



Stabilité, Bifurcations et Classification des Elastica

Mémoire réalisé par Bertrand DESMONS
pour l'obtention du diplôme de Master en sciences mathématiques

Année académique 2008–2009

Directeur : C. TROESTLER

Rapporteurs : C. MICHAUX
C. GRUMIAU



Je tiens à remercier tous ceux qui m'ont soutenu pour la rédaction de ce mémoire. Je remercie en particulier Christophe Troestler pour le soutien, autant mathématique que personnel, qu'il a pu apporter tout au long de cette année.

Table des matières

Introduction	5
I Stabilité	9
I.1 Définitions générales, et le cas des équations autonomes	10
I.2 Stabilité linéaire	14
I.2.1 Solutions générales de l'équation	15
I.2.2 Points stationnaires et étude de stabilité	18
I.2.3 Espaces stable, instable, central	20
I.3 Stabilité générale — linéarisation	20
I.3.1 Théorème du rang constant	21
I.3.2 Transformation de solutions d'équations différentielles via un difféomorphisme	23
I.3.3 Linéarisation au voisinage d'un point non-singulier	25
I.3.4 Linéarisation au voisinage d'un point singulier, résonances	28
I.3.5 Variétés stable, instable, centrale	31
II Aspects théoriques des bifurcations	37
II.1 Equations différentielles linéaires avec paramètre	37
II.1.1 Le problème dans \mathbb{R}	38
II.1.2 Etude de la dimension 2	39
II.2 Equations non-linéaires avec paramètres	41
II.2.1 Bifurcations locales en dimension 1	43
II.2.2 Une bifurcation locale en dimension 2 : la bifurcation de Hopf	46
II.2.3 Généralisation en dimension supérieure	47
III Algorithmes de continuation	49
III.1 Continuation aux ordres 0 et 1	49
III.2 Continuation par pseudo-longueur d'arc	52
III.2.1 Continuation par longueur d'arc avec paramètre de contrôle	58
III.2.2 Un paramètre de contrôle supplémentaire	59

IV Le problème de l'elastica	61
IV.1 Historique du problème	61
IV.2 Formulation précise du problème	62
IV.3 Etude de la courbure d'un elastica	68
IV.3.1 Courbure d'elastica et équation différentielle	68
IV.3.2 Plan de phases pour la courbure	69
IV.3.3 Quelques notions sur les fonctions elliptiques	71
IV.3.4 Formule explicite de la courbure	75
IV.4 Classification des elastica	78
IV.5 Généralisations du problème	82
IV.5.1 Le problème des courbes dans l'espace	82
IV.5.2 Membranes élastiques sur un substrat mou	83
A Calcul des exponentielles matricielles	85
A.1 Calcul des solutions	85
A.1.1 Cas de matrices à valeurs propres réelles	86
A.1.2 Cas de matrices diagonalisables à valeurs propres complexes	86
A.1.3 Cas de matrices Jordanisables à valeurs propres complexes	88
A.2 Conséquences sur les solutions aux problèmes de Cauchy	91
A.2.1 Vecteurs propres de valeur propre réelle	93
A.2.2 Vecteurs propres généralisés de valeur propre réelle	93
A.2.3 Vecteurs propres de valeurs propres complexes	96
A.2.4 Vecteurs propres généralisés de valeurs propres complexes	97
A.2.5 Vecteurs quelconques	100
B Variétés différentiables	103
B.1 Définitions et propriétés élémentaires	103
B.2 Une proposition concernant le cercle unité	106
B.3 Applications différentiables, vecteurs tangents, différentielles	107

Introduction

De nombreux problèmes physiques, chimiques ou biologiques, peuvent être modélisés à l'aide d'équations différentielles. Dans la majorité des cas, ces équations font intervenir des paramètres extérieurs. Or, lors d'expériences, très souvent, ces paramètres sont quelque peu perturbés : il est en général impossible de réaliser une expérience où les paramètres entrant en jeu sont fixés exactement. Une question naturelle se pose alors : est-ce que le résultat obtenu est « le même » ? Plus précisément, une modification légère du ou des paramètres entraîne-t-elle une modification légère des résultats obtenus, de telle sorte que les propriétés que l'on peut en déduire soient identiques ?

En toute généralité, la réponse à ces questions est négative : par exemple, dans un cadre physique, poser un crayon verticalement sur sa pointe sur une surface plane, entraîne inévitablement sa chute, bien que théoriquement, s'il était parfaitement posé, il devrait rester en position verticale. Une légère variation dans la position du crayon lui donne donc un comportement qui n'est pas similaire sans la variation. Il existe bien une position du crayon de telle sorte qu'il reste en place, mais il est impossible de se rapprocher de cette position. Ce point particulier, stationnaire, n'est donc pas stable. Cependant, si dans l'exemple précédent, le crayon n'est pas bien taillé, de telle sorte que sa pointe présente un « plat », alors il est possible de le faire tenir sur sa pointe (cela est d'autant plus difficile que le plat en question est petit). La position verticale se révèle alors stable, puisque des positions proches de la verticale amènent le crayon en cette position.

Les positions de repos du crayon correspondent, d'un point de vue mathématique, aux solutions constantes de l'équation différentielle associée. La stabilité de telles positions s'exprime donc à travers les solutions ; l'étude mathématique de l'équation différentielle associée permet alors de caractériser ce phénomène. Dans ce mémoire, nous nous proposons de donner quelques outils permettant cette étude. Dans un premier temps, nous considérons les équations différentielles ne faisant pas intervenir de paramètre ; c'est-à-dire, les équations de la forme $\partial u = f(u)$. (Dans l'exemple précédent, cela revient à fixer le crayon, ou plus exactement, d'avoir toujours la même pointe.) Nous définissons diverses notions de stabilité et nous étudions ces notions pour les solutions constantes, d'abord dans

un cadre linéaire. Nous énonçons alors des critères précis de stabilité dans ce cas ; ces critères nous permettent alors d'aborder la stabilité dans le cas des équations différentielles ordinaires non-linéaires. La stabilité étant une notion locale (puisque l'on discute de positions proches), on espère pouvoir étudier la stabilité d'un système décrit par une fonction f générale en considérant le système donné par l'approximation linéaire de f au point stationnaire considéré. Cependant, l'information ainsi reçue n'est pas toujours fiable ; parfois même elle est erronée. Nous donnons donc des conditions sous lesquelles un système non-linéaire est équivalent, localement, à un système linéaire.

Dans un deuxième temps, nous nous concentrons sur les équations faisant intervenir un paramètre réel ; elles ont la forme générale $\partial u = f(u, \mu)$. A paramètre fixé, l'étude précédemment développée nous informe de la stabilité des points stationnaires ; nous sommes intéressés dans ce chapitre à comprendre l'évolution des solutions obtenues en fonction du paramètre. Nous étudions principalement le cas des dimensions 1 et 2 commençant par une étude du cas linéaire. Le comportement des solutions peut, dans ces cas, être décrit de manière détaillée. Le cas non-linéaire fait appel à un outil essentiel, le théorème des fonctions implicites, qui permet de décrire une courbe de points stationnaires comme graphe d'une fonction dépendant du paramètre et défini au voisinage d'une valeur de paramètre fixée. Mais en certains points, il se peut que ce théorème ne s'applique pas. Des problèmes de différents types peuvent alors survenir, entre autres : perte de stabilité d'un point stationnaire, ou pire, apparition ou disparition de certains d'entre eux quand le paramètre évolue. Cela est la base de la théorie des bifurcations. De tels comportements arrivent même lorsque la fonction f est à valeurs dans \mathbb{R} ; nous étudions en particulier trois phénomènes de bifurcation :

- apparition de deux points stationnaires lors de l'évolution du paramètre : *bifurcation selle-nœud*, section II.2.1.1 ;
- croisement de deux points stationnaires et échange de stabilité à ce croisement : *bifurcation transcritique*, section II.2.1.2 ;
- perte de stabilité d'un point stationnaire et apparition de deux nouveaux points stationnaires : *bifurcation en fourche*, section II.2.1.3.

Dans chacune des situations, nous décrivons explicitement la situation au voisinage d'une valeur de paramètre où l'on ne peut suivre un point stationnaire. Quand f est à valeurs dans \mathbb{R}^2 , d'autres comportements sont possibles ; nous énonçons un résultat supplémentaire de bifurcation, décrivant un comportement particulier dans ce cas : c'est la *bifurcation de Höpf* (section II.2.2).

L'étude théorique ci-avant développée donne des critères d'existence de points stationnaires au voisinage d'une certaine valeur du paramètre ; cependant, nous ne pouvons en général connaître en pratique la position d'un point stationnaire en fonction du paramètre. Nous pouvons déterminer l'existence, certes, ainsi qu'une localisation d'un tel point en fonction du paramètre, sans toutefois pouvoir donner

la valeur exacte. Dans le troisième chapitre, nous proposons des algorithmes qui permettent de déterminer ces valeurs ; plus précisément, ces algorithmes de continuation partent d'un point stationnaire connu pour une valeur de paramètre fixée et essaient de le faire évoluer en fonction du paramètre, de telle sorte que l'on décrive une courbe de points stationnaires. Dans les différents algorithmes, nous employons une même démarche : à partir d'un point stationnaire connu, nous commençons par approximer un nouveau point, plus loin sur la courbe ; nous affinons ensuite cette approximation de telle sorte que le nouveau point obtenu soit assez proche d'un point stationnaire. Nous présentons deux algorithmes similaires dont l'idée est de faire progresser le paramètre par petits pas. Un troisième algorithme, plus évolué, est également présenté ; il essaie de suivre la courbe décrite par les solutions constantes en ne supposant plus cette fois la courbe décrite comme fonction du paramètre. L'idée principale de continuation pour ce dernier algorithme est une approximation via le vecteur tangent à la courbe de points stationnaires.

Nous nous intéressons enfin à un problème physique dont l'étude peut être réalisée grâce à la théorie précédemment développée ; il s'agit du problème de l'elastica. Plus une courbe est pliée, plus l'énergie potentielle qu'elle contient est importante ; le problème consiste alors à trouver la forme des courbes planes de longueur fixée, qui se plient le moins (donc, qui minimisent l'énergie associée), sous certaines conditions au bord (on fixe les extrémités de la courbe ainsi que les vecteurs tangents en ces points). Or, une quantité relative à la courbe nous permet de caractériser ce plissage ; il s'agit de sa courbure, laquelle rend compte de la variation des vecteurs tangents en tout point. Dans ce chapitre, nous obtenons une formule explicite pour la courbure de telles courbes, ce qui nous permet, grâce aux conditions au bord, d'obtenir les courbes répondant au problème. Par ailleurs, on peut considérer le problème sous un autre angle : partant de la position initiale de la ligne droite, on rapproche les extrémités en faisant varier aussi les vecteurs tangents jusqu'à obtenir la situation qui nous intéresse. Cette vision permet d'employer les algorithmes de continuation : le paramètre est lié à la distance des points aux bords ainsi qu'à l'orientation des vecteurs tangents.

Nous terminons en mentionnant brièvement deux généralisations, d'une part les courbes dans l'espace minimisant leur énergie de courbure, d'autre part les films élastiques posés sur divers substrats, dont une modélisation (dans le cas où le substrat est de l'eau) revient à étudier des courbes planes minimisant une énergie liée à leur courbure et au déplacement du substrat en-dessous du film. Les solutions analytiques pour ce dernier problème sont, à ce jour et à notre connaissance, inconnues. De plus, la question d'une modélisation de ce problème dans un cadre plus général (où le substrat sous le film est un gel), est lui aussi ouvert. Le service de P. Damman travaille actuellement sur le sujet afin de proposer diverses possibilités. La validation de ces modèles peut être traitée avec les algorithmes de continuation.

Chapitre I

Stabilité

[Faire une intro disant ce que tu vas faire (motivation et plan), de quel bouquin/article tu t'inspires et éventuellement pointer ta contribution personnelle. Ceci est vrai pour tous les chapitres.]

Dans de nombreuses applications physiques, on rencontre des équations différentielles dépendant de paramètres ; faire varier ces paramètres donnent alors des configurations différentes. Une question naturelle qui se pose est de savoir si le comportement des solutions à cette équation, pour des paramètres proches, sont similaires. Mais avant de pouvoir parler de « variation de comportement », il faut déterminer ce que l'on entend par le comportement de solutions ; ce premier chapitre répond à cette attente. Nous attacherons une importance particulière aux solutions constantes et aux solutions qui, une fois suffisamment proches d'une telle solution, restent toujours proches de cette solution ; cela donne lieu aux définitions de stabilité, par lesquelles nous débutons.

Ces définitions sont données dans un cadre général, mais nous nous restreignons dans la suite aux équations autonomes, plus faciles à manipuler. Nous étudions en détail la stabilité dans le cas d'équations linéaires ; de nombreux calculs relatifs à cette étude sont développés dans l'annexe A. Cette étude réalisée, nous permet d'étudier le cas général des équations autonomes ; nous remarquons que le comportement de solutions de telles équations au voisinage d'un point stationnaire peut être appréhendé par l'étude d'une équation linéaire.

Pour ce chapitre, nous nous inspirons principalement des ouvrages de M. Demazure [5] et de P. Glendinning [6]. La preuve du théorème de la variété stable est inspirée du livre de L. Perko [11].

I.1 Définitions générales, et le cas des équations autonomes

On considère l'équation différentielle

$$\partial_t u = f(t, u) \quad (\text{I.1})$$

où $f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ est une fonction (continue), et où l'on recherche des fonctions $u : \mathbb{R} \rightarrow \mathbb{R}^n$. Il est connu que, pour f localement Lipschitz, le problème de Cauchy :

$$\begin{cases} \partial_t u = f(t, u) \\ u(t_0) = u_0 \end{cases} \quad (\text{I.2})$$

admet une solution (maximale) unique, pour $t_0 \in \mathbb{R}$ et $u_0 \in \mathbb{R}^n$ fixés ; ce résultat, dû à Cauchy et Lipschitz, est présenté par exemple, dans le livre de J. Mahwin et N. Rouche [10]. Dans ce qui suit, les fonctions constantes solutions de cette équation joueront un rôle important.

Définition 1. On dit d'un point $u_0 \in \mathbb{R}^n$ qu'il est *point stationnaire de l'équation (I.1)* si la solution $u : \mathbb{R} \rightarrow \mathbb{R}^n$ de (I.2) (avec $t_0 \in \mathbb{R}$ quelconque) est constante ; en d'autres termes, u_0 est point stationnaire si $f(t, u_0) = 0$ pour tout $t \in \mathbb{R}$.

Supposons que $u_0 \in \mathbb{R}^n$ soit un point stationnaire ; une question naturelle est la suivante : si l'on démarre d'une condition initiale v_0 proche de u_0 en t_0 , et si v est la solution au problème de Cauchy associé, quel est le comportement de v par rapport à la fonction constante envoyant tout réel sur u_0 ? Cela conduit aux deux définitions suivantes, données dans le cas général :

Définition 2. Soit u une solution de (I.1), définie sur un intervalle ouvert $I \subset \mathbb{R}$. On dit qu'elle est *stable au sens de Lyapounov* si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que : pour toute solution v de (I.1), définie sur un intervalle J (non-disjoint de I),

$$\begin{aligned} &\text{si } \|u(t_0) - v(t_0)\| < \delta \text{ pour un certain } t_0 \in I \cap J, \\ &\text{alors } \forall t \in I \cap J, t \geq t_0, \|u(t) - v(t)\| < \varepsilon. \end{aligned}$$

Définition 3. Soit u une solution de (I.1), définie sur un intervalle ouvert $I \subset \mathbb{R}$ non-borné supérieurement. On dit qu'elle est *quasi-asymptotiquement stable* s'il existe $\delta > 0$ tel que : pour toute solution v de (I.1), définie sur un intervalle $[t_0, +\infty[\subset I$,

$$\text{si } \|u(t_0) - v(t_0)\| < \delta, \text{ alors } \|u(t) - v(t)\| \xrightarrow[t \rightarrow +\infty]{} 0.$$

Leurs idées respectives sont les suivantes. La première définition de stabilité dit que toute solution, proche de u en un certain point t_0 , restera proche au-delà de ce point. La seconde donne plutôt un résultat « à l'infini » ; elle dit que toute solution, proche de u en un certain point t_0 , est aussi proche de u qu'on le souhaite, pour des t assez grands (et peut-être bien plus grands que t_0). En quelque sorte, la stabilité au sens de Lyapounov nous assure un contrôle tout le long de la trajectoire, mais sans pouvoir contrôler la proximité des fonctions ; à l'opposé de la stabilité quasi-asymptotique, qui permet un certain contrôle de la proximité, mais seulement « assez loin ».

Remarquons qu'il s'agit là de deux définitions distinctes de stabilité ; en particulier l'une n'implique pas l'autre en général. En effet, un exemple simple, l'équation différentielle $\partial_t u = 0$, montre que les solutions (les fonctions constantes) sont toutes stables au sens de Lyapounov, mais en aucun cas quasi-asymptotiquement stables. Pour l'autre sens, prenons pour exemple l'équation différentielle

$$\partial^2 u = \left(\frac{2}{t^2} - 1 \right) u;$$

la fonction identiquement nulle (définie sur $\mathbb{R}^+ \setminus \{0\}$) est clairement point stationnaire de cette équation ; on montre par ailleurs que toutes les autres solutions s'écrivent sous la forme $u(t) = \frac{1}{t}(a \sin t + b \cos t)$, pour $a, b \in \mathbb{R}$. De telles solutions convergent vers le point stationnaire 0 lorsque t tend vers l'infini ; 0 se révèle donc être quasi-asymptotiquement stable. Par contre, on montre qu'il n'est pas stable au sens de Lyapounov. En prenant la négation de la définition, fixons $\varepsilon = \frac{1}{10}$. Soit $\delta > 0$; considérons alors la fonction $u(t) = \frac{1}{t} \sin t$. C'est une solution de l'équation différentielle précitée, qui s'annule en $t_0 = \pi$; et donc $|u(t_0)| < \delta$. Cependant, il existe des réels t plus grands que π , pour lesquels $|u(t)|$ est plus grand que $\varepsilon = \frac{1}{10}$. Par exemple, pour $t = \frac{3}{2}\pi$, on a $u(t) = -\frac{2}{3\pi}$, donc $|u(t)| = \frac{2}{3\pi} > \frac{2}{20} = \varepsilon$.

Nous allons donc dans la suite, considérer les deux types de stabilité précitées à la fois ; nous définissons :

Définition 4. Une solution u de (I.1), définie sur un intervalle ouvert $I \subset \mathbb{R}$ non borné supérieurement, est dite *asymptotiquement stable* si elle est à la fois stable au sens de Lyapounov, et quasi-asymptotiquement stable.

On peut par ailleurs regarder ce qui se passe dans le passé ; c'est-à-dire, connaissant ce qui se passe pour une certaine valeur t_0 , pouvoir caractériser ce qui se passe pour des $t < t_0$. Remarquons toutefois que, si u est solution du problème de Cauchy (I.2), alors la fonction v qui, à t associe $v(t) := u(-t)$, est solution du problème

$$\begin{cases} \partial_t v = f(-t, v) \\ v(-t_0) = u_0 \end{cases} \quad (\text{I.3})$$

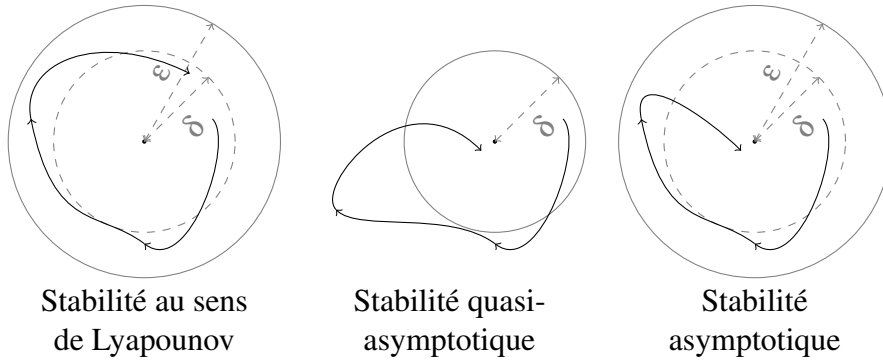


FIG. I.1 – Illustrations des différentes notions de stabilité. Les boules dessinées sont centrées au point 0 ; les courbes représentent la différence des fonctions en question, c'est-à-dire $u(t) - v(t)$.

et l'on peut discuter de la « stabilité inversée » de u (on parlera d'*instabilité*¹) en regardant la stabilité de v . Formellement, on a

Définition 5. On dit d'une solution u de (I.1), définie sur un intervalle $I =]a, b[$ (avec $-\infty \leq a < b \leq +\infty$) qu'elle est *instable au sens de Lyapounov* (resp. *quasi-asymptotiquement instable*, en supposant $a = -\infty$) si la fonction $t \mapsto u(-t)$ (définie sur $] -b, -a[$) est stable au sens de Lyapounov (resp. quasi-asymptotiquement stable). Cela est équivalent à demander que, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :

$$\begin{aligned} &\text{pour toute solution } v \text{ de (I.1), si } \|u(t_0) - v(t_0)\| < \delta \text{ pour un certain } t_0 \in I, \\ &\text{alors } \forall t \leq t_0, \|u(t) - v(t)\| < \varepsilon \end{aligned}$$

(resp. il existe $\delta > 0$ tel que :

$$\begin{aligned} &\text{pour toute solution } v \text{ de (I.1), si } \|u(t_0) - v(t_0)\| < \delta \text{ pour un certain } t_0 \in I, \\ &\text{alors } \|u(t) - v(t)\| \xrightarrow[t \rightarrow -\infty]{} 0). \end{aligned}$$

Si $a = -\infty$, on dira que u est *asymptotiquement instable* si elle est à la fois instable au sens de Lyapounov et quasi-asymptotiquement instable.

Concentrons-nous maintenant sur les équations différentielles autonomes. Rappelons qu'il s'agit des équations différentielles où la fonction f en jeu ne

¹Il faut être très prudent avec ce terme : dans la littérature, certains auteurs parlent de solution instable pour signifier qu'elle n'est pas stable ; mais n'ont pas de terme spécifique pour la « stabilité passée » ou « inversée ». D'autres en revanche adoptent notre convention (par exemple, M. Demazure [5]), et parlent (comme nous le ferons) de solution non-stable. Il est donc important de voir quelle définition d'« instabilité » est utilisée dans tout document.

fait pas intervenir le temps t ; elles sont donc de la forme

$$\partial_t u = f(u). \quad (\text{I.4})$$

Dans ce cas, on a la proposition suivante :

Proposition 1. *Si u est une solution de (I.4), et si v est définie par :*

$$v(t) := u(t - t_0)$$

pour $t_0 \in \mathbb{R}$ fixé, alors v est aussi solution de (I.4). De plus, les notions de stabilité pour u sont conservées par v .

Dès lors, les notions de point stationnaire et de stabilité sont simplifiées. En effet, les points stationnaires correspondent alors aux racines de f ; et, concernant la stabilité, on peut remplacer les t_0 arbitraires par 0.

Grâce à cette proposition, on peut voir les solutions de ce type d'équations sous un autre angle. Au lieu de fixer u_0 , considérer la solution u provenant du problème de Cauchy, et évaluer son image en t , on peut fixer les paramètres globalement, et donc considérer la fonction

$$\begin{aligned} \Phi : \mathbb{R}^n \times \mathbb{R} &\rightarrow \mathbb{R}^n \\ (u_0, t) &\mapsto u(t) \end{aligned} \quad (\text{I.5})$$

où u est l'unique solution maximale au problème de Cauchy (I.2), pour lequel $u(0) = u_0$. Notons toutefois que Φ peut ne pas être définie partout ; toutefois, elle jouit de propriétés intéressantes dans le cadre des équations autonomes :

Proposition 2. *La fonction Φ définie grâce à l'égalité (I.5) vérifie les faits suivants :*

1. $\text{Dom } f \times \{0\} \subset \text{Dom } \Phi$, et pour tout $x \in \text{Dom } f$, $\Phi(x, 0) = x$;
2. pour tout $x \in \mathbb{R}^n$ et tous $s, s' \in \mathbb{R}$, si $(x, s + s') \in \text{Dom } \Phi$, alors $(x, s) \in \text{Dom } \Phi$, le couple $(\Phi(x, s), s')$ se trouve aussi dans $\text{Dom } \Phi$ et on a l'égalité $\Phi(x, s + s') = \Phi(\Phi(x, s), s')$;
3. $\Phi(\cdot, s)$ est « injective » dans le sens suivant : pour $s \in \mathbb{R}$ fixé, si y et z sont deux points tels que $(y, s) \in \text{Dom } \Phi$, $(z, s) \in \text{Dom } \Phi$ et pour lesquels $\Phi(y, s) = \Phi(z, s)$, alors $y = z$.
4. Enfin, pour tous $x, y \in \mathbb{R}^n$ et tout $s \in \mathbb{R}$, si $(y, s) \in \text{Dom } \Phi$ et $\Phi(y, s) = x$, alors $(x, -s) \in \text{Dom } \Phi$, et $y = \Phi(x, -s)$:

Démonstration. Il s'agit en réalité d'une réécriture de faits concernant les équations autonomes. Le premier point est évident. Pour prouver le deuxième point, posons $u := \Phi(x, \cdot)$; c'est une solution de l'équation (I.4), et on peut considérer la

fonction v définie par $v(t) := u(t + s)$, également solution au vu de la proposition précédente. On obtient, d'une part, $v(s') = \Phi(y, s')$ avec $y = v(0) = u(s)$, mais par définition de u :

$$v(s') = \Phi(\Phi(x, s), s');$$

par ailleurs, en utilisant la définition de v et de u , on a :

$$v(s') = u(s + s') = \Phi(x, s + s')$$

ce qui donne l'égalité attendue. Le troisième point peut s'exprimer de la manière suivante : si u et v sont deux solutions de (I.4), dont les conditions initiales sont respectivement $u(0) = y$ et $v(0) = z$, et si l'on a $u(s) = v(s)$, alors $u(0) = v(0)$. Cela se déduit par unicité de la solution au problème de Cauchy (I.2). Enfin, le dernier point s'exprime en termes de solutions de l'équation (I.4), par le fait que si u est solution et vérifie $u(0) = y$ et $u(s) = x$, alors on a une solution v qui vérifie $v(0) = x$ et $v(-s) = y$; c'est une simple application de la proposition 1. \square

De façon générale, on définit :

Définition 6. Un *flux* sur \mathbb{R}^n est une fonction $\Phi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ continue en les deux variables et vérifiant les propriétés 1, 2 et 4 de la proposition précédente. Le *flux associé à l'équation différentielle (I.4)* est la fonction Φ définie en (I.5).

Grâce à ces définitions, on peut exprimer les notions de stabilité plus simplement :

Proposition 3. Soit u une solution de l'équation (I.4), définie sur un intervalle $I =]a, b[$ (avec $-\infty \leq a < b \leq +\infty$). On a équivalence entre « u est stable au sens de Lyapounov » et le fait suivant :

$$\begin{aligned} & \text{pour tout } \varepsilon > 0, \text{ il existe } \delta > 0 \text{ tel que,} \\ & \forall x \in B(u(0), \delta), \forall t \geq 0, \|u(t) - \Phi(x, t)\| < \varepsilon \end{aligned}$$

De même, en fixant $b = +\infty$, on a équivalence entre « stabilité quasi-asymptotique de u » et ce qui suit :

$$\begin{aligned} & \text{il existe } \delta > 0 \text{ tel que,} \\ & \forall x \in B(u(0), \delta), \|u(t) - \Phi(x, t)\| \xrightarrow{t \rightarrow -\infty} 0. \end{aligned}$$

I.2 Stabilité linéaire

[Il doit être clair d'après l'intro du chapitre le rôle que ceci va jouer.] On s'intéresse maintenant aux équations différentielles linéaires ; i.e. celles de la forme

$$\partial_t u = Au \tag{I.6}$$

où A est une matrice carrée d'ordre n (à coefficients réels).

I.2.1 Solutions générales de l'équation

I.2.1.1 Formes des solutions

En dimension 1, les solutions sont évidemment les exponentielles d'exposant At à constante multiplicative près, i.e. les fonctions de la forme $u(t) = e^{At} u_0$ avec $u_0 \in \mathbb{R}$. Une idée pour obtenir toutes les solutions de (I.6) est de définir une « exponentielle matricielle » ; cela peut être fait grâce aux séries. On définit pour une matrice A carrée d'ordre n ,

$$e^{At} = \sum_{n=0}^{+\infty} \frac{(At)^n}{n!} \quad (\text{I.7})$$

Cela est bien défini. En effet, si l'on considère la norme matricielle comme celle associée aux applications linéaires², on obtient que

$$\left\| \sum_{i=0}^n \frac{(At)^i}{i!} \right\| \leq \sum_{i=0}^n \frac{\|(At)^i\|}{i!} \leq \sum_{i=0}^n \frac{|t|^i \|A\|^i}{i!}$$

et donc, que

$$\|e^{At}\| \leq e^{\|A\||t|}. \quad (\text{I.8})$$

Pour tout t , la série définie par (I.7) est donc majorée, en norme, par un réel ; cela implique sa convergence.

Grâce à cette définition, on obtient le théorème suivant, similaire à ce qui se passe en dimension 1 :

Théorème 4. Soit $A \in \mathbb{R}^{n \times n}$ et $u_0 \in \mathbb{R}^n$. Le problème de Cauchy

$$\begin{cases} \partial_t u = Au \\ u(0) = u_0 \end{cases} \quad (\text{I.9})$$

admet une unique solution définie sur tout \mathbb{R} ; celle-ci est donnée par $u(t) = e^{At} u_0$.

Démonstration. La fonction donnée est bien solution du problème de Cauchy : en effet, on a $\partial_t \text{Id} = 0$ et, pour tout $n \geq 1$,

$$\partial_t \left(\frac{(At)^n}{n!} \right) = \frac{1}{n!} A^n n t^{n-1} = A \frac{A^{n-1} t^{n-1}}{(n-1)!}$$

²Pour rappel, elle est définie par $\|A\| := \sup_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|}$.

d'où l'on tire, pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\partial_t \left(\sum_{i=0}^n \frac{(At)^i}{i!} \right) &= \sum_{i=0}^n \partial_t \left(\frac{(At)^i}{i!} \right) && \text{(par linéarité de la différentielle)} \\
&= \sum_{i=1}^n A \frac{A^{i-1} t^{i-1}}{(i-1)!} && \text{(grâce aux égalités ci-dessus - - noter que le terme pour } i=0 \text{ est nul)} \\
&= A \sum_{i=0}^{n-1} \frac{A^i t^i}{i!} && \text{(par linéarité et changement d'indice)}
\end{aligned}
\tag{I.10}$$

et en passant à la limite pour $n \rightarrow +\infty$:

$$\lim_{n \rightarrow +\infty} \partial_t \left(\sum_{i=0}^n \frac{(At)^i}{i!} \right) = A \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} = A e^{At}.$$

On prouve alors que $\partial_t \left(\sum_{i=0}^{\infty} \frac{(At)^i}{i!} \right)$ existe et vaut la limite ci-dessus. Mais on remarque que la limite ci-dessus est uniforme en t sur tout intervalle borné, grâce à l'inégalité (I.8) ; par conséquent nous venons de prouver que e^{At} est dérivable et est solution de (I.6).

Prouvons maintenant l'unicité de la solution. Soit $x(t)$ une solution du problème de Cauchy ci-dessus. Posons $y(t) := e^{-At} x(t)$. On obtient alors

$$\begin{aligned}
\partial_t y(t) &= \partial_t (e^{-At}) x(t) + e^{-At} \partial_t x(t) \\
&= -A e^{-At} x(t) + e^{-At} (Ax(t)) && \text{par définition de } x \text{ et par ci-dessus} \\
&= 0 && \text{car } A \text{ et } e^{-At} \text{ commutent}
\end{aligned}$$

donc y est une fonction constante égale à un certain c . Par conséquent, quel que soit $t \in \mathbb{R}$, $c = e^{-At} x(t)$ i.e. $x(t) = c e^{At}$, et l'on obtient $c = u_0$ en considérant $x(0)$. \square

I.2.1.2 Calcul des solutions

Afin de déterminer les conditions sur A pour lesquelles nous avons des points stationnaires, leur nombre, puis l'étude de leur stabilité, nous devons déterminer comment se comporte la solution $u(t) = e^{At} u_0$ du problème de Cauchy (I.9) selon la forme de A .

L'annexe A prouve en détail les deux théorèmes suivants. Le premier donne un résultat d'invariance des solutions obtenues :

Théorème 5. Soit F un sous-espace vectoriel de \mathbb{R}^n s'écrivant comme somme directe d'espaces propres généralisés de A . Toute solution du problème de Cauchy dont la condition initiale est dans F , a son image contenue dans F .

Intuitivement, F s'écrit comme une somme directe de F_i ; si bien que tout vecteur u_0 s'écrit de manière unique comme somme de u_i , avec $u_i \in F_i$ pour tout i . La solution du problème de Cauchy (I.9) de condition initiale u_0 est alors la somme des solutions du même problème pour les conditions initiales u_i . Or, on montre que chacune de ces solutions reste dans son espace F_i ; les détails de ce fait sont expliqués dans l'annexe A, en section A.2. Dès lors, leur somme reste dans l'espace somme, qui n'est autre que F .

Le second théorème exprime une équivalence entre la convergence des solutions vers 0 et une condition sur les valeurs propres; plus précisément, on peut prouver :

Théorème 6. Supposons que \mathbb{R}^n soit décomposé en somme directe :

$$\mathbb{R}^n = \bigoplus_{i=1,\dots,k} F_i \oplus \bigoplus_{j=1,\dots,m} G_j,$$

où les espaces F_i et G_j vérifient les propriétés suivantes :

- chaque F_i est espace propre réel généralisé de A de valeur propre λ_i ;
- chaque G_j est engendré par des vecteurs x_l et y_l de telle sorte que tous les $x_l + \mathbf{i}y_l$ soient vecteurs propres complexes généralisés de A , pour une même valeur propre z_j ;
- pour tous i et j compris entre 1 et k , si $i \neq j$, alors $\lambda_i \neq \lambda_j$;
- pour tous i et j compris entre 1 et m , si $i \neq j$, alors $z_i \neq z_j$.

Soient $I \subset \{1, \dots, k\}$ et $J \subset \{1, \dots, m\}$; posons alors

$$F = \bigoplus_{i \in I} F_i \oplus \bigoplus_{j \in J} G_j,$$

Si tous les λ_i , pour $i \in I$, sont strictement négatifs, ainsi que toutes les parties réelles des z_j , pour $j \in J$, alors toutes les solutions du problème de Cauchy dont les conditions initiales sont dans F , convergent vers 0 quand $t \rightarrow +\infty$. De même, si tous les λ_i sont strictement positifs, ainsi que toutes les parties réelles des z_i , alors toutes les solutions du problème de Cauchy dont les conditions initiales sont dans F , convergent vers 0 quand $t \rightarrow -\infty$.

Réciproquement, soit $u_0 \in F$; si la solution du problème de Cauchy converge vers 0 lorsque $t \rightarrow +\infty$, alors tous les λ_i et toutes les parties réelles des z_i sont strictement négatives; de même si la solution tend vers 0 pour $t \rightarrow -\infty$ alors tous les λ_i et toutes les parties réelles des z_i sont strictement positives.

De façon très intuitive, on décompose u_0 en somme de vecteurs u_i et v_j , chacun de ces vecteurs appartenant respectivement à F_i et G_j . La solution u associée à u_0 est alors la somme des solutions du problème de Cauchy, de conditions initiales u_i et v_j . On montre alors que chacune de ces solutions fait apparaître un facteur $e^{\lambda_i t}$ ou $e^{\Re z_j t}$ respectivement ; les autres facteurs étant polynômiaux en t . Dès lors, le comportement asymptotique de u est déterminé par ces exponentielles, et les comportements de ces dernières sont données par les signes des λ_i et $\Re z_j$. Dès lors, nous obtenons les critères annoncés dans le théorème. Tous les détails de ce qui précède, se trouve également en section A.2 de l'annexe A.

I.2.2 Points stationnaires et étude de stabilité

Les points stationnaires de l'équation (I.6) sont simples à obtenir : il s'agit des fonctions constantes, qui annulent la matrice A . On obtient facilement que, si A est inversible, alors seule la fonction constante nulle est point stationnaire. Dans le cas contraire, on remarque que 0 est valeur propre de A , et l'espace propre associé à cette valeur propre est exactement l'ensemble des points stationnaires ; c'est aussi le noyau de A .

I.2.2.1 Cas des matrices inversibles

Nous étudions donc, lorsque A est inversible, la stabilité asymptotique de 0 comme unique point stationnaire. Soit v la solution au problème de Cauchy (I.9) où l'on pose à la place de $u_0 = 0$, un vecteur v_0 non-nul. Nous souhaitons voir si $v(t) \rightarrow 0$ quand $t \rightarrow +\infty$. En appliquant le théorème précédent, puisque v_0 est quelconque, nous avons cette condition si et seulement si toutes les valeurs propres de A ont des parties réelles strictement négatives. De la même façon, nous aurons $v(t) \rightarrow 0$ quand $t \rightarrow -\infty$ si et seulement si toutes les valeurs propres de A ont des parties réelles strictement positives. Cela garantit la stabilité ou l'instabilité quasi-asymptotiques, selon le cas considéré. Pour obtenir la stabilité (ou l'instabilité) asymptotique, nous devons montrer qu'à condition initiale v_0 suffisamment proche de 0, la solution reste aussi proche de 0 qu'on le souhaite. Nous considérons simplement le cas de la stabilité, l'instabilité s'obtenant de manière similaire.

I.2.2.2 Cas général

Supposons à présent que A ne soit plus inversible. Son noyau n'est alors pas trivial, et on note V un espace complémentaire à celui-ci (de telle sorte que $\mathbb{R}^n = V \oplus \ker A$). Soit v_1 un vecteur quelconque de \mathbb{R}^n ; on décompose alors $v_1 = v_0 + x$ avec $v_0 \in V$ et $x \in \ker A$. Si \bar{v} et v désignent respectivement les solutions aux

problèmes de Cauchy (I.9) de conditions initiales respectives v_1 et v_0 , alors, pour tout $t \in \mathbb{R}$:

$$\begin{aligned}\bar{v}(t) &= e^{At} v_1 \\ &= e^{At} (v_0 + x) \\ &= e^{At} v_0 + x && \text{car } x \in \ker A \\ &= v(t) + x\end{aligned}$$

si bien qu'étudier la stabilité de x revient à étudier la stabilité de 0. Mais, on ne peut avoir stabilité ni instabilité asymptotiques : en effet, deux points stationnaires, aussi proches soient-ils, resteront toujours à même distance ; d'où, la stabilité (et l'instabilité) quasi-asymptotiques sont mis en défaut. Par contre, la stabilité et l'instabilité au sens de Lyapounov peuvent être étudiés ; et comme annoncé, l'étude de cette stabilité en 0 suffit. En considérant \tilde{A} comme A , mais restreinte sur l'espace V , on remarque que \tilde{A} est inversible ; dès lors, nous pouvons étudier la stabilité de l'unique point stationnaire $0 \in V$ pour \tilde{A} comme précédemment, puis ramener les résultats obtenus dans l'espace $\mathbb{R}^n = V \oplus \ker A$ pour connaître l'éventuelle stabilité ou instabilité au sens de Lyapounov.

En conclusion, on obtient le résultat suivant :

Proposition 7 (Caractérisation de la stabilité asymptotique linéaire). *Soit $A \in \mathbb{R}^{n \times n}$. Si les valeurs propres de A sont toutes à partie réelle strictement négative, alors le point 0 est le seul point stationnaire, et il est asymptotiquement stable. Réciproquement si 0 est point stationnaire unique du système $\partial u = Au$, et s'il est asymptotiquement stable, alors A est inversible et toutes ses valeurs propres ont une partie réelle strictement négative.*

De même on a équivalence entre les deux points suivants :

1. *0 est le seul point stationnaire du système $\partial u = Au$, et il est asymptotiquement instable ;*
2. *A est inversible et n'admet que des valeurs propres à partie réelle strictement positive.*

On obtient un critère semblable pour la stabilité au sens de Lyapounov :

Proposition 8 (Caractérisation de la stabilité linéaire au sens de Lyapounov). *Soit $A \in \mathbb{R}^{n \times n}$. Si toutes les valeurs propres de A ont une partie réelle négative (ou nulle), alors tous les points se trouvant dans $\ker A$ sont des points stationnaires stables au sens de Lyapounov, et ce sont les seuls.*

Réciproquement si tous les points stationnaires du système $\partial u = Au$, se trouvant dans $\ker A$, sont stables au sens de Lyapounov, alors A n'a que des valeurs propres dont la partie réelle est négative ou nulle.

On a de même la proposition analogue caractérisant l'instabilité au sens de Liapounov à partir de valeurs propres positives ou nulles.

I.2.3 Espaces stable, instable, central

Pour $A \in \mathbb{R}^{n \times n}$, on peut définir les espaces suivants :

Définition 7. L'espace stable associé à A au point stationnaire 0 est l'espace engendré par les vecteurs propres généralisés de A de valeur propre à partie réelle négative ; on le note $E^s(0)$.

L'espace instable associé à A au point stationnaire 0 est l'espace engendré par les vecteurs propres généralisés de A de valeur propre à partie réelle positive ; on le note $E^u(0)$.

L'espace central associé à A au point stationnaire 0 est l'espace engendré par les vecteurs propres généralisés de A de valeur propre à partie réelle nulle ; on le note $E^c(0)$.

On obtient alors une décomposition de \mathbb{R}^n en somme directe :

$$\mathbb{R}^n = E^s(0) \oplus E^u(0) \oplus E^c(0)$$

et le théorème :

Théorème 9. Si $E^c(0) = \{0\}$, alors

$$E^s(0) = \{x \in \mathbb{R}^n \mid e^{At} x \xrightarrow[t \rightarrow +\infty]{} 0\}$$

$$E^u(0) = \{x \in \mathbb{R}^n \mid e^{At} x \xrightarrow[t \rightarrow -\infty]{} 0\}$$

Démonstration. La condition du théorème implique que tout vecteur v de \mathbb{R}^n se décompose de façon unique en une somme $x + y$, avec $x \in E^s(0)$ et $y \in E^u(0)$. La proposition précédente nous permet de dire que $e^{At} x \rightarrow 0$ lorsque $t \rightarrow +\infty$, alors que $e^{At} y \rightarrow 0$ lorsque $t \rightarrow -\infty$; par ailleurs $\|e^{At} x\| \geq \|x\|$ pour $t \leq 0$, tandis que $\|e^{At} y\| \geq \|y\|$ pour $t \geq 0$. Dès lors, la convergence de $e^{At} v$ vers 0 lorsque $t \rightarrow +\infty$ est possible si et seulement si $y = 0$, auquel cas, $v \in E^s(0)$; on obtient la deuxième conclusion de façon similaire. \square

I.3 Stabilité générale — linéarisation

Notre but ici est de trouver les points stationnaires de l'équation générale (I.4) :

$$\partial_t u = f(u)$$

avec $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ qu'on suppose de classe C^∞ , puis d'en étudier la stabilité. Pour trouver les points stationnaires, il suffit de chercher les racines de f . Une fois ces

points obtenus, on peut étudier leur stabilité. Or, on remarque que la stabilité d'un point stationnaire est une propriété locale, puisque l'on considère, pour l'étudier, des points proches comme conditions initiales. Dès lors, dans la suite de cette section, et sans perte de généralité, on suppose que 0 est un point stationnaire de (I.4), quitte à considérer une translation convenable.

I.3.1 Théorème du rang constant

Une idée pour déterminer la stabilité de ce point est d'utiliser un développement limité pour f :

$$f(u) = f(0) + \partial f(0) \cdot u + {}^t u \cdot g(u) \cdot u \quad \text{avec } g(0) = \partial^2 f(0), \quad (\text{I.11})$$

d'où on voit que, si l'on souhaite se ramener à un système linéaire (dont on vient d'étudier la stabilité de points stationnaires), on voudrait prendre pour matrice $A := \partial f(0)$. Or, en général, cela pose problème ; en effet, supposons que $\partial f(0) = 0$. Le système linéaire n'admet alors comme solutions que les fonctions constantes ; chaque condition initiale est alors un point stationnaire. Or ce n'est peut-être pas le cas pour le système initial, à cause des ordres supérieurs. Dans ce cas, la fonction f n'est pas constante dans un voisinage de 0, ce qui est équivalent à ce que sa jacobienne ∂f n'est pas identiquement nulle lorsqu'elle est évaluée en des points non-nuls au voisinage de 0. Prenons pour exemple l'équation différentielle $\partial u = 1 + u^2$. On peut montrer que $u(t) = \tan t$ est une solution non-constante de cette équation ; or, le système linéarisé donnerait, au voisinage de 0, $\partial u = 0$, puisque la dérivée de $1 + u^2$ s'annule en 0.

De manière générale, le rang de la matrice jacobienne au voisinage d'un point (ici 0), est toujours supérieur (ou égal) au rang de la jacobienne évaluée en 0. En effet, on montre que le rang d'une matrice vaut au moins k si et seulement s'il existe un mineur d'ordre k dans cette matrice qui ne s'annule pas ; connaissant cette équivalence, et le fait que $\text{GL}_k(\mathbb{R})$, l'ensemble des matrices inversibles, est un ouvert dans $\mathbb{R}^{k \times k}$, on déduit que si le mineur considéré, évalué en 0, est non-nul, alors il le reste sur un voisinage de 0. Le théorème qui suit donne une première réponse à la possibilité de pouvoir se ramener à une équation différentielle linéaire, en conservant les mêmes points stationnaires :

Théorème 10 (Théorème du rang constant). *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction de classe C^∞ ; soit U un voisinage de 0 sur lequel le rang de la jacobienne ∂f est supposé constant. Il existe deux fonctions $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et $\psi : \mathbb{R}^m \rightarrow \mathbb{R}^m$, toutes deux C^∞ sur des voisinages V et W de 0, inclus respectivement à U et \mathbb{R}^m , et telles que :*

$$\text{pour tout } x \in V, f \circ \phi(x) = \psi(\partial f(0) \cdot x).$$

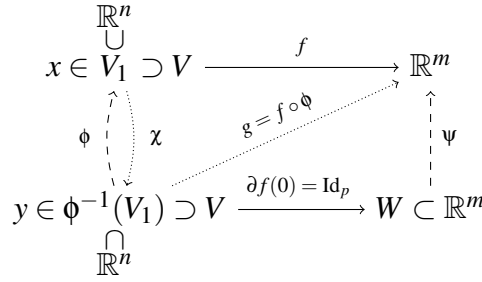


FIG. I.2 – Construction des fonctions entrant en jeu dans le théorème 10.

Démonstration. On pose p le rang de $\partial f(0)$; par hypothèse, il est constant sur U . Sans perte de généralité, on peut supposer que $\partial f(0)$ est l'identité sur les p premiers vecteurs de base et identiquement nulle sur le sous-espace complémentaire (les matrices de passages éventuelles n'affectant en rien le comportement C^∞ de ϕ et ψ). De plus, on peut supposer que $f(0) = 0$, quitte à effectuer une translation des images. Pour des facilités de notation, on écrira un vecteur x de \mathbb{R}^n sous la forme (\bar{x}, \tilde{x}) avec $\bar{x} \in \mathbb{R}^p$, l'espace orthogonal à $\ker \partial f(0)$.

Définissons $\chi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de la façon suivante :

$$\chi(x) := (f_1(x), \dots, f_p(x), \tilde{x})$$

Alors, $\partial \chi(0) = \text{Id}_n$; en particulier, c'est une matrice inversible. Pour tout x dans un voisinage V_1 de 0 (dans \mathbb{R}^n), la matrice $\partial \chi(x)$ a donc un déterminant non-nul ; cela implique que χ est un difféomorphisme local en 0 . On pose alors ϕ la fonction réciproque de χ . Puisque $f(0) = 0$, on a $\chi(0) = 0$, et $\phi : \phi^{-1}(V_1) \rightarrow V_1$, envoie donc 0 sur 0 .

Soit maintenant $g : \phi^{-1}(V_1) \rightarrow \mathbb{R}^n$ la fonction définie par

$$\begin{aligned} g(y) &:= f(\phi(y)) \\ &= (\bar{y}, f_{p+1}(\phi(y)), \dots, f_m(\phi(y))) \end{aligned}$$

Cette fonction est différentiable sur son domaine, et admet pour différentielle

$$\partial g(y) = \partial f|_{\phi(y)} \cdot \partial \phi(y)$$

Pour $y \in \phi^{-1}(V_1)$, la matrice $\partial \phi(y)$ est inversible, et sur $U \cap V_1$, $\partial f|_{\phi}$ est de rang constant p ; donc, sur un voisinage V de 0 (contenu dans $U \cap V_1 \cap \phi^{-1}(V_1) \subset \mathbb{R}^n$), ∂g est de rang constant p ; or vu l'expression de g , on a, pour $x = \phi(y)$:

$$\partial g(y) = \begin{pmatrix} \text{Id}_p & 0 \\ \partial_{\bar{y}} f_{p+1}(x) & \partial_{\bar{y}} f_{p+1}(x) \\ \vdots & \vdots \\ \partial_{\bar{y}} f_m(x) & \partial_{\bar{y}} f_m(x) \end{pmatrix}$$

si bien que les différentielles $\partial_{\bar{y}} f_i$, pour $i = p + 1, \dots, m$, sont nécessairement nulles (sans quoi, le rang serait strictement supérieur à p). Toutes les fonctions f_i sont constantes vis-à-vis de $\bar{y} = \bar{x}$ et s'écrivent donc $f_i(\bar{x})$.

On pose alors $\psi(\bar{y}, \tilde{y}) := (\bar{y}, \tilde{f}(\phi(\bar{y})) + \tilde{y})$, où $\tilde{f} = (f_{p+1}, \dots, f_m)$. Ainsi définie, sur $W := \partial f(0)(V) \times \mathbb{R}^{m-p}$, ψ est clairement C^∞ (on remarque que $\partial\psi$ peut être exprimée comme somme de Id_m avec une matrice composée des dérivées partielles selon \bar{x} des f_i pour $i = p + 1, \dots, m$); et on a

$$\begin{aligned} f(\phi(y)) &= g(\bar{y}, \tilde{y}) \\ &= (\bar{y}, f_{p+1}(\phi(\bar{y})), \dots, f_m(\phi(\bar{y}))) \\ &= \psi(\bar{y}) \\ &= \psi(\partial f(0) \cdot y) \end{aligned} \quad \square$$

Une première condition à la possibilité de linéariser (en 0) l'équation (I.4), est donc que le rang de la jacobienne (évaluée en 0) soit constant sur un voisinage (de 0). Cependant, cela ne suffit pas pour étudier la stabilité; une condition sur les applications ϕ et ψ doit être imposée. En effet, supposons que nous ayons une solution u de (I.4). Cette solution s'écrit $\phi \circ v$ pour une certaine fonction v , pour des t proches de 0, disons sur un intervalle I . On a alors :

$$\begin{aligned} \partial(\phi \circ v) &= f(\phi(v)) \\ \partial\phi|_v \cdot \partial v &= \psi(\partial f(0) \cdot v). \end{aligned}$$

le tout sur I . Bien que cela permet l'étude des points stationnaires, cela ne garantit en rien que leur stabilité sera conservée.

I.3.2 Transformation de solutions d'équations différentielles via un difféomorphisme

Au lieu de passer directement au système linéaire, fixons ϕ un difféomorphisme local (disons $\phi : U \rightarrow V$ avec U et V ouverts), et regardons quelles sont les conditions à imposer sur une fonction $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ pour que la condition suivante soit vérifiée :

$$\begin{aligned} &\text{si } u \text{ est une solution de } \partial u = f(u), \\ &\text{alors la fonction } v = \phi \circ u \text{ est solution de } \partial v = g(v). \end{aligned} \quad (\text{I.12})$$

[Notations pas géniales : $\phi \circ u$...] Il faut et il suffit pour cela que $\partial(\phi \circ u) = g(\phi(u))$, mais le premier membre, après utilisation de la dérivation composée et de la définition de u , vaut $\partial\phi|_u \cdot f(u)$. Cela impose, en posant $v = \phi(u)$, que

$$g(v) = \partial\phi|_{\phi^{-1}(v)} \cdot f(\phi^{-1}(v)). \quad (\text{I.13})$$

De la même façon, en inversant les rôles de f et g et en prenant le difféomorphisme inverse, on a

$$f(u) = \partial\phi_{|\phi(u)}^{-1} \cdot g(\phi(u))$$

et cela conduit aux définitions suivantes :

Définition 8. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction et ϕ un difféomorphisme local défini sur un ouvert U et à valeurs dans un ouvert V . On dit que :

- g est le *push-forward* de f par ϕ si g est défini par l'équation (I.13), pour tout $v \in V$;
- g est le *pull-back* de f par ϕ si, pour tout $u \in U$,

$$g(u) = \partial\phi_{|\phi(u)}^{-1} \cdot f(\phi(u)). \quad (\text{I.14})$$

Le push-forward de f par ϕ est noté $\phi_* f$ et le pull-back de f par ϕ , est noté $\phi^* f$.

Comme conséquence immédiate, on obtient que si g est le push-forward de f par le difféomorphisme local ϕ , alors la condition (I.12) est vérifiée ; toutes les solutions du système sont alors « poussées » vers le nouveau système. Par ailleurs, cela impose une condition supplémentaire sur le difféomorphisme ψ qu'on retrouve dans le théorème du rang constant 10 : en effet, si l'on souhaite passer de f à $\partial f(0)$ tout en conservant les solutions (localement), on doit avoir, d'une part, grâce au théorème 10 :

$$f(\phi(x)) = \psi(\partial f(0) \cdot x),$$

et d'autre part, par définition de push-forward :

$$\partial f(0) \cdot u = \partial\phi_{|\phi^{-1}(u)} \cdot f(\phi^{-1}(u));$$

dès lors, on obtient la condition :

$$f \circ \phi(x) = \psi \left(\partial\phi_{|\phi^{-1}(x)} \cdot f(\phi^{-1}(x)) \right).$$

Les points stationnaires du système initial sont transférés en points stationnaires dans le nouveau système ; mais cette fois, on peut discuter de la stabilité, grâce à la proposition suivante :

Proposition 11. Soit u un point stationnaire de l'équation (I.4) :

$$\partial u = f(u).$$

Supposons que v une solution du même système, et que $v(t) \xrightarrow[t \rightarrow +\infty]{} u$. Alors $\phi(u)$ est point stationnaire pour l'équation

$$\partial v = \phi_* f(v), \quad (\text{I.15})$$

et $\phi(v(t)) \xrightarrow[t \rightarrow +\infty]{} \phi(u)$.

Le théorème suivant est une conséquence immédiate de la proposition.

Théorème 12. *Si u est un point stationnaire stable de l'équation (I.4), alors $\phi(u)$ est point stationnaire stable de l'équation (I.15).*

Puisque les mêmes arguments peuvent être employés pour la preuve de ce théorème et celle de la proposition précédente, nous ne montrons que celle du théorème.

Démonstration. On sait qu'il existe $\varepsilon > 0$ tel que, pour toute solution w de (I.4) vérifiant $|w(0) - u(0)| < \varepsilon$, on a $w(t) \xrightarrow[t \rightarrow +\infty]{} u(0)$. On doit montrer qu'il existe $\delta > 0$ tel que pour toute solution \tilde{w} de (I.15) vérifiant $|\tilde{w}(0) - \phi(u(0))| < \delta$, on ait $\tilde{w}(t) \xrightarrow[t \rightarrow +\infty]{} \phi(u(0))$. Ce δ nous est fourni par la continuité de ϕ en $u(0)$; en effet, puisque ϕ est un difféomorphisme qui maintient les solutions, toute solution \tilde{w} de (I.15) s'exprime sous la forme $\phi(w)$ avec w solution de (I.4). Vu la convergence de $w(t)$ vers $u(0)$ dès que $w(0)$ est distant de $u(0)$ d'au plus ε , on obtient le résultat par continuité de ϕ en $u(0)$. \square

L'idée est donc simple : on part à la recherche des difféomorphismes ϕ pour lesquels le push-forward de f par ϕ , est linéaire. Evidemment, comme déjà remarqué en début de section, certaines fonctions f n'admettront pas de tels difféomorphismes ; cela conduit à la définition :

Définition 9. Une fonction f est dite C^r -linéarisable en un point $a \in \mathbb{R}^n$ s'il existe un difféomorphisme local en a , $\phi : U \rightarrow V$, de classe C^r , s'annulant en a , et tel que ϕ_*f soit linéaire.

I.3.3 Linéarisation au voisinage d'un point non-singulier

Un résultat important est le suivant : si la fonction f ne s'annule pas en un point $a \in \mathbb{R}^n$, alors le difféomorphisme ϕ ci-dessus peut être choisi de telle sorte que ϕ_*f soit constante. Pour le montrer, nous allons utiliser la version suivante du théorème de Cauchy- Lipschitz, théorème que nous avons déjà évoqué en début de chapitre :

Lemme 13 (Théorème de Cauchy-Lipschitz). *Soient $a \in \mathbb{R}^n$, et V un ouvert contenant a . Supposons que $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ soit une fonction de classe C^r sur V , pour un certain $r \geq 1$. Il existe une fonction $\Phi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$, de classe C^r , définie dans un voisinage $U \times I$ du couple $(a, 0)$, avec $U \subset V$ et vérifiant les deux conditions : $\Phi(\cdot, 0) = \text{Id}$ sur U ; $\partial_t \Phi(x, t) = f(\Phi(x, t))$.*

[Pour plus de clarté, on pourrait l'énoncer en début de chapitre ; d'autant plus qu'il est évoqué dans le texte (et non sous forme d'une prop).] Le livre de M. Demazure [5], donne une démonstration directe de ce fait, utilisant le théorème des fonctions implicites. Ce théorème, ainsi énoncé, nous permet de prouver le résultat suivant :

Théorème 14. *Soit U un ouvert de \mathbb{R}^n . Si $f : U \rightarrow \mathbb{R}^n$ est de classe C^r , pour un certain $r \geq 1$, et $a \in U$ est un point pour lequel f est non-nulle, alors il existe deux ouverts $U' \subset U$ et V de \mathbb{R}^n contenant respectivement a et 0 , ainsi qu'un difféomorphisme local $\phi : U' \rightarrow V$, tel que si u est une solution de (I.4), avec condition initiale $u(0) \in U'$, alors $\phi^{-1} \circ u$ est solution de $\partial v = (1, 0, \dots, 0)$, avec $v(0) \in V$.*

D'un point de vue plus géométrique, nous sommes en train d'affirmer que l'on peut, au voisinage d'un point où f ne s'annule pas, penser que le flux est constitué de lignes droites ; plus formellement, le flux est « topologiquement équivalent » à un réseau de droites.

Démonstration. Sans perte de généralité, on peut supposer que $a = 0$. Le lemme précédent nous fournit une application $\Phi : O \times I \rightarrow \mathbb{R}^n$, avec I intervalle ouvert, et O ouvert contenant 0 et inclus à U . Pour pouvoir construire notre difféomorphisme ϕ , l'idée est évidemment de se baser sur Φ ; or les espaces de départ ne peuvent coïncider, puisque Φ prend $n + 1$ arguments (n arguments liés à un vecteur de O , plus un argument « de temps », dans I). Pour pouvoir définir ϕ , on doit donc « tuer » une dimension ; cela va être possible grâce au vecteur $f(0)$, supposé non-nul.

Considérons E un hyperplan de \mathbb{R}^n ne contenant pas $f(0)$. On pose $\tilde{O} := E \cap O$; et $\tilde{\Phi}$ la restriction de Φ sur $I \times \tilde{O}$. On obtient alors, pour tout $t \in I$ et tout $x \in \tilde{O}$:

$$\partial_t \tilde{\Phi}(x, t) = f(\Phi(x, t))$$

d'où, en évaluant en $(0, 0)$:

$$\partial_t \tilde{\Phi}(0, 0) = f(\Phi(0, 0)) = f(0) \neq 0.$$

Par ailleurs, puisque $\tilde{\Phi}$ et Φ sont identiques sur $I \times \tilde{O}$, et que $\Phi(x, 0) = x$ pour tout $x \in O$, on obtient

$$\partial_x \tilde{\Phi}(0, 0) = \begin{pmatrix} 0 \\ \text{Id}_{n-1} \end{pmatrix}$$

Fixons une base (e_2, \dots, e_n) pour E et $e_1 = f(0)$, et définissons

$$\begin{aligned} \phi : I \times \tilde{O} &\rightarrow \mathbb{R}^n \\ (t, x) &\mapsto \phi(t, x) := \tilde{\Phi}(x, t). \end{aligned} \tag{I.16}$$

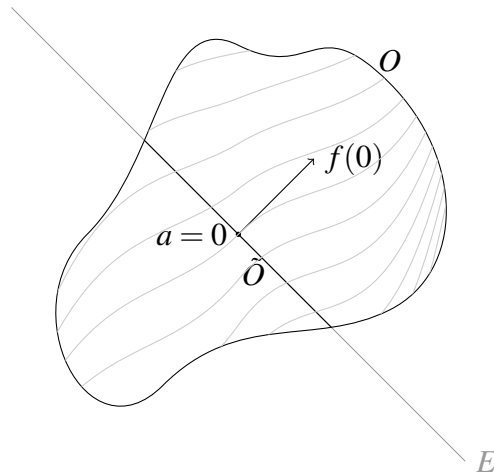


FIG. I.3 – Illustration du théorème 14 ; il affirme que les lignes grises (représentant le flux) peuvent être vues comme un réseau de droites.

Alors, la matrice Jacobienne de ϕ , évaluée en $(0, 0)$, est l'identité. On peut alors trouver deux ouverts $U' \subset I \times \tilde{O}$ et $V \subset \mathbb{R}^n$ tels que la différentielle de ϕ soit non nulle sur U' et donc, que ϕ soit un difféomorphisme entre U' et V . Sans perte de généralité, quitte à restreindre U' , on peut supposer que, pour tout $x \in U'$, on ait $f(x) \cdot f(0) \neq 0$.

Nous montrons à présent que ϕ convient au problème. Soit v une fonction, supposons que $\phi \circ v$ soit solution de l'équation $\partial u = f(u)$, et que $\phi \circ v(0) \in U'$. Alors, $\phi \circ v(t) = \Phi(v(0), t)$, pour t dans un intervalle ouvert I' contenant 0. Vu l'hypothèse sur U' , on peut affirmer que $v(t) \cdot f(0) \neq 0$ pour tout $t \in I'$; il existe alors $t_0 \in I'$ tel que $v(t_0) \in \tilde{O}$. La fonction $w = v(\cdot + t_0)$ possède aussi les propriétés précédentes, en particulier, $\phi \circ w(t) = \Phi(w(0), t)$; mais par définition de ϕ , on a $\phi(w(t)) = \Phi(\tilde{w}(t), w_1(t))$ (lorsque w s'écrit (w_1, \tilde{w})). On en déduit que $\tilde{w}(t) = \Phi(w(0), t - w_1(t))$ et ce, pour tout t ; par définition, cette relation a lieu dans \tilde{O} . Or, toute solution de l'équation $\partial u = f(u)$ s'exprime à travers Φ et doit nécessairement traverser \tilde{O} ; cela a pour conséquence que $\tilde{w}(t)$ est constante, donc $w_1(t) = x_1 + t$ pour un certain $x_1 \in \mathbb{R}$ et pour tout $t \in I'$. Cette constante x_1 s'avère être nulle en considérant $w_1(0)$, et on trouve que $w(t) = \tilde{x} + t(1, 0, \dots, 0)$ pour un certain $\tilde{x} \in \tilde{U}$. Retournant à v , on a $v(t) = \tilde{x} + (t - t_0)(1, 0, \dots, 0)$. On voit clairement que v et w sont solutions de l'équation $\partial v = (1, 0, \dots, 0)$. \square

La construction ci-dessus ne fonctionne pas pour les points en lesquels la fonction f s'annule. On va voir que, dans ce cas, des problèmes supplémentaires se posent.

I.3.4 Linéarisation au voisinage d'un point singulier, résonances

Cette fois, on considère des conditions sur la différentielle de f pour pouvoir linéariser. En fait, on définit une notion sur les applications linéaires :

Définition 10. Soit L un endomorphisme de \mathbb{R}^n . On dit qu'elle est C^r -rigide si, pour toute fonction f de classe C^r au voisinage de 0, si $f(0) = 0$ et si $\partial f(0) = L$, alors il existe un difféomorphisme local ϕ de classe C^r , s'annulant en 0, de différentielle égale à Id, et telle que $\phi^* f = L$.

Cette définition implique la proposition suivante :

Proposition 15. Soit f une fonction de classe C^r . Si $\partial f(0)$ est C^r -rigide, alors f est C^r -linéarisable.

Démonstration. Par définition de C^r -rigidité, on a l'existence d'un difféomorphisme local ϕ vérifiant $\phi^* f = L$. Ce difféomorphisme est celui qui convient pour linéariser f . \square

Il suffit donc d'étudier la rigidité des applications linéaires pour en déduire une éventuelle linéarisation de la fonction f . Pour ce faire, nous allons étudier le comportement de ϕ vis-à-vis de f .

Soit L un endomorphisme de \mathbb{R}^n , et f une fonction de classe C^r . La condition $\phi^* L = f$ équivaut à

$$\partial\phi(x) \cdot f(x) = L(\phi(x)). \quad (\text{I.17})$$

Afin de déterminer des conditions nécessaires ou suffisantes sur L pour qu'elle soit C^r -rigide, on va écrire les développements de Taylor de f et ϕ jusqu'à l'ordre r , en 0. Désignons par $P_n^{(i)}$ l'ensemble des applications polynômes en n variables, homogènes de degré i (c'est-à-dire, les polynômes qui définissent ces applications n'ont que monômes de degré exactement i). Remarquons au passage que $P_n^{(1)}$ est l'ensemble des endomorphismes linéaires de \mathbb{R}^n .

On peut alors écrire :

$$f(x) = Lx + f_2(x) + \cdots + f_r(x) + R_f(x) \quad (\text{I.18})$$

$$\phi(x) = x + \phi_2(x) + \cdots + \phi_r(x) + R_\phi(x) \quad (\text{I.19})$$

pour tout x au voisinage de 0, car par hypothèse, $f(0) = \phi(0) = 0$, $\partial f(0) = L$ et $\partial\phi(0) = \text{Id}$; chaque f_i et chaque ϕ_i est un élément de $P_n^{(i)}$.

Avec cela, l'égalité (I.17) se réécrit :

$$\begin{aligned} \partial(\text{Id} + \phi_2 + \cdots + \phi_r + R_\phi)(x) \cdot (Lx + f_2(x) + \cdots + f_r(x) + R_f(x)) \\ = L(\text{Id}(x) + \phi_2(x) + \cdots + \phi_r(x) + R_\phi(x)), \end{aligned}$$

soit, en distribuant :

$$\begin{aligned}
& \partial \text{Id}(x) \cdot (Lx + f_2(x) + \cdots + f_r(x) + R_f(x)) \\
& \quad + \partial \phi_2(x) \cdot (Lx + f_2(x) + \cdots + f_r(x) + R_f(x)) + \cdots \\
& \quad + \partial \phi_r(x) \cdot (Lx + f_2(x) + \cdots + f_r(x) + R_f(x)) \\
& \quad + \partial R_\phi(x) \cdot (Lx + f_2(x) + \cdots + f_r(x) + R_f(x)) \\
& = Lx + L(\phi_2(x)) + \cdots + L(\phi_r(x)) + L(R_\phi(x)).
\end{aligned}$$

En égalant les termes degré par degré, on obtient :

- pour les termes de degré 1 : $Lx = Lx$;
- pour les termes de degré 2 : $f_2(x) + \partial \phi_2(x)(Lx) = L(\phi_2(x))$, ou encore $L(\phi_2(x)) - \partial \phi_2(x)(Lx) = f_2(x)$;
- pour les termes de degré 3 : $f_3(x) + \partial \phi_2(x)(f_2(x)) + \partial \phi_3(x)(Lx) = L(\phi_3(x))$, d'où $L(\phi_3(x)) - \partial \phi_3(x)(Lx) = f_3(x) + \partial \phi_2(x)(f_2(x))$;
- ...
- pour les restes : $R(x) = LR_\phi(x)$, où R est la somme de tous les termes du membre de gauche faisant intervenir R_f ou R_ϕ ainsi que tous les termes polynômiaux homogènes de degré strictement supérieur à r .

De façon générale, définissons, pour tout $k = 2, \dots, r$, l'application A_k , endomorphisme de $P_n^{(k)}$ de la façon suivante :

$$\begin{aligned}
A_k : P_n^{(k)} &\rightarrow P_n^{(k)} \\
p &\mapsto A_k(p) \\
\text{avec } A_k(p) : x &\mapsto L(p(x)) - \partial p(x) \cdot Lx.
\end{aligned} \tag{I.20}$$

Les équations déduites ci-dessus, s'écrivent alors sous une même forme, à savoir : pour tout $k = 2, \dots, r$, on a l'égalité $A_k(\phi_k) = f_k + F_{k-1}(f, \phi)$, où $F_{k-1}(f, \phi)$ ne fait intervenir que les $\partial \phi_i$ et f_i , pour $i < k$. Par exemple, on a $F_1(f, \phi)(x) = 0$, $F_2(f, \phi)(x) = \partial \phi_2(x)(f_2(x))$, et ainsi de suite.

De ces égalités, on peut déduire certaines informations. Supposons, pour commencer, qu'il existe une valeur $k \geq 2$ pour laquelle A_k ne soit pas surjective. Soit u une application polynôme, homogène de degré k , se trouvant en-dehors de $\text{Im} A_k$. Si f est précisément la fonction $L + u$, alors les équations pour ce cas s'écrivent $A_i(\phi_i) = 0$, pour $i < k$ (puisque tous les f_i sont supposés nuls dans ce cas) ; et $A_k(\phi_k) = u$. Un tel ϕ_k ne pourra pas être trouvé ; dès lors, on vient de montrer qu'il n'existe pas de transformation ϕ permettant d'amener les solutions de l'équation $\partial u = f(u)$ vers celle de l'équation linéarisée.

Réciproquement, si tous les endomorphismes A_k sont surjectifs, alors on peut résoudre les égalités en ϕ_k , donc obtenir une application ϕ convenable. Par ailleurs, puisque chaque A_k est un endomorphisme linéaire, et vu l'hypothèse, il s'avère être bijectif ; cela détermine ϕ de façon unique jusqu'à l'ordre r .

En conséquence, on a le fait suivant : *Pour que L soit C^r -rigide, il faut et il suffit que les applications A_k définies en (I.20) soient surjectives.*

Or, ces endomorphismes sont surjectifs si et seulement s'ils n'admettent pas 0 comme valeur propre, cela suggère l'étude de ces applications. Pour la suite, fixons $(e_i)_{i=1,\dots,n}$ une base de \mathbb{R}^n ; et remarquons le fait suivant : si, pour un multi-indice $m = (m_1, \dots, m_n)$, de degré total $|m| = m_1 + \dots + m_n = k$, et un naturel $i \in \{1, \dots, n\}$, on définit $p_{m,i}(x_1 e_1 + \dots + x_n e_n) := x_1^{m_1} \dots x_n^{m_n} e_i$, alors la suite $(p_{m,i})_{\substack{i \in \{1, \dots, n\} \\ |m|=k}}$, forme une base de $P_n^{(k)}$.

Supposons dans un premier temps que L soit diagonalisable, et que les e_i soient des vecteurs propres de L . Alors, on a clairement $L(p_{m,i}(x)) = \lambda_i p_{m,i}(x)$ pour tous m, i et $x \in \mathbb{R}^n$. De plus, pour tous m, i, j et $x = x_1 e_1 + \dots + x_n e_n \in \mathbb{R}^n$, on a $x_j \partial_j p_{m,i}(x) = m_j p_{m,i}(x)$; dès lors,

$$\begin{aligned} \partial p_{m,i}(x)(Lx) &= \sum_j \partial_j p_{m,i}(x)(Lx)_j \\ &= \sum_j \partial_j p_{m,i}(x)(\lambda_j x_j) \\ &= \sum_j \lambda_j m_j p_{m,i}(x). \end{aligned}$$

On en déduit dès lors que

$$A_k(p_{m,i}) = (\lambda_i - \sum_j \lambda_j m_j) p_{m,i};$$

cela donnant immédiatement les valeurs et vecteurs propres de A_k . En conclusion, on obtient que A_k est surjective si et seulement si aucune des relations $\lambda_i - \sum_j \lambda_j m_j$ n'est nulle. Cela motive la définition suivante :

Définition 11. Soit L une application linéaire ; soient $\lambda_1, \dots, \lambda_n$ ses valeurs propres. On dit qu'elle *admet une résonance d'ordre k* s'il existe n naturels m_i , tels que $\sum_{i=1}^n m_i = k$, et tels qu'il existe $j \in \{1, \dots, n\}$ pour lequel

$$\lambda_j = \sum_{i=1}^n m_i \lambda_i.$$

Immédiatement, s'en déduit la proposition :

Proposition 16. *Une application linéaire diagonalisable est C^r -rigide si et seulement si elle n'admet aucune résonance d'ordre k , pour tout $k \leq r$.*

On peut prouver que le résultat reste vrai dans le cas des applications linéaires Jordanisables. Dès lors, on a une condition de non-linéarisation :

Proposition 17. *Soit f une fonction de classe C^r de \mathbb{R}^n vers \mathbb{R}^n . S'il existe un entier $k \leq r$ pour lequel $\partial f(0)$ admet une résonance, alors f n'est pas C^k -linéarisable.*

I.3.5 Variétés stable, instable, centrale

Les notions d'espaces stable et instable, rencontrés dans le cas linéaire, peuvent se généraliser au cas non-linéaire en termes de variétés ; les notions nécessaires sont rappelées dans l'annexe B. La généralisation nous est assurée par le théorème suivant :

Théorème 18 (Théorème de la variété stable). *Supposons que 0 soit un point stationnaire de l'équation (I.4), avec f de classe C^r , avec $r \geq 1$. Notons Φ le flux associé. Supposons que $\partial f(0)$ soit inversible. Alors il existe une variété différentiable $V^s(0)$, de classe C^r , dont le plan tangent en 0 est l'espace stable associé à $\partial f(0)$, et vérifiant la condition :*

$$\text{pour tout } x \in V^s(0), \Phi(x, t) \xrightarrow[t \rightarrow +\infty]{} 0.$$

De même, il existe une variété différentiable $V^u(0)$ de classe C^r , dont le plan tangent en 0 est l'espace instable associé à $\partial f(0)$, et vérifiant la condition :

$$\text{pour tout } x \in V^u(0), \Phi(x, t) \xrightarrow[t \rightarrow -\infty]{} 0.$$

Définition 12. Lorsque 0 est point stationnaire, les variétés $V^s(0)$ et $V^u(0)$, dont l'existence nous est assurée par le théorème précédent, sont respectivement appelées *variété stable en 0* et *variété instable en 0*.

La preuve qui suit ne montre que l'existence de la variété stable, on obtient facilement l'existence de la variété instable en « inversant la variable t », comme nous l'avons fait pour obtenir les définitions d'instabilité. Nous nous inspirons ici du livre de L. Perko [11].

Idée de preuve. Il suffit de prouver les deux points suivants. Premièrement, en tout point x dont le flux généré tend vers 0, il existe un ensemble U , composé exclusivement de tels points, et une application $\phi : U \rightarrow \mathbb{R}^k$ (k étant à déterminer) de telle sorte que $\phi(U)$ soit ouvert dans \mathbb{R}^k et ϕ soit de classe C^r . Deuxièmement, on doit montrer que toute courbe passant par le point 0, admet un vecteur tangent en ce point, qui est contenu dans $E^s(0)$.

Au vu du second point à prouver, le naturel k doit nécessairement valoir la dimension de $E^s(0)$. Quitte à modifier la base, on peut écrire $\partial f(0)$ sous la forme

$$\partial f(0) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix} = \widehat{A} + \widehat{B}$$

de telle sorte que A et B soient Jordanisées (ou quasi-Jordanisées réelles — voir les transformations en annexe A), A n'ayant que des valeurs propres dont la partie

réelle est strictement négative et B n'ayant que des valeurs propres dont la partie réelle est strictement positive. Le flux linéaire $e^{\partial f(0)t}$ peut alors se décomposer en somme $U(t) + V(t)$, avec

$$U(t) = e^{\hat{A}t} \quad \text{et} \quad V(t) = e^{\hat{B}t} \quad (\text{I.21})$$

Par ailleurs, le développement de Taylor à l'ordre 1 de $f = (\bar{f}, \tilde{f})$ au voisinage de 0 donne les deux égalités

$$\bar{f}(x) = A\bar{x} + \bar{g}(x) \quad \text{et} \quad \tilde{f}(x) = B\tilde{x} + \tilde{g}(x)$$

avec $g(x) = (\bar{g}(x), \tilde{g}(x)) = o(\|x\|)$; en particulier, $\partial g(0) = 0$. Résoudre l'équation $\partial u = f(u)$ sur ce voisinage revient alors à résoudre l'équation

$$\partial u = \partial f(0)\bar{u} + g(u)$$

De plus, g est de classe C^1 au moins; on peut donc majorer, pour tout ε , la norme de $\partial g(x)$ par ε , pour x suffisamment proche de 0. Cela implique le fait suivant :

$$\begin{aligned} &\text{pour tout } \varepsilon > 0, \text{ il existe } \delta > 0 \text{ tel que :} \\ &\forall x \neq y, \|x\| < \delta, \|y\| < \delta, \|g(x) - g(y)\| < \varepsilon\|x - y\|. \end{aligned} \quad (\text{I.22})$$

En effet, si x et y sont donnés, considérons $h : [0, 1] \rightarrow \mathbb{R}^n$ définie par $h(t) = g(tx + (1-t)y)$; le théorème des accroissements finis appliqué à chacune des fonctions h_i (qui sont de classe C^1) donne l'existence d'un $t_i \in]0, 1[$ tel que $|h_i(1) - h_i(0)| \leq |\partial_t h_i(t_i)|$; mais cette dérivée vaut précisément $\partial g_i(t_i(y-x) + x) \cdot (y-x)$. Dès lors, pour tout $i = 1, \dots, n$, on a l'inégalité

$$\begin{aligned} |g_i(x) - g_i(y)| &\leq |\partial g_i(t_i(y-x) + x) \cdot (y-x)| \\ &\leq \|\nabla g_i(t_i(y-x) + x)\|_2 \|y-x\|_2 \end{aligned}$$

et chacun des gradients apparaissant ci-dessus peut être majoré strictement, dans un voisinage de x , par ε . Le résultat annoncé s'en déduit alors.

On fixe à présent $\alpha > 0$ de telle sorte que $-\alpha$ majore toutes les parties réelles des valeurs propres négatives. Ce choix fait, on peut trouver $K > 0$ suffisamment grand et $\sigma > 0$ suffisamment petit, de telle sorte que l'on ait les inégalités

$$\begin{aligned} \|U(t)\|_\infty &\leq K e^{-(\alpha+\sigma)t}, \quad \text{pour tout } t \geq 0 \\ \text{et } \|V(t)\|_\infty &\leq K e^{\sigma t}, \quad \text{pour tout } t \leq 0. \end{aligned}$$

A ε fixé, considérons à présent l'équation intégrale en Φ :

$$\Phi(x, t) = U(t)x + \int_0^t U(t-s)g(\Phi(x, s)) ds - \int_t^{+\infty} V(t-s)g(\Phi(x, s)) ds \quad (\text{I.23})$$

pour $x \in B(0, \eta)$, où η sera choisi plus petit que δ , ce dernier étant fourni par (I.22). On peut montrer que si $\Phi(x, t)$ est une solution continue de cette équation, alors il s'agit d'une solution de l'équation différentielle initiale.

On note que l'équation ci-dessus équivaut aux égalités

$$\begin{aligned}\bar{\Phi}(x, t) &= e^{At} \bar{x} + \int_0^t e^{A(t-s)} \bar{g}(\Phi(x, s)) ds \\ \tilde{\Phi}(x, t) &= - \int_t^{+\infty} e^{B(t-s)} \tilde{g}(\Phi(x, s)) ds,\end{aligned}$$

cela donne en particulier les conditions initiales :

$$\bar{\Phi}(x, 0) = \bar{x} \quad \text{et} \quad \tilde{\Phi}(x, 0) = - \int_0^{+\infty} V(-s)g(\Phi(\bar{x}, s)) ds.$$

Grâce à cela, on peut définir $\psi : \mathbb{R}^k \rightarrow \mathbb{R}^{n-k}$ par $\psi(y) = \tilde{\Phi}((y, 0), 0)$, et obtenir le fait suivant : si $y = \Phi(\bar{x}, 0)$, on a $\bar{y} = \bar{x}$ et $\tilde{y} = \tilde{\Phi}(\bar{x}, 0) = \psi(\bar{x})$; cela implique les $n - k$ égalités

$$y_j = \psi_j(\bar{y})$$

si bien que, en posant $X := \{y \in \mathbb{R}^k \mid (0, y) \in B(0, \eta)\}$, l'ensemble $V^s(0) := X \times \psi(X) \subset \mathbb{R}^n$ se trouve être une variété de dimension k (on définit les cartes comme étant les projections sur X).

Nous montrons que (I.23) admet bien une solution continue ; de sorte que le raisonnement précédent conduit bien à la construction d'une variété différentiable. Nous donnons ensuite les principaux arguments qui permettent de déduire qu'il s'agit bien de la bonne variété.

Nous trouvons une solution de l'équation intégrale (I.23) par approximations successives, comme suit. On fixe $\phi_0(x, t) := 0$; puis

$$\phi_{m+1}(x, t) = U(t)x + \int_0^t U(t-s)g(\phi_m(x, s)) ds - \int_t^{+\infty} V(t-s)g(\phi_m(x, s)) ds;$$

notons que chaque ϕ_m est de on prouve alors la majoration suivante pour tout n , pour tout x et pour tout $t \geq 0$:

$$\|\phi_{m+1}(x, t) - \phi_m(x, t)\| \leq \frac{K\|x\| e^{-\alpha t}}{2^m} \quad (\text{I.24})$$

Pour $m = 0$, puisque $\phi_1(x, t) = U(t)x$, l'égalité est vérifiée. Supposons-la vraie pour tout $m' \leq m$; alors on a les inégalités suivantes :

$$\begin{aligned}
& \|\phi_{m+1}(x,t) - \phi_m(x,t)\| \\
&= \left\| \int_0^t U(t-s) (g(\phi_m(x,s)) - g(\phi_{m-1}(x,s))) ds \right. \\
&\quad \left. - \int_t^{+\infty} V(t-s) (g(\phi_m(x,s)) - g(\phi_{m-1}(x,s))) ds \right\| \\
&\leq \int_0^t \|U(t-s) (g(\phi_m(x,s)) - g(\phi_{m-1}(x,s)))\| ds \\
&\quad + \int_t^{+\infty} \|V(t-s) (g(\phi_m(x,s)) - g(\phi_{m-1}(x,s)))\| ds \\
&\leq \int_0^t \|U(t-s)\|_{\infty} \varepsilon \|\phi_m(x,s) - \phi_{m-1}(x,s)\| ds \\
&\quad + \int_t^{+\infty} \|V(t-s)\|_{\infty} \varepsilon \|\phi_m(x,s) - \phi_{m-1}(x,s)\| ds \\
&\leq \varepsilon \int_0^t K e^{-(\alpha+\sigma)(t-s)} \frac{K\|x\| e^{-\alpha s}}{2^{m-1}} ds + \varepsilon \int_t^{+\infty} K e^{\sigma s} \frac{K\|x\| e^{-\alpha s}}{2^{m-1}} \\
&\leq \frac{\varepsilon K^2 \|x\| e^{-\alpha t}}{\sigma 2^{m-1}} + \frac{\varepsilon K^2 \|x\| e^{-\alpha t}}{\sigma 2^{m-1}}
\end{aligned}$$

La majoration de cette somme par la fraction $\frac{K\|x\| e^{-\alpha t}}{2^m}$ est possible pour autant que l'on ait $\frac{\varepsilon K}{\sigma} < \frac{1}{4}$; il faut donc choisir ε plus petit que $\frac{\sigma}{4K}$. Une fois un tel ε fixé, la condition (I.22) est vérifiée pour tous les points situés dans $B(0, \delta)$. Nous voulons appliquer cette condition sur les $\phi_m(x, t)$; par conséquent nous devons imposer $\phi_m(x, t) < \delta$ de sorte que l'inégalité puisse s'appliquer. Mais nous avons une borne supérieure sur la norme de $\phi_1(x, t)$; au vu de (I.24), cette dernière peut être majorée par $K\|x\|$. En imposant que cette constante soit inférieure à $\frac{\delta}{2}$, la première majoration (c'est-à-dire, celle qui permet de prouver le fait pour $m = 1$) est valide. Puisqu'on suppose $\|x\| < \eta$, il suffit de prendre $\eta := \frac{\delta}{2K}$. Prouvons que les constantes ε et η ainsi fixées, nous permettent de conclure. Il suffit pour cela de voir que, pour tout m , $\phi_m(x, t)$ est en norme plus petit que δ . Mais

$$\begin{aligned}
\|\phi_m(x, t)\| &\leq \sum_{i=1}^m \|\phi_i(x, t) - \phi_{i-1}(x, t)\| && \text{(par inégalité triangulaire ;} \\
&&& \text{notons que } \phi_0(x, t) = 0) \\
&\leq \sum_{i=1}^m \frac{K\|x\| e^{-\alpha t}}{2^i} && \text{(par hypothèse de récurrence)} \\
&\leq K\eta \sum_{i=1}^m \frac{1}{2^i} && \text{(par majoration de } \|x\| \text{ et de} \\
&&& \text{l'exponentielle)} \\
&\leq \frac{\delta}{2} \left(2 - \frac{2}{2^{m+1}}\right) && \text{(par définition de } \eta \text{ et somme} \\
&&& \text{géométrique)}
\end{aligned}$$

et cette dernière expression est bien strictement inférieure à δ , pour tout m .

Le fait d'avoir majoré suffisamment chaque distance entre deux itérées, assure que la suite de fonctions $(\phi_m(x, t))_{m \in \mathbb{N}}$, définies pour tout $t \geq 0$ et tout x de norme plus petite que η est de Cauchy. En effet, si $\varepsilon > 0$ nous est donné, en choisissant N de telle sorte que $\frac{\eta^K}{2^{N-1}} < \varepsilon$, on a, pour tous m et $m' > m$ plus grands que N , pour tout $t \geq 0$ et pour tout $x \in B(0, \eta)$:

$$\begin{aligned} \|\phi_{m'}(x, t) - \phi_m(x, t)\| &\leq \sum_{i=m}^{m'-1} \|\phi_{i+1}(x, t) - \phi_i(x, t)\| \\ &\leq \sum_{i=m}^{m'-1} \frac{K\|x\| e^{-\alpha t}}{2^i} \\ &\leq K\eta \sum_{i=N}^{m'-1} \frac{1}{2^i} = \frac{K\eta}{2^N} \sum_{i=0}^{m'-N-1} \frac{1}{2^i} \\ &\leq \frac{K\eta}{2^{N-1}} \end{aligned}$$

car la somme apparaissant dans l'avant-dernière inégalité vaut $2(1 - \frac{1}{2^{m'-N}})$. En conséquence, elle converge uniformément vers une fonction Φ , définie et continue sur $B(0, \eta) \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$. Mais par cette convergence uniforme, en prenant les limites dans chacun des membres de (I.23), on voit que Φ est solution de cette équation, donc aussi de l'équation différentielle initiale. L'uniforme convergence implique aussi que Φ est différentiable sur son domaine ; et l'inégalité prouvée par récurrence montre que

$$|\Phi(x, t)| \leq 2K\|x\| e^{-\alpha t}, \quad (\text{I.25})$$

pour tout $t \geq 0$ et tout x de norme plus petite que η .

De plus, si v est une solution de l'équation différentielle, avec condition initiale $v(0) \in V^s(0)$, alors $v(t) = \Phi(x, t)$ pour un certain x , d'où par unicité des solutions, $v(t) = \Phi(x, t)$. La majoration connue sur Φ donne alors que $v(t) \rightarrow 0$ quand $t \rightarrow \infty$. Par ailleurs, on peut montrer que c'est équivalent³ ; c'est-à-dire, si $v(0) \notin V^s(0)$, alors $v(t) \not\rightarrow 0$ pour $t \rightarrow \infty$. La proposition (1) relative aux translatées de solutions d'équations autonomes, donne alors que $v(t) \in V^s(0)$, pour tout $t \geq 0$.

Le fait que cette variété soit tangente en 0 à $E^s(0)$ est obtenu par l'annulation de la différentielle $\partial\psi$. \square

Il n'existe pas de généralisation naturelle, unique, pour l'espace central. Toutefois, un théorème similaire au théorème de la variété stable peut être prouvé :

Théorème 19 (Théorème de la variété centrale). *Si 0 est un point stationnaire de l'équation (I.4), avec f de classe C^r au voisinage de 0, alors il existe une variété*

³ L. Perko suggère la référence de E. Coddington et N. Levinson [4] pour une preuve de ce fait.

différentiable $V^c(0)$ de classe C^r , dont le plan tangent en 0 est $E^c(0)$. Par ailleurs, si la variété instable est réduite à $\{0\}$, alors il existe une fonction $h : \mathbb{R}^c \rightarrow \mathbb{R}^s$, c et s étant les dimensions respectives de $E^c(0)$ et de $E^s(0)$, telle que les équations

$$\partial u = f(u) \text{ sur } V^c(0) \quad \text{et} \quad \partial v = \partial f(0)|_{E^c(0)} v + f(v, h(v)) \text{ sur } \mathbb{R}^c \quad (\text{I.26})$$

fournissent des flux topologiquement équivalents au voisinage de 0.

A la différence des variétés stable et instable, la variété centrale n'est pas unique. Cela est d'autant plus vrai pour les fonctions h qui apparaissent dans le théorème. Nous référons au livre de J. Carr [3] pour une preuve détaillée.

Nous admettons enfin le théorème suivant, qui nous sera fort utile dans l'étude des bifurcations :

Théorème 20. *Si 0 est point stationnaire de (I.4), et si la variété instable qui lui est associée est réduite à ce point, alors sa stabilité est induite par celle qu'a 0 dans le second système de (I.26).*

Dès lors, lorsque la variété centrale n'est pas réduite à 0, le comportement qu'a le flux dessus détermine la stabilité du point stationnaire 0.

Chapitre II

Aspects théoriques des bifurcations

Dans le premier chapitre, nous avons étudié la stabilité de solutions d'équations différentielles autonomes, et en particulier des solutions particulières constantes, c'est-à-dire, des points stationnaires. Considérons à présent, plus généralement, l'équation différentielle

$$\partial_t u = f(u, \mu) \tag{II.1}$$

où $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ est une fonction qu'on supposera, au minimum, continue en μ et localement Lipschitz en u . Pour μ fixé, on peut déterminer les points stationnaires et discuter de leur stabilité, comme réalisé au chapitre précédent.

Qu'en est-il maintenant si le paramètre μ varie ? C'est l'essence même de la théorie des bifurcations. Nous allons comprendre comment les points stationnaires peuvent varier ; d'abord dans le cas linéaire, puis dans le cas non-linéaire, où des comportements plus variés apparaissent. Parmi ces comportements, certains méritent une attention particulière ; car non seulement ils modifient le comportement des points stationnaires (en terme de stabilité), mais également leur nombre. Nous étudierons quelques-uns de ces phénomènes.

Nous nous inspirons des livres de M. Demazure [5] et de P. Glendinning [6] pour l'étude du cas linéaire ; les théorèmes de bifurcations sont essentiellement inspirés par P. Glendinning [6] et par le livre de S. Wiggins [15].

II.1 Equations différentielles linéaires avec paramètre

On considère l'équation différentielle

$$\partial_t u = A(\mu)u + B(\mu) \tag{II.2}$$

où $A : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ est une application à valeurs matricielles et $B : \mathbb{R} \rightarrow \mathbb{R}^n$ est une application à valeurs vectorielles. Pour μ fixé, les solutions sont connues ; comme rappelé au chapitre précédent, ce sont des exponentielles matricielles (avec, dans ce cas-ci, un terme indépendant). On peut également déterminer les points stationnaires et discuter de leur stabilité.

Le problème qui nous préoccupe est de connaître leur comportement lorsque le paramètre varie. Pour bien comprendre les différents phénomènes, nous allons commencer par considérer la dimension 1. Nous parlerons également de la dimension deux, où des comportements plus variés surviennent. La dimension n en général, donne des comportements encore plus divers ; elle amène des discussions selon le nombre de valeurs propres qui changent de signe en même temps, offrant alors des structures de flots très variés. Il est difficile de pouvoir appréhender les diverses possibilités, et nous n'en ferons donc pas d'étude détaillée.

II.1.1 Le problème dans \mathbb{R}

Dans ce cadre, nous voulons connaître le comportement des fonctions u_μ , solutions de

$$\partial_t u = A(\mu)u + B(\mu)$$

lorsque μ varie. A et B sont ici simplement des fonctions réelles. Les solutions sont, pour $A(\mu) \neq 0$:

$$u_\mu(t) = e^{A(\mu)t} u_0 - \frac{B(\mu)}{A(\mu)}$$

et si $A(\mu) = 0$, elles deviennent

$$u_\mu(t) = B(\mu)t + u_0$$

pour $u_0 \in \mathbb{R}$. Lorsque μ varie, c'est donc principalement A qui régit le comportement des u_μ ; ce dernier change dès que l'on passe sur une racine de A . En particulier, si A est continue, et que l'on considère une valeur de μ qui n'est pas racine de A , alors, au voisinage de ce μ , les solutions de l'équation différentielle auront un comportement similaire aux u_μ .

Pour $A(\mu) \neq 0$, le seul point stationnaire est $-\frac{B(\mu)}{A(\mu)}$ (obtenu lorsque $u_0 = 0$). Concernant sa stabilité asymptotique, une solution proche en 0 tendra vers $-\frac{B(\mu)}{A(\mu)}$ pour $t \rightarrow +\infty$ si et seulement si $e^{A(\mu)t} u_0$ tend vers 0 pour $t \rightarrow +\infty$; ce qui est vérifié si et seulement si $A(\mu) < 0$. On obtient de même que $-\frac{B(\mu)}{A(\mu)}$ est instable si et seulement si $A(\mu) > 0$. Pour $A(\mu) = 0$, le nombre de points stationnaires est déterminé par la valeur que prend $B(\mu)$; si ce dernier est nul, alors seules les fonctions constantes sont solutions, et elles sont toutes points stationnaires ; dans le cas contraire, les solutions sont des fonctions affines avec même pente et aucune

ne peut être point stationnaire. La discussion sur la stabilité ne concerne donc que le cas $A(\mu) = B(\mu) = 0$; or dans ce cas, il ne peut y avoir stabilité ou instabilité asymptotique puisque toutes les solutions sont constantes (et donc gardent toujours la même distance en tout $t \in \mathbb{R}$). Néanmoins, tous les points stationnaires sont à la fois stables et instables au sens de Lyapounov.

II.1.2 Etude de la dimension 2

Dans le cas de la dimension 2, on peut écrire en tout μ , la matrice $A(\mu)$ sous la forme

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

La fonction g , qui à μ , associe le polynôme $\det(A(\mu) - X \text{Id})$, donne en tout μ un polynôme du second degré, qu'on peut écrire explicitement en fonction de A ; on a

$$\begin{aligned} g(\mu)(X) &= X^2 - (a(\mu) + d(\mu))X + (a(\mu)d(\mu) - b(\mu)c(\mu)) \\ &= X^2 - \text{tr}A(\mu)X + \det A(\mu). \end{aligned}$$

Le comportement de g est donc entièrement déterminé en fonction du déterminant et de la trace de A . Pour tout μ , g admettra soit deux racines réelles (distinctes ou non), soit deux racines complexes conjuguées. Nous serons dans le cas de racines réelles si et seulement si $D(\mu) := (\text{tr}A(\mu))^2 - 4\det A(\mu) \geq 0$. Si $D(\mu) > 0$, alors $A(\mu)$ admet deux valeurs propres réelles, dont la somme vaut la trace et le produit vaut le déterminant. Si par contre, $D(\mu) = 0$, $A(\mu)$ admet alors une valeur propre réelle double; dans ce cas, la trace en est son double et le déterminant, son carré. Ce n'est donc qu'un cas particulier des deux faits précédents.

Connaître la somme et le produit des valeurs propres nous permet d'en déduire leur signe et donc, la stabilité des points stationnaires de l'équation (II.2). En effet, à μ fixé :

- si $\det A(\mu) < 0$, alors les deux valeurs propres sont non-nulles et de signes opposés; il n'y a qu'un seul point stationnaire qu'on note S . Si E est l'espace propre associé à la valeur propre négative, alors toute solution de (II.2), dont la condition initiale est de la forme $u(0) = S + v$, $v \in E$, convergera vers S ; tandis que toutes les autres s'approcheront de l'espace affine $S + E$ sans jamais l'atteindre. Par ailleurs, si w est une telle fonction, alors la fonction $t \mapsto \|w(t)\|$ tend vers $+\infty$ lorsque $t \rightarrow +\infty$. Le point stationnaire S est donc un point de selle.
- si $\det A(\mu) > 0$, alors les deux valeurs propres sont non-nulles et de même signe; il n'y a toujours qu'un seul point stationnaire noté S . Le signe des valeurs propres est donc le même que la trace : si celle-ci est positive, alors

S est instable ; si elle est négative, alors S est stable. On parle dans ce cas respectivement de nœud instable et stable. Remarquons que la trace ne peut pas être nulle, sans quoi les valeurs propres se trouveraient opposées, donc leur produit, qui est le déterminant, serait négatif (ou nul).

- si $\det A(\mu) = 0$, on est en un point où A n'est pas inversible ; l'existence de points stationnaires est alors déterminé par la valeur de $B(\mu)$. Si on applique sur $B(\mu)$ les transformations qui permettent de diagonaliser (ou de jordaniser) $A(\mu)$, obtenant ainsi un vecteur $v(\mu)$, une condition nécessaire et suffisante pour assurer cette existence est que l'un des deux points suivants soit vérifié :

- soit $A(\mu)$ et $B(\mu)$ sont tous deux nuls,
- soit $A(\mu)$ est de rang 1 et la seconde composante de $v(\mu)$ est nulle.

Sous cette seconde hypothèse, l'équation (II.2) admet toute une droite de points stationnaires.

Toujours si $\det A(\mu) = 0$, au moins l'une des valeurs propres de $A(\mu)$ est nulle. L'autre valeur propre est alors donnée explicitement par la trace. Si elle est non-nulle, son signe détermine si les solutions de (II.2) convergent vers un point de la droite lorsque t tend vers $+\infty$ ou $-\infty$.

Si maintenant $D(\mu) < 0$, alors $A(\mu)$ admet deux valeurs propres complexes conjuguées (et est donc inversible). A nouveau, leur somme vaut la trace ; mais cette somme vaut aussi deux fois la partie réelle de l'une d'entre elles. Par conséquent, la stabilité de l'unique point stationnaire $S = A^{-1}(\mu)B(\mu)$ est caractérisée par le signe de la trace de $A(\mu)$ de la façon suivante :

- Pour $\operatorname{tr} A(\mu) < 0$, les solutions de (II.2) effectuent des spirales qui convergent vers S ; c'est donc un point stable, on parle en particulier de foyer attracteur.
- Pour $\operatorname{tr} A(\mu) > 0$, les solutions de (II.2) effectuent des spirales qui s'éloignent de plus en plus de S ; c'est donc un point instable, on parle aussi de foyer répulseur.
- Enfin, si $\operatorname{tr} A(\mu) = 0$, alors les parties réelles des deux valeurs propres sont nulles ; les solutions décrivent alors des cercles. Seul le point $(0, 0)$ est stationnaire, et il se révèle non-stable de façon asymptotique. On parle dans ce cas de centre.

Cela caractérise complètement les solutions en termes des fonctions $h : \mu \mapsto \det A(\mu)$ et $k : \mu \mapsto \operatorname{tr} A(\mu)$. On peut alors considérer la courbe paramétrée (h, k) dans le plan, et selon la position des points, discuter de la stabilité des solutions, comme le montre la figure II.1.

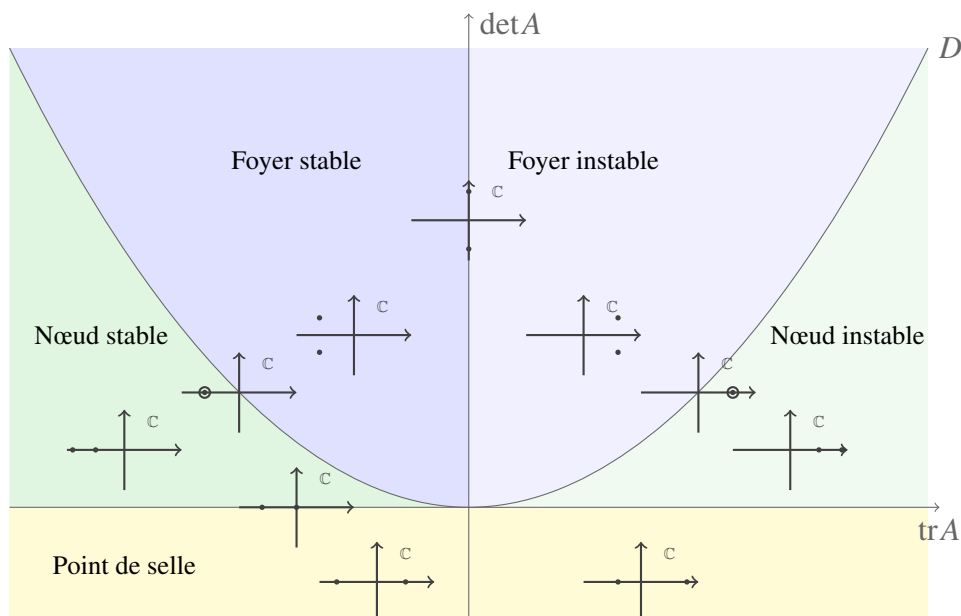


FIG. II.1 – Comportement des valeurs propres de A en fonction de son déterminant et sa trace. Les zones plus foncées représentent les endroits où l'unique point stationnaire est stable.

II.2 Equations non-linéaires avec paramètres

On retourne au problème de comprendre le comportement des points stationnaires de l'équation (II.1) lorsque μ varie. Nous cherchons donc, pour μ quelconque, d'une part, connaître les points u_μ pour lesquels

$$f(u_\mu, \mu) = 0$$

et d'autre part, en déterminer leur stabilité.

Mais, au voisinage d'un point (u_0, μ_0) où f s'annule, on peut décrire une courbe $u(\mu)$ moyennant des hypothèses sur f . Afin d'énoncer plus aisément les théorèmes qui suivent, nous supposons, sauf mention, que f est C^1 en sa variable u et continue en μ .

Décrire localement une courbe sur laquelle f s'annule, est possible grâce au théorème des fonctions implicites que nous rappelons ici sous une forme générale :

Théorème 21 (Théorème des fonctions implicites). *Soient X, Y, Z trois espaces de Banach, et $F : X \times Y \rightarrow Z$ une fonction continue définie au voisinage U d'un point $(x_0, y_0) \in X \times Y$. Supposons que $\partial_y F$ existe en tant que différentielle de Fréchet, et qu'elle soit continue et bijective en (x_0, y_0) . Alors il existe une application $h : X \rightarrow$*

Y , définie et continue sur un voisinage de x_0 , et vérifiant, pour tout $(x, y) \in U$:

$$F(x, y) = 0 \iff y = h(x).$$

Pour notre cas, en prenant $X = \mathbb{R}$, $Y = Z = \mathbb{R}^n$, on obtient le résultat suivant :

Corollaire 22. *Si $\partial_u f$ existe en tant que différentielle de Fréchet, et si elle est continue et bijective en (u_0, μ_0) , alors il existe un $r > 0$ tel que l'on puisse décrire les points stationnaires au moyen d'une fonction continue définie sur $] \mu_0 - r, \mu_0 + r [$, valant u_0 en μ_0 ; plus formellement, pour tout $\mu \in] \mu_0 - r, \mu_0 + r [$, il existe $u_\mu \in \mathbb{R}^n$, proche de u_0 , et point stationnaire de l'équation (II.1) pour le paramètre μ .*

Si la condition citée dans ce corollaire est vérifiée pour tous les points stationnaires en une valeur du paramètre μ_0 , alors on conserve le même nombre de points stationnaires lorsque le paramètre évolue à proximité de μ_0 ; le seul changement de comportement qu'il peuvent subir concerne donc leur stabilité.

Or, sous la condition du corollaire, aucune valeur propre de $\partial_u f(u_0, \mu_0)$ n'est nulle; puisqu'on suppose f de classe C^1 sur sa première variable, $\partial_u f(u, \mu)$ est continue, au moins au voisinage de (u_0, μ_0) ; les valeurs propres de $\partial_u f(u, \mu)$ sont donc assez proches de celles obtenues en (u_0, μ_0) , et, quitte à restreindre le voisinage, on peut faire en sorte qu'elles ne s'annulent pas, et donc ne changent pas de signe sur ce voisinage. Par conséquent, les informations obtenues quant à la stabilité du point stationnaire (u_0, μ_0) seront les mêmes pour les points stationnaires proches.

Là où la différentielle de f par rapport à u est bijective, connaître le comportement des points stationnaires (leur nombre ainsi que la stabilité de chacun) ne pose donc pas de difficulté majeure. L'outil principal est le théorème des fonctions implicites, qui nous permet de décrire chaque point stationnaire en fonction du paramètre. La perte de bijectivité de cette différentielle est fatale : on ne peut alors plus suivre, même localement, la variation du point stationnaire concerné. Nous allons voir ci-après que le comportement d'un tel point est alors très variable. L'analyse de ce comportement étant assez compliquée, nous allons commencer par le cas où $n = 1$. Nous énonçons et démontrons trois théorèmes qui permettent de déterminer le comportement de points stationnaires au voisinage d'un point où la dérivée selon u est non-bijection, donc nulle. En considérant le cas $n = 2$, un nouveau comportement vient s'ajouter; nous énonçons le résultat décrivant cette nouvelle variation possible.

La généralisation en dimension n quelconque se réalise grâce aux théorèmes de la variété centrale et de réduction sur cette variété (théorèmes 19 et 20); nous évoquons quelque peu comment on retrouve les bifurcations étudiées en dimension supérieure. Nous n'irons pas plus loin dans cette étude, l'augmentation dans les dimensions apportant encore d'autres comportements possibles et très variés.

Comme le note S. Wiggins, plus le nombre de valeurs propres situées sur l'axe $i\mathbb{R}$ est important, plus il y a de comportements exotiques ([15], p. 254).

II.2.1 Bifurcations locales en dimension 1

Nous supposons dans cette partie, outre le fait que f soit C^1 en u , qu'elle admette des dérivées partielles d'ordre 2 au voisinage du point que l'on considère. Sans perte de généralité, nous pouvons supposer que ce point est $(0,0)$.

Définition 13. On dit que f admet une bifurcation locale en $(0,0)$ s'il existe $\varepsilon > 0$ et $r > 0$ tels que, pour tout $\mu \in [-\varepsilon, \varepsilon]$, les flux engendrés par $f(\cdot, -\mu)$ et $f(\cdot, \mu)$ ne sont pas topologiquement équivalents ; c'est-à-dire, il existe des solutions de l'une des équations, dont le comportement n'est pas reflété parmi les solutions de l'autre équation.

La propriété suivante se déduit immédiatement des discussions ci-dessus :

Proposition 23. Si la dérivée partielle de f selon u ne s'annule pas en $(0,0)$, alors f ne subit pas de bifurcation en ce point.

Nous travaillons donc, dans toute cette partie, sous l'hypothèse

$$f(0,0) = 0; \quad \partial_u f(0,0) = 0.$$

Les énoncés des théorèmes qui suivent sont inspirés du livre de P. Glendinning [6], les preuves étant inspirées davantage par le livre de S. Wiggins [15]. Les théorèmes énoncent à chaque fois un certain nombre de points stationnaires par région et en donnent la stabilité ; nous ne prouvons dans ce qui suit l'existence de ces points.

II.2.1.1 Bifurcation selle-nœud

Théorème 24. Supposons que f ainsi que $\partial_u f$ s'annulent tous deux en $(0,0)$. Si $A := \partial_{\mu} f(0,0) \neq 0$ et $B := \partial_{uu}^2 f(0,0) \neq 0$, alors il existe une courbe de points stationnaires de l'équation $\partial u = f(\mu, u)$ au voisinage de $(0,0)$, tangente en ce point à la droite d'équation $\mu = 0$, et on a la caractérisation suivante sur le nombre de points stationnaires et leur stabilité :

- si $AB < 0$, alors, pour μ suffisamment petit,
 - pour $\mu < 0$, il n'y a aucun point stationnaire proche de $(0,0)$;
 - pour $\mu > 0$, on trouve deux points stationnaires proches de $(0,0)$, hyperboliques, le plus grand étant stable et le plus petit instable lorsque $B < 0$; et inversement si $B > 0$;

- si au contraire, $AB > 0$, alors les conclusions précédentes s'inversent : on a 2 points stationnaires pour $\mu < 0$ et aucun pour $\mu > 0$.

Démonstration. Puisque $\partial_\mu f(0,0)$ est non-nulle, en tant qu'application linéaire de \mathbb{R} dans \mathbb{R} , elle est inversible ; et le théorème des fonctions implicites s'applique sur f de la façon suivante : il existe une fonction h , définie sur un intervalle ouvert V contenant 0, telle que :

$$\text{pour tout } u \in V, \quad f(u, h(u)) = 0.$$

La fonction g définie sur V par $g(u) = f(u, h(u))$ est donc identiquement nulle ; ses dérivées le sont également. Or, on a $\partial g(u) = \partial_u f(u, h(u)) + \partial_\mu f(u, h(u)) \partial h(u)$, pour tout $u \in V$; en particulier, en 0, les hypothèses sur les dérivées partielles de f impliquent que $\partial h(0) = 0$; dès lors, la droite d'équation $\mu = 0$ est tangente au graphe $\{(u, h(u)) \mid u \in V\}$, qui représente la courbe de points stationnaires de f au voisinage de 0. De plus, on a $\partial^2 g(u) = \partial_{uu}^2 f(u, h(u)) + 2\partial_{u\mu} f(u, h(u)) \partial h(u) + \partial_\mu f(u, h(u)) \partial^2 h(u)$, l'évaluation en 0 et nos hypothèses donnent alors

$$\partial^2 h(0) = -\frac{\partial_{uu}^2 f(0,0)}{\partial_\mu f(0,0)} \neq 0.$$

Sur un voisinage V' inclus à V , la fonction h a alors une concavité fixée, déterminée par les signes de $A = \partial_\mu f(0,0)$ et $B = \partial_{uu}^2 f(0,0)$. Le point 0 est donc un extremum de h , et il s'agit d'un minimum si et seulement si les signes de A et B sont opposés. Sous cette hypothèse, on obtient l'existence de points stationnaires pour f uniquement pour des valeurs du paramètre μ positives.

Par ailleurs, ∂h peut être développé de façon limitée :

$$\partial h(u) = \partial^2 h(0) \frac{u^2}{2} + o(u);$$

et s'avère être non-nulle au voisinage de 0, en des points autres que 0. Dès lors, on peut affirmer que h est un difféomorphisme local sur deux intervalles $]0, \varepsilon[$ et $] -\varepsilon, 0[$ inclus à V' ; et l'on note u^+ et u^- les difféomorphismes inverses correspondants.

Nous avons dès lors, pour $u \in] -\varepsilon, \varepsilon[$ et $\mu \in h(] -\varepsilon, \varepsilon[)$, les équivalences :

$$f(u, \mu) = 0 \quad \iff \quad \mu = h(u) \quad \iff \quad (u = u^+(\mu) \text{ ou } u = u^-(\mu));$$

les difféomorphismes locaux u^+ et u^- décrivent donc les points stationnaires de f au voisinage de 0, en fonction du paramètre μ .

□

II.2.1.2 Bifurcation transcritique

Théorème 25. *Supposons que f ainsi que ∂f s'annulent tous deux en $(0,0)$. Supposons également que, pour tout μ au voisinage de 0, on ait $f(\mu,0) = 0$. Si $A := \partial_{uu}^2 f(0,0) \neq 0$ et $C := \partial_{u\mu}^2 f(0,0) \neq 0$, alors il existe une courbe de points stationnaires de l'équation $\partial u = f(\mu,u)$ au voisinage de $(0,0)$, qui intersecte transversalement la droite $u = 0$ à l'origine.¹ De plus, pour μ suffisamment petit mais non-nul, on trouve deux points stationnaires 0 et u_μ , dont la stabilité est caractérisée comme suit :*

- si $C < 0$, alors, pour $\mu < 0$, le point stationnaire le plus grand est instable, l'autre est stable ; et il y a échange du phénomène de stabilité pour $\mu > 0$;
- si au contraire, $C > 0$, alors la conclusion s'inverse selon le signe de μ .

Démonstration. Sur le voisinage de $(0,0)$ où l'on a $f(\mu,0) = 0$, on peut l'écrire sous la forme $f(u,\mu) = ug(u,\mu)$, en définissant

$$g(u,\mu) := \begin{cases} \frac{f(u,\mu)}{u} & \text{si } u \neq 0 \\ \partial_u f(0,\mu) & \text{si } u = 0. \end{cases} \quad (\text{II.3})$$

Dès lors, $g(0,0) = \partial_u f(0,0) = 0$. Par ailleurs, les dérivées partielles de g sont, pour u non-nul, $\partial_\mu g = \frac{\partial_\mu f}{u}$ et $\partial_u g = \frac{\partial_u f}{u} - g$; et pour $u = 0$, la définition de g implique $\partial_u g(0,\mu) = \partial_{uu}^2 f(0,\mu)$ et $\partial_\mu g(0,\mu) = \partial_{u\mu}^2 f(0,\mu)$. Or, vu l'hypothèse réalisée sur $\partial_{uu}^2 f(0,0)$, on a $\partial_\mu g(0,0)$; le théorème des fonctions implicites donne une fonction h définie sur un voisinage V de 0, pour laquelle $g(u,h(u)) = 0$ (et donc aussi $f(u,h(u)) = 0$). Des calculs identiques à ceux réalisés dans la preuve précédente montrent alors que

$$\partial_u h(0) = -\frac{\partial_u g(0,0)}{\partial_\mu g(0,0)} = -\frac{\partial_{uu}^2 f(0,0)}{\partial_{u\mu}^2 f(0,0)}$$

et nos hypothèses sur les dérivées partielles de f donnent que $\partial h(0) \neq 0$ (et est bien définie). Quitte à prendre un voisinage V plus petit, on peut supposer que $\partial_u h(u)$ est de signe constant sur V ; h se trouve être un difféomorphisme local dont l'inverse est noté u^* . Cette fonction donne alors l'autre point stationnaire pour μ au voisinage de 0. □

II.2.1.3 Bifurcation en fourche, dite « pitchfork »

Théorème 26. *Supposons que f ainsi que ∂f s'annulent tous deux en $(0,0)$. Supposons aussi que $f(0,\mu) = 0$ sur un voisinage de $(0,0)$. [utile ?] Si $\partial_{uu}^2 f(0,0) = 0$,*

¹Cela signifie que la tangente à la courbe en $(0,0)$ n'est pas la droite $u = 0$.

$B := \partial_{\mu\mu}^2 f(0,0) \neq 0$ et $C := \partial_{uuu}^3 f(0,0) \neq 0$, alors il existe deux courbes de points stationnaires de l'équation $\partial u = f(\mu, u)$ au voisinage de $(0,0)$, l'une tangente en ce point à la droite d'équation $\mu = 0$, l'autre coupant cette droite transversalement à l'origine. De plus, on a la caractérisation suivante sur le nombre de points stationnaires et leur stabilité :

- si $BC < 0$, alors, pour μ suffisamment petit,
 - pour $\mu < 0$, il y a un seul point stationnaire proche de $(0,0)$, stable si $C < 0$ et instable si $C > 0$;
 - pour $\mu > 0$, on trouve trois points stationnaires proches de $(0,0)$, les deux points les plus éloignés étant stables et celui au milieu instable lorsque $C < 0$, et inversement si $C > 0$.
- si au contraire, $BC > 0$, alors les conclusions précédentes s'inversent : on a 3 points stationnaires pour $\mu < 0$ puis un seul pour $\mu > 0$.

Démonstration. Comme pour la preuve précédente, on définit une fonction auxiliaire g , à l'aide de (II.3) ; cette fois g répond exactement aux critères énoncés au théorème de bifurcation selle-nœud ; si bien que l'on a deux courbes de points stationnaires u^+ et u^- , non-nulles, au voisinage de $(0,0)$, pour $\mu > 0$ lorsque $\partial_{\mu}g(0,0)\partial_{uu}^2g(0,0) < 0$, et pour $\mu < 0$ lorsque $\partial_{\mu}g(0,0)\partial_{uu}^2g(0,0) > 0$. Or, par définition de g , les deux valeurs apparaissant dans les produits sont respectivement $\partial_{\mu\mu}^2 f(0,0)$ et $\partial_{uuu}^3 f(0,0)$, d'où les conditions apparaissant dans le théorème. \square

II.2.2 Une bifurcation locale en dimension 2 : la bifurcation de Hopf

Il est important de noter un type de bifurcation particulier en dimension 2. Il s'agit ici d'une condition sur les valeurs propres complexes de la différentielle. On fixe toujours le point $(0,0)$ comme point stationnaire, et l'on suppose ici que $\partial_u(0,0)$ admet exactement deux valeurs propres complexes de partie réelle nulle.

Théorème 27 (Théorème de Poincaré-Andronov-Hopf). *Supposons que $0 \in \mathbb{R}^2$ soit un point stationnaire pour l'équation $\partial u = f(u, \mu)$, pour tous les μ au voisinage de 0. Supposons également que $\partial_u f(0, \mu)$ admet deux valeurs propres complexes non-réelles au voisinage de 0, dont on note $r(\mu)$ leur partie réelle. Si $r(0) = 0$ et si $\mu r'(\mu) > 0$ pour $\mu \neq 0$, alors :*

- le point stationnaire 0 est stable pour $\mu < 0$ et instable pour $\mu > 0$;
- en tout $\mu > 0$, il existe un cercle centré en 0 tel que :
 - toute solution du système dont un point est sur le cercle, a toute son image sur ce cercle ;
 - toutes les autres solutions convergent en effectuant des spirales vers ce cercle.

II.2.3 Généralisation en dimension supérieure

Les quatre types de bifurcations rencontrés ci-avant peuvent se retrouver en dimension supérieure. Etant donnée l'équation générale (II.1) :

$$\partial u = f(u, \mu),$$

l'on peut étudier les points stationnaires et leur stabilité à μ fixé. Comme déjà remarqué, l'étude de stabilité passe par l'étude des valeurs propres de la différentielle (selon u) évaluée au point stationnaire considéré. Comme dans le cas de la dimension 1, on peut se contenter d'étudier les comportements au voisinage de $(0, 0)$, considéré tel que $u = 0$ soit point stationnaire pour le paramètre $\mu = 0$

Si la différentielle de f selon u évaluée en $(0, 0)$ est bijective, le théorème des fonctions implicites nous assure de l'existence d'un point stationnaire $u(\mu)$ proche de 0, pour des valeurs de paramètre proche de 0. Supposons que ce ne soit pas le cas ; alors, cette différentielle admet 0 comme valeur propre. Si elle est de multiplicité 1, la variété centrale s'avère être aussi de dimension 1 ; faisant alors appel au théorème 20, pour autant que les autres valeurs propres soient de partie réelle strictement négative, on peut conclure sur la stabilité de certains points stationnaires.

Nous supposons donc dans la suite de cette partie, les faits suivants :

- Le point $0 \in \mathbb{R}^n$ est stationnaire pour (II.1), pour la valeur du paramètre 0.
- La différentielle $\partial_u f(0, 0)$ admet 0 comme valeur propre, avec multiplicité 1 ; on note s le nombre de valeurs propres dont la partie réelle est positive, et u le nombre de valeurs propres dont la partie réelle est négative. (On a alors $s + u + 1 = n$.)
- L'espace propre associé à la valeur propre 0 est engendré par un vecteur v .

Alors, par exemple, le théorème de bifurcation selle-nœud s'exprime de façon générale :

Théorème 28 (Théorème de bifurcation selle-nœud). *Si $A := \partial_\mu f(0, 0) \cdot v \neq 0$ et $B := \partial_{uu}^2 f(0, 0) \cdot (v, v) \neq 0$, alors il existe une courbe de points stationnaires de l'équation $\partial u = f(\mu, u)$ au voisinage de $(0, 0)$, tangente en ce point à la droite d'équation $\mu = 0$, et on a la caractérisation suivante sur le nombre de points stationnaires et leur stabilité :*

- si $AB < 0$, alors, pour μ suffisamment petit,
 - pour $\mu < 0$, il n'y a aucun point stationnaire proche de $(0, 0)$;
 - pour $\mu > 0$, on trouve deux points stationnaires proches de $(0, 0)$, hyperboliques², l'un étant stable et l'autre instable.

²Un point stationnaire est dit hyperbolique si la différentielle partielle selon u , évaluée en ce point, n'admet aucune valeur propre de partie réelle nulle.

- *si au contraire, $AB > 0$, alors les conclusions précédentes s'inversent : on a 2 points stationnaires pour $\mu < 0$ et aucun pour $\mu > 0$.*

Chapitre III

Aspects numériques des bifurcations : algorithmes de continuation

Reprenons l'équation de base :

$$\partial_t u = f(u, \mu) \quad (\text{III.1})$$

et la question principale du chapitre précédent : comment les solutions varient lorsque le paramètre μ varie ? Cette fois, l'idée n'est pas de les décrire théoriquement, mais d'avoir une bonne approximation numérique de leur comportement. La question est donc plus du type : comment « suivre » ces solutions ? C'est tout le but de ce chapitre : fournir des outils numériques permettant de suivre des solutions, et détecter d'éventuels « accidents » dus à des bifurcations.

Le but de ces algorithmes n'est pas de connaître l'évolution de toutes les solutions, mais plutôt de déterminer comment varient les points stationnaires. Comme on l'a vu dans les chapitres précédents, ils sont la clé de la compréhension du comportement des solutions. L'idée de départ est la suivante : on suppose ces points connus pour une valeur fixée μ et on s'intéresse à connaître leur évolution lorsque μ évolue.

III.1 Continuation aux ordres 0 et 1

La première idée, naïve, consiste à faire évoluer μ par petits pas. Etant donné un point stationnaire connu $u(\mu)$ et un μ' proche (disons pour fixer les idées $\mu' - \mu = \varepsilon > 0$, on s'intéresse à connaître l'évolution du point pour μ' , c'est-à-dire, connaître $u(\mu')$. Or, comme déjà remarqué au chapitre I, un point stationnaire pour

une équation autonome est une racine pour la fonction apparaissant au second membre de l'équation différentielle ; ici, le problème est donc :

Problème 1. Pour μ fixé, connaissant $u(\mu) \in \mathbb{R}^n$ vérifiant $f(u(\mu), \mu) = 0$, trouver, pour $\mu' = \mu + \varepsilon$, un point $u(\mu')$ satisfaisant

$$f(u(\mu'), \mu') = 0.$$

Ayant donc fixé μ' , on voudrait pouvoir écrire f sous la forme :

$$f(u', \mu') = f(u, \mu') + \partial_u f(u, \mu') \cdot (u' - u) + o(\|u' - u\|),$$

ce qui impose que f soit \mathcal{C}^1 en sa variable u , ce que nous supposons par la suite. Notons $g(u) := f(u, \mu')$; on cherche donc à résoudre :

$$g(u) + \partial g(u) \cdot (u' - u) = 0$$

ce qui implique la résolution du système

$$\partial g(u) \cdot (u' - u) = -g(u)$$

en l'inconnue u' . Cela suggère d'utiliser une méthode itérative ; par exemple, la méthode de Newton :

$$\begin{aligned} x_0 &:= u(\mu) \\ x_n &:= x_{n-1} + \Delta_n, \quad \text{avec } \Delta_n \text{ solution de } \partial g(x_n) \cdot \Delta = -g(x_n). \end{aligned} \tag{III.2}$$

Si la suite $(x_n)_{n \in \mathbb{N}}$ converge, on prend $u(\mu')$ comme étant la limite de cette suite ; dans le cas contraire, il sera nécessaire de prendre un μ' plus proche de μ .

Cela pose cependant quelques soucis si l'on souhaite programmer directement cette routine. Deux problèmes arrivent, il s'agit d'une part, « comment détecter la convergence de la suite (x_n) », et d'autre part, « il est nécessaire d'assurer un arrêt éventuel s'il y a divergence pour des μ' très proches de μ ; quelle condition imposer ? ». Pour le premier, remarquons que (x_n) converge si et seulement si, la suite (Δ_n) tend vers 0. Considérer de ne plus itérer dès que les pas Δ_n deviennent assez petits est donc une solution acceptable. Pour le second problème, une démarche similaire est applicable : on décide d'arrêter la routine dès que la proximité entre un μ' et μ s'avère trop importante. Il est évident que la courbe ne sera pas suivie indéfiniment ; ainsi, on se fixera une borne maximale μ^* , plus grande que μ , telle que l'algorithme finisse avec succès une fois cette borne atteinte.

Jusqu'à présent, nous avons traité indifféremment la continuation à l'ordre 0 et celle à l'ordre 1 ; la différence entre les deux algorithmes réside en un seul fait : lorsqu'une estimation u_n a été trouvée pour une valeur de μ_n , la continuation à l'ordre 0 va se contenter d'estimer u_{n+1} en initialisant (III.2) avec $x_0 := u_n$;

pour la continuation à l'ordre 1, on approxime davantage le point u_{n+1} souhaité en prenant comme estimation x_0 le point w tel que (w, μ_{n+1}) se trouve sur la tangente à la courbe $f = 0$ au point (u_n, μ_n) . Cette tangente est donnée par l'équation paramétrique $(x, y) = (u_n, \mu_n) + tv, t \in \mathbb{R}$, où v est un vecteur tangent à la courbe. Mais, puisqu'on souhaite rester sur la courbe de niveau 0 de f , si $v = (\bar{v}, \tilde{v})$ est un vecteur tangent, on doit avoir

$$\nabla f(u_{n-1}, \mu_{n-1}) \cdot v = 0, \quad (\text{III.3})$$

où ∇f est le gradient de f . Par définition du gradient, et par la décomposition de ∂f , cette égalité devient :

$$\partial_u f(u_{n-1}, \mu_{n-1}) \cdot \bar{v} + \partial_\mu f(u_{n-1}, \mu_{n-1}) \tilde{v} = 0. \quad (\text{III.4})$$

Cette équation cache un système de n équations en $n + 1$ inconnues ; par conséquent les vecteurs v solutions de l'équation précédente se trouvent dans un sous-espace vectoriel de dimension 1. Or, ce qui nous intéresse, c'est de trouver w tel que $(w - u_n, \varepsilon) = tv$ pour un certain $t \in \mathbb{R}$; en fixant $t = \varepsilon$ et $\tilde{v} = 1$, le problème revient à résoudre $\partial_u f(u_{n-1}, \mu_{n-1}) \cdot \bar{v} = -\partial_\mu f(u_{n-1}, \mu_{n-1})$, puis de prendre $w = u_n + \varepsilon \bar{v}$. C'est précisément ce qui sera réalisé par l'algorithme de continuation à l'ordre 1.

Puisque les algorithmes sont identiques à estimateur près, on peut écrire ce dernier sous forme d'un algorithme externe. Les autres entrées de l'algorithme seront : une fonction $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ et un point (u, μ) où elle s'annule ; une estimation ε qui servira de pas initial ; un critère S qui dira à quel moment l'estimation qu'on est en train d'effectuer est suffisante ; ainsi qu'une borne maximale μ^* au-delà de laquelle l'algorithme s'arrête. Nous avons également besoin d'un moyen de résoudre les systèmes de la forme $\partial_u f v = -f$ et $\partial_u f v = -\partial_\mu f$. Dans l'algorithme, nous exprimons ces moyens par les termes « solver1 » et « solver2 » respectivement.

L'algorithme se présente alors ainsi :

◆ **Algorithme CONTINUE01** ($f, (u, \mu), \varepsilon, S, \mu^*$)

$\mu' \leftarrow \mu + \varepsilon$;

if $\mu' > \mu^*$ **then**

$\mu' \leftarrow \mu^*$

end if

$u' \leftarrow \text{ESTIMATE}(u, \text{solver2})$

$u'' \leftarrow \text{NEWTON}(f(\cdot, \mu'), u', S, \text{solver1})$

if u'' est bien défini ET $S(u'')$ est vérifié **then**

print "Estimation pour $\mu' : u''$ " ;

if $\mu' = \mu^*$ **then**

```

    return (Fin de l'algorithme avec succès)
else
     $u_{\text{new}} \leftarrow u''$ ;
     $\mu_{\text{new}} \leftarrow \mu'$ ;
     $\varepsilon_{\text{new}} \leftarrow \text{AUGMENTER}(\varepsilon)$ 
end if
else
     $\varepsilon_{\text{new}} \leftarrow \text{REDUIRE}(\varepsilon)$ ;
     $u_{\text{new}} \leftarrow u$ ;
     $\mu_{\text{new}} \leftarrow \mu$ 
end if
if  $\varepsilon_{\text{new}}$  est trop petit then
    return (Fin de l'algorithme avec échec)
end if
CONTINUE01( $f, (u_{\text{new}}, \mu_{\text{new}}), \varepsilon_{\text{new}}, S, \mu^*$ )

```

L'algorithme ESTIMATE est la seule différence entre la continuation à l'ordre 0 et celle à l'ordre 1 ; elle se contente de retourner le u qu'on lui fournit à l'ordre 0, mais pour l'ordre 1, elle lui fait subir un déplacement, comme expliqué plus haut. Nous faisons appel à un algorithme effectuant la méthode de Newton, mais tout autre algorithme de recherche de racines pourrait convenir. Enfin, nous faisons appel à deux mini-algorithmes, AUGMENTER et REDUIRE, dont le rôle est de contrôler le pas selon que les approximations peuvent être trouvées. On peut par exemple décider d'augmenter ε en le doublant et le réduire en divisant par deux ; cette augmentation et cette réduction peuvent être choisis en fonction du problème que l'on souhaite aborder.

III.2 Continuation par pseudo-longueur d'arc

Les algorithmes de continuation précédents montrent leurs limites lorsque u varie beaucoup par rapport au paramètre ; ou même pire, si la courbe que l'on essaie de suivre « rebrousse chemin » en une certaine valeur μ' . Ces algorithmes suivront la courbe de points stationnaires jusqu'à proximité de μ' mais ne trouveront jamais la suite.

Pour pallier à ce manque, l'idée est de considérer la courbe, non pas comme graphe d'une fonction $u(\mu)$, mais comme résultant d'une paramétrisation $s \mapsto (u(s), \mu(s))$. Le problème consiste alors à :

Problème 2. Pour s_0 fixé, connaissant $u(s_0)$ et $\mu(s_0)$ vérifiant $f(u(s_0), \mu(s_0)) = 0$,

trouver, pour $s = s_0 + \varepsilon$, un point $u(s)$ et un réel $\mu(s)$ satisfaisant

$$f(u(s), \mu(s)) = 0.$$

Comme précédemment, on peut essayer de développer f sous la forme :

$$f(u', \mu') = f(u, \mu) + \partial f(u, \mu) \cdot \begin{pmatrix} u' - u \\ \mu' - \mu \end{pmatrix} + o(\|u' - u\| + |\mu' - \mu|),$$

ce qui impose cette fois que f soit C^1 en ses deux variables. Avec les notations du problème ci-dessus, on souhaite avoir

$$0 = f(u(s), \mu(s)) + \partial f(u(s_0), \mu(s_0)) \cdot \begin{pmatrix} u(s) - u(s_0) \\ \mu(s) - \mu(s_0) \end{pmatrix} + o(\|u(s) - u(s_0)\| + |\mu(s) - \mu(s_0)|).$$

Mais cela n'implique pas l'unicité de la solution car la matrice $\partial f(u(s_0), \mu(s_0))$ est de taille $n + 1 \times n$, ce qui crée un système de n équations à $n + 1$ inconnues. Il nous faut donc ajouter une égalité de telle sorte que le système admette une solution unique. Cette égalité va nous être donnée par le fait que l'on souhaite une paramétrisation par longueur d'arc ; dès lors, on impose la condition :

$$\|\partial_s u(s)\|^2 + (\partial_s \mu(s))^2 = 1;$$

qui peut être approximée, pour des s proches de s_0 (disons, $s - s_0 = \varepsilon$), avec la contrainte [\[Expliquer\]](#)

$$\|\partial_s u(s_0)\| \cdot (u(s) - u(s_0)) + \partial_s \mu(s_0)(\mu(s) - \mu(s_0)) - \varepsilon = 0.$$

Cette contrainte peut encore être exprimée sous la forme

$$\begin{pmatrix} \partial_s u(s_0) \\ \partial_s \mu(s_0) \end{pmatrix} \cdot \begin{pmatrix} u(s) - u(s_0) \\ \mu(s) - \mu(s_0) \end{pmatrix} = \varepsilon;$$

si bien que l'approximation « vecteur tangent de norme 1 en s » que l'on fait n'est autre que de demander que l'approximation $(u(s), \mu(s))$ soit suffisamment proche de $(u(s_0), \mu(s_0))$ dans la direction du vecteur tangent.

Dès lors, si l'on considère la matrice $A(u, \mu)$ définie par blocs comme :

$$A(u, \mu) = \begin{pmatrix} \partial_u f(u, \mu) & \partial_\mu f(u, \mu) \\ {}^t \partial_s u(s_0) & \partial_s \mu(s_0) \end{pmatrix},$$

le système défini par

$$A(u(s_0), \mu(s_0)) \begin{pmatrix} u(s) - u(s_0) \\ \mu(s) - \mu(s_0) \end{pmatrix} = \begin{pmatrix} -f(u(s_0), \mu(s_0)) \\ \varepsilon \end{pmatrix}$$

donne $n + 1$ équations à $n + 1$ inconnues. Comme précédemment, on peut résoudre numériquement ce système via la méthode de Newton :

$$\begin{aligned}
u_0 &:= u(s_0); \mu_0 := \mu(s_0) \\
\begin{pmatrix} u_{n+1} \\ \mu_{n+1} \end{pmatrix} &:= \begin{pmatrix} u_n \\ \mu_n \end{pmatrix} + \Delta_n \\
&\text{avec } \Delta_n \text{ solution de} \\
A(u_n, \mu_n) \Delta &= \begin{pmatrix} -f(u_n, \mu_n) \\ \varepsilon \end{pmatrix} \tag{III.5}
\end{aligned}$$

Cependant, résoudre directement ce système est, en général, coûteux. Il faudrait, pour chaque itération, calculer $\partial f(u_n, \mu_n)$, lui ajouter, en dernière ligne, l'équivalent des dérivées de u et μ ; puis seulement tenter de résoudre le système. Il est possible de faire mieux, en utilisant uniquement la matrice $\partial_u f$ pour résoudre le système initial. Si, en tout (u, μ) , la matrice $\partial f(u, \mu)$ possède de bonnes propriétés permettant une résolution rapide des systèmes dans laquelle elle intervient, deux résolutions rapides seront plus efficaces qu'une seule résolution du système (III.5) pour laquelle nous n'avons pas d'information.

Proposition 29. Notons $J := \partial f(u_n, \mu_n)$. Soient v et w , respectivement des solutions de $Jv = -f(u_n, \mu_n)$ et de $Jw = -\partial_\mu f(u_n, \mu_n)$. Si l'on définit :

$$z := \frac{\varepsilon - \partial_s u \cdot v}{\partial_s \mu + \partial_s u \cdot w} \tag{III.6}$$

$$\text{puis } \Delta_n := \begin{pmatrix} v + wz \\ z \end{pmatrix}, \tag{III.7}$$

alors Δ_n est solution de (III.5).

Démonstration. On montre que $J(v + wz) + \partial_\mu f(u_n, \mu_n)z = -f(u_n, \mu_n)$, puis que $\partial_s u \cdot (v + wz) + \partial_s \mu z = \varepsilon$. En développant le premier membre de la seconde égalité, on a

$$\begin{aligned}
\partial_s u \cdot (v + wz) + \partial_s \mu z &= (\partial_s u \cdot v + \partial_s \mu z) + \partial_s u \cdot wz \\
&= \varepsilon - \partial_s u \cdot v + \partial_s \mu z + \partial_s u \cdot wz \\
&= \varepsilon;
\end{aligned}$$

par ailleurs, le premier membre de la première relation donne

$$J(v + wz) + \partial_\mu f(u_n, \mu_n)z = (Jv + \partial_\mu f(u_n, \mu_n)z) + Jwz;$$

or, puisque $Jv = -f(u_n, \mu_n)$ et $Jw = -\partial_\mu f(u_n, \mu_n)$, l'égalité est vérifiée. \square

L'algorithme se présente alors comme suit : pour chaque n , on calcule $\partial_u f(u_n, \mu_n)$ et $\partial_\mu f(u_n, \mu_n)$; on résout les systèmes intermédiaires énoncés dans la proposition, on effectue les calculs nécessaires, puis on trouve u_{n+1} comme annoncé dans la méthode de Newton.

Il reste néanmoins un point à éclaircir : nous avons besoin, lorsque u et μ sont vus comme courbes paramétrées, du vecteur tangent en tout point de la courbe, c'est-à-dire des dérivées $\partial_s u$ et $\partial_s \mu$. Or, nous n'avons d'information que sur les itérées u_n ; et nous ne voulons pas, en pratique, connaître l'expression exacte de la paramétrisation (si on l'avait, le problème se réduirait à faire varier s pour obtenir tous les points !). L'idée est alors de prendre une estimation correcte de ce vecteur. Pour cela, il suffit de reprendre le raisonnement qui a conduit aux équations (III.3) et (III.4) ; la différence majeure est qu'ici, le sous-espace déduit de cette seconde équation peut être généré par un vecteur de la forme $(w, 0)$.

Mais, si c'est le cas, cela implique qu'on cherche une solution de

$$\partial_u f(u_n, \mu_n) \cdot w = 0,$$

et que la courbe se révèle être tangente en (u_n, μ_n) à l'hyperplan d'équation $\mu = 0$. Ce point s'avère donc être un point de bifurcation, et nécessite une analyse spéciale. Toutefois, en pratique, c'est un cas qui ne se produira jamais ; on peut cependant être proche de cette situation, mais nous n'allons considérer que le second cas et discuter selon la norme du vecteur w obtenu ; si elle s'avère trop importante (selon un critère qu'on se fixe à l'avance), alors on considère être proche d'un point de bifurcation et cet algorithme s'arrête.

Supposons donc que l'espace est généré par un vecteur de la forme $(w, 1)$; on peut alors résoudre le système

$$\partial_u f(u_n, \mu_n) \cdot w = -\partial_\mu f(u_n, \mu_n);$$

qui est par ailleurs la même équation que dans la proposition 29. Cela nous donne donc un vecteur tangent à la courbe d'équation $f(u, \mu) = 0$ en $(u, \mu) = (u_n, \mu_n)$. Le fait d'imposer une paramétrisation par longueur d'arc revient à trouver \tilde{v} tel que le vecteur $(\tilde{v}, \tilde{v}) = \tilde{v}(w, 1)$ soit de norme 1, et bien orienté (de telle sorte qu'on continue bien de suivre la courbe, et que l'on ne revienne pas en arrière). La première condition est facilement obtenue en divisant le vecteur $(w, 1)$ par sa norme. Pour la seconde condition, l'idée est de se baser sur les estimations précédentes, afin de déterminer dans quel sens les paramètres ont évolué. Ainsi, on prendra pour estimation de $\partial_s u$ le vecteur normalisé tel quel si l'estimation obtenue μ_{n-1} est plus grande que l'estimation précédente μ_{n-2} ; dans le cas contraire on prend son opposé pour estimation. Quant au réel $\partial_s \mu$, son signe est déterminé comme pour le vecteur $\partial_s u$ (donc, même signe que $\mu_{n-1} - \mu_{n-2}$), et sa valeur absolue vaut la norme du vecteur $(w, 1)$. Par construction, le vecteur $(\partial_s u, \partial_s \mu)$ est de norme 1.

Mais cela soulève un autre problème : nous avons à présent besoin de *deux* points initiaux pour commencer ! La résolution est cependant simple, dans le cas où seul un point nous est donné : on commence simplement avec une itération de continuation à l'ordre 0, de telle sorte à bien avoir deux points initiaux proches.

Quelles sont les conditions d'arrêt de l'algorithme ? Il faut remarquer que demander que les paramètres dépassent une certaine valeur n'est pas suffisant, puisque l'on peut à présent suivre la courbe avec une décroissance des paramètres. L'idée est donc de se fixer deux bornes, une inférieure et une supérieure ; on peut prendre, pour borne inférieure, la valeur initiale du paramètre μ_0 . On décide alors d'arrêter l'algorithme avec succès si, lors d'une estimation trouvée (u_n, μ_n) , le paramètre μ_n est plus petit que μ_0 ou plus grand qu'une borne supérieure μ^* fixée. Les entrées de l'algorithme sont donc identiques à celles pour la continuation par paramètre, à la différence près que le paramètre ε gère la distance entre les points (u_i, μ_i) au lieu de la variation du paramètre.

Les algorithmes se présentent alors comme ci-dessous. Nous commençons par l'algorithme qui donne une approximation du vecteur tangent en la courbe $f = 0$ au point (u_2, μ_2) . Nous supposons que ce point est le résultat de la dernière itération, et que l'avant-dernière nous ait donné un point (u_1, μ_1) .

◆ **Algorithme** VECTEURTANGENT($\text{solver2}, (u_1, \mu_1), (u_2, \mu_2)$)

Résoudre l'équation $\partial_u f(u_2, \mu_2) \cdot w = \partial_\mu f(u_2, \mu_2)$ à l'aide de solver2 ;

$d \leftarrow \sqrt{1 + \|w\|^2}$;

if d est trop grand **then**

return (arrêt avec échec)

else

$e \leftarrow \mu_2 - \mu_1$;

$v \leftarrow \frac{e}{d} w$;

return $(v, \frac{e}{d})$

end if

Cela nous donne alors une approximation de $(\partial_s u, \partial_s \mu)$; dès lors, on peut estimer un point un peu plus loin sur la courbe ; grâce à la proposition 29, l'algorithme suivant permet de trouver une racine, pour ε petit :

◆ **Algorithme** TROUVERACINE($((u', \mu'), (\partial_s u, \partial_s \mu)S, \varepsilon, \text{solver1}, \text{solver2})$)

$u_0 \leftarrow u'$; $\mu_0 \leftarrow \mu'$

$n \leftarrow 0$;

repeat

 Résoudre les systèmes de la proposition 29 au point (u_n, μ_n) en utilisant solver1 et solver2 .

$v \leftarrow$ solution du premier système ;

$w \leftarrow$ solution du second système ;

```

 $z \leftarrow$  Expression (III.6);
 $\Delta_n \leftarrow$  Expression (III.7);
 $(u_{n+1}, \mu_{n+1}) \leftarrow (u_n, \mu_n) + \Delta_n$ ;
 $n \leftarrow n + 1$ 
until ( $\Delta_n$  soit assez proche de 0 ET  $S(u_n, \mu_n)$ ) OU  $\Delta_n$  trop loin de 0;
if  $\Delta_n$  trop loin de 0 then
    return (Fin de l'algorithme avec échec)
else
    return  $(u_n, \mu_n)$ 
end if

```

Nous avons alors toutes les clés en main pour effectuer la continuation par pseudo-longueur d'arc. L'algorithme final se présente ainsi :

```

◆ Algorithme CONTINUEPSEUDOARCLENGTH( $f, (u, \mu), (u', \mu'), \varepsilon, S, \mu_0, \mu^*$ )
 $t \leftarrow$  VECTEURTANGENT(solver2,  $(u, \mu), (u', \mu')$ )
if  $t$  n'est pas défini then
    return (Point de bifurcation proche — arrêt de l'algorithme avec échec)
else
     $(u'', \mu'') \leftarrow$  TROUVERACINE( $(u', \mu'), t, S, \varepsilon$ , solver1, solver2);
    if  $(u'', \mu'')$  bien défini then
        if  $\mu''$  dépasse les bornes  $\mu_0$  et  $\mu^*$  then
            return (Arrêt de l'algorithme avec succès)
        else
             $X_1 \leftarrow (u', \mu'); X_2 \leftarrow (u'', \mu'')$ ;
             $\varepsilon \leftarrow$  AUGMENTER( $\varepsilon$ );
        end if
    else
         $X_1 \leftarrow (u, \mu); X_2 \leftarrow (u', \mu')$ ;
         $\varepsilon \leftarrow$  REDUIRE( $\varepsilon$ );
    end if
    if  $\varepsilon$  trop petit then
        return (Arrêt de l'algorithme avec échec)
    else
        CONTINUEPSEUDOARCLENGTH( $f, X_1, X_2, \varepsilon, S, \mu_0, \mu^*$ )
    end if
end if

```

III.2.1 Continuation par longueur d'arc avec paramètre de contrôle

Il se peut que le problème de continuation soit « mal posé » pour l'algorithme précédent ; c'est notamment le cas lorsqu'une légère évolution du paramètre fait varier fortement les points stationnaires. Afin de remédier à ce problème, on pourrait définir une nouvelle fonction, disons g , de telle sorte que

$$g(\mu, u) = f\left(\frac{\mu}{\theta}, u\right)$$

avec θ assez grand de telle sorte que la variation de $\frac{\mu}{\theta}$ soit assez faible ; et appliquer l'algorithme ci-avant décrit. Cependant, à chaque fois qu'on a besoin d'évaluer la fonction g , on devra passer par f ; de même, il faudrait exprimer les différentielles de g en termes de celles de f . Toutefois, on va montrer que le raisonnement ayant abouti à l'algorithme de continuation peut être modifié de telle sorte qu'il intègre cette nouvelle constante θ souhaitée.

Par définition de g , ses différentielles s'expriment en termes de celles de f par :

$$\begin{aligned}\partial_u g(\mu, u) &= \partial_u f\left(\frac{\mu}{\theta}, u\right) \\ \partial_\mu g(\mu, u) &= \frac{1}{\theta} \partial_\mu f\left(\frac{\mu}{\theta}, u\right).\end{aligned}$$

Le point clé de l'algorithme réside dans la résolution des systèmes

$$\partial_u g(\mu, u) y_1 = -g(\mu, u) \quad \text{et} \quad \partial_u g(\mu, u) y_2 = \partial_\mu g(\mu, u);$$

mais cela revient exactement à résoudre les systèmes liés à f :

$$\begin{aligned}\partial_u f\left(\frac{\mu}{\theta}, u\right) v_1 &= -f\left(\frac{\mu}{\theta}, u\right) \\ \partial_u f\left(\frac{\mu}{\theta}, u\right) v_2 &= \partial_\mu f\left(\frac{\mu}{\theta}, u\right)\end{aligned}$$

puis prendre $y_1 = v_1$ et $y_2 = \frac{v_2}{\theta}$. Comme déjà vu dans les précédents algorithmes, le vecteur $(y_2, 1)$ que l'on peut construire, est un vecteur tangent à la courbe $g = 0$ en (u, μ) . Sa norme vaut, en termes de v_2 et θ :

$$\|(y_2, 1)\| = \sqrt{\frac{\|v_2\|^2}{\theta^2} + 1} = \frac{1}{\theta} \sqrt{\|v_2\|^2 + \theta^2}$$

si bien que la construction du vecteur $(\partial_s u, \partial_s \mu)$ tangent à la courbe $g = 0$ en un point (u, μ) ne nécessite aucun calcul sur la fonction g elle-même. Le signe à choisir pour le sens correct du vecteur est déterminé comme précédemment.

L'algorithme présente donc peu de modifications par rapport à l'algorithme de continuation ; cependant, il permet un meilleur contrôle de la différence de variation entre le paramètre et les points estimés.

III.2.2 Un paramètre de contrôle supplémentaire

Comme nous l'avons déjà remarqué, les problèmes numériques surviennent surtout lorsque les points stationnaires évoluent rapidement par rapport au paramètre. Un contrôle supplémentaire peut être imposé si l'on travaille en dimension $n \geq 2$. Étant donnés deux itérations successives u_{n-1} et u_n d'un algorithme de continuation, on peut regarder l'angle que font les vecteurs obtenus. Une façon simple d'en obtenir une idée est de considérer le quotient

$$\tau_n := \frac{u_{n-1} \cdot u_n}{\|u_{n-1}\| \|u_n\|}$$

lorsque les deux itérations considérées sont non-nulles. le nombre τ_n représente le cosinus de l'angle que font les deux vecteurs u_{n-1} et u_n ; par conséquent, plus il est proche de 1, moins la variation (en angle) des deux itérées considérées est forte ; a contrario, si τ_n est proche de 0, alors les deux vecteurs sont « presque orthogonaux », cela suggère une forte variation qu'il faut étudier plus profondément. Ainsi, on peut, par exemple, refuser l'estimation u_n ainsi produite si l'on considère que l'angle qu'elle fait avec la précédente, est trop grand (c'est-à-dire, si le cosinus est trop petit) ; et exiger alors de recommencer l'estimation avec un pas plus petit. A nouveau, nous n'écrivons pas l'algorithme, qui présente peu de modifications.

Chapitre IV

Membranes élastiques : le problème de l'elastica

Le problème de l'elastica a été posé au XVII^e siècle. Après un aperçu historique, on formule le problème de façon rigoureuse en termes de courbes paramétrées par leur longueur d'arc. On voit alors apparaître la courbure comme élément essentiel de l'étude des courbes minimisantes, on l'étudie plus en profondeur par la suite. Cette étude permet de donner une expression explicite à la courbure, ce qui permet d'en déduire les courbes répondant au problème. Nous parlons en fin de chapitre d'une généralisation naturelle de ce problème, consistant à déposer une membrane élastique sur un substrat mou ; nous évoquons principalement le cas de l'eau.

Pour des raisons techniques, nous utilisons la notation f' pour désigner la dérivée d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$; la notation ∂ sera alors utilisée, sans ambiguïté, sur des fonctionnelles. Nous nous inspirons de l'article de R. Levien [9] pour l'historique, de l'article de G. Brunnett [2] pour la formulation et la généralisation du problème, et des articles de D. A. Singer [14] et de L. Pocivavsek [12] pour les généralisations.

IV.1 Historique du problème

On peut attribuer la première formulation rigoureuse du problème de l'elastica à James Bernoulli, en 1691. Partant d'une barre verticale, maintenue à sa base, il attache une masse en son sommet de telle sorte qu'elle se plie, et ce, suffisamment pour que la tangente au sommet soit perpendiculaire à la tangente à la base.

Le problème qu'il pose est le suivant : étant donné cette configuration, comment se comporte la barre ? Cela induit deux questions, qui sont, d'une part, trouver les équations régissant la position de la barre, et d'autre part, résoudre ces

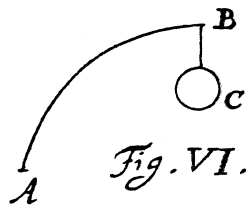


FIG. IV.1 – Dessin original de Bernoulli, réalisé en 1691.

équations pour obtenir sa forme.

Un an plus tard, il obtient une résolution partielle du problème, obtenant les équations. Mais ce n'est qu'en 1694 qu'il publie une première solution. Certains émettent des critiques dès cette publication. Huygens notamment, fait remarquer que Bernoulli n'a traité qu'un cas particulier d'elastica. En effet, rien n'interdit, a priori, de fixer une masse autre que celle amenant la barre arquée en angle droit ! Bernoulli accepte cette critique ; et trouve rapidement une formulation plus générale de son équation. Cette publication semble toutefois oubliée par la suite.

Il faudra attendre environ 50 ans pour que le problème de l'elastica intéresse à nouveau. Cette fois, c'est le neveu de James Bernoulli, prénommé Daniel, qui se penche sur la question. En 1742, il propose la technique du calcul des variations pour l'attaquer ; mais lui, n'y arrive pas. Il s'adresse alors à Leonhard Euler pour lui soumettre son idée. Ce dernier arrive, dans l'année qui suit, à une résolution complète qui lui permet aussi de classer les courbes obtenues. Cette classification, que nous étudierons ci-après, a été réalisée à partir des équations exprimées sous forme cartésienne ; cependant, une formulation plus simple est possible lorsque l'on considère les courbes paramétrées par longueur d'arc. Dans la suite, nous utiliserons une telle paramétrisation. *[Naissance des fonctions elliptiques !]*

IV.2 Formulation précise du problème

Puisqu'il s'agit de courbes planes, il suffit de regarder les paramétrisations à valeurs dans \mathbb{R}^2 . On fixe l'intervalle de paramétrisation $[0, L]$ pour $L > 0$, ainsi que deux points P et Q dans \mathbb{R}^2 (servant de conditions au bord pour la courbe) et deux vecteurs V et W , qu'on peut supposer de norme 1, eux aussi dans \mathbb{R}^2 (afin de prescrire les vecteurs tangents aux extrémités). Nous supposons que les paramétrisations choisies ont des vecteurs tangents en tout point ; si γ est l'une de ces paramétrisations, cela impose que, pour tout $x \in [0, L]$, $\gamma'(x) \neq 0$. Dès lors, on peut montrer qu'on ne perd pas en généralité en supposant que $\gamma'(x)$ est de norme 1 en tout x ; cela revient à considérer uniquement les paramétrisations par

longueur d'arc. Cela implique aussi que la longueur de la courbe est la longueur de l'intervalle de définition, qui est L .

Le problème se pose donc sur l'espace

$$\Gamma := \{\gamma \in \mathcal{C}^2([0, L], \mathbb{R}^2) \mid \gamma(0) = P, \gamma(L) = Q, \gamma'(0) = V, \gamma'(L) = W, \\ \text{et } \forall s \in [0, L], \|\gamma'(s)\| = 1\}$$

et se définit comme suit :

Problème 3. *Minimiser la fonctionnelle $\int_0^L (\kappa(s))^2 ds$, sur l'espace Γ , où $\kappa(s)$ est la courbure (signée) de la fonction γ choisie.*

En tout s , $\kappa(s)$ est, en valeur absolue, égal à l'inverse du rayon de courbure. Puisqu'on travaille en paramétrisation normale, la courbure signée est aussi égale en valeur absolue à $\|\gamma''\|$. Le signe de $\kappa(s)$ est donné par le sens de variation des vecteurs tangents aux points de la courbe au voisinage de s ; $\kappa(s)$ sera positif si et seulement si ces vecteurs tangents font en sorte que le cercle approximant la courbe (à l'ordre 2) doit être parcouru dans le sens anti-horlogique.

Le fait que l'on impose, sur Γ , que les vecteurs tangents soient de norme 1, permet de caractériser ces derniers de manière univoque en termes de l'angle qu'ils font avec un vecteur de référence; par exemple, le vecteur $(1, 0)$. Ainsi, si $\gamma \in \Gamma$, en tout $s \in [0, L]$, on peut définir $\Psi(s)$ de telle sorte que $\gamma'(s) = (\cos \Psi(s), \sin \Psi(s))$. Formellement, la construction de Ψ est fournie par la proposition 39 de l'annexe B. En dérivant cette expression, on obtient aisément que

$$\gamma''(s) = \Psi'(s)(-\sin \Psi(s), \cos \Psi(s)), \quad \text{d'où } \|\gamma''(s)\| = |\kappa(s)| = |\Psi'(s)|.$$

La courbure signée de γ est donc aussi égale à Ψ' ; en effet, quand Ψ' est positif, l'angle que forme les vecteurs tangents croît

Les conditions au bord sur les vecteurs tangents s'expriment facilement en termes de Ψ (cela revient à poser des conditions au bord sur Ψ). En réalité, Ψ se voit naturellement imposer des conditions au bord, toujours par la même proposition de l'annexe B. On obtient $\Psi(0)$ comme mesure principale de l'angle entre le vecteur V et le vecteur $(1, 0)$, mais $\Psi(L)$ peut ne pas être la mesure principale de l'angle que forme W avec $(1, 0)$; cela est dû au fait que la courbe effectue une ou plusieurs boucles. Les conditions sur les vecteurs tangents s'expriment donc par deux réels Ψ_0 et Ψ_L tels que le vecteur $(\cos \Psi_0, \sin \Psi_0)$ soit égal à V et $(\cos \Psi_L, \sin \Psi_L)$ soit égal à W . Les conditions au bord sur la courbe elle-même s'expriment via l'égalité $Q = P + \int_0^L \gamma'(s) ds$, ce qui permet de les formuler également en termes de Ψ , en faisant apparaître une intégrale.

Le problème initial se ramène donc sur l'espace

$$X_1 := \{\Psi \in \mathcal{C}^1([0, L]) \mid \Psi(0) = \Psi_0, \Psi(L) = \Psi_L\}.$$

et s'énonce comme suit :

Problème 4. *Minimiser, sur l'espace X_1 , la fonctionnelle \mathcal{E} définie par*

$$\mathcal{E}(\Psi) = \int_0^L (\Psi'(s))^2 ds \quad (\text{IV.1})$$

sous la contrainte que Ψ vérifie :

$$Q = P + \int_0^L \begin{pmatrix} \cos \Psi(s) \\ \sin \Psi(s) \end{pmatrix} ds. \quad (\text{IV.2})$$

Appliquer une même translation sur P et Q ne modifie en rien la contrainte ; sans perte de généralité, on peut supposer que $P = (0, 0)$. De plus, ajouter une même quantité ψ aux angles de bord Ψ_0 et Ψ_L , puis appliquer la rotation d'angle ψ sur Q , donne un problème identique ; on peut donc faire en sorte que Q se retrouve sur la demi-droite des points à abscisse positive et ordonnée nulle. Cependant, n'oublions pas que P , choisi comme l'origine, et Q , sont les points extrémaux d'une courbe dont la longueur n'excède pas L ; donc l'abscisse de Q peut être exprimée sous la forme $L(1 - \delta)$, pour $\delta \in [0, 1]$.

Nous sommes donc amenés à minimiser \mathcal{E} sous les contraintes

$$\int_0^L \cos \Psi(s) ds = L(1 - \delta) \quad \text{et} \quad \int_0^L \sin \Psi(s) ds = 0. \quad (\text{IV.3})$$

Afin de pouvoir appliquer le principe des multiplicateurs de Lagrange, nous allons changer notre espace $C^1([0, L])$ pour un espace plus gros, de telle sorte que l'on puisse disposer du produit scalaire usuel sur les fonctions :

$$\langle f, g \rangle := \int_0^L f(x)g(x) d\mu(x)$$

C'est dans les espaces de Sobolev (et en particulier, dans H^1), que nous allons trouver le bon espace. Pour un intervalle I ouvert (ici, nous allons prendre $I =]0, L[$), $H^1(I)$ se définit comme l'espace des fonctions f de $L^2(I)$, telles que l'on ait la propriété suivante. Soit g une fonction de classe C^∞ à support compact ; on peut alors définir $\int_0^L f(x)g'(x) dx$. Si f est dérivable, vu la condition sur g , cette intégrale est égale à $-\int_0^L f'(x)g(x) dx$. Si f n'est pas dérivable au sens classique, on peut néanmoins de cette manière lui définir une dérivée au sens des distributions ; c'est-à-dire, poser f' comme étant la distribution, qui, à g fonction C^∞ à support compact, associe le réel $-\int_0^L f(x)g'(x) dx$. Puisque $L^2(I)$ est un espace de Hilbert, on peut imposer que f' soit un élément de $L^2(I)$. C'est précisément ce que l'on impose à f pour être élément de $H^1(I)$.

Un théorème important, que nous admettrons, permet la transposition du problème sans grandes difficultés :

Théorème 30. Soit $f \in H^1(I)$. Il existe une fonction g continue sur I et telle que $f = g$ au sens de $L^2(I)$.

Le problème se pose alors sur l'espace

$$X := \{\Psi \in H^1(]0, L[) \mid \Psi(0) = \Psi_0, \Psi(L) = \Psi_L\}.$$

Grâce au théorème précédent, les conditions au bord peuvent être correctement définies. On peut alors exprimer le problème sous cette forme :

Problème 5. Minimiser, sur l'espace X , la fonctionnelle \mathcal{E} définie par (IV.1), sous les contraintes que Ψ vérifie (IV.3).

On peut prouver que le problème admet au moins une solution, grâce à la méthode directe du calcul des variations ; par ailleurs, on peut montrer que les contraintes forment une fonction dont la différentielle, évaluée en les extrema sur la contrainte, est surjective.

Le principe des multiplicateurs de Lagrange peut alors s'appliquer à notre problème ; le résoudre revient alors à trouver les points critiques de la fonctionnelle

$$\int_0^L (\Psi'(s))^2 + \lambda \cos(\Psi(s)) + \mu \sin(\Psi(s)) - \lambda(1 - \delta) ds$$

sur l'espace $X \times \mathbb{R} \times \mathbb{R}$. On peut voir (λ, μ) comme un point de \mathbb{R}^2 , si bien qu'il peut être considéré en coordonnées polaires. En écrivant $\lambda = 2a \cos \phi$ et $\mu = 2a \sin \phi$, l'on reconnaît dans le deuxième et le troisième termes, une formule trigonométrique ; dès lors la fonctionnelle précédente se récrit :

$$\mathcal{F}(\Psi, a, \phi) := \int_0^L (\Psi'(s))^2 + 2a \cos(\Psi(s) - \phi) - 2a(1 - \delta) \cos \phi ds \quad (\text{IV.4})$$

Pour étudier les points critiques de \mathcal{F} , nous devons en étudier la différentielle en tout $(\Psi, a, \phi) \in X \times \mathbb{R}^+ \times \mathbb{R}$. Les résultats des dérivations partielles par rapport à a et ϕ étant connues (ce sont les contraintes fixées en (IV.2)), il nous suffit de regarder la différentielle partielle de \mathcal{F} suivant $\Psi \in X$. Fixons donc une telle fonction, et considérons Φ un vecteur tangent de X . On montre alors que Φ se trouve dans $H_0^1(]0, L[)$, l'espace des fonctions de $H^1(]0, L[)$ qui s'annulent en 0 et en L . Pour ce faire, remarquons que X est un espace affiné d'espace vectoriel sous-jacent $H_0^1(]0, L[)$. En effet, si l'on note

$$T(s) := \Psi_0 + \frac{s}{L}(\Psi_L - \Psi_0)$$

on a l'égalité

$$X = T + H_0^1(]0, L[);$$

autrement dit, toute fonction de X peut être vue comme somme d'une fonction s'annulant aux bords avec la fonction T ; réciproquement, il est clair que si l'on ajoute T à une fonction $\Phi \in H_0^1(]0, L[)$, alors la fonction $T + \Phi$ vaut Ψ_0 en 0 et Ψ_L en L ; c'est donc un élément de X . Dès lors, nous obtenons que X est affiné dans $H^1(]0, L[)$ et que les vecteurs tangents en tout $\Psi \in X$ sont des fonctions de $H_0^1(]0, L[)$.

Pour toutes fonctions $\Psi \in X$ et $\Phi \in H_0^1$, la différentielle (partielle sur X) de \mathcal{F} en Ψ , suivant la direction Φ , vaut :

$$\begin{aligned}
& \partial_{\Psi} \mathcal{F}(\Psi, a, \phi) \cdot \Phi \\
&= \lim_{t \rightarrow 0} \frac{\mathcal{F}(\Psi + t\Phi, a, \phi) - \mathcal{F}(\Psi, a, \phi)}{t} \\
&= \lim_{t \rightarrow 0} \frac{1}{t} \left(\int_0^L (\Psi'(s) + t\Phi'(s))^2 - (\Psi'(s))^2 \right. \\
&\quad \left. - 2a(\cos(\Psi(s) + t\Phi(s) - \phi) - \cos(\Psi(s) - \phi)) ds \right) \\
&= \lim_{t \rightarrow 0} \int_0^L \frac{1}{t} (2t\Psi'(s)\Phi'(s) - t^2(\Phi'(s))^2) \\
&\quad - \frac{2a}{t} (\cos(\Psi(s) + t\Phi(s) - \phi) - \cos(\Psi(s) - \phi)) ds \\
&= \lim_{t \rightarrow 0} \int_0^L 2\Psi'(s)\Phi'(s) - t(\Phi'(s))^2 \\
&\quad - \frac{2a}{t} (\cos(\Psi(s) + t\Phi(s) - \phi) - \cos(\Psi(s) - \phi)) ds \\
&= \int_0^L 2\Psi'(s)\Phi'(s) - \frac{2}{a} \sin(\Psi(s) - \phi)\Phi(s) ds.
\end{aligned}$$

En utilisant une intégration par parties sur le premier terme de l'intégrale, et en remarquant que $\Psi'(s)\Phi(s)$ s'annule en 0 et en L , on obtient que

$$\partial_{\Psi} \mathcal{F}(\Psi, a, \phi) \cdot \Phi = \int_0^L -2\Psi''(s)\Phi(s) - 2a \sin(\Psi(s) - \phi)\Phi(s) ds.$$

où Ψ'' est à prendre au sens des distributions, comme nous l'avons fait pour pouvoir définir H^1 .

A paramètres a et ϕ fixés, les points critiques Ψ de $\mathcal{F}(\cdot, a, \phi)$ sont exactement ceux qui annulent $\partial_{\Psi} \mathcal{F}(\Psi, a, \phi) \cdot \Phi$, quelle que soit la direction $\Phi \in H_0^1$ choisie. En particulier, si l'on considère, en tout $s \in]0, L[$, des fonctions $\Phi_{s,k}$ de classe C^∞ sur $[0, L]$ (pour $k \in \mathbb{N} \setminus \{0\}$) d'intégrale égale à 1, et de support contenant s de plus en plus petit lorsque k grandit (par exemple, pour tout $k > 0$, on prend un support inclus à $I_{s,k} :=]x - \frac{1}{k}, x + \frac{1}{k}[\cap]0, L[$), on doit avoir, pour tout $s \in]0, L[$ et pour tout

$k > 0$, l'égalité

$$\int_{I_{s,k}} -2\Psi''(s)\Phi_{s,k}(s) - 2a \sin(\Psi(s) - \phi)\Phi_{s,k}(s) ds = 0$$

On peut montrer que l'on obtient alors

$$2\Psi''(s) = -2a \sin(\Psi(s) - \phi), \quad (\text{IV.5})$$

pour tout $s \in [0, L]$. *[Cela demande $\Psi \in L^p$, ce qu'on ne sait pas.]* Cette égalité implique alors que Ψ'' peut être considérée comme fonction continue, et même de classe C^1 , car composée de \sin qui est C^∞ avec Ψ qui peut être choisie de classe C^1 . Dès lors, on peut même choisir Ψ de classe C^3 . L'égalité donne alors que Ψ'' est aussi de classe C^3 ; dès lors Ψ peut être choisie 5 fois continûment différentiable. Par récurrence, on montre facilement que l'on peut en fait choisir $\Psi \in C^\infty$.

Si l'on multiplie les deux membres de l'égalité (IV.5) par $\Psi'(s)$, on reconnaît une égalité entre deux fonctions facilement primitivables, donc leurs primitives seront égales à une constante près; soit

$$(\Psi'(s))^2 = 2a \cos(\Psi(s) - \phi) - A, \quad A \in \mathbb{R}. \quad (\text{IV.6})$$

Tous les points critiques de \mathcal{F} vérifient donc cette condition; par ailleurs si Ψ est une fonction vérifiant (IV.6), alors elle vérifie aussi l'égalité (IV.5), et l'on en déduit

$$\partial_\Psi \mathcal{F}(\Psi, a, \phi) \cdot \Phi = \int_0^L 2a \sin(\Psi(s) - \phi)\Phi(s) - 2a \sin(\Psi(s) - \phi)\Phi(s) ds = 0.$$

Donc, un triplet (Ψ, a, ϕ) est point critique de \mathcal{F} si et seulement si la condition (IV.6) est vérifiée.

Les minima locaux pour \mathcal{E} se trouvent parmi ces triplets; plus exactement, si Ψ est un minimum pour \mathcal{E} , alors il existe a et ϕ , deux réels, pour lesquels la condition (IV.6) est satisfaite.

Mais, si l'on a un tel triplet (Ψ, a, ϕ) , alors on a l'égalité

$$\mathcal{E}(\Psi) = \int_0^L 2a \cos(\Psi(s) - \phi) - A ds = \int_0^L 2a \cos \Psi(s) \cos \phi + 2a \sin \Psi(s) \sin \phi - A ds;$$

si Ψ se trouve sur la contrainte définie par (IV.3), le deuxième terme de l'intégrand s'annule; quant au premier, il vaut $2a(1 - \delta) \cos \phi$. Donc, Ψ est un minimum local de \mathcal{E} sous les contraintes quand A vaut précisément $2a(1 - \delta) \cos \phi$, auquel cas $\mathcal{E}(\Phi) = 0$. Dès lors, fixer l'un des paramètres parmi A et δ , détermine univoquement l'autre. Pour des raisons techniques, nous choisissons comme paramètre $-A$.

La condition (??) va nous permettre de définir les elastica :

Définition 14. On appelle *elastica de paramètre de tension* $A \in \mathbb{R}$, toute courbe $\gamma \in \Gamma$ dont la courbure (signée) κ vérifie, pour certains $a \geq 0$ et $\phi \in \mathbb{R}$, et pour tout $s \in [0, L]$:

$$\kappa^2(s) = 2a \cos(\Psi(s) - \phi) + A \quad \text{où } \Psi(s) = \int_0^s \kappa(t) dt + \Psi_0.$$

IV.3 Etude de la courbure d'un elastica

IV.3.1 Courbure d'elastica et équation différentielle

Théorème 31. Soient $\kappa \in C^2(\mathbb{R})$ et $A \in \mathbb{R}$. Si la condition suivante sur κ est vérifiée :

il existe $a \geq 0$ et $\phi \in \mathbb{R}$ pour lesquels

$$\kappa^2(s) = 2a \cos(\Psi(s) - \phi) + A \quad (\text{IV.7})$$

(où $\Psi(s) = \int_0^s \kappa(t) dt + \Psi_0$), alors κ est solution de l'équation différentielle

$$\kappa'' = -\frac{1}{2}\kappa(\kappa^2 - A). \quad (\text{IV.8})$$

Réciproquement, si κ est une solution de (IV.8), alors κ vérifie la condition (IV.7).

Grâce à ce théorème, on peut considérer la courbure comme fonction définie sur tout \mathbb{R} .

Démonstration. On prouve d'abord le sens direct, plus simple. En dérivant la relation (IV.7), on obtient $2\kappa\kappa' = -2a\Psi' \sin(\Psi - \phi)$; sur l'ouvert U où $\kappa = \Psi'$ ne s'annule pas, on a $\kappa' = -a \sin(\Psi - \phi)$. En dérivant à nouveau, cela donne $\kappa'' = -a\Psi' \cos(\Psi - \phi)$, soit la relation attendue en utilisant (IV.7). Par continuité des fonctions entant en jeu, l'égalité se prolonge sur la fermeture de U . Si cette fermeture n'est pas \mathbb{R} , il existe alors un intervalle $I \subset \mathbb{R}$, ouvert, sur lequel $\kappa = \Psi'$ s'annule. On remarque alors que κ'' s'annule aussi sur cet intervalle; dès lors, l'égalité (IV.8) se réduit à l'identité triviale $0 = 0$ sur I .

Prouvons à présent la réciproque. Si l'on multiplie les deux membres de l'équation (IV.8) par $2\kappa'$, on obtient l'équation

$$2\kappa'\kappa'' = -\kappa'\kappa^3 + A\kappa'\kappa.$$

La différence des fonctions $(\kappa')^2$ et $-\frac{1}{4}\kappa^4 + \frac{A}{2}\kappa^2$, primitives de chacun des deux membres, est donc égale à une constante que nous appellerons, pour des raisons de calcul, $B - \frac{1}{4}A^2$. Dès lors, on obtient l'égalité

$$(\kappa')^2 + \frac{(\kappa^2 - A)^2}{4} = B. \quad (\text{IV.9})$$

Puisque κ est supposée être une solution réelle, le membre de gauche est toujours positif ou nul ; ce sera aussi le cas pour B . On déduit alors de (IV.9) que, pour tout s , le point $(\frac{\kappa^2(s)}{2}, \kappa'(s))$ se trouve sur un cercle centré en $(\frac{A}{2}, 0)$ et de rayon \sqrt{B} .

Si $B = 0$, on déduit que κ' est nulle partout, et que κ^2 est constante égale à A . En prenant $a = \phi = 0$, la condition (IV.7) est vérifiée.

Supposons maintenant $B > 0$, et posons $a := \sqrt{B}$. On peut alors paramétriser les points $(\frac{1}{2}\kappa^2(s), \kappa'(s))$ par une fonction f de classe C^1 telle que

$$\kappa'(s) = -a \sin f(s) \quad \text{et} \quad \kappa^2(s) = 2a \cos f(s) + A. \quad (\text{IV.10})$$

(le résultat technique permettant de définir f est la proposition 39 de l'annexe B.) En dérivant la seconde égalité, on a $2\kappa\kappa' = -2a \sin f f'$; c'est-à-dire, en simplifiant les deux membres par $-2a$:

$$\kappa \sin f = f' \sin f$$

On montre à présent que $\kappa = f'$. Cela terminera la preuve, car, au vu de l'égalité $\kappa = \Psi'$, les deux fonctions f et Ψ seront égales à une constante ϕ près.

L'égalité que l'on souhaite est aisément obtenue sur l'ensemble ouvert où $\sin f$ ne s'annule pas (ce qui est équivalent à demander que κ' ne s'annule pas). Par continuité, l'égalité est obtenue aussi sur la fermeture de cet ensemble. Si cette fermeture n'est pas \mathbb{R} , alors κ' s'annule sur un intervalle ouvert I ; dans ce cas, f est un multiple fixé de π sur I , et κ^2 est constante sur cet intervalle (et vaut $A + 2a$ ou $A - 2a$, le signe du deuxième terme dépendant de f). Mais si κ' s'annule sur I , il en est de même pour sa dérivée ; si bien que κ est solution de (IV.8) si et seulement si

$$0 = -\frac{1}{2}\kappa(\pm 2a)$$

donc, κ ne peut qu'être la solution nulle. Puisque f est supposée constante, l'égalité $\kappa = f'$ est vérifiée. \square

Etudier les courbures des elastica revient donc à étudier les solutions de l'équation différentielle (IV.8). Pour ce faire, nous allons regarder le plan de phases associé à cette équation.

IV.3.2 Plan de phases pour la courbure

Le diagramme de phases de l'équation (IV.8) est donné par les courbes de niveau de la fonction F définie par

$$F(\kappa, \kappa') = (\kappa')^2 + \frac{(\kappa^2 - A)^2}{4}; \quad (\text{IV.11})$$

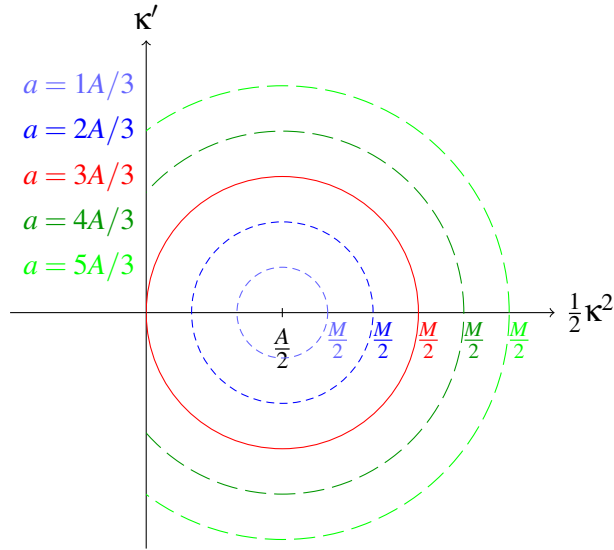


FIG. IV.2 – Courbes liant $\frac{\kappa^2}{2}$ et κ' pour différentes valeurs de a .

c'est équivalent à faire varier B , le membre de droite de l'équation (IV.9). Comme déjà discuté dans la preuve précédente, si $B < 0$, aucun point (κ, κ') ne se trouve dans la courbe de niveau B . On doit donc considérer les courbes de niveau pour $B \geq 0$. Pour $B = 0$, on ne retrouve que deux points dans la courbe de niveau : en \sqrt{A} et en $-\sqrt{A}$. Ce sont deux solutions constantes. La solution nulle est la seule autre solution constante ; on note que $F(0,0) = \frac{A^2}{4}$. Lorsque B est compris strictement entre 0 et cette valeur, on obtient comme courbes de niveau deux « œufs », centrés aux points d'abscisses \sqrt{A} et $-\sqrt{A}$, ceux-ci étant disjointes (ce sont des courbes réalisées à l'aide de tirets courts de la figure IV.3). En effet, tous les points d'abscisse nulle ont une image par F supérieure à $\frac{A^2}{4}$. C'est précisément pour cette valeur que les deux ellipsoïdes se rejoignent pour former un « huit » (courbe en trait plein dans le diagramme de phase), lequel se transforme en une figure ressemblant à un masque de plongée lorsque B dépasse cette valeur (courbes formées de longs tirets). De cette figure, on peut déduire des faits importants sur κ , et sur κ^2 . Premièrement, on voit que κ admet toujours un extrema global, et κ^2 un maximum global, qu'on notera M . Deuxièmement, cette valeur est toujours strictement plus grande que A , sauf précisément en cas de solution constante, auquel cas on a l'égalité. Au vu du graphique liant $\frac{\kappa^2}{2}$ et κ' , une égalité particulière se dégage entre M et A : c'est l'égalité $M = 2A$, qui correspond également à l'endroit où les deux ellipsoïdes se rejoignent dans le diagramme des phases. On montrera alors qu'une condition nécessaire et suffisante pour que κ admette des zéros est précisément le fait que M soit strictement plus grand que $2A$; dans le cas contraire

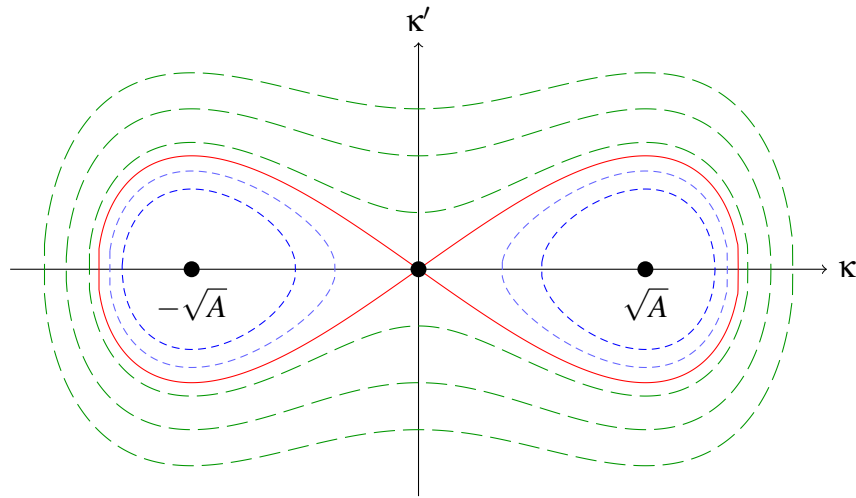


FIG. IV.3 – Courbes de niveau de la fonction F , définie par (IV.11). Il s'agit également du diagramme de phases de l'équation (IV.8), où le sens de parcours est le sens horlogique pour toutes les courbes formées de tirets ; c'est également le cas pour la courbe en trait plein, à ceci près que l'on ne passe jamais par $(0, 0)$, solution constante.

on prouvera que κ^2 est toujours plus grand que $M - 2A$. Ces faits ainsi obtenus, nous permettront de donner explicitement κ en fonction de s , comme fonction elliptique.

IV.3.3 Quelques notions sur les fonctions elliptiques

Avant de prouver formellement les points annoncés, voyons donc les notions de base des fonctions elliptiques. Nous allons pour ce faire suivre une exposition générale, en voyant les fonctions elliptiques de Jacobi comme généralisations des fonctions trigonométriques. Cette exposition suit de près un exposé de Joseph Schwalm (voir [13]), lui-même basé sur un cours de W.M. Kinnersley.

Pour définir les fonctions trigonométriques, on peut s'y prendre de la manière suivante. Considérons le cercle centré à l'origine et de rayon 1. A tout point A de coordonnées (x, y) de ce cercle, on peut associer un angle θ ; c'est l'angle que forme la demi-droite d'origine $O = (0, 0)$ et passant par $I = (1, 0)$ avec la demi-droite de même origine et passant par A . On peut d'ailleurs l'obtenir avec les

relations ¹ :

$$\theta = \begin{cases} \chi(x) & \text{si } y \geq 0; \\ 2\chi(1) - \chi(x) & \text{si } y < 0 \end{cases} \quad \text{avec } \chi(x) = \int_0^x \frac{1}{\sqrt{1-t^2}} dt$$

en remarquant que $\theta(1) = \frac{\pi}{2}$. Les fonctions sin et cos sont alors définies, sur l'intervalle $[0, 2\pi[$, par

$$\sin \theta = y \quad \text{et} \quad \cos \theta = x;$$

puis sont définies sur tout \mathbb{R} par périodicité.

Généralisons la construction précédente en considérant, non plus le cercle unité, mais une ellipse d'équation

$$\left(\frac{x}{a}\right)^2 + y^2 = 1. \quad (\text{IV.12})$$

où a est choisi, sans perte de généralité, plus grand que 1. Soit $A = (x, y)$ un point de cet ellipse ; et notons $r := \sqrt{x^2 + y^2}$. L'excentricité de l'ellipse est donnée par $k = \sqrt{1 - \frac{1}{a^2}}$; remarquons qu'elle est nulle dans le cas d'un cercle (c-à-d, quand $a = 1$). On peut montrer cette fois que la longueur de la courbe d'une ellipse est donnée, pour $x > 0$, par :

$$\chi(x) := \int_0^x \frac{1}{\sqrt{1-t^2}\sqrt{1-k^2t^2}} dt.$$

Si, au lieu de prendre les abscisses pour intégrer, on prend les angles, on effectue la transformation $t = \sin \phi$, d'où l'on a $dt = \cos \phi d\phi$; par ailleurs $1 - t^2 = \cos^2 \phi$, si bien que, en désignant toujours par θ l'angle formé par les demi-droites OI et OA , l'intégrale précédente est égale à :

$$F(\theta, k) := \int_0^\theta \frac{\cos \phi}{|\cos \phi| \sqrt{1 - k^2 \sin^2 \phi}} d\phi. \quad (\text{IV.13})$$

Cette intégrale est appelée *intégrale elliptique incomplète de première espèce*. En restreignant sa première variable sur l'intervalle $[0, \frac{\pi}{2}]$, le cosinus étant toujours positif sur cet intervalle, l'expression de F prend une forme plus simple. On convient par ailleurs de noter

$$K(k) := F\left(\frac{\pi}{2}, k\right) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - k^2 \sin^2 \phi}} d\phi. \quad (\text{IV.14})$$

¹On peut montrer que la relation définissant χ est précisément celle qui permet de calculer la longueur de la courbe du cercle unité sur sa partie à ordonnée positive ; dès lors χ peut permettre de définir la fonction arcsin.

qui, elle, est appelée *intégrale elliptique complète de première espèce*. Cette intégrale, ainsi que la précédente, sont dûes à Legendre. On peut définir $F(\cdot, k)$ en terme de sa forme simple, d'abord sur $[\frac{\pi}{2}, \pi]$ par $F(\psi) = 2K(k) - F(\pi - \psi, k)$; puis sur $[\pi, 2\pi]$ par $F(\psi, k) = 2K(k) - F(\psi - \pi, k)$; enfin sur tout \mathbb{R} en remarquant que $F(2\pi, k) = F(0, k) = 0$, et en prolongeant donc par périodicité.

Pour $k = 0$, on trouve $F(\theta, k) = \theta$, et ce, pour tout $\theta \in \mathbb{R}$; mais cette égalité n'est pas vraie en général. Par conséquent, en général, $F(\theta, k)$ n'est pas égal à l'angle formé par les demi-droites OI et OA .

Pour k fixé, la fonction $F(\cdot, k)$ est strictement croissante sur l'intervalle $[0, \frac{\pi}{2}]$, donc injective. En effet, les cosinus se simplifiant (car positifs), l'intégrand est alors une fonction strictement positive sur l'intervalle $[0, \frac{\pi}{2}]$ [*en $\frac{\pi}{2}$ la fonction explose => vérifier que F n'explose pas.*] De plus, pour tout θ dans cet intervalle, on a $F(\theta, k) \in [0, K(k)]$. Sur ces intervalles, à k fixé, F admet donc une réciproque, définie comme étant la fonction *amplitude*, et notée $\text{am}(\cdot | k^2)$. On a donc, pour tout $k \in [0, 1]$, et pour tout $u \in [0, 2K(k)]$, l'égalité $F(\text{am}(u | k^2), k) = u$.

Grâce à l'amplitude, on peut définir les fonctions de Jacobi. On définit, d'abord sur l'intervalle $[0, K(k)]$, pour $m = k^2 \in [0, 1]$:

$$\begin{aligned} \text{sn}(u|m) &= \sin \text{am}(u|m), \\ \text{cn}(u|m) &= \cos \text{am}(u|m) \quad \text{et} \\ \text{dn}(u|m) &= \partial_u \text{am}(u|m). \end{aligned}$$

puis, comme précédemment, l'on prolonge ces définitions sur tout \mathbb{R} . Ces fonctions sont respectivement appelées *sinus elliptique*, *cosinus elliptique*, et *delta elliptique*. On peut définir les fonctions elliptiques autrement, grâce à la proposition suivante :

Proposition 32. *Soit $a > 0$, fixons A un point de coordonnées (x, y) situé sur l'ellipse d'équation (IV.12). On note $m = 1 - \frac{1}{a^2}$ le carré de son excentricité ; soit θ l'angle formé par les droites OI et OA . Si u est une valeur dont l'amplitude (de paramètre m) vaut θ , alors on a les égalités :*

$$\text{sn}(u|m) = y, \quad \text{cn}(u|m) = \frac{x}{a}, \quad \text{et} \quad \text{dn}(u|m) = \frac{r}{a},$$

où $r = \sqrt{x^2 + y^2}$.

[preuve ? Ou réf.] Remarquons alors que le delta elliptique, est toujours strictement positif, quel que soit son argument et son paramètre.

Grâce aux égalités de la proposition précédente, on déduit des formules liant les trois fonctions elliptiques : on a, pour tout u ,

$$\text{sn}(u|m)^2 + \text{cn}(u|m)^2 = 1. \quad (\text{IV.15})$$

De plus, on obtient $\operatorname{dn}(u|m)^2 = \operatorname{cn}(u|m)^2 + \frac{1}{a^2} \operatorname{sn}(u|m)^2$; en remplaçant cn^2 en accord avec la formule ci-dessus, on a $\operatorname{dn}(u|m)^2 - \frac{1}{a^2} \operatorname{sn}(u|m)^2 + \operatorname{sn}(u|m)^2 = 1$, soit, par définition de k et m :

$$\operatorname{dn}(u|m)^2 + m \operatorname{sn}(u|m)^2 = 1. \quad (\text{IV.16})$$

Puisque la fonction dn est toujours positive quelle que soit le paramètre, cette égalité est équivalente à $\operatorname{dn}(u|m) = \sqrt{1 - m \operatorname{sn}(u|m)^2}$.

Grâce aux définitions, on peut calculer facilement les dérivées de chacune des fonctions elliptiques. On obtient :

$$\begin{aligned} \partial_u \operatorname{sn}(u|m) &= \operatorname{cn}(u|m) \operatorname{dn}(u|m) \\ \partial_u \operatorname{cn}(u|m) &= -\operatorname{sn}(u|m) \operatorname{dn}(u|m) \\ \partial_u \operatorname{dn}(u|m) &= -m \operatorname{sn}(u|m) \operatorname{cn}(u|m). \end{aligned}$$

Pour la première égalité, on a

$$\begin{aligned} \partial_u \operatorname{sn}(u|m) &= \partial \sin(\operatorname{am}(u|m)) \partial_u \operatorname{am}(u|m) \\ &= \cos(\operatorname{am}(u|m)) \operatorname{dn}(u|m) = \operatorname{cn}(u|m) \operatorname{dn}(u|m). \end{aligned}$$

La deuxième égalité se prouve de façon similaire. Pour la troisième, utilisant l'égalité $\operatorname{dn}(u|m) = \sqrt{1 - m \operatorname{sn}(u|m)^2}$, on obtient

$$\begin{aligned} \partial_u \operatorname{dn}(u|m) &= \partial_u \sqrt{1 - m \operatorname{sn}(u|m)^2} \\ &= \frac{-2m \operatorname{sn}(u|m) \partial_u \operatorname{sn}(u|m)}{2\sqrt{1 - m \operatorname{sn}(u|m)^2}} \\ &= -m \frac{\operatorname{sn}(u|m) \operatorname{cn}(u|m) \operatorname{dn}(u|m)}{\operatorname{dn}(u|m)} = -m \operatorname{sn}(u|m) \operatorname{cn}(u|m). \end{aligned}$$

Si l'on élève la première de ces relations au carré, et que l'on utilise les deux égalités précédentes (IV.15) et (IV.16), on obtient l'égalité

$$(\partial_u \operatorname{sn}(u|m))^2 = (1 - \operatorname{sn}(u|m)^2)(1 - m \operatorname{sn}(u|m)^2) \quad (\text{IV.17})$$

qui sera utilisée par la suite.

Terminons cette partie en considérant deux types de fonctions elliptiques particulières ; il s'agit de celles dont le paramètre vaut respectivement 0 et 1. Pour le cas où $m = 0$, on a déjà remarqué que $F(\cdot, 0)$ était l'identité ; sa réciproque, $\operatorname{am}(\cdot|0)$ envoie donc tout u sur u , et les définitions des fonctions elliptiques se réduisent à $\operatorname{sn}(u|0) = \sin u$, $\operatorname{cn}(u|0) = \cos u$ et $\operatorname{dn}(u|0) = 1$. Nous retrouvons donc la trigonométrie usuelle. Le cas $m = 1$ est tout aussi particulier ; on peut le voir

comme un cas limite où l'on fait tendre a vers $+\infty$. On peut montrer dans ce cas que les fonctions elliptiques de paramètre 1 sont connues ; en fait, nous avons les égalités $\operatorname{sn}(u|1) = \tanh u$; $\operatorname{cn}(u|1) = \operatorname{sech} u$; $\operatorname{dn}(u|1) = \operatorname{sech} u$, où la fonction sech est la sécante hyperbolique, fonction inverse du cosinus hyperbolique (\tanh désignant la tangente hyperbolique).

Il est possible d'en dire beaucoup plus sur ces fonctions ; beaucoup de livres y sont consacrés. Cependant, nous ne développerons pas davantage cette théorie, les éléments nécessaires pour la suite étant tous compris dans ce résumé.

IV.3.4 Formule explicite de la courbure

Théorème 33. Soit $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ une solution non-nulle de l'équation différentielle (IV.8). Supposons que κ^2 admet un maximum global M en \tilde{s} . Alors :

1. $M \geq A$; et $M = A$ si et seulement si κ^2 est constante, égale à A ;
2. κ admet des zéros si et seulement si $M > 2A$; si au contraire $M \leq 2A$, alors κ^2 se trouve toujours dans l'intervalle $[M - 2A, M]$;
3. lorsque $M > 2A$, on a

$$\kappa(s) = \operatorname{sign} \kappa(\tilde{s}) \sqrt{M} \operatorname{cn} \left(\sqrt{\frac{M-A}{2}} (s - \tilde{s}) \mid \frac{M}{2(M-A)} \right) ;$$

4. lorsque $A < M \leq 2A$, on a

$$\kappa(s) = \sqrt{M} \operatorname{dn} \left(\sqrt{M} (s - \tilde{s}) \mid \frac{2(M-A)}{M} \right).$$

Remarquons avant de commencer la preuve, que, malgré la découpe en différents cas, il y a une continuité vis-à-vis du paramètre M . En effet, quand $M = 2A$, on obtient $2(M-A) = M$ donc les deux fractions, servant de paramètres pour les fonctions elliptiques, sont toutes deux égales à 1. Mais, comme remarqué précédemment, $\operatorname{cn}(\cdot|1) = \operatorname{dn}(\cdot|1) = \operatorname{sech}$. De même, quand $M = A$, le paramètre du delta elliptique vaut 0 et dans ce cas, on a l'égalité $\operatorname{dn}(\cdot|0) = 1$.

Démonstration. On a $\kappa'(\tilde{s}) = 0$; dès lors, l'expression (IV.9) donne, en \tilde{s} :

$$\frac{(\kappa^2(\tilde{s}) - A)^2}{4} = \frac{(M - A)^2}{4} = B.$$

Par conséquent, l'expression (IV.9) peut se récrire

$$\begin{aligned} (\kappa')^2 &= \frac{(M - A)^2}{4} - \frac{(\kappa^2 - A)^2}{4} \\ &= \frac{(M - \kappa^2)(M + \kappa^2 - 2A)}{4}. \end{aligned} \tag{IV.18}$$

κ étant une solution réelle, $M - \kappa^2$ étant toujours positif, on en déduit que $M + \kappa^2 - 2A$ est aussi positif, et ce, quel que soit s ; par conséquent $M \geq A$. Si de plus κ n'est pas constant, κ' admet des valeurs non-nulles, et donc $(\kappa')^2$ est strictement positif; cela n'est possible que pour $\kappa^2 < M$, et pour $M + \kappa^2 > 2A$, d'où l'inégalité stricte. Cela prouve le point 1.

Pour le point 2, supposons d'abord que $\kappa(s) = 0$ pour un certain $s \in \mathbb{R}$; dès lors, $(\kappa'(s))^2 = \frac{M}{4}(M - 2A)$; puisque le membre de gauche est positif, le membre de droite doit l'être aussi, d'où $M \geq 2A$. Si l'on avait l'égalité, alors κ' devrait être nul, et donc κ aussi car constant; mais cela contredit notre hypothèse de base. En conséquence, on a l'inégalité stricte $M > 2A$. Réciproquement, à l'aide du diagramme de phases, on peut montrer que, si on a cette inégalité, alors κ admet des zéros. En effet, si l'inégalité a lieu, alors le couple κ, κ' décrit l'une des courbes en tirets longs, et l'on peut montrer que le couple la suit effectivement sans effectuer d'arrêt.

Par ailleurs, le second facteur du produit apparaissant dans (IV.18) devant être positif pour tout $s \in \mathbb{R}$, si on suppose que $M \leq 2A$, alors $M - 2A \leq 0$ et κ^2 doit donc être plus grand que $2A - M$.

Prouvons à présent que κ s'exprime explicitement. Supposons d'abord que $M > 2A$; et définissons la fonction z comme suit :

$$z(s) := \sqrt{\frac{M - \kappa^2(s)}{M}}$$

Elle est bien définie, puisque M est la borne supérieure de $\text{Im } \kappa^2$. En élevant au carré, puis en dérivant, on obtient :

$$2zz' = (z^2)' = \frac{1}{M}(-2\kappa\kappa')$$

soit, en simplifiant et en élevant à nouveau au carré :

$$\begin{aligned} z^2(z')^2 &= \frac{1}{M^2}\kappa^2(\kappa')^2 \\ &= \frac{1}{4M^2}\kappa^2(M - \kappa^2)(M + \kappa^2 - 2A) \\ &= \frac{1}{4M^2}(M - Mz^2)(Mz^2)(M + M - Mz^2 - 2A) \\ &= \frac{1}{4}z^2(1 - z^2)(2(M - A) - Mz^2) \\ &= \frac{1}{2}z^2(M - A)(1 - z^2)\left(1 - \frac{M}{2(M - A)}z^2\right) \end{aligned}$$

Sur l'ouvert où la fonction z n'est pas nulle, on peut simplifier les deux membres par z^2 . On prolonge par continuité à la fermeture de cet ouvert. Si ce dernier

ensemble n'est pas \mathbb{R} , il existe un intervalle où z est nulle ; alors κ atteint son extrema sur cet intervalle et $\kappa^2 = M$ est donc constante sur cet intervalle. Par le point 1, on déduit $M = A$, ce qui contredit notre hypothèse $M > 2A$.

L'équation ainsi obtenue est alors, à constante multiplicative près, celle que vérifie le sinus elliptique (voir (IV.17)) ; on remarque que $2(M - A) = 2M - 2A > M$ si bien que le paramètre de la fonction est bien entre 0 et 1. On obtient

$$z(s) = \operatorname{sn}\left(\sqrt{\frac{M-A}{2}}(s - \tilde{s}) \mid \frac{M}{2(M-A)}\right)$$

donc, par l'égalité $z^2 + \frac{\kappa^2}{M} = 1$ et la propriété (IV.15) liant sinus et cosinus elliptiques, on obtient (3).

On effectue un travail similaire pour M dans l'intervalle $]A, 2A]$; en posant

$$z(s) := \sqrt{\frac{M - \kappa^2(s)}{2(M-A)}}$$

qui est bien définie, on a

$$2zz' = (z^2)' = \frac{1}{2(M-A)}(-2\kappa\kappa')$$

puis en élevant au carré :

$$\begin{aligned} z^2(z')^2 &= \frac{1}{4(M-A)^2} \kappa^2 \frac{1}{4} (M - \kappa^2)(M + \kappa^2 - 2A) \\ &= \frac{1}{16(M-A)^2} (M - 2(M-A)z^2) (2(M-A)z^2) (2M - 2(M-A)z^2 - 2A) \\ &= \frac{1}{8(M-A)} (M - 2(M-A)z^2) z^2 (2(M-A) - 2(M-A)z^2) \\ &= \frac{1}{4} z^2 \left(1 - \frac{2(M-A)}{M} z^2\right) (1 - z^2) \end{aligned}$$

et on procède avec le même argument que précédemment pour simplifier z^2 .

Ici, $2(M - A) = 2M - 2A \leq M$, donc $\frac{M}{2(M-A)}$ est bien entre 0 et 1 ; par l'égalité (IV.17), z s'exprime alors par :

$$z(s) = \operatorname{sn}\left(\sqrt{\frac{M-A}{2}}(s - \tilde{s}) \mid \frac{M}{2(M-A)}\right)$$

et on déduit l'expression de κ telle que dans l'équation (4), en considérant cette fois l'égalité $\frac{2(M-A)}{M} z^2 + \frac{\kappa^2}{M} = 1$ et la propriété (IV.16) liant sinus et delta elliptiques. \square

Le fait de connaître explicitement les courbures possibles permet de déterminer les angles que font les vecteurs tangents en tout point, via l'expression $\Psi(s) = \int_0^s \kappa(s) ds + \Psi_0$; grâce à cette expression, on en déduit une formule explicite pour les elastica.

IV.4 Classification des elastica

En faisant varier le paramètre A dans les équations précédentes, on fait varier les formulations de la courbure ; et donc le type de courbes obtenues. Nous présentons ici les principales formes des courbes, en précisant pour quelles valeurs du paramètre ces courbes sont obtenues. Sans perte de généralité, on suppose que le maximum M du carré de la courbure est égal à 1.

Remarquons cependant les faits suivants. D'abord, si M est fixé à 1, alors le nombre a intervenant dans l'expression de la courbure d'un elastica (voir définition 14) est fixé en fonction de A : en effet, le maximum de κ^2 vaut exactement $2a + A$; cette quantité étant fixée à 1, on obtient l'égalité $a = \frac{1-A}{2}$. Ensuite, comme κ est solution d'une équation autonome, on peut supposer que son maximum (ou son minimum) est atteint en 0 ; par conséquent, on peut fixer également dans la définition 14, $\phi = 0$.

Dès lors, au vu de l'égalité $2a \cos \phi (1 - \delta) = A$, on obtient grâce à ci-dessus, $(1 - A)(1 - \delta) = A$; soit l'égalité

$$1 - \delta + A(\delta - 2) = 0$$

ou encore, puisque $\delta \in [0, 1]$:

$$A = 1 - \frac{1}{2 - \delta}.$$

Vu la variation possible pour δ , il suffit de faire varier A entre 0 et 1 pour obtenir tous les types de courbes.

La figure IV.4 représente l'elastica de base, celui qui a été étudié de prime abord par J. Bernoulli. On remarque d'ailleurs sur la courbe, la forme particulière, à ceci près qu'on impose ici le passage de la courbe à l'origine avec une tangente nulle. La figure IV.7 présente une courbe plane fermée ; on peut montrer que c'est la seule valeur de A (autre que 1) pour laquelle la courbe élastique obtenue est fermée.

La figure IV.9 est l'une des plus particulières. Contrairement à toutes les autres, la courbure, pour cette valeur du paramètre, n'atteint son maximum qu'une seule fois. On a déjà noté que cette courbure est la fonction sécante hyperbolique.

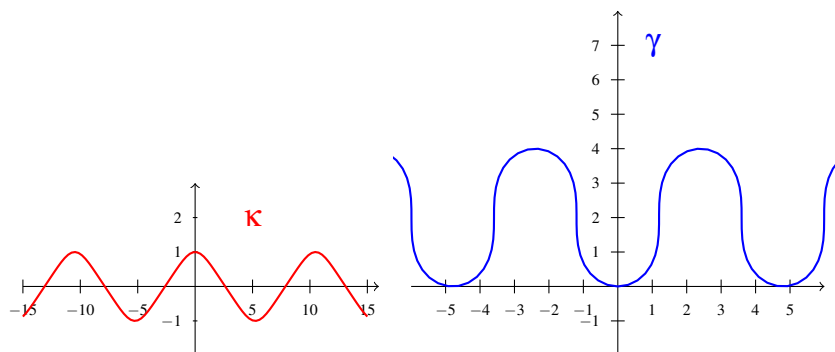


FIG. IV.4 – Courbe élastique obtenue avec paramètre $A = 0$.

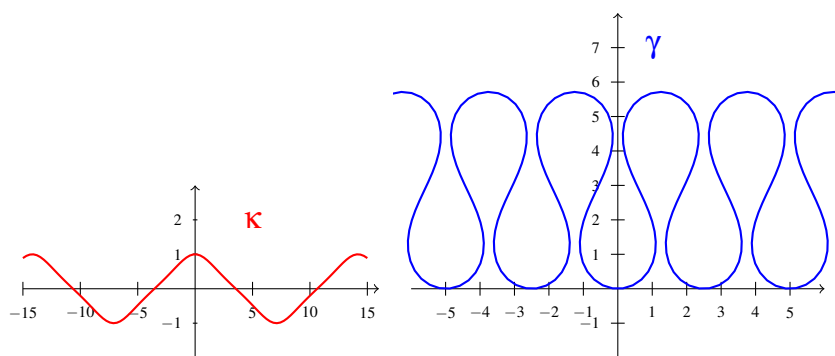


FIG. IV.5 – Courbe élastique obtenue avec paramètre $A = 0.3$.

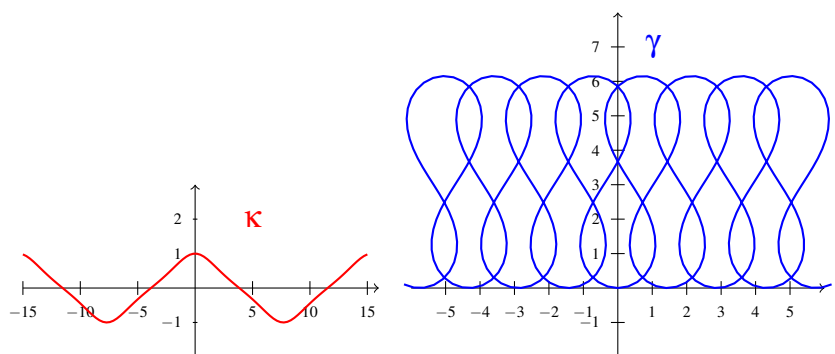


FIG. IV.6 – Courbe élastique obtenue avec paramètre $A = 0.35$.

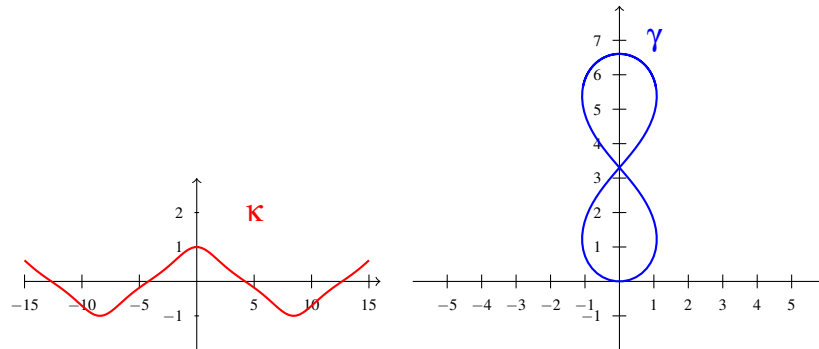


FIG. IV.7 – Courbe élastique obtenue avec paramètre $A = 0.394757$.

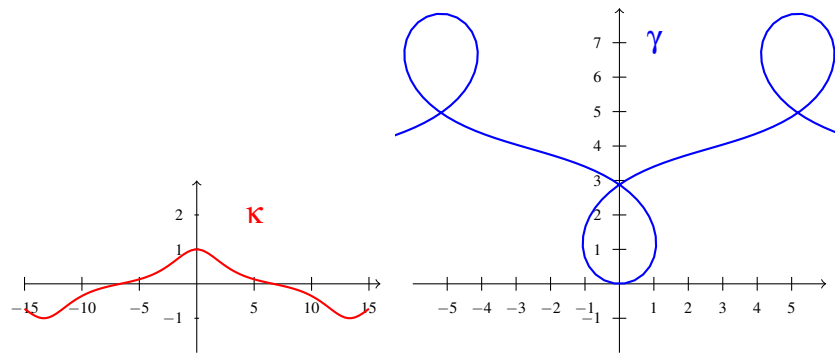


FIG. IV.8 – Courbe élastique obtenue avec paramètre $A = 0.49$.

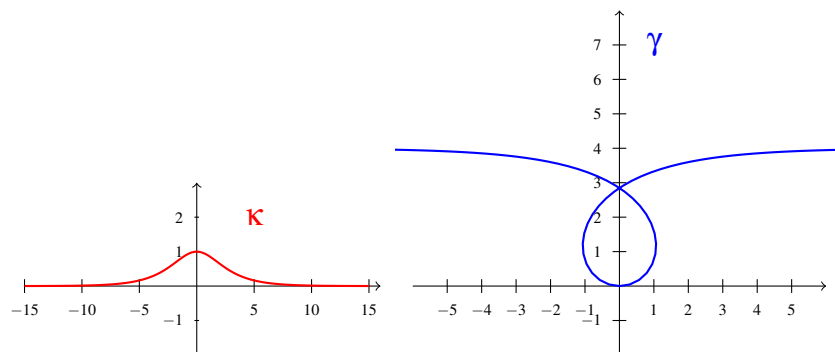


FIG. IV.9 – Courbe élastique obtenue avec paramètre $A = 0.5$.

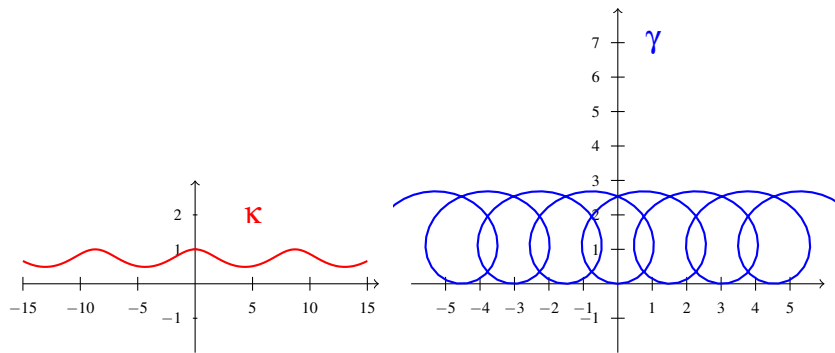


FIG. IV.10 – Courbe élastique obtenue avec paramètre $A = 0.62$.

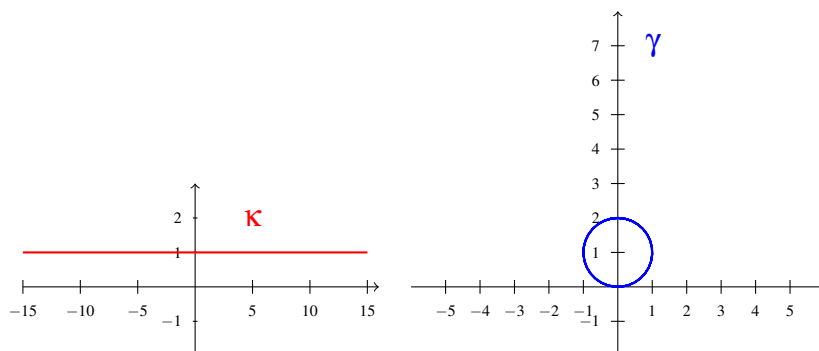


FIG. IV.11 – Courbe élastique obtenue avec paramètre $A = 1$.

IV.5 Généralisations du problème

Outre le problème tel que posé en 1691 par Bernoulli, il est possible de le généraliser de différentes manières. Nous discutons deux voies possibles.

IV.5.1 Le problème des courbes dans l'espace

La première consiste à modifier l'espace de base, comme le fait D. Singer [14]. Au lieu de considérer des courbes planes, on s'intéresse à minimiser la même fonctionnelle, mais cette fois pour des courbes dans l'espace. En utilisant la technique du calcul des variations, il obtient deux équations que doivent vérifier la courbure et la torsion d'une telle courbe ; l'une de ces équations est similaire à l'équation (IV.8) si ce n'est qu'un terme faisant intervenir la torsion est ajouté. Ainsi, si κ est la courbure, et τ désigne la torsion, les équations suivantes sont obtenues :

$$\begin{aligned}\kappa'' &= -\frac{1}{2}\kappa(\kappa^2 - A - 2\tau^2) \\ \kappa\tau' + 2\kappa'\tau &= 0\end{aligned}$$

Remarquons qu'à torsion nulle, la seconde équation est toujours vérifiée, tandis que la première est exactement (IV.8). Résoudre ce système est équivalent à résoudre les deux équations

$$\frac{(\kappa^2 - A)^2}{4} + \kappa'^2 + \kappa^2\tau^2 = c_1 \quad \text{et} \quad \kappa^2\tau = c_2$$

où c_1, c_2 sont deux constantes ; cela peut se voir en multipliant, dans les deux premières égalités, par κ' dans la première équation, et par κ dans la seconde. Pour autant que $\kappa \neq 0$, ces deux équations donnent lieu à l'égalité

$$\frac{(\kappa^2 - A)^2}{4} + \kappa'^2 + \frac{c_2^2}{\kappa^2} = c_1.$$

D. Singer en déduit alors les solutions pour le carré de la courbure, sous une forme faisant intervenir le sinus elliptique ; cette forme est similaire à celle que nous avons obtenu pour pouvoir en déduire les expressions en termes de cosinus et delta elliptiques.

La grande différence entre les courbes élastiques planes et les courbes dans l'espace réside sans doute dans le théorème suivant, prouvé par Langer et Singer, en 1983 (et dont l'énoncé exact, ainsi que la preuve, est dans [14]) :

Théorème 34. *Il existe une infinité de courbes dans l'espace, qui minimisent leur énergie de courbure, et qui sont fermées.*

IV.5.2 Membranes élastiques sur un substrat mou

L'autre généralisation possible est de considérer des films posés sur un substrat et pressés latéralement. S'il n'y a tout simplement pas de substrat, on retombe dans le cas des courbes planes ; ici, un terme d'énergie potentielle s'ajoute à l'énergie que l'on souhaite minimiser ; l'expression de l'énergie est alors modélisée, pour des substrats très mous (par exemple, l'eau), par :

$$\mathcal{E} : \gamma \mapsto \frac{1}{2}B \int_0^L \|\partial_s^2 \gamma(s)\|^2 ds + \frac{1}{2}K \int_0^L (\gamma_2(s))^2 \partial_s \gamma_1(s) ds,$$

avec B et K deux constantes positives. On retrouve dans le premier terme l'énergie de courbure ; le second terme est un terme d'énergie potentielle, lié au déplacement du substrat sous le film. L'espace sous-jacent à cette étude sera constitué de courbes planes deux fois différentiables, dont les conditions au bord sont fixées, ainsi que les vecteurs tangents ; comme pour l'étude lorsqu'il n'y a pas de substrat, on peut se restreindre aux courbes paramétrées par longueur d'arc. L'espace X que l'on considère est donc le suivant :

$$X = \left\{ \gamma \in H^2(]0, L[; \mathbb{R}^2) \mid \forall s \in]0, L[, \|\partial_s \gamma(s)\| = 1 \text{ et} \right. \\ \left. \begin{aligned} \gamma(0) &= (0, 0), \quad \partial_s \gamma(0) = (1, 0), \\ \gamma(L) &= (L - \delta, 0), \quad \partial_s \gamma(L) = (1, 0) \end{aligned} \right\}.$$

A notre connaissance, il n'existe pas à l'heure actuelle, de résolution formelle de ce problème ; L. Pocivavsek [12] tente cependant de trouver une approximation linéaire aux solutions, et obtient des résultats numériques sans expliquer réellement la méthode employée.

Néanmoins, l'existence d'une solution peut être prouvée grâce à la méthode directe du calcul des variations. Les expérimentations physiques mettent en évidence certaines propriétés que devraient posséder les solutions de ce problème. Une piste de recherche est donc de prouver rigoureusement ces propriétés ; notamment de symétries et d'unicité.

Lorsque le substrat n'est pas totalement mou (comme un gel, par exemple), les expériences physiques montrent que la courbe se comporte assez différemment. Alors que pour l'eau, la courbe concentre son énergie en un pic, pour un gel, elle en forme plusieurs, certains devenant plus prononcés que d'autres lorsque le film est suffisamment pressé. Dans ce cadre, la forme que doit avoir l'énergie n'est, pour l'instant, pas connue. Cela ouvre une nouvelle porte de recherches, en étudiant les expérimentations physiques effectuées, ou en utilisant des algorithmes de continuation.

On le voit, ce problème relativement ancien suscite encore beaucoup de questions, dont la plupart sont encore ouvertes.

Annexe A

Calcul des exponentielles matricielles

Cette partie est consacrée aux détails techniques de l'étude de stabilité linéaire. On garde les mêmes notations que dans le texte principal, à savoir que A désigne une matrice carrée d'ordre n à coefficients réels, et que l'on considère le système (I.6) :

$$\partial_t u = Au.$$

A condition initiale $u(0) = u_0$ fixée, on a montré que la solution est $u(t) = e^{At} u_0$; c'est le théorème 4.

A.1 Calcul des solutions

Si A est diagonale, son exponentielle l'est aussi ; et pour la calculer, il suffit de prendre l'exponentielle de chaque coefficient diagonal. Pour une matrice A diagonalisable (ou Jordanisable) dans \mathbb{R} , si P est une matrice de passage et si $B = P^{-1}AP$ est diagonale (resp. la Jordanisée correspondante), alors :

$$e^{At} = P e^{Bt} P^{-1}.$$

Malheureusement, la Jordanisation dans \mathbb{R} n'est pas toujours possible. Il faut, dans certains cas, passer aux complexes (il s'agit en fait des cas où la matrice admet des valeurs propres complexes). Dans un premier temps, nous discutons de ce qui se passe dans le cas de matrices à valeurs propres réelles. Puis, nous discutons des matrices admettant des valeurs propres complexes ; d'abord le cas diagonalisable, ensuite le cas « Jordanisable ».

A.1.1 Cas de matrices à valeurs propres réelles

Supposons donc que A n'admette que des valeurs propres réelles. Si ces valeurs propres sont toutes distinctes, l'espace \mathbb{R}^n est décomposé en somme directe de n sous-espaces propres de dimension 1, chacun lié à une valeur propre. Dans le cas contraire, tout dépend de la matrice selon qu'elle soit diagonalisable ou non.

Si elle est diagonalisable, on obtient quelque chose de similaire au cas de valeurs propres distinctes : l'espace \mathbb{R}^n se scinde en plusieurs espaces propres, chacun ayant pour dimension la multiplicité algébrique de la valeur propre qui lui est associée. Etant donnée une base de \mathbb{R}^n constituée exclusivement de vecteurs propres, chacun de ces espaces peut être à nouveau scindé en sous-espaces, chacun de dimension 1 (car engendré par un seul vecteur). *[On convient d'appeler sous-espaces propres chacun de ces sous-espaces.]*

Si la matrice n'est que « Jordanisable », on n'obtient plus de découpe aussi claire. Précisons ce qui se passe dans ce cas. Soit λ une valeur propre de A , posons m sa multiplicité algébrique et d la dimension de l'espace propre associé. Alors, si $d = m$, la matrice est diagonalisable sur le bloc correspondant : on ne trouvera pas, sur sa Jordanisée, un bloc de Jordan composé de λ de taille plus grande que 1. Puisqu'on suppose A non-diagonalisable, on a, pour certaines valeurs propres λ , l'inégalité stricte $d < m$; pour de telles valeurs propres, on obtient d blocs de Jordan, dont au moins l'un d'entre eux est de taille plus grande que 1. Si m_1, \dots, m_d désignent les tailles respectives des blocs, alors $\sum_{i=1}^d m_i = m$; par ailleurs, on obtient une découpe en d espaces, le i^{e} ayant pour dimension m_i . Ces d espaces *[qu'on appelle sous-espaces propres généralisés.]* engendrent un espace de dimension m , qui porte le nom d'*espace propre généralisé*. A ne pas confondre avec les espaces propres donc, qui sont en général de dimension strictement inférieure.

A.1.2 Cas de matrices diagonalisables à valeurs propres complexes

Il est facile de voir que si un complexe z est valeur propre d'une matrice à coefficients réels, alors \bar{z} , son conjugué, est aussi valeur propre ; en effet, le polynôme caractéristique est à coefficients réels.

Supposons dans un premier temps, que la matrice A soit diagonalisable. Alors, il existe une matrice diagonale B , semblable à A , qui s'écrit sous la forme

$$B = \text{diag}(\lambda_1, \dots, \lambda_k, z_1, \bar{z}_1, \dots, z_m, \bar{z}_m) = \text{diag}(\lambda_1, \dots, \lambda_k, D_1, \dots, D_m) \quad (\text{A.1})$$

$$\text{où, pour tout } i = 1, \dots, m, D_i := \begin{pmatrix} z_i & 0 \\ 0 & \bar{z}_i \end{pmatrix}$$

On va montrer que les matrices D_i sont semblables à des matrices à coefficients réels. Soit v_i un vecteur propre de D_i de valeur propre z_i . On note x_i et y_i respectivement les vecteurs partie réelle et partie imaginaire de v_i ; donc, $v_i = x_i + \mathbf{i}y_i$. Alors,

$$\begin{aligned} Av_i &= z_iv_i \\ A(x_i + \mathbf{i}y_i) &= z_i(x_i + \mathbf{i}y_i) \\ Ax_i + \mathbf{i}Ay_i &= z_ix_i + \mathbf{i}z_iy_i \end{aligned}$$

et si r_i et s_i désignent respectivement les parties réelle et imaginaire de z_i , on obtient :

$$\begin{cases} Ax_i = r_ix_i - s_iy_i \\ Ay_i = s_ix_i + r_iy_i \end{cases} \quad (\text{A.2})$$

On peut faire un raisonnement similaire avec un vecteur propre de valeur propre \bar{z}_i , et obtenir des relations comme celles ci-dessus. Remarquons par ailleurs que $y_i + \mathbf{i}x_i = \mathbf{i}(\overline{x_i + \mathbf{i}y_i})$ est un de ces vecteurs propres, conduisant exactement aux relations (A.2).

Pour chaque valeur de i , on peut alors remplacer le bloc complexe D_i par un bloc réel défini par

$$\tilde{D}_i := \begin{pmatrix} r_i & -s_i \\ s_i & r_i \end{pmatrix} \quad (\text{A.3})$$

et donc, remplacer la matrice B diagonale, mais à coefficients complexes, par une matrice \tilde{B} , presque diagonale, antisymétrique sur les éléments non-diagonaux, à coefficients réels ; en fait,

$$\tilde{B} = \text{diag}(\lambda_1, \dots, \lambda_k, \tilde{D}_1, \dots, \tilde{D}_m). \quad (\text{A.4})$$

Remarquons que A reste semblable à \tilde{B} , mais maintenant, on peut prendre une matrice de passage réelle ; en effet, si w_i est un vecteur propre de valeur propre λ_i , et si l'on définit

$$X := (w_1, \dots, w_k, x_1, y_1, \dots, x_m, y_m)$$

on constate que $AX = \tilde{B}X$. Par ailleurs, si l'on fixe un w_i , on a $\tilde{B}w_i = \lambda_iw_i$; et si l'on fixe une paire de vecteurs (x_i, y_i) , on a $\tilde{B}(x_i, y_i) = (r_ix_i - s_iy_i, r_iy_i + s_ix_i) = (x_i, y_i) \cdot \tilde{D}_i$. Ces égalités montrent alors que X commute avec \tilde{B} ; d'où l'on obtient $X^{-1}AX = \tilde{B}$.

Dans cete situation, l'espace \mathbb{R}^n est décomposé en $k + m$ sous-espaces, les k premiers sous-espaces étant de dimension 1, et engendrés par les w_i ; les m autres étant de dimension 2, et engendrés par une paire de vecteurs (x_i, y_i) . On a évidemment l'égalité $k + 2m = n$.

A.1.3 Cas de matrices Jordanisables à valeurs propres complexes

Passons maintenant au cas général, d'une matrice Jordanisable, non diagonalisable. Le sous-espace correspondant aux valeurs propres réelles est déjà traité : en effet, en se restreignant à ce sous-espace, on peut appliquer ce que l'on a fait précédemment. Il nous faut donc regarder ce qui se passe pour les sous-espaces propres généralisés de valeurs propres complexes. On se place dans le cas où A n'est pas diagonalisable, et qu'il existe une valeur propre complexe z pour laquelle la multiplicité algébrique m est strictement supérieure à la dimension de l'espace propre associé d . Il existe alors des vecteurs w (complexes), qui annulent l'opérateur $(A - z\text{Id})^m$, mais qui ne sont pas vecteurs propres (i.e., ils n'annulent pas $A - z\text{Id}$). Cela crée d sous-espaces non triviaux, *[les définir !]* un par valeur propre, dont au moins un est de dimension strictement supérieure à 1. La construction de Jordan permet de les construire un à un, de la façon suivante.

Posons, pour tout $k = 0, \dots, m-1$, $U_k := \{w \in \mathbb{R}^n \mid (A - z\text{Id})^{k+1}w = 0 \text{ et } (A - z\text{Id})^k w \neq 0\}$, et $V_k := \cup_{i \leq k} U_i$. Alors $U_0 = V_0$ est l'espace propre usuel ; c'est un sous-espace vectoriel de dimension d . Par ailleurs, tous les V_k sont des sous-espaces vectoriels, par construction. Notre hypothèse fait que U_1 est non-vide ; et par la même occasion, qu'on a l'inclusion stricte $V_1 \subsetneq V_0$. Posons alors m_1 l'indice maximal pour lequel U_m est non-vide ; soit $v_0^{(1)}$ un vecteur de U_{m_1} . On construit alors la suite $v_0^{(1)}, \dots, v_{m_1}^{(1)}$, qu'on définit par la relation $v_i^{(1)} = Av_{i-1}^{(1)} - zv_{i-1}^{(1)} = (A - z\text{Id})^i v_0^{(1)}$. Remarquons que chaque $v_i^{(1)}$ se trouve alors dans U_{m_1-i} ; en particulier, $v_0^{(1)}$ est un vecteur propre. Par ailleurs, on obtient, pour tout $i = 0, \dots, m_1 - 1$, les égalités $Av_i^{(1)} = (A - z\text{Id})v_i^{(1)} + zv_i^{(1)} = zv_i^{(1)} + v_{i+1}^{(1)}$. L'espace E_1 engendré par tous les $v_i^{(1)}$ est un espace vectoriel ; qui est sous-espace vectoriel de V_{m-1} .

Cependant, $E_1 \subsetneq V_{m-1}$ en général, puisque E_1 ne contient qu'un vecteur propre (c'est $v_0^{(1)}$), alors que V_{m-1} les contient tous. L'égalité des espaces est obtenue précisément s'il n'y a qu'un vecteur propre associé à z , c'est-à-dire, quand $d = 1$; et dans ce cas, $m_1 = m$. Si l'inclusion est stricte, alors on a $d \geq 2$; il existe donc $k \geq 0$ pour lequel U_k contient d'autres vecteurs que les multiples (non-nuls) de $v_k^{(1)}$. Dans ce cas, on réitère la construction, pour obtenir des vecteurs $v_k^{(2)}$, pour $k = 0, \dots, m_2$; ces vecteurs engendrent un nouvel espace E_2 . On poursuit de la sorte de façon à obtenir, comme annoncé, d sous-espaces non-triviaux E_1, \dots, E_d , dont un est au moins de dimension 2. Remarquons néanmoins que la construction ci-dessus fonctionne pour toutes les matrices à coefficients dans un corps ; nous l'avons présentée ici dans le cadre des complexes.

Pour simplifier les notations pour la suite, posons E l'un des sous-espaces E_i

ci-avant définis, et notons k sa dimension. Alors, l'espace \bar{E} , défini par

$$\bar{E} := \{y + \mathbf{i}x \mid x + \mathbf{i}y \in E\},$$

est un sous-espace qui a même dimension que E . Ces deux sous-espaces sont en réalité sous-ensembles de \mathbb{C}^n . Remarquons les faits suivants :

- si $v = x + \mathbf{i}y \in E$ vérifie $Av = zv$, alors le vecteur correspondant dans \bar{E} , qui est $w = y + \mathbf{i}x$, vérifie $Aw = \bar{z}w$;
- si $v = x + \mathbf{i}y$ et $v' = x' + \mathbf{i}y'$ sont deux vecteurs de E tels que $Av = zv + v'$, alors les vecteurs correspondants dans \bar{E} , respectivement $w = y + \mathbf{i}x$ et $w' = y' + \mathbf{i}x'$, vérifient $Aw = \bar{z}w + w'$.

Il suffit pour les montrer, de prendre les conjugués sur les égalités connues, puis de multiplier les deux membres obtenus par le scalaire \mathbf{i} . On en déduit alors que si $(v_i), i = 1, \dots, k$ est une base de vecteurs propres généralisés de E , avec $v_i = x_i + \mathbf{i}y_i$, alors les vecteurs correspondants $w_i = y_i + \mathbf{i}x_i, i = 1, \dots, k$, forment une base de vecteurs propres généralisés de \bar{E} .

Considérons $(v_i), i = 1, \dots, k$ base de vecteurs propres généralisés de E de valeur propre $z = r + \mathbf{i}s$; comme pour la construction de E , on pose que v_k est un vecteur propre de A et que pour tous les autres indices i , on a $Av_i = zv_i + v_{i+1}$. Si $v_i = x_i + \mathbf{i}y_i$, on obtient, de façon similaire au cas Jordanisable à valeurs propres réelles, et au cas diagonalisable à valeurs propres complexes, les relations

$$\begin{cases} Ax_i = rx_i - sy_i + x_{i+1} \\ Ay_i = sx_i + ry_i + y_{i+1} \end{cases} \quad (\text{A.5})$$

Ceci suggère la transformation suivante : si J est une Jordanisée complexe de A , elle contient un bloc correspondant aux espaces E et \bar{E} ; ce bloc, de taille $2k \times 2k$, est de la forme

$$D := \begin{pmatrix} z & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & z & 1 & \cdots & 0 & 0 & \vdots & \ddots & & & & & \vdots \\ 0 & 0 & z & \ddots & 0 & 0 & \vdots & & \ddots & & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \cdots & z & 1 & \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & z & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \bar{z} & 1 & 0 & \cdots & 0 & 0 & & \\ \vdots & \ddots & & & \vdots & 0 & \bar{z} & 1 & \cdots & 0 & 0 & & \\ \vdots & & \ddots & & \vdots & 0 & 0 & \bar{z} & \ddots & 0 & 0 & & \\ \vdots & & & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \\ \vdots & & & & \ddots & \vdots & 0 & 0 & 0 & \cdots & \bar{z} & 1 & \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \bar{z} & \end{pmatrix} \quad (\text{A.6})$$

lorsque la base \mathbf{a} , dans cet ordre, les vecteurs $v_i = x_i + \mathbf{i}y_i$ ($i = 1, \dots, k$) puis les vecteurs $w_i = y_i + \mathbf{i}x_i$. Si l'on change l'ordre pour que chaque v_i soit suivi de son w_i correspondant (voir le cas des matrices diagonalisables), on obtient le bloc \bar{D} suivant :

$$\bar{D} := \begin{pmatrix} z & 0 & 1 & 0 & 0 & \dots & & & & \\ 0 & \bar{z} & 0 & 1 & 0 & \dots & & & & \\ & 0 & z & 0 & 1 & 0 & \dots & & & \\ & 0 & 0 & \bar{z} & 0 & 1 & \dots & & & \\ 0 & 0 & z & 0 & \dots & & & & & \\ 0 & 0 & 0 & \bar{z} & \dots & & & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ & & & & & & \dots & z & 0 & 1 & 0 \\ & & & & & & \dots & 0 & \bar{z} & 0 & 1 \\ & & & & & & \dots & & & & \\ & & & & & & \dots & 0 & z & 0 & \\ & & & & & & \dots & & 0 & 0 & \bar{z} \end{pmatrix}$$

Puisque les v_i et les w_i forment une base, leurs parties réelles, les x_i et y_i , forment également une base. Au vu des égalités (A.5), en passant à cette nouvelle base, on obtient le bloc ci-après :

$$\tilde{D} := \begin{pmatrix} r & -s & 1 & 0 & 0 & \dots & & & & \\ s & r & 0 & 1 & 0 & \dots & & & & \\ & 0 & r & -s & 1 & 0 & \dots & & & \\ & 0 & s & r & 0 & 1 & \dots & & & \\ 0 & 0 & r & -s & \dots & & & & & \\ 0 & 0 & s & r & \dots & & & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ & & & & & & \dots & r & -s & 1 & 0 \\ & & & & & & \dots & s & r & 0 & 1 \\ & & & & & & \dots & & & & \\ & & & & & & \dots & 0 & r & -s & \\ & & & & & & \dots & & s & r \end{pmatrix} \quad (\text{A.7})$$

Soit \tilde{J} la matrice J Jordanisée de A où l'on a remplacé le bloc D par le bloc \tilde{D} . Alors, si X est la matrice de type $n \times 2k$ définie colonne par colonne par

$$X := (x_1, y_1, \dots, x_k, y_k)$$

on obtient l'égalité $AX = \tilde{J}X$. Effectuant cette opération pour chacun des espaces propres généralisés, on ramène toute Jordanisée complexe en une matrice réelle, qu'on peut qualifier de « triangulaire par blocs 2×2 ». On peut montrer, de manière identique au cas de valeurs propres réelles, que A est semblable à une matrice réelle, où tous les blocs de Jordan sont remplacés par des blocs tels que \tilde{D} , et que la matrice de passage permettant d'effectuer cette transformation, est réelle.

A.2 Conséquences sur les solutions aux problèmes de Cauchy

Nous reprenons maintenant les différentes conclusions obtenues lors de la diagonalisation (ou de la Jordanisation) de A , la ramenant sous une forme « quasi-diagonale », et nous voyons ce que cela implique, pour un choix de $u_0 \in \mathbb{R}^n$, sur la solution du problème de Cauchy (I.9) :

$$\begin{cases} \partial_t u = Au \\ u(0) = u_0 \end{cases}$$

En particulier, nous prouvons les deux théorèmes suivants. Le premier donne un résultat d'invariance sur les solutions. Nous commençons par énoncer une proposition qui nous aide à prouver le théorème :

Proposition 35. *Si $u_0 \in \mathbb{R}^n$ se trouve dans un espace propre (généralisé ou non) F , de valeur propre réelle, alors la solution au problème de Cauchy avec u_0 comme condition initiale, a son image entièrement contenue dans F .*

De même, si u_0 est combinaison linéaire de vecteurs x_i et y_i , de telle sorte que tous les $x_i + \mathbf{i}y_i$ soient vecteurs propres d'une même valeur propre complexe, et si F l'espace vectoriel engendré par les x_i et les y_i , alors nous conservons la même conclusion.

Il en est de même si l'on fixe une valeur propre complexe, dont le sous-espace propre est engendré par des vecteurs propres généralisés ; dans ce cas, si k est la dimension de l'espace propre associé, si $(x_i + \mathbf{i}y_i)_{i=1}^k$ en est une base, et si, pour tout $i = 1, \dots, k$, l'on fixe :

- E_i le sous-espace propre généralisé, contenant $x_i + \mathbf{i}y_i$, tel que construit en A.1.3 ;
- $(x_j^{(i)} + \mathbf{i}y_j^{(i)})_{j=1}^{m_i}$ une base de vecteurs propres généralisés de E_i , de telle sorte que $x_{m_i}^{(i)} + \mathbf{i}y_{m_i}^{(i)} = x_i + \mathbf{i}y_i$
- et r_i un entier compris entre 1 et m_i

alors, posons F l'espace engendré par les vecteurs $x_j^{(i)}$ et $y_j^{(i)}$, avec $i \in \{1, \dots, k\}$ et $j \in \{r_i, \dots, m_i\}$. Toute solution du problème de Cauchy, dont la condition initiale est dans F , a son image entièrement contenue dans F .

Le théorème s'énonce alors comme suit :

Théorème 36. *Soit F un sous-espace vectoriel de \mathbb{R}^n s'écrivant comme somme directe d'espaces propres généralisés de A . Toute solution du problème de Cauchy dont la condition initiale est dans F , a son image contenue dans F .*

Moyennant des conditions sur les valeurs propres (en fait, sur le signe de leurs parties réelle), on peut déduire des notions de convergence sur les solutions obtenues lorsque la condition initiale est bien choisie. A nouveau, nous énonçons une proposition :

Proposition 37. *Soit $u_0 \in \mathbb{R}^n$ un vecteur propre généralisé de valeur propre réelle ou complexe z . Si la partie réelle de z est strictement négative, alors la solution au problème de Cauchy de condition initiale u_0 , converge vers 0 lorsque $t \rightarrow \infty$; et la réciproque est également vraie. De même, on a équivalence entre « partie réelle de z positive » et « la solution converge vers 0 pour $t \rightarrow -\infty$ ».*

Cette proposition admet aussi une généralisation ; pour pouvoir l'énoncer simplement, fixons quelques notations. Supposons que \mathbb{R}^n soit décomposé en somme directe :

$$\mathbb{R}^n = \bigoplus_{i=1, \dots, k} F_i \oplus \bigoplus_{j=1, \dots, m} G_j,$$

où les espaces F_i et G_j vérifient les propriétés suivantes :

- chaque F_i est espace propre réel généralisé de A de valeur propre λ_i ;
- chaque G_j est engendré par des vecteurs x_l et y_l de telle sorte que tous les $x_l + iy_l$ soient vecteurs propres complexes généralisés de A , pour une même valeur propre z_j ;
- pour tous i et j compris entre 1 et k , si $i \neq j$, alors $\lambda_i \neq \lambda_j$;
- pour tous i et j compris entre 1 et m , si $i \neq j$, alors $z_i \neq z_j$.

L'espace \mathbb{R}^n est alors scindé en somme directe d'espaces propres de A . On considère alors un sous-espace F de \mathbb{R}^n de telle sorte qu'il contienne certains espaces F_i et G_j . Formellement, soient $I \subset \{1, \dots, k\}$ et $J \subset \{1, \dots, m\}$; posons alors

$$F = \bigoplus_{i \in I} F_i \oplus \bigoplus_{j \in J} G_j.$$

La généralisation de la proposition précédente s'énonce alors comme suit :

Théorème 38. *Si tous les λ_i , pour $i \in I$, sont strictement négatifs, ainsi que toutes les parties réelles des z_j , pour $j \in J$, alors toutes les solutions du problème de Cauchy dont les conditions initiales sont dans F , convergent vers 0 quand $t \rightarrow +\infty$. De même, si tous les λ_i sont strictement positifs, ainsi que toutes les parties réelles des z_i , alors toutes les solutions du problème de Cauchy dont les conditions initiales sont dans F , convergent vers 0 quand $t \rightarrow -\infty$.*

Réciproquement, soit $u_0 \in F$; si la solution du problème de Cauchy converge vers 0 lorsque $t \rightarrow +\infty$, alors tous les λ_i et toutes les parties réelles des z_i sont strictement négatives ; de même si la solution tend vers 0 pour $t \rightarrow -\infty$ alors tous les λ_i et toutes les parties réelles des z_i sont strictement positives.

Nous prouvons ces résultats en considérant d'abord le cas où u_0 est un vecteur propre (généralisé ou non). Les quatre cas précédemment discutés nous ont permis de transformer A en une matrice qui lui est semblable, réelle, et « presque triangulaire supérieure » (tous les éléments sont nuls à l'exception éventuelle des éléments situés sur la diagonale, de la bande située en-dessous de celle-ci, et des trois bandes situées juste au-dessus de la diagonale); cela nous permet de donner de façon explicite la solution $u(t)$ du problème de Cauchy avec condition initiale u_0 . Ces discussions prouvent alors les propositions 35 et 37. Nous considérons enfin le cas d'un vecteur u_0 quelconque; ce vecteur se décompose alors en une somme de vecteurs propres sur lesquels on peut appliquer la proposition. Cela permet alors de prouver les théorèmes 36 et 38.

A.2.1 Vecteurs propres de valeur propre réelle

Soit λ une valeur propre réelle de A , F le sous-espace propre associé et $u_0 \in F$. On a vu que ce vecteur est aussi vecteur propre de e^A avec pour valeur propre e^λ . Le problème de Cauchy admet alors une solution unique qu'on note $u(t)$; le théorème (4) donne $u(t) = e^{At} u_0$. Puisque u_0 est un vecteur propre de A , il l'est aussi pour tous les $t^n A^n$, admettant alors pour valeur propre $t^n \lambda^n$. Dès lors, on obtient, pour tout $t \in \mathbb{R}$, $u(t) = e^{\lambda t} u_0 \in F$. On vient ainsi de montrer que *tous les espaces et sous-espaces propres sont invariants* dans le sens où, pour toute solution u de l'équation différentielle (I.6), si elle prend une valeur dans un certain espace ou sous-espace propre, alors son image y est entièrement dans cet espace. La proposition 35 est donc vérifiée dans ce cas; la proposition 37 est aussi vraie vu l'expression de $u(t)$.

A.2.2 Vecteurs propres généralisés de valeur propre réelle

Soit maintenant u_0 un vecteur dans un sous-espace propre généralisé F de valeur propre λ . Soit $d = \dim F$ et $(v_i)_{i=1}^d$ une base de F telle que v_d soit vecteur propre et pour tous les autres i , $Av_i = \lambda v_i + v_{i+1}$. On écrit $u_0 = \sum_{i=1}^d x_i v_i$; cela conduit aux relations :

$$\begin{aligned} Au_0 &= \lambda u_0 + u_1 \\ A^2 u_0 &= \lambda(Au_0) + Au_1 \\ &= \lambda^2 u_0 + \lambda u_1 + (\lambda u_1 + u_2) \\ &= \lambda^2 u_0 + 2\lambda u_1 + u_2 \\ A^3 u_0 &= \lambda^3 u_0 + \lambda^2 u_1 + 2\lambda(\lambda u_1 + u_2) + (\lambda u_2 + u_3) \\ &= \lambda^3 u_0 + 3\lambda^2 u_1 + 3\lambda u_2 + u_3 \end{aligned}$$

$$\begin{aligned}
A^4 u_0 &= \lambda^4 u_0 + \lambda^3 u_1 + 3\lambda^2(\lambda u_1 + u_2) + 3\lambda(\lambda u_2 + u_3) + (\lambda u_3 + u_4) \\
&= \lambda^4 u_0 + 4\lambda^3 u_1 + 6\lambda^2 u_2 + 4\lambda u_3 + u_4
\end{aligned}$$

où, pour tout i , on montre que

$$u_i = \sum_{j=i}^d x_{j-i} v_j. \quad (\text{A.8})$$

On voit se dégager la formule suivante :

$$A^i u_0 = \sum_{j=0}^i \binom{i}{j} \lambda^{i-j} u_j \quad (\text{A.9})$$

Prouvons-la par récurrence. Nous avons déjà montré cas de base ci-dessus. Pour l'induction, on a :

$$\begin{aligned}
A^{i+1} u_0 &= \sum_{j=0}^i \binom{i}{j} \lambda^{i-j} A u_j \\
&= \sum_{j=0}^i \binom{i}{j} \lambda^{i-j} (\lambda u_j + u_{j+1}) \\
&= \sum_{j=0}^i \binom{i}{j} \lambda^{i-j+1} u_j + \sum_{j=0}^i \binom{i}{j} \lambda^{i-j} u_{j+1};
\end{aligned}$$

en sortant les termes extrêmes pour $j = 0$ dans la première somme et $j = i$ dans la seconde, on a

$$A^{i+1} u_0 = \sum_{j=1}^i \binom{i}{j} \lambda^{i-j+1} u_j + \lambda^{i+1} u_0 + \sum_{j=0}^{i-1} \binom{i}{j} \lambda^{i-j} u_{j+1} + u_{i+1};$$

en réordonnant, et en modifiant les indices de la seconde somme :

$$\begin{aligned}
A^{i+1} u_0 &= \lambda^{i+1} u_0 + u_{i+1} + \sum_{j=1}^i \binom{i}{j} \lambda^{i-j+1} u_j + \sum_{j=1}^i \binom{i}{j-1} \lambda^{i-j+1} u_j \\
&= \lambda^{i+1} u_0 + u_{i+1} + \sum_{j=1}^i \binom{i+1}{j} \lambda^{i+1-j} u_j \\
&= \sum_{j=0}^{i+1} \binom{i+1}{j} \lambda^{i+1-j} u_j \quad (\text{A.10})
\end{aligned}$$

Cette formule est valable même pour les $i > d$; mais dans ce cas, $u_i = 0$.

Conservons le même u_0 que ci-dessus ; on considère le problème de Cauchy (I.9), et sa solution u . Au vu de ci-dessus, on a pour tout t :

$$\begin{aligned}
u(t) &= e^{At} u_0 = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n u_0 \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} \left(\sum_{j=0}^n \binom{n}{j} \lambda^{n-j} u_j \right) && \text{par la formule (A.9) démontrée ci-dessus} \\
&= \sum_{n=0}^d \frac{t^n}{n!} \left(\sum_{j=0}^n \binom{n}{j} \lambda^{n-j} u_j \right) \\
&\quad + \sum_{n=d+1}^{\infty} \frac{t^n}{n!} \left(\sum_{j=0}^d \binom{n}{j} \lambda^{n-j} u_j \right) && \text{car pour } n > d, u_n = 0,
\end{aligned}$$

et, en intervertissant dans les deux termes, les symboles sommatoires :

$$\begin{aligned}
u(t) &= \sum_{j=0}^d \sum_{n=j}^d \frac{t^n}{n!} \binom{n}{j} \lambda^{n-j} u_j + \sum_{j=0}^d \sum_{n=d+1}^{\infty} \frac{t^n}{n!} \binom{n}{j} \lambda^{n-j} u_j \\
&= \sum_{j=0}^d \left(\sum_{n=j}^{\infty} \binom{n}{j} \frac{t^n}{n!} \lambda^{n-j} \right) u_j \\
&= \sum_{j=0}^d \left(\sum_{n=j}^{\infty} \frac{n!}{j!(n-j)!} \frac{t^j t^{n-j}}{n!} \lambda^{n-j} \right) u_j \\
&= \sum_{j=0}^d \frac{t^j}{j!} \left(\sum_{n=j}^{\infty} \frac{t^{n-j}}{(n-j)!} \lambda^{n-j} \right) u_j \\
&= \sum_{j=0}^d \frac{t^j}{j!} e^{\lambda t} u_j && \text{(A.11)}
\end{aligned}$$

Puisque les u_i sont dans le même espace F que u_0 (voir leur expression — (A.8)), on obtient que pour tout $t, u(t) \in F$. Dès lors, la proposition 35 est vérifiée, et les espaces propres généralisés sont invariants. Par ailleurs, l'expression A.11 donne

$$u(t) = e^{\lambda t} \sum_{j=0}^d \frac{t^j}{j!} u_j \quad \text{(A.12)}$$

et la conclusion de la proposition 37 s'en déduit.

A.2.3 Vecteurs propres de valeurs propres complexes

Pour rappel, on a transformé une matrice diagonale semblable à A à coefficients parfois complexes en une matrice « presque diagonale », mais à coefficients réels (voir la relation (A.1) devenue (A.4) grâce aux blocs réels \tilde{D}_i — cf. (A.3)). Or, chacun de ces blocs peut s'écrire :

$$\tilde{D} = \begin{pmatrix} r & -s \\ s & r \end{pmatrix} = r\text{Id} + s \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

(où l'on a retiré les indices pour un souci de lisibilité). Notons

$$S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{A.13}$$

Alors, on a

$$\begin{aligned} SS &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\text{Id} \\ S^3 &= -S \\ S^4 &= \text{Id} \end{aligned}$$

Dès lors, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} e^{\tilde{D}t} &= e^{rt\text{Id} + stS} \\ &= e^{rt\text{Id}} e^{stS} \end{aligned}$$

et le deuxième facteur donne :

$$\begin{aligned} e^{stS} &= \sum_{n=0}^{\infty} \frac{(st)^n S^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(st)^{4n} S^{4n}}{(4n)!} \left(\text{Id} + \frac{stS}{4n+1} + \frac{(st)^2 S^2}{(4n+1)(4n+2)} \right. \\ &\quad \left. + \frac{(st)^3 S^3}{(4n+1)(4n+2)(4n+3)} \right) \\ &= \sum_{n=0}^{\infty} \frac{(st)^{4n} \text{Id}^n}{(4n)!} \left(\text{Id} + \frac{stS}{4n+1} - \frac{(st)^2 \text{Id}}{(4n+1)(4n+2)} \right. \\ &\quad \left. - \frac{(st)^3 S}{(4n+1)(4n+2)(4n+3)} \right) \\ &= \text{Id} \left(\sum_{n=0}^{\infty} \frac{(st)^{4n}}{(4n)!} - \sum_{n=0}^{\infty} \frac{(st)^{4n+2}}{(4n+2)!} \right) + S \left(\sum_{n=0}^{\infty} \frac{(st)^{4n+1}}{(4n+1)!} - \sum_{n=0}^{\infty} \frac{(st)^{4n+3}}{(4n+3)!} \right) \\ &= \text{Id} \sum_{n=0}^{\infty} (-1)^n \frac{(st)^{2n}}{(2n)!} + S \sum_{n=0}^{\infty} (-1)^n \frac{(st)^{2n+1}}{(2n+1)!} \end{aligned}$$

car, pour la première somme, on somme sur tous les indices pairs, mais de façon alternée ; et un argument similaire est valable pour la seconde somme. On remarque ainsi que la première somme est le développement en série de $\cos(st)$; tandis que la seconde est le développement de $\sin(st)$; si bien que

$$e^{stS} = \begin{pmatrix} \cos(st) & -\sin(st) \\ \sin(st) & \cos(st) \end{pmatrix}, \quad \text{d'où } e^{\tilde{D}t} = e^{rt} \begin{pmatrix} \cos(st) & -\sin(st) \\ \sin(st) & \cos(st) \end{pmatrix}. \quad (\text{A.14})$$

Supposons que $u_0 = ax_0 + by_0$ soit un vecteur pour lequel $x_0 + iy_0$ est vecteur propre de A (dans \mathbb{C}^n), de valeur propre $z = r + is$. Les relations (A.2) impliquent $Ax_0 = rx_0 - sy_0$ et $Ay_0 = sx_0 + ry_0$; ou sous une forme concise, $A(x_0, y_0) = (x_0, y_0)\tilde{D}$. Dès lors, on obtient, pour tout $t \in \mathbb{R}$, l'égalité $e^{At}(x_0, y_0) = (x_0, y_0)e^{\tilde{D}t}$. Les calculs réalisés ci-dessus pour $e^{\tilde{D}t}$ donnent alors les deux relations :

$$\begin{cases} e^{At}x_0 = e^{rt}(\cos(st)x_0 - \sin(st)y_0) \\ e^{At}y_0 = e^{rt}(\cos(st)y_0 + \sin(st)x_0) \end{cases} \quad (\text{A.15})$$

Dès lors, $e^{At}u_0 = ae^{At}x_0 + be^{At}y_0$, et les égalités ci-dessus donnent :

$$e^{At}u_0 = e^{rt}((a\cos(st) - b\sin(st))x_0 + ((b\cos(st) + a\sin(st))y_0), \quad (\text{A.16})$$

d'où l'on tire que, pour tout t , $e^{At}u_0$ est une combinaison linéaire des vecteurs x_0 et y_0 ; la proposition 35 est à nouveau vérifiée. Par ailleurs, on remarque que l'expression obtenue pour $e^{At}u_0$ est majorée en norme par $e^{rt}(\|x_0\| + \|y_0\|)$; nous avons donc bien convergence vers 0 pour $r < 0$. Réciproquement, si l'on suppose $r \geq 0$, alors, pour tout $t \geq 0$, on a

$$\|e^{rt}(\cos(st)x_0 - \sin(st)y_0)\| \geq \|\cos(st)x_0 - \sin(st)y_0\|,$$

cette dernière prenant la valeur constante $\|y_0\| \neq 0$ en tous les t de la forme $\frac{4k\pi + \pi}{2s}$, $k \in \mathbb{Z}$, et ne pouvant donc tendre vers 0 pour $t \rightarrow \infty$. Effectuant un raisonnement similaire pour la convergence lorsque $t \rightarrow -\infty$, cela prouve, dans ce cas, la proposition 37.

A.2.4 Vecteurs propres généralisés de valeurs propres complexes

Soit $(v_i), i = 1, \dots, k$ une base de vecteurs propres généralisés de valeur propre $z = r + is$ pour un espace E ; on pose que v_k est un vecteur propre de A et que pour tous les autres indices i , on a $Av_i = zv_i + v_{i+1}$. (La construction de tels vecteurs est

rappelé en section A.1.3.) En posant $v_i = x_i + iy_i$, on a déjà obtenu les relations (A.5) :

$$\begin{cases} Ax_i = rx_i - sy_i + x_{i+1} \\ Ay_i = sx_i + ry_i + y_{i+1} \end{cases}$$

Rappelons aussi que, au vu de ces égalités, on a modifié les blocs de Jordan de la forme (A.6) en blocs réels (A.7). Le raisonnement qui suit est similaire au cas Jordanisable réel ; quelques adaptations sont cependant nécessaires. Soit F l'espace réel engendré par les vecteurs x_i et y_i , et $u_0 \in F$. Posons $u_0 = \sum_{i=0}^k (a_i x_i + b_i y_i)$, et \tilde{D} la matrice définie en (A.3).

On remarque alors les égalités suivantes :

$$\begin{aligned} A(x_0, y_0) &= (rx_0 - sy_0 + x_1, sx_0 + ry_0 + y_1) \\ &= (x_0, y_0)\tilde{D} + (x_1, y_1) \\ A^2(x_0, y_0) &= A(x_0, y_0)\tilde{D} + A(x_1, y_1) \\ &= (x_0, y_0)\tilde{D}^2 + 2(x_1, y_1)\tilde{D} + (x_2, y_2). \end{aligned}$$

Si l'on pose $w_i = (x_i, y_i)$, on peut prouver par récurrence

$$A^i w_0 = \sum_{j=0}^i \binom{i}{j} w_{i-j} \tilde{D}^j;$$

un raisonnement similaire à (A.10) fonctionne (il suffit de remplacer tout λ par la matrice \tilde{D} , et tous les u_i par les w_i ; il faut cependant faire attention à l'ordre dans lequel les objets sont manipulés). Pour tout t , on en déduit alors l'expression de $w^{(0)}(t) = e^{At} w_0$, à nouveau les calculs réalisés au développement (A.11) restent valables, avec le même remplacement ; ce qui conduit à

$$w^{(0)}(t) = (x^{(0)}(t), y^{(0)}(t)) = \sum_{j=0}^d \frac{t^j}{j!} w_j e^{\tilde{D}t}$$

De la même façon, on montre que, pour tout $i = 0, \dots, k$:

$$w^{(i)}(t) = (x^{(i)}(t), y^{(i)}(t)) = \sum_{j=i}^d \frac{t^j}{j!} w_j e^{\tilde{D}t}$$

Grâce à la relation (A.14), l'égalité précédente peut se récrire :

$$\begin{aligned} (x^{(i)}(t), y^{(i)}(t)) &= \sum_{j=i}^d \frac{t^j}{j!} e^{rt} (\cos(st)x_j - \sin(st)y_j, \cos(st)y_j + \sin(st)x_j) \\ &= e^{rt} \sum_{j=i}^d \frac{t^j}{j!} (\cos(st)x_j - \sin(st)y_j, \cos(st)y_j + \sin(st)x_j) \end{aligned}$$

et l'on a donc les deux égalités

$$\begin{aligned}
x^{(i)}(t) &= e^{rt} \left(\cos(st) \sum_{j=i}^d \frac{t^j}{j!} x_j - \sin(st) \sum_{j=i}^d \frac{t^j}{j!} y_j \right) \\
&= e^{rt} \left(\cos(st) \bar{x}^{(i)}(t) - \sin(st) \bar{y}^{(i)}(t) \right) \\
y^{(i)}(t) &= e^{rt} \left(\cos(st) \sum_{j=i}^d \frac{t^j}{j!} y_j + \sin(st) \sum_{j=i}^d \frac{t^j}{j!} x_j \right) \\
&= e^{rt} \left(\cos(st) \bar{y}^{(i)}(t) + \sin(st) \bar{x}^{(i)}(t) \right)
\end{aligned}$$

Nous obtenons alors toutes les solutions lorsque la condition initiale est un vecteur de base. La solution $u(t)$ du problème de Cauchy avec condition initiale u_0 s'obtient alors par combinaison linéaire : puisque $u_0 = \sum_{i=0}^k (a_i x_i + b_i y_i)$, on a

$$\begin{aligned}
u(t) &= \sum_{i=0}^k (a_i x^{(i)}(t) + b_i y^{(i)}(t)) \\
&= \sum_{i=0}^k \left(a_i e^{rt} (\cos(st) \bar{x}^{(i)}(t) - \sin(st) \bar{y}^{(i)}(t)) \right. \\
&\quad \left. + b_i e^{rt} (\cos(st) \bar{y}^{(i)}(t) + \sin(st) \bar{x}^{(i)}(t)) \right) \\
&= e^{rt} \left(\sum_{i=0}^k \bar{x}^{(i)}(t) (a_i \cos(st) - b_i \sin(st)) \right. \\
&\quad \left. + \sum_{i=0}^k \bar{y}^{(i)}(t) (a_i \sin(st) + b_i \cos(st)) \right);
\end{aligned}$$

chacun des termes se développe de la façon suivante :

$$\begin{aligned}
\sum_{i=0}^k \bar{x}^{(i)}(t) (a_i \cos(st) - b_i \sin(st)) &= \sum_{i=0}^k (a_i \cos(st) - b_i \sin(st)) \left(\sum_{j=i}^d \frac{t^j}{j!} x_j \right) \\
&= \sum_{i=0}^k \sum_{j=i}^d (a_i \cos(st) - b_i \sin(st)) \frac{t^j}{j!} x_j \\
&= \sum_{j=0}^d \sum_{i=j}^k (a_i \cos(st) - b_i \sin(st)) \frac{t^j}{j!} x_j \\
&= \sum_{j=0}^d (\tilde{a}_j \cos(st) - \tilde{b}_j \sin(st)) \frac{t^j}{j!} x_j \quad (\text{A.17})
\end{aligned}$$

où $\tilde{a}_j = \sum_{i=j}^k a_i$ et $\tilde{b}_j = \sum_{i=j}^k b_i$; de la même façon, on a l'égalité

$$\sum_{i=0}^k \bar{y}^{(i)}(t)(a_i \sin(st) + b_i \cos(st)) = \sum_{j=0}^d (\tilde{b}_j \cos(st) + \tilde{a}_j \sin(st)) \frac{t^j}{j!} y_j.$$

Dès lors, l'expression de $u(t)$ est donnée par

$$\begin{aligned} u(t) &= e^{rt} \sum_{j=0}^d \frac{t^j}{j!} (\cos(st)(\tilde{a}_j x_j + \tilde{b}_j y_j) + \sin(st)(\tilde{a}_j y_j - \tilde{b}_j x_j)) \\ &= e^{rt} \cos(st) \sum_{j=0}^d \frac{t^j}{j!} (\tilde{a}_j - \tilde{b}_j) x_j + e^{rt} \sin(st) \sum_{j=0}^d \frac{t^j}{j!} (\tilde{a}_j + \tilde{b}_j) y_j \\ &= e^{rt} (\cos(st) \bar{x}(t) + \sin(st) \bar{y}(t)); \end{aligned} \tag{A.18}$$

nous utiliserons cette écriture concise pour le cas général. Notons toutefois que $\bar{x}(t)$ n'est autre que $\bar{x}^{(0)}(t)$, et de même $\bar{y}(t) = \bar{y}^{(0)}(t)$.

Nous pouvons alors conclure sur la proposition 35, puisque cette dernière relation met en évidence que, pour tout t , $u(t) \in F$. De plus, si les premiers a_i et b_i sont nuls, disons les r premiers, l'adaptation du calcul (A.17) fait apparaître que la somme pour u démarre à $r + 1$ au lieu de 0; en conséquence, on obtient que, si u_0 est engendré par une partie des vecteurs (disons, les $2s$ derniers), alors $u(t)$ reste toujours dans l'espace engendré par ces $2s$ vecteurs.

Un raisonnement similaire à ce qui a été fait dans le cas « diagonalisable à valeurs propres complexes », montre que la proposition 37 est aussi vraie.

A.2.5 Vecteurs quelconques

Finissons par le cas général d'un vecteur u_0 qui ne soit ni vecteur propre, ni vecteur propre généralisé de A . Supposons dans un premier temps que A soit diagonalisable; alors u_0 s'exprime de manière unique sous la forme

$$u_0 = \sum_{i=0}^k v_i + \sum_{j=0}^m (b_j x_j + c_j y_j), \tag{A.19}$$

avec $v_i \in F_i$ pour tout i , F_i étant un espace propre de A de valeur propre λ_i ; et $x_j \in G_j$ et $y_j \in G_j$ sont tels que $x_j + \mathbf{i}y_j$ soit le seul parmi les $x_l + \mathbf{i}y_l$ à être vecteur propre de A de valeur propre complexe $z_j = r_j + \mathbf{i}s_j$.

Les calculs réalisés précédemment donnent explicitement la solution au problème de Cauchy pour chacun des v_i , x_j et y_j ; cela nous permet d'en déduire celle pour la solution de condition initiale u_0 par linéarité: nous avons

$$\begin{aligned} u(t) &= \sum_{i=0}^k a_i e^{\lambda_i t} v_i + \sum_{j=0}^m e^{r_j t} ((b_j \cos(s_j t) + c_j \sin(s_j t)) x_j \\ &\quad + (c_j \cos(s_j t) - b_j \sin(s_j t)) y_j. \end{aligned}$$

Si F est la somme directe des espaces F_i et G_j , alors $u_0 \in F$ et, par l'égalité ci-dessus, $u(t) \in F$ pour tout $t \geq 0$.

Si maintenant A n'est que Jordanisable, nous conservons l'égalité (A.19) pour v_i et $x_j + iy_j$ vecteurs propres généralisés ; cependant, la forme que prend $u(t)$ doit être modifiée. Mais, si l'on pose respectivement $v_{(i)}$ et $w_{(j)}$ les solutions obtenues pour le problème de Cauchy lorsque la condition initiale est v_i , ou respectivement $b_j x_j + c_j y_j$, la linéarité implique

$$u(t) = \sum_{i=0}^k a_i v_{(i)}(t) + \sum_{j=0}^m w_{(j)}(t);$$

mais, grâce aux égalités (A.12) et (A.18), les $v_{(i)}(t)$ et les $w_{(j)}(t)$ s'expriment de la façon suivante :

$$\begin{cases} v_{(i)}(t) = e^{\lambda_i t} \bar{v}_i(t) \\ w_{(j)}(t) = e^{r_j t} (\cos(s_j t) \bar{x}_j(t) + \sin(s_j t) \bar{y}_j(t)) \end{cases}$$

Dès lors, l'expression de $u(t)$ devient :

$$u(t) = \sum_{i=0}^k a_i e^{\lambda_i t} \bar{v}_i(t) + \sum_{j=0}^m e^{r_j t} (\cos(s_j t) \bar{x}_j(t) + \sin(s_j t) \bar{y}_j(t)).$$

De par la construction des fonctions $\bar{v}_i(t)$, $\bar{x}_j(t)$ et $\bar{y}_j(t)$, le résultat d'invariance général, énoncé au théorème 36, s'en déduit. Par ailleurs, à nouveau grâce à la forme des fonctions $\bar{v}_i(t)$, $\bar{x}_j(t)$ et $\bar{y}_j(t)$ (qui sont polynômiales en t), le résultat de convergence (théorème 38) peut être prouvé.

Annexe B

Variétés différentiables

Cette annexe couvre tous les points ayant trait aux variétés différentiables. De façon concise, on peut dire qu'une variété différentiable est un ensemble qui se comporte localement comme \mathbb{R}^n , pour $n \geq 1$ fixé; ce n est alors la dimension de la variété. Nous énonçons les principaux faits concernant les variétés différentiables, et nous prouvons un résultat que nous utilisons par deux fois dans le chapitre consacré aux *elastica*. Pour approfondir davantage, on pourra consulter, par exemple, les ouvrages de D. Barden et C. Thomas [1], de S. Lang [8] ou de S. T. Hu [7].

B.1 Définitions et propriétés élémentaires

Définition 15. Soit X un ensemble. Une *carte sur X* est la donnée d'un couple (U, ϕ) où U est un sous-ensemble de X qu'on appellera *domaine* de la carte et ϕ une application définie sur U à valeurs dans \mathbb{R}^n , pour un certain $n \geq 1$.

Une carte définit dès lors, localement sur X , une topologie : la topologie initiale provenant de la topologie usuelle sur \mathbb{R}^n . On souhaite maintenant « recoller » les morceaux, afin de pouvoir définir une topologie sur tout X . Cela passe par des conditions sur les cartes dont les domaines ont une partie commune.

Définition 16. Un *atlas* sur X est une collection de cartes $\{(U_i, \phi_i)\}_{i \in I}$, chaque ϕ_i étant à valeurs dans \mathbb{R}^n , cette collection vérifiant de plus les conditions suivantes :

- $\bigcup_{i \in I} U_i = X$;
- pour tous $i, j \in I$, si U_i et U_j sont d'intersection non-vide, alors la fonction $\phi_i \circ \phi_j^{-1}$ définie sur $\phi_j(U_i \cap U_j)$ et à valeurs dans $\phi_i(U_i \cap U_j)$, est un homéomorphisme entre deux ouverts de \mathbb{R}^n .

Si de plus, il existe un naturel r pour lequel toutes les applications $\phi_i \circ \phi_j^{-1}$ sont de classe C^r , alors l'atlas est dit *de classe C^r* .

On définit alors que deux atlas (de même classe C^r) sont équivalents si leur union est encore un atlas ; c'est équivalent à imposer le fait suivant : si (U, ϕ) est une carte du premier atlas, et (V, ψ) une carte du second, tels que $U \cap V \neq \emptyset$, alors la fonction $\psi \circ \phi^{-1}$, définie sur les ensembles adéquats, est de classe C^r .

On fixe alors ce qui est communément appelé structure différentiable sur X :

Définition 17. Une *structure différentiable sur X* est la donnée d'une classe d'équivalence d'atlas, pour la relation d'équivalence évoquée ci-dessus.

Un argument traditionnel, utilisant le lemme de Zorn, montre que toute structure différentiable admet un atlas maximal ; il consiste précisément en l'union de tous les atlas présents dans la structure différentiable. Dès lors, on se soucie peu de la technique évoquée, utilisant simplement, localement, une fonction ϕ permettant de passer de l'ensemble X vers \mathbb{R}^n . Mais, formellement, allons jusqu'au bout :

Définition 18. Une *variété différentiable* est la donnée d'un ensemble X , que l'on munit d'une structure différentiable. On dit qu'elle est de *dimension n* et de *classe C^r* si tous les atlas de la structure différentiable sont de classe C^r et chacune des fonctions ϕ_i est à valeurs dans \mathbb{R}^n .

Illustrons ces définitions dans des exemples concrets. D'abord, considérons le cercle unité, dans \mathbb{R}^2 : c'est une variété différentiable, de dimension 1. Quatre applications suffisent pour décrire la variété, par exemple :

$$\begin{aligned} \text{pour les points d'abscisse positive : } & \phi_1(x, y) = y, \phi_1^{-1}(y) = (\sqrt{1-y^2}, y), \\ \text{pour les points d'abscisse négative : } & \phi_2(x, y) = y, \phi_2^{-1}(y) = (-\sqrt{1-y^2}, y), \\ \text{pour les points d'ordonnée positive : } & \phi_3(x, y) = x, \phi_3^{-1}(x) = (x, \sqrt{1-x^2}), \\ \text{pour les points d'ordonnée négative : } & \phi_4(x, y) = x, \phi_4^{-1}(x) = (x, -\sqrt{1-x^2}), \end{aligned}$$

et l'on montre, par exemple, que $\phi_3 \circ \phi_1^{-1}$ est un difféomorphisme de classe C^∞ sur l'intervalle $]0, 1[$; il s'agit en fait de la fonction $y \mapsto \sqrt{1-y^2}$. Pour prouver la proposition qui suit, nous allons utiliser une autre paramétrisation. On considère les ensembles

$$\begin{aligned} U_1 &= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \text{ et } x > \frac{1}{2}\}, \\ U_2 &= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \text{ et } y > \frac{1}{2}\}, \\ U_3 &= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \text{ et } x < -\frac{1}{2}\}, \\ U_4 &= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \text{ et } y < -\frac{1}{2}\}, \end{aligned} \tag{B.1}$$

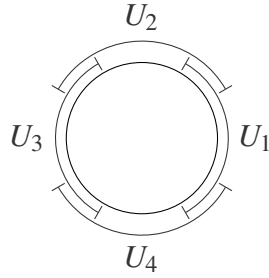


FIG. B.1 – Paramétrisation du cercle avec les ensembles U_i .

dont l'union recouvre bien le cercle. On définit, sur chacun de ces ensembles, une application ψ_i définie comme suit :

$$\begin{aligned} \text{pour } U_1, \psi_1(x, y) &= \arcsin y, \\ \text{pour } U_2, \psi_2(x, y) &= \arccos x, \\ \text{pour } U_3, \psi_3(x, y) &= \pi - \arcsin y, \\ \text{pour } U_4, \psi_4(x, y) &= 2\pi - \arccos x. \end{aligned}$$

Il est facile de constater que les applications inverses sont toutes de la forme $t \mapsto (\cos t, \sin t)$; par ailleurs les applications permettant les changements de cartes sont exactement l'identité, sauf celles qui permettent le passage de U_1 à U_4 , qui, elles, translatent de 2π (formellement, $\psi_1 \circ \psi_4^{-1}$ est l'application $x \mapsto x - 2\pi$). Les atlas définis d'une part par les applications ϕ_i , et d'autre part, par les applications ψ_j , sont équivalents. Sans le prouver formellement, signalons quand même que $\phi_1 \circ \psi_1^{-1}$ n'est autre que le sinus (dont le domaine est restreint de manière adéquate), et que, par exemple, $\phi_3 \circ \psi_1^{-1}$ est le cosinus (à nouveau restreint sur le bon domaine) ; les autres fonctions de passage sont similaires.

Considérons, pour autre exemple, la sphère unité, dans \mathbb{R}^3 (et plus généralement, dans \mathbb{R}^n). On peut la projeter dans \mathbb{R}^2 (resp. dans \mathbb{R}^{n-1}) d'une façon similaire au cercle. En considérant des parties de la sphère de la même manière que précédemment, on définit des cartes ; en général, pour les points de la sphère dont une composante x_i est positive, on considère l'application qui omet cette composante ; l'inverse de cette fonction récupère alors la valeur de x_i par la relation $x_i = \sqrt{1 - \sum_{j=1, \dots, n; j \neq i} x_j^2}$. Dans le cas de la sphère de \mathbb{R}^3 , cette relation devient par exemple, si $i = 1$: $x_1 = \sqrt{1 - x_2^2 - x_3^2}$.

B.2 Une proposition concernant le cercle unité

Proposition 39. Soient f et g deux fonctions $\mathbb{R} \rightarrow \mathbb{R}$, de classe C^r , $r \geq 0$, telles que, pour tout $x \in \mathbb{R}$, $(f(x))^2 + (g(x))^2 = 1$. Il existe une fonction $\psi \in C^r(\mathbb{R}, \mathbb{R})$ et un réel $\psi_0 \in [0, 2\pi[$ tels que $\psi(0) = \psi_0$ et, pour tout x , $f(x) = \cos(\psi(x))$ et $g(x) = \sin(\psi(x))$

Démonstration. Commençons par choisir ψ_0 . Le point $(f(0), g(0))$ est sur le cercle unité et donc, dans l'un des U_i définis par les relations (B.1). S'il se trouve dans U_2, U_3 , ou U_4 , on fixe $\psi_0 = \psi_i(f(0), g(0))$; les définitions de ces fonctions montrent alors que ψ_0 est dans l'intervalle désiré. Si par contre le point se trouve dans U_1 , il faut mieux choisir l'angle. On peut supposer que le point n'est que dans U_1 et non dans un autre U_i , sans quoi l'argument précédent convient. Sous cette condition, c'est le signe de $g(0)$ qui détermine l'angle initial. Si $g(0)$ est positif (ou nul), on prend $\psi_0 = \psi_1(f(0), g(0)) \geq 0$; dans le cas contraire, on ajoute 2π , de telle sorte que $\psi_0 = \psi_1(f(0), g(0)) + 2\pi < 2\pi$.

L'angle initial ainsi choisi, on voudrait définir ψ de façon locale, via les cartes, puis montrer que globalement, cela définit bien ψ , et qu'elle est de bonne classe. Cependant, à cause des définitions de ψ_1 et ψ_4 , (la transformation $\psi_1 \circ \psi_4^{-1}$ n'étant pas l'identité), cela peut engendrer des discontinuités; celles-ci surviennent lorsque $(f(x), g(x))$ décrit un tour complet du cercle. L'idée est donc de compter ce nombre de tours; nous le ferons de manière relative, en considérant qu'en 0, nous n'avons effectué aucun tour.

Localement donc, on peut définir une fonction $\tilde{\psi}$ de la façon suivante. Soit $x \in \mathbb{R}$; le point $(f(x), g(x))$ se trouve dans l'un des U_i ; éventuellement, il peut aussi se trouver dans un U_j , et l'on peut supposer $j > i$. Si ce n'est pas le cas, on prend $\tilde{\psi}(x) = \psi_i(f(x), g(x))$. Sinon, si $j = i + 1$ (de telle sorte que x se trouve à l'intersection des ensembles U_1 et U_2 , ou U_2 et U_3 , ou U_3 et U_4), on prend la même définition de $\tilde{\psi}$. Puisque $\psi_i \circ \psi_j^{-1}$ est l'identité sur les ouverts adéquats, clairement $\psi_i = \psi_j$ sur $U_i \cap U_j$, et la définition de $\tilde{\psi}$ n'est donc pas ambiguë. Il reste donc le cas $x \in U_1 \cap U_4$ à traiter. La continuité des fonctions f et g implique l'existence de deux intervalles ouverts I et J contenant x et tels que tout point de I (resp. de J) est envoyé, par (f, g) , dans U_1 (resp. dans U_4). Sans perte de généralité, supposons-les maximaux pour leur propriétés respectives; un argument courant faisant appel au lemme de Zorn permet de le montrer. La définition que nous allons prendre pour $\tilde{\psi}$ dépendra de la façon dont (f, g) est venu dans l'intersection $U_1 \cap U_4$ et comment il en sort; cela peut être connu par la position relative des intervalles I et J par rapport à leur intersection, qui est aussi un intervalle ouvert.

Considérons $I_- := \{y \in I | y < x\}$, $I_+ := \{y \in I | y > x\}$, $J_- := \{y \in J | y < x\}$, $J_+ := \{y \in J | y > x\}$. Nous définissons $\tilde{\psi}(x)$ (et, plus généralement, $\tilde{\psi}(y)$ pour $y \in I \cap J$) de la façon suivante.

- Si $I_- \subsetneq J_-$, alors (f, g) vient de U_4 ; on définit : $\tilde{\psi}(x) := \psi_4(x) = 2\pi + \psi_1(x)$.
- si $I_- \supsetneq J_-$, alors (f, g) vient de U_1 ; on définit : $\tilde{\psi}(x) := \psi_1(x) = \psi_4(x) - 2\pi$.
- Si $I_- = J_-$, alors, c'est le comportement futur qui impose le choix de $\tilde{\psi}$:
 - si $I_+ \subsetneq J_+$, (f, g) sort par U_4 , et l'on pose alors $\tilde{\psi} = \psi_4(x) = 2\pi + \psi_1(x)$;
 - si $I_+ \supsetneq J_+$, (f, g) sort par U_1 , et l'on pose alors $\tilde{\psi} = \psi_1(x) = \psi_4(x) - 2\pi$;
 - si $I_+ = J_+$, (f, g) reste toujours dans $U_1 \cap U_4$, et l'on pose alors $\tilde{\psi} = \psi_4(x) = 2\pi + \psi_1(x)$.

$\tilde{\psi}$ ainsi définie, s'avère être de même classe que f et g , à l'exception éventuelle des points x pour lesquels f vaut $\frac{1}{2}$ et g est négative, ou g vaut $-\frac{1}{2}$ et f est positive (cela correspond aux points où l'on quitte l'intersection $U_1 \cap U_4$). On remarque que si $\tilde{\psi}$ est discontinue en un point x , alors on peut donner explicitement la valeur de $\tilde{\psi}(x)$: si nous sommes dans la première configuration, alors $\tilde{\psi}(x) = -\frac{\pi}{6}$, si nous sommes dans la seconde configuration, alors $\tilde{\psi}(x) = \frac{10\pi}{6}$.

On définit enfin une fonction k pour compter le nombre de tours, de telle sorte que $\psi = \tilde{\psi} + 2\pi k$ soit continue et de même classe que f et g , comme suit :

- $k(0) = 0$;
- k est constant sur tout intervalle où $\tilde{\psi}$ est continue ;
- en tout point x de discontinuité de $\tilde{\psi}$, soit $\varepsilon > 0$ tel que $\tilde{\psi}$ soit continue à la fois sur $]x - \varepsilon, x[$ et $]x, x + \varepsilon[$. Si k a été définie sur $]x - \varepsilon, x[$, on l'étend sur $]x, x + \varepsilon[$ de la façon suivante :
 - si $\tilde{\psi}(x) = -\frac{\pi}{6}$, alors $k(y) = k(x) = k(x - \frac{\varepsilon}{2}) + 1$, pour tout $y \in [x, x + \varepsilon[$;
 - sinon, $\tilde{\psi}(x) = \frac{10\pi}{6}$, et l'on définit $k(y) = k(x) = k(x - \frac{\varepsilon}{2}) - 1$, pour tout $y \in [x, x + \varepsilon[$.
- De la même façon, si k a été définie sur $]x, x + \varepsilon[$, on l'étend sur $]x - \varepsilon, x[$ de la façon suivante :
 - si $\tilde{\psi}(x) = -\frac{\pi}{6}$, alors $k(y) = k(x) = k(x + \frac{\varepsilon}{2}) - 1$, pour tout $y \in [x, x + \varepsilon[$;
 - sinon, $\tilde{\psi}(x) = \frac{10\pi}{6}$, et l'on définit $k(y) = k(x) = k(x + \frac{\varepsilon}{2}) + 1$, pour tout $y \in [x, x + \varepsilon[$.

On montre à présent qu'avec ces définitions, ψ est bien continue. Elle l'est en tout point où $\tilde{\psi}$ l'est, puisque k est constant sur un voisinage d'un tel point. Aux points x où $\tilde{\psi}$ est discontinue, k effectue un saut également ; mais par construction de k , la continuité de ψ est assurée. \square

B.3 Applications différentiables, vecteurs tangents, différentielles

Etant donnés M et N deux variétés différentiables, on peut s'intéresser aux applications $f : M \rightarrow N$. On peut alors demander une forme de régularité sur f grâce à la structure de M et N .

Définition 19. On dit que $f : M \rightarrow N$ est *différentiable en* $p \in M$ s'il existe deux cartes (U, ϕ) et (V, ψ) , respectivement de M et de N , telles que :

- $p \in U$ et $f(p) \in V$;
- la fonction $\psi \circ f \circ \phi^{-1} : \phi(U \cap f^{-1}(V)) \rightarrow \psi(f(U) \cap V)$ est différentiable en $\phi(p)$.

L'ensemble des fonctions différentiables est noté $\mathcal{F}(M, N)$, ou plus simplement, lorsque N est \mathbb{R} muni de la structure différentiable usuelle, $\mathcal{F}(M)$.

Il est facile de montrer que cette définition est correcte, c'est-à-dire, que la différentiabilité en un point ne dépend pas des cartes choisies en ce point. Dès lors, demander que ce soit vrai seulement pour deux cartes ou pour toutes les cartes, est équivalent.

Un exemple simple de fonction différentiable en un point est précisément, une application associée à une carte, dont le domaine contient ce point. En effet, si (U, ϕ) est une carte d'une variété différentiable, avec $p \in U$, alors ϕ est à valeurs dans \mathbb{R}^n ; on voit ce dernier ensemble comme variété différentiable, dont une carte est donnée par $(\mathbb{R}^n, \text{Id})$. Dès lors, ϕ est différentiable en p puisque, en choisissant précisément les cartes (U, ϕ) et $(\mathbb{R}^n, \text{Id})$, la fonction $\text{Id} \circ \phi \circ \phi^{-1} = \text{Id}|_{\phi(U)}$ est différentiable en $\phi(p)$.

Par ailleurs, si (U, ϕ) est une carte sur une variété différentiable et si $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ désigne la projection sur la i^{e} coordonnée, alors les fonctions $\phi_i := \pi_i \circ \phi$ s'avèrent elles aussi différentiables. Elles portent le nom d'*applications coordonnées* sur U .

Comme pour les fonctions $\mathbb{R}^n \rightarrow \mathbb{R}$ qui sont différentiables, on voudrait différentier en un point et dans une certaine direction. Prendre un point sur une variété est simple ; cependant, il n'est pas aussi aisé de définir une direction, du moins aussi facilement que dans \mathbb{R}^n (où prendre un vecteur non-nul suffit). Pour pallier à ce manque, on peut à nouveau se baser sur les cartes. Si p est un point de M , on peut choisir une carte (U, ϕ) telle que $p \in U$ et $\phi(p) = 0$. Dans \mathbb{R}^n cette fois, on veut exprimer un vecteur de telle sorte qu'on puisse le récupérer avec ϕ^{-1} ; cela est possible moyennant l'astuce suivante. Au lieu de penser au vecteur seul, on va le considérer comme vecteur tangent d'une courbe paramétrée par une fonction $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ telle qu'elle passe à l'origine en 0. Dès lors, notre vecteur tangent est donné par $\gamma'(0)$. On remarque sans surprise que, pour une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $v \in \mathbb{R}^n$ et $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n, t \rightarrow tv$, les égalités suivantes sont vérifiées :

$$\partial f(0) \cdot v = \lim_{t \rightarrow 0} \frac{f(tv) - f(0)}{t} = \lim_{t \rightarrow 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t} = (f \circ \gamma)'(0)$$

et que $\gamma'(0) = v$, ce qui correspond à la « chain rule ». Remarquons que, si l'on remplace γ par une autre courbe paramétrée par δ , mais telle que $\gamma(0) = 0 = \delta(0)$

et $\gamma'(0) = v = \delta'(0)$, le réel obtenu ci-dessus est aussi égal à $(f \circ \delta)'(0)$; En conséquence, nous allons, avant de remonter à la variété, identifier ces deux courbes. Notons également que nous n'avons pas besoin de définir les paramétrisations sur tout \mathbb{R} , mais seulement sur un intervalle ouvert contenant 0.

Définition 20. Deux courbes γ et δ définies sur des intervalles respectifs I et J contenant 0, seront dites *équivalentes en 0* si leurs images en 0 sont les mêmes et si on a $\gamma'(0) = \delta'(0)$.

Il est facile de voir que la relation entre les courbes, définie ci-dessus, est une relation d'équivalence. On peut alors voir un vecteur de \mathbb{R}^n comme une classe d'équivalence de courbes ; cette façon, certes compliquée, de voir les vecteurs, permet néanmoins de les « remonter » vers la variété M .

En effet, reprenons la carte (U, ϕ) pour laquelle $\phi(p) = 0$. Prenons un vecteur v de \mathbb{R}^n , choisissons γ une courbe qui la représente (c'est-à-dire, $\gamma'(0) = v$). Puisqu'elle vaut 0 en 0, il existe un intervalle ouvert I tel que $\gamma(I) \subset \phi(U)$. On considère alors la courbe $\phi^{-1} \circ \gamma$ définie sur I et à valeurs dans M ; par définition de ϕ , elle passe par p . Le vecteur tangent en p associé à v se définit alors par classe d'équivalence :

Définition 21. Un *vecteur tangent en un point p sur une variété différentiable M* est une classe d'équivalence de courbes $v : I \rightarrow M$, où I est un intervalle ouvert contenant 0, et où la relation d'équivalence est définie par :

v et μ sont équivalents s'il existe une carte (U, ϕ) , $p \in U$, telle que $(\phi \circ v)(0) = \phi(p) = (\phi \circ \mu)(0)$ et $(\phi \circ v)'(0) = (\phi \circ \mu)'(0)$. L'ensemble des vecteurs tangents en p sur M est noté TM_p .

On montre, comme pour les fonctions différentiables, que la définition est bien fondée, c-à-d. que la relation d'équivalence ne dépend pas de la carte choisie. Remarquons par ailleurs, pour la variété usuelle \mathbb{R}^n , le plan tangent en tout point n'est autre que \mathbb{R}^n lui-même.

On peut alors considérer les différentielles des fonctions différentiables dont le domaine et l'image sont des variétés :

Définition 22. Si $f : M \rightarrow N$ est une fonction différentiable, la *différentielle de f au point $p \in M$* est l'application :

$$f_*(p) : TM_p \rightarrow TN_{f(p)}, \quad [v] \mapsto [f \circ v]$$

Quelques remarques s'imposent. Premièrement, si les variétés M et N sont respectivement \mathbb{R}^n et \mathbb{R}^m , alors les classes s'identifient aux vecteurs ; nous sommes alors en train d'associer, à un vecteur $v \in \mathbb{R}^n$, le vecteur tangent en $f(p)$ défini par

la classe d'équivalence dont un représentant est la courbe $t \mapsto f(p + tv)$. La dérivée (en t) de cette courbe, évaluée en 0, n'est autre que $\partial f(p) \cdot v$, ce qui correspond bien à notre intuition de base.

Prenons un autre cas particulier. On peut, grâce à cette définition, considérer les différentielles des courbes $v : I \rightarrow M$. Au réel 1, seul vecteur tangent de \mathbb{R} intéressant (les autres s'obtenant par multiples), on associe le vecteur tangent dont une courbe représentative est donnée par... v elle-même. On redécouvre que la différentielle de v en tout point est la dérivée en ce point...

A nouveau, comme toute définition faisant appel à une relation d'équivalence, on montre que la différentielle d'une fonction en un point est bien définie. Si l'on prend deux courbes v et μ définissant le même vecteur tangent en un point $p \in M$, alors on a, pour une carte (U, ϕ) , $p \in U$, l'égalité $(\phi \circ v)(0) = (\phi \circ \mu)(0) = \phi(p)$, mais aussi $(\phi \circ v)'(0) = (\phi \circ \mu)'(0)$. Dès lors, si l'on fixe sur N , une carte (V, ψ) avec $f(p) \in V$, on a

$$\begin{aligned}(\psi \circ f \circ v)'(0) &= \partial(\psi \circ f \circ \phi^{-1})(\phi(\mu(0))) \cdot (\phi \circ \mu)'(0) \\(\psi \circ f \circ \mu)'(0) &= \partial(\psi \circ f \circ \phi^{-1})(\phi(\mu(0))) \cdot (\phi \circ \mu)'(0)\end{aligned}$$

et notre hypothèse fait que les deux membres de droite sont égaux ; en conséquence de quoi, les vecteurs tangents que définissent $f \circ v$ et $f \circ \mu$ sont les mêmes.

Cela termine l'exposé des notions que nous utilisons dans le texte principal.

Bibliographie

- [1] D. Barden and C. Thomas. *An introduction to differential manifolds*. Imperial College Press, 2003.
- [2] G. Brunnett. The curvature of plane elastic curves. Technical report, Naval Postgraduate School, March 1993.
- [3] J. Carr. *Applications of centre manifold theory*. Springer, 1981.
- [4] E. A. Coddington and Norman Levinson. *Theory of Ordinary Differential Equations*. Krieger, 1985.
- [5] M. Demazure. *Bifurcations and catastrophes – Geometry of Solutions to Nonlinear Problems*. Springer - Verlag, 2000.
- [6] P. Glendinning. *Stability, Instability and Chaos – an Introduction to the theory of nonlinear differential equations*. Cambridge University Press, 1994.
- [7] S.-T. Hu. *Differentiable manifolds*. Holt, Rinehart and Winston New York, 1969.
- [8] S. Lang. *Introduction to differentiable manifolds*. Springer, 2002.
- [9] R. Levien. The elastica : a mathematical history. Technical Report UCB/EECS-2008-103, EECS Department, University of California, Berkeley, Aug 2008.
- [10] J. Mahwin and N. Rouche. *Equations différentielles ordinaires - tome 1 : Théorie générale*. Masson et Cie, 1973.
- [11] L. Perko. *Differential equations and dynamical systems*. Springer, 2001.
- [12] L. Pocivavsek and al. Stress and fold localization in thin elastic membranes. *Science*, 320 :912–916, 2008.
- [13] W. A. Schwalm. Elliptic functions sn, cn, dn, as trigonometry, October 2002.
- [14] D. A. Singer. Lectures on elastic curves and rods. *Curvature and Variational Modeling in Physics and Biophysics. AIP Conference Proceedings.*, 1002 :3–32, 2008.
- [15] S. Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*. Springer-Verlag, 1990.