

Mathématiques Générales

Jean-Claude Hausmann

*Université de Genève
Edition 2002*

Table des matières

| | |
|---|-----------|
| Avant-propos : Mathématiques et Sciences | 1 |
| 1 Simulations mathématiques : systèmes dynamiques | 3 |
| 1.1 Systèmes dynamiques discrets | 4 |
| 1.2 La croissance exponentielle | 6 |
| 1.3 États stationnaires - Cycles - Stabilité | 8 |
| 1.4 La croissance logistique | 10 |
| 1.5 Déterminisme, prévisibilité, hasard | 13 |
| 2 Algèbre Linéaire | 14 |
| 2.1 Structure d'espace vectoriel sur \mathbb{R}^n | 15 |
| 2.1.1 Définitions | 15 |
| 2.1.2 Produit scalaire – Norme – Angles | 15 |
| 2.1.3 Indépendance linéaire - Bases - Dimension d'un espace vectoriel | 17 |
| 2.2 Matrices | 19 |
| 2.2.1 Définitions | 19 |
| 2.2.2 Déterminants | 20 |
| 2.2.3 Interprétation géométrique du déterminant | 25 |
| 2.2.4 Systèmes d'équations linéaires | 26 |
| 2.3 Applications linéaires | 29 |
| 2.4 Changements de bases | 31 |
| 2.5 Valeurs propres - vecteurs propres | 32 |
| 2.5.1 Définitions | 32 |
| 2.5.2 Diagonalisation | 34 |
| 2.5.3 Matrices symétriques : | 37 |
| 2.6 Systèmes dynamiques linéaires | 37 |
| 2.6.1 Stabilité de l'origine | 37 |
| 2.6.2 Populations avec pyramide d'âges | 39 |
| 2.6.3 Le système de Fibonacci | 41 |
| 3 Calcul différentiel | 44 |
| 3.1 Applications différentiables à une variable | 44 |
| 3.1.1 Différentiabilité | 45 |
| 3.1.2 Comportement de f au voisinage d'un point | 46 |

| | | |
|----------|---|-----------|
| 3.1.3 | Formule de Taylor à une variable | 49 |
| 3.2 | Applications différentiables à plusieurs variables | 51 |
| 3.2.1 | Dérivées partielles - Gradient | 51 |
| 3.2.2 | Déférentiablité | 53 |
| 3.2.3 | Dérivée de fonctions composées | 55 |
| 3.2.4 | Application : calcul d'erreurs | 56 |
| 3.2.5 | Formule de Taylor à plusieurs variables | 56 |
| 3.2.6 | Gradient et surfaces de niveau | 58 |
| 3.3 | Problèmes d'extrema | 59 |
| 3.3.1 | Fonctions d'une variable | 59 |
| 3.3.2 | Fonctions de plusieurs variables | 60 |
| 3.3.3 | Méthode des moindres carrés - Régression linéaire | 62 |
| 3.4 | Stabilité des états stationnaires d'un système dynamique | 64 |
| 4 | Calcul intégral | 65 |
| 4.1 | Intégrales et primitives | 65 |
| 4.2 | Calcul de primitives et d'intégrales | 67 |
| 4.3 | Intégrales de surface et de volume | 71 |
| 4.4 | Quelques applications | 75 |
| 5 | Systèmes dynamiques continus – Équations différentielles | 78 |
| 5.1 | Champs de vecteurs - trajectoires. | 78 |
| 5.2 | Trajectoires et équations différentielles | 80 |
| 5.3 | États stationnaires - Stabilité | 80 |
| 5.4 | Quelques techniques de résolution d'équations différentielles | 82 |
| 5.5 | Quelques systèmes dynamiques continus | 87 |
| 5.5.1 | Croissance exponentielle | 87 |
| 5.5.2 | Croissance logistique | 88 |
| 5.5.3 | Proies–Prédateurs | 89 |
| A | Notes techniques et compléments | 91 |
| A.1 | Notations | 91 |
| A.2 | Entiers | 92 |
| A.3 | Nombres rationnels | 93 |
| A.4 | Nombres réels | 94 |
| A.5 | Entre les rationnels et les réels | 96 |
| A.6 | Nombres complexes | 97 |
| A.7 | Polynômes | 99 |
| A.8 | Espaces vectoriels – Produits scalaires | 101 |
| A.9 | Composantes principales | 102 |
| | Bibliographie | 104 |
| | Index | 105 |

Mathématiques et Sciences

Activité fondamentale de l'esprit humain, les mathématiques se sont développées sans discontinuité depuis l'Antiquité. Leur construction se poursuit chaque jour, donnant lieu à des milliers d'articles par an. Les plus grands esprits y ont laissé leur empreinte (Euclide, Newton, Gauss, Poincaré, etc).

Les concepts mathématiques, même les plus usités (nombres, cercles, droites), n'ont pas d'existence ailleurs que dans notre imagination. Le plus souvent, ils ont été élaborés par pure curiosité intellectuelle et pas en vue d'une application. On peut ainsi s'étonner que de telles abstractions puissent être utiles, notamment pour les sciences naturelles et la technologie. Nous savons que c'est pourtant le cas : depuis Galilée et Kepler, une grande partie de l'activité scientifique consiste à décrire, *en termes mathématiques*, des phénomènes naturels, convenablement isolés et idéalisés. L'aspect opératoire et déductif des mathématiques est alors mis en jeu pour découvrir et démontrer de nouvelles "lois de la nature", que l'on peut ensuite soumettre à l'expérience. Les concepts mathématiques sont ainsi des *outils de pensée, de découverte et de compréhension* ("mathema" = "comprendre", en grec) au service de toutes les sciences. L'usage de l'ordinateur et la numérisation de l'information accroissent sans cesse l'usage possible et le besoin de modèles mathématiques.

Concrètement, voici quelques outils mathématiques utiles dans la pratique scientifique et dont on parlera dans ce cours :

- des *nombres*, pour exprimer les grandeurs que l'on mesure et effectuer des calculs. Les principaux systèmes numériques (nombres entiers, rationnels, réels, complexes) seront discutés dans les § A.2 à A.6, notamment du point de vue de leur usage correct sur ordinateur.
- des *concepts géométriques* pour maîtriser intellectuellement l'espace (calcul vectoriel, algèbre linéaire, voir ch. 2).
- le *calcul infinitésimal*, outil intellectuel d'une valeur inestimable sans lequel la science et la technologie moderne n'existeraient pas (ch. 3-5).

Il se décompose en deux aspects :

- le *calcul différentiel* (ch. 3), notamment pour les problèmes d’optimisation.
- le *calcul intégral* (ch. 4), pour le calcul d’aires, volume, débits, valeurs moyennes, etc.
- des modèles pour décrire et prédire l’évolution des phénomènes naturels (*systèmes dynamiques, équations différentielles*, ch. 1 et 5). Grâce à l’essor de l’ordinateur, les systèmes dynamiques constituent le courant principal des mathématiques appliquées actuelles.
- des raisonnements efficaces, souples et rigoureux et dont les conclusions sont fiables.
- les *probabilités et statistiques* (semestre d’été).

En résumé, les mathématiques jouent un rôle fondamental dans la connaissance scientifique et son développement. Ce sont de puissants outils intellectuels qui permettent de formuler les lois scientifiques de manière précise, maniable et subtile. Elles débouchent sur un usage efficient de l’ordinateur, notamment dans l’organisation d’expériences “virtuelles”. Enfin, les mathématiques sont un cadre de raisonnement rigoureux et opératoire tel qu’il est partout nécessaire en science.

Après cette introduction, l’étudiant en sciences devrait être convaincu du bien fondé de s’intéresser au plus possible de mathématiques. Il s’agit d’un investissement important qui lui donnera un avantage substantiel dans la suite de ses études et de sa carrière.

Chapitre 1

Simulations mathématiques : systèmes dynamiques

Comment s'est formée la terre ? Comment va évoluer la population mondiale ? Que va-t-il se passer dans telle réaction chimique ? Quel temps fera-t-il dans quelques jours ?

L'un des buts principaux de la science est de décrire l'évolution des phénomènes naturels. L'essor de l'ordinateur a fait exploser cet aspect de l'activité scientifique, permettant de s'attaquer à toutes sortes de problèmes et de tenir compte d'une multitude de paramètres (plusieurs dizaines de millions en météorologie). Les simulations d'évolutions "virtuelles" ouvrent chaque jour de nouveaux champs d'investigation, remplaçant des expériences impossibles, dangereuses ou trop coûteuses (écologie, processus industriels, etc.).

Mais, pour mettre en scène une évolution virtuelle sur un ordinateur, il faut d'abord en avoir un *modèle mathématique*. L'étude mathématique des modèles d'évolution est la théorie des *systèmes dynamiques* (popularisée sous le nom de *théorie du chaos*). C'est aujourd'hui la partie principale des mathématiques appliquées et elle influence notre vision de ce qu'est un modèle scientifique. C'est pourquoi nous avons trouvé intéressant de prendre, dans ce cours, les systèmes dynamiques comme *fil conducteur* et comme prétexte pour apprendre les diverses techniques utiles dans toutes les sciences.

On distingue deux types de systèmes dynamiques¹ :

1. les systèmes *discrets*, dans lesquels le temps varie par saut d'une unité.
Ce sont ceux qui servent aux simulations sur ordinateurs. On en verra de très simples dans ce chapitre et d'autres dans les chapitres 2 et 3.

¹ Il s'agit ici de systèmes *déterministiques*, où une cause détermine un effet ; nous n'abor-derons pas les systèmes *probabilistiques*, où le hasard entre en jeu (voir cependant p. 13).

2. les systèmes *continus*, dans lesquels le temps varie continûment. Plus anciens et importants du point de vue théorique, ils relèvent de la théorie des équations différentielles. Ils seront traités au chapitre 5.

1.1 Systèmes dynamiques discrets

Par définition, un **système dynamique discret** se compose de :

1. un ensemble E (l'ensemble des **états** possibles du système)
2. une application $f : E \rightarrow E$, de E dans lui-même (la **loi d'évolution** du système).

Le temps varie *discrètement*, c'est-à-dire *par saut d'une unité*. Notons $s(t)$ l'état du système au temps t . L'évolution du système, à partir d'un l'état initial $s_0 = s(0)$, est donc décrite par une suite d'éléments de E :

$$s_0 = s(0), s(1), s(2), s(3) \dots \quad (1.1)$$

On appelle cette suite la **trajectoire** (ou l'*orbite*) de l'état s_0 . On suppose que cette trajectoire est *déterminée par l'application f* , grâce à l'**équation d'évolution**

$$s(t+1) = f(s(t)) \quad (1.2)$$

Autrement dit : si le système est dans l'état s , il sera, après une unité de temps, dans l'état $f(s)$.

Au temps $t+2$, l'état du système sera $s(t+2) = f(f(s(t))) = f^2(s(t))$. Pour connaître l'état $s(t+n)$, après n étapes, il suffit ainsi d'**itérer** n fois f , c'est-à-dire de composer n fois l'application f avec elle même :

$$s(t+n) = f^n(s(t)) := \underbrace{f \circ f \circ \dots \circ f}_{n \text{ fois}}(s(t)) \quad (n^{\text{ième itération de }} f).$$

(Le symbole $f^{\circ n}$ peut être utilisé s'il y a risque de confusion avec une puissance.) La trajectoire (1.1) de s_0 s'écrit ainsi

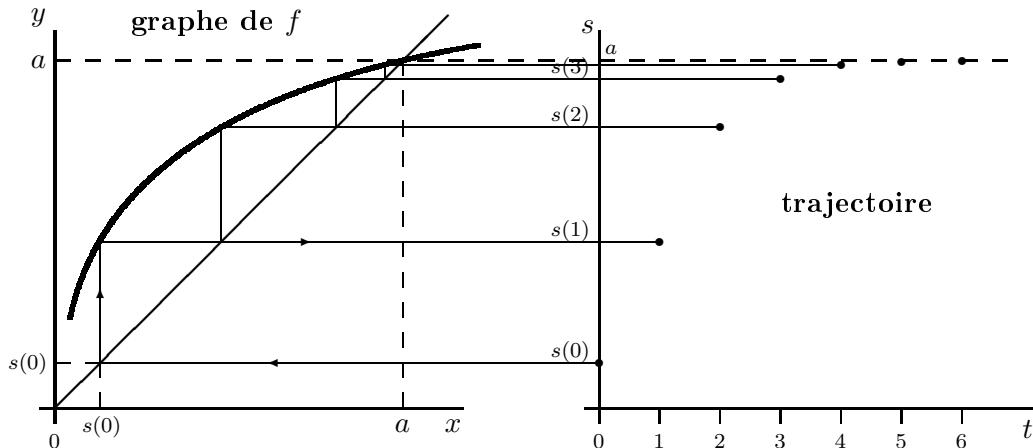
$$s_0, f(s_0), f^2(s_0), f^3(s_0) \dots \quad (1.3)$$

Dans la pratique, l'information sur l'état du système est *numérisée*, c'est-à-dire concentrée sous forme de m nombres réels. Ces nombres peuvent être des pressions, des températures, des positions, vitesses, concentrations de substances chimiques, nombre de cellules etc. Un état s est ainsi un m -uple $s = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m$ de nombres réels ($\mathbb{R} :=$ ensemble des nombres

réels ; voir p. 94) et l'ensemble E des états du système est un sous-ensemble de \mathbb{R}^m . Quant au choix de l'unité de temps, il dépend du phénomène à modéliser : seconde, année, temps d'une division cellulaire, etc.

L'itération d'une fonction est idéale pour le traitement informatique, grâce à des programmes simples et rapides du type "boucle". Ceci explique pourquoi les systèmes dynamiques discrets sont à la base de toutes les simulations sur ordinateurs. Dans MAPLE, la n^{e} itération f^n de f s'obtient par la commande `f@@n`.

Itération graphique : Lorsque l'ensemble des états E est l'ensemble \mathbb{R} des nombres réels, la fonction f peut être visualisée par son graphe. La trajectoire de $s(0)$ peut aussi s'obtenir géométriquement à partir de ce graphe. Pour cela, il faut placer côté à côté le plan (x,y) (où l'on dessine le graphe de f) et le plan (t,s) (où l'on dessine le graphe de la trajectoire de $s_0 = s(0)$). L'axe des x et celui des t doivent être placés à la même hauteur et les axes des x , des y et des s gradués à la même échelle.

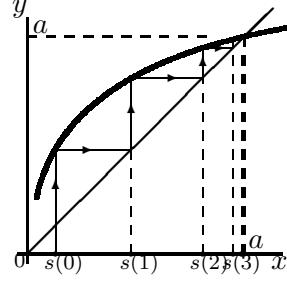


Les points de la trajectoire cherchée sont obtenus de la façon suivante : partant du point $(0, s(0))$ sur l'axe des s , on se déplace horizontalement vers la gauche jusqu'à la droite à 45° $\{y = x\}$. On monte alors verticalement jusqu'au graphe de f puis on revient horizontalement jusqu'à la droite $t = 1$. On a ainsi trouvé le point $(1, s(1))$. Cette opération peut se répéter à partir de n'importe quel point $(t, s(t))$ du graphe de la trajectoire et donne le point $(t + 1, s(t + 1))$. On peut ainsi dessiner le graphe de la trajectoire de s_0 .

Par exemple, dans la figure ci-dessus, on voit que la trajectoire tend vers l'état d'équilibre $s = a$ qui est un état stationnaire : $f(a) = a$ (voir p. 8).

Remarques :

- On ne dessine souvent que la partie gauche de la figure précédente, c'est-à-dire le cheminement en escalier entre la droite $\{x = y\}$ et le graphe de f . Cette partie contient en effet toute l'information sur la trajectoire.
- L'intérêt de cette amusante méthode graphique est surtout théorique, pour l'aide au raisonnement. On ne peut pas l'utiliser pour des résultats quantitatifs car l'imprécision des traits de crayon s'amplifie trop rapidement lors de l'itération.



1.2 La croissance exponentielle

Ce système dynamique est l'un des plus simples et des plus fréquents. Il modélise l'évolution d'une population en phase de croissance ou de décroissance non-freinée. Un état consiste en une seule valeur numérique $x \in \mathbb{R}$ et la loi d'évolution $f : \mathbb{R} \rightarrow \mathbb{R}$ est

$$f(x) := ax \quad (a \in \mathbb{R} \text{ constante}) \quad (1.4)$$

dont le graphe est une droite de pente a passant par l'origine.

Exemples :

1. Capital placé à taux d'intérêt constant τ . L'état x est le montant du capital et $a = 1 + \tau/100$.
2. Populations de cellules, de bactéries, etc. L'état x est simplement la masse ou la densité (=masse/volume). On a $a > 1$.
3. Échantillon d'un élément chimique radioactif. Dans ce cas x peut être la masse ou l'intensité du rayonnement de l'échantillon. On a $0 < a < 1$.
4. (Newton) Refroidissement d'un corps dans un fluide. L'état du système est la différence de température entre le corps et le fluide.

Si $x(0)$ est l'état initial au temps $t = 0$, sa trajectoire $x(n)$ sera donnée par

$$x(n) = a^n x(0) \quad (1.5)$$

Remarques :

1. si $a = -1$, on a une évolution périodique de période 2. L'état du système oscille indéfiniment entre $-x(0)$ et $x(0)$. Plus généralement, lorsque $a < 0$, l'état du système alterne à chaque étape entre les nombres positifs et négatifs.
2. La valeur $|x(t)|$ varie exponentiellement en fonction de t avec un **co-efficien de croissance** $|a|$.
3. Si $a > 0$, notre système est la discréétisation d'un système où le temps varie continûment et où l'état $x(t)$ est donné par

$$x(t) = x(t_0) a^{t-t_0} = x(t_0) e^{(t-t_0) \ln a}. \quad (1.6)$$

Si $a > 1$, on décrit souvent le système par le **temps de doublement** : c'est le temps T pour lequel on a $x(t+T) = 2x(t)$. Pour une culture de cellules, il coïncide avec le temps de division cellulaire. Analogiquement, si $0 < a < 1$, le système sera déterminé par sa **demi-vie** : le temps T nécessaire pour que $x(t+T) = \frac{1}{2}x(t)$. Dans les deux cas, on déduit de l'équation (1.6) que le rapport entre T et a est

$$T = \frac{\ln 2}{|\ln a|}. \quad (1.7)$$

Toujours par (1.6), si on connaît $x(t)$ et $x(t_0)$, on peut calculer T :

$$T = \frac{(t-t_0) \ln 2}{|\ln x(t) - \ln x(t_0)|} = \frac{(t-t_0) \ln 2}{|\ln \frac{x(t)}{x(t_0)}|}. \quad (1.8)$$

Applications :

- La concentration d'une population de bactéries s'est multipliée par 2,9 en 5 heures. Utilisant l'équation (1.8), le temps de doublement de la population est

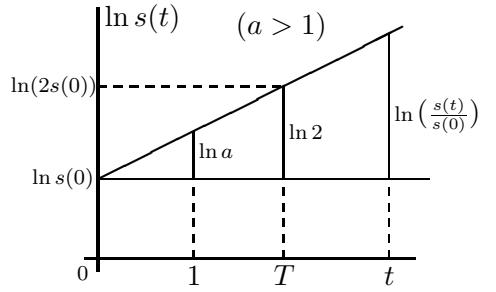
$$T = \frac{5 \ln 2}{\ln 2.9} \cong 3.26 \text{ heures}$$

- Le rayonnement dû au carbone 14 d'un échantillon de bois est 0.7 fois celui d'un même échantillon de ce bois vivant. La demi-vie T du C^{14} étant de 5730 ans, on déduit de (1.8) qu'il s'est écoulé

$$t = \frac{5730 |\ln(0.7)|}{\ln 2} \cong 2949 \text{ ans}$$

depuis la mort de l'arbre.

Graphique logarithmique : On reconnaît graphiquement un système à croissance exponentielle $s(t+1) = a s(t)$ en portant les points $(m, s(t+m))$ sur un graphique dont l'ordonnée est en échelle logarithmique. Ces points seront alignés sur une droite. Plus précisément, sur un graphique standard, les points $(m, \ln s(t+m))$ seront alignés sur une droite de pente $\ln a$. La figure ci-contre illustre ce fait (avec $t_0 = 0$ et $a > 1$). On retrouve géométriquement les formules (1.7) et (1.8) en utilisant le théorème de Thalès.



1.3 États stationnaires - Cycles - Stabilité

Dans la théorie des systèmes dynamiques, les notions suivantes sont importantes car elle permettent un discours *qualitatif* sur l'évolution d'un système :

Etats stationnaires : Un état s est dit **stationnaire** si $f(s) = s$. La trajectoire d'un état stationnaire est constante : $s(n) = s(0)$. Le système est en *état d'équilibre*. Un état stationnaire s est **stable** si la trajectoire des points voisins de s tend vers s . Autrement dit, s est stable s'il existe un voisinage V de s dans E tel que $s(0) \in V$ entraîne que $s(n) \rightarrow s$ lorsque $n \rightarrow \infty$. Le voisinage V des points qui sont ainsi attirés par s est le **bassin d'attraction** de s (par analogie avec le “bassin” d'un fleuve : l'ensemble des rivières qui s'y jettent).

Par exemple, dans la croissance exponentielle (§ 1.2, p. 6), l'état $x = 0$ est le seul état stationnaire. Il est stable si et seulement si $|a| < 1$. Le bassin d'attraction de 0 est alors tout \mathbb{R} .

Etats périodiques : Un état s est dit **périodique** s'il existe un entier $p \geq 1$ tel que $f^p(s) = s$. Le plus petit de ces entiers p est la **période** de s . Par exemple, un état stationnaire est périodique de période 1. La trajectoire d'un état périodique (de période p) s'appelle un **cycle** (de période p). Un cycle A est **stable** s'il existe un voisinage V de A dans E (son *bassin d'attraction*), tel que les trajectoires partant de V convergent vers A . Autrement dit, si $s(0) \in V$, la distance entre $s(n)$ et l'ensemble A tend vers 0 quand $n \rightarrow \infty$.

Exemple 1 : $E := \mathbb{R}$ et $f(s) = -s^3$. L'état $s = 0$ est stationnaire stable, son bassin étant $]-1, 1[$. L'ensemble $\{-1, 1\}$ est un cycle de période 2 non-stable.

Exemple 2 : $E := \mathbb{R}$ et $f(s) = -\sqrt[3]{s}$. Cette fois, c'est le cycle $\{-1, 1\}$ qui est stable, avec pour bassin d'attraction $\mathbb{R} - \{0\}$. L'état stationnaire $s = 0$ n'est pas stable.

Un état stationnaire ou un cycle stables sont des cas particuliers d'**attracteurs** - sous-systèmes admettant un bassin d'attraction. La notion de stabilité ou d'attracteur est très importante pour la modélisation scientifique, les phénomènes donnant lieu à des théories jouissant généralement d'une certaine stabilité. Les états que l'on rencontre sont alors très proches d'un attracteur (appelons-le A). De petites perturbations, ne changeront pas cette situation puisque les trajectoires seront rapidement ramenées vers A . Cela donne une robustesse à la théorie puisque l'évolution ne sera pas affectée par de légères perturbations, par exemple dues aux facteurs négligés par le modèle. Par contre, de plus grandes perturbations peuvent amener dans le bassin d'un autre attracteur.

Par exemple, les météorologues voient le climat actuel de la terre, relativement tempéré, comme un attracteur du système de l'atmosphère terrestre. Un autre attracteur est le climat glaciaire. Pendant les 100000 dernières années, notre système atmosphérique a passé plusieurs fois d'un attracteur à l'autre.

L'algèbre linéaire et le calcul différentiel fournissent des conditions suffisantes pour qu'un état stationnaire soit stable (voir §3.4, p. 64). Dans le cas d'un système à une variable (\mathbb{R}, f) , avec f dérivable, on a :

Condition sur la dérivée : Soit $a \in \mathbb{R}$ tel que $f(a) = a$ (état stationnaire). Alors

1. si $|f'(a)| < 1$, a est stable.
2. si $|f'(a)| > 1$, a est non-stable. ^a

^aLe cas $|f'(a)| = 1$ ne permet aucune conclusion sans information supplémentaire.

Exemple : Prenons la fonction $f(x) = x + \sin x$. Le point π est un état stationnaire et $f'(\pi) = 0$. Il est donc stable par la condition sur la dérivée. En partant de $x(0) = 2$, on obtient 9 décimales de π après 4 itérations :

| n | $x(n)$ |
|-----|-------------|
| 0 | 2.000000000 |
| 1 | 2.909297427 |
| 2 | 3.139509133 |
| 3 | 3.141592652 |
| 4 | 3.141592654 |
| 5 | 3.141592654 |

Obtenu par le programme MAPLE :

```
> x := 2 : f := x -> x+sin(x) :
for n from 0 to 5 do
[n,evalf((f@n)(x))] ; od;
```

La condition sur la dérivée ci-dessus se généralise pour un cycle $A := \{a_1, \dots, a_p\}$. On entend par là que $f(a_1) = a_2$, $f(a_2) = a_3$, etc et $f(a_p) = a_1$. Ceci est équivalent à ce que $f^p(a_1) = a_1$, donc que a_1 soit un état stationnaire pour f^p . La formule de dérivation d'une fonction composée donne $(f^p)'(a_1) = f'(a_1)f'(a_2)\cdots f'(a_p)$. On en déduit :

Proposition 1.3.1 *Soit $A := \{a_1, \dots, a_p\}$ un cycle de f . Alors :*

1. si $|f'(a_1)f'(a_2)\cdots f'(a_p)| < 1$, le cycle A est stable.
2. si $|f'(a_1)f'(a_2)\cdots f'(a_p)| > 1$, le cycle A est instable.

La détermination des bassins d'attraction est, en général, un problème difficile. Lorsque $f : \mathbb{R} \rightarrow \mathbb{R}$ est un polynôme, le théorème ci-dessous, démontré au début du siècle, permet de détecter la présence de cycles stables en donnant un point de leur bassin d'attraction.

Théorème 1.3.2 (Théorème de Fatou) *Supposons que la loi d'évolution $f : \mathbb{R} \rightarrow \mathbb{R}$ d'un système dynamique soit un polynôme de degré $d \geq 2$. Supposons que le système admette un cycle A stable. Alors le bassin d'attraction de A contient au moins un point x tel que $f'(x) = 0$.*

La dérivée f' de f étant un polynôme de degré $d - 1$, il y a au plus $d - 1$ points où f' s'annule (voir prop. A.7.1, p. 99). On en déduit que le système (\mathbb{R}, f) , avec f un polynôme de degré $d \geq 2$, admet au plus $d - 1$ cycles stables (il peut y en avoir moins, ou même pas du tout). Nous utiliserons ce théorème dans l'étude de la croissance logistique, au paragraphe suivant.

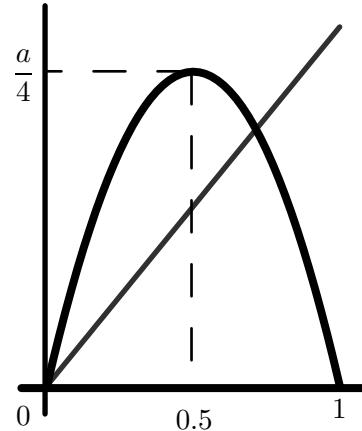
1.4 La croissance logistique

La croissance exponentielle étudiée au paragraphe 1.2 a peu de chance de pouvoir se poursuivre indéfiniment. Le système de la **croissance logistique** est l'un des plus courants pour modéliser la dynamique d'une population à plus long terme (voir [9, pp. 41–46], [11, p. 124–132], [12]). Il suppose que la population est limitée par l'environnement à une valeur maximale K . Il est alors naturel de remplacer le nombre d'individus s par $u := s/K$. Ainsi, la

quantité maximale d'individus est 1 et l'état du système est le nombre réel $u \in [0, 1]$. La loi d'évolution $f_a : [0, 1] \rightarrow [0, 1]$ pour la croissance logistique avec coefficient de fertilité a est la suivante :

$$f_a(u) := au(1 - u) = au - au^2 \quad (1.9)$$

dont le graphe est une parabole (voir ci-contre). Lorsque u est petit, le terme en u^2 est négligeable et on a $f_a(u) \approx au$. Approximativement, la population croît exponentiellement avec coefficient de croissance a . Lorsque u s'approche de 1, le terme $a(1 - u)$ devient plus petit que 1 et la population décroît.



Contrairement à la croissance exponentielle, l'état $u(n)$ n'aura pas d'expression simple telle que (1.5), la fonction f^2 étant un polynôme compliqué de degré 4, f^3 de degré 8, etc (essayez avec MAPLE). La formule devient rapidement impossible à écrire. On peut, en revanche, pratiquer une analyse qualitative des trajectoires avec les notions du paragraphe précédent.

Le graphe de f_a indique qu'il y a deux états stationnaires. Algébriquement, l'équation $f_a(u) = u$ a bien 2 solutions :

$$u = 0 \quad \text{et} \quad u = \frac{a - 1}{a} = 1 - \frac{1}{a}. \quad (1.10)$$

La dérivée $f'_a(u) = a - 2au$ de f_a en ces points est :

$$f'_a(0) = a \quad \text{et} \quad f'_a\left(\frac{a - 1}{a}\right) = 2 - a. \quad (1.11)$$

On suppose $a > 1$. La conditions sur la dérivée (voir p. 9) montre que l'état stationnaire $u = 0$ est instable pour $a > 1$. Quant à $u = 1 - 1/a$, il est stable si $1 < a < 3$ et instable lorsque $a > 3$. La population va ainsi évoluer de la manière suivante :

- 1) si $1 < a < 3$, quelque soit la population de départ u_0 , la population $u(t)$ tend vers l'état d'équilibre $u = 1 - 1/a$.
- 2) Lorsque $a > 3$, les deux états stationnaires sont instables et les trajectoires sont en quelque sorte repoussées de l'un vers l'autre. A mesure que a croît, des phénomènes de plus en plus compliqués se produisent. Pour certaines

valeurs de a , on va trouver des cycles stables et la population tendra vers un comportement cyclique. Observons que le point $u = 1/2$ est le seul où la dérivée de f_a s'annule. Par le théorème de Fatou (p. 10), si un cycle stable existe, le point $u = 1/2$ sera obligatoirement dans son bassin. On peut démontrer qu'en augmentant a , on obtient des attracteurs périodiques de toute période : d'abord 2, puis 4, 8, etc, et finalement 7, 5 puis 3 (l'ordre exact d'apparition a été découvert par le mathématicien ukrainien Sharkowskii en 1964). Ces situations sont illustrées dans le tableau suivant.

| Croissance logistique $u(n)$ pour $n \geq 50$ avec $u_0 = 0.5$ | | | | | |
|--|-----------|-----------|------------|-------------|-------------|
| n | $a = 3.3$ | $a = 3.5$ | $a = 3.55$ | $a = 3.739$ | $a = 3.833$ |
| 50 | 0.479427 | 0.382820 | 0.354798 | 0.499626 | 0.153411 |
| 51 | 0.823603 | 0.826941 | 0.812653 | 0.934749 | 0.497815 |
| 52 | 0.479427 | 0.500884 | 0.540480 | 0.228052 | 0.958232 |
| 53 | 0.823603 | 0.874997 | 0.881683 | 0.658230 | 0.153411 |
| 54 | 0.479427 | 0.382820 | 0.370329 | 0.841137 | 0.497815 |
| 55 | 0.823603 | 0.826941 | 0.827809 | 0.499626 | 0.958232 |
| 56 | 0.479427 | 0.500884 | 0.506021 | 0.934749 | 0.153411 |
| 57 | 0.823603 | 0.874997 | 0.887371 | 0.228052 | 0.497815 |
| 58 | 0.479427 | 0.382820 | 0.354800 | 0.658230 | 0.958232 |
| 59 | 0.823603 | 0.826941 | 0.812655 | 0.841137 | 0.153411 |
| 60 | 0.479427 | 0.500884 | 0.540477 | 0.499626 | 0.497815 |
| 61 | 0.823603 | 0.874997 | 0.881684 | 0.934749 | 0.958232 |
| 62 | 0.479427 | 0.382820 | 0.370328 | 0.228052 | 0.153411 |
| 63 | 0.823603 | 0.826941 | 0.827807 | 0.658230 | 0.497815 |
| 64 | 0.479427 | 0.500884 | 0.506026 | 0.841137 | 0.958232 |
| 65 | 0.823603 | 0.874997 | 0.887371 | 0.499626 | 0.153411 |
| 66 | 0.479427 | 0.382820 | 0.354800 | 0.934749 | 0.497815 |
| période | 2 | 4 | 8 | 5 | 3 |

3) Lorsque u s'approche de 4, le système devient **chaotique**. Les trajectoires semblent évoluer aléatoirement dans tout l'intervalle $[0,1]$ ou dans certains sous-intervalles. Le système présente alors une **sensibilité aux conditions initiales** : deux trajectoires, mêmes issues d'états initiaux très proches, divergent rapidement l'une de l'autre et ne présentent plus aucune similitude (comme dans un écoulement très turbulent). Ci-contre, pour $a = 3.9$, les 12 premiers états des trajectoires de 0.600 et 0.601.

| $a = 3.9$ | | |
|-----------|-------|-------|
| 0 | 0.600 | 0.601 |
| 1 | 0.936 | 0.935 |
| 2 | 0.234 | 0.236 |
| 3 | 0.698 | 0.704 |
| 4 | 0.822 | 0.813 |
| 5 | 0.571 | 0.593 |
| 6 | 0.955 | 0.941 |
| 7 | 0.167 | 0.215 |
| 8 | 0.543 | 0.658 |
| 9 | 0.968 | 0.877 |
| 10 | 0.122 | 0.421 |
| 11 | 0.417 | 0.950 |
| 12 | 0.948 | 0.183 |

1.5 Déterminisme, prévisibilité, hasard

Les travaux sur les systèmes dynamiques de ces 30 dernières années, par exemple les phénomènes vus au paragraphe précédent concernant la croissance logistique, ont changé notre façon de comprendre les concepts de *déterminisme* et de *prévisibilité*. Ils montrent que, contrairement à ce que l'on croyait, ces notions sont fondamentalement différentes.

En effet, dans la croissance logistique, l'évolution d'une population initiale $u(0)$ est *déterminée*, par la loi assez simple $u(n+1) = au(n)(1 - u(n))$. En revanche, si, comme cela est le cas en pratique, $u(0)$ est donné par une approximation numérique ou résulte de mesures expérimentales, les trajectoires sont foncièrement *imprévisibles* (lorsque a est grand). La marge d'erreur ne permet de prévoir que quelques étapes. Améliorer la prévisibilité est très coûteux car le besoin en précision sur la donnée initiale augmente exponentiellement en fonction du nombre d'étapes que l'on veut contrôler.

Ces cas de *sensibilité aux conditions initiales* semblent très fréquents. Par exemple, il paraissent inhérents aux systèmes dynamiques utilisés en météorologie, ce qui rend coûteuses et peut-être impossible les prévisions météorologiques à plus de quelques jours.

De même, on a vu que des systèmes dynamiques déterminés peuvent avoir des trajectoires de nature aléatoire (on utilise d'ailleurs des systèmes de ce type pour générer des nombres aléatoires dans les ordinateurs). Le fait que l'on observe un phénomène qui, apparemment, évolue au hasard ne prouve donc pas que ce phénomène n'obéisse à aucune loi.

Pour en savoir plus sur ces questions, voir [13].

Chapitre 2

Algèbre Linéaire

Supposons que des causes x, y, \dots déterminent des effets $f(x), f(y), \dots$. On dit que l'on **raisonne linéairement** si l'on suppose que l'addition de deux causes additionne les effets correspondants et qu'amplifier une cause d'un certain facteur amplifie l'effet correspondant dans la même proportion. En formule :

$$f(x + y) = f(x) + f(y) \quad \text{et} \quad f(\lambda x) = \lambda f(x) \quad (\lambda \in \mathbb{R}). \quad (2.1)$$

On dit alors que l'application $f : E \rightarrow F$ qui associe à une cause son effet est **linéaire**. Il faut évidemment que les expressions $x + y$, λx , etc, aient un sens et que ces procédés d'addition et d'amplification jouissent de propriétés raisonnables. Ceci est le cas lorsque E et F sont des *espaces vectoriels* (p. 101), par exemple \mathbb{R}^n .

Le raisonnement linéaire est le plus simple dont on dispose. S'il ne correspond pas à la réalité, on cherche souvent à l'utiliser en première approximation.

L'**algèbre linéaire** est la branche des mathématiques qui formalise le raisonnement linéaire, étudiant les espaces vectoriels et les applications linéaires. Ses outils et ses résultats sont utilisés dans d'innombrables applications des mathématiques et dans les statistiques. De plus, le calcul vectoriel est la façon opératoire de faire de la géométrie et donc important pour maîtriser l'espace. C'est pourquoi il est important que l'étudiant en sciences se familiarise dès que possibles avec les notions d'algèbre linéaire les plus courantes (vecteurs, matrices, déterminants, valeurs et vecteurs propres).

2.1 Structure d'espace vectoriel sur \mathbb{R}^n

2.1.1 Définitions

Deux points de \mathbb{R}^n peuvent être additionnés : si $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$, leur somme est définie par

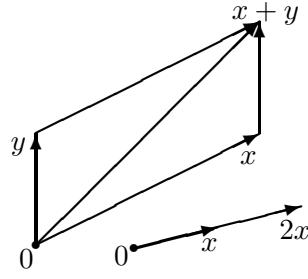
$$x + y := (x_1 + y_1, \dots, x_n + y_n).$$

D'autre part, on peut multiplier $x = (x_1, \dots, x_n)$ par $\lambda \in \mathbb{R}$ selon la règle

$$\lambda \cdot x := (\lambda x_1, \dots, \lambda x_n).$$

Ces deux opérations munissent \mathbb{R}^n d'une **structure d'espace vectoriel (réel)**. Pour la définition précise d'un espace vectoriel, voir p. 101.

Les éléments de \mathbb{R}^n peuvent être vus soit comme des points dans un espace, soit comme des vecteurs. Lorsqu'on imagine un point $p \in \mathbb{R}^n$ comme un vecteur, il est visualisé comme une flèche partant de l'origine jusqu'au point p . L'addition des vecteurs est comme celles des forces en statique. La multiplication par un "scalaire" $\lambda > 0$ est une homothétie de rapport λ .



2.1.2 Produit scalaire – Norme – Angles

On ne multiplie pas deux vecteurs $x, y \in \mathbb{R}^n$, mais on peut faire leur **produit scalaire**, qui est le nombre réel $x \cdot y$ défini par

$$x \cdot y := \sum_{i=1}^n x_i y_i.$$

Exemples : 1) dans \mathbb{R}^3 , le produit scalaire de $(1, 2, -1)$ et $(3, 0, 2)$ est :

$$(1, 2, -1) \cdot (3, 0, 2) = 1 \cdot 3 + 2 \cdot 0 + (-1) \cdot 2 = 1.$$

2) Si $a \in \mathbb{R}$, on définit $\mathbf{a} \in \mathbb{R}^n$ comme le vecteur dont toutes les composantes valent a : $\mathbf{a} := (a, \dots, a)$. Le produit scalaire de $x = (x_1, \dots, x_n)$ avec $\mathbf{1}$ ou $\frac{\mathbf{1}}{n}$ donne la somme ou la moyenne des nombres x_i :

$$\mathbf{1} \cdot x = \sum_{i=1}^n x_i \quad \text{et} \quad \frac{\mathbf{1}}{n} \cdot x = \frac{1}{n} \sum_{i=1}^n x_i.$$

Observons que

$$x \cdot x = \sum_{i=1}^n x_i^2 \geq 0.$$

Ceci permet de définir la **norme** $\|x\|$ de $x \in \mathbb{R}^n$ par

$$\|x\| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^n x_i^2} \geq 0.$$

Par le théorème de Pythagore, la norme de x représente la longueur du vecteur x , c'est-à-dire la distance entre le point x et l'origine.

Proposition 2.1.1 *Les propriétés du produit scalaire et de la norme sont*

- 1) $(x + y) \cdot z = x \cdot z + y \cdot z$.
- 2) $(\lambda x) \cdot y = \lambda(x \cdot y)$.
- 3) $x \cdot y = y \cdot x$.
- 4) $\|x\| \geq 0$ et ($\|x\| = 0 \Leftrightarrow x = 0$).
- 5) $\|\lambda x\| = |\lambda| \|x\|$.
- 6) $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2x \cdot y$.
- 7) $|x \cdot y| \leq \|x\| \|y\|$ (**inégalité de Schwarz**).
- 8) $\|x + y\| \leq \|x\| + \|y\|$ (**inégalité du triangle**).

La démonstration de ces propriétés est donnée dans l'annexe A.8 p. 102.

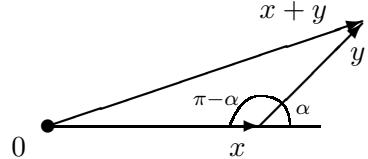
Angles : Par l'inégalité de Schwarz, le quotient $|x \cdot y| / (\|x\| \|y\|)$ est ≤ 1 .

Il existe donc un unique $\alpha \in [0, \pi]$ tel que

$$\cos \alpha = \frac{x \cdot y}{\|x\| \|y\|} \quad (\alpha \in [0, \pi]). \quad (2.2)$$

On dira, par définition, que α est l'**angle entre les vecteurs** x et y de \mathbb{R}^n . Lorsque $n = 2, 3$, cela correspond bien à l'angle usuel. En effet, le théorème du cosinus implique que

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos(\pi-\alpha) = \\ &= \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| \cos \alpha. \end{aligned}$$



En comparant cette dernière équation avec la propriété 6) de la proposition 2.1.1, on a bien la formule (2.2).

En particulier, deux vecteurs $x, y \in \mathbb{R}^n$ seront dits **orthogonaux** si et seulement si $x \cdot y = 0$. Dans ce cas on a le théorème de Pythagore :

$$x \cdot y = 0 \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2. \quad (2.3)$$

2.1.3 Indépendance linéaire - Bases - Dimension d'un espace vectoriel

Soit $\mathcal{V} := \{v_1, \dots, v_k\} \subset V$ une famille de vecteurs dans un espace vectoriel V . Un vecteur $x \in V$ qui peut s'écrire sous la forme

$$x = \sum_{i=1}^k \lambda_i v_i \quad (\lambda_i \in \mathbb{R}), \quad (2.4)$$

s'appelle une **combinaison linéaire** des vecteurs v_i . L'ensembles de toutes les combinaisons linéaires de vecteurs de \mathcal{V} forme un sous-espace vectoriel $\text{EV}(\mathcal{V})$ de V ; c'est le **sous-espace vectoriel engendré par \mathcal{V}** .

On dit que la famille \mathcal{V} est **libre** si le sous-espace vectoriel $\text{EV}(\mathcal{V})$ ne peut pas être engendré par une sous-famille $\mathcal{V}' \subset \mathcal{V}$ non-égale à \mathcal{V} . Les vecteurs d'une famille libre sont dits **linéairement indépendants**.

Exemples :

1. Une famille libre ne peut pas contenir le vecteur nul.
2. Deux vecteurs sont linéairement indépendants s'ils ne sont pas alignés.
3. Trois vecteurs sont linéairement indépendants s'ils ne sont pas coplanaires.

On démontre facilement la proposition suivante (exercice) :

Proposition 2.1.2 *Pour une famille $\mathcal{V} := \{v_1, \dots, v_k\} \subset V$, les conditions suivantes sont équivalentes :*

1. \mathcal{V} est libre.
2. aucun vecteur de \mathcal{V} n'est combinaison linéaire des autres.
3. l'équation

$$\sum_{i=1}^k \lambda_i v_i = 0 \quad (\lambda_i \in \mathbb{R}) \quad (2.5)$$

n'est possible que si $\lambda_i = 0$ pour tout i (Cette dernière condition est souvent prise pour définition d'une famille libre dans la littérature).

Supposons que $\mathcal{V} := \{v_1, \dots, v_k\} \subset V$ soit une famille libre. Alors, tout vecteur $x \in \text{EV}(\mathcal{V})$ s'écrit de manière *unique* $x = \sum x_i v_i$. En effet, si

$$x = \sum_{i=1}^k x_i v_i = \sum_{i=1}^k x'_i v_i,$$

on aurait

$$\sum_{i=1}^k (x'_i - x_i) v_i = 0$$

et donc $x'_i - x_i = 0$ puisque les v_i sont linéairement indépendants.

Une famille libre \mathcal{B} qui engendre un sous-espace vectoriel W de V s'appelle une **base** de W . Si $x \in W$, les uniques scalaires $x_i \in \mathbb{R}$ tels que $x = \sum_{i=1}^k x_i v_i$ s'appellent les **coordonnées de x dans la base $\{v_1, \dots, v_n\}$** .

On peut démontrer que toutes les bases d'un espace vectoriel V ont le même nombre d'éléments. Ce nombre s'appelle la **dimension** de V . Par exemple, les vecteurs

$$e_1 := (1, 0, \dots, 0), \quad e_2 := (0, 1, \dots, 0), \quad \dots, \quad e_n := (0, 0, \dots, 1)$$

forment une base de \mathbb{R}^n . En effet l'équation

$$(x_1, \dots, x_n) = \sum_{i=1}^n x_i e_i \tag{2.6}$$

montre à la fois que les e_i engendent \mathbb{R}^n et qu'ils sont linéairement indépendants. L'espace \mathbb{R}^n est donc de dimension n et toutes ses bases auront n éléments.

La base $\{e_1, \dots, e_n\}$ s'appelle la **base standard** de \mathbb{R}^n . Si $x = (x_1, \dots, x_n)$, l'équation (2.6) montre que les scalaires x_i sont les coordonnées de x dans la base standard.

Il y a beaucoup d'autres bases de \mathbb{R}^n . Par exemple, pour \mathbb{R}^3 , on peut prendre

$$(1, 0, 0), \quad (1, 1, 0), \quad (1, 1, 1).$$

Une façon courante de voir que des vecteurs sont linéairement indépendants, par le calcul d'un déterminant, sera vue à la proposition 2.2.7, p. 24.

2.2 Matrices

2.2.1 Définitions

Une **matrice** de taille $(p \times q)$, ou $(p \times q)$ -**matrice**, est une application

$$A : \{1, \dots, p\} \times \{1, \dots, q\} \rightarrow \mathbb{R}$$

qui à chaque couple d'entiers (i, j) fait correspondre un **coefficent** $a_{ij} \in \mathbb{R}$. Cette information peut se visualiser par un tableau à p lignes et q colonnes dans lequel le coefficient a_{ij} occupe l'intersection de la i^{e} ligne avec la j^{e} colonne :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix}.$$

On peut **additionner deux matrices** de même taille, en additionnant coefficient par coefficient :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 4 \\ 5 & 6 & 6 \end{pmatrix}.$$

On peut aussi **multiplier une matrice par un nombre réel λ** ; cela consiste à multiplier tous les coefficients par λ :

$$2 \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}.$$

Avec ces deux opérations, l'ensemble $\mathcal{M}_{p \times q}$ des $(p \times q)$ -matrices est un espace vectoriel de dimension pq .

Une matrice peut être vue comme un empilement de “vecteurs ligne” ou une juxtaposition de “vecteurs colonne”. Le $i^{\text{ème}}$ **vecteur ligne** $L_i(A)$ et le $j^{\text{ème}}$ **vecteur colonne** $C_j(A)$ de $A \in \mathcal{M}_{p \times q}$ sont

$$L_i(A) := (a_{i1} \cdots a_{iq}) \quad \text{et} \quad C_j(A) := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{pj} \end{pmatrix}.$$

Le **produit AB de deux matrices** A et B est défini si le nombre de colonnes de A est égal au nombre de lignes de B . Si $A \in \mathcal{M}_{p \times q}$ et $B \in \mathcal{M}_{q \times r}$ la matrice produit $C = AB \in \mathcal{M}_{p \times r}$ est définie par

$$c_{ij} := \sum_{k=1}^q a_{ik} b_{kj}.$$

Autrement dit, c_{ij} est le produit scalaire de $L_i(A)$ avec $C_j(B)$. Exemple :

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}. \quad (2.7)$$

On peut démontrer que les produits de matrices, lorsqu'ils sont définis, jouissent des propriétés suivantes :

- 1) $A(BC) = (AB)C$.
- 2) $A(B + C) = AB + AC$.
- 3) $(A + B)C = AC + BC$.

En revanche, le produit n'est pas commutatif ($AB \neq BA$, en général). Premièrement, AB peut être défini mais pas BA , comme dans (2.7) ci-dessus. Deuxièmement, AB et BA peuvent être définis mais pas de même taille :

$$(1 \ 2) \begin{pmatrix} 2 \\ 0 \end{pmatrix} = (2) \quad \text{et} \quad \begin{pmatrix} 2 \\ 0 \end{pmatrix} (1 \ 2) = \begin{pmatrix} 2 & 4 \\ 0 & 0 \end{pmatrix}.$$

Troisièmement, même dans le cas de **matrices carrées**, où la taille est conservée, on a $AB \neq BA$ en général :

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ce dernier exemple montre aussi que l'on peut obtenir la matrice nulle par produit de deux matrices qui ne le sont pas.

Soit A une matrice $p \times q$. La **transposée** A^T de A est la $(q \times p)$ -matrice dont les vecteurs colonnes sont les vecteurs lignes de A . Exemple :

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \implies A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Si le produit AB de deux matrices existe, alors $B^T A^T$ existe et on a $(AB)^T = B^T A^T$.

2.2.2 Déterminants

A toute matrice carrée A , on associe un nombre réel, son **déterminant**, que l'on note $\det A$. Si A est présentée sous forme d'un tableau, $\det A$ peut

aussi être noté par le même tableau entre barres verticales :

$$\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

La définition rigoureuse du déterminant nécessite un certain bagage théorique (voir [2, ch. IV]). Nous nous contenterons ici d'expliquer les algorithmes pour le calculer qui sont les suivants :

1. si (a) est une matrice (1×1) , alors $\det(a) := a$.
2. pour une matrice (2×2) , on définit

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc$$

3. si $A = (a_{ij})$ est une matrice $(n \times n)$, on choisit n'importe quelle ligne de A (disons la i ^{ème}). On peut alors calculer $\det A$ par la formule :

$$\det A = \sum_{k=1}^n (-1)^{i+k} a_{ik} \cdot D_{ik} \quad (2.8)$$

où D_{ik} (le (i, k) ^{ème} **mineur** de la matrice A) est le déterminant de la $(n-1) \times (n-1)$ -matrice obtenue en supprimant la i ^{ème} ligne et la k ^{ème} colonne de A . On dit que l'on a calculé $\det A$ par **développement par rapport à la i ^{ème} ligne**.

On peut aussi choisir une colonne (disons la j ^{ème}) et développer par rapport à cette colonne :

$$\det A = \sum_{k=1}^n (-1)^{j+k} a_{kj} \cdot D_{kj}. \quad (2.9)$$

Ces formules permettent le calcul du déterminant de n'importe quelle matrice en se ramenant successivement au calcul de déterminants de matrices plus petites. On peut démontrer que le résultat obtenu ne dépend pas de la ligne ou la colonne choisie.

Remarque : Le signe $(-1)^{i+k}$ est la mise en formule du schéma en damier :

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & & & & \end{pmatrix}.$$

Exemples : 1. Développement par rapport à la 3^e ligne :

$$\begin{vmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \\ -1 & 1 & 0 \end{vmatrix} = (-1) \begin{vmatrix} 3 & 0 \\ 1 & 4 \end{vmatrix} - \begin{vmatrix} 1 & 0 \\ 2 & 4 \end{vmatrix} = -12 - 4 = -16.$$

Si l'on développe par rapport à la 3^e colonne , on obtient :

$$\begin{vmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \\ -1 & 1 & 0 \end{vmatrix} = -4 \begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = -16.$$

On voit qu'il est avantageux de choisir les lignes ou colonnes comportant beaucoup de zéros.

2. Le déterminant de matrices **triangulaires** est le produit des éléments de leur diagonale :

$$\begin{vmatrix} \lambda_1 & & * & & \\ & \ddots & & & \\ 0 & & & & \\ & & & \ddots & 0 \\ & & & & \lambda_n \end{vmatrix} = \begin{vmatrix} \lambda_1 & & & & \\ * & \ddots & & & \\ & & \ddots & & 0 \\ & & & & \lambda_n \end{vmatrix} = \lambda_1 \cdots \lambda_n.$$

(Un gros zéro indique une région (ici : dessus ou dessous la diagonale) dont tous les coefficients sont nuls. Un astérisque indique des coefficients quelconques dont on ne se préoccupe pas. Une matrice triangulaire avec * = 0 est dite **diagonale**.)

Propriétés diverses du déterminant :

2.2.1 Soit A' la matrice obtenue de A en ajoutant à une ligne de A (disons la i ^{ème}) une combinaison linéaire des autres lignes de A :

$$L_i(A') = L_i(A) + \sum_{\substack{k=1 \\ (k \neq i)}}^n \lambda_k L_k(A) \quad , \quad L_r(A') = L_r(A) \text{ si } r \neq i$$

Alors $\det A' = \det A$. De même, on ne change pas le déterminant en ajoutant à une colonne de A une combinaison linéaire des autres colonnes.

Cette propriété est souvent utilisée pour simplifier le calcul de $\det A$, en augmentant le nombre de coefficients nuls.

2.2.2 Si A' est la matrice obtenue de A en échangeant deux lignes de A , alors $\det A' = -\det A$. Il en est de même si l'on échange deux colonnes.

2.2.3 Si A' est la matrice obtenue de A en multipliant tous les éléments d'une ligne de A par $\lambda \in \mathbb{R}$, alors $\det A' = \lambda \cdot \det A$. Il en est de même si l'on multiplie par λ tous les éléments d'une colonne.

2.2.4 Le déterminant d'une matrice est égal au déterminant de sa transposée :

$$\det A = \det A^T.$$

Cette propriété explique pourquoi chaque énoncé faisant intervenir des lignes a son analogue avec les colonnes.

2.2.5 $\det(AB) = \det A \cdot \det B$.

Inverse d'une matrice carrée : La $(n \times n)$ -matrice diagonale

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

s'appelle la **matrice identité** (d'ordre n). Elle joue le rôle d'élément neutre pour le produit des matrices $n \times n$:

$$I_n X = X I_n = X \quad \forall X \in \mathcal{M}_{n \times n}. \quad (2.10)$$

Une matrice $A \in \mathcal{M}_{n \times n}$ est dite **inversible** s'il existe une matrice $A^{-1} \in \mathcal{M}_{n \times n}$ telle que

$$A A^{-1} = A^{-1} A = I_n. \quad (2.11)$$

Une condition nécessaire pour que A soit inversible est que $\det A \neq 0$. En effet, par 2.2.5, on a :

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

(on aura donc $\det(A^{-1}) = (\det A)^{-1}$). Il se trouve que la condition $\det A \neq 0$ est aussi suffisante pour que A soit inversible. En effet, on peut démontrer ([2, p. 153]) :

Théorème 2.2.6 Une $(n \times n)$ -matrice A est inversible si et seulement si $\det A \neq 0$. Dans ce cas, son inverse $A^{-1} = (b_{kl})$ se calcule avec les mineurs D_{ij} de A par la formule :

$$b_{kl} = \frac{(-1)^{k+l}}{\det A} D_{lk}.$$

Exemples :

1. Pour une matrice 2×2 :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2. Pour inverser une matrice diagonale, on inverse simplement les coefficients :

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix}.$$

3. Pour la (3×3) -matrice dont on a calculé le déterminant à la p. 22 :

$$\begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \\ -1 & 1 & 0 \end{pmatrix}^{-1} = \frac{-1}{16} \begin{pmatrix} -4 & 0 & 12 \\ -4 & 0 & -4 \\ 3 & -4 & -5 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & 0 & \frac{-3}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{-3}{16} & \frac{1}{4} & \frac{5}{16} \end{pmatrix}.$$

Déterminants et indépendance linéaire : Si l'un des vecteurs colonne d'une matrice A est combinaison linéaire des autres, on peut par l'opération 2.2.1 modifier A sans changer son déterminant de manière que cette colonne soit nulle. On en déduit que si les vecteurs colonnes de A sont linéairement dépendants, alors $\det A = 0$. Il en est de même pour les vecteurs lignes. En fait, on peut démontrer l'équivalence ([2, p. 156]) :

Proposition 2.2.7 *Soit A une $(n \times n)$ -matrice. Alors :*

$$\begin{array}{ccc} \text{les vecteurs} & & \text{les vecteurs ligne} \\ \text{colonne} & \iff & \text{de } A \text{ sont} \\ \text{de } A \text{ sont} & & \text{linéairement} \\ \text{linéairement} & & \text{indépendants} \\ \text{indépendants} & & \end{array} \iff \det A \neq 0.$$

Plus généralement soit A une $(p \times q)$ -matrice. Toutes les $(r \times r)$ -matrices que l'on peut obtenir à partir de A en supprimant des lignes et/ou des colonnes s'appellent des **mineurs d'ordre r** de A . On peut démontrer ([2, p. 156]) :

Proposition 2.2.8 (et définition) *Soit A une $(p \times q)$ -matrice. Les trois nombres suivants sont égaux :*

1. le nombre maximal de vecteurs colonne de A qui sont linéairement indépendants
 2. le nombre maximal de vecteurs ligne de A qui sont linéairement indépendants
 3. le plus grand ordre d'un mineur de A dont le déterminant est non-nul.
- Par définition, ce nombre s'appelle le **rang** de la matrice A .

Exemple : Le rang de la matrice

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

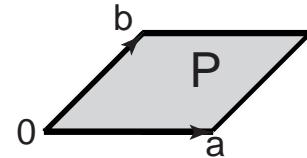
est égal à 2. En effet, il y a au plus deux vecteurs ligne linéairement indépendants puisque le premier et le troisième sont égaux. Il y en a au moins deux puisque le déterminant du mineur encadré est non nul.

2.2.3 Interprétation géométrique du déterminant

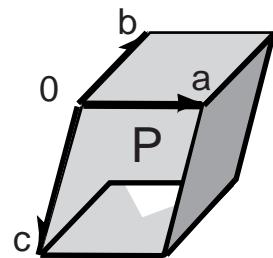
Le déterminant d'une $(n \times n)$ -matrice A est une mesure de l'indépendance linéaire des vecteurs ligne ou colonne (voir proposition 2.2.7). Pour $n = 2, 3$, on a l'interprétation géométrique suivante :

1) $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$ est, en valeur absolue, l'aire du parallélogramme P engendré par les vecteurs ligne

$$a = (a_1, a_2) \text{ et } b = (b_1, b_2).$$



2) $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$ est, en valeur absolue, le volume du parallélépipède P engendré par les vecteurs ligne $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$ et $c = (c_1, c_2, c_3)$.



Le signe du déterminant est positif si les vecteurs ligne forment une base d'orientation positive de \mathbb{R}^2 (sens trigonométrique) ou de \mathbb{R}^3 (règle du tire-bouchon). Dans la figure ci-dessus, le déterminant 2×2 est positif et le déterminant 3×3 est négatif.

2.2.4 Systèmes d'équations linéaires

Ici, “système” ne veut pas dire “système dynamique” mais ensemble d'équations. Considérons le **système** \mathcal{E} :

$$\mathcal{E} : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases} \quad (2.12)$$

de n **équations à p -inconnues** x_1, \dots, x_p . Les coefficients a_{ij} et b_i sont des nombres réels. Une solution de \mathcal{E} est un vecteur (x_1, \dots, x_p) satisfaisant à toutes les équations.

La matrice

$$A := A_{\mathcal{E}} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \in \mathcal{M}_{n \times p}$$

s'appelle la **matrice du système**. Si l'on définit les matrices colonnes

$$X := \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p \times 1}, \quad B := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{M}_{n \times 1},$$

le système \mathcal{E} est équivalent à l'équation matricielle

$$\mathcal{E} : AX = B. \quad (2.13)$$

Système de Cramer : Un système de Cramer¹ est un système carré ($p = n$) tel que $\det A \neq 0$. La matrice A est alors inversible et l'application $X \mapsto AX$ est une bijection de $\mathcal{M}_{n \times 1}$ sur lui-même. Un système de Cramer a donc une unique solution. Elle est donnée par

$$X = A^{-1}B. \quad (2.14)$$

En utilisant le calcul de la matrice inverse (2.2.6, p. 23), l'équation (2.14) est équivalente aux n équations

$$x_k = \frac{1}{\det A} \sum_{i=1}^n (-1)^{i+k} D_{ik} b_i \quad (k = 1, \dots, n). \quad (2.15)$$

¹Gabriel CRAMER (1704–1752) fut le premier professeur de mathématiques de l'Université de Genève, quand une chaire de cette discipline y fut créée en 1724.

Or, si l'on remplace dans la matrice A le k^{e} vecteur colonne par B et que l'on calcule le déterminant en développant par rapport à cette colonne, on obtient

$$\left| \begin{array}{cccccc} a_{11} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{array} \right| = \sum_{i=1}^n (-1)^{i+k} D_{ik} b_i. \quad (2.16)$$

La comparaison de (2.15) et (2.16) donne la règle de Cramer :

Théorème 2.2.9 (Règle de Cramer) *Un système de Cramer $AX = B$ a une solution unique donnée par*

$$x_k = \frac{1}{\det A} \left| \begin{array}{cccccc} a_{11} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{array} \right| \quad (k = 1, \dots, n).$$

Autres systèmes : Soit $\text{Sol } \mathcal{E} \subset \mathbb{R}^p$ l'ensemble des solutions du système \mathcal{E} . Lorsque l'on n'a pas un système de Cramer, par exemple lorsque $n \neq p$, il n'admet peut-être pas de solution ($\text{Sol } \mathcal{E} = \emptyset$). Par exemple :

$$\begin{cases} x + y = 1 \\ x + y = 2 \end{cases} \quad (2.17)$$

Pour un système $\mathcal{E} : AX = B$, on introduit les deux notions suivantes :

1. le **rang** du système est le rang de la matrice A
2. le système $AX = 0$ s'appelle le **système homogène** associé à \mathcal{E} . On le note \mathcal{E}_0 .

Supposons que le système \mathcal{E} admette une solution X_0 . L'ensemble $\text{Sol } \mathcal{E}$ satisfait alors à la propriété suivante :

Proposition 2.2.10 1. $\text{Sol } \mathcal{E} = \{X_0 + X \mid X \in \text{Sol } \mathcal{E}_0\}$.

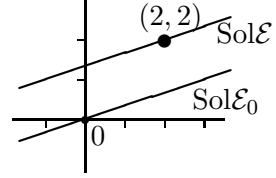
2. $\text{Sol } \mathcal{E}_0$ est un sous-espace vectoriel de \mathbb{R}^p de dimension $p - r$, où r est le rang de \mathcal{E} .

Remarque : La propriété 1) dit que pour trouver toutes les solutions de \mathcal{E} , il suffit de prendre une solution particulière X_0 et de lui ajouter toutes les solutions du système homogène associé. Grâce à 2), il suffit d'en trouver $p - r$ linéairement indépendantes et d'en faire toutes les combinaisons linéaires.

Exemple 1 : Considérons le système à 1 équation et 2 inconnues

$$\mathcal{E} : x - 3y = -4.$$

$\text{Sol } \mathcal{E}_0$ est la droite $y = \frac{x}{3}$ (sous-espace vectoriel de dimension 1). Prenons $X = (2, 2)$ (c'est-à-dire $x = y = 2$) comme solution particulière de \mathcal{E} . L'ensemble $\text{Sol } \mathcal{E}$ est donc une droite de pente $1/3$ passant par $(2, 2)$.



PREUVE DE 2.2.10 : La première affirmation est facile : si $X \in \text{Sol } \mathcal{E}$, alors $X - X_0 \in \text{Sol } \mathcal{E}_0$. De plus il est clair que $\text{Sol } \mathcal{E}_0$ est un espace vectoriel.

Pour la deuxième affirmation, on peut supposer, en renumérotant au besoin les équations et les inconnues, que le mineur d'ordre r du coin supérieur gauche de A est de déterminant non-nul :

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0. \quad (2.18)$$

Les r premiers vecteurs lignes L_1, \dots, L_r de A forment donc une base de l'espace vectoriel engendré par tous les vecteurs ligne. Les autres lignes, comme L_{r+1} , sont alors des combinaisons linéaires de L_1, \dots, L_r . Le système homogène \mathcal{E}_0 est ainsi équivalent au système des r premières équations, que l'on peut écrire :

$$\left\{ \begin{array}{lcl} a_{11}x_1 + \cdots + a_{1r}x_r & = & -a_{1,r+1}x_{r+1} - \cdots - a_{1p}x_p \\ \vdots & & \\ a_{r1}x_1 + \cdots + a_{rr}x_r & = & -a_{r,r+1}x_{r+1} - \cdots - a_{rp}x_p \end{array} \right. \quad (2.19)$$

A chaque valeur de (x_{r+1}, \dots, x_p) correspond une unique solution du système (2.19) car c'est un système de Cramer. On vérifie que cette correspondance donne un isomorphisme entre \mathbb{R}^{p-r} et $\text{Sol } \mathcal{E}_0$.

Pour trouver une solution particulière de \mathcal{E} , on pourra poser $x_{r+1} = \cdots = x_p = 0$. Pour trouver $p-r$ solutions indépendantes de \mathcal{E}_0 , il suffira de poser successivement $(x_{r+1}, \dots, x_p) = (1, 0, \dots, 0), (0, 1, \dots, 0), \dots (0, 0, \dots, 1)$. \square

Exemple :

$$\left\{ \begin{array}{lcl} x + y + z + t & = & 1 \\ y - z + t & = & 2 \\ x + 2y & & + 2t = 3 \end{array} \right. \quad (2.20)$$

Le rang du système est 2 puisque la troisième équation est la somme des deux premières et que le mineur d'ordre 2 du coin supérieur gauche est de déterminant 1. La solution particulière de \mathcal{E} de type $(x, y, 0, 0)$ est $(-1, 2, 0, 0)$. La solution particulière de \mathcal{E}_0 de la forme $(x, y, 1, 0)$ est $(-2, 1, 1, 0)$ et celle de la forme $(x, y, 0, 1)$ est $(0, -1, 0, 1)$. L'ensemble $\text{sol } \mathcal{E}_0$ est ainsi l'espace vectoriel de dimension 2 engendré par $(-2, 1, 1, 0)$ et $(0, -1, 0, 1)$ et on a :

$$\text{sol } \mathcal{E} = \{(-1, 2, 0, 0) + \lambda(-2, 1, 1, 0) + \mu(0, -1, 0, 1) \mid \lambda, \mu \in \mathbb{R}\}.$$

2.3 Applications linéaires

Définition : Une application $f : V \rightarrow W$ entre deux espaces vectoriels est dite **linéaire** si, pour tout $x, y \in V$ et $\lambda \in \mathbb{R}$, on a

1. $f(x + y) = f(x) + f(y)$ et
2. $f(\lambda x) = \lambda f(x)$.

Exemples :

- L'application nulle $f(x) = 0$ est linéaire.
- Toute application linéaire $f : \mathbb{R} \rightarrow \mathbb{R}$ est la multiplication par une constante : $f(x) = \alpha x$ (le graphe de f est donc une droite passant par 0). En effet : $f(x) = f(x \cdot 1) = xf(1)$; la constante α est $f(1)$.
- Plus généralement, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ donnée par $f(x) = \alpha x$, où $\alpha \in \mathbb{R}$, est linéaire. C'est l'**homothétie de rapport** α .

Matrices et applications linéaires : Une $(p \times q)$ -matrice A détermine une application linéaire de \mathbb{R}^q dans \mathbb{R}^p . C'est l'application $X \mapsto AX$ (on regarde les éléments de \mathbb{R}^q et \mathbb{R}^p comme des matrices à une colonne). Par exemple, prenons la matrice

$$A := \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \end{pmatrix}. \quad (2.21)$$

L'application $X \mapsto AX$ est

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y \\ x + 3y - z \end{pmatrix}.$$

L'application linéaire de \mathbb{R}^3 dans \mathbb{R}^2 déterminée par la matrice A est donc

$$(x, y, z) \mapsto (2x + y, x + 3y - z). \quad (2.22)$$

Comme l'énonce le théorème ci-dessous, ce procédé donne une bijection entre les applications linéaires de \mathbb{R}^q dans \mathbb{R}^p et les $(p \times q)$ -matrices. Rappeons que e_1, \dots, e_q désigne la base standard de \mathbb{R}^n (p. 18).

Théorème 2.3.1 *Soit $f : \mathbb{R}^q \rightarrow \mathbb{R}^p$ une application linéaire. Alors il existe une unique matrice $M_f \in \mathcal{M}_{p \times q}$ telle que f soit l'application linéaire déterminée par M_f (On dira que M_f est la **matrice de f**). La matrice M_f est construite de la manière suivante : son k ème vecteur colonne est $f(e_k)$.*

Exemple : Pour l'application linéaire $f(x, y, z) := (2x + y, x + 3y - z)$ de (2.22), on a

$$f(e_1) = (2, 1) \quad , \quad f(e_2) = (1, 3) \quad , \quad f(e_3) = (0, -1)$$

On retrouve bien les colonnes de la matrice (2.21).

PREUVE DE 2.3.1 : Désignons par $g : \mathbb{R}^q \rightarrow \mathbb{R}^p$ l'application linéaire déterminée par la matrice M_f construite dans le théorème 2.3.1. Nous allons montrer que $g = f$.

Soit E_k la matrice colonne $q \times 1$ correspondant au vecteur e_k . En faisant le produit $M_f E_k$, on obtient justement le k ème vecteur colonne de M_f , c'est-à-dire $f(e_k)$. On a donc $g(e_k) = f(e_k)$. Mais tout $x \in \mathbb{R}^q$ s'écrit

$$x = x_1 e_1 + \cdots + x_q e_q.$$

Puisque f et g sont linéaires, on aura

$$f(x) = x_1 f(e_1) + \cdots + x_q f(e_q) = x_1 g(e_1) + \cdots + x_q g(e_q) = g(x). \quad \square$$

La correspondance $f \mapsto M_f$ entre applications linéaires et matrices se comporte bien pour la composition des applications. En effet, on peut démontrer :

Proposition 2.3.2 *Soit $\mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R}^p$ deux applications linéaires. Alors, la composition $g \circ f$ est linéaire et sa matrice $M_{g \circ f}$ est donnée par le produit matriciel*

$$M_{g \circ f} = M_g M_f.$$

En particulier, la matrice de f^m (f itérée m fois) est M_f^m .

Exemple : La **rotation** ρ_α d'angle α autour de 0 dans \mathbb{R}^2 est une application linéaire $\rho_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dont la matrice est

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

La composition de deux rotations additionne simplement les angles : $\rho_\alpha \circ \rho_\beta = \rho_{\alpha+\beta}$. Par la Proposition 2.3.2, cela donne l'égalité matricielle :

$$\begin{aligned} \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} = \\ &= \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -\sin \alpha \cos \beta - \cos \alpha \sin \beta \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta & \cos \alpha \cos \beta - \sin \alpha \sin \beta \end{pmatrix}. \end{aligned}$$

Ceci démontre les formules classiques :

$$\begin{aligned} \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin(\alpha + \beta) &= \sin \alpha \cos \beta + \cos \alpha \sin \beta. \end{aligned}$$

2.4 Changements de bases

Problème : Soit $\mathcal{A} := \{a_1, \dots, a_n\}$ une base de \mathbb{R}^n . Soit $x := (x_1, \dots, x_n)$ un vecteur de \mathbb{R}^n . Quelles sont les coordonnées $(\tilde{x}_1, \dots, \tilde{x}_n)$ de x dans la base \mathcal{A} ?

Notons par $a_i = (a_{1i}, \dots, a_{ni})$ les coordonnées de a_i dans la base standard de \mathbb{R}^n . Cette notation d'indice est choisie pour que les a_i deviennent les vecteurs colonne d'une $(n \times n)$ -matrice

$$P := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \quad (2.23)$$

appelée **matrice de passage** pour la base \mathcal{A} . Comme les vecteurs a_i sont linéairement indépendants, $\det P \neq 0$ par la proposition 2.2.8, p. 24. On en déduit que la matrice P est inversible (théorème 2.2.6 p. 23). On peut démontrer ([2, p. 97]) que les coordonnées $(\tilde{x}_1, \dots, \tilde{x}_n)$ dans la base \mathcal{A} de $x = (x_1, \dots, x_n)$ sont données par l'équation matricielle :

$$\begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \quad (2.24)$$

Changement de matrice pour une application linéaire : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application linéaire dont la matrice est M_f . Si l'on travaille avec les coordonnées dans la base \mathcal{A} , l'application f sera donnée par une matrice \widetilde{M}_f . On dit que \widetilde{M}_f est la **matrice de f relativement à la base \mathcal{A}** . On peut démontrer ([2, p. 98-99]) que, si P est la matrice de passage (2.23) pour la base \mathcal{A} , alors, la matrice \widetilde{M}_f est obtenue à l'aide de M_f par la formule

$$\boxed{\widetilde{M}_f = P^{-1} M_f P.} \quad (2.25)$$

2.5 Valeurs propres - vecteurs propres

Valeurs et vecteurs propres sont parmi les concepts les plus importants de l'algèbre linéaire. Par exemple, ils sont essentiels pour la formulation de la mécanique quantique. Nous en rencontrerons plusieurs applications : systèmes dynamiques (§ 2.6, 3.4 et 5.3), extrema (p. 61), résolution d'équations différentielles (p. 84), composantes principales (p. 103).

2.5.1 Définitions

Soit $f : V \rightarrow V$ une application linéaire d'un espace vectoriel V dans lui-même. On dit que $\lambda \in \mathbb{R}$ est une **valeur propre** pour f s'il existe un vecteur $x \neq 0$ dans V tel que

$$f(x) = \lambda x. \quad (2.26)$$

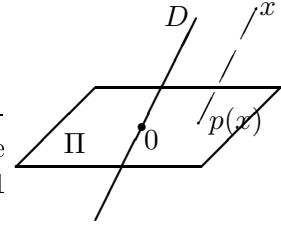
Un tel vecteur x est appelé un **vecteur propre**² pour la valeur propre λ .

Une $(n \times n)$ -matrice M représente une application linéaire de $\mathbb{R}^n \rightarrow \mathbb{R}^n$ qui aura d'éventuelles valeurs propres et vecteurs propres. On parlera alors des **valeurs propres et vecteurs propres de la matrice M** .

Exemples :

- L'application nulle $f(x) = 0$ a une seule valeur propre : $\lambda = 0$. Tous les vecteurs *non-nuls* sont vecteurs propres.
- λ est la seule valeur propre d'une homothétie de rapport λ . Dans ce cas tous les vecteurs non-nuls sont vecteurs propres.
- une rotation d'angle α dans le plan n'a pas de valeur propre, à moins que $\alpha = 0$ ou π .

²En anglais : valeur propre = eigenvalue, vecteur propre = eigenvector. Dans certains ouvrages de statistique, on utilise les termes "latent root" et "latent vector".



- dans \mathbb{R}^3 , la projection sur un plan Π parallèlement à une droite D . Les vecteurs non-nuls de Π sont des vecteurs propres de valeur propre 1 et ceux de D de valeur propre 0.

Pour trouver les valeurs propres de $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, on pratique de la manière suivante. Si λ est une valeur propre pour f , on a l'équation $f(x) - \lambda x = 0$. Matriciellement, utilisant la matrice M_f de f , cela peut s'écrire

$$(M_f - \lambda I)X = 0. \quad (2.27)$$

où I est la matrice identité. Par 2.2.9 Une telle équation a une solution avec $X \neq 0$ si et seulement si

$$\det(M_f - \lambda I) = 0. \quad (2.28)$$

Posons $P_f(\lambda) := \det(M_f - \lambda I)$. P_f est un polynôme de degré n en λ appelé le **polynôme caractéristique** de f . Le polynôme P_f a, au plus, n racines (solutions de l'équation $P_f(\lambda) = 0$, voir p. 99) qui sont des nombres complexes. L'ensemble $\text{Sp}_f \subset \mathbb{C}$ des racines du polynôme caractéristique P_f s'appelle le **spectre** de f . On définit également le **spectre** d'une $(n \times n)$ -matrice M comme le spectre de l'application linéaire de $\mathbb{R}^n \rightarrow \mathbb{R}^n$ que détermine M .

Le spectre d'une application linéaire est une information très importante sur cette application. Pour l'instant, intéressons-nous aux racines réelles qui, par l'équation 2.28, sont les valeurs propres de f :

Proposition 2.5.1 *Les valeurs propres de f sont les racines réelles du polynôme caractéristique P_f de f .*

Exemple 1 : Prenons $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par $f(x, y) = (y, x)$. Sa matrice M_f est

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

Son polynôme caractéristique P_f est

$$P_f := \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1$$

Le spectre de f sera donc l'ensemble des solutions de l'équation $\lambda^2 - 1 = 0$, c'est-à-dire $\text{Sp}_f = \{\pm 1\}$. Comme $\text{Sp}_f \subset \mathbb{R}$, on a deux valeurs propres ± 1 .

Les vecteurs propres pour $\lambda = -1$ sont les solutions non-nulles de l'équation matricielle

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix},$$

c'est-à-dire les solutions non-nulles du système d'équations

$$\begin{cases} y = -x \\ x = -y \end{cases}$$

On voit que l'on peut prendre $v_- = (1, -1)$ comme vecteur propre et que tous les autres vecteurs propres pour la valeur propre -1 sont les multiples non-nuls de v_- .

De même, on trouve que les vecteurs propres pour la valeur propre $+1$ sont les multiples non-nuls de $v_+ := (1, 1)$.

Dans la base $\{v_-, v_+\}$, la matrice de f est diagonale : $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

Exemple 2 : Soit f une rotation d'angle α dans le plan. Sa matrice M_f est

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Son polynôme caractéristique est

$$P_f(\lambda) = \begin{vmatrix} \cos \alpha - \lambda & -\sin \alpha \\ \sin \alpha & \cos \alpha - \lambda \end{vmatrix} = \lambda^2 - 2\lambda \cos \alpha + 1.$$

On a donc $Sp_f = \{\cos \alpha \pm i \sin \alpha = e^{\pm i\alpha}\}$. Si $\alpha \neq 0, \pi$, $Sp_f \cap \mathbb{R} = \emptyset$ et il n'y a pas de valeur propre. Cependant, on voit que Sp_f porte toute l'information sur f , puisqu'on peut en extraire l'angle α .

2.5.2 Diagonalisation

Une application linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est dite **diagonalisable** s'il existe une base de \mathbb{R}^n formée de vecteurs propres de f . Relativement à cette base \mathcal{A} , la matrice $M_{f,\mathcal{A}}$ de f sera diagonale avec comme coefficients diagonaux les valeurs propres de f .

Soit M une $(n \times n)$ -matrice. Elle représente une application linéaire de $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Si cette application est diagonalisable, on dira que la **matrice M est diagonalisable**. Par la proposition 2.25 p. 32, cela signifie qu'il existe une matrice inversible P (la matrice de passage vers une base \mathcal{A} de

vecteurs propres de f) telle que la matrice $P^{-1}MP$ est diagonale (avec, sur la diagonale, les valeurs propres de M). Les colonnes de P contiennent donc les coordonnées des vecteurs propres de f .

On voit que pour qu'une application linéaire (ou matrice) soit diagonalisable, il faut qu'elle ait assez de valeurs propres. En fait, on a :

Proposition 2.5.2 *Si une application linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est diagonalisable, alors $\text{Sp}_f \subset \mathbb{R}$. (toutes les racines du polynôme caractéristique P_f de f sont réelles).*

Le fait que $\text{Sp}_f \subset \mathbb{R}$ n'implique pas forcément que f soit diagonalisable (exemple 1.2 ci-dessous). Cependant, c'est le cas si l'on a n valeurs propres distinctes :

Proposition 2.5.3 *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application linéaire. Si f admet n valeurs propres deux-à-deux distinctes, alors f est diagonalisable.*

PREUVE: Il suffit de démontrer que les vecteurs propres v_i pour k valeurs propres $\lambda_1, \dots, \lambda_k$ distinctes sont linéairement indépendants (n vecteurs linéairement indépendants de \mathbb{R}^n formeront une base). Cela se démontre par récurrence sur k . C'est le cas si $k = 1$ puisqu'un vecteur propre, par définition, est non-nul.

Supposons que $x := \sum_{i=1}^k \alpha_i v_i = 0$. On calcule $f(x) - \lambda_k x$:

$$0 = f(x) - \lambda_k x = \sum_{i=1}^{k-1} \alpha_i (\lambda_i - \lambda_k) v_i$$

Par hypothèse de récurrence, on aura $\alpha_i(\lambda_i - \lambda_k) = 0$. Comme $\lambda_i \neq \lambda_k$, cela implique $\alpha_i = 0$ pour $i = 1, \dots, k-1$, et donc aussi $\alpha_k = 0$.

L'équation $\sum_{i=1}^k \alpha_i v_i = 0$ n'est donc possible que si $\alpha_i = 0$. Par la proposition 2.1.2 p. 17, ceci est équivalent à ce que les vecteurs v_i sont linéairement indépendants. \square

Exemple 1 : Pour une application linéaire $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, le polynôme caractéristique est du 2^e degré. On sait qu'un tel polynôme a deux racines λ_1 et λ_2 . On a les trois possibilités suivantes :

1. λ_1 et λ_2 réelles distinctes. Ayant deux valeurs propres distinctes, f est diagonalisable par la proposition 2.5.3 ci-dessus. On a donc une base \mathcal{A} de vecteurs propres relativement à laquelle la matrice de f est

$$M_{f,\mathcal{A}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

2. $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$. On a donc une seule valeur propre. On peut alors démontrer que f est diagonalisable si et seulement si f est une homothétie de rapport λ . Dans le cas contraire, on peut montrer qu'il existe une base \mathcal{A} de \mathbb{R}^2 telle que

$$M_{f,\mathcal{A}} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

3. $\text{Sp}_f = \{\rho e^{\pm i\alpha}\}$, avec $\alpha \neq 0, \pi$. Dans ce cas, f n'a pas de valeur propre et f n'est pas diagonalisable. On peut alors démontrer qu'il existe une base \mathcal{A} de \mathbb{R}^2 relativement à laquelle la matrice de f est de la forme

$$M_{f,\mathcal{A}} = \begin{pmatrix} \rho \cos \alpha & -\rho \sin \alpha \\ \rho \sin \alpha & \rho \cos \alpha \end{pmatrix}.$$

Exemple 2 : Considérons $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la projection sur un plan Π parallèlement à une droite D (voir p. 33). Supposons que Π est engendré par les vecteurs $a := (1, 2, -1)$ et $b := (0, 2, 1)$ et que D est engendrée par $d := (1, 1, -1)$. La base $\mathcal{A} := \{a, b, d\}$ est une base de vecteurs propres dans laquelle la matrice de f est diagonale :

$$\mathcal{M}_{f,\mathcal{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

La matrice P de passage pour la base \mathcal{A} est

$$P := \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}.$$

Par (2.25) la relation entre M_f et $M_{f,\mathcal{A}}$ est $M_{f,\mathcal{A}} = P^{-1}M_fP$. La matrice M_f de la projection f dans la base standard de \mathbb{R}^3 est donc

$$M_f = P M_{f,\mathcal{A}} P^{-1} = \begin{pmatrix} -3 & 1 & -2 \\ -4 & 2 & -2 \\ 4 & -1 & 3 \end{pmatrix}.$$

Exemple 3 : Un autre cas de diagonalisabilité est donné par les matrices symétriques (voir paragraphe suivant).

2.5.3 Matrices symétriques :

Une matrice carrée A est dite **symétrique** si elle est égale à sa transposée : $A^T = A$. Du point de vue des coefficients, cela veut dire que $a_{ij} = a_{ji}$. Les matrices symétriques jouissent de nombreuses propriétés, dont nous ne verrons que quelques unes ici, et apparaissent dans beaucoup d'applications : matrice Hessienne (p. 56), composantes principales (p. 103) géométrie des distances (stéréochiomie) etc. Une matrice symétrique est diagonalisable (Théorème 2.5.4 ci-dessous) et des algorithmes rapides permettent le calcul numérique de ses valeurs propres. De tels algorithmes sont implantés sur plusieurs logiciels (comme MATLAB).

Proposition 2.5.4 *Soit A une matrice symétrique. Alors*

1. $\text{Sp}_A \subset \mathbb{R}$.
2. *A est diagonalisable. En fait, A admet une base de vecteurs propres deux-à-deux orthogonaux.*

Remarquons que 2) implique 1) par la proposition 2.5.2, p. 35. Pour une preuve de 2.5.4, voir [2, p. 292].

Le résultat ci-dessous est utile pour savoir, sans avoir besoin de les calculer, si toutes les valeur propres d'une matrice symétrique A sont positives. Notons A_k le $k^{\text{ème}}$ **mineur principal** de A : c'est le mineur obtenu en prenant les k premières lignes et les k premières colonnes de A :

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1k} & \\ \vdots & & \vdots & \cdots \\ a_{k1} & \cdots & a_{kk} & \\ \vdots & & & \end{array} \right) \quad k^{\text{ème}} \text{ mineur principal}$$

Proposition 2.5.5 *Soit A une matrice symétrique. Alors, les deux conditions sont équivalentes :*

1. *toutes les valeurs propres de A sont > 0 .*
2. *le déterminant de chaque mineur principal est > 0 :*

$$\det A_k > 0 \quad \forall k = 1, 2, \dots, n.$$

2.6 Systèmes dynamiques linéaires

2.6.1 Stabilité de l'origine

Les systèmes dynamiques les mieux connus sont ceux l'ensemble des états est \mathbb{R}^n et la loi d'évolution $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est linéaire. Le point 0 est donc un

état stationnaire. Le théorème principal de ce paragraphe est une condition nécessaire et suffisante pour que 0 soit un état stationnaire stable.

Si v est un vecteur propre de f pour la valeur propre $\lambda \in \mathbb{R}$, la droite $\{\alpha v \mid \alpha \in \mathbb{R}\}$ constitue un sous-système à croissance exponentielle, avec coefficient de croissance λ . Dans ce sous-système, 0 est stable si et seulement si $|\lambda| < 1$. D'où une condition nécessaire pour que 0 soit un état stable pour le système (\mathbb{R}^n, f) est qu'il n'y ait pas de valeur propre λ avec $|\lambda| \geq 1$. Pour obtenir une condition nécessaire et suffisante, il faut prendre en compte non seulement les valeurs propres mais aussi tout le spectre de f (voir p. 33). Définissons le **rayon spectral** de f comme étant le maximum des modules des éléments de $\text{Sp}(f)$.

Théorème 2.6.1 *Le point $0 \in \mathbb{R}^n$ est un état stationnaire stable si et seulement si le rayon spectral de f est < 1 .*

La preuve de ce théorème dépasse le cadre de ce cours. Il y a néanmoins un cas facile : celui où f est diagonalisable (voir p. 34). En effet, on a alors une base $\mathcal{V} := \{v_1, \dots, v_n\}$ de vecteurs propres et, pour les coordonnées dans cette base, la matrice M_f de f est diagonale :

$$M_f = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ 0 & & & \lambda_n \end{pmatrix}$$

où les λ_i sont les valeurs propres de f . Quand on itère f m -fois, on a :

$$M_{f^m} = (M_f)^m = \begin{pmatrix} \lambda_1^m & & & 0 \\ & \ddots & & \\ 0 & & & \lambda_n^m \end{pmatrix}.$$

On voit que $(M_f)^m \rightarrow 0$ quand $m \rightarrow \infty$ si et seulement si tous les λ_i satisfont $|\lambda_i| < 1$.

Supposons que $\lambda_1 > \lambda_i$ pour $i \geq 2$. Soit $x := (x_1, \dots, x_n)$ (coordonnées dans la base de vecteurs propres \mathcal{V}). Si $x_1 \neq 0$, on a :

$$f^m(x) = (\lambda_1^m x_1, \dots, \lambda_n^m x_n) = \lambda_1^m x_1 p_m$$

où p_m est le vecteur

$$p_m := (1, (\frac{\lambda_2}{\lambda_1})^m \frac{x_2}{x_1}, \dots, (\frac{\lambda_n}{\lambda_1})^m \frac{x_n}{x_1}) \rightarrow (1, 0, \dots, 0) = v_1 \quad (m \rightarrow \infty).$$

Le vecteur normalisé $\frac{f^m(x)}{\|f^m(x)\|}$ tend donc vers le vecteur propre $\pm v_1$:

$$\frac{f^m(x)}{\|f^m(x)\|} \rightarrow \frac{x_1}{\|x_1\|} \frac{p_m}{\|p_m\|} = \pm v_1.$$

Plus généralement, on peut ainsi montrer le résultat suivant :

Proposition 2.6.2 *Soit $f : V \rightarrow V$ une application linéaire. Supposons que f admette une valeur propre $\lambda_{\max} \in \mathbb{R}_{\geq 0}$ qui soit strictement plus grande que le module de toute autre racine du polynôme caractéristique de f (en particulier, λ_{\max} est le rayon spectral de f). Soit v_{\max} un vecteur propre pour λ_{\max} avec $\|v_{\max}\| = 1$. Alors, pour presque tout état $z \in V$, on a*

$$\lim_{m \rightarrow \infty} \frac{f^m(z)}{\|f^m(z)\|} = \pm v_{\max} \quad \text{et} \quad \lim_{m \rightarrow \infty} \frac{\|f^m(z)\|}{\|f^{m-1}(z)\|} = \lambda_{\max}.$$

Cette proposition dit qu'à long terme, le système se comporte comme un système à croissance exponentielle avec coefficient de croissance λ_{\max} . Elle peut être utilisée pour trouver λ_{\max} , l'itération de f étant facile pour un ordinateur. Pour une application de la proposition 2.6.2 en dynamique des populations, voir l'exemple c), p. 41.

2.6.2 Populations avec pyramide d'âges

On présente ici un système dynamique discret pour l'évolution d'une population divisée en m classes d'âge C_1, \dots, C_m , des plus jeunes aux plus âgés. Ce modèle, qui remonte à Euler (1760), est utilisé pour toutes sorte de populations : populations humaines ([10, ch. 4],[8, § 1.2.2]) ou d'animaux dont on désire comprendre et/ou influencer le développement (repeuplement, lutte contre insectes ou parasites).

L'état de la population est un vecteur $(x_1, \dots, x_m) \in \mathbb{R}^m$, où x_i est le nombre d'individus de la classe C_i . On ne compte souvent que les femelles en âge d'avoir des descendants (c'est la donnée importante pour le renouvellement d'une population). Par exemple, pour une population humaine, on comptera le nombre de femmes entre 0 et 50 ans, par tranche de 5 ans, ce qui donne 10 classes (0-5 ans, 5-10 ans, ..., 45-50 ans). L'état du système est ainsi un vecteur $(x_1, \dots, x_{10}) \in \mathbb{R}^{10}$ et l'unité de temps est de 5 ans.

La loi d'évolution $f(x_1, \dots, x_m) = (x'_1, \dots, x'_m)$ est définie de la façon suivante : les femelles C_i donnent naissance à des femelles C_1 avec un coefficient de natalité β_i , d'où

$$x'_1 = \sum_{i=1}^m \beta_i x_i.$$

D'autre part, les femelles C_i passent dans la classe C_{i+1} avec un coefficient de survie α_i , ce qui se traduit par

$$x'_{i+1} = \alpha_i x_i \quad \text{si } i \geq 1.$$

Par convention, $\alpha_m = 0$, les individus de C_m sortant de la population considérée. En résumé,

$$f(x_1, \dots, x_m) = \left(\sum_{i=1}^m \beta_i x_i, \alpha_1 x_1, \alpha_2 x_2, \dots, \alpha_{m-1} x_{m-1} \right). \quad (2.29)$$

On voit que la loi d'évolution $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ est une application linéaire. Sa matrice M_f , pour $m = 2, 3, 4, \dots$, est

$$\begin{pmatrix} \beta_1 & \beta_2 \\ \alpha_1 & 0 \end{pmatrix}, \quad \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 \\ \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \end{pmatrix}, \quad \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & \alpha_3 & 0 \end{pmatrix}, \dots \quad (2.30)$$

En général, M_f contient les coefficients de natalité sur sa 1^{ère} ligne et les coefficients de survie sur sa 1^{ère} diagonale inférieure. Ces matrices portent le nom de *matrices de Leslie*.

Soit $p_k := \alpha_1 \cdots \alpha_k$; c'est la probabilité pour un nouveau-né de survivre à k étapes. Le nombre moyen de descendantes qu'une femelle aura pendant sa vie entière sera donc

$$E := \beta_1 + p_1 \beta_2 + p_2 \beta_3 + \cdots + p_{m-1} \beta_m. \quad (2.31)$$

On peut démontrer (voir [10, § 4.4]) :

Théorème 2.6.3 *a) f a une seule de ses valeurs propres qui est réelle positive. Appelons-la λ_{pos} .*

- b) λ_{pos} est le rayon spectral de f.*
- c) si $E < 1$, alors $\lambda_{\text{pos}} < 1$. Si $E > 1$, alors $\lambda_{\text{pos}} > 1$.*
- d) $E = 1 - (-1)^m P_f(1)$.*

En utilisant le théorème 2.6.1 de la p. 38 le théorème 2.6.3 ci-dessus donne un moyen de savoir, par la seule connaissance du nombre E , si la population va s'éteindre ou se développer :

Théorème 2.6.4 *a) si $E < 1$, alors le nombre d'individus de la population décroît et tend vers 0.*

- b) si $E > 1$, alors le nombre d'individus de la population croît et tend vers l'infini.*

Exemples : a) Si $m = 2$, et $\beta_1=0$, on a $M_f = \begin{pmatrix} 0 & \beta_2 \\ \alpha_1 & 0 \end{pmatrix}$ et $E = \alpha_1\beta_2$. Le polynôme caractéristique de f étant $\lambda^2 - \alpha_1\beta_2$, les valeurs propres de f sont $\pm\sqrt{\alpha_1\beta_2}$. Cela illustre le point c) du théorème 2.6.3.

b) Le cas $E = 1$ ne permet pas de conclusion sans information supplémentaire. Il donne parfois lieu à des comportements cycliques, comme par exemple avec la matrice $M_f = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (les individus ne meurent pas pendant la période considérée et donnent naissance à 1 descendant après une étape). L'évolution ici est cyclique de période 2, puisque $M_f^2 = I$.

c) Si $E > 1$, cas où la population croît, on peut décrire la pyramide d'âges à long terme si $\lambda_{\text{pos}} > |\lambda_i|$ pour toutes les autres racines réelles ou complexes λ_i du polynôme caractéristique P_f . Il existe alors un unique vecteur propre $v_{\text{pos}} = (v_1, \dots, v_m)$ pour λ_{pos} tel que $v_1 = 1$. Vu la forme simple de la matrice M_f , il est facile de trouver v_{pos} . On a

$$\lambda_{\text{pos}}v_2 = \alpha_1v_1 = p_1, \quad \lambda_{\text{pos}}v_3 = \alpha_2v_2 = \frac{\alpha_2\alpha_1v_1}{\lambda_{\text{pos}}} = \frac{p_2}{\lambda_{\text{pos}}}, \quad \text{etc.} \quad (2.32)$$

d'où $v_{\text{pos}} = (1, p_1/\lambda_{\text{pos}}, \dots, p_{m-1}/\lambda_{\text{pos}}^{m-1})$. Par la proposition 2.6.2 de la p. 39, l'état de la population $x(t) := (x_1(t), \dots, x_m(t))$, une fois divisé par $x_1(t)$, tend vers le vecteur propre v_{pos} :

$$\frac{1}{x_1(t)}x(t) \rightarrow (1, \frac{p_1}{\lambda_{\text{pos}}}, \frac{p_2}{\lambda_{\text{pos}}^2}, \dots, \frac{p_{m-1}}{\lambda_{\text{pos}}^{m-1}}) \quad (t \rightarrow \infty) \quad (2.33)$$

et le taux de croissance à long terme est λ_{pos} .

2.6.3 Le système de Fibonacci

C'est un cas particulier de population avec pyramide d'âges (voir § précédent) où l'on a 2 classes d'âge (jeunes,vieux). Un état est donc un couple $s := (x, y) \in \mathbb{R}^2$. La loi d'évolution est par définition $f(x, y) := (x + y, x)$ dont la matrice est

$$M := M_f = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Ce système a été imaginé vers l'an 1200 par le mathématicien italien Fibonacci (1170–1230, aussi connu sous le nom de Leonardo da Pisa) pour modéliser la croissance de populations de lapins.

Pour l'état initial $s(0) := (1, 1)$, la trajectoire est

| n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | ... |
|--------|--------|--------|--------|--------|--------|---------|----------|-----|
| $s(n)$ | (1, 1) | (2, 1) | (3, 2) | (5, 3) | (8, 5) | (13, 8) | (21, 13) | ... |

Le nombre x_n de jeunes individus à l'étape n augmente donc selon la suite

$$1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \quad 21 \quad 34 \quad 55 \quad 89 \quad \dots \quad (2.34)$$

appelée **suite de Fibonacci**³, caractérisée par l'équation $x_n = x_{n-1} + x_{n-2}$. De toute façon, comme $E = 2$, la population tend vers l'infini par le théorème 2.6.4 de la p. 40. Pour l'évolution qualitative (proportion jeunes/vieux), on se réfère à l'exemple c) de la p. 41. Il faut chercher la valeur propre λ_{pos} et le vecteur propre v_{pos} . Le polynôme caractéristique de f étant

$$P_f = \begin{vmatrix} 1-X & 1 \\ 1 & -X \end{vmatrix} = X^2 - X - 1,$$

les valeurs propres de f sont les deux solutions de l'équation $X^2 - X - 1 = 0$. On trouve

$$\lambda_{\text{pos}} := \frac{1 + \sqrt{5}}{2} \approx 1.618033989 \quad \text{et} \quad \lambda_{\text{neg}} := \frac{1 - \sqrt{5}}{2} \approx -.618033989.$$

Ceci illustre bien le théorème 2.6.3 de la p. 40. On découvre que la valeur propre λ_{pos} est le **nombre d'or**, connu depuis l'antiquité, et qui joue un rôle en architecture et en botanique [13, ch.9]. Pour trouver v_{pos} , on résout l'équation matricielle :

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_{\text{pos}} x \\ \lambda_{\text{pos}} y \end{pmatrix}. \quad (2.35)$$

Les solutions sont les couples (x, y) tels que $x = \lambda_{\text{pos}}y$ ce qui forme une droite \mathcal{D}_{pos} . Pour v_{pos} , on a $x = 1$ d'où $v_{\text{pos}} = (1, 1/\lambda_{\text{pos}})$, comme prévu par la formule (2.33). De même, on vérifie que la droite $\mathcal{D}_{\text{neg}} := \{(x, y) \mid x = \lambda_{\text{neg}}y\}$ contient les vecteurs propres pour la valeur propre λ_{neg} . D'après l'exemple c) de la p. 41, l'état $s(t) = (x(t), y(t))$ tendra, lorsque $t \rightarrow \infty$, vers la droite \mathcal{D}_{pos} , c'est-à-dire que

$$\frac{x(t)}{y(t)} \rightarrow \lambda_{\text{pos}} \quad (t \rightarrow \infty)$$

La proportion jeunes/vieux, de même que le taux de croissance, se stabilisera vers le nombre d'or $\frac{1+\sqrt{5}}{2} \approx 1.618033989$. Voici un calcul avec l'état initial $s(0) = (1, 1)$:

³Pourquoi de nombreuses fleurs ont-elles un nombre de pétales appartenant à la suite de Fibonacci ? voir [14] et [13, ch. 9].

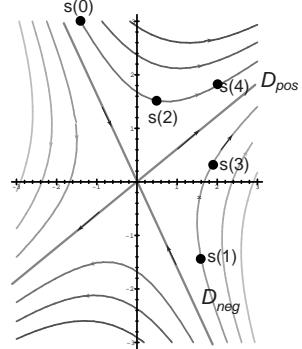
| n | $x(n)/y(n)$ |
|-----|-------------|
| 0 | 1.000000000 |
| 5 | 1.625000000 |
| 10 | 1.617977528 |
| 15 | 1.618034448 |
| 20 | 1.618033985 |
| 25 | 1.618033989 |
| 30 | 1.618033989 |
| 35 | 1.618033989 |
| 40 | 1.618033989 |

Obtenu par le programme MAPLE :

```
s:=(1,1):
f:=(x,y)->(x+y,x): g:=(x,y)->x/y:
for n from 0 to 40 by 5
do [n,evalf((g@(f@@n))(s))]; od;
```

Remarques : 1) le bassin de la droite \mathcal{D}_{pos} , qui est un attracteur, est $\{0\} \cup (\mathbb{R}^2 - \mathcal{D}_{\text{neg}})$. La droite \mathcal{D}_{neg} est un sous-système non-stable qui sépare le bassin de la demi-droite $\mathcal{D}_{\text{pos}} \cap \{x > 0\}$ (trajectoires tendant vers (∞, ∞)) du bassin de la demi-droite $\mathcal{D}_{\text{pos}} \cap \{x < 0\}$ (trajectoires tendant vers $(-\infty, -\infty)$) ; on ne les rencontrera pas pour des problèmes de populations où x et y sont des nombres positifs).

2) la fonction $h(x, y) = x^2 - xy - y^2$ satisfait $h(s) = -h(f(s))$. L'état $s(n)$ oscille donc entre deux courbes de niveau de la fonction h , les courbes $h(s) = h(s(0))$ et $h(s) = -h(s(0))$. Ces courbes de niveau $x^2 - xy - y^2 = \text{constante}$ sont des hyperboles dont les asymptotes sont \mathcal{D}_{pos} et \mathcal{D}_{neg} (figure ci-contre).



3) Si l'état initial $s(0) = (x(0), y(0))$ appartient à la droite \mathcal{D}_{neg} , la trajectoire tendra vers 0. Il est intéressant de noter que ce cas *ne peut pas être observé numériquement* ! En effet, le nombre λ_{neg} est irrationnel car $\sqrt{5}$ n'est pas une fraction (voir annexe A.3). Lorsque $x(0)$ et $y(0)$ sont des valeurs *numériques*, qui sont forcément des fractions (voir p. 93), on aura $x(0)/y(0) \in \mathbb{Q}$ d'où $s(0) \notin \mathcal{D}_{\text{neg}}$ et $s(0)$ sera dans le bassin de \mathcal{D}_{pos} .

Autrement dit : la trajectoire de $s_0 := (\frac{1-\sqrt{5}}{2}, 1)$ tend vers 0 mais tout essai numérique, par exemple avec $\tilde{s}_0 = (-.618033989, 1)$, donne, quelque soit la précision de l'approximation, une trajectoire qui tend vers l'infini !

r

Chapitre 3

Calcul différentiel

Elaboré à partir du XVII^e siècle, le calcul différentiel et intégral est l'un des outils intellectuels les plus puissants que l'homme possède. C'est le langage incontournable de la physique (mécanique, thermodynamique, électromagnétisme, relativité, mécanique quantique, etc). C'est cet outil qui a permis à Maxwell de dégager l'existence des ondes électromagnétiques, si importantes dans notre vie quotidienne et qui n'auraient peut-être jamais été trouvées autrement. On peut donc dire sans exagérer que le calcul infinitésimal a changé la face du monde et est l'un des facteurs fondamentaux de la révolution scientifique et technologique à laquelle nous continuons d'assister.

Ce chapitre présente quelques outils du calcul différentiel à une et plusieurs variables, dans le but de résoudre les problèmes suivants :

- trouver les extrema d'une fonction d'une ou plusieurs variables (optimisation) ;
- décider si un état stationnaire d'un système dynamique est stable.

3.1 Applications différentiables à une variable

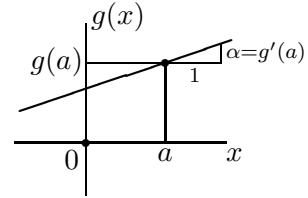
Les étudiants sont supposés avoir déjà rencontré la notion de dérivée d'une application de \mathbb{R} dans \mathbb{R} . Les formules usuelles de dérivation et les dérivées des fonctions courantes se trouvent dans les tables comme [5] ou dans MAPLE. Ces différentes techniques de dérivation seront revues aux exercices.

3.1.1 Différentiabilité

Une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ est dite **affine** si son graphe est une droite. Si $a \in \mathbb{R}$, on a alors

$$g(x) = g(a) + \alpha(x - a). \quad (3.1)$$

Le nombre α est la *pente* de la droite $y = g(x)$. On a aussi que $\alpha = g'(a)$, la dérivée de g en a .



Les fonctions affines sont parmi les plus simples à comprendre et à calculer. Si l'on a une fonction f quelconque définie au voisinage d'un point $a \in \mathbb{R}$, il est donc naturel de chercher à l'approcher par une fonction affine

$$f(x) \approx f(a) + \alpha(x - a) \quad (\text{au voisinage de } a). \quad (3.2)$$

Lorsqu'une "bonne" approximation existe, on dit que f est différentiable (en a). Pour remplacer l'égalité approximative (3.2) par une vraie égalité, il faut introduire un *reste* (ou *erreur*) $R(x)$:

$$f(x) = f(a) + \alpha(x - a) + R(x).$$

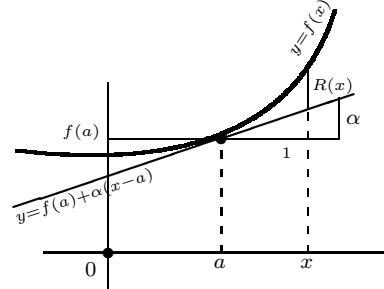
L'approximation affine sera "bonne" si le reste $R(x)$ se développe suffisamment lentement. La définition précise est

Définition :

f est **differentiable en a** s'il existe $\alpha \in \mathbb{R}$ tel que

$$f(x) = f(a) + \alpha(x - a) + R(x) \quad (3.3)$$

$$\text{avec } \frac{R(x)}{x - a} \rightarrow 0 \text{ quand } x \rightarrow a$$



Géométriquement, la différentiabilité de f en a correspond à ce que le graphe de f possède une tangente en a , d'équation $y = f(a) + \alpha(x - a)$.

Le coefficient α est facile à calculer. Quand $x \neq a$, on a

$$\alpha = \frac{f(x) - f(a)}{x - a} - \frac{R(x)}{x - a}. \quad (3.4)$$

Le membre de gauche est une fonction constante ; sa limite lorsque $x \rightarrow a$ est α . Par (3.3), le membre de droite aura une limite lorsque $x \rightarrow a$ si et

seulement si $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$ existe. On sait que cela veut dire que f est **dérivable** en a et la limite sera la dérivée

$$f'(a) := \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}$$

de f en a . On a démontré :

Théorème 3.1.1 *La fonction f est différentiable en a si et seulement si elle est dérivable en a . Dans ce cas, on a*

$$\begin{aligned} f(x) &= f(a) + f'(a)(x - a) + R(x) \\ \text{avec } \frac{R(x)}{x - a} &\rightarrow 0 \text{ quand } x \rightarrow a \end{aligned} \tag{3.5}$$

Remarque : Le théorème (3.1.1) montre que les notions “différentiable” et “dérivable” coïncident. Ce n’est pas le cas pour des fonctions de plusieurs variables (voir p. 55). Dans les deux cas, c’est la différentiabilité et la formule (3.5) qui sont importantes car, comme nous le verrons, elles déterminent le comportement de f au voisinage de a .

3.1.2 Comportement de f au voisinage d’un point

Intuitivement, la différentiabilité de f en a implique que, au voisinage de a , f se comporte à peu près comme la fonction affine $g(x) = f(a) + f'(a)(x - a)$. En effet, le graphe de f vient se “coller” sur celui de g . Un exemple de ce phénomène est la proposition suivante :

Proposition 3.1.2 (lemme de traversement) *Soit f une fonction différentiable en a . Supposons que $f'(a) \neq 0$. Alors le graphe de f traverse en a la droite horizontale $y = f(a)$.*

PREUVE: Supposons que $f'(a) > 0$. Comme f est différentiable en a , on a $f(x) = f(a) + f'(a)(x - a) + R(x)$ avec $\frac{R(x)}{x - a} \rightarrow 0$ quand $x \rightarrow a$. Il existe donc un voisinage U de a tel que $|\frac{R(x)}{x - a}| < f'(a)$ ce qui implique que

$$\frac{f(x) - f(a)}{x - a} = f'(a) + \frac{R(x)}{x - a} > 0. \tag{3.6}$$

Cette dernière inégalité entraîne que

$$\begin{aligned} x < a &\Rightarrow f(x) < f(a) \\ x > a &\Rightarrow f(x) > f(a), \end{aligned}$$

donc le graphe de f a traversé la droite $y = f(a)$ (ici, de bas en haut). Le cas $f'(a) < 0$ se traite de la même manière et la droite $y = f(a)$ sera traversée de haut en bas. \square

Définition : Un point $a \in \mathbb{R}$ est un **minimum local** pour la fonction f s'il existe un voisinage U de a tel que $f(x) > f(a)$ lorsque $a \neq x \in U$. De même, on définit la notion de **maximum local**. Un point qui est un minimum ou un maximum local s'appelle un **extremum local**.

La recherche d'extrema locaux fera l'objet du paragraphe 3.3. Pour l'instant, le lemme de traversement 3.1.2 implique immédiatement le résultat suivant :

Corollaire 3.1.3 Soit f une fonction différentiable en a . Supposons que a soit un extremum local de f . Alors $f'(a) = 0$. \square

Le corollaire 3.1.3 dit que la condition $f'(a) = 0$ est *nécessaire* pour que a soit un extremum local de f . Elle n'est pas *suffisante*, comme le montre la fonction $f(x) = x^3$ qui, bien que $f'(0) = 0$, est strictement croissante.

Voici maintenant quelques théorèmes classiques pour les fonctions différentiables. On dit que $f : [a, b] \rightarrow \mathbb{R}$ est **differentiable** si elle est différentiable en tout point de l'intervalle $[a, b]$.

Théorème 3.1.4 (Théorème des accroissements finis) Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction différentiable. Alors il existe $x \in]a, b[$ tel que

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

PREUVE: On démontre d'abord le cas particulier où $f(a) = f(b)$ (*théorème de Rolle*). On utilise que l'image d'un intervalle fermé par une application continue est un intervalle fermé. L'application f est donc constante ou a des extrema dans $]a, b[$. On applique alors la proposition 3.1.3.

Le cas général s'obtient en appliquant le théorème de Rolle à la fonction

$$h(x) := f(x) - (x - a) \frac{f(b) - f(a)}{b - a}. \quad \square$$

Corollaire 3.1.5 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction différentiable. Si $f'(x) > 0$ pour tout $x \in]a, b[$, alors f est strictement croissante. Si $f'(x) < 0$ pour tout $x \in]a, b[$, alors f est strictement décroissante.

PREUVE: Soient $u, v \in [a, b]$ avec $u < v$. Par le théorème des accroissements finis appliqué à l'intervalle $[u, v]$, il existe $x \in]u, v[$ avec $\frac{f(v) - f(u)}{v - u} = f'(x)$. Si $f'(x) > 0$, cela implique que $f(v) > f(u)$ (et le contraire si $f'(x) < 0$).

Corollaire 3.1.6 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction différentiable. Alors, $f'(x) = 0$ pour tout $x \in]a, b[$ si et seulement si f est constante.

PREUVE: Si $f(x) \neq f(a)$, on aurait, par le théorème des accroissements finis, un point $y \in]a, x[$ avec $f'(y) \neq 0$. \square

Enfin, l'une des fameuses règles de L'Hospital :

Proposition 3.1.7 (Règle de l'Hospital) Soient f et g dérivables sur un intervalle ouvert $]a, a + \epsilon[$ (ou $]a - \epsilon, a[$). On suppose que

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

$$\text{Alors, } \lim_{x \rightarrow a} \frac{f(x)}{g(x)} \text{ existe et } \boxed{\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}}.$$

PREUVE: La démonstration, pour un énoncé aussi général, repose sur le théorème des accroissements finis (voir [1, p. 242-244]). Nous allons donner une démonstration plus simple dans le cas particulier où f et g sont différentiables en a et où $g'(a) \neq 0$. Il s'agit de montrer qu'alors

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

On applique la formule (3.5) à f et g :

$$\frac{f(x)}{g(x)} = \frac{f'(a)(x-a) + R(x)}{g'(a)(x-a) + S(x)} \quad (3.7)$$

avec

$$\lim_{x \rightarrow a} \frac{R(x)}{x-a} = \lim_{x \rightarrow a} \frac{S(x)}{x-a} = 0.$$

En divisant le membre de droite de (3.7) haut et bas par $x-a$, on obtient

$$\frac{f(x)}{g(x)} = \frac{\frac{f'(a)}{x-a} + \frac{R(x)}{x-a}}{\frac{g'(a)}{x-a} + \frac{S(x)}{x-a}} \rightarrow \frac{f'(a)}{g'(a)}. \quad \square$$

Remarques :

1. La règle de l'Hospital est aussi vraie si $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \pm\infty$.
 2. On peut appliquer la règle de l'Hospital dans le cas où $x \rightarrow \pm\infty$.
- Exemple :*

$$\lim_{x \rightarrow \infty} \frac{e^{ax}}{x^n} = \lim_{x \rightarrow \infty} \frac{ae^{ax}}{nx^{n-1}} = \cdots = \lim_{x \rightarrow \infty} \frac{a^n e^{ax}}{n!} = \infty \text{ si } a > 0. \quad (3.8)$$

Cela prouve que la fonction exponentielle e^{ax} avec $a > 0$ croît plus vite que tout polynôme.

3.1.3 Formule de Taylor à une variable

Nous avons vu au paragraphe précédent qu'une fonction différentiable en a admet une bonne approximation par un polynôme du 1^{er} degré

$$f(x) \approx P_1(x) = \alpha_0 + \alpha_1(x - a) \quad (3.9)$$

si $\alpha_0 = f(a)$ et $\alpha_1 = f'(a)$. Observons que

$$P_1(a) = f(a) \quad \text{et} \quad P'_1(a) = f'(a). \quad (3.10)$$

L'approximation est **à l'ordre 1** en ce sens que

$$\lim_{x \rightarrow a} \frac{f(x) - P_1(x)}{x - a} = 0. \quad (3.11)$$

On s'intéresse à des approximations plus fines de f , avec des polynômes P_n de degré $n > 1$. En s'inspirant de (3.10), on cherche un polynôme P_n tel que

$$P_n(a) = f(a) , \quad P'_n(a) = f'(a) , \quad \dots , \quad P_n^{(n)}(a) = f^{(n)}(a) \quad (3.12)$$

où $f^{(n)}$ désigne la $n^{\text{ième}}$ dérivée de f . L'approximation sera alors **à l'ordre n** . En effet :

Proposition 3.1.8 *Soient f et g deux fonctions n -fois dérивables en a . Supposons que $f^{(k)}(a) = g^{(k)}(a)$ pour $0 \leq k \leq n$. Alors*

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

PREUVE: On applique la règle de L'Hospital (Proposition 3.1.7) n fois à l'envers :

$$\begin{aligned} 0 &= \frac{f^{(n)}(a) - g^{(n)}(a)}{1} = \lim_{x \rightarrow a} \frac{f^{(n-1)}(x) - g^{(n-1)}(x)}{x - a} = \\ &= 2 \lim_{x \rightarrow a} \frac{f^{(n-2)}(x) - g^{(n-2)}(x)}{(x - a)^2} = \dots = n! \lim_{x \rightarrow a} \frac{f(x) - g(x)}{(x - a)^n} \quad \square \end{aligned}$$

Reste à trouver le polynôme P_n . On utilise le lemme suivant, dont la démonstration se fait par calcul direct :

Lemme 3.1.9 *Considérons la fonction polynomiale*

$$P(x) := \alpha_0 + \alpha_1(x - a) + \dots + \alpha_n(x - a)^n.$$

Alors, pour tout $k \geq 0$, on a

$$\alpha_k = \frac{1}{k!} P^{(k)}(a)$$

En mettant tous ces résultats ensemble, on obtient la formule de Taylor :

Théorème 3.1.10 (Formule de Taylor) *Soit f une fonction n -fois dérivable en a . Alors*

$$\begin{aligned} f(x) &= f(a) + \frac{f'(a)}{1!}(x - a) + \cdots \\ &\quad \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n(x) \\ (\text{en résumé : } f(x) &= \sum_{k=0}^n \frac{1}{k!} f^{(k)}(a)(x - a)^k + R_n(x)) \end{aligned}$$

avec

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

Remarque :

1. Quand $a = 0$, la formule de Taylor porte parfois le nom de *formule de MacLaurin*.
2. Le reste $R_n(x)$ est, dans la littérature, parfois remplacé par les expressions $o((x - a)^n)$ (petit “O” de $(x - a)^n$) ou, comme dans MAPLE, $O((x - a)^{n+1})$ (grand “O” de $(x - a)^{n+1}$). Ces expressions signifient que

$$\lim_{x \rightarrow a} \frac{R_n(x)}{(x - a)^n} = 0 \quad \text{et} \quad \lim_{x \rightarrow a} \frac{R_n(x)}{(x - a)^{n+1}} = \text{constante.}$$

3. Lorsque f est dérivable $(n + 1)$ fois, le **reste** $R_n(x)$ dans la formule de Taylor est donné par différentes expressions ([1, pp. 251–253], [3, pp. 541–543]). Par exemple :

$$R_n(x) = \frac{1}{n!} \int_a^x (x - t)^n f^{(n+1)}(t) dt. \quad (3.13)$$

On a aussi le **reste de Lagrange** :

$$R_n(x) = \frac{(x - a)^{n+1}}{(n + 1)!} f^{(n+1)}(\xi) \quad (3.14)$$

où ξ est un certain point entre a et x . En conséquence, si l'on sait que $|f^{(n+1)}(t)| \leq M$ pour t entre a et x , on aura, utilisant que $|\xi - a| \leq |x - a|$, l'estimation

$$|R_n(x)| \leq \frac{M}{(n + 1)!} |x - a|^{n+1}. \quad (3.15)$$

Exemple : Le développement de Taylor en 0 à l'ordre 6 pour $f(x) = \sin x$ est

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + R_6(x).$$

On peut utiliser la borne $|f^{(n)}(x)| \leq 1$ puisque $f^{(n)}(x) = \pm \sin x$ ou $\pm \cos x$. Par (3.15), cela donne

$$|R_6(x)| \leq \frac{1}{7!}|x|^7 \leq 0.002|x|^7$$

Lorsque $|x| \leq \pi/4$, on obtient $|R_6(x)| \leq 0.0000366$. Le calcul exact montre qu'en fait $|R_6(\pi/4)| \leq 0.0000363$.

3.2 Applications différentiables à plusieurs variables

3.2.1 Dérivées partielles - Gradient

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Notons, comme à la p. 18

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, \dots, 0), \dots, e_n = (0, 0, \dots, 1)$$

les vecteurs de la base standard de \mathbb{R}^n . En fixant toutes les variables $x_k := a_k$ sauf la $i^{\text{ème}}$ que l'on laisse varier, on obtient une fonction d'une variable $t \mapsto f(a + te_i)$. Si cette fonction est dérivable en $t = a_i$, sa dérivée s'appelle la **dérivée partielle de f en a par rapport à la variable x_i** . Elle se note

$$\frac{\partial f}{\partial x_i}(a) := \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h}$$

Si ces dérivées partielles existent en tout a , on obtient n nouvelles fonctions de \mathbb{R}^n dans \mathbb{R} :

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$$

Exemple : $f(x, y, z) = x^2 e^{yz}$ (on a $n = 3$ et les variables sont notées x, y, z). Alors :

$$\frac{\partial f}{\partial x} = 2xe^{yz}, \quad \frac{\partial f}{\partial y} = x^2ze^{yz}, \quad \frac{\partial f}{\partial z} = x^2ye^{yz}.$$

On voit que les dérivées partielles se calculent comme les dérivées usuelles, les variables que l'on ne dérive pas étant traitées comme des constantes.

Les dérivées partielles d'ordre 2 se définissent en dérivant les dérivées partielles premières. Les notations sont :

$$\frac{\partial^2 f}{\partial x_i^2} := \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right)$$

si l'on dérive deux fois par rapport à la variable x_i et

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)$$

si l'on dérive d'abord par rapport à x_j puis par rapport à x_i .

Dans l'exemple précédent $f(x, y, z) = x^2 e^{yz}$, on a

$$\frac{\partial^2 f}{\partial x^2} = 2e^{yz}, \quad \frac{\partial^2 f}{\partial y^2} = x^2 z^2 e^{yz}, \quad \frac{\partial^2 f}{\partial y \partial z} = x^2 z y e^{yz}, \text{ etc.}$$

A priori on aurait 9 dérivées partielles d'ordre 2. Mais on voit par exemple que

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

Cette indépendance de l'ordre de dérivation n'est pas toujours vraie [1, (4.2) p. 316] mais elle a lieu si, comme dans l'exemple, les dérivées partielles sont des fonctions continues [1, th. 4.3] :

Proposition 3.2.1 *Supposons que les dérivées partielles $\frac{\partial^2 f}{\partial x_i^2}$, $\frac{\partial^2 f}{\partial x_j^2}$ et $\frac{\partial^2 f}{\partial x_i \partial x_j}$ existent dans un voisinage de a et sont continues en a . Alors $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existe et l'on a*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Les n dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en a peuvent être vues comme les composantes d'un vecteur ; ce vecteur s'appelle le **gradient** de f en a :

$$\text{grad } f(a) := \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right). \quad (3.16)$$

L'application $\text{grad } f : a \mapsto \text{grad } f(a)$ est un champ de vecteurs, c'est-à-dire une application de \mathbb{R}^n dans \mathbb{R}^n . Ce champ de vecteurs est appelé le **champ de gradient** ou simplement **gradient** de f . La signification géométrique du gradient sera discutée au § 3.2.6, p. 58. En physique, il s'interprète comme la

force que subit une particule soumise à un potentiel f (exemple : particule chargée dans un potentiel électrique). Voir aussi p. 76

Le cas d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ de n variables à valeur dans \mathbb{R}^p n'est pas vraiment plus difficile puisqu'une telle fonction consiste en p fonctions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)). \quad (3.17)$$

On a donc $(p \times n)$ dérivées partielles $\frac{\partial f_i}{\partial x_j}(a)$ qui constituent une $(p \times n)$ -matrice, la **matrice jacobienne** $Df(a)$ de f en a :

$$Df(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}. \quad (3.18)$$

La i ème ligne de $Df(a)$ est le vecteur $\text{grad } f_i(a)$.

Exemple : Supposons que $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ soit une application linéaire de matrice $M_f := (b_{ij})$. Avec les notations de (3.17), on a $f_i(x_1, \dots, x_n) = \sum_{j=1}^n b_{ij}x_j$ d'où

$$\frac{\partial f_i}{\partial x_j} = b_{ij}.$$

Ceci montre que $Df(a) = M_f$ quelque soit $a \in \mathbb{R}^n$. C'est la généralisation du fait familier à une variable que la dérivée de $f(x) = bx$ est constante égale à b .

3.2.2 Différentiabilité

Nous allons suivre le plan du § 3.1 pour définir la notion de différentiabilité pour les fonctions de n variables (fonctions de \mathbb{R}^n dans \mathbb{R} ou $\mathbb{R}^n \rightarrow \mathbb{R}^p$). On dira qu'une fonction définie au voisinage de $a \in \mathbb{R}^n$ est **differentiable** en a si elle admet une bonne approximation par une fonction affine, les notions de "fonction affine" et de "bonne approximation" restant à préciser.

Une fonction $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est dite **affine** si elle est la somme d'une constante et d'une fonction linéaire :

$$g(x) = c + L(x). \quad (3.19)$$

Si $a = (a_1, \dots, a_n)$ est un point de \mathbb{R}^n , on peut écrire

$$g(x) = c + L(a) + L(x - a) = g(a) + L(x - a). \quad (3.20)$$

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est une fonction non-affine, on aura

$$f(x) = f(a) + L(x - a) + R(x). \quad (3.21)$$

Pour chaque fonction linéaire $L : \mathbb{R}^n \rightarrow \mathbb{R}^p$, l'équation (3.21) définit simplement une nouvelle fonction "reste" : $R : \mathbb{R}^n \rightarrow \mathbb{R}^p$. Le problème est de trouver L pour que cette fonction perturbatrice se développe suffisamment lentement au voisinage de a , c'est-à-dire :

$$\frac{\|R(x)\|}{\|x - a\|} \rightarrow 0 \quad \text{quand } \|x - a\| \rightarrow 0. \quad (3.22)$$

Nous sommes prêts pour la définition d'une application différentiable, analogue à celle des fonctions d'une variable de la p. 45.

Définition :

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ une application et $a \in \mathbb{R}^n$. On dit que f est **differentiable** en a s'il existe une application linéaire $L : \mathbb{R}^n \rightarrow \mathbb{R}^p$ telle que

$$\begin{aligned} f(x) &= f(a) + L(x - a) + R(x) \\ \text{avec } &\frac{\|R(x)\|}{\|x - a\|} \rightarrow 0 \text{ quand } \|x - a\| \rightarrow 0. \end{aligned} \quad (3.23)$$

La matrice $M_L \in \mathcal{M}_{p \times n}(\mathbb{R})$ se calcule de la même façon que pour une fonction d'une variable (p. 46). On regarde le cas $x = (a_1, a_2, \dots, a_i + h, \dots, a_n)$ et on en déduit que le coefficient $(M_L)_{ki}$ est la dérivée partielle

$$(M_L)_{ki} = \frac{\partial f_k}{\partial x_i}(a). \quad (3.24)$$

Autrement dit, $M_L = Df(a)$, la matrice jacobienne de f et

$$L(x - a) = Df(a)(x - a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix}. \quad (3.25)$$

En particulier, lorsque $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ($p = 1$), la matrice ligne $Df(a)$ peut être vue comme le vecteur gradient et on peut utiliser le produit scalaire :

$$L(x - a) = \text{grad } f(a) \cdot (x - a). \quad (3.26)$$

L'existence de dérivées partielles est donc une condition nécessaire pour que f soit différentiable. Contrairement au cas d'une variable (théorème 3.1.1, p. 46), ce n'est pas une condition suffisante [1, §IV.3]. La condition suffisante la plus courante est la continuité des dérivées partielles [1, théorème 3.6, ch. IV]. On peut donc énoncer les analogues du théorème 3.1.1, p. 46 pour les fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}$ et $\mathbb{R}^n \rightarrow \mathbb{R}^p$:

Théorème 3.2.2 *Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $a \in \mathbb{R}^n$. Supposons que les dérivées partielles d'ordre 1 de f existent au voisinage de a et soient continues en a . Alors f est différentiable en a et on a*

$$\begin{aligned} f(x) &= f(a) + \text{grad } f(a) \cdot (x - a) + R(x) \\ \text{avec } &\frac{\|R(x)\|}{\|x - a\|} \rightarrow 0 \text{ quand } \|x - a\| \rightarrow 0. \end{aligned} \tag{3.27}$$

Théorème 3.2.3 *Soit $f := (f_1, \dots, f_p) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $a \in \mathbb{R}^n$. Supposons que les dérivées partielles d'ordre 1 de f_i ($i = 1, \dots, p$) existent au voisinage de a et soient continues en a . Alors f est différentiable en a et on a*

$$\begin{aligned} f(x) &= f(a) + \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \cdots & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix} + R(x) \\ \text{avec } &\frac{\|R(x)\|}{\|x - a\|} \rightarrow 0 \text{ quand } \|x - a\| \rightarrow 0. \end{aligned} \tag{3.28}$$

3.2.3 Dérivée de fonctions composées

Pour deux fonctions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ on connaît la règle de dérivation de la composition $g \circ f$: si f est différentiable en a et g est différentiable en $f(a)$, alors

$$(g \circ f)'(a) = g'(f(a)) f'(a). \tag{3.29}$$

Ce résultat se généralise aux fonctions de plusieurs variables de la manière suivante [1, p. 306]

Proposition 3.2.4 *Soient*

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^p \xrightarrow{g} \mathbb{R}^q$$

des fonctions différentiables. Alors $g \circ f$ est différentiable et la matrice jacobienne $D(g \circ f)$ en $a \in \mathbb{R}^n$ est le produit matriciel

$$D(g \circ f)(a) = Dg(f(a)) \cdot Df(a). \quad (3.30)$$

Exemple : Prenons le cas particulier $n = 1 = q$. On regarde $c : \mathbb{R} \rightarrow \mathbb{R}^p$ comme une courbe différentiable

$$c : t \mapsto c(t) = (c_1(t), \dots, c_p(t)).$$

Le vecteur

$$\dot{c}(t) := (\dot{c}_1(t), \dots, \dot{c}_p(t))$$

s'appelle le **vecteur vitesse** de la courbe c . La formule (3.30) implique que la dérivée de la composition $g \circ c$ est donnée par le produit scalaire

$$(g \circ c)'(t) = \text{grad } g(c(t)) \cdot \dot{c}(t) \quad (3.31)$$

3.2.4 Application : calcul d'erreurs

Supposons que l'on ait n données x_1, \dots, x_n et que x_i est connue avec une imprécision (“erreur”) $\pm \Delta x_i$. Notons $\Delta x := (\Delta x_1, \dots, \Delta x_n) \in \mathbb{R}^n$.

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$. L'erreur Δx produit une erreur Δf sur $f(x_1, \dots, x_n)$. Cette erreur est souvent estimée à l'aide de la formule (3.27) :

$$|\Delta f| \cong |\text{grad } f(x_1, \dots, x_n) \cdot \Delta x| \leq \|\text{grad } f(x_1, \dots, x_n)\| \|\Delta x\|, \quad (3.32)$$

la dernière inégalité étant l'inégalité de Schwarz (p. 16). La formule (3.32) est à appliquer avec précaution. Elle donne des résultats raisonnables lorsque Δx est très petit et que $\text{grad } f$ ne varie pas trop autour de x .

3.2.5 Formule de Taylor à plusieurs variables

Les formules (3.27) et (3.28) du paragraphe précédent sont les développements de Taylor de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ou $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ à l'ordre 1. Le développement de Taylor pour une fonction à n variables donne lieu à des expressions assez compliquées (voir [1, p. 320]). Nous nous bornerons à donner le terme d'ordre 2 pour une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ qui sera utilisé pour les problèmes d'extrema. Introduisons la **matrice hessienne** $\mathcal{H}f(a) \in \mathcal{M}_{n \times n}(\mathbb{R})$ de f en a :

$$\mathcal{H}f(a) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \quad \mathcal{H}f(a)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}. \quad (3.33)$$

Si on se place dans les hypothèses de la proposition 3.2.1, p. 52, la matrice hessienne sera une matrice symétrique.

Pour $x \in \mathbb{R}^n$, considérons le produit matriciel

$$T_2(x) := (x_1 - a_1, \dots, x_n - a_n) \mathcal{H}f(a) \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix}. \quad (3.34)$$

$T_2(x)$ est une (1×1) -matrice que l'on considérera, par abus de notation, comme un nombre réel. C'est le terme d'ordre 2 de la formule de Taylor.

Théorème 3.2.5 *Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $a \in \mathbb{R}^n$. Supposons que les dérivées partielles d'ordre ≤ 2 de f existent au voisinage de a et soient continues en a . Alors, on a la formule*

$$\begin{aligned} f(x) &= f(a) + \text{grad } f(a) \cdot (x - a) + \frac{1}{2} T_2(x) + R_2(x) \\ \text{avec } &\frac{\|R_2(x)\|}{\|x - a\|^2} \rightarrow 0 \text{ quand } \|x - a\| \rightarrow 0. \end{aligned} \quad (3.35)$$

où $T_2(x)$ est défini par l'équation (3.34).

Exemple : $f(x, y) := x + y \sin x$ et $a = (0, 0)$. On a

$$\frac{\partial f}{\partial x} = 1 + y \cos x, \quad \frac{\partial f}{\partial y} = \sin x.$$

D'où

$$\text{grad } f(0) = (1, 0) \quad \text{et} \quad \text{grad } f(0) \cdot (x, y) = x.$$

Les dérivées secondes sont

$$\frac{\partial^2 f}{\partial x^2} = -y \sin x, \quad \frac{\partial^2 f}{\partial y^2} = 0, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \cos x.$$

La matrice hessienne $\mathcal{H}f(0)$ est donc

$$\mathcal{H}f(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

et

$$(x, y) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x, y) \begin{pmatrix} y \\ x \end{pmatrix} = (2xy)$$

d'où $T_2(x, y) = 2xy$. La formule de Taylor à l'ordre 2 donne

$$f(x, y) = x + xy + R_2(x, y).$$

3.2.6 Gradient et surfaces de niveau

Le graphe d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est difficile à visualiser si $n = 2$, malgré les dessins en perspective fournis par les ordinateurs. Cela est impossible si $n \geq 3$. On obtient quand même des informations graphiques en utilisant les courbes ou surfaces de niveau (isothermes, isobares, etc).

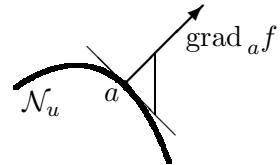
Soit $u \in \mathbb{R}$. La **surface de niveau** u est

$$\mathcal{N}_u := f^{-1}(\{u\}) = \{x \in \mathbb{R}^n \mid f(x) = u\}.$$

Le mot "surface" est utilisé par analogie du cas d'une fonction à 3 variables. Par exemple, si $f(x, y, z) = x^2 + y^2 + z^2$, alors \mathcal{N}_u est la sphère d'équation $x^2 + y^2 + z^2 = u$. Pour une fonction de \mathbb{R}^2 dans \mathbb{R} , ce sont des **courbes de niveau** (des cercles pour l'exemple $f(x, y) = x^2 + y^2$).

Ce paragraphe explique les relations entre le champ de gradient et les surfaces niveau. Les démonstrations peuvent se trouver dans [1, p. 305], [7, ch. 3].

Proposition 3.2.6 *Soit f une fonction différentiable. Le champ de gradient $\text{grad } f$ est orthogonal aux surfaces de niveau. Il pointe dans la direction où f s'accroît le plus rapidement.*



Exemple : Dans l'exemple ci-dessus $f(x, y, z) = x^2 + y^2 + z^2$, on a $\text{grad } f = (2x, 2y, 2z)$. On a donc $\text{grad } f(a) = 2a$. Le gradient est colinéaire avec les rayons et donc bien perpendiculaire aux sphères $x^2 + y^2 + z^2 = u$. Il pointe vers l'extérieur, où f s'accroît. La proposition 3.2.6 dit que la façon d'accroître f le plus rapidement est de s'éloigner radialement de l'origine.

Remarque : Observons que, dans l'exemple ci-dessus, l'ensemble \mathcal{N}_0 n'est pas une surface mais est réduit au point 0. Cette situation dégénérée ne se produit que si $\text{grad } f(a) = 0$. Lorsque $\text{grad } f(a) \neq 0$, on peut démontrer que \mathcal{N}_u est bien une surface au voisinage de a (ou une courbe, si $n = 2$).

La signification de la proposition 3.2.6 peut être précisée de la manière suivante. Soit $v \in \mathbb{R}^n$ avec $|v| = 1$. Considérons la courbe différentiable $c_v : \mathbb{R} \rightarrow \mathbb{R}^n$ définie par

$$c_v(t) := a + tv.$$

Le point $c_v(t)$ s'éloigne radialement de a dans la direction v à vitesse horaire 1 (son vecteur vitesse est v , de norme 1). La dérivée de $f \circ c_v$ en 0 est la **dérivée directionnelle** de f en a dans la direction v :

$$D_v f(a) := (f \circ c_v)'(0) = \lim_{h \rightarrow 0} \frac{f(a + hv) - f(a)}{h}. \quad (3.36)$$

En utilisant la formule (3.31) de la p. 56, on obtient

$$D_v f(a) := \text{grad } f(a) \cdot v = |\text{grad } f(a)| \cos \alpha \quad (3.37)$$

où $\alpha \in [0, \pi]$ est l'angle entre v et $\text{grad } f(a)$ (la dernière égalité utilise que $|v| = 1$).

On voit que

1. $D_v f(a) = 0$ si $\alpha = \pi/2$. C'est que v est tangente à la surface de niveau.
2. $D_v f(a)$ est maximale si $\alpha = 0$, c'est-à-dire si v est colinéaire à $\text{grad } f(a)$ et pointe dans la même direction que lui. C'est la direction de plus forte croissance pour f et on a $D_v f(a) = |\text{grad } f(a)|$.
3. $D_v f(a)$ est minimale si $\alpha = \pi$, c'est-à-dire si v pointe dans la direction opposée à $\text{grad } f(a)$. On a alors $D_v f(a) = -|\text{grad } f(a)|$. C'est la direction de plus forte décroissance pour f .

3.3 Problèmes d'extrema

La recherche d'extrema locaux est très importante en sciences car les états d'équilibre des phénomènes naturels sont souvent ceux qui minimisent ou maximalisent localement une fonction (par exemple l'énergie).

La définition d'un extremum local pour une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est la même que pour une fonction d'une variable (p. 47). Un point $a \in \mathbb{R}^n$ est un **minimum local** pour la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ s'il existe un voisinage U de a dans \mathbb{R}^n tel que $f(x) > f(a)$ lorsque $a \neq x \in U$. De même, on définit la notion de **maximum local**. Un point qui est un minimum ou un maximum local s'appelle un **extremum local**.

3.3.1 Fonctions d'une variable

Le fait que a soit ou non un extremum local dépend du **premier terme non-nul dans la formule de Taylor pour f en a** :

Théorème 3.3.1 Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ et $a \in \mathbb{R}$. Supposons que

$$f'(a) = f''(a) = \dots = f^{(p-1)}(a) = 0 \quad \text{et} \quad f^{(p)}(a) \neq 0.$$

Alors, a est un extremum local si et seulement si p est pair. Dans ce cas, a est un minimum local si $f^{(p)}(a) > 0$ et un maximum local si $f^{(p)}(a) < 0$.

PREUVE: Si $f'(a) = f''(a) = \dots = f^{(p-1)}(a) = 0$, la formule de Taylor à l'ordre p (théorème 3.1.10, p. 50) prend la forme $f(x) = f(a) + \frac{1}{p!} f^{(p)}(a)(x-a)^p + R_p(x)$ avec $\frac{R_p(x)}{(x-a)^p} \rightarrow 0$ quand $x \rightarrow a$.

Voyons tout d'abord le cas p impair. Supposons que $f^{(p)}(a) > 0$. Il existe donc un voisinage U de a tel que $|\frac{R_p(x)}{(x-a)^p}| < f^{(p)}(a)$ ce qui implique que

$$\frac{f(x) - f(a)}{(x-a)^p} = f^{(p)}(a) + \frac{R(x)}{(x-a)^p} > 0, \quad \forall x \in U. \quad (3.38)$$

Puisque p est impair, on en déduit que

$$\begin{aligned} x < a &\Rightarrow f(x) < f(a) \\ x > a &\Rightarrow f(x) > f(a), \end{aligned}$$

Le graphe de f traverse donc, au dessus de a , la droite horizontale $y = f(a)$ et a ne peut donc être un extremum local. Le cas $f^{(p)}(a) < 0$ se traite de la même manière (cette partie de la démonstration est identique à celle du lemme de traversement (p. 46)).

Pour avoir un extremum local, il faut donc que p soit pair. Dans ce cas, si $f^{(p)}(a) > 0$, la formule (3.38) implique que $f(x) > f(a)$ lorsque $x \in U$ et $x \neq a$ (puisque $(x-a)^p > 0$). Le point a est donc un minimum local. Le cas $f^{(p)}(a) < 0$ se traite de la même manière, donnant lieu à un maximum local. \square

Remarque :

1. Le lemme de traversement (p. 46) est un cas particulier du théorème 3.3.1 ($p = 1$). En particulier, on retrouve que si a est un extremum local, alors $f'(a) = 0$.
2. On trouve fréquemment dans la littérature la condition suffisante d'extremum local portant sur la dérivée seconde : $f'(a) = 0$ et $f''(a) \neq 0$. C'est le cas particulier du théorème 3.3.1 où $p = 2$.

3.3.2 Fonctions de plusieurs variables

Tout au long de ce paragraphe, les fonctions sont supposées avoir des dérivées partielles continues jusqu'à l'ordre 2 et donc $\frac{\partial^2}{\partial x_i \partial x_j} = \frac{\partial^2}{\partial x_j \partial x_i}$ (proposition 3.2.1, p. 52). En conséquence, la matrice hessienne de f (p. 56) :

$$(\mathcal{H}f(a))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

est une matrice symétrique.

Nous avons d'abord une condition *nécessaire* pour qu'un point a soit un extremum local.

Proposition 3.3.2 *Si $a \in \mathbb{R}^n$ est un extremum local pour $f : \mathbb{R}^n \rightarrow \mathbb{R}$, alors $\text{grad } f(a) = 0$.*

PREUVE: Si $a = (a_1, \dots, a_n)$ est un extremum local, la fonction $t \mapsto f(t, a_2, \dots, a_n)$ aura un extremum local en $t = a_1$. Sa dérivée sera donc nulle et cette dérivée est $\partial f / \partial x_1(a)$. En pratiquant de même avec les autres variables, on voit que

$$\text{grad } f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) = 0. \quad \square$$

La condition $\text{grad } f(a) = 0$ n'est pas suffisante pour que a soit un extremum local. Les conditions suffisantes les plus simples portent sur les valeurs propres de la matrice hessienne $\mathcal{H}f(a)$. Etant donné les hypothèses indiquées au début de ce paragraphe, $\mathcal{H}f(a)$ est une matrice symétrique et donc ses valeurs propres sont réelles (Proposition 2.5.4 p. 37). On peut démontrer le résultat suivant.

Théorème 3.3.3 *Soit $a \in \mathbb{R}^n$ tel que $\text{grad } f(a) = 0$. Alors*

- 1) *si toutes les valeurs propres de $\mathcal{H}f(a)$ sont > 0 , le point a est un minimum local pour f .*
- 2) *si toutes les valeurs propres de $\mathcal{H}f(a)$ sont < 0 , le point a est un maximum local pour f .*
- 3) *si $\mathcal{H}f(a)$ a des valeurs propres positives et d'autres négatives, le point a n'est pas un extremum local pour f .*
- 4) *si l'on n'est pas dans l'un des cas ci-dessus, on ne peut pas savoir, sans information supplémentaire, si a est un extremum local ou non.*

En utilisant la proposition 2.5.5 p. 37 et le point 1) du théorème ci-dessus, on obtient :

Corollaire 3.3.4 *Soit $a \in \mathbb{R}^n$ tel que $\text{grad } f(a) = 0$. Alors, si tous les mineurs principaux de $\mathcal{H}f(a)$ sont de déterminant > 0 , le point a est un minimum local pour f .*

Exemple 1 : Pour une fonction à deux variables, les cas du théorème 3.3.3 peuvent être illustrés par les exemples suivants ($a = (0, 0)$) :

| $f(x, y)$ | $\mathcal{H}f(0)$ | conclusion |
|--------------|--|---------------------------------|
| $x^2 + y^2$ | $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ | minimum |
| $-x^2 - y^2$ | $\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ | maximum |
| $x^2 - y^2$ | $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ | non-extremum (point “selle”) |

Exemple 2 : $f(x, y, z) = x^2 + y^2 + z^2 + axy + byz$. On a

$$\text{grad } f = (2x + ay, 2y + ax + bz, 2z + by)$$

donc $\text{grad } f(0) = 0$. La matrice hessienne de f en 0 est

$$\mathcal{H}f(0) = \begin{pmatrix} 2 & a & 0 \\ a & 2 & b \\ 0 & b & 2 \end{pmatrix}.$$

Les déterminants des mineurs principaux sont 2, $4 - a^2$ et $8 - 2(a^2 + b^2)$. Le point 0 est donc un minimum si $a^2 + b^2 < 4$. En fait, les valeurs propres de $\mathcal{H}f(0)$ (calculées par Maple) sont

$$2, 2 + \sqrt{a^2 + b^2}, 2 - \sqrt{a^2 + b^2}.$$

Lorsque $a^2 + b^2 > 4$, le point 0 n'est pas un extremum local.

Pour le polynôme du 2^e degré le plus général en x, y, z , :

$$f(x, y, z) = x^2 + y^2 + z^2 + axy + byz + cxz$$

on obtient, pour que 0 soit un minimum local, la condition nettement plus compliquée

$$\det \mathcal{H}f(0) = 8 - 2(a^2 + b^2 + c^2) + 2acb > 0. \quad (3.39)$$

3.3.3 Méthode des moindres carrés - Régression linéaire

Supposons que l'on ait observé des données numériques $(x_1, y_1), \dots, (x_n, y_n)$ et, qu'en théorie, il devrait exister une fonction affine $f(x) = ax + b$ telle que $y_i = f(x_i)$ pour tout i . En pratique, les égalités $y_i = ax_i + b$ n'auront pas exactement lieu à cause des imprécisions ou des facteurs négligés par

la théorie. La **méthode des moindres carrés** ou **Régression linéaire** permet de trouver la fonction affine la plus vraisemblable.

Pour mesurer l'écart entre les y_i et $f(x_i)$, on utilise la fonction :

$$L(a, b) := \sum_{i=1}^n (y_i - f(x_i))^2 = \sum_{i=1}^n (y_i - ax_i - b)^2 \quad (3.40)$$

On a $L(a, b) = 0$ si et seulement si $y_i = ax_i + b$ pour tout i .

On cherche donc a, b tels que $L(a, b)$ soit minimale. Par 3.3.2, p. 61, il faut que $\text{grad } L(a, b) = 0$. Les paramètres a et b cherchés doivent donc satisfaire aux deux équations :

$$\frac{\partial L}{\partial a} = 0 \quad \text{et} \quad \frac{\partial L}{\partial b} = 0. \quad (3.41)$$

Ces équations s'écrivent

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i + b n = \sum_{i=1}^n y_i \end{cases} \quad (3.42)$$

On regarde ceci comme un système de deux équations en les inconnues a et b . Le déterminant du système est

$$n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2.$$

L'inégalité de Schwarz (Proposition 2.1.1, p. 16) pour les vecteurs (x_1, \dots, x_n) et $\mathbf{1} := (1, \dots, 1)$ implique que ce déterminant est non-nul (sauf si les x_i étaient tous égaux...). Le système est donc de Cramer (p. 26) et admet une unique solution :

$$\begin{aligned} a &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \\ b &= \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \end{aligned} \quad (3.43)$$

Observons que $L(a, b) \rightarrow \infty$ quand $|(a, b)| \rightarrow \infty$. Comme il n'y a qu'un point tel que $\text{grad } L = 0$, donc au plus un extremum, c'est forcément un minimum.

On peut toujours modifier l'origine sur l'axe des x et des y de manière que $\sum_{i=1}^n x_i = 0$ et $\sum_{i=1}^n y_i = 0$ (données *centrées*). En posant $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$, les formules (3.43) deviennent simplement

$$b = 0 \quad \text{et} \quad a = \frac{x \cdot y}{\|x\|^2}. \quad (3.44)$$

3.4 Stabilité des états stationnaires d'un système dynamique

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et $a \in \mathbb{R}^n$ tel que $f(a) = a$. Le point a est donc un état stationnaire pour le système dynamique discret (\mathbb{R}^n, f) . Supposons que f soit différentiable en a . La matrice jacobienne $Df(a)$ de f en a détermine une application linéaire de \mathbb{R}^n dans lui-même, d'où un système dynamique linéaire $(\mathbb{R}^n, Df(a))$. Le principe général est le suivant : les conditions de stabilité de l'état stationnaire a du système (\mathbb{R}^n, f) sont *presque* les mêmes que pour l'état 0 dans le système linéaire $(\mathbb{R}^n, Df(a))$. Les conditions de stabilité de 0 pour un système linéaire sont données par le rayon spectral (§ 2.6, p. 37). On peut démontrer :

Théorème 3.4.1 *Soit a un état stationnaire pour le système $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Supposons que f soit différentiable en a . Alors*

1. *si le rayon spectral de $Df(a)$ est < 1 , l'état stationnaire a est stable.*
2. *si le rayon spectral de $Df(a)$ est > 1 , l'état stationnaire a est instable.*

Remarques : 1) Quand le rayon spectral est 1, on ne peut tirer aucune conclusion sans information supplémentaire. C'est la différence avec un système linéaire.

2) Le cas $n = 1$ correspond à la "condition sur la dérivée" vue à la p. 9.

Chapitre 4

Calcul intégral

4.1 Intégrales et primitives

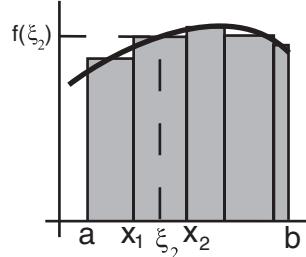
Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction (continue). Son intégrale $\int_a^b f$ est définie de la façon suivante. On partage l'intervalle $[a, b]$ en n sous-intervalles

$$a = x_0 < x_1 < \cdots < x_n = b.$$

On choisit, dans chaque intervalle $[x_i, x_{i+1}]$ un point $\xi_i \in [x_i, x_{i+1}]$. On considère la somme

$$S_n := \sum_{i=1}^n f(\xi_i) \Delta x_i$$

où $\Delta x_i := (x_i - x_{i-1})$. Le nombre S_n mesure (lorsque f est positive) l'aire de la réunion des rectangles de base $[x_i, x_{i+1}]$ et de hauteur $f(\xi_i)$.



Par définition, l'intégrale $\int_a^b f$ de f sur l'intervalle $[a, b]$ est définie par

$$\int_a^b f := \lim S_n , \quad (4.1)$$

la limite étant prise pour $n \rightarrow \infty$ de sorte que le maximum des Δx_i tende vers 0. La notation historique $\int_a^b f(x) dx$ est souvent utilisée pour $\int_a^b f$. Son avantage est d'indiquer par rapport à quelle variable on intègre (si f est une fonction de plusieurs variables).

Lorsque $f(x) \geq 0$, l'intégrale $\int_a^b f$ mesure donc l'aire comprise entre le graphe de f et l'axe horizontal. En général, $\int_a^b f$ représente aussi cette aire à

condition de compter positivement les portions où $f(x) \geq 0$ et négativement celles où $f(x) \leq 0$. En particulier,

$$\text{moyenne}_{[a,b]}(f) := \frac{1}{b-a} \int_a^b f \quad (4.2)$$

est la **valeur moyenne** de f sur l'intervalle $[a, b]$. Si f est continue, il existe $u \in [a, b]$ tel que $f(u) = \text{moyenne}_{[a,b]}(f)$ (une fonction continue prend toutes les valeurs entre ses valeurs extrémales). Ceci démontre le

Théorème 4.1.1 (Théorème de la moyenne) *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors, il existe $u \in [a, b]$ tel que*

$$\int_a^b f = (b-a)f(u).$$

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Une fonction $F : [a, b] \rightarrow \mathbb{R}$ dérivable telle que $F'(x) = f(x)$ s'appelle une **primitive de f** . Par exemple, $F(x) := \frac{1}{2}e^{2x} + 3$ est une primitive de $f(x) = e^{2x}$. Par le corollaire 3.1.6 p. 47, deux primitives F et \bar{F} de f diffèrent par l'addition d'une constante : $\bar{F}(x) = F(x) + C$.

Si $F : [a, b] \rightarrow \mathbb{R}$ est une fonction, on note $F|_a^b := F(b) - F(a)$.

Théorème 4.1.2 (Théorème fondamental du calcul intégral) *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Soit $F : [a, b] \rightarrow \mathbb{R}$ une primitive de f . Alors*

$$\int_a^b f = F|_a^b = F(b) - F(a).$$

PREUVE: On démontre tout d'abord que la fonction

$$G(x) := \int_a^x f$$

est une primitive de f . En effet :

$$G'(x) = \lim_{h \rightarrow 0} \frac{\int_a^{x+h} f - \int_a^x f}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f. \quad (4.3)$$

Par le théorème de la moyenne 4.1.1, on a $\int_x^{x+h} f = h f(u)$ pour un certain $u \in [x, x+h]$, donc

$$G'(x) = \lim_{h \rightarrow 0} f(u) = f(x).$$

Si maintenant F est une primitive de f , on a $G(x) = F(x) + \text{cte}$. Comme $G(a) = \int_a^a f = 0$, on voit que cte $= -F(a)$. On a donc bien

$$\int_a^b f = G(b) = F(b) - F(a). \quad \square$$

NOTATION : la notation $\int f$ ou $\int f(x)dx$ désigne une primitive quelconque de f . Par exemple : $\int \sin(x)dx = -\cos x + C$ où C est une constante.

4.2 Calcul de primitives et d'intégrales

La primitive d'une fonction ne peut pas toujours s'exprimer à l'aide de fonctions que l'on connaît déjà. Par exemple, les primitives

$$\int e^{-x^2}dx , \quad \int \frac{e^x}{x}dx , \quad \int \frac{\sin x}{x}dx , \quad \int \frac{dx}{\ln x} ,$$

sont de nouvelles fonctions qui n'ont pas d'expression plus élémentaire (en utilisant des polynômes, des fonctions trigonométriques, exponentielles, ou des logarithmes). Par exemple, MAPLE appelle $\text{Si}(x)$ la fonction $\int_0^x \frac{\sin t}{t}dt$ qui est une primitive de $\frac{\sin x}{x}$ (voir preuve du théorème 4.1.2). Expérience : contempler le graphe de la fonction Si (`plot(Si(x),x=0..20)`).

De nombreuses fonctions usuelles ont cependant une primitive calculable. Pour la trouver, on peut :

- faire usage d'une table de primitives (exemple : [5, p. 76]).
- utiliser MAPLE. La commande `int(f,x)` donne une primitive de l'expression f tandis que `int(f,x=a..b)` donne $\int_a^b f(x)dx$.

Exemple : `int(sin(x),x=0..Pi/2)` donne la réponse 1.

L'usage de MAPLE est une aide bienvenue mais il n'est pas inutile de confirmer à la main et de réarranger les résultats qui sont parfois mis sous une forme compliquée. Par exemple, si l'on demande à MAPLE de dériver une primitive qu'il a trouvé, il arrive que l'on ne reconnaît pas l'expression de départ.

- décomposer ou transformer la fonction à intégrer en fonctions plus simples. Il y a bien entendu la linéarité de l'intégrale :

$$\int(f+g) = \int f + \int g.$$

Les deux autres méthodes principales sont l'*intégration par substitution* et l'*intégration par partie*.

Intégration par substitution : Supposons que l'on cherche $F(x) = \int f(x)dx$. Si $x = x(u)$ est une fonction de u , la règle de dérivation des

fonctions composées donne

$$\frac{d}{du}(F(x(u))) = \frac{dF}{dx}(x(u)) \frac{dx}{du} = f(x(u)) x'(u).$$

On a donc la formule

$$F(x(u)) = \int f(x(u)) x'(u) du. \quad (4.4)$$

Si maintenant la fonction $u \mapsto x(u)$ est bijective, admettant donc un inverse $x \mapsto u(x)$, on pourra, en remplaçant u par $u(x)$, obtenir $F(x)$ à l'aide de $F(x(u))$. Quant au calcul de $\int_a^b f(x) dx$, il s'obtient dans ce cas par

$$\int_a^b f(x) dx = \int_{u(a)}^{u(b)} f(x(u)) x'(u) du.$$

(4.5)

EXEMPLE : On veut calculer $\int_{-r}^r \sqrt{r^2 - x^2} dx$, ce qui est l'aire d'un demi-disque de rayon r .

On cherche donc $F(x) := \int \sqrt{r^2 - x^2} dx$. Le fait que $r^2 - r^2 \sin^2 x = r^2 \cos^2 x$ suggère la substitution $x(u) = r \sin u$. On aura donc :

$$F(x(u)) = \int r \cos u \sqrt{r^2 - r^2 \sin^2 u} du = r^2 \int \cos^2 u du. \quad (4.6)$$

La primitive de $\cos^2 u$ se trouve dans les tables :

$$\int \cos^2 u du = \frac{1}{2} (u + \cos u \sin u).$$

(On peut calculer $\int \cos^2 u du$ et $\int \sin^2 u du$ en utilisant que $\cos^2 u + \sin^2 u = 1$ et $\cos^2 u - \sin^2 u = \cos(2u)$). La fonction $u \mapsto r \sin u$ est une bijection de $[-\pi/2, \pi/2]$ sur $[-r, r]$, d'inverse $u(x) = \arcsin(\frac{x}{r})$. En substituant dans 4.6, on obtient

$$\int \sqrt{r^2 - x^2} dx = \frac{r^2}{2} \arcsin\left(\frac{x}{r}\right) + \frac{x}{2} \sqrt{r^2 - x^2}. \quad (4.7)$$

L'intégrale cherchée peut donc se calculer de deux façons. La première à l'aide de la primitive que l'on vient de trouver :

$$\int_{-r}^r \sqrt{r^2 - x^2} dx = \left(\frac{r^2}{2} \arcsin\left(\frac{x}{r}\right) + \frac{x}{2} \sqrt{r^2 - x^2} \right) \Big|_{-r}^r = \frac{1}{2} \pi r^2,$$

la seconde en remplaçant les limites d'intégration comme dans (4.5) :

$$\int_{-r}^r \sqrt{r^2 - x^2} dx = r^2 \int_{\arcsin(-1)}^{\arcsin 1} \cos^2 u du = \frac{r^2}{2} (u + \cos u \sin u) \Big|_{-\pi/2}^{\pi/2} = \frac{1}{2} \pi r^2.$$

On a ainsi démontré que l'aire d'un disque de rayon r est πr^2 .

Intégration par parties : La formule de dérivation d'un produit $(fg)' = f'g + fg'$ donne les formules d'**intégration par partie** :

$$\int f'(x) g(x) dx = f(x) g(x) - \int f(x) g'(x) dx \quad (4.8)$$

et

$$\int_a^b f'(x) g(x) dx = f(x) g(x) \Big|_a^b - \int_a^b f(x) g'(x) dx. \quad (4.9)$$

EXEMPLE : Pour calculer $\int x^n \ln x dx$, on pose $f'(x) = x^n$ et $g(x) = \ln x$. La formule (4.9) donne

$$\int x^n \ln x dx = \frac{x^{n+1}}{n+1} \ln x - \frac{1}{n+1} \int x^n dx = \frac{x^{n+1}}{n+1} \ln x - \frac{x^{n+1}}{(n+1)^2} + C.$$

Cette dernière formule est vraie pour $n \neq -1$ (on a $\int \frac{\ln x}{x} dx = \frac{1}{2}(\ln x)^2$). En particulier, si $n = 0$:

$$\int \ln x dx = x \ln x - x + C.$$

Intégrales improches : Ce sont des intégrales du type

$$\int_a^\infty f \quad , \quad \int_{-\infty}^b f \quad , \quad \int_{-\infty}^\infty f \quad ,$$

etc. Elles sont définies comme des limites, par exemple

$$\int_a^\infty f := \lim_{b \rightarrow \infty} \int_a^b f \quad ,$$

etc, lorsque ces limites existent. Exemple, si $n > 1$:

$$\int_1^\infty \frac{dx}{x^n} = \lim_{b \rightarrow \infty} \frac{-1}{(n-1)(x^{n-1})} \Big|_1^b = \frac{1}{n-1} \quad , \quad (n > 1).$$

En revanche $\int_1^b \frac{dx}{x} = \ln b \rightarrow \infty$ ($b \rightarrow \infty$).

Intégration numérique : S'il n'est pas possible de connaître une primitive de f , on peut calculer des valeurs approchées de $\int_a^b f$. Voici quelques méthodes possibles :

- la *méthode des chimistes* : on dessine le graphe Γ_f de f , on découpe la portion entre Γ_f et l'axe horizontal (dont l'aire est $\int_a^b f$) et on pèse le papier.
- en utilisant MAPLE, par la commande `evalf(int(f, x=a..b))`;
- en estimant $\int_a^b g$ où g est une fonction proche de f dont l'intégrale est facile à calculer. Le plus souvent, g est obtenue par interpolation entre des points $x_0 = a < x_1 < \dots < x_n = b$ où l'on connaît la valeur de la fonction. Par exemple, si $x_i - x_{i-1} = \frac{b-a}{n}$ (points également répartis), l'interpolation linéaire donne la **formule des trapèzes** (voir figure (4.2)) :

$$\int_a^b f(x) dx \approx \frac{b-a}{2n} [f(a) + f(b) + 2 \sum_{i=1}^{n-1} f(x_i)].$$

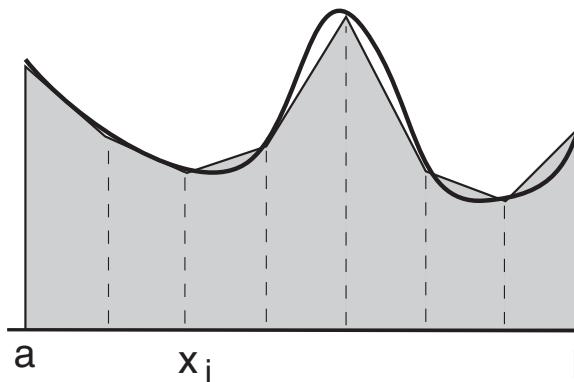


FIGURE (4.2)
Méthode des trapèzes.

Une meilleure interpolation (quadratique) donne la **formule de Simpson**. Ces formules, avec estimation de l'erreur maximale, sont dans les tables ([5, p. 90]). Voir aussi [1, pp. 129–130].

L'avantage de ces méthodes est de ne nécessiter qu'une quantité finie d'information numérique sur la fonction f (sa valeur en un certain nombre de points). En science expérimentale, il se peut que ce soit la seule chose dont on dispose.

4.3 Intégrales de surface et de volume

Soit Σ une surface dans l'espace et soit $f : \Sigma \rightarrow \mathbb{R}$ une fonction continue. Son intégrale $\int_{\Sigma} f$ est définie de manière analogue à l'intégrale d'une fonction sur un intervalle. On partage Σ en une union de domaines A_i ($i = 1 \dots, n$), qui sont disjoints ou ne se rencontrent que sur leur frontière. On note ΔA_i l'aire du domaine A_i et on définit :

$$\int_{\Sigma} f := \lim \sum_{i=1}^n f(\xi_i) \Delta A_i$$

où $\xi_i \in A_i$ et la limite est prise sur des partages de plus en plus fins de manière que $\max \{\text{diamètre}(A_i)\} \rightarrow 0$.

Si Σ est un domaine du plan, le graphe Γ_f de f est une surface dans \mathbb{R}^3 au dessus de Σ . Le nombre $\int_{\Sigma} f$ est alors le volume du solide compris entre Σ et Γ_f (compté positivement pour les parties où $f(x) > 0$ et négativement autrement).

Comme à une variable (voir (4.2), p. 66), on a que

$$\text{moyenne } (f) := \frac{1}{\text{Aire } (\Sigma)} \int_{\Sigma} f$$

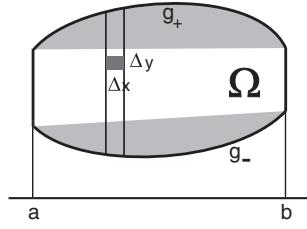
est la **valeur moyenne de f sur Σ** .

Calcul d'une intégrale de surface :

Considérons tout d'abord le cas où $\Sigma = \Omega$ est un domaine du plan délimité par

$$a \leq x \leq b \quad \text{et} \quad g_-(x) \leq y \leq g_+(x)$$

où g_- et g_+ sont deux fonctions dérivables de $[a, b]$ dans \mathbb{R} (voir figure ci-contre).



Intuitivement, le calcul de $\int_{\Omega} f$ repose sur le raisonnement suivant : on promène dans Ω un petit pavé d'aire $\Delta x \Delta y$. On fixe d'abord x et on fait tendre Δy vers 0. On obtient une fonction $H(x) = \int_{g_-(x)}^{g_+(x)} f(x, y) dy$. En laissant ensuite varier x de a à b et en faisant tendre Δx vers 0, l'intégrale $\int_{\Omega} f$ cherchée sera $\int_a^b H(x) dx$. Ceci permet de montrer que $\int_{\Omega} f$ s'obtient par deux intégrales successives :

$$\int_{\Omega} f = \int_a^b \left(\int_{g_-(x)}^{g_+(x)} f(x, y) dy \right) dx$$

(4.10)

Les parenthèses sont souvent omises et on écrit simplement

$$\int_{\Omega} f = \int_a^b \int_{g_-(x)}^{g_+(x)} f(x, y) dy dx.$$

EXEMPLE : Soit $\Omega := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}$ le disque de rayon r centré en 0. Calculons $\int_{\Omega} f$, où $f(x, y) := \sqrt{r^2 - x^2 - y^2}$, ce qui donnera le volume d'une demi-boule de rayon r dans \mathbb{R}^3 .

Le domaine Ω peut être décrit par

$$-r \leq x \leq r \quad \text{et} \quad g_-(x) \leq y \leq g_+(x)$$

avec $g_{\pm}(x) = \pm\sqrt{r^2 - x^2}$. Comme pour (4.7) p. 68 on a la primitive

$$\int \sqrt{r^2 - x^2 - y^2} dy = \frac{r^2 - x^2}{2} \arcsin\left(\frac{y}{\sqrt{r^2 - x^2}}\right) + \frac{y}{2} \sqrt{r^2 - x^2 - y^2}$$

d'où

$$\int_{g_-(x)}^{g_+(x)} f(x, y) dy = \frac{r^2 - x^2}{2} (\arcsin(1) - \arcsin(-1)) = \frac{\pi}{2} (r^2 - x^2).$$

Le calcul de $\int_{\Omega} f$ donne, par (4.10)

$$\int_{\Omega} f = \int_{-r}^r \int_{g_-(x)}^{g_+(x)} f(x, y) dy dx = \frac{\pi}{2} \int_{-r}^r (r^2 - x^2) dx = \frac{\pi r^3}{2} \left(2 - \frac{2}{3}\right) = \frac{2}{3} \pi r^3.$$

On ainsi démontré que le volume d'une boule de rayon r est $\frac{4}{3}\pi r^3$.

Dans le cas général, pour le calcul de $\int_{\Sigma} f$, on paramétrise la surface Σ par un domaine du plan Ω :

$$\Sigma = \{\alpha(u, v) \mid (u, v) \in \Omega\}$$

où $\alpha(u, v) = (\alpha_1(u, v), \alpha_2(u, v), \alpha_3(u, v))$ est une application différentiable injective de Ω dans \mathbb{R}^3 (dont l'image est Σ). On définit

- les u -courbes : $\alpha(u, v_0)$ (v_0 est fixe et u varie)

- les v -courbes : $\alpha(u_0, v)$ (u_0 est fixe et v varie)

qui sont des courbes sur Σ . Leurs vecteurs vitesses sont

$$\alpha_u := \left(\frac{\partial \alpha_1}{\partial u}, \frac{\partial \alpha_2}{\partial u}, \frac{\partial \alpha_3}{\partial u} \right) \quad \text{et} \quad \alpha_v := \left(\frac{\partial \alpha_1}{\partial v}, \frac{\partial \alpha_2}{\partial v}, \frac{\partial \alpha_3}{\partial v} \right).$$

On définit les fonctions $E, F, G : \Omega \rightarrow \mathbb{R}$ par les produits scalaires :

$$E := \alpha_u \cdot \alpha_u \quad , \quad F := \alpha_u \cdot \alpha_v \quad , \quad G := \alpha_v \cdot \alpha_v. \quad (4.11)$$

L'idée du calcul de $\int_{\Sigma} f$ est d'utiliser comme élément d'aire le parallélogramme engendré par les vecteurs vitesse α_u et α_v . Si θ est l'angle entre ces deux vecteurs, le carré de cette aire vaut

$$\|\alpha_u\|^2 \|\alpha_v\|^2 \sin^2 \theta = \|\alpha_u\|^2 \|\alpha_v\|^2 (1 - \cos^2 \theta) = EG - F^2.$$

Ceci permet de remplacer $\int_{\Sigma} f$ par une intégrale sur Ω que l'on peut calculer par (4.10) :

$$\boxed{\int_{\Sigma} f = \int_{\Omega} (f \circ \alpha) \sqrt{EG - F^2}} \quad (4.12)$$

En particulier, si l'on intègre sur Σ la fonction constante égale à 1, on obtient l'aire de Σ :

$$\boxed{\text{Aire}(\Sigma) = \int_{\Omega} \sqrt{EG - F^2}} \quad (4.13)$$

EXEMPLE : 1. On veut calculer l'aire de la sphère Σ_r de rayon r centrée en 0. On peut paramétriser cette sphère par des coordonnées sphériques :

$$\alpha(u, v) = (r \cos v \cos u, r \cos v \sin u, r \sin v)$$

avec $\Omega := [0, 2\pi] \times [-\pi/2, \pi/2]$. Les u -courbes sont les parallèles et les v -courbes les méridiens de la sphère. Leur vecteurs vitesse sont

$$\alpha_u(u, v) = (-r \cos v \sin u, r \cos v \cos u, 0)$$

et

$$\alpha_v(u, v) = (-r \sin v \cos u, -r \sin v \sin u, r \cos v).$$

Les fonctions E, F, G définie en (4.11) sont donc

$$E(u, v) = r^2 \cos^2 v \quad , \quad F(u, v) = 0 \quad , \quad G(u, v) = r^2$$

d'où $EG - F^2 = r^4 \cos^2 v$. Par (4.13) et (4.10), l'aire de Σ_r vaudra

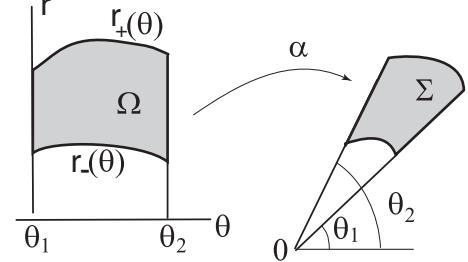
$$\text{Aire}(\Sigma_r) = r^2 \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos v \, du = r^2 \int_0^{2\pi} \sin v \Big|_{-\pi/2}^{\pi/2} \, du = 2r^2 \int_0^{2\pi} \, du = 4\pi r^2.$$

2. *Coordonnées polaires* : la surface Σ peut aussi être un domaine du plan auquel cas l'équation (4.12) fait office de formule de changement de variables. Par exemple, supposons que Σ est le domaine du plan limité en coordonnées polaires r, θ par

$$\theta_1 \leq \theta \leq \theta_2 \quad \text{et} \quad r_-(\theta) \leq r(\theta) \leq r_+(\theta). \quad (4.14)$$

Considérons le plan en coordonnées cartésiennes avec θ en abscisse et r en ordonnées. Soit Ω le domaine donné par les conditions (4.14). Le domaine Σ est alors paramétré par $\alpha : \Omega \rightarrow \Sigma$ où $\alpha(\theta, r) = (r \cos \theta, r \sin \theta)$. On a

$$\alpha_\theta = (-r \sin \theta, r \cos \theta) \quad \text{et} \quad \alpha_r = (\cos \theta, \sin \theta).$$



On en déduit que $E = r^2$, $F = 0$, $G = 1$ et donc $\sqrt{EG - F^2} = r$. La formule (4.12) donne donc

$$\int_{\Sigma} f = \int_{\theta_1}^{\theta_2} \int_{r_-(\theta)}^{r_+(\theta)} r f(\theta, r) dr d\theta. \quad (4.15)$$

Intégrales de volume : Soit V un solide dans l'espace et soit $f : V \rightarrow \mathbb{R}$ une fonction continue. Son intégrale $\int_V f$ est définie comme pour l'intégrale d'une fonction sur une surface. On partage V en une union de petits solides V_i ($i = 1, \dots, n$), qui sont disjoints ou ne se rencontrent que sur leur frontière. On note ΔV_i le volume de V_i et on définit :

$$\int_V f := \lim \sum_{i=1}^n f(\xi_i) \Delta V_i$$

où $\xi_i \in V_i$ et la limite est prise sur des partages de plus en plus fins de manière que $\max \{\text{diamètre}(V_i)\} \rightarrow 0$.

Exemples :

1. Comme pour les intervalles et les surfaces, l'expression

$$\text{moyenne}_V(f) := \frac{1}{\text{vol } V} \int_V f$$

est la valeur moyenne de la fonction f sur V . On calcule ainsi des pressions moyennes, des températures moyennes, etc.

2. Si $f(x)$ est la densité du solide en x , $\int_V f$ est la masse totale de V .
3. Si $f(x)$ est la densité du solide en x , les coordonnées (x_0, y_0, z_0) du centre de gravité de V sont

$$x_0 = \frac{1}{\text{vol } V} \int_V x f \quad , \quad y_0 = \frac{1}{\text{vol } V} \int_V y f \quad , \quad z_0 = \frac{1}{\text{vol } V} \int_V z f. \quad (4.16)$$

Calcul de $\int_V f$: Si le solide V est limité par

$$a \leq x \leq b \quad , \quad g_-(x) \leq y \leq g_+(x) \quad , \quad h_-(x, y) \leq z \leq h_+(x, y)$$

le calcul de $\int_V f$ peut s'effectuer par trois intégrales successives

$$\int_V f = \int_a^b \left[\int_{g_-(x)}^{g_+(x)} \left(\int_{h_-(x, y)}^{h_+(x, y)} f(x, y, z) dz \right) dy \right] dx. \quad (4.17)$$

(On ommet souvent les parenthèses et les crochets). Les changements de coordonnées se font de la manière suivante. Pour les coordonnées sphériques

$$x = r \cos \varphi \cos \theta \quad , \quad y = r \cos \varphi \sin \theta \quad , \quad z = r \sin \varphi$$

l'élément de volume $dz dy dx$ doit être remplacé par $r^2 \cos \varphi dr d\varphi d\theta$. pour les coordonnées cylindriques

$$x = r \cos \theta \quad , \quad y = r \sin \theta \quad , \quad z = z$$

on remplace $dx dy dz$ par $rd\theta dz$.

EXEMPLE : Le volume de la boule B_R centrée en 0 et de rayon R est l'intégrale sur B_R de la fonction constante égale à 1. En coordonnées sphériques :

$$\begin{aligned} \int_{B_R} 1 &= \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \int_0^R r^2 \cos \varphi dr d\varphi d\theta = \\ &= \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \frac{R^3}{3} \cos \varphi d\varphi d\theta = \int_0^{2\pi} \frac{2R^3}{3} d\theta = \frac{4\pi R^3}{3}. \end{aligned}$$

4.4 Quelques applications

Longueur de courbe : Soit $c : [a, b] \in \mathbb{R}^n$ une application différentiable paramétrant une courbe Γ dans \mathbb{R}^n . Pendant l'intervalle de temps $[t, t + \Delta t]$, le point de $c(t)$ parcourt un arc dont la longueur est approximativement

$\|\dot{c}(t)\| \Delta t$, où $\|\dot{c}(t)\|$ est la norme du vecteur vitesse $\dot{c}(t)$. La longueur totale de Γ sera ainsi donnée par l'intégrale

$$\text{longueur}(\Gamma) = \int_a^b \|\dot{c}(t)\| dt. \quad (4.18)$$

EXEMPLE :

1. On paramétrise le cercle de rayon r centré à l'origine dans le plan par $c(t) := (r \cos t, r \sin t)$ ($t \in [0, 2\pi]$). On a donc $\dot{c}(t) = (-r \sin t, r \cos t)$ et $\|\dot{c}(t)\|$ est constante égale à r . On retrouve que la longueur du cercle est $\int_0^{2\pi} r dt = 2\pi r$ (ce qui est rassurant car c'est la définition de π).
2. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction dérivable. On paramétrise le graphe Γ_f dans \mathbb{R}^2 par $c(t) = (t, f(t))$ ($t \in [a, b]$). On a donc $\dot{c}(t) = (1, f'(t))$ et la longueur de Γ_f sera

$$\text{longueur}(\Gamma_f) = \int_a^b \sqrt{1 + f'(t)^2} dt. \quad (4.19)$$

Travail : Soit $c : [a, b] \in \mathbb{R}^n$ une application différentiable paramétrant une courbe Γ dans \mathbb{R}^n . On suppose que le point $c(t)$ est soumis à un champ de forces A (champ de vecteurs $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$). Pendant l'intervalle de temps $[t, t + \Delta t]$ s'effectue un **travail** valant approximativement $(A(c(t)) \cdot \dot{c}(t))\Delta t$. Le travail total T effectué par le point $c(t)$ sur la courbe Γ est ainsi donnée par l'intégrale

$$T = \int_a^b (A(c(t)) \cdot \dot{c}(t)) dt.$$

Un cas particulier important est celui où la champ A est le champ de gradient $A = \text{grad } U$ d'une fonction $U : \mathbb{R}^n \rightarrow \mathbb{R}$ (potentiel). C'est, par exemple, le cas pour un champ gravitationnel ou un champ électrique. Le travail T effectué sur la courbe Γ est alors, utilisant la formule (3.24) de la p. 51 :

$$T = \int_a^b (\text{grad } U(c(t)) \cdot \dot{c}(t)) dt = \int_a^b (U \circ c)'(t) dt = U(c(b)) - U(c(a)).$$

On voit que T est la **différence de potentiel** entre les point extrémités $c(a)$ et $c(b)$ de Γ . En particulier, si la courbe Γ est fermée ($c(a) = c(b)$), le travail T sera nul.

Réciproquement, si un champ de forces A a la propriété que le travail sur toute courbe fermée est nul alors il existe un potentiel U , unique à une constante près, tel que $A = \text{grad } U$: on choisit un point $B \in \mathbb{R}^n$ et on définit $U(x)$ comme le travail effectué sur une courbe quelconque Γ joignant B à x .

Débit : On considère un champ de vecteurs A au voisinage d'une surface Σ dans \mathbb{R}^3 . On interprète A comme le champ de vitesse des particules d'un fluide. On veut calculer le **débit** D_A du fluide passant au travers de Σ (mesuré en unités de volume par unité de temps).

Pour cela, on munit Σ d'un champ de vecteurs $N : \Sigma \rightarrow \mathbb{R}^3$ variant continûment tel que $\|N(x)\| = 1$ et $N(x)$ est orthogonal au plan tangent à Σ en x . Par exemple, si Σ est une surface de niveau d'une fonction $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ (voir paragraphe 3.2.6 p. 55), on peut prendre pour N le gradient normé : $N(x) = \text{grad } f(x)/\|\text{grad } f(x)\|$. Le débit D_A au travers de Σ , dans le sens de N , est alors donné par l'intégrale de surface

$$D_A = \int_{\Sigma} A \cdot N \quad (4.20)$$

Le produit scalaire, dans la formule 4.20, s'impose car, en ce qui concerne le débit, seule compte la composante de N orthogonale à Σ ($A \cdot N$ est, au signe près, la norme de cette composante).

Intégrale d'une gaussienne : La fonction gaussienne $f(x) = e^{-x^2}$ est utile pour le calcul des probabilités. Bien que sa primitive n'ait pas d'expression élémentaire, on peut calculer son intégrale $A := \int_{-\infty}^{\infty} e^{-x^2} dx$ en faisant intervenir une intégrale double.

$$\begin{aligned} A^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) = \int_{-\infty}^{\infty} e^{-x^2} \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) dx = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dy dx = \int_{\mathbb{R}^2} e^{-(x^2+y^2)}. \end{aligned}$$

L'expression $x^2 + y^2$ suggère de passer en coordonnées polaires pour le calcul de l'intégrale de surface $\int_{\mathbb{R}^2} e^{-(x^2+y^2)}$. Avec la formule (4.15) p. 74 :

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)} = \int_0^{2\pi} \int_0^{\infty} r e^{-r^2} dr d\theta = \int_0^{2\pi} \underbrace{\left(-\frac{e^{-r^2}}{2} \Big|_0^{\infty} \right)}_{= 1/2} d\theta = \pi$$

ce qui montre la formule classique

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}. \quad (4.21)$$

Chapitre 5

Systèmes dynamiques continus – Équations différentielles

Ce cours a commencé par les systèmes dynamiques discrets, qui modélisent l'évolution d'un phénomène où le temps varie par saut d'une unité. Dans ce chapitre, on introduit les systèmes dynamiques continus où le temps varie continuûment ($t \in \mathbb{R}$).

Les systèmes dynamiques discrets se sont développés depuis une vingtaine d'années. On peut parler sur leur avenir car ils constituent la base des simulations sur ordinateur. Leur développement influence aujourd'hui la pensée scientifique.

Beaucoup plus anciens, les systèmes continus sont rattachés à la notion d'équation différentielle et leur origine remonte aux débuts du calcul infinitésimal et de la science moderne. Ils ont profondément influencé le développement des mathématiques et de la physique et gardent une importance primordiale pour la formalisation théorique. En revanche, ils ne sont pas bien adaptés au calcul numérique : les méthodes numériques pour résoudre des équations différentielles (voir [1, §II.9]) sont en fait des approximations de systèmes dynamiques continus par des systèmes discrets.

5.1 Champs de vecteurs - trajectoires.

Soit E un domaine de \mathbb{R}^n . Un **champ de vecteurs** dans E est une application différentiable $A : E \rightarrow \mathbb{R}^n$. On le visualise en dessinant le vecteur $A(x)$ partant du point x . Nous avons déjà rencontré des champs de vecteurs : le gradient $\text{grad } f$ d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (voir p. 52 et p. 58).

Une **trajectoire** d'un champ de vecteur A est une courbe différentiable

$s : J \rightarrow E$ définie sur un intervalle $J \subset \mathbb{R}$ telle que son vecteur vitesse $\dot{s}(t)$ soit en tout temps donné par le champ de vecteurs A :

$$\dot{s}(t) = A(s(t)) \quad \forall t \in J. \quad (5.1)$$

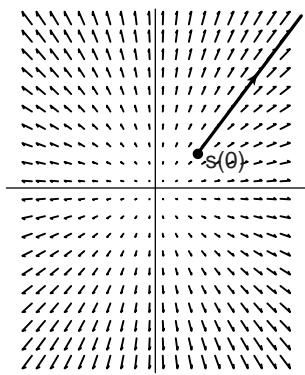


FIGURE (5.2)
le champ $A(x, y) = (2x, 2y)$
et l'une de ses trajectoires.

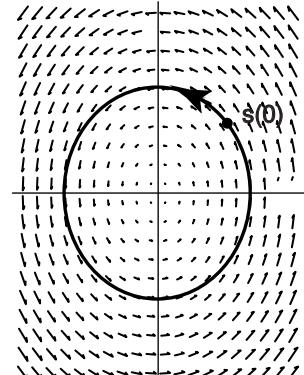


FIGURE (5.3)
le champ $A(x, y) = (-2y, 2x)$
et l'une de ses trajectoires.

Exemples :

1. Le champ $A := \text{grad } f$ sur \mathbb{R}^2 où $f(x, y) := x^2 + y^2$. On a donc $A(x, y) = (2x, 2y)$. La trajectoire de A telle que $s(0) = (x_0, y_0)$ est $s(t) = (x_0 e^{2t}, y_0 e^{2t})$. Voir figure (5. 2).
2. Le champ $A(x, y) := (-2y, 2x)$ sur \mathbb{R}^2 de la figure (5. 3). La trajectoire de A telle que $s(0) = (x_0, y_0)$ est

$$s(t) = (x_0 \cos 2t - y_0 \sin 2t, x_0 \sin 2t + y_0 \cos 2t).$$

3. Le champ $A(x) := 1 + x^2$ sur \mathbb{R} . La trajectoire $x(t)$ de A telle que $x(0) = x_0$ est $x(t) = \operatorname{tg}(t + \arctg(x_0))$.

Dans les exemples ci-dessus, nous parlons de “la” trajectoire de A telle que $s(0) = s_0$. On peut en effet démontrer que *si A est différentiable et si $s_0 \in E$ et $t_0 \in \mathbb{R}$, il existe une unique trajectoire, définie sur un intervalle maximal J autour de t_0 , telle que $s(t_0) = s_0$* . L’intervalle J est maximal en ce sens que la trajectoire ne peut pas, en général, être prolongée au delà de J parce qu’elle sortirait de E . Ceci peut se produire même si $E = \mathbb{R}^n$. On voit en effet que, pour le champ $A(x) := 1 + x^2$ sur \mathbb{R} de l’exemple 3 ci-dessus, la trajectoire $x(t) = \operatorname{tg}(t)$ telle que $x(0) = 0$ s’en va à l’infini lorsque $t \rightarrow \pm\pi/2$.

Systèmes dynamiques continus : La modélisation de l'évolution d'un phénomène en temps continu se fait de la manière suivante. Un **système dynamique continu** est une paire (E, A) où

1. E est un domaine de \mathbb{R}^n (l'ensemble des **états** du système).
2. $A : E \rightarrow \mathbb{R}^n$ est un champ de vecteurs (**dynamique** ou **loi d'évolution** du système).

L'évolution du système s'exprime de la manière suivante : *si le système est dans un état $s(t_0) = s_0$ au temps t_0 , son état au temps t sera $s(t)$, où s est la trajectoire du champs de vecteurs A telle que $s(t_0) = s_0$.* Comme dans le cas des systèmes dynamiques discrets, on parlera de la **trajectoire** de l'état s_0 . On peut y penser comme la trajectoire d'une molécule d'un fluide dont le champ de vitesses est A .

5.2 Trajectoires et équations différentielles

Une **équation différentielle** est une équation faisant intervenir une fonction $x(t)$, certaines de ses dérivées \dot{x} , \ddot{x} , etc et éventuellement la variable t . Exemple : $\dot{x} = 2x + t$, $\ddot{x} = -x$, etc. L'**ordre** de l'équation différentielle est l'ordre de la plus grande dérivée de x qui y figure. Par exemple, $\dot{x} = 2x + t$ est d'ordre 1, $\ddot{x} = -x$ d'ordre 2, etc.

Soit $A : E \rightarrow \mathbb{R}^n$ un champ de vecteur sur un domaine E de \mathbb{R}^n . Si $s(t) := (s_1(t), \dots, s_n(t))$ et $A(x) = (A_1(x), \dots, A_n(x))$, l'équation vectorielle (5.1) est équivalente à

$$\begin{cases} \dot{s}_1(t) &= A_1(s(t)) \\ \vdots \\ \dot{s}_n(t) &= A_n(s(t)). \end{cases} \quad (5.2)$$

Ceci est un **système de n équations différentielles du premier ordre**. En particulier, si $n = 1$, le champ de vecteur est simplement une fonction dérivable $A : E \rightarrow \mathbb{R}$ et sa trajectoire $x(t)$ est la solution de l'équation différentielle du premier ordre $\dot{x} = A(x)$.

5.3 États stationnaires - Stabilité

Ces notions sont définies de manière analogue au cas discret (§ 1.3, p. 8). Soit (E, A) un système dynamique continu.

Un état x est dit **stationnaire** si $A(x) = 0$. La trajectoire d'un état stationnaire est constante : $s(t) = s(0) = x$. Le système est en *état d'équilibre*.

Une trajectoire est dite **périodique** s'il existe un nombre réel $T > 0$ tel que $s(t+T) = s(t)$. Un état est périodique si sa trajectoire est périodique. Le plus petit de ces nombres T est la **période** de la trajectoire (ou de l'état). Par convention, un état stationnaire est périodique de période 0. Une trajectoire périodique (de période T) s'appelle aussi un **cycle** (de période T).

Un sous-ensemble E' de E est **invariant** si, pour tout $x \in E'$ la trajectoire de x reste dans E' . Un point stationnaire, un cycle ou une trajectoire sont des sous-ensembles invariants.

Un sous-ensemble invariant E' de E est **stable**, ou est un **attracteur**, s'il existe un voisinage V de E' dans E (son **bassin d'attraction**) tel que toutes les trajectoires partant d'états dans V convergent vers E' . Par exemple : un point stationnaire stable, un cycle stable, etc.

Comme au § 3.4.1, p. 64, on peut obtenir des conditions suffisantes pour qu'un état stationnaire b soit stable. Ceci est particulièrement important car il est en général impossible de trouver des formules explicites pour les trajectoires. Ces conditions portent sur la matrice jacobienne $DA(b)$ du champ A en b :

$$DA(b) := \begin{pmatrix} \frac{\partial A_1}{\partial x_1}(b) & \dots & \frac{\partial A_1}{\partial x_n}(b) \\ \vdots & & \vdots \\ \frac{\partial A_n}{\partial x_1}(b) & \dots & \frac{\partial A_n}{\partial x_n}(b) \end{pmatrix} \in \mathcal{M}_{n \times n} \quad (5.3)$$

Cette $n \times n$ -matrice a des valeurs propres. On peut démontrer :

Théorème 5.3.1 *Soit b un état stationnaire pour le système (\mathbb{R}^n, A) . On suppose que le champ de vecteurs A est différentiable en b . Alors, si la partie réelle de chaque valeur propre de $DA(b)$ est < 0 , le point b est stable.*

Exemple : Supposons que $A = \text{grad } f$ pour $f : E \rightarrow \mathbb{R}$. La matrice DA est la matrice hessienne $(\frac{\partial^2 f}{\partial x_i \partial x_j})$ qui est symétrique (voir § 3.2.1). Ses valeurs propres sont donc réelles. La condition du théorème 5.3.1 dit que l'état stationnaire b est stable si b est un maximum local de f (voir théorème 3.3.3 p. 61). Observons du reste que si $s(t)$ est une trajectoire de $\text{grad } f$ alors, par la formule 3.31, p. 56, on a

$$\frac{d(f \circ s)}{dt} = \dot{s}(t) \cdot \text{grad } f(s(t)) = \|\text{grad } f(s(t))\|^2.$$

On voit que $f \circ s$ est strictement croissante en dehors des états stationnaires (la valeur de f croît le long des trajectoires de $\text{grad } f$).

Le cas particulier $n = 1$ dans le théorème 5.3.1 mérite un énoncé pour lui-même. On a une seule équation différentielle $\dot{x} = A(x)$ qui détermine l'évolution d'une grandeur $x(t)$ en fonction de t . La matrice jacobienne $DA(x)$

est une matrice (1×1) dont l'unique coefficient est $A'(x)$, la dérivée de la fonction A . Pour un état stationnaire b ($A(b) = 0$), le théorème 5.3.1 prend donc la forme

Théorème 5.3.2 *Soit b un état stationnaire pour l'équation différentielle $\dot{x} = A(x)$, où $A : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable en b . Alors, si $A'(b) < 0$, le point b est stable. Si $A'(b) > 0$, le point b est instable.*

5.4 Quelques techniques de résolution d'équations différentielles

Trouver une formule explicite pour les solutions d'une équation différentielle (et donc pour les trajectoires d'un champ de vecteurs) est, en général, impossible. Par exemple, l'équation différentielle $\dot{x} = f(t)$ a pour solution toute primitive $x(t) = \int f(t)dt$ qui, en général, est une nouvelle fonction que l'on ne sait pas exprimer. Il y a cependant quelques cas exceptionnels où des expressions explicites des solutions existent et nous allons en voir ci-dessous les plus simples.

Équations “à variables séparables” : Il s'agit des équations du type

$$\frac{dx}{dt} = f(x)g(t). \quad (5.4)$$

On utilise le “truc”

$$\frac{dx}{f(x)} = g(t) dt. \quad (5.5)$$

d'où, en intégrant, une relation

$$\int \frac{dx}{f(x)} = \int g(t) dt + C, \quad C = \text{constante}, \quad (5.6)$$

entre x et t qui, dans les bons cas, permet d'extraire $x(t)$. La constante C se détermine par la condition initiale $x(t_0) = x_0$.

EXEMPLE : Cherchons la solution $x(t)$ de l'équation différentielle

$$\frac{dx}{dt} = t(1+x) \quad (5.7)$$

telle que $x(0) = 3$. En séparant les variables, on obtient

$$\int \frac{dx}{1+x} = \int t dt + C \quad (5.8)$$

donc

$$\ln |1+x| = \frac{t^2}{2} + C \quad (5.9)$$

qui est équivalent à

$$|1+x| = K e^{\frac{t^2}{2}} \quad (5.10)$$

où $K = e^C > 0$. On peut alors enlever la valeur absolue en permettant $K \in \mathbb{R}$. L'équation (5.10) pour $t = 0$ donne $K = 4$, puisque $x(0) = 3$. La solution cherchée est donc

$$x(t) = 4e^{\frac{t^2}{2}} - 1.$$

Équations “linéaires inhomogènes” : Il s'agit des équations du type

$$\frac{dx}{dt} = g(t)x + h(t). \quad (5.11)$$

On introduit le “facteur intégrant”

$$R(t) := e^{-P(t)} \quad (5.12)$$

où P est une primitive de g . On vérifie que la solution générale de (5.11) est alors

$$x(t) = \frac{1}{R(t)} \left[\int R(t) h(t) dt + C \right]. \quad (5.13)$$

Comme pour les variables séparables, la constante C se détermine par la condition initiale $x(t_0) = x_0$.

PREUVE: On dérive $x = \frac{1}{R} [C + \int Rh]$:

$$\dot{x} = \frac{1}{R^2} \left[R^2 h - \dot{R}(C + \int Rh) \right] = h - \frac{\dot{R}}{R} x.$$

Or

$$-\frac{\dot{R}}{R} = -(\ln R)' = P' = g. \quad \square$$

EXEMPLE : Cherchons la solution $x(t)$ de l'équation différentielle

$$\dot{x} = -\frac{x}{t} + t^2 \quad (5.14)$$

telle que $x(1) = 2$. On a donc $g(t) = -1/t$ et $h(t) = t^2$. On peut prendre pour primitive de g la fonction $P(t) = -\ln t$, d'où $R(t) = e^{-P(t)} = t$. La solution générale donnée par (5.13) est

$$x(t) = \frac{1}{t} \left[\int t^3 dt + C \right] = \frac{t^3}{4} + \frac{C}{t} \quad (5.15)$$

et la condition initiale $x(1) = 2$ implique $C = 7/4$.

Systèmes d'équations linéaires : On cherche les trajectoires $\dot{x} = A(x)$ dans \mathbb{R}^n pour un champ $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ qui est linéaire. L'application A est donc donné par une matrice $A := (a_{ij}) \in \mathcal{M}_{n \times n}$ et l'on doit résoudre un système d'équations différentielles

$$\begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \quad (5.16)$$

On observe que si x et \bar{x} sont deux solutions du système (5.16), alors $x + \bar{x}$ et ax ($a \in \mathbb{R}$ constante) sont aussi des solutions. L'ensemble des solutions forme donc donc un espace vectoriel. *On peut démontrer que cet espace vectoriel est de dimension n .* Pour trouver toutes les solutions, il suffit donc d'en trouver n qui soient linéairement indépendantes et d'en faire des combinaisons linéaires. Une façon de trouver des solutions linéairement indépendantes est d'utiliser les valeurs propres de A :

Théorème 5.4.1 *Supposons que la matrice A ait n valeurs propres réelles distinctes $\alpha_1, \dots, \alpha_n$. Alors, la solution générale du système (5.16) est*

$$x(t) = C_1 e^{\alpha_1 t} \xi_1 + \cdots + C_n e^{\alpha_n t} \xi_n$$

où ξ_i est un vecteur propre pour la valeur propre α_i . Les constantes C_i se déterminent par la condition initiale $x(0) = x_0$.

PREUVE: Si $x(t) = C_i e^{\alpha_i t} \xi_i$, alors

$$\dot{x} = C_i e^{\alpha_i t} \alpha_i \xi_i = A(C_i e^{\alpha_i t} \xi_i) = A(x).$$

Donc les $C_i e^{\alpha_i t} \alpha_i \xi_i$ sont des solutions de (5.16). Comme les valeurs propres α_i sont distinctes les vecteurs propres ξ_i sont linéairement indépendants (voir la preuve de la proposition 2.5.3, p. 35.). \square

Lorsque les valeurs propres ne sont pas deux-à-deux distinctes ou qu'elles sont complexes, la situation est plus compliquée.

EXEMPLE : Considérons le système

$$\begin{cases} \dot{x} = y \\ \dot{y} = x \end{cases} \quad (5.17)$$

Matriciellement, il est équivalent à :

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Les valeurs propres de $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ sont 1 (vecteur propre $(1, 1)$) et -1 (vecteur propre $(1, -1)$). La solution générale du système (5.17) est, d'après le théorème 5.4.1

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = C_1 e^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + C_2 e^{-t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

c'est-à-dire :

$$\begin{cases} x(t) = C_1 e^t + C_2 e^{-t} \\ y(t) = C_1 e^t - C_2 e^{-t}. \end{cases} \quad (5.18)$$

Les constantes C_1 et C_2 sont déterminées par les conditions initiales $x(0) = x_0$ et $y(0) = y_0$. En effet, en posant $t = 0$ dans (5.18), on obtient $x_0 = C_1 + C_2$ et $y_0 = C_1 - C_2$. La solution du système (5.17) avec conditions initiales $x(0) = x_0$ et $y(0) = y_0$ sera donc

$$\begin{cases} x(t) = \frac{x_0 + y_0}{2} e^t + \frac{x_0 - y_0}{2} e^{-t} \\ y(t) = \frac{x_0 + y_0}{2} e^t - \frac{x_0 - y_0}{2} e^{-t}. \end{cases}$$

Équations d'ordre supérieur : Une équation différentielle d'ordre $n \geq 2$ est souvent équivalente à un système de n équations du premier ordre. Par exemple, si l'on a une équation du 2^e ordre

$$\ddot{x} = F(x, \dot{x}), \quad (5.19)$$

on rajoute la fonction $y = \dot{x}$ et l'équation (5.19) est équivalente au système

$$\begin{cases} \dot{x} = y \\ \dot{y} = F(x, y). \end{cases} \quad (5.20)$$

Un exemple important est celui des *équations à coefficients constants*

$$\ddot{x} + a\dot{x} + bx = 0 \quad (a, b \in \mathbb{R}). \quad (5.21)$$

En posant $y := \dot{x}$ comme ci-dessus, l'équation (5.21) devient équivalente au système linéaire

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (5.22)$$

Par le paragraphe précédent, les solutions de (5.22) forment un espace vectoriel de dimension 2 qui dépend des valeurs propres de la (2×2) -matrice dans (5.22). Ces valeurs propres sont les solutions λ_1, λ_2 de l'équation du 2^e degré

$$\lambda^2 + a\lambda + b = 0 \quad (5.23)$$

qui s'appelle l'**équation caractéristique** de l'équation différentielle (5.21) (observer l'analogie entre les équations (5.21) et (5.23)). On peut démontrer le théorème suivant :

Théorème 5.4.2 *La solution générale de l'équation différentielle $\ddot{x} + a\dot{x} + bx = 0$ s'exprime, en fonction des solutions λ_1, λ_2 de son équation caractéristique $\lambda^2 + a\lambda + b = 0$, de la façon suivante :*

1. si $\lambda_1 \neq \lambda_2$ et $\lambda_i \in \mathbb{R}$, les solutions de (5.21) sont de la forme

$$C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$$

2. si $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$, les solutions de (5.21) sont de la forme

$$C_1 e^{\lambda t} + C_2 t e^{\lambda t}$$

3. si $\lambda_{1,2} = \alpha \pm \beta i \notin \mathbb{R}$ les solutions de (5.21) sont de la forme

$$e^{\alpha t} (C_1 \cos(\beta t) + C_2 \sin(\beta t)).$$

EXEMPLE 1 : L'OSCILLATEUR HARMONIQUE. C'est l'équation différentielle de base des mouvements oscillatoires qui est

$$\ddot{x} = -\omega^2 x \quad (\omega \in \mathbb{R}). \quad (5.24)$$

Le nombre $x(t)$ s'interprète physiquement comme la position d'un mobile se déplaçant sur une droite et subissant une force, tendant à le ramener à l'origine, force qui est proportionnelle à $x(t)$ (par la loi de Newton, la dérivée seconde (accélération) est proportionnelle à la force subie). On peut imaginer une masse entre deux ressorts tendus.

L'équation caractéristique de (5.24) est $\lambda^2 = -\omega^2$ dont les racines sont $\lambda_{1,2} = \pm i\omega$. Les solutions de l'équation différentielle (5.24) sont donc, par le théorème 5.4.2, toutes de la forme

$$C_1 \cos(\omega t) + C_2 \sin(\omega t). \quad (5.25)$$

Observons que la fonction $A \sin(\omega t + \varphi)$ (A, φ constantes) est aussi solution de $\ddot{x} = -\omega^2 x$. On doit donc pouvoir la mettre sous la forme de (5.25). En effet :

$$A \sin(\omega t + \varphi) = A \cos \varphi \sin(\omega t) + A \sin \varphi \cos(\omega t).$$

On aura donc $C_1 = A \sin \varphi$ et $C_2 = A \cos \varphi$. De même, toute somme $C_1 \cos(\omega t) + C_2 \sin(\omega t)$ peut se mettre sous forme $A \sin(\omega t + \varphi)$ avec $A = \sqrt{C_1^2 + C_2^2}$ et $\varphi = \arctg(C_1/C_2)$. On a ainsi montré que toute solution de $\ddot{x} = -\omega^2 x$ est un mouvement oscillatoire $A \sin(\omega t + \varphi)$ d'*amplitude* A , de *période* $2\pi/\omega$ avec une *phase* (décalage) φ . Observons que ω est déterminé par l'équation différentielle tandis que les constantes A et φ sont déterminées par la façon dont le mouvement démarre : la position initiale $x(0) = x_0$ et la vitesse initiale $\dot{x}(0) = v_0$.

EXEMPLE 2 : L'OSCILLATEUR HARMONIQUE AMORTI. Si, dans l'exemple de l'oscillateur harmonique, le mobile subit une force de friction proportionnelle à sa vitesse \dot{x} , qui le ralentit, l'équation différentielle du mouvement sera

$$\ddot{x} = -2a\dot{x} - \omega^2 x \quad (\omega \in \mathbb{R}, a > 0) \quad (5.26)$$

(le facteur 2 simplifiera l'expression des solutions). Son équation caractéristique est $\lambda^2 = -2a\lambda - \omega^2$ dont les racines sont $\lambda_{1,2} = -a \pm \sqrt{a^2 - \omega^2}$. Lorsque $a < \omega$, ces racines sont complexes conjuguées : $\lambda_{1,2} = -a \pm i\tilde{\omega}$ où $\tilde{\omega} := \sqrt{\omega^2 - a^2}$. Par le théorème 5.4.2, les solutions de l'équation différentielle (5.26) sont de la forme

$$e^{-at} (C_1 \cos(\tilde{\omega}t) + C_2 \sin(\tilde{\omega}t)) = Ae^{-at} \sin(\tilde{\omega}t + \varphi). \quad (5.27)$$

On a donc un mouvement oscillant comme dans l'exemple 1) mais de période $\frac{2\pi}{\tilde{\omega}} > \frac{2\pi}{\omega}$ qui est de plus amorti par la fonction e^{-at} .

5.5 Quelques systèmes dynamiques continus

5.5.1 Croissance exponentielle

C'est le système (\mathbb{R}, A) où $A(x) := \alpha x$ ($\alpha \in \mathbb{R}$ constante). Il correspond donc à l'équation différentielle

$$\boxed{\dot{x} = \alpha x.} \quad (5.28)$$

On vérifie immédiatement que

$$x(t) = x_0 e^{\alpha t} \quad (5.29)$$

satisfait à l'équation 5.28 avec $x(0) = x_0$. C'est donc la trajectoire de l'état x_0 . En posant $a := e^\alpha$, cette trajectoire s'écrit

$$x(t) = x_0 a^t. \quad (5.30)$$

On voit que c'est la solution du système dynamique discret $x(t+1) = a x(t)$ (croissance exponentielle avec coefficient de croissance a ; voir § 1.2).

Le seul état stationnaire est $x_0 = 0$. C'est un attracteur si et seulement si $\alpha < 0$. Cela illustre le théorème 5.3.1 p. 81. En effet, si $A(x) = \alpha x$ alors $DA = (\alpha)$ a pour seule valeur propre α .

REMARQUES : 1) l'équation (5.28) est à variables séparables. Exercice : trouver la solution (5.29) par la méthode de la p. 82.

2) Le champ $A(x) = \alpha x$ est linéaire. La solution (5.29) est un cas particulier du théorème 5.4.1 p. 84.

5.5.2 Croissance logistique

Comme dans le cas des systèmes dynamiques discrets, la croissance logistique, proposée par Verhulst en 1836, modélise l'évolution d'une population animée par une croissance exponentielle mais limitée à K individus par l'environnement. L'équation différentielle de l'évolution du système est

$$\boxed{\dot{x} = \alpha x(1 - \frac{x}{K})} \quad (5.31)$$

(correspondant au système (\mathbb{R}, A) avec $A(x) = \alpha x(1 - \frac{x}{K})$). Il y a deux états stationnaires : $x = 0$ et $x = K$. On a $A'(x) = \alpha(1 - 2x/K)$, donc $A'(0) = \alpha$ et $A'(K) = -\alpha$. Si $\alpha > 0$, on a donc que 0 est instable et K est un attracteur.

Le système dynamique continu de la croissance logistique est ainsi nettement plus simple que son analogue discret vu au § 1.4.

L'équation différentielle (5.31) de la croissance logistique est à variable séparables ce qui permet de trouver, par la méthode de la p. 82 une formule explicite de la trajectoire $x(t)$ telle que $x(0) = x_0$

$$x(t) = \frac{x_0 K e^{\alpha t}}{K + x_0(e^{\alpha t} - 1)}. \quad (5.32)$$

Par la règle de l'Hospital, on voit que $x(t) \rightarrow K$ quand $t \rightarrow \infty$ (si $\alpha > 0$).

5.5.3 Proies–Prédateurs

Ce système a été proposé par Volterra en 1926 pour modéliser l’interaction d’une population de proies avec leurs prédateurs (voir [9, p. 63]). C’est le système (E, A) où

1. $E := \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$. Le nombre x est la quantité de proies et y celle de prédateurs.
2. le champ $A = (A_1, A_2)$ est $A(x, y) := (x(\alpha - \beta y), y(\gamma x - \delta))$.

Les trajectoires sont donc les solutions du système d’équations différentielles :

$$\begin{cases} \dot{x} = x(\alpha - \beta y) \\ \dot{y} = y(\gamma x - \delta). \end{cases} \quad (5.33)$$

Ceci rend compte des hypothèses suivantes :

1. en l’absence de prédateurs, les proies grandissent exponentiellement, avec taux de croissance α .
2. l’effet de la prédation est de réduire la quantité de proies d’un facteur proportionnel à x et à y . C’est le terme $-\beta xy$.
3. en l’absence de proies, les prédateurs, qui n’ont pas de nourriture, décroissent exponentiellement avec un taux δ .
4. l’accroissement des prédateurs est proportionnel à leur quantité et à celle des proies. C’est le terme γxy .

Si $x > 0$ et $y > 0$, on a un unique état stationnaire $(\frac{\delta}{\gamma}, \frac{\alpha}{\beta})$.

On peut faire le changement de variables

$$u := \frac{\gamma}{\delta}x \quad , \quad v := \frac{\beta}{\alpha}y. \quad (5.34)$$

Avec ces nouvelles coordonnées (u, v) , le système (5.33) devient :

$$\begin{cases} \dot{u} = \alpha u(1 - v) \\ \dot{v} = \delta v(u - 1). \end{cases} \quad (5.35)$$

et l’état stationnaire est $(1, 1)$. Le système d’équations différentielles (5.35) correspond au champ de vecteurs $A(u, v) := (\alpha u(1 - v), \delta v(u - 1))$.

La matrice jacobienne de $DA(1, 1)$ est

$$DA(1, 1) = \begin{pmatrix} 0 & -\alpha \\ \delta & 0 \end{pmatrix}$$

Son polynôme caractéristique est $X^2 + \alpha\delta$. Le spectre de $DA(1, 1)$ est ainsi $\pm i\sqrt{\alpha\delta}$. La partie réelle de ces nombres complexes étant nulle, le théorème 5.3.1 ne donne aucune information sur la stabilité de l'état stationnaire $(1, 1)$.

En fait, les trajectoires sont périodiques. En effet, considérons la fonction

$$H(u, v) := \frac{\delta}{\alpha}(u - \ln u) + v - \ln v$$

On a

$$\text{grad } H = \left(\frac{\delta}{\alpha} \frac{u-1}{u}, \frac{v-1}{v} \right).$$

d'où

$$\text{grad } H \cdot A = 0. \quad (5.36)$$

Le gradient de H est donc orthogonal au champ A . Comme $\text{grad } H$ est orthogonal aux courbes de niveau de H , on en déduit que les trajectoires $s(t)$ de A sont des paramétrisations des courbes de niveau de H (et que $H(s(t)) = \text{cte}$).

Or, $\text{grad } H(1, 1) = 0$ et la matrice hessienne de H en $(1, 1)$ est

$$\begin{pmatrix} \frac{\delta}{\alpha} & 0 \\ 0 & 1 \end{pmatrix}.$$

Ses valeurs propres sont positives. Par le théorème 3.3.4, p. 59 le point $(1, 1)$ est un minimum local pour H . Les courbes de niveau autour d'un minimum local sont des courbes fermées.

Annexe A

Notes techniques et compléments

A.1 Notations

Logique – Ensembles :

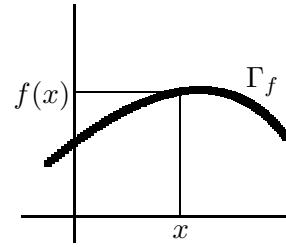
| | |
|-------------------|--|
| \forall | pour tout, quel que soit |
| \exists | il existe, on peut trouver |
| \Rightarrow | implique, entraîne |
| \Leftrightarrow | est équivalent à , si et seulement si |
| $::=$ | égal par définition à |
| $x \in E$ | x appartient à l'ensemble E ; x est élément de E . |
| $E \subset F$ | E est inclus dans F ; E est un sous-ensemble de F |
| $\{x \mid p(x)\}$ | ensemble des x tels que la propriété $p(x)$ est vraie |
| \emptyset | ensemble vide |
| \cup | union |
| \cap | intersection |
| $E \times F$ | Ensemble produit ; $E \times F := \{(x, y) \mid x \in E \text{ et } y \in F\}$ |
| E^n | $E \times \cdots \times E$ (n fois). Les éléments de E^n sont les n -uples (x_1, \dots, x_n) avec $x_i \in E$. |

Application – Fonctions Une **application** f définie sur l'ensemble E à valeurs dans l'ensemble F est une correspondance qui associe à tout élément $x \in E$ un élément $y = f(x) \in F$. On peut l'imaginer comme une cause x qui détermine un effet $f(x)$. On note $f : E \rightarrow F$ et/ou $f : x \mapsto y$.

L'application f est déterminée par son **graphé** Γ_f :

$$\Gamma_f := \{(x, y) \in E \times F \mid y = f(x)\} \subset E \times F.$$

Lorsque $E = F = \mathbb{R}$, le graphé de f est un sous-ensemble du plan que l'on utilise pour visualiser f . Si $E = \mathbb{R}^2$ et $F = \mathbb{R}$, cette visualisation peut se faire en perspective (graphé 3D).



Les mots *application* et *fonction* sont synonymes. Le terme “fonction” est plutôt réservé aux applications à valeurs numériques ($f : E \rightarrow \mathbb{R}$) mais nous l'avons aussi employé pour des applications à valeur dans \mathbb{R}^n .

Une application $f : E \rightarrow F$ est :

- *injective* si $f(x) = f(x') \Rightarrow x = x'$ (au plus une préimage).
- *surjective* si $\forall y \in F, \exists x \in E$ tel que $f(x) = y$ (au moins une préimage).
- *bijective* si elle est injective et surjective (exactement une préimage).

A.2 Entiers

Les entiers sont :

- les **entiers naturels** : $\mathbb{N} := \{0, 1, 2, 3, \dots\}$
- les **entiers relatifs** : $\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$.

Le calcul avec les nombres entiers est, par nature, un calcul machinal. Depuis l'Antiquité, on se sert d'abaques ou de bouliers. Notre écriture de position (unités, dizaines, etc), introduite par les Hindous et diffusée par les Arabes dans le Haut-Moyen-Age, est un codage graphique de ces procédés. Le logiciel d'une calculette ou d'un ordinateur est élaboré sur le même principe que l'abaque ou le boulier.

Les logiciels de calcul scientifique comme MAPLE sont censés calculer exactement en nombre entier sans limite a priori. Les restrictions ne proviennent que des capacités de mémoire et de rapidité de la machine sur laquelle ils sont implantés. En revanche, de nombreux autres programmes imposent des limites sur la grandeur des nombres entiers.

Des calculs avec des entiers de plus de 100 chiffres sont quotidiennement effectués en cryptographie pour la sécurité des données informatiques et leur transmission sur l'internet.

A.3 Nombres rationnels

L'ensemble de **nombres rationnels** \mathbb{Q} est l'ensemble des fractions

$$\mathbb{Q} := \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z} \text{ et } q \neq 0 \right\}$$

muni de l'addition et de la multiplication usuelles

$$\frac{p}{q} + \frac{p'}{q'} = \frac{pq' + qp'}{qq'} \quad , \quad \frac{p}{q} \cdot \frac{p'}{q'} = \frac{pp'}{qq'}.$$

Avec ces opérations, \mathbb{Q} est un **corps**, c'est-à-dire que pour tout $x, y, z \in \mathbb{Q}$, on a :

- 1) $(x + y) + z = x + (y + z)$.
- 2) $0 + x = x + 0 = x$.
- 3) $\exists -x \in \mathbb{Q}$ tel que $x + (-x) = 0$.
- 4) $x + y = y + x$.
- 5) $(xy)z = x(yz)$
- 6) $1 \cdot x = x \cdot 1 = x$.
- 7) $x(y + z) = xy + xz$.
- 8) $xy = yx$.
- 9) les éléments non-nuls sont inversibles : si $x \neq 0$, il existe (un unique) x^{-1} avec $xx^{-1} = 1$

La propriété 9) est la seule qui ne soit pas vraie pour \mathbb{Z} . Toutes les propriétés algébriques de \mathbb{Q} découlent de celles ci-dessus. Par exemple, $0x = 0$ provient de $0x = (0 + 0)x = 0x + 0x$.

Un nombre rationnel p/q est un couple d'entier (p, q) avec la relation d'équivalence

$$(p, q) \sim (p', q') \iff pq' = p'q.$$

Les nombres rationnels se prêtent donc au calcul exact à la machine, comme les nombres entiers.

Dans la vie quotidienne, les fractions sont utilisées pour la mesure approximative des grandeurs continues (longueurs, temps, etc). Par exemple $x = 0.354$ veut dire $x = 354/1000$. Cependant, comme l'ont découvert les Pythagoriciens (V^e siècle avant JC), les nombres rationnels sont impropre à l'expression exacte des grandeurs continues et on ne peut pas les mettre en bijection avec tous les points d'une droite. Par exemple, la longueur d de la diagonale d'un carré de côté 1 satisfait l'équation $d^2 = 2$ par le théorème de Pythagore. Or, il n'existe *aucune* fraction dont le carré vaut 2. En effet :

Proposition A.3.1 *Soient k et m des entiers positifs. Supposons que l'équation $x^k = m$ n'ait pas de solution dans \mathbb{Z} . Alors, elle n'en a pas non plus dans \mathbb{Q} .*

PREUVE: La fraction p/q qui satisferait $(p/q)^k = m$ n'est pas un entier par l'hypothèse. Donc $q \neq 1$. On suppose que p/q est une fraction réduite, c'est-à-dire qu'il n'y a pas de possibilité de simplification entre p et q (exemple $3/4$). Dans ce cas, la fraction

$$\left(\frac{p}{q}\right)^k = \frac{p \cdots p}{q \cdots q}$$

est aussi réduite, car on ne peut toujours rien simplifier, et $q^k \neq 1$. Donc la puissance k^e d'une fraction qui n'est pas un entier n'est pas non plus un entier. \square

A.4 Nombres réels

Considérons les deux affirmations suivantes :

1. Si la température d'une pièce passe de 18° à 21° en une heure, il y a eu au moins un instant pendant cette heure où cette température a été égale à 20° .
2. Quand une balle de tennis passe d'un côté à l'autre du court, il y a un instant où son centre de gravité est exactement au dessus du filet.

Ces affirmations relèvent du **principe de continuité** qui reflète notre impression sur la nature "continue" de grandeurs comme le temps, la température ou la position d'un objet dans l'espace. Elles ne sont guère vérifiables expérimentalement. On accumulerait plutôt des preuves du contraire ! En visionnant le film d'une partie de tennis image par image, la probabilité est nulle d'en trouver une où le centre de la balle soit exactement au dessus du filet (à comparer avec la remarque 2.6.3, p. 43). Le principe de continuité n'est qu'une vue de l'esprit, une sorte d'exigence du raisonnement.

Reste à trouver un système de nombres maniable et logiquement cohérent dans lequel le principe de continuité soit vrai. Historiquement, ce fut une tâche difficile qui a duré plus de 2000 ans et ne fut achevée qu'à la fin du XIX^e siècle [6, ch. 2].

Le corps des **nombres réels** \mathbb{R} est l'aboutissement de ces longs efforts. Il est exactement construit pour que le principe de continuité fonctionne¹. Les mathématiciens en ont plusieurs constructions équivalentes ([1, p. 177], [6, ch. 2]). Intuitivement, un nombre réel est une limite de fractions.

Nous nous bornerons ici à expliquer par un exemple le maniement pratique des nombres réels. Considérons la grandeur x^2 . Comme $1^2 = 1 < 2$ et $2^2 = 4 > 2$, le principe de continuité dit qu'il existe au moins un nombre réel

¹L'adjectif "réel" est usurpé puisque le principe de continuité n'existe que dans notre imagination. Cette confusion est source de blocages psychologiques envers les nombres complexes, arbitrairement considérés eux comme "imaginaires".

$1 < r < 2$ tel que $r^2 = 2$. Comme $x \mapsto x^2$ est strictement croissante, il n'y en aura qu'un. Cherchons-le :

| | | | | | |
|--------------|-----|-----|-----|-----|-----|
| x | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 |
| $x^2 \leq 2$ | oui | oui | oui | oui | non |

Donc $1.4 < r < 1.5$. On recommence avec un pas plus fin :

| | | | |
|--------------|------|------|------|
| x | 1.40 | 1.41 | 1.42 |
| $x^2 \leq 2$ | oui | oui | non |

donc $1.41 < r < 1.42$, etc. On obtiendra ainsi, successivement, de plus en plus de décimales de $\sqrt{2}$.

Cette procédure illustre deux des façons de s'imaginer un nombre réel :

1. comme une suite d'intervalles emboîtés d'extrémités rationnelles dont la longueur tend vers 0.
2. comme un développement décimal illimité. Si le nombre est rationnel, son développement est fini ou devient périodique. Mais un nombre irrationnel, comme $\sqrt{2}$, a un côté inatteignable : la connaissance de ses décimales ne pourra pas s'obtenir en un nombre fini d'opérations.

La suite d'intervalles circonscrivant un nombre réel n'est évidemment pas unique. En général, son écriture décimale ne l'est pas non plus. Par exemple : $0.999\dots = 1.000\dots$ En effet, si $n := 0.999\dots$, on a $10n = 9.999\dots$, d'où $10n - n = 9$ et donc $n = 1$.

En pratique, et avec les machines, on travaille avec des nombres réels "arrondis" qui correspondent à de petits intervalles de nombres rationnels fermés à gauche :

$$x \approx 0.453 \iff 0.4525 \leq x < 0.4535$$

Dans MAPLE, la commande `evalf(X,n)` renvoie l'évaluation de l'expression X arrondie à n chiffres, en principe sans limite sur n due au logiciel. Mais attention aux multiples pièges ! Par exemple :

```
> v := evalf(sqrt(2),40);
v := 1.414213562373095048801688724209698078570
> v^2 - 2;
0
```

Cette réponse $v^2 - 2 = 0$ est manifestement fausse. On a $v = \sqrt{2} + \varepsilon$ avec $\varepsilon \approx 10^{-39}$ (on a demandé 40 chiffre, il reste 39 décimales). D'où

$$v^2 - 2 = 2 + 2\varepsilon + \varepsilon^2 - 2 \approx 2\varepsilon.$$

On s'attend ainsi à ce que $v^2 - 2 \approx 10^{-39}$ ou 10^{-38} . Il se passe qu'en effectuant l'élévation au carré de v , MAPLE a repris sa précision par défaut, qui est de 10 chiffres. Cette précision est contrôlée par la commande `Digits` et, si l'on veut garder nos 40 chiffres, il faut le demander *avant de commencer le calcul* :

```
> Digits := 45 : v^2 - 2;
.92808 10-39
```

Cette fois, la réponse est plausible. De tels avatars ne sont pas dus à des imperfections du logiciel ; ils sont inhérents au maniement des nombres réels sur une machine et imposent de sévères limites aux simulations numériques.

A.5 Entre les rationnels et les réels

Si la construction des nombres réels ne s'est achevée qu'à la fin du XIX^e siècle, on n'a pas attendu jusque-là pour faire des calculs ! Les Arabes, puis les Italiens, ont développé au Moyen-Age le calcul algébrique que l'on connaît avec des nombres rationnels et des expressions comme \sqrt{u} , etc. Il s'agit de calcul *formel*. On ne s'occupe pas de savoir ce que désigne le symbole \sqrt{u} . Il est traité comme une variable de polynôme, sauf qu'il satisfait à la relation $(\sqrt{u})^2 = u$. Cela permet, chaque fois que l'on rencontre $(\sqrt{u})^2$, de le remplacer par u . De même, on peut calculer avec $\sqrt[3]{u}$, qui satisfait à $(\sqrt[3]{u})^3 = u$, ou avec d'autres symboles satisfaisant à des relations algébriques plus compliquées.

Certains nombres, comme π ou e sont en revanche **transcendants** : il ne satisfont à aucune équation polynomiale à coefficients rationnels. Dans les calculs, ils se comportent comme des variables, sans autres simplifications possibles que celles usuelles dans les polynômes.

Contrairement à l'approche numérique des nombres réels, les algorithmes de ce calcul symbolique se prêtent au calcul exact à la machine. Cela permet de caluler *exactement* avec de petites parties de \mathbb{R} . Voici deux exemples avec MAPLE :

```
> v := (1+sqrt(2))^4; w := (1+Pi)^4;
```

$$v := \left(1 + \sqrt{2}\right)^4$$

$$w := (1 + \pi)^4$$

```
> expand(v); expand(w);
```

$$17 + 12\sqrt{2}$$

$$1 + 4\pi + 6\pi^2 + 4\pi^3 + \pi^4$$

A.6 Nombres complexes

Le passage des nombres réels \mathbb{R} aux nombres complexes \mathbb{C} relève du procédé expliqué dans le § A.5 : on calcule algébriquement avec les nombres réels et une expression $\sqrt{-1}$ que l'on note généralement i . Un **nombre complexe** est donc un polynôme en une variable i mais, comme i^2 se remplace par -1 , on n'aura que des polynômes de degré 1. Un nombre complexe est donc une expression $a + bi$. La somme et le produit se font comme pour des polynômes, en remplaçant chaque i^2 par -1 . Exemple :

$$(3 + 4i)(1 - 2i) = 3 + 4i - 6i - 8i^2 = 11 - 2i$$

$$i^3 = i^2i = -i.$$

Avec ces opérations, l'ensemble \mathbb{C} des nombres complexes est un corps (p. 93) : la seule chose non-banale à vérifier est que si $a + bi \neq 0$, il admet un inverse. Celui-ci est donné par

$$(a + bi)^{-1} = \frac{a - bi}{a^2 + b^2}. \quad (\text{A.1})$$

Se donner un nombre complexe $z = a + bi \in \mathbb{C}$ revient à se donner un couple $(a, b) \in \mathbb{R}^2$ de nombres réels. On met ainsi en bijection \mathbb{C} avec \mathbb{R}^2 , donnant l'interprétation géométrique des nombres complexes comme les points du plan (les nombres réels devenant les points de l'axe horizontal). Si $z = a + bi \in \mathbb{C}$, alors

1. a s'appelle la **partie réelle** et b la **partie imaginaire** de z .
2. Le nombre complexe $\bar{z} := a - ib$ s'appelle le **conjugué** de z . Dans le plan, z et \bar{z} s'obtiennent l'un de l'autre par réflexion par rapport à l'axe horizontal. On a $\overline{z + z'} = \bar{z} + \bar{z}'$ et $\overline{zz'} = \bar{z}\bar{z}'$

3. Observons que $z\bar{z} = a^2 + b^2 \in \mathbb{R}_{\geq 0}$. Le nombre réel

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2} \geq 0$$

s'appelle le **module** de z . Par le théorème de Pythagore, $|z|$ est la distance entre z et l'origine 0. La formule A.1 peut se réécrire :

$$z^{-1} = \frac{\bar{z}}{|z|^2}.$$

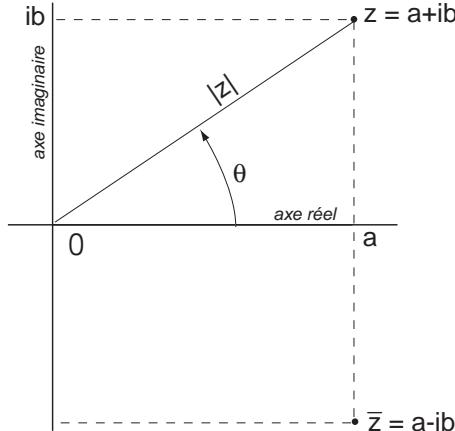
4. les coordonnées polaires permettent d'écrire $z = a + ib$ sous **forme trigonométrique** :

$$z = \rho(\cos \theta + i \sin \theta) \quad (\text{A.2})$$

On a donc $a = \rho \cos \theta$, $b = \rho \sin \theta$ et $\rho = |z|$. L'angle θ , l'**argument** de z , est défini à un multiple entier de 2π près (et non-défini si $z = 0$). La relation $\operatorname{tg} \theta = \frac{b}{a}$ a lieu. La formule pour θ en utilisant la fonction arctg dont les valeurs sont dans $[-\frac{\pi}{2}, \frac{\pi}{2}]$ est

$$\theta = \begin{cases} \operatorname{arctg}(\frac{b}{a}) & \text{si } a \geq 0 \\ \operatorname{arctg}(\frac{b}{a}) + \pi & \text{si } a < 0 \end{cases} \quad (\text{A.3})$$

Par exemple : l'argument de $1 + i$ est $\pi/4$ et celui de $-1 + i$ est $3\pi/4$.



Notation exponentielle des nombres complexes : En comparant les développements de Taylor (voir p. 50) de $\cos u$, $\sin u$ et e^u :

$$\cos u = 1 - \frac{1}{2}u^2 + \frac{1}{4!}u^4 + \dots$$

$$\sin u = u - \frac{1}{3!}u^3 + \dots$$

$$e^u = 1 + u + \frac{1}{2}u^2 + \frac{1}{3!}u^3 + \frac{1}{4!}u^4 + \dots$$

et en remplaçant u par $i\theta$ dans la dernière équation, Euler a trouvé que :

$$e^{i\theta} = \cos\theta + i \sin\theta \quad (\text{A.4})$$

En combinant avec (A.2), on obtient la notation exponentielle des nombres complexes :

$$z = \rho e^{i\theta} \quad (\text{A.5})$$

où $\rho = |z|$ est le module et θ l'argument de z . La notation exponentielle rend le produit des nombres complexes particulièrement simple :

$$zz' = \rho e^{i\theta} \rho' e^{i\theta'} = \rho \rho' e^{i(\theta+\theta')} \quad (\text{A.6})$$

On voit que lors d'un produit, les modules se multiplient et les arguments s'additionnent. En particulier, la multiplication par i est une rotation de $\pi/2$. En restreignant aux nombres réels ($\theta = 0$ ou π) cette interprétation géométrique du produit des nombres complexes jette un nouvel éclairage sur la règle des signes des nombres réels (moins par moins = plus). Quant à l'addition des nombres complexes, c'est, géométriquement, l'addition des vecteurs dans le plan.

A.7 Polynômes

Un **polynôme de degré n** est une expression

$$P(x) := a_n x^n + \cdots + a_1 x + a_0 \quad (\text{A.7})$$

avec $a_n \neq 0$. Les coefficients a_k sont dans un corps \mathbb{K} , pour nous \mathbb{R} ou \mathbb{C} . Un polynôme définit donc une fonction de \mathbb{K} dans \mathbb{K} . Un nombre $u \in \mathbb{K}$ est une **racine** du polynôme P si $P(u) = 0$.

Il est avantageux de pouvoir décomposer un polynôme P en produit de polynômes de degré plus petit. Une bonne situation est d'avoir des facteurs du 1^{er} degré $x - u$. Cela implique que u est une racine de P ; on a équivalence :

Proposition A.7.1 $u \in \mathbb{K}$ est une racine du polynôme P si et seulement si P est divisible par $x - u$, c'est-à-dire qu'il existe un polynôme Q tel que

$$P(x) = (x - u)Q(x).$$

En particulier, un polynôme de degré n a au plus n racines.

PREUVE: Comme pour les nombres entiers, le calcul des polynômes comprend la **division avec reste** : on peut diviser P par un polynôme $S \neq 0$, obtenant un quotient Q et un reste R :

$$P(x) = S(x)Q(x) + R(x)$$

avec $\deg(R) < \deg(S)$. Si $S = x - u$, le degré de S vaut 1 et donc $R = r_0$ est un polynôme constant. Mais comme $P(u) = 0$, on aura $r_0 = 0$. \square

Le corps des nombres complexes jouit de la propriété remarquable que tout polynôme se factorise de façon maximale. En effet, on peut démontrer :

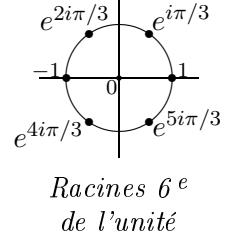
Théorème A.7.2 (Théorème fondamental de l'algèbre) Soit $P(x) := a_n x^n + \dots + a_1 x + a_0$ un polynôme à coefficients dans \mathbb{C} . Alors, P se décompose en produit de polynômes du 1er degré :

$$P(x) = a_n \prod_{k=1}^n (x - u_k)$$

(Remarque : les racines $u_k \in \mathbb{C}$ de P ne sont pas forcément toutes distinctes.)

Exemple : l'équation $z^n = 1$ admet n solutions distinctes : $\{e^{2i\pi k/n} \mid k = 0, \dots, n-1\}$. Le module de ces nombres complexes est 1 et leur argument est $2\pi k/n$. On les appelle les **racines n ème de l'unité**. Géométriquement, elles constituent les sommets d'un polygone régulier à n cotés sur le cercle de rayon 1. Ce sont donc les racines du polynôme $z^n - 1$ qui se factorise en

$$z^n - 1 = (z - 1)(z - e^{2i\pi/n})(z - e^{4i\pi/n}) \cdots (z - e^{2i\pi(n-1)/n}).$$



Dans le corps des nombres réels certains polynômes ne se factorisent pas. Par exemple, un polynômes du 2^e degré $P(x) = ax^2 + bx + c$ ne se factorise que si son discriminant $b^2 - 4ac \geq 0$. Sinon, il n'a pas de racines réelles et P est *irréductible*. Il se trouve que c'est le seul cas :

Proposition A.7.3 Tout polynôme à coefficients réels se factorise en produit de polynômes de degré 1 et de polynômes de degré 2 dont le discriminant est strictement négatif.

PREUVE: Soit P un polynôme à coefficients réels. Considéré comme un polynôme à coefficients complexes il se factorise complètement, par le théorème A.7.2 :

$$P(x) = a_n \prod_{k=1}^n (x - u_k)$$

avec $u_k \in \mathbb{C}$. On peut conjuguer dans \mathbb{C} cette dernière équation :

$$\overline{P(x)} = \overline{a_n} \prod_{k=1}^n (\overline{x} - \overline{u_k}).$$

Si $x \in \mathbb{R}$, on a $\overline{P(x)} = P(x)$ puisque les coefficients de P sont réels. Cela entraîne que si $u_k := \rho(\cos \theta + i \sin \theta)$ est une racine de P non réelle, alors $\overline{u_k}$ est aussi racine de P . Dans ce cas, on peut regrouper ces deux racines conjuguées

$$(x - u_k)(x - \overline{u_k}) = x^2 + 2\rho \cos \theta + \rho^2$$

ce qui donne un polynôme du 2^e degré à coefficients réels et à discriminant négatif. \square

A.8 Espaces vectoriels – Produits scalaires

Définition : Un **espace vectoriel** (réel) est un ensemble V muni de deux lois

$$\begin{array}{ccc} V \times V & \xrightarrow{+} & V \\ (x, y) & \mapsto & x + y \end{array} \quad \text{et} \quad \begin{array}{ccc} \mathbb{R} \times V & \xrightarrow{\cdot} & V \\ (\lambda, y) & \mapsto & \lambda \cdot y \end{array}.$$

satisfaisant aux axiômes suivants : il existe un élément $0 \in V$ (**le vecteur nul**) et, pour tout $x, y, z \in V$ et $\lambda, \mu \in \mathbb{R}$, on a

- | | |
|---|---|
| 1) $(x + y) + z = x + (y + z).$ | 5) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y.$ |
| 2) $0 + x = x + 0 = x.$ | 6) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x.$ |
| 3) $\exists (-x) \in V$ tel que $x + (-x) = 0.$ | 7) $(\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x).$ |
| 4) $x + y = y + x.$ | 8) $1 \cdot x = x.$ |

Les éléments de V s'appellent des **vecteurs** et ceux de \mathbb{R} des **scalaires**. Si l'on remplace \mathbb{R} par \mathbb{C} , on aura la définition d'un **espace vectoriel complexe**. Pour alléger l'écriture, on notera αx pour $\alpha \cdot x$.

Les axiomes 1) à 8) ci-dessus impliquent d'autres propriétés, par exemple : $0 \cdot x = 0$. En effet, par 6), on a

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$$

En soustrayant $0 \cdot x$ (possible par 3)), il reste $0 \cdot x = 0$. De même, on a $\alpha \cdot 0 = 0$ et $(-1) \cdot x = -x$.

Soit V un espace vectoriel et $W \subset V$. On dit que W est un **sous-espace vectoriel** de V si $0 \in W$ et si

$$x, y \in W \Rightarrow x + y \in W \text{ et } \lambda x \in W \quad \forall \lambda \in \mathbb{R}.$$

Par exemple, dans l'espace \mathbb{R}^3 , une droite ou un plan passant par 0 sont des sous-espaces vectoriels.

Produit scalaire : Voici la démonstration de la proposition 2.1.1 p. 16 dont nous recopions l'énoncé :

Proposition 2.1.1 *Les propriétés du produit scalaire et de la norme sont*

- 1) $(x + y) \cdot z = x \cdot z + y \cdot z.$
- 2) $(\lambda x) \cdot y = \lambda(x \cdot y).$
- 3) $x \cdot y = y \cdot x.$
- 4) $\|x\| \geq 0$ et ($\|x\| = 0 \Leftrightarrow x = 0$).
- 5) $\|\lambda x\| = |\lambda| \|x\|.$
- 6) $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2x \cdot y.$
- 7) $|x \cdot y| \leq \|x\| \|y\|$ (inégalité de Schwarz).
- 8) $\|x + y\| \leq \|x\| + \|y\|$ (inégalité du triangle).

PREUVE: Les propriétés 1) à 5) découlent directement des définitions du produit scalaire et de la norme. L'énoncé 6) vient de 1) et 3) :

$$\|x + y\|^2 = (x + y) \cdot (x + y) = x \cdot x + y \cdot y + x \cdot y + y \cdot x = \|x\|^2 + \|y\|^2 + 2x \cdot y.$$

L'inégalité du triangle provient de 6) et de l'inégalité de Schwarz :

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2x \cdot y \leq \|x\|^2 + \|y\|^2 + 2|x \cdot y| \leq \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2. \end{aligned}$$

Il reste à démontrer l'inégalité de Schwarz. La démonstration est assez intéressante. Si $x \cdot y = 0$, il n'y a rien à démontrer. Supposons que $x \cdot y \neq 0$, ce qui implique $x \neq 0$ et, par la propriété 4), $\|x\| \neq 0$.

On calcule $\|\alpha x + y\|^2$ pour $\alpha \in \mathbb{R}$:

$$0 \leq \|\alpha x + y\|^2 = (\alpha x + y) \cdot (\alpha x + y) = \alpha^2 \|x\|^2 + 2\alpha(x \cdot y) + \|y\|^2.$$

Le membre de droite est un polynôme du deuxième degré en α :

$$0 \leq A\alpha^2 + B\alpha + C.$$

Cela montre que l'équation $0 = A\alpha^2 + B\alpha + C$ a au plus une solution, d'où $B^2 - 4AC \leq 0$. En remplaçant A , B et C par leurs valeurs, cette dernière inégalité devient :

$$4(x \cdot y)^2 \leq 4\|x\|^2\|y\|^2.$$

En prenant les racines carrées, on obtient l'inégalité de Schwarz. \square

A.9 Composantes principales

Imaginons que l'on fasse n mesures de p grandeurs. Les données recueillies constituent une matrice $(p \times n)$:

$$R := \begin{pmatrix} r_{11} & \cdots & \cdots & r_{1n} \\ \vdots & & & \vdots \\ r_{p1} & \cdots & \cdots & r_{pn} \end{pmatrix} \quad (\text{A.8})$$

la i ème mesure donnant le i ème vecteur colonne R_i de la matrice R . On regarde R comme un nuage de n points dans \mathbb{R}^p . L'*analyse en composante principales* [15] résout la question suivante : comment projeter R dans un sous-espace Π de dimension $k < p$ dans \mathbb{R}^p en produisant le minimum de distorsion ? La distorsion entre deux nuages R et R' se mesure par le nombre $\sum_{i=1}^n \|R'_i - R_i\|^2$. Les motivations sont les suivantes :

1. compression de l'information pour stockage ou transmission.
2. représentation graphique des données ($k = 2$).
3. élimination de petites perturbations dues aux imprécisions ou à un bruit de fond.
4. mise en évidence de certaines combinaisons linéaires des données qui sont (presque) constantes, en vue de dégager des "lois de conservation" (analogues à la loi de conservation de l'énergie en physique).

L'idée pour trouver Π est la suivante. On suppose que $\sum_{i=1}^n R_i = 0$, ce qui s'obtient en translatant l'origine au centre de gravité $\frac{1}{n} \sum_{i=1}^n R_i = 0$ du nuage R . On forme alors la **matrice de Gram** des vecteurs ligne :

$$S := RR^T \in \mathcal{M}_{p \times p}.$$

Les coefficients s_{ij} de S sont les produits scalaires des vecteurs ligne de R ; S est donc une matrice symétrique. Supposons, pour simplifier, que S ait p valeurs propres distinctes $\lambda_1 > \dots > \lambda_p$ (on peut montrer que les valeurs propres d'une matrice de Gram sont ≥ 0). On peut montrer :

Proposition A.9.1

1. *Le k -sous-espace vectoriel Π cherché est engendré par les vecteurs propres correspondant aux valeurs propres $\lambda_1, \dots, \lambda_k$ de la matrice S .*
2. *la distorsion introduite en projetant orthogonalement R sur Π est $\lambda_{k+1} + \dots + \lambda_p$. Par exemple, $\lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_p = 0$ si et seulement si R est déjà contenu dans un sous-espace vectoriel de dimension k (c'est-à-dire $\text{rang } R = k$).*

Bibliographie

- [1] Hairer G. & Wanner G. L'analyse au fil de l'histoire. *Springer-Verlag* 2001.
- [2] Grifone J. Algèbre linéaire. *Cepadues-Editions* 1990.
- [3] Callahan J., Hoffman K. & al. Calculus in context. *W.H. Freeman and Co.* 1995.
- [4] Giroux A. Mathématiques pour chimistes. *Presses de l'Université de Montréal* 1983.
- [5] Commission romande de mathématique, de physique et de chimie. Formulaires et Tables. *Editions du Tricorne* 1985. (*tables de l'enseignement secondaire romand*).
- [6] Ebbinghaus H-D. & al. Numbers. *Springer-Verlag* 1991.
- [7] Hubbard J.H. & B.B. Vector calculus, linear algebra and differential forms : an unified approach. *Prentice-Hall* 1999.
- [8] Hoppensteadt C.S. & Peskin C.S. Mathematics in medecine and life sciences. *Springer-Verlag* 1992.
- [9] Murray J.D. Mathematical biology. *Springer-Verlag* 1990.
- [10] Pollard J.H. Mathematical models for the growth of human populations. *Cambridge University Press* 1973.
- [11] Sandefur T. Discrete dynamical systems, theory and applications. *Calderon Press, Oxford* 1990.
- [12] Devaney R. Chaos, fractals and dynamics. *Adison-Wesley* 1990.
- [13] Stewart I. La nature et les nombres : l'irréelle réalité des mathématiques. *Hachette* 1998.
- [14] Douady S. & Couder Y. Phyllotaxis as a physical self-organized growth process. *Physical Review Letters* 68 (1992) 2098-2101.
- [15] K.I. Diamantaris & S.Y. Yung. Principal component, Neural Networks, theory and applications. John Wyley & sons, Inc. 1996.

MAPLE :

- [16] Levine A. Introduction à MAPLE. *Ellipses* 1994.
- [17] Garvan F. MAPLE V primer. *CRC Press* 1997. (*format de poche*).
- [18] Cornill J-M. & Testud Ph. MAPLE : Introduction raisonnée à l'usage de l'étudiant, de l'ingénieur et du chercheur. *Springer-Verlag* 1995.

Index

- accroissements finis, théorème des 47
- affine, fonction 45, 53
- angle entre deux vecteurs 16
- application 91
 - graphie d'une 92
 - argument, d'un nombre complexe 98
 - attracteur 9, 81
 - périodique, voir cycle stable 10, 12
- base 18
 - standard de \mathbb{R}^n 18
 - changement de, 31
- bassin (d'attraction) 8, 81
- bijective, application 92
- caractéristique,
 - polynôme 33
 - équation (d'une éq. différentielle) 86
- carbonne 14, datation au 7
- carrée, matrice 20
- champ de vecteurs 78
- changement de base 31
- chaotique, système 12
- coefficient de croissance 7
- coefficients d'une matrice 19
- combinaison linéaire 17
- complexes, nombres 97
- composantes principales 103
- conjugué, d'un nombre complexe 97
- continuité, principe de 94
- coordonnées 18
 - dans une base quelconque, 31
 - polaires 74
- corps 93
- cosinus, théorème du 16
- Cramer, règle de 27
- Cramer, système de 26
- cryptographie 92
- cycle 8, 81
 - stable 8
- débit 77
- degré, d'un polynôme 99
- demi-vie 7
- dérivable, fonction 46
- dérivée
 - condition sur la 9
 - directionnelle 59
- déterminant 20
 - interprétation géométrique 25
- diagonale, matrice 22
- diagonalisable
 - application linéaire 34
 - matrice 34
- différentiable, fonction
 - d'une variable, en un point 45,
 - d'une variable, globalement 47
 - de plusieurs variables 53, 54
- dimension, d'un espace vectoriel 18
- directionnelle, dérivée 59
- entiers naturels 92
- entiers relatifs 92
- équation différentielle 80
 - à variables séparables 53
 - équation caractéristique d'une 86
 - linéaire inhomogène 83
 - d'ordre supérieur 85
 - ordre d'une 80, 84
 - système d' 80, 84
- équilibre, état d' 8, 80
- espace vectoriel
 - réel 101
 - complexe 101
- états d'un système dynamique 4, 80
- Euler 39, 99
- exponentielle, croissance 6, 87
- extremum local 47, 59

Fatou, théorème de 10
 Fibonacci, suite de 42
 Fibonacci, système de 41
 fonction 91
 formel, calcul 96
 gaussienne 77
 gradient 52
 Gram, matrice de 103
 graphe, d'une application 92
 gravité, centre de 75
 harmonique, oscillateur 86
 hessienne, matrice 56
 homogène, système d'équations 27
 homothétie 29, 32
 identité, matrice 23
 imaginaire, partie 97
 imprévisibilité 13
 indépendants, vecteurs linéairement 17
 inégalité de Schwarz 16, 102
 inégalité du triangle 16
 injective, application 92
 intégrale 65
 d'une gaussienne 77
 de surface 71
 de volume 74
 intégration
 numérique 70
 par parties 69
 par substitution 67
 invariant, ensemble 81
 inverse d'une matrice 23
 inversible, matrice 23
 itération 4
 graphique 5
 jacobienne, matrice 53, 81
 latent root : notice p. 32
 Leslie, matrice de 40
 L'Hospital, règles de 48
 libre, famille 17
 linéaire
 algèbre 14
 application 14, 29
 combinaison 17
 raisonnement 14
 régression 63
 linéairement indépendants, vecteurs 17
 logistique, croissance 10, 88
 loi d'évolution 4, 80
 longueur d'une courbe 75
 MacLaurin, formule de 50
 matrice 19
 carrée 20
 de Gram 103
 de Leslie 40
 de passage 31
 d'une application linéaire 30
 relativement à une base 32
 identité 23
 symétrique 37
 transposée 20
 d'un système d'équations 26
 maximum local 47, 59
 météorologie 13, 12
 mineur 21
 ordre d'un 24
 principal 37
 minimum local 47, 59
 module, d'un nombre complexe 98
 moindres carrés, méthode des 63
 moyenne, théorème de la 66
 moyenne, valeur 66, 71, 74
 naturels, entiers 92
 niveau, surface, courbe 58
 nombre d'or 42
 nombres
 complexes 97
 rationnels 93
 réels 94
 transcendants 96
 norme (d'un vecteur) 16
 or, nombre d' 42
 orbite 4
 ordre, d'une approximation 49
 orthogonaux, vecteurs 17
 oscillateur harmonique 86
 amorti 87
 partielle, dérivée 51
 parties, intégration par 69
 passage, matrice de 31
 pente 45

périodique, état 8, 81
 polaires, coordonnées 74
 polynôme 99
 polynôme
 degré d'un 99
 division avec reste d'un 100
 racine d'un 99
 potentiel 53, 76
 différence de 76
 primitive d'une fonction 66
 principal, mineur 37
 principales, composantes 103
 produit de matrices 19
 produit scalaire 15
 proies-prédateurs, système 89
 projection sur un plan 33, 36
 propre, valeur
 d'une application linéaire 32
 d'une matrice 32
 propre, vecteur
 d'une application linéaire 32
 d'une matrice 32
 racine
 de l'unité 100
 d'un polynôme 99
 radioactivité 6
 rang
 d'un système 27
 d'une matrice 25,
 rationnels, nombres 93
 rayon spectral 38
 réelle, partie 97
 réels, nombres 94
 régression linéaire 63
 relatifs, entiers 92
 rotation 31, 32, 34
 scalaire 101
 Schwarz, inégalité de 16, 102
 sensibilité aux conditions initiales 12
 Simpson, formule de 70
 spectral, rayon 38
 spectre
 d'une application linéaire 33
 d'une matrice carrée 33
 sous-espace vectoriel 101
 stable,
 cycle 8
 ensemble invariant 81
 état stationnaire 8
 standard, base de \mathbb{R}^n 18
 stationnaire, état 8, 80
 substitution, intégration par 67
 surjective, application 92
 symétrique, matrice 37
 système dynamique
 continu 80
 discret 3
 linéaire 37
 taille d'une matrice 19
 Taylor
 reste 50
 formule de (1 variable) 50
 formule de (n variables) 57
 temps de doublement 7
 tire-bouchon, règle du 25
 trapèzes, formule des 70
 trajectoire
 d'un champ de vecteurs 78
 d'un système continu 80
 d'un système discret 4
 transcendants, nombres 96
 transposée, matrice 20
 travail 76
 traversement, lemme de 46
 triangle, inégalité du 16
 triangulaire, matrice 22
 unité, racines de l' 100
 variables séparables 53
 vecteur 101
 colonne 19
 ligne 19
 vitesse 56
 vectoriel, espace, voir "espace vectoriel"