Antonin Guilloux

1M002 : SUITES ET INTÉGRALES, ALGÈBRE LINÉAIRE
GROUPE MIPI 25 – ANNÉE 2017
PARTIE II : CALCUL MATRICIEL

Antonin Guilloux

1M002 : SUITES ET INTÉGRALES, ALGÈBRE LINÉAIRE
GROUPE MIPI 25 – ANNÉE 2017
PARTIE II : CALCUL MATRICIEL

Antonin Guilloux

Avertissement

Ce polycopié suit mon cours 1M002 dispensé au groupe MIPI 25 en 2017. Ce cours de mathématiques générale du second semestre du L1 2016/2017 à l'UPMC couvre les thèmes des suites et intégrales et une introduction à l'algèbre linéaire.

Différentes sources d'inspiration m'ont servi pour rédiger ce polycopié, notamment plusieurs notes de cours de collègues (par exemple Patrick Polo, Claire David, Sophie Chemla, Jean-Lin Journé...). Je tiens à remercier tout particulièrement Patrick Polo, dont les conseils m'ont été très utiles. Certaines sections sont en grande partie reprises du polycopié qu'il a lui-même rédigé en 2013/2014. Les discussions avec Laurent Charles et Pierre Charollois m'ont aussi permis d'améliorer ce texte.

Je remercie aussi toutes les personnes qui m'ont signalé des coquilles dans des versions précédentes de ce texte.

Les questions de cours sont indiquées par le symbole (QC) en marge : (QC)

TABLE DES MATIÈRES

CHAPITRE 1

INTRODUCTION : UN PEU DE LOGIQUE

Les mathématiques demandent une grande rigueur dans l'exposition des résultats et des démonstrations. Si l'intuition et la persuasion sont des outils indispensables dans la recherche de solutions à des problèmes, si l'aspect de jeu d'énigmes peut-être une motivation pour faire des maths, il n'en reste pas moins que l'étape finale d'un travail est une rédaction rigoureuse des résultats et des raisonnements.

Le but de cette rigueur est de permettre une vérification au lecteur. Dans l'idéal, il ne doit pas y avoir d'ambiguïtés dans la rédaction et le lecteur doit savoir exactement ce que l'auteur du texte a en tête. Un point qui n'est pas complètement évident au premier abord est qu'une rédaction rigoureuse dépend du lecteur. Par exemple, considérons le texte suivant

> La fonction $x \mapsto x^2 - 2$ est définie et continue sur $\mathbf{R}$, vaut -1 pour $x = 1$ et 2 pour $x = 2$. D'après le théorème des valeurs intermédiaires, elle s'annule sur l'intervalle $[1, 2]$.

Normalement, il doit vous apparaître comme rigoureux et vous devez être convaincus. Mais, bien sûr, si je dis la même chose à un élève de seconde, il ne peut pas considérer cette rédaction comme rigoureuse, car il ne connaît pas les définitions et le théorème auxquels elle fait référence. Il faudra que je trouve un autre moyen de le convaincre de l'existence de $\sqrt{2}$ [1]. Dans le cadre de ce cours, je considère que vous connaissez les définitions et résultats du cours 1M001. Au fur et à mesure, nous verrons de nouvelles définitions et

1. Par exemple, d'après le théorème de Pythagore, c'est la longueur de la diagonale du carré de côté 1. C'est rigoureux pour l'élève de seconde – et même pour nous, ça donne beaucoup plus d'informations et d'intuition !

de nouveaux résultats qui permettrons de dire de nouvelles choses de manière rigoureuse.

Cette rigueur repose notamment sur une utilisation très précise du langage (et on parle d'ailleurs de langage mathématique pour le différencier du langage courant), sur des règles de logique et de raisonnements bien établies et bien comprises et sur une grande attention aux définitions des objets et des cadres de travail et à l'exposition des démonstrations.

Cette présente introduction cherche à expliciter les deux premiers points : quelle est cette précision dans le langage qu'on recherche, et quels sont les outils logiques dont on dispose. Quant au troisième point, il sera illustré tout au long du semestre par la rédaction même du cours.

Une des dimensions de votre formation en mathématiques est d'apprendre à maîtriser cette utilisation du langage et de la logique dans votre pratique des mathématiques. Vous devez vous approprier ces façons de mettre en forme vos raisonnements pour pouvoir les expliquer à d'autres.

1.1. Énoncés mathématiques

1.1.1. Propositions et énoncés. — Une proposition est tout simplement une affirmation grammaticalement correcte :

Exemples 1.1.1. —

1. x est positif.
2. La fonction est continue.

On ne peut pas décider si ces affirmations sont vraies ou fausses : tous les termes ne sont pas bien définis. Par exemple, pour la première des deux, ça dépend de qui est x.

Un énoncé est une affirmation dont tous les mots sont bien définis (il ne doit pas y avoir d'ambiguïté). Il est vrai ou faux. Par exemple :

Exemples 1.1.2. —

1. Tout nombre réel positif est le carré d'un nombre réel.
2. Considérons un nombre rationnel positif. Alors il est le carré d'un nombre rationnel.
3. Toute fonction continue sur $\mathbf{R}$ est positive.

Pour construire des énoncés ou propositions plus compliqués, nous utilisons des *connecteurs logiques* : la conjonction ("et"), la disjonction ("ou"), l'implication, l'équivalence et la négation.

E1 / E2	Vrai	Faux
Vrai	Vrai	Faux
Faux	Faux	Faux

TABLE 1. Table de vérité de « E1 et E2 »

Considérons donc deux énoncés $E1$ et $E2$.

1.1.2. Conjonction. — La conjonction de E1 et E2 est l'énoncé « E1 et E2 ». Par exemple :

Exemples 1.1.3. —

1. Tout nombre réel positif est le carré d'un nombre réel et tout nombre réel est le cube d'un nombre réel.
2. Tout nombre réel est le carré d'un nombre réel et tout nombre réel est le cube d'un nombre réel.

La conjonction « E1 et E2 »est vraie quand les deux énoncés sont vrais. Si l'un des deux est faux ou les deux sont faux, la conjonction est fausse. On peut résumer ceci par la table de vérité de la conjonction 1. Parmi les deux exemples, certains sont-ils vrais ?

1.1.3. Disjonction. — La disjonction de E1 et E2 est l'énoncé « E1 ou E2 ». Par exemple :

Exemples 1.1.4. —

1. Tout nombre réel est positif ou tout nombre réel est négatif.
2. Tout nombre réel est le carré d'un nombre réel ou tout nombre réel est le cube d'un nombre réel.

La disjonction « E1 ou E2 »est vraie quand l'un des deux énoncés est vrai ou les deux sont vrais. Si les deux sont faux, la disjonction est fausse. On pourra se reporter la la table 2 de vérité de la disjonction. Parmi les deux exemples, certains sont-ils vrais ?

1.1.4. Implication et équivalence. — L'implication est l'énoncé « Si E1, alors E2 » qu'on peut dire aussi « E1 implique E2 »(noté « E1 $\Rightarrow$ E2 »).

Exemples 1.1.5. —

1. 2 est positif implique 2 est un carré.

E2 \ E1	Vrai	Faux
Vrai	Vrai	Vrai
Faux	Vrai	Faux

TABLE 2. Table de vérité de « E1 ou E2 »

E2 \ E1	Vrai	Faux
Vrai	Vrai	Vrai
Faux	Faux	Vrai

TABLE 3. Table de vérité de « E1 implique E2 »

E2 \ E1	Vrai	Faux
Vrai	Vrai	Faux
Faux	Faux	Vrai

TABLE 4. Table de vérité de « E1 équivaut à E2 »

2. 3 est négatif implique -2 est un carré.
3. 3 est positif implique -2 est un carré.

L'implication « E1 implique E2 » est vraie si E1 et E2 sont vrais ou si E1 est faux (peu importe la vérité de E2) : du faux on peut déduire n'importe quoi. On se reporter à la table de vérité 3

Attention, sur l'implication nous commençons à voir des divergences entre le langage courant et le langage mathématique : en général, l'intuition nous commande de dire que le deuxième exemple ci-dessus est faux. Or, il est vrai, car le premier énoncé est faux. En plus, les exemples 2 et 3 sont très étranges car il n'y a pas de rapport entre l'énoncé de gauche et celui de droite, et ça heurte notre intuition de l'implication. Il se trouve que pour un raisonnement logique rigoureux, la définition donnée est la bonne.

L'équivalence est l'énoncé « (E1 implique E2) et (E2 implique E1) », qu'on peut dire aussi « E1 équivaut à E2 » (noté « E1 $\Leftrightarrow$ E2 »).

Vues les règles données ci-dessus, l'équivalence est vraie si les énoncés sont tous les deux vrais ou tous les deux faux. Dans les autres cas, elle est fausse, voir la table 4.

1.1.5. Négation. — La négation de l'énoncé E1 est l'énoncé « non E1 ». La négation est vraie si E1 est faux ; elle est fausse si E1 est vrai.

Ca semble simple dit comme ça, mais vous pouvez méditer sur les règles (surtout la troisième) :

1. « non(E1 et E2) » est (équivalent à) l'énoncé « non(E1) ou non(E2) ».
2. « non(E1 ou E2) » est (équivalent à) l'énoncé « non(E1) et non(E2) ».
3. « non(E1 $\Rightarrow$ E2) » est (équivalent à) l'énoncé « E1 et non(E2) ».

Pour vérifier ces règles, une bonne façon est de dresser un tableau de vérité : faire un tableau avec toutes les valeurs possibles de vérité pour E1 et E2 et verifier que dans chaque cas les deux énoncés proposés sont en même temps vrais et en même temps faux.

1.1.6. Variables et quantificateurs. — Regardons le texte suivant [2] :

> Lorsque le cube et les choses, pris ensembles, sont égaux à un nombre discret, [3], on trouve deux nombres qui diffèrent de celui-là [4] tels que leur produit soit toujours égal au cube du tiers des choses nettes [5]. Le reste alors, en règle générale, de la soustraction bien réalisée de leurs racines cubiques est égal à ta chose principale [6].
>
> Dans le second de ces actes [7], lorsque le cube reste seul [8], tu observeras ces autres accords : tu diviseras immédiatement le nombre en deux parties, de sorte que l'une multipliée par l'autre donne clairement et exactement le cube du tiers de la chose [9]. Ensuite, de ces deux parties, selon une règle habituelle, tu prendras les racines cubiques ajoutées ensembles, et cette somme sera ton résultat [10].
> [...]

Il est très difficile de le comprendre, en tous cas pour nous lecteurs modernes. Nous préférerions lire et écrire :

---

2. Tiré d'une lettre de 1546 de Tartaglia à Cardan, traduction personnelle.
3. C'est à dire qu'on veut résoudre une équation $x^3 + px = q$.
4. On cherche u et v tels que $u - v = q$.
5. On doit avoir aussi $uv = \left(\frac{p}{3}\right)^3$.
6. Autrement dit, une solution est $\sqrt[3]{u} - \sqrt[3]{v}$.
7. On étudie un deuxième cas, dans une liste décrite plus haut dans la lettre.
8. C'est à dire que l'équation est maintenant $x^3 = px + q$. Les nombres négatif étaient suspects ! p et q désignent toujours des nombres positifs.
9. Maintenant, on cherche u et v tels que $u + v = q$ et $uv = \left(\frac{p}{3}\right)^3$.
10. La solution est alors $\sqrt[3]{u} + \sqrt[3]{v}$.

Pour résoudre l'équation $z^3 + pz = q$, on peut chercher u et v tels que :

$$\begin{cases} u - v & = q \\ uv & = \left(\frac{p}{3}\right)^3 \end{cases}$$

Alors (dans le cas où u et v sont réels), la différence de leurs racines cubiques donnera une racine de l'équation initiale.

Pour résoudre l'équation $z^3 = pz + q$, on peut chercher u et v tels que :

$$\begin{cases} u + v & = q \\ uv & = \left(\frac{p}{3}\right)^3 \end{cases}$$

Alors (dans le cas où u et v sont réels), la somme de leurs racines cubiques donnera une racine de l'équation initiale.

Exercice : essayer de comprendre que c'est effectivement une traduction du texte ci-dessus. Ce texte décrit partiellement une méthode pour résoudre les équations du troisième degré, appelée méthode de Cardan. Elle est dûe à Cardan qui repère une erreur dans la méthode ci-dessus et la corrige.

Nous avons utilisé des variables : nous avons donné un nom à des objets qu'il est long de décrire avec des mots. N'oubliez jamais que ce n'est que ça : un nom commode donné à des objets qu'on pourrait décrire avec des mots. Les énoncés mathématiques sont avant tout des phrases !

Une autre abréviation dont nous nous servons couramment sont les quantificateurs : nous utilisons le symbole $\forall$ pour dire « pour tout »et le symbole $\exists$ pour dire « il existe ».

Exemples 1.1.6. —

1. $\exists x \in \mathbf{R}, x > 0$ se lit « il existe un nombre réel positif ».
2. $\forall y \in [0, +\infty[, \exists z \in \mathbf{R}, y = z^2$ se lit « pour tout nombre réel positif, il existe un nombre réel dont le premier est le carré » et surtout se *comprend* comme l'affirmation « tout nombre réel positif est un carré ».

Avec toutes ces règles (et la connaissance de quelques lettres grecques), vous pouvez lire la définition de la continuité en un point x_0 d'une fonction f définie sur un intervalle I de $\mathbf{R}$:

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon$$

Un bon exercice : une fois que vous l'avez lue, essayez de la comprendre, au sens de la dire avec une phrase française la plus compréhensible possible.

1.1.7. Négation, encore. — Il nous faut revenir sur la négation, pour comprendre comment nier des phrases compliquées. Tout d'abord un avertissement : il n'est pas si facile de nier une phrase. Cependant, il y a une méthode automatique, qui n'aide pas beaucoup la compréhension mais qui permet d'arriver au bon résultat.

Les règles données plus haut expliquaient comment nier des phrases liées par un connecteur : si on sait nier E1 et nier E2, alors on sait nier « E1 et E2 » : c'est l'énoncé « non(E1) ou non(E2) ».

Pour nier un énoncé qui commence par un quantificateur, il suffit de l'échanger avec l'autre quantificateur, sans changer la condition sur la variable s'il y en a une. Si P est une proposition et A un ensemble, la négation de « $\forall x \in A$, P » est « $\exists x \in A$, non(P) »et inversement.

Par exemple, la suite des étapes nécessaires pour nier la définition de la continuité donnée plus haut est :

1. non$(\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$ est (équivalent à) :
2. $\exists \varepsilon > 0$, non$(\exists \eta > 0, \forall x \in I, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$ qui est (équivalent à) :
3. $\exists \varepsilon > 0, \forall \eta > 0$, non$(\forall x \in I, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$ qui est (équivalent à) :
4. $\exists \varepsilon > 0, \forall \eta > 0, \exists x \in I$, non$(|x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$ qui est (équivalent à) :
5. $\exists \varepsilon > 0, \forall \eta > 0, \exists x \in I$, $|x - x_0| \leq \eta$ et non$(|f(x) - f(x_0)| \leq \varepsilon)$ qui est (équivalent à) :
6. $\exists \varepsilon > 0, \forall \eta > 0, \exists x \in I$, $|x - x_0| \leq \eta$ et $|f(x) - f(x_0)| > \varepsilon$.

Il ne reste maintenant qu'à comprendre ce qui signifie cet ultime énoncé !

1.2. Raisonnements

Nous passons ici en revue certains types de raisonnements qui nous serons utiles. Nous en verrons plusieurs exemples au cours du semestre.

1.2.1. Par récurrence. — Vous le connaissez : il s'agit de démontrer un énoncé du type « pour tout entier n, $P(n)$ » où P est une propriété de n.

Pour ça on *initialise* : on montre $P(0)$.

Ensuite, on montre l'*hérédité* ou *propagation* : on montre pour tout entier n l'implication $P(n) \Rightarrow P(n+1)$.

Alors l'énoncé « pour tout entier n, $P(n)$ » est prouvé.

Nous verrons beaucoup d'exemples de tels raisonnements.

1.2.2. Par la contraposée. — On a deux énoncés A et B et on veut démontrer que A $\Rightarrow$ B.

Pour ce faire, on remarque (par un tableau de vérité !) que l'énoncé A $\Rightarrow$ B est équivalent à l'énoncé non(B) $\Rightarrow$ non(A).

Et c'est ce dernier qu'on cherche à montrer.

Pour valider ce raisonnement, on écrit la table de vérité des deux implications et on constate qu'elles sont identiques

Exemple 1.2.1. — Montrons pour tout entier n que : n^2 est pair $\Rightarrow$ n est pair.

On écrit l'implication contraposée : n est impair $\Rightarrow$ n^2 est impair. C'est cela maintenant qu'on veut montrer.

C'est facile : si n est impair, on peut écrire $n = 2k + 1$ pour k un entier. Ainsi $n^2 = 4k^2 + 4k + 1$ est bien un nombre impair.

1.2.3. Par l'absurde. — C'est un raisonnement très proche de la contraposée. On veut montrer, sous certaines hypothèses, l'énoncé A.

Pour ça, on suppose, sous les mêmes hypothèses, que l'énoncé non(A) est vrai, et on cherche une contradiction.

Exemple 1.2.2. — Montrons que $\sqrt{2}$ est irrationnel.

On veut montrer qu'il est impossible de trouver deux rationnels p et q sans facteurs communs tels que $\sqrt{2} = \frac{p}{q}$.

Pour ça, on raisonne par l'absurde : on suppose qu'il existe deux entiers p et q, qui ne sont pas tous les deux pairs car sans facteurs communs, tels que $\sqrt{2} = \frac{p}{q}$.

En mettant au carré, on obtient $q^2 = 2p^2$. Donc q^2 est pair et d'après l'exemple précédent, q est pair : $q = 2q_0$ pour un entier q_0. En remplaçant, on obtient $p^2 = 2q_0^2$, donc p^2 est pair. Finalement q et p sont pairs : c'est la contradiction recherchée.

1.3. Quelques notations sur les ensembles

Passons en revue quelques notations sur les ensembles, sans vraiment nous poser la question de ce qu'est un ensemble – ça nous emmènerait trop loin ! On suppose donc qu'on a deux ensembles A et B.

1. L'intersection de A et B est l'ensemble $A \cap B$ des éléments qui sont à la fois dans A et dans B. Autrement dit, pour tout x, $x \in A \cap B$ équivaut à $x \in A$ et $x \in B$.
2. L'union de A et B est l'ensemble $A \cup B$ des éléments qui sont dans A ou dans B (ou les deux). Autrement dit, pour tout x, $x \in A \cup B$ équivaut à $x \in A$ ou $x \in B$.
3. L'ensemble A est inclus dans B, noté $A \subset B$, si tout élément de A est aussi élément de B. Autrement dit $A \subset B$ équivaut à : pour tout x, $x \in A \Rightarrow x \in B$.

Exemples 1.3.1. —

1. $]-\infty, 1] \cap [0, +\infty[$ est l'ensemble $[0, 1]$.
2. $]-\infty, 1] \cup [0, +\infty[$ est l'ensemble $\mathbf{R}$.
3. $]1, 2[\subset]0, +\infty[$ est vraie.
4. $\{k\pi, \text{ pour } k \in \mathbf{Z}\} \subset [0, +\infty[$ est faux. Notons que le premier ensemble est l'ensemble des multiples (positifs ou négatifs) de π.

PARTIE I

LES SUITES

CHAPITRE 2

SUITES RÉELLES ET COMPLEXES

Les suites sont un objet fondamental à la fois en mathématiques et dans l'application des mathématiques aux autres sciences. Nous verrons dans ce cours et les travaux dirigés divers exemples : approximation d'un nombre irrationnel par des décimaux ; suite de Syracuse ; algorithme de Newton pour approcher les racines d'un polynôme ; modélisation des prêts bancaires.

Ce chapitre commence principalement par des rappels. Ce sera le prétexte pour réintroduire sur divers exemples les nombres complexes et leurs propriétés.

2.1. Suites réelles ou complexes ; opérations sur les suites

On fixe le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$.

Définition 2.1.1. — Une suite à valeurs[1] dans $\mathbf{K}$ est une application

$$u \,:\, \mathbf{N} \to \mathbf{K}.$$

On note $u = (u_n)_{n\in\mathbf{N}}$ ou $u = (u_0, u_1, \ldots, u_n, \ldots)$.
L'ensemble des suites est noté $\mathcal{S}_{\mathbf{K}}$ ou $\mathbf{K}^{\mathbf{N}}$.
Le nombre u_n est appelé n-ième terme de la suite u.

Il arrive parfois que la suite ne soit pas définie pour tous les entiers de $\mathbf{N}$, mais seulement pour un sous-ensemble d'indices $I \subset \mathbf{N}$. Une suite définie sur I est une application $u \,:\, I \to \mathbf{K}$. On la note $u = (u_n)_{n\in I}$.

On peut définir une suite par :

– une formule ; par exemple la suite définie par $u_n = n^2 + \cos(n)$.

1. On dit plus simplement suite réelle si $\mathbf{K} = \mathbf{R}$ et complexe si $\mathbf{K} = \mathbf{C}$.

– un processus de construction (suite définie par récurrence) ; dans ce cas, on donne u_0, puis on explique comment construire u_{n+1} à partir de u_n. Autrement dit, la suite vérifie une relation de récurrence $u_{n+1} = f(u_n)$. Par exemple, les suites de Syracuse : on choisit u_0 égal à un entier quelconque et on pose $u_{n+1} = \frac{u_n}{2}$ si u_n est pair et $u_{n+1} = 3u_n + 1$ si u_n est impair. Ces suites peuvent être très compliquées à étudier. Par exemple, pour les suites de Syracuse, on conjecture [2] que pour tout choix de u_0, la suite passera par 1 (c'est-à-dire qu'il existe n avec $u_n = 1$). Par exemple, si $u_0 = 4$, alors la suite est $(4, 2, 1, 4, 2, 1 \ldots)$. En revanche, si $u_0 = 15$, alors la suite est

$$(15, 46, 23, 70, 35, 106, 53, 160, 80, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, \ldots).$$

– Une récurrence étendue : on se donne les p premiers termes, et on explique comment construire u_{n+p} à partir de u_n, u_{n+1}, …, u_{n+p-1}.

2.1.1. Somme et produit par un scalaire. —

Définition 2.1.2. — Soit u et v deux éléments de $\mathcal{S}_{\mathbf{K}}$ et $\lambda \in \mathbf{K}$. On définit :

– La suite $u + v$ par $u + v = (u_n + v_n)_{n\in\mathbf{N}}$.
– La suite $\lambda \cdot u$ par $\lambda \cdot u = (\lambda u_n)_{n\in\mathbf{N}}$

On réinterprétera plus tard cette définition en disant que $\mathcal{S}_{\mathbf{K}}$ est un *espace vectoriel*

2.1.2. Produit de suites. — On sait aussi multiplier les suites entre elles. On sort ici du cadre des espaces vectoriels pour entrer dans celui des $\mathbf{K}$-*algèbres*.

Définition 2.1.3. — Soient u et v deux suites à valeurs dans $\mathbf{K}$. On définit leur produit uv par $uv = (u_n v_n)_{n\in\mathbf{N}}$.

2.2. Quelques exemples

2.2.1. Les suites arithmétiques. —

Définition 2.2.1. — Une suite $u = (u_n)_{n\in\mathbf{N}}$ est dite *arithmétique* de raison $r \in \mathbf{K}$ si elle vérifie la relation de récurrence :

(QC)

$$u_{n+1} = u_n + r.$$

2. C'est-à-dire on pense que c'est vrai, mais qu'on ne sait pas le prouver. Pour ces suites, on l'a vérifié numériquement pour tous les u_0 jusqu'à au moins 10^{18} ! Le lecteur pourra aller voir la suite d'articles consacrés à cette suite sur le magnifique site Image des mathématiques : premier article, deuxième article, troisième article.

On montre alors, par récurrence :

Proposition 2.2.2. — *Soit (u_n) une suite arithmétique de raison r.* (QC)

- *Son terme général est $u_n = u_0 + nr$.*
- *La somme de termes consécutifs est donnée par :*

$$\sum_{k=p}^{n} u_k = (n-p+1)\frac{u_p + u_n}{2}.$$

Exemples 2.2.3. —

1. La suite définie par $u_n = n$ est arithmétique de raison 1. La somme des n premiers entiers est $1 + \ldots + n = \frac{n(n+1)}{2}$. Une preuve géométrique est aussi donnée en TD.
2. On peut représenter graphiquement dans le plan complexe l'évolution de la suite arithmétique de premier terme $u_0 = 1$ et de raison $r = 1+i$. Voir à cette adresse pour une animation.

2.2.2. Les suites géométriques. —

Définition 2.2.4. — Une suite $u = (u_n)_{n\in\mathbf{N}}$ est dite *géométrique* de raison $q \in \mathbf{K}$ si elle vérifie la relation de récurrence : (QC)

$$u_{n+1} = qu_n.$$

On montre alors, par récurrence :

Proposition 2.2.5. — *Soit (u_n) une suite géométrique de raison q.* (QC)

- *Son terme général est $u_n = u_0 q^n$.*
- *Si $q \neq 1$, la somme de termes consécutifs est donnée par :*

$$\sum_{k=p}^{n} u_k = u_p \frac{1-q^{n-p+1}}{1-q}.$$

Exemple 2.2.6. — On peut représenter graphiquement l'évolution d'une suite géométrique de premier terme $u_0 = 1$ et de raison $q \in \mathbf{C}$. Voir cette adresse pour une animation interactive. On peut traiter le cas $q = 1 + i$.

Tout d'abord, pour multiplier par un nombre complexe, il vaut mieux le mettre sous notation exponentielle : on remarque que

$$1 + i = \sqrt{2}\left(\cos\left(\frac{\pi}{4}\right) + i\sin\left(\frac{\pi}{4}\right)\right) = \sqrt{2}e^{i\frac{\pi}{4}}.$$

Donc la multiplication par $1 + i$ se traduit géométriquement par une similitude : on fait une homothétie de rapport $\sqrt{2}$ puis une rotation d'un quart de tour.

2.2.3. Les suites arithmético-géométriques. —

Définition 2.2.7. — Une suite $u = (u_n)_{n\in\mathbf{N}}$ est dite *arithmético-géométrique* de raisons $r, q \in \mathbf{K}$ si elle vérifie la relation de récurrence :

$$u_{n+1} = qu_n + r.$$

Proposition 2.2.8. — *On suppose $q \neq 1$* [3]*. Le terme général d'une suite arithmético-géométrique de raisons q et r est :*

$$u_n = q^n u_0 + \frac{r(q^n - 1)}{q - 1}.$$

Démonstration. — Modifions la suite u de manière à trouver une suite géométrique : définissons la suite v par

$$v_n = u_n + \frac{r}{q-1}.$$

Alors on a

$$v_{n+1} = u_{n+1} + \frac{r}{q-1} = qu_n + r + \frac{r}{q-1} = qu_n + \frac{qr}{q-1} = q\left(u_n + \frac{r}{q-1}\right) = qv_n.$$

Donc la suite (v_n) est géométrique de raison q et $v_n = q^n v_0 = q^n(u_0 + \frac{r}{q-1})$. On en déduit que $u_n = v_n - \frac{r}{q-1} = q^n u_0 + \frac{r(q^n-1)}{q-1}$. □

Exemple 2.2.9. — Un bon exemple de suites arithmético-géométriques est donné par les prêts bancaires. Imaginons qu'on emprunte 10 000 euros au taux annuel de 3% et qu'on décide de rembourser 100 euros par mois. On veut savoir combien de mois on va mettre à rembourser le prêt.

Notons u_n la somme ("le capital") restant due à la banque après n mois ($u_n = 0$ si on a fini de rembourser). On a $u_0 = 10000$. Pour calculer u_{n+1} en fonction de u_n, la règle est la suivante : les 100 euros de remboursement servent d'abord à payer les intérêts du mois sur la somme u_n, puis à rembourser le capital. Le taux mensuel d'intérêt est $3/12 = 0,25\%$. Par exemple, pour le premier remboursement, on doit commencer par payer $0,0025 * 10000 = 25$ euros d'intérêts, et on rembourse donc 75 euros de capital. Ainsi, $u_1 = 10000 + 25 - 100 = 9925$. Plus généralement, on obtient la relation de récurrence :

$$u_{n+1} = u_n + 0,0025u_n - 100 = 1,0025u_n - 100.$$

La suite $(u_n)_{n\in\mathbf{N}}$ est donc arithmético-géométrique, de raison $q = 1,0025$ et $r = -100$, tant qu'on n'a pas fini de rembourser. La proposition précédente nous donne la formule :

$$\begin{aligned} u_n = (1,0025)^n u_0 - 100\frac{1,0025^n - 1}{1,0025 - 1} &= (1,0025)^n 10000 - 40000(1,0025^n - 1) \\ &= 40000 - 30000 \times 1,0025^n. \end{aligned}$$

3. Sinon, la suite est arithmétique.

Le nombre de mois nécessaire au remboursement est le plus petit entier n tel que $40000 - 30000 \times 1,0025^n \leq 0$. Ce nombre est négatif pour $n \geq \frac{\ln(4/3)}{\ln(1,0025)} \simeq 115,2$ (exercice !). Donc le nombre de mois nécessaire au remboursement est 116 (presque 10 ans).

On observe que si on décide de rembourser 200 euros par mois ($r = -200$), la durée est 54 mois – soit moins de la moitié. En revanche, si on ne rembourse que 50 euros par mois, alors la durée est 278 mois – soit plus du double. Et on a un problème si on ne veut rembourser que 25 euros par mois : chaque mois, ces 25 euros partent intégralement pour payer les intérêts et on ne rembourse jamais le capital.

2.3. Suites bornées, convergentes

2.3.1. Suites bornés, majorées, monotones. — On rappelle que le *module* d'un nombre complexe $z = a + ib$ est $|z| = \sqrt{a^2 + b^2}$. Si z est réel (donc $z = a$), alors son module est sa valeur absolue. Le module vérifie quelques propriétés utiles :

- on a $|a| \leq |z|$ et $|b| \leq |z|$;
- on a, si λ est réel, $|\lambda z| = |\lambda||z|$;
- on a l'*inégalité triangulaire* : pour tous $z = a+ib$, $z' = a'+ib'$ complexes, $|z + z'| \leq |z| + |z'|$ et $|z - z'| \geq |z| - |z'|$.

Définition 2.3.1. —
- Une suite réelle ou complexe u_n est dite *bornée* si $\exists M \in \mathbf{R}, \forall n \in \mathbf{N}, |u_n| \leq M$. (QC)
- Une suite réelle est majorée si $\exists M \in \mathbf{R}, \forall n \in \mathbf{N}, u_n \leq M$.
- Une suite réelle est minorée si $\exists m \in \mathbf{R}, \forall n \in \mathbf{N}, u_n \geq m$.

Proposition 2.3.2. — *Si u et v sont deux suites à valeurs dans $\mathbf{K}$ bornées, et $\lambda \in \mathbf{K}$, alors les suites $u + v$, $\lambda \cdot u$ et uv sont encore bornées.*

Remarque 2.3.3. — Dans le langage qu'on mettra en place tout au long de ce semestre, les deux premières propriétés s'expriment en disant que l'espace des suites bornées est un sous-espace vectoriel de $\mathcal{S}_{\mathbf{K}}$.

Démonstration. — Prouvons la première assertion, c'est à dire que $u + v$ est bornée, les autres sont similaires : Soit u et v deux suites complexes bornées.

On veut montrer que $u + v$ est bornée, c'est-à-dire qu'on cherche un réel M tel que pour tout entier n on ait $|u_n + v_n| \leq M$.

Or u et v sont bornées. Il existe donc deux réels M_u et M_u tels que pour tout n, $|u_n| \leq M_u$ et $|v_n| \leq M_u$. Considérons la suite $u + v$. Le module de son n-ième terme vérifie, grâce à l'inégalité triangulaire :

$$|u_n + v_n| \leq |u_n| + |v_n| \leq M_u + M_v.$$

Le réel $M = M_u + M_v$ convient donc. La suite $u+v$ est donc bien bornée. □

De plus, une suite complexe est bornée si et seulement si sa partie imaginaire et sa partie réelle sont bornées :

Proposition 2.3.4. — *Soit $u = (u_n)_{n\in\mathbf{N}}$ une suite complexe, et notons $a = (a_n = \mathrm{Re}(u_n))_{n\in\mathbf{N}}$ sa partie réelle et $b = (b_n = \mathrm{Im}(u_n))_{n\in\mathbf{N}}$ sa partie imaginaire.*

(QC)

Alors u est bornée si et seulement si a et b le sont.

Démonstration. — Supposons u bornée, et soit $M \in \mathbf{R}$ tel que pour tout n, $|u_n| \leq M$. Comme pour tout n on a $|a_n| \leq |u_n| \leq M$ et $|b_n| \leq |u_n| \leq M$, les suites a et b sont bornées.

Réciproquement, supposons a et b bornées. Alors il existe un réel $M > 0$ tel que pour tout n on a $|a_n| \leq M$ et $|b_n| \leq M$. Or, $|u_n| = \sqrt{a_n^2 + b_n^2} \leq \sqrt{M^2 + M^2} \leq \sqrt{2}M$. Donc u est bornée. □

Rappelons la définition de suites monotones (seulement dans le cas des suites réelles, ça n'a pas de sens pour les suites complexes).

Définition 2.3.5. — Pour une suite réelle u, on dit que :

- u est croissante si pour tout $n \geq 0$, $u_{n+1} \geq u_n$;
- u est décroissante si pour tout $n \geq 0$, $u_{n+1} \leq u_n$;
- u est monotone si elle est croissante ou décroissante.

On ajoute l'adjectif *strictement* si les inégalités sont strictes.

2.3.2. Suites convergentes. — Vous connaissez la définition :

Définition 2.3.6. — Une suite réelle ou complexe u est dite *convergente* si il existe un nombre réel ou complexe l tel que :

(QC)

$$\forall \varepsilon > 0, \exists N \in \mathbf{N}, \forall n \geq N, |u_n - l| \leq \varepsilon.$$

Le nombre $l = \lim(u)$ est appelé la limite de la suite u. On dit aussi que u tend vers l, noté $u_n \xrightarrow{n\to\infty} l$.

Proposition 2.3.7. — *Soit $(u_n)_{n\in\mathbf{N}}$ une suite de $\mathcal{S}_\mathbf{K}$ convergente vers $l \in \mathbf{K}$. Alors*

1. *La suite est bornée.*
2. *La suite $(u_{n+1} - u_n)_{n\in\mathbf{N}}$ tend vers 0.*

Démonstration. — 1. On utilise la définition de convergence avec $\varepsilon = 1$: il existe N entier tel que pour tout $n \geq N$, $|u_n - l| \leq 1$. En utilisant l'inégalité triangulaire, on obtient $|u_n| \leq |l| + 1$. Donc la suite est bornée par $\max\{|u_0|, |u_1|, \ldots, |u_N|, |l| + 1\}$.

2. Soit $\varepsilon > 0$. Soit N tel que pour tout $n \geq N$, $|u_n - l| \leq \varepsilon$. Alors, pour tout $n \geq N$, on a $|u_{n+1} - u_n| \leq |(u_{n+1} - l) + (l - u_n)| \leq |u_{n+1} - l| + |u_n - l| \leq 2\varepsilon$. Ca montre bien la convergence vers 0.

□

Quand une suite n'est pas convergente, on dit qu'elle est divergente. En niant la définition précédente, on voit qu'une suite est divergente si pour tout l, il existe $\varepsilon > 0$ tel que pour tout $N \in \mathbf{N}$, il existe $n \geq N$ tel que $|u_n - l| > \varepsilon$.

Exemple 2.3.8. — Prenons le cas des suites géométriques complexes. Soient $u_0 \neq 0 \in \mathbf{C}$, $q \in \mathbf{C}$ et u_n la suite géométrique de premier terme u_0 et de raison q. Trois cas d'études se présentent :

- si $|q| > 1$, alors $|u_n| = |u_0 q^n| = |u_0||q|^n$ tend vers $+\infty$. Donc la suite u_n diverge, par contraposée du premier point de la proposition ci-dessus.
- si $|q| < 1$, alors $|u_n| = |u_0 q^n| = |u_0||q|^n$ tend vers 0. Ainsi $|u_n - 0| \to 0$, ce qui d'après la définition signifie que $u_n \to 0$.
- si $|q| = 1$: on remarque que $|u_{n+1} - u_n| = |u_0 q^n (q - 1)| = |u_0||q - 1|$ est constant. Si $q = 1$, alors la suite est constante et donc convergente.

  Si $q \neq 1$, on voit que $u_{n+1} - u_n$ ne tend pas vers 0 : la suite n'est pas convergente (à nouveau par contraposée du second point de la proposition ci-dessus).

Proposition 2.3.9. — *Soit u une suite complexe. Notons $a = (a_n = \mathrm{Re}(u_n))_{n \in \mathbf{N}}$ et $b = (b_n = \mathrm{Im}(u_n))_{n \in \mathbf{N}}$. Alors u tend vers $l = r + is$ si et seulement si a tend vers r et b tend vers s.* (QC)

Démonstration. — Supposons que $u_n \xrightarrow{n \to \infty} l$ et soit $\varepsilon > 0$. Utilisons la définition de $u_n \xrightarrow{n \to \infty} l$ pour cet ε : il existe un entier N tel que pour tout $n \geq N$, on a $|u_n - l| \leq \varepsilon$. Or on a déjà vu que $|a_n - r| \leq |u_n - l|$ et $|b_n - s| \leq |u_n - l|$. Donc, pour tout $n \geq N$, $|a_n - r| \leq \varepsilon$ et $|b_n - s| \leq \varepsilon$. Ça montre que a et b convergent, vers r et s respectivement.

Réciproquement, supposons que $a_n \xrightarrow{n \to \infty} r$ et $b_n \xrightarrow{n \to \infty} s$. Soit $\varepsilon > 0$. Utilisons la définition de la convergence de a et b pour le réel $\frac{\varepsilon}{\sqrt{2}}$: il existe $N \in N$ tel que pour tout $n \geq N$, on a $|a_n - r| \leq \frac{\varepsilon}{\sqrt{2}}$ et $|b_n - s| \leq \frac{\varepsilon}{\sqrt{2}}$. On en

déduit, pour $n \geq N$:

$$\begin{aligned} |u_n - (r+is)| &= \sqrt{(a_n - r)^2 + (b_n - s)^2} \\ &\leq \sqrt{\frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2}} \\ &\leq \varepsilon \end{aligned}$$

Ça prouve bien que u tend vers $l = r + is$. □

La limite est unique : si $u_n \to l$ et $u_n \to l'$, alors $l = l'$ (vous l'avez vu pour les suites réelles, la proposition précédente le montre pour les suites complexes).

On a aussi un résultat sur les opérations sur les limites :

(QC)

Proposition 2.3.10. — *Soient u et v deux suites convergentes, de limite respectivement l et l', et $\lambda \in \mathbf{K}$.*

Alors les suites $u + v$, $\lambda \cdot u$ et uv sont convergentes, de limites respectives $l + l'$, $\lambda \cdot l$ et ll'.

Démonstration. — Vous le savez déjà pour les suites réelles. La proposition énoncée plus haut permet de le faire pour les suites complexes. Faisons par exemple le cas du produit : si la suite $(u_n = a_n + ib_n)$ tend vers $l = a + ib$ et la suite $(v_n = c_n + id_n)$ tend vers $l' = c + id$, alors on a les quatre convergences de suites réelles $a_n \to a$, $b_n \to b$, $c_n \to c$ et $d_n \to d$. Mais le terme général de $u_n v_n$ est $a_n c_n - b_n d_n + i(a_n d_n + b_n c_n)$. Par les théorèmes d'opérations sur les limites de suites réelles et la proposition 2.3.9, $u_n v_n$ tend vers $ac - bd + i(ad + bc) = ll'$. □

Pour le cas des suites *réelles*, vous connaissez déjà des théorèmes. Tout d'abord, le théorème de convergence des suites monotones bornées :

Théorème 2.3.11. — *Toute suite réelle croissante et majorée, ou décroissante et minorée, converge.*

Et les théorèmes sur le passage à la limite dans les inégalités (attention, elles deviennent larges !) et le théorème des gendarmes :

Théorème 2.3.12. — *Soit $u = (u_n)$, $v = (v_n)$ et $w = (w_n)$ trois suites à valeurs dans $\mathbf{R}$. Alors :*

- *Si $u \xrightarrow{n\to\infty} l$, $v_n \xrightarrow{n\to\infty} l'$ et pour tout $n \in \mathbf{N}$, $u_n \leq v_n$, alors $l \leq l'$.*
- *Si $u \xrightarrow{n\to\infty} l$, $v_n \xrightarrow{n\to\infty} l'$ et pour tout $n \in \mathbf{N}$, $u_n < v_n$, alors $l \leq l'$.*
- *Si u et v convergent vers la même limite l et que pour tout $n \in \mathbf{N}$, $u_n \leq w_n \leq v_n$, alors w est convergente, de limite l.*

Enfin, un autre théorème connu est le théorème des suites adjacentes :

Théorème 2.3.13. — *Soient u et v deux suites réelles, u croissante, v décroissante, telles que pour tout n, $u_n \leq v_n$ et $v_n - u_n \to 0$.*

Alors u et v sont convergentes, et ont la même limite.

Remarquons que dans l'énoncé ci-dessus, l'hypothèse $u_n \leq v_n$ est en fait inutile : la suite $v_n - u_n$ est décroissante et tend vers 0, donc ne peut être que positive. Je préfère cependant insister sur ce point.

C'est un bon exercice de vérifier que vous savez prouver ce théorème à partir des résultats rappelés ci-dessus.

2.3.3. Calculs de limites. — Vous avez déjà déterminé la limites de certaines suites. Rappelons quelques résultats et méthodes utiles pour calculer des limites.

1. Si $u_n = \frac{P(n)}{Q(n)}$ est un rapport de deux polynômes, alors la limite (finie ou infinie) en $n \to +\infty$ est la limite du rapport des monômes de plus haut degré. Par exemple,
$$\lim_{n\to+\infty} \frac{3n^2+2}{4n^2+3n+2} = \lim_{+\infty} \frac{3n^2}{4n^2} = \frac{3}{4}.$$
2. Si u_n est de la forme $f(v_n)$ où f est continue et v_n admet une limite l dans le domaine de définition de f, alors u_n tend vers $f(l)$. Dans le cas où v_n tend vers une limite finie ou infinie l et que f admet une limite (finie ou infinie) L en l, alors u_n tend vers L. Par exemple :
$$\lim_{n\to+\infty} e^{-n} = 0.$$
$$\lim_{n\to+\infty} \ln\left(1+\frac{1}{n}\right) = 0.$$
3. Utilisation des développements limités : dans le cas d'une forme indéterminée, par exemple $u_n = n\ln\left(1+\frac{1}{n}\right)$, on peut utiliser les développements limités pour préciser l'information dont on dispose. Par exemple, on sait que pour tout $x > -1$, on peut écrire $\ln(1+x) = x + x\varepsilon(x)$, où $\varepsilon(x) \xrightarrow{x\to 0} 0$. En prenant $x = \frac{1}{n}$, on obtient
$$\ln\left(1+\frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n}\varepsilon\left(\frac{1}{n}\right).$$
Donc on peut écrire :
$$u_n = 1 + \varepsilon\left(\frac{1}{n}\right) \xrightarrow{n\to+\infty} 1.$$

D'autres exemples seront vus en TD.

2.4. Sous-suites et le théorème de Bolzano-Weierstrass

2.4.1. Sous-suites. —

Définition 2.4.1. — Soit $u = (u_n)_{n\in\mathbf{N}} \in \mathcal{S}_\mathbf{K}$ une suite. Une *suite extraite*, ou sous-suite, de u est une suite $u' = (u_{\varphi(n)})_{n\in\mathbb{N}}$ pour une fonction $\varphi : \mathbf{N} \to \mathbf{N}$ strictement croissante.

Une telle fonction s'appelle une *extraction*. Montrons que pour toute extraction, on a

$$\varphi(n) \geq n.$$

En effet, par récurrence, $\varphi(0) \geq 0$ et, si $\varphi(n) \geq n$, alors on a $\varphi(n+1) > \varphi(n) \geq n$. On obtient donc $\varphi(n+1) > n$ ce qui implique $\varphi(n+1) \geq n+1$.

Exemple 2.4.2. — Considérons la suite complexe u définie par $u_n = i^n$. On a $u = (1, i, -1, -i, 1, i, -1, -i, \ldots)$. La sous-suite des éléments d'indices divisibles par 4 est la suite $(u_{4n})_{n\in\mathbf{N}}$. Ici, l'extraction est la fonction $\varphi(n) = 4n$. Cette sous-suite est constante égale à 1.

2.4.2. Le théorème de Bolzano-Weierstrass. —

Théorème 2.4.3 **(Bolzano-Weierstrass)**. — *De toute suite réelle ou complexe bornée, on peut extraire une sous-suite convergente*

(QC)

Avertissement : la preuve qui suit diffère de celle présentée en amphi. On prouve ici d'abord le cas des suites réelles (avec la même idée que dans la preuve faite en amphi) et ensuite on étend le résultat aux suites complexes. Une raison est la suivante : la preuve présentée en amphi est efficace et se présente bien sur un dessin, mais impose des notations lourdes qu'il serait fastidieux d'utiliser à l'écrit. On prouve donc d'abord le théorème dans le cas des suites réelles :

Démonstration. — On considère donc une suite réelle bornée. Soient a un minorant et b un majorant. On construit par récurrence deux suites $(a_k)_{k\in\mathbf{N}}$ et $(b_k)_{k\in\mathbf{N}}$ telles que :

1. pour tout k, on a $a_k \leq b_k$.
2. pour tout $k \geq 1$, on a $a_{k+1} \geq a_k$ et $b_{k+1} \leq b_k$.
3. pour tout k, $b_k - a_k = \frac{b-a}{2^k}$
4. Pour tout k entier, une infinité de termes de la suite $(u_n)_{n\in\mathbf{N}}$ se trouvent dans l'intervalle $[a_k, b_k]$.

Pour ça, on commence par poser $a_0 = a$, $b_0 = b$. Ça vérifie bien les quatre conditions ci-dessus.

Supposons donc a_k et b_k construits, et construisons a_{k+1} et b_{k+1}. Pour ça, on coupe l'intervalle $[a_k, b_k]$ en deux sous intervalles : $[a_k, \frac{a_k+b_k}{2}]$ et $[\frac{a_k+b_k}{2}, b_k]$. Comme une infinité de termes de la suite $(u_n)_{n\in\mathbf{N}}$ sont dans $[a_k, b_k]$, alors dans au moins un des deux sous-intervalles, il y a une infinité de termes de la suite. Si c'est le cas pour le premier sous-intervalles, alors on pose $a_{k+1} = a_k$ et $b_{k+1} = \frac{b_k+a_k}{2}$. Sinon, on pose $a_{k+1} = \frac{b_k+a_k}{2}$ et $b_{k+1} = b_k$.

Dans les deux cas, les 4 conditions sont facilement vérifiées.

Les deux suites construites sont adjacentes : elles convergent vers une limite commune $l \in [a, b]$. Construisons maintenant une sous-suite de $(u_n)_{n\in\mathbf{N}}$ qui converge aussi vers l. Pour ça on construit par récurrence une extraction φ telle que pour tout k, $u_{\varphi(k)} \in [a_k, b_k]$:

- On pose $\varphi(0) = 0$.
- Si $\varphi(k)$ est construite, alors on définit $\varphi(k + 1)$ comme le plus petit indice $n > \varphi(k)$ tel que $u_n \in [a_{k+1}, b_{k+1}]$. C'est possible, car l'intervalle considéré contient une infinité de termes de la suite.

Alors, φ est une fonction strictement croissante de $\mathbf{N}$ dans $\mathbf{N}$; c'est donc une extraction.

Comme pour tout k, on a $a_k \leq u_{\varphi(k)} \leq b_k$, que $a_k \to l$ et $b_k \to l$, le théorème des gendarmes nous garantit que la sous-suite $(u_{\varphi(k)})_{k\in\mathbf{N}}$ converge vers l. La preuve du théorème de Bolzano-Weierstrass dans le cas réel est donc terminée. □

Pour le cas complexe, on commence par la proposition suivante :

Proposition 2.4.4. — *Soit (u_n) une suite convergente de limite l. Alors toute sous-suite de (u_n) est convergente de limite l.*

Démonstration. — Soit $u = (u_n)$ une suite qui tend vers l. Soit $(u_{\varphi(n)})$ une suite extraite ; la fonction $\varphi : \mathbf{N} \to \mathbf{N}$ est donc une fonction strictement croissante. Soit $\varepsilon > 0$. On cherche un entier N tel que pour tout $n \geq N$, $|u_{\varphi(n)} - l| \leq \varepsilon$. Pour ça, utilisons la définition de $u_n \xrightarrow{n\to\infty} l$ pour cet ε : il existe un entier M tel que pour tout $n \geq M$, $|u_n - l| \leq \varepsilon$.

On a vu que $\varphi(M) > M$. Alors, pour tout $n \geq M$, on a $\varphi(n) \geq \varphi(M) > M$ et donc $|u_{\varphi(n)} - l| \leq \varepsilon$. Ça conclut la preuve. □

On peut conclure la preuve du théorème de Bolzano-Weierstrass complexe :

Démonstration. — Considérons $u = (u_n = x_n + iy_n)$ une suite complexe et bornée. Alors, comme on l'a déjà vu, les deux suites (x_n) et (y_n) sont bornées. D'après le cas réel du théorème de Bolzano Weierstrass, on peut extraire une sous-suite convergente de (x_n) : il existe une fonction strictement croissante $\varphi : \mathbf{N} \to \mathbf{N}$ tel que $(x_{\varphi(n)})$ est convergente.

Considérons maintenant la suite $y' = (y'_n = y_{\varphi(n)})_{n\in\mathbf{N}}$. C'est une suite réelle bornée. Donc, à nouveau, on peut en extraire une sous-suite convergente. Notons la $(y'_{\psi(n)} = y_{\varphi\circ\psi(n)})_{n\in\mathbf{N}}$.

Mais la suite $(x_{\varphi\circ\psi(n)})_{n\in\mathbf{N}}$ est une suite extraite de la suite $(x_{\varphi(n)})_{n\in\mathbf{N}}$. D'après la proposition précédente, elle est donc convergente.

Ainsi, la suite $(u_{\varphi\circ\psi(n)})_{n\in\mathbf{N}}$ a une partie réelle et une partie imaginaire convergente ; on a vu que ça implique qu'elle est convergente. □

Remarque 2.4.5. — Comme application de ce théorème, on peut donner une preuve du fait qu'une fonction continue sur un intervalle $[a, b]$ compact est bornée et atteint ses bornes :

Soit $f : [a, b] \to \mathbf{R}$ une fonction continue. Notons $M = \sup\{f(x),\ x \in [a, b]\}$ (c'est un nombre réel ou $+\infty$). Par les propriétés du sup, il existe une suite (x_n) d'élément de $[a, b]$ telle que $f(x_n) \to M$. Or, la suite $(x_n)_{n\in\mathbf{N}}$ étant bornée, on peut en extraire une sous-suite convergente $x_{\varphi(n)} \to x \in [a, b]$. Par continuité de f, on a $f(x_{\varphi(n)}) \to f(x)$. Mais d'autre part, on a $f(x_{\varphi(n)}) \to M$. Donc $M = f(x)$: le sup est un nombre réel et est atteint. On raisonne de la même façon pour l'inf.

Le théorème de Bolzano-Weierstrass permet aussi de montrer que toute fonction continue sur un intervalle $[a, b]$ compact est uniformément continue.

CHAPITRE 3

SUITES NUMÉRIQUES RÉCURRENTES

Dans tout ce chapitre, on considérera une fonction réelle f, de domaine de définition D, et on étudiera les suites :

$$u_0 \in D \ ; \ u_{n+1} = f(u_n).$$

On se posera d'abord le problème de la définition de cette suite (peut-on vraiment définir tous les termes ?) puis de sa convergence. On donnera une application à la recherche de valeurs approchés de nombres réels, via l'algorithme de Newton.

3.1. Représentation graphique

On peut faire une représentation graphique intéressante de ce type de suites. Plaçons nous dans le plan euclidien avec un repère orthonormé, traçons les deux axes de coordonnées Ox, Oy, la droite $\mathcal{D}$ d'équation $y = x$ (dite *première bissectrice*) et le graphe $\mathcal{G}$ de la fonction f.

Considérons le point P_0 du graphe $\mathcal{G}$ d'abscisse u_0. Son ordonnée est alors $f(u_0) = u_1$. On cherche alors le point P_0' de $\mathcal{D}$ de même ordonnée que P_0, en traçant la droite horizontale qui passe par P_0 et en observant où elle intersecte $\mathcal{D}$. Par construction, les coordonnées de P_0' sont alors (u_1, u_1). Puis on cherche le point P_1 de $\mathcal{G}$ qui a la même abscisse que P_0', en traçant la droite verticale qui passe par P_0' et en observant où elle intersecte $\mathcal{G}$. L'abscisse de P_1 est u_1 ; donc son ordonnée est $f(u_1) = u_2$. On continue comme ça à construire les points P_i et P_i'. Les coordonnées de P_i sont (u_i, u_{i+1}). Notamment, en projetant sur l'axe des abscisses ou des ordonnées, on voit les termes consécutifs de la suite. (QC)

Exemple 3.1.1. — Prenez $f(x) = x^2$. Et appliquez la méthode précédente avec $u_0 = \frac{1}{2}$ ou $u_0 = 2$. Voir l'animation présentée en cours.

Prenez $f(x) = \frac{1}{x+1}$. Et appliquez la méthode précédente avec $u_0 = \frac{1}{2}$. Voir la deuxième animation présentée en cours.

3.2. Définition, monotonie

La première précaution à prendre est de s'assurer que la suite est bien définie pour tout n. Pour ça, on définit :

Définition 3.2.1. — Un intervalle I de $\mathbf{R}$ est dit *stable* par f si $I \subset D$ et $f(I) \subset I$.

On obtient alors par une récurrence immédiate :

Proposition 3.2.2. — *Si I est stable par f et si $u_0 \in I$, alors il existe une unique suite $(u_n)_{n\in\mathbf{N}}$ d'éléments de I telle que pour tout $n \geq 0$, $u_{n+1} = f(u_n)$.*

À partir de maintenant, on fixe une fonction réelle f et un intervalle I stable par f. Remarquons que si I est borné, la suite (u_n) est bornée.

On peut alors relier la monotonie de la suite (u_n) aux propriétés de f :

(QC) ***Proposition 3.2.3***. — *Si $f(x) \geq x$ pour tout $x \in I$, la suite (u_n) est croissante (strictement si l'inégalité est stricte).*
Si $f(x) \leq x$ pour tout $x \in I$, la suite (u_n) est décroissante (strictement si l'inégalité est stricte).

Démonstration. — Traitons le premier cas : pour tout n, $u_{n+1} = f(u_n) \geq u_n$. La suite (u_n) est donc bien croissante. L'autre cas est similaire. □

Plus intéressant :

(QC) ***Proposition 3.2.4***. —
1. *Si la fonction f est croissante, la suite (u_n) est monotone. Elle est croissante si $u_1 \geq u_0$, décroissante sinon.*
2. *Si la fonction f est décroissante, les deux suites extraites (u_{2n}) et (u_{2n+1}) sont monotones, de sens contraire.*

Attention à bien lire l'énoncé : le sens de variation de (u_n) n'est pas celui de f. Le premier exemple vu dans la section représentation graphique montre effectivement que le sens de variation dépend de u_0.

Démonstration. — Le premier point se montre par récurrence. Le sens de monotonie de la suite est déterminé par l'ordre entre u_0 et u_1. Supposons par exemple $u_1 \leq u_0$ et montrons par récurrence que pour tout $n \in \mathbf{N}$, $u_{n+1} \leq u_n$.

L'initialisation est faite. Si $u_{n+1} \leq u_n$, alors comme la fonction f est croissante, on a $f(u_{n+1}) \leq f(u_n)$, soit $u_{n+2} \leq u_{n+1}$. Ainsi la suite (u_n) est décroissante.

Pour le deuxième point, observons que la fonction $f \circ f$ est croissante [1]. Or les deux suites extraites (u_{2n}) et (u_{2n+1}) vérifient la récurrence $v_{n+1} = f \circ f(v_n)$. Donc elles sont toutes les deux monotones, et leur sens de variation dépend des deux premiers termes. Donc on peut écrire :
(u_{2n}) croissante $\Leftrightarrow u_2 \geq u_0 \Leftrightarrow$ (car f décroissante) $u_3 = f(u_2) \leq u_1 = f(u_0)$ $\Leftrightarrow (u_{2n+1})$ décroissante. □

3.3. Convergence et points fixes

On se pose maintenant le problème de la convergence de ces suites. On ajoute à partir de maintenant les deux hypothèses : f est continue sur I et I est un intervalle fermé [2]. Les limites possibles sont les *points fixes* de f, c'est-à-dire les points $x \in I$ tels que $f(x) = x$.

Proposition 3.3.1. — *Si f est continue sur I et $u_n \to l$, alors l appartient à I et l est point fixe de f : $f(l) = l$.* (QC)

Démonstration. — Premièrement, comme $u_n \to l$ et que I est fermé, on a $l \in I$.

Ensuite, considérons la suite $(f(u_n))_{n \in \mathbf{N}}$. Par continuité de f, elle tend vers $f(l)$. Mais par définition de (u_n), cette suite est la suite extraite (u_{n+1}). Elle tend donc vers l. Par unicité de la limite, on a $f(l) = l$. □

Donc pour étudier une éventuelle convergence de la suite (u_n), on commence par chercher des points fixes de f. Dans le cas où I est borné, on est assuré de leur existence, grâce au théorème des valeurs intermédiaires :

Proposition 3.3.2. — *Soit I un intervalle fermé borné stable par la fonction f continue sur I. Alors f a un point fixe dans I.*

Démonstration. — La fonction $x \mapsto f(x) - x$ est continue sur I, positive en la borne gauche de I et négative en la borne droite. Donc elle s'annule sur I ; c'est-à-dire que f a un point fixe. □

1. En effet, si $x \leq y$, on a par décroissance $f(x) \geq f(y)$ puis par décroissance encore $f \circ f(x) \leq f \circ f(y)$.

2. C'est-à-dire du type $[a, b]$, $]-\infty, b]$, $[a, +\infty[$ ou $[a, b]$.

3.4. Convergence et points fixes, Cas des fonctions dérivables

On suppose maintenant que f est continûment dérivable sur I. En faisant quelques dessins, il semble que la convergence de (u_n) vers un point fixe l n'est possible que si $|f'(l)| \leq 1$. Étudions ça plus précisément. D'abord le cas où $|f'(l)| > 1$:

(QC)

Proposition 3.4.1. — *Soit l un point fixe de f en lequel $|f'(l)| > 1$. Supposons que $u_n \to l$. Alors, il existe $N \in \mathbf{N}$ tel que pour $n \geq N$, $u_n = l$.*

Dans ce cas, on dit que la suite est *stationnaire*.

Démonstration. — On le montre par *contraposée* : montrons que si pour tout N il existe $n \geq N$ tel que $u_n \neq l$, alors u_n ne tend pas vers l. Cette dernière phrase veut dire

$$\exists \varepsilon > 0,\ \forall N \in \mathbf{N},\ \exists n \geq N,\ |u_n - l| > \varepsilon.$$

On cherche donc ce ε.

Soit $1 < K < |f'(l)|$. Par continuité de $|f'|$, il existe un intervalle J autour de l tel que pour tout $x \in J$, $|f'(x)| > K$. Soit maintenant $\varepsilon > 0$ tel que l'intervalle $[l - \varepsilon; l + \varepsilon]$ est inclus dans J. Montrons que cet ε convient. Pour ça, on fixe $N \in \mathbf{N}$ et on cherche un entier $n \geq N$ tel que $|u_n - l| > \varepsilon$.

Soit donc $p \geq N$ tel que $u_p \neq l$ (c'est possible par hypothèse). Si $|u_p - l| > \varepsilon$, alors on peut prendre $n = p$. Sinon, par définition de ε, on a $u_p \in J$. Par l'inégalité des accroissements finis, on a :

$$|u_{p+1} - l| = |f(u_p) - f(l)| > K|u_p - l|.$$

Et tant qu'on ne sort pas de l'intervalle J, on a $|u_{p+r} - l| > K^r|u_p - l|$. Comme $K > 1$ et $|u_p - l| > 0$, au bout d'un moment, on a $|u_{p+r} - l| > \varepsilon$. On peut alors prendre $n = p + r$. □

Exemple 3.4.2. — C'est le cas pour le point fixe 1 de la fonction f définie par $f(x) = x^2$

En revanche, si $|f'(l)| < 1$, alors pour x suffisamment proche de l, on a $|f(x) - l| < |x - l|$ donc la suite u_n semble devoir tendre vers l. Pour ça, on définit la notion de fonction contractante :

Définition 3.4.3. — Soit $0 < k < 1$. Une fonction $f : I \to I$ est dite *k-contractante* si pour tout $x, y \in I$, $|f(x) - f(y)| \leq k|x - y|$.

Une fonction f est dite *contractante* s'il existe $0 < k < 1$ tel que f est k-contractante.

Remarque 3.4.4. — Une fonction contractante est uniformément continue. En effet, soit f une fonction contractante sur I. Alors on a pour tous x et y dans I : si $|x-y| \leq \varepsilon$, alors $|f(x)-f(y)| \leq |x-y| \leq \varepsilon$.

Une bonne façon d'être contractante est d'avoir une dérivée strictement majorée par 1 sur un intervalle borné :

Proposition 3.4.5. — *On suppose f continue sur I et $\sup_{x\in I} |f'(x)| < 1$. Alors f est contractante.*

Démonstration. — Notons $k = \sup_{x\in I} |f'(x)|$. Par l'inégalité des accroissements finis sur I, on a pour tous x et y de I : $|f(y)-f(x)| \leq k|y-x|$. Comme $k<1$, on obtient bien que la fonction est k-contractante. □

Remarque 3.4.6. — Si f est continûment dérivable et que I est borné, on sait que $|f'|$ atteint ses bornes sur I. Donc il suffit d'avoir pour tout $x \in I$, $|f'(x)| < 1$ pour entrer dans le cadre de la proposition précédente.

Théorème 3.4.7. — *On suppose que I est fermé et f est continue et k-contractante sur I. Alors :* (QC)

1. *f admet un unique point fixe l dans I (théorème du point fixe de Banach).*
2. *Pour tout $u_0 \in I$, la suite définie par $u_{n+1} = f(u_n)$ converge vers l, et on a :*
$$|u_n - l| \leq k^n |u_0 - l|.$$

Démonstration. — On traite ici le cas de I fermé borné [3]. On sait alors déjà que f a un point fixe dans I. Si l et l' sont points fixes, on a $|l-l'| = |f(l)-f(l')| \leq k|l-l'|$. Comme $k<1$, ce n'est possible que si $l=l'$. Le point fixe est donc unique.

Par définition, on a $|u_{n+1}-l| = |f(u_n)-f(l)| < k|u_n-l|$. On montre ainsi par récurrence que $|u_n-l| \leq k^n|u_0-l|$. Ainsi, (u_n) tend vers l. □

Exemple 3.4.8. — Considérons la fonction $f(x) = \frac{1}{2}\left(x + \frac{3}{x}\right)$ et choisissons le premier terme $u_0 = 2$.

L'intervalle $I = [\sqrt{3}, +\infty[$ est stable. La dérivée y est comprise entre 0 et $\frac{1}{2}$. Donc la fonction est $\frac{1}{2}$-contractante. $\sqrt{3}$ est le seul point fixe sur I.

Donc la suite converge vers $\sqrt{3}$, et $|u_n - \sqrt{3}| \leq \left(\frac{1}{2}\right)^n |2-\sqrt{3}| \leq \left(\frac{1}{2}\right)^n$.

En pratique, en partant de $u_0 = 2$: on calcule $u_1 = 7/4 = 1,75$, $u_2 = 1,7321428...$, $u_3 = 1,73205081...$, $u_4 = 1,7320508075688...$ et $\sqrt{3} = 1,7320508075688...$

3. Le cas général général se traite avec une notion que nous n'avons pas vu cette année : celle de suite de Cauchy.

(u_4 coincide avec $\sqrt{3}$ à 17 chiffres après la virgule). On peut se reporter à animation

3.5. Méthode de Newton

Cette partie n'a été que brièvement présentée en amphi.

Le but est de généraliser l'exemple précédent pour trouver de bonnes approximations d'une solution de $F(x) = 0$, où F est une fonction C^2 définie sur I.

Commençons par décrire la méthode : on part d'un point de départ $u_0 \in I$. Si on connaît u_n, le terme u_{n+1} est l'abscisse de l'intersection de l'axe des abscisses et de la droite tangente au graphe de F en $(u_n, F(u_n))$. On peut se référer à cette animation.

En général, l'équation de la tangente au graphe en $(x_0, F(x_0))$ est

$$y = F(x_0) + F'(x_0)(x - x_0).$$

Son intersection avec l'axe des abscisses a lieu en $x = x_0 - \frac{F(x_0)}{F'(x_0)}$. Notons f la fonction $x \mapsto x - \frac{F(x)}{F'(x)}$, définie et dérivable sur $\{x \in I,\, F'(x) \neq 0\}$. La suite (u_n) est donc définie par $u_0 \in I$ et $u_{n+1} = f(u_n)$.

Proposition 3.5.1. — *Si la suite (u_n) est définie pour tout n et converge vers l, alors $F(l) = 0$.*

Démonstration. — Si la suite (u_n) converge vers l, alors $u_{n+1} = u_n - \frac{F(u_n)}{F'(u_n)}$ converge aussi vers l. Or u_n converge vers l, donc on doit avoir $\frac{F(u_n)}{F'(u_n)} \to 0$. Or $F(u_n)$ tend vers $F(l)$. La seule possibilité est donc que $F(l) = 0$ (quel que soit le comportement de $F'(u_n)$). □

Proposition 3.5.2. — *Soit l un zéro de F tel que $F'(l) \neq 0$. Alors, pour u_0 suffisamment proche de l, la suite converge vers l.*

Démonstration. — En effet, f est dérivable en 0 et

$$f'(l) = 1 - \frac{(F'(l))^2 - F(l)F''(l)}{(F'(l))^2} = 1 - 1 + 0 = 0.$$

Donc, par continuité de f', sur un certain intervalle J autour de l, $|f'(x)| < \frac{1}{2}$. Donc sur cet intervalle, f est $\frac{1}{2}$-contractante, le théorème précédent s'applique. Ainsi pour tout $u_0 \in J$, u_n converge vers l. □

L'exemple présenté à la fin de la section précédente s'obtient comme l'application de cette méthode à la fonction F définie par $F(x) = x^2 - 3$.

Un commentaire pour conclure : si on ne fait pas une étude préalable de la fonction, il est difficile de prévoir vers quoi va tendre la suite (u_n), et même si elle est bien définie. Si une étude de la fonction permet de déterminer un intervalle comme dans la dernière proposition, alors on peut sans problème prendre un u_0 dans J. Sinon, on peut *essayer* d'utiliser cette méthode en espérant converger, ce qui avec la première proposition assure qu'on converge vers un 0 de F.

PARTIE II

CALCUL MATRICIEL

CHAPITRE 4

MATRICES

On fixe dans ce chapitre le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$.

Un des objectifs de ce semestre est l'introduction à l'algèbre linéaire, avec la définition des notions d'espaces vectoriels, dimension, base, application linéaire... Ces notions seront définies en fin de semestre. Pour l'instant, nous nous concentrons sur les aspects concrets de l'algèbre linéaire, notamment le calcul matriciel.

L'utilisation des matrices et des opérations qu'on peut faire dessus est très répandue, notamment dans toutes les applications informatiques. Les raisons en sont très certainement multiples, mais on peut remarquer que ces objets allient une grande facilité et efficacité des calculs (on verra par exemple le pivot de Gauss) à une puissante théorie sous-jacente. Nous tenterons d'éclairer quelques applications à l'étude et l'exploitation du Web qui ont connu depuis une vingtaine d'années un grand développement.

4.1. Matrices

Définition 4.1.1. — L'ensemble, noté $\mathcal{M}_{p,q}(\mathbf{K})$, des *matrices à coefficients dans* $\mathbf{K}$ *à* p *lignes et* q *colonnes* est l'ensemble (QC)

$$\mathcal{M}_{p,q}(\mathbf{K}) = \left\{ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & \dots & \dots & a_{pq} \end{pmatrix} \text{ où les } a_{ij} \text{ sont dans } \mathbf{K} \right\}.$$

Une telle matrice est notée $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$. Les nombres a_{ij} s'appellent les coefficients de A. L'indice i désigne le numéro de sa ligne et l'indice j désigne celui de sa colonne.

Quand $p = 1$, on parle de vecteurs lignes. Inversement, quand $q = 1$, on parle de vecteurs colonnes.

Quand $p = q$, on parle de matrices carrées de taille p ; on note plus simplement $\mathcal{M}_p(\mathbf{K})$ l'ensemble des matrices carrées de taille p.

Les matrices nous seront très utiles comme exemples et outils pour l'algèbre linéaire. Cependant on peut noter que d'autres utilisations sont possibles :

Exemple 4.1.2. — Les matrices d'adjacences ou de transition d'un graphe :

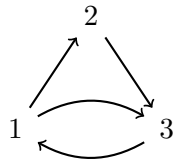

FIGURE 1. Notre graphe

Un graphe est un ensemble de flèches entre des points qu'on numérote de 1 à n (sur l'exemple, c'est de 1 à 3). On définit la matrice d'adjacence 1 comme la matrice carrée de $\mathcal{M}_n(\mathbf{K})$ dont le coefficient i, j vaut le nombre de flèches de i vers j. Sur notre exemple, la matrice est

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

La matrice de transition T est une variante de la matrice d'adjacence : c'est la matrice carrée de $\mathcal{M}_n(\mathbf{K})$ dont le coefficient i, j vaut le nombre de flèches de i vers j divisé par le nombre total de flèches partant de i. Sur notre exemple :

$$T = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

On peut ainsi imaginer les "matrices du Web" en imaginant le Web comme un gigantesque graphe (une page est un sommet, un lien crée une flèche entre deux sommets) et en considérant les deux matrices ci-dessous. C'est

par exemple le point de départ des algorithmes utilisés par les moteurs de recherche.

Exemple 4.1.3. — La « matrice d'Amazon » est la matrice qui comporte autant de lignes que d'utilisateurs d'Amazon et autant de colonnes que d'objet vendus et dont le coefficient i, j est l'appréciation (un entier entre 0 et 5) de l'utilisateur i sur l'objet j.

4.2. Somme de matrices et produit par un scalaire

Définition 4.2.1. — On définit la somme de deux matrices de $\mathcal{M}_{p,q}(\mathbf{K})$ et le produit d'une matrice par un scalaire $t \in \mathbf{K}$ par : (QC)

- $(a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}} + (b_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}} = (a_{ij} + b_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$
- $t \cdot (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}} = (ta_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$

Ces opérations se comportent de manière sympathique ; par exemple, pour tous scalaires s et t et toutes matrices A et B, on vérifie directement avec la formule : (QC)

$$(s+t)(A+B) = sA + tA + sB + tB.$$

On verra en fin de semestre qu'avec ces opérations, $\mathcal{M}_{p,q}(\mathbf{K})$ est un $\mathbf{K}$-espace vectoriel.

La matrice dont tous les coefficients sont nuls est appelée la *matrice nulle*.

On appelle *matrices élémentaires* de $\mathcal{M}_{p,q}(\mathbf{K})$ les matrices dont un seul coefficients vaut 1 et tous les autres sont nuls. Par exemple, quand $p = q = 2$, les matrices élémentaires sont les quatre matrices :
$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ et $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.
On vérifie alors :

$$a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Autrement dit, toute matrice est *combinaison linéaire* (c'est-à-dire somme avec des coefficients) des matrices élémentaires, et cette écriture est unique. On verra plus tard que (la taille étant fixée) la famille des matrices élémentaires forme une base de l'espace des matrices.

4.3. Produits de matrices

On va définir le produit AB d'une matrice $A \in \mathcal{M}_{p,r}(\mathbf{K})$ et d'une matrice $B \in \mathcal{M}_{n,q}(\mathbf{K})$ quand $r = n$, c'est-à-dire :
On peut multiplier deux matrices quand le nombre de colonnes de celle de gauche égale le nombre de ligne de celle de droite.
Le résultat est alors une matrice de $\mathcal{M}_{p,q}(\mathbf{K})$, c'est-à-dire :
Le produit de deux matrices a autant de lignes que celle de gauche et de colonnes que celle de droite.

4.3.1. Cas des vecteurs lignes et colonnes. — Pour commencer, on regarde le cas $p = 1$, $q = 1$ et $r = n$. Alors la matrice A est un vecteur ligne à n coefficients et la matrice B un vecteur colonne à n coefficients. On les note

$$A = (a_{\star 1}, \ldots, a_{\star n}) \quad \text{et} \quad B = \begin{pmatrix} b_{1\bullet} \\ \vdots \\ b_{n\bullet} \end{pmatrix}.$$

On pose alors

$$AB = (a_{\star 1}, \ldots, a_{\star n}) \begin{pmatrix} b_{1\bullet} \\ \vdots \\ b_{n\bullet} \end{pmatrix} = (a_{\star 1} b_{1\bullet} + a_{\star 2} b_{2\bullet} + \ldots + a_{\star n} b_{n\bullet}) = \sum_{k=1}^{n} a_{\star k} b_{k\bullet}.$$

On remarque alors une double propriété de linéarité :

Remarque 4.3.1. — Pour tous scalaire s et $t \in \mathbf{K}$, vecteurs lignes A et A' dans $\mathcal{M}_{1,n}(\mathbf{K})$ et vecteurs colonnes B et B' dans $\mathcal{M}_{n,1}(\mathbf{K})$, on a

- (linéarité par rapport à la variable de gauche) $(sA+tA')B = sAB+tA'B$
- (linéarité par rapport à la variable de droite) $A(sB+tB') = sAB+tAB'$.

On dit que ce produit est *bilinéaire*. La preuve de ces égalités est un calcul à partir de la définition. Faisons-le d'abord dans le premier cas :

$$\begin{aligned}(sA+tA')B &= \sum_{k=1}^{n}(sA+tA')_{\star k}b_{k\bullet}\\ &= \sum_{k=1}^{n}(sa_{\star k}+ta'_{\star k})b_{k\bullet}\\ &= \sum_{k=1}^{n}sa_{\star k}b_{k\bullet}+ta'_{\star k}b_{k\bullet}\\ &= s\sum_{k=1}^{n}a_{\star k}b_{k\bullet}+t\sum_{k=1}^{n}a'_{\star k}b_{k\bullet}\\ &= sAB+tA'B\end{aligned}$$

Dans l'autre cas, le calcul est similaire :

$$\begin{aligned}A(sB+tB') &= \sum_{k=1}^{n}a_{\star k}(sB+tB')_{k\bullet}\\ &= \sum_{k=1}^{n}a_{\star k}(sb_{k\bullet}+tb'_{k\bullet})\\ &= s\sum_{k=1}^{n}a_{\star k}b_{k\bullet}+t\sum_{k=1}^{n}a_{\star k}b'_{k\bullet}\\ &= sAB+tAB'\end{aligned}$$

4.3.2. Cas général. — Nous pouvons maintenant définir le produit de matrices dans le cas général :

Définition 4.3.2. — Le produit de la matrice $A=(a_{ij})_{\substack{1\leq i\leq p\\1\leq j\leq n}}\in\mathcal{M}_{p,n}(\mathbf{K})$ et de la matrice $B=(b_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq q}}\in\mathcal{M}_{n,q}(\mathbf{K})$ est la matrice $P=(p_{ij})_{\substack{1\leq i\leq p\\1\leq j\leq q}}=AB\in\mathcal{M}_{p,q}(\mathbf{K})$ définie par : (QC)
p_{ij} est le produit de la i-ième ligne de A par la j-ième colonne de B, ou encore

$$p_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}.$$

Notamment, le produit AC d'une matrice $A\in\mathcal{M}_{p,n}(\mathbf{K})$ et d'une matrice colonne $C\in\mathcal{M}_{n,1}(\mathbf{K})$ est encore une matrice colonne. De plus, si on note C_j la j-ième colonne de B, alors le produit AB est la matrice $(AC_1,AC_2,\dots AC_q)$ (la juxtaposition de ces q colonnes). De même le produit LB d'une matrice

ligne $L \in \mathcal{M}_{1,n}(\mathbf{K})$ et d'une matrice $B \in \mathcal{M}_{n,q}(\mathbf{K})$ est encore une matrice ligne. De plus, si on note L_i la i-ième ligne de A, alors le produit AB est la matrice $\begin{pmatrix} L_1B \\ L_2B \\ \vdots \\ L_pB \end{pmatrix}$ (la juxtaposition de ces p lignes). Autrement dit, on peut calculer le produit matriciel ligne par ligne, ou bien colonne par colonne.

Proposition 4.3.3. — *On fixe $n, p, q, r \in \mathbf{N}^*$.*

(QC)

1. *Le produit matriciel $\mathcal{M}_{p,n}(\mathbf{K}) \times \mathcal{M}_{n,q}(\mathbf{K}) \to \mathcal{M}_{p,q}(\mathbf{K})$ est bilinéaire : pour tous s et $t \in \mathbf{K}$, A et $A' \in \mathcal{M}_{p,n}(\mathbf{K})$ et B et $B' \in \mathcal{M}_{n,q}(\mathbf{K})$, on a*

$$(sA + tA')B = sAB + tA'B \qquad \text{et} \qquad A(sB + tB') = sAB + tAB'.$$

2. *Le produit matriciel est associatif : pour $A \in \mathcal{M}_{p,n}(\mathbf{K})$, $B \in \mathcal{M}_{n,r}(\mathbf{K})$ et $C \in \mathcal{M}_{r,q}(\mathbf{K})$, on a $(AB)C = A(BC)$.*

Démonstration. — Le premier point découle de la remarque 4.3.1 : le coefficient d'indice ij de $(sA + tA')B$ est le produit de la i-ième ligne de $sA + tA'$ et de la j-ième colonne de B. En notant L_i et L'_i les i-ièmes lignes de A et A' et C_j la j-ième colonne de B, on a donc :

$$((sA + tA')B)_{ij} = (sL_i + tL'_i)C_j = sL_iC_j + tL'_iC_j = s(AB)_{ij} + t(A'B)_{ij}.$$

Donc le coefficient d'indice i, j de $(sA + tA')B$ vaut $s(AB)_{ij} + t(A'B)_{ij}$. On en déduit $(sA + tA')B = sAB + tA'B$. L'autre relation est similaire.

Le deuxième point se prouve par un calcul. Remarquons d'abord que nous avons bien le droit de former ces deux produits et qu'ils donnent tous les deux une matrice à p lignes et q colonnes. Maintenant, pour $1 \leq i \leq j$, on écrit :

$$((AB)C)_{ij} = \sum_{k=1}^{r} (AB)_{ik}C_{kj} = \sum_{k=1}^{r} (\sum_{l=1}^{n} A_{il}B_{lk})C_{kj} = \sum_{\substack{1 \leq k \leq r \\ 1 \leq l \leq n}} A_{il}B_{lk}C_{kj}$$

et

$$(A(BC))_{ij} = \sum_{l=1}^{n} A_{il}(BC)_{lj} = \sum_{l=1}^{n} A_{il}(\sum_{k=1}^{r} B_{lk}C_{kj}) = \sum_{\substack{1 \leq k \leq r \\ 1 \leq l \leq n}} A_{il}B_{lk}C_{kj}$$

Les coefficients de ces deux matrices sont égaux, c'est donc qu'elles sont égales. □

Notamment, grâce au deuxième point, on écrira des produits de plusieurs matrices sans parenthèses.

Le produit de matrices n'est pas commutatif ; tout d'abord parce que on ne peut pas forcément former le produit BA si on peut former AB. Mais même quand les deux existent, ils peuvent être différents (voir TD).

4.3.3. Transposée de matrice. —

Définition 4.3.4. — Soit $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$ une matrice de $\mathcal{M}_{p,q}(\mathbf{K})$. On définit sa *transposée*, notée ${}^tA$, dans $\mathcal{M}_{q,p}(\mathbf{K})$ par : (QC)

$$({}^tA)_{ij} = a_{ji}.$$

Les lignes de A deviennent les colonnes de ${}^tA$. Par exemple, si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, sa transposée est ${}^tA = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

4.4. Cas des matrices carrées

4.4.1. Puissance. — On remarque que si A et B sont deux matrices carrées dans $\mathcal{M}_p(\mathbf{K})$ on peut toujours former le produit AB et que c'est encore une matrice carré de taille p. Notamment, on peut définir la puissance d'une matrice carrée :

Définition 4.4.1. — Soit $A \in \mathcal{M}_p(\mathbf{K})$. On définit A^n comme le produit $AA \cdots A$, où A apparaît n fois. (QC)

Exemple 4.4.2. — Revenons à notre matrice d'adjacence de graphe. Montrons par récurrence que le coefficient d'indices i,j de A^n, qu'on notera $a_{i,j}(n)$, est le nombre de chemins dans le graphe de n étapes qui vont de i à j.

On procède par récurrence. Pour $n = 1$, c'est la définition de A. Supposons que la propriété soit vraie pour un entier n. On calcule les coefficients de A^{n+1} en utilisant $A^{n+1} = A^n A$. En utilisant la formule du produit, on a :

$$a_{i,j}(n+1) = \sum_{k=1}^{3} a_{i,k}(n) a_{k,j}.$$

Regardons le produit $a_{i,k}(n)a_{k,j}$. Par hypothèse de récurrence, le premier terme est le nombre de chemins en n étapes de i à k. Le deuxième est le nombre de chemins en une étape de k à j. Donc ce produit est le nombre de chemins qui vont de i à j en $n+1$ étapes en passant par k à l'étape n. Quand on fait la somme sur tous les k, c'est à dire toutes les possibilités de dernière étape, on

obtient bien que $a_{i,j}(n+1)$ est le nombre de chemins possibles entre i et j en $n+1$ étapes.

4.4.2. Matrice identité et matrices diagonales. — Dans le cas des matrices carrées, on a une matrice remarquable pour le produit : la *matrice identité*, notée I_n, dont les tous les coefficients diagonaux valent 1 et les autres 0 : le coefficient vaut 1 si le numéro de la ligne est le même que celui de la colonne, 0 sinon. On la représente de cette façon :

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Cette matrice est remarquable, car c'est un élément neutre pour le produit :

Proposition 4.4.3. — *Soit A une matrice de $\mathcal{M}_n(\mathbf{K})$. Alors $AI_n = I_nA = A$.*

Nous verrons la preuve la semaine prochaine. Exercice : vérifiez le dans le cas $n = 2$.

Démonstration. — Prouvons par exemple $I_nA = A$: le coefficient i, j de I_nA est $\sum_{k=1}^{n}(I_n)_{ik}A_{kj}$. Tous les termes de cette somme sont nuls sauf pour $k = i$, auquel cas $(I_n)_{ik} = 1$. Donc ce coefficient vaut A_{ij}. Ça prouve l'égalité voulue. □

(QC)

Définition 4.4.4. — Plus généralement, on appelle *matrice diagonale* une matrice $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq p}}$ dont seuls coefficients non nuls sont les a_{ii}.Une telle matrice se représente sous la forme $A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & a_{pp} \end{pmatrix}$.

Proposition 4.4.5. — *Si D est diagonale, de coefficients diagonaux $(d_1, \dots d_p)$, alors D^n est encore diagonale, de coefficients diagonaux $(d_1^n, \dots, d_p^n)$.*

La preuve est laissée en exercice au lecteur.

On appelle matrice triangulaire supérieure une matrice dont tous les coefficients sous la diagonale sont nuls (le coefficient i, j vaut 0 si $i > j$). Inversement, on appelle matrice triangulaire inférieure une matrice dont tous les coefficients au dessus de la diagonale sont nuls (le coefficient i, j vaut 0 si $i < j$).

4.4.3. Matrices inversibles. — Un rôle crucial sera joué par la notion de *matrice inversible*.

Définition 4.4.6. — Une matrice A de $\mathcal{M}_p(\mathbf{K})$ est dite inversible si il existe une matrice B de $\mathcal{M}_p(\mathbf{K})$ telle que $AB = BA = I_p$. (QC)

Cette matrice B est alors unique, s'appelle l'inverse de A et est notée A^{-1}.

La preuve de l'unicité est une simple manipulation : si B et B' sont deux candidates, on peut écrire (en utilisant $AB = I_n$ et $B'A = I_p$) :

$$B' = B'I_p = B'(AB) = (B'A)B = I_pB = B.$$

Exemple 4.4.7. — On laisse le lecteur vérifier les exemples suivants :

- La matrice $\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$ est inversible d'inverse $\begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$.
- La matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est inversible d'inverse $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$
- La matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ n'est pas inversible : tout produit de la forme BA a sa première colonne nulle, donc ne peut pas valoir I_2.
- La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ n'est pas inversible : tout produit de la forme BA a ses deux colonnes égales, donc ne peut pas valoir I_2.

On verra plus tard, dans deux chapitres, comment déterminer si une matrice est ou non inversible, et si oui comment l'inverser, grâce au déterminant. Pour une matrice inversible A, la puissance A^n est définie pour tout $n \in \mathbf{Z}$: en effet, pour $n > 0$, ça a déjà été fait. Pour $n = 0$, on pose par convention $A^n = I_p$; pour $n < 0$, $A^n = (A^{-1})^n$.

4.4.4. Trace. — On définit encore la *trace* d'une matrice carrée $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq p}}$, notée $\mathrm{Tr}(A)$ comme la somme des coefficients diagonaux : (QC)

$$\mathrm{Tr}(A) = \sum_{k=1}^{p} a_{kk}.$$

Proposition 4.4.8. — *La trace d'une matrice carrée et de sa transposée sont égales*

Démonstration. — En transposant une matrice carrée, on ne change pas sa diagonale, donc pas sa trace. □

CHAPITRE 5

RÉSOLUTION DES SYSTÈMES LINÉAIRES

5.1. Systèmes linéaires, aspects théoriques

Fixons p et n deux entiers non nuls. Un *système linéaire à p équations et n inconnues* est un système $(\mathcal{S})$ de la forme :

$$(\mathcal{S}) : \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & & & \vdots & = & \vdots \\ a_{p1}x_1 & + & a_{p2}x_2 & + & \dots & + & a_{pn}x_n & = & b_p \end{cases}$$

où les a_{ij} et les b_i sont des éléments de $\mathbf{K}$, et les x_j sont des inconnues.

L'ensemble des solutions d'un tel système est l'ensemble des $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ qui vérifient toutes les équations.

Interprétation matricielle : Notons $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} \in \mathcal{M}_{pn}(\mathbf{K})$, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$ et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Alors le système $(\mathcal{S})$ se réécrit comme l'équation matricielle $AX = B$. En effet, le produit AX est un vecteur colonne à p lignes et on peut la calculer :

$$AX = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{p1}x_1 + \ldots + a_{pn}x_n \end{pmatrix}.$$

On appelle A la *matrice du système* $(\mathcal{S})$.

On appelle *matrice augmentée* de $(\mathcal{S})$ la matrice, notée $(A|b)$ formée de A et de b en dernière colonne. La donnée de la matrice augmentée est équivalente à celle du système, chaque ligne codant une équation du système. On parlera donc indifféremment du système ou de la matrice augmentée.

Si $b = 0$, on dit que le système est homogène. On note $(\mathcal{S}_0)$ le système homogène associé à $(\mathcal{S})$, c'est-à-dire le système $AX = 0$. Notons $L_1, \ldots, L_p$ les lignes du système homogène $(\mathcal{S}_0)$. On dit qu'un certains nombres de ces lignes sont liées si il existe une combinaison linéaire (dont tous les coefficients ne sont pas nuls) – c'est-à-dire une somme avec coefficients – de ces lignes qui mène à l'équation nulle. Si un ensemble de lignes n'est pas lié, on dit que les équations sont libres, ou indépendantes.

On appelle rang de $(\mathcal{S})$, noté $\mathrm{rg}(\mathcal{S})$ le nombre maximal de lignes indépendantes.

Un premier théorème important est le suivant :

(QC)

Théorème 5.1.1. — *Soit $(\mathcal{S})$ un système $AX = b$ et on suppose A inversible. Alors pour tout b, le système a une unique solution, donnée par $X = A^{-1}b$.*

Démonstration. — On voit ici l'efficacité de la notation matricielle :

Montrons d'abord que $A^{-1}b$ est solution : on doit montrer que $A(A^{-1}b) = b$. Mais on a vu que $A(A^{-1}b) = (AA^{-1})b = I_p b = b$. Ce vecteur est donc bien une solution.

Réciproquement, si X est solution, alors $AX = b$. En multipliant les deux termes par A^{-1}, on obtient $A^{-1}(AX) = A^{-1}b$. Or, comme ci-dessus, on calcule $A^{-1}(AX) = (A^{-1}A)X = I_p X = X$. On a donc bien $X = A^{-1}b$. □

Le théorème précédent est important théoriquement. Cependant nous devons mettre en garde le lecteur : pour résoudre un système $AX = b$, il est très coûteux d'inverser la matrice A. En effet, si A est une matrice carrée de taille p, il faut environ p^3 opérations pour l'inverser (on en reparlera au chapitre suivant). En revanche, la méthode que nous allons présenter ne nécessite "que" p^2 opérations. Donc le théorème suivant ne s'utilise que quand on connaît déjà l'inverse de A, ou bien quand on doit résoudre le système pour un grand nombre de b différents.

Ce résultat théorique nous guidera pour la résolution d'un tel système :

Proposition 5.1.2. — *Supposons qu'on ait une solution X_0 du système $(\mathcal{S})$. Alors l'ensemble des solutions est :*

$$\{X_0 + X \text{ pour } X \text{ solution du système homogène } (\mathcal{S}_0)\}.$$

Démonstration. — On a la suite d'équivalence :

$$\begin{aligned} Y \text{ solution de } (\mathcal{S}) &\Leftrightarrow AY = B \\ &\Leftrightarrow AY = AX_0 \\ &\Leftrightarrow A(Y - X_0) = 0 \\ &\Leftrightarrow X = Y - X_0 \text{ est solution de } (\mathcal{S}_0). \end{aligned}$$

□

Il reste à pouvoir utiliser en pratique ce théorème : comment fait on pour déterminer si il existe une solution ? et comment décrire l'ensemble des solutions du système homogène ? Pour ça, on utilise le pivot de Gauss.

5.2. Systèmes linéaires : le pivot de Gauss

5.2.1. Introduction. — On cherche à décrire un *algorithme*, c'est-à-dire une méthode automatique qui mène au résultat voulu. Notamment, un algorithme est une méthode qu'on peut « enseigner » à un ordinateur, pour lui faire résoudre le problème.

Notre problème ici est de résoudre les systèmes linéaires. Pour les petits systèmes, on sait qu'on peut y arriver par des manipulations sur les équations : on les ajoute entre elles, on les multiplie par des scalaires... Le but cette section est de formaliser ces opérations. L'objectif est double : d'une part, comme évoqué, de pouvoir les faire faire à un ordinateur pour résoudre de gros systèmes linéaires (quelques centaines, voire milliers, voire plus, d'inconnues et d'équations, ça peut arriver vraiment dans les applications). D'autre part, s'assurer que les opérations qu'on fait ne changent pas la nature du système (notamment l'ensemble des solutions).

Exemple 5.2.1. — Pour se convaincre qu'il faut faire attention, considérons le système

$$(\mathcal{S}) : \begin{cases} 2x_1 + 2x_2 + x_3 = 0 \\ x_1 + 2x_2 + 2x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}$$

On remplace la ligne L_1 par $L_2 - L_1$ pour éliminer x_2, L_2 par $L_2 - L_3$ pour éliminer x_1 et L_3 par $L_3 - L_1$ pour éliminer x_3 ; on obtient le système

$$(\mathcal{S}') : \begin{cases} x_1 - x_3 = 0 \\ x_2 + x_3 = 0 \\ x_1 + x_2 = 0 \end{cases}$$

Autrement dit, les solutions de $(\mathcal{S}')$ sont tous les vecteurs vérifiant $x_1 = x_3 = -x_2$.

Or on vérifie que $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ n'est pas solution de $(\mathcal{S})$, alors qu'il l'est de $(\mathcal{S}')$. Nos opérations n'étaient pas autorisées.

(Exercice : déterminer l'ensemble des solutions de $(\mathcal{S})$, y compris en utilisant l'algorithme qu'on va décrire dans la suite de ce cours.)

5.2.2. Opérations élémentaires sur les lignes. — Soit donc $(\mathcal{S})$ un système $AX = b$, avec $A \in \mathcal{M}_{p,n}(\mathbf{K})$ et $b \in \mathbf{K}^p$.

On définit trois opérations élémentaires sur les lignes de la matrice augmentée (donc sur les équations de $(\mathcal{S})$) :

Dilatation : Remplacer la i-ème ligne L_i par $L_i' = \alpha L_i$, pour un scalaire α *non nul*. On le note $L_i \leftarrow \alpha L_i$. Ça revient à multiplier la matrice augmentée à gauche par la matrice $D_i(\alpha)$ diagonale dont tous les coefficients diagonaux valent 1, sauf le i-ème qui vaut α.

Transvection : Remplacer la i-ème ligne L_i par $L_i' = L_i + \lambda L_j$, pour un $j \neq i$ et λ un scalaire. On le note $L_i \leftarrow L_i + \lambda L_j$. Ca revient à multiplier la matrice augmentée à gauche par la matrice $T_{ij}(\lambda) = I_p + \lambda E_{ij}$.

Permutation : Échanger les lignes i et j : $L_i' = L_j$ et $L_j' = L_i$. On le note $L_i \leftrightarrow L_j$. Ça revient à multiplier la matrice augmentée à gauche par la matrice $P_{ij} = I_p - E_{ii} - E_{jj} + E_{ij} + E_{ji}$.

Si X est une solution du système avant l'opération, il est encore solution du système modifié. De plus, on voiten outre que ces opérations sont réversibles :

Proposition 5.2.2. — *Si un système $(\mathcal{S}')$ est obtenu à partir de $(\mathcal{S})$ par une de ces trois opérations, alors le contraire est aussi vrai : $(\mathcal{S})$ est obtenu à partir de $(\mathcal{S}')$ par une de ces trois opérations.*

Notamment les matrices $D_i(\alpha)$ (pour $\alpha \neq 0$), $T_{ij}(\lambda)$ et P_{ij} sont inversibles.

Démonstration. — En effet, si on a effectué $L_i \leftarrow \alpha L_i$, on revient au système initial en effectuant $L_i \leftarrow \frac{1}{\alpha} L_i$; si on a effectué $L_i \leftarrow L_i + \alpha L_j$, on revient au système initial en effectuant $L_i \leftarrow L_i - \alpha L_j$; enfin, si on a effectué $L_i \leftrightarrow L_j$ on revient au système initial en effectuant la même opération.

L'inverse de $D_i(\alpha)$ est $D_i(\frac{1}{\alpha})$, celui de $T_{ij}(\lambda)$ est $T_{ij}(-\lambda)$ et P_{ij} est son propre inverse. □

Finalement on obtient la remarque cruciale suivante :

Proposition 5.2.3. — *Si un système $(\mathcal{S}')$ est obtenu à partir de $(\mathcal{S})$ par une des trois opérations élémentaires, alors ils ont les mêmes solutions. On dit qu'ils sont* équivalents.

De plus, ils ont même rang.

Démonstration. — On a déjà remarqué que toute solution de $(\mathcal{S})$ est solution de $(\mathcal{S}')$. Or $(\mathcal{S})$ est aussi obtenu à partir de $(\mathcal{S}')$ par une de ces trois opérations. Donc toute solution de $(\mathcal{S}')$ est aussi solution de $(\mathcal{S})$.

Pour le rang, le rang de $(\mathcal{S})$ est la dimension de l'espace vectoriel engendré par les lignes de A. Cette dimension ne change pas en effectuant une des trois opérations.

Pour la dernière remarque sur le rang, nous y reviendrons quand nous aurons un peu plus étudié les problèmes de dimension. □

5.2.3. L'algorithme. — Décrivons maintenant l'algorithme. Soit $(A|b)$ la matrice augmentée d'un système. On note a_{ij} les coefficients de A. Le but est à chaque étape de faire disparaître une inconnue de toutes les équations sauf une puis de recommencer avec une matrice plus petite. Voilà la procédure :

On regarde la première colonne de la matrice $(A|b)$:

1. Si la première colonne de A est nulle, on appelle A' la matrice déduite de A en enlevant la première colonne. Si A' est vide, on s'arrête. Sinon, on recommence avec la matrice augmentée $(A'|b)$ sans toucher au reste de la matrice $(A|b)$.
2. Sinon, soit i le plus petit indice tel que $a_{i1} \neq 0$.
   a) Si $i \neq 1$, on fait $L_i \leftrightarrow L_1$. Maintenant la matrice augmentée, qu'on note toujours $(A|b)$, a un coefficient a_{11} non nul.
   b) Si $a_{11} \neq 1$, on fait $L_1 \leftarrow \frac{1}{a_{11}} L_1$. Maintenant la matrice augmentée, qu'on note toujours $(A|b)$, a un coefficient $a_{11} = 1$. On appelle ce coefficient « le pivot » ; pour plus de clarté, on l'encadre dans la matrice augmentée.
   c) Pour $i = 2, 3, \ldots, p$, on fait $L_i \leftarrow L_i - a_{i1} L_1$. Maintenant la matrice augmentée, qu'on note toujours $(A|b)$, a sa première colonne égale à $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

On note alors $(A|b) = \left(\begin{array}{c|c} 1 \quad * \ldots * & b_1 \\ 0 & \\ \vdots \quad A' & b' \\ 0 & \end{array}\right)$. Si A' est pas vide, on s'arrête ; sinon on recommence avec la matrice augmentée $(A'|b')$ sans toucher au reste de la matrice $(A|b)$.

On remarque qu'à chaque fois qu'on recommence, la matrice A' est strictement plus petite, donc cette algorithme termine : au bout d'un nombre fini d'étapes, la matrice A' est vide et on s'arrête. Comme à chaque étape, on fait une des trois opérations élémentaires, les solutions du système obtenu à la fin de cet algorithme sont les mêmes que celle du système de départ. Comme question de cours, on peut vous demander d'appliquer cet algorithme sur un exemple.

(QC)

Dernière étape optionnelle : en appliquant des transvections, on annule les coefficients au dessus des pivots : dans chaque colonne k contenant un pivot (en commençant par la gauche), on effectue pour tout $i < k$ l'opération $L_i \leftarrow L_i - a_{ik}L_k$.

Exemple 5.2.4. — On part de la matrice

$$\left(\begin{array}{cccccc|c} 1 & 1 & 0 & 0 & 1 & 1 & b_1 \\ 2 & 2 & 1 & 1 & 3 & 4 & b_2 \\ 5 & 5 & 2 & 2 & 8 & 8 & b_3 \end{array}\right).$$

On applique l'algorithme : la première colonne n'est pas nulle, donc on est dans le deuxième cas. Le coefficient en haut à gauche est déjà un 1, c'est notre pivot ; on l'entoure :

$$\left(\begin{array}{cccccc|c} \boxed{1} & 1 & 0 & 0 & 1 & 1 & b_1 \\ 2 & 2 & 1 & 1 & 3 & 4 & b_2 \\ 5 & 5 & 2 & 2 & 8 & 8 & b_3 \end{array}\right).$$

Ensuite, on fait les deux opérations $L_2 \leftarrow L_2 - 2L_1$ et $L_3 \leftarrow L_3 - 5L_1$, pour obtenir

$$\left(\begin{array}{cccccc|c} \boxed{1} & 1 & 0 & 0 & 1 & 1 & b_1 \\ 0 & 0 & 1 & 1 & 1 & 2 & b_2 - 2b_1 \\ 0 & 0 & 2 & 2 & 3 & 3 & b_3 - 5b_1 \end{array}\right).$$

La matrice $(A'|b')$ est alors composée des deux dernières lignes moins la première colonne. Elle commence par une colonne de 0 ; on enlève cette colonne,

et on obtient un nouveau pivot :

$$\left(\begin{array}{cccccc|c} \boxed{1} & 1 & 0 & 0 & 1 & 1 & b_1 \\ 0 & 0 & \boxed{1} & 1 & 1 & 2 & b_2 - 2b_1 \\ 0 & 0 & 2 & 2 & 3 & 3 & b_3 - 5b_1 \end{array}\right).$$

On applique $L_3 \leftarrow L_3 - 2L_2$, et on obtient :

$$\left(\begin{array}{cccccc|c} \boxed{1} & 1 & 0 & 0 & 1 & 1 & b_1 \\ 0 & 0 & \boxed{1} & 1 & 1 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & 1 & -1 & b_3 - 2b_2 - b_1 \end{array}\right).$$

La matrice $(A'|b')$ est alors composée des 4 derniers éléments de la dernière ligne ; elle commence par un zéro qu'on enlève, puis on obtient le dernier pivot :

$$\left(\begin{array}{cccccc|c} \boxed{1} & 1 & 0 & 0 & 1 & 1 & b_1 \\ 0 & 0 & \boxed{1} & 1 & 1 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & \boxed{1} & -1 & b_3 - 2b_2 - b_1 \end{array}\right).$$

L'algorithme est terminé : toutes les lignes commencent par un 1. On peut appliquer l'étape optionnelle : on effectue les opérations $L_1 \leftarrow L_1 - L_3$ et $L_2 \leftarrow L_2 - L_3$. On obtient :

$$\left(\begin{array}{cccccc|c} \boxed{1} & 1 & 0 & 0 & 0 & 2 & 2b_1 + 2b_2 - b_3 \\ 0 & 0 & \boxed{1} & 1 & 0 & 3 & -b_1 + 3b_2 - b_3 \\ 0 & 0 & 0 & 0 & \boxed{1} & -1 & b_3 - 2b_2 - b_1 \end{array}\right).$$

En partant d'une matrice $(A|b)$ quelconque, on arrive sur une matrice $(A'|b')$ où A' est échelonnée, et même échelonnée réduite si on applique l'étape optionnelle :

Définition 5.2.5. — Une matrice $A \in \mathcal{M}_{p,n}(\mathbf{K})$ est dite *échelonnée à r lignes* si ses r lignes non nulles sont les r premières et si, pour $1 \leq i \leq r$, le premier coefficient non nul de L_i vaut 1 et est en position k_i, avec $1 \leq k_1 < k_2 < \ldots < k_r \leq n$.

Elle est échelonnée *réduite* si en plus dans les colonnes des pivots tous les coefficients sauf le pivot sont nuls.

Il vaut mieux se représenter cette notion ; une matrice échelonnée ressemble à :

$$\begin{pmatrix} \boxed{1} & * & * & \dots & * & * & \dots \\ 0 & 0 & \boxed{1} & \dots & * & * & \dots \\ \vdots & \vdots & & & & & \\ 0 & 0 & 0 & \dots & 0 & \boxed{1} & \dots \\ 0 & & & & \dots & 0 & \dots \\ \vdots & & & & \dots & & \vdots \end{pmatrix}.$$

Une matrice échelonnée réduite à :

$$\begin{pmatrix} \boxed{1} & * & 0 & \dots & * & 0 & \dots \\ 0 & 0 & \boxed{1} & \dots & * & 0 & \dots \\ \vdots & \vdots & & & & & \\ 0 & 0 & 0 & \dots & 0 & \boxed{1} & \dots \\ 0 & & & & \dots & 0 & \dots \\ \vdots & & & & \dots & & \vdots \end{pmatrix}.$$

On note toujours k_1, …, k_r les numéros de colonnes où apparaissent les pivots. Notons j_1, …, j_{n-r} les autres.

Remarque 5.2.6. — Un système $(\mathcal{S})$ dont la matrice est échelonnée à r ligne est de rang r. En effet, les r premières lignes sont indépendantes, mais dès qu'on rajoute une ligne nulle, il n'y a plus indépendance.

En partant d'une matrice $(A|b)$ quelconque, l'algorithme termine sur une matrice $(A'|b')$ où b' est un vecteur dont les coordonnées sont de la forme $b'_i = f_i(b_1, \dots b_p)$ où f_i est linéaire. Autrement dit, après application de l'algorithme, on arrive à un système :

$$\begin{cases} \boxed{x_1} + & x_2 + & \dots & 0 & + \dots & 0 & \dots & * & = f_1(b_1, \dots, b_p) \\ & & & \boxed{x_{k_2}} & + \dots & 0 & \dots & * & = f_2(b_1, \dots, b_p) \\ & & & & & \vdots & & \vdots & \vdots \\ & & & & & \boxed{x_{k_r}} & \dots & * & = f_r(b_1, \dots, b_p) \\ & & & & & & & 0 & = f_{r+1}(b_1, \dots, b_p) \\ & & & & & & & & \vdots \\ & & & & & & & 0 & = f_n(b_1, \dots, b_p) \end{cases}$$

On voit donc que pour que ce système ait une solution, il faut que $f_{r+1}(b_1, \dots, b_p) = \dots = f_p(b_1, \dots, b_p) = 0$. On a donc obtenu une équation pour les b tels que le

système a une solution :

$$(\mathcal{S}) \text{ a une solution} \Leftrightarrow f_{r+1}(b_1, \ldots, b_p) = \ldots = f_p(b_1, \ldots, b_p) = 0.$$

On peut maintenant énoncer le théorème :

Théorème 5.2.7. — *Soient* $A \in \mathcal{M}_{p,n}(\mathbf{K})$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{K}^n$ *et* $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in \mathbf{K}^p$. *Soit* $(\mathcal{S})$ *le système* $AX = b$ *et* $(A|b)$ *sa matrice augmentée.* (QC)

1. *Après application du pivot de Gauss, le système devient* $(\mathcal{S}')$ *de matrice augmentée* $(A'|b')$ *où* A' *est échelonnée réduite à* r *lignes non nulles, les pivots étant en colonnes* $1 \leq k_1 < \ldots < k_r \leq n$. *De plus les coefficients de* b' *sont de la forme* $b'_i = f_i(b_1, \ldots, b_r)$ *où les* f_i *sont des applications linéaires (sommes avec coefficients des coordonnées de* b*).*
2. *Le système a des solutions si et seulement si* $f_{r+1}(b_1, \ldots, b_p) = \ldots = f_p(b_1, \ldots, b_p) = 0$.
3. *Si le système a une solution et qu'il y a autant de pivots que de colonnes, le système a une unique solution.*
4. *Si le système a une solution et qu'il y a moins de pivots que de colonnes, notons* j_1, …, j_{n-r} *les numéros de colonnes où n'apparaît pas de pivot. On appelle* x_{j_1}, …, $x_{j_{n-r}}$ *les* variables libres. *Pour chaque valeur des variables libres, le système a une unique solution.*

Ce théorème, avec le théorème concluant l'étude théorique, permet donc de décrire explicitement les solutions de $(\mathcal{S})$ après avoir appliqué le pivot de Gauss.

Démonstration. — Le premier point a déjà été vu.

Le deuxième point découle de la forme du système $(A'|b')$ ci-dessus : $b \in \mathrm{im}(A)$ implique qu'il existe X tel que $A'X = b'$. Ça implique donc que

$$f_{r+1}(b_1, \ldots, b_p) = \ldots = f_p(b_1, \ldots, b_p) = 0.$$

Montrons la réciproque sur un exemple : supposons que

$$(A'|b') = \left(\begin{array}{ccccc|c} \boxed{1} & 1 & 0 & 1 & 0 & f_1(b) \\ 0 & 0 & \boxed{1} & 1 & 0 & f_2(b) \\ 0 & 0 & 0 & 0 & \boxed{1} & f_3(b) \\ 0 & 0 & 0 & 0 & 0 & f_4(b) \end{array}\right)$$

et que f_4 vaut zéro. Remarquons alors que $k_1 = 1$, $k_2 = 3$ et $k_3 = 5$ et que les variables libres sont celles d'indice $j_1 = 2$ et $j_2 = 4$. Une solution du système

$(A'|b')$ est alors une solution du système [1] :

$$(\mathcal{S}_1) \ : \ \begin{cases} x_{k_1} & & & = & f_1(b) - x_{j_1} - x_{j_2} \\ & x_{k_2} & & = & f_2(b) - x_{j_2} \\ & & x_{k_3} & = & f_3(b) \end{cases}$$

On trouve donc une unique solution à ce système pour chaque valeur de x_{j_1} et x_{j_2}. Et s'il n'y pas de variable libre, il n'y a pas de paramètres, donc une unique solution. □

Dans le cas où il y a autant de variables que d'équations (la matrice A est carrée) et que le pivot de Gauss mène à une matrice échelonnée réduite sans ligne nulle (les pivots sont sur la diagonale), la résolution du système $AX = b$ mène toujours à une unique solution. C'est le cas où la matrice A est inversible, déjà abordé en début de chapitre. Nous verrons dans le prochain chapitre que cette méthode nous a en fait permis de calculer l'inverse de A.

5.2.4. Corollaires. — Une conséquence de l'algorithme est la suivante :

Proposition 5.2.8. — *Soit A une matrice de $\mathcal{M}_{pn}(\mathbf{K})$. Alors il existe une matrice $P \in \mathcal{M}_p(\mathbf{K})$, inversible, telle que PA est échelonnée réduite.*

On peut montrer que la matrice échelonnée réduite est unique, mais nous ne l'utiliserons pas dans ce cours.

Démonstration. — On applique l'algorithme à la matrice A : à chaque étape, on multiplie à gauche par une des matrices $D_\alpha, T_{ij}(\lambda)$ ou P_{ij}. Notons M_1, $M_2, \ldots,$ M_N ces matrices inversibles. Ainsi on a obtenu que $M_N M_{N-1} \ldots M_1 A$ est échelonnée réduite.

Considérons $P = M_N M_{N-1} \ldots M_1$. C'est bien une matrice inversible, car son inverse est $M_1^{-1} M_2^{-1} \ldots M_N^{-1}$. La proposition est démontrée □

Dans le cas des matrices carrées, on obtient :

Proposition 5.2.9. — *Soit A une matrice carrée et P une matrice inversible telle que PA est échelonnée réduite. Deux cas sont possibles :*

- *soit $PA = I_p$ et alors $P = A^{-1}$*
- *soit PA a au moins une ligne nulle et alors A n'est pas inversible.*

1. On ignore la dernière ligne car on a supposé que $f_4(b) = 0$.

Démonstration. — Dans le premier cas, on a $PA = I_p$. On verra au chapitre suivant que c'est suffisant pour que A soit inversible et $P = A^{-1}$.

Dans le second cas, le système $AX = 0$ a une infinité de solution : en effet, PA a une ligne nulle, donc n'a qu'au plus $p-1$ pivots. Il y a donc une variable libre, ce qui implique une infinité de solutions au système. Or on a vu que si A était inversible, le système aurait une unique solution. Par contraposée, A n'est pas inversible. □

Ça donne même un algorithme pour déterminer si $A \in \mathcal{M}_n(\mathbf{K})$ est inversible et si oui calculer son inverse :
Première étape : Considérer la matrice $(A|I_p)$ (qui a p lignes et $2p$ colonnes) et en agissant sur ses lignes par la méthode du pivot de Gauss aboutir à une matrice $(A'|B)$ où A' est échelonné réduite.

On vient de voir qu'agir sur les lignes revient à multiplier la matrice par une matrice P inversible. On a donc $(A'|B) = P(A|I_p) = (PA|P)$.

Deux cas sont alors possibles :
Premier cas : La matrice $A' = PA$ a une ligne nulle. Alors A n'est pas inversible.
Deuxième cas : Si A' n'a pas de ligne nulle, A est inversible et son inverse est $P = B$.

Remarquons que si au cours de l'application de l'algorithme (même avant d'avoir fini), on crée une ligne nulle dans A', alors on peut tout de suite s'arrêter : cette ligne restera nulle jusqu'à la fin de l'algorithme et donc A n'est pas inversible. En pratique, il est souvent pratique de commencer par faire la première étape de l'algorithme de Gauss, pour arriver à une matrice échelonnée. A ce moment, on peut discuter de l'inversibilité ou non. Si la matrice est inversible, on continue les opérations sur les lignes pour arriver à $A' = I_p$.

Exemple 5.2.10. — Considérons la matrice $A_x = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 1 & x \end{pmatrix}$. On mène alors les opérations sur les lignes :

$$(A_x|I_3) \xrightarrow[L_3 \leftarrow L_3 - L_1]{L_2 \leftarrow L_2 - L_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & x-1 & -1 & 0 & 1 \end{array}\right)$$

La matrice de gauche est échelonnée :

$$A'_x = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & x-1 \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Premier cas : $x = 1$, alors A'_1 a une ligne nulle, donc A_1 n'est pas inversible.
Deuxième cas : $x \neq 1$. Alors, on passe à la deuxième étape de l'algorithme :

$$\begin{aligned}
(A'_x|P) \quad &\xrightarrow{L_4 \leftarrow \frac{1}{x-1} L_4} \quad \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{x-1} & 0 & \frac{1}{x-1} \end{array}\right) \\
&\xrightarrow{L_1 \leftarrow L_1 - L_2} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 2 & -1 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{x-1} & 0 & \frac{1}{x-1} \end{array}\right) \\
&\xrightarrow[L_2 \leftarrow L_2 - 3L_3]{L_1 \leftarrow L_1 + 2L_3} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2x-4}{x-1} & -1 & \frac{2}{x-1} \\ 0 & 1 & 0 & \frac{-x+4}{x-1} & 1 & \frac{-3}{x-1} \\ 0 & 0 & 1 & -\frac{1}{x-1} & 0 & \frac{1}{x-1} \end{array}\right)
\end{aligned}$$

La matrice de gauche est bien I_3. L'inverse de A_x, pour $x \neq 1$ est donc :

$$A_x^{-1} = \begin{pmatrix} \frac{2x-4}{x-1} & -1 & \frac{2}{x-1} \\ \frac{-x+4}{x-1} & 1 & \frac{-3}{x-1} \\ -\frac{1}{x-1} & 0 & \frac{1}{x-1} \end{pmatrix}.$$

On laisse le lecteur vérifier qu'on a bien $A_x A_x^{-1} = I_4$.

CHAPITRE 6

DÉTERMINANTS

Dans ce chapitre $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$ est un corps. On travaillera dans les espaces de matrices carrées $\mathcal{M}_n(\mathbf{K})$.

6.1. Cas des matrices 2×2

Définition 6.1.1. — Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Son *déterminant* est le scalaire (QC)

$$\text{dét}(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Faisons une série de remarque à propos de ce nombre. Certaines sont des questions de cours : on donnera un énoncé général quand on aura défini le déterminant en toute dimension, mais il est bon de le vérifier dans le cas plus concret des matrices de taille 2. On peut déjà remarquer que la valeur absolue de dét(A) est l'aire du parallélogramme de côtés les vecteurs $\Big(a, c\Big)$ et $\Big(b, d\Big)$.

1. dét$(AB) =$ dét(A)dét(B).
2. Considérons la matrice $A' = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Alors on calcule [1] :

$$AA' = A'A = \text{dét}(A) I_2.$$

Proposition 6.1.2. — *A est inversible si et seulement si* dét$(A) \neq 0$. *Dans ce cas,* $A^{-1} = \frac{1}{\text{dét}(A)} A'$. (QC)

1. Rappelons que $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ est la matrice identité.

Démonstration. — Supposons A inversible. Le premier point ci-dessus permet d'écrire :

$$\text{dét}(A)\text{dét}(A^{-1}) = \text{dét}(AA^{-1}) = \text{dét}(I_2) = 1.$$

On en déduit que $\text{dét}(A) \neq 0$.

Réciproquement, si $\text{dét}(A) \neq 0$, en utilisant le point 2 ci -dessus, on a :

$$\frac{1}{\text{dét}(A)} A'A = A \frac{1}{\text{dét}(A)} A' = I_2.$$

A est donc inversible, d'inverse $A^{-1} = \frac{1}{\text{dét}(A)} A'$. □

D'autres propriétés :

1. On peut constater directement que $\text{dét}(A) = \text{dét}({}^tA)$.
2. Si les deux colonnes de A sont égales, alors $\text{dét}(A) = 0$. On dit que le déterminant est *alterné*.
3. Le déterminant de A est *multilinéaire* : pour tous scalaires λ et λ', on a

$$\begin{vmatrix} \lambda a + \lambda' a' & b \\ \lambda c + \lambda' c' & d \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \lambda' \begin{vmatrix} a' & b \\ c' & d \end{vmatrix} \quad \text{et} \quad \begin{vmatrix} a & \lambda b + \lambda' b' \\ c & \lambda d + \lambda' d' \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \lambda' \begin{vmatrix} a & b' \\ c & d' \end{vmatrix}.$$

6.2. Cas général

Le but de ce chapitre est de construire et d'étudier une fonction dét : $\mathcal{M}_n(\mathbf{K}) \to \mathbf{K}$ qui ressemble à la fonction introduite dans le cas des matrices de taille 2 ; notamment qui soit capable de déceler si A est inversible (et qui permette d'écrire une formule pour son inverse). Pour ça, on donne une définition par récurrence.

Précisons quelques définitions.

Définition 6.2.1. — Une application $D : \mathcal{M}_n(\mathbf{K}) \to \mathbf{K}$ est dite *linéaire par rapport aux colonnes* ou *multilinéaire.* si elle vérifie que pour toute matrice $B = (C_1|C_2|\dots|C_n) \in \mathcal{M}_n(\mathbf{K})$, pour tout $1 \leq j \leq n$, pour tout vecteur colonne $C'_j \in \mathcal{M}_{n,1}(\mathbf{K})$ et tout $\alpha \in \mathbf{K}$, on a [2]

$$\begin{aligned} D(C_1|\dots|\alpha C_j + C'_j|\dots|C_n) &= \alpha D(C_1|\dots|C_j|\dots|C_n) + D(C_1|\dots|C'_j|\dots|C_n) \\ &= \alpha D(B) + D(C_1|\dots|C'_j|\dots|C_n) \end{aligned}$$

Elle est dite *alternée* si quand deux colonnes de A sont égales, $\varphi(A) = 0$.

Si $A = (a_{ij})$ est une matrice carrée de taille n, on note $A - L_i - C_j$ la matrice construite en enlevant de A la i-ème ligne et la j-ème colonne. C'est une matrice carrée de taille $n-1$. Supposons donc qu'on sait défini le déterminant en taille

2. Dans la formule suivante, seule la j-ième colonne change.

$n-1$ et notons $\Delta_{ij} = (-1)^{i+j}\text{dét}(A - L_i - C_j)$. On appelle Δ_{ij} le cofacteur d'indice i, j de A. Par exemple, si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, alors $A - L_2 - C_3 = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix}$.

Et $\Delta_{23} = (-1)^{2+3}\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = -(8-14) = 6$.

On énonce maintenant le théorème général sur le déterminant, qui permet de le définir par récurrence sur la dimension :

Théorème 6.2.2. — *Soit $n \in \mathbf{N}^*$. Il existe une unique application, appelée le déterminant et notée* dét*, de $\mathcal{M}_n(\mathbf{K})$ dans $\mathbf{K}$ multilinéaire et alternée telle que* $\text{dét}(I_n) = 1$. *De plus toute fonction D multilinéaire est alternée vérifie $D(A) = D(I_n)\text{dét}(A)$.*

Enfin, on a les formules suivantes.

– *Développement du déterminant selon la ligne i :* (QC)

$$\text{dét}(A) = \sum_{j=1}^{n} a_{ij}\Delta_{ij}.$$

– *Développement du déterminant selon la colonne j :*

$$\text{dét}(A) = \sum_{i=1}^{n} a_{ij}\Delta_{ij}.$$

On notera parfois $|A|$ le déterminant de A.

On renvoie à la section annexe de ce chapitre (non traitée en amphi – à la fin du chapitre) pour la preuve générale. Nous traitons dans la section suivante le cas des matrices de taille 3, qui contient les idées essentielles

6.2.1. Cas des matrices de taille 3. — Dans cette section, on considère une fonction D de $\mathcal{M}_3(\mathbf{K})$ dans $\mathbf{K}$ qui est multilinéaire et alternée et telle que $D(I_3) = 1$. On va montrer qu'il n'y a qu'une seule possibilité, qu'on appelera le déterminant. Cette section sert aussi d'exemple pour les notions abstraites (multilinéarité, cofacteurs...) définies ci-dessus.

Soit donc $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. On veut montrer qu'on n'a pas le choix pour déterminer $D(A)$. On commence par utiliser la linéarité par rapport à la première colonne.

Notons $C_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = a_{11} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_{21} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_{31} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. La linéarité de D par rapport à la première colonne permet d'écrire :

$$D(A) = a_{11} D \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} + a_{21} D \begin{pmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} + a_{31} D \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 1 & a_{32} & a_{33} \end{pmatrix}$$

Étudions le premier de ces termes : $D \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix}$. On peut maintenant décomposer la deuxième colonne en somme :

$$C_2 = a_{12} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_{32} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

En utilisant la linéarité par rapport à cette colonne, on arrive à :

$$D \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} = a_{12} D \begin{pmatrix} 1 & 1 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} + a_{22} D \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} + a_{31} D \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 1 & a_{33} \end{pmatrix}$$

Mais, dans le premier de ces termes, les deux premières colonnes sont égales, donc on peut utiliser la propriété d'alternance : $D \begin{pmatrix} 1 & 1 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} = 0$.

On a donc $D \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} = a_{22} D \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} + a_{31} D \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 1 & a_{33} \end{pmatrix}$. En utilisant la linéarité par rapport à la troisième colonne et l'alternance quand deux colonnes sont égales, on obtient successivement :

$$\begin{aligned} D \begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} &= a_{13} D \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{23} D \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} + a_{33} D \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= a_{33} D \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et} \end{aligned}$$

$$D\begin{pmatrix}1&0&a_{13}\\0&0&a_{23}\\0&1&a_{33}\end{pmatrix} = a_{13}D\begin{pmatrix}1&0&1\\0&0&0\\0&1&0\end{pmatrix} + a_{23}D\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix} + a_{33}D\begin{pmatrix}1&0&0\\0&0&0\\0&1&1\end{pmatrix}$$
$$= a_{23}D\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$$

On a donc réussi à calculer le premier terme de la décomposition de $D(A)$:

$$a_{11}D\begin{pmatrix}1&a_{12}&a_{13}\\0&a_{22}&a_{23}\\0&a_{32}&a_{33}\end{pmatrix} = a_{11}a_{22}a_{33}D\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + a_{11}a_{32}a_{23}D\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$$

En raisonnant de même sur les 2 autres termes, on trouve :

$$\begin{aligned} D(A) = {} & a_{11}a_{22}a_{33}D\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + a_{11}a_{32}a_{23}D\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix} \\ & + a_{21}a_{12}a_{33}D\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix} + a_{21}a_{32}a_{13}D\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix} \\ & + a_{31}a_{12}a_{23}D\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix} + a_{31}a_{22}a_{13}D\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix} \end{aligned}$$

On voit donc que D ne dépend que de ses valeurs sur les 6 matrices avec des 0 et des 1 ci-dessus. Pour conclure, il suffit de comprendre ces 6 valeurs. Pour ça, la propriété générale suivante :

Proposition 6.2.3. — *Soient $D : \mathcal{M}_n(\mathbf{K}) \to \mathbf{K}$ multilinéaire et alternée, $A \in \mathcal{M}_n(\mathbf{K})$ et A' obtenue à partir de A en échangeant deux colonnes de A. Alors on a $D(A) = -D(A')$.*

Démonstration. — Notons $A = (C_1|\cdots|C_i|\cdots|C_j|\cdots|C_n)$ et
$A' = (C_1|\cdots|C_j|\cdots|C_i|\cdots|C_n)$ (on a échangé les colonnes i et j).

Considérons la matrice $B = (C_1|\cdots|C_i+C_j|\cdots|C_i+C_j|\cdots|C_n)$. Elle a deux colonnes égales, donc $D(B) = 0$. Mais en utilisant la linéarité par rapport à la colonne i, on a

$$D(B) = D(C_1|\cdots|C_i|\cdots|C_i+C_j|\cdots|C_n) + D(C_1|\cdots|C_j|\cdots|C_i+C_j|\cdots|C_n).$$

On utilise dans les deux termes ci-dessus la linéarité par rapport à la colonne j pour obtenir :

$$\begin{aligned} D(B) &= D(C_1|\cdots|C_i|\cdots|C_i|\cdots C_n) + D(C_1|\cdots|C_i|\cdots|C_j|\cdots C_n) \\ &\quad + D(C_1\cdots|C_j|\cdots|C_i|\cdots C_n) + D(C_1\cdots|C_j|\cdots|C_j|\cdots C_n) \\ &= 0 + D(A) + D(A') + 0 \end{aligned}$$

Les deux zéros sont obtenus car deux colonnes des matrices correspondantes sont égales.

Finalement $D(A) + D(A') = D(B) = 0$, donc $D(A) = -D(A')$. □

Revenons donc à nos six matrices. On obtient :

$$D\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix} = D\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix} = D\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix} = -D(I_3)$$

car on peut passer de chacune de ces trois matrices à I_3 en échangeant deux colonnes.

En revanche $D\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix} = D\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix} = D(I_3)$ car il faut 2 échanges pour revenir à I_3.

En se rappelant qu'on a supposé $D(I_3) = 1$ on obtient bien l'unicité (et donc nous revenons à la notation dét) :

$$\text{dét}(A) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.$$

Cette formule s'appelle la formule de Sarrus. Attention à son utilisation dans les calculs, elle est souvent mal appliquée.

On peut maintenant comprendre les formules de développement suivant les lignes ou les colonnes. Par exemple, pour la première colonne, on groupe les termes :

$$\begin{aligned} \text{dét}(A) &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \\ &= a_{11}\text{dét}\begin{pmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{pmatrix} - a_{21}\text{dét}\begin{pmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{pmatrix} + a_{31}\text{dét}\begin{pmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{pmatrix} \end{aligned}$$

On reconnait des cofacteurs, comme définis au début de cette section : $\text{dét}(A) = a_{11}\Delta_{11}(A) + a_{21}\Delta_{21}(A) + a_{31}\Delta_{31}(A)$.

En factorisant d'abord les termes de la première ligne, on trouve : $\text{dét}(A) = a_{11}\Delta_{11}(A) + a_{12}\Delta_{12}(A) + a_{13}\Delta_{13}(A)$.

6.2.2. Retour au cas général : quelques propriétés du déterminant. —

Proposition 6.2.4. — *Si une matrice* $A = \begin{pmatrix} a_1 & & & \\ \star & a_2 & & \\ \vdots & \ddots & \ddots & \\ \star & \dots & \star & a_n \end{pmatrix}$ *est triangulaire inférieure, alors* $\det(A) = a_1 \dots a_n$.

Démonstration. — On le prouve directement par récurrence sur n : par définition du déterminant $\begin{vmatrix} a_1 & & & \\ \star & a_2 & & \\ \vdots & \ddots & \ddots & \\ \star & \dots & \star & a_n \end{vmatrix} = a_1 \begin{vmatrix} a_2 & & \\ \ddots & \ddots & \\ \dots & \star & a_n \end{vmatrix} + 0 = a_1 a_2 \cdots a_n$

□

On en tire du théorèmes les deux conséquences suivantes :

Proposition 6.2.5. — *Pour toutes matrices A et B dans $\mathcal{M}_n(\mathbf{K})$, on a* (QC)

$$\det(AB) = \det(A)\det(B) \quad \textit{et}$$
$$\det({}^tA) = \det(A).$$

Démonstration. — Fixons $A \in \mathcal{M}_n(\mathbf{K})$ et considérons la fonction D définie par $D(B) = \det(AB)$. Montrons qu'elle est multilinéaire et alternée. Remarquons pour ça que si $B = (C_1|\dots|C_n)$, alors $AB = (AC_1|\dots[AC_n)$.

L'alternance est alors immédiate : si deux colonnes de B sont égales, alors deux colonnes de AB sont égales et donc (par alternance de dét), $D(B) = \det(AB) = 0$.

Pour la multilinéarité, fixons V un vecteur colonne, $1 \leq j \leq n$ et $\alpha \in \mathbf{K}$. Modifions la matrice B' en modifiant sa j-ème colonne. Notons $B' = (C_1|\dots|\alpha C_j + V|\dots|C_n$ et $B'' = (C_1|\dots|V|\dots|C_n)$. On veut écrire : $D(B') = \alpha D(B) + D(B'')$. Remarquons que, par linéarité de $C \to AC$, on a :

$$AB' = (AC_1|\dots|A(\alpha C_j + V)|\dots|C_n) = (AC_1|\dots|\alpha AC_j + AV)|\dots|C_n).$$

Donc, par multilinéarité de dét, on obtient la multilinéarité :

$$\begin{aligned} D(B') = \det(AB') & \\ &= \alpha\det(AC_1|\dots|AC_j|\dots|AC_n) + \det(AC_1|\dots|AV|\dots|AC_n) \\ &= \alpha\det(AB) + \det(AB'') \\ &= \alpha D(B) + D(B''). \end{aligned}$$

On peut donc appliquer la partie unicité du théorème : pour toute matrice B, on a $D(B) = D(I_n)\text{dét}(B)$. Or $D(I_n) = \text{dét}(AI_n) = \text{dét}(A)$. On a bien prouvé

$$\text{dét}(AB) = \text{dét}(A)\text{dét}(B).$$

Pour la transposée : on procède par récurrence sur la taille n de la matrice. C'est évident si $n = 1$: ${}^tA = A$.

Supposons que ce soit vrai au rang n. Développons le déterminant de ${}^tA$ le long de la première ligne :

$$\text{dét}({}^tA) = \sum_{k=1}^{n} ({}^tA)_{1k}(-1)^{1+k}\Delta_{1k}({}^tA).$$

Or $({}^tA)_{1k} = a_{k1}$ et $\Delta_{1k}({}^tA)$ est le déterminant de la matrice ${}^tA - L_1({}^tA) - C_k({}^tA)$. Comme la première ligne de ${}^tA$ est la première colonne de A et la k-ième colonne de ${}^tA$ est la k-ième ligne de A, on a l'égalité de matrices :

$${}^tA - L_1({}^tA) - C_k({}^tA) = {}^t(A - C_1(A) - L_k(A)).$$

Donc, par hypothèse de récurrence, les déterminants de ces deux matrices sont égaux. Autrement dit on a :

$$\Delta_{1k}({}^tA) = \Delta_{k1}(A).$$

On obtient ainsi

$$\text{dét}({}^tA) = \sum_{k=1}^{n} a_{k1}(-1)^{1+k}\Delta_{k1}(A).$$

On reconnaît dans le membre de droite le développement le long de la première colonne de dét(A). On obtient bien ainsi l'égalité voulue :

$$\text{dét}({}^tA) = \text{dét}(A).$$

□

Corollaire 6.2.6. — *Si une matrice est triangulaire supérieure, son déterminant est le produit des coefficients diagonaux.*

Démonstration. — Si A est triangulaire supérieure, alors ${}^tA$ est triangulaire inférieure et a les mêmes coefficients diagonaux. Donc dét$(A) =$ dét$({}^tA)$ est le produit des coefficients diagonaux. □

6.3. Déterminant et inversion

Définition 6.3.1. — Soit $A \in \mathcal{M}_n(\mathbf{K})$. Soient i et j deux entiers entre 1 et n. Alors on appelle *comatrice de* A la matrice des cofacteurs, notée $\mathrm{com}(A)$, c'est-à-dire que le coefficient d'indice i, j de $\mathrm{com}(A)$ est le cofacteur d'indice i, j :

$$(\mathrm{com}(A))_{ij} = \Delta_{ij}(A).$$

Proposition 6.3.2. — *Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors on a*

$$A^t\mathrm{com}(A) = \mathrm{dét}(A)I_n = {}^t\,\mathrm{com}(A)A.$$

Démonstration. — Calculons le coefficient d'indice i, j du produit $A^t\mathrm{com}(A)$:

$$\begin{aligned}(A^t\mathrm{com}(A))_{ij} &= \sum_{k=1}^{n} a_{ik}({}^t\mathrm{com}(A))_{kj} \\ &= \sum_{k=1}^{n} a_{ik}(\mathrm{com}(A))_{jk} \\ &= \sum_{k=1}^{n} a_{ik}(-1)^{j+k}\Delta_{jk}(A).\end{aligned}$$

On reconnaît dans cette dernière formule le développement par rapport à la ligne j du déterminant de la matrice obtenue à partir de A en remplaçant sa j-ème ligne par la ligne L_i.

Ainsi, si $i = j$, le coefficient vaut $\mathrm{dét}(A)$. En revanche si $i \neq j$, la matrice dont on calcule le déterminant a deux lignes égales, donc son déterminant est nul.

On a bien montré que les coefficients de $A^t\mathrm{com}(A)$ sont nuls, sauf les coefficients diagonaux qui valent tous $\mathrm{dét}(A)$. Autrement dit, $A^t\mathrm{com}(A) = \mathrm{dét}(A)I_n$.

L'autre égalité se montre de la même manière, mais en reconnaissant un développement le long d'une colonne. □

Grâce à ce calcul, on obtient le théorème :

Théorème 6.3.3. — *Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A est inversible si et seulement si* $\mathrm{dét}(A) \neq 0$. *Dans ce cas, on a :* (QC)

$$A^{-1} = \frac{1}{\mathrm{dét}(A)}{}^t\mathrm{com}(A) \text{ et } \mathrm{dét}(A^{-1}) = \frac{1}{\mathrm{dét}(A)}.$$

Démonstration. — Supposons A inversible. Alors on a : $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1$. Ça montre que $\det(A) \neq 0$ et que $\det(A^{-1}) = \frac{1}{\det(A)}$.

Supposons maintenant que $\det(A) \neq 0$. Alors, la proposition précédente permet d'écrire :

$$A\left(\frac{1}{\det(A)}\,{}^t\mathrm{com}(A)\right) = I_n = \left(\frac{1}{\det(A)}\,{}^t\mathrm{com}(A)\right)A.$$

Ça montre bien que A est inversible d'inverse $\frac{1}{\det(A)}\,{}^t\mathrm{com}(A)$. □

6.4. Calcul de déterminants

Dans cette section, A est une matrice de $\mathcal{M}_n(\mathbf{K})$, L_i sont ses lignes et C_i ses colonnes.

Prenons le temps de définir une notation qui nous simplifiera les écritures par la suite :

Pour $\alpha \in \mathbf{K}$ et $1 \leq i \neq j \leq n$, on notera $A' \xleftarrow{L_i \leftarrow L_i + \alpha L_j} A$ pour signifier que A' est obtenue à partir de A en remplaçant la ligne L_i par $L_i + \alpha L_j$.

De même, si $\alpha \neq 0$, on notera $A' \xleftarrow{L_i \leftarrow \alpha L_i} A$ pour signifier que A' est obtenue à partir de A en remplaçant la ligne L_i par αL_i ; et on notera $A' \xleftarrow{L_i \leftrightarrow L_j} A$ pour signifier que A' est obtenue à partir de A en échangeant les lignes i et j.

Enfin, on notera $A' \xleftarrow{C_i \leftarrow C_i + \alpha C_j} A$ pour signifier que A' est obtenue à partir de A en remplaçant la colonne C_i par $C_i + \alpha C_j$.

Enfin, si $\alpha \neq 0$, on notera $A' \xleftarrow{C_i \leftarrow \alpha C_i} A$ pour signifier que A' est obtenue à partir de A en remplaçant la colonne C_i par αC_i ; et on notera $A' \xleftarrow{C_i \leftrightarrow C_j} A$ pour signifier que A' est obtenue à partir de A en échangeant les colonnes i et j.

Proposition 6.4.1. — *Soient A et A' deux matrices de $\mathcal{M}_n(\mathbf{K})$.*

1. *Si $A' \xleftarrow{L_i \leftarrow L_i + \alpha L_j} A$ ou $A' \xleftarrow{C_i \leftarrow C_i + \alpha C_j} A$, alors* $\det(A) = \det(A')$.
2. *Si $A' \xleftarrow{L_i \leftarrow \alpha L_i} A$ ou $A' \xleftarrow{C_i \leftarrow \alpha C_i} A$ (avec $\alpha \neq 0$), alors* $\det(A') = \alpha\det(A)$.
3. *Si $A' \xleftarrow{L_i \leftrightarrow L_j} A$ ou $A' \xleftarrow{C_i \leftrightarrow C_j} A$, alors* $\det(A') = -\det(A)$.

Démonstration. — Dans le cas des colonnes le dernier point a déjà été vu, le deuxième est conséquence direct de la linéarité. Il reste à montrer le premier : soit $i \neq j$, $\alpha \in \mathbf{K}$, et $A = (C_1|\dots|C_j|\dots|C_n)$. Modifions la j-ème colonne

pour définir la matrice $A' = (C_1|\dots|C_j + \alpha C_i|\dots|C_n)$. Par multilinéarité, on a

$$D(A') = D(A) + \alpha D(C_1|\dots|C_i|\dots|C_n).$$

Or la deuxième des matrices apparaissant à droite contient deux fois la colonne C_i : comme i-ème colonne, mais aussi comme j-ème colonne. Par alternance, D s'annule donc. Ainsi $D(A') = D(A)$.

Pour montrer le cas des lignes, on utilise $\det(A') = \det({}^tA')$ et $\det(A) = \det({}^tA)$ et que si on passe de A à A' par une opération élémentaire sur les lignes, on passe de ${}^tA$ à ${}^tA'$ par la même opération élémentaire, mais sur les colonnes. □

Remarque 6.4.2. — Notamment, dans tous les cas, $\det(A) = 0 \Leftrightarrow \det(A') = 0$.

Exemple 6.4.3. — **Premier exemple :**

$$\begin{vmatrix} 1 & 3 & 2 \\ 1 & 4 & 3 \\ 1 & 5 & 4 \end{vmatrix} = 0.$$

En effet, on remarque $\begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$. Donc on a

$$\begin{vmatrix} 1 & 3 & 2 \\ 1 & 4 & 3 \\ 1 & 5 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \\ 1 & 1 & 4 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \\ 1 & 4 & 4 \end{vmatrix}.$$

Or les deux matrices à droite ont deux colonnes égales, donc leur déterminant est nul.

Remarquons que le fait que les colonnes soient liées implique que le rang de A est $< n$. On peut aussi utiliser la prochaine proposition qui dit que son déterminant est nul.

Deuxième exemple : Calculons $D = \begin{vmatrix} 1 & 2 & 4 \\ 8 & 7 & n \\ 4 & 2 & 1 \end{vmatrix}$. Les deux opérations $C_2 \leftarrow C_2 - 2C_1$, et $C_3 \leftarrow C_3 - 4C_1$ ne changent pas le déterminant. On obtient :

$$D = \begin{vmatrix} 1 & 0 & 0 \\ 8 & -9 & n-32 \\ 4 & -6 & -15 \end{vmatrix} = 1 \cdot \begin{vmatrix} -9 & n-32 \\ -6 & -15 \end{vmatrix} = 9 \times 15 + 6 \times (n-32) = 6n - 57.$$

Annexe : preuve de l'existence du déterminant ; Non traité en amphi

On montre en fait le théorème plus précis suivant :

Théorème 6.4.4. — *Soit $n \in \mathbf{N}^*$. Le déterminant est une application de $\mathcal{M}_n(\mathbf{K})$ dans $\mathbf{K}$ multilinéaire et alternée*

(Multilinéarité) dét *est linéaire par rapport aux colonnes ;*

(Alternance) dét *s'annule quand deux colonnes de la matrice sont égales*[3] *;*

(Normalisation) $\text{dét}(I_n) = 1$.

De plus une telle fonction est unique. Mieux : si D est une application de $\mathcal{M}_n(\mathbf{K})$ dans $\mathbf{K}$ multilinéaire et alternée, alors pour toute matrice $A \in \mathcal{M}_n(\mathbf{K})$, on a :

$$D(A) = D(I_n)\text{dét}(A).$$

Les formules de développement suivant les lignes colonnes seront montrées au cours de la preuve.

Avant de se lancer dans la preuve du théorème, observons une conséquence des hypothèses de multilinéarité et d'alternance :

Proposition 6.4.5. — *Soit D une application de $\mathcal{M}_n(\mathbf{K})$ dans $\mathbf{K}$ multilinéaire et alternée. Alors on a*

- *D change de signe si on échange deux colonnes.*
- *D ne change pas si on ajoute à une colonne un multiple d'une autre.*

Remarque 6.4.6. — Quand on aura prouvé le théorème, cette proposition s'appliquera notamment à la fonction dét. Donc on aura l'énoncé « le déterminant change de signe quand on échange deux colonnes et ne change pas si on ajoute à une colonne un multiple d'une autre. » C'est le cas de toute la série de proposition qu'on montrera au cours de la preuve du théorème.

Démonstration. — On a déjà montré ces propriétés. □

On va montrer le théorème 6.2.2 par récurrence sur la dimension n.

6.4.1. Initialisation. — Dans le cas $n = 1$, on a $\mathcal{M}_1(\mathbf{K}) = \{a \cdot I_1,\, a \in \mathbf{K}\}$. La condition de multilinéarité se réduit à de la linéarité (il n'y a qu'une seule colonne) ; la condition d'alternance est vide.

3. On dit qu'une telle application est *alternée*.

Donc l'application définie par $\text{dét}_1(a \cdot I_1) = a$ vérifie bien les trois conditions. De plus la partie unicité du théorème découle de la linéarité : si $D : \mathcal{M}_1(\mathbf{K}) \to \mathbf{K}$ est linéaire, alors

$$D(a \cdot I_1) = aD(I_1) = D(I_1)\text{dét}(a \cdot I_1).$$

6.4.2. Hérédité. — Soit $n \geq 1$. On suppose maintenant que le théorème est vrai pour l'entier n : il existe une application $\text{dét}_n : \mathcal{M}_n(\mathbf{K}) \to \mathbf{K}$ multilinéaire, alternée et normalisée. De plus, si $D : \mathcal{M}_n(\mathbf{K}) \to \mathbf{K}$ est multilinéaire et alternée, alors pour tout $A \in \mathcal{M}_n(\mathbf{K})$, $D(A) = D(I_n)\text{dét}_n(A)$.

On veut montrer que le théorème est encore vrai en dimension $n+1$. La proposition suivante montre que la partie unicité est vérifiée.

Proposition 6.4.7. — *Soit $D : \mathcal{M}_{n+1}(\mathbf{K}) \to \mathbf{K}$ multilinéaire et alternée. Alors, pour tout $A \in \mathcal{M}_{n+1}(\mathbf{K})$, pour tout $1 \leq j \leq n+1$, on a :*

$$D(A) = D(I_{n+1}) \sum_{i=1}^{n+1} a_{ij}\Delta_{ij}(A).$$

Démonstration. — Soit D une application multilinéaire et alternée. Notons

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, E_{n+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

la base canonique des vecteurs lignes. Notons $A = (C_1|\ldots|C_j|\ldots|C_{n+1})$. On a alors $C_j = \sum_{i=1}^{n+1} a_{ij}E_i$. Par linéarité de D, on obtient :

$$D(A) = \sum_{i=1}^{n+1} a_{ij}D(C_1|\ldots|E_i|\ldots|C_{n+1}).$$

On veut montrer que pour tout $1 \leq i \leq n+1$, on a $D(C_1|\ldots|E_i|\ldots|C_{n+1}) = (-1)^{i+j}\text{dét}_n(A - L_i - C_j)$.

Faisons-le sur un exemple pour éviter les notations trop lourdes. Il suffit à comprendre le raisonnement général. On suppose donc que $n = 5$, $j = 4$ et $i = 2$. Alors, si $A = (a_{ij})_{\substack{1 \leq i \leq 5 \\ 1 \leq j \leq 5}}$, on a

$$(C_1|\ldots|E_i|\ldots|C_{n+1}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & a_{15} \\ a_{21} & a_{22} & a_{23} & 1 & a_{25} \\ a_{31} & a_{32} & a_{33} & 0 & a_{35} \\ a_{41} & a_{42} & a_{43} & 0 & a_{45} \\ a_{51} & a_{52} & a_{53} & 0 & a_{55} \end{pmatrix}.$$

On a vu que D ne change pas si on ajoute à une colonne le multiple d'une autre. On peut donc remplacer les 4 colonnes C_k, pour $k \neq 4$ par $C_k - a_{2k}C_4$. On obtient :

$$D(C_1|\dots|E_i|\dots|C_{n+1}) = D\begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & a_{15} \\ 0 & 0 & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & a_{35} \\ a_{41} & a_{42} & a_{43} & 0 & a_{45} \\ a_{51} & a_{52} & a_{53} & 0 & a_{55} \end{pmatrix}.$$

Notons maintenant ψ l'application de $\mathcal{M}_4(\mathbf{K})$ dans $\mathbf{K}$, définie pour $B = (b_{ij})_{\substack{1 \leq i \leq 4 \\ 1 \leq j \leq 4}}$ par :

$$\psi(B) = D\begin{pmatrix} b_{11} & b_{12} & b_{13} & 0 & b_{14} \\ 0 & 0 & 0 & 1 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & b_{24} \\ b_{31} & b_{32} & b_{33} & 0 & b_{34} \\ b_{41} & b_{42} & b_{43} & 0 & b_{44} \end{pmatrix}.$$

On notera $\tilde{B}$ la matrice qui apparaît à droite. On vient de voir que $D(C_1|\dots|E_i|\dots|C_{n+1}) = \psi(A - L_i - C_j)$. Montrons que ψ est alternée et multilinéaire :

- Alternée : c'est clair ; si $B = (B_1|B_2|B_3|B_4)$ a deux colonnes égales, c'est aussi le cas de $\tilde{B}$.
- Multilinéaire : montrons la linéarité par rapport à la première colonne, les autres sont identiques. Soit B' la matrice obtenue en remplaçant la première colonne B_1 de B par $\alpha B_1 + V$ ($V = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$ est un vecteur colonne) et B'' celle obtenue en remplaçant B_1 par V ; alors la première colonne de $\tilde{B'}$ est $\alpha\tilde{B}_1 + \begin{pmatrix} v_1 \\ 0 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$. Par linéarité de D par rapport à la première colonne, on obtient $D(\tilde{B'}) = \alpha D(\tilde{B}) + D(\tilde{B''})$. En termes de la fonction ψ, on a bien

$$\psi(B') = \alpha\psi(B) + \psi(B'').$$

Ainsi, par unicité en dimension n, on a $\psi(B) = \psi(I_4)\text{dét}_n(B)$. Montrons que $\psi(I_4) = (-1)^{2+4}D(I_5)$. Tout d'abord, on a

$$\tilde{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Or $\psi(I_4) = D(\tilde{I}_4)$. On peut transformer $\tilde{I}_4$ en I_5 en échangeant les colonnes 3 et 4, puis 2 et 3. Il faut donc 2 échanges (en général, il en faut $j-i$, je continue le raisonnement avec ce facteur). On a vu que D changeait de signe quand on échangeait deux colonnes. Donc $\psi(I_4) = (-1)^{j-i}D(I_4) = (-1)^{i+j}D(I_5)$. La dernière égalité vient du fait que $j-i$ et $i+j$ diffèrent d'un nombre pair $(2i)$, donc $(-1)^{i+j} = (-1)^{j-i}(-1)^{2i} = (-1)^{j-i}$.

Finalement, on a obtenu $\psi(B) = (-1)^{i+j}D(I_5)\text{dét}(B)$. Donc $\psi(A - L_i - C_j) = (-1)^{i+j}D(I_5)\Delta_{ij}(A)$. C'est bien ce qu'on voulait montrer. □

Maintenant qu'on est assuré de la partie unicité, montrons l'existence en proposant une formule :

Proposition 6.4.8. — *Soit* $1 \leq i \leq n+1$. *On définit l'application* $D : \mathcal{M}_{n+1}(\mathbf{K}) \to \mathbf{K}$ *pour* $A = (a_{ij})_{\substack{1 \leq i \leq n+1 \\ 1 \leq j \leq n+1}}$ *par :*

$$D_i(A) = \sum_{j=1}^{n+1} (-1)^{i+j} a_{ij} \Delta_{ij}(A).$$

Alors D_i *est multilinéaire, alternée et normalisée.*

Remarque : la définition proposée dans la section précédente est cette formule pour le $i = 1$. On voit qu'on aurait pu choisir n'importe quelle ligne pour le définir.

Démonstration. —

Normalisée : Calculons $D_i(I_{n+1})$. Dans la somme, le seul indice j pour lequel a_{ij} est non nul est $j = i$. Dans ce cas-là, il vaut 1 et $(-1)^{i+j} = (-1)^{2i} = 1$. Donc $D_i(I_{n+1}) = \Delta_{ii}(I_{n+1})$. Or on vérifie que $I_{n+1} - L_i - C_i = I_n$. Donc, $\Delta_{ii}(I_{n+1}) = \text{dét}_n(I_n) = 1$.

Alternée : Supposons que les colonnes d'indices p et q de A soient égales, avec pour fixer les notations $p < q$. Alors, pour tous $j \neq p, q$, $A - L_i - C_j$ a

deux colonnes égales : celles correspondantes à p et q. Donc dans ce cas, $\Delta_{ij}(A) = 0$. De plus, $a_{ip} = a_{iq}$. Donc on est ramené à :

$$D_i(A) = a_{ip}((-1)^{i+p}\Delta_{ip}(A) + (-1)^{i+q}\Delta_{iq}(A)).$$

Or la colonne $C = C_p = C_q$ est à la place p dans $A - L_i - C_q$ et à la place $q-1$ dans $A - L_i - C_p$. Les autres colonnes étant égales, on passe de $A - L_i - C_q$ à $A - L_i - C_p$ en effectuant $q-1-p$ échanges de colonnes. Donc $\Delta_{ip}(A) = (-1)^{q-p-1}\Delta_{iq}(A)$. On en déduit :

$$\begin{aligned} D_i(A) &= a_{ip}\Delta_{iq}(A)((-1)^{i+p}(-1)^{q-p-1} + (-1)^{i+q}) \\ &= a_{ip}\Delta_{iq}(A)((-1)^{i+q-1} + (-1)^{i+q}) = 0. \end{aligned}$$

Multilinéaire : on fixe un indice de colonne l, et on montre la linéarité par rapport à cette colonne. Soit donc $\alpha \in \mathbf{K}$ et $V \in \mathbf{K}^n$ un vecteur colonne. Soit A' la matrice déduite de A en remplaçant la colonne C_l par $\alpha C_l + V$ et A'' celle où on remplace C_l par V. Calculons $D_i(A')$. Pour $j \neq l$, on a $a_{ij} = a'_{ij} = a''_{ij}$ et $\Delta_{ij}(A') = \alpha\Delta_{ij}(A) + \Delta_{ij}(A'')$. Réciproquement, si $j = l$, on a $a'_{il} = \alpha a_{il} + a''_{il}$ et $\Delta_{il}(A') = \Delta_{il}(A) = \Delta_{il}(A'')$. En tous cas, on obtient :

$$a'_{ij}\Delta_{ij}(A') = \alpha a_{ij}\Delta_{ij}(A) + \Delta_{ij}(A'').$$

En sommant sur j, on obtient bien $D_i(A') = \alpha D_i(A) + D_i(A'')$.

□

On peut maintenant finir la preuve du théorème :

Démonstration. — On peut poser $\text{dét}_{n+1} = D_i$. D'après la deuxième proposition ci-dessus, cette fonction est multilinéaire, alternée et normalisée. De plus, d'après la première proposition, il y a une seule telle fonction. Notamment, ça prouve que toutes les D_i sont égales, ce qui n'est pas évident.

La partie unicité du théorème est exactement le contenu de la première proposition. □

Donc, nous sommes rassurés : le déterminant existe. La preuve du théorème nous a même donné les formules de développement selon les lignes/colonnes pour le calculer.

PARTIE III

INTÉGRATION

CHAPITRE 7

INTÉGRATION

Avertissement : ce chapitre est accompagné en cours de nombreuses illustrations. Par manque de temps, elles ne sont pas (encore) reprises dans ce polycopié. L'étudiant doit se référer à ses notes de cours d'amphi pour les illustrations, presque indispensables à une bonne compréhension.

7.1. Primitives et intégrales

Étant donnée une fonction continue f sur un intervalle I, on cherche une primitive de f :

Définition 7.1.1. — Soit f une fonction continue sur un intervalle I. On appelle *primitive* de f toute fonction F dérivable sur I, de dérivée f.

On remarque que si F est une primitive et $C \in \mathbf{R}$, alors $F + C$ est encore une primitive. Ce sont les seules :

Proposition 7.1.2. — *Soit f une fonction continue sur I et F une primitive de f. Alors l'ensemble des primitives de f est l'ensemble des fonctions $F + C$ pour $C \in \mathbf{R}$.*

Démonstration. — On a déjà remarqué que toutes les fonctions $F + C$ sont des primitives. Montrons que ce sont les seules : soit G une primitive. Alors $G - F$ est dérivable sur I, de dérivée $G' - F' = f - f = 0$. D'après le théorème des accroissements finis, $G - F$ est constante. Notons C sa valeur. On a alors $G = F + C$. □

En observant un dessin, on peut se convaincre que si on arrive à définir précisément la notion d'*aire sous le graphe* de f, nous aurons notre primitive.

Cette opération est appelée *intégration*. Nous verrons plus tard dans ce chapitre que c'est possible : toute fonction continue admet une primitive. Nous commençons par admettre ce résultat et nous en tirons quelques propriétés de l'intégration qui permettent son calcul effectif pour de nombreuses fonctions usuelles :

Théorème 7.1.3. — *Soit f une fonction continue sur un intervalle I de* $\mathbf{R}$. *Alors :*

(QC)

- *Il existe une primitive F de f définie sur I.*
- *Soit a et b dans I. Alors la valeur de $F(b)-F(a)$ ne dépend pas du choix de la primitive F de f. On note ce nombre $\int_a^b f(t)dt$ et on l'appelle* intégrale de f entre a et b.

Remarques 7.1.4. —

- Le nom de la variable importe peu dans la notation de l'intégrale : $\int_a^b f(x)dx = \int_a^b f(t)dt$. Cette notation est en fait utilisée pour désigner, dans le cas où f dépend de plusieurs variables ou paramètres, quelle est la variable d'intégration.
- Le deuxième point découle de la proposition précédente : en effet, si on choisit une autre primitive G, alors il existe un réel c tel que $G = F + c$. On a donc :
$$G(b) - G(a) = (F(b) + c) - (F(a) + c) = F(b) - F(a).$$
- On obtient immédiatement $\int_a^b f(t)dt = -\int_b^a f(t)dt$ et $\int_a^a f(t)dt = 0$.
- Fixons $a \in I$ et observons la fonction $x \mapsto \int_a^x f(t)dt$ définie sur I. Si F est une primitive de f, on peut la réécrire $x \mapsto F(x) - F(a)$. C'est donc la primitive de f qui s'annule en a. On utilisera donc la notation $\int f(x)dx$ (sans spécifier les bornes) pour désigner une primitive.

Quand nous prouverons ce théorème, nous l'étendrons en fait à une classe plus large de fonctions, dites *fonctions intégrables* qui contient notamment les fonctions continues par morceaux.

Exemple 7.1.5. — Considérons la fonction f définie sur $\mathbf{R}$ par $f(x) = x+1$. Une primitive de f est la fonction F définie par $F(x) = \frac{x^2}{2} + x + 1$ [1]. On calcule donc :
$$\int_0^4 (x+1)dx = F(4) - F(0) = 13 - 1 = 12.$$

1. Une autre primitive est définie par la formule $\frac{x^2}{2} + x$!

Sur cet exemple, il n'y a pas de problème pour définir l'aire sous le graphe de f : c'est l'aire d'un trapèze. On observe bien que cette aire vaut 12.

7.2. Propriétés de l'intégrale des fonctions continues

Du théorème précédent découlent des propriétés de l'intégrale qui nous permettent de la calculer pour de nombreuses fonction usuelles. On fixe f et g deux fonctions continues sur I, a, b et c trois éléments de I et λ un réel.

Proposition 7.2.1 **(Relation de Chasles)**. —

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt. \tag{QC}$$

Démonstration. — Soit en effet F une primitive de f (elle existe d'après le théorème). Alors on a, par définition de la notation intégrale :

$$\int_a^b f(t)dt = F(b)-F(a) = (F(b)-F(c))+(F(c)-F(a)) = \int_a^c f(x)dx+\int_c^b f(x)dx.$$

□

Proposition 7.2.2 **(Linéarité)**. — *On a :*

$$\begin{aligned} &- \int_a^b (f+g)(t)dt = \int_a^b f(t)dt + \int_a^b g(t)dt. \\ &- \int_a^b (\lambda f)(t)dt = \lambda \int_a^b f(t)dt. \end{aligned} \tag{QC}$$

Démonstration. — En effet, si F est une primitive de f et G une primitive de G, alors $F+G$ est une primitive de $f+g$ et λF est une primitive de λf. On a alors :

$$\begin{aligned} &- \int_a^b (f+g)(t)dt = (F+G)(b)-(F+G)(a) = F(b)-F(a)+G(b)-G(a) = \\ &\quad \int_a^b f(t)dt + \int_a^b g(t)dt. \\ &- \int_a^b (\lambda f)(t)dt = (\lambda F)(b) - (\lambda F)(a) = \lambda(F(b)-F(a)) = \lambda \int_a^b f(t)dt. \end{aligned}$$

□

Proposition 7.2.3 **(Positivité)**. — *On suppose ici $a \leq b$.*

1. *Si $f \geq 0$ sur $[a,b]$, alors* $\displaystyle\int_a^b f(t)dt \geq 0$.

(QC)

2. *De plus, si $a < b$ et s'il existe un élément c de $[a, b]$ tel que $f(c) > 0$, alors* $\int_a^b f(t)dt > 0$.

Démonstration. — Soit F une primitive de f. Sa dérivée est positive, donc (par le théorème des accroissements finis) F est croissante. Comme $b > a$, on a donc $F(b) \geq F(a)$. L'intégrale est donc positive.

Si de plus $f(c) > 0$ pour un c dans $[a, b]$, alors par continuité de f il existe un intervalle $[d, e] \subset [a, b]$ contenant c sur lequel f est strictement positive. Remarquons qu'on a $a \leq d < e \leq b$. F est strictement croissante sur $[d, e]$ et $F(e) > F(d)$. On a donc (en utilisant la relation de Chasles) :

$$\int_a^b f(t)dt = \int_a^d f(t)dt + \int_d^e f(t)dt + \int_e^b f(t)dt.$$

Le premier et le dernier terme de la somme de droite sont ≥ 0 d'après le premier point de la proposition. On a déjà vu que le deuxième est > 0. Donc la somme est strictement positive. □

On en déduit :

Corollaire 7.2.4. — *On a pour $a \leq b$:*

(QC)

1. *si $f \leq g$,* $\int_a^b f(t)dt \leq \int_a^b g(t)dt$.
2. *de plus,* $\left|\int_a^b f(t)dt\right| \leq \int_a^b |f(t)|dt$.

Démonstration. —

1. On a $g - f \geq 0$, donc $\int_a^b (g - f)(t)dt \geq 0$. Or, en utilisant la linéarité, on a $\int_a^b (g - f)(t)dt = \int_a^b g(t)dt - \int_a^b f(t)dt$. On obtient bien l'inégalité voulue.
2. On part de l'inégalité $-|f| \leq f \leq |f|$. En utilisant le point précédent, on a :
$$\int_a^b (-|f|)(t)dt \leq \int_a^b f(t)dt \leq \int_a^b |f(t)|dt.$$
Or, par linéarité, le terme de gauche vaut $-\int_a^b |f(t)|dt$. Donc on peut écrire :
$$-\int_a^b |f(t)|dt \leq \int_a^b f(t)dt \leq \int_a^b |f(t)|dt.$$
On obtient bien l'inégalité voulue.

□

7.3. Calculs de primitives et d'intégrales

Pour calculer une intégrale, l'idée est la suivante : on connaît une primitive pour beaucoup de fonctions usuelles. Par linéarité, certaines fonctions se ramènent à des fonctions usuelles : par exemple, un polynôme est une combinaison linéaire de monômes pour lesquels on connaît une primitive. En plus, on ajoutera deux outils pour se ramener à des fonctions usuelles : l'intégration par parties et le changement de variables.

7.3.1. Quelques primitives. — Le point de départ à tout calcul est de connaître une liste de primitives des fonctions usuelles.

Proposition 7.3.1 **(Les primitives des fonctions usuelles)**
On a (C est une constante réelle) :

1. *Si* $\alpha \neq -1$, *alors* $\displaystyle\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C.$
2. $\displaystyle\int \frac{dx}{x} = \ln(|x|) + C.$ (QC)
3. $\displaystyle\int e^x dx = e^x + C.$
4. $\displaystyle\int \cos(x)dx = \sin(x) + C$ *et* $\displaystyle\int \sin(x)dx = -\cos(x) + C.$
5. $\displaystyle\int \frac{dx}{1+x^2} = \arctan(x) + C.$

Démonstration. — Pour prouver cette proposition, il suffit de dériver la fonction à droite de l'égalité et de vérifier qu'on obtient bien la fonction à gauche. On laisse le lecteur vérifier. □

7.3.2. L'intégration par parties. — On introduit une nouvelle notation utile pour toute fonction f définie sur l'intervalle $[a, b]$:

$$[f(x)]_a^b = f(b) - f(a).$$

Considérons f et g deux fonctions dérivables sur un intervalle $[a, b]$. On rappelle la formule de la dérivée d'une fonction produit :

$$(fg)'(x) = f'g(x) + fg'(x).$$

On en déduit les deux versions de l'intégration par parties :

Proposition 7.3.2 **(Intégration par parties)**. — *Pour toutes fonctions dérivables f et g sur un intervalle* $[a, b]$, *on a :*

Primitive $\displaystyle\int (f'g)(x)dx = fg - \int (fg')(x)dx$; (QC)

Intégrale $\int_a^b (f'g)(x)dx = [(fg)(x)]_a^b - \int_a^b (fg')(x)dx$.

Démonstration. — D'après la formule de dérivation d'une fonction produit, une primitive de $f'g$ est aussi une primitive de $(fg)' - fg'$ car les deux fonctions sont égales. Autrement dit,

$$\int (f'g)(x)dx = \int ((fg)' - fg')(x)dx = \int (fg)'(x)dx - \int (fg')(x)dx.$$

La dernière égalité s'obtient par linéarité. On remarque de plus que, par définition, $\int (fg)'(x)dx$ est une primitive de $(fg)'$: on peut prendre la fonction fg !

La deuxième version découle de la première en faisant la différence des valeurs en b et en a. □

En pratique, le jeu est de reconnaître dans une fonction à intégrer qui sera f' et qui sera g, c'est-à-dire qui on va intégrer ou dériver. Le but étant évidemment de simplifier l'expression !

(QC)

Exemple 7.3.3. — Déterminons une primitive $\int \ln(x)dx$ sur $]0, +\infty[$: posons $f'(x) = 1$, $g(x) = \ln(x)$, et donc $f(x) = x$, $g'(x) = \frac{1}{x}$:

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

Exemple 7.3.4. — Calculons $\int_0^{\frac{\pi}{2}} x^2\cos(x)dx$. On pose dans un premier temps $g(x) = x^2$ (en dérivant, le degré baissera) et $f'(x) = \cos(x)$. On a alors $f(x) = \sin(x)$ et $g'(x) = 2x$. On obtient

$$\int_0^{\frac{\pi}{2}} x^2\cos(x)dx = \left[x^2\sin(x)\right]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} 2x\sin(x)dx.$$

On calcule $\left[x^2\sin(x)\right]_0^{\frac{\pi}{2}} = \frac{\pi^2}{4}$. Il reste à calculer

$$\int_0^{\frac{\pi}{2}} 2x\sin(x)dx = 2\int_0^{\frac{\pi}{2}} x\sin(x)dx.$$

On fait encore une intégration par parties (en veillant à ne pas défaire ce qu'on a fait !) : on pose $g(x) = x$ et $f'(x) = \sin(x)$, on a alors $g'(x) = 1$ et $f(x) = -\cos(x)$. Ainsi :

$$\int_0^{\frac{\pi}{2}} x\sin(x)dx = [-x\cos(x)]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} -\cos(x)dx = 1.$$

Ainsi, $\int_0^{\frac{\pi}{2}} x^2 \cos(x)dx = \frac{\pi^2}{4} - 2.$

7.3.3. Changement de variable. —

7.3.3.1. Par translation. — Tout part de la remarque suivante : si F est une primitive de f, alors $x \mapsto F(x + x_0)$ est une primitive de $x \mapsto f(x + x_0)$. Autrement dit :

Proposition 7.3.5. — *On a :*

$$\begin{aligned} \text{Primitive } & \int f(x+x_0)dx = F(x+x_0) + C. \\ \text{Intégrale } & \int_a^b f(x+x_0)dx = \int_{a+x_0}^{b+x_0} f(x)dx. \end{aligned} \tag{QC}$$

La deuxième propriété est par ailleurs claire en se rappelant que l'intégrale est l'aire sous le graphe : en changeant x en $x+x_0$, on ne fait qu'une translation horizontale au graphe (de $-x_0$), ce qui ne change pas l'aire en dessous.

On peut appliquer la proposition précédente à tous les primitives de bases :

Proposition 7.3.6. — *On a pour tout $x_0 \in \mathbf{R}$:*

1. *Si $\alpha \neq -1$, alors* $\int (x+x_0)^\alpha dx = \frac{(x+x_0)^{\alpha+1}}{\alpha+1} + C.$
2. $\int \frac{dx}{x+x_0} = \ln(|x+x_0|) + C.$
3. $\int \ln(x+x_0)dx = (x+x_0)\ln(|x+x_0|) - (x+x_0) + C.$
4. $\int e^{x+x_0}dx = e^{x+x_0} + C.$
5. $\int \cos(x+x_0)dx = \sin(x+x_0)+C$ *et* $\int \sin(x+x_0)dx = -\cos(x+x_0)+C.$
6. $\int \frac{dx}{1+(x+x_0)^2} = \arctan(x+x_0) + C.$

Exemple 7.3.7. — Calculons $\int_{-1}^{0} \frac{x^2}{(x-1)^3}dx$. Pour ça, posons $u = x-1$. Quand x varie entre -1 et 0, u varie entre -2 et -1. On a alors :

$$\begin{aligned}
\int_{-1}^{0} \frac{x^2}{(x-1)^3}dx &= \int_{-2}^{-1} \frac{(u+1)^2}{u^3}du = \int_{-2}^{-1} \frac{u^2+2u+1}{u^3}du \\
\text{(par linéarité)} \qquad &= \int_{-2}^{-1} \frac{u^2}{u^3}du + 2\int_{-2}^{-1} \frac{u}{u^3}du + \int_{-2}^{-1} \frac{1}{u^3}du \\
&= \int_{-2}^{-1} \frac{1}{u}du + 2\int_{-2}^{-1} \frac{1}{u^2}du + \int_{-2}^{-1} \frac{1}{u^3}du \\
&= [\ln|u|]_{-2}^{-1} + 2\left[-\frac{1}{u}\right]_{-2}^{-1} + \left[-\frac{1}{2u^2}\right]_{-2}^{-1} \\
&= (\ln(1) - \ln(2)) + 2(1 - \frac{1}{2}) + (-\frac{1}{2} + \frac{1}{8}) \\
&= -\ln(2) + 1 - \frac{3}{8} = -\ln(2) + \frac{5}{8}
\end{aligned}$$

7.3.3.2. Affine. — On veut faire un changement de variable un tout petit peu plus compliqué : transformer x en $px + x_0$ (où $p \neq 0$), c'est-à-dire un changement de variable affine. Pour ça on remarque que si F est une primitive de f, alors $x \mapsto \frac{1}{p}F(px + x_0)$ est une primitive de $x \mapsto f(px + x_0)$. On en déduit :

Proposition 7.3.8. — *On a :*

(QC)
Primitive $\int f(px + x_0)dx = \frac{1}{p}F(px + x_0) + C$.
Intégrale $\int_a^b f(px + x_0)dx = \frac{1}{p}\int_{pa+x_0}^{pb+x_0} f(x)dx$.

Démonstration. — Il peut être utile de montrer la seconde forme de la proposition : d'après le lien entre primitive et intégrale, on peut écrire :

$$\begin{aligned}
\int_a^b f(px + x_0)dx &= \left[\frac{1}{p}F(px + x_0)\right]_a^b = \frac{1}{p}F(pb + x_0) - \frac{1}{p}F(pa + x_0) \\
&= \frac{1}{p}\left[F(x)\right]_{pa+x_0}^{pb+x_0} = \frac{1}{p}\int_{pa+x_0}^{pb+x_0} f(x)dx.
\end{aligned}$$

□

Exemple 7.3.9. — Calculons $\int \frac{dx}{x^2 + 4x + 8}$. Pour ça, ramenons le dénominateur à une forme dite "normale" :

$$x^2 + 4x + 8 = 4\left(\left(\frac{x+2}{2}\right)^2 + 1\right).$$

Posons donc $u = \frac{x+2}{2}$. On a alors, en utilisant les notations physiciennes [2], $\frac{du}{dx} = \frac{1}{2}$, ou encore $du = \frac{dx}{2}$ ce qui s'écrit aussi $dx = 2du$. On a alors :

$$\begin{aligned}\int \frac{dx}{x^2+4x+8} &= \int \frac{2du}{4(u^2+1)} = \frac{1}{2}\int \frac{du}{u^2+1} \\ &= \frac{1}{2}\arctan(u) + C\end{aligned}$$

Il ne reste plus qu'à revenir à la variable x :

$$\int \frac{dx}{x^2+4x+8} = \frac{1}{2}\arctan\left(\frac{x+2}{2}\right) + C.$$

7.3.3.3. Cas général. — Le cas général est donné par la considération de la dérivation d'une fonction composée :

$$(f \circ g)'(x) = g'(x)f'(g(x)).$$

On en déduit :

Proposition 7.3.10. — *On a :*

Primitive $\int f'(g(x))g'(x)dx = f \circ g(x) + C.$ (QC)

Intégrale *Si f est continue et g est C^1, on a*

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

Attention à la deuxième forme : si g est décroissante, les bornes après changement de variable ne sont pas dans le bon ordre.

En particulier, on peut appliquer la première forme aux primitives de base :

Proposition 7.3.11. — *On a pour toute fonction f dérivable :*

1. *Si $\alpha \neq -1$, alors* $\int f'(x)f^\alpha(x)dx = \frac{f^{\alpha+1}(x)}{\alpha+1} + C.$
2. $\int \frac{f'(x)dx}{f(x)} = \ln(|f(x)|) + C.$
3. $\int e^{f(x)}f'(x)dx = e^{f(x)} + C.$

Exemple 7.3.12. — Déterminons une primitive de $\tan(x)$ sur un intervalle de définition. On remarque que $\tan(x) = \frac{\sin(x)}{\cos(x)}$ et que $\sin(x)$ est (au signe

2. Il est possible de justifier formellement l'usage de ces notations. Dans ce cours, nous les utiliserons surtout comme un moyen efficace de procéder au changement de variable dans les intégrales. C'est d'ailleurs à cause de ça aussi que la notation dx apparaît dans l'intégrale.

près) la dérivée de cos en x. On obtient avec le deuxième cas de la proposition ci-dessus :

$$\int \tan(x)dx = -\ln(|\cos(x)|) + C.$$

Traitons un autre exemple, où la fonction $g'(x)$ n'apparaît pas clairement à gauche. Dans ce cas, on fait le changement de variable désiré, et on compense l'absence de ce terme. La notation physicienne est alors très efficace :

Exemple 7.3.13. — Calculons $\int_0^4 \frac{dx}{1+\sqrt{x}}$. Pour ça, le dénominateur semble pénible, donc on est tenté de poser $u = 1 + \sqrt{x}$. On a alors $x = (u-1)^2$ et $dx = 2(u-1)du$. Quand x varie entre 0 et 4, u varie entre 1 et 3. On écrit donc [3] :

$$\int_0^4 \frac{dx}{1+\sqrt{x}} = \int_1^3 \frac{2u-2}{u}du = 2\int_1^3 du - 2\int_1^3 \frac{1}{u}du = 4 - 2\ln(3)$$

A titre d'exercice, le lecteur pourra montrer $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin(x) + C$ en effectuant le changement de variable $x = \sin(u)$.

7.4. Intégration des fonctions rationnelles

Une fraction rationnelle est un quotient de polynômes. À une fraction rationnelle est associée une fonction (qui consiste à évaluer le quotient des deux fonctions polynômes), appelée *fonction rationnelle*, définie partout sauf en les racines du dénominateur. Il est bon de savoir qu'on peut intégrer toutes les fonctions rationnelles.

---

3. Essayons de nous convaincre que le calcul que nous faisons relève du changement de variables. Pour ça décortiquons-le. On part d'une intégrale de la forme $\int_0^4 f(x)dx$. On écrit $u = \varphi(x)$ (ici $\varphi(x) = 1 + \sqrt{x}$) et $f(x) = g(u)$. Ici φ est une bijection, donc on l'inverse : $x = \varphi^{-1}(u)$. Quand on calcule dx en fonction de u et du, on écrit en réalité

$$\frac{dx}{du} = \left(\varphi^{-1}\right)'(u).$$

De plus, on a par définition $g(u) = f(\varphi^{-1}(u))$

Quand on remplace dx par sa valeur $\left(\varphi^{-1}\right)'(u)du$ et $f(x)$ par $g(u) = f(\varphi^{-1}(u))$, on écrit l'égalité :

$$\int_0^4 f(x)dx = \int_{\varphi(0)}^{\varphi(4)} \left(\varphi^{-1}\right)'(u) f(\varphi^{-1}(u))du.$$

Mais si on lit l'égalité de droite à gauche, on voit la formule de changement de variable, pour $x = \varphi^{-1}(u)$. Ce calcul est donc possible tant que φ^{-1} est dérivable, c'est-à-dire tant qu'on peut calculer dx en fonction de u et du.

Nous allons dans cette section présenter rapidement cette technique d'intégration. Commençons par deux exemples.

7.4.1. Exemples. —

Exemple 7.4.1. — Commençons par déterminer une primitive sur $]1,+\infty[$ de la fonction $f(x)=\frac{2x^2-1}{x^3-x^2}$.

On peut manipuler l'expression de $f(x)$ pour arriver à l'égalité $f(x)=\frac{1}{x}+\frac{1}{x^2}+\frac{1}{x-1}$. Or, maintenant, f est exprimée comme une somme de fonctions dont on connaît une primitive. Une primitive de f sur $[1,+\infty[$ est donc :

$$F(x)=\ln(x)-\frac{1}{x}+\ln(x-1).$$

Exemple 7.4.2. — Deuxième exemple : déterminons $\displaystyle\int_1^2 \frac{1}{t(2t^2+t+1)}dt$. À nouveau, une première étape est de manipuler l'expression de la fonction. On peut vérifier l'égalité :

$$\frac{1}{t(2t^2+t+1)}=\frac{1}{t}-\frac{2t+1}{2t^2+t+1}.$$

Remarquons que le dénominateur $2t^2+t+1$ n'a pas de racines réelles. Nous verrons plus tard que c'est pour cette raison que nous ne cherchons pas à simplifier encore la formule.

On sait déterminer l'intégrale du premier des deux termes à droite. On a en effet $\displaystyle\int_1^2 \frac{1}{t}dt=\ln(2)$.

Pour le second terme, on essaye d'abord de se ramener à une expression du type $\frac{u'}{u}$. La dérivée du dénominateur est $4t+1$. On écrit alors le numérateur sous la forme $2t+1=\frac{1}{2}(4t+1)+\frac{1}{2}$. On a donc

$$\begin{aligned}\int_1^2 \frac{2t+1}{2t^2+t+1}dt &= \frac{1}{2}\int_1^2 \frac{4t+1}{2t^2+t+1}dt+\frac{1}{2}\int_1^2 \frac{1}{2t^2+t+1}dt\\ &= \frac{1}{2}\Big[\ln(2t^2+t+1)\Big]_1^2+\frac{1}{2}\int_1^2 \frac{1}{2t^2+t+1}dt\\ &= \frac{1}{2}\left(\ln(11)-\ln(4)\right)+\frac{1}{2}\int_1^2 \frac{1}{2t^2+t+1}dt.\end{aligned}$$

Il reste à calculer l'intégrale de $\frac{1}{2t^2+t+1}$. Ce calcul est similaire à celui de l'exemple 7.3.9 : on essaye de se ramener à la fonction $\frac{1}{u^2+1}$ qu'on sait intégrer avec l'arctangente. On écrit le polynôme du second degré sous la forme : $2t^2+$

$t+1=\frac{7}{8}\left(\left(\frac{4x+1}{\sqrt{7}}\right)^2+1\right)$. En utilisant le changement de variable $u=\frac{4x+1}{\sqrt{7}}$,

$$\int_1^2 \frac{1}{2t^2+t+1}dt=\frac{2}{\sqrt{7}}\left[\text{Arctan}\left(\frac{4x+1}{\sqrt{7}}\right)\right]_1^2.$$

Finalement nous obtenons :

$$\begin{aligned}\int_1^2 \frac{1}{t(2t^2+t+1)}dt &= \ln(2)-\frac{\ln\left(\frac{11}{4}\right)}{2}-\frac{1}{\sqrt{7}}\left(\text{Arctan}\left(\frac{9}{\sqrt{7}}\right)-\text{Arctan}\left(\frac{5}{\sqrt{7}}\right)\right)\\ &= \frac{1}{2}\ln\left(\frac{16}{11}\right)-\frac{1}{\sqrt{7}}\left(\text{Arctan}\left(\frac{9}{\sqrt{7}}\right)-\text{Arctan}\left(\frac{5}{\sqrt{7}}\right)\right)\end{aligned}$$

7.4.2. Décomposition en éléments simples. — Il existe une technique algorithmique, appelée *décomposition en éléments simples*, pour procéder à la première étape des deux exemples précédents : décomposer une fonction rationnelle comme somme de fonctions dont on sait calculer une primitive.

Un *élément simple* est une fonction rationnelle d'une des formes suivantes :

- $\frac{a}{x-b}$ où a et b sont des réels.
- $\frac{a}{(x-b)^n}$ où a et b sont des réels et $n\geq 2$ un entier.
- $\frac{x+a}{x^2+bx+c}$ où a, b, c sont des réels et le polynôme x^2+bx+c n'a pas de racines réelles.
- $\frac{x+a}{(x^2+bx+c)^n}$ où $n\geq 2$ est un entier, a, b, c sont des réels et le polynôme x^2+bx+c n'a pas de racines réelles.

Le théorème de décomposition en éléments simples [4] énonce que toute fonction rationnelle f peut s'écrire, de manière unique et algorithmique, comme somme d'un polynôme et d'éléments simples. De plus, il stipule que les éléments simples qui apparaissent sont à chercher parmi ceux dont le dénominateur divise le dénominateur de f (ce qui ne laisse qu'un nombre fini de possibilités). Enfin le polynôme est le reste de la division euclidienne du dénominateur par le numérateur. Quand le degré du numérateur est inférieur au degré du dénominateur, ce terme polynomial n'apparaît pas.

Exemple 7.4.3. — 1. Considérons la fraction $\frac{1}{x(x^2+1)}$. Le théorème de décomposition en éléments simples énonce qu'il existe un réel a et deux réels b et c tels que :

$$\frac{1}{x(x^2+1)}=\frac{a}{x}+\frac{bx+c}{x^2+1}.$$

4. On renvoie le lecteur intéressé à la page de les-mathématiques.net pour plus d'informations sur ce théorème.

On peut déterminer a, b et c en réduisant au même dénominateur :

$$\frac{1}{x(x^2+1)} = \frac{a(x^2+1)+x(bx+c)}{x(x^2+1)} = \frac{(a+b)x^2+cx+a}{x(x^2+1)}.$$

On obtient donc le système $a=1$, $c=0$, $a+b=0$. Finalement, la décomposition en éléments simples de $\frac{1}{x(x^2+1)}$ est :

$$\frac{1}{x(x^2+1)} = \frac{1}{x} - \frac{x}{x^2+1}.$$

Une primitive est donc $\ln(|x|) - \frac{1}{2}\ln(x^2+1)$.

2. Considérons la fraction $\frac{x}{(x+1)^2(x-1)}$. Le théorème dit qu'il existe trois nombres réels a, b, c tels que :

$$\frac{x}{(x+1)^2(x-1)} = \frac{a}{x+1} + \frac{b}{(x+1)^2} + \frac{c}{x-1}.$$

À nouveau, on met au même dénominateur pour obtenir :

$$\frac{x}{(x+1)^2(x-1)} = \frac{a(x^2-1)+b(x-1)+c(x^2+2x+1)}{(x+1)^2(x-1)} = \frac{(a+c)x^2+(b+2c)x+(c-b-a)}{(x+1)^2(x-1)}$$

On obtient encore une fois un système : $a+c=0$, $b+2c=1$, $c-b-a=0$. L'unique solution est $a=-\frac{1}{4}$, $b=\frac{1}{2}$ et $c=\frac{1}{4}$. La décomposition en éléments simples de $\frac{x}{(x+1)^2(x-1)}$ est :

$$\frac{x}{(x+1)^2(x-1)} = \frac{-1}{4(x+1)} + \frac{1}{2(x+1)^2} + \frac{1}{4(x-1)}.$$

Une primitive est donc $\frac{1}{4}\left(\ln(|x-1|) - \ln(|x+1|) - 2\frac{1}{x+1}\right)$.

À ce stade, le lecteur devrait être convaincu que pour les trois premiers types d'éléments simples on sait calculer une primitive – pour le troisième grâce à l'arctangente et un changement de variable affine. Le cas du quatrième élément simple est moins clair. Cependant, on peut trouver une formule par récurrence (sur n) pour calculer ce type de primitive, à l'aide d'une intégration par parties judicieuse. Nous ne traitons pas ici ce calcul.

CHAPITRE 8

ÉTUDE THÉORIQUE DE L'INTÉGRALE

Le but de ce chapitre est de revenir sur la définition même de l'intégrale, laissée en suspens dans le chapitre précédent. Nous étudierons aussi quelques propriétés théoriques de l'intégrale.

8.1. Intégration des fonctions en escalier

On passe par les fonctions en escalier pour définir l'intégrale des fonctions continues. C'est en grande mesure un détour technique.

8.1.1. Subdivision. — Soit $a < b$ deux réels.

Définition 8.1.1. — Une subdivision de l'intervalle $[a, b]$ est un ensemble fini $\sigma = \{x_0 = a < x_1 < x_2 < \ldots < x_n = b\}$.

Une subdivision τ qui contient (comme ensemble) une subdivision σ est dite *plus fine* que σ.

Si σ et σ' sont deux subdivisions, on note $\sigma \vee \sigma'$ la subdivision qui est l'union de σ et σ'.

Le *pas* d'une subdivision $\sigma = \{x_0 < x_1 < \ldots < x_n\}$, noté $|\sigma|$, est le $\max\{x_i - x_{i-1}, \text{ pour } 1 \leq i \leq n\}$.

Une subdivision est dite *régulière* si pour tout $1 \leq i \leq n$, on a $x_i - x_{i-1} = |\sigma|$.

Exemple 8.1.2. — L'ensemble $\{0 = \frac{0}{n} < \frac{1}{n} < \ldots < \frac{n-1}{n} < \frac{n}{n} = 1\}$ est une subdivision régulière de $[0, 1]$.

Remarques 8.1.3. — – Si σ et σ' sont deux subdivisions, alors $\sigma \vee \sigma'$ est une subdivision plus fine que σ et que σ'.

– Si $\sigma = \{x_0 < x_1 < \ldots < x_n\}$ est une subdivision régulière de $[a, b]$, alors on a :

$$b - a = \sum_{i=1}^{n} x_i - x_{i-1} = \sum_{i=1}^{n} |\sigma| = n|\sigma|.$$

Donc $|\sigma| = \frac{b-a}{n}$ et $x_i = a + \sum_{k=1}^{i} x_i - x_{i-1} = a + i\frac{b-a}{n}$.

8.1.2. Fonctions en escalier. —

(QC)

Définition 8.1.4. — Une fonction $f : [a, b] \to \mathbf{R}$ est dite *en escalier* si il existe une subdivision $\sigma = \{x_0 < x_1 < \ldots < x_n\}$ telle que pour tout $1 \leq i \leq n$, f est constante sur l'intervalle $]x_{i-1}, x_i[$.

Attention, la définition ne dit rien sur le comportement de f en les points x_i de la subdivision.

Définition 8.1.5. — Si f est en escalier et σ est une subdivision de $[a, b]$ vérifiant la propriété de la définition précédente, on dit que σ est adaptée à f.

Attention il n'y a pas une unique subdivision adaptée à une fonction en escalier donnée. Notamment, si τ est plus fine qu'une subdivision adaptée, alors τ est encore adaptée.

Proposition 8.1.6. — *Soient f et g deux fonctions de l'espace* $\mathrm{Esc}(a, b)$ *des fonctions en escalier sur $[a, b]$ et λ un nombre réel. Alors $f + g$ et λf sont dans* $\mathrm{Esc}(a, b)$.

Autrement dit, l'espace des fonctions en escalier est un espace vectoriel.

Démonstration. — Soient f et g deux fonctions en escalier et $\lambda \in \mathbf{R}$.

Soient σ une subdivision adaptée à f et σ' une subdivision adaptée à g. Considérons la subdivision $\tau = \sigma \vee \sigma'$. Montrons que τ est adaptée à $f + g$. Notons $\tau = \{a = x_0 < x_1 < \ldots < x_n = b\}$.

Soit $1 \leq i \leq n$. Montrons que $f+g$ est constante sur $]x_{i-1}, x_i[$. Effectivement, τ est adaptée à f (car elle est plus fine que σ) et à g (car elle est plus fine que σ'). Donc sur cet intervalle, f et g sont constantes, donc $f + g$ aussi.

Ainsi τ est adaptée à $f + g$, donc cette dernière fonction est en escalier.

De plus, σ est adaptée à λf, donc cette dernière fonction est en escalier. □

8.1.3. Intégration. — Soit f une fonction en escalier et $\sigma = \{a = x_0 < x_1 < \ldots < x_n = b\}$ une subdivision adaptée. Notons y_i la valeur de f sur l'intervalle $]x_{i-1}, x_i[$. On note alors $I(f, \sigma)$ le nombre :

$$I(f, \sigma) = \sum_{i=1}^{n} (x_i - x_{i-1}) y_i.$$

Proposition 8.1.7. — *Si σ et σ' sont adaptées à f, alors $I(f, \sigma) = I(f, \sigma')$.*

Démonstration. — On montre d'abord le lemme :

Lemme 8.1.8. — *Si τ est plus fine que σ, alors $I(f, \tau) = I(f, \sigma)$.*

Démonstration. — Notons $\sigma = \{a = x_0 < x_1 < \ldots < x_n = b\}$ et $\tau = \{a = x'_0 < x'_1 < \ldots < x'_N = b\}$. Chaque élément de σ est aussi un élément de τ, donc notons k_i l'entier tel que $x_i = x'_{k_i}$. On a $k_0 = 0$ et $k_n = N$.

Pour tout $1 \leq i \leq n$, on a

$$x_i - x_{i-1} = \sum_{j=k_{i-1}+1}^{k_i} x'_j - x'_{j-1}.$$

De plus, soit y_i la valeur de f sur l'intervalle $]x_{i-1}, x_i[$. Alors, pour tout $k_{i-1} + 1 \leq j \leq k_i$, le nombre y_i est la valeur de la fonction f sur l'intervalle $]x'_j, x'_{j+1}[$. Donc on a :

$$\begin{aligned} I(f, \sigma) &= \sum_{i=1}^{n} (x_i - x_{i-1}) y_i \\ &= \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} (x'_j - x'_{j-1}) y_i \\ &= \sum_{j=1}^{N} (x'j - x'_{j-1}) y_i \\ &= I(f, \tau). \end{aligned}$$

□

Concluons la preuve de la proposition : si σ et σ' sont deux décompositions adaptées, alors on peut considérer la décomposition $\sigma \vee \sigma'$. Elle est plus fine que σ et σ', donc d'après le lemme, on a :

$$I(f, \sigma) = I(f, \sigma \vee \sigma') = I(f, \sigma').$$

□

Donc pour toutes les décompositions adaptées, la valeur de $I(f,\sigma)$ est la même. On note ce nombre :

$$\int_a^b f(t)dt.$$

8.1.4. Propriétés de l'intégrale. — L'intégrale que nous venons de définir vérifie un certain nombre de propriétés :

***Proposition 8.1.9* (Relation de Chasles).** — *Si $f \in \mathrm{Esc}(a,b)$ et $a < c < b$, alors on a :*

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Démonstration. — Soit σ une subdivision adaptée à f et $\tau = \sigma \cup \{c\}$. Alors τ est plus fine que σ, donc est adaptée à f. Utilisons τ pour calculer l'intégrale de f. Pour ça notons $\tau = \{a = x_0 < \ldots < x_k = c < \ldots < x_n = b\}$. On a (avec les notations habituelles) :

$$\begin{aligned}\int_a^b f(t)dt = I(f,\tau) &= \sum_{i=1}^{n}(x_i - x_{i-1})y_i \\ &= \sum_{i=1}^{k}(x_i - x_{i-1})y_i + \sum_{i=k+1}^{n}(x_i - x_{i-1})y_i\end{aligned}$$

Or $\{a = x_0 < x_1 < \ldots < x_k = c\}$ est une subdivision de $[a,c]$ qui est adaptée à f. De même, $\{c = x_k < x_{k+1} < \ldots < x_n = b\}$ est une subdivision de $[c,b]$ qui est adaptée à f. On en déduit que les deux derniers termes sont respectivement $\int_a^c f(t)dt$ et $\int_c^b f(t)dt$.

On a bien montré l'égalité voulue. □

***Proposition 8.1.10* (Linéarité).** — *L'application $\mathrm{Esc}(a,b) \to \mathbf{R}$ qui associe à f son intégrale $\int_a^b f(t)dt$ est linéaire :*

$$\int_a^b (f+g)(t)dt = \int_a^b f(t)dt + \int_a^b g(t)dt$$

$$\text{et } \int_a^b \lambda f(t)dt = \lambda \int_a^b f(t)dt$$

pour toutes fonctions en escalier f et g et pour tout réel λ.

Démonstration. — Soient f, g deux fonctions en escalier et $\lambda \in \mathbf{R}$. On veut montrer :

$$\int_a^b (f+g)(t)dt = \int_a^b f(t)dt + \int_a^b g(t)dt.$$

Prenons, comme dans la preuve de la proposition 8.1.6, une subdivision $\sigma = \{x_0 < \ldots < x_n\}$ de $[a, b]$ adaptée à la fois à f, à g et à $f + \lambda g$. Alors, sur chaque intervalle $]x_{i-1}, x_i[$, f est constante de valeur y_i, g est constante de valeur z_i et $f + g$ est constante de valeur $y_i + z_i$. On a donc :

$$\begin{aligned}\int_a^b (f+g)(t)dt = I(f+g,\sigma) &= \sum_{i=1}^{n}(x_i - x_{i-1})(y_i + z_i)\\ &= \sum_{i=1}^{n}(x_i - x_{i-1})y_i + \sum_{i=1}^{n}(x_i - x_{i-1})z_i\\ &= I(f,\sigma) + I(g,\sigma)\\ &= \int_a^b f(t)dt + \int_a^b g(t)dt.\end{aligned}$$

De plus, $\displaystyle\int_a^b \lambda f(t)dt = I(\lambda f,\sigma) = \sum_{i=1}^{n}(x_i - x_{i-1})y_i\lambda = \lambda\left(\sum_{i=1}^{n}(x_i - x_{i-1})y_i\right) = \lambda I(f,\sigma) = \lambda\int_a^b f(t)dt.$ □

***Proposition 8.1.11* (Positivité).** — *Si $f \in \mathrm{Esc}(a, b)$ est positive, alors on a $\displaystyle\int_a^b f(t)dt \geq 0$.*

Cette proposition est claire : dans la formule pour définir $I(f, \sigma)$ tous les termes sont positifs.

Comme pour les fonctions continues, on en tire le corollaire :

Corollaire 8.1.12. — *Soient f et g deux fonctions en escaliers sur $[a, b]$. On a :*

1. *Si $f \leq g$, alors $\int_a^b f(t)dt \leq \int_a^b g(t)dt$;*
2. *$\left|\int_a^b f(t)dt\right| \leq \int_a^b |f|(t)dt$.*

8.2. Intégration des fonctions continues par morceaux

8.2.1. Intégrales supérieure et inférieure. — Soit $f : [a,b] \to \mathbf{R}$ une fonction bornée. On note :

$$\begin{aligned} I_+(f,a,b) &= \inf\{\int_a^b \varphi(t)dt \text{ pour } \varphi \in \mathrm{Esc}(a,b),\ \varphi \geq f\} \\ &\quad \text{(l'intégrale supérieure de } f) \\ \text{et } I_-(f,a,b) &= \sup\{\int_a^b \varphi(t)dt \text{ pour } \varphi \in \mathrm{Esc}(a,b),\ \varphi \leq f\} \\ &\quad \text{(l'intégrale inférieure de } f) \end{aligned}$$

Proposition 8.2.1. — *Définissons* $m = \inf\{f(x)$ *pour* $x \in [a,b]\}$ *et* $M = \sup\{f(x)$ *pour* $x \in [a,b]\}$. *Alors,* $I_-(f,a,b)$ *et* $I_+(f,a,b)$ *sont des réels et on a :*

$$(b-a)m \leq I_-(f,a,b) \leq I_+(f,a,b) \leq (b-a)M.$$

Démonstration. — On remarque que φ_m définie par $\varphi_m(x) = m$ est une fonction en escalier inférieure à f. De même, ψ_M définie par $\psi_M(x) = M$ est une fonction en escalier supérieure à f.

Soit maintenant φ et ψ deux fonctions en escalier, avec $\varphi \leq f$ et $\psi \geq f$. Comme $\varphi \leq \psi$, on a vu que $\int_a^b \varphi(t)dt \leq \int_a^b \psi(t)dt$.

En appliquant l'inégalité précédente avec $\varphi = \varphi_m$, on obtient que pour tout ψ en escalier supérieure à f, $(b-a)m \leq \int_a^b \psi(t)dt$. Donc, l'ensemble dont $I_+(f,a,b)$ est l'inf est minoré, donc $I_+(f,a,b)$ est bien défini. De même $I_-(f,a,b)$ est bien défini.

Comme $\varphi_m \leq f$, on obtient $I_-(f,a,b) \geq \int_a^b \varphi_m(t)dt = (b-a)m$. De même, $I_+(f,a,b) \leq \int_a^b \psi_M(t)dt = (b-a)M$. Il reste à comparer $I_-(f,a,b)$ et $I_+(f,a,b)$: reprenons notre inégalité

$$\int_a^b \varphi(t)dt \leq \int_a^b \psi(t)dt.$$

En passant au sup sur les fonctions φ en escalier inférieure à f, on obtient $I_-(f,a,b) \leq \int_a^b \psi(t)dt$. En passant maintenant à l'inf sur les fonctions ψ en escalier supérieure à f, on conclut :

$$I_-(f,a,b) \leq I_+(f,a,b).$$

□

Il nous sera utile d'avoir une relation de Chasles pour I_+ et I_- :

Proposition 8.2.2. — *Si $a < c < b$, on a :*

$$\begin{aligned} I_+(f,a,b) &= I_+(f,a,c) + I_+(f,c,b) \\ I_-(f,a,b) &= I_-(f,a,c) + I_-(f,c,b) \end{aligned}$$

Démonstration. — Montrons la première des deux égalités, l'autre est similaire. On montre une double inégalité :

Pour $\varphi_1 \in \mathrm{Esc}(a,c)$ et $\varphi_2 \in \mathrm{Esc}(c,b)$ supérieures à f, on note φ la fonction définie sur $[a,b]$ telle que $\varphi(x) = \varphi_1(x)$ si $x \in [a,c]$ et $\varphi(x) = \varphi_2(x)$ si $x \in]c,b]$. Alors φ est en escalier sur $[a,b]$ et est supérieure à f. De plus, d'après la relation de Chasles pour les fonction en escalier, on a :

$$\int_a^c \varphi_1(t)dt + \int_c^b \varphi_2(t)dt = \int_a^b \varphi(t)dt.$$

En prenant l'inf sur φ_1 et φ_2, on obtient

$$I_+(f,a,c) + I_+(f,c,b) \leq I_+(f,a,b).$$

Pour l'autre inégalité, si φ est en escalier sur $[a,b]$ et supérieure à f, on note φ_1 sa restriction à $[a,c]$ et φ_2 sa restriction à $[c,b]$. Alors $\varphi_1 \in \mathrm{Esc}(a,c)$ et $\varphi_2 \in \mathrm{Esc}(c,b)$ sont des fonctions supérieures à f. On a encore

$$\int_a^c \varphi_1(t)dt + \int_c^b \varphi_2(t)dt = \int_a^b \varphi(t)dt.$$

En prenant cette fois l'inf sur φ, on obtient

$$I_+(f,a,c) + I_+(f,c,b) \geq I_+(f,a,b).$$

L'égalité est bien démontrée. □

8.2.2. Fonctions intégrables. —

Définition 8.2.3. — Une fonction f bornée de $[a,b]$ dans $\mathbf{R}$ est dite *intégrable* sur $[a,b]$ si $I_+(f,a,b) = I_-(f,a,b)$.

Dans ce cas, on note $\displaystyle\int_a^b f(t)dt$ la valeur commune de $I_+(f,a,b) = I_-(f,a,b)$.

Remarque 8.2.4. — Remarquons que, par définition des bornes inférieures et supérieurs, f est intégrable si et seulement si pour tout $\varepsilon > 0$, il existe φ et $\psi \in \mathrm{Esc}(a,b)$ avec $\varphi \leq f \leq \psi$ et

$$\int_a^b \psi(t)dt - \int_a^b \varphi(t)dt \leq \varepsilon.$$

On peut en déduire que si f est intégrable sur $[a,b]$, elle est intégrable sur tout intervalle $[c,d] \subset [a,b]$.

Le résultat le plus important concerne les fonctions continues par morceaux :

Définition 8.2.5. — Une fonction $f : [a, b] \to \mathbf{R}$ est dite *continue par morceaux* si il existe une subdivision $a = x_0 < x_1 < \ldots < x_n = b$ telle que f est
(QC) continue sur tous les intervalles $]x_{i-1}, x_i[$ et admet des limites à gauche et à droite en les x_i.

Une autre façon de l'exprimer est de dire que la restriction de f à $]x_{i-1}, x_i[$ se prolonge par continuité au segment $[x_{i-1}, x_i]$. On en déduit deux propriétés des fonctions continues par morceaux :

- Elles sont bornées (sur chaque $]x_{i-1}, x_i[$, f l'est car elle se prolonge continuement au segment. Il y a un nombre fini de tels intervalles, et il reste un nombre fini de points : les x_i).
- Elles admettent en tout point une limite à gauche et une limite à droite.

Théorème 8.2.6. — *Soit f une fonction continue par morceaux sur $[a, b]$. Alors, f est intégrable sur $[a, b]$ et même, pour tout $a \leq x \leq b$, $I_+(f, a, x) =$*
(QC) *$I_-(f, a, x)$.*

De plus, si f est continue, la fonction $x \mapsto I_+(f, a, x)$ est une primitive de f.

On note $\int_a^b f(t)dt$ la valeur commune $I_+(f, a, b) = I_-(f, a, b)$. On a $\int_a^b f(t)dt = I_+(f, a, b) - I_+(f, a, a)$ (le dernier terme est nul). Si f est continue, c'est donc la différence des valeurs en b et a d'une primitive de f, vu le dernier point du théorème. On est donc cohérent avec la notation introduite au début du chapitre précédent. De plus, le théorème précédent montre le premier théorème 7.1.3 de ce chapitre.

Démonstration. — On va en réalité montrer que toute fonction bornée et qui admet des limites à gauche et à droite est intégrable. En effet soit f une telle fonction. Définissons les deux fonctions F_+ et F_- sur $[a, b]$ par $F_+(x) = I_+(f, a, x)$ et $F_-(x) = I_-(f, a, x)$.

On constate que $F_+(a) = F_-(a) = 0$. On veut montrer que les deux fonctions $F_+(x) = I_+(f, a, x)$ et $F_-(x) = I_-(f, a, x)$ sont égales. Pour ça, montrons que la différence $F_+ - F_-$ est constante en montrant qu'elle est dérivable, de dérivée nulle.

Commençons par montrer que F_+ est dérivable à gauche en tout point x_0, de dérivée $l_g(x_0) = \lim_{x \to x_0^-} f(x)$. Pour tout $x < x_0$, la relation de Chasles pour I_+ permet d'écrire $F_+(x_0) - F_+(x) = \int_x^{x_0} f(t)dt$.

Fixons $\varepsilon > 0$. Par définition de $l_g(x_0)$ il existe $y < x_0$ tel que pour tout $y < x < x_0$, on a :

$$l_g(x_0) - \varepsilon \leq f(x) \leq l_g(x_0) + \varepsilon.$$

En utilisant l'encadrement pour I_+, on obtient :

$$(x_0 - x)(l_g(x_0) - \varepsilon) \leq I_+(f, x, x_0) \leq (x_0 - x)(l_g(x_0) + \varepsilon).$$

Autrement dit, pour tout $y < x < x_0$, on a :

$$l_g(x_0) - \varepsilon \leq \frac{F_+(x_0) - F_+(x)}{x_0 - x} \leq l_g(x_0) + \varepsilon.$$

On a bien montré que F_+ est dérivable à gauche en x_0, de dérivée la limite à gauche de f.

On montre de même que :

- F_+ est dérivable à droite en x_0, de dérivée la limite à droite de f.
- F_- est dérivable à gauche en x_0, de dérivée la limite à gauche de f.
- F_- est dérivable à droite en x_0, de dérivée la limite à droite de f.

Revenons à notre différence $F_+ - F_-$: elle est dérivable à gauche et à droite en tout point, de dérivée à gauche et à droite nulle. Elle est donc dérivable, de dérivée nulle.

Ainsi $F_+(x) = F_-(x)$ pour tout $a \leq x \leq b$, ce qui était l'objectif !

Enfin, si f est continue, sa limite à gauche et à droite en tout point sont égales à la valeur de f en ce point. Autrement dit, $F+$ est dérivable, de dérivée f. □

8.3. Quelques propriétés supplémentaires de l'intégrale des fonctions continues

8.3.1. Parité, imparité et intégrale. —

Proposition 8.3.1. — (QC)

- *Si $f : [-a, a] \to \mathbf{R}$ est impaire et continue, alors $\int_{-a}^{a} f(t)dt = 0$.*
- *Si $f : [-a, a] \to \mathbf{R}$ est paire et continue, alors $\int_{-a}^{a} f(t)dt = 2\int_{0}^{a} f(t)dt$.*

Démonstration. — Pour le premier point, on effectue le changement de variable $u = -t$. On obtient :

$$\int_{-a}^{a} f(t)dt = -\int_{a}^{-a} f(-u)du = \int_{-a}^{a} f(-u)du = -\int_{-a}^{a} f(u)du.$$

Donc l'intégrale est égale à son opposée : elle est nulle.

Pour le deuxième point on utilise d'abord la relation de Chasles avant d'effectuer le changement de variable $u = -t$ dans la première intégrale :

$$\int_{-a}^{a} f(t)dt = \int_{-a}^{0} f(t)dt + \int_{0}^{a} f(t)dt = \int_{0}^{a} f(-u)du + \int_{0}^{a} f(t)dt.$$

Or $\int_0^a f(-u)du = \int_0^a f(u)du$ par parité. On obtient bien :

$$\int_{-a}^{a} f(t)dt = 2\int_{0}^{a} f(t)dt.$$

□

(QC)

Proposition 8.3.2. — *Si $f : \mathbf{R} \to \mathbf{R}$ est continue et périodique de période $T > 0$, pour tout $a \in \mathbf{R}$ on a :*

$$\int_{a}^{a+T} f(t)dt = \int_{0}^{T} f(t)dt.$$

Démonstration. — Soit k l'entier tel que $kT \leq a < (k+1)T$. On écrit alors :

$$\begin{array}{rcl} \int_a^{a+T} f(t)dt & = & \int_a^{(k+1)T} f(t)dt + \int_{(k+1)T}^{a+T} f(t)dt \\ \text{(chgmt variable } u = t - T) & = & \int_a^{(k+1)T} f(t)dt + \int_{kT}^{a} f(u)du \\ \text{(Chasles)} & = & \int_{kT}^{(k+1)T} f(t)dt \\ \text{(chgmt variable } u = t - kT) & = & \int_0^T f(t)dt \end{array}$$

□

8.3.2. Inégalité et formule de la moyenne. —

(QC)

Proposition 8.3.3. — *Soit f une fonction continue et g une fonction positive et intégrable, toutes deux définies sur un intervalle $[a, b]$. Notons $m = \min_{[a,b]}(f)$ et $M = \max_{[a,b]}(f)$ Alors on a*

1. *$m \int_a^b g(x)dx \leq \int_a^b fg(x)dx \leq M \int_a^b g(x)dx$.*
2. *il existe $c \in [a, b]$ tel que $\int_a^b fg(x)dx = f(c) \int_a^b g(x)dx$.*

Démonstration. — Le deuxième point se déduit du premier par le théorème des valeurs intermédiaires appliqué à la fonction $x \mapsto \left(\int_a^b g(t)dt\right) f(x)$.

Le premier : on a par définition $m \leq f(t) \leq M$ pour tout $a \leq t \leq b$. On peut multiplier cette inégalité par $g(t) \geq 0$:

$$mg(t) \leq f(t)g(t) \leq Mg(t).$$

Par croissance de l'intégrale, on peut intégrer cette inégalité :

$$m\int_{a}^{b} g(t)dt \leq \int_{a}^{b} f(t)g(t)dt \leq M\int_{a}^{b} g(t)dt.$$

□

Corollaire 8.3.4. — *En appliquant avec $g = 1$, on obtient pour une fonction f continue :*

1. $m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$.
2. *il existe* $c \in [a,b]$ *tel que* $\int_a^b f(x)dx = f(c)(b-a)$.

Démonstration. — C'est clair, une fois avoir remarqué que, comme $g(t) = 1$, $\int_a^b g(t)dt = b - a$. □

8.4. Formule de Taylor avec reste intégral

Théorème 8.4.1 **(Formule de Taylor avec reste intégral)**

Soit f une fonction de classe C^{n+1} sur un intervalle I de $\mathbf{R}$ et $a \in I$. Alors, pour tout $b \in I$, on a :

$$f(b) = f(a) + f'(a)(b-a) + \ldots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}dt.$$

Démonstration. — On procède par récurrence sur n.

Commençons par le cas $n = 0$: soit f une fonction C^1 ; on veut montrer $f(b) = f(a) + \int_a^b f'(t)dt$. Or le lien entre primitive et intégrale nous donne $\int_a^b f'(t)dt = f(b) - f(a)$. Donc la formule est démontrée au rang 0.

Montrons le rang $n = 1$ pour bien comprendre la propagation : on part de la formule précédente et on fait une intégration par parties dans l'intégrale : on pose $u = f'$, $v' = 1$. Alors on peut prendre $u' = f''$ et $v(t) = -(b-t)$ (la primitive qui s'annule en b. On obtient alors :

$$f(b) = f(a) + \left[f'(t)(-(b-t))\right]_a^b - \int_a^b f''(t)(-(b-t))dt.$$

Après avoir simplifié les signes, on obtient :

$$f(b) = f(a) + f'(a)(b-a) + \int_a^b f''(t)(b-t)dt.$$

La formule est bien démontrée au rang 1.

Supposons qu'elle est démontrée au rang $n-1$ et montrons-la au rang n. Soit f une fonction de classe C^{n+1}. Elle est aussi de classe C^n, donc on peut lui appliquer l'hypothèse de récurrence :

$$f(b) = f(a) + f'(a)(b-a) + \ldots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} + \int_a^b f^{(n)}(t)\frac{(b-t)^{n-1}}{(n-1)!}dt.$$

Faisons alors une intégration par parties dans l'intégrale : cette fois-ci, on pose $u = f^{(n)}$ et $v'(t) = \frac{(b-t)^{n-1}}{(n-1)!}$. On peut choisir $u' = f^{(n+1)}$ et $v = -\frac{(b-t)^n}{n!}$. Il

vient alors :

$$\begin{aligned}\int_a^b f^{(n)}(t)\frac{(b-t)^{n-1}}{(n-1)!}dt &= \left[f^{(n)}\left(-\frac{(b-t)^n}{n!}\right)\right] - \int_a^b f^{(n+1)}(t)\left(-\frac{(b-t)^n}{n!}\right)dy \\ &= \frac{f^{(n)}(a)}{n!}(b-a)^n + \int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}dt.\end{aligned}$$

En utilisant cette égalité dans la formule donnée par la récurrence, on obtient bien :

$$f(b) = f(a) + f'(a)(b-a) + \ldots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}dt.$$

□

Grâce à la formule de la moyenne, on en déduit :

Théorème 8.4.2. — *Formule de Taylor-Lagrange Soit f une fonction de classe C^{n+1} sur un intervalle I de $\mathbf{R}$ et $a \in I$. Alors, pour tout $b > a$ dans I, il existe $c \in [a,b]$ tel que :*

$$f(b) = f(a) + f'(a)(b-a) + \ldots + \frac{f^{(n)}(a)}{n!}(b-a)^n + f^{(n+1)}(c)\frac{(b-a)^{n+1}}{(n+1)!}.$$

Ce développement n'a pas été traité en cours.

Démonstration. — On applique la formule de la moyenne à $\int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}dt$: la fonction $f^{(n+1)}$ est intégrable (car continue) et $t \mapsto \frac{(b-t)^n}{n!}$ est intégrable et positive sur $[a,b]$. Donc il existe $c \in [a,b]$ tel que

$$\int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!} = f^{(n+1)}(c)\int_a^b \frac{(b-t)^n}{n!}dt = f^{(n+1)}(c)\frac{(b-t)^{n+1}}{(n+1)!}.$$

□

Un commentaire sur ces deux formules : elles ont l'avantage par rapport à la formule classique, de donner une formulation explicite de l'erreur entre f et son approximation polynomiale de degré n. Notamment, ces formules peuvent être utilisée pour b pas forcément « très proche » de a.

8.5. Inégalité de Cauchy-Schwarz – Non traité en amphi

L'inégalité qui suit est fondamentale et n'est pas propre au monde de l'intégrale ; elle relève en fait plutôt du monde de l'algèbre bilinéaire. Une autre justification de sa présence dans ce cours est l'élégance de sa preuve.

Théorème 8.5.1. — *Inégalité de Cauchy-Schwarz Soit f et g deux fonctions intégrables sur $[a,b]$. Alors :*

$$\left(\int_a^b fg(t)dt\right)^2 \leq \left(\int_a^b f^2(t)dt\right)\left(\int_a^b g^2(t)dt\right).$$

Avant de prouver la formule, donnons en un corollaire :

Corollaire 8.5.2. — *Si f est intégrable sur $[0,1]$, alors*

$$\left(\int_0^1 f(t)dt\right)^2 \leq \int_0^1 f^2(t)dt.$$

Démonstration. — Dans l'inégalité de Cauchy-Schwarz, on pose $a=0$, $b=1$ et g constante égale à 1. On a alors $\int_0^1 g^2(t)dt = 1$. Donc on conclut :

$$\left(\int_0^1 f(t)dt\right)^2 \leq \int_0^1 f^2(t)dt.$$

□

Prouvons maintenant l'inégalité de Cauchy-Schwarz :

Démonstration. — Soient f et g intégrables, et pour $X \in \mathbf{R}$ posons

$$P(X) = \int_a^b (f+Xg)^2(t)dt.$$

On remarque d'abord que pour tout X, $P(X)$ est positif : c'est l'intégrale d'une fonction positive. D'un autre côté, on peut développer :

$$P(X) = \left(\int_a^b g^2(t)dt\right)X^2 + 2\left(\int_a^b fg(t)dt\right)X + \left(\int_a^b f^2(t)dt\right).$$

C'est donc un polynôme en X, de degré 2 qui est toujours positif. Donc il ne change pas de signe, et donc son discriminant Δ est négatif ou nul. Donc :

$$\Delta = 4\left(\int_a^b fg(t)dt\right)^2 - 4\left(\int_a^b f^2(t)dt\right)\left(\int_a^b g^2(t)dt\right) \leq 0.$$

On a bien l'inégalité voulue ! □

On va déterminer quand on a égalité dans l'inégalité. On utilisera la propriété suivante, déjà vue :

Proposition 8.5.3. — *Soit $f : [a,b] \to \mathbf{R}$ positive et continue. Si $\int_a^b f(t)dt = 0$, alors $f = 0$.* (QC)

Démonstration. — Considérons F la primitive de f qui s'annule en a : $F(x) = \int_a^x f(t)dt$. Alors, la dérivée de F est f, donc est positive : F est croissante. Si $\int_a^b f(t)dt = 0$, alors $F(b) = F(a) = 0$. Donc F est constant et sa dérivé f est nulle. □

CHAPITRE 9

SOMMES DE RIEMANN ET CALCUL APPROCHÉ D'INTÉGRALES

9.1. Sommes de Riemann

Cette section est très liée à l'exercice 1 de la troisième feuille sur l'intégration.

Définition 9.1.1. — Une *subdivision marquée* (σ, θ) d'un intervalle $[a, b]$ est la donnée d'une subdivision $\sigma = (a = x_0 < x_1 < \ldots < x_n = b)$ de $[a, b]$ et d'un marquage $\theta = (y_0, y_1, \ldots, y_{n-1})$ où chaque y_k est dans l'intervalle $[x_k, x_{k+1}]$.

Autrement dit, on choisit un point dans chaque intervalle défini par la subdivision.

Remarque 9.1.2. — En cours, on n'a traité pour alléger les notations que le cas de $y_i = x_i$, c'est à dire que le point choisi dans chaque intervalle est la borne de gauche.

Définition 9.1.3. — Soit (σ, θ) une subdivision marquée de $[a, b]$. Alors la *somme de Riemann* de f associée à (σ, θ) est :

$$S(f, \sigma, \theta) = \sum_{k=0}^{n-1} (x_{k+1} - x_k) f(y_k).$$

Donnons un peu de sens à cette somme : $S(f, \sigma, \theta)$ est l'intégrale de la fonction en escalier qui vaut $f(y_k)$ sur l'intervalle $[x_k, x_{k+1}]$.

Remarquons tout de suite un cas particulier intéressant : si $\sigma = (0 < \frac{1}{n} < \frac{2}{n} < \ldots < \frac{n-1}{n} < 1$ est la subdivision régulière de $[0, 1]$ en n intervalles, et $y_k = \frac{k}{n}$, alors on a comme en TD :

$$S(f, \sigma, \theta) = \frac{1}{n} \sum_{k=0}^{n-1} f(\frac{k}{n}).$$

Théorème 9.1.4. — *Soit f une fonction intégrable sur $[a,b]$ et (σ_n, θ_n) une suite de subdivisions marquées de $[a,b]$. On suppose que le pas de σ_n tend vers 0. Alors, on a*

$$S(f, \sigma_n, \theta_n) \xrightarrow{n\to\infty} \int_a^b f(t)dt.$$

Nous ne démontrerons pas ce théorème sous cette forme, mais plutôt le théorème plus précis qu'on obtient en supposant de plus que f est C^1 :

(QC)

Théorème 9.1.5. — *Soit f une fonction C^1 sur $[a,b]$ et (σ, θ) une subdivision marquée de $[a,b]$. On note ℓ le pas de σ. Alors, on a :*

$$\left| S(f, \sigma, \theta) - \int_a^b f(t)dt \right| \leq \ell(b-a) \sup_{[a,b]} |f'|.$$

Démonstration. — Soit φ la fonction en escalier dont la valeur sur l'intervalle $[x_k, x_{k+1}]$ est $f(y_k)$. On a :

$$\left| S(f, \sigma, \theta) - \int_a^b f(t)dt \right| = \left| \int_a^b (\varphi - f)(t)dt \right| \leq \int_a^b |\varphi - f|(t)dt.$$

Montrons que $|\varphi - f| \leq \ell \sup_{[a,b]} |f'|$: soit $x \in [a,b]$ et k tel que $x \in [x_k - x_{k+1}[$. Alors on a $|\varphi - f|(x) = |\varphi(x) - f(x)| = |f(y_k) - f(x)|$. D'après l'inégalité des accroissements finis, cette dernière grandeur est inférieure à $\sup_{[a,b]} |f'||y_k - x|$. Or y_k et x sont dans le même intervalle de la subdivision σ, donc à distance inférieure à ℓ. Ainsi, $|\varphi - f| \leq \ell \sup_{[a,b]} |f'|$. En intégrant entre a et b, on obtient :

$$\left| S(f, \sigma, \theta) - \int_a^b f(t)dt \right| \leq \int_a^b |\varphi - f|(t)dt \leq (b-a)\ell \sup_{[a,b]} |f'|.$$

□

En particulier pour le cas des subdivisions régulières :

(QC)

Corollaire 9.1.6. — *Soit f une fonction C^1 sur $[0,1]$. Alors on a*

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) - \int_0^1 f(t)dt \right| \leq \frac{\sup_{[0,1]} |f'|}{n}.$$

En particulier $\frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \xrightarrow{x\to\infty} \int_0^1 f(t)dt$.

Le résultat précédent a un grand intérêt théorique : il permet de trouver la limite de sommes qu'on ne saurait pas traiter autrement. On renvoie au td pour des exemples.

9.2. Calculs approchés d'intégrales

Un problème important est de savoir en pratique calculer des valeurs approchées d'intégrales. Pour ça, on approxime la fonction par des fonctions en escalier, ou affine par morceaux, ou plus complexes... Le point de départ est le théorème précédent sur les sommes de Riemann : une somme de Riemann n'est pas forcément difficile à calculer et peut fournir une approximation (en $\frac{1}{n}$) de l'intégrale. On renvoie à la première partie de l'exercice 1 de la troisième feuille de TD sur l'intégration.

On peut, sans augmenter le nombre de calculs à faire (calculer n fois une valeur de la fonction), obtenir une approximation en $\frac{1}{n^2}$: en approximant la fonction à intégrer par des fonctions affines par morceaux, on effectue alors la méthode des trapèzes, où l'erreur est en $\frac{1}{n^2}$ (pour $n+1$ valeurs calculées). On renvoie à la deuxième partie de l'exercice 1 de la troisième feuille de TD sur l'intégration.

9.3. Épilogue : Intégrale de fonctions à valeurs complexes

Étendons rapidement la définition de l'intégrale aux fonctions à valeurs complexes :

Définition 9.3.1. — Soit $f : [a, b] \to \mathbf{C}$ une fonction. Notons $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ ses parties réelle et imaginaire.

- On dit que f est intégrable si et seulement si $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.
- Dans ce cas, on note $\int_a^b f(t)dt = \int_a^b \mathrm{Re}(f)(t)dt + i \int_a^b \mathrm{Im}(f)(t)dt$.

On déduit sans difficultés des propriétés de l'intégrale des fonctions à valeurs réelles :

Proposition 9.3.2. — *La relation de Chasles est toujours vérifiée : si $f : [a, b] \to \mathbf{C}$ est intégrable et $c \in [a, b]$, on a :*

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

De plus la propriété de linéarité est toujours vérifiée.

PARTIE IV

ALGÈBRE LINÉAIRE

CHAPITRE 10

ESPACES VECTORIELS

La notion d'espace vectoriel est omniprésente en mathématiques et dans ses applications. Un espace vectoriel est un ensemble d'éléments appelés vecteurs qu'on sait additionner et qu'on sait multiplier par un nombre réel (ou complexe) – appelé scalaire. On a déjà rencontré des exemples d'espaces dans lesquels on sait additionner et multiplier par un scalaire. Ces exemples sont à garder en tête :

- les espaces de la géométrie euclidienne : la droite $\mathbf{R}$, le plan $\mathbf{R}^2$, l'espace $\mathbf{R}^3$ et plus généralement $\mathbf{R}^n$;
- l'espace des suites réelles $\mathcal{S}_{\mathbf{R}}$,
- l'espace $\mathcal{F}(\mathbf{R}, \mathbf{R})$ des fonctions de $\mathbf{R}$ dans $\mathbf{R}$
- l'espace des polynômes $\mathbf{R}[X]$.
- l'espace $\mathcal{M}_{p,q}(\mathbf{R})$ des matrices de taille $p \times q$.

On doit penser aussi à leurs analogues sur le corps des complexes : $\mathbf{C}$, $\mathbf{C}^2$, $\mathbf{C}^3$, $\mathbf{C}^n$, $\mathbf{C}^{\mathbf{N}}$, $\mathcal{F}(\mathbf{R}, \mathbf{C})$, $\mathbf{C}[X]$, $\mathcal{M}_{p,q}(\mathbf{C})$.

J'encourage le lecteur à vérifier avant de lire la suite qu'il sait effectivement additionner dans ces espaces et qu'il sait multiplier par un nombre réel (ou complexe pour les analogues). On peut remarquer qu'on sait faire d'autres opérations : multiplier des suites, composer des fonctions... Nous ne développerons pas dans ce cours les structures sous-jacentes.

La première section peut sembler abrupte et abstraite : nous formalisons le fait que les espaces de la liste ci-dessus partagent les mêmes propriétés, et nous arrivons à la définition d'espace vectoriel. Cette section n'a été que survolée en cours et le lecteur peut, dans un premier temps, la survoler. Nous essayons de l'illustrer par de nombreux exemples.

10.1. Groupes abéliens, corps, espaces vectoriels

Nous avons dit qu'un espace vectoriel est un ensemble dans lequel nous savons faire deux choses :

- Additionner des vecteurs. C'est formalisé par la définition de *groupe abélien.*
- Multiplier par un scalaire. L'espace des scalaires est un *corps*, dont nous donnons la définition.

Enfin, il y a des liens entre ces opérations. Le lecteur vérifiera sur les exemples ci-dessus que la multiplication par un scalaire se distribue sur l'addition. La formalisation de ces liens mène à la définition d'*espace vecoriel.*

10.1.1. Groupes abéliens. —

Définition 10.1.1. — Un groupe abélien $(A, +)$ est un ensemble A avec une loi de composition interne $+ : A \times A \to A$ telle que :

- Il existe un élément neutre 0_A vérifiant : $\forall a \in A$, $0_A + a = a + 0_A = a$.
- Tout élément $a \in A$ possède un symétrique $(-a) \in A$ vérifiant $a+(-a) = (-a) + a = 0_A$.
- $+$ est associative : $\forall a, b, c \in A, (a + b) + c = a + (b + c)$.
- $+$ est commutative : $\forall a, b \in A, a + b = b + a$.

Si on enlève le dernier axiome, on définit ce qu'on appelle simplement un groupe.

Il découle de la définition que

- L'élément neutre est unique [1]. Si le contexte est clair, on le notera simplement 0.
- Le symétrique d'un élément a est unique [2]. On introduit donc la notation de la soustraction : $a - b = a + (-b)$.
- Grâce à l'associativité et la commutativité, on définit sans ambiguïté la somme $\sum_{i=1}^{n} a_i$ d'une famille d'éléments $(a_i)_{1 \leq i \leq n}$ d'éléments de A.

Exemple 10.1.2. — Les objets suivants sont des groupes abéliens :

- $(\mathbf{Z}, +)$, $(\mathbf{Q}, +)$, $(\mathbf{R}, +)$, $(\mathbf{C}, +)$ [3].
- $(\mathbf{R}^*, \times)$, $(\mathbf{C}^*, \times)$, $(\{z \in \mathbf{C}, |z| = 1\}, \times)$ [4].

1. Si 0_A et $0'_A$ vérifient tous les deux la définition d'élément neutre, on écrit $0_A = 0'_A + 0_A = 0'_A$ en utilisant cette définition une fois pour $0'_A$ et une fois pour 0_A.
2. Si b et c vérifient $a+b = a+c = 0_A$, on a $c = 0_A + c = (b+a)+c = b+(a+c) = b+0_A = b$.
3. Attention, $(\mathbf{N}, +)$ n'est pas un groupe abélien : 1 n'a pas de symétrique.
4. Attention, $(\mathbf{R}, \times)$ n'est pas un groupe abélien : 0 n'a pas de symétrique

- $(\mathbf{R}^2, +)$, $(\mathbf{R}^3, +)$ et plus généralement $(\mathbf{R}^n = \{(x_1, \ldots, x_n), x_i \in \mathbf{R}\}, +)$.
- L'espace des fonctions $\mathcal{F}(\mathbf{R}, \mathbf{R})$ de $\mathbf{R}$ dans $\mathbf{R}$ [5] avec l'addition des fonctions. L'élément neutre est la fonction nulle.
- L'espace $\mathcal{S}_{\mathbf{R}}$ des suites réelles avec leur addition. L'élément neutre est la suite nulle.
- L'espace $\mathbf{R}[X]$ des polynômes à coefficients réels avec leur addition. L'élément neutre est le polynôme nul.
- L'espace $\mathcal{M}_{p,q}(\mathbf{R})$ des matrices réelles de taille $p \times q$ avec leur addition. L'élément neutre est la matrice nulle.
- Les analogues complexes $\mathbf{C}^2$, $\mathbf{C}^3$, $\mathbf{C}^n$, $\mathcal{F}(\mathbf{R}, \mathbf{C})$, $\mathbf{C}^{\mathbf{N}}$ et $\mathbf{C}[X]$ chacun muni de son addition usuelle.

10.1.2. Corps. —

Définition 10.1.3. — Un corps $(\mathbf{K}, +, \times)$ est un ensemble muni de deux lois de compositions internes $+$ (l'addition) et $\times$ (la multiplication) telles que :

- $(\mathbf{K}, +)$ est un groupe abélien. On note 0_K son élément neutre.
- $(\mathbf{K} \setminus \{0\}, \times)$ est un groupe abélien. On note 1_K son élément neutre.
- La multiplication est distributive par rapport à l'addition : $\forall x, y, z \in \mathbf{K}, x \times (y + z) = x \times y + x \times z$ et $(y + z) \times x = y \times x + z \times x$

Les exemples connus du lecteur sont $\mathbf{Q}$, $\mathbf{R}$ et $\mathbf{C}$. Dans le cadre de ce cours, le lecteur peut penser, en lisant « Soit $\mathbf{K}$ un corps... », à « Soit $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$... ».

Remarque 10.1.4. — Si on ne demande pas l'existence d'un inverse (i.e. un symétrique pour la loi $\times$) des éléments non nuls, on définit un *anneau commutatif*. Par exemple, $\mathbf{Z}$ est un anneau commutatif. On rencontrera dans ce cours des anneaux non-commutatifs (l'anneau $\mathcal{M}_n(\mathbf{R})$ des matrices carrées de taille n) : la multiplication n'est pas commutative (mais l'addition le reste).

10.1.3. Espaces vectoriels. — Soit $\mathbf{K}$ un corps.

Définition 10.1.5. — Un $\mathbf{K}$-espace vectoriel (abrégé en $\mathbf{K}$-ev) $(E, +)$ est un groupe abélien avec une loi de composition externe

$$\begin{array}{rccc} \cdot & : & \mathbf{K} \times E & \to & E \\ & & (\lambda, v) & \mapsto & \lambda \cdot v \end{array}$$

telle que :

5. Plus généralement, $\mathcal{F}(X, A)$ où X est un ensemble et A un groupe abélien.

- $\forall v \in E$, $1_{\mathbf{K}} \cdot v = v$; $\forall \lambda, \lambda' \in \mathbf{K}$, $v \in E$, $\lambda \cdot (\lambda' \cdot v) = (\lambda\lambda') \cdot v$.
- Il y a une double distributivité de la loi $\cdot$ par rapport aux additions dans $\mathbf{K}$ et E :

$$\forall \lambda, \lambda' \in \mathbf{K},\ v, v' \in E, (\lambda + \lambda') \cdot (v + v') = \lambda \cdot v + \lambda' \cdot v + \lambda \cdot v' + \lambda' \cdot v'.$$

On notera 0_E parfois comme le vecteur nul $\vec{0}$ et les élements neutre du corps 0 et 1. Le corps $\mathbf{K}$ est appelé corps des scalaires de E.

Les axiomes impliquent que :

Proposition 10.1.6. — *Soit E un $\mathbf{K}$-ev. Alors :*

- *pour tous vecteurs u, v, w dans E, $u + v = u + w \Rightarrow v = w$. (Cette propriété est en fait vrai dans tout groupe abélien.)*
- *pour tout $v \in E$, $0 \cdot v = \vec{0}$ et $(-1) \cdot v = -v$.*
- *Pour tous $\lambda \in \mathbf{K}$ et $v \in E$, on a $\lambda \cdot v = \vec{0} \Leftrightarrow (\lambda = 0$ ou $v = \vec{0})$.*

Démonstration. —
- On ajoute $(-u)$ à $u + v = u + w$ pour obtenir $v = w$.
- pour tout $v \in E$, $0 \cdot v = \vec{0}$: en effet, on a $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$. En ajoutant $-(0 \cdot v)$, on obtient $0 \cdot v = \vec{0}$. De même $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = \vec{0}$. Donc $(-1) \cdot v = -v$.
- Pour le sens $\Leftarrow$: on a déjà vu que $0 \cdot v = \vec{0}$. Le fait que $\lambda \cdot \vec{0} = \vec{0}$ découle de $\vec{0} = \vec{0} + \vec{0}$: on en déduit $\lambda \cdot \vec{0} = \lambda \cdot \vec{0} + \lambda \cdot \vec{0}$, puis on soustrait $\lambda \cdot \vec{0}$ à gauche et à droite pour obtenir $\lambda \cdot \vec{0} = \vec{0}$.

  Pour l'implication $\Rightarrow$: soit λ et v tels que $\lambda \cdot v = \vec{0}$. Supposons que $\lambda \neq 0$. Alors on considère l'inverse λ^{-1} de λ et on a :

$$\vec{0} = \lambda^{-1} \cdot \vec{0} = \lambda^{-1} \cdot (\lambda \cdot v) = (\lambda^{-1}\lambda) \cdot v = 1 \cdot v = v.$$

  Donc $v = \vec{0}$. L'implication est prouvée.

□

Exemple 10.1.7. — La liste d'exemples de groupes abéliens donnés plus haut donnent des $\mathbf{K}$-ev :

- $\mathbf{R}^2$, $\mathbf{R}^3$, $\mathbf{R}^n$, $\mathcal{F}(\mathbf{R}, \mathbf{R})$, $\mathbf{R}^{\mathbf{N}}$, $\mathbf{R}[X]$ et $\mathcal{M}_{p,q}(\mathbf{R})$ sont des $\mathbf{R}$-ev.
- $\mathbf{C}^2$, $\mathbf{C}^3$, $\mathbf{C}^n$, $\mathcal{F}(\mathbf{R}, \mathbf{C})$, $\mathbf{C}^{\mathbf{N}}$, $\mathbf{C}[X]$ et $\mathcal{M}_{p,q}(\mathbf{C})$ sont des $\mathbf{C}$-ev.
- $\mathbf{C}$ est un $\mathbf{C}$-ev, mais aussi un $\mathbf{R}$-ev. Tout élément z de $\mathbf{C}$ s'écrit $z = \mathrm{Re}(z) \cdot 1 + \mathrm{Im}(z) \cdot i$.

10.2. Combinaisons linéaires, Sous-espaces vectoriels

Soit $\mathbf{K}$ un corps et E un $\mathbf{K}$-ev.

10.2.1. Combinaisons linéaires. —

Définition 10.2.1. — Si on a des vecteurs $v_1, \dots v_n$ dans E et des scalaires $\lambda_1, \dots \lambda_n$ dans $\mathbf{K}$, la combinaison linéaire des v_i de coefficients λ_i est la somme :

$$\sum_{i=1}^{n} \lambda_i \cdot v_i.$$

Exemple 10.2.2. —
- L'ensemble des combinaisons linéaires d'un vecteur v non nul est la droite des vecteurs qui lui sont proportionnels.
- Dans $\mathbf{R}^2$, tout vecteur $v = \begin{pmatrix} x \\ y \end{pmatrix}$ est une combinaison linéaire des deux vecteurs « de base » $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$:
$$v = x \cdot v_1 + y \cdot v_2.$$
- (Autre interprétation du même phénomène) On a vu (exemples 10.1.7) que dans le $\mathbf{R}$-ev $\mathbf{C}$, tout vecteur z est combinaison linéaire des vecteurs 1 et i.

Il y a un lien entre combinaison linéaire et système linéaire. En effet, considérons un système linéaire $AX = b$, où A est une matrice de taille $p \times n$, X un vecteur d'inconnus, et b un vecteur colonne de taille p, autrement dit un vecteur de $\mathbf{R}^p$. Notons C_1, …, C_n les colonnes de A : ce sont aussi des vecteurs de $\mathbf{R}^p$.

Soit maintenant une solution $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. On peut écrire le produit AX sous la forme $AX = x_1C_1 + \dots x_nC_n$. L'équation $AX = b$ se réécrit donc :

$$x_1C_1 + \dots x_nC_n = b.$$

Autrement dit, une solution du système est une façon d'écrire b comme combinaison linéaire des colonnes de A.

10.2.2. Sous-espaces vectoriels. —

Définition 10.2.3. — On appelle sous-espace vectoriel (abrégé en sous-ev) un sous-ensemble F non-vide d'un $\mathbf{K}$-ev E qui est stable par combinaison linéaire : Si des vecteurs $v_1, \dots, v_n$ sont dans F et si $\lambda_1, \dots, \lambda_n$ sont des scalaires de $\mathbf{K}$, alors la combinaison linéaire $\sum_{i=0}^{n} \lambda_i \cdot v_i$ est encore dans F. (QC)

Pour vérifier qu'un sous-ensemble est bien un sous-ev, on utilisera en pratique la condition plus simple suivante :

Proposition 10.2.4. — *Pour un sous-ensemble F non vide de E, on a l'équivalence :*

F est un sous-espace vectoriel $\Leftrightarrow$ pour tous u, v dans F et λ dans $\mathbf{K}$, $u+\lambda\cdot v$ appartient à F.

Démonstration. — Le sens $\Rightarrow$ est immédiat.

Dans l'autre sens, on montre par récurrence sur le nombre n de vecteurs qu'une combinaison linéaire est dans F. Notons d'abord que la deuxième condition implique que $\vec{0} \in F$: on prend un élément $u \in F$ (F est supposé non vide) et on forme $u + (-1) \cdot u = \vec{0}$. De plus, si v est dans F, $-v = \vec{0} + (-1) \cdot v$ est encore dans F. Faisons maintenant le raisonnement par récurrence :

- si $n = 1$, on n'a qu'un vecteur $v \in F$ et un scalaire λ, et $\lambda \cdot v = \vec{0} + \lambda$ appartient bien à F par la deuxième condition.
- Supposons que toute combinaison linéaire de n vecteurs de F est dans F. Considérons une combinaison linéaire $w = \sum_{i=1}^{n+1} \lambda_i \cdot v_i$ de vecteurs de F. Alors, par hypothèse de récurrence la combinaison linéaire $u = \sum_{i=1}^{n} \lambda_i \cdot v_i$ des n premiers vecteurs est dans F. Et par la condition de droite de l'équivalence, on a $w = u + \lambda_{n+1} \cdot v_{n+1} \in F$.

□

Exemple 10.2.5. — Les sous-ensembles $\{\vec{0}\}$ et E sont des sous-espaces vectoriels de E (appelés parfois « triviaux »).

L'ensemble des solutions d'un système linéaire homogène $AX = 0$ avec A de taille $p \times n$ est un sous-espace vectoriel de $\mathbf{R}^n$.

Remarquons qu'on a montré qu'un sous-espace vectoriel contient toujours le vecteur nul et le symétrique de chaque vecteur. L'addition de E se restreint alors en une addition dans F, qui vérifie les axiomes de groupe abélien. De même, la multiplication par un scalaire sur E se restreint à F et vérifie les axiomes de compatibilité de la définition des $\mathbf{K}$-ev. En d'autres termes, on a :

Théorème 10.2.6. — *Tout sous-espace vectoriel F est un $\mathbf{K}$-ev pour les lois d'addition et de multiplication par un scalaire induite par celles de E.*

Exemple 10.2.7. —

- Dans $\mathbf{R}^2$, la droite horizontale $\{\binom{x}{0}, x \in \mathbf{R}\}$ est un sous-ev.
- Dans $\mathcal{S}_{\mathbf{R}}$, l'ensemble des suites bornées est un sous-espace vectoriel ainsi que celui des suites convergentes.

– Dans $\mathcal{F}(\mathbf{R}, \mathbf{R})$, les ensembles de fonctions bornées, ou continues, ou intégrables ou dérivables ou infiniment dérivables sont des sous-ev.

Exercice : démontrer ces affirmations (voir aussi le TD).

Soit $(v_1, \ldots, v_n)$ une famille de vecteurs de E. L'ensemble, noté $\mathrm{Vect}(v_1, \ldots, v_n)$, des combinaisons linéaires des v_i est stable par combinaison linéaire ; on obtient donc le théorème :

Théorème 10.2.8. — *L'ensemble* $\mathrm{Vect}(v_1, \ldots v_n)$ *est un sous-espace vectoriel de* E *; c'est le plus petit sous-espace vectoriel contenant les vecteurs* v_i*.* (QC)

On appelle cet ensemble le sous-espace vectoriel engendré par $(v_1, \ldots, v_n)$.

Démonstration. — Si $u = \sum_{i=1}^{n} \alpha_i \cdot v_i$ et $w = \sum_1^n \beta_i \cdot v_i$ sont deux éléments de $\mathrm{Vect}(v_1 \ldots, v_n)$, et si $\lambda \in \mathbf{K}$, on a :

$$u + \lambda \cdot w = \sum_1^n (\alpha_i + \lambda\beta_i) \cdot v_i \in \mathrm{Vect}(v_1, \ldots, v_n).$$

Ça prouve que $\mathrm{Vect}(v_1, \ldots, v_n)$ est un sous-ev. Maintenant, si un sous-espace vectoriel contient tous les v_i, par définition il contient toutes leurs combinaisons linéaires, donc tout $\mathrm{Vect}(v_1, \ldots, v_n)$. □

On peut remarquer que dans le devoir sur les suites récurrentes linéaires d'ordre 2 (voir la section **??**), on a montre que l'espace des suites solutions d'une récurrence linéaire d'ordre 2 est un sous-espace vectoriel de l'espace des suites et qu'il est engendré par deux vecteurs.

10.2.3. Intersection de sous-espaces vectoriels. —

Proposition 10.2.9. — *Soit* E *un* $\mathbf{K}$*-ev et* F*,* F' *deux sous-ev. Alors l'intersection* $F \cap F'$ *est un sous-ev.*

Démonstration. — Si u et v sont dans $F \cap F'$ et λ est dans $\mathbf{K}$, alors $u + \lambda \cdot v$ sont dans F et dans F' car chacun des deux est un sous-ev. Donc $u + \lambda \cdot v$ est dans $F \cap F'$. □

10.3. Application linéaire

10.3.1. Définition. — Soient E et E' deux $\mathbf{K}$-espaces vectoriels.

Définition 10.3.1. — Une application f de E dans E' est dite linéaire si elle préserve la somme et la multiplication par un scalaire : (QC)

$$\forall u, v \in E,\ \lambda \in \mathbf{K},\ f(u + \lambda \cdot v) = f(u) + \lambda \cdot f(v).$$

On note $\mathcal{L}(E, E')$ l'ensemble des applications linéaires de E dans E'.

Exemple 10.3.2. —
- Les rotations et homothéties du plan de centre 0 sont des applications linéaires.
- L'application $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x - y + z$ est linéaire de $\mathbf{C}^3$ dans $\mathbf{C}$.
- Sur l'espace des suites réelles convergentes, l'application « limite » à valeurs dans $\mathbf{R}$, qui à une suite associe sa limite, est linéaire.

Comme dans le cas des sous-espaces vectoriels, on montre par récurrence :

Proposition 10.3.3. — *Une application de E dans E' est linéaire si et seulement si pour toute combinaison linéaire $\sum_1^n \lambda_i \cdot v_i$ d'éléments de E, on a*

$$f\left(\sum_1^n \lambda_i \cdot v_i\right) = \sum_1^n \lambda_i \cdot f(v_i).$$

Notamment, $f(0_E) = 0_{E'}$ et $f(-u) = -f(u)$.

Remarque 10.3.4. — On définit naturellement la somme de deux applications linéaires : $(f + f')(v) = f(v) + f'(v)$ et la multiplication par un scalaire : $(\lambda \cdot f)(v) = \lambda \cdot f(v)$. De plus, l'application nulle qui envoie tous les vecteurs de E sur le vecteur nul de E' est linéaire. C'est l'élément neutre de l'addition. Avec ces opérations, $\mathcal{L}(E, E')$ est un $\mathbf{K}$-espace vectoriel. Exercice : le vérifier.

Un peu de vocabulaire (les mots viennent du grec : morphé veut dire "forme" et homo "semblable", endon "en dedans", iso "même") :
- Une application linéaire entre deux espaces est appelée un *(homo)morphisme* d'espaces vectoriels.
- Une application linéaire bijective est appelée un *isomorphisme* d'espaces vectoriels.
- Une application linéaire de E dans E est appelée un *endomorphisme* de E.
- Un endomorphisme qui est bijectif est appelé un *automorphisme* de E.

Exemple 10.3.5. — Si E est un espace vectoriel, on note id_E l'application identité [6] de E. C'est un automorphisme de E.

La composition de deux applications linéaires est encore linéaire :

Proposition 10.3.6. — *Soient $f \in \mathcal{L}(E, E')$ et $g \in \mathcal{L}(E', E'')$ deux applications linéaires. Alors la composée $g \circ f$ est une application linéaire de E dans E''.* (QC)

Démonstration. — Soient u et v deux vecteurs de E et $\lambda \in \mathbf{K}$. Alors on a $f(u + \lambda \cdot v) = f(u) + \lambda \cdot f(v)$ par linéarité de f. Par linéarité de g, on a :

$$g \circ f(u + \lambda \cdot v) = g(f(u) + \lambda \cdot f(v)) = g(f(u)) + \lambda \cdot g(f(v)).$$

On obtient bien $g \circ f(u+\lambda \cdot v) = g \circ f(u) + \lambda \cdot g \circ f(v)$; donc $g \circ f$ est linéaire. □

10.3.2. Application linéaire associée à une matrice. — Soit p, n deux entiers. On note les vecteurs de $\mathbf{K}^n$ et $\mathbf{K}^p$ comme des vecteurs colonnes (autrement dit, on considère un élément de $\mathcal{M}_{n,1}(\mathbf{K})$ comme un élément de $\mathbf{K}^n$). Remarquons que si A est dans $\mathcal{M}_{p,n}(\mathbf{K})$ et $X \in \mathbf{K}^n$ est un vecteur colonne, alors AX est encore un vecteur colonne de taille p, donc un élément de $\mathbf{K}^p$.

Proposition 10.3.7. — *Soit $A \in \mathcal{M}_{p,n}(\mathbf{K})$; alors l'application*

$$f_A : \begin{cases} \mathbf{K}^n & \to & \mathbf{K}^p \\ X & \mapsto & AX \end{cases}$$

est linéaire. On l'appelle application linéaire associée à A.

Démonstration. — C'est une conséquence de la bilinéarité du produit matriciel :

$$f_A(X + tX') = A(X + tX') = AX + tAX' = f_A(X) + tf_A(X').$$

□

Exemple 10.3.8. — Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, alors on a

$$f_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 4x + 5y + 6z \end{pmatrix}.$$

6. On rappelle que, si X est un ensemble, id_X est la fonction qui envoie tout élément x sur x.

10.3.3. Noyau, Image. —

Définition 10.3.9. — Soit $f \in \mathcal{L}(E, E')$. On appelle :

(QC)
- le noyau de f, noté[7] $\ker(f)$, est $\{v \in E, f(v) = 0\}$.
- l'image de f, notée $\mathrm{im}(f)$, est $\{v' \in E', \exists v \in E, f(v) = v'\}$.

(QC) ***Proposition 10.3.10***. — *Le noyau de f est un sous-espace vectoriel de E, l'image un sous-espace vectoriel de E'.*

Démonstration. — Ça découle de la définition. Par exemple : si u et v sont dans $\ker(f)$ et $\lambda \in \mathbf{K}$, $f(u + \lambda \cdot v) = f(u) + \lambda \cdot f(v) = 0$. Donc $u + \lambda \cdot v$ est dans $\ker(f)$. □

Le noyau permet de mesurer l'injectivité de f :

(QC) ***Proposition 10.3.11***. — *Une application linéaire f est injective si et seulement si* $\ker(f) = \{\vec{0}\}$.

Démonstration. — Si f est injective, comme $f(\vec{0}) = \vec{0}$, $f(v) \neq \vec{0}$ pour tout $v \neq \vec{0}$. Donc $\ker(f) = \{0\}$.

Réciproquement : soit $v \neq v' \in E$. Alors $v - v' \neq \vec{0}$. Donc $v - v' \notin \ker(f)$; on en déduit $\vec{0} \neq f(v - v') = f(v) - f(v')$. On conclut : $f(v) \neq f(v')$. □

Encore une fois, on peut faire un retour sur les systèmes linéaires : soit $AX = b$ un système linéaire. D'abord, ce système a une solution si et seulement si il existe un vecteur X tel que $AX = f_A(X) = b$. Autrement dit, ce système a des solutions si et seulement si $b \in \mathrm{im}(f_A)$.

De plus, l'ensemble des solutions du système homogène $AX = 0$ est exactement le noyau de l'application linéaire f_A associée à A. Cette remarque donne le moyen de calculer le noyau d'une application linéaire f_A associée à une matrice A : c'est l'ensemble des solutions du système linéaire $AX = 0$, qu'on sait résoudre par pivot de Gauss.

10.4. Bases, dimensions

On considère toujours E un $\mathbf{K}$-espace vectoriel.

10.4.1. Familles libres, bases. —

(QC) ***Définition 10.4.1***. — Soient $(v_1, \ldots, v_n)$ une famille[8] d'éléments de E.

7. De l'allemand kernel, qui veut dire noyau
8. C'est-à-dire un ensemble numéroté d'éléments.

1. On dit que la famille est *génératrice* si $\mathrm{Vect}(v_1, \ldots, v_n) = E$.
2. On dit que la famille est *liée* s'il existe des scalaires $\lambda_1, \ldots, \lambda_n$ non tous nuls tels que $\sum_1^n \lambda_i \cdot v_i = 0$. Une telle relation est appelée relation de liaison entre les v_i.
3. On dit que la famille est *libre* si elle n'est pas liée.

Exemple 10.4.2. — La famille de $\mathbf{R}^n$ suivante est à la fois libre et génératrice.

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, v_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Exercice : vérifier que c'est bien une base.

Définition 10.4.3. — Si une famille $(v_1, \ldots v_n)$ est à la fois libre et génératrice, on dit qu'elle est une *base* de E.

Remarques 10.4.4. —
- Si la famille $(v_1, \ldots v_n)$ est liée, toute famille plus grande $(v_1, \ldots, v_p)$ avec $p > n$ est liée : on prend les $(\lambda_i)_{1 \leq i \leq n}$ donné par une relation de liaison (ils sont donc non tous nuls). On définit $\lambda_i = 0$ pour $n + 1 \leq i \leq p$. Alors $\sum_1^p \lambda_i \cdot v_i = 0$ est une relation de liaison pour $(v_1, \ldots, v_p)$.
- Par contraposée, toute sous-famille d'une famille libre est libre.
- Une famille est libre ssi on a : $\sum_1^n \lambda_i \cdot v_i = 0 \Rightarrow$ tous les λ_i sont nuls. Ou encore : si deux combinaisons linéaires des v_i sont égales, elles ont les mêmes coefficients. En effet : si $\sum_1^n \lambda_i \cdot v_i = \sum_1^n \mu_i \cdot v_i$, alors $\sum_1^n (\lambda_i - \mu_i) \cdot v_i = 0$, donc $\lambda_i = \mu_i$ pour tout i.
- Finalement, $(v_1, \ldots, v_n)$ est une base si et seulement si tout vecteur de E s'écrit de manière unique comme combinaison linéaire des vecteurs v_i.

On peut construire une base en enlevant des éléments redondants dans une famille génératrice :

Proposition 10.4.5. — *On peut extraire une base de E de toute famille génératrice finie $(u_1, \ldots, u_p)$.*

Démonstration. — On enlève les vecteurs redondants tant qu'on peut :

Si la famille est libre, on a bien une base.

Sinon, elle est liée : un des vecteurs est combinaison linéaire des autres : on peut l'enlever. On obtient une sous-famille de cardinal $p - 1$ qui est toujours génératrice. Et on continue à enlever les vecteurs inutiles jusqu'à obtenir une famille libre.

On arrivera toujours à une famille libre : sinon, ça veut dire qu'on peut enlever tous les vecteurs. Mais la famille vide n'est pas génératrice. □

On aura besoin d'un lemme à propos des familles libres :

Lemme 10.4.6 **(Lemme d'extension des familles libres)**
Soit $(v_1, \dots, v_n)$ *une famille libre de* E *et* $v \in E$ *un vecteur qui n'est pas dans* $\mathrm{Vect}(v_1, \dots, v_n)$.
Alors la famille $(v_1, \dots, v_n, v)$ *est encore libre.*

Démonstration. — Considérons une combinaison linéaire nulle $\sum_{i=1}^{n} \lambda_i \cdot v_i + \lambda \cdot v = 0_E$. Il s'agit de montrer que tous les coefficients λ_i et λ sont nuls.

Tout d'abord, si $\lambda \neq 0$, alors on peut écrire $v = \sum_{i=1}^{n} \frac{-\lambda_i}{\lambda} \cdot v_i$ comme une combinaison linéaire des vecteurs v_i. Mais on a supposé que v n'est pas dans l'espace vectoriel engendré par les v_i. Donc $\lambda = 0$.

Ainsi, la combinaison linéaire s'écrit $\sum_{i=1}^{n} \lambda_i \cdot v_i = 0$. Comme la famille des v_i est libre, on obtient bien que tous les λ_i sont nuls. □

Enfin, une application de l'injectivité des applications linéaires nous sera utile :

Proposition 10.4.7. — *Soit* $f : E \to E'$ *une application linéaire injective, et* v_1, …, v_n *une famille de vecteurs de* E.
Alors v_1, … v_n *est liée si et seulement si* $f(v_1)$, …, $f(v_n)$ *est liée.*

Démonstration. — L'implication directe : si $\sum \lambda_i v_i = 0$ avec des λ_i non tous nuls, alors on a $f(\sum \lambda_i v_i) = f(0) = 0$. Or, $f(\sum \lambda_i v_i) = \sum \lambda_i f(v_i)$ par linéarité. On a bien obtenu une relation de liaison entre les $f(v_i)$.

Réciproquement : si $\sum \lambda_i f(v_i) = 0$, avec les λ_i non tous nuls, alors on peut écrire par linéarité $f(\sum \lambda_i v_i) = 0$. Or f est supposée injective, donc on peut en déduire que $\sum \lambda_i v_i = 0$. Nous avons bien une relation de liaison. □

10.4.2. Coordonnées dans une base. — Soit E un $\mathbf{K}$-ev est $\mathcal{B} = (u_1, \dots u_n)$ une base de E.

Définition 10.4.8. — Pour $v \in E$, on appelle *coordonnées de* v *dans la base* $\mathcal{B}$ le vecteur $\begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix}$ tel que $v = \sum_1^n \lambda_i \cdot u_i$.

Remarque 10.4.9. — On a vu qu'un vecteur s'écrivait de manière unique comme combinaison linéaire des vecteurs d'une base. Donc les coordonnées sont bien définies.

Proposition 10.4.10. — *L'application « coordonnées dans la base $\mathcal{B}$ », de E dans $\mathbf{K}^n$, qui envoie un vecteur v sur le vecteur de ses coordonnées, est un isomorphisme d'espace vectoriel.*

Démonstration. — Si u est de coordonnées $\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}$ et v de coordonnées $\begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}$, alors $u + \lambda \cdot v = \sum_1^n (\alpha_i + \lambda\beta_i) \cdot u_i$ est de coordonnées

$$\begin{pmatrix}\alpha_1 + \lambda\beta_1\\ \vdots \\ \alpha_n + \lambda\beta_n\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} + \lambda \cdot \begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}.$$

L'application est donc bien linéaire.

Elle est surjective : si $\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_n\end{pmatrix}$ est un vecteur de $\mathbf{K}^n$, c'est l'image du vecteur $\sum_1^n \lambda_i \cdot u_i$. Elle est injective : son noyau est l'ensemble des vecteurs de coordonnées nulles ; il est donc réduit à $\left\{\begin{pmatrix}0\\ \vdots \\ 0\end{pmatrix}\right\}$ □

10.4.3. Dimension. —

Définition 10.4.11. — Un $\mathbf{K}$-espace vectoriel est dit de dimension finie s'il admet une famille génératrice finie $(v_1, \ldots, v_n)$.

Dans le cas du plan $\mathbf{R}^2$, on se convainc qu'une base est toujours composée de deux vecteurs non colinéaires. On va montrer que c'est toujours vrai : toutes les bases ont même cardinal. Rappelons que le cardinal d'une famille est le nombre de vecteurs qui la composent. Le but de ce paragraphe est en effet de montrer le théorème fondamental suivant :

Théorème 10.4.12. — *Soit E un espace vectoriel de dimension finie. Alors :*

1. *toutes les bases ont même cardinal. On appelle ce cardinal la* dimension *de E, noté* $\dim(E)$ *;*
2. *de plus toute famille libre est de cardinal $p \leq \dim(E)$ et, si elle est de cardinal* $\dim(E)$*, c'est une base ;*
3. *enfin toute famille génératrice est de cardinal $p \geq \dim(E)$ et, si elle est de cardinal* $\dim(E)$*, c'est une base.*

(QC)

Exemple 10.4.13. —
- Les $\mathbf{R}$-ev $\mathbf{R}^n$ sont de dimension n : on connaît la base canonique.
- L'espace $\mathbf{C}_d[X]$ des polynômes de $\mathbf{C}[X]$ de degré inférieur ou égal à d est un sous-ev, de dimension $d+1$. Une base est $1, X, \ldots, X^d$.

Commençons par un lemme disant que si une famille a plus de vecteurs qu'une base, alors elle est liée :

Lemme 10.4.14. — *Soit $(u_1, \ldots, u_n)$ une base. Alors toute famille $(v_1, \ldots, v_p)$ avec $p > n$ est liée.*

Une autre façon de le dire (en contraposant l'implication) est que toute famille libre est de cardinal inférieur ou égal à n. Ce lemme est le cœur du théorème précédent.

Démonstration. — Considérons C_1, …, C_p les vecteurs de coordonnées des v_i dans la base des u_j. Ce sont des vecteurs colonnes de $\mathbf{R}^n$.

Remarquons que si nous montrons que les vecteurs C_1, …, C_p sont liés, alors v_1, …, v_p sont eux-mêmes liés : c'est la proposition 10.4.7. De plus, on a vu que montrer que ces vecteurs sont liés revient à trouver une solution non nulle au système $AX = 0$, où A est la matrice de taille $n \times p$ dont les colonnes sont les C_i.

Mais on sait que $n < p$, donc ce système a plus d'inconnues que d'équations : il y a forcément des variables libres, donc forcément des solutions non nulles. □

On peut se lancer dans la preuve du théorème :

Démonstration. —
1. Soient $\mathcal{B} = (u_1, \ldots, u_n)$ et $\mathcal{B}' = (v_1, \ldots, v_p)$ deux bases. On veut montrer que $n = p$. Pour ça on utilise deux fois le lemme précédent : d'une part $\mathcal{B}$ est une base et $\mathcal{B}'$ est libre, donc le lemme précédent donne que $p \leq n$. D'autre part, $\mathcal{B}'$ est une base et $\mathcal{B}$ est libre, donc le lemme précédent donne que $n \leq p$. Au final toutes les bases ont même cardinal $n = \dim(E)$.
2. Soit $(v_1, \ldots, v_p)$ une famille génératrice. On sait qu'on peut en extraire une base, qui sera composée de n vecteurs d'après le premier point. Donc d'une part $p \geq n$, d'autre part si $p = n$, on n'a pas besoin d'extraire : la famille est déjà une base.
3. Une famille libre $(u_1, \ldots, u_p)$ est de cardinal inférieur au cardinal de n'importe quelle base comme le prouve le lemme précédent. Or on a une base de cardinal n. Ainsi $p \leq n$.

Enfin si une famille libre est de cardinal n, alors elle est génératrice. En effet, dans le cas contraire, on utilise le lemme d'extension des familles libres pour rajouter un vecteur et obtenir une famille libre de cardinal $n+1$, ce qui est impossible.

□

On peut toujours rajouter des vecteurs à une famille libre pour construire une base :

***Théorème 10.4.15* (Théorème de la base incomplète)**
Toute famille libre peut-être complétée en une base. (QC)

Démonstration. — Par récurrence descendante sur le cardinal de la famille libre : si c'est n, alors c'est une base. On suppose que le théorème est vrai pour toute famille libre de cardinal $p+1 \leq n$. Soit $(u_1, \dots u_p)$ une famille libre. Comme $p < n$, ce n'est pas une base. Soit donc u qui n'est pas dans l'espace vectoriel engendré par les (u_i). D'après le lemme d'extension des familles libres, la famille $(u, u_1, \dots, u_p)$ est encore libre. On conclut par hypothèse de récurrence. □

Proposition 10.4.16. — *Soient E un $\mathbf{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel. Alors F est de dimension finie et* $\dim(F) \leq \dim(E)$.

Démonstration. — Une famille libre de F est aussi une famille libre de E. Donc son cardinal est inférieur à $\dim(E)$. Soit $(u_1, \dots, u_n)$ une famille libre de F de cardinal maximal. On a donc $n \leq \dim(E)$. Montrons que cette famille engendre F : sinon, on utilise le lemme d'extension des familles libres pour construire une famille libre de F de cardinal $n+1$. C'est impossible. Donc $(u_1, \dots, u_n)$ est une base de F et $\dim(F) = n \leq \dim(E)$. □

10.5. Le rang

10.5.1. Cas des applications linéaires. —

Proposition 10.5.1. — *Soient E un $\mathbf{K}$-espace vectoriel de dimension n, E' un $\mathbf{K}$-espace vectoriel de dimension p et $f : E \to E'$ une application linéaire.* (QC)

Alors $\mathrm{im}(f)$ *est de dimension finie ; on appelle rang de f, noté* $\mathrm{rg}(f)$*, cette dimension. On a alors* $\mathrm{rg}(f) \leq \min(n, p)$.

Démonstration. — Soit e_1, …, e_n une base de E. Tout élément de l'image de f s'écrit donc $f(x_1e_1 + \dots x_ne_n) = x_1f(e_1) + \dots x_nf(e_n)$. Autrement dit, $\mathrm{im}(f) = \mathrm{Vect}(f(e_1), \dots, f(e_n))$ et donc est de dimension finie [9] inférieure à n.

De plus, $\mathrm{im}(f)$ est un sous-espace vectoriel de E'. La dimension de $\mathrm{im}(f)$ est donc bien inférieure à p. □

On conclut cette section par le théorème du rang et une conséquence :

(QC)

***Théorème 10.5.2* (Théorème du rang).** — *Soient E un $\mathbf{K}$-espace vectoriel de dimension n, E' un $\mathbf{K}$-espace vectoriel de dimension p et $f : E \to E'$ une application linéaire.*

Alors on a : $\mathrm{rg}(f) + \dim(\ker(f)) = n$.

Démonstration. — Le noyau $\ker(f)$ est de dimension finie $d \leq n$, car c'est un sous-espace vectoriel de E qui est de dimension n. Soit $(v_1, \dots, v_d)$ une base de $\ker(f)$. Complétons-la, grâce au théorème de la base incomplète, en une base $(v_1, \dots v_d, v_{d+1}, \dots, v_n)$ de E. Montrons que la famille $f(v_{d+1})$, …, $f(v_n)$ forme une base de $\mathrm{im}(f)$.

Tout d'abord, c'est une famille génératrice : en effet, un vecteur v de $\mathrm{im}(f)$ s'écrit par définition $v = f(u)$. On écrit $u = x_1v_1 + \dots + x_nv_n$, et on obtient $v = x_1f(v_1) + \dots + x_nf(v_n)$. Or les vecteurs v_1, …, v_d sont dans le noyau. Donc $v = f(u) = x_{d+1}f(v_{d+1}) + \dots + x_nf(v_n)$. La famille est bien génératrice.

De plus c'est une famille libre : supposons qu'on ait des scalaires λ_{d+1}, …, λ_n tels que $\lambda_{d+1}f(v_{d+1}) + \dots + \lambda_nf(v_n) = 0$; montrons que ces scalaires sont nuls. Pour ça, on réécrit l'égalité précédente $f(\lambda_{d+1}v_{d+1} + \dots + \lambda_nv_n) = 0$, par linéarité de f. Autrement dit, $\lambda_{d+1}v_{d+1} + \dots + \lambda_nv_n$ est un vecteur de $\ker(f)$. Comme v_1, …, v_d est une base de $\ker(f)$, on peut écrire $\lambda_{d+1}v_{d+1} + \dots + \lambda_nv_n = \lambda_1v_1 + \dots \lambda_dv_d$. Autrement dit, on a une relation de liaison entre les v_i. Or ils forment une base, donc notamment sont libres. Donc tous les coefficients sont nuls. Notamment, $\lambda_{d+1} = \dots = \lambda_n = 0$. La famille est bien libre.

Ainsi, on a une base de $\mathrm{im}(f)$ de cardinal $n - d$. On obtient bien :

$$\mathrm{rg}(f) + \dim(\ker(f)) = (n - d) + d = n.$$

□

On en déduit un point crucial de l'algèbre linéaire :

9. On peut remarquer que pour ça, on n'a pas utilisé que E' est de dimension finie.

Proposition 10.5.3. — *Soient E et E' deux $\mathbf{K}$-espaces vectoriels* de même dimension n *et* $f : E \to E'$ *une application linéaire.* (QC)

Alors f est injective $\Leftrightarrow$ f est surjective $\Leftrightarrow$ f est bijective.

Démonstration. — On a la suite d'équivalence :

$$\begin{aligned} f \text{ est injective} &\Leftrightarrow \ker(f) = \{0\} \\ &\Leftrightarrow \dim(\ker(f)) = 0 \\ &\Leftrightarrow \mathrm{rg}(f) = n \text{ (grâce au théorème du rang)} \\ &\Leftrightarrow \mathrm{im}(f) = E' \\ &\Leftrightarrow f \text{ est surjective.} \end{aligned}$$

Ainsi, f surjective implique f injective et surjective, donc f bijective. Et bien sûr, f bijective implique f surjective. On a bien toutes les équivalences voulues. □

10.5.2. Cas des matrices. — On définit pour les matrices ce qu'on a défini pour les applications linéaires. On note $(e_j)_{1\leq j\leq n}$ la base canonique de $\mathbf{K}^n$.

Définition 10.5.4. — On appelle noyau et image d'une matrice $A \in \mathcal{M}_{pn}(\mathbf{K})$ le noyau et l'image de l'application linéaire associée f_A ; on les note $\ker(A)$ et $\mathrm{im}(A)$. Ce sont des sous-espaces vectoriels, respectivement de $\mathbf{K}^n$ et $\mathbf{K}^p$. (QC)

On peut comprendre un peu mieux cette application linéaire.

Proposition 10.5.5. — *Soit $A \in \mathcal{M}_{pn}(\mathbf{K})$. Notons C_1, …, C_n les colonnes de A. Alors on a pour tout j :*

$$f_A(e_j) = C_j.$$

Démonstration. — $f_A(e_j)$ est le produit Ae_j. C'est donc un vecteur colonne à p coordonnées. Pour $1 \leq i \leq p$, la i-ème coordonnée est $L_i e_j$, où $L_i = (a_{i1}, \ldots, a_{in})$ est la i-ème ligne de A. Par définition du produit et de e_j, on a $L_i e_j = a_{ij}$. Donc les coordonnées de Ae_j sont les coefficients apparaissant dans la j-ème colonne de A : on a bien $Ae_j = C_j$. □

On en déduit l'expression générale de l'image d'un vecteur :

Proposition 10.5.6. — *L'image du vecteur* $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *de* $\mathbf{K}^n$ *est :*

$$f_A(v) = x_1 C_1 + \ldots + x_n C_n.$$

Démonstration. — En effet, on peut écrire $v = x_1e_1 + \ldots + x_ne_n$. Par linéarité, on obtient $f_A(v) = x_1f_A(e_1) + \ldots + x_nf_A(e_n)$. La proposition précédente permet de conclure : $f_A(v) = x_1C_1 + \ldots + x_nC_n$. □

Exemple 10.5.7. — Grâce à ces propriétés, on peut déterminer l'application linéaire associée la matrice identité $I_n \in \mathcal{M}_n(\mathbf{K})$: c'est l'identité de $\mathbf{K}^n$.

La proposition suivante permet de définir le rang d'une matrice, notion qui sera importante plus loin.

Proposition 10.5.8. — *Soit $A \in \mathcal{M}_{pn}(\mathbf{K})$. Alors* $\mathrm{im}(A)$ *est un espace vectoriel de dimension finie engendré par les colonnes de la matrice ; on appelle rang de A, noté* $\mathrm{rg}(A)$*, cette dimension. On a alors* $\mathrm{rg}(A) \leq \min(n, p)$.

(QC)

Cette proposition est la conséquence directe de la proposition analogue pour les applications linéaires. On a aussi l'analogue matriciel du théorème du rang :

(QC)

Théorème 10.5.9. — *Soit $A \in \mathcal{M}_{pn}(\mathbf{K})$. Alors on a :*

$$\mathrm{rg}(A) + \dim(\ker(A)) = n.$$