

Introduction aux équations différentielles ordinaires

Pierre Bernhard

12 novembre 2007

1 Introduction

De l'eau s'écoule dans un plan d'eau aux formes complexes, et on connaît la direction et la vitesse du courant en tous les points du plan d'eau. On met des petits bateaux à flotter en divers endroits et on regarde le chemin que chacun suit. Déterminer ces chemins —ou trajectoires— s'appelle intégrer une équation différentielle ordinaire. Les trouver, en examiner les propriétés, est l'objet de ce cours.

1.1 Exemple : flot dans le plan

Examinons donc de plus près ce qu'il s'agit de faire. La vitesse de l'eau en chaque point du plan d'eau peut être vue comme un vecteur indiquant la direction du déplacement de l'eau en cet endroit, et dont la longueur représente la vitesse absolue du déplacement. La donnée d'un vecteur en chaque point d'un domaine (ici le plan d'eau) s'appelle un *champ de vecteurs*. C'est une fonction qui à tout point du domaine fait correspondre un vecteur. Si nous commençons (enfin) à écrire des symboles mathématiques, et pour notre problème situé sur un plan, un point du domaine sera repéré par ses coordonnées x_1 et x_2 , deux nombres réels qu'on pourra considérer conjointement sous le nom collectif de x , et un vecteur sera aussi la donnée de deux nombres, ses composantes sur les mêmes axes, que nous noterons v_1 et v_2 pour rappeler qu'il s'agit de vitesses, et noterons collectivement v . Un champ de vecteur d'un domaine, disons Ω , du plan \mathbb{R}^2 est donc une fonction vectorielle $v = f(x)$ (de Ω dans \mathbb{R}^2) qu'il faut comprendre comme

$$\begin{aligned}v_1 &= f_1(x_1, x_2), \\v_2 &= f_2(x_1, x_2).\end{aligned}$$

Nos petits bateaux se déplacent à la vitesse de l'eau au point où ils sont, de sorte que si on note

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

la position à l'instant t de l'un d'eux qui était au point x_0 (attention c'est un vecteur) à l'instant initial, le chemin qu'il suit satisfait

$$\frac{dx}{dt}(t) = f(x(t)), \tag{1}$$

$$x(0) = x_0. \tag{2}$$

Trouver le chemin $x(t)$ suivi revient donc bien à trouver la solution d'une équation différentielle, et même une équation différentielle vectorielle.

On voit bien que les chemins —on dit les *trajectoires*— suivis par nos petits bateaux sont des courbes partout tangentes à la vitesse v . Intégrer une équation différentielle, c'est donc trouver des courbes partout tangentes au champ de vecteurs, et aussi à quelle vitesse elles sont parcourues. Pour un problème dans un plan comme celui évoqué ici, on peut imaginer de dessiner le vecteur v en tous les points d'une grille suffisamment fine placée sur le domaine considéré, puis, à la main, dessiner de telles courbes. En un temps où les calculateurs capables d'exploiter les méthodes numériques que nous évoquerons bientôt n'existaient pas, cette méthode a été utilisée avec succès. Ainsi résolvait-on des équations différentielles du plan par une méthode graphique.

Avant de formaliser un peu plus, il est temps de commenter un terme du titre de ce cours : *ordinaires*. Nous avons recherché deux fonctions d'une seule variable t . L'équation fait intervenir la dérivée par rapport à cette variable. Ce sont des dérivées *ordinaires* par opposition à des dérivées

partielles qui interviendraient si nos fonctions dépendaient de plusieurs variables, donnant lieu à des équations aux dérivées partielles (EDP). La théorie des équations aux dérivées partielles, dans une certaine mesure plus difficile, mais aussi moins précise dans ses conclusions, est à la théorie des équations différentielles ordinaires ce que la mécanique des milieux continus est à la mécanique rationnelle.

1.2 Orientation du cours

De très nombreux phénomènes physiques, naturels ou artefacts, sont naturellement modélisés par des systèmes d'équations différentielles. Historiquement, la mécanique rationnelle —la mécanique céleste en particulier— a été un moteur puissant du développement de cette théorie. Mais des applications autres ont été considérées dès l'aube de l'analyse, notamment du fait que le calcul des variations, une préoccupation d'ingénieur puisqu'il s'agissait de minimiser un coût, la longueur d'une route, etc, décrivait sa solution en termes d'une équation différentielle.

Le cours qui suit va s'intéresser aux problèmes naturels en ingénierie et modélisation. Ainsi nous ignorerons des résultats d'intérêt surtout mathématique (tels qu'analyticité des solutions par exemple), mais nous accorderons de l'importance aux méthodes qui permettent de montrer des propriétés qualitatives des solutions, (positivité de certaines variables, stabilité, ...) parce que ce sont souvent des propriétés nécessaires d'un modèle.

Il est important, quand on a construit un modèle mathématique d'un phénomène physique, de vérifier les propriétés du modèle, et avant toutes choses qu'il possède les propriétés nécessaires pour qu'il soit une image fidèle du phénomène modélisé. Ainsi, l'existence d'une solution aux équations différentielles écrites est un test, non pas des propriétés du système physique, mais de l'adéquation du modèle mathématique. Il en va souvent de même de la positivité de certaines variables, par exemple si elles représentent une masse, une concentration, ... C'est seulement quand un modèle a été *validé* en vérifiant ses propriétés fondamentales, et que ses "prédictions" sont conformes à l'expérience dans suffisamment de cas, qu'il peut être utilisé pour prédire le résultat d'expériences pas faites, indésirables ou infaisables, pour tester, essayer, optimiser.

1.3 Équations différentielles de \mathbb{R}^n

1.3.1 Forme générale et forme canonique

Toute la théorie que nous développerons porte sur des équations différentielles *vectérielles*, c'est à dire des systèmes d'équations différentielles scalaires couplées.

Pour généraliser l'exemple ci-dessus, on supposera dorénavant que la fonction inconnue à trouver est une fonction vectorielle à n composantes, (ce qui est la même chose que n fonctions scalaires), dépendant d'une seule variable, dite *indépendante*, le temps dans l'exemple des petits bateaux, qu'on appellera généralement le temps, bien que dans certaines applications il puisse s'agir d'une toute autre grandeur. Reprenant la notation de Newton pour les fluxions, on notera d'habitude \dot{x} la dérivée première de $x(t)$ par rapport à t , \ddot{x} , $x^{(3)}$, $x^{(4)}$, ... ses dérivées successives d'ordre supérieur. Ainsi, une équation différentielle très générale pourrait-elle s'écrire

$$F(t, x, \dot{x}, \ddot{x}, \dots, x^{(p)}) = 0,$$

où $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ est une fonction vectorielle donnée. Le plus haut ordre de dérivation apparaissant dans l'équation, p ici, est appelé l'*ordre* de l'équation, le nombre de fonctions inconnues, i.e. la dimension de x , ici n , sa *dimension*.

La forme normale d'une équation d'ordre p est (en appelant y la fonction vectorielle recherchée)

$$y^{(p)} = F_N(t, y, \dot{y}, \dots, y^{(p-1)}). \quad (3)$$

On sait que d'après le théorème des fonctions implicites, la forme générale peut se ramener à cette forme normale au voisinage de tout $(n + 2)$ -uplet $(t, x_0, x_1, \dots, x_p)$ qui annule F pourvu que la dérivée partielle $\partial F / \partial x_p$ (qui est une matrice carrée) soit inversible. L'étude de ce qui se passe au voisinage des points où cette matrice cesse d'être inversible sort du cadre de ce cours.

On va montrer que moyennant une augmentation de la dimension, on peut toujours la ramener à la forme canonique (1).

$$\dot{x} = f(x).$$

Équation d'ordre supérieur Soit donc une équation différentielle de la forme (3), où y est un vecteur de dimension m . (Et donc aussi ses dérivées.) Posons

$$x = \begin{pmatrix} y \\ \dot{y} \\ \vdots \\ y^{(p-1)} \end{pmatrix}.$$

Le vecteur x est donc de dimension mp , puisque formé de p sous-vecteurs de dimension m , qui peuvent être notés comme $y^{(k)} = x_{k+1}$. L'équation différentielle s'écrit ainsi

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3, \\ &\vdots \\ \dot{x}_p &= F_N(t, x_1, x_2, \dots, x_p). \end{aligned}$$

On a ainsi une équation différentielle du premier ordre de dimension $mp = n$. Elle est encore *non autonome*, c'est à dire de la forme

$$\dot{x} = f(t, x) \quad (4)$$

Équation non autonome Soit donc une équation différentielle non autonome, c'est à dire de la forme (4) ci-dessus. Il reste à poser

$$x = \begin{pmatrix} x \\ t \end{pmatrix},$$

c'est à dire $x_k = x_k$ pour tout $k \in \{1, 2, \dots, n\}$ et $x_{n+1} = t$. Adoptons la notation $x_{(1,n)}$ pour désigner les composantes 1 à n de x , et l'équation différentielle devient

$$\dot{x} = \begin{pmatrix} f(x_{n+1}, x_{(1,n)}) \\ 1 \end{pmatrix} = f(x).$$

Dans certains cas, il peut être préférable de garder la forme non autonome explicitement.

Grâce à cette transformation, de tout théorème établi pour la forme canonique on pourra déduire la forme qu'il prend pour un système non autonome.

1.3.2 Conditions initiales

L'exemple des petits bateaux montre suffisamment que tant qu'on n'a pas précisé où on a déposé le bateau au début de sa navigation, il est vain de vouloir déterminer sa position aux instants ultérieurs. Donc une équation différentielle n'a de solution bien définie qu'une fois qu'une *condition initiale* est spécifiée, ou tout autre condition qui singularise une trajectoire parmi toutes les trajectoires possibles. On remarque que, dans les transformations ci-dessus, la connaissance d'une condition initiale complète pour l'équation (4), exige d'avoir les valeurs de y et de ses $p - 1$ premières dérivées pour déterminer une trajectoire.

Réversibilité du temps Soulignons que, contrairement au temps physique, le temps d'une équation différentielle ordinaire est parfaitement réversible. La formule (6) ci-dessous peut être utilisée pour le montrer. Dans la suite, motivé par les applications physiques, nous montrerons tous les résultats en temps positif, on laisse au lecteur d'exprimer les résultats en temps négatif. C'est pourquoi nous prendrons toujours l'origine des temps à l'intérieur de l'intervalle d'étude.

Dans le cas d'une équation différentielle autonome, il n'y a pas de perte de généralité à toujours supposer que la condition initiale est donnée à l'instant $t = 0$. En effet, si on connaît une condition à un autre instant, disons t_0 , soit $x(t_0) = \xi$, il suffit de poser $\tilde{x}(t) = x(t_0 + t)$. On voit immédiatement que \tilde{x} satisfait l'équation différentielle $\dot{\tilde{x}} = f(\tilde{x})$, $\tilde{x}(0) = \xi$.

L'étude de l'ensemble des trajectoires possibles, pour toutes les conditions initiales possibles, est une question intéressante. On appelle *flot* la fonction $\phi(t)$ qui à tout $\xi \in \mathbb{R}^n$ fait correspondre la solution (si elle est bien définie) $x(t) = \phi(t)(\xi)$ de l'équation différentielle initialisée en $x(0) = \xi$.

Dans le cas non autonome, la donnée d'une condition initiale doit spécifier à quel instant t_0 elle est donnée. Le flot est alors donné par la *semi-groupe* $\phi(t, t_0, \cdot)$ qui donne la solution $x(t)$ de l'équation différentielle initialisée en $x(t_0) = \xi$ comme $x(t) = \phi(t, t_0, \xi)$. On peut déjà annoncer un résultat (banal) le concernant :

Proposition 1.1 *Si le semi groupe $\phi(\cdot, \cdot, \cdot)$ est bien défini sur un intervalle \mathcal{T} de la droite réelle, c'est à dire si l'équation différentielle a une unique solution pour toute condition initiale et sur tout sous-intervalle de \mathcal{T} , il satisfait les relations*

$$\begin{aligned} \forall t \in \mathcal{T}, \quad \phi(t, t, \cdot) &= I, \text{ l'identité,} \\ \forall t_0, t_1, t_2 \in \mathcal{T}, \quad \phi(t_2, t_1, \phi(t_1, t_0, \xi)) &= \phi(t_2, t_0, \xi). \end{aligned} \tag{5}$$

La démonstration est laissée en exercice. Le fait qu'une équation différentielle puisse être intégrée en temps positif ou rétrograde (si elle peut être intégrée) justifie que la deuxième relation, dite *relation de semi-groupe* soit valide indépendamment de l'ordre des temps t_0, t_1 et t_2 .

1.3.3 Graphe et portrait de phase

À une solution $t \mapsto x(t)$ de l'équation différentielle, on peut naturellement faire correspondre son graphe. Si x est scalaire, ce graphe se dessine dans le plan (t, x) et constitue une représentation complète de la solution.

À une équation différentielle de dimension 2 autonome, on fait souvent correspondre le tracé des trajectoires dans le plan des x . Ce tracé donne l'aspect géométrique des trajectoires, mais non la vitesse à laquelle elles sont parcourues.

On peut remarquer que le graphe de la solution d'une équation différentielle scalaire est le tracé de la trajectoire de l'équation différentielle augmentée pour la rendre autonome. En particulier, la tangente au graphe au point (t, x) est le vecteur de coordonnées $(1, f(t, x))$

On appelle *portrait de phase* le dessin de l'ensemble des trajectoires, renseigné avec les divers éléments que nous étudierons dans la suite de ce cours : équilibres, bassin d'attraction des équilibres stables, etc. Le cas plan fait l'objet d'une théorie très complète que nous ne ferons qu'esquisser ici, la théorie de Poincaré Bendixon.

2 Existence, unicité, et toutes ces sortes de choses

2.1 Préliminaires

2.1.1 Définition

Soit \mathcal{T} un intervalle de la droite réelle contenant 0. On pourrait être tenté de dire qu'on appelle *solution sur l'intervalle \mathcal{T}* de l'équation différentielle une fonction dérivable $t \mapsto x(t)$ qui satisfait (2) et (1) pour tout $t \in \mathcal{T}$. Toutesfois, il est utile d'étendre très légèrement cette définition. Sans chercher à donner la définition la plus générale possible (dite de *Carathéodory*), nous posons la définition suivante.

Définition 2.1 Soit Ω un ouvert de \mathbb{R}^n . Soit $\Omega_0, \Omega_1, \dots$ une partition ouverte finie de Ω . (C'est à dire que ce sont des ouverts disjoints, et que l'union de leurs fermetures recouvre Ω .) Soit $f : \Omega \rightarrow \mathbb{R}^n$ une fonction continue dans chacun des Ω_i . Soit $x_0 \in \Omega_0$, et soit \mathcal{T} un intervalle ouvert de \mathbb{R} contenant 0. On appelle *solution sur l'intervalle \mathcal{T} de l'équation différentielle (1)(2)* une fonction $t \mapsto x(t)$ de \mathcal{T} dans Ω , satisfaisant (2), telle que $x(t)$ appartient à un des Ω_i sauf en un nombre fini d'instant t_k , (globalement) continue, et satisfaisant (1) dans tout Ω_i .

On semble s'être un peu compliqué la vie... On veut juste permettre à f d'avoir des discontinuités, et donc à $x(\cdot)$ de n'être que dérivable par morceaux, tout en insistant dans ce cas sur son caractère *continu* aux endroits où sa pente est discontinue.

On élève à la dignité de théorème l'affirmation suivante, très simple mais essentielle :

Théorème 2.1 Si $x(\cdot)$ est une solution de (1)(2) sur \mathcal{T} , alors

$$\forall t \in \mathcal{T}, \quad x(t) = x_0 + \int_0^t f(x(s)) ds. \quad (6)$$

Réciproquement, toute fonction $x(\cdot)$ continue satisfaisant (6) est une solution de (1)(2).

Démonstration Soit $x(\cdot)$ une solution de (1)(2). Étant presque partout dérivable, elle est absolument continue, et donc, pour tout t , $x(t) - x(0) = \int \dot{x}(s) ds$. Réciproquement, soit $x(\cdot)$ une fonction satisfaisant (6). Alors elle satisfait évidemment (2). En outre, elle est dérivable en tout point s où $f(x(s))$ est continue, elle est donc absolument continue, et par dérivation de la formule (6) satisfait (1). ■

En conséquence, la formule (6) est *équivalente* à (1)(2).

2.1.2 Contre-exemples

Déterminer la solution, ou une solution, de l'équation (1)(2) s'appelle le *problème de Cauchy*. Ni l'existence ni l'unicité de la solution ne sont évidentes. Donnons tout de suite des contre-exemples très simples. Ici x est scalaire. On considère l'équation différentielle

$$\dot{x} = \begin{cases} -1 & \text{si } x \geq 0, \\ 1 & \text{si } x < 0. \end{cases}, \quad x(0) = 0.$$

Cette équation différentielle n'a pas de solution (au sens ordinaire) : il n'y a pas de fonction dérivable (ou dérivable par morceaux) $x(t)$ qui satisfasse ces conditions.

Quand il y a existence, il n'y a pas nécessairement unicité. À nouveau prenons un x scalaire, et considérons l'équation différentielle

$$\begin{aligned}\dot{x} &= 2\sqrt{|x|}, \\ x(0) &= 0.\end{aligned}$$

On voit immédiatement qu'elle admet les deux solutions $x(t) = 0$ d'une part, et d'autre part $x(t) = t^2$ si $t \geq 0$ et $x(t) = -t^2$ si $t < 0$. En fait elle admet une infinité de solutions. (On laisse en exercice de les déterminer.)

2.1.3 Lignes d'Euler et théorème de Peano

Une intuition naturelle pour étudier l'équation différentielle est la suivante. Choisissons un *pas de temps* très petit δt , et gardons la vitesse constante un temps δt . Donc, pour $t \leq \delta t$, approximations la solution par la fonction $\tilde{x}(t) = x_0 + f(x_0)t$. Posons $t_k = k\delta t$. On a ainsi $\tilde{x}(t_1) = x_0 + f(x_0)\delta t =: x_1$. Puis pour $t \in [t_1, t_2]$, on prend $\tilde{x}(t) = x_1 + (t - \delta t)f(x_1)$, et ainsi de suite. Pour tout k entier, et $s \in [0, \delta t]$, on a donc $\tilde{x}(t_k + s) = x_k + sf(x_k)$.

On a construit ainsi une "ligne brisée" (une fonction affine par morceaux) qui a toutes les chances d'être, si f est assez régulière, une bonne approximation d'une solution. Cette procédure s'appelle la "méthode d'Euler", la fonction engendrée une "ligne d'Euler". Elle sera étudiée plus en détail dans le chapitre 4.2. La raison pour l'introduire ici, outre son caractère naturel et la possibilité qu'elle donne de faire des expériences numériques dès à présent, est que, si f est continue, on peut démontrer¹ par utilisation du théorème d'Arzela Ascoli que, si on fait tendre δt vers zéro, une sous-suite des lignes d'Euler converge uniformément sur un intervalle de temps contenant l'origine vers une solution de (1)(2) sur cet intervalle. On en déduit un premier théorème :

Théorème 2.2 (Peano) *Si la fonction f est continue, il existe au moins une solution de (1)(2) sur un intervalle contenant l'origine.*

Remarque 2.1 *Deux remarques s'imposent :*

1. *Ce théorème est local, c'est à dire qu'il ne dit pas si cette solution peut être étendue à l'intervalle désiré, un contre exemple est l'équation différentielle $\dot{x} = 1 + x^2$, $x(0) = 0$ dont l'unique solution, $x(t) = \tan(t)$ ne peut être prolongée au-delà de l'intervalle $(-\pi/2, \pi/2)$,*
2. *il ne garantit absolument pas l'unicité de la solution évoquée.*

2.2 Théorème de Picard

2.2.1 Équation de point fixe

Soit $\mathcal{T} \subset \mathbb{R}$ un voisinage de l'origine, et \mathcal{X} l'ensemble des fonctions continues de \mathcal{T} dans \mathbb{R}^n . Considérons la fonction \mathcal{F} de \mathcal{X} dans lui-même définie par,

$$\mathcal{F}(x(\cdot))(t) = x_0 + \int_0^t f(x(s)) ds, \quad (7)$$

alors (6) peut s'écrire

$$x(\cdot) = \mathcal{F}(x(\cdot)) \quad (8)$$

¹Mais cette démonstration est trop difficile pour être donnée ici

Une équation de cette forme s'appelle *équation de point fixe*. Elle fait l'objet du théorème fondamental de l'analyse non linéaire, que nous rappelons après la définition suivante :

Définition 2.2 Une application F d'un espace normé X dans lui-même est appelée une contraction si elle est Lipschitz-continue avec un module de Lipschitz strictement inférieur à un, c'est à dire s'il existe un nombre positif $\alpha < 1$ tel que

$$\forall x, y \in X, \quad \|F(x) - F(y)\| \leq \alpha \|x - y\|. \quad (9)$$

Théorème 2.3 (Banach) Soit X un espace vectoriel normé complet —on appelle cela un espace de Banach—, et \mathcal{F} une contraction de X dans lui-même, alors l'équation (8) admet une solution et une seule dans X , et toute suite engendrée par la récurrence $x^{(n+1)} = \mathcal{F}(x^{(n)})$ converge vers ce point.

Démonstration Rappelons qu'un espace vectoriel normé est dit *complet* si les suites de Cauchy y sont toujours convergentes.

Montrons d'abord que si un point fixe existe, il est unique. Supposons donc que x et y sont deux points fixes. De la définition (8) d'un point fixe et de (9), on déduit que

$$\|x - y\| \leq \alpha \|x - y\|,$$

ce qui implique donc que $\|x - y\| = 0$.

Montrons ensuite la convergence de la suite engendrée par $x^{(n+1)} = \mathcal{F}(x^{(n)})$ depuis un $x^{(0)}$ quelconque. En utilisant cette équation et (9), il vient

$$\begin{aligned} \|x^{(2)} - x^{(1)}\| &\leq \alpha \|x^{(1)} - x^{(0)}\| \\ \|x^{(3)} - x^{(2)}\| &\leq \alpha \|x^{(2)} - x^{(1)}\| \leq \alpha^2 \|x^{(1)} - x^{(0)}\| \\ &\vdots \\ \|x^{(n)} - x^{(n-1)}\| &\leq \alpha \|x^{(n-1)} - x^{(n-2)}\| \leq \alpha^{n-1} \|x^{(1)} - x^{(0)}\| \end{aligned}$$

Prenons deux entiers $m < n$, on a

$$\|x^{(n)} - x^{(m)}\| \leq \|x^{(n)} - x^{(n-1)}\| + \|x^{(n-1)} - x^{(n-2)}\| + \dots + \|x^{(m+1)} - x^{(m)}\|,$$

soit, en utilisant les majorations ci-dessus

$$\|x^{(n)} - x^{(m)}\| \leq (\alpha^{n-1} + \alpha^{n-2} + \dots + \alpha^m) \|x^{(1)} - x^{(0)}\| \leq \frac{\alpha^{n-1}}{1 - \alpha} \|x^{(1)} - x^{(0)}\|.$$

Rappelons que $n > m$ et que $\alpha < 1$, de sorte que le membre de droite de l'inégalité ci-dessus tend vers zéro quand $m \rightarrow \infty$. Donc cette suite est de Cauchy.

Par hypothèse, les suites de Cauchy de X convergent, donc il existe x tel que $x^{(n)} \rightarrow x$. Et en prenant la limite des deux membres de (8), on voit immédiatement que $x = \mathcal{F}(x)$. ■

La suite des $x^{(n)}$ est appelée *suite de Picard* associée à l'équation de point fixe (8).

2.2.2 Théorème de Picard

On affirme alors le lemme :

Lemme 2.1 Soit $\mathcal{T} = [a, b] \subset \mathbb{R}$ un voisinage de l'origine, et Ω un domaine borné de \mathbb{R}^n , $x_0 \in \Omega$ donné. Soit \mathcal{V} l'espace des fonctions continues de \mathcal{T} dans Ω doté de la norme du sup.

Si la fonction f est lipshitz continue dans Ω , et pour tout intervalle $[a, b] \subset \mathbb{R}$ suffisamment petit, la fonction (7) définit une contraction de \mathcal{V} dans lui-même.

Démonstration Rappelons que pour deux fonctions continues x et y de $[a, b]$ dans \mathbb{R}^n , la norme du sup est $\|x - y\|_\infty = \sup_{t \in [a, b]} \|x(t) - y(t)\|$. La convergence au sens de cette norme est la convergence uniforme. Et rappelons aussi que l'espace ainsi défini est complet (de Banach). Nous nous restreignons ensuite au sous-espace affine des fonctions satisfaisant $x(0) = x_0$ pour x_0 fixé.

Il faut d'abord se convaincre que si f est Lipschitz continue dans un voisinage Ω , si $[a, b]$ est un intervalle de temps suffisamment court, \mathcal{F} envoie \mathcal{V} dans lui-même. D'abord, il est clair que $\mathcal{F}(x)(0) = x_0$, et que $\mathcal{F}(x)$ est une fonction continue de $[a, b]$ dans \mathbb{R}^n . Il reste à vérifier que $\mathcal{F}(x)(t)$ reste dans Ω . La fonction f étant continue, la norme $\|f(x)\|$ est bornée sur Ω , disons par γ . De ce fait,

$$\|\mathcal{F}(x)(t) - x_0\| \leq \gamma|t| \leq \gamma \max\{-a, b\},$$

de sorte que si $[a, b]$ est un intervalle suffisamment petit, on peut garantir que $\mathcal{F}(x)(t) \in \Omega$.

Évaluons alors la différence ci-dessous, (nous faisons le calcul pour $t > 0$ et laissons le soin au lecteur de vérifier le cas $t < 0$)

$$\begin{aligned} \|\mathcal{F}(x)(t) - \mathcal{F}(y)(t)\| &= \left\| \int_0^t [f(x(s)) - f(y(s))] ds \right\| \\ &\leq \int_0^t \|f(x(s)) - f(y(s))\| ds \\ &\leq \int_0^t \alpha \|x(s) - y(s)\| ds \\ &\leq \int_0^t \alpha \|x - y\|_\infty ds = \alpha t \|x - y\|_\infty. \end{aligned}$$

Ainsi, si $t \in [0, b]$ avec $\alpha b = \beta$, on a une borne uniforme pour la norme dans \mathbb{R}^n de la différence : $\|\mathcal{F}(x)(t) - \mathcal{F}(y)(t)\| \leq \beta \|x - y\|_\infty$ où la dernière norme est déjà celle du sup, d'où on déduit bien, en norme du sup des deux côtés maintenant, $\|\mathcal{F}(x) - \mathcal{F}(y)\|_\infty \leq \beta \|x - y\|_\infty$, et donc que \mathcal{F} est une contraction dès lors que $\beta = \alpha b < 1$. ■

On en déduit le théorème le plus utilisé en matière d'existence d'une solution :

Théorème 2.4 (Picard) *Si la fonction f est Lipschitz continue dans un voisinage de x_0 , il existe, dans un voisinage de $t = 0$, une unique solution $t \mapsto x(t)$ au problème de Cauchy (1)(2).*

Démonstration C'est un corollaire immédiat de ce que les équations (1)(2) sont équivalentes à (6) laquelle est une équation de point fixe pour \mathcal{F} dont on vient de voir que, pour $[a, b]$ suffisamment petit c'est une contraction de \mathcal{V} —qui est complet— dans lui-même, et donc ce point fixe existe et est unique. ■

En écho aux remarques précédentes, remarquons qu'il s'agit encore d'un résultat local en temps, mais cette fois avec unicité. On dit qu'il y a existence et unicité locales de la solution.

On rappelle aussi le fait suivant :

Proposition 2.1 *Si $f \in C^1$, elle est Lipschitz continue sur tout domaine borné. On la dira localement Lipschitz continue, ou L.L.C.*

En effet, ses dérivées partielles étant toutes continues sont bornées sur tout borné. Ainsi la norme d'opérateur de sa matrice jacobienne sera bornée, et son sup est une constante de Lipschitz pour f . (Exercice).

Ainsi, pour $f \in C^0$ on a existence locale, mais pas nécessairement unicité de la solution. Si $f \in C^1$, on a existence et unicité locales.

2.3 Existence globale

2.3.1 Majoration a priori

La question de l'existence globale est liée à celle de savoir comment une solution qui existe localement —i.e. sur un intervalle, disons, $[0, b]$ —, peut cesser d'exister. Or de ce point de vue, les équations différentielles ont un comportement très simple. Le seul phénomène qui puisse faire perdre l'existence est que la solution diverge en temps fini. (On dit parfois qu'elle “explose”.) Nous avons donné un exemple de ce phénomène ci dessus avec l'équation $\dot{x} = 1 + x^2$, $x(0) = 0$.

Nous formalisons ce fait dans le théorème suivant.

Théorème 2.5 *Soit f une fonction L.L.C. de \mathbb{R}^n dans lui-même. Soit $\mathcal{T} \subset \mathbb{R}$ un voisinage ouvert de l'origine qui peut être infini. Soit une famille $K(t)$ de compacts de \mathbb{R}^n définie pour tout $t \in \mathcal{T}$. Si pour tout $t \in \mathcal{T}$ tel que la solution de (1)(2) existe sur $[0, t)$, et pour tout $s < t$, $x(s) \in K(t)$, alors la solution $x(t)$ existe pour tout $t \in \mathcal{T}$.*

Démonstration Supposons donc que la solution $x(\cdot)$ de (1)(2), qui existe est unique localement, satisfait en outre les conditions du théorème. Soit aussi $t^* = \sup\{t \mid \exists x(s), s \in [0, t)\}$. (On ferait de même avec l'inf.) Et supposons que $t^* \in \mathcal{T}$. Nous affirmons que $x(t)$ a une limite bien définie quand $t \uparrow t^*$. Soit en effet une suite $\{t_n\}$ d'instants croissants tendant vers t^* . Cette suite est donc de Cauchy. Remarquons que $\dot{x}(s) = f(x(s))$ est borné sur $[0, t^*)$, puisque f étant continue, elle est bornée sur $K(t^*)$. Soit $\gamma \geq \|\dot{x}(s)\|$. Alors, $\|x(t_k) - x(t_\ell)\| \leq \gamma|t_k - t_\ell|$. Donc la suite des $x(t_k)$ est elle aussi de Cauchy. Donc elle converge. Et cette limite est unique (indépendante de la suite $\{t_k\}$ choisie) car deux suites d'instants tendant vers t^* peuvent être fondues en une seule. En outre, cette limite appartient à $K(t^*)$. Notons la x^* . La fonction f est L.L.C. au voisinage de x^* . Donc il existe un intervalle de \mathbb{R} , $[t^*, t^* + \varepsilon]$, $\varepsilon > 0$, tel que l'équation différentielle ordinaire (1) initialisée en $x(t^*) = x^*$ ait une solution sur cet intervalle. Mais alors, la concaténation de la solution $x(\cdot)$ sur $[0, t^*)$ avec cette dernière constitue une solution de (1)(2) sur $[0, t^* + \varepsilon]$, contredisant la définition de t^* . ■

Quand on a montré qu'une solution satisfait les conditions du théorème, on dit qu'on dispose d'une *majoration a priori*. (On a borné $\|x\|$ sans connaître la solution $x(t)$.) L'utilisation la plus courante de ce théorème est avec $\mathcal{T} = [0, \infty)$ et K constant. Insistons pourtant sur le fait suivant :

Remarque 2.2 *L'énoncé du théorème ci-dessus n'interdit pas que la famille de compacts $K(t)$ soit non bornée quand $t \rightarrow \infty$.*

Corollaire 2.1 *L'une quelconque des deux conditions ci-dessous suffit à assurer l'existence globale de la solution, la première sur \mathbb{R} la seconde sur $[t_0, \infty)$ (f est toujours supposée LLC) :*

1. $\exists a, b > 0 : \forall x, \|f(x)\| \leq a\|x\| + b$,
2. $\exists a, b > 0 : \forall x, \langle x, f(x) \rangle \leq a\|x\|^2 + b$.

(Au demeurant, la première condition ci-dessus implique la deuxième avec des a et b différents.)

Démonstration On laisse le lecteur appliquer le lemme de Gronwall, à $\|x(t)\|$ pour la première condition, à $\|x(t)\|^2$ pour la seconde, sur un intervalle $[0, T]$. ■

2.3.2 Champ rentrant

Une méthode pour montrer que la solution d'une équation différentielle ordinaire ne quitte pas un domaine donné, borné ou pas, est celle du "champ rentrant".

On donne d'abord une définition.

Définition 2.3 On dit qu'un ensemble $\Omega \subset \mathbb{R}^n$ est invariant par (1) si, s'il existe t_1 tel que $x(t_1) \in \Omega$, alors, $\forall t \geq t_1, x(t) \in \Omega$.

Imaginons qu'une variété différentiable \mathcal{S} de dimension $n - 1$ (une courbe du plan, une surface de \mathbb{R}^3 , une hypersurface en dimension supérieure) sépare l'espace ambiant en deux régions disjointes, Ω^- et Ω^+ . Pour tout $x \in \mathcal{S}$, soit $\nu(x)$ une normale à \mathcal{S} , continue et jamais nulle. Elle pointe donc toujours dans la même région définie par \mathcal{S} , disons la région Ω^+ . Alors,

Proposition 2.2 Si

$$\forall x \in \mathcal{S}, \quad \langle \nu(x), f(x) \rangle > 0,$$

alors Ω^+ est invariant par (1).

On hésite à donner une preuve de ce fait tant il est évident si on a compris la notion de courbe intégrale. Esquissons pourtant une preuve. On construit une fonction $V(x)$ de classe C^1 définie dans un voisinage de \mathcal{S} , telle que $\forall x \in \Omega^-, V(x) < 0$, et $\forall x \in \Omega^+, V(x) > 0$. Donc \mathcal{S} est sa surface $V(x) = 0$. On exige en outre que le gradient ∇V soit non nul sur \mathcal{S} . On peut par exemple prendre pour V la distance orientée à \mathcal{S} . Alors, $\forall x \in \mathcal{S}, \nabla V(x)$ est colinéaire à $\nu(x)$, et orienté dans le même sens. Il en découle que

$$\forall x \in \mathcal{S}, \quad \frac{dV(x(t))}{dt} = \langle \nabla V(x), f(x) \rangle > 0.$$

Donc si pour un certain $t_1, x(t_1) \in \Omega^+$, c'est à dire $V(x(t_1)) > 0$, le fait que $x(t)$ sorte de Ω^+ veut dire que $V(x(t))$ change de signe. Mais $V(x(t))$ est continue. Donc elle passe par 0, ce qui est une façon compliquée de dire que pour que $x(t)$ sorte de Ω^+ , il faut qu'il traverse \mathcal{S} . Or, si $x(t_2) \in \mathcal{S}$, alors $V(x(t_2)) = 0$ et $dV(x(t))/dt > 0$ en $t = t_2$. Donc dans un voisinage droit de $t_2, V(x(t)) > 0$, donc $x(t) \in \Omega^+$. (Cette preuve est beaucoup plus compliquée à comprendre que le fait qu'elle démontre.)■

Exemple Un exemple typique d'utilisation de cette notion est comme suit : soit une équation différentielle (1) telle que pour $x_1 = 0$, et pour tout (x_2, \dots, x_n) , on ait $f_1(x) > 0$. Alors si $x_1(0) \geq 0$, $x_1(t)$ reste positif pour tout t (aussi longtemps que la solution $x(t)$ de l'équation différentielle existe).

3 Dépendance des données

Les données de (1)(2) sont f et x_0 . On va s'intéresser à la dépendance de la solution par rapport à l'une et à l'autre. Mais auparavant, il nous faut démontrer un résultat intermédiaire, l'inégalité de Gronwall, qui est un outil fondamental de ce chapitre.

3.1 Inégalité de Gronwall

Lemme 3.1 (Gronwall) Soit $t \mapsto r(t)$ une fonction positive continue de $[0, T]$ dans \mathbb{R}_+ . Si l'inégalité suivante est satisfaite pour tout $t \in [0, T]$:

$$r(t) \leq \rho + \lambda \int_0^t r(s) ds \tag{10}$$

alors, la fonction $r(t)$ vérifie aussi pour tout $t \in [0, T]$

$$r(t) \leq \rho e^{\lambda t}.$$

Démonstration Comme $r(t)$ est par hypothèse continue, elle est bornée sur le compact $[0, T]$. Soit c une telle borne : $\forall s \in [0, T], r(s) \leq c$. En reportant cette majoration dans (10), il vient

$$r(t) \leq \rho + \lambda ct,$$

ou $r(s) \leq \rho + \lambda cs$. Reportons cette deuxième majoration dans (10) de nouveau. Il vient

$$r(t) \leq \rho + \lambda \int_0^t (\rho + \lambda cs) ds = \rho(1 + \lambda t) + c \frac{\lambda^2 t^2}{2}.$$

Reportons à nouveau cette borne (exprimée pour $r(s)$) dans (10), et itérons n fois ce processus. On obtient

$$r(t) \leq \rho \left(1 + \lambda t + \frac{(\lambda t)^2}{2} + \dots + \frac{(\lambda t)^n}{n!} \right) + c \frac{(\lambda t)^{n+1}}{(n+1)!}.$$

Cette inégalité étant vraie pour tout n , il en va de même avec la limite du membre de droite quand $n \rightarrow \infty$, ce qui donne l'inégalité annoncée. ■

3.2 Continuité par rapport à f

La solution d'une équation différentielle ordinaire est continue par rapport au deuxième membre au sens précis suivant.

Théorème 3.1 Soit f une fonction L.L.C. de \mathbb{R}^n dans lui-même. Soit $x_0 \in \mathbb{R}^n$, et $[a, b]$ un voisinage de l'origine dans \mathbb{R} , tel que la solution de (1)(2) est définie sur $[a, b]$. Soit $\{f_n\}$ une suite de fonctions L.L.C. tendant vers f , uniformément sur tout compact. Pour chaque n , soit $x_n(\cdot)$ la solution de l'équation différentielle $\dot{x} = f_n(x)$, $x(0) = x_0$. Alors

- pour n assez grand, les solutions $x_n(\cdot)$ sont définies sur tout $[a, b]$,
- et elles tendent uniformément (sur tout compact si $[a, b]$ est infini) vers la solution $x(\cdot)$ de l'équation (1)(2).

Démonstration On commence par "localiser" le raisonnement de la façon suivante : soit $x(\cdot)$ l'unique solution de (1)(2) sur $[a, b]$. Soit γ un majorant de $\|f(x(t))\|$ pour $t \in [a, b]$. Soit B la boule de \mathbb{R}^n de centre 0 et de rayon 2γ , et λ le module de Lipschitz de f dans B . Soit $T > 0 \in [a, b]$.

On utilise la formule (6) pour écrire, pour tout $t \in [0, T]$,

$$x_n(t) - x(t) = \int_0^t (f_n(x_n(s)) - f(x(s))) ds.$$

Posons $r_n(t) = \|x_n(t) - x(t)\|$. On a

$$r_n(t) \leq \int_0^t \|f_n(x_n(s)) - f(x_n(s))\| + \int_0^t \|f(x_n(s)) - f(x(s))\| ds.$$

Choisissons n suffisamment grand pour que pour tout $x \in B$, $\|f_n(x) - f(x)\| < \delta$. Aussi longtemps que $x_n(t) \in B$, on a

$$r_n(t) \leq \int_0^t \delta \, ds + \int_0^t \lambda r_n(s) \, ds \leq \delta T + \int_0^t \lambda r_n(s) \, ds.$$

Par le lemme de Gronwall, on en déduit

$$r_n(t) \leq \delta T e^{\lambda t} \leq \delta T e^{\lambda T}.$$

Il était loisible de choisir $\delta = \gamma/T \exp(\lambda T)$, ce qui donne $r_n(t) \leq \gamma$, garantissant que l'unique solution $x_n(\cdot)$ reste bien dans B . Mais aussi, nous avons une majoration a priori qui, par le théorème 2.5 garantit l'existence de la solution $x_n(\cdot)$ sur tout $[a, b]$.

Soit $\varepsilon > 0$ donné. On peut choisir $\delta = \varepsilon/T \exp(\lambda T)$, ce qui donnera $r_n(t) \leq \varepsilon$, ce qui démontre la convergence uniforme de $x_n(\cdot)$ vers $x(\cdot)$ sur $[0, T]$. ■

Ce résultat permet aussi d'étudier la continuité de la solution par rapport à un paramètre intervenant dans f . Toutefois, nous utiliserons plutôt le résultat suivant à cette fin.

3.3 Régularité par rapport à la condition initiale

On considère maintenant une fonction f L.L.C. fixée, et on s'intéresse à l'unique solution $x(\cdot)$ de (1)(2), sur un intervalle où elle existe. On note parfois $x(t) = \phi(t)(x_0)$ ou $x(t) = \phi(t, x_0)$ cette solution. Comme déjà indiqué, on se réfère au *semi-groupe* $t \mapsto \phi(t)$ et au *flot* $x \mapsto \phi(t)(x)$. On va considérer la régularité de l'un comme de l'autre.

Dans ce qui suit nous aurons besoin de la matrice jacobienne de f notée

$$Df(x) = F(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}.$$

On aurait aussi besoin des dérivées d'ordre supérieur. On laisse le lecteur qui voudrait le faire se débattre avec les notations.

On hésite à ériger au rang de théorème la remarque évidente suivante :

Remarque 3.1 *Le semigroupe est par définition dérivable, puisque $d\phi(t)(x_0)/dt = f(\phi(t)(x_0))$. De plus, si f est C^1 , c'est à dire qu'il existe F décrite ci-dessus, continue, alors le semigroupe est deux fois dérivable :*

$$\ddot{x}(t) = F(x(t))f(x(t)), \quad \text{soit} \quad \frac{d^2}{dt^2}\phi(t)(x_0) = F(\phi(t)(x_0))f(\phi(t)(x_0)).$$

On pourrait continuer avec les dérivées d'ordre supérieur, en introduisant des notations (tensorielles) adaptées. On laisse le lecteur faire les 4 premiers ordres pour le cas scalaire $n = 1$.

On s'intéresse dorénavant au flot, c'est à dire à la dépendance par rapport aux conditions initiales.

3.3.1 Continuité

On va démontrer que la solution $x(\cdot)$ de (1)(2) dépend continûment de x_0 . Ceci dépend de manière critique de l'unicité de la solution pour les conditions initiales voisines de x_0 (donc du caractère L.L.C. de f), comme le contre-exemple suivant le montre.

Contre exemple Soit l'équation différentielle scalaire (1)(2) avec $f(x) = 2x/\sqrt{|x|}$ si $x \neq 0$, et $f(0) = 0$. On notera que f est continue, en 0 comme ailleurs. La seule solution pour $t \geq 0$ est $x(t) = (x_0 - t)^2$ pour $x_0 < 0$, et $x(t) = (x_0 + t)^2$ pour $x_0 > 0$. Ainsi, pour tout $t > 0$, $x(t)$ subit une discontinuité quand x_0 franchit 0. Mais en $x_0 = 0$, f n'est pas Lipschitz continue. Et l'équation différentielle admet les deux solutions ci-dessus, ainsi que la solution $x(t) = 0$ et une infinité d'autres.

Voici donc le résultat annoncé.

Théorème 3.2 Soit f une fonction L.L.C. de \mathbb{R}^n dans lui-même, $[a, b]$ un intervalle contenant 0 sur lequel la solution de (1)(2) est définie. Quelque soit $t \in [a, b]$, l'application $x_0 \mapsto x(t)$ est continue.

Démonstration On localise le raisonnement comme au théorème 3.1. Soit x_0 et y_0 dans B , et $x(\cdot)$ et $y(\cdot)$ les solutions de l'équation différentielle (1) initialisée en x_0 et y_0 respectivement. Soit $[0, T]$ un intervalle sur lequel ces deux solutions existent. Posons $r_0 = \|y_0 - x_0\|$ et $r(t) = \|y(t) - x(t)\|$. On a, pour tout $t \in [0, T]$,

$$r(t) \leq r_0 + \int_0^t \|f(y(s)) - f(x(s))\| ds,$$

d'où, en utilisant le caractère Lipschitz continu de f , et aussi longtemps que $y(s) \in B$,

$$r(t) \leq r_0 + \lambda \int_0^t r(s) ds,$$

et toujours par l'inégalité de Gronwall, $r(t) \leq r_0 \exp(\lambda T)$. Pour r_0 suffisamment petit, ceci nous fournit une majoration a priori de $y(t)$, garantissant son existence sur $[0, T]$ et qu'il reste bien dans B . En outre, en faisant tendre y_0 vers x_0 , donc r_0 vers 0, on voit que $r(t) \rightarrow 0$, uniformément sur $[0, T]$. ■

3.3.2 Dérivabilité

Nous allons maintenant établir que si f est de classe C^1 , le flot est dérivable par rapport à x_0 , et donner un moyen de calculer cette dérivée. C'est un résultat de "sensibilité" très important dans beaucoup d'applications.

Théorème 3.3 Soit $x \mapsto f(x)$ une fonction de classe C^1 de \mathbb{R}^n dans lui-même, et $F(x)$ sa matrice jacobienne, supposée L.L.C. Soit \mathcal{T} un voisinage ouvert de l'origine dans \mathbb{R} , et $x(t) = \phi(t)(x_0)$ la solution de l'équation différentielle (1)(2) supposée exister sur \mathcal{T} . Alors $\phi(t)(\cdot)$ est dérivable pour tout $t \in \mathcal{T}$, et sa matrice jacobienne $X(t)$ en $x(t) = \phi(t)(x_0)$ peut être obtenue comme solution de l'équation différentielle suivante :

$$\dot{X}(t) = F(x(t))X(t), \quad X(0) = I. \quad (11)$$

Démonstration Soit z un vecteur de \mathbb{R}^n , et $\varepsilon > 0$ destiné à tendre vers zéro. Soit $x_\varepsilon(\cdot)$ la solution de l'équation différentielle initialisée en $x_\varepsilon(0) = x_0 + \varepsilon z$. Pour tout $t \in \mathcal{T}$, on sait que cette solution existe sur $[0, t]$ pour ε suffisamment petit. En utilisant la forme (6), on a

$$x_\varepsilon(t) - x(t) = \varepsilon z + \int_0^t (f(x_\varepsilon(s)) - f(x(s))) ds.$$

Utilisons la formule des accroissements finis pour exprimer l'intégrande : il existe $x'(s) \in [x(s), x_\varepsilon(s)]$ tel que

$$x_\varepsilon(t) - x(t) = \varepsilon z + \int_0^t \frac{\partial f}{\partial x}(x'(s))(x_\varepsilon(s) - x(s)) ds.$$

Divisons par ε en posant $y_\varepsilon(t) := (x_\varepsilon(t) - x(t))/\varepsilon$. On peut aussi utiliser la notation $F(x')$. On obtient ainsi

$$y_\varepsilon(t) = z + \int_0^t F(x'(s))y_\varepsilon(s) ds.$$

On reconnaît que $y_\varepsilon(\cdot)$ satisfait donc l'équation différentielle

$$\dot{y}_\varepsilon(t) = F(x'(t))y_\varepsilon(t), \quad y_\varepsilon(0) = z.$$

Faisons tendre ε vers zéro. Nous savons par le théorème de continuité 3.2 que $x_\varepsilon(\cdot)$ tend uniformément vers $x(\cdot)$, et donc aussi $x'(\cdot)$. Donc la dérivée partielle $F(x'(\cdot))$, qui est L.L.C. par hypothèse, converge uniformément vers $F(x(\cdot))$. Par application du théorème de continuité 3.1, on en déduit que $y_\varepsilon(\cdot)$ converge uniformément sur $[0, t]$ vers la solution $y(\cdot)$ de l'équation différentielle

$$\dot{y}(t) = F(x(t))y(t), \quad y(0) = z.$$

Utilisons ce résultat avec pour z successivement les n vecteurs de base canoniques e_i de \mathbb{R}^n . Appelons X_i le vecteur y obtenu comme ci-dessus pour $z = e_i$. La matrice $X(t)$ dont les $X_i(t)$ sont les colonnes est la matrice jacobienne recherchée. Elle satisfait bien les conditions (11) annoncées. ■

Remarque 3.2 *On notera que l'équation différentielle (11) est non autonome, car la matrice $F(x(t))$ dépend du temps t . Par contre, si on considère simultanément les équation différentielle (1)(2) et (11), alors on a un système autonome.*

On peut donc appliquer ce même théorème au système ainsi obtenu, et de montrer l'existence de dérivées d'ordre supérieur sous la condition que f soit plusieurs fois continûment dérivable.

Nous laissons au lecteur le soin d'exprimer un résultat de dérivée seconde.

3.4 Régularité par rapport à un paramètre

On considère ici le cas d'une équation différentielle ordinaire dont le deuxième membre dépend d'un paramètre, disons p :

$$\dot{x} = f(x, p), \quad x(0) = x_0.$$

On voudrait connaître des résultats de régularité semblables aux précédents, pour la dépendance en p . Le réflexe naturel est d'utiliser le premier paragraphe de ce chapitre, en convertissant une variation de p en une variation de f . Il se trouve qu'on obtiendra des résultats plus complets en utilisant au contraire le deuxième paragraphe, et l'astuce suivante : on fait de p une des variables d'"état", de dérivée nulle. Ainsi l'équation devient

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} f(x, p) \\ 0 \end{pmatrix}, \quad \begin{pmatrix} x(0) \\ p(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ p \end{pmatrix}.$$

qui est une équation du type (1)(2). Des variations de p sont maintenant à traiter comme des variations de condition initiale. On dispose donc de tous les théorèmes de régularité établis les concernant. On

laisse le lecteur faire l'exercice utile de vérifier qu'alors, si on pose $\partial x(t)/\partial p = \xi(t)$, et avec les notations précédentes par ailleurs, on a

$$\dot{\xi} = F(x(t), p)\xi + \frac{\partial f}{\partial p}(x(t), p).$$

4 Résolution des équations différentielles ordinaires

4.1 Résolution analytique

4.1.1 équations linéaires scalaires

On va donner ici quelques faits très simples concernant les équations différentielles scalaires, parce qu'ils sont utilisés abondamment dans toute la théorie. Par contre, nous repoussons au chapitre 6 une étude plus complète des équations différentielles linéaires vectorielles, ou d'ordre plus élevé que un (ce qui est la même chose).

Équations homogènes Le résultat le plus élémentaire, mais absolument fondamental, de la théorie, concerne les équations différentielles linéaires *scalaires*. Érigeons en proposition ce résultat que les lecteurs connaissent :

Proposition 4.1 *L'équation différentielle scalaire*

$$\dot{x} = ax, \quad x(0) = x_0 \in \mathbb{R}$$

admet une solution unique définie sur tout \mathbb{R} :

$$x(t) = x_0 e^{at}.$$

En fait, cette affirmation est un cas particulier de l'équation différentielle non autonome (qu'il vaut mieux ici garder sous sa forme non autonome)

$$\dot{x} = a(t)x, \quad x(t_0) = x_0 \in \mathbb{R},$$

qui a une solution pour tout t tel que $a(\cdot)$ soit intégrable sur $[t_0, t]$, solution qui s'écrit

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

Ce résultat simplissime sera utilisé pour obtenir une formule de représentation moins intuitive concernant une équation différentielle ordinaire de la forme

$$\dot{x} = xg(x), \quad x(0) = x_0 \in \mathbb{R}, \tag{12}$$

où g est supposée bornée et Lipshitz continue dans un intervalle Ω de \mathbb{R} . On a la propriété suivante :

Lemme 4.1 *L'équation différentielle scalaire (12) admet une solution aussi longtemps que $x(t)$ reste dans Ω , solution qui admet la formule de représentation suivante :*

$$x(t) = x_0 e^{\int_{t_0}^t g(x(s)) ds}.$$

En particulier, $x(t)$ ne change jamais de signe.

Équations non homogènes Le cas non homogène le plus banal est celui de l'équation faisant intervenir un terme constant b :

$$\dot{x} = ax + b, \quad x(t_0) = x_0.$$

En effet, il suffit de remarquer qu'en posant $y = x + b/a$, cette équation est l'équation homogène $\dot{y} = ay$ du paragraphe précédent.

Un cas un peu plus intéressant est celui où a et b sont des fonctions $a(t)$ et $b(t)$. Les auteurs parlent parfois de "variation des constantes" à propos de la formule ci-dessous, qui donne l'unique solution de cette équation différentielle, comme on peut s'en convaincre directement en dérivant sous le signe somme. Posons

$$\phi(t_2, t_1) = e^{\int_{t_1}^{t_2} a(s) ds}.$$

La formule annoncée est

$$x(t) = \phi(t, t_0)x_0 + \int_{t_0}^t \phi(t, s)b(s) ds.$$

À nouveau, on renvoie le lecteur au chapitre 6 pour un développement plus complet.

4.1.2 Équations particulières

Il y a toute une ménagerie d'équations différentielles ordinaires particulières pour lesquelles une méthode ad hoc a permis de trouver une solution en termes soit des fonctions élémentaires, soit d'intégrales. Nous donnons en annexe un excellent texte d'Éric Benoit, professeur à l'université de La Rochelle, qui fait un point très exhaustif de cette question.

4.1.3 Séparation des variables, cas scalaire

Un cas particulier de la méthode dite de "séparation des variables" (cf. annexe) doit être mentionné. Considérons à nouveau l'équation (1)(2), scalaire. On peut encore l'écrire

$$\frac{dx}{f(x)} = dt$$

Intégrant de 0 à t , on obtient

$$\int_{x_0}^{x(t)} \frac{1}{f(y)} dy = t.$$

Si l'intégrale peut être exprimée à l'aide des fonctions élémentaires, cela donne la fonction réciproque de $x(\cdot)$... et si cette fonction admet elle-même une réciproque simple...

4.1.4 Intégrale première

À défaut de pouvoir trouver une solution explicite, des informations utiles sur la solution peuvent parfois être déduites de la mise en évidence d'une *intégrale première* (voire de plusieurs) :

Définition 4.1 On appelle *intégrale première* d'une équation différentielle ordinaire (1) (ou "du mouvement") une fonction des variables x_i qui reste constante le long de toute solution $x(\cdot)$ de l'équation différentielle.

Une propriété évidente est ainsi :

Proposition 4.2 Une fonction $V(x)$ de classe C^1 de \mathbb{R}^n dans \mathbb{R} est une intégrale première si et seulement si

$$\forall x \in \mathbb{R}^n, \quad DV(x) \cdot f(x) = 0,$$

qu'il faut bien sur lire

$$\langle \nabla V(x), f(x) \rangle = \sum_{i=1}^n \frac{\partial V}{\partial x_i}(x) f_i(x) = 0.$$

Un exemple particulièrement élémentaire peut illustrer ce propos : Soit A une matrice carrée antisymétrique, c'est à dire telle que $A^t = -A$. Et considérons l'équation différentielle $\dot{x} = Ax$. On vérifie trivialement que la norme $V(x) = \|x\|^2$ est une intégrale première, car

$$DV(x)f(x) = 2x^t Ax = 0.$$

Ainsi les trajectoires de ce système restent sur la sphère de centre l'origine sur laquelle se trouve la condition initiale. Le cas particulier du plan ($n = 2$) correspond à l'oscillateur harmonique

$$\begin{aligned} \dot{x}_1 &= -\omega x_2, \\ \dot{x}_2 &= \omega x_1. \end{aligned}$$

dont les trajectoires sont les cercles centrés à l'origine, ici parcourus dans le sens trigonométrique.

4.2 Résolution numérique

Dans beaucoup de cas, notamment en ingénierie, ce qu'on cherche est une réponse numérique : la connaissance de la trajectoire d'une équation différentielle ordinaire pour une condition initiale particulière. Dans ce cas, disposer d'une solution explicite est un intermédiaire parfois commode, mais somme toute pas très utile, puisque in fine, pour avoir des nombres il faut se référer à des tables. Dans ces cas, une résolution numérique, approximant la solution avec la précision souhaitée, est aussi utile.

Dans tout ce chapitre, f est toujours supposée continûment dérivable le nombre de fois qu'il faut pour que le calcul conduit soit correct.

4.2.1 Lignes d'Euler

La première idée pour approximer la solution d'une équation différentielle ordinaire est celle que nous avons évoquée dès le premier chapitre, des lignes d'Euler. On voit facilement qu'elle n'est pas très satisfaisante. Imaginons que la solution recherchée soit un cercle, disons la solution de l'équation différentielle

$$\begin{aligned} \dot{x} &= -y, \\ \dot{y} &= x. \end{aligned}$$

Chaque segment de notre ligne polygonale est tangent au cercle de centre l'origine passant par son point initial. Donc nous dessinons une spirale qui s'éloigne de l'origine.

Nous allons examiner ce qu'on peut dire de la précision de cette méthode.

Auparavant, faisons remarquer que puisqu'on sait approximer numériquement les intégrales, la suite de Picard fournit un algorithme qu'on peut programmer. Cette méthode n'est pourtant guère usitée.

4.2.2 Ordre d'approximation

Nous allons considérer divers schémas d'approximation, le schéma d'Euler n'étant que le premier d'entre eux. Nous adoptons les notations suivantes, pour alléger (un peu...) des calculs vite fastidieux. L'approximation obtenue pour un pas de temps h sera notée $x^h(t)$. De plus, nous poserons $x(kh) = x_k$ et $x^h(kh) = x_k^h$. De même, nous écrirons $f(x_k) = f_k$ et $f(x_k^h) = f_k^h$. De même pour les dérivées de f . Remarquons que si sa dérivée première f' est la matrice jacobienne bien connue, les dérivées ultérieures sont des tenseurs d'ordre plus élevé. Nous éviterons ici de faire appel à des notations de calcul tensoriel.

Nous avons donc pour le schéma d'Euler

$$\begin{aligned}x_0^h &= x_0, \\x_{k+1}^h &= x_k^h + hf_k^h.\end{aligned}$$

Toutes les comparaisons que nous ferons sont fondées sur les développements de Taylor. En l'occurrence, notons que

$$x_{k+1} = x_k + hf_k + O(h^2).$$

Posons $x_k^h - x_k = \varepsilon_k$. En prenant la différence des deux dernières récurrences, il vient

$$\varepsilon_{k+1} = \varepsilon_k + h(f_k^h - f_k) + O(h^2).$$

Utilisons la formule des accroissements finis : il existe $x'_k \in [x_k, x_k^h]$ tel que

$$\varepsilon_{k+1} = \varepsilon_k + hf'(x'_k)\varepsilon_k + O(h^2).$$

Utilisons cette expression sur k pas, pour obtenir

$$\begin{aligned}\varepsilon_k &= [I + hf'(x'_{k-1})][I + hf'(x'_{k-2})] \cdots [I + hf'(x'_1)][I + hf'(x'_0)]\varepsilon_0 \\ &\quad + [I + hf'(x'_{k-1})][I + hf'(x'_{k-2})] \cdots [I + hf'(x'_1)]O(h^2) \\ &\quad \vdots \\ &\quad + [I + hf'(x'_{k-1})][I + hf'(x'_{k-2})]O(h^2) \\ &\quad + [I + hf'(x'_{k-1})]O(h^2) \\ &\quad + O(h^2).\end{aligned}$$

Notons alors que $\varepsilon_0 = 0$ par hypothèse. De plus, f étant supposée de classe C^1 , sa dérivée f' est globalement bornée dans le domaine (borné) considéré. Il en ressort que chaque ligne du tableau ci-dessus après la première est d'ordre $O(h^2)$. Or il y en a k . Donc ε_k est de l'ordre de kh^2 . Pour intégrer sur un intervalle $[0, T]$ donné, il faut faire T/h pas. Donc l'erreur $\varepsilon(T)$ est de l'ordre de $(T/h)h^2$, soit de l'ordre de h .

On dit que la méthode d'Euler est du premier ordre. En divisant le pas par 2, on réduit à peu près de moitié l'erreur commise. Mais nous allons voir plus loin qu'outre le temps de calcul, recourir à un pas extrêmement fin peut avoir des répercussions néfastes sur la précision par un autre mécanisme pervers.

Refaire ces calculs pour les ordres supérieurs serait absolument obscène... Retenons donc la leçon : si l'erreur sur un pas est de l'ordre de h^ℓ , la méthode est globalement d'ordre $\ell - 1$, parce que'on accumule de l'ordre de $1/h$ erreurs. On ferait un théorème précis de cette affirmation en se servant du théorème de dérivabilité de la solution par rapport aux conditions initiales.

4.2.3 Stabilité

Une idée naturelle pour améliorer la méthode d'Euler est la suivante. Le schéma d'Euler est fondé sur l'approximation

$$\dot{x}_k \simeq \frac{x_{k+1} - x_k}{h},$$

qui est d'ordre 2, d'où l'ordre 1 de la méthode. On pourrait avantageusement remplacer cette approximation par l'approximation d'ordre 3 :

$$\dot{x}_k \simeq \frac{x_{k+1} - x_{k-1}}{2h},$$

qui conduit au schéma parfois appelé "saute mouton" :

$$x_{k+1}^h = x_{k-1}^h + 2hf_k^h.$$

On peut vérifier que ce schéma est bien d'ordre 2. Certes, il faut l'initialiser sur un pas avec une autre méthode. Disons qu'on fait dix pas d'une méthode d'Euler avec un pas 10 fois plus fin.

Programmons-le, et résolvons l'équation différentielle scalaire $x(0) = 10$, $\dot{x} = -0,1x$ avec un pas, disons, de 10^{-3} . Au bout de quelques centaines de pas, la "solution" calculée se met à exhiber une divergence oscillatoire violente. Pour une approximation à l'ordre 2 d'une exponentielle décroissante, voilà qui a de quoi surprendre. Regardons donc l'algorithme précis que nous avons programmé :

$$\begin{aligned}x_0^h &= 10, \\x_1^h &= \text{une approximation de } 10 \exp(-10^{-4}), \\x_{k+1}^h &= x_{k-1}^h - 2 \times 10^{-4} x_k^h.\end{aligned}$$

La dernière équation est une récurrence linéaire du deuxième ordre. On sait comment exprimer sa solution générale : calculer les racines r_1 et r_2 de son équation caractéristique, et la solution s'écrit

$$x_k^h = \alpha_1 r_1^k + \alpha_2 r_2^k.$$

Il reste à ajuster α_1 et α_2 avec les deux premières valeurs x_0^h et x_1^h . L'équation caractéristique est ici

$$z^2 + 2 \times 10^{-4} z - 1 = 0.$$

Les racines ont -1 pour produit. Donc l'une des deux est de module supérieur à 1. C'est la racine négative $r_2 = -10^{-4} - \sqrt{1 + 10^{-8}}$. Ce sont ses puissances successives qu'on a vu dominer la solution et produire une oscillation.

Que se passe-t-il ? Comment concilier cette divergence avec notre affirmation que la méthode est "d'ordre 2" ?

La méthode *serait* d'ordre 2 si les calculs que nous faisons étaient exactement les opérations arithmétiques que nous avons programmées. Si x_1^h était exactement égal à $10 \exp(-10^{-4})$, et si nos calculs n'étaient entachés d'aucune erreur d'arrondi, le coefficient α_2 serait exactement 0, et la racine négative n'aurait aucun effet sur l'approximation numérique.

Mais d'une part, x_1^h n'a été qu'approximé, et surtout, même s'il avait été connu avec une précision infinie, nos calculs sont tous entachés d'erreurs d'arrondi, dues à la précision finie de la représentation des nombres en machine. De ce fait, très vite, les deux x_j^h passés sont tels que notre récurrence récupère un coefficient α_2 non nul. Au bout d'un moment, il devient suffisant pour que l'effet de la racine plus négative que -1 se voie.

Ainsi, notre estimation de l'ordre ignorait les *erreurs d'arrondi*, et ne se préoccupait que de ce qu'on appelle les *erreurs de troncature*. C'est à dire les erreurs que fait notre méthode en supposant qu'on fait les calculs décrits exactement. Mais il y a aussi des erreurs d'arrondi, dont le nombre augmente avec le nombre de pas de calcul. La méthode saute mouton, programmée avec quelques dizaines de pas peut donner d'excellents résultats. Mais avec beaucoup de pas, les erreurs d'arrondi la rendent inutilisable.

On montre que la question de l'accumulation —ou de l'évanouissement— des erreurs d'arrondi peut se juger sur l'équation différentielle obtenue en remplaçant f par un développement au premier ordre autour de la solution théorique, ce qui permet de se ramener à l'étude d'une équation différentielle linéaire. Les erreurs d'arrondi ne s'accumuleront pas si cette équation différentielle est *stable*. Dans ce cas, l'effet des erreurs d'arrondi lointaines s'estompe, et il n'y a pas accumulation.

À une époque où on fait couramment des calculs de 10^{12} opérations arithmétiques, cette question de la stabilité du schéma est primordiale.

4.2.4 Méthodes de Runge Kutta

De nombreux schémas d'ordre plus élevé que 1, mais stables, ont été proposés. Ils diffèrent par des qualités annexes, mais importantes, comme le volume de calculs, la capacité à résoudre des équation différentielle "raides", c'est à dire où l'ordre de grandeur de $\|f\|$ est susceptible de varier très vite dans certaines régions de l'espace, etc. Ce n'est pas l'objet de ce cours de faire un point complet. Nous nous contentons d'indiquer ce que sont les méthodes de Runge Kutta d'ordre 2 et 4, les plus usitées. (Pour les équation différentielle non linéaires, il n'y a pas de méthode de Runge Kutta d'ordre plus élevé que 4. On utilise parfois la méthode de Gear d'ordre 5.)

RK2 La méthode de Runge Kutta d'ordre 2 peut s'écrire

$$\begin{aligned}k_1 &= hf_k^h, \\k_2 &= hf(x_k^h + k_1), \\x_{k+1}^h &= x_k^h + \frac{1}{2}(k_1 + k_2).\end{aligned}$$

soit encore

$$x_{k+1}^h = x_k^h + \frac{h}{2}[f_k^h + f(x_k^h + hf_k^h)]$$

Vérifions son ordre : on peut développer $f(x_k^h + hf_k^h) = f_k^h + hf'(x_k^h)f_k^h + O(h^2)$ pour obtenir

$$x_{k+1}^h = x_k^h + hf_k^h + \frac{h^2}{2}f'(x_k^h)f_k^h + O(h^3).$$

Si on remarque que $f'f = \ddot{x}$, on voit que c'est un développement de Taylor exact à l'ordre deux. L'erreur à chaque pas est donc en h^3 , et, en application de la règle vue plus haut, la méthode est donc bien d'ordre deux.

RK4 La méthode de Runge Kutta d'ordre 4 peut être décrite comme suit :

$$\begin{aligned}k_1 &= hf_k^h, \\k_2 &= hf(x_k^h + \frac{1}{2}k_1),\end{aligned}$$

$$\begin{aligned}
k_3 &= hf(x_k^h + \frac{1}{2}k_2), \\
k_4 &= hf(x_k^h + k_3), \\
x_{k+1}^h &= x_k^h + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).
\end{aligned}$$

On laisse le lecteur vérifier sur le cas scalaire ($n = 1$) qu'elle est alors bien d'ordre 4. Les calculs sont fastidieux. L'ordre est aussi de 4 pour le cas vectoriel. Les calculs sont (bien) pires !

5 Stabilité

Une question importante concernant les équations différentielles ordinaires est leur comportement asymptotique quand $t \rightarrow \infty$. Dans les applications à l'automatique, par exemple, stabiliser le système commandé sera un objectif majeur. Nous allons développer ici la théorie de Lyapunov.

5.1 Préliminaires

La stabilité se jugera toujours au voisinage d'un *équilibre* de l'équation différentielle, c'est à dire un point où l'état peut rester immobile. Qui dit immobile dit dérivée nulle. On donne donc une première définition :

Définition 5.1 On appelle *équilibre* de l'équation différentielle (1) tout point \bar{x} tel que $f(\bar{x}) = 0$.

Pour simplifier la suite de l'exposé, on supposera toujours que 0 est un équilibre, et on étudiera la stabilité à son voisinage. Il n'y a en fait pas de perte de généralité à faire ainsi. Soit en effet \bar{x} un équilibre au voisinage duquel on veut examiner la stabilité de l'équation différentielle. Posons $\tilde{x} = x - \bar{x}$, et $\tilde{f}(\tilde{x}) = f(\tilde{x} + \bar{x})$. L'équation différentielle s'écrit $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$. Et l'équilibre qui nous intéresse est bien $\tilde{x} = 0$.

Nous donnons ensuite des définitions précises des concepts de stabilité que nous allons utiliser.

Définition 5.2 L'équation différentielle ordinaire (1) est dite *Lyapunov stable* (au voisinage de 0) si, quelque soit $\varepsilon > 0$, il existe $\delta > 0$ tel que, si $\|x_0\| < \delta$, alors, pour tout $t > 0$, $\|x(t)\| < \varepsilon$.

Cette première définition dit juste que $x(t)$ reste dans un voisinage de l'origine, aussi petit que l'on veut pourvu qu'il parte assez près de l'origine. Pour exiger qu'en outre il tende vers zéro, on utilise une deuxième définition :

Définition 5.3 L'équation différentielle (1) est dite *Lyapunov asymptotiquement stable* (au voisinage de 0) si elle est Lyapunov stable, et en outre $x(t) \rightarrow 0$ quand $t \rightarrow \infty$.

On s'intéresse enfin aux états à partir desquels la stabilité asymptotique est assurée.

Définition 5.4

1. Pour une équation différentielle (1) Lyapunov asymptotiquement stable en 0, on appelle *bassin d'attraction* de 0 l'ensemble des états initiaux x_0 pour lesquels la solution de l'équation différentielle avec $x(0) = x_0$ converge vers 0 quand $t \rightarrow \infty$.
2. Si le bassin d'attraction de 0 est tout \mathbb{R}^n , on dit que l'équation différentielle est *globalement Lyapunov asymptotiquement stable*. Dans le cas contraire, on dit qu'elle est *localement Lyapunov asymptotiquement stable*.

5.2 Premiers théorèmes de Lyapunov

On donne ici la version classique, “globale” de la théorie de Lyapunov. On laisse le soin au lecteur de se faire des versions “locales” en tant que de besoin. En particulier, les hypothèses que nous faisons sur le comportement à l’infini des fonctions de Lyapunov peuvent être évitées.

L’intuition de la théorie de Lyapunov vient de la considération de l’énergie emmagasinée dans un système électrique ou mécanique. Souvent, des équations de l’électricité ou de la mécanique, on peut déduire directement que cette énergie décroît au cours du temps ; d’où on peut déduire que le système va vers le repos. On va étendre cette idée, la “fonction de Lyapunov” que nous allons introduire jouant le rôle de l’énergie.

On a besoin d’une définition technique :

Définition 5.5 Une fonction réelle ϕ de \mathbb{R}_+ dans lui-même est dite de classe K si elle possède les propriétés suivantes :

1. $\phi(0) = 0$,
2. ϕ est strictement croissante,
3. $\phi(r) \rightarrow \infty$ quand $r \rightarrow \infty$.

(Remarquons que donc $\forall r > 0, \phi(r) > 0$.)

Puis on introduit notre fonction “énergie” :

Définition 5.6 On dira qu’une fonction V de \mathbb{R}^n dans \mathbb{R}_+ est une fonction candidate Lyapunov si il existe deux fonctions de classe K , α et β telles que

$$\forall x \in \mathbb{R}^n, \quad \alpha(\|x\|) \leq V(x) \leq \beta(\|x\|).$$

Cette définition implique notamment que V est nulle à l’origine, et strictement positive partout ailleurs.

On considèrera dans la suite la fonction

$$W(x) = V'(x)f(x) = \langle \nabla V(x), f(x) \rangle. \quad (13)$$

Son rôle est que, si $x(\cdot)$ désigne une solution de (1), alors

$$\frac{dV(x(t))}{dt} = W(x(t)). \quad (14)$$

On énonce alors le premier théorème de Lyapunov :

Théorème 5.1 (Lyapunov) Si $\forall x \in \mathbb{R}^n, W(x) \leq 0$, alors l’équation différentielle (1)(2) admet une solution sur $(0, \infty)$ quelque soit x_0 , et cette solution est Lyapunov stable.

Démonstration

Existence globale La remarque fondamentale est qu’en vertu de (14), $W(x) \leq 0$ implique que $V(x(t))$ est décroissante le long des trajectoires. Donc $\alpha(\|x(t)\|) \leq V(x(t)) \leq V(x(0))$. Mais comme elles tendent vers l’infini à l’infini, les fonctions de classe K ont des ensembles de niveau compacts. Ainsi $\alpha(\|x(t)\|) \leq V(x(0))$ constitue une estimation a priori qui garantit l’existence globale de la solution.

Stabilité Soit $\varepsilon > 0$ donné, et v_0 un nombre positif inférieur à $\alpha(\varepsilon)$ de sorte que $\alpha(r) \leq v_0$ implique $r \leq \varepsilon$. Soit alors δ tel que $\beta(\delta) = v_0$. On affirme qu si $\|x_0\| \leq \delta$, alors $\forall t \geq 0, \|x(t)\| \leq \varepsilon$. En effet, $\|x_0\| \leq \delta$ implique que $V(x_0) \leq \beta(\|x_0\|) \leq v_0$, Mais par hypothèse, $V(x(t))$ est décroissante, de sorte que pour $t > 0$, $V(x(t)) \leq V(x_0) \leq v_0$. Et donc $\alpha(\|x(t)\|) \leq V(x(t)) \leq v_0$. Et v_0 a été choisi pour que ceci implique $\|x(t)\| \leq \varepsilon$. Ceci achève de démontrer le théorème. ■

On renforce la conclusion dans le deuxième de Lyapunov :

Théorème 5.2 (Lyapunov) *S'il existe une fonction de classe $K \gamma$ telle que,*

$$\forall x \in \mathbb{R}^n, \quad W(x) \leq \gamma(\|x\|),$$

alors l'équation différentielle est (Lyapunov) asymptotiquement stable.

Ce théorème sera un corollaire du théorème de Lasalle que nous démontrons maintenant.

5.3 Théorème de Lasalle

Théorème 5.3 (Lasalle) *Sous les mêmes hypothèses qu'au théorème 5.1 (i.e. $\forall x \in \mathbb{R}^n, W(x) \leq 0$), la solution $x(t)$ tend, quand t tend vers l'infini, vers Ω_∞ défini comme le plus grand ensemble invariant par (1) contenu dans $\mathcal{K} = \{x \mid W(x) = 0\}$.*

Toute la suite de ce chapitre est la démonstration de ce théorème.

5.3.1 Existence d'un ensemble limite

Rappelons que nous savons désormais que pour toute condition initiale x_0 , la solution correspondante de (1)(2) existe pour tout $t \geq 0$, et qu'en outre elle reste dans le compact $\mathcal{K}(x_0) = \{x \in \mathbb{R}^n \mid \alpha(\|x\|) \leq V(x_0)\}$. Donc, pour toute suite $\{t_n\}$ (croissante et) tendant vers l'infini, la suite $\{x(t_n)\}$ a des points d'accumulation.

À toute condition initiale x_0 , nous associons l'ensemble $\Omega_\infty(x_0)$ de tous les points d'accumulation des suites $x(t_n)$ pour toutes les suites $\{t_n\}$ croissantes tendant vers l'infini. Nous affirmons :

Proposition 5.1

1. $\Omega_\infty(x_0)$ est fermé.
2. Quelque soit l'ouvert \mathcal{V} contenant Ω_∞ , il existe $T > 0$ tel que, $\forall t \geq T, x(t) \in \mathcal{V}$.

Démonstration

1. Soit $x_n^* \rightarrow \bar{x}$ une suite convergente de $\Omega_\infty(x_0)$. À chaque x_n^* , on peut associer un $x(t_n)$ qui en est à une distance inférieure à $1/n$, et en choisissant $t_n > n$ (et $t_n > t_{n+1}$). Cette suite de $x(t_n)$ converge vers \bar{x} , qui appartient donc à $\Omega_\infty(x_0)$.
2. Si la deuxième proposition est violée, pour tout n il existe $t_n > n$ tel que $x(t_n) \notin \mathcal{V}$. Donc $x(t_n) \in \mathcal{V}^c \cap \mathcal{K}(x_0)$ qui est compact. Donc elle admet un point d'accumulation qui n'est pas dans $\Omega_\infty(x_0)$, ce qui est contradictoire avec la définition de $\Omega_\infty(x_0)$. ■

5.3.2 Invariance de l'ensemble limite

Proposition 5.2 *L'ensemble limite $\Omega_\infty(x_0)$ est invariant par (1).*

Démonstration Soit $x^* \in \Omega_\infty(x_0)$, et $y(t)$ la solution de (1) initialisée en $y(0) = x^*$. Soit $\tau > 0$. Nous allons montrer que $y(\tau) \in \Omega_\infty(x_0)$.

Par définition de $\Omega_\infty(x_0)$, il existe une suite $\{t_n\}$ tendant vers l'infini telle que la solution de (1)(2) satisfasse $x(t_n) \rightarrow x^*$. Considérons $z_n(\cdot)$ la solution de (1) initialisée en $z_n(0) = x(t_n)$. Alors, par l'unicité de la solution de (1), $z_n(\tau) = x(t_n + \tau)$. Mais par la continuité des solutions par rapport aux conditions initiales, $z_n(\tau) \rightarrow y(\tau)$. Donc $x(t_n + \tau) \rightarrow y(\tau)$, et donc $y(\tau) \in \Omega_\infty(x_0)$. ■

On montrerait facilement que pour tout x_0 , $\Omega_\infty(x_0)$ est connexe. On peut donc voir cet ensemble comme *une trajectoire limite* vers laquelle tend $x(t)$.

5.3.3 Invariance de V dans $\Omega_\infty(x_0)$

Proposition 5.3 *La fonction V est constante dans $\Omega_\infty(x_0)$.*

Démonstration Supposons que V ne soit pas constante dans $\Omega_\infty(x_0)$. Soient donc y et z dans $\Omega_\infty(x_0)$, et $V(y) - V(z) = \delta > 0$. V est continue. On peut donc trouver des ouverts $\mathcal{V} \ni y$ et $\mathcal{W} \ni z$ tels que $\forall x \in \mathcal{V}, |V(x) - V(y)| \leq \delta/3$ et $\forall x \in \mathcal{W}, |V(x) - V(z)| \leq \delta/3$. Comme il existe une suite de $x(t_n) \rightarrow z$, il existe t' tel que $x(t') \in \mathcal{W}$. Il existe aussi une suite $t_k \rightarrow \infty$ telle que $x(t_k) \rightarrow y$, et donc un $t'' > t'$ tel que $x(t'') \in \mathcal{V}$. On en déduit que $V(x(t'')) \leq V(z) + \delta/3$ et $V(x(t'')) \geq V(y) - \delta/3$. Donc $V(x(t'')) - V(x(t')) \geq \delta/3 > 0$, ce qui est contradictoire avec le fait que $t'' > t'$ et que V est non croissante le long de toute trajectoire. ■

En conséquence, V est constante dans $\Omega_\infty(x_0)$. Soit donc $y^* \in \Omega_\infty(x_0)$. Considérons la trajectoire de (1) issue de y^* , dont on sait qu'elle reste dans $\Omega_\infty(x_0)$. Elle satisfait $V(y(t)) = V(y^*)$, donc $W(y(t)) = 0$ pour tout $t \geq 0$, (tout t et non "presque tout" par continuité) et en particulier $W(y^*) = 0$.

5.3.4 Conclusion

On a donc montré que

Proposition 5.4 *Pour tout x_0 , $\Omega_\infty(x_0)$ est un ensemble invariant par (1) et contenu dans l'ensemble $\mathcal{K} := \{x \in \mathbb{R}^n \mid W(x) = 0\}$.*

Soit $\Omega_\infty = \cup_{x_0 \in \mathbb{R}^n} \Omega_\infty(x_0)$. Alors, pour tout x_0 , la solution de (1)(2) —c'est à dire toute solution de (1)— tend vers Ω_∞ . Par construction, $\Omega_\infty \in \mathcal{K}$. Et comme union d'ensembles invariants par (1) il est lui-même invariant. (On peut voir Ω_∞ comme l'ensemble des trajectoires de (1) restant dans \mathcal{K} .)
Q.E.D.

Démonstration deuxième théorème de Lyapunov Sous les hypothèses du théorème 5.2, $\mathcal{K} = \{0\}$.

5.4 Théorème de Poincaré

Il s'agit ici d'un des innombrables théorèmes de Poincaré, le fondateur de la théorie qualitative des équations différentielles ordinaires, dont Lyapunov est un émule. Ce théorème affirme qu'au voisinage d'un équilibre, si l'équation différentielle linéaire tangente est asymptotiquement stable, il en va de même localement de l'équation différentielle non linéaire.

Soit donc f telle que $f(0) = 0$, et appelons $A = f'(0)$ son jacobien en 0. Nous affirmons le théorème :

Théorème 5.4 (Poincaré) *Si l'équation différentielle $\dot{x} = Ax$ est asymptotiquement stable, l'équation différentielle (1) est localement asymptotiquement stable.*

Démonstration Cette preuve facile va être l'occasion de voir une version "locale" de la théorie de Lyapunov. On pourrait "localiser" bien plus.

On montrera au chapitre 6 que si l'équation différentielle linéaire est asymptotiquement stable, alors l'équation algébrique en P

$$A^t P + PA = -2I$$

a une solution unique P qui est une matrice symétrique définie positive. Soient $\mu \geq \nu > 0$ sa plus grande et plus petite valeur propre respectivement. Prenons $V(x) = \langle x, Px \rangle$ pour fonction de Lyapunov. On a bien $\nu\|x\|^2 \leq V(x) \leq \mu\|x\|^2$. C'est donc une candidate fonction de Lyapunov valide (avec $\alpha(r) = \nu r^2$ et $\beta(r) = \mu r^2$.) Calculons sa dérivée le long d'une trajectoire, en développant $f(x) = Ax + g(x)$ où $g(x)$ tend vers zéro avec x comme $\|x\|^2$. On obtient

$$W(x) = x^t(A^t P + PA)x + 2x^t P g(x) = -2\|x\|^2 + 2x^t g(x).$$

Le dernier terme ci-dessus tend vers zéro avec x comme $\|x\|^3$. Donc, il existe $\eta > 0$ tel que pour $\|x\| \leq \eta$, on ait $2x^t g(x) < \|x\|^2$, et donc $W(x) < -\|x\|^2$.

Soit alors $\delta = \sqrt{\nu/\mu} \eta \leq \eta$, de sorte que $\mu\delta^2 = \nu\eta^2 =: v$. Le domaine $\Omega := \{x \mid V(x) \leq v\}$ est contenu dans la boule de rayon η . En effet $x \in \Omega$ implique notamment $\nu\|x\|^2 \leq v = \nu\eta^2$ soit $\|x\| \leq \eta$. On voit alors par un argument de champ rentrant qu'il est invariant par (1). En effet, si $V(x(t_1)) = v$, $W(x(t_1)) < 0$ et $V(x(t))$ est strictement décroissante, et donc $V(x(t)) < v$ pour t dans un voisinage droit de t_1 .

On peut dès lors appliquer le deuxième théorème de Lyapunov. ■

5.5 Équations différentielles du plan : aperçu de la théorie de Poincaré Bendixon

Pour les équations différentielles de \mathbb{R}^2 , il existe une théorie qualitative très complète appelée théorie de Poincaré Bendixon. Le fait essentiel lié à la dimension 2 est qu'alors une trajectoire, une courbe, sépare (localement) l'espace en deux parties disjointes. Nous citons ici les deux faits essentiels, et renvoyons le lecteur au document d'Eduardo Sontag en annexe pour une description plus complète.

5.5.1 Classification des équilibres

On rappelle qu'un équilibre est un x^* tel que $f(x^*) = 0$. On regarde alors la matrice A jacobienne de f en x^* . C'est une matrice 2×2 . Elle a donc deux valeurs propres. Celles-ci nous renseignent sur l'aspect du champ de trajectoires au voisinage de l'équilibre :

Proposition 5.5

- Si les valeurs propres ont une partie imaginaire non nulle, les trajectoires sont des spirales, convergentes si la partie réelle est négative, divergentes si la partie réelle est positive.
- Si les valeurs propres sont réelles de même signe, l'équilibre est un "nœud", stable ou instable suivant que ce signe est négatif ou positif.

– Si les signes des valeurs propres réelles sont opposés, l'équilibre est un col.

L'aspect des trajectoires au voisinage d'un nœud ou d'un col est représenté dans la figure ci-dessous, où les axes (obliques) dessinés sont les directions propres. Dans le dessin pour les nœuds, l'axe horizontal correspond à la valeur propre de plus petit module.

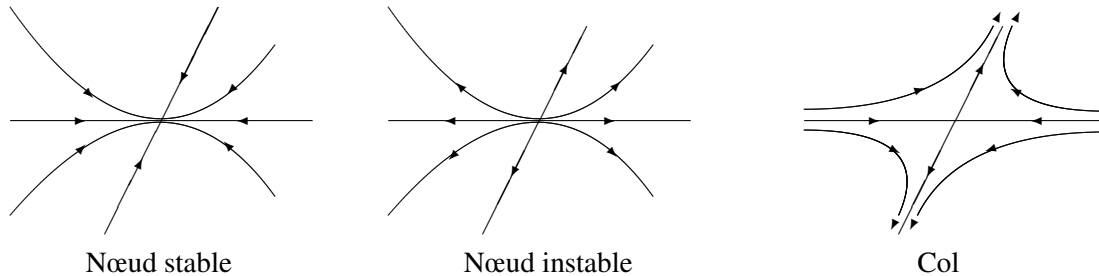


FIG. 1 – Aspect du champ de trajectoires au voisinage d'un équilibre avec valeurs propres réelles

5.5.2 Théorème de Poincaré Bendixon

Supposons qu'un argument de champ rentrant, par exemple, ou du type Lyapunov (voir le chapitre 5) permet de confiner les solutions dans un domaine borné du plan. Alors, le comportement asymptotique des solutions est simple :

Théorème 5.5 (Poincaré Bendixon) *Pour une équation différentielle du plan, si la solution ne quitte pas un domaine borné Ω , alors, quand t tend vers l'infini, soit $x(t)$ tend vers un équilibre stable contenu dans Ω , soit il tend vers un cycle limite stable.*

6 Équations différentielles linéaires

On va, dans ce chapitre, s'intéresser aux équations différentielles linéaires en l'état, et, pour certains résultats, *non autonomes*, c'est à dire où la variable indépendante (que nous appelons toujours conventionnellement le temps, désigné par t) intervient explicitement dans le second membre. Nous avons indiqué dans le premier chapitre qu'on peut toujours se ramener au cas autonome en faisant de t une variable d'état supplémentaire, de dérivée unité. Pourtant, nous ne recourons pas à cet artifice ici, parce que la linéarité en t ne sert à rien. Soit, il y a une dépendance, et elle peut aussi bien être non linéaire, soit il n'y en a pas, et on a des résultats plus complets.

6.1 Équations non autonomes

6.1.1 Existence et unicité

Soit $t \mapsto A(t)$ une fonction de \mathbb{R} (ou d'un intervalle \mathcal{T} de \mathbb{R}) dans $\mathbb{R}^{n \times n}$, l'ensemble des matrices carrées réelles de type $n \times n$. Nous la supposons continue par morceaux. (Mesurable suffirait, mais nous ne cherchons à aucun moment la généralité maximum.) On s'intéresse à l'équation différentielle

$$\dot{x} = A(t)x. \tag{15}$$

On sait que tant qu'elle n'est pas dotée d'une condition initiale du type (2) par exemple, elle n'est pas complètement spécifiée. Mais précisément, notre première préoccupation va être de décrire l'ensemble des trajectoires (ou lignes intégrales) possibles.

Nous vérifions d'abord l'existence des solutions, pour toute condition initiale.

Théorème 6.1 *L'équation différentielle (15) admet une solution globale (sur \mathcal{T}) pour toute condition initiale.*

Démonstration Remarquons d'abord que le second membre est localement lipschitzien en x . Donc on a l'existence locale et l'unicité. Pour montrer l'existence globale, posons $\|x(t)\|^2 = r(t)$. On a

$$\dot{r}(t) = 2x^t(t)\dot{x}(t) = x^t(t)(A(t) + A^t(t))x(t).$$

Comme $A(\cdot)$ est continue par morceaux, les valeurs propres de $A + A^t$ sont bornées sur tout intervalle $[0, t]$. Soit $\mu(t)$ une telle borne. On a alors

$$\forall s \leq t, \quad \dot{r}(s) \leq \mu(s)r(s),$$

soit

$$r(t) \leq r(0) + \int_0^t \mu(s)r(s) ds.$$

Par le lemme de Gronwall 3.1, on en déduit que pour tout t , $r(t) \leq r(0) \exp(\mu(t)t)$, ce qui constitue une majoration a priori et permet d'appliquer le théorème d'existence globale. ■

Insistons sur une conséquence de l'unicité, évidente mais néanmoins essentielle

Corollaire 6.1 *Soit $x(\cdot)$ une trajectoire de (15). Si pour une valeur quelconque de t_1 on a $x(t_1) = 0$, alors $x(t) = 0$ pour tout t .*

6.1.2 Matrice de transition

On s'intéresse maintenant à la famille de toutes les trajectoires possibles, solutions de (15) pour toutes les conditions initiales possibles.

Théorème 6.2 *L'ensemble des trajectoires de l'équation différentielle (15) forme un espace vectoriel de dimension n .*

Démonstration Soient $x_1(\cdot)$ et $x_2(\cdot)$ deux trajectoires, et λ_1 et λ_2 deux réels. On vérifie aisément que $x_\lambda(\cdot) = \lambda_1 x_1(\cdot) + \lambda_2 x_2(\cdot)$ est encore une trajectoire de (15). Donc l'ensemble des trajectoires est bien un espace vectoriel. Montrons que sa dimension est n . Soient donc $n + 1$ trajectoires $x_i(\cdot)$, $i = 1, 2, \dots, n + 1$. Soit t_1 un instant fixé. Les $n + 1$ vecteurs $x_i(t_1)$ sont nécessairement linéairement dépendants. Il existe donc $n + 1$ réels, non tous nuls, λ_i , $i = 1 \dots, n + 1$, tels que $\sum_{i=1}^n \lambda_i x_i(t_1) = 0$. Mais alors la fonction $t \mapsto \sum_{i=1}^n \lambda_i x_i(t)$ est une trajectoire de (15), nulle en t_1 , donc toujours nulle d'après le corollaire ci-dessus. Donc, en tant que fonctions du temps, les $n + 1$ fonctions $t \mapsto x_i(t)$ sont linéairement dépendantes sur \mathbb{R} . ■

On va se donner un moyen d'utiliser ce résultat dans des calculs en dotant cet espace vectoriel d'une base. Soit $\phi_i(\cdot, t_0)$ la solution de (15) dotée de la condition initiale $x_i(t_0) = \phi_i(t_0, t_0) = e_i$, le i -ième vecteur de base de \mathbb{R}^n .

Proposition 6.1 *Les $\phi_i(t, t_0)$ sont indépendants pour tout t .*

Démonstration Si à un instant t_1 , les $\phi_i(t_1, t_0)$ étaient linéairement dépendants, on formerait une combinaison linéaire nulle en t_1 , mais comme chaque $\phi_i(\cdot, t_0)$ est une trajectoire de (15), il en est de même de cette combinaison linéaire (à coefficients constants), et d'après le corollaire elle resterait nulle pour tout t , ce qui est impossible puisqu'en t_0 les ϕ_i sont linéairement indépendants. ■

On a donc n trajectoires indépendantes de (15). Elles forment une base, et donc toute autre trajectoire est une combinaison linéaire à coefficients constants des ces n là.

Définition 6.1 On appelle matrice de transition la fonction matricielle de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}^{n \times n}$ suivante :

$$\Phi(t, t_0) = \left(\phi_1(t, t_0) \mid \phi_2(t, t_0) \mid \cdots \mid \phi_n(t, t_0) \right).$$

C'est une matrice carrée, toujours inversible, dont nous récapitulons les propriétés qui découlent immédiatement de la façon dont nous l'avons construite :

$$\forall t_0, \quad \Phi(t_0, t_0) = I, \quad (16)$$

$$\forall t, t_0, \quad \frac{d\Phi(t, t_0)}{dt} = A(t)\Phi(t, t_0), \quad (17)$$

et nous érigeons en théorème ce résultat maintenant évident :

Théorème 6.3 La matrice définie par (16)(17) est définie de manière unique, et l'unique solution de (15) dotée de la condition initiale $x(t_0) = x_0$ est

$$x(t) = \Phi(t, t_0)x_0. \quad (18)$$

Nous donnons maintenant la propriété peut-être essentielle de la matrice de transition, propriété dite de *semi-groupe* (on verra en fait qu'ici, c'est un groupe).

Proposition 6.2 La matrice de transition Φ satisfait la propriété de groupe

$$\forall t_0, t_1, t_2, \quad \Phi(t_2, t_1)\Phi(t_1, t_0) = \Phi(t_2, t_0). \quad (19)$$

ainsi que

$$\Phi(t_1, t_0)^{-1} = \Phi(t_0, t_1), \quad (20)$$

$$\frac{d\Phi(t_1, t_0)}{dt_0} = -\Phi(t_1, t_0)A(t_0). \quad (21)$$

Démonstration Intégrons (15) depuis $(t_0, x(t_0) = x_0)$ jusqu'à t_1 . Alors $x(t_1) = \Phi(t_1, t_0)x_0 = x_1$. Puis intégrons de $(t_1, x(t_1) = x_1)$ jusqu'à t_2 . On a alors $x(t_2) = \Phi(t_2, t_1)x_1 = \Phi(t_2, t_1)\Phi(t_1, t_0)x_0$. Mais ce faisant, on a tout simplement intégré l'équation différentielle de $(t_0, x(t_0) = x_0)$ à t_2 , de sorte qu'on a $x(t_2) = \Phi(t_2, t_0)x_0$. Comme ces calculs sont corrects quelque soit x_0 , on en déduit (19).

La propriété (20) s'obtient en prenant $t_2 = t_0$ dans (19) et en utilisant (16). La propriété (21) s'obtient en dérivant la précédente par rapport à t_0 . ■

Exercice On laisse en exercice le résultat suivant :

$$\frac{d \det \Phi(t, t_0)}{dt} = \text{tr} A(t) \det \Phi(t, t_0).$$

Remarquons enfin qu'on aurait pu ne manipuler qu'une matrice à une seule variable temps. Supposons que $0 \in \mathcal{T}$, et posons $X(t) = \Phi(t, 0)$. Il découle de (19) et son corollaire (20) que $\Phi(t_1, t_0) = X(t_1)X^{-1}(t_0)$. Cette matrice $X(t)$ est souvent appelée dans les textes de mathématiques *matrice fondamentale* de (15). L'expression "matrice de transition" est plus celle de l'automatique.

6.1.3 Équation non homogène

La formule (22) qui suit est aussi importante que les précédentes. Mais il n'y a pas d'effort à faire pour la retenir, car nous allons en voir une interprétation naturelle.

Nous considérons maintenant l'équation différentielle

$$\begin{aligned}\dot{x} &= A(t)x + b(t), \\ x(t_0) &= x_0,\end{aligned}$$

où $b(\cdot)$ est une fonction vectorielle mesurable donnée. On vérifie par simple dérivation que sa solution existe et peut s'écrire

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^t \Phi(t, s)b(s) ds. \quad (22)$$

Nous allons interpréter (expliquer ?) cette formule comme une *formule de superposition*. D'abord, le membre de droite est la somme (la superposition) du terme $\Phi(t, t_0)x_0$ qui est la solution de l'équation différentielle sans le terme d'excitation b , c'est à dire la contribution de la condition initiale, et de l'intégrale, qui est la contribution du terme d'excitation $b(\cdot)$, la solution de la même équation différentielle à état initial nul. Cette intégrale elle-même peut être vue comme la superposition (la somme) de tous les petits incréments de "condition initiale" $b(s) ds$ à chaque instant s .

6.2 Équation autonome

6.2.1 Équation scalaire d'ordre n

Cas homogène Rappelons en commençant le résultat élémentaire célèbre concernant l'équation différentielle ordinaire scalaire linéaire d'ordre n

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = 0. \quad (23)$$

En essayant des solutions de la forme $y(t) = \exp(\lambda t)$, on trouve que λ doit être une racine du *polynôme caractéristique* $\chi(z)$ de cette équation, défini comme

$$\chi(z) = z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n = 0. \quad (24)$$

Et toutes les combinaisons linéaires de ces exponentielles sont encore des solutions de l'équation différentielle. Mais ceci n'épuise pas le sujet quand il y a moins de n racines distinctes, car alors on a ainsi moins de n solutions indépendantes, et on n'a donc pas une base de l'espace des trajectoires. Le contre exemple

$$\ddot{y} - 2\dot{y} + y = 0,$$

qui admet pour trajectoires toutes les fonctions de la forme

$$y(t) = (at + b)e^t,$$

illustre ce point.

C'est aussi ce que l'on trouve si on considère la transformée de Laplace $Y(s)$ de $y(t)$: l'équation différentielle devient $\chi(s)Y(s) =$ un terme de conditions initiales α , soit $Y(s) = \alpha/\chi(s)$, qu'on peut ensuite développer en éléments simples. Cependant, ce développement en éléments simples ne sera pas ... simple si les racines de χ ne sont pas toutes simples (excusez cette allitération !).

La suite de ce chapitre montrera que si les racines de χ sont les $\lambda_k, k = 1, \dots, \ell$, et que chacune est d'ordre r_k alors les solutions de l'équation différentielle autonome homogène ci-dessus sont toutes de la forme

$$y(t) = \sum_{k=1}^{\ell} p_k(t) e^{\lambda_k t},$$

où les $p_k(\cdot)$ sont des polynômes de degré $r_k - 1$. On a alors juste assez de coefficients à ajuster pour régler y et ses $n - 1$ premières dérivées en 0.

Cas non homogène particulier Étendons enfin l'analyse élémentaire à cette formule célèbre, concernant l'équation différentielle (23) avec un second membre non holonome d'une forme particulière. Convenons de noter l'équation différentielle (23) à l'aide de l'opérateur symbolique de dérivation D comme

$$\chi(D)y = 0.$$

Nous considérons maintenant une équation de la forme

$$\chi(D)y = \sum_{k=1}^{\ell} b_k e^{\mu_k t} \quad (25)$$

où les μ_k sont des constantes. On n'oubliera pas non-plus que les sinus et cosinus sont des combinaisons linéaires d'exponentielles (complexes). Ce résultat concerne donc aussi les deuxièmes membres de la forme d'un "polynôme trigonométrique"

$$\sum_{k=1}^{\ell} b_k e^{\sigma_k t} \cos(\omega_k t + \phi_k).$$

On appelle $Y(\cdot)$ une solution de l'équation homogène (23). On vérifie par simple substitution que

$$y(t) = Y(t) + \sum_{k=1}^{\ell} \frac{b_k}{\chi(\mu_k)} e^{\mu_k t}$$

est solution de (25), et toutes les solutions sont obtenues de cette façon.

Matrice compagnon En faisant la transformation proposée au premier chapitre pour se ramener à une équation différentielle vectorielle du premier ordre, l'équation (23) se met sous la forme $\dot{x} = Ax$ avec

$$x = \begin{pmatrix} y \\ \dot{y} \\ \vdots \\ y^{n-2} \\ y^{n-1} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix}.$$

Cette forme de matrice est appelée *matrice compagnon*. On invite le lecteur à vérifier que son polynôme caractéristique (de matrice) est précisément χ donné par (24). On est ainsi ramené à ce qui suit.

6.2.2 Équation vectorielle stationnaire

Matrice de transition exponentielle On s'intéresse maintenant au cas autonome, appelé plutôt, en automatique, équation différentielle linéaire stationnaire :

$$\dot{x} = Ax. \quad (26)$$

Il se produit dans ce cas une simplification considérable, **qui n'a pas d'équivalent**, même avec une exponentielle d'intégrale, **dans le cas non stationnaire**.

Théorème 6.4 Pour l'équation différentielle linéaire stationnaire (26),

- d'une part, $\Phi(t, t_0)$ n'est fonction que de la différence $t - t_0$, de sorte qu'on peut écrire $\Phi(t - t_0)$ avec une matrice Φ fonction d'une seule variable,
- d'autre part, $\Phi(t)$ est donnée par la formule

$$\Phi(t) = e^{At} = \sum_{k=0}^{\infty} A^k \frac{t^k}{k!}.$$

(Bien sûr, $A^0 = I$.)

On vérifie comme pour le cas scalaire que cette série est toujours normalement convergente (la série des normes converge), et qu'il en va de même pour la série dérivée terme à terme. De sorte que la série des dérivées converge bien vers la dérivée de la somme de la série. Et ceci donne une preuve du théorème par simple inspection.

Il faut pourtant se méfier qu'une exponentielles matricielle peut avoir un comportement inattendu ; par exemple, elle peut se comporter comme t (et non pas une exponentielle en $\exp(at)$) quand $t \rightarrow \infty$. Un exemple type est

$$\exp \left[\begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Un calcul direct permet de vérifier que $d \exp(At)/dt = A \exp(At)$. Ce qui empêche la formule $\exp(\int_{t_0}^t A(s) ds)$ de représenter la matrice de transition d'une équation différentielle non stationnaire est le défaut de commutation des produits $A(t)A(t_0) \neq A(t_0)A(t)$.

Il y a au moins deux façons (d'essayer) de calculer une exponentielle matricielle $\exp(At)$. L'une est la série, particulièrement commode si la matrice A est nilpotente. L'autre est... d'intégrer l'équation différentielle.

Forme générale de la solution Notons d'abord l'effet d'un changement de base dans l'espace des x . Si on pose $y = Tx$ avec T constante inversible, on voit qu'on a

$$\dot{y} = TAT^{-1}y.$$

Or il est immédiat en examinant la série exponentielle que

$$\exp(TAT^{-1}t) = T \exp(At)T^{-1}.$$

En particulier, si le polynôme χ (24) a toutes ses racines distinctes, la matrice compagnon associée à (23) est diagonalisable, et on peut retrouver ainsi le résultat élémentaire donnant y comme combinaison linéaire des exponentielles $\exp(\lambda_k t)$. Nous étendons ici ce résultat, y compris pour le cas de valeurs propres multiples, à toutes les équation différentielle linéaires stationnaires.

Théorème 6.5 Les solutions de l'équation différentielle (26) sont toutes de la forme

$$x(t) = \sum_{k=1}^{\ell} p_k(t) e^{\lambda_k t} \quad (27)$$

où ℓ est le nombre de valeurs propres distinctes, et p_k est un polynôme vectoriel de degré inférieur à l'ordre de multiplicité de λ_k .

Un polynôme de degré d ayant $d + 1$ coefficients, il y a juste n coefficients libres, à ajuster pour satisfaire les conditions initiales.

Démonstration On sait que toute matrice carrée peut être mise par changement de base sous sa forme de Jordan qui est une forme diagonale par blocs,

$$A = T \operatorname{diag}(B_1, B_2, \dots, B_r) T^{-1}$$

où les blocs diagonaux sont eux-mêmes de la forme

$$B_k = \lambda_{\ell_k} I_{m_k} + J_{m_k}.$$

Ici,

- le complexe (ou réel) λ_{ℓ_k} est une valeur propre de A , deux indices ℓ_k différents pouvant correspondre à la même valeur propre (les λ_{ℓ_k} ne sont pas nécessairement tous distincts),
- l'entier m_k est la taille du bloc,
- la somme des m_k correspondant à la même valeur propre est égale à la multiplicité de cette valeur propre (en tant que racine du polynôme caractéristique),
- pour tout entier m , I_m désigne la matrice identité de dimension m ,
- et J_m une matrice carrée de dimension m comportant des zéros partout sauf une sur-diagonale de 1.

Les puissances successives d'une matrice diagonale par blocs sont elles-mêmes diagonales par bloc, ces blocs étant les mêmes puissances des blocs d'origine. Il en va donc de même de l'exponentielle. Il nous faut donc examiner

$$\exp(B_m t) = \sum_k \frac{t^k}{k!} \lambda (I_m + J_m)^k.$$

La remarque importante est que les matrices J_m sont nilpotentes : $J_m^m = 0$. Ainsi, en utilisant la formule du binôme pour développer les sommes $(I_m + J_m)^k$, ce qui est loisible car les deux matrices commutent, on a, sur la diagonale, des termes $\lambda^k t^k / k!$, et hors de la diagonale seulement des termes de puissance inférieure à m . En regardant les puissances successives de J_m , on voit même qu'on aura des termes en t^p sur la p -ème surdiagonale, $p = 1, \dots, m - 1$, et rien d'autre.

Quoi qu'il en soit, les termes diagonaux sommeront en exponentielles $\exp(\lambda t)$, les autres seront des polynômes en t , de degré inférieur à l'ordre de multiplicité de la valeur propre λ . En recombinaison linéairement par les matrices T et T^{-1} , on arrive à la formule annoncée. ■

6.3 Stabilité

6.3.1 Condition spectrale

Une conséquence immédiate de la formule (27) est :

Théorème 6.6 *L'équation différentielle (26) est asymptotiquement stable (vers 0) si et seulement si les valeurs propres de A sont toutes à partie réelle strictement négative.*

En effet, dans ce cas quelque soient les coefficients des polynômes, leur comportement à l'infini est dominé par la décroissance vers zéro des exponentielles. Dans tout autre cas, certaines conditions initiales conduisent à des solutions soit oscillantes (paire valeurs propres imaginaires pures simples) soit divergentes.

6.3.2 Théorie de Lyapunov linéaire

Cependant, ce résultat n'épuise pas la question de la stabilité des systèmes stationnaires. On va voir que la théorie de Lyapunov prend ici une forme renforcée, et donne un résultat très utile.

Essayons une fonction de Lyapunov quadratique, c'est à dire de la forme

$$V(x) = \langle x, Px \rangle$$

où P est une matrice symétrique définie positive. Cette dernière condition implique que toutes les valeurs propres sont (réelles puisque P est, sans perte de généralité, choisie symétrique) positives. Soit α et β la plus petite et la plus grande valeur propre de P , on a $\alpha\|x\|^2 \leq V(x) \leq \beta\|x\|^2$. Donc V est une candidate fonction de Lyapunov valide. Posons alors

$$PA + AB + Q = 0. \tag{28}$$

Cette équation est appelée *équation de Lyapunov*. On voit qu'on a, avec les notations du chapitre 5, $W(x) = -\langle x, Qx \rangle$. On va montrer les résultats suivants :

Théorème 6.7

1. *S'il existe une matrice définie positive P telle que la matrice Q définie par (28) soit définie positive, l'équation différentielle (26) est asymptotiquement stable.*
2. *Réciproquement, si l'équation différentielle (26) est asymptotiquement stable (c'est à dire si la matrice A a toutes ses valeurs propres à partie réelle négative), quelque soit la matrice Q définie positive, l'équation de Lyapunov (28) admet une unique solution P , qui est définie positive.*

Démonstration La première affirmation est une conséquence immédiate du deuxième théorème de Lyapunov. Il nous faut donc démontrer la seconde.

Supposons donc que A a toutes ses valeurs propres à partie réelle négative. (On dit qu'elle est *Hurwicz*.) La forme (27) des trajectoires de notre équation différentielle, qui sont par ailleurs de la forme $x(t) = \Phi(t)x_0$ montre que la norme de la matrice $\Phi(t)$ tend exponentiellement vers zéro. En conséquence, l'intégrale

$$P := \int_0^\infty \Phi^t(s)Q\Phi(s) ds$$

est absolument (ou plutôt "normalement") convergente. Montrons qu'elle satisfait l'équation de Lyapunov. On peut aussi écrire

$$P := \int_t^\infty \Phi^t(s-t)Q\Phi(s-t) ds$$

comme le changement de variable $s \rightarrow s - t$ le montre. Dérivons de part et d'autre par rapport à t , en dérivant sous le signe somme, comme la convergence normale (de l'intégrale dérivée) nous le permet. On obtient

$$0 = -A^t P - PA - Q,$$

qui est l'équation (28).

Reste à montrer l'unicité de la solution P . C'est un théorème d'algèbre linéaire connu (mais pas facile) que l'équation entre matrices carrées $XA + BX + C = 0$ admet une solution X unique pour tout C si et seulement si B n'a pas de valeur propre opposée à une valeur propre de A . Or dans l'équation (28), la matrice qui joue le rôle de B c'est A^t , qui a les mêmes valeurs propres que A . Or, A ne peut pas avoir de paire de valeurs propres opposées puisque, par hypothèse ici, les parties réelles de ses valeurs propres sont toutes (strictement) négatives. ■

Remarque 6.1 *Le théorème faux serait d'affirmer que pour toute matrice P définie positive, la matrice $Q = -A^t P - PA$ est définie positive, comme le montre le contre exemple suivant :*

$$A = \begin{pmatrix} -\frac{1}{4} & -1 \\ 1 & -\frac{1}{4} \end{pmatrix}, \quad P = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 1 \\ 1 & \frac{1}{2} \end{pmatrix}.$$

Manifestement Q n'est pas définie positive, puisque son déterminant est $-1/2$. Pourtant, les valeurs propres de la matrice A sont $-1/4 \pm i$, et donc elle est Hurwicz.