

Équations différentielles ordinaires

Sylvie Benzoni

11 mai 2007

Chapitre I

Introduction

La forme la plus générale d'une *équation différentielle ordinaire* (en abrégé *É.D.O.*) est

$$F(t, u, u', \dots, u^{(k)}) = 0,$$

où u est une fonction inconnue de la variable réelle t à valeurs dans \mathbb{R}^n ou plus généralement dans un espace de Banach X , $u', \dots, u^{(k)}$ désignent les dérivées successives de u , et F est une fonction donnée, supposée « régulière » (on précisera comment par la suite) sur $I \times U \times U^1 \times \dots \times U^k$ où I est un intervalle ouvert de \mathbb{R} , U, U^1, \dots, U^k sont des ouverts connexes de X . On ne s'intéressera dans ce cours qu'à des équations différentielles *résolues*, pour lesquelles il existe une fonction G , régulière sur $I \times U \times U^1 \times \dots \times U^{k-1}$ telle que

$$F(t, u, u', \dots, u^{(k)}) = 0 \quad \Leftrightarrow \quad u^{(k)} = G(t, u, u', \dots, u^{(k-1)}).$$

On observe de plus que

$$u^{(k)} = G(t, u, u', \dots, u^{(k-1)}) \quad \Leftrightarrow \quad U' = \mathbb{G}(t, U),$$

$$U := \begin{pmatrix} u \\ u' \\ \vdots \\ u^{(k-1)} \end{pmatrix}, \quad \mathbb{G}(t, U) := \begin{pmatrix} 0 & I & \dots & 0 \\ & \ddots & \ddots & \vdots \\ & & & I \\ 0 & \dots & & 0 \end{pmatrix} U + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ G(t, u, u', \dots, u^{(k-1)}) \end{pmatrix},$$

la fonction \mathbb{G} étant évidemment aussi régulière que G . On supposera donc sans perte de généralité $k = 1$.

Désormais, on considère « une »¹ équation dite d'ordre 1, de la forme

$$\frac{du}{dt} = f(t, u)$$

où u est une fonction inconnue de la variable réelle t à valeurs dans un espace de Banach X , et f est une fonction donnée sur $I \times U$, ouvert connexe non vide de $\mathbb{R} \times X$. Lorsque f ne dépend pas de t , l'équation différentielle est dite *autonome*.

Remarque I.1 *On peut toujours se ramener, par une astuce, à une équation autonome. En effet, il suffit de considérer l'équation étendue*

$$\frac{d}{ds} \begin{pmatrix} u \\ t \end{pmatrix} = \begin{pmatrix} f(t, u) \\ 1 \end{pmatrix}.$$

¹Les guillemets sont là pour souligner qu'en général il s'agit en fait d'un *système* d'équations !

Cette approche, parfois utile, est malgré tout artificielle. Il faut savoir étudier directement certaines propriétés des équations dites à *coefficients variables*, où f dépend vraiment de t . Ce sera le cas au moins dans les deux premiers chapitres.

Exemples et contre-exemples. De nombreux modèles physiques (en mécanique, électricité, chimie, écologie, etc.) s'expriment au moyen d'équations différentielles ordinaires en dimension finie. Citons simplement l'équation de la mécanique des points matériels :

$$m \frac{d^2x}{dt^2} = F(x),$$

qui s'écrit de façon équivalente dans le *plan de phase*

$$\begin{cases} \frac{dx}{dt} = v, \\ \frac{dv}{dt} = \frac{1}{m} F(x). \end{cases}$$

Dans cette équation, posée dans \mathbb{R}^2 et autonome, $F(x)$ représente la résultante des forces appliquées au point x , supposé de masse m . Des exemples d'équations différentielles en dimension infinie peuvent venir de la discrétisation d'*équations aux dérivées partielles*. Prenons par exemple l'*équation de la chaleur*

$$\partial_t v = \partial_{xx}^2 v.$$

Une façon d'approcher les solutions de cette équation est de chercher $v(j\Delta x, t) \simeq u_j(t)$ (où Δx est un *pas* de discrétisation et $j \in \mathbb{Z}$) avec

$$\frac{du_j}{dt} = \frac{1}{\Delta x^2} (u_{j+1} - 2u_j + u_{j-1}).$$

Ceci est une équation différentielle ordinaire dans l'espace de suites $\ell^p(\mathbb{Z})$, qui est un espace de Banach quel que soit $p \in [1, \dots, +\infty]$. Attention, pour voir l'équation de la chaleur elle-même comme une équation différentielle ordinaire, il faudrait disposer d'un espace fonctionnel qui soit un espace de Banach stable par ∂_{xx}^2 ! De façon générale, la théorie des équations différentielles ordinaires ne s'applique pas aux équations aux dérivées partielles. (Toutefois, les équations aux dérivées partielles d'évolution *linéaires* posées dans *tout l'espace* se ramènent à des équations différentielles ordinaires grâce à la transformation de Fourier...)

« Résolution » des équations différentielles. Dans le chapitre II, on va s'intéresser à l'existence, à l'unicité, et à la dépendance des solutions par rapport aux « conditions initiales » $u(t_0) = u_0$. On (re)verra des résultats essentiellement théoriques, car il y a très peu d'équations différentielles dont on connaît explicitement les solutions, en dehors des équations linéaires à coefficients constants (dont les solutions s'expriment à l'aide de l'exponentielle de matrice) et des équations scalaires d'ordre 1 autonomes (dont le calcul des solutions se ramène à un calcul de primitive). Le chapitre III sera consacré aux propriétés spécifiques des équations linéaires, en insistant sur le cas à coefficients variables. À partir du chapitre IV, on s'attaquera aux *propriétés qualitatives* des équations différentielles : sans prétendre résoudre ces équations, on peut en effet obtenir beaucoup d'informations sur le comportement de leurs solutions. (Cette idée générale remonte à Poincaré.) On étudiera notamment l'existence et la stabilité de solutions particulières, comme les solutions stationnaires et les solutions périodiques (qui jouent un grand rôle dans les applications).

Un outil indispensable. La théorie des équations différentielles utilise abondamment le lemme (ou l'inégalité) de Gronwall. Ce chapitre introductif est l'occasion de le rappeler.

Parfois, on appelle (à tort) lemme de Gronwall le fait suivant : si une fonction u à valeurs dans \mathbb{R} , de classe C^1 , satisfait une inégalité différentielle du type :

$$u'(t) \leq b(t) + a(t)u(t), \quad \text{pour tout } t \in [0, T]$$

alors

$$u(t) \leq e^{\int_0^t a(s) ds} u(0) + \int_0^t b(\tau) e^{\int_\tau^t a(s) ds} d\tau, \quad \text{pour tout } t \in [0, T].$$

En effet, l'inégalité différentielle implique

$$\frac{d}{dt} \left(e^{-\int_0^t a(s) ds} u(t) \right) \leq e^{-\int_0^t a(s) ds} b(t),$$

et donc par intégration entre 0 et t on obtient immédiatement l'inégalité annoncée. Le lemme de Gronwall est un peu plus subtil, puisqu'il suppose une inégalité intégrale et non une inégalité différentielle (la seconde impliquant la première mais pas l'inverse). Or les *estimations a priori* que l'on obtient en général sont plutôt du type intégral, d'où l'intérêt de ce lemme, dont la preuve est néanmoins élémentaire.

Lemme I.1 (Gronwall) Si $u \in C([0, T]; \mathbb{R}^+)$ est telle qu'il existe a et $b \in C([0, T]; \mathbb{R}^+)$ avec

$$u(t) \leq b(t) + \int_0^t a(\tau) u(\tau) d\tau, \quad \text{pour tout } t \in [0, T],$$

alors

$$u(t) \leq b(t) + \int_0^t b(\tau) a(\tau) e^{\int_\tau^t a(s) ds} d\tau, \quad \text{pour tout } t \in [0, T].$$

Dém. La seule astuce consiste à majorer l'intégrale du second membre

$$v(t) = \int_0^t a(\tau) u(\tau) d\tau$$

par la méthode décrite précédemment. Comme

$$v'(t) = a(t)u(t) \leq a(t)(b(t) + v(t))$$

par hypothèse², on a donc

$$\frac{d}{dt} \left(e^{-\int_0^t a(s) ds} v(t) \right) \leq e^{-\int_0^t a(s) ds} a(t) b(t),$$

d'où après intégration (notez que $v(0) = 0$ par définition) :

$$v(t) \leq \int_0^t a(\tau) b(\tau) e^{\int_\tau^t a(s) ds} d\tau, \quad \text{pour tout } t \in [0, T].$$

En majorant de cette façon $v(t)$ dans l'inégalité de départ, on obtient le résultat. \square

²Attention, ici intervient de façon cruciale le fait que a soit positive !

Cette version du lemme de Gronwall est donnée surtout pour mettre en évidence la méthode de calcul. On pourrait donner une autre version de l'inégalité obtenue, en intégrant par parties si b est dérivable :

$$u(t) \leq b(0) e^{\int_0^t a(s) ds} + \int_0^t b'(\tau) e^{\int_\tau^t a(s) ds} d\tau, \quad \text{pour tout } t \in [0, T].$$

En particulier, si b est constante, on obtient simplement :

$$u(t) \leq b e^{\int_0^t a(s) ds}.$$

En pratique, il est conseillé de refaire rapidement le calcul pour éviter les erreurs.

Chapitre II

Le problème de Cauchy général

Soit U un ouvert connexe non vide d'un \mathbb{R} -espace de Banach X , I un intervalle ouvert de \mathbb{R} , et $f : I \times U \rightarrow X$ une fonction *continue*. On appelle *solution* de l'équation différentielle

$$(1) \quad \frac{du}{dt} = f(t, u)$$

une fonction u de classe \mathcal{C}^1 sur un intervalle $J \subset I$ et à valeurs dans U , dont la dérivée vérifie

$$u'(t) = f(t, u(t)) \quad \text{pour tout } t \in J.$$

(Il est parfois d'usage de noter par une lettre différente la fonction inconnue dans (1) et les solutions.) On appelle *condition initiale* une égalité de la forme $u(t_0) = u_0$ avec $(t_0, u_0) \in I \times U$. On appelle *problème de Cauchy* le système d'équations

$$(2) \quad \begin{cases} u'(t) = f(t, u(t)), & t \in J, \\ u(t_0) = u_0. \end{cases}$$

Résoudre le problème de Cauchy (localement) revient à trouver un intervalle $J \subset I$ contenant t_0 et une fonction u de classe \mathcal{C}^1 sur J satisfaisant (3).

Lorsque f est « seulement » continue et X est de dimension infinie, on ne peut rien dire. Si X est de dimension finie en revanche, le théorème de Cauchy-Arzela-Peano (dont la démonstration repose sur le théorème d'Ascoli) affirme l'existence d'au moins une solution de (3), quel que soit $(t_0, u_0) \in I \times U$. Mais il n'y a pas unicité en général.

On supposera dans la suite un peu plus de régularité sur f , de sorte que l'on ait à la fois existence et unicité de solutions locales.

1 Existence et unicité locale

Théorème II.1 (Cauchy-Lipschitz) *On suppose $f \in \mathcal{C}(I \times U; X)$ et localement Lipschitzienne par rapport à u .*

Existence : *Quel que soit $(t_0, u_0) \in I \times U$, il existe $\tau > 0$ et $u \in \mathcal{C}^1([t_0 - \tau, t_0 + \tau]; U)$ solution de (3) avec $J = [t_0 - \tau, t_0 + \tau]$.*

Unicité : *si v est une autre solution, elle coïncide avec u sur un intervalle d'intérieur non vide inclus dans $[t_0 - \tau, t_0 + \tau]$.*

Régularité : *Si de plus f est de classe \mathcal{C}^r , $r \geq 1$, alors u est de classe \mathcal{C}^{r+1} .*

Dém. Pour éviter de « traîner » des valeurs absolues, on fait la démonstration en remplaçant l'intervalle centré $[t_0 - \tau, t_0 + \tau]$ par $[t_0, t_+]$ avec $t_+ = t_0 + \tau > t_0$ (le cas de l'intervalle $[t_0 - \tau, t_0]$ s'en déduisant par le changement de variable $t \mapsto (t_0 - t)$).

L'étape 0 consiste à choisir un voisinage de (t_0, u_0) de la forme

$$C(t_+, R) = [t_0, t_+] \times \bar{B}(u_0, R) \subset I \times U$$

avec $t_+ > t_0$ et $R > 0$, où $\bar{B}(u_0, R)$ désigne la boule fermée de centre u_0 et de rayon R , dans lequel f est bornée, disons par une constante M :

$$\|f(t, u)\| \leq M \quad \text{quel que soit } (t, u) \in [t_0, t_+] \times \bar{B}(u_0, R),$$

et est Lipschitzienne, disons avec une constante de Lipschitz L :

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{pour tous } t \in [t_0, t_+], x \text{ et } y \in \bar{B}(u_0, R).$$

La nœud de la démonstration consiste alors à réécrire le problème de Cauchy

$$\begin{cases} u'(t) = f(t, u(t)), \\ u(t_0) = u_0 \end{cases}$$

sous la forme intégrale :

$$u(t) = u_0 + \int_{t_0}^t f(\tau, u(\tau)) \, d\tau,$$

et à chercher une solution comme limite de la suite $(u^n)_{n \in \mathbb{N}}$ définie par

$$u^0 = u_0, \quad u^{n+1}(t) = u_0 + \int_{t_0}^t f(\tau, u^n(\tau)) \, d\tau.$$

L'étape 1 est très facile. Elle consiste à vérifier que pour t_+ assez proche de t_0 , la suite $(u^n)_{n \in \mathbb{N}}$ est bien définie dans $[t_0, t_+]$ et à valeurs dans $\bar{B}(u_0, R)$. En effet, la fonction u^0 constante est bien à valeurs dans $\bar{B}(u_0, R)$. Supposons qu'on ait construit u^n continue sur $[t_0, t_+]$ et à valeurs dans $\bar{B}(u_0, R)$. Alors u^{n+1} est bien définie et de plus :

$$\|u^{n+1}(t) - u_0\| \leq \int_{t_0}^t \|f(\tau, u^n(\tau))\| \, d\tau \leq (t_+ - t_0) M.$$

Donc il suffit que

$$(t_+ - t_0) M \leq R$$

pour que u^{n+1} soit aussi à valeurs dans $\bar{B}(u_0, R)$.

L'étape 2 est de montrer que, quitte à diminuer t_+ , la suite $(u^n)_{n \in \mathbb{N}}$ est de Cauchy et donc convergente dans l'espace de Banach $\mathcal{C}([t_0, t_+]; X)$ (muni de la norme uniforme). Pour $n \geq 1$ on a

$$\|u^{n+1}(t) - u^n(t)\| \leq \int_{t_0}^t \|f(\tau, u^n(\tau)) - f(\tau, u^{n-1}(\tau))\| \, d\tau \leq (t_+ - t_0) L \sup_{[t_0, t_+]} \|u^n - u^{n-1}\|.$$

Donc, par exemple si t_+ est tel que

$$(t_+ - t_0) L \leq \frac{1}{2}$$

on a par une récurrence immédiate :

$$\sup_{[t_0, t_+]} \|u^{n+1} - u^n\| \leq \frac{1}{2^n} \sup_{[t_0, t_+]} \|u^1 - u^0\|.$$

On en déduit aisément que la suite $(u^n)_{n \in \mathbb{N}}$ est de Cauchy.

L'étape 3 est quasi-immédiate. Il faut vérifier que la limite u est solution de notre problème. Or, par passage à la limite dans l'égalité

$$u^{n+1}(t) = u_0 + \int_{t_0}^t f(\tau, u^n(\tau)) \, d\tau$$

on voit que

$$u(t) = u_0 + \int_{t_0}^t f(\tau, u(\tau)) \, d\tau.$$

Comme u est continue ainsi que f , le second membre de cette égalité est de classe \mathcal{C}^1 et donc u aussi. (Plus généralement, si f est de classe \mathcal{C}^r , $r \geq 0$, u est de classe \mathcal{C}^{r+1} .) Et u vérifie :

$$u'(t) = f(t, u(t)) \quad \text{pour tout } t \in [t_0, t_+].$$

Ceci achève la preuve de l'existence d'une solution.

La preuve de l'unicité locale est très facile grâce au lemme de Gronwall. En effet, si v est une autre solution du même problème de Cauchy, elle est à valeurs dans $\bar{B}(u_0, R)$ sur un intervalle $[t_0, t_1]$, $0 < t_1 \leq t_+$. Et on a

$$u(t) - v(t) = \int_{t_0}^t (f(\tau, u(\tau)) - f(\tau, v(\tau))) \, d\tau.$$

et donc

$$\|u(t) - v(t)\| \leq L \int_{t_0}^t \|u(\tau) - v(\tau)\| \, d\tau.$$

Par conséquent, une version simplissime du lemme de Gronwall implique :

$$\|u(t) - v(t)\| \leq e^{Lt} \|u(t_0) - v(t_0)\| = 0$$

pour tout $t \in [t_0, t_1]$. □

Remarques.

- On peut affiner la démonstration de l'existence, en supprimant la restriction sur le temps d'existence $(t_+ - t_0) L \leq \frac{1}{2}$. En effet, on montre facilement par récurrence que la suite u^n vérifie en fait l'inégalité :

$$\|u^{n+1}(t) - u^n(t)\| \leq \frac{L^n (t - t_0)^n}{n!} \sup_{[t_0, t_+]} \|u^1 - u^0\| \quad \text{pour } t \in [t_0, t_+].$$

Comme la série $\sum \frac{L^n (t_+ - t_0)^n}{n!}$ est convergente, on en déduit que (u^n) est de Cauchy dans $\mathcal{C}([t_0, t_+])$ sans avoir à diminuer t_+ .

- Le théorème s'applique en particulier aux fonctions f de classe \mathcal{C}^1 , qui sont localement lipschitziennes (d'après le théorème des accroissements finis).

2 Solutions maximales

Le théorème II.1 fournit des solutions locales. L'unicité de ces solutions permet de démontrer le résultat suivant.

Lemme II.1 *Sous les hypothèses du théorème II.1, si $u_1 \in \mathcal{C}^1(J_1; U)$ et $u_2 \in \mathcal{C}^1(J_2; U)$ sont deux solutions sur des intervalles J_1 et J_2 respectivement, et s'il existe $t_0 \in J_1 \cap J_2$ tel que $u_1(t_0) = u_2(t_0)$, alors*

$$u_1(t) = u_2(t), \quad \text{pour tout } t \in J_1 \cap J_2.$$

Dém. C'est une application classique de la notion de *connexité*. On remarque que $J_1 \cap J_2$ est un intervalle non vide par hypothèse et on considère l'ensemble :

$$A = \{t \in J_1 \cap J_2; u_1(t) = u_2(t)\}.$$

D'après l'unicité locale des solutions, A est un ouvert. De plus, il est clairement fermé par continuité de u_1 et u_2 . Donc A est égal à $J_1 \cap J_2$. \square

Ce lemme montre qu'il existe un plus grand intervalle J sur lequel le problème de Cauchy (3) admet une solution, et que cette solution est unique. Cette solution est appelée *solution maximale* : par définition, on ne peut pas la prolonger à $I \setminus J$. Lorsque $J = I$ on dit que cette solution est *globale*.

La question naturelle est ensuite de savoir à quelle(s) condition(s) une solution maximale est globale.

Théorème II.2 *Sous les hypothèses du théorème II.1, soit $u \in \mathcal{C}^1(J; U)$ une solution maximale. Notons b la borne supérieure de I et β la borne supérieure de J . Alors ou bien $\beta = b$ ou bien « u sort de tout compact » de U , c'est-à-dire que pour tout compact $K \subset U$, il existe $\eta < \beta$ tel que*

$$u(t) \in U \setminus K, \quad \text{pour } t \geq \eta \text{ avec } t \in J.$$

Et on a le résultat analogue pour les bornes inférieures.

Dém. Supposons $\beta < b$ et raisonnons par l'absurde. S'il existait un compact K et une suite t_n tendant vers β telle que $u(t_n) \in K$ pour tout n , quitte à en extraire une sous-suite, on pourrait supposer qu'elle converge vers $\underline{u} \in K$. Soient $\underline{\tau} > 0$ et $R > 0$ tel que f soit bornée et Lipschitzienne par rapport à u dans $[\beta - 2\underline{\tau}, \beta + 2\underline{\tau}] \times \bar{B}(\underline{u}, 2R)$. Alors en vertu du lemme II.2 ci-après, il existe $\tau \leq \underline{\tau}$ tel que pour tout $(t_0, u_0) \in [\beta - \tau, \beta + \tau] \times \bar{B}(\underline{u}, R)$ la solution maximale du problème de Cauchy (3) est définie sur un intervalle contenant $[t_0 - \tau, t_0 + \tau]$. Or, pour n assez grand, $(t_n, u(t_n)) \in [\beta - \tau, \beta + \tau] \times \bar{B}(\underline{u}, R)$, et $t_n + \tau > \beta$. Ceci contredit le fait que u soit une solution maximale. \square

Lemme II.2 *Supposons que f soit continue, bornée et Lipschitzienne par rapport à u dans $[\underline{t} - 2\underline{\tau}, \underline{t} + 2\underline{\tau}] \times \bar{B}(\underline{u}, 2R)$ pour $\underline{\tau} > 0$ et $R > 0$. Alors il existe $\tau \in]0, \underline{\tau}]$ tel que pour tout $(t_0, u_0) \in [\underline{t} - \tau, \underline{t} + \tau] \times \bar{B}(\underline{u}, R)$, la solution maximale du problème de Cauchy (3) soit définie sur un intervalle contenant $[t_0 - \tau, t_0 + \tau]$.*

Dém. Il suffit de reprendre la démonstration du théorème II.1 : en effet, la méthode suivie permet de construire une solution dans l'intervalle $[t_0 - \tau, t_0 + \tau]$ qui soit à valeurs dans $\bar{B}(u_0, R)$ (donc *a fortiori* à valeurs dans $\bar{B}(\underline{u}, 2R)$ par l'inégalité triangulaire), pour tout $\tau > 0$ et inférieur à

$$\min\left(\tau, \frac{R}{M}, \frac{1}{2L}\right),$$

où M est une borne de f et L une constante de Lipschitz de f sur $[\underline{t} - 2\tau, \underline{t} + 2\tau] \times \bar{B}(\underline{u}, 2R)$. (Comme on l'a remarqué plus haut, la contrainte $\tau \leq 1/(2L)$ n'est en fait pas nécessaire : mais le fait que la constante de Lipschitz L existe est crucial !) \square

3 Existence globale

On peut parfois montrer que toutes les solutions maximales sont globales. C'est le cas si la fonction f est définie sur X tout entier et si elle est *globalement* Lipschizienne : car alors il n'y a pas de risque de sortir de son domaine de définition, ni du domaine de validité de sa constante de Lipschitz.

Théorème II.3 *On suppose $f \in \mathcal{C}(I \times X; X)$ et globalement Lipschizienne par rapport à u . Alors, quel que soit $(t_0, u_0) \in I \times X$, il existe (un unique) $u \in \mathcal{C}^1(I; X)$ solution de (3).*

Dém. Reprendre la démonstration du théorème II.1 et la remarque qui le suit : en utilisant la constante de Lipschitz globale de f , quels que soient a, b tels que $t_0 \in [a, b] \subset I$, on construit une suite de solutions approchées (u^n) qui est de Cauchy dans $\mathcal{C}([a, b]; X)$. \square

Attention ! L'hypothèse de ce théorème est très forte. En général, il ne s'applique pas à des fonctions f non-linéaires. (Penser à l'exemple de l'équation de Riccati, où $I = \mathbb{R}$, $X = \mathbb{R}$ et $f(t, u) = u^2$: seule la solution nulle est globale.) En revanche, il s'applique aux fonctions f affines : si $b \in \mathcal{C}(I; X)$ et $A \in \mathcal{C}(I; \mathcal{L}(X))$ (où $\mathcal{L}(X)$ désigne l'espace des application *linéaires continues* de X dans X), toutes les solutions maximales de

$$\frac{du}{dt} = A(t)u + b(t)$$

sont globales. On reviendra plus en détails sur ce type d'équations au chapitre III.

4 Dépendance par rapport aux données initiales

On a entrevu dans le lemme II.2 ce que l'on peut attendre lorsqu'on fait varier la condition initiale. En fait, on peut préciser ce lemme, pour comparer effectivement deux solutions de données initiales distinctes.

Lemme II.3 *Dans le cadre du lemme II.2, il existe $C > 0$ tel que si $(t_i, v_i) \in [\underline{t} - \tau, \underline{t} + \tau] \times \bar{B}(\underline{u}, R)$ pour $i = 1$ ou 2 , avec $|t_2 - t_1| \leq \tau$, les solutions u_i des problèmes de Cauchy correspondants vérifient l'estimation :*

$$\|u_1(t) - u_2(t)\| \leq C \|v_1 - v_2\| + C |t_1 - t_2| \quad \text{pour tout } t \in [t_1 - \tau, t_1 + \tau] \cap [t_2 - \tau, t_2 + \tau].$$

Dém. Par hypothèse, l'intervalle $[t_1 - \tau, t_1 + \tau] \cap [t_2 - \tau, t_2 + \tau]$ est non vide et contient l'intervalle d'extrémités t_1 et t_2 . En faisant la différence des deux relations :

$$u_i(t) = v_i + \int_{t_i}^t f(\tau, u_i(\tau)) \, d\tau,$$

on obtient l'inégalité :

$$\|u_1(t) - u_2(t)\| \leq \|v_1 - v_2\| + M|t_1 - t_2| + L \int_{t_1}^t \|u_1(s) - u_2(s)\| \, ds,$$

d'où par le lemme I.1 :

$$\|u_1(t) - u_2(t)\| \leq e^{L\tau} (\|v_1 - v_2\| + M|t_1 - t_2|).$$

□

Théorème II.4 (du flot) *On suppose f de classe C^2 sur $I \times U$, où I est un intervalle ouvert de \mathbb{R} et U est un ouvert de X . Soit $(t_0, u_0) \in I \times U$. Il existe un voisinage $\mathcal{W} \times \mathcal{V} = [t_0 - \tau, t_0 + \tau] \times \bar{B}(u_0, R)$ de (t_0, u_0) dans $I \times U$ et une unique application $\phi^{t_0} \in C^1(\mathcal{W} \times \mathcal{V}; U)$ telle que*

$$\begin{cases} \frac{\partial \phi^{t_0}}{\partial t}(t, v) = f(t, \phi^{t_0}(t, v)), & \text{pour tout } (t, v) \in \mathcal{W} \times \mathcal{V}, \\ \phi^{t_0}(t_0, v) = v, & \text{pour tout } v \in \mathcal{V}. \end{cases}$$

La fonction ϕ^{t_0} est appelée *flot* (local) au point t_0 de l'équation différentielle. En particulier, ce théorème affirme que pour toute donnée initiale v proche de u_0 l'application $u : t \mapsto \phi^{t_0}(t, v)$ est solution du problème de Cauchy

$$\begin{cases} u'(t) = f(t, u(t)), & \text{pour tout } t \in [t_0 - \tau, t_0 + \tau], \\ u(t_0) = v, \end{cases}$$

et qu'en plus elle dépend de manière C^1 de v . Lorsqu'il n'y a pas d'ambiguïté, on écrira simplement ϕ au lieu de ϕ^{t_0} .

La démonstration du théorème II.4 repose sur le calcul suivant, formel pour l'instant. Supposons que l'on ait construit ϕ et que l'on puisse dériver l'égalité

$$\frac{\partial \phi}{\partial t}(t, v) = f(t, \phi(t, v))$$

par rapport à v . Désignons par D la différentiation dans X , et soit $\psi(t, v) = D\phi(t, v)$ (par définition, c'est une application linéaire continue X dans X). Alors, en permutant $\frac{\partial}{\partial t}$ avec D (on suppose les fonctions suffisamment régulières pour pouvoir appliquer le lemme de Schwarz), on obtient :

$$\frac{\partial \psi}{\partial t}(t, v) = Df(t, \phi(t, v)) \circ \psi(t, v).$$

En supposant ϕ connue, ceci est une équation différentielle *linéaire* par rapport à ψ ! De plus, en dérivant la relation $\phi(t_0, v) = v$ on obtient la condition initiale

$$\psi(t_0, v) = \mathbf{I}_X.$$

Dém. [Théorème du flot] Pour $v \in \mathcal{V} := \bar{B}(u_0, R)$ avec $\bar{B}(u_0, 2R) \subset U$, notons

$$\begin{aligned} \phi(\cdot, v) : \mathcal{W} := [t_0 - \tau, t_0 + \tau] &\rightarrow \bar{B}(v, R) \\ t &\mapsto \phi(t, v) = u(t) \end{aligned}$$

la solution de l'équation différentielle construite dans le lemme II.2 pour la donnée initiale v . Alors ϕ est de classe \mathcal{C}^1 par rapport à t , et d'après le lemme II.3 elle est Lipschitzienne par rapport à v avec une constante de Lipschitz uniforme sur $\mathcal{W} \times \mathcal{V}$. Elle est en particulier continue comme fonction de (t, v) . Donc l'application

$$(t, v) \mapsto Df(t, \phi(t, v))$$

est aussi continue. Par suite, pour tout $v \in \mathcal{V}$ l'équation différentielle *linéaire*, appelée *équation aux variations* :

$$(3) \quad \frac{d\Psi}{dt} = Df(t, \phi(t, v)) \circ \Psi,$$

admet une solution unique sur \mathcal{W} tout entier satisfaisant la condition initiale $\Psi(t_0) = I_X$. Notons $\psi(t, v) = \Psi(t)$ cette solution à l'instant t . On veut montrer que ϕ est continûment différentiable par rapport à v et $\psi(t, v) = D\phi(t, v)$.

La première étape de la démonstration consiste à montrer que ψ dépend continûment de (t, v) . D'après l'inégalité triangulaire, on a

$$\|\psi(t, v) - \psi(\sigma, w)\| \leq \|\psi(t, v) - \psi(\sigma, v)\| + \|\psi(\sigma, v) - \psi(\sigma, w)\|.$$

Le premier terme se majore directement par :

$$\begin{aligned} \|\psi(t, v) - \psi(\sigma, v)\| &\leq |t - \sigma| \max_{s \in [\sigma, t]} \|Df(s, \phi(s, v))\| \max_{s \in [\sigma, t]} \|\psi(s, v)\| \\ &\leq |t - \sigma| a(\sigma, t; v) e^{a(\sigma, t; v) \max(|\sigma|, |t|)} \end{aligned}$$

par lemme de Gronwall, avec

$$a(\sigma, t; v) := \max_{s \in [\sigma, t]} \|Df(s, \phi(s, v))\|.$$

D'autre part, on peut aussi majorer le second terme grâce au lemme de Gronwall. Après calcul (voir le lemme II.4 ci-après), on obtient :

$$\|\psi(\sigma, v) - \psi(\sigma, w)\| \leq b(\sigma; v, w) |\sigma| e^{\alpha(\sigma; v, w) |\sigma|}$$

où

$$\begin{aligned} b(\sigma; v, w) &:= \|Df(\sigma, \phi(\sigma, v)) - Df(\sigma, \phi(\sigma, w))\|, \\ \alpha(\sigma; v, w) &:= \max(\|Df(\sigma, \phi(\sigma, v))\|, \|Df(\sigma, \phi(\sigma, w))\|). \end{aligned}$$

Par continuité de Df et de ϕ , les fonctions a et α sont localement bornées, tandis que $b(\sigma; v, w)$ converge vers 0 lorsque v tend vers w . Donc $\|\psi(t, v) - \psi(\sigma, w)\|$ tend vers 0 lorsque (t, v) tend vers (σ, w) .

La seconde étape est de montrer que $D\phi(t, v)$ existe et coïncide avec $\psi(t, v)$. Considérons pour cela l'accroissement :

$$\theta(t, h) := \phi(t, v + h) - \phi(t, v).$$

D'après le lemme II.3,

$$\|\theta(t, h)\| \leq C \|h\|$$

si v et $v + h$ sont dans \mathcal{V} . Et on a par définition de ϕ ,

$$\theta(t_0, h) = v + h - v = h.$$

Donc

$$\theta(t, h) = h + \int_{t_0}^t (f(s, \phi(s, v + h)) - f(s, \phi(s, v))) \, ds.$$

D'autre part, par définition de ψ :

$$\psi(t, v) \cdot h = h + \int_{t_0}^t Df(s, \phi(s, v)) \psi(s, v) \cdot h \, ds.$$

D'où, par soustraction et inégalité triangulaire :

$$\begin{aligned} \|\theta(t, h) - \psi(t, v) \cdot h\| &\leq \int_{t_0}^t \|f(s, \phi(s, v + h)) - f(s, \phi(s, v)) - Df(s, \phi(s, v)) \cdot \theta(s, h)\| \, ds \\ &\quad + \int_{t_0}^t \|Df(s, \phi(s, v)) (\theta(s, h) - \psi(s, v) \cdot h)\| \, ds. \end{aligned}$$

Le premier terme se majore en utilisant le fait que f est de classe \mathcal{C}^2 (on aurait besoin seulement de \mathcal{C}^1 en dimension finie, puisqu'alors Df serait localement uniformément continue). Quitte à réduire R on peut en effet supposer D^2f bornée sur $[t_0 - \tau, t_0 + \tau] \times \bar{B}(u_0; R)$, et donc, d'après le théorème des accroissements finis, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour $x, y \in \bar{B}(u_0; 2R)$, $\|y - x\| \leq \eta$ entraîne

$$\|f(s, x) - f(s, y) - Df(s, y) \cdot (x - y)\| \leq \varepsilon \|x - y\|$$

pour tout $s \in [t_0, t]$. Donc, pour $\|h\| \leq \eta/C$, on a

$$\|f(s, \phi(s, v + h)) - f(s, \phi(s, v)) - Df(s, \phi(s, v)) \cdot \theta(s, h)\| \leq \varepsilon C \|h\|.$$

D'où pour tout $t \geq t_0$:

$$\|\theta(t, h) - \psi(t, v) \cdot h\| \leq \varepsilon C \|h\| (t - t_0) + a(t_0, t; v) \int_{t_0}^t \|\theta(s, h) - \psi(s, v) \cdot h\| \, ds$$

avec la même signification pour a que dans la première étape. On appliquant une nouvelle fois le lemme de Gronwall, on déduit pour $t \in [t_0 - \tau, t_0 + \tau]$:

$$\|\theta(t, h) - \psi(t, v) \cdot h\| \leq \varepsilon C \|h\| \tau e^{a(t_0 - \tau, t_0 + \tau; v) \tau}.$$

Ceci montre que

$$\lim_{h \rightarrow 0} \frac{1}{\|h\|} (\phi(t, v + h) - \phi(t, v) - \psi(t, v) \cdot h) = 0,$$

c'est-à-dire que ϕ est différentiable par rapport à v et $D\phi(t, v) = \psi(t, v)$.

En conclusion, ϕ est continûment dérivable par rapport à t et à v et donc de classe \mathcal{C}^1 . \square

En prime, on a montré que la différentielle de ϕ par rapport à v , $D\phi$ est de classe \mathcal{C}^1 et solution de l'équation linéaire (3) avec la condition initiale

$$D\phi(t_0, v) = \mathbf{I}_X.$$

Pour compléter, montrons maintenant le lemme utilisé dans la démonstration.

Lemme II.4 Si $A : [t_0 - \tau, t_0 + \tau] \times \mathcal{V} \rightarrow \mathcal{L}(X)$ (l'espace des applications linéaires continues de X dans X) est continue (par rapport à sa première variable t), alors les solutions $t \mapsto \psi(t, v)$ de l'équation linéaire

$$\frac{\partial \psi}{\partial t}(t, v) = A(t, v) \psi(t, v)$$

telles que $\psi(t_0, v) = \psi_0$ satisfont l'estimation

$$\begin{aligned} \|\psi(t, v) - \psi(t, w)\| &\leq \|\psi_0\| \tau \max_{s \in [t_0 - \tau, t_0 + \tau]} \|A(s, v) - A(s, w)\| \\ &\quad \times \exp\left(\tau \max_{s \in [t_0 - \tau, t_0 + \tau]} \max(\|A(s, v)\|, \|A(s, w)\|)\right) \end{aligned}$$

quels que soient $t \in [t_0 - \tau, t_0 + \tau]$, et $v, w \in \mathcal{V}$.

Dém. Quitte à faire une translation en t on peut supposer $t_0 = 0$. Et on peut faire la preuve pour $t \geq t_0$ (le cas $t \leq t_0$ s'en déduisant par symétrie). Enfin, pour simplifier les écritures, introduisons les notations

$$a(t, v) := \max_{s \in [0, t]} \|A(s, v)\| \quad \text{et} \quad b(t, v, w) := \max_{s \in [0, t]} \|A(s, v) - A(s, w)\|.$$

Puisque par définition,

$$\psi(t, v) - \psi(t, w) = \int_0^t A(s, v) \psi(s, v) - A(s, w) \psi(s, w) ds,$$

on a pour tout $t \in [0, \tau]$,

$$\|\psi(t, v) - \psi(t, w)\| \leq a(\tau, v) \int_0^t \|\psi(s, v) - \psi(s, w)\| ds + b(\tau, v, w) \int_0^t \|\psi(s, w)\| ds.$$

De plus,

$$\|\psi(t, w)\| \leq \|\psi_0\| + a(\tau, w) \int_0^t \|\psi(s, w)\| ds.$$

Donc, par une première application du lemme de Gronwall,

$$\|\psi(t, w)\| \leq \|\psi_0\| e^{t a(\tau, w)}.$$

On en déduit

$$\|\psi(t, v) - \psi(t, w)\| \leq \int_0^t a(\tau, v) \|\psi(s, v) - \psi(s, w)\| + b(\tau, v, w) e^{t a(\tau, w)} \|\psi_0\| ds.$$

En appliquant à nouveau le lemme de Gronwall, ceci donne

$$\begin{aligned} \|\psi(t, v) - \psi(t, w)\| &\leq b(\tau, v, w) \|\psi_0\| \int_0^t e^{s a(\tau, w)} e^{(t-s) a(\tau, v)} ds \\ &\leq b(\tau, v, w) \|\psi_0\| \tau e^{\tau \max(a(\tau, w), a(\tau, v))}, \end{aligned}$$

ce qui est l'inégalité annoncée (avec $t \in [t_0, t_0 + \tau]$ et $t_0 = 0$). \square

5 Dépendance par rapport à des paramètres

Connaissant le théorème du flot, on obtient très facilement un résultat pour des équations dépendant d'un paramètre :

Théorème II.5 *On suppose f de classe C^1 sur $I \times U \times \Lambda$, où I est un intervalle ouvert de \mathbb{R} , U est un ouvert de X et Λ un ouvert d'un autre espace de Banach Y . Les solutions de l'équation différentielle à paramètre :*

$$\frac{du}{dt} = f(t, u, \lambda)$$

sont continûment différentiables par rapport à λ .

Dém. Considérons le système *augmenté* :

$$\begin{cases} \frac{du}{dt} = f(t, u, \mu), \\ \frac{d\mu}{dt} = 0. \end{cases}$$

Soit Φ^{t_0} le flot local en t_0 de ce système. Alors la solution du problème de Cauchy :

$$\begin{cases} u'(t) = f(t, u(t), \lambda), \\ u(t_0) = v \end{cases}$$

est donnée par

$$u(t) = P(\Phi^{t_0}(t, v, \lambda))$$

où $P : X \times Y \rightarrow X$ désigne la projection sur X . D'après le théorème II.4, c'est une fonction continûment différentiable par rapport à (v, λ) et donc en particulier par rapport à λ . \square

Chapitre III

Équations linéaires

Dans ce chapitre, on s'intéresse exclusivement aux équations différentielles du type

$$(1) \quad \frac{du}{dt} = A(t)u + b(t),$$

avec $b \in \mathcal{C}(I; X)$ et $A \in \mathcal{C}(I; \mathcal{L}(X))$ (où $\mathcal{L}(X)$ désigne l'espace des application *linéaires continues* dans l'espace de Banach X). On dit que sont des équations « linéaires » (car $A(t)$ est linéaire) avec « terme source » (b). On qualifie parfois (1) simplement d'équation linéaire (par abus de langage, le terme correct étant plutôt affine).

1 Existence globale

D'après le théorème II.3, les solutions de (1) sont globales.

Théorème III.1 *Pour $b \in \mathcal{C}(I; X)$ et $A \in \mathcal{C}(I; \mathcal{L}(X))$, quel que soit $(t_0, u_0) \in I \times X$, il existe une unique fonction $u \in \mathcal{C}^1(I; X)$ solution de (1) telle que $u(t_0) = u_0$.*

On peut aussi donner une démonstration de ce théorème en faisant appel au théorème de point fixe suivant :

Théorème III.2 *Soit \mathcal{T} un opérateur sur un espace de Banach dont une puissance est contractante. Alors \mathcal{T} admet un point fixe unique.*

Dém. Ce résultat est une conséquence facile du théorème de point fixe de Banach-Picard. Par hypothèse, il existe un entier k tel que \mathcal{T}^k soit contractant et donc admette un point fixe unique u . Alors

$$\mathcal{T}^k(\mathcal{T}(u)) = \mathcal{T}(\mathcal{T}^k(u)) = \mathcal{T}(u).$$

Donc $\mathcal{T}(u)$ est aussi un point fixe de \mathcal{T}^k . À cause de l'unicité de ce point fixe on a ainsi $\mathcal{T}(u) = u$. De plus, tout point fixe de \mathcal{T} étant évidemment un point fixe de \mathcal{T}^k , u est l'unique point fixe de \mathcal{T} . \square

Démonstration du théorème III.1. Pour résoudre le problème de Cauchy de donnée « initiale » $u(t_0) = u_0$, on cherche u tel que

$$u(t) = u_0 + \int_{t_0}^t (A(s) u(s) + b(s)) ds$$

c'est-à-dire de façon équivalente

$$u(t) = v(t) + u_0, \quad \text{avec} \quad v(t) = \int_{t_0}^t (A(s) (v(s) + u_0) + b(s)) ds.$$

Pour simplifier, on suppose (sans perte de généralité, il suffit de translater toutes les fonctions en jeu) $t_0 = 0$. Soient α et β tels que $\alpha < 0 < \beta$ et $[\alpha, \beta] \subset I$. Définissons alors l'opérateur :

$$\begin{aligned} \mathcal{T} : \mathcal{C}([\alpha, \beta]; X) &\rightarrow \mathcal{C}([\alpha, \beta]; X) \\ v &\mapsto \mathcal{T}v; (\mathcal{T}v)(t) = \int_0^t (A(s) (v(s) + u_0) + b(s)) ds. \end{aligned}$$

Quels que soient v et w dans $\mathcal{C}([\alpha, \beta]; X)$ on a

$$(\mathcal{T}v)(t) - (\mathcal{T}w)(t) = \int_0^t A(s) (v(s) - w(s)) ds$$

et donc

$$\|(\mathcal{T}v)(t) - (\mathcal{T}w)(t)\| \leq C t \max_{s \in [0, t]} \|v - w\|, \quad \text{avec} \quad C := \max_{s \in [\alpha, \beta]} \|A(s)\|.$$

Montrons par récurrence que

$$\|(\mathcal{T}^n v)(t) - (\mathcal{T}^n w)(t)\| \leq C^n \frac{t^n}{n!} \max_{s \in [0, t]} \|v - w\|$$

pour tout $t \in [\alpha, \beta]$. C'est vrai à l'ordre 1 comme on vient de le voir. Supposons l'inégalité vraie à l'ordre n . Alors

$$\begin{aligned} \|(\mathcal{T}^{n+1} v)(t) - (\mathcal{T}^{n+1} w)(t)\| &= \|\mathcal{T}((\mathcal{T}^n v)(t)) - \mathcal{T}((\mathcal{T}^n w)(t))\| \\ &\leq C \int_0^t \|(\mathcal{T}^n v)(s) - (\mathcal{T}^n w)(s)\| ds \leq C \max_{s \in [0, t]} \|v - w\| \int_0^t C^n \frac{s^n}{n!} ds \end{aligned}$$

d'après l'hypothèse de récurrence. En évaluant l'intégrale on obtient immédiatement l'inégalité à l'ordre $n + 1$:

$$\|(\mathcal{T}^{n+1} v)(t) - (\mathcal{T}^{n+1} w)(t)\| \leq C^{n+1} \frac{t^{n+1}}{(n+1)!} \|v - w\|.$$

Par conséquent, on a

$$\|\mathcal{T}^n v - \mathcal{T}^n w\| \leq C^n \frac{(\beta - \alpha)^n}{n!} \|v - w\|$$

quels que soient v et w . Donc \mathcal{T}^n est contractant pour n assez grand et d'après le théorème III.2, \mathcal{T} admet un point fixe unique. \square

2 Résolvante

Considérons l'équation différentielle linéaire dans $\mathcal{L}(X)$:

$$(2) \quad \frac{dM}{dt} = A(t) \circ M.$$

D'après le théorème III.1, ses solutions maximales sont globales.

Définition III.1 On appelle résolvante de l'équation différentielle

$$\frac{du}{dt} = A(t)u$$

l'application

$$R : I \times I \rightarrow \mathcal{L}(X) \\ (t, t_0) \mapsto R(t, t_0)$$

où $t \mapsto R(t, t_0)$ est la solution du problème de Cauchy pour (2) avec la condition initiale $R(t_0, t_0) = \mathbf{I}_X$.

On observe que pour tout $u_0 \in X$, l'application $u : t \mapsto R(t, t_0)u_0$ est la solution du problème de Cauchy pour (1) et la condition initiale $u(t_0) = u_0$. Autrement dit, le flot de (1) est donné par

$$\phi^{t_0}(t, u_0) = R(t, t_0)u_0.$$

Par construction, R est continûment différentiable par rapport à t . De plus, le lemme II.3 (application simple du lemme de Gronwall) montre que R est Lipschitzienne par rapport à t_0 . Donc en particulier, R est une fonction continue de (t, t_0) . Comme l'ensemble des isomorphismes de $\mathcal{L}(X)$ est un ouvert et $R(t_0, t_0) = \mathbf{I}_X$ est trivialement un isomorphisme, $R(t, t_0)$ est un isomorphisme pour tout t voisin de t_0 . En fait, ceci est vrai pour tout $t \in I$ comme conséquence de la :

Proposition III.1 La résolvante est telle que

$$R(t, s) \circ R(s, t_0) = R(t, t_0) \quad \text{quels que soient } t, s, t_0 \in I.$$

Dém. Soit $u_0 \in X$ et $u(t) = R(t, t_0)u_0$. Alors la solution du problème de Cauchy

$$\begin{cases} \frac{dv}{dt} = A(t)v \\ v(s) = u(s) \end{cases}$$

n'est autre que u , mais elle s'exprime aussi comme

$$v : t \mapsto v(t) = R(t, s)u(s) = R(t, s) \circ R(s, t_0)u_0.$$

D'où

$$R(t, t_0)u_0 = R(t, s) \circ R(s, t_0)u_0.$$

Ceci étant vrai quel que soit u_0 , on en déduit la formule voulue. \square

Corollaire III.1 Pour tout $(t, t_0) \in I \times I$, $R(t, t_0)$ est un isomorphisme, d'inverse $R(t_0, t)$. De plus, R est continûment différentiable comme fonction de deux variables, et ses dérivées partielles satisfont

$$\frac{\partial R}{\partial t}(t, s) = A(t) \circ R(t, s) \quad \text{et} \quad \frac{\partial R}{\partial s}(t, s) = -R(t, s) \circ A(s).$$

Dém. La première partie est conséquence immédiate de la proposition III.1 :

$$R(t_0, t) \circ R(t, t_0) = R(t_0, t_0) = \mathbf{I}_X \quad \text{et} \quad R(t, t_0) \circ R(t_0, t) = R(t, t) = \mathbf{I}_X.$$

L'expression de la dérivée partielle de R par rapport à sa première variable découle directement de la définition de R : elle est continue comme composée de fonctions continues (on rappelle que R est continue comme fonction des deux variables). Pour montrer que R est dérivable par rapport à sa seconde variable, on utilise la première partie : comme $R(t, s) = R(s, t)^{-1}$, R est dérivable par rapport à s par composition de $s \mapsto R(s, t)$ et de l'application

$$\begin{aligned} \text{Isom}(X) &\rightarrow \text{Isom}(X) \\ M &\mapsto M^{-1}, \end{aligned}$$

dont on sait qu'elle est (continûment) différentiable. Pour exprimer effectivement la dérivée partielle de R par rapport à sa seconde variable, on dérive simplement l'identité

$$R(t, s) \circ R(s, t) = \mathbf{I}_X$$

en utilisant l'expression de la dérivée partielle de R par rapport à sa première variable, ce qui donne :

$$\frac{\partial R}{\partial s}(t, s) \circ R(s, t) + R(t, s) \circ A(s) \circ R(s, t) = 0.$$

Comme $R(s, t)$ est un isomorphisme, on en déduit l'expression annoncée. \square

Attention ! En général, on ne connaît pas explicitement la résolvante.

2.1 Cas des équations linéaires autonomes

Si A est indépendant de t , la résolvante de (1) s'exprime à l'aide de l'exponentielle, dont on rappelle la

Définition III.2 Pour tout $A \in \mathcal{L}(X)$,

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

(La série ci-dessus est normalement convergente dans $\mathcal{L}(X)$.)

Lemme III.1 Pour tout $A \in \mathcal{L}(X)$, l'application $t \mapsto e^{tA}$ est continûment dérivable (et même de classe C^∞), et

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A.$$

La démonstration est laissée en exercice. Elle utilise de façon cruciale la propriété de l'exponentielle :

$$e^{(t+s)A} = e^{tA} e^{sA}.$$

Le lemme III.3 montre que la résolvente de l'équation différentielle linéaire *autonome*

$$\frac{du}{dt} = Au$$

est donnée par

$$R(t, s) = e^{(t-s)A}.$$

2.2 Cas des équations linéaires en dimension finie

Dimension 1. Les équations différentielles (d'ordre 1) linéaires scalaires, pour lesquelles $X = \mathbb{R}$, se résolvent explicitement : la résolvente de

$$\frac{du}{dt} = a(t)u$$

où l'inconnue u est à valeurs réelles et $a : I \rightarrow \mathbb{R}$ est continue est donnée par

$$R(t, s) = e^{\int_s^t a(\tau) d\tau}.$$

Dimension 2 et plus. Il est tentant de généraliser la formule précédente. Cependant, si les matrices $A(t)$ (ou les endomorphismes) ne commutent pas deux à deux, une telle formule (que l'on se garde volontairement d'écrire) est *fausse* : ceci est lié au fait que e^{A+B} n'est *pas* le produit de e^A et e^B lorsque les matrices A et B ne commutent pas.

Tout ce que l'on peut dire est que la résolvente est liée à une base de solutions.

Proposition III.2 Soit $A : I \rightarrow \mathbf{M}_n(\mathbb{R})$ continue, et $R : I \rightarrow \mathbf{M}_n(\mathbb{R})$ la résolvente de

$$(3) \quad \frac{du}{dt} = A(t)u,$$

c'est-à-dire que pour tout $(t, s) \in I \times I$,

$$\frac{\partial R}{\partial t}(t, s) = A(t)R(t, s) \quad \text{et} \quad R(s, s) = \mathbf{I}_n.$$

Si (e_1, \dots, e_n) est une base de \mathbb{R}^n , alors les applications $t \mapsto R(t, t_0)e_i$ (pour $i \in \{1, \dots, n\}$) forment une famille indépendante de solutions de (3). Inversement, si u_1, \dots, u_n est une famille indépendante de solutions de (3), soit $B : t \in I \mapsto B(t) \in \mathbf{GL}_n(\mathbb{R})$ dont les vecteurs colonnes sont $u_1(t), \dots, u_n(t)$. Alors la résolvente de (3) est donnée par

$$R(t, s) = B(t)B(s)^{-1}.$$

Dém. La première partie découle de la définition et du corollaire III.1, qui montre que $R(t, s) \in \mathbf{GL}_n(\mathbb{R})$. La preuve de la seconde partie consiste simplement à remarquer que l'application $t \mapsto B(t)B(s)^{-1}$ est solution du même problème de Cauchy que $t \mapsto R(t, s)$. \square

Remarque. On peut aussi montrer que $R(t, s)$ est inversible en vérifiant que son déterminant satisfait une équation différentielle scalaire linéaire (d'ordre 1). C'est l'objet du résultat suivant, parfois appelé formule de Liouville.

Proposition III.3 Soit $\Delta(t, s) := \det(R(t, s))$. Alors

$$\Delta(t, s) = e^{\int_s^t \operatorname{tr}(A(\tau)) d\tau}.$$

Dém. C'est un petit calcul qui utilise la différentielle du déterminant, dont on rappelle qu'elle est donnée par :

$$d(\det)(M) \cdot H = \operatorname{tr}((\operatorname{com}M)^t H).$$

Par conséquent, la règle de dérivation des fonctions composées montre que :

$$\frac{\partial \Delta}{\partial t}(t, s) = \operatorname{tr}\left((\operatorname{com}R(t, s))^t \frac{\partial R}{\partial t}(t, s)\right).$$

D'où

$$\frac{\partial \Delta}{\partial t}(t, s) = \operatorname{tr}\left((\operatorname{com}R(t, s))^t A(t) R(t, s)\right) = \operatorname{tr}\left((\operatorname{com}R(t, s))^t R(t, s) A(t)\right) = \Delta(t, s) \operatorname{tr}(A(t))$$

d'après la formule $(\operatorname{com}R)^t R = (\det R) \mathbf{I}$. □

3 Formule de Duhamel

Ce paragraphe est consacré à un résultat très facile à démontrer et néanmoins fondamental. Il exprime le principe de superposition bien connu selon lequel : *la solution générale d'une équation linéaire avec terme source est donnée par la solution générale de l'équation homogène plus une solution particulière de l'équation avec terme source.*

Proposition III.4 (Formule de Duhamel) Soit $A \in \mathcal{C}(I; \mathcal{L}(X))$ et R la résolvante de l'équation différentielle homogène :

$$\frac{du}{dt} = A(t)u$$

Soit $b \in \mathcal{C}(I; X)$ et ϕ^{t_0} le flot en $t_0 \in I$ de l'équation

$$\frac{du}{dt} = A(t)u + b(t).$$

Alors pour tout $(t, v) \in I \times X$,

$$\phi^{t_0}(t, v) = R(t, t_0) \cdot v + \int_{t_0}^t R(t, s) \cdot b(s) ds.$$

Dém. On vérifie par un calcul facile que l'application

$$t \mapsto \int_{t_0}^t R(t, s) \cdot b(s) ds$$

est solution de l'équation avec le terme source b . Comme $t \mapsto R(t, t_0) \cdot v$ est solution de l'équation homogène, la somme des deux est, par linéarité, solution de l'équation avec terme source. De plus elle vaut v à $t = t_0$. Donc c'est l'unique solution cherchée. □

4 Équations linéaires autonomes

L'objectif ici est d'introduire certains outils utiles dans la suite du cours. On commence par le cadre plus élémentaire des équations en dimension finie.

4.1 Dimension finie

On considère dans ce paragraphe $A \in \mathbf{M}_n(\mathbb{R})$ et l'équation différentielle linéaire homogène associée

$$(4) \quad \frac{du}{dt} = Au.$$

On sait que les solutions sont définies sur \mathbb{R} tout entier. Que peut-on dire de leur comportement à l'infini ? Cela dépend à l'évidence des propriétés de la matrice A , que l'on va pour l'occasion identifier à un endomorphisme de \mathbb{C}^n . L'analyse de (4) est intimement liée à l'algèbre linéaire, dont voici quelques rappels (pour information le vocabulaire anglophone, lorsqu'il n'est pas intuitif, est donné en note).

Définition III.3 Les valeurs propres¹ de A sont les nombres complexes λ pour lesquels le noyau de $(A - \lambda \mathbf{I})$ est non trivial. Si λ est une valeur propre de A , on dit que l'espace $\text{Ker}(A - \lambda \mathbf{I})$ est le sous-espace propre² associé. Si k est le plus grand entier tel que

$$\text{Ker}(A - \lambda \mathbf{I})^{k-1} \subset \text{Ker}(A - \lambda \mathbf{I})^k \quad \text{avec inclusion stricte,}$$

l'espace $\text{Ker}(A - \lambda \mathbf{I})^k$ est appelé sous-espace caractéristique³ de λ .

- La dimension du sous-espace propre $\text{Ker}(A - \lambda \mathbf{I})$ est appelée multiplicité géométrique de λ .
- La dimension du sous-espace caractéristique $\text{Ker}(A - \lambda \mathbf{I})^k$ est appelée multiplicité algébrique de λ .
- Lorsque les multiplicités géométrique et algébrique coïncident, la valeur propre λ est dite semi-simple.
- Lorsque les multiplicités géométrique et algébrique coïncident et valent 1, la valeur propre λ est dite simple.

Théorème III.3

- Les valeurs propres de A sont les zéros du polynôme caractéristique :

$$P(\lambda) := \det(\lambda \mathbf{I} - A).$$

La multiplicité algébrique d'une valeur propre est égale à son ordre comme zéro du polynôme caractéristique. Les sous-espaces caractéristiques sont en somme directe et leur somme est \mathbb{C}^n .

- **(Cayley-Hamilton)** Les valeurs propres de A sont aussi les zéros du polynôme minimal, défini comme le générateur unitaire de l'idéal des polynômes annulateurs de A .
- Si toutes les valeurs propres sont semi-simples, les sous-espaces caractéristiques et les sous-espaces propres sont égaux, et la matrice A est diagonalisable.

Pour la démonstration, se reporter à un ouvrage d'algèbre linéaire...

¹« eigenvalues »

²« eigenspace »

³« generalized eigenspace »

Attention ! Le vocabulaire se télescope un peu car on appelle aussi *résolvante* de A la matrice $(A - \lambda \mathbf{I})^{-1}$ lorsque λ n'est pas une valeur propre de A . Le terme est donc à manipuler avec précaution.

Le comportement asymptotique des solutions de (4) est donné par le résultat suivant, que l'on donne sur \mathbb{R}^+ et qui admet évidemment un analogue sur \mathbb{R}^- (changer t en $-t$ et A en $-A$).

Proposition III.5 *L'application $t \mapsto e^{tA}$ est bornée sur \mathbb{R}^+ si et seulement si les valeurs propres de A sont toutes de partie réelle négative ou nulle et si les valeurs propres imaginaires pures sont semi-simples. Elle tend vers 0 en $+\infty$ si et seulement si les valeurs propres de A sont toutes de partie réelle strictement négative.*

Dém. Soit $\{\lambda_1, \dots, \lambda_p\}$ l'ensemble des valeurs propres de A (avec $p \leq n$ et $\lambda_i \neq \lambda_j$ pour $i \neq j$). À chaque valeur propre λ_j on associe le sous-espace caractéristique E_j , et Π_j la projection sur E_j parallèlement à $\bigoplus_{i \neq j} E_i$. Par définition de E_j , il existe un entier k_j tel que

$$(A - \lambda_j \mathbf{I})^{k_j} \Pi_j = 0,$$

et d'après le théorème III.3, on a

$$\mathbf{I} = \sum_{j=1}^p \Pi_j.$$

Par suite

$$e^{tA} = \sum_{j=1}^p e^{tA} \Pi_j = \sum_{j=1}^p e^{t\lambda_j} e^{t(A - \lambda_j \mathbf{I})} \Pi_j = \sum_{j=1}^p e^{t\lambda_j} \sum_{k=0}^{k_j-1} \frac{t^k}{k!} (A - \lambda_j \mathbf{I})^k \Pi_j,$$

ce qui tend bien vers 0 en $+\infty$ si les λ_j sont toutes de partie réelle strictement négative. Par ailleurs, si les λ_j sont de partie réelle négative ou nulle et si les k_j valent tous 1, alors e^{tA} est bornée sur \mathbb{R}^+ . Les réciproques sont faciles à vérifier :

- S'il existe λ de partie réelle strictement positive tel que $Av = \lambda v$ pour un vecteur v non nul, alors

$$e^{tA} v = e^{t\lambda} v$$

est non borné.

- Si $\lambda \in i\mathbb{R}$ est une valeur propre non semi-simple, il existe des vecteurs non nuls v et w tels que

$$Av = \lambda v, \quad Aw = \lambda w + v.$$

On montre par récurrence que

$$A^m w = \lambda^m w + m\lambda^{m-1} v$$

pour tout entier naturel m , et donc

$$e^{tA} w = e^{t\lambda} (w + tv)$$

est de norme supérieure ou égale à $t\|v\| - \|w\|$ pour $t \geq 0$, ce qui est évidemment non borné.

- Enfin, si $\lambda \in i\mathbb{R}$ est une valeur propre même simple, $Av = \lambda v$ avec v non nul entraîne que

$$e^{tA} v = e^{t\lambda} v$$

est de norme égale à $\|v\|$, ce qui ne tend pas vers 0. □

Définition III.4 Une matrice n'ayant pas de valeur propre imaginaire pure est dite hyperbolique.

Proposition III.6 Soit $A \in \mathbf{M}_n(\mathbb{R})$ une matrice hyperbolique. Alors les ensembles

$$E^s = \{ w \in \mathbb{R}^n ; \lim_{t \rightarrow +\infty} e^{tA} w = 0 \}, \quad E^u = \{ w \in \mathbb{R}^n ; \lim_{t \rightarrow -\infty} e^{tA} w = 0 \}$$

sont des sous-espaces vectoriels de \mathbb{R}^n stables par A , appelés respectivement sous-espace stable et sous-espace instable, et il existe des projecteurs Π^u et Π^s commutant avec A tels que $\Pi^u + \Pi^s = \mathbf{I}$ satisfaisant les estimations suivantes. En notant $\sigma(A)$ l'ensemble des valeurs propres de A et

$$\beta = -\max\{\operatorname{Re} \lambda < 0 ; \lambda \in \sigma(A)\} \quad \text{et} \quad \gamma = \min\{\operatorname{Re} \lambda > 0 ; \lambda \in \sigma(A)\},$$

pour tout $\varepsilon > 0$, il existe des constantes b_ε et c_ε telles que

$$\|e^{tA} \Pi^s\| \leq b_\varepsilon e^{-(\beta-\varepsilon)t} \quad \forall t \geq 0 \quad \text{et} \quad \|e^{tA} \Pi^u\| \leq c_\varepsilon e^{(\gamma-\varepsilon)t} \quad \forall t \leq 0.$$

Dém. Le fait que E^s et E^u soient des sous-espaces vectoriels de \mathbb{R}^n stables par A est évident (rappelons que e^{tA} commute avec A). On va commencer par définir les projecteurs Π^u et Π^s , et on vérifiera ensuite que $E^s = \operatorname{Im}(\Pi^s)$ et $E^u = \operatorname{Im}(\Pi^u)$. On reprend pour cela les mêmes notations que dans la preuve de la proposition III.5. Soient

$$\Pi^s := \sum_{j; \operatorname{Re} \lambda_j < 0} \Pi_j \quad \text{et} \quad \Pi^u := \sum_{j; \operatorname{Re} \lambda_j > 0} \Pi_j.$$

Noter que ces projecteurs sont bien à valeurs dans \mathbb{R}^n : en effet, les valeurs propres non réelles de A sont deux à deux conjuguées, et $\Pi_{j'} = \overline{\Pi_j}$ si $\lambda_{j'} = \overline{\lambda_j}$ donc $\Pi_j + \Pi_{j'}$ est à valeurs dans \mathbb{R}^n .

Comme les Π_j , Π^s et Π^u sont bien des projecteurs commutant avec A . Et comme A n'a pas de valeur propre imaginaire pure on a

$$\Pi^u + \Pi^s = \sum_{j=1}^p \Pi_j = \mathbf{I}.$$

De plus, on a l'expression explicite :

$$e^{tA} \Pi^s = \sum_{j; \operatorname{Re} \lambda_j < 0} e^{t\lambda_j} \sum_{k=0}^{k_j-1} \frac{t^k}{k!} (A - \lambda_j \mathbf{I})^k \Pi^s,$$

d'où, pour $t \geq 0$,

$$\|e^{tA} \Pi^s\| \leq p \max_{j \leq p, k \leq K} \|(A - \lambda_j \mathbf{I})^k\| \|\Pi^s\| e^{-\beta t} \sum_{k=0}^K \frac{t^k}{k!}$$

avec $K = \max_j k_j - 1$. Le dernier terme, polynôme en t , étant majoré par une constante fois $e^{\varepsilon t}$ pour $\varepsilon > 0$, on en déduit l'estimation de $\|e^{tA} \Pi^s\|$. L'estimation de $\|e^{tA} \Pi^u\|$ s'obtient exactement de la même manière. On déduit directement de ces estimations les inclusions

$$\operatorname{Im}(\Pi^s) \subset E^s \quad \text{et} \quad \operatorname{Im}(\Pi^u) \subset E^u.$$

Réciproquement, soit par exemple $w \in E^s$. Comme Π^u commute avec A , il commute aussi avec e^{tA} , donc

$$\|\Pi^u w\| = \|e^{tA} \Pi^u e^{-tA} w\| \leq c_\varepsilon e^{(\gamma-\varepsilon)t} \|e^{-tA} w\|$$

pour tout $t \leq 0$. En prenant $\varepsilon < \gamma$, le second membre tend vers 0 lorsque tend t vers $-\infty$ (pour le dernier terme, on utilise simplement la définition de E^s). Donc $\Pi^u w = 0$, c'est-à-dire que $w = \Pi^s w \in \text{Im}(\Pi^s)$. Ainsi $E^s \subset \text{Im}(\Pi^s)$. On montre de façon analogue que E^u est inclus dans $\text{Im}(\Pi^u)$. \square

4.2 Dimension infinie

Pour simplifier la présentation, on suppose dans ce paragraphe que X est un \mathbb{C} -espace de Banach, et que A est un opérateur linéaire et continu sur X , ce que l'on note $A \in \mathcal{L}(X)$.

On a vu au paragraphe précédent des résultats sur le comportement de e^{tA} lorsque X est de dimension finie, qui reposent pour une large part sur la réduction des matrices. Lorsque X est de dimension infinie, tout devient plus délicat. Tout d'abord, le spectre de A n'est plus exclusivement constitué de valeurs propres.

Définition III.5 On appelle spectre de $A \in \mathcal{L}(X)$ l'ensemble $\sigma(A)$ des nombres complexes λ tels que $(A - \lambda I_X) \notin \text{Isom}(X)$ (ensemble des isomorphismes de X). Le spectre ponctuel est inclus dans le spectre : c'est l'ensemble des valeurs propres, c'est-à-dire des nombres $\lambda \in \mathbb{C}$ pour lesquels $(A - \lambda I_X)$ est non injectif.

Remarques.

- L'ensemble des isomorphismes étant un ouvert, l'ensemble résolvant de A :

$$\rho(A) := \mathbb{C} \setminus \sigma(A)$$

est ouvert, par continuité de l'application

$$\begin{aligned} \mathbb{C} &\rightarrow \mathcal{L}(X) \\ \lambda &\mapsto (A - \lambda I_X). \end{aligned}$$

- Le spectre d'un opérateur $A \in \mathcal{L}(X)$ est borné :

$$\sigma(A) \subset \{\lambda; |\lambda| \leq \|A\|\}.$$

En effet, si $|\lambda| > \|A\|$, $(A - \lambda I_X)$ est inversible d'inverse

$$(A - \lambda I_X)^{-1} = -\frac{1}{\lambda} \sum_{m=0}^{+\infty} \frac{1}{\lambda^m} A^m,$$

cette série étant normalement convergente. Plus précisément, on a (voir la preuve ci-après)

$$(5) \quad \sigma(A) \subset \{\lambda; |\lambda| \leq r_A := \lim \|A^m\|^{1/m}\} \quad \text{et} \quad \sigma(A) \cap \{\lambda; |\lambda| = r_A\} \neq \emptyset.$$

- Ainsi, $\sigma(A)$ est un compact non vide de \mathbb{C} .

Attention ! Si A était un opérateur non borné (c'est-à-dire linéaire mais pas continu), on pourrait encore définir son spectre mais il pourrait être vide ou non borné.

Démonstration de (5). C'est un peu subtil, car rien ne dit *a priori* que la limite de $\|A^m\|^{1/m}$ existe. Nous allons le prouver, en même temps que la *formule du rayon spectral* :

$$(6) \quad \lim_{m \rightarrow +\infty} \|A^m\|^{1/m} = \sup\{|\lambda|; \lambda \in \sigma(A)\}.$$

(Si on savait que $\sigma(A)$ est exclusivement constitué de valeurs propres, la preuve serait plus facile.) On sait déjà que $\sigma(A)$ est un compact, inclus dans le disque de centre 0 et de rayon $\|A\|$. Le critère de Cauchy sur la convergence des séries montre que

$$\{\lambda; |\lambda| > \limsup \|A^m\|^{1/m}\} \subset \rho(A),$$

d'où

$$\sigma(A) \subset \{\lambda; |\lambda| \leq \limsup \|A^m\|^{1/m}\}.$$

En fait, on a mieux presque sans effort. En effet, si $\lambda \in \sigma(A)$ alors $\lambda^m \in \sigma(A^m)$. Ceci résulte simplement de la formule

$$(A^m - \lambda^m I_X) = (A - \lambda I_X) \sum_{k=0}^{m-1} \lambda^k A^{m-1-k}.$$

On en déduit

$$|\lambda| \leq \|A^m\|^{1/m} \quad \text{pour tout } \lambda \in \sigma(A) \quad \text{et pour tout } m \in \mathbb{N},$$

d'où

$$\sigma(A) \subset \{\lambda; |\lambda| \leq \liminf_{m \rightarrow +\infty} \|A^m\|^{1/m}\}.$$

On aura effectivement démontré (6), si l'on vérifie que :

- premièrement le spectre $\sigma(A)$ est non vide, ce qui permet de définir

$$s_A := \sup\{|\lambda|; \lambda \in \sigma(A)\},$$

- deuxièmement, $\limsup_{m \rightarrow +\infty} \|A^m\|^{1/m} \leq \sigma$ pour tout $\sigma > s_A$.

Car alors on aura la suite d'inégalités

$$s_A \leq \liminf_{m \rightarrow +\infty} \|A^m\|^{1/m} \leq \limsup_{m \rightarrow +\infty} \|A^m\|^{1/m} \leq s_A,$$

d'où $s_A = \liminf_{m \rightarrow +\infty} \|A^m\|^{1/m} = \limsup_{m \rightarrow +\infty} \|A^m\|^{1/m}$. Ceci implique évidemment (5) (pour la seconde partie de (5), on dit simplement que le sup de la fonction continue $\lambda \mapsto |\lambda|$ est nécessairement atteint sur le compact non vide $\sigma(A)$; autrement dit : il existe $\lambda \in \sigma(A)$ sur le cercle de centre 0 et de rayon $s_A = r_A$!)

La démonstration des deux propriétés énoncées ci-dessus utilise le

Lemme III.2 *L'application*

$$\begin{aligned} \mathbf{R}_A : \rho(A) &\rightarrow \mathcal{L}(X) \\ \lambda &\mapsto (A - \lambda I_X)^{-1} \end{aligned}$$

est holomorphe (c'est-à-dire dérivable comme fonction de variable complexe).

Dém. On sait déjà que cette application est continue, comme composée des applications continues $\lambda \in \mathbb{C} \mapsto A - \lambda \mathbf{I}_X$ et $u \in \text{Isom}(X) \mapsto u^{-1}$. De plus, pour λ et λ_0 dans l'ensemble résolvant $\rho(A)$,

$$(7) \quad (A - \lambda \mathbf{I}_X)^{-1} - (A - \lambda_0 \mathbf{I}_X)^{-1} = (\lambda - \lambda_0) (A - \lambda \mathbf{I}_X)^{-1} (A - \lambda_0 \mathbf{I}_X)^{-1}.$$

Ceci résulte du calcul facile suivant :

$$\begin{aligned} & (\lambda - \lambda_0) (A - \lambda \mathbf{I}_X)^{-1} (A - \lambda_0 \mathbf{I}_X)^{-1} = \\ & (A - \lambda \mathbf{I}_X)^{-1} (A - \lambda_0 \mathbf{I}_X) (A - \lambda_0 \mathbf{I}_X)^{-1} - (A - \lambda \mathbf{I}_X)^{-1} (A - \lambda \mathbf{I}_X) (A - \lambda_0 \mathbf{I}_X)^{-1}. \end{aligned}$$

On déduit de la formule (7) que

$$\frac{d}{d\lambda} (A - \lambda \mathbf{I}_X)^{-1} = (A - \lambda \mathbf{I}_X)^{-2}.$$

□

Comme R_A est à valeurs dans l'algèbre de Banach $\mathcal{L}(X)$, il faut prendre quelques précautions si l'on veut appliquer la théorie classique des fonctions de variable complexe. L'astuce consiste à considérer les applications

$$\begin{aligned} f : \rho(A) &\rightarrow \mathbb{C} \\ \lambda &\mapsto \varphi((A - \lambda \mathbf{I}_X)^{-1}), \end{aligned}$$

avec $\varphi \in \mathcal{L}(X)'$. Le théorème de Hahn-Banach montre qu'il existe au moins un $\varphi \in \mathcal{L}(X)'$ pour lequel f n'est pas identiquement nul. De plus, f est holomorphe (par composition), et elle admet le développement

$$f(\lambda) = -\lambda^{-1} \sum_{m=0}^{+\infty} \lambda^{-m} \varphi(A^m)$$

pour $|\lambda| \geq \|A\|$. En particulier, on a $\lim_{|\lambda| \rightarrow +\infty} |f(\lambda)| = 0$. Le théorème de Liouville montre donc que $\rho(A)$ ne peut pas être égal à \mathbb{C} tout entier. Autrement dit, $\sigma(A)$ est non vide. Soit donc $s_A = \sup\{|\lambda|; \lambda \in \sigma(A)\}$ comme annoncé. L'extérieur du disque centré en 0 et rayon s_A est donc inclus dans $\rho(A)$, et f est une fonction holomorphe de $1/\lambda$ pour $|\lambda| > s_A$. Par suite, le développement en série entière

$$f(1/z) = -z \sum_{m=0}^{+\infty} z^m \varphi(A^m)$$

est convergent dans le disque de rayon $1/s_A$. En particulier, on doit donc avoir

$$\sup_m |\lambda^{-m} \varphi(A^m)| < \infty \quad \text{pour } |\lambda| > s_A.$$

Grâce au résultat d'analyse fonctionnelle donné dans le théorème III.4 ci-après (appliqué à l'espace $\mathcal{B} = \mathcal{L}(X)$ et à la famille $B_m = \lambda^{-m} A^m$), on en déduit qu'il existe des nombres réels positifs $C(\lambda)$ tels que

$$\sup_m |\lambda|^{-m} \|A^m\| \leq C(\lambda) \quad \text{pour } |\lambda| > s_A,$$

d'où

$$\limsup_m \|A^m\|^{1/m} \leq |\lambda| \quad \text{pour } |\lambda| > s_A.$$

Ceci achève la démonstration de (5). □

Théorème III.4 Si (B_m) est une famille d'éléments d'un \mathbb{C} -espace de Banach \mathcal{B} tel que pour tout $\varphi \in \mathcal{B}'$,

$$\sup_m |\varphi(B_m)| < \infty$$

alors

$$\sup_m \|B_m\| < \infty.$$

Dém. Cela revient simplement à combiner le théorème de Hahn-Banach avec le théorème de Banach-Steinhaus : le premier montre en effet que

$$\|B_m\|_{\mathcal{B}} = \|b_m\|_{\mathcal{L}(\mathcal{B}'; \mathbb{C})} \quad \text{où} \quad \langle b_m, \varphi \rangle := \varphi(B_m) \quad \forall \varphi \in \mathcal{B}',$$

et le second dit que

$$\sup_m |\langle b_m, \varphi \rangle| < \infty \quad \forall \varphi \in \mathcal{B}' \quad \Rightarrow \quad \sup_m \|b_m\|_{\mathcal{L}(\mathcal{B}'; \mathbb{C})} < \infty.$$

□

Grâce à l'astuce mentionnée plus haut (utiliser des formes linéaires continues φ pour revenir dans \mathbb{C}), on peut faire passer les résultats classiques de la théorie des fonctions de variable complexe aux fonctions à valeurs dans une algèbre de Banach. En particulier, on a le

Théorème III.5 Soit Ω un ouvert connexe non vide de \mathbb{C} et \mathcal{B} une algèbre de Banach. Si $f : \Omega \rightarrow \mathcal{B}$ est holomorphe elle est analytique, c'est-à-dire développable en série entière (et réciproquement). De plus, si Γ et Γ' sont des lacets homotopes dans Ω , alors

$$\int_{\Gamma} f(z) dz = \int_{\Gamma'} f(z) dz.$$

Remarque. Dans le cas de l'algèbre $\mathcal{B} = \mathcal{L}(X)$ et de la fonction R_A , la première partie de l'énoncé est facile à vérifier. On peut même préciser l'observation faite au lemme III.2 en

Proposition III.7 L'application R_A est analytique. De plus, R_A n'admet pas de prolongement analytique en dehors de $\rho(A)$.

Dém. Pour $\lambda, \lambda_0 \in \rho(A)$, la formule (7) s'écrit aussi

$$R_A(\lambda) = (I_X - (\lambda - \lambda_0) R_A(\lambda_0)) R_A(\lambda_0).$$

Par suite, on a

$$R_A(\lambda) = \sum_{n=0}^{+\infty} (\lambda - \lambda_0)^n (R_A(\lambda_0))^{n+1},$$

cette série étant convergente pour $|\lambda - \lambda_0| < 1/\|R_A(\lambda_0)\|$. Ceci montre que R_A est analytique dans $\rho(A)$. « Inversement » si la série ci-dessus converge, alors sa somme est

$$S_A(\lambda) := \sum_{n=0}^{+\infty} (\lambda - \lambda_0)^n (R_A(\lambda_0))^{n+1} = (I_X - (\lambda - \lambda_0) R_A(\lambda_0))^{-1} R_A(\lambda_0).$$

Or

$$R_A(\lambda_0) (A - \lambda I_X) = I_X - (\lambda - \lambda_0) R_A(\lambda_0).$$

Donc $S_A(\lambda)(A - \lambda I_X) = I_X$. On montre de la même façon que $(A - \lambda I_X)S_A(\lambda) = I_X$. D'où $\lambda \in \rho(A)$ et $R_A(\lambda) = S_A(\lambda)$. Par conséquent, R_A n'admet pas de prolongement analytique en dehors de $\rho(A)$.

Pour s'en convaincre, on fait un raisonnement classique mais un peu acrobatique. Supposons qu'une fonction analytique f prolonge R_A au voisinage d'un point $\underline{\lambda}$ du bord de $\rho(A)$. Par définition, $\underline{\lambda}$ n'appartient pas à l'ouvert $\rho(A)$. Soit \underline{r} le rayon de convergence du développement en série entière de f au point $\underline{\lambda}$. Il existe un point $\lambda_0 \in \rho(A) \cap \{\lambda; |\lambda - \underline{\lambda}| < \underline{r}/2\}$, et le développement en série entière de f au point λ_0 a un rayon au moins égal à $\underline{r} - |\lambda_0 - \underline{\lambda}| > |\lambda_0 - \underline{\lambda}|$. En particulier, $\underline{\lambda}$ est dans le disque de convergence. Or, f coïncide avec R_A au voisinage de λ_0 et son développement en série entière est donc

$$f(\lambda) = \sum_{n=0}^{+\infty} (\lambda - \lambda_0)^n (R_A(\lambda_0))^{n+1}.$$

D'après ce qui précède, tous les points du disque de convergence devraient être dans $\rho(A)$, mais ce n'est pas le cas de $\underline{\lambda}$. \square

Nous sommes maintenant armés pour aborder l'extension de la proposition III.6 en dimension infinie. Comme en dimension finie, on dit d'un opérateur qu'il est hyperbolique s'il n'a pas de spectre imaginaire pur.

Théorème III.6 *Si A est un opérateur hyperbolique, il existe des projecteurs continus Π^s et Π^u comme dans la proposition III.6.*

Dém. On procède en deux étapes : **1)** on construit les projecteurs Π^s et Π^u de sorte que $A\Pi^s$ ait son spectre de partie réelle négative et $A\Pi^u$ ait son spectre de partie réelle positive ; **2)** on démontre l'estimation de e^{tA} lorsque A a son spectre de partie réelle négative ; l'estimation « symétrique » s'en déduit par le changement de variables $t \mapsto -t$.

1) Le spectre de A étant un compact disjoint de l'axe imaginaire pur, il existe des courbes fermées simples Γ_{\pm} incluses dans les demi-plans de droite/gauche et qui entourent les ensembles

$$\Sigma_{\pm} := \sigma(A) \cap \{\lambda; \operatorname{Re} \lambda \gtrless 0\}$$

(par exemple, on peut prendre le bord de demi-disques de rayon $\|A\| + \varepsilon$). Soient alors

$$\Pi^s = \frac{1}{2i\pi} \int_{\Gamma_-} (zI_X - A)^{-1} dz, \quad \Pi^u = \frac{1}{2i\pi} \int_{\Gamma_+} (zI_X - A)^{-1} dz.$$

Prenons par exemple le cas de Π^s , et pour alléger les notations on omet l'exposant s ainsi que l'indice $-$ à Γ .

On commence par vérifier que Π est un projecteur. Pour cela, on remarque que d'après le théorème III.5 Π est invariant par un petit déplacement de Γ (puisque l'ensemble résolvant est ouvert). Soit Γ' un chemin fermé simple entourant Γ et ne rencontrant pas l'axe imaginaire. Alors

$$\begin{aligned} \Pi \circ \Pi &= \frac{1}{(2i\pi)^2} \int_{\Gamma} \int_{\Gamma'} (zI_X - A)^{-1} (z'I_X - A)^{-1} dz dz' \\ &= \frac{1}{(2i\pi)^2} \int_{\Gamma} \int_{\Gamma'} (z' - z)^{-1} (zI_X - A)^{-1} - (z' - z)^{-1} (z'I_X - A)^{-1} dz dz' \end{aligned}$$

d'après la formule (7). Or, puisque Γ' est extérieur à Γ , l'indice des points de Γ' par rapport à Γ est nul, tandis que l'indice des points de Γ par rapport à Γ' vaut 1. Donc il reste seulement le premier terme, qui est égal à Π .

Pour vérifier que Π commute avec A , on utilise le fait que les résolvantes commutent entre elles (d'après la formule (7)). Ceci implique, pour tout $\lambda \in \mathbb{C} \setminus \sigma(A)$ (ensemble non vide !),

$$(A - \lambda I_X)^{-1} \Pi = \Pi (A - \lambda I_X)^{-1},$$

d'où

$$\Pi (A - \lambda I_X) = (A - \lambda I_X) \Pi \quad \text{et donc} \quad \Pi A = A \Pi.$$

Considérons alors A^- la restriction de A au sous-espace $X^- := \text{Im}(\Pi)$ (stable par A) et A^+ la restriction de A au sous-espace $X^+ := \text{Im}(I_X - \Pi)$ (également stable par A). On a $A^\pm \in \mathcal{L}(X^\pm)$ et les ensembles résolvants de ces endomorphismes contiennent celui de A , car pour tout $\lambda \notin \sigma(A)$, on a simplement

$$(A^\pm - \lambda I_{X^\pm})^{-1} = (A - \lambda I_X)^{-1}|_{X^\pm}.$$

En fait, on va montrer que leurs spectres sont exactement

$$\sigma(A^\pm) = \Sigma_\pm.$$

Pour cela, on commence par observer que pour tout $\lambda \notin \sigma(A)$ qui n'est pas sur Γ ,

$$\begin{aligned} (A - \lambda I_X)^{-1} \Pi &= -\frac{1}{2i\pi} \int_{\Gamma} (A - \lambda I_X)^{-1} (A - z I_X)^{-1} dz \\ &= \frac{1}{2i\pi} (A - \lambda I_X)^{-1} \int_{\Gamma} \frac{dz}{z - \lambda} - \frac{1}{2i\pi} \int_{\Gamma} (A - z I_X)^{-1} \frac{dz}{z - \lambda} \end{aligned}$$

d'après la formule (7).

- Pour λ à l'extérieur de Γ , le premier terme est nul, et il reste donc :

$$(A - \lambda I_X)^{-1} \Pi = \frac{1}{2i\pi} \int_{\Gamma} (A - z I_X)^{-1} \frac{dz}{\lambda - z}.$$

Le membre de droite est analytique en λ , quel que soit λ à l'extérieur de Γ , et fournit donc un prolongement analytique de $(A - \lambda I_X)^{-1} \Pi$ à tout l'extérieur de Γ . Ceci montre que $\sigma(A^-)$ est à l'intérieur de Γ , et donc $\sigma(A^-) \subset \Sigma_-$.

- Pour λ à l'intérieur de Γ , on a

$$(A - \lambda I_X)^{-1} \Pi = (A - \lambda I_X)^{-1} + \frac{1}{2i\pi} \int_{\Gamma} (A - z I_X)^{-1} \frac{dz}{\lambda - z}$$

et donc $(A - \lambda I_X)^{-1} (I_X - \Pi)$ se prolonge analytiquement à tout l'intérieur de Γ . Ceci montre que $\sigma(A^+)$ est à l'extérieur de Γ , et donc $\sigma(A^+) \subset \Sigma_+$.

Enfin, $\sigma(A) \cap \rho(A^-) \cap \rho(A^+) = \emptyset$ car pour tout $\lambda \in \rho(A^-) \cap \rho(A^+)$, on a

$$\begin{aligned} & (R_{A^-}(\lambda) \Pi + R_{A^+}(\lambda) (I_X - \Pi)) (A - \lambda I_X) \\ &= (A - \lambda I_X) (R_{A^-}(\lambda) \Pi + R_{A^+}(\lambda) (I_X - \Pi)) = I_X \end{aligned}$$

et donc $\lambda \in \rho(A)$. Par suite, $\Sigma_\pm \subset \sigma(A^\pm)$.

2) Il s'agit maintenant d'obtenir une estimation de $e^{tA}\Pi = e^{tA\Pi}$. Pour simplifier, on note A à la place de $A\Pi$, ce qui revient à supposer que le spectre de A est tout entier à l'intérieur d'une courbe fermée simple Γ , elle-même incluse dans un demi-plan $\{\lambda; \operatorname{Re} \lambda \leq \gamma\}$ avec $\gamma \in]-\beta, 0[$. On veut montrer qu'il existe $b > 0$ tel que pour tout $t \geq 0$,

$$\|e^{tA}\| \leq b e^{\gamma t}.$$

Là encore, il faut faire preuve d'imagination car la démonstration faite en dimension finie ne s'adapte pas. L'idée repose sur la *transformation de Laplace*, qui suggère qu'au moins *formellement*

$$e^{tA} = \frac{1}{2i\pi} \int_{\operatorname{Re} z = \gamma} e^{zt} (zI_X - A)^{-1} dz.$$

(Noter que par construction la droite verticale $\{z; \operatorname{Re} z = \gamma\}$ est incluse dans l'ensemble résolvant $\rho(A)$.) Toutefois, rien n'assure que cette intégrale soit convergente ! En effet, l'exponentielle est oscillante mais pas décroissante à l'infini, et $(zI_X - A)^{-1} = -R_A(z)$ se comporte *a priori* en $1/z$ à l'infini. En revanche, si l'on modifie convenablement le chemin d'intégration on peut rendre l'intégrale absolument convergente : en l'occurrence (puisque le spectre de A est borné) on va voir qu'on peut tout simplement remplacer la droite $\{z; \operatorname{Re} z = \gamma\}$ par la courbe fermée Γ (ce qui supprime tout problème de convergence) !

Soit en effet

$$S(t) := \frac{1}{2i\pi} \int_{\Gamma} e^{zt} (zI_X - A)^{-1} dz.$$

Comme $\operatorname{Re} z \leq \gamma$ le long de Γ , il existe évidemment $C > 0$ tel que pour tout $t \geq 0$,

$$\|S(t)\| \leq C e^{\gamma t}.$$

D'après l'étape 1) on sait que $S(0) = I_X$. De plus, par dérivation sous le signe \int , on a

$$\begin{aligned} S'(t) &= \frac{1}{2i\pi} \int_{\Gamma} z e^{zt} (zI_X - A)^{-1} dz \\ &= \frac{1}{2i\pi} \int_C e^{zt} (zI_X - A + A) (zI_X - A)^{-1} dz = AS(t) \end{aligned}$$

puisque

$$\int_{\Gamma} e^{zt} dz = 0$$

(la fonction $z \mapsto e^{zt}$ n'ayant évidemment pas de singularité à l'intérieur de Γ). Autrement dit, S est solution du problème de Cauchy :

$$\frac{dS}{dt} = AS, \quad S(0) = I_X.$$

C'est donc que $S(t) = e^{tA}$. La formule ainsi obtenue :

$$e^{tA} = \frac{1}{2i\pi} \int_{\Gamma} e^{zt} (zI_X - A)^{-1} dz$$

est appelée *intégrale de Dunford-Taylor*. □

5 Équations différentielles linéaires à coefficients périodiques

Soit X un espace de Banach (sur \mathbb{R} ou \mathbb{C}) et $A \in \mathcal{C}(\mathbb{R}; X)$ périodique de période T . On considère l'équation différentielle linéaire

$$\frac{du}{dt} = A(t)u,$$

et $R : \mathbb{R} \times \mathbb{R} \rightarrow \text{Isom}(X)$ sa résolvante :

$$\frac{\partial R}{\partial t}(t, s) = A(t) \circ R(t, s) \quad \text{et} \quad R(s, s) = \mathbf{I}_X.$$

Définition III.6 *Quel que soit $s \in \mathbb{R}$, on appelle l'opérateur*

$$C(s) := R(s + T, s)$$

opérateur de monodromie (au point s).

Proposition III.8 *Les opérateurs de monodromie sont tous conjugués. Pour qu'il existe une solution T -périodique non-triviale de l'équation différentielle, il faut et il suffit que 1 soit valeur propre de ces opérateurs.*

Dém. On commence par observer que

$$R(t + T, s + T) = R(t, s)$$

pour tout $(t, s) \in \mathbb{R} \times \mathbb{R}$. (Les deux membres, vus comme fonctions de t , sont solutions du même problème de Cauchy.) Grâce à la formule de la proposition III.1, on a donc, quels que soient t et s :

$$C(t)R(t, s) = R(t + T, s) = R(t + T, s + T)C(s) = R(t, s)C(s).$$

Par ailleurs, l'existence d'une solution périodique $u \neq 0$ équivaut à l'existence de $u_0 \neq 0$ tel que $R(T, 0)u_0 = u_0$. \square

Le résultat suivant montre que, en dimension finie, les équations différentielles linéaires à coefficients périodiques sont en quelque sorte conjuguées à des équations différentielles autonomes.

Théorème III.7 (Floquet-Lyapunov) *Si X est un \mathbb{C} -espace vectoriel de dimension finie et $A \in \mathcal{C}(\mathbb{R}; \mathcal{L}(X))$ est T -périodique, il existe une application T -périodique $Q : \mathbb{R} \rightarrow \text{Isom}(X)$ et $B \in \mathcal{L}(X)$ telle que u est solution de*

$$\frac{du}{dt} = A(t)u$$

si et seulement si $t \mapsto v(t) := Q(t)u(t)$ est solution de

$$\frac{dv}{dt} = Bv.$$

Dém. La dimension finie permet d'appliquer le lemme III.3 i) ci-après : il existe donc $B \in \mathcal{L}(X)$ tel que

$$e^{TB} = C(0) = R(T, 0).$$

Soit alors Q définie par

$$Q(t) = e^{tB} R(0, t).$$

On a bien $Q(t) \in \text{Isom}(X)$, et d'une part

$$Q(t) R(t, s) u(s) = e^{tB} R(0, s) u(s)$$

est bien solution de l'équation associée à B et d'autre part,

$$Q(t+T) = e^{tB} R(T, 0) R(0, t+T) = e^{tB} R(T, t+T) = e^{tB} R(0, t) = Q(t).$$

□

Remarque. Dans le cadre de ce théorème, on a inversement $u(t) = P(t) v(t)$, avec

$$P(t) = Q(t)^{-1} = R(t, 0) e^{-tB}.$$

La version réelle du théorème est le

Théorème III.8 (Floquet-Lyapunov) *Si X est un \mathbb{R} -espace vectoriel de dimension finie et $A \in \mathcal{C}(\mathbb{R}; \mathcal{L}(X))$ est T -périodique, il existe une application $2T$ -périodique $Q : \mathbb{R} \rightarrow \text{Isom}(X)$ et $B \in \mathcal{L}(X)$ telle que u est solution de*

$$\frac{du}{dt} = A(t) u$$

si et seulement si $t \mapsto v(t) := Q(t) u(t)$ est solution de

$$\frac{dv}{dt} = B v.$$

Dém. D'après le lemme III.3 ii), il existe $B \in \mathcal{L}(X)$ tel que

$$e^{2TB} = C(0)^2 = R(T, 0)^2.$$

On définit à nouveau Q par

$$Q(t) = e^{tB} R(0, t).$$

Elle est $2T$ -périodique car

$$\begin{aligned} Q(t+2T) &= e^{tB} R(T, 0)^2 R(0, t+2T) = e^{tB} R(T, 0) R(T, t+2T) \\ &= e^{tB} R(T, t+T) = e^{tB} R(0, t) = Q(t). \end{aligned}$$

□

Lemme III.3 (Inversion de l'exponentielle)

- i). Pour tout $C \in \text{GL}_n(\mathbb{C})$, il existe $B \in \text{M}_n(\mathbb{C})$ tel que $e^B = C$.
- ii). Pour tout $C \in \text{GL}_n(\mathbb{R})$, il existe $B \in \text{M}_n(\mathbb{R})$ tel que $e^B = C^2$.

Dém. i) Par réduction aux sous-espaces caractéristiques, on peut supposer que $C = \lambda \mathbf{I}_n + N$, où N est nilpotent d'ordre $k \leq n - 1$. Comme C est inversible, λ est non nul : il existe donc $\beta \in C$ tel que $\lambda = e^\beta$, et l'on peut écrire

$$C = e^\beta (\mathbf{I}_n + \lambda^{-1} N).$$

Tout revient donc à trouver \tilde{B} tel que

$$e^{\tilde{B}} = \mathbf{I}_n + \tilde{N}, \quad \tilde{N} := \lambda^{-1} N.$$

(Noter que \tilde{N} est nilpotent comme N .) Car alors il suffira de poser $B = \beta \mathbf{I}_n + \tilde{B}$. Désormais on omet les tildes (ce qui revient à supposer $\lambda = 1$). Par analogie avec le développement en série entière de $\ln(1 + t)$ au point $t = 0$, un candidat est

$$B = \sum_{m=1}^{+\infty} \frac{(-1)^{m+1}}{m} N^m = \sum_{m=1}^k \frac{(-1)^{m+1}}{m} N^m.$$

La vérification de l'égalité $e^B = \mathbf{I}_n + N$ n'est toutefois pas immédiate. Pour cela on va utiliser quelques propriétés de l'exponentielle et notamment le fait que l'exponentielle est un difféomorphisme local en $\mathbf{0}_n$: il existe une boule \mathcal{V}_0 centrée en $\mathbf{0}_n$ et une boule \mathcal{V}_1 centrée en \mathbf{I}_n telles que pour tout $C_1 \in \mathcal{V}_1$ il existe un unique $M \in \mathcal{V}_0$ tel que $C_1 = e^M$. (Ceci résout pour le moment notre problème seulement lorsque la matrice N est « petite ») On va de plus montrer que l'exponentielle réalise un homéomorphisme de l'ensemble des matrices nilpotentes d'ordre k , que l'on note \mathcal{N}_k , sur l'ensemble $\mathbf{I}_n + \mathcal{N}_k$. La bijection réciproque sera bien sûr celle qu'on attend, à savoir :

$$\mathcal{L}_k : C \mapsto \sum_{m=1}^k \frac{(-1)^{m-1}}{m} (C - \mathbf{I}_n)^m.$$

Montrons tout d'abord que le difféomorphisme qui est la réciproque locale de l'exponentielle est

$$\mathcal{L} : C \mapsto \sum_{m=1}^{+\infty} \frac{(-1)^{m+1}}{m} (C - \mathbf{I}_n)^m.$$

On a $\mathcal{L}(e^{\mathbf{0}_n}) = \mathcal{L}(\mathbf{I}_n) = \mathbf{0}_n$. De plus, pour tout $B \in \mathcal{V}_0$ et pour tout $t \in [0, 1]$:

$$\frac{d}{dt} \mathcal{L}(e^{tB}) = \sum_{m=1}^k (-1)^{m-1} (e^{tB} - \mathbf{I}_n)^{m-1} e^{tB} B = \sum_{m=0}^{+\infty} (\mathbf{I}_n - e^{tB})^m e^{tB} B = B.$$

Donc $\mathcal{L}(e^B) = B$. On en déduit que pour tout $M \in \mathcal{V}_1$, $e^{\mathcal{L}(M)} = M$. Cette formule se traduit en particulier, pour $M \in \mathcal{V}_1 \cap (\mathbf{I}_n + \mathcal{N}_k)$, par

$$e^{\mathcal{L}_k(M)} = M.$$

Soit $N \in \mathcal{N}_k$. Regardons l'application

$$t \mapsto e^{\mathcal{L}_k(\mathbf{I}_n + tN)} - \mathbf{I}_n - tN.$$

Elle est *polynômiale* (car l'application $t \mapsto \mathcal{L}_k(\mathbf{I}_n + tN)$ est polynômiale et à valeurs dans \mathcal{N}_k , donc il ne reste qu'un nombre fini de termes lorsqu'on compose par l'exponentielle), et nulle au voisinage de 0. Elle est donc identiquement nulle. En particulier,

$$e^{\mathcal{L}_k(\mathbf{I}_n + N)} = \mathbf{I}_n + N.$$

ii) Voir par exemple [6].

□

Théorème III.9 (« spectral mapping theorem ») Soit $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Si $C \in \mathbf{GL}_n(\mathbb{K})$ s'écrit $C = e^B$, alors les valeurs propres de C coïncident avec les exponentielles des valeurs propres de B , avec les mêmes multiplicités.

Définition III.7 Les valeurs propres des opérateurs de monodromie sont appelées multiplicateurs caractéristiques. Les nombres μ tels que $e^{\mu T}$ est un multiplicateur caractéristique sont appelés exposants de Floquet.

Grâce aux théorèmes III.7 et III.9, la proposition III.5 implique le

Théorème III.10 Les solutions de

$$\frac{du}{dt} = A(t)u$$

sont bornées sur \mathbb{R}^+ si et seulement si les multiplicateurs caractéristiques sont tous de module inférieur ou égal à 1 et ceux de modules 1 sont semi-simples. Toutes les solutions tendent vers 0 en $+\infty$ si et seulement si les multiplicateurs caractéristiques sont tous de module strictement inférieur à 1.

Considérons maintenant l'équation avec terme source :

$$\frac{du}{dt} = A(t)u + b(t),$$

avec b périodique de même période que A .

Proposition III.9 Si 1 n'est pas dans le spectre des opérateurs de monodromie de l'équation homogène, l'équation avec terme source admet au moins une solution périodique.

Remarque. Lorsque $b \equiv 0$ la solution périodique en question est identiquement nulle !

Dém. Soit R la résolvante de l'équation homogène. D'après la formule de Duhamel, toute solution de l'équation avec terme source vérifie

$$u(T) = R(T, 0)u(0) + \int_0^T R(T, s)b(s) ds.$$

Si l'opérateur $\mathbf{I}_X - R(T, 0)$ est inversible, il existe un unique $u_0 \in X$ tel que

$$(\mathbf{I}_X - R(T, 0))u_0 = \int_0^T R(T, s)b(s) ds.$$

La solution du problème de Cauchy de donnée initiale u_0 en $t = 0$ vérifie donc $u(T) = u(0)$. Ceci suffit à prouver qu'elle est T -périodique, car sa translatée $t \mapsto u(t + T)$ est solution du même problème de Cauchy (on utilise ici le fait que le terme source ait la même périodicité que A). \square

Théorème III.11 Supposons X de dimension finie. Pour que l'équation

$$\frac{du}{dt} = A(t)u + b(t)$$

avec A et b périodiques de même période, admette une solution périodique, il suffit qu'elle admette une solution bornée sur \mathbb{R}^+ .

Dém. Soit P l'application affine définie par

$$P(v) = R(T, 0) \left(v + \int_0^T R(0, s) b(s) ds \right),$$

où R est la résolvante de l'équation homogène. Toute solution u de l'équation avec terme source vérifie d'après la formule de Duhamel (itérée), pour tout $n \in \mathbb{N}$, $u(nT) = P^n(u_0)$ où $u_0 = u(0)$. Donc s'il existe une solution bornée, c'est qu'il existe $u_0 \in X$ tel que la suite $P^n(u_0)$ soit bornée.

On va raisonner par l'absurde et montrer que la non-existence de solutions périodiques empêche $P^n(u_0)$ d'être bornée. S'il n'y a pas de solution périodique c'est que l'opérateur P n'a pas de point fixe. Autrement dit, le vecteur $y := R(T, 0) \int_0^T R(0, s) b(s) ds$ n'appartient pas à $E := \text{Im}(\mathbf{I}_X - R(T, 0))$, qui est fermé puisqu'on a supposé l'espace de dimension finie. Donc il existe $\varphi \in E^\perp$ tel que $\varphi(y) \neq 0$ (sinon on aurait $y \in (E^\perp)^\perp = \overline{E} = E$). Par construction, ceci implique

$$\varphi(v - R(T, 0)v) = 0 \quad \text{pour tout } v \in X,$$

c'est-à-dire $\varphi(v) = \varphi(R(T, 0)v)$ et donc $\varphi(v) = \varphi(R(T, 0)^n v)$ pour tout entier n . Or une récurrence facile montre que

$$P^n(v) = R(T, 0)^n v + \sum_{k=0}^{n-1} R(T, 0)^k y.$$

On en déduit

$$\varphi(P^n(v)) = \varphi(v) + n\varphi(y),$$

ce qui n'est pas borné puisque $\varphi(y) \neq 0$. □

Remarque. En dimension infinie, la même démonstration fonctionne, à condition de *supposer*

$$\overline{\text{Im}(\mathbf{I}_X - R(T, 0))} = \text{Im}(\mathbf{I}_X - R(T, 0)).$$

Chapitre IV

Équations autonomes

On revient maintenant à des équations non-linéaires, supposées autonomes c'est-à-dire de la forme

$$(1) \quad \frac{du}{dt} = f(u),$$

où f est une fonction de classe C^1 sur un ouvert U d'un espace de Banach X dans X . On dit qu'une telle fonction f est un *champ de vecteurs*. L'étude de l'équation différentielle (1) revient à l'étude des champs de vecteurs.

1 Champs de vecteurs

Définition IV.1 On appelle *intégrale première* d'un champ de vecteur f sur U une fonction $E \in C^1(U, \mathbb{R})$ telle que

$$dE(u) \cdot f(u) = 0 \quad \text{quel que soit } u \in U.$$

On dit aussi que E est une *intégrale première* de l'équation différentielle (1).

Par définition, quel que soit $u \in C^1(J; U)$ solution de (1), si E est une intégrale première de (1) alors $E \circ u$ est indépendant de t .

Exemple. Le champ de vecteurs associé à la loi de Newton en mécanique est de la forme $f : (x, v) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto (v, F(x))$. Lorsque F « dérive d'un potentiel V » c'est-à-dire $F = -\text{grad}V$, alors l'énergie totale

$$E(x, v) = \frac{1}{2} \|v\|^2 + V(x)$$

est une intégrale première.

La connaissance d'intégrales premières est cruciale pour espérer résoudre explicitement l'équation différentielle ; et elle est toujours utile pour montrer des propriétés qualitatives de cette équation.

Définition IV.2 Un champ de vecteurs f est dit *complet* si et seulement toutes les solutions maximales de (1) sont globales, c'est-à-dire définies sur \mathbb{R} tout entier.

Attention ! La régularité d'un champ de vecteurs n'a rien à voir avec le fait qu'il soit complet ou non, comme le montre l'exemple très simple de l'équation de Riccati, où $X = \mathbb{R}$, $f(u) = u^2$ (donc f est analytique !) et les solutions non triviales « explosent » en temps fini.

Pour démontrer qu'un champ de vecteurs est complet, on peut faire appel au

Théorème IV.1 *Supposons que pour tout intervalle ouvert borné J , toute solution $u \in \mathcal{C}^1(J; U)$ de*

$$\frac{du}{dt} = f(u)$$

soit à valeurs dans un compact K_J de U . Alors le champ f est complet.

Dém. C'est une conséquence du théorème des bouts (théorème II.2), un peu subtile car le compact K_J peut *a priori* dépendre dangereusement du diamètre de J . En fait, on va appliquer le théorème des bouts à un système (équivalent) *augmenté*.

Soit $u \in \mathcal{C}(J; U)$ une solution maximale de

$$\frac{du}{dt} = f(u).$$

Alors l'application $s \in J \mapsto (u, t) = (u(s), s)$ est une solution maximale du système augmenté :

$$\begin{cases} \frac{du}{ds} = f(u), \\ \frac{dt}{ds} = 1. \end{cases}$$

Supposons que $\beta = \sup J$ et $\alpha = \inf J$ soient finis. Choisissons $T > \max(\beta, -\alpha)$. D'après le théorème II.2, la solution maximale (u, t) du système augmenté doit sortir du compact $K_J \times [-T, T]$. Comme u reste dans K_J , c'est que t sort de l'intervalle $[-T, T]$ et donc de J ; ce qui est absurde puisque $t = s \in J$.

Donc l'une au moins des bornes de J est infinie. Et en fait les deux le sont. On s'en convainc aisément en reprenant la démonstration du théorème des bouts : si on avait par exemple $J =]-\infty, \beta[$ avec $\beta < +\infty$, alors, pour tout $\gamma < \beta$ et $T > \max(\beta, -\gamma)$, la solution $(u, t) \in \mathcal{C}^1(] \gamma, \beta[)$ du système augmenté devrait sortir du compact $K_{] \gamma, \beta[} \times [-T, T]$ lorsque $s \nearrow \beta$, ce qui est impossible. \square

En dimension finie, vérifier les hypothèses de ce théorème revient simplement à trouver des *estimations a priori*, c'est-à-dire une majoration des solutions sur les intervalles bornés.

Exemple. Le champ de vecteurs de la mécanique $f : (x, v) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto (v, -\text{grad}V(x))$ est complet pourvu que V soit une fonction positive.

En effet, rappelons que

$$E(x, v) = \frac{1}{2} \|v\|^2 + V(x)$$

est une intégrale première. Si V est positive, si $t \mapsto (x(t), v(t))$ est une solution de l'équation différentielle associée à f , alors $v(t) = x'(t)$, d'où $x(t) = x(0) + \int_0^t v(s) ds$, et

$$\frac{1}{2} \|v(s)\|^2 \leq E(x(s), v(s)) = E(x(0), v(0)),$$

et donc

$$\|x(t)\| \leq \|x(0)\| + T \sqrt{2E(x(0), v(0))}$$

pour $t \in]-T, T[$. Ainsi la restriction de (x, v) à tout intervalle borné est à valeurs dans un ensemble fermé borné de \mathbb{R}^{2n} . Donc on peut appliquer le théorème IV.1.

Définition IV.3 *Les trajectoires des solutions maximales de (1), c'est-à-dire les courbes dans X de la forme $\{u(t); t \in J\}$ où $u \in \mathcal{C}^1(J; U)$ est une solution maximale de (1), sont appelées courbes intégrales du champ de vecteurs f .*

Remarque. D'après l'unicité dans le théorème de Cauchy-Lipschitz, les courbes intégrales d'un champ de vecteurs sont deux à deux disjointes !

2 Flot

Les flots locaux ϕ^{t_0} et ϕ^{t_1} d'une équation différentielle autonome (1) se déduisent l'un de l'autre par translation. En effet, les applications $t \mapsto \phi^{t_0}(t, v)$ et $t \mapsto \phi^{t_1}(t + t_1 - t_0, v)$ sont par définition deux solutions de (1) et elles prennent la même valeur v en $t = t_0$, donc elles sont égales.

C'est pourquoi on fixe $t_0 = 0$ une fois pour toutes. On notera désormais $\phi_t(v) = \phi^0(t, v)$.

2.1 Orbites

Si le champ de vecteurs f est complet, le flot est global, c'est-à-dire que pour tout $v \in U$, l'application $t \mapsto \phi_t(v)$ est définie sur \mathbb{R} tout entier. De plus, l'ensemble

$$\{ \phi_t \in \mathcal{C}^1(U; U) ; t \in \mathbb{R} \}$$

muni de la loi de composition \circ est un *groupe commutatif*. En effet, il contient $\mathbf{I}_U = \phi_0$ et en utilisant à nouveau l'unicité du théorème de Cauchy-Lipschitz on montre que

$$\phi_s \circ \phi_t = \phi_t \circ \phi_s = \phi_{t+s}$$

quels que soient s et t . Cette structure algébrique justifie la définition suivante.

Définition IV.4 *Pour tout $v \in U$ on appelle orbite de v (pour l'équation différentielle (1)) la courbe intégrale de f passant par v , c'est-à-dire l'ensemble*

$$\{ \phi_t(v) ; t \in J \}$$

où est l'intervalle maximal d'existence de ϕ au point v .

Bien sûr, les orbites sont deux à deux disjointes.

Remarque. Si l'on dispose d'une intégrale première E , toute orbite est évidemment incluse dans un ensemble de niveau de E .

Proposition IV.1 *L'orbite d'un point v est réduite au singleton $\{v\}$ si seulement si $f(v) = 0$. Si ce n'est pas le cas et s'il existe $t_0 \neq s_0$ tel que $\phi_{t_0}(v) = \phi_{s_0}(v)$ alors $t \mapsto \phi_t(v)$ est une solution globale périodique et l'orbite de v est une courbe fermée simple.*

Dém. Supposons que $\phi_t(v)$ soit défini pour $t \in [0, T[$ avec $0 \leq s_0 < t_0 < T < +\infty$. Soit m le plus grand entier tel que

$$m(t_0 - s_0) < T.$$

L'application

$$t \mapsto \phi_{t+t_0-s_0}(v)$$

est solution de l'équation différentielle et vaut

$$\phi_{t_0}(v) = \phi_{s_0}(v)$$

en $t = s_0$. En utilisant une fois de plus l'unicité des solutions, on en déduit

$$\phi_{t+t_0-s_0}(v) = \phi_t(v) \quad \text{pour tout } t \in [0, T - (t_0 - s_0)[.$$

En particulier,

$$\phi_{m(t_0-s_0)}(v) = \phi_0(v) = v.$$

Par conséquent, l'application u définie par

$$u(t) = \phi_{t-m(t_0-s_0)}(v) \quad \text{pour } t \in [m(t_0 - s_0), (m+1)(t_0 - s_0)]$$

est solution du problème de Cauchy :

$$\begin{cases} u'(t) = f(u(t)), \\ u(m(t_0 - s_0)) = v \end{cases}$$

et prolonge $t \mapsto \phi_t(v)$ au delà de T . C'est donc que ϕ_t peut être prolongée à tout \mathbb{R} . Et elle est périodique de période $(t_0 - s_0)$.

L'ensemble $\{t > 0; \phi_t(v) = v\}$ contient en particulier $t_0 - s_0$. On peut donc définir

$$T := \inf\{t > 0; \phi_t(v) = v\}.$$

Si v n'est pas un point fixe, T est strictement positif. En effet, il existe $\varphi \in X'$ tel que $\varphi(f(v)) \neq 0$. Supposons par exemple $\varphi(f(v)) > 0$. Par continuité de $t \mapsto \varphi(f(\phi_t(v)))$ on a donc, dans un intervalle $[0, \eta]$ avec $\eta > 0$ assez petit, $\varphi(f(\phi_t(v))) \geq \varepsilon > 0$. Par suite,

$$\varphi(\phi_t(v) - v) = \int_0^t \varphi(f(\phi_s(v))) ds \geq t\varepsilon$$

pour $t \in [0, \eta]$. Donc $\phi_t(v) - v$ ne peut s'annuler dans $[0, \eta]$ en dehors de $t = 0$. Ceci montre que T est strictement positif. De plus T est atteint par continuité de $t \mapsto \phi_t(v)$.

Avec cette définition de T , l'application $t \mapsto \phi_t(v)$ est injective sur $[0, T[$, et l'orbite de v est la courbe fermée $\{\phi_t(v); t \in [0, T[$, qui est sans point double car $\phi_t(v) = \phi_s(v)$ implique $\phi_{t-s}(v) = v$ et donc $t - s \geq T$ si $t > s$. \square

Définition IV.5 On appelle point stationnaire ou point fixe ou point d'équilibre un point $v \in U$ dont l'orbite est réduite à $\{v\}$, ce qui équivaut à ce que le champ de vecteurs s'annule au point v : on dit aussi que v est un point singulier du champ de vecteurs.

Définition IV.6 On appelle cycle une orbite fermée.

Définition IV.7 On appelle orbite hétérocline une courbe intégrale globale

$$\{\phi_t(v); t \in \mathbb{R}\}$$

reliant deux points fixes différents en $+\infty$ et $-\infty$. On appelle orbite homocline une orbite

$$\{\phi_t(v); t \in \mathbb{R}\}$$

reliant un même point fixe en $+\infty$ et $-\infty$.

2.2 Redressement du flot

Définition IV.8 On appelle section locale du champ f en un point non singulier y (c'est-à-dire tel que $f(y) \neq 0$), un ouvert S d'une hypersurface contenant y tel que pour tout $u \in S$, $f(u) \notin T_u S$ (l'espace tangent à S au point u) : on dit aussi que $f(u)$ est transverse à $T_u S$.

En particulier, il existe toujours des sections planes, c'est-à-dire incluses dans un hyperplan affine. En effet, si $f(y) \neq 0$, il existe un hyperplan (vectoriel) fermé H tel que

$$\mathbb{R} f(y) \oplus H = X.$$

(En dimension infinie, ceci repose sur le théorème de Hahn-Banach.) Par continuité de f , l'ensemble $\{u \in U; f(u) \notin H\}$ est ouvert. Donc il existe $r > 0$ tel que pour tout $u \in B(y, r)$ (la boule ouverte de centre y et de rayon r), $f(u) \notin H$. Par conséquent, l'intersection de l'hyperplan affine $y + H$ (dont l'espace tangent en tout point est précisément H) avec $B(y, r)$ est une section locale en y .

Théorème IV.2 (de redressement du flot) Si $\mathbb{R} f(y) \oplus H = X$, alors il existe $\tau > 0$ et $r > 0$ tels que

$$S_r := y + H_r \quad \text{avec} \quad H_r := \{h \in H; \|h\|_X < r\}$$

soit une section locale en y et l'application

$$\begin{aligned} \Phi :]-\tau, \tau[\times H_r &\rightarrow \mathcal{B} := \Phi(]-\tau, \tau[\times S_r) \\ (t, h) &\mapsto \phi_t(y + h), \end{aligned}$$

soit un difféomorphisme. En outre, pour tout $v \in \mathcal{B}$, il existe un unique $t \in]-\tau, \tau[$ tel que $\phi_t(v) \in S_r$.

L'ouvert \mathcal{B} est appelé une boîte à flot.

Dém. On a déjà vu la première partie. Il s'agit ensuite d'appliquer le théorème d'inversion locale. La fonction Φ est de classe \mathcal{C}^1 pour τ et r assez petits. De plus, la différentielle de Φ au point $(0, 0)$ est donnée par

$$d\Phi(0, 0) \cdot (s, k) = s f(y) + k \quad \text{pour tout } (s, k) \in \mathbb{R} \times H.$$

C'est une bijection de $\mathbb{R} \times H$ sur X puisque $\mathbb{R}f(y) \oplus H = X$, et elle est évidemment continue. Donc sa réciproque est aussi continue d'après le théorème de l'application ouverte : on peut aussi le voir de façon plus « élémentaire » (bien qu'il y ait derrière le théorème de Hahn-Banach), en écrivant $H = \text{Ker } \varphi$ avec $\varphi \in X'$ tel que $\varphi(f(y)) = \|f(y)\|^2$, de sorte que

$$s f(y) + k = x \quad \Leftrightarrow \quad s = \frac{\varphi(x)}{\varphi(f(y))}, \quad k = x - \frac{\varphi(x)}{\varphi(f(y))} f(y).$$

Par conséquent, Φ est un difféomorphisme local, et donc un difféomorphisme de $] -\tau, \tau[\times H_r$ sur son image, quitte à réduire τ et r .

Ceci implique que pour tout $v \in \mathcal{B}$, il existe un unique $(t, h) \in] -\tau, \tau[\times H_r$ tel que

$$\phi_t(y + h) = v,$$

d'où $\phi_{-t}(v) = y + h \in S_r$. Inversement, si $s \in] -\tau, \tau[$ est tel que $\phi_s(v) = y + h' \in S_r$, alors $v = \phi_{-s}(y + h')$ et donc $s = -t$ et $h = h'$. \square

Remarque IV.1 *Le difféomorphisme inverse Φ^{-1} opère un redressement du flot. En effet, l'image réciproque d'une orbite $\{\phi_t(v); t \in] -\tau, \tau[\}$ avec $v = y + h \in S_r$ n'est autre que le segment de droite $\{(t, h); t \in] -\tau, \tau[\}$.*

3 Portraits de phase

Définition IV.9 *On appelle portrait de phases d'un champ de vecteurs f , ou de l'équation différentielle associée (1), la partition de U en orbites.*

Obtenir des informations sur le portrait de phase ne nécessite pas de savoir résoudre explicitement l'équation différentielle !

Le point de départ consiste à repérer les orbites *stationnaires*, ce qui revient à résoudre l'équation algébrique $f(v) = 0$. On verra plus loin (Chapitre V) comment trouver l'allure des orbites au voisinage des points stationnaires (où précisément le théorème de redressement du flot ne s'applique pas !)

Il y a d'autres orbites remarquables à rechercher. La recherche de cycles et d'orbites homocline ou hétéroclines est en général non triviale. On verra dans la suite du cours quelques outils pour cela.

Remarque. Si l'on dispose d'une intégrale première E , les orbites sont évidemment incluses dans des ensembles de niveau de E . En dimension finie n , si l'on dispose de $n - 1$ intégrales premières indépendantes, les intersections de leurs ensembles de niveaux forment des courbes, qui fournissent le portrait de phase complet. (C'est une situation idéale et rare !)

Quoi qu'il en soit, en dimension finie on obtient des informations intéressantes sur l'allure des orbites par une simple étude du signe des composantes du champ de vecteurs.

Les portraits de phase en dimension 1 sont essentiellement triviaux : ils reposent seulement sur le tableau de variations de la fonction f . En dimension supérieure à 3 ils sont difficiles à représenter et d'une grande complexité en général. Le cas de la dimension 2 est à la fois riche et pas trop compliqué... (On en verra une raison profonde avec le théorème de Poincaré-Bendixson.)

Il est important de remarquer que le portrait de phases dépend en fait seulement de la direction du champ de vecteurs.

Proposition IV.2 *Le portrait de phases d'un champ de vecteurs f coïncide avec le portrait de phases de tout champ de vecteurs de la forme λf , où $\lambda \in \mathcal{C}^1(U; \mathbb{R}^{+*})$.*

Dém. C'est une simple question de changement de paramétrage. Soit $v \in U$ et $u \in \mathcal{C}^1(]s_-, s_+[)$ la solution maximale de

$$u'(s) = \lambda(u(s)) f(u(s)), \quad u(0) = v.$$

Soit alors

$$T(s) = \int_0^s \lambda(u(\sigma)) d\sigma,$$

qui est une fonction strictement croissante puisque λ est par hypothèse à valeurs positives, et soit S la fonction réciproque de T , définie sur

$$]t_-, t_+[, \quad \text{où } t_{\pm} := \int_0^{s_{\pm}} \lambda(u(\sigma)) d\sigma.$$

On a

$$\frac{d}{dt} u(S(t)) = u'(S(t)) \frac{dS}{dt} = \lambda(u(S(t))) f(u(S(t))) \frac{1}{\lambda(u(S(t)))} = f(u(S(t)))$$

et $u(S(0)) = u(0) = v$. Donc $U = u \circ S$ est solution du problème de Cauchy :

$$\begin{cases} U'(t) = f(U(t)), \\ U(0) = v. \end{cases}$$

De plus, U est maximale. Supposons en effet que l'on puisse prolonger U à $]t_-, t_+ + \varepsilon]$, $\varepsilon > 0$. Prolongeons alors S à $]t_-, t_+ + \varepsilon]$ par

$$S(t) = \int_0^t \frac{d\tau}{\lambda(U(\tau))}.$$

On peut ainsi prolonger T à $]s_-, s_+ + \eta]$ avec

$$\eta = \int_{t_+}^{t_+ + \varepsilon} \frac{d\tau}{\lambda(U(\tau))} > 0.$$

Et on trouve que $U \circ T$ est solution du même problème de Cauchy que u sur l'intervalle $]s_-, s_+ + \eta]$, contenant strictement $]s_-, s_+[$. Cela contredit le fait que u soit maximale. C'est donc que U est aussi une solution maximale (pour le champ f). Autrement dit, l'orbite de v sous le champ λf , c'est-à-dire la courbe

$$\{u(s); s \in]s_-, s_+[\} = \{u \circ S(t); t \in]t_-, t_+[\}$$

est exactement l'orbite de v sous le champ f . □

Remarque IV.2 *Quel que soit le champ de vecteurs $f \in \mathcal{C}^1(U)$, il existe $\lambda \in \mathcal{C}^1(U; \mathbb{R}^{+*})$ tel que $g := \lambda f$ soit complet. En effet, il suffit de poser $\lambda(u) := 1/(1 + \|f(u)\|^2)$ car alors $g = \lambda f$ est borné et le théorème IV.1 montre que ce champ g est complet.*

4 Ensembles ω -limite

Dans ce qui suit f est un champ de vecteurs de classe \mathcal{C}^1 sur l'espace de Banach X tout entier pour simplifier.

4.1 Propriétés générales

Définition IV.10 Soit $v \in X$. Si l'intervalle J de définition de la courbe intégrale de f passant par v est tel que $\sup J = +\infty$, on appelle ensemble ω -limite de v l'ensemble $L_\omega(v)$ des valeurs d'adhérence de $\{\phi_t(v)\}$ lorsque $t \rightarrow +\infty$. Si $\inf J = -\infty$, on appelle ensemble α -limite de v l'ensemble $L_\alpha(v)$ des valeurs d'adhérence de $\{\phi_t(v)\}$ lorsque $t \rightarrow -\infty$.

(Ces noms ont été choisis parce que α et ω sont respectivement la première et la dernière lettre de l'alphabet grec.)

Exemples.

- Si v appartient à une orbite hétérocline, alors par définition il existe v_- et $v_+ \in U$ tels que $\lim_{t \rightarrow \pm\infty} \phi_t(v) = v_\pm$. Par conséquent, $L_\alpha(v) = \{v_-\}$ et $L_\omega(v) = \{v_+\}$.
- Si v appartient à un cycle γ , alors $L_\alpha(v) = L_\omega(v) = \gamma$.

Proposition IV.3 Les ensembles ω -limite et α -limite d'un point v sont fermés et invariants par le flot. Ils sont de plus inclus dans l'ensemble de niveau $\{w; E(w) = E(v)\}$ si E est une intégrale première. Plus généralement, si $F : X \rightarrow \mathbb{R}$ est une fonction continue et monotone le long des courbes intégrales, les ensembles ω -limite et α -limite sont inclus dans des ensembles de niveau de F .

Dém. Ces propriétés sont presque contenues dans la définition. Si $w \in L_\omega(v)$, il existe une suite t_n tendant vers $+\infty$ telle que $\phi_{t_n}(v)$ tend vers w . Alors, quel que soit t ,

$$\phi_t(w) = \lim_{n \rightarrow +\infty} \phi_{t+t_n}(v)$$

appartient aussi à $L_\omega(v)$. Donc cet ensemble est bien invariant par le flot. D'autre part, le complémentaire de l'ensemble ω -limite est par définition l'ensemble des w tels qu'il existe $\varepsilon_0 > 0$ et $t_0 > 0$ avec

$$\|\phi_t(v) - w\| \geq \varepsilon_0$$

pour tout $t \geq t_0$. Or pour un tel w , on a évidemment

$$\|\phi_t(v) - z\| \geq \frac{1}{2}\varepsilon_0$$

pour tout $t \geq t_0$ et pour tout $z \in B(w, \frac{1}{2}\varepsilon_0)$. Donc le complémentaire de l'ensemble ω -limite est ouvert. Si $F : X \rightarrow \mathbb{R}$ est continue et telle que $t \mapsto F(\phi_t(v))$ soit simplement monotone, alors il existe $a \in [-\infty, +\infty]$ tel que

$$\lim_{t \rightarrow +\infty} F(\phi_t(v)) = a.$$

Si $y \in L_\omega(v)$, il existe une suite $(t_n)_{n \in \mathbb{N}}$ tendant vers $+\infty$ telle que

$$\lim_{n \rightarrow +\infty} \phi_{t_n}(v) = y.$$

Par continuité de F , on en déduit

$$a = \lim_{n \rightarrow +\infty} F(\phi_{t_n}(v)) = F(y)$$

(en particulier, $|a| < \infty$). Ceci prouve que $L_\omega(v) \subset \{y; F(y) = a\}$. \square

Proposition IV.4 *Soit $v \in X$. Si l'orbite $\{\phi_t(v)\}$ est définie pour tout $t \in \mathbb{R}^+$ et relativement compacte, alors $L_\omega(v)$ est compact non vide et connexe.*

Dém. La compacité de $L_\omega(v)$ est quasi-immédiate : c'est un fermé inclus dans le compact

$$\overline{\{\phi_t(v); t \in \mathbb{R}^+\}}.$$

Il est non vide puisque l'orbite $\{\phi_t(v); t \in \mathbb{R}^+\}$ admet au moins une valeur d'adhérence. La connexité de $L_\omega(v)$ demande un peu plus de travail. Raisonnons par l'absurde et supposons qu'il existe deux ouverts disjoints U_1 et U_2 tels que

$$L_\omega(v) \subset U_1 \cup U_2, \quad L_\omega(v) \cap U_i \neq \emptyset.$$

Par définition de $L_\omega(v)$, pour tout $n \in \mathbb{N}$, il existe donc une suite (τ_n) strictement croissante telle que $\phi_{\tau_{2n}}(v) \in U_2$, $\phi_{\tau_{2n+1}}(v) \in U_1$. Comme l'ensemble $\{\phi_t(v); \tau_{2n} \leq t \leq \tau_{2n+1}\}$ est connexe (image d'un connexe par une application continue !), il ne peut être inclus dans $U_1 \cup U_2$. Donc il existe $t_n \in]\tau_{2n}, \tau_{2n+1}[$ tel que $\phi_{t_n}(v) \notin U_1 \cup U_2$. Or la suite $\phi_{t_n}(v)$ admet une valeur d'adhérence $a \in L_\omega(v)$, et cette limite appartient au complémentaire de $U_1 \cup U_2$ (car c'est un fermé). Ceci contredit l'inclusion $L_\omega(v) \subset U_1 \cup U_2$. C'est donc que de tels ouverts U_1 et U_2 n'existent pas, ce qui montre que $L_\omega(v)$ est connexe. \square

Remarque IV.3 *Inversement, si $L_\omega(v)$ est compact et non vide, l'orbite*

$$\{\phi_t(v); t \in \mathbb{R}^+\}$$

est relativement compacte, puisqu'elle admet au moins une valeur d'adhérence, et toutes ses valeurs d'adhérence sont incluses dans le compact $L_\omega(v)$.

4.2 Équations dans le plan

Dans \mathbb{R}^2 , les sections locales "planes" sont des segments de droite, en particulier orientables. De plus, on dispose du

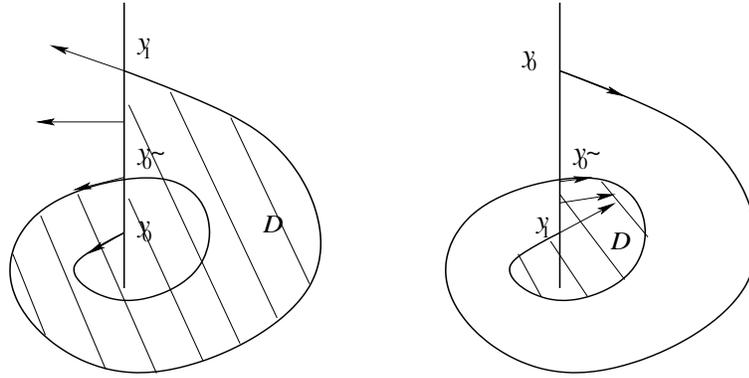
Théorème IV.3 (de Jordan) *Une courbe fermée simple divise le plan en deux régions connexes, l'une bornée et l'autre non bornée.*

Grâce à ces deux ingrédients, on peut démontrer le

Théorème IV.4 (de Poincaré-Bendixson) *Pour un champ de vecteurs dans \mathbb{R}^2 , tout ensemble ω -limite compact non vide et sans point fixe est un cycle.*

Il existe d'autres versions de ce théorème (sans l'hypothèse « sans point fixe » notamment).

La démonstration n'est pas difficile mais un peu longue. Elle utilise les résultats préliminaires que voici.

Figure IV.1: Flot et section ; composante D en hachuré.

Lemme IV.1 Si S est une section locale, si (t_n) est une suite croissante telle que $y_n = \phi_{t_n}(v) \in S$, alors la suite (y_n) est aussi monotone le long de S , c'est-à-dire

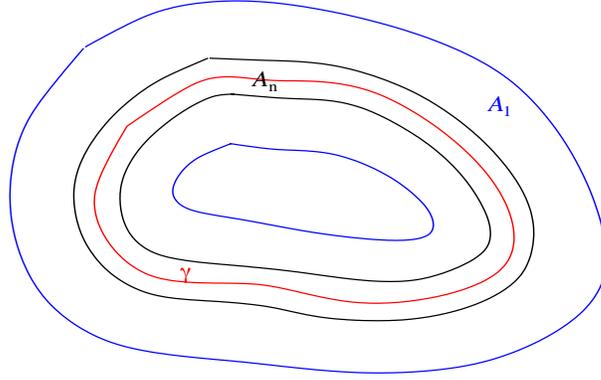
$$\det(y_{n+1} - y_n, y_n - y_{n-1}) \geq 0.$$

Dém. Il suffit de faire la preuve pour trois points y_0, y_1, y_2 . Fixons donc $y_0 = \phi_{t_0}(v)$ et $y_1 = \phi_{t_1}(v)$ appartenant à S avec $t_0 < t_1$ et $y_0 \neq y_1$. On peut supposer que l'orbite de v ne coupe pas S dans l'intervalle $I := [y_0, y_1]$: si tel était le cas, on remplacerait y_0 par $\tilde{y}_0 = \phi_{\tilde{t}_0}(y_1)$ le premier point appartenant à S en partant de y_1 dans le sens des $t < t_1$ (cf Figure IV.1), car montrer que y_0, y_1 et $y_2 = \phi_t(v), t > t_1$ sont ordonnés revient à montrer que \tilde{y}_0, y_1 et y_2 le sont.

Soit C la courbe formée de la réunion de I et de $\{\phi_s(v) ; s \in [t_0, t_1]\}$. C'est une courbe fermée, simple grâce à la précaution que l'on a prise et au fait qu'une orbite ne s'intersecte pas elle-même. Donc d'après le théorème IV.3, elle divise le plan en deux composantes connexes. Soit D celle qui est bornée. Puisque S est une section, le champ f pointe vers le même demi-plan le long de S , et en particulier le long de I . En particulier, il pointe soit vers D soit vers $\mathbb{R}^2 \setminus D$. Quitte à renverser le temps (et échanger y_0 et y_1), on peut supposer qu'on est dans le second cas. Alors D est négativement invariant par le flot, c'est-à-dire que pour tout $y \in D$, $\phi_t(y) \in D$ pour $t < 0$. En effet, le flot ne peut sortir ni par I ni par $\{\phi_t(v) ; t \in [t_0, t_1]\}$. Ceci implique en particulier que $\phi_t(y_1)$ appartient à $\mathbb{R}^2 \setminus D$ pour tout $t > t_1$. De plus, $S \setminus I$ est constitué de deux intervalles I_0 et I_1 contenant respectivement y_0 et y_1 dans leur bord. Or on peut joindre tout point de I_0 assez proche de y_0 à $\phi_{-\varepsilon}(y_0)$ avec $\varepsilon > 0$ assez petit, point appartenant à D , sans passer par le bord de D , on en déduit que I_0 est dans D . Donc, si $y_2 = \phi_{t_2}(v) \in S$ et $t > t_2$, y_2 appartient nécessairement à I_1 . \square

Corollaire IV.1 Si S est une section locale, pour tout $v \in U$, $L_\omega(v) \cap S$ contient au plus un point.

Dém. Supposons que $L_\omega(v) \cap S$ contienne deux points distincts, y_1 et y_2 . Soient alors des boîtes à flot disjointes, $\mathcal{V}_1 \ni y_1$ et $\mathcal{V}_2 \ni y_2$. Comme ces points sont dans $L_\omega(z)$, l'orbite z repasse une infinité de fois dans chacune de ces boîtes, et donc aussi par chacun des intervalles $I_1 = \mathcal{V}_1 \subset S$ et $I_2 = \mathcal{V}_2 \subset S$. Plus précisément, il existe une suite (t_n) , croissante et tendant vers $+\infty$ avec n , telle $\phi_{t_{2n+1}}(v) \in I_1$ et $\phi_{t_{2n}}(v) \in I_2$. Comme I_1 et I_2 sont disjoints, ceci contredit le lemme IV.1. \square

Figure IV.2: Voisinages tubulaires A_n de γ .

Dém. [Théorème de Poincaré-Bendixson] Supposons $L_\omega(v)$ compact et $y \in L_\omega(v)$. On va montrer que l'orbite de y , qui est incluse dans $L_\omega(v)$, est un cycle. On considère pour cela $L_\omega(y)$, qui est un sous-ensemble non vide de $L_\omega(v)$. Soit alors $z \in L_\omega(y)$, S une section locale en z et \mathcal{Z} une boîte à flot associée. Il existe une suite (t_n) tendant vers $+\infty$ telle que $\phi_{t_n}(y)$ appartienne au voisinage \mathcal{Z} de z pour tout n . En particulier, il existe des instants t_n et t_m distants de plus de 2τ , et des instants s_n et $s_m \in]-\tau, \tau[$ (avec τ comme dans le théorème IV.2) tels que $\phi_{t_n+s_n}(y)$ et $\phi_{t_m+s_m}(y)$ appartiennent à S . D'après le corollaire IV.1, ceci implique

$$\phi_{t_n+s_n}(y) = \phi_{t_m+s_m}(y).$$

Comme $t_n + s_n \neq t_m + s_m$, et que y n'est pas un équilibre, l'orbite de y est donc un cycle γ d'après la proposition IV.1), et $\gamma \subset L_\omega(v)$.

Pour prouver que $L_\omega(v)$ est égal à au cycle γ , raisonnons par l'absurde. Supposons γ strictement inclus dans $L_\omega(v)$. Soit $y_0 \in L_\omega(v)$, $y_0 \notin \gamma$, et n_0 tel que

$$1 < 2^{n_0} \text{dist}(y_0, \gamma).$$

Notons, pour tout $n \in \mathbb{N}$,

$$A_n := \bigcup_{z \in \gamma} B(z, 1/2^n).$$

Alors, pour tout $n \geq n_0$, $y_0 \notin \overline{A_n}$. De plus, pour tout $n \geq n_0$, il existe y_n tel que

$$y_n \in L_\omega(v) \cap A_n \quad \text{et} \quad y_n \notin \gamma.$$

En effet, supposons que pour tout $y \in L_\omega(v)$, ou bien $y \in \gamma$ ou bien $y \notin A_n$. Comme le complémentaire de $\overline{A_{n+1}}$ contient le complémentaire de A_n (ça va mieux avec un dessin !, cf Figure IV.2), et $\gamma \subset A_{n+2}$, $y_0 \in L_\omega(v) \cap (U \setminus \overline{A_{n+1}})$, on aurait alors

$$L_\omega(v) \subset A_{n+2} \cup (U \setminus \overline{A_{n+1}}), \quad L_\omega(v) \cap A_{n+2} \neq \emptyset, \quad L_\omega \cap (U \setminus \overline{A_{n+1}}) \neq \emptyset.$$

Les ensembles A_{n+2} et $(U \setminus \overline{A_{n+1}})$ étant ouverts, cela contredirait la connexité de $L_\omega(v)$ (voir la proposition IV.4).

Maintenant, par compacité de $L_\omega(v)$, la suite (y_n) admet une sous-suite convergente vers un point z et par construction de (y_n) , $\text{dist}(z, \gamma) = 0$, donc $z \in \gamma$ puisque γ est compact. Pour simplifier, on note encore (y_n) la suite extraite. Soit S une section locale en z et \mathcal{Z} une boîte à flot

associée. Pour n assez grand, $y_n \in \mathcal{Z}$, donc il existe $s_n \in]-\tau, \tau[$ tel que $x_n := \phi_{s_n}(y_n) \in S$, et $x_n \neq z$ puisque $y_n \notin \gamma$. Comme $y_n \in L_\omega(v)$, on a aussi $x_n \in L_\omega(v)$. Autrement dit, z et x_n sont deux points distincts de $L_\omega(v) \cap S$. Ceci est en contradiction avec le corollaire IV.1.

C'est donc que notre hypothèse était absurde : on a nécessairement

$$\gamma = L_\omega(v).$$

□

Chapitre V

Stabilité des solutions stationnaires

On va s'intéresser ici à deux types de stabilité : l'une concernant le comportement asymptotique des solutions pour des données initiales proches d'une solution stationnaire ; l'autre concernant le portrait de phase autour d'un point fixe lorsqu'on perturbe l'équation (en faisant varier des paramètres par exemple). Dans le premier cas, on parle de stabilité asymptotique (ou de stabilité tout court), que l'on peut étudier grâce à la théorie de Lyapunov. Dans le second cas, on parle de stabilité structurelle, ce qui amène aux notions de variétés stable/ instable/centrale et à des problèmes de bifurcations.

1 Théorie de Lyapunov

Définition V.1 *Un point singulier v d'un champ de vecteurs $f \in \mathcal{C}^1(U)$ est dit stable s'il existe un voisinage \mathcal{V}_0 de v dans U tel que*

- i). le flot $\phi_t(w)$ est défini pour tout $t \geq 0$ et pour tout $w \in \mathcal{V}_0$,*
- ii). pour tout voisinage \mathcal{W} de v dans U , il existe un voisinage $\mathcal{V} \subset \mathcal{V}_0$ tel que $\phi_t(w)$ appartient à \mathcal{W} pour tout $t \geq 0$ et pour tout $w \in \mathcal{V}$.*

Si de plus, on a

- iii). pour tout $w \in \mathcal{V}_0$,*

$$\lim_{t \rightarrow +\infty} \phi_t(w) = v$$

alors le point fixe v est dit asymptotiquement stable.

On dit d'un point non stable qu'il est instable.

Concernant la stabilité des équilibres pour les équations linéaires, la proposition III.5 se reformule ainsi.

Théorème V.1 *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Le point 0 est stable pour l'équation*

$$\frac{du}{dt} = Au$$

si et seulement si

- i). les valeurs propres de A sont toutes de partie réelle négative ou nulle,*
- ii). les valeurs propres imaginaires pures de A sont semi-simples.*

Le point 0 est asymptotiquement stable si et seulement si

iii). les valeurs propres de A sont toutes de partie réelle strictement négative.

Pour les équations non-linéaires, la stabilité des équilibres est plus délicate à étudier (en dimension finie et a fortiori en dimension infinie). La théorie de Lyapunov repose sur l'existence de fonctions (de Lyapunov !) qui permettent de « contrôler », dans une certaine mesure, le comportement asymptotique des solutions pour des données initiales proches de l'équilibre.

Théorème V.2 (Lyapunov n°1) Soit v un point d'équilibre du champ $f \in C^1(U)$, U étant un ouvert dans un espace de Banach X . On suppose qu'il existe un voisinage \mathcal{V} de v dans U et une fonction $F \in C^2(\mathcal{V})$, telle que

- le point v est un minimum local de F , d'où nécessairement $DF(v) = 0$, et de plus il existe $\alpha > 0$ tel que $D^2F(v) \geq \alpha \text{Id}$,
- et la fonction F est décroissante le long des orbites issues de \mathcal{V} , c'est-à-dire $DF(u) \cdot f(u) \leq 0 \quad \forall u \in \mathcal{V}$.

Alors v est un point d'équilibre stable.

Dém. Soit $\eta_0 > 0$ tel que f soit bornée, disons par M , et Lipschitzienne dans la boule $B(v, 2\eta_0) \subset \mathcal{V}$. La démonstration du théorème de Cauchy-Lipschitz montre que le flot $\phi_t(w)$ est défini pour tout $w \in B(v, \eta_0)$ et $t \in [0, \tau]$ avec $\tau := \frac{\eta_0}{2M}$.

D'après l'hypothèse sur F et la formule de Taylor à l'ordre 2, il existe $\eta \leq \eta_0$ tel que

$$F(v+h) - F(v) \geq \frac{\alpha}{4} \|h\|^2$$

pour tout $h \in X$ avec $\|h\| \leq \eta$. Par suite,

$$\inf_{\|u-v\|=\eta} F(u) \geq F(v) + \frac{\alpha}{4} \eta^2 =: m.$$

Par continuité de F ,

$$\mathcal{V}_\eta := \{u \in B(v, \eta); F(u) < m\}$$

est un voisinage (ouvert) de v dans U . De plus, si $w \in \mathcal{V}_\eta$,

$$F(\phi_t(w)) \leq F(w) < m$$

pour $t \in [0, \tau]$ et donc $\|\phi_t(w) - v\|$ reste strictement inférieur à $\eta \leq \eta_0$. En particulier, $\phi_\tau(w)$ appartient à $\mathcal{V}_\eta \subset B(v, \eta_0)$. Donc on peut prolonger $\phi_t(w)$ à $t \in [\tau, 2\tau]$, et de proche en proche à tout intervalle $[k\tau, (k+1)\tau]$ ($k \in \mathbb{N}$), tant en restant dans \mathcal{V}_η . En conclusion, pour tout $w \in \mathcal{V}_\eta$, $\phi_t(w)$ est défini pour tout $t \geq 0$ et à valeurs dans \mathcal{V}_η . \square

Définition V.2 La fonction F du théorème V.2 est appelée une fonction de Lyapunov.

Exemple : une intégrale première ayant un minimum local (au sens strict de l'énoncé du théorème V.2) en v est un fonction de Lyapunov !

Dans les modèles physiques, les fonctions de Lyapunov sont souvent reliées à une *énergie*.

Remarque V.1 En dimension finie, on peut affaiblir l'hypothèse sur F , en demandant que v soit un minimum local strict sans nécessairement avoir d'estimation de $D^2F(v)$.

Si l'on dispose d'une fonction de Lyapunov dotée d'une propriété supplémentaire adéquate, on peut obtenir la stabilité asymptotique de l'équilibre.

Théorème V.3 (Lyapunov n°2) Soit v un point d'équilibre du champ $f \in \mathcal{C}^1(U)$, U étant un ouvert dans un espace de Banach X . On suppose qu'il existe un voisinage \mathcal{V} de v dans U et une fonction de Lyapunov $F \in \mathcal{C}^2(\mathcal{V})$ telle que

$$\bullet \text{ il existe } \beta > 0 \text{ tel que } DF(u) \cdot f(u) \leq -\beta (F(u) - F(v)) \quad \forall u \in \mathcal{V}.$$

Alors v est un point d'équilibre asymptotiquement stable.

Dém. On sait déjà d'après la preuve du théorème V.2 qu'il existe un voisinage \mathcal{W} de v inclus dans \mathcal{V} tel que

$$\phi_t(\mathcal{W}) \subset \mathcal{W}$$

pour tout $t \geq 0$. L'hypothèse renforcée sur F permet facilement de voir que

$$\lim_{t \rightarrow +\infty} \phi_t(w) = v$$

pour tout $w \in \mathcal{W}$. En effet,

$$\frac{d}{dt} (F(\phi_t(w)) - F(v)) = DF(\phi_t(w)) \cdot f(\phi_t(w)) \leq -\beta (F(\phi_t(w)) - F(v))$$

implique par une intégration immédiate :

$$(F(\phi_t(w)) - F(v)) \leq (F(w) - F(v)) e^{-\beta t}.$$

Or

$$F(\phi_t(w)) - F(v) \geq \frac{\alpha}{4} \|\phi_t(w) - v\|^2$$

par construction de \mathcal{W} . □

Définition V.3 La fonction F du théorème V.3 est appelée une fonction de Lyapunov forte.

Remarque V.2

i). Bien sûr, une intégrale première n'est pas une fonction de Lyapunov forte. Cependant, on peut parfois trouver une fonction de Lyapunov forte comme intégrale première d'un système "approché". C'est le cas par exemple pour l'équation du pendule avec amortissement :

$$\frac{d^2x}{dt^2} + k \frac{dx}{dt} + \sin x = 0, \quad k > 0.$$

Soit E l'intégrale première connue pour $k = 0$:

$$E(x, x') = \frac{1}{2} (x')^2 - \cos x.$$

Alors pour une solution du système amorti on a :

$$\frac{d}{dt} E(x(t), x'(t)) = x'(t) (x''(t) + \sin x(t)) = -k (x'(t))^2 < 0$$

pour $x'(t) \neq 0$. La fonction E est donc une fonction de Lyapunov aux points de la forme $(2k\pi, 0)$; ce n'est pas tout à fait une fonction de Lyapunov forte, mais elle permet de montrer que ces points sont asymptotiquement stables (ceci est laissé en exercice).

ii). Des systèmes admettant une fonction de Lyapunov forte évidente sont les systèmes gradient, de la forme

$$\frac{du}{dt} = -\text{grad}V(u).$$

Si v est un minimum local strict de V alors V est une fonction de Lyapunov forte au point v .

Un résultat remarquable de la théorie de Lyapunov, en dimension finie, est l'équivalence de la stabilité asymptotique pour l'équation non-linéaire et pour sa version linéarisée au point d'équilibre. Il peut s'énoncer ainsi.

Théorème V.4 (Lyapunov n°3) Soit v un point d'équilibre du champ $f \in C^1(U)$, $U \subset \mathbb{R}^n$. On suppose que la matrice jacobienne $Df(v)$ a toutes ses valeurs propres de partie réelle strictement négative. Alors v est asymptotiquement stable.

On dit aussi en abrégé que la *stabilité spectrale* (c'est-à-dire que le spectre du linéarisé est inclus dans $\{\lambda \in \mathbb{C}; \text{Re } \lambda < 0\}$) implique la stabilité non-linéaire. C'est un résultat spécifique de la dimension finie¹, qui équivaut d'après le théorème V.1 à l'assertion suivante. Si v est un point d'équilibre de f et si 0 est un point d'équilibre *asymptotiquement stable* du système *linéarisé* autour de ce point d'équilibre :

$$\frac{du}{dt} = Df(v)u$$

alors v est un point d'équilibre asymptotiquement stable du système *non-linéaire*.

Pour démontrer le théorème V.4, on aura besoin du

Lemme V.1 Soit $A \in \mathcal{M}_n(\mathbb{R})$ dont les valeurs propres sont toutes de partie réelle strictement négative. Alors l'équation

$$\frac{du}{dt} = Au$$

admet une fonction de Lyapunov forte qui est une forme quadratique.

Dém. On commence par "diagonaliser A à ε près", c'est-à-dire à trouver, pour tout $\varepsilon > 0$, une matrice $P_\varepsilon \in \text{GL}_n(\mathbb{C})$ et une matrice $C_\varepsilon \in \mathcal{M}_n(\mathbb{C})$ de norme (subordonnée à la norme hermitienne sur \mathbb{C}^n) inférieure à ε telles que

$$P_\varepsilon^{-1} A P_\varepsilon = D + C_\varepsilon$$

avec D diagonale de coefficients $\{\lambda_1, \dots, \lambda_n\}$ les valeurs propres de A . Pour cela, il suffit de trigonaliser A de façon particulière : si a est l'endomorphisme de matrice A dans la base canonique de \mathbb{C}^n , on fabrique une base (e_1, \dots, e_n) dans laquelle la matrice de a soit une matrice triangulaire supérieure $D + B$, c'est-à-dire que pour tout $j \in \{1, \dots, n\}$:

$$a(e_j) = \lambda_j e_j + \sum_{i=1}^{j-1} b_{ij} e_i,$$

¹Démontrer le même résultat en dimension infinie demande en général beaucoup d'efforts supplémentaires.

puis on dilate les vecteurs de base en $\tilde{\mathbf{e}}_j = \eta^j \mathbf{e}_j$, de sorte que

$$a(\tilde{\mathbf{e}}_j) - \lambda_j \tilde{\mathbf{e}}_j = \sum_{i=1}^{j-1} b_{ij} \eta^{j-i} \tilde{\mathbf{e}}_i.$$

Autrement dit, on a $P^{-1} A P = D + C$ avec P la matrice de vecteurs colonnes $\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n$ et C la matrice triangulaire supérieure de coefficients $c_{ij} := b_{ij} \eta^{j-i}$ pour $i \leq j-1$ (et $c_{ij} = 0$ pour $i \geq j$). Cette matrice est de norme inférieure à ε pourvu que $\eta < \min(1, \varepsilon/\|B\|)$.

Soit alors

$$q(u) = u^* Q^* Q u = \|Q u\|^2, \quad Q = P_\varepsilon^{-1}.$$

On a

$$q(0) = 0, \quad Dq(u) \cdot h = u^* Q^* Q h + h^* Q^* Q u, \quad D^2 q(u) = 2 Q^* Q.$$

Comme $Q^* Q$ est hermitienne définie positive, la fonction q admet bien un minimum local strict en 0 (on a la minoration évidente $Q^* Q \geq \alpha \mathbf{I}_n$ avec $\alpha = \min_{\|h\|=1} \|Qh\|^2$). De plus,

$$Dq(u) A u = 2 \operatorname{Re} (u^* Q^* Q A u) = 2 \operatorname{Re} (u^* Q^* D Q u) + 2 \operatorname{Re} (u^* Q^* C Q u).$$

D'après l'inégalité de Cauchy-Schwarz, on a

$$\operatorname{Re} (u^* Q^* C Q u) \leq \|Q u\| \|C Q u\| \leq \varepsilon \|Q u\|^2$$

par construction de C . Donc, si $\beta = -\max_j \operatorname{Re} \lambda_j$, on a

$$Dq(u) A u \leq -2(\beta - \varepsilon) q(u),$$

ce qui implique que q est une fonction de Lyapunov forte pourvu qu'on ait choisi $\varepsilon < \beta$. \square

Dém. [Théorème V.4] On peut appliquer le lemme V.1 à la matrice jacobienne $A = Df(v)$. On a ainsi une forme quadratique $q(u) = \|Q u\|^2$, telle que

$$q(0) = 0, \quad Dq(0) = 0, \quad Dq(0) \geq \alpha \mathbf{I}_n$$

et

$$Dq(u) A u \leq -\beta q(u)$$

avec α et $\beta > 0$. Alors $F(w) = q(w - v)$ définit une fonction de Lyapunov forte pour le champ f . En effet, on a de façon évidente :

$$F(v) = 0, \quad DF(v) = 0, \quad DF(v) \geq \alpha \mathbf{I}_n.$$

De plus, d'après la formule de Taylor avec reste intégral,

$$f(w) = Df(v)(w - v) + B(v, w) \cdot (w - v, w - v)$$

où $B(v, w)$ est l'application bilinéaire définie par :

$$B(v, w) = \frac{1}{2} \int_0^1 (1 - \theta) D^2 f(v + \theta(w - v)) d\theta.$$

En particulier, il existe $c > 0$ tel que pour $\|w - v\| \leq 1$:

$$\|B(v, w) \cdot (w - v, w - v)\| \leq c \|w - v\|^2 \leq \frac{2c}{\alpha} \|Q(w - v)\|^2.$$

Par suite

$$\begin{aligned} DF(w) f(w) &= 2 \operatorname{Re} ((w - v)^* Q^* Q f(w)) \\ &= Dq(w - v) A(w - v) + 2 \operatorname{Re} ((w - v)^* Q^* Q B(v, w) \cdot (w - v, w - v)) \end{aligned}$$

d'où par l'inégalité de Cauchy-Schwarz

$$DF(w) f(w) \leq -\beta F(w) + \frac{4c}{\alpha} \|Q\| \|Q(w - v)\|^3 \leq -\left(\beta - \frac{4c}{\alpha} \|Q\|^2 \|w - v\|\right) F(w).$$

Pour $\|w - v\| \leq \frac{\alpha\beta}{8c\|Q\|^2}$ on a donc

$$DF(w) f(w) \leq -\frac{\beta}{2} F(w).$$

Donc F est une fonction de Lyapunov forte pour f au point v , et l'équilibre v est asymptotiquement stable d'après le le théorème V.3. \square

Pour un "gros" système, il peut parfois être plus facile de trouver une fonction de Lyapunov que de montrer que le système linéarisé a ses valeurs propres sont de partie réelle négative. En dimension infinie, le spectre du linéarisé est encore plus difficile à localiser et il ne suffit pas à prouver la stabilité. La recherche de fonctions de Lyapunov est donc très importante.

2 Points fixes hyperboliques

La classification des points fixes pour les équations différentielles linéaires dans le plan (cf TDs) montre que les centres (correspondant à une matrice $A \in M_2(\mathbb{R})$ avec des valeurs propres imaginaires pures) jouent un rôle particulier, au sens où ils ne sont pas structurellement stables : une modification arbitrairement petite de la matrice

$$\begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} \quad \text{en} \quad \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

avec $|\alpha| \ll 1$ modifie dramatiquement le portrait de phase au voisinage de 0 ! Ces points sont les prototypes de points fixes *non hyperboliques*.

Définition V.4 *Un point singulier v du champ f est dit hyperbolique si le spectre de l'opérateur $Df(v)$ n'intersecte pas l'axe imaginaire.*

Les points fixes hyperboliques sont précisément ceux qui sont structurellement stables, au sens topologique du terme. C'est l'objet du théorème suivant, que nous admettrons.

Théorème V.5 (Hartman-Grobman) *Si v est un point fixe hyperbolique de l'équation différentielle*

$$\frac{du}{dt} = f(u),$$

avec $f \in C^1(U)$, U un ouvert de \mathbb{R}^n , il existe un voisinage \mathcal{V} de v dans U , un voisinage \mathcal{O} de 0 dans \mathbb{R}^n et un homéomorphisme h de \mathcal{V} sur \mathcal{O} tel que le flot $\phi_t(w)$ vérifie :

$$h(\phi_t(w)) = e^{tDf(v)} h(w)$$

pour tout $w \in \mathcal{V}$ et $t \in \mathbb{R}$ tel que $\phi_t(w)$ soit bien défini.

Ce théorème signifie que l'on peut déformer *continûment* les orbites au voisinage de v pour obtenir les orbites du système linéarisé

$$\frac{du}{dt} = Df(v)(u - v),$$

et réciproquement. Il ne dit rien sur les propriétés de nature différentielle (tangentes, convexité, etc.) de ces courbes. En particulier, il ne fait pas la distinction entre les différents nœuds en dimension 2. Le théorème dit de la variété stable ci-après ne fait pas non plus de distinction entre ces points, mais il donne des informations sur les propriétés différentielles des orbites, souvent utiles dans les applications.

Théorème V.6 (de la variété stable locale) *Si v est un point fixe hyperbolique de l'équation différentielle*

$$\frac{du}{dt} = f(u)$$

(avec $f \in \mathcal{C}^1(U)$, U un ouvert de \mathbb{R}^n), dont le flot est noté $\phi_t(w)$, il existe un voisinage \mathcal{V} de v dans U tel que les ensembles

$$W_{\text{loc}}^s = \{w \in \mathcal{V}; t \mapsto \phi_t(w) \in \mathcal{C}_b^1(\mathbb{R}^+)\}, \quad W_{\text{loc}}^u = \{w \in \mathcal{V}; t \mapsto \phi_t(w) \in \mathcal{C}_b^1(\mathbb{R}^-)\}$$

soient des variétés (à bord) contenant v , tangentes respectivement à E^s et E^u , les sous-espaces stable et instable de $Df(v)$, au point v : autrement dit, W_{loc}^s et W_{loc}^u sont des graphes de fonctions différentiables sur des voisinages de v dans les sous-espaces affines $v + E^s$ et $v + E^u$ respectivement, dont la différentielle s'annule en v .

De plus, W_{loc}^s est positivement invariante par le flot, c'est-à-dire $\phi_t(W_{\text{loc}}^s) \subset W_{\text{loc}}^s$ pour tout $t \geq 0$, tandis que $\phi_t(W_{\text{loc}}^u) \subset W_{\text{loc}}^u$ pour tout $t \leq 0$, et on a

$$W_{\text{loc}}^s = \{w \in \mathcal{V}; \lim_{t \rightarrow +\infty} \phi_t(w) = v\}, \quad W_{\text{loc}}^u = \{w \in \mathcal{V}; \lim_{t \rightarrow -\infty} \phi_t(w) = v\}.$$

Définition V.5 Les ensembles W_{loc}^s et W_{loc}^u sont appelés respectivement variété stable locale et variété instable locale.

Nous allons voir une démonstration du théorème de la variété stable (locale) à l'aide de la méthode de Lyapunov-Schmidt.

Méthode de Lyapunov-Schmidt. Cette méthode permet de résoudre un problème non-linéaire du type

$$F(x, \xi) = 0,$$

lorsque le théorème des fonctions implicites ne s'applique pas, c'est-à-dire lorsque la différentielle de F par rapport à x n'est pas inversible. Plus précisément, considérons F de la forme

$$F(x, \xi) = Bx - G(x, \xi)$$

pour $(x, \xi) \in X \times \mathbb{R}^p$, où X est un espace de Banach, B est un opérateur continu de X dans un autre espace de Banach Z et $G \in \mathcal{C}^1(X \times \mathbb{R}^p; Z)$ est telle que²

$$G(0, 0) = 0, \quad \partial_x G(0, 0) = 0.$$

La méthode de Lyapunov-Schmidt est fondée sur une conséquence du

²il suffit en fait que G soit de classe \mathcal{C}^1 au voisinage de $(0, 0)$

Théorème V.7 (de l'application ouverte) *Si B est une application linéaire continue surjective d'un espace de Banach X dans un espace de Banach Y , alors B est ouverte, c'est-à-dire pour tout ouvert Ω de X , $B(\Omega)$ est un ouvert de Y . Par suite, si B est bijective, son inverse B^{-1} est continu.*

(La démonstration de ce théorème se trouve par exemple dans le livre d'analyse fonctionnelle de [5], p. 19–20.) On sait qu'une application linéaire injective admet un inverse à gauche au sens algébrique, de même qu'une application linéaire surjective admet un inverse à droite. Du point de vue de l'analyse fonctionnelle, l'existence d'un inverse à droite (continu) pour une application linéaire continue surjective demande en plus que son noyau admette un *supplémentaire topologique*³, tandis que l'existence d'un inverse à gauche (continu) pour une application linéaire continue injective demande que son image soit fermée et admette un supplémentaire topologique (voir par exemple [5], p. 23). Le lemme suivant rassemble en quelque sorte ces deux résultats en un seul.

Lemme V.2 *Soit B un opérateur linéaire continu d'un espace de Banach \mathcal{X} dans un espace de Banach \mathcal{Z} . Si $\mathcal{X}_0 := \text{Ker}B$ admet un supplémentaire topologique $\mathcal{X}_1 = \text{Ker}P$, où P est un projecteur continu d'image \mathcal{X}_0 , et si $\text{Im}B$ est fermé, alors il existe un opérateur K linéaire continu de $\text{Im}B$ dans \mathcal{X}_1 tel que*

$$BK = \text{Id}_{\text{Im}B} \quad \text{et} \quad KB = \text{Id}_{\mathcal{X}} - P.$$

Dém. La restriction de B à \mathcal{X}_1 est injective d'image $\text{Im}B$. Donc, pour tout $z \in \text{Im}B$, il existe un unique $x \in \mathcal{X}_1$ tel que $z = Bx$. Si on définit l'opérateur K sur $\text{Im}B$, qui est un espace de Banach comme sous-espace fermé d'un espace de Banach, par $Kz = x$, K est évidemment linéaire, et continu d'après le théorème de l'application ouverte ; il vérifie par construction l'identité $KBx = x$ pour tout $x \in \mathcal{X}_1$. De plus, n'importe quel $x \in \mathcal{X}$ s'écrit de façon unique sous la forme $x = x_0 + x_1$ avec $x_0 = Px \in \text{Ker}B$ et $x_1 \in \mathcal{X}_1$. On a alors $Bx = Bx_1$, d'où $KBx = KBx_1 = x_1 = x - Px$ par définition de K . \square

Ainsi donc, si B est un opérateur satisfaisant les hypothèses de ce lemme, l'équation $Bx = G(x, \xi)$ (avec $G \in \mathcal{C}^1(\mathcal{X} \times \mathbb{R}^p; \mathcal{Z})$) équivaut à $G(x, \xi) \in \text{Im}B$ et $B(x_0 + x_1) = G(x_0 + x_1, \xi)$ avec $x_0 = Px \in \mathcal{X}_0 = \text{Ker}B$ et $x_1 \in \mathcal{X}_1$. Si de plus il existe Q un projecteur continu sur $\text{Im}B$, ceci équivaut à $G(x_0 + x_1, \xi) = QG(x_0 + x_1, \xi)$ et $Bx_1 = QG(x_0 + x_1, \xi)$. En composant à gauche par K , cette dernière équation implique $x_1 = x_1 - Px_1 = KQG(x_0 + x_1, \xi)$. On arrive ainsi au système :

$$(1) \quad \begin{cases} x_1 = KQG(x_0 + x_1, \xi), \\ 0 = (\text{Id}_{\mathcal{Z}} - Q)G(x_0 + x_1, \xi), \end{cases}$$

Réciproquement, si on a une solution $(x_1, x_0) \in \mathcal{X}_1 \times \mathcal{X}_0$ de ce système, alors

$$B(x_0 + x_1) = Bx_1 = BKQG(x_0 + x_1, \xi) = QG(x_0 + x_1, \xi) = G(x_0 + x_1, \xi).$$

Or la première équation du système (1) :

$$x_1 = KQG(x_0 + x_1, \xi)$$

³Par définition, un supplémentaire topologique d'un sous-espace *fermé* est un supplémentaire au sens algébrique, qui est de plus fermé et sur lequel il existe un projecteur continu.

se résout par le théorème des fonctions implicites. En effet, par hypothèse sur G , la fonction

$$\begin{aligned} F_1 : \mathcal{X}_1 \times \mathcal{X}_0 \times \mathbb{R}^p &\rightarrow \mathcal{X}_1 \\ (x_1, x_0, \xi) &\mapsto x_1 - K Q G(x_0 + x_1, \xi) \end{aligned}$$

est telle que $F_1(0, 0, 0) = 0$ et $\partial_{x_1} F_1(0, 0, 0) = \text{Id}_{\mathcal{X}_1}$. Donc il existe une application χ_1 de classe \mathcal{C}^1 au voisinage de $(0, 0)$ telle que $F_1(x_1, x_0, \xi) = 0$ avec (x_1, x_0, ξ) voisin de $(0, 0, 0)$ équivaut à $x_1 = \chi_1(x_0, \xi)$. En substituant cette relation dans la deuxième équation du système (1), on voit que la résolution de l'équation de départ $F(x, \xi) = 0$ au voisinage de $(0, 0)$ équivaut à la résolution de

$$(\text{Id}_{\mathcal{Z}} - Q) G(x_0 + \chi_1(x_0, \xi), \xi) = 0$$

au voisinage de $(0, 0)$. Cette dernière équation est trivialement satisfaite dans le cas où $Q = \text{Id}_{\mathcal{Z}}$, c'est-à-dire si $\text{Im} B = \mathcal{Z}$. Ce sera le cas dans la démonstration du théorème de la variété stable par cette méthode.

Dém. [Théorème de la variété stable locale] Sans perte de généralité, on suppose $v = 0$, et on note $A = Df(0)$, $g(u) = f(u) - Au$. Noter que le théorème III.6 s'applique à A , fournissant ainsi des projecteurs continus Π^s et Π^u sur les sous-espaces stable et instable de A respectivement.

On considère les espaces $\mathcal{X} = \mathcal{C}_b^1(\mathbb{R}^-; X)$ et $\mathcal{Z} = \mathcal{C}_b(\mathbb{R}^-; X)$ munis des normes « naturelles » :

$$\|x\|_{\mathcal{X}} = \max(\|x\|_{\infty}, \|x'\|_{\infty}), \quad \|z\|_{\mathcal{Z}} = \|z\|_{\infty} := \max_{t \in \mathbb{R}^-} \|z(t)\|_X.$$

Alors

$$B : x \in \mathcal{X} \mapsto \frac{dx}{dt} - Ax$$

définit un opérateur linéaire continu de \mathcal{X} dans \mathcal{Z} . D'après le théorème III.6, on voit que

$$\text{Ker} B = \text{Im} S$$

où

$$\begin{aligned} S : E^u &\rightarrow \mathcal{X} \\ \xi &\mapsto S\xi : t \in \mathbb{R}^- \mapsto e^{tA} \xi. \end{aligned}$$

Un projecteur sur $\text{Ker} B$ est simplement donné par

$$Px = S \circ \Pi^u x(0).$$

On a bien $P \circ P = P$ puisque $(Px)(0) = \Pi^u x(0)$ et donc

$$P(Px) = S \circ \Pi^u (Px)(0) = S \circ \Pi^u \circ \Pi^u x(0) = S \circ \Pi^u x(0) = Px.$$

D'autre part, $\text{Im} B = \mathcal{Z}$. Soit en effet $z \in \mathcal{Z}$. Trouver $x \in \mathcal{X}$ tel que $z = Bx$ équivaut à résoudre l'équation différentielle avec terme source :

$$\frac{dx}{dt} = Ax + z$$

assortie d'une condition à la limite en $-\infty$ (par définition de \mathcal{X} , x doit être bornée ainsi que sa dérivée en $-\infty$). Une telle solution s'obtient grâce à une formule de type Duhamel :

$$(2) \quad x(t) = \int_{-\infty}^t e^{(t-\tau)A} \Pi^s z(\tau) d\tau - \int_t^0 e^{(t-\tau)A} \Pi^u z(\tau) d\tau, \quad t \leq 0.$$

La première intégrale est bien convergente, et on a la majoration :

$$\left| \int_{-\infty}^t e^{(t-\tau)A} \Pi^s z(\tau) d\tau \right| \leq b_\varepsilon \|z\|_{\mathcal{Z}} \int_{-\infty}^t e^{-(\beta-\varepsilon)(t-\tau)} d\tau = \frac{b_\varepsilon}{\beta-\varepsilon} \|z\|_{\mathcal{Z}}.$$

On a aussi une majoration pour la seconde intégrale :

$$\begin{aligned} \left| \int_t^0 e^{(t-\tau)A} \Pi^u z(\tau) d\tau \right| &\leq c_\varepsilon \|z\|_{\mathcal{Z}} \int_t^0 e^{(\gamma-\varepsilon)(t-\tau)} d\tau = \frac{c_\varepsilon}{\gamma-\varepsilon} \|z\|_{\mathcal{Z}} (1 - e^{(\gamma-\varepsilon)t}) \\ &\leq \frac{c_\varepsilon}{\gamma-\varepsilon} \|z\|_{\mathcal{Z}}. \end{aligned}$$

Donc la formule (2) définit bien une fonction continue bornée sur \mathbb{R}^- ; le fait que $x'(t) = Ax(t) + z(t)$ découle d'un calcul facile, en se rappelant que $\Pi^u + \Pi^s = I_X$; et on déduit une borne pour x' . Donc x appartient à \mathcal{X} et vérifie $Bx = z$.

Étant donné $w \in X$ proche de 0, on peut reformuler à l'aide des opérateurs B et S le fait que $x : t \mapsto \phi_t(w)$ appartienne à \mathcal{X} . En effet, par définition du flot, x doit être solution du problème de Cauchy :

$$(3) \quad \begin{cases} \frac{dx}{dt} = Ax + g(x), \\ x(0) = w. \end{cases}$$

Ce problème a une solution $x \in \mathcal{X}$ si et seulement si $Px = S\xi$ avec $\xi = \Pi^u w$ et $x_1 = x - Px \in \mathcal{X}_1 := \text{Ker}P$ est solution de

$$Bx_1 = g(S\xi + x_1).$$

D'après le lemme V.2, il existe un opérateur K linéaire continu de \mathcal{Z} dans \mathcal{X}_1 tel que

$$BK = \text{Id}_{\mathcal{Z}} \quad \text{et} \quad KB = \text{Id}_{\mathcal{X}} - P.$$

D'après le calcul fait plus haut, on a même une expression explicite de K donnée par la formule (2) :

$$(Kz)(t) = \int_{-\infty}^t e^{(t-\tau)A} \Pi^s z(\tau) d\tau - \int_t^0 e^{(t-\tau)A} \Pi^u z(\tau) d\tau, \quad t \leq 0.$$

Soit alors

$$\begin{aligned} F_1 : \mathcal{X}_1 \times E^u &\rightarrow \mathcal{X}_1 \\ (x_1, \xi) &\mapsto x_1 - Kg(S\xi + x_1). \end{aligned}$$

On a $F_1(0, 0) = 0$ et $d_{x_1} F_1(0, 0) = \text{Id}_{\mathcal{X}_1}$. Donc on peut appliquer le théorème des fonctions implicites à F_1 au voisinage de $(0, 0)$. On en déduit que l'équation $Bx_1 = g(S\xi + x_1)$ admet une solution unique $x_1 = \chi_1(\xi)$ dans un voisinage $\mathcal{V}_1 \times \mathcal{W} \subset \mathcal{X}_1 \times E^u$ de $(0, 0)$, avec χ de classe \mathcal{C}^1 sur \mathcal{W} et $\chi_1(0) = 0$, $d_\xi \chi_1(0) = 0$ (puisque $Dg(0) = 0$). Par suite, pour tout

$w \in (\Pi^u)^{-1}(\mathcal{W})$, le problème de Cauchy (3) admet une unique solution $x \in S\mathcal{W} + \mathcal{V}_1 \subset \mathcal{X}$, donnée par

$$x = S\xi + \chi_1(\xi), \quad \xi = \Pi^u w.$$

En particulier, on a $w = x(0) = \xi + \chi_1(\xi)(0)$. On peut noter que $\chi_1(\xi)(0)$ appartient à E^s , car $x_1 = \chi_1(\xi)$ appartient à $\text{Ker } P$ et donc $S\Pi^u x_1(0) = 0$, ce qui implique $\Pi^u x_1(0) = 0$. L'application

$$\begin{aligned} h : \mathcal{W} \subset E^u &\rightarrow E^s \\ \xi &\mapsto h(\xi) = \chi_1(\xi)(0) \end{aligned}$$

est de classe \mathcal{C}^1 et vérifie $h(0) = 0$, $d_\xi h(0) = 0$. Cela signifie exactement que

$$\{w \in (\Pi^u)^{-1}(\mathcal{W}); t \mapsto \phi_t(w) \in \mathcal{X}\} = \{w = \xi + h(\xi); \xi \in \mathcal{W}\}$$

est une variété tangente à E^u .

Il reste à vérifier que le voisinage $\mathcal{V} = (\Pi^u)^{-1}(\mathcal{W})$, est tel que

$$\phi_t(\mathcal{V}) \subset \mathcal{V}$$

pour tout $t \leq 0$, et qu'on a même

$$\lim_{t \rightarrow -\infty} \phi_t(w) = 0$$

pour tout $w \in \mathcal{V}$. Pour cela on va appliquer le théorème de Lyapunov n°3 à une équation différentielle réduite, pour laquelle 0 est asymptotiquement stable.

Considérons en effet, pour $w \in \mathcal{V}$, $y(t) := \Pi^u(\phi_t(w))$: au moins au voisinage de $t = 0$ on a $y(t) \in \mathcal{W}$ et donc $\phi_t(w) = y(t) + h(y(t))$, ce qui montre que y est solution dans E^u du problème de Cauchy :

$$\begin{cases} \frac{dy}{dt} = A^u y + \Pi^u \circ g(y + h(y)), \\ y(0) = \Pi^u w, \end{cases}$$

où $A^u = Df(v) \circ \Pi^u$. Par construction, l'application $y \mapsto \Pi^u \circ g(y + h(y))$ a une différentielle nulle en $y = 0$. Comme A^u a tout son spectre de partie réelle strictement positive, le théorème V.4 (en renversant le temps) montre que, quitte à réduire \mathcal{W} , $y(t)$ reste dans \mathcal{W} pour tout $t \leq 0$ et de plus que

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

Ceci implique évidemment $\phi_t(w) \in \mathcal{W}$ pour tout $t \leq 0$ et, par continuité de h (qui vaut 0 en 0!),

$$\lim_{t \rightarrow -\infty} \phi_t(w) = 0.$$

La preuve concernant la variété stable est identique : il suffit de changer f en $-f$. \square

Remarque V.3 À partir des variétés stable et instable locale on peut définir les variétés globales :

$$W^s = \bigcup_{t \leq 0} \phi_t(W_{\text{loc}}^s) \quad \text{et} \quad W^u = \bigcup_{t \geq 0} \phi_t(W_{\text{loc}}^u).$$

Cependant, rien ne dit que ce sont des graphes.

3 Variétés invariantes

Plus généralement, on peut trouver des variétés invariantes par le flot d'une équation différentielle en séparant de façon arbitraire le spectre de l'équation linéarisée : ceci ne demande plus l'hyperbolicité du point fixe !

De plus, on peut construire des « variétés globales », définies comme des graphes de fonctions définies sur des sous-espaces entiers (invariants par l'équation linéarisée), à condition que la partie non-linéaire du champ de vecteurs soit globalement assez petite.

C'est l'objet du résultat suivant, que l'on admettra mais dont on montrera comment il implique en particulier le théorème V.6.

Théorème V.8 *Soit $f \in \mathcal{C}^1(\mathbb{R}^n)$, $v \in \mathbb{R}^n$ tel que $f(v) = 0$, et E^- , E^+ des sous-espaces supplémentaires dans \mathbb{R}^n , invariants par $A := Df(v)$ tels que*

$$\sigma(A|_{E^-}) \subset \{\lambda; \operatorname{Re} \lambda \leq -\beta\}, \quad \sigma(A|_{E^+}) \subset \{\lambda; \operatorname{Re} \lambda \geq \gamma\},$$

où $-\beta < \gamma$ (sans rien supposer a priori sur leur signe). En notant $g(x) = f(x) - Df(v) \cdot (x - v)$, si $\|g\|_{\mathcal{C}_b^1(\mathbb{R}^n)}$ est assez petit, il existe un unique $h \in \mathcal{C}^1(E^+; E^-)$ tel que

$$h(0) = 0, \quad Dh(0) = 0, \quad \sup_{\xi \in E^+} \|Dh(\xi)\| < +\infty,$$

et le graphe de h :

$$W^+ := \{v + \xi + h(\xi); \xi \in E^+\}$$

soit invariant par le flot ϕ de l'équation différentielle $u' = f(u)$. De plus, pour tout $w = v + \xi + h(\xi) \in W^+$ et pour tout $\gamma' > \gamma$ il existe $C > 0$ tel que

$$\|\phi_t(w) - v\| \leq C e^{\gamma' t} \|\xi\|.$$

Dans cet énoncé, si les deux paramètres β et γ sont strictement positifs, cela implique en particulier que le point fixe est hyperbolique, et la variété obtenue est précisément la variété instable.

Voyons comment en déduire précisément le théorème V.6. Pour fixer les idées, disons qu'il s'applique pour $\|g\|_{\mathcal{C}^1(\mathbb{R}^n)} \leq \varepsilon$. En général, on ne peut espérer cette inégalité globale. Cependant on peut s'y ramener, quitte à modifier f , et donc aussi g en dehors d'un voisinage de v . En effet, il existe une boule $B(v; \eta)$ dans \mathbb{R}^n telle que

$$\sup_{x \in B(v; \eta)} \|Dg(x)\| \leq \varepsilon/3, \quad \text{d'où} \quad \sup_{x \in B(v; \eta)} \|g(x)\| \leq \eta \varepsilon/3$$

par le théorème des accroissements finis. Soit alors $\chi \in \mathcal{C}^\infty(\mathbb{R}^n; \mathbb{R}^+)$, valant identiquement 1 dans $B(v; \eta/3)$, identiquement 0 en dehors de $B(v; \eta)$ telle que

$$\sup_{x \in \mathbb{R}^n} \|\chi(x)\| = 1, \quad \sup_{x \in \mathbb{R}^n} \|D\chi(x)\| \leq 2/\eta.$$

(L'existence d'un tel χ , appelé *fonction de troncature* est laissée en exercice.) Considérons \tilde{f} définie par

$$\tilde{f}(x) = Df(v) \cdot (x - v) + \tilde{g}(x), \quad \tilde{g} := \chi g.$$

Évidemment on a $\tilde{f} = f$ dans $B(v; \eta/3)$, et donc en particulier $D\tilde{f}(v) = Df(v)$. De plus

$$\|D\tilde{g}(x)\| \leq \|D\chi(x)\| \sup_{x \in B(v; \eta)} \|g(x)\| + \|Dg(x)\| \leq \frac{2}{\eta} \frac{\eta\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Donc le théorème V.8 s'applique à l'équation différentielle $u' = \tilde{f}(u)$, dont le flot coïncide avec celui de $u' = f(u)$ dans $B(v; \eta/3)$.

Définition V.6 Lorsque $\gamma = 0$ dans l'énoncé du théorème V.8, la variété obtenue W^+ est appelée variété centrale-instable (au point v) et généralement notée W^{cu} . On obtient la variété centrale-stable W^{cs} par renversement du temps ($t \mapsto -t$). On définit alors la variété centrale (au point v) par $W^c := W^{cu} \cap W^{cs}$.

Quitte à diminuer ε dans l'énoncé du théorème V.8, on peut démontrer que W^c est le graphe d'une fonction de classe \mathcal{C}^1 au dessus de $E^c := E^{cu} \cap E^{cs}$ où E^{cu} et E^{cs} désignent naturellement les espaces tangents respectivement à W^{cu} et W^{cs} au point v . D'après le théorème V.8, si $w \in W_c$, pour tout $\alpha > 0$ il existe $C > 0$ tel que

$$\forall t \in \mathbb{R}, \quad \|\phi_t(w) - v\| \leq C e^{\alpha t}.$$

Autrement dit, les solutions de problèmes Cauchy avec donnée initiale dans la variété centrale W_c ne peuvent croître que de façon modérée (tout en restant à valeurs dans W_c du fait de son invariance par le flot).

Attention, on rappelle que le théorème V.8 ne s'applique qu'aux champs de vecteurs dont la partie non-linéaire est assez petite dans $\mathcal{C}_b^1(\mathbb{R}^n)$. En fait, on peut démontrer (notamment à l'aide d'une troncature comme ci-dessus), l'existence d'une *variété centrale locale*.

Théorème V.9 (de la variété centrale locale) Si v est un point fixe de l'équation différentielle

$$\frac{du}{dt} = f(u)$$

(avec $f \in \mathcal{C}^1(U)$, U un ouvert de \mathbb{R}^n), dont le flot est noté $\phi_t(w)$, si E^c , le sous-espace central de $Df(v)$, c'est-à-dire la somme des sous-espaces caractéristiques associés aux valeurs propres imaginaires pures de $Df(v)$, est non trivial, il existe un voisinage \mathcal{V} de v dans U et une variété locale W_{loc}^c , tangente à E^c au point v , c'est-à-dire que

$$W_{loc}^c = \{w = v + \xi + h(\xi); \xi \in \mathcal{O} \subset E^c\},$$

où \mathcal{O} est un voisinage de 0 dans E^c et $h : \mathcal{O} \rightarrow E^s \oplus E^u$ (somme des sous-espaces caractéristiques associés aux valeurs propres de parties réelles respectivement strictement négatives et strictement positives) est telle que $h(0) = 0$, $dh(0) = 0$, vérifiant les deux propriétés suivantes :

i). Invariance par le flot : quel que soit $w \in W_{loc}^c$, quel que soit $t \in \mathbb{R}$,

$$\phi_t(w) \in \mathcal{V} \quad \Rightarrow \quad \phi_t(w) \in W_{loc}^c.$$

ii). Lieu des orbites globales proches de v : quel que soit $w \in \mathcal{V}$, si $\phi_t(w)$ est défini et à valeurs dans \mathcal{V} quel que soit $t \in \mathbb{R}$, alors $w \in W_{loc}^c$.

Attention, ce théorème ne prétend pas montrer l'unicité de W_{loc}^c : il peut exister plusieurs variétés centrales locales, tangentes à E^c et vérifiant les propriétés *i) ii)*. Un exemple très simple est celui du système

$$\begin{cases} x' = x^2, \\ y' = -y, \end{cases}$$

dont on peut calculer explicitement les solutions et constater que les orbites sont incluses dans les courbes ayant une équation de la forme

$$y = C \exp(1/x),$$

où C est une constante arbitraire. Le portrait de phase montre ainsi une infinité de courbes tangentes à l'axe Ox et vérifiant les propriétés *i) ii)* : ce sont précisément les graphes de toutes les fonctions de la forme

$$x \mapsto \begin{cases} 0 & x \geq 0, \\ C \exp(1/x) & x < 0. \end{cases} .$$

(Voir la figure V.3.)

4 Introduction aux bifurcations

On va s'intéresser ici aux bifurcations dites de co-dimension 1, c'est-à-dire concernant des champs de vecteurs dépendant d'un paramètre scalaire et dont le portrait de phase (local) change brutalement lorsque le paramètre traverse une valeur critique. D'après le théorème de Hartman-Grobman, ceci ne peut se produire qu'au voisinage d'un point fixe *non hyperbolique*. Nous allons donc considérer

$$\begin{aligned} f : U \times \Lambda \subset \mathbb{R}^n \times \mathbb{R} &\rightarrow \mathbb{R}^n \\ (u, \lambda) &\mapsto f(u, \lambda), \end{aligned}$$

de classe \mathcal{C}^2 , tel qu'il existe $(u_0, \lambda_0) \in U \times \Lambda$ pour lequel $f(u_0, \lambda_0) = 0$ et le spectre de $Df(u_0, \lambda_0)$ (où D désigne la différentielle partielle par rapport à u) intersecte l'axe imaginaire. Plus précisément, nous allons nous placer successivement dans les cas les plus simples :

- i).* la seule valeur propre imaginaire pure de $Df(u_0, \lambda_0)$ est 0 ;
- ii).* la jacobienne $Df(u_0, \lambda_0)$ admet exactement deux valeurs propres imaginaires pures (et donc forcément conjuguées l'une à l'autre).

Sous des hypothèses « génériques », le premier cas correspond à la bifurcation élémentaire dite *col-nœud*, et le second à une *bifurcation de Hopf*.

4.1 Bifurcation col-nœud

Si $Df(u_0, \lambda_0)$ a seulement 0 comme valeur propre imaginaire pure, alors le système *augmenté* :

$$(4) \quad \begin{cases} \frac{du}{dt} = f(u, \lambda), \\ \frac{d\lambda}{dt} = 0, \end{cases}$$

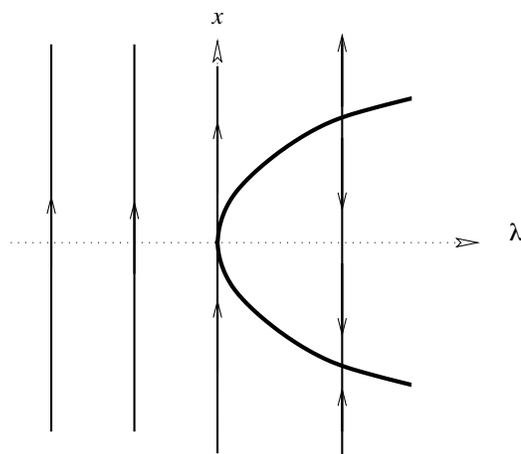


Figure V.1: Portrait de phase du système (5) lorsque $\partial_\lambda g(0, \lambda_0) > 0$ et $\partial_x g(0, \lambda_0) > 0$.

admet une variété centrale locale en (u_0, λ_0) de dimension 2. Le système réduit sur cette variété est de la forme

$$(5) \quad \begin{cases} \frac{dx}{dt} = g(x, \lambda), \\ \frac{d\lambda}{dt} = 0, \end{cases}$$

où l'inconnue x est scalaire, $g(0, \lambda_0) = 0$ et $\partial_x g(0, \lambda_0) = 0$. Supposons en outre que

$$\partial_\lambda g(0, \lambda_0) \neq 0, \quad \partial_x g(0, \lambda_0) \neq 0.$$

Pour fixer les idées, on supposera ces deux quantités positives. Dans ce cas, pour $\lambda < \lambda_0$, la fonction $g(\cdot, \lambda)$ ne s'annule pas au voisinage de 0, pour $\lambda = \lambda_0$ elle s'annule exactement au point 0, et pour $\lambda > \lambda_0$ elle s'annule en deux points de signes opposés. Plus précisément, au voisinage de $(0, \lambda_0)$ dans le plan $\{(x, \lambda) \in \mathbb{R}^2\}$, l'ensemble des zéros de g est une courbe, tangente à l'axe Ox en $(0, \lambda_0)$, avec la concavité tournée vers $\lambda > \lambda_0$. Le portrait de phases du système (5) est particulièrement simple à tracer, car sur les droites d'équation $\lambda = \text{constante}$, il se ramène au portrait de phase d'équations différentielles *scalaires* ! Il est d'usage de tracer ce portrait de phases avec le paramètre (λ) en abscisse, voir la figure V.1.

Le nom de cette bifurcation provient du portrait de phase de l'équation scalaire

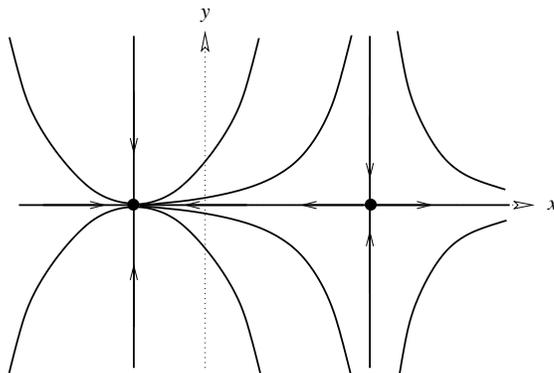
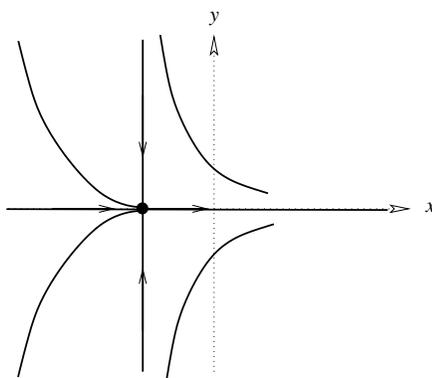
$$\frac{dx}{dt} = g(x, \lambda)$$

augmentée d'une équation hyperbolique, par exemple

$$\frac{dy}{dt} = -y.$$

Le portrait de phase du système plan ainsi obtenu est qualitativement le même que celui du système modèle :

$$(6) \quad \begin{cases} \frac{dx}{dt} = x^2 - \lambda + \lambda_0, \\ \frac{dy}{dt} = -y, \end{cases}$$

Figure V.2: Portrait de phase du système (6) pour $\lambda > \lambda_0$.Figure V.3: Portrait de phase du système (6) pour $\lambda = \lambda_0$.

(aussi appelé forme réduite). Il change radicalement lorsque λ traverse la valeur critique λ_0 , voir les figures V.2, V.3, V.4.

4.2 Bifurcation de Hopf

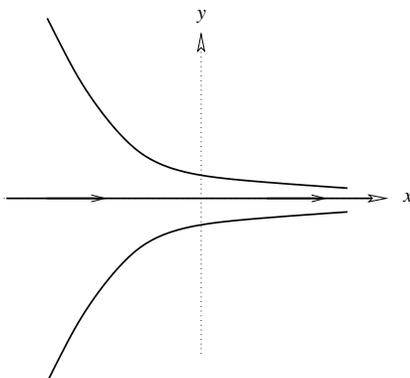
On revient maintenant à un champ f tel que $f(u_0, \lambda_0) = 0$ et l'on suppose que $Df(u_0, \lambda_0)$ admet exactement deux valeurs propres imaginaires pures, $\pm i\omega$ avec $\omega \neq 0$. Ainsi, 0 n'est pas valeur propre, et donc $Df(u_0, \lambda_0)$ est inversible. Le théorème des fonctions implicites montre donc que l'ensemble des zéros de f voisins de (u_0, λ_0) est paramétré par λ : il existe une fonction u_* , définie et de classe \mathcal{C}^2 au voisinage de λ_0 , telle que $u_*(\lambda_0) = u_0$ et

$$f(v, \lambda) = 0 \iff v = u_*(\lambda).$$

À nouveau le système augmenté (4) admet une variété centrale locale en (u_0, λ_0) . Elle est ici de dimension 3, et contient nécessairement la courbe de points fixes $\{(u_*(\lambda), \lambda)\}$. Le système réduit sur cette variété se ramène à un système plan dépendant du paramètre λ . C'est pourquoi on suppose dans ce qui suit $n = 2$.

Théorème V.10 (bifurcation de Hopf) *On suppose que $Df(u_*(\lambda), \lambda)$ admet deux valeurs propres complexes conjuguées valant $\pm i\omega$ en λ_0 et dont la partie réelle change de signe en λ_0 . Alors trois cas sont possibles au voisinage de λ_0 :*

- i). *le point fixe $u_*(\lambda)$ est un centre : le portrait de phase du champ de vecteurs (non-linéaire !) $f(\cdot, \lambda)$ est constitué de cycles concentriques ;*

Figure V.4: Portrait de phase du système (6) pour $\lambda < \lambda_0$.

- ii). on a une bifurcation surcritique : pour $\lambda < \lambda_0$, le champ de vecteurs $f(\cdot, \lambda)$ n'admet aucun cycle autour de $u_*(\lambda)$, et pour $\lambda > \lambda_0$, il admet exactement un cycle autour (et assez proche) de $u_*(\lambda)$; le diamètre de ce cycle est en $\mathcal{O}(\sqrt{\lambda - \lambda_0})$.
- iii). on a une bifurcation souscritique : pour $\lambda > \lambda_0$, le champ de vecteurs $f(\cdot, \lambda)$ n'admet aucun cycle autour de $u_*(\lambda)$, et pour $\lambda < \lambda_0$, il admet exactement un cycle autour (et assez proche) de $u_*(\lambda)$; le diamètre de ce cycle est en $\mathcal{O}(\sqrt{\lambda_0 - \lambda})$.

La démonstration de ce théorème repose sur la théorie des formes normales, qui sort du cadre de ce cours. Nous nous contenterons d'observer ce qui se passe sur la forme réduite de Poincaré-Andronov :

$$\begin{cases} x' = -y + \lambda - x(x^2 + y^2), \\ y' = x + \lambda - y(x^2 + y^2). \end{cases}$$

La valeur critique du paramètre est ici évidemment $\lambda = 0$. En passant en coordonnées polaires, ce système se ramène à

$$\begin{cases} r' = \lambda r - r^3, \\ \theta' = 1, \end{cases}$$

c'est-à-dire finalement à l'équation scalaire

$$\frac{dr}{d\theta} = \lambda r - r^3.$$

Si $\lambda < 0$ cette équation a un seul point fixe, 0, et si $\lambda > 0$, elle en a trois : 0 et $\pm\sqrt{\lambda}$. Attention, $\sqrt{\lambda}$ et $-\sqrt{\lambda}$ ne sont pas des points fixes du système de départ, mais $\sqrt{\lambda}$ est précisément le rayon d'un cycle.

Index

- équation
 - augmentée, 16, 40
 - aux variations, 13
- équation différentielle
 - autonome, 3
 - ordinaire, 3
 - résolue, 3
- analytique, 29
- caractéristique
 - multiplicateur, 36
 - polynôme, 23
 - sous-espace, 23
- champ de vecteurs, 39
- courbe intégrale, 41
- estimation a priori, 40
- Floquet
 - exposant de, 36
- flot, 12
- holomorphe, 27
- hyperbolique
 - matrice, 25
- intégrale première, 39
- Lyapunov-Schmidt
 - méthode de, 57
- minimal
 - polynôme, 23
- multiplicité
 - algébrique, 23
 - géométrique, 23
- opérateur de monodromie, 33
- orbite, 41
 - hétérocline, 43
 - homocline, 43
- phase
 - plan de, 4
 - portrait de, 44
- point
 - d'équilibre, 43
 - fixe, 43
 - singulier, 43
 - stationnaire, 43
- polynôme
 - caractéristique, 23
 - minimal, 23
- propre
 - sous-espace, 23
 - valeur, 23
- résolvant
 - ensemble, 26
- redressement
 - du flot, 44
- semi-simple
 - valeur propre, 23
- simple
 - valeur propre, 23
- solution
 - globale, 10
 - locale, 10
 - maximale, 10
- sous-espace
 - caractéristique, 23
 - propre, 23
 - sous-espace instable, 25
 - sous-espace stable, 25
- spectre, 26
- stabilité, 51
 - asymptotique, 51
 - non-linéaire, 54
 - spectrale, 54
- système
 - gradient, 54

- valeur propre, 23
 - semi-simple, 23
 - simple, 23
- variété
 - centrale, 63
 - centrale locale, 63
 - centrale-instable, 63
 - centrale-stable, 63
 - instable locale, 57
 - stable locale, 57

Bibliographie

- [1] V. Arnold. *Équations différentielles ordinaires*. Éditions Mir, Moscow, 1974. Champs de vecteurs, groupes à un paramètre, difféomorphismes, flots, systèmes linéaires, stabilités des positions d'équilibre, théorie des oscillations, équations différentielles sur les variétés, Traduit du russe par Djilali Embarek.
- [2] V. Arnold. *Chapitres supplémentaires de la théorie des équations différentielles ordinaires*. “Mir”, Moscow, 1984. Translated from the Russian by Djilali Embarek, Reprint of the 1980 edition.
- [3] A. Avez. *Calcul différentiel*. Collection Maîtrise de Mathématiques Pures. Masson, Paris, 1983.
- [4] Martin Braun. *Differential equations and their applications*, volume 11 of *Texts in Applied Mathematics*. Springer-Verlag, New York, fourth edition, 1993. An introduction to applied mathematics.
- [5] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, 1983. Théorie et applications.
- [6] Carmen Chicone. *Ordinary differential equations with applications*, volume 34 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999.
- [7] Shui Nee Chow and Jack K. Hale. *Methods of bifurcation theory*, volume 251 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York, 1982.
- [8] Michel Crouzeix and Alain L. Mignot. *Analyse numérique des équations différentielles*. Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, 1984.
- [9] Jean-Pierre Demailly. *Analyse numérique et équations différentielles*. Collection Grenoble Sciences. Presses Universitaires de Grenoble, Grenoble, 1996.
- [10] Morris W. Hirsch and Stephen Smale. *Differential equations, dynamical systems, and linear algebra*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Pure and Applied Mathematics, Vol. 60.
- [11] Gérard Iooss and Daniel D. Joseph. *Elementary stability and bifurcation theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1990.
- [12] J. D. Murray. *Mathematical biology. I*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2002. An introduction.
- [13] Lawrence Perko. *Differential equations and dynamical systems*, volume 7 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 2001.
- [14] Hervé Reinhard. *Équations différentielles*. Proceedings of Symposia in Pure Mathematics. Dunod, Paris, 1982. Fondements et applications. [Foundations and applications.].

- [15] M. Roseau. *Équations différentielles*. Masson, Paris, 1976. Maîtrise de Mathématiques.
- [16] David Ruelle. *Elements of differentiable dynamics and bifurcation theory*. Academic Press Inc., Boston, MA, 1989.
- [17] Stephen Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 2 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1990.

Table des matières

I	Introduction	3
II	Le problème de Cauchy général	7
1	Existence et unicité locale	7
2	Solutions maximales	10
3	Existence globale	11
4	Dépendance par rapport aux données initiales	11
5	Dépendance par rapport à des paramètres	16
III	Équations linéaires	17
1	Existence globale	17
2	Résolvante	19
2.1	Cas des équations linéaires autonomes	20
2.2	Cas des équations linéaires en dimension finie	21
3	Formule de Duhamel	22
4	Équations linéaires autonomes	23
4.1	Dimension finie	23
4.2	Dimension infinie	26
5	Équations différentielles linéaires à coefficients périodiques	33
IV	Équations autonomes	39
1	Champs de vecteurs	39
2	Flot	41
2.1	Orbites	41
2.2	Redressement du flot	43
3	Portraits de phase	44
4	Ensembles ω -limite	46
4.1	Propriétés générales	46
4.2	Équations dans le plan	47
V	Stabilité des solutions stationnaires	51
1	Théorie de Lyapunov	51
2	Points fixes hyperboliques	56
3	Variétés invariantes	62
4	Introduction aux bifurcations	64
4.1	Bifurcation col-nœud	64
4.2	Bifurcation de Hopf	66