de discrétisation, c'est-à-dire de  $e_n = |y(t_n) - y_n|$ , on peut alors raisonner comme pour le schéma d'Euler explicite (un point important est d'avoir une borne sur  $y_n$  car  $f$  n'est que localement lipschitzienne). On obtient

$$e_n = |y_n - y(t_n)| \leq \tilde{D} h^2, \text{ pour tout } n \in \llbracket 0, N \rrbracket,$$

où  $\tilde{D}$  ne dépend que de  $y_0, f$  et  $\alpha$ .

### 4.2.3 Schéma RK4

Une fois de plus le but de ce type de schéma numérique est de minimiser l'erreur afin que l'approximation de la solution soit encore plus proche de la solution exacte. Nous souhaitons avoir une erreur de la forme suivante  $e_n \leq \bar{D} h^4$  (le schéma est alors d'ordre au moins 4). On rappelle que la forme intégrale de l'équation s'écrit

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} \underbrace{f(s, y(s))}_{y'(s)} ds$$

On pose alors  $g = y'$ . Le schéma d'Euler explicite consistait à écrire

$$\int_a^{a+h} g(s) ds = hg(a) + \varepsilon_1$$

avec  $|\varepsilon_1| \leq h^2 M_2$ . Dans le cas du schéma RK2 nous avons utilisé

$$\int_a^{a+h} g(s) ds = \frac{h}{2} (g(a) + g(a+h)) + \varepsilon_2,$$

où  $|\varepsilon_2| \leq h^3 M_3$ . Pour le schéma RK4 nous allons utiliser une approximation encore plus précise de l'intégrale

$$\int_a^{a+h} g(s) ds = \frac{h}{6} \left( g(a) + 4g\left(a + \frac{h}{2}\right) + g(a+h) \right) + \varepsilon_3,$$

où  $|\varepsilon_3| \leq h^5 M_5$  avec  $M_5$  ne dépendant que des dérivées de  $g$  jusqu'à l'ordre 5. Ceci nous suggère l'idée suivante

$$y(t_{n+1}) = y(t_n) + \frac{h}{6} \left[ f(t_n, y(t_n)) + 4f\left(t_n + \frac{h}{2}, y\left(t_n + \frac{h}{2}\right)\right) + f(t_{n+1}, y(t_{n+1})) \right] + \bar{\varepsilon}_n,$$

où  $|\bar{\varepsilon}_n| \leq M_5 h^5$  ( $M_5$  ne dépendant que des dérivées de  $y$  jusqu'à l'ordre 5).

Comme on souhaite avoir un schéma explicite, on montre qu'il est possible, pour la partie droite de cette formule, de remplacer les quantités  $y(t_n + h/2)$  et  $y(t_{n+1})$  par des quantités calculées à partir de  $y(t_n)$ , en ayant toujours une erreur de consistance en  $h^5$  (calcul fastidieux).

On obtient alors le schéma RK4 qui s'écrit :

$$y_{n+1} = y_n + \frac{h}{6} \left[ f(t_n, y_n) + 2f\left(t_n + \frac{h}{2}, y_{n,1}\right) + 2f\left(t_n + \frac{h}{2}, y_{n,2}\right) + f(t_{n+1}, y_{n,3}) \right], \quad (4.21)$$

où  $y_{n,1} = y_n + \frac{h}{2} f(t_n, y_n)$ ,  $y_{n,2} = y_n + \frac{h}{2} f\left(t_n + \frac{h}{2}, y_{n,1}\right)$ ,  $y_{n,3} = y_n + h f\left(t_n + \frac{h}{2}, y_{n,2}\right)$ .

Pour ce schéma, avec les mêmes méthodes que précédemment (développement de Taylor, bornes sur  $y_n$ ) on montre que l'erreur de discrétisation (notée  $e_n$ ) est en  $h^4$  (le schéma est donc d'ordre au moins 4).

### 4.3 Schémas numériques implicites

Nous allons introduire les schémas numériques implicites par le moyen d'exemples.

#### 4.3.1 Premier exemple : une équation linéaire

Considérons le problème de Cauchy suivant, avec  $a > 0$  :

$$\begin{cases} y(t) = -ay(t), & t > 0, \\ y(0) = 1 \end{cases} \quad (4.22)$$

La solution est  $y(t) = e^{-at}$ , pour tout  $t > 0$ . Cette équation modélise l'évolution d'une quantité notée  $y$ .

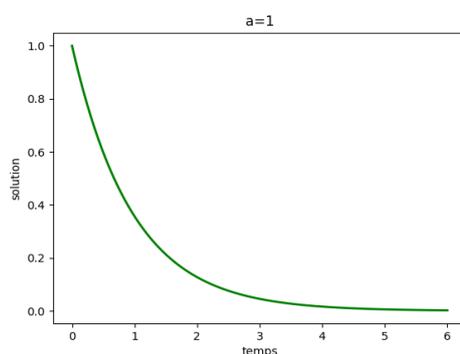


Fig. 4.1: Solution de  $y'(t) = -y(t)$  pour  $y_0 = 1$

Pour de nombreux problèmes cette quantité est positive et il est alors souhaitable que le schéma numérique donne aussi une solution positive. La figure 4.1 nous montre l'allure de la courbe pour  $a = 1$ . Nous étudions tout d'abord le schéma d'Euler explicite.

Le schéma d'Euler explicite s'écrit de la manière suivante :

$$\begin{cases} \forall n \in \llbracket 0, N-1 \rrbracket, & y_{n+1} = y_n - ah y_n, \\ y_0 = y(0). \end{cases}$$

On a donc en faisant un raisonnement par récurrence :

$$\forall n \in \llbracket 0, N \rrbracket, \quad y_n = y_0(1 - ah)^n.$$

On distingue deux cas :  $h \leq 1/a$  ou  $h > 1/a$  (le cas  $h = 1/a$  donne une solution nulle pour tout temps). On remarque que le schéma explicite ne donne pas, pour  $h > 1/a$ , une solution qualitativement bonne puisque  $y_n < 0$  pour  $n$  impair (alors que l'on souhaite une solution positive pour tout  $n$ ). La figure 4.2, confirme bien ce comportement insatisfaisant du schéma d'Euler explicite si  $h$  est trop grand.

Le schéma d'Euler implicite s'écrit de la manière suivante :

$$\begin{cases} \forall n \in \llbracket 0, N-1 \rrbracket, & y_{n+1} = y_n - ah y_{n+1}, \\ y_0 = y(0). \end{cases}$$

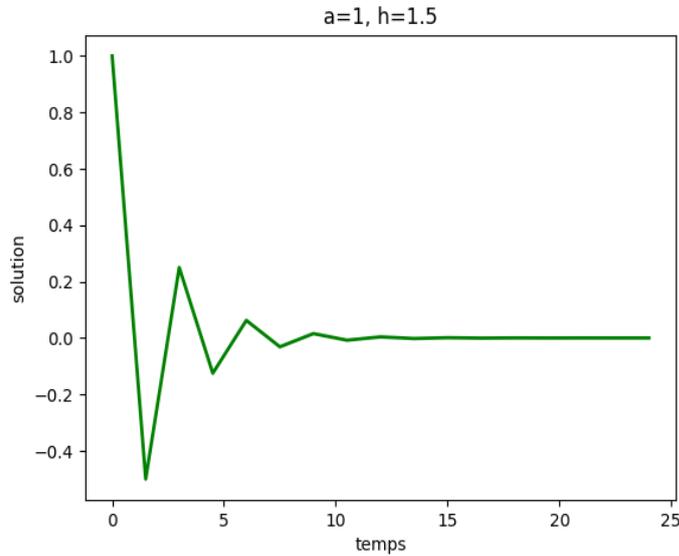


Fig. 4.2: Solution approchée avec la schéma explicite pour  $y'(t) = -y(t)$ ,  $y_0 = 1$ ,  $h = 1.5$

On a donc en faisant un raisonnement par récurrence

$$y_{n+1} = \frac{y_n}{1 + ah} \text{ et donc } y_n = \frac{y_0}{(1 + ah)^n}.$$

On constate que quelque soit  $h > 0$ , la suite  $(y_n)_{n \in \mathbb{N}}$  ne prend que des valeurs positives comme la suite  $(y(t_n))_{n \in \mathbb{N}}$ . Cet exemple montre l'intérêt des schémas implicites, car ici lorsque  $a$  est grand le schéma d'Euler explicite doit avoir un pas  $h$  inférieur à  $1/a$  pour avoir une solution approchée positive, alors que pour Euler implicite peu importe la valeur de  $h$  ou de  $a$ .

On ne peut pas toutefois déduire de cet exemple que le schéma implicite est meilleur que le schéma explicite. En effet, considérons maintenant le problème

$$\begin{cases} y'(t) = ay(t), & t > 0, \\ y(0) = 1, \end{cases} \quad (4.23)$$

avec  $a > 0$ , dont la solution est  $y(t) = e^{at}$ , pour tout  $t > 0$ .

Le schéma d'Euler implicite donne ici  $y_n = 1/(1 - ah)^n$  alors que le schéma d'Euler explicite donne  $y_n = (1 + ah)^n$  (pour tout  $n$ ).

Dans ce cas, c'est le schéma d'Euler implicite qui donne des valeurs négatives si  $h > 1/a$ .

### 4.3.2 Deuxième exemple : une équation non linéaire

Considérons le problème suivant

$$\begin{cases} y'(t) = -\sqrt{y(t)}, \\ y(0) = 1. \end{cases} \quad (4.24)$$

La solution est  $y(t) = \left(1 - \frac{t}{2}\right)^2$  pour  $t \leq 1/2$  puis  $y(t) = 0$  pour  $t > 1/2$ .

Le schéma d'Euler explicite s'écrit, pour  $y_n \geq 0$ .

$$y_{n+1} = y_n - h\sqrt{y_n}.$$

Ce schéma ne donnera pas une solution pour tout  $n$  (car cette équation donne  $y_{n+1} < 0$  si  $0 < y_n < h^2$ ). Le schéma d'Euler implicite s'écrit

$$y_{n+1} = y_n - h\sqrt{y_{n+1}}, \text{ avec } y_{n+1} > 0.$$

Ceci donne  $y_{n+1} = z_{n+1}^2$  avec  $z_{n+1} = (1/2)(-h + \sqrt{h^2 + 4y_n})$ . Noter que  $z_{n+1} > 0$  si  $y_n > 0$ . Dans la figure 4.3, La solution exacte est en bleu, la solution donnée par le schéma d'Euler implicite est en vert et la solution donnée par le schéma d'Euler explicite est en rouge, cette dernière devient négative pour  $t = 1.8$  lorsque  $h = 3/10$ .

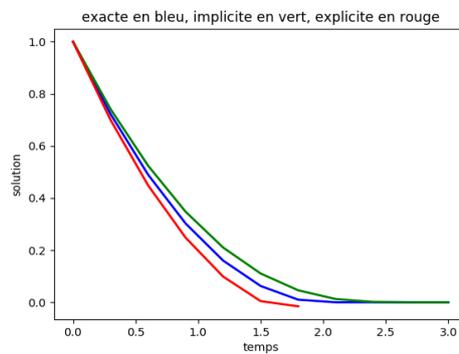


Fig. 4.3:  $y'(t) = -\sqrt{y(t)}$ ,  $h = 3/10$ , la solution explicite est négative pour  $t = 1.8$

## 4.4 Détermination numérique de l'ordre d'un schéma

Le TP1 illustre les différentes propriétés des schémas d'Euler, RK2 et RK4. Dans ce TP, il est demandé de calculer l'ordre du schéma. Si le schéma est d'ordre  $k$ , le maximum des erreurs de discrétisation  $e$  est tel qu'il existe  $C_1$  et  $C_2 \in \mathbb{R}_+$  tels que  $C_1 h^k \leq e \leq C_2 h^k$ . On s'attend donc à ce que, lorsque l'on divise le pas de discrétisation par 2, on divise l'erreur à peu près par  $2^k$ . Si le schéma est d'ordre 1, l'erreur est divisée à peu près par 2; si le schéma est d'ordre 2, l'erreur est divisée à peu près par 4; si le schéma est d'ordre 4, l'erreur est divisée à peu près par 16. . .

On peut aussi tracer l'erreur en fonction du pas de discrétisation dans un graphique log-log. En effet, si  $e \approx C h^k$ , on a  $\ln e \approx \ln C + k \ln h$  et donc l'ordre  $k$  est la pente de la droite ainsi obtenue (si le schéma est d'ordre  $k$ ).

## Cours 5. Équations différentielles linéaires du second ordre

Comme leur nom l'indique, les équations différentielles du second ordre font intervenir la dérivée seconde de la fonction recherchée.

**Définition 9** (Équation différentielle linéaire du second ordre). *Une équation différentielle linéaire du second ordre s'écrit :*

$$y''(t) + a(t)y'(t) + b(t)y(t) = f(t), \quad t \in I \quad (5.1)$$

où  $I$  est un intervalle ouvert de  $\mathbb{R}$ , (par exemple  $I = ]A, B[$ ), et les fonctions  $f$ ,  $a$  et  $b$ , sont connues et continues de  $\mathbb{R}$  dans  $\mathbb{R}$ . On cherche alors les fonctions  $y$  de classe  $C^2$  qui vérifient (5.1).

Le problème de Cauchy associé fait intervenir une condition initiale qui porte sur le couple  $(y(t_0), y'(t_0))$  en un point  $t_0$  donné; il s'écrit :

$$y''(t) + a(t)y'(t) + b(t)y(t) = f(t), \quad t \in ]t_0, T[ \quad (5.2a)$$

$$y(t_0) = y_0 \quad (5.2b)$$

$$y'(t_0) = z_0 \quad (5.2c)$$

où  $t_0, T, f, a, b, y_0$  et  $z_0$  sont donnés,  $t_0 < T \leq +\infty$ ,  $f, a, b$  sont continues de  $]t_0, T[$  à valeurs dans  $\mathbb{R}$ ,  $y_0, z_0 \in \mathbb{R}$ . On cherche  $y$  de classe  $C^2$  sur  $]t_0, T[$  (à valeurs dans  $\mathbb{R}$ ), continue ainsi que sa dérivée sur  $]t_0, T[$  et qui vérifie (5.2)

**Théorème 5.** *Soient  $t_0, T, a, b, f, y_0, z_0$  donnés, vérifiant les hypothèses ci dessus. alors il existe une et une seule fonction  $y$  solution de (5.2)*

Ce théorème est une conséquence du résultat d'existence et d'unicité pour les systèmes différentiels, vus ultérieurement. En effet le problème (5.2) peut s'écrire sous la forme d'un système différentiel d'ordre un par l'introduction d'une nouvelle variable  $z = y'$ . Le problème (5.2) revient alors à résoudre le problème

suivant :

$$\begin{aligned}y' &= z, \\z'(t) + a(t)z(t) + b(t)y(t) &= f(t), \quad t \in ]t_0, T[ = I, \\y(t_0) &= y_0, \\z(t_0) &= z_0.\end{aligned}$$

## 5.1 Équation homogène du second ordre

**Théorème 6.** *L'équation homogène associée à l'EDO linéaire du second ordre (5.1) consiste à prendre un second membre nul (c'est-à-dire  $f(t) = 0$  pour tout  $t$ ); elle s'écrit donc*

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0, \quad t \in I \quad (5.3)$$

*Sous les hypothèses de la définition 9, l'ensemble des solutions  $y \in C^2(I, \mathbb{R})$  de l'EDO homogène (5.3) est un espace vectoriel (réel) de dimension 2.*

*Démonstration.* Supposons  $y_1, y_2$  solutions de (5.3) alors  $y_1 + y_2$  est encore solution de (5.3). Si  $y$  est solution de (5.3) alors, pour  $\alpha \in \mathbb{R}$ ,  $\alpha y$  est encore solution de (5.3). L'ensemble des solutions de (5.3) est donc un espace vectoriel. Montrons maintenant que la dimension de l'ensemble des solutions de (5.1) est égale à 2.

On choisit un point  $t_0$  de  $I$ . Une conséquence du théorème 5 est que pour tout  $y_0$  et  $z_0$  dans  $\mathbb{R}$ , il existe une et une seule solution de (5.2). Soit :

- $y_1$  la solution de (5.2) avec  $f \equiv 0, y_0 = 0, z_0 = 1$ ,
- $y_2$  la solution de (5.2) avec  $f \equiv 0, y_0 = 1, z_0 = 0$ .

La solution correspondant à  $y_0 = \alpha$  et  $z_0 = \beta$  est alors  $\alpha y_2 + \beta y_1$ . La dimension de l'ensemble des solutions de (5.3) est donc égale à 2 et  $\{y_1, y_2\}$  est une base de l'ensemble des solutions.  $\square$

Nous allons nous intéresser maintenant aux 3 questions suivantes :

1) Comment calculer l'ensemble des solutions de (5.3)?

Nous allons pour cela chercher deux solutions linéairement indépendantes "simples".

2) Comment calculer l'ensemble des solutions de (5.1)?

Une solution possible est de trouver une solution particulière  $y_p$  et ensuite trouver la solution générale à l'aide de la solution  $y_h$  de l'équation homogène (5.3) :  $y = y_h + y_p$ .

3) Comment calculer les solutions de (5.2)?

Pour répondre à cette question, on pose  $y = \alpha y_1 + \beta y_2 + y_p$  (solution générale de (5.1)) et déterminer les coefficients  $\alpha$  et  $\beta$  grâce aux conditions initiales, en posant  $y(t_0) = y_0$  et  $y'(t_0) = z_0$ .

On répond dans un premier temps à ces questions dans le cas des solutions de l'équation homogène à coefficients constants. On considère  $a$  et  $b$  des fonctions constantes dans (5.3) (et donc  $a, b \in \mathbb{R}$ ) et  $I = \mathbb{R}$ . On cherche  $y$  sous la forme  $y(t) = e^{rt}$ . on a donc :

$$y(t) = e^{rt}, \quad y'(t) = r e^{rt}, \quad y''(t) = r^2 e^{rt}.$$

En réinjectant  $y$  dans (5.3) on a :

$$r^2 e^{rt} + ar e^{rt} + be^{rt} = 0. \quad (5.4)$$

L'équation (5.4) suggère alors l'équation dite caractéristique :

$$r^2 + ar + b = 0 \quad (5.5)$$

dont le discriminant s'écrit  $\Delta = a^2 - 4b$ . On remarque alors que  $r$  est solution de (5.5) si et seulement si  $e^{rt}$  est solution de (5.3). Distinguons 3 cas :

- 1) l'équation caractéristique (5.5) a deux racines réelles distinctes,
- 2) l'équation caractéristique (5.5) a deux racines complexes (conjuguées)
- 3) l'équation caractéristique (5.5) a une racine double

### 5.1.1 Équation à coefficients constants, cas de deux racines réelles

Dans ce cas,  $\Delta = a^2 - 4b > 0$ , et les deux racines réelles sont  $r_1 = \frac{-a - \sqrt{\Delta}}{2}$  et  $r_2 = \frac{-a + \sqrt{\Delta}}{2}$  (noter que  $r_1 - r_2 < 0$ ). Les fonctions  $y_1$  et  $y_2$  définies par

$$y_1(t) = e^{r_1 t}, \quad y_2(t) = e^{r_2 t}, \quad t \in \mathbb{R}$$

sont alors solutions de (5.3). Montrons que  $y_1$  et  $y_2$  sont linéairement indépendantes. En effet, soient  $\alpha, \beta \in \mathbb{R}$  tels que  $\alpha y_1 + \beta y_2 = 0$ . On a alors

$$\alpha e^{r_1 t} + \beta e^{r_2 t} = [\alpha e^{(r_1 - r_2)t} + \beta] e^{r_2 t} = 0 \text{ pour tout } t \in \mathbb{R}.$$

Comme  $r_1 - r_2 < 0$ , on a  $\lim_{t \rightarrow +\infty} \alpha e^{(r_1 - r_2)t} + \beta = \beta = 0$ . On a donc aussi  $\alpha = 0$  ce qui montre que les solutions  $y_1$  et  $y_2$  sont bien linéairement indépendantes. L'ensemble des solutions de (5.3) est engendré par  $y_1$  et  $y_2$ , et la solution générale de (5.3) est

$$y_h(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t), \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

**Exemple 11.** *Considérons l'équation suivante :*

$$y''(t) - 3y'(t) + 2y(t) = 0 \quad (5.6)$$

L'équation caractéristique s'écrit :

$$r^2 - 3r + 2 = 0 \quad (5.7)$$

On a donc  $\Delta = 1$ , d'où  $r_1 = 1$  et  $r_2 = 2$ ; la solution générale de (5.6) est de la forme :

$$y(t) = \alpha e^t + \beta e^{2t}, \quad \alpha, \beta \in \mathbb{R} \quad (5.8)$$

### 5.1.2 Équation à coefficients constants, cas de deux racines complexes

Dans ce cas  $\Delta < 0$  et en posant  $\lambda = \frac{-a}{2}$  et  $\mu = \frac{\sqrt{-\Delta}}{2}$  les racines sont  $r_1 = \lambda + i\mu$  et  $r_2 = \lambda - i\mu$  et sont bien complexes conjuguées. Même si on est intéressé (à terme) par des solutions réelles, on commence par remarquer qu'il existe des solutions à valeurs complexes sous la forme suivante :

$$\begin{aligned} y_c(t) &= \alpha e^{r_1 t} + \beta e^{r_2 t} = \alpha e^{\lambda t} e^{i\mu t} + \beta e^{\lambda t} e^{-i\mu t} \\ &= \alpha e^{\lambda t} (\cos(\mu t) + i \sin(\mu t)) + \beta e^{\lambda t} (\cos(\mu t) - i \sin(\mu t)), \text{ pour tout } t \in \mathbb{R}, \end{aligned} \quad (5.9)$$

avec  $\alpha, \beta \in \mathbb{C}$ .

Grâce à la linéarité de l'équation (5.3), les parties réelle et imaginaire de  $y_c$  sont aussi des solutions de (5.3). En prenant les parties réelle et imaginaire de  $y_c$  avec  $\alpha = 1$  et  $\beta = 0$ , on obtient donc les solutions suivantes de (5.3), qui sont à valeurs réelles :

$$y_1 : t \mapsto e^{\lambda t} \cos(\mu t) \text{ et } y_2 : t \mapsto e^{\lambda t} \sin(\mu t).$$

Comme les fonctions  $y_1$  et  $y_2$  sont linéairement indépendantes, et que l'ensemble des solutions de (5.3) est un espace vectoriel de dimension 2, la solution générale de (5.3) est :

$$y : t \mapsto y(t) = \gamma y_1(t) + \delta y_2(t), \quad \gamma, \delta \in \mathbb{R}. \quad (5.10)$$

**Exemple 12** (Cas de deux racines complexes : exemple du pendule pesant). *Considérons le système suivant :*

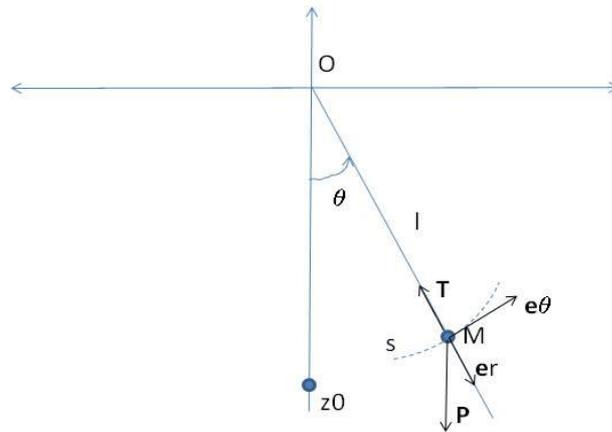


Fig. 5.1: Système du pendule pesant

Nous allons considérer l'équation du pendule linéarisée. On suppose que  $g/l = 1$  où  $g$  est l'intensité de la gravité et  $l$  la longueur du pendule), et on approche  $\sin \theta$  par  $\theta$  en considérant que l'angle  $\theta$  est petit. Le mouvement du pendule est alors décrit par l'équation (où on a posé  $y = \theta$ ) :

$$y''(t) + y(t) = 0 \quad (5.11)$$

L'équation caractéristique s'écrit :  $r^2 + 1 = 0$ , et les racines  $r_1 = i$  et  $r_2 = -i$  sont imaginaires pures. Avec les notations ci-dessus  $\lambda = 0$  et  $\mu = 1$ . On a vu plus haut que la solution générale dans  $\mathbb{C}$  est de la forme

$$y(t) = \alpha e^{r_1 t} + \beta e^{r_2 t}, \quad \alpha, \beta \in \mathbb{C}$$

et la solution générale (dans  $\mathbb{R}$ ) est donc de la forme

$$y(t) = \gamma \cos t + \delta \sin t, \quad \gamma, \delta \in \mathbb{R}.$$

### 5.1.3 Équation à coefficients constants, cas d'une racine double

On a donc  $\Delta = a^2 - 4b = 0$ , d'où  $r = -\frac{a}{2}$ . Une base de l'ensemble des solutions est alors donnée par les fonctions  $y_1$  et  $y_2$  définies par

$$y_1(t) = e^{rt}, \quad y_2(t) = te^{rt}.$$

La solution générale de (5.3) est de la forme

$$\alpha e^{rt} + t\beta e^{rt}, \quad \alpha, \beta \in \mathbb{R} \quad (5.12)$$

On peut se demander comment on trouve cette fonction  $y_2$ . Un moyen de trouver  $y_2$  à partir de  $y_1$  est de poser  $y_2(t) = z(t)y_1(t)$ . Comme  $y_1(t) = e^{rt}$ , ceci donne

$$y_2'(t) = z'(t)e^{rt} + rz(t)e^{rt}; \quad y_2''(t) = z''(t)e^{rt} + rz'(t)e^{rt} + rz'(t)e^{rt} + r^2z(t)e^{rt}.$$

L'équation sur  $y_2$  qui est  $y_2'' + ay_2' + by_2 = 0$  devient,

$$z''(t)e^{rt} + 2rz'(t)e^{rt} + r^2z(t)e^{rt} + a(z'(t)e^{rt} + rz(t)e^{rt}) + bz(t)e^{rt} = 0,$$

ce qui est équivalent à

$$z''(t) + (2r + a)z'(t) + (r^2 + ar + b)z(t) = 0,$$

ou encore, comme  $r^2 + ar + b = 0$ ,

$$z''(t) + (2r + a)z'(t) = 0,$$

On obtient ainsi une équation de premier ordre sur  $z'$ . On appelle d'ailleurs cette méthode "méthode de réduction d'ordre" car c'est une équation du premier ordre sur  $z'$ . Elle fonctionne dès que  $y_1$  est une solution de l'équation différentielle considérée et que l'équation est linéaire, quelque soit son ordre.

Ici comme  $r$  est racine double, on a  $2r + a = 0$  et donc l'équation sur  $z$  est  $z'' = 0$ .

Une solution possible est donc  $z(t) = t$ . Ce qui donne  $y_2 = te^t$  (qui est bien une solution linéairement indépendante de  $y_1$ ).

**Exemple 13.** Soit

$$y''(t) - 2y'(t) + y(t) = 0 \quad (5.13)$$

Nous avons alors  $r^2 - 2r + 1 = 0$ , on a une racine double  $r = 1$ , d'où la solution générale

$$y(t) = \alpha e^t + \beta te^t, \quad \alpha, \beta \in \mathbb{R} \quad (5.14)$$

## 5.2 Équation non homogène

On se limite ici au cas des coefficients constants, *i.e.* on considère  $a, b$  constants dans (5.1). On détermine la solution générale de la même manière que pour les EDO du premier ordre. Il faut d'abord trouver la solution de l'équation homogène associée à (5.1) qui est de la forme  $t \mapsto y_h(t) = \alpha y_1(t) + \beta y_2(t)$  où  $\{y_1, y_2\}$  est une base de l'espace vectoriel des solutions de l'équation homogène, puis trouver une solution particulière. Plusieurs méthodes sont possibles pour cela.

**Méthode 1 : recherche d'une solution particulière** On "devine" une solution particulière  $y_p$ , et on exprime la solution générale de (5.1) sous la forme :

$$y(t) = \underbrace{\alpha y_1(t) + \beta y_2(t)}_{\text{solution de l'équation homogène (5.3)}} + y_p(t) \quad (5.15)$$

**Méthode 2 : Méthode de la variation des constantes** On cherche une solution particulière ou, plus directement, l'ensemble des solutions de l'équation non homogène, sous la forme  $\alpha_1(t)y_1(t) + \alpha_2(t)y_2(t)$  (où les fonctions  $y_1$  et  $y_2$  engendrent l'ensemble des solutions de l'équation homogène). Cette méthode est décrite dans l'exercice 3 du partiel de 2019. (exercice 3).

Nous allons développer la méthode 1 avec des équations du second ordre dont les équations homogènes associées ont été étudiées dans les exemples 11, 12 et 13.

**Exemple 14.** On considère ici l'équation suivante :

$$y''(t) - 3y'(t) + 2y(t) = e^{-t} \quad (5.16)$$

dont l'équation homogène associée est l'équation (5.6) qui a été vue dans l'exemple 11. On a vu que la solution générale de l'équation homogène associée (5.6) est de la forme

$$y(t) = \alpha e^t + \beta e^{2t}, \quad \alpha, \beta \in \mathbb{R}.$$

On cherche alors une solution particulière sous la forme  $y_p(t) = \gamma e^{-t}$ <sup>1</sup>. On a donc :

$$y_p'(t) = -\gamma e^{-t}; \quad y_p''(t) = \gamma e^{-t}$$

En réinjectant  $y_p$  dans (5.16), il vient

$$y_p''(t) - 3y_p'(t) + 2y_p(t) = \gamma e^{-t} + 3\gamma e^{-t} + 2\gamma e^{-t} \quad (5.17)$$

$$= 6\gamma e^{-t} = e^{-t}. \quad (5.18)$$

Ce qui donne comme solution particulière  $y_p(t) = \frac{1}{6}e^{-t}$ . La solution générale de (5.16) est donc

$$y(t) = \alpha e^t + \beta e^{2t} + \frac{1}{6}e^{-t}, \quad \alpha, \beta \in \mathbb{R}.$$

**Remarque 5** (Cas d'un second membre appartenant à l'e.v. des solutions de l'équation homogène associée). En prenant dans l'exemple précédent comme second membre  $f(t) = e^t$ , une solution particulière serait cherchée sous la forme  $y_p(t) = \gamma t e^t$ .

**Exemple 15** (Le pendule linéarisé avec second membre).

Reprenons l'exemple 12 mais maintenant avec un terme de forçage qui donne un second membre non nul; on obtient l'équation suivante :

$$y''(t) + y(t) = \sin t \quad (5.19)$$

La solution générale de l'équation homogène associée est

$$y_h(t) = \alpha \sin t + \beta \cos t, \quad \alpha, \beta \in \mathbb{R}.$$

On cherche alors la solution particulière sous la forme  $y_p(t) = \gamma t \sin t + \delta t \cos t$ .

On a donc :

$$y_p'(t) = (\gamma - \delta t) \sin t + (\delta + \gamma t) \cos t, \quad y_p''(t) = (2\gamma - \delta t) \cos t + (-\gamma t - 2\delta) \sin t$$

Alors, en réinjectant  $y_p$  dans (5.19) :

$$y_p''(t) + y_p(t) = \gamma t \sin t + \delta t \cos t + (2\gamma - \delta t) \cos t + (-\gamma t - 2\delta) \sin t \quad (5.20)$$

$$= -2\delta \sin t + 2\gamma \cos t = \sin t \quad (5.21)$$

1. Nous cherchons une solution sous cette forme  $\gamma e^{-t}$ , car le second membre est aussi de cette forme.

On en déduit alors que  $\delta = -\frac{1}{2}$  et  $\gamma = 0$  d'où  $y_p(t) = -\frac{1}{2}t \cos t$ . La solution générale de l'équation (5.19) est donc

$$t \mapsto y(t) = y_h(t) + y_p(t) = \alpha \sin t + \beta \cos t - \frac{1}{2}t \cos t.$$

**Remarque 6** (Résonance). En raison du terme de forçage, la solution du problème de l'équation (5.19) est non bornée. C'est ce qu'on appelle un cas de résonance. Sans l'approximation des petits angles, l'équation du pendule est en réalité non linéaire :

$$y''(t) + \sin(y(t)) = f(t). \quad (5.22)$$

Pour cette équation non linéaire, on peut montrer, mais cela est difficile, qu'il existe  $\varepsilon > 0$  tel que  $|f(t)| \leq \varepsilon$  pour tout  $t$  implique que  $y$  est bornée.

Les problèmes de résonance sont à l'origine de l'effondrement de certaines structures, voir par exemple le cas du Tacoma Bridge :

[https://fr.wikipedia.org/wiki/Pont\\_du\\_détroit\\_de\\_Tacoma\\_\(1940\)](https://fr.wikipedia.org/wiki/Pont_du_détroit_de_Tacoma_(1940))

**Exemple 16.** Soit

$$y''(t) - 2y'(t) + y = \sin t \quad (5.23)$$

On a vu à l'exemple 12 que l'équation caractéristique à une racine double,  $r = 1$  et que la solution de l'équation homogène est de la forme.

$$y_h(t) = \alpha e^t + \beta t e^t \quad (5.24)$$

On cherche alors  $y_p$  sous la forme  $y_p(t) = \gamma \sin t + \delta \cos t$ . On a donc

$$y_p'(t) = \gamma \cos t - \delta \sin t; \quad y_p''(t) = -\gamma \sin t - \delta \cos t$$

On réinjecte  $y_p$  dans (5.23), on obtient alors :

$$\begin{aligned} y_p''(t) - 2y_p'(t) + y_p(t) &= -\gamma \sin t - \delta \cos t - 2\gamma \cos t + 2\delta \sin t + \gamma \sin t + \delta \cos t \\ &= -2\gamma \cos t + 2\delta \sin t \end{aligned}$$

On en déduit alors que  $\gamma = 0$  et  $\delta = \frac{1}{2}$ . La solution générale de l'équation (5.23) est donc

$$t \mapsto y(t) = y_h(t) + y_p(t) = \alpha e^t + \beta t e^t + \frac{1}{2} \cos t.$$

## Cours 6. Systèmes différentiels linéaires

On aborde ici les systèmes d'équations différentielles, qui font intervenir le même nombre d'équations que d'inconnues, et qui sont, de plus, linéaires (ou plus exactement, affines, dans le cas non homogène).

### 6.1 Définitions

On note  $n \geq 1$  le nombre d'équations et d'inconnues du système différentiel linéaire, qui s'écrit

$$x'_i(t) = a_{i,1}(t)x_1(t) + \dots + a_{i,n}(t)x_n(t) + g_i(t), \quad t \in I, \quad 1 \leq i \leq n \quad (6.1)$$

où pour tout  $i, j \in \llbracket 1, n \rrbracket$ ,  $a_{i,j}, g_i \in \mathcal{C}(I, \mathbb{R})$  et  $I$  un ouvert de  $\mathbb{R}$ . Les inconnues sont les fonctions  $x_i \in C^1(I, \mathbb{R})$ .

Le problème de Cauchy associé (que nous appellerons “problème de Cauchy classique”) consiste à prendre  $I = ]t_0, T[$  ( $T > t_0$ ) et à ajouter à l'équation (6.1) que  $x$  est continue en  $t_0$  et prend une valeur donnée en  $t_0$ . C'est donc le problème :

$$x'_i(t) = a_{i,1}(t)x_1(t) + \dots + a_{i,n}(t)x_n(t) + g_i(t), \quad 1 \leq i \leq n, \quad t \in ]t_0, T[ \quad (6.2a)$$

$$x_i(t_0) = x_i^{(0)}, \quad 1 \leq i \leq n. \quad (6.2b)$$

Si  $x$  est solution du problème 6.2, du fait que  $x$  est de classe  $C^1$  sur  $]t_0, T[$  et continue en  $t_0$ , on peut déduire de (6.1) que  $x$  est dérivable à droite en  $t_0$  et vérifie (6.1) avec cette dérivée à droite (cette démonstration est laissée à titre d'exercice).

En inversant le sens du temps on peut également considérer le problème de Cauchy “à gauche” consistant à prendre  $I = ]T, t_0[$  ( $T < t_0$ ) et à ajouter (à l'équation (6.1)) que  $x$  est continue en  $t_0$  et prend une valeur donnée en  $t_0$ . Ici encore de (6.1) on déduit alors que  $x$  est dérivable à gauche en  $t_0$  et vérifie (6.1) avec cette dérivée à gauche.

La résolution de ces deux problèmes de Cauchy (problème de Cauchy classique et problème de Cauchy à gauche) est alors équivalente à la résolution du problème de Cauchy général qui consiste à étudier le système (6.1) avec des conditions initiales sur les fonctions inconnues  $x_i$  en un point  $t_0 \in I$  donné :

$$x'_i(t) = a_{i,1}(t)x_1(t) + \dots + a_{i,n}(t)x_n(t) + g_i(t), \quad 1 \leq i \leq n, \quad t \in I \quad (6.3a)$$

$$x_i(t_0) = x_i^{(0)}, \quad 1 \leq i \leq n. \quad (6.3b)$$

Le problème de Cauchy général (6.3) peut s'écrire sous forme matricielle :

$$X'(t) = A(t)X(t) + G(t), \quad t \in I \quad (6.4a)$$

$$X(t_0) = X^{(0)} \quad (6.4b)$$

où  $A \in \mathcal{M}_n(\mathbb{R})$ <sup>1</sup> est définie par .

$$A(t) = \begin{bmatrix} a_{1,1}(t) & \dots & a_{1,n}(t) \\ a_{2,1}(t) & \dots & a_{2,n}(t) \\ \vdots & \ddots & \vdots \\ a_{n,1}(t) & \dots & a_{n,n}(t) \end{bmatrix},$$

et  $X'(t)$ ,  $X(t)$ ,  $G(t)$  et  $X^{(0)}$  sont des matrices colonne à  $n$  lignes, que l'on identifie à des vecteurs de  $\mathbb{R}^n$ , définis par

$$X'(t) = \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix}, X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, G(t) = \begin{bmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{bmatrix}, \text{ et } X^{(0)} = \begin{bmatrix} x_1^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}.$$

**Exemple 17** (Équations du second ordre). Une application directe des systèmes linéaires du premier ordre est la résolution d'une équation différentielle linéaire du second ordre de la forme

$$x''(t) + a(t)x'(t) + b(t)x(t) = c(t), t \in I \quad (6.5)$$

où  $a, b$  et  $c$  sont des fonctions continues de  $I$  (intervalle ouvert de  $\mathbb{R}$ ) dans  $\mathbb{R}$ . Soit  $x$  une fonction (de  $I$  dans  $\mathbb{R}$ ) de classe  $C^2$  solution de (6.5). On pose :

$$\begin{aligned} x_1(t) &= x(t) \\ x_2(t) &= x'(t). \end{aligned}$$

Les fonctions  $x_1, x_2$  sont de classe  $C^1$  et on a :

$$\begin{aligned} x'_1(t) &= x'(t) = x_2(t) \\ x'_2(t) &= x''(t) = -ax'(t) - bx(t) + c(t) = -ax_2(t) - bx_1(t) + c(t). \end{aligned}$$

On peut alors écrire l'EDO (6.5) sous la forme d'un système différentiel linéaire de deux équations à deux inconnues :

$$X'(t) = AX(t) + B(t) \text{ avec } X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, A = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \text{ et } B(t) = \begin{bmatrix} 0 \\ c(t) \end{bmatrix}$$

Notons que de la manière plus générale, toutes les équations différentielles linéaires d'ordre  $n$  peuvent se mettre sous la forme d'un système différentiel linéaire du premier ordre, de  $n$  équations à  $n$  inconnues.

## 6.2 Systèmes homogènes

Avec les notations précédentes, le système homogène associé au système différentiel (6.1) consiste à prendre  $G(t) = 0$ . Il s'écrit alors

$$X'(t) = A(t)X(t), t \in I. \quad (6.6)$$

Soit  $E_n$  l'ensemble des solutions de (6.6) :

$$E_n = \{X \in \mathcal{C}^1(I, \mathbb{R}^n), X \text{ solution de (6.6)}\} \quad (6.7)$$

1.  $\mathcal{M}_n(\mathbb{R})$  désigne l'ensemble des matrices carrées d'ordre  $n \times n$  à coefficients réels

Comme dans le cas des équations ( $n = 1$ ), l'ensemble  $E_n$  est un espace vectoriel sur  $\mathbb{R}$ . En effet, si  $X, Y \in E_n$  alors  $X + Y \in E_n$  car

$$\begin{aligned} X'(t) + Y'(t) &= A(t)X(t) + A(t)Y(t) \\ &= A(t)(X(t) + Y(t)) \end{aligned}$$

D'autre part, si  $X \in E_n$  et  $\alpha \in \mathbb{R}$  alors  $\alpha X \in E_n$  car

$$(\alpha X)'(t) = \alpha X'(t) = A(t)(\alpha X(t))$$

Nous avons vu dans les chapitres précédents que dans les cas  $n = 1$  et  $n = 2$ , cet espace est de dimension respective 1 et 2. En dimension  $n$ , est-il de dimension  $n$ ? La réponse découle d'un théorème d'existence et d'unicité pour le problème de Cauchy (général) associé que nous admettrons pour le moment (voir cours 9 pour la démonstration).

**Théorème 7** (Existence et unicité). Soient  $I$  un ouvert de  $\mathbb{R}$ ,  $t_0 \in I$ ,  $A \in C(I, \mathcal{M}_n(\mathbb{R}))$  (c'est-à-dire que les composantes de  $A$  sont continues de  $I$  à valeurs dans  $\mathbb{R}$ ) et  $X^0 \in \mathbb{R}^n$  donnés. On considère le problème de Cauchy (général) :

$$X'(t) = A(t)X(t), \quad t \in I \tag{6.8a}$$

$$X(t_0) = X^0. \tag{6.8b}$$

Alors il existe une et une seule solution au problème de Cauchy (6.8)

**Remarque 7.** Le théorème 7 reste vrai si on remplace  $A(t)X(t)$  par  $A(t)X(t) + G(t)$  avec  $G \in C(I, \mathbb{R})$ .

**Corollaire 1** (Dimension de  $E_n$ ). L'espace vectoriel  $E_n$  des solutions du système différentiel (6.6) est de dimension  $n$ .

*Démonstration.* Soit  $t_0 \in I$ , et soit  $X^{(0)}$  le vecteur

$$X^{(0)} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Le théorème 7 entraîne qu'il existe une fonction  $X^{(1)}$  solution de (6.8). Plus généralement, pour tout  $1 \leq i \leq n$ , il existe une fonction  $X^{(i)}$  solution de (6.8) pour la donnée initiale

$$X^{(0)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

où le 1 est la  $i$ -ème composante du vecteur  $X^{(0)}$ . Les fonctions  $X^{(1)}, \dots, X^{(n)}$  sont  $n$  solutions de (6.6). Montrons que la famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est une base de  $E_n$  (et donc que  $\dim(E_n) = n$ ). Considérons  $X$  solution de (6.6),  $X$  est donc solution de (6.8) avec

$$X^{(0)} = X(t_0) = \begin{bmatrix} X_1(t_0) \\ \vdots \\ X_n(t_0) \end{bmatrix}$$

On pose maintenant  $Y = X - \sum_{i=1}^n X_i(t_0)X^{(i)}$ . La fonction  $Y$  est encore une solution de (6.6) et elle vérifie  $Y(t_0) = 0$ . La partie unicité du théorème 7 donne alors que  $Y$  est la fonction nulle et donc  $X = \sum_{i=1}^n X_i(t_0)X^{(i)}$ . Ceci montre que la famille de fonctions  $\{X^{(1)}, \dots, X^{(n)}\}$  engendre  $E_n$ .

Montrons maintenant que  $\{X^{(1)}, \dots, X^{(n)}\}$  est une famille libre.

Soient :  $\alpha_i \in \mathbb{R}$ ,  $i \in \llbracket 1, n \rrbracket$ , tels que  $\alpha_1 X^{(1)} + \dots + \alpha_n X^{(n)} = 0$ . On a alors :

$$\forall t \in I, \sum_{i=1}^n \alpha_i X^{(i)}(t) = 0.$$

Cette fonction est donc nulle en particulier pour  $t_0 = 0$  et donc :

$$\sum_{i=1}^n \alpha_i X^{(i)}(t_0) = 0. \quad (6.9)$$

Or, on sait que :

$$X^{(i)}(t_0) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{-ème composante}$$

On déduit donc de (6.9) que  $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ . Ceci montre bien que la famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est une famille libre. La famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est donc une base et on a ainsi prouvé que  $\dim E_n = n$ .  $\square$

**Proposition 4** (CNS d'indépendance des solutions). Soient  $X^{(1)}, \dots, X^{(n)}$   $n$  solutions de (6.6), c'est-à-dire  $n$  éléments de  $E_n$ , et soit  $t_0 \in I$ . La famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est libre de l'espace vectoriel  $\mathcal{C}(I, \mathbb{R})$  si et seulement si la famille  $\{X^{(1)}(t_0), \dots, X^{(n)}(t_0)\}$  est une famille libre de l'espace vectoriel  $\mathbb{R}^n$ .

*Démonstration.* On suppose que  $\{X^{(1)}, \dots, X^{(n)}\}$  est une famille libre. Soient  $\alpha_1, \dots, \alpha_n$ , tels que

$$\sum_{i=1}^n \alpha_i X^{(i)}(t_0) = 0.$$

Comme  $E_n$  est un espace vectoriel,  $\sum_{i=1}^n \alpha_i X^{(i)}$  est solution de (6.6) et  $\sum_{i=1}^n \alpha_i X^{(i)}(t_0) = 0$ . La partie

unicité du théorème (7) donne alors que  $\sum_{i=1}^n \alpha_i X^{(i)} = 0$  (c'est-à-dire que la fonction  $\sum_{i=1}^n \alpha_i X^{(i)}$  est la

fonction nulle). Comme la famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est une famille libre, on en déduit que  $\alpha_i = 0$  pour tout  $i \in \{1, \dots, n\}$ , et donc que la famille (de vecteurs)  $\{X^{(1)}(t_0), \dots, X^{(n)}(t_0)\}$  est libre.

La réciproque est plus simple; en particulier elle n'utilise pas que les fonctions  $X^{(i)}$  sont solutions de (6.6) On suppose que la famille (de vecteurs)  $\{X^{(1)}(t_0), \dots, X^{(n)}(t_0)\}$  est libre. On veut montrer que la famille (de fonctions)  $\{X^{(1)}, \dots, X^{(n)}\}$  est libre. Soient  $\alpha_1, \dots, \alpha_n$ , tels que :  $\sum_{i=0}^n \alpha_i X^{(i)} = 0$ ; alors

$$\text{en particulier : } \sum_{i=0}^n \alpha_i X^{(i)}(t_0) = 0.$$

Comme la famille  $\{X^{(1)}(t_0), \dots, X^{(n)}(t_0)\}$  est libre, ceci implique que  $\alpha_i = 0$  pour tout  $i \in \{1, \dots, n\}$ , et donc que la famille (de fonctions)  $\{X^{(1)}, \dots, X^{(n)}\}$  est libre.  $\square$

### 6.3 Cas des systèmes différentiels linéaires autonomes

Calculons une base de  $E_n$  dans le cas où  $A$  ne dépend pas de  $t$ ; on dit dans ce cas que le système différentiel linéaire est autonome. Pour l'étude de ce cas, on va considérer  $I = \mathbb{R}$ . Le système s'écrit donc

$$X'(t) = AX(t), \quad t \in \mathbb{R} \quad (6.10)$$

avec  $n \geq 1$  et  $A \in \mathcal{M}_n(\mathbb{R})$ . On note  $a_{i,j}$  ( $i, j \in \llbracket 1, n \rrbracket$ ) les coefficients de  $A$ .

Rappelons quelques résultats fondamentaux d'algèbre linéaire. Soit  $A \in \mathcal{M}_n(\mathbb{R})$ .

- 1)  $[\exists X \neq 0 \mid AX = 0] \Leftrightarrow [\det A = 0] \Leftrightarrow A$  est non inversible.
- 2)  $\lambda \in \mathbb{R}$  ou  $\mathbb{C}$ ,  $\exists X \neq 0 \mid AX = \lambda X \Leftrightarrow \det(A - \lambda I_n) = 0$ ,  $I_n$  est la matrice identité de  $\mathcal{M}_n(\mathbb{R})$ .

Trois cas sont à distinguer pour déterminer une base de l'espace vectoriel  $E_n$  des solutions de (6.10).

- 1) La matrice  $A$  est diagonalisable dans  $\mathbb{R}$ . Dans ce cas, il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  et  $\varphi_1, \dots, \varphi_n \in \mathbb{R}^n$  tels que  $\{\varphi_1, \dots, \varphi_n\}$  est une base de  $\mathbb{R}^n$  et  $A\varphi_i = \lambda_i \varphi_i$  pour tout  $i$ .
- 2) La matrice  $A$  est diagonalisable dans  $\mathbb{C}$ . Dans ce cas, il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$  et  $\varphi_1, \dots, \varphi_n$  une base dans  $\mathbb{C}^n$ , tel que  $A\varphi_i = \lambda_i \varphi_i$ . Remarquons que si  $A\varphi = \lambda\varphi$  alors  $A\bar{\varphi} = \bar{\lambda}\bar{\varphi}$ .
- 3) La matrice  $A$  n'est pas diagonalisable.

#### 6.3.1 Cas d'une matrice diagonalisable dans $\mathbb{R}$

Soit  $\{\varphi_1, \dots, \varphi_n\}$  une base de  $\mathbb{R}^n$  formée de vecteurs propres de  $A$ . Pour  $1 \leq i \leq n$ ,  $A\varphi_i = \lambda_i \varphi_i$ ,  $\lambda_i \in \mathbb{R}$ . On va chercher la fonction  $X$  de  $\mathbb{R}$  dans  $\mathbb{R}^n$ , de classe  $C^1$ , solution du système différentiel (6.10) en décomposant  $X(t)$  (pour tout  $t \in \mathbb{R}$ ) dans la base  $\{\varphi_1, \dots, \varphi_n\}$ . On cherche donc des fonctions

$\alpha_i$  de  $\mathbb{R}$  dans  $\mathbb{R}$ , de classe  $C^1$ , telles que la fonction  $X$  définie par  $X(t) = \sum_{i=1}^n \alpha_i(t) \varphi_i$  soit solution de (6.10). Comme

$$X'(t) = \sum_{i=1}^n \alpha_i'(t) \varphi_i \quad \text{et}$$

$$AX(t) = A\left(\sum_{i=1}^n \alpha_i(t) \varphi_i\right) = \sum_{i=1}^n \alpha_i(t) A\varphi_i = \sum_{i=1}^n \alpha_i(t) \lambda_i \varphi_i.$$

Le système différentiel (6.10)  $X'(t) = AX(t)$  est alors équivalent à

$$\alpha_i'(t) = \lambda_i \alpha_i(t) \quad \text{pour tout } t \in \mathbb{R}, \quad \text{pour tout } i \in \llbracket 1, n \rrbracket.$$

Le système (6.10) est ainsi ramené à  $n$  équations découplées. La fonction  $X$  ainsi définie est solution de (6.10) si et seulement si il existe  $c_1, \dots, c_n \in \mathbb{R}$  tels que

$$\alpha_i(t) = c_i e^{\lambda_i t} \text{ pour tout } t \in \mathbb{R}, \text{ pour tout } i \in \llbracket 1, n \rrbracket.$$

On obtient ainsi un ensemble de solutions du système différentiel (6.10) sous la forme

$$X(t) = \sum_{i=1}^n c_i e^{\lambda_i t} \varphi_i, \text{ pour tout } t \in \mathbb{R},$$

avec  $c_1, \dots, c_n \in \mathbb{R}$ . Cet ensemble est de dimension  $n$ . En effet, en posant  $X^{(i)}(t) = e^{\lambda_i t} \varphi_i$  pour  $i \in \llbracket 1, n \rrbracket$ , la famille (de fonctions)  $\{X^{(1)}, \dots, X^{(n)}\}$  est une famille libre car les vecteurs  $\varphi_1, \dots, \varphi_n$  sont indépendants (c'est-à-dire que la famille  $\{\varphi_1, \dots, \varphi_n\}$  est libre) et que  $X^{(i)}(0) = \varphi_i$  (voir la proposition 4). La famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est donc une base de  $E_n$ .

### 6.3.2 Cas d'une matrice diagonalisable dans $\mathbb{C}$ (mais non dans $\mathbb{R}$ )

Supposons ici  $n = 2$  pour simplifier. Il existe  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ , tel que  $\lambda = \alpha + i\beta$ , où  $\beta \in \mathbb{R}^*$  et  $\alpha \in \mathbb{R}$ , et il existe  $\varphi \in \mathbb{C}^n$  tel que  $A\varphi = \lambda\varphi$ . Comme  $A$  est à valeurs réelles, on a aussi  $A\bar{\varphi} = \bar{\lambda}\bar{\varphi}$ . On pose alors  $\varphi = \psi + i\chi$ , avec  $\psi, \chi \in \mathbb{R}^n$ , de sorte que

$$A\varphi = \lambda\varphi \Leftrightarrow A(\psi + i\chi) = (\alpha + i\beta)(\psi + i\chi).$$

On peut remarquer que  $\chi \neq 0$  (car  $A\psi \in \mathbb{R}^n$  et  $\lambda\psi \notin \mathbb{R}^n$  car  $\beta \neq 0$ ). De même  $\psi \neq 0$  (car  $A\chi \in \mathbb{R}^n$  et  $\lambda\chi \notin \mathbb{R}^n$  car  $\beta \neq 0$ ). En fait, les vecteurs  $\psi, \chi$  sont même indépendants (si  $\chi = \gamma\psi$  avec  $\gamma \in \mathbb{R}$ , on obtient une contradiction, par exemple, en multipliant l'équation  $A\varphi = \lambda\varphi$  par  $(1 - i\gamma)$ ). On obtient donc une solution de (6.10) à valeurs complexes, notée  $X$ , en posant  $X(t) = e^{\lambda t} \varphi$ . Pour obtenir des solutions de (6.10) à valeurs réelles, on décompose  $X(t)$  en partie réelle et partie imaginaire :

$$\begin{aligned} X(t) = e^{\lambda t} \varphi &= e^{(\alpha + i\beta)t} \varphi \\ &= e^{\alpha t} e^{i\beta t} (\psi + i\chi) \\ &= e^{\alpha t} (\cos(\beta t) + i \sin(\beta t)) (\psi + i\chi) \\ &= e^{\alpha t} (\cos(\beta t)\psi - \sin(\beta t)\chi) + i e^{\alpha t} (\sin(\beta t)\psi + \cos(\beta t)\chi). \end{aligned}$$

On remarque maintenant que les parties réelle et imaginaire de  $X$  sont des solutions (à valeurs réelles) de (6.10). En effet,

$$\begin{aligned} X(t) &= \Re(X(t)) + i\Im(X(t)) \\ X'(t) &= \Re(X'(t)) + i\Im(X'(t)) \\ AX(t) &= A\Re(X(t)) + iA\Im(X(t)) \\ X'(t) = AX(t) &\Leftrightarrow (\Re(X(t)))' = A\Re(X(t)) \text{ et } (\Im(X(t)))' = A\Im(X(t)) \end{aligned}$$

On obtient donc deux solutions de (6.10) en posant :

$$\begin{cases} X^{(1)}(t) = e^{\alpha t} (\cos(\beta t)\psi - \sin(\beta t)\chi) \\ X^{(2)}(t) = e^{\alpha t} (\sin(\beta t)\psi + \cos(\beta t)\chi) \end{cases}$$

Les fonctions  $X^{(1)}$  et  $X^{(2)}$  sont linéairement indépendantes (c'est-à-dire que la famille  $\{X^{(1)}, X^{(2)}\}$  est libre) car les vecteurs  $\psi, \chi$  sont linéairement indépendants (et  $X^{(1)}(0) = \psi, X^{(2)}(0) = \chi$ ). On en déduit qu'elles engendrent  $E_2$  (car on sait que  $E_2$  est de dimension 2).

Le cas  $n > 2$  se traite de manière similaire en distinguant les valeurs propres réelles et complexes de  $A$ .

**Exemple 18** (Calcul dans le cas d'un système  $2 \times 2$ ). *Considérons le système (6.10) avec*

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

*Les valeurs propres de la matrices  $A$  sont  $\lambda = 1 \pm i$ , c'est-à-dire avec les notations précédentes  $\alpha = 1$  et  $\beta = 1$ . On a  $\lambda_1 = 1 + i$  et  $\lambda_2 = 1 - i$ . Un vecteur propre associé à  $\lambda_1$  est  $\varphi = \psi + i\chi$  avec*

$$\psi = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \chi = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

*(c'est à dire que  $A\varphi = \lambda_1\varphi$  avec  $\varphi = \psi + i\chi$ ). On obtient alors une base de  $E_2$  en posant*

$$\begin{cases} X^{(1)}(t) = e^t \left( \cos t \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \sin t \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \\ X^{(2)}(t) = e^t \left( \sin t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \cos t \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \end{cases}$$

*ce qui s'écrit aussi*

$$\begin{cases} X^{(1)}(t) = e^t \begin{bmatrix} -\sin t \\ \cos t \end{bmatrix} \\ X^{(2)}(t) = e^t \begin{bmatrix} \cos t \\ \sin t \end{bmatrix} \end{cases}$$

*Pour la matrice  $A$  définie ci-dessus, la solution générale de (6.10) est alors  $c_1 X^{(1)}(t) + c_2 X^{(2)}(t)$  avec  $c_1, c_2 \in \mathbb{R}$ .*

### 6.3.3 Cas d'une matrice non diagonalisable

Commençons par un exemple. Considérons le système suivant pour  $t \in \mathbb{R}$

$$\begin{cases} x_1'(t) = x_2(t) \\ x_2'(t) = -2x_2(t) - x_1(t) \end{cases}$$

En posant

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix},$$

le système s'écrit sous forme matricielle :

$$X'(t) = AX(t) \tag{6.11}$$

Le polynôme caractéristique de  $A$  est donc  $P_A(\lambda) = (\lambda + 1)^2$ . La valeur propre  $\lambda = -1$  est double. Mais le sous espace propre associé à  $\lambda$  est de dimension 1. Il est engendré par le vecteur propre  $\phi = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ . La matrice  $A$  n'est pas diagonalisable. Puisqu'on a exhibé un vecteur propre, on sait déjà qu'on a trouvé une solution de (6.11). C'est la fonction  $t \mapsto e^{-t}\phi$ . On la note  $X^{(1)}$  et on a, bien sûr,  $X^{(1)} \in E_2$ .

Pour trouver une deuxième solution (linéairement indépendante de la première), on peut ici (pour cet exemple) se ramener à une équation du deuxième ordre. On remarque que

$$x_1''(t) = x_2'(t) = -2x_2(t) - x_1(t) = -2x_1'(t) - x_1(t),$$

d'où

$$x_1''(t) + 2x_1'(t) + x_1(t) = 0. \quad (6.12)$$

Les solutions de (6.12) sont alors de la forme :

$$x_1(t) = \alpha e^{-t} + \beta t e^{-t}, \quad \alpha, \beta \text{ deux constantes réelles arbitraires}$$

On peut en déduire une valeur pour  $x_2$

$$x_2(t) = x_1'(t) = -\alpha e^{-t} - \beta t e^{-t} + \beta e^{-t}.$$

Finalement

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \alpha e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta \begin{bmatrix} t e^{-t} \\ (1-t)e^{-t} \end{bmatrix}. \quad (6.13)$$

On obtient ainsi toutes les solutions de  $E_2$ .

**Remarque 8** (Structure de la solution). *Le vecteur  $\begin{bmatrix} t e^{-t} \\ (1-t)e^{-t} \end{bmatrix}$  s'écrit sous la forme  $e^{-t}\psi + t e^{-t}\phi$ , où  $\psi = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  et  $\phi = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ . La solution générale (6.13) est donc de la forme*

$$X(t) = e^{-t} [\alpha \phi + \beta(\psi + t\phi)]$$

où  $\phi$  est vecteur propre associé à la valeur propre (double)  $-1$ .

**Remarque 9** (Vers l'exponentielle de matrice). *La solution générale donnée par (6.13) s'écrit aussi*

$$X(t) = \begin{bmatrix} e^{-t} & t e^{-t} \\ -e^{-t} & (1-t)e^{-t} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

*Cet exemple sera généralisé dans le cours suivant en introduisant la fonction exponentielle d'une matrice.*

# Cours 7. Partiels corrigés

## Partiel du vendredi 25 octobre 2019

L'examen contient 3 exercices. Le barème est sur 26 points, il n'est donc pas demandé de tout faire pour avoir 20. Les documents (polycopié du cours, notes de TD, notes personnelles) sont autorisés.

**Exercice 1** (Deux équations linéaires du 1er ordre, barème 7 points).

Soient  $a > 0$  et  $y_0 > 0$ . On s'intéresse au problème de Cauchy suivant :

$$y'(t) = -ay(t), \quad t > 0, \quad (7.1)$$

$$y(0) = y_0. \quad (7.2)$$

1) Donner la solution du problème (7.1)-(7.2)

(c'est-à-dire la fonction  $y \in C([0, +\infty[, \mathbb{R}) \cap C^1(]0, +\infty[, \mathbb{R})$  solution de (7.1)-(7.2)).

*Corrigé – La solution est donnée par  $y(t) = y_0 e^{-at}$ .*

Soit  $b > 0$ . On s'intéresse maintenant au problème de Cauchy suivant, avec  $y$  donnée à la question 1 :

$$z'(t) = -bz(t) + ay(t), \quad t > 0, \quad (7.3)$$

$$z(0) = 0. \quad (7.4)$$

2) On suppose, dans cette question, que  $a \neq b$ .

(a) Donner la solution de (7.3)-(7.4).

*Corrigé – L'EDO à résoudre en  $z$  s'écrit  $z'(t) = -bz(t) + ay_0 e^{-at}$ . Les solutions de l'équation homogène associée sont de la forme  $z_h(t) = Ce^{-bt}$  avec  $C \in \mathbb{R}$ . On cherche une solution particulière sous la forme  $z_p(t) = \alpha e^{-at}$  où  $\alpha \in \mathbb{R}$  est à déterminer. On a alors  $z'_p(t) = -a\alpha e^{-at}$  et donc le coefficient  $\alpha$  doit satisfaire :  $-a\alpha = -b\alpha + ay_0$ . Ceci est possible car on a supposé  $a \neq b$  et on a donc*

$$\alpha = \frac{ay_0}{b-a}.$$

*Ecrivons alors la condition initiale pour déterminer  $C$  :*

$$z(0) = z_h(0) + z_p(0) = C + \frac{ay_0}{b-a} = 0,$$

*et donc  $C = ay_0/(a-b)$ , ce qui donne*

$$z(t) = \frac{ay_0}{b-a}(e^{-at} - e^{-bt}) \quad (7.5)$$

(b) Donner les valeurs de  $\lim_{t \rightarrow +\infty} z(t)$  et  $\int_0^{+\infty} z(t) dt$ .

Corrigé – En passant à la limite sur (7.5), on obtient  $\lim_{t \rightarrow +\infty} z(t) = 0$ . En intégrant (7.5) entre 0 et  $+\infty$ , on obtient :

$$\int_0^{+\infty} z(t) dt = \frac{ay_0}{b-a} \left( \frac{1}{a} - \frac{1}{b} \right) = \frac{y_0}{b}.$$

3) On suppose, dans cette question, que  $a = b$ .

(a) Donner la solution de (7.3)-(7.4).

Corrigé – La solution particulière obtenue à la question précédente n'est possible que pour  $a \neq b$ . Dans le cas  $a = b$ , l'EDO s'écrit  $z'(t) + az(t) = ay_0 e^{-at}$  et le second membre appartient à l'espace vectoriel des solutions de l'équation homogène associée. On cherche alors une solution particulière de la forme  $z_p(t) = \beta t e^{-at}$  où  $\beta \in \mathbb{R}$  est à déterminer. Une solution particulière est donc  $z_p(t) = ay_0 t e^{-at}$ . Cette solution vérifie aussi la condition initiale  $z(0) = 0$  et par unicité de la solution, on en déduit que

$$z(t) = ay_0 t e^{-at}.$$

(b) Donner les valeurs de  $\lim_{t \rightarrow +\infty} z(t)$  et  $\int_0^{+\infty} z(t) dt$ .

Corrigé – On a  $\lim_{t \rightarrow +\infty} z(t) = 0$ . Par intégration par parties,

$$\int_0^{+\infty} z(t) dt = [-y_0 t e^{-at}]_0^{+\infty} + \int_0^{+\infty} y_0 e^{-at} dt = \frac{y_0}{a}.$$

**Exercice 2** (Une équation non linéaire du 1er ordre, barème 5 points).

On s'intéresse au problème de Cauchy suivant :

$$y'(t) = 1 + t + \sin(t + y^2(t)), \quad t > 0, \quad (7.6)$$

$$y(0) = 0. \quad (7.7)$$

1) En appliquant les théorèmes du cours, montrer que le problème (7.6)-(7.7) admet une solution maximale.

Corrigé – La fonction  $f : (t, y) \mapsto 1 + t + \sin(t + y^2)$  est de classe  $C^\infty$  et donc localement lipschitzienne en sa deuxième variable. On peut donc appliquer le théorème de Cauchy-Lipschitz et le théorème d'unicité pour dire que le problème de Cauchy admet une solution maximale.

On note  $y$  la solution maximale donnée à la question 1,  $y \in C([0, T_m[, \mathbb{R}) \cap C^1(]0, T_m[, \mathbb{R})$ .

2) Montrer que  $y(t) \geq 0$  pour tout  $t \in [0, T_m[$ .

Corrigé – On a  $y(0) = 0$  et  $y'(t) \geq 0$  pour tout  $t \in ]0, T_m[$ , donc  $y(t) \geq 0$  pour tout  $t \in [0, T_m[$ .

3) Montrer que  $T_m = +\infty$ .

Corrigé – On a  $y'(t) = 1 + t + \sin(t + y(t)^2) \leq 2 + t$ . Donc  $y(t) \leq 2t + \frac{t^2}{2} \leq 2T_m + \frac{T_m^2}{2}$  pour tout  $t \in [0, T_m[$ , ce qui montre l'existence globale par le théorème sur la solution maximale.

**Exercice 3** (Une équation linéaire du deuxième ordre, barème 14 points). Soient  $I$  un intervalle ouvert de  $\mathbb{R}$  et  $a, b \in C(I, \mathbb{R})$ . On s'intéresse à l'équation, pour  $y \in C^2(I, \mathbb{R})$ ,

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0, \quad t \in I. \quad (7.8)$$

- 1) Rappeler brièvement pourquoi l'ensemble des solutions de l'équation (7.8) forme un espace vectoriel (sur  $\mathbb{R}$ ). Quelle est la dimension de cette espace vectoriel ?

*Corrigé – Le fait que l'ensemble des solutions de l'équation (7.8) forme un espace vectoriel est dû à la linéarité de l'équation (7.8). En effet, si les fonctions  $y$  et  $z$  sont solutions de (7.8), la fonction  $y + z$  est encore solution de l'équation (7.8). De même si  $y$  est solution de (7.8) et  $\alpha \in \mathbb{R}$ , la fonction  $\alpha y$  est encore solution de l'équation (7.8).*

*La dimension de cette espace vectoriel est 2.*

Pour la suite de l'exercice, on note  $E$  l'ensemble des solutions de l'équation (7.8).

On se donne deux solutions de l'équation (7.8), notées  $y_1$  et  $y_2$  et on définit la fonction  $z$  de  $I$  dans  $\mathbb{R}$  par

$$z(t) = y_1'(t)y_2(t) - y_2'(t)y_1(t), \text{ pour tout } t \in I.$$

- 2) Donner (en fonction des données du problème) l'équation différentielle du 1er ordre satisfaite par  $z$  sur  $I$ .

*Corrigé – On remarque que, pour tout  $t \in I$ ,*

$$\begin{aligned} z'(t) &= y_1''(t)y_2(t) - y_2''(t)y_1(t) = (-a(t)y_1'(t) - b(t)y_1(t))y_2(t) + (a(t)y_2'(t) + b(t)y_2(t))y_1(t) \\ &= -a(t)z(t). \end{aligned}$$

- 3) Montrer que  $z(t) \neq 0$  pour tout  $t \in I$  sauf si  $z$  est la fonction identiquement nulle.

*Corrigé – La fonction  $z$  est solution de l'équation différentielle du 1er ordre  $z'(t) = -a(t)z(t)$ ,  $t \in I$ . La fonction identiquement nulle est solution de cette équation différentielle. Comme les trajectoires des solutions de cette équation différentielle ne se rencontrent pas, on en déduit bien que  $z(t) \neq 0$  pour tout  $t \in I$  si  $z$  n'est pas la fonction identiquement nulle.*

*Une autre façon de répondre à la question est de remarquer que la solution générale de l'équation différentielle satisfaite par  $z$  est  $z(t) = Ce^{A(t)}$  ou  $A$  est une primitive de  $a$  et  $C \in \mathbb{R}$ . Cette fonction de s'annule pas sauf si  $C = 0$  (et on a alors  $z \equiv 0$ ).*

- 4) On suppose dans cette question que  $z$  est la fonction identiquement nulle. Soit  $t_0 \in I$ . On considère le système de deux équations à deux inconnues réelles (notées  $\alpha_1$  et  $\alpha_2$ )

$$\begin{aligned} \alpha_1 y_1(t_0) + \alpha_2 y_2(t_0) &= 0, \\ \alpha_1 y_1'(t_0) + \alpha_2 y_2'(t_0) &= 0. \end{aligned}$$

En utilisant  $z(t_0) = 0$ , montrer que ce système n'est pas inversible et donc qu'il existe  $(\alpha_1, \alpha_2) \in \mathbb{R}^2$  solution de ce système et tel que  $(\alpha_1, \alpha_2) \neq (0, 0)$ .

En déduire que pour tout  $t \in I$ ,  $\alpha_1 y_1(t) + \alpha_2 y_2(t) = 0$ . (Les fonctions  $y_1$  et  $y_2$  sont donc dépendantes.)

[Utiliser le théorème d'unicité vu en cours pour les équations du deuxième ordre]

*Corrigé – la matrice de ce système est*

$$A = \begin{bmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{bmatrix}.$$

*Comme  $\det(A) = -z(t_0) = 0$ , le système n'est pas inversible. Ceci prouve que  $\text{Ker}(A) \neq \{0\}$  et donc qu'il existe  $(\alpha_1, \alpha_2) \in \mathbb{R}^2$  solution de ce système et tel que  $(\alpha_1, \alpha_2) \neq (0, 0)$ .*

*On pose, pour tout  $t \in I$ ,  $y(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t)$ . La fonction  $y$  est solution de (7.8). Elle est nulle ainsi que sa dérivée en 0. Le théorème d'unicité vu en cours pour les équations du second ordre donne alors  $y \equiv 0$ .*

Pour la suite de l'exercice, on suppose que  $z$  n'est pas la fonction identiquement nulle.

5) Soient  $t_0 \in I$  et  $c, d \in \mathbb{R}$ . A l'équation (7.8), on ajoute la condition

$$y(t_0) = c, \quad y'(t_0) = d. \quad (7.9)$$

(On rappelle qu'il existe une unique solution à (7.8)-(7.9).)

En utilisant  $z(t_0) \neq 0$ , montrer qu'il existe  $\alpha_1, \alpha_2 \in \mathbb{R}$  tels que la solution de (7.8)-(7.9) s'écrit

$$y(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t), \quad \text{pour tout } t \in I. \quad (7.10)$$

En déduire que les fonctions  $y_1, y_2$  forment une base de  $E$ .

*Corrigé –* Soient  $\alpha_1, \alpha_2 \in \mathbb{R}$ . La fonction  $y$  définie par (7.10) est solution de (7.8)-(7.9) si et seulement si

$$\begin{aligned} \alpha_1 y_1(t_0) + \alpha_2 y_2(t_0) &= c, \\ \alpha_1 y_1'(t_0) + \alpha_2 y_2'(t_0) &= d. \end{aligned}$$

Comme  $z(t_0) \neq 0$ , ce système (dont les inconnues sont  $\alpha_1, \alpha_2$ ) est inversible et admet donc une (unique) solution. On obtient ainsi la solution de (7.8)-(7.9).

Quand  $c$  et  $d$  décrivent  $\mathbb{R}$ , on obtient toutes les solutions de (7.8). Ceci prouve que les fonctions  $y_1, y_2$  engendrent  $E$  et, comme  $\dim E = 2$ , forment une base de  $E$ .

Soit  $g \in C(I, \mathbb{R})$ . On s'intéresse maintenant à l'équation

$$y''(t) + a(t)y'(t) + b(t)y(t) = g(t), \quad t \in I. \quad (7.11)$$

6) On cherche une solution de (7.11) sous la forme  $y(t) = \alpha_1(t)y_1(t) + \alpha_2(t)y_2(t)$  avec la condition  $\alpha_1'(t)y_1(t) + \alpha_2'(t)y_2(t) = 0$  pour tout  $t \in I$ .

Montrer que  $y$  est solution de (7.11) si et seulement si  $\alpha_1'(t)y_1'(t) + \alpha_2'(t)y_2'(t) = g(t)$  pour tout  $t \in I$ .

En déduire qu'il est effectivement possible de trouver des fonctions  $\alpha_1$  et  $\alpha_2$  telles que  $y$  soit solution de (7.11) et qu'on obtient ainsi toutes les solutions de (7.11).

*Corrigé –*

Comme  $y'(t) = \alpha_1(t)y_1'(t) + \alpha_2(t)y_2'(t)$  et  $y''(t) = \alpha_1'(t)y_1'(t) + \alpha_2'(t)y_2'(t) + \alpha_1(t)y_1''(t) + \alpha_2(t)y_2''(t)$  et que  $y_1, y_2$  sont solutions de (7.8), on obtient que  $y$  est solution de (7.11) si et seulement si  $\alpha_1'(t)y_1'(t) + \alpha_2'(t)y_2'(t) = g(t)$  (pour tout  $t \in I$ ).

Il suffit donc de trouver des fonctions  $\alpha_1$  et  $\alpha_2$  telles que, pour tout  $t \in I$ ,

$$\begin{aligned} \alpha_1'(t)y_1(t) + \alpha_2'(t)y_2(t) &= 0, \\ \alpha_1'(t)y_1'(t) + \alpha_2'(t)y_2'(t) &= g(t). \end{aligned}$$

Comme  $z(t) \neq 0$ , ce système est inversible et on trouve donc les fonctions  $\alpha_1'$  et  $\alpha_2'$  en fonction de  $y_1$  et  $y_2$ . Ceci donne aussi (en prenant les primitives) les fonctions  $\alpha_1$  et  $\alpha_2$ .

Comme les primitives sont définies à une constante près, on obtient bien toutes les solutions de (7.11).

7) Dans cette question, on prend  $I = \mathbb{R}$ ,  $a(t) = 0$ ,  $b(t) = 1$  et  $g(t) = \sin t$ .

(a) Donner deux fonctions  $y_1, y_2$  formant une base de  $E$ .

*Corrigé –* L'équation caractéristique pour cette équation est  $r^2 + 1 = 0$ . Les solutions sont  $r = \pm i$ . Une base de  $E$  est obtenue en prenant  $y_1(t) = \cos t$  et  $y_2(t) = \sin t$

(b) Donner toutes les solutions de (7.11) (il est autorisé de deviner la forme d'une solution particulière, sinon utiliser la méthode de la question 6...).

Corrigé – Comme  $g = y_2$ , on cherche une solution particulière sous la forme  $y(t) = \alpha t \cos t + \beta t \sin t$  (avec  $\alpha, \beta \in \mathbb{R}$ ). Comme

$$y'(t) = -\alpha t \sin t + \beta t \cos t + \alpha \cos t + \beta \sin t,$$

et

$$y''(t) = -y(t) - 2\alpha \sin t + 2\beta t \cos t,$$

la fonction  $y$  est solution de (7.11) si seulement si  $\alpha = -1/2$  et  $\beta = 0$ . Une solution particulière est donc

$$y(t) = -\frac{1}{2} t \cos t, \text{ pour tout } t \in \mathbb{R}.$$

La solution générale de (7.11) est donc, pour  $\alpha_1, \alpha_2 \in \mathbb{R}$ ,

$$y(t) = -\frac{1}{2} t \cos t + \alpha_1 \cos t + \alpha_2 \sin t, \text{ pour tout } t \in \mathbb{R}.$$

### Partiel du vendredi 23 octobre 2020

Le partiel contient 3 exercices. Le barème est sur 23 points, il n'est donc pas demandé de tout faire pour avoir 20. Les documents (polycopié du cours, notes de TD, notes personnelles) sont autorisés.

#### Exercice 4 (Dynamique des populations, barème 10 points).

Soient  $\alpha, a \in C(\mathbb{R}_+, \mathbb{R})$  telles que  $0 < \alpha(t) < 1$  et  $a(t) > 0$  pour tout  $t \in \mathbb{R}$ . Pour  $y_0 > 0$ , on s'intéresse au problème de Cauchy

$$y'(t) = y(t)(\alpha(t) - y(t))a(t), \quad t > 0, \quad (7.12)$$

$$y(0) = y_0. \quad (7.13)$$

- 1) Montrer que le problème (7.12)-(7.13) admet une unique solution maximale. Cette solution est définie sur l'intervalle  $[0, T_m[$ . Montrer que  $y(t) > 0$  pour tout  $t \in [0, T_m[$ .

Corrigé – La fonction  $f : t, x \mapsto x(\alpha(t) - x)a(t)$  est continue sur  $\mathbb{R}_+ \times \mathbb{R}$  dans  $\mathbb{R}$ . Sa dérivée par rapport à son deuxième argument est continue. La fonction  $f$  est donc localement lipschitzienne par rapport à son deuxième argument (proposition 1 du 2ème cours). Les théorèmes du cours donnent alors l'existence et l'unicité d'une solution maximale.

La fonction nulle est la trajectoire d'une solution de l'équation. Comme les trajectoires ne se rencontrent pas et que  $y_0 > 0$ , on en déduit que  $y(t) > 0$  pour tout  $t \in [0, T_m[$ .

- 2) On suppose dans cette question que la fonction  $\alpha$  est constante. Montrer que la solution maximale trouvée à la question 1 est globale (c'est-à-dire que  $T_m = +\infty$ ).

Corrigé – Si  $y_0 = \alpha$ ,  $y(t) = \alpha$  pour tout  $t$ , la solution est globale.

Si  $y_0 < \alpha$ , la solution maximale reste entre les deux points stationnaires 0 et  $\alpha$ . Elle n'explose pas en  $T_m$  et donc  $T_m = +\infty$ .

Si  $y_0 > \alpha$ , la solution reste au dessus de  $\alpha$ . On a donc  $y'(t) < 0$  pour tout  $t$ . Ceci montre qu'elle est décroissante et donc qu'elle reste entre  $y_0$  et le point stationnaire  $\alpha$ . Elle n'explose pas en  $T_m$  et donc  $T_m = +\infty$ .

- 3) On ne suppose plus que la fonction  $\alpha$  est constante.

- (a) on suppose dans cette question que  $y_0 < 1$ . Montrer que  $y(t) < 1$  pour tout  $t \in [0, T_m[$ .

[On pourra raisonner par contradiction. Poser  $t^* = \min\{t \in \mathbb{R}_+ \text{ tel que } y(t) \geq 1\}$  (de sorte que  $t^* > 0$  et  $y(t) < 1$  si  $t < t^*$ ) et montrer que  $y(t^*) = 1$ ,  $y'(t^*) \geq 0$  et  $y'(t^*) < 0$ , ce qui est impossible].

En déduire que  $T_m = +\infty$ .

Corrigé – On suppose qu'il existe  $t > 0$  tel que  $y(t) \geq 1$ . On pose alors  $t^* = \inf\{t \in \mathbb{R}_+ \text{ tel que } y(t) \geq 1\}$ . Par continuité de  $y$ , on a  $y(t^*) \geq 1$  (ceci montre que "inf" est aussi "min"). Si  $y(t^*) > 1$ , par le théorème des valeurs intermédiaires (et le fait que  $y(0) < 1$ ) il existe  $t \in ]0, t^*[$  tel que  $y(t) = 1$ , ce qui contredit la définition de  $t^*$ . Donc,  $y(t^*) = 1$ .

On remarque maintenant que  $y'(t^*) \geq 0$  car  $y(t) - y(t^*) = y(t) - 1 < 0$  si  $t < t^*$  et que  $y'(t^*) = (\alpha(t^*) - 1)a(t^*) < 0$ . On obtient bien une contradiction.

Comme la solution reste bloquée entre 0 et 1, on en déduit que  $T_m = +\infty$ .

- (b) on suppose dans cette question que  $y_0 \geq 1$ . Montrer que  $y(t) < y_0$  pour tout  $t \in ]0, T_m[$ .  
En déduire que  $T_m = +\infty$ .

Corrigé – Comme  $y_0 \geq 1$  et  $\alpha(0) < 1$ ,  $y_0(\alpha(0) - y_0) < 0$ . Par continuité de  $y$ , il existe alors  $\varepsilon > 0$  tel que  $y(t)(\alpha(t) - y(t)) < 0$  pour  $t \in [0, \varepsilon]$ . On en déduit que  $y(t) < y_0$  pour  $t \in [0, \varepsilon]$ .

On raisonne maintenant de nouveau par contradiction. Si il existe  $t \in ]0, T_m[$  tel que  $y(t) \geq y_0$ , on pose  $t^* = \inf\{t \in ]0, +\infty[ \text{ tel que } y(t) \geq y_0\}$ . On remarque que  $t^* \geq \varepsilon > 0$ . Par continuité de  $y$ , on a  $y(t^*) \geq y_0$ . (En fait, on peut aussi montrer que  $y(t^*) = y_0$ .)

Comme dans la question précédente, on montre que  $y'(t^*) \geq 0$  (car  $y(t) < y_0 \leq y(t^*)$  si  $t < t^*$ ) et  $y'(t^*) = y(t^*)(\alpha - y(t^*))a(t^*) < 0$ , ce qui est impossible.

Comme la solution reste bloquée entre 0 et  $y_0$ , on en déduit que  $T_m = +\infty$ .

- 4) On suppose à nouveau, dans cette question, que la fonction  $\alpha$  est constante et on note  $y$  la solution de (7.12)-(7.13)

- (a) On suppose dans cette question qu'il existe  $a_0 > 0$  tel que  $a(t) \geq a_0$  pour tout  $t > 0$ .

Montrer que  $\lim_{t \rightarrow +\infty} y(t) = \alpha$ .

Corrigé – Si  $y_0 = \alpha$ ,  $y(t) = \alpha$  pour tout  $t$  (et donc  $\lim_{t \rightarrow +\infty} y(t) = \alpha$ ).

Si  $y_0 < \alpha$ , la fonction  $y$  est strictement croissante majorée par  $\alpha$ . Elle a donc une limite que l'on note  $\ell$  avec  $\ell \in ]y_0, \alpha]$ .

Si  $y_0 > \alpha$ , la fonction  $y$  est strictement décroissante minorée par  $\alpha$ . Elle a donc une limite que l'on note aussi  $\ell$  avec  $\ell \in [\alpha, y_0[$ .

Dans les deux cas, on peut noter que  $\ell \neq 0$ .

On utilise le théorème des accroissements finis entre  $t$  et  $t + 1$ , il existe  $\theta_t \in ]t, t + 1[$  tel que

$$y(t + 1) - y(t) = y'(\theta_t),$$

et donc

$$a_0 \geq a(\theta_t) = \frac{y(t + 1) - y(t)}{y(\theta_t)(\alpha - y(\theta_t))}.$$

Quand  $t \rightarrow +\infty$ ,  $\theta_t \rightarrow +\infty$  et  $y(\theta_t) \rightarrow \ell$ . On en déduit que le terme de droite de cette inégalité tend vers 0 (en contradiction avec  $a_0 > 0$ ) sauf si  $\ell = \alpha$ .

- (b) (On ne suppose plus l'existence de  $a_0$ .) On suppose  $y_0 \neq \alpha$ .

Montrer que  $\lim_{t \rightarrow +\infty} y(t) = \alpha$  si et seulement si  $\lim_{t \rightarrow +\infty} \int_0^t a(s) ds = +\infty$ .

Corrigé – Il est possible, de calculer la solution avec la méthode des variables séparables, Si  $y_0 \neq \alpha$ , on a alors  $y(t) \neq 0$  et  $y(t) \neq \alpha$  pour tout  $t$  et l'équation s'écrit

$$\frac{y'(t)}{y(t)(\alpha - y(t))} = a(t) \text{ pour tout } t > 0.$$

Ce qui donne (avec une décomposition en éléments simples)

$$\frac{y'(t)}{y(t)} + \frac{y'(t)}{\alpha - y(t)} = \alpha a(t) \text{ pour tout } t > 0.$$

Dans le cas  $y(t) < \alpha$  pour tout  $t$ , ceci donne

$$\left(\ln \frac{y}{\alpha - y}\right)' = \alpha a.$$

En notant  $A$  la primitive de  $a$  s'annulant en 0, c'est-à-dire  $A(t) = \int_0^t a(s)ds$ , on obtient que la fonction  $\ln \frac{y}{\alpha - y} - \alpha A$  est constante. Il existe donc  $C \in \mathbb{R}$  tel que

$$\frac{y}{\alpha - y} = Ce^{\alpha A}, \quad y = \frac{\alpha C e^{\alpha A}}{1 + C e^{\alpha A}}.$$

La valeur de  $C$  se calcule en utilisant  $y(0) = y_0$ ,

$$y(t) = \frac{\alpha y_0}{y_0 + (\alpha - y_0)e^{-\alpha A(t)}}$$

On trouve la même formule si  $y(t) > \alpha$  pour tout  $t$ .

Si  $\lim_{t \rightarrow +\infty} A(t) = +\infty$ , on déduit de cette formule  $\lim_{t \rightarrow +\infty} y(t) = \alpha$ .

Si  $\lim_{t \rightarrow +\infty} A(t) = \gamma < +\infty$ , on déduit de cette formule

$$\lim_{t \rightarrow +\infty} y(t) = \frac{\alpha y_0}{y_0 + (\alpha - y_0)e^{-\alpha \gamma}} \neq \alpha.$$

### Exercice 5 (Petits exemples, barème 6 points).

1) On considère l'équation différentielle du deuxième ordre

$$y''(t) + 2y'(t) + y(t) = \psi(t), \quad t \in \mathbb{R}. \quad (7.14)$$

(a) Donner la solution générale de l'équation homogène associée

*Corrigé –* Le polynôme caractéristique est  $r^2 + 3r + 1$  dont la racine double est  $-1$

La solution générale de l'équation homogène associée est donc la fonction

$$t \mapsto y_h(t) = C_1 e^{-t} + C_2 t e^{-t},$$

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

(b) Donner une solution particulière de (7.14) avec  $\psi(t) = e^{2t}$ .

*Corrigé –* On cherche une solution particulière associée à  $e^{2t}$  sous la forme  $y_1 = \alpha e^{2t}$ . La condition sur  $\alpha$  est alors  $4\alpha + 4\alpha + \alpha = 9\alpha = 1$ , c'est-à-dire  $\alpha = 1/9$ ,  $y_1(t) = (1/9)e^{2t}$ .

(c) Donner une solution particulière de (7.14) avec  $\psi(t) = e^{-t}$ .

*Corrigé –* On cherche une solution particulière associée à  $e^{-t}$  sous la forme  $y_2(t) = \alpha t^2 e^{-t}$  (car les fonctions  $e^{-t}$  et  $t e^{-t}$  sont solutions de l'équation homogène).

Ceci donne

$$y_2'(t) = -\alpha t^2 e^{-t} + \alpha 2t e^{-t},$$

$$y_2''(t) = \alpha t^2 e^{-t} - \alpha 4t e^{-t} + \alpha 2e^{-t}.$$

La condition sur  $\alpha$  est alors  $2\alpha = 1$ , c'est-à-dire  $\alpha = 1/2$ ,  $y_2(t) = (1/2)t^2 e^{-t}$ .

(d) Donner la solution générale de (7.14) avec  $\psi(t) = e^{2t} + e^{-t}$ .

*Corrigé –* La solution générale de l'équation non homogène est donc la fonction

$$t \mapsto y_h(t) = C_1 e^{-t} + C_2 t e^{-t} + (1/9)e^{2t} + (1/2)t^2 e^{-t}.$$

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

2) On considère l'équation différentielle du premier ordre

$$y'(t) - 2ty(t) = e^{t^2}, \quad t \in \mathbb{R}.$$

(a) Donner la solution générale de l'équation homogène associée.

*Corrigé – La solution générale de l'équation homogène s'obtient avec la primitive de  $-2t$  c'est la fonction*

$$t \mapsto Ce^{t^2},$$

avec  $C \in \mathbb{R}$ .

(b) Donner la solution générale de l'équation complète.

*Corrigé – On peut chercher une solution particulière avec la méthode de réduction d'ordre (appelée aussi ici "variation de la constante"), c'est-à-dire sous la forme  $y_p(t) = z(t)e^{t^2}$ .*

*Comme  $y_p'(t) = z'(t)e^{t^2} + 2tz(t)e^{t^2}$ , la condition sur  $z$  est  $z'(t) = 1$ . on peut donc prendre  $z(t) = t$  et  $y_p(t) = te^{t^2}$ .*

*La solution générale de l'équation complète est donc la fonction*

$$t \mapsto Ce^{t^2} + te^{t^2},$$

avec  $C \in \mathbb{R}$ .

**Exercice 6** (Une équation linéaire du 2ème ordre, barème 7 points).

On s'intéresse à l'équation, pour  $y \in C^2(\mathbb{R}_+^*, \mathbb{R})$ ,

$$t^2 y''(t) + 3ty'(t) + y(t) = 0, \quad t > 0. \quad (7.15)$$

1) Montrer que les théorèmes vus en cours permettent de montrer que l'ensemble des solutions de l'équation (7.15) forme un espace vectoriel (sur  $\mathbb{R}$ ) de dimension 2.

*Corrigé – En divisant l'équation par  $t^2$ , on est ramené à une équation du type  $y''(t) + a(t)y'(t) + b(t)y(t) = 0$ ,  $\in I = \mathbb{R}_+^*$ . Les fonctions  $a$  et  $b$  sont continues sur  $I$  et un théorème du cours s'applique pour montrer que l'ensemble des solutions de l'équation (7.15) forme un espace vectoriel (sur  $\mathbb{R}$ ) de dimension 2.*

2) Montrer que pour  $\alpha \in \mathbb{R}$  bien choisi la fonction  $y_1 : t \mapsto t^\alpha$  est solution de (7.15).

*Corrigé – La fonction  $y_1$  est solution si et seulement si  $\alpha(\alpha - 1) + 3\alpha + 1 = \alpha^2 + 2\alpha + 1 = 0$ , c'est-à-dire  $\alpha = -1$*

3) On cherche maintenant une seconde solution de (7.15), notée  $y_2$ , linéairement indépendante de la fonction  $y_1$  trouvée à la question précédente. On cherche  $y_2$  en utilisant la technique de "réduction d'ordre", c'est-à-dire que l'on pose, pour tout  $t > 0$ ,  $y_2(t) = z(t)y_1(t)$ .

(a) Montrer que  $y_2$  est solution de (7.15) si et seulement si  $tz''(t) + z'(t) = 0$  pour tout  $t > 0$ .

*Corrigé –*

$$y_2 = zy_1,$$

$$y_2' = zy_1' + z'y_1,$$

$$y_2'' = zy_1'' + 2z'y_1' + z''y_1.$$

*La fonction  $y_2$  est donc solution de (7.15) si et seulement si  $t^2 y_1 z'' + (2t^2 y_1' + 3ty_1)z' = 0$ , c'est-à-dire*

$$tz''(t) + z'(t) = 0, \quad t > 0.$$

- (b) Dédurre de la question précédente, une fonction  $y_2$ , linéairement indépendante de la fonction  $y_1$ , solution de (7.15).

Corrigé – En supposant  $z' > 0$  on obtient  $z''(t)/z'(t) = -1/t$ , c'est-à-dire  $(\ln(z'))'(t) = -1/t$ . Une solution est donc obtenue en prenant  $\ln(z'(t)) = -\ln(t)$ , et donc

$$z'(t) = e^{-\ln(t)} = \frac{1}{t}.$$

Une solution possible est  $z(t) = \ln(t)$  et

$$y_2(t) = \frac{\ln(t)}{t} \text{ pour tout } t > 0.$$

Les deux solutions  $y_1$  et  $y_2$  sont bien linéairement indépendantes (car la fonction  $t \mapsto \ln(t)$  est non constante).

- 4) Donner l'ensemble des solutions de (7.15).

Corrigé – La solution générale de l'équation (7.15) est la fonction  $t \mapsto (C_1 + C_2 \ln(t))/t$  avec  $C_1, C_2 \in \mathbb{R}$ .

- 5) Donner l'ensemble des solutions de l'équation non homogène

$$t^2 y''(t) + 3t y'(t) + y(t) = \frac{1}{t^2}, \quad t > 0.$$

Corrigé – On cherche une solution particulière sous la forme  $y_p(t) = \gamma/t^2$ . Comme  $y_p'(t) = -2\gamma t^{-3}$  et  $y_p''(t) = 6\gamma t^{-4}$ , La condition sur  $\gamma$  est alors

$$6\gamma - 6\gamma + \gamma = 1,$$

c'est-à-dire  $\gamma = 1$  et  $y_p = 1/t^2$ .

La solution générale de l'équation non homogène est donc la fonction

$$t \mapsto \frac{C_1 + C_2 \ln(t)}{t} + \frac{1}{t^2},$$

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

## Cours 8. Exponentielle de matrices

On s'intéresse à nouveau dans ce cours au système autonome homogène qui s'écrit :

$$X'(t) = AX(t), \quad t \in \mathbb{R} \quad (8.1)$$

où  $A \in \mathcal{M}_n(\mathbb{R})$ . On rappelle que l'ensemble  $E_n$  des solutions de (8.1) est un espace vectoriel de dimension  $n$ . On cherche une base de  $E_n$  c'est-à-dire une famille de  $n$  fonctions (de  $\mathbb{R}$  dans  $\mathbb{R}^n$ ) qu'on note  $X^{(1)}, \dots, X^{(n)}$ , linéairement indépendantes et solutions de (8.1). La famille  $\{X^{(1)}, \dots, X^{(n)}\}$  est alors une base de  $E_n$  et les solutions de (8.1) sont les combinaisons linéaires des fonctions  $X^{(1)}, \dots, X^{(n)}$ ; elles sont donc de la forme

$$X(t) = \sum_{i=1}^n c_i X^{(i)}(t), \quad t \in \mathbb{R}, \quad c_1, \dots, c_n \in \mathbb{R}.$$

Une telle solution peut s'écrire sous forme matricielle en introduisant pour  $t \in \mathbb{R}$  la matrice  $M(t) \in \mathcal{M}_n(\mathbb{R})$  définie par

$$M(t) = \begin{bmatrix} X^{(1)}(t) & \dots & X^{(n)}(t) \end{bmatrix}, \quad \text{pour } t \in \mathbb{R} \quad (8.2)$$

dont la  $i$ -ème colonne est donc  $X^{(i)}(t)$ . Il est clair que

$$M'(t) = AM(t). \quad (8.3)$$

Pour  $(c_1, \dots, c_n) \in \mathbb{R}$ , on pose

$$C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}),$$

de sorte que

$$M(t)C = \sum_{i=1}^n c_i X^{(i)}(t).$$

Noter que cette égalité résulte simplement de la définition du produit matrice-vecteur : le vecteur  $M(t)C$  est une combinaison linéaire des colonnes de  $M(t)$  pour  $t \in \mathbb{R}$ , dont les coefficients sont les composantes du vecteur. La solution générale de (8.1) est donc  $M(t)C$ ,  $t \in \mathbb{R}$ , avec  $C$  arbitraire dans  $\mathbb{R}^n$ . L'objectif est de trouver une matrice  $M(t)$ . Il est assez facile de la trouver dans le cas où la matrice  $A$  est diagonalisable, et nous commencerons par ce cas, après avoir rappelé quelques résultats fondamentaux d'algèbre linéaire. Nous traiterons ensuite le cas où la matrice  $A$  est non diagonalisable. On note  $Sp(A)$  l'ensemble des valeurs propres de  $A$  et  $P_A$  le polynôme caractéristique de  $A$  défini par  $P_A(y) = \det(A - yI)$  où  $I$  est la matrice identité. On commence par rappeler la notion, fondamentale, de multiplicité géométrique et algébrique d'une valeur propre.

**Définition 10** (Multiplicité géométrique et algébrique). Soit  $A \in \mathcal{M}_n(\mathbb{R})$ ,  $\lambda \in Sp(A)$ , on note  $m_g(\lambda)$  la multiplicité géométrique de  $\lambda$  et  $m_a(\lambda)$  la multiplicité algébrique de  $\lambda$ . On rappelle que  $m_g(\lambda) \in \mathbb{N}$  est défini par :

$$m_g(\lambda) = \dim(\text{Ker}(A - \lambda I))$$

et  $m_a(\lambda) \in \mathbb{N}$  par

$$P_A(y) = (y - \lambda)^{m_a(\lambda)} Q(y),$$

où  $P_A$  le polynôme caractéristique de  $A$  et  $Q$  est un polynôme de degré  $n - m_a(\lambda)$  tel que  $Q(\lambda) \neq 0$ . On a forcément  $m_a(\lambda) \geq 1$  et  $m_g(\lambda) \geq 1$  puisque  $\lambda$  est une valeur propre, et donc  $\dim(\text{Ker}(A - \lambda I)) \geq 1$  et  $\lambda$  est une racine de  $P_A$ . De plus, on note qu'on a toujours  $1 \leq m_g(\lambda) \leq m_a(\lambda)$  et  $\sum_{\lambda \in Sp(A)} m_a(\lambda) = n$ .

La matrice  $A$  est diagonalisable dans  $\mathbb{C}$  si et seulement si  $m_a(\lambda) = m_g(\lambda)$  pour tout  $\lambda$ .

La matrice  $A$  est diagonalisable dans  $\mathbb{R}$  si  $A$  est diagonalisable dans  $\mathbb{C}$  et  $\lambda \in \mathbb{R}$  pour tout  $\lambda \in Sp(A)$ .

## 8.1 Premier cas : La matrice $A$ est diagonalisable...

### 8.1.1 ... dans $\mathbb{R}$

Pour  $\lambda \in Sp(A)$  on choisit  $\varphi \in \mathbb{R}^n$  vecteur propre de  $A$  associé à  $\lambda$ , c'est-à-dire  $\varphi \neq 0$  et  $A\varphi = \lambda\varphi$ . La fonction  $t \mapsto e^{\lambda t}\varphi$  est alors solution de (8.1). Comme  $A$  est diagonalisable dans  $\mathbb{R}$ , on obtient ainsi  $n$  solutions linéairement indépendantes de (8.1), notées  $X^{(i)}$ ,  $i = 1, \dots, n$ . Plus précisément,  $X^{(i)}(t) = e^{\lambda_i t}\varphi_i$ , avec  $A\varphi_i = \lambda_i\varphi_i$ ,  $\lambda_i \in \mathbb{R}$  et  $\varphi_i \neq 0$ . Soit  $M(t)$  la matrice définie par (8.2). On rappelle que la solution de générale de (8.1) est alors la fonction  $t \mapsto M(t)C$ , avec  $C \in \mathbb{R}^n$ .

### 8.1.2 ... dans $\mathbb{C}$

Supposons maintenant que  $\lambda \in Sp(A)$  et  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ . On pose  $\lambda = \alpha + i\beta$ ,  $\beta \neq 0$ . Soit  $\varphi \in \mathbb{C}^n$  où  $\varphi = \psi + i\chi$ , avec  $\psi, \chi \in \mathbb{R}^n$  et  $A\varphi = \lambda\varphi$ . On rappelle (voir section 6.3.2) que l'on obtient deux solutions linéairement indépendantes de (8.1) en posant :

$$\begin{cases} X^{(1)}(t) = e^{\alpha t}(\cos(\beta t)\psi - \sin(\beta t)\chi) \\ X^{(2)}(t) = e^{\alpha t}(\sin(\beta t)\psi + \cos(\beta t)\chi) \end{cases}$$

Pour montrer que ces deux solutions sont linéairement indépendantes, il suffit grâce à la proposition 4 de remarquer que leurs valeurs en 0 sont linéairement indépendantes, c'est-à-dire que  $\psi$  et  $\chi$  sont des vecteurs indépendants. Commençons par montrer que  $\psi$  et  $\chi$  sont non nuls. En effet si  $\chi = 0$ , alors

$$\underbrace{A\psi}_{\in \mathbb{R}} = \underbrace{(\alpha + i\beta)\psi}_{\notin \mathbb{R}},$$

ce qui est donc impossible. De même si  $\psi = 0$  alors :

$$\underbrace{Ai\chi}_{\in i\mathbb{R}} = \underbrace{(\alpha + i\beta)(i\chi)}_{\notin i\mathbb{R}},$$

ce qui est également impossible. Maintenant si  $\psi$  et  $\chi$  sont liés, alors il existe  $c \in \mathbb{R}^*$  tel que  $\chi = c\psi$ , on a alors

$$A((1 + ic)\psi) = (\alpha + i\beta)((1 + ic)\psi).$$

En multipliant cette égalité par  $(1 - ic)$  on obtient

$$\underbrace{A(1 + c^2)\psi}_{\in \mathbb{R}} = \underbrace{(\alpha + i\beta)(1 + c^2)\psi}_{\notin \mathbb{R}}$$

ce qui est impossible. On en déduit que  $\psi$  et  $\chi$  sont linéairement indépendants.

Pour chaque valeur propre complexe non réelle, de  $A$ , on obtient donc deux solutions linéairement indépendantes de (8.1) construites à partir d'un vecteur propre  $\varphi$ . Si  $\lambda$  est de multiplicité géométrique  $m_g$  (le conjugué de  $\lambda$  est alors aussi de multiplicité  $m_g$ ), on construit ainsi  $2m_g$  solutions linéairement indépendantes de (8.1). (On obtiendra le même ensemble de solutions avec la valeur propre  $\bar{\lambda}$ .)

Si  $\lambda \in Sp(A)$ ,  $\lambda \in \mathbb{R}$  et  $\lambda$  est de multiplicité géométrique  $m_g$ , on construit  $m_g$  solutions linéairement indépendantes de (8.1), comme dans le cas réel traité au paragraphe 8.1.1.

Comme on suppose que  $A$  est diagonalisable dans  $\mathbb{C}$  (on a donc  $m_g(\lambda) = m_a(\lambda)$  pour tout  $\lambda \in Sp(A)$ ), on obtient ainsi  $n$  solutions linéairement indépendantes de (8.1). (Pour montrer que ces solutions sont linéairement indépendantes, il suffit de remarquer que leurs valeurs en 0 sont linéairement indépendantes, ce point n'est pas détaillé ici).

En conclusion des paragraphes 8.1.1-8.1.2, si  $A$  est diagonalisable dans  $\mathbb{R}$  ou  $\mathbb{C}$  alors (8.1), il existe  $n$  solutions linéairement indépendantes de (8.1).

## 8.2 Cas général

On va étudier maintenant le cas général, où  $A$  peut ne pas être diagonalisable, c'est-à-dire qu'il existe  $\lambda \in Sp(A)$  telle que  $m_g(\lambda) < m_a(\lambda)$ , comme dans l'exemple suivant.

**Exemple 19** (Une matrice non diagonalisable). Soit

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \text{ et donc } P_A(y) = \begin{vmatrix} -y & 1 \\ 0 & -y \end{vmatrix} = y^2.$$

On a donc  $Sp(A) = \{0\}$ ,  $m_a(0) = 2$  et  $m_g(0) = 1$ .

On rappelle un résultat d'algèbre linéaire qui va permettre de trouver une base de l'ensemble des solutions du système différentiel (8.1).

**Lemme 8** (Multiplicité algébrique et dimension du sous-espace propre). Soient  $A \in \mathcal{M}_n(\mathbb{R})$  et  $\lambda \in Sp(A)$ . Alors, il existe  $d \in \mathbb{N}$  tel que  $1 \leq d \leq m_a(\lambda)$  et  $\dim(\text{Ker}(A - \lambda I)^d) = m_a(\lambda)$ . (En particulier  $\dim(\text{Ker}(A - \lambda I)^{m_a(\lambda)}) = m_a(\lambda)$ .)

**Exemple 20** (Cas  $m_a(\lambda) = 2$  et  $m_g(\lambda) = 1$ ). Soient  $A \in \mathcal{M}_n(\mathbb{R})$  et  $\lambda \in Sp(A)$  où  $\lambda \in \mathbb{R}$ . Supposons que  $m_g(\lambda) = 1$  et  $m_a(\lambda) = 2$ , alors on a  $\dim(\text{Ker}(A - \lambda I)) = 1$  et  $\dim(\text{Ker}(A - \lambda I)^2) = 2$  (ici  $d = 2$  avec les notations du lemme 8). Remarquons que comme la dimension de l'espace propre associé à  $\lambda$  est égal à 1, on n'a pas deux vecteurs propres associés à  $\lambda$  linéairement indépendants. Soit  $\varphi$  un vecteur propre associé à  $\lambda$ ; on a donc  $\varphi \in \text{Ker}(A - \lambda I)$ . Une première solution de l'équation (8.1) est la fonction  $X_\varphi$  définie par  $X_\varphi(t) = e^{\lambda t} \varphi$  (pour tout  $t \in \mathbb{R}$ ). Comme  $\dim(\text{Ker}(A - \lambda I)^2) = 2$ , il existe un vecteur  $\psi \in \text{Ker}(A - \lambda I)^2$  indépendant de  $\varphi$ . Une deuxième solution de l'équation (8.1) s'obtient alors en posant

$$X_\psi(t) = e^{\lambda t} (\psi + t \underbrace{(A - \lambda I)\psi}_{\neq 0})$$

En effet,

$$\begin{aligned}(X_\psi(t))' &= e^{\lambda t}(\lambda\psi + \lambda t(A - \lambda I)\psi + (A - \lambda I)\psi) \\ &= e^{\lambda t}(A\psi + \lambda t(A - \lambda I)\psi)\end{aligned}$$

Or  $\psi \in \text{Ker}(A - \lambda I)^2$  et donc  $\lambda(A - \lambda I)\psi = A(A - \lambda I)\psi$ . On a donc

$$\begin{aligned}(X_\psi(t))' &= e^{\lambda t}(A\psi + tA(A - \lambda I)\psi) \\ &= Ae^{\lambda t}(\psi + t(A - \lambda I)\psi) \\ &= AX_\psi(t).\end{aligned}$$

**Question :** peut-on généraliser la procédure pour trouver la solution de cet exemple ?

La réponse est oui, l'exemple 20 peut être généralisé, voici comment. Soient  $A \in \mathcal{M}_n(\mathbb{R})$  et  $\lambda \in \text{Sp}(A)$  avec  $\lambda \in \mathbb{R}$  (le cas  $\lambda \in \mathbb{C}$  peut aussi être traité). Par le lemme 8, il existe  $d \geq 1$  tel que  $\dim(\text{Ker}(A - \lambda I)^d) = m_a(\lambda)$ . Soit  $\psi \in \text{Ker}(A - \lambda I)^d$ . On obtient une solution de l'équation (8.1) en prenant la fonction  $X$  définie par

$$X_\psi(t) = e^{\lambda t} \left[ \psi + t(A - \lambda I)\psi + \dots + \frac{t^{d-1}}{(d-1)!}(A - \lambda I)^{d-1}\psi \right].$$

Comme la dimension de  $\text{Ker}(A - \lambda I)^d$  est  $m_a(\lambda)$ , on construit ainsi  $m_a(\lambda)$  solutions linéairement indépendantes. Enfin, comme  $\sum_{\lambda \in \text{Sp}(A)} m_a(\lambda) = n$ , on construit ainsi  $n$  solutions linéairement indépendantes

de l'équation (8.1) ce qui permet de construire une matrice  $M(t)$  vérifiant (8.3).

Ce résultat peut être retrouvé en construisant une autre matrice  $M(t)$  possible à partir d'une définition de la fonction exponentielle de  $\mathcal{M}_n(\mathbb{R})$  dans  $\mathcal{M}_n(\mathbb{R})$ .

On rappelle que pour  $n = 1$  le système (8.1) se ramène à l'équation différentielle suivante,

$$x'(t) = ax(t), \quad t \in \mathbb{R},$$

avec  $a \in \mathbb{R}$ , dont la solution générale est  $x(t) = ce^{at}$ , avec  $c \in \mathbb{R}$ . L'idée est alors de définir une matrice notée  $e^{At}$  de manière à pouvoir écrire  $X(t) = e^{At}C^{-1}$ , avec  $C \in \mathbb{R}^n$  dans le cas  $n \geq 1$ . Dans ce cas, la matrice  $e^{At}$  est donc une matrice  $M(t)$  qui vérifie (8.3), mais ce n'est bien sûr pas la seule.)

**Définition 11** (Exponentielle de matrice). Soit  $A \in \mathcal{M}_n(\mathbb{R})$  de coefficients  $(a_{i,j})$ ,  $i, j \in \{1, \dots, n\}$ . On pose

$$e^A = \sum_{k=0}^{+\infty} \frac{A^k}{k!}.$$

On montre (voir lemme 9) que cette définition a un sens car cette série est convergente. Si  $t \in \mathbb{R}$  on définit le produit  $tA$  (qu'on notera aussi parfois  $At$  par abus de notation) par

$$\forall i, j \in \llbracket 1, n \rrbracket, \quad (tA)_{i,j} = (At)_{i,j} = ta_{i,j}$$

de sorte que

$$\forall t \in \mathbb{R}, \quad e^{tA} = e^{At} = \sum_{k=0}^{+\infty} \frac{(tA)^k}{k!} = \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!}.$$

1. Noter qu'on écrit  $X(t) = e^{At}C$  et non  $X(t) = Ce^{At}$ . En effet,  $e^{At} \in \mathcal{M}_n(\mathbb{R})$ , et  $C \in \mathbb{R}^n$  et donc  $Ce^{At}$  n'a aucun sens.

Quelles sont les questions à se poser à propos de cette définition ?

- Q1. La première question posée par la définition de  $e^A$  est de montrer que la série est bien convergente, c'est-à-dire de montrer que, pour chaque  $i, j \in \llbracket 1, n \rrbracket$ , la série (à valeurs réelles)  $\sum_{k=0}^{+\infty} \frac{(A^k)_{i,j}}{k!}$  est convergente.
- Q2. A-t-on  $(e^{tA})' = Ae^{tA}$  ? Noter que la dérivée de la fonction à valeurs matricielles  $t \mapsto e^{tA}$  est définie par les dérivées de chacune des composantes de  $e^{tA}$  (qui sont des fonctions à valeurs réelles).
- Q3. Si  $A, B \in \mathcal{M}_n(\mathbb{R})$ , a-t-on  $e^{A+B} = e^A e^B$  ?

### 8.2.1 Bien-fondé de la définition de l'exponentielle

Le lemme suivant répond à la question Q1.

**Lemme 9** (Convergence de la série). *Soit  $A \in \mathcal{M}_n(\mathbb{R})$  de coefficients  $(a_{i,j})$ ,  $i, j \in \{1, \dots, n\}$ . Alors, la série  $\sum_{k=0}^{+\infty} \frac{A^k}{k!}$  est convergente.*

*Démonstration.* Posons

$$a = \max\{|a_{i,j}|, i, j \in \llbracket 1, n \rrbracket\},$$

et montrons par récurrence que

$$\forall i, j \in \llbracket 1, n \rrbracket, \quad \forall k \in \mathbb{N}^*, \quad |(A^k)_{i,j}| \leq n^{k-1} a^k. \quad (8.4)$$

En effet,

- 1) d'après la définition de  $a$ , (8.4) est vérifiée au rang  $k = 1$ .
- 2) Supposons que l'hypothèse de récurrence (8.4) soit vérifiée jusqu'au rang  $k$ , et montrons qu'elle l'est encore au rang  $k + 1$ . Par inégalité triangulaire, et par hypothèse de récurrence, on a, pour tout  $i, j \in \llbracket 1, n \rrbracket$ ,

$$\begin{aligned} |(A^{k+1})_{i,j}| &\leq \sum_{l=0}^n |(A^k)_{i,l} A_{l,j}| \\ &\leq n n^{k-1} a^k a = n^k a^{k+1} \end{aligned}$$

ce qui termine la récurrence. On pose alors

$$u_k = \frac{n^{k-1} a^k}{k!} \geq \left| \frac{(A^k)_{i,j}}{k!} \right|.$$

On remarque alors que

$$\left| \frac{u_{k+1}}{u_k} \right| = \left| \frac{n^k a^{k+1}}{(k+1)!} \frac{k!}{n^{k-1} a^k} \right| = \left| \frac{na}{k+1} \right| \xrightarrow{k \rightarrow +\infty} 0,$$

ce qui prouve que la série  $\sum_{k=0}^{+\infty} \frac{A^k}{k!}$  est donc absolument convergente. □

On a ainsi montré que les matrices  $e^A$  et  $e^{tA}$  sont bien définies. On peut de plus remarquer que pour tout  $T > 0$ , la série  $\sum_{k=0}^{+\infty} \frac{t^k A^k}{k!}$  est uniformément convergente sur  $[-T, T]$ .

### 8.2.2 Dérivée de $t \mapsto e^{tA}$

Nous allons répondre à la question Q2 grâce au lemme suivant :

**Lemme 10** (Dérivée d'une limite).

Soit  $(F_n)_{n \in \mathbb{N}}$  une suite de fonctions appartenant à  $C^1(\mathbb{R}, \mathbb{R})$ . On suppose que

- 1)  $F_n$  converge simplement vers  $F$  (c'est-à-dire que pour tout  $s \in \mathbb{R}$ ,  $\lim_{n \rightarrow +\infty} F_n(s) = F(s)$ ),
- 2)  $F'_n$  converge localement uniformément vers  $G$ , c'est-à-dire que pour tout  $T > 0$ ,

$$\lim_{n \rightarrow +\infty} \sup_{-T \leq s \leq T} |F'_n(s) - G(s)| = 0.$$

Alors  $F \in C^1(\mathbb{R}, \mathbb{R})$  et  $F' = G$ .

*Démonstration.* On remarque tout d'abord que  $G$  est continue comme limite localement uniforme de fonctions continues. Puis, on fixe  $a \in \mathbb{R}$ , soit  $s \in \mathbb{R}$ , alors

$$F_n(s) - F_n(a) = \int_a^s F'_n(t) dt$$

et donc, quand  $n \rightarrow +\infty$ ,

$$F(s) - F(a) = \int_a^s G(t) dt$$

car  $F'_n$  converge uniformément vers  $G$  sur  $[a, s]$ . Ceci montre que  $F$  est de classe  $C^1$  et que  $F' = G$ .  $\square$

En appliquant le lemme 10 aux sommes partielles d'une série, on obtient le corollaire suivant :

**Corollaire 2** (Dérivée d'une série). Soient  $(f_k)_{k \in \mathbb{N}}$  une suite de fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$ , de classe  $C^1$ . On suppose que la série  $\sum_{k \geq 0} f_k$  est convergente et que la série des dérivées  $\sum_{k \geq 0} f'_k$  est localement uniformément convergente.

On pose  $f = \sum_{k \geq 0} f_k$ . Alors,  $f$  est de classe  $C^1$  et  $f' = \sum_{k \geq 0} f'_k$ .

La preuve de ce corollaire consiste simplement à appliquer le lemme 10 à  $F_n = \sum_{p=0}^n f_p$ .

La proposition suivante donne la réponse à la question Q2 :

**Proposition 5** (Dérivée de l'exponentielle). Soit  $A \in \mathcal{M}_n(\mathbb{R})$  et  $t \in \mathbb{R}$ , alors

$$(e^{tA})' = Ae^{tA}.$$

*Démonstration.* Appliquons le corollaire 2 à l'application  $f : t \mapsto e^{tA}$  et aux fonctions  $f_k : t \mapsto \frac{(tA)^k}{k!}$  (plus précisément à chaque composante de ces applications à valeurs vectorielles). Par définition,

$$e^{tA} = \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!},$$

et donc

$$(e^{tA})' = \sum_{k=1}^{+\infty} \frac{t^{k-1} A^k}{(k-1)!} = A \left( \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!} \right) = Ae^{tA}.$$

$\square$

Une conséquence de ce résultat est que la solution du problème de Cauchy :

$$X'(t) = AX(t) \quad (8.5a)$$

$$X(0) = X_0, \quad X_0 \in \mathbb{R}^n \quad (8.5b)$$

est  $X(t) = e^{tA}X_0$ . En effet,  $(e^{tA}X_0)' = A(e^{tA}X_0)$ , et on a bien  $X(0) = X_0$ .

L'espace vectoriel  $E_n$  des solutions de (8.1) est donc formé par les fonctions  $t \mapsto e^{tA}C$ , avec  $C \in \mathbb{R}^n$ . Autrement dit, en notant  $X^{(i)}(t)$  la  $i$ -ième colonne de  $e^{At}$  on a

$$e^{At} = \begin{bmatrix} X^{(1)}(t) & \dots & X^{(n)}(t) \end{bmatrix}.$$

La matrice  $e^{At}$  est l'une des matrices  $M(t)$  permettant d'écrire toutes les solutions de (8.1) sous la forme  $M(t)C$  avec  $C \in \mathbb{R}^n$ .

**Remarque 10** (La matrice exponentielle est inversible). Soit  $t \in \mathbb{R}$ . Comme les fonctions  $X^{(i)}$  sont linéairement indépendantes (et solutions de (8.1)), les vecteurs  $X^{(i)}(t)$  sont linéairement indépendants pour tout  $t \in \mathbb{R}$ , voir la proposition 4). En particulier, l'image de la matrice  $e^{At}$  (qui est l'espace vectoriel engendré par ses vecteurs colonnes) est de dimension  $n$ . La matrice  $e^{At}$  est donc inversible en vertu du théorème du rang<sup>2</sup>.

Pour conclure la réponse à la question Q2, on peut aussi remarquer que l'application  $t \mapsto e^{tA}$  est l'unique solution du problème différentiel à valeurs matricielles

$$M'(t) = AM(t), \quad t \in \mathbb{R}, \quad (8.6)$$

$$M(0) = I. \quad (8.7)$$

### 8.2.3 Produit de matrices exponentielles

Le lemme suivant répond à la question Q3.

**Lemme 11** (Sur le produit d'exponentielles de matrices). Soit  $A, B \in \mathcal{M}_n(\mathbb{R})$ , alors  $e^{A+B} = e^A e^B$  si  $AB = BA$ .

*Démonstration.* Supposons que  $AB = BA$ , alors :

- 1)  $A^k B = BA^k$  (par récurrence sur  $k$ ),
- 2)  $e^A B = B e^A$  (en utilisant la continuité du produit de matrices),
- 3) et donc  $\forall t \in \mathbb{R}, \quad e^{tA} B = B e^{tA}$  (en changeant  $A$  en  $tA$ )
- 4) Soient  $U : t \mapsto U(t)$  et  $V : t \mapsto V(t)$  des applications continues dérivables de  $\mathbb{R}$  dans  $\mathcal{M}_n(\mathbb{R})$  alors :  $(U(t)V(t))' = U(t)V'(t) + U'(t)V(t)$ . En effet, pour tous  $i, j \in \llbracket 1, n \rrbracket$  on a  $(U(t)V(t))_{i,j} = \sum_{\ell=1}^n U_{i,\ell}(t)V_{\ell,j}(t)$  et donc  $(U(t)V(t))'_{i,j} = \sum_{\ell=1}^n (U_{i,\ell}(t)V_{\ell,j}(t))'(t) = (U'(t)V(t))_{i,j} + (U(t)V'(t))_{i,j}$

En prenant  $U(t) = e^{tA}$  et  $V(t) = e^{tB}$  et en utilisant l'item 3 :

$$(e^{tA}e^{tB})' = Ae^{tA}e^{tB} + e^{tA}Be^{tB} \quad (8.8)$$

$$= Ae^{tA}e^{tB} + Be^{tA}e^{tB} \quad (8.9)$$

2. Pour toute matrice  $B \in \mathcal{M}_n(\mathbb{R})$ ,  $\dim(\text{Ker}(B)) + \dim(\text{Im}(B)) = n$ .

En posant  $X(t) = e^{At}e^{Bt}$  on a donc

$$\begin{aligned} X'(t) &= (A + B)X(t), \\ X(0) &= I, \end{aligned}$$

c'est-à-dire que l'application  $t \mapsto e^{tA}e^{tB}$  est solution de (8.6)-(8.7) avec  $A + B$  au lieu de  $A$ . Or, on sait que l'unique solution de ce problème est l'application  $t \mapsto e^{t(A+B)}$ . On donc  $e^{tA}e^{tB} = e^{t(A+B)}$  pour tout  $t$  et donc, pour  $t = 1$ ,  $e^Ae^B = e^{(A+B)}$ . □

On rappelle maintenant que l'objectif de cette section 8.2 est de construire une matrice  $M(t)$  telle que la solution générale de (8.1) soit l'ensemble des fonctions  $t \mapsto M(t)C$  avec  $C$  arbitraire dans  $\mathbb{R}^n$ . On a trouvé une matrice  $M(t)$  convenable en prenant  $M(t) = e^{tA}$ . Malheureusement, le calcul de  $e^{tA}$  à partir de la définition 11 n'est pas simple sauf dans des cas particuliers (comme  $A$  diagonale ou nilpotente). On va montrer maintenant que l'on peut trouver une autre matrice  $M(t)$  convenable, plus facile à calculer (et ceci généralise la construction de la solution dans l'exemple 20).

Soit  $\psi \in \mathbb{R}^n$ , l'application  $t \mapsto (e^{At}\psi)$  est solution de (8.1) (c'est la solution qui vaut  $\psi$  en 0). Le lemme 12 nous permet de calculer cette application pour  $n$  vecteurs  $\psi$  particuliers (et ceci nous donnera une nouvelle matrice  $M(t)$ ).

**Lemme 12.** Soient  $A \in \mathcal{M}_n(\mathbb{R})$  et  $\lambda \in Sp(A)$ . On suppose (pour simplifier) que  $\lambda \in \mathbb{R}$ . On note  $d$  le plus petit entier tel que  $\dim(\text{Ker}(A - \lambda I)^d) = m_a(\lambda)$  (voir lemme 8). Soit  $\psi \in \text{Ker}(A - \lambda I)^d$  alors

$$e^{At}\psi = e^{\lambda t} \left[ \psi + t(A - \lambda I)\psi + \dots + \frac{t^{d-1}}{(d-1)!} (A - \lambda I)^{d-1}\psi \right].$$

*Démonstration.* Comme les matrices  $\lambda I$  et  $A - \lambda I$  commutent et comme la multiplication par la matrice  $e^{\lambda I}$  est identique à la multiplication par le scalaire  $e^{\lambda t}$ , on a

$$\begin{aligned} e^{At}\psi &= e^{At+\lambda It-\lambda It}\psi \\ &= e^{\lambda t} e^{(A-\lambda I)t}\psi \\ &= e^{\lambda t} \left[ \psi + t(A - \lambda I)\psi + \dots + \frac{t^{d-1}}{(d-1)!} (A - \lambda I)^{d-1}\psi \right]. \end{aligned}$$

□

Le lemme 12 est intéressant pour  $m_g(\lambda) < m_a(\lambda)$ , il permet de calculer  $e^{At}\psi$  sans calculer  $e^{At}$  (si  $m_g(\lambda) = m_a(\lambda)$ , on peut prendre  $d = 1$ , on retrouve les solutions des paragraphes 8.1.1 et 8.1.2). Si  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ , Le même lemme est vrai pour des solutions à valeurs dans  $\mathbb{C}^n$ . En prenant, comme dans le paragraphe 8.1.2 les parties réelle et imaginaire, on trouve des solutions à valeurs dans  $\mathbb{R}^n$ .

Grâce au lemme 12, pour tout  $\lambda \in Sp(A)$ , on construit ainsi  $m_a(\lambda)$  solutions linéairement indépendantes de (8.1). Et comme  $\sum_{\lambda \in Sp(A)} m_a(\lambda) = n$ , on obtient  $n$  solutions linéairement indépendantes de (8.1) et

donc une matrice  $M(t)$  vérifiant (8.3) permettant d'écrire que l'espace  $E_n$  des solutions de (8.1) est  $E_n = \{t \mapsto M(t)C, C \in \mathbb{R}^n\}$ .

### 8.3 Calcul de l'exponentielle d'une matrice

Le calcul de  $e^A$  peut être fait avec la définition 11 lorsque la matrice  $A$  est "simple", par exemple diagonale ou nilpotente. Il est difficile, avec la définition 11, dans le cas général. A vrai dire, le calcul de  $e^A$  (ou de  $e^{tA}$ ) n'est pas utile pour la résolution du système (8.1). Ce qui est utile est de trouver  $n$  solutions linéairement indépendantes de (8.1) comme expliqué dans la section 8.2.

Mais, dès que nous avons  $n$  solutions linéairement indépendantes de (8.1), il est facile de calculer  $e^{tA}$  et donc  $e^A$ . En effet, soient  $X^{(1)}, \dots, X^{(n)}$ ,  $n$  des solutions indépendantes de (8.1). Alors l'espace vectoriel  $E_n$  des solutions de (8.1) est l'ensemble  $\{t \mapsto M(t)C, C \in \mathbb{R}^n\}$  des combinaisons linéaires des colonnes de la matrice  $M(t)$ , où  $M(t)$  est la matrice dont la  $i$ -ème colonne est  $X^{(i)}(t)$ . La solution de (8.5) est alors  $X(t) = M(t)C$  avec  $M(0)C = X_0$ , c'est-à-dire  $C = M(0)^{-1}X_0$  (on rappelle que  $M(t)$  est inversible pour tout  $t$ ). Mais cette solution est aussi  $X(t) = e^{At}X_0$ . On a donc  $e^{At} = M(t)M^{-1}(0)$  pour tout  $t \in \mathbb{R}$  et donc  $e^A = M(1)M^{-1}(0)$ .

### 8.4 Résolution des systèmes non homogènes

Soit le système suivant :

$$X'(t) = AX(t) + G(t) \quad (8.10)$$

où  $G \in \mathcal{C}(\mathbb{R}, \mathbb{R}^n)$  et  $A \in \mathcal{M}_n(\mathbb{R})$ . Comme dans le cas  $n = 1$ , nous allons procéder en 2 étapes pour résoudre ce problème,.

Etape 1 On résout d'abord l'équation homogène associée à (8.10), c'est-à-dire lorsque  $G = 0$ . On trouve une matrice  $M(t)$  permettant d'exprimer à l'instant  $t$  toutes les solutions du système homogène, qui vérifie donc (8.3).

Etape 2 On recherche ensuite une solution particulière en devinant sa forme (méthode 1) ou en utilisant la méthode de la variation de la constante (méthode 2) consistant à poser  $X(t) = M(t)C(t)$ ,  $C(t) \in \mathbb{R}^n$ . Comme d'habitude, cette deuxième méthode consiste à réduire l'ordre du système, c'est-à-dire ici à passer d'un système d'ordre 1 (sur l'inconnue  $X$ ) à un système d'ordre 0 sur la nouvelle fonction inconnue  $C'$ . En effet, si  $X(t) = M(t)C(t)$ , alors  $X'(t) = M'(t)C(t) + M(t)C'(t) = AM(t)C(t) + M(t)C'(t)$ . Et donc  $X$  est solution si  $M(t)C'(t) = G(t)$ , c'est-à-dire  $C'(t) = M(t)^{-1}G(t)$ .

Cette dernière formule donne  $C(t) = C(0) + \int_0^t M(s)^{-1}G(s)ds$ . Puis comme  $X(t) = M(t)C(t)$ ,  $M(0)C(0) = X_0$  et  $e^{At} = M(t)M^{-1}(0)$ , on a pour tout  $t$

$$X(t) = M(t)C(0) + M(t) \int_0^t M(s)^{-1}G(s)ds,$$

$$X(t) = M(t)M(0)^{-1}X_0 + \int_0^t M(t)M(x)^{-1}G(s)ds,$$

$$X(t) = e^{At}X_0 + \int_0^t e^{A(t-s)}G(s)ds.$$

Cette dernière formule s'appelant "formule de Duhamel".

Nous donnons ci après deux exemples.

**Exemple 21** (Un système non homogène très simple). Soit  $n = 2$ , on considère le système :

$$X'(t) = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} X(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{2t}$$

Les valeurs propres de la matrice en question sont  $\lambda_1 = 1 - \sqrt{2}$  et  $\lambda_2 = 1 + \sqrt{2}$ . La solution du système homogène est donc  $X_h(t) = c_1 e^{\lambda_1 t} \phi_1 + c_2 e^{\lambda_2 t} \phi_2$  où  $\phi_1$  (resp.  $\phi_2$ ) est un vecteur propre associé à  $\lambda_1$  (resp.  $\lambda_2$ ). On cherche alors une solution particulière sous la forme  $X_p(t) = \begin{bmatrix} a \\ b \end{bmatrix} e^{2t}$ . Par identification, on obtient  $a = -1$  et  $b = -2$ . La solution générale est donc

$$X(t) = X_h(t) + X_p(t) = c_1 e^{\lambda_1 t} \phi_1 + c_2 e^{\lambda_2 t} \phi_2 + \begin{bmatrix} a \\ b \end{bmatrix} e^{2t}.$$

**Exemple 22** (Un système non homogène un peu moins simple). Soit

$$X'(t) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} X(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^t \quad (8.11)$$

On pose  $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$  qui peut être décomposée comme suit :

$$A = I + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

et donc

$$e^{tA} = e^{(I+B)t} = e^{tI} e^{tB},$$

car les matrices  $I$  et  $B$  commutent. Comme  $B^2$  est la matrice nulle, on a :

$$e^{tB} = \sum_{k=0}^{+\infty} \frac{t^k B^k}{k!} = I + tB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 2t & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2t & 1 \end{bmatrix}$$

On peut alors en déduire que :

$$e^{(I+B)t} = \begin{bmatrix} 1 & 0 \\ 2t & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^t \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ 2te^t & e^t \end{bmatrix}$$

En posant

$$M(t) = e^{(I+B)t} = \begin{bmatrix} e^t & 0 \\ 2te^t & e^t \end{bmatrix}$$

La solution générale du système homogène est donc  $M(t)C$ , avec  $C$  arbitraire dans  $\mathbb{R}^2$ .

Pour trouver la solution générale du système non homogène nous allons procéder avec les deux méthodes proposées (pour les comparer).

**Solution particulière** Ici comme 1 est valeur propre algébriquement double et géométriquement simple de la matrice du système et que le terme non homogène est  $e^t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  (on cumule deux difficultés), on est amené à chercher une solution particulière sous la forme :

$$X_p(t) = \begin{bmatrix} a \\ b \end{bmatrix} e^t + t \begin{bmatrix} c \\ d \end{bmatrix} e^t + t^2 \begin{bmatrix} e \\ f \end{bmatrix} e^t$$

En écrivant que  $X_p$  est solution du système (8.11), on trouve  $c = 1$ ,  $2a = d$ ,  $e = 0$  et  $f = c$ . Une solution possible est donc  $a = b = d = e = 0$ ,  $c = f = 1$ , c'est-à-dire

$$X_p(t) = \begin{bmatrix} t \\ t^2 \end{bmatrix} e^t$$

La solution générale du système non homogène est alors  $X_p(t) + M(t)C$  avec  $C$  arbitraire dans  $\mathbb{R}^2$ .

**Variation de la constante** On cherche  $X_p$  sous la forme  $X_p(t) = M(t)C(t)$ . L'équation sur  $C'$  est alors  $C'(t) = M(t)^{-1}G(t)$  où  $G$  est le terme non homogène du système.

Comme

$$M(t) = \begin{bmatrix} e^t & 0 \\ 2te^t & e^t \end{bmatrix} = e^t \begin{bmatrix} 1 & 0 \\ 2t & 1 \end{bmatrix},$$

un calcul simple donne

$$M(t)^{-1} = e^{-t} \begin{bmatrix} 1 & 0 \\ -2t & 1 \end{bmatrix}.$$

et donc

$$C'(t) = \begin{bmatrix} 1 & 0 \\ -2t & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2t \end{bmatrix}.$$

Une solution est donc

$$C(t) = \begin{bmatrix} t \\ -t^2 \end{bmatrix}$$

c'est-à-dire

$$X_p(t) = M(t)C(t) = \begin{bmatrix} t \\ t^2 \end{bmatrix} e^t.$$

On retrouve, et c'est heureux, la même solution générale du système (8.11) que dans la première méthode, c'est-à-dire

$$X(t) = \begin{bmatrix} c_1 + t \\ 2tc_1 + c_2 + t^2 \end{bmatrix} e^t,$$

avec  $c_1, c_2$  arbitraires dans  $\mathbb{R}$ .

## Cours 9. Systèmes non linéaires

Soit  $I$  est un intervalle ouvert de  $\mathbb{R}$ ,  $f \in \mathcal{C}(I \times \mathbb{R}^n, \mathbb{R}^n)$ ,  $n \geq 1$ ; on cherche des solutions  $X \in \mathcal{C}^1(I, \mathbb{R}^n)$  du système différentiel non linéaire suivant :

$$X'(t) = f(t, X(t)), \quad t \in I, \quad (9.1)$$

auquel on peut associer un problème de Cauchy en lui adjoignant une condition initiale :

Soit  $t_0 < T \leq +\infty$ ,  $I = [t_0, T[$ ,  $n \geq 1$ ,  $f \in \mathcal{C}(I \times \mathbb{R}^n, \mathbb{R}^n)$  et  $X^{(0)} \in \mathbb{R}^n$ . On cherche alors  $X \in \mathcal{C}([t_0, T[, \mathbb{R}^n) \cap \mathcal{C}^1(]t_0, T[, \mathbb{R}^n)$  solution du problème de Cauchy :

$$X'(t) = f(t, X(t)), \quad t_0 < t < T, \quad (9.2a)$$

$$X(t_0) = X^{(0)}. \quad (9.2b)$$

On donne ci-après deux exemples qui seront étudiés en TD et TP.

**Exemple 23** (Modèle de Lotka-Volterra). Soient  $a > 0$ ,  $b > 0$ ,  $c > 0$ ,  $d > 0$ , le modèle prédateur-proie de Lotka-Volterra s'écrit

$$x_1'(t) = ax_1(t) - bx_1(t)x_2(t), \quad t > 0, \quad (9.3a)$$

$$x_2'(t) = -cx_2(t) + dx_1(t)x_2(t), \quad t > 0. \quad (9.3b)$$

L'inconnue est la fonction  $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  de  $\mathbb{R}_+$  à valeurs dans  $\mathbb{R}^2$ . L'inconnue  $x_1$  désigne la population de proies, qui a un taux de croissance de  $a$  en l'absence des prédateurs  $x_2$  alors que la population  $x_2$  dépend pour sa survie des proies  $x_1$  et a un taux de décroissance  $c$  en leur absence.

Dans le cas général et pour la plupart des conditions initiales, on ne sait pas calculer de solution explicite de ce système. On peut toutefois l'effectuer de manière approchée à l'aide d'un schéma numérique (voir TP4).

**Exemple 24** (Le pendule pesant non linéarisé). Le modèle est ici

$$x''(t) + \sin(x(t)) = 0, \quad t > 0. \quad (9.4)$$

L'équation (9.4) peut s'écrire sous la forme d'un système différentiel non linéaire. On pose  $y_1(t) = x(t)$  et  $y_2(t) = x'(t)$ . On a donc

$$y_1'(t) = y_2(t), \quad t > 0, \quad (9.5a)$$

$$y_2'(t) = -\sin(y_1(t)), \quad t > 0. \quad (9.5b)$$

Là encore, on ne sait pas en général calculer de solution explicite de ce système, mais la résolution de ce système peut aussi s'effectuer de manière approchée à l'aide d'un schéma numérique (voir le projet de fin d'UE).

Les questions relatives à l'étude des systèmes linéaires se posent également pour les systèmes non linéaires :

- 1) La solution existe-elle ? est-elle unique ? est-elle globale ?
- 2) Quels sont les points d'équilibre du système ?
- 3) Qu'en est-il de la stabilité de la solution ?
- 4) Quel est le comportement de la solution lorsque  $t \rightarrow \infty$  ?
- 5) Comment peut-on calculer les solutions ?

Nous allons y répondre dans les sections suivantes.

## 9.1 Existence, unicité, solution maximale

Considérons le problème de Cauchy en ajoutant l'hypothèse que  $f$  est localement lipschitzienne par rapport à son deuxième argument, c'est-à-dire que

$$\forall \eta : t_0 < t_0 + \eta < T, \forall A \in \mathbb{R}_+^*, \exists C_{\eta,A} : \|f(t, X) - f(t, Y)\| \leq C_{\eta,A} \|X - Y\| \\ \forall t \in [t_0, t_0 + \eta], \quad \forall X, Y \in B_A = \{z \in \mathbb{R}^n, \|z\| \leq A\}.$$

Dans cette définition  $\|\cdot\|$  désigne une norme sur  $\mathbb{R}^n$ . La définition ne dépend pas de la norme choisie (car toutes les normes sur  $\mathbb{R}^n$  sont équivalentes). Un choix possible (et c'est ce choix que nous ferons dans la suite) consiste à prendre :

$$\text{Pour } X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \|X\| = \max\{|x_i|, i \in \{1, \dots, n\}\}.$$

L'existence locale de la solution est donnée par le théorème de Cauchy-Lipschitz, dont la démonstration est identique au cas  $n = 1$  en remplaçant  $|\cdot|$  par  $\|\cdot\|$ , voir section 3.1.

**Théorème 8** (Existence locale : Cauchy-Lipschitz). *Soit :*

$$\begin{cases} I = [t_0, T[ \quad t_0 < T \leq +\infty, \\ f \in \mathcal{C}(I \times \mathbb{R}^n, \mathbb{R}^n). \end{cases} \quad (9.6)$$

*On suppose que  $f$  est localement lipschitzienne par rapport à son deuxième argument. Soit  $X^{(0)} \in \mathbb{R}^n$ . Alors il existe  $\alpha > 0$  et il existe une fonction  $X \in \mathcal{C}([t_0, t_0 + \alpha[, \mathbb{R}^n) \cap \mathcal{C}^1([t_0, t_0 + \alpha[, \mathbb{R}^n)$  telle que*

$$X'(t) = f(t, X(t)), \quad \forall t \in ]t_0, t_0 + \alpha[ \quad (9.7a)$$

$$X(t_0) = X^{(0)}. \quad (9.7b)$$

*Il y a donc existence locale d'une solution.*

Il y a aussi unicité de la solution par le théorème d'unicité énoncé ci-dessous. Là encore, la démonstration est identique au cas  $n = 1$ , en remplaçant  $|\cdot|$  par  $\|\cdot\|$  (voir théorème 2 et section 3.2).

**Théorème 9** (Unicité). *Sous les hypothèses du théorème 8, pour tout  $\alpha > 0$  tel que  $t_0 + \alpha \leq T$ , le problème (9.7) admet au plus une solution sur l'intervalle  $[t_0, t_0 + \alpha[$ .*

De même, l'existence de la solution maximale se démontre comme dans le cas  $n = 1$ .

**Théorème 10** (Solution maximale). *Toujours sous les hypothèses du théorème 8, le problème de Cauchy (9.7) admet une unique solution maximale  $X$  définie sur  $[t_0, T_m[$ ,  $t_0 < T_m \leq T$ . Par définition,  $T_m$  s'appelle le temps maximal d'existence, et il est défini comme le maximum des temps  $t_0 + \alpha$  tel que le problème (9.7) admette une solution. (Le problème n'admet donc pas de solution sur  $[t_0, \tilde{T}[$  si  $\tilde{T} > T_m$ .)  
De plus, Si  $T_m < T$  alors  $\lim_{t \rightarrow T_m} \|X(t)\| = +\infty$ .*

Dans le cas de l'exemple 23 du système de Lotka-Volterra, si on rajoute au système différentiel (9.3) les conditions initiales

$$x_1(0) = x_1^{(0)}, \quad (9.8)$$

$$x_2(0) = x_2^{(0)}, \quad (9.9)$$

les théorèmes d'existence locale et d'unicité s'appliquent assez facilement. Il existe donc une solution maximale. L'existence globale est vraie mais demande un peu plus de travail. (Voir l'examen de janvier 2020.)

Voyons maintenant l'exemple 24 du pendule avec  $I = [0, +\infty[$  et une donnée initiale sur  $y_1$  et  $y_2$  en 0. On obtient là encore, par application des théorèmes précédents, existence locale et unicité. Il existe donc une solution maximale. Reste à savoir si la solution est globale. On note  $T_m$  le temps d'existence de la solution maximale.

On rappelle que  $y_2' = -\sin(y_1)$ , et donc pour  $0 \leq t \leq T_m$ ,  $y_2(t) = y_2(0) + \int_0^t -\sin(y_1(s)) ds$ .

On en déduit que pour tout  $0 \leq t \leq T_m$ , on a  $y_2(0) - t \leq y_2(t) \leq y_2(0) + t$ , et donc  $|y_2(t)| \leq |y_2^{(0)}| + t$ . Ceci montre que, si  $T_m < +\infty$ , la fonction  $y_2$  est bornée. Comme  $y_1' = y_2$ , la fonction  $y_1$  est aussi bornée. On en déduit (toujours si  $T_m < +\infty$ ) que la solution n'explose pas. On en déduit donc (par le théorème 10 que la solution est globale (c'est-à-dire  $T_m = +\infty$ ).

Lorsque la fonction  $f$  est définie sur un intervalle ouvert  $]S, T[$  contenant  $t_0$ , on peut aussi introduire (comme nous l'avons vu lors d'un cours précédent) le problème de Cauchy "à gauche" consistant à chercher des solutions au problème (9.7) sur  $]S, t_0[$  au lieu de  $]t_0, T[$ , avec une donnée en  $t_0$ . En regroupant ces deux problèmes (problème de Cauchy classique et problème de Cauchy à gauche) on obtient des résultats d'existence locale et d'unicité pour le problème de Cauchy général. On est surtout intéressé pour la suite par le théorème d'unicité que l'on donne maintenant.

**Théorème 11** (Unicité, problème de Cauchy général). *Soient  $I$  un intervalle ouvert de  $\mathbb{R}$  et  $f \in \mathcal{C}(I \times \mathbb{R}^n, \mathbb{R}^n)$ . On suppose que  $f$  est localement lipschitzienne par rapport à son deuxième*

argument, c'est-à-dire que pour tout  $J$  intervalle compact inclus dans  $I$  et tout  $A \in \mathbb{R}$ , il existe  $C_{J,A}$  t.q.

$$\|f(t, x) - f(t, y)\| \leq C_{J,A} \|x - y\|, \quad \forall t \in J, \forall x, y \in B_A = \{z \in \mathbb{R}^n, \|z\| \leq A\}.$$

Soient  $\alpha, \beta > 0$ ,  $t_0 - \alpha, t_0 + \beta \in I$  et  $X^{(0)} \in \mathbb{R}^n$ . Alors, le problème (9.7) admet au plus une solution sur l'intervalle  $]t_0 - \alpha, t_0 + \beta[$  (c'est-à-dire avec, dans (9.7a),  $t \in ]t_0 - \alpha, t_0 + \beta[$  au lieu de  $t \in ]t_0, t_0 + \alpha[$ ).

On donne maintenant deux corollaires de ce théorème d'unicité (théorème 11) portant sur le fait que les trajectoires des solutions de (9.1) ne se rencontrent pas. On attire l'attention sur le fait que le terme "trajectoire" est employé, dans ces deux corollaires, en deux sens distincts. La premier sens est celui qu'on a déjà vu pour les équations différentielles non linéaires : la trajectoire de la solution  $y$  est le graphe de  $y$ , c'est-à-dire de l'application  $t \mapsto y(t)$  (avec  $t \in I \subset \mathbb{R}$  et  $y(t) \in \mathbb{R}^n$ ).

**Corollaire 3** (Les trajectoires ne se rencontrent pas). Soit  $I$  un intervalle ouvert de  $\mathbb{R}$  et  $\bar{t} \in I$ . On suppose que  $f$  vérifie les hypothèses du théorème 11. Soit  $X, Y \in \mathcal{C}^1(I, \mathbb{R}^n)$  solutions de (9.1). On suppose  $X(\bar{t}) = Y(\bar{t})$ , alors :

$$\forall t \in I, \quad X(t) = Y(t)$$

*Démonstration.* Le corollaire 3 est une conséquence immédiate du théorème 11 en prenant  $t_0 = \bar{t}$ .  $\square$

Le deuxième corollaire concerne le cas autonome, c'est-à-dire le cas du système (9.1) lorsque la fonction  $f$  ne dépend pas de  $t$  (et donc  $f$  est une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$ ) et concerne la notion de trajectoire comme courbe paramétrée par le temps.

**Définition 12** (Trajectoire dans  $\mathbb{R}^n$ ). Soit  $I$  un intervalle de  $\mathbb{R}$  et  $X$  une solution de (9.1). On appelle trajectoire de  $X$  dans l'espace  $\mathbb{R}^n$  (ou encore "trajectoire dans  $\mathbb{R}^n$ "), qu'on note  $\gamma(X)$ , la courbe paramétrée  $t \in I \mapsto X(t)$  :

$$\gamma(X) = \{X(t), t \in I\},$$

qui est bien une courbe dans l'espace  $\mathbb{R}^n$ .

**Remarque 11** (Orbite). En mécanique, lorsque la trajectoire dans  $\mathbb{R}^n$  de la solution  $X$  est fermée, l'ensemble  $\gamma(X)$  est aussi appelé "orbite" de  $X$ .

**Corollaire 4** (Trajectoires dans  $\mathbb{R}^n$  pour les systèmes autonomes). Soient  $X$  et  $Y$  deux solutions de (9.1) avec  $I = \mathbb{R}$ ,  $f$  indépendante de  $t$  (le système est donc autonome) et  $f$  localement lipschitzienne. On suppose qu'il existe  $t_1, t_2 \in \mathbb{R}$  tels que  $X(t_1) = Y(t_2)$  alors

$$\{X(t), t \in \mathbb{R}\} = \{Y(t), t \in \mathbb{R}\},$$

c'est-à-dire  $\gamma(X) = \gamma(Y)$ .

Le corollaire 4 montre que deux trajectoires dans  $\mathbb{R}^n$  ne peuvent pas se rencontrer : si deux trajectoires se rencontrent en un point alors les deux trajectoires sont confondues.

*Démonstration.* Soient  $u$  et  $v$  les fonctions de  $\mathbb{R}$  dans  $\mathbb{R}^n$  définies par

$$t \in \mathbb{R} \mapsto u(t) = X(t + t_1), \quad t \in \mathbb{R} \mapsto v(t) = Y(t + t_2).$$

On a  $u'(t) = X'(t + t_1) = f(X(t + t_1)) = f(u(t))$ ,  $v'(t) = Y'(t + t_2) = f(Y(t + t_2)) = f(v(t))$ . De plus,  $u(0) = X(t_1) = Y(t_2) = v(0)$ . Les fonctions  $u$  et  $v$  sont donc solutions de (9.1) et prennent la même valeur en 0. Le théorème d'unicité (théorème 11) ou directement son corollaire (Corollaire 3) nous donne alors  $u(t) = v(t)$  pour tout  $t \in \mathbb{R}$ . On a donc  $\gamma(u) = \gamma(v)$  et comme  $\gamma(u) = \gamma(X)$  et  $\gamma(v) = \gamma(Y)$  on a bien montré que  $\gamma(X) = \gamma(Y)$ .  $\square$

**Remarque 12** (Périodicité). *Dans le corollaire 4, si  $X = Y$  et  $t_2 > t_1$  on obtient que  $X$  est périodique de période  $t_2 - t_1$  car, dans la démonstration, on a prouvé  $u(t) = v(t)$  pour tout  $t \in \mathbb{R}$ , c'est-à-dire  $X(t + t_1) = Y(t + t_2)$  pour tout  $t \in \mathbb{R}$  et donc  $X(t) = X(t + t_2 - t_1)$  pour tout  $t \in \mathbb{R}$ .*

**Remarque 13.** *Dans le corollaire 4, on peut remplacer  $I = \mathbb{R}$  par  $I = [t_0, +\infty[$ . En supposant  $t_2 \geq t_1$ , dans la démonstration du corollaire on obtient  $u(t) = v(t)$  pour tout  $t \geq t_0 - t_1$  (la fonction  $u$  n'est pas définie pour  $t < t_0 - t_1$ ). La conclusion est donc que  $X(t + t_1) = Y(t + t_2)$  pour tout  $t \geq t_0 - t_1$ , c'est-à-dire que  $X(t) = Y(t + t_2 - t_1)$  pour tout  $t \geq t_0$ . En particulier ceci donne  $\gamma(X) \subset \gamma(Y)$ .*

## 9.2 Stabilité des systèmes autonomes

Dans cette section on considère que  $f$  ne dépend pas de  $t$  et est toujours localement lipschitzienne de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$ . On considère alors le système suivant

$$X'(t) = f(X(t)), \tag{9.10}$$

et le problème de Cauchy associé, pour  $X^{(0)} \in \mathbb{R}^n$ ,

$$X'(t) = f(X(t)), \quad 0 < t < +\infty, \tag{9.11a}$$

$$X(0) = X^{(0)}. \tag{9.11b}$$

Soit  $a \in \mathbb{R}^n$ . Le point  $a$  est point d'équilibre du système  $X'(t) = f(X(t))$  si la fonction constante et égale à  $a$  est solution de ce système. Elle est donc solution de (9.11) avec  $X^{(0)} = a$ . (On dit aussi que la fonction constante et égale à  $a$  est une solution stationnaire du système.)

Bien sûr, une condition nécessaire et suffisante pour que le point  $a$  soit point d'équilibre est que  $f(a) = 0$ .

**Définition 13.** *Soit  $a \in \mathbb{R}^n$  un point d'équilibre du système (9.10).*

1) *On dit que  $a$  est un point d'équilibre (uniformément) stable de (9.10) si pour tout  $\varepsilon > 0$  il existe  $\delta > 0$  tel que*

$$\|X^{(0)} - a\| \leq \delta \Rightarrow \begin{cases} \text{Le problème (9.11) a une solution globale notée } X \\ \text{et } \sup_{t \in [0, +\infty[} \|X(t) - a\| \leq \varepsilon. \end{cases}$$

2) *On dit que  $X_0$  est instable si  $X_0$  n'est pas stable.*

3) *On dit que  $X_0$  est asymptotiquement stable si  $X_0$  est stable et si il existe  $\delta > 0$  tel que*

$$\|X^{(0)} - a\| \leq \delta \Rightarrow \begin{cases} \text{Le problème (9.11) a une solution globale notée } X \\ \text{et } \lim_{t \rightarrow +\infty} \|X(t) - a\| = 0. \end{cases}$$

Dire que le point d'équilibre est uniformément stable revient à dire que si l'on part suffisamment proche de ce point d'équilibre, on reste proche de ce point d'équilibre (c'est l'exemple du pendule).

Dire que le point d'équilibre est asymptotiquement stable revient à dire que si l'on part suffisamment proche de ce point d'équilibre, on tend vers ce point d'équilibre quand  $t \rightarrow +\infty$  (c'est l'exemple du pendule amorti).

### 9.2.1 Systèmes autonomes linéaires

Soit  $A \in \mathcal{M}_n(\mathbb{R})$ ; on considère le système suivant :

$$X'(t) = AX(t), \quad t > 0 \quad (9.12)$$

La fonction identiquement nulle est solution de (9.12). Autrement dit 0 est un point (de  $\mathbb{R}^n$ ) d'équilibre pour le système (9.12) (on dit aussi que 0 est un point stationnaire). Cette solution est-elle stable ou instable ? La réponse est apportée par le théorème suivant, dont la preuve sera donnée au prochain cours.

**Théorème 12** (Stabilité d'un système différentiel linéaire). Soit  $A \in \mathcal{M}_n(\mathbb{R})$ . On note  $Sp(A)$  l'ensemble des valeurs propres de  $A$ . Si  $\lambda$  est valeur propre de  $A$ , on note  $\Re(\lambda)$  sa partie réelle. On distingue plusieurs cas :

- 1) Si  $\forall \lambda \in Sp(A), \Re(\lambda) < 0$ , alors 0 est stable et même asymptotiquement stable.
- 2) Si  $\exists \lambda \in Sp(A)$  tel que  $\Re(\lambda) > 0$ , alors 0 est instable.
- 3) Si  $\forall \lambda \in Sp(A), \Re(\lambda) \leq 0$  et il existe  $\lambda \in Sp(A)$  avec  $\Re(\lambda) = 0$ , alors on distingue deux sous-cas (on rappelle que  $m_a(\lambda)$  (resp.  $m_g(\lambda)$ ) désigne la multiplicité algébrique (resp. géométrique) d'une valeur propre  $\lambda$ , voir définition 10) :
  - (a) Si  $\forall \lambda \in Sp(A)$  tel que  $\Re(\lambda) = 0$  on a  $m_a(\lambda) = m_g(\lambda)$ , alors 0 est stable, mais non asymptotiquement stable;
  - (b) Si  $\exists \lambda \in Sp(A)$  tel que  $\Re(\lambda) = 0$  et  $m_a(\lambda) > m_g(\lambda)$ , alors 0 est instable.

**Exemple 25** (Un cas  $2 \times 2$ ). Soit le système  $\begin{cases} y_1'(t) = y_2(t), & t > 0 \\ y_2'(t) = -y_1(t), & t > 0, \end{cases}$  qui s'écrit  $Y' = AY$  avec

$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ . Le polynôme caractéristique associé à  $A$  est  $P_A(X) = X^2 + 1$ . Les valeurs propres sont donc  $\lambda = \pm i$ . On est dans le cas 3a du théorème 12, le point d'équilibre 0 est stable mais n'est pas asymptotiquement stable.

### 9.2.2 Systèmes autonomes non linéaires

Nous allons étudier la stabilité d'un point d'équilibre du système non-linéaire suivant

$$X'(t) = f(X(t)), \quad t > 0. \quad (9.13)$$

On suppose que  $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$  et  $f(a) = 0$  pour un point  $a \in \mathbb{R}^n$ . (A vrai dire, on a simplement besoin dans la suite que  $f$  soit de classe  $C^1$  dans un voisinage de  $a$ ).

Le point d'équilibre  $a$  est-il stable ou instable ?

D'après la définition de la dérivée d'une fonction, si  $n = 1$ ,

$$f(x) = f(a) + (x - a)f'(a) + |x - a|\varepsilon(x),$$

avec  $\lim_{x \rightarrow a} \varepsilon(x) = 0$ .

Si  $n > 1$  on a, de manière similaire,

$$f(X) = f(a) + A(X - a) + \|X - a\|\varepsilon(X),$$

où  $\lim_{X \rightarrow a} \varepsilon(X) = 0$  et  $A$  est la matrice jacobienne de  $f$  au point  $a$ .

(On rappelle que l'application  $y \mapsto Ay$  est la dérivée de  $f$  au point  $a$ .)

Prenons par exemple  $n = 2$ . On note  $f_1$  et  $f_2$  les deux composantes de  $f$  et  $x_1$  et  $x_2$  les deux composantes de  $X$ , de sorte que

$$f(X) = \begin{bmatrix} f_1(X) \\ f_2(X) \end{bmatrix}, \quad A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) \end{bmatrix}$$

On note  $A = J_f(a)$  la matrice jacobienne en  $a$  de  $f$ .

Si  $a$  est un point d'équilibre de (9.13), on a  $f(a) = 0$  et

$$f(X) = \underbrace{f(a)}_{=0} + A(X - a) + \|X - a\|\varepsilon(X)$$

On pose  $X - a = Y$ , on a alors :

$$f(Y + a) = AY + Y\varepsilon(Y), \quad \lim_{Y \rightarrow 0} \varepsilon(Y) = 0$$

Le système (9.13) devient, avec  $g(Y) = Y\varepsilon(Y)$ ,

$$Y'(t) = AY(t) + g(Y(t)). \tag{9.14}$$

Ceci permet de ramener l'étude de la stabilité du point  $a$  pour le système (9.13) à l'étude de la stabilité de 0 pour le système (9.14).

**Théorème 13** (Stabilité du système non linéaire par l'étude du système linéarisé).

Soit  $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$  et  $a \in \mathbb{R}^n$  tel que  $f(a) = 0$ . On considère le système différentiel non linéaire autonome (9.13). On note  $A = J_f(a)$  la matrice jacobienne de  $f$  en  $a$ .

- 1) Si  $\forall \lambda \in Sp(A)$ ,  $\Re(\lambda) < 0$ , alors  $a$  est asymptotiquement stable.
- 2) Si  $\exists \lambda \in Sp(A)$  tel que  $\Re(\lambda) > 0$ , alors  $a$  est instable.
- 3) Si  $\forall \lambda \in Sp(A)$ ,  $\Re(\lambda) \leq 0$  et il existe  $\lambda \in Sp(A)$  avec  $\Re(\lambda) = 0$ , on ne peut pas conclure par l'étude de  $J_f(a)$ .

*Démonstration.* Nous donnons ici seulement la démonstration pour  $n = 1$ . Elle est beaucoup plus simple que celle du cas  $n > 1$  qui sera faite au cours suivant. Supposons donc  $n = 1$ . La matrice jacobienne de  $f$  au point  $a$  est alors le nombre réel  $f'(a)$  (identifié à un élément de  $\mathcal{M}_1(\mathbb{R})$ ).

- 1) *Premier cas* :  $f'(a) < 0$ . Le premier cas du théorème 13 correspond à  $f'(a) < 0$ . Dans ce cas, il existe  $\delta_0 > 0$  tel que  $f'(z) < 0$  pour tout  $z$  tel que  $|a - z| \leq \delta_0$ . Par le théorème des accroissements finis, il existe  $\theta$  entre  $a$  et  $z$  tel que  $f(z) - f(a) = f'(\theta)(z - a)$ . Comme  $f(a) = 0$ , on en déduit

$$f(z) = f(z) - f(a) > 0 \text{ si } a - \delta_0 \leq z < a, \quad (9.15)$$

$$f(z) = f(z) - f(a) < 0 \text{ si } a < z < a + \delta_0. \quad (9.16)$$

Soit  $x$  la solution maximale de (9.13) correspondant à la condition initiale  $x(0)$ .

Si  $a - \delta_0 \leq x(0) < a$ , la fonction  $x$  commencera par être strictement croissante, grâce à (9.15) (et  $x'(t) = f(x(t))$ ), et restera strictement croissante tant que  $x(t) < a$ . Mais comme la fonction constante égale à  $a$  est solution de (9.13), il est impossible qu'il existe  $t$  tel que  $x(t) = a$ . On en déduit que la solution  $x$  est globale et que  $a - \delta_0 \leq x(0) < x(t) < a$  pour tout  $t \in ]0, +\infty[$ .

Si  $a < x(0) \leq a + \delta_0$ , un raisonnement analogue, utilisant (9.16), donne que  $x$  est strictement décroissante et  $a < x(t) < x(0) \leq a + \delta_0$  pour tout  $t \in ]0, +\infty[$ .

Ceci montre la stabilité de  $a$ . En effet pour tout  $\varepsilon > 0$ , il suffit de prendre  $\delta = \min\{\varepsilon, \delta_0\}$  et on obtient

$$|x(0) - a| \leq \delta \Rightarrow |x(t) - a| \leq \varepsilon \text{ pour tout } t \geq 0.$$

Pour montrer la stabilité asymptotique de  $a$ , on remarque que si  $|x(0) - a| \leq \delta_0$ , la fonction  $x(t)$  est monotone bornée. Elle a donc une limite en  $+\infty$  notée  $c$ . Cette limite appartient à l'intervalle  $]a - \delta_0, a + \delta_0[$ . En remarquant que pour tout  $t > 0$  il existe  $\theta_t \in [t, t+1]$  tel que  $x(t+1) - x(t) = x'(\theta_t) = f(x(\theta_t))$ , on en déduit que  $f(c) = 0$ , ce qui prouve que  $c = a$  (en effet, grâce à (9.15)-(9.16)  $a$  est seul point de l'intervalle  $]a - \delta_0, a + \delta_0[$  où  $f$  s'annule). Ceci montre bien la stabilité asymptotique de  $a$ .

- 2) *Second cas* :  $f'(a) > 0$ . Un raisonnement analogue au précédent nous donne l'existence de  $\delta_0 > 0$  tel que

$$f(z) = f(z) - f(a) < 0 \text{ si } a - \delta_0 \leq z < a, \quad (9.17)$$

$$f(z) = f(z) - f(a) > 0 \text{ si } a < z < a + \delta_0. \quad (9.18)$$

Soit  $x$  la solution maximale de (9.13) correspondant à la condition initiale  $x(0)$ . Cette solution est définie sur l'intervalle  $[0, T_m[$ . On va montrer que si  $x(0) \neq a$ ,

$$\sup_{t \in [0, T_m[} |x(t) - a| \geq \delta_0.$$

(Ceci montre bien l'instabilité de  $a$ .)

Il suffit, bien sûr, de considérer le cas où  $|x(0) - a| < \delta_0$ . (Car  $|x(0) - a| \leq \sup_{t \in [0, T_m[} |x(t) - a|$ .)

Si  $a - \delta_0 < x(0) < a$ , la fonction  $x$  commencera par être strictement décroissante grâce à (9.17) (et  $x'(t) = f(x(t))$ ) et restera strictement décroissante tant que  $a - \delta_0 \leq x(t)$ . Il est impossible que  $a - \delta_0 \leq x(t)$  pour tout  $t \in [0, T_m[$  car on aurait alors  $T_m = +\infty$  (car il n'y aurait pas explosion en  $T_m$ ) et (comme dans le premier cas) on aurait  $\lim_{t \rightarrow +\infty} x(t) = c$  avec  $a - \delta_0 \leq c < a$  et  $f(c) = 0$  en contradiction avec (9.17). Ceci montre qu'il existe  $t_1 > 0$  tel que  $x(t_1) < a - \delta_0$  et donc  $\sup_{t \in [0, T_m[} |x(t) - a| \geq |x(t_1) - a| > \delta_0$ .

On peut d'ailleurs aussi remarquer que  $x(t) < a - \delta_0$  pour tout  $t > t_1$  (ceci peut se montrer par l'absurde, si  $t_2$  est le premier point après  $t_1$  pour lequel  $x(t_2) = a - \delta_0$ , on a alors  $x'(t_2) \geq 0$ , mais  $x'(t_2) = f(x(t_2)) = f(a - \delta_0) < 0$  grâce à (9.17). C'est donc impossible).

Un raisonnement analogue peut être fait si  $a < x(0) < a + \delta_0$ .

Il donne aussi  $\sup_{t \in [0, T_m[} |x(t) - a| > \delta_0$ . On a bien montré l'instabilité de  $a$ .

Ceci termine la démonstration du théorème 13 dans le cas  $n = 1$ .

□

## Cours 10. Preuves des théorèmes de stabilité

Rappelons que nous étudions le problème (9.2) et que nous nous intéressons à la stabilité d'un point d'équilibre, c'est-à-dire d'un point  $a$  de  $\mathbb{R}^n$  tel que  $f(a) = 0$ .

La section suivante est dédiée à la preuve du théorème 12, c'est-à-dire au cas où  $f(X) = AX$  avec  $A \in \mathcal{M}_n(\mathbb{R})$  et  $a = 0$ . Le système considéré est alors le système (9.12). On étudie la stabilité de la solution d'équilibre  $X \equiv 0$ . On rappelle que par définition,

1) 0 est stable si :

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tel que } \|X^{(0)}\| \leq \delta \Rightarrow \sup_{t \in [0, +\infty[} \|X(t)\| \leq \varepsilon \quad (10.1)$$

2) 0 est asymptotiquement stable si 0 est stable et  $\lim_{t \rightarrow +\infty} \|X(t)\| = 0$

### 10.1 Preuve du théorème de stabilité linéaire

On étudie ici le problème (9.12). L'objet de cette section est de donner la démonstration du théorème 12 dans le cas  $n = 1$  et  $n = 2$ . Nous utiliserons les notations suivantes :

$$A \in \mathcal{M}_n(\mathbb{R}), \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

$$\|X\| = \max\{|x_i|, i \in \{1, \dots, n\}\}, \quad \|A\| = \max\{|A_{i,j}|, i, j \in \{1, \dots, n\}\}.$$

Nous allons d'abord traiter le cas  $n = 1$  (qui est simple) puis le cas  $n = 2$  (qui contient l'essentiel des difficultés du cas  $n > 1$ ).

#### 10.1.1 Cas linéaire, une seule équation ( $n = 1$ )

Supposons dans un premier temps que  $n = 1$ . Le problème s'écrit (pour  $a$  et  $x^{(0)}$  donnés) :

$$x'(t) = ax, \quad t > 0 \quad (10.2a)$$

$$x(0) = x^{(0)} \quad (10.2b)$$

La solution de (10.2) est  $x(t) = x^{(0)}e^{at}$ .

- 1) Si  $a \leq 0$ , alors  $|x(t)| \leq |x^{(0)}|$  et 0 est donc solution stable (uniformément).
- 2) Si  $a < 0$ ,  $\lim_{t \rightarrow +\infty} x(t) = 0$  et 0 est donc solution asymptotiquement stable.
- 3) Si  $a > 0$ ,  $x^{(0)} \neq 0$  alors  $\sup_{t \in [0, +\infty[} |x(t)| = +\infty$ , l'équilibre 0 est donc instable.

### 10.1.2 Cas linéaire, système, $n = 2$

On considère le problème de Cauchy associé au problème (9.12), qui s'écrit :

$$X'(t) = AX, \quad t > 0 \quad (10.3a)$$

$$X(0) = X^{(0)} \quad (10.3b)$$

À partir de maintenant on suppose  $n = 2$  et on note  $\lambda_1, \lambda_2$  les deux racines du polynôme caractéristique de  $A$  (avec éventuellement  $\lambda_1 = \lambda_2$ ). On étudie successivement les trois cas du théorème 12.

**Cas 1,**  $\Re(\lambda_1) < 0, \Re(\lambda_2) < 0$ .

Nous allons démontrer ici, en distinguant trois sous-cas, que 0 est asymptotiquement stable.

- 1) *Cas où  $A$  diagonalisable dans  $\mathbb{R}$ .* Il existe donc  $\lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 \leq \lambda_2 < 0$  et  $\{\varphi_1, \varphi_2\}$  base de  $\mathbb{R}^n$  tels que

$$A\varphi_1 = \lambda_1\varphi_1, \quad A\varphi_2 = \lambda_2\varphi_2.$$

La solution générale du système (10.3a) est

$$X(t) = c_1 X^{(1)}(t) + c_2 X^{(2)}(t) \text{ avec } X^{(1)}(t) = e^{\lambda_1 t} \varphi_1, \quad X^{(2)}(t) = e^{\lambda_2 t} \varphi_2$$

On a donc

$$X(t) = M(t)C \text{ avec } C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad M(t) = \begin{bmatrix} X^{(1)}(t) & X^{(2)}(t) \end{bmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

La solution de (10.3a)-(10.3b) est  $X(t) = M(t)M(0)^{-1}X(0)$ . (Noter que la matrice  $M(0)$  est bien inversible car  $\varphi_1, \varphi_2$  sont linéairement indépendantes).

Le lemme suivant est laissé à titre d'exercice (utiliser  $\lambda_1 \leq \lambda_2 < 0$ ).

**Lemme 13** (Bornes sur la solution). *Avec les notations précédentes, on a*

(a)  $\|M(t)\| \leq ce^{\lambda_2 t}$ , et  $c > 0$  ne dépend que de  $\varphi_1$  et  $\varphi_2$

(b)  $\|X(t)\| \leq b\|X(0)\|e^{\lambda_2 t}$  et  $b > 0$  ne dépend que de  $\varphi_1$  et  $\varphi_2$

En posant  $\delta = \varepsilon/b$  on obtient la stabilité (définition (10.1)). Puis, comme  $\lim_{t \rightarrow +\infty} \|X(t)\| = 0$ , on obtient même la stabilité asymptotique.

- 2) *Cas où  $A$  est diagonalisable dans  $\mathbb{C}$  et non dans  $\mathbb{R}$ .* Dans ce cas  $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta$  avec  $\beta \neq 0$ . Soit  $\varphi = \psi + i\chi$  tel que  $A\varphi = \lambda_1\varphi$ .

On a déjà vu que  $\{\psi, \chi\}$  est une base de  $\mathbb{R}^2$ .

La solution générale du système (10.3a) est

$$X(t) = c_1 X^{(1)}(t) + c_2 X^{(2)}(t)$$

avec

$$X^{(1)}(t) = e^{\alpha t}(\cos(\beta t)\psi - \sin(\beta t)\chi), \quad X^{(2)}(t) = e^{\alpha t}(\sin(\beta t)\psi + \cos(\beta t)\chi)$$

On a donc  $X(t) = M(t)C$  avec

$$C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad M(t) = \begin{bmatrix} X^{(1)}(t) & X^{(2)}(t) \end{bmatrix} \in \mathcal{M}_n(\mathbb{R})$$

La solution de (10.3a)-(10.3b) est  $X(t) = M(t)M(0)^{-1}X(0)$ . (Ici encore, on peut noter que la matrice  $M(0)$  est bien inversible car  $\psi, \chi$  sont linéairement indépendantes).

Un lemme analogue au lemme 13 est laissé en exercice.

**Lemme 14** (Bornes sur la solution). Avec les notations précédentes, on a

(a)  $\|M(t)\| \leq ce^{\alpha t}$ , et  $c > 0$  ne dépend que de  $\varphi_1$  et  $\varphi_2$

(b)  $\|X(t)\| \leq b\|X(0)\|e^{\alpha t}$  et  $b > 0$  ne dépend que de  $\varphi_1$  et  $\varphi_2$

Comme  $\alpha \leq 0$ , on obtient la stabilité en posant  $\delta = \varepsilon/b$  (définition (10.1)). Puis, Grâce à  $\alpha < 0$ , on obtient même la stabilité asymptotique.

- 3) *Cas où A est non diagonalisable.* Si  $A$  est non diagonalisable, cela signifie que  $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$  et que  $m_a(\lambda) = 2, m_g(\lambda) = 1$ . La solution générale du système (10.3a) est alors

$$X(t) = c_1 X^{(1)}(t) + c_2 X^{(2)}(t) \text{ avec } X^{(1)}(t) = e^{\lambda t} \varphi, \quad X^{(2)}(t) = e^{\lambda t}(\psi + t\varphi),$$

où  $A\varphi$  est un vecteur propre associé à  $\lambda$ , i.e.  $A\varphi = \lambda\varphi$ , et  $\psi$  est un vecteur tel que  $\varphi, \psi$  forment une base de  $\text{Ker}(A - \lambda I)^2$  (et donc ici une base de  $\mathbb{R}^2$ ).

On a ici encore  $X(t) = M(t)C$  avec

$$C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad M(t) = \begin{bmatrix} e^{\lambda t} \varphi & e^{\lambda t}(\psi + t\varphi) \end{bmatrix} = e^{\frac{\lambda t}{2}} \begin{bmatrix} e^{\frac{\lambda t}{2}} \varphi & e^{\frac{\lambda t}{2}}(\psi + t\varphi) \end{bmatrix}$$

La solution de (10.3a)-(10.3b) est toujours  $X(t) = M(t)M(0)^{-1}X(0)$ .

Comme  $\lambda < 0$ , les fonctions  $t \mapsto e^{\frac{\lambda t}{2}}$  et  $t \mapsto te^{\frac{\lambda t}{2}}$  sont bornées sur  $\mathbb{R}_+$ . On en déduit que  $\|M(t)\| \leq ae^{\frac{\lambda t}{2}}$  avec  $a$  ne dépendant que de  $\varphi, \psi$  et  $\lambda$ . On a également  $\|X(t)\| \leq be^{\frac{\lambda t}{2}}\|X(0)\|$  avec  $b$  ne dépendant que de  $\varphi, \psi$  et  $\lambda$ . Ceci donne la stabilité asymptotique de 0.

**Cas 2 : il existe  $\lambda \in Sp(A)$  tel que  $\Re(\lambda) > 0$ .** On distingue les cas  $\lambda \in \mathbb{R}$  et  $\lambda \in \mathbb{C}$ .

- 1) *Cas  $\lambda \in \mathbb{R}$ .* Soit  $\varphi \in \mathbb{R}^2, \varphi \neq 0$  tel que  $A\varphi = \lambda\varphi$ , de sorte que la fonction définie par  $X^{(1)}(t) = e^{\lambda t}\varphi$  est une solution de (10.3a).

Pour tout  $\varepsilon > 0$ , la fonction  $X(t) = \varepsilon e^{\lambda t}\varphi$  est la solution de (10.3a)-(10.3b) avec la donnée initiale  $X(0) = \varepsilon\varphi$ .

Comme  $\sup_{t>0} \|\varepsilon X(t)\| = +\infty$  et que  $\|X(0)\|$  est arbitrairement petit, ceci montre que 0 est un équilibre instable.

- 2) *Cas  $\lambda \in \mathbb{C}$  (et non dans  $\mathbb{R}$ ).* Supposons maintenant que  $\lambda$  soit complexe :  $\lambda = \alpha + i\beta, \alpha > 0, \beta \neq 0$ . Soit  $\varphi$  tel que  $A\varphi = \lambda\varphi, \varphi = \psi + i\chi$ .

La fonction

$$X^{(1)}(t) = e^{\alpha t}(\cos(\beta t)\psi - \sin(\beta t)\chi)$$

est solution de (10.3a). On utilise maintenant le fait que  $\psi \neq 0$ .

Pour tout  $\varepsilon > 0$ , la fonction  $X(t) = \varepsilon e^{\alpha t}(\cos(\beta t)\psi - \sin(\beta t)\chi)$  est la solution de (10.3a)-(10.3b) avec la donnée initiale  $X(0) = \varepsilon\psi$ .

On a alors :

$$\sup_{t>0} \|\varepsilon X(t)\| = +\infty, \quad \|\varepsilon X(0)\| \leq \varepsilon\|\psi\|$$

Ceci prouve que 0 est un équilibre instable.

**Cas 3 :**  $\Re(\lambda_1) \leq 0$ ,  $\Re(\lambda_2) \leq 0$  et  $\Re(\lambda_1)\Re(\lambda_2) = 0$ .

On peut supposer sans perte de généralité que  $\Re(\lambda_1) = 0$ . (Noter qu'il est possible que  $\lambda_1 = \lambda_2$ .)

1) Cas 3(a) :  $\forall \lambda \in Sp(A)$  tel que  $\Re(\lambda) = 0$  on a  $m_a(\lambda) = m_g(\lambda)$ .

Si  $\lambda_1 \neq \lambda_2$  on a donc aussi  $m_a(\lambda_2) = m_g(\lambda_2)$  (plus précisément  $m_a(\lambda_2) = m_g(\lambda_2) = 1$ ).

Si  $\lambda_1 = \lambda_2$  (c'est-à-dire  $m_a(\lambda_1) = 2$ ), on a alors  $m_a(\lambda_1) = m_g(\lambda_1) = 2$ .

La matrice est donc toujours diagonalisable (dans  $\mathbb{R}$  ou  $\mathbb{C}$ ). On a profité ici du fait que  $n = 2$  (pour  $n > 2$ , la situation serait un peu plus complexe).

On distingue maintenant les cas  $A$  diagonalisable dans  $\mathbb{R}$  et  $A$  diagonalisable dans  $\mathbb{C}$ .

(a) Cas  $A$  est diagonalisable dans  $\mathbb{R}$ . Comme  $\lambda_1, \lambda_2 \in \mathbb{R}$ , la solution générale de (10.3a) est donnée par les fonctions  $X^{(1)}$  et  $X^{(2)}$  (avec  $A\varphi_i = \lambda_i\varphi_i$ ,  $\varphi_i \neq 0$ ,  $i = 1, 2$ )

$$X^{(1)}(t) = e^{\lambda_1 t} \varphi_1 \quad X^{(2)}(t) = e^{\lambda_2 t} \varphi_2$$

On a donc  $M(t) = [e^{\lambda_1 t} \varphi_1 \quad e^{\lambda_2 t} \varphi_2]$  et, comme  $\lambda_i \leq 0$  pour  $i = 1, 2$ , il existe  $a$  et  $b$  ne dépendant que de  $\varphi_1$  et  $\varphi_2$  tels que  $\|M(t)\| \leq a$ , et  $\|X(t)\| \leq b\|X(0)\|$  pour tout  $t > 0$ . Le point 0 est donc un équilibre stable.

Mais 0 n'est pas asymptotiquement stable car  $\lambda_1 = 0$  et donc, pour tout  $\varepsilon > 0$ , la fonction constante égale à  $\varepsilon\varphi_1$  est solution de (10.3a).

(b) Cas  $A$  est diagonalisable dans  $\mathbb{C}$  (et non dans  $\mathbb{R}$ )

Supposons maintenant que  $\lambda_1 = \alpha + i\beta = i\beta$  et  $\lambda_2 = \alpha - i\beta = -i\beta$  (avec  $\beta \neq 0$ ). Avec les notations habituelles, la solution générale de (10.3a) est donnée par les fonctions

$$X^{(1)}(t) = \cos(\beta t)\psi - \sin(\beta t)\chi \quad X^{(2)}(t) = \sin(\beta t)\psi + \cos(\beta t)\chi$$

On a donc :  $\|M(t)\| \leq a$ , avec  $a$  ne dépendant que de  $\psi, \chi$  et  $\|X(t)\| \leq b\|X(0)\|$  avec  $b$  ne dépendant que de  $\psi, \chi$ . Il y a équilibre stable en 0. Par contre, comme dans le cas  $A$  diagonalisable dans  $\mathbb{R}$ , 0 n'est pas asymptotiquement stable.

2) Cas 3(b) :  $\exists \lambda \in Sp(A)$  tel que  $\Re(\lambda) = 0$  et  $m_a(\lambda) > m_g(\lambda)$ .

Dans le cas  $n = 2$  considéré ici, ce cas correspond à  $0 \in Sp(A)$  et  $m_a(0) = 2$ ,  $m_g(0) = 1$ . La solution générale de (10.3a) est donnée par les fonctions

$$X^{(1)}(t) = \varphi \quad X^{(2)}(t) = \psi + t\varphi,$$

avec  $\varphi, \psi \in \mathbb{R}^2$ , non nuls. La fonction  $X^{(1)}$  est bornée mais  $\sup_{t \in \mathbb{R}_+} \|X^{(2)}(t)\| = +\infty$ .

Il y a donc équilibre instable en 0, car  $\|\varepsilon X^{(2)}(0)\| = |\varepsilon|\|\psi\|$  est aussi petit que l'on veut (et non nul) alors que  $\sup_{t \in \mathbb{R}_+} \|\varepsilon X^{(2)}(t)\| = +\infty$  pour  $\varepsilon \neq 0$ .

Ceci termine la démonstration du théorème 12.

**Exemple 26** (Pendule pesant linéarisé). *Considérons à nouveau l'exemple du pendule pesant linéarisé. On rappelle que l'équation est*

$$x''(t) + x(t) = 0, \quad t > 0. \quad (10.4)$$

On pose  $x_1(t) = x(t)$ ,  $x_2(t) = x'(t)$ ,  $X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$  et  $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ .

L'équation (du second ordre) (10.4) se ramène donc au système  $X'(t) = AX$ . Les valeurs propres de la matrice  $A$  sont  $\pm i$ , le théorème 12 donne donc la stabilité de 0 (car  $m_a(\lambda) = m_g(\lambda)$  pour les deux valeurs propres).

On peut retrouver ce résultat directement, les solutions du système sont

$$x_1(t) = \alpha \cos t + \beta \sin t \quad x_2(t) = -\alpha \sin t + \beta \cos t, \quad \alpha = x_1(0), \quad \beta = x_2(0)$$

On a alors :  $\|X(t)\| \leq |\alpha| + |\beta| \leq 2\|X(0)\|$  pour tout  $t$ . Ce qui prouve la stabilité de 0.

## 10.2 Preuve partielle du théorème de stabilité, cas non linéaire

On se place maintenant sous les hypothèses du théorème 13 et on considère le système différentiel non linéaire (9.13), et  $a \in \mathbb{R}^n$  tel que  $f(a) = 0$ . Le cas  $n = 1$  a été démontré dans le cours précédent. On va s'intéresser ici au cas  $n = 2$  (qui contient l'essentiel des difficultés du cas  $n > 1$ ) et on ne démontre que l'item 1 du théorème, c.à.d. que si  $\Re(\lambda) < 0$  pour tout  $\lambda \in Sp(A)$ , alors  $a$  est asymptotiquement stable. En posant  $Y = X - a$  l'équation sur  $Y$  est  $Y'(t) = f(Y(t) + a) = h(Y(t))$  avec  $h$  définie par  $h(z) = f(z + a)$ . (Noter que  $h \in C^1(\mathbb{R}^n, \mathbb{R}^n)$  car  $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ .) On rappelle que l'on note  $J_f(a)$  la matrice jacobienne en  $a$  de  $f$  et  $J_h(0)$  la matrice jacobienne en 0 de  $h$ . Comme  $J_h(0) = J_f(a)$ , il suffit de démontrer le théorème 13 pour une fonction  $f$  telle que  $f(0) = 0$ , avec  $a = 0$ . Comme on suppose que  $n = 2$ , on note  $x_1, x_2$  les composantes de  $X$ , et le système (9.13) s'écrit

$$\begin{aligned}x'_1(t) &= f_1(x_1(t), x_2(t)) \quad t > 0, \\x'_2(t) &= f_2(x_1(t), x_2(t)), \quad t > 0.\end{aligned}$$

La dérivabilité de  $f$  au point 0 donne, pour tout  $Z \in \mathbb{R}^2$ ,

$$f(Z) = \underbrace{f(0)}_{=0} + \underbrace{J_f(0)Z}_A + \underbrace{\|Z\|\varepsilon(Z)}_{g(Z)} = AZ + g(Z),$$

avec  $A = J_f(0) = \begin{bmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{bmatrix}$ ,  $g(Z) = \|Z\|\varepsilon(Z)$  et  $\lim_{Z \rightarrow 0} \varepsilon(Z) = 0$ .

Le système (9.13) s'écrit donc  $X'(t) = AX(t) + g(X(t))$ . Noter aussi que  $g \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ . Par hypothèse, on a  $\Re(\lambda_1) < 0$  et  $\Re(\lambda_2) < 0$  pour  $\lambda_1, \lambda_2 \in Sp(A)$ , et on veut montrer que 0 est asymptotiquement stable. On commence par un calcul semblable à celui que l'on fait avec la technique "variation de la constante". Soit  $X$  la solution maximale de 9.13 avec donnée initiale  $X(0)$  (cette solution est donc définie sur  $[0, T_m[$ ). On pose  $C(t) = e^{-At}X(t)$  et on a donc  $X(t) = e^{At}C(t)$ . On en déduit que

$$\begin{aligned}X'(t) &= Ae^{At}C(t) + e^{At}C'(t) \\ &= AX(t) + g(X(t)) = Ae^{At}C(t) + g(X(t)).\end{aligned}$$

Ceci donne que  $C'(t) = e^{-At}g(X(t))$ . Comme

$$C(t) = \int_0^t C'(s)ds + C(0) \text{ et } C(0) = X(0),$$

on obtient

$$C(t) = X(0) + \int_0^t e^{-As}g(X(s))ds.$$

On en déduit la formule suivante, dite formule de Duhamel :

$$X(t) = e^{At}X(0) + \int_0^t e^{A(t-s)}g(X(s))ds, \text{ pour tout } t \in [0, T_m[.$$

Pour la suite, on suppose sans perte de généralité, que  $\Re(\lambda_1) \leq \Re(\lambda_2) = 2\alpha < 0$ . En reprenant la démonstration du théorème 12, cas 1 (section 10.1.2) et en utilisant le fait que  $e^{At} = M(t)M(0)^{-1}$ , il existe  $\tilde{b}$ , ne dépendant que de  $\varphi_1, \varphi_2$  (vecteurs propres associés à  $\lambda_1$  et  $\lambda_2$ ) tel que  $\|e^{At}\| \leq \tilde{b}e^{\alpha t}$  pour

tout  $t \geq 0$  (on peut par ailleurs remarquer que  $\tilde{b} \geq 1$ ). Compte tenu de la définition de la norme (dans  $\mathbb{R}^2$  et dans  $\mathcal{M}_2(\mathbb{R})$ ), ceci donne, avec  $b = 2\tilde{b}$ ,  $\|e^{At}Z\| \leq be^{\alpha t}\|Z\|$  pour tout  $Z \in \mathbb{R}^2$ . On a donc

$$\|X(t)\| \leq be^{\alpha t}\|X(0)\| + \int_0^t be^{\alpha(t-s)}\|g(X(s))\|ds \text{ pour tout } t \in [0, T_m].$$

On utilise maintenant le fait que  $g(Z) = \|Z\|\varepsilon(Z)$ . Comme  $\varepsilon(Z) \rightarrow 0$  lorsque  $Z \rightarrow 0$ , on a

$$\forall \tilde{\varepsilon} > 0, \exists \delta > 0 : \|X\| \leq \delta \Rightarrow \|\varepsilon(X)\| \leq \tilde{\varepsilon}.$$

Soit donc  $\tilde{\varepsilon} > 0$  et  $0 < T < T_m$ . Si

$$\forall s \in [0, T], \|X(s)\| \leq \delta, \quad (10.5)$$

on a alors, pour tout  $s \in [0, T]$ ,  $\|g(X(s))\| \leq \|X(s)\|\tilde{\varepsilon}$  et donc, pour tout  $0 \leq t \leq T$ ,

$$e^{-\alpha t}\|X(t)\| \leq b\|X(0)\| + \int_0^t be^{-\alpha s}\|X(s)\|\tilde{\varepsilon}ds$$

On applique une technique de Gronwall (voir lemme 4); en posant

$$\varphi(t) = b\|X(0)\| + \int_0^t be^{-\alpha s}\|X(s)\|\tilde{\varepsilon}ds,$$

l'inégalité précédente donne

$$\varphi'(t) = b\tilde{\varepsilon}\|X(t)\|e^{-\alpha t} \leq b\tilde{\varepsilon}\varphi(t).$$

On a donc  $\varphi'(t) - b\tilde{\varepsilon}\varphi(t) \leq 0$  pour tout  $0 \leq t \leq T$ . En posant  $\psi(t) = \varphi(t)e^{-b\tilde{\varepsilon}t}$ , on obtient

$$\psi'(t) = (\varphi'(t) - \tilde{\varepsilon}b\varphi(t))e^{-b\tilde{\varepsilon}t} \leq 0.$$

La fonction  $\psi$  est décroissante et donc  $\psi(t) \leq \psi(0)$  pour tout  $0 \leq t \leq T$ . On en déduit

$$\varphi(t)e^{-b\tilde{\varepsilon}t} \leq \varphi(0) \text{ et donc } \varphi(t) \leq b\|X(0)\|e^{b\tilde{\varepsilon}t}.$$

Comme  $\|X(t)\| \leq e^{\alpha t}\varphi(t)$ , on a  $\|X(t)\| \leq b\|X(0)\|e^{b\tilde{\varepsilon}t}e^{\alpha t}$  pour tout  $0 \leq t \leq T$ . En choisissant  $\tilde{\varepsilon} = -\frac{\alpha}{2b}$  (ce qui est possible car  $\alpha < 0$ ), on obtient, à condition que (10.5) soit vérifiée,

$$\|X(t)\| \leq b\|X(0)\|e^{\frac{\alpha}{2}t} \text{ pour tout } 0 \leq t \leq T. \quad (10.6)$$

On va montrer maintenant que si  $\|X(0)\| \leq \frac{\delta}{b}$ , alors la condition (10.5) est vérifiée pour tout  $T < T_m$ , que  $T_m = +\infty$  et (10.6) est vraie pour tout  $t > 0$  (ce qui donne la stabilité asymptotique de 0).

On suppose donc que  $\|X(0)\| \leq \frac{\delta}{b}$ . Ceci donne, en particulier,  $\|X(0)\| < \delta$  (car on a déjà vu que  $b > 1$ ).

On montre par l'absurde que  $\|X(t)\| < \delta$  pour tout  $t < T_m$ . En effet, sinon on pose

$$T = \inf\{t \in [0, T_m], \|X(t)\| = \delta\}.$$

On a  $T > 0$  car  $\|X(0)\| < \delta$  et (10.5) est vérifié. On a donc (10.6), ce qui donne pour  $t = T$ ,  $\|X(T)\| \leq b\|X(0)\|e^{\frac{\alpha}{2}T} < \delta$ , en contradiction avec la définition de  $T$ . On a donc montré que  $\|X(t)\| < \delta$  pour tout  $t < T_m$ . Ceci donne d'une part que  $T_m = +\infty$  (il n'y a pas "explosion" en  $T_m$ ) et d'autre part que

$$\|X(t)\| \leq b\|X(0)\|e^{\frac{\alpha}{2}t} \text{ pour tout } 0 \leq t < +\infty.$$

On en déduit que 0 est asymptotiquement stable.

## 10.3 Exemples

### 10.3.1 Pendule non linéarisé

Revenons maintenant à l'exemple du pendule et considérons le modèle du pendule non linéarisé.

$$x''(t) + \frac{g}{\ell} \sin(x(t)) = 0 \quad (10.7)$$

Dans l'équation (10.7)  $g$  correspond à l'intensité de la pesanteur et  $\ell$  la longueur du fil, on considère la Figure 5.1 page 41. On suppose que  $\frac{g}{\ell} = 1$ . On pose  $x_1 = x$  et  $x_2 = x'$ . Alors :

$$X'(t) = \begin{bmatrix} x_2 \\ -\sin(x_1) \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

On sait que  $(0, 0)$  ainsi que  $(\pi, 0)$  sont deux points d'équilibre. La matrice jacobienne s'écrit :

$$J_f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ -\cos(x_1) & 0 \end{bmatrix}.$$

Si  $x_1 = x_2 = 0$  alors  $J_f(0, 0) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$  dont les valeurs propres sont  $\pm i$ . On ne peut donc pas conclure avec le théorème 13.

Si  $x_1 = \pi$  et  $x_2 = 0$ ,  $J_f(\pi, 0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ , dont les valeurs propres sont  $\pm 1$ . L'équilibre est donc instable.

### 10.3.2 Modèle de Lotka-Volterra

Le système de Lotka-Volterra (9.3) peut s'écrire :

$$\begin{aligned} x_1'(t) &= x_1(t)(a - bx_2(t)), \\ x_2'(t) &= x_2(t)(-c + dx_1(t)). \end{aligned}$$

On rappelle que  $a > 0$ ,  $b > 0$ ,  $c > 0$ ,  $d > 0$ . Les points d'équilibre sont  $(0, 0)$  et  $(\frac{c}{d}, \frac{a}{b})$ . La matrice jacobienne est :

$$J_f(x_1, x_2) = \begin{bmatrix} a - bx_2 & -bx_1 \\ dx_2 & -c + dx_1 \end{bmatrix},$$

Si  $x_1 = x_2 = 0$  alors  $J_f(0, 0) = \begin{bmatrix} a & 0 \\ 0 & -c \end{bmatrix}$ , et le point  $(0, 0)$  est instable.

Si  $x_1 = \frac{c}{d}$  et  $x_2 = \frac{a}{b}$  alors  $J_f\left(\frac{c}{d}, \frac{a}{b}\right) = \begin{bmatrix} 0 & -\frac{bc}{d} \\ \frac{ad}{b} & 0 \end{bmatrix}$ . Le polynôme caractéristique de cette matrice est

$\lambda^2 + \frac{adbc}{bc} = \lambda^2 + ac$ . Les valeurs propres sont  $\pm i\sqrt{ac}$ . On est dans le cas 3) du théorème 13 et on ne peut pas conclure.

Il est en fait possible de démontrer que les solutions sont périodiques. Ceci est démontré dans l'examen de janvier 2020 et est illustré à l'aide d'un schéma numérique dans l'exercice 2 du TP4. On peut estimer cette période assez précisément à l'aide du schéma RK4.

# Cours 11. Schémas numériques pour les systèmes

On présente ici les schémas numériques pour l'approximation des systèmes différentiels abordés dans les chapitres précédents. Soit  $f \in C([0, T[ \times \mathbb{R}^n, \mathbb{R}^n)$ , localement lipschitzienne par rapport à son deuxième argument. On cherche une approximation de la fonction  $X \in \mathcal{C}^1([0, T_m[, \mathbb{R}^n) \cap \mathcal{C}([0, T_m[, \mathbb{R}^n)$  solution de

$$X'(t) = f(t, X(t)), \quad t > 0, \quad (11.1a)$$

$$X(0) = X^{(0)}, \quad (11.1b)$$

où  $T_m$  est le temps maximal d'existence de la solution. En TP4 on calculera ainsi, par exemple, des solutions approchées du système de Lotka-Volterra (9.3).

## 11.1 Exemples de schémas numériques

On note  $X_k$  l'approximation recherchée de  $X(t_k)$  où  $X$  est solution de (11.1),  $t_k = kh$ , où  $h$  est le pas du temps et  $k \in \mathbb{N}$ . Comme il s'agit d'un système, l'inconnue  $X$  à l'instant  $t$  est un vecteur (colonne) ayant  $n$  composantes, et donc  $X_k$  est un vecteur de  $\mathbb{R}^n$ . Voici quelques schémas déjà vus dans le cadre des équations ( $n = 1$ ) :

1) Schéma d'Euler explicite :

$$\frac{X_{k+1} - X_k}{h} = f(t_k, X_k).$$

2) Schéma d'Euler implicite :

$$\frac{X_{k+1} - X_k}{h} = f(t_{k+1}, X_{k+1}).$$

3) Schéma de Crank Nicolson :

$$\frac{X_{k+1} - X_k}{h} = \frac{1}{2}f(t_k, X_k) + \frac{1}{2}f(t_{k+1}, X_{k+1}).$$

Ce schéma fait en quelque sorte une moyenne entre le schéma d'Euler explicite et le schéma d'Euler implicite. Nous verrons que ce schéma donne une excellente approximation pour certains problèmes<sup>1</sup>, mais il est implicite, comme le schéma d'Euler implicite.

4) Schéma de Heun :

$$\frac{X_{k+1} - X_k}{h} = \frac{1}{2}f(t_k, X_k) + \frac{1}{2}f(t_{k+1}, X_k + hf(t_k, X_k))$$

---

1. Voir TP4

5) Schéma Runge Kutta d'ordre 4 (RK4) :

$$\begin{aligned} X_{k,1} &= X_k + \frac{h}{2} f(t_k, X_k), \\ X_{k,2} &= X_k + \frac{h}{2} f\left(t_k + \frac{h}{2}, X_{k,1}\right), \\ X_{k,3} &= X_k + hf\left(t_k + \frac{h}{2}, X_{k,2}\right), \\ \frac{X_{k+1} - X_k}{h} &= \frac{1}{6} [f(t_k, X_k) + 2f(t_k, X_{k,1}) + 2f(t_k, X_{k,2}) + f(t_{k+1}, X_{k,3})] \end{aligned}$$

On doit se poser quelques questions lors du choix du schéma numérique :

- 1) Quelles sont les propriétés souhaitées ?
- 2) Le schéma est-il convergent ? Peut on démontrer une estimation d'erreur ?
- 3) Comment calculer  $X_{k+1}$  quand  $X_{k+1}$  dépend non linéairement de  $X_k$  ?

Les réponses à ces questions sont abordées dans les paragraphes suivants.

## 11.2 Propriétés de conservation d'un schéma

Il peut être important de conserver la positivité de la solution ou de conserver une grandeur physique (l'énergie totale par exemple pour un problème de mécanique).

### 11.2.1 Conservation de la positivité

Considérons le problème de Cauchy suivant :

$$\begin{cases} y'(t) = -f(y), \\ y(0) = y_0, \quad y_0 > 0, \end{cases} \quad (11.2)$$

avec  $f(y) = y$  ou  $f(y) = y^2$  ou  $f(y) = \sqrt{y}$ . La quantité  $y$  peut représenter une concentration ou une masse volumique, grandeurs qui doivent rester positives. Dans le cadre de modèles issus de problèmes liés à la sûreté nucléaire, par exemple, Il peut être important de conserver certaines grandeurs positives tels que l'énergie, la masse volumique d'eau, la concentration en éléments radioactif... Si la solution approchée (obtenue avec l'aide d'un schéma numérique) venait à ne pas respecter cette positivité, cela pourrait amener des erreurs de calcul importantes pouvant entraîner une solution approchée très éloignée de la réalité (l'idée naïve consistant à mettre à 0 l'approximation numérique lorsque le schéma fournit une valeur négative peut avoir des conséquences catastrophiques).

Commençons par un exemple avec  $n = 1$ , et prenons  $f(y) = \sqrt{y}$ . Le schéma d'Euler explicite s'écrit alors

$$y_{k+1} = y_k - hf(y_k).$$

On a  $y_{k+1} < 0$  si  $y_k - hf(y_k) < 0$ . Donc pour  $f(y) = \sqrt{y}$  on aura  $y_k - hf(y_k) < 0$  dès que  $\sqrt{y_k} < h$ , ce qui risque d'arriver rapidement car on sait que la solution exacte tend vers 0 lorsque  $t \rightarrow +\infty$  (voir exemple 4.3.2).

Examinons maintenant le schéma d'Euler implicite, qui s'écrit, pour  $f(y) = \sqrt{y}$

$$y_{k+1} = y_k - h\sqrt{y_{k+1}}.$$

On a donc

$$y_k = h\sqrt{y_{k+1}} + y_{k+1}.$$

On part de  $y_k > 0$  connu. En posant  $z_{k+1} = \sqrt{y_{k+1}}$ ,  $z_{k+1}$  est donc solution de l'équation du second degré  $z^2 + hz = y_k$ . L'étude de la fonction  $g : z \mapsto g(z) = z^2 + hz - y_k$  est facile et permet de s'assurer qu'il existe une unique solution strictement positive à l'équation  $g(z) = 0$ . On note  $z_{k+1}$  cette solution et on déduit  $y_{k+1} = \sqrt{z_{k+1}}$ .

Pour  $f(y) = \sqrt{y}$  et avec le schéma d'Euler implicite, on aura donc toujours (pour tout  $h > 0$  et  $y_k > 0$ )  $y_{k+1} > 0$ .

Pour cet exemple précis, la conservation de la positivité sera donc assurée par le choix du schéma d'Euler implicite, et pas par celui du schéma d'Euler explicite.

### 11.2.2 Conservation de l'énergie

Prenons l'exemple du système de Coriolis (voir le TP4). Il s'écrit :

$$\left. \begin{aligned} u'(t) &= -v(t), \\ v'(t) &= u(t), \end{aligned} \right\} t > 0,$$

$$u'(0) = u_0, \quad v'(0) = v_0,$$

avec  $u_0, v_0 \in \mathbb{R}$  donnés. On doit avoir conservation de l'énergie cinétique c'est-à-dire

$$E_c(t) = \frac{1}{2}(u^2(t) + v^2(t)) = \frac{1}{2}(u_0^2 + v_0^2) \text{ et donc } E'_c(t) = 0 \text{ pour tout } t > 0$$

Le TP4 montre qu'avec Euler explicite l'énergie associée à la solution approchée croît très vite, alors qu'avec Euler implicite cette même énergie est décroissante. Avec le schéma RK4 elle décroît très légèrement. Finalement avec l'approximation de Crank-Nicolson nous constatons que l'énergie est conservée<sup>2</sup>.

### 11.2.3 Fonction de Lyapunov

**Définition 14** (Fonction de Lyapunov). *On appelle fonction de Lyapunov associée au système différentiel (11.1a) une fonction  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  telle que  $\frac{d}{dt}(g(X(t))) \leq 0$  pour toute fonction  $X$  solution de (11.1a).*

Parmi les fonctions de Lyapunov possibles, on peut avoir l'énergie ou l'entropie. On va alors chercher un schéma numérique qui respecte cette fonction de Lyapunov et qui soit donc tel que  $g(X_{k+1}) \leq g(X_k)$  pour tout  $k$ .

L'exercice 3 du td 8 s'intéresse au système suivant :

$$\begin{cases} x'(t) = y(t) - f(x(t)), & t > 0 \\ y'(t) = -x(t), & t > 0 \end{cases} \quad (11.3)$$

La fonction  $f$  vérifie  $f(z)z > 0$  pour tout  $z \neq 0$ . Une fonction de Lyapunov est ici  $g(x, y) = x^2 + y^2$ .

On a bien  $\frac{d}{dt}g(x(t), y(t)) \leq 0$ .

Il y a deux intérêts majeurs à l'existence d'une fonction de Lyapunov :

- 1) conserver une propriété importante du modèle,
- 2) obtenir un résultat de stabilité qui peut permettre de montrer la convergence du schéma numérique.

2. Pour des problèmes aux dérivées partielles on fait appel soit à Euler ou à Heun mais très rarement à RK4.

### 11.3 Convergence et estimation d'erreur.

Nous allons nous intéresser ici à l'erreur de discrétisation ainsi qu'à l'erreur de consistance d'un schéma donné. On s'intéresse à la solution du problème de Cauchy (11.1) sur l'intervalle  $[0, T]$ ,  $T < T_m$  (où  $T_m$  est le temps d'existence de la solution maximale). On note  $X$  cette solution. Pour un schéma numérique donné et un pas de temps  $h$  tel que  $Nh = T$ ,  $N \in \mathbb{N}^*$ , on définit l'erreur de discrétisation par

$$e_h = \max\{\|X_k - X(t_k)\|, 0 \leq k \leq N\},$$

où  $t_k = kh$  et  $X_k$  est l'approximation de  $X(t_k)$  donnée par le schéma numérique.

On se demande alors si  $e_h \xrightarrow{h \rightarrow 0} 0$  (convergence du schéma).

Puis, on se demande s'il existe  $c > 0$  et  $k \in \mathbb{N}^*$  tels que  $e_h \leq ch^k$  (estimation d'erreur). On dit alors que le schéma est au moins d'ordre  $k$ .

Cette étude de convergence et d'estimation d'erreur s'effectue en règle générale en deux parties :

- 1) Consistance du schéma.
- 2) Stabilité du schéma (au sens "estimation sur la solution approchée indépendante de  $h$ ").

#### 11.3.1 Estimation d'erreur

Soit  $\bar{X}_k = X(t_k)$  la solution exacte<sup>3</sup> au point  $t_k$ . Comme nous l'avons vu dans le cas  $n = 1$  (section 4.2.1), le schéma d'Euler explicite, par exemple, nous donne :

$$\frac{\bar{X}_{k+1} - \bar{X}_k}{h} = f(t_k, \bar{X}_k) + \varepsilon_k.$$

On rappelle que  $\varepsilon_k$  est l'erreur de consistance et si  $f$  est de classe  $C^1$  (et donc  $X$  est de classe  $C^2$ ) alors  $\|\varepsilon_k\| < Ch$ , où  $C$  dépend des dérivées secondes de  $X$ , mais pas de  $h$ . (On a choisi ici une norme sur  $\mathbb{R}^n$ , notée  $\|\cdot\|$ .)

Ici encore, comme nous l'avons vu dans la section 4.2.1, on peut obtenir, en utilisant le schéma numérique, des estimations sur  $X_k$  pour tout  $k \leq N$  (avec  $Nh = \alpha$  en reprenant les notations de la section 4.2.1,  $\alpha$  peut être plus petit que  $T_m$ ).

En utilisant la majoration de l'erreur de consistance et l'estimation sur la solution approchée, on en déduit (toujours comme dans la section 4.2.1) une majoration de l'erreur de discrétisation,

$$e_h = \max\{\|X_k - X(t_k)\|, 0 \leq k \leq N\} \leq \tilde{C}h.$$

Pour les schémas de Heun et de Crank-Nicolson, si  $f$  est de classe  $C^2$ , on aura  $\|\varepsilon_k\| \leq Ch^2$  et

$$e_h = \max\{\|X_k - X(t_k)\|, 0 \leq k \leq N\} \leq \tilde{C}h^2.$$

Pour RK4, si  $f$  de classe  $C^4$  alors  $\|\varepsilon_k\| \leq Ch^4$  et

$$e_h = \max\{\|X_k - X(t_k)\|, 0 \leq k \leq N\} \leq \tilde{C}h^4.$$

---

3. Attention de ne pas confondre  $\bar{X}_k = X(t_k)$  avec  $X_k$  ; en général la suite  $(\bar{X}_k)_k$  n'est pas solution du schéma numérique

### 11.3.2 Convergence sans régularité sur la solution exacte

Au paragraphe précédent, nous avons montré la convergence du schéma grâce à une estimation d'erreur, qui nécessite l'existence de la solution exacte et qui nécessite aussi une régularité plus que  $C^1$  de cette solution exacte. Nous allons donner dans cette section un autre moyen de prouver la convergence de la solution approchée vers la solution exacte quand  $h \rightarrow 0$  qui n'utilise pas d'hypothèse de régularité sur la solution exacte : celle-ci est seulement de classe  $C^1$  (en fait, en utilisant le caractère  $C^1$  de  $X$ , la méthode de la section 11.3.1 donne quand même un résultat d'estimation d'erreur, voir la remarque 14). Le plus grand intérêt de la méthode présentée maintenant est qu'elle n'utilise pas l'existence de la solution exacte, et elle permet même de prouver l'existence d'une solution exacte au problème (11.1) (elle peut donc remplacer la preuve du théorème 8).

Cette méthode est plus difficile que celle de section 11.3.1. Nous donnons ici seulement les idées principales.

Dans la section 11.3.1, on a essentiellement utilisé deux ingrédients :

- 1) La solution exacte est "presque" solution du schéma numérique (c'est l'erreur dite de consistance).
- 2) On a une estimation sur la solution approchée.

Pour cette nouvelle méthode, les ingrédients sont :

- 1) La solution approchée est "presque" solution du problème (11.1) (c'est bien une sorte d'erreur de consistance).
- 2) On a des estimations sur la solution approchée<sup>4</sup>.

Pour cette méthode, nous allons utiliser la forme intégrale du problème (11.1) exposée dans la proposition 2. On rappelle que la fonction  $f$  est continue de  $[0, T[ \times \mathbb{R}^n$  dans  $\mathbb{R}^n$ , localement lipschitzienne par rapport à son deuxième argument. Pour tout  $0 < \alpha < T$ , une fonction  $X$  de  $[0, \alpha]$  dans  $\mathbb{R}^n$  est solution de (11.1) si et seulement si

$$X \in C([0, \alpha], \mathbb{R}^n), \quad (11.4)$$

$$X(t) = X^{(0)} + \int_0^t f(s, X(s)) ds \text{ pour tout } 0 \leq t \leq \alpha. \quad (11.5)$$

Nous allons exposer brièvement la méthode dans le cas du schéma d'Euler explicite.

Pour  $\alpha$  fixé ( $0 < \alpha < T$ ) et  $h = \alpha/N$ ,  $N \in \mathbb{N}^*$ , le schéma s'écrit

$$\begin{aligned} X_0 &= X^{(0)}, \\ \frac{X_{k+1} - X_k}{h} &= f(t_k, X_k) \text{ pour } 0 \leq k \leq N-1. \end{aligned}$$

(On rappelle que  $t_k = kh$ .) On peut alors définir une fonction  $\chi_h$  (continue, affine par morceaux) sur tout l'intervalle  $[0, \alpha]$  en posant

$$\chi_h(t) = X_k + hf(t_k, X_k)(t - t_k), \quad t_k < t \leq t_{k+1}, \quad k \in 0, \dots, N-1.$$

Noter que la fonction  $\chi_h$  dépend de  $N$  (car  $h = \alpha/N$ ). Avec cette définition de  $\chi_h$ , la solution approchée vérifie "presque" (11.4)-(11.5); en effet, en posant  $g_h(s) = f(t_k, X_k)$  pour  $t_k < s \leq t_{k+1}$ , on a pour  $0 \leq t \leq \alpha$  :

$$\begin{aligned} \chi_h(t) &= X^{(0)} + \int_0^t g_h(s) ds \\ &= X^{(0)} + \int_0^t f(s, \chi_h(s)) ds + \int_0^t (g_h(s) - f(s, \chi_h(s))) ds \end{aligned} \quad (11.6)$$

4. Ces estimations peuvent être obtenues à partir du schéma numérique dont la solution approchée est solution; c'est ce qui est fait dans la section 11.3.1, mais aussi du fait que la solution approchée est "presque" solution du problème continu (11.1).

Or si on avait  $\chi_h(t) = X^{(0)} + \int_0^t f(s, \chi_h(s)) ds$ ,  $\chi_h$  serait solution du problème continu (11.1). L'erreur de consistance est donc due ici à la différence entre  $f(t_k, X_k)$  et  $f(s, \chi_h(s))$  lorsque  $t_k < s \leq t_{k+1}$ .

Il s'agit maintenant d'avoir des estimations sur la fonction  $\chi_h$  (c'est-à-dire sur les valeurs de  $X_k$ ). Ceci peut se faire comme dans la section 11.3.1 (et dans la section 4.2.1). On obtient l'existence d'un certain  $\alpha > 0$ , pour lequel, quelque soit  $N$  (on rappelle que  $h = \alpha/N$ ),

$$\|\chi_h(t)\| \leq C_1 \text{ pour tout } 0 \leq t \leq \alpha,$$

où  $C_1$  dépend des données du problème ( $f$  et  $X^{(0)}$ ) mais pas de  $h$ .

En utilisant la définition de  $g$ , on a donc aussi

$$\|g_h(t)\| \leq C_2 \text{ pour tout } 0 \leq t \leq \alpha,$$

où  $C_2$  dépend des données du problème ( $f$  et  $X^{(0)}$ ) mais pas de  $h$ .

Pour continuer, nous avons besoin d'un théorème d'analyse fonctionnelle, le théorème d'Ascoli, que l'on rappelle maintenant.

**Théorème 14** (Théorème d'Ascoli). *Soit  $\alpha > 0$ . Soit  $(u_N)_{N \in \mathbb{N}^*}$  une suite de fonctions définies sur  $[0, \alpha]$  et à valeurs dans  $\mathbb{R}^n$  telle que :*

1) *La suite  $(u_N)_{N \in \mathbb{N}^*}$  est uniformément bornée, c'est-à-dire qu'il existe  $M > 0$  tel que*

$$\forall N \in \mathbb{N}, \forall t \in [0, \alpha], \|u_N(t)\| \leq M,$$

2) *La suite  $(u_N)_{N \in \mathbb{N}^*}$  est uniformément équicontinue, c'est-à-dire*

$$\forall \varepsilon > 0, \exists \eta > 0 \text{ tel que } \forall N \in \mathbb{N}, \forall (t, s) \in [0, \alpha]^2, |t - s| < \eta \Rightarrow \|u_N(t) - u_N(s)\| < \varepsilon.$$

*Alors, il existe une sous-suite de la suite  $(u_N)_{N \in \mathbb{N}^*}$  qui converge uniformément sur  $[0, \alpha]$ .*

Les estimations sur  $\chi_h$  et  $g_h$  permettent de montrer que les hypothèses du théorème 14 sont vérifiées pour la suite  $(u_N)_{N \in \mathbb{N}^*}$ , avec  $u_N = \chi_h$ ,  $h = \alpha/N$ . Il existe donc une suite de pas de temps tendant vers 0 pour laquelle  $\chi_h$  converge uniformément vers une fonction  $X$ .

Il est alors possible de passer à la limite dans l'équation (11.6) pour montrer que  $X$  vérifie (11.5). On a ainsi prouvé l'existence de  $X$  solution de (11.1).

Notons que le théorème d'Ascoli donne une convergence "à une sous-suite près", c'est-à-dire qu'il existe une sous suite de la suite des solutions approchées qui converge vers  $X$  solution de (11.1). On obtient la convergence de toute la suite en utilisant le théorème d'unicité (théorème 9) et en raisonnant, par exemple, par l'absurde : on suppose que la suite ne converge pas vers  $X$ , Il existe alors  $\varepsilon > 0$  et une sous suite qui reste à une distance supérieure à  $\varepsilon$  de  $X$ . Par le théorème d'Ascoli, on extrait de cette sous suite une sous-suite qui converge. La limite de cette nouvelle sous suite est encore solution de (11.1), c'est donc (par le théorème d'unicité) la fonction  $X$ , ce qui amène à une contradiction.

**Remarque 14** (Estimation d'erreur). *On a ainsi démontré l'existence de la solution exacte et la convergence de la solution approchée vers la solution exacte (en utilisant le théorème d'unicité, théorème 9). La convergence ayant lieu pour la norme de la convergence uniforme, on en déduit*

$$e_h = \max\{\|X_k - X(t_k)\|, 0 \leq k \leq N\} \leq \eta(h), \quad (11.7)$$

avec  $\lim_{h \rightarrow 0} \eta(h) = 0$ .

En fait, en reprenant la méthode de la section 4.2.1 (qui suppose l'existence de la solution exacte) il est aussi possible de prouver une estimation d'erreur du type (11.7) avec une fonction  $h \mapsto \eta(h)$ , déterminée par la solution exacte et les estimations sur la solution approchée.

## 11.4 Résolution des systèmes non linéaires

Dans le cas des schémas implicites (Euler implicite, Crank Nicolson...), la valeur de  $X_{k+1}$  nécessite la résolution d'un système linéaire ou non linéaire, selon la nature du système différentiel. Prenons comme exemple le schéma d'Euler implicite. On rappelle que :

$$X_{k+1} = X_k + hf(t_{k+1}, X_{k+1}),$$

où  $X_k, t_k, t_{k+1}$  sont connus. On pose  $y = X_k$  et on définit la fonction  $g$  (de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$ ) par  $g(z) = f(t_{k+1}, z)$ . On cherche donc  $X_{k+1}$  tel que

$$X_{k+1} = X_k + hg(X_{k+1}),$$

ou encore, en posant  $F(z, h) = z - hg(z) - y$ ,

$$F(X_{k+1}, h) = 0.$$

Noter que  $F(y, 0) = 0$ .

Si  $f$  est de classe  $C^1$ , le théorème des fonctions implicites (qui sera vu en cours de calcul différentiel) énoncé ci-après permet de montrer que pour  $h > 0$  assez petit, il existe un et un seul  $X_{k+1}$  proche de  $y$  (c'est-à-dire proche de  $X_k$ ) tel que  $F(X_{k+1}, h) = 0$ .

**Théorème 15** (Théorème des fonctions implicites). Soit  $F \in C^1(\mathbb{R}^n \times \mathbb{R})$ . Soit  $y \in \mathbb{R}^n$  tel que  $F(y, 0) = 0$ . On suppose que la dérivée de  $F$  par rapport à son premier argument au point  $(y, 0)$  est inversible, alors :

$$\exists \varepsilon > 0 \exists \delta > 0, \forall h \in ]-\varepsilon, \varepsilon[ \exists ! z \in B(y, \delta) \text{ tel que } F(z, h) = 0.$$

(On rappelle que  $B(y, \delta) = \{z \in \mathbb{R}^n, \|y - z\| < \delta\}$ .)

La fonction  $F$  est bien ici de classe  $C^1$ . En notant  $D_1F$  la dérivée de  $F$  par rapport à son premier argument, on a  $D_1F(z, y) = I - hDg(z)$  de sorte que  $D_1F(y, 0) = I$ . On peut donc appliquer le théorème des fonctions implicites. Il donne l'existence de  $X_{k+1}$  au moins si  $h$  assez petit (il donne aussi l'unicité si on se limite à chercher  $X_{k+1}$  proche de  $X_k$ ).

Il reste maintenant à voir comment on peut effectivement calculer  $X_{k+1}$ . Une méthode possible est la méthode de Newton. On suppose maintenant  $h$  fixé. En posant  $G(z) = F(z, h)$ , on cherche donc  $X_{k+1}$  tel que  $G(X_{k+1}) = 0$ . La méthode de Newton pour la résolution de l'équation  $G(z) = 0$  s'écrit :

initialisation :  $z_0 = y$ ,

itérations :  $DG(z_\ell)(z_{\ell+1} - z_\ell) = -G(z_\ell)$  pour  $\ell \geq 0$ .

Un avantage considérable de la méthode de Newton est qu'elle converge très rapidement. Un inconvénient est que la convergence n'est assurée que si l'initialisation est faite en un point assez près de la solution. Dans le cas du schéma d'Euler implicite, on initialise le calcul de  $X_{k+1}$  par  $X_k$ , qui *a priori* devrait être proche de  $X_k$ .

**Exemple 27** (Exemple du système du projet de fin de semestre). On considère le système différentiel suivant (problème du pendule pesant avec un “forcing” périodique) :

$$\begin{aligned}x_1'(t) &= x_2(t) \\x_2'(t) &= -\sin(x_1(t)) + \sin t\end{aligned}$$

Le schéma d’Euler implicite s’écrit pour  $k \geq 0$ ,

$$\begin{cases}x_{1,k+1} = x_{1,k} + hx_{2,k+1} \\x_{2,k+1} = x_{2,k} - h\sin(x_{1,k+1}) + h\sin(t_{k+1})\end{cases}$$

En posant  $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ , ceci s’écrit  $X_{k+1} = X_k + hF_k(X_{k+1})$  avec  $F_k(X) = \begin{bmatrix} x_2 \\ -\sin(x_1) + \sin(t_{k+1}) \end{bmatrix}$ .

En posant  $G(X) = X - X_k - hF_k(X)$ , on cherche donc  $X_{k+1}$  tel que  $G(X_{k+1}) = 0$ . On a

$$G(X) = \begin{bmatrix} x_1 - x_{1,k} - hx_2 \\ x_2 - x_{2,k} + h\sin(x_1) - h\sin(t_{k+1}) \end{bmatrix},$$

la dérivée de  $G$  au point  $X$  est donnée par sa matrice jacobienne :

$$J_G(X) = I - h \begin{bmatrix} 0 & 1 \\ -\cos(x_1) & 0 \end{bmatrix} = \begin{bmatrix} 1 & -h \\ h\cos(x_1) & 1 \end{bmatrix}$$

La méthode de Newton s’écrit donc :

initialisation :  $Z_0 = X_k$ ,

itérations :  $J_G(Z_\ell)(Z_{\ell+1} - Z_\ell) = -G(Z_\ell)$  pour  $\ell \geq 0$ .

La valeur de  $X_{k+1}$  est alors  $Z_\ell$  pour un indice  $\ell \geq 0$  choisi avec un critère d’arrêt, par exemple  $\|Z_\ell - Z_{\ell-1}\| \leq \varepsilon$  pour un  $\varepsilon > 0$ . En général, la méthode de Newton ne nécessite que quelques itérations.

## Cours 12. Comportement des solutions pour $n = 2$

Le but de ce chapitre est de montrer comment se comportent les solutions d'un système différentiel autonome, de deux équations à deux inconnues, grâce à l'étude de leurs trajectoires dans  $\mathbb{R}^2$ . Les connaissances acquises dans les chapitres précédents sont utiles tout au long de cette étude.

On suppose que  $n = 2$ , et on rappelle le problème de Cauchy étudié :

$$X'(t) = f(X(t)), \quad t > 0, \quad (12.1a)$$

$$X(0) = X^{(0)}, \quad (12.1b)$$

où la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  est localement lipschitzienne et  $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ .

On note  $\gamma$  la trajectoire (dans  $\mathbb{R}^2$ , c'est-à-dire dans le plan de phase) de la solution maximale, qui est définie par :

$$\gamma(X) = \{X(t), t \in [0, T_m[)\}.$$

où  $T_m$  est le temps d'existence de  $X$ . L'objet de ce cours est d'étudier l'allure de  $\gamma$ .

### 12.1 Comportement des solutions pour le cas linéaire autonome

On s'intéresse dans cette section au cas linéaire. Dans le cas linéaire, le système (12.1a) s'écrit  $X'(t) = AX(t)$ , avec  $A \in \mathcal{M}_2(\mathbb{R})$ . On distingue trois cas :  $A$  diagonalisable dans  $\mathbb{R}$ ,  $A$  non diagonalisable et  $A$  diagonalisable dans  $\mathbb{C}$ .

#### 12.1.1 $A$ diagonalisable dans $\mathbb{R}$

Si  $A$  diagonalisable dans  $\mathbb{R}$ , on note  $\lambda_1, \lambda_2 \in \mathbb{R}$  les valeurs propres et  $\varphi_1, \varphi_2$  les vecteurs propres associés. La solution du système (12.1a) est alors de la forme :

$$X(t) = C_1 e^{\lambda_1 t} \varphi_1 + C_2 e^{\lambda_2 t} \varphi_2, \quad C_1, C_2 \in \mathbb{R} \quad (12.2a)$$

$$\text{avec } A\varphi_i = \lambda_i \varphi_i, \quad \varphi_i \neq 0, \quad i \in \{1, 2\}, \quad (12.2b)$$

L'allure des solutions dépend des positions de  $\lambda_1$  et  $\lambda_2$  par rapport à 0.

Dans les figures d'illustration qui suivent, les vecteurs propres sont  $\varphi_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  et  $\varphi_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ .

Pour chaque cas (sauf le dernier), quatre conditions initiales différentes sont choisies, qui donnent les quatre trajectoires en rouge, bleu, vert et violet sur chaque figure.

- 1)  $A$  diagonalisable dans  $\mathbb{R}$  avec  $\lambda_1 < \lambda_2 < 0$  : figure 12.1. Comme les valeurs propres de  $A$  sont négatives, la solution tend vers 0 quand  $t \rightarrow +\infty$ . Comme  $\lambda_1 < \lambda_2 < 0$ , le terme  $e^{\lambda_1 t}$  tend vers 0 plus vite que  $e^{\lambda_2 t}$ , la trajectoire tend à être parallèle à la droite engendrée par  $\varphi_2$  (si  $C_2 \neq 0$ ).

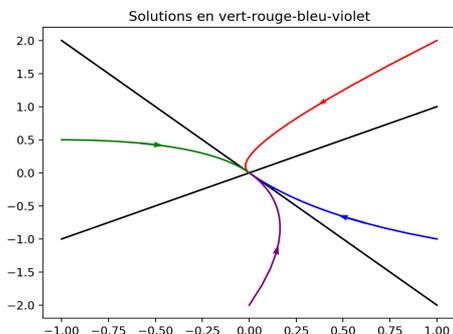


Fig. 12.1:  $\lambda_1 = -2, \lambda_2 = -1$

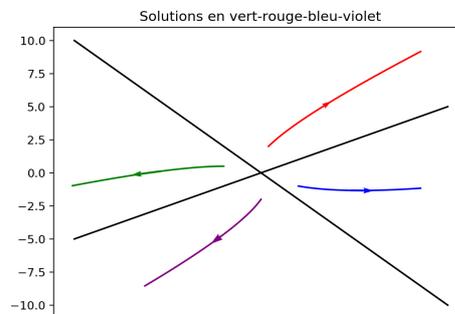


Fig. 12.2:  $\lambda_1 = 2, \lambda_2 = 1$

- 2)  $A$  diagonalisable dans  $\mathbb{R}$  avec  $0 < \lambda_2 < \lambda_1$  : figure 12.2. Dans ce cas, les solutions sont non bornées (quand  $t \rightarrow +\infty$ ). Comme  $\lambda_1 > \lambda_2$ , le terme  $e^{\lambda_1 t}$  tend vers  $+\infty$  plus vite que  $e^{\lambda_2 t}$ , la trajectoire tend à être parallèle à la droite engendrée par  $\varphi_1$  (si  $C_1 \neq 0$ ).
- 3)  $A$  est diagonalisable dans  $\mathbb{R}$  avec  $\lambda_1 = \lambda_2$  : figures 12.3 et 12.4.

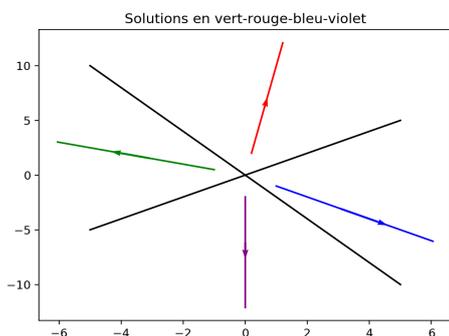


Fig. 12.3:  $\lambda_1 = \lambda_2 = 1$

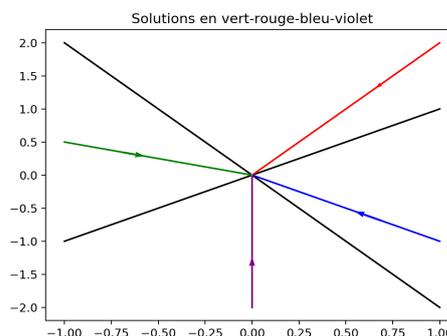


Fig. 12.4:  $\lambda_1 = \lambda_2 = -1$

Dans le cas  $\lambda_1 = \lambda_2 > 0$  (figure 12.3), les trajectoires partent à l'infini. Dans le cas  $\lambda_1 = \lambda_2 < 0$  (figure 12.4), elles tendent vers 0.

- 4)  $A$  diagonalisable dans  $\mathbb{R}$  avec  $\lambda_1 < 0 < \lambda_2$  : figure 12.5. Les solutions sont ici non bornées si  $C_2 \neq 0$  et les trajectoires tendent (quand  $t \rightarrow +\infty$ ) à être parallèles à la droite engendrée par  $\varphi_2$ .
- 5)  $A$  est diagonalisable dans  $\mathbb{R}$  avec  $\lambda_1 = 0 \neq \lambda_2$  : figure 12.6. Les trajectoires sont parallèles à  $\varphi_2$ . Elles partent à l'infini si  $\lambda_2 > 0$  et tendent vers une constante si  $\lambda_2 < 0$ .

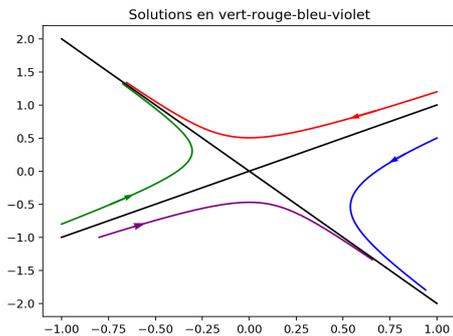


Fig. 12.5:  $\lambda_1 = -2, \lambda_2 = 1$

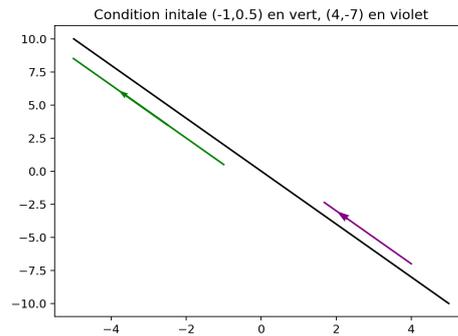


Fig. 12.6:  $\lambda_1 = 0, \lambda_2 = 1$  (vert) et  $-1$  (violet)

### 12.1.2 $A$ non diagonalisable

On suppose maintenant que  $A$  non diagonalisable. On a donc avec  $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$  et la solution générale est ( $C_1, C_2$  arbitraires dans  $\mathbb{R}$ ) :

$$C_1 e^{\lambda t} \varphi + C_2 e^{\lambda t} (\psi + t\varphi),$$

avec  $A\varphi = \lambda\varphi$  et  $A\psi - \lambda\psi = \varphi, \varphi \neq 0$ .

Trois cas sont à étudier, selon la position de  $\lambda$  par rapport à 0.

Dans les figures d'illustration qui suivent, le vecteur propre est  $\varphi = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ .

Comme précédemment, pour les deux premiers cas, quatre conditions initiales différentes sont choisies, qui donnent les quatre trajectoires en rouge, bleu, vert et violet sur chaque figure.

- 1)  $A$  non diagonalisable,  $\lambda < 0$  : figure 12.7. Dans ce cas, les solutions tendent vers 0 et les trajectoires tendent (quand  $t \rightarrow +\infty$ ) à être parallèles à la droite engendrée par  $\varphi$ .

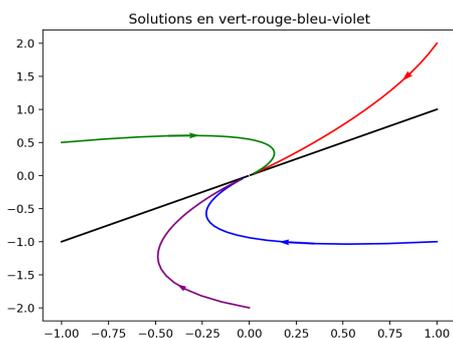


Fig. 12.7:  $\lambda = -1, A$  non diagonalisable

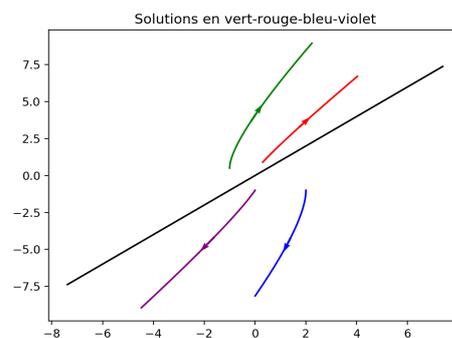


Fig. 12.8:  $\lambda = 1, A$  non diagonalisable

- 2)  $A$  non diagonalisable,  $\lambda > 0$  : figure 12.8. Dans ce cas, les solutions sont non bornées et tendent, quand  $t \rightarrow +\infty$ , à être parallèles à la droite engendrée par  $\varphi$  (si  $C_2 \neq 0$ ).

- 3)  $A$  non diagonalisable,  $\lambda = 0$ . Dans ce cas, la trajectoire (toujours au sens de “trajectoire dans  $\mathbb{R}^2$ ”) est alors une demi-droite parallèle à  $\varphi$  car la solution générale est alors

$$C_1\varphi + C_2(\psi + t\varphi).$$

### 12.1.3 $A$ diagonalisable dans $\mathbb{C}$

Les valeurs propres de  $A$  sont alors  $\lambda$  et  $\bar{\lambda}$  avec  $\lambda = \alpha + i\beta$ ,  $\beta \neq 0$ . Soit  $\varphi$  tel que  $A\varphi = \lambda\varphi$  avec  $\varphi = \varphi_1 + i\varphi_2$ ,  $\varphi_1, \varphi_2 \in \mathbb{R}^2$  non nuls. La solution générale du système  $X'(t) = AX(t)$  est alors, avec  $C_1, C_2$  arbitraires dans  $\mathbb{R}$  :

$$C_1e^{\alpha t}(\cos(\beta t)\varphi_1 - \sin(\beta t)\varphi_2) + C_2e^{\alpha t}(\sin(\beta t)\varphi_1 + \cos(\beta t)\varphi_2).$$

Ici encore, nous distinguons encore 3 cas selon la position de  $\alpha$  par rapport à 0.

- 1)  $A$  diagonalisable dans  $\mathbb{C}$  avec  $\alpha = 0$  : figure 12.9. La trajectoire (dans le plan  $x_1, x_2$ ) est une ellipse.

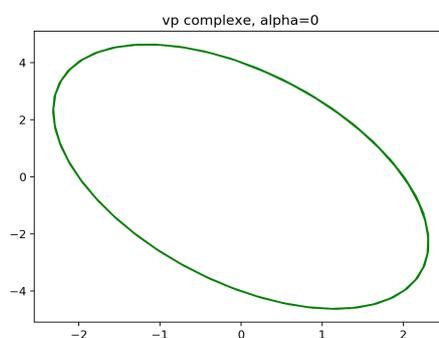


Fig. 12.9:  $A$  diagonalisable dans  $\mathbb{C}$  avec  $\lambda = i\beta$ ,  $\beta \neq 0$

- 2)  $A$  diagonalisable dans  $\mathbb{C}$  avec  $\alpha > 0$  : figure 12.10 La trajectoire a la forme d’une spirale. La norme de la solution tend vers  $+\infty$  quand  $t \rightarrow +\infty$ .

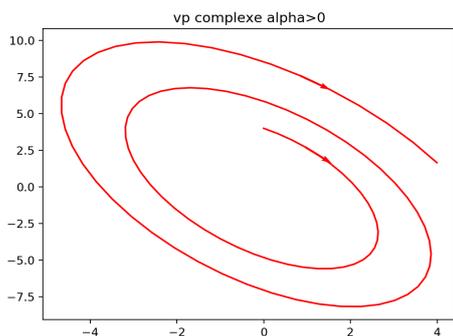


Fig. 12.10:  $A$  diagonalisable dans  $\mathbb{C}$ ,  $\alpha > 0$

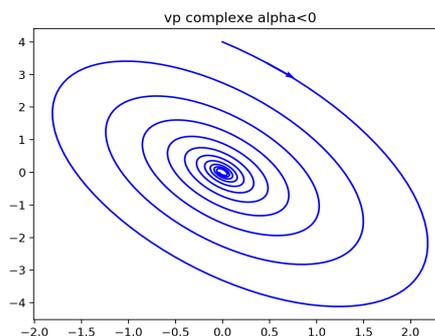


Fig. 12.11:  $A$  diagonalisable dans  $\mathbb{C}$ ,  $\alpha < 0$

- 3)  $A$  diagonalisable dans  $\mathbb{C}$  avec  $\alpha < 0$  : figure 12.11. La trajectoire a ici aussi la forme d'une spirale. La solution tend vers 0 quand  $t \rightarrow +\infty$ .

**Remarque 15** (Remarque sur la discrétisation dans le  $A$  diagonalisable dans  $\mathbb{C}$ ,  $\alpha = 0$ ). Le système différentiel est toujours  $X'(t) = AX(t)$ . Les valeurs propres de  $A$  sont  $\lambda$  et  $\bar{\lambda}$  avec  $\lambda = i\beta \neq 0$ . On note  $\varphi$  un vecteur propre associé à  $\lambda$ , donc  $A\varphi = i\beta\varphi$ ,  $\varphi \neq 0$ ,  $\varphi \in \mathbb{C}^2$ . Pour une condition initiale  $X^{(0)} \neq 0$ , la trajectoire de la solution exacte est une ellipse (figure 12.9) En particulier, ceci implique que  $\|X(t)\|$  ne tend pas vers 0 et ne tend pas vers  $+\infty$  quand  $t \rightarrow +\infty$ .

La discrétisation par le schéma d'Euler explicite, avec pas de temps  $h$ , s'écrit  $\frac{X_{n+1} - X_n}{h} = AX_n$ . Ceci donne, pour tout  $n \geq 0$ ,  $X_{n+1} = (I + hA)X_n$  et donc  $X_n = (I + hA)^n X_0$ . En prenant  $X_0 = \varphi$  (on triche un peu car  $\varphi \notin \mathbb{R}^2$ ), on obtient  $X_n = (1 + ih\beta)^n \varphi$  et donc, en choisissant une norme sur  $\mathbb{C}^2$ ,  $\|X_n\| = |1 + ih\beta|^n \|\varphi\| = (\sqrt{1 + h^2\beta^2})^n \|\varphi\| \rightarrow +\infty$  quand  $n \rightarrow +\infty$ , comme dans le cas  $\alpha > 0$ , figure 12.10.

Si on choisit la discrétisation par le schéma d'Euler implicite, avec pas de temps  $h$ ,  $\frac{X_{n+1} - X_n}{h} = AX_{n+1}$  et donc, pour tout  $n \geq 0$ ,  $X_{n+1} = (I - hA)^{-1} X_n$ . En prenant  $X_0 = \varphi$ , on obtient  $X_n = (\frac{1}{1 - ih\beta})^n \varphi$ ,  $\|X_n\| = |\frac{1}{1 - ih\beta}|^n \|\varphi\| = (\frac{1}{\sqrt{1 + h^2\beta^2}})^n \|\varphi\|$ . On en déduit que  $\|X_n\| \rightarrow 0$  quand  $n \rightarrow +\infty$ , comme dans le cas  $\alpha < 0$ , figure 12.11. Un exemple explicite est donné dans le tp4 et le td10.

Pour conclure cette étude du cas linéaire, on constate que, pour un système linéaire, il existe trois possibilités pour la solution  $X$  :

- 1)  $\|X(t)\| \xrightarrow{t \rightarrow +\infty} +\infty$ ,
- 2)  $\|X(t)\| \xrightarrow{t \rightarrow +\infty} 0$ ,
- 3)  $X$  est périodique (ce cas contenant le cas d'une solution stationnaire).

Dans la section suivante, nous allons voir que d'autres cas sont possibles si le système est non linéaire.

## 12.2 Comportement des solutions pour le cas non linéaire autonome

Considérons maintenant le problème de Cauchy (12.1) dans le cas non linéaire. La trajectoire dans  $\mathbb{R}^2$  est toujours notée  $\gamma(X) = \{X(t), t > 0\}$ . On s'intéresse au cas où la solution est globale ( $T_m = +\infty$ ) et bornée, c'est-à-dire qu'il existe  $A \in \mathbb{R}$  tel que  $\|X(t)\| \leq A$  pour tout  $t > 0$ .

### 12.2.1 La solution $X$ est convergente

**Proposition 6.** Soit  $X$  une solution de (12.1). On suppose qu'il existe  $a \in \mathbb{R}^2$  tel que  $\lim_{t \rightarrow +\infty} X(t) = a$ , alors  $a$  est un point d'équilibre, autrement dit  $f(a) = 0$ .

*Démonstration.* On note  $f_1$  et  $f_2$  les deux composantes de  $f$ ,  $x_1$  et  $x_2$  les deux composantes de  $X$  et  $a_1$  et  $a_2$  les deux composantes de  $a$ . Avec ces notations, le système s'écrit

$$\begin{aligned} x_1'(t) &= f_1(x_1(t), x_2(t)) \\ x_2'(t) &= f_2(x_1(t), x_2(t)), \end{aligned}$$

Par le théorème des accroissements finis, pour tout  $t > 0$ , il existe  $\theta_{1,t}, \theta_{2,t} \in ]0, 1[$  tels que :

$$\begin{aligned}x_1(t+1) - x_1(t) &= x_1'(t + \theta_{1,t}) = f_1(X(t + \theta_{1,t})), \\x_2(t+1) - x_2(t) &= x_2'(t + \theta_{2,t}) = f_2(X(t + \theta_{2,t})).\end{aligned}$$

Comme

$$f_1(x_1(t + \theta_{1,t}), x_2(t + \theta_{1,t})) \xrightarrow{t \rightarrow +\infty} f_1(a_1, a_2) = f_1(a)$$

et

$$f_2(x_1(t + \theta_{2,t}), x_2(t + \theta_{2,t})) \xrightarrow{t \rightarrow +\infty} f_2(a_1, a_2) = f_2(a)$$

on en déduit que  $f(a) = 0$ . □

### 12.2.2 La solution $X$ est bornée non convergente

On suppose toujours que la solution  $X$  est bornée et on suppose aussi qu'il n'y a pas de point d'équilibre. La solution est donc non convergente.

D'autre part, on sait que la trajectoire ne peut pas se recouper (car nous sommes dans le cas autonome). L'intuition suggère que la trajectoire va décrire une sorte de spirale sans converger vers un point (car il n'y a pas de point d'équilibre) et sans aller vers l'infini (car la solution est bornée), comme sur la figure 12.12.

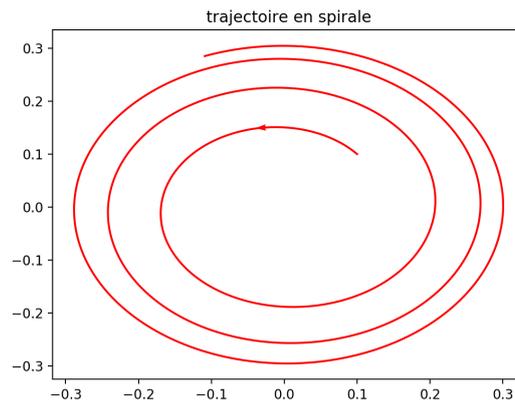


Fig. 12.12: Trajectoire en spirale

Ceci est précisé dans le théorème de Poincaré-Bendixson, qu'on admettra (la démonstration dépasse un peu le cadre de ce cours). Ce théorème donne un résultat qualitatif très intéressant sur les systèmes différentiels pour  $n = 2$ . Il concerne les systèmes du type (12.1a).

**Théorème 16** (Poincaré-Bendixson).

Soit  $X$  solution (maximale) de (12.1) et soit  $D$  un ensemble fermé borné de  $\mathbb{R}^2$ . On suppose que :

- 1)  $\{X(t), 0 \leq t < T_m\} \subset D$ .
- 2)  $D$  ne contient aucun point d'équilibre.

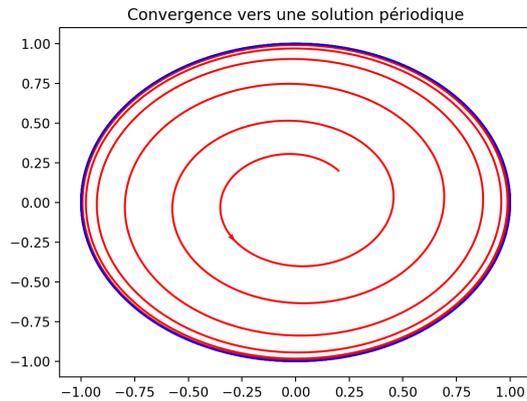


Fig. 12.13: Trajectoire en spirale convergeant vers une solution périodique

Alors,  $T_m = +\infty$  et  $X(t)$  “converge” vers une courbe qui est la trajectoire d’une solution périodique de (12.1), au sens suivant :

1) Il existe  $\tilde{X}$  solution périodique non constante de (12.1). On rappelle que “périodique” signifie

$$\exists T > 0, \tilde{X}(t + T) = \tilde{X}(t), \forall t > 0.$$

2)  $d(X(t), \gamma(\tilde{X})) = \inf\{|X(t) - y|, y \in \gamma(\tilde{X})\} \xrightarrow{t \rightarrow +\infty} 0.$

Plus exactement, on définit l’ensemble “ $\omega$ -limite” de  $X$ , noté  $\omega(X)$ , par :

$y \in \omega(X)$  si il existe une suite  $(t_n)_{n \in \mathbb{N}}$  telle que  $\lim_{n \rightarrow +\infty} t_n = +\infty$  et  $\lim_{n \rightarrow +\infty} X(t_n) = y,$

et on montre qu’il existe  $\tilde{X}$  solution périodique non constante telle que  $\omega(X) = \gamma(\tilde{X}).$

Noter que le fait que  $T_m = +\infty$  découle de l’hypothèse  $\{X(t), 0 \leq t < T_m\} \subset D$  (qui entraîne qu’il n’y a pas explosion en temps fini). Noter aussi que la solution périodique  $\tilde{X}$  est non constante car  $D$  ne contient aucun point d’équilibre.

La figure 12.13 illustre cette situation : la trajectoire de la solution (en rouge) converge vers la trajectoire d’une solution périodique (en bleu).

Voici maintenant quatre exemples de systèmes différentiels non linéaires autonomes pour  $n = 2.$

**Exemple 28** (Le pendule non linéaire). Rappelons l’équation du pendule :

$$x''(t) + \frac{g}{\ell} \sin(x(t)) = 0, t > 0. \quad (12.3)$$

Dans l’équation (12.3),  $g > 0$  est la constante de gravité et  $\ell > 0$  la longueur du fil (voir figure 5.1 page 41). On pose  $x_1 = x$  et  $x_2 = x'.$

L’équation (12.3) est alors équivalente au système suivant :

$$X'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = f(X(t)) = \begin{bmatrix} x_2(t) \\ -(g/\ell) \sin(x_1(t)) \end{bmatrix}.$$

On sait que  $(0, 0)$  est un point d'équilibre de ce système. La matrice jacobienne de  $f$  au point  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  est

$$J_f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ -\frac{g}{\ell} \cos(x_1) & 0 \end{bmatrix}.$$

Si  $x_1 = x_2 = 0$ ,

$$J_f(0, 0) = \begin{bmatrix} 0 & 1 \\ -\frac{g}{\ell} & 0 \end{bmatrix}.$$

Le polynôme caractéristique est  $\lambda^2 + \frac{g}{\ell} = 0$  et les valeurs propres sont  $\lambda = \pm i\sqrt{g/\ell}$ .

L'étude du problème linéarisé ne nous permet donc pas de conclure sur la stabilité ou l'instabilité du point  $(0, 0)$ . Cette étude peut être faite en montrant la conservation de l'énergie.

En multipliant (12.3) par  $x'(t)$  on obtient :

$$\frac{1}{2} \frac{d}{dt} (x'(t))^2 - \frac{d}{dt} \left( \frac{g}{\ell} \cos(x(t)) \right) = 0 \quad (12.4)$$

L'énergie totale du pendule est la somme de l'énergie cinétique et de l'énergie potentielle. A une constante multiplicative près, cette énergie est (à l'instant  $t$ )

$$E(t) = \frac{1}{2} (x'(t))^2 - \frac{g}{\ell} \cos(x(t)).$$

L'équation (12.4) exprime donc la conservation de l'énergie totale, c'est-à-dire  $E'(t) = 0$ .

Pour retrouver le fait que l'énergie est la somme de l'énergie potentielle et de l'énergie cinétique, on pose :

$$p(t) = \begin{bmatrix} \ell \sin(x(t)) \\ \ell \cos(x(t)) \end{bmatrix} \text{ et donc } p'(t) = \begin{bmatrix} \ell \cos(x(t))x'(t) \\ -\ell \sin(x(t))x'(t) \end{bmatrix}$$

Le point  $p(t)$  est la position du corps pesant à l'instant  $t$  (c'est la position du point  $M$  dans la figure 5.1). La vitesse du corps est  $p'(t)$ . On peut ainsi calculer l'énergie cinétique et l'énergie potentielle :

$$E_c(t) = \frac{1}{2} m \ell^2 (x'(t))^2 \text{ et } E_p(t) = -mg\ell \cos(x(t)).$$

L'énergie totale est la somme de  $E_c$  et  $E_p$ ,

$$E_{tot}(t) = E_c(t) + E_p(t) = \frac{1}{2} m \ell^2 (x'(t))^2 + mg\ell \cos(x(t)) \quad (12.5)$$

$$= m \ell^2 \left( \frac{1}{2} (x'(t))^2 - \frac{g}{\ell} \cos(x(t)) \right) \quad (12.6)$$

$$= m \ell^2 E(t) \quad (12.7)$$

En résumé, on a  $E(t) = E(0)$  pour tout  $t > 0$ . Ceci montre que la solution  $X$  est à chaque instant sur la courbe d'équation  $(1/2)x_2^2 - (g/\ell) \cos(x_1) = a$  avec  $a$  donné par la condition initiale.

Si  $x_2(0) = 0$  et  $|x_1(0)| < \pi/2$ , on peut aussi montrer que la solution est bornée et périodique (mais ce n'est pas fait ici).

Il est intéressant de voir aussi ce que donnent les schémas numériques pour ce système dans le cas  $x_2(0) = 0$  et  $|x_1(0)| < \pi/2$ . On note  $h$  le pas de temps.

- 1) Avec le schéma d'Euler explicite, pour  $g/\ell = 1$ ,  $x(0) = \pi/4$ ,  $x'(0) = 0$  et  $h = 0.1$ , la solution semble "exploser" pour  $T = 50$ . Mais la solution approchée converge néanmoins vers la solution exacte quand  $h \rightarrow 0$ .
- 2) Avec le schéma d'Euler implicite, toujours pour  $g/\ell = 1$ ,  $x(0) = \pi/4$ ,  $x'(0) = 0$  et  $h = 0.1$ , la solution semble tendre vers 0 pour  $T = 50$ . Ici aussi la solution approchée converge vers la solution exacte quand  $h \rightarrow 0$ . Noter qu'avec ce schéma implicite, la méthode de résolution (à chaque pas de temps) par l'algorithme de Newton fonctionne très bien.
- 3) Avec le schéma de Heun, la solution approchée est très nettement meilleure que pour les schémas d'Euler. La solution approchée est déjà très proche de la solution exacte pour  $g/\ell = 1$ ,  $x(0) = \pi/4$ ,  $x'(0) = 0$  et  $h = 0.1$ .

**Exemple 29** (Équations de Lotka-Volterra, système proie-prédateur). Soient  $a, b, c, d \in \mathbb{R}_+^*$ , et  $x_0, y_0 \in \mathbb{R}_+^*$ ; on considère le système différentiel suivant :

$$x'(t) = ax(t) - bx(t)y(t), \quad t > 0, \quad (12.8)$$

$$y'(t) = -cy(t) + dx(t)y(t), \quad t > 0, \quad (12.9)$$

avec la condition initiale

$$x(0) = x_0, \quad (12.10)$$

$$y(0) = y_0. \quad (12.11)$$

On peut montrer que les solutions sont périodiques (cela faisait partie de l'examen de janvier 2020). Les schémas de Heun et de RK4 donnent une excellente approximation sans avoir besoin de prendre un pas de temps très petit.

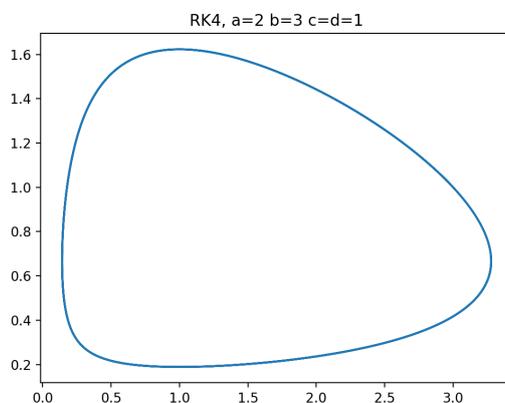


Fig. 12.14: Système proie-prédateur

La figure 12.14 donne la trajectoire dans  $\mathbb{R}^2$  (plan  $(x_1, x_2)$ ) de la solution périodique autour du point d'équilibre  $\left(\frac{c}{d}, \frac{a}{b}\right)$  pour  $x_0 = 3$  et  $y_0 = 1$ .

**Exemple 30** ( Effet tunnel (TD8)).

On considère le problème de Cauchy suivant :

$$x_1'(t) = x_2(t) + x_1(t) - x_1^3(t), t > 0 \quad (12.12a)$$

$$x_2'(t) = -x_1(t), t > 0 \quad (12.12b)$$

On pose  $X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$  et

$$f(X(t)) = \begin{bmatrix} x_2(t) + x_1(t) - x_1^3(t) \\ -x_1(t) \end{bmatrix},$$

de sorte que le système s'écrit  $X'(t) = f(X(t))$ . Le seul point d'équilibre est 0 (c'est-à-dire  $x_1 = x_2 = 0$ ). La matrice jacobienne de  $f$  au point  $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  est

$$J_f(x_1, x_2) = \begin{bmatrix} 1 - 3x_1^2 & 1 \\ -1 & 0 \end{bmatrix},$$

et donc

$$J_f(0) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}.$$

Les valeurs propres de  $J_f(0)$  sont complexes conjuguées,  $\lambda = \frac{1 \pm i\sqrt{3}}{2}$ . Le point d'équilibre 0 est donc instable. Plus précisément, soit  $X$  une solution maximale du système 12.12a-12.12b, on peut montrer qu'il existe  $\varepsilon > 0$  t.q.  $\|X(0)\| \geq \varepsilon$  implique  $\|X(t)\| \geq \varepsilon$  pour tout  $t \in [0, T_m[$  (où  $T_m$  est le temps d'existence de cette solution maximale).

Il est assez facile de montrer que  $T_m = +\infty$ . En effet, On pose

$$E(t) = \frac{1}{2}x_1^2(t) + \frac{1}{2}x_2^2(t).$$

En multipliant (12.12a) par  $x_1$  et (12.12b) par  $x_2$ , on obtient

$$E'(t) = x_1^2(t) - x_1^4(t) = x_1^2(t)(1 - x_1^2(t))(1 + x_1(t)). \quad (12.13)$$

De l'égalité (12.13), on déduit que  $E'(t) \leq 1$  (car le maximum de la fonction  $y \mapsto y^2(1 - y)(1 + y)$  est inférieur à 1). On en déduit  $E(t) \leq E(0) + t$  et donc  $x_1^2(t) + x_2^2(t) \leq 2E(0) + 2t$  pour tout  $t \in [0, T_m[$ . La solution est donc globale (c'est-à-dire  $T_m = +\infty$ ).

Nous pouvons aussi déduire de (12.13) le signe de  $E'(t)$  :

$$E'(t) \geq 0 \text{ si } x_1 \in ]-1, 1[ \text{ et } E'(t) \leq 0 \text{ si } x_1 \notin ]-1, 1[.$$

Donc si  $|x_1| > 1$  alors  $E$  décroît et si  $|x_1| < 1$   $E$  croît.

Un travail supplémentaire (non détaillé ici) permet de montrer que la solution est bornée. Il existe  $A > 0$  tel que  $\|X(t)\| \leq A$  pour tout  $t \geq 0$ . En prenant  $D = \{Y \in \mathbb{R}^2, \varepsilon \leq \|Y\| \leq A\}$ , le théorème 16 s'applique, et  $X(t)$  "converge" vers la trajectoire d'une solution périodique. Ceci prouve, en particulier, l'existence d'une solution périodique.

Un exemple de solution est donné dans la figure 12.15 (dans le plan  $(x_1, x_2)$ ). La condition initiale est  $x_1(0) = 0.5, x_2(0) = 0$ .

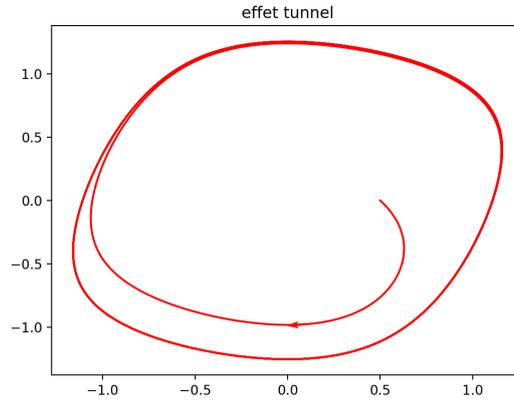


Fig. 12.15: Effet tunnel, solution avec  $x_1(0) = 0.5, x_2(0) = 0$

**Exemple 31** (Exemple tiré du livre de Martin Braun).

Considérons le problème suivant :

$$x_1'(t) = -x_2(t) + x_1(t)(1 - x_1(t)^2 - x_2(t)^2), \quad t > 0 \quad (12.14a)$$

$$x_2'(t) = x_1(t) + x_2(t)(1 - x_1(t)^2 - x_2(t)^2), \quad t > 0. \quad (12.14b)$$

En multipliant (12.14a) par  $x_1$  et (12.14b) par  $x_2$  on obtient :

$$\frac{1}{2}(x_1^2 + x_2^2)' = (x_1^2 + x_2^2)(1 - x_1^2 - x_2^2)$$

On pose alors  $r^2 = x_1^2 + x_2^2$ ,  $r \geq 0$ , et on obtient  $\frac{1}{2}(r^2)' = r^2(1 - r^2)$ . On effectue un autre changement de variable en posant  $r^2 = z$ . On obtient ainsi une nouvelle équation :

$$z' = 2z(1 - z)$$

On sait résoudre cette équation (ceci peut se faire, par exemple, en décomposant en éléments simples l'expression  $1/z(1 - z)$ ) on trouve ainsi  $z$  puis  $r$ . En notant  $r_0$  la valeur de  $r$  pour  $t = 0$ , la solution est, si  $r_0 \neq 1$ ,

$$r(t) = \frac{r_0}{\sqrt{r_0^2 + e^{-2t}(1 - r_0^2)}} \xrightarrow{t \rightarrow +\infty} 1.$$

(En particulier, ceci montre que les solutions maximales de 12.14a-12.14b sont globales). La figure 12.16 donne la solution pour  $r$  (en fonction de  $t$ ) lorsque  $r_0 = 0.5$  (courbe bleue) et lorsque  $r_0 = 1.5$  (courbe rouge).

Pour le système 12.14a-12.14b, le seul point d'équilibre est 0. Lorsque  $r_0 > 0$ , comme  $\lim_{t \rightarrow +\infty} r(t) = 1$  (et que  $r(t) \neq 0$  pour tout  $t$ ) la solution reste dans une partie  $D$  fermée bornée de  $\mathbb{R}^2$ , le théorème 16 s'applique et la trajectoire de cette solution converge vers la trajectoire d'une solution périodique. La trajectoire de cette solution périodique est le cercle de rayon 1.

En fait, il est possible de résoudre complètement le problème de Cauchy pour le système 12.14a-12.14b. On a déjà trouvé la solution pour  $r$ . Un petit calcul donne que  $x_1(t) = r(t) \cos(\theta(t))$  et  $x_2(t) = r \sin(\theta(t))$  avec  $\theta'(t) = 1$  et donc  $\theta(t) = \theta(0) + t$  pour tout  $t \geq 0$ .

La figure 12.17 ci-dessous donne la trajectoire dans  $\mathbb{R}^2$  pour la condition initiale  $x_1(0) = x_2(0) = 0.1$ .

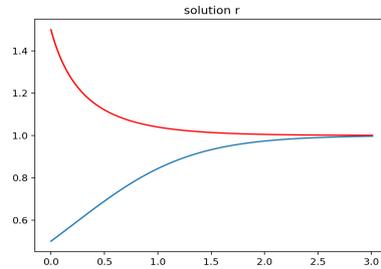


Fig. 12.16: Graphe de  $r = \sqrt{x_1^2 + x_2^2}$ ,  $(x_1, x_2)$  solution de (12.14) pour  $r_0 = 0.5$  (en bleu) et  $r_0 = 1.5$  (en rouge)

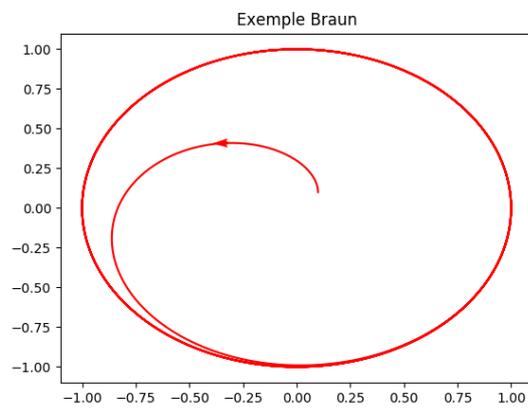


Fig. 12.17: Trajectoire pour le problème (12.14), avec condition initiale  $x_1(0) = x_2(0) = 0.1$