

Nils-Peter Skoruppa

Mathématiques pour les sciences
naturelles



Polycopié

Premier Cycle — Université Bordeaux 1

Version: Id: poly.tex,v 1.2 2003/11/20 23:10:59 fenrir Exp

Avertissement

Ce polycopié est une version un peu élaborée de mes notes au cours du même nom que j'ai assuré pendant l'hiver 1996.

Talence, le 24 décembre 1997

Table des Matières

1	Fonctions dérivables	1
1.1	Rappels	1
1.2	Les Accroissements Finies	3
1.3	Quelques conséquences des Accroissements Finis	4
1.4	Formule de Taylor	6
1.5	Etude locale d'une fonction	9
1.6	Développement limité	13
2	Courbes	17
2.1	Rappels sur $\mathbb{R}^n$	17
2.2	Courbes paramétrées	19
2.3	La courbure d'une courbe plane	21
3	Fonctions en deux variables	27
3.1	Vocabulaire de base	27
3.2	Dérivées de fonctions en 2 variables	29
3.3	Etudes locales	33
4	Equations différentielles du premier ordre	39
4.1	Existence et unicité des solutions	40
4.2	Variation de la constante	41
4.3	Séparation des variables	43
4.4	Equation différentielle de Bernoulli	45
4.5	Changement de variable	46
5	Déterminants, valeurs et vecteurs propres	49
5.1	La définition du déterminant	50
5.2	Règle de calcul pour les déterminants	51
5.3	Applications aux équations linéaires	55
5.4	Valeurs et vecteurs propres	56

5.5	Matrices triangonalisables	59
6	Systèmes linéaires d'équations	
	différentielles à coefficients constants	61
6.1	Fonctions à valeurs complexes	62
6.2	La théorie générale	63
6.3	Calcul d'un système fondamental	66
6.4	Equations différentielles linéaires de 2me ordre	
	à coefficients constants	70

Chapitre 1

Fonctions dérivables

1.1 Rappels

Un intervalle fermé est un ensemble de nombres réels de la forme

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

ou de la forme

$$[a, \infty) = \{x \in \mathbb{R} : a \leq x\}, \quad (-\infty, b] = \{x \in \mathbb{R} : x \leq b\}.$$

Un intervalle du premier type est aussi appelé compact ou fermé et borné.

Un intervalle de la forme

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}$$

est appelé ouvert. Ici on admet pour a, b les symboles $\pm\infty$. Ainsi par exemple $(-\infty, +\infty) = \mathbb{R}$ est un intervalle ouvert.

Le mot fonction indique dans ce premier chapitre toujours une application $f : D \rightarrow \mathbb{R}$ où le domaine D de f est un sous-ensemble des nombres réels. On dit qu'une fonction est définie sur ou dans un ensemble E si E est un sous-ensemble du domaine de f .

Définition. Une fonction f définie sur un intervalle I est dite *continue en* $a \in I$ si

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Remarque. On rappelle que pour une fonction f définie sur D l'écriture

$$b = \lim_{x \rightarrow a, x \in D} f(x)$$

signifie qu'il existe au moins une série (a_n) dans D telle que $a = \lim_n a_n$, et que pour toute telle série on a

$$b = \lim_n f(a_n).$$

On suppose que la notion d'une limite d'une série est bien connue. Ici a n'appartient pas nécessairement au domaine de f . Si il est évident du contexte quel D il faut prendre on supprime le D dans l'écriture, i.e. on écrit simplement

$$b = \lim_{x \rightarrow a} f(x).$$

Remarque. En conséquence, si une fonction est continue dans un intervalle I (i.e. continue en tout $a \in I$), alors on a

$$\lim_n f(a_n) = f(\lim_n a_n)$$

pour toute série convergente dans I et avec limite dans I , et vice versa. Brèf : Continuité de f indique que l'on peut échanger 'limite' et 'l'application f '.

Nous rappelons sans preuve

Théorème 1.1. (Théorème principale sur les fonction continues).
Soit f continue sur l'intervalle compact I , alors $f(I)$ est compact.

En conséquence, si f est continue sur l'intervalle compact I , alors il existe un minimum et un maximum absolu de f , i.e. il existe $x, y \in I$ tel que $f(x) \leq f(\xi) \leq f(y)$ pour tout $\xi \in I$. En plus, pour tout c entre $f(x)$ et $f(y)$ il existe un ξ tel que $c = f(\xi)$.

Définition. Une fonction f définie sur un intervalle ouvert I est appelée dérivable en $a \in I$ si la limite

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

existe et est finie. La limite est dite la dérivée de f en a , noté $f'(a)$.

Remarque. Si f est dérivable en a , alors la fonction

$$\Delta_a f : I \rightarrow \mathbb{R}, \quad (\Delta_a f)(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} & \text{si } x \neq a \\ f'(a) & \text{si } x = a \end{cases}$$

est continue en a .

Notation. Si f est dérivable dans I (i.e. f est dérivable pour tout $a \in I$) on note $f' : I \rightarrow \mathbb{R}$ la fonction $x \mapsto f'(x)$. On dit que f est n -fois dérivable sur I si f est dérivable, f' est dérivable, $f(2) := (f')'$ est dérivable, $\dots$, si $f^{(n-1)}$ est dérivable, et donc $f^{(n)} := (f^{(n-1)})'$ existe. On pose aussi $f^{(0)} = f$.

Théorème 1.2. *Si $f : I \rightarrow \mathbb{R}$ est dérivable en a , alors f est continue en a .*

Démonstration. On a

$$f(x) = f(a) + (x - a) \cdot (\Delta_a f)(x)$$

et donc

$$\lim f(x) = f(a) + (\lim(x - a)) \cdot \lim(\Delta_a f)(x) = f(a) + 0 \cdot f'(a) = f(a).$$

□

On suppose que les règles usuelles pour calculer la dérivée d'une fonction sont connues.

1.2 Les Accroissements Finies

Théorème 1.3. *Soit $f : (a, b) \rightarrow \mathbb{R}$ dérivable en $c \in (a, b)$. Si f possède un maximum locale ou minimum locale en c , alors $f'(c) = 0$.*

Remarque. On parle d'un maximum locale en c si il existe un $\varepsilon > 0$ tel que $f(c) \geq f(x)$ pour tout $c - \varepsilon < x < c + \varepsilon$. Un minimum locale sera défini analoguement.

Démonstration. Supposons que f possède un minimum locale en c . On a $(\Delta_c f)(c - \frac{1}{n}) \leq 0$ pour $n \in \mathbb{N}$ assez grand, d'où $f'(c) = \lim(\Delta_c f)(c - \frac{1}{n}) \leq 0$. De même, on a $(\Delta_c f)(c + \frac{1}{n}) \geq 0$ pour $n \in \mathbb{N}$ assez grand, d'où $f'(c) = \lim(\Delta_c f)(c + \frac{1}{n}) \geq 0$. La seule possibilité est donc que $f'(c) = 0$. Le cas que f atteint un maximum locale en c est analogue. □

Théorème 1.4. (Théorème de Rolle). *Soit $f : [a, b] \rightarrow \mathbb{R}$ continu sur $[a, b]$ et dérivable dans (a, b) . Soit $f(a) = f(b)$. Alors il existe un $\xi \in (a, b)$ tel que*

$$f'(\xi) = 0.$$

Démonstration. Soit x et y tel que $f([a, b]) = [f(x), f(y)]$. De tels x et y existent d'après le théorème principal des fonctions continues. Donc f prend son max et min en x et y respectivement.

Si $x \in (a, b)$, alors $f'(x) = 0$. Si $y \in (a, b)$, alors $f'(y) = 0$. Si $\{x, y\} \subseteq \{a, b\}$, alors $f(x) = f(y)$, et donc f est constant et puis $f' \equiv 0$. □

Théorème 1.5. (Théorème des accroissements finis) Soit $f: [a, b] \rightarrow \mathbb{R}$ continu sur $[a, b]$ et dérivable dans (a, b) . Alors il existe un $\xi \in (a, b)$ tel que

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

Démonstration. On définit $F: [a, b] \rightarrow \mathbb{R}$ par

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Alors $F(a) = F(b) = f(a)$. D'après Rôle on a $F'(\xi) = 0$ pour un $\xi \in (a, b)$. Mais $F'(\xi) = 0$ est équivalent à la formule donnée. $\square$

Théorème 1.6. (Accroissements à la Cauchy) Soient $f, g: [a, b] \rightarrow \mathbb{R}$ continues sur $[a, b]$ et dérivables dans (a, b) . Alors il existe un $\xi \in (a, b)$ tel que

$$(f(b) - f(a)) g'(\xi) = (g(b) - g(a)) f'(\xi).$$

Remarque. 1. Pour $g(x) = x$ on retrouve les accroissements simples. 2. Si on suppose que $g'(x) \neq 0$ pour tout $x \in (a, b)$, alors on a $g(b) - g(a) \neq 0$ (d'après les accroissements simples), et on peut écrire les accroissements à la Cauchy sous la forme

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Démonstration. On pose pour $x \in [a, b]$:

$$F(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)].$$

On a $F(a) = f(a)g(b) - g(a)f(b) = F(b)$. Appliquer Rolle à F . $\square$

1.3 Quelques conséquences des Accroissements Finis

Corollaire 1.6.1. Soit $f: [a, b] \rightarrow \mathbb{R}$ continu et dérivable dans (a, b) tel que $f'(x) = 0$ pour tout $x \in (a, b)$. Alors f est constante.

Démonstration. Fixons un $c \in (a, b)$. Alors pour tout $x \in [a, b]$, $x \neq c$, il existe un ξ entre c et x tel que $\frac{f(c) - f(x)}{c - x} = f'(\xi)$. Mais $f'(\xi) = 0$, donc $f(x) = f(c)$. $\square$

1.3. QUELQUES CONSÉQUENCES DES ACCROISSEMENTS FINIS 5

Corollaire 1.6.2. Soient $f, g: [a, b] \rightarrow \mathbb{R}$ continus et dérivables dans (a, b) tels que $f'(x) = g'(x)$ pour tout $x \in I$. Alors il existe une constante C tel que $f(x) = g(x) + C$ pour tout $x \in [a, b]$.

Démonstration. Appliquer le corollaire précédent à $f - g$. □

Corollaire 1.6.3. Soit $f: (a, b) \rightarrow \mathbb{R}$ dérivable et $f'(x) \geq 0$ pour tout $x \in (a, b)$. Alors f est croissante.

Démonstration. Soit $x < y$. Il existe un $x < \xi < y$ tel que

$$\frac{f(y) - f(x)}{y - x} = f'(\xi).$$

Car $f'(\xi) \geq 0$ on a donc $f(y) - f(x) \geq 0$. □

Remarque. De même on peut montrer que f est strictement croissante ou décroissante ou strictement décroissante si $f'(x)$ est > 0 , $\leq$ ou < 0 respectivement.

Exemple. Montrer : (i) Si f est croissante sur (a, b) , alors $f'(x) \geq 0$ pour $x \in (a, b)$. (ii) La proposition suivante est fausse : Si f est strictement croissante sur (a, b) , alors $f'(x) > 0$ pour $x \in (a, b)$. (Un contre-exemple est $f(x) = x^3$: la fonction f est strictement croissante sur tout $\mathbb{R}$, mais $f'(0) = 0$.)

Théorème 1.7. (Règle de l'Hôpital) Soient f, g dérivables dans un intervalle ouvert I et a une borne de I . On suppose que $\lim_{x \rightarrow a} f(x) = 0$ et $\lim_{x \rightarrow a} g(x) = 0$, et que $\lim_{x \rightarrow a} f'(x)/g'(x)$ existe. Alors $\lim_{x \rightarrow a} f(x)/g(x)$ existe, et on a

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.$$

Démonstration. Cas 1 : $a \in \mathbb{R}$. En prolongeant f et g , i.e. en définissant $f(a) = g(a) := 0$, on peut supposer que a appartient au domaine de définition de f et g et que f et g sont continues. Par les accroissement à la Cauchy on obtient ainsi

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi)}{g'(\xi)}$$

avec un $\xi = \xi(x)$ entre x et a . Pour $x \rightarrow a$ on a $\xi = \xi(x) \rightarrow a$ (car ξ est entre x et a), et d'où le théorème.

Cas 2 : $a = \pm\infty$. Poser $F(x) = f(1/x)$ et $G(x) = g(1/x)$. D'après cas 1 on a

$$\begin{aligned}\lim_{x \rightarrow a} \frac{f(x)}{g(x)} &= \lim_{x \rightarrow a} \frac{F(1/x)}{G(1/x)} = \lim_{y \rightarrow 0} \frac{F(y)}{G(y)} \\ &= \lim_{y \rightarrow 0} \frac{F'(y)}{G'(y)} = \lim_{y \rightarrow 0} \frac{f'(y) \cdot \left(\frac{-1}{y^2}\right)}{g'(y) \cdot \left(\frac{-1}{y^2}\right)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.\end{aligned}$$

□

Exemple. $\lim_{x \rightarrow 0} \sin x/x = \lim_{x \rightarrow 0} \cos x = 1$.

1.4 Formule de Taylor

Comme autre conséquence fondamentale des accroissements à la Cauchy on obtient :

Théorème 1.8. (Taylor) Soit $f: I \rightarrow \mathbb{R}$ une fonction $n + 1$ -fois dérivable sur l'intervalle ouvert I , et soit $a \in I$. On pose

$$p_{n,a}(x) = f(a) + f'(a)(x - a) + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n$$

(n -me polynôme de Taylor de f en a). Alors pour tout $x \in I$ il existe un ξ entre a et x tel que

$$\frac{f(x) - p_{n,a}(x)}{(x - a)^{n+1}} = \frac{f^{(n+1)}(\xi)}{(n + 1)!}.$$

Démonstration. On fixe x , et on considère les fonctions de la variable t qui sont définies par

$$\begin{aligned}F(t) &= [f(t) + f'(t)(x - t) + \cdots + \frac{f^{(n)}(t)}{n!}(x - t)^n] \\ G(t) &= (x - t)^{n+1}.\end{aligned}$$

D'après les accroissements à la Cauchy on a

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(\xi)}{G'(\xi)}$$

avec un ξ entre x et a . Or, on a

$$F(x) = f(x), \quad F(a) = p_{n,a}, \quad G(x) = 0, \quad G(a) = (x - a)^{n+1}.$$

De plus,

$$G'(\xi) = -(n+1)(x-\xi)^n,$$

et

$$\begin{aligned} F'(\xi) &= f'(\xi) + f''(\xi)(x-\xi) + \cdots + \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n \\ &\quad - [f'(\xi) + f''(\xi)(x-\xi) + \cdots + \frac{f^{(n)}(\xi)}{(n-1)!}(x-\xi)^{n-1}] \\ &= \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n. \end{aligned}$$

D'où le théorème. □

Remarque. Pour $n = 0$ on a $p_{n,a}(x) = f(a)$, et donc on reobtient le théorème des accroissements finis.

Remarque. On peut écrire Taylor sous la forme

$$f(x) = p_{n,a}(x) + R(x), \quad R(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Donc

$$f(x) \approx p_{n,a}(x)$$

si "l'erreur" $R(x)$ est petit. Par exemple, si $f^{(n+1)}$ est continue en a , et si $x \rightarrow a$, alors ξ (comme fonction de x) tend vers a , et donc

$$\lim_{x \rightarrow a} f^{(n+1)}(\xi) = f^{(n+1)}(a).$$

Donc $\lim_{x \rightarrow a} R(x)/(x-a)^n = 0$, et on peut espérer que $R(x)$ tend vers 0 le plus rapide le plus n est élevé. Donc pour n assez grand le polynôme $p_{n,a}(x)$ sera très souvent une bonne approximation de $f(x)$ pour $x \approx a$.

Il existe beaucoup d'autres formules pour l'erreur $R(x)$ (voir les livres), même dans le cas où f est seulement n -fois dérivable.

Exercice. Montrer que $p_{n,a}^{(k)}(a) = f^{(k)}(a)$ pour $0 \leq k \leq n$.

Exemple. On considère $f(x) = \exp(x)$ et $a = 0$. (Rappel : $\exp(x)$ est l'unique fonction dérivable sur $\mathbb{R}$ telle que $\exp' = \exp$ et $\exp(0) = 1$. Nous utilisons dans ce qui suit que $\exp$ est croissante et $e := \exp(1) < 3$.) On a $\exp^{(n)}(0) = 1$ pour tout n , et donc

$$p_{n,0} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}.$$

D'après Taylor on a $\exp(x) = p_{n,0}(x) + R(x)$ avec

$$|R(x)| = \left| \frac{\exp(\xi)}{(n+1)!} x^{n+1} \right| \leq \frac{\exp(|x|)}{(n+1)!} |x|^{n+1}$$

(avec un $\xi = \xi(x)$).

Par exemple, on trouve ainsi

$$e = \exp(1) = p_{n,0}(1) + E = 1 + \frac{1}{2} + \frac{1}{6} + \cdots + \frac{1}{10!} + E = 2.7182818... + E,$$

avec

$$E \leq \exp(1)/11! \leq \frac{3}{11!} \leq 10^{-7}.$$

Si on fixe maintenant x et si on utilise $\lim_n \frac{|x|^{n+1}}{(n+1)!} = 0$, on obtient la formule

$$\exp(x) = \lim_n \sum_{k=0}^n \frac{x^k}{k!}.$$

Comme dans l'exemple précédent on peut montrer en général

Théorème 1.9. *Soit f une fonction infiniment dérivable dans un intervalle ouvert I . On fixe $a \in I$. On suppose qu'il existe un constant M tel que*

$$|f^{(n)}(\xi)| \leq M$$

pour tout n et tout $\xi \in I$. Alors, pour tout $x \in I$ on a

$$f(x) = \lim_n \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k.$$

Remarque. Les hypothèses du théorème sont satisfaites par exemple pour $\cos$ et $\sin$, pour $\log x$ avec $I = (\varepsilon, +\infty)$ (pour tout $\varepsilon > 0$ fixé) et pour beaucoup d'autres des fonctions usuelles.

Attention. La formule pour f du dernier théorème n'est pas toujours valable, i.e. sans hypothèses supplémentaires sur f . Contre-exemple

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

est infiniment dérivable sur $\mathbb{R}$. Mais $p_{n,0} \equiv 0$ pour tout n (exercice).

Exercice. Montrer :

$$\begin{aligned} \cos(x) &= \lim_n \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} \mp \cdots + \frac{(ix)^{2n}}{(2n)!} \right), \\ \sin(x) &= \lim_n \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \cdots + i \frac{(ix)^{2n+1}}{(2n+1)!} \right) \end{aligned}$$

pour tout nombre réel x . Ici $i = \sqrt{-1}$, i.e. $i^2 = -1$.

1.5 Etude locale d'une fonction

On peut résumer le théorème de Taylor qualitativement en disant que toute fonction suffisamment dérivable se comporte localement comme un polynôme. Pour préciser cette phrase il faut une précision mathématique de la notion d'un comportement qualitatif d'une fonction proche d'un point.

Notation. (Les O de Landau) Soient f et g des fonctions définies sur un intervalle I , et soit $a \in I$ soit une borne de I (par exemple $a = +\infty$ et $I = \mathbb{R}$). Alors on écrit

$$f(x) = o(g(x)) \text{ pour } x \rightarrow a$$

si il existe un voisinage E de a et une fonction ε définie sur $E \cap I$ et avec $\lim_{x \rightarrow a} \varepsilon(x) = 0$ telle que

$$|f(x)| \leq \varepsilon(x)|g(x)|$$

pour tout $x \in E \cap I$. On écrit

$$f(x) = O(g(x)) \text{ pour } x \rightarrow a$$

si il existe un voisinage E de a et constante A (qui ne dépend de x) telle que

$$\left| \frac{f(x)}{g(x)} \right| \leq A.$$

pour tout $x \in E \cap I$

Remarque. Un voisinage de a est un intervalle ouvert contenant a si a est un nombre, et est un intervalle non vide de la forme $(s, +\infty)$ resp. $(-\infty, s)$ si $a = +\infty$ resp. $a = -\infty$.

Si $g(x) \neq 0$ pour tout x dans un voisinage de x , alors $f(x) = o(g(x))$ est équivalente à dire que

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0.$$

Exemple. $x + 1 = o(x^2)$, $\sin(x) = O(1)$, $\log x = o(x)$ ($x \rightarrow \infty$), $x + 1 = O(1)$ ($x \rightarrow 0$).

Remarque. On remarque que $o(g)$ entraîne et est plus fort que $O(g)$.

Utilisant cette notation on peut résumer Taylor qualitativement en disant (voir la remarque après le théorème de Taylor) :

Théorème 1.10. Soit f $(n+1)$ -fois dérivable dans un intervalle ouvert I et $f^{(n+1)}$ continue en $a \in I$. Alors

$$f(x) = p_{n,a}(x) + O((x-a)^{n+1})$$

pour $x \rightarrow a$.

Remarque. On observe que le théorème implique

$$f(x) = p_{n,a}(x) + o((x-a)^n) \quad (x \rightarrow a).$$

Une autre écriture de cette formule est

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(n)}(a)}{n!}h^n + o(h^n)$$

pour $h \rightarrow 0$.

Nous allons montrer que les deux dernières formules sont vraies même si f est seulement n -fois dérivable (l'hypothèse qui est en tout cas nécessaire pour parler du polynôme de Taylor).

Théorème 1.11. Soient f, g n -fois dérivable dans l'intervalle ouvert I , et soit $a \in I$. Alors les propriétés suivantes sont équivalentes :

1. $f(x) = g(x) + o((x-a)^n)$ pour $x \rightarrow a$.
2. $f^{(k)}(a) = g^{(k)}(a)$ pour $0 \leq k \leq n$.

Remarque. On dit que deux fonctions $f, g : I \rightarrow \mathbb{R}$ sont égales de l'ordre n en $a \in I$ si la propriété 1 est satisfaite, i.e. si

$$\lim_{x \rightarrow a} \frac{f(x) - g(x)}{(x-a)^n} = 0.$$

Démonstration. On remplaçant $f(x)$ par $f(x) - g(x)$ on peut supposer que $g(x) = 0$.

On suppose 2. En appliquant n fois la règle de l'Hôpital on obtient

$$\begin{aligned} \lim_{x \rightarrow a} \frac{f(x)}{(x-a)^n} &= \lim_{x \rightarrow a} \frac{f'(x)}{n(x-a)^{n-1}} = \cdots \\ &= \lim_{x \rightarrow a} \frac{f^{(n-1)}(x)}{n!(x-a)} = \lim_{x \rightarrow a} \frac{f^{(n)}(x)}{n!} = 0, \end{aligned}$$

i.e. propriété 1.

On suppose maintenant 1. Soit $p_{n,a}$ le n -me polynôme de Taylor de f en a . On vérifie que $p_{n,a}^{(k)}(a) = f^{(k)}(a)$ pour $0 \leq k \leq n$. On a déjà démontré dans la première partie de la preuve que donc $f(x) = p_{n,a}(x) + o((x-a)^n)$. Avec l'hypothèse $f(x) = o((x-a)^n)$ on en conclut

$$p_{n,a}(x) = o((x-a)^n) \quad (x \rightarrow a),$$

i.e.

$$f(a) + f'(a)(x-a) + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n = o((x-a)^n).$$

On en déduit $f(a) = 0$, et en divisant par $x-a$ ensuite

$$f'(a) + \frac{f''(a)}{2!} + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^{n-1} = o((x-a)^{n-1}).$$

Encore en déduit $f'(a) = 0$, et donc en divisant par $x-a$ que

$$\frac{f''(a)}{2!} + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^{n-2} = o((x-a)^{n-2}),$$

et puis $f''(a) = 0$. Par récurrence sur n on trouve ainsi 2. □

On a utilisé dans la démonstration que $p_{n,a}^{(k)}(a) = f^{(k)}(a)$ pour $0 \leq k \leq n$. Il est facile à montrer que $p_{n,a}(x)$ est l'unique polynôme avec cette propriété : En fait, on peut écrire tout polynôme $q(x)$ de l'ordre n sous la forme

$$q(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \dots + a_n(x-a)^n$$

avec des nombres réels a_k convenables (exercice), et on vérifie immédiatement que les a_k vérifient $a_k = q^{(k)}(a)/k!$. Donc comme corollaire du théorème précédent on obtient

Théorème 1.12. *Soit f une fonction n -fois dérivable dans l'intervalle ouvert I , et soit $a \in I$. Alors $f(x)$ est égale à $p_{n,a}$ de l'ordre n (i.e. $f(x) = p_{n,a}(x) + o(x^n)$), et $p_{n,a}$ est le seul polynôme d'ordre n avec cette propriété.*

Comme application du dernier théorème on montre :

Théorème 1.13. *Soit f une fonction infiniment dérivable dans un intervalle ouvert I . On suppose que $a \in I$ est un point stationnaire (i.e. que $f'(a) = 0$), et qu'il existe un $n \geq 2$ tel que $f^{(n)}(a) \neq 0$. Soit n le plus petit de tels nombres naturels. Alors on a trois possibilités :*

1. Si n est pair est $f^{(n)}(a) > 0$, alors f possède un minimum local chez a .

2. Si n est pair est $f^{(n)}(a) < 0$, alors f possède un maximum local chez a .
3. Si n est impair, alors f ne possède ni un maximum ni un minimum locale chez a .

Démonstration. On déduit avec les hypothèses que

$$f(a+h) = f(a) + \frac{f^{(n)}(a)}{n!}h^n + o(h^n),$$

i.e.

$$\frac{f(a+h) - f(a)}{h^n} = \frac{f^{(n)}(a)}{n!} + o(1)$$

pour $h \rightarrow 0$. Car $f^{(n)}(a) \neq 0$ il existe un $\varepsilon > 0$ tel que

$$\text{signe de } \frac{f(a+h) - f(a)}{h^n} = \text{signe de } \frac{f^{(n)}(a)}{n!}$$

pour tout $|h| < \varepsilon$. Si n est pair on en déduit

$$f(a+h) - f(a) = \text{signe de } \frac{f^{(n)}(a)}{n!}$$

pour tout $|h| < \varepsilon$, si n est impair alors

$$f(a+h) - f(a) = \begin{cases} -\text{signe de } \frac{f^{(n)}(a)}{n!} & \text{si } h < 0 \\ +\text{signe de } \frac{f^{(n)}(a)}{n!} & \text{si } h > 0 \end{cases}$$

pour tout $|h| < \varepsilon$. D'où le théorème. $\square$

Exemple. Soit $f(x) = x^8$. Alors $x = 0$ est le seul point stationnaire. On a $f^{(k)}(0) = 0$ pour $1 \leq k < 8$ et $f^{(8)}(0) = 8! > 0$. Alors f a un minimum locale chez 0.

Par contre, $f(x) = x^7$ satisfait a $f^{(k)}(0) = 0$ pour $1 \leq k < 7$ et $f^{(7)}(0) = 7! \neq 0$, et on fait $f(x)$ change le signe quand x passe 0.

Attention. Il se peut que $f^{(n)}(a) = 0$ pour tout $n > 0$. Dans ce cas on ne peut rien dire : Exemple :

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

a un minimum locale en 0 et $f^{(n)}(0) = 0$ pour tout n . Par contre,

$$g(x) = \begin{cases} \exp(-\frac{1}{x^2}) & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -\exp(-\frac{1}{x^2}) & \text{si } x < 0 \end{cases}.$$

change signe quand x passe 0 et $g^{(n)}(0) = 0$ pour tout n .

1.6 Développement limité

Si on veut étudier le comportement qualitatif d'une fonction chez un point il n'est pas toujours pratique de calculer les dérivées successives pour en obtenir le polynôme de Taylor. Par exemple, si f est définie via des opérations élémentaires sur les fonctions (somme, produit, quotient et composition) à l'aide de fonctions dont on connaît les polynômes de Taylor, on peut passer immédiatement aux polynômes de Taylor de f sans calculer ses dérivées. Nous expliquons ce procédé dans un cadre un peu plus général.

Soit $\phi_0, \phi_1, \phi_2, \dots$ une suite de fonction définies dans un intervalle I telle que

$$\phi_{n+1} = o(\phi_n) \quad (n = 0, 1, \dots)$$

pour $x \rightarrow a$. Ici $a \in I$ ou a est une borne de I (en particulier, on admet explicitement $a = \pm\infty$).

Notation. Soit f une fonction définie sur I , et soit $a_0, a_1, a_2, \dots$ une suite réelle. On écrit

$$f(x) \sim a_0\phi_0(x) + a_1\phi_1(x) + a_2\phi_2(x) + \dots \text{ pour } x \rightarrow a \quad (*)$$

si pour tout n on a

$$f(x) = a_0\phi_0(x) + a_1\phi_1(x) + a_2\phi_2(x) + a_n\phi_n(x) + o(\phi_n)$$

pour $x \rightarrow a$. On appelle le côté droit de $(*)$ un développement limité (DL) de f (en a et par rapport aux ϕ_n). La somme partielle

$$a_0\phi_0(x) + a_1\phi_1(x) + a_2\phi_2(x) + a_n\phi_n(x)$$

est appelée le DL d'ordre $o(\phi_n)$ de f en a (en a et p.r. aux ϕ_n).

Attention. Ce que est expliqué ici est mieux appelé en anglais “asymptotic expansion”. Cette notion comprend tous les français “développements limités” que j’ai trouvés (en fait, j’en ai trouvé plusieurs : “le polynôme de Taylor, des petites généralisations, aussi certains développements à l’infinie ...”).

Exemple. Le théorème de Taylor donne un tel développement limité : ici $\phi_n(x) = (x - a)^n$ (en fait $(x - a)^n = o((x - a)^{n+1})$ pour $x \rightarrow a$), et on a vu que si f est infiniment dérivable dans I et $a \in I$, alors

$$f(x) \sim f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \dots \quad (x \rightarrow a).$$

Mais attention : cette identité n'entraîne pas que $f(x)$ est égal à

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k.$$

(voir ci-dessus le contre-exemple avec la fonction $\exp(-1/x^2)$).

Exercice. Montrer

$$1. \log(1-x) \sim -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots (\sim \sum_{n=1}^{\infty} \frac{x^n}{n}),$$

$$2. (1+x)^\alpha \sim 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \dots (\sim \sum_{n=0}^{\infty} \binom{\alpha}{n}x^n)$$

pour $x \rightarrow 0$. Rappel : $y^\alpha = \exp(\alpha \log y)$ pour $y > 0$ et

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}.$$

En particulier, on a

$$\sqrt{1+x} \sim 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \frac{7}{256}x^5 - \frac{21}{1024}x^6 + \dots$$

quand $x \rightarrow 0$. (En fait, les DL de cet exercice sont des identités pour $|x| < 1$ si on interprète les côtés droits comme sommes infinies — une notion à préciser).

Exercice. On considère $\phi_0(x) = 1/x$, $\phi_1(x) = 1/x^2$, $\phi_2(x) = 1/x^3 \dots$ et $a = \infty$. Comme exemple d'un DL par rapport à cette suite on peut prendre

$$\log(1 + \frac{1}{x}) \sim \frac{1}{x} - \frac{1}{2x^2} + \frac{1}{3x^3} - \frac{1}{4x^4} \pm \dots \quad (x \rightarrow \infty).$$

(Indication : remplacer x par $-1/x$ dans le DL de $\log(1-x)$ pour $x \rightarrow 0$).

Théorème 1.14. *Les coefficients a_n dans un DL d'une fonction f par rapport à une suite ϕ_n sont uniquement déterminés par f .*

Démonstration. On montre cette unicité par récurrence sur N en utilisant l'identité

$$a_N = \lim_{n \rightarrow \infty} \frac{f(x) - a_0\phi_0(x) - \dots - a_{N-1}\phi_{N-1}(x)}{\phi_N},$$

qui est une conséquence immédiate de la définition d'un développement limité. $\square$

Ce qui concerne la déduction d'un DL nous montrons par des exemples qu'il est souvent possible de calculer un DL d'une fonction qui est définie par des opérations élémentaires à l'aide de fonctions plus simples.

On suppose que R, S sont des nombres naturels et que

$$\begin{aligned} f(x) &\sim x^{-R}(a_0 + a_1x + a_2 \dots) \\ g(x) &\sim x^{-S}(b_0 + b_1x + b_2 \dots) \end{aligned}$$

pour $x \rightarrow 0$. Il s'agit ici d'un DL par rapport à la suite $\phi_0(x) = x^{-S}$, $\phi_1(x) = x^{1-S}$, $\phi_2(x) = x^{2-S}$, $\dots$. On pose $P(x) = a_0 + a_1x + \dots + a_nx^n$ et $Q(x) = b_0 + b_1x + \dots + b_nx^n$. Alors, on obtient un DL pour $x \rightarrow 0$ de l'ordre $o(x^n)$ de la fonction

1. $f(x) + g(x)$ en faisant la somme des DL d'ordre $o(x^n)$ de f et g ,
2. $x^{R+S}f(x)g(x)$ si $b_0 \neq 0$ en effectuant le produit $P \cdot Q$ et ne conservant que les termes de degré $\leq n$.
3. $\frac{x^R f(x)}{x^S g(x)}$ en effectuant la division de P par Q selon les puissances croissantes jusqu'à l'ordre n , disons $P(x) = A(x)Q(x) + x^{n+1} \dots$; alors $x^R f(x)/(x^S g(x)) = A(x) + o(x^{n+1})$.
4. $g(x)^R f(g(x))$ si $g(x) = o(x)$ en remplaçant le x dans $P(x)$ par $Q(x)$ et en supprimant tous les termes d'ordre $\leq n$.

Il est clair, que le DL de $x^{R+S}f(x)g(x)$ de l'ordre $o(x^n)$ entraîne le DL de $f(x)g(x)$ de l'ordre $o(x^{n-R-S})$ en divisant par x^{R+S} , et analogue pour le quotient et la composition.

Pour montrer par exemple 2 on procède comme ce qui suit :

$$\begin{aligned}
 (P(x) + o(x^n))(Q(x) + o(x^n)) &= P(x)Q(x) + P(x)o(x^n) + Q(x)o(x^n) + o(x^{2n}) \\
 &= P(x)Q(x) + o(x^n) + o(x^n) + o(x^n) \\
 &= \text{les termes de degré } \leq n \text{ de } a(x)b(x) + o(x^n).
 \end{aligned}$$

Ici on a utilisé les faits évidents que $x^m o(x^n) = o(x^{m+n})$, qu'un $o(x^m)$ est en particulier un $o(x^n)$ si $m \geq n$, et que la somme (finie) de plusieurs $o(x^n)$ est un $o(x^n)$.

La démonstration des autres règles est pareilles. On peut également généraliser de telles règles facilement à des DL p.r. aux suites $\phi_n(x) = (x-a)^{n-S}$ (avec $x \rightarrow a$), ou $\phi_n(x) = x^{S-n}$ (avec $x \rightarrow \infty$).

Exercice. Montrer, en utilisant $\cos(x) \sim 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots$ et $\log(1+x) \sim x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \dots$ que l'on a

1. $\frac{\cos(x) \log(1+x)}{x} = 1 - \frac{1}{2}x - \frac{1}{6}x^2 + o(x^3)$
2. $\frac{\cos(x)}{x \log(1+x)} = \frac{1}{x^2} + \frac{1}{2} \frac{1}{x} - \frac{7}{12} - \frac{5}{24}x + \frac{41}{720}x^2 + \frac{3}{160}x^3 + o(x^3),$
3. $\frac{\cos(\log(1+x))}{\log(1+x)} = \frac{1}{x} + \frac{1}{2} - \frac{7}{12}x + \frac{7}{24}x^2 - \frac{109}{720}x^3 + o(x^3)$

pour $x \rightarrow 0$.

Chapitre 2

Courbes

2.1 Rappels sur $\mathbb{R}^n$

Nous utilisons $\mathbb{R}^n$ pour l'ensemble des vecteurs à lignes

$$x = (x_1, x_2, \dots, x_n) \quad (x_i \in \mathbb{R}).$$

Pour deux vecteurs x et y nous posons

$$x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n, \quad |x| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

On appelle $|x|$ la longueur de x . Pour $n = 2$ on a en particulier

$$|x| = \sqrt{x_1^2 + x_2^2},$$

qui est la longueur euclidienne de x d'après le théorème de Pythagore. En particulier on en déduit que

$$S^1 := \{x \in \mathbb{R}^2 : |x| = 1\}$$

est le cercle de rayon 1 et de centre $0 = (0, 0, \dots, 0)$. Car pour tout nombre réel t on a $\cos^2 t + \sin^2 t = 1$ on observe que

$$\phi(t) := (\cos t, \sin t)$$

est un élément du cercle S^1 . Donc on a une application

$$\phi : [0, 2\pi) \rightarrow S^1.$$

Théorème 2.1. *L'application ϕ est bijective, i.e. pour tout $x \in S^1$ il existe un et un seul $t \in [0, 2\pi)$ tel que $x = \phi(t)$.*

Démonstration. Soit $x = (a, b) \in S^1$. On sait que $\cos([0, \pi]) = [-1, +1]$. Donc il existe un et un seul $t \in [0, \pi]$ tel que $a = \cos t$. On a

$$b^2 = 1 - a^2 = 1 - \cos^2 t = \sin^2 t,$$

i.e. $b = \pm \sin t$. Cas 1 : $b \geq 0$. Alors $b = \sin t$ car $\sin t \geq 0$ pour $t \in [0, \pi]$. Donc $x = \phi(t)$ et ce t est unique car $\sin u < 0$ pour $\pi < u < 2\pi$. Cas 2 : $b < 0$. Alors on pose $t' = 2\pi - t$, et on trouve $a = \cos t = \cos t'$ et $b = -\sin t = \sin t'$. Donc ici $x = \phi(t')$ et ce t' est encore unique car $\sin u \geq 0$ pour $0 \leq u \leq \pi$. $\square$

Corollaire 2.1.1. (coordonnées polaires) *Tout x de $\mathbb{R}^2$ peut être écrit sous la forme*

$$x = r \cdot (\cos t, \sin t)$$

avec des nombres réels $r \geq 0$ et $0 \leq t < 2\pi$. Si $x \neq 0$, alors r et t sont uniques.

Remarque. On appelle t l'angle orienté entre le vecteur $(1, 0)$ et x .

Démonstration. Si $x = r(\cos t, \sin t)$, alors $|x| = r$, et — si $x \neq 0$ —, alors $x/|x| \in S^1$, et donc t est l'unique t tel que $\phi(t) = x/|x|$. Réciproquement, si on définit r et t par les formules précédentes on obtient l'écriture de x comme dans le corollaire. $\square$

En utilisant les coordonnées polaires on obtient l'interprétation suivantes du produit scalaire $x \cdot y$ pour $x, y \in \mathbb{R}^2$. Soit $x = r(\cos t, \sin t)$ et $y = s(\cos u, \sin u)$, alors

$$x \cdot y = rs(\cos t \cos u + \sin t \sin u) = |x| \cdot |y| \cdot \cos(t - u).$$

Donc $x \cdot y / (|x| \cdot |y|)$ est le cosinus de l'angle (non-orienté) entre x et y . En particulier, $x \cdot y = 0$ si et seulement si x est perpendiculaire à y .

Finalement, pour un $x = (a, b) \in \mathbb{R}^2$ on pose

$$\times x = (-b, a).$$

Evidemment $x \cdot (\times x) = 0$. En plus, pour obtenir $\times x$ dans le dessin usuel de $\mathbb{R}^2$ il faut tourner x à gauche (dans le sens anti-montre) par l'angle 90° .

2.2 Courbes paramétrées

Définition. Une courbe (paramétrée) est une application $c : I \rightarrow \mathbb{R}^n$ d'un intervalle I dans $\mathbb{R}^n$.

Remarque. On peut écrire $c(t) = (c_1(t), \dots, c_n(t))$ avec des fonctions usuelles $c_i : I \rightarrow \mathbb{R}$. Dans ce qui suit on suppose toujours que I est un intervalle ouvert, et que les c_i sont infiniment dérivables; bref : c est infiniment dérivable. On pose

$$c^{(n)} = (c_1^{(n)}, \dots, c_n^{(n)}).$$

Par la même façon, i.e. composé par composé, on peut définir l'intégrale d'une courbe en posant

$$\int_{t_0}^{t_1} c(t) dt = \left(\int_{t_0}^{t_1} c_1, \dots, \int_{t_0}^{t_1} c_n \right).$$

Les règles de calcul pour les fonctions avec valeur dans $\mathbb{R}$ se généralisent d'une façon évidente. Par exemple, on a

$$(c \cdot d)' = c' \cdot d + c \cdot d'$$

pour deux courbes c et d . Comme un autre exemple, on a pour toute courbe c

$$c(T) = \int_{t_0}^T c'(t) dt + c(t_0).$$

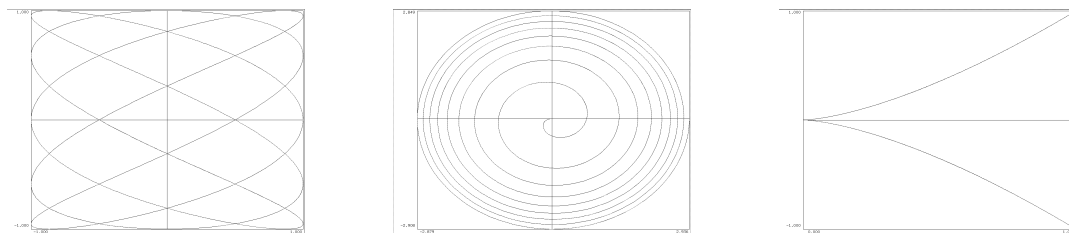
En fait, la courbe définie par le côté droit et c ont la même dérivée, donc sont égales à l'addition d'un vecteur constant a près. En plus, ils passent le même point pour l'argument t_0 , d'où $a = 0$.

Parfois on écrit $\dot{c}$ et $\ddot{c}$ au lieu de $c' = c^{(1)}$ et $c'' = c^{(2)}$.

Exemple. On fixe un point $a \in \mathbb{R}^2$, un vecteur $b \in \mathbb{R}^2$, et des nombres réels $R > 0$ et $l \neq 0$. Ici quelques exemples de courbes :

1. La droite $c(t) = a + bt$ ($t \in \mathbb{R}$).
2. Le cercle (ou l'arc) $c(t) = a + R(\cos t, \sin t)$ ($t \in \mathbb{R}$).
3. La spirale logarithmique $c(t) = \log t (\cos t, \sin t)$ ($t \in (0, +\infty)$).
4. Un "quart huit" (lemniscate) $c(t) = r(\sqrt{\frac{1+r^2}{2}}, \sqrt{\frac{1-r^2}{2}})$ ($0 \leq t \leq 1$).
Le lemniscate est le lieu géométrique des points dont le produit des distances à $(\frac{1}{\sqrt{2}}, 0)$ et à $(-\frac{1}{\sqrt{2}}, 0)$ est égal à $\frac{1}{2}$.
5. L'hélice $c(t) = (R \cos(t), R \sin(t), lt)$ ($t \in \mathbb{R}$).

6. Pour $m, n \in \mathbb{Z}$ non nuls, la courbe de Lissajou $c(t) = (\cos(mt), \sin(nt))$ ($t \in \mathbb{R}$).
7. La pointe $c(t) = (t^2, t^3)$ ($t \in \mathbb{R}$).



Lissajou, spirale et pointe

Remarque. L'exemples 2. (avec $a = 0$) et 3. sont des cas spéciaux des courbes qui sont données sous la forme

$$c(t) = \rho(t) (\cos \theta(t), \sin \theta(t))$$

avec des fonctions infiniment dérivables $\rho(t) \geq 0$ et $\theta(t)$. Donc $\rho(t)$ est la distance de $(0, 0)$ à $c(t)$, et $\theta(t)$ est l'angle orienté entre le vecteur $(1, 0)$ et le vecteur de $(0, 0)$ à $c(t)$ (à l'addition d'un multiple de 2π près). Autrement dit, $\rho(t)$ et $\theta(t)$ sont les coordonnées polaires du point $c(t)$.

Notation. On appelle $c'(t)$ (la direction de) la tangente à la courbe c . L'interprétation géométrique est immédiate de la formule

$$c'(t) = \lim_{h \rightarrow 0} \frac{c(t+h) - c(t)}{h}.$$

Si on s'imagine t comme le temps et $c(t)$ comme chemin parcouru par un particle, alors on appelle $c'(t)$ aussi la vitesse et $c''(t)$ l'accélération.

Définition. La longueur d'une courbe c de t_0 à t_1 est définie comme

$$L(c|_{[t_0, t_1]}) = \int_{t_0}^{t_1} |c'(t)| dt.$$

Exemple. Pour le cercle $c(t) = (\cos(t), \sin(t))$ ($0 < t < \alpha$) on trouve $|c'(t)| = (\sin^2 t + \cos^2 t)^{\frac{1}{2}} = 1$ et donc

$$L(c|_{[0, \alpha]}) = \alpha.$$

Notation. Une courbe est dite de vitesse 1 si $|c'(t)| = 1$ pour tout $t \in I$.

Remarque. Pour une courbe de vitesse 1 on a toujours

$$L(c|_{[\alpha, \beta]}) = \beta - \alpha.$$

Donc une telle courbe est paramétrisée par sa propre longueur.

2.3 La courbure d'une courbe plane

Une courbe paramétrée est un chemin parcouru par une particule. Dans la vie quotidienne on décrit souvent un chemin par des informations “locales” : Rapidement à gauche, puis lentement tout droit, puis encore plus lentement à droite etc. Ces descriptions locales admettent de trouver le chemin global et vice versa. Il s'agit de faire une théorie pareille pour les courbes paramétrées c . Les mots “rapidement”, “lentement” se traduisent facilement par la vitesse (scalaire) $|c'(t)|$ — seulement la valeur absolue, pas la direction.

Pour traduire les “gauche” et “droite” il faut introduire d'abord un sens de droite et gauche si on marche le long de la courbe. La bonne notion est le repère de Frenet. Pour le définir il nous faut quelques hypothèses supplémentaires.

Dèsormais nous supposons que toutes nos courbes sous considération sont infiniment dérivables. En plus, nous supposons qu'elles sont régulières.

Notation. Une courbe $c : I \rightarrow \mathbb{R}^n$ est appelée régulière si $c'(t) \neq 0$ pour tout $t \in I$.

Comme exemple d'une courbe qui n'est pas régulière on peut considérer la pointe (en $t=0$; voir ci-dessus).

Définition. Soit $c : I \rightarrow \mathbb{R}^2$ une courbe régulière. La famille $\{e_1, e_2\}$ de courbes $e_1, e_2 : I \rightarrow \mathbb{R}^2$ définies par

$$e_1(t) = \frac{c'(t)}{|c'(t)|}, \quad e_2(t) = \times e_1(t)$$

est appelée le repère de Frenet associé à la courbe c .

Remarque. Il faut s'imaginer $e_1(t), e_2(t)$ comme vecteurs attachés à $c(t)$, et donc comme un repère de $\mathbb{R}^2$ qui séjourne avec le particle dont le voyage est décrit par $c(t)$. On remarque que e_2 donne toujours à gauche par rapport à la direction de mouvement du particle; d'où un sens de gauche et droite donné au mouvement.

Remarque. $e_1(t)$ et $e_2(t)$ sont infiniment dérivables.

Théorème 2.2. (Identité de Frenet) $\omega(t) := e_1'(t) \cdot e_2(t) = -e_2'(t) \cdot e_1(t)$.

Remarque. On a

$$e_1'(t) = \omega(t) e_2(t).$$

En effet, e_1' est un multiple de e_2 : prendre la dérivée de $1 = e_1 \cdot e_1$ donne $e_1 \cdot e_1' = 0$. En écrivant $e_1' = \lambda e_2$ avec un λ convenable et en multipliant par e_2 on trouve $e_1 \cdot e_2 = \lambda$. De même on peut vérifier que

$$e_2' = -\omega e_1.$$

Démonstration. On a $e_1(t) \cdot e_2(t) = 0$, d'où — en prenant la dérivée —

$$e_1'(t) \cdot e_2(t) + e_1(t) \cdot e_2'(t) = 0.$$

□

Théorème 2.3. Soient $c : I \rightarrow \mathbb{R}^2$ et $\tilde{c} : \tilde{I} \rightarrow \mathbb{R}^2$ des courbes régulières avec des repères de Frenet e_j et $\tilde{e}_j$ respectivement. On pose $\omega = e_1 \cdot e_2$ et $\tilde{\omega} = \tilde{e}_1 \cdot \tilde{e}_2$. On suppose qu'il existe une application strictement croissante et infiniment dérivable $\phi : I \rightarrow \tilde{I}'$ telle que $c(t) = \tilde{c}(\phi(t))$ pour tout $t \in I$. Alors on a

$$\frac{\omega(t)}{|c'(t)|} = \frac{\tilde{\omega}(\phi(t))}{|\tilde{c}'(\phi(t))|}.$$

Remarque. Dans la situation décrite dans le théorème on dit que c provient de $\tilde{c}$ par un changement de paramètre.

Démonstration. On a $c' = (\tilde{c} \circ \phi) \phi'$, donc $|c'| = |\tilde{c}' \circ \phi| |\phi'|$ et $e_1 = \tilde{e}_1 \circ \phi$, d'où

$$\omega e_2 = e_1' = (\tilde{e}_1' \circ \phi) \phi' = (\tilde{\omega} \circ \phi) (\tilde{e}_2 \circ \phi) \phi' = (\tilde{\omega} \circ \phi) (\tilde{e}_2 \circ \phi) \frac{|c'|}{|\tilde{c}' \circ \phi|}$$

En comparant les valeurs absolues on en obtient le résultat. □

Définition. Soit c une courbe régulière. La quantité

$$\kappa(t) := \frac{\omega(t)}{|c'(t)|}$$

est appelée la courbure de c .

Remarque. Avec les notations du dernier théorème on a

$$\kappa(t) = \tilde{\kappa}(\phi(t)).$$

Donc la courbure est une quantité “invariante” de la courbe c , i.e. une quantité qui ne change pas si on change le paramètre.

Remarque. Si c est de vitesse 1, alors les définitions simplifient et on a

$$c' = e_1, \quad c'' = \kappa e_2.$$

En particulier, on a

$$|c''| = |\kappa|.$$

Remarque. Une **première interprétation géométrique de la courbure** est la suivante : Le vecteur c'' donne dans la direction du changement de c' (regarder dans un dessin $c''(t) \approx \frac{c'(t+h)-c'(t)}{h}$ pour un petit h). Mais $c'' = (|c'|e_1)' = |c'|'e_1 + |c'|e_1'$, i.e.

$$c'' = |c'|'e_1 + |c'|^2\kappa e_2,$$

d'où $c'' \cdot e_2 = |c'|^2\kappa$. Donc, si $\kappa(t) > 0$, alors l'angle entre $c''(t)$ et $e_2(t)$ est acut ($< \frac{\pi}{2}$), i.e. $c''(t)$ donne à gauche : la courbe tourne à gauche. Si $\kappa(t) < 0$, alors l'angle entre $c''(t)$ et $e_1(t)$ est obtuse (i.e. $> \frac{\pi}{2}$) et c'' donne à droite : la courbe tourne à droite. Si $\kappa(t) = 0$, mais $\kappa'(t) \neq 0$, alors κ change signe en t , et donc la courbe change la direction en t : on parle d'un point d'inflexion en t .

Exemple. On considère la courbe $c(t) = (t, \sin t)$ pour $t \in]0, \pi[$. Alors

$$c'(t) = (1, \cos t), \quad e_1(t) = \frac{(1, \cos t)}{\sqrt{1 + \cos^2 t}}, \quad e_2(t) = \frac{(-\cos t, 1)}{\sqrt{1 + \cos^2 t}},$$

$$\kappa(t) = e_1'(t) \cdot e_2(t) = \frac{-\sin t}{(1 + \cos^2 t)^{\frac{3}{2}}}$$

Donc $\kappa(t) < 0$ pour $t \in]0, \pi[$ et $\kappa(t) > 0$ pour $t \in]\pi, 0[$. En $t = \pi$ on a un point d'inflexion.

Comme l'exemple précédent montre (so on fait le calcul de $\kappa(t)$) il est assez pénible de calculer la courbure directement selon sa définition. Une formule plus confortable est la suivante :

Théorème 2.4. (Formule pour la courbure)

$$\kappa(t) = \frac{c_1'(t)c_2''(t) - c_2'(t)c_1''(t)}{|c'(t)|^3}.$$

Remarque. En utilisant les déterminants (et en supprimant t) on peut écrire cette formule plus courte sous la forme

$$\kappa = \frac{\det(c'c'')}{|c'|^3}$$

Démonstration. Comme ci-dessus $c'' = |c'|'e_1 + |c'|^2\kappa e_2$, d'où $\det(c'c'') = \det(|c'|e_1, |c'|^2\kappa e_2) = \kappa |c'|^3$. $\square$

Exemple. Soit $c(t) = R(\cos t, \sin t)$ un cercle de rayon R . Alors on a

$$\kappa(t) = \frac{1}{R}.$$

Remarque. Cet exemple donne la **deuxième interprétation géométrique de la courbure** : Le cercle avec centre $c(t) + \frac{1}{\kappa(t)}e_2(t)$ et rayon $\frac{1}{|\kappa(t)|}$ touche la courbe c en $P = c(t)$. Sa tangente en P et $e_1(t)$ ont la même direction (exercice), et en P sa courbure est la même que la courbure de c . Donc localement il donne une bonne approximation de c en P . C'est pour ça qu'on appelle $1/|\kappa(t)|$ aussi le rayon de courbure de c en P .

Théorème 2.5. *Soit c une courbe régulière. On suppose que c' est donnée sous la forme*

$$c'(t) = |c'(t)| (\cos \theta(t), \sin \theta(t))$$

avec une fonction dérivable $\theta(t)$. Alors on a

$$\kappa = \frac{\theta'}{|c'|}.$$

Remarque. Ce théorème donne la **troisième interprétation géométrique de la courbure** : $\theta(t)$ est l'angle entre le vecteur $(1, 0)$ et la vitesse. Donc, pour une courbe de vitesse 1 on peut considérer $\kappa(t)$ comme le taux de changement de la direction de la courbe en temps t , i.e. le taux de changement de cet angle en temps t .

Démonstration. Par calcul direct on trouve, en posant $v := |c'|$,

$$c'' = v' (\cos \theta, \sin \theta) + v\theta' (-\sin \theta, \cos \theta), \quad \kappa = \det(c'c'')/v^3 = \theta'/v.$$

□

On rappelle qu'une droite dans $\mathbb{R}^2$ est un ensemble de la forme $\{x \in \mathbb{R}^2 : x \cdot a = \alpha\}$ avec un $0 \neq a \in \mathbb{R}^2$ et un $\alpha \in \mathbb{R}$ donnés.

Théorème 2.6. (Courbures zéro) *Soit c une courbe régulière. Alors $\kappa \equiv 0$ si et seulement si $a \cdot c(t) = \alpha$ avec $0 \neq a \in \mathbb{R}^2$ et $\alpha \in \mathbb{R}$ convenables.*

Démonstration. On utilise la formule

$$c(T) = \int_{t_0}^T c'(t) dt + c(t_0)$$

avec un t_0 fixé.

Si $\kappa = 0$, alors $e_1' = 0$, i.e. e_1 est un vecteur constant. Donc, en écrivant $c'(t) = |c'(t)|e_1(t_0)$, on obtient

$$c(T) = e_1(t_0) \int_{t_0}^T |c'(t)| dt + c(t_0),$$

et donc $e_2(t_0) \cdot c(t) = -e_2(t_0) \cdot c(t_0)$ pour tout t .

Réciproquement, si $a \cdot c = \alpha$ avec des constantes $a \neq 0$ et α , alors $a \cdot c' = 0$, $a \cdot c'' = 0$, donc c' et c'' sont des multiples de $\times a$ (ici on utilise $a \neq 0$), et d'où $\kappa = 0$ par la formule pour la courbure. $\square$

Théorème 2.7. (Courbure constante non nulle) *Soit $c : I \rightarrow \mathbb{R}^2$ une courbe régulière, et soit R un nombre réel $R > 0$. Alors les 2 propositions suivantes sont équivalentes :*

1. $|\kappa(t)| = 1/R$ pour tout $t \in I$.
2. Il existe $a \in \mathbb{R}^2$ tel que $|c(t) - a| = R$ pour tout $t \in I$.

Démonstration. Si $(c - a) \cdot (c - a) = R^2$, alors — en prenant la dérivée de cette identité — $c' \cdot (c - a) = 0$, et donc $c - a$ est un multiple de e_2 . Car sa longueur est R on a $c - a = \varepsilon R e_2$ avec un $\varepsilon \in \{\pm 1\}$. Prenant la dérivée de cette identité donne, en utilisant les identités de Frenet,

$$c' = \varepsilon R e_2' = -\varepsilon R \kappa |c'| e_1,$$

et en considérant les valeurs absolues $1 = R |\kappa|$.

Réciproquement $|\kappa| = 1/R$ entraîne

$$e_2' = -\varepsilon |c'| e_1 / R$$

avec $\varepsilon = \pm 1$ convenable. Autrement dit

$$0 = |c'| e_1 + \varepsilon R e_2' = c' + \varepsilon R e_2',$$

i.e. $c + \varepsilon R e_2 = a$ pour un vecteur constant a , ce qui entraîne $|c - a| = R$. $\square$

Les deux théorèmes précédents sont vrai en plus de généralité, i.e. la courbure détermine la courbe uniquement (si la vitesse scalaire $|c'(t)|$ est quelques valeurs initiales sont connues).

Théorème 2.8. *Soient $c, \tilde{c} : I \rightarrow \mathbb{R}^2$ deux courbes régulières. On suppose $c(t_0) = \tilde{c}(t_0)$ et $c'(t_0) = \tilde{c}'(t_0)$ pour un $t_0 \in I$, et que $|c'(t)| = |\tilde{c}'(t)|$ et $\kappa(t) = \tilde{\kappa}(t)$ pour tout $t \in I$. Alors $c = \tilde{c}$.*

Démonstration. Avec une fonction convenable on peut écrire $c'(t)$ sous la forme

$$c'(t) = |c'(t)| (\cos \theta(t), \sin \theta(t))$$

pour tout $t \in I$. On utilise sans démonstration le fait que l'on peut même choisir θ dérivable. On a vu que $\theta' = \kappa |c'|$, et donc

$$\theta(t) = \int_{t_0}^t \kappa(s) |c'(s)| ds + \theta(t_0).$$

En résumant on obtient ainsi

$$c(T) = \int_{t_0}^T |c'(t)| \left(\cos \left[\int_{t_0}^t \kappa(s) |c'(s)| ds + \theta(t_0) \right], \right. \\ \left. \sin \left[\int_{t_0}^t \kappa(s) |c'(s)| ds + \theta(t_0) \right] \right) dt + c(t_0).$$

Cette formule exprime $c(T)$ en fonction de κ , $|c'|$ et les valeurs initiales $c(t_0)$ et $c'(t_0)$. D'où le théorème. $\square$

Théorème 2.9. (Résultat principale sur la courbure) Soit $\kappa(t)$ une fonction continue dans l'intervalle ouvert I , et soient $t_0 \in I$ et $a, b \in \mathbb{R}^2$, $|b| = 1$. Alors il existe une et une seule courbe $c : I \rightarrow \mathbb{R}^2$ de vitesse 1 avec courbure égal à κ et telle que $c(t_0) = a$ et $c'(t_0) = b$.

Démonstration. L'existence : On vérifie que la courbe donnée par la formule dans la démonstration du dernier théorème avec $|c'| = 1$, i.e.

$$c(T) = \int_{t_0}^T \left(\cos \left[\int_{t_0}^t \kappa(s) ds + \theta_0 \right], \sin \left[\int_{t_0}^t \kappa(s) ds + \theta_0 \right] \right) dt + a.$$

(avec θ_0 tel que $b = (\cos \theta_0, \sin \theta_0)$) satisfait aux propriétés en question. L'unicité est clair d'après le dernier théorème. $\square$

Chapitre 3

Fonctions en deux variables

3.1 Vocabulaire de base

Un rectangle R est un sous-ensemble de $\mathbb{R}^2$ de la forme

$$R = I \times J = \{(x, y) : x \in I, y \in J\}$$

où I et J sont des intervalles usuels de nombres réels. On appelle R fermé ou ouvert si I et J sont tous les deux fermés ou ouverts respectivement. Donc, R est fermé si R est de la forme $R = [a_1, a_2] \times [b_1, b_2]$ avec des nombres réels a_j, b_k .

Une suite $\{a_n\}$ dans $\mathbb{R}^2$ est une famille de points

$$a_n = (a_n^{(1)}, a_n^{(2)})$$

avec des suites réelles usuelles de nombres réels $\{a_n^{(1)}\}$ et $\{a_n^{(2)}\}$. On définit

$$\lim_n a_n := (\lim_n a_n^{(1)}, \lim_n a_n^{(2)}),$$

si les deux limites à gauche existent.

Une fonction en deux variable est une application

$$f : R \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y)$$

d'un rectangle R dans l'ensemble des nombres réels.

Si a est un point sur la borne de R on écrit

$$\beta = \lim_{(x,y) \rightarrow a} f(x, y)$$

si $\beta = \lim_n f(a_n)$ pour toute suite $\{a_n\}$ dans R avec $\lim_n a_n = a$.

Exercice. Montrer

$$\lim_{(x,y) \rightarrow 0} \frac{\sin(x^2 + y^3)}{x^2 + y^3} = 1.$$

On appelle f continue en $a \in R$ si

$$\lim_{(x,y) \rightarrow a} f(x, y) = f(a),$$

et on appelle f continue sur R si f est continue en tout a dans R .

Remarque. Plus court on peut dire que F est continue dans R si pour toute suite $\{a_n\}$ avec limite dans R on a

$$\lim_n f(a_n) = f\left(\lim_n a_n\right) :$$

“On peut échanger la limite et l’application f .”

Exercice. Montrer que le produit, la somme, le quotient (si le dénominateur n’annule pas) et la composition (si les domaines conviennent) de fonctions continues sont continues.

En particulier les polynômes en deux variables sont continus sur tout $\mathbb{R}^2$. Un polynôme est une fonction de la forme

$$f(x, y) = \sum_{i=0}^m \sum_{j=0}^n a_{i,j} x^i y^j$$

avec des constantes (les coefficients) $a_{i,j}$. En fait, une telle fonction est somme de produits des fonctions constantes et $g(x, y) = x$ et $h(x, y) = y$ qui sont évidemment continues.

Théorème 3.1. Soit $f : R \rightarrow \mathbb{R}$ continue sur l’intervalle fermé R . Alors $f(R)$ est un interval fermé dans $\mathbb{R}$.

Remarque. Le théorème est faux si R n’est pas fermé. Contre-exemple : $f(x, y) = \frac{x}{y}$ et $R = [0, 1] \times (0, 1]$. Le théorème assure en particulier qu’une fonction continue possède un maximum et un minimum globale sur un rectangle fermé.

Le théorème peut être facilement réduit au théorème analogue pour les fonctions continues en une variable en considérant des “tranches” qui seront introduites ci-dessous. On admet la démonstration.

Pour la représentation graphique d’une fonction $f(x, y)$ (défini sur un rectangle R) on a au moins trois possibilités ‘:

1. On trace (la projection sur le plan de) le graphe $G_f := \{(x, y, f(x, y)) : (x, y) \in R\} \subset \mathbb{R}^3$. On observe que le graphe est une “surface” dans l’espace en 3 dimensions.
2. On trace dans R les courbes de niveau de f , i.e. on trace les courbes $f(x, y) = \alpha$ pour plusieurs constantes α . Il existe un théorème (“sur les fonctions implicites”) qui assure que sous certain hypothèse raisonnable les ensembles des $(x, y) \in R$ tels que $f(x, y) = \alpha$ sont des images de courbes paramétrées.
3. Comme dans un atlas on colorie les points (x, y) du rectangle R avec des couleurs variées selon les valeurs $f(x, y)$.

L’étude d’une fonction $f(x, y)$ en deux variables peut souvent être réduit à l’étude des fonctions d’une seule variable en considérant des tranches de f . La tranche de f le long d’une courbe paramétrée $c(t)$ (sur un intervalle I avec image dans R) est la fonction numérique

$$f_c : I \rightarrow \mathbb{R}, \quad t \mapsto f(c(t)).$$

Les tranches les plus simples et plus utiles sont les tranches le long des droites:

$$t \mapsto f(a + bt)$$

avec $a, b \in R$. On peut considérer le graphe d’une tranche comme une “tranche” du graphe de f dans le sens ordinaire.

3.2 Dérivées de fonctions en 2 variables

Soit f une fonction définie sur un rectangle ouvert R dans $\mathbb{R}^2$.

Définition. On dit que f est dérivable en $a \in R$ si il existe un vecteur $\delta \in \mathbb{R}^2$ tel que

$$\lim_{h \rightarrow 0} \frac{|f(a + h) - f(a) - \delta \cdot h|}{|h|} = 0.$$

On appelle δ la dérivée de f en a , noté $f'(a)$ (ou $\text{grad } f(a)$ ou $\nabla f(a)$).

Théorème 3.2. Si f est dérivable en a , alors f est continue en a .

Démonstration. Copier la démonstration pour les fonctions d’une seule variables, □

Pour calculer $f'(a)$ on introduit la notion des dérivées partielles.

Définition. On dit que $f(x, y)$ est partiellement dérivable en $a = (a_1, a_2) \in R$ par rapport à x si la tranche $\phi(x) := f(x, a_2)$ est dérivable en a_1 . La dérivée est notée

$$\frac{\partial f}{\partial x}(a).$$

Analogue on pose

$$\frac{\partial f}{\partial y}(a) = \lim_{y \rightarrow a_2} \frac{f(a_1, y) - f(a_1, a_2)}{y - a_2},$$

si cette limite existe. On dit que f est partiellement dérivable en a (resp. dans R) si toutes les deux dérivées partielles existe en a (resp. en tout $a \in R$).

Remarque. Si f est partiellement dérivable dans le rectangle R on a ainsi deux nouvelles fonctions définies sur R : les dérivées partielles de f définies par

$$\frac{\partial f}{\partial x} : (x, y) \mapsto \frac{\partial f}{\partial x}(x, y), \quad \frac{\partial f}{\partial y} : (x, y) \mapsto \frac{\partial f}{\partial y}(x, y).$$

Théorème 3.3. Soit f dérivable en a . Alors f est partiellement dérivable par rapport à x et y et on a

$$f'(a) = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a) \right).$$

Démonstration. Immédiatement de la définition de $f'(a)$ on a avec $f'(a) = \delta = (\delta_1, \delta_2)$

$$\lim_{h_1 \rightarrow 0} \frac{|f(a + h_1, a_2) - f(a_1, a_2) - \delta_1 h_1|}{|h_1|} = 0.$$

Donc on a aussi

$$\lim_{h_1 \rightarrow 0} \frac{f(a + h_1, a_2) - f(a_1, a_2) - \delta_1 h_1}{h_1} = 0,$$

i.e.

$$\delta_1 = \lim_{h_1 \rightarrow 0} \frac{f(a + h_1, a_2) - f(a_1, a_2)}{h_1} = \frac{\partial f}{\partial x}(a).$$

La preuve pour y est la même. □

Réciproquement on a (on admet la démonstration) :

Théorème 3.4. On suppose que f est partiellement dérivable dans un rectangle ouvert contenant a , et que les dérivées partielles sont continue en a . Alors f est dérivable en a .

Remarque. Si les dérivées partielles de f existent dans R , alors on ne peut pas conclure sans hypothèses supplémentaires que f est dérivable en a . Il se peut que f est partiellement dérivable mais n'est pas même continue en a . Contre-exemple : La fonction

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) \neq 0 \\ 0 & \text{si } (x, y) = 0 \end{cases}$$

n'est pas continue en 0 car

$$\lim_n f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{1}{2} \neq f(0, 0).$$

Mais elle est partiellement dérivable dans $\mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} \frac{y(y^2-x^2)}{(x^2+y^2)^2} & \text{si } y \neq 0 \\ 0 & \text{si } y = 0 \end{cases}, \quad \frac{\partial f}{\partial y}(x, y) = \begin{cases} \frac{x(x^2-y^2)}{(x^2+y^2)^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

Si f est partiellement dérivable dans R il se peut que ses dérivées partielles sont encore partiellement dérivable, donc on peut considérer

$$\begin{aligned} \frac{\partial^2}{\partial x^2} &:= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right), & \frac{\partial^2}{\partial y \partial x} &:= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right), \\ \frac{\partial^2}{\partial x \partial y} &:= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right), & \frac{\partial^2}{\partial y^2} &:= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right), \end{aligned}$$

On peut continuer à prendre les dérivées partielles (si ils existent). Donc il se peut que l'on obtient la fonction

$$\frac{\partial^n f}{\partial x^{n_1} \partial y^{n_2} \dots \partial x^{n_{p-1}} \partial y^{n_p}},$$

en dérivant f partiellement n_p -fois par rapport à y , puis n_{p-1} fois par rapport à x etc, où $n_1 + \dots + n_p = n$. Cette fonction est appelée une dérivée partielle d'ordre n .

Exercice. Combien de dérivées partielles d'ordre n existent-ils a priori ?

Définition. On dit que f est n fois continuellement dérivable, si toutes les dérivées partielles jusqu'à l'ordre n existent et sont continues dans R .

Pour manipuler les dérivées partielles de produit, sommes, quotients de fonctions ou d'une composition de la forme $g(f(x, y))$ avec une fonction en une variable g , on applique les règles pour les fonctions d'une seule variable : c'est possible car les dérivées partielles sont des dérivées de tranches, donc de fonctions en une variable. Il y a deux règles supplémentaires “

Théorème 3.5. Soit f deux fois continuellement dérivables dans R . Alors

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Démonstration. Admise. □

Théorème 3.6. Soient g, h partiellement dérivable dans le rectangle ouvert R et avec valeurs dans un rectangle ouvert S , et soit f dérivable dans S . Alors la composition

$$f(g(x, y), h(x, y))$$

est partiellement dérivable dans R , et on a

$$\begin{aligned} \frac{\partial}{\partial x} f(g, h) &= \frac{\partial f}{\partial x}(g, h) \cdot \frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}(g, h) \cdot \frac{\partial h}{\partial x}, \\ \frac{\partial}{\partial y} f(g, h) &= \frac{\partial f}{\partial x}(g, h) \cdot \frac{\partial g}{\partial y} + \frac{\partial f}{\partial y}(g, h) \cdot \frac{\partial h}{\partial y}. \end{aligned}$$

Si g et h sont dérivables dans R , alors $f(g, h)$ est dérivable dans R aussi.

Démonstration. Exercice. □

Remarque. Les formules s'écrivent plus simples sous la forme

$$\nabla(f(g, h)) = (\nabla f)(g, h) \begin{pmatrix} \nabla g \\ \nabla h \end{pmatrix}.$$

Remarque. Un cas spécial important est le cas d'une composition $f(g, h)$ avec des fonctions $g(x)$ et $h(x)$ d'une seule variable. En les considère comme fonction en deux variable en posant $\tilde{g}(x, y) = g(x)$ et analogue pour h . On a ainsi $\frac{\partial \tilde{g}}{\partial x}(x, y) = g'(x)$, et donc

$$\frac{d}{dt} f(g, h) = \frac{\partial f}{\partial x}(g, h) g'(x) + \frac{\partial f}{\partial y}(g, h) h'(x).$$

Exemple. Pour calculer la dérivée de

$$\phi(t) = \int_{\log t}^{t^2} \sin(s) ds$$

on pose

$$f(x, y) = \int_x^y \sin(s) ds,$$

et donc

$$\phi'(t) = \frac{d}{dt} f(\log t, t^2) = -\sin(\log t) \cdot \frac{1}{t} + \sin(t^2) \cdot 2t.$$

On peut vérifier le résultat en calculant directement la dérivée en utilisant

$$\phi(t) = \cos(\log t) - \cos(t^2).$$

3.3 Etudes locales

Soit f dérivable dans le rectangle ouvert R . Si v est un vecteur de $\mathbb{R}^2$ et $a \in R$, on pose

$$\frac{\partial f}{\partial v} = f'(a) \cdot v,$$

et on l'appelle dérivée de f dans la direction v . On a d'après les règles pour calculer des dérivées

$$\frac{\partial f}{\partial v} = \frac{d}{dt} f(a + tv) \Big|_{t=0}.$$

Donc $\frac{\partial f}{\partial v}(a)$ est la pente en a de la tranche de f le long de la droite $t \mapsto a + tv$. Dans quelle direction v , disons $|v| = 1$, est cette pente maximal ou minimal ?

Théorème 3.7. *Soit $f'(a) \neq 0$ et $u = f'(a)/|f'(a)|$. Alors, pour tout $v \in \mathbb{R}^2$, $|v| = 1$ et $v \neq u$, on a*

$$\frac{\partial f}{\partial u}(a) = |f'(a)| > \frac{\partial f}{\partial v}(a).$$

Démonstration. On a pour $|v| = 1$

$$\frac{\partial f}{\partial v}(a) = f'(a) \cdot v = |f'(a)| \cos \theta,$$

ou $0 \leq \theta < 2\pi$ est l'angle entre $f'(a)$ et v . Si $v \neq u$ on a $\theta \neq 0$, i.e. $\cos \theta < 1$, et donc

$$\frac{\partial f}{\partial v}(a) < |f'(a)| = \frac{\partial f}{\partial u}(a).$$

D'où le théorème □

Pour étudier f dans un petit voisinage d'un point on peut s'imaginer f comme polynôme. C'est justifié par le théorème de Taylor :

Théorème 3.8. (Formule de Taylor.) *Soit f une fonction $(n+1)$ -fois continuellement dérivable dans un rectangle ouvert R , soit $a \in R$. Alors pour tout h tel que $a + h \in R$ on a*

$$\begin{aligned} f(a + h) = H_{a,0} + H_{a,1}(h) + \frac{1}{2!} H_{a,2}(h) + \frac{1}{3!} H_{a,3}(h) + \cdots + \\ + \frac{1}{n!} H_{a,n}(h) + \frac{1}{n+1!} H_{a+\xi h, n+1}(h), \end{aligned}$$

avec un $\xi \in (0, 1)$, et où, pour un b dans R et un nombre naturel k , on a

$$H_{b,k}(h) = \sum_{p=0}^k \binom{k}{p} \frac{\partial f}{\partial x^p \partial y^{k-p}}(b) h_1^p h_2^{k-p} = \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y} \right)^k f(b).$$

Démonstration. On applique la formule de Taylor pour les fonctions d'une seule variable à la tranche

$$\phi(t) = f(a + th).$$

On a donc

$$f(a + h) = \phi(1) = \phi(0) + \phi'(0) + \frac{1}{2!}\phi''(0) + \cdots + \frac{1}{n!}\phi^{(n)}(0) + \frac{1}{(n+1)!}\phi^{(n)}(\xi),$$

avec un $\xi \in (0, 1)$. Par récurrence sur k on montre que

$$\phi^{(k)}(t) = \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y} \right)^k f(a + th).$$

D'où le théorème. □

Remarque. On observe que les $H_{a,k}$ pour a fixé sont des polynômes en deux variables h_1 et h_2 . Comme corollaire on obtient la formule un peu plus faible

$$f(a + h) = H_{a,0} + H_{a,1}(h) + \frac{1}{2!}H_{a,2}(h) + \cdots + \frac{1}{n!}H_{a,n}(h) + o(|h|^n)$$

pour $|h| \rightarrow 0$.

Exercice. Montrer la formule précédent, i.e. montrer

$$\lim_{h \rightarrow 0} \frac{|H_{a+\xi h, n+1}(h)|}{|h|^n} = 0.$$

Remarque. Un cas particulièrement intéressant est le cas $n = 2$. Ici on peut écrire la formule de Taylor sous la forme

$$f(a + h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h^t \cdot \begin{pmatrix} \frac{\partial f}{\partial x^2}(a) & \frac{\partial f}{\partial x \partial y}(a) \\ \frac{\partial f}{\partial y \partial x}(a) & \frac{\partial f}{\partial y^2}(a) \end{pmatrix} \cdot h + o(|h|^2).$$

Les points signifient multiplication matricielle. La matrice carrée est appelée Hessian de f .

Théorème 3.9. *Soit f dérivable dans le rectangle R . Si f possède un minimum ou maximum locale en $a \in R$, alors $f'(a) = 0$.*

Remarque. On appelle a un minimum locale si il existe un rectangle ouvert R' dans R contenant a tel que $f(a) \leq f(x, y)$ pour tout $(x, y) \in R'$. Analogie on définit un maximum local.

Démonstration. Si f a un minimum local en a alors en particulier les tranches $x \mapsto f(x, a_2)$ et $y \mapsto f(a_1, y)$ possèdent un minimum local en a_1 et a_2 respectivement. Donc leurs dérivées s'annulent. $\square$

Comme dans le cas des fonctions d'une seule variable on ne peut pas conclure de $f'(a)$ que f possède un maximum ou minimum locale en a . Comme pour une variable il faut considérer les dérivées partielles d'ordre supérieur à 1 aussi.

Théorème 3.10. *Soit f infiniment dérivable dans le rectangle ouvert R . On suppose qu'il existe un $n \geq 1$ tel que $H_{a,1} \equiv H_{a,2} \equiv \dots \equiv H_{a,n-1} \equiv 0$ et tel que $P := H_{a,n} \not\equiv 0$. Pour $\theta \in \mathbb{R}$ on pose*

$$f_\theta(t) = f(a + t(\cos \theta, \sin \theta)).$$

Soit $p \geq 0$ le nombre des racines de l'équation $P(\cos \theta, \sin \theta) = 0$ dans l'intervalle $[0, \pi)$, et, si $p \geq 1$ soient $0 \leq \alpha_1 < \dots < \alpha_p < \pi$ les racines. On a :

1. *Si n est pair et $p = 0$, alors f possède un maximum ou minimum locale en a (voir fig. 1 (a)).*
2. *Si n est impair, alors pour tout $\theta \neq \alpha_j$ ($1 \leq j \leq p$) $f_\theta(t)$ a ni un minimum local ni un maximum local en $t = 0$ (voir fig. 1 (b)).*
3. *Si n est pair et $p \geq 1$, alors pour chaque intervalle I de la forme $[0, \alpha_1)$, (α_1, α_2) , $\dots$, (α_{p-1}, α_p) ou (α_p, π) la fonction $f_\theta(t)$ possède en $t = 0$ soit un minimum local pour tout $\theta \in I$, soit un maximum local pour tout $\theta \in I$ (voir fig. 2 (b)).*

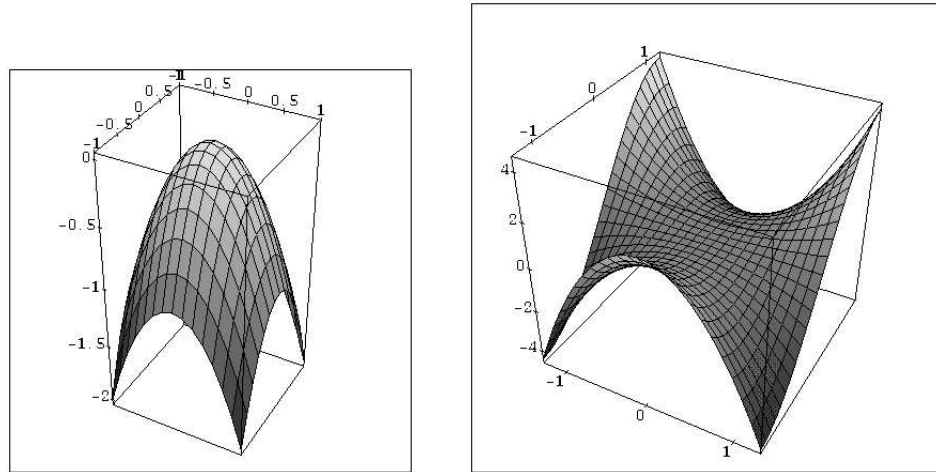
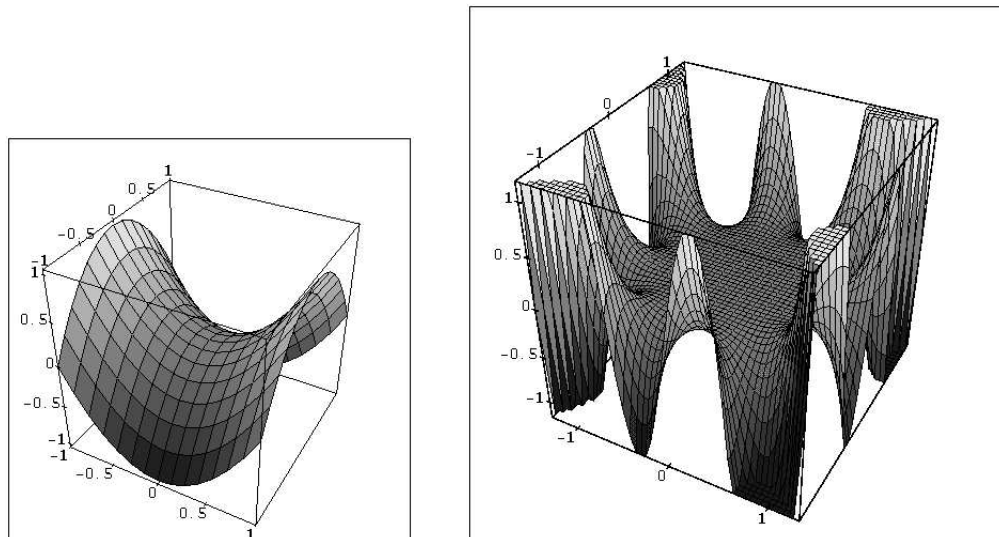
Remarque. On remarque qu'il n'existe qu'un nombre fini de racines de $P(\cos \theta, \sin \theta) = 0$ dans l'intervalle $[0, 2\pi)$. En fait, pour $\theta \neq 0$ on a

$$\begin{aligned} P(\cos \theta, \sin \theta) = 0 &\iff \sin^n(\theta) P(\cot \theta, 1) = 0 \\ &\iff \exists \rho \in \mathbb{R} : P(\rho, 1) = 0 \text{ et } \theta = \text{acot}(\rho). \end{aligned}$$

Or $P(t, 1)$ est un polynôme de degré $\leq n$, donc possède au plus n racines réelles.

Démonstration. On a comme dans la démonstration de la formule de Taylor

$$f_\theta^{(k)}(0) = H_{a,k}(\cos \theta, \sin \theta).$$

Figure 3.1: (a) $x^2 + y^2$ et (b) $y(3x^2 - y^2)$ Figure 3.2: (a) $x^2 - y^2$ et (b) $x^7y - 7x^5y^3 + 7x^3y^5 - xy^7$

Donc en particulier

$$f_{\theta}^{(k)}(0) = 0 \quad (1 \leq k < n), \quad f_{\theta}^{(n)}(0) = P(\cos \theta, \sin \theta).$$

En appliquant le théorème correspondant sur les fonctions d'une seule variable à $f_{\theta}^{(n)}(a)$ on obtient le théorème. On utilise pour 3. que l'on a soit $f_{\theta}^{(n)}(0) > 0$ pour tout $\theta \in I$, soit $f_{\theta}^{(n)}(0) < 0$ pour tout $\theta \in I$, car $f_{\theta}^{(\theta)}$, en tant que fonction de θ est continu et ne possède pas de zéros dans I . Pour 1. on observe que f possède un maximum local en a si et seulement si tous les $f_{\theta}(t)$ possède un maximum locale en $t = 0$, et analogue avec minimum au lieu de maximum. $\square$

Remarque. Un cas important est le cas $f'(a) = 0$ et que le hessien de f en a est régulière, i.e. que

$$\Delta := \frac{\partial^2 f}{\partial x^2}(a) \frac{\partial^2 f}{\partial y^2}(a) - \left(\frac{\partial^2 f}{\partial x \partial y}(a) \right)^2 \neq 0.$$

Ici

$$\begin{aligned} & P(\cos \theta, \sin \theta) \\ &= \frac{\partial^2 f}{\partial x^2}(a) \cos^2(\theta) + 2 \frac{\partial^2 f}{\partial x \partial y}(a) \cos(\theta) \sin(\theta) + \frac{\partial^2 f}{\partial y^2}(a) \sin^2(\theta). \end{aligned}$$

Donc, si $\Delta > 0$, alors $P(\cos \theta, \sin \theta) = 0$ ne possède pas de racines réelles, et en conséquence f possède un maximum locale en a si

$$\frac{\partial^2 f}{\partial x^2}(a) = P(1, 0) > 0,$$

et un minimum sinon (voir fig. 1 (a)). Si $\Delta < 0$, alors on a exactement deux racines $0 \leq \alpha_1 < \alpha_2 < \pi$, et donc $f_{\theta}(t)$ a un minimum locale en $t = 0$ pour θ (strictement) entre α_1 et α_2 , et un maximum locale pour les theta dans $[0, \alpha_1)$ ou (α_2, π) , ou vice versa. On appelle a un point de selle de f (voir fig. 2 (a)).

Chapitre 4

Equations différentielles du premier ordre

Soit f une fonction en deux variables, définie sur un sous-ensemble G de $\mathbb{R}^2$. On dit qu'une fonction dérivable $\phi : I \rightarrow \mathbb{R}$ définie sur un intervalle ouvert I est une solution de

$$y' = f(x, y),$$

si

1. le graphe $\{(x, \phi(x)) : x \in I\}$ est contenu dans G ,
2. et $\phi'(x) = f(x, \phi(x))$ pour tout $x \in I$.

L'équation $y' = f(x, y)$ est appelée équation différentielle (explicite) du premier ordre. Les mots “premier ordre” signifient que l'équation contient y' mais pas $y^{(n)}$ avec $n > 1$, et le mot “explicite” indique que l'équation est résolu par rapport à y' au contraire par exemple à une équation différentielle plus générale de la forme $f(y', y, x) = 0$.

L'équation $y' = f(x, y)$ possède une simple interprétation géométrique. A chaque point (x, y) du domaine G on attache la droite avec pente $f(x, y)$. On appelle le dessin résultant le champs de direction de l'équation $y' = f(x, y)$. Si ϕ est une solution, alors la courbe $c(x) = (x, \phi(x))$ convient à ce champs de direction dans le sens que la tangente $c'(x)$ donne justement dans la direction de la droite attaché au point $c(x)$. Réciproquement, toute courbe qui convient au champs de direction de $y' = f(x, y)$ et qui est de la forme $c(x) = (x, \phi(x))$ représente une solution de l'équation différentielle en question. La restriction aux équations explicites n'est pas trop sérieux, car — d'après le théorème sur les fonctions implicites de l'analyse supérieure — on peut toujours localement résoudre une équation $f(y', y, x) = 0$ pour y' sous certains hypothèses naturels et “génériques”.

Dans ce qui suit on utilise des lettres comme y ou $\tilde{y}$, y_1 etc. pour les solutions au lieu de ϕ , car c'est plus suggestif et naturel.

4.1 Existence et unicité des solutions

Définition. On dit que f satisfait à une condition de Lipschitz par rapport à y (brèf : CL) sur G si il existe une constante L telle que

$$|f(x, y) - f(x, y_1)| \leq L |y - y_1|$$

pour tout (x, y) et (x, y_1) dans G .

Théorème 4.1. Soit G un rectangle ouvert. On suppose que $f_y := \frac{\partial f}{\partial y}$ existe dans G , et que f_y est continue. Alors f satisfait à une CL sur tout sous-rectangle fermé de G .

Démonstration. Soit R un sous-rectangle fermé. La fonction continue $|f_y|$ prend son maximum sur R , donc est majoré par une constante L sur R . D'autre part, pour tout (x, y) et (x, y_1) il existe d'après les accroissements finis un ξ entre y et y_1 tel que

$$f(x, y) - f(x, y_1) = f_y(x, \xi)(y - y_1).$$

D'où le théorème. □

Théorème 4.2. (Unicité des solutions) On suppose que f satisfait à une CL sur G . Soient ϕ, ψ deux solutions de $y' = f(x, y)$ sur l'intervalle ouvert I . On suppose que $\phi(x_0) = \psi(x_0)$ pour un $x_0 \in I$. Alors $\psi \equiv \phi$.

Remarque. Autrement dit tout point de G se trouve sur le graphe d'au plus une solution de $y' = f(x, y)$.

Démonstration. On pose $g = \phi - \psi$. On a

$$|g'(x)| = |f(x, \phi(x)) - f(x, \psi(x))| \leq L |\phi(x) - \psi(x)| = L |g(x)|,$$

avec une constante L . En plus, $g(x_0) = 0$. Le théorème est maintenant une conséquence du lemme suivant. □

Lemme. Soit G une fonction dérivable sur un intervalle ouvert I . On suppose qu'il existe une constante L telle que $|g'(x)| \leq L |g(x)|$ pour tout $x \in I$, et un $x_0 \in I$ tel que $g(x_0) = 0$. Alors $g \equiv 0$.

Démonstration. On suppose que g n'est pas identiquement 0. Soit $\epsilon > 0$ tel que $L\epsilon < 1$. Car g est continue est possède au moins un zéro (par exemple x_0) il existe un x_1 tel que $g(x_1) \neq 0$, mais $g(x_1 - \epsilon) = 0$ (exercice). Soit $|g(a)|$ le maximum de $|g|$ sur $[x_1 - \epsilon, x_1]$. D'après les accroissements finis on trouve un ξ entre a et $x_1 - \epsilon$ tel que

$$|g(a)| = |g(a) - g(x_1 - \epsilon)| = |g'(\xi)(a - x_1 + \epsilon)| \leq L |g(a)| \epsilon < |g(a)|,$$

une contradiction □

Remarque. La condition CL n'est pas nécessaire pour assurer l'unicité comme dans le théorème (mais confortable à vérifier, voir le avant-dernier théorème). D'autrepart elle n'est pas superflue. Contre-exemple :

$$f(x, y) = \sqrt{|y|}, \quad G = \mathbb{R}^2$$

possède les deux solutions $y_1 \equiv 0$ et $y_2(x) = \frac{1}{4} \text{sign}(x)x^2$ qui satisfont toutes les deux à $y_1(0) = y_2(0) = 0$. (Bien-sûr, f ne satisfait à une CL.)

Théorème 4.3. (Existence) *Soit f continue sur un rectangle ouvert G , et soit $(x_0, y_0) \in G$. Alors il existe un intervalle ouvert I contenant x_0 et une solution ϕ de $y' = f(x, y)$ sur I avec $\phi(x_0) = y_0$.*

Remarque. Mais en générale, i.e. sans hypothèses supplémentaires, la solution ϕ n'est pas unique.

L'idée. On considère des suite de segment qui conviennent à peu près le champs de direction. Fixer $\epsilon > 0$. On commence en $P_0 := (x_0, y_0)$ et marche sur la droite attaché à P_0 jusqu'à ce qu'on arrive à un point de la forme $P_1 := (x_0 + \epsilon, \dots)$. Maintenant on marche sur la droite attaché à P_1 jusqu'à un point $P_2 := (x_0 + 2\epsilon, \dots)$, puis sur la droite attaché à ce point etc. On considère la suite de segment résultant comme graphe d'une fonction ϕ_ϵ . On peut montrer que ϕ_ϵ "converge" vers une solution de $y' = f(x, y)$ pour $\epsilon \rightarrow 0$. □

Dans les sections suivantes nous présentons quelques méthodes pour résoudre explicitement une équation différentielle donnés.

4.2 Variation de la constante

Soient a et b deux fonctions définies et continues dans un intervalle ouvert I . Nous considérons l'équation différentielles linéaire

$$y' = ay + b,$$

ou, conformément à la notation de l'introduction,

$$y' = f(x, y), \quad f(x, y) = a(x)y + b(x), \quad G = I \times \mathbb{R}.$$

Evidemment f est continue et satisfait à une condition de Lipschitz sur tout rectangle fermé dans $I \times \mathbb{R}$, donc on peut appliquer les théorèmes de l'unicité et de l'existence de la section précédente. Mais ici on peut analyser l'équation d'une manière plus élémentaire en la résolvant "explicitement". Nous considérons d'abord le cas $b \equiv 0$, i.e. l'équation homogène $y' = ay$.

Théorème 4.4. *Soit A une primitive de a . Alors, ils sont équivalentes :*

1. $y : I \rightarrow \mathbb{R}$ est une solution de $y' = ay$.
2. $y = C \exp(A)$ pour une constante C .

Remarque. En particulier, les solutions de $y' = ax$ forment un sous-espace vectoriel dans l'espace vectoriel des fonctions sur I de dimension 1.

Démonstration. Par calcul direct on vérifie que $y = C \exp(A)$ satisfait à $y' = ay$.

On suppose 1.. La fonction $g := y \exp(-A)$ satisfait à

$$g' = y' \exp(-A) - A' y \exp(-A) = 0.$$

Donc $g \equiv C$ avec une constante C , i.e. $y = C \exp(A)$. □

Exemple. Soit $a(x) = |x|$ et $I = \mathbb{R}$. On peut prendre $A(x) = \text{sign}(x) \frac{1}{2} x^2$, et donc

$$y = C \exp(\text{sign}(x) \frac{1}{2} x^2)$$

parcours avec C variant les solutions de

$$y' = |x|y.$$

Le cas général peut être traité par la variation de constante.

Théorème 4.5. *Soit A une primitive de a , et soit B une primitive de $b \exp(-A)$. Alors, ils sont équivalentes :*

1. $y : I \rightarrow \mathbb{R}$ est une solution de $y' = ay + b$.
2. $y = (B + K) \exp(A)$ avec une constante K .

Remarque. Brèf : Les solution de $y' = ay + b$ sont

$$y = \exp\left(\int a\right) \int (b \exp(-\int a)).$$

Démonstration (Variation de la constante). On verifie directement que 2. entraîne 1.

Reciproquement, soit y une solution. On écrit y sous la forme $y = C \exp(A)$ (i.e. on pose $C = y \exp(-A)$) avec une fonction C . Donc on “varie la constante C ” que l’on a vu dans le th^éorème précédent. Car y est une solution de $y' = ay + b$, alors

$$C' \exp(A) + CA' \exp(A) = y' = ay + b = aC \exp(A) + b,$$

donc $C' = b \exp(-A)$, i.e. $C = B + K$ pour une constante K , et c’est 2.. $\square$

Une conséquence évidente est

Corollaire 4.5.1. *Soit $(x_0, y_0) \in I \times \mathbb{R}$. Alors il existe une et une seule solution ϕ de $y' = ax + b$ sur I avec $\phi(x_0) = y_0$.*

Exemple. Nous considérons le problème de résoudre

$$y' = xy + \cos(x), \quad y(0) = 0.$$

La solution générale est

$$y = (B + K)e^A,$$

où

$$A(x) = \frac{1}{2}x^2, \quad B(x) = \int_0^x e^{-\frac{t^2}{2}} \cos(t) dt.$$

Avec $y(0) = 0$ on trouve donc

$$y = e^{\frac{1}{2}x^2} \int_0^x e^{-\frac{t^2}{2}} \cos(t) dt.$$

4.3 Séparation des variables

On suppose que l’on peut séparer les variables dans $f(x, y)$, i.e. que l’on peut écrire

$$f(x, y) = a(x)b(y)$$

avec des fonctions a et b définies dans des intervalles ouverts I_1 et I_2 respectivement. Nous supposons aussi que $b(y) \neq 0$ pour tout $y \in I_2$, et que a et b sont continues. Donc $y' = f(x, y)$ devient

$$\frac{y'}{b(y)} = a(x),$$

ou, en écrivant $dy = y' dx$, plus suggestif

$$\frac{dy}{b} = a dx.$$

Théorème 4.6. (Séparation de variables) Soit $P(y)$ une primitive de $1/b(y)$ et $A(x)$ une primitive de $a(x)$. Alors, P est monotone et possède donc une fonction inverse P^{-1} . Ils sont équivalent :

1. $\phi(x)$ est une solution de $y' = a(x)b(y)$.
2. Il existe une constante C telle que

$$\phi = P^{-1}(A + C).$$

Démonstration. On peut montrer que P est strictement monotone (car $1/b$ n'a aucune racine) et donc P^{-1} existe et est dérivable. Si ϕ est une solution, on a

$$P'(\phi) \phi' = \frac{\phi'}{b(\phi)} = a,$$

et donc

$$P(\phi) = A + C$$

avec une constante C convenable. C'est 2. Réciproquement, on peut vérifier par calcul direct que toute fonction de la forme 2. est une solution de 1. $\square$

Remarque. On peut résumer la méthode utiliser dans la démonstration symboliquement sous la forme :

$$\begin{aligned} y' &= a(x)b(y) \\ \triangleright \frac{dy}{b(y)} &= a(x)dx \triangleright P(y) := \int \frac{dy}{b(y)} = \int a(x)dx \\ &\triangleright y = P^{-1}\left(\int a(x)dx\right). \end{aligned}$$

Remarque. On peut aussi interpréter la méthode comme un changement de variable : On pose

$$\eta = P(y), \quad y = P^{-1}(\eta),$$

et on trouve

$$\eta' = P'(y) y' = a,$$

ce qu'on sait résoudre.

Exemple. On considère $y' = x^2/y$ pour $x \in \mathbb{R}$ et $y > 0$:

$$y' = \frac{x^2}{y}$$

$$\triangleright ydy = x^2 dx \triangleright (P(y) :=) \frac{1}{2}y^2 = \frac{13^3}{x} + C$$

$$\triangleright y = \sqrt{\frac{1}{3}x^3 + C}.$$

Exemple. Nous considérons $y' = b + y^2$ avec une constante $b > 0$ (et pour $G = \mathbb{R} \times \mathbb{R}$)

$$y' = b + y^2$$

$$\triangleright \frac{dy}{b + y^2} = dx \triangleright P(y) := \int \frac{dy}{b + y^2} = x + C \triangleright P(y) = \frac{1}{\sqrt{b}} \arctan\left(\frac{y}{\sqrt{b}}\right) = x + C$$

$$\triangleright y = \sqrt{b} \tan(\sqrt{b}x + C)$$

Exercice. Discuter et résoudre $y' = b + y^2$ dans le cas $b < 0$.

4.4 Equation différentielle de Bernoulli

On considère l'équation

$$y' = ay + by^\alpha.$$

Ici a et b sont des fonctions continues dans un intervalle ouvert I , et $f(x, y) = ay + by^\alpha$ est considéré comme fonction sur $G = I \times \mathbb{R}_{>0}$. Finalement α est un nombre réel, et nous rappelons la définition

$$y^\alpha = \exp(\alpha \log y).$$

Car on suppose $y > 0$ le nombre $\log y$ est définie. Pour $\alpha = 1$ on obtient l'équation linéaire de la section précédente.

On suppose $\alpha \neq 1$. Soit y une solution de l'équation de Bernoulli dans I . Car $y(x) > 0$ pour tout $x \in I$ on peut poser

$$y = z^{\frac{1}{1-\alpha}} \quad z = y^{1-\alpha}.$$

On vérifie que

$$y' = \frac{1}{1-\alpha} z^{\frac{\alpha}{1-\alpha}} z' = ay + by^\alpha = az^{\frac{1}{1-\alpha}} + bz^{\frac{\alpha}{1-\alpha}},$$

et donc

$$z' = (1 - \alpha)(az + b).$$

Cette équation peut être résolue par la méthode de la section précédente. Réciproquement on vérifie que toute solution de l'équation pour z donne une solution de l'équation pour y .

Exemple.

$$\begin{aligned} y' &= y + xy^2 \\ \triangleright z = \frac{1}{y}, \quad z' &= -z - x \triangleright z = \left(- \int_0^x e^t t dt + K \right) e^{-x} = (1 - xe^x + K) e^{-x} \\ \triangleright y &= \frac{1}{1 - x + Ke^{-x}} \end{aligned}$$

4.5 Changement de variable

Les méthodes utilisées dans les deux sections précédentes sont des cas spéciaux de la méthode “changement de variable” : Pour transformer l'équation

$$y' = f(x, y) \quad (G = I \times J)$$

à une équation différentielle (que l'on sait résoudre) on pose

$$x = \alpha(u), \quad y = \beta(z)$$

avec des fonctions dérivables $\alpha : I' \rightarrow I$ et $\beta : J' \rightarrow J$. On obtient formellement

$$dy = \beta'(z) dz = f(\alpha(u), \beta(z)) \alpha'(u) du,$$

et donc l'équation pour y devient

$$z' = g(u, z), \quad g(u, z) := f(\alpha(u), \beta(z)) \frac{\alpha'(u)}{\beta'(z)},$$

i.e. une équation différentielle en z avec g définie sur $I' \times J'$.

Théorème 4.7. *Soient α et β strictement monotone et $\alpha, \alpha^{-1}, \beta, \beta^{-1}$ continuellement dérivables (i.e. dérivables avec dérivées continues). Alors ϕ est une solution de $y' = f(x, y)$ sur I si et seulement si $\beta^{-1} \circ \phi \circ \alpha$ est une solution de $z' = g(u, z)$ sur I' .*

Démonstration. Exercice : par calcul direct. □

Exemple. On considère

$$y' = \frac{1}{x^2} + y^2$$

pour $x, y > 0$. Ici on trouve

$$\begin{aligned} y = 1/z, \quad dy = -\frac{dz}{z^2}, \quad x = u \triangleright z' = -\left(\frac{z^2}{u^2} + 1\right) \\ \triangleright z_1 := \frac{z(u)}{u}, \quad z' = z_1' u + z_1 \triangleright z_1' = -(z_1^2 + z_1 + 1)/u \\ \triangleright \frac{dz_1}{z_1^2 + z_1 + 1} = -\frac{du}{u} \triangleright P(z_1) := \int_0^{z_1} \frac{dt}{t^2 + t + 1} = -\log u + C \\ \triangleright y = \frac{1}{xP^{-1}(-\log x + C)} \end{aligned}$$

Exercice. Calculer explicitement $P(z_1)$ et vérifier que le y dans la dernière ligne est vraiment une solution de l'équation différentielle en question.

Remarque. L'équation précédente et dans l'exemple de la dernière section sont des cas spéciaux de l'équation différentielle de Riccati :

$$y' = a(x) + b(x)y + c(x)y^2.$$

Chapitre 5

Déterminants, valeurs et vecteurs propres

Dans ce qui suit nous désignons par K soit les nombres complexes, soit les nombres réels. Nous notons K^n l'espace vectoriel des vecteurs à colonnes de longueur n et avec éléments dans K . De plus, nous utilisons $M_n(K)$ pour l'ensemble des matrices carrées à n lignes et à n colonnes, et avec des éléments dans K . Nous rappelons que on peut calculer la somme et le produit de deux matrices dans $M_n(K)$, et que l'on peut multiplier les matrices avec un scalaire. De plus, une matrice $A \in M_n(K)$ est dite inversible, si et seulement si il existe une matrice B tel que $AB = E$. Ici E est la matrice unité. Un tel B , si il existe, est unique et noté A^{-1} . On a pour une matrice inversible aussi $A^{-1}A = 1$.

Finalement nous rappelons la forme matricielle d'un système d'équations linéaires en n inconnues x_i

$$\begin{aligned}a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\&\vdots \\a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n &= b_n.\end{aligned}$$

Si on considère les $a_{i,j}$ comme éléments d'une matrice A et les x_i et b_i comme éléments de vecteurs x et b respectivement, on peut écrire ce système plus court sous la forme

$$Ax = b.$$

5.1 La définition du déterminant

Si A est une matrice de $M_n(K)$ et si $1 \leq p, q \leq n$ sont des nombres naturels nous notons $A_{p,q}$ la matrice de $M_{n-1}(K)$ que l'on obtient en supprimant la ligne p et la colonne q de A (ici on suppose bien-sûr $n > 1$).

Définition. Le déterminant d'une matrice carrée A avec n lignes est défini par récurrence sur n comme suivant : Si $n = 1$ et $A = (a)$ avec un nombre a , alors $\det(A) = a$. Si $n > 1$, alors

$$\begin{aligned}\det(A) &= a_{1,1} \det(A_{1,1}) - a_{1,2} \det(A_{1,2}) + a_{1,3} \det(A_{1,3}) - \cdots \\ &\quad + (-1)^{n+1} a_{1,n} \det(A_{1,n}) \\ &= \sum_{p=1}^n (-1)^{p+1} a_{1,p} \det(A_{1,p}).\end{aligned}$$

Exemple. Pour les matrices de taille 2 et 3 on trouve ainsi

$$\begin{aligned}\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= ad - bc, \\ \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} &= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.\end{aligned}$$

Ici nous avons utilisé les bars verticales pour indiquer les déterminants.

Souvent nous considérons le déterminant comme application qui associe à n vecteurs à colonnes $a_1, \dots, a_n$ de longueur n le nombre

$$\det(a_1, \dots, a_n) := \det(A),$$

où A est la matrice avec colonnes $a_1, \dots, a_n$. Donc nous avons pour tout n les deux application

$$\det : M_n(K) \rightarrow K$$

et

$$\det : K^n \times \cdots \times K^n \text{ (} n\text{-fois) } \rightarrow K.$$

Les trois théorèmes suivants sont les règles de bases dans la théorie des déterminants. On pourrait démontrer sans difficulté les deux premiers par récurrence sur n . La preuve du troisième serait facile si on avait une certaine maîtrise dans les sommes multiples. Nous admettons leur démonstration.

Théorème 5.1. *Le déterminant est une application alternée, i.e. : si $A \in M_n(K)$ et si A' est la matrice qu'on obtient en échangeant dans A deux colonnes alors*

$$\det(A) = -\det(A').$$

Théorème 5.2. *Le déterminant est une application multilinéaire, i.e. : pour tous $a_1, \dots, a_{i-1}, a_i, a'_i, a_{i+1}, \dots, a_n \in K^n$ et $\lambda, \mu \in K$ on a*

$$\begin{aligned} & \det(a_1, \dots, a_{i-1}, \lambda a_i + \mu a'_i, a_{i+1}, \dots, a_n) \\ &= \lambda \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) + \mu \det(a_1, \dots, a_{i-1}, a'_i, a_{i+1}, \dots, a_n). \end{aligned}$$

Réciproquement on a

Théorème 5.3. (Unicité) *Soit $\Delta : M_n(K) \rightarrow K$ une application alternée est multilinéaire, alors*

$$\Delta(A) = \Delta(E) \cdot \det(A)$$

pour tout $A \in M_n(K)$ (où E note la matrice unité).

5.2 Règle de calcul pour les déterminants

Théorème 5.4. *Pour tout $A, B \in M_n(K)$ on a*

$$\det(AB) = \det(A) \det(B).$$

Démonstration. On considère l'application

$$\Delta : M_n(K) \rightarrow K$$

définie par $\Delta(B) = \det(AB)$. Considérée comme application sur des vecteurs on a la formule

$$\Delta(b_1, \dots, b_n) = \det(Ab_1, \dots, Ab_n),$$

où Ab_i est le produit usuel de la matrice A avec le vecteur b_i . Vu de cette formule il est clair que Δ est alternée est multilinéaire car $\det$ l'est. Par le théorème d'unicité on donc

$$\Delta(B) = \Delta(E) \det(B)$$

pour toute $B \in M_n(K)$. Or $\Delta(E) = \det(AE) = \det(A)$. □

Théorème 5.5. *Une matrice $A \in M_n(K)$ est inversible si et seulement si on a $\det(A) \neq 0$.*

Démonstration. Soit A inversible. Alors $E = AA^{-1}$, et donc

$$1 = \det(E) = \det(AA^{-1}) = \det(A) \det(A^{-1})$$

En particulier, on a $\det(A) \neq 0$.

Réciproquement, si $\det(A) \neq 0$, alors les colonnes a_i de A sont linéairements indépendants. En fait, si les colonnes sont linéairements dépendants, alors $\det(A) = 0$. C'est clair, si deux colonnes sont égales car dans ce cas la matrice A' que l'on obtient en échangeant ces deux colonnes est égal à A , et car on a d'autrepart $\det(A') = -\det(A)$. Dans le cas général il existe un i , disons $i = 1$ pour faciliter les notations, tel que $a_1 = \lambda_2 a_2 + \dots + \lambda_n a_n$ avec des $\lambda_i \in K$ convenables. Donc on a

$$\det(A) = \det\left(\sum_{i=2}^n \lambda_i a_i, a_2, \dots, a_n\right) = \sum_{i=2}^n \lambda_i \det(a_i, a_2, \dots, a_n).$$

Or chaque terme de la somme à droite est 0 d'après le cas déjà discuté.

Donc, si $\det(A) \neq 0$ les a_i forment une base de K^n , et pour tout $1 \leq h \leq n$ il existe donc un vecteur b_h tel que

$$Ab_h = e_h,$$

où e_h est le vecteur avec tout élément 0 sauf le h me élément qui est égal à 1. Autrement dit, si B est la matrice avec colonnes b_h , alors $AB = E$. Donc A est inversible. $\square$

Remarque. La preuve a montré en fait que

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

pour toute matrice inversible. Elle montre aussi qu'une matrice A est inversible si et seulement si ses colonnes sont linéairements indépendants.

Théorème 5.6. Soient $a_1, \dots, a_n \in K^n$. Alors $\det(a_1, \dots, a_n) = 0$ si et seulement si les $a_1, \dots, a_n$ sont linéairement dépendants.

Démonstration. Du théorème précédent en utilisant qu'une matrice est inversible si et seulement si ses colonnes sont linéairement indépendants (voir la remarque après le théorème). $\square$

Théorème 5.7. Soient $A \in M_m(K)$ et $B \in M_n(K)$. Alors

$$\det \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det(A) \det(B).$$

Remarque. En particulier on a

$$\det \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$$

pour $\lambda_i \in K$.

Démonstration. Par récurrence sur n . Exercice. $\square$

Théorème 5.8. *Pour tout $A \in M_n(K)$ on a*

$$\det(A^t) = \det(A).$$

Ici A^t est la matrice transposée de A .

Démonstration. On admet la démonstration qui est un peu pénible. Mais on peut vérifier comme exercice le théorème dans le cas $n = 2, 3$ avec les formules explicites données ci-dessus. $\square$

Comme corollaire du fait que $\det(A^t) = \det(A)$ nous avons

Corollaire 5.8.1. *Si A' provient d'une matrice A par échange de deux lignes, alors $\det(A') = -\det(A)$.*

Démonstration. La matrice transposée $(A')^t$ provient de A^t par échange de deux colonnes, et donc on a $\det((A')^t) = \det(A^t)$. D'autrepart le théorème précédent dit $\det((A')^t) = \det(A')$ et $\det(A^t) = \det(A)$. $\square$

En itérant ce corollaire on obtient

Corollaire 5.8.2. *Soient $l_1, \dots, l_n$ les lignes de la matrice A . Alors*

$$\det \begin{pmatrix} l_i \\ l_1 \\ \vdots \\ l_{i-1} \\ l_{i+1} \\ \vdots \\ l_n \end{pmatrix} = (-1)^{i+1} \det(A).$$

Théorème 5.9. (Règle de Lagrange) Pour tout $1 \leq i, j \leq n$ et $A \in M_n(K)$ on a

$$\det(A) = \sum_{p=1}^n (-1)^{p+j} a_{p,j} \det(A_{p,j}) = \sum_{p=1}^n (-1)^{i+p} a_{i,p} \det(A_{i,p}).$$

Remarque. La deuxième formule avec $i = 1$ correspond à notre définition du déterminant.

Démonstration. Formule 2 entraîne formule 1 : en fait, on a en appliquant formule 2 à A^t et en utilisant $\det((A^t)_{i,p}) = \det(A_{p,i})$

$$\det(A) = \det(A^t) = \sum_{p=1}^n (-1)^{i+p} a_{p,i} \det((A^t)_{i,p}) = \sum_{p=1}^n (-1)^{i+p} a_{p,i} \det(A_{p,i}),$$

qui est la formule 1 pour i .

Pour formule 2 soit $\tilde{A} = (\tilde{a}_{p,q})$ la matrice qui provient de A en mettant la i ème ligne de A comme première ligne. Alors d'après la définition du déterminant et avec $\tilde{a}_{1,p} = a_{i,p}$ et $\tilde{A}_{1,p} = A_{i,p}$ on a

$$\begin{aligned} \det(A) &= (-1)^{i+1} \det(\tilde{A}) = (-1)^{i+1} \sum_{p=1}^n (-1)^{1+p} \tilde{a}_{1,p} \det(\tilde{A}_{1,p}) \\ &= \sum_{p=1}^n (-1)^{i+p} a_{i,p} \det(A_{i,p}), \end{aligned}$$

i.e. formule 2. □

Théorème 5.10. Pour une matrice $A \in M_n(K)$ soit $A^* = (a_{i,j}^*) \in M_n(K)$ la matrice telle que

$$a_{i,j}^* = (-1)^{i+j} \det(A_{j,i}).$$

Alors

$$AA^* = A^*A = \det(A) E.$$

Remarque. En particulier, si A est inversible, i.e. si $\det(A) \neq 0$, alors

$$A^{-1} = \frac{1}{\det(A)} A^*.$$

Démonstration. Posons $AA^* = (c_{p,q})$, $A^* = (a_{p,q}^*)$, et soient l_i les lignes de A . En appliquant la formule de Lagrange pour la ligne q on trouve

$$c_{p,q} = \sum_{j=1}^n a_{p,i} a_{j,q}^* = \sum_{j=1}^n a_{p,i} (-1)^{j+q} \det(A_{q,j}) = \det \begin{pmatrix} l_1 \\ \vdots \\ l_{q-1} \\ l_p \\ l_{q+1} \\ \vdots \\ l_n \end{pmatrix}.$$

Si $p = q$ la matrice à droite est égale à A , et on obtient donc $\det(A)$, et si $p \neq q$, alors la matrice à droite a deux mêmes lignes et son déterminant est donc 0. La démonstration pour $A^*A = \det(A)E$ est pareille. $\square$

Exemple. Pour une matrice carrée à 2 lignes

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

on a

$$\det(A) = ad - bc$$

et, si elle est inversible, alors

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Exercice. Montrer par calcul direct que l'on a toujours pour toute matrice inversible ou non

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \det(A) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

5.3 Applications aux équations linéaires

Nous considérons un système d'équations linéaires de la forme

$$Ax = b,$$

où A est dans $M_n(K)$, où $b \in K^n$, et où le vecteur x est l'inconnu.

Théorème 5.11. *Le système d'équations linéaires $Ax = 0$ possède une solution non-triviale (i.e. une solution $x \neq 0$) si et seulement si $\det(A) = 0$.*

Démonstration. L'équation $Ax = 0$ a une solution non-triviale si et seulement si les colonnes de A sont linéairement dépendants. On a déjà vu que c'est équivalent à $\det(A) = 0$. $\square$

Si A est inversible, alors il existe une et une seule solution x . C'est $x = A^{-1}b$. Soit

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

On a une formule explicite pour le h me élément x_h de x pour tout $1 \leq h \leq n$.

Théorème 5.12. (Règle de Kramer) Soient A_h la matrice que l'on obtient de A en remplaçant la h me colonne de A par b . Alors

$$x_h = \frac{\det(A_h)}{\det(A)}.$$

Démonstration. Si a_i est la i me colonne de A , alors on peut écrire $Ax = b$ sous la forme

$$\sum_{i=1}^n a_i x_i = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b.$$

Donc

$$\begin{aligned} \det(A_h) &= \det(a_1, \dots, a_{h-1}, b, a_{h+1}, \dots, a_n) \\ &= \sum_{i=1}^n x_i \det(a_1, \dots, a_{h-1}, a_i, a_{h+1}, \dots, a_n) \\ &= x_h \det(A). \end{aligned}$$

Pour la 2me identité on a utilisé que $\det$ est multilinéaire, et pour la dernière qu'il est alternée. D'où la formule de Kramer. $\square$

5.4 Valeurs et vecteurs propres

Soit $A \in M_n(K)$. Un nombre $\lambda \in K$ est dite valeur propre de A si il existe un vecteur $x \neq 0$ dans K^n tel que

$$Ax = \lambda x.$$

Un tel vecteur est dit vecteur propre de A par rapport à λ . On note que l'ensemble

$$S(\lambda) := \{x \in K^n : Ax = \lambda x\}$$

est un sous-espace vectoriel de K^n .

Pour un nombre $t \in K$ on pose

$$\chi_A(t) := \det(tE - A).$$

De la définition du déterminant il est clair que $\chi_A(t)$ est un polynôme de degré n .

Exemple. Pour une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ on a

$$\chi_A(t) = (t - a)(t - d) - bc = t^2 - (a + d)t + \det(A).$$

Théorème 5.13. *Un λ est une valeur propre de A si et seulement si $\chi_A(\lambda) = 0$.*

Démonstration. Pour $\lambda \in K$ l'ensemble $S(\lambda)$ est l'ensemble des solutions du système d'équations linéaires $(\lambda E - A)x = 0$. Il existe une solution non-triviale si et seulement si $\chi(\lambda) = \det(\lambda E - A) = 0$. $\square$

Théorème 5.14. *Soient λ_j ($1 \leq j \leq p$) des valeurs propres de A , deux à deux différents, et pour tout $1 \leq j \leq p$ soit x_j un vecteur propre par rapport à λ_j . Alors les vecteurs $x_1, \dots, x_p$ sont linéairement indépendants.*

Démonstration. Nous traitons seulement le cas $p = 2$ dont la preuve contient déjà l'idée pour le cas générale. Supposons que $x_1 = \mu x_2$ pour un $\mu \in K$. Multiplions par A nous en obtenons

$$\lambda_1 x_1 = Ax_1 = \mu Ax_2 = \mu \lambda_2 x_2,$$

et donc $\lambda_1 \mu = \mu \lambda_2$, i.e. $\lambda_1 = \lambda_2$, qui est absurde. $\square$

On appelle A diagonalisable sur K si et seulement si il existe une matrice inversible $P \in M_n(K)$ telle que

$$A = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1}$$

avec des $\lambda_j \in K$. Soient $x_1, \dots, x_n$ les colonnes de P . Alors l'identité ci-dessus est équivalente à $AP = PD$, où D est la matrice diagonale, ou bien à

$$Ax_j = \lambda_j x_j \quad (1 \leq j \leq n).$$

On observe aussi que

$$\begin{aligned}\chi_A(t) &= \det(tE - PDP^{-1}) = \det(P(tE - D)P^{-1}) \\ &= \det(P) \det(tE - D) \det(P)^{-1} = \chi_D(t),\end{aligned}$$

où D est la matrice diagonale avec les λ_j sur le diagonale. Donc

$$\chi_A(t) = (t - \lambda_1) \cdots (t - \lambda_n).$$

En particulier, les λ_j et leurs multiplicités sont unique (à l'ordre près) : ces sont les racines de $\chi_A(t) = 0$, et donc les valeurs propres de A . Les colonnes x_j de la “matrice de passage” P sont des vecteurs propres par rapport aux λ_j .

Théorème 5.15. *La matrice $A \in M_n(K)$ est diagonalisable si et seulement si elle possède n vecteurs propres qui sont linéairement indépendants.*

Démonstration. Soit A diagonalisable. Alors, avec les notations précédentes les colonnes $x_1, \dots, x_n$ de P forment un système de n vecteurs propres linéairement indépendants. Réciproquement si on a n vecteurs propres linéairement indépendants x_j , alors avec la matrice P formée par les x_j comme colonnes on trouve que $AP = PD$ avec la matrice diagonale avec les valeurs propres des x_j sur le diagonale. $\square$

Corollaire 5.15.1. *Si $A \in M_n(K)$ possède n valeurs propres deux à deux différents, alors elle est diagonalisable.*

Exemple. Soit $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Le polynôme caractéristique de A est

$$\chi_A(t) = t^2 - 1 = (t - 1)(t + 1).$$

Donc la matrice est diagonalisable sur $\mathbb{R}$. Comme solution de

$$(A - E)x_1 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} x_1 = x_1, \quad (A + E)x_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} x_2 = x_2$$

on trouve

$$x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

respectivement. D'où

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \quad \left(\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \right).$$

Exercice. Montrer que la matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ n'est pas diagonalisable sur $\mathbb{R}$, mais que elle est bien diagonalisable sur $\mathbb{C}$.

Il n'est pas vrai que toute matrice A carrée est diagonalisable sur $\mathbb{C}$.

Exemple. Pour la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on a

$$\chi_A(P) = t^2.$$

Donc, si A était diagonalisable alors

$$A = P0_2P^{-1} = 0_2,$$

où 0_2 est la matrice avec tout élément égal à 0. Mais cette identité est absurde.

Il existe des critères simples pour décider si une matrice est diagonalisable ou non. Ici nous contentons au suivant que l'on peut facilement déduire de la discussion ci-dessus : Soient λ_j ($1 \leq j \leq p$) les valeurs propres différents de A , i.e. les racines différents de $\chi_A(t) = 0$. Pour chaque λ_j soit $m_j = \dim S(\lambda_j)$, i.e.

$$m_j = n - \text{rang}(A - \lambda_j E).$$

Alors A est diagonalisable si et seulement si

$$n = m_1 + \cdots + m_p.$$

5.5 Matrices triangonalisables

Nous avons vu une matrice triangulaire qui n'est pas diagonalisable sur les nombres complexes. Dans un sens c'est l'exemple le plus méchant. Nous appelons une matrice triangonalisable si et seulement si il existe une matrice inversible P telle que $P^{-1}AP = T$ est une matrice triangulaire $T \in M_n(K)$. Par une matrice triangulaire nous indiquons toujours une matrice avec tout élément au-dessous le diagonale égal à 0.

Théorème 5.16. *La matrice $A \in M_n(K)$ est triangonalisable si et seulement si toute racine de $\chi_A(t) = 0$ est contenue dans K .*

Corollaire 5.16.1. *Toute matrice $A \in M_n(\mathbb{C})$ est diagonalisable.*

Démonstration. Si $A = PTP^{-1}$ avec une matrice triangulaire $T \in M_n(K)$, alors

$$\chi_A(t) = \chi_T(t) = (t - \lambda_1) \cdots (t - \lambda_n),$$

où les λ_j sont les éléments sur la diagonale de T . Donc ces éléments sont les racines de $\chi_A(t) = 0$.

Pour montrer la direction réciproque nous procédons par récurrence sur n . Pour $n = 1$ le théorème est évident. Supposons que toute matrice B de taille $n - 1$ avec toutes racines de $\chi_B(t) = 0$ dans K est triangulisable. Alors, soit $\lambda \in K$ tel que $\chi_A(\lambda) = 0$. Il existe donc un $0 \neq x \in K^n$ tel que $Ax = \lambda x$. Complétons x à une base $c_1 = x, c_2, \dots, c_n$ de K^n , et posons P_1 la matrice avec colonnes égales aux c_i . On a

$$AP_1 = P_1 \begin{pmatrix} \lambda & b \\ 0_{n-1} & B \end{pmatrix}$$

avec un vecteur à ligne b , avec une matrice carrée de taille $n - 1$, et avec 0_{n-1} indiquant le vecteur 0 dans K^{n-1} . Or $\chi_A(t) = (t - \lambda)\chi_B(t)$. Donc toutes les racines de $\chi_B(t)$ sont de racines de $\chi_A(t) = 0$, et d'après l'hypothèse sur A dans K . D'après l'hypothèse de récurrence il existe une matrice inversible P_2 de taille $n - 1$ telle que

$$B = P_2 U P_2^{-1}$$

avec une matrice triangulaire U . En résumant on a

$$A = P_1 \begin{pmatrix} 1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} \lambda & bP_2 \\ 0 & U \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} P_1^{-1},$$

(exercice), et d'où que A est triangulisable. □

Exemple. La matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ n'est pas diagonalisable sur $\mathbb{R}$ (i.e. elle ne peut pas être écrite sous la forme PTP^{-1} avec des matrice réelles P et T et T diagonale) car $\chi_A(t) = (t - i)^2$. Par contre, elle est bien-sûr diagonalisable sur $\mathbb{C}$.

Exercice. "Triagonaliser" la matrice

$$\begin{pmatrix} 58 & 121 \\ -25 & -52 \end{pmatrix}$$

sur $\mathbb{R}$.

Chapitre 6

Systèmes linéaires d'équations différentielles à coefficients constants

Nous considérons une équations différentielle de la forme

$$y' = Ay,$$

où $A \in M_n(K)$, et où on cherche un vecteur

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

de fonctions définies sur $\mathbb{R}$ qui satisfait à l'équation donné. Aurement dit on cherche des fonctions $y_1, \dots, y_n$ telle que

$$\begin{aligned} y_1' &= a_{1,1}y_1 + a_{1,2}y_2 + \dots + a_{1,n}y_n \\ &\vdots \\ y_n' &= a_{n,1}y_1 + a_{n,2}y_2 + \dots + a_{n,n}y_n. \end{aligned}$$

Une application principale de nos résultats seront les équations différentielles ordinaires d'ordre n à coefficients constants :

$$\phi^{(n)} + a_{n-1}\phi^{(n-1)} + \dots + a_1\phi' + a_0\phi = 0.$$

Ici les a_j sont des constantes, et "l'inconnu" ϕ est une fonction numérique

ordinaires avec dérivée supérieures $\phi^{(k)}$. On pose

$$A = \begin{pmatrix} a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} \phi^{(n-1)} \\ \phi^{(n-2)} \\ \vdots \\ \phi' \\ \phi \end{pmatrix},$$

et puis l'équation différentielle pour ϕ devient équivalente au système d'équations différentielles $y' = Ay$.

Si les éléments $a_{p,q}$ de A sont réels on cherche naturellement des fonctions à valeurs réelles, i.e. des fonctions $y_j : \mathbb{R} \rightarrow \mathbb{R}$. Mais pour avoir une théorie claire il est nécessaire d'admettre aussi des matrices A à éléments complexes et comme solutions des fonctions à valeurs complexes

$$y_j : \mathbb{R} \rightarrow \mathbb{C}.$$

6.1 Fonctions à valeurs complexes

Il faut donc expliquer f' pour une fonction $f : \mathbb{R} \rightarrow \mathbb{C}$. Pour cela on écrit

$$f = f_{\text{re}} + if_{\text{im}},$$

où f_{re} et f_{im} sont les parties réelles et complexes de f respectivement, i.e. on a $f_{\text{re}}(t) = \text{Re}(f(t))$ et $f_{\text{im}}(t) = \text{Im}(f(t))$ pour tout $t \in \mathbb{R}$. On appelle f dérivable si f_{re} et f_{im} le sont, et on pose

$$f' := f'_{\text{re}} + if'_{\text{im}}$$

si f est dérivable.

Exercice. Vérifier les règles

$$(fg)' = f'g + fg', \quad \left(\frac{1}{f}\right)' = -\frac{f'}{f^2}$$

pour les fonctions à valeurs complexes.

Nous devons aussi prolonger la fonction exponentielle à une fonction définie sur l'ensemble des nombres complexes. Pour un $z \in \mathbb{C}$, disons $z = x + iy$ avec x et y réel, on pose

$$\exp(z) = \exp(x) (\cos(y) + i \sin(y)).$$

On utilise aussi e^z pour $\exp(z)$.

Théorème 6.1. *Pour tout $z, w \in \mathbb{C}$ on a*

$$\exp(z + w) = \exp(z) \exp(w).$$

Démonstration. Cette formule pour $z \in \mathbb{R}$ est bien connue, et pour $z \in i\mathbb{R}$ elle est équivalente aux lois d'addition pour les fonctions cosinus et sinus. Pour z générale elle est une combinaison de tous les trois. Exercice : Vérifier cette argument en détail. $\square$

Théorème 6.2. *Pour tout $\lambda \in \mathbb{C}$ la fonction $f(t) := \exp(t\lambda)$ (en une variable réelle t) est dérivable, et on a*

$$f'(t) = \lambda \exp(\lambda t).$$

Démonstration. Exercice. $\square$

6.2 La théorie générale

Soit $A \in M_n(K)$. Il est clair que l'ensemble des solutions

$$\mathcal{L}(A) = \{y : \mathbb{R} \rightarrow K \mid y' = Ay\}$$

est un sous-espace vectoriel du K -espace vectoriel des fonctions dérivables sur $\mathbb{R}$ à valeurs dans K . Nous allons montrer que $\mathcal{L}$ est de dimension finie, et comment on peut calculer une base.

L'équation $y' = Ay$ avec $A \in M_n(K)$ rassemble beaucoup à l'équation différentielle simple $f' = af$ avec une constante f ; en fait si $n = 1$ elle est de cette forme. La solution générale de cette équation simple est $f(t) = \exp(at)c$ avec une constante c . On peut penser à généraliser cette idée au cas d'une matrice A . Supposons il existe une matrice $F(t)$ carrée de taille n dont les éléments sont de fonctions définies et dérivables sur $\mathbb{R}$ à valeurs dans K , telle que

$$F'(t) = AF(t) = F(t)A, \quad F(0) = E.$$

On peut penser de $F(t)$ comme d'une fonction

$$F : \mathbb{R} \rightarrow M_n(K).$$

Nous allons voir dans la prochaine section qu'une telle fonction $F(t)$ existe.

Exemple. Soit

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad F(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Ici on a

$$F'(t) = \begin{pmatrix} -\sin(t) & -\cos(t) \\ \cos(t) & -\sin(t) \end{pmatrix},$$

et on vérifie que l'on a en fait $F' = AF = FA$ et $F(0) = E$.

Nous montrons

Théorème 6.3. *F est unique. On a $F(t+u) = F(t)F(u)$ pour tout $t, u \in \mathbb{R}$. En particulier $F(t)$ est inversible pour tout t et $F(t)^{-1} = F(-t)$.*

Démonstration. On a

$$\frac{d}{dt} F(t)F(-t) = F'(t)F(-t) - F(t)F'(-t) = AF(t)F(-t) - F(t)AF(-t) = 0.$$

Donc $F(t)F(-t)$ est constante, et, en prenant $t = 0$, en fait égal à E .

Soit $G(t)$ tel que $G'(t) = AG(t) = G(t)A$. En dérivant on vérifie par le même raisonnement comme avant que $F(-t)G(t) = F(t)^{-1}G(t)$ et constante, et égal à E , i.e. $G = F$.

La fonction $G(t) := F(t+u)F(u)^{-1}$ (pour u fixé) satisfait aussi à $G' = AF = FA$ et $G(0) = E$, et donc est égale à $F(t)$. $\square$

Nous utilisons désormais les notations

$$\exp(tA), \quad e^{tA} \quad e_A(t)$$

au lieu de $F(t)$. Dons on peut écrire par exemple très suggestivement

$$e^{(t+u)A} = e^{tA}e^{uA}.$$

Théorème 6.4. *Pour tout $c \in K^n$ la fonction $y(t) = \exp(tA)c$ est une solution de $y' = Ay$. L'application*

$$K^n \rightarrow \mathcal{L}(A), \quad c \mapsto y = e_A c$$

est un isomorphisme de K -espaces vectoriels.

Remarque. L'application inverse est $y \mapsto y(0)$.

Démonstration. Par calcul direct on vérifie que $y = e_A c$ est une solution. Réciproquement, si $y' = Ay$, alors

$$\frac{d}{dt} \exp(-tA)y(t) = -A \exp(-tA)y(t) + \exp(-tA)y'(t) = 0,$$

donc $e_A^{-1}y$ est une constante $c \in K^n$.

Il reste à montrer que l'application est injective. Soit c dans le noyau de l'application, i.e. $\exp(tA)c = 0$ pour tout t . Car $\exp(0 \cdot A) = E$ on en déduit $c = 0$. $\square$

Corollaire 6.4.1. $\dim_K \mathcal{L}(A) = n$.

Corollaire 6.4.2. Soit $t_0 \in \mathbb{R}$. Alors des solutions $y_1, \dots, y_k$ de $y' = Ay$ sont linéairement indépendents si et seulement si $y_1(t_0), \dots, y_k(t_0)$ le sont.

Démonstration. En fait, les y_j sont linéairement indépendents si et seulement les vecteurs $y_j(0)$ le sont, qui est équivalent à dire que les vecteurs $\exp(t_0 A)y_j(0) = y_j(t_0)$ sont indépendents. $\square$

Les colonnes de $\exp(tA)$ sont n solutions linéairement indépendents de $y' = Ay$. En fait, si e_j est la base canonique de K^n , alors les fonctions $\exp(tA)e_j$ (i.e. les colonnes de $\exp(tA)$) forment une base de $\mathcal{L}(A)$.

Supposons réciproquement que $y_1, \dots, y_n$ sont des solutions linéairement indépendents de $y' = Ay$. On appelle les y_j “un système fondamental de l'équation $y' = Ay$ ”, et on appelle la matrice

$$W = W(y_1, \dots, y_n) := (y_1, \dots, y_n)$$

“le wronskien” du système fondamental $y_1, \dots, y_n$ de $y' = Ay$. D'après le théorème on a

$$W(t) = \exp(tA) M$$

pour une matrice constante M . Considérons $t = 0$ et remarquons que $W(t)$ est inversible pour tout t , on obtient que $M = W(0)$ est inversible. Donc on a

$$\exp(tA) = W(t)W^{-1}(0),$$

ce qui dit que le calcul de $\exp(tA)$ équivalent à déterminer n solutions linéairement indépendents. Nous allons expliquer dans les sections suivantes comment on peut trouver de telles solutions.

On peut formuler le théorème principal ci-dessus en disant que l'application

$$K^n \rightarrow \mathcal{L}(A), \quad c \mapsto Wc$$

est un isomorphisme d'espace vectoriel (sur K). En pratique on calcule plutôt avec un wronskien quelconque et pas necessairement avec le wronskien spécial $\exp(tA)$. C'est pour ça que nous formulons les deux théorèmes suivants en utilisant un wronskien général.

Théorème 6.5. Soit $y_0 \in K^n$ et $t_0 \in \mathbb{R}$. Alors il existe une et une solution de $y' = Ay$ telle que $y(t_0) = y_0$. Si W est un wronskien de l'équation $y' = Ay$, alors $y(t) = W(t)W(t_0)^{-1}y_0$.

Démonstration. Toute solution est de la forme $y = Wc$ avec un unique $c \in K^n$. On a $y(t_0) = y_0$ si et seulement si $W(t_0)c = y_0$, i.e. $c = W(t_0)^{-1}y_0$. $\square$

Théorème 6.6. Soit $b : \mathbb{R} \mapsto K^n$ continue. Alors, pour toute valeur initiale $t_0 \in \mathbb{R}$ et $y_0 \in K^n$, il existe une et une seule $y : \mathbb{R} \rightarrow K^n$ dérivable tel que

$$y' = Ay + b, \quad y(t_0) = y_0.$$

Si W est le wronskien d'un système fondamental de $y' = Ay$, alors on a

$$y(t) = y_0 + W(t) \int_{t_0}^t W(t)^{-1} b(t) dt.$$

Démonstration. On procède par variation de constante. Si y est une solution, alors on pose $y = Wv$ avec une fonction $v = v(t)$ (i.e. on pose $v := W^{-1}y$). On vérifie que $Wv' = b$, et donc

$$v(t) = c + \int_{t_0}^t W(s)^{-1} b(s) ds$$

avec un vecteur constant c convenable. Pour $t = t_0$ on trouve $W(t_0)c = y_0$, et d'où la formule du théorème.

Réciproquement il est clair que la fonction donnée dans le théorème est une solution comme cherchée. $\square$

6.3 Calcul d'un système fondamental

Dans cette section nous allons montrer que la fonction matricielle $\exp(tA)$ existe en fait et comment on peut la calculer explicitement. Nous considérons quand-même seulement deux cas : le cas où A est diagonalisable et, en toute généralité, le cas $n = 2$. Le cas général n'est pas plus compliqué si on connaît un peu de l'algèbre linéaire avancé (plus précisément : forme normale de Jordan). Un traitement sans ces outils ferait la théorie apparaître obscure. Heureusement $n = 2$ est A diagonalisable convient à la pratique.

A diagonalisable

Soit $A \in M_n(\mathbb{C})$. Rappelons que A est diagonalisable si le polynôme caractéristique

$$\chi_A(t) = \det(tE - A)$$

possède exactement n racines deux à deux différentes. D'après l'hypothèse il existe une matrice P complexe inversible telle que

$$A = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1}$$

avec des nombres complexes (les racines de $\chi_A(t) = 0$) convenables. Appelons D la matrice diagonale au milieu.

Théorème 6.7. *Pour t réel soit*

$$F(t) := P \begin{pmatrix} \exp(t\lambda_1) & & \\ & \ddots & \\ & & \exp(t\lambda_n) \end{pmatrix} P^{-1}.$$

Alors $F(0) = E$ et $F'(t) = AF(t) = F(t)A$. En plus, si A est à éléments réels, alors $F(t)$, pour tout t , l'est aussi.

Remarque. Donc la fonction matricielle que nous avons notée $\exp(tA)$ dans la section précédente existe et est égale au $F(t)$ du théorème.

Démonstration. Soit $\Delta(t)$ la matrice au milieu avec les $\exp(t\lambda_j)$ sur la diagonale. On observe que $\Delta'(t) = D\Delta(t) = \Delta(t)D$. D'où

$$F' = P\Delta'P^{-1} = PD\Delta P^{-1} = PDP^{-1}P\Delta P^{-1} = AF.$$

Analogue on vérifie $\Delta' = FA$. Il est clair que $\Delta(0) = E$.

Nous admettons la démonstration que $F(t)$ est réel pour réel A . □

Exercice. Calculer $\exp\left(t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right)$.

Les colonnes de P dans le formule pour $\exp(tA)$ sont des vecteurs propres de A . Il est très facile à vérifier directement :

Théorème 6.8. *Soit y_0 un vecteur propre de A avec valeur propre λ . Alors $y(t) := y_0 \exp(\lambda t)$ est une solution de $y' = Ay$.*

Démonstration. Par calcul direct on vérifie

$$Ay(t) = \exp(\lambda t)Ay_0 = \exp(\lambda t)\lambda y_0 = y'(t).$$

□

Dans des calculs pratiques on ne s'intéresse pas vraiment pour $\exp(tA)$, mais plutôt pour le wronskien

$$W(t) = P \begin{pmatrix} \exp(t\lambda_1) & & \\ & \ddots & \\ & & \exp(t\lambda_n) \end{pmatrix}$$

(pour éviter le calcul de P^{-1}). Pour le calculer, il faut donc déterminer les racines de $\chi_A(t) = 0$ et, disons qu'ils sont deux à deux différentes, calculer pour chaque racine un vecteur propre pour en obtenir la matrice P formée avec ces vecteurs propres comme colonnes. Très souvent les éléments de ce wronskien sont des fonctions à valeurs complexes, même si A (et donc aussi $\exp(tA)$) est réel. Néanmoins $W(t)$ est utile : si on considère par exemple pour A et y_0 réels la solution de $y' = Ay$ et $y(t_0) = y_0$ à valeurs réelles, alors il est banale mais utile à remarquer qu'elle est identique à la solution à valeurs complexes (par unicité de cette solution à valeurs complexes). Donc comme règle générale on calcule toujours avec des y à valeurs complexes en sachant par la théorie que l'on va trouver à la fin automatiquement les solutions à valeurs réelles si ils existent.

Dans le cas où A n'est pas diagonalisable on ne peut pas trouver n vecteurs propres qui sont linéairement indépendents. Ici on utilise :

Théorème 6.9. *Soit λ une racine de $\chi_A(t)$ avec multiplicité m (i.e. m est le nombre naturel le plus grand tel que $\chi_A(t) = (t - \lambda)^m p(t)$ avec un polynôme convenable $p(t)$). Alors il existe n solutions linéairement indépendents de $y' = Ay$ qui sont de la forme*

$$y(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{pmatrix} e^{\lambda t},$$

où les $p_j(t)$ sont des polynômes en t de degré $\leq m - 1$.

Nous admettons la démonstration et nous n'entrons pas dans les détails du calcul des $p_j(t)$ etc., mais nous allons vérifier et appliquer ce théorème implicitement dans le cas $n = 2$ par calcul direct.

Le cas A non-diagonalisable pour $n = 2$

Soit

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

On a

$$\chi_A(t) = t^2 - (a + d)t + \det(A).$$

Le cas où A est diagonalisable était traité dans la section ci-dessus. Donc supposons que A ne soit pas diagonalisable, i.e. que $\chi_A(t) = 0$ a une seule

racine (double) est que A n'est pas diagonale. On remarque que avoir une seule racine est équivalent à dire que

$$(a + d)^2 - 4 \det(A) = 0$$

(car à gauche c'est le discriminant du polynôme quadratique $\chi_A(t)$). La racine est

$$\lambda = \frac{a + d}{2},$$

et donc

$$\chi_A(t) = (t - \frac{a + d}{2})^2.$$

Soit c_1 un vecteur propre de A avec valeur propre λ , soit c_2 un vecteur quelconque, mais tel que c_1 et c_2 forment une base de $\mathbb{C}^2$, et soit finalement P la matrice avec colonnes c_1 et c_2 . On vérifie que

$$A = P \begin{pmatrix} \lambda & x \\ 0 & \lambda \end{pmatrix} P^{-1}$$

avec un x convenable (En fait, on peut toujours modifier P tel que l'on a $x = 1$ (exercice), mais c'est sans importance ici). On a

$$\exp(tA) = \exp(t\lambda) P \begin{pmatrix} 1 & tx \\ 0 & 1 \end{pmatrix} P^{-1},$$

i.e. si on appelle la fonction matricielle à droite $F(t)$, alors :

Théorème 6.10. *On a $F' = AF = FA$ et $F(0) = E$. Si A est réel, alors $F(t)$ l'est aussi pour tout t .*

Démonstration. Par calcul direct; par exemple :

$$\begin{aligned} F'(t) &= \exp(t\lambda) P \begin{pmatrix} \lambda & (\lambda t + 1)x \\ 0 & \lambda \end{pmatrix} P^{-1} \\ &= \exp(t\lambda) P \begin{pmatrix} \lambda & x \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & tx \\ 0 & 1 \end{pmatrix} P^{-1} = AF(t). \end{aligned}$$

□

Remarques

Il existe une formule explicite pour $\exp(tA)$ pour A arbitraire. Elle n'est pas tellement importante pour la pratique mais très utile dans des considérations théoriques. Cette formule est

$$\exp(tA) = E + \frac{t}{1!}A + \frac{t^2}{2!}A^2 + \frac{t^3}{3!}A^3 + \dots = \lim_{N \rightarrow \infty} \sum_{n=0}^N \frac{t^n}{n!}A^n.$$

La limite de matrices est à prendre élément par élément.

6.4 Equations différentielles linéaires de 2me ordre à coefficients constants

Nous considérons l'équation différentielle

$$\phi'' + b\phi' + a\phi = 0,$$

avec des constantes b et c . Cette équation est équivalente à

$$\frac{d}{dt} \begin{pmatrix} \phi' \\ \phi \end{pmatrix} = \begin{pmatrix} -b & -c \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \phi' \\ \phi \end{pmatrix}.$$

Appelons A la matrice à droite. On vérifie

$$\chi_A(t) = t^2 + bt + c.$$

Soient λ et μ les deux racines du polynôme caractéristique, i.e. soit

$$t^2 + bt + c = (t - \lambda)(t - \mu).$$

Deux racines différentes : $\lambda \neq \mu$

On vérifie par calcul direct que

$$f_\lambda(t) = \exp(t\lambda), \quad f_\mu(t) = \exp(t\mu)$$

sont des solutions de $\phi'' + b\phi' + a\phi = 0$, et qu'ils sont linéairements indépendents. Leur wronskien associé est

$$W(t) = \begin{pmatrix} f'_\lambda(t) & f'_\mu(t) \\ f_\lambda(t) & f_\mu(t) \end{pmatrix} = \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \exp(\lambda t) & 0 \\ 0 & \exp(\mu t) \end{pmatrix}$$

En particulier, pour chercher la solution générale de

$$\phi'' + b\phi' + a\phi = g(t),$$

où $g(t)$ est une fonction continue, on applique la théorie générale à

$$\frac{d}{dt} \begin{pmatrix} \phi' \\ \phi \end{pmatrix} = A \begin{pmatrix} \phi' \\ \phi \end{pmatrix} + \begin{pmatrix} g(t) \\ 0 \end{pmatrix},$$

et on trouve

$$\begin{pmatrix} \phi' \\ \phi \end{pmatrix} = y_0 + W(t) \int_{t_0}^t W(s)^{-1} \begin{pmatrix} g(s) \\ 0 \end{pmatrix} ds.$$

Exercice. Ecrire cette identité élément par élément.

Racine double : $\lambda = \mu$

Ici $\lambda = -\frac{b}{2}$ est la seule racine du polynôme caractéristique $\chi_A(t) = t^2 + bt + c$.
On vérifie que

$$f_1(t) = e^{\lambda t}, \quad f_2(t) = te^{\lambda t}$$

sont deux solutions linéairement indépendants de $\phi'' + b\phi' + c\phi = 0$: par exemple

$$f_2''(t) + bf_2'(t) + cf_2(t) = e^{\lambda t}2\lambda + \lambda^2 t + b(1 + \lambda t) + c = 0$$

(utiliser $\lambda = -\frac{b}{2}$). On peut facilement calculer le wronskien associé à ces solutions et puis appliquer la théorie générale pour résoudre tout problème posé.