

Notes du cours Equations Différentielles
Math318

Thierry Ramond
Mathématiques
Université Paris Sud
e-mail: thierry.ramond@math.u-psud.fr

version du 5 mars 2015

Table des matières

| | | |
|----------|--|----------|
| 1 | Exemples et Notions de base | 6 |
| 1.1 | Motivations | 6 |
| 1.2 | La plus simple des équations différentielles | 7 |
| 1.2.1 | Résoudre une équation différentielle | 7 |
| 1.2.2 | Résolution des EDL scalaires d'ordre 1 homogènes | 8 |
| 1.2.3 | Solutions à valeurs complexes | 9 |
| 1.2.4 | Résolution des EDL scalaires d'ordre 1 avec second membre | 10 |
| 1.2.5 | Equations non résolues | 11 |
| 1.3 | Une équation d'ordre 2 | 11 |
| 1.4 | Une équation non-linéaire | 14 |
| 1.4.1 | Des solutions de l'équation $y' = -y^2$ | 14 |
| 1.4.2 | Retour sur la notion de solution | 14 |
| 1.5 | Solutions maximales | 15 |
| 1.6 | Problème de Cauchy | 16 |
| 1.6.1 | Mettons un peu d'ordre | 16 |
| 1.6.2 | Le théorème de Cauchy-Lipschitz. Première visite | 18 |
| 1.6.3 | Exemple : résolution de l'équation $y' = -y^2$ | 19 |
| 1.6.4 | Exemple : résolution de $y' = ky - y^2$ | 19 |
| 1.7 | Equations différentielles d'ordre supérieur | 21 |
| 1.7.1 | Qu'est-ce qu'une équation différentielle d'ordre n ? | 21 |
| 1.7.2 | Se ramener à un système de n équations différentielles d'ordre 1 | 22 |
| 1.7.3 | Au fait : c'est quoi un système différentiel d'ordre 1 ? | 23 |

| | | |
|----------|---|-----------|
| 1.7.4 | Le théorème de Cauchy-Lipschitz pour les systèmes, et pour les équations d'ordre n . Première visite. | 24 |
| 1.8 | Systèmes linéaires | 25 |
| 1.8.1 | Qu'est-ce qu'un système linéaire ? | 25 |
| 1.8.2 | Structure de l'ensemble des solutions | 26 |
| 1.8.3 | Dimension de l'espace des solutions | 27 |
| 1.8.4 | Méthode de variation de la constante | 28 |
| 2 | Réduction des matrices | 31 |
| 2.1 | Matrice d'une application linéaire | 31 |
| 2.1.1 | Définitions | 31 |
| 2.1.2 | Changement de base | 32 |
| 2.1.3 | Sous-espaces stables | 33 |
| 2.2 | Valeurs propres et vecteurs propres | 34 |
| 2.2.1 | Définitions | 34 |
| 2.2.2 | Polynôme caractéristique | 37 |
| 2.2.3 | Le théorème de Cayley-Hamilton | 38 |
| 2.3 | Espaces propres généralisés | 39 |
| 2.3.1 | Définition et propriétés | 39 |
| 2.3.2 | Condition nécessaire et suffisante de diagonalisabilité | 42 |
| 2.4 | Forme normale de Jordan | 43 |
| 2.4.1 | Matrices nilpotentes | 43 |
| 2.4.2 | Forme normale de Jordan | 44 |
| 2.4.3 | Méthode pratique pour le calcul | 46 |
| 3 | Systèmes linéaires | 48 |
| 3.1 | Exponentielle de matrice | 49 |
| 3.1.1 | Norme d'une matrice | 49 |
| 3.1.2 | Définition | 50 |
| 3.1.3 | La fonction $t \mapsto e^{tA}$ | 51 |
| 3.1.4 | Quelques propriétés | 53 |
| 3.1.5 | Calcul de e^{tA} | 55 |

| | | |
|----------|---|-----------|
| 3.2 | Solutions des systèmes linéaires à coefficients constants | 56 |
| 3.2.1 | Calcul effectif de $e^{tA}X_0$ | 57 |
| 3.2.2 | Solutions réelles | 58 |
| 3.2.3 | Variation de la constante | 59 |
| 3.3 | Systèmes linéaires à coefficients variables | 60 |
| 3.3.1 | Résolvante | 60 |
| 3.3.2 | Système fondamental de solutions - Wronskien | 62 |
| 3.3.3 | Méthode de variation de la constante | 64 |
| 4 | Existence et unicité | 65 |
| 4.1 | Le cas linéaire | 65 |
| 4.1.1 | Formulation intégrale | 65 |
| 4.1.2 | Schéma de Picard | 66 |
| 4.1.3 | Unicité | 67 |
| 4.2 | Le cas autonome | 68 |
| 4.2.1 | Fonctions Lipschitziennes | 69 |
| 4.2.2 | Fonctions localement Lipschitziennes | 69 |
| 4.2.3 | Existence dans le cas autonome | 70 |
| 4.3 | Le cas général | 71 |
| 4.3.1 | Le théorème du point fixe | 72 |
| 4.3.2 | Tonneau de sécurité | 73 |
| 4.3.3 | Existence et unicité locale | 74 |
| 4.3.4 | Solution maximale | 76 |
| 4.4 | Existence globale | 76 |
| 4.4.1 | Un critère d'existence globale | 76 |
| 4.4.2 | Qu'est-ce qui empêche une solution d'être globale ? | 78 |
| 4.4.3 | Le lemme de Gronwall | 80 |
| 5 | Stabilité et linéarisation | 82 |
| 5.1 | Portraits de phase des systèmes linéaires homogènes 2×2 à coefficients constants | 83 |
| 5.2 | Notions de stabilité | 86 |

| | | |
|----------|---|-----------|
| 5.2.1 | Définitions | 86 |
| 5.2.2 | Cas des systèmes linéaires à coefficients constants | 87 |
| 5.3 | Fonction de Lyapounov | 88 |
| 5.4 | Linéarisation et stabilité | 91 |
| 5.4.1 | Système linéarisé | 91 |
| 5.4.2 | Le critère de Routh | 92 |
| 5.4.3 | Le Théorème d'Hartman-Grobman | 93 |
| 6 | Systèmes prédateurs-proies | 95 |
| 6.1 | Existence et unicité. Temps d'existence | 95 |
| 6.2 | Points d'équilibre | 96 |
| 6.3 | Comportement global | 97 |
| 6.4 | Une fonction de Lyapounov | 98 |

Préambule

Ce cours s'adresse à des étudiants de L3 en mathématiques, dans le cadre de l'UE Math318 du Département de Mathématiques d'Orsay. Peut-être pourra-t-il également intéresser un public plus large d'étudiants en sciences (physique, biologie. . .).

Le matériel présenté ici est inspiré de plusieurs sources, dont en particulier le livre de Jean-Pierre Demailly "Analyse Numérique des équations différentielles", dont je ne peux que recommander très fortement la lecture. Il est également très proche de ce que l'on peut trouver dans les notes de cours préparées par Françoise Issard-Roch.

Ces notes sont encore sous forme préliminaire. Tous les commentaires et les suggestions sont les bienvenus. Elles contiennent encore trop peu d'exemples, et le lecteur devra se reporter aux exercices proposés en Travaux Dirigés pour des illustrations des résultats présentés ici.

Chapitre 1

Exemples et Notions de base

1.1 Motivations

La modélisation mathématique des phénomènes du monde réel passe très souvent par l'écriture d'une équation différentielle, qui décrit les variations de la quantité que l'on veut étudier en fonction d'un paramètre. C'est probablement Isaac Newton qui a le premier écrit et étudié des équations différentielles. Son principe fondamental de la dynamique s'écrit par exemple

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

où $t \mapsto x(t)$ décrit la trajectoire du centre de gravité d'un objet de masse m , soumis au champ de force $x \mapsto F(x)$. Il est d'ailleurs clair que pour connaître la position $x(t)$ de l'objet à l'instant t , il faut connaître non seulement la règle qui gouverne les variations de cette fonction, mais aussi la position et la vitesse initiale du centre de masse. Dans le cas d'un objet se déplaçant verticalement sous l'action de l'attraction terrestre, supposée constante dans la région où le mouvement a lieu, et compte tenu de la résistance de l'air, l'équation différentielle ci-dessus s'écrit

$$mx''(t) = -kx'(t) + mg,$$

où k, g sont des constantes. Un autre exemple célèbre est celui du pendule : une masse m est suspendue à l'extrémité d'une corde de longueur ℓ dont l'autre extrémité est fixe. L'angle $\theta(t)$ que fait, à l'instant t , le fil avec la verticale vérifie la relation

$$\theta''(t) + \frac{g}{\ell} \sin(\theta(t)) = 0.$$

Cette équation n'a pas de solution que l'on puisse exprimer simplement avec des fonctions connues. Il faut avoir à l'esprit que c'est le cas de la plupart des équations différentielles, même si les exercices proposés aux étudiants consistent souvent à trouver une solution explicite - c'est plus facile. L'étude d'une équation différentielle consiste la plupart du temps à décrire certaines propriétés des solutions sans les calculer, et c'est cela que ce cours a pour ambition de vous enseigner.

On trouve des équation différentielles dans tous les domaines de la physique. Par exemple la charge électrique $q(t)$ d'un condensateur dans un circuit RLC (résistance-bobine-condensateur),

alimentée par une source de courant alternatif, doit vérifier l'équation

$$Lq''(t) + Rq'(t) + \frac{1}{C}q(t) = U \cos(\omega t).$$

La biologie est aussi une source importante d'équations différentielles. Si l'on est amené à modéliser l'évolution d'une population de bactéries, on arrive naturellement à l'idée que le taux de croissance de leur nombre est proportionnel à ce nombre. Celui-ci est alors décrit par l'équation différentielle

$$N'(t) = kN(t),$$

où k est une constante, que l'on pourra ajuster de sorte que la solution de l'équation coïncide au mieux avec les données expérimentales. La modélisation des systèmes prédateurs-proies, due à Lotka et Volterra, est particulièrement éclairante. On étudiera des systèmes différentiels de la forme

$$\begin{cases} y_1'(t) = ay_1(t) - by_1(t)y_2(t), \\ y_2'(t) = -cy_2(t) + dy_1(t)y_2(t), \end{cases}$$

Ici a, b, c et d sont des constantes positives, et Volterra a postulé que les solutions $(y_1(t), y_2(t))$ décrivent l'évolution au cours du temps de la population de proies (des sardines, des lapins...) ($y_1(t)$) et de leurs prédateurs ($y_2(t)$) (des requins, des loups...). Ce type de système est également très utilisé en économie,

Signalons enfin que les mathématiques elles-mêmes peuvent être source d'équations différentielles. Si l'on cherche les fonctions dérivables y telles que $y(a+b) = y(a)y(b)$ pour tous réels a, b , on arrive très vite à l'équation $y' = ky$ où $k = y'(0)$. Il est d'ailleurs important de noter que résoudre l'équation

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

c'est trouver les courbes $(t, y(t))$ dont le vecteur tangent au point d'abscisse t est $(1, f(t, y(t)))$.

1.2 La plus simple des équations différentielles

On s'intéresse tout d'abord à l'équation différentielle très simple

$$(1.1) \quad y' = a(t)y$$

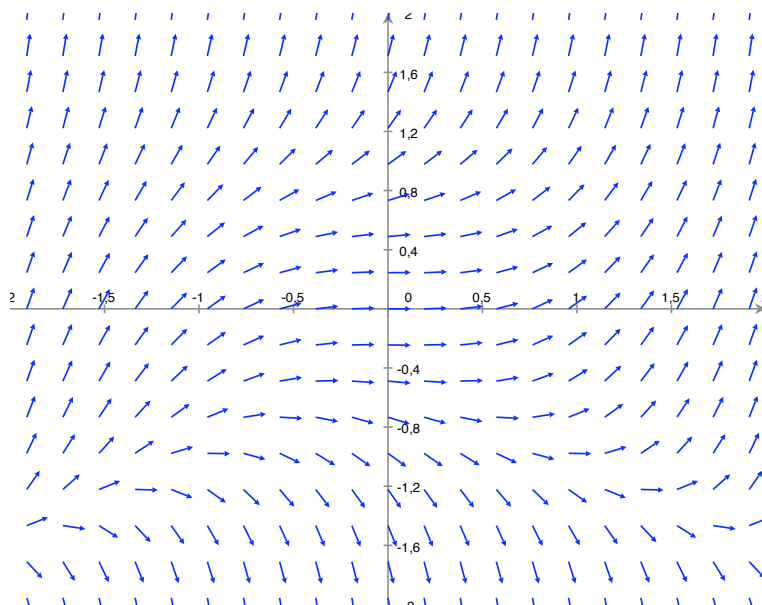
qui modélise les phénomènes dont la vitesse de croissance est proportionnelle à la taille, avec un taux (le coefficient de proportionnalité) $a(t)$ éventuellement variable dans le temps. On peut penser par exemple à l'évolution

- d'une population de bactéries,
- d'une quantité de matière radioactive,
- d'une somme d'argent placée à taux variable $a(t)$.

1.2.1 Résoudre une équation différentielle

On veut résoudre cette équation, c'est-à-dire trouver (soyons ambitieux) toutes les solutions de l'équation. Qu'est-ce que cela signifie ? On entend par là trouver toutes les fonctions $t \mapsto y(t)$ telles que pour tout t ,

$$y'(t) = a(t)y(t).$$

FIGURE 1.1 – Le champ de vecteurs $f(t, y) = (1, t^2 + y^3)$

Tout d'abord, pour que cette équation ait un sens, il est indispensable que y soit dérivable. Elle sera donc automatiquement continue. On va se placer dans le cas où la fonction a est continue. Si y est une solution de l'équation, y' sera donc continue. On peut alors préciser la question :

Soit a une fonction continue. Résoudre l'équation $y' = a(t)y$, c'est trouver toutes les fonctions de classe C^1 telles que $y'(t) = a(t)y(t)$ pour tout t .

Reste un point : pour quels t ? Si a est définie et continue sur l'intervalle ouvert I , il semble raisonnable de chercher toutes les fonctions $y \in C^1(I)$ telle que, pour tout $t \in I$, $y'(t) = a(t)y(t)$. Pour cette équation, on va voir dans un instant que c'est effectivement possible. Pour d'autres équations, on va rapidement voir qu'il faut être moins gourmand. Au passage, l'ensemble $C^1(I)$ étant beaucoup plus simple à définir et à manipuler lorsque I est ouvert (plutôt que fermé), on se placera dans ce cadre.

1.2.2 Résolution des EDL scalaires d'ordre 1 homogènes

Allons-y ! Dans le cas où $a(t)$ est une fonction constante sur l'intervalle I , on doit trouver toutes les fonctions $y \in C^1(I)$ telles que

$$(1.2) \quad \forall t \in I, \quad y'(t) - ay(t) = 0.$$

Il se trouve que lorsque $y \in C^1(I)$,

$$(e^{-at}y(t))' = e^{-at}(-ay(t) + y'(t)).$$

Donc (1.2) équivaut à

$$\forall t \in I, (e^{-at}y(t))' = 0.$$

Nous voilà arrivé à une équation différentielle que l'on sait résoudre !

$$(1.2) \iff \exists C \in \mathbb{R}, e^{-at}y(t) = C \iff \exists C \in \mathbb{R}, y(t) = Ce^{at}$$

Autrement dit, on a prouvé que

Soit $a \in \mathbb{R}$. L'ensemble \mathcal{S} des solutions de l'équation $y' = ay$ sur l'intervalle ouvert I est

$$\mathcal{S} = \{t \mapsto Ce^{at}, C \in \mathbb{R}\}.$$

Passons au cas général où $a(t)$ n'est pas forcément constante sur l'intervalle ouvert I . On essaye de procéder de la même manière, c'est-à-dire trouver une fonction $A(t)$ telle que

$$\forall t \in I, y'(t) - ay(t) = 0 \iff \forall t \in I, (e^{-A(t)}y(t))' = 0.$$

Puisque $(e^{-A(t)}y(t))' = e^{-A(t)}(-A'(t)y(t) + y'(t))$, on voit qu'il suffit de prendre pour A n'importe quelle primitive de la fonction a . Cela tombe bien : puisque a est continue sur I , la fonction a admet des primitives sur cet intervalle. On a donc obtenu

Soit $a : I \rightarrow \mathbb{R}$ une fonction continue sur l'intervalle ouvert I . L'ensemble \mathcal{S} des solutions de l'équation $y' = a(t)y$ sur I est

$$\mathcal{S} = \{t \mapsto Ce^{A(t)}, C \in \mathbb{R}\},$$

où $A : I \rightarrow \mathbb{R}$ est une primitive de a sur I .

1.2.3 Solutions à valeurs complexes

Jusqu'ici, on n'a considéré que des fonctions à valeurs réelles. Que ce soit le coefficient $a(t)$ où les solutions $y(t)$ de l'équation (1.1). Il peut arriver que l'on cherche les solutions à valeurs dans \mathbb{C} , en particulier (mais pas exclusivement) quand $a(t)$ est lui-même à valeurs complexes. En relisant ce qui précède dans ce cadre, on arrive très facilement à

Soit $a : I \rightarrow \mathbb{C}$ une fonction continue sur l'intervalle ouvert I . L'ensemble \mathcal{S} des solutions à valeurs complexes de l'équation $y' = a(t)y$ sur I est

$$\mathcal{S} = \{t \mapsto Ce^{A(t)}, C \in \mathbb{C}\},$$

où $A : I \rightarrow \mathbb{R}$ est une primitive de a sur I .

On peut être un peu inquiet à propos de la notion de primitive d'une fonction à valeurs complexes. On aurait tort. Si $a : I \rightarrow \mathbb{C}$, la fonction a s'écrit $a(t) = a_1(t) + ia_2(t)$ où $a_{1,2} : I \rightarrow \mathbb{R}$ sont deux fonctions à valeur réelles, respectivement la partie réelle et la partie imaginaire de la fonction a . Elles sont définies par

$$a_1(t) = \frac{a(t) + \overline{a(t)}}{2} \text{ et } a_2(t) = \frac{a(t) - \overline{a(t)}}{2i}.$$

Ce sont donc des fonctions continues sur I , et les primitives de a sur I sont les fonctions $A_1 + iA_2$ où A_1 et A_2 sont des primitives respectives de a_1 et a_2 .

Exercice 1.2.1 Donner les primitives de $t \mapsto e^{3it}$.

Exercice 1.2.2 Montrer que si a est à valeurs réelles, et y une solution de (1.1) à valeurs complexes, alors $y_1 = \operatorname{Re} y$ et $y_2 = \operatorname{Im} y$ sont aussi des solutions. En déduire que

Lorsque a est à valeurs réelles, il suffit de connaître les solutions de (1.1) à valeurs réelles pour connaître toutes les solutions de (1.1) à valeurs complexes.

1.2.4 Résolution des EDL scalaires d'ordre 1 avec second membre

On complique un peu les choses. On s'intéresse maintenant aux équations différentielles de la forme

$$(1.3) \quad y' = a(t)y + b(t),$$

où a et b sont deux fonctions continues sur l'intervalle ouvert $I =]\alpha, \beta[$.

Proposition 1.2.3 L'ensemble des solutions à valeurs complexes de l'équation (1.3) sur I est

$$\mathcal{S} = \{t \mapsto e^{A(t)}(C + \int_{\alpha}^t e^{-A(s)}b(s)ds), C \in \mathbb{C}\},$$

où A est une primitive de a sur I .

Preuve: Soit A une primitive de a sur I . Supposons que $y : I \rightarrow \mathbb{C}$ soit une solution de (1.3). On pose, pour $t \in I$, $w(t) = e^{-A(t)}y(t)$. On a

$$\begin{aligned} w'(t) &= -a(t)e^{-A(t)}y(t) + e^{-A(t)}y'(t) \\ &= -a(t)e^{-A(t)}y(t) + e^{-A(t)}(a(t)y(t) + b(t)) \\ &= e^{-A(t)}b(t). \end{aligned}$$

Donc il existe une constante $C \in \mathbb{C}$ telle que

$$w(t) = \int_{\alpha}^t e^{-A(s)}b(s)ds + C,$$

et

$$y(t) = e^{A(t)}w(t) = e^{A(t)}\left(C + \int_{\alpha}^t e^{-A(s)}b(s)ds\right).$$

Réciproquement, on vérifie facilement que toutes les fonctions $y : t \mapsto e^{A(t)}\left(C + \int_{t_0}^t e^{-A(s)}b(s)ds\right)$ sont des solutions de (1.3) sur I . \square

On notera que le choix d'intégrer à partir de α est arbitraire, et n'importe quel autre élément de $[\alpha, \beta]$ ferait l'affaire (cela revient à changer la valeur de la constante C).

Remarque 1.2.4 Plus que la formule ci-dessus, il faut retenir la méthode qui y conduit, en particulier l'idée de changer de fonction inconnue pour se ramener à une équation que l'on sait résoudre.

1.2.5 Equations non résolues

On désigne ces équations sous le nom d'équations différentielles d'ordre 1, linéaires, inhomogènes :

- d'ordre 1 parce que c'est le degré de dérivation le plus élevé qui apparait dans l'équation,
- linéaire parce que la partie homogène (celle où apparaissent y et ses dérivées) est linéaire par rapport à y , on y reviendra,
- inhomogène en raison de la présence de $b \neq 0$: la fonction nulle n'est pas solution de l'équation.

On dit parfois que les équations différentielles de la forme (1.3) sont des équations différentielles résolues, parce que la dérivée de plus haut degré s'écrit comme fonction des dérivées d'ordre inférieur, et l'on s'autorise à considérer comme équations différentielles, d'ordre 1 par exemple, des équations de la forme

$$(1.4) \quad u(t)y' + v(t)y = g(t),$$

où u, v et g sont des fonctions continues. Bien entendu, sur chaque intervalle où la fonction u ne s'annule pas, (1.4) est équivalente à l'équation différentielle

$$y' = a(t)y + b(t),$$

avec $a(t) = -v(t)/u(t)$ et $g(t) = f(t)/u(t)$, que l'on sait maintenant résoudre. Par exemple l'équation $(t+2)y' = 2 - y$ n'est pas une équation différentielle résolue sur \mathbb{R} . Mais l'on peut tout de même se demander quelles sont les fonctions de classe \mathcal{C}^1 sur \mathbb{R} qui vérifient cette équation.

1.3 Une équation d'ordre 2

On étudie maintenant les équations du type de celle obtenue pour décrire l'évolution de la charge dans un circuit RLC, c'est-à dire de la forme

$$(1.5) \quad y'' + p(t)y' + q(t)y = f(t)$$

où p, q et f sont des fonctions continues sur l'intervalle ouvert $I =]\alpha, \beta[$.

Dans le cas où p et q sont des fonctions constantes, on peut comme pour l'équation (1.3) décrire très explicitement l'ensemble des solutions. On se concentre sur le cas où $f = 0$, autrement dit on considère l'équation

$$(1.6) \quad y'' + py' + qy = 0,$$

et on traite le cas (le plus important) où $p, q \in \mathbb{R}$. On reviendra plus tard, mieux armés, sur le cas où $f \neq 0$.

On peut deviner le résultat qui suit en essayant de trouver une solution de la forme $y : t \mapsto e^{rt}$. En calculant les dérivées de cette fonction, on voit que l'on doit nécessairement avoir

$$r^2 + pr + q = 0.$$

Cette relation est souvent appelée *équation caractéristique* de (1.6).

Proposition 1.3.1 Soit $p, q \in \mathbb{R}$, et $P(r) = r^2 + pr + q$. On note $\Delta = p^2 - 4q$ le discriminant du polynôme du second degré P . Soit \mathcal{S} l'ensemble des solutions sur \mathbb{R} à valeurs réelles de l'équation (1.6).

i) Si $\Delta > 0$, notant $r_1, r_2 \in \mathbb{R}$ les racines de $P(r)$,

$$\mathcal{S} = \{t \mapsto y_{c_1, c_2}(t), y_{c_1, c_2}(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}, c_1, c_2 \in \mathbb{R}\},$$

ii) Si $\Delta = 0$, notant $r_0 \in \mathbb{R}$ la racine de $P(r)$,

$$\mathcal{S} = \{t \mapsto y_{c_1, c_2}(t), y_{c_1, c_2}(t) = e^{r_0 t}(c_1 t + c_2), c_1, c_2 \in \mathbb{R}\},$$

iii) Si $\Delta < 0$, notant $\delta \in \mathbb{R}^+$ un réel tel que $\delta^2 = -\Delta$,

$$\mathcal{S} = \{t \mapsto y_{c_1, c_2}(t), y_{c_1, c_2}(t) = e^{-pt/2}(c_1 \cos(\delta t/2) + c_2 \sin(\delta t/2)), c_1, c_2 \in \mathbb{R}\},$$

Preuve: — On commence par le cas où $\Delta = 0$. Soit $y(t)$ une solution de (1.6). On pose $z(t) = e^{-r_0 t} y(t) = e^{pt/2} y(t)$. On a

$$\begin{aligned} z'(t) &= \frac{p}{2} e^{pt/2} y(t) + e^{pt/2} y'(t) \\ z''(t) &= \frac{p^2}{4} e^{pt/2} y(t) + p e^{pt/2} y'(t) + e^{pt/2} y''(t) = e^{pt/2} (qy + py' + y'') = 0. \end{aligned}$$

Donc $z(t) = c_1 t + c_2$ pour certains réels c_1, c_2 , et $y(t)$ est de la forme annoncée. Réciproquement, toutes les fonctions de cette forme sont des solutions de (1.6), ce qui prouve la proposition dans ce cas.

— On considère maintenant le cas où $\Delta > 0$. Soit $y(t)$ une solution de (1.6). Notant $r_1 =$

$\frac{-p-\sqrt{\Delta}}{2}$ la plus petite racine de P , on pose $z(t) = e^{-r_1 t} y(t)$. On a

$$\begin{aligned} z'(t) &= r_1 e^{-r_1 t} y(t) + e^{-r_1 t} y'(t) \\ z''(t) &= (r_1)^2 e^{-r_1 t} y(t) - 2r_1 e^{-r_1 t} y'(t) + e^{-r_1 t} y''(t) \\ &= [(r_1)^2 - q] e^{-r_1 t} y(t) - [2r_1 + p] e^{-r_1 t} y'(t), \end{aligned}$$

où on a utilisé le fait que $y'' = -py' - qy$. On a $2r_1 + p = -\sqrt{\Delta}$ et

$$(r_1)^2 - q = r_1^2 + (r_1^2 + pr_1) = r_1(2r_1 + p) = -r_1 \sqrt{\Delta}.$$

Donc

$$z''(t) = \sqrt{\Delta} [-r_1 e^{-r_1 t} y(t) + e^{-r_1 t} y'(t)] = \sqrt{\Delta} z'(t).$$

Posant $v = z'$, on a $v' = \sqrt{\Delta} v$, donc $v(t) = A e^{\sqrt{\Delta} t}$, et on en déduit que

$$z(t) = c_2 e^{\sqrt{\Delta} t} + c_1,$$

pour certains réels c_1, c_2 . Du coup

$$y(t) = c_2 e^{\sqrt{\Delta} t} e^{r_1 t} + c_1 e^{r_1 t} = c_1 e^{r_1 t} + c_2 e^{r_2 t},$$

comme annoncé. Réciproquement, toutes les fonctions de cette forme vérifient (1.6), et on a terminé la preuve de la proposition dans le cas $\Delta > 0$.

— Dans le cas $\Delta < 0$ on commence par chercher les solutions à valeurs dans \mathbb{C} . Soit $\delta > 0$ tel que $\delta^2 = -\Delta$. On note $r_1 = \frac{-p-i\delta}{2}$ et $r_2 = \frac{-p+i\delta}{2}$ les racines complexes de P . Comme dans le cas où $\Delta > 0$, on pose $z(t) = e^{-r_1 t} y(t)$, et on obtient

$$z''(t) = i\delta z'(t).$$

On en déduit de la même manière que $z(t) = z_2 e^{i\delta t} + z_1$ pour certaines constantes $z_1, z_2 \in \mathbb{C}$, et

$$y(t) = (z_2 e^{i\delta t} + z_1) e^{r_1 t} = z_2 e^{r_2 t} + z_1 e^{r_1 t}.$$

On peut vérifier facilement que toutes les fonctions de cette forme vérifient l'équation (1.6). Autrement dit, on a établi que l'ensemble des solutions de (1.6) à valeurs dans \mathbb{C} est

$$\mathcal{S}_{\mathbb{C}} = \{(\mathbb{R}, y_{z_1, z_2}), y_{z_1, z_2}(t) = z_1 e^{r_1 t} + z_2 e^{r_2 t}, z_1, z_2 \in \mathbb{C}\}.$$

On cherche maintenant les fonctions $y_{z_1, z_2} \in \mathcal{S}_{\mathbb{C}}$ qui sont à valeurs réelles. Notant $z_1 = a_1 + ib_1$ et $z_2 = a_2 + ib_2$, on a

$$y_{z_1, z_2}(t) = e^{-pt/2} [(a_1 + ib_1)(\cos(\delta t/2) - i \sin(\delta t/2)) + (a_2 + ib_2)(\cos(\delta t/2) + i \sin(\delta t/2))],$$

et donc

$$\operatorname{Im} y_{z_1, z_2}(t) = e^{-pt/2} [(a_2 - a_1) \sin(\delta t/2) + (b_1 + b_2) \cos(\delta t/2)].$$

Pour que ce nombre soit nul pour tout $t \in \mathbb{R}$, il faut et il suffit que $a_2 = a_1$ et que $b_2 = -b_1$, autrement dit que $z_2 = \overline{z_1}$ et dans ce cas,

$$y_{z_1, z_2}(t) = e^{-pt/2} (2a_1 \cos(\delta t/2) + 2b_1 \sin(\delta t/2)),$$

ce qui est bien la forme annoncée (prendre $c_1 = 2a_1$ et $c_2 = 2a_2$). \square

1.4 Une équation non-linéaire

On reprend le modèle d'évolution du nombre de bactéries. L'équation $N'(t) = kN(t)$ conduit à $N(t) = Ce^{kt}$, donc à une population dont le nombre croît indéfiniment. Ce n'est évidemment pas très satisfaisant : les bactéries ont besoin de se nourrir, et la nourriture disponible est toujours en quantité limitée. Autrement dit, le nombre de bactéries doit se stabiliser (au moins) quand il devient grand. On est conduit à imaginer que ce nombre vérifie en fait l'équation différentielle

$$(1.7) \quad y' = ky - y^2.$$

Le terme en y^2 est négligeable devant ky pour y petit -donc y devrait croître exponentiellement tant qu'elle reste assez petite, mais devient prépondérant pour y grand - dans ce cas, la population devrait avoir tendance à diminuer.

1.4.1 Des solutions de l'équation $y' = -y^2$

On reviendra sur la résolution de l'équation (1.7) en général, mais on se concentre pour l'instant sur le cas $k = 0$. Autrement dit, on veut résoudre l'équation

$$(1.8) \quad y' = -y^2$$

N'ayant pas grand chose dans notre boîte à outil, on cherche parmi les fonctions que l'on connaît celles qui vérifient cette relation. On se souvient que, si $a \in \mathbb{R}$, et pour $t \neq a$, $(\frac{1}{t-a})' = -\frac{1}{(t-a)^2}$. Autrement dit, notant

$$\left\{ \begin{array}{ccc} y_a^+ :]a, +\infty[& \rightarrow & \mathbb{R} \\ t & \mapsto & \frac{1}{t-a} \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{ccc} y_a^- :]-\infty, a[& \rightarrow & \mathbb{R} \\ t & \mapsto & \frac{1}{t-a} \end{array} \right.$$

on voit que pour tout $a \in \mathbb{R}$, y_a^+ est une solution de (1.8) sur $]a, +\infty[$, et que y_a^- est une solution de (1.8) sur $] -\infty, a[$. Chose un peu étrange, ces fonctions ne peuvent pas être prolongées en des fonctions \mathcal{C}^1 sur \mathbb{R} , donc en des solutions de (1.8) sur \mathbb{R} , alors que l'équation a un sens pour tout t dans \mathbb{R} (les coefficients ne dépendent même pas de t). Pour ajouter à la confusion, on pourra remarquer aussi que la fonction nulle est elle solution de (1.8) sur \mathbb{R} .

Pour clarifier les choses, on est donc contraint de préciser ce qu'on entend par équation différentielle, et de modifier un peu le sens que l'on a donné à la notion de solution.

1.4.2 Retour sur la notion de solution

Définition 1.4.1 Une équation différentielle scalaire d'ordre 1 s'écrit

$$(1.9) \quad y' = f(t, y),$$

où f est une fonction continue sur $I \times U$ à valeurs dans \mathbb{R} , I et U étant des intervalles ouverts de \mathbb{R} .

On dit que le couple (J, y) , constitué d'un intervalle $J \subset I$ et d'une fonction $y : J \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , est une solution de l'équation lorsque

- i) pour tout $t \in J$, on a $y(t) \in U$.
- ii) pour tout $t \in J$, on a $y'(t) = f(t, y(t))$.

On est bien sûr obligé de vérifier la condition (i) pour que la condition (ii) ait un sens.

Exemple 1.4.2 i) L'équation (1.3) : $y' = a(t)y + b(t)$ s'écrit sous la forme (1.9) avec $f(t, x) = a(t)x + b(t)$ définie et continue sur $I \times \mathbb{R}$ quand a et b sont continues sur I .

Les solutions que nous avons mises en évidence s'écrivent alors (I, y_c) , avec $y_c(t) = \dots$

ii) L'équation (1.4) : $u(t)y' + v(t)y = f(t)$ est une équation non-résolue, qui ne s'écrit pas automatiquement sous la forme (1.9). On peut s'y ramener sur chaque intervalle où u ne s'annule pas.

iii) L'équation (1.5) : $y'' + p(t)y' + q(t)y = f(t)$ est une équation du second ordre, donc ne s'écrit pas sous la forme (1.9).

iv) L'équation (1.7) : $y' = ky - y^2$ s'écrit sous la forme (1.9) avec $f(t, x) = kx - x^2$, qui est une fonction continue (et même \mathcal{C}^∞) sur $\mathbb{R} \times \mathbb{R}$, et qui ne dépend pas de t .

Dans le cas particulier $k = 0$, on a trouvé trois types de solutions : $(\mathbb{R}, 0)$, $(]-\infty, a[, y_a^-)$ et $(]a, +\infty[, y_a^+)$.

Pour distinguer les différents cas que l'on vient de rencontrer, on utilise la

Définition 1.4.3 On dit que (J, y) est une solution globale de l'équation différentielle (1.9) lorsque $J = I$.

1.5 Solutions maximales

Supposons que l'on ait déterminé une solution (J, y) de l'équation différentielle (1.9) par un procédé quelconque, et que ce ne soit pas une solution globale ($J \neq I$). Il peut être possible de trouver un intervalle J' plus grand que J sur lequel la fonction y , ou plus exactement son prolongement, est encore solution de (1.9)...

Considérons par exemple l'équation différentielle scalaire d'ordre 1

$$(1.10) \quad y' = f(t, y) = \sqrt{y}$$

avec $f : \mathbb{R} \times [0, +\infty[\rightarrow \mathbb{R}$.

- $(\mathbb{R}, 0)$ est une solution globale de (1.10) : il n'est pas question de vouloir la prolonger.
- Pour $c \in \mathbb{R}$, je note y_c la fonction définie par $y_c(t) = (t - c)^2/4$ pour $t \geq c$. On vérifie facilement que $]c, +\infty[, y_c)$ est une solution de (1.10). On voit aussi que la fonction y_c^- définie par $y_c^-(t) = (t - c)^2/4$ pour $t < c$ n'est pas solution de l'équation sur $] - \infty, c[$. Par contre, considérons la fonction \tilde{y}_c définie sur \mathbb{R} par

$$\tilde{y}_c(t) = \begin{cases} (t - c)^2/4 & \text{pour } t \geq c, \\ 0 & \text{pour } t < c. \end{cases}$$

Cette fonction est \mathcal{C}^1 sur \mathbb{R} (il faut vérifier que \tilde{y}_c est continue en 0, dérivable en 0 et que sa dérivée est continue en 0). De plus $(\mathbb{R}, \tilde{y}_c)$ est une solution de (1.10). Puisque \tilde{y}_c coïncide avec y_c là où cette dernière est définie, on dit que la solution $(\mathbb{R}, \tilde{y}_c)$ prolonge $]c, +\infty[, y_c)$. De manière générale :

Définition 1.5.1 Soient (J_1, y_1) et (J_2, y_2) deux solutions de l'équation différentielle (1.9). On dit que (J_2, y_2) prolonge (où est un prolongement de) (J_1, y_1) lorsque J_2 contient J_1 , et y_2 coïncide avec y_1 sur J_1 :

$$\forall t \in J_1, y_2(t) = y_1(t).$$

On dit que (I, y) est une solution maximale de (1.9) lorsqu'elle n'admet pas d'autre prolongement qu'elle-même.

Bien entendu, toute solution globale est maximale : que peut-on vouloir de plus qu'une solution définie sur tout l'intervalle où l'équation à un sens ? La réciproque est fausse : il y a des solutions maximales qui ne sont pas globales.

Dans le cas de l'équation $y' = -y^2$, on a vu que la solution $]a, +\infty[, t \mapsto \frac{1}{t-a}$ ne peut pas être prolongée : s'il existait un prolongement strict (\tilde{J}, \tilde{y}) , on aurait nécessairement $a \in \tilde{J}$, et \tilde{y} continue (\mathcal{C}^1 même) en a . Or

$$\lim_{t \rightarrow a^+} \tilde{y}(t) = \lim_{t \rightarrow a^+} y(t) = -\infty,$$

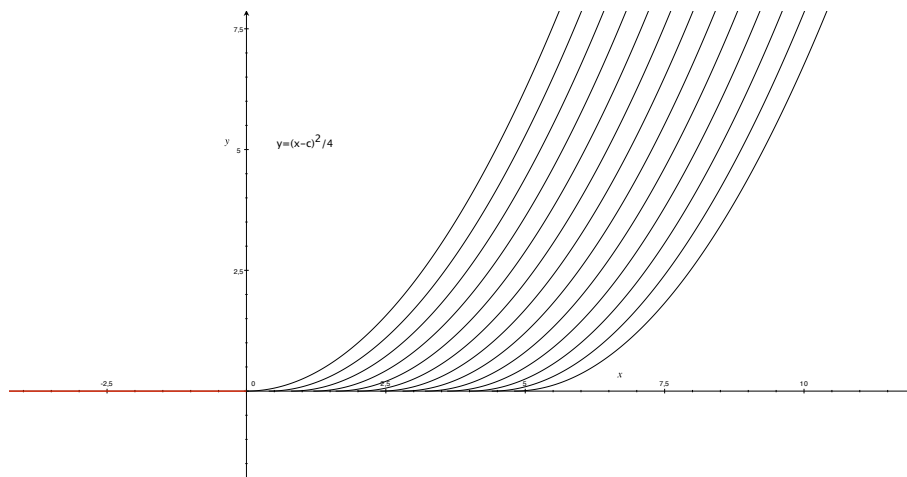
ce qui est absurde. Donc il s'agit d'une solution maximale, bien qu'elle ne soit pas globale.

Lorsqu'une solution n'est pas maximale, elle peut avoir plusieurs prolongements maximaux distincts. Reprenons l'équation différentielle (1.10) : $y' = \sqrt{y}$ sur $\mathbb{R} \times [0, +\infty[$. $(\mathbb{R}, 0)$ est une solution globale (donc maximale), de même que $(\mathbb{R}, \tilde{y}_c)$. En particulier, la solution $] - \infty, 0[, 0)$ admet comme prolongements maximaux $(\mathbb{R}, 0)$ et (\mathbb{R}, y_c) pour tout $c \geq 0$.

1.6 Problème de Cauchy

1.6.1 Mettons un peu d'ordre

On vient de le voir, une équation différentielle a en général beaucoup de solutions, même si l'on ne s'intéresse qu'aux solutions maximales. D'un autre côté, on peut raisonnablement penser

FIGURE 1.2 – Plusieurs prolongements maximaux de la solution $(]-\infty, 0], 0)$

que lorsqu'une équation différentielle décrit l'évolution au cours du temps d'une quantité physique, celle-ci est entièrement déterminée par les conditions initiales. Par exemple, l'équation différentielle qui décrit l'évolution du nombre de bactéries dans une boîte de culture doit avoir une seule solution, dès lors qu'on connaît la population initiale de bactéries dans la boîte. Ceci conduit à la notion de problème de Cauchy pour l'équation différentielle (1.9), pour lequel on espère avoir une solution unique.

Définition 1.6.1 Soit f une fonction continue de $I \times U$ dans \mathbb{R} , où I et U sont des intervalles ouverts de \mathbb{R} . Soit $t_0 \in I$ et $y_0 \in \mathbb{R}^n$. Résoudre le problème de Cauchy (en t_0)

$$(1.11) \quad \begin{cases} y' = f(t, y), \\ y(t_0) = y_0, \end{cases}$$

c'est trouver toutes les solutions maximales (J, y) de l'équation différentielle $y' = f(t, y)$ telles que $t_0 \in J$ et $y(t_0) = y_0$.

Exemple 1.6.2 Dans le cas d'une équation linéaire scalaire du premier ordre, le problème de Cauchy s'écrit donc, pour $a, b \in \mathcal{C}^0(I)$,

$$\begin{cases} y' = a(t)y + b(t), \\ y(t_0) = y_0, \end{cases}$$

On connaît toutes les solutions maximales de l'équation. Ce problème est donc équivalent à la question de trouver parmi les fonctions

$$y_c : t \mapsto e^{A(t)} \left(c + \int_{\alpha}^t e^{-A(s)} b(s) ds \right).$$

celles qui vérifient $y(t_0) = y_0$. Il y en a une, et une seule ! C'est y_{c_0} où

$$e^{A(t_0)}(c_0 + \int_{\alpha}^{t_0} e^{-A(s)}b(s)ds) = y_0,$$

c'est-à-dire

$$c_0 = y_0 e^{-A(t_0)} - \int_{\alpha}^{t_0} e^{-A(s)}b(s)ds.$$

En pratique, on a intérêt à remplacer α par t_0 dans l'expression générale de la solution, et on a, plus simplement

$$c_0 = y_0 e^{-A(t_0)}.$$

Mieux, si on choisit $A(t) = \int_{t_0}^t a(s)ds$ comme primitive de a , on a simplement

$$c_0 = y_0.$$

L'exemple suivant montre que néanmoins, la situation n'est pas tout à fait aussi simple qu'on pourrait le souhaiter : il arrive qu'un problème de Cauchy ait plusieurs solutions.

Exemple 1.6.3 La fonction nulle est une solution sur \mathbb{R} de l'équation différentielle $y' = 3|y|^{2/3}$. La fonction $x \mapsto x^3$ est également une solution sur \mathbb{R} , et le problème de Cauchy

$$(1.12) \quad \begin{cases} y' = 3|y|^{2/3}, \\ y(0) = 0, \end{cases}$$

admet donc au moins deux solutions.

1.6.2 Le théorème de Cauchy-Lipschitz. Première visite

On verra plus tard cependant que sous des hypothèses relativement faibles sur la fonction f , il existe un intervalle ouvert J inclus dans I et contenant t_0 , tel que le problème (1.11) admette une et une seule solution dans J . Ce célèbre résultat porte le nom de Théorème de Cauchy-Lipschitz. C'est une des pierres angulaires de la théorie des équations différentielles, et nous passerons du temps à le démontrer dans toute sa généralité, et à décortiquer ses corollaires.

Pour pouvoir travailler confortablement, on admet provisoirement la

Proposition 1.6.4 (Théorème de Cauchy-Lipschitz - première version) Soit $f : I \times U \rightarrow \mathbb{R}$ une fonction continue, où I et U sont des intervalles ouverts de \mathbb{R} . Soit aussi $t_0 \in I$ et $y_0 \in U$. Si f est de classe \mathcal{C}^1 sur $I \times U$, alors le problème de Cauchy (1.11) admet une unique solution maximale (J, y) .

Il faut noter que cet énoncé n'est certainement pas le meilleur possible. Il donne une condition suffisante d'existence et d'unicité, mais on verra que l'on peut l'affaiblir. D'ailleurs dans le cas des équations linéaires, on a vu que le problème de Cauchy admet une solution unique, même si la fonction $f(t, x) = a(t)x + b(t)$ est supposée seulement continue par rapport à t .

Exercice 1.6.5 Pourquoi le Théorème de Cauchy-Lipschitz ne s'applique-t-il pas dans le cas de l'équation (1.10) ? et dans le cas de l'équation (1.12) ?

1.6.3 Exemple : résolution de l'équation $y' = -y^2$

Pour illustrer l'une des façons d'utiliser ce théorème, on revient sur l'équation (1.8) : $y' = -y^2$, que l'on va résoudre.

On commence par une remarque simple mais extrêmement importante. La fonction $f(t, x) = -x^2$ est \mathcal{C}^1 sur $\mathbb{R} \times \mathbb{R}$. Autrement dit le théorème de Cauchy-Lipschitz s'applique pour le problème de Cauchy

$$\begin{cases} y' = y^2, \\ y(t_0) = y_0, \end{cases}$$

et ce pour tout $(t_0, y_0) \in \mathbb{R}^2$. En particulier, pour tout t_0 , il existe une unique solution maximale pour le problème $y' = y^2, y(t_0) = 0$. On la connaît : c'est $(\mathbb{R}, 0)$.

Supposons alors que (J, y) soit une solution de l'équation (1.8). On a l'alternative suivante :

- ou bien y est la fonction nulle (et $J = \mathbb{R}$),
- ou bien y ne s'annule pas sur J .

En effet si y s'annule sur J , par exemple en t_0 , alors y est la fonction nulle en raison de ce qui précède.

Soit donc (J, y) une solution maximale qui n'est pas la solution nulle. Pour tout $t \in J$, on a

$$y'(t) = -y^2(t) \iff -\frac{y'(t)}{y^2(t)} = 1 \iff \exists C \in \mathbb{R}, \frac{1}{y(t)} = t + C \iff \exists C \in \mathbb{R}, y(t) = \frac{1}{t + C}$$

On a utilisé le fait que y ne s'annule pas sur J pour diviser l'équation par $y(t)^2$, puis on a intégré l'équation obtenue.

Donc, nécessairement, pour tout $t \in J$, $y(t) = 1/(t + C)$ où C est un réel. Comme y est \mathcal{C}^1 sur J , J est un intervalle qui ne contient pas $-C$. Puisque (J, y) est maximale, on a forcément $J =]-\infty, -C[$ ou $J =]-C, +\infty[$. On a déjà vu que $(]-\infty, a[, y_-^a)$ et $(]a, +\infty[, y_+^a)$ sont des solutions de (1.8). On a donc montré que

L'ensemble des solutions maximales de l'équation $y' = y^2$ sur \mathbb{R} est

$$\mathcal{S} = \{(\mathbb{R}, 0), (]-\infty, a[, y_-^a), (]a, +\infty[, y_+^a), a \in \mathbb{R}\}.$$

1.6.4 Exemple : résolution de $y' = ky - y^2$

On résout maintenant l'équation (1.7) : $y' = ky - y^2$, avec $k > 0$. Là encore, la fonction $f(t, x) = kx - x^2$ est \mathcal{C}^1 sur \mathbb{R}^2 , donc le théorème de Cauchy-Lipschitz s'applique. En s'inspirant de ce qui précède, on cherche tout d'abord s'il y a des solutions constantes (la solution nulle

pour $y' = -y^2$). La fonction constante $y(t) = a$ vérifie l'équation sur un intervalle J si et seulement si

$$ka - a^2 = 0 \iff a = 0 \text{ ou } a = k.$$

On a donc deux solutions constantes, forcément globales, $(\mathbb{R}, 0)$ et (\mathbb{R}, k) .

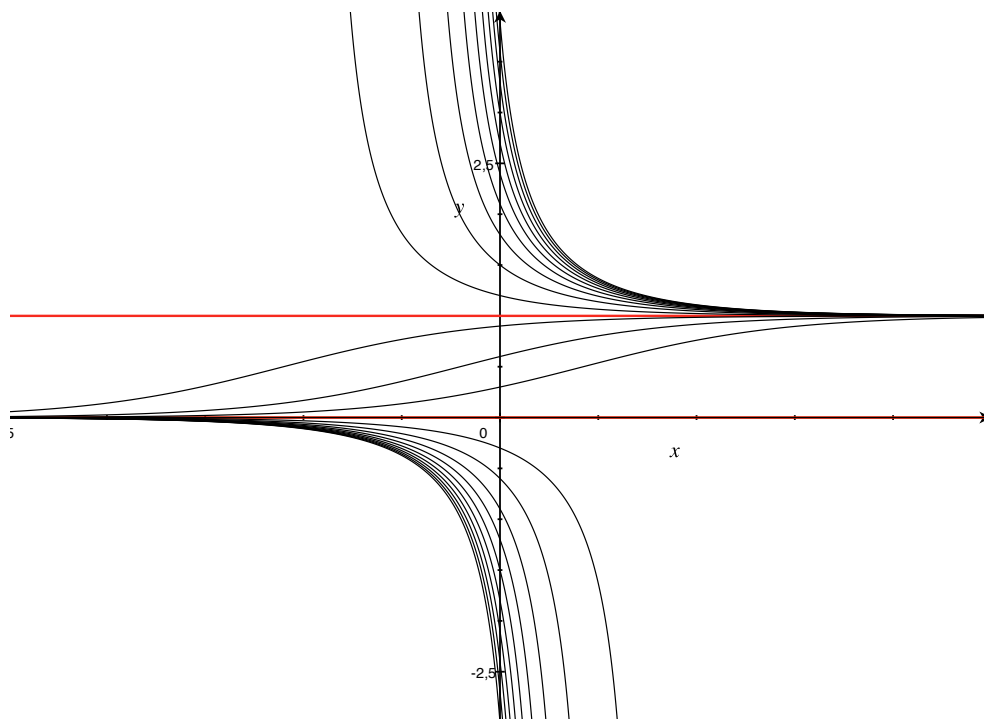


FIGURE 1.3 – Des solutions de l'équation $y' = ky - y^2$, avec $k > 0$

Soit maintenant (J, y) une autre solution. S'il existe $t_0 \in J$ tel que $y(t_0) = 0$ ou k , alors y est solution sur J du problème de Cauchy

$$\begin{cases} y' = ky - y^2, \\ y(0) = 0 \text{ ou } k. \end{cases}$$

Par unicité de Cauchy, y serait l'une des deux solutions constantes. C'est absurde, donc $y(t) \neq 0$ pour tout $t \in J$ et $y(t) \neq k$ pour tout $t \in J$. On a mieux : puisque la fonction y est continue, on a

- (a) ou bien $y(t) < 0$ pour tout $t \in J$,
- (b) ou bien $y(t) \in]0, k[$ pour tout $t \in J$,
- (c) ou bien $y(t) > k$ pour tout $t \in J$.

Graphiquement, ce phénomène est particulièrement parlant. Il s'agit bien sûr d'un principe général :

Lorsque le théorème de Cauchy-Lipschitz s'applique, les courbes représentatives de deux solutions distinctes d'une équation différentielle scalaire d'ordre 1 ne se coupent pas.

Exercice 1.6.6 Le démontrer.

Ce fait a aussi un intérêt pour le calcul des solutions. Si (J, y) est une solution de (1.7) qui n'est pas constante, alors, pour tout $t \in J$,

$$(k - y(t))y(t) = ky(t) - y^2(t) \neq 0.$$

Ainsi (J, y) est une solution de (1.7) si et seulement si

$$\forall t \in J, y'(t) = ky(t) - y^2(t) \iff \forall t \in J, \frac{y'(t)}{k - y(t)} + \frac{y'(t)}{y(t)} = k,$$

et en intégrant

$$\exists C \in \mathbb{R}, \forall t \in J, \ln \frac{|y(t)|}{|k - y(t)|} = kt + C,$$

ou encore

$$(1.13) \quad \exists C > 0, \forall t \in J, \frac{|y(t)|}{|k - y(t)|} = Ce^{kt}.$$

Pour conclure, il faut maintenant discuter suivant les trois cas (a), (b) et (c). Plaçons nous dans le cas (a). On a $y(t) < 0$ et $k - y(t) > 0$ pour tout $t \in J$, donc (1.13) équivaut à

$$\exists C > 0, \forall t \in J, \frac{y(t)}{y(t) - k} = Ce^{kt} \iff \exists C \in \mathbb{R}, \forall t \in J, y(t) = k \frac{Ce^{kt}}{Ce^{kt} - 1}.$$

On notera de plus que l'on doit avoir $y(t) < 0$ pour tout $t \in J$, donc $Ce^{kt} < 1$ pour tout $t \in J$, c'est-à-dire $t < -\frac{1}{k} \ln C$. Autrement dit, si (J, y) est une solution maximale de (1.7) telle que $y(t) < 0$ pour tout $t \in J$, il existe $C > 0$ tel que

$$J =] - \infty, -\frac{1}{k} \ln C[\text{ et } y(t) = k \frac{Ce^{kt}}{Ce^{kt} - 1}.$$

On laisse au lecteur le soin de considérer les deux autres cas (b) et (c), et de confronter ses résultats avec la Figure 1.3.

1.7 Equations différentielles d'ordre supérieur

1.7.1 Qu'est-ce qu'une équation différentielle d'ordre n ?

Jusqu'à présent, on s'est préoccupé essentiellement d'équations différentielles scalaires d'ordre 1, à l'exception des équations différentielles linéaires scalaires d'ordre 2 à coefficients constants. De manière générale, voilà ce que l'on nomme équation différentielle scalaire d'ordre n .

Définition 1.7.1 Une équation différentielle d'ordre n s'écrit

$$(1.14) \quad y^{(n)} = F(x, y, y', y'', \dots, y^{(n-1)})$$

où F est une fonction continue sur $I \times U$ à valeurs dans \mathbb{R}^n , I étant un intervalle ouvert de \mathbb{R} , et U un ouvert de \mathbb{R}^n .

On dit que le couple (J, y) , constitué d'un intervalle $J \subset I$ et d'une fonction $y : J \rightarrow \mathbb{R}$ de classe \mathcal{C}^n , est une solution de l'équation (1.14) lorsque

- i) pour tout $t \in J$, on a $(y(t), y'(t), y''(t), \dots, y^{(n-1)}(t)) \in U$.
- ii) pour tout $t \in J$, on a $y^{(n)}(t) = F(y(t), y'(t), y''(t), \dots, y^{(n-1)}(t))$.

Exercice 1.7.2 Donner la fonction F , I et U dans les exemples de l'Introduction.

1.7.2 Se ramener à un système de n équations différentielles d'ordre 1

Pour énoncer des résultats théoriques, mais aussi parfois pour résoudre les équations d'ordre n , on préfère ramener une équation différentielle d'ordre n à un système de n équations différentielles d'ordre 1.

Commençons par l'exemple de l'équation (1.5) : $y'' + p(t)y' + q(t)y = f(t)$. Supposons que (J, y) en soit une solution, et posons

$$Y : J \rightarrow \mathbb{R}^2, Y(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}.$$

Puisque y est nécessairement \mathcal{C}^2 , la fonction Y est \mathcal{C}^1 sur J et

$$Y'(t) = \begin{pmatrix} y'(t) \\ y''(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -p(t) & -q(t) \end{pmatrix} Y(t) + \begin{pmatrix} 0 \\ f(t) \end{pmatrix}$$

Autrement dit, (J, Y) est solution du système différentiel

$$Y'(t) = G(t, Y(t))$$

où $G : I \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est la fonction définie par

$$G(t, X) = \begin{pmatrix} 0 & 1 \\ -p(t) & -q(t) \end{pmatrix} X + \begin{pmatrix} 0 \\ f(t) \end{pmatrix} = \begin{pmatrix} x_2 \\ -p(t)x_1 - q(t)x_2 + f(t) \end{pmatrix},$$

avec $X = (x_1, x_2) \in \mathbb{R}^2$. De manière générale

Proposition 1.7.3 Soit $F : I \times U \rightarrow \mathbb{R}$ une fonction continue, où I est un intervalle ouvert de \mathbb{R} et U un ouvert de \mathbb{R}^n . On note $G : I \times U \rightarrow \mathbb{R}^n$ la fonction définie par

$$G(t, x_1, \dots, x_n) = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ F(t, x_1, \dots, x_n) \end{pmatrix}.$$

i) Si (J, y) est une solution de l'équation différentielle d'ordre n

$$(1.15) \quad y^{(n)} = F(x, y, y', y'', \dots, y^{(n-1)}),$$

alors (J, Y) , avec $Y(t) = \begin{pmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(n)}(t) \end{pmatrix}$, est une solution du système différentiel d'ordre 1

$$(1.16) \quad Y' = G(t, Y),$$

ii) Réciproquement, si (J, Y) est une solution de (1.16), avec $Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}$, alors (J, y_1) est une solution de (1.15).

La preuve de cette proposition est facile, et on laisse au lecteur le soin de la rédiger. On remarquera en particulier que si Y est \mathcal{C}^1 et vérifie le système (1.16), sa première composante y_1 est bien une fonction \mathcal{C}^n .

1.7.3 Au fait : c'est quoi un système différentiel d'ordre 1 ?

On formalise maintenant la notion de système différentiel d'ordre 1. La différence avec les équations scalaires d'ordre 1, c'est que l'inconnue est une fonction vectorielle $Y : I \rightarrow \mathbb{R}^n$, ou, si l'on préfère, un n -uplet de fonctions (y_1, \dots, y_n) (les composantes de la fonction Y).

Définition 1.7.4 Un système différentiel d'ordre 1 s'écrit

$$(1.17) \quad Y' = F(t, Y),$$

où F est une fonction continue sur $I \times U$ à valeurs dans \mathbb{R}^n , I étant un intervalle ouvert de \mathbb{R} , et U un ouvert de \mathbb{R}^n .

On dit que le couple (J, Y) , constitué d'un intervalle $J \subset I$ et d'une fonction $Y : J \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 , est une solution de l'équation lorsque

- i) pour tout $t \in J$, on a $Y(t) \in U$.
- ii) pour tout $t \in J$, on a $Y'(t) = F(t, Y(t))$.

(1.17) est bien d'un système d'équations. En effet, notant

$$Y(t) = (y_1(t), \dots, y_n(t))$$

les n coordonnées de la fonction $Y : J \rightarrow \mathbb{R}^n$, et

$$F(t, x) = (F_1(t, x), \dots, F_n(t, x))$$

les n coordonnées de la fonction $F : I \times U \rightarrow \mathbb{R}^n$, l'équation (1.17) s'écrit plus explicitement

$$(1.18) \quad \begin{cases} y_1'(t) = F_1(t, y_1(t), \dots, y_n(t)), \\ y_2'(t) = F_2(t, y_1(t), \dots, y_n(t)), \\ \vdots = \vdots \\ y_n'(t) = F_n(t, y_1(t), \dots, y_n(t)). \end{cases}$$

Par exemple, le système prédateurs-proies de Lotka et Volterra

$$\begin{cases} y_1'(t) = ay_2(t) - by_1(t)y_2(t), \\ y_2'(t) = -cy_1(t) + dy_1(t)y_2(t), \end{cases}$$

s'écrit $Y' = F(t, Y)$ avec

$$F : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, F(t, x_1, x_2) = (ax_2 - bx_1x_2, -cx_1 + dx_1x_2).$$

1.7.4 Le théorème de Cauchy-Lipschitz pour les systèmes, et pour les équations d'ordre n . Première visite.

Pour les systèmes d'ordre 1, on démontrera le même théorème de Cauchy-Lipschitz que pour les équations scalaires :

Proposition 1.7.5 (Théorème de Cauchy-Lipschitz - première version- cas des systèmes) Soit $F : I \times U \rightarrow \mathbb{R}^n$ une fonction continue, où I est un intervalle ouvert de \mathbb{R} et U un ouvert de \mathbb{R}^n . Soit aussi $t_0 \in I$ et $Y_0 \in U$. Si F est de classe \mathcal{C}^1 sur $I \times U$, alors le problème de Cauchy

$$\begin{cases} Y' = F(t, Y) \\ Y(t_0) = Y_0 \end{cases}$$

admet une unique solution maximale (J, Y) .

En rapprochant la Proposition 1.7.3 et ce qui précède, on comprend que la bonne notion de problème de Cauchy pour une équation différentielle d'ordre n est la suivante :

Définition 1.7.6 Soit F une fonction de $\mathbb{R} \times \mathbb{R}^n$ dans \mathbb{R} , définie et continue sur $I \times U$, où I est un intervalle ouvert de \mathbb{R} , et U un ouvert de \mathbb{R}^n . Soit $t_0 \in I$ et $(y_0, y_1, \dots, y_{n-1}) \in \mathbb{R}^n$. Résoudre le problème de Cauchy (en t_0)

$$(1.19) \quad \begin{cases} y' = F(t, y) \\ y(t_0) = y_0, y'(t_0) = y_1, \dots, y^{(n-1)}(t_0) = y_{n-1}, \end{cases}$$

c'est trouver toutes les solutions (maximales) (J, y) de l'équation différentielle $y' = F(t, y)$ telles que $t_0 \in J$ et $y(t_0) = y_0, y'(t_0) = y_1, \dots, y^{(n-1)}(t_0) = y_{n-1}$.

Le théorème de Cauchy-Lipschitz (la Proposition 1.7.5), appliqué au système d'ordre 1 correspondant, dit en effet que, lorsque F est \mathcal{C}^1 , ce problème admet une unique solution maximale. Typiquement, ce résultat dit que l'équation de Newton (une équation différentielle d'ordre 2) détermine de manière unique la trajectoire $y(t)$ d'un mobile à condition que l'on spécifie sa position initiale $y(0)$ et sa vitesse initiale $y'(0)$.

1.8 Systèmes linéaires

L'étude des systèmes linéaires va occuper une bonne partie de ce cours. Il y a au moins deux excellentes raisons pour cela. Tout d'abord, on va réussir à donner une expression explicite de toutes les solutions de ces systèmes. On verra ensuite que les propriétés des solutions de systèmes différentiels quelconques sont souvent (et on précisera dans quel cas) proches de celles de systèmes linéaires à coefficients constants bien choisis.

1.8.1 Qu'est-ce qu'un système linéaire ?

Il est temps de préciser pourquoi on parle d'*équations linéaires* pour les équations différentielles de la forme (1.3) : $y' = a(t)y + b(t)$. Pour chaque t fixé, posons $\ell_t(x) = a(t)x + b(t)$. La fonction $\ell_t : \mathbb{R} \rightarrow \mathbb{R}$ est affine. Autrement dit $h_t : x \mapsto \ell_t(x) - \ell_t(0)$ est linéaire :

$$\forall \alpha_1, \alpha_2 \in \mathbb{C}, \forall x, x_2 \in \mathbb{R}, h_t(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 h_t(x_1) + \alpha_2 h_t(x_2).$$

On étend cette notion aux systèmes. Soit $F : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction affine par rapport à sa variable X . Cela signifie que pour tout t fixé, la fonction $H(t, X) = F(t, X) - F(t, 0)$ est linéaire par rapport à X . Autrement dit, il existe une matrice $A(t)$ telle que $H(t, X) = A(t)X$, et finalement, notant $B(t) = F(t, 0)$,

$$F(t, X) = A(t)X + B(t).$$

Définition 1.8.1 Soit $F : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continue, où I est un intervalle ouvert de \mathbb{R} . On dit que le système différentiel

$$(1.20) \quad Y' = F(t, Y)$$

est linéaire lorsqu'il existe deux fonctions $A : I \rightarrow \mathcal{M}_n(\mathbb{R})$ et $B : I \rightarrow \mathbb{R}^n$, forcément continues, telles que

$$F(t, X) = A(t)X + B(t)$$

Pour les connaisseurs, un système différentiel (ou une équation scalaire) est dit(e) linéaire lorsque la fonction $X \mapsto F(t, X)$ est affine pour tout t ...

Pour simplifier les prochains énoncés, nous allons admettre -provisoirement- le résultat suivant (cf. le Corollaire 4.4.3). On l'a d'ailleurs déjà démontré pour les équations linéaires scalaires d'ordre 1.

Proposition 1.8.2 Toutes les solutions maximales du système linéaire (1.20) sont globales.

Parlant d'une solution (J, Y) d'un système linéaire, on ne précisera donc plus l'intervalle J : on ne considère désormais que les solutions maximales, donc $J = I$.

1.8.2 Structure de l'ensemble des solutions

Les résultats qui suivent sont énoncés pour les systèmes linéaires, mais valent bien sûr aussi pour les équations linéaires scalaires. Ce sont des remarques tout à fait simples, mais cruciales.

Le système différentiel

$$(1.21) \quad Y' = A(t)Y$$

est appelée *système différentiel homogène* associée à (1.20).

Proposition 1.8.3 Si Y_1 et Y_2 sont deux solutions du système $Y' = A(t)Y$, alors pour tous $\alpha_1, \alpha_2 \in \mathbb{C}$, la fonction $\alpha_1 Y_1 + \alpha_2 Y_2$ est aussi une solution. Autrement dit, l'ensemble des solutions d'un système linéaire homogène est un sous-espace vectoriel de $\mathcal{C}^1(I, \mathbb{R}^n)$.

Preuve.— Il suffit de remarquer que pour tout $t \in I$, $(\alpha_1 Y_1 + \alpha_2 Y_2)'(t) = \alpha_1 Y_1'(t) + \alpha_2 Y_2'(t)$ et que, notant $H(t, X) = A(t)X$,

$$H(t, \alpha_1 Y_1(t) + \alpha_2 Y_2(t)) = \alpha_1 H(t, Y_1(t)) + \alpha_2 H(t, Y_2(t)).$$

□

Proposition 1.8.4 Si Y_1 et Y_2 sont deux solutions du système linéaire $Y' = F(t, Y)$, alors $Z = Y_1 - Y_2$ est une solution du système linéaire homogène associé.

Preuve.— Pour $t \in I$, on a

$$Z'(t) = Y_1'(t) - Y_2'(t) = F(t, Y_1(t)) - F(t, Y_2(t)) = H(t, Y_1(t)) - H(t, Y_2(t)) = H(t, Z(t)).$$

□

On en déduit le résultat connu sous le nom de principe de superposition :

Proposition 1.8.5 Soit Y_0 une solution de l'équation (1.20). Toute solution Y de (1.20) s'écrit $Y = Y_0 + Z$, où Z est solution de l'équation différentielle homogène associée (1.21).

Autrement dit, pour trouver **TOUTES** les solutions de (1.20), il suffit de

1. trouver toutes les solutions Z de l'équation homogène associée (1.21) ;
2. trouver **UNE** solution Y_0 de (1.20).

Pour les connaisseurs encore, le principe de superposition dit que l'ensemble des solutions d'un système linéaire est un espace affine, dont la direction est l'espace vectoriel des solutions de l'équation homogène associée.

1.8.3 Dimension de l'espace des solutions

On suppose toujours que $A : I \rightarrow \mathcal{M}_n(\mathbb{R})$ est une fonction continue. On a vu que l'ensemble \mathcal{S}_H des solutions du système linéaire homogène

$$(1.22) \quad Y' = A(t)Y$$

est un sous-espace vectoriel de $\mathcal{C}^1(I, \mathbb{R}^n)$. Dans le cas des équations scalaires ($n = 1$), on a vu aussi que

$$\mathcal{S}_H = \{t \mapsto Ce^{A(t)}, C \in \mathbb{R}\},$$

et qu'il s'agit donc d'un espace vectoriel de dimension 1, engendré par la fonction $t \mapsto e^{A(t)}$ par exemple. Grâce au théorème de Cauchy-Lipschitz, on va démontrer la

Proposition 1.8.6 Soit $A : I \rightarrow \mathcal{M}_n(\mathbb{R})$ est une fonction continue. L'espace vectoriel \mathcal{S}_H des solutions du système linéaire homogène $Y' = A(t)Y$ est de dimension n .

Preuve: Soit $t_0 \in I$. D'après la Proposition 1.7.5 (en fait la Proposition 4.3.7, puisque $F(t, X) = A(t)X$ n'est pas supposée de classe \mathcal{C}^1), pour chaque $X_0 \in \mathbb{R}^n$ le problème

$$(1.23) \quad \begin{cases} Y' = A(t)Y, \\ Y(t_0) = X_0, \end{cases}$$

admet une unique solution, que je note (I, Y_{X_0}) . L'application $\phi : X_0 \mapsto Y_{X_0}$ de \mathbb{R}^n dans \mathcal{S}_H est donc bijective :

- elle est injective : si $Y_{X_1} = Y_{X_2}$ on a nécessairement $X_1 = X_2$.
- elle est surjective : si $Y \in \mathcal{S}_H$, on a $Y = \phi_{Y(t_0)}$.

Montrons maintenant que ϕ est linéaire. Pour $X_1, X_2 \in \mathbb{R}^n$ et $\alpha_1, \alpha_2 \in \mathbb{R}$,

$$Y_{\alpha_1 X_1 + \alpha_2 X_2} = \alpha_1 Y_{X_1} + \alpha_2 Y_{X_2}.$$

Puisque $Y_{\alpha_1 X_1 + \alpha_2 X_2}$ est l'unique solution Y du système qui vérifie $Y(t_0) = \alpha_1 X_1 + \alpha_2 X_2$, il suffit de montrer que $\alpha_1 Y_{X_1} + \alpha_2 Y_{X_2}$ est aussi une solution du même problème de Cauchy. Le fait que $\alpha_1 Y_{X_1} + \alpha_2 Y_{X_2}$ est une solution a déjà été démontré : \mathcal{S}_H est un espace vectoriel. Enfin

$$(\alpha_1 Y_{X_1} + \alpha_2 Y_{X_2})(t_0) = \alpha_1 Y_{X_1}(t_0) + \alpha_2 Y_{X_2}(t_0) = \lambda_1 X_1 + \lambda_2 X_2.$$

La bijection $\phi : \mathbb{R} \rightarrow \mathcal{S}_H$ est donc un isomorphisme d'espace vectoriel, et

$$\dim \mathcal{S}_H = \dim \phi(\mathbb{R}^n) = \dim \mathbb{R}^n = n.$$

□

En traduisant ce résultat dans le cas où le système d'ordre 1 que l'on considère provient d'une équation différentielle scalaire d'ordre n , on obtient la

Proposition 1.8.7 L'ensemble des solutions d'une équation différentielle linéaire scalaire d'ordre n résolue à coefficient continus est un sous-espace vectoriel de dimension n de $\mathcal{C}^n(I, \mathbb{R})$.

1.8.4 Méthode de variation de la constante

On a élucidé la structure de l'espace des solutions des solutions de l'équation homogène $Y' = A(t)Y$. Pour décrire l'ensemble des solutions de l'équation complète $Y' = A(t)Y + B(t)$, il reste à en trouver UNE solution.

Voici un procédé bien connu pour les équations linéaires d'ordre 1 : la méthode de variation de la constante. On verra plus loin que cette procédure marche aussi pour les systèmes linéaires.

On considère donc l'équation, avec $a, b : I \rightarrow \mathbb{R}$ continues,

$$(1.24) \quad y' = a(t)y + b(t).$$

On sait que les solutions de l'équation homogène associée

$$y' = a(t)y$$

sont les fonctions de la forme $y(t) = Ce^{A(t)}$, où $A(t)$ est une primitive de $a(t)$ sur I .

L'idée est la suivante : on cherche une solution y de l'équation sous la forme

$$y(t) = c(t)e^{A(t)}$$

où c est une fonction \mathcal{C}^1 à déterminer. De manière un peu rapide, on dit que l'on fait varier la constante c qui apparaît dans l'expression de la solution de l'équation homogène associée à (1.24).

Pour que cette fonction soit une solution, il faut et il suffit que

$$c'(t)e^{A(t)} + a(t)c(t)e^{A(t)} = y'(t) = a(t)y(t) + b(t) = a(t)c(t)e^{A(t)} + b(t),$$

c'est-à-dire

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

et il suffit donc de prendre pour $c(t)$ une primitive de $t \mapsto b(t)e^{-A(t)}$, c'est-à-dire

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

où $t_0 \in I$ est un point quelconque. On en déduit que

$$y(t) = e^{A(t)} \int_{t_0}^t b(s)e^{A(s)} ds$$

est une solution de (1.24), ce qui, compte tenu de la Proposition 1.8.5, fournit une autre preuve de la Proposition 1.2.3.

On décrit maintenant rapidement la méthode de variation de la constante pour les équations linéaires d'ordre 2. Là encore, elle permet, lorsque l'on connaît une base, donc deux solutions indépendantes de l'espace \mathcal{S}_H de l'équation homogène associée, de trouver une solution de l'équation complète.

Supposons que y_1 et y_2 soient deux solutions indépendantes de l'équation

$$(1.25) \quad y'' + p(t)y' + q(t)y = 0$$

On sait que toutes les solutions de cette équation s'écrivent

$$y = \alpha_1 y_1 + \alpha_2 y_2.$$

On cherche deux fonctions $\alpha_1(t)$ et $\alpha_2(t)$ de classe \mathcal{C}^2 , de sorte que la fonction

$$(1.26) \quad y = \alpha_1(t)y_1 + \alpha_2(t)y_2$$

soit solution de

$$y'' + p(t)y' + q(t)y = f(t).$$

On suppose aussi que $\alpha_1(t)$ et $\alpha_2(t)$ vérifient la relation

$$(1.27) \quad \alpha_1'(t)y_1 + \alpha_2'(t)y_2 = 0.$$

Cette seconde hypothèse ne semble pas très naturelle, mais on verra plus loin que l'on peut affirmer à priori que de telles fonctions existent. L'hypothèse (1.27) a pour conséquence de simplifier un peu l'expression de y' et de y'' :

$$\begin{aligned} y'(t) &= \cancel{\alpha_1'(t)y_1 + \alpha_2'(t)y_2} + \alpha_1(t)y_1' + \alpha_2(t)y_2', \\ y''(t) &= \alpha_1'(t)y_1' + \alpha_2'(t)y_2' + \alpha_1(t)y_1'' + \alpha_2(t)y_2''. \end{aligned}$$

On forme alors l'équation différentielle que y doit satisfaire :

$$\begin{aligned} f &= (\alpha'_1(t)y'_1 + \alpha'_2(t)y'_2 + \alpha_1(t)y''_1 + \alpha_2(t)y''_2) + p(\alpha_1(t)y'_1 + \alpha_2(t)y'_2) + q(\alpha_1(t)z_1 + \alpha_2(t)z_2), \\ &= \alpha_1(t)(z''_1 + pz'_1 + qz_1) + \alpha_2(t)(z''_2 + pz'_2 + qz_2) + \alpha'_1(t)y'_1 + \alpha'_2(t)y'_2 \end{aligned}$$

et l'on se souvient que y_1 et y_2 sont des solutions de (1.25). On obtient

$$(1.28) \quad f = \alpha'_1(t)y'_1 + \alpha'_2(t)y'_2.$$

Il reste alors à résoudre le système linéaire (1.27), (1.28) d'inconnues α'_1 et α'_2 , ce qui donne une expression pour ces fonctions. On peut alors prendre pour α_1 et α_2 une primitive quelconque des fonctions obtenues.

Au passage, on voit que l'existence de solutions pour le système (1.27), (1.28) est liée à la non-annulation du déterminant associé : pour que (1.27), (1.28) admette une solution, il faut et il suffit que

$$\forall t \in I, \mathcal{W}(y_1, y_2)(t) := \det \begin{pmatrix} y_1(t) & y_2(t) \\ y'_1(t) & y'_2(t) \end{pmatrix} \neq 0$$

La quantité $\mathcal{W}(y_1, y_2)(t) = y_1(t)y'_2(t) - y'_1(t)y_2(t)$ est appelée Wronskien des fonctions y_1 et y_2 . Nous reverrons cette quantité plus loin.

Chapitre 2

Réduction des matrices

Comme on l'a fait pour les équations linéaires scalaires d'ordre 1, on est capable de donner l'ensemble des solutions des systèmes linéaires à coefficients constants homogènes ou non, ce que l'on fera dans le chapitre suivant. Pour ce faire, il faut d'abord revoir et approfondir quelques notions d'algèbre linéaire, permettant de ramener un système quelconque

$$(2.1) \quad Y' = AY, \text{ avec } A \in \mathcal{M}_n(\mathbb{C}),$$

à une forme plus simple, avec laquelle on saura se débrouiller. L'idée est que si $P \in \mathcal{M}_n(\mathbb{C})$ est une matrice inversible, en posant $Z = P^{-1}Y$, l'équation (2.1) est équivalente à l'équation

$$Z' = BZ, \text{ avec } B = P^{-1}AP.$$

En effet, puisque $Z' = (P^{-1}Y)' = P^{-1}Y'$, on a

$$Z' = BZ \iff P^{-1}Y' = P^{-1}AP.P^{-1}Y \iff Y' = AY.$$

On a donc tout intérêt à trouver une matrice P telle que $P^{-1}AP$ soit la plus simple possible. Cette question est celle de la réduction des matrices, que nous reprenons en détail.

2.1 Matrice d'une application linéaire

2.1.1 Définitions

On note \mathbb{K} l'un des corps \mathbb{R} ou \mathbb{C} . La plupart des définitions et des énoncés ci-dessous sont valables dans ces deux cas. Soit $a \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^n)$ une application linéaire, et $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ une base de \mathbb{K}^n . La matrice de a dans la base \mathcal{B} est le tableau de nombres $A = (a_{i,j})$ à n lignes et n colonnes, dont les colonnes sont les coordonnées des vecteurs $a(e_j)$ dans la base \mathcal{B} . Autrement dit, les $a_{i,j}$ sont entièrement déterminés par la relation

$$a(e_j) = \sum_{i=1}^n a_{i,j} e_i.$$

Réciproquement, une matrice $A \in \mathcal{M}_n(\mathbb{K})$, c'est-à-dire tableau à n lignes et n colonnes d'éléments de \mathbb{K} , peut être vu comme la matrice d'une application linéaire, modulo le choix d'une base de \mathbb{K}^n .

On rappelle aussi que si A et B sont les matrices de deux applications linéaires a et b dans une base \mathcal{B} , la matrice de $a \circ b$ dans la base \mathcal{B} est la matrice produit $C = AB$ définie par

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}.$$

2.1.2 Changement de base

On veut calculer la matrice A' de a dans la base $\mathcal{B}' = \{e'_1, e'_2, \dots, e'_n\}$ de \mathbb{K}^n lorsque l'on connaît la matrice A de a dans la base $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$.

Définition 2.1.1 On appelle matrice de passage de la base \mathcal{B} à la base \mathcal{B}' la matrice $P = (p_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ dont les colonnes sont les coordonnées des vecteurs e'_j dans la base \mathcal{B} :

$$e'_j = \sum_{i=1}^n p_{i,j} e_i$$

Une matrice de passage P est toujours inversible : puisque ses vecteurs colonnes sont indépendants, le rang de P est n , ce qui, en utilisant le théorème du rang, entraîne que P est injective et bijective. Malgré son nom, la matrice de passage de \mathcal{B} dans \mathcal{B}' permet de calculer les coordonnées d'un vecteur dans l'ancienne base \mathcal{B} , connaissant ses coordonnées dans la nouvelle base \mathcal{B}' :

Proposition 2.1.2 Soit $u \in \mathbb{K}^n$ un vecteur, de coordonnées $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ dans la base \mathcal{B} , et de coordonnées $\begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix}$ dans la base \mathcal{B}' . On a $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = P \begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix}$.

Preuve: Il suffit de vérifier la proposition pour les vecteurs de la base \mathcal{B}' . Le résultat pour les autres vecteurs en découlera par linéarité. Les coordonnées du vecteur e'_k dans la base \mathcal{B}' sont

$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, où le 1 figure sur la k -ième ligne. On calcule donc $P \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} p_{1,k} \\ p_{2,k} \\ \vdots \\ p_{n,k} \end{pmatrix}$, ce qui est bien, par définition de P , le vecteur colonne des coordonnées de e'_k dans la base \mathcal{B} . \square

En particulier, cette proposition permet de montrer que P^{-1} est la matrice de passage de la base \mathcal{B}' à la base \mathcal{B} .

Proposition 2.1.3 Soit $a \in \mathcal{L}(\mathbb{K}^n)$, A la matrice de a dans la base \mathcal{B} , et A' sa matrice dans la base \mathcal{B}' . Si P est la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' , on a

$$A' = P^{-1}AP$$

Preuve: Si u a pour coordonnées $\begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix}$ dans la base \mathcal{B}' , on a vu que $X = P \begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix}$ est le vecteur colonne des coordonnées de u dans la base \mathcal{B} . Alors $Y = AX$ est le vecteur colonne des coordonnées de $a(u)$ dans la base \mathcal{B} et finalement $P^{-1}Y$ est le vecteur colonne des coordonnées de $a(u)$ dans la base \mathcal{B}' . \square

Définition 2.1.4 On dit que deux matrices A et B sont semblables sur \mathbb{K} lorsqu'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ telle que $A = PBP^{-1}$.

De manière équivalente, deux matrices sont semblables si et seulement si elles représentent la même application linéaire.

2.1.3 Sous-espaces stables

La réduction d'une matrice passe d'abord par la décomposition de \mathbb{K}^n en une somme directe de sous-espaces stables.

Définition 2.1.5 On dit qu'un sous-espace vectoriel F de \mathbb{K}^n est stable par A lorsque $A(F) \subset F$.

En effet, supposons qu'il existe une famille F_1, F_2, \dots, F_p de sous-espaces vectoriels de \mathbb{K}^n , stables par A et tels que

$$\mathbb{K}^n = F_1 \oplus F_2 \oplus \dots \oplus F_p.$$

Pour chaque $j \in \{1, \dots, p\}$, on note $\mathcal{B}_j = \{e_1^j, e_2^j, \dots, e_{n_j}^j\}$ une base de F_j . On a donc au passage $n_j = \dim F_j$ et $n_1 + \dots + n_p = n$.

Puisque \mathbb{K}^n est la somme directe des F_j , la famille $\mathcal{B} = \cup_{j=1 \dots p} \mathcal{B}_j = \{e_\ell^j, j = 1, \dots, p, \ell = 1 \dots n_j\}$ est une base de \mathbb{K}^n . Soit $e_{\ell_0}^{j_0}$ un vecteur de cette base. Puisque chaque F_j est stable par A , $Ae_{\ell_0}^{j_0} \in F_{j_0}$, et les coordonnées de $Ae_{\ell_0}^{j_0}$ sur les vecteurs e_ℓ^j avec $j \neq j_0$ sont toutes nulles. Autrement dit, la matrice de A dans la base \mathcal{B} est diagonale par blocs :

$$(2.2) \quad A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_p \end{pmatrix},$$

où chaque bloc $A_j \in \mathcal{M}_{n_j}(\mathbb{K})$ est la matrice dans la base \mathcal{B}_j de la restriction de A à F_j .

Dans la suite, on va construire, pour une matrice A donnée, une famille de sous-espaces invariants supplémentaires. Il s'agira ensuite de choisir pour chacun de ces sous-espaces une base qui rend le bloc correspondant le plus simple possible.

2.2 Valeurs propres et vecteurs propres

2.2.1 Définitions

Soit $A \in \mathcal{M}_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}).

Définition 2.2.1 On dit que $\lambda \in \mathbb{C}$ est une valeur propre complexe de A lorsqu'il existe un vecteur non-nul $u \in \mathbb{C}^n$ tel que $Au = \lambda u$. On note $\text{sp}_{\mathbb{C}}(A)$ l'ensemble des valeurs propres complexes de A (c'est le spectre complexe de A).

Pour $\lambda \in \text{sp}(A)$, l'espace vectoriel $E_\lambda = \text{Ker}(A - \lambda I) \subset \mathbb{C}^n$ est le sous-espace propre (complexe) associé à λ . Sa dimension $\mu(\lambda)$ est appelée multiplicité géométrique de λ .

Lorsque $A \in \mathcal{M}_n(\mathbb{R})$, on a une définition équivalente pour le spectre réel de A :

Définition 2.2.2 On dit que $\lambda \in \mathbb{R}$ est une valeur propre réelle de A lorsqu'il existe un vecteur non-nul $u \in \mathbb{R}^n$ tel que $Au = \lambda u$. On note $\text{sp}_{\mathbb{R}}(A)$ l'ensemble des valeurs propres réelles de A (c'est le spectre réel de A).

Pour $\lambda \in \text{sp}_{\mathbb{R}}(A)$, l'espace vectoriel $E_{\lambda} = \text{Ker}(A - \lambda I) \subset \mathbb{R}^n$ est le sous-espace propre (réel) associé à λ . Sa dimension $\mu(\lambda)$ est appelée multiplicité géométrique de λ .

Remarque 2.2.3 Si $A \in \mathcal{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{C} \setminus \mathbb{R}$ une valeur propre complexe de A associée au vecteur propre $u \in \mathbb{C}^n$, alors $\bar{\lambda}$ est une valeur propre de A associée au vecteur propre \bar{u} . On a en effet $Au = \lambda u$ donc $\overline{Au} = \bar{\lambda} \bar{u}$, et puisque $A \in \mathcal{M}_n(\mathbb{R})$, $\overline{Au} = A\bar{u}$.

Les sous-espaces propres sont stables par A : si $u \in E_{\lambda}$, $Au = \lambda u \in E_{\lambda}$. Lorsqu'ils sont supplémentaires, la matrice A est diagonalisable, c'est-à-dire semblable à une matrice diagonale. En effet si

$$(2.3) \quad \text{Ker}(A - \lambda_1 I) \oplus \cdots \oplus \text{Ker}(A - \lambda_p I) = \mathbb{C}^n,$$

la décomposition par blocs (2.2) ci-dessus est

$$A = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \ddots & \\ & & & \lambda_p I \end{pmatrix},$$

puisque la restriction de A à $\text{Ker}(A - \lambda_j I)$ est $\lambda_j I$. Autrement dit, il existe une base dans laquelle l'application linéaire dont A est la matrice est diagonale. Le problème est que (2.3) n'a pas toujours lieu.

Néanmoins, deux sous-espaces propres distincts ont une intersection réduite au vecteur nul : si $Au = \lambda u = \lambda' u$, on a $u = 0$ ou $\lambda = \lambda'$. On a même la

Proposition 2.2.4 Soit $k \in \mathbb{N}^*$. Si u_1, u_2, \dots, u_k sont k vecteurs propres associés chacun à des valeurs propres 2 à 2 distinctes d'une même matrice A , la famille $\{u_1, u_2, \dots, u_k\}$ est libre.

Preuve: Par récurrence sur k .

- Pour $k = 1$, la proposition découle du fait qu'un vecteur propre ne peut pas être le vecteur nul.

- Soit $k \geq 2$. Supposons que la proposition soit vraie pour $k - 1$ vecteurs propres. Soit alors u_1, \dots, u_{k-1} des vecteurs propres associés respectivement aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ de A , deux à deux distinctes. Supposons que $u_k = \sum_{j=1}^{k-1} \alpha_j u_j$. Alors

$$\sum_{j=1}^{k-1} \alpha_j \lambda_k u_j = \lambda_k u_k = A u_k = \sum_{j=1}^{k-1} \alpha_j A u_j = \sum_{j=1}^{k-1} \alpha_j \lambda_j u_j$$

Donc

$$\sum_{j=1}^{k-1} \alpha_j (\lambda_k - \lambda_j) u_j = 0.$$

Par hypothèse de récurrence, tous les coefficients $\alpha_j (\lambda_k - \lambda_j)$ sont nuls. Mais puisque $u_k \neq 0$, il existe au moins un j_0 tel que $\alpha_{j_0} \neq 0$. On a donc $\lambda_k = \lambda_{j_0}$ ce qui est absurde.

□

Autrement dit, la somme (2.3) est toujours directe, mais n'est en général pas égale à \mathbb{C}^n : la somme des multiplicités géométriques des valeurs propres est en général $< n$. On retient aussi que A est diagonalisable si et seulement si la relation (2.3) est vraie :

Définition 2.2.5 On dit que $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable sur \mathbb{R} (resp. sur \mathbb{C}) s'il existe une base de $\{\mathbb{R}^n, \mathbb{C}^n\}$ constituée de vecteurs propres de A .

De manière équivalente, une matrice $A \in \mathcal{M}_n(\mathbb{C})$ (par exemple) est diagonalisable si tout vecteur de \mathbb{C}^n peut s'écrire comme somme de vecteurs propres de A , c'est-à-dire

$$(2.4) \quad \mathbb{C}^n = \text{Ker}(A - \lambda_1 I) \oplus \dots \oplus \text{Ker}(A - \lambda_d I),$$

où $\text{sp}(A) = \{\lambda_1, \dots, \lambda_d\}$.

Proposition 2.2.6 $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable si et seulement si il existe une matrice P inversible et une matrice D diagonale telles que

$$A = P D P^{-1}.$$

Preuve: Si A est diagonalisable, la matrice P de passage de la base canonique à la base de vecteurs propres de A convient. Réciproquement, si $A = P D P^{-1}$ les colonnes de P forment un système de n vecteurs propres indépendants. □

2.2.2 Polynôme caractéristique

Voici un moyen commode de calculer les valeurs propres d'une matrice A .

Proposition 2.2.7 Les valeurs propres de A sont les racines du polynôme de degré n

$$P_A(\lambda) = \det(\lambda I - A),$$

appelé polynôme caractéristique de la matrice A . La multiplicité d'une valeur propre λ comme racine de P_A est appelée multiplicité algébrique de λ . On la note $m(\lambda)$.

Preuve: Dire que $Au = \lambda u$ pour un $u \neq 0$ équivaut à dire qu'il existe un vecteur non-nul dans $\text{Ker}(A - \lambda I)$, et donc que $A - \lambda I$ n'est pas inversible, ou encore que $\det(A - \lambda I) = 0$. \square

Puisque tout polynôme admet au moins une racine dans \mathbb{C} , le spectre d'une matrice n'est jamais vide. Par contre il se peut qu'une matrice à coefficients réels n'admette aucune valeur propre réelle.

Proposition 2.2.8 Si A et B sont deux matrices semblables, elles ont même polynôme caractéristique, et donc même spectre.

Preuve: Soit P inversible telle que $A = PBP^{-1}$. On a

$$\begin{aligned} P_A(\lambda) &= \det(\lambda I - A) = \det(P) \det(\lambda I - A) \det(P^{-1}) \\ &= \det(P(\lambda I - A)P^{-1}) = \det(\lambda I - PAP^{-1}) = P_B(\lambda). \end{aligned}$$

\square

On peut donc en particulier parler du polynôme caractéristique et du spectre d'une application linéaire a : il s'agit du polynôme caractéristique et du spectre de n'importe quelle représentation matricielle de a .

Proposition 2.2.9 Soit λ_0 une valeur propre de A . On a

$$1 \leq \mu(\lambda_0) \leq m(\lambda_0).$$

Preuve: On note pour simplifier $\mu = \mu(\lambda_0)$. Soit (e_1, e_2, \dots, e_μ) une base de E_{λ_0} . On peut la compléter en une base $\mathcal{B} = (e_1, \dots, e_\mu, e_{\mu+1}, \dots, e_n)$ de \mathbb{K}^n . Soit P la matrice de passage de la base canonique à la base \mathcal{B} , c'est-à-dire la matrice dont les colonnes sont les coordonnées des vecteurs e_j dans la base canonique. On peut facilement vérifier (par exemple en appliquant chacune des deux matrices aux vecteurs de la base canonique) que

$$P^{-1}AP = \begin{pmatrix} \lambda_0 I_\mu & C \\ 0 & D \end{pmatrix},$$

où I_μ est la matrice identité de $\mathcal{M}_\mu(\mathbb{K})$, et C et D sont deux matrices qu'il est inutile de calculer. En effet on peut écrire

$$(2.5) \quad \begin{pmatrix} \lambda_0 I_\mu & C \\ 0 & D \end{pmatrix} = \begin{pmatrix} I_\mu & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} \lambda_0 I_\mu & C \\ 0 & I_{n-\mu} \end{pmatrix},$$

ce qui montre (par exemple en développant ces déterminants selon les lignes) que

$$P_A(\lambda) = \det(A - \lambda I) = \det((\lambda_0 - \lambda)I_\mu) \det D = (\lambda_0 - \lambda)^\mu \det D.$$

Ainsi λ_0 est une racine de P_A d'ordre au moins μ . □

2.2.3 Le théorème de Cayley-Hamilton

Rappelons qu'on dit qu'un polynôme est scindé sur \mathbb{K} s'il s'écrit comme produit de facteurs de degré 1 à coefficients dans \mathbb{K} . Dans \mathbb{C} , tout polynôme est scindé, mais ce n'est pas le cas dans \mathbb{R} , comme par exemple $P(x) = x^2 + 1$.

Proposition 2.2.10 (Condition nécessaire et suffisante de trigonalisabilité) Une matrice A est semblable dans \mathbb{K} à une matrice triangulaire si et seulement si son polynôme caractéristique est scindé sur \mathbb{K} . En particulier dans \mathbb{C} , toute matrice est semblable à une matrice triangulaire.

Preuve: Soit $A \in \mathcal{M}_n(K)$. Puisque le polynôme caractéristique de A est scindé, A admet au moins une valeur propre λ . Soit $e_1 \in \mathbb{K}^n$ un vecteur propre associé à λ , et F_1 le sous-espace vectoriel engendré par e_1 . Puisque $Ae_1 = \lambda e_1$, F_1 est stable par A . Soit alors F'_1 un supplémentaire de F_1 dans \mathbb{K}^n , et \mathcal{B}'_1 une base de F'_1 . Dans la base $\mathcal{B} = \{e_1\} \cup \mathcal{B}'_1$, la matrice de A est de la forme

$$\left(\begin{array}{c|ccc} \lambda & * & \dots & * \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array} \right),$$

où $A_1 \in \mathcal{M}_n(\mathbb{K})$. On raisonne alors par récurrence. La proposition est trivialement vraie pour $n = 1$, et si elle est vraie pour les matrices de taille $n - 1$, le raisonnement précédent montre qu'elle est vraie pour les matrices de taille n .

Réciproquement, si $A = PBP^{-1}$ avec B triangulaire supérieure, on a $P_A(x) = P_B(x)$ grâce à la Proposition 2.2.8, et en développant $\det(\lambda I - B)$ par rapport à ses colonnes par exemple, on voit que

$$P_B(\lambda) = \prod_{j=1}^n (\lambda - b_{j,j}),$$

où les $b_{j,j}$ sont les coefficients diagonaux de B . Donc P_A est scindé, et les valeurs propres de A (et de B) sont les $b_{j,j}$. □

Proposition 2.2.11 (Théorème de Cayley-Hamilton) Soit $A \in \mathcal{M}_n(\mathbb{K})$, et $P_A \in \mathbb{K}^n[x]$ son polynôme caractéristique. On a

$$P_A(A) = 0.$$

Preuve: Soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de \mathbb{C}^n dans laquelle A est triangulaire supérieure. On note $F_0 = \{0\}$, $F_1 = \langle \{e_1\} \rangle$, $F_2 = \langle \{e_1, e_2\} \rangle$, ..., $F_n = \langle \{e_1, \dots, e_n\} \rangle = \mathbb{C}^n$. Les F_j sont stables par A , et vérifient

$$F_0 \subsetneq F_1 \subsetneq \dots \subsetneq F_n = \mathbb{C}^n.$$

On note λ_j les éléments diagonaux de B : ce sont les valeurs propres de A rangées dans un certain ordre, répétées suivant leur multiplicité algébrique. Puisque B est triangulaire supérieure, pour tout $j \in \{1, \dots, n\}$ on a $(A - \lambda_j I)F_j \subset F_{j-1}$, et donc

$$P_A(A)(\mathbb{C}^n) = \prod_{j=1}^n (A - \lambda_j I)F_n \subset \prod_{j=1}^{n-1} (A - \lambda_j I)F_{n-1} \subset \dots \subset (A - \lambda_1 I)F_1 = \{0\},$$

ce qui prouve le théorème pour $\mathbb{K} = \mathbb{C}$. Dans le cas où $\mathbb{K} = \mathbb{R}$, on voit aussi que

$$P_A(A)(\mathbb{R}^n) = 0,$$

lorsque P_A est le polynôme caractéristique de A écrit sous forme scindée dans $\mathbb{C}[x]$. Si \tilde{P}_A est le polynôme caractéristique de A à coefficients réels, on a de toutes façons, pour $x \in \mathbb{R}$, $\tilde{P}_A(x) = P_A(x)$, donc $\tilde{P}_A(A) = 0$. \square

2.3 Espaces propres généralisés

On a compris qu'une matrice n'est pas diagonalisable lorsque ses sous-espaces propres sont trop petits. On introduit maintenant une autre famille de sous-espaces stables associés aux valeurs propres de A , qui contiennent le sous-espace propre correspondant.

2.3.1 Définition et propriétés

Soit $A \in \mathcal{M}_n(\mathbb{K})$, et

$$P_A(\lambda) = \prod_{j=1}^d (\lambda - \lambda_j)^{m_j}$$

le polynôme caractéristique de A - qu'on a écrit sous forme scindé : les λ_j peuvent être dans \mathbb{C} . On a noté par commodité $m_j = m(\lambda_j)$ la multiplicité algébrique de λ_j .

Définition 2.3.1 On appelle espace propre généralisé associé à la valeur propre λ_j le sous-espace vectoriel

$$G_j = \text{Ker}((A - \lambda_j I)^{m_j})$$

Proposition 2.3.2 Les sous-espaces propres généralisés ont les propriétés suivantes :

- i) Chaque G_j est stable par A .
- ii) $E_j \subset G_j$, où E_j est le sous-espace vectoriel propre associé à λ_j .
- iii) $\mathbb{C}^n = G_1 \oplus G_2 \oplus \dots \oplus G_d$.
- iv) Pour tout j , $\dim G_j = m_j$.

Preuve:

i) Soit $u \in G_j$. On a $(A - \lambda_j I)^{m_j}(Au) = A(A - \lambda_j I)^{m_j}(u) = 0$, ce qui montre que $Au \in G_j$.

ii) Si $m_j = 1$, $E_j = G_j$. Sinon $m_j > 1$, et pour $u \in E_j$,

$$(A - \lambda_j I)^{m_j}u = (A - \lambda_j I)^{m_j-1}(A - \lambda_j I)u = 0,$$

donc $u \in G_j$.

iii) On démontre d'abord le résultat général suivant :

Lemme 2.3.3 (Lemme des Noyaux) Soit $P \in \mathbb{C}^n[x]$ un polynôme, tel que $P = P_1 P_2 \dots P_k$ où les polynômes P_j sont deux à deux premiers entre eux, alors

$$\text{Ker } P(A) = \text{Ker } P_1(A) \oplus \text{Ker } P_2(A) \oplus \dots \oplus \text{Ker } P_k(A).$$

Preuve: On n'écrit la preuve que dans le cas de deux polynômes ($k = 2$). Le cas général s'en déduit par une récurrence facile. D'après le théorème de Bézout, il existe deux polynômes $Q_1, Q_2 \in \mathbb{C}^n[x]$ tels que

$$1 = P_2(x)Q_2(x) + P_1(x)Q_1(x),$$

et donc tels que

$$I = P_2(A)Q_2(A) + P_1(A)Q_1(A).$$

Tout vecteur $U \in \mathbb{C}^n$ s'écrit donc

$$(2.6) \quad U = P_2(A)Q_2(A)U + P_1(A)Q_1(A)U.$$

Pour $U \in \text{Ker } P_1(A) \cap \text{Ker } P_2(A)$, (2.6) donne, puisque les polynôme d'une même matrice A commutent entre eux,

$$P_1(A)(Q_1(A)U) + P_2(A)(Q_2(A)U) = Q_1(A)(P_1(A)U) + Q_2(A)(P_2(A)U) = 0,$$

ce qui montre que

$$(2.7) \quad \text{Ker } P_1(A) \cap \text{Ker } P_2(A) = \{0\}.$$

Soit alors $U \in \text{Ker } P(A)$. On a

$$\begin{aligned} P_1(A)(P_2(A)Q_2(A)U) &= Q_2(A)(P_1(A)P_2(A)U) = Q_2(A)P(A)U = 0, \\ P_2(A)(P_1(A)Q_1(A)U) &= Q_1(A)(P_1(A)P_2(A)U) = Q_1(A)P(A)U = 0, \end{aligned}$$

et donc, d'après (2.6) et (2.7),

$$\text{Ker } P(A) \subset \text{Ker } P_1(A) \oplus \text{Ker } P_2(A).$$

Enfin, si $U \in \text{Ker } P_1(A)$ ou $U \in \text{Ker } P_2(A)$, on a

$$P(A)U = P_1(A)P_2(A)U = P_2(A)P_1(A)U = 0,$$

ce qui montre que

$$\text{Ker } P_1(A) \oplus \text{Ker } P_2(A) \subset \text{Ker } P(A),$$

et qui termine la preuve du Lemme. \square

Pour prouver le point (iii), on applique ce résultat au polynôme caractéristique $P = P_A$ de A , que l'on peut écrire $P(x) = \prod_{j=1}^d (x - \lambda_j)^{m_j}$. On notera que l'on a utilisé le fait que le polynôme P est automatiquement scindé sur \mathbb{C}^n . On a $P = P_1 P_2 \dots P_d$ avec $P_j(x) = (x - \lambda_j)^{m_j}$. Les polynômes P_j sont deux à deux premiers entre eux, et le théorème de Cayley-Hamilton donne $P(A) = 0$. On a donc

$$\mathbb{C}^n = \text{Ker } P(A) = \text{Ker } P_1(A) \oplus \dots \oplus \text{Ker } P_d(A) = G_1 \oplus \dots \oplus G_d.$$

- iv) Pour tout $j \in \{1, \dots, d\}$, on note $\alpha_j = \dim G_j$ et $\mathcal{B}_j = \{e_{j,1}, e_{j,2}, \dots, e_{j,\alpha_j}\}$ une base de G_j . D'après le point (iii) ci-dessus, $\mathcal{B} = \mathcal{B}_1 \cup \dots \mathcal{B}_d$ est une base de \mathbb{C}^n , et la matrice de l'application linéaire associée à A dans cette base est diagonale par blocs, c'est-à-dire de la forme :

$$(2.8) \quad A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_d \end{pmatrix},$$

avec $A_j \in \mathcal{M}_{\alpha_j}(\mathbb{C})$. Le polynôme caractéristique de A s'écrit donc (en raisonnant comme dans (2.5)),

$$P_A(x) = \prod_{j=1}^d P_{A_j}(x).$$

Soit λ une valeur propre de la matrice A_j . On $P_{A_j}(\lambda) = 0$ donc $P_A(\lambda) = 0$, et $\lambda = \lambda_k$ pour un $k \in \{1, \dots, d\}$. Soit alors $u_j \in \mathbb{C}^{\alpha_j} \setminus \{0\}$ un vecteur propre de A_j associé à $\lambda = \lambda_k$, et $U = 0 \oplus \dots \oplus u_j \oplus \dots \oplus 0 \in \mathbb{C}^n$. On a $U \in G_j$, et $AU = \lambda_k U$, d'où

$$0 \neq U \in G_j \cap \text{Ker}(A - \lambda_k I),$$

ce qui entraîne que $j = k$. Ainsi la seule valeur propre de A_j est λ_j , ce qui donne

$$P_{A_j}(x) = (x - \lambda_j)^{\alpha_j},$$

et

$$P_A(x) = \prod_{j=1}^d (x - \lambda_j)^{m_j} = \prod_{j=1}^d (x - \lambda_j)^{\alpha_j} = \prod_{j=1}^d P_{A_j}(x).$$

On a donc nécessairement $\alpha_j = m_j$.

□

2.3.2 Condition nécessaire et suffisante de diagonalisabilité

On est maintenant capable de donner une condition nécessaire et suffisante, facile à vérifier, pour qu'une matrice soit diagonalisable. La situation est un peu différente suivant que l'on cherche à diagonaliser sur \mathbb{R} ou sur \mathbb{C} :

Proposition 2.3.4 Soit $A \in \mathcal{M}_n(\mathbb{R})$. La matrice A est diagonalisable sur \mathbb{R} si et seulement si

- i) Le polynôme caractéristique de A est scindé sur \mathbb{R} ,
- ii) Pour toute valeur propre λ de A , $\mu(\lambda) = m(\lambda)$.

Tout polynôme est scindé sur \mathbb{C} . Pour la diagonalisabilité sur \mathbb{C} , la condition nécessaire et suffisante devient

Proposition 2.3.5 Soit $A \in \mathcal{M}_n(\mathbb{K})$. La matrice A est diagonalisable sur \mathbb{C} si et seulement si pour toute valeur propre λ de A , on a $\mu(\lambda) = m(\lambda)$.

Preuve: Soit $\lambda_1, \lambda_2, \dots, \lambda_d$ les valeurs propres de A , de multiplicité algébrique m_j et de multiplicité géométrique μ_j .

Si $\mu_j = m_j$, puisque $E_j \subset G_j$ d'après le point (ii) ci-dessus, et $\dim E_j = \mu_j = m_j = \dim G_j$, on a $E_j = G_j$, et le point (iii) donne

$$\mathbb{C}^n = E_1 \oplus \dots \oplus E_d,$$

donc A est diagonalisable.

Si A est diagonalisable, on a

$$G_1 \oplus \cdots \oplus G_d = \mathbb{C}^n = E_1 \oplus \cdots \oplus E_d,$$

donc en particulier $n = \sum_{j=1}^d m_j = \sum_{j=1}^d \mu_j$. D'après la Proposition 2.2.9, on a $\mu_j \leq m_j$ pour tout $j \in \{1, \dots, d\}$. Il est donc impossible que l'on ait $\mu_k < m_k$ pour un certain k . \square

2.4 Forme normale de Jordan

On a donc rempli la première partie du programme : en choisissant une base \mathcal{B} de \mathbb{C}^n qui est la réunion de bases \mathcal{B}_j de chaque G_j , la matrice de A dans cette base s'écrit sous la forme diagonale par bloc (2.2), où chaque bloc A_j est la restriction de A au sous-espace G_j . Posant

$$N_j = A_j - \lambda_j I,$$

on a d'ailleurs $N_j^{m_j} = 0$. Autrement dit $A_j = \lambda_j I + N_j$ avec N_j matrice nilpotente d'ordre $\leq m_j$.

2.4.1 Matrices nilpotentes

Définition 2.4.1 Une matrice $N \neq 0$ est nilpotente lorsqu'il existe $s \in \mathbb{N}^*$ tel que $N^s = 0$. Le plus petit entier $s \geq 1$ tel que $N^s = 0$ est l'ordre de nilpotence de N .

Proposition 2.4.2 Soit $N \in \mathcal{M}_n(\mathbb{C})$. Les assertions suivantes sont équivalentes

- i) N est nilpotente.
- ii) La seule valeur propre de N est 0.
- iii) Le polynôme caractéristique de N est $P_N(x) = x^n$.

Preuve: Supposons que N est nilpotente d'ordre p . Si λ est une valeur propre de N , associée au vecteur propre $U \in \mathbb{C}^n$, on a

$$0 = N^p u = N^{p-1} N u = \lambda N^{p-1} u = \cdots = \lambda^p u,$$

donc $\lambda = 0$. Le polynôme caractéristique de P_N est un polynôme scindé dans $\mathbb{C}^n[x]$, de degré n , dont 0 est la seule racine. On a donc $P_N(x) = cx^n$ pour un certain $c \in \mathbb{C}$. Or $P_N(x) = \det(xI - N)$, donc $c = 1$. On a prouvé que (i) \Rightarrow (ii) \Rightarrow (iii). Réciproquement, si $P_N(x) = x^n$, le théorème de Cayley-Hamilton donne $N^n = P_N(N) = 0$. Donc N est nilpotente d'ordre inférieur ou égal à n . \square

Remarque 2.4.3 — L'ordre de nilpotence de $N \in \mathcal{M}_n(\mathbb{C})$ est toujours inférieur à n .
 — Si N est une matrice nilpotente et diagonalisable, alors N est semblable à la matrice nulle, donc est nulle.

L'exponentielle d'une matrice nilpotente est, en principe, facile à calculer. C'est en particulier un polynôme en t , puisque

Si N est nilpotente d'ordre p ,

$$e^{tN} = I + tN + \frac{t^2}{2!}N^2 + \cdots + \frac{t^{(p-1)}}{(p-1)!}N^{(p-1)}.$$

2.4.2 Forme normale de Jordan

Puisque l'on peut toujours triangulariser une matrice (sur \mathbb{C}), et qu'alors la diagonale contient les valeurs propres de la matrice (cf. la Proposition 2.2.10), on voit que toute matrice nilpotente est semblable à une matrice de la forme

$$\begin{pmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

où les $*$ désignent un coefficient complexe dont on ignore tout à priori. On peut cependant démontrer le résultat suivant :

Proposition 2.4.4 Soit $N \in \mathcal{M}_k(\mathbb{C})$ une matrice nilpotente. La matrice N est semblable à une matrice diagonale par blocs de la forme

$$(2.9) \quad N = \begin{pmatrix} N_1 & & & & \\ & N_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & N_r \end{pmatrix},$$

où chaque bloc $N_\ell \in \mathcal{M}_{k_\ell}(\mathbb{R})$, avec $1 \leq k_\ell \leq k$, et $k_1 + \cdots + k_r = k$, est de la forme

$$N_j = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & 0 & \\ \vdots & & & & \ddots & 1 \\ 0 & \cdots & & & \cdots & 0 \end{pmatrix}$$

Revenant à la matrice A initiale, on obtient sa forme normale de Jordan :

Proposition 2.4.5 Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est semblable à une matrice de la forme

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_d \end{pmatrix},$$

avec $A_j \in \mathcal{M}_{m_j}(\mathbb{C})$ donnée par

$$A_j = \begin{pmatrix} T_1^j & & & \\ & T_2^j & & \\ & & \ddots & \\ & & & T_{r_j}^j \end{pmatrix},$$

et $T_\ell^j \in \mathcal{M}_{k_\ell}(\mathbb{C})$ donnée par

$$T_\ell^j = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & \dots & 0 \\ \vdots & \lambda_j & 1 & 0 & \dots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & 0 & \\ \vdots & & & & \ddots & 1 \\ 0 & \dots & & & \dots & \lambda_j \end{pmatrix}$$

On notera les deux cas extrêmes : on peut avoir $r_j = 1$ (la matrice N_j est d'un seul bloc), et dans ce cas

$$A_j = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & \dots & 0 \\ \vdots & \lambda_j & 1 & 0 & \dots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & 0 & \\ \vdots & & & & \ddots & 1 \\ 0 & \dots & & & \dots & \lambda_j \end{pmatrix}.$$

On peut aussi avoir $r_j = m_j$ (la matrice N_j est constitué de m_j blocs de taille 1), ce qui signifie que A_j s'écrit

$$A_j = \begin{pmatrix} \lambda_j & 0 & 0 & \dots & \dots & 0 \\ \vdots & \lambda_j & 0 & 0 & \dots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & & \ddots & 0 \\ \vdots & & & & \ddots & 0 \\ 0 & \dots & & \dots & \dots & \lambda_j \end{pmatrix}.$$

Exercice 2.4.6 Dresser la liste de tous les blocs A_j possibles pour $A \in \mathcal{M}_2(\mathbb{C})$, $A \in \mathcal{M}_3(\mathbb{C})$, puis pour $A \in \mathcal{M}_4(\mathbb{C})$.

2.4.3 Méthode pratique pour le calcul

Voici un algorithme pour calculer la forme de Jordan d'une matrice A :

1) Déterminer les valeurs propres et les vecteurs propres de A . Si la matrice A est diagonalisable, c'est fini. Sinon :

2) Déterminer une base adaptée de chacun des sous-espaces propres généralisés. Pour ce faire, on peut procéder ainsi : pour chaque valeur propre λ_j ,

- On choisit un vecteur propre $e_{1,j}$ de A associé à la valeur propre λ_j (déterminé à l'étape 1). On cherche alors $e_{2,j}$ tel que (il s'agit de résoudre un système linéaire dont les inconnues sont les coordonnées de $e_{2,j}$)

$$(A - \lambda_j I)e_{2,j} = e_{1,j}.$$

Pour un tel vecteur on aura en effet

$$(A - \lambda_j I)^2 e_{2,j} = (A - \lambda_j I)e_{1,j} = 0,$$

donc $e_{2,j} \in G_j$. De plus $e_{1,j}$ et $e_{2,j}$ seront forcément linéairement indépendants : sinon $e_{2,j}$ serait aussi un vecteur propre associé à λ_j et $(A - \lambda_j I)e_{2,j} = 0$ ce qui est absurde. Enfin

$$Ae_{2,j} = e_{1,j} + \lambda_j e_{2,j},$$

ce qui est la forme cherchée.

- On répète ce procédé, au plus $m_j - 1$ fois : on cherche $e_{k+1,j}$ tel que $(A - \lambda_j I)e_{k+1,j} = e_{k,j}$. On peut alors être confronté au problème suivant : s'il existe $p < m_j$ tel que $(A - \lambda_j I)^p = 0$, le procédé s'arrête après $p - 1$ étapes : on ne fabrique donc pas assez de vecteurs pour former une base de G_j . On doit alors recommencer le processus avec un autre vecteur propre pour la valeur propre λ_j .

Voici un exemple : Soit $A \in \mathcal{M}_4(\mathbb{C})$ la matrice définie par

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

- Son polynôme caractéristique est $P_A(x) = ((x-1)^2 + 1)^2$ (développer par rapport à la première colonne par exemple). La matrice A a donc deux valeurs propres $\lambda_1 = 1+i$ et $\lambda_2 = 1-i$, de multiplicités algébriques respectives $m_1 = m_2 = 2$. Les sous-espaces propres associés sont de dimension 1, engendrés par $e_{1,1} = (1, 0, -i, 0)$ et $e_{2,1} = (1, 0, i, 0)$ respectivement. La matrice A n'est donc pas diagonalisable, et on veut l'écrire sous forme de Jordan.

- On cherche un vecteur $e_{1,2} = (x_1, x_2, x_3, x_4)$ tel que $(A - \lambda_1)e_{1,2} = e_{1,1}$. Le système correspondant s'écrit

$$\begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -i \\ 0 \end{pmatrix}.$$

On trouve $e_{1,2} = (0, 1, 0, -i)$. Puisque $m_1 = 2$ on sait que $(e_{1,1}, e_{1,2})$ est une base du sous-espace propre généralisé G_1 associé à λ_1 .

- On procède de la même manière pour trouver une base de G_2 . On cherche un vecteur $e_{2,2}$ tel que $(A - \lambda_1)e_{2,2} = e_{2,1}$. Tous calculs faits, on trouve $e_{2,2} = (0, 1, 0, i)$.

Finalement, notant P la matrice de passage de la base canonique à la base \mathcal{B} , c'est à dire

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -i & 0 & i & 0 \\ 0 & -i & 0 & i \end{pmatrix},$$

on a

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

Il reste à calculer P^{-1} - par exemple avec la méthode de Gauss. On trouve

$$P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 0 & i & 0 \\ 0 & 1 & 0 & i \\ 1 & 0 & -i & 0 \\ 0 & 1 & 0 & -i \end{pmatrix}.$$

Chapitre 3

Systèmes linéaires

On veut décrire complètement et explicitement l'ensemble des solutions des systèmes linéaires à coefficients constants

$$Y' = AY, A \in \mathcal{M}_n(\mathbb{R}).$$

On verra à la fin de ce chapitre ce qui persiste de cette étude dans le cas des systèmes linéaires à coefficients variables.

$$Y' = A(t)Y + B(t).$$

Pour commencer, on revient sur le cas de l'équation scalaire linéaire d'ordre 1 homogène à coefficients constants. On considère le problème de Cauchy

$$\begin{cases} y' = ay, \\ y(0) = 1, \end{cases}$$

dont on a noté $y : t \mapsto e^{at}$ la solution. Quelles informations l'équation donne-t-elle sur cette fonction ?

Une fonction $y \in \mathcal{C}^1$ est solution du problème de Cauchy ci-dessus si et seulement si

$$y(t) = a \int_0^t y'(s) ds + y(0) = 1 + a \int_0^t y(s) ds.$$

Il semble qu'on n'ait rien gagné à écrire ainsi l'équation : la fonction inconnue y se trouve des deux côtés de l'égalité. L'idée géniale (n'ayons pas peur des mots) est de définir une suite de fonctions (y_n) par récurrence, en posant $y_0 : t \mapsto 1$ et

$$y_{n+1} : t \mapsto 1 + a \int_0^t y_{n-1}(s) ds.$$

Si la suite (y_n) converge uniformément vers une fonction y , on doit avoir, pour chaque t

$$y(t) = \lim_{n \rightarrow +\infty} y_{n+1}(t) = \lim_{n \rightarrow +\infty} \left(1 + \int_0^t y_{n-1}(s) ds \right) = 1 + \int_0^t \lim_{n \rightarrow +\infty} y_{n-1}(s) ds = 1 + \int_0^t y(s) ds.$$

Autrement dit si la suite (y_n) converge uniformément vers y , la fonction y est une solution du problème de Cauchy.

On peut calculer les fonctions y_n de proche en proche :

$$y_1(t) = 1 + a \int_0^t 1 ds = 1 + at, \quad y_2(t) = 1 + \int_0^t (1 + as) ds = 1 + at + \frac{(at)^2}{2}, \dots y_n(t) = \sum_{k=0}^n \frac{(at)^k}{k!}.$$

Sous réserve de convergence uniforme (qui a bien lieu), on a donc

$$e^{at} = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{(at)^k}{k!}.$$

Dans ce qui suit nous allons donner un sens à la fonction

$$t \mapsto e^{tA} X_0 = \left(\lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{(tA)^k}{k!} \right) X_0,$$

pour n'importe quelle matrice $A \in \mathcal{M}_n(\mathbb{R})$ (ou même $A \in \mathcal{M}_n(\mathbb{C})$), et n'importe quel vecteur $X_0 \in \mathbb{R}^n$ (ou même $X_0 \in \mathbb{C}^n$). Nous verrons que cette fonction est la solution du problème de Cauchy

$$\begin{cases} Y' = AY, \\ Y(0) = X_0. \end{cases}$$

3.1 Exponentielle de matrice

3.1.1 Norme d'une matrice

Pour simplifier l'exposé, on travaille sur le corps \mathbb{R} . Jusqu'à mention explicite du contraire, ce qui suit peut être lu en remplaçant \mathbb{R}, \mathbb{R}^n et $\mathcal{M}_n(\mathbb{R})$ par \mathbb{C}, \mathbb{C}^n et $\mathcal{M}_n(\mathbb{C})$ respectivement. Pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on note $\|x\|$ la quantité

$$\|x\| = \max_{j=1, \dots, n} |x_j|$$

La fonction $x \mapsto \|x\|$ est une norme sur \mathbb{R}^n (ou \mathbb{C}^n), et $(\mathbb{R}^n, \|\cdot\|)$ est un espace vectoriel normé complet (on dit aussi un espace de Banach).

Pour $A \in \mathcal{M}_n(\mathbb{R})$, on note $\|A\|$ la quantité

$$\|A\| = \sup\{\|Ax\|, x \in \mathbb{R}^n, \|x\| = 1\}$$

Proposition 3.1.1 La fonction $A \mapsto \|A\|$ est une norme sur $\mathcal{M}_n(\mathbb{R})$, et $(\mathcal{M}_n(\mathbb{R}), \|\cdot\|)$ est un espace vectoriel normé complet. On a aussi, pour toutes matrices $A, B \in \mathcal{M}_n(\mathbb{R})$, $\|AB\| \leq \|A\| \|B\|$.

Preuve: D'abord l'ensemble Σ des $x \in \mathbb{R}^n$ tels que $\|x\| = 1$ est un compact de \mathbb{R}^n , et la fonction $x \mapsto \|Ax\|$ est continue sur \mathbb{R}^n . Donc elle admet une borne supérieure sur Σ , et $\|A\|$ désigne bien un élément de \mathbb{R}^+ .

Si $\|A\| = 0$, on a $\|Ax\| = 0$ pour tout $x \in \Sigma$, donc $A = 0$. Bien sûr $\|A\| = 0$ si $A = 0$. Soit $\lambda \in \mathbb{R}$. On a

$$\|\lambda A\| = \sup\{\|(\lambda A)x\|, x \in \Sigma\} = \sup\{|\lambda| \|Ax\|, x \in \Sigma\} = |\lambda| \|A\|.$$

Enfin, pour $x \in \Sigma$, on a

$$\|(A + B)x\| \leq \|Ax\| + \|Bx\| \leq \sup_{x \in \Sigma} \|Ax\| + \sup_{x \in \Sigma} \|Bx\| \leq \|A\| + \|B\|,$$

ce qui montre, en prenant la borne supérieure du membre de gauche, que l'on a bien l'inégalité triangulaire

$$\|A + B\| \leq \|A\| + \|B\|.$$

On vient de prouver que $(\mathcal{M}_n(\mathbb{R}), \|\cdot\|)$ est un espace vectoriel normé. Le fait qu'il est complet peut être vu de différentes manières. On peut remarquer par exemple que pour tout coefficient $a_{i,j}$ de A , on a

$$|a_{i,j}| \leq \|A\|.$$

Il suffit pour s'en convaincre de prendre $x = e_j = (0, \dots, 0, 1, 0, \dots, 0)$ (un 1 à la j -ème place). On a alors

$$\|A\| \geq \|Ax\| = \left\| \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix} \right\| \geq |a_{i,j}|$$

Supposons alors que (A_n) soit une suite de Cauchy dans $(\mathcal{M}_n(\mathbb{R}), \|\cdot\|)$. Toutes les suites $(a_{i,j}^n)$ de coefficients sont des suites de Cauchy de \mathbb{R} , donc convergent vers un $a_{i,j}$. Notant $A = (a_{i,j})_{i,j}$ la matrice correspondante, il est facile de voir que (A_n) tend vers A .

On a aussi, pour $x \in \mathbb{R}^n \setminus \{0\}$, notant que $y = x/\|x\| \in \Sigma$,

$$\|Ax\| = \|x\| \|A(\frac{x}{\|x\|})\| \leq \|x\| \sup_{y \in \Sigma} \|Ay\| \leq \|A\| \|x\|.$$

Du coup, pour $x \in \Sigma$,

$$\|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\|,$$

ce qui, en considérant la borne supérieure du membre de gauche donne bien $\|AB\| \leq \|A\| \|B\|$. \square

3.1.2 Définition

Proposition 3.1.2 Soit $A \in \mathcal{M}_n(\mathbb{R})$. La série de matrice de terme général $(\frac{1}{k!} A^k)$ est convergente.

Preuve: On note d'abord que la série numérique de terme général $(\frac{1}{k!} \|A\|^k)$ est convergente, puisqu'il s'agit d'une série entière de rayon de convergence infini. Comme dans tout espace vectoriel normé complet, la série normalement convergente est convergente : pour $q > p$, on a

$$\left\| \sum_{k=0}^q \frac{1}{k!} A^k - \sum_{k=0}^p \frac{1}{k!} A^k \right\| \leq \sum_{k=p+1}^q \frac{1}{k!} \|A\|^k.$$

Or le membre de droite peut être rendu aussi petit que souhaité en prenant p et q assez grand, puisque, la série $\sum_{k \geq 0} \frac{1}{k!} A^k$ étant convergente, c'est une série de Cauchy. Ainsi la série de terme général $\frac{1}{k!} A^k$ est une série de Cauchy, donc converge. \square

Définition 3.1.3 Pour $A \in \mathcal{M}_n(\mathbb{R})$, on note $\exp(A)$ (ou e^A) la somme de la série

$$\exp(A) = \sum_{k \geq 0} \frac{1}{k!} A^k.$$

C'est un élément de $\mathcal{M}_n(\mathbb{R})$.

Exemples 3.1.4 i) Notant $O \in \mathcal{M}_n(\mathbb{R})$ la matrice nulle, on a $\exp(O) = I + O + O^2/2! + \dots = I$.

ii) Si $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ est une matrice diagonale, on a $A^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$, donc $e^A = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.

3.1.3 La fonction $t \mapsto e^{tA}$

Dans l'introduction de ce cours, on a caractérisé la fonction exponentielle $t \mapsto e^{at}$ de \mathbb{R} dans \mathbb{R} comme la solution du problème de Cauchy

$$\begin{cases} y' = ay, \\ y(0) = 1. \end{cases}$$

On montre maintenant que la fonction $t \mapsto e^{tA}$ à valeurs dans $\mathcal{M}_n(\mathbb{R})$ joue le même rôle. On commence par la

Proposition 3.1.5 Soit $A \in \mathcal{M}_n(\mathbb{R})$. La fonction $\phi : t \mapsto e^{tA}$ est \mathcal{C}^∞ de \mathbb{R} dans $\mathcal{M}_n(\mathbb{R})$. De plus

- i) Pour tous $t_1, t_2 \in \mathbb{R}$, $\phi(t_1 + t_2) = \phi(t_1)\phi(t_2)$.
- ii) Pour tout $t \in \mathbb{R}$, $\phi(t)$ est inversible, d'inverse $(\phi(t))^{-1} = \phi(-t)$.
- iii) Pour tout $n \in \mathbb{N}$, et tout $t \in \mathbb{R}$, $\phi^{(n)}(t) = A^n e^{tA}$.

Preuve: On prouve d'abord le point (i). En multipliant terme à terme les séries normalement convergentes qui définissent $\phi(t_1)$ et $\phi(t_2)$ on trouve

$$\begin{aligned}\phi(t_1)\phi(t_2) &= \left(\sum_{p \geq 0} \frac{t_1^p}{p!} A^p\right) \left(\sum_{q \geq 0} \frac{t_2^q}{q!} A^q\right) = \sum_{p,q \geq 0} \frac{t_1^p t_2^q}{p! q!} A^{p+q} = \sum_{k \geq 0} \left(\sum_{p+q=k} \frac{t_1^p t_2^q}{p! q!}\right) A^k \\ &= \sum_{k \geq 0} \frac{1}{k!} \left(\sum_{p=0}^k \frac{k!}{p!(k-p)!} t_1^p t_2^{k-p}\right) A^k = \sum_{k \geq 0} \frac{1}{k!} (t_1 + t_2)^k A^k = \phi(t_1 + t_2).\end{aligned}$$

De ce fait $\phi(t)\phi(-t) = \phi(0) = I = \phi(-t)\phi(t)$, ce qui montre le point (ii). On passe à la preuve du point (iii). Soit $\phi_k(t) = \frac{t^k}{k!} A^k$. La fonction ϕ_k est clairement \mathcal{C}^∞ . On a $\phi'_0(t) = 0$ et pour $1 \leq k$, $\phi'_k(t) = \frac{t^{k-1}}{(k-1)!} A^k = A\phi_{k-1}(t)$. La série de terme général $\phi'_k(t)$ est uniformément convergente sur tout compact $[-m, m]$ de \mathbb{R} : pour $t \in [-m, m]$ on a

$$\sum_{k=0}^n \|\phi'_k(t)\| = \|A\| \sum_{k=1}^n \frac{|t|^{k-1}}{(k-1)!} \|A\|^{k-1} \leq \|A\| \|e^{mA}\|.$$

Puisque, de plus, la série de fonctions de terme général $\phi_k(t)$ converge simplement, sa fonction somme $\phi : t \mapsto e^{tA}$ est dérivable, et sa dérivée s'obtient en dérivant terme à terme :

$$\phi'(t) = \sum_{k \geq 0} \phi'_k(t) = A \sum_{k \geq 0} \phi_k(t) = Ae^{tA}.$$

On conclut par récurrence que ϕ est \mathcal{C}^∞ et que ses dérivées successives sont bien données par (iii). \square

Proposition 3.1.6 Soit $A \in \mathcal{M}_n(K)$, et $X_0 \in \mathbb{R}^n$. La fonction $\phi_{X_0} : \mathbb{R} \rightarrow \mathbb{R}^n$ donnée par $\phi_{X_0}(t) = e^{tA} X_0$ est l'unique solution du problème de Cauchy

$$(3.1) \quad \begin{cases} Y' = AY, \\ Y(0) = X_0. \end{cases}$$

Preuve: On a d'abord $\phi_{X_0}(0) = X_0$. On vérifie aussi que $\phi'_{X_0}(t) = A\phi_{X_0}(t)$, en écrivant le développement limité à l'ordre 1 de ϕ_{X_0} au point t :

$$\phi_{X_0}(t+h) = e^{(t+h)A} X_0 = e^{tA} e^{hA} X_0 = e^{tA} (I + hA + h^2 R_2(A, h)) X_0 = X_0 + h e^{tA} X_0 + O(h^2),$$

ou $R_2(A, h) = \sum_{j \geq 0} \frac{h^j}{(j+2)!} A^{2+j} = A^2 e^{hA}$ est bornée pour tout h assez petit.

Supposons maintenant que $Y(t)$ soit une solution de (3.1), et posons $X(t) = e^{-tA} Y(t)$. On calcule

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

Or, en revenant à la définition on a

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

donc $X'(t) = 0$, et la fonction X est nécessairement une fonction constante : pour tout $t \in \mathbb{R}$, $X(t) = X(0) = Y(0) = X_0$, ce qui donne $Y(t) = e^{tA}X_0 = \phi_{X_0}(t)$. \square

On notera que cette proposition n'est autre que le théorème de Cauchy-Lipschitz dans le cas des systèmes linéaires homogènes. Comme on l'a déjà dit, le théorème de Cauchy-Lipschitz n'a d'intérêt que lorsque l'on ne sait pas donner explicitement les solutions.

3.1.4 Quelques propriétés

On regroupe ici quelques propriétés de l'exponentielle de matrice qui se révéleront utiles pour leur calcul effectif.

Proposition 3.1.7 Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{R})$. Si A et B commutent, on a $e^{A+B} = e^A e^B$.

Preuve: On suppose que A et B commutent. On sait que $\phi_{X_0}(t) = e^{t(A+B)}X_0$ est l'unique solution du problème de Cauchy

$$(3.2) \quad \begin{cases} Y' = (A+B)Y, \\ Y(0) = X_0. \end{cases}$$

On va montrer que $\psi(t) = e^{tA}e^{tB}X_0$ vérifie aussi (3.2). On a d'abord clairement $\psi(0) = X_0$. On calcule

$$\psi'(t) = Ae^{tA}e^{tB}X_0 + e^{tA}Be^{tB}X_0.$$

Lorsque A et B commutent, A^k et B commutent pour tout $k \in \mathbb{N}$, et l'on a

$$e^{tA}B = \left(\sum_{k \geq 0} \frac{t^k}{k!} A^k \right) B = \sum_{k \geq 0} \frac{t^k}{k!} A^k B = \sum_{k \geq 0} \frac{t^k}{k!} A^k B = B \left(\sum_{k \geq 0} \frac{t^k}{k!} A^k \right) = Be^{tA}.$$

On a donc $\psi'(t) = Ae^{tA}e^{tB}X_0 + e^{tA}Be^{tB}X_0 = (A+B)e^{tA}e^{tB}X_0 = (A+B)\psi(t)$. Finalement $\psi(t) = \phi_{X_0}(t)$ pour tout $t \in \mathbb{R}$ et tout $X_0 \in \mathbb{R}^n$, ce qui, en prenant $t = 1$, donne la relation annoncée. \square

► On peut trouver des matrices A et B pour lesquelles $e^{A+B} \neq e^A e^B$. Prenons par exemple $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Les matrices A et B ne commutent pas puisque l'on a $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. On voit facilement que $A^2 = B^2 = 0$, et donc que

$$\exp(A) = I + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \exp(B) = I + B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

d'où

$$\exp(A)\exp(B) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

D'un autre côté, $A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, et on va montrer dans la section suivante que

$$\exp(A + B) = \frac{1}{2} \begin{pmatrix} e + 1/e & e - 1/e \\ e - 1/e & e + 1/e \end{pmatrix}.$$

On peut aussi montrer une réciproque affaiblie de cette proposition. Supposons que $e^{t(A+B)} = e^{tA}e^{tB}$ pour tout $t \in \mathbb{R}$. En dérivant deux fois cette égalité, on obtient

$$(A + B)e^{t(A+B)} = Ae^{tA}e^{tB} + e^{tA}Be^{tB}$$

puis

$$(A + B)^2 e^{t(A+B)} = A^2 e^{tA}e^{tB} + 2Ae^{tA}Be^{tB} + e^{tA}B^2 e^{tB},$$

ce qui donne, pour $t = 0$,

$$(A + B)^2 = A^2 + 2AB + B^2,$$

et donc $BA = AB$.

On verra plus loin l'importance de la Proposition 3.2 pour le calcul effectif des exponentielles de matrice. Voici d'ores et déjà un exemple. Soit $A \in \mathcal{M}_3(\mathbb{R})$ la matrice donnée par

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La matrice A n'est pas diagonalisable (voir plus loin), et le calcul successifs des puissances de A pour former e^A n'est pas simple. On remarque alors que $A = D + N$, avec

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

et que D et N commutent : $DN = ND = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. On a donc $e^A = e^D e^N$, et l'on sait

calculer $e^D = \text{diag}(e^2, e^2, e^1)$. Pour le calcul de e^N , on remarque que $N^2 = 0$, donc $e^N = I + N$. On peut donc calculer

$$e^A = \begin{pmatrix} e^2 & e^2 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^1 \end{pmatrix}.$$

Proposition 3.1.8 Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{R})$. S'il existe $P \in \mathcal{M}_n(\mathbb{R})$ inversible tel que $A = PBP^{-1}$, on a

$$e^A = Pe^BP^{-1}.$$

Preuve: On remarque que, pour tout $k \in \mathbb{N}$, $(PBP^{-1})^k = PB^kP^{-1}$, et donc

$$e^A = \sum_{k \geq 0} \frac{1}{k!} A^k = \sum_{k \geq 0} \frac{1}{k!} (PBP^{-1})^k = \sum_{k \geq 0} \frac{1}{k!} (PB^kP^{-1}) = P \left(\sum_{k \geq 0} \frac{1}{k!} B^k \right) P^{-1} = Pe^BP^{-1}.$$

□

3.1.5 Calcul de e^{tA}

On reprend le théorème de forme normale de Jordan sous une forme un peu moins précise, mais bien adaptée au calcul de l'exponentielle d'une matrice.

Proposition 3.1.9 (Décomposition de Dunford) Soit $A \in \mathcal{M}_n(\mathbb{C})$. Il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{C})$ et un couple de matrices (D, N) de $\mathcal{M}_n(\mathbb{C})$ telles que

i) D est diagonale, N est nilpotente.

ii) $DN = ND$.

iii) $A = P(D + N)P^{-1}$.

De plus, notant $P_A(x) = \prod_{j=1}^d (\lambda - \lambda_j)^{m_j}$ le polynôme caractéristique de A ,

$$(3.3) \quad D + N = \begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_d \end{pmatrix} \quad \text{où } T_j = D_j + N_j = \begin{pmatrix} \lambda_j & * & 0 & \cdots & 0 \\ & \lambda_j & \ddots & & \vdots \\ & & \ddots & & 0 \\ & & & \ddots & * \\ & & & & \lambda_j \end{pmatrix} \in \mathcal{M}_{m_j}(\mathbb{C}),$$

avec $D_j = \lambda_j I$ et $*$ = 0 ou 1.

On obtient finalement le résultat suivant pour e^{tA} :

Corollaire 3.1.10 Soit $A \in \mathcal{M}_n(\mathbb{C})$, et $t \in \mathbb{R}$. La matrice $\exp(tA)$ est semblable à la matrice

$$\begin{pmatrix} e^{tT_1} & & \\ & \ddots & \\ & & e^{tT_d} \end{pmatrix},$$

avec

$$e^{tT_j} = e^{tD_j} e^{tN_j} = e^{t\lambda_j} \left(I + tN_j + \frac{t^2}{2!} N_j^2 + \cdots + \frac{t^{(m_j-1)}}{(m_j-1)!} N_j^{(m_j-1)} \right),$$

où les $T_j = D_j + N_j$ sont les matrices données dans la Proposition 3.1.9.

Preuve.— On sait que dans une base bien choisie tA s'écrit sous la forme par blocs (3.3). Puisque le produit de matrices diagonales par blocs s'obtient en faisant le produit par blocs, il suffit de calculer e^{tT_j} pour un bloc quelconque. Puisque D_j et N_j commutent, on a

$$e^{tT_j} = e^{t(D_j + N_j)} = e^{tD_j} e^{tN_j},$$

et puisque $N_j \in \mathcal{M}_{m_j}(\mathbb{C})$ est nilpotente, $(N_j)^k = 0$ pour tout $k \geq m_j$, donc

$$e^{tT_j} = e^{t\lambda_j} \left(I + tN_j + \frac{t^2}{2!} N_j^2 + \cdots + \frac{t^{(m_j-1)}}{(m_j-1)!} N_j^{(m_j-1)} \right).$$

□

Reprenons l'exemple de la matrice

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

On sait que

$$A = P \left(\begin{array}{cc|cc} 1+i & 1 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ \hline 0 & 0 & 1-i & 1 \\ 0 & 1 & 0 & 1-i \end{array} \right) P^{-1},$$

et on va calculer séparément l'exponentielle des deux blocs diagonaux $T_1 = \begin{pmatrix} 1+i & 1 \\ 0 & 1+i \end{pmatrix}$ et $T_2 = \begin{pmatrix} 1-i & 1 \\ 0 & 1-i \end{pmatrix}$. On a $T_1 = \lambda_1 I + N$ et $T_2 = \lambda_2 I + N$, et le seul calcul à faire est celui de $e^{tN} = I + tN = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$. On obtient

$$e^{tA} = P \left(\begin{array}{cc|cc} e^{(1+i)t} & e^{(1+i)t}t & 0 & 0 \\ 0 & e^{(1+i)t} & 0 & 0 \\ \hline 0 & 0 & e^{(1-i)t} & e^{(1-i)t}t \\ 0 & 1 & 0 & e^{(1-i)t} \end{array} \right) P^{-1}.$$

3.2 Solutions des systèmes linéaires à coefficients constants

On revient maintenant à la résolution des systèmes linéaires à coefficients constants

$$(3.4) \quad Y' = AY(t) + B(t),$$

où $A \in \mathcal{M}_n(\mathbb{K})$ et $t \mapsto B(t)$ est une application continue d'un intervalle ouvert $I \subset \mathbb{R}$ dans \mathbb{K}^n . On sait que l'ensemble \mathcal{S} des solutions est l'ensemble des fonctions de la forme

$$Y(t) = e^{tA}X_0 + Y_1(t),$$

où $X_0 \in \mathbb{K}^n$, et $Y_1(t)$ est une solution de l'équation (3.4). Il reste deux problèmes :

- i) calculer efficacement $e^{tA}Y_0$,
- ii) trouver une solution Y_1 .

3.2.1 Calcul effectif de $e^{tA}X_0$

Pour le point (i), on utilise évidemment les résultats de la section précédente : on trouve une base $\mathcal{B} = \{V_1, \dots, V_n\}$ dans laquelle A s'écrit sous forme de Dunford (Jordan) $D + N$. Notant P la matrice de passage de la base canonique à la base \mathcal{B} , on a

$$(3.5) \quad e^{tA}X_0 = Pe^{t(D+N)}P^{-1}X_0.$$

Cette formule permet en principe le calcul, mais il est plus efficace de procéder de la manière suivante. Soit $(\alpha_1, \dots, \alpha_n)$ les coordonnées de X_0 dans la base \mathcal{B} . Autrement dit

$$X_0 = \alpha_1 V_1 + \dots + \alpha_n V_n.$$

On a

$$e^{tA}X_0 = \alpha_1 e^{tA}V_1 + \dots + \alpha_n e^{tA}V_n.$$

Par ailleurs, le calcul de $e^{tA}V_k$ se fait à vue, compte tenu de la forme donnée dans le Corollaire 3.1.10 : si $V_k \in G_j$,

$$(3.6) \quad e^{tA}V_k = e^{\lambda_j t} \left(I + tN_j + \frac{t^2}{2!}N_j^2 + \dots + \frac{t^{(m_j-1)}}{(m_j-1)!}N_j^{(m_j-1)} \right) V_k.$$

On pourra noter que $(\alpha_1, \dots, \alpha_n)$ sont les composantes du vecteur $P^{-1}X_0$, ce qui montre que cette approche ne diffère pas réellement de la formule (3.5). A ceci près qu'on n'a pas eu besoin de calculer P^{-1} !

Reprenons encore notre exemple :

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

On sait que

$$e^{tA} = P \left(\begin{array}{cc|cc} e^{(1+i)t} & te^{(1+i)t} & 0 & 0 \\ 0 & e^{(1+i)t} & 0 & 0 \\ \hline 0 & 0 & e^{(1-i)t} & te^{(1-i)t} \\ 0 & 1 & 0 & e^{(1-i)t} \end{array} \right) P^{-1},$$

mais on supposons que l'on veuille seulement calculer $e^{tA}X_0$, avec $X_0 = (2, 0, -2i, -i)$. On peut bien sûr utiliser la formule ci-dessus, et commencer par calculer $P^{-1}X_0$. Cependant, notant que $X_0 = 2e_{1,1} - e_{2,2}$, on voit beaucoup plus rapidement que

$$e^{tA}X_0 = 2e^{tA}e_{1,1} - e^{tA}e_{2,2} = 2e^{(1+i)t}e_{1,1} - e^{(1-i)t}(te_{2,1} + e_{2,2}) = \begin{pmatrix} 2e^{(1+i)t} - te^{(1-i)t} \\ -e^{(1-i)t} \\ -2e^{(1+i)t} - ite^{(1-i)t} \\ -ie^{(1-i)t} \end{pmatrix}$$

De la formule (3.6), on peut tirer facilement le résultat de structure suivant, qui n'a pas beaucoup d'intérêt autre que théorique.

Proposition 3.2.1 Les solutions à valeurs complexes du système différentiel $Y' = AY$ s'écrivent comme combinaisons linéaires de fonctions de la forme

$$t \mapsto e^{t\lambda_j} t^k u_{j,k},$$

où λ_j est une valeur propre de A , $k \leq m_j - 1$ un entier et $u_{j,k}$ un vecteur de \mathbb{C}^n .

3.2.2 Solutions réelles

On peut avoir envie de ne considérer que les solutions réelles de \mathcal{S}_H . En particulier, si $A \in \mathcal{M}_n(\mathbb{R})$ et $X_0 \in \mathbb{R}^n$, la fonction $t \mapsto e^{tA} X_0$ est à valeurs réelles : sa partie imaginaire $I(t)$ vérifie en effet le problème de Cauchy

$$\begin{cases} I' = AI, \\ I(0) = 0, \end{cases}$$

dont l'unique solution est la fonction nulle. Pourtant si les valeurs propres de A ne sont pas réelles, il n'est pas clair du tout que l'expression de $e^{tA} X_0$ donne une fonction à valeurs dans \mathbb{R}^n . On se souvient alors de la remarque suivante : si $A \in \mathcal{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{C} \setminus \mathbb{R}$ une valeur propre complexe de A associée au vecteur propre $u \in \mathbb{C}^n$, alors $\bar{\lambda}$ est une valeur propre de A associée au vecteur propre \bar{u} . On a en effet $Au = \lambda u$ donc $\overline{Au} = \bar{\lambda} \bar{u}$, et puisque $A \in \mathcal{M}_n(\mathbb{R})$, $\overline{Au} = A\bar{u}$.

L'idée consiste alors à regrouper les valeurs propres complexes conjuguées, et de choisir pour $G_{\bar{\lambda}}$ la base complexe conjuguée de celle de G_{λ} . Considérons l'exemple suivant : Soit $a, b \in \mathbb{R}$, avec $b \neq 0$, et $A \in \mathcal{M}_2(\mathbb{R})$ la matrice

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Les valeurs propres de A sont $\lambda = a + ib$ et $\bar{\lambda}$. Dans \mathbb{C} , A est semblable à la matrice $\text{diag}(\lambda, \bar{\lambda})$, et e^{tA} est semblable à la matrice

$$\begin{pmatrix} e^{t(a+ib)} & 0 \\ 0 & e^{t(a-ib)} \end{pmatrix}.$$

On se souvient que si $u = u_1 + iu_2$, avec $u_1, u_2 \in \mathbb{R}^2$ est un vecteur propre de A pour la valeur propre λ , $\bar{u} = u_1 - iu_2$ est un vecteur propre associé à $\bar{\lambda}$. Pour $X_0 = \alpha_1 u + \alpha_2 \bar{u} \in \mathbb{R}^2$, avec $\alpha_1, \alpha_2 \in \mathbb{C}$, on doit avoir $\alpha_2 = \overline{\alpha_1}$ et

$$e^{tA} X_0 = \alpha_1 e^{t(a+ib)} u + \alpha_2 e^{t(a-ib)} \bar{u}.$$

On ne voit pas clairement sur cette expression que $Y(t) \in \mathbb{R}^2$. Mais en écrivant $u = u_1 + iu_2$, avec $u_1, u_2 \in \mathbb{R}^2$, on obtient

$$e^{tA} X_0 = e^{ta} (c_1 \cos(bt) - c_2 \sin(bt)) u_1 + e^{ta} (c_1 \sin(bt) + c_2 \cos(at)) u_2.$$

Cet exemple est tout à fait général. En fait si $\lambda_j = a_j + ib_j$ est une valeur propre de $A \in \mathcal{M}_n(\mathbb{R})$, alors $\bar{\lambda}_j$ aussi et le bloc e^{tT_j} associé à $G_{\lambda_j} \oplus G_{\bar{\lambda}_j}$ s'écrit, dans une base convenable

$$e^{tT_j} = e^{a_j t} \begin{pmatrix} R_j & tR_j & \frac{t^2}{2!}R_j & \dots & \frac{t^{m-1}}{(m-1)!}R_j \\ 0 & R_j & tR_j & \dots & \frac{t^{m-2}}{(m-2)!}R_j \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ & & & & tR_j \\ 0 & \dots & & & R_j \end{pmatrix},$$

où $R_j = \begin{pmatrix} \cos b_j t & -\sin b_j t \\ \sin b_j t & \cos b_j t \end{pmatrix}$. On pourra retenir la

Proposition 3.2.2 Soit $A \in \mathcal{M}_n(\mathbb{R})$. Les solutions à valeurs réelles du système différentiel $Y' = AY$ s'écrivent comme combinaisons linéaires de fonctions de la forme

$$t \mapsto e^{ta_j} \cos(b_j t) t^k u_{j,k}, \quad t \mapsto e^{ta_j} \sin(b_j t) t^k v_{j,k},$$

où $\lambda_j = a_j + ib_j$ est une valeur propre de A , $k \leq m_j - 1$ un entier et $u_{j,k}, v_{j,k} \in \mathbb{R}^n$.

3.2.3 Variation de la constante

Comme pour les équations scalaires, il existe un algorithme pour trouver UNE solution $Y_1(t)$ de l'équation (3.4). Il consiste à chercher une solution $Y_1(t)$ de la forme

$$Y_1(t) = e^{tA} U(t).$$

En formant l'équation que doit satisfaire Y_1 , on obtient

$$e^{tA} U'(t) = B(t),$$

ou encore (attention : le produit des matrices n'est pas commutatif),

$$U'(t) = e^{-tA} B(t).$$

On peut donc choisir

$$U(t) = \int_{t_0}^t e^{-sA} B(s) ds,$$

et on obtient comme solution

$$Y_1(t) = e^{tA} \int_{t_0}^t e^{-sA} B(s) ds.$$

Bien entendu, on est encore confronté au problème du calcul effectif de $e^{-tA} B(t)$. Là encore, la meilleure solution est d'écrire $B(t)$ dans la base $\mathcal{B} = \{V_1, \dots, V_n\}$ dans laquelle A s'écrit sous forme Dunford (Jordan) : $B(t) = \alpha_1(t)V_1 + \dots + \alpha_n(t)V_n$. On a alors comme ci-dessus

$$e^{-tA} B(t) = \alpha_1(t) e^{-tA} V_1 + \dots + \alpha_n(t) e^{-tA} V_n.$$

3.3 Systèmes linéaires à coefficients variables

On a vu que les solutions du système linéaire homogène à coefficients constants $Y' = AY$ sont les fonctions de la forme $t \mapsto e^{tA}X_0$, où X_0 est la valeur de la solution en $t = 0$, et l'on peut calculer explicitement $e^{tA}X_0$. Il ne persiste qu'assez peu de choses de l'étude précédente dans le cas des systèmes homogènes à coefficients variables $Y' = A(t)Y$. En particulier, on ne doit pas s'attendre à avoir une expression explicite des solutions. Cependant, la linéarité de l'équation a des conséquences qu'il est important de connaître, et que l'on décrit ci-dessous.

3.3.1 Résolvante

On considère l'équation différentielle linéaire homogène

$$(3.7) \quad Y' = A(t)Y,$$

où $t \mapsto A(t)$ est une fonction continue d'un intervalle ouvert $I \subset \mathbb{R}$ dans $\mathcal{M}_n(\mathbb{R})$. Dans ce cadre on a admis (et on démontrera plus tard) que, pour tout $(t, X_0) \in I \times \mathbb{R}^n$, le problème de Cauchy

$$(3.8) \quad \begin{cases} Y' = A(t)Y, \\ Y(t_0) = X_0, \end{cases}$$

admet une unique solution maximale, qui est globale. On sait aussi que l'ensemble \mathcal{S} des solutions maximales de l'équation $Y' = A(t)Y$ est un espace vectoriel de dimension n , et on a noté $\Phi_s : \mathcal{S} \rightarrow \mathbb{R}^n$ l'application linéaire bijective qui à une solution Y de l'équation associe sa valeur $Y(s)$ en s .

Définition 3.3.1 On appelle résolvante de l'équation différentielle (3.7), la fonction de $I \times I$ dans $\mathcal{M}_n(\mathbb{R})$ définie par

$$R(t, t_0) = \Phi_t \circ (\Phi_{t_0})^{-1}.$$

Autrement dit, $R(t, t_0)$ est l'application linéaire qui à un vecteur $X \in \mathbb{R}^n$ associe la valeur en t de la solution de l'équation différentielle qui vaut X à l'instant t_0 .

Cette définition a un sens même si $t \mapsto A(t) = A$ est une matrice constante. Dans ce cas $(\Phi_s)^{-1} = e^{-sA}$, et $R(t, t_0) = e^{(t-t_0)A}$.

On rassemble dans la proposition suivante les propriétés essentielles de la résolvante. Cette proposition est à rapprocher de la Proposition 3.1.5 où les mêmes propriétés sont établies pour $t \mapsto e^{(t-t_0)A}$.

Proposition 3.3.2 Soit $R : I \times I \rightarrow \mathcal{M}_n(\mathbb{R})$ la résolvante de l'équation différentielle (3.7). On a

- i) $\forall t \in I, R(t, t) = Id.$
- ii) $\forall t_0, t_1, t_2 \in I, R(t_2, t_1)R(t_1, t_0) = R(t_2, t_0).$
- iii) $\forall (t, t_0) \in I, \partial_t(R(t, t_0)) = A(t)R(t, t_0).$

Preuve: Les deux premières assertions découlent directement de la Définition 3.3.1, et l'on se concentre sur (iii). Soit $X_0 \in \mathbb{R}^n$ et $Y(t) = R(t, t_0)X_0$ la solution du problème de Cauchy (3.8). On a

$$\partial_t(R(t, t_0))X_0 = \partial_t(R(t, t_0)X_0) = Y'(t) = A(t)Y(t) = A(t)R(t, t_0)X_0.$$

Ce calcul étant valable pour tout $X_0 \in \mathbb{R}^n$, on a bien la propriété annoncée. \square

On pourra noter que $R(t, t_0)$ est une matrice inversible. En effet (i) et (ii) donnent

$$R(t, t_0)R(t_0, t) = Id = R(t_0, t)R(t, t_0),$$

donc $R(t, t_0)^{-1} = R(t_0, t).$

Il est également utile de noter que les propriétés (i) et (iii) peuvent se résumer en disant que $R(t, t_0)$ est une solution du problème de Cauchy dans $\mathcal{M}_n(\mathbb{R}) = \mathbb{R}^{n^2}$

$$(3.9) \quad \begin{cases} M' = A(t)M, \\ M(t_0) = Id, \end{cases}$$

On sait d'ailleurs que ce problème admet une unique solution. En particulier si $Y_1, Y_2 \dots Y_n$ sont n solutions respectives dans \mathbb{R}^n des n problèmes de Cauchy

$$(3.10) \quad \begin{cases} Y' = A(t)Y, \\ Y(t_0) = e_j, \end{cases}$$

où les e_j sont les vecteurs de la base canonique de \mathbb{R}^n , il est simple de voir que la matrice $M(t)$ dont les vecteurs colonnes sont les $Y_j(t)$ vérifie (3.9). Autrement dit cette matrice n'est autre que $R(t, t_0).$

► Insistons encore : il n'y a pas de procédé général permettant de calculer $R(t, t_0)$. On a cependant le résultat suivant, qui n'a pas réellement d'intérêt en lui-même et peut être considéré comme un exercice.

Proposition 3.3.3 Si $A(t_1)A(t_2) = A(t_2)A(t_1)$ pour tous $t_1, t_2 \in I$, alors

$$R(t, t_0) = \exp \left(\int_{t_0}^t A(s)ds \right).$$

Preuve: On note $M(t) = \exp \left(\int_{t_0}^t A(s) ds \right)$, et on calcule

$$M(t+h) = \exp \left(\int_{t_0}^t A(s) ds + \int_t^{t+h} A(s) ds \right).$$

Or, pour tous $a, b, c, d \in I$, on a

$$\begin{aligned} \int_a^b A(s_1) ds_1 \int_c^d A(s_2) ds_2 &= \int_{[a,b] \times [c,d]} A(s_1) A(s_2) ds_1 ds_2 \\ &= \int_{[a,b] \times [c,d]} A(s_2) A(s_1) ds_1 ds_2 = \int_c^d A(s_2) ds_2 \int_a^b A(s_1) ds_1. \end{aligned}$$

Donc, grâce à la Proposition 3.1.7,

$$M(t+h) = \exp \left(\int_t^{t+h} A(s) ds \right) M(t).$$

Or la formule de Taylor-Young donne

$$\exp \left(\int_t^{t+h} A(s) ds \right) = Id + hA(t) + o(h),$$

donc

$$M(t+h) - M(t) = hA(t)M(t) + o(h)M(t),$$

et

$$\lim_{h \rightarrow 0} \frac{M(t+h) - M(t)}{h} = A(t)M(t).$$

Autrement dit, $t \mapsto M(t)$ est une solution de l'équation différentielle $Y' = A(t)Y$ dans $\mathcal{M}_n(\mathbb{R}) = \mathbb{R}^{n^2}$, et comme $M(t_0) = Id$, le théorème de Cauchy-Lipschitz (unicité) donne bien $M(t) = R(t, t_0)$. \square

3.3.2 Système fondamental de solutions - Wronskien

On considère toujours l'équation différentielle linéaire homogène à coefficients variables (3.7). Si l'on connaît la résolvante $t \mapsto R(t, t_0)$, on a à disposition un système de n solutions de l'équation différentielle. En effet, notant $(e_j)_{j=1,n}$ la base canonique de \mathbb{R}^n , la fonction $Y_j(t) = R(t, t_0)e_j$ est la solution vérifiant $Y_j(t_0) = e_j$. Autrement dit, les vecteurs colonnes de $R(t, t_0)$ sont n solutions de l'équation différentielle. Puisque l'espace des solutions est de dimension n , on peut se demander si ces n fonctions en forment une base. Si c'est le cas, on dit que les fonctions Y_j forment un système fondamental de solutions. Plus précisément :

Définition 3.3.4 Soient Y_1, Y_2, \dots, Y_n des solutions de l'équation différentielle linéaire homogène $Y' = A(t)Y$. On dit que $\{Y_1, Y_2, \dots, Y_n\}$ est un système fondamental de solution lorsque les Y_j forment un système libre de fonctions, i.e.

$$\sum_{j=1}^n a_j Y_j(t) = 0 \text{ pour tout } t \in I \Rightarrow a_1 = a_2 = \dots = a_n = 0.$$

Attention! Deux fonction Y_1 et Y_2 peuvent être indépendantes même s'il existe un (et même plusieurs) t où $Y_1(t) = Y_2(t)$ par exemple. C'est le cas des fonction $Y_1(t) = \begin{pmatrix} 1 \\ \cos t \end{pmatrix}$ et $Y_2(t) = \begin{pmatrix} \sin t \\ 0 \end{pmatrix}$ bien que $Y_1(\pi/2) = Y_2(\pi/2)$.

Définition 3.3.5 Soient Y_1, Y_2, \dots, Y_n des solutions de l'équation différentielle linéaire homogène $Y' = A(t)Y$. On appelle wronskien des fonctions $Y_j, j = 1 \dots n$, la fonction

$$t \mapsto \mathcal{W}(t) = \det |Y_1(t) \dots Y_n(t)|.$$

Proposition 3.3.6 Soit $\mathcal{W}(t) = \det R(t, t_0)$. On a

$$\mathcal{W}(t) = \exp\left(\int_{t_0}^t \text{Tr } A(s) ds\right).$$

Preuve: On a

$$\mathcal{W}(t+h) = \det R(t+h, t) \mathcal{W}(t).$$

Or la formule de Taylor-Young à l'ordre 1 donne $R(t+h, t) = Id + hA(t) + o(h)$, donc

$$\mathcal{W}(t+h) = \det(Id + hA(t) + o(h)) \mathcal{W}(t) = (1 + h \text{Tr } A(t) + o(h)) \mathcal{W}(t),$$

et \mathcal{W} est la solution de l'équation différentielle

$$\mathcal{W}' = (\text{Tr } A(t)) \mathcal{W},$$

qui vérifie $\mathcal{W}(t_0) = 1$. □

Par conséquent, les vecteurs colonnes de la résolvante forment un système fondamental de solutions de l'équation : le déterminant de la résolvante - leur wronskien- n'est jamais nul.

3.3.3 Méthode de variation de la constante

Pour les équations homogènes à coefficients non-constants, on a vu que la résolvante $R(t, t_0)$ joue le même rôle que la fonction $t \mapsto e^{(t-t_0)A}$. Pour les équations inhomogènes, on peut aussi l'utiliser pour obtenir une solution particulière de l'équation. On reprend la méthode de variation de la constante. On veut donc trouver une solution de l'équation

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

On sait que les solutions de l'équation homogène associée s'écrivent $R(t, t_0)X_0$, et l'on cherche une solution sous la forme

$$Y_1(t) = R(t, t_0)U(t).$$

En formant l'équation que doit satisfaire Y_1 , on obtient

$$R(t, t_0)U'(t) = B(t),$$

ou encore, puisque $R(t, t_0)$ est inversible d'inverse $R(t_0, t)$,

$$U'(t) = R(t_0, t)B(t).$$

On peut donc choisir

$$U(t) = \int_{t_0}^t R(t_0, s)B(s)ds,$$

et on obtient comme solution

$$Y_1(t) = R(t, t_0) \int_{t_0}^t R(t_0, s)B(s)ds.$$

Chapitre 4

Existence et unicité

On attaque maintenant la preuve du théorème de Cauchy-Lipschitz. On va en donner plusieurs variantes. Il s'agit de donner des conditions suffisantes pour l'existence et l'unicité locale de la solution du problème de Cauchy

$$\begin{cases} Y' = F(t, Y), \\ Y(t_0) = X_0, \end{cases}$$

où $F : I \times U \rightarrow \mathbb{R}^m$ est continue, I étant un intervalle ouvert de \mathbb{R} contenant t_0 et U un ouvert de \mathbb{R}^m contenant X_0 .

4.1 Le cas linéaire

On commence par le cas linéaire. D'une part parce que l'on a admis - et utilisé - le résultat correspondant depuis plusieurs cours, et d'autre part parce que la démonstration dans ce cadre, tout en étant plus simple que dans le cas général, en fait apparaître les idées principales. On s'intéresse donc au problème de Cauchy

$$(4.1) \quad \begin{cases} Y' = A(t)Y + B(t), \\ Y(t_0) = X_0, \end{cases}$$

où $A : I \rightarrow \mathcal{M}_n(\mathbb{C})$ et $B : I \rightarrow \mathbb{C}^n$ sont des fonctions continues sur l'intervalle ouvert $I \subset \mathbb{R}$, et $X_0 \in \mathbb{C}^n$ est donné.

4.1.1 Formulation intégrale

On commence par écrire le problème (4.1) sous une forme équivalente, dite forme intégrale, qui s'avère souvent être plus maniable.

Proposition 4.1.1 La fonction $Y : J \rightarrow \mathbb{R}^m$ est une solution de (4.1) sur l'intervalle ouvert $J \subset I$ si et seulement si pour tout $t \in J$, on a

$$(4.2) \quad Y(t) = X_0 + \int_{t_0}^t A(s)Y(s) + B(s)ds.$$

Preuve: Si Y est une solution de (4.1), on a pour tout $s \in J$

$$Y'(s) = A(s)Y(s) + B(s).$$

L'équation (4.2) s'obtient alors en intégrant cette égalité entre t_0 et t , sachant qu'on a aussi $Y(t_0) = X_0$.

Réciproquement, si (4.2) est vérifiée, la fonction Y est automatiquement \mathcal{C}^1 sur J puisque primitive de la fonction continue $s \mapsto A(s)Y(s) + B(s)$, et on obtient $Y'(t) = A(t)Y(t) + B(t)$ en dérivant. \square

4.1.2 Schéma de Picard

On considère alors la suite (Y_n) de fonctions définies par la relation de récurrence - appelée schéma de Picard :

$$(4.3) \quad \begin{cases} Y_0(t) = X_0, \\ Y_{n+1}(t) = X_0 + \int_{t_0}^t A(s)Y_n(s) + B(s)ds. \end{cases}$$

L'idée que l'on a en tête est de montrer que cette suite de fonctions est bien définie, et qu'elle converge uniformément vers une fonction Y sur J , que l'on veut être de classe \mathcal{C}^1 , sur un certain intervalle J contenant t_0 . Si c'est le cas, on aura, pour $t \in J$,

$$Y(t) = \lim_{n \rightarrow +\infty} Y_{n+1}(t) = X_0 + \int_{t_0}^t \lim_{n \rightarrow +\infty} A(s)Y_n(s) + B(s)ds = X_0 + \int_{t_0}^t A(s)Y(s) + B(s)ds.$$

Donc (J, Y) sera bien une solution de (4.2).

Remarque 4.1.2 Pour l'équation différentielle scalaire $y' = ay$, où $a \in \mathbb{R}^*$. Si l'on cherche la solution qui vérifie $y(0) = x_0$, les premières fonctions de la suite fournie par le schéma de Picard sont

$$\begin{aligned} y_0(t) &= x_0, \\ y_1(t) &= x_0 + \int_0^t ay_0(s)ds = x_0(1 + at), \\ &\vdots \\ y_n(t) &= x_0\left(1 + at + \frac{(at)^2}{2!} + \cdots + \frac{(at)^n}{n!}\right) \\ &\vdots \end{aligned}$$

On sait que la suite (y_n) converge vers $y : t \mapsto x_0 e^{at}$ qui est bien la solution du problème.

Pour simplifier les notations, on va supposer que $t_0 = 0$, et on travaille désormais dans un intervalle de la forme $[0, T]$ où la seule limitation concernant $T > 0$ est que $[0, T] \subset I$. Il est clair que si Y_n est \mathcal{C}^1 sur $[0, T]$, alors la fonction Y_{n+1} est bien définie et elle-même \mathcal{C}^1 sur $[0, T]$. Puisque que c'est le cas de Y_0 , on a démontré par récurrence que la suite (Y_n) est bien définie par (4.5), et que les fonctions de cette suite sont \mathcal{C}^1 sur $[0, T]$.

On pose $W_n = Y_n - Y_{n-1}$, pour tout $n \geq 1$. On a tout d'abord, pour tout $t \in [0, T]$,

$$\|W_1(t)\| \leq \int_0^t \|A(s)X_0 + B(s)\| ds \leq Mt,$$

où

$$M = \sup_{s \in [0, T]} \|A(s)X_0 + B(s)\| < +\infty$$

puisque A et B sont continue sur le compact $[0, T]$. Par récurrence, on obtient alors pour tout $n \geq 1$, pour tout $t \in [0, T]$,

$$\|W_n(t)\| \leq A^{n-1} M \frac{t^n}{n!},$$

où $A = \sup_{s \in [0, T]} \|A(s)\|$. En effet

$$\|W_{n+1}(t)\| \leq \int_0^t \|A(s)W_n(s)\| ds \leq A \int_0^t \|W_n(s)\| ds.$$

En particulier, la série de fonctions $\sum_{n \geq 0} W_n(t)$ est normalement convergente, donc converge uniformément sur $[0, T]$. Or

$$\sum_{n=0}^N W_n(t) = Y_{N+1}(t) - Y_0(t),$$

donc la suite (Y_n) converge uniformément, et sa limite Y est continue sur $[0, T]$. On a vu que dans ces conditions, $([0, T], Y)$ est une solution de (4.2).

On a donc démontré l'existence d'une solution du problème de Cauchy (4.2) sur tout intervalle de la forme $[0, T]$ inclus dans I . On peut démontrer de la même manière l'existence d'une solution sur tout intervalle de la forme $[-T, 0]$ inclus dans I , et finalement sur I puisque, notant $I =]a, b[$, on a

$$I = \cup_{k > 0} [a + \frac{1}{k}, b - \frac{1}{k}].$$

4.1.3 Unicité

Il reste à montrer l'unicité de la solution de (4.2). On reprend la même idée : si $([0, T], Y_1)$ et $([0, T], Y_2)$ sont deux solutions de (4.2) (toujours avec $t_0 = 0$ pour simplifier), on a

$$\|Y_1(t) - Y_2(t)\| \leq \int_0^t \|A(s)(Y_1(s) - Y_2(s))\| ds \leq AMt,$$

où $M = \sup_{s \in [0, T]} \|Y_1(s) - Y_2(s)\|$. Par récurrence on obtient alors, pour tout $n \geq 1$,

$$\|Y_1(t) - Y_2(t)\| \leq M \frac{A^n t^n}{n!}.$$

Comme le membre de droite tend vers 0 quand $n \rightarrow +\infty$, on obtient bien $Y_1 = Y_2$ sur $[0, T]$.

On a donc démontré le

Théoreme 4.1.3 Soit $A : I \rightarrow \mathcal{M}_n(\mathbb{C})$ et $B : I \rightarrow \mathbb{C}^n$ des fonctions continues sur l'intervalle ouvert $I \subset \mathbb{R}$, et $X_0 \in \mathbb{C}^n$. Le problème de Cauchy

$$\begin{cases} Y' = A(t)Y + B(t), \\ Y(t_0) = X_0, \end{cases}$$

admet une unique solution maximale, qui est globale.

4.2 Le cas autonome

Pour mettre en évidence une difficulté qui n'apparaît pas dans le cas linéaire, tout en restant dans un cadre assez simple, on s'intéresse maintenant au cas des équations différentielles autonomes

$$Y' = F(Y),$$

où la fonction F ne dépend pas de t . Encore une fois, on veut prouver un résultat d'existence et d'unicité locale pour le problème de Cauchy associé

$$(4.4) \quad \begin{cases} Y' = F(Y), \\ Y(t_0) = X_0, \end{cases}$$

où $F : U \rightarrow \mathbb{R}^m$ est une fonction continue de l'ouvert U de \mathbb{R}^m dans \mathbb{R}^m , et $X_0 \in U$ est donné.

Cette fois (J, Y) est solution de (4.4) si et seulement si

i) Y est continue sur J , et $Y(t) \in U$ pour tout $t \in J$.

ii) Pour tout $t \in J$, on a

$$Y(t) = X_0 + \int_{t_0}^t F(Y(s)) ds.$$

On veut définir une suite (Y_n) de fonctions par la relation de récurrence

$$(4.5) \quad \begin{cases} Y_0(t) = X_0, \\ Y_{n+1}(t) = X_0 + \int_{t_0}^t F(Y_n(s)) ds. \end{cases}$$

On est alors confronté au problème suivant : la fonction F n'étant continue que sur U , la suite (Y_n) ne sera bien définie qu'à condition que pour chaque n , $Y_n(s)$ reste dans U pour tout $s \in [t_0, t] \subset J$. Cela conduit naturellement à imposer que la fonction F soit Lipschitzienne.

4.2.1 Fonctions Lipschitziennes

Définition 4.2.1 Soit $F : U \rightarrow \mathbb{R}^m$ une fonction définie sur l'ouvert U de \mathbb{R}^m et $k > 0$ un réel. On dit que F est lipschitzienne de rapport k sur U lorsque

$$\forall x, y \in U, \|F(x) - F(y)\| \leq k\|x - y\|.$$

On peut visualiser cette propriété pour les fonctions d'une variable : une fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ est lipschitzienne de rapport k lorsque les cordes tracées entre le point $(x, F(x))$ et un autre $(y, F(y))$ de la courbe représentative de F ont toutes une pente comprise en $-k$ et k .

Par exemple la fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ définie par $F(x) = \sin x$ est lipschitzienne de rapport 1 ;

$$|\sin x - \sin y| = |(x - y) \int_0^1 \cos(ty + (1 - t)x) dt| \leq |x - y|.$$

Par contre la fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ définie par $F(x) = x^2$ n'est pas lipschitzienne : puisque $F(x) - F(y) = x^2 - y^2 = (x - y)(x + y)$, la majoration $|F(x) - F(y)| \leq k|x - y|$ est mise en défaut dès que $|x + y| > k$. On introduit donc une notion un peu moins exigeante.

4.2.2 Fonctions localement Lipschitziennes

Définition 4.2.2 Soit $F : U \rightarrow \mathbb{R}^m$ une fonction définie sur l'ouvert U de \mathbb{R}^m . On dit que F est localement lipschitzienne sur U lorsque, pour tout $x_0 \in U$, il existe $r_0 > 0$ et $k_0 > 0$ tels que

$$\forall x, y \in B(x_0, r_0), \|F(x) - F(y)\| \leq k_0\|x - y\|.$$

Cette fois, on dispose d'une classe plus grande de fonctions. En particulier :

Proposition 4.2.3 Si $F : U \rightarrow \mathbb{R}^m$ est \mathcal{C}^1 , alors F est localement lipschitzienne sur U .

Preuve: Soit $x_0 \in U$. Puisque U est ouvert, il existe $r_0 > 0$ tel que $B(x_0, r_0) \subset U$. On note $Q = B(x_0, r_0/2) \subset U$, et $M \geq 0$ la borne supérieure de la fonction continue dF sur le compact \overline{Q} . Si $x, y \in Q$, le segment $[x, y] = \{ty + (1-t)x, t \in [0, 1]\}$ est inclus dans le convexe Q , et l'on peut considérer la fonction f définie sur $[0, 1]$ par

$$f(t) = F(ty + (1-t)x).$$

Cette fonction est \mathcal{C}^1 et l'on a

$$F(y) - F(x) = f(1) - f(0) = \int_0^1 g'(t)dt = \int_0^1 dF((ty + (1-t)x)) \cdot (y - x)dt$$

d'où

$$\|F(y) - F(x)\| \leq \|y - x\| \int_0^1 \|dF((ty + (1-t)x))\|dt \leq M\|y - x\|,$$

ce qui montre que F est Lipschitzienne de rapport M sur Q . □

Réciproquement, on voit facilement que toute fonction localement lipschitzienne est continue, mais il existe des fonctions localement lipschitziennes qui ne sont pas \mathcal{C}^1 : par exemple $x \mapsto |x|$ est lipschitzienne de rapport 1 sur \mathbb{R} mais n'est pas dérivable en 0.

4.2.3 Existence dans le cas autonome

On démontre l'existence d'une solution (J, Y) du problème (4.4) en supposant que la fonction $F : U \rightarrow \mathbb{R}^m$ est localement Lipschitzienne. Encore une fois, on prend $t_0 = 0$ pour simplifier les notations.

Puisque U est ouvert, il existe $r_0 > 0$ tel que $B(X_0, r_0) \subset U$. Puisque F est localement lipschitzienne, quitte à diminuer r_0 , il existe $k_0 > 0$ tel que F est lipschitzienne de rapport k_0 sur $B(X_0, r_0)$.

Soit alors $Q = B(X_0, r_0/2)$. Sur le compact \overline{Q} , la fonction continue F est bornée, et l'on pose

$$M = \sup\{\|F(X)\|, X \in \overline{Q}\} < +\infty.$$

Etape 1 : Soit $T = \frac{r_0}{2M}$. On démontre par récurrence que le schéma de Picard définit bien une suite de fonctions continues sur $[-T, T]$. On considère la propriété :

$\mathcal{P}(n)$: La fonction Y_n est définie et continue sur $[-T, T]$, et

$$\sup_{t \in [-T, T]} \|Y_n(t) - X_0\| \leq \frac{r_0}{2}.$$

La propriété $\mathcal{P}(0)$ est vraie pour n'importe quelle valeur de T , puisque Y_0 est la fonction constante, égale à X_0 .

Supposons que $\mathcal{P}(n)$ soit vraie. Puisque Y_n est continue sur $[-T, T]$ à valeurs dans \bar{Q} , $F \circ Y_n$ est continue sur $[-T, T]$, et Y_{n+1} est bien définie et continue sur $[-T, T]$. De plus

$$\|Y_{n+1}(t) - X_0\| = \left\| \int_0^t F(Y_n(s)) ds \right\| \leq \int_0^t \|F(Y_n(s))\| ds \leq MT,$$

puisque $T \leq \frac{r_0}{2M}$, $\mathcal{P}(n+1)$ est vraie.

La propriété est héréditaire, et puisqu'elle est vraie pour $n = 0$, elle est vraie pour tout $n \in \mathbb{N}$.

Etape 2 : Comme dans le cas linéaire, on montre que la suite (Y_n) converge uniformément sur $[-T, T]$. On a d'abord, pour $t > 0$,

$$\|Y_1(t) - Y_0(t)\| \leq \int_0^t \|F(Y_0(s))\| ds \leq Mt$$

Puis, en utilisant le fait que F est k_0 -lipschitzienne sur \bar{Q} ,

$$\|Y_2(t) - Y_1(t)\| \leq \int_0^t \|F(Y_1(s)) - F(Y_0(s))\| ds \leq k_0 \int_0^t \|Y_1(s) - Y_0(s)\| ds \leq Mk_0 \frac{t^2}{2}.$$

On en déduit par récurrence que

$$\|Y_n - Y_{n-1}\|_\infty \leq M \frac{(k_0 T)^n}{n!},$$

et on conclut comme dans le cas linéaire.

Etape 3 : On montre que la fonction Y est une solution du problème de Cauchy (4.4).

On a, pour tout n ,

$$Y_{n+1}(t) = X_0 + \int_{t_0}^t F(Y_n(s)) ds,$$

et donc, en passant à la limite, en utilisant la convergence uniforme de la suite $(F \circ Y_n)$,

$$Y(t) = X_0 + \lim_{n \rightarrow +\infty} \int_{t_0}^t F(Y_n(s)) ds = X_0 + \int_{t_0}^t \lim_{n \rightarrow +\infty} F(Y_n(s)) ds = X_0 + \int_{t_0}^t F(Y(s)) ds.$$

D'autre part Y est continue, et $Y(t) = \lim Y_n(t) \in \bar{Q} \subset U$, donc Y vérifie les deux conditions (i) et (ii) ci-dessus. La fonction Y est bien une solution du problème de Cauchy (4.4) sur $[-T, T]$.

4.3 Le cas général

Cette fois on considère le problème de Cauchy pour une équation différentielle d'ordre 1

$$(4.6) \quad \begin{cases} Y' = F(t, Y), \\ Y(t_0) = X_0, \end{cases}$$

où $F : I \times U \rightarrow \mathbb{R}^m$ est une fonction continue de l'ouvert $I \times U$ de $\mathbb{R} \times \mathbb{R}^m$ dans \mathbb{R}^m , et $X_0 \in U$ est donné.

Pour éviter de rencontrer des difficultés techniques à chaque étape du raisonnement, qui pourraient faire perdre de vue la simplicité de la preuve, on va formaliser un peu les idées mises en oeuvre dans les deux cas précédents : à chaque fois, on a fait apparaître la solution du problème de Cauchy comme point fixe de l'application

$$\Phi : Y \in \mathcal{C}^0(J) \mapsto X_0 + \int_{t_0}^t F(s, Y(s)) ds.$$

4.3.1 Le théorème du point fixe

Définition 4.3.1 Soit E un espace vectoriel normé, et $\Phi : E \rightarrow E$ une application. On dit que Φ est contractante lorsque Φ est Lipschitzienne de rapport < 1 , c'est-à-dire s'il existe une constante $0 < k < 1$ telle que

$$\forall x, y \in E, \quad \|\Phi(x) - \Phi(y)\| \leq k\|x - y\|.$$

Proposition 4.3.2 (Théorème du point fixe de Banach) Soit E un espace de Banach. Si $\Phi : E \rightarrow E$ est contractante, alors Φ admet un unique point fixe.

Preuve.— On règle d'abord la question de l'unicité : si Φ admet deux points fixes $x_1 \neq x_2$, on a

$$\|x_1 - x_2\| \leq \|\Phi(x_1) - \Phi(x_2)\| \leq k\|x_1 - x_2\| < \|x_1 - x_2\|,$$

ce qui est absurde.

Soit (x_n) la suite définie par $x_0 \in E$ et $x_{n+1} = \Phi(x_n)$. Si la suite (x_n) converge vers un $x \in E$, on aura, puisque Φ est automatiquement continue,

$$x = \lim x_{n+1} = \lim \Phi(x_n) = \Phi(\lim x_n) = \Phi(x).$$

Autrement dit, la limite éventuelle de la suite (x_n) est un point fixe de Φ . Pour montrer que cette suite converge, on va montrer qu'il s'agit d'une suite de Cauchy. Or, pour $n \geq 1$,

$$\|x_{n+1} - x_n\| = \|\Phi(x_n) - \Phi(x_{n-1})\| \leq k\|x_n - x_{n-1}\| \leq \dots \leq k^n\|x_1 - x_0\|.$$

Donc, pour $q > p > n$,

$$\|x_q - x_p\| \leq \sum_{j=p}^q k^{j-1} \|x_1 - x_0\| \leq \frac{k^n}{1-k} \|x_1 - x_0\|.$$

Puisque $k < 1$, ceci montre que (x_n) est une suite de Cauchy dans l'espace de Banach E , et finalement converge. \square

On aura besoin du raffinement suivant :

Corollaire 4.3.3 Soit $\Phi : E \rightarrow E$ une application sur l'espace de Banach E . Si Φ possède une itérée contractante, alors Φ admet un unique point fixe.

Preuve.— On sait que $\Phi^{\circ m}$ admet un unique point fixe a . Si b est un point fixe de Φ , on a $\Phi^{\circ m}(b) = b$, donc $b = a$ nécessairement, ce qui montre que Φ a au plus un point fixe. De plus

$$\Phi^{\circ m}(\Phi(a)) = \Phi^{\circ(m+1)}(a) = \Phi(\Phi^{\circ m}(a)) = \Phi(a),$$

donc $\Phi(a) = a$ puisque c'est un point fixe de $\Phi^{\circ m}$. \square

4.3.2 Tonneau de sécurité

Pour appliquer le théorème du point fixe à l'application Φ définie par

$$Y \mapsto \Phi(Y) : t \mapsto X_0 + \int_{t_0}^t F(s, Y(s)) ds,$$

il faut d'abord trouver un espace de Banach E , inclus dans $\mathcal{C}^0(I, U)$ tel que $\Phi(E) \subset E$. On cherche E sous la forme

$$E = \mathcal{C}^0([t_0 - T, t_0 + T], \overline{B(X_0, r_0)})$$

Puisque I et U sont ouverts et contiennent t_0 et X_0 , il existe un cylindre $C_0 = [t_0 - T, t_0 + T] \times \overline{B(X_0, r_0)}$ centré en (t_0, X_0) et contenu dans $I \times U$.

Définition 4.3.4 On dit qu'un cylindre $C_T = [t_0 - T, t_0 + T] \times \overline{B(X_0, r_0)} \subset C_0$ centré en (t_0, X_0) , est un tonneau de sécurité pour (4.6) lorsque toute solution éventuelle Y de (4.6) sur $J = [t_0 - T, t_0 + T]$ vérifie

$$\forall t \in J, Y(t) \in \overline{B(X_0, r_0)}.$$

Le fait essentiel est qu'il existe des tonneaux de sécurité. Autrement dit, en ne supposant que la continuité de F , on peut assurer a priori que le graphe d'une éventuelle solution ne s'écarte pas trop vite du point (t_0, X_0) . Prouvons-le : puisque F est continue sur $I \times U$, elle est bornée sur le compact C_0 et on note

$$(4.7) \quad M = \sup\{\|F(t, X)\|, (t, X) \in C_0\} < +\infty.$$

On a alors la

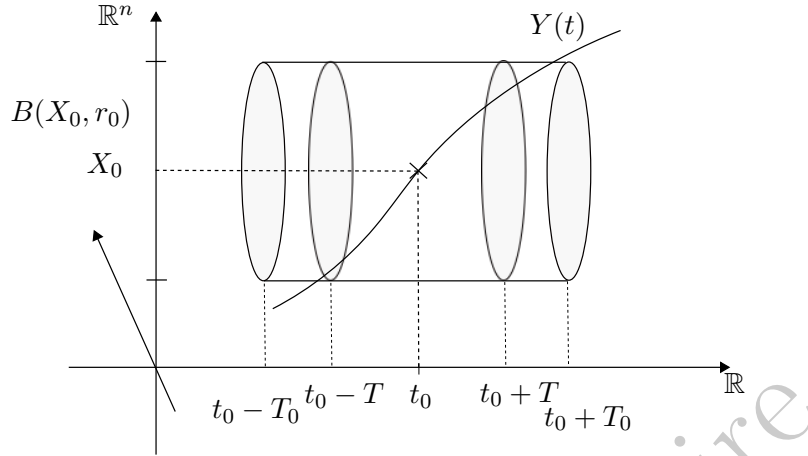


FIGURE 4.1 – Tonneau de sécurité

Proposition 4.3.5 Si $T \leq \min(T_0, \frac{r_0}{M})$, le cylindre C_T est un cylindre de sécurité pour (4.6).

Preuve: On raisonne par contraposée : on suppose que C_T n'est pas un cylindre de sécurité, donc qu'il existe une solution Y de (4.6) qui quitte le cylindre C_T . Soit alors t le plus petit temps dans $[t_0, t_0 + T]$ tel que $Y(t) \notin B(X_0, r_0)$, plus précisément

$$t = \inf\{s \in [t_0, t_0 + T], \|Y(s) - X_0\| > r_0\}.$$

On a $\|Y(s) - X_0\| \leq r_0$ pour tout $s < t$, donc par continuité

$$\|Y(t) - X_0\| = r_0.$$

D'autre part, $Y(s) \in \overline{B(X_0, r_0)}$ pour tout $s \in [t_0, t]$ donc

$$\forall s \in [t_0, t], \|Y'(s)\| \leq \|F(s, Y(s))\| \leq M,$$

et

$$\|Y(t) - X_0\| \leq \int_{t_0}^t \|Y'(s)\| ds \leq M(t - t_0).$$

On a donc nécessairement $T > t - t_0 > \frac{M}{r_0}$. □

4.3.3 Existence et unicité locale

On commence par donner la notion de fonction lipschitzienne adaptée au fait que l'équation (4.6) n'est pas autonome.

Définition 4.3.6 Soit $F : I \times U \rightarrow \mathbb{R}^m$ une fonction continue. On dit que F est localement lipschitzienne par rapport à sa seconde variable lorsque pour tout $(t_0, X_0) \in I \times U$, il existe un cylindre $C = [t_0 - T, t_0 + T] \times \overline{B}(X_0, r_0) \subset I \times U$ et une constante $k_0 = k_0(t_0, X_0)$ telle que

$$\forall (t, X_1), (t, X_2) \in C, \|F(t, X_1) - F(t, X_2)\| \leq k_0 \|X_1 - X_2\|.$$

Le lecteur montrera sans difficulté l'analogue de la Proposition 4.2.3 : si F est continue et admet des dérivées partielles continues par rapport à sa seconde variable $X \in U$, alors F est localement lipschitzienne par rapport à sa seconde variable.

Proposition 4.3.7 (Théorème de Cauchy-Lipschitz) Soit $F : I \times U \rightarrow \mathbb{R}^m$ une fonction continue, localement lipschitzienne par rapport à sa seconde variable. Il existe $T > 0$ tel que le problème de Cauchy (4.6) admet une unique solution dans $[t_0 - T, t_0 + T]$.

Preuve.— Soit $r_0 > 0$ tel que $\overline{B}(X_0, r_0) \subset U$. Soit aussi $0 < T < \frac{r_0}{M}$ tel que $C_T = [t_0 - T, t_0 + T] \times \overline{B}(X_0, r_0)$ soit un tonneau de sécurité pour (4.6). On note $E = \mathcal{C}^0([t_0 - T, t_0 + T], \overline{B}(X_0, r_0))$. Muni de la norme

$$\|Y\|_\infty = \sup_{[t_0 - T, t_0 + T]} \|Y(t)\|,$$

E un espace de Banach.

- Soit alors Φ l'application définie sur E par

$$\forall t \in [t_0 - T, t_0 + T], \Phi(Y)(t) = X_0 + \int_{t_0}^t F(s, Y(s)) ds.$$

La fonction $\Phi(Y)$ est continue sur $[t_0 - T, t_0 + T]$, et, pour tout $t \in [t_0 - T, t_0 + T]$,

$$\|\Phi(Y)(t) - X_0\| \leq \int_{t_0}^t \|F(s, Y(s))\| ds \leq MT \leq r_0.$$

Donc Φ envoie E dans E .

- Montrons que la fonction Φ possède une itérée contractante. Soit $Y, Z \in E$. Quitte à diminuer T et r_0 , on peut supposer que F est k_0 -lipschitzienne par rapport à sa seconde variable dans C_T . On a donc

$$\|\Phi(Y) - \Phi(Z)\| \leq \int_{t_0}^t \|F(s, Y(s)) - F(s, Z(s))\| ds \leq k_0 \int_{t_0}^t \|Y(s) - Z(s)\| ds \leq k_0 |t - t_0| \|Y - Z\|_\infty$$

Par récurrence, on en déduit que, pour tout $p \geq 1$,

$$\|\Phi^p(Y) - \Phi^p(Z)\|_\infty \leq \frac{(k_0)^p}{p!} T^p \|Y(s) - Z(s)\|.$$

Puisque $\frac{(k_0)^p}{p!} T^p \rightarrow 0$ quand $p \rightarrow +\infty$, il existe $p \in \mathbb{N}$ tel que Φ^p est contractante.

Ainsi Φ admet un unique point fixe, qui est l'unique solution de (4.6) sur $[t_0 - T, t_0 + T]$. \square

4.3.4 Solution maximale

On peut déduire assez facilement du théorème d'existence et d'unicité locale le résultat suivant

Corollaire 4.3.8 Sous les conditions du théorème de Cauchy-Lipschitz, le problème (4.6) admet une unique solution maximale.

Preuve.— Soit (J, Y_1) et (J, Y_2) deux solutions de l'équation $Y' = F(t, Y)$ définies sur un même intervalle J . S'il existe $t_0 \in J$ tel que $Y_1(t_0) = Y_2(t_0)$, alors Y_1 et Y_2 coïncident sur J . Sinon l'ensemble $\{t \in J, t > t_0, Y_1(t) \neq Y_2(t)\}$ n'est pas vide, donc admet une borne inférieure \tilde{t}_0 . Par continuité, on a $Y_1(\tilde{t}_0) = Y_2(\tilde{t}_0)$. Mais le théorème d'unicité locale donne l'existence d'un $T > 0$ tel que $Y_1(t) = Y_2(t)$ sur $[\tilde{t}_0 - T, \tilde{t}_0]$, ce qui contredit la définition de \tilde{t}_0 .

Maintenant si (J_1, Y_1) et (J_2, Y_2) sont deux solutions maximales de (4.6), elles coïncident sur $J_1 \cap J_2$. Comme elles sont maximales, on a $J_1 \cap J_2 = J_1 = J_2$. \square

On peut se demander si la taille de l'intervalle d'existence J de la solution maximale de (4.6) peut varier beaucoup pour des données de Cauchy proches les unes des autres. Voilà une réponse rassurante, que nous admettons.

Proposition 4.3.9 Soit $F \in \mathcal{C}^0(I \times U, \mathbb{R}^m)$ avec I un intervalle ouvert de \mathbb{R} , et U un ouvert de \mathbb{R}^m , localement lipschitzienne par rapport à sa seconde variable. Soit K un compact de $I \times U$. Il existe $T_K > 0$ tel que, pour toute donnée initiale $(t_0, X_0) \in K$, la solution maximale de (4.6) est définie au moins sur $]t_0 - T_K, t_0 + T_K[$.

4.4 Existence globale

4.4.1 Un critère d'existence globale

Alors que le théorème de Cauchy-Lipschitz permet d'établir l'existence locale d'une solution, on a vu des cas - en particulier celui des équations linéaires à coefficients constants - où toutes les solutions maximales sont des solutions globales. Voici un critère assez général qui donne une condition suffisante pour que ce soit le cas.

Proposition 4.4.1 Soit $F : I \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ une fonction continue, où I est un intervalle ouvert de \mathbb{R} . On suppose qu'il existe une fonction $k : I \rightarrow \mathbb{R}^+$ continue, telle que, pour tout $t \in I$ fixé, la fonction $X \mapsto F(t, X)$ est $k(t)$ -lipschitzienne sur \mathbb{R}^m . Alors la solution maximale du problème de Cauchy (4.6) est globale.

Preuve.— On reprend la preuve du Théorème de Cauchy-Lipschitz. Soit $[t_0 - T_-, t_0 + T_+]$ un intervalle compact quelconque inclus dans I , et

$$E = \mathcal{C}^0([t_0 - T_-, t_0 + T_+], \mathbb{R}^m).$$

On sait que $(E, \|\cdot\|_\infty)$ est un espace de Banach. Pour Y dans E on note $\Phi(Y)$ la fonction définie par

$$\Phi(Y) : t \mapsto X_0 + \int_{t_0}^t F(s, Y(s)) ds.$$

La fonction $\Phi(Y)$ appartient bien à E , et on va montrer que l'application $\Phi : E \rightarrow E$ admet une itérée contractante.

Soit $K = \max_{t \in [t_0 - T_-, t_0 + T_+]} k(t)$. Puisque $t \mapsto k(t)$ est continue et l'intervalle compact, on a $K < +\infty$. L'application F est K -lipschitzienne par rapport à sa seconde variable sur $[t_0 - T_-, t_0 + T_+]$, et donc, pour $Y, Z \in E$,

$$\|\Phi(Y) - \Phi(Z)\| \leq \int_{t_0}^t \|F(s, Y(s)) - F(s, Z(s))\| ds \leq K|t - t_0| \|Y - Z\|.$$

Par récurrence on obtient

$$\|\Phi^{\circ p}(Y) - \Phi^{\circ p}(Z)\| \leq \frac{K^p |t - t_0|^p}{p!} \|Y - Z\| \leq \frac{K^p (\max(T_+, T_-))^p}{p!} \|Y - Z\|,$$

et l'on peut choisir p assez grand pour que $\frac{K^p (\max(T_+, T_-))^p}{p!} < 1$.

L'application Φ possède donc un point fixe unique, ce qui signifie que le problème de Cauchy 4.6 admet une unique solution sur $[t_0 - T_-, t_0 + T_+]$. Puisque cet intervalle est quelconque dans I , on a démontré la proposition. \square

Exemple 4.4.2 La fonction $F : \mathbb{R} \times \mathbb{R}$ définie par $F(t, x) = \sin(tx)$ vérifie l'hypothèse de la Proposition ci-dessus. On a en effet, en appliquant le théorème des accroissements finis,

$$|\sin(tx_1) - \sin(tx_2)| = |t(x_1 - x_2) \cos(c)| \leq |t| |x_1 - x_2|.$$

On peut donc prendre $k(t) = |t|$, qui est bien une fonction continue.

Cette proposition s'applique aussi dans le cas des systèmes linéaires à coefficients variables, où $F(t, X) = A(t)X + B(t)$, avec $t \mapsto A(t) \in \mathcal{M}_n(\mathbb{C})$ et $t \mapsto B(t) \in \mathbb{C}^n$ continues. On a en effet

$$\|F(t, X_1) - F(t, X_2)\| = \|A(t)\| \|X_1 - X_2\|,$$

et l'on peut prendre $k(t) = \|A(t)\|$, qui est bien continue. On obtient donc le

Corollaire 4.4.3 Toutes les solutions maximales du système linéaire $Y' = A(t)Y + B(t)$, avec A et B continues, sont globales.

On peut aussi se demander si la condition ci-dessus est nécessaire pour qu'il y ait existence globale. La réponse est négative comme le montre l'exemple suivant.

Exemple 4.4.4 Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ la fonction définie par $F(x) = e$ si $x \leq 0$, et $F(x) = x \ln x$ pour $x \geq 0$. La fonction F est continue, mais n'est pas Lipschitzienne au voisinage de $x = 0$. Cependant les solutions maximales de l'équation $y' = F(y)$ sont globales.

4.4.2 Qu'est-ce qui empêche une solution d'être globale ?

On a vu que sous des hypothèses relativement peu contraignantes, le problème de Cauchy

$$(4.8) \quad \begin{cases} Y' = F(t, Y), \\ Y(t_0) = X_0, \end{cases}$$

admet une unique solution maximale (J, Y) . Par contre on sait qu'il existe des cas où cette solution maximale n'est pas globale. On veut comprendre ce qui empêche de prolonger la solution (J, Y) . On pourrait imaginer diverses raisons, comme par exemple le fait que Y cesse d'être dérivable aux extrémités de J . La réponse porte le nom de théorème des bouts : c'est toujours le comportement de Y aux extrémités de J qui pose problème. Voici d'abord un résultat général connu sous le nom de Théorème d'échappement fort.

Proposition 4.4.5 Soit $F \in \mathcal{C}^0(I \times U, \mathbb{R}^m)$, avec I un intervalle ouvert de \mathbb{R} , et U un ouvert de \mathbb{R}^m , localement lipschitzienne par rapport à sa seconde variable. Soit aussi K un compact de U , et $(t_0, X_0) \in K$. Si $(]a, b[, Y)$ est la solution maximale de (4.8), il existe $[t_-, t_+] \subset]a, b[$ tel que

$$\forall t \notin [t_-, t_+], Y(t) \notin K.$$

Autrement dit, il existe un instant t_+ à partir duquel le graphe $\{(t, Y(t)), t \in]a, b[\}$ de la solution quitte le compact K et n'y revient pas. De même, il existe un instant t_- avant lequel ce graphe n'est jamais entré dans K .

Preuve: On montre l'existence de t_+ en raisonnant par l'absurde (celle de t_- se démontre exactement de la même manière). On suppose donc que

$$\forall t_+ \in]a, b[, \exists t > t_+, Y(t) \in K.$$

On peut alors choisir une suite $(t_n) \subset]a, b[$ telle que

$$\lim_{n \rightarrow +\infty} t_n = b, \text{ et } Y(t_n) \in K.$$

Puisque K est compact, on peut supposer (quitte à extraire une sous-suite) que la suite $(Y(t_n))$ converge vers un certain $X_\infty \in K$.

D'après la Proposition 4.3.9, il existe $T_\delta > 0$ tel que pour tout $(t_0, X_0) \in B((b, X_\infty), \delta)$, la solution maximale de (4.8) est définie au moins sur $]t_0 - T_\delta, t_0 + T_\delta[$. Or pour n assez grand, $(t_n, Y(t_n))$ appartient à $B((b, X_\infty), \delta)$, et $t_n + \delta > b$. Donc la solution associée à $(t_n, Y(t_n))$ est un prolongement strict de $(]a, b[, Y)$ ce qui contredit le caractère maximal de cette solution. \square

Le théorème des bouts est une variante de la Proposition 4.4.5 dans le cas où $U = \mathbb{R}^m$. Dans ce cas, lorsque la solution maximale de l'équation différentielle n'est pas globale, c'est qu'elle "explose" avant d'atteindre le bord de I . De manière plus précise

Proposition 4.4.6 Soit $F \in \mathcal{C}^0([a, b[\times \mathbb{R}^m, \mathbb{R}^m)$ avec $a < b$, localement lipschitzienne par rapport à sa seconde variable. Soit $(t_0, X_0) \in]a, b[\times U$ et $(]T_*, T^*[, Y)$ la solution maximale de (4.8). On a l'alternative suivante :

ou bien $T^* = b$,

ou bien $T^* < b$ et $\lim_{t \rightarrow T^*} \|Y(t)\| = +\infty$.

De même

ou bien $T_* = a$,

ou bien $T_* > a$ et $\lim_{t \rightarrow T_*} \|Y(t)\| = +\infty$.

Preuve: Supposons que $T^* < b$. Pour tout $R > 0$, on sait qu'il existe $t_+(R) < T^*$ tel que, pour tout $t > t_+(R)$, $\|Y(t) - X_0\| > 2R$, et donc $\|Y(t)\| \geq R$ pourvu que R soit assez grand ($R > \|X_0\|$). \square

Corollaire 4.4.7 Sous les hypothèses précédentes, si (J, Y) est une solution maximale, alors J est un intervalle ouvert.

Preuve.— Supposons par exemple que $J =]T_*, T^*]$. Puisque Y est continue sur J , on a $\lim_{t \rightarrow T^*} \|Y(t)\| = \|Y(T^*)\| < +\infty$, ce qui contredit la proposition. \square

Ce théorème des bouts sert souvent sous la forme suivante. Soit (J, Y) la solution maximale de (4.8). S'il existe $M > 0$ tel que $\|Y(t)\| \leq M$ pour tout $t \in J$, alors (J, Y) est une solution globale. Par exemple on a la

Proposition 4.4.8 Soit $F \in \mathcal{C}^0(I \times \mathbb{R}^m, \mathbb{R}^m)$ avec I intervalle ouvert de \mathbb{R} , localement lipschitzienne par rapport à sa seconde variable. Si F est bornée, alors les solutions maximales de l'équation $Y' = F(t, Y)$ sont globales.

Preuve: Soit (J, Y) une solution maximale. Pour $t \in J$ on a

$$\|Y'(t)\| \leq \|F(t, Y(t))\| \leq M$$

Donc

$$\|Y(t)\| \leq \left\| \int_{t_0}^t Y'(s) ds \right\| + \|X_0\| \leq M|t - t_0| + \|X_0\|.$$

En particulier, si J est borné, la fonction Y est bornée sur J . D'après le théorème des bouts ci-dessus, on a alors $J = I$. \square

Dans le cas où F croît au plus linéairement par rapport à t , on a le même résultat :

Proposition 4.4.9 Soit $F \in \mathcal{C}^0(I \times \mathbb{R}^m, \mathbb{R}^m)$ avec I intervalle ouvert de \mathbb{R} , localement lipschitzienne par rapport à sa seconde variable. S'il existe A, B tels que $\|F(t, X)\| \leq At + B$ pour tous (t, X) , alors les solutions maximales de l'équation $Y' = F(t, Y)$ sont globales.

Preuve: Soit (J, Y) une solution maximale. Pour $t \in J$ on a

$$\|Y'(t)\| \leq \|F(t, Y(t))\| \leq At + B$$

Donc

$$\|Y(t)\| \leq \left\| \int_{t_0}^t Y'(s) ds \right\| + \|X_0\| \leq \frac{A}{2}|t - t_0|^2 + B|t - t_0| + \|X_0\|.$$

En particulier, si J est borné, la fonction Y est bornée sur J . D'après le théorème des bouts ci-dessus, on a alors $J = I$. \square

4.4.3 Le lemme de Gronwall

Voici un résultat concernant des inéquations différentielles. On peut le voir en particulier comme un moyen de montrer l'existence globale de solutions. Supposons par exemple que $y'(t) \leq 0$. Pour une telle fonction, on sait que $y(t) \leq y(t_0)$ pour tout $t \geq t_0$: il ne peut pas y avoir explosion pour les $t > t_0$. Outre le résultat, on pourra retenir l'idée de la preuve : en changeant de fonction inconnue (par exemple à l'aide d'un facteur intégrant) on se ramène à l'inéquation différentielle $y'(t) \leq 0$.

On commence par une version intégrale.

Proposition 4.4.10 Soit ϕ, g deux fonctions continues sur $[t_0, t_0 + T]$ à valeurs réelles, avec $g \geq 0$, et $a \in \mathbb{R}$. Si

$$\forall t \in [t_0, t_0 + T], \quad \phi(t) \leq a + \int_{t_0}^t g(s)\phi(s)ds,$$

alors

$$\forall t \in [t_0, t_0 + T], \quad \phi(t) \leq a \exp\left(\int_{t_0}^t g(s)ds\right)$$

Preuve: Soit $w(t) = a + \int_{t_0}^t g(s)\phi(s)ds$. On a $w'(t) = g(t)\phi(t)$, donc, puisque $g(t) \geq 0$ pour tout t ,

$$w'(t) \leq g(t)w(t).$$

Soit alors $G(t) = \int_{t_0}^t g(s)ds$.

$$(w(t)e^{-G(t)})' = (w'(t) - g(t)w(t))e^{-G(t)} \leq 0,$$

donc $w(t)e^{-G(t)} \leq w(t_0)e^{-G(t_0)} = a$ et

$$a \exp\left(\int_{t_0}^t g(s)ds\right) = ae^{G(t)} \geq w(t) \geq \phi(t).$$

□

Voici maintenant une version "équation différentielle".

Proposition 4.4.11 Soit $\phi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe C^1 telle que $\phi(0) = 0$ et

$$\forall t \in \mathbb{R}, |\phi'(t)| \leq C + M|\phi(t)|$$

où $C, M > 0$ sont des constantes données. Alors

$$\forall t \in \mathbb{R}, |\phi(t)| \leq \frac{C}{M}(e^{M|t|} - 1).$$

Preuve.— Soit $t \geq 0$. On a

$$\begin{aligned} |\phi(t)| &= |\phi(t) - \phi(0)| = \left| \int_0^t \phi'(s)ds \right| \leq \int_0^t |\phi'(s)|ds \\ &\leq \int_0^t C + M|\phi(s)|ds \leq Ct + M \int_0^t |\phi(s)|ds = w(t). \end{aligned}$$

Soit alors $f(t) = e^{-Mt}w(t)$. On a

$$f'(t) = e^{-Mt}w'(t) - Me^{-Mt}w(t) \leq e^{-Mt}(C + Mw(t)) - Me^{-Mt}w(t) = Ce^{-Mt}.$$

Donc

$$f(t) = \int_0^t f'(s)ds \leq C \int_0^t e^{-Ms}ds \leq \frac{C}{M}(1 - e^{-Mt}).$$

On obtient le résultat en multipliant les deux membres de l'inégalité par e^{Mt} .

Il reste à montrer le résultat pour $t \in \mathbb{R}^-$. □

Exercice 4.4.12 Soit $A : \mathbb{R} \rightarrow \mathcal{M}_n(\mathbb{R})$ une application continue. On suppose qu'il existe $a > 0$ tel que $\|A(t)\| \leq a$ pour tout $t \in \mathbb{R}$. On note (J, Y) la solution maximale du problème de Cauchy

$$\begin{cases} Y'(t) = A(t)Y(t), \\ Y(0) = X_0, \end{cases}$$

où $X_0 \in \mathbb{R}^n$ est donné.

i) Pourquoi peut-on d'affirmer que $J = \mathbb{R}$?

ii) En posant $v(t) = \|Y(t)\|$, où $\|\cdot\|$ désigne la norme euclidienne, démontrer que :

$$\forall t \geq 0, \|Y(t)\| \leq e^{at}\|X_0\|.$$

iii) Démontrer que $\forall t \leq 0, \|Y(t)\| \leq e^{-at}\|X_0\|$.

On a donc, pour tout $t \in \mathbb{R}$, $\|Y(t)\| \leq e^{a|t|}\|X_0\|$.

Chapitre 5

Stabilité et linéarisation

On se pose maintenant la question du comportement asymptotique, c'est à dire pour des temps très long, des solutions d'équations différentielles - en particulier non-linéaires. Pour simplifier l'exposé, on se concentre sur les équations autonomes

$$(5.1) \quad \begin{cases} Y' = F(Y), \\ Y(0) = X_0, \end{cases}$$

où $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$ est supposée continue et localement lipschitzienne, de sorte que ce problème de Cauchy admet une unique solution maximale. On peut par exemple avoir en tête le système prédateurs-proies de Lotka-Volterra

$$\begin{cases} y_1' = ay_1 - cy_1y_2, \\ y_2' = -by_2 + dy_1y_2, \end{cases}$$

que l'on a déjà évoqué. La question que l'on se pose est de prédire, en fonction de la donnée initiale, l'évolution à long terme de chacune des populations de proies $y_1(t)$ et de prédateurs $y_2(t)$, même si on ne peut pas écrire les solutions de ce système explicitement.

Il y a un cas où la réponse est simple : celui où la donnée initiale X_0 est un point d'équilibre pour l'équation.

Définition 5.0.13 On dit que $X_0 \in U$ est un point d'équilibre pour l'équation différentielle $Y' = F(Y)$ lorsque $F(X_0) = 0$.

Dans ce cas en effet, la solution du problème est la fonction constante $Y(t) = X_0$, et le comportement de $Y(t)$ pour des t grands n'est pas difficile à décrire.

Se pose alors naturellement la question de savoir si l'on peut encore prédire le comportement de $Y(t)$ lorsque la condition initiale X_0 est suffisamment proche d'un point d'équilibre \tilde{X}_0 pour l'équation. On peut espérer en particulier que $Y(t)$ reste pour tout temps assez proche de la fonction constante $\tilde{Y} : t \mapsto \tilde{X}_0$.

5.1 Portraits de phase des systèmes linéaires homogènes 2×2 à coefficients constants

Pour les systèmes linéaires homogènes à coefficients constants $Y' = AY$, $X_0 = 0$ est toujours un point fixe. Pour avoir une idée des différents comportements que l'on peut rencontrer, on va examiner l'allure des solutions des systèmes 2×2 homogènes à coefficients constants $Y' = AY$, où $A \in \mathcal{M}_2(\mathbb{R})$.

Pour les équations différentielles scalaires ($m = 1$), il est possible de représenter les courbes intégrales des solutions $t \mapsto y(t)$: ce sont les courbes représentatives de fonctions de \mathbb{R} dans \mathbb{R} . Dès que l'on considère des systèmes, le dessin d'une courbe intégrale est plus difficile : pour $m = 2$ par exemple il s'agit de représenter dans \mathbb{R}^3 le graphe d'une fonction de \mathbb{R} dans \mathbb{R}^2 . On préfère dans ce cas oublier la description de la courbe au cours du temps, et ne regarder que la trace obtenue sur \mathbb{R}^2 . De manière générale

Définition 5.1.1 On appelle orbite de l'équation $Y' = F(Y)$ toute courbe \mathcal{O} de \mathbb{R}^m définie par

$$\mathcal{O} = \{Y(t), t \in \mathbb{R}\},$$

où Y est une solution de l'équation.

Autrement dit, une orbite peut être vue comme la projection sur \mathbb{R}^m de la courbe intégrale d'une solution. C'est aussi l'image dans \mathbb{R}^m de la courbe paramétrée $t \mapsto Y(t)$.

Proposition 5.1.2 Deux orbites \mathcal{O}_1 et \mathcal{O}_2 distinctes d'une même équation différentielle autonome ne se coupent jamais.

Preuve: Supposons qu'il existe $M \in \mathcal{O}_1 \cap \mathcal{O}_2$, et notons (I_1, Y_1) et (I_2, Y_2) les deux solutions auxquelles \mathcal{O}_1 et \mathcal{O}_2 sont associées. Il existe deux réels t_1 et t_2 tels que

$$Y_1(t_1) = Y_2(t_2) = M.$$

Soit alors $T = t_1 - t_2$ et $\tilde{Y}_2 = \tau_T Y_2$ la fonction définie sur $\tilde{I}_2 = T + I_2$ par $\tilde{Y}_2(t) = Y_2(t - (t_1 - t_2))$. Parce que l'équation est autonome, $(\tilde{I}_2, \tilde{Y}_2)$ est une solution aussi : pour tout $t \in \tilde{I}_2$, on a $t - T \in I_2$ et

$$\tilde{Y}_2'(t) = Y_2'(t - T) = F(Y_2(t - T)) = F(\tilde{Y}_2(t)).$$

De plus $\tilde{Y}_2(t_1) = Y_2(t_2) = Y_1(t_1)$. Donc $\tilde{Y}_2 = Y_1$, et \mathcal{O}_2 , qui est aussi la courbe associée à \tilde{Y}_2 , coïncide avec \mathcal{O}_1 . \square

Avec cette proposition, on comprend que l'on aura une bonne idée de l'allure des orbites de toutes les solutions à partir du dessin de quelques unes seulement. Ce genre de dessin porte le nom de portrait de phase pour l'équation différentielle.

On retiendra aussi la

Proposition 5.1.3 Si Y est une solution de l'équation autonome $Y' = F(Y)$, alors la fonction $\tilde{Y} : t \mapsto Y(t - \tau)$ est aussi une solution, et son orbite est la même que celle de Y .

On distingue quatre cas disjoints :

Au moins une des valeurs propres est nulle

- Si les deux valeurs propres sont nulles, et A n'est pas diagonalisable, A semblable à la matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dans une base adéquate, le système s'écrit $x'_1 = x_2, x'_2 = 0$, donc $x_2 = c_2$ et $x_1 = c_2 t + c_1$.
- Si une seule des valeurs propres est nulle, la matrice est diagonalisable, et le système s'écrit, toujours dans un système de coordonnées bien choisi, $x'_1 = \lambda x_1, x'_2 = 0$, ce qui donne $x_1 = c_1 e^{\lambda t}$ et $x_2 = c_2$.

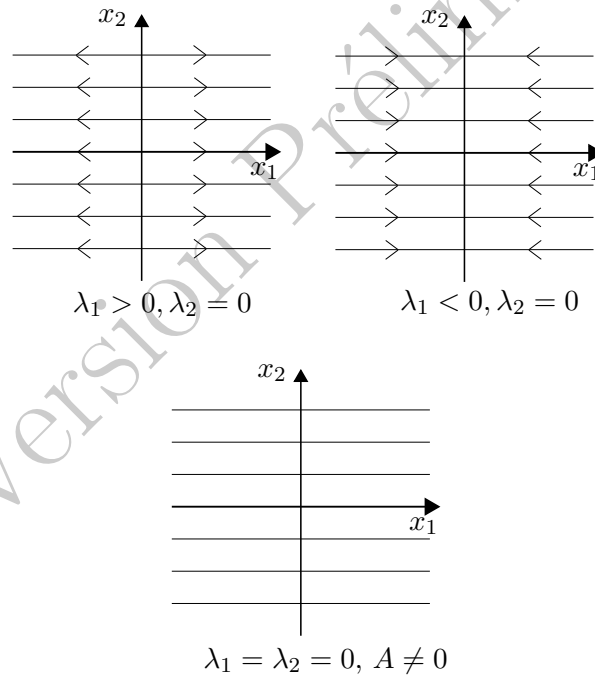


FIGURE 5.1 – Cas dégénérés

Cas de deux racines réelles distinctes

Dans ce cas la matrice A est diagonalisable, et quitte à travailler dans une base de vecteurs propres, le système $Y' = AY$ équivaut à

$$\begin{cases} x'_1 = \lambda_1 x_1, \\ x'_2 = \lambda_2 x_2, \end{cases}$$

avec $|\lambda_1| > |\lambda_2|$. Les solutions sont $(x_1(t), x_2(t)) = (c_1 e^{\lambda_1 t}, c_2 e^{\lambda_2 t})$, où (c_1, c_2) est un vecteur constant. Pour un (c_1, c_2) donné, l'orbite correspondante est la courbe

$$x_2 = C x_1^{\lambda_2/\lambda_1}$$

On a alors deux portraits de phases différents :

- i) si les valeurs propres sont de même signe, on a $0 < \lambda_2/\lambda_1 < 1$, et le portrait de phase est appelé "noeud" (stable si elles sont négatives, instable si elles sont positives).
- ii) si les valeurs propres sont de signes opposés, on a $-1 < \lambda_2/\lambda_1 < 0$ et le portrait de phase est appelé "selle".

Exercice 5.1.4 Faire les dessins correspondants.

Cas où A n'est pas diagonalisable

Dans ce cas, A possède une valeur propre double réelle λ , et l'équation se ramène après changement de base à

$$\begin{cases} x_1'(t) = \lambda x_1(t) + x_2(t), \\ x_2'(t) = \lambda x_2(t). \end{cases}$$

Les solutions sont $(x_1(t), x_2(t)) = (e^{\lambda t}(c_1 + c_2 t), e^{\lambda t} c_2)$, et le portrait de phase correspondant est appelé noeud impropre.

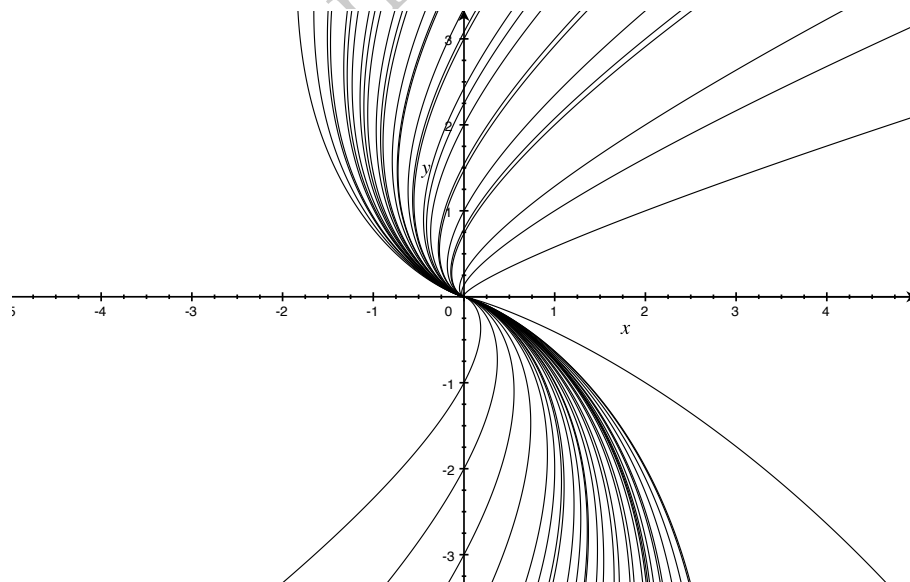


FIGURE 5.2 – Noeud impropre

Cas de deux racines complexes distinctes

Si A a deux valeurs propres complexes conjuguées distinctes, $\lambda = ae^{i\delta}$ et $\bar{\lambda} = ae^{-i\delta}$, l'équation $Y' = AY$ s'écrit dans une base adéquate

$$\begin{cases} x'(t) = a \cos(\delta)x(t) - a \sin(\delta)y(t), \\ y'(t) = a \sin(\delta)x(t) + a \cos(\delta)y(t). \end{cases}$$

On cherche $(x(t), y(t))$ sous la forme $x(t) = r(t) \cos(\theta(t))$, $y(t) = r(t) \sin(\theta(t))$. On obtient alors le système

$$\begin{cases} r'(t) = (a \cos \delta)r(t) \\ \theta'(t) = -(a \sin \delta) \end{cases}$$

Les solutions sont donc données par

$$r(t) = r_0 e^{(a \cos \delta)t}, \theta(t) = \theta_0 - (a \sin \delta)t$$

Ce sont des spirales logarithmiques si $\cos(\delta) \neq 0$. On parle de "foyer", stable si $\cos(\delta) < 0$, et instable si $\cos(\delta) > 0$. Le sens de rotation des spirales est déterminé par le signe de $\sin(\delta)$: trigonométrique si ce signe est positif. Enfin si $\cos(\delta) = 0$ les orbites sont des cercles (plutôt des ellipses : le repère dans lequel on travaille, constitué de vecteurs propres de A , n'est en général pas orthonormé), et l'on parle de "centre".

5.2 Notions de stabilité

La notion d'équilibre stable ou instable nous est familière pour les systèmes mécaniques. On peut penser par exemple au mouvement d'une bille soumise à l'action de la pesanteur sur un terrain présentant des bosses et des creux. Si on lache la bille exactement au fond d'un creux, celle-ci ne bouge pas : le fond d'un creux est un point d'équilibre. Si on la lache près du fond d'un creux, la bille va osciller de part et d'autre du minimum sans sortir du creux : le fond d'un creux est un point d'équilibre stable.

Que se passe-t-il si l'on lache la bille exactement au sommet d'une bosse ? L'expérience est difficile à réaliser : en général on rate le sommet, et la bille part très loin d'un côté ou de l'autre de la bosse. L'écriture de l'équation différentielle obtenue à partir du principe de Newton nous apprend pourtant que le sommet d'une bosse est aussi un point d'équilibre. Cette fois la trajectoire de la bille dépend très fortement de la position initiale de celle-ci : il s'agit d'un équilibre instable. Même si l'on n'est pas maladroit, on rate en fait aussi le fond du creux : mais contrairement au cas du sommet, la stabilité de ce point d'équilibre fait que la bille reste proche du fond, et que l'on ne s'en aperçoit pas.

5.2.1 Définitions

Voici deux notions de stabilité pour un point d'équilibre. On pourrait formuler des définitions du même type pour la stabilité d'une solution quelconque, mais l'on se contente ici de parler de la stabilité de la solution constante $Y : t \mapsto X_0$ où X_0 est un point d'équilibre. De manière imagée, on dit que cette orbite est stable si toutes les orbites restent aussi proches que l'on

veut du point d'équilibre, quitte à considérer des conditions initiales suffisamment proches de X_0 .

Dans toute la suite, il sera commode de noter $\Phi_t(X)$ la valeur à l'instant t de la solution du problème de Cauchy

$$\begin{cases} Y' = F(Y), \\ Y(0) = X. \end{cases}$$

Définition 5.2.1 Soit $X_0 \in \mathbb{R}^m$ un point d'équilibre pour l'équation $Y' = F(Y)$. On dit que X_0 est stable (au sens de Lyapounov) lorsque pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que, pour tout $X \in \mathbb{R}^m$ avec $\|X_0 - X\| < \delta$,

- i) $\Phi_t(X)$ est défini pour tout $t \in \mathbb{R}^+$,
- ii) $\forall t > 0, \|\Phi_t(X) - X_0\| < \epsilon$.

Exercice 5.2.2 Pour les systèmes linéaires 2×2 , dans quels cas $X_0 = 0$ est-il un point fixe stable ?

L'examen des orbites obtenues pour les systèmes 2×2 conduit aussi à la définition suivante.

Définition 5.2.3 Soit $X_0 \in \mathbb{R}^m$ un point d'équilibre stable pour l'équation $Y' = F(Y)$. On dit que X_0 est asymptotiquement stable lorsqu'il existe $\delta > 0$ tel que

$$(5.2) \quad \|X_0 - X\| < \delta \Rightarrow \lim_{t \rightarrow +\infty} \|\Phi_t(X) - X_0\| = 0.$$

Exercice 5.2.4 Pour les systèmes linéaires 2×2 , dans quels cas $X_0 = 0$ est-il un point fixe asymptotiquement stable ?

Contrairement à ce que l'on pourrait penser de prime abord, la propriété (5.2) n'entraîne pas la stabilité : on peut trouver des orbites qui partent loin du point d'équilibre avant de tendre vers ce point quand $t \rightarrow +\infty$.

5.2.2 Cas des systèmes linéaires à coefficients constants

Dans le cas des systèmes linéaires (pas seulement les systèmes 2×2), c'est à dire lorsque $F(X) = AX$ pour $A \in \mathcal{M}_m(\mathbb{R})$, 0 est un point d'équilibre (c'est le seul si et seulement si A est inversible), et on a le résultat suivant.

Proposition 5.2.5 Soit $A \in \mathcal{M}_m(\mathbb{R})$, et $\lambda_j, j = 1 \dots p$ ses valeurs propres.

- 0 est asymptotiquement stable si et seulement si tous les λ_j ont une partie réelle strictement négative.
- 0 est stable si et seulement si tous les λ_j ont une partie réelle négative ou nulle et les blocs de Jordan correspondants à un λ_j de partie réelle nulle sont diagonaux.

Preuve: On utilise la forme explicite des solutions obtenue dans la Proposition 3.2.1. Si A est diagonalisable, les solutions de l'équation $Y' = AY$ s'écrivent

$$Y(t) = \sum_{j=1}^m e^{t\lambda_j} U_j,$$

où les λ_j sont les valeurs propre de A répétées suivant leur multiplicité, et U_j des vecteurs propres associés. On a donc

$$\|Y(t)\| \leq \sum_{j=1}^m e^{t \operatorname{Re} \lambda_j} \|U_j\|,$$

et la proposition en découle facilement dans ce cas. Lorsque A n'est pas diagonalisable, les solutions de l'équation $Y' = AY$ s'écrivent

$$Y(t) = \sum_{j=1}^p \sum_{k=0}^{q_j} \alpha_{j,k} e^{t\lambda_j} t^k U_{j,k},$$

où les λ_j sont les valeurs propres distinctes de A , et les $U_{j,k}$ des vecteurs de \mathbb{C}^n . On a donc

$$\|Y(t)\| \leq \sum_{j=1}^p e^{t \operatorname{Re} \lambda_j} \sum_{k=0}^{q_j} |\alpha_{j,k}| t^k \|U_{j,k}\|.$$

Les termes de la somme qui correspondent à des λ_j de partie réelle strictement négative tendent vers 0 quand $t \rightarrow +\infty$. Par contre si $\operatorname{Re} \lambda_j = 0$, le seul cas où le terme correspondant $\sum_{k=0}^{q_j} |\alpha_{j,k}| t^k \|U_{j,k}\|$ ne tend pas vers l'infini quand $t \rightarrow +\infty$ est celui où $q_j = 0$. Autrement dit, si le bloc correspondant à la valeur propre λ_j dans la décomposition de Jordan de A n'est pas diagonal, il existe des solutions $Y(t)$ telles que $Y(t) \rightarrow +\infty$ quand $t \rightarrow +\infty$. Dans le cas contraire où tous les blocs correspondants à des valeurs propres de partie réelle nulle sont diagonaux, la fonction $Y(t)$ est bornée, et il suffit de prendre $\|Y(0)\|$ (les $\alpha_{j,k}$) suffisamment petits pour que $\|Y(t)\|$ reste, pour tout $t > 0$, inférieur à un $\epsilon > 0$ fixé. \square

5.3 Fonction de Lyapounov

On donne maintenant un critère qui permet d'affirmer qu'un point d'équilibre est stable.

Définition 5.3.1 Soit X_0 un point d'équilibre pour l'équation $Y' = F(Y)$. Soit V un voisinage ouvert et borné de X_0 , et $L : V \rightarrow \mathbb{R}^+$ une fonction \mathcal{C}^1 . On dit que L est une fonction de Lyapunov sur V pour l'équation $Y' = F(Y)$ lorsque

- i) $L(X_0) = 0$,
- ii) $\forall X \in \bar{V} \setminus \{X_0\}$, on a $L(X) > 0$,
- iii) $\forall X \in V$, $\nabla L(X) \cdot F(X) \leq 0$.

Remarque 5.3.2 Lorsque la condition (iii) ci-dessus est une inégalité stricte, c'est-à-dire lorsque L vérifie (i), (ii) et

$$\forall X \in V \setminus \{X_0\}, \nabla L(X) \cdot F(X) < 0,$$

on dit que L est une fonction de Lyapounov stricte.

Proposition 5.3.3 Soit X_0 un point d'équilibre pour l'équation $Y' = F(Y)$. Soit aussi $L : V \rightarrow \mathbb{R}^+$ une fonction \mathcal{C}^1 qui ne s'annule qu'en X_0 . La fonction L est une fonction de Lyapounov (resp. une fonction de Lyapounov stricte) pour X_0 sur V si et seulement si, pour toute solution (J, Y) de l'équation telle que $Y(t) \in V$ pour tout $t \in J$, la fonction $t \mapsto L(Y(t))$ est décroissante (resp. strictement décroissante).

Preuve: Supposons que L est une fonction de Lyapounov sur V pour l'équation $Y' = F(Y)$. Si (J, Y) est une solution de l'équation telle que $Y(t) \in V$ pour tout $t \in J$, on a

$$\partial_t(L(Y(t))) = \nabla L(Y(t)) \cdot Y'(t) \leq 0,$$

donc la fonction $t \mapsto L(Y(t))$ est décroissante.

Réciproquement : soit $X \in V$ et $Y(t) = \Phi_t(X)$ la solution maximale de l'équation $Y' = F(Y)$ telle que $Y(0) = X$. Puisque Y est continue et puisque V est ouvert, il existe $T > 0$, tel que $Y(t) \in V$ pour tout $t \in]-T, T[$. Donc $t \mapsto L(Y(t))$ est décroissante sur $] - T, T[$, et puisqu'il s'agit d'une fonction dérivable, on a, pour tout $t \in] - T, T[$,

$$0 \geq \partial_t(L(Y(t))) = \nabla L(Y(t)) \cdot Y'(t).$$

Pour $t = 0$ on obtient $0 \geq \nabla L(X) \cdot F(X)$. □

Voici le principal résultat de ce chapitre, connu sous le nom de (premier) Théorème de Lyapounov, et que l'on amettra.

Proposition 5.3.4 S'il existe une fonction de Lyapounov (resp. une fonction de Lyapounov stricte) pour le point d'équilibre X_0 , celui-ci est stable (resp. asymptotiquement stable).

Preuve.— Il existe $r > 0$ tel que $L(X) > L(X_0)$ pour tout $X \in \overline{B(X_0, r)} \setminus \{X_0\}$. De plus, pour tout $\epsilon < r$, il existe $\alpha > 0$ tel que, notant

$$U_\alpha = \{X \in B(X_0, r), L(X) < L(X_0) + \alpha\}$$

on ait $U_\alpha \subset B(X_0, \epsilon)$.

En effet sinon, pour tout $\alpha = \frac{1}{n}$, on pourrait trouver $X_n \in B(X_0, r)$ telle que

$$L(X_n) < L(X_0) + \frac{1}{n} \text{ et } \|X_n - X_0\| \geq \epsilon.$$

Puisque la suite (X_n) est bornée, elle admet une sous-suite convergente vers un certain X_1 tel que $X_1 \in \overline{B(X_0, r)} \setminus B(X_0, \epsilon)$ et $L(X_1) \leq L(X_0)$ ce qui contredit le fait que X_0 est un minimum strict de L dans $\overline{B(X_0, r)}$.

Soit $(]T_*, T^*[, Y(t))$ une solution maximale de l'équation, avec $Y(0) = X \in U_\alpha$. On a

$$\partial_t(L(Y(t))) = \nabla L(Y(t)) \cdot Y'(t) = -\|\nabla L(Y(t))\|^2 \leq 0,$$

donc la fonction $t \mapsto L(Y(t))$ est décroissante. Comme $L(X) < L(X_0) + \delta$, on a $L(Y(t)) < L(X_0) + \alpha$ pour tout $t > 0$ donc $L(t) \in U_\alpha \subset B(X_0, \epsilon)$ pour tout $t \in [0, T^*[$. Le théorème des bouts (la Proposition 4.4.6) entraîne alors $T^* = +\infty$.

Enfin, puisque L est continue, U_α est ouvert et contient X_0 , donc il existe $\delta > 0$ tel que $B(X_0, \delta) \subset U_\alpha$, et pour tout $X \in B(X_0, \delta)$ on a, pour tout $t > 0$, $\Phi_t(X) \in B(X_0, \epsilon)$. Donc X_0 est un point d'équilibre stable.

Supposons maintenant de plus que $L : V \rightarrow \mathbb{R}^+$ est une fonction de Lyapounov stricte. On sait que X_0 est stable. Soit alors $\epsilon > 0$ tel que $B(X_0, \epsilon) \subset V$, et $\delta > 0$ tels que pour tout $X \in B(X_0, \delta)$, la solution maximale $\Phi_t(X)$ issue de X est définie pour tout $t \geq 0$ et vérifie $\Phi_t(X) \in B(X_0, \epsilon)$.

Pour $X \in B(X_0, \delta)$, on note $g(t) = L(\Phi_t(X))$. La fonction g est décroissante et minorée par 0, donc il existe $\ell \geq 0$ tel que $g(t) \rightarrow \ell$ quand $t \rightarrow +\infty$. Supposons que $\ell > 0$. Puisque l'ensemble $\{\Phi_t(X), t > 0\}$ est bornée, il existe une suite (t_n) telle que $t_n \rightarrow +\infty$ et $(\Phi_{t_n}(X))$ converge, vers un certain $X_1 \in \overline{B(X_0, \epsilon)}$. Pour $s > 0$, on a

$$L(\Phi_s(X_1)) = L(\Phi_s(\lim_{n \rightarrow +\infty} \Phi_{t_n}(X))) = \lim_{n \rightarrow +\infty} L(\Phi_{s+t_n}(X)) = \lim_{t \rightarrow +\infty} g(t) = \ell = L(X_1),$$

donc L est constante sur la trajectoire issue de X_1 , ce qui contredit l'hypothèse que L est une fonction de Lyapounov stricte si $X_1 \neq X_0$. Donc $X_1 = X_0$, i.e. pour tout $X \in B(X_0, \delta)$, $\Phi_t(X) \rightarrow X_1$. \square

Exemple 5.3.5 On considère le système

$$\begin{cases} y_1'(t) = y_2 - y_1(y_1^2 + y_2^2), \\ y_2'(t) = -y_1 - y_2(y_1^2 + y_2^2). \end{cases}$$

Il s'écrit $Y' = F(Y)$ avec $F(X_1, X_2) = (X_2 - X_1(X_1^2 + X_2^2), -X_1 - X_2(X_1^2 + X_2^2))$. Puisque F est \mathcal{C}^∞ et

$$\partial_2 F_1(X_1, X_2) = 1 - 2X_1X_2 \neq -1 - 2X_1X_2 = \partial_1 F_2(X_1, X_2),$$

le champ F n'est pas un champ de gradient. Le seul point d'équilibre est $X_0 = (0, 0)$, et si $(y_1(t), y_2(t))$ est une solution, on a

$$\partial_t \left(\frac{y_1^2(t) + y_2^2(t)}{2} \right) = y_1(t)y_1'(t) + y_2(t)y_2'(t) = -(y_1^2(t) + y_2^2(t))^2 \leq 0.$$

On considère donc la fonction $L : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ définie par $L(X_1, X_2) = X_1^2 + X_2^2$. On vérifie facilement qu'il s'agit d'une fonction de Lyapounov stricte pour X_0 , et donc que X_0 est un point d'équilibre asymptotiquement stable.

Exemple 5.3.6 On reprend l'exemple vu plus haut du système $Y' = F(Y)$ avec $F(X_1, X_2) = (-X_2^3, X_1^3)$, dont le seul point d'équilibre est $X_0 = (0, 0)$. Ce système s'écrit

$$\begin{cases} y_1'(t) = -y_2^3, \\ y_2'(t) = y_1^3, \end{cases}$$

et l'on peut remarquer que si $(y_1(t), y_2(t))$ en est une solution, on a

$$\partial_t \left(\frac{y_1^4(t) + y_2^4(t)}{4} \right) = y_1(t)^3 y_1'(t) + y_2(t)^3 y_2'(t) = 0.$$

Donc la fonction $L : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ définie par $L(X_1, X_2) = X_1^4 + X_2^4$ est une fonction de Lyapounov pour X_0 , mais cette fois pas une fonction de Lyapounov stricte. Le point d'équilibre $X_0 = (0, 0)$ est stable, mais le théorème de Lyapounov ne dit pas s'il est asymptotiquement stable ou non. On peut cependant voir facilement que ce n'est pas le cas. Puisque L est constante sur les courbes intégrales de cette équation, chacune de ces courbes est incluse dans une courbe de niveau de L , c'est-à-dire un des ensembles

$$L^{-1}(c) = \{(X_1, X_2) \in \mathbb{R}^2, X_1^4 + X_2^4 = c\}.$$

Ces courbes de niveau sont fermées et restent loin de $X_0 = (0, 0)$ pour $c > 0$. Donc aucune trajectoire non triviale $(y_1(t), y_2(t))$ ne peut tendre vers X_0 quand $t \rightarrow +\infty$.

5.4 Linéarisation et stabilité

5.4.1 Système linéarisé

Pour un système linéaire à coefficients constants, on peut déterminer facilement si un point d'équilibre est stable, asymptotiquement stable ou instable. Dans le cas général où X_0 est un point d'équilibre d'une équation $Y' = F(Y)$ par exemple non-linéaire, il est donc naturel d'essayer de se ramener à une équation linéaire, et d'approcher F près de X_0 par sa différentielle $d_{X_0}F$ en X_0 .

Définition 5.4.1 Soit $F : U \rightarrow \mathbb{R}^m$ une fonction de classe \mathcal{C}^1 sur l'ouvert $U \subset \mathbb{R}^m$, et $X_0 \in U$ un point d'équilibre pour l'équation $Y' = F(Y)$. On appelle linéarisé de l'équation en X_0 le système linéaire à coefficients constants

$$Z' = AZ, \text{ avec } A = d_{X_0}F.$$

5.4.2 Le critère de Routh

Pour un système linéaire $Y' = AY$, on a montré que le point d'équilibre $X_0 = 0$ est asymptotiquement stable lorsque les valeurs propres de A ont une partie réelle strictement négative. On a utilisé pour cela le calcul explicite des solutions de ce système obtenue dans le Chapitre 3. Dans l'esprit de ce qui précède, on peut aussi exhiber une fonction de Lyapounov stricte pour ce système.

Proposition 5.4.2 Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice dont toutes les valeurs propres ont une partie réelle strictement négative. Pour tout $X \in \mathbb{R}^n$, l'intégrale généralisée

$$\int_0^{+\infty} \|e^{sA}X\|^2 ds$$

converge, et la fonction $L(X) = \int_0^{+\infty} \|e^{sA}X\|^2 ds$ est une fonction de Lyapounov stricte pour l'équation $Y' = AY$.

Preuve: On a vu que pour tout $X \in \mathbb{R}^m$, on peut écrire

$$e^{sA}X = \sum_{j=1}^p \sum_{k=0}^{q_j} X_{j,k} e^{t\lambda_j} t^k U_{j,k}.$$

Soit alors

$$m = \max\{\operatorname{Re} \lambda, \lambda \in \operatorname{sp}(A)\} < 0.$$

Puisque, pour tout $j \in \mathbb{N}$, $e^{m/2}t^j \rightarrow 0$ quand $t \rightarrow +\infty$, il existe une constante $C > 0$ telle que

$$\|e^{sA}X\| \leq C e^{ms/2},$$

ce qui entraîne la convergence de l'intégrale.

Soit $X, U \in \mathbb{R}^m$ et $t \in \mathbb{R}$. On a

$$\begin{aligned} L(X + tU) &= \int_0^{+\infty} \|e^{sA}(X + tU)\|^2 ds = \int_0^{+\infty} (e^{sA}X + te^{sA}U) \cdot (e^{sA}X + te^{sA}U) ds \\ &= \int_0^{+\infty} \|e^{sA}X\|^2 + 2t(e^{sA}X) \cdot (e^{sA}U) + t^2\|e^{sA}U\|^2 ds \end{aligned}$$

Donc

$$(5.3) \quad \nabla L(X) \cdot U = d_X L(U) = \lim_{t \rightarrow 0} \frac{L(X + tU) - L(X)}{t} = 2 \int_0^{+\infty} (e^{sA} X) \cdot (e^{sA} U) ds.$$

En particulier, si $Y(t) = e^{tA} X_0$ est une solution de l'équation $Y' = AY$, on a

$$(5.4) \quad \begin{aligned} \nabla L(Y(t)) \cdot Y'(t) &= 2 \int_0^{+\infty} (e^{sA} Y(t)) \cdot (e^{sA} Y'(t)) ds = 2 \int_0^{+\infty} (e^{sA} Y(t)) \cdot (e^{sA} AY(t)) ds \\ &= \int_0^{+\infty} \partial_s (\|e^{sA} Y(t)\|^2) ds = -\|Y(t)\|^2. \end{aligned}$$

□

Il s'avère que la fonction L définie ci-dessus reste une fonction de Lyapounov pour de petites perturbation de l'équation $Y' = AY$, c'est-à-dire des équations de la forme $Y' = F(Y)$ où $F(Y)$ est "proche" de AY . On obtient en particulier le critère suivant, que l'on attribue en théorie des systèmes à Routh.

Proposition 5.4.3 Soit X_0 un point d'équilibre pour l'équation $Y' = F(Y)$, et $A = d_{X_0} F$. Si toutes les valeurs propres de A ont une partie réelle strictement négative, alors X_0 est asymptotiquement stable.

Preuve: Pour simplifier les notations, on suppose que $X_0 = 0$. On a alors $F(X) = AX + G(X)$, où $G(X) = o(X)$ quand $X \rightarrow 0$. D'après (5.4), (en prenant $t = 0$ et $Y(0) = X$) on a donc

$$\nabla L(X) \cdot F(X) = \nabla L(X) \cdot AX + \nabla L(X) \cdot G(X) = -\|X\|^2 + \nabla L(X) \cdot G(X).$$

Or il existe $C > 0$ tel que $\|\nabla L(X)\| < C\|X\|$ (cf. (5.3)), donc pour $\|X\|$ suffisamment petit, on a

$$\|G(X)\| \leq \frac{1}{2C} \|X\|,$$

et

$$\|\nabla L(X) \cdot F(X)\| \leq -\frac{1}{2} \|X\|^2.$$

Donc L est bien une fonction de Lyapounov stricte en X_0 pour l'équation. □

5.4.3 Le Théorème d'Hartman-Grobman

On termine par un résultat profond, qu'il serait déraisonnable de vouloir démontrer ici. En fait, lorsqu'aucune valeur propre de $d_{X_0} F$ n'a une partie réelle nulle, le comportement des solutions du système $Y' = F(Y)$ près de X_0 "ressemble" à celui des solutions du système linéarisé :

Proposition 5.4.4 (Théorème d'Hartman-Grobman) Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction \mathcal{C}^1 , telle que $F(0) = 0$. Si la matrice $A = d_0 F$ n'a pas de valeur propre de partie réelle nulle, alors il existe deux voisinages ouverts U et V de 0, et un homéomorphisme $h : U \rightarrow V$ tel que h transforme les orbites de l'équation $Y' = F(y)$ en les orbites de l'équation linéarisée $Y' = AY$. Plus précisément

$$\forall X \in U, \exists I \subset \mathbb{R}, \forall t \in I, h(\Phi_t(X)) = e^{tA}h(X).$$

Version Préliminaire

Chapitre 6

Systèmes prédateurs-proies

On revient comme promis au tout début du cours sur les systèmes différentiels de Lotka et Volterra, introduits pour décrire l'évolution conjointe de deux populations : les prédateurs et les proies. Il s'agit des systèmes de la forme

$$(6.1) \quad \begin{cases} y_1'(t) = ay_1(t) - by_1(t)y_2(t), \\ y_2'(t) = -cy_2(t) + dy_1(t)y_2(t), \end{cases}$$

Ici a, b, c et d sont des constantes strictement positives. L'idée est que si les prédateurs ($y_2(t)$) et les proies ($y_1(t)$) évoluaient indépendamment l'une de l'autre, le nombre de proies croîtrait de manière exponentielle (loi de croissance Malthusienne) : $y_1(t) = C_1 e^{at}$, et le nombre de prédateurs diminuerait de manière exponentielle aussi $y_2(t) = C_2 e^{-ct}$. Ces deux fonctions vérifieraient alors le système

$$\begin{cases} y_1'(t) = ay_1(t) \\ y_2'(t) = -cy_2(t), \end{cases}$$

L'interaction entre les deux populations se traduit, d'une part par une diminution du taux de croissance de la population de proies proportionnelle au nombre de rencontres entre elles, et d'autre part, par une augmentation du taux de croissance de la population de prédateurs, également proportionnelle au nombre de rencontres entre les deux populations.

6.1 Existence et unicité. Temps d'existence

On considère le problème de Cauchy (6.1) avec pour condition initiale $y_1(0) = x_1, y_2(0) = x_2$, où x_1 et x_2 sont deux réels positifs ou nuls. L'équation (6.1) s'écrit

$$Y' = F(Y), \text{ avec } F(x_1, x_2) = (ax_1 - bx_1x_2, -cx_2 + dx_1x_2).$$

La fonction $F(X)$ est indépendante de t (il s'agit d'un système autonome), et est localement lipschitzienne par rapport à X puisqu'elle est C^∞ . Le théorème de Cauchy-Lipschitz s'applique, et permet d'affirmer que ce problème admet une unique solution maximale (J, Y) .

Voyons quelques cas particuliers :

- si $x_1 = x_2 = 0$, la solution est la fonction constante nulle $Y(t) = (0, 0)$ et $J = \mathbb{R}$.
- si $x_1 = 0, x_2 > 0$, on voit que la fonction $Y(t) = (0, x_2 e^{-ct})$ est une solution, donc la solution, et $J = \mathbb{R}$ aussi.
- si $x_1 > 0, x_2 = 0$, là encore la solution est $Y(t) = (x_1 e^{at}, 0)$, avec $J = \mathbb{R}$.

On sait que les orbites de ce système ne peuvent pas se couper. On en déduit donc que lorsque $x_1 > 0, x_2 > 0$, la solution $(J, (y_1, y_2))$ correspondante vérifie $y_1(t) > 0$ et $y_2(t) > 0$ pour tout $t \in J$. Dans ce cas on a, pour tout $t \in J$, $y_1'(t) \leq ay_1(t)$, donc $y_1(t) \leq x_1 e^{at}$, et $y_2'(t) \leq dy_1(t)y_2(t) \leq f(t)y_2(t)$ où $f(t) = dx_1 e^{at}$. En intégrant, on obtient,

$$y_2(t) \leq x_2 + \int_0^t f(s)y_2(s)ds$$

et le Lemme de Gronwall (Proposition 4.4.10) donne, pour $t > 0$,

$$y_2(t) \leq x_2 \exp\left(\int_0^t f(s)ds\right).$$

Du coup la solution $Y(t) = (y_1(t), y_2(t))$ reste bornée sur tout intervalle borné, et le théorème des bouts (Proposition 4.4.6) entraîne donc que $[0, +\infty[\subset J$. On peut montrer aussi que $] -\infty, 0] \subset J$: on a $y_2'(t) \geq -cy_2(t)$ donc $y_2(t) \leq x_2 e^{-ct}$ pour $t < 0$, puis $y_1'(t) \geq -by_1(t)y_2(t)$, et on conclut de la même manière.

En conclusion, toutes les solutions du système, correspondant à des données initiales x_1, x_2 positives, sont globales.

6.2 Points d'équilibre

Les points d'équilibre éventuels du système vérifient l'équation

$$F(x_1, x_2) = 0 \Leftrightarrow \begin{cases} x_1(a - bx_2) = 0, \\ x_2(-c + dx_1) = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = 0, \\ x_2 = 0, \end{cases} \quad \text{ou} \quad \begin{cases} x_2 = a/b, \\ x_1 = c/d. \end{cases}$$

Il y a donc deux points d'équilibre : $O(0, 0)$ et $A(c/d, a/b)$. La différentielle de F au point (x_1, x_2) est

$$dF(x_1, x_2) = \begin{pmatrix} a - bx_2 & -bx_1 \\ dx_2 & -c + dx_1 \end{pmatrix},$$

donc

$$dF(0) = \begin{pmatrix} a & 0 \\ 0 & -c \end{pmatrix} \quad \text{et} \quad dF(A) = \begin{pmatrix} 0 & -bc/d \\ da/b & 0 \end{pmatrix}.$$

Le système linéarité en 0 est donc un point selle, alors que le système linéarité en A est un centre. Le théorème de Hartman-Grobman permet d'affirmer que les orbites du système complet auront, près de O , "la même allure" que celles du système linéarité, mais aucun des résultats précédents ne permet de prédire l'allure des orbites du système complet près de A .

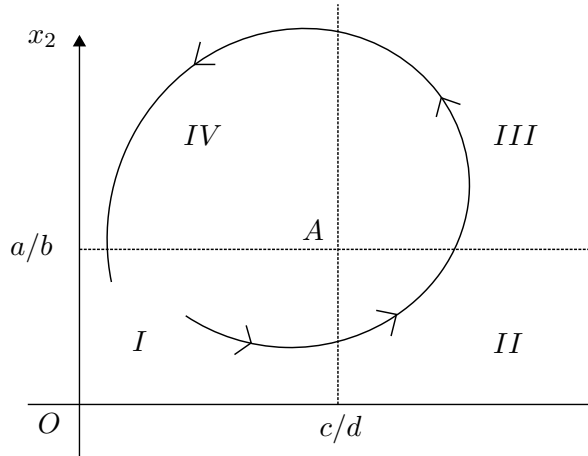


FIGURE 6.1 – Régionnement du plan pour le système prédateurs-proies

6.3 Comportement global

On découpe le premier quadrant $\{(x_1, x_2) \in \mathbb{R}^2, x_1 > 0, x_2 > 0\}$ en quatre régions : $I =]0, c/d[\times]0, a/b[$, $II =]c/d, +\infty[\times]0, a/b[$, $III =]c/d, +\infty[\times]a/b, +\infty[$ et $IV =]0, c/d[\times]a/b, +\infty[$.

On montre d'abord qu'une trajectoire issue de la région I passe nécessairement dans II . On commence par la

Proposition 6.3.1 Soit $([a, +\infty[, Y)$ une solution du système linéaire autonome $Y' = F(Y)$. Si $Y(t) \rightarrow X$ quand $t \rightarrow +\infty$, X est un point d'équilibre du système, i.e. $F(X) = 0$.

Preuve.— Ce résultat repose sur le

Lemme 6.3.2 Soit f une fonction \mathcal{C}^1 sur $[a, +\infty[$. S'il existe $\ell \in \mathbb{R}$ et $\ell' \in \mathbb{R} \cup \{-\infty, +\infty\}$ tels que $f(t) \rightarrow \ell$ et $f'(t) \rightarrow \ell'$ quand $t \rightarrow +\infty$, alors $\ell' = 0$.

Preuve du lemme : Supposons par exemple que $\ell' > 0$ (resp. $\ell' = +\infty$). Soit alors $m = \ell'/2$ (resp. $m > 0$ quelconque). Il existe $T > 0$ tel que $f'(t) > m$ pour tout $t > T$. Alors, pour tout $t > T$,

$$f(t) - f(T) = \int_T^t f'(s) ds \geq m(t - T) \rightarrow +\infty \text{ quand } t \rightarrow +\infty,$$

ce qui contredit le fait que $f(t) \rightarrow \ell$ quand $t \rightarrow +\infty$. Le cas $\ell' < 0$ se traite de la même manière. \square

Soit alors $Y = (y_1, y_2, \dots, y_n)$ une solution du système qui converge vers X . On applique la proposition aux fonctions $y_1(t), \dots, y_n(t)$ qui ont toutes une limite quand $t \rightarrow +\infty$. Puisque F est continue, l'équation $Y' = F(Y)$ dit alors que les fonctions $y'_j(t)$ ont aussi une limite,

donc que ces limites sont nulles. Du coup, en passant à la limite $t \rightarrow +\infty$ dans l'équation $Y'(t) = F(Y(t))$, on obtient $F(X) = 0$. \square

Soit $(x_1^0, x_2^0) \in I$, et $(y_1(t), y_2(t))$ la solution du problème de Cauchy correspondant. Supposons que $(y_1(t), y_2(t)) \in I$ pour tout $t \in [0, +\infty[$. Alors $y_1'(t) > 0$ et $y_2'(t) < 0$ pour tout $t > 0$. Comme y_1 et y_2 sont alors monotones et bornées, elles convergent toutes les deux vers $x_1^\infty \geq x_1^0 > 0$ et $x_2^\infty \leq x_2^0 < a/b$ respectivement. Or la limite d'une courbe intégrale est nécessairement un point fixe de F ce qui est absurde puisque $x_1^\infty > 0$ et $x_2^\infty < a/b$.

Récapitulons : la trajectoire issue de $(x_1^0, x_2^0) \in I$ quitte nécessairement la région I . Comme $y_2(t)$ est décroissante, elle ne peut sortir de I qu'en un point $(y_1(T), y_2(T))$ du segment $\{x_1 = c/d, x_2 \in]0, a/b[\}$. En un tel point on a $y_1'(T) > 0$ donc la courbe rentre dans la région II .

De la même manière, on montre que toute trajectoire issue de la région II passe dans la région III . Soit $(x_1^0, x_2^0) \in II$, et $(y_1(t), y_2(t))$ la solution du problème de Cauchy correspondant. Supposons que $(y_1(t), y_2(t)) \in II$ pour tout $t \in [0, +\infty[$. Alors $y_1'(t) > 0$ et $y_2'(t) > 0$ pour tout $t > 0$. La fonction y_2 est bornée par hypothèse, et deux cas peuvent se présenter pour y_1 :

- ou bien y_1 est bornée : dans ce cas $(y_1(t), y_2(t))$ tend vers une limite X , qui est un point fixe de F , mais c'est impossible puisque $y_1(0) > c/d$ compte tenu de la croissance de y_1 .
- ou bien y_1 tend vers $+\infty$. Dans ce cas, la deuxième équation du système $y_2'(t) = -cy_2(t) + dy_1(t)y_2(t)$ donne $y_2'(t) \rightarrow +\infty$ quand $t \rightarrow +\infty$, ce qui contredit le fait que y_2 a une limite.

Donc la trajectoire quitte la région II , et compte tenu du sens de variation de y_1 , elle entre nécessairement dans la région III .

On prouve avec les mêmes arguments que toute trajectoire issue de la région III entre dans la région IV , et que toute trajectoire issue de la région IV rentre dans la région I . On cherche à savoir si les trajectoires sont périodiques, c'est-à-dire si une trajectoire issue de $(x_1^0, x_2^0) \in I$ repasse par $(x_1^0, x_2^0) \in I$ après avoir fait le tour du point fixe A .

6.4 Une fonction de Lyapounov

On va montrer que la fonction $H : (x_1, x_2) \mapsto dx_1 - c \ln x_1 + bx_2 - a \ln x_2$ est une fonction de Lyapounov pour le système au point A . En fait H est constante le long des trajectoires : si $(y_1(t), y_2(t))$ est une solution du système, pour des données initiales strictement positives, on a $y_1(t), y_2(t) > 0$ et

$$\begin{aligned} \partial_t(H(y_1(t), y_2(t))) &= \partial_1 H(y_1(t), y_2(t))y_1'(t) + \partial_2 H(y_1(t), y_2(t))y_2'(t) \\ &= (d - \frac{c}{y_1(t)})y_1(t)(a - by_2(t)) + (b - \frac{a}{y_2(t)})y_2(t)(-c + dy_1(t)) \\ &= 0. \end{aligned}$$

Le théorème de Lyapounov nous apprend que A est un point d'équilibre stable pour le système prédateurs-proies, mais on peut être plus précis : dire que H est constante le long des tra-

jectoires équivaut à dire que chaque trajectoire est incluse dans une courbe de niveau de la fonction H . Or ces courbes de niveaux sont fermées, donc les trajectoires sont périodiques.

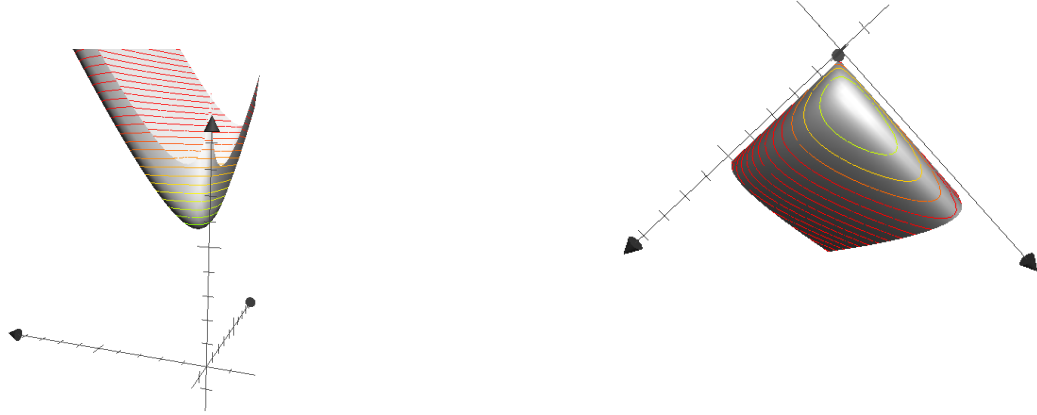


FIGURE 6.2 – La fonction H : vue de côté, vue du dessous

On peut le voir de la manière suivante. Soit $(x_1^0, x_2^0) \in I$ et $Y(t) = (y_1(t), y_2(t))$ la trajectoire correspondante. On sait que la trajectoire revient dans la région I, et puisqu'elle ressort dans la région II, il existe $0 < T_1 < T_2$ tels que $y_1(T_j) = c/d$ et $y_2(T_j) \in]0, a/b[$. Or sur le segment $\{x_1 = c/d, x_2 \in]0, a/b[\}$, la fonction H est injective :

$$H(c/d, x_2) = f(x_2) \text{ avec } f(x) = c - c \ln c/d + b - a \ln x,$$

et $f'(x) = b - a/x > 0$ pour $x \in]0, a/b[$. Comme $H(Y(T_1)) = H(Y(T_2))$, on a donc $Y(T_1) = Y(T_2)$. On peut alors en déduire que Y est périodique de période $T = T_2 - T_1$: la fonction $\tilde{Y}(t) = Y(t + T)$ est une solution (puisque l'équation est autonome), et $\tilde{Y}(T_1) = Y(T_2) = Y(T_1)$. Donc $\tilde{Y}(t) = Y(t)$ pour tout t , c'est-à-dire $Y(t + T) = Y(t)$.

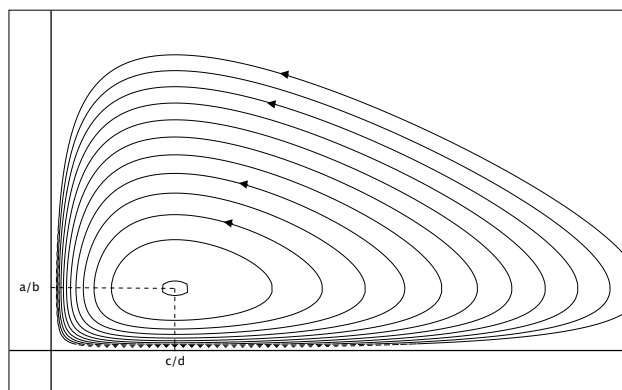


FIGURE 6.3 – Portrait de phase pour le système prédateurs-proies

Bibliographie

- [1] Jean-Pierre Demailly
- [2] Dieudonné
- [3] Perko
- [4] Mawhin Rouché
- [5] Viterbo

Version Préliminaire