

Laurent BERNIS

# COURS DE MATHÉMATIQUES MP\*

## Chapitre I

# LES PROBABILITÉS, RÉVISIONS DE MPSI

Le présent chapitre couvre le programme de première année en probabilité. Le cours complet, est émaillé d'exercices qui en assurent l'assimilation.

## INTRODUCTION

Les probabilités se distinguent du reste des mathématiques en ce qu'elles consistent non seulement en une théorie mathématique, mais encore en une activité de modélisation, activité que les autres domaines des mathématiques une fois fondés et développés ont rejetée en dehors de leurs préoccupations, se coupant ainsi des activités scientifiques qui les ont parfois fait naître, et qu'ils abandonnent volontiers aux autres sciences. Ainsi l'étude des équations différentielles mène-t-elle une existence autonome, indépendamment des problèmes physiques qui l'alimentent. Cette situation s'explique aisément, puisque c'est précisément l'impuissance de la science à apporter une réponse prédictive à un phénomène qui fait que l'on recourt aux probabilités. Les lois de la mécanique ne sauraient prédire le résultat d'un lancé de dé, puisqu'il dépend de façon sûrement non continue de conditions initiales nombreuses et inconnues. Bref, les probabilités apparaissent là où la science renonce à exercer son pouvoir, et le travail de modélisation, peu conséquent et très variable d'une situation à l'autre, a atterri naturellement dans les mains des mathématiciens.

De cette double essence, les probabilités héritent d'une terminologie propre qui se réfère au contexte concret qu'elles modélisent qui se substitue au langage mathématique, ainsi là où un mathématicien parlerait de partie d'un ensemble, de complémentaire, le probabiliste use des termes *événement*, *événement contraire*. Cette contamination langagière constitue la principale difficulté dans l'abord de cette discipline.

L'objet des probabilités est d'attribuer à un événement qui n'est jugé probable d'aucune science un nombre compris entre 0 et 1 et qui est sensé traduire la fréquence d'apparition de l'événement si l'on répétait un grand nombre de fois les circonstances qui l'ont produit. Dans les sciences on part d'une circonstance connue avec précision (conditions initiales) et on en déduit avec certitude la situation qui en résultera. La position et la vitesse d'un satellite en un instant donné fournissent sa position future. Les probabilités partent d'un événement précisément connu et lui accorde une plausibilité plus ou moins grande, sa probabilité ; ce sont cette fois les conditions initiales qui échappent à notre connaissances, ou qui sont tellement complexes qu'elles en deviennent inexploitables : lorsque on jette un dé on sait seulement que le dé roulera s'arrêtera sur une face. La probabilité d'un événement se veut proportionnelle à la fréquence de réalisation du dit événement si l'on reproduisait un grand nombre de fois les circonstances qui l'ont créé, circonstances dont les détails échappent à notre savoir et conduisent à chaque expérience à une issue *a priori* différente. Ce faisant on suppose implicitement qu'il y a une certaine régularité dans l'apparition de l'événement.

Les probabilités sont nées au XVII<sup>e</sup> siècle, à la suite des travaux de Blaise Pascal, Pierre de Fermat et Christian Huygens motivés par les jeux de hasard. Le terme *probabilité* prend peu à peu son sens actuel avec les développements du traitement mathématique du sujet par Jakob Bernoulli. Il faut attendre le début du XX<sup>e</sup> siècle et les travaux de Kolmogorov, pour voir se développer une théorie axiomatisée des probabilités, théorie qui se fonde dans la théorie de l'intégration de Lebesgue.

Pour ce qui nous concerne, obéissant au programme nous traiterons des probabilités finies c'est-à-dire pour lesquels une expérience conduit à un nombre fini de résultats possibles. Dans ce cadre la part de la modélisation est importante, celle du traitement mathématique reste fort élémentaire.

On signalera enfin les deux jolis ouvrages [?] et [?] dont le début est abordable.

# PROBABILITÉ, ESPACE PROBABILISÉ

Nous allons commencer par décrire clairement le cadre des probabilités et en fixer la terminologie. Nous ne séparerons pas la présentation des définitions premières du rôle qu'elles jouent dans la modélisation d'une situation concrète.

## 2.1 Expérience aléatoire et univers

Notre but, comme nous l'avons dit dans l'introduction, est d'étudier des phénomènes concrets : jet de dé, tirage du loto, déplacement d'une microbille en suspension dans un fluide sous l'action des collisions avec les molécules du fluide, entre deux instants, prix à venir d'une action etc., phénomènes que les sciences sont impuissantes à prévoir et qui cependant présentent une certaine régularité dans la répétition. On parle pour un tel phénomène d'*expérience aléatoire* ou *épreuve aléatoire*. Pour modéliser l'ensemble des résultats d'une expérience aléatoire on recourt à un ensemble, appelé *univers* que la coutume désigne par  $\Omega$ . Cet ensemble est de natures diverses, donnons des exemples :

**Exemple 2.1.1.** —

1. Dans le cas d'un lancé d'un dé (à six face), l'univers est tout naturellement  $\{1, 2, 3, 6\}$  le résultat d'un lancé est représenté par le nombre marqué sur la face tournée vers le ciel.
2. Dans le cas d'un lancé de deux dés, l'univers est  $\{1, 2, 3, 6\}^2$  le résultat d'un lancé est représenté par le couple formé par le nombre indiqué par le premier dé, puis celui indiqué par le second.
3. Dans le cas de la distribution dans un jeu de 52 cartes d'une main de 5 cartes,  $\Omega$  est l'ensemble des parties à 5 éléments de l'ensemble  $\{1, \dots, 52\}$ , on a numéroté les cartes de 1 à 52.
4. On dispose de deux urnes contenant chacune  $N$  boules blanches ou noires, (le nombre de boules d'un couleur pouvant varier d'une urne à l'autre), on choisit une urne puis on tire une boule de l'urne choisie, ici  $\Omega$  est  $\{0, 1\} \times \{1, \dots, N\}$ , la première composante d'un élément de  $\Omega$  désigne le numéro de l'urne, la seconde celui attribué aux  $N$  boules de chaque urne.
5. On regarde la suite des lancés successifs d'une pièce,  $\Omega = \{P, F\}^{\mathbb{N}^*}$  le  $n^{\text{e}}$  terme d'une suite élément de  $\Omega$  est  $P$  si au  $n^{\text{e}}$  lancé on obtient pile,  $F$  sinon.
6. Dans le cas du déplacement entre deux instants d'une microbille livrée aux chocs des molécules d'un fluide,  $\Omega = \mathbb{R}^3$ .

**Tout univers est par convention non vide et dans la suite de ce chapitre on ne s'intéressera qu'à des univers  $\Omega$  finis.**

L'univers étant fixé, on peut étudier différents événements qui peuvent advenir suite à l'expérience aléatoire. la réalisation d'un événement peut se modéliser par une partie de  $\Omega$ . Par exemple dans le cadre du premier exemple 2.1.1., l'événement « le résultat du lancé est pair » est modélisé par la partie  $\{0, 2, 4, 6\}$ . Dans l'exemple 5, l'événement « les 3 premier lancés donnent le même résultat » se modélise par la partie  $\{(P, P, P, x_4, x_5, \dots) | (x_4, x_5, \dots) \in \{P, F\}^{\mathbb{N}}\} \cup \{(F, F, F, x_4, x_5, \dots) | (x_4, x_5, \dots) \in \{P, F\}^{\mathbb{N}}\}$ . Si la réalisation d'un premier événement se modélise par une partie  $A$ , de  $\Omega$  la réalisation d'un second par une partie  $B$ , alors la partie  $A \cup B$  modélise la réalisation du premier événement ou du second, la partie  $A \cap B$  modélise la réalisation du premier événement et du second, enfin le complémentaire de  $A$  modélise la réalisation de l'événement contraire au premier.

Pratiquement on a tendance à confondre la situation réelle modélisée et le modèle mathématique, si bien que le langage courant lui-même en vient à contaminer le langage mathématique, conduisant à la terminologie suivante :

**Terminologie 2.1.1.** —

- Les éléments de l'univers  $\Omega$  sont appelés *issues* ou *résultats possibles* ou *réalisations* ;
- Les parties de  $\Omega$  sont appelées *événements* ;
- On appelle *événement élémentaire* toute partie de  $\Omega$  qui est un singleton ;
- Si  $A$  et  $B$  sont des événements on appelle *événement  $A$  ou  $B$*  l'événement  $A \cup B$ , *événement  $A$  et  $B$*  l'événement  $A \cap B$  et l'*événement contraire de  $A$* , le complémentaire (dans  $\Omega$ ) de  $A$ , noté  $\bar{A}$  ;
- Des éléments  $A$  et  $B$  sont dit *incompatibles* si ils sont disjoints ( $A \cap B = \emptyset$ ) ;
- L'élément  $\Omega$  est dit *événement certain*, l'événement  $\emptyset$  est dit *événement impossible*.
- Un *système complet d'événements* est une famille de parties de  $\Omega$ ,  $(A_1, \dots, A_n)$  telle que  $\{A_1, \dots, A_n\}$  soit une partition de  $\Omega$ <sup>1</sup>.

La porosité entre le réel et le modèle fait qu'on pourra par exemple écrire à propos de l'exemple 2.1.1.-2 : « soit  $A$  l'événement "le résultat du lancé de dé est pair" ». Dans cette phrase  $A = \{2, 4, 6\}$  événement au sens mathématique est entièrement confondu avec l'événement réel qu'il décrit, à savoir le fait que le lancé de dé donne un

1. Rappelons que  $\{A_1, \dots, A_n\}$  est une partition de  $\Omega$  si par définition aucun des  $A_i$  n'est vide, les  $A_i$  sont deux à deux disjoints, et leur réunion est  $\Omega$ .

nombre pair. Cette négligence coupable est coutumière en probabilité et cette hérésie ne semble guère émouvoir ceux-là mêmes qui distinguent scrupuleusement le modèle mathématique du réel qu'il représente, dès lors qu'il s'agit d'autres sciences.

## 2.2 Probabilité

Nous en arrivons à l'objet essentiel de ce cours : définir la probabilité d'un événement. Soit un événement  $A$  quelconque. Il sera réalisé si l'issue de l'expérience est un élément de  $A$ . Par exemple dans le cadre de l'exemple 1 de 2.1.1 l'événement  $\{2, 4, 6\}$  qui modélise un score pair, est réalisé dès que l'issue de l'expérience est 2, 4, ou 6. On cherche à attribuer à  $A$  un nombre réel compris entre 0 et 1 qui mesure le degré de plausibilité de la réalisation de  $A$ , ce nombre étant d'autant plus proche de 1 que la réalisation de  $A$  est vraisemblable. Pour avoir une idée intuitive de ce que doit être ce nombre et de ce que seront ses propriétés on peut imaginer que l'on répète  $n$  fois la même expérience et que l'on observe la fréquence  $f_n(A)$  de la réalisation de  $A$  (nombre de fois où l'issue de l'expérience est élément de  $A$  divisé par  $n$ ). L'acte de fondation des probabilités est que  $f_n(A)$  tend en un certain sens vers une limite lorsque  $n$  tend vers  $+\infty$ , c'est cette limite que nous appellerons probabilité de l'événement  $A$  et noterons  $\mathbf{P}(\{A\})$ .

Les propriétés des fréquences d'un événement imposent alors immédiatement à l'application  $\mathbf{P}$  que nous cherchons à définir, les propriétés suivantes :

- i.  $0 \leq P(A) \leq 1$ .
- ii.  $\mathbf{P}(\Omega) = 1$ .
- iii.  $\mathbf{P}(\emptyset) = 0$ .
- iv. Si  $B$  est un élément tel que  $A \cap B = \emptyset$ , alors  $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ .
- v.  $\mathbf{P}(A) = \sum_{a \in A} \mathbf{P}(\{a\})$ .

Ces propriétés ne sont pas indépendantes les unes des autres. On pourrait pour une définition rigoureuse de la fonction  $\mathbf{P}$ , se donner la probabilité des événements élémentaires et supposer i., ii. et v., Cette démarche serait évidemment la plus naturelle et redonnerait iii. et iv. Mais une telle définition n'est pas généralisable au cas d'un univers  $\Omega$  de cardinal infini. Aussi allons-nous adopter la définition d'une probabilité sur  $\Omega$  suivante :

### Définition 2.2.1. — PROBABILITÉ —

On appelle probabilité sur un univers fini  $\Omega$ , toute application  $\mathbf{P}$  de  $\mathcal{P}(\Omega)$  dans  $[0, 1]$  qui jouit des propriétés suivantes

1.  $\mathbf{P}(\Omega) = 1$ .
2. Pour tout  $A$  et tout  $B$  parties disjointes de  $\Omega$  (événements incompatibles),

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B).$$

On appelle espace probabilisé fini tout couple  $(\mathbf{P}, \Omega)$  où  $\Omega$  est un univers fini et  $\mathbf{P}$  une probabilité sur  $\Omega$ .

Examinons les propriétés d'une probabilité

**Proposition 2.2.2.** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A_1, A_2, \dots, A_n$  des événements deux à deux incompatibles (disjoints si l'on préfère), alors

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbf{P}(A_i).$$

**Corollaire 2.2.3.** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A$  un événement (une partie si l'on veut) de  $\Omega$ . Alors

$$\mathbf{P}(A) = \sum_{x \in A} \mathbf{P}(\{x\}).$$

*Preuve de la proposition 2.2.2.* — Notons  $\mathbf{H}_n$  la propriété à prouver

- $\mathbf{H}_2$  est vraie par définition d'une probabilité.
- Soit un entier  $m \geq 2$ . On suppose  $\mathbf{H}_2$  vraie. Soient  $A_1, A_2, \dots, A_{m+1}$  des événements deux à deux incompatibles.

$$\left(\bigcup_{i=1}^n A_i\right) \cap A_{m+1} = \bigcup_{i=1}^n (A_i \cap A_{m+1}) = \bigcup_{i=1}^n \emptyset = \emptyset.$$

Donc par  $\mathbf{H}_2$ ,

$$\mathbf{P}\left(\bigcup_{i=1}^{m+1} A_i\right) = \mathbf{P}\left(\left(\bigcup_{i=1}^m A_i\right) \cup A_{m+1}\right) = \mathbf{P}\left(\bigcup_{i=1}^m A_i\right) + \mathbf{P}(A_{m+1}),$$

et donc d'après  $\mathbf{H}_m$ ,

$$\mathbf{P}\left(\bigcup_{i=1}^{m+1} A_i\right) = \left(\sum_{i=1}^m \mathbf{P}(A_i)\right) + A_{m+1} = \sum_{i=1}^{m+1} \mathbf{P}(A_i).$$

Voici  $\mathbf{H}_{m+1}$  prouvée.

Par récurrence nous venons de montrer la propriété  $\mathbf{H}_n$  pour tout entier  $n \geq 2$ . □

*Preuve du corollaire 2.2.3.* — L'univers  $\Omega$  étant fini il s'écrit  $\{x_1, x_2, \dots, x_n\}$  (où les  $x_i$  sont deux à deux distincts). Si  $n = 1$  alors il n'y a rien à prouver sinon, les événements élémentaires  $\{x_1\}, \{x_2\}, \dots, \{x_n\}$  étant incompatibles et leur réunion  $\Omega$ , 2.2.2. donne le résultat. □

Venons en aux propriétés d'une probabilité, dont certaines ont été intuitées dans l'approche heuristique par les fréquence.

Le corollaire 2.2.3 montre que la probabilité  $\mathbf{P}$  est définie par sa valeur sur les événements élémentaires. On a un peu mieux comme le montre l'exercice suivant.

**Exercice 2.2.4.** — Soit  $\Omega$  un ensemble fini. Soit  $(p_x)_{x \in \Omega}$  une famille de réels.

1. Montrer qu'il existe au plus une probabilité  $\mathbf{P}$  sur l'univers  $\Omega$  telle que, pour tout élément  $x$  de  $\Omega$ ,

$$\mathbf{P}(\{x\}) = p_x.$$

2. Montrer qu'il existe une probabilité  $\mathbf{P}$  sur  $\Omega$  telle que, pour tout élément  $x$  de  $\Omega$ ,

$$\mathbf{P}(\{x\}) = p_x$$

si et seulement si, pour tout  $x \in \Omega$ ,  $p_x \geq 0$  et  $\sum_{x \in \Omega} p_x = 1$ .

**Proposition 2.2.5.** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A$  et  $B$  des événements de  $\Omega$ .

1.  $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$ .
2.  $\mathbf{P}(\emptyset) = 0$ .
3.  $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$ .
4. Si  $A \subset B$  alors  $\mathbf{P}(A) \leq \mathbf{P}(B)$ .

**Remarque** — le premier point de cette proposition revêt un caractère pratique important. Il est parfois, en effet, plus facile de calculer la probabilité de l'événement contraire que de l'événement lui même (cf. exercice 2.2.7.).

*Preuve de la proposition 2.2.5.* —

1. Les événements  $A$  et  $\bar{A}$  sont incompatibles donc d'après la définition d'une probabilité,  $1 = \mathbf{P}(\Omega) = \mathbf{P}(A \cup \bar{A}) = \mathbf{P}(A) + \mathbf{P}(\bar{A})$ , et donc

$$\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A).$$

2. Résulte du point 1 appliqué à  $\Omega$ .

3. Les événements  $A$  et  $B \cap \bar{A}$  sont disjoints et leur réunion est  $A \cup B$ , donc :

$$\mathbf{P}(A \cup B) = \mathbf{P}(A \cup (B \cap \bar{A})) = \mathbf{P}(A) + \mathbf{P}(B \cap \bar{A}). \quad (\text{I.1})$$

D'autre part les événements  $B \cap \bar{A}$  et  $B \cap A$  sont disjoints de réunion  $B$ , donc :

$$\mathbf{P}(B) = \mathbf{P}((B \cap \bar{A}) \cup (B \cap A)) = \mathbf{P}(B \cap \bar{A}) + \mathbf{P}(B \cap A).. \quad (\text{I.2})$$

Finalement grâce à (I.1) et (I.2),

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(B \cap \bar{A})$$

4. Si  $A \subset B$ , alors  $A \cup B = B$  et (I.1) s'écrit :

$$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B \cap \bar{A}) \geq \mathbf{P}(A)$$

Ainsi s'achève la preuve de 2.2.4. □

Lors de la modélisation d'une expérience aléatoire réelle, le choix de la probabilité  $\mathbf{P}$  peut être plus ou moins expérimentale, il peut aussi se fonder sur les symétries du problème. Ainsi lorsque l'expérience est le jet d'un dé, la symétrie même du dé fait qu'il n'y a pas plus de raison qu'il fasse un chiffre plutôt qu'un autre, dans un jeu de carte, une main n'a pas plus de raison de sortir qu'une autre. Pour ce genre d'expérience aléatoire on introduit la probabilité uniforme

qui accorde à chaque événement élémentaire la même probabilité. C'est par défaut la probabilité que l'on considère sans indication particulière.

**Proposition-définition 2.2.6.** — **PROBABILITÉ UNIFORME** — *Soit  $\Omega$  un ensemble fini. Il existe une et une seule probabilité  $\mathbf{P}$  sur l'univers  $\Omega$  qui soit constante sur les événements élémentaires.*

Elle est donnée par :

$$\mathbf{P}(A) = \frac{\text{card}A}{\text{card}\Omega},$$

pour toute partie  $A$  de  $\Omega$ .

Elle vérifie :

$$\mathbf{P}(\{x\}) = \frac{1}{\text{card}\Omega},$$

pour tout élément  $x$  de  $\Omega$ .

*Preuve de la proposition 2.2.6.* —

- Supposons que  $\mathbf{P}$  soit une probabilité sur l'univers  $\Omega$ , constante sur les événements élémentaires, de valeur  $p$ , alors comme  $\Omega$  est la réunion disjointe de tous ses événements élémentaires, d'après le corollaire 2.2.3.,

$$1 = \mathbf{P}(\Omega) = \sum_{x \in \Omega} \mathbf{P}(\{x\}) = \sum_{x \in \Omega} p = \text{card}(\Omega) p.$$

Soit  $p = \frac{1}{\text{card}(\Omega)}$ . Toujours d'après 2.2.3. on a aussi, pour toute partie  $A$  de  $\Omega$

$$\mathbf{P}(A) = \sum_{x \in A} \mathbf{P}(\{x\}) = \sum_{x \in A} p = \text{card}(A) p = \frac{\text{card}(A)}{\text{card}(\Omega)}. \quad (\text{I.3})$$

Il existe donc au plus une probabilité constante sur les singletons de  $\Omega$ .

- Réciproquement l'application

$$\mathbf{P} : \mathcal{P}(\Omega) \rightarrow \mathbf{R}; A \mapsto \frac{\text{card}(A)}{\text{card}(\Omega)},$$

est une probabilité. En effet :

- Les propriétés des cardinaux assurent que  $\mathbf{P}$  est à valeurs dans  $[0, 1]$  ;
- $\mathbf{P}(\Omega) = \frac{\text{card}(\Omega)}{\text{card}(\Omega)} = 1$  ;
- Soit enfin  $A$  et  $B$  des événements disjoints,

$$\mathbf{P}(A \cup B) = \frac{\text{card}(A \cup B)}{\text{card}(\Omega)} = \frac{\text{card}(A) + \text{card}(B)}{\text{card}(\Omega)} = \frac{\text{card}(A)}{\text{card}(\Omega)} + \frac{\text{card}(B)}{\text{card}(\Omega)} = \mathbf{P}(A) + \mathbf{P}(B).$$

Il existe donc une et une seule probabilité  $\mathbf{P}$  sur l'univers  $\Omega$  qui soit constante sur les événements élémentaires. On a montré dans la preuve de l'unicité qu'elle vérifiait les propriétés annoncées.  $\square$

**Exercices 2.2.7.** —

1. Soit  $(\Omega, p)$  un espace probabilisé fini. Soient  $A$  et  $B$  des parties de  $\Omega$  telles que  $\mathbf{P}(A) = \frac{3}{4}$  et  $\mathbf{P}(B) = \frac{1}{3}$ . Montrer que  $\frac{1}{12} \leq \mathbf{P}(A \cap B) \leq \frac{1}{3}$ .
2. Soit un entier  $n \geq 2$ .
  - (a) Quelle est la probabilité qu'une famille à  $n$  enfants soit constituée d'enfants des deux sexes ?
  - (b) Quelle est la probabilité pour qu'une famille ait au plus une fille ?
  - (c) On désigne par  $A$  l'événement décrit dans la première question et par  $B$  celui décrit dans la seconde. Comparez  $\mathbf{P}(A \cap B)$  et  $\mathbf{P}(A)\mathbf{P}(B)$ .

Réponse :  $1 - \frac{1}{2^{n-1}}$  ;  $\frac{n+1}{2^n}$  ;  $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$  si et seulement si  $n = 3$ .

3. Soit un entier  $n \leq 365$ . On considère une classe de  $n$  élèves nés une même année (non bisextile). Quelle est la probabilité qu'au moins deux étudiants soit nés le même jour ?

Réponse :  $1 - \frac{A_{365}^n}{365^n}$ .

4. Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A$ ,  $B$ , et  $C$  des parties de  $\Omega$ . Montrer que

$$\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C).$$

Pour une généralisation (Formule de Poincaré) voir exercices.

## 2.3 Conditionnement

On se pose maintenant la question suivante : Dans une expérience aléatoire lorsque l'on sait qu'un événement  $B$  est réalisé comment mesurer les chances qu'un élément  $A$  le soit ? Intuitivement on conçoit bien que la réalisation de  $B$  change les chances de voir  $A$  se réaliser, par exemple dans le cas de deux jets successifs d'un dé l'événement  $A$  : « la somme des deux jets est supérieure ou égale à 10 » est modifié par la réalisation de l'événement  $B$  « le premier jet donne  $k$  ». Si  $k$  est inférieur ou égal à 3, on voit bien qu'il n'y a aucune chance que  $A$  soit réalisé, dans le cas contraire on sent bien intuitivement que plus  $k$  est grand plus augmentent les chances de voir  $A$  se réaliser. Pour définir la plausibilité de  $A$  sachant que  $B$  est réalisé, reprenons l'approche déjà faite pour définir une probabilité en termes de fréquences. Effectuons  $n$  expériences successives, afin de voir la plausibilité de  $A$  sachant  $B$  réalisé, regardons la fréquence d'apparition de  $A$  lorsque  $B$  est réalisé, c'est-à-dire le rapport du nombre d'expériences à l'issue desquelles  $A$  et  $B$  sont réalisés et du nombre d'expériences à l'issue desquelles  $B$  l'est. Avec les notations du 2.2., ce rapport vaut

$$\frac{nf_n(A \cap B)}{nf_n(B)} = \frac{f_n(A \cap B)}{f_n(B)}$$

Par un passage à la limite, acte fondateur des probabilités, cette quantité tend vers :

$$\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

C'est ce rapport qui va définir la probabilité d'avoir  $A$  sachant  $B$  réalisé.

**Définition 2.3.1.** — Soit  $(\Omega, \mathbf{P})$  un espace probabilisé fini,  $B$  une partie de  $\Omega$  telle que  $\mathbf{P}(B) \neq 0$ . Pour toute partie  $A$  de  $\Omega$ , on appelle probabilité conditionnelle de l'événement  $A$  sachant l'événement  $B$  et on note  $\mathbf{P}(A|B)$  le réel :

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

On retiendra

$$\boxed{\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\mathbf{P}(B)} \quad (\text{I.4})$$

**Proposition 2.3.2.** — Soit  $(\Omega, \mathbf{P})$  un espace probabilisé fini,  $B$  une partie de  $\Omega$  telle que  $\mathbf{P}(B) \neq 0$ , la probabilité conditionnelle sachant  $B$ ,

$$\mathbf{P}(\cdot|B) : \mathcal{P}(\Omega) \rightarrow \mathbf{R}; A \mapsto \mathbf{P}(A|B)$$

est une probabilité sur  $\Omega$ . On la note souvent  $\mathbf{P}_B$ .

*Preuve de la proposition 2.3.2.* —

- Comme  $\mathbf{P}$  est à valeurs positives,  $(\mathbf{P}(\cdot|B))$  l'est aussi et comme, pour toute partie  $A$  de  $\Omega$ ,  $(A \cap B) \subset B$ , la propriété 4 de 2.2.5., assure que  $(\mathbf{P}(\cdot|B))$  est à valeurs dans  $[0, 1]$ ;
- $\mathbf{P}(\Omega|B) = \frac{\mathbf{P}(\Omega \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B)}{\mathbf{P}(B)} = 1$ ;
- Enfin soient  $A_1$  et  $A_2$ , des parties de  $\Omega$  disjointes.

$$\begin{aligned} \mathbf{P}(A_1 \cup A_2|B) &= \frac{\mathbf{P}((A_1 \cup A_2) \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}((A_1 \cap B) \cup (A_2 \cap B))}{\mathbf{P}(B)} = \\ &= \frac{\mathbf{P}(A_1 \cap B) + \mathbf{P}(A_2 \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A_1 \cap B)}{\mathbf{P}(B)} + \frac{\mathbf{P}(A_2 \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A_1|B) + \mathbf{P}(A_2|B). \end{aligned}$$

Ainsi  $\mathbf{P}(\cdot|B)$  est-elle une probabilité. □

**Exercice d'application** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A$ ,  $B$  et  $C$  des événements de  $\Omega$ . On suppose  $\mathbf{P}(C) \neq 0$ . Montrer que  $\mathbf{P}(A \cup B|C) = \mathbf{P}(A|C) + \mathbf{P}(B|C) - \mathbf{P}(A \cap B|C)$ .

Nous allons maintenant généraliser (I.4).

**Proposition 2.3.3** — THÉORÈME DES PROBABILITÉS COMPOSÉES —

Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $n$  un entier supérieur ou égal à 2. Pour toute famille  $(A_k)_{k=1, \dots, n}$  de parties de  $\Omega$ , telle que  $\mathbf{P}\left(\bigcap_{k=1}^{n-1} A_k\right) \neq 0$ . Alors :

$$\mathbf{P}\left(\bigcap_{k=1}^n A_k\right) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2) \dots \mathbf{P}(A_{n-1}|A_1 \cap A_2 \dots \cap A_{n-2}) \mathbf{P}(A_n|A_1 \cap A_2 \dots \cap A_{n-1}).$$

*Preuve de 2.3.3.* — Observons que puisque  $\mathbf{P}\left(\bigcap_{k=1}^{n-1} A_k\right) \neq 0$ , d'après 2.2.5.-4,  $\mathbf{P}\left(\bigcap_{k=1}^k A_k\right) \neq 0$  pour  $k = 1, 2, \dots, n-1$ , ce qui confère un sens à la formule à prouver.

La propriété à prouver notée  $(H_n)$  se démontre sans surprises par récurrence sur  $n$ .

- La propriété  $(H_2)$  n'est autre que (I.4).
- Soit un entier  $m \geq 2$ . Supposons  $(H_m)$ . Considérons alors  $(A_k)_{k=1, \dots, m+1}$  une famille de parties de  $\Omega$ , telle que

$$\mathbf{P}\left(\bigcap_{k=1}^m A_k\right) \neq 0. \text{ Par (I.4);}$$

$$\mathbf{P}\left(\bigcap_{k=1}^{m+1} A_k\right) = \mathbf{P}(A_1 \cap A_2 \cdots \cap A_m) \mathbf{P}(A_{m+1} | A_1 \cap A_2 \cdots \cap A_m).$$

Donc, compte tenu de  $(H_m)$ ,

$$\mathbf{P}\left(\bigcap_{k=1}^n A_k\right) = \mathbf{P}(A_1) \mathbf{P}(A_2 | A_1) \dots \mathbf{P}(A_{m-1} | A_1 \cap A_2 \cdots \cap A_{m-2}) \mathbf{P}(A_m | A_1 \cap A_2 \cdots \cap A_{m-1}) \mathbf{P}(A_{m+1} | A_1 \cap A_2 \cdots \cap A_m).$$

Voici  $(H_m)$  prouvée.

Donc pour tout entier  $n \geq 2$ ,  $(H_n)$  est vraie. □

**Exercice 2.3.4.** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini,  $A$  et  $B$  des parties de  $\Omega$ . On suppose  $\mathbf{P}(A) \neq 0$ .

1. On suppose que pour tout  $x \in \Omega$ ,  $\mathbf{P}(\{x\}) \neq 0$ . Montrer que  $\mathbf{P}(B|A) = 1$  si et seulement si  $B \subset A$ . Si l'on ne suppose plus que pour tout  $x \in \Omega$ ,  $\mathbf{P}(\{x\}) \neq 0$ , qu'elle implication reste vraie dans l'implication précédente ?

**Proposition 2.3.5.** — FORMULE DES PROBABILITÉS TOTALES —

Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $(B_1, \dots, B_n)$  un système complet d'événements, tous de probabilité non nulle. Alors, pour toute partie  $A$  de  $\Omega$ ,

$$\mathbf{P}(A) = \sum_{i=1}^n \mathbf{P}(A|B_i) \mathbf{P}(B_i).$$

*Preuve de la proposition 2.3.5.* — Les  $A \cap B_i$ ,  $i = 1, \dots, n$  sont deux à deux distincts de réunion  $A$ . Donc :

$$\mathbf{P}(A) = \sum_{i=1}^n \mathbf{P}(A \cap B_i) = \sum_{i=1}^n \mathbf{P}(A|B_i) \mathbf{P}(B_i).$$

□

**Exemple 2.3.6.** — PROBLÈME DU CHEVALIER DE MÉRÉ —

Donnons à titre d'exemple un problème historique à l'origine du développement des probabilités. Il fut posé par le Chevalier de Méré à Blaise Pascal. Le voici. Deux joueurs jouent à un jeu de hasard (sans parties nulles), le gagnant et le premier à remporter trois parties. Les joueurs doivent mettre fin à leur rencontre avant la fin, à un moment où le premier joueur a gagné deux parties, le second une seule. Comment partager équitablement la mise entre les deux protagonistes ?

Dit en langage probabiliste, la question sous-entend que la part de la mise que doit empocher chaque joueur est proportionnelle à la probabilité qu'il a de gagner. Voyons ça. On peut faire comme si les joueurs jouaient quoiqu'il arrive encore deux parties, en considérant que si le premier joueur emporte la première de ces deux parties, comme il a gagné, la suivante est jouée pour le plaisir. Ceci permet de prendre comme univers  $\Omega = \{0, 1\}^2$ , un élément  $(x, y)$  de  $\Omega$  représente la situation où la première partie a été gagnée par le joueur  $x$  la suivante par le joueur  $y$ . On munira  $\Omega$  de la probabilité uniforme.

Remarquons que pour un tel univers à chaque partie les deux joueurs ont bien la même probabilité de gagner. En effet l'événement  $B_i$  modélisant « le joueur  $i$  emporte, disons la première partie, » est pour  $i = 1, 2$ ,  $\{B_i = \{(i, y), y \in \{1, 2\}\}\}$  ensemble de cardinal 2. Du reste nous aurions pu ne pas expliciter  $\Omega$  et nous contenter de cette propriété.

La famille  $(B_1, B_2)$  est un système complet d'événements. Soit l'événement  $A$  le premier joueur l'emporte la rencontre. La formule des probabilités totales dit :

$$\mathbf{P}(A) = \mathbf{P}(B_1) \mathbf{P}(A|B_1) + \mathbf{P}(B_2) \mathbf{P}(A|B_2).$$

Nous avons mentionné que  $\mathbf{P}(B_1) = \mathbf{P}(B_2) = \frac{1}{2}$ . Par ailleurs,  $\mathbf{P}(A|B_1) = 1$  (le duel voit quoi qu'il arrive la victoire du premier joueur), tandis que  $\mathbf{P}(A|B_2) = \frac{1}{2}$ , puisque le premier joueur gagne la rencontre si et seulement si il gagne la seconde partie.



Au total

$$\mathbf{P}(A) = \frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{2} = \frac{3}{4}.$$

Le premier joueur doit légitimement empocher les  $\frac{3}{4}$  de la mise.

On retiendra que sur cet exemple on n'a pas besoin réellement d'expliciter  $\Omega$ .

Passons à ce que l'on appelle la probabilité des causes.

**Proposition 2.3.7.** — FORMULE DE BAYES OU DE PROBABILITÉ DES CAUSES —

Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $(A_1, \dots, A_n)$  un système complet d'événements de probabilité non nulle. Alors pour toute partie  $B$  de  $\Omega$  de probabilité non nulle, et tout  $j \in \{1, \dots, n\}$

$$\mathbf{P}(A_j|B) = \frac{\mathbf{P}(B|A_j)\mathbf{P}(A_j)}{\sum_{i=1}^n \mathbf{P}(B|A_i)\mathbf{P}(A_i)}.$$

Souvent dans la pratique les événements  $A_i$  précèdent chronologiquement l'événement  $B$  et peuvent même être pensé comme des causes possibles de cette événement. La formule de Bayes permet de déterminer la probabilité de chaque cause possible, d'où son nom. D'un pure point de vue mathématique, il ne s'agit que d'un calcul de probabilité conditionnelle et l'aspect temporel n'a pas de réalité.

*Preuve de la proposition 2.3.7.* — Soient  $B$  une partie de  $\Omega$  de probabilité non nulle et  $j$  un élément de  $\{1, \dots, n\}$ . Par définition des probabilités conditionnelles,

$$\mathbf{P}(A_j|B)\mathbf{P}(B) = \mathbf{P}(A_j \cap B) = \mathbf{P}(B|A_j)\mathbf{P}(A_j),$$

cf. (I.4). En appliquant alors à  $B$  la formule des probabilités totales, on a bien le résultat.  $\square$

**Exemple 2.3.8.** — On reprend l'exemple 2.1.1.-4. On note  $b_i$  le nombre de boules blanches supposé non nul, dans l'urne numéro  $i$ ,  $n_i = N - b_i$  désigne le nombre de boules noires. On suppose avoir tiré une boule blanche. Quelle est *a posteriori* la probabilité que le tirage ait eu lieu dans la première urne? Autrement dit notons  $A_i$  l'événement « on a choisi pour le tirage l'urne numéro  $i$  » pour  $i=1,2$  et  $B$  l'événement « on a tiré une boule blanche » ; déterminons  $\mathbf{P}(A_1|B)$ .

$(A_1, A_2)$  est un système complet d'événements, la formule de Bayes affirme :

$$\mathbf{P}(A_1|B) = \frac{\mathbf{P}(B|A_1)\mathbf{P}(A_1)}{\mathbf{P}(B|A_1)\mathbf{P}(A_1) + \mathbf{P}(B|A_2)\mathbf{P}(A_2)}.$$

On peut munir l'univers  $\Omega$ , défini au 2.1.1., d'une probabilité uniforme, ce n'est pas ici l'essentiel, ce qui est primordial c'est que pour ce choix, la probabilité de tirer une urne ou une autre est la même et que la probabilité de tirer une boule dans chaque urne est uniforme. On aurait pu se limiter à cette hypothèse sans préciser ni  $\Omega$ , ni la probabilité dont on l'équipe, comme dans l'exemple 2.2.6.

Ainsi  $\mathbf{P}(B|A_i) = \frac{b_i}{N}$ , et  $\mathbf{P}(A_i) = \frac{1}{2}$ . Donc :

$$\mathbf{P}(A_1|B) = \frac{\frac{b_1}{N} \times \frac{1}{2}}{\frac{b_1}{N} \times \frac{1}{2} + \frac{b_2}{N} \times \frac{1}{2}} = \frac{b_1}{b_1 + b_2}.$$

**Exercice 2.3.9.** — Un client achète une ampoule dans un magasin. Dans ce magasin 30% des ampoules proviennent d'une usine  $F_1$  le reste d'une usine  $F_2$ . A la sortie de l'usine  $F_1$  2 % des ampoules sont défectueuses, seulement 1% à la sortie de l'usine  $F_2$ . Sachant que l'ampoule achetée par le client est défectueuse, quelle est la probabilité qu'elle ait été fabriquée par la première usine.

Réponse :  $\frac{6}{13}$ .

## 2.4 Événements indépendants

Commençons par une démarche heuristique. Cherchons à modéliser le fait que dans une expérience aléatoire des événements sont indépendants, c'est-à-dire que la réalisation de l'un ne modifie en rien les chances de réalisation de l'autre. Supposons la situation modélisée par un espace probabilisé fini  $(\Omega, \mathbf{P})$  et les deux événements par des parties  $A$  et  $B$  de  $\Omega$ . Admettons momentanément que ces événements sont de probabilités non nulles. Dire que les chances de réalisation du premier événement  $A$  ne sont pas modifiées par la réalisation du second événement se traduit naturellement par le fait que  $\mathbf{P}(A|B) = \mathbf{P}(A)$ . Ce qui s'écrit encore  $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$ . Cette dernière expression

est symétrique en  $A$  et  $B$  et va servir de définition à l'indépendance de deux événements (de probabilité non nulle ou pas).

**Définition 2.4.1.** — Soit  $(\Omega, \mathbf{P})$  un espace probabilisé fini. On dit que des événements (parties)  $A$  et  $B$  de  $\Omega$  sont indépendants si

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).$$

Soient un entier  $m \geq 2$  et  $A_1, A_2, \dots, A_m$  des événements de  $\Omega$ , on dit qu'ils sont deux à deux indépendants si, pour tout couple  $(i, j)$  d'éléments distincts de  $\{1, \dots, m\}$ ,

$$\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j).$$

De la définition d'une probabilité conditionnelle il vient immédiatement :

**Proposition 2.4.1.** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A$  et  $B$  des événements de  $\Omega$ . On suppose de surcroît que  $\mathbf{P}(B) \neq 0$ . Alors  $A$  et  $B$  sont indépendants si et seulement si  $\mathbf{P}(A|B) = \mathbf{P}(A)$ .

**Exemple 2.4.2.** —

- Reprenons l'exemple 2.1.1.-5 sur les  $N$  lancers successifs d'un dé. On regarde l'événement  $A$  « le premier lancé donne face » et l'événement  $B$  « le second lancé donne face ». Montrons qu'ils sont indépendants.

$$A = \{(F, X_2, X_3, \dots, X_N) | (X_2, X_3, \dots, X_N) \in \{P, F\}^{N-1}\},$$

$$B = \{(X_1, F, X_3, \dots, X_N) | (X_1, X_3, \dots, X_N) \in \{P, F\}^{N-1}\}.$$

Le cardinal de  $A$  et  $B$  est  $2^{N-1}$ . Donc  $\mathbf{P}(A) = \mathbf{P}(B) = \frac{2^{N-1}}{\text{card}(\Omega)} = \frac{2^{N-1}}{2^N} = \frac{1}{2}$ .

Par ailleurs  $A \cap B = \{(F, F, X_3, \dots, X_N) | (X_3, \dots, X_N) \in \{P, F\}^{N-2}\}$  et donc :

$$\mathbf{P}(A \cap B) = \frac{2^{N-2}}{2^N} = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = \mathbf{P}(A)\mathbf{P}(B).$$

Les événements  $A$  et  $B$  sont bien indépendants, ce qui est conforme à notre intuition le fait de faire face au premier lancé et de faire face au deuxième non aucun rapport l'un avec l'autre.

- Prenons à présent l'exemple 2.3.8 où l'on modélisait le choix d'une urne puis le tirage d'une boule dans l'urne choisie. Considérons l'événement  $B$  « la boule tirée est blanche » et l'événement  $U_1$  « l'urne choisie est la première ». Intuitivement on conçoit fort bien que la probabilité d'obtenir une boule blanche va dépendre du choix de l'urne si les deux urnes n'ont pas le même nombre de boules blanches. Voyons cela.

$$U_1 = \{(1, y), y \in \{1, \dots, N\}\}.$$

Supposons que dans les deux urnes nous ayons numéroté d'abord les boules blanches (dans l'urne numéro  $i$ , les boules blanches sont celles numérotées de 1 à  $n_i$ ). L'événement  $B$  vaut alors

$$B = \{(1, y), y \in \{1, n_1\}\} \cup \{(2, y), y \in \{1, n_2\}\}.$$

Donc,  $B$  étant écrit comme la réunion de deux événements incompatibles

$$\mathbf{P}(B) = \mathbf{P}(\{(1, y), y \in \{1, n_1\}\}) + \mathbf{P}(\{(2, y), y \in \{1, n_2\}\}) = \frac{n_1}{\text{card}\Omega} + \frac{n_2}{\text{card}\Omega} = \frac{n_1 + n_2}{2 \times N}; \quad \mathbf{P}(U_1) = \frac{1}{2}.$$

L'événement  $B \cap U_1 = \{(1, y), y \in \{1, n_1\}\}$ , donc  $\mathbf{P}(B \cap U_1) = \frac{n_1}{2N}$ .

Ainsi a-t-on :  $\mathbf{P}(B \cap U_1) = \mathbf{P}(B)\mathbf{P}(U_1)$  si et seulement si :

$$\frac{n_1}{2N} = \frac{1}{2} \times \frac{n_1 + n_2}{2 \times N},$$

soit si et seulement si  $\frac{n_1}{2} = \frac{n_1 + n_2}{2}$ . Finalement  $B$  et  $U_1$  sont indépendants si et seulement si  $n_1 = n_2$  ce qui est conforme à notre intuition.

**Exercices 2.4.3.** — Soit  $A$  et  $B$  des événements d'un espace probabilisé fini  $(\Omega, \mathbf{P})$ .

1. Montrer que si  $A$ , et  $B$  sont indépendants, alors  $A$  et  $\bar{B}$  le sont, ainsi que  $\bar{A}$  et  $\bar{B}$ .
2. On suppose que  $A \cap B = \emptyset$ . Montrer que  $A$  et  $B$  sont indépendants si et seulement si  $\mathbf{P}(A) = 0$  ou  $\mathbf{P}(B) = 0$ .
3. On suppose  $A$  et  $\bar{A}$  indépendants. Que dire de la probabilité de  $A$  ?

**Définition 2.4.4.** — soit  $(\Omega, \mathbf{P})$  un espace probabilisé fini. Soient un entier  $m \geq 2$  et  $A_1, A_2, \dots, A_m$  des événements de  $\Omega$ , on dit qu'ils sont (mutuellement) indépendants si, pour toute partie finie  $J$  de  $\{1, \dots, m\}$  non vide,

$$\mathbf{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbf{P}(A_j).$$

Pour vérifier concrètement la mutuelle indépendance de  $m$  éléments il faut vérifier  $2^m - m - 1$  conditions (le nombre de parties de  $\{1, \dots, m\}$  ayant strictement plus d'un élément).

Des événements de  $\Omega$ , mutuellement indépendants, sont deux à deux indépendants, il suffit pour le voir de prendre en particulier dans la définition de la mutuellement indépendance, les parties  $J$  à 2 éléments. En revanche la réciproque est fautive, comme le montre l'exemple suivant. Prenons  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$  et munissons cette ensemble de la probabilité uniforme. Posons  $A = \{\omega_1, \omega_2\}$ ,  $B = \{\omega_1, \omega_3\}$ ,  $C = \{\omega_1, \omega_4\}$ .

$$\mathbf{P}(A \cap B) = \mathbf{P}(A \cap C) = \mathbf{P}(B \cap C) = \mathbf{P}(\{\omega_1\}) = \frac{1}{4}.$$

Les événements  $A$ ,  $B$  et  $C$  sont donc bien deux à deux indépendants. En revanche :  $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(\{\omega_1\}) = \frac{1}{4}$  tandis que  $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \left(\frac{2}{4}\right)^3 = \frac{1}{8}$ .

Dans le même ordre d'idées, on peut avoir  $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$  sans que  $A, B, C$  soient indépendants deux à deux et donc *a fortiori* mutuellement indépendants, comme le lecteur pourra s'en convaincre en prenant  $A = B$ , avec  $\mathbf{P}(A) \in ]0, 1[$ .

**Exercice 2.4.5.** — Soient  $(\Omega, \mathbf{P})$  un espace probabilisé fini et  $A_1, \dots, A_n$  des événements de  $\Omega$ .

- Supposons les événements  $A_1, \dots, A_n$  (mutuellement) indépendants. Montrer que les événements  $A_1$  et  $\bigcap_{i=2}^n A_i$  sont indépendants.
- Montrer que  $A_1, \dots, A_n$  sont (mutuellement) indépendants si et seulement si, pour toute partie  $J$  non vide de  $\{1, \dots, n\}$  telle que  $\mathbf{P}\left(\bigcap_{j \in J} A_j\right) \neq 0$ , et tout élément  $i$  de  $\{1, \dots, n\}$ , si  $i \notin J$  alors :

$$\mathbf{P}\left(A_i \mid \bigcap_{j \in J} A_j\right) = \mathbf{P}(A_i).$$

Comparer avec 2.4.1.

## VARIABLES ALÉATOIRES

Lorsque l'on a modélisé une expérience aléatoire simple (par exemple par la loi uniforme), il est alors possible d'étudier des événements plus complexes qui sont fonctions des issues de notre expérience... C'est le rôle des variables aléatoires, qui sont, contrairement à ce que leur nom laisse accroire des fonctions définies sur l'univers. Elles vont nous permettre d'associer une probabilité mesurant des événements complexes et difficiles à modéliser.

### 3.1 Premières définitions

**Définition 3.1.1.** — Soit  $(\Omega, \mathbf{P})$  un espace probabilisé. On appelle variable aléatoire sur  $(\Omega, \mathcal{A}, \mathbf{P})$ , toute application  $X$  de  $\Omega$  dans un ensemble quelconque  $E$ .

Si de plus  $E$  est une partie de  $\mathbf{R}$ , on parle de variable aléatoire réelle.

**Exemples 3.1.2.** — On reprend l'espace probabilisé fini défini en 2.1.1-2 qui modélise le lancé de deux dés. On peut définir sur  $(\Omega, \mathcal{A}, \mathbf{P})$  les variables aléatoires suivantes :

- (a)  $X_1 : \{1, \dots, 6\}^2 \rightarrow \mathbf{R}; (x, y) \mapsto x$  (résultat du premier dé),  
 (b)  $X_2 : \{1, \dots, 6\}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \max\{x, y\}$  (résultat maximum des deux dés),  
 (c)  $X_3 : \{1, \dots, 6\}^2 \rightarrow \mathbf{R}; (x, y) \mapsto x + y$  (somme des résultats des deux dés),  
 (d)  $X_4 : \{1, \dots, 6\}^2 \rightarrow \mathbf{R}; (x, y) \mapsto 1$  si  $x \geq y$ , 2 sinon (numéro du dé affichant le plus grand nombre),
- On considère l'espace probabilisé fini  $(S_n, \mathbf{P})$  où  $S_n$  est le groupe symétrique à  $n$  éléments et  $p$  la probabilité uniforme sur  $S_n$ .

- (a) On définit sur cet espace probabilisé la variable aléatoire  $N$  qui à un élément  $\sigma$  de  $S_n$  associe le nombre de ses points fixes.
- (b) On peut aussi considérer la variable aléatoire  $M$  qui à un élément  $\sigma$  de  $S_n$  associe la matrice de permutation  $P_\sigma$ , c'est-à-dire l'élément de  $\mathcal{M}_n(\mathbf{R})$  dont la  $j^{\text{e}}$  colonne est le  $\sigma(j)^{\text{e}}$  vecteur de la base canonique de  $\mathcal{M}_{n,1}(\mathbf{R})$  (un 1 en  $\sigma(j)^{\text{e}}$  ligne, des 0 ailleurs).
3. Soient  $X_1, X_2, \dots, X_m$  des variables aléatoires définies sur un espace probabilisé  $(\Omega, \mathbf{P})$  à valeurs respectivement dans des espaces  $E_1, E_2, \dots, E_m$ , alors on dispose d'une variable aléatoire

$$Z : \Omega \rightarrow E_1 \times \dots \times E_m ; \omega \mapsto (X_1(\omega), X_2(\omega), \dots, X_m(\omega)).$$

On la notera abusivement  $(X_1, \dots, X_m)$ .

Dans ces exemples, les variables aléatoires  $M$  et  $Z$  ne sont pas réelles, les autres le sont.

**Notations** — Les probabilistes recourent sans vergogne pour les variables aléatoires à des notations abusives proscrites partout ailleurs. Voyons cela :

Soient  $X$  une variable aléatoire définie sur un espace probabilisé fini  $(\Omega, \mathbf{P})$ , à valeurs dans un ensemble  $E$ ,  $A$  et  $B$  des parties de  $E$  et  $a$  un élément de  $E$ .

NOTATIONS RIGOUREUSES	NOTATIONS PROBABILISTES
$X^{-1}(A)$ ou $\{\omega \in \Omega   X(\omega) \in A\}$	$\{X \in A\}$ ou $(X \in A)$
$X^{-1}(\{a\})$	$\{X = a\}$ ou $(X = a)$
Pour $E = \mathbf{R}$ , $X^{-1}([a, +\infty[)$	$\{X \geq a\}$ ou $(X \geq a)$
$\mathbf{P}(X^{-1}(A))$	$\mathbf{P}(X \in A)$
$\mathbf{P}(X^{-1}(\{a\}))$	$\mathbf{P}(X = a)$
Pour $E = \mathbf{R}$ , $\mathbf{P}(X^{-1}([a, +\infty[))$	$\mathbf{P}(X \geq a)$
$\vdots$	$\vdots$

**Définition 3.1.3.** — Soient  $X$  une variable aléatoire sur un espace probabilisé  $(\Omega, \mathbf{P})$  à valeurs dans  $E$  et  $f$  une application de  $E$  dans un ensemble  $F$ . L'application  $f \circ X$  définit alors une variable aléatoire sur  $(\Omega, \mathcal{A}, \mathbf{P})$  à valeurs dans  $F$ , on l'appelle image de  $X$  par  $f$  et on la note abusivement  $f(X)$ .

## 3.2 Loi d'une variable aléatoire

Nous allons maintenant associer à une variable aléatoire une probabilité définie sur son image.

**Proposition-définition 3.2.1.** — LOI D'UNE VARIABLE ALÉATOIRE —

Soit  $X$  une variable aléatoire définie sur un espace probabilisé fini  $(\Omega, \mathbf{P})$  à valeurs dans un ensemble  $E$ . Alors l'application

$$\mathbf{P}_X : \mathcal{P}(X(\Omega)) \rightarrow [0, 1] ; A \mapsto \mathbf{P}(X \in A)$$

est une probabilité sur  $X(\Omega)$ .

On appelle cette probabilité, loi de la variable aléatoire  $X$ .

*Preuve de la proposition 3.2.1.* —

Remarquons pour commencer que  $X(\Omega)$  est fini puisque  $\Omega$  l'est, ce, que  $E$  le soit ou non. Par ailleurs  $\mathbf{P}_X$  est bien à valeurs dans  $[0, 1]$  puisque  $\mathbf{P}$  l'est.

- $\{X \in X(\Omega)\} = \Omega$ , donc  $\mathbf{P}_X(X(\Omega)) = \mathbf{P}(\Omega) = 1$ .
- Soient  $A$  et  $B$  des parties disjointes de  $X(\Omega)$ .

D'une part, par définition de  $\mathbf{P}_X$ ,

$$\mathbf{P}_X(A \cup B) = \mathbf{P}(X \in A \cup B) = \mathbf{P}(X^{-1}(A \cup B)) = \mathbf{P}(X^{-1}(A) \cup X^{-1}(B)) = \mathbf{P}(\{X \in A\} \cup \{X \in B\}).$$

D'autre part,

$$\{X \in A\} \cap \{X \in B\} = X^{-1}(A) \cap X^{-1}(B) = X^{-1}(A \cap B) = X^{-1}(\emptyset) = \emptyset,$$

donc  $\{X \in A\}$  et  $\{X \in B\}$  sont deux événements de  $\Omega$  disjoints. Donc

$$\mathbf{P}_X(A \cup B) = \mathbf{P}(\{X \in A\}) + \mathbf{P}(\{X \in B\}) = \mathbf{P}_X(A) + \mathbf{P}_X(B).$$

De ces deux points il vient que  $\mathbf{P}_X$  est une probabilité. □

**Remarque :** On pourrait songer à prolonger  $\mathbf{P}_X$  en une probabilité définie sur  $E$ , par  $(\mathbf{P}_X(A) = \mathbf{P}(X = A)$  pour tout  $A$  dans  $E$ ) mais comme  $E$  n'est pas nécessairement fini, on sort du cadre du programme de première année et l'on verra cette année que la chose nécessite quelques précautions.

On écrira

Soient  $X$  une variable aléatoire définie sur un espace probabilisé fini  $(\Omega, \mathbf{P})$  à valeurs dans un ensemble  $E$  et  $f$  une application de  $E$  dans un ensemble  $F$ . Soit  $B$  une partie de  $f(X(\Omega))$ , alors :

$$\mathbf{P}_{f(X)}(B) = \mathbf{P}(f(X) \in B) = \mathbf{P}(X \in f^{-1}(B)) = \mathbf{P}_X(f^{-1}(B)).$$

Le corollaire 2.2.3. nous a appris qu'une probabilité sur un univers fini est entièrement définie par ses valeurs sur les singletons, il en résulte le résultat suivant.

**Proposition 3.2.2.** — *Soit  $X$  une variable aléatoire définie sur un espace probabilisé fini  $(\Omega, \mathbf{P})$  à valeurs dans un ensemble  $E$ . Alors la loi de  $X$  est entièrement déterminée par les valeurs  $\mathbf{P}(X = x)$ , où  $x$  décrit  $X(\Omega)$ .*

Plus précisément pour toute partie  $A$  de  $X(\Omega)$ ,

$$\mathbf{P}_X(A) = \sum_{x \in A} \mathbf{P}(X = x).$$

Nous allons donner maintenant des lois de variables aléatoires qui modélisent de nombreuses situations.

Nous supposons qu'est donné un espace probabilisé fini  $(\Omega, \mathbf{P})$ . En vertu de 3.2.2., nous définirons les lois des variables aléatoires  $X$  par les valeurs  $\mathbf{P}(X = x)$ .

LOI UNIFORME —

**Définition 3.2.3.** — *On dit qu'une variable aléatoire  $X$  à valeurs dans  $E$  ensemble fini suit la loi uniforme si, pour tout  $x \in E$ ,*

$$\mathbf{P}(X = x) = \frac{1}{|E|}$$

On note :

$$X \sim \mathcal{U}(E).$$

C'est la situation dégénérée où  $\mathbf{P}$  est la loi uniforme et  $X$  l'identité, mais pas que. Prenons pour  $\Omega$ , l'ensemble  $\{0, \dots, 99\}$  que l'on munit de la probabilité uniforme, et  $X$  la variable qui à un élément  $\omega$  de  $\Omega$  associe le dernier chiffre dans l'écriture décimale de  $\omega$ , ici  $E = \{0, \dots, 9\}$  et  $X \sim \mathcal{U}(E)$ , puisque pour  $i = 0, \dots, 9$ ,

$$\{X = i\} = \{i, \overline{1i}, \overline{2i}, \dots, \overline{9i}\},$$

et donc

$$\mathbf{P}_X(\{i\}) = \mathbf{P}(X = i) = \frac{|\{X = i\}|}{|E|} = \frac{10}{100} = \frac{1}{10} = \frac{1}{|E|}.$$

LOI DE BERNOULLI —

Commençons par un exemple. Le lancer d'une pièce non équilibrée conduit à deux issues possibles *pile* ou *face*, l'univers est donc  $\{P, F\}$ , la probabilité de  $\{P\}$  est prise égale à  $p$  élément de  $]0, 1[$ , donc celle de  $\{F\}$  à  $1 - p$ . La variable aléatoire  $X$  qui associe 1 à  $P$  et 0 à  $F$  suit la loi :

$$\mathbf{P}_X(\{1\}) = p, \mathbf{P}_X(\{0\}) = 1 - p.$$

De façon générale :

**Définition 3.2.4.** — *Soit  $p$  un élément de  $[0, 1]$ . On dit qu'une variable aléatoire  $X$  à valeurs dans  $\{0, 1\}$  suit la loi de Bernoulli de paramètre  $p$  si,  $\mathbf{P}(X = 1) = p$  (donc  $\mathbf{P}(X = 0) = 1 - p$ ) On note :*

$$X \sim \mathcal{B}(p).$$

Notons que  $X \sim \mathcal{B}(\frac{1}{2})$  ne signifie rien d'autre que  $X \sim \mathcal{U}(\{0, 1\})$ .

La loi de Bernoulli est très utile : dans toute expérience à deux issues, (pile ou face, victoire ou défaite...) la variable aléatoire qui renvoie 1 pour une issue, 0 pour l'autre, suit une loi de Bernoulli. Par abus on dit encore que la variable aléatoire qui renvoie le résultat de l'expérience (qui n'est autre que l'identité) suit une loi de Bernoulli, ce faisant on identifie implicitement les deux issues à 0 et 1.

Une variable aléatoire qui suit une loi de Bernoulli de paramètre  $p$  est dite simplement *variable de Bernoulli de paramètre  $p$* .

**Proposition 3.2.5** — Soit  $A$  un événement de  $(\Omega, \mathcal{A}, \mathbf{P})$  de probabilité  $p$ . La fonction indicatrice (caractéristique) de  $A$ ,  $\mathbf{1}_A$  est une variable aléatoire de Bernoulli de paramètre  $p$ .

*Preuve de 2.2.5.* — On a  $\mathbf{P}(\mathbf{1}_A = 1) = \mathbf{P}(A) = p$  et  $\mathbf{P}(\mathbf{1}_A = 0) = \mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A) = 1 - p$ .

## LOI BINOMIALE

Commençons là encore par un exemple.

On considère une urne contenant  $N$  boules blanches et  $M$  boules noires. On tire  $n$  boules avec remise. L'univers est bien sûr, après numérotation des boules,  $\{1, \dots, N + M\}^n$  et il est naturellement muni de la probabilité uniforme. On note  $X$  la variable aléatoire qui à un élément de  $\Omega$ , associe le nombre de boules blanches obtenues, mathématiquement le nombre de ses composantes qui sont le numéro d'une boule blanche. Soit  $k \in \{0, \dots, n\}$ . Etant donné une partie  $J$  à  $k$  éléments de  $\{1, \dots, n\}$  le nombre d'éléments  $(x_1, \dots, x_n)$  de  $\Omega$ , tels que pour tout  $i \in \{1, \dots, n\}$ ,  $x_i$  soit le numéro d'une boule blanche si  $j \in J$  et d'une boule noire sinon, vaut  $N^k M^{n-k}$ . Or il existe  $\binom{n}{k}$  parties de  $\{1, \dots, n\}$  à  $k$  éléments donc le cardinal de  $\{X = k\}$  est  $\binom{n}{k} N^k M^{n-k}$ . Comme le cardinal de  $\Omega$  est  $(N + M)^n$  on a :

$$\mathbf{P}(X = k) = \binom{n}{k} \left( \frac{N}{N + M} \right)^k \left( \frac{M}{N + M} \right)^{n-k}.$$

Posons  $p = \frac{N}{N + M}$ , alors :

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Plus généralement :

**Définition 3.2.6.** — Soit un entier  $n \geq 1$  et  $p$  un élément de  $[0, 1]$ . On dit qu'une variable aléatoire  $X$  à valeurs dans  $\{0, \dots, n\}$  suit la loi binomiale de taille  $n$  et de paramètre  $p$  (ou plus négligemment, de paramètres  $n$  et  $p$ ) si,

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

On note :

$$X \sim \mathcal{B}(n, p).$$

Une variable aléatoire qui suit une loi binomiale de paramètre  $n$  et  $p$  est dite *variable binomiale de paramètres  $n$  et  $p$* .

## Remarques —

1. On retrouve que  $\mathbf{P}_X(\{1, \dots, n\}) = 1$  grâce au binôme de Newton, qui donne son nom à la loi :

$$\mathbf{P}_X(\{1, \dots, n\}) = \sum_{k=0}^n \mathbf{P}(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1^n = 1.$$

2. La loi binomiale de paramètre 1 et  $p$  est la loi de Bernoulli de paramètre  $p$ .

## Exercice —

Soit  $X$  une variable aléatoire.

1. On suppose que  $X \sim \mathcal{U}(\{1, \dots, n\})$  identifier la loi de  $n - X$ .
2. On suppose que  $X \sim \mathcal{B}(n, p)$  identifier la loi de  $n - X$ .

**Exercice 3.2.7.** — Une urne contient  $N$  boules blanches et  $M$  boules noires. On tire  $n$  boules sans remise, donc  $n \leq N + M$ . On note  $X$  le nombre de boules blanches obtenues.

1. Décrire l'espace probabilisé  $(\Omega, \mathbf{P})$  qui modélise cette expérience aléatoire.
2. Soit  $k \in \{0, n\}$ . Déterminer  $\mathbf{P}(X = k)$ .
3. On suppose que  $N$  et  $M$  tendent vers  $+\infty$  de telle sorte que  $\frac{N}{N+M}$  tende vers une limite finie  $p$ . Montrer que  $\mathbf{P}(X = k)$  tend vers  $\binom{n}{k} p^k (1-p)^{n-k}$ . Interpréter ce résultat.

*Solution de l'exercice 3.2.7.*—

1. Comme on tire les boules sans remise et que l'on ne se soucie pas de l'ordre dans lequel elles ont été prélevées, il est loisible de considérer que l'on tire les  $n$  boules d'un coup et de prendre pour  $\Omega$  l'ensemble des parties à  $n$  éléments de  $\{1, \dots, N+M\}$ , son cardinal est  $\binom{N+M}{n}$ . On munit naturellement  $\Omega$  de la probabilité uniforme.
2. On dispose de  $X$  qui est une variable aléatoire sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . Si  $k > N$  ou  $n-k > M$  alors l'ensemble  $\{X = x\}$  est vide, sinon  $|\{X = x\}| = \binom{N}{k} \binom{M}{n-k}$ . Donc :

$$\mathbf{P}(X = k) = \begin{cases} \frac{\binom{N}{k} \binom{M}{n-k}}{\binom{N+M}{n}}, & \text{si } k \leq N \text{ et } n-k \leq M, \\ 0, & \text{sinon.} \end{cases}$$

3. Calcul élémentaire. Lorsque le nombre de boules blanches et noires devient très grand le fait de ne pas remettre ou de remettre les boules ne change pas grand chose, ce qui explique le résultat.

### 3.3 Couples et $n$ uplets de variables aléatoires

Dans tout ce paragraphe nous est donné un espace probabilisé fini  $(\Omega, \mathbf{P})$ . Toutes les variables aléatoires sont définies sur cet espace.

On a vu que si l'on dispose de deux variables aléatoires  $X_1$  et  $X_2$  à valeurs respectivement dans  $E_1$  et  $E_2$ , d'après 3.1.3.-3  $(X_1, X_2)$  est une variable aléatoire à valeurs dans  $E_1 \times E_2$ , la loi de  $(X_1, X_2)$  s'appellera loi conjointe de  $X_1$  et  $X_2$ . Réciproquement si  $X$  est une variable aléatoire à valeurs dans un produit cartésien  $E_1 \times E_2$ , la première et la seconde composante de  $X$  sont des variables aléatoires, leurs lois sont appelées respectivement première et seconde loi marginale de  $X$ . Nous allons donner une définition plus générale dans le cas de  $p$  variables aléatoires.

Jusqu'à la fin du paragraphe  $p$  désigne un entier supérieur ou égal à 2, et  $E_1, E_2, \dots, E_p$  des ensembles.

**Définition 3.3.1.** — Soient  $X_1, X_2, \dots, X_p$  des variables aléatoires à valeurs respectivement dans les ensembles  $E_1, E_2, \dots, E_p$ . On appelle loi conjointe (ou jointe) de ces variables (prises dans cet ordre) la loi de la variable aléatoire  $(X_1, X_2, \dots, X_p)$ . Autrement dit, en notant  $X = (X_1, \dots, X_p)$ , pour  $(\omega_1, \omega_2, \dots, \omega_p)$  un élément de  $X(\Omega)$ ,

$$\mathbf{P}_X(\{(\omega_1, \omega_2, \dots, \omega_k)\}) = \mathbf{P}(X_1 = \omega_1, \dots, X_p = \omega_p).$$

Si  $X$  est une variable aléatoire à valeurs dans  $E_1 \times E_2 \times \dots \times E_p$ , pour  $k = 1, 2, \dots, p$ , on appelle  $k^{\text{e}}$  loi marginale de  $X$  la loi de  $X_k$ ,  $k^{\text{e}}$  composante de  $X$ .

On peut déduire facilement une loi marginale de la loi d'une variable :

**Proposition 3.3.2** — Soit  $X = (X_1, \dots, X_p)$  une variable aléatoire à valeurs dans  $E_1 \times \dots \times E_p$ . Pour  $k$  élément de  $\{1, \dots, p\}$ , la  $k^{\text{e}}$  loi marginale de  $X_k$ ,  $P_{X_k}$  est donnée par :

$$\mathbf{P}_{X_k}(A) = \mathbf{P}\left(X \in (E_1 \times \dots \times E_{k-1} \times A \times E_{k+1} \times \dots \times E_p)\right),$$

pour tout  $A \subset X_k(\Omega)$ .

**Preuve de la proposition 3.3.2.** —

Soit  $A \subset X_k(\Omega)$ . L'ensemble  $\{X_k \in A\} = \{X \in E_1 \times \dots \times E_{k-1} \times A \times E_{k+1} \times \dots \times E_p\}$ , le résultat en résulte instantanément.  $\square$

**Illustration lorsque  $p = 2$ ,** — Soient  $X$  et  $Y$  des variables aléatoires indépendantes et qui suivent la même loi binomiale d'ordre  $n$  et de paramètre  $p$ . On note  $U = \min\{X, Y\}$  et  $V = \max\{X, Y\}$ .

Déterminons pour commencer la loi du couple  $(U, V)$ . En premier lieu

$$(U, V)(\Omega) = \{(i, j) \in \{0, \dots, n\}^2 | i \leq j\}.$$

Soit alors  $(i, j) \in (U, V)(\Omega)$ . Deux cas.

— Premier cas  $i = j$ .

$$\mathbf{P}(U = i, V = j) = \mathbf{P}(X = i, Y = i) = \mathbf{P}(X = i)\mathbf{P}(Y = i) = \left(\binom{n}{i} p^i (1-p)^{n-i}\right)^2,$$

car  $X$  et  $Y$  sont indépendantes.

— second cas  $i < j$

$$\mathbf{P}(U = i, V = j) = \mathbf{P}\left(\left((X = i) \cap (Y = j)\right) \cup \left((X = j) \cap (Y = i)\right)\right) = \mathbf{P}(X = i, Y = j) + \mathbf{P}(X = j, Y = i),$$

Car les événements  $((X = i) \cap (Y = j))$  et  $((X = j) \cap (Y = i))$  sont incompatibles, or  $X$  et  $Y$  sont indépendantes, donc,

$$\mathbf{P}(U = i, V = j) = \mathbf{P}(X = i)\mathbf{P}(Y = j) + \mathbf{P}(X = j)\mathbf{P}(Y = i) = 2 \binom{n}{i} p^i (1-p)^{n-i} \binom{n}{j} p^j (1-p)^{n-j}.$$

Au total,

$$\mathbf{P}(U = i, V = j) = \begin{cases} \binom{n}{i}^2 p^{2i} (1-p)^{2n-2i}, & \text{si } i = j, \\ 2 \binom{n}{i} \binom{n}{j} p^{i+j} (1-p)^{2n-i-j}, & \text{si } i < j. \end{cases}$$

Etudions par exemple la première loi marginale de  $(U, V)$ .

Soit  $i \in \{0, \dots, n\}$ .

$$\mathbf{P}_U(\{i\}) = \mathbf{P}((U, V) \in \{i\} \times \{1, \dots, n\}) = \sum_{j=0}^n \mathbf{P}((U, V) = (i, j)) = \sum_{j=i}^n \mathbf{P}_{(U,V)}(i, j),$$

en effet les événements  $\{(U, V) = (i, j)\}$ ,  $j = 0, \dots, n$  sont deux à deux disjoints de réunion  $\{(U, V) = (i, j)\}$ .  
Donc

$$\mathbf{P}_U(\{i\}) = \binom{n}{i}^2 p^{2i} (1-p)^{2n-2i} + \sum_{j=i+1}^n 2 \binom{n}{i} \binom{n}{j} p^{i+j} (1-p)^{2n-i-j},$$

En convenant qu'une somme vide est nulle.

On peut présenter synthétiquement ces résultats dans un tableau.

	0	1	...	$j$	...	$n$	
0	$\mathbf{P}_{(U,V)}(\{0,0\})$	$\mathbf{P}_{(U,V)}(\{0,1\})$	...	$\mathbf{P}_{(U,V)}(\{0,j\})$	...	$\mathbf{P}_{(U,V)}(\{0,n\})$	$\rightarrow \mathbf{P}_U(\{0\})$
1	$\mathbf{P}_{(U,V)}(\{1,0\})$	$\mathbf{P}_{(U,V)}(\{1,1\})$	...	$\mathbf{P}_{(U,V)}(\{1,j\})$	...	$\mathbf{P}_{(U,V)}(\{1,n\})$	$\rightarrow \mathbf{P}_U(\{1\})$
$\vdots$							$\vdots$
$i$	$\mathbf{P}_{(U,V)}(\{i,0\})$	$\mathbf{P}_{(U,V)}(\{i,1\})$	...	$\mathbf{P}_{(U,V)}(\{i,j\})$	...	$\mathbf{P}_{(U,V)}(\{i,n\})$	$\rightarrow \mathbf{P}_U(\{i\})$
$\vdots$							$\vdots$
$n$	$\mathbf{P}_{(U,V)}(\{n,0\})$	$\mathbf{P}_{(U,V)}(\{n,1\})$	...	$\mathbf{P}_{(U,V)}(\{n,j\})$	...	$\mathbf{P}_{(U,V)}(\{n,n\})$	$\rightarrow \mathbf{P}_U(\{n\})$
	$\downarrow$	$\downarrow$		$\downarrow$		$\downarrow$	
	$\mathbf{P}_V(\{0\})$	$\mathbf{P}_V(\{1\})$	...	$\mathbf{P}_V(\{j\})$	...	$\mathbf{P}_V(\{n\})$	

Les sommes en lignes donnent  $\mathbf{P}_U$ , celles en colonne  $\mathbf{P}_V$ .

On a noté abusivement  $\mathbf{P}((U, V) = (i, j))$  plus simplement  $\mathbf{P}_{(U,V)}(\{i, j\})$  même si  $(i, j)$  n'est pas élément de  $(U, V)(\Omega)$ .

En revanche, avec les notations de 3.3.1., la donnée pour une variable  $X = (X_1, \dots, X_p)$  de ses  $p$  lois marginales ne permet pas en général de connaître la loi de  $X$ . Donnons un contre-exemple.

**Contre-exemple 3.3.3.** — Une urne contient 4 boules numérotées de 1 à 4. Les deux premières sont blanches les deux dernières noires. Dans une première expérience on tire deux boules avec remise. L'univers  $\Omega_1$  est donc  $\{1, \dots, 4\}^2$  muni, faute de plus d'information, de la probabilité uniforme. On considère la variable aléatoire  $X = (X_1, X_2)$  à valeurs dans  $\{B, N\}^2$ , qui à un élément  $(x_1, x_2)$  de  $\Omega_1$  associe le couple  $(C_1, C_2)$  où  $C_i$  est la couleur de  $x_i$  pour  $i = 1, 2$ .



Ainsi  $X(1, 4) = (B, N)$ ,  $X(3, 3) = (N, N)$ . Les deux variables  $X_1$  et  $X_2$  suivent naturellement une loi de Bernoulli de paramètre  $\frac{1}{2}$ . En effet on a par exemple :

$$\mathbf{P}(X_1 = B) = \frac{|\{(B, B), (B, N)\}|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}.$$

Dans une seconde expérience on prélève sans remise deux boules dans notre urne, l'univers  $\Omega_2$  est alors l'ensemble des couples de 2 éléments distincts de  $\{1, \dots, 4\}^1$ , toujours muni de la probabilité uniforme. Le cardinal de  $\Omega_2$  est le nombre d'arrangements à 2 éléments de  $\{1, \dots, 4\}$ , soit  $4 \times 3 = 12$ . On considère la variable aléatoire  $Y = (Y_1, Y_2)$  à valeurs dans  $\{B, N\}^2$ , qui à un élément  $(y_1, y_2)$  de  $\Omega_2$  associe comme dans la première expérience le couple des couleurs correspondantes.

La variable  $Y_1$  suit clairement une loi de Bernoulli de paramètre  $\frac{1}{2}$ . L'étude de  $Y_2$  est plus délicate. D'après 2.3.5,

$$\mathbf{P}(Y_2 = B) = \mathbf{P}(Y_2 = B | Y_1 = B) \mathbf{P}(Y_1 = B) + \mathbf{P}(Y_2 = B | Y_1 = N) \mathbf{P}(Y_1 = N) = \frac{1}{3} \times \frac{1}{2} + \frac{2}{3} \times \frac{1}{2} = \frac{1}{2}.$$

Ainsi,  $Y_2$  suit-elle aussi une loi de Bernoulli de paramètre  $\frac{1}{2}$ .

Les variables  $X$  et  $Y$  ont mêmes lois marginales. Mais

$$\mathbf{P}(X = (B, B)) = \frac{|\{(1, 1); (1, 2); (2, 1); (2, 2)\}|}{4^2} = \frac{4}{4^2} = \frac{1}{4},$$

tandis que  $\mathbf{P}(Y = (B, B)) = \frac{|\{(1, 2), (2, 1)\}|}{12} = \frac{1}{6}$ . Ainsi donc  $X$  et  $Y$  n'ont-elles pas la même loi.

Remarque : on a noté de la même façon par  $\mathbf{P}$  les probabilités sur  $\Omega_1$  et sur  $\Omega_2$  afin de ne pas alourdir les notations

La loi conjointe permet également de découvrir les lois conditionnelles de variables aléatoires.

**Proposition 3.3.4** — Soient  $X = (X_1, \dots, X_p)$  une variable aléatoire à valeurs dans  $E_1 \times \dots \times E_p$  et  $k$  élément de  $\{1, \dots, p\}$  et  $(A_1, \dots, A_p)$  un élément de  $(X_1(\Omega), \dots, X_p(\Omega))$  tel que :

$$\mathbf{P}(X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1}, \dots, X_p \in A_p) \neq 0.$$

Alors

$$\mathbf{P}(X_k \in A_k | X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1} \dots X_p \in A_p) = \frac{\mathbf{P}_X(A_1 \times \dots \times A_p)}{\mathbf{P}_X(A_1 \times \dots \times A_{k-1} \times E_k \times A_{k+1} \times \dots \times A_p)}.$$

pour tout  $A \in E_1$ .

*Preuve de la proposition 3.3.4.* —

Par définition  $\mathbf{P}(X_k \in A_k | X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1} \dots X_p \in A_p)$  vaut :

$$\frac{\mathbf{P}\left((X_k \in A_k) \cap (X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1} \dots X_p \in A_p)\right)}{\mathbf{P}(X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1}, \dots, X_p \in A_p)}$$

D'une part,  $\mathbf{P}((X_k \in A_k) \cap (X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1} \dots X_p \in A_p))$  vaut :

$$\mathbf{P}(X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_k \in A_k, X_{k+1} \in A_{k+1} \dots X_p \in A_p) = P_X(A_1, \dots, A_p).$$

D'autre part,  $\mathbf{P}(X_1 \in A_1, \dots, X_{k-1} \in A_{k-1}, X_{k+1} \in A_{k+1}, \dots, X_p \in A_p)$  vaut :

$$\mathbf{P}(X \in (A_1 \times \dots \times A_{k-1} \times E_k \times A_{k+1} \times \dots \times A_p)) = \mathbf{P}_X(A_1 \times \dots \times A_{k-1} \times E_k \times A_{k+1} \times \dots \times A_p).$$

D'où le résultat. □

### 3.4 Variables aléatoires indépendantes

Dans tout ce paragraphe on dispose d'un espace probabilisé  $(\Omega, \mathbf{P})$ . Toutes les variables aléatoires sont définies sur cet espace.

1. Contrairement à la situation de l'exercice 3.2.7., l'ordre de tirage à une importance pour la suite et l'on ne se contente pas de prendre pour issue de notre expérience une partie à 2 éléments

**Définition 3.4.1.** — Des variables aléatoires  $X$  et  $Y$  à valeurs respectivement dans des ensembles  $E_1$  et  $E_2$  sont dites indépendantes si, pour toute partie  $A$  de  $E_1$  et toute partie  $B$  de  $E_2$ ,

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\mathbf{P}(Y \in B).$$

autrement dit si les événements  $\{X \in A\}$  et  $\{Y \in B\}$  sont indépendants.

Plus généralement des variables aléatoires  $X_1, \dots, X_p$  à valeurs respectivement dans des ensembles  $E_1, \dots, E_p$  sont dites mutuellement indépendantes si, pour tout  $(A_1, \dots, A_p)$  tel que  $A_1 \subset E_1, \dots, A_p \subset E_p$ ,

$$\mathbf{P}(X_1 \in A_1, \dots, X_p \in A_p) = \mathbf{P}(X_1 \in A_1) \times \dots \times \mathbf{P}(X_p \in A_p).$$

autrement dit si les événements  $\{X_1 \in A_1\}, \dots, \{X_p \in A_p\}$  sont mutuellement indépendants.

Des variables aléatoires  $X_1, \dots, X_p$  à valeurs respectivement dans des ensembles  $E_1, \dots, E_p$  sont dites deux à deux indépendantes si, pour tout couple  $(i, j)$  d'éléments distincts de  $\{1, \dots, p\}$ ,  $X_i$  et  $X_j$  sont indépendantes.

Si avec les notations de la précédente définition,  $X_1, \dots, X_p$  sont mutuellement indépendantes alors elles sont deux à deux indépendantes, en effet considérons sans restreindre la généralité les variables  $X_1$  et  $X_2$ . Pour toute partie  $A_1$  de  $E_1$  et  $A_2$  de  $E_2$ ,

$$\mathbf{P}_{(X_1, X_2)}(A_1 \times A_2) = \mathbf{P}(X_1 \in A_1, X_2 \in A_2) = \mathbf{P}(X_1 \in A_1, X_2 \in A_2, X_3 \in E_3, \dots, X_p \in E_p)$$

Donc par mutuelle indépendance,

$$\mathbf{P}_{(X_1, X_2)}(A_1 \times A_2) = \mathbf{P}(X_1 \in A_1)\mathbf{P}(X_2 \in A_2)\mathbf{P}(X_3 \in E_3) \dots \mathbf{P}(X_p \in E_p) = \mathbf{P}(X_1 \in A_1)\mathbf{P}(X_2 \in A_2) \times 1^{p-2}$$

et finalement

$$\mathbf{P}_{(X_1, X_2)}(A_1 \times A_2) = \mathbf{P}(X_1 \in A_1)\mathbf{P}(X_2 \in A_2).$$

En revanche, on peut avoir  $X_1, \dots, X_p$ , pour  $p \geq 3$  deux à deux indépendantes sans être mutuellement indépendantes.

**Remarque** : Soient des variables aléatoires  $X_1, \dots, X_p$  à valeurs respectivement dans des ensembles  $E_1, \dots, E_p$ . On a vu que les lois des variables  $(X_1, \dots, X_p)$  ne permettent pas en général d'obtenir leur loi conjointe, cf. contre-exemple 3.3.3. Par contre si les variables sont indépendantes, la chose est possible, puisqu'en notant  $X = (X_1, \dots, X_p)$ ,

$$\mathbf{P}_X(A_1 \times \dots \times A_p) = \mathbf{P}(X_1 \in A_1, \dots, X_p \in A_p) = \mathbf{P}(X_1 \in A_1) \times \dots \times \mathbf{P}(X_p \in A_p) = \mathbf{P}_{X_1}(A_1) \times \dots \times \mathbf{P}_{X_p}(A_p).$$

C'est précisément la situation décrite dans la première expérience du contre exemple 3.3.3. Les variables  $X_1$  et  $X_2$  sont indépendantes, par exemple :

$$\mathbf{P}(X_1 \in \{B\}, X_2 \in \{N\}) = \frac{|\{(1, 3); (1, 4); (2, 3); (2, 4)\}|}{4^2} = \frac{1}{4},$$

tandis que

$$\mathbf{P}(X_1 \in \{B\}) \times \mathbf{P}(X_2 \in \{N\}) = \frac{|\{(1, i); (2, i)|i \in \{1, 4\}\}|}{4^2} \times \frac{|\{(i, 3); (i, 4)|i \in \{1, 4\}\}|}{4^2} = \frac{8}{16} \times \frac{8}{16} = \frac{1}{4}.$$

On étudie de même tous les autres cas. Ceci est intuitivement rassurant le second tirage n'est pas influencé par le résultat du premier. On aurait du reste pu modéliser l'expérience par un couple de deux variables aléatoires  $X'_1$  et  $X'_2$  indépendantes qui suivent une loi de Bernoulli (à valeurs dans  $\{B, N\}$ ).

Plus généralement lorsqu'une expérience est en fait la répétition de  $N$  expériences toutes identiques et sans influence les unes sur les autres, il est possible de la modéliser par un  $N$ -uplet de variables aléatoires mutuellement indépendantes suivant la même loi. Prenons un exemple.

On jette une pièce non nécessairement équilibrée  $N$  fois ( $N$  désigne un entier supérieur ou égal à 2). On pourrait définir pour univers  $\{P, F\}^N$  et définir une probabilité sur cette univers, dans le cas où la pièce est équilibrée, ce serait la probabilité uniforme, dans les autres cas c'est un peu plus ardu et il est plus simple de considérer  $N$  variables aléatoires  $(X_1, \dots, X_N)$  mutuellement indépendantes et qui suivent toute la même loi de Bernoulli (à valeurs dans  $\{P, F\}$ ) et dont le paramètre  $p$  dépend de la pièce. La  $i^{\text{e}}$  variable donne le résultat du  $i^{\text{e}}$  lancer. Ainsi

$$\mathbf{P}(X_1 = P, X_2 = P, \dots, X_N = P) = \mathbf{P}(X_1 = P) \times \dots \times \mathbf{P}(X_N = P)$$

donne la probabilité pour qu'à chaque coup la pièce donne pile.

En général on ne précise pas l'espace probabilisé sur lequel sont définies les variables  $X_1, \dots, X_N$  et l'on ne s'inquiète même pas de son existence. Toutefois il est possible de justifier cette négligence. Les plus curieux trouveront dans l'exercice suivant une présentation rigoureuse de cette façon de procéder.

**Exercice 3.4.2.** — Soit  $(E_1, \mathbf{P}_1), \dots, (E_n, \mathbf{P}_n)$  des espaces probabilisés. On note  $E = E_1 \times \dots \times E_n$  et pour  $i = 1, \dots, n$ ,  $X_i$  la  $i^{\text{e}}$  projection de  $E$ ,

$$X_i : E \rightarrow E_i; (x_1, \dots, x_n) \mapsto x_i.$$

1. Montrer qu'il existe une probabilité  $\mathbf{P}$  et une seule sur  $E$  telle que pour tout élément  $(x_1, \dots, x_N)$  de  $\mathbf{E}$ ,

$$\mathbf{P}(x_1, \dots, x_N) = \mathbf{P}_1(x_1)\mathbf{P}_2(x_2) \times \mathbf{P}_N(x_N).$$

2. Montrer que les variables  $X_i$ ,  $i = 1, \dots, N$  sont mutuellement indépendantes et que pour  $i = 1, \dots, N$  et tout élément  $A_i$  de  $E_i$ ,  $\mathbf{P}(X_i \in A_i) = \mathbf{P}_i(A_i)$ . pour  $i = 1, \dots, N$
3. Application montrer qu'il existe  $N$  variables de Bernoulli toutes de paramètre  $p$ , définies sur un même espace probabilisé et mutuellement indépendantes.

**Proposition 3.4.3.** — Soient  $X_1$  et  $X_2$  des variables aléatoires indépendantes à valeurs respectivement dans des ensembles  $E_1, E_2$  Soient par ailleurs, pour  $i = 1, 2$   $f_i$  une application de  $E_i$  dans une ensemble  $F_i$ . Alors les variables aléatoires  $f_1(X_1)$  et  $f_2(X_2)$  sont indépendantes.

*Preuve de la proposition 3.4.3* — Soient  $A_1 \in F_1$  et  $A_2 \in F_2$ , alors :

$$\mathbf{P}(f_1(X_1) \in A_1, f_2(X_2) \in A_2) = \mathbf{P}(X_1 \in f_1^{-1}(A_1), X_2 \in f_2^{-1}(A_2)),$$

donc par indépendance de  $X_1$  et  $X_2$ ,

$$\mathbf{P}(f_1(X_1) \in A_1, f_2(X_2) \in A_2) = \mathbf{P}(X_1 \in f_1^{-1}(A_1))\mathbf{P}(X_2 \in f_2^{-1}(A_2)) = \mathbf{P}(f_1(X_1) \in A_1)\mathbf{P}(f_2(X_2) \in A_2).$$

Donc les variables aléatoires  $f_1(X_1)$  et  $f_2(X_2)$  sont indépendantes. □

Plus généralement on peut prouver le résultat suivant :

*Proposition 3.4.4.* — Soient un entier  $N \geq 2$ ,  $X_1, X_2, \dots, X_N$  des variables aléatoires mutuellement indépendantes à valeurs respectivement dans des ensembles  $E_1, \dots, E_N$  et  $k$  un entier tel que  $1 \leq k \leq N-1$ . Si  $f$  est une application définie sur  $E_1 \times E_2 \cdots \times E_k$  et  $g$  une application définie sur  $E_{k+1} \times E_{k+2} \cdots \times E_N$  alors les variables  $f(X_1, \dots, X_k)$  et  $g(X_{k+1}, \dots, X_N)$  sont indépendantes.

Attention de ne pas généraliser trop hâtivement la proposition 3.4.4., comme nous le montre l'exercice suivant :

**Exercice 3.4.5.** — Soient deux variables aléatoires à valeurs dans  $\{-1, 1\}$ ,  $U$  et  $V$ , définies sur un même espace probabilisé  $(\Omega, \mathbf{P})$  indépendantes et de même loi :

$$\mathbf{P}_U(-1) = \mathbf{P}_V(-1) = \frac{1}{3}; \quad \mathbf{P}_U(1) = \mathbf{P}_V(1) = \frac{2}{3}.$$

On définit sur  $(\Omega, \mathcal{A}, \mathbf{P})$ , les variables aléatoires sur  $X$  et  $Y$  par :

$$X = U, Y = \text{sign}(U)V.$$

1. Quelle est la loi de la variable aléatoire  $(X, Y)$  ?
2. Les variables aléatoires  $X$  et  $Y$  sont elles indépendantes ?
3. Les variables  $X^2$  et  $Y^2$  sont elles indépendantes ?

*Réponses*

1. La loi de  $(X, Y)$  est entièrement définie par les valeurs de  $P(X = x, Y = y)$ , pour tout couple  $(x, y)$  d'éléments de  $\{-1, 1\}$ , données ci dessous :

	Y	-1	1
X	· · ·		
-1		$\frac{2}{9}$	$\frac{1}{9}$
1		$\frac{2}{9}$	$\frac{4}{9}$

2. Non, par exemple  $\mathbf{P}(X = 1)\mathbf{P}(Y = 1) = \frac{2}{3} \times \left(\frac{1}{3} \times \frac{1}{3} + \frac{2}{3} \times \frac{2}{3}\right) = \frac{10}{27} \neq \mathbf{P}(X = 1, Y = 1)$ .
3. Oui d'après 3.3.4, puisque  $X^2 = U^2$  et  $Y^2 = V^2$ .

Etudions à présent la somme de  $N$  variables aléatoires de Bernoulli indépendantes. Cette étude est d'une grande utilité pratique. Considérons par exemple un joueur qui lance  $N$  fois une pièce non nécessairement équilibrée. Nous avons vu qu'une telle expérience se modélise par  $N$  variables aléatoires indépendantes  $X_1, \dots, X_N$  qui suivent toutes une même loi de Bernoulli, en convenant par exemple que pile est représenté par 1 et face par 0. Alors en notant  $S_N = X_1 + X_2 + \dots + X_N$ ,  $\mathbf{P}(S_n) = k$  représente pour tout entier  $k$  la probabilité que le joueur fasse au cours de ses  $N$  lancers,  $k$  fois pile. La loi de  $S_n$  est remarquable :

**Proposition 3.4.6.** — Soient un entier  $N \geq 1$  et  $N$  variables aléatoires indépendantes sur  $(\Omega, \mathcal{A}, \mathbf{P})$ ,  $X_1, \dots, X_N$  qui suivent toutes la loi de Bernoulli de paramètre  $p$ , pour  $i = 1, \dots, N$ ,

$$X_i \sim \mathcal{B}(p).$$

Alors la variable aléatoire somme  $S_N$ , définie par  $S_N = X_1 + X_2 + \dots + X_N$  suit une loi binomiale de paramètres  $N$  et  $p$  :

$$S_N \sim \mathcal{B}(N, p).$$

*Preuve de la proposition 3.4.6.* —

Commençons par une preuve par récurrence sur  $N$ , c'est sans doute la preuve la plus simple, mais elle masque le sens du résultat.

Notons  $(P_N)$  la propriété à prouver.

- $(P_1)$  est vraie puisque, comme nous l'avons déjà dit la loi binomiale  $\mathcal{B}(1, p)$  n'est rien d'autre que la loi de Bernoulli de paramètre  $p$ .
- Soit un entier  $M \geq 1$  tel que  $(P_M)$  soit vraie. Prenons  $X_1, \dots, X_{M+1}$  des variables aléatoires indépendantes qui suivent toutes la loi de Bernoulli de paramètre  $p$  et posons :

$$S_{M+1} = X_1 + \dots + X_{M+1}, \quad S_M = X_1 + \dots + X_M.$$

Notons que  $S_{M+1}$  est à valeurs dans  $\{1, \dots, M+1\}$ , puisque les  $X_i, i = 1, \dots, M+1$  sont à valeurs dans  $\{0, 1\}$ . Soit alors  $k \in \{0, \dots, M+1\}$ .

— *Premier cas* :  $k \geq 1$ . Puisque  $(\{X_{M+1} = 0\}; \{X_{M+1} = 1\})$  est un système complet d'événements,

$$\mathbf{P}(S_{M+1} = k) = \mathbf{P}(S_{M+1} = k | X_{M+1} = 1) \mathbf{P}(X_{M+1} = 1) + \mathbf{P}(S_{M+1} = k | X_{M+1} = 0) \mathbf{P}(X_{M+1} = 0).$$

Soit

$$\mathbf{P}(S_{M+1} = k) = \mathbf{P}(S_M = k-1)p + \mathbf{P}(S_M = k)(1-p).$$

Donc d'après  $(P_M)$ ,

$$\mathbf{P}(S_{M+1} = k) = \binom{M}{k-1} p^{k-1} (1-p)^{M-k+1} p + \binom{M}{k} p^k (1-p)^{M-k} (1-p) = \left( \binom{M}{k-1} + \binom{M}{k} \right) p^k (1-p)^{M+1-k}$$

et finalement par la formule de Pascal,

$$\mathbf{P}(S_{M+1} = k) = \binom{M+1}{k} p^k (1-p)^{M+1-k}$$

— *Second cas* :  $k = 0$

$$\mathbf{P}(S_{M+1} = 0) = \mathbf{P}(X_1 = 0, X_2 = 0, \dots, X_{M+1} = 0).$$

L'indépendance mutuelle des variables  $X_1, \dots, X_{M+1}$  assure donc :

$$\mathbf{P}(S_{M+1} = 0) = (1-p)^{M+1} = \binom{M+1}{0} p^0 (1-p)^{M+1-0}.$$

De ces deux cas il vient que  $(P_{M+1})$  est vraie.

Ainsi a-t-on prouvé par récurrence la propriété  $(P_{M+1})$ .

Donnons à présent une preuve ayant du sens.

Soit  $k$  un élément de  $\{1, \dots, N\}$ . Désignons par  $\mathcal{P}_k$  l'ensemble des parties de  $\{1, \dots, N\}$  à  $k$  éléments.

$$\mathbf{P}(S_N = k) = \mathbf{P} \left( \bigcup_{J \in \mathcal{P}_k} \left( \bigcap_{j \in J} \{X_j = 1\} \cap \bigcap_{j \in \bar{J}} \{X_j = 0\} \right) \right).$$

Les ensembles  $A_J = \left( \bigcap_{j \in J} \{X_j = 1\} \cap \bigcap_{j \in \bar{J}} \{X_j = 0\} \right)$ ,  $J \in \mathcal{P}_k$  sont deux à deux disjoints, en effet  $\omega \in A_J$  si et seulement si  $X_j(\omega) = 1$ , pour  $j \in J$  et  $X_j(\omega) = 0$  sinon. Donc :

$$\mathbf{P}(S_N = k) = \sum_{J \in \mathcal{P}_k} \mathbf{P} \left( \bigcap_{j \in J} \{X_j = 1\} \cap \bigcap_{j \in \bar{J}} \{X_j = 0\} \right).$$

l'indépendance mutuelle des  $X_j$ ,  $j = 1, \dots, n$  assure alors que :

$$\mathbf{P}(S_N = k) = \sum_{J \in \mathcal{P}_k} \prod_{j \in J} \mathbf{P}(X_j = 1) \times \prod_{j \in \bar{J}} \mathbf{P}(X_j = 0) = \sum_{J \in \mathcal{P}_k} p^k (1-p)^{N-k}.$$

finalement :

$$\mathbf{P}(X = k) = \binom{N}{k} p^k (1-p)^{N-k}.$$

Pour ceux qui déploieraient la sècheresse de ce calcul donnons une approche moins rigoureuse mais plus intuitive. Pour une partie  $J$  à  $k$  éléments donnée la probabilité de  $\{X_j = 1\}$  pour tous les éléments de  $J$  et  $\{X_j = 0\}$  pour les autres est  $p^k (1-p)^{N-k}$ . Il y a  $\binom{N}{k}$  parties à  $k$  éléments donc :  $\mathbf{P}(S_N = k) = \binom{N}{k} p^k (1-p)^{N-k}$ .  $\square$

### 3.5 Moments

Dans toute ce paragraphe on considère un espace probabilisé fini  $(\Omega, \mathbf{P})$  sur lequel seront définies, sauf mention contraire, toutes les variables aléatoires.

#### ESPÉRANCE

Commençons par une approche heuristique. On effectue une expérience quelconque et l'on considère une grandeur réelle  $X$  qui dépend de l'issue de l'expérience, autrement dit  $X$  est une fonction définie sur l'ensemble des issues possibles, à valeurs dans une partie  $E$  de  $\mathbf{R}$ . On imagine que l'on effectue un grand nombre  $n$  de fois l'expérience et on fait la moyenne des valeurs prises par  $X$ , on trouve

$$\frac{1}{n} \sum_{\omega \in \Omega} \text{Nb}(\{\omega\}) X(\omega),$$

où  $\text{Nb}(\{\omega\})$  représente le nombre de fois où l'événement élémentaire  $\omega$  se produit. Notons qu'avec les notations de l'introduction,  $\frac{1}{n} \text{Nb}(\{\omega\})$  vaut  $f_n(\{\omega\})$ , fréquence de l'événement  $\{\omega\}$  si bien que la moyenne des valeurs prises par  $X$ , vaut :

$$\sum_{\omega \in \Omega} f_n(\{\omega\}) X(\omega).$$

D'après l'introduction de la probabilité en terme de fréquences, cette quantité tend avec  $n$  en un certain sens vers :

$$\sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) X(\omega).$$

Cette quantité qui mesure la valeur moyenne de  $X$  s'appellera espérance de  $X$ .

**Définition 3.5.1** — Soit  $X$  une variable aléatoire réelle définie sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . On appelle espérance de  $X$  et l'on note  $E(X)$ , la quantité

$$E(X) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) X(\omega).$$

L'espérance est donc la moyenne des images par  $X$  des éléments  $\omega$  de  $\Omega$ , pondérée par la probabilité de l'événement  $\{\omega\}$ . On peut donner une autre forme de l'espérance.

**Proposition 3.5.2.** — Soit  $X$  une variable aléatoire définie sur  $(\Omega, \mathcal{A}, \mathbf{P})$ , à valeur dans une partie  $E$  de  $\mathbf{R}$ . Alors

$$E(X) = \sum_{x \in E} \mathbf{P}(X = x) x = \sum_{x \in X(\Omega)} \mathbf{P}(X = x) x.$$

L'espérance de  $X$  apparait donc comme la moyenne des valeurs  $x$  prises par  $X$  pondérées par la probabilité que  $X$  prenne la valeur  $x$ , en effet  $\sum_{x \in X(\Omega)} \mathbf{P}(X = x) = \sum_{x \in X(\Omega)} \mathbf{P}_X(\{x\}) = 1$ .

*Preuve de la proposition 3.5.2* — Que  $\sum_{x \in E} \mathbf{P}(X = x) x = \sum_{x \in X(\Omega)} \mathbf{P}(X = x) x$  résulte de ce que  $\mathbf{P}(X = x) = 0$ , pour tout élément  $x$  de  $E$  qui n'est pas élément de  $X(\Omega)$ . On notera au passage que  $\Omega$  étant fini, les sommes de la formule sont bien définies, la première comme somme d'un nombre fini de termes non nuls, l'autre comme somme finie. L'ensemble  $\{X = x\}_{x \in X(\Omega)}$  est une partition de  $\Omega$ . Donc :

$$E(X) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) X(\omega) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} \mathbf{P}(\{\omega\}) X(\omega),$$

Soit

$$E(X) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} \mathbf{P}(\{\omega\})x = \sum_{x \in X(\Omega)} \left( \mathbf{P} \left( \bigcup_{\omega \in \{X=x\}} \{\omega\} \right) \right) x,$$

puisque  $(\{\omega\})_{\omega \in \{X=x\}}$  est une famille d'éléments deux à deux incompatibles, leur réunion étant de plus  $\{X=x\}$ ,

$$E(X) = \sum_{x \in X(\Omega)} \mathbf{P}(X=x) x.$$

Voici la formule prouvée. □

**Exemple 3.5.3.** — Un jeu consiste à miser un franc sur un des numéros 1, 2, 3, 4, 5 et 6. On lance alors un dé et si le numéro sur lequel on a misé sort, le joueur empoche 5 francs, sinon il ne gagne rien, dans tous les cas il laisse sa mise. Modélisons le jeu et le gain que peut espérer le joueur. L'univers est  $\{1, \dots, 6\}$  que l'on munira de la probabilité uniforme. Le gain est représenté par une variable aléatoire  $G$  défini comme suit. On note  $k$  le numéro sur lequel le joueur a misé, alors pour tout  $x \in \Omega$ ,  $G(x) = -1$  si  $x \neq k$ , et  $G(k) = 5 - 1 = 4$ . L'espérance de  $G$  vaut donc :

$$E(G) = \mathbf{P}(G = -1)(-1) + \mathbf{P}(G = 4)4 = \mathbf{P}(\{1, \dots, 6\} \setminus \{k\})(-1) + \mathbf{P}(\{k\})4 = \frac{5}{6}(-1) + \frac{1}{6}4 = -\frac{1}{6}.$$

En conclusion, puisque l'espérance est négative, si l'on a foi en les probabilités, on peut s'attendre à ce qu'un joueur qui jouerait un grand nombre de parties, perde de l'argent (en gros  $\frac{1}{6}$  de franc).

**Remarque** — L'intérêt de la formule de la proposition 3.4.2. est qu'il est nul besoin de connaître l'espace probabilisé  $(\Omega, \mathbf{P})$  pour calculer l'espérance de  $X$ , seule la loi de  $X$  est nécessaire. En fait l'espérance de  $X$  ne dépend que de sa loi.

Donnons donc l'espérance d'une variables aléatoire réelle  $X$  lorsqu'elle obéit à des lois connues.

- **Variable constante** : la variable  $X$  prend une valeur constante  $c$ , donc  $\mathbf{P}(X=x) = \delta_{x,c}$ , pour tout réel  $x$ .

$$E(x) = 1 \cdot c = c.$$

- **Loi uniforme** : on suppose que  $X$  suit la loi uniforme sur  $\{1, \dots, n\}$ .

$$E(X) = \sum_{k=1}^n \mathbf{P}(X=k)k = \sum_{k=1}^n \frac{1}{n}k,$$

donc

si  $X \sim \mathcal{U}(\{1, \dots, n\})$  alors  $E = \frac{n+1}{2}$

- **Loi de Bernoulli** : on suppose que  $X$  suit la loi de Bernoulli de paramètre  $p$ .

$$E(X) = \mathbf{P}(X=1)1 + \mathbf{P}(X=0)0 = p \times 1 + (1-p) \times 0,$$

Donc

$$\text{si } X \sim \mathcal{B}(p) \text{ alors } E = p$$

- **Loi binomiale** : on suppose que  $X$  suit la loi de Binomiale de paramètre  $(n, p)$ . Le calcul direct est assez délicat, nous allons donner dans la suite une façon rapide de trouver l'espérance, pour lors, nous développons néanmoins une méthode de calcul

$$E(X) = \sum_{k=0}^n \mathbf{P}(X=k)k = \sum_{k=1}^n \binom{n}{k} p^k (1-p)^{n-k} k.$$

Par ailleurs on a l'égalité dans  $\mathbf{R}[T]$  :

$$\sum_{k=0}^n \binom{n}{k} T^k (1-p)^{n-k} = (T + (1-p))^n.$$

Par dérivation formelle :

$$\sum_{k=1}^n \binom{n}{k} k T^{k-1} (1-p)^{n-k} = n(T + (1-p))^{n-1}.$$

Donc en substituant à l'indéterminé  $T$  dans l'égalité précédente le réel  $p$  et en multipliant par  $p$ , on a :

$$\sum_{k=1}^n \binom{n}{k} k p^k (1-p)^{n-k} = p n (p + (1-p))^{n-1},$$

et donc

si  $X \sim \mathcal{B}(n, p)$ . alors  $E = pn$

Passons aux propriétés de l'espérance.

**Proposition 3.5.4.** — Soient  $X$  et  $Y$  des variables aléatoires réelles définies sur  $(\Omega, \mathcal{A}, \mathbf{P})$ , et  $(a, b)$  un couple de réels. L'espérance jouit des propriétés suivantes :

1. Linéarité :  $E(aX + bY) = aE(X) + bE(Y)$ , autrement dit  $E$  est une application linéaire de l'espace vectoriel des variables aléatoires réelles définies sur  $(\Omega, \mathcal{A}, \mathbf{P})$ , dans  $\mathbf{R}$ .
2. Positivité : si  $X$  est à valeurs positives, alors  $E(X) \geq 0$ .
3. Croissance : si  $X \geq Y$ , alors  $E(X) \geq E(Y)$ .

*Preuve de la proposition 3.5.4* —

1.

$$E(aX + bY) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\})(aX + bY)(\omega) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\})(aX(\omega) + bY(\omega))$$

Donc

$$E(aX + bY) = a \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\})X(\omega) + b \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\})Y(\omega) = aE(X) + bE(Y).$$

2. Supposons  $X \geq 0$ .

$$E(X) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\})X(\omega).$$

Donc  $E(X) \geq 0$ , puisque tous les termes de la somme précédente sont positifs.

3. Comme  $X - Y \geq 0$ , le deuxième point dit que  $E(X - Y) \geq 0$ , ce qui compte tenu du premier s'écrit :  $E(X) - E(Y) \geq 0$ .

□

**Application** — La proposition précédente redonne l'espérance d'une variable qui suit une loi binomiale. En effet si  $X_1, \dots, X_n$  sont des variables aléatoires sur  $(\Omega, \mathcal{A}, \mathbf{P})$  qui suivent toutes la loi de Bernoulli de paramètre  $p$ , alors on a vu que  $X_1 = X_2 + \dots + X_n$  suit la loi binomiale de paramètre  $(n, p)$ , cf. 3.3.6. Mais  $E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n) = np$  d'après le calcul déjà fait de l'espérance d'une variable de Bernoulli.

**Exercice 3.5.5.** — Généralisons le troisième point de 3.5.4. Soient  $X$  et  $Y$  des variables aléatoires réelles définies sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . On suppose que  $\mathbf{P}(X \geq Y) = 1$ . Montrer que  $E(X) \geq E(Y)$ .

*Solution de l'exercice 3.5.5.* — Notons  $\Omega' = \{X \geq Y\}$ . Par hypothèse  $\mathbf{P}(\Omega') = 1$ , donc  $\mathbf{P}(\bar{\Omega}') = 0$  et donc, pour tout  $\omega \in \bar{\Omega}'$ ,  $\mathbf{P}(\{\omega\}) = 0$ . Par ailleurs

$$E(X) = \sum_{\omega \in \Omega'} X(\omega)\mathbf{P}(\{\omega\}) + \sum_{\omega \in \bar{\Omega}'} X(\omega)\mathbf{P}(\{\omega\}) = \sum_{\omega \in \Omega'} X(\omega)\mathbf{P}(\{\omega\})$$

Donc

$$E(X) \geq \sum_{\omega \in \Omega'} Y(\omega)\mathbf{P}(\{\omega\}) = \sum_{\omega \in \Omega'} Y(\omega)\mathbf{P}(\{\omega\}) + \sum_{\omega \in \bar{\Omega}'} Y(\omega)\mathbf{P}(\{\omega\}) = E(Y).$$

**Définition-proposition 3.5.6.** — Une variable aléatoire réelle  $X_0$  est dite centrée si par définition :  $E(X_0) = 0$ .

Si  $X$  est une variable aléatoire réelle, alors la variable aléatoire  $X - E(X)$  est centrée.

**Preuve de 3.5.6.** — Résulte immédiatement de la linéarité de l'espérance et du calcul de l'espérance d'une variable constante. □

On étudie maintenant l'espérance d'un produit de composition.

**Proposition 3.5.7.** — FORMULE DE TRANSFERT —

Soit  $X$  une variable aléatoire sur  $(\Omega, \mathcal{A}, \mathbf{P})$  à valeurs dans un espace  $E$  et  $f$  une application de  $X(\Omega)$  à valeurs réelles. Alors

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbf{P}(X = x) = \sum_{x \in X(\Omega)} f(x)\mathbf{P}_X(\{x\})$$

Autrement dit,  $E(f(X))$  est l'espérance de  $f$  sur l'espace probabilisé fini  $(X(\Omega), \mathbf{P}_X)$ .

**Remarque :** on peut écrire encore avec quelques abus,  $E(f(X)) = \sum_{x \in E} f(x)\mathbf{P}(X = x)$ , puisque pour tout  $x \in E$  si  $x \notin X(\omega)$ , alors  $\mathbf{P}(X = x) = 0$ , ( $f$  est alors prolongée à  $E$  de quelconque manière).

*Preuve de la proposition 3.5.7.* —

L'ensemble  $\{X = x\}_{x \in X(\Omega)}$  est une partition de  $\Omega$ . Donc :

$$E(f(X)) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) f(X(\omega)) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} \mathbf{P}(\{\omega\}) f(X(\omega)),$$

Soit

$$E(f(X)) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} \mathbf{P}(\{\omega\}) f(x) = \sum_{x \in X(\Omega)} \left( \mathbf{P} \left( \bigcup_{\omega \in \{X=x\}} \{\omega\} \right) \right) f(x),$$

puisque  $(\{\omega\})_{\omega \in \{X=x\}}$  est une famille d'éléments deux à deux incompatibles, leur réunion étant de plus  $\{X = x\}$ ,

$$E(f(X)) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x) f(x) = \sum_{x \in X(\Omega)} \mathbf{P}_X(\{x\}) f(x).$$

Voici la formule prouvée. □

Etudions à présent le comportement de l'espérance vis à vis du produit.

**Proposition 3.5.8.** — *soient  $X_1, X_2, \dots, X_p$  des variables aléatoires réelles mutuellement indépendantes définies sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . Alors :*

$$E(X_1 \times X_2 \times \dots \times X_p) = E(X_1) \times E(X_2) \times \dots \times E(X_p).$$

*Preuve de la proposition 3.5.8.* — Posons  $X = (X_1, X_2, \dots, X_p)$  et  $Z = X_1 X_2 \dots X_p$ , notons que  $Z = f(X)$ , où  $f : \mathbf{R}^p \rightarrow \mathbf{R}; (x_1, \dots, x_p) \mapsto x_1 \times \dots \times x_p$ . Par la remarque qui suit 3.5.7.,

$$E(Z) = \sum_{(x_1, x_2, \dots, x_p) \in \mathbf{R}^p} \mathbf{P}(X = (x_1, x_2, \dots, x_p)) f(x_1, x_2, \dots, x_p).$$

Donc compte tenu de la mutuelle indépendance de  $X_1, \dots, X_p$ ,

$$E(Z) = \sum_{(x_1, x_2, \dots, x_p) \in \mathbf{R}^p} \mathbf{P}(X = x_1) \mathbf{P}(X = x_2) \dots \mathbf{P}(X = x_p) x_1 x_2 \dots x_p.$$

Donc

$$E(Z) = \sum_{x_1 \in \mathbf{R}} \mathbf{P}(X = x_1) x_1 \sum_{x_2 \in \mathbf{R}} \mathbf{P}(X = x_2) x_2 \dots \sum_{x_p \in \mathbf{R}} \mathbf{P}(X = x_p) x_p.$$

Finalement :  $E(Z) = E(X_1) E(X_2) \dots E(X_p)$ . □

**Attention !**

- le résultat 3.5.8. est faux sans l'hypothèse de mutuelle indépendance,
- que l'espérance du produit de variables aléatoires soit le produit de leurs espérances ne prouve pas la mutuelle indépendance de ces variables.

Lorsque une variable aléatoire  $X$  prend des valeurs entières il est souvent plus aisé de déterminer  $\mathbf{P}(X \geq x)$  que  $\mathbf{P}(X = x)$ , l'expression de  $\mathbf{P}(X = x)$  qui s'écrit  $\mathbf{P}(X = x) = \mathbf{P}(X \geq x) - \mathbf{P}(X \geq x + 1)$  prend une forme souvent compliquée. Qu'importe l'exercice qui suit donne l'expression de la variance en fonction de  $\mathbf{P}(X \geq x)$ .

**Exercice** — Soit  $X$  une variable aléatoire définie sur  $\Omega$  à valeurs dans  $\{0, n\}$ .

Montrer que :  $E(X) = \sum_{x=1}^n \mathbf{P}(X \geq x)$ .

*Solution* — Pour tout élément  $x$  de  $\{0, \dots, n\}$ ,  $\{X \geq x\}$  est la réunion disjointe de  $\{X \geq x + 1\}$  et de  $\{X = x\}$ , si bien que :

$$\mathbf{P}(X = x) = \mathbf{P}(X \geq x) - \mathbf{P}(X \geq x + 1)$$

Donc

$$E(X) = \sum_{x=0}^n x \mathbf{P}(X = x) = \sum_{x=0}^n x (\mathbf{P}(X \geq x) - \mathbf{P}(X \geq x + 1)) = \sum_{x=0}^n x \mathbf{P}(X \geq x) - \sum_{x=0}^n x \mathbf{P}(X \geq x + 1)$$

Grâce au changement d'indice de sommation «  $y = x + 1$  » dans l'ultime somme, on obtient donc :

$$E(X) = \sum_{x=0}^n x \mathbf{P}(X \geq x) - \sum_{y=1}^{n+1} (y - 1) \mathbf{P}(X \geq y) = .$$



Mais  $X$  étant à valeur dans  $\{1, \dots, n\}$ ,  $\mathbf{P}(X \geq n+1) = 0$ , donc

$$\mathbf{E}(X) = \sum_{x=1}^n x \mathbf{P}(X \geq x) - \sum_{x=1}^n (x-1) \mathbf{P}(X \geq x) = \sum_{x=1}^n \mathbf{P}(X \geq x).$$

### ÉCART TYPE

L'espérance d'une variable aléatoire, donne on l'a vu une idée de la valeur moyenne des valeurs prises par  $X$  sur un grand nombre d'expériences aléatoires. Par contre elle ne donne aucune information sur les fluctuations de ces valeurs autour de cette moyenne. Pour mesurer l'écart moyenne entre les valeurs prises par  $X$  et  $\mathbf{E}(X)$ , il est d'abord naturel de considérer  $\mathbf{E}(|X - \mathbf{E}(X)|)$ . Mais cette quantité est peu propice aux calculs et l'on préfère introduire une quantité, appelée variance de  $X$ .

**Définition 3.5.9.** — VARIANCE — Soit  $X$  une variable aléatoire réelle. On appelle variance de  $X$  et l'on note  $V(X)$ , la quantité :

$$V(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$$

La variance est donc donnée par les formules :

$$V(X) = \sum_{\omega \in \Omega} (X(\omega) - \mathbf{E}(X))^2 \mathbf{P}(\{\omega\}); \quad V(x) = \sum_{x \in X(\omega) \text{ (où } E)} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x)$$

$\mathbf{E}((X - \mathbf{E}(X))^2)$  est en quelque sorte la moyenne des écarts au carré entre  $X(\omega)$  et  $\mathbf{E}(X)$ , pondérée par la probabilité de  $\{\omega\}$ . Ce faisant, on diminue par rapport  $\mathbf{E}(|X - \mathbf{E}(X)|)$  l'importance des petits écarts (inférieurs à 1) et l'on augmente celle des plus importants, mais ceci n'a que peu d'importance face au gain en matière de calcul et à la richesse mathématique de cette notion.

Donnons quelques propriétés de la variance utiles à son calcul pratique.

**Proposition 3.5.10.** — Soient  $X$  une variable aléatoire réelle,  $a$  et  $b$  des réels, alors :

$$V(aX + b) = a^2 V(X).$$

*Preuve de la proposition 3.5.10* — Par linéarité de l'espérance et le fait que l'espérance d'une variable constante est précisément cette constante,

$$V(aX + b) = \mathbf{E}((aX + b - \mathbf{E}(aX + b))^2) = \mathbf{E}((aX + b - (a\mathbf{E}(X) + b))^2) = \mathbf{E}(a^2(X - \mathbf{E}(X))^2) = a^2 \mathbf{E}((X - \mathbf{E}(X))^2) = a^2 V(X).$$

□

La variance est donc *homogène de degré 2* il en résulte que si  $X$  représente une grandeur physique, des mètres par exemple  $V(X)$  ne représente pas la même grandeur, mais son carré. C'est pourquoi, on définit un nouveau indicateur de dispersion, l'*écart type*, qui est la racine carrée de la variance.

**Définition 3.5.11** — Soit  $X$  une variable aléatoire réelle. On appelle écart-type  $X$  et l'on note  $\sigma(X)$ , la quantité :

$$\sigma(x) = \sqrt{V(X)}.$$

**Définition, proposition 3.5.12.** — Soit  $X_0$  une variable aléatoire réelle. On dit que  $X_0$  est réduite si  $V(X_0) = 1$ . Si  $X$  est une variable aléatoire réelle d'écart-type non nul, alors  $\frac{X - \mathbf{E}(X)}{\sigma(X)}$  est une variable centrée et réduite.

*Preuve de la proposition 3.5.12* — Résulte directement de la linéarité de l'espérance et de 3.4.10. □

La formule qui suit est comme nous le verrons utile pour calculer la variance dans de nombreux cas.

**Proposition 3.5.13.** — FORMULE DE KÖNIG-HUYGHENS — Soit  $X$  une variable aléatoire réelle. Alors :

$$V(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2.$$

*Preuve de 3.5.13.* —

$$V(X) = \mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{E}(X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2).$$

Donc par linéarité de l'espérance, et le fait que l'espérance d'une variable constante est cette constante,

$$V(X) = E(X^2) - 2E(x)E(X) + E(X)^2 = E(X^2) - E(X)^2.$$

□

On peut utiliser ce résultat dans le calcul la variance d'une variables aléatoire réelle  $X$  lorsqu'elle obéit à des lois connues.

- **Variable constante** : la variable  $X$  prend une valeur constante  $c$ ,

$$V(X) = E(X^2) - (E(X))^2 = c^2 - c^2 = 0$$

- **Loi de Bernoulli** :

$$V(X) = E(X^2) - (E(X))^2 = E(X) - E(X)^2 = p - p^2$$

Donc

$$\text{si } X \sim \mathcal{B}(p), \text{ alors } V(X) = p(1 - p)$$

La formule de König-Huyghens assure en particulier que  $E(X^2) \geq E(X)^2$ . L'exercice suivant généralise ce résultat.

**Exercice 3.5.14.** — Soient  $X$  et  $Y$  des variables aléatoires. Montrer que :

$$E(XY)^2 \leq E(X)^2 E(Y)^2$$

On retrouve  $E(X^2) \geq E(X)^2$  en choisissant  $Y$  constante égale à 1

*Indication* : utiliser l'inégalité de Cauchy-Schwarz.

On peut aussi pour les calculs de variance utiliser le résultat suivant.

**Proposition 3.5.15.** — Soient  $X$  et  $Y$  des variables aléatoires réelles. Si elle sont indépendantes, alors :

$$V(X + Y) = V(X) + V(Y)$$

*Preuve de la proposition 3.5.15.* — Là encore, 3.5.13. et la linéarité de l'espérance nous sauve :

$$\begin{aligned} V(X+Y) &= E((X+Y)^2) - E(X+Y)^2 = E(X^2 + Y^2 + 2XY) - (E(X) + E(Y))^2 = \\ &= E(X^2) + E(Y^2) + 2E(XY) - E(X)^2 - E(Y)^2 - 2E(X)E(Y). \end{aligned}$$

Donc, toujours 3.5.13,

$$V(X+Y) = V(X) + V(Y) + 2(E(XY) - E(X)E(Y)) \quad (\text{I.5})$$

Mais si  $X$  et  $Y$  sont indépendantes, alors, d'après 3.4.7. :

$$V(X+Y) = V(X) + V(Y) + 2(E(X)E(Y) - E(X)E(Y)) = V(X) + V(Y).$$

□

On déduit de ce résultat la variance d'une loi binomiale.

**Loi binomiale** — La somme de  $n$  variables de Bernoulli indépendantes de paramètre  $p$  suit une loi binomiale de paramètre  $(n, p)$  donc sa variance est par applications itératives de 3.5.15.,  $n(p(1-p))$ .

$$\boxed{\text{si } X \sim \mathcal{B}(n, p), \text{ alors } V(X) = np(1-p)}$$

#### INÉGALITÉS DE MARKOV ET BIENAYMÉ-TCHEBYCHEV

Nous allons étudier la probabilité qu'une variable s'écarte de son espérance.

On appelle inégalité de Markov diverses inégalités qui reposent sur un raisonnement simple à retenir.

Prenons  $X$  une variable aléatoire réelle et  $h$  une application de  $\mathbf{R}$  dans  $\mathbf{R}$  à valeurs positives. Evaluons l'espérance de  $h(X)$ . Soient un réel  $a > 0$  et  $\Omega' := \{h(X) \geq a\}$ .

$$E(h(X)) = \sum_{\omega \in \Omega} h(X(\omega))\mathbf{P}(\{\omega\}) = \sum_{\omega \in \Omega'} h(X(\omega))\mathbf{P}(\{\omega\}) + \sum_{\omega \in \bar{\Omega}'} h(X(\omega))\mathbf{P}(\{\omega\}).$$

Comme  $h$  est à valeur positive, on a donc

$$E(h(X)) \geq \sum_{\omega \in \Omega'} h(X(\omega))\mathbf{P}(\{\omega\}) \geq \sum_{\omega \in \Omega'} a\mathbf{P}(\{\omega\}) = a\mathbf{P}(\Omega')$$

finalement

$$E(h(X)) \geq a\mathbf{P}(h(X) \geq a) \quad (\text{I.6})$$

On en déduit immédiatement pour  $h = |\cdot|$ , l'inégalité de Markov classique :

$$\boxed{\mathbf{P}(|X| \geq a) \leq \frac{E(|X|)}{a}} \quad (\text{inégalité de Markov})$$

Cette égalité a surtout une importance théorique et intervient dans la preuve de la loi faible des grands nombres.

Par le même genre de techniques on traitera l'exercice suivant.

**Exercice 3.5.16.** — Soit  $X$  une variable aléatoire réelle

1. Soit  $g$  une application de  $\mathbf{R}_+$  dans  $\mathbf{R}_+$ , strictement croissante. Montrer que pour tout réel  $a > 0$ ,

$$\mathbf{P}(|X| \geq a) \leq \frac{E(g(|X|))}{g(a)}$$

2. Soient un réel  $\alpha > 0$  et une application  $h : \mathbf{R} \rightarrow [0, \alpha]$ . Montrer que pour tout réel  $a$  tel que  $0 \leq a < \alpha$ , on a

$$\mathbf{P}(h(X) \geq a) \geq \frac{E(h(X)) - a}{\alpha - a}.$$

L'inégalité de Bienaymé-Tchebychev montre bien comment la variance contrôle l'écart d'une variable à son espérance. Soit un réel  $a > 0$ . Appliquons (I.6) à la variable aléatoire  $X - E(X)$  avec  $h : x \mapsto x^2$  :

$$\mathbf{P}((X - E(X))^2 \geq a^2) \leq \frac{E((X - E(X))^2)}{a^2}.$$

Mais comme  $\{|X - E(X)| \geq a\} = \{(X - E(X))^2 \geq a^2\}$ , on obtient la formule suivante :

$$\boxed{\mathbf{P}(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}} \quad (\text{inégalité de Bienaymé-Tchebychev})$$

## COVARIANCE, CORRÉLATION

On cherche à présent à mesurer la « ressemblance » entre deux variables aléatoires  $X$  et  $Y$ . Plus exactement étant donné  $X$  et  $Y$  variables aléatoires on cherche à comparer les variables centrées réduites  $X_0 = \frac{X - \mathbf{E}(X)}{\sigma(X)}$  et  $Y_0 = \frac{Y - \mathbf{E}(Y)}{\sigma(Y)}$ , ce afin de comparer la dispersion de  $X$  et  $Y$  autour de leurs espérances respectives en les ramenant à des quantités adimensionnées.

Pour ce faire remarquons que, dans le cas où aucun événement élémentaire n'est de probabilité nulle,

$$\mathbf{R}^\Omega \times \mathbf{R}^\Omega : (X_1, X_2) \mapsto E(X_1 X_2) = \sum_{\omega \in \Omega} X_1(\omega) X_2(\omega) \mathbf{P}\{\omega\}$$

est un produit scalaire, noté  $\langle \cdot | \cdot \rangle$ , sur l'espace vectoriel des variables aléatoires réelles sur  $(\Omega, \mathcal{A}, \mathbf{P})$ , la variance d'une variable aléatoire  $Z$  n'est alors que la norme au carré de  $Z - \mathbf{E}(Z)$ . Si des événements élémentaires  $\{\omega\}$  sont de probabilité nulle ce n'est plus un produit scalaire mais une forme bilinéaire symétrique positive.

L'inégalité de Cauchy-Schwarz (valable dans tous les cas) donne alors

$$|E(X_0 Y_0)| \leq \sqrt{V(X_0) V(Y_0)} = 1$$

Plus précisément en nous plaçant dans le cas où aucun événement élémentaire n'est de probabilité nulle,  $X_0$  et  $Y_0$  sont normés,  $-1 \leq E(X_0 Y_0) \leq 1$  et d'après la condition d'égalité dans l'inégalité de Cauchy-Schwarz  $E(X_0 Y_0) = 1$  si et seulement si  $X_0 = Y_0$ ,  $E(X_0 Y_0) = -1$  si et seulement si  $X_0 = -Y_0$ ,  $E(X_0 Y_0) = 0$  si et seulement si  $X_0$  est orthogonal à  $Y_0$ .  $E(X_0 Y_0)$  mesure donc la ressemblance entre  $X_0$  et  $Y_0$  si cette quantité est proche de 1 alors  $X_0$  et  $Y_0$  se ressemblent, si elle est proche de  $-1$  alors  $X_0$  est proche de  $-Y_0$  et enfin, si cette quantité est voisine de 0, alors  $X_0$  et  $Y_0$  sont sans rapport.

Notons que  $E(X_0 Y_0) = \frac{E((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))}{\sigma(X)\sigma(Y)}$  et adoptons les définitions suivantes :

**Définition, proposition 3.5.17** — Soient  $X$  et  $Y$  des variables aléatoires réelles sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . On appelle covariance de  $X$  et  $Y$  et l'on note  $\text{Cov}(X, Y)$  la quantité

$$\text{Cov}(X, Y) = E((X - \mathbf{E}(X))(Y - \mathbf{E}(Y))),$$

lorsque  $\sigma(X)\sigma(Y) \neq 0$ , on appelle coefficient de corrélation de  $X$  et  $Y$  et l'on note  $\rho(X, Y)$  la quantité

$$\rho(X, Y) = \frac{E((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))}{\sigma(X)\sigma(Y)}$$

Le coefficient de corrélation de  $X$  et  $Y$  est élément de  $[-1, 1]$ .

On dit que  $X$  et  $Y$  sont non corrélées si  $\text{Cov}(X, Y) = 0$ , ou si l'on préfère, lorsque  $\sigma(X)\sigma(Y) \neq 0$ , si  $\rho(X, Y) = 0$ .

L'interprétation en terme de forme bilinéaire symétrique de la covariance, assure les propriétés suivantes :

$$\boxed{\text{Cov}(X, X) = V(X)}$$

$$\boxed{V(X + Y) = V(X) + 2\text{Cov}(X, Y) + V(Y); \text{Cov}(X, Y) = \frac{1}{2}(V(X + Y) - V(X) - V(Y))}$$

On déduit de ces formules l'expression suivante de la covariance qui généralise 3.5.13.

**Proposition 3.5.18.** — Soient  $X$  et  $Y$  des variables aléatoires réelles sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . Alors

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

*Preuve de 3.5.18.* —

$$\text{Cov}(X, Y) = \frac{1}{2}(V(X + Y) - V(X) - V(Y)),$$

donc compte tenu de 3.5.13,

$$\frac{1}{2}(E((X + Y)^2) - E(X + Y)^2 - E(X^2) + E(X)^2 - E(Y^2) + E(Y)^2) = E(XY) - E(X)E(Y).$$



Passons au lien entre non corrélation et indépendance.

**Proposition 3.5.19.** — Soient  $X$  et  $Y$  des variables aléatoires réelles sur  $(\Omega, \mathcal{A}, \mathbf{P})$ . Si  $X$  et  $Y$  sont indépendantes, alors  $X$  et  $Y$  sont non corrélées.

*Preuve de la proposition 3.5.19.* — Résulte directement de l'expression de la covariance de 3.5.18. et de 3.5.8.  $\square$

La réciproque est fautive comme le montre le contre-exemple suivant.

**Contre-exemple 3.5.20.** — On prend pour espace probabilisé  $\{-1, 0, 1\}$  muni de la probabilité uniforme. Soient les variables aléatoires  $X$  qui est l'identité, et  $Y$  définie par  $Y = 1$ , si  $X = 0$ , 0 sinon. On a  $X \sim \mathcal{U}(\{-1, 0, 1\})$ ,  $XY$  nulle et  $\mathbf{E}(X) = 0$  donc :

$$\text{Cov}(XY) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = 0 - 0 = 0.$$

Les variables  $X$  et  $Y$  sont non corrélées.

Mais  $\mathbf{P}(X = 1, Y = 1) = 0$  puisque  $XY = 0$ , et  $\mathbf{P}(X = 1) = \frac{1}{3}$  et  $\mathbf{P}(Y = 1) = \frac{1}{3}$ , ainsi :

$$\mathbf{P}(X = 1, Y = 1) \neq \mathbf{P}(X = 1)\mathbf{P}(Y = 1),$$

et donc  $X$  et  $Y$  ne sont pas indépendantes.

#### MOMENTS

L'intérêt des moments apparaîtra dans le cours de spé. lorsque  $\Omega$  ne sera plus fini.

Brutalement, la définition :

**Définition 3.5.21.** — Soient  $X$  une variable aléatoire réelle sur  $(\Omega, \mathcal{A}, \mathbf{P})$  et un entier  $k \geq 0$ . On appelle moment d'ordre  $k$  de  $X$  la quantité  $\mathbf{E}(X^k)$ .

Ainsi, l'espérance de  $X$  est son moment d'ordre 0, sa variance le moment d'ordre 2 de  $X - \mathbf{E}(X)$ .

En particulier le moment d'ordre 2 de  $X$  vaut :

$$\sum_{\omega \in \Omega} X(\omega)^2 \mathbf{P}(\omega)$$

On notera l'analogie avec le moment<sup>1</sup> d'inertie  $\mathcal{M}$  par rapport à un point  $O$  de  $n$  points matériels  $\omega_1, \dots, \omega_n$  de masses respectives  $\mathbf{M}(x_1), \dots, \mathbf{M}(x_n)$ , distants respectivement de  $X(\omega_1), \dots, X(\omega_n)$  de 0 :

$$\mathcal{M} = \sum_{i=1}^n X(\omega_i)^2 \mathbf{M}(\omega_i)$$

La variance de  $X$  est à rapprocher du moment d'inertie par rapport au centre de masse.

1. Moment, vient du latin *momentum*-i, impulsion, poids...

## Chapitre II

# RÉVISIONS D'ALGÈBRE LINÉAIRE DE SUP. DIAGONALISATION, TRIGONALISATION

Le présent chapitre a deux objectifs : réviser les notions d'algèbres linéaires de math. sup. et introduire la diagonalisation des matrices et endomorphismes.

La révision du programme de MPSI se fera par un survol du cours de première année, nous ne respecterons pas la chronologie et l'ordre naturel de présentation des notions, de même nous ne prétendons pas reprendre de façon exhaustive la totalité du programme de math. sup. Le présent cours est plutôt une promenade agrémentée d'exemples et d'exercices au travers des connaissances d'algèbres linéaires indispensables à tout taupin.

Nous avons choisi d'introduire dès ce chapitre la diagonalisation et la trigonalisation des matrices et des endomorphismes, en effet ces notions cruciales sont nécessaires pour aborder rapidement des sujets de concours. Nous axerons notre étude et les exercices sur les aspects les plus pratiques de la théorie et sur ses utilisations (suites à récurrence linéaire, systèmes différentiels). Un prochain chapitre traitera de la réduction des endomorphismes et abordera entre autre la diagonalisation et la trigonalisation sous un jour plus théorique. Le présent chapitre a donc aussi vocation à préparer les difficultés à venir.

Dans tout ce chapitre  $\mathbf{K}$  désigne un corps. Dans la pratique le programme nous demande de nous limiter  $\mathbf{R}$  ou  $\mathbf{C}$ . C'est ce que nous ferons le plus souvent même si en quelques rares occasions nous pourrions considérer d'autre corps ( $\mathbf{Q}$ ,  $\mathbf{Z}/2\mathbf{Z}$ ...).

## ESPACES VECTORIELS

Nous allons, dans cette partie, reprendre rapidement les résultats principaux de sup. sur les espaces vectoriels. Nous supposons connu l'ensemble du cours de première année qui nous servira, entre autre, à illustrer ce cours d'exemples.

Dans la partie 1,  $\mathbf{E}$  désigne un espace vectoriel sur le corps  $\mathbf{K}$ . On ne fait aucune hypothèse de dimension sur  $\mathbf{E}$ , on ne suppose pas même que la notion de dimension ait été définie.

Nous noterons les vecteurs de  $\mathbf{E}$  par des minuscules latines italiques surmontés d'une flèche :  $\vec{x}$ ,  $\vec{y}$ ,  $\vec{z}$ ... Les éléments de  $\mathbf{K}$  seront eux notés par des lettres minuscules grecques :  $\alpha$ ,  $\beta$ ,  $\gamma$ ... L'élément neutre de  $\mathbf{E}$  sera noté  $\vec{0}_{\mathbf{E}}$ . Les applications linéaires par de lettres dépourvues de flèche  $\ell$ ,  $f$ ,...

### 1.1 Familles libres et génératrices, bases

**Définition 1.1.1.** — Soit  $(\vec{x}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$ , où  $I$  est un ensemble quelconque. Soit  $(\alpha_i)_{i \in I}$  une famille presque nulle<sup>1</sup>, c'est-à-dire que l'ensemble des éléments  $i$  de  $I$ , tels que  $\alpha_i$  soit non nul est fini. Soit  $\vec{x}$  un élément de  $\mathbf{E}$ . On dit que  $\vec{x}$  est la combinaison linéaire de la famille  $(\vec{x}_i)_{i \in I}$ , associée à la famille de coefficients  $(\alpha_i)_{i \in I}$ , si par définition

$$\vec{x} = \sum_{i \in I} \alpha_i \vec{x}_i.$$

1. On peut employer de façon concurrente l'expression à support fini

**Remarques** —

- Cette dernière somme est la somme d'un nombre fini de vecteurs de  $\mathbf{E}$ , puisque la famille  $(\alpha_i)_{i \in I}$  est presque nulle, ainsi a-t-elle bien un sens.
- Si  $I$  est l'ensemble vide, nous conviendrons que la somme précédente est nulle (c'est-à-dire égale à  $\vec{0}_{\mathbf{E}}$ ).
- Notons que si  $I$  est fini, nous retrouvons le cas le plus fréquemment rencontré en MPSI.

**Définition 1.1.2.** — Soit  $(\vec{x}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$ . Cette famille est dite libre si, par définition, la seule combinaison linéaire de la famille qui soit nulle est celle associée à la famille nulle. Autrement dit, pour toute famille  $(\alpha_i)_{i \in I}$  d'éléments de  $\mathbf{K}$ , presque nulle, si  $\sum_{i \in I} \alpha_i \vec{x}_i = \vec{0}_{\mathbf{E}}$ , alors  $(\alpha_i)_{i \in I}$  est une famille nulle.

Si la famille  $(\vec{x}_i)_{i \in I}$  est libre on dit aussi parfois que les vecteurs  $\vec{x}_i$ ,  $i \in I$  sont indépendants.

Une famille qui n'est pas libre est dite liée.

**Exemples** —

- Dans  $\mathcal{M}_{n,1}(\mathbf{R})$ , pour  $n \in \mathbf{N}^*$ , la famille  $(E_1, E_2, \dots, E_n)$ , où pour  $i = 1, 2, \dots, n$ ,  $E_i$  est le vecteur colonne dont tous les coefficients sont nuls excepté celui de la  $i^{\text{e}}$  ligne, est libre
- Dans  $\mathbf{R}^{\mathbf{N}}$  la famille  $(\delta_i)_{i \in \mathbf{N}}$ , où pour tout entier  $i \geq 0$ ,

$$\delta_i = (\delta_{i,n})_{n \in \mathbf{N}} = (\underbrace{0, \dots, 0}_{i+1 \text{ termes}}, 1, 0, \dots)$$

est libre<sup>1</sup>.

- Dans  $\mathbf{K}[X]$ , espace vectoriel des polynômes à coefficients dans  $\mathbf{K}$ , la famille  $(X^n)_{n \in \mathbf{N}}$  est libre.
- Dans  $\mathbf{C}(X)$ , espace vectoriel des fractions rationnelles à coefficients dans  $\mathbf{C}$ , la famille  $\left(\frac{1}{X-a}\right)_{a \in \mathbf{C}}$  est libre (cf. unicité de la décomposition en éléments simples).

**Exercice**

1. On considère la famille de vecteurs de  $\mathbf{R}^{\mathbf{R}}$ , espace des applications de  $\mathbf{R}$  dans  $\mathbf{R}$

$$(| \cdot - 1 |, | \cdot - 2 |, | \cdot - 3 |).$$

Cette famille est-elle libre ?

2. Même question pour la famille  $(1, \cos(2 \cdot), \cos^2)$ .
3. Montrer que dans  $\mathbf{R}^{\mathbf{R}}$  la famille  $(\exp(a \cdot))_{a \in \mathbf{R}}$  est libre.

**Solution** — Donnons la solution du 3.

Soit  $(\alpha_a)_{a \in \mathbf{R}}$  une famille presque nulle de réels telle que  $\sum_{a \in \mathbf{R}} \alpha_a \exp(a \cdot) = 0_{\mathbf{R} \rightarrow \mathbf{R}}$ . Supposons la famille  $(\alpha_a)_{a \in \mathbf{R}}$  non nulle. L'ensemble  $\{a \in \mathbf{R} | \alpha_a \neq 0\}$  est non vide et fini, à ce titre il admet un plus grand élément  $a_0$ . Alors

$$0 = \sum_{a \in \mathbf{R}} \alpha_a \exp(at) \underset{t \rightarrow +\infty}{\sim} \alpha_{a_0} \exp(a_0 t)$$

Voilà qui est absurde puisque  $\alpha_{a_0} \exp(a_0 \cdot)$  n'est pas nulle au voisinage de  $+\infty$ . Donc la famille  $(\alpha_a)_{a \in \mathbf{R}}$  est nulle et donc  $(\exp(a \cdot))_{a \in \mathbf{R}}$  est libre.

**Définition 1.1.3.** — Soit  $(\vec{x}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$ . Cette famille est dite génératrice si, par définition, tout vecteur de  $\mathbf{E}$  est combinaison linéaire de la famille  $(\vec{x}_i)_{i \in I}$ .

**Exemples** —

- Dans  $\mathbf{E}$  la famille identité indicée par  $\mathbf{E}$ , c'est-à-dire  $(\vec{x}_{\vec{a}})_{\vec{a} \in \mathbf{E}}$ , où pour tout élément  $\vec{a}$  de  $\mathbf{E}$ ,  $\vec{x}_{\vec{a}} = \vec{a}$ , est génératrice. On la note plus simplement  $(\vec{a})_{\vec{a} \in \mathbf{E}}$ .
- Dans  $\mathcal{M}_{n,1}(\mathbf{R})$ , pour  $n \in \mathbf{N}^*$ , la famille  $(E_1, E_2, \dots, E_n)$ , est génératrice.
- Dans  $\mathbf{K}[X]$ , espace vectoriel des polynômes à coefficients dans  $\mathbf{K}$ , la famille  $(X^n)_{n \in \mathbf{N}}$  est génératrice.

1. On aura reconnu le fameux symbole de Kronecker,  $\delta_{i,n}$ , cette quantité vaut 1 pour  $i = n$ , 0 dans les autres cas.

**Définition 1.1.4.** — On appelle *base* de  $\mathbf{E}$  toute famille de vecteurs de  $\mathbf{E}$  qui est libre et génératrice.

Soit  $(\vec{x}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$ . Notons  $F(I)$  l'ensemble des familles presque nulles d'éléments de  $\mathbf{K}$  indicées par  $I$ , on remarque sans mal que  $F(I)$  est un sous-espace vectoriel de l'espace vectoriel  $\mathcal{F}(I, \mathbf{E})$  des familles d'éléments de  $\mathbf{K}$  indicées par  $I$ . Soit alors l'application

$$\Phi : F(I) \rightarrow \mathbf{E}; (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i \vec{x}_i.$$

Les règles de calcul dans un espace vectoriel donnent immédiatement que  $\Phi$  est linéaire. On a alors la proposition suivante, dont la preuve facile est laissée en exercice :

**Proposition 1.1.5.** —

1. La famille  $(\vec{x}_i)_{i \in I}$  est libre si et seulement si le noyau de  $\Phi$  est réduit à la famille nulle, c'est-à-dire si et seulement si  $\Phi$  est injective.
2. La famille  $(\vec{x}_i)_{i \in I}$  est génératrice si et seulement si l'image de  $\Phi$  est  $\mathbf{E}$  tout entier, c'est-à-dire si et seulement si  $\Phi$  est surjective.
3. La famille  $(\vec{x}_i)_{i \in I}$  est une base si et seulement si  $\Phi$  est une bijection, c'est-à-dire si et seulement si, pour tout élément  $\vec{x}$  de  $\mathbf{E}$ , il existe une et une seule famille presque nulle d'éléments de  $\mathbf{K}$ ,  $(\alpha_i)_{i \in I}$  telle que :

$$\vec{x} = \sum_{i \in I} \alpha_i \vec{x}_i.$$

### Exemples

- Dans  $\mathcal{M}_{n,1}(\mathbf{R})$ , pour  $n \in \mathbf{N}^*$ , la famille  $(E_1, E_2, \dots, E_n)$  est une base de  $\mathcal{M}_{n,1}(\mathbf{R})$ , dite *base canonique*.
- Soient  $n$  et  $p$  des entiers strictement positifs. Pour tout élément  $i$  de  $\{1, \dots, n\}$  et tout élément  $j$  de  $\{1, \dots, p\}$ , on note  $E_{i,j}$  l'élément de  $\mathcal{M}_{n,p}(\mathbf{K})$  dont tous les coefficients sont nuls, à l'exception de celui de la  $i^{\text{e}}$  ligne et  $j^{\text{e}}$  colonne qui vaut 1. La famille  $(E_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,p}}$  est une base de  $\mathcal{M}_{n,p}(\mathbf{K})$ , dite *base canonique*.
- La famille  $(X^n)_{n \in \mathbf{N}}$  est une base de  $\mathbf{K}[X]$ . On l'appelle *base canonique* de  $\mathbf{K}[X]$ .
- Donnons une base de l'espace vectoriel des fractions rationnelles à coefficients dans  $\mathbf{C}$ ,  $\mathbf{C}(X)$ . Posons  $\tilde{C} := \mathbf{C} \cup \{\omega\}$ , où  $\omega$  n'est pas un élément de  $\mathbf{C}$ . Soit alors la famille d'éléments de  $\mathbf{C}(X)$ ,  $(R_{(n,a)})_{(n,a) \in \mathbf{N}^* \times \tilde{C}}$ , où pour tout élément  $n$  de  $\mathbf{N}^*$  et tout complexe  $a$ , on a :  $R_{n,\omega} = X^{n-1}$ ,  $R_{n,a} = \frac{1}{(X-a)^n}$ . Cette famille est une base de  $\mathbf{C}(X)$ , comme permettant de le prouver 1.1.5–3 et l'unicité de la décomposition en éléments simples d'une fraction rationnelle dans  $\mathbf{C}$ .

**Définition 1.1.6.** — Soit  $(\vec{e}_i)_{i \in I}$  une base de  $\mathbf{E}$ , notée  $\mathcal{B}$ . pour tout élément  $\vec{x}$  de  $\mathbf{E}$ , on appelle *famille des composantes* (ou des *coordonnées*) dans la base  $\mathcal{B}$  de  $\vec{x}$ , l'unique famille presque nulle  $(\alpha_i)_{i \in I}$  d'éléments de  $\mathbf{K}$  telle que :  $\vec{x} = \sum_{i \in I} \alpha_i \vec{e}_i$ , cf. 1.1.5–3. Pour tout élément  $i$  de  $I$ ,  $\alpha_i$  s'appelle la *composante* (ou *coordonnée*) de  $\vec{x}$  sur le vecteur  $\vec{e}_i$  dans la base  $\mathcal{B}$ .

**Remarque** — Un vecteur n'a toujours qu'un nombre fini de composantes non nulles.

**Remarques 1.1.7.** — Dans les définitions de famille libre, famille génératrice, base, l'ordre des vecteurs n'intervient pas ; plus précisément si  $(\vec{x}_i)_{i \in I}$  est une famille libre, (resp. une famille génératrice, une base) alors pour toute permutation  $\sigma$  de  $I$ ,  $(\vec{x}_{\sigma(i)})_{i \in I}$  est une famille libre, (resp. une famille génératrice, une base), comme le montre par exemple, la proposition 1.1.5.

D'aucuns saisissent l'occasion pour parler de *parties libres*, de *partie génératrice*. Pour donner des bases sérieuses à la chose nous dirons qu'une partie  $A$  de  $\mathbf{E}$  est libre (resp. génératrice), si, par définition, la famille  $(\vec{a})_{\vec{a} \in A}$  est libre, génératrice. Nous n'abuserons pas de ces notions...

### Rappel — ESPACE VECTORIEL DE DIMENSION FINIE —

Un espace vectoriel  $\mathbf{E}$  est dit de *dimension finie*, si il admet une famille génératrice finie, c'est-à-dire indexée par un ensemble fini.

Si c'est le cas, alors il admet une base finie  $\mathcal{B}_0$  et toute base de  $\mathbf{E}$  a même cardinal que  $\mathcal{B}_0$ <sup>1</sup>. le cardinal commun à toutes les bases s'appelle *dimension* de  $\mathbf{E}$ .

Dans un espace vectoriel de dimension finie, une base est une famille libre de cardinal maximum et une famille génératrice de cardinal minimum.

1. Le cardinal d'une famille finie est par définition le cardinal de la famille qui l'indexe, dans le cas d'une base, c'est aussi le cardinal de l'image de la famille.



Dans un espace de dimension  $n$ , tout famille libre de cardinal  $n$  est une base, toute famille génératrice de cardinal  $n$  est une base.

Dans un espace vectoriel de dimension finie, de toute famille génératrice on peut extraire une sous-famille qui soit une base.

Dans un espace vectoriel de dimension finie, si  $(\vec{e}_1, \dots, \vec{e}_p)$  est une famille libre non génératrice, de  $\mathbf{E}$  et  $(\vec{f}_1, \dots, \vec{f}_m)$  une famille génératrice, avec donc, d'après le point précédent,  $m \geq p$ , alors il existe une base de la forme

$$(\vec{e}_1, \dots, \vec{e}_p, \vec{f}_{i_1}, \vec{f}_{i_2}, \dots, \vec{f}_{i_k}),$$

où  $i_1, i_2, \dots, i_k$  sont éléments de  $\{1, \dots, m\}$ . Dit de façon moins rigoureuse mais plus parlante : on peut compléter toute famille libre en une base par des vecteurs d'une famille génératrice (théorème de la base incomplète). En particulier toute famille libre peut se compléter en une base.

Étudions l'effet d'une application linéaire sur les familles libres ou génératrice et sur les bases. Soit  $\mathbf{F}$  un espace vectoriel sur le même corps  $\mathbf{K}$  que  $\mathbf{E}$  et  $\ell$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ . Soit  $(\vec{e}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$ . Pour tout élément  $i$  de  $I$  on pose  $\vec{f}_i := \ell(\vec{e}_i)$ . La famille  $(\vec{f}_i)_{i \in I}$  s'appelle l'image de la famille  $(\vec{e}_i)_{i \in I}$  par  $\ell$ . Introduisons avec des notations issues de 1.1.5., les applications :

$$\Phi : F(I) \rightarrow \mathbf{E}; (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i \vec{e}_i,$$

$$\Phi' : F(I) \rightarrow \mathbf{F}; (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i \vec{f}_i.$$

La linéarité de  $\ell$  assure que :  $\Phi' = \ell \circ \Phi$ .

Si  $(\vec{e}_i)_{i \in I}$  est une famille libre, alors  $\Phi$  est injective, si de plus  $\ell$  est injective alors par composition,  $\Phi'$  est injective, et donc  $(\vec{f}_i)_{i \in I}$  est libre (1.1.5.).

Si  $(\vec{e}_i)_{i \in I}$  est une famille génératrice, alors  $\Phi$  est surjective, si de plus  $\ell$  est surjective alors par composition,  $\Phi'$  est surjective, et donc  $(\vec{f}_i)_{i \in I}$  est génératrice (1.1.5.).

Si  $(\vec{e}_i)_{i \in I}$  est une base, alors  $\Phi$  est bijective, si de plus  $\ell$  est bijective alors par composition,  $\Phi'$  est bijective, et donc  $(\vec{f}_i)_{i \in I}$  est une base (1.1.5.).

Enfin toujours dans le cas où  $(\vec{e}_i)_{i \in I}$  est une base de  $\mathbf{E}$ ,  $\ell(E) = \ell(\Phi(F(I))) = \Phi'(F(I)) = \text{vect}((\vec{f}_i)_{i \in I})$ .

On vient de prouver le résultat suivant :

**Proposition 1.1.8.** —

1. L'image par une application linéaire injective d'une famille libre est une famille libre.
2. L'image par une application linéaire surjective d'une famille génératrice est une famille génératrice.
3. L'image par une application linéaire bijective d'une base est une base.
4. L'image d'une application linéaire, est l'espace vectoriel engendré par l'image d'une base de l'ensemble de départ.

**Proposition, définition 1.1.9.** — Soit  $(\vec{x}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$ . L'ensemble des combinaisons linéaires de la famille  $(\vec{x}_i)_{i \in I}$  est un sous-espace vectoriel de  $\mathbf{E}$ . On l'appelle sous-espace vectoriel engendré par  $(\vec{x}_i)_{i \in I}$ , on le note  $\text{vect}((\vec{x}_i)_{i \in I})$ .

Le sous-espace vectoriel  $\text{vect}((\vec{x}_i)_{i \in I})$  est le plus petit sous-espace vectoriel contenant  $\{\vec{x}_i | i \in I\}$ .

*Preuve de la proposition 1.1.9.* — Avec les notations de 1.1.5., l'ensemble des combinaisons linéaires de la famille  $(\vec{x}_i)_{i \in I}$  est l'image de  $\Phi$ , c'est donc un sous-espace vectoriel de  $\mathbf{E}$ ; il contient évidemment  $\{\vec{x}_i | i \in I\}$ .

Tout sous-espace vectoriel de  $\mathbf{E}$  qui contient  $\{\vec{x}_i | i \in I\}$ , stable par combinaison linéaire, contient notamment n'importe quelle combinaison linéaire de  $\{\vec{x}_i | i \in I\}$ , donc contient  $\text{vect}((\vec{x}_i)_{i \in I})$ .

Donc  $\text{vect}((\vec{x}_i)_{i \in I})$  est le plus petit sous-espace vectoriel de  $\mathbf{E}$  contenant  $\{\vec{x}_i | i \in I\}$ . □

**Remarques** —

- La terminologie *sous-espace vectoriel engendré* par  $(\vec{x}_i)_{i \in I}$  est en accord avec le fait que par définition  $(\vec{x}_i)_{i \in I}$  en est une famille *génératrice*;
- Si  $A$  est une partie de  $\mathbf{E}$  on appelle sous-espace vectoriel engendré par  $A$ , le sous-espace vectoriel engendré par la famille  $(\vec{a})_{\vec{a} \in A}$ ; c'est donc le plus petit sous-espace vectoriel de  $\mathbf{E}$  qui contienne  $A$ . On le note  $\text{vect}(A)$ .

**Définition 1.1.10.** — Soit  $(\vec{x}_1, \dots, \vec{x}_p)$  une famille de  $\mathbf{E}$ ,  $\mathbf{K}$ -espace vectoriel de dimension quelconque, notée  $x$ . On appelle rang de  $x$ , noté  $\text{rg}(x)$ , la dimension de l'espace vectoriel engendré par  $(x)$ .

$$\text{rg}(x) = \dim(\text{vect}(x)).$$

On rappelle sans démonstration un résultat important dans la suite de ce chapitre.

**Proposition 1.1.11.** — On ne modifie pas le rang d'une famille finie de vecteurs en :

- multipliant un vecteur de la famille par un élément non nul de  $\mathbf{K}$ ;
- permutant l'ordre des vecteurs;
- ajoutant à l'un de ses vecteurs une combinaison linéaire des autres vecteurs.

## 1.2 Bases et applications linéaires

$\mathbf{E}$  et  $\mathbf{F}$  désignent des espaces vectoriels sur  $\mathbf{K}$ .

**Proposition 1.2.1.** — Supposons que  $\mathbf{E}$  admette une base  $(\vec{e}_i)_{i \in I}$ . l'application  $C$  de  $\mathbf{E}$  dans  $\mathcal{F}(I, \mathbf{K})$ , qui à un élément  $\vec{x}$  de  $\mathbf{E}$  associe la famille  $(x_i)_{i \in I}$  de ses coordonnées dans la base  $(\vec{e}_i)_{i \in I}$ , est linéaire

*Preuve de la proposition 1.2.1.* — La preuve directe de ce résultat qui est sans malice est laissée en exercice.

On peut aussi remarquer que  $C$  est à valeur dans  $F(I)$  et induit une application de  $\mathbf{E}$  dans  $F(I)$  qui n'est autre que la bijection réciproque de

$$\Phi : F(I) \rightarrow \mathbf{E}; (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i \vec{e}_i,$$

cf. 1.1.5. La bijection réciproque d'une application linéaire en est une autre, d'où le résultat. □

Le résultat qui va suivre est crucial et est à l'origine de la représentation matricielle des applications linéaires.

**Proposition 1.2.2.** — Supposons que  $\mathbf{E}$  admette une base  $(\vec{e}_i)_{i \in I}$ . Soit  $(\vec{f}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{F}$ . Il existe une et une seule application linéaire  $\ell$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , telle que pour tout élément  $i$  de  $I$ ,

$$\ell(\vec{e}_i) = \vec{f}_i. \quad (\text{II.1})$$

*Preuve de la proposition 1.2.2.* —

- Supposons qu'il existe une application linéaire  $\ell$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , vérifiant (II.1). Pour tout élément  $\vec{x}$  de  $\mathbf{E}$ , en notant  $(x_i)_{i \in I}$  la famille de ses coordonnées dans la base  $(\vec{e}_i)_{i \in I}$ , on a nécessairement

$$\ell(\vec{x}) = \ell\left(\sum_{i \in I} x_i \vec{e}_i\right) = \sum_{i \in I} x_i \ell(\vec{e}_i) = \sum_{i \in I} x_i \vec{f}_i.$$

- Réciproquement, soit  $\ell$  l'application de  $\mathbf{E}$  dans  $\mathbf{F}$  qui à tout élément  $\vec{x}$  de  $\mathbf{E}$  associe  $\sum_{i \in I} x_i \vec{f}_i$ , où  $(x_i)_{i \in I}$  est la famille de ses coordonnées dans la base  $(\vec{e}_i)_{i \in I}$ . Montrons que  $\ell$  convient.
  - $\ell$  est linéaire. En effet c'est  $\Phi' \circ C$  où

$$\Phi' : F(I) \rightarrow \mathbf{F}; (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i \vec{f}_i,$$

et  $C$  l'application composante dans  $(\vec{e}_i)_{i \in I}$  (cf. 1.2.1.), donc la composée de deux applications linéaires.

- D'autre part pour tout élément  $j$  de  $I$ , la famille de ses coordonnées dans  $(\vec{e}_i)_{i \in I}$  est  $(\delta_{i,j})_{i \in I}$  et donc  $\ell(\vec{e}_j) = \vec{f}_j$ .

Donc  $\ell$  convient.

Finalement il existe une et une seule application linéaire  $\ell$  vérifiant (II.1). □

**Remarque** — Supposons que  $\mathbf{F}$  soit de dimension finie non nulle  $p$  et prenons  $(\vec{f}_1, \vec{f}_2, \dots, \vec{f}_p)$  une de ses base. D'après 1.2.2., il existe une unique application linéaire  $\Phi$  de  $\mathbf{K}^p$  dans  $\mathbf{F}$ , qui au  $i^{\text{e}}$  vecteur de la base canonique de  $\mathbf{K}^p$ ,  $(\underbrace{0, 0, \dots, 1, 0, \dots, 0}_{i \text{ termes}})$ , associe  $\vec{f}_i$ , pour  $i = 1, 2, \dots, p$ ; c'est, d'après ce qui précède :

$$\Phi : \mathbf{K}^p \rightarrow \mathbf{F}; (\alpha_1, \alpha_2, \dots, \alpha_p) \mapsto \sum_{i=1}^p \alpha_i \vec{f}_i.$$

On retrouve l'application  $\Phi$  étudiée en 1.1.5., avec pour  $I$  l'ensemble fini  $\{1, 2, \dots, p\}$ .  $\Phi$  est, puisque  $(\vec{f}_1, \vec{f}_2, \dots, \vec{f}_p)$  est une base, bijective.

On vient de montrer :

*tout  $\mathbf{K}$ -espace vectoriel de dimension finie non nulle  $p$  est isomorphe à  $\mathbf{K}^p$ .*

### 1.3 Sommes directes

Dans ce paragraphe  $\mathbf{E}$  désigne un espace vectoriel sur un sous-corps  $\mathbf{K}$  de  $\mathbf{C}$ .

**Définition 1.3.1.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie<sup>1</sup> de sous-espaces vectoriels de  $\mathbf{E}$ . On appelle somme de cette famille, l'ensemble, noté  $\sum_{i \in I} \mathbf{E}_i$  des éléments  $\vec{x}$  de  $\mathbf{E}$  qui sont de la forme :

$$\vec{x} = \sum_{i \in I} \vec{x}_i,$$

où  $\vec{x}_i$  est, pour tout  $i \in I$ , un élément de  $\mathbf{E}_i$ .

**Remarque** — Si  $I = \{i_1, i_2, \dots, i_n\}$ ,  $\sum_{i \in I} \mathbf{E}_i$  se note aussi  $\mathbf{E}_{i_1} + \mathbf{E}_{i_2} + \dots + \mathbf{E}_{i_n}$ , on l'appelle aussi *somme des espaces*  $\mathbf{E}_{i_1}, \mathbf{E}_{i_2}, \dots$  et  $\mathbf{E}_{i_n}$ .

**Proposition 1.3.2.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$ . La somme de cette famille,  $\sum_{i \in I} \mathbf{E}_i$ , est un sous-espace vectoriel de  $\mathbf{E}$ .

Nous renvoyons le lecteur à son cours de sup. pour la preuve très facile de ce résultat.

**Définition 1.3.3.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$ . La somme de cette famille,  $\sum_{i \in I} \mathbf{E}_i$ , est dite directe si par définition chacun de ses éléments s'écrit de manière unique comme une somme d'éléments des  $\mathbf{E}_i$ ,  $i \in I$ ; c'est-à-dire si pour tout élément  $\vec{x}$  de  $\sum_{i \in I} \mathbf{E}_i$ , si  $\vec{x} = \sum_{i \in I} \vec{x}_i$ , et  $\vec{x} = \sum_{i \in I} \vec{y}_i$ , avec  $\vec{x}_i$  et  $\vec{y}_i$  éléments de  $\mathbf{E}_i$  pour tout  $i \in I$ , alors  $\vec{x}_i = \vec{y}_i$ , pour tout  $i \in I$ .

Si la somme de la famille est directe  $\sum_{i \in I} \mathbf{E}_i$  se note  $\bigoplus_{i \in I} \mathbf{E}_i$ .

**Remarque** — Si  $I = \{i_1, i_2, \dots, i_n\}$ , et si  $\sum_{i \in I} \mathbf{E}_i$  est directe on la note aussi

$$\mathbf{E}_{i_1} \oplus \mathbf{E}_{i_2} \oplus \dots \oplus \mathbf{E}_{i_n}$$

**Proposition 1.3.4.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$ . La somme de cette famille est directe, si et seulement si, pour toute famille  $(\vec{x}_i)_{i \in I}$  d'éléments de  $\mathbf{E}$ , telle que pour tout élément  $i$  de  $I$ ,  $\vec{x}_i \in \mathbf{E}_i$ , si  $\sum_{i \in I} \vec{x}_i = \vec{0}_{\mathbf{E}}$ , alors pour tout élément  $i$  de  $I$ ,  $\vec{x}_i = \vec{0}_{\mathbf{E}}$ .

*Preuve 1.3.4.* —

- Supposons la somme de la famille directe.

Soit  $(\vec{x}_i)_{i \in I}$  une famille d'éléments de  $\mathbf{E}$  telle que :

$$\sum_{i \in I} \vec{x}_i = \vec{0}_{\mathbf{E}};$$

— pour tout  $i \in I$ ,  $\vec{x}_i \in \mathbf{E}_i$ .

Or par ailleurs,

$$\sum_{i \in I} \vec{0}_{\mathbf{E}} = \vec{0}_{\mathbf{E}};$$

— pour tout  $i \in I$ ,  $\vec{0}_{\mathbf{E}} \in \mathbf{E}_i$ .

Donc par définition même d'une somme directe,  $\vec{x}_i = \vec{0}_{\mathbf{E}}$ , pour tout élément  $i$  de  $I$ .

- *Hypothèse* : pour toute famille  $(\vec{x}_i)_{i \in I}$  d'éléments de  $\mathbf{E}$ , telle que pour tout élément  $i$  de  $I$ ,  $\vec{x}_i \in \mathbf{E}_i$ , si  $\sum_{i \in I} \vec{x}_i = \vec{0}_{\mathbf{E}}$ ,

alors la famille  $(\vec{x}_i)_{i \in I}$  est nulle.

Soient  $(\vec{x}_i)_{i \in I}$  et  $(\vec{y}_i)_{i \in I}$  des familles d'éléments de  $\mathbf{E}$  telles que :

$$\sum_{i \in I} \vec{x}_i = \sum_{i \in I} \vec{y}_i;$$

— pour tout  $i \in I$ ,  $\vec{x}_i \in \mathbf{E}_i$  et  $\vec{y}_i \in \mathbf{E}_i$ .

Par différence, il vient :

1. « Finie » signifie ici que  $I$  est fini.

- $\sum_{i \in I} \vec{x}_i - \vec{y}_i = \vec{0}$ ;
  - pour tout  $i \in I$ ,  $\vec{x}_i - \vec{y}_i \in I$ .
- Donc par hypothèse,  $\vec{x}_i - \vec{y}_i = \vec{0}_{\mathbf{E}}$  soit  $\vec{x}_i = \vec{y}_i$ , pour tout  $i \in I$ .

D'où le résultat. □

**Proposition 1.3.5.** — Soit  $\mathbf{E}_1$  et  $\mathbf{E}_2$  des sous-espaces vectoriels de  $\mathbf{E}$ . Leur somme  $\mathbf{E}_1 + \mathbf{E}_2$  est directe si et seulement si leur intersection,  $\mathbf{E}_1 \cap \mathbf{E}_2$  est réduite à  $\{0_{\mathbf{E}}\}$ .

*Preuve de la proposition 1.4.4.* —

- Supposons que  $\mathbf{E}_1 + \mathbf{E}_2 = \mathbf{E}_1 \oplus \mathbf{E}_2$ .  
Soit  $\vec{x}$  un élément de  $\mathbf{E}_1 \cap \mathbf{E}_2$ .

$$\underbrace{\vec{x}}_{\in \mathbf{E}_1} + \underbrace{(-\vec{x})}_{\in \mathbf{E}_2} = \vec{0}_{\mathbf{E}}.$$

- Or  $\vec{x}$  est élément de  $\mathbf{E}_1$ ,  $-\vec{x}$  de  $\mathbf{E}_2$ , donc d'après la proposition précédente  $\vec{x} = \vec{0}_{\mathbf{E}}$ . Donc  $\mathbf{E}_1 \cap \mathbf{E}_2$  est réduite à  $\{0_{\mathbf{E}}\}$ .
- Supposons  $\mathbf{E}_1 \cap \mathbf{E}_2 = \{0_{\mathbf{E}}\}$ .  
Supposons qu'il existe  $\vec{x}_1, \vec{y}_1$  éléments de  $\mathbf{E}_1$ ,  $\vec{x}_2, \vec{y}_2$  éléments de  $\mathbf{E}_2$  tels que :

$$\vec{x}_1 + \vec{x}_2 = \vec{y}_1 + \vec{y}_2.$$

Alors  $\vec{x}_1 - \vec{y}_1 = \vec{y}_2 - \vec{x}_2$ , or  $\vec{x}_1 - \vec{y}_1 \in \mathbf{E}_1$  et  $\vec{y}_2 - \vec{x}_2 \in \mathbf{E}_2$ , donc  $\vec{x}_1 - \vec{y}_1$  et  $\vec{y}_2 - \vec{x}_2$  sont éléments de  $\mathbf{E}_1 \cap \mathbf{E}_2$ , ensemble qui est réduit à  $\{0_{\mathbf{E}}\}$ . Donc  $\vec{x}_1 = \vec{x}_2$  et  $\vec{y}_1 = \vec{y}_2$ . La somme est donc directe.

D'où l'équivalence annoncée. □

**!** **Mise en garde** — La propriété précédente ne se généralise pas au cas de trois sous-espaces vectoriels ou plus. S'il est vrai que l'intersection de sous-espaces vectoriels de  $\mathbf{E}$  en somme directe est encore nulle, ce, quel que soit leur nombre, il ne faut pas croire que  $p$  sous-espaces vectoriels de  $\mathbf{E}$ ,  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_p$ , avec  $p \geq 3$ , tels que  $\mathbf{E}_1 \cap \mathbf{E}_2 \cap \dots \cap \mathbf{E}_p = \{0_{\mathbf{E}}\}$ , soient nécessairement en somme directe. Donnons un exemple. Dans l'espace vectoriel  $\mathbf{R}^2$ , on considère les sous-espaces  $\mathbf{E}_1 = \text{vect}((1, 0))$ ,  $\mathbf{E}_2 = \text{vect}((0, 1))$  et  $\mathbf{E}_3 = \text{vect}((1, 1))$ . De toute évidence  $\mathbf{E}_1 \cap \mathbf{E}_2 \cap \mathbf{E}_3 = \{0_{\mathbf{E}}\}$ , pourtant la somme de ces trois espaces n'est pas directe puisque :

$$\underbrace{(1, 0)}_{\in \mathbf{E}_1} + \underbrace{(0, 1)}_{\in \mathbf{E}_2} + \underbrace{(-1, -1)}_{\in \mathbf{E}_3} = (0, 0) \text{ (cf. 1.4.4.)}$$

Par contre nous avons le résultat suivant d'utilité pratique très limitée :

**Proposition 1.3.6.** — Soient  $p$  un entier supérieur ou égal à 2 et  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_p$  des sous-espaces vectoriels de  $\mathbf{E}$ . La somme  $\mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_p$  est directe si et seulement si, pour tout élément  $j$  de  $\{1, 2, \dots, p\}$ ,

$$\mathbf{E}_j \cap \sum_{\substack{i=1, \dots, p \\ i \neq j}} \mathbf{E}_i = \{0_{\mathbf{E}}\}.$$

*Preuve de la proposition 1.3.6. —*

**Remarque** — Dans la pratique on montre souvent que  $n$  sous-espaces vectoriels  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$  sont en somme directe par récurrence. En effet si  $\mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_{n-1} = \mathbf{E}_1 \oplus \mathbf{E}_2 \oplus \dots \oplus \mathbf{E}_{n-1}$  et si  $\mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_{n-1}$  et  $\mathbf{E}_n$  sont en somme directe, alors  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$  sont en somme directe comme le montre la définition même d'une somme directe.

**Définition 1.3.7.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille de sous-espaces vectoriels de  $\mathbf{E}$ . On dit que les sous-espaces  $\mathbf{E}_i, i \in I$ , sont supplémentaires si, par définition les deux conditions suivantes sont réalisées :

1. La somme de la famille  $(\mathbf{E}_i)_{i \in I}$  est directe,  $\sum_{i \in I} \mathbf{E}_i = \bigoplus_{i \in I} \mathbf{E}_i$ .
2. La somme de la famille  $(\mathbf{E}_i)_{i \in I}$  vaut  $\mathbf{E}$ ,  $\sum_{i \in I} \mathbf{E}_i = \mathbf{E}$ .

On écrit alors :  $\mathbf{E} = \bigoplus_{i \in I} \mathbf{E}_i$ .

**Exemples 1.3.8.** —

- Prenons ici pour  $\mathbf{E}$ ,  $\mathcal{F}(\mathbf{R}, R)$ , ensemble des applications de  $\mathbf{R}$  dans  $\mathbf{R}$ . L'espace vectoriel des applications de  $\mathbf{R}$  dans  $\mathbf{R}$  paires, noté  $\mathcal{P}(\mathbf{R}, \mathbf{R})$  et l'espace vectoriel des applications de  $\mathbf{R}$  dans  $\mathbf{R}$  impaires, noté  $\mathcal{I}(\mathbf{R}, \mathbf{R})$ , sont deux sous espaces vectoriels de  $\mathcal{F}(\mathbf{R}, R)$  supplémentaires :

$$\mathcal{F}(\mathbf{R}, R) = \mathcal{P}(\mathbf{R}, R) \oplus \mathcal{I}(\mathbf{R}, R).$$

- Soit  $n$  un entier naturel non nul. Prenons ici pour  $\mathbf{E}$ ,  $\mathcal{M}_n(\mathbf{R})$ . L'ensemble des éléments de  $\mathcal{M}_n(\mathbf{R})$  symétriques, noté  $\mathcal{S}_n(\mathbf{R})$  et l'ensemble des éléments de  $\mathcal{M}_n(\mathbf{R})$  antisymétriques, noté  $\mathcal{A}_n(\mathbf{R})$ , sont des sous-espaces vectoriels de  $\mathcal{M}_n(\mathbf{R})$  supplémentaires :

$$\mathcal{M}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R}) \oplus \mathcal{A}_n(\mathbf{R}).$$

- Soit  $n$  un entier naturel. Prenons ici pour  $\mathbf{E}$ ,  $\mathbf{K}[X]$ . Soit  $P$  un élément de  $\mathbf{K}[X]$  de degré  $n+1$ . Désignons par  $\mathbf{K}[X]_n$  l'ensemble des éléments de  $\mathbf{K}[X]$  de degré inférieur ou égal à  $n$ , et par  $P\mathbf{K}[X]$  celui de ses éléments multiples de  $P$ . Alors  $\mathbf{K}[X]_n$  et  $P\mathbf{K}[X]$  sont des sous-espaces vectoriels de  $\mathbf{K}[X]$  supplémentaires :

$$\mathbf{K}[X] = \mathbf{K}[X]_n \oplus P\mathbf{K}[X].$$

Justifions.....

Revenons sur une notion de importante de MPSI : celle de projection. Soient  $\mathbf{F}$  et  $\mathbf{G}$  des sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires. Pour tout élément  $\vec{x}$  de  $\mathbf{E}$ , il existe un seul couple  $(\vec{y}, \vec{z})$  de  $\mathbf{F} \times \mathbf{G}$ , tel que :  $\vec{x} = \vec{y} + \vec{z}$  (cf. 1.4.3.) le vecteur  $\vec{y}$  s'appelle projeté de  $\vec{x}$  sur  $\mathbf{F}$  suivant  $\mathbf{G}$  et l'application  $p$  qui à un élément de  $\mathbf{E}$  associe son projeté sur  $\mathbf{F}$  suivant  $\mathbf{G}$ , s'appelle projection sur  $\mathbf{F}$  suivant  $\mathbf{G}$ <sup>1</sup>.

On appelle symétrique de  $\vec{x}$  par rapport à  $\mathbf{F}$  suivant  $\mathbf{G}$ , l'élément  $\vec{y} - \vec{z}$  et l'application  $s$  qui à un élément de  $\mathbf{E}$  associe son symétrique par rapport à  $\mathbf{F}$  suivant  $\mathbf{G}$ , s'appelle symétrie par rapport à  $\mathbf{F}$  suivant  $\mathbf{G}$ .

**Exercice** —

1. On reprend le première exemple de 1.3.8. Déterminer la projection de  $\exp$  sur  $\mathcal{P}$  suivant  $\mathcal{I}$  et celle sur  $\mathcal{I}$  suivant  $\mathcal{P}$ .
2. On reprend le deuxième exemple de 1.3.8. Identifier la symétrie par rapport à  $\mathcal{S}_n(\mathbf{R})$  suivant  $\mathcal{A}_n(\mathbf{R})$ .

*Solution* —

Maintenant généralisons !

---

1. Ne pas oublier de préciser suivant  $G$ , en effet  $p$  dépend essentiellement de  $G$ .

**Proposition-définition 1.3.9.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires :

$$\bigoplus_{i \in I} \mathbf{E}_i = \mathbf{E}$$

- Pour tout élément  $j$  de  $I$ ,  $\mathbf{E}_j$  et  $\sum_{i \in I - \{j\}} \mathbf{E}_i$  sont deux sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires.
- Pour tout élément  $j$  de  $I$ , la projection sur  $\mathbf{E}_j$  suivant  $\sum_{i \in I - \{j\}} \mathbf{E}_i$ , notée  $p_j$ , est appelée projection sur  $\mathbf{E}_j$ , relativement à la décomposition de  $\mathbf{E}$  en sous-espaces supplémentaires  $\mathbf{E} = \bigoplus_{i \in I} \mathbf{E}_i$ .

*Preuve de la proposition 1.3.9.* — La démonstration est laissée en exercice, c'est une conséquence directe de 1.3.6.

**Remarque** Gardons les hypothèses de la proposition précédente. Pour tout élément  $\vec{x}$  de  $\mathbf{E}$  il existe, une unique famille  $(\vec{x}_i)_{i \in I}$  d'éléments de  $\mathbf{E}$ , telle que :

$$\begin{cases} \vec{x} = \sum_{i \in I} \vec{x}_i \\ \vec{x}_i \in \mathbf{E}_i, \text{ pour tout élément } i \text{ de } I \end{cases}$$

Comme, pour tout élément  $j$  de  $I$ ,  $\vec{x}$  s'écrit encore :

$$\vec{x} = \underbrace{\vec{x}_j}_{\in \mathbf{E}_j} + \underbrace{\left( \sum_{i \in I - \{j\}} \vec{x}_i \right)}_{\in \sum_{i \in I - \{j\}} \mathbf{E}_i}$$

On en déduit que  $\vec{x}_j = p_j(\vec{x})$ .

La proposition suivante en découle :

**Proposition 1.3.10.** — Avec les notations de la proposition précédente, nous obtenons :

- pour tout élément  $j$  de  $I$ ,  $p_j \circ p_j = p_j$  ;
- pour tout  $j$  et tout  $i$  éléments distincts de  $I$ ,  $p_j \circ p_i = 0_{\mathcal{L}(\mathbf{E})}$  ;
- enfin  $\sum_{i \in I} p_i = id_{\mathbf{E}}$ .

**Proposition 1.3.11.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires, soit  $\mathbf{F}$  un espace vectoriel sur le corps  $\mathbf{K}$  et soit, pour tout élément  $i$  de  $I$ ,  $u_i$  une application linéaire de  $\mathbf{E}_i$  dans  $\mathbf{F}$ . Alors il existe une unique application linéaire  $u$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , telle que, pour tout élément  $i$  de  $I$ ,  $u_i$  soit la restriction de  $u$  à  $\mathbf{E}_i$ .

*Preuve 1.3.11. —*

- Soit  $\vec{u}$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ , telle que, pour tout élément  $i$  de  $I$ ,  $\vec{u}_i$  soit la restriction de  $u$  à  $\mathbf{E}_i$ . Soit  $\vec{x}$  élément de  $\mathbf{E}$ . En utilisant le troisième point de 1.4.10. et ses notations, il vient :

$$u(\vec{x}) = \left( u \circ \sum_{i \in I} p_i \right) (\vec{x}) = u \left( \sum_{i \in I} p_i(\vec{x}) \right) = \sum_{i \in I} u(p_i(\vec{x})).$$

Or pour tout élément  $i$  de  $I$ ,  $p_i(\vec{x}) \in \mathbf{E}_i$ , donc  $u(p_i(\vec{x})) = u_i(p_i(\vec{x}))$ . D'où :

$$u(\vec{x}) = \sum_{i \in I} u_i(p_i(\vec{x})).$$

Nécessairement  $u$  est l'application  $\sum_{i \in I} u_i \circ p_i$ .

- Réciproquement posons  $u := \sum_{i \in I} u_i \circ p_i$ . Evidemment  $u$  est linéaire comme somme de composées d'applications linéaires. Soient maintenant  $j$  élément de  $I$  et  $\vec{x}$  élément de  $\mathbf{E}_j$ . Comme  $p_j(\vec{x}) = \vec{x}$ ,

$$u(\vec{x}) = \sum_{i \in I} u_i \circ p_i \circ p_j(\vec{x}).$$

En utilisant 1.3.10., on obtient :

$$u(\vec{x}) = u_j \circ p_j(\vec{x}) = u_j(\vec{x}).$$

Ainsi la restriction de  $u$  à  $\mathbf{E}_j$  est-elle bien  $u_j$ . L'application  $u$  ainsi définie, convient.

D'où le résultat. □

**Proposition 1.3.12.** — BASES ET SOMMES DIRECTES —

On suppose ici  $\mathbf{E}$  de dimension finie. Soit  $p$  un entier supérieur ou égal à 1 et  $(\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_p)$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires. Pour tout élément  $i$  de  $I$ ,  $n_i$  désigne la dimension de  $\mathbf{E}_i$ , supposée non nulle, et  $(\vec{e}_{i,1}, \vec{e}_{i,2}, \dots, \vec{e}_{i,n_i})$  une base de  $\mathbf{E}_i$ , notée  $\mathcal{B}_i$ . Alors la famille  $\mathcal{B}$ , concaténée des bases  $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_p$ , c'est-à-dire la famille :

$$\mathcal{B} = (\vec{e}_{1,1}, \vec{e}_{1,2}, \dots, \vec{e}_{1,n_1}, \vec{e}_{2,1}, \vec{e}_{2,2}, \dots, \vec{e}_{2,n_2}, \dots, \vec{e}_{i,1}, \vec{e}_{i,2}, \dots, \vec{e}_{i,n_i}, \dots, \vec{e}_{p,1}, \vec{e}_{p,2}, \dots, \vec{e}_{p,n_p})$$

est une base de  $\mathbf{E}$ .

*Preuve de la proposition 1.3.12.* — C'est bougrement facile... (le caractère générateur de  $\mathcal{B}$  résulte des définitions 1.4.3. et 1.4.7., la liberté de 1.4.4.). Nous laissons au lecteur le soin et le plaisir de rédiger cette preuve.

Cette proposition admet le corollaire immédiat suivant :

**Proposition 1.3.13** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$ , supposé ici de dimension finie. Si la somme de la famille  $(\mathbf{E}_i)_{i \in I}$  est directe, alors :

$$\dim \left( \sum_{i \in I} \mathbf{E}_i \right) = \sum_{i \in I} \dim(\mathbf{E}_i).$$

**Corollaire (du corollaire) 1.3.14.** — Soit  $(\mathbf{E}_i)_{i \in I}$  une famille finie de sous-espaces vectoriels de  $\mathbf{E}$ , supposé ici de dimension finie, dont la somme est directe. Les sous-espaces,  $\mathbf{E}_i$ ,  $i \in I$  sont supplémentaires si et seulement si :

$$\sum_{i \in I} \dim(\mathbf{E}_i) = \dim(\mathbf{E}).$$

## 1.4 Autour du rang d'un endomorphisme

Où nous allons revenir sur un théorème d'algèbre linéaire de sup. fondamental : *le théorème du rang*.

Dans toute cette sous partie  $\mathbf{E}$  et  $\mathbf{F}$  désignent des espaces vectoriels sur un corps  $\mathbf{K}$ , et  $u$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ . L'élément neutre de  $\mathbf{E}$  sera noté  $\vec{0}_{\mathbf{E}}$ , celui de  $\mathbf{F}$ ,  $\vec{0}_{\mathbf{F}}$ .

**Proposition 1.4.1.** — THÉORÈME DU RANG — Soit  $\mathbf{E}'$  un supplémentaire du noyau de  $u$ ,  $\text{Ker}(u)$ , alors l'application,

$$u' : \mathbf{E}' \rightarrow \text{Im}(u) ; x \mapsto u(\vec{x}).$$

Est un isomorphisme de  $\mathbf{E}'$  sur l'image de  $u$ .



*Preuve de la proposition 1.4.1. —*

- LINÉARITÉ

$u'$  hérite de la linéarité de  $\vec{u}$ .

- SURJECTIVITÉ

Soit  $\vec{y}$  un élément de  $\text{Im}(u)$ . Montrons qu'il a un antécédent par  $u'$ . Par définition même de  $\text{Im}(u)$ , on dispose d'un élément  $\vec{x}$  de  $\mathbf{E}$  tel que  $\vec{y} = u(\vec{x})$ .  $\vec{x}$  n'est peut être pas élément de  $\mathbf{E}'$ , mais, puisque par hypothèse,  $\mathbf{E} = \mathbf{E}' \oplus \text{Ker}(u)$ , on dispose  $\vec{x}'$  élément de  $\mathbf{E}'$  et  $\vec{z}$  de  $\text{Ker}(u)$  tels que  $\vec{x} = \vec{x}' + \vec{z}$ . Mais alors  $u(\vec{x}) = u(\vec{x}') + u(\vec{z}) = u(\vec{x}')$  et donc  $\vec{y} = u(\vec{x}')$ . Mais comme  $\vec{x}'$  est élément de  $\mathbf{E}'$ ,  $\vec{y} = u'(\vec{x}')$ . Ainsi  $\vec{y}$ , élément quelconque de  $\text{Im}(u)$  admet-il un antécédent  $\vec{x}'$  par  $u'$  et partant,  $u'$  est-elle surjective.

- INJECTIVITÉ

$\text{Ker}(u') = \text{Ker}(u) \cap \mathbf{E}'$ . Or  $\text{Ker}(u)$  et  $\mathbf{E}'$  sont supplémentaires donc  $\text{Ker}(u) \cap \mathbf{E}' = \{\vec{0}_{\mathbf{E}}\}$ . Le noyau de  $u'$  est donc réduit à  $\{\vec{0}_{\mathbf{E}}\}$ , Donc  $u'$  est injective.

Finalement  $u'$  est un isomorphisme de  $\mathbf{E}'$  sur  $\text{Im}(u)$ . □

**Remarque** – L'essentiel de cette proposition s'énonce ainsi :

*L'image d'une application linéaire est isomorphe à tout supplémentaire de son noyau.*

C'est ce résultat qu'il convient de bien retenir, les corollaires classiques du théorème du rang, vus en sup., en découlent immédiatement.

**Corollaire 1.4.2.** — FORMULE DU RANG —

$\mathbf{E}$  est dans ce corollaire supposé de dimension finie ( $\mathbf{F}$  reste de dimension quelconque). On note  $\text{rg}(u)$  le rang de  $u$  qui, par définition est la dimension de l'image de  $u$ . Alors :

$$\text{rg}(u) + \dim(\text{Ker}(u)) = \dim \mathbf{E}.$$

*Preuve du corollaire 1.4.2.* – c'est une conséquence immédiate de 1.4.1. et 1.3.13. □

**Corollaire (du corollaire) 1.4.3.** —  $\mathbf{E}$  et  $\mathbf{F}$  sont, dans ce corollaire, supposés de dimension finie, et de même dimension.

Alors, on a l'équivalence des trois propositions suivantes :

1.  $u$  est bijective.
2.  $u$  est injective.
3.  $u$  est surjective.

*Preuve du corollaire 1.4.3.* — L'application  $u$  est injectif si et seulement si  $\dim(\text{Ker}(u)) = 0$ . Par ailleurs,  $u$  est surjectif si et seulement si  $\text{rg}(u) = \dim \mathbf{F}$ , donc ici si et seulement si  $\text{rg}(u) = \dim \mathbf{E}$ . Or, d'après 1.5.2.,  $\dim(\text{Ker}(u)) = 0$ , équivaut à  $\text{rg}(u) = \dim \mathbf{E}$ . D'où l'équivalence de 2. et 3. L'injectivité et la surjectivité ensemble, équivalent à la bijectivité, donc 1. est équivalent à 2. (ou 3.). □

Le cas où :  $\mathbf{E} = \mathbf{F}$  est très important dans la pratique, 1.4.3. s'écrit alors :

**Corollaire (du corollaire du corollaire) 1.4.4.** — Un endomorphisme d'un espace vectoriel de dimension finie est bijectif si et seulement si il est injectif ou surjectif.

**Remarque** — le résultat est faux en dimension infinie. Donnons un exemple d'endomorphisme surjectif non injectif, un exemple d'endomorphisme injectif non surjectif.

**Proposition 1.4.5.** —  $\mathbf{E}$  est ici de dimension quelconque. soient  $\mathbf{G}$  un sous-espace vectoriel de  $\mathbf{E}$  et  $\mathbf{F}$  et  $\mathbf{F}'$  des supplémentaires de  $\mathbf{G}$ , c'est-à-dire que

$$\mathbf{E} = \mathbf{G} \oplus \mathbf{F}' \text{ et } \mathbf{E} = \mathbf{G} \oplus \mathbf{F}$$

Alors  $\mathbf{F}$  et  $\mathbf{F}'$  sont isomorphes. En particulier si  $\mathbf{G}$  admet un supplémentaire de dimension finie  $p$ , alors tous ses supplémentaires sont de dimension finie  $p$ .<sup>1</sup>

Preuve de la proposition 1.4.5. —

□

Voici un résultat très utile dans la suite.

**Proposition 1.4.6.** — Soient  $\mathbf{E}$ ,  $\mathbf{F}$ ,  $\mathbf{E}'$  et  $\mathbf{F}'$  des espaces vectoriels sur  $\mathbf{K}$  de dimensions finies ou non. Soit  $u$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ .

- Soit  $\varphi$  une application linéaire injective de  $\mathbf{F}$  dans  $\mathbf{F}'$ . Alors :

$$\operatorname{rg}(\varphi \circ u) = \operatorname{rg}(u).$$

- Soit  $\psi$  une application linéaire surjective de  $\mathbf{E}'$  sur  $\mathbf{E}$ . Alors :

$$\operatorname{rg}(u \circ \psi) = \operatorname{rg}(u).$$

Preuve de la proposition 1.4.6. —

- $\operatorname{Im}(\varphi \circ u) = (\varphi \circ u)(\mathbf{E}) = \varphi(u(\mathbf{E}))$ ,  $\varphi$  étant injective, elle induit un isomorphisme de  $u(\mathbf{E})$  sur  $\varphi(u(\mathbf{E}))$ . Donc  $\dim(\varphi(u(\mathbf{E}))) = \dim(u(\mathbf{E}))$ , soit

$$\operatorname{rg}(\varphi \circ u) = \operatorname{rg}(u).$$

- $\operatorname{Im}(u \circ \psi) = u \circ \psi(\mathbf{E}')$ . Or, puisque  $\psi$  est surjective,  $\psi(\mathbf{E}') = \mathbf{E}$ . Donc :  $\operatorname{Im}(u \circ \psi) = u(\mathbf{E})$ . Donc  $\dim(\operatorname{Im}(u \circ \psi)) = \dim(u(\mathbf{E}))$  soit :

$$\operatorname{rg}(u \circ \psi) = \operatorname{rg}(u).$$

D'où le résultat.

□

**Remarque 1.4.7.** — On applique souvent cette proposition dans le cas particulier où  $\varphi$  et  $\psi$  sont des isomorphismes : La composition à droite ou à gauche par un isomorphisme ne change pas le rang d'une application linéaire.

Nous allons donner une application très classique du théorème du rang. Il s'agit d'un résultat du cours qui figure en tant que tel au programme. Il convient de plus de bien en retenir le raisonnement réutilisable dans des circonstances voisines.

**Application** — POLYNÔMES D'INTERPOLATION DE LAGRANGE—

Nous nous intéressons à l'approximation d'une application  $f$  par des polynômes. Plus précisément, nous allons chercher s'il existe un polynôme  $P$ , de degré inférieur ou égal à un entier  $n$ , qui coïncide avec  $f$  en certains points.

Si nous nous donnons deux réels distincts  $a_0$  et  $a_1$ , nous pouvons trouver un et un seul polynôme de degré inférieur ou égal à 1 qui coïncide avec  $f$  en ces points. En effet le graphe d'un polynôme de degré inférieur ou égal à 1 est une droite, et nous savons de longue date que par les deux points  $(a_0, f(a_0))$ ,  $(a_1, f(a_1))$ , passe une et une seule droite. Ce résultat se généralise, nous allons montrer :

Soit  $n$  un entier naturel, soient  $a_0, a_1, \dots, a_n$ ,  $n+1$  réels (resp. complexes) distincts et soient  $b_0, b_1, \dots, b_n$  des réels (resp. complexes) distincts ou non. Alors il existe un et un seul polynôme  $Q$ , élément de  $\mathbf{R}[X]$  (resp. de  $\mathbf{C}[X]$ ), de degré inférieur ou égal à  $n$  tel que, pour tout élément  $i$  de  $\{0, 1, \dots, n\}$ ,  $Q(a_i) = b_i$ .

Prouvons le résultat.

1. On dit parfois que  $p$  est la codimension de  $\mathbf{G}$ .

Soit l'application

$$\vartheta : \mathbf{K}[X]_n \rightarrow \mathbf{K}^{n+1}; P \mapsto (P(a_0), P(a_1), \dots, P(a_n)).$$

L'application  $\vartheta$  est clairement linéaire. Déterminons son noyau. Soit  $P$  élément de  $\text{Ker}\vartheta$ .  $P$  admet  $n+1$  racines distinctes :  $a_0, a_1, \dots, a_n$ , or  $\deg P \leq n$ , donc  $P$  est nul. Le noyau de  $\vartheta$  est réduit à  $\{0\}$ , donc  $\vartheta$  est injective. Mais comme  $\mathbf{K}[X]_n$  et  $\mathbf{K}^{n+1}$  ont même dimension (à savoir  $n+1$ ),  $\vartheta$  est bijective (cf. 1.4.3.). Donc il existe un unique élément  $Q$  de  $\mathbf{K}[X]_n$  dont l'image par  $\vartheta$  soit  $b_0, b_1, \dots, b_n$ , à savoir  $\vartheta^{-1}(b_0, b_1, \dots, b_n)$ . C'est-à-dire qu'il existe un unique élément de  $\mathbf{K}[X]_n$  qui prenne comme valeur  $b_i$  en  $a_i$ , pour  $i = 0, 1, \dots, n$ .

Grâce à cette proposition on montre, qu'étant donnée une application  $f$  d'un intervalle  $I$  de  $\mathbf{R}$  dans  $\mathbf{K}$  et  $n+1$  points distincts de  $I$ ,  $a_0, a_1, \dots, a_n$ , il existe un unique polynôme  $Q_f$  de degré inférieur ou égal à  $n$ , qui coïncide avec  $f$  en  $a_0, a_1, \dots, a_n$ ; c'est  $\vartheta^{-1}(f(a_0), f(a_1), \dots, f(a_n))$ .

$Q_f$  s'appelle le polynôme d'interpolation de  $f$  au points  $a_0, a_1, \dots, a_n$ . Historiquement il servait à déterminer une valeur approchée de  $f$  en un point  $t$ , connaissant les valeurs de  $f$  en des points  $a_0, a_1, \dots, a_n$ . On verra en T.D. comment, sous certaines hypothèses, majorer la quantité  $|Q_f - f|$ .

L'interpolation sert de nos jours aux calculs approchés d'intégrales ou à la résolution approchée de certaines équations différentielles (cf. T.D.)

Donnons une formule explicite de  $Q_f$ .

Revenons au problème initial dont nous gardons les hypothèses et les notations :  $Q = \vartheta^{-1}(b_0, b_1, \dots, b_n)$ . Pour connaître  $\vartheta^{-1}$ , il suffit de connaître ses valeurs sur une base de  $\mathbf{K}^{n+1}$ , nous choisirons bien sûr, comme base, la base canonique,

$(\vec{e}_0, \vec{e}_1, \dots, \vec{e}_n)$  où pour tout élément  $i$  de  $0, 1, \dots, n$ , on note  $\vec{e}_i = \left( \underbrace{0, 0, \dots, 0}_{i+1}, 1, 0, \dots, 0 \right)$ . Posons pour tout élément  $i$  de  $0, 1, \dots, n$ ,  $L_i := \vartheta^{-1}(\vec{e}_i)$ . De  $(b_0, b_1, \dots, b_n) = b_0\vec{e}_0 + b_1\vec{e}_1 + \dots + b_n\vec{e}_n$ , on déduit par linéarité de l'application  $\vartheta^{-1}$ ,

$$Q = \vartheta^{-1}(b_0\vec{e}_0 + b_1\vec{e}_1 + \dots + b_n\vec{e}_n) = b_0L_0 + b_1L_1 + \dots + b_nL_n$$

Reste à déterminer, pour  $i$  élément de  $\{0, 1, \dots, n\}$ ,  $L_i$ . Comme  $\vartheta(L_i) = \vec{e}_i$ ,  $L_i$  est donc le polynôme élément de  $\mathbf{K}[X]_n$  caractérisé par :

- $L_i(a_i) = 1$  ;
- $L_i(a_j) = 0$ , pour tout élément  $j$  de  $\{0, 1, \dots, n\}$ , distinct de  $i$ .

$L_i$  admet donc comme racines les  $a_j$ ,  $j$  distinct de  $i$ , donc est de la forme

$$L_i = c \prod_{\substack{j=0, \dots, n \\ j \neq i}} (X - a_j),$$

où  $c$  est un élément de  $\mathbf{K}$ . De  $L_i(a_i) = 1$ , on déduit  $c = \frac{1}{\prod_{\substack{j=0, \dots, n \\ j \neq i}} (a_i - a_j)}$ .

Finalement :

$$L_i = \frac{\prod_{\substack{j=0, \dots, n \\ j \neq i}} (X - a_j)}{\prod_{\substack{j=0, \dots, n \\ j \neq i}} (a_i - a_j)}, \text{ pour } i = 0, 1, \dots, n \text{ et } Q = b_0L_0 + b_1L_1 + \dots + b_nL_n$$

Notons que  $(L_0, L_1, \dots, L_n)$  est l'image par l'isomorphisme  $\vartheta^{-1}$  de la base canonique de  $\mathbf{K}^n$ . C'est donc, à ce titre, une base de  $\mathbf{K}[X]_n$  (cf. 1.1.10.).

Soit  $P$  un élément quelconque de  $\mathbf{K}[X]_n$ . Décomposons  $P$  dans la base  $(L_0, L_1, \dots, L_n)$ .  $P$  s'écrit :

$$P = c_0L_0 + c_1L_1 + \dots + c_nL_n,$$

avec  $c_j$  élément de  $\mathbf{K}$ , pour  $j = 0, 1, \dots, n$ . Soit  $i$  un élément quelconque de  $\{0, 1, \dots, n\}$ . En substituant  $a_i$  à l'indéterminée  $X$ , dans l'égalité précédente, il vient :

$$P(a_i) = c_0 \underbrace{L_0(a_i)}_0 + c_1 \underbrace{L_1(a_i)}_0 + \dots + c_i \underbrace{L_i(a_i)}_1 + \dots + c_n \underbrace{L_n(a_i)}_0 = c_i.$$

D'où :

$$P = P(a_0)L_0 + P(a_1)L_1 + \dots + P(a_n)L_n$$

## 1.5 Formes linéaires, hyperplans

Pour commencer un rappel.

**Définition-proposition 1.5.1.**—On appelle *forme linéaire* sur  $\mathbf{E}$  toute application linéaire de  $\mathbf{E}$  dans  $\mathbf{K}$ . L'ensemble des formes linéaires est un sous-espace vectoriel de l'espace vectoriel  $K^{\mathbf{E}}$  des applications de l'ensemble  $\mathbf{E}$  dans l'espace vectoriel  $\mathbf{K}$ . On le note fréquemment, pour simplifier  $\mathbf{E}^*$

**Exemple** —

Supposons que  $\mathbf{E}$  admette  $(\vec{e}_i)_{i \in I}$  pour base, l'application qui à un vecteur  $\vec{x}$  de  $\mathbf{E}$  associe sa coordonnée sur  $\vec{e}_j$ , pour un élément quelconque de  $j$  de  $I$ , est une application clairement linéaire, c'est du reste la composante sur  $\vec{e}_j$  de l'isomorphisme réciproque de

$$\Phi : F(I) \rightarrow \mathbf{E}; (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i \vec{e}_i,$$

On l'appelle *forme coordonnée* sur  $\vec{e}_j$  dans la base  $e$ , et on la note  $e_j^*$ .

*Cas de la dimension finie*

On se limite au cas particulier d'un espace vectoriel  $\mathbf{E}_n$  sur  $\mathbf{K}$  de dimension finie  $n$  non nulle. Soit  $(\vec{e}_1, \dots, \vec{e}_n)$  une base de  $\mathbf{E}_n$  notée plus simplement  $e$ .

Nous allons exhiber une base de  $\mathbf{E}_n^*$  particulière. Soit  $\ell$  un élément de  $\mathbf{E}_n^*$  et  $\vec{x}$  un vecteur de  $\mathbf{E}_n$ .

$$\ell(\vec{x}) = \dots$$

Donc la famille  $(e_1^*, \dots, e_n^*)$  est une *famille génératrice* de  $\mathbf{E}_n^*$ .

Montrons sa liberté. Soient  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  un élément de  $\mathbf{K}^n$  tel que :

$$\sum_{i=1}^n \dots$$

Alors...

La famille est donc libre.

On a prouvé le résultat suivant :

**Proposition-définition 1.5.2.** — La famille  $(e_1^*, e_2^*, \dots, e_n^*)$  est une base de  $\mathbf{E}_n^*$ . On l'appelle *base duale* de la base  $e$ , et on la note traditionnellement  $e^*$ .

On déduit de la proposition 1.5.2. que la dimension de  $\mathbf{E}_n^*$  est  $n$  : Ce résultat se retrouve aussi par la dimension de l'ensemble des applications linéaires d'un espace vectoriel de dimension  $n$  dans un espace de dimension  $p$  qui vaut ...

Revenons à notre espace  $\mathbf{E}$  de dimension quelconque.

**Proposition 1.5.3.** — On appelle *hyperplan* de  $\mathbf{E}$ , tout sous-espace vectoriel de  $\mathbf{E}$  qui est le noyau d'une forme linéaire non nulle.

**Exemple 1.5.4** — L'ensemble des matrices éléments de  $\mathcal{M}_n(\mathbf{K})$  de trace nulle est un hyperplan de  $\mathcal{M}_n(\mathbf{K})$ , puisque, on le revera bientôt la trace est une forme linéaire.

Examinons le cas d'un espace vectoriel  $\mathbf{E}_n$  sur  $\mathbf{K}$  de dimension finie  $n$  non nulle.

• Soit  $\ell$  une forme linéaire de  $\mathbf{E}_n$  non nulle. Son image est un sous-espace vectoriel de  $\mathbf{K}$ , c'est donc  $\mathbf{K}$  ou  $\{0\}$ . Comme  $\ell$  est non nulle, c'est  $\mathbf{K}$ , donc  $\text{rg}(\ell) = 1$  par suite la formule du rang dit qu'un hyperplan est de dimension  $n - 1$ .

• Réciproquement.....

Finalement on a :

**Proposition 1.5.5.** — *Les hyperplans d'un espace vectoriel de dimension finie  $n$  non nulle sont exactement ses sous-espaces de dimension  $n - 1$*

Remarquons que dans le cas où la dimension de l'espace est 3, cadre naturel de la géométrie élémentaire, ses hyperplans sont donc ses plans, d'où le nom d'*hyperplan*, notion qui généralise celle de plan d'un espace de dimension 3, à des espaces de dimension quelconque.

Généralisons de même à la dimension quelconque le résultat évident qui dit que dans un espace de dimension 3 les supplémentaires d'un plan sont exactement les droites non incluses dans celui-ci.

**Proposition 1.5.6.** — *Soit  $\mathbf{H}$  un hyperplan de  $\mathbf{E}$  et  $\mathbf{D}$  une droite de  $\mathbf{E}$  non incluse dans  $\mathbf{H}$ , alors  $\mathbf{D}$  et  $\mathbf{H}$  sont supplémentaires.*

*Preuve de 1.5.6.* — On dispose par définition d'une forme linéaire  $\ell$  non nulle telle que  $\mathbf{H} = \text{Ker}(\ell)$  et l'on se donne un vecteur  $\vec{e}$  qui dirige  $\mathbf{D}$ .....



Donc d'après 1.4.5., tout supplémentaire d'un hyperplan  $\mathbf{H}$  est une droite, évidemment non incluse dans  $\mathbf{H}$ . Au tour de la réciproque

**Proposition 1.5.7.** — *Tout supplémentaire d'une droite est un hyperplan.*

*Preuve de 1.5.7.* — Soit  $\mathbf{H}$  le supplémentaire d'une droite  $\mathbf{D}$  dirigée par  $\vec{e}$ .....

D'où le résultat □

**Exemple** — D'après le troisième point de 1.3.8., l'ensemble des polynômes éléments de  $\mathbf{K}[X]$ , divisibles par  $X$ ,  $X\mathbf{K}[X]$ , est un hyperplan de  $\mathbf{K}[X]$ , c'est en effet un sous-espace vectoriel qui admet  $\mathbf{K}[X]_0 = \text{vec}(X^0)$ , espace de dimension 1, comme supplémentaire.

*Équation d'un hyperplan*

Soit  $\mathbf{H}$  un hyperplan de  $\mathbf{E}$ . On dispose donc d'une forme linéaire  $\ell$  non nulle telle que  $H = \text{Ker}(\ell)$ . On dit alors qu'une équation de  $\mathbf{H}$  est  $\ell(\vec{x}) = 0$ , on note

$$\mathbf{H} : \ell(\vec{x}) = 0.$$

Bien sûr l'équation d'un hyperplan  $\mathbf{H}$  de  $\mathbf{E}$  n'est pas unique, mais on dispose du résultat d'unicité suivant :

**Proposition 1.5.8.** — *Soient  $\mathbf{H}$  un hyperplan de  $\mathbf{E}$ , d'équation*

$$\mathbf{H} : \ell(\vec{x}) = 0,$$

*et  $\ell'$  une forme linéaire sur  $\mathbf{E}$ . Alors  $\mathbf{H}$  admet l'équation*

$$\mathbf{H} : \ell'(\vec{x}) = 0.$$

*si et seulement si il existe un élément  $\lambda$  non nul de  $\mathbf{K}$  tel que  $\ell' = \lambda\ell$ .*

**Preuve de la proposition 1.5.8.** —

• **HYPOTHÈSE** *il existe un élément  $\lambda$  non nul de  $\mathbf{K}$  tel que  $\ell' = \lambda\ell$*

.....

- HYPOTHÈSE  $\mathbf{H}$  est le noyau de  $\ell'$

.....

D'où le résultat. □

Examinons le cas de la dimension finie. Soit  $\mathbf{H}$  un hyperplan d'un espace vectoriel  $\mathbf{E}_n$  de dimension finie  $n$  non nulle, d'équation

$$\mathbf{H} : \ell(\vec{x}) = 0.$$

Equipons  $\mathbf{E}_n$  d'une base  $e = (\vec{e}_1, \dots, \vec{e}_n)$  et introduisons la base  $e^*$ , définie en 1.5.2.

Soit un hyperplan défini par une forme linéaire  $\ell$  non nulle,

$$\mathbf{H} : \ell(\vec{x}) = 0.$$

La forme linéaire  $\ell$  se décompose dans  $e^*$  :

$$\ell = a_1 e_1^* + a_2 e_2^* + \dots + a_n e_n^*,$$

où  $a_1, a_2, \dots, a_n$  sont des éléments de  $\mathbf{K}^n$  non tous nuls. Donc si l'on désigne pour  $i = 1, 2, \dots, n$ ,  $x_i$  la  $i^{\text{e}}$  coordonnée d'un vecteur générique  $\vec{x}$  de  $\mathbf{E}_n$  dans la base  $e$ , alors

$$\ell(\vec{x}) = a_1 e_1^*(\vec{x}) + a_2 e_2^*(\vec{x}) + \dots + a_n e_n^*(\vec{x}) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n.$$

On obtient ainsi, ce que l'on appelle l'équation de  $\mathbf{H}$  dans la base  $e$  :

$$\mathbf{H} : a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0.$$

Réciproquement, pour tout  $n$ -uplet  $(a_1, a_2, \dots, a_n)$  non nul d'éléments de  $K$ , l'équation

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0$$

définit un hyperplan, le noyau de la forme linéaire  $a_1 e_1^* + a_2 e_2^* + \dots + a_n e_n^*$ .

Dans le cas de la dimension 3, on retrouve en particulier l'équation d'un plan vectoriel telle qu'elle a été vue dans les petites classes, et dans le cas de la dimension 2 celle d'une droite vectorielle.

D'après 1.5.8., Si  $\mathbf{H}$  est un hyperplan de  $\mathbf{E}$  d'équation

$$\mathbf{H} : a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0,$$

alors

$$a'_1 x_1 + a'_2 x_2 + \dots + a'_n x_n = 0$$

est une équation de  $\mathbf{H}$  si et seulement si, il existe un élément  $\lambda$  non nul de  $\mathbf{K}$  tel que :

$$(a'_1, \dots, a'_n) = \lambda(a_1, \dots, a_n).$$

Restons en dimension finie non nulle  $n$ . Et étudions l'intersection de  $m$  hyperplans.

**Proposition 1.5.9.** — Soient  $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m$  des hyperplans d'un espace vectoriel  $\mathbf{E}_n$  de dimension finie  $n$  non nulle. Alors l'intersection de  $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m$  est un espace vectoriel de dimension au moins  $n - m$  :

$$\dim \bigcap_{i=1}^m \mathbf{H}_i \geq n - m.$$

**Remarque** —

- Cette proposition n'a d'intérêt que si  $2 \leq m < n$ .
- On peut avoir égalité dans la proposition 1.5.9. Considérons  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base de  $\mathbf{E}_n$ .....

*Preuve de la proposition 1.5.9.* — Nous allons donner une preuve par récurrence, toutefois nous proposerons dans les feuilles d'exercices une preuve directe plus instructive. Soit  $(R_n)$  la propriété<sup>1</sup> :

« Dans un espace de dimension non nulle  $n$ , pour tout entier  $m \geq 2$  l'intersection de  $m$  hyperplans est un espace vectoriel de dimension au moins  $n - m$ . »

- $(R_1)$  est trivialement vraie d'après la remarque.
- Soit  $d \geq 1$ . On suppose  $(R_d)$ . Soient alors un entier  $m \geq 2$  et  $\mathbf{H}_1, \dots, \mathbf{H}_m$ , des hyperplans d'un espace vectoriel  $\mathbf{E}_{d+1}$  de dimension  $d + 1$ .....

D'où  $(R_{d+1})$

Ainsi, vient-on de prouver par récurrence que  $(R_n)$  est vraie pour tout entier  $n \geq 1$ . D'où le résultat □

**Remarque** — On garde les notations de 1.5.9. et l'on note pour  $i = 1, \dots, m$ ,  $\ell_i$  une forme linéaire dont le noyau est  $\mathbf{H}_i$ . Alors  $\dim \bigcap_{i=1}^m \mathbf{H}_i = n - m$  si et seulement si les formes linéaires  $\ell_1, \dots, \ell_m$  sont indépendantes. Ce résultat sera proposé en exercice.

**Proposition 1.5.10.** — Soit  $\mathbf{F}$  un sous espace vectoriel de  $\mathbf{E}_n$ ,  $\mathbf{K}$ -espace vectoriel de dimension  $n$  non nul. On suppose que  $\dim(\mathbf{F}) = n - m$ , avec  $m$  entier compris au sens large entre 1 et  $m$ . Alors il existe  $m$  hyperplans  $\mathbf{H}_1, \dots, \mathbf{H}_m$  de  $\mathbf{E}_m$  dont l'intersection est égale à  $\mathbf{F}$  :

$$\mathbf{F} = \bigcap_{i=1}^m \mathbf{H}_i.$$

1. On peut aussi faire une récurrence sur le nombre d'hyperplans



*Preuve de la proposition 1.5.10. —*

.....



**Exemple** — Ainsi dans un espace de dimension 3, une droite est-elle l'intersection de deux plans.

# MATRICES, ENDOMORPHISMES

## 2.1 Introduction

Dans cette sous-partie,  $n$  et  $p$  désignent des entiers naturels non nuls.

Pour simplifier l'écriture nous désignerons systématiquement un élément  $(a_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,p}}$  de  $\mathcal{M}_{n,p}(\mathbf{K})$ , par  $A$ , un élément  $(b_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,p}}$ , par  $B$ , un élément  $(c_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,p}}$ , par  $C$ , etc. Nous noterons aussi le terme de la  $i^{\text{e}}$  ligne et de la  $j^{\text{e}}$  colonne d'un élément  $M$  de  $\mathcal{M}_{n,p}(\mathbf{K})$ ,  $M[i, j]$ , cette notation sera privilégiée pour des éléments de  $\mathcal{M}_{n,p}(\mathbf{K})$  exprimés au moyens d'autres matrices, par exemple :  $({}^tAB)[i, j]^1$ .

De plus on se donne dans cette sous partie  $\mathbf{E}$  un  $\mathbf{K}$ -espace vectoriel de dimension  $p$ , muni d'une base  $(\vec{e}_1, \dots, \vec{e}_p)$ , notée  $e$  et  $\mathbf{F}$  un  $\mathbf{K}$ -espace vectoriel de dimension  $n$ , muni d'une base  $(\vec{f}_1, \dots, \vec{f}_n)$ , notée  $f$ .

Nous allons livrer presque sans preuve les principaux résultats de MPSI.

L'an passé l'ensemble des matrices a été muni « naturellement » d'une addition «  $+$  » et d'une opération de  $\mathbf{K}$  sur  $\mathcal{M}_{n,p}(\mathbf{K})$ , «  $\cdot$  ». Muni de ces deux lois  $\mathcal{M}_{n,p}(\mathbf{K})$  est un espace vectoriel<sup>2</sup>.

Les matrices interviennent dans de nombreux domaines des sciences, pour représenter un graphe, une image numérisée, un produit scalaire etc., mais une utilisation capitale est la représentation des applications linéaires. En effet d'après 1.2.2., une application linéaire est entièrement définie par la donnée des images d'une base, information qui peut être représentée par une matrice. Précisément :

**Définition 2.1.1.** — MATRICE D'UNE APPLICATION LINÉAIRE —

Soit  $u$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ . On appelle matrice de  $u$  dans les bases  $e$  et  $f$ , l'élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ , dont le coefficient situé sur la  $i^{\text{e}}$  ligne et de la  $j^{\text{e}}$  colonne est.....

pour  $i = 1, \dots, n$  et  $j = 1, \dots, p$ .

La matrice de  $u$  dans les bases  $e$  et  $f$  sera notée  $\text{Mat}_{e,f}(u)$  et dans le cas où  $\mathbf{F} = \mathbf{E}$  et  $f = e$ , plus simplement,  $\text{Mat}_e(u)$ .

Résumons :

$$\text{Mat}_{e,f}(u) = \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,p} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,p} \\ \vdots & \vdots & & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,p} \end{pmatrix}$$

$$u(\vec{e}_j) = \sum_{\dots} \dots\dots\dots$$

**Exemple** —

1. Soit  $u$  l'application linéaire  $u$  de  $\mathbf{R}^3$  dans  $\mathbf{R}^2$  suivante :

$$u : \mathbf{R}^3 \rightarrow \mathbf{R}^2 ; (x, y, z) \mapsto (x + y + z, 2x + 2y).$$

On munit  $\mathbf{R}^3$  et  $\mathbf{R}^2$  de leur bases canoniques respectives,  $\mathcal{B}_3$  et  $\mathcal{B}_2$ . Alors

1. Une autre notation possible et peut-être plus naturelle, est  $M(i, j)$ , puisque, rappelons le, un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$  est une application de  $\llbracket 1, p \rrbracket \times \llbracket 1, n \rrbracket$  dans  $\mathbf{K}$  et  $M(i, j)$  n'est rien d'autre que l'image de  $(i, j)$  par  $M$ .

2. Ces deux lois sont les lois usuelles sur l'ensemble des applications de  $\llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$  dans  $\mathbf{K}$  qu'est  $\mathcal{M}_{n,p}(\mathbf{K})$

$$\text{Mat}_{\mathcal{B}_3, \mathcal{B}_2}(u) = \dots$$

2. Soient un entier  $n \geq 1$  et l'endomorphisme de  $\mathbf{R}[X]_n$   $P \mapsto P(X+1)$ . La matrice  $M$  de  $u$  dans la base canonique de  $\mathbf{R}[X]_n$  est :

3. Soit  $\mathcal{B}_c$  la base canonique de  $\mathcal{M}_2(\mathbf{R})$  indexée par  $\{1, 2, 3, 4\}$  comme suit :

$$\mathcal{B}_c = \left( \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

La matrice dans la base  $\mathcal{B}_c$  de l'endomorphisme  $\mathcal{T}$  de  $\mathcal{M}_2(\mathbf{R})$  qui à une matrice associe sa transposée (transposition) est :

Il est alors remarquable que les lois «  $+$  » et «  $\cdot$  » naturelles définies sur  $\mathcal{M}_{n,p}(\mathbf{K})$  soient compatibles avec l'addition des applications linéaires et la multiplication d'une application linéaire par un élément de  $\mathbf{K}$ , de façon précise et théorique :

**Proposition 2.1.2.** — *Avec les notations de 2.1.1., l'application*

$$\text{Mat}_{e,f} : \mathcal{L}(\mathbf{E}, \mathbf{F}) \rightarrow \mathcal{M}_{n,p}(\mathbf{K}) ; v \mapsto \text{Mat}_{e,f}(v)$$

*est .....*

Concrètement pour tout  $u_1$  et tout  $u_2$  éléments de  $\mathcal{L}(\mathbf{E}, \mathbf{F})$ , tout  $\lambda$  élément de  $\mathbf{K}$ ,

$$\boxed{\text{Mat}_{e,f}(u_1 + u_2) = \text{Mat}_{e,f}(u_1) + \text{Mat}_{e,f}(u_2)}$$

$$\boxed{\text{Mat}_{e,f}(\lambda u_1) = \lambda \text{Mat}_{e,f}(u_1)}$$

$$\boxed{\text{Mat}_{e,f} \text{ est une bijection}}$$

De plus la dimension de  $\mathcal{L}(\mathbf{E}, \mathbf{F})$  est donc celle de  $\mathcal{M}_{n,p}(\mathbf{K})$ , soit ....., ce que donne le décompte des vecteurs de la base canonique de  $\mathcal{M}_{n,p}(\mathbf{K})$  vue en 1.1.

Il est alors naturel de définir une multiplication entre matrices qui soit compatible avec la composition des applications linéaires.

Soit un entier  $q \geq 1$ . Soit  $A$  un élément de  $\mathcal{M}_{q,n}(\mathbf{K})$  et  $B$  un élément  $\mathcal{M}_{n,p}(\mathbf{K})$ .

**Définition 2.1.3** — On appelle *produit de la matrice  $A$  par  $B$* , noté  $A \times B$  ou plus simplement  $AB$ , l'élément  $C$  de  $\mathcal{M}_{q,p}(\mathbf{K})$ , défini par

$$c_{i,j} = \sum_{\dots} \dots\dots\dots$$

pour tout  $i \in \{1, \dots, q\}$  et tout  $j \in \{1, \dots, p\}$ .

Avec cette définition si  $\mathbf{G}$  est un  $\mathbf{K}$ -espace vectoriel de dimension  $q$ , muni d'une base  $g$ , et  $v \in \mathcal{L}(\mathbf{F}, \mathbf{G})$ ,  $u \in \mathcal{L}(\mathbf{E}, \mathbf{F})$ , alors :

$$\boxed{\text{Mat}_{e,g}(v \circ u) = \text{Mat}_{f,g}(v) \text{Mat}_{e,f}(u)}$$

$$\mathbf{E} \longrightarrow \mathbf{F} \longrightarrow \mathbf{G}$$

Grossièrement dit, la matrice du produit (de composition) est le produit des matrices

On en déduit en particulier pour les matrices carrées le théorème de structure suivant :

**Proposition 2.1.4.** —

1.  $(\mathcal{M}_n(\mathbf{K}), +, \times)$  est un anneau (non commutatif).

2. L'application

$$\text{Mat}_e : \mathcal{L}(\mathbf{E}) \rightarrow \mathcal{M}_n(\mathbf{K}); v \mapsto \text{Mat}_e(v)$$

est un .....

3. Pour tout endomorphisme  $u$  de  $\mathbf{E}$ ,  $\text{Mat}_e(u)$  est inversible si et seulement si  $u$  est un automorphisme et si c'est le cas :

$$(\text{Mat}_e(u))^{-1} = \text{Mat}_e(u^{-1}).$$

$\text{Mat}_e$  induit donc ...

*Preuve de 2.1.4.* —

On se bornera à rappeler ce qui doit être prouvé, pour les détails voir cours de sup.

Pour montrer que  $(\mathcal{M}_n(\mathbf{K}), +, \times)$  est un anneau il faut prouver :

Pour montrer que  $\text{Mat}_e$  est un.....

Enfin...

### Exercice —

Soient  $n$  un entier et  $M$  l'élément de  $\mathcal{M}_n(\mathbf{R})$  suivant :

$$M = \left( \binom{j-1}{i-1} \right)_{\substack{i=1,\dots,n \\ j=1,\dots,n}},$$

où l'on convient que pour tout couple  $(k, n)$  d'entiers naturels, si  $k > n$ , alors le coefficient binomial  $\binom{n}{k}$  est nul. Déterminer  $M^{-1}$ .

*Solution* —

**Remarque** — L'isomorphisme entre  $\mathcal{L}(\mathbf{E})$  et  $\mathcal{M}_p(\mathbf{K})$  et 1.4.4. permet de montrer pour un élément  $M$  de  $\mathcal{M}_p(\mathbf{K})$  l'équivalence des trois propriétés suivantes :

- i.  $M$  est inversible.
- ii.  $M$  est inversible à droite.
- iii.  $M$  est inversible à gauche.

On vient de voir :

- 1.  $(\mathcal{M}_n(\mathbf{K}), +, \cdot)$  est un  $\mathbf{K}$ -espace vectoriel.
- 2.  $(\mathcal{M}_n(\mathbf{K}), +, \times)$  est un anneau.
- 3. De plus pour tout  $A$  tout  $B$  éléments de  $\mathcal{M}_n(\mathbf{K})$  et tout  $\lambda$  élément de  $\mathbf{K}$ ,

$$\lambda \cdot (A \times B) = (\lambda \cdot A) \times B = A \times (\lambda \cdot B)$$

La dernière propriété autorise à noter  $\lambda \cdot (A \times B)$ , plus simplement  $\lambda AB$ .

Les trois propriétés font de  $(\mathcal{M}_n(\mathbf{K}), +, \times, \cdot)$  ce que l'on appelle une  **$\mathbf{K}$ -algèbre** (cf. récapitulatif sur les structures algébriques).

On montre en utilisant le cours de sup., que de même  $(\mathcal{L}(\mathbf{E}), +, \circ, \cdot)$  est une algèbre.

**Exercice** — Donner d'autres exemples d'algèbres rencontrées en MPSI.

*Réponse* —

L'application  $\text{Mat}_e$  est on l'a vu à la fois un morphisme d'anneaux et une application linéaire de  $\mathcal{L}(\mathbf{E})$  dans  $\mathcal{M}_n(\mathbf{K})$ , on dit alors que c'est un *morphisme d'algèbres* et même puisque cette application est bijective, *un isomorphisme*

*d'algèbres*. Pour montrer que c'est un morphisme d'algèbres, il suffit de vérifier les propriétés suivantes :

L'isomorphisme entre  $\mathcal{L}(\mathbf{E})$  et  $\mathcal{M}_p(\mathbf{K})$  permet de raisonner indifféremment sur les endomorphismes ou sur les matrices associées dans une base donnée, suivant ce qui semble le plus simple.

La matrice d'une application linéaire contient toute les informations sur cette application linéaire, elle permet le calcul de l'image d'un vecteur quelconque.

Soit toujours  $u$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ . Soit un élément  $\vec{x}$  de  $\mathbf{E}$  dont l'image par  $u$  sera notée  $\vec{y}$ . On note  $X$  le vecteur colonne des coordonnées de  $\vec{x}$  dans la base  $e$  et  $Y$  celui de  $\vec{y}$  dans  $f$ . alors en notant  $M = \text{Mat}_{e,f}(u)$  :

$$\boxed{Y = MX} \quad (\text{II.2})$$

**Définition, Proposition 2.1.5** — ENDOMORPHISME CANONIQUEMENT ASSOCIÉ — Soit  $M$  un élément de  $\mathcal{M}_p(\mathbf{K})$ . On appelle endomorphisme de  $\mathcal{M}_{p,1}(\mathbf{K})$ , canoniquement associé à  $M$ , l'endomorphisme

$$m : \mathcal{M}_{p,1}(\mathbf{K}) \rightarrow \mathcal{M}_{p,1}(\mathbf{K}) ; X \mapsto MX,$$

*c'est l'endomorphisme de  $\mathcal{M}_{p,1}(\mathbf{K})$  dont la matrice dans la base canonique est  $M$ .*

*Preuve de la proposition 2.1.5.* — Soient  $j \in \{1, \dots, p\}$  et  $E_j$  le  $j^{\text{e}}$  vecteur de la base canonique.

$$m(E_j) = ME_j = \dots\dots$$

Donc  $m(E_j)$  est la  $j^{\text{e}}$  colonne de  $M$  qui est dans la base canonique de  $\mathcal{M}_{p,1}(\mathbf{K})$  son propre vecteur coordonnées !

Donc la matrice de  $m$  dans la base canonique est  $M$ . □

Il arrive que l'on identifie sans vergogne  $M$  et l'endomorphisme  $m$  de  $\mathcal{M}_{p,1}(\mathbf{K})$ , canoniquement associé, du reste on parle de l'image de  $M$  et du noyau de  $M$  qui sont par définition respectivement l'image de  $m$  et le noyau de  $m$ .

CHANGEMENT DE BASES.

Commençons par une définition. On se donne au préalable  $(\vec{e}'_1, \dots, \vec{e}'_p)$  une (autre) base de  $\mathbf{E}$ , notée  $e'$ .

**Définition 2.1.6.** — On appelle matrice de passage de la base  $e$  à la base  $e'$  l'élément de  $\mathcal{M}_p(\mathbf{K})$  dont le coefficient situé sur la  $i^{\text{e}}$  ligne et de la  $j^{\text{e}}$  colonne est.....

*pour  $i = 1, \dots, p$  et  $j = 1, \dots, p$ .*

*La matrice de passage de la base  $e$  à la base  $e'$  sera notée  $P_e^{e'}$ .*

$$P_e^{e'} = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,p} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,p} \\ \vdots & \vdots & & \vdots \\ c_{p,1} & c_{p,2} & \cdots & c_{p,p} \end{pmatrix}$$

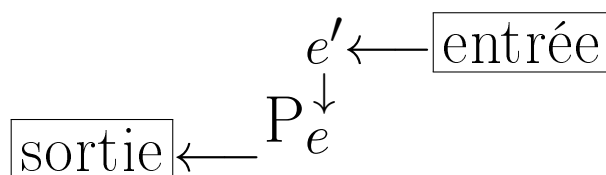
$$\vec{e}'_j = \sum_{\dots} \dots\dots\dots$$

On remarque alors que  $P_e^{e'}$  n'est rien d'autre que la matrice.....

Donc si  $X$  est le vecteur colonne des coordonnées d'un vecteur  $\vec{x}$  de  $\mathbf{E}$  dans la base  $e$  et  $X'$  le vecteur colonne de ses coordonnées dans la base  $e'$ , alors d'après.....

$$X = P_e^{e'} X'$$

Illustrons :



Le second point de la proposition 2.1.4. nous apprend que si  $e''$  est une base de  $\mathbf{E}$ , alors

$$P_{e''}^{e'} P_{e'}^e = P_{e''}^e$$

$$\underbrace{\begin{matrix} \mathbf{E} & \xrightarrow{\text{id}_{\mathbf{E}}} & \mathbf{E} & \xrightarrow{\text{id}_{\mathbf{E}}} & \mathbf{E} \\ e & & e & & e \\ & P_e^e & & P_e^e & \\ & \underbrace{\hspace{1.5cm}} & & & \\ & P_e^e & & & \end{matrix}}_{P_e^e}$$

Le troisième point lui, nous apprend alors que  $P_e^{e'}$  est inversible et :

$$\left(P_e^{e'}\right)^{-1} = P_{e'}^e$$

ce que donne aussi la précédente formule et le fait évident que  $P_e^e = I_n$



Considérons à présent une (autre) base  $f$  de  $\mathbf{F}$ , toujours une application linéaire  $u$  de  $\mathbf{E}$  dans  $\mathbf{F}$ . On pose :

$$M = \text{Mat}_{e,f}(u) ; M' = \text{Mat}_{e',f'}(u).$$

Alors l'interprétation des matrices de passage en termes de matrice de l'identité donne :

$$\boxed{M' = P \cdots MP \cdots}$$

$$\underbrace{\begin{array}{ccccc} \mathbf{E} & \xrightarrow{\text{id}_{\mathbf{E}}} & \mathbf{E} & \xrightarrow{u} & \mathbf{F} & \xrightarrow{\text{id}_{\mathbf{F}}} & \mathbf{F} \\ e' & & e & & f & & f' \\ & & P_e^e & & M & & P_f^f \end{array}}_{M'}$$

En particulier, si  $\phi$  est un endomorphisme de  $\mathbf{E}$ ,  $M = \text{Mat}_e(\phi)$ ,  $M' = \text{Mat}_{e'}(\phi)$  et  $P = P_e^{e'}$ . Alors :

$$\boxed{M' = P \ M P}$$

## 2.2 Trace

Nous allons commencer par rappeler la définition de la trace d'une matrice.

**Définition 2.2.1.** — Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . On appelle trace de  $M$  la quantité  $\sum_{i=1}^n m_{i,i}$ . On notera  $\text{Tr}(M)$  la trace de  $M^1$ .

Il a été vu en sup., le résultat élémentaire suivant.

**Proposition 2.2.2.** — l'application  $\text{Tr} : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathbf{R} ; M \mapsto \text{Tr}(M)$  est une forme linéaire sur  $\mathcal{M}_n(\mathbf{R})$ .

La trace joue un rôle important en algèbre, comme nous allons le découvrir à présent, mais elle joue également un rôle important en calcul différentiel.

Maintenant une définition essentielle.

**Définition 2.2.3.** — Soient  $A$  et  $B$  des éléments de  $\mathcal{M}_n(\mathbf{K})$ . On dit que  $A$  est semblable à  $B$ , si par définition, il existe un élément  $P$  de  $\text{GL}_n(\mathbf{R})$  tel que :  $B = PAP^{-1}$ .

Vient alors la proposition suivante dont la preuve, évidente, sera laissée en exercice :

**Proposition 2.2.4.** — La relation sur  $\mathcal{M}_n(\mathbf{R})$ , « être semblable » (relation de similitude) est une relation d'équivalence.

Il découle de cette proposition et du caractère symétrique d'une relation d'équivalence que l'on peut employer des expressions du type « les matrices  $A$  et  $B$  sont semblables ».

La relation de similitude, apparaît naturellement lors de changements de bases dans l'écriture matricielle d'un endomorphisme, précisément :

**Proposition 2.2.5.** — Soit  $u$  un endomorphisme d'un  $\mathbf{K}$ -espace vectoriel  $\mathbf{E}$  de dimension  $n$ , dont la matrice dans une base  $\mathcal{B}$  sera notée  $A$ . Soit  $B$  une matrice, élément de  $\mathcal{M}_n(\mathbf{K})$ . Les matrices  $A$  et  $B$  sont semblables si et seulement si, il existe une base  $\mathcal{B}'$  de  $\mathbf{E}$ , telle que  $B = \text{Mat}_{\mathcal{B}'}(u)$ .

*Preuve de la proposition 2.2.5.* —

• Supposons qu'il existe une base  $\mathcal{B}'$  de  $\mathbf{E}$ , telle que  $B = \text{Mat}_{\mathcal{B}'}(u)$ . Alors la formule de changement de base pour un endomorphisme que nous venons de revoir, donne que  $B = PAP^{-1}$ , où  $P = P_{\mathcal{B}'}^{\mathcal{B}}$ . Donc  $B$  est semblable à  $A$ .

1. Le programme autorise une seconde notation  $\text{tr}(M)$ .

- Supposons que  $B$  soit semblable à  $A$ . On dispose donc d'un élément  $P$  de  $\mathrm{GL}_n(\mathbf{K})$  tel que  $B = PAP^{-1}$ .

....

On retiendra le résultat utilisé ici : *Toute matrice inversible peut être vue comme une matrice de passage d'une base donnée à une certaine base.*

Donnons un résultat technique que nous allons exploiter par la suite.

**Proposition 2.2.6.** — *Soient  $A$  et  $B$  des éléments de  $\mathcal{M}_n(\mathbf{R})$ . On a :*

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

*Preuve de la proposition 2.2.6.* —

On retiendra :

$$\mathrm{Tr}(AB) = \sum$$

La propriété 2.2.6. caractérise à une constant multiplicative près la trace, comme nous allons le voir dans l'exercice qui suit.

**Exercice 2.2.7.** —

1. Montrer que l'application  $\varphi : \mathcal{M}_n(\mathbf{R}) \times \mathcal{M}_n(\mathbf{R}) \rightarrow \mathbf{R}; (A, B) \mapsto \mathrm{Tr}({}^tAB)$  est un produit scalaire sur  $\mathcal{M}_n(\mathbf{R})$ .
2. Soit  $\ell$  une forme linéaire sur  $\mathcal{M}_n(\mathbf{R})$ , telle que pour tout  $A$  et tout  $B$  éléments de  $\mathcal{M}_n(\mathbf{R})$ ,  $\ell(AB) = \ell(BA)$ . Montrer qu'il existe un réel  $\lambda$  tel que  $\ell = \lambda \mathrm{Tr}$ .

*Solution de l'exercice 2.2.7. —*

On retiendra à propos de cet exercice le résultat au programme de MPSI :

*Pour tout  $i$  tout  $j$  tout  $k$  et tout  $\ell$ , éléments de  $\{1, \dots, p\}$ ,*

$$E_{i,j}E_{k,\ell} =$$

**Proposition 2.2.8.** — *Des éléments  $A$  et  $B$  de  $\mathcal{M}_n(\mathbf{K})$ , semblables ont même trace.*

*Preuve de la proposition 2.2.8. —*

**Mise en garde** — La réciproque de la proposition 2.2.8. est fausse : des éléments de  $\mathcal{M}_n(\mathbf{K})$ , de même trace, ne sont pas nécessairement semblables. Par exemple, ...

Les deux propositions 2.2.5. et 8. conduisent à la définition qui suit.

**Proposition-définition 2.2.9.** — *Soit  $u$  un endomorphisme d'un  $\mathbf{K}$ -espace vectoriel  $\mathbf{E}$  de dimension  $n$ . Toutes les matrices de  $u$  ont même trace; la valeur commune de leur trace s'appelle trace de  $u$  et se note  $\text{Tr}(u)$  (ou  $\text{tr}(u)$ ).*

**Proposition 2.2.10.** — *Soit  $E$  un  $\mathbf{K}$ -espace vectoriel de dimension  $n$ . Soient  $u$  et  $v$  des éléments de  $\mathcal{L}(\mathbf{E})$*

1.  $\text{Tr} : \mathcal{L}(\mathbf{E}) \rightarrow \mathbf{K}; w \mapsto \text{Tr}(w)$  est linéaire.
2.  $\text{Tr}(u \circ v) = \text{Tr}(v \circ u)$ .

*Preuve de la proposition 2.2.10. —*

Le premier point résulte de la linéarité de la trace des matrices, le second vient de la propriété analogue sur les matrices et de 2.2.6. et de ce que la matrice du produit de composition de deux endomorphismes est le produit de leurs matrices.

□

**Exercice 2.2.11.** — Soit  $\mathbf{E}$  un  $\mathbf{K}$ -espace vectoriel de dimension  $n$ . Soient  $\mathbf{F}$  et  $\mathbf{G}$  des sous-espaces de  $\mathbf{E}$  supplémentaires. Soient enfin  $s$  la symétrie par rapport à  $\mathbf{F}$  suivant  $\mathbf{G}$  et  $\pi$  la projection sur  $\mathbf{F}$  suivant  $\mathbf{G}$ . Déterminer la trace de  $s$  et de  $\pi$ .

*Solution de l'exercice 2.2.11.* —

## 2.3 Rang

Dans ce paragraphe  $n$  et  $p$  désignent des entiers naturels non nuls.

**Notations** — Pour tout entier naturel  $k$  non nul nous noterons  $\mathcal{B}_k$  la base canonique de  $\mathcal{M}_{k,1}(\mathbf{K})$ . Pour  $A, B, M, D_1$  etc. éléments de  $\mathcal{M}_{k_1,k_2}(\mathbf{K})$ , où  $k_1$  et  $k_2$  sont des éléments de  $\mathbf{N}^*$  distincts ou non, nous noterons par la minuscule correspondante :  $a, b, m, d_1$  etc., les applications linéaires de  $\mathcal{M}_{k_2,1}(\mathbf{K})$  dans  $\mathcal{M}_{k_1,1}(\mathbf{K})$ , dont ils sont les matrices,  $\mathcal{M}_{k_2,1}(\mathbf{K})$  étant muni de  $\mathcal{B}_{k_2}$  et  $\mathcal{M}_{k_1,1}(\mathbf{K})$  de  $\mathcal{B}_{k_1}$ . Concrètement, par exemple,

$$m : \mathcal{M}_{k_2,1}(\mathbf{K}) \rightarrow \mathcal{M}_{k_1,1}(\mathbf{K}) ; X \mapsto MX$$

D'abord rappelons brutalement une définition :

**Définition 2.3.1.** — On appelle *rang* d'un élément  $M$  de  $\mathcal{M}_{n,p}(\mathbf{K})$ , le...

Le rang de  $M$  se note  $\text{rg}(M)$ .

Presque immédiatement vient le résultat suivant.

**Proposition 2.3.2.** — Soit  $M$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ . Le rang de  $M$  est égal au rang de l'application linéaire

$$m : \mathcal{M}_{p,1}(\mathbf{K}) \rightarrow \mathcal{M}_{n,1}(\mathbf{K}) ; X \mapsto MX.$$

*Preuve de la proposition 2.3.2* — Notons dans l'ordre  $C_1, C_2, \dots, C_p$  les  $p$  colonnes de  $M$ .

**Corollaire 2.3.3.** — Soient  $\mathbf{E}$  et  $\mathbf{F}$  des  $\mathbf{K}$ -espaces vectoriels de dimensions respectives  $p$  et  $n$  et  $u$  un élément de  $\mathcal{L}(\mathbf{E}, \mathbf{F})$ . On considère  $\mathcal{B}$  et  $\mathcal{B}'$  des bases respectivement de  $\mathbf{E}$  et de  $\mathbf{F}$  et l'on note  $M := \text{Mat}_{\mathcal{B}, \mathcal{B}'}(u)$ . Le rang de la matrice  $M$  est égal au rang de  $u$ .

*Preuve du corollaire 2.3.3.* — Soit  $\varphi_p$  l'isomorphisme de  $\mathbf{E}$  sur  $\mathcal{M}_{p,1}(\mathbf{K})$  qui à un vecteur de  $\mathbf{E}$  associe le vecteur colonne de ses coordonnées dans  $\mathcal{B}$  et  $\varphi_n$  l'isomorphisme de  $\mathbf{F}$  sur  $\mathcal{M}_{n,1}(\mathbf{K})$  qui à un vecteur de  $\mathbf{F}$  associe le vecteur colonne de ses coordonnées dans  $\mathcal{B}'$ .

Avec les notations de 2.3.2.,

$$u = \dots$$

Donc finalement  $\text{rg}(u) = \text{rg}(m)$ . □

Ce corollaire permet dans la pratique d'étudier le rang des applications linéaires en étudiant le rang de leur matrice. Par exemple :

**Exercice 2.3.4.** —

1. Soit  $\mathbf{E}$  un  $\mathbf{K}$ -espace vectoriel de dimension  $n$ . Soient  $\mathbf{F}$  et  $\mathbf{G}$  des sous-espaces de  $\mathbf{E}$  supplémentaires. Soient enfin la projection sur  $\mathbf{F}$  suivant  $\mathbf{G}$ . Déterminer le rang de  $\pi$ .
2. Soient  $A$  et  $A'$  des éléments de  $\mathcal{M}_{n,1}(\mathbf{K})$  non nul. On pose  $M := A {}^t A'$ . Quel est le rang de  $M$ .

*Solution de l'exercice 2.3.4.* —

Nous avons vu la notion de matrices carrées semblables, c'est-à-dire de matrices carrées représentant le même endomorphisme. Nous allons définir une relation plus générale entre éléments de  $\mathcal{M}_{n,p}(\mathbf{K})$  qui traduit le fait qu'ils représentent la même application linéaire. Mettons cela en forme.

**Définition 2.3.5.** — Soient  $A$  et  $B$  des éléments de  $\mathcal{M}_{n,p}(\mathbf{K})$ . On dit que  $A$  est équivalente à  $B$ , si par définition, il existe un élément  $P$  de  $\text{GL}(\mathbf{R})$  et un élément  $Q$  de  $\text{GL}(\mathbf{R})$  tels que :  $B = PAQ^{-1}$  (ou, si l'on préfère,  $BQ = PA$ ). .

Evidemment on a :

**Proposition 2.3.6.** — *La relation sur  $\mathcal{M}_{n,p}(\mathbf{R})$ , « être équivalente » est, comme son nom l'indique, une relation d'équivalence.*

Il découle de cette proposition et du caractère symétrique d'une relation d'équivalence que l'on peut employer des expressions du type « les matrices  $A$  et  $B$  sont équivalentes ».

**Remarque** — Des matrices, éléments de  $\mathcal{M}_n(\mathbf{K})$  semblables sont équivalentes. La réciproque est fausse, nous verrons que  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  et  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  sont équivalentes, cf. 2.3.10., elles ne sont pas semblables puisque de traces distinctes.

**Proposition 2.3.7.** — *Soit  $u$  une application linéaire d'un  $\mathbf{K}$ -espace vectoriel  $\mathbf{E}$  de dimension  $p$  dans un  $\mathbf{K}$ -espace vectoriel  $\mathbf{F}$  de dimension  $n$ . Une base  $e$  de  $\mathbf{E}$  et une base  $f$  de  $\mathbf{F}$  ayant été choisies, on pose  $A := \text{Mat}_{e,f}(u)$ . Soit  $B$  une matrice, élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ . Les matrices  $A$  et  $B$  sont équivalentes si et seulement si, il existe une base  $e'$  de  $\mathbf{E}$  et une base  $f'$  de  $\mathbf{F}$ , telles que  $B := \text{Mat}_{e',f'}(u)$ .*

*Preuve de la proposition 2.3.7.* —

• Supposons qu'il existe une base  $e'$  de  $\mathbf{E}$  et une base  $f'$  de  $\mathbf{F}$ , telles que  $B := \text{Mat}_{e',f'}(u)$ . Alors la formule de changement de bases pour une application linéaire donne :

$$B = PAQ^{-1},$$

où  $P = P_{f'}^f$ , et  $Q = P_{e'}^e$  ; donc  $B$  est équivalente à  $A$ .

Supposons  $B$  équivalente à  $A$ . On dispose donc d'un élément  $P$  de  $\text{GL}(\mathbf{R})$  et un élément  $Q$  de  $\text{GL}(\mathbf{R})$  tel que :  $B = PAQ^{-1}$ .

D'après la remarque finale faite dans la preuve de la proposition 2.2.5.,  $Q^{-1}$  peut être vue comme la matrice de passage de la base  $e$  à une certaine base  $e'$ , tandis que  $P^{-1}$  peut être vue comme la matrice de passage de  $f$  à une certaine base  $f'$ . Il en résulte que  $B$  est la matrice de  $u$  avec  $e'$  comme base de départ et  $f'$  comme base d'arrivée.

D'où le résultat. □

**Proposition 2.3.8.** — *Des éléments  $A$  et  $B$  de  $\mathcal{M}_{n,p}(\mathbf{K})$  équivalents ont même rang.*

*Preuve de la proposition 2.3.8.* —

On déduit de ce résultat, qu'en particulier des matrices carrées semblables sont de même rang. La réciproque est évidemment fausse :

Nous allons voir que la réciproque de la proposition 3.2.6. est vraie. pour ce faire nous aurons besoin du résultat

technique et plus précis suivant :

**Proposition 2.3.9.** — Soient  $M$  un élément  $\mathcal{M}_{n,p}(\mathbf{K})$  et  $r$  un entier naturel. La matrice  $M$  est de rang  $r$  si et seulement si elle est équivalente à la matrice  $J_{n,p,r}$ , où :

$$J_{n,p,r} := \begin{pmatrix} I_r & 0_{r,p-r} \\ 0_{n-r,r} & 0_{n-r,p-r} \end{pmatrix},$$

en convenant bien sûr que  $J_{n,p,0} = 0_{n,p}$ .

On note souvent la matrice  $J_{n,p,r}$ , plus simplement  $J_r$ .

*Preuve de la proposition 2.3.9.* —

- Supposons  $M$  et  $J_r$  équivalents.

- Supposons  $M$  de rang  $r$ .

**Exercice**— Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$  de rang 1. Montrer qu'il existe des éléments  $A$  et  $A'$  de  $\mathcal{M}_{n,1}(\mathbf{K})$  tels que  $M = A {}^t A'$  (réciproque de 2.3.4-2.).

*Solution* —

Comme annoncé on a le résultat :

**Proposition 2.3.10.** — *Des éléments  $A$  et  $B$  de  $\mathcal{M}_{n,p}(K)$  sont équivalents si et seulement si ils ont même rang.*

*Preuve de la proposition 2.3.10.* — Nous avons déjà vu que des matrices équivalentes ont même rang (3.2.8.). Supposons réciproquement que  $A$  et  $B$  aient le même rang  $r$ . Alors d'après 3.2.9.,  $A$  et  $B$  sont toutes deux équivalente à  $J_r$ . La relation « être équivalent » étant transitive (cf; 3.2.6.) on en déduit que  $A$  et  $B$  sont équivalentes.

La proposition 2.3.9. admet une cascade de corollaires évidents.

**Corollaire 2.3.11.** — *Une matrice et sa transposée ont même rang.*

*Preuve du corollaire 2.3.11.* —

**Corollaire du corollaire 2.3.12.** — *Le rang d'une matrice est le rang de ses vecteurs lignes.*

**Corollaire du corollaire du corollaire 2.3.13.** — *Soit  $M$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ . Alors :*

$$\text{rg}(M) \leq \min(n, p).$$

**Proposition 2.3.13.** — Soit  $M$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ . Le rang de  $M$  est l'ordre maximum d'une sous-matrice carrée de  $M$  inversible.

*Preuve de la proposition 2.3.13.* —

Notons  $r$  le rang de la matrice  $M$ ,  $C_1, C_2, \dots, C_p$  ses colonnes et  $L_1, L_2, \dots, L_n$  ses lignes.

- Soit  $A$  une sous-matrice carrée d'ordre  $k$  de  $M$  inversible.  $A$  s'écrit

$$A = (a_{i,j})_{\dots\dots\dots}$$



Donc  $k \leq r$ .

•

Donc  $M$  admet une sous-matrice carrée inversible d'ordre  $r$ .

D'où le résultat. □

## 2.4 Opérations élémentaires sur les matrices

On désigne encore par  $n$  et  $p$  des entiers naturels non nuls.

Considérons le cas général d'un espace vectoriel  $\mathbf{E}$  et d'une famille  $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ , notée  $\mathcal{F}$ .

Il a été vu dans la première partie, que l'on ne modifie pas le rang de la famille en :

- Ajoutant à un des vecteur de la famille une combinaison linéaire des autres vecteurs de la famille ;
- Permutant les vecteurs de la famille ;
- Multipliant un des vecteur par un scalaire non nul.

Examinons les conséquences de cette remarque quant au rang d'une matrice. Soit  $M$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ . Nous noterons, pour  $j = 1, 2, \dots, p$ ,  $C_j$  la  $j^{\text{e}}$  colonne de  $M$  et pour  $i = 1, 2, \dots, n$ ,  $L_i$  la  $i^{\text{e}}$  ligne de  $M$ .

Le rang de la matrice  $M$  étant celui de la famille de ses colonnes, il résulte directement de la remarque préliminaire que les trois types d'opérations suivants sur ses colonnes ne modifient pas son rang.

1. Remplacer  $C_j$  par  $C_j + \lambda C_i$ , où  $j$  et  $i$  sont des éléments *distincts* de  $\{1, \dots, p\}$  et  $\lambda$  un élément de  $\mathbf{K}$  non nul.  
On notera cette opération :

$$C_j \leftarrow C_j + \lambda C_i$$

Notons qu'en itérant cette opération on peut ajouter à la  $j^{\text{e}}$  colonne une combinaison linéaire des autres colonnes.

2. Pour  $\sigma$  élément de  $S_p$ , remplacer simultanément  $C_1$  par  $C_{\sigma(1)}$ ,  $C_2$  par  $C_{\sigma(2)}$ ...  $C_p$  par  $C_{\sigma(p)}$ .  
On notera cette opération :

$$C_j \leftarrow C_{\sigma(j)}, \text{ pour } j = 1, 2, \dots, p$$

3. Remplacer la colonne  $C_j$  par  $\lambda C_j$ , où  $j$  est un élément de  $\{1, \dots, p\}$  et  $\lambda$  un scalaire *non nul*.  
On notera cette opération :

$$C_j \leftarrow \lambda C_j$$

Le rang de la matrice  $M$  étant aussi celui de la famille de ses lignes, il résulte directement de la remarque préliminaire que les trois types d'opérations suivants sur ses lignes ne modifient pas son rang.

1. Remplacer  $L_i$  par  $L_i + \lambda L_j$ , où  $i$  et  $j$  sont des éléments *distincts* de  $\{1, \dots, n\}$  et  $\lambda$  un élément de  $\mathbf{K}$  non nul. On notera cette opération :

$$L_i \leftarrow L_i + \lambda L_j$$

Notons qu'en itérant cette opération on peut ajouter à la  $i^e$  ligne une combinaison linéaire des autres lignes.

2. Pour  $\sigma$  élément de  $S_n$ , remplacer simultanément  $L_1$  par  $L_{\sigma(1)}$ ,  $L_2$  par  $L_{\sigma(2)}$ ...  $L_p$  par  $L_{\sigma(p)}$ . On notera cette opération

$$L_i \leftarrow L_{\sigma(i)}, \text{ pour } i = 1, 2, \dots, n$$

3. Remplacer la ligne  $C_i$  par  $\lambda C_i$ , où  $i$  est un élément de  $\{1, \dots, n\}$  et  $\lambda$  un scalaire *non nul*. On notera cette opération

$$L_i \leftarrow \lambda L_i$$

Nous allons maintenant, en complément, montrer que ces opérations élémentaires sur les colonnes ou les lignes correspondent à des multiplications par des matrices. Les différents types de matrices que nous allons étudier ne figurent pas explicitement au programme, néanmoins elles font l'objet de beaucoup d'exercices et problèmes de concours.

**Notations** — Par  $(E_1, E_2, \dots, E_p)$  nous désignerons la base canonique de  $\mathcal{M}_{p,1}(\mathbf{K})$ , notée  $\mathcal{B}_p$ ; la base canonique de  $\mathcal{M}_{n,1}(\mathbf{K})$  sera elle notée  $\mathcal{B}_n$ . Pour  $A, B, M, D_1$  etc. éléments de  $\mathcal{M}_{n,p}(\mathbf{K})$ , nous noterons par la minuscule correspondante :  $a, b, m, d_1$  etc., les applications linéaires de  $\mathcal{M}_{p,1}(\mathbf{K})$  dans  $\mathcal{M}_{n,1}(\mathbf{K})$ , dont ils sont les matrices,  $\mathcal{M}_{p,1}(\mathbf{K})$  étant muni de  $\mathcal{B}_p$  et  $\mathcal{M}_{n,1}(\mathbf{K})$  de  $\mathcal{B}_n$ . Enfin  $(E_{p,i,j})_{\substack{i=1\dots p \\ j=1\dots p}}$  désigne la base canonique de  $\mathcal{M}_p(\mathbf{K})$ ,  $E_{p,i,j}$ , est rappelons le, l'élément de  $\mathcal{M}_p(\mathbf{K})$  ayant un 1 sur la  $i^e$  ligne et la  $j^e$  colonne et des 0 ailleurs, (on le notera plus simplement si le contexte est clair  $E_{i,j}$ ).

MATRICE DE TRANSVECTION —

Commençons par une définition.

**Définition 2.4.1.** — On appelle *matrice de transvection d'ordre  $p$*  toute matrice de la forme :

$$I_p + \lambda E_{p,i,j},$$

où  $\lambda$  est un élément non nul de  $\mathbf{K}$ ,  $j$  et  $i$  des éléments distincts de  $\{1, \dots, p\}$ .

On notera ici une telle matrice  $T_{p,i,j}(\lambda)$ , ou plus simplement s'il n'y a pas d'ambiguïté  $T_{i,j}(\lambda)$ .

$$T_{p,i,j}(\lambda) = \left( \begin{array}{c} \phantom{\lambda} \end{array} \right)$$

Etudions l'effet de la multiplication par une matrice de transvection à droite :  $MT_{p,i,j}(\lambda)$ .

**Proposition 2.4.2.** — *Multiplier  $M$  à droite par la matrice de transvection  $T_{p,i,j}$  revient à...*

$M \leftarrow MT_{p,i,j}(\lambda) /$	$\leftarrow$
--------------------------------------	--------------

Etudions à présent l'effet de la multiplication par une matrice de transvection à gauche :  $T_{n,i,j}(\lambda)M$ .

**Proposition 2.4.3.** — *Multiplier  $M$  à gauche par la matrice de transvection  $T_{n,i,j}$  revient à...*

$M \leftarrow T_{n,i,j}(\lambda)M /$	$\leftarrow$
--------------------------------------	--------------

**Remarque** — *Le déterminant d'une matrice de transvection vaut...*

On retrouve ainsi le fait que la multiplication (à droite ou à gauche) par une matrice de transvection ne modifie pas le rang.

**Exercice 2.4.4.** — *Soient  $\lambda$  un élément non nul de  $\mathbf{K}$ ,  $j$  et  $i$  des éléments distincts de  $\{1, \dots, p\}$ . Déterminer l'inverse de la matrice  $T_{p,i,j}(\lambda)$ .*

*Solution de l'exercice 2.4.4.* —

MATRICE DE PERMUTATION —

**Définition 2.4.5.** — *Soit  $\sigma$  une permutation élément du groupe symétrique à  $p$  éléments  $S_p$ . On appelle matrice de permutation associée à  $\sigma$  la matrice  $P_\sigma = (p_{\sigma,i,j})_{\substack{i=1,\dots,p \\ j=1,\dots,p}}$  où, pour tout  $i$  et tout  $j$  éléments de  $1, \dots, p$ ,*

$$p_{\sigma,i,j} = \begin{cases} 1, & \text{si } i = \sigma(j), \\ 0, & \text{si } i \neq \sigma(j). \end{cases}$$

$$P_\sigma = \left( \begin{array}{c} \phantom{0} \end{array} \right)$$

On observe que dans chaque colonne d'une matrice de permutation il y a un et un seul 1 et que des 0. De même dans

chaque ligne, il y a un et un seul 1 et que des 0, ceci résulte . . .

Etudions l'effet de la multiplication par une matrice de permutation à droite :  $MP_\sigma$ .

**Proposition 2.4.6.** — *Multiplier  $M$  à droite par la matrice de permutation  $P_\sigma$  revient à . .*

$M \leftarrow MP_\sigma /$	$\leftarrow$
----------------------------	--------------

Etudions à présent l'effet de la multiplication par une matrice de permutation à gauche. Soit  $s$  une permutation

élément  $S_n$ , Etudions  $P_s M$ .

**Proposition 2.4.7.** — Multiplier  $M$  à gauche par la matrice de permutation  $P_\sigma$  revient à...

$$M \leftarrow P_\sigma M / \quad \leftarrow$$

**Remarque** — Le déterminant d'une matrice de permutation vaut...

On retrouve ainsi le fait que la multiplication (à droite ou à gauche) par matrice de permutation ne modifie pas le rang.

Concluons notre étude des matrices de permutation par quelques propriétés données au grés d'un exercice

**Exercice 2.4.8.** —

1. Soit  $\sigma_1$  et  $\sigma_2$  des éléments de  $S_p$ . Montrer que  $P_{\sigma_1} P_{\sigma_2}$  est une matrice de permutation à déterminer.
2. En déduire l'inverse d'une matrice de permutation  $P_\sigma$ , pour  $\sigma \in S_p$ . Montrer que l'ensemble des matrices de permutations d'ordre  $p$  est un sous-groupe de  $\text{GL}_p(\mathbf{K})$ . On le note dans la suite  $\mathcal{P}_p(\mathbf{K})$ .
3. Montrer que l'application  $\varphi; S_p \rightarrow \mathcal{P}_p(\mathbf{K}); \sigma \mapsto P_\sigma$  est un morphisme de groupes. En déduire un système de générateurs de  $\mathcal{P}_p(\mathbf{K})$ .
4. Soit  $\sigma$  un élément de  $S_p$ . Déterminer  ${}^t P_\sigma$ .

*Solution de l'exercice 2.4.8. —*

**Remarque** — De la question 3. de l'exercice, il ressort que pour permuter les colonnes (resp. les lignes) d'une matrice, on peut échanger un certain nombre de fois deux colonnes (resp. deux lignes) ce qui revient à multiplier à droite (resp. à gauche) un certain nombre de fois par des matrices de transposition.

MATRICE DE DILATATION —

On se propose de réaliser l'opération sur les colonnes de  $M$ ,  $C_j \leftarrow \lambda C_j$ . Brutalement, le résultat évident :

**Proposition, définition 2.4.9.** — Soient  $j$  un élément de  $\{1, \dots, p\}$  et  $\lambda$  un élément non nul de  $\mathbf{K}$ . Remplacer la  $j^e$  colonne de  $M$ ,  $C_j$  par  $\lambda C_j$  revient à multiplier  $M$  à droite par la matrice  $D_{p,j}(\lambda)$ , où

$$D_{p,j}(\lambda) = \begin{pmatrix} & \\ & \end{pmatrix}.$$

La matrice  $D_{p,j}(\lambda)$  est appelée matrice de dilatation.

$M \leftarrow MD_{p,j}(\lambda) / \quad \leftarrow$
---

**Proposition 2.4.10.** — Soient  $\lambda$  un élément non nul de  $\mathbf{K}$  et  $i$  de  $\{1, \dots, n\}$ . La multiplication de  $M$  à gauche par la matrice  $D_{n,i}(\lambda)$  revient à...

$M \leftarrow D_{n,i}(\lambda)M / \quad \leftarrow$
---

*Preuve de la proposition 2.4.10. —*

**Remarque** — *Le déterminant de la matrice de dilatation  $D_{p,j}(\lambda)$  vaut...*

On retrouve ainsi le fait que la multiplication (à droite ou à gauche) par une matrice de dilatation ne modifie pas le rang.

**Remarque** — Il ressort de ce qui précède que toute transformation élémentaire sur les colonnes (resp. les lignes) d'une matrice revient à multiplier à droite (resp. à gauche) par une matrice inversible, (produit de matrices élémentaires).

Remarquons que la matrice qui par multiplication à gauche (resp. droite) effectue une transformation élémentaire sur les lignes (resp. colonnes) ne dépend pas du nombre de colonnes (resp. lignes) de  $M$

Donc on peut retrouver la matrice qui par multiplication effectue cette transformation : soit  $P$  la matrice sensée effectuer par multiplication à droite (resp. gauche) une transformation  $T$  sur les colonnes (resp. lignes) alors on obtient  $P$  en effectuant sur la matrice  $I_p$  (resp.  $I_n$ ) la transformation  $T$ , en effet  $I_p P = P$ ; ( resp.  $P I_n = P$ )

**Exemple** —

La matrice élémentaire qui par multiplication à... ajoute 2 fois la 3<sup>e</sup> ligne à la 5<sup>e</sup> pour un élément de  $\mathcal{M}_{n,p}(\mathbf{R})$  est

Celle qui par multiplication à... ajoute 5 fois la 2<sup>e</sup> colonne à la 6<sup>e</sup> pour un élément de  $\mathcal{M}_{n,p}(\mathbf{R})$  est

Il découle aussi de la remarque précédente le résultat suivant :

**Proposition 2.4.11.** — *Soit  $M$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ .*

1. *Soit  $M'$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$  déduit de  $M$  par opérations élémentaires sur les colonnes, alors l'image de  $M'$  est égale à l'image de  $M$ .*
2. *Soit  $M''$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$  déduit de  $M$  par opérations élémentaires sur les lignes, alors le noyau de  $M''$  est égale au noyau de  $M$ .*



*Preuve de la proposition 2.4.11. —*

1. D'après la remarque on peut trouver  $P \in \text{GL}_p(\mathbf{K})$  telle que  $M' = MP...$

2. De même peut-on trouver  $P \in \text{GL}_n(\mathbf{K})$  telle que  $M' = QM$ .



## PIVOTS DE GAUSS

Nous allons rappeler les méthodes d'échelonnement de Gauss et leurs applications.

Définissons pour commencer les matrices échelonnées en lignes ou en colonnes.

**Définition 2.4.12.** — On dit qu'un élément  $B$  de  $\mathcal{M}_{n,p}(\mathbf{K})$  est échelonné en ligne si elle vérifie la condition suivante : Pour  $i = 1, \dots, n-1$ ,

- Si la  $i^{\text{e}}$  ligne est nulle alors la  $i+1^{\text{e}}$  est nulle également ;
- Sinon, si la  $i+1^{\text{e}}$  ligne est non nulle l'indice  $j$  du premier coefficient  $b_{i+1,j}$  non nul dans la ligne  $i+1$  est strictement supérieur à celui du premier coefficient non nul dans la ligne  $i$ .

**Définition 2.4.13.** — On dit qu'un élément  $B$  de  $\mathcal{M}_{n,p}(\mathbf{K})$  est échelonné en colonnes si elle vérifie la condition suivante : Pour  $j = 1, \dots, p-1$ ,

- Si la  $j^{\text{e}}$  colonne est nulle alors la  $j+1^{\text{e}}$  est nulle également ;
- Sinon, si la  $j+1^{\text{e}}$  colonne est non nulle, alors l'indice  $i$  du premier coefficient  $b_{i,j+1}$  non nul dans la colonne  $j+1$  est strictement supérieur à celui du premier coefficient non nul dans la colonne  $j$ .

**Remarques** —

- Une matrice est échelonnée en colonnes si et seulement si sa transposée est échelonnée en lignes.
- Une matrice échelonnée en lignes est triangulaire supérieure.
- Une matrice triangulaire supérieure, ou même diagonale peut ne pas être échelonnée en lignes.
- Une matrice échelonnée en colonnes est triangulaire inférieure.
- Une matrice triangulaire inférieure, ou même diagonale peut ne pas être échelonnée en lignes.
- La matrice  $J_{n,p,r}$  est échelonnée en lignes et en colonnes.

Le pivot de Gauss fournit une méthode algorithmique pour transformer, par des transformations élémentaires sur les lignes (resp. colonnes) une matrice en une matrice échelonnée en lignes, (resp. colonnes). Précisément

**Proposition 2.4.14.** — PIVOT DE GAUSS —

Soit  $M$  un élément de  $\mathcal{M}_{n,p}(\mathbf{K})$ .

1. Il existe un élément  $U$  de  $\mathrm{GL}_n(\mathbf{K})$ , produit de matrices de transvections et de permutations (transpositions) telle que  $UM$  soit échelonnée en lignes.

*Autrement dit on peut en ajoutant à une ligne un multiple d'une autre et en permutant des lignes, un certain nombre de fois transformer  $M$  en une matrice échelonnée en ligne.*

2. Il existe un élément  $V$  de  $\mathrm{GL}_p(\mathbf{K})$ , produit de matrices de transvections et de permutations (transpositions) telle que  $MV$  soit échelonnée en colonnes.

*Autrement dit on peut en ajoutant à une colonne un multiple d'une autre et en permutant des colonnes, un certain nombre de fois transformer  $M$  en une matrice échelonnée en colonne.*

*Preuve de la proposition 2.4.14.* — Nous allons montrer le premier point. Le second s'en déduit par transposition. Nous allons donner un raisonnement par récurrence sur  $n$  qui fournit du reste, presque instantanément, un algorithme récursif.

Notons  $(P_n)$  la propriété : « Pour tout entier  $p \geq 1$  et tout élément  $A$  de  $\mathcal{M}_{n,p}(\mathbf{K})$  il existe un élément  $U$  de  $\mathrm{GL}_n(\mathbf{K})$ , produit de matrices de transvections et de transpositions telle que  $UA$  soit échelonnée en lignes. » .

- Clairement la propriété  $(P_1)$  est vraie, puisque toute matrice ligne est (déjà) échelonnée.
- Soit  $m \in \mathbf{N}^*$ . Supposons  $(P_m)$ . Soit  $B \in \mathcal{M}_{m+1,p}(\mathbf{K})$ ... Afin d'alléger la rédaction nous noterons  $b_{i,j}$  le terme d'indice

$(i, j)$  des différentes matrices obtenues par transformation à partir de  $B$ .

Il est possible de donner l'algorithme sous forme non récursive. Dans le cas d'une matrice carré d'ordre  $n$  le nombre d'opérations est dominé par  $n^3$ . Nous renvoyons au cours d'informatique pour plus de détails, pour notre part nous allons donner des exemples simples traités à la main.

Il est pratique pour un pivot sur les lignes d'écrire la matrice  $I_n$  à droite de  $M$  et de lui faire subir les mêmes transformations sur les lignes qu'à la matrice  $M$ , on obtient ainsi en fin de pivot la matrice  $U$  à la place de  $I_n$

Il est pratique pour un pivot sur les colonnes d'écrire la matrice  $I_p$  sous la matrice  $M$  et de lui faire subir les mêmes transformations sur les colonnes qu'à la matrice  $M$ , on obtient ainsi en fin de pivot la matrice  $V$  à la place de  $I_p$

**Exemple** — Echelonnons la matrice suivante en lignes, puis en colonnes.

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 4 & 8 & 6 \\ 0 & 1 & -2 & 3 \\ 0 & 4 & 2 & 6 \end{pmatrix}$$

**Applications —**

L'échelonnement d'une matrice peut être utilisé pour calculer le **rang** d'une matrice. En effet, avec les notations de la proposition 2.4.14.,

$$\mathbf{rg}(M) = \mathbf{rg}(UM) = \mathbf{rg}(MV)$$

Or le rang d'une matrice échelonnée en lignes (resp. colonnes) est le nombre de lignes (resp. colonnes) non nulles.

On peut aussi en déduire dans le cas où  $n = p$  le **déterminant**, en effet

Enfin l'échelonnement en colonnes fournit une base de l'image et du noyau de  $M$ .

**Proposition 2.4.15.** — *On garde les notations de 2.2.14. et l'on note  $r$  le rang de  $M$ . Alors :*

1. ( ) est une base de l'image de  $M$ .

2. ( ) est une base du noyau de  $M$ .

Reprenons l'exemple précédent.

### Calcul de l'inverse

Commençons par observer que si dans la proposition 2.3.15. on autorise  $U$  à être non seulement produit de matrices de transvections et de transpositions mais aussi de dilatations, alors on peut imposer à la matrice  $UM$  d'être échelonnée en ligne avec pour chaque ligne non nulle le premier terme non nul égal à 1.

Dans le cas où  $p = n$  et où  $M$  est inversible, on montre sans mal que l'on peut construire de façon algorithmique une matrice  $U'$ , produit de matrices de transvections telle que :

$$U'UM = I_n.$$

On a alors  $U'U = M^{-1}$ .

Nous ne détaillerons pas d'avantage et passons directement à un exemple :

$$\begin{pmatrix} 2 & 2 & 2 \\ 4 & 0 & 2 \\ 8 & 8 & 0 \end{pmatrix}$$

Dans un registre très voisin, on peut construire deux matrices  $P \in \mathrm{GL}_n(\mathbf{K})$ ,  $Q \in \mathrm{GL}_p(\mathbf{K})$ , produits de matrices élémentaires telles que

$$PMQ = J_{n,p,r},$$

où  $r$  désigne le rang de  $M$  et retrouver ainsi une autre preuve de 2.3.9.

Pour finir un exercice.

**Exercice 2.4.17.** —

1. On suppose dans cette question que  $M$  est élément de  $\mathrm{GL}_n(\mathbf{K})$ , ( $n = p$ ). Montrer qu'il existe  $P$  et  $R$  éléments de  $\mathrm{GL}_n(\mathbf{K})$ , produits de matrices de transvections d'ordre  $n$  tels que :

$$PMR = \mathrm{diag}(1, 1, \dots, 1, d),$$

où  $d$  est le déterminant de  $M$ .

2. En déduire que l'ensemble des transvections engendre  $SL_n(\mathbf{K})$ , ensemble des éléments de  $\mathcal{M}_n(\mathbf{K})$  de déterminant 1.
3. **Réservé  $\frac{5}{2}$ .** Déduire de 2., que  $\mathrm{GL}_n^+(\mathbf{R})$ , ensemble des éléments de  $\mathcal{M}_n(\mathbf{R})$  de déterminant strictement positif, est connexe par arcs et que  $\mathrm{GL}_n^-(\mathbf{R})$ , ensemble des éléments de  $\mathcal{M}_n(\mathbf{R})$  de déterminant strictement négatif, l'est également.

*Solution de l'exercice 2.4.16.* —

## DIAGONALISATION

### 3.1 Introduction

L'objet de cette partie est de trouver, pour une matrice  $M$ , lorsque la chose est possible une matrice semblable diagonale. Outre les applications théoriques que nous découvrirons, l'intérêt pour mener des calculs se comprend très vite. Donnons un exemple

**Exemple 3.1.1.** — Soit  $M$  la matrice, élément de  $\mathcal{M}_2(\mathbf{R})$ , suivante :

$$M = \begin{pmatrix} 3 & 7 \\ 5 & 5 \end{pmatrix}.$$

On se propose de calculer pour tout entier  $n \geq 0$ ,  $M^n$ .

On sait calculer sans mal la puissance  $n^e$  d'une matrice diagonale, qui est la matrice diagonale dont les termes diagonaux sont les puissances  $n^e$  de ceux de la matrice initiale. Cherchons à nous ramener à cette situation. Soit  $m$  l'endomorphisme de  $\mathcal{M}_{2,1}(\mathbf{R})$  canoniquement associé à  $M$ . Pour que  $M$  soit semblable à une matrice diagonale, il faut trouver une base  $(U_1, U_2)$  de  $\mathcal{M}_{2,1}(\mathbf{R})$  telle que.....

Donc en posant  $P = \begin{pmatrix} & \\ & \end{pmatrix}$  et  $\Delta = \text{diag}(\dots\dots\dots)$ , on a :

$$M = P \cdots \Delta P \cdots.$$

Ainsi, par récurrence a-t-on :

$$M^n =$$

pour tout entier  $n \geq 0$ .

Pour mesure d'avantage l'intérêt de déterminer une matrice diagonale semblable à une matrice donnée traitons un exercice :

**Exercice 3.1.2.** —

1. Déterminer les suite réelles  $(u_n)_{n \in \mathbf{N}}$  et  $(v_n)_{n \in \mathbf{N}}$  telles que

$$\begin{cases} u_{n+1} = 6u_n + 4v_n, \\ v_{n+1} = 11u_n - v_n, \end{cases}, \quad (\text{II.3})$$

pour tout entier  $n \geq 0$ .

2. Déterminer les couples d'applications de  $\mathbf{R}$  dans  $\mathbf{R}$  de classe  $\mathcal{C}^1$ ,  $(\varphi, \psi)$  tels que :

$$\begin{cases} \varphi' = 6\varphi + 4\psi, \\ \psi' = 11\varphi - \psi, \end{cases}, \quad (\text{II.4})$$

*Solution de l'exercice 3.1.2. —*

## 3.2 Préliminaires techniques

Nous désignons par  $\mathbf{K}$  indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$  et par  $\mathbf{E}$  un espace vectoriel sur le corps  $\mathbf{K}$ , de dimension quelconque. Là encore ce qui suit est vrai pour  $\mathbf{K}$  corps quelconque.

**Définition 3.2.1.** — Soit  $u$  un endomorphisme de  $\mathbf{E}$  et soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$  stable par  $u$  ( $u(\mathbf{F}) \subset \mathbf{F}$ ). On dispose alors de l'application

$$\mathbf{F} \rightarrow \mathbf{F}; \quad \vec{x} \mapsto u(\vec{x}).$$

Cette application qui est évidemment un endomorphisme de  $\mathbf{F}$ , s'appelle endomorphisme induit par  $u$  sur  $\mathbf{F}$ .

**Proposition 3.2.2.** — Soient  $u$  et  $v$  des éléments de  $\mathcal{L}(\mathbf{E})$ . Si  $u$  et  $v$  commutent ( $u \circ v = v \circ u$ ), alors  $\text{Im}(v)$  et  $\text{Ker}(v)$  sont stables par  $u$ .

*Preuve de la proposition 3.2.2. —*



- Soit  $\vec{y}$  un élément de  $\text{Im}(v)$ . Il existe  $\vec{x}$  élément de  $\mathbf{E}$  tel que  $v(\vec{x}) = \vec{y}$ .

$$u(\vec{y}) = u(v(\vec{x})) = u \circ v(\vec{x}) = v \circ u(\vec{x}) = v(u(\vec{x})).$$

Donc  $u(\vec{y}) \in \text{Im}(v)$ . Donc,  $\vec{y}$  étant quelconque, on a prouvé la stabilité de  $\text{Im}(v)$  par  $u$ .

- Soit  $\vec{x}$  un élément de  $\text{Ker}(v)$

$$v(u(\vec{x})) = v \circ u(\vec{x}) = u \circ v(\vec{x}) = u(v(\vec{x})) = u(\vec{0}_E) = \vec{0}_E.$$

Donc  $u(\vec{x}) \in \text{Ker}(v)$ . Donc  $\vec{x}$  étant constant,  $\text{Ker}(v)$  est stable par  $u$ . □

**Proposition 3.2.3.** — On suppose ici que  $\mathbf{E}$  est de dimension finie non nulle  $n$ . Soient  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$  de dimension non nulle  $p$  et  $u$  un élément de  $\mathcal{L}(\mathbf{E})$ . Soit enfin  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base de  $\mathbf{E}$  notée  $\mathcal{B}$ , adaptée à  $\mathbf{F}$ , c'est-à-dire que  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  est une base de  $\mathbf{F}$ .

—  $\mathbf{F}$  est stable par  $u$  si et seulement si la matrice de  $u$  dans  $\mathcal{B}$  est de la forme

$$\begin{pmatrix} A & B \\ 0_{n-p,p} & C \end{pmatrix},$$

avec  $A \in \mathcal{M}_p(\mathbf{K})$ ,  $B \in \mathcal{M}_{p,n-p}(\mathbf{K})$ ,  $C \in \mathcal{M}_{n-p}(\mathbf{K})$ .

— Si c'est le cas,  $A = \text{mat}_{(\vec{e}_1, \dots, \vec{e}_p)}(v)$ , où  $v$  est l'endomorphisme induit par  $u$  sur  $\mathbf{F}$ .

*Preuve de la proposition 3.2.3.* —

**Proposition 3.2.4.** — GÉNÉRALISATION — On suppose ici que  $\mathbf{E}$  est de dimension finie non nulle  $n$ . Soit  $(\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_q)$  une famille de  $q$  sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires. Soit  $\mathcal{B}$  une base de  $\mathbf{E}$  adaptée à la décomposition de  $\mathbf{E} = \bigoplus_{i=1}^q \mathbf{F}_i$ , c'est-à-dire que  $\mathcal{B}$  est la concaténée de  $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$ , où pour  $i = 1, 2, \dots, q$ ,  $\mathcal{B}_i$  est une base de  $\mathbf{F}_i$ . Soit enfin  $u$  un endomorphisme de  $\mathbf{E}$ .

— Les  $q$  sous-espaces  $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_q$  sont stables par  $u$  si et seulement si, la matrice de  $u$  dans  $\mathcal{B}$  est diagonale par blocs de la forme :

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & A_3 & \\ & & & 0 \\ & 0 & & \ddots & \\ & & & & A_q \end{pmatrix}, \quad (\text{II.5})$$

avec, pour  $i = 1, 2, \dots, q$ ,  $A_i \in \mathcal{M}_{n_i}(\mathbf{K})$ , où  $n_i$  désigne la dimension de  $\mathbf{F}_i$ .

— Si c'est le cas alors, pour  $i = 1, 2, \dots, q$ ,  $A_i = \text{mat}_{\mathcal{B}_i}(u_i)$ , où  $u_i$  est l'endomorphisme induit par  $u$  sur  $\mathbf{F}_i$ .

*Preuve de la proposition 3.2.4.* — (exercice)

Terminons par le rappel d'un résultat au programme de MPSI sur les déterminants de matrices de la forme (II.5) qui sera utile dans la suite.

**Proposition 3.2.5.** — DÉTERMINANTS PAR BLOCS — Soient  $n$  un entier naturel supérieur ou égal à 2,  $p$  un élément de  $\{1, \dots, n-1\}$  et  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$  de la forme

$$\begin{pmatrix} A & B \\ 0_{n-p,p} & C \end{pmatrix},$$

avec  $A \in \mathcal{M}_p(\mathbf{K})$ ,  $B \in \mathcal{M}_{p,n-p}(\mathbf{K})$ ,  $C \in \mathcal{M}_{n-p}(\mathbf{K})$ . Alors

$$\text{Det}(M) = \text{Det}(A)\text{Det}(C).$$

*Preuve de la proposition 3.2.5.* — Commençons par rappeler des résultats fondamentaux de la théorie des déterminants.

Soit  $\mathbf{F}$  un espace vectoriel sur  $\mathbf{K}$  de dimension non nulle  $p$  ;

- l'ensemble des formes  $p$ -linéaires alternées sur  $\mathbf{F}$  est un espace vectoriel de dimension 1 ;
- soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  une base de  $\mathbf{F}$ , notée  $\mathcal{B}$ . Il existe donc une et une seule forme  $p$ -linéaire alternée sur  $\mathbf{F}$  qui prenne la valeur 1 en  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ . On l'appelle déterminant dans la base  $\mathcal{B}$ , on la note  $\det_{\mathcal{B}}$  ;
- donc pour toute forme  $p$ -linéaire alternée sur  $\mathbf{F}$ ,  $\varphi$ , il existe  $k$  élément de  $\mathbf{K}$  tel que :  $\varphi = k \det_{\mathcal{B}}$  ;
- le déterminant d'une matrice  $P$  d'ordre  $p$  est par définition le déterminant dans la base canonique de  $\mathcal{M}_p(\mathbf{K})$  de ses vecteurs colonnes, on le note  $\det P$ .

Passons à la preuve proprement dite de la proposition.

Soit l'application :

$$\varphi : (\mathcal{M}_{p,1}(\mathbf{K}))^p \rightarrow \mathbf{K}; (X_1, X_2, \dots, X_p) \mapsto \text{Det} \begin{pmatrix} X_1 X_2 \dots X_p & B \\ 0_{n-p,p} & C \end{pmatrix}.$$

$\varphi$  est une forme  $p$ -linéaire alternée sur  $\mathcal{M}_{p,1}(\mathbf{K})$ , espace de dimension  $p$ . Il existe donc un élément  $k$  de  $\mathbf{K}$  tel que :

$$\varphi = k \det_{\mathcal{B}_c},$$

où  $\mathcal{B}_c$  est la base canonique de  $\mathcal{M}_{p,1}(\mathbf{K})$ .

$$\varphi \left( \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right) = k \times 1.$$

Or, par définition

$$\varphi \left( \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right) = \text{Det} \begin{pmatrix} 1 & & & B \\ & \ddots & & \\ & & 1 & \\ 0_{n-p,p} & & & C \end{pmatrix}$$

En développant ce dernier déterminant  $p$  fois par rapport à la première colonne il vient :

$$\varphi \left( \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right) = \text{Det} C.$$

Donc  $k = \text{Det} C$ . Mais  $\text{Det} M = \varphi(A_1, A_2, \dots, A_p)$ , où, pour  $i = 1, 2, \dots, p$ ,  $A_i$  désigne la  $i^{\text{e}}$  colonne de  $A$ . Donc  $\text{Det} M = k \det_{\mathcal{B}_c}(A_1, A_2, \dots, A_p) = \text{Det} C \text{Det} A$ . D'où le résultat.  $\square$

### 3.3 Valeurs propres, vecteurs propres

Nous désignons encore par  $\mathbf{K}$  indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$ , par  $\mathbf{E}$  un espace vectoriel sur le corps  $\mathbf{K}$ , de dimension quelconque,  $u$  désignera lui un endomorphisme de  $\mathbf{E}$ .

Nous allons étudier les vecteurs de  $\mathbf{E}$  qui sont transformés par  $u$  en un vecteur qui leur est colinéaire.

**Définition 3.3.1.** — VALEURS PROPRES — On appelle valeur propre de  $u$ , tout élément  $\lambda$  de  $\mathbf{K}$  tel qu'il existe un élément  $\vec{x}$  de  $\mathbf{E}$ , **non nul**, tel que :

$$u(\vec{x}) = \lambda \vec{x}.$$

Si c'est le cas, le vecteur  $\vec{x}$  est dit vecteur propre de  $u$  associé à la valeur propre  $\lambda$ .

! Par définition un vecteur propre est non nul !

**Remarque** — Un vecteur propre est associé à une et une seule valeur propre.

On a la proposition évidente, mais importante suivante :

**Proposition 3.3.2.** — Soit  $\lambda$  élément de  $\mathbf{K}$ . Les trois propositions suivantes sont équivalentes :

- i)  $\lambda$  est valeur propre de  $u$  ;
- ii)  $\text{Ker}(\text{Id}_{\mathbf{E}} - u)$  est non réduit à  $\{\vec{0}_{\mathbf{E}}\}$  ;
- iii) L'endomorphisme  $\text{Id}_{\mathbf{E}} - u$  n'est pas injectif.

**Remarques** —

- 0 est valeur propre de  $u$  si et seulement si  $u$  est non injectif.
- Si  $\mathbf{E}$  est de dimension finie, alors  $\lambda$  est valeur propre si et seulement si  $u - \text{Id}_{\mathbf{E}}$  est non surjectif, ou encore si et seulement si il est non bijectif.

**Définition 3.3.3.** — SPECTRE — Dans le cas où  $\mathbf{E}$  est de dimension finie, on appelle spectre de  $u$ , l'ensemble de ses valeurs propres, on le note  $\text{sp}(u)$ .

**Définition 3.3.4.** — ESPACES PROPRES — Soit  $\lambda$  une valeur propre de  $u$ . On appelle espace propre de  $u$  associé à  $\lambda$ , l'ensemble  $\text{Ker}(\text{Id}_{\mathbf{E}} - u)$ . On note cet ensemble  $\mathbf{E}_{\lambda}(u)$  ou, lorsqu'il n'y a pas de confusion possible sur l'endomorphisme,  $\mathbf{E}_{\lambda}$ .

**Remarques** —

- Un espace propre est donc un sous-espace vectoriel de  $\mathbf{E}$  non réduit à  $\{\vec{0}_{\mathbf{E}}\}$ .
- $\mathbf{E}_{\lambda}$  est la réunion de l'ensemble de tous les vecteurs propres associés à  $\lambda$  et de  $\{\vec{0}_{\mathbf{E}}\}$ .

**Interprétation « géométrique »** — Soit  $\lambda$  une valeur propre de  $u$ .  $\mathbf{E}_{\lambda}$  est stable par  $u$ , comme le montre un calcul élémentaire. De plus l'endomorphisme  $u$  induit sur  $\mathbf{E}_{\lambda}$  une homothétie, celle de rapport  $\lambda$ .

La stabilité des espaces propres se généralise :

**Proposition 3.3.5.** — Soient  $u$  et  $v$  des éléments, de  $\mathcal{L}(\mathbf{E})$  qui commutent ( $u \circ v = v \circ u$ ). Alors, tout espace propre de l'un est stable par l'autre.

*Preuve de la proposition 3.3.5.* — prenons par exemple  $\lambda$  une valeur propre de  $u$ . Les endomorphismes  $v$  et  $u - \text{Id}_{\mathbf{E}}$  commutent, puisque  $v$  commute avec  $u$  et  $\text{Id}_{\mathbf{E}}$ . Donc d'après la proposition 3.2.2.,  $\text{Ker}(\text{Id}_{\mathbf{E}} - u) = \mathbf{E}_{\lambda}(u)$  est stable par  $v$ . □

Voilà maintenant un des résultats essentiels de ce chapitre.

**Théorème 3.3.6.** — Soit  $u$  élément de  $\mathcal{L}(u)$ .

- i) Soient  $p$  un élément de  $\mathbf{N}^*$  et  $(\vec{v}_i)_{i \in \{1, \dots, p\}}$  une famille de vecteurs propres de  $u$ , associés à des valeurs propres deux à deux distinctes. Alors  $(\vec{v}_i)_{i \in \{1, \dots, p\}}$  est libre.
- ii) Soient  $q$  un entier supérieur ou égal à 2 et  $(\mathbf{E}_i)_{i \in \{1, \dots, q\}}$  une famille de sous-espaces propres de  $u$  associés à des valeurs propres deux à deux distinctes. Alors la somme  $\sum_{i=1}^q \mathbf{E}_i$  est directe :

$$\sum_{i=1}^q \mathbf{E}_i = \bigoplus_{i=1}^q \mathbf{E}_i.$$

*Preuve du théorème 3.3.6.* —

### Applications

1. Pour tout réel  $\alpha$ , on pose

$$e_\alpha : \mathbf{R} \rightarrow \mathbf{R}; x \mapsto e^{\alpha x}.$$

Montrons que  $(e_\alpha)_{\alpha \in \mathbf{R}}$  est une famille libre de  $\mathcal{F}(\mathbf{R}, \mathbf{R})$ . Soit  $D$  l'opérateur linéaire de « dérivation » :

$$D : \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}) \rightarrow \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}); f \mapsto f'$$

Pour tout réel  $\alpha$ ,  $e_\alpha$  est un vecteur propre de  $D$ , associé à la valeur propre  $\alpha$ . Soit  $(\mu_\alpha)_{\alpha \in \mathbf{R}}$  une famille presque nulle de réels telle que

$$\sum_{\alpha \in \mathbf{R}} \mu_\alpha e_\alpha = 0.$$

Supposons  $(\mu_\alpha)_{\alpha \in \mathbf{R}}$  non nulle. Notons  $\mu_{\alpha_1}, \mu_{\alpha_2}, \dots, \mu_{\alpha_n}$ , les éléments non nuls distincts de  $\{\mu_\alpha, \alpha \in \mathbf{R}\}$ .

$$\sum_{i=1}^n \mu_{\alpha_i} e_{\alpha_i} = 0.$$

Or d'après le théorème 1.2.6.,  $(e_{\alpha_1}, e_{\alpha_2}, \dots, e_{\alpha_n})$  est libre, d'où la contradiction. Donc la famille  $(\mu_\alpha)_{\alpha \in \mathbf{R}}$  est nulle. d'où la liberté de la famille  $(e_\alpha)_{\alpha \in \mathbf{R}}$ .

2. Pour tout  $n$ , entier naturel non nul on pose :

$$c_n : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \cos(nt), \quad s_n : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \sin(nt)$$

Montrons que la famille  $(c_n, s_n)_{n \in \mathbf{N}} = (c_1, s_1, c_2, s_2, \dots, c_n, s_n, \dots)$  est libre.

### Cas matriciel —

Soient  $n$  un entier naturel non nul et  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . On introduit l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$

$$m : \mathcal{M}_{n,1}(\mathbf{K}) \rightarrow \mathcal{M}_{n,1}(\mathbf{K}); X \mapsto MX,$$

appelé endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  *canoniquement associé* à  $M$ . On identifiera  $M$  à  $m$  et l'on pourra donc parler du *noyau* de  $M$ , de *image* de  $M$ , de *valeurs propres* de  $M$  et d'espaces propres de  $M$ , qui sont donc respectivement le noyau de  $m$ , l'image de  $m$ , les valeurs propres de  $m$  et les vecteurs propres de  $m$ .

Ainsi un élément  $\lambda$  de  $\mathbf{K}$  est-il valeur propre de  $M$  s'il existe  $X$  élément de  $\mathcal{M}_{n,1}(\mathbf{K})$  non nul, tel que  $MX = \lambda X$ . Si tel est le cas, l'espace propre de  $M$  associé à  $\lambda$ ,  $\mathbf{E}_\lambda(M)$ , est le sous-espace vectoriel  $\text{Ker}(M - I_n)$ .

**Proposition 3.3.7.** — Soient  $u$  un élément de  $\mathcal{L}(\mathbf{E})$ ,  $\mathcal{B}$  une base de  $\mathbf{E}$  et  $M$  la matrice de  $u$  dans  $\mathcal{B}$ . Soit  $\vec{x}$  un vecteur de  $\mathbf{E}$  et  $X$  le vecteur colonne des coordonnées de  $\vec{x}$  dans  $\mathcal{B}$ . Le vecteur  $\vec{x}$  est vecteur propre de  $u$  associée à une valeur propre  $\lambda$  si et seulement si  $X$  est vecteur propre de  $M$  associé à  $\lambda$ .

*Preuve de la proposition 3.3.7.* — On a que  $\vec{x}$  est vecteur propre de  $u$  associée à la valeur propre  $\lambda$  si et seulement si

$$\vec{x} \neq \vec{0}_{\mathbf{E}} \text{ et } u(\vec{x}) = \lambda \vec{x}, \text{ ce qui équivaut à } X \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \text{ et } MX = \lambda X. \text{ Donc } \vec{x} \text{ est vecteur propre de } u \text{ associée à la valeur}$$

propre  $\lambda$  si et seulement si  $X$  est vecteur propre de  $M$  associé à  $\lambda$ . □

Une conséquence de cette proposition est que, dans l'étude pratique des espaces propres d'un endomorphisme, on se ramène souvent au cas matriciel.

**Mise en garde** — Dans la pratique une matrice est souvent donnée par un tableau de réels voire d'entiers. Il convient donc de préciser pour l'étude des valeurs propres dans lequel des espaces  $\mathcal{M}_n(\mathbf{R})$  ou  $\mathcal{M}_n(\mathbf{C})$  on travaille, ce choix est loin d'être indifférent. Prenons l'exemple de la matrice  $M$  suivante :

$$M = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

où  $\theta$  est un réel non congru à 0 modulo  $\pi$ . Soit  $\lambda$  un élément de  $\mathbf{K}$ .  $\lambda$  est valeur propre de  $M$  si et seulement si  $\text{Ker}(M - \lambda \text{id}_{\mathbf{E}})$  est non nul, soit si et seulement si

$$\det(\lambda I_2 - M) = \lambda^2 - 2 \cos \theta \lambda + 1 = 0.$$

- Si  $M$  est considérée comme un élément de  $\mathcal{M}_2(\mathbf{C})$ , alors elle admet deux valeurs propres  $e^{i\theta}$  et  $e^{-i\theta}$
- Si  $M$  est considérée comme un élément de  $\mathcal{M}_2(\mathbf{R})$ , alors elle n'a pas de valeur propre<sup>1</sup>.

### Exemples —

#### 1. HOMOTHÉTIES —

Soient  $k$  un réel et  $h$  l'homothétie de rapport  $k$  (c'est-à-dire que  $h = k \text{id}_{\mathbf{E}}$ ). On montre sans mal que  $h$  admet une et une seule valeur propre,  $\cdot$ , l'espace propre associé est  $\cdot$ .

#### 2. PROJECTIONS ET SYMÉTRIES—

Soient  $\mathbf{F}$  et  $\mathbf{G}$  deux sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires. On désigne par  $p$  la projection de  $\mathbf{E}$  sur  $\mathbf{F}$  suivant  $\mathbf{G}$  et par  $s$  la symétrie de  $\mathbf{E}$  par rapport à  $\mathbf{F}$  suivant  $\mathbf{G}$ . On exclut les cas triviaux  $\mathbf{F} = \mathbf{E}$  et  $\mathbf{F} = \{\vec{0}_{\mathbf{E}}\}$ , qui nous ramènent au premier exemple.

- *Cas de la projection  $p$*

- *Cas de la symétrie  $s$*

#### 3. AFFINITÉS —

Soit  $k$  un réel distinct de 1 et de 0. On garde les notations de l'exemple précédent et l'on note  $a$  l'affinité de  $\mathbf{E}$  par rapport à  $\mathbf{F}$  suivant  $\mathbf{G}$  de rapport  $k$ . Rappelons que  $a$  est définie par, pour tout élément  $x$  de  $\mathbf{E}$ ,

$$a(x) = p(x) + k(x - p(x)).$$

---

1. Qu'une matrice de rotation ne transforme aucun vecteur en un vecteur colinéaire n'est guère fait pour nous surprendre !

### 3.4 Polynôme caractéristique

Dans ce paragraphe  $\mathbf{K}$  désigne indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$  et  $\mathbf{E}$  un  $\mathbf{K}$ -espace vectoriel, supposé ici *de dimension finie, non nulle*  $n$ . On considère encore, dans tout le 3.4., un endomorphisme  $u$  de  $\mathbf{E}$ .

Soit  $\lambda$  un élément de  $\mathbf{K}$ . On a signalé que  $\lambda$  est valeur propre de  $u$  si et seulement si  $\lambda \text{id}_{\mathbf{E}} - u$  est non injectif, donc  $\lambda$  est valeur propre de  $u$  si et seulement si  $\det(\lambda \text{id}_{\mathbf{E}} - u) = 0$ . D'où l'intérêt d'étudier l'application :

$$\chi_u : \mathbf{K} \rightarrow \mathbf{K}; \lambda \mapsto \det(\lambda \text{id}_{\mathbf{E}} - u).$$

Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ .  $\lambda$  est valeur propre de  $M$  si et seulement si  $\det(\lambda I_n - M) = 0$ . On note  $\chi_M$  l'application

$$\chi_M : \mathbf{K} \rightarrow \mathbf{K}; \lambda \mapsto \det(\lambda I_n - M).$$

Autrement dit  $\chi_M$  est aussi l'application associée, comme au dessus, à l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé à  $M$ . On a alors la remarque évidente suivante :

**Remarque 3.4.1.** — Soit  $\mathcal{B}$  une base de  $\mathbf{E}$  et  $M$  la matrice de  $u$  dans  $\mathcal{B}$ . Alors pour tout élément  $\lambda$  de  $\mathbf{K}$ ,

$$\chi_M(\lambda) = \chi_u(\lambda).$$

**Proposition 3.4.2.** — Soient  $A$  et  $B$  des éléments de  $\mathcal{M}_n(\mathbf{K})$  semblables. Alors :

$$\chi_A = \chi_B$$

*Preuve de la proposition 3.4.2.* — Par hypothèse il existe un élément  $P$  de  $\text{GL}_n(\mathbf{K})$  tel que  $B = PAP^{-1}$ . Soit alors un élément  $\lambda$  de  $\mathbf{K}$ ,

$$\chi_B(\lambda) = \det(\lambda I_n - B) = \det(\lambda PP^{-1} - PAP^{-1}) = \det(P(\lambda I_n - A)P^{-1}).$$

Donc

$$\chi_B(\lambda) = \det(P)\det(\lambda I_n - A)\det(P^{-1}) = \det(P)\det(P^{-1})\det(\lambda I_n - A) = \chi_A(\lambda).$$

Comme  $\lambda$  est quelconque on a :  $\chi_A = \chi_B$ .

On aurait pu également prouver ce résultat en montrant que  $B$  est la matrice de l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé à  $A$ , dans une certaine base et utiliser la remarque précédente.  $\square$

Étudions maintenant  $\chi_u$ .

**Proposition-définition 3.4.3.** — Soit  $u$  un endomorphisme de  $\mathbf{E}$  et  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . L'application  $\chi_u$  (resp.  $\chi_M$ ) est polynomiale.

Le polynôme associé encore noté  $\chi_u$  (resp.  $\chi_M$ ) est appelé polynôme caractéristique de  $u$  (resp.  $M$ ).

*Preuve de la proposition 3.4.3.* — D'après la remarque 3.4.1., il suffit de traiter le cas matriciel. Notons  $(m_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$  la matrice  $M$  et pour tout élément  $\lambda$  de  $\mathbf{K}$ ,  $(a_{i,j}(\lambda))_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$  la matrice  $\lambda I_n - M$ . Ainsi a-t-on, pour tout  $(i,j) \in \{1,\dots,n\}$  et tout élément  $\lambda$  de  $\mathbf{K}$  :

$$\begin{cases} a_{i,j}(\lambda) = -m_{i,j}, & \text{pour } i \neq j, \\ a_{i,i}(\lambda) = \lambda - m_{i,i}. \end{cases}$$

L'application  $a_{i,j} : \mathbf{K} \rightarrow \mathbf{K}; \lambda \mapsto a_{i,j}(\lambda)$  est donc, pour tout  $(i,j) \in \{1,\dots,n\}$ , polynomiale.

Or pour tout élément  $\lambda$  de  $\mathbf{K}$ ,

$$\chi_M(\lambda) = \sum \dots$$

#### Étude des coefficients du polynôme caractéristique —

Observons pour commencer que le degré de  $\chi_M$  est *a priori* inférieur ou égal à  $n$ , en effet :

— ETUDE DU TERME DE DEGRÉ  $n$  :

$\chi_M$  est de degré  $n$  et son coefficient dominant est 1.

— ETUDE DU TERME DE DEGRÉ  $n - 1$  :

— ETUDE DU TERME DE DEGRÉ 0 :

Au total :

$$\chi_M = X^n + \dots \quad X^{n-1} + \dots \quad ,$$

Et donc

$$\chi_u = X^n + \dots \quad X^{n-1} + \dots \quad .$$

Dans le cas particulier où  $n = 2$  on a donc

$$\chi_M = X^2 - \text{Tr}(M)X + \det(M).$$

**Remarques** —

- Une matrice d'ordre  $n$ , ou un endomorphisme défini sur un espace vectoriel de dimension  $n$ , admette au plus  $n$  valeurs propres.
- Pour retrouver rapidement et facilement ces relations, on peut raisonner avec  $M$  diagonale,  $M = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ .  
Pour tout élément  $\lambda$  de  $\mathbf{K}$ ,

$$\det(\lambda I_n - M) = \prod_{i=1}^n (\lambda - \lambda_i) = (\lambda)^n + \sum_{i=1}^n -\lambda_i (\lambda)^{n-1} + \dots + \prod_{i=1}^n -\lambda_i.$$

- Dans la littérature notamment française on trouve une autre définition du polynôme caractéristique :

$$\tilde{\chi}_M : \mathbf{K} \rightarrow \mathbf{K}; \lambda \mapsto \det(M - \lambda I_n).$$

Ce polynôme est relié au précédent par :  $\tilde{\chi}_M = (-1)^n \chi_M$ ,



Désignons toujours par  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . Les valeurs propres de  $u$  (resp.  $M$ ) sont, on l'a vu en début de paragraphe, les racines dans  $\mathbf{K}$  de  $\chi_u$  (resp.  $\chi_M$ ).

**Exemple** — La matrice  $M$  est supposée triangulaire supérieure, alors les valeurs propres de  $M$  sont  $\dots$

**Définition 3.4.4.** — On appelle (ordre de) multiplicité d'une valeur propre  $\lambda$  de  $u$  (resp. de  $M$ ) l'ordre de multiplicité de  $\lambda$  en tant que racine de  $\chi_u$  (resp.  $\chi_M$ )

**Cas d'un polynôme caractéristique scindé** — On suppose que  $\chi_M$  est scindé (c'est notamment le cas lorsque  $\mathbf{K} = \mathbf{C}$ ),  $M$  possède donc  $n$  valeurs propres distinctes ou non,  $\lambda_1, \lambda_2, \dots, \lambda_n$ .

D'une part on a vu :

$$\chi_M = X^n + \dots + X^{n-1} + \dots + \dots . \quad (\text{II.6})$$

D'autre part  $\chi_M$ , polynôme unitaire, vaut :

$$\chi_M = \prod_{i=1}^n (X - \lambda_i). \quad (\text{II.7})$$

En développant cette dernière égalité, on obtient :

$$\chi_M = \dots . \quad (\text{II.8})$$

Par comparaison de (II.6) et (II.8), on obtient les relations :

$$\sum_{i=1}^n \lambda_i = \dots ,$$

$$\prod_{i=1}^n \lambda_i = \dots .$$

**Remarques** — La preuve de ces deux dernières formules est immédiate pour une matrice diagonale, ce qui permet de les retrouver rapidement

Etudions maintenant le polynôme caractéristique d'un endomorphisme induit.

**Proposition 3.4.5.** — Soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$  stable par  $u$ . Notons  $v$  l'endomorphisme induit par  $u$  sur  $\mathbf{F}$ . Alors :  $\chi_v$  divise  $\chi_u$ .

*Preuve de la proposition 3.4.5.—*

**Remarques complémentaires —**

On déduit directement de 3.4.5. que

- les valeurs propres de  $v$  sont des valeurs propres de  $u$  ;
- l'ordre de multiplicité d'une valeur propre  $\lambda$  de  $v$  est inférieur à l'ordre de multiplicité de  $\lambda$ , considéré comme une valeur propre de  $u$  ;
- pour  $\lambda$  valeur propre de  $v$ ,

$$\mathbf{E}_\lambda(v) = \text{Ker}(\lambda \text{id}_{\mathbf{F}} - v) = \mathbf{F} \cap \text{Ker}(\lambda \text{id}_{\mathbf{E}} - u) = \mathbf{F} \cap \mathbf{E}_\lambda(u).$$

Généralisons la proposition précédente.

**Proposition 3.4.6.** — *Soit  $(\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_q)$  une famille de  $q$  sous-espaces vectoriels de  $\mathbf{E}$  supplémentaires et stables par  $u$ . Notons, pour  $i = 1, 2, \dots, q$ ,  $u_i$  l'endomorphisme induit par  $u$  sur  $\mathbf{F}_i$ . Alors :*

$$\chi_u = \cdots$$

*Preuve de la proposition 3.4.6.—*

### 3.5 Diagonalisation

$\mathbf{E}$  est toujours un  $\mathbf{K}$ -espace vectoriel de dimension finie non nulle  $n$  et  $u$  un endomorphisme de  $\mathbf{E}$ .

**Définition 3.5.1.** — On dit que  $u$  est diagonalisable si, par définition, il existe une base  $\mathcal{B}$  de  $\mathbf{E}$ , constituée de vecteurs propres de  $u$ .

**Remarque** Si  $u$  est diagonalisable, alors la matrice de  $u$  dans la base  $\mathcal{B}$ , donnée par la définition précédente, est diagonale; ce qui justifie la terminologie.

**Définition 3.5.2.** — Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . On dit que  $M$  est diagonalisable, si par définition, elle est semblable à une matrice diagonale.

**Proposition 3.5.3.** — Soit  $M$  la matrice de  $u$  dans une base  $\mathcal{B}$  de  $\mathbf{E}$ . Alors  $M$  est diagonalisable, si et seulement si l'endomorphisme  $u$  est diagonalisable.

*Preuve de la proposition 3.5.3.* —

- Supposons que  $u$  soit diagonalisable.

On peut alors trouver une base  $\mathcal{B}'$  de  $\mathbf{E}$  dans laquelle la matrice de  $u$ , est diagonale.  $M$ , matrice de  $u$  dans la base  $\mathcal{B}$  est donc semblable à une matrice diagonale.

- Supposons que  $M$  soit diagonalisable.

La matrice  $M$  est semblable à une matrice diagonale,  $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ , notée  $\Delta$ . On dispose alors d'un élément  $P$  de  $\text{GL}_n(\mathbf{K})$  tel que :

$$\Delta = P^{-1}MP. \quad (\text{II.9})$$

Notons pour  $i = 1, 2, \dots, n$ ,  $P_i$  la  $i^{\text{e}}$  colonne de  $P$  et  $\vec{e}_i$  le vecteur de  $\mathbf{E}$  dont le vecteur colonne des coordonnées dans la base  $\mathcal{B}$  est  $P_i$ .

$$\det_{\mathcal{B}}(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = \text{Det}P \neq 0.$$

Donc  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  est une base de  $\mathcal{M}_{n,1}(\mathbf{K})$  et  $P$  est la matrice de passage de  $\mathcal{B}$  à  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  :  $P = P_{\mathcal{B}}^{(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)}$ . Donc, d'après (II.9),  $\Delta$  est la matrice de  $u$  dans  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ . Donc pour  $i = 1, 2, \dots, n$ ,  $m(\vec{e}_i) = \lambda_i \vec{e}_i$ . La base  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  est donc constituée de vecteurs propres de  $u$ . Autrement dit  $u$  est diagonalisable.

□

On déduit de cette proposition le corollaire immédiat suivant, qui vient renforcer l'identification d'une matrice élément de  $\mathcal{M}_n(\mathbf{K})$  et de l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé.

**Corollaire 3.5.4.** — Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . La matrice  $M$  est diagonalisable, si et seulement si l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé à  $M$  est diagonalisable.

**Remarque** — Il résulte de la proposition 1.1.3 et de son corollaire, que dans la pratique, on étudie surtout la diagonalisation de matrices, le cadre des endomorphismes étant plutôt réservé aux exercices théoriques. C'est dans cet esprit que nous énoncerons les résultats pour des endomorphismes, laissant au lecteur le soin d'en donner une traduction matricielle.

Donnons le premier critère du caractère diagonalisable d'un endomorphisme.

**Théorème 3.5.5.** — L'endomorphisme  $u$  est diagonalisable si et seulement si la somme de ses espaces propres est égale à  $E$ .

**Remarque** — On a vu (3.3.6.) que les espaces propres de  $u$  sont en somme directe, donc le théorème précédent peut s'énoncer : L'endomorphisme  $u$  est diagonalisable si et seulement si ses espaces propres sont des sous-espaces supplémentaires.

*Preuve du théorème 3.5.5.* —

- L'endomorphisme  $u$  est diagonalisable.

On note  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$  les sous-espaces propres deux à deux distincts de  $u$ . Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base de  $\mathbf{E}$  constituée de vecteurs propres de  $u$  que nous noterons  $\mathcal{B}$ . Remarquons que l'existence de  $\mathcal{B}$  justifie celle d'au moins un espace propre.

- La somme des espaces propres de  $u$  est égale à  $\mathbf{E}$

**Remarque 3.5.6.** — Supposons  $u$  diagonalisable, avec les notations précédentes on a donc :

$$\mathbf{E} = \bigoplus_{j=1}^k \mathbf{E}_j.$$

Notons pour  $j = 1, 2, \dots, k$ ,  $p_j$  la projection sur  $\mathbf{E}_j$  associée à cette décomposition de  $\mathbf{E}$ . On a vu dans le cours sur les espaces vectoriels que

$$\sum_{j=1}^k p_j = \text{id}_E. \quad (\text{II.10})$$

Soit  $\vec{x}$  un élément de  $\mathbf{E}$ , d'après (XII.5),

$$\vec{x} = \sum_{j=1}^k p_j(\vec{x}),$$

donc

$$u(\vec{x}) = \sum_{j=1}^k u(p_j(\vec{x})) = \sum_{j=1}^k \lambda_j p_j(\vec{x}),$$

où pour  $j = 1, 2, \dots, k$ ,  $\lambda_j$  est la valeur propre associée à  $\mathbf{E}_j$ . Donc,  $\vec{x}$  étant quelconque,

$$u = \sum_{j=1}^k \lambda_j p_j.$$

Par ailleurs nous avons observé que pour  $j = 1, 2, \dots, k$ ,  $\mathbf{E}_j$  est stable par  $u$  et que  $u$  induit sur  $\mathbf{E}_j$  une homothétie. Au total  $\mathbf{E}$  est la somme directe de sous-espaces stables sur lesquels  $u$  induit une homothétie.

Tirons du théorème 3.5.5. un critère plus pratique donnant le caractère diagonalisable d'un endomorphisme.

Les sous-espaces propres d'un endomorphisme  $u$  sont en somme directe, donc la dimension de leur somme est la somme de leur dimension. Donc on déduit de 3.5.5 le critère suivant :

**Proposition 3.5.7.** — *L'endomorphisme  $u$  est diagonalisable si et seulement si la somme des dimensions de ses sous-espaces propres est la dimension de  $\mathbf{E}$ .*

Nous voilà en mesure de comprendre mieux pourquoi une matrice, comme nous le verrons, peut ne pas être diagonalisable. Tout repose sur la proposition suivante

**Proposition 3.5.8.** — *L'ordre de multiplicité d'une valeur propre est supérieur ou égal à la dimension de l'espace propre associé*

*Preuve de la proposition 3.5.8.* — Soit  $\lambda$  une valeur propre de  $u$ . Notons  $m$  son ordre de multiplicité et notons  $d$  la dimension de l'espace propre associé.

Déduisons de cette proposition un nouveau critère du caractère diagonalisable d'un endomorphisme :

**Proposition 3.5.9.** — *Supposons que  $\text{sp}(u) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ , (les valeurs propres  $\lambda_1, \lambda_2, \dots, \lambda_k$  étant deux à deux distinctes). Pour  $i = 1, 2, \dots, k$ , notons  $m_i$  la multiplicité de  $\lambda_i$  et  $d_i$  la dimension de l'espace propre associé,  $\mathbf{E}_{\lambda_i}$ . Alors  $u$  est diagonalisable si et seulement si*

- $\chi_u$  est scindé dans  $\mathbf{K}[X]$  ;
- Pour  $i = 1, 2, \dots, k$ ,  $d_i = m_i$ .

*Preuve de la proposition 3.5.9.* —

**Corollaire 3.5.10.** — *Si  $\chi_u$  est scindé dans  $\mathbf{K}[X]$  à racines simples, alors  $u$  est diagonalisable.*

*Preuve du corollaire 3.5.10.* — Avec les notations de la proposition 1.4.10, on a  $k = n$  et, pour  $i = 1, 2, \dots, k$ ,  $m_i = 1$ . Pour tout élément  $i$  de  $\{1, 2, \dots, n\}$ , d'une part  $d_i \leq m_i = 1$  (cf. 1.4.9.), et d'autre part  $d_i \geq 1$ , car  $\mathbf{E}_{\lambda_i} \neq \{\vec{0}_{\mathbf{E}}\}$ . Donc, pour tout élément  $i$  de  $\{1, 2, \dots, n\}$ ,  $m_i = d_i (= 1)$  et de plus  $\chi_u$  est scindé, donc d'après 1.4.10,  $u$  est diagonalisable.

□

On donnera une autre preuve de ce résultat dans la partie suivante.

#### PRATIQUE DE LA DIAGONALISATION D'UNE MATRICE

Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . Dans la pratique on se limite à  $n = 2, 3$  ou très exceptionnellement 4.

- On calcul  $\chi_M$  et on détermine  $\lambda_1, \lambda_2, \dots, \lambda_k$  les valeurs propres de  $M$  deux à deux distinctes.
- Si  $\chi_M$  est non scindé, alors  $M$  n'est pas diagonalisable, sinon on détermine une base  $\mathcal{B}_i$  de chaque espace propre,  $E_{\lambda_i}(M)$ , pour  $i = 1, 2, \dots, k$ .
- Pour tout élément  $i$  de  $\{1, 2, \dots, k\}$ , on note  $d_i$  la dimension de  $E_{\lambda_i}(M)$  et  $m_i$  l'ordre de multiplicité de  $\lambda_i$ . S'il existe un élément  $i$  de  $\{1, 2, \dots, k\}$  tel que  $d_i < m_i$  alors  $M$  n'est pas diagonalisable. Sinon  $M$  est diagonalisable et  $\bigoplus_{i=1}^k \mathbf{E}_{\lambda_i}(M) = \mathbf{E}$ . Donc  $\mathcal{B}$  concaténée des bases  $\mathcal{B}_1, \mathcal{B}_1, \dots, \mathcal{B}_k$ ,

$$\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_1, \dots, \mathcal{B}_k),$$

est une base de  $\mathcal{M}_{n,1}(\mathbf{K})$ . Cette base est constituée de vecteurs propres de  $M$ . On pose

$$P = \mathcal{P}_{\mathcal{B}_c, \mathcal{B}},$$

$\mathcal{B}_c$  étant la base canonique de  $\mathcal{M}_{n,1}(\mathbf{K})$ . On pose également

$$\Delta = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{d_1}, \underbrace{\lambda_2, \dots, \lambda_2}_{d_2}, \dots, \underbrace{\lambda_k, \dots, \lambda_k}_{d_k}).$$

La matrice  $\Delta$  est la matrice de l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé à  $M$  dans la base  $\mathcal{B}$ , donc

$$M = P\Delta P^{-1}.$$

Au travail!

Déterminer si les matrices suivantes sont diagonalisables et diagonaliser celles qui le sont.

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

**Remarque** — Pour une matrice donnée par un tableau de réels, on doit préciser si l'on considère la matrice comme élément de  $\mathcal{M}_n(\mathbf{R})$  ou de  $\mathcal{M}_n(\mathbf{C})$ , en effet elle peut être diagonalisable dans le second espace sans l'être dans le premier.

Reprenons l'exemple :  $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ .

**Remarque** — Pour  $n \geq 2$ , il existe des matrices non diagonalisables même dans  $\mathcal{M}_n(\mathbf{C})$ . Soit en effet la matrice

élément de  $\mathcal{M}_n(\mathbf{C})$  :

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

Terminons par quelques remarques utiles.

**Remarques** —

- Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{R})$ , dont le polynôme caractéristique n'est pas scindé dans  $\mathbf{R}$  mais qui est diagonalisable dans  $\mathcal{M}_n(\mathbf{C})$ .  $M$  admet donc des valeurs propres complexes non réelles ; soit  $\lambda$  une d'entre elles. La matrice  $M$  étant à coefficients réels,  $\chi_M$  est à coefficients réels. Il admet donc  $\bar{\lambda}$  comme racine complexe. Soit  $X$  un vecteur propre élément de  $\mathcal{M}_{n,1}(\mathbf{C})$ , associé à la valeur propre  $\lambda$ . Par définition,  $MX = \lambda X$ . La conjugaison étant un automorphisme du corps  $\mathbf{C}$ , on a par conjugaison  $\overline{MX} = \overline{\lambda X} = \bar{\lambda} \bar{X}$ . La matrice  $M$  étant à coefficients réels on a donc :

$$M\bar{X} = \bar{\lambda}\bar{X}.$$

Autrement dit  $\bar{X}$ , qui est non nul, est un vecteur propre de  $M$  associé à la valeur propre  $\bar{\lambda}$ . Cette remarque permet de diagonaliser plus rapidement une matrice réelle dans  $\mathcal{M}_n(\mathbf{C})$ .

- Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$  diagonalisable. Alors  ${}^tM$  est diagonalisable. En effet il existe un élément  $P$  de  $\text{GL}_n(\mathbf{K})$  et un élément  $\Delta$  de  $\mathcal{M}_n(\mathbf{K})$  diagonal tels que

$$M = P\Delta P^{-1}.$$

Donc par transposition

$${}^tM = {}^t(P^{-1}) \Delta {}^tP.$$



Or  ${}^t(P^{-1}) = ({}^tP)^{-1}$ , donc  ${}^tM$  est diagonalisable. Pour une matrice  $M$  non nécessairement diagonalisable, on a, dans le même ordre d'idée, que  $M$  et  ${}^tM$  ont même polynôme caractéristique; ce qui résulte directement de la définition du polynôme caractéristique et du fait que le déterminant d'une matrice est égal à celui de sa transposée.

- Dans certains cas il est totalement inutile de calculer le polynôme caractéristique (et du reste bien difficile) pour déterminer les éléments propres d'une matrice. Donnons un exemple. Soit  $M = A^tA$ , où  $A$  est un élément de  $\mathcal{M}_{n,1}(\mathbf{R})$  (cas particulier de l'exemple 2.3.4-2).

Nous verrons d'autres exemples, un en fin de chapitre, ainsi qu'en exercice.

# TRIGONALISATION

## 4.1 Définitions, Premières propriétés

$\mathbf{E}$  est toujours un  $\mathbf{K}$ -espace vectoriel de dimension finie non nulle  $n$ ,  $u$  un endomorphisme de  $\mathbf{E}$  et  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . Comme toujours, le programme nous limite à  $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{C}$ , sans aucune nécessité mathématiques.

Il arrive, on l'a vu qu'une matrice ou qu'un endomorphisme ne soit pas diagonalisable. Nous allons donner une forme de réduction plus faible que la diagonalisation

**Définition 4.1.1.** — *On dit que  $u$  est trigonalisable si, par définition, il existe une base  $\mathcal{B}$  de  $\mathbf{E}$ , telle que la matrice de  $u$  dans  $\mathcal{B}$  soit triangulaire supérieure.*

**Définition 4.1.2.** — *On dit que  $M$  est trigonalisable, si par définition, elle est semblable à une matrice triangulaire supérieure.*

**Proposition 4.1.3.** — *Si  $M$  est la matrice de  $u$  dans une base  $\mathcal{B}$  de  $\mathbf{E}$ . Alors  $M$  est trigonalisable, si et seulement si l'endomorphisme  $u$  est trigonalisable.*

La preuve de cette proposition est similaire à celle de la proposition 3.5.3.

On déduit de cette proposition le corollaire immédiat suivant, qui vient renforcer l'identification d'une matrice élément de  $\mathcal{M}_n(\mathbf{K})$  et de l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé.

**Corollaire 4.1.4.** — *La matrice  $M$  est trigonalisable, si et seulement si l'endomorphisme de  $\mathcal{M}_{n,1}(\mathbf{K})$  canoniquement associé à  $M$  est trigonalisable.*

**Remarque** — Il résulte de la proposition 4.1.3. et de son corollaire, que dans la pratique, on étudie surtout la trigonalisation de matrices, le cadre des endomorphismes étant plutôt réservé aux exercices théoriques.

Donnons un critère simple du caractère trigonalisable.

**Théorème 4.1.5** — CRITÈRE DE TRIGONALISATION —

*L'endomorphisme  $u$  est trigonalisable si et seulement si son polynôme caractéristique est scindé dans  $\mathbf{K}[X]$ .*

*La matrice  $M$  est trigonalisable si et seulement si son polynôme caractéristique est scindé dans  $\mathbf{K}[X]$ .*

On déduit de ce théorème le corollaire immédiat suivant :

**Corollaire 4.1.6.** — *Tout endomorphisme d'un  $\mathbf{C}$ -espace vectoriel de dimension finie non nulle est trigonalisable. Tout élément de  $\mathcal{M}_n(\mathbf{C})$  est trigonalisable*

*Preuve du théorème 4.1.6.* —

- HYPOTHÈSE *On suppose  $M$  trigonalisable.*

La matrice  $M$  est semblable à une matrice triangulaire supérieure  $T$  (cf. 3.1.2.).  $T$  est de la forme

$$T = \begin{pmatrix} a_0 & \times & \dots & \dots & \times \\ 0 & a_1 & \times & \dots & \times \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n-1} & \times \\ 0 & \dots & \dots & 0 & a_n \end{pmatrix},$$

avec  $a_1, a_2, \dots, a_n$  éléments de  $\mathbf{K}$ . Donc  $\chi_M = \dots$

- RÉCIPROQUE

...

## PRATIQUE DE LA TRIGONALISATION D'UNE MATRICE

La preuve du caractère trigonalisable d'une matrice à polynôme caractéristique scindé fournit une méthode algorithmique de trigonalisation, toutefois la pratique de la trigonalisation n'est pas un objectif du programme. A l'exception de cas fort simples, comme le cas des matrices d'ordre 2 et dans certains cas celles d'ordre 3.

• Prenons  $M$  un élément de  $\mathcal{M}_2(\mathbf{K})$ , dont le polynôme caractéristique est scindé (quitte éventuellement à se placer dans  $\mathcal{M}_2(\mathbf{C})$ ). Si  $M$  a deux valeurs propres distinctes, alors elle est diagonalisable. Prenons les cas où  $M$  n'a qu'une valeur propre double,  $\lambda$ . Si  $\mathbf{E}_\lambda$  est de dimension 2 alors  $M = \lambda I_2$  et l'on a dû s'en rendre compte! sinon en prenant  $E_1$ , vecteur propre de  $M$ , et si l'on complète la famille  $(E_1)$  en une base  $(E_1, E_2)$  (Dans la pratique il est souvent plus simple de prendre pour  $E_2$  un des deux vecteurs de la base canonique). La matrice de l'endomorphisme canoniquement associée à  $M$  dans la base  $(E_1, E_2)$  est de la forme :

$$\begin{pmatrix} \lambda & a \\ 0 & b \end{pmatrix}.$$

Notons de plus que  $b = \dots$

Autrement dit  $M$  est semblable à  $\begin{pmatrix} \lambda & a \\ 0 & \dots \end{pmatrix}$ .

**NB.** La valeur de  $a$  dépend du choix de la base!

**Exemple** — Trigonalisons l'élément  $A$  de  $\mathcal{M}_2(\mathbf{R})$  suivant.

$$A = \frac{1}{2} \begin{pmatrix} 9 & 5 \\ -5 & -1 \end{pmatrix}$$

• Prenons  $M$  un élément de  $\mathcal{M}_3(\mathbf{K})$  dont le polynôme caractéristique est scindé (quitte éventuellement à se placer dans  $\mathcal{M}_3(\mathbf{C})$ ). Si les trois valeurs propres sont distinctes alors  $M$  est diagonalisable.

- Sinon commençons par le cas où la matrice admet deux valeurs propres  $\lambda_1$  d'ordre 1 et  $\lambda_2$  d'ordre 2. On exclut le cas où  $\mathbf{E}_{\lambda_2}$  est de dimension 2, cas où la matrice est diagonalisable. Donc  $\mathbf{E}_{\lambda_1}$  et  $\mathbf{E}_{\lambda_2}$  sont tous deux de dimension 1

et il est loisible de choisir un vecteur propre  $E_1$  dirigeant  $\mathbf{E}_{\lambda_1}$ , un vecteur propre  $E_2$  dirigeant  $\mathbf{E}_{\lambda_2}$ . Alors

**Exemple** — Trigonaliser l'élément  $B$  de  $\mathcal{M}_3(\mathbf{R})$  suivant :

$$B = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 2 & 0 \end{pmatrix}$$

- Envisageons le cas où  $M$  admet une seule valeurs propre  $\lambda$  d'ordre de multiplicité 3. Oublions le cas où  $\dim(\mathbf{E}_\lambda) = 3$ , la matrice est alors diagonale! et regardons le cas où  $\dim(\mathbf{E}_\lambda) = 2$ .....

**Exemple** – Trigonalisons l'élément  $C$  de  $\mathcal{M}_3(\mathbf{R})$  suivant :

$$C = \frac{1}{2} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ -1 & 1 & 4 \end{pmatrix} \tag{II.11}$$

Reste le cas où  $\dim(\mathbf{E}_\lambda) = 1$ , beaucoup plus pénible, le programme semble suggérer de dédaigner cette occupation et de faire appel le cas échéant à un logiciel de calcul formel.

## 4.2 Applications de la trigonalisation

Outre des applications théoriques qui seront vu lors d'un prochain chapitre, la trigonalisation permet, lorsque la diagonalisation est impossible, même en se plaçant dans  $\mathbf{C}$ , de résoudre des système différentiels. Nous étudierons la chose en détail lors d'un chapitre sur les équations différentielles. Dans certains cas la trigonalisation permet le calcul des puissances de matrices, mais nous verrons que ce n'est pas toujours le cas. Elle sert également à la détermination de suites données par des relations de récurrence

SYSTÈME DIFFÉRENTIEL

### Exercice 4.2.1. —

Déterminer les triplets d'applications  $(\phi, \chi, \psi)$  de  $\mathbf{R}$  dans  $\mathbf{R}$ , de classe  $\mathcal{C}^1$  tels que :

$$\begin{cases} \phi' = 2\phi + 1\chi - \psi, \\ \chi' = 2\chi, \\ \psi' = \phi + 2\chi. \end{cases}$$

*Solution de l'exercice 4.2.1. —*

PUISSANCES DE MATRICES ET SUITES À RÉCURRENCES LINÉAIRES

**Exercice 4.2.2.** — Déterminer pour tout  $n \in \mathbf{N}^*$  la puissance  $n^e$  de la matrice suivante

$$A = \frac{1}{2} \begin{pmatrix} 9 & 5 \\ -5 & -1 \end{pmatrix}$$

*Solution de l'exercice 4.2.2. —*

**Exercice 4.2.3.** — Déterminer toutes les suites réelles  $(u_n)_{n \in \mathbf{N}}$ ,  $(v_n)_{n \in \mathbf{N}}$ ,  $(w_n)_{n \in \mathbf{N}}$  telles que :

$$\begin{cases} 2u_{n+1} = u_n + v_n + 2w_n, \\ 2v_{n+1} = u_n + v_n - 2w_n, \\ 2w_{n+1} = -u_n + v_n + 4w_n, \end{cases} \quad (\text{II.12})$$

pour tout entier  $n \geq 0$ .

*Solution de l'exercice 4.2.3.* — Soient  $(u_n)_{n \in \mathbf{N}}$ ,  $(v_n)_{n \in \mathbf{N}}$ ,  $(w_n)_{n \in \mathbf{N}}$  des suites réelles. Posons pour tout entier  $n \geq 0$ ,

$$U_n = \quad \text{et } M =$$

La matrice  $M$  a déjà été trigonalisée dans le précédent paragraphe.

Donnons deux méthodes.

**Mise en garde** — En général trigonaliser une matrice ne suffit pas au calcul de ses puissances, .....

On verra en exercice dans un chapitre prochain une réduction plus forte qui permet elle, ce calcul.



Par contre la trigonalisation donne la diagonale des puissances  $n^e$  d'une matrice, ce qui, on le verra, n'est pas sans intérêt. Précisément :

Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$  (donc de  $\mathcal{M}_n(\mathbf{C})$ ) On note  $\lambda_1, \dots, \lambda_n$  les  $n$  valeur propres complexes, distinctes ou non de  $M$ . Alors  $M$  est dans  $\mathcal{M}_n(\mathbf{C})$  semblable à une matrice  $T$  de la forme

$$T = \begin{pmatrix} \lambda_1 & \times & \times & \cdots & \times \\ 0 & \lambda_2 & \times & \cdots & \times \\ 0 & 0 & \lambda_3 & & \times \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Notons que l'on retrouve ainsi, la trace et le déterminant étant des invariants par la relation de similitude, les formules vue en 3.4.

$$\mathrm{Tr}(M) = \lambda_1 + \cdots + \lambda_n ; \det(M) = \lambda_1 \times \cdots \times \lambda_n.$$

On montre par une récurrence immédiate que  $T^m$  est de la forme :

$$T^m = \begin{pmatrix} & \times & \times & \cdots & \times \\ 0 & & \times & \cdots & \times \\ 0 & 0 & & & \times \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \end{pmatrix},$$

pour tout entier  $m \geq 0$ . Donc pour tout entier  $m \geq 0$ ,

$\mathrm{Tr}(M^m) =$	$;$	$\det(M^m) =$
----------------------	-----	---------------

Voici un exercice qui peut utiliser ce résultat.

**Exercice 4.2.4.** — Déterminer les valeurs propres de la matrice suivante. Est-elle diagonalisable ?

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

*Solution de l'exercice 4.2.4.* —

Pour rebondir sur la mise en garde donnons quelques résultats sur les matrices et les endomorphismes nilpotents. Nous leur consacrerons lors d'un prochain chapitre un peu plus d'intérêt

### 4.3 Endomorphismes et matrices nilpotentes

Rappelons pour commencer les définitions.

**Définition 4.3.1.** — Un endomorphisme  $u$  de  $\mathbf{E}$  est dit nilpotent, si par définition il existe un entier  $k \geq 1$  tel que  $u^k = 0_{\mathcal{L}(\mathbf{E})}$ . Si  $u$  est nilpotent on appelle ordre de nilpotence de  $u$ , le plus petit entier naturel  $k$  tel que  $u^k = 0_{\mathcal{L}(\mathbf{E})}$ .

Une matrice  $M$  élément de  $\mathcal{M}_n(\mathbf{K})$  est dite nilpotente, si par définition il existe un entier  $k \geq 1$  tel que  $u^k = 0_n$ . Si  $M$  est nilpotent on appelle ordre de nilpotence de  $M$ , le plus petit entier naturel  $k$  tel que  $u^k = 0_n$ .

#### Remarques

1. L'ordre de nilpotence d'un endomorphisme nilpotent (ou d'une matrice) est bien défini puisque  $\{k \in \mathbf{N}^* \mid u^k = 0_{\mathcal{L}(\mathbf{E})}\}$  est une partie de  $\mathbf{N}$  non vide.
2. Un endomorphisme (ou une matrice) est nilpotent d'ordre 1 si et seulement si il nul.
3. Si pour un entier  $k \geq 0$ ,  $u^k = 0_{\mathcal{L}(\mathbf{E})}$ , alors pour tout entier  $k' \geq k$ ,  $u^{k'} = 0_{\mathcal{L}(\mathbf{E})}$ . Donc un entier  $p \geq 1$  est l'ordre de nilpotence d'un endomorphisme  $u$  si et seulement si :

$$u^p = 0_{\mathcal{L}(\mathbf{E})} \text{ et } u^{p-1} \neq 0_{\mathcal{L}(\mathbf{E})}.$$

Voici le lien entre matrice nilpotente et endomorphisme nilpotent.

**Proposition 4.3.2.** — Soient  $u$  un endomorphisme de  $\mathbf{E}$  dont la matrice dans une base  $\mathcal{B}$  de  $\mathbf{E}$  est  $M$ , et un entier  $p \geq 1$ . Alors  $u$  est nilpotent d'ordre  $p$  si et seulement si  $M$  est nilpotente d'ordre  $p$ .

*Preuve de la proposition 4.3.2.* — Comme  $\text{Mat}_{\mathcal{B}}$  est un isomorphisme de  $\mathbf{E}$  sur  $\mathcal{M}_n(\mathbf{K})$ , on a pour tout entier naturel  $k$  que  $M^k$  est la matrice de  $u^k$  dans  $\mathcal{B}$  et est nul si et seulement si  $u^k$  l'est : le résultat en découle immédiatement.

□

**Proposition 4.3.3.** — L'ordre de nilpotence d'un endomorphisme  $u$  de  $\mathbf{E}$  (ou d'un élément de  $\mathcal{M}_n(\mathbf{K})$ ) est inférieur ou égal à  $n$ , dimension de  $\mathbf{E}$

*Preuve de la proposition 4.3.3.* —

□

Nous donnerons dans un prochain chapitre une preuve plus fulgurante.

L'ordre d'un endomorphisme nilpotent de  $\mathbf{E}$  (ou d'un élément de  $\mathcal{M}_n(\mathbf{K})$ ) peut être égal à  $n$ . Ainsi la matrice

$$N =$$

est d'ordre  $n$ .

La preuve de 4.3.3. montre que tout endomorphisme de  $\mathbf{E}$  d'ordre de nilpotence égal à  $n$ , dimension de  $\mathbf{E}$  admet  $N$

comme matrice dans une certaine base.

Passons au lien avec la trigonalisation.

**Proposition 4.3.4.** — Une matrice  $M$ , élément de  $\mathcal{M}_n(\mathbf{K})$  est nilpotente si et seulement si elle trigonalisable et a toutes ses valeurs propres nulles.

*Preuve de la proposition 4.3.4.* —



## Chapitre III

# ESPACES VECTORIELS NORMÉS

Dans tout ce chapitre  $\mathbf{K}$  désigne indifféremment le corps des nombres complexes ou celui des nombres réels, dans le premier cas,  $|\cdot|$  désigne le module, dans le second la valeur absolue.

Nous noterons en général les éléments des  $\mathbf{K}$ -espaces vectoriels que nous rencontrerons en minuscules italiques, surmontées d'une flèche :  $\vec{x}, \vec{y}, \vec{z}, \dots$ , les éléments de  $\mathbf{K}$  en caractères minuscules italiques :  $r, \alpha, \beta, \dots x, y, z, \dots$ . Les ensembles sans structure particulière seront notés par des majuscules italiques :  $A, B, E, F, \dots$  les espaces vectoriels, sous-espaces vectoriels, par des majuscules romanes grasses :  $\mathbf{E}, \mathbf{F}, \dots$

## RUDIMENTS DE TOPOLOGIE

La topologie du grec  $\tau\omicron\pi\omicron\varsigma$  lieu et  $\lambda\omicron\gamma\omicron\varsigma$  discours, formalise les notions intuitives « d'être proche » de « tendre vers » et fournit des bases rigoureuses aux définitions de la limite d'une suite, d'une application, ou de la continuité. Conformément au programme, nous nous bornerons ici à ne définir et étudier que celles des topologies qui proviennent d'une norme, c'est-à-dire les plus conformes au sens commun. Le plan peut servir assez bien de support à notre intuition en tout cas tant que nous travaillons en dimension finie, il est donc souhaitable dans l'élaboration des raisonnements de recourir assez systématiquement à une figure plane. On ne fait jamais assez de dessins !

### 1.1 Distances, normes

Commençons par définir une notion de distance fort conforme au sens commun.

**Définition 1.1.1.** — DISTANCE —

On appelle distance sur un ensemble  $E$ , toute application  $d$ , de  $E \times E$  dans  $\mathbf{R}$ , à valeurs positives, qui jouit des propriétés suivantes :

1. Quels que soient  $x$  et  $y$  éléments de  $E$ ,  $d(x, y) = d(y, x)$  (symétrie).
2. Quels que soient  $x$ , et  $y$  éléments de  $E$ ,  $d(x, y) = 0$  si et seulement si  $x = y$  (séparation).
3. Quels que soient  $x, y$  et  $z$  éléments de  $E$ ,  $d(x, y) \leq d(x, z) + d(z, y)$  (inégalité triangulaire).

**Exemples 1.1.2.** — On vérifie sans mal que les applications suivantes sont des distances sur  $\mathbf{R}$  :

1.  $d : \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}; (x, y) \mapsto |y - x|$ .
2.  $\tilde{d} : \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} 1, & \text{pour } x \neq y \\ 0, & \text{pour } x = y \end{cases}$

La première distance est particulière en ce sens qu'elle provient d'une *norme*, en l'occurrence la valeur absolue ; plus précisément :

**Définition 1.1.3.** — NORME —

Soit  $\mathbf{E}$  un espace vectoriel sur le corps  $\mathbf{K}$ . On appelle norme sur  $\mathbf{E}$ , toute application  $N$  de  $\mathbf{E}$  dans  $\mathbf{R}$  à valeurs positives, qui jouit des propriétés suivantes : pour tout  $\vec{x}$  et tout  $\vec{y}$  éléments de  $\mathbf{E}$  et tout  $\lambda$  élément de  $\mathbf{K}$ ,

1.  $N(\vec{x}) = 0$  si et seulement si  $\vec{x} = \vec{0}_{\mathbf{E}}$ .
2.  $N(\lambda \vec{x}) = |\lambda| N(\vec{x})$  (homogénéité).
3.  $N(\vec{x} + \vec{y}) \leq N(\vec{x}) + N(\vec{y})$  (inégalité triangulaire).

Par exemple, la valeur absolue sur  $\mathbf{R}$ , le module sur  $\mathbf{C}$  sont des normes.

Par analogie avec ces exemples, on préfère noter la norme d'un vecteur  $\vec{x}$  de  $\mathbf{E}$ ,  $|\vec{x}|$ ,  $\|\vec{x}\|$ ,  $\|\vec{x}\|_{\mathbf{E}}$ ,  $\|\vec{x}\|$  etc. plutôt que  $N(\vec{x})$ . L'application  $N$  se note alors  $|\cdot|$ ,  $\|\cdot\|$ ,  $\|\cdot\|_{\mathbf{E}}$ ,  $\|\cdot\|$  ...

**Proposition 1.1.4.**— Soit  $N$  une norme sur un  $\mathbf{K}$ -espace vectoriel  $\mathbf{E}$ . Pour tout  $\vec{x}$  et tout  $\vec{y}$  éléments de  $\mathbf{E}$ ,

$$|N(\vec{x}) - N(\vec{y})| \leq N(\vec{x} - \vec{y}).$$

Ce résultat, parfois appelé *deuxième inégalité triangulaire*, est connu de longue date pour la valeur absolue, la preuve est analogue dans le cas général. On verra dans la suite de ce cours que cette propriété ne dit rien d'autre que la norme est une application 1-lipschitzienne de  $(\mathbf{E}, N)$  dans  $(\mathbf{R}, |\cdot|)$ .

Associons à chaque norme une distance, grâce au résultat suivant :

**Proposition 1.1.5.** — Soit  $N$  une norme sur un  $\mathbf{K}$ -espace vectoriel  $\mathbf{E}$ . L'application

$$d : \mathbf{E} \times \mathbf{E} \rightarrow \mathbf{R}; (\vec{x}, \vec{y}) \mapsto N(\vec{x} - \vec{y})$$

est une distance sur  $\mathbf{E}$ . On l'appelle distance associée à la norme  $N$ . On dit encore que la distance  $d$  provient de la norme  $N$ .

*Preuve de la proposition 1.1.5.* —

La distance  $\tilde{d}$  de l'exemple 1.1.2. ne provient pas d'une norme. En effet, une distance sur un espace vectoriel  $\mathbf{E}$  qui provient d'une norme se comporte bien vis à vis des translations et des homothéties ou, si l'on préfère, des lois de l'espace vectoriel ; plus précisément, si la distance  $d$  provient d'une norme alors pour tout  $\vec{x}$  tout  $\vec{y}$  tout  $\vec{z}$  vecteurs de  $\mathbf{E}$  et tout  $\lambda$  élément de  $\mathbf{R}$ , on a :

$$d(\lambda\vec{x}, \lambda\vec{y}) = |\lambda| d(\vec{x}, \vec{y}), \quad (\text{III.1})$$

$$d(\vec{x} + \vec{z}, \vec{y} + \vec{z}) = d(\vec{x}, \vec{y}). \quad (\text{III.2})$$

Le lecteur établira à titre d'exercice une réciproque : *une distance qui satisfait à (III.1–III.2) provient d'une norme.*

Donnons divers exemples de normes.

**Exemples 1.1.7.** —

1. NORMES SUR  $\mathbf{K}^n$

Désignons par  $n$  un entier naturel non nul.

(a) L'application

$$n_{\infty} : \mathbf{K}^n \rightarrow \mathbf{R}; (x_1, x_2, \dots, x_n) \mapsto \sup \{|x_i|, i \in \{1, 2, \dots, n\}\}$$

définit une norme sur  $\mathbf{K}^n$ .

(b) Pour  $p$  réel supérieur ou égal à 1, l'application

$$n_p : \mathbf{K}^n \rightarrow \mathbf{R}; (x_1, x_2, \dots, x_n) \mapsto \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}$$

définit une norme sur  $\mathbf{K}^n$ .

Les cas  $p = 1$  et  $p = 2$  sont au programme. Le cas général est hors programme et fait l'objet d'un des exercices.

## 2. NORMES SUR $\mathcal{C}^0([a, b], K)$

Dans la suite  $[a, b]$  désigne un segment de  $\mathbf{R}$  non réduit à un point et  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([a, b], K)$  des applications de  $[a, b]$  dans  $\mathbf{K}$  continues.

(a) L'application

$$\mathbf{E} \rightarrow \mathbf{R} : f \mapsto \sup\{|f(x)|, x \in [a, b]\}$$

est une norme sur  $\mathbf{E}$ .

(b) Pour  $p$  réel supérieur ou égal à 1, l'application

$$\mathbf{E} \rightarrow \mathbf{R} : f \mapsto \left( \int_a^b |f(x)|^p dx \right)^{1/p}$$

est une norme sur  $\mathbf{E}$ .

Là encore, seuls les cas  $p = 1$  et  $p = 2$  sont au programme et le cas général fait l'objet d'un des exercices.

Dans le premier cas on appelle cette norme, *norme de la convergence en moyenne* dans le second *norme de la convergence en moyenne quadratique*.

## 3. NORMES SUR UN PRODUIT D'ESPACES VECTORIELS

Désignons par  $n$  un entier supérieur ou égal à 1. Soient  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$  des espaces vectoriels munis respectivement des normes  $|\cdot|_1, |\cdot|_2, \dots, |\cdot|_n$ . Par  $\mathbf{E}$  on désigne l'espace vectoriel produit  $\mathbf{E}_1 \times \mathbf{E}_2 \times \dots \times \mathbf{E}_n$ . Les applications suivantes sont des normes sur  $\mathbf{E}$  :

(a)  $n'_\infty : \mathbf{E} \rightarrow \mathbf{R}; (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \mapsto \sup\{|\vec{x}_i|_i, i \in \{1, 2, \dots, n\}\}.$

(b)  $n'_1 : \mathbf{E} \rightarrow \mathbf{R}; (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \mapsto \sum_{i=1}^n |\vec{x}_i|_i.$

(c)  $n'_2 : \mathbf{E} \rightarrow \mathbf{R}; (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \mapsto \left( \sum_{i=1}^n |\vec{x}_i|_i^2 \right)^{1/2}.$

En principe on prend par défaut la première norme :  $n'_\infty$ .

**Preuve** —

- $n_\infty$  EST UNE NORME SUR  $\mathbf{K}^n$

- $n_1$  EST UNE NORME SUR  $\mathbf{K}^n$

- $n_2$  EST UNE NORME SUR  $\mathbf{K}^n$

Dans le cas où  $\mathbf{K} = \mathbf{R}$ , on a vu en MPSI que  $n_2$  est la norme associée au produit scalaire canonique sur  $\mathbf{R}^n$ .

Bien sûr seule l'inégalité triangulaire présente quelque intérêt...

- $N_\infty$  EST UNE NORME SUR  $\mathbf{E}$

- $N_1$  EST UNE NORME SUR  $\mathbf{E}$

Contentons nous de montrer le premier point, la preuve de l'homogénéité et de l'inégalité triangulaire est laissée en exercice.

- $N_2$  EST UNE NORME SUR  $\mathbf{E}$

Le résultat a été prouvé en MPSI dans le cas où  $\mathbf{K} = \mathbf{R}$ , en effet on alors que  $N_2$  est la norme associée au produit scalaire

$$\mathbf{E} \times \mathbf{E} \rightarrow \mathbf{R}; (f, g) \mapsto \dots$$



**Définition 1.1.8.** — On appelle *espace vectoriel normé*, en abrégé *e.v.n.* tout  $\mathbf{K}$ -espace vectoriel muni d'une norme, c'est-à-dire de façon plus précise tout couple  $(\mathbf{E}, \|\cdot\|)$ , où  $\mathbf{E}$  désigne un espace vectoriel et  $\|\cdot\|$  une norme sur  $\mathbf{E}$ .

Il arrive que, par abus de langage, on écrive, lorsque la norme dont on équipe  $\mathbf{E}$  ne fait pas de doute, « l'e.v.n  $\mathbf{E}$  », là où l'on aurait dû écrire « l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$  ». Cet abus n'est pas pire que celui qui nous fait écrire : « le corps  $\mathbf{R}$  », au lieu de « le corps  $(\mathbf{R}, +, \times)$  ».

Lorsque l'on parle d'une « partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  » on entend que  $A$  est une partie de l'ensemble  $\mathbf{E}$ , en sous-entendant cette précision supplémentaire, que, sauf mention contraire, les différentes propriétés topologiques de  $A$  seront relatives à la norme  $\|\cdot\|$ . On appelle souvent « points d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  », les éléments de  $\mathbf{E}$ .

Les définitions suivantes sont alors indispensables à la suite du cours.

**Définition 1.1.9.** — Soient  $(\mathbf{E}, \|\cdot\|)$  un espace vectoriel normé,  $\vec{a}$  un élément de  $\mathbf{E}$  et  $r$  un réel strictement positif. On appelle *boule ouverte* (respectivement *fermée*) de centre  $\vec{a}$  de rayon  $r$ , l'ensemble des éléments  $\vec{x}$  de  $\mathbf{E}$  tels que

$$\|\vec{x} - \vec{a}\| < r \text{ (respectivement } \|\vec{x} - \vec{a}\| \leq r \text{ )}.$$

On la notera dans ce cours  $B_o(\vec{a}, r)$  (respectivement  $B_f(\vec{a}, r)$ ).

Dans le cas où  $\mathbf{E}$  est de dimension 2, on parle plutôt de *disques* que de boules. Dans l'e.v.n.  $(\mathbf{R}, |\cdot|)$  les boules ouvertes (respectivement fermées) de centre  $a$  et de rayon  $r$  sont les intervalles ouverts (respectivement fermés) centrés en  $a$  et de largeur  $2r$ .

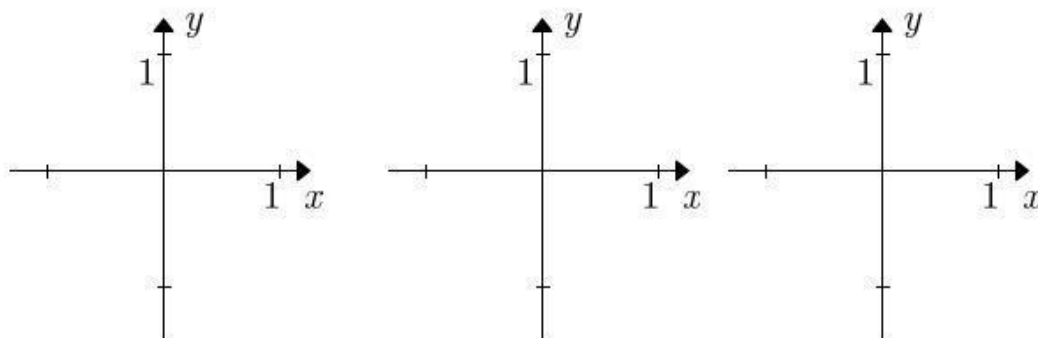


FIGURE III.1 — Ces trois figures représentent dans  $\mathbf{R}^2$  les boules de centre  $(0,0)$  et de rayon 1, pour, dans l'ordre, les normes  $n_2$ ,  $n_1$  et  $n_\infty$ .

**Exercice** — Déterminer les boules de  $\mathbf{R}^3$  de centre  $(0,0)$  et de rayon 1 (boules unité), pour les normes  $n_2$ ,  $n_1$  et  $n_\infty$ .

**Définition 1.1.10.** — Une partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  est dite *bornée* si par définition elle est contenue dans une boule fermée.

On peut dans la définition précédente fixer le centre de la boule :

**Proposition 1.1.11.** — Une partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  est bornée, si et seulement si, pour tout élément  $\vec{a}$  de  $\mathbf{E}$ , elle est contenue dans une boule de centre  $\vec{a}$ .

*Preuve de la proposition 1.1.11.* —

- Si pour tout élément  $\vec{a}$  de  $\mathbf{E}$ , elle est contenue dans une boule fermée de centre  $\vec{a}$ , alors elle est, par définition, bornée.
- Réciproquement, supposons  $A$  bornée et prenons  $\vec{a}$  un élément de  $\mathbf{E}$ . Montrons que  $A$  est incluse dans une boule fermée de centre  $\vec{a}$ . Par définition, il existe un élément  $\vec{a}_0$  de  $\mathbf{E}$  et un réel strictement positif  $R$  tels que :  $A \subset B_f(\vec{a}_0, R)$ . La figure VI.2 laisse accroire que :

$$A \subset B_f(\vec{a}, R + \|\vec{a} - \vec{a}_0\|). \quad (\text{III.3})$$

Vérifions le !

Soit  $\vec{x}$  élément de  $A$ . L'inégalité triangulaire donne

$$\|\vec{x} - \vec{a}\| \leq \|\vec{x} - \vec{a}_0\| + \|\vec{a}_0 - \vec{a}\|.$$

Comme  $A \subset B_f(\vec{a}_0, R)$ ,  $\|\vec{x} - \vec{a}_0\| \leq R$ , et donc  $\|\vec{x} - \vec{a}\| \leq R + \|\vec{a}_0 - \vec{a}\|$ , autrement dit,

$$\vec{x} \in B_f(\vec{a}, R + \|\vec{a} - \vec{a}_0\|).$$

L'élément  $\vec{x}$  étant quelconque, on a montré (64). D'où,  $\vec{a}$  étant quelconque, le résultat. □

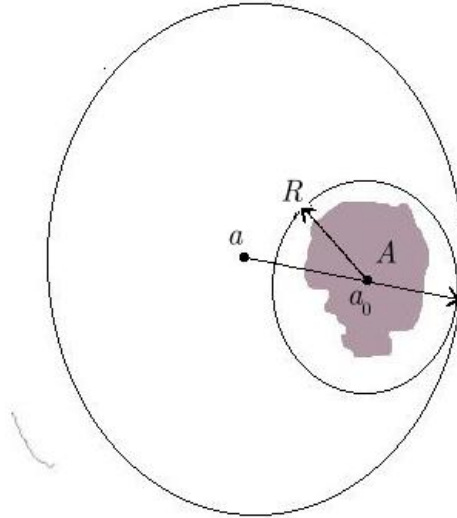


FIGURE III.2 –

Remarquons que le fait d'être bornée pour une partie d'un espace vectoriel  $\mathbf{E}$  dépend de la norme choisie. Prenons par exemple pour  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([0, 1], \mathbf{R})$  et pour  $A$  l'ensemble  $\{f_n, n \in \mathbf{N}\}$ , où pour tout entier naturel  $n$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; t \mapsto (n+1)t^n.$$

$A$  est bornée dans  $(\mathcal{C}^0([0, 1], \mathbf{R}), N_1)$ , en effet pour tout entier naturel  $n$ ,  $N_1(f_n) = 1$ ; en revanche  $N_\infty(f_n) = n+1$ , et donc  $A$  ne saurait être bornée dans  $(\mathcal{C}^0([0, 1], \mathbf{R}), N_\infty)$ .

**Proposition-définition 1.1.12.** —

- Soit  $X$  un ensemble non vide et  $(\mathbf{E}, \|\cdot\|)$  un e.v.n. Une application  $\vec{f}$  de  $X$  dans  $\mathbf{E}$  est dite bornée pour la norme  $\|\cdot\|$ , si par définition, l'ensemble  $\vec{f}(X)$  est une partie bornée de  $(\mathbf{E}, \|\cdot\|)$ .
- L'ensemble des applications de  $X$  dans  $\mathbf{E}$ , bornées pour  $\|\cdot\|$  est un sous-espace vectoriel de  $\mathcal{F}(X, \mathbf{E})$  espace vectoriel des applications de  $X$  dans  $\mathbf{E}$ . Cet ensemble sera noté  $\mathcal{B}(X, \mathbf{E})$ .
- Pour tout élément  $f$  de  $\mathcal{B}(X, \mathbf{E})$ , l'ensemble  $\{\|\vec{f}(x)\|, x \in X\}$  admet une borne supérieure (élément de  $\mathbf{R}$ ). On la note  $\sup_{x \in X} \|\vec{f}(x)\|$ .
- L'application

$$N_\infty : \mathcal{B}(X, \mathbf{E}) \rightarrow \mathbf{R}; \vec{f} \mapsto \sup_{x \in X} \|\vec{f}(x)\|$$

est une norme sur  $\mathcal{B}(X, \mathbf{E})$ . On la note encore  $\|\cdot\|_\infty$ .

*Preuve de la proposition 1.1.12.* — Que l'ensemble des applications de  $X$  dans  $\mathbf{E}$ , bornées soit un sous-espace-vectoriel de  $\mathcal{F}(X, \mathbf{E})$  est un résultat facile laissé en exercice; l'existence de la borne supérieure de  $\{\|\vec{f}(x)\|, x \in X\}$ , pour  $\vec{f} \in \mathcal{B}(X, \mathbf{E})$ , résulte de ce que cet ensemble est non vide majoré; détaillons le dernier point.

Le premier point et le troisième (inégalité triangulaire) pour  $N_\infty$  définie sur  $\mathcal{B}(X, \mathbf{E})$  se prouve comme dans le cas de  $N_\infty$  définie sur  $\mathcal{C}^0([a, b], \mathbf{R})$ , (cf. 1.1.7.2.(a)), étudions l'homogénéité.

Soient  $\vec{f} \in \mathcal{B}(X, \mathbf{E})$  et  $\lambda \in \mathbf{K}$ .

Notons qu'en vertu de 1.1.11. une application  $\vec{f}$  de  $X$  dans  $\mathbf{E}$  est bornée si et seulement si  $\vec{f}(X)$  est inclus dans un boule fermée de centre  $\vec{0}_{\mathbf{E}}$ . Donc si et seulement s'il existe un réel strictement positif  $R$ , tel que pour tout élément  $x$  de  $X$ ,

$$\|\vec{f}(x)\| \leq R.$$

**Définition 1.1.13.** — Soient  $(\mathbf{E}, \|\cdot\|)$  un e.v.n.,  $A$  une partie non vide de  $(\mathbf{E}, \|\cdot\|)$  et  $\vec{a}$  un point de  $(\mathbf{E}, \|\cdot\|)$ . On appelle distance de  $\vec{a}$  à  $A$ , le réel  $\inf\{\|\vec{x} - \vec{a}\|, \vec{x} \in A\}$ <sup>1</sup>. On le note  $d(\vec{a}, A)$ .

**Remarque**

- Il ne faut pas croire que la distance d'un point  $\vec{a}$  à une partie  $A$  soit toujours « atteinte », c'est-à-dire qu'il existe un élément  $\vec{x}_0$  de  $A$  tel que  $\|\vec{x}_0 - \vec{a}\| = d(\vec{a}, A)$ . Prenons par exemple  $\mathbf{R}$  muni de la valeur absolue, dans cet e.v.n.  $d(0, ]1, 2])$  vaut clairement 1, cependant, pour tout élément  $x$  de  $]0, 1]$ ,  $|x - 0| > 1$ .
- Il a été prouvé en sup que pour un  $\mathbf{R}$ -espace vectoriel muni d'une norme euclidienne, c'est-à-dire qui provient d'un produit scalaire, (comme  $n_2$  ou  $N_2$ ) la distance d'un vecteur  $\vec{a}$  à un sous-espace vectoriel de dimension finie  $\mathbf{F}$  est atteinte en un unique point de  $\mathbf{F}$  (le projeté orthogonal de  $\vec{a}$  sur  $\mathbf{F}$ ). Il en est tout autrement des normes quelconques (cf. figure VI.3). Cette propriété qui distingue les normes euclidiennes, leur font jouer un

---

1. L'existence de cette borne inférieure est assurée par la non vacuité de  $\{\|\vec{x} - \vec{a}\|, \vec{x} \in A\}$  et le fait que cet ensemble est minoré par 0.

rôle essentiel en analyse dans de nombreux théorèmes d'existence et d'unicité. Il existe du reste, dans le cadre des espaces vectoriels sur  $\mathbf{C}$ , la généralisation de cette propriété, mais elle c'est envolée du programme.

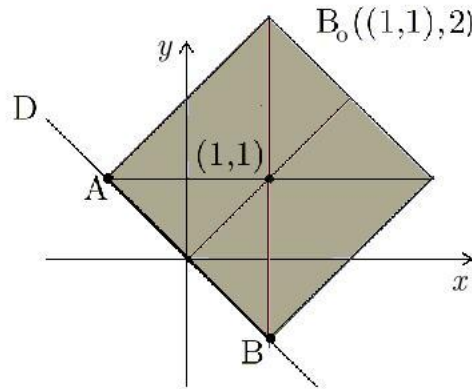


FIGURE III.3 – Dans  $\mathbf{R}^2$  muni de la norme  $n_1$ ,  $(1,1)$  est à une distance 2 de la droite d'équation  $x + y = 0$ ; cette distance est atteinte en tout point du segment  $[A, B]$ , et seulement en ces points.

**Proposition 1.1.14.** — Soient  $(\mathbf{E}, \|\cdot\|)$  un e.v.n. et  $A$  une partie non vide de  $(\mathbf{E}, \|\cdot\|)$ . Pour tout  $\vec{x}$  et tout  $\vec{y}$  éléments de  $\mathbf{E}$ ,

$$|d(x, A) - d(y, A)| \leq \|\vec{x} - \vec{y}\|.$$

On verra par la suite que cette propriété se traduit en disant que l'application de  $(\mathbf{E}, \|\cdot\|)$  dans  $(\mathbf{R}, |\cdot|)$ , qui à un élément  $\vec{x}$  associe  $d(\vec{x}, A)$  est 1-lipschitzienne.

*Preuve de la proposition 1.1.14.* —

## 1.2 Ouverts fermés et autres facéties

Nous allons maintenant formaliser la notion de partie qui contient son « bord », dite *fermée* ou qui ne contient aucune partie de son « bord » dite *ouverte*. Intuitivement une partie  $A$  ne contient aucune partie de son « bord », si aucun de ses points n'est sur le « bord », c'est-à-dire si autour de chaque point de  $A$  il n'y a que des points de  $A$ . Nous adopterons donc la définition suivante :

**Définition 1.2.1.** — Une partie  $A$  d'un e.v.n. est dite *ouverte*, si par définition, pour tout élément  $\vec{x}$  de  $A$ , il existe un réel strictement positif  $r$ , tel que la boule de centre  $\vec{x}$ , de rayon  $r$ ,  $B_o(\vec{x}, r)$ , soit incluse dans  $A$ .

**Remarque** —

1. On parle souvent d'*ouvert* pour désigner une partie ouverte.
2. Une boule ouverte est ... ouverte! (cf. figure (III.4)).
3. Les intervalles ouverts au sens ordinaire, sont ouverts au sens de la définition 1.2.1 dans  $(\mathbf{R}, |\cdot|)$ , s'ils sont bornés se sont du reste des boules ouvertes. Enfin que ce sont les seuls intervalles à être ouverts au sens de 1.2.1.

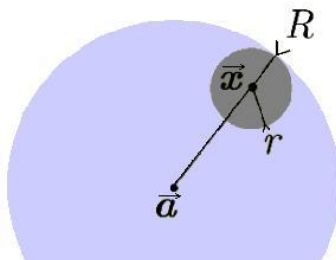


FIGURE III.4 – si  $\vec{x}$  est élément de  $B_o(\vec{a}, R)$ , alors  $B_o(\vec{x}, r)$  est incluse dans  $B_o(\vec{a}, R)$ , pour, par exemple,  $r = R - \|\vec{x} - \vec{a}\|$ .

Les parties ouvertes jouissent des propriétés suivantes<sup>1</sup> :

**Proposition 1.2.2.** — *Soit  $(\mathbf{E}, \|\cdot\|)$  un e.v.n.*

1. *L'ensemble vide et l'ensemble  $\mathbf{E}$  sont des parties ouvertes.*
2. *Toute intersection finie de parties ouvertes de  $\mathbf{E}$  est ouverte.*
3. *Toute réunion (finie ou non) de parties ouvertes de  $\mathbf{E}$  est ouverte.*

*Preuve de la proposition 1.2.2.*—

- Le point 1. est un exercice insignifiant laissé au lecteur.
- Montrons 2.  
Soient  $U_1, U_2, \dots, U_n$  des parties ouvertes de  $(\mathbf{E}, \|\cdot\|)$ . Soit  $\vec{x}$  un élément éventuel de  $U_1 \cap U_2 \cap \dots \cap U_n$ . Pour  $i = 1, 2, \dots, n$ ,  $x \in U_i$ ...

L'élément  $\vec{x}$  étant quelconque il en résulte que  $U_1 \cap U_2 \cap \dots \cap U_n$  est ouvert.

- Montrons 3.  
Soit  $(U_i)_{i \in I}$  une famille de parties ouvertes de  $(\mathbf{E}, \|\cdot\|)$ , indexée par un ensemble  $I$  fini ou non. Soit  $\vec{x}$  un

1. L'ensemble des ouverts s'appelle topologie de  $\mathbf{E}$ . Plus généralement on appelle topologie d'un ensemble  $E$ , indépendamment de toute norme ( $E$  n'est pas nécessairement un espace vectoriel) ou distance, tout ensemble  $\mathcal{T}$  de parties de  $E$ , appelées ouverts, qui satisfont les trois propriétés 1.2.2.1–3. Cette étude générale est hors programme.

éventuelle élément de  $\bigcup_{i \in I} U_i$ .

L'élément  $\vec{x}$  étant quelconque il en résulte que  $\bigcup_{i \in I} U_i$  est ouvert.

□

**Remarque** — La propriété 2. ne se généralise pas à une intersection quelconque d'ouverts, en effet, par exemple pour tout naturel non nul  $n$ ,  $]-\frac{1}{n}, \frac{1}{n}[$  est un ouvert de  $(\mathbf{R}, |\cdot|)$ , tandis que  $\bigcap_{n \in \mathbf{N}^*} ]-\frac{1}{n}, \frac{1}{n}[$  ne l'est évidemment pas puisque cet ensemble est le singleton  $\{0\}$  !

**Exercice 1.2.3.** — Montrer qu'une partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , non vide<sup>1</sup>, est ouverte, si et seulement si elle est la réunion de boules ouvertes.

**Solution** —

- Si  $A$  est réunion de boules ouvertes elle est, puisque les boules ouvertes sont ouvertes, réunion d'ouverts. Mais alors, d'après 1.2.2.3. elle est ouverte.
- Réciproquement, Si  $A$  est ouverte pour tout élément  $\vec{x}$  de  $A$  on peut choisir une boule ouverte de centre  $\vec{x}$  incluse dans  $A$ , nous la noterons  $B_{\vec{x}}$ .

On a bien que  $A$  est la réunion de boules ouvertes.

**Remarque 1.2.4.** — Soit  $(\mathbf{E}, \|\cdot\|)$  un e.v.n. Quels que soient  $\vec{a}$  et  $\vec{b}$  éléments distincts de  $\mathbf{E}$ , il existe deux ouverts disjoints l'un contenant  $\vec{a}$ , l'autre  $\vec{b}$ .

On peut, par exemple, prendre  $B_0\left(\vec{a}, \frac{\|\vec{b}-\vec{a}\|}{2}\right)$  et  $B_0\left(\vec{b}, \frac{\|\vec{b}-\vec{a}\|}{2}\right)$ .

□

Voyons maintenant les parties qui contiennent leur « bord » dites *fermées*. Une partie contient son « bord » si son complémentaire ne contient aucune partie du sien, d'où la définition :

**Définition 1.2.5.** — Une partie  $A$  d'un espace vectoriel normé  $(\mathbf{E}, \|\cdot\|)$  est dite *fermée*, si, par définition, son complémentaire est ouvert.

1. Si l'on convient qu'une réunion « vide » de parties est l'ensemble vide, cette hypothèse est inutile.

On déduit alors, par passage au complémentaire de 1.2.2., les propriétés suivantes :

**Proposition 1.2.6.** — *Soit un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .*

1. *L'ensemble  $\mathbf{E}$  et l'ensemble vide sont des parties fermées<sup>1</sup>.*
2. *Toute réunions finie de parties fermées de  $\mathbf{E}$  est fermée.*
3. *Toute intersection (finie ou non) de parties fermées de  $\mathbf{E}$  est fermée.*

**Exercice** — Montrer que 2. ne se généralise pas à une réunion quelconque.

**Remarque** —

1. La figure III.5, montre que les boules fermées sont ... fermées !
2. Les singleton sont fermés.
3. En particulier les intervalles fermés sont fermés dans  $(\mathbf{R}, |\cdot|)$  au sens de la définition 1.2.5., puisque ce sont des boules fermées. Ce sont du reste les seuls intervalles bornés à être fermés au sens de 1.2.5.
4. les intervalles semi-ouverts de la forme  $] - \infty, a]$  et  $[a, +\infty[$ , avec  $a \in \mathbf{R}$ , sont fermés au sens de 1.2.5.

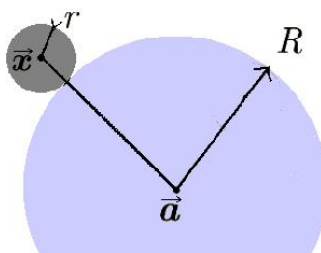


FIGURE III.5 — si  $\vec{x}$  est élément du complémentaire de  $B_f(\vec{a}, R)$ , alors  $B_o(\vec{x}, r)$  est incluse dans le complémentaire  $B_f(\vec{a}, R)$ , pour, par exemple,  $r = \|\vec{x} - \vec{a}\| - R$ .

**Définition 1.2.7.** — VOISINAGE —

On appelle voisinage d'un élément  $\vec{a}$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , toute partie  $V$  de  $\mathbf{E}$  telle qu'il existe un ouvert  $U$  vérifiant :

$$\{\vec{a}\} \subset U \subset V.$$

Dans la suite nous noterons  $\mathcal{V}_{\vec{a}}$  l'ensemble des voisinages de  $\vec{a}$ .

On déduit sans mal de 1.2.2 et de 1.2.4 la proposition suivante :

**Proposition 1.2.8.** — *Soit un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  et  $\vec{a}$  un point de  $\mathbf{E}$ .*

1. *Toute partie de  $\mathbf{E}$  qui contient un voisinage de  $\vec{a}$  est un voisinage de  $\vec{a}$ .*
2. *Toute intersection finie de voisinages de  $\vec{a}$  est un voisinage de  $\vec{a}$ .*
3. *Quels que soient  $\vec{a}_1$  et  $\vec{a}_2$  éléments distincts de  $\mathbf{E}$ , il existe un voisinage de  $\vec{a}_1$  et un voisinage de  $\vec{a}_2$  disjoints.*

Montrer que 2. ne se généralise pas à un intersection quelconque.

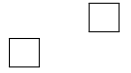
**Proposition 1.2.9.** — *Une partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  est ouverte si et seulement si elle est un voisinage de chacun de ses points.*

*Preuve de la proposition 1.2.9.* —

- Supposons  $A$  ouverte. Soit  $\vec{x}$  un éventuel point de  $A$ , comme :  $\{\vec{x}\} \subset A \subset A$ ,  $A$  est par définition voisinage de  $\vec{x}$ .

1. On observe donc que *fermé* n'est pas le contraire de *ouvert*, puisque l'ensemble vide et  $\mathbf{E}$  sont à la fois ouverts et fermés. Par ailleurs, certaines parties sont ni ouvertes ni fermées, comme  $[0, 1[$  dans  $\mathbf{R}$  muni de la valeur absolue.

- Réciproquement, supposons  $A$  voisinage de chacun de ses points.



On retiendra de cette preuve que tout voisinage  $V$  d'un point  $\vec{x}$  contient une boule ouverte de centre  $\vec{x}$  (cf. figure III.6).

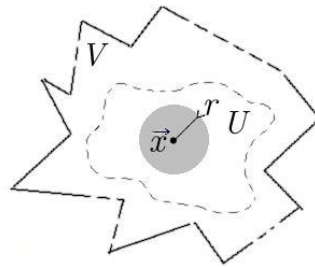


FIGURE III.6 – Un voisinage  $V$  d'un point  $\vec{x}$  contient une boule ouverte de centre  $\vec{x}$ .

Définissons les points d'une partie quelconque qui ne sont pas sur son « bord », dits intérieurs :

**Définition 1.2.10.** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . Un élément  $\vec{a}$  de  $A$  est dit intérieur à  $A$  si, par définition, il existe une boule ouverte de centre  $\vec{a}$  incluse dans  $A$ . L'ensemble des points intérieurs à  $A$  s'appelle l'intérieur de  $A$ . L'intérieur de  $A$  se note  $\overset{\circ}{A}$ .

**Remarque** — Dans cette définition, on peut remplacer *boule ouverte* par *boule fermée*, puisque toute boule ouverte contient une boule fermée de même centre et vice versa.

**Proposition 1.2.11.** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . Un élément  $\vec{a}$  de  $A$  est intérieur à  $A$  si et seulement si  $A$  est un voisinage de  $\vec{a}$ .

*Preuve de la proposition 1.2.11.* —

- Si  $\vec{a}$  est intérieur à  $A$ , il existe une boule ouverte  $B$  de centre  $\vec{a}$  incluse dans  $A$ . on a  $\{\vec{a}\} \subset B \subset A$  et donc par définition,  $A$  est voisinage de  $\vec{a}$ .
- Réciproquement si  $A$  est un voisinage de  $\vec{a}$ , il existe, comme on l'a vu, une boule ouverte de centre  $\vec{a}$  incluse dans  $A$ ,  $\vec{a}$  est intérieur à  $A$ .

**Proposition 1.2.12.** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . L'intérieur de  $A$ ,  $\overset{\circ}{A}$ , est la réunion de tous les ouverts inclus dans  $A$ .

$\overset{\circ}{A}$  est donc un ouvert (cf. 1.2.2.2.). C'est même le plus grand ouvert au sens de l'inclusion, inclus dans  $A$ .

*Preuve de la proposition 1.1.12* —

- Montrons que la réunion de tous les ouverts inclus dans  $A$  est incluse dans  $\overset{\circ}{A}$ . Soit  $U$  un ouvert inclus dans  $A$ . Soit  $\vec{x}$  un point quelconque de  $U$ . L'ensemble  $U$  étant ouvert, il existe une boule ouverte  $B$  de centre  $\vec{x}$  incluse dans  $U$ , donc dans  $A$ , donc  $\vec{x}$  est intérieur à  $A$ . Voila prouvé :  $U \subset \overset{\circ}{A}$ .  $U$  étant un ouvert quelconque inclus dans  $A$ , la réunion de tous les ouverts inclus dans  $A$  est incluse dans  $\overset{\circ}{A}$ .
- Montrons l'inclusion inverse. Soit  $\vec{x}$  un point quelconque de  $\overset{\circ}{A}$ . Par définition, il existe une boule ouverte  $B$ , de centre  $\vec{x}$ , incluse dans  $A$ , or  $B$  est un ouvert inclus dans  $A$  qui, bien évidemment contient  $\vec{x}$ , donc  $\vec{x}$  est élément de la réunion de tous les ouverts inclus dans  $A$ . Donc finalement  $\overset{\circ}{A}$  est incluse dans la réunion de tous les ouverts inclus dans  $A$ .



D'où, par double inclusion, l'égalité demandée.



**Exercice 1.2.13.** — Déterminer l'intérieur des parties suivantes :

1.  $\mathbf{Q}$  dans  $(\mathbf{R}, |\cdot|)$ .
2.  $\mathbf{R} - \mathbf{Q}$  dans  $(\mathbf{R}, |\cdot|)$ .
3.  $[0, 1]$  dans  $(\mathbf{R}, |\cdot|)$ .
4. Une droite affine  $D$  dans  $(\mathbf{R}^2, n_2)$ .
5. La boule fermée  $B_f(\tilde{a}, r)$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .

*Solution de l'exercice 1.2.13.*—

- 1.
- 2.
- 3.
- 4.
- 5.

Définissons maintenant, pour une partie  $A$ , les points qui sont dans  $A$  ou sur son « bord », ou comme on dit, *adhérents* à la partie.

**Définition 1.2.14.** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . Un élément  $\tilde{a}$  de  $\mathbf{E}$  est dit *adhérent* à  $A$  si, par définition, toute boule ouverte de centre  $\tilde{a}$  rencontre  $A$ . L'ensemble des points adhérents à  $A$  s'appelle l'*adhérence* de  $A$ . L'adhérence de  $A$  se note  $\bar{A}$ .

**Remarque** — Dans cette définition on peut remplacer *boule ouverte* par *boule fermée*, puisque toute boule ouverte contient une boule fermée de même centre et *vice versa*.

Comme toute boule ouverte de centre  $\tilde{a}$  est un voisinage de  $\tilde{a}$  et que tout voisinage de  $\tilde{a}$  contient une boule ouverte de centre  $\tilde{a}$  il vient :

**Proposition 1.2.15.** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . Un élément  $\tilde{a}$  de  $E$  est adhérent à  $A$  si et seulement si tout voisinage de  $\tilde{a}$  rencontre  $A$ .

**Proposition 1.2.16** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . L'adhérence de  $A$ ,  $\bar{A}$ , est l'intersection de tous les fermés contenant  $A$ .

$\bar{A}$  est donc un fermé (1.2.6.3.). C'est même le plus petit fermé, au sens de l'inclusion, contenant  $A$ .

*Preuve de la proposition 1.2.16. —*

**Exercice 1.2.17.** — Déterminer l'adhérence des parties suivantes :

1.  $\mathbf{Q}$  dans  $(\mathbf{R}, |\cdot|)$ .
2.  $\mathbf{R} - \mathbf{Q}$  dans  $(\mathbf{R}, |\cdot|)$ .
3.  $[0, 1]$  dans  $(\mathbf{R}, |\cdot|)$ .
4. Une droite affine  $D$  dans  $(\mathbf{R}^2, n_2)$ .
5. La boule ouverte  $B_o(\tilde{a}, r)$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .

*Solution de l'exercice 1.2.17. —*

- 1.
- 2.
- 3.
- 4.
- 5.

La définition suivante est cruciale en analyse :

**Définition 1.2.18.** — Une partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  est dite dense si, par définition  $\bar{A} = \mathbf{E}$ .

**Exemple** — Dans  $(\mathbf{R}, |\cdot|)$ ,  $\mathbf{Q}$  et  $\mathbf{R} - \mathbf{Q}$  sont denses.

Il résulte de 1.2.12. et de 1.2.16., la proposition suivante :

**Proposition 1.2.19.** — Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .

- $A$  est ouverte si et seulement si  $\overset{\circ}{A} = A$ .
- $A$  est fermée si et seulement si  $\bar{A} = A$ .

On peut enfin définir le « bord » d'une partie que nous appellerons *frontière* :

**Définition 1.2.20.** — FRONTIÈRE —

On appelle frontière d'une partie  $A$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , l'ensemble  $\bar{A} - \overset{\circ}{A}$ . On le note  $\text{Fr}(A)$ .

D'après 1.2.20.,

$$\text{Fr}(A) = \bar{A} \cap (\mathbf{E} - \overset{\circ}{A}) = \bar{A} \cap \overline{\mathbf{E} - A}.$$

La frontière d'une partie est donc fermée.

**Exercice 1.2.21.** — Déterminer la frontière des parties suivantes :

1.  $\mathbf{Q}$  dans  $(\mathbf{R}, |\cdot|)$ .
2.  $\mathbf{R} - \mathbf{Q}$  dans  $(\mathbf{R}, |\cdot|)$ .
3.  $[0, 1]$  dans  $(\mathbf{R}, |\cdot|)$ .
4. Une droite affine  $D$  dans  $(\mathbf{R}^2, n_2)$ .
5. La boule ouverte  $B_o(\vec{a}, r)$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .
6. La boule fermée  $B_f(\vec{a}, r)$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .

*Solution de l'exercice 1.2.22.* —

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.

### 1.3 Topologie relative à une partie d'un e.v.n.

Dans ce paragraphe  $(\mathbf{E}, \|\cdot\|)$  désigne un espace vectoriel normé sur  $\mathbf{K}$ ,  $d$  la distance associée à la norme  $\|\cdot\|$  et, enfin,  $A$  une partie quelconque de  $\mathbf{E}$ .

Dans un premier temps le lecteur pressé, ou soucieux de ménager ses forces pour la suite du programme, pourrait se contenter d'adopter comme définitions :

- Une partie de  $A$  est dite ouverte (respectivement fermée) relativement à  $A$  si, par définition c'est l'intersection d'un ouvert (respectivement d'un fermé) et de  $A$ .
- On appelle voisinage d'un point  $\vec{a}$  de  $A$  relativement à  $A$ , tout intersection d'un voisinage de  $\vec{a}$  et de  $A$ .

Une partie ouverte (respectivement fermée) relativement à  $A$ , s'appelle plus simplement un ouvert (respectivement un fermé) relatif de  $A$ . Il convient de ne pas confondre avec un ouvert (respectivement un fermé) inclus dans  $A$ .

Il ne resterait plus à voir que les fermés de  $A$  sont les complémentaires dans  $A$  des ouverts de  $A$ . Afin de satisfaire les esprits plus curieux, ou plus courageux et de leur offrir une définition plus profonde donnons une autre présentation.

La restriction de  $d$  à  $A$  est une distance sur  $A$  que l'on notera encore  $d$  par un abus consacré.

Commençons par une définition.

**Définition 1.3.1.** — Soient  $\vec{a}$  un point de  $A$  et  $r$  un réel strictement positif. On appelle boule ouverte de  $A$  de centre  $\vec{a}$ , de rayon  $r$ , l'ensemble des points de  $A$  dont la distance à  $\vec{a}$  est strictement inférieure à  $r$ . On la notera dans ce cours  $B_{A,o}(\vec{a}, r)$ .

Notons que tout simplement :  $B_{A,o}(\vec{a}, r) = B_o(\vec{a}, r) \cap A$  !

On définit alors un ouvert relativement à  $A$  comme on a défini un ouvert en 1.2.1.

**Définition 1.3.2.** — Une partie  $U_A$  de  $A$  est dite ouverte relativement à  $A$  (ou ouvert relatif de  $A$ ) si, par définition, pour tout élément  $\vec{x}$  de  $U_A$ , il existe un réel strictement positif  $r$ , tel que  $B_{A,o}(\vec{x}, r)$  soit incluse dans  $U_A$ .

Dans le cas où  $A$  est égal à  $\mathbf{E}$  on retrouve la définition 1.1.2.

Remarquons que si  $U_A$  est un ouvert relatif de  $A$ , alors, pour tout élément  $\vec{x}$  de  $A$ , il existe un réel strictement positif, noté  $r_{\vec{x}}$ , tel que  $B_{A,o}(\vec{x}, r_{\vec{x}})$  soit incluse dans  $U_A$ . On montre alors sans mal, en raisonnant comme dans l'exercice I.2.3. que  $U_A = \bigcup_{\vec{x} \in U_A} B_{A,o}(\vec{x}, r_{\vec{x}})$ . Soit encore :

$$U_A = \bigcup_{\vec{x} \in U_A} (B_o(\vec{x}, r_{\vec{x}}) \cap A) = \left( \bigcup_{\vec{x} \in U_A} B_o(\vec{x}, r_{\vec{x}}) \right) \cap A.$$

Or  $\bigcup_{\vec{x} \in U_A} B_o(\vec{x}, r_{\vec{x}})$  est, d'après I.2.2.3., un ouvert (de  $\mathbf{E}$ ), donc  $U_A$  est l'intersection de  $A$  et d'un ouvert (de  $\mathbf{E}$ ). Réciproquement Si  $U$  est un ouvert de  $\mathbf{E}$ , alors  $U \cap A$  est un ouvert relatif de  $A$ . En effet,  $U$  étant ouvert pour tout élément  $\vec{x}$  de  $U \cap A$  il existe un réel strictement positif  $r$  tel que  $B_o(\vec{x}, r)$  soit incluse dans  $U$ , donc tel que  $B_{A,o}(\vec{x}, r)$  soit incluse dans  $U \cap A$ . Bref, on a montré :

**Proposition 1.3.3.** — *les ouverts relatif de  $A$  sont les « traces sur  $A$  » des ouverts (de  $\mathbf{E}$ ), c'est-à-dire les intersections des ouverts et de  $A$ .*

**Exemples** — Dans  $(\mathbf{R}, |\cdot|)$ , prenons  $A = [0, 3[$ . Alors,  $]1, 2[$ ,  $[0, 2[$ ,  $]0, 2[$ , sont des ouverts de  $A$ ;  $[1, 2[$ ,  $[1, 2]$  n'en sont pas.

Il résulte immédiatement des propositions 1.3.3. et 1.2.2. :

**Proposition 1.3.4.**<sup>1</sup> —

1. *L'ensemble vide et l'ensemble  $A$  sont des ouverts relatifs de  $A$ .*
2. *Toute intersection finie d'ouverts de  $A$  est un ouvert relatif de  $A$ .*
3. *Toute réunion (finie ou non) d'ouverts de  $A$  est un ouvert relatif de  $A$ .*

Définissons les fermés relativement à  $A$ .

**Définition 1.3.5.** — *Une partie  $F_A$  de  $A$  est dite fermée relativement à  $A$  (ou fermé relatif de  $A$ ) si, par définition, son complémentaire dans  $A$  est un ouvert relatif de  $A$ , c'est-à-dire si  $A \setminus F_A$  est un ouvert relatif de  $A$ .*

Dans le cas où  $A$  est  $\mathbf{E}$ , on retrouve la définition 1.2.5.

C'est alors un très bel exercice de théorie des ensembles que de montrer :

**Proposition 1.3.6.** — *les fermés relatifs de  $A$  sont les « traces sur  $A$  » des fermés (de  $\mathbf{E}$ ).*

**Exemples** — Dans  $(\mathbf{R}, |\cdot|)$  prenons  $A = [0, 3[$ . Alors,  $[1, 2]$ ,  $[1, 3[$  sont des fermés de  $A$ ;  $]0, 1]$ ,  $]2, 3[$  n'en sont pas.

**Proposition 1.3.7.** —

1. *Les ouverts relatifs de  $A$  coïncident avec les ouverts (de  $\mathbf{E}$ ) inclus dans  $A$  si et seulement si  $A$  est ouvert.*
2. *Les fermés relatifs de  $A$  coïncident avec les fermés (de  $\mathbf{E}$ ) inclus dans  $A$  si et seulement si  $A$  est fermé.*

*Preuve de la proposition 1.3.7* — Montrons par exemple le premier point.

- Si les ouverts relatifs de  $A$  coïncident avec les ouverts inclus dans  $A$  alors, comme  $A$  est un ouvert relatif de  $A$  (cf. 1.3.4.1.),  $A$  est ouvert.
  - Réciproquement supposons  $A$  ouvert.
    - Un ouvert  $U$  inclus dans  $A$  est évidemment un ouvert de  $A$  puisque  $U = U \cap A$  (ceci est du reste vrai même lorsque  $A$  n'est pas ouvert).
    - Un ouvert relatif  $U_A$  de  $A$  est l'intersection d'un ouvert  $U$  et de  $A$ , mais  $A$  étant ouvert  $U_A$  est ouvert comme intersection de deux ouverts.  $U_A$  est donc un ouvert inclus dans  $A$ .
- Les ouverts relatifs de  $A$  sont donc les ouverts (de  $\mathbf{E}$ ) inclus dans  $A$ . □

**Définition 1.3.8.** — *On appelle voisinage relativement à  $A$  d'un élément  $\vec{a}$  de  $A$  toute partie  $V_A$  de  $A$  telle qu'il existe un ouvert relatif  $U_A$  de  $A$ , vérifiant :*

$$\{\vec{a}\} \subset U_A \subset V_A.$$

Dans le cas où  $A$  est  $\mathbf{E}$ , on retrouve la définition 1.2.7.

**Proposition 1.3.9.** — *Les voisinages relativement à  $A$  d'un élément  $\vec{a}$  de  $A$  sont les traces sur  $A$  des voisinages de  $\vec{a}$ .*

*Preuve de la proposition 1.3.9.* —

- Soit  $V$  un voisinage de  $\vec{a}$ . Montrons que  $V \cap A$  est un voisinage de  $\vec{a}$  relativement à  $A$ .  $V$  étant un voisinage de  $\vec{a}$ , on déduit l'existence d'un ouvert  $U$  tel que  $\{\vec{a}\} \subset U \subset V$ . Mais puisque  $\vec{a}$  est élément de  $A$ ,

$$\{\vec{a}\} \subset U \cap A \subset V \cap A.$$

Or d'après 1.3.3.  $U \cap A$  est un ouvert relatif de  $A$ , donc  $V \cap A$  est un voisinage de  $\vec{a}$  relativement à  $A$ .

- Réciproquement, soit  $V_A$  un voisinage de  $\vec{a}$  relativement à  $A$ . Montrons que  $V_A$  est l'intersection d'un voisinage de  $\vec{a}$  et de  $A$ . Par définition il existe un ouvert relatif de  $A$ ,  $U_A$ , vérifiant  $\{\vec{a}\} \subset U_A \subset V_A$ .  $U_A$  étant un ouvert relativement à  $A$ , d'après 1.3.3., il existe un ouvert  $U$  (de  $\mathbf{E}$ ) tel que  $U_A = U \cap A$ . Alors si l'on pose  $V = U \cup U_A$ , on a :

$$\{\vec{a}\} \subset U \subset V.$$

Autrement dit,  $V$  est un voisinage de  $\vec{a}$ . Par ailleurs,

$$V \cap A = (U \cup U_A) \cap A = (U \cap A) \cup (U_A \cap A) = U_A \cup V_A = V_A.$$

$V_A$  est la trace de  $V$ , voisinage de  $\vec{a}$  sur  $A$ .

---

1. D'après la terminologie de la note de la proposition I.2.2, cette proposition dit que l'ensemble des ouverts relativement à  $A$  est une topologie sur  $A$ , topologie qui est définie sans recourir directement à une norme mais à une distance. On appelle cette topologie la topologie induite sur  $A$  par celle de  $\mathbf{E}$ .



**Remarque 1.3.10.** — Si  $\mathbf{F}$  est un sous-espace vectoriel de  $\mathbf{E}$ , la restriction de  $\|\cdot\|$  à  $\mathbf{F}$  est une norme sur  $\mathbf{F}$ , la distance associée à cette norme est la restriction de  $d$  à  $\mathbf{F} \times \mathbf{F}$ . Les ouverts de  $\mathbf{F}$ , les fermés de  $\mathbf{F}$  et les voisinages relativement à  $\mathbf{F}$  définis en 1.3.2., 1.3.5. et 1.3.8. coïncident donc respectivement avec les ouverts, les fermés et les voisinages de l'espace vectoriel normé  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  où  $\|\cdot\|_{\mathbf{F}}$  désigne la restriction de  $\|\cdot\|$  à  $\mathbf{F}$ .

Donnons une définition plus générale de la densité que celle vue en 1.2.18., notion que nous rencontrerons souvent cette année :

**Définition 1.3.11.** — Soient  $D$  une partie de  $A$ . La partie  $D$  est dite dense dans  $A$  si, par définition,  $A \subset \bar{D}$ .

- La définition précédente, en fait, nous dit que le plus petit fermé relativement à  $A$  contenant  $D$  est  $A$ . le plus petit fermé relativement à  $A$  contenant  $D$ , est ce que l'on appelle l'adhérence de  $D$  relativement à  $A$ , ainsi retrouve-t-on une formulation similaire à 1.2.18.

## LIMITES, CONTINUITÉ

### 2.1 Limite d'une suite

Dans ce paragraphe  $(\mathbf{E}, \|\cdot\|)$  désigne un espace vectoriel sur le corps  $\mathbf{K}$  normé ( $\mathbf{K}$  désigne toujours  $\mathbf{R}$  ou  $\mathbf{C}$ ).

Nous allons étudier les suites à valeurs dans un e.v.n. et pour l'essentiel généraliser des résultats de MPSI sur les suites réelles ou complexes.

**Définition 2.1.1.** — On appelle suite à valeurs dans  $\mathbf{E}$  toute famille  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeurs dans  $\mathbf{E}$  indexée par  $\mathbf{N}$ . A la rigueur, on appelle encore suite, une famille indexée par un intervalle de  $\mathbf{N}$  du type  $\{n_0, n_0 + 1, n_0 + 2, \dots\}$ . On peut, dans ce cas, noter la suite  $(\vec{x})_{n \geq n_0}$ .

En tant qu'ensemble d'applications d'un ensemble dans un espace vectoriel, l'ensemble des suites à valeurs dans  $\mathbf{E}$ , noté  $\mathbf{E}^{\mathbf{N}}$ , se munit d'une structure d'espace vectoriel :

- la somme de deux suites  $(\vec{x}_n)_{n \in \mathbf{N}}$  et  $(\vec{y}_n)_{n \in \mathbf{N}}$  est la suite  $(\vec{x}_n + \vec{y}_n)_{n \in \mathbf{N}}$ ,
- $\alpha \cdot (\vec{x}_n)_{n \in \mathbf{N}}$ , pour  $\alpha$  élément de  $\mathbf{K}$ , est la suite  $(\alpha \cdot \vec{x}_n)_{n \in \mathbf{N}}$

**Définition-proposition 2.1.2** —

- Une suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeurs dans  $\mathbf{E}$  est dite bornée, pour la norme  $\|\cdot\|$ , si, en tant qu'application de  $\mathbf{N}$  dans  $\mathbf{E}$ , elle est bornée, c'est-à-dire si l'ensemble  $\{\|\vec{x}_n\|, n \in \mathbf{N}\}$  est une partie bornée de  $(\mathbf{E}, \|\cdot\|)$ .
- L'ensemble des suites à valeurs dans  $\mathbf{E}$  bornées est un sous espace vectoriel de  $\mathbf{E}^{\mathbf{N}}$ . On le note  $\ell^\infty(\mathbf{N})$  ou plus simplement  $\ell^\infty$ .
- Si  $(\vec{x}_n)_{n \in \mathbf{N}}$  est bornée, alors  $\{\|\vec{x}_n\|, n \in \mathbf{N}\}$  admet une borne supérieure (dans  $\mathbf{R}$ ) et l'application, notée  $N_\infty$ , qui à toute suite bornée  $(\vec{x}_n)_{n \in \mathbf{N}}$  associe  $\sup\{\|\vec{x}_n\|, n \in \mathbf{N}\}$ , est une norme sur l'espace vectoriel des suites bornées.

*Preuve de la proposition 2.1.2.* — il s'agit d'un cas particulier de 1.1.12.

Donnons la définition de la limite d'une suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeurs dans  $\mathbf{E}$ , comme cela a été fait pour une suite à valeurs dans  $\mathbf{K}$ , l'an passé.

Soit donc  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $\mathbf{E}$ , de façon informelle, nous dirons qu'elle converge vers  $\vec{\ell}$ , élément de  $\mathbf{E}$ , si et seulement si  $\vec{x}_n$  est aussi voisin de  $\vec{\ell}$  que nous le désirons, pourvu que nous ayons eu soin de choisir  $n$  supérieur à un certain rang  $n_0$ ; ce qui se traduit mathématiquement de façon rigoureuse par la définition suivante :

**Définition 2.1.3.** — La suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers l'élément  $\vec{\ell}$  de  $\mathbf{E}$ , dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , si, pour tout voisinage  $V$  de  $\vec{\ell}$ , il existe un entier naturel  $n_0$ , tel que pour tout entier naturel  $n$ , si  $n \geq n_0$  alors  $\vec{x}_n$  appartienne à  $V$ .

**Remarque** — On peut, dans cette définition, quitte à « changer de  $n_0$  » remplacer  $n \geq n_0$  par  $n > n_0$ .

On dit encore que  $x_n$  tend vers  $\vec{\ell}$ , lorsque  $n$  tend vers l'infini, ou que  $\vec{\ell}$  est limite de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$ . On note :

$$\vec{x}_n \xrightarrow{n \rightarrow +\infty} \vec{\ell}, \text{ ou encore } \lim_{n \rightarrow +\infty} \vec{x}_n = \vec{\ell}.$$

Une suite qui admet une limite est dite *convergente*, sinon elle est dite *divergente*.

**Proposition 2.1.4.** — La limite, si elle existe, est unique, plus précisément, si la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers l'élément  $\vec{\ell}$  de  $\mathbf{E}$  et converge vers l'élément  $\vec{\ell}'$  de  $\mathbf{E}$ , dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , alors  $\vec{\ell} = \vec{\ell}'$ .

Cette proposition autorise des expressions du type « **La** limite de la suite ».

*Preuve de la proposition 2.1.4.* —

Puisque toute boule ouverte (resp. fermée) de centre  $\vec{\ell}$  de rayon non nul est un voisinage de  $\vec{\ell}$  et que, d'autre part, tout voisinage de  $\vec{\ell}$  contient une boule ouverte (resp. fermée) de centre  $\vec{\ell}$  de rayon non nul, En désignant par  $\varepsilon$  le rayon des boules en question, la définition (2.1.3) donne :

**Proposition 2.1.5.** — La suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers l'élément  $\vec{\ell}$  de  $\mathbf{E}$  dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$  si et seulement si, pour tout réel strictement positif  $\varepsilon$ , il existe un entier naturel  $n_0$  tel que pour tout entier naturel  $n$ , si  $n \geq n_0$  alors  $\|\vec{x}_n - \vec{\ell}\| < \varepsilon$  (resp.  $\|\vec{x}_n - \vec{\ell}\| \leq \varepsilon$ ).

On retrouve ainsi dans le cas particulier où l'e.v.n. est  $(\mathbf{R}, |\cdot|)$ , la définition usuelle donnée dans les petites classes.

Notons également que :

**Proposition 2.1.6.** — La suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers l'élément  $\vec{\ell}$  de  $\mathbf{E}$  dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , si et seulement si la suite  $(\|\vec{x}_n - \vec{\ell}\|)_{n \in \mathbf{N}}$  converge vers 0 dans l'e.v.n.  $(\mathbf{R}, |\cdot|)$ .

**Remarque** — Nous avons eu soin de préciser dans la définition de  $\vec{x}_n \xrightarrow{n \rightarrow +\infty} \vec{\ell}$ , « dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$  ». Cette précision est nécessaire, en effet, si  $\|\cdot\|'$  désigne une autre norme sur  $\mathbf{E}$ , la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  peut très bien converger vers  $\vec{\ell}$ , dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$  et ne pas converger vers  $\vec{\ell}$ , dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|')$ . Prenons l'exemple suivant : On désigne par  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([0, 1], \mathbf{R})$  ; considérons la suite  $(f_n)_{n \in \mathbf{N}}$  où pour tout  $n \in \mathbf{N}$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R} ; t \mapsto t^n.$$

Dans  $(\mathbf{E}, N_1)$ , la suite converge vers  $0_{[0, 1]}$ , fonction nulle de  $[0, 1]$ , ce d'après 2.1.6., en effet, pour tout  $N \in \mathbf{N}$ ,  $N_1(f_n - 0_{[0, 1]}) = \frac{1}{n+1}$  et tend donc vers zéro dans  $(\mathbf{R}, |\cdot|)$ . En revanche, pour tout  $N \in \mathbf{N}$ ,  $N_\infty(f_n - 0_{[0, 1]})$  vaut 1 ! Donc dans  $(\mathbf{E}, N_\infty)$  la suite  $(f_n)_{n \in \mathbf{N}}$  ne converge pas vers  $0_{[0, 1]}$  (toujours 2.1.6. ).

Bien sûr, lorsqu'il n'y a pas d'ambiguïté, nous écrirons tout simplement «  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers  $\vec{\ell}$  » . Souvent, du reste pour lever toute ambiguïté nous parlerons de suite « à valeurs dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$  » ; ce qui signifie que la suite est à valeurs dans l'ensemble  $\mathbf{E}$  et que la convergence éventuelle de la suite s'étudie dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$  .

Etudions maintenant les propriétés des limites. Le passage à la limite se comporte bien vis à vis des lois de  $\mathbf{E}^{\mathbf{N}}$  ; plus précisément :

**Proposition 2.1.7.** — Soient  $(\vec{x}_n)_{n \in \mathbf{N}}$  et  $(\vec{y}_n)_{n \in \mathbf{N}}$  des suites à valeurs dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , admettant respectivement comme limites  $\vec{\ell}$  et  $\vec{\ell}'$ . Soit  $\alpha$  un élément de  $\mathbf{K}$ . Alors :

1.  $(\vec{x}_n)_{n \in \mathbf{N}} + (\vec{y}_n)_{n \in \mathbf{N}}$  converge vers  $\vec{\ell} + \vec{\ell}'$ .
2.  $\alpha \cdot (\vec{x}_n)_{n \in \mathbf{N}}$  converge vers  $\alpha \cdot \vec{\ell}$ .

Preuve de la proposition 2.1.7. —

Dans le cas particulier où  $(\mathbf{E}, \|\cdot\|)$  vaut  $(\mathbf{K}, |\cdot|)$  on a de plus :

**Proposition 2.1.8.** — : Soient  $(x_n)_{n \in \mathbf{N}}$  et  $(y_n)_{n \in \mathbf{N}}$  des suites à valeurs dans l'e.v.n.  $(\mathbf{K}, |\cdot|)^1$ , admettant respectivement comme limites  $\vec{\ell}$  et  $\vec{\ell}'$ , alors

1.

$$x_n \times y_n \xrightarrow{n \rightarrow +\infty} \ell \times \ell'.$$

2. Si de plus  $x_n$  est différent de zéro pour tout élément  $n$  de  $\mathbf{N}$  et si  $\vec{\ell}$  est non nul, alors :

$$\frac{1}{x_n} \xrightarrow{n \rightarrow +\infty} \frac{1}{\ell}$$

Preuve de la proposition 2.1.8. — Montrons par exemple, le premier point.

Soit  $\varepsilon$  un réel arbitraire strictement positif. Evaluons, pour tout entier naturel  $n$ ,  $|\ell\ell' - x_n y_n|$ . Pour ce faire, regardons la figure ci-dessous, où le produit  $\ell\ell'$  se représente par un rectangle de côtés  $\ell$  et  $\ell'$  et qui nous suggère d'écrire :

$$\ell\ell' - x_n y_n = (\ell(\ell' - y_n)) + (y_n(\ell - x_n)).$$

Ce qui conduit à la majoration :

$$|\ell\ell' - x_n y_n| \leq |\ell(\ell' - y_n)| + |y_n(\ell - x_n)|$$

Posons pour tout entier naturel  $n$ ,  $A_n := \ell(\ell' - y_n)$ ,  $B_n := y_n(\ell - x_n)$ .

- Majorons pour commencer  $|A_n|$ .

Puisque  $(y_n)_{n \in \mathbf{N}}$  tend vers  $\ell'$ , il existe un entier naturel  $n_1$ , tel que, pour tout entier naturel  $n$ , si  $n > n_1$ , alors

$$|\ell' - y_n| < \frac{\varepsilon}{2(|\ell| + 1)},$$

(le « +1 » au dénominateur est là juste au cas où  $\ell$  serait nul). Donc pour tout entier  $n > n_1$ ,

$$|A_n| < |\ell| \frac{\varepsilon}{2(|\ell| + 1)} < \frac{\varepsilon}{2}. \quad (\text{III.4})$$

- Majorons maintenant  $|B_n|$ .

Pour tout entier  $n > n_1$ ,  $|\ell' - y_n| < \frac{\varepsilon}{2(|\ell| + 1)}$ , et donc  $|y_n| < |\ell'| + \frac{\varepsilon}{2(|\ell| + 1)}$ . Appelons  $D$  la quantité strictement positive  $|\ell'| + \frac{\varepsilon}{2(|\ell| + 1)}$ . Puisque  $(x_n)_{n \in \mathbf{N}}$  tend vers  $\ell$ , il existe un entier naturel  $n_2$ , tel que, pour tout entier naturel  $n$ , si  $n \geq n_2$ , alors  $|\ell - x_n| < \frac{\varepsilon}{2D}$ . Posons  $N_0 = \max\{n_0, n_1\}$ . Pour tout entier naturel  $n > N_0$ ,

$$|B_n| < D \frac{\varepsilon}{2D} = \frac{\varepsilon}{2}. \quad (\text{III.5})$$

1. le choix de la norme sur  $\mathbf{K}$  est, comme nous le verrons, sans importance.

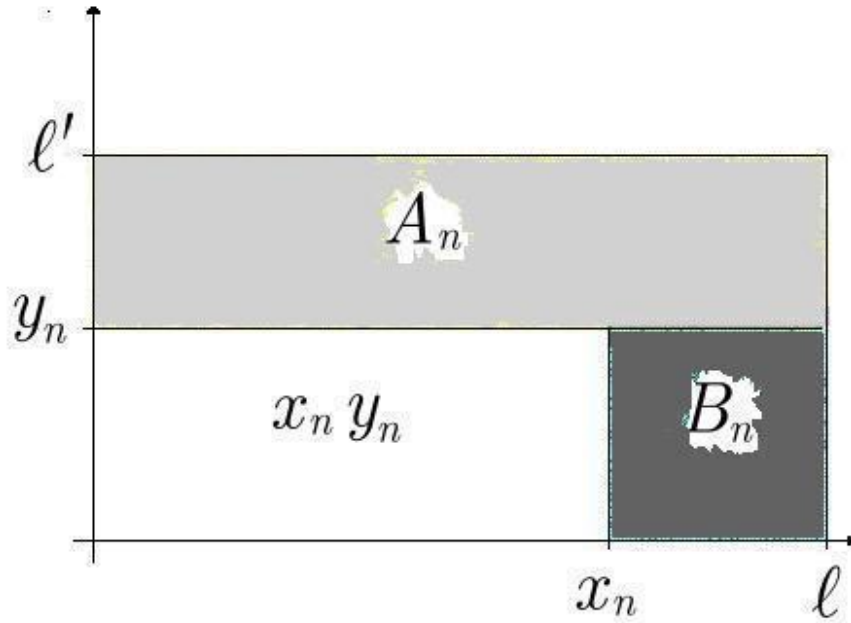


FIGURE III.7 –

Il résulte de (III.4–III.5) que pour tout entier naturel  $n$ , si  $n > N_0$ , alors

$$|\ell\ell' - x_n y_n| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Le réel  $\varepsilon$  étant arbitraire, la suite  $(x_n \times y_n)_{n \in \mathbf{N}}$  converge vers  $\ell \times \ell'$ . □

**Exercice 2.1.9.** — Soient  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $\mathbf{E}$  admettant comme limite  $\vec{\ell}$  dans  $(\mathbf{E}, \|\cdot\|)$  et  $(\alpha_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $\mathbf{K}$  admettant  $\lambda$  comme limite dans  $(\mathbf{R}, |\cdot|)$ . Montrer, en s’inspirant de 2.1.8., que  $(\alpha_n \cdot \vec{x}_n)_{n \in \mathbf{N}}$  converge vers  $\lambda \cdot \vec{\ell}$  dans  $(\mathbf{E}, \|\cdot\|)$ .

**Proposition 2.1.10.** — Soit  $(\mathbf{E}_1, |\cdot|_1), (\mathbf{E}_2, |\cdot|_2), \dots, (\mathbf{E}_p, |\cdot|_p)$  des e.v.n. On munit l’espace vectoriel produit  $\mathbf{E} = \mathbf{E}_1 \times \mathbf{E}_2 \times \dots \times \mathbf{E}_p$  de la norme  $n'_\infty$  définie au 1.1.7.5. par,

$$n'_\infty(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = \sup_{i=1,2,\dots,p} (|\vec{x}_i|_i).$$

Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $(\mathbf{E}, n'_\infty)$ . On notera, pour  $i = 1, 2, \dots, p$  et tout  $n \in \mathbf{N}$ ,  $\vec{x}_{i,n}$  la  $i^{\text{e}}$  composante de  $\vec{x}_n$ . Soit  $\vec{\ell}$  un élément de  $\mathbf{E}$ ,  $\vec{\ell} = (\vec{\ell}_1, \vec{\ell}_2, \dots, \vec{\ell}_p)$ .

La suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers  $\vec{\ell}$  si et seulement si la suite  $(x_{i,n})_{n \in \mathbf{N}}$  converge vers  $\vec{\ell}_i$  dans  $(\mathbf{E}_i, |\cdot|_i)$ , pour  $i = 1, 2, \dots, p$ .



*Preuve de la proposition 2.1.10. —*

**Proposition 2.1.11.** — *Toute suite à valeur dans un e.v.n. convergente y est bornée.*

*Preuve de la proposition 2.1.11. —* Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , convergeant vers  $\vec{\ell}$ . Il existe un entier naturel  $n_0$  tel que, pour tout entier naturel  $n$ , si  $n > n_0$  alors  $\|\vec{x}_n - \vec{\ell}\| \leq 1$ , donc tel que pour tout entier naturel  $n > n_0$   $\|\vec{x}_n\| \leq \|\vec{\ell}\| + 1$ . Donc pour tout entier naturel  $n$ ,

$$\|\vec{x}_n\| \leq \max\{\|\vec{\ell}\| + 1, \|\vec{x}_0\|, \|\vec{x}_1\|, \dots, \|\vec{x}_{n_0}\|\}.$$

Donc pour tout entier naturel  $n$ ,  $\vec{x}_n$  est élément de la boule  $B_f(\vec{0}_{\mathbf{E}}, R)$ , où  $R = \max\{\|\vec{\ell}\| + 1, \|\vec{x}_0\|, \|\vec{x}_1\|, \dots, \|\vec{x}_{n_0}\|\}$ .

Donc la suite est bornée. □

**Contre-exemple** —  $((-1)^n)_{n \in \mathbf{N}}$  est bornée dans  $(\mathbf{R}, \|\cdot\|)$ , mais non convergente : la réciproque de 2.1.11. est fausse.

**Remarque** — Dans le cas où  $\mathbf{E} = \mathbf{R}$  il a été étudié, l'an passé, une notion de limite généralisée dans  $\bar{\mathbf{R}}^1$ , le lecteur est prié de se reporter au cours de MPSI.

Les suites constituent un outil précieux pour caractériser les points de l'adhérence d'une partie. On a :

**Proposition 2.1.12.** — *Soit  $A$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . Un point  $\vec{x}$  de  $(\mathbf{E}, \|\cdot\|)$ , est adhérent à  $A$ , si et seulement si il est la limite dans  $(\mathbf{E}, \|\cdot\|)$  d'une suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeurs dans  $A$  (c'est-à-dire telle que  $\vec{x}_n$  soit élément de  $A$  pour tout entier naturel  $n$ ).*

---

1. La définition de la limite dans  $\mathbf{R}$  est du reste la même que celle du 2.1.3., si l'on munit  $\bar{\mathbf{R}}$  de la topologie, au sens de la note de la proposition 1.2.2., définie de la façon suivante : Les ouverts de  $\bar{\mathbf{R}}$  sont les parties  $A$  telles que pour tout élément  $x$  de  $A$ , il existe un intervalle  $I$  de la forme  $]a, b[$  ou  $[-\infty, b[$  ou  $]a, +\infty]$ , avec  $-\infty < a < b < +\infty$ , tel que  $\{x\} \subset I \subset A$ . Le lecteur vérifiera qu'est bien ainsi définie une topologie (c'est-à-dire qu'est satisfaite la proposition 1.2.2.). Exemples d'ouverts de  $\bar{\mathbf{R}}$  :  $[+\infty, 3[$ ,  $]-\infty, 3[$ ,  $]3, +\infty]$ ,  $\bar{\mathbf{R}} - \{5\}$ ...

*Preuve de la proposition 2.1.12 —*

La proposition 2.1.12. est très utile pour démontrer qu'une partie  $F$  d'un e.v.n. est fermée.

**Proposition 2.1.13.** — *Soit  $F$  une partie d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . La partie  $F$  est fermée si et seulement si la limite de toute suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeur dans  $F$  convergente est dans  $F$ .*

*Soit  $A$  une partie de  $\mathbf{E}$  et  $F_A$  une partie de  $A$ . La partie  $F_A$  est un fermé relatif de  $A$  si et seulement si la limite de toute suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeur dans  $F_A$  convergent vers un élément de  $A$ , est dans  $F$ .*

**Preuve de la proposition 2.1.13.** —

- PREMIER POINT

- SECOND POINT

**Exemples —**

1. Soit  $\mathbf{E}$  l'ensemble des applications de  $[0, 1]$  dans  $\mathbf{R}$  continues, muni de la norme  $N_1$ . Soit  $F$  l'ensemble des éléments de  $\mathbf{E}$  qui prennent en 0 la valeur 1. Montrons que  $F \dots$

2. Soit  $(\mathbf{E}, \|\cdot\|)$  un espace euclidien<sup>1</sup> de dimension 2, plutôt considéré ici comme un plan affine, c'est-à-dire, concrètement que nous noterons ses éléments comme des points  $A, B, M, \dots$   $U$  un ouvert de  $(\mathbf{E}, \|\cdot\|)$ , et  $A$  un point de  $U$ . On note  $F$  l'ensemble des éléments  $M$  de  $U$  qui peuvent être joints à  $A$  par une ligne brisée incluse

---

1. la norme est évidemment celle associée au produit scalaire)

dans  $U$ . Montrons que  $F$  est un fermé relatif de  $U$ .

Les sous-suites se définissent comme dans le cas réel.

**Définition 2.1.14.** — SUITE EXTRAITES —

Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans un espace vectoriel  $\mathbf{E}$ . On appelle sous-suite de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$ , toute suite  $(\vec{y}_p)_{p \in \mathbf{N}}$ , telle qu'il existe une application  $\varphi$  de  $\mathbf{N}$  dans  $\mathbf{N}$ , strictement croissante, telle que pour tout élément  $p$  de  $\mathbf{N}$ ,  $\vec{y}_p = \vec{x}_{\varphi(p)}$ . On dit encore que la suite  $(\vec{y}_p)_{p \in \mathbf{N}}$  est une suite extraite de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$ , et on la note souvent  $(\vec{x}_{\varphi(p)})_{p \in \mathbf{N}}$ .

On montre comme en MPSI dans le cas réel que :

**Proposition 2.1.15.** — Toute suite extraite d'une suite convergente converge vers la même limite.

**Remarque** — Une suite peut diverger alors même que certaines de ses sous-suites convergent. Ainsi  $((-1)^n)_{n \in \mathbf{N}}$  est-elle une suite divergente alors que  $((-1)^{2p})_{p \in \mathbf{N}}$  et  $((-1)^{2p+1})_{p \in \mathbf{N}}$  convergent.

Du nouveau !

**Définition 2.1.16.** — Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , soit  $\vec{\ell}$  un élément de  $\mathbf{E}$ . On dit que  $\vec{\ell}$  est une valeur d'adhérence de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$ , si, pour tout voisinage  $V$  de  $\vec{\ell}$ , l'ensemble des entiers  $n$  tels que  $\vec{x}_n$  est élément de  $V$  est infini.

On peut donner deux formulations moins parlantes mais plus utiles dans la pratique de cette définition. En gardant les notations de 2.1.15 :

**Proposition 2.1.17.** — L'élément  $\vec{\ell}$  est une valeur d'adhérence de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  si et seulement si, pour tout voisinage  $V$  de  $\vec{\ell}$  et tout entier naturel  $n_0$ , il existe un entier naturel  $n$  supérieur<sup>1</sup> à  $n_0$ , tel que  $\vec{x}_n \in V$ .

**Proposition 2.1.18.** — L'élément  $\vec{\ell}$  est une valeur d'adhérence de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  si et seulement si, pour tout réel strictement positif  $\varepsilon$  et tout entier naturel  $n_0$ , il existe un entier naturel  $n$  supérieur à  $n_0$ , tel que  $\|\vec{x}_n - \vec{\ell}\| < \varepsilon$  (ou  $\|\vec{x}_n - \vec{\ell}\| \leq \varepsilon$ ).

*Preuve des propositions 2.1.17–8* — La preuve de ces propositions est laissée en exercice.

**Exemple** — Dans  $(\mathbf{R}, |\cdot|)$  la suite  $((-1)^n)_{n \in \mathbf{N}}$  admet deux valeurs d'adhérence 1 et  $-1$ .

Une suite convergeant vers  $\vec{\ell}$ , admet  $\vec{\ell}$  comme valeur d'adhérence. On a le résultat plus fort suivant :

**Proposition 2.1.19.** — Une suite convergente admet une unique valeur d'adhérence, sa limite.

1. Au sens strict ou large, c'est comme on préfère !

*Preuve de la proposition 2.1.19. —*

### Remarques

- Il résulte de 2.1.19. qu'en particulier, une suite ayant deux valeurs d'adhérences distinctes diverge. Par exemple  $((-1)^n)_{n \in \mathbf{N}}$  diverge.
- La réciproque de 2.1.19. est fausse, ainsi la suite

diverge et admet une seule valeur d'adhérence. Par contre on a la proposition suivante.

**Proposition 2.1.20.** — *Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  et  $\vec{\ell}$  un élément de  $\mathbf{E}$ .  $\vec{\ell}$  est valeur d'adhérence de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  si et seulement si il existe une sous-suite de la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  qui converge vers  $\vec{\ell}$ .*

*Preuve de la proposition 2.1.20. —*

Définissons pour des suites à valeurs vectorielles, comme l'an passé pour des suites réelles, les relations de *prépondérance* de *domination* et d'*équivalence*. Nous nous bornerons à redonner dans le cas général au programme les définitions déjà rencontrées en Sup.

**Définition 2.1.21.** — Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , et  $(\alpha_n)_{n \in \mathbf{N}}$  une suite à valeurs réelles. On dit que la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  est *négligeable* devant la suite  $(\alpha_n)_{n \in \mathbf{N}}$ , où que la suite  $(\alpha_n)_{n \in \mathbf{N}}$  est *prépondérante* devant la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  si, pour tout réel  $\varepsilon$ , strictement positif, il existe un entier positif  $n_0$  tel que, pour tout entier naturel  $n$ , si  $n \geq n_0$ , alors

$$\|\vec{x}_n\| < \varepsilon |\alpha_n|.$$

On note  $\vec{x}_n \underset{n \rightarrow +\infty}{=} o(\alpha_n)$ , (on dit  $\vec{x}_n$  est un « petit o de  $\alpha_n$  »).

**Remarque 2.1.22.** — Observons que dans le cas où la suite  $(\alpha_n)_{n \in \mathbf{N}}$  ne s'annule pas  $(\vec{x}_n)_{n \in \mathbf{N}}$  est donc négligeable devant la suite  $(\alpha_n)_{n \in \mathbf{N}}$ , si et seulement si, dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ ,

$$\frac{1}{\alpha_n} \cdot \vec{x}_n \underset{n \rightarrow +\infty}{\rightarrow} \vec{0}_{\mathbf{E}}.$$

**Exemples —**

1. Dans
- $(\mathbf{C}, |\cdot|)$
- ,

$$\frac{8}{1+n^2} e^{in\theta} \underset{n \rightarrow +\infty}{=} o\left(\frac{1}{n+1}\right).$$

2. Pour tout entier naturel
- $n$
- , considérons l'application

$$f_n : [0, 1] \rightarrow \mathbf{R}; t \mapsto t^n.$$

On a dans  $(\mathcal{C}^0([0, 1], \mathbf{R}), N_\infty)$  :

$$f_n \underset{n \rightarrow +\infty}{=} o(\sqrt{n}).$$

**Définition 2.1.23.** — Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , et  $(\alpha_n)_{n \in \mathbf{N}}$  une suite à valeurs réelles. On dit que la suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  est dominée par la suite  $(\alpha_n)_{n \in \mathbf{N}}$ , si il existe un réel strictement positif  $M$  et un entier positif  $n_0$ <sup>1</sup> tels que, pour tout entier naturel  $n$ , si  $n \geq n_0$ , alors

$$\|\vec{x}_n\| < M|\alpha_n|.$$

On note  $\vec{x}_n \underset{n \rightarrow +\infty}{=} O(\alpha_n)$ , (on dit  $\vec{x}_n$  est un  $\ll$  grand  $o$  de  $\alpha_n \gg$ ).

**Remarque 2.1.24.** — Observons que dans le cas où la suite  $(\alpha_n)_{n \in \mathbf{N}}$  ne s'annule pas  $(\vec{x}_n)_{n \in \mathbf{N}}$  est donc dominée par la suite  $(\alpha_n)_{n \in \mathbf{N}}$ , si et seulement si, la suite  $\left(\frac{1}{\alpha_n} \cdot \vec{x}_n\right)_{n \in \mathbf{N}}$  est bornée dans l'e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .

**Exemples —**

1. Dans
- $(\mathbf{C}, |\cdot|)$
- ,

$$\frac{8}{1+n^2} e^{in\theta} \underset{n \rightarrow +\infty}{=} O\left(\frac{1}{n^2+1}\right).$$

2. Avec les notations des exemples précédent Dans
- $(\mathcal{C}^0([0, 1], \mathbf{R}), N_\infty)$
- :
- $(f_n)_{n \in \mathbf{N}}$
- est dominée par toute suite réelle constante

**Définition 2.1.25.** — Soient  $(\vec{x}_n)_{n \in \mathbf{N}}$  et  $(\vec{y}_n)_{n \in \mathbf{N}}$  des suites à valeurs dans un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , on dit que  $(\vec{x}_n)_{n \in \mathbf{N}}$  est équivalente à  $(\vec{y}_n)_{n \in \mathbf{N}}$  et l'on note  $\vec{x}_n \underset{n \rightarrow +\infty}{\sim} \vec{y}_n$  si  $\vec{x}_n - \vec{y}_n \underset{n \rightarrow +\infty}{=} o(\|\vec{x}_n\|)$ .

**Remarque 2.1.26.** — Observons que dans le cas de deux suites réelles  $(x_n)_{n \in \mathbf{N}}$  et  $(y_n)_{n \in \mathbf{N}}$  si la suite  $(y_n)_{n \in \mathbf{N}}$  ne s'annule pas, alors  $x_n \underset{n \rightarrow +\infty}{\sim} y_n$  si et seulement si la suite  $\left(\frac{x_n}{y_n}\right)_{n \in \mathbf{N}}$  converge vers 1 dans  $(\mathbf{R}, |\cdot|)$ .

**Proposition 2.1.27.** — La relation  $\ll$  être équivalente  $\gg$ , sur l'ensemble des suites à valeurs dans un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ , est une relation... d'équivalence!

*Preuve de la proposition 2.1.27.* — Laissée en exercice.

**2.2 Limite d'une application, continuité**

Par le passé on a défini la limite d'une application de  $\mathbf{R}$  dans  $\mathbf{R}$  au point  $a$  :

$f(x)$  tend vers  $\ell$  lorsque  $x$  tend vers  $a$ , si pour tout réel  $\varepsilon$  strictement positif, il existe un réel strictement positif  $\eta$ , tel que pour tout élément  $x$  de  $\mathbf{R}$ , si  $|x - a| < \eta$  alors  $|f(x) - \ell| \leq \varepsilon$ .

C'est-à-dire, plus prosaïquement, que  $f(x)$  est aussi voisin de  $\ell$  que l'on veut, pourvu que  $x$  soit suffisamment voisin de  $a$ . Cette dernière formulation nous suggère la généralisation de la définition de limite aux applications d'un e.v.n. dans un autre.

**Définition 2.2.1.** — Soient  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  et  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  des e.v.n., soit  $\vec{f}$  une application d'une partie  $D$  de  $\mathbf{E}$  dans  $\mathbf{F}$ . Soit  $\vec{a}$  un élément de  $\mathbf{E}$  adhérent à  $D$ , soit enfin  $\vec{\ell}$  un élément de  $\mathbf{F}$ . On dit que  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ , si pour tout voisinage  $V$  de  $\vec{\ell}$ , il existe un voisinage  $U$  de  $\vec{a}$  tel que, pour tout élément  $\vec{x}$ , de  $D$ , si  $x \in U$  alors  $\vec{f}(\vec{x}) \in V$  (ou encore  $\vec{f}(U \cap D) \subset V$ ).

Les précisions  $\ll$  tend vers  $\vec{\ell}$   $\gg$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  et  $\ll$  lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$   $\gg$  sont nécessaires, car, comme pour la convergence d'une suite, la convergence d'une application dépend du choix des normes sur  $\mathbf{E}$  et  $\mathbf{F}$ . Lorsqu'il n'y a pas d'ambiguïté sur les normes dont on équipe  $\mathbf{E}$  et  $\mathbf{F}$ , on omet cette précision ; pour lever toute ambiguïté du reste,

1. On peut alors montrer qu'il est loisible de prendre  $n_0 = 0$

on peut écrire « soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  ».

On dit encore que  $\vec{\ell}$  est limite de  $\vec{f}$  en  $\vec{a}$  ou que  $\vec{f}$  converge vers  $\vec{\ell}$  en  $\vec{a}$ . On note  $\vec{f}(\vec{x}) \xrightarrow{\vec{x} \rightarrow \vec{a}} \vec{\ell}$  ou  $\lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x}) = \vec{\ell}$ , ou encore  $\lim_{\vec{a}} \vec{f} = \vec{\ell}$ .

**Remarque** — Nous avons pris dans la définition  $\vec{a}$  adhérent à  $D$ , c'est une obligation en effet ...

Puisque toute boule de rayon non nul ouverte (resp. fermée) de centre  $\vec{\ell}$  est un voisinage particulier de  $\vec{\ell}$  et que d'autre part tout voisinage de  $\vec{\ell}$  contient une boule ouverte (resp. fermée) de centre  $\vec{\ell}$  de rayon non nul, on peut remplacer dans la définition 2.2.1. « tout voisinage  $V$  de  $\vec{\ell}$  » par « pour toute boule ouverte (resp. fermée) de centre  $\vec{\ell}$  ». De même peut-on remplacer « il existe un voisinage  $U$  de  $\vec{a}$  » par « il existe une boule ouverte (resp. fermée) de centre  $\vec{a}$  ». En désignant par  $\varepsilon$  le rayon des boules de centre  $\vec{\ell}$  et par  $\eta$  celui des boules de centre  $\vec{a}$ , la définition 2.2.1. devient :

**Proposition 2.2.2.** — Avec les notation de 2.2.1,  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ , si et seulement pour tout réel strictement positif  $\varepsilon$ , il existe un réel strictement positif  $\eta$ , tel que pour tout  $\vec{x}$  élément de  $D$ ,  $\|\vec{x} - \vec{a}\|_{\mathbf{E}} < (\leq) \eta$  alors  $\|\vec{f}(\vec{x}) - \vec{\ell}\|_{\mathbf{F}} < (\leq) \varepsilon$ .

Il est pratique d'introduire la notion plus générale de limite selon une partie  $P$  :

**Définition 2.2.3.** — soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  soit  $P$  une partie de  $D$  et soit  $\vec{a}$  un élément de  $\mathbf{E}$  adhérent à  $P$ , soit enfin  $\vec{\ell}$  un élément de  $\mathbf{F}$ . l'élément  $\vec{\ell}$  est la limite de  $\vec{f}$  en  $\vec{a}$  selon  $P$ , si, pour tout voisinage  $V$  de  $\vec{\ell}$ , il existe un voisinage  $U$  de  $\vec{a}$  tel que

$$\vec{f}(U \cap P) \subset V.$$

Ce qui ne signifie rien d'autre que  $\vec{\ell}$  est la limite de la restriction de  $\vec{f}$  à  $P$  en  $\vec{a}^1$ . On note  $\vec{f}(\vec{x}) \xrightarrow[\substack{\vec{x} \rightarrow \vec{a} \\ x \in P}} \vec{\ell}$  ou  $\lim_{\substack{\vec{x} \rightarrow \vec{a} \\ x \in P}} \vec{f}(\vec{x}) = \vec{\ell}$ .

**Remarque 2.2.4.** — Notons que si  $\vec{a}$  est élément de  $D$  et si  $\vec{f}$  admet  $\vec{\ell}$  comme limite en  $\vec{a}$  alors nécessairement  $\vec{\ell} = \vec{f}(\vec{a})$ . En effet...

C'est pourquoi, un choix utile est fréquent de la partie  $P$  dans 2.2.3. est  $P = D - \{\vec{a}\}$ . On dit alors que  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  par valeurs distinctes et l'on parle de limite *épointée*; on note :  $\vec{f}(\vec{x}) \xrightarrow[\substack{\vec{x} \rightarrow \vec{a} \\ x \neq \vec{a}}]{} \vec{\ell}$  ou  $\lim_{\substack{\vec{x} \rightarrow \vec{a} \\ x \neq \vec{a}}} \vec{f}(\vec{x}) = \vec{\ell}$ .

$\lim_{\substack{\vec{x} \rightarrow \vec{a} \\ x \neq \vec{a}}} \vec{f}(\vec{x}) = \vec{\ell}$  se traduit en utilisant 2.2.2. comme suit.

Pour tout réel strictement positif  $\varepsilon$ , il existe un réel strictement positif  $\eta$ , tel que pour tout  $\vec{x}$  élément de  $D$  distinct de  $\vec{a}$ ,  $\|\vec{x} - \vec{a}\|_{\mathbf{E}} < (\leq) \eta$  alors  $\|\vec{f}(\vec{x}) - \vec{\ell}\|_{\mathbf{F}} < (\leq) \varepsilon$ .

Un choix également fort fréquent de la partie  $P$ , lorsque  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  est  $(\mathbf{R}, |\cdot|)$  est

$$P = D \cap ]-\infty, a[ \quad (\text{resp. } P = D \cap ]a, +\infty[).$$

On parle alors de *limite à gauche* (resp. de limite à droite en  $a$ ). On note :

$$\vec{f}(x) \xrightarrow[\substack{x \rightarrow a \\ x < a}]{} \vec{\ell} \quad (\text{resp. } \vec{f}(x) \xrightarrow[\substack{x \rightarrow a \\ x > a}]{} \vec{\ell}) \quad \text{ou} \quad \vec{f}(x) \xrightarrow{x \rightarrow a^-} \vec{\ell} \quad (\text{resp. } \vec{f}(x) \xrightarrow{x \rightarrow a^+} \vec{\ell})$$

1. C'est pourquoi, dans la suite, nous nous limiterons à énoncer certains résultats pour des limites ordinaires sans partie filtrante  $P$ , qui sont donc vrais avec une partie filtrante.



ou encore

$$\lim_{\substack{x \rightarrow a \\ x < a}} \vec{f}(x) = \vec{\ell} \text{ ( resp. } \lim_{\substack{x \rightarrow a \\ x > a}} \vec{f}(x) = \vec{\ell} \text{ ) ou } \lim_{x \rightarrow a^-} \vec{f}(x) = \vec{\ell} \text{ ( resp. } \lim_{x \rightarrow a^+} \vec{f}(x) = \vec{\ell} \text{ )}.$$

On écrit encore parfois  $\vec{\ell} = \vec{f}(a_{-0})$  (resp.  $\vec{\ell} = \vec{f}(a_{+0})$ )

Donnons maintenant une caractérisation séquentielle de la limite sous les hypothèses de 2.2.3.

**Proposition 2.2.5.** —  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$  dans l'e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ , selon  $P$ , si et seulement si pour toute suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  à valeurs dans  $P$  de limite  $\vec{a}$ , dans  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ , la suite  $(\vec{f}(\vec{x}_n))_{n \in \mathbf{N}}$  converge vers  $\vec{\ell}$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ .

*Preuve de la proposition 2.2.5.* —

**Corollaire 2.2.6.** — La limite est unique, plus précisément, avec les notations de la définition II.2.1., si  $\lim_{\vec{a}} \vec{f} = \vec{\ell}$  et  $\lim_{\vec{a}} \vec{f} = \vec{\ell}'$  alors  $\vec{\ell} = \vec{\ell}'$ .

*Preuve du corollaire 2.2.6.* — Résulte immédiatement de 2.2.5. et de l'unicité de la limite d'une suite (proposition 2.1.4.).

**Corollaire 2.2.7.** — Soient  $\vec{f}$  et  $\vec{g}$  des applications d'une partie  $D$  d'une e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ ,  $\alpha$  un élément de  $\mathbf{K}$ ,  $\vec{a}$  un point de  $\bar{A}$ ,  $\vec{\ell}$  et  $\vec{\ell}'$  des éléments de  $\mathbf{F}$ . On suppose que :

$$\lim_{\vec{a}} \vec{f} = \vec{\ell} \text{ et } \lim_{\vec{a}} \vec{g} = \vec{\ell}'.$$

Alors :

1.

$$\lim_{\vec{a}} \vec{f} + \vec{g} = \vec{\ell} + \vec{\ell}'.$$

2.

$$\lim_{\vec{a}} \alpha \cdot \vec{f} = \alpha \cdot \vec{\ell}.$$

3. Si de plus  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  est  $(\mathbf{K}, |\cdot|)$ ,

$$\lim_{\vec{a}} f \times g = \ell \times \ell'.$$

4. Enfin, si  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  est  $(\mathbf{K}, |\cdot|)$  et si  $\ell' \neq 0$ , alors  $\vec{a}$  est adhérent à  $D'$ , où  $D'$  est le domaine de définition de  $\frac{f}{g}$  et

$$\lim_{\vec{a}} \frac{f}{g} = \frac{\ell}{\ell'}.$$

*Preuve du corollaire 2.2.7.* —

**Proposition 2.2.8.** — Soient  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ ,  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  et  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$  des e.v.n. Soit  $\vec{f}$  une application d'une partie  $D$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , Soit  $\vec{g}$  une application d'une partie  $D'$  de  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  dans  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$ . On suppose que  $\vec{f}(D) \subset D'$ , de sorte que l'on peut définir sur  $D$  l'application  $\vec{g} \circ \vec{f}$ . Soit  $\vec{a}$  un élément adhérent à  $D$ . Supposons enfin qu'il existe  $\vec{\ell}$ , élément de  $\mathbf{F}$  tel que  $\lim_{\vec{a}} \vec{f} = \vec{\ell}$ . Alors :

1. L'élément  $\vec{\ell}$  est adhérent à  $D'$ .

2. Si, de plus, il existe  $\vec{\ell}'$ , élément de  $\mathbf{G}$ , tel que  $\lim_{\vec{\ell}} \vec{g} = \vec{\ell}'$ , alors  $\lim_{\vec{a}} \vec{g} \circ \vec{f} = \vec{\ell}'$ .

*Preuve de la proposition 2.2.8. —*

**Proposition 2.29.** — *Soit  $(\mathbf{F}_1, |\cdot|_1), (\mathbf{F}_2, |\cdot|_2), \dots, (\mathbf{F}_p, |\cdot|_p)$  des e.v.n. On munit l'espace vectoriel produit  $\mathbf{F} = \mathbf{F}_1 \times \mathbf{F}_2 \times \dots \times \mathbf{F}_p$ , comme à l'accoutumée de la norme  $n'_\infty$  définie au 1.1.7.5. par,*

$$n'_\infty(\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n) = \sup_{i=1,2,\dots,p} (|\vec{y}_i|_i).$$

*Soit  $f$  une application d'une partie  $D$  d'une e.v.n.  $(\mathbf{E}, \|\cdot\|)$  dans  $(\mathbf{F}, n'_\infty)$ . Pour  $i = 1, \dots, p$ ,  $\vec{f}_i$  désigne la  $i^{\text{e}}$  composante de  $\vec{f}$  ( $f = (\vec{f}_1, \dots, \vec{f}_p)$ ). Soit  $\vec{a}$  un point de  $\mathbf{E}$  adhérent à  $D$  et  $\vec{\ell} = (\vec{\ell}_1, \dots, \vec{\ell}_p)$  un élément de  $\mathbf{F}$ . Alors,  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$  lorsque  $\vec{x}$  tend vers  $\vec{a}$ , si et seulement si,  $\vec{f}_i(\vec{x})$  tend vers  $\vec{\ell}_i$  lorsque  $\vec{x}$  tend vers  $\vec{a}$ , dans  $(\mathbf{F}_i, |\cdot|_i)$ , pour  $i=1,2,\dots,p$ .*

*Preuve de la proposition 2.2.9 —*

## LIMITE GÉNÉRALISÉE

Il a été vu en sup. des définitions de *limite généralisée* d'une application de  $\mathbf{R}$  dans  $\mathbf{R}$ , Rappelons en quelques unes :

- Soient  $f$  une application d'un intervalle  $I$  dans  $\mathbf{R}$  et  $a$  une extrémité de  $I$ . On dit que  $f$  admet  $+\infty$  comme limite en  $a$  si pour tout réel  $B$ , il existe un réel  $\eta > 0$  tel que pour tout  $x \in I$ , si  $|x - a| \leq \eta$ , alors :

$$f(x) \geq B$$

- Soient  $f$  une application d'un intervalle  $I$  dont l'extrémité supérieure est  $+\infty$  dans  $\mathbf{R}$  et  $\ell$  un réel. On dit que  $f$  admet  $\ell$  comme limite en  $+\infty$  si pour tout réel  $\varepsilon > 0$ , il existe un réel  $A$  tel que pour tout  $x \in I$ , si  $x \geq A$ , alors :

$$|f(x) - \ell| \leq \varepsilon.$$

- Soit  $f$  une application d'un intervalle  $I$  dont l'extrémité supérieure est  $+\infty$  dans  $\mathbf{R}$ . On dit que  $f$  admet  $\ell$  comme limite en  $+\infty$  si...

Afin d'unifier toutes ces définitions et de les généraliser encore à des e.v.n. adoptons la définition suivante d'un voisinage de  $\pm\infty$ .

**Définition 2.2.10.** — On appelle voisinage de  $+\infty$  (respectivement  $-\infty$ ) dans  $\mathbf{R}$ , toute partie  $V$  de  $\mathbf{R}$  qui contient un intervalle de la forme  $[A, +\infty[$  (respectivement  $] -\infty, A]$ ) où  $A$  est un réel.

On dit que  $+\infty$  (resp.  $-\infty$ ) est adhérent à une partie  $D$  de  $\mathbf{R}$  si par définition tout voisinage de  $+\infty$ , (resp.  $-\infty$ ) rencontre  $D$ , ou si l'on préfère, si pour tout  $A \in \mathbf{R}$ ,  $D \cap [A, +\infty[ \neq \emptyset$  (resp.  $D \cap ] -\infty, A] \neq \emptyset$ ).

Les trois définitions rappelées sont alors équivalentes à la définition 2.2.1.,

30 secondes de réflexions.....

On peut de ce fait, avec la définition 2.2.9, toujours grâce à 2.2.1. donner la définition de limite dans des situations très générales et diverses

On se place dans le cadre suivant :

On prend  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans un  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . Par  $\vec{a}$  on désigne :

- soit un élément de  $\mathbf{E}$  adhérent à  $D$  ;
- soit, dans le cas où  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  est à  $(\mathbf{R}, |\cdot|)$ ,  $+\infty$  ou  $-\infty$ , supposé adhérent à  $D$  (cf. 2.2.9.).

Par  $\vec{\ell}$  on désigne :

- soit un élément de  $\mathbf{F}$ ,
- soit, dans le cas où  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  et égal à  $(\mathbf{R}, |\cdot|)$ ,  $+\infty$  ou  $-\infty$

Dans tous les cas, on dit que  $\vec{f}$  admet  $\vec{\ell}$  comme limite en  $\vec{a}$  si elle satisfait à la définition 2.2.1.

Explicitons quelques cas :

$f$  EST À VALEURS RÉELLE ET  $\vec{a}$  EST UN POINT DE  $\mathbf{E}$ ,  $\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = -\infty$

Pour tout.....

$\vec{f}$  EST DÉFINIE SUR  $\mathbf{R}$  ET  $\vec{\ell}$  EST UN POINT DE  $\mathbf{F}$ ,  $\lim_{x \rightarrow -\infty} \vec{f}(x) = \vec{\ell}$

Pour tout.....

Reste une dernière généralisation un peu particulière :

**Proposition 2.2.11.** — Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans un  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . On suppose que pour tout réel  $R > 0$ , il existe un élément de  $D$  de norme supérieure ou égale à  $R$ . Soit enfin  $\vec{\ell}$  un élément de  $\mathbf{F}$ . On dit que  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$ , lorsque  $\|\vec{x}\|_{\mathbf{E}}$  tend vers  $+\infty$ , si....

**Proposition 2.2.12.** — Soit  $f$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans  $(\mathbf{R}, |\cdot|)$ . On suppose que pour tout réel  $R > 0$ , il existe un élément de  $D$  de norme supérieure ou égale à  $R$ . On dit que  $f(\vec{x})$  tend vers  $+\infty$

(resp.  $-\infty$ ), lorsque  $\|\vec{x}\|_{\mathbf{E}}$  tend vers  $+\infty$ , si...

Définissons maintenant une notion fondamentale en analyse, celle de continuité.

**Définition 2.2.13.** — Soient deux e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  et  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . Soit  $\vec{f}$  une application d'une partie  $D$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$ ,  $\vec{a}$  un élément de  $D$ . On dit que  $\vec{f}$  est continue en  $\vec{a}$  (pour les normes  $\|\cdot\|_{\mathbf{E}}$  et  $\|\cdot\|_{\mathbf{F}}$ ) si, par définition, pour tout voisinage  $V$  de  $\vec{f}(\vec{a})$ , il existe un voisinage  $U$  de  $\vec{a}$  tel que pour tout élément  $\vec{x}$  de  $D$ , si  $\vec{x}$  appartient à  $U$ , alors  $\vec{f}(\vec{x})$  appartient à  $V$ .

Notons que  $\vec{f}$  doit être définie en  $\vec{a}$  pour pouvoir y être continue. Si  $\vec{f}$  est continue en tout point d'une partie  $A$  de  $D$  on dit que  $\vec{f}$  est continue sur  $A$ , si  $\vec{f}$  est continue sur  $D$  on dit plus simplement que  $\vec{f}$  est continue.

• LIEN ENTRE LIMITE ET CONTINUITÉ—

Dire que  $\vec{f}$  est continue en  $\vec{a}$  signifie donc simplement que  $\vec{f}$  admet une limite en  $\vec{a}$  (qui alors ne peut être que  $\vec{f}(\vec{a})$ ) (cf. 2.2.4.)<sup>1</sup>.

• LIEN ENTRE LIMITE ÉPOINTÉE ET CONTINUITÉ—

Notons pour commencer que  $\vec{f}$  est continue en tout point  $\vec{a}$  de  $D$ , isolé, c'est-à-dire tel qu'il existe un voisinage  $U$  de  $\vec{a}$  tel que  $U \cap D = \{\vec{a}\}$ , en effet on a alors  $\vec{f}(U \cap D) = \{\vec{f}(\vec{a})\}$ , cet ensemble est donc inclus dans tout voisinage  $V$  de  $\vec{f}(\vec{a})$ .

Si  $\vec{a}$  n'est pas un point isolé de  $D$ , alors on a que  $\vec{a}$  est adhérent à  $D - \{\vec{a}\}$  et l'on peut étudier la limite épointée de  $\vec{f}$  en  $\vec{a}$ . On a alors que  $\vec{f}$  est continue en  $\vec{a}$  si et seulement si

$$\vec{f}(\vec{x}) \xrightarrow[\substack{\vec{x} \rightarrow \vec{a} \\ \vec{x} \neq \vec{a}}]{} \vec{f}(\vec{a}).$$

Le lien existant entre la limite et la continuité permet de traduire la continuité de  $\vec{f}$  en utilisant 2.2.2.

Exploitions les résultats 2.2.7–8. pour énoncer des résultats sur les opérations portant sur les fonctions continues.

**Proposition 2.2.14.** — Soient  $\vec{f}$  et  $\vec{g}$  des applications d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . Soit  $\vec{a}$  un élément de  $D$ , et  $\alpha$  et  $\beta$  des éléments de  $\mathbf{K}$ .

1. Supposons que  $\vec{f}$  et  $\vec{g}$  soient continues en  $\vec{a}$  alors  $\alpha\vec{f} + \beta\vec{g}$  est continue en  $\vec{a}$ . De plus si  $f$  et  $g$  sont à valeurs dans  $\mathbf{K}$ ,  $f \times g$  est continue en  $\vec{a}$ .
2. L'ensemble des applications continues de  $D$  dans  $\mathbf{F}$  est un sous-espace vectoriel de l'espace vectoriel des applications de l'ensemble  $D$  dans l'espace vectoriel  $\mathbf{F}$ . On le note  $\mathcal{C}^0(D, \mathbf{F})$ . Pour  $\mathbf{F} = \mathbf{R}$ , L'ensemble des applications continues de  $D$  dans  $\mathbf{K}$ , est une sous-algèbre de l'algèbre des applications de l'ensemble  $D$  dans l'algèbre  $\mathbf{K}$ .

1. On voit donc, de nouveau l'intérêt de la limite épointée lorsque l'on étudie  $\vec{f}$  en un point où elle est définie.

*Preuve de la proposition 2.2.14.* — Le premier point résulte immédiatement de 2.2.7. Passons au second.

**Proposition 2.2.15.** — *La composée de deux applications continues est continue.*

*Preuve de la proposition 2.2.15.* — Elle résulte immédiatement de 2.2.8.

**Remarque 2.2.16.** — Les propositions 2.2.14–15. permettent de prouver la continuité de nombreuses applications.

Par exemple on montre que les applications polynomiales de  $\mathbf{K}^n$ , muni d'une norme quelconque dans  $(\mathbf{K}, |\cdot|)$  sont continues.

L'application  $\mathbf{K}^* \rightarrow \mathbf{K}^*$ ;  $z \mapsto \frac{1}{z}$ , étant continue de  $(\mathbf{K}, |\cdot|)$  dans  $(\mathbf{K}, |\cdot|)$  (cf. sup). Les résultat 2.2.14–15., assure alors la continuité des applications rationnelles de  $n$  variables sur leur domaine de définition.

Voilà un théorème essentiel de caractérisation de la continuité.

**Théorème fondamental 2.2.17.** — *Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . L'application  $\vec{f}$  est continue si et seulement si l'image réciproque de tout fermé de  $\mathbf{F}$  est un fermé relatif de  $D$ , c'est-à-dire l'intersection d'un fermé de  $\mathbf{E}$  et de  $D$ .*

**Corollaire 2.2.18.** — *Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ .  $\vec{f}$  est continue si et seulement si l'image réciproque de tout ouvert de  $\mathbf{F}$  est un ouvert relatif de  $D$ , c'est-à-dire l'intersection d'un ouvert de  $\mathbf{E}$  et de  $D$ .*

*Preuve du corollaire 2.2.18.* — Soit  $F$  un fermé de  $\mathbf{F}$ .

*Preuve de la proposition 2.2.17.* —

- Supposons  $f$  continue.

- Réciproquement, supposons que l'image réciproque de tout fermé de  $\mathbf{F}$  soit un fermé relatif de  $D$ .

Les résultats 2.2.17–18. permettent de montrer que de nombreux ensembles sont ouverts ou fermés.

**Exemple 2.2.19**

1. Soit  $O = \{(x, y) \in \mathbf{R}^2 / x > y, y > x^2\}$  est ouvert dans  $(\mathbf{R}^2, n_2)$ .



En effet

2. Soit  $F = \{(x, y) \in \mathbf{R}^2 / x \geq 3, y + x \leq 1\}$  est fermé dans  $(\mathbf{R}^2, n_2)$ . En effet

Soient  $n$  et  $p$  des entiers strictement positifs. Considérons des applications continues  $f_1, f_2, \dots, f_p$  de  $(\mathbf{K}^n; \|\cdot\|)$  dans  $(\mathbf{R}, |\cdot|)$  (où  $\|\cdot\|$  est une norme sur  $\mathbf{K}^n$ ) et  $a_1, b_1, a_2, b_2, \dots, a_p, b_p$  des réels alors l'ensemble  $U$ , défini par

$$U := \{(x_1, x_2 \dots x_n) \in \mathbf{K}^n / a_i < f_i(x_1, x_2 \dots x_n) < b_i, i = 1, 2, \dots, p\},$$

est un ouvert de  $(\mathbf{K}^n; \|\cdot\|)$ . En effet

3. Avec les mêmes notations, on montre que l'ensemble  $G$ , défini par

$$G := \{(x_1, x_2 \dots x_n) \in \mathbf{K}^n / a_i \leq f_i(x_1, x_2 \dots x_n) \leq b_i, i = 1, 2, \dots, p\},$$

est un fermé de  $(\mathbf{K}^n; \|\cdot\|)$ .

4. On munit  $\mathcal{M}_n(\mathbf{R})$  de la norme  $n_\infty$  définie par,

$$n_\infty \left( (m_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,n}} \right) = \sup_{\substack{i=1,\dots,n \\ j=1,\dots,n}} |m_{i,j}|,$$

pour tout élément  $(m_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$  de  $\mathcal{M}_n(\mathbf{R})$ . Montrer que  $\mathrm{GL}_n(\mathbf{R})$  est un ouvert de  $\mathcal{M}_n(\mathbf{R})$ .

5. Montrer que la sphère unité de  $\mathbf{R}^n$  est un .....de  $(\mathbf{R}^n, n_\infty)$ .

Ce dernier exemple est un cas particulier du résultat suivant :

**Corollaire 2.2.20.** — *L'ensemble des « zéros » d'une application  $\vec{f}$  d'une partie  $D$  d'un e.v.n  $(\mathbf{E}, \|\cdot\|)$ , à valeurs dans un e.v.n  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , continue, est un fermé de  $D$ .*

**Preuve du corollaire 2.2.21.** —

**Corollaire (du corollaire) 2.2.21.** — *Soient  $\vec{f}$  et  $\vec{g}$  des applications d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ , à valeurs dans un e.v.n  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  continues. Soient  $A$  une partie dense dans  $D$ . Si  $\vec{f}$  et  $\vec{g}$  coïncident sur  $A$  alors elles sont égales.*

*Preuve du corollaire 2.2.21.* —

**Exercice d'application** — Déterminer les applications  $f$  de  $(\mathbf{R}, |\cdot|)$  dans  $(\mathbf{R}, |\cdot|)$ , continues, telles que pour tous réels  $x$  et  $y$ ,

$$f(x+y) = f(x) + f(y).$$

*Solution* —

**Proposition-définition 2.2.22.** — Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  continue. Soit  $\vec{a}$  un élément adhérent à  $D$ , mais non élément de  $D$ .

— Si  $\vec{f}$  admet en  $\vec{a}$ , une limite  $\vec{\ell}$ , alors il existe une unique application continue  $\vec{f}$  qui prolonge  $\vec{f}$ , à  $D \cup \{\vec{a}\}$ , elle est définie par :

$$\vec{f} : D \cup \{\vec{a}\} \rightarrow \mathbf{F}; \vec{x} \mapsto \begin{cases} \vec{f}(\vec{x}), & \text{pour } \vec{x} \neq \vec{a}, \\ \vec{\ell}, & \text{pour } \vec{x} = \vec{a}. \end{cases}$$

— Si  $\vec{f}$  n'admet pas en  $\vec{a}$  de limite, alors il n'existe aucune application continue qui prolonge  $\vec{f}$  à  $D \cup \{\vec{a}\}$ .  
—  $\vec{f}$  s'appelle prolongement par continuité de  $\vec{f}$  à  $D \cup \{\vec{a}\}$ .

*Preuve de la proposition 2.2.22. —*

**Proposition 2.2.23.** — Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , continue. Soit  $D'$  une partie de  $D$  alors la restriction de  $\vec{f}$  à  $D'$  est continue.

*Preuve de la proposition 2.2.23. —*

**Remarque 2.2.24.** — La réciproque de la proposition 2.2.20. est fausse. La restriction de  $\vec{f}$  à une partie  $D'$  peut être continue sans que pour autant  $\vec{f}$  soit continue en tout point de  $D'$ . Donnons des exemples.

Par contre on peut montrer :

**Exercice** — Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  Soit  $D'$

une **Ouvert** relatif de  $D$ . On suppose que la restriction de  $\vec{f}$  à  $D'$  est continue. Montrer que  $f$  est continue en tout point  $D'$ .

**Solution** —

Nous allons maintenant donner les définitions, pour des applications, des relations de domination de prépondérance et d'équivalence vues pour les suites. soit  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  et  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  des e.v.n.,  $\vec{a}$  désigne un élément de  $\mathbf{E}$ , ou si  $(\mathbf{E}, \|\cdot\|) = (\mathbf{R}, |\cdot|)$ ,  $a$  peut désigner aussi  $+\infty$  ou  $-\infty$ . Si l'on est dans le premier cas,  $P$  désigne, une partie de  $\mathbf{E}$ , à laquelle  $\vec{a}$  est adhérent, si  $a = +\infty$  (resp.  $-\infty$ ),  $P$  désigne  $\mathbf{R}$  entier ou une partie qui rencontre tout intervalle de la forme  $]c, +\infty[$  (resp.  $] - \infty, c[$ ), avec  $c$  réel.

Notons ici  $\mathcal{F}_{\vec{a}}(\mathbf{F})$  l'ensemble des fonctions  $f$  à valeurs dans  $\mathbf{F}$ , pour lesquelles il existe un voisinage  $V$  de  $\vec{a}$ , tel que  $\vec{f}$  soit définie sur  $V \cap P$ . On parle de fonctions à valeurs dans  $\mathbf{F}$  définies au voisinage de  $\vec{a}$  sur  $P$ . Pour  $(\mathbf{E}, \|\cdot\|) = (\mathbf{R}, |\cdot|)$  et  $a$  réel on prend souvent pour  $P$ ,  $]a, +\infty[$  (resp.  $] - \infty, a[$ ), on dit que les éléments de  $\mathcal{F}_a(\mathbf{F})$  sont les fonctions à valeurs dans  $\mathbf{F}$ , définies au voisinage de  $a$  à droite (resp. à gauche).

**Définition 2.2.24** — Soient  $\vec{f}$  un élément de  $\mathcal{F}_{\vec{a}}(\mathbf{F})$  et  $g$  un élément de  $\mathcal{F}_{\vec{a}}(\mathbf{R})$ , on dit que  $\vec{f}$  est négligeable devant  $g$  au voisinage de  $\vec{a}$  selon  $P$ , où que la fonction  $g$  est prépondérante devant  $\vec{f}$  au voisinage de  $\vec{a}$  selon  $P$  si, pour tout réel  $\varepsilon$  strictement positif, il existe un voisinage  $U$  de  $\vec{a}$  tel que pour tout  $\vec{x}$  élément de  $U \cap P$ , on ait :

$$\|\vec{f}(\vec{x})\| \leq \varepsilon |g(\vec{x})|.$$

On note  $\vec{f}(\vec{x}) = o_{\vec{x} \rightarrow \vec{a}}(g(\vec{x}))$  ou  $\vec{f}(\vec{x}) = o(g(\vec{x}))$ , au voisinage de  $\vec{a}$  selon  $P$ .

Bien sur si  $a$  est un réel pour  $P$  égal à  $]a, +\infty[$  (resp.  $] - \infty, a[$ ) on parle de négligeable au voisinage de  $a$ , à droite (resp. à gauche). Pour  $P = \mathbf{E}$ , la précision « selon  $P$  » est inutile.

**Remarque 2.2.25.** — Observons que dans le cas où il existe un voisinage  $U$  de  $\vec{a}$  tel que  $g$  ne s'annule pas sur  $U \cap P$ , alors  $\frac{1}{g}\vec{f}$  est définie au moins sur  $U \cap P$ ; et  $\vec{f}$  est négligeable devant  $g$  au voisinage de  $\vec{a}$  selon  $P$  si et seulement si  $\frac{1}{g}\vec{f}$  tend vers zéro lorsque  $\vec{x}$  tend vers  $\vec{a}$  selon  $U \cap P$ .

**Définition 2.2.27.** — Soit  $\vec{f}$  un élément de  $\mathcal{F}_{\vec{a}}(\mathbf{F})$  et  $g$  un élément de  $\mathcal{F}_{\vec{a}}(\mathbf{R})$  on dit que  $\vec{f}$  est dominée par  $g$  au voisinage de  $\vec{a}$  selon  $P$ , si il existe un réel  $M$  et un voisinage  $U$  de  $\vec{a}$  tel que pour tout  $\vec{x}$  élément de  $U \cap P$  on ait :

$$\|\vec{f}(\vec{x})\| \leq M |g(\vec{x})|.$$

On note :  $\vec{f}(\vec{x}) = O_{\vec{x} \rightarrow \vec{a}}(g(\vec{x}))$  ou  $\vec{f}(\vec{x}) = O(g(\vec{x}))$ , au voisinage de  $\vec{a}$  selon  $P$ .

Bien sur si  $a$  est un réel pour  $P$  égal à  $]a, +\infty[$  (resp.  $] - \infty, a[$ ) on parle de domination au voisinage de  $a$ , à droite (resp. à gauche). Pour  $P = \mathbf{E}$ , la précision « selon  $P$  » est inutile.

**Remarque 2.2.30.** — Observons que dans le cas où il existe un voisinage  $U$  de  $\vec{a}$  tel que  $g$  ne s'annule pas sur  $U \cap P$ , alors  $\frac{1}{g}\vec{f}$  est définie au moins sur  $U \cap P$ ; et  $\vec{f}$  est dominée par  $g$  au voisinage de  $\vec{a}$  selon  $P$  si et seulement si il existe un voisinage  $U'$  de  $\vec{a}$  tel que  $\frac{1}{g}\vec{f}$  soit définie et bornée sur  $U' \cap P$ .

**Définition 2.2.31.** — Soit  $\vec{f}$  et  $\vec{g}$  des éléments de  $\mathcal{F}_{\vec{a}}(\mathbf{F})$ , on dit que  $\vec{f}$  est équivalente à  $\vec{g}$  au voisinage de  $\vec{a}$  selon  $P$ , si,

$$\vec{f}(\vec{x}) - \vec{g}(\vec{x}) = \underset{\substack{\vec{x} \rightarrow \vec{a} \\ \vec{x} \in P}}{o}(\|\vec{g}(\vec{x})\|_{\mathbf{F}}).$$

On note  $\vec{f}(\vec{x}) \underset{\substack{\vec{x} \rightarrow \vec{a} \\ \vec{x} \in P}}{\sim} \vec{g}(\vec{x})$  ou  $\vec{f}(\vec{x}) \sim \vec{g}(\vec{x})$ , au voisinage de  $\vec{a}$  selon  $P$ .

**Remarque 2.2.32.** — Observons que dans le cas  $f$  et  $g$  sont à valeurs dans  $\mathbf{R}$  et où il existe un voisinage  $U$  de  $a$  tel que  $g$  ne s'annule pas sur  $U \cap P$ , alors  $\frac{f}{g}$  est définie au moins sur  $U \cap P$  et  $f$  est équivalent à  $g$  au voisinage de  $a$  selon  $P$ , si et seulement si

$$\frac{f(x)}{g(x)} \underset{\substack{x \rightarrow a \\ x \in P}}{\rightarrow} 1.$$

On montre sans mal :

**Proposition 2.2.33** — la relation sur  $\mathcal{F}_{\vec{a}}(\mathbf{F})$ , « être équivalente au voisinage de  $\vec{a}$  selon  $P$  à » est une relation d'équivalence.

## 2.3 Continuité uniforme

**Définition 2.3.1.** — soit Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . On dit que  $\vec{f}$  est uniformément continue, si, par définition, pour tout réel  $\varepsilon$ , strictement positif, il existe un réel  $\eta$  strictement positif tel que, pour tout  $\vec{x}$  et tout  $\vec{y}$  élément de  $D$ , si  $\|\vec{x} - \vec{y}\|_{\mathbf{E}} \leq \eta$  alors  $\|\vec{f}(\vec{x}) - \vec{f}(\vec{y})\|_{\mathbf{F}} \leq \varepsilon$ . On dit que  $\vec{f}$  est uniformément continue sur une partie  $A$  de  $D$  si, par définition,  $\vec{f}|_A$  restriction de  $\vec{f}$  à  $A$  est uniformément continue.

**Remarques** —

- La notion de continuité uniforme, contrairement à la continuité, est une notion *globale* : on ne peut pas parler de continuité uniforme en un point
- On peut dans la définition remplacer les inégalités strictes par des inégalités larges (cf. continuité).

Gardons les notation de 2.1.1., et comparons les propriétés :  $\vec{f}$  est continue et  $\vec{f}$  est uniformément continue.

$\vec{f}$  est continue s'écrit :

Pour tout  $\vec{x} \in D$  et pour tout réel  $\varepsilon > 0$ , il existe un réel  $\eta > 0$  tel que pour tout  $\vec{y}$  élément de  $D$ , si  $\|\vec{x} - \vec{y}\|_{\mathbf{E}} \leq \eta$  alors  $\|\vec{f}(\vec{x}) - \vec{f}(\vec{y})\|_{\mathbf{F}} \leq \varepsilon$ .

La continuité uniforme donne donc clairement la continuité. En revanche la continuité ne semble pas suffire à assurer la continuité uniforme ; en effet

Illustrons par un exemple.

**Exemple 2.3.2.** — Soit l'application

$$f_1 : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto x^2.$$

Donnons, pour justifier notre intuition, un critère de non continuité uniforme.

**Proposition 2.3.3.** — CRITÈRE DE NON CONTINUITÉ UNIFORME —

*Sous les hypothèses de 2.3.1.,  $\vec{f}$  n'est pas uniformément continue, si et seulement si, il existe deux suites  $(\vec{x}_n)_{n \in \mathbf{N}}$  et  $(\vec{y}_n)_{n \in \mathbf{N}}$  à valeurs dans  $D$ , telles que :*

$$1. \quad \vec{x}_n - \vec{y}_n \underset{n \rightarrow +\infty}{\rightarrow} 0_{\mathbf{E}}.$$

$$2. \quad \vec{f}(\vec{x}_n) - \vec{f}(\vec{y}_n) \text{ ne tend pas vers } \vec{0}_{\mathbf{F}}, \text{ lorsque } n \text{ tend vers } +\infty.$$

*preuve de la proposition 2.3.3. —*

**Remarque 2.3.4.** — On a en fait prouvé plus que la proposition 2.3.3., On peut en effet y remplacer 2. par :

2. **bis.** Il existe  $\varepsilon_0 > 0$  tel que, pour tout  $n \in \mathbf{N}$ ,  $\|\vec{f}(\vec{x}_n) - \vec{f}(\vec{x}_n)\| \geq \varepsilon_0$ .

Appliquons ce critère pour prouver que l'application  $f_1$  de l'exemple 2.3.2. n'est pas uniformément continue.

Traisons un autre exemple



**Exemple 2.3.4.** — Soit l'application

$$f_2 : \mathbf{R}_+^* \rightarrow \mathbf{R}; x \mapsto \frac{1}{x^2}.$$

**Définition 2.3.5.** — Soient Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  et  $k$  un réel positif ou nul. On dit que  $\vec{f}$  est lipschitzienne de constante  $k$ , ou  $k$ -lipschitzienne, si, par définition, pour tout  $\vec{x}$  et tout  $\vec{y}$  éléments de  $D$ ,

$$\|\vec{f}(\vec{x}) - \vec{f}(\vec{y})\|_{\mathbf{F}} \leq k \|\vec{x} - \vec{y}\|_{\mathbf{E}}.$$

Si de plus  $0 \leq k < 1$ , on dit que  $\vec{f}$  est  $k$ -contractante.

**Exemples** —

1. D'après l'inégalité des accroissements finis, une application d'un intervalle  $I$  de  $(\mathbf{R}, |\cdot|)$  à valeurs dans  $(\mathbf{R}, |\cdot|)$ , dérivable et telle que pour tout  $x \in I$ ,  $|f'(x)| \leq k$  est  $k$ -lipschitzienne.
2. Soit  $(\mathbf{E}, \|\cdot\|)$  un espace vectoriel normé. La norme  $\|\cdot\|$  est une application de  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans  $(\mathbf{R}, |\cdot|)$ , 1-lipschitzienne, cf. 1.1.4.
3. Soit  $A$  une partie non vide d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . L'application  $d$  de  $(\mathbf{E}, \|\cdot\|)$  dans  $(\mathbf{R}, |\cdot|)$ , qui à un élément  $\vec{x}$  de  $\mathbf{E}$  associe sa distance à  $A$  est 1-lipschitzienne, cf. 1.1.14.

On a alors le résultat évident

**Proposition 2.3.6.** — Une application  $k$ -lipschitzienne est uniformément continue.

On verra que la réciproque est fautive : il existe des applications uniformément continues non lipschitzienne.

## COMPACTITÉ

### 3.1 Généralité

Dans toute ce paragraphe  $(\mathbf{E}, \|\cdot\|)$  désigne un espace vectoriel normé sur le corps  $\mathbf{K}$ . (rappelons que  $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{K} = \mathbf{C}$ ).

**Définition 3.1.1.** — Soit  $K$  une partie de  $(\mathbf{E}, \|\cdot\|)$ . On dit que  $K$  est une partie compacte ou un compact de  $(\mathbf{E}, \|\cdot\|)$  si, par définition toute suite à valeurs dans  $K$  admet une sous-suite qui converge vers un élément de  $K$ .

**Remarque** — Compte tenu de la proposition 2.1.19., On peut dire que  $K$  est compact si et seulement si toute suite à valeurs dans  $K$  admet une valeur d'adhérence élément de  $K$ .

**Exemples** —

1. L'ensemble vide est un compact de  $(\mathbf{E}, \|\cdot\|)$ .
2. Toute partie finie de  $\mathbf{E}$  est un compact de  $(\mathbf{E}, \|\cdot\|)$ .

**Proposition 3.1.2.** — Toute partie compacte de  $(\mathbf{E}, \|\cdot\|)$  est fermée et bornée.

*Preuve de la proposition 3.1.2.* —

Attention la réciproque de 3.1.2. est fausse. Il existe des parties fermées bornées non compactes. Toutefois nous verrons, que si  $\mathbf{E}$  est de dimension finie, les compacts se confondent avec les parties fermées bornées.

On a vu qu'une suite qui converge admet une et une seule valeur d'adhérence (2.1.18.), mais que la réciproque est fausse. Dans le cas d'une suite à valeur dans un compact la réciproque est vraie :

**Proposition 3.1.3.** — *Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $K$ , compact d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ . la suite converge si et seulement si elle admet une et une seule valeur d'adhérence.*

*Preuve de la proposition 3.1.3. —*

- Que, si  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge alors elle admet une et une seule valeur d'adhérence, a été prouvé en 2.1.18., comme nous venons de le rappeler.
- Supposons que  $(\vec{x}_n)_{n \in \mathbf{N}}$  admette une et une seule valeur d'adhérence  $\vec{\ell}$ . Supposons que  $(\vec{x}_n)_{n \in \mathbf{N}}$  ne converge pas vers  $\vec{\ell}$ . Alors il existe un réel strictement positif,  $\varepsilon_0$  tel que.....

Voici maintenant deux lemmes techniques au programme fort utiles pour prouver que des parties sont compactes.

**Lemme 3.1.4.** — *Soit  $K$  une partie compact de  $(\mathbf{E}, \|\cdot\|)$ . Toute partie  $A$  fermée incluse dans  $K$  est compact.*

*Preuve du lemme 3.1.4. —*

**Lemme 3.1.5.** — *Soit  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  et  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$  des espaces vectoriels normés. On munit  $\mathbf{F} \times \mathbf{G}$  de la norme  $n'_{\infty}$  (cf. 1.1.7.3.(a)),*

$$n'_{\infty} : \mathbf{F} \times \mathbf{G} \rightarrow \mathbf{R}; (\vec{x}, \vec{y}) \mapsto \sup(\|\vec{x}\|_{\mathbf{F}}, \|\vec{y}\|_{\mathbf{G}}).$$

*Soient  $K$  un compact de  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  et  $H$  un compact de  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$ . Alors  $K \times H$  est un compact de  $(\mathbf{F} \times \mathbf{G}, n'_{\infty})$ .*

*Brutalement : le produit de deux compacts est compact.*

*Preuve du lemme 3.1.5. —*

### 3.2 Compacts de $(\mathbf{K}^n, n_\infty)$

Commençons par étudier les compacts de  $(\mathbf{R}, |\cdot|)$ .

**Proposition 3.2.1.** — *Tout segment de  $(\mathbf{R}, |\cdot|)$  est compact.*

**Preuve de la proposition 3.2.1.** —

**Corollaire 3.2.2.** — *Toute partie fermée bornée de  $(\mathbf{R}, |\cdot|)$  est compacte.*

**Preuve du corollaire 3.2.2.** —

Passons au cas général.

**Théorème 3.2.3.** — *Les parties compactes de  $(\mathbf{K}^n, n_\infty)$  sont exactement les parties fermées bornées.*

*Preuve du théorème 3.2.3.* — Nous avons vu au 3.1.2. que tout compact est fermé borné. Il nous suffit donc de montrer que toute partie fermée bornée est compacte.

- PREMIER CAS  $\mathbf{K} = \mathbf{R}$

- SECOND CAS  $\mathbf{K} = \mathbf{C}$

**Corollaire 3.2.4.** — BOLZANO WEIERSTRASS —

*Toute suite bornée de  $(\mathbf{K}^n, n_\infty)$  admet une sous-suite convergente.*

### 3.3 Compacité et continuité

**Proposition 3.3.1.** — *Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , continue. Soit  $K$  une partie de  $D$ .*

*Si  $K$  est compact alors  $\vec{f}(K)$  est compact. Autrement dit, l'image directe d'un compact par une application continue est compacte.*

**Remarque** — L'image réciproque d'un compact par une application continue, n'est pas en général compact. Prenons

par exemple ...

*Preuve de la proposition 3.3.1.* — Supposons que  $K$  soit un compact. Montrons que  $\vec{f}(K)$  est compact.

**Corollaire 3.3.2.** — Soit  $f$  une application continue d'une partie compacte  $K$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans l'e.v.n.  $(\mathbf{R}, |\cdot|)$ . Alors :

- L'application  $f$  est bornée.
- Si de plus,  $K$  est non vide, alors il existe  $\vec{x}_M$  et  $\vec{x}_m$  éléments de  $K$ , tels que :

$$f(\vec{x}_M) = \sup_{\vec{x} \in K} f(\vec{x}); \quad f(\vec{x}_m) = \inf_{\vec{x} \in K} f(\vec{x}),$$

ou, comme on dit, la borne supérieure et la borne inférieure sont atteintes.

On en déduit le résultat plus général suivant :

**Corollaire (du corollaire) 3.3.3.** — Soit  $f$  une application continue d'une partie compacte  $K$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . Alors :

- L'application  $\vec{f}$  est bornée.
- Si de plus,  $K$  est non vide, alors il existe  $\vec{x}_M$  et  $\vec{x}_m$  éléments de  $K$ , tels que :

$$\|\vec{f}(\vec{x}_M)\|_{\mathbf{F}} = \sup_{\vec{x} \in K} \|\vec{f}(\vec{x})\|_{\mathbf{F}}; \quad \|\vec{f}(\vec{x}_m)\|_{\mathbf{F}} = \inf_{\vec{x} \in K} \|\vec{f}(\vec{x})\|_{\mathbf{F}}.$$

*Preuve du corollaire 3.3.3. —*

Prouvons maintenant le corollaire 3.3.2.

*Preuve du corollaire 3.3.2. —*

**Application 3.3.4.** — Soit  $K$  un compact d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  et  $\vec{a}$  un élément de  $\mathbf{E}$ . Montrons qu'il existe un



élément  $\vec{x}_0$  de  $K$  tel que  $\|\vec{x}_0 - \vec{a}\| = d(\vec{a}, K)$ .

Le théorème de Heine, vu en MPSI pour  $(\mathbf{R}, |\cdot|)$ , se généralise à des e.v.n. quelconques :

**Proposition 3.3.5.** — THÉORÈME DE HEINE —

*Toute application  $\vec{f}$  d'une partie compacte  $K$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  à valeurs dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , continue est uniformément continue.*

*Preuve de la proposition 3.3.5. —*

Donnons en exercice quelques applications du théorème de Heine.

**Exercice 3.3.6.** — Montrer que l'application  $f : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto \sqrt{x}$  est uniformément continue.

*Solution de l'exercice 3.3.6. —*

Voilà, comme promis, une application uniformément continue, non lipschitzienne.

**Exercice 3.3.7. —**

1. Montrer qu'une application  $f$  continue de  $(\mathbf{R}, |\cdot|)$  dans  $(\mathbf{R}, |\cdot|)$  qui admet une limite  $\ell$  en  $+\infty$  et  $\ell'$  en  $-\infty$  est uniformément continue.
2. Montrer qu'une application  $g$  continue de  $(\mathbf{R}, |\cdot|)$  dans  $(\mathbf{R}, |\cdot|)$  et  $T$ -périodique ( $T > 0$ ) est uniformément continue.

*Solution de l'exercice 3.3.7. —*

## CONNEXITÉ PAR ARCS

Nous allons modéliser la notion de partie en « un seul morceau », dites *connexes par arcs*. Elle joue un rôle crucial en analyse.

Dans tout cette partie, on se donne un e.v.n.  $(\mathbf{E}, \|\cdot\|)$  sur le corps  $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{C}$ .

### 4.1 Définition, premières propriétés

Commençons par une définition.

**Définition 4.1.1.** On appelle chemin de  $\mathbf{E}$ , toute application de  $[0, 1]^1$  dans  $\mathbf{E}$ , continue.

L'ensemble  $\vec{\varphi}([0, 1])$  est alors appelé trajectoire ou support du chemin,  $\vec{\varphi}(0)$  est appelé l'extrémité initiale du chemin,  $\vec{\varphi}(1)$  l'extrémité finale.

Un chemin ayant pour extrémité initiale un point  $\vec{a}$  de  $\mathbf{E}$ , pour extrémité finale un point  $\vec{b}$  de  $\mathbf{E}$ , est dit joindre les points  $\vec{a}$  et  $\vec{b}$ .

! Nous avons fait le choix de convenir qu'un chemin est automatiquement continue. Ce choix n'est pas universel, il convient donc à l'écrit ou à l'oral de bien préciser la chose.

**Définition 4.1.2.** — une partie  $A$  de  $(\mathbf{E}, \|\cdot\|)$  est dite connexe par arcs si par définition, pour tout couple de points  $(\vec{a}, \vec{b})$  de  $A$  il existe un chemin de support inclus dans  $A$  et joignant  $\vec{a}$  à  $\vec{b}$ .

Interprétons dans la définition d'un chemin  $\vec{\varphi}$  le rôle de la continuité. La continuité de  $\vec{\varphi}$  dit que  $\vec{\varphi}(t)$  est aussi près de  $\vec{\varphi}(t_0)$  que l'on veut, pourvu que  $t$  soit suffisamment proche de  $t_0$ . Autrement dit le support du chemin  $\vec{\varphi}$  ne présente pas de « sauts ». Si l'on préfère, dans le cas où  $\mathbf{E}$  est un plan on peut tracer la trajectoire de  $\vec{\varphi}$  sans lever la main.

La définition de la connexité par arcs est donc conforme à notre intuition : une partie  $A$  est en un seul morceau, si l'on peut se rendre de n'importe quel point de  $A$  à n'importe quel point  $\vec{b}$  de  $A$  sans avoir à faire de « saut ».

**Exemples niais** —

- L'ensemble vide est connexe par arcs.
- L'ensemble  $\mathbf{E}$  est connexe par arcs.
- Un singleton de  $\mathbf{E}$  est connexe par arcs.

Étudions des notions plus fortes que la connexité par arcs qui fournissent des exemples de parties connexes par arcs.

**Définition 4.1.3** — Soient  $\vec{a}$  et  $\vec{b}$  des éléments de  $\mathbf{E}$ . On appelle segment d'extrémités  $\vec{a}$  et  $\vec{b}$ , noté  $[\vec{a}, \vec{b}]$ , la partie de  $\mathbf{E}$  :

$$\left\{ \vec{a} + \lambda(\vec{b} - \vec{a}), \lambda \in [0, 1] \right\}.$$

Donnons nous  $\vec{a}$  et  $\vec{b}$  et remarquons que l'on a immédiatement :

$$[\vec{a}, \vec{b}] = \{ \lambda \vec{b} + (1 - \lambda) \vec{a}, \lambda \in [0, 1] \} = \{ t \vec{a} + (1 - t) \vec{b}, t \in [0, 1] \}.$$

On a donc  $[\vec{a}, \vec{b}] = [\vec{b}, \vec{a}]$ .

1. Plus généralement on parle encore de chemin lorsque l'application est définie sur un segment  $[\alpha, \beta]$  d'intérieur non vide.

Le segment  $[\vec{a}, \vec{b}]$  apparaît comme le support du chemin

$$[0, 1] \rightarrow \mathbf{E}; \lambda \mapsto \lambda \vec{b} + (1 - \lambda) \vec{a}, \quad (\text{III.6})$$

(extrémité initiale ....., extrémité finale .....), ou du chemin

$$[0, 1] \rightarrow \mathbf{E}; t \mapsto t \vec{a} + (1 - t) \vec{b}. \quad (\text{III.7})$$

(extrémité initiale ....., extrémité finale .....).

Lorsque l'on considère  $\mathbf{E}$  comme affine ou que l'on travaille dans un sous-espace affine de  $\mathbf{E}$ , on a avec des notations plus adaptées à cette situation :

$$[p, q] = p + [0, 1] \vec{pq}$$

pour tout  $p$  et tout  $q$ , points de  $\mathbf{E}$ ,

Le segment  $[p, q]$  est alors une partie de la droite affine  $p + \mathbf{R} \vec{pq}$ , avec des notions oubliées de géométrie élémentaire,  $[p, q]$  est l'ensemble des barycentres des points  $p$  et  $q$  affectés de coefficients positifs. On se reportera à l'annexe pour une définition précise des barycentres.

**Définition 4.1.4.** — Une partie  $A$  de  $E$  est dite convexe, si, par définition, pour tout couple  $(\vec{a}, \vec{b})$  d'éléments de  $A$ , le segment  $[\vec{a}, \vec{b}]$  est inclus dans  $A$ .

**Exemples 4.1.5.** —

- Tout sous-espace vectoriel de  $(\mathbf{E}, \|\cdot\|)$  est convexe.
- Tout sous-espace affine de  $(\mathbf{E}, \|\cdot\|)$  est convexe.
- Pour  $\mathbf{E}$  de dimension 2, tout demi-plan affine de  $(\mathbf{E}, \|\cdot\|)$  est convexe.
- Toute boule ouverte ou fermée de  $(\mathbf{E}, \|\cdot\|)$  est convexe.
- Si  $\mathbf{E}$  n'est pas nul, alors sphère de  $\mathbf{E}$  n'est pas convexe.

Autrement dit une partie  $A$  de  $\mathbf{E}$ , vu comme un espace affine, est convexe si les barycentres à coefficients positifs de deux de ses points est encore élément de  $A$ . Cette propriété qui définit un convexe peut se généraliser.

**Proposition 4.1.6.** — *Une partie  $A$  de  $\mathbf{E}$  (vu comme un espace affine), est convexe si et seulement si les barycentres à coefficients positifs d'un nombre quelconque de points de  $A$  sont éléments de  $A$ .*

*Preuve de 4.1.6.* — La condition est évidemment suffisante et généralise la définition exprimée en terme de barycentre. Prouvons sa suffisance par récurrence.

Soit  $A$  un convexe de  $\mathbf{E}$ ,

- La partie  $A$  est stable par passage au barycentre à coefficients positifs de deux de ses points<sup>1</sup>, par définition de la convexité.
- Soit un entier  $n \geq 2$ . On suppose  $A$  stable par passage au barycentre à coefficients positif de  $n$  de ses points...

Donc on a montré par récurrence la stabilité de  $A$  par passage au barycentre à coefficients positifs d'un nombre quelconque de ses points.

D'où l'équivalence annoncée



**Exercice** Soit  $A$  une partie de  $\mathbf{E}$  vu comme un espace affine. On note  $\text{Conv}(A)$  l'ensemble des barycentres d'un nombre quelconque de points de  $A$  affectés de coefficients positifs. Montrer que  $\text{Conv}(A)$  est le plus petit convexe qui contienne  $A$ .

**Proposition 4.1.7.** - - *Toute partie convexe  $A$  de  $\mathbf{E}$  est un connexe par arcs de  $(\mathbf{E}, \|\cdot\|)$ .*

*Preuve de la proposition 4.1.7.* —

**Proposition 4.1.8.** — *Une partie  $A$  de  $\mathbf{E}$  est dite étoilée s'il existe un élément  $\vec{a}$  de  $A$  tel que pour tout élément  $\vec{x}$  de  $A$ , le segment  $[\vec{a}, \vec{x}] \subset A$ .*

---

1. On aurait pu, bien sûr, initialiser à 1, mais la notion de barycentre d'un point n'est pas une notion très exaltante!

On dit alors que  $A$  est étoilé par rapport à  $\vec{a}$ .

### Exemples —

- Tout convexe de  $\mathbf{E}$  est étoilé (par rapport à chacun de ses points).
- La réunion de 2 droites affines distinctes de  $\mathbf{E}$  s'intersectant en  $\vec{a}$  est étoilée mais non convexe.
- Pour  $\mathbf{E} = \mathbf{R}^2$ , posons  $L = D_f((0, 0), 1) - D_f((1, 0), 1)$ .

- Pour  $\mathbf{E} = \mathbf{R}^2$ ,  $\mathbf{R}^2 \setminus \{(0, 0)\}$  .....

**Proposition 4.1.9.** — *Toute partie étoilée de  $\mathbf{E}$  est un connexe par arcs de  $(\mathbf{E}, \|\cdot\|)$ .*

*Preuve de la proposition 4.1.9. —*

Donnons à présent un exemple non totalement trivial de partie connexe par arcs, qui n'est pas étoilée.

On suppose que  $\mathbf{E}$  est un espace vectoriel de dimension finie  $n \geq \dots$ . Soit  $S$  la sphère unité de  $(\mathbf{E}, \|\cdot\|)$ .

- La partie  $S$  de  $\mathbf{E}$  n'est pas étoilée
- Montrons que  $S$  est un connexe par arcs de  $(\mathbf{E}, \|\cdot\|)$ .



On notera l'importance de faire un dessin dans la preuve de la connexité par arcs. Du reste dans les cas simples le programme suggère d'accepter comme preuve un dessin clair.

## 4.2 Composantes connexes par arcs

Dit grossièrement, nous allons maintenant pour des parties de  $\mathbf{E}$  quelconques, nous intéresser aux morceaux qui la constituent.

Dans toute cette sous-partie  $A$  est une partie non vide d'un e.v.n.  $(\mathbf{E}, \|\cdot\|)$ .

**Définition 4.2.1.** — On définit sur  $A$  une relation  $\mathcal{R}_A$  de la façon suivante : des éléments  $\vec{a}$  et  $\vec{b}$  sont en relation par  $\mathcal{R}$ , si il existe un chemin d'extrémité initiale  $\vec{a}$  et d'extrémité finale  $\vec{b}$  dont le support est inclus dans  $A$ .

**Proposition 4.2.2.** — La relation  $\mathcal{R}_A$  définie en 4.2.1. est une relation d'équivalence.

*Preuve de la proposition 4.2.2.* —

D'où le résultat. □

**Proposition 4.2.3.** — On appelle composantes connexes par arcs de  $A$  toute classe d'équivalence pour la relation  $\mathcal{R}_A$

Il résulte des propriétés des relation d'équivalence que :

- les composantes connexes de  $A$  (non vide), sont non vides,
- deux composantes connexes distinctes sont d'intersection vide,
- la réunion de toutes les composante connexe de  $A$  est  $A$ .

On traduit ces propriétés en disant que l'ensemble des classes d'équivalences de  $A$  forme une partition de  $A$ .

**Cas du vide :** On convient que l'ensemble vide admet une et une seule composante connexe l'ensemble vide !

**Proposition 4.2.4.** — Soit  $\vec{a}$  un élément de  $A$ . La composante connexe par arcs de  $\vec{a}$  est le plus grand (pour l'inclusion) des connexes par arcs inclus dans  $A$  qui contiennent  $\vec{a}$ .

Cette proposition reste vraie pour l'ensemble vide avec la convention faite au dessus.

*Preuve de la proposition 4.2.4.* —

- Montrons que  $C_{\vec{a}}$ , la composante connexe par arcs de  $\vec{a}$ , est connexe par arcs et inclus dans  $A$ .

- Soit  $C'$  un connexe par arcs inclus dans  $A$  qui contient  $\vec{a}$ . Montrons que :  $C' \subset C_{\vec{a}}$ .

Donc on a montré que  $C_{\vec{a}}$  est le plus grand des connexes par arcs qui contiennent  $\vec{a}$ . □

La partie  $A$  est on l'a vu donc réunion de ses composantes connexes par arcs, qui sont donc des parties de  $A$  connexes par arcs maximales.

**Exercice**— Déterminer sans donner de preuve les composantes connexes par arcs de  $A$  dans les cas suivants :

1. Dans  $(\mathbf{R}^2, n_2)$ ,  $A = \mathbf{R}^2 - C((0, 0), 1)$ .
2. Dans  $(\mathbf{R}^2, n_2)$ ,  $A = \{(x, y) \in \mathbf{R}^2, xy = 1\}$ .
3. Dans  $(\mathbf{R}, | \cdot |)$   $A = \mathbf{N}$ .

*Solution* —

### 4.3 Théorème de la valeur intermédiaire

Nous allons un instant nous placer dans  $(\mathbf{R}, |\cdot|)$  pour comprendre ce que la notion de connexité par arcs généralise de nos connaissances sur l'ensemble des réels.

**Proposition 4.3.1.** — *Les parties de  $\mathbf{R}$ , connexes par arcs, sont les intervalles*

*Preuve de la proposition 4.3.1.* —

On a vu en MPSI que l'image par une application de  $\mathbf{R}$  dans  $\mathbf{R}$  d'un intervalle est un intervalle et donné comme corollaire de ce résultat le théorème de la valeur intermédiaire. Nous allons généraliser ces deux résultats aux cas des parties connexes par arcs.

**Proposition 4.3.2.** — *Soit  $\vec{f}$  une application d'une partie  $D$  de  $(\mathbf{E}, \|\cdot\|)$  dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , continue. L'image directe par  $\vec{f}$  de toute partie  $A$  de  $D$ , connexe par arcs est connexe par arcs.*

*Preuve de la proposition 4.3.2.* —

Passons à la généralisation du théorème de la valeur intermédiaire.

**Proposition 4.3.3.** — Soit  $g$  une application d'une partie  $A$  de  $(\mathbf{E}, \|\cdot\|)$ , connexe par arcs, dans  $\mathbf{R}$ , continue. Soient  $\vec{x}_1$  et  $\vec{x}_2$  des éléments de  $A$  tels que  $g(\vec{x}_1) \leq g(\vec{x}_2)$ . Pour tout élément  $t$  de  $[g(\vec{x}_1), g(\vec{x}_2)]$  il existe un élément  $\vec{y}$  de  $A$  tel que :  $g(\vec{y}) = t$ .

*Preuve de la Proposition 4.3.3.* —

Terminons par quelques exercices qui montrent l'importance du théorème de la valeur intermédiaire.

**Exercice 4.3.4** — Soient  $A_1$  et  $A_2$  des parties de  $\mathbf{E}$  connexes par arcs, d'intersection non vide. Montrer que  $A_1 \cup A_2$  est connexe par arcs.

*Solution de l'exercice 4.3.4.* —

Voilà maintenant un exercice donnant un résultat très important en analyse, nous illustrerons cette importance par deux applications spectaculaires.

**Exercice 4.3.5.** —

Soit  $A$  une partie connexe par arcs de  $(\mathbf{E}; \|\cdot\|)$ .

1. Soit  $B$  une partie de  $A$  qui est ouverte relativement à  $A$  et fermée relativement à  $A$ . Montrer que  $B = A$  ou  $B = \emptyset$ .
2. Soit  $\vec{f}$  une application continue de  $A$  dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . On suppose que pour tout élément  $\vec{x}$  de  $A$ , il existe un voisinage  $V$  de  $\vec{x}$ , relativement à  $A$  tel que  $\vec{f}$  soit constante sur  $V$ , on dit que  $\vec{f}$  est relativement constante. Montrer que  $\vec{f}$  est constante.
3. On suppose que  $A$  est ouverte. Montrer que pour tout couple  $(\vec{a}, \vec{b})$  d'élément de  $A$ , il existe une ligne brisée incluse dans  $A$ , d'extrémité  $\vec{a}, \vec{b}$ .

*Solution de l'exercice 4.3.5. —*

**Exercice d'application à l'algèbre linéaire —**

- Soit  $n$  un entier naturel non nul. Montrer que  $\mathrm{GL}_n(\mathbf{R})$  est une partie ouverte dense non connexe par arcs de l'e.v.n  $(\mathcal{M}_n(\mathbf{R}), n_\infty)$ .
- Montrer que  $\mathrm{O}_n(\mathbf{R})$  est une partie non connexe par arcs de l'e.v.n  $(\mathcal{M}_n(\mathbf{R}), n_\infty)$ .
- Montrer que  $\mathrm{SO}_2(\mathbf{R})$  est une partie connexe par arcs de l'e.v.n  $(\mathcal{M}_n(\mathbf{R}), n_\infty)$ .
- Déterminer les composantes connexes par arcs de  $\mathrm{GL}_2(\mathbf{R})$  dans l'e.v.n  $(\mathcal{M}_n(\mathbf{R}), n_\infty)$ .

— Montrer que  $\mathrm{GL}_n(\mathbf{C})$  est une partie connexe par arcs de l'e.v.n  $(\mathcal{M}_n(\mathbf{C}), n_\infty)$ .

*Solution* —



#### 4.4 Applications linéaires continues

Dans toute cette partie  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  et  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  désignent des espaces vectoriels normés sur le corps  $\mathbf{K}$ , ( $\mathbf{K}$  désigne toujours le corps des réels ou celui des complexes).

Soit  $\vec{\ell}$  une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ . La linéarité de  $\vec{\ell}$  dit notamment que  $\vec{\ell}(\lambda\vec{x}) = \lambda\vec{\ell}(\vec{x})$ , pour tout élément  $\vec{x}$  de  $\mathbf{E}$  et tout élément  $\lambda$  de  $\mathbf{K}$ . Ainsi, la connaissance de  $\vec{\ell}$  au voisinage de  $\vec{0}_{\mathbf{E}}$  donne-elle la connaissance de  $\vec{\ell}$  sur l'espace  $\mathbf{E}$  entier. Il en résulte la proposition fondamentale suivante.

**Proposition 5.1.1.** — *Soit  $\vec{\ell}$  une application linéaire de  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . Les quatre propositions suivantes sont équivalentes :*

- i. L'application  $\vec{\ell}$  est continue (sur  $\mathbf{E}$ ).*
- ii. L'application  $\vec{\ell}$  est continue en  $\vec{0}_{\mathbf{E}}$ .*
- iii. Il existe un réel  $K$  tel que pour tout élément  $\vec{x}$  de  $\mathbf{E}$ ,*

$$\|\vec{\ell}(\vec{x})\|_{\mathbf{F}} \leq K\|\vec{x}\|_{\mathbf{E}}.$$

- iv. L'application  $\vec{\ell}$  est lipschitzienne de  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ .*



*Preuve de la proposition 5.1.1. —*

**Proposition 5.1.2.** — *L'ensemble des applications linéaires continues de  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$  dans  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ , est un sous-espace vectoriel de  $\mathcal{L}(\mathbf{E}, \mathbf{F})$ . On le note  $\mathcal{L}_c(\mathbf{E}, \mathbf{F})$ .*

*Preuve de la proposition 5.1.2. —*

**Proposition 5.1.3.** —  $\mathcal{L}_c(\mathbf{E}, \mathbf{E})$ , est une sous algèbre de  $(\mathcal{L}(\mathbf{E}, \mathbf{E}), +, \cdot, \circ)$ .

*Preuve de la proposition 5.1.3. —*

Donnons quelques exemples élémentaires d'étude de la continuité d'applications linéaires.

**Exemple 5.1.4.** — Dans ces exemples  $\mathbf{E}$  désigne  $\mathcal{C}^0([0, 1], \mathbf{R})$  espace vectoriel des applications de  $[0, 1]$  dans  $\mathbf{R}$ .

- Soit l'application linéaire :

$$\ell_0 : \mathbf{E} \rightarrow \mathbf{R}, f \mapsto f(0).$$

— On munit  $\mathbf{E}$  de la norme  $N_\infty$  et  $\mathbf{R}$  de la valeur absolue. Etudions la continuité de  $\ell_0$

— Reprenons la même étude en munissant  $\mathbf{E}$  de la norme  $N_1$ .

- Soit l'application linéaire :  $P_0$  de  $\mathbf{E}$  dans  $\mathbf{E}$  qui tout élément  $f$  de  $\mathbf{E}$  associe ;

$$[0, 1] \rightarrow [0, 1]; x \mapsto \int_0^x f(t) dt.$$

— On munit  $\mathbf{E}$  de la norme  $N_\infty$ . Etudions la continuité de  $P_0$

— Reprenons la même étude en munissant  $\mathbf{E}$  de la norme  $N_1$ .

- Soit l'application linéaire :  $D : \mathcal{C}^1([0, 1], \mathbf{R}) \rightarrow \mathbf{E}$ ,  $f \mapsto f'$ .
  - On munit  $\mathbf{E}$  de la norme  $N_\infty$  et  $\mathcal{C}^1([0, 1], \mathbf{R})$  de la restriction de  $N_\infty$ . Etudions la continuité de  $D$

— Reprenons la même étude en munissant toujours  $\mathbf{E}$  de la norme  $N_\infty$  et  $\mathcal{C}^1([0, 1], \mathbf{R})$  de la norme,

$$N : \mathcal{C}^1([0, 1] \rightarrow \mathbf{R}; f \mapsto N_\infty(f) + N_\infty(f').$$

## 4.5 Normes équivalentes

Dans tout ce paragraphe,  $\mathbf{E}$  désigne un espace vectoriel sur le corps  $\mathbf{K}$ , ( $\mathbf{K}=\mathbf{R}$  ou  $\mathbf{C}$ ).

Les notions topologiques que nous venons d'étudier reposent sur celles, première d'*ouvert*, il est donc naturel d'adopter la définition suivante :

**Définition 5.3.1.** — Soient  $N$  et  $N'$  des normes sur  $\mathbf{E}$ . On dit que  $N'$  est équivalente à  $N$  si, par définition, elle définit les mêmes ouverts que  $N$ , dit de façon plus précise, si toute partie de  $\mathbf{E}$ , est ouverte dans l'e.v.n.  $(\mathbf{E}, N')$  si et seulement si elle est ouverte dans l'e.v.n.  $(\mathbf{E}, N)$ .

Il est évident que la relation « être équivalente à », sur l'ensemble des normes définies sur  $\mathbf{E}$ , est une relation d'équivalence, on peut donc dire que des normes *sont équivalentes*.

Nous allons donner des critères d'équivalence pour des normes. Pour ce faire, on va devoir commencer par une définition qui repose sur la précision terminologique suivante.

Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ , dans une partie  $D'$  d'un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$ . Bien que  $f$  ne soit pas à valeurs dans un e.v.n., et ne relève donc pas de la définition 2.2.10., on peut quant-même parler de la continuité de  $\vec{f}$ , en convenant que  $\vec{f}$  est continue si par définition l'application

$$D \rightarrow \mathbf{F}; \vec{x} \mapsto \vec{f}(\vec{x})$$

l'est. Avec cette convention donnons la définition suivante.

**Définition 5.3.2.** — HOMÉOMORPHISME —

Soit  $\vec{f}$  une application d'une partie  $D$  d'un e.v.n.  $(\mathbf{E}_1, \|\cdot\|_{\mathbf{E}_1})$ , dans une partie  $D'$  d'un e.v.n.  $(\mathbf{E}_2, \|\cdot\|_{\mathbf{E}_2})$ . On dit que  $\vec{f}$  est un homéomorphisme de  $D$  sur  $D'$ , si par définition

1. L'application  $\vec{f}$  est une bijection de  $D$  sur  $D'$ .
2. L'application  $\vec{f}$  est continue.
3. La bijection réciproque de  $\vec{f}$ ,  $\vec{f}^{-1}$ , est continue.

**Proposition 5.3.3.** — Soient  $N$  et  $N'$  des normes sur  $\mathbf{E}$ . Les normes  $N$  et  $N'$  sont équivalentes si et seulement si l'application  $\text{id}_{\mathbf{E}}$  de  $(\mathbf{E}, N)$  dans  $(\mathbf{E}, N')$  est un homéomorphisme.

*Preuve de la proposition 5.3.3.—*

Donnons maintenant un critère *pratique* d'équivalence entre normes. C'est ce critère que l'on emploie pour prouver concrètement que des normes sont équivalentes.

**Proposition 5.3.4.** — *Soient  $N$  et  $N'$  des normes sur  $\mathbf{E}$ .  $N$  et  $N'$  sont équivalente si et seulement si, il existe des réels strictement positifs,  $h$  et  $k$  tels que :*

$$hN' \leq N \leq kN',$$

*c'est-à-dire, tels que pour tout élément  $\vec{x}$  de  $\mathbf{E}$ ,  $hN'(\vec{x}) \leq N(\vec{x}) \leq kN'(\vec{x})$ .*

*Preuve de la proposition 5.3.4. —*

Soient  $N$  et  $N'$  des normes sur  $\mathbf{E}$  équivalentes. On sait qu'il existe des réels strictement positifs,  $h$  et  $k$  tels que :

$$hN' \leq N \leq kN'. \quad (\text{III.8})$$

Pour tout élément  $\vec{a}$  de  $\mathbf{E}$  et  $r$  de  $\mathbf{R}_+^*$ , on note :

- $B_f(\vec{a}, r)$  (rep.  $B_o(\vec{a}, r)$ ) la boule fermée (resp. ouverte) de centre  $\vec{a}$  et de rayon  $r$  de l'e.v.n.  $(\mathbf{E}, N)$  ;
- $B'_f(\vec{a}, r)$  (rep.  $B'_o(\vec{a}, r)$ ) la boule fermée (resp. ouverte) de centre  $\vec{a}$  et de rayon  $r$  de l'e.v.n.  $(\mathbf{E}, N')$ .

Prenons un élément  $\vec{a}$  de  $\mathbf{E}$  et  $r$  un élément de  $\mathbf{R}_+^*$ . On a alors, d'après (III.8),

$$B'_f(\vec{a}, \quad) \subset B_f(\vec{a}, r) \subset B'_f(\vec{a}, \quad) \quad (\text{III.9})$$

$$B'_o(\vec{a}, \quad) \subset B_o(\vec{a}, r) \subset B'_o(\vec{a}, \quad) \quad (\text{III.10})$$

*Toute boule de centre  $\vec{a}$  pour une norme contient et est contenue dans une boule de même centre pour l'autre norme.*

**Remarque 5.3.5.** — L'équivalence des normes est une propriété très forte. Toute les propriétés topologiques vues à cette heure, vont rester invariantes par changement de normes équivalentes. Précisément, soient  $N$  et  $N'$  des normes équivalentes sur  $\mathbf{E}$ , et  $n$  et  $n'$  des normes équivalentes sur un  $\mathbf{K}$ -espace vectoriel  $\mathbf{F}$ .

Par définition :

- Les ouverts de  $(\mathbf{E}, N)$  et de  $(\mathbf{E}, N')$  coïncident.

Donc par passage au complémentaire :

- Les fermés de  $(\mathbf{E}, N)$  et de  $(\mathbf{E}, N')$  coïncident.

L'intérieure, respectivement l'adhérence, d'une partie  $A$  est définie comme le plus grand ouvert inclus  $A$ , respectivement comme le plus petit fermé contenant  $A$ . Donc :

- L'intérieure, l'adhérence et donc la frontière d'une partie  $A$  de  $\mathbf{E}$  dans  $(\mathbf{E}, N)$  et dans  $(\mathbf{E}, N')$  coïncident

Les ouverts servent à définir les voisinages, donc :

- Les voisinages de  $(\mathbf{E}, N)$  et de  $(\mathbf{E}, N')$  coïncident.

La notion de voisinage sert à définir la notion de convergence et de valeur d'adhérence d'une suite, de limite d'une application et de continuité d'une application. D'où :

- Une suite d'éléments de  $\mathbf{E}$  converge vers un élément  $\vec{\ell}$  de  $\mathbf{E}$ , dans  $(\mathbf{E}, N)$  si et seulement si elle converge dans  $(\mathbf{E}, N')$  vers  $\vec{\ell}$ .
- Un élément  $\vec{\ell}$  de  $\mathbf{E}$  est valeur d'adhérence d'une suite d'éléments de  $\mathbf{E}$  dans  $(\mathbf{E}, N)$  si et seulement si il est valeur d'adhérence de cette suite dans  $(\mathbf{E}, N')$ .
- Soit  $\vec{f}$  une application d'une partie  $D$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$ .  $\vec{f}(\vec{x})$  tend vers un élément  $\vec{\ell}$  de  $\mathbf{F}$  dans  $(\mathbf{F}, n)$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  (point adhérent à  $D^1$ ) dans  $(\mathbf{E}, N)$  si et seulement si  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$  dans  $(\mathbf{F}, n')$  lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}, N')$ .
- Une application d'une partie  $D$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$  est continue lorsque  $\mathbf{E}$  est muni de  $N$  et  $\mathbf{F}$  de  $n$  si et seulement si elle est continue lorsque  $\mathbf{E}$  est muni de  $N'$  et  $\mathbf{F}$  de  $n'$ .

La notion de valeur d'adhérence d'une suite et celle de suite convergente étant invariantes par changement de normes équivalentes, on obtient :

- Les compacts de  $(\mathbf{E}, N)$  et de  $(\mathbf{E}, N')$  coïncident.

Le critère d'équivalence 5.3.4. assure également les propriétés suivantes.

- Une suite à valeurs dans  $\mathbf{E}$  est bornée dans  $(\mathbf{E}, N)$  si et seulement si elle est bornée dans  $(\mathbf{E}, N')$ .
- Une application d'une partie  $D$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$  est bornée, lorsque  $\mathbf{F}$  est muni de  $n$ , si et seulement si elle est bornée, lorsque  $\mathbf{F}$  est muni de  $n'$ .
- Une application d'une partie  $D$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$  est uniformément continue, lorsque  $\mathbf{E}$  est muni de  $N$  et  $\mathbf{F}$  de  $n$ , si et seulement si elle est uniformément continue, lorsque  $\mathbf{E}$  est muni de  $N'$  et  $\mathbf{F}$  de  $n'$ .

---

1. L'adhérence de  $D$  est prise dans  $(\mathbf{E}, N)$  ou  $(\mathbf{E}, N')$ , puisque d'après ce qui précède c'est la même chose !

- Une application d'une partie  $D$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$  est lipschitzienne, lorsque  $\mathbf{E}$  est muni de  $N$  et  $\mathbf{F}$  de  $n$  si et seulement si elle est lipschitzienne, lorsque  $\mathbf{E}$  est muni de  $N'$  et  $\mathbf{F}$  de  $n'^1$ .
- Une application du segment  $[0, 1]$  de  $(\mathbf{R}, |\cdot|)$  dans  $\mathbf{E}$  étant continue lorsque l'on munit  $\mathbf{E}$  de  $N$  si et seulement si elle lorsque l'on munit  $\mathbf{E}$  de  $N'$ , une partie  $A$  de  $E$  est un connexe par arcs de  $(\mathbf{E}, N)$  si et seulement si c'est un connexe par arcs de  $(\mathbf{E}, n')$ .

**Exercice 5.3.6.** — DEMI-ÉQUIVALENCE —

Soient  $N$  et  $N'$  des normes sur  $\mathbf{E}$ . On suppose qu'il existe un élément  $k$  de  $\mathbf{R}_+^*$  tel que :

$$N \leq kN'.$$

Alors, si une suite à valeurs dans  $\mathbf{E}$  converge dans  $(\mathbf{E}, N)$ , alors elle converge dans  $(\mathbf{E}, N')$ .

*Solution de l'exercice 5.3.6.* —

Nous allons maintenant donner des exemples de normes équivalentes et de normes non équivalentes.

**Exemples 5.3.7.**— Soit  $n$  un entier supérieur ou égal à 1. Les normes sur  $\mathbf{R}^n$ ,  $n_1$ ,  $n_2$ , et  $n_\infty$  sont équivalentes.

- $n_1$  ET  $n_\infty$  SONT ÉQUIVALENTES  
Soit  $(x_1, x_2, \dots, x_n)$  un élément de  $\mathbf{R}^n$ .

1. Attention la constante de lipschitzité peut, elle, varier.



- $n_1$  ET  $n_2$  SONT ÉQUIVALENTES

Soit  $(x_1, x_2, \dots, x_n)$  un élément de  $\mathbf{R}^n$ .

- Par transitivité  $n_2$  et  $n_\infty$  sont équivalentes. Le lecteur est invité à donner une preuve directe de ce fait en utilisant 5.3.4. et en exhibant les meilleures constantes  $h$  et  $k$ .

**Exemples 5.3.8.**— On note  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([a, b], \mathbf{R})$ , où  $[a, b]$  désigne un segment non réduit à un point.

Montrons qu'aucun couple d'éléments de  $\{N_1, N_2, N_\infty\}$  n'est un couple de normes équivalentes.

Nous venons de voir que les normes classique sur  $\mathcal{C}^0[a, b]$ ,  $\mathbf{R}$  ne sont pas équivalentes en revanche, les normes usuelles sur  $\mathbf{R}^n$  sont équivalentes. En fait on dispose du résultat incroyablement plus fort suivant

**Théorème 5.3.9.** — *On suppose que  $\mathbf{E}$  est un espace vectoriel de dimension finie. Alors, toutes les normes sur  $\mathbf{E}$  sont équivalentes.*

*Preuve du théorème 5.3.9.* —

Nous allons montrer que toute norme est équivalente à une norme particulière que nous allons définir maintenant. La relation sur l'ensemble des normes définies sur  $\mathbf{E}$  « être équivalente » étant une relation d'équivalence, on aura l'équivalence de toutes les normes.

Excluons le cas trivial où  $\mathbf{E}$  est de dimension nulle. Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  une base de  $\mathbf{E}$  notée  $\mathcal{B}$ .

Le théorème 5.3.9. et la remarque 5.3.5. font, que lorsque, en dimension finie, on parle de notions topologiques, on ne précise pas les normes utilisées. Ainsi, écrit-on par exemple :

« Soit  $K$  un compact de  $\mathcal{M}_3(\mathbf{R})$  » ou encore « Soit  $F$  une application continue de  $\mathbf{K}^6$  dans  $\mathcal{M}_3(\mathbf{R})$ . » ...

Nous allons maintenant exploiter cette propriété des espaces vectoriels de dimension finie pour étudier leur propriétés topologiques, propriétés, répétons le, qui ne dépendent pas des normes dont on les équipe.

## ESPACES VECTORIELS DE DIMENSION FINIE

Dans toute cette partie  $\mathbf{E}$  désigne un  $\mathbf{K}$ -espace vectoriel de dimension finie  $d$  non nulle ( $\mathbf{K}=\mathbf{R}$  ou  $\mathbf{C}$ ). On considérera une base quelconque de  $\mathbf{E}$ ,  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_d)$ , que nous noterons  $\mathcal{B}$ .

### 5.1 Limites de suites et d'application

Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite d'éléments de  $\mathbf{E}$ . Pour tout  $i$  élément de  $\{1, 2, \dots, d\}$  et pour tout entier naturel  $n$ , on note  $x_{i,n}$  la  $i^{\text{e}}$  coordonnée de  $\vec{x}_n$  dans  $\mathcal{B}$ . On dispose donc de suite réelles  $(x_{i,n})_{n \in \mathbf{N}}$ ,  $i = 1, 2, \dots, d$ . Soit  $\vec{\ell}$  un élément de  $\mathbf{E}$ . On note  $(\ell_1, \ell_2, \dots, \ell_d)$  le  $d$ -uplet des coordonnées de  $\vec{\ell}$  dans  $\mathcal{B}$ . On a alors le résultat

**Proposition 6.1.1.** — *La suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  converge vers  $\vec{\ell}$  si et seulement si, pour tout élément  $i$  de  $\{1, 2, \dots, d\}$ ,  $x_{i,n} \xrightarrow{n \rightarrow +\infty} \ell_i$ .*

*Preuve de la proposition 6.1.1.* — Commençons par introduire un résultat utile dans la suite. On note  $\Phi$  l'application de  $\mathbf{E}$  dans  $\mathbf{K}^d$ , qui à un vecteur  $\vec{x}$  de  $\mathbf{E}$  associe le  $d$ -uplet de ses coordonnées dans  $\mathcal{B}$ . Le cours de sup. nous apprend que  $\Phi$  est un isomorphisme de l'espace vectoriel  $\mathbf{E}$  sur l'espace vectoriel  $\mathbf{K}^d$ . on notera  $\Psi$  l'isomorphisme réciproque,

$$\Psi : \mathbf{K}^d \rightarrow \mathbf{E}; (x_1, x_2, \dots, x_d) \mapsto \sum_{i=1}^d x_i \vec{e}_i.$$

On peut définir sur  $\mathbf{E}$ , la norme suivante :

$$n_{\mathcal{B},\infty} : \mathbf{E} \rightarrow \mathbf{R}; \vec{x} \mapsto \sup_{i=1,2,\dots,d} |x_i|,$$

où  $x_i$  désigne la  $i^{\text{e}}$  coordonnée de  $\vec{x}$  dans  $\mathcal{B}$ , pour  $i = 1, 2, \dots, d$ . Si l'on préfère  $n_{\mathcal{B},\infty} = n_\infty \circ \Phi$ , ce qui permet de vérifier rapidement que  $n_{\mathcal{B},\infty}$  est une norme. Par définition, pour tout élément  $\vec{x}$  de  $\mathbf{E}$ ,  $n_{\mathcal{B},\infty}(\vec{x}) = n_\infty(\Phi(\vec{x}))$ .<sup>1</sup> Donc...

On a donc prouvé

**lemme 6.1.2.** — *L'application  $\Phi$  est un homéomorphisme de  $\mathbf{E}$  sur  $\mathbf{K}^d$ .*

*L'application  $\Psi$  est un homéomorphisme de  $\mathbf{K}^d$  sur  $\mathbf{E}$ .*

Terminons la preuve de 6.1.1.

D'où le résultat<sup>2</sup>. □

La caractérisation séquentielle de la limite d'une application (cf. 2.2.5.) permet de déduire de ce résultat un résultat analogue pour les applications.

1. On dit que  $\Phi$  est une isométrie de  $(\mathbf{E}, n_{\mathcal{B},\infty})$  dans  $(\mathbf{K}^d, n_\infty)$ .

2. On notera que bien que toutes les normes sur  $\mathbf{E}$  soit équivalentes, certaines telles  $n_{\mathcal{B},\infty}$  sont plus pratiques que d'autres.

Soient  $(\mathbf{E}_0, N_0)$  un e.v.n. de dimension quelconque et  $\vec{f}$  une application d'une partie  $D$  de  $(\mathbf{E}_0, N_0)$  à valeurs dans  $\mathbf{E}$ . On note pour  $i = 1, 2, \dots, d$ ,  $f_i$  la  $i^{\text{e}}$  application composante de  $\vec{f}$  dans  $\mathcal{B}$ . Soit  $\vec{a}$  un point de  $\bar{D}$  et  $\vec{\ell}$  un élément de  $\mathbf{E}$ . On note  $(\ell_1, \ell_2, \dots, \ell_d)$ ,  $\Phi(\vec{\ell})$ . On a alors :

**Proposition 6.1.3.** —  $\vec{f}(\vec{x})$  tend vers  $\vec{\ell}$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}_0, N_0)$  si et seulement si  $f_i(\vec{x})$  tend vers  $\ell_i$ , lorsque  $\vec{x}$  tend vers  $\vec{a}$  dans  $(\mathbf{E}_0, N_0)$ , pour  $i = 1, 2, \dots, d$ .

On a alors, en conservant les notations précédentes, le corollaire immédiat suivant

**Proposition 6.1.4.** —

- Soit  $\vec{x}_0$  un élément de  $D$ . L'application  $\vec{f}$  est continue, en  $\vec{x}_0$ , si et seulement si pour tout élément  $i$  de  $\{1, 2, \dots, d\}$ ,  $f_i$  est continue en  $\vec{x}_0$ .
- L'application  $\vec{f}$  est continue, si et seulement si pour tout élément  $i$  de  $\{1, 2, \dots, d\}$ ,  $f_i$  est continue

Cette proposition nous dit que pour étudier la continuité d'une application à valeurs dans un espace vectoriel de dimension  $d$ , il suffit d'étudier la continuité de  $d$  applications à valeurs dans  $K$ . On fera bien attention qu'il n'existe aucun résultat analogue pour une application définie sur un espace vectoriel de dimension  $d$ .

**Ce qui marche...**

Soit l'application

$$M : \mathbf{R} \rightarrow \mathcal{M}_2(\mathbf{R}); t \mapsto \begin{pmatrix} 2 \cos t & -\sin t \\ \sin t & \frac{1}{2} \cos t \end{pmatrix}.$$

Les applications de  $\mathbf{R}$  dans  $\mathbf{R}$ ,  $t \mapsto 2 \cos t$ ,  $t \mapsto \sin t$ ,  $t \mapsto -\sin t$  et  $t \mapsto \frac{1}{2} \cos t$  sont continues; ce sont les applications coordonnées de  $M$ , dans la base canonique de  $\mathcal{M}_2(\mathbf{R})$ , donc  $M$  est continue.

**Ce qui ne marche pas...**

Soit l'application

$$f : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} 0, & \text{pour } x \neq y, \\ 1, & \text{pour } x = y \text{ et } x \neq 0, \\ 0, & \text{pour } x = y = 0. \end{cases}$$

Les applications

$$f(\cdot, 0) : \mathbf{R} \rightarrow \mathbf{R}; x \mapsto f(x, 0),$$

$$f(0, \cdot) : \mathbf{R} \rightarrow \mathbf{R}; y \mapsto f(0, y),$$

sont continues en 0, car...

. Mais l'application  $f$  n'est pas continue en  $(0, 0)$ ; en effet .

## 5.2 Compacité

En dimension finie la réciproque de la proposition 4.1.2. est vraie :

**PROPOSITION 6.2.1.** — Les parties compactes de l'espace vectoriel de dimension finie  $\mathbf{E}$  sont exactement ses parties fermées bornées

*Preuve de la proposition 6.2.1.*— On a vu qu'en dimension quelconque, tout compact est fermé borné. Il nous suffit donc de prouver qu'une partie  $K$  de  $\mathbf{E}$ , fermée et bornée est compacte.

Exercice 6.2.2. — Montrer que  $\mathcal{O}_n(\mathbf{R})$  est une partie compacte de  $\mathcal{M}_n(\mathbf{R})$ .

*Solution de l'exercice 6.2.2. —*

**Théorème 6.2.3.** — BOLZANO-WEIERSTRASS —

*Toute suite bornée, à valeurs dans un espace de dimension finie, admet une suite extraite convergente, ou, si l'on préfère, une valeur d'adhérence*

*Preuve du théorème 6.2.3.* — Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite bornée d'éléments de  $\mathbf{E}$ . Par définition du caractère bornée, il existe une boule fermée  $B$  de  $\mathbf{E}$  telle que  $(\vec{x}_n)_{n \in \mathbf{N}}$  soit à valeurs dans  $B$ . Mais la boule  $B$ , fermée et bornée dans  $\mathbf{E}$ , de dimension finie, est compacte. D'où le résultat.  $\square$

Il résulte de la proposition 3.1.3., que

**Proposition 6.2.4.** — *Une suite bornée de  $E$  converge si et seulement si elle admet une et une seule valeur d'adhérence.*

**Exercice** — *Soit  $(\vec{x}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $\mathbf{E}$  (espace vectoriel de dimension finie), qui converge vers un élément  $\vec{\ell}$  de  $\mathbf{E}$ . Alors  $\{\vec{x}_n, n \in \mathbf{N}\} \cup \{\vec{\ell}\}$  est un compact.*

*Solution* —

### 5.3 Applications linéaires et multilinéaires

**Proposition 6.3.1.** — *Toute application linéaire  $\vec{\ell}$  de  $\mathbf{E}$  dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  est continue.*

*Preuve de la proposition 6.3.1.* —



**Proposition 6.3.2.** — Soient  $\mathbf{E}_1$  et  $\mathbf{E}_2$  des espaces vectoriels de dimensions finies. Toute application  $\vec{b}$  de  $\mathbf{E}_1 \times \mathbf{E}_2$  dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  (de dimension quelconque), bilinéaire est continue.

*Preuve de la proposition 6.3.2.* —

On laissera en exercice la généralisation suivante de la proposition 6.4.3.

**Proposition 6.3.3.** — Soit  $d$  un entier supérieur ou égal à 2. Soient  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_p$  des espaces vectoriels de dimensions finies. Toute application  $\vec{b}$  de  $\mathbf{E}_1 \times \mathbf{E}_2 \times \dots \times \mathbf{E}_p$  dans un e.v.n.  $(\mathbf{F}, \|\cdot\|_{\mathbf{F}})$  (de dimension quelconque),  $d$ -linéaire est continue.

### Exemples

- Soit  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$  un espace euclidien. Le produit scalaire est continue sur  $\mathbf{E} \times \mathbf{E}$ . Si de plus  $\mathbf{E}$  est de dimension 3 orienté, Le produit vectoriel est continue.
- De nouveau  $\mathbf{E}$  est un  $\mathbf{K}$ -espace vectoriel quelconque de dimension finie  $d$ , et  $\mathcal{B}$  désigne une de ses bases. L'application  $\det_{\mathcal{B}}$  est continue sur  $\mathbf{E}^d$ .
- On a déjà utilisé la continuité de  $\det$  sur  $\mathcal{M}_n(K)$ . Rappelons .....

## ANNEXE : BARYCENTRES

Brève mise au point sur les espaces affines, et les barycentres, cette annexe est utile pour l'étude des parties convexes.

Donnons rapidement quelques précisions sur les espaces affines (entre)vus en MPSI. Commençons par examiner le cas d'un sous-espace affine d'un espace vectoriel  $\mathbf{E}$ , c'est-à-dire d'un ensemble  $\mathcal{F}$  de la forme

$$\mathcal{F} = \vec{x}_0 + \mathbf{F},$$

où  $\mathbf{F}$  est un sous-espace vectoriel de  $\mathbf{E}$  (appelé direction de  $\mathcal{F}$ ). On observe que la somme deux éléments de  $\mathcal{F}$  n'est pas en général un élément de  $\mathcal{F}$ , par contre la somme d'un élément de  $\mathcal{F}$  et d'un élément de  $\mathbf{F}$  est élément de  $\mathcal{F}$ , notons aussi que la différence entre deux éléments de  $\mathcal{F}$  est un élément de  $\mathbf{F}$ , enfin pour tout couple d'éléments de  $\mathcal{F}$ ,  $(p_1, p_2)$  il existe un et un seul élément de  $\mathbf{F}$ ,  $\vec{f}$  tel que  $p_1 + \vec{f} = p_2$ . Ces propriétés vont inspirer la définition générale d'espace affine.

Soit  $\mathcal{E}$  un ensemble, on notera ses éléments  $m, p, o, q$  etc., on les appellera points, et  $\mathbf{E}$  un espace vectoriel, on notera ses éléments  $\vec{x}, \vec{y}, \vec{v}$  etc. On dit que  $\mathcal{E}$  est un espace affine de direction  $\mathbf{E}$  si il existe une opération de  $\mathbf{E}$  sur  $\mathcal{E}$ ,

$$\mathbf{E} \times \mathcal{E} \rightarrow \mathcal{E}; (p, \vec{x}) \mapsto p \oplus \vec{x}$$

(à un point de  $\mathcal{E}$  on ajoute un vecteur de  $\mathbf{E}$  pour obtenir un point de  $\mathcal{E}$ ) qui jouit des propriétés suivantes :

1. Pour tout  $p \in \mathcal{E}$ ,  $p \oplus \vec{0}_{\mathbf{E}} = p$ .
2. Pour tout  $p \in \mathcal{E}$  et  $(\vec{x}, \vec{y}) \in \mathbf{E}^2$ ,  $p \oplus (\vec{x} + \vec{y}) = (p \oplus \vec{x}) \oplus \vec{y}$
3. Pour tout couple  $(p, q)$  d'éléments de  $\mathcal{E}$ , il existe un et un seul élément  $\vec{x}$  de  $\mathbf{E}$  tel que :  $q = p \oplus \vec{x}$ . On le note  $\vec{pq}$  ou  $q - p$ .

On ne distingue plus typographiquement dans la suite l'addition  $+$  de deux vecteurs de  $\mathbf{E}$  et celle d'un vecteur à un point  $\oplus$ .

Ainsi par exemple un sous-espace affine  $\mathcal{F}$  de direction  $\mathbf{F}$  d'un espace vectoriel  $\mathbf{E}$  est-il un espace affine de direction  $\mathbf{F}$ . Un cas particulier intéressant est le sous-espace affine de  $\mathbf{E}$ ,  $\vec{0}_{\mathbf{E}} + \mathbf{E}$ . Autrement dit  $\mathbf{E}$  se munit naturellement d'une structure d'espace affine, dite canonique.

De façon imagée et intuitive munir  $\mathbf{E}$  de sa structure d'espace affine canonique revient à considérer ses éléments, tantôt comme des points (notés  $m, p, q...$  ou aussi  $M, A, B...$  et représentés par des croix) tantôt comme des vecteurs (notés  $\vec{x}, \vec{y}, \vec{v}...$  et représentés par des flèches. La loi  $\oplus$  et  $+$  se confondent en fait dans ce cas, et la distinction entre l'espace affine  $\mathbf{E}$  et l'espace vectoriel n'est guère que typographique.

**Définition** — On appelle repère d'un espace affine  $\mathcal{E}$  dirigé par  $\mathbf{E}$ , tout couple  $(o, \mathcal{B})$  où  $o$  est un point de  $\mathcal{E}$  et  $\mathcal{B}$  une base de  $\mathbf{E}$ . On appelle cordonnée d'un point  $p$  de  $\mathcal{E}$  dans le repère  $(o, \mathcal{B})$  les coordonnées de  $\vec{op}$  dans  $\mathcal{B}$ .

Dans la suite  $\mathcal{E}$  est un espace affine de direction  $\mathbf{E}$ , éventuellement  $\mathbf{E}$  lui-même muni de sa structure euclidienne canonique.

Soient  $n$  un entier supérieur ou égal à 2<sup>1</sup>,  $(p_1, p_2, \dots, p_n)$  un  $n$ -uplet de points de  $\mathcal{E}$ ,  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  un  $n$ -uplet de réels dont la somme est non nulle :

$$\sum_{i=1}^n \lambda_i \neq 0.$$

On a alors le résultat suivant.

**Proposition-définition A.1.** — *Soit  $o$  un point de  $\mathbf{E}$ , alors le point  $g$  définit par :*

$$g = o + \frac{1}{\sum_{i=1}^n \lambda_i} \sum_{i=1}^n \lambda_i \overrightarrow{op_i}$$

*est indépendant du point  $o$ .*

*On l'appelle barycentre de la famille de points  $(p_1, p_2, \dots, p_n)$  pondérée par  $(\lambda_1, \lambda_2, \dots, \lambda_n)$ .*

Le barycentre  $g$  des points pondérés  $(p_1, \lambda_1), (p_2, \lambda_2), \dots, (p_n, \lambda_n)$  est caractérisé par :

$$\sum_{i=1}^n \lambda_i \overrightarrow{gp_i} = \vec{0}_{\mathbf{E}}.$$

**Exemple** — Le barycentre  $m$  de deux points  $p$  et  $q$  munis tous deux du coefficient  $\frac{1}{2}$ , vérifie  $\overrightarrow{mp} + \overrightarrow{mq} = \vec{0}_{\mathbf{E}}$ , on l'appelle milieu de  $p$  et  $q$ .

**Remarques** —

1. Dans  $\mathbf{E}$ , muni de sa structure d'espace affine canonique le barycentre de la famille de points  $(p_1, p_2, \dots, p_n)$  pondérée par  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  s'écrit encore plus simplement

$$g = \frac{1}{\sum_{i=1}^n \lambda_i \neq 0} \sum_{i=1}^n \lambda_i p_i \neq 0,$$

comme on le voit en choisissant comme point  $o$  le vecteur nul de  $\mathbf{E}$ .

2. Si  $\mathcal{E}$  est  $\mathbf{R}$  muni de sa structure affine canonique, la notion de barycentre se confond avec celle de moyenne.
3. Si  $(\{1, 2, \dots, n\}, P)$  est un espace probabilisé et  $X$  une variable aléatoire définie sur cet espace à valeurs dans  $\mathbf{E}$ , alors on pourrait définir l'espérance de  $X$ , comme pour une variable réelle, comme le barycentre de  $(X(1), \dots, X(n))$  pondérée par  $(P(1), \dots, P(n))$
4. Lorsque  $\mathbf{E}$  est de dimension 3, Le barycentre de la famille de points  $(p_1, p_2, \dots, p_n)$  pondérée par  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  peut modéliser le centre de gravité de  $n$  masses ponctuelles  $\lambda_1, \lambda_2, \dots, \lambda_n$  situées en des points, représentés par  $p_1, p_2, \dots, p_n$ , de l'espace physique.

*Preuve de A.1.* — Soit  $o'$  un point de  $\mathcal{E}$ ...

Le barycentre est évidemment invariant par permutation des points, précisément, si  $\sigma \in S_n$  alors le barycentre de la famille  $(p_1, p_2, \dots, p_n)$  pondérée par  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  est aussi celui de la famille  $(p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(n)})$  pondérée par  $(\lambda_{\sigma(1)}, \lambda_{\sigma(2)}, \dots, \lambda_{\sigma(n)})$ . On parle donc plus simplement du barycentre des points  $p_1, p_2, \dots, p_n$  pondérés respectivement par  $\lambda_1, \lambda_2, \dots, \lambda_n$  ou encore des points pondérés  $(p_1, \lambda_1), (p_2, \lambda_2), \dots, (p_n, \lambda_n)$ . Le réel  $\lambda_i$ , est appelé, pour  $i = 1, \dots, n$ , coefficient ou poids du point  $p_i$ .

**Remarque** — Le barycentre  $g$  d'une famille de points  $(p_1, p_2, \dots, p_n)$ , pondérée par  $(\lambda_1, \lambda_2, \dots, \lambda_n)$ ,  $(\sum_{i=1}^n \lambda_i \neq 0)$ , est aussi pour tout réel  $\alpha$  non nul le barycentre de  $(p_1, p_2, \dots, p_n)$ , pondérée par  $(\alpha\lambda_1, \alpha\lambda_2, \dots, \alpha\lambda_n)$ . Le cas particulier

1. Le cas  $n = 1$  est envisageable mais d'un intérêt pratique totalement nul

$\alpha = \frac{1}{\sum_{i=1}^n \lambda_i}$ , dit que  $g$  est barycentre de  $(p_1, p_2, \dots, p_n)$  coefficientée par une famille de somme 1. Ainsi l'ensemble des barycentres de deux points  $p$  et  $q$  est-il l'ensemble

$$\{o + t\vec{op} + (1-t)\vec{oq}, t \in \mathbf{R}\}.$$

Dans le calcul d'un barycentre, comme pour une moyenne on peut associer des points.

**Proposition A.2.** — ASSOCIATIVITÉ DU BARYCENTRE —

Soient  $(p_1, \lambda_1), (p_2, \lambda_2), \dots, (p_n, \lambda_n)$  des points pondérés, avec  $\sum_{i=1}^n \lambda_i \neq 0$  et  $\{I_1, I_2, \dots, I_k\}$  une partition de  $\{1, \dots, n\}$  telle que pour  $j = 1, 2, \dots, k$ ,  $\sum_{i \in I_j} \lambda_i \neq 0$ . Notons pour  $j = 1, \dots, k$ ,  $g_j$  le barycentre des points pondérés  $(p_i, \lambda_i), i \in I_j$ . Alors le barycentre  $g$  des points pondérés  $(p_1, \lambda_1), \dots, (p_n, \lambda_n)$  est aussi le barycentre des points pondérés

$$\left(g_1, \sum_{i \in I_1} \lambda_i\right), \left(g_2, \sum_{i \in I_2} \lambda_i\right), \dots, \left(g_k, \sum_{i \in I_k} \lambda_i\right).$$

Le coefficient dans le calcul de  $g$  dont on affecte le barycentre partiel  $g_j$  est la somme des coefficients des points dont il est le barycentre.

*Preuve de A.2.* —

Notons que d'un point de vue pratique le calcul des coordonnées du barycentre dans un repère est très simple.

Soit  $g$  le barycentre des points pondérés  $(p_1, \lambda_1), \dots, (p_n, \lambda_n)$ . Soit  $\mathcal{R}$  un repère de  $\mathcal{E}$ , supposé de dimension  $d$ . notons pour  $i = 1, \dots, n$  et  $j = 1, \dots, d$ ,  $p_{j,i}$  la  $j^{\text{e}}$  coordonnée de  $p_i$  et  $g_j$  la  $j^{\text{e}}$  coordonnée de  $g$  dans le repère  $\mathcal{R}$ . Alors, de façon évidente, pour  $j = 1, \dots, d$  :

$$g_j = \frac{1}{\sum_{i=1}^n \lambda_i} \sum_{i=1}^n \lambda_i p_{j,i}$$

autrement dit la  $j^{\text{e}}$  coordonnée de  $g$  est la moyenne des  $j^{\text{e}}$  coordonnées de  $p_1, \dots, p_n$ , pondérées respectivement par  $\lambda_1, \dots, \lambda_n$ .



## Chapitre IV

# SÉRIES, FAMILLES SOMMABLES

## GÉNÉRALITÉ SUR LES SÉRIES À VALEURS DANS UN ESPACE VECTORIEL DE DIMENSION FINIE

Le présent chapitre a plusieurs objectifs de consolider les notions de sup sur les séries réelles et les compléter, étendre les connaissances sur les séries réelles aux séries à valeurs dans un espace vectoriel de dimension finie, enfin introduire dans le seul but de l'étude des probabilités les familles sommables.

Dans toute cette partie  $\mathbf{K}$  désigne indifféremment le corps des nombres complexes ou celui des nombres réels et  $\mathbf{E}$  un espace vectoriel de dimension finie  $d$  non nulle sur  $\mathbf{K}$ . L'équivalence des normes en dimension finie nous permet de ne pas préciser, sauf cas particulier, la norme dont on l'équipera et qui sera notée  $\|\cdot\|$ .

### 1.1 Définitions, premières propriétés

**Définition 1.1.1.** — Soit  $(\vec{u}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $(\mathbf{E}, \|\cdot\|)$ . On pose pour tout entier naturel  $n$ ,  $\vec{S}_n := \sum_{k=0}^n \vec{u}_k$ .

- On dit que la série de terme général  $\vec{u}_n$ , notée  $\sum \vec{u}_n$ , converge si, par définition, la suite  $(\vec{S}_n)_{n \in \mathbf{N}}$  à valeurs dans  $\mathbf{E}$ , converge. Dans le cas contraire, on dit que la série diverge.
- Dans le cas où la série de terme général  $\vec{u}_n$  converge, la limite de la suite  $(\vec{S}_n)_{n \in \mathbf{N}}$  s'appelle somme de la série, on la note :

$$\sum_{n=0}^{+\infty} \vec{u}_n.$$

- Pour tout entier naturel  $n$ ,  $\vec{S}_n$  s'appelle somme partielle d'ordre  $n$  de la série.

La précédente définition appelle les remarques suivantes.

**Remarques** —

- Nous n'avons pas défini réellement ce qu'est une série, l'essence mathématique de cet objet ne sera révélée que dans la suite et pour l'essentiel il n'est nul besoin d'en savoir plus que ce qui est dit dans la définition 1.1.1.
- Ne nous y trompons pas la « somme » d'une série, n'en n'est pas véritablement une, il serait faux de lui appliqué des propriétés déduites hâtivement de celles d'une vraie somme, telles l'associativité ou la commutativité.

Dans la pratique il est parfois utile d'étudier des séries associées à des suites indicées non par  $\mathbf{N}$  entier, mais par une partie de  $\mathbf{N}$ , de la forme  $(\vec{u}_n)_{n \geq n_0}$  où  $n_0$  est un élément de  $\mathbf{N}^*$ . On note  $\sum_{n \geq n_0} \vec{u}_n$  la série associée, cette série sera dite

convergente si par définition la suite  $(\vec{S}_n)_{n \geq n_0}$  converge, où pour tout entier  $n \geq n_0$ ,  $S_n := \sum_{k=n_0}^n \vec{u}_k$ . Dans le cas de

la convergence, la somme de la série sera naturellement notée  $\sum_{n=n_0}^{+\infty} \vec{u}_n$ . Ceci étant dit associons à la suite  $(\vec{u}_n)_{n \geq n_0}$  la suite  $(\vec{u}_n^*)_{n \in \mathbf{N}}$  définie par :

$$\vec{u}_n^* := \begin{cases} \vec{0}_{\mathbf{E}}, & \text{pour } n < n_0, \\ \vec{u}_n & \text{pour } n \geq n_0. \end{cases}$$

La somme partielle d'ordre  $n$  de la série  $\sum_{n \geq n_0} \vec{u}_n$  et celle de la série  $\sum \vec{u}_n^*$  coïncident, pour tout entier  $n \geq n_0$ , il en résulte que  $\sum_{n \geq n_0} \vec{u}_n$  converge si et seulement si  $\sum \vec{u}_n^*$  converge et, quand c'est le cas, on a l'égalité des sommes :

$\sum_{n=n_0}^{+\infty} \vec{u}_n = \sum_{n=0}^{+\infty} \vec{u}_n^*$ . Il est donc possible de ramener l'étude de toute série à celle d'une série de la forme donnée par la définition 1.1.1.

*C'est pourquoi dans tout ce chapitre, afin d'alléger l'écriture, nous nous limiterons à des séries du type 1.1.1., laissant dans les cas pratiques, le soin au lecteur de transcrire les résultats de ce cours à des séries plus générales.*

Soient  $\sum \vec{u}_n$  une série convergente de terme général élément de  $(\mathbf{E}, \|\cdot\|)$  et  $m$  un entier naturel. Alors la série  $\sum_{n \geq m+1} \vec{u}_n$  converge. En effet la suite de leurs sommes partielles diffère d'une constante, puisque pour tout entier  $n \geq m+1$ , avec les notation de 1.1.1.,

$$\sum_{k=0}^n \vec{u}_k = \vec{S}_m + \sum_{k=m+1}^n \vec{u}_k. \quad (\text{IV.1})$$

D'où la définition suivante

**Définition 1.1.2.** — RESTE D'UNE SÉRIE CONVERGENTE —

On appelle *reste d'ordre  $m$  de la série convergente  $\sum \vec{u}_n$* , noté  $\vec{R}_m$ , la somme de la série  $\sum_{n \geq m+1} \vec{u}_n$ , autrement dit :

$$\vec{R}_m := \sum_{k=m+1}^{+\infty} \vec{u}_k.$$

La formule (IV.1) assure alors par passage à la limite sur  $n$  :

$$\boxed{\sum_{k=0}^{+\infty} \vec{u}_k = \vec{S}_m + \vec{R}_m.}$$

Autrement dit  $\vec{R}_m$  est la quantité qui *reste* à ajouter à la somme partielle d'ordre  $m$  pour obtenir la somme de la série, d'où son nom.

La nature d'une série, c'est-à-dire sa convergence ou sa divergence n'est pas altérée si l'on modifie le terme général de la série,  $\vec{u}_n$  pour un nombre fini de valeurs de  $n$ , plus précisément on a le résultat suivant.

**Proposition 1.1.3.** — Soient  $(\vec{u}_n)_{n \in \mathbf{N}}$  et  $(\vec{v}_n)_{n \in \mathbf{N}}$  des suites à valeurs dans  $(\mathbf{E}, \|\cdot\|)$  telles que l'ensemble des entiers naturels  $n$  tels que  $\vec{u}_n \neq \vec{v}_n$  soit fini. Alors les séries  $\sum \vec{u}_n$  et  $\sum \vec{v}_n$  sont de même nature.

*Preuve de la proposition 1.1.3.* — l'ensemble des entiers naturels  $n$  tels que  $\vec{u}_n \neq \vec{v}_n$  étant une partie finie de  $\mathbf{N}$ , elle admet un plus grand élément, notons le  $n_0$ . Notons par ailleurs  $U_n$  et  $V_n$  les sommes partielles d'ordre  $n$  respectives de  $\sum \vec{u}_n$  et  $\sum \vec{v}_n$ . On a pour tout entier naturel  $n \geq n_0 + 1$ ,

$$V_n = U_n + \sum_{k=0}^{n_0} v_k - u_k.$$

Les suites des sommes partielles des deux séries diffèrent donc à partir d'un certain rang d'une constante, elles sont donc de même nature, d'où le résultat.  $\square$

**Exemples** —

1. Soit  $\vec{a}$  un élément de  $\mathbf{E}$ . Etudions la série  $\sum \vec{a}$ .

2. Même question pour la série  $\sum_{n \geq 1} \frac{1}{n(n+1)} \cdot \vec{a}$

3. SÉRIES GÉOMÉTRIQUES — On prend pour  $\mathbf{E}, \mathbf{C}$ . Etudions la série  $\sum \rho^n$ , où  $\rho$  est un nombre complexe.

4. On prend pour  $\mathbf{E}, \mathcal{M}_2(\mathbf{R})$ . Etudions suivant la valeur du réel  $a$ , la série  $\sum M^n$ , où  $M = \begin{pmatrix} a & 2a \\ 2a & a \end{pmatrix}$ .

**Proposition 1.1.4.** — Soit  $(\vec{u}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $\mathbf{E}$ . Si la série  $\sum \vec{u}_n$  converge alors  $\vec{u}_n \xrightarrow{n \rightarrow +\infty} \vec{0}_{\mathbf{E}}$ . Autrement dit, une condition nécessaire pour qu'une série converge est que son terme général tende vers zéro.

!

Cette condition n'est pas suffisante, il existe des séries dont le terme général tend vers zéro et qui divergent.



Ainsi la série — dite série *harmonique* —  $\sum_{n \geq 1} \frac{1}{n}$  diverge.

**Terminologie** — Une série dont le terme général ne tend pas vers 0 est donc divergente, on dit qu'elle est *grossièrement divergente*;

*Preuve de la proposition 1.1.4.* —

D'où le résultat. □

### À quoi servent les séries ?

La définition 1.1.1. laisse accroire que l'étude des séries se ramène à celle des suites, ce qui clorait le chapitre. En fait c'est presque l'inverse : on va développer des techniques puissantes d'étude directe des séries qui permettront d'étudier des suites. Explications !

- Pour montrer que  $\frac{e^n}{n!} \xrightarrow{n \rightarrow +\infty} 0$ , on peut montrer que la série  $\sum \frac{e^n}{n!}$  converge, ce qui se fait, comme on le verra bientôt, en deux lignes.
- Toute suite peut, dans le même ordre d'idée, être vue comme la suite des sommes partielles d'une série, précisément :

**Proposition 1.1.5** — Soit  $(\vec{s}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $(\mathbf{E}, \|\cdot\|)$ . Il existe une et une seule série  $\sum \vec{u}_n$ , dont  $(\vec{s}_n)_{n \in \mathbf{N}}$  soit la suite des sommes partielles, c'est la série définie par :

$$\begin{aligned} \vec{u}_0 &= \cdots \\ \vec{u}_n &= \cdots, \text{ pour tout } n \geq 1. \end{aligned}$$

### Remarques 1.1.6 —

1. La série  $\sum \vec{u}_n$  ainsi définie est dite *téléscopique*, ou *série domino*. Les séries téléscopiques comptent avec les séries géométriques, parmi les rares séries dont on sache calculer la suite des sommes partielles.
2. On notera dans le même ordre d'idées, que l'opération, qui à une suite  $(u_n)_{n \in \mathbf{N}}$ , par exemple de réels, associe la suite  $(u_n - u_{n-1})_{n \in \mathbf{N}^*}$  et celle qui lui associe  $\left(\sum_{k=0}^n u_k\right)$ , sont, à peu de choses près, des opérations inverses l'une de l'autre. Elles jouent pour les suites des rôles analogues à ceux tenus pour les applications, respectivement par la dérivation et la primitivation. Cette analogie est très fructueuse dans l'étude des séries ; elle permet de comprendre de nombreuses solutions à des problèmes qui, sans elles, eussent immensément paru artificielles et parachutées. Par exemple, en premier lieu, la facilité que l'on a à sommer des

séries télescopiques, est la traduction, en terme de suites, de la simplicité du calcul de la primitive d'une dérivée.

$f$	$(u_n)_{n \in \mathbf{N}}$
$f'$	$(u_n - u_{n-1})_{n \in \mathbf{N}^*}$
$x \mapsto \int_0^x f(t) dt$	$\left( \sum_{k=0}^n u_k \right)_{n \in \mathbf{N}}$

## 1.2 Convergence absolue

Il a été vu l'an passé que pour une série réelle,  $\sum u_n$ , si  $\sum |u_n|$  converge alors  $\sum u_n$  converge. La convergence de la série des valeurs absolues de  $|u_n|$  est appelée naturellement convergence absolue. C'est cette notion que nous allons étendre à présent aux séries à valeurs dans  $\mathbf{E}$ .

Au préalable rapelons un théorème de comparaison vue en MPSI dont nous avons besoin et que nous reverrons en détail dans la partie suivante.

**Rappel** — Soient  $(u_n)_{n \in \mathbf{N}}$  et  $(v_n)_{n \in \mathbf{N}}$  des suites de réels positifs et  $k$  un élément de  $\mathbf{R}_+^*$ . On suppose que pour tout entier naturel  $n$ ,

$$0 \leq u_n \leq kv_n$$

Alors si  $\sum v_n$  converge alors  $\sum u_n$  converge.

C'est ce théorème qui a permis de montrer pour une suite réelle  $(u_n)_{n \in \mathbf{N}}$  que  $\sum |u_n|$  converge alors  $\sum u_n$  converge. Petit retour en arrière. On définit les suites de réels positifs  $(u_n^+)_{n \in \mathbf{N}}$  et  $(u_n^-)_{n \in \mathbf{N}}$  par :

$$u_n^+ = \begin{cases} u_n & \text{si } u_n \geq 0, \\ 0 & \text{si } u_n < 0, \end{cases}$$

$$u_n^- = \begin{cases} -u_n & \text{si } u_n \leq 0, \\ 0 & \text{si } u_n > 0, \end{cases}$$

pour tout entier  $n \geq 0$ .

Remarquons alors que

pour tout entier  $n \geq 0$ ,  $u_n = \dots\dots\dots$ ,  $|u_n| = \dots\dots\dots$  et en particulier

$$0 \leq u_n^\pm \leq |u_n|$$

Donc si  $\sum |u_n|$  converge alors  $\dots\dots\dots$

Supposons à présent que  $(u_n)_{n \in \mathbf{N}}$  soit une suite de nombre complexes et que  $\sum |u_n|$  converge...

Donc  $\sum u_n$  converge.

Généralisons !

**Définition, proposition 1.2.1.** — CONVERGENCE ABSOLUE —

Soit  $(\vec{u}_n)_{n \in \mathbf{N}}$  une suite à valeurs dans  $\mathbf{E}$ . Si pour une norme  $\|\cdot\|$ ,  $\sum \|\vec{u}_n\|$  converge alors, pour toute norme  $\|\cdot\|_{\mathbf{E}}$  sur  $\mathbf{E}$  la série  $\sum \|\vec{u}_n\|_{\mathbf{E}}$  converge.

Si c'est le cas, on dit que la série  $\sum \vec{u}_n$  converge absolument.

**Preuve de la proposition 1.2.1.** —

Supposons la convergence de la série  $\sum \|\vec{u}_n\|$ .

□

**Proposition 1.2.2.** — Soit  $(\vec{u}_n)_{n \in \mathbf{N}}$  une suite d'éléments de  $\mathbf{E}$ .

- si la série  $\sum \vec{u}_n$  converge absolument alors elle converge.
- de plus si tel est le cas, on a :

$$\left\| \sum_{n=0}^{+\infty} \vec{u}_n \right\| \leq \sum_{n=0}^{+\infty} \|\vec{u}_n\|.$$

*Preuve de la proposition 1.2.2.* —

Supposons que la série  $\sum \vec{u}_n$  converge absolument. la proposition 1.2.1. nous laisse le choix de la norme, prenons  $n_{\mathcal{B}, \infty}$ , où  $\mathcal{B}$  est une base quelconque de  $\mathbf{E}$ . Notons pour tout  $n \in \mathbf{N}$ ,  $u_{i,n}$  la  $i^e$  composante de  $\vec{u}_n$  dans  $\mathcal{B}$  et  $S_{i,n}$  la somme

partielle d'ordre  $n$  de la série  $\sum u_{i,n}$ .



**Cas particulier** — Soit  $(u_n)_{n \in \mathbf{N}}$  une suite de nombres complexes. Si  $\sum |u_n|$  converge, alors  $\sum u_n$  converge.

**Remarque** — Il existe des séries convergentes non absolument convergentes. Par exemple.....

La proposition 1.2.1. nous incite donc à étudier les séries à terme général élément de  $\mathbf{R}_+$ , ou, comme on dit, les séries à termes positifs

## SÉRIES À TERMES POSITIFS

### 2.1 Convergence des séries à termes positifs

**Définition 2.1.1.** — On appelle série à termes positifs (resp. strictement positifs), toute série à termes réels,  $\sum u_n$  telle que pour tout entier naturel  $n$ ,  $u_n \in \mathbf{R}_+$  (resp.  $u_n \in \mathbf{R}_+^*$ ).

On a le résultat immédiat mais crucial suivant :

**Proposition 2.1.2.** — Une série à termes positifs  $\sum u_n$  converge si et seulement si la suite de ses sommes partielles  $(S_n)_{n \in \mathbf{N}}$  est majorée.

*Preuve de la proposition 2.1.2. —*

**Remarque** — Si une série quelconque converge alors la suite de ses sommes partielles est majorée, car convergente, la réciproque est évidemment fautive comme le montre l'étude de la série à termes réels  $\sum \dots$

Donnons pour finir un exemple d'étude de série qui n'utilise que ce dernier résultat.

**Exemple** — Pour tout entier naturel  $n$ ,  $p_n$  désigne le nombre de chiffres dans l'écriture décimale de  $n$ . Étudions la série

$$\sum_{n \geq 0} p_n^{-p_n}.$$

## 2.2 Théorèmes de comparaison directe

Commençons par le théorème de comparaison déjà mentionné.

Proposition 2.2.1. — THÉORÈME DE COMPARAISON DIRECTE —

Soient  $\sum u_n$  et  $\sum v_n$  des séries à termes positifs telles que pour tout entier naturel  $n$ ,

$$u_n \leq v_n.$$

Alors on a :

- Si la série  $\sum v_n$  converge, alors  $\sum u_n$  converge ;
- Si la série  $\sum u_n$  diverge, alors  $\sum v_n$  diverge.

*Preuve de la proposition 2.2.1* — Le second point est la contraposée du premier, il suffit donc de prouver le premier. Supposons donc que  $\sum v_n$  converge.

D'où le résultat. □

Ce résultat est fondamental, pour autant il ne s'applique tel quel que rarement dans des cas concrets. Nous allons donc en déduire une succession de corollaires d'un usage mieux adapté aux cas pratiques.

**Proposition 2.2.2.** — GÉNÉRALISATION DU THÉORÈME DE COMPARAISON DIRECTE —

Soient  $(u_n)_{n \in \mathbb{N}}$  et  $(v_n)_{n \in \mathbb{N}}$  des suites de réels pour lesquelles il existe un entier naturel  $n_0$  et un réel  $k$ , tels que pour tout entier  $n \geq n_0$ ,

- a)  $u_n \geq 0$  et  $v_n \geq 0$  ;
- b)  $u_n \leq kv_n$ .

Si la série  $\sum v_n$  converge, alors  $\sum u_n$  converge, (contraposée : Si la série  $\sum u_n$  diverge, alors  $\sum v_n$  diverge).

**Terminologie** — On traduit souvent de façon concise le a) en disant que les séries  $\sum u_n$  et  $\sum v_n$  sont *positives à partir d'un certain rang* ou plus joliment *ultimement positives* et le b) en disant que  $u_n \leq kv_n$  à *partir d'un certain rang*.

*Preuve de la proposition 2.2.2.* — Supposons que  $\sum v_n$  converge. Alors  $\sum kv_n$  converge puisque la suite des sommes partielles de cette série est égale à  $k$  fois celle de  $\sum v_n$ . Enfin d'après la proposition 1.1.3., on ne modifie pas la nature des séries  $\sum u_n$  et  $\sum kv_n$  en remplaçant les  $n_0$  premiers termes de ces séries par 0, ainsi modifiées ces séries relèvent de la proposition 2.2.2., qui assure la convergence de  $\sum u_n$ . □

Pour être plus maniable que la proposition 2.2.1. ce résultat reste néanmoins d'une utilisation difficile, les corollaires successifs suivants sont d'un usage plus agréables.

**Corrolaire 2.2.3.** — Soient  $(u_n)_{n \in \mathbb{N}}$  et  $(v_n)_{n \in \mathbb{N}}$  des suites de réels, positives à partir d'un certain rang, telles que

$$u_n = O_{n \rightarrow \infty}(v_n).$$

Alors on a :

- Si la série  $\sum v_n$  converge, alors  $\sum u_n$  converge ;
- Si la série  $\sum u_n$  diverge, alors  $\sum v_n$  diverge.

*Preuve du corollaire 2.2.3.* —

**Corrolaire 2.2.4.** — Soient  $(u_n)_{n \in \mathbb{N}}$  et  $(v_n)_{n \in \mathbb{N}}$  des suites de réels, positives à partir d'un certain rang, telles que

$$u_n = o_{n \rightarrow \infty}(v_n).$$

Alors on a :

- Si la série  $\sum v_n$  converge, alors  $\sum u_n$  converge ;

— Si la série  $\sum u_n$  diverge, alors  $\sum v_n$  diverge.

Preuve du corollaire 2.2.4. —

**Corrolaire 2.2.5.** — Soient  $(u_n)_{n \in \mathbb{N}}$  et  $(v_n)_{n \in \mathbb{N}}$  des suites de réels, positives à partir d'un certain rang, telles que

$$u_n \underset{n \rightarrow \infty}{\sim} v_n.$$

Alors les séries  $\sum u_n$  et  $\sum v_n$  sont de même nature.

Preuve du corollaire 2.2.5. —

**Remarque** —

1. Dans la proposition 2.2.5. il suffit de supposer qu'une seule des deux suites est positive à partir d'un certain rang ; l'autre l'est alors aussi, puisque  $u_n \underset{n \rightarrow \infty}{\sim} v_n$ . Cette remarque est très utile, comme on le verra, dans la pratique.
2. Dans les résultats 2.2.2–5., on peut, bien évidemment remplacer le mot *positif* par *négatif*, en effet on passe du cas positif au cas négatif en remplaçant le terme général des séries par son opposé... En fait ces résultats s'appliquent aux séries *de signe constant à partir d'un certain rang*.

Pour utiliser ces résultats de comparaison, encore faut-il connaître la nature de certaines séries *de référence* auxquelles comparer les séries à étudier. Le programme en retient deux types : les *séries géométriques* et les *séries de Riemann*

SÉRIES GÉOMÉTRIQUES —

Soit  $\rho$  un réel positif. Étudions la série associée à la suite géométrique  $(\rho^n)_{n \in \mathbb{N}}$ , on parle encore de série géométrique. Pour tout entier naturel  $n$ ,

Donc

$$\sum \rho^n \text{ converge si et seulement si } \dots$$

**Exemple** —

Étudions la série suivante  $\sum_{n \geq 1} \left( \frac{1}{2} + \frac{\cos(3n)}{8} + \frac{\ln(n)}{n} \right)^n$ .

SÉRIES DE RIEMANN —

Soit  $\alpha$  un réel. Etudions la série

$$\sum_{n \geq 1} \frac{1}{n^\alpha},$$

appelée série de Riemann.

On a vu que pour  $\alpha = 1$  la série diverge. Par ailleurs, on peut observer que pour  $\alpha \leq 0$ , la série est grossièrement divergente.

Nous allons dans la suite supposer que  $\alpha \neq 1$  et rechercher une série télescopique de terme général équivalent à celui de la série de Riemann.

En conclusion :

$\sum_{n \geq 1} \frac{1}{n^\alpha} \text{ converge si et seulement si } \dots$
---

**Exemples** —



1. Etudions la série suivante

$$\sum_{n \geq 1} \frac{1}{n^2 + 1 + \frac{1}{2} \cos(e^n)}.$$

2. Etudions la série suivante

$$\sum_{n \geq 1} \frac{\ln n}{n}.$$

3. Etudions la série suivante

$$\sum_{n \geq 2} \frac{1}{\sqrt{n} \ln n}.$$

4. Etudions la série suivante

$$\sum_{n \geq 2} \frac{1}{n^2 \ln n}.$$

5. Etudions la série suivante

$$\sum_{n \geq 2} \frac{(\ln n)^{12}}{n^2}.$$

Nous allons donner sous les hypothèses des corollaires 2.2.3–5 des renseignements complémentaires sur les restes et les sommes partielles des séries comparées.

**Proposition 2.2.6.** — Soient  $(u_n)_{n \in \mathbb{N}}$  et  $(v_n)_{n \in \mathbb{N}}$  des suites de réels, positives à partir d'un certain rang.

1. RELATION DE DOMINATION

On suppose que  $u_n = \underset{n \rightarrow \infty}{\mathrm{O}}(v_n)$ .

- Cas convergent

On suppose que  $\sum v_n$  converge, alors, on l'a vu,  $\sum u_n$  converge et les restes de ces séries vérifient :

$$\sum_{k=n+1}^{+\infty} u_k = \underset{n \rightarrow \infty}{\mathrm{O}} \left( \sum_{k=n+1}^{+\infty} v_k \right);$$

- Cas divergent

On suppose que  $\sum u_n$  diverge, alors, on l'a vu,  $\sum v_n$  diverge et les sommes partielles de ces séries vérifient :

$$\sum_{k=0}^n u_k = \underset{n \rightarrow \infty}{\mathrm{O}} \left( \sum_{k=0}^n v_k \right).$$

2. RELATION DE PRÉPONDÉRENCE

On suppose que  $u_n = \underset{n \rightarrow \infty}{\mathrm{o}}(v_n)$ .

- Cas convergent

On suppose que  $\sum v_n$  converge, alors, on l'a vu,  $\sum u_n$  converge et les restes de ces séries vérifient :

$$\sum_{k=n+1}^{+\infty} u_k = \underset{n \rightarrow \infty}{\mathrm{o}} \left( \sum_{k=n+1}^{+\infty} v_k \right);$$

- Cas divergent

On suppose que  $\sum u_n$  diverge, alors, on l'a vu,  $\sum v_n$  diverge et les sommes partielles de ces séries vérifient :

$$\sum_{k=0}^n u_k = \underset{n \rightarrow \infty}{\mathrm{o}} \left( \sum_{k=0}^n v_k \right).$$

3. RELATION D'ÉQUIVALENCE

On suppose que  $u_n \underset{n \rightarrow \infty}{\sim} v_n$ . Les séries  $\sum u_n$  et  $\sum v_n$  sont de même nature ;

- Cas convergent

On suppose que  $\sum u_n$  et  $\sum v_n$  convergent, alors les restes de ces séries vérifient :

$$\sum_{k=n+1}^{+\infty} u_k \underset{n \rightarrow \infty}{\sim} \sum_{k=n+1}^{+\infty} v_k;$$

- Cas divergent

On suppose que  $\sum u_n$  et  $\sum v_n$  divergent, alors les sommes partielles de ces séries vérifient :

$$\sum_{k=0}^n u_k \underset{n \rightarrow \infty}{\sim} \sum_{k=0}^n v_k.$$

**Remarque** — Dans le cas convergent cette proposition ne donne aucun renseignement sur les sommes partielles, dans le cas divergent les restes des séries n'existent même pas.

Récapitulons. Pour tout naturel  $n$ , notons respectivement  $S_n$ ,  $S'_n$  les sommes partielles de  $\sum u_n$  et  $\sum v_n$ , et sous réserve d'existence  $R_n$ ,  $R'_n$  leurs restes respectifs.

	suites	restes	sommes partielles
séries convergentes			
	$u_n = O(v_n)$	$R_n = O(R'_n)$	rien
	$u_n = o(v_n)$	$R_n = o(R'_n)$	rien
	$u_n \sim v_n$	$R_n \sim R'_n$	rien
séries divergentes			
	$u_n = O(v_n)$	non définis	$S_n = O(S'_n)$
	$u_n = o(v_n)$	non définis	$S_n = o(S'_n)$
	$u_n \sim v_n$	non définis	$S_n \sim S'_n$

*Preuve de la proposition 2.2.6. —*



**Corollaire 2.2.7.** — Si  $(v_n)_{n \in \mathbf{N}}$  est une suite de réels, positives à partir d'un certain rang et  $(\vec{u}_n)$  une suite d'éléments de  $\mathbf{E}$ . On suppose que  $\vec{u}_n = \underset{n \rightarrow \infty}{\mathrm{O}}(v_n)$  (resp.  $\underset{n \rightarrow \infty}{\mathrm{O}}(v_n)$ ).

- Cas convergent

On suppose que  $\sum v_n$  converge, alors,  $\sum \vec{u}_n$  converge et les restes de ces séries vérifient :

$$\sum_{k=n+1}^{+\infty} \vec{u}_k = \underset{n \rightarrow \infty}{\mathrm{O}} \left( \sum_{k=n+1}^{+\infty} v_k \right) \text{ (resp. } \underset{n \rightarrow \infty}{\mathrm{O}} \left( \sum_{k=n+1}^{+\infty} v_k \right));$$

- Cas divergent

On suppose que  $\sum \vec{u}_n$  diverge, alors, on l'a vu,  $\sum v_n$  diverge et les sommes partielles de ces séries vérifient :

$$\sum_{k=0}^n \vec{u}_k = \underset{n \rightarrow \infty}{\mathrm{O}} \left( \sum_{k=0}^n v_k \right) \text{ (resp. } \underset{n \rightarrow \infty}{\mathrm{O}} \left( \sum_{k=0}^n v_k \right)).$$

*Preuve du corollaire 2.3.7. —*

**Application** — THÉORÈMES D'ERNESTO CESÀRO — Les théorèmes de Cesàro déjà rencontrés, ne sont pas au programme, toutefois la proposition 2.2.6. redonne dans le cas d'une suite réelle les différentes versions du théorème de Cesàro. Voyons cela !

Soit  $(u_n)_{n \in \mathbf{N}}$  une suite de réels. Soit la suite de réels  $(v_n)_{n \in \mathbf{N}}$  (suite des moyennes) où, pour tout  $n \in \mathbf{N}$ ,

$$v_n = \frac{1}{n+1} \sum_{k=0}^n u_k.$$

1. On suppose que  $(u_n)_{n \in \mathbf{N}}$  admet une limite  $\ell$ . Montrer que  $(v_n)_{n \in \mathbf{N}}$  admet  $\ell$  comme limite.
2. On suppose que  $u_n \xrightarrow{n \rightarrow +\infty} +\infty$  (resp.  $-\infty$ ). Montrer que  $v_n \xrightarrow{n \rightarrow +\infty} +\infty$  (resp.  $-\infty$ ).
3. Soit  $\sum \alpha_n$  une série à termes strictement positifs divergente. Soit la suite de réels  $(w_n)_{n \in \mathbf{N}}$  (suite des moyennes pondérées de  $(u_n)_{n \in \mathbf{N}}$ ) où, pour tout  $n \in \mathbf{N}$ ,

$$w_n = \frac{1}{\sum_{k=0}^n \alpha_k} \sum_{k=0}^n \alpha_k u_k.$$

(a) On suppose que  $(u_n)_{n \in \mathbf{N}}$  admet une limite  $\ell$ . Montrer que  $(w_n)_{n \in \mathbf{N}}$  admet  $\ell$  comme limite.

(b) On suppose que  $u_n \xrightarrow{n \rightarrow +\infty} +\infty$  (resp.  $-\infty$ ). Montrer que  $w_n \xrightarrow{n \rightarrow +\infty} +\infty$  (resp.  $-\infty$ ).

*Solution* —

**Exercice 2.2.8.** — Utilisons la proposition 2.2.6. pour déterminer des équivalents simples de sommes partielles et de restes de séries.

- EQUIVALENT D'UN RESTE

Soit un réel  $a > 1$ . Donner un équivalent simple de  $\sum_{k=n}^{+\infty} \frac{1}{k^a}$ , lorsque  $n$  tend vers  $+\infty$ .

- EQUIVALENTS DE SOMME PARTIELLES

- Utilisation de « primitives discrètes »

Donner un équivalent simple de  $\sum_{k=1}^n \frac{1}{k}$ , lorsque  $n$  tend vers  $+\infty$ .

Soit un réel  $a < 1$ . Donner un équivalent simple de  $\sum_{k=1}^n \frac{1}{k^a}$ , lorsque  $n$  tend vers  $+\infty$ .

— « *Croissance rapide* »

Donner un équivalent simple de  $\sum_{k=1}^n k^k$ , lorsque  $n$  tend vers  $+\infty$ .



## 2.3 Règles de d'Alembert

**Proposition 2.3.1.** — RÈGLE DE D'ALEMBERT —

Soit  $\sum u_n$  une série à termes strictement positifs.

1. Si il existe un entier naturel  $n_0$ , tel que pour tout entier  $n \geq n_0$ ,

$$\frac{u_{n+1}}{u_n} \geq 1,$$

alors  $\sum u_n$  diverge.

2. Si il existe un entier naturel  $n_0$  et un élément  $k$  de  $[0, 1[$ , tels que pour tout entier  $n \geq n_0$ ,

$$\frac{u_{n+1}}{u_n} \leq k,$$

alors  $\sum u_n$  converge.

Preuve de la proposition 2.3.1. —

!

Dans le second point de la règle de d'Alembert, il ne faut surtout pas oublier le réel  $k$ . On peut très bien avoir, pour tout  $n \in \mathbf{N}$ ,  $\frac{u_{n+1}}{u_n} < 1$ , sans que la série  $\sum u_n$  converge, comme c'est le cas pour la série  $\sum \dots\dots\dots$

**Remarque** — Dans la proposition 2.3.1., on notera que l'on peut se contenter de supposer que  $u_n$  est strictement positif, à partir du rang  $n_0$ .

Il existe une forme asymptotique de la règle de d'Alembert d'un emploi souvent plus aisé.

**Proposition 2.3.2.** — RÈGLE DE D'ALEMBERT, SECONDE FORME —

Soit  $\sum u_n$  une série à termes strictement positifs à partir d'un certain rang. On suppose que  $\frac{u_{n+1}}{u_n}$  tend vers un élément  $\ell$  de  $\mathbf{R} \cup \{+\infty\}$ , lorsque  $n$  tend vers  $+\infty$ . Alors :

1. Si  $\ell < 1$  alors la série  $\sum u_n$  converge.
2. Si  $\ell > 1$  alors la série  $\sum u_n$  diverge.

**Remarques** —

- Dans le cas où  $\ell = 1$ , il n'y a pas de résultats généraux. La série peut diverger ou converger, comme par exemple  $\dots$

Remarquons toutefois que si  $\frac{u_{n+1}}{u_n}$  tend vers 1, par valeurs supérieures, lorsque  $n$  tend vers  $+\infty$ , alors la série

diverge, en effet :

- On se gardera de croire que, pour toute série  $\sum u_n$ , à termes strictement positifs,  $\frac{u_{n+1}}{u_n}$  tend vers une limite, lorsque  $n$  tend vers  $+\infty$

*Preuve de la proposition 2.3.2. —*

La proposition 2.3.2. est bien adaptée aux séries dont le terme général contient des produits, étudions par exemple la série  $\sum \frac{2^n}{n!}$ .

**Exercice 2.3.3.** — Soit  $\sum u_n$  et  $\sum v_n$  des séries. On suppose qu'il existe un entier naturel  $n_0$  tel que, pour tout entier  $n \geq n_0$ ,

- $u_n > 0$  et  $v_n > 0$  (séries à termes strictement positifs à partir d'un certain rang).
- $\frac{u_{n+1}}{u_n} \leq \frac{v_{n+1}}{v_n}$ .

- Montrer que si la série  $\sum v_n$  converge alors la série  $\sum u_n$  converge.
- Quel résultat retrouve-t-on lorsque  $v_n$  est une série géométrique.
- étudier la série de terme général

$$a_n = \sqrt{n!} \prod_{p=1}^n \sin\left(\frac{1}{\sqrt{p}}\right).$$

*solution de l'exercice 2.3.3. —*

## 2.4 Comparaison à une intégrale

Observons que le calcul de  $\int_0^n \frac{dt}{t^2}$  est une vaste trivialité, par contre on ne sait pas évaluer  $\sum_{k=1}^n \frac{1}{k^2}$ . Cette facilité que l'on a à évaluer des primitives, dont on a déjà, par ailleurs, mentionné l'analogie avec les sommes partielles fait l'intérêt du résultat qui va suivre.

**Proposition 2.4.1.** — *Soit  $f$  une application de  $\mathbf{R}_+$  dans  $\mathbf{R}$ , continue par morceaux. On suppose que*

1. *L'application  $f$  est à valeur positives ;*
2. *L'application  $f$  est décroissante.*

*Alors la série  $\sum f(n)$  converge si et seulement si l'intégrable  $\int_0^{+\infty} f(t)dt$  converge (c'est-à-dire que  $\int_0^x f(t)dt$  admet une limite lorsque  $x$  tend vers  $+\infty$ ).*

On laisse le soin au lecteur d'adapter ce résultat à une série associée à une suite indexée par  $\llbracket n_0, +\infty \rrbracket$ .

*Preuve de la proposition 2.4.1. —*

**Remarque** — Il existe au programme un théorème plus fort que 2.4.1., nous le présenterons dans le cours sur l'intégrale.

Outre le théorème 2.4.1. qu'il faut connaître, la comparaison série/intégrale permet de déterminer des équivalents de sommes partielles et de restes de séries, offrant ainsi des méthodes concurrentes à celles développées dans le paragraphe 2. Illustrons notre propos.

**Exemples** —

## • EQUIVALENTS DE SOMMES PARTIELLES

— *Cas d'une application décroissante*

Déterminer un équivalent simple de  $\sum_{k=1}^n \frac{1}{k}$ , lorsque  $n$  tend vers  $+\infty$ .

La clef du succès réside dans la réalisation de **deux** beaux dessins.

— *Cas d'une application croissante*

Déterminer un équivalent simple de  $\sum_{k=1}^n k^{20015}$ , lorsque  $n$  tend vers  $+\infty$ .

- EQUIVALENT D'UN RESTE

Déterminer un équivalent simple de  $\sum_{k=n}^{+\infty} \frac{1}{k^2}$ , lorsque  $n$  tend vers  $+\infty$ .

## 2.5 Un exemple récapitulatif

Nous allons étudier la série

$$\sum_{n \geq 2} \frac{1}{n^a (\ln(n))^b},$$

où  $a$  et  $b$  sont des réels.

L'étude de cette série est donnée à titre *d'exercice*, elle ne figure pas au programme. Il faut donc savoir retrouver très

vite ce qui va suivre, mais on ne peut l'utiliser au concours.





## 2.6 Applications des théorèmes de comparaisons

Désignons dans ce paragraphe par  $(\mathbf{E}, \|\cdot\|)$  un e.v.n. sur le corps  $\mathbf{K}$  ( $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{C}$ ).

**Définition 2.6.1.** — Soient  $\sum \vec{u}_n$  et  $\sum \vec{v}_n$  des séries de termes généraux éléments de  $\mathbf{E}$  et  $\lambda$  un élément de  $\mathbf{K}$ .

- On appelle série somme des séries  $\sum \vec{u}_n$  et  $\sum \vec{v}_n$ , la série de terme général  $\vec{u}_n + \vec{v}_n$ .  
On la note  $\sum \vec{u}_n + \sum \vec{v}_n$ .
- On appelle produit de la série  $\sum \vec{u}_n$  par le scalaire  $\lambda$ , la série de terme général  $\lambda \vec{u}_n$ .  
On la note  $\lambda \cdot \sum \vec{u}_n$ .

Le moment est enfin venu de se pencher sur la définition précise d'une série. On appelle série d'éléments de  $\mathbf{E}$ , tout élément  $(\vec{u}_n, \vec{U}_n)_{n \in \mathbf{N}}$  de  $\mathcal{F}(\mathbf{N}, \mathbf{E} \times \mathbf{E})$  (c'est-à-dire toute suite à valeurs dans  $\mathbf{E} \times \mathbf{E}$ ) tel que, pour tout entier naturel  $n$ ,

$$U_n = \sum_{k=0}^n u_k.$$

Intuitivement la donnée d'une série  $(\vec{u}_n, \vec{U}_n)_{n \in \mathbf{N}}$  consiste en la donnée d'une suite  $(\vec{u}_n)_{n \in \mathbf{N}}$  et d'un procédé de sommation de cette suite, la suite  $(\vec{U}_n)_{n \in \mathbf{N}}$ .

On vérifie sans mal

**Proposition 2.6.2.** — L'ensemble des séries d'éléments de  $\mathbf{E}$  est un sous-espace vectoriel de  $\mathcal{F}(\mathbf{N}, \mathbf{E} \times \mathbf{E})$ .

On remarque alors, que la somme de deux séries et le produit d'une série par un élément de  $\mathbf{K}$  ne sont rien d'autres que les lois induites sur l'ensemble des séries par les lois de l'espace vectoriel  $\mathcal{F}(\mathbf{N}, \mathbf{E} \times \mathbf{E})$ .

Reste à interpréter la convergence d'une série. Une série  $\sum \vec{u}_n$  converge aux sens de la définition 1.1.1., si et seulement si en tant que suite  $(\vec{u}_n, \vec{U}_n)_{n \in \mathbf{N}}$  à valeurs dans  $(\mathbf{E} \times \mathbf{E}, n'_\infty)^1$  elle converge. En effet

**Proposition 2.6.3.** — L'ensemble des séries d'éléments de  $\mathbf{E}$  convergentes est un sous-espace vectoriel de l'ensemble des séries d'éléments de  $\mathbf{E}$ .

De plus pour tout  $\alpha$  et tout  $\beta$  éléments de  $\mathbf{K}$  et pour toute série  $\sum \vec{u}_n$  et toute série  $\sum \vec{v}_n$  convergentes, la somme de la séries  $\alpha \cdot \sum \vec{u}_n + \beta \cdot \sum \vec{v}_n$  vaut  $\alpha \cdot \sum_{n=0}^{+\infty} \vec{u}_n + \beta \cdot \sum_{n=0}^{+\infty} \vec{v}_n$ .

*Preuve de la proposition 2.6.3.* — Nous laissons cette preuve facile en exercice.

**Proposition 2.6.4.** — L'ensemble des séries d'éléments de  $\mathbf{E}$  absolument convergentes est un sous-espace vectoriel de l'ensemble des séries d'éléments de  $\mathbf{E}$ . et même un sous-espace vectoriel de l'ensemble des séries d'éléments de  $\mathbf{E}$  convergentes.

1.  $n'_\infty$  désigne ici bien sûr, la norme définie par, pour tout couple  $(\vec{x}, \vec{y})$  d'éléments de  $\mathbf{E}$ ,  $n'_\infty(\vec{x}, \vec{y}) = \sup\{\|\vec{x}\|, \|\vec{y}\|\}$ .

*Preuve de la proposition 2.6.4. —*

Soient une série  $\sum u_n$  convergente et une série  $\sum v_n$  divergente. Notons leur somme  $\sum w_n$ .

On a  $\sum v_n = \sum w_n - \sum u_n$ , donc  $\sum w_n$  diverge, en effet sinon,  $\sum v_n$  convergerait (cf. 2.6.3.).

De même 2.6.4. montre que la somme d'une série absolument convergente et d'une série non absolument convergente n'est pas absolument convergente.

On se gardera de conclure quant à la nature de la somme de deux séries divergentes, qui peut être convergente ou divergente. Donner des exemples dans chaque cas...

Récapitulons

$\sum u_n$	$\sum v_n$	$\sum w_n$
<i>cv</i>	<i>cv</i>	<i>cv</i>
<i>div</i>	<i>cv</i>	<i>div</i>
<i>cv</i>	<i>div</i>	<i>div</i>
<i>div</i>	<i>div</i>	?

**Exercice** — Soit  $\sum_{n \geq 1} u_n$  une série à termes positifs convergente. Quelle est la nature de la série  $\sum_{n \geq 1} \frac{\sqrt{u_n}}{n}$ .

## SÉRIES RÉELLES

### 3.1 Introduction

Commençons par une remarque concernant une série  $\sum \vec{u}_n$  à termes éléments d'un espace  $K$ -vectoriel  $\mathbf{E}$  de dimension finie  $d$  non nulle ( $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{C}$ ). Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_d)$  une base de  $\mathbf{E}$  notée  $\mathcal{B}$ . Notons pour  $i = 1, \dots, d$   $((u_{i,n})_{n \in \mathbf{N}})$  la  $i^{\text{e}}$  suite composante de la suite  $(\vec{u}_n)_{n \in \mathbf{N}}$  dans la base  $\mathcal{B}$ . On a alors le résultat presque évident suivant, qui découle de l'étude de la convergence d'une suite à valeurs dans un espace vectoriel de dimension finie :

**Proposition 3.1.1** — *La série  $\sum \vec{u}_n$  converge si et seulement si, pour tout élément  $i$  de  $\{1, 2, \dots, d\}$ , la série d'élément de  $\mathbf{K}$ ,  $\sum u_{i,n}$  converge.*

Si tel est le cas, alors :

$$\sum_{n=0}^{+\infty} \vec{u}_n = \sum_{i=0}^d \left( \sum_{n=0}^{+\infty} u_{i,n} \vec{e}_i \right).$$

L'étude d'une série d'éléments de  $\mathbf{E}$  se ramène donc à l'étude de  $d$  séries d'éléments de  $\mathbf{K}$ .

Dans le cas où  $\mathbf{K} = \mathbf{C}$  la proposition précédente, appliquée à  $\mathbf{C}$  considéré comme un  $\mathbf{R}$ -espace vectoriel de dimension 2, muni de la base  $(1, i)$  affirme que pour tout élément  $i$  de  $\{1, \dots, d\}$ ,  $\sum u_{i,n}$  converge si et seulement si  $\sum \operatorname{Re}(u_{i,n})$  et  $\sum \operatorname{Im}(u_{i,n})$  convergent, et si c'est le cas,

$$\sum_{n=0}^{+\infty} u_{i,n} = \sum_{n=0}^{+\infty} \operatorname{Re}(u_{i,n}) + i \sum_{n=0}^{+\infty} \operatorname{Im}(u_{i,n}).$$

Bref, on l'aura compris, l'étude d'une série d'éléments de  $\mathbf{E}$  se ramène à l'étude de séries réelles. D'où l'importance de telles séries, et du reste de cette partie.

### 3.2 Plan d'étude d'une série réelle

Soit  $\sum u_n$  une série réelle.

1. ETUDE DU TERME GÉNÉRAL : A-t-on  $u_n \xrightarrow{n \rightarrow +\infty} 0$  ?

- (a) Si NON, alors La série diverge.
- (b) Si OUI, alors on passe au point 2.

2. ETUDE DE LA CONVERGENCE ABSOLUE : La série  $\sum |u_n|$  converge-t-elle ?

Pour répondre à cette question on dispose notamment de théorèmes de comparaisons de séries à termes positifs et de la règle de D'Alembert (si  $|u_n|$  ne s'annule pas).

- (a) Si OUI, alors La série converge.

(b) Si NON, alors on passe au point 3.

### 3. ETUDE DIRECTE DE $\sum u_n$ .

Il n'y a pas de route bien tracée. On recourt à divers stratagèmes, on peut parfois, comme on le verra dans le paragraphe 3., utiliser un développements limités de  $u_n$ . La convergence dans le cas de non convergence absolue n'est toutefois pas un objectif du programme, il ne devrait pas en principe faire l'objet d'exercices d'oraux, exception faite des séries alternées qui sont toujours au programme et dont nous allons, sur le champs, parler.

L'étude d'une série par développement limité de son terme général, fait apparaître des séries d'un type particulier appelées séries alternées, que nous allons étudier dans le paragraphe suivant.

## 3.3 Séries alternées

**Définition 3.3.1.** — On dit qu'une série réelle  $\sum u_n$  est alternée si par définition pour tout entier naturel  $n$ ,

$$u_{n+1}u_n \leq 0.$$

Autrement dit, une série alternée est une série dont le terme général change de signe à chaque rang.

**Remarque 3.3.3.** — Soit  $\sum u_n$  une série alternée. posons, pour tout entier naturel  $n$ ,  $a_n := |u_n|$ . Alors :

$$\begin{array}{c} \text{pour tout } n \in \mathbf{N}, u_n = (-1)^n a_n, \\ \text{ou} \\ \text{pour tout } n \in \mathbf{N}, u_n = (-1)^{n+1} a_n. \end{array}$$

Quitte à multiplier la série par  $-1$ , on peut toujours supposer que l'on est dans le premier cas.

**Proposition 3.3.3.** — THÉORÈME SPÉCIAL SUR LES SÉRIES ALTERNÉES —

Soit  $\sum u_n$  une série alternée. On suppose que :

1.  $u_n \xrightarrow{n \rightarrow +\infty} 0$ .
2. La suite  $(|u_n|)_{n \in \mathbf{N}}$  décroît.

Alors :

- La série  $\sum u_n$  converge.
- Pour tout entier naturel  $n$ ,  $|R_n| \leq |u_{n+1}|$ , où  $R_n$  désigne le reste d'ordre  $n$  de la série  $\sum u_n$ . On dit parfois que la valeur absolue du reste est majorée par celle du premier terme négligé.

*Preuve de la proposition 3.3.3.* — Pour fixer les idées on supposera que la série est de la première forme donnée par

la remarque 3.3.2. On note pour tout entier naturel  $n$ ,  $S_n$  la somme partielle d'ordre  $n$  de la série  $\sum u_n$ .

**Remarque 3.3.4.** —

- Le résultat de convergence de la proposition précédente demeure si l'on suppose que la série est alternée à partir d'un certain rang et que la suite  $(|u_n|)_{n \in \mathbf{N}}$  décroît à partir d'un certain rang.
- On retiendra de la démonstration de 3.3.3. que pour tout  $n \in \mathbf{N}$ ,

$$S_1 \leq S_3 \leq S_5 \leq \dots S_{2p+1} \leq \sum_{n=0}^{+\infty} u_n \leq S_{2n} \leq \dots \leq S_6 \leq S_4 \leq S_2.$$

Ces inégalités sont à inverser si l'on suppose que la série est de la seconde forme donnée par la remarque 3.3.2.

**Exemple** — les séries  $\sum_{n \geq 1} \frac{(-1)^n}{n}$ ,  $\sum_{n \geq 1} \frac{(-1)^n}{\sqrt{n}}$  sont convergentes, mais pas absolument convergentes.

**Exercice** — Soit  $n$  un élément de  $\mathbf{N}^*$ . Montrer que pour tout  $k \in \{0, \dots, n\}$ ,

$$1 - \frac{k}{n} + \frac{k^2}{2n^2} - \frac{k^3}{6n^3} \leq \exp\left(-\frac{k}{n}\right) \leq 1 - \frac{k}{n} + \frac{k^2}{2n^2}.$$

*Solution* —

#### UN EXEMPLE ÉDIFIANT

On se propose d'étudier la série  $\sum_{n \geq 1} u_n$ , où pour tout entier  $n \geq 1$

$$u_n = (-1)^n \left( \frac{1}{\sqrt{n}} + \frac{(-1)^n}{n} \right).$$

La suite

La suite  $(|u_n|)_{n \in \mathbf{N}}$  ne décroît donc pas.

Ce résultat appelle les remarques suivantes qui sont autant de démentis cinglants à des préjugés, hélas fort répandus.

- Le théorème spécial sur les séries alternées est faux si l'on ne suppose pas la décroissance de  $(|u_n|)_{n \in \mathbf{N}}$ .
- Deux suites de termes généraux équivalents n'ont pas forcément la même monotonie. On n'étudie jamais la monotonie d'une suite en prenant un équivalent de son terme.
- Deux séries à termes équivalents peuvent être de natures différentes : la positivité est essentielle dans le théorème de comparaison.

### 3.4 Exemples d'utilisation de développements limités

Donnons quelques exemples d'études par développement limité. Ce genre d'étude n'est pas un objectif du programme et ne devrait pas en principe faire l'objet d'exercices d'oraux ou de questions d'écrit, cependant vu l'ancienneté du sujet et l'inertie de l'esprit humain il est à redouter que l'on continue à en voir au concours.

EXEMPLES NE CONTENANT PAS DE PARAMÈTRE

**Règle d'or :** On effectuera toujours des développements limités au sens fort qui sont plus économiques que les développements classiques. On s'arrêtera lorsque apparaît un terme de la forme  $O(\alpha_n)$ , où  $\alpha_n$  est le terme général d'une série positive convergente.

— *Premier exemple*

Etudier la série  $\sum_{n \geq 2} u_n$ , où pour tout entier naturel  $n \geq 2$ ,  $u_n = \frac{(-1)^n}{\sqrt{n + (-1)^n}}$ .

— *Second exemple*

Etudier la série  $\sum_{n \geq 1} u_n$ , où pour tout entier naturel  $n \geq 1$ ,  $u_n = \exp\left((-1)^n \frac{\ln n}{n}\right) - 1$ .

EXEMPLES À PARAMÈTRE

On utilise plutôt des développements limités au sens faibles.

— *Premier exemple*



Etudier la série  $\sum_{n \geq 2} u_n$ , où pour tout entier naturel  $n \geq 2$ ,  $v_n = \ln \left( 1 + \frac{(-1)^n}{n^\alpha} \right)$

— *Second exemple*

Etudier la série  $\sum_{n \geq 2} u_n$ , où pour tout entier naturel  $n \geq 2$ ,  $w_n = \sin \left( \frac{(-1)^n}{n^\alpha} + \frac{k}{n^{5\alpha}} \right)$

## SÉRIES CLASSIQUES À VALEURS DANS $\mathcal{M}_d(\mathbf{K})$

Par  $\mathbf{E}$  on persiste à désigner un  $\mathbf{K}$ -espace vectoriel de dimension finie  $d$  (non nulle).

Comme à l'accoutumée on note  $M$  un élément  $(m_{i,j})_{\substack{i=1,\dots,d \\ j=1,\dots,d}}$  de  $\mathcal{M}_d(\mathbf{K})$ .

### 4.1 Séries géométrique

Le programme ne contient aucun résultat spécifique sur les séries géométrique de matrices, mais suggère que la généralisation des séries réelles aux séries à valeurs dans des espaces vectoriels est principalement destinée aux séries de matrices, aussi allons nous, sous forme d'exercices, étudier les séries géométriques de matrices.

Revenons à l'exemple 4 du début de ce chapitre, pour le généraliser à une matrice diagonalisable. Soit  $A$  un élément de  $\mathcal{M}_d(\mathbf{K})$  diagonalisable dans  $\mathcal{M}_d(\mathbf{C})$ . On note  $\lambda_1, \lambda_2, \dots, \lambda_d$  ses  $d$  valeurs propres distinctes ou non.

**Exercice 4.1.1.** — Montrer que la série  $\sum A^n$  converge si et seulement si  $\max_{i=1,\dots,d} |\lambda_i| < 1$ .

*Solution de l'exercice 4.1.1.* —

Donnons un exemple de série de matrice non diagonalisable. Soit la matrice :

$$B = \begin{pmatrix} \frac{1}{3} & 1 & 1 \\ 0 & \frac{1}{3} & 1 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

La matrice  $B$  ne saurait être diagonalisable car .....

Etudions à présent la série géométrique  $\sum B^n$ .

Passons au cas général.

**Exercice 4.1.2.** — On muni  $\mathcal{M}_d(\mathbf{R})$  de la norme déjà rencontrée (norme de Frobenius)

$$N : \mathcal{M}_d(\mathbf{R}) \rightarrow \mathbf{R}_+; M \mapsto \sqrt{\text{Tr}({}^tMM)},$$

cette norme est en faite la norme euclidienne canonique si l'on identifie  $\mathcal{M}_d(\mathbf{R})$  à  $\mathbf{R}^{d^2}$ ,

1. Montrer que pour tout  $A$  et tout  $B$  éléments de  $\mathcal{M}_d(\mathbf{K})$ ,  $N(AB) \leq N(A)N(B)$ .
2. On suppose que  $A$  est un élément de  $\mathcal{M}_d(\mathbf{K})$  tel que  $N(A) < 1$ . Montrer la convergence de la série  $\sum A^n$ .

*Solution de l'exercice 4.1.2. —*

**Remarque** — Nous verrons en cours d'année comment généraliser la norme de Frobenius à des matrices complexes. L'important du reste n'est pas le choix de la norme, mais qu'il en existe une telle que le produit des normes de deux matrices majore la norme de leur produit.

Revenons un instant au cas complexe. Soit  $|z|$  un nombre complexe. La série  $\sum z^n$  converge si et seulement si  $|z| < 1$ . Si c'est le cas, puisque pour tout entier naturel  $n$ ,

$$1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z},$$

il vient  $\sum_{n=0}^{+\infty} z^n = \frac{1}{1 - z}$ .

Nous allons généraliser à  $\mathcal{M}_d(\mathbf{K})$  ou  $\mathcal{L}(\mathbf{E})$  ce résultat.

**Exercice 4.1.3.** — Soit  $A$  un élément de  $\mathcal{M}_d(\mathbf{R})$  tel que  $N(A) < 1$ .

1. Montrer que  $I_d + A$  est inversible et exprimer son inverse comme la somme d'une série.

2. En déduire que  $\mathrm{GL}_d(\mathbf{K})$  est un ouvert de  $\mathcal{M}_d(\mathbf{K})$ .

Donnons en exercice un cas particulièrement simple et fort semblable.

**Exercice** —

Soit  $M$  l'élément de  $\mathcal{M}_3(\mathbf{R})$ ,

$$M = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Calculer  $M^{-1}$ .

*Solution* —

## 4.2 Exponentielle

La fonction exponentielle définie sur  $\mathbf{R}$  peut être introduite de pas mal de façons. On peut la définir comme la fonction réciproque du logarithme lui-même défini comme la primitive de  $x \mapsto \frac{1}{x}$  nulle en 1. On pourrait voir aussi l'exponentielle comme LA solution sur  $\mathbf{R}$  valant 1 en 0 de l'équation différentielle  $y' = y$ , ce qui exige hélas des connaissances techniques sur les équation différentielles, on peut enfin construire l'exponentielle par équation fonctionnelle : c'est une des applications  $f$  continue de  $\mathbf{R}$  dans  $\mathbf{R}$  telle que pour tout  $(x, y) \in \mathbf{R}^2$ ,  $f(x + y) = f(x)f(y)$ , cette méthode exige un raisonnement par densité de  $\mathbf{Q}$  dans  $\mathbf{R}$ .

Ensuite il est loisible d'introduire l'exponentielle complexe à partir de l'exponentielle réelle et des fonctions cosinus et sinus elles-mêmes arrivées de je ne sais où...

Nous allons pour notre part introduire l'exponentielle réelle et complexe par les séries. Puis nous définirons l'exponentielle d'une matrice et d'un endomorphisme.

**Proposition, définition 4.2.1.** — *Soit  $a$  un élément de  $\mathbf{C}$  Alors : La série  $\sum \frac{a^n}{n!}$  converge absolument, (donc converge). La somme de cette série est appelée exponentielle de  $a$  et notée  $\exp(a)$  ou  $e^a$*

$$\exp(a) = \sum_{n=0}^{+\infty} \frac{a^n}{n!}.$$

Nous verrons que cette exponentielle définie bien celle, pas ou mal définie, que l'on pratique depuis des lustres.

*Preuve de la proposition 4.2.1.* —

Généralisons !

On pourrait munir  $\mathcal{M}_d(\mathbf{R})$  de la norme de Frobenius vue en exercice, mais pour lors sa généralisation à  $\mathcal{M}_d(\mathbf{C})$  n'est pas immédiate, alors choisissons par exemple sur  $\mathcal{M}_d(\mathbf{K})$  ( $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{C}$ ) la norme  $\|\cdot\|_\infty$  que nous définîmes dans un précédent chapitre par :

$$\|M\|_\infty = \sup_{\substack{i=1,\dots,d \\ j=1,\dots,d}} |m_{i,j}|,$$

pour tout  $M \in \mathcal{M}_d(\mathbf{R})$ .

Observons alors, et nous n'aurons pas besoin de plus dans la suite, que

$$\|AB\| \leq,$$

pour tout  $A$  et tout  $B$ , éléments de  $\mathcal{M}_d(\mathbf{K})$ , en effet....

Passons au résultat :

**Proposition, définition 4.2.2.** — Soit  $A$  un élément de  $\mathcal{M}_d(\mathbf{K})$  Alors : La série  $\sum \frac{A^n}{n!}$  converge absolument, (donc converge). La somme de cette série est appelée exponentielle de  $A$  et notée  $\exp(A)$  ou  $e^A$

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$

*Preuve de la proposition 4.2.2.* —

**Remarque** — On peut définir de même l'exponentielle d'un endomorphisme  $u$  de  $\mathbf{E}$ , par

$$\exp(u) = \sum_{n=0}^{+\infty} \frac{u^n}{n!},$$

la convergence de cette série se démontrant par exemple en utilisant l'isomorphisme entre  $\mathcal{L}(\mathbf{E})$  et  $\mathcal{M}_d(\mathbf{K})$ ,  $\text{Mat}_{\mathcal{B}}$ , où  $\mathcal{B}$  est une base de  $\mathbf{E}$ , isomorphisme qui pour un bon choix de normes est une isométrie....

**Exemple** — Soit  $M$  l'élément de  $\mathcal{M}_3(\mathbf{R})$ ,

$$M = \begin{pmatrix} 10 & 4 & 0 \\ 7 & 7 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Calculer  $\exp(M)$ .

*Solution* —

Notons à propos de cet exercice que si  $\lambda_1, \dots, \lambda_d$  sont des éléments de  $\mathbf{K}$  alors :

$$\exp \left( \text{diag}(\lambda_1, \dots, \lambda_d) \right) = \dots\dots\dots$$

**Exercice** —

1. Montrer que l'exponentielle d'une matrice  $A \in \mathcal{M}_d(\mathbf{R})$  symétrique est symétrique.
2. Soit  $B \in \mathcal{M}_d(\mathbf{K})$  une matrice antisymétrique, que vaut  ${}^t \exp(B)$  !



*Solution* —

## FAMILLES SOMMABLES

Nous allons nous intéresser à un objet mathématique plus général que les séries. Les séries pour commencer sont associées à des suites indicées par  $\mathbf{N}$ , nous nous affranchirons, du moins en apparence de cette contrainte en considérant des familles indicées par des ensembles dénombrables quelconques. Ensuite une série est, on la voit, une suite et un mode de sommation : on ajoute dans l'ordre, successivement les termes de la suite. Pour les familles sommables nous allons nous libérer de cette exigence pour ne plus être obligé de tenir compte d'un ordre dans la sommation, ordre qui n'a du reste plus tellement de sens lorsque la famille n'est pas indicée par  $\mathbf{N}$  ou un ensemble naturellement ordonné.

La première sous-partie introduit de façon rigoureuse le concept de dénombrabilité. La deuxième veut introduire les familles sommables de réels positifs, la suivante l'étend aux familles de nombres complexes. La quatrième et dernière partie donnera des applications des familles sommables, notamment aux séries doubles et aux produits de Cauchy de deux séries. L'objectif de toute cette parties est de fournir un outil indispensable au cours de probabilités.

### 5.1 Dénombrabilité

Un ensemble peut être infini ou fini, mais de même que les ensembles finis peuvent différer par leur cardinaux, les ensembles infinis peuvent être de natures différentes. Pour certains il est possible de numéroté (indéfiniment) leurs éléments ce sont les ensembles dénombrables, tels  $\mathbf{N}$ ,  $\mathbf{Z}$ , et de façon plus surprenante  $\mathbf{Q}$ , pour d'autres tels  $\mathbf{R}$ , la chose est impossible.

**Définition 5.1.1.** — *Un ensemble  $A$  est dit dénombrable si, par définition, il peut être mis en bijection avec  $\mathbf{N}$ .*

Un ensemble  $A$  dénombrable peut donc s'écrire  $A = \{x_n\}_{n \in \mathbf{N}}$ , où  $(x_n)_{n \in \mathbf{N}}$  est une suite injective.

**Exemples niais** —

- $\mathbf{N}$  est dénombrable
- L'ensemble  $\mathbf{P}$  des entiers naturels pairs est dénombrable, en effet

$$\phi : \mathbf{P} \rightarrow \mathbf{N}; p \mapsto \dots\dots\dots$$

est une bijection de  $\mathbf{P}$  sur  $\mathbf{N}$

Notons que  $\mathbf{P}$  est une partie stricte de  $\mathbf{N}$ .

- *A contrario* l'ensemble  $\mathbf{Z}$  est dénombrable en effet....

Notons que  $\mathbf{Z}$  contient strictement  $\mathbf{N}$ .

Le cas des entiers naturels pairs est un cas particulier du résultat suivant

**Proposition 5.1.2** — Toute partie  $A$  de  $\mathbf{N}$  infinie est dénombrable.

*Preuve de la proposition 5.1.2.* — Soit  $A$  une partie infinie de  $\mathbf{N}$ . Nous allons « numéroté » les éléments de  $A$  par ordre croissant et construire une bijection de  $\mathbf{N}$  sur  $A$  en associant à un entier naturel  $n$  le  $n + 1^{\text{e}}$  élément de  $A$ . De façon précise et maniaque on pose :

$$\phi(0) = \min(A);$$

et pour tout entier  $n \geq 0$ ,

$$\phi(n+1) = \min(A \setminus \{\phi(0), \phi(1), \dots, \phi(n)\}).$$

Notons que la non vacuité de  $A$  assure que la définition de  $\phi(0)$  est valide, et que le caractère infini de  $A$  assure la non vacuité de  $A \setminus \{\phi(0), \phi(1), \dots, \phi(n)\}$ , qui rend licite la définition de  $\phi(n+1)$ .

On a ainsi construit une application de  $\mathbf{N}$  dans  $A$ . C'est une bijection, en effet :

- Par construction  $\phi$  est strictement croissante donc injective.
- Soit  $a$  un élément de  $A$ , par stricte croissance de  $\phi$ ,  $a < \phi(a+1) = \min(A \setminus \{\phi(0), \phi(1), \dots, \phi(n)\})$  donc  $a \in \{\phi(0), \phi(1), \dots, \phi(n)\}$ , et donc voilà  $\phi$  surjective.

□

**Proposition 5.1.3.** — *Un ensemble  $A$  est fini ou dénombrable (on dit au plus dénombrable), si et seulement si il est en bijection avec une partie de  $\mathbf{N}$ .*

*Preuve de la proposition 5.1.3.* — Soit  $A$  un ensemble.

- HYPOTHÈSE : *On suppose qu'il existe une bijection  $\phi$  de  $A$  sur une partie  $B$  de  $\mathbf{N}$ .*  
Si  $B$  est finie, alors  $A$  est finie, sinon, d'après 5.1.2, il existe alors une bijection  $\psi$  de la partie infinie  $B$  de  $\mathbf{N}$  sur  $\mathbf{N}$ , donc  $\psi \circ \phi$  est une bijection de  $A$  sur  $\mathbf{N}$  et donc  $A$  est *dénombrable*.
- HYPOTHÈSE : *On suppose que  $A$  est au plus dénombrable.*  
Alors dans le cas où  $A$  est dénombrable il est en bijection avec  $\mathbf{N}$  en personne, dans le cas contraire il est fini et se met en bijection avec  $\{1, 2, \dots, |A|\}$  de façon naturelle (et de  $|A|!$  manières!).

D'où le résultat □

On peut exprimer le résultat 5.1.2. en disant qu'un ensemble  $A$  est au plus dénombrable si et seulement si il existe une injection de  $A$  dans  $\mathbf{N}$ .

**Corollaire 5.1.4.** — Soit  $A$  un ensemble dénombrable. Toute partie  $B$  de  $A$  est au plus dénombrable.

**Preuve du corollaire 5.1.4.** — Soit  $B$  une partie de  $A$ . Par hypothèse on dispose d'une bijection  $\phi$  de  $A$  sur  $\mathbf{N}$ . Celle-ci induit une bijection de  $B$  sur  $\phi(B)$  et la proposition 5.1.3. clame que  $B$  est au plus dénombrable □

Voici des exemples plus surprenants.

**Exemple 5.1.5.** —

1.  $\mathbf{N}^2$  est dénombrable. Donnons l'idée de la preuve de ce résultat.

2.  $\mathbf{Q}$  est dénombrable.

L'exemple 5.1.5-2 se généralise dans la proposition suivante que nous admettrons.

**Proposition 5.1.6.** — *Un produit fini d'ensembles dénombrables est dénombrable*

Attention, un produit dénombrable d'ensembles dénombrables ou même finis n'est pas dénombrable. Montrons par

exemple que l'ensemble des suites à valeurs entières,  $\mathbf{N}^{\mathbf{N}}$  n'est pas dénombrable.

On montre de la même façon que  $\mathbf{R}$  n'est pas dénombrable. En effet on a vu en exercices. que l'application qui à un élément  $x$  de  $[0, 1[$  associe la suite standard de ses décimales est une bijection de  $[0, 1[$  sur  $\mathcal{S}$ , ensemble des suites d'éléments de  $\{0, \dots, 9\}$ , (indiquée par  $\mathbf{N}^*$ ) non ultimement constantes à 9.

Nous admettrons également

**Proposition 5.1.7.** — *Soit  $\Lambda$  un ensemble fini ou dénombrable et  $(I_\lambda)_{\lambda \in \Lambda}$  une famille d'ensembles, telle que pour tout  $\lambda \in \Lambda$ ,  $I_\lambda$  soit au plus dénombrable. Alors  $\bigcup_{\lambda \in \Lambda} I_\lambda$  est au plus dénombrable.*

*Dit brutalement : une réunion finie ou dénombrable d'ensembles finis ou dénombrables est fini ou dénombrable*

## 5.2 Famille sommable de réels positifs

Dans toute cette sous-partie  $I$  désigne une partie dénombrable<sup>1</sup>, et  $(u_i)_{i \in I}$  et  $(v_i)_{i \in I}$  une familles de réels positifs ou nuls. .

Dans tout le 5.2., pour tout  $i \in I$ ,  $\boxed{u_i \geq 0, v_i \geq 0}$

On notera  $\mathcal{F}$  l'ensemble des parties finies de  $I$ . Alors  $\left\{ \sum_{i \in F} u_i, F \in \mathcal{F} \right\}$  est un ensemble non vide de réels. Si il est majoré, alors il admet une borne supérieure. On a alors la définition suivante :

**Définition 5.2.1.** —

*On dit que la famille  $(u_i)_{i \in I}$  est sommable si, par définition,  $\left\{ \sum_{i \in F} u_i, F \in \mathcal{F} \right\}$  est majoré.*

1. Cette restriction en fait est à la fois inutile et n'en n'est pas vraiment une : on peut définir pour des familles de réels positifs indicées par un ensemble  $I$  quelconque une notion de sommabilité, mais si une famille est sommable alors l'ensemble des indices pour lesquels son terme est non nul est nécessairement au plus dénombrable.

Si la famille  $(u_i)_{i \in I}$  est sommable, par définition sa somme est le réel

$$\sup \left( \left\{ \sum_{i \in F} u_i, F \in \mathcal{F} \right\} \right).$$

Sinon la somme de la famille  $(u_i)_{i \in I}$  est par définition  $+\infty$ .

Dans les deux cas la somme de la famille  $(u_i)_{i \in I}$  est notée  $\sum_{i \in I} u_i$ .

$(u_i)_{i \in I}$	$\sum_{i \in I} u_i$
Sommable	$\sup \left( \left\{ \sum_{i \in F} u_i, F \in \mathcal{F} \right\} \right)$
Non sommable	$+\infty$

**Remarque** — On peut étendre la définition précédente au cas d'une famille indexée par un ensemble  $J$  fini. La famille est alors toujours sommable et sa somme au sens des familles sommables se confond avec la somme au sens ordinaire des éléments de son image. La chose n'a donc que peu d'intérêt mathématique mais elle nous simplifiera l'énoncé de certains résultats.

### Exemples élémentaires —

1. La famille constante égale à 1 ( $u_i = 1$  pour tout  $i \in I$ ) est ....

2. On suppose que  $\{u_i, i \in I | u_i \neq 0\}$  est fini.

3. On prend pour  $I = \mathbf{Z}^*$  et pour tout  $i \in \mathbf{Z}^*$ ,  $u_i = \frac{1}{i^2}$ .

Dans le cas particulier où  $I = \mathbf{N}$  la notion de sommabilité d'une famille de réels positifs se confond avec celle de série convergente.

**Proposition 5.2.2.** — Soit  $(u_n)_{n \in \mathbf{N}}$  une suite (famille) de réels positifs ou nuls. On a l'équivalence des deux propriétés suivantes :

1. La série à termes positifs ou nuls  $\sum u_n$  converge.
2. La famille de réels positifs ou nuls  $(u_n)_{n \in \mathbf{N}}$  est sommable.

Si c'est le cas alors la somme de la famille sommable  $(u_n)_{n \in \mathbf{N}}$  est égale à celle de la série  $\sum u_n$  :

$$\sum_{n \in \mathbf{N}} u_n = \sum_{n=0}^{+\infty} u_n.$$

*Preuve de la proposition 5.2.2. —*

- HYPOTHÈSE  $\sum u_n$  converge.

Soit  $F$  une partie finie de  $\mathbf{N}$ , par positivité des termes de la suite  $(u_n)_{n \in \mathbf{N}}$ ,

$$\sum_{n \in F} u_n \leq \sum_{n=0}^{\max(F)} u_n \leq \sum_{n=0}^{+\infty} u_n$$

Donc  $F$  étant quelconque,  $(u_n)_{n \in \mathbf{N}}$  est sommable.

De plus, la borne supérieure étant le *plus petit* des majorants :

$$\sum_{n \in \mathbf{N}} u_n \leq \sum_{n=0}^{+\infty} u_n. \quad (\text{IV.2})$$

- HYPOTHÈSE  $(u_n)_{n \in \mathbf{N}}$  est sommable.

Alors pour tout entier  $N \geq 0$ , la borne supérieure d'un ensemble étant en particulier un de ses majorants,

$$\sum_{n=0}^N u_n = \sum_{n \in \{0, \dots, N\}} u_n \leq \sum_{n \in \mathbf{N}} u_n.$$

La suite des sommes partielles de la série  $\sum u_n$  est donc majorée, cette série étant à termes positifs, elle converge donc. De plus laissant tendre  $N$  vers  $+\infty$  dans l'inégalité précédente,

$$\sum_{n=0}^{+\infty} u_n \leq \sum_{n \in \mathbf{N}} u_n. \quad (\text{IV.3})$$

D'où l'équivalence de 1. et de 2.

Les inégalités (IV.2-IV.3) assurent que lorsque 1. et/ou 2. sont vraies, alors

$$\sum_{n=0}^{+\infty} u_n = \sum_{n \in \mathbf{N}} u_n.$$

□

Dans le cas où  $I$  est quelconque il est encore possible de ramener l'étude de la sommabilité d'une famille à celle d'une série, c'est l'objet du lemme qui va suivre. Ce résultat n'est pas au programme mais nous servira par la suite à prouver un résultat du programme. Il repose sur le fait que la dénombrabilité de  $I$ , nous permet de choisir une bijection  $j$  de  $\mathbf{N}$  sur  $I$ . On dispose alors de la famille sommable  $(u_i)_{i \in I}$  et de la série  $\sum_{n \geq 0} u_{j(n)}$ . Alors :

**Lemme 5.2.3. —** *On a l'équivalence des deux propriétés suivantes :*

1. La série à termes positifs ou nuls  $\sum_{n \geq 0} u_{j(n)}$  converge.

2. La famille de réels positifs ou nuls  $(u_i)_{i \in \mathbf{N}}$  est sommable.

Si c'est le cas alors la somme de la famille sommable  $(u_i)_{i \in I}$  est égale à celle de la série  $\sum_{n \geq 0} u_{j(n)}$  :

$$\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} u_{j(n)}.$$

*Preuve du lemme 5.2.3. —*

- HYPOTHÈSE  $\sum u_{j(n)}$  converge.

Soit  $F$  une partie finie de  $I$ .  $F$  est de la forme  $\{j(n_1), j(n_2), \dots, j(n_q)\}$ , où  $n_1, \dots, n_q$  sont les entiers naturels images par  $j^{-1}$  des  $q$  éléments de  $F$ . Posons  $N = \max\{n_1, \dots, n_q\}$ . Alors  $F \subset j(\{0, \dots, N\})$  et donc, par positivité des termes de la famille  $(u_i)_{i \in I}$ ,

$$\sum_{i \in F} u_i \leq \sum_{n=0}^N u_{j(n)} \leq \sum_{n=0}^{+\infty} u_{j(n)}.$$

La partie  $F$  étant quelconque, voici prouvée la sommabilité de  $(u_i)_{i \in I}$ . La borne supérieure étant le *plus petit* des majorants,

$$\sum_{i \in I} u_i \leq \sum_{n=0}^{+\infty} u_{j(n)}. \quad (\text{IV.4})$$

- HYPOTHÈSE  $(u_i)_{i \in I}$  est sommable.

Pour tout  $N \in \mathbf{N}$ ,  $\{j(0), j(1) \dots j(N)\}$  est une partie finie de  $I$  et comme la borne supérieure est un majorant,

$$\sum_{n=0}^N u_{j(n)} \leq \sum_{i \in I} u_i. \quad (\text{IV.5})$$

La suite des sommes partielles de  $\sum_{n \geq 0} u_n$ , est donc majorée, cette série étant à termes positifs elle converge et de plus par passage à la limite dans (IV.5),

$$\sum_{n=0}^{+\infty} u_{j(n)} \leq \sum_{i \in I} u_i. \quad (\text{IV.6})$$

D'où l'équivalence de 1. et de 2. Les inégalités (VI.4) et (VI.5) assurent alors  $\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} u_{j(n)}$ . □

Voici quelques propriétés des sommes de familles sommables de réels positifs ou nul.

On suppose que  $(u_i)_{i \in I}$  est sommable. Soit  $A$  une partie de  $I$ . D'après 5.1.4.  $A$  est au plus dénombrable. Par définition de la somme d'une famille sommable, si  $A$  est finie, alors :

$$\sum_{i \in A} u_i \leq \sum_{i \in I} u_i.$$

Ce résultat va maintenant s'étendre au cas où  $A$  est dénombrable.

**Proposition 5.2.4.** — Soit  $A$  une partie de  $I$  dénombrable. Si  $(u_i)_{i \in I}$  est sommable, alors  $(u_i)_{i \in A}$  est sommable et

$$\sum_{i \in A} u_i \leq \sum_{i \in I} u_i.$$

*Preuve de la proposition 5.2.4.* —

Soit  $F_A$  une partie finie de  $A$ , c'est *a fortiori* une partie finie de  $I$ , et donc la borne supérieure étant un majorant,

$$\sum_{i \in F_A} u_i \leq \sum_{i \in I} u_i.$$

Donc,  $F_A$  étant quelconque, la sommabilité de  $(u_i)_{i \in A}$ . De plus, la borne supérieure étant le *plus petit* des majorants :

$$\sum_{i \in A} u_i \leq \sum_{i \in I} u_i.$$

□

**Proposition 5.2.5.** —

1. Si pour tout  $i \in I$ ,  $0 \leq u_i \leq v_i$  alors, si  $(v_i)_{i \in I}$  est sommable alors  $(u_i)_{i \in I}$  est sommable et

$$\sum_{i \in I} u_i \leq \sum_{i \in I} v_i. \quad (\text{IV.7})$$

2. Soit un réel  $a \geq 0$ . Si  $(u_i)_{i \in I}$  est sommable alors  $(au_i)_{i \in I}$  est sommable et

$$\sum_{i \in I} au_i = a \sum_{i \in I} u_i. \quad (\text{IV.8})$$

3. Si  $(u_i)_{i \in I}$  et  $(v_i)_{i \in I}$  sont sommables alors  $(u_i + v_i)_{i \in I}$  est sommable et

$$\sum_{i \in I} (u_i + v_i) = \sum_{i \in I} u_i + \sum_{i \in I} v_i. \quad (\text{IV.9})$$

*Preuve de la proposition 5.2.5. —*

Nous utiliserons sans cesse le lemme 5.2.3. Par  $j$  nous désignerons une bijection de  $I$  sur  $\mathbf{N}$ .

1. Supposons la famille  $(v_i)_{i \in I}$  est sommable donc la série  $\sum v_{j(n)}$  converge (5.2.3.). Mais par le théorème de comparaison de séries à termes positifs la série  $\sum u_{j(n)}$  et donc la famille  $(u_i)_{i \in I}$  est sommable (5.2.3. encore).

Par ailleurs le cours sur les séries a su nous convaincre que  $\sum_{n=0}^{+\infty} u_{j(n)} \leq \sum_{n=0}^{+\infty} v_{j(n)}$  ce qui grâce à 5.2.3., s'écrit :

$$\sum_{i \in I} u_i \leq \sum_{i \in I} v_i.$$

2. Supposons la famille  $(u_i)_{i \in I}$  est sommable donc la série  $\sum u_{j(n)}$  converge, et donc  $\sum a u_{j(n)}$  converge donc la famille  $(a u_i)_{i \in I}$  est sommable (par deux applications de plus de 5.2.3.). Par ailleurs :

$$\sum_{i \in I} a u_i = \sum_{n=0}^{+\infty} a u_{j(n)} = a \sum_{n=0}^{+\infty} u_{j(n)} = a \sum_{i \in I} u_i.$$

3. Là encore on se ramène au résultat analogue pour les séries positives par le lemme 5.2.3.

□

**Remarque** — Dans la proposition 5.2.5. les l'inégalités (IV.7) et les égalités (h-IV.9) sont sans aucune hypothèse de sommabilité vraies, à condition d'étendre de façon naturelle l'addition et la multiplication de réels et la relation  $\leq$  à  $\mathbf{R} \cup \{+\infty\}$  par :

$$x + (+\infty) = (+\infty) + x = +\infty, \quad (\text{IV.10})$$

$$(+\infty) + (+\infty) = +\infty, \quad (\text{IV.11})$$

$$a \times (+\infty) = +\infty, \quad (\text{IV.12})$$

$$0 \times (+\infty) = 0, \quad (\text{IV.13})$$

$$x \leq +\infty, \quad (\text{IV.14})$$

$$+\infty \leq +\infty, \quad (\text{IV.15})$$

pour tout réel  $x$  et tout réel  $a > 0$ .

Voici un lemme technique totalement évident dans le cas d'une partie  $A$  de  $I$  finie.

**Proposition 5.2.6.** — *Soit  $A$  une partie de  $I$  (donc au plus dénombrable). Alors, avec les conventions (14) et (15),*

$$\sum_{i \in A} u_i = \sum_{i \in I} \mathbf{1}_A(i) u_i.$$

*Preuve de la proposition 5.2.6. —* Le cas fini ne présente pas d'intérêt. Supposons  $A$  dénombrable.

Soit  $F_A$  une partie finie de  $A$ ,

$$\sum_{i \in F_A} u_i = \sum_{i \in F_A} \mathbf{1}_A(i) u_i.$$

Or  $F_A$  est aussi une partie finie de  $I$ , donc

$$\sum_{i \in F_A} u_i \leq \sum_{i \in I} \mathbf{1}_A(i) u_i.$$

D'où avec les conventions (14) et (15),

$$\sum_{i \in A} u_i \leq \sum_{i \in I} \mathbf{1}_A(i) u_i \leq +\infty.$$

Soit maintenant  $G$  une partie finie de  $I$

$$\sum_{i \in G} \mathbf{1}_A(i) u_i = \sum_{i \in G \cap A} u_i \leq \sum_{i \in A} u_i.$$

Et donc toujours avec les conventions (14) et (15),

$$\sum_{i \in I} \mathbf{1}_A(i) u_i \leq \sum_{i \in A} u_i \leq +\infty.$$



Finalement :

$$\sum_{i \in I} \mathbf{1}_A(i) u_i = \sum_{i \in A} u_i \leq +\infty.$$

□

Voici une propriété essentielle pour le cours de probabilité.

**Proposition 5.2.7.** — Soient  $A$  et  $B$  des parties de  $I$  disjointes. Alors :

$$\sum_{i \in A \cup B} u_i = \sum_{i \in A} u_i + \sum_{i \in B} u_i.$$

*Preuve de la proposition 5.2.7.* — La disjonction de  $A$  et de  $B$  donne  $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B$ . Alors par 5.2.6.,

$$\sum_{i \in A \cup B} u_i = \sum_{i \in I} \mathbf{1}_{A \cup B}(i) u_i = \sum_{i \in I} \mathbf{1}_A(i) u_i + \mathbf{1}_B(i) u_i$$

En faisant appel à 5.2.5 -3,

$$\sum_{i \in A \cup B} u_i = \sum_{i \in I} \mathbf{1}_A(i) u_i + \sum_{i \in I} \mathbf{1}_B(i) u_i,$$

puis à 5.2.6.,

$$\sum_{i \in A \cup B} u_i = \sum_{i \in A} u_i + \sum_{i \in B} u_i (\leq +\infty).$$

□

Passons au théorème central de cette partie que le programme nous demande d'admettre.

**Proposition 5.2.8.** — SOMMATION PAR PAQUETS POUR LES FAMILLES DE RÉELS POSITIFS OU NULS —

Soit  $\{I_n\}_{n \in \mathbf{N}}$  une partition de  $I$ . Alors la famille  $(u_i)_{i \in I}$  est sommable si et seulement si :

— pour tout entier naturel  $n$ , la famille  $(u_i)_{i \in I_n}$  est sommable,

— la série  $\sum_{n \geq 0} \left( \sum_{i \in I_n} u_i \right)$  converge.

Si c'est le cas :

$$\sum_{n=0}^{+\infty} \left( \sum_{i \in I_n} u_i \right) = \sum_{i \in I} u_i.$$

**Exemple** — Reprenons l'exercice du 2.1. : l'étude de  $\sum p_n^{-p_n}$ , où  $p_n$  est le nombre de chiffres dans l'écriture décimale de  $n$ .

Donnons une application aux cas des familles indicées par  $\mathbf{N}^2$

Soit  $(u_{p,q})_{(p,q) \in \mathbf{N}^2}$  une famille (« suite double ») de réels positifs ou nuls.

pour tout  $(p, q) \in \mathbf{N}^2$ ,  $\boxed{u_{p,q} \geq 0}$

Une partition évidente de  $\mathbf{N}^2$  est  $\{I_p\}_{p \in \mathbf{N}}$  avec pour tout entier naturel  $p$ ,

$$I_p = \{p\} \times \mathbf{N} = \{(p, q), q \in \mathbf{N}\}.$$

D'après 5.2.2. et le théorème de sommation par paquets donne :

**Proposition 5.2.9.** — *La famille de réels positifs ou nuls  $(u_{p,q})_{(p,q) \in \mathbf{N}^2}$  est sommable si et seulement si :*

- *pour tout élément  $p$  de  $\mathbf{N}$ , la série  $\sum_{q \geq 0} u_{p,q}$  converge,*

- *la série  $\sum_{p \geq 0} \left( \sum_{q=0}^{+\infty} u_{p,q} \right)$  converge.*

Si tel est le cas alors :

$$\sum_{p=0}^{+\infty} \left( \sum_{q=0}^{+\infty} u_{p,q} \right) = \sum_{i \in I} u_i$$

Remarquons alors que  $p$  et  $q$  jouent des rôles symétrique et que l'on aurait pu prendre comme partition de  $\mathbf{N}^2$ ,

$$\{\mathbf{N} \times \{q\}\}_{q \in \mathbf{N}}.$$

On obtient donc le théorème suivant :

**Proposition 5.2.10.** — THÉORÈME DE FUBINI TONELLI —

Soient  $(u_{p,q})_{(p,q) \in \mathbf{N}^2}$  une famille de réels POSITIFS ou nuls. Si pour tout élément  $p$  de  $\mathbf{N}$ ,  $\sum_{q \geq 0} u_{p,q}$  converge et si

$\sum_{p \geq 0} \left( \sum_{q=0}^{+\infty} u_{p,q} \right)$  converge, alors la famille est sommable et pour tout élément  $q$  de  $\mathbf{N}$ ,  $\sum_{p \geq 0} u_{p,q}$  converge et  $\sum_{q \geq 0} \left( \sum_{p=0}^{+\infty} u_{p,q} \right)$  converge de plus

$$\sum_{p=0}^{+\infty} \left( \sum_{q=0}^{+\infty} u_{p,q} \right) = \sum_{(p,q) \in \mathbf{N}^2} u_{p,q} = \sum_{q=0}^{+\infty} \left( \sum_{p=0}^{+\infty} u_{p,q} \right).$$

**Remarque** — La réciproque est vraie, mais moins employée et peut être vue comme un cas particulier de 5.3.12.

**Exercice 5.2.11.** — Montrer que pour tout entier  $n \geq 2$ ,  $\sum_{k \geq 2} \frac{1}{k^n}$  converge. montrer que :  $\sum_{n \geq 2} \left( \sum_{k=2}^{+\infty} \frac{1}{k^n} \right)$  converge et donner sa somme.

*Solution de l'exercice 5.2.11.* —

### 5.3 Famille sommable de réels ou de complexes

Nous allons étendre la notion de sommabilité des familles de nombres réels ou complexes. Comme dans la sous-partie précédente,  $I$  désigne un ensemble dénombrable. Maintenant  $(u_i)_{i \in I}$  désignera une famille de nombres réels ou complexes.

**Définition 5.3.1.** — *Nous dirons que la famille  $(u_i)_{i \in I}$  est sommable si par définition la famille de réels positifs ou nuls,  $(|u_i|)_{i \in I}$  est sommable.*

Cette notion prolonge bien la notion de sommabilité des familles de réels positifs ou nuls : une famille de réels positifs ou nuls sommable est encore sommable comme famille de réels. Il nous faut maintenant définir la somme d'une telle série.

#### CAS RÉELS

Prenons  $(u_i)_{i \in I}$  à valeurs réelles. On définit, comme nous l'avons fait pour les suites dans le 1.2., les familles de réels positifs  $(u_i^+)_{i \in I}$  et  $(u_i^-)_{i \in I}$ , partie positive et partie négative de  $(u_i)_{i \in I}$ , par :

$$u_i^+ = \begin{cases} u_i & \text{si } u_i \geq 0, \\ 0 & \text{si } u_i < 0, \end{cases}$$

$$u_i^- = \begin{cases} -u_i & \text{si } u_i \leq 0, \\ 0 & \text{si } u_i > 0, \end{cases}$$

pour tout élément  $i$  de  $I$ . On a pour tout  $i \in I$ ,

$$u_i = u_i^+ - u_i^- ; |u_i| = u_i^+ + u_i^-.$$

Remarquons que pour une partie finie  $F$  de  $I$ ,

$$\sum_{i \in F} u_i = \sum_{i \in F} u_i^+ - \sum_{i \in F} u_i^-.$$

Inspirons nous de cette égalité pour définir la somme de la famille  $(u_i)_{i \in I}$  lorsqu'elle est sommable.

**Proposition, définition 5.3.2.** — *On suppose la famille de réels  $(u_i)_{i \in I}$  sommable. Alors les familles de réels positifs  $(u_i^+)_{i \in I}$  et  $(u_i^-)_{i \in I}$  sont sommables.*

*Par définition on appelle somme de la famille  $(u_i)_{i \in I}$  et on note  $\sum_{i \in I} u_i$  le réel :*

$$\sum_{i \in I} u_i^+ - \sum_{i \in I} u_i^-.$$

#### Remarques —

1. Si  $(u_i)_{i \in I}$  est à valeur dans  $\mathbf{R}_+$  sa somme en tant que famille de réels positifs ou nuls et celle en tant que famille de réels coïncident, puisque pour tout  $i \in I$ ,  $u_i = u_i^+$  et  $u_i^- = 0$ .
2. La somme d'une famille de réels non positifs ou nuls n'existe que dans le cas où elle est sommable (dans le cas où ni  $(u_i^+)_{i \in I}$  ni  $(u_i^-)_{i \in I}$  ne sont sommables, la précédente définition conduirait à  $+\infty - (+\infty)$  quantité indéterminé(able)).
3. Si l'on étend comme pour les familles de réels positifs ou nuls la notion de sommabilité dans le cas où  $I$  est fini, la somme que l'on vient de définir coïncide opportunément avec la somme ordinaire de ses termes.

*Preuve de 5.3.2.* —



#### CAS COMPLEXE

Prenons  $(u_i)_{i \in I}$  à valeurs complexe. La définition de la somme d'une famille sommable de complexes utilise celle donnée à l'instant pour les familles sommables de réels :

**Proposition, définition 5.3.3.** — *On suppose la famille de réels  $(u_i)_{i \in I}$  sommable. Alors les familles de réels  $(\operatorname{Re}(u_i))_{i \in I}$  et  $(\operatorname{Im}(u_i))_{i \in I}$  sont sommables.*

*Par définition on appelle somme de la famille  $(u_i)_{i \in I}$  et on note  $\sum_{i \in I} u_i$  le complexe :*

$$\sum_{i \in I} \operatorname{Re}(u_i) + i \sum_{i \in I} \operatorname{Im}(u_i).$$

#### Remarques —

1. Si  $(u_i)_{i \in I}$  est sommable sa somme peut encore s'écrire :

$$\sum_{i \in I} u_i = \sum_{i \in I} (\operatorname{Re}(u_i))^+ - \sum_{i \in I} (\operatorname{Re}(u_i))^- + i \sum_{i \in I} (\operatorname{Im}(u_i))^+ - i \sum_{i \in I} (\operatorname{Im}(u_i))^- \quad (\text{IV.16})$$

2. Si  $(u_i)_{i \in I}$  est à valeurs dans  $\mathbf{R}$  sa somme en tant que famille de réels et celle en tant que famille de complexes coïncident.

3. La somme d'une famille de complexes non positifs ou nuls n'existe que dans le cas où elle est sommable.

4. Si l'on étend comme pour les familles de réels positifs ou nuls la notion de sommabilité dans le cas où  $I$  est fini, la somme que l'on vient de définir coïncide opportunément avec la somme ordinaire de ses termes.

*Preuve de 5.3.3. —*

**Dans la suite  $(u_i)_{i \in I}$  est à valeurs indifféremment dans  $\mathbf{R}$  ou  $\mathbf{C}$ .**

Il résulte immédiatement de la définition et de 5.2.2., la proposition suivante.

**Proposition 5.3.4. —** *Soit  $(u_n)_{n \in \mathbf{N}}$  une famille (suite) de nombres complexes. Alors sont équivalentes les deux propositions suivantes :*

1. la série  $\sum u_n$  est absolument convergente.
2. La famille  $(u_n)_{n \in \mathbf{N}}$  est sommable. Si c'est le cas alors la somme de la famille sommable  $(u_n)_{n \in \mathbf{N}}$  est égale à celle de la série  $\sum u_n$  :

$$\sum_{n \in \mathbf{N}} u_n = \sum_{n=0}^{+\infty} u_n.$$

Donnons à présent un lemme qui ne figure pas au programme mais utile dans des preuves à venir.

Considérons comme en 5.2.3. une bijection  $j$  de  $\mathbf{N}$  sur  $I$ . On a le résultat suivant.

**Lemme 5.3.5. —** *On suppose que la famille  $(u_i)_{i \in I}$  est sommable. Alors la série  $\sum_{n \geq 0} u_{j(n)}$  converge absolument et sa somme est égale à la somme de la famille sommable  $(u_i)_{i \in I}$  :*

$$\sum_{n=0}^{+\infty} u_{j(n)} = \sum_{i \in I} u_i.$$

*Preuve du lemme 5.2.5 —*

Par définition de la sommabilité d'une famille de complexes,  $(|u_i|)_{i \in I}$  est sommable et donc, d'après le lemme 5.2.3., la série  $\sum_{n \geq 0} |u_{j(n)}|$  converge, et donc  $\sum_{n \geq 0} u_{j(n)}$  converge absolument.

La sommabilité de  $(u_i)_{i \in I}$  assure par ailleurs, comme nous venons de le voir celles des 4 familles de réels positifs  $((\operatorname{Re}(u_i))^\pm)_{i \in I}$  ;  $((\operatorname{Im}(u_i))^\pm)_{i \in I}$ , donc, toujours d'après 5.2.3, la convergence des 4 séries

$$\sum_{n \geq 0} \operatorname{Re}(u_{j(n)})^\pm ; \sum_{n \geq 0} \operatorname{Im}(u_{j(n)})^\pm$$

et l'égalité des sommes des 4 familles et des séries correspondantes.

Donc, d'après (IV.16) :

$$\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} (\operatorname{Re}(u_{j(n)}))^+ - \sum_{n=0}^{+\infty} (\operatorname{Re}(u_{j(n)}))^- + i \sum_{n=0}^{+\infty} (\operatorname{Im}(u_{j(n)}))^+ - i \sum_{n=0}^{+\infty} (\operatorname{Im}(u_{j(n)}))^-$$

Donc

$$\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} u_{j(n)}.$$

□

Donnons quelques propriétés des familles sommables de complexes.

**Proposition 5.3.6.** — *Si  $(u_i)_{i \in I}$  est sommable, alors pour tout partie  $A$  de  $I$ , la famille  $(u_i)_{i \in A}$  est sommable.*

*Si  $A$  et  $B$  sont deux parties de  $I$  disjointes alors :*

$$\sum_{i \in A \cup B} u_i = \sum_{i \in A} u_i + \sum_{i \in B} u_i.$$

Dans cette proposition on étend le cas échéant la notion de sommabilité et de somme à des familles finies, comme déjà mentionné.

*Preuve de la proposition 5.3.6.* — Résulte directement de 5.2.7. appliqué à  $((\operatorname{Re}(u_i))^\pm)_{i \in I}$  ;  $((\operatorname{Im}(u_i))^\pm)_{i \in I}$ .

□

**Proposition 5.3.7.** — *L'ensemble  $S$  des familles de réels (resp. complexes) indicées par  $I$  sommables, est un sous-espace vectoriel de  $\mathbf{R}^I$  (resp.  $\mathbf{C}^I$ ). De plus l'application qui à un élément de  $S$  associe sa somme :*

$$(u_i)_{i \in I} \mapsto \sum_{i \in I} u_i$$

*est linéaire.*

*Preuve de la proposition 5.3.7.* —

Le lemme 5.3.5 prouve ce que suggère la définition même de la sommabilité : la permutation des termes d'une famille ne modifie pas sa sommabilité.

**Proposition 5.3.8.** — *Soit  $\sigma$  une permutation de  $I$ . Si  $(u_i)_{i \in I}$  est sommable alors  $(u_{\sigma(i)})_{i \in I}$  est sommable et de même somme.*

*Preuve de la proposition 5.3.8. —*

**Corollaire 5.3.9.** — Soit  $\sigma$  une permutation de  $\mathbf{N}$  et  $\sum a_n$  une série. Si  $\sum a_n$  converge absolument alors  $\sum a_{\sigma(n)}$  converge absolument et  $\sum_{n=0}^{+\infty} a_n = \sum_{n=0}^{+\infty} a_{\sigma(n)}$ . Dit brutalement : on ne modifie pas la somme d'une série absolument convergente en permutant ses termes.

*Preuve du corollaire 5.3.9. —*

Supposons que  $\sum a_n$  converge absolument alors la famille  $(a_n)_{n \in \mathbf{N}}$  est sommable et  $\sum_{n=0}^{+\infty} a_n = \sum_{n \in \mathbf{N}} a_n$ , par 5.3.4. Donc 5.3.8. affirme alors que  $(a_{\sigma(n)})_{n \in \mathbf{N}}$  est sommable et que  $\sum_{n \in \mathbf{N}} a_{\sigma(n)} = \sum_{n \in \mathbf{N}} a_n$ . Enfin 5.3.4. assure que  $\sum a_{\sigma(n)}$  converge absolument et que :

$$\sum_{n=0}^{\mathbf{N}} a_{\sigma(n)} = \sum_{n \in \mathbf{N}} a_n.$$

Finalement  $\sum_{n=0}^{\mathbf{N}} a_{\sigma(n)} = \sum_{n=0}^{\mathbf{N}} a_n$ . □

**Mise en garde !** La somme et la nature d'une série convergente, non absolument convergente, sont affectées par la permutation de ses termes. Pire, si  $\sum u_n$  est une série réelle convergente, non absolument convergente, on peut montrer qu'il existe une permutation  $\sigma$  telle que  $\sum u_{\sigma(n)}$  diverge et pour tout réel  $\ell$ , une permutation  $\sigma'$  telle que  $\sum u_{\sigma'(n)}$  converge de somme  $\ell$ .

Donnons un théorème de sommation par paquets.

**Proposition 5.3.10.** — SOMMATION PAR PAQUETS POUR LES FAMILLES DE COMPLEXES —

Soit  $\{I_n\}_{n \in \mathbf{N}}$  une partition de  $I$ . On suppose que la famille  $(u_i)_{i \in I}$  de complexes est SOMMABLE. Alors

- pour tout entier naturel  $n$ , la famille  $(u_i)_{i \in I_n}$  est sommable,
- la série  $\sum_{n \geq 0} \left( \sum_{i \in I_n} u_i \right)$  converge,
- enfin  $\sum_{n=0}^{+\infty} \left( \sum_{i \in I_n} u_i \right) = \sum_{i \in I} u_i$ .

**Remarque 5.3.11** —

1. Notons bien que par hypothèse  $(u_i)_{i \in I}$  est sommable.
2. La sommabilité pour tout  $n \in \mathbf{N}$ , de la famille  $(u_i)_{i \in I_n}$  et la convergence (même absolue) de la série  $\sum_{n \geq 0} \left( \sum_{i \in I_n} u_i \right)$ , ne suffit pas à prouver la sommabilité de  $(u_i)_{i \in I}$ . Prenons par exemple la famille  $((-1)^k)_{k \in \mathbf{N}}$  et pour tout  $n \in \mathbf{N}$ ,

$$I_n = \{2n, 2n+1\} \dots$$

3. En revanche pour montrer la sommabilité de  $(u_i)_{i \in I}$ , très souvent dans la pratique on utilise 5.2.8., théorème de sommation pour des familles positives, appliqué à  $(|u_i|)_{i \in I}$ . Concrètement on montre que :
- pour tout entier naturel  $n$ , la famille  $(|u_i|)_{i \in I_n}$  est sommable,
  - la série  $\sum_{n \geq 0} \left( \sum_{i \in I_n} |u_i| \right)$  converge.

Traisons le cas particulier des séries doubles. Le théorème de sommation par paquets donne en prenant comme partition de  $\mathbf{N}^2$ ,  $\{\{p\} \times \mathbf{N}\}_{p \in \mathbf{N}}$  et  $\{\mathbf{N} \times \{q\}\}_{q \in \mathbf{N}}$ ,

**Proposition 5.3.12.** — THÉORÈME DE FUBINI LEBESGUE POUR LES SUITES DOUBLES COMPLEXES —

Soient  $(u_{p,q})_{(p,q) \in \mathbf{N}^2}$  une famille de réels ou de complexes. On suppose que la famille  $(|u_{p,q}|)_{(p,q) \in \mathbf{N}^2}$  est SOMMABLE. Alors :

1. Pour tout élément  $p$  de  $\mathbf{N}$ ,  $\sum_{q \geq 0} u_{p,q}$  converge ;
2.  $\sum_{p \geq 0} \left( \sum_{q=0}^{+\infty} u_{p,q} \right)$  converge ;
3. Pour tout élément  $q$  de  $\mathbf{N}$ ,  $\sum_{p \geq 0} u_{p,q}$  converge ;
4.  $\sum_{q \geq 0} \left( \sum_{p=0}^{+\infty} u_{p,q} \right)$  converge ;
5. Enfin :  $\sum_{p=0}^{+\infty} \left( \sum_{q=0}^{+\infty} u_{p,q} \right) = \sum_{(p,q) \in \mathbf{N}^2} u_{p,q} = \sum_{q=0}^{+\infty} \left( \sum_{p=0}^{+\infty} u_{p,q} \right)$ .

On notera que l'on fait l'hypothèse cruciale : la famille  $(|u_{p,q}|)_{(p,q) \in \mathbf{N}^2}$  est sommable, hypothèse mise en défaut dans l'exemple 5.3.13.

Pour vérifier la sommabilité de  $(|u_{p,q}|)_{(p,q) \in \mathbf{N}^2}$ , on doit montrer que, pour tout élément  $p$  de  $\mathbf{N}$ ,  $\sum_{q \geq 0} |u_{p,q}|$  converge et  $\sum_{p \geq 0} \left( \sum_{q=0}^{+\infty} |u_{p,q}| \right)$  converge, ou bien que pour tout élément  $q$  de  $\mathbf{N}$ ,  $\sum_{p \geq 0} |u_{p,q}|$  converge et  $\sum_{q \geq 0} \left( \sum_{p=0}^{+\infty} |u_{p,q}| \right)$  converge. Le choix de l'ordre de sommation, nous l'avons dit dans le paragraphe précédent, n'est alors pas en général indifférent dans la pratique.

Retenons en résumé, que si l'on peut sommer dans un ordre quelconque la famille des modules, alors on peut sans les modules, sommer dans l'ordre qui nous plaît et sans que l'ordre influe sur le résultat.

**Exemple 5.3.13.** —

La famille  $(u_{p,q})_{(p,q) \in \mathbf{N}^2}$  est donnée par le tableau suivant :

Nous sommes dans la situation où  $\sum_{p=0}^{+\infty} \left( \sum_{q=0}^{+\infty} u_{p,q} \right)$  et  $\sum_{q=0}^{+\infty} \left( \sum_{p=0}^{+\infty} u_{p,q} \right)$  ont un sens mais où ces deux quantités sont différentes. l'hypothèse de sommabilité dans la proposition 5.3.12. est donc cruciale.

Traisons un exercice.

**Exercice 5.3.14.** — Montrer que pour tout entier  $n \geq 2$ ,  $\sum_{k \geq 2} \frac{(-1)^n}{k^n}$  converge. Montrer que :  $\sum_{n \geq 2} \left( \sum_{k=2}^{+\infty} \frac{(-1)^n}{k^n} \right)$  converge et donner sa somme.



*Solution de l'exercice 5.3.14. —*

## 5.4 Produit de Cauchy de deux séries absolument convergente

Etant donné des séries  $\sum u_n$  et  $\sum v_n$  d'éléments de  $\mathbf{K}$ , convergentes on recherche une série convergente dont la somme serait le produit des sommes des séries  $\sum u_n$  et  $\sum v_n$ .

Commençons par une démarche heuristique, purement formelle.

DÉMARCHE HEURISTIQUE

On a formellement

$$\sum_{p=0}^{+\infty} u_p \sum_{q=0}^{+\infty} v_q = \sum_{(p,q) \in \mathbf{N}^2} u_p v_q.$$

Regroupons formellement les termes de cette dernière « somme », en s'inspirant de la figure ci-dessous.

$$\sum_{p=0}^{+\infty} u_p \sum_{q=0}^{+\infty} v_q = \sum \left( \sum \right).$$

On intuite que

$$\sum_{p=0}^{+\infty} u_p \sum_{q=0}^{+\infty} v_q = \sum_{n=0}^{+\infty} c_n,$$

avec  $c_n = \sum$  .

Revenons à une étude rigoureuse qui donnera sous certaines hypothèses un sens à ce qui précède. La propriété à prouver est une sommation par paquets d'une suite double  $(u_p v_q)_{(p,q) \in \mathbf{N}^2}$ . Nous avons vu dans la sous-partie précédente qu'une telle opération exige la sommabilité de la famille  $(u_p v_q)_{(p,q) \in \mathbf{N}^2}$ , cette dernière propriété nécessite des hypothèses : nous prendrons l'absolue convergence des séries  $\sum u_n$  et  $\sum v_n$ .

**Définition 5.4.1.**— Soient  $\sum u_n$  et  $\sum v_n$  des séries de nombres réelles ou complexes. On appelle produit de Cauchy de ces séries, ou série produit, la série  $\sum c_n$ , où pour tout entier naturel  $n$ ,

$$c_n = \sum$$

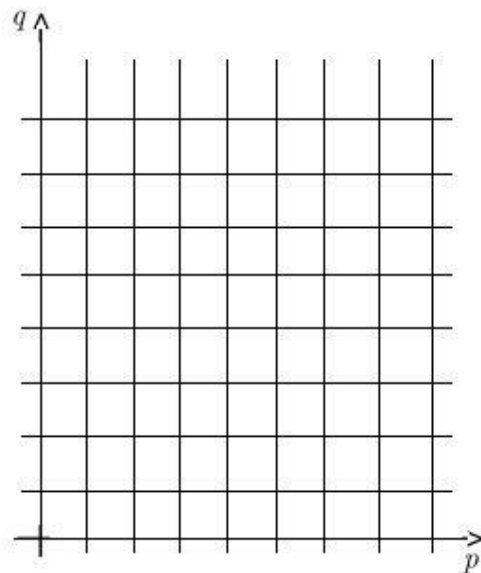


FIGURE IV.1 –

**Remarque** —

- Avec les notations précédentes, notons que, pour tout entier naturel  $n$ ,  $c_n = \sum_{k=0}^n u_k v_{n-k} = \sum_{k=0}^n v_k u_{n-k}$ .
- les formules donnant le produit de Cauchy montre que le produit de Cauchy est commutatif.

On dispose alors du résultat suivant :

**Proposition 5.4.2.**— Soient  $\sum u_n$  et  $\sum v_n$  des séries de nombres réelles ou complexes absolument convergentes. La série produit de ces séries,  $\sum c_n$  est absolument convergente et

$$\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} u_n \sum_{n=0}^{+\infty} v_n,$$

autrement dit, la somme de la série produit est le produit des sommes.

**Remarque** — Si l'on suppose que  $\sum u_n$  et  $\sum v_n$  convergent mais ne convergent pas absolument, la série produit

peut diverger, par exemple le produit de la série  $\sum \frac{(-1)^n}{\sqrt{n+1}}$  par elle-même diverge grossièrement.

*Preuve de la proposition 5.4.2. —*



Montrons comme promis que l'exponentielle définie 4.2. est bien celle que l'on pratique depuis des lustres.

**Proposition 5.4.3.** — *Soit  $z_1$  et  $z_2$  des éléments de  $\mathbf{K}$ .*

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2).$$

**Remarque** — Le résultat est faux dans une algèbre normée de dimension finie quelconque ; il demeure vrai cependant, lorsque les deux éléments  $z_1$  et  $z_2$  commutent, mais ce fait est hors programme.

*Preuve de la proposition 5.4.3. —*



L'exponentiel est donc rigoureusement définie. La proposition précédente assure que pour  $x$  et  $y$  réels,

$$\exp(x + iy) = \exp(x) \exp(iy).$$

On peut alors poser pour tout  $x$  réel,

$$\cos(x) = \operatorname{Re}(\exp(ix)), \sin(x) = \operatorname{Im} \exp(ix)$$

Et l'on retrouve les formules de MPSI...

## 5.5 Application aux probabilités

Nous allons donner quelques éléments de probabilités sur un ensemble dénombrable qui généraliseront les connaissances de sup. sur un ensemble fini. Il ne s'agit pas d'un cours de probabilité, celui-ci cloturera l'année, mais plus une initiation aux probabilités qui permettra tout au long de l'année de se familiariser avec le singulier langage des probabilités en traitant des exercices en relations avec les chapitres traités et rendra plus facile la mise en place finale du cours de probabilité proprement dit.

Par  $\Omega$  on désigne un ensemble dénombrable que nous nomerons *univers* ses parties *événements* et qui modélise l'ensemble des issues possibles d'une expérience aléatoire. Par exemple si l'on jette une pièce un nombre quelconque de fois, on modélise l'expérience par l'univers qui est l'ensemble des suites finies d'éléments de  $\{P, F\}$ ,

$$\omega = \{(); (P); (F); (P, P); (P, F); (F, F); (P, P, P); (P, P, F); \dots\}.$$

On définit alors comme l'an passé sur  $\mathcal{P}(\Omega)$  une probabilité :

**Définition 5.5.1.** — *On appelle probabilité, toute application  $\mathbf{P}$  de  $\mathcal{P}(\Omega)$  dans  $[0, 1]$ , qui jouit des propriétés suivantes :*

1.  $\mathbf{P}(\Omega) = 1$ ;
2. Pour toute suite de  $(A_n)_{n \in \mathbf{N}}$  de parties (événements) de  $\Omega$  deux à deux disjointes,

$$\mathbf{P}\left(\bigcup_{n \in \mathbf{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n), \text{ (}\sigma\text{-additivité).}$$

La propriété 2. sous-entend la convergence de la série à terme positifs  $\sum \mathbf{P}(A_n)$  ou si l'on préfère la sommabilité de la famille  $(\mathbf{P}(A_n))_{n \in \mathbf{N}}$  (cf. 5.2.3.)

Si  $\mathbf{P}$  est une probabilité, alors on a comme dans le cas où  $\Omega$  est fini que

$$\boxed{\mathbf{P}(\emptyset) = 0}$$

Il suffit d'appliquer 2. à la suite de parties constamment égale au vide. On en déduit alors que pour toute famille finie  $(A_0, A_1, \dots, A_n)$  de parties de  $\Omega$ ,

$$\mathbf{P}\left(\bigcup_{k=0}^n A_k\right) = \sum_{n=0}^n \mathbf{P}(A_k).$$

C'est une conséquence de 2. appliqué à la suite  $(A_0, A_1, \dots, A_n, \emptyset, \emptyset, \dots)$

Cette propriété appelée additivité est strictement plus faible que la  $\sigma$ -additivité, bien que plus intuitive elle ne permet pas dans le cas dénombrable de traiter de façon satisfaisante, certaines questions, même simples.

Une conséquence immédiate, comme dans le cas d'un univers fini, est qu'une probabilité est entièrement définie par ses valeurs sur les événements élémentaires (singletons).

**Proposition 5.5.2.** — *La famille  $(\mathbf{P}(\{\omega\}))_{\omega \in \Omega}$  est sommable et pour toute partie  $A$  de  $\Omega$ ,*

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\})$$

*Preuve de la proposition 5.5.2. —*

**Proposition 5.5.3.** — *Soit  $(p_\omega)_{\omega \in \Omega}$  une famille de réels, il existe une probabilité  $\mathbf{P}$  sur  $\Omega$  telle que pour tout  $\omega \in \Omega$ ,  $\mathbf{P}(\{\omega\}) = p_\omega$  si et seulement si :*

1. *Pour tout  $\omega \in \Omega$ ,  $p_\omega \geq 0$ .*
2. *La famille  $(p_\omega)_{\omega \in \Omega}$  est sommable de somme 1.*

*Si  $(p_\omega)_{\omega \in \Omega}$  satisfait ces deux conditions, il existe une et une seule probabilité  $\mathbf{P}$  sur  $\omega$  telle que pour tout  $\omega \in \Omega$ ,  $\mathbf{P}(\{\omega\}) = p_\omega$ , elle est donnée par :*

$$\mathbf{P}(A) = \sum_{\omega \in A} p_\omega,$$

*pour toute partie  $A$  de  $\Omega$ .*

*Preuve de la proposition 5.5.3. —*

Une probabilité étant définie de façon unique par la donnée de ses valeurs sur les événements élémentaires, dans la pratique c'est ainsi qu'elle est définie, la famille sommable de réels positifs de somme 1 la définissant est parfois appelé « loi » de la probabilité.

**Exemple** — LOI DE POISSON —

Soit un réel  $\lambda > 0$ . La famille  $(p_n)_{n \in \mathbf{N}}$ , où pour tout entier  $n \geq 0$ ,  $p_n = e^{-\lambda} \frac{\lambda^n}{n!}$  est la loi d'une probabilité sur  $\mathbf{N}$ , appelé loi de Poisson de paramètre  $\lambda$ . En effet...

**Remarque** — Les propriétés d'une probabilités vue en sup. demeurent notamment celles sur la probabilité de la réunion.

**Dans la suite on note  $p_\omega = \mathbf{P}(\{\omega\})$ , pour tout élément  $\omega$  de  $\Omega$**

Passons aux variables aléatoires définies comme dans le cas d'un univers fini.

**Définition 5.5.4.** — *On appelle variables aléatoires  $X$  sur  $\Omega$  toute application  $X$  de  $\Omega$  dans un ensemble  $\mathbf{E}$ . Si  $\mathbf{E}$  est une partie de  $\mathbf{R}$  on parle de variable aléatoire réelle.*

On se donne jusqu'à la fin une variable aléatoire  $X$  à valeurs dans un ensemble  $\mathbf{E}$ .

L'ensemble  $\mathbf{E}$  peut ne pas être dénombrable, mais l'image de  $\Omega$  par  $X$  est un ensemble, que nous noterons  $\mathbf{E}'$ , qui est fini ou dénombrable.

Pour toute partie  $A$  de  $\mathbf{E}'$  on note, comme dans le cours de sup. et avec les notations abusives qui y sont définies,

$$\mathbf{P}_X(A) = \mathbf{P}(X \in A).$$

On a encore le résultat suivant :

**Proposition 5.5.5.** — L'application

$$\mathbf{P}_X : \mathcal{P}(E') \rightarrow [0, 1]; A \mapsto \mathbf{P}_X(A)$$

est une probabilité sur  $\mathbf{E}'$ .

*Preuve 5.5.5.* —

**Remarque** — La probabilité  $\mathbf{P}_X$  est en vertu de 5.5.2., entièrement définie par la famille

$$(\mathbf{P}_X(\{x\}))_{x \in E'}.$$

Notons que pour tout  $x \in E'$ ,

$$\mathbf{P}_X(\{x\}) = \sum_{\omega \in \{X=x\}} p_\omega.$$

On appelle loi de la variable  $X$  aussi bien la probabilité  $\mathbf{P}_X$  que la famille  $(\mathbf{P}_X(\{x\}))_{x \in E'}$ , qui la définit.

Définissons l'espérance.

**Définition 5.5.6.** — Soit  $X$  une variable aléatoire discrète définie sur  $\Omega$ . On appelle espérance de  $X$  et l'on note  $E(X)$  le nombre

$$\sum_{\omega \in \Omega} X(\omega)p_{\omega}$$

lorsqu'il est bien défini. L'espérance est donc définie dans les deux cas suivants :

1. La famille  $(X(\omega)p_{\omega})_{\omega \in \Omega}$  est sommable;
2. La famille  $(X(\omega)p_{\omega})_{\omega \in \Omega}$  est non sommable et  $X$  est à valeurs positives ou nulles.

Dans le second cas  $E(X) = +\infty$ .

Comme dans le cas fini, l'espérance d'une variable ne dépend que de sa loi, voici ce que nous murmure la proposition suivante.

**Proposition 5.5.7.** — Soit  $X$  une variable aléatoire réelle définie sur  $\Omega$  ayant une espérance finie si et seulement si  $(xP(X=x))_{x \in E'}$  est sommable et si tel est le cas :

$$E(X) = \sum_{x \in E'} xP(X=x).$$

En fin d'année lorsque nous étudierons les variables aléatoires définies sur un univers non dénombrable, nous ne pourrons faire autrement que de définir l'espérance par cette dernière formule.

*Preuve de la proposition 5.5.7.* —

**Exercice** — Soit un réel  $\lambda$ . Montrer qu'une variable aléatoire qui suit une loi de Poisson de paramètre  $\lambda$  admet une espérance à déterminer.

*Solution* —

Pour finir donnons une expression de l'espérance très utile dans la pratique et dont la preuve est à rapprocher de la transformation d'Abel vue en exercice.

**Exercice** — Soit  $X$  une variable aléatoire définie sur  $\Omega$  à valeurs dans  $\mathbf{N}$ . Montrer que :

$$E(X) = \sum_{n \in \mathbf{N}} \mathbf{P}(X > n) \leq +\infty.$$

*Solution* —

## Chapitre V

# RÉVISIONS SUR LES FONCTIONS À VALEURS RÉELLES ; FONCTIONS CONVEXES

Dans ce bref chapitre nous rappellerons quelques résultats de MPSI le plus souvent sans démonstration, mais agrémentés de remarques et d'exercices. Il ne se veut pas une reprise exhaustive du cours de première année, mais une promenade au travers du programme de MPSI sur les fonctions, prétexte à introduire en fin de chapitre la notion de fonction convexe, qui, il faut bien le dire, vient dans le programme de MP\* comme un cheveu dans la soupe.

Dans tout le chapitre  $f$  désignera une application définie sur un intervalle  $I$ , non réduit à un point, et à valeurs réels. On note  $a$  l'extrémité inférieure de  $I$ ,  $b$  l'extrémité supérieure :  $-\infty \leq a < b \leq +\infty$ .

## RÉVISIONS DE MPSI

### 1.1 Limite-continuité

La plupart des propriétés des limites vues en MPSI ont été revues cette année dans le cours de topologie dans le cas général des e.v.n. Toutefois le fait que la topologie de  $\mathbf{R}$  est liée à l'ordre (les boules ouvertes sont des intervalles donc définies par des inégalités), donne un théorème propre aux fonctions à valeurs réelles : le théorème de la limite monotone.

**Proposition 1.1.1** — THÉORÈME DE LA LIMITE MONOTONE —

*On suppose  $f$  monotone alors  $f$  admet une limite à droite et une limite à gauche finies, en tout point intérieur à  $I$ , une limite généralisée dans  $\bar{\mathbf{R}}$  à gauche en  $a$ , une limite généralisée à droite dans  $\bar{\mathbf{R}}$  en  $b$ .*

Donnons un résultat moins concis et plus précis. On se limitera dans l'énoncé au cas où  $f$  est croissante, à charge pour le lecteur de traduire dans le cas décroissant ces résultats qui se déduisent avec une goutte de bon sens des précédents et se prouvent en raisonnant sur  $-f$ .

**Proposition 1.1.2.** — *On suppose  $f$  croissante. Pour tout réel  $x_0$  intérieur à  $I$ ,*

$$\sup_{x \in ]-\infty, x_0[ \cap I} f(x) = \lim_{x \rightarrow x_0^-} f(x) \leq f(x_0) \leq \lim_{x \rightarrow x_0^+} f(x) = \inf_{x \in ]x_0, +\infty[ \cap I} f(x).$$

*Si  $f$  est définie en  $a$  alors :*

$$f(a) \leq \lim_{x \rightarrow a^+} f(x) = \inf_{x \in ]a, +\infty[ \cap I} f(x).$$

*Si  $f$  n'est pas définie en  $a$  alors,*

$$\begin{cases} \lim_{x \rightarrow a^+} f(x) = \inf_{x \in ]a, +\infty[ \cap I} f(x) & \text{si } f \text{ est minorée,} \\ -\infty & \text{sinon.} \end{cases}$$

*Si  $f$  est définie en  $b$  alors :*

$$\sup_{x \in ]-\infty, b[ \cap I} f(x) = \lim_{x \rightarrow b^-} f(x) \leq f(b).$$



Si  $f$  n'est pas définie en  $b$  alors,

$$\begin{cases} \lim_{x \rightarrow b^-} f(x) = \sup_{x \in ]-\infty, b[ \cap I} f(x) & \text{si } f \text{ est majorée,} \\ +\infty & \text{sinon .} \end{cases}$$

**Fig. 1**

cas :  $f$  croissante

*Preuve de la proposition 1.1.2.* — Traitons par exemple le cas où  $x_0$  est un point intérieur à  $I$  et montrons :

$$\lim_{x \rightarrow x_0^-} f(x) = \sup_{x \in ]-\infty, x_0[ \cap I} f(x),$$

Les autres cas sont du même métal.



Un petit exercice d'application !

**Exercice 1.1.3.** —

1. Soit  $\phi$  une application de  $\mathbf{R}_+^*$  dans  $\mathbf{R}$  croissante, telle que l'application

$$g : \mathbf{R}_+^* \rightarrow \mathbf{R}; x \mapsto \frac{\phi(x)}{x}$$

soit décroissante.

Montrer que  $\phi$  est continue.

2. Soient  $f_1$  et  $f_2$  des applications croissantes dont la somme est continue. Montrer que  $f_1$  et  $f_2$  sont continues

*Solution de l'exercice 1.1.3. —*

Ce lien entre limite et monotonie interdit à une fonction monotone d'être trop discontinue...

**Exercice 1.1.4.** — Soit  $\phi$  une application croissante d'un segment  $[a, b]$  à valeurs réelles. Montrer que l'ensemble des points de discontinuité de  $\phi$  est au plus dénombrable.

*Solution de l'exercice 1.1.4. —*

La continuité a abondamment été étudiée dans le chapitre sur les e.v.n. Le résultat principal de MPSI, le théorème de la valeur intermédiaire a été généralisé dans l'étude des parties connexes par arcs. Nous allons faire de rapides rappels sur la question.

Proposition 1.1.5. — THÉORÈME DE LA VALEUR INTERMÉDIAIRE —

*Si  $f$  est continue, alors l'image par  $f$  de tout intervalle inclus dans  $I$  est un intervalle.*

*De façon moins condensée si  $x_1$  et  $x_2$  sont des éléments de  $I$  tels que  $x_1 < x_2$ , alors pour tout réel  $y$  compris (au sens large) entre  $f(x_1)$  et  $f(x_2)$ , il existe un élément  $\zeta$  de  $[x_1, x_2]$  tel que  $f(\zeta) = y$  ( $y$  est une valeur de  $f$ )*

Ainsi par exemple si  $f$  est à valeurs rationnelles, alors elle est constante.

**Remarque** — Un cas particulier du théorème de la valeur intermédiaire s'énonce ainsi : si il existe  $x_1, x_2$  éléments de  $I$  tels que  $f(x_1)$  et  $f(x_2)$  ne soient pas de même signe ( $f(x_1)f(x_2) \leq 0$ ) alors  $f$  s'annule entre  $x_1$  et  $x_2$  (au sens large). Il est alors remarquable que cette propriété est équivalente au théorème de la valeur intermédiaire en effet....

**Exercice 1.1.6.** — On suppose que  $I = [a, b]$

1. On suppose que  $f$  est continue et  $f([a, b]) \subset [a, b]$ . Montrer que  $f$  admet un point fixe.
2. On suppose que  $f$  est continue et  $[a, b] \subset f([a, b])$ . Montrer que  $f$  admet un point fixe.
3. On suppose que  $f$  est croissante et  $f([a, b]) \subset [a, b]$ . Montrer que  $f$  admet un point fixe. Le résultat demeure-t-il si l'on remplace croissant par décroissant.

*Solution de l'exercice 1.1.6. —*

Remarquons que la proposition 1.1.5. n'assure pas l'unicité de l'élément  $\zeta$ , pour ce faire il faut de surcroît une hypothèse de stricte monotonie, comme nous allons voir maintenant.

La stricte monotonie d'une application assure clairement son injectivité. La réciproque est fausse ainsi...

est-elle une application injective non strictement monotone. Par contre on peut montrer que si  $f$  est *continue* et injective alors elle est strictement monotone. En effet...

Avec la continuité on peut aller plus loin. Si  $f$  est continue sur  $I$  et strictement monotone alors elle est injective et, compte tenue du théorème de la valeur intermédiaire,  $f$  réalise une bijection de  $I$  sur  $f(I)$  qui est un intervalle, on a même :

**Proposition 1.1.7.** — *Si  $f$  est continue sur l'intervalle  $I$  et strictement monotone alors elle réalise une bijection de  $I$  sur l'intervalle  $f(I)$ , de plus la bijection réciproque est continue. On dit que  $f$  réalise un homéomorphisme de  $I$  sur  $f(I)$ .*

**Remarques**

1. Les trois mots *continue*, *intervalle* et *strictement monotone* sont cruciaux.

2. On peut préciser la forme de  $f(I)$  en utilisant au besoin le théorème de la limite monotone. Par exemple

—  $f$  croissant : si  $I = [a, b[$  alors,  $f(I) = \dots$   
 si  $I = ]-\infty, b[$  alors,  $f(I) = \dots$

—  $f$  décroissant : si  $I = [a, b]$  alors,  $f(I) = \dots$   
 si  $I = [b, +\infty[$  alors,  $f(I) = \dots$

— Nous rappellerons en 1.2.6. une amélioration de ce résultat dans le cas d'une application de classe  $\mathcal{C}^k$ .

**Exercice 1.1.8.** — Existe-t-il une application  $\phi$  de  $\mathbf{R}$  dans  $\mathbf{R}$  continue telle que  $\phi \circ \phi = -\text{id}_{\mathbf{R}}$  ?

*Solution de l'exercice 1.1.8.* —

Qu'une application continue sur un segment soit bornée et atteigne ses bornes a été revu dans le cours de topologie ainsi que le théorème de Heine.

Terminons cette sous-partie par un exercice.

**Exercice 1.1.9.** —

On suppose que  $I = [a, b]$  et que  $f$  est continue sur  $I$ . Soient  $x_1, x_2, \dots, x_n$  des éléments de  $I$ . montrer qu'il existe un élément  $c$  de  $I$  tel que  $f(c) = \frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n}$ .

*Solution de l'exercice 1.1.9.* —

**Exercice 1.1.10.** —

Soient  $f_1$  et  $f_2$  des applications continues sur  $[a, b]$ . On suppose que  $f_1 > f_2$ . Montrer qu'il existe un réel  $c > 0$  tel que  $f_1 \geq f_2 + c$ .

*Solution de l'exercice 1.1.10.* —

## 1.2 Dérivation

Nous laissons de côté les définitions et propriétés de base de la dérivation, que nous reverrons dans le prochain chapitre généralisées à des fonctions à valeurs dans un espace vectoriel de dimension finie. Nous nous limiterons aux propriétés propres aux fonctions à valeurs réelles en premier le théorème de Rolle et ses innombrables conséquences.

Commençons par un lemme crucial.

**Lemme 1.2.1.** — *Supposons que  $f$  soit dérivable sur  $I$  et que  $f$  admette un extrémum local en un point  $x_0$  intérieur à  $I$ , alors,*

$$f'(x_0) = 0.$$

### Remarque

- La réciproque de 1.2.1 est fausse, comme le montre...
- Que  $x_0$  soit un point intérieur est une hypothèse essentielle, en effet...

On en déduit le lemme de Rolle :

**Proposition 1.2.2.** — LEMME DE ROLLE —

*Supposons que  $I$  soit le segment  $[a, b]$ , que  $f$  soit continue sur  $[a, b]$  et dérivable en tout point de  $]a, b[$ . Alors si  $f(a) = f(b)$  alors il existe un élément  $c$  de  $]a, b[$  tel que  $f'(c) = 0$ .*

*Preuve de la proposition 1.2.2.* —

Cherchons à utiliser 1.1.1.

L'application  $f$  étant continue sur le segment  $[a, b]$ , elle atteint sa borne supérieure en un point  $x_M$  de  $[a, b]$ .

- Si  $x_M \in ]a, b[$ , alors  $f$  admet *a fortiori* un extrémum local en  $x_M$ , donc d'après 1.1.1.

$$f'(x_M) = 0;$$

- sinon, toujours par continuité sur le segment  $[a, b]$ ,  $f$  atteint sa borne inférieure en un point  $x_m$  de  $[a, b]$ , et :
  - si  $x_m \in ]a, b[$  alors, toujours par 1.1.1.,

$$f'(x_m) = 0,$$

- sinon, on est dans le cas où  $f$  atteint sa borne supérieure et sa borne inférieure en  $a$  et  $b$  ( $f(a) = f(b)$ ), et donc est constante et donc, par exemple,

$$f'\left(\frac{a+b}{2}\right) = 0.$$

Dans tous les cas,  $f'$  s'annule en un point intérieur à  $I$ . □

**Application 1.2.3.** — Soit un entier  $n \geq 1$ . Supposons  $f$  de classe  $\mathcal{C}^n$  sur l'intervalle  $I$  et qu'il existe  $n+1$  réels distincts  $x_0, x_1, \dots, x_n$  en lesquels  $f$  s'annule, alors la dérivée  $n^e$  de  $f$  s'annule au moins en un point de  $I$  (et même du plus petit intervalle contenant les  $x_i$ ,  $i = 0, \dots, n$ ).

Commençons par donner une preuve relâchée...

Donons une preuve rigoureuse par récurrence.

On peut faire mieux !

**Exercice** — Supposons encore que soit  $f$  de classe  $\mathcal{C}^n$  et qu'elle admette  $k$  zéros distincts  $x_1, x_2, \dots, x_k$  d'ordres respectifs  $p_1, \dots, p_k$  de sorte que  $p_1 + \dots + p_k = n + 1$ . Montrer que la dérivée  $n^{\text{e}}$  de  $f$  s'annule au moins en un point de  $I$ .

**Application 1.2.3.** — Soit  $P$  un polynôme de degré  $n \geq 1$ , réel, scindé à racines simples. Montrons que pour  $i = 0, 1, \dots, n - 1$ ,  $P^{(i)}$  est scindé à racines simples...

On déduit de ce résultat qu'un polynôme ayant deux coefficients successifs nuls, n'est pas simplement scindé. En

effet...

Donnons en exercice une généralisation classique du lemme de Rolle :

**Exercice 1.2.4.** — On suppose que  $f$  est dérivable sur  $I = \mathbf{R}$  et que  $f$  admet en  $+\infty$  et  $-\infty$  une même limite  $\ell$ . Montrer que la dérivée de  $f$  s'annule.

*Solution de l'exercice 1.2.4.* —

Une conséquence importante du lemme de Rolle est l'égalité des accroissements finis :

**Proposition 1.2.5.** — —ÉGALITÉ DES ACCROISSEMENTS FINIS —

Supposons que  $I$  soit le segment  $[a, b]$ , que  $f$  soit continue sur  $[a, b]$  et dérivable en tout point de  $]a, b[$ . Alors il existe un élément  $c$  de  $]a, b[$  tel que :

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

Ce résultat nous dit en substance qu'un véhicule, lors d'un trajet, verra sa vitesse instantanée prendre au moins une fois la valeur de sa vitesse moyenne. Plus géométriquement il existe un point  $c$  tel que la tangente au graphe de  $f$  en  $(c, f(c))$  ait comme pente celle de la corde du graphe entre  $(a, f(a))$  et  $(b, f(b))$ . Cette dernière remarque inspire la preuve de ce résultat.

Une première conséquence de l'égalité des accroissements finis est l'étude des variations de  $f$  par le signe de la dérivée, nous renvoyons au cours de MPSI.

**Attention !** pour qu'une application dérivable soit strictement croissante il n'est pas nécessaire que sa dérivée soit strictement positive (ni même strictement positive sauf en un nombre fini de points...)

**Exercice** — Donner une condition nécessaire et suffisante sur la dérivée de  $f$  pour que la fonction  $f$  soit strictement monotone.

En particulier dans le théorème de l'homéomorphisme croissant 1.1.7., la stricte monotonie peut résulter de l'étude de la dérivée : pour une fonction de classe  $\mathcal{C}^1$ , si la dérivée ne s'annule pas alors elle est de signe constant et l'application est strictement monotone. Reste alors à rappeler le résultat suivant sur la dérivée d'une bijection réciproque pour donner sous des hypothèses plus fortes une nouvelle version de 1.1.7.

**Proposition 1.2.6** — Supposons que  $f$  soit dérivable et induise un homéomorphisme de  $I$  sur  $f(I)$ ; pour que l'homéomorphisme réciproque  $g$  soit dérivable en  $f(t_0)$ , où  $t_0$  est un point de  $I$ , il faut et il suffit que  $f'(t_0)$  soit non nulle.

Si tel est le cas, alors  $g'(f(t_0)) = \frac{1}{f'(t_0)}$ .

D'où :

**Proposition 1.2.7.** — supposons  $f$  de classe  $\mathcal{C}^k$ , où  $k$  désigne un entier supérieur ou égal à 1. Si  $f'$  ne s'annule pas alors  $f$  induit une bijection de  $I$  sur  $f(I)$  qui est un intervalle, de plus la bijection réciproque est de classe  $\mathcal{C}^k$ . On dit que  $f$  induit un  $\mathcal{C}^k$ -difféomorphisme<sup>1</sup> de  $I$  sur  $f(I)$ .

Voici une autre application de l'égalité des accroissements finis.

**Proposition 1.2.8.** — On suppose  $f$  continue sur  $I$  et dérivable en tout point de  $I \setminus \{x_0\}$ . Si

$$f'(x) \xrightarrow[x \neq x_0]{x \rightarrow x_0} \ell,$$

où  $\ell$  est un élément de  $\bar{\mathbf{R}}$ , alors

$$\frac{f(x) - f(x_0)}{x - x_0} \xrightarrow[x \neq x_0]{x \rightarrow x_0} \ell.$$

En particulier si  $\ell$  est réel, alors  $f$  est dérivable en  $x_0$  et la dérivée de  $f$  est continue en  $x_0$ .

Une autre conséquence importante de 1.2.5. est le théorème de prolongement de la dérivabilité :

**Corollaire 1.2.9.** — THÉORÈME DE PROLONGEMENT DU CARACTÈRE  $\mathcal{C}^n$  —

Soit un entier  $n \geq 1$ . On suppose  $f$  continue. Si la restriction de  $f$  à  $I \setminus \{x_0\}$  est de classe  $\mathcal{C}^n$ , et si pour  $i = 0, 1, \dots, n$ ,  $f^{(i)}(x)$  admet une limite réelle,  $\ell_i$ , lorsque  $x$  tend vers  $x_0$  par valeurs distinctes, alors  $f$  est de classe  $\mathcal{C}^n$  et pour  $i = 0, \dots, n$ , on a  $f^{(i)}(x_0) = \ell_i$ .

**Exemple 1.2.10.** — UNE FONCTION ILLUSTRE ET UTILE — Soit l'application

$$\phi : \mathbf{R} \rightarrow \mathbf{R}; x \mapsto \begin{cases} \exp\left(-\frac{1}{x}\right), & \text{si } x > 0, \\ 0, & \text{sinon.} \end{cases}$$

Montrons que  $\phi$  est  $\mathcal{C}^\infty$ .

Pour aller plus loin le lecteur pourra traiter l'exercice suivant

**Exercice** —

1. En utilisant l'exemple précédent, construire une application de classe  $\mathcal{C}^\infty$  nulle en dehors du segment  $[-1, 1]$ .

---

1. Le terme n'est plus au programme



2. Construire une application de classe  $\mathcal{C}^\infty$ , à valeurs dans  $[0, 1]$ , nulle en dehors de  $[-1, 1]$  constante égale à 1 sur  $[-\frac{1}{2}, \frac{1}{2}]$  (fonction plateau).

Étudions à présent la continuité des dérivées. Il est assez difficile d'exhiber des fonctions dérivables qui ne soient pas de classe  $\mathcal{C}^1$ , la raison en est assez profonde, d'abord la discontinuité d'une dérivée ne peut correspondre à l'image la plus simple que l'on se fait d'une discontinuité en un point  $x_0$  d'une application  $g$  : l'existence d'une limite à droite, d'une limite à gauche et les quantités  $g(x_0)$ ,  $\lim_{x \rightarrow x_0^+} g(x)$  et  $\lim_{x \rightarrow x_0^-} g(x)$  non toutes les trois égales, on parle de discontinuité de première espèce.

Une application  $g$  définie sur  $I$  et qui admet en  $x_0$  point intérieur à  $I$ , une discontinuité de première espèce n'est pas la dérivée d'une application...

Donnons maintenant l'exemple classique d'application dérivable et non continûment dérivable.

**Exemple 1.2.11.** —

Pour tout entier  $n \geq 0$  on définit les applications

$$S_n : \mathbf{R} \rightarrow \mathbf{R} \quad x \mapsto \begin{cases} x^n \sin\left(\frac{1}{x}\right), & \text{si } x \neq 0, \\ 0 \text{ sinon,} \end{cases}$$

$$C_n : \mathbf{R} \rightarrow \mathbf{R} \quad x \mapsto \begin{cases} x^n \cos\left(\frac{1}{x}\right), & \text{si } x \neq 0, \\ 0 \text{ sinon,} \end{cases}$$

Pour aller plus loin...

**Exercice** — Déterminer pour tout  $n \in \mathbf{N}$  la régularité optimale de  $S_n$  et  $C_n$ .

La discontinuité de  $S_2$  est comparable à celle  $C_0$ . Nous allons voir en exercice, que le caractère complexe des discontinuités s'explique par le fait que ces dernières satisfont à la propriété de la valeur intermédiaire.

**Exercice 1.2.12** — THÉORÈME DE DARBOUX — On suppose  $f$  dérivable. Soient  $x_1$  et  $x_2$  des éléments de  $I$  tels que  $x_1 < x_2$  et  $f'(x_1)f'(x_2) \leq 0$ . Montrer qu'il existe  $\zeta \in [x_1, x_2]$  tel que  $f'(\zeta) = 0$ .

*Solution de l'exercice 1.2.12.* —

Nous allons donner une preuve proche de celle du lemme de Rolle, pour d'autres preuves voir les feuilles

d'exercices.

Le théorème de Darboux confère aux dérivées une propriété proche de la continuité. En fait on a bâti une application dont la dérivée est non continue en un point, et il serait possible de l'utiliser pour en construire d'autres dont les dérivées sont non continues en plusieurs points, voire sur un ensemble dénombrable de points, mais on montre qu'une dérivée est toujours continue sur un « très gros » ensemble (cf. feuilles d'exercices).

A l'inverse des fonctions continues qui sont pour la plupart nulle part dérivables, les fonctions dérivables sont elles « presque  $C^1$  ».

L'inégalité des accroissements finis, les formules de Taylors seront revues dans le prochain chapitre dans un cadre plus général. Les développements limités ont été révisés en colles.

## FONCTIONS CONVEXES

### 2.1 Définition, première propriétés

La notion de fonction convexe est cruciale en analyse, notamment en optimisation.

On continue comme dans la partie précédente à considérer l'application  $f$  de  $I$  dans  $\mathbf{R}$ .

Commençons par la définition brutale.

**Définition 2.1.1.** — *L'application  $f$  est dite convexe si pour tout couple  $(x, y)$  d'éléments de  $I$ , et tout élément  $\lambda$  de  $[0, 1]$ ,*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (\text{V.1})$$

*Si  $-f$  est convexe, on dit que  $f$  est concave.*

De façon plus parlante  $f$  est convexe si l'image par  $f$  de la moyenne à coefficients positifs de deux points est inférieure à la moyenne des images. Il arrive que certains parlent de barycentre plutôt que de moyenne, ce qui est rigoureusement

la même chose lorsque l'on se place dans  $\mathbf{R}$ . Les moyennes à coefficients positifs de deux réels  $x$  et  $y$  décrivent le segment  $[x, y]$ .

**Formulation géométrique** — Soient  $x$  et  $y$  des éléments de  $I$ . Pour tout  $\lambda \in [0, 1]$ , posons  $x_\lambda = \lambda x + (1 - \lambda)y$ . Appelons *corde* du graphe de  $f$  entre  $x$  et  $y$ , le segment de  $\mathbf{R}^2$ ,  $[(x, f(x)); (y, f(y))]$ .

$$[(x, f(x)); (y, f(y))] = \{\lambda(x, f(x)) + (1 - \lambda)(y, f(y)), \lambda \in [0, 1]\} = \{(x_\lambda, \lambda f(x) + (1 - \lambda)f(y)), \lambda \in [0, 1]\}.$$

Dire que  $f$  est convexe c'est dire que .....

**Remarque** —

1. L'inégalité (V.1) est vérifiée pour toute application pour  $\lambda = 0$  et  $\lambda = 1$ . C'est pourquoi dans la définition de la convexité on aurait pu remplacer « tout élément  $\lambda$  de  $[0, 1]$  » par « tout élément  $\lambda$  de  $]0, 1[$  ».
2. On peut définir la stricte convexité par :  $f$  est dite strictement convexe si pour tout  $(x, y) \in I$ , et tout élément  $\lambda$  de  $]0, 1[$ ,

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y). \quad (\text{V.2})$$

Cette notion ne semble pas être au programme.

**Exemples**

- L'identité de  $\mathbf{R}$  dans  $\mathbf{R}$  est convexe, (mais pas strictement) et concave.

- L'application  $|\cdot|$  de  $\mathbf{R}$  dans  $\mathbf{R}$  est convexe, (mais pas strictement).

La convexité a été définie grâce à la moyenne de deux points, mais la propriété la définissant se généralise à un nombre quelconque de points.

**Proposition 2.1.2.** — INÉGALITÉ DE JANSEN —

Supposons que  $f$  soit convexe. Alors pour tout  $n$ -uplet  $(x_1, x_2, \dots, x_n)$  d'éléments de  $I$ , pour tout  $n$ -uplet  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  de réels POSITIFS tel que :  $\sum_{i=1}^n \lambda_i = 1$ , on a :

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i). \quad (\text{V.3})$$

De façon imagée, l'image de la moyenne à coefficients positifs est inférieure à la moyenne des images. Si l'on n'impose pas à la somme des  $\lambda_i$  de valoir 1, l'inégalité (V.3) prend la forme plus lourde :

$$f\left(\frac{1}{\sum_{i=1}^n \lambda_i} \sum_{i=1}^n \lambda_i x_i\right) \leq \frac{1}{\sum_{i=1}^n \lambda_i} \sum_{i=1}^n \lambda_i f(x_i).$$

*Preuve de la proposition 2.1.2.* —

□

Donnons le lien entre fonction convexe et partie convexe.

**Définition 2.1.3.** — On appelle épigraphe de  $f$  l'ensemble noté  $\mathcal{E}_f$ ,

$$\mathcal{E}_f = \{(x, y) | x \in I \text{ et } y \geq f(x)\}$$

L'épigraphe est la partie, comme son nom l'indique, située « au dessus » du graphe.

**Proposition 2.1.4.** — *L'application  $f$  est convexe si et seulement si...*

*Preuve de la proposition 2.1.4.* —

**Exercices 2.1.5.** — Soient  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications convexes sur  $I$ . On suppose définie l'application  $g = \sup_{n \in \mathbf{N}} f_n$ . Montrer que  $g$  est convexe.

*Solution de l'exercice 2.1.5.* —

Etudions à présent la pente d'une application convexe.

**Proposition 2.1.6.** — LEMME DES TROIS PENTES —

*Si l'application  $f$  est convexe alors pour tout triplet  $(x, y, z)$  d'éléments de  $I$  tels que  $x < y < z, \dots$*

*Preuve de la proposition 2.1.6. —*

**Corolaire 2.1.7.** — CROISSANCE DE LA PENTE —

*L'application  $f$  est convexe si et seulement si pour tout élément  $x_0$  de  $I$ , l'application*

$$I \setminus \{x_0\} \rightarrow \mathbf{R}; x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$$

*est croissante.*

*Preuve de la proposition 2.1.7. —*

La régularité des fonctions convexes n'est pas au programme. Les élèves intéressés sont invités à regarder les feuilles d'exercices. Dans la partie suivante nous supposerons les applications régulières et nous nous intéresserons à leur convexité.

## 2.2 Fonctions convexes dérivables, deux fois dérivables

**Proposition 2.2.1.** — *On suppose  $f$  dérivable. Alors  $f$  est convexe si et seulement si  $f'$  croît.*

**Remarques** —

1. Il existe des applications convexes non dérivables, telle  $|\cdot|$  sur  $\mathbf{R}$ .
2. Cette proposition montre que pour une fonction dérivable, la convexité au voisinage de chaque point assure la convexité. Ce résultat est encore vraie pour une fonction non supposée dérivable.

*Preuve de 2.2.1. —*

- HYPOTHÈSE :  $f$  est convexe.

- HYPOTHÈSE :  $f'$  croît.

Montrons la croissance des pentes. Soit  $x_0 \in I$ . Considérons l'application

$$\theta_{x_0} : I \setminus \{x_0\} \rightarrow \mathbf{R}; x \mapsto \frac{f(x) - f(x_0)}{x - x_0}.$$





**Exercice** — On suppose  $f$  convexe et dérivable. Montrer que si  $f'$  s'annule en  $x_0$  alors  $f$  atteint son minimum absolue en  $x_0$ .

*Solution* —

**Corollaire 2.2.2.** — On suppose  $f$  deux fois dérivable. Alors  $f$  est convexe si et seulement si  $f''$  est positive ou nulle.

*Preuve du corollaire 2.2.2.* —

**Exercice** — Soit  $\alpha$  un réel positif ou nul. On pose

$$f_\alpha : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto x^\alpha,$$

en convenant que  $0^\alpha = 0$ . Etudiez la convexité (ou la concavité) de  $f_\alpha$  en fonction de  $\alpha$ , de la fonction exponentielle et de la fonction logarithme.

*Solution* —

Passons à une interprétation géométrique.

**Proposition 2.2.3** — *Supposons  $f$  dérivable et convexe. Alors le graphe de  $f$  est au dessus de ses tangentes, c'est-à-dire que pour tout élément  $x_0$  de  $I$ , l'application*

$$I \rightarrow \mathbf{R}; \dots$$

*est positive.*

*Preuve de la proposition 2.2.3.* —

## 2.3 Exemple d'inégalités de convexité

Donnons quelques inégalités classiques utilisant la convexité.

Les deux inégalités suivantes et leur preuve sont à connaître par coeur.

pour tout réel $x > -1$ , $\ln(1+x) \leq x$
---

pour tout réel $x$ , $\exp(x) \geq 1+x$
---

En effet...

**Remarques** —

1. Ces inégalités peuvent évidemment être démontrées aussi par une étude de fonction.
2. Ces inégalités sont en fait strictes pour  $x \neq 0$ , il faut pour le prouver, soit recourir à la stricte convexité (hors programme), soit recourir à une étude de fonction plus couteuse.

Voici un encadrement classique du sinus à connaître.

$$\text{pour tout } x \in \left[0, \frac{\pi}{2}\right], \quad \frac{2}{\pi}x \leq \sin(x) \leq x$$

En effet..

On sait par la positivité de  $(x - y)^2$ , inégalité fort utile déjà rencontrée, que  $xy \leq \frac{1}{2}(x^2 + y^2)$ . Nous allons généraliser.

Soit  $p$  et  $q$  de réels strictement supérieurs à 1 tels que  $\frac{1}{p} + \frac{1}{q} = 1$ . Alors pour tout couple  $(x, y)$  de réels positifs,

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q \text{ (inégalité de Young)}$$

En effet...

**Exercice** — INÉGALITÉ DE HÖLDER — Soient  $(a_1, \dots, a_n)$  et  $(b_1, \dots, b_n)$  des  $n$ -uplets de réels positifs,  $p$  et  $q$  sont tels que dans l'inégalités de Young. Montrer que

$$\sum_{i=1}^n a_i b_i \leq \left( \sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n b_i^q \right)^{\frac{1}{q}}$$

Quelle inégalité obtient-on lorsque  $p = q = 2$  ?



## Chapitre VI

# FONCTIONS D'UNE VARIABLE RÉELLE, APPROXIMATION

## APPROXIMATION UNIFORME

On va chercher à définir de diverses manières la convergence de suites d'applications. Pour des suites d'applications bornées d'un ensemble  $X$  à valeurs dans  $\mathbf{R}$  on peut, on l'a vu, considérer la convergence dans l'e.v.n.,  $(\mathcal{B}(X, \mathbf{R}), N_\infty)$ . On a également considéré la convergence de suite d'éléments de  $\mathcal{C}^0([a, b], \mathbf{R})$  pour les normes  $N_1$ ,  $N_2$  ou  $N_\infty$ . Dans cette partie nous allons donner des notions plus générales de convergence qui pourront s'appliquer à des applications non continues ni bornées.

Dans cette partie  $\mathbf{K}$  désignera indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$ .

### 1.1 Convergence simple et uniforme

Dans ce paragraphe  $A$  désigne une partie d'un  $\mathbf{K}$ -espace vectoriel de dimension finie  $\mathbf{E}$ , cet espace sera muni d'une norme notée  $\|\cdot\|$ . Toutes les applications considérées seront à valeurs dans  $\mathbf{K}$  ou, éventuellement dans un  $\mathbf{K}$ -espace vectoriel de dimension finie  $\mathbf{F}$ . Dans les preuves, pour ne pas les alourdir, nous traiterons que du cas d'applications à valeurs dans  $\mathbf{K}$ , le cas d'applications à valeurs dans  $\mathbf{F}$  se traitant de même, il suffit de remplacer  $|\cdot|$ , module ou valeur absolue par une norme quelconque sur  $\mathbf{F}$ ,  $\|\cdot\|_{\mathbf{F}}$ .

Dans toute la suite, sauf mention contraire,  $(f_n)_{n \in \mathbf{N}}$  désigne une suite d'applications de  $A$  dans  $\mathbf{K}$  (ou éventuellement  $\mathbf{F}$ ).

Les résultats énoncés seront évidemment généralisables à des suites indicées, non par  $\mathbf{N}$ , mais par une partie de  $\mathbf{N}$  de la forme  $\{n \in \mathbf{N}, n \geq n_0\}$ , suites qui se ramènent par réindexation à celles du type précédent. Dans les exercices et exemples nous rencontrerons de telles suites.

Commençons par la notion la plus élémentaire de convergence la convergence *simple ou ponctuelle*.

**Définition 1.1.1.** — CONVERGENCE SIMPLE OU PONCTUELLE —

On dit que la suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement, si, par définition, pour tout élément  $\vec{x}$  de  $A$ , la suite d'éléments de  $\mathbf{K}$  (ou de  $\mathbf{F}$ )  $(f_n(\vec{x}))_{n \in \mathbf{N}}$  converge.

Si la suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement, alors, en notant pour tout élément  $\vec{x}$  de  $A$ ,  $f(\vec{x})$  la limite de la suite  $(f_n(\vec{x}))_{n \in \mathbf{N}}$ , on dispose d'une application

$$f : A \rightarrow \mathbf{K} \text{ ( ou } \mathbf{F} \text{ ) ; } \vec{x} \mapsto f(\vec{x}).$$

Cette application est appelée *limite simple* de la suite  $(f_n)_{n \in \mathbf{N}}$ . On dit encore que la suite  $(f_n(\vec{x}))_{n \in \mathbf{N}}$  converge simplement vers  $f$ .

Une conséquence immédiate de l'unicité de la limite d'une suite d'éléments de  $\mathbf{K}$  (ou  $\mathbf{F}$ ) est que si  $(f_n)_{n \in \mathbf{N}}$  admet  $f$  et  $g$  comme limites simples, alors  $f = g$  (unicité de la limite simple), on peut donc parler de **LA** limite simple d'une suite d'applications.

**Exemple 1.1.2.** — On prend  $A = [0, 1]$ ,  $\mathbf{K} = \mathbf{R}$  et pour tout entier naturel  $n$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; x \mapsto x^n.$$

La suite  $(f_n)_{n \in \mathbf{N}}$  converge donc simplement vers

$$f : [0, 1] \rightarrow \mathbf{R}; x \mapsto$$

**Exemple 1.1.3.** — On prend  $A := \mathbf{C}$ ,  $\mathbf{K} = \mathbf{C}$  et pour tout entier naturel  $n$ ,

$$f_n : \mathbf{C} \rightarrow \mathbf{C}; z \mapsto \frac{z^n}{n!}.$$

La suite  $(f_n)_{n \in \mathbf{N}}$  converge donc simplement vers

Passons à la convergence simple des séries d'applications.

**Définition 1.1.3.** — CONVERGENCE SIMPLE D'UNE SÉRIE —

- On dit que la série d'applications, de terme général  $f_n$ , notée  $\sum f_n$  converge simplement si, par définition, pour tout élément  $\vec{x}$  de  $A$ , la série d'éléments de  $\mathbf{K}$  (de  $\mathbf{F}$ ),  $\sum f_n(\vec{x})$  converge.
- Pour tout élément  $n$  de  $\mathbf{N}^*$ , on note  $S_n$  l'application de  $A$  dans  $\mathbf{K}$  ( $\mathbf{F}$ ) définie par

$$S_n := \sum_{k=0}^n f_k,$$

quantité appelée (application) somme partielle d'ordre  $n$  de la série.

- Enfin si la série converge simplement, on dispose des applications

$$S : A \rightarrow \mathbf{K}(\mathbf{F}); \vec{x} \mapsto \sum_{k=0}^{+\infty} f_k(\vec{x})$$

et, pour tout naturel  $n$ ,

$$R_n : A \rightarrow \mathbf{K}(\mathbf{F}); \vec{x} \mapsto \sum_{k=n+1}^{+\infty} f_k(\vec{x}).$$

La première s'appelle (application) somme on la note  $\sum_{n=0}^{+\infty} f_n$ , la seconde (application) reste d'ordre  $n$ , de la série d'applications  $\sum f_n$ . Pour tout entier naturel  $n$ ,  $S = S_n + R_n$ .

De façon évidente, on peut reformuler la convergence simple d'une série en terme de suites :

**Proposition 1.1.4.** — La série d'applications  $\sum f_n$  converge simplement si et seulement si, la suite  $(S_n)_{n \in \mathbf{N}}$  de ses sommes partielles converge simplement.

Si c'est le cas la somme  $S$  de la série est alors la limite simple de  $(S_n)_{n \in \mathbf{N}}$  et la suite  $(R_n)_{n \in \mathbf{N}}$  des restes de la série converge simplement vers l'application nulle sur  $A$ .

**Exemple 1.1.5.** —

1. On prend d'abord, pour tout entier naturel  $n$ ,

$$f_n : \mathbf{C} \rightarrow \mathbf{C}; z \mapsto \frac{z^n}{n!}.$$

La série  $\sum f_n$  converge simplement vers

2. Soit  $m$  un élément de  $\mathbf{N}^*$ . On note  $\|\cdot\|$  la norme usuelle définie sur  $\mathcal{M}_m(\mathbf{R})$  par, pour tout  $M \in \mathcal{M}_m(\mathbf{R})$ ,

$$\|M\| = \sup_{X \in \mathcal{M}_{m,1}(\mathbf{R}), \|X\|_1 \leq 1} \|MX\|_1.$$

On note  $A$  l'ensemble des éléments  $M$  de  $\mathcal{M}_m(\mathbf{R})$  tels que  $\|M\| < 1$  et pour tout  $n \in \mathbf{N}$ , on pose

$$f_n : A \rightarrow \mathcal{M}_m(\mathbf{R}); M \mapsto M^n.$$

La série  $\sum f_n$  converge simplement et sa somme est l'application

$$f : A \rightarrow \mathcal{M}_m(\mathbf{R}); M \mapsto \dots$$

Donnons une notion plus forte et très importante de convergence : la *convergence uniforme*. Donnons brutalement, pour commencer, la définition, de nombreux commentaires suivront.

**Définition 1.1.6.** — CONVERGENCE UNIFORME —

Soit  $f$  une application de  $A$  dans  $\mathbf{K}$  ( $\mathbf{F}$ ). On dit que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$  si, par définition, pour tout réel strictement positif  $\varepsilon$ , il existe un entier naturel  $n_0$  tel que pour tout élément  $n$  de  $\mathbf{N}$  et tout élément  $\vec{x}$  de  $A$ , si  $n \geq n_0$ , alors

$$|f(\vec{x}) - f_n(\vec{x})| < \varepsilon.$$

L'application  $f$  est appelée limite uniforme de la suite  $(f_n)_{n \in \mathbf{N}}$ .

On dit que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$  sur une partie  $B$  de  $A$  si, par définition, la suite des restrictions des  $f_n$  à  $B$ ,  $(f_n|_B)_{n \in \mathbf{N}}$ , converge uniformément vers  $f|_B$ , restriction de  $f$  à  $B$ .

**Remarque** — On peut évidemment dans la définition de la convergence uniforme, remplacer «  $< \varepsilon$  » par «  $\leq \varepsilon$  » et «  $n \geq n_0$  » par «  $n > n_0$  ».

#### COMPARAISON CONVERGENCE SIMPLE ET UNIFORME

La suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers  $f$  si, par définition, pour tout élément  $\vec{x}$  de  $A$ , et pour tout réel strictement positif  $\varepsilon$ , il existe un entier naturel  $n_0$  tel que, pour tout  $n \in \mathbf{N}$  si  $n \geq n_0$ , alors

$$|f(\vec{x}) - f_n(\vec{x})| < \varepsilon.$$



Quelles différence avec la définition de la convergence uniforme ?

Il ressort que :

**Proposition 1.1.7.** — *Si la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$  alors elle converge simplement vers la même application  $f$ .*

De « l'unicité de la limite simple » et de la proposition précédente on déduit « l'unicité d'une limite uniforme », précisément :

**Proposition 1.1.8.** — *Si la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$  et vers une application  $g$  alors,  $f = g$ .*

On peut donc parler de **LA** limite uniforme d'une suite

Il semble que la réciproque de 1.1.7. soit fausse. Pour mieux comprendre ce qui pourrait empêcher une suite convergeant simplement de converger uniformément, nous allons examiner soigneusement deux exemples.

**Exemples 1.1.9.** —

1. On suppose que  $A$  est un ensemble fini non vide  $\{a_1, a_2, \dots, a_p\}$  et que  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers une application  $f$ . Etudions la convergence uniforme.

2. On considère la suite  $(f_n)_{n \in \mathbf{N}^*}$ , où, pour tout  $n \in \mathbf{N}^*$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; x \mapsto \begin{cases} 1 - nx, & \text{pour } 0 \leq x \leq \frac{1}{n}, \\ 0, & \text{pour } \frac{1}{n} < x. \end{cases}$$

La suite  $(f_n)_{n \in \mathbf{N}^*}$  converge simplement vers

$$f : [0, 1] \rightarrow \mathbf{R}; x \mapsto \dots$$

Etudions la convergence uniforme. La convergence simple assure en particulier, que pour tout  $x \in [0, 1]$ , il existe un entier  $n_0$  tel que, pour tout entier naturel  $n \geq 0$ ,

$$|f(\vec{x}) - f_n(\vec{x})| < \frac{1}{2} \tag{VI.1}$$

Notons, pour tout  $x \in [0, 1]$ ,  $N_x$  l'ensemble des entiers naturels  $n_0$  tels que (VI.1) soit satisfaite pour tout  $n \geq n_0$  et  $n_x$  le plus petit élément de cette partie non vide de  $\mathbf{N}$ .

Étudions  $n_x \dots$

Il n'y a donc pas convergence uniforme.

On verra dans la suite que pour prouver qu'il n'y a pas convergence uniforme, on procède de manière bien différente et plus simple.

Passons à la convergence uniforme des séries.

**Définition 1.1.10.** — On dit que la série d'applications  $\sum f_n$  converge uniformément si, par définition la suite de ses sommes partielles  $(S_n)_{n \in \mathbf{N}}$  converge uniformément.

**Remarque 1.1.11.** — On suppose que  $\sum f_n$  converge uniformément.

1. *A fortiori*  $(S_n)_{n \in \mathbf{N}}$  converge simplement et donc la série  $\sum f_n$  converge simplement. En notant  $S$  la somme de la série  $\sum f_n$ , la suite  $(S_n)_{n \in \mathbf{N}}$  converge donc uniformément vers  $S$ .
2. La définition même de la convergence uniforme de  $(S_n)_{n \in \mathbf{N}}$  vers  $S$  dit que la suite des restes  $(R_n)_{n \in \mathbf{N}}$  converge uniformément vers l'application nulle sur  $A$  (en effet, pour tout  $n \in \mathbf{N}$ ,  $S - S_n = R_n$ ).

#### ÉTUDE PRATIQUE DE LA CONVERGENCE UNIFORME

Pour étudier la suite d'applications  $(f_n)_{n \in \mathbf{N}}$  on procède souvent en deux temps.

1. *Étude de la convergence simple.*  
On détermine la limite simple de la suite notée  $f$ .
2. *Étude de la convergence uniforme vers  $f$*   
On prend  $\vec{x}$  élément de  $A$  et on cherche à majorer, pour tout  $n \in \mathbf{N}$ ,  $|f(\vec{x}) - f_n(\vec{x})|$  par une quantité  $a_n$  indépendante de  $\vec{x}$  et telle que  $a_n \xrightarrow{n \rightarrow +\infty} 0$ . En effet on a le résultat suivant :

**Proposition 1.1.12.** — *Si, il existe une suite de nombres réels  $(a_n)_{n \in \mathbf{N}}$  telle que*

- 1. Pour tout entier naturel  $n$ , et tout élément  $\vec{x}$  de  $A$ ,*

$$|f(\vec{x}) - f_n(\vec{x})| \leq a_n.$$

- 2. La suite  $(a_n)_{n \in \mathbf{N}}$  converge vers 0.*

*Alors la suite d'applications  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$ .*

*Preuve de la proposition 1.1.12. —*

On peut parfois étudier une série d'applications en appliquant la précédente démarche à la suite de ses sommes partielles.

Passons à la pratique :

**Exemples 1.1.13.** —

1. On reprend l'exemple 1.1.2. Pour tout  $n \in \mathbf{N}$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; x \mapsto x^n.$$

- (a) *Etude de la convergence simple.*

On a vu que la suite convergeait simplement vers l'application

$$f : [0, 1] \rightarrow \mathbf{R}; x \mapsto \dots$$

- (b) *Etude de la convergence uniforme.*

Soit  $x$  élément de  $[0, 1]$ .

Donc la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément sur  $\dots$

On montrera qu'il n'y a pas de convergence uniforme sur  $[0, 1]$ .

2. On reprend l'exemple 1.1.5. On étudie la série d'applications  $\sum f_n$ , où, pour tout entier naturel  $n$ ,

$$f_n : \mathbf{C} \rightarrow \mathbf{C}; z \mapsto \frac{z^n}{n!}.$$

(a) *Etude de la convergence simple.*

On a vu que la série convergeait simplement de somme  $\dots$

(b) *Etude de la convergence uniforme.*

Soit  $z$  élément de  $\mathbf{C}$ .

Donc la série  $\sum f_n$  converge uniformément sur  $\dots$

On montrera qu'il n'y a pas de convergence uniforme sur  $\mathbf{C}$ .

**Exercice** – Montrer que dans les exemples précédents, il n'y a pas convergence uniforme sur le domaine de définition entier.

*Solution* — Traitons par exemple l'exemple  $\dots$

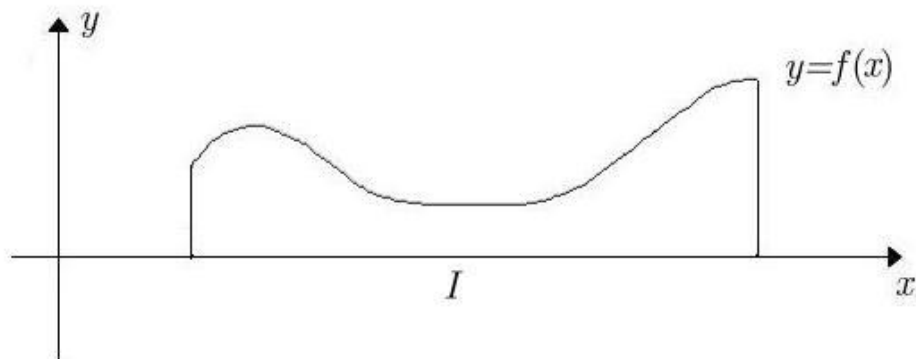


FIGURE VI.1 –

La suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$  se traduit par : *Pour tout réel strictement positif  $\varepsilon$ , il existe un entier  $n_0$  tel que ...*

On peut donner une image qui quoique caricaturale est assez parlante, de la convergence simple non uniforme (fig. VI.2), puis de la convergence uniforme (fig. VI.3) vers l'application nulle.

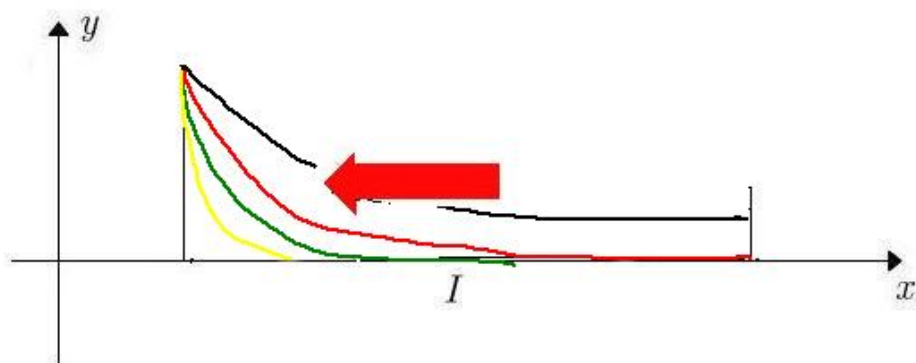


FIGURE VI.2 –

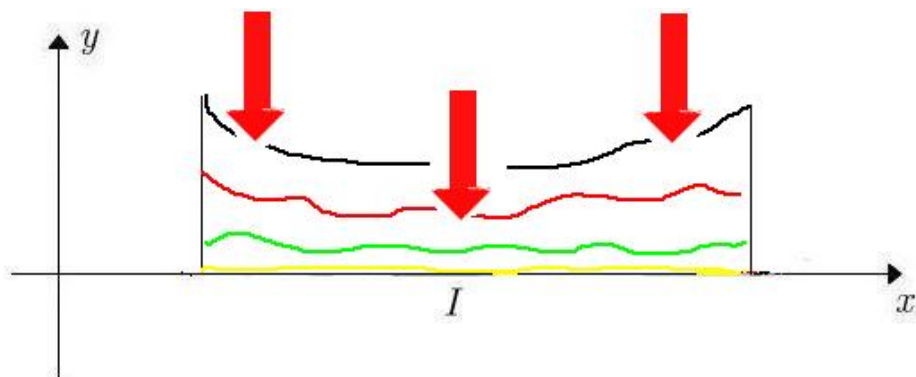


FIGURE VI.3 –

#### CONTINUITÉ DE LA LIMITE

Dans l'exemple 1.1.9.2. Pour tout élément  $n$  de  $\mathbf{N}^*$ ,  $f_n$  est continue. La suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers une application non continue. Donnons un autre exemple :

**Exemple 1.1.14.** —

On considère la suite  $(f_n)_{n \in \mathbf{N}^*}$ , où, pour tout  $n \in \mathbf{N}^*$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; x \mapsto \begin{cases} 2nx, & \text{pour } 0 \leq x \leq \frac{1}{2n}, \\ 2 - 2nx, & \text{pour } \frac{1}{2n} \leq x \leq \frac{1}{n}, \\ 0, & \text{pour } \frac{1}{n} < x. \end{cases}$$

pour tout élément  $n$  de  $\mathbf{N}^*$ ,  $f_n$  est continue. La suite  $(f_n)_{n \in \mathbf{N}^*}$  converge simplement vers l'application...

La convergence n'est pas uniforme

On vient de voir, à travers cet exemple et l'exemple 1.1.9.2, qu'une limite simple non uniforme d'une suite d'applications continues peut être, ou ne pas être, continue. Nous allons voir qu'il n'en est pas de même pour une limite uniforme.

**Proposition 1.1.15.** — *On suppose que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$ . Alors :*

1. *Soit  $\vec{x}_0$  un élément de  $A$ . Si pour tout entier naturel  $n$ ,  $f_n$  est continue en  $\vec{x}_0$ , alors  $f$  est continue en  $\vec{x}_0$ .*
2. *Si pour tout entier naturel  $n$ ,  $f_n$  est continue, alors  $f$  est continue.*

*Preuve de la Proposition 1.1.15.* — Le second point découle directement du premier que nous allons maintenant prouver.

Soit  $\varepsilon$  un réel strictement positif. Pour tout élément  $\vec{x}$  de  $A$  et pour tout entier naturel  $n$ ,

$$|f(\vec{x}) - f(\vec{x}_0)| \leq \dots$$

Le programme de MP offre un théorème plus général, qui porte sur la notion de limite. Le voila.

**Proposition 1.1.16.** — *On suppose que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$ .*

*Soit  $\vec{a}$  un point de  $\mathbf{E}$  adhérent à  $A$ . On suppose que pour tout entier naturel  $n$ ,  $f_n$  admet en  $\vec{a}$  une limite notée  $\ell_n$ . Alors*

- 1. La suite  $(\ell_n)_{n \in \mathbf{N}}$  converge vers un élément noté  $\ell$ .*
- 2. L'application  $f$  admet  $\ell$  comme limite en  $\vec{a}$ .*

*Dans le cas où  $\mathbf{E} = \mathbf{R}$  et où tout voisinage de  $+\infty$  (resp.  $-\infty$ ) rencontre  $A$ , le résultat demeure en remplaçant  $\vec{a}$  par  $+\infty$  (resp.  $-\infty$ ).*

*Commentaires :* De façon brutale mais parlante le théorème s'écrit :

$$\boxed{\lim_{\vec{a}} \left( \lim_{\vec{n}} f_n \right) = \lim_n \left( \lim_{\vec{a}} f_n \right)}$$

$$\begin{array}{ccc}
f_0(\vec{x}) & \xrightarrow{\vec{x} \rightarrow \vec{a}} & \ell_0 \\
f_1(\vec{x}) & \xrightarrow{\vec{x} \rightarrow \vec{a}} & \ell_1 \\
f_2(\vec{x}) & \xrightarrow{\vec{x} \rightarrow \vec{a}} & \ell_2 \\
\vdots & & \vdots \cdot \\
f_n(\vec{x}) & \xrightarrow{\vec{x} \rightarrow \vec{a}} & \ell_n \\
\text{CU} \downarrow n \rightarrow \infty & & \\
f(\vec{x}) & & 
\end{array}$$

Sans l'hypothèse de convergence uniforme ce résultat peut être mis en défaut. Par exemple on considère la suite  $(f_n)_{n \in \mathbf{N}}$  où pour tout entier naturel  $n$ ,

$$f_n : \quad \quad \rightarrow \mathbf{R}; \quad x \mapsto$$

*Preuve de la proposition 1.1.16. —*



Les propositions 1.1.15–16. peuvent s'adapter aux séries

**Proposition 1.1.17.** — *On suppose que la série  $\sum f_n$  converge uniformément, de somme  $S$ . Alors :*

1. *Soit  $\vec{x}_0$  un élément de  $A$ . Si pour tout entier naturel  $n$ ,  $f_n$  est continue en  $\vec{x}_0$ , alors  $S$  est continue en  $\vec{x}_0$ .*
2. *Si pour tout entier naturel  $n$ ,  $f_n$  est continue, alors  $S$  est continue.*

*Preuve de la proposition 1.1.17. —*

**Proposition 1.1.18.** — *On suppose que la série  $\sum f_n$  converge uniformément, de somme  $S$ .*

*Soit  $\vec{a}$  un point de  $\mathbf{E}$  adhérent à  $A$ . On suppose que pour tout entier naturel  $n$ ,  $f_n$  admet en  $\vec{a}$  une limite notée  $\ell_n$ . Alors*

1. *La*

2. *L'application  $S$  admet* *comme limite en  $\vec{a}$ .*

*Dans le cas où  $\mathbf{E} = \mathbf{R}$  et où tout voisinage de  $+\infty$  (resp.  $-\infty$ ) rencontre  $A$ , le résultat demeure en remplaçant  $\vec{a}$  par  $+\infty$  (resp.  $-\infty$ ).*

*Preuve de la proposition 1.1.18. —*

Montrons sur des exemples comment étudier la continuité de la somme d'une suite ou d'une série d'applications. Dans la pratique on a rarement la convergence uniforme sur tout le domaine de définition mais sur des parties de ce domaine. On peut néanmoins parfois montrer la continuité de la limite ou de la somme.

**Exemple 1.1.19.** — Soit la série  $\sum_{n \geq 1} f_n$ , où pour tout  $n \in \mathbf{N}^*$ ,

$$f_n : [0, 1[ \rightarrow \mathbf{R}; t \mapsto \frac{t^n}{\sqrt{n}\sqrt{n+1}}.$$

Montrer que la série converge simplement. Sa somme est-elle continue ?

**Exemple 1.1.20.** —

Soit la série  $\sum f_n$ , où pour tout  $n \in \mathbf{N}$ ,  $f_n : \mathcal{M}_n(\mathbf{K}) \rightarrow \mathcal{M}_n(\mathbf{K}) ; M \mapsto \frac{M^n}{n!}$ . On sait (cf. cours sur les séries) que la série  $\sum f_n$  converge simplement vers l'application exponentielle :

$\mathcal{M}_m(\mathbf{K}) \rightarrow \mathcal{M}_m(\mathbf{K}); M \mapsto \exp(M)$ . Montrons la continuité de cette application.

**Exemple 1.1.21.** — Soit la série  $\sum_{n \geq 1} f_n$ , où pour tout  $n \in \mathbf{N}^*$ ,

$$f_n : [0, 1[ \times \mathbf{R}_+^* \rightarrow \mathbf{R}; (x, y) \mapsto \frac{x}{yn^2} + \frac{x^n \ln(n)}{\sqrt{n}}.$$

Montrons que la série  $\sum_{n \geq 1} f_n$  converge simplement et que sa somme est continue.

Avec les mêmes idées que celles employées dans ces exemples on peut montrer le résultat suivant, que l'on pourra au choix utiliser tel quel, ou redémontrer dans les applications.

**Proposition 1.1.22.** — *On suppose que la suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers une application  $f$ . On suppose de plus que, pour tout  $n \in \mathbf{N}$ ,  $f_n$  est continue et que pour tout point  $\vec{a}$  de  $A$  il existe un voisinage  $V_a$  relativement  $A$  de  $a$  tel que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément sur  $V_a$ , alors l'application  $f$  est continue.*

**Exercice** — (Difficile) Montrer que dans le cas général, si  $(f_n)_{n \in \mathbf{N}}$  est une suite d'applications continues sur  $A$  qui

converge uniformément sur tout compact de  $A$ , alors  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers une application  $f$  continue.

L'étude de la régularité de la limite permet de démontrer qu'une suite ou qu'une série ne converge pas uniformément. Cette faculté est assez anecdotique et peu dans le sens du programme, mais vue sa simplicité donnons néanmoins un exemple.

**Exemples** —

— Reprenons l'exemple 1.1.9.2.

— Soit la suite  $(f_n)_{n \geq 1}$ , où pour tout élément  $n \in \mathbf{N}^*$ ,

$$f_n : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto \sin\left(\frac{nx}{n+x}\right).$$

Montrez que la suite converge simplement mais non uniformément.

## LIEN ENTRE LA CONVERGENCE UNIFORME ET LA CONVERGENCE POUR LA NORME $N_\infty$ .

**Rappel** — Notons  $\mathcal{B}(A, \mathbf{K})$  (resp.  $\mathcal{B}(A, \mathbf{F})$ ) l'ensemble des applications bornées de  $A$  dans  $\mathbf{K}$  (resp.  $\mathbf{F}$ ). On dispose sur cet ensemble de la norme  $\|\cdot\|_\infty$ , définie par

$$\|g\|_\infty = \sup_{\vec{x} \in A} |g(\vec{x})| \quad (\text{resp.} \quad \sup_{\vec{x} \in A} \|g(\vec{x})\|_{\mathbf{F}}).$$

La convergence uniforme et la convergence en norme infinie, bien qu'étroitement liées diffèrent en ce sens que la première s'applique à des situations plus générales que la seconde. En effet la convergence en norme infinie n'intéresse que les applications bornées, la convergence uniforme non. Voyons cela sur un exemple.

**Exemple 1.1.23.** — Pour tout entier naturel  $n$ , non nul, posons

$$f_n : \mathbf{R}_+ \rightarrow \mathbf{R}; t \mapsto \dots$$

On ne peut pas parler de la norme infinie des  $f_n$  car ces applications ne sont pas bornées. En revanche la suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement et uniformément vers  $\dots$

Par contre, on dispose du résultat suivant.

**Proposition 1.1.24.** — *Supposons que pour tout entier naturel  $n$  l'application  $f_n$  soit bornée. Alors*

- *Si  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$ , alors*
  - i. *L'application  $f$  est bornée.*
  - ii.  *$(f_n)_{n \in \mathbf{N}}$  converge vers  $f$  dans l'espace vectoriel normé  $(\mathcal{B}(A, \mathbf{K}), \|\cdot\|_\infty)$ .*
- *Réciproquement, si  $(f_n)_{n \in \mathbf{N}}$  converge dans  $(\mathcal{B}(A, \mathbf{K}), \|\cdot\|_\infty)$  vers une application  $f$ , alors  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$ .*

**Remarques** —

1. Dans cette proposition on peut bien sûr, comme dans toute cette partie, remplacer  $\mathbf{K}$  par  $\mathbf{F}$ .
2. Il résulte de cette proposition que la norme infinie s'appelle parfois *norme de la convergence uniforme*.

*Preuve de la proposition 1.1.24.* —

PREMIER POINT : La suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$

SECOND POINT :  $(f_n)_{n \in \mathbf{N}}$  converge dans  $(\mathcal{B}(A, \mathbf{K}), \|\cdot\|_\infty)$  vers  $f$

Dans le cas où les applications  $f_n$  ne sont pas nécessairement bornées, on dispose néanmoins du résultat suivant qui lie encore la convergence uniforme et celle en norme infinie.

**Proposition 1.1.25.** —

- Si  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$ , alors pour  $n$  suffisamment grand, l'application  $f - f_n$  est bornée et

$$\|f - f_n\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

- Réciproquement, si  $f_n - f$  est bornée pour  $n$  suffisamment grand et si  $\|f - f_n\|_\infty \xrightarrow{n \rightarrow +\infty} 0$ , alors  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers  $f$ .

*Preuve de la proposition 1.1.25.* — La preuve se calque sur celle de la précédente proposition, le détail en est laissé au lecteur.

**Application** — On reprend l'exemple 1.1.14.

Si l'on suppose de plus que  $A$  est compact, alors toute application continue sur  $A$  à valeurs dans  $\mathbf{K}$  est, on l'a vu dans le premier chapitre, bornée, autrement dit :

$$\mathcal{C}^0(A, \mathbf{K}) \subset \mathcal{B}(A, \mathbf{K})$$

On a de plus

**Corollaire 1.1.26.** — Si  $A$  est compact alors l'ensemble des applications continues de  $A$  dans  $\mathbf{K}$ ,  $\mathcal{C}^0(A, \mathbf{K})$ , est une partie fermée de  $(\mathcal{B}(A, \mathbf{K}), \|\cdot\|_\infty)$ .



*Preuve de 1.1.26. —*

## 1.2 Convergence normale des séries d'applications

Dans cette partie, comme dans la précédente,  $A$  désigne une partie non vide de l'espace vectoriel de dimension finie  $\mathbf{E}$  et  $(f_n)_{n \in \mathbf{N}}$  est toujours une suite d'applications de  $A$  dans  $\mathbf{K}$  ou dans  $\mathbf{F}$ .

Brutalement la définition :

**Définition 1.2.1** *On dit que la série d'applications  $\sum f_n$  converge normalement si, par définition :*

1. *Pour tout entier naturel  $n$ ,  $f_n$  est bornée.*
2. *La série à termes positifs  $\sum \|u_n\|_\infty$  converge.*

Historiquement, *normal* signifie *ordinaire*, car il s'agit dans les cas concrets du mode de convergence le plus banal. On peut aussi, pour se souvenir de la définition, considérer que la convergence normale est la convergence de « séries des normes ».

Il existe un critère simple de convergence normale :

**Proposition 1.2.2.** — *la série d'applications  $\sum f_n$  converge normalement si et seulement si, il existe une suite de réels positifs  $(a_n)_{n \in \mathbf{N}}$  telle que :*

1. *Pour tout élément  $n$  de  $\mathbf{N}$  et tout élément  $\vec{x}$  de  $A$ ,*

$$|u_n(\vec{x})| \leq a_n.$$

2. *La série de termes positifs  $\sum a_n$  converge.*

*Preuve de la proposition 1.2.2. —*

Pour montrer dans la pratique que  $\sum f_n$  converge normalement, très souvent, on ne calcule pas  $\|f_n\|_\infty$ , puisque le critère, d'un usage plus simple, s'applique.

**Exemple** — Pour tout  $n \in \mathbf{N}^*$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; x \mapsto \frac{\sin(nx)^3 + {}^n\sqrt{x} - \sin x}{1 + n^2 + \sin(x)}.$$

Pour montrer qu'il n'y a pas de convergence normale, on doit souvent calculer  $\|f_n\|_\infty$ .

**Exemple** — Pour tout  $n \in \mathbf{N}^*$ ,

$$f_n : [0, 1] \rightarrow \mathbf{R}; x \mapsto \frac{x^n \exp(-x)}{\sqrt{n}}.$$

La convergence normale est très utile car elle constitue, comme nous allons voir dans la proposition suivante, le mode de convergence pour une série le plus fort, et que, malgré cela, son étude est très facile.

**Proposition 1.2.3.** — *On suppose que la série d'applications  $\sum f_n$  converge normalement. Alors :*

1. *la série d'applications  $\sum f_n$  converge uniformément.*
2. *la série d'applications  $\sum |f_n|$  converge uniformément (convergence uniforme absolue de  $\sum f_n$ ).*

*Preuve de la proposition 1.2.3. —*

**Remarque pratique** — Pour montrer qu'il y a convergence uniforme on commence par essayer de prouver qu'il y a convergence normale. En effet, comme on l'a signalé l'étude de la convergence normale est aisée et il est rare dans la pratique qu'il y ait convergence uniforme et pas convergence normale surtout pour des applications  $f_n$  à valeurs positives.

En fait il existe deux grandes catégories de cas où la convergence est uniforme mais pas normale. Découvrons les à travers des exemples.

- Séries à termes alternés  
Pour tout  $n \in \mathbf{N}$ ,

$$f_n : [-1, 0] \rightarrow \mathbf{R}; x \mapsto \frac{x^n}{\sqrt{n}\sqrt{n+1}}.$$

*Convergence simple :*

- Séries à termes à supports disjoints  
Soit  $k$  un réel positif ou nul. Pour tout  $n \in \mathbf{N}$  on considère,

$$f_n : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto \begin{cases} 0, & \text{si } x \notin [n\pi, (n+1)\pi], \\ \frac{|\sin x|}{(n+1)^k} & \text{sinon .} \end{cases}$$

*Convergence simple :*



### 1.3 Théorèmes de densité

dans ce paragraphe nous allons donner des théorèmes d'approximation uniforme, d'applications par des applications « plus simples ». Ces résultats sont d'une importance cruciale en analyse, puisqu'ils permettent d'étendre des résultats, faciles à prouver pour les applications approximantes, aux applications qu'elles approchent.

On désigne toujours par  $\mathbf{F}$  un espace vectoriel de dimension finie sur le corps  $\mathbf{K}$  des nombres réels ou complexes. Dans tout ce paragraphe on désignera par  $[a, b]$  un segment de  $\mathbf{R}$  non dégénéré ( $a < b$ ).

APPLICATIONS EN ESCALIER, APPLICATIONS CONTINUES PAR MORCEAUX.

Commençons par quelques définitions, déjà rencontrées en MPSI dans un cadre plus restrictif.

**Définition 1.3.1.** — On appelle subdivision du segment  $[a, b]$  toute suite finie  $(a_0, a_1, \dots, a_n)$  de réels où  $n$  est un entier non nul et

$$a = a_0 < a_1 < \dots < a_n = b.$$

**Définition 1.3.2.** — Une application  $\vec{\varphi}$  de  $[a, b]$  dans  $\mathbf{F}$  est dite en escalier si, par définition il existe une subdivision de  $[a, b]$ ,  $(a_0, a_1, \dots, a_n)$ , telle que pour  $i = 0, 1, \dots, n-1$ , la restriction de  $\vec{\varphi}$  à  $]a_i, a_{i+1}[$  soit constante.

Si  $\vec{\varphi}$  est une application de  $[a, b]$  dans  $\mathbf{F}$  en escalier, toute subdivision  $(a_0, a_1, \dots, a_n)$  de  $[a, b]$  telle que la restriction de  $\vec{\varphi}$  à  $]a_i, a_{i+1}[$  soit constante, pour  $i = 0, 1, \dots, n-1$ , est dite adaptée ou subordonnée à  $\vec{\varphi}$ .



FIG. 4

**Exemple** —

- L'application  $E$  de  $[a, b]$  dans  $\mathbf{R}$  qui à  $x$  associe sa partie entière est une fonction en escalier ;
- La restriction de la fonction caractéristique de  $\mathbf{N}$  à  $[a, b]$  est une fonction en escalier ;
- L'application de  $[0, 1]$  dans  $\mathbf{R}$  qui à un élément  $x$  de  $[0, 1]$  associe 1, si  $x$  est de la forme  $\frac{1}{n}$ , avec  $n \in \mathbf{N}$  et 0 sinon, n'est pas une application en escalier.

**Définition 1.3.3.** — Une application  $\vec{\varphi}$  de  $\mathbf{R}$  dans  $\mathbf{F}$  est dite en escalier si, par définition il existe un segment  $[c, d]$  telle que l'on ait :

1.  $\vec{\varphi}$  est nulle sur le complémentaire de  $[c, d]$ .
2. La restriction de  $\vec{\varphi}$  à  $[c, d]$  est en escalier.

**Remarque** — L'application partie entière n'est pas une application en escalier sur  $\mathbf{R}$ .

Nous laissons en exercice le résultat suivant dont la preuve est identique à celle donnée en MPSI dans le cas réel.

**Proposition 1.3.4.** — L'ensemble des applications en escalier de  $[a, b]$  (resp.  $\mathbf{R}$ ) dans  $\mathbf{F}$  est un sous-espace vectoriel de  $\mathcal{F}([a, b], \mathbf{F})$  (resp.  $\mathcal{F}(\mathbf{R}, \mathbf{F})$ ). On le note  $\mathcal{E}([a, b], \mathbf{F})$  (resp.  $\mathcal{E}(\mathbf{R}, \mathbf{F})$ ).

**Définition 1.3.5.** — Une application  $\vec{f}$  de  $[a, b]$  dans  $\mathbf{F}$  est dite continue par morceaux si, par définition, il existe une subdivision de  $[a, b]$ ,  $(a_0, a_1, \dots, a_n)$ , telle que pour  $i = 0, 1, \dots, n-1$ , la restriction de  $\vec{f}$  à  $]a_i, a_{i+1}[$  se prolonge à  $[a_i, a_{i+1}]$  en une application  $\vec{f}_i$  continue.

Si  $\vec{f}$  est une application de  $[a, b]$  dans  $\mathbf{F}$  continue par morceaux, toute subdivision  $(a_0, a_1, \dots, a_n)$  de  $[a, b]$  telle que la restriction de  $\vec{f}$  à  $]a_i, a_{i+1}[$  se prolonge à  $[a_i, a_{i+1}]$  en une application  $\vec{f}_i$  continue, est dite adaptée ou subordonnée à  $\vec{f}$ .

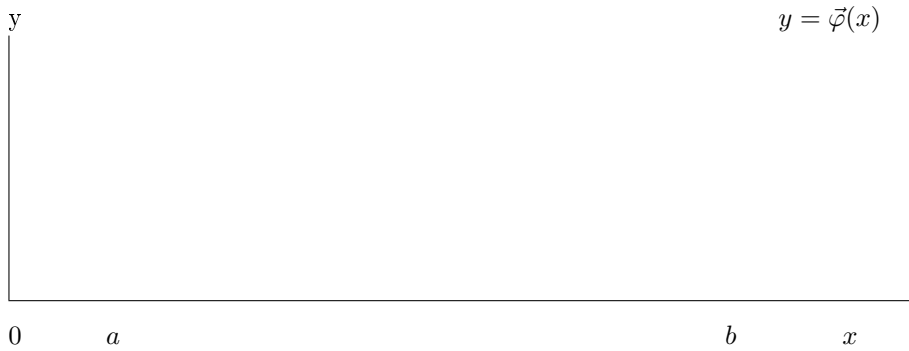


FIG. 5

**Exemples —**

- Toute application de  $[a, b]$  dans  $\mathbf{R}$  continue est continue par morceaux.
- Toute application de  $[a, b]$  dans  $\mathbf{R}$  en escalier est continue par morceaux.
- L'application  $[-1, 1] \rightarrow \mathbf{R}; x \mapsto \begin{cases} \frac{1}{x}, & \text{pour } x \neq 0; \\ 0, & \text{pour } x = 0; \end{cases}$  n'est pas continue par morceaux.

**Remarque —** Une application de  $[a, b]$  dans  $\mathbf{R}$  continue par morceaux a en particulier, une limite à droite en  $a$ , une limite à gauche en  $b$  et une limite à droite et à gauche en tout point de  $]a, b[$ .

**Définition 1.3.6. —** Une application  $\vec{f}$  de  $\mathbf{R}$  dans  $\mathbf{F}$  est dite continue par morceaux si, par définition, la restriction de  $\vec{f}$  à tout segment de  $\mathbf{R}$  est continue par morceaux.

**Exemple —** l'application de  $\mathbf{R}$  dans  $\mathbf{R}$ , partie entière est continue par morceaux.

**Proposition 1.3.7. —** L'ensemble des applications de  $[a, b]$  (resp.  $\mathbf{R}$ ) dans  $\mathbf{F}$  continue par morceaux est un sous-espace vectoriel de  $\mathcal{F}([a, b], \mathbf{F})$  (resp.  $\mathcal{F}(\mathbf{R}, \mathbf{F})$ ). On le note  $\mathcal{C}_m^0([a, b], \mathbf{F})$  (resp.  $\mathcal{C}_m^0(\mathbf{R}, \mathbf{F})$ ).

De plus toute application continue par morceaux sur un segment est bornée :  $\mathcal{C}_m^0([a, b], \mathbf{F})$  est inclus dans  $\mathcal{B}([a, b], \mathbf{F})$ .

*Preuve de la proposition 1.3.7. —* Le premier point est laissé en exercice... Passons au second. Soit  $\vec{f} \in \mathcal{C}_m^0([a, b], \mathbf{F})$ . Soit  $(a_0, a_1, \dots, a_n)$  une subdivision de  $[a, b]$  adaptée à  $\vec{f}$ . Pour  $i = 1, 2, \dots, n-1 \dots$

Il est temps d'énoncer le résultat principal.

**Proposition 1.3.8. —** Pour toute application  $\vec{f}$  de  $[a, b]$  dans  $\mathbf{F}$ , continue par morceaux, il existe une suite  $(\vec{\varphi}_n)$  d'éléments de  $\mathcal{E}([a, b], \mathbf{F})$  qui converge uniformément vers  $\vec{f}$ ; autrement dit  $\vec{f}$  est limite uniforme d'une suite d'applications en escalier.

Nous avons observé précédemment que  $\mathcal{E}([a, b], \mathbf{F}) \subset \mathcal{C}_m^0([a, b], \mathbf{F}) \subset \mathcal{B}([a, b], \mathbf{F})$ . Plaçons nous dans l'espace vectoriel normé  $(\mathcal{B}([a, b], \mathbf{F}), \|\cdot\|_\infty)$ . Compte tenue de La proposition 1.3.8. peut se reformuler de la façon suivante :

**Proposition 1.3.8. bis —** Dans l'espace vectoriel normé  $(\mathcal{B}([a, b], \mathbf{F}), \|\cdot\|_\infty)$ , ...

Passons à la preuve.

*Preuve de la proposition 1.3.8. —*

CAS PARTICULIER :  $\vec{f}$  est continue.





FIG. 6

CAS GÉNÉRAL : On ne suppose plus  $\vec{f}$  continue. La continuité par morceaux de  $\vec{f}$  assure l'existence d'une subdivision  $(b_0, b_1, \dots, b_n)$  de  $[a, b]$  adaptée à  $\vec{f}$ . Pour  $i = 1, 2, \dots, n-1$ , on note  $\vec{f}_i$  l'application continue sur  $[b_i, b_{i+1}]$  qui prolonge  $f|_{]b_i, b_{i+1}[}$ .

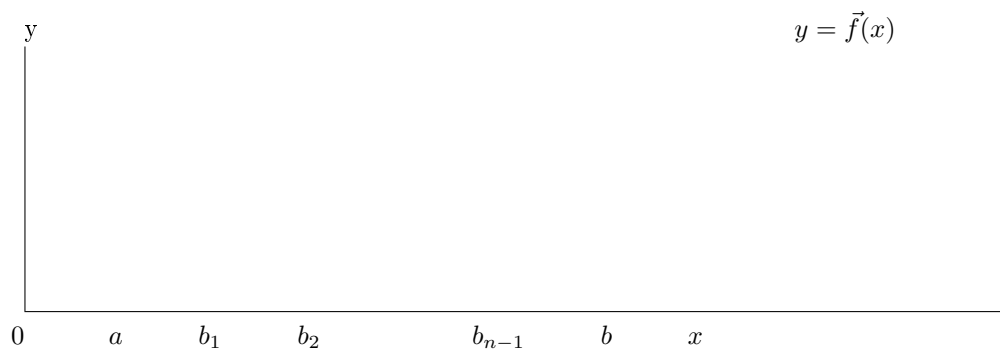


FIG. 7

Donnons un exemple où, pour prouver un résultat pour une application continue par morceaux, on commence par le prouver, ce qui est beaucoup plus simple, pour une application en escalier.

**Exercice 1.3.9.** — LEMME DE LEBESGUE, RIEMANN — Soit  $f$  un élément de  $\mathcal{C}_m^0([a, b], \mathbf{C})$ . Posons pour tout entier naturel  $n$ ,  $I_n(f) := \int_a^b f(t) \exp(int) dt$ . Montrer que  $I_n(f) \xrightarrow{n \rightarrow +\infty} 0$ .

*Solution de l'exercice 1.3.9.—*

#### APPLICATIONS POLYNOMIALES, THÉORÈME DE WIERSTRASS

**Proposition 1.3.10.** — THÉORÈME DE WIERSTRASS —

*pour toute application  $f$  de  $[a, b]$  dans  $\mathbf{K}$ , continue il existe une suite d'applications polynomiales (associées à des éléments de  $\mathbf{K}[X]$ ) qui converge sur  $[a, b]$  uniformément vers  $f$ . Autrement dit, l'espace vectoriel des applications polynomiales définies sur  $[a, b]$  est dense dans  $(\mathcal{C}^0([a, b], \mathbf{K}), \|\cdot\|_\infty)$ .*

La preuve de ce résultat n'est pas au programme, nous donnerons en TD. ou en exercice diverses preuves de la proposition 1.3.10.

**Remarque** — Dans le théorème de Weierstraß le fait de se placer sur un segment est essentiel. Par exemple, il existe sur  $]0, 1[$  des applications continues qui ne sont pas limite uniforme d'une suite d'applications polynomiales. Prenons l'application...

On a même le résultat suivant :

**Exercice** — Soit  $f$  une application de  $\mathbf{R}$  dans  $\mathbf{R}$  qui est limite uniforme, sur  $\mathbf{R}$  d'une suite d'applications polynomiales. Montrer que  $f$  est polynomiale.

*Solution* —

Nous allons conclure ce paragraphe en prouvant derechef le lemme de Riemann-Lebesgue (exercice 1.3.9), pour une application continue, par un raisonnement utilisant la densité des fonctions polynomiales dans  $(\mathcal{C}^0([a, b], \mathbf{C}), \|\cdot\|_\infty)$ .

**Exercice** — LEMME DE LEBESGUE, RIEMANN, LE RETOUR — Soit  $f$  un élément de  $\mathcal{C}^0([a, b], \mathbf{C})$ . Posons pour tout entier naturel  $n$ ,  $I_n := \int_a^b f(t) \exp(int) dt$ .

1. On suppose dans cette question que  $f$  est de classe  $\mathcal{C}^1$ . Montrer que  $I_n \xrightarrow{n \rightarrow +\infty} 0$ .
2. En déduire que dans le cas général ( $f$  continue)  $I_n \xrightarrow{n \rightarrow +\infty} 0$ .

*Solution de l'exercice* —

## DÉRIVATION

Nous allons brièvement généraliser la notion de dérivée et les résultats sur la dérivation vus en première année, à des applications à valeurs vectorielles.

Dans toute la partie II,  $\mathbf{F}$  désignera un  $\mathbf{K}$ -espace vectoriel de dimension finie non nulle  $n$ ;  $\mathbf{K}$  est toujours le corps des nombres réels ou celui des nombres complexes.  $I$  désignera un intervalle non réduit à un point.

### 2.1 Dérivées, fonctions de classe $\mathcal{C}^1$

**Définition 2.1.1.** — Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ , soit  $a$  un point de  $I$ . On dit que  $\vec{f}$  est dérivable en  $a$ , si, par définition, il existe un élément  $\vec{\ell}$  de  $\mathbf{F}$  tel que

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{1}{x - a} \cdot (\vec{f}(x) - \vec{f}(a)) = \vec{\ell}.$$

Si c'est le cas,  $\vec{\ell}$  s'appelle dérivée de  $\vec{f}$  au point  $a$ ; on la note  $\vec{f}'(a)$ ,  $D\vec{f}(a)$  ou encore  $\frac{d\vec{f}}{dx}(a)$ .

**Remarque** — L'unicité de la limite assure l'unicité, quand elle existe, de la dérivée en  $a$ , ce qui justifie bien les notations précédentes.

**Proposition 2.1.2.** — On a l'équivalence des deux propositions suivantes :

- i)  $\vec{f}$  est dérivable au point  $a$  et sa dérivée en  $a$  vaut  $\vec{\ell}$ ,
- ii)  $\vec{f}(x) = \vec{f}(a) + (x - a) \cdot \vec{\ell} + \vec{o}_{x \rightarrow a}(x - a)$ .

Ce résultat se prouve comme pour une application à valeurs réelles (cf. cours de première année), et on dira encore, quand l'égalité ii) est satisfaite, que  $\vec{f}$  admet un développement limité à l'ordre 1 au voisinage de  $a$ .

**Remarque** — ii) permet souvent dans la pratique de prouver qu'une application est dérivable (cf. proposition 2.1.12), il donne aussi le corollaire suivant :

**Corollaire 2.1.3.** — Si  $\vec{f}$  est dérivable en un point  $a$  de  $I$ , alors  $\vec{f}$  est continue en  $a$ .

*Preuve du théorème 2.1.3.* — Si  $\vec{f}$  est dérivable en  $a$ , alors,

$$\vec{f}(x) = \vec{f}(a) + (x - a) \cdot \vec{f}'(a) + \vec{o}_{x \rightarrow a}(x - a),$$

et donc

$$\vec{f}(x) = \vec{f}(a) + \underset{x \rightarrow a}{o}(1).$$

D'où la continuité de  $\vec{f}$  en  $a$ .

Comme pour les applications à valeurs réelles il est possible de définir la notion de dérivée droite ou à gauche pour une application vectorielle.

**Définition 2.1.4.** — Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ , soit  $a$  un point de  $I$ . On dit que  $\vec{f}$  est dérivable à droite (resp. à gauche) en  $a$ , si,  $]a, +\infty[ \cap I$  (resp.  $] - \infty, a[ \cap I$ ) par définition, est non vide et si il existe un élément  $\vec{\ell}$  de  $\mathbf{F}$  tel que

$$\lim_{\substack{x \rightarrow a \\ x \geq a \text{ (resp. } \leq a)}} \frac{1}{x - a} \cdot (\vec{f}(x) - \vec{f}(a)) = \vec{\ell}.$$

Si c'est le cas,  $\vec{\ell}$  s'appelle dérivée de  $\vec{f}$  à droite (resp. à gauche) au point  $a$ , on la note  $\vec{f}'_d(a)$  (resp.  $\vec{f}'_g(a)$ ), .

Rappelons dans le cas réel l'interprétation géométrique de la dérivée.

**Interprétation géométrique dans le cas réel** —

Soit  $g$  une application de  $I$  dans  $\mathbf{R}$ , supposée dérivable en un point  $a$  de  $I$ . Pour tout point  $x$  de  $I$  distinct de  $a$ , on note  $D_x$  la droite de  $\mathbf{R}^2$ , passant par les points  $(a, g(a))$  et  $(x, g(x))$ . Pour tout point  $x$  de  $I$  la pente de  $D_x$  vaut  $\frac{g(x) - g(a)}{x - a}$ ; cette quantité tend vers  $g'(a)$  lorsque  $x$  tend vers  $a$ . La droite passant par  $(a, g(a))$  de pente  $g'(a)$  est, en ce sens, la « limite » de  $D_x$ , lorsque  $x$  tend vers  $a$  (voir figure 4). On l'appelle *tangente* au graphe de  $g$  au point  $(a, g(a))$ . L'équation de  $T$  au point  $(a, g(a))$  est donc :

$$T : y = g(a) + (x - a)g'(a),$$

or d'après 2.1.2.

$$g(x) = g(a) + (x - a)g'(a) + \underset{x \rightarrow a}{o}(x - a),$$

donc la distance  $d_x$ , du point du graphe de  $g$ ,  $(x, g(x))$  à  $T$ , est négligeable devant  $(x - a)$  lorsque  $x$  tend vers  $a$  (voir figure 5),

$$d_x = \underset{x \rightarrow a}{o}(x - a).$$

Cette propriété caractérise la tangente parmi toutes les droites passant par  $(a, g(a))$ .

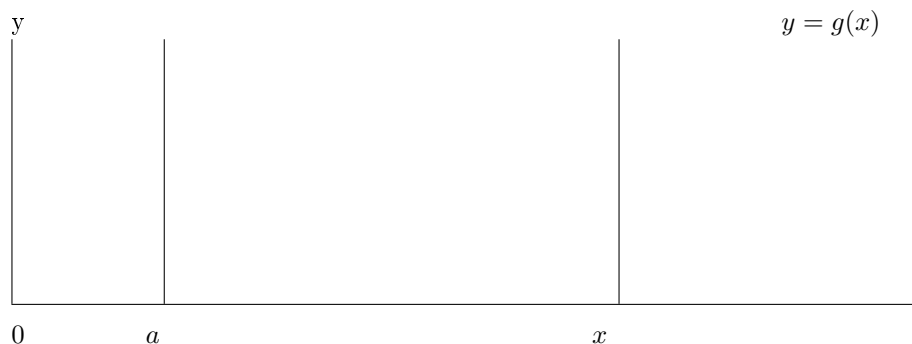


FIG. 9

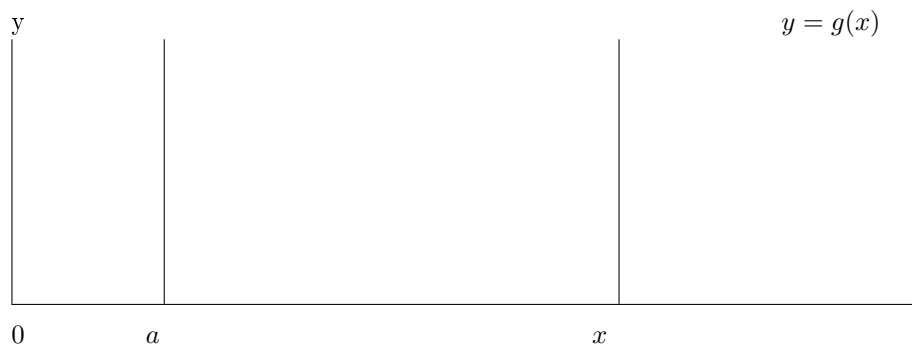


FIG. 10

Nous allons dans le cas d'une application vectorielle donner une interprétation géométrique de la dérivée. Pour ce faire introduisons une notion importante que nous rencontrerons plusieurs fois cette année, celle d'arc paramétré.

## ARC PARAMÉTRÉ —

Dans ce paragraphe  $\mathbf{F}$  est muni de sa structure affine canonique, c'est en tout cas la situation la plus fréquente lorsque l'on étudie des arcs paramétrés, concrètement on considèrera les éléments de  $\mathbf{F}$  comme de points et on les notera comme tels, leur différences comme des vecteurs. On notera ici les points de  $\mathbf{F}$ , et les applications à valeurs dans  $\mathbf{F}$ , par des majuscules maigres romaines, on notera par des majuscules maigres romaines surmontées d'une flèche, les vecteurs de  $\mathbf{F}$  ou les applications à valeurs dans cet espace vectoriel.

Un arc paramétré n'est rien d'autre qu'une application  $F$  de  $I$  dans  $\mathbf{F}$ , de façon plus précise, c'est suivant la coutume le couple  $(I, F)$ . La variable qui décrit l'intervalle  $I$  est appelé paramètre, ce qui donne la terminologie arc paramétré. Dans la suite de ce paragraphe  $(I, F)$  désignera un arc paramétré.

**Définition 2.1.5.** — On appelle support de l'arc paramétré  $(I, F)$ , ou courbe associée à l'arc paramétré  $(I, F)$  l'ensemble  $F(I)$ .

Le point de paramètre  $t$ , où  $t$  est un élément de  $I$  est le couple  $(t, F(t))$ .

Le point de paramètre  $t$  porte à la fois la donnée d'un point géométrique  $M = F(t)$  et la valeur du paramètre  $t$  pour laquelle  $F$  prend la valeur  $M$ , Le point  $M$  pouvant être l'image de plusieurs éléments de  $I$ .

**Notation** — Lorsque  $\mathbf{F}$  est muni d'un repère et que  $f_i$  est la  $i^{\text{e}}$  fonction composante dans ce repère on note encore l'arc  $(I, F)$  de la façon suivante :

$$\begin{cases} x_1 = f_1(t), \\ x_2 = f_2(t), \\ \dots \\ x_d = f_d(t), \end{cases} \quad t \in I,$$

où  $(x_1, \dots, x_d)$  est l'écriture des coordonnées d'un point générique de  $\mathbf{F}$ . Ainsi si  $\mathbf{F} = \mathbf{R}^2$  muni de sa structure euclidienne canonique, considérons pour  $R \in \mathbf{R}_+^*$  l'arc géométrique :

$$\begin{cases} x = R \cos(t), \\ y = R \sin(t), \end{cases} \quad t \in [0, 2\pi].$$

Le support de cet arc est.....

Le point de paramètre  $\frac{\pi}{2}$  est .....

**Interprétation cinématique** —

Souvent  $t$  représente le « temps »,  $F(t)$  représente pour un point matériel donné sa *position* au temps  $t$ ,  $F$  s'appelle alors la *loi horaire* et le support de l'arc  $(I, F)$  représente la *trajectoire* du point matériel.

Supposons dans la suite que  $F$  soit dérivable en un point  $a$  de  $I$ , On notera  $\vec{F}'(a)$  la dérivée de  $F$  au point  $a$ ,  $\vec{F}'(a)$  est la limite de  $\frac{1}{t-a} \cdot (F(t) - F(a))$ , lorsque  $t$  tend vers  $a$  par valeurs distinctes, c'est donc un vecteur. Notons que c'est aussi pour tout point  $O$  de  $\mathbf{F}$  la dérivée au point  $a$  de l'application

$$\vec{F}_O : I \rightarrow \mathbf{F}; t \mapsto \overrightarrow{OF(t)}.$$

en effet..

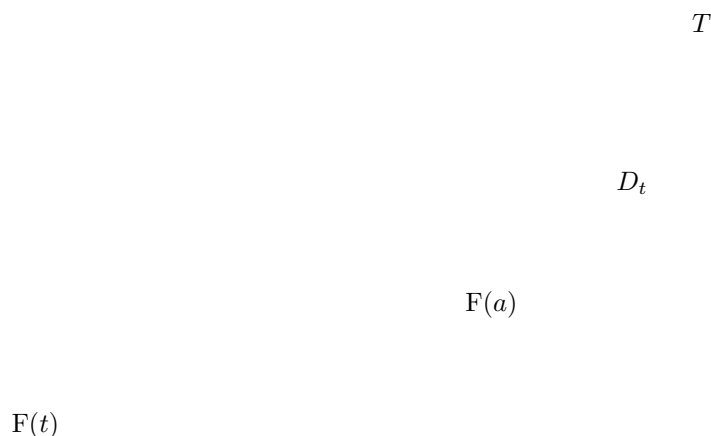


FIG. 11

On suppose que  $\vec{F}'(a)$  existe et est non nulle. On dit que le point de paramètre  $t$  est régulier. soit  $T$  la droite passant par le point  $F(a)$  et dirigée par  $\vec{F}'(a)$ . D'après 2.1.2

$$F(t) = F(a) + (t - a)\vec{F}'(a) + o_{t \rightarrow a}(t - a), \quad (\text{VI.2})$$

Donc puisque  $\vec{F}'(a) \neq \vec{0}$ , il existe un voisinage  $V$  de  $a$ , tel que pour tout élément  $t$  de  $V \cap I$ , distinct de  $a$ ,  $F(t) \neq F(a)$ , pour un tel élément  $t$ , on note  $D_t$  la droite passant par les deux points  $F(t)$  et  $F(a)$ .  $D_t$  est dirigée par le vecteur  $\frac{1}{t-a} \cdot \overrightarrow{F(a)F(t)}$ , vecteur qui tend vers  $\vec{F}'(a)$ , lorsque  $t$  tend vers  $a$ , par valeurs distinctes. En ce sens,  $T$  est « limite » de  $D_t$ , lorsque  $t$  tend vers  $a$  par valeurs distinctes.  $T$  s'appelle la *tangente* à l'arc  $(I, F)$  au point de paramètre  $a$ .

Notons que l'on retrouve la définition de la tangente du point précédent pour une application  $g$  à valeurs réelles, comme cas particulier : Pour  $\mathbf{F} = \mathbf{R}^2$  et pour  $F : t \mapsto (t, g(t))$ , on montre facilement que la tangente à l'arc  $(I, F)$ , au point de paramètre  $a$ , est la tangente au graphe de  $g$  au point  $(a, g(a))$ .

Notons pour tout  $t \in I$ ,  $M(t)$  le point de  $T$ ,  $F(a) + (t - a)\vec{F}'(a)$ , d'après (VI.2),  $\|F(t) - M(t)\| = o(t - a)$  lorsque  $t$  tend vers  $a$ , la droite  $T$  est donc en ce sens très proche du support de l'arc paramétré, d'où son nom<sup>1</sup>. On peut montrer que cette propriété caractérise la tangente parmi les droites passant par  $F(a)$ .

**Interprétation cinématique** — En cinématique  $\vec{F}'(a)$  s'appelle *vecteur vitesse instantanée au temps  $a$* . Si de plus  $\mathbf{F}$  est muni d'une structure euclidienne, sa norme,  $\|\vec{F}'(a)\|$ , s'appelle *vitesse instantanée au temps  $a$* . En anticipant sur la suite si  $F$  est deux fois dérivable en  $a$ , alors  $\vec{F}''(a)$  s'appelle l'accélération au temps  $a$ .

Définissons maintenant les applications dérivables et de classe  $\mathcal{C}^1$ .

**Définition 2.1.6.** — Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ .  $\vec{f}$  est dite *dérivable*, si, par définition, elle est dérivable en tout point de  $I$ . Si  $\vec{f}$  est dérivable, l'application

$$\vec{f}' : I \rightarrow \mathbf{F}; x \mapsto \vec{f}'(x)$$

s'appelle l'application *dérivée*, ou plus simplement, la *dérivée* de  $\vec{f}$ . On la note encore  $D\vec{f}$  ou  $\frac{d\vec{f}}{dx}$ .

**Définition 2.1.7.** — Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ .  $\vec{f}$  est dite *continûment dérivable*, ou encore (de classe)  $\mathcal{C}^1$ , si par définition, elle est dérivable et  $\vec{f}'$  est continue.

1. En latin *tango-is-ere* signifie toucher.

**Définition 2.1.8.** — Soient  $\vec{f}$  et  $\vec{g}$  des applications de  $I$  dans  $\mathbf{F}$  toutes deux dérivables en un point  $a$  de  $I$ . Soient  $\lambda$  et  $\mu$  des éléments de  $\mathbf{K}$ . Alors  $\lambda\vec{f} + \mu\vec{g}$  est dérivable en  $a$  et de plus

$$D(\lambda\vec{f} + \mu\vec{g})(a) = \lambda D\vec{f}(a) + \mu D\vec{g}(a).$$

La preuve est analogue à celle donnée en première année pour les applications à valeurs réelles. Il admet le corollaire immédiat suivant.

**Corollaire 2.1.9.** — L'ensemble des applications de  $I$  dans  $\mathbf{F}$ , continûment dérivables est un sous-espace vectoriel de l'espace vectoriel  $\mathcal{F}(I, \mathbf{F})$ , des applications de  $I$  dans  $\mathbf{F}$ . On le note  $\mathcal{C}^1(I, \mathbf{F})$  (ou  $\mathcal{C}^1(I)$  si  $\mathbf{F} = \mathbf{R}$ ). De plus l'application

$$D : \mathcal{C}^1(I, \mathbf{F}) \rightarrow \mathcal{F}(I, \mathbf{F}); \vec{f} \mapsto D\vec{f}$$

est linéaire à valeurs dans  $\mathcal{C}^0(I, \mathbf{F})$ .

**Proposition 2.1.10.** — Soit  $(\mathbf{G}, \|\cdot\|)$  un espace vectoriel normé. soit  $\vec{\ell}$  une application linéaire de  $\mathbf{F}$  dans  $\mathbf{G}$ ; soit enfin  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ , dérivable en un point  $a$  de  $I$ . Alors,  $\vec{\ell} \circ \vec{f}$  est dérivable en  $a^1$  et de plus,

$$D(\vec{\ell} \circ \vec{f})(a) = \vec{\ell}(D\vec{f}(a)).$$

*Preuve de la proposition 2.1.10.* — Pour tout élément  $x$  de  $I$  distinct de  $a$ ,  $\vec{\ell}$  étant linéaire,

$$\frac{1}{x-a} \cdot (\vec{\ell} \circ \vec{f}(x) - \vec{\ell} \circ \vec{f}(a)) = \vec{\ell} \left( \frac{1}{x-a} \cdot (\vec{f}(x) - \vec{f}(a)) \right).$$

Or d'une part  $\frac{1}{x-a} \cdot (\vec{f}(x) - \vec{f}(a)) \xrightarrow[x \neq a]{x \rightarrow a} D\vec{f}(a)$ , et d'autre part  $\vec{\ell}$  est continue, puisque  $\mathbf{F}$  est de dimension finie, donc,

$$\frac{1}{x-a} \cdot (\vec{\ell} \circ \vec{f}(x) - \vec{\ell} \circ \vec{f}(a)) \xrightarrow[x \neq a]{x \rightarrow a} \vec{\ell}(D\vec{f}(a)).$$

D'où le résultat. □

La proposition 2.1.10., admet une succession de corollaires importants

**Corollaire 2.1.11.** — Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base de  $\mathbf{F}$  notée  $\mathcal{B}$ .  $\vec{f}$  désigne toujours une application de  $I$  dans  $\mathbf{F}$  et l'on note, pour  $i = 1, 2, \dots, n$  la  $i^{\text{e}}$  application composante de  $\vec{f}$  dans la base  $\mathcal{B}$ .  $\vec{f}$  est dérivable en  $a$ , élément de  $I$ , si et seulement si, pour tout élément  $i$  de  $\{1, 2, \dots, n\}$ ,  $f_i$  est dérivable en  $a$ .

De plus, si c'est le cas,  $\vec{f}'(a) = \sum_{i=1}^n f'_i(a) \vec{e}_i$ .

*Preuve du corollaire 2.1.11.* —

— Supposons que pour tout élément  $i$  de  $\{1, 2, \dots, n\}$ ,  $f'_i$  soit dérivable en  $a$ . Pour tout élément  $x$  de  $I$ , distinct de  $a$ ,

$$\frac{1}{x-a} \cdot (\vec{f}(x) - \vec{f}(a)) = \sum_{i=1}^n \frac{1}{x-a} (f_i(x) - f_i(a)) \vec{e}_i.$$

Comme pour  $i = 1, 2, \dots, n$   $\frac{1}{x-a} (f_i(x) - f_i(a)) \xrightarrow[x \neq a]{x \rightarrow a} f'_i(a)$ , les résultats sur les limites nous apprennent que

$$\frac{1}{x-a} \cdot (\vec{f}(x) - \vec{f}(a)) \xrightarrow[x \neq a]{x \rightarrow a} \sum_{i=1}^n f'_i(a) \vec{e}_i.$$

Donc  $\vec{f}$  est dérivable en  $a$  et  $\vec{f}'(a) = \sum_{i=1}^n f'_i(a) \vec{e}_i$ . D'où le résultat.

1.  $\vec{\ell} \circ \vec{f}$  est à valeurs dans l'espace vectoriel de dimension finie  $\vec{\ell}(\mathbf{F})$  ( $\mathbf{F}$  est de dimension finie), donc l'étude de sa dérivabilité ne sort pas du cadre de ce cours.



- Réciproquement, supposons  $\vec{f}$  dérivable en  $a$ . Notons pour  $i = 1, 2, \dots, n$ ,  $e_i^*$  la  $i^{\text{e}}$  forme coordonnée dans la base  $\mathcal{B}$ ; remarquons que  $f_i = e_i^* \circ \vec{f}$ . Donc d'après 2.1.8.,  $f_i$  est dérivable en  $a$  et  $f'_i(a) = e_i^*(f'(a))$ , pour  $i = 1, 2, \dots, n$ . Or

$$\vec{f}'(a) = \sum_{i=1}^n e_i^*(f'(a)) \vec{e}_i,$$

donc

$$\vec{f}'(a) = \sum_{i=1}^n f'_i(a) \vec{e}_i.$$

D'où le résultat. □

**Corollaire (du corollaire) 2.1.12.** — Soit  $f$  une application de  $I$  dans  $\mathbf{C}$ .  $f$  est dérivable en un point  $a$  de  $I$  si et seulement si,  $\operatorname{Re}(f)$  et  $\operatorname{Im}(f)$  sont dérivables en  $a$ . Si tel est le cas, alors,  $f'(a) = (\operatorname{Re}(f))'(a) + i(\operatorname{Im}(f))'(a)$ .

*Preuve du corollaire 2.1.12.* — On considère  $\mathbf{C}$  comme un  $\mathbf{R}$ -espace vectoriel de dimension 2 et on applique le corollaire précédent avec  $(1, i)$  comme base, les applications composantes de  $f$  dans cette base sont alors  $f_1 = \operatorname{Re}(f)$  et  $f_2 = \operatorname{Im}(f)$ . D'où le résultat. □

**Corollaire 2.1.13.** — Soit  $f$  une application de  $I$  dans  $\mathbf{C}$ .  $f$  est dérivable en un point  $a$  de  $I$  si et seulement si,  $\bar{f}$  est dérivable en  $a$ , et si tel est le cas, alors

$$D\bar{f}(a) = \overline{Df(a)}.$$

*Preuve du corollaire 2.1.13.* — la conjugaison est une application linéaire de  $\mathbf{C}$  dans  $\mathbf{C}$ , vu comme un  $\mathbf{R}$  espace vectoriel de dimension 2. La dérivabilité de  $f$  entraîne donc celle de  $\bar{f}$  d'après la proposition 2.1.10. La réciproque résulte de ce que  $\bar{\bar{f}} = f$ . La proposition 2.1.8. assure aussi la formule  $D\bar{f}(a) = \overline{Df(a)}$ . On peut aussi pour prouver ce résultat utiliser le corollaire 2.1.10. et la proposition 2.1.6, puisque  $\bar{f} = \operatorname{Re}(f) - i\operatorname{Im}(f)$ .

Dans la pratique on vient de voir que le calcul d'une dérivée d'une application à valeurs dans  $\mathbf{F}$  se fait en calculant la dérivée de chaque composantes.

**Exercice** — Soit l'arc paramétré de  $\mathbf{R}^2$ , muni de sa structure euclidienne canonique,

$$\begin{cases} x = 2t^3, \\ y = 3t^2, \end{cases} \quad t \in \mathbf{R}.$$

- Déterminer l'ensemble  $D$  des réels  $t$  tels que le point de paramètre  $t$  de l'arc soit régulier. Pour un élément  $t_0$  de  $D$  donner une équation cartésienne de la tangente  $T$  et de la normale au point de paramètre  $t_0$  (c'est-à-dire de la droite passant par  $(2t_0^3, 3t_0^2)$  et orthogonale à  $T$ ).
- Montrer que pour tout  $t \in D$  il existe un et un seul élément  $t'$  de  $D$  tel que les tangentes à l'arc aux points de paramètres  $t$  et  $t'$  soient orthogonale.
- Montrer qu'il existe deux et seulement deux éléments de  $D$ ,  $t_1$  et  $t_2$  tels que les tangentes aux points de paramètres  $t_1$  et  $t_2$  soient aussi des normales à la courbe.

*Solution* — On note  $F$  l'application  $\mathbf{R} \rightarrow \mathbf{R}^2$ ;  $t \mapsto (2t^3, 3t^2)$ .

D'abord une belle figure.

**Proposition 2.1.14.** — Soient  $\mathbf{F}_1$  et  $\mathbf{F}_2$  des espaces vectoriels de dimensions finies, soit  $(\mathbf{G}, \|\cdot\|)$  un espace vectoriel normé, soit  $\vec{B}$  une application bilinéaire de  $\mathbf{F}_1 \times \mathbf{F}_2$  dans  $\mathbf{G}$ , soient  $\vec{f}_1$  et  $\vec{f}_2$  des applications définies sur  $I$ , à valeurs respectivement dans  $\mathbf{F}_1$  et  $\mathbf{F}_2$ , soit enfin l'application

$$\vec{b} : I \rightarrow \mathbf{G}; t \mapsto \vec{B}(\vec{f}_1(t), \vec{f}_2(t)).$$

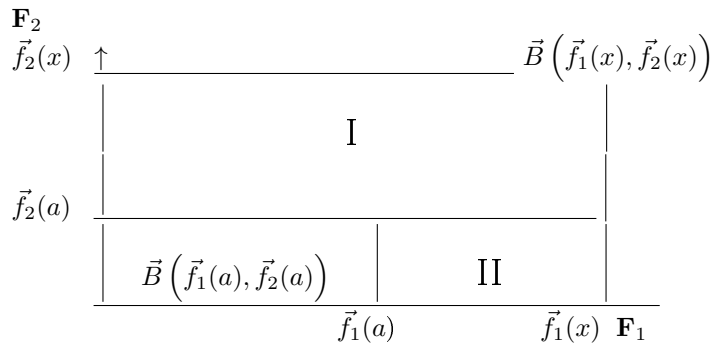
Si  $\vec{f}_1$  et  $\vec{f}_2$  sont dérivables en un point  $a$  de  $I$ , alors,  $\vec{b}$  l'est aussi et

$$\vec{b}'(a) = \vec{B}(\vec{f}_1'(a), \vec{f}_2(a)) + \vec{B}(\vec{f}_1(a), \vec{f}_2'(a)).$$

*Preuve de la proposition 2.1.14.* — Pour  $x$  élément de  $I$  évaluons  $\vec{b}(x) - \vec{b}(a)$  en s'inspirant de la figure ci dessous.

---

1. Notons que  $\vec{b}$  prend en fait ses valeurs dans l'espace vectoriel de *dimension finie*  $\vec{B}(\mathbf{F}_1 \times \mathbf{F}_2)$ , l'étude de sa dérivabilité ne sort donc pas du cadre de ce cours.



$$\vec{b}(x) - \vec{b}(a) = \underbrace{\vec{B}\left(\vec{f}_1(x), \vec{f}_2(x) - \vec{f}_2(a)\right)}_{\text{I}} + \underbrace{\vec{B}\left(\vec{f}_1(x) - \vec{f}_1(a), \vec{f}_2(a)\right)}_{\text{II}}$$

Afin d'étudier séparément chacun des deux termes du second membre, remarquons que, puisque pour  $i = 1, 2$ ,  $\vec{f}_i$  est dérivable en  $a$ , d'après la proposition 2.1.2,

$$\vec{f}_i(x) = \vec{f}_i(a) + (x - a) \cdot \vec{f}'_i(a) + (x - a) \cdot \vec{\varepsilon}_i(x), \quad (\text{VI.3})$$

où  $\vec{\varepsilon}_i$  est une application de  $I$  dans  $\mathbf{F}_i$ , telle que  $\lim_{x \rightarrow a} \vec{\varepsilon}_i(x) = \vec{0}_{\mathbf{F}_i}$ .

— ETUDE DU TERME II.

Donc finalement,

$$\vec{B}\left(\vec{f}_1(x) - \vec{f}_1(a), \vec{f}_2(a)\right) = (x - a) \cdot \vec{B}\left(\vec{f}'_1(a), \vec{f}_2(a)\right) + \vec{o}_{x \rightarrow a}(x - a). \quad (\text{VI.4})$$

$$\vec{B}(\vec{f}_1(x), \vec{f}_2(x) - \vec{f}_2(a)) = (x - a) \cdot \vec{B}(\vec{f}_1(a), \vec{f}_2'(a)) + \vec{o}_{x \rightarrow a}(x - a). \quad (\text{VI.5})$$

Il résulte de (VI.4-VI.5), que

$$\vec{b}(x) - \vec{b}(a) = (x - a) \cdot \left( \vec{B}(\vec{f}_1'(a), \vec{f}_2(a)) + \vec{B}(\vec{f}_1(a), \vec{f}_2'(a)) \right) + \vec{o}_{x \rightarrow a}(x - a),$$

ce qui d'après 2.1.2. assure que  $\vec{b}$  est dérivable en  $a$ , de dérivée en  $a$ ,  $\vec{B}(\vec{f}_1'(a), \vec{f}_2(a)) + \vec{B}(\vec{f}_1(a), \vec{f}_2'(a))$ .  $\square$

**Remarque** — Si en plus  $\vec{f}_1$  et  $\vec{f}_2$  sont de classe  $\mathcal{C}^1$ , alors, la continuité de  $\vec{B}$  de  $\vec{f}_1$ , de  $\vec{f}_1'$ , de  $\vec{f}_2$  et de  $\vec{f}_2'$ , assure que  $\vec{b}$  est de classe  $\mathcal{C}^1$ .

### Applications, illustrations—

— CAS SIMPLE

En prenant pour  $\mathbf{F}_1$ ,  $\mathbf{F}_2$ , et  $\mathbf{G}$  l'ensemble  $\mathbf{R}$  et pour  $\vec{B}$ , le produit de deux réels, on retrouve le résultat de terminale :

$$(f_1 f_2)'(a) = f_1'(a) f_2(a) + f_1(a) f_2'(a).$$

— PRODUIT SCALAIRE, MOUVEMENT CIRCULAIRE

Soient  $(\mathbf{E}, (\cdot|\cdot))$  un espace euclidien et  $\vec{f}_1$  et  $\vec{f}_2$  des applications de  $I$  dans  $\mathbf{E}$ , dérivables. Alors d'après 2.1.12. l'application

$$b : I \rightarrow \mathbf{R}; t \mapsto (\vec{f}_1(t) | \vec{f}_2(t))$$

est dérivable et pour tout élément  $t$  de  $I$ ,

$$b'(t) =$$

Soit  $\|\cdot\|$  la norme associée au produit scalaire  $(\cdot|\cdot)$ , on déduit en particulier de ce qui précède que l'application  $n : I \rightarrow \mathbf{R}; t \mapsto \|\vec{f}_1(t)\|^2$ , est dérivable de dérivée

$$n' : I \rightarrow \mathbf{R}; t \mapsto$$

Supposons que de plus  $n$  soit constante, de valeur  $R$ , réel strictement positif. On a prouvé qu'alors, pour tout élément  $t$  de  $I$ ,  $\vec{f}_1(t)$  est orthogonal à  $\vec{f}_1'$ .

Particularisons ce dernier résultat. Soit  $F$  une application de  $I$  dans  $\mathbf{E}_2$ , plan euclidien muni de sa structure affine canonique, dérivable. Soient  $O$  un point de  $\mathcal{E}_2$  et  $\mathbf{R}$  un élément de  $\mathbf{R}_+^*$ . On suppose que pour tout élément  $t$  de  $I$ ,  $\|\overrightarrow{OF(t)}\| = R$ , autrement dit,  $t \mapsto F(t)$  représente la loi horaire d'un mouvement circulaire de centre  $O$ . On a pour tout élément  $t$  de  $I$ ,  $\vec{F}'(t) \perp \overrightarrow{OF(t)}$ .

*Dans un mouvement circulaire, le vecteur vitesse est, à chaque instant, orthogonal au rayon.*

## — PRODUIT VECTORIEL, MOUVEMENT À ACCÉLÉRATION CENTRALE

On désigne par  $\mathbf{E}_3$  un espace euclidien de dimension 3 orienté, muni de sa structure affine canonique. Soit  $F$  une application de  $I$  dans  $\mathbf{E}_3$  dérivable, et telle que  $\vec{F}'$  soit elle-même dérivable (on note  $(\vec{F}')', \vec{F}''$ ). On suppose qu'il existe un point  $O$  de  $\mathbf{E}_3$  tel que pour tout élément  $t$  de  $I$ ,  $\overrightarrow{OF(t)}$  soit colinéaire à  $\vec{F}''$ , autrement dit,  $t \mapsto F(t)$  est la loi horaire d'un mouvement à *accélération centrale, de centre*  $O$ .

Posons  $\vec{\sigma} : I \rightarrow \mathbf{E}_3$ ;  $t \mapsto \overrightarrow{OF(t)} \wedge \vec{F}'(t)$  (moment cinétique). D'après la proposition précédente  $\vec{\sigma}$  est dérivable et pour tout élément  $t$  de  $I$ ,

$$\vec{\sigma}'(t) = \dots$$

On en déduit, comme on le verra dans la proposition 2.1.13., que  $\vec{\sigma}$  est constant ; on notera  $\vec{\Sigma}$  la valeur constante de  $\vec{\sigma}$ , et l'on supposera dans la suite que  $\vec{\Sigma}$  est *non nul*.

Pour tout élément  $t$  de  $I$ ,  $F(t)$  est dans le plan  $P$  passant par  $O$  est orthogonal à  $\vec{\Sigma}$ . En effet, pour tout  $t$ , élément de  $I$ ,

Conclusion : *un mouvement à accélération centrale est plan.* (traiter en exercice le cas où  $\vec{\Sigma}$  est nul.)

Terminons par un résultat très important, déjà rencontré pour les applications à valeurs réelles.

**Proposition 2.1.13.** — *Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ , dérivable.  $\vec{f}$  est constante si et seulement si  $\vec{f}'$  est nulle.*

*Preuve de la proposition 2.1.13*— Si  $\vec{f}$  est constante, alors, la définition de la dérivée donne immédiatement la nullité de  $\vec{f}'$ . Montrons la réciproque. Supposons que  $\vec{f}'$  soit nulle, montrons que  $\vec{f}$  est constante. Le résultat est acquis pour une application à valeurs réelles depuis l'an passé et va être utilisé en faisant apparaître les composantes de  $\vec{f}$  dans une base. soit  $\mathcal{B}$  une base de  $\mathbf{F}$ ; notons  $f_i$  la  $i^{\text{e}}$  composante de  $\vec{f}$  dans  $\mathcal{B}$ , pour  $i = 1, 2, \dots, n$ .

— ÉTAPE 1 :  $\mathbf{F}$  est un espace vectoriel sur le corps  $\mathbf{R}$ .

$f_i$  est dérivable de dérivée la  $i^{\text{e}}$  composante de  $\vec{f}'$  (cf. 2.1.9.), donc  $f'_i$  est nulle et donc, comme on l'a montré en MPSI, grâce à l'égalité des accroissements finis,  $f_i$  est constante sur l'intervalle  $I$ , ce pour  $i = 1, 2, \dots, n$ . Donc  $\vec{f}$  est constante.

— ÉTAPE 2 :  $\mathbf{F} = \mathbf{C}$ .

$\vec{f}$  est constante grâce à l'étape 1 ; en effet,  $\mathbf{C}$  est un  $\mathbf{R}$ -espace vectoriel de dimension 2.

— ÉTAPE 3 :  $\mathbf{F}$  est un espace vectoriel sur le corps  $\mathbf{C}$ .

On raisonne comme à l'étape 1, l'étape 2 assurant la constance des applications composantes de  $\vec{f}$  dans une base,  $f_1, f_2, \dots, f_n$ , applications à valeurs complexes. Donc  $\vec{f}$  est constante.

Dans tous les cas,  $\vec{f}$  est constante. □

L'intégrale vue, on pourra donner une preuve qui ne recourt pas l'usage des fonctions composantes.

## 2.2 Dérivées d'ordres supérieurs, fonctions de classe $\mathcal{C}^n$

Définissons comme en première année la notion de dérivée d'ordre supérieur ; nous allons procéder par récurrence.

**Définition 2.2.1.** — *Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ . définissons par récurrence, pour tout entier naturel  $n$ , la dérivée d'ordre  $n$  de  $\vec{f}$  en un point  $a$  de  $I$ , notée  $\vec{f}^{(n)}(a)$ .*

- On appelle dérivée d'ordre 0, de  $\vec{f}$  en un point  $a$  de  $I$ , notée  $\vec{f}^{(0)}(a)$ , la quantité  $\vec{f}(a)$ .
- Si  $\vec{f}$  est dérivable en un point  $a$  de  $I$ ,  $\vec{f}'(a)$  s'appelle la dérivée d'ordre 1 en  $a$  de  $\vec{f}$ ; on la note  $\vec{f}^{(1)}(a)$ .
- Si pour un point  $a$  de  $I$ , il existe un réel strictement positif  $\eta$ , tel que pour tout élément  $t$  de  $I \cap ]a - \eta, a + \eta[$ ,  $\vec{f}^{(1)}(t)$  existe et si l'application

$$\vec{\phi}_1 : I \cap ]a - \eta, a + \eta[ \rightarrow \mathbf{F}; t \mapsto \vec{f}^{(1)}(t)$$

est dérivable en  $a$  on dit que  $\vec{\phi}_1'(a)$  est la dérivée d'ordre 2 de  $\vec{f}$  en  $a$ ; on la note  $\vec{f}^{(2)}(a)$ <sup>1</sup>.

- Plus généralement, soit  $n$  un élément de  $\mathbf{N}^*$ . Supposons avoir défini la dérivée d'ordre  $n$ . Si pour un point  $a$  de  $I$ , il existe un réel strictement positif  $\eta$ , tel que pour tout élément  $t$  de  $I \cap ]a - \eta, a + \eta[$ ,  $\vec{f}^{(n)}(t)$  existe et si l'application

$$\vec{\phi}_n : I \cap ]a - \eta, a + \eta[ \rightarrow \mathbf{F}; t \mapsto \vec{f}^{(n)}(t)$$

est dérivable en  $a$  on dit que  $\vec{\phi}_n'(a)$  est la dérivée d'ordre  $n + 1$  de  $\vec{f}$  en  $a$ ; on la note  $\vec{f}^{(n+1)}(a)$ <sup>2</sup>.

**Définition 2.2.2.** — Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$ . On dit que  $f$  est (de classe)  $\mathcal{C}^n$ , où  $n$  est un entier naturel, si par définition, pour tout élément  $t$  de  $I$ ,  $\vec{f}^{(n)}(t)$  existe, et si l'application, appelée (application) dérivée d'ordre  $n$ ,

$$\vec{f}^{(n)} : I \rightarrow \mathbf{F}; t \mapsto \vec{f}^{(n)}(t)$$

est continue.

On dit que  $\vec{f}$  est de classe  $\mathcal{C}^\infty$  si, par définition,  $\vec{f}$  est de classe  $\mathcal{C}^k$ , pour tout entier naturel  $k$ .

Remarquons que pour  $n = 0$ , on retrouve la définition d'application de classe  $\mathcal{C}^0$ , que la définition 2.2.2 généralise donc.

**Remarque 2.2.3.** — Soit  $n$  un entier naturel. Sous réserve d'existence,

- on note pour  $a$  point de  $I$ ,  $\vec{f}^{(2)}(a)$ ,  $\vec{f}''(a)$  et  $\vec{f}^{(3)}(a)$ ,  $\vec{f}'''(a)$ ...
- La dérivée d'ordre  $n$  en un point  $a$  de  $I$ ,  $\vec{f}^{(n)}(a)$ , se note  $\frac{d^n \vec{f}}{dt^n}(a)$ , ou encore  $D^n \vec{f}(a)$ , l'application dérivée d'ordre  $n$ ,  $\vec{f}^{(n)}$ , se note elle,  $\frac{d^n \vec{f}}{dt^n}$ , ou encore  $D^n \vec{f}$ ,
- $\vec{f}^{(n+1)} = (\vec{f}^{(n)})' = (\vec{f}')^{(n)}$  (se prouve par récurrence),
- si  $f$  est de classe  $\mathcal{C}^n$ , alors  $f$  est de classe  $\mathcal{C}^k$  pour tout élément  $k$  de  $\{0, 1, \dots, n\}$ .

**Proposition 2.2.4.** — Soit  $n$  un élément de  $\mathbf{N} \cup \{+\infty\}$ . L'ensemble des applications de  $I$  dans  $\mathbf{F}$  de classe  $\mathcal{C}^n$ , est un sous-espace vectoriel de  $\mathcal{F}(I, \mathbf{F})$ , on le note  $\mathcal{C}^n(I, \mathbf{F})$  (ou  $\mathcal{C}^n(I)$  si  $\mathbf{F} = \mathbf{R}$ ). De plus, dans le cas où  $n$  est un naturel,

$$D^n : \mathcal{C}^n(I, \mathbf{F}) \rightarrow \mathcal{F}(I, \mathbf{F}); \vec{f} \mapsto D^n \vec{f}$$

est linéaire à valeurs dans  $\mathcal{C}^0(I, \mathbf{F})$ .

*Preuve de la proposition 2.2.4.* — Exercice! très facile... et très pénible!

Terminons par quelques résultats dont certains ont été vus en première année.

**Proposition 2.2.5.** — FORMULE DE LEIBNITZ —

Soit  $n$  un naturel et soient  $f$  et  $g$  des applications de  $I$  dans  $\mathbf{K}$  admettant des dérivées d'ordre  $n$  en un point  $a$  de  $I$ . Alors le produit  $f \times g$  admet une dérivée d'ordre  $n$  en  $a$  et :

$$(f \times g)^{(n)}(a) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(a) g^{(n-k)}(a).$$

**Corollaire 2.2.6.** — Soit  $n$  un élément de  $\mathbf{N} \cup \{+\infty\}$ . Le produit de deux applications de classe  $\mathcal{C}^n$  à valeurs réelles ou complexes et de classe  $\mathcal{C}^n$ .

L'ensemble des applications de  $I$  dans  $\mathbf{K}$  de classe  $\mathcal{C}^n$  est une sous-algèbre de  $\mathcal{F}(I, \mathbf{K})$ .

**Exercice 2.2.7.** — GÉNÉRALISATION DE LA FORMULE DE LEIBNITZ —

Soient  $\mathbf{F}_1$  et  $\mathbf{F}_2$  des espace vectoriels de dimensions finies, soit  $(\mathbf{G}, \|\cdot\|)$  un espace vectoriel normé, soit  $\vec{B}$  une application bilinéaire de  $\mathbf{F}_1 \times \mathbf{F}_2$  dans  $\mathbf{G}$ , soient  $\vec{f}_1$  et  $\vec{f}_2$  des applications définies sur  $I$ , à valeurs respectivement dans  $\mathbf{F}_1$  et  $\mathbf{F}_2$ , soit enfin l'application

$$\vec{b} : I \rightarrow \mathbf{G}; t \mapsto \vec{B}(\vec{f}_1(t), \vec{f}_2(t)).$$

1. La définition de la dérivée comme limite, assure que cette quantité ne dépend pas du choix de  $\eta$ .

2. La définition de la dérivée comme limite, assure que cette quantité ne dépend pas du choix de  $\eta$ .

Montrer que, si  $\vec{f}_1$  et  $\vec{f}_2$  admettent une dérivée d'ordre  $n$  en un point  $a$  de  $I$ , alors,  $\vec{b}$  admet une dérivée d'ordre  $n$  en  $a$ , et que

$$\vec{b}^n(a) = \sum_{k=0}^n \binom{n}{k} \vec{B} \left( \vec{f}^{(k)}(a), \vec{g}^{(n-k)}(a) \right).$$

*Indication* — Ce résultat, qui semble hors programme, se prouve, comme la formule de Leibnitz ordinaire, par récurrence, le cas  $n = 1$  a été prouvé en 2.1.12. Donnons en exercice un exemple d'utilisation.

**Exercice** — On munit ici  $\mathbf{R}^2$  de sa structure euclidienne canonique orientée. Soit  $\vec{u}$  et  $\vec{v}$  les applications

$$\vec{u} : \mathbf{R} \rightarrow \mathbf{R}^2; \theta \mapsto (\cos \theta, \sin \theta),$$

$$\vec{v} : \mathbf{R} \rightarrow \mathbf{R}^2; \theta \mapsto (-\sin \theta, \cos \theta).$$

1. Montrer que  $\vec{u}$  et  $\vec{v}$  sont de classe  $\mathcal{C}^\infty$  et préciser leurs dérivées.
2. Soit  $\rho$  une application de  $\mathbf{R}$  dans  $\mathbf{R}$  de classe  $\mathcal{C}^\infty$ , et

$$F : \mathbf{R} \rightarrow \mathbf{R}^2; \theta \mapsto \rho(\theta) \cdot \vec{u}(\theta).$$

Montrer que  $F$  est de classe  $\mathcal{C}^\infty$ , et donner  $F^{(4)}$ .

**Proposition 2.2.8.** — Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base de  $\mathbf{F}$  notée  $\mathcal{B}$ .  $\vec{f}$  désigne toujours une application de  $I$  dans  $\mathbf{F}$  et l'on note, pour  $i = 1, 2, \dots, n$ ,  $f_i$ , la  $i^e$  application composante de  $\vec{f}$  dans la base  $\mathcal{B}$ .

$\vec{f}$  admet en  $a$ , élément de  $I$ , une dérivée d'ordre  $n$ , si et seulement si, pour tout élément  $i$  de  $\{1, 2, \dots, n\}$ ,  $f_i$  admet en  $a$  une dérivée d'ordre  $n$ .

De plus, si c'est le cas,  $\vec{f}^{(n)}(a) = \sum_{i=1}^n f_i^{(n)}(a) \vec{e}_i$ .

*Preuve de la proposition 2.2.8.* — C'est une conséquence immédiate de 2.1.9 et de la définition d'une application de classe  $\mathcal{C}^n$ .

Soit  $\vec{f}$  une application de  $I$  dans  $F$  de classe  $\mathcal{C}^n$ , où  $n$  est entier naturel. On déduit de la proposition 2.2.7., en appliquant à chaque composante de  $\vec{f}$  dans une base  $\mathcal{B}$  la formule de Taylor-Young, la formule suivante.

**Proposition 2.2.9.** — FORMULE DE TAYLOR-YOUNG VECTORIELLE —

Soit  $\vec{f}$  une application de  $I$  dans  $F$  de classe  $\mathcal{C}^n$ , où  $n$  est entier naturel. et soit  $a$  un point de  $I$ . alors

$$\vec{f}(x) = \sum_{k=0}^n \frac{(x-a)^k}{k!} \cdot \vec{f}^{(k)}(a) + \vec{\sigma}_{x \rightarrow a} \left( (x-a)^n \right).$$

**Remarque** —

- On peut déduire cette formule de la formule de Taylor avec reste intégrale, que nous donnerons dans la partie suivante pour les fonctions à valeurs vectorielles, et éviter ainsi le recours peu élégant aux coordonnées.
- La formule est encore vraie sous l'hypothèse plus faible suivante : la dérivée  $n^e$  de  $f$  au point  $a$  existe, toutefois elle ne semble pas au programme sous cette forme, sauf évidemment pour  $n = 0$  ou  $n = 1$ , cf. 2.1.2.,
- Cette formule est une formule *locale*, elle s'utilise pour des calculs de limites.

**Théorème 2.2.10.** — Soit  $J$  un intervalle non réduit à un point. Soit  $\varphi$  une application de  $J$  dans  $I$  de classe  $\mathcal{C}^n$ . Alors  $\vec{f} \circ \varphi$  est de classe  $\mathcal{C}^n$ .

*Preuve du Proposition 2.2.11.* — Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base de  $\mathbf{F}$  notée  $\mathcal{B}$ . Pour  $i = 1, 2, \dots, n$ ,  $f_i$ ,  $i^e$  application composante de  $\vec{f}$  dans la base  $\mathcal{B}$ . D'après 2.2.8.  $f_i$  est de classe  $\mathcal{C}^n$ , et donc d'après le cours de sup.  $f_i \circ \varphi$  est de classe  $\mathcal{C}^n$ . Or  $f_i \circ \varphi$  est la  $i^e$  composante de  $\vec{f} \circ \varphi$  dans la base  $\mathcal{B}$ , la proposition 2.2.8. assure donc que  $\vec{f} \circ \varphi$  est de classe  $\mathcal{C}^n$ .  $\square$





# BRÈVE EXTENSION DE L'INTÉGRALE AUX FONCTIONS VECTORIELLES

Nous allons définir brièvement l'intégrale d'applications à valeurs dans un espace vectoriel de dimension finie.

Le programme suggère de définir l'intégrale, sur un segment, d'une application vectorielle, comme le vecteur dont la  $i^{\text{e}}$  composante dans une base est l'intégrale de la  $i^{\text{e}}$  application composante, ce qui permet, dans presque toutes les situations de se ramener à l'intégrale des fonctions à valeurs réelles ou complexes vue en sup. Cette présentation, outre qu'elle nécessite de montrer que l'intégrale ainsi construite est indépendante du choix de la base, ne livre aucune possibilité d'extension à des espaces de dimensions quelconque.

C'est pourquoi nous avons choisi de donner brièvement une construction intrinsèque de l'intégrale qui n'utilise pas le choix d'une base, et qui du reste pourrait s'appliquer à certains espaces de dimension infinie. Cette construction repose sur un des théorèmes de densité du 1.3., et c'est ce qui a motivé notre choix : nous pourrions en effet ainsi nous familiariser avec des raisonnements par densité. Notons enfin que la construction proposée n'est pas plus difficile que celle de l'intégrale d'une application réelle vue en sup. que réutiliserait lâchement et hypocritement le recours à une base.

Le lecteur peu sensible à ces considérations élevées pourra ignorer dans un premier temps le 3.2 et se rendre directement au 3.3.

Dans la première partie nous étendons de façon succincte aux *applications en escalier*, à valeurs vectorielles la notion d'intégrale vue en MPSI pour les applications numériques. Dans la seconde partie nous construirons l'intégrale à proprement parler des applications *continues par morceaux* à valeurs vectorielles en utilisant la densité des applications en escalier dans l'ensemble des applications continues par morceaux. Enfin, une dernière partie sera consacrée aux propriétés de l'intégrale ainsi construite (pour l'essentiel celles déjà rencontrées en sup., pour l'intégrale des applications numériques).

Dans toute la partie 3,  $\mathbf{F}$  désignera encore un  $\mathbf{K}$ -espace vectoriel de dimension finie non nulle  $d$ ;  $\mathbf{K}$  est toujours le corps des nombres réels ou celui des nombres complexes.  $J$  désignera un segment non réduit à un point.

## 3.1 Intégrales des applications en escalier

Nous définirons l'intégrale d'une fonction en escalier à valeurs dans  $F$ , comme cela a été fait en MPSI, pour une application en escalier à valeurs réelles. Plus précisément

**Proposition-définition 3.1.1.** — Soit  $\vec{\varphi}$  un élément de  $\mathcal{E}(J, \mathbf{F})$ . Soit  $(a_i)_{i=0, \dots, n}$  une subdivision de  $J$ , subordonnée à  $\vec{\varphi}$ . On notera  $\vec{c}_i$  la valeur constante de  $\vec{\varphi}$  sur  $]a_i, a_{i+1}[$ , pour  $i = 0, \dots, n-1$ .

La quantité  $\sum_{i=0}^{n-1} (a_{i+1} - a_i) \cdot \vec{c}_i$ , élément de  $\mathbf{F}$ , ne dépend que de  $\vec{\varphi}$ , et pas de la subdivision subordonnée à  $\varphi$  choisie.

On l'appelle *intégrale de  $\vec{\varphi}$  (sur  $J$ )*, on la note,  $\int_J \vec{\varphi}$ .

*Preuve de la proposition 3.1.1.* — C'est la même que celle de sup. pour des applications à valeurs réelles.

**Remarque** — Dans le cas d'une application  $\varphi$  à valeurs dans  $\mathbf{R}_+$ ,  $\int_J \varphi$  s'interprète comme l'aire comprise entre l'axe des abscisses et le graphe de l'application  $\varphi$ .

Donnons maintenant sans démonstrations des propriétés élémentaires vues en MPSI dans le cas des applications à valeurs réelles.

**Théorème 3.1.2.** — LINÉARITÉ DE L'INTÉGRALE DES APPLICATIONS EN ESCALIER

L'application,  $\mathcal{E}(J, \mathbf{F}) \rightarrow \mathbf{F}$ ;  $\vec{\varphi} \mapsto \int_J \vec{\varphi}$  est linéaire.

**Théorème 3.1.3.** — INTÉGRATION DE LA NORME —

Soit  $\vec{\varphi}$  un élément de  $\mathcal{E}(J, \mathbf{F})$  et  $\|\cdot\|$  une norme sur  $\mathbf{F}$ . Alors  $\|\vec{\varphi}\|$  est élément de  $\mathcal{E}(J, \mathbf{R})$  et

$$\left\| \int_J \vec{\varphi} \right\| \leq \int_J \|\vec{\varphi}\|.$$

**Théorème 3.1.4.** — EFFET D'UNE APPLICATION LINÉAIRE —

Soit  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$  un espace vectoriel normé, soit  $\vec{\ell}$  une application linéaire de  $\mathbf{F}$  dans  $\mathbf{G}$  et soit  $\vec{\varphi}$  un élément de  $\mathcal{E}(J, \mathbf{F})$  et  $\|\cdot\|$  une norme sur  $\mathbf{F}$ . Alors

$$i) \quad \vec{\ell} \circ \vec{\varphi} \in \mathcal{E}\left(J, \vec{\ell}(\mathbf{F})\right)^1.$$

$$ii) \quad \int_J \vec{\ell} \circ \vec{\varphi} = \vec{\ell} \left( \int_J \vec{\varphi} \right).$$

**Proposition 3.1.5.** — Soit  $\varphi$  un élément de  $\mathcal{E}(J, \mathbf{R})$ . Soient  $m$  et  $M$  des réels tels que, pour tout élément  $x$  de  $J$ ,  $m \leq \varphi \leq M^2$ , alors,

$$m|J| \leq \int_J \varphi \leq M|J|,$$

$|J|$  désigne ici la longueur du segment  $J$ .

## 3.2 Intégrales des applications continues par morceaux

Pour définir l'intégrale d'une application continue par morceaux  $\vec{f}$ , nous allons nous souvenir que, d'après 1.3, elle est limite uniforme d'une suite  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  d'applications en escalier. Nous aimerions définir l'intégrale de  $\vec{f}$  comme la limite de la suite des intégrales de  $\vec{\varphi}_n$ . Pour ce faire, nous commencerons par prouver que cette suite converge, c'est l'objet lemme 3.2.1., puis, nous montrerons que cette limite ne dépend que de  $\vec{f}$  et non du choix de la suite  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  (lemme 3.2.2.). Nous pourrions définir alors l'intégrale de  $\vec{f}$  comme la limite commune des suites des intégrales d'applications en escalier qui convergent uniformément vers  $\vec{f}$ .

**Lemme 3.2.1.** — Soit  $\vec{f}$  un élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ . Soit  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  une suite d'éléments de  $\mathcal{E}(J, \mathbf{F})$  qui converge uniformément vers  $\vec{f}$ . Posons, pour tout naturel  $n$ ,  $\vec{I}_n = \int_J \vec{\varphi}_n$ ; alors la suite  $(\vec{I}_n)_{n \in \mathbf{N}}$  converge.

*Preuve du lemme 3.2.1.* —

L'espace vectoriel  $\mathbf{F}$  étant de dimension finie, toutes les normes sont équivalentes, nous désignerons par  $\|\cdot\|$ , l'une d'elles.

La suite  $(\vec{I}_n)_{n \in \mathbf{N}}$  est bornée.

En effet la suite  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  converge uniformément vers  $\vec{f}$  donc converge vers  $\vec{f}$  dans  $(\mathcal{B}(J, \mathbf{F}), \|\cdot\|_{\infty})$ , elle est donc bornée. On peut donc choisir un réel  $M \geq 0$  tel que pour tout  $n \in \mathbf{N}$ ,  $\|\vec{\varphi}_n\|_{\infty} \leq M$  et donc pour tout  $n \in \mathbf{N}$ , d'après 3.1.3. et 3.1.5.,

$$\|\vec{I}_n\| \leq \int_J \|\vec{\varphi}_n\| \leq \|\vec{\varphi}_n\|_{\infty} |J| \leq M|J|$$

La suite  $(\vec{I}_n)_{n \in \mathbf{N}}$  admet une valeur d'adhérence  $\ell$ .

C'est une conséquence directe du théorème de Bolzano-Weierstrass, puisque  $\mathbf{F}$  est de dimension finie.

Nous allons bien sûr montrer que  $(\vec{I}_n)_{n \in \mathbf{N}}$  converge vers  $\ell$ .

Soit un réel  $\varepsilon > 0$ .

La convergence uniforme de la suite  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  nous fournit un entier  $n_0 \geq 0$  tel que pour tout entier  $n$ , si  $n \geq 0$ , alors

$$\|\vec{\varphi}_n - \vec{f}\|_{\infty} \leq \frac{\varepsilon}{4|J|}$$

et donc pour tout couple d'entiers  $(p, q)$  tels que  $p \geq q \geq n_0$ , on a  $\|\vec{\varphi}_p - \vec{\varphi}_q\|_{\infty} \leq \frac{\varepsilon}{2|J|}$  et donc :

$$\|I_p - I_q\| \leq \int_J \|\vec{\varphi}_p - \vec{\varphi}_q\| \leq \|\vec{\varphi}_p - \vec{\varphi}_q\|_{\infty} |J| \leq \frac{\varepsilon}{2}, \quad (\text{VI.6})$$

encore 3.1.3. et 3.1.5. Par ailleurs, par caractérisation d'une valeur d'adhérence on peut trouver un entier  $n_1 \geq n_0$ , tel que

$$\|\vec{I}_{n_1} - \ell\| \leq \frac{\varepsilon}{2} \quad (\text{VI.7})$$

1. Notons que  $\vec{\ell}(\mathbf{F})$  est, comme  $\mathbf{F}$ , de dimension finie.

2. Une application en escalier étant bornée de tels réels existent bien.

Pour tout entier naturel  $n$ , si  $n \geq n_1$  alors,

$$\|\vec{I}_n - \ell\| \leq \|\vec{I}_n - \vec{I}_{n_1}\| + \|\vec{I}_{n_1} - \ell\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d'après (VI.6) et (VI.7).

Donc  $(\vec{I}_n)_{n \in \mathbf{N}}$  converge vers  $\ell$ . □

**Lemme 3.2.2.** — Soit  $\vec{f}$  un élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ . Soient  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  et  $(\vec{\psi}_n)_{n \in \mathbf{N}}$  des suites d'éléments de  $\mathcal{E}(J, \mathbf{F})$  qui convergent uniformément vers  $\vec{f}$ . Posons, pour tout naturel  $n$ ,  $\vec{I}_n = \int_J \vec{\varphi}_n$  et  $\vec{J}_n = \int_J \vec{\psi}_n$ ; alors les suites  $(\vec{I}_n)_{n \in \mathbf{N}}$  et  $(\vec{J}_n)_{n \in \mathbf{N}}$  ont même limite.

*Preuve du lemme 3.2.2.* —

Soit  $\varepsilon$  un élément de  $\mathbf{R}_+^*$ . Soit  $m$  un entier naturel. D'après 3.1.2.

$$\|\vec{I}_m - \vec{J}_m\| = \left\| \int_J \vec{\varphi}_m - \int_J \vec{\psi}_m \right\| = \left\| \int_J (\vec{\varphi}_m - \vec{\psi}_m) \right\|.$$

Donc d'après 3.1.3.

$$\|\vec{I}_m - \vec{J}_m\| \leq \int_J \|\vec{\varphi}_m - \vec{\psi}_m\|.$$

Or  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  et  $(\vec{\psi}_n)_{n \in \mathbf{N}}$  convergent uniformément vers  $\vec{f}$ , donc  $(\vec{\varphi}_n - \vec{\psi}_n)_{n \in \mathbf{N}}$  converge uniformément vers l'application nulle de  $J$  dans  $\mathbf{F}$ . en particulier il existe un naturel  $n_0$  tel que pour tout naturel  $n$ , si  $n \geq n_0$  alors pour tout élément  $x$  de  $J$ ,

$$\|\vec{\varphi}_n(x) - \vec{\psi}_n(x)\| \leq \frac{\varepsilon}{|J|}.$$

Donc, d'après 3.1.5., si  $m \geq n_0$ , alors

$$\|\vec{I}_m - \vec{J}_m\| \leq \int_J \|\vec{\varphi}_m - \vec{\psi}_m\| \leq |J| \frac{\varepsilon}{|J|} = \varepsilon.$$

$\varepsilon$  étant quelconque,  $(\vec{I}_n - \vec{J}_n)_{n \in \mathbf{N}}$  converge vers  $0_{\mathbf{F}}$ , les suites  $(\vec{I}_n)_{n \in \mathbf{N}}$  et  $(\vec{J}_n)_{n \in \mathbf{N}}$ , qui convergent d'après le lemme 3.2.1., ont donc même limite. □

**Définition 3.2.3.** — INTÉGRALE D'UNE APPLICATION CONTINUE PAR MORCEAUX —

Soit  $\vec{f}$  un élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ . D'après le 1.3.,  $\vec{f}$  est limite de suites de fonctions en escalier  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$ . L'intégrales sur  $J$  de  $\varphi_n$  tend, lorsque  $n$  tend vers  $+\infty$ , vers un élément de  $\mathbf{F}$ , ne dépendant que de  $\vec{f}$  et pas de la suite  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$ , cet élément est par définition l'intégrale sur  $J$  de  $\vec{f}$ .

**Remarque** — Soit  $\vec{\varphi}$  un élément de  $\mathcal{E}(J, \mathbf{F})$ .  $\vec{\varphi}$  est aussi élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ , on peut donc l'intégrer sur  $J$ , au sens du 1., en tant qu'élément de  $\mathcal{E}(J, \mathbf{F})$ , au sens du 2., en tant qu'élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ . Montrons que l'intégrale de  $\vec{\varphi}$  est, dans les deux sens, la même. La suite d'éléments de  $\mathcal{E}(J, \mathbf{F})$  constamment égale à  $\vec{\varphi}$  converge uniformément vers  $\vec{\varphi}$ , élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ . Donc l'intégrale de  $\varphi$  au sens des applications continues par morceaux et la limite de la suite constamment égale à l'intégrale de  $\vec{\varphi}$ , au sens des applications en escalier, donc l'intégrale de  $\vec{\varphi}$  est dans les deux sens la même.

l'application de  $\mathcal{C}_m^0(J, \mathbf{F})$  dans  $\mathbf{F}$  qui à une application associe son intégrale, *prolonge* donc l'application  $\mathcal{E}(J, \mathbf{F}) \rightarrow \mathbf{F}$ ;  $\vec{\varphi} \mapsto \int_J \vec{\varphi}$ .

On peut donc, sans risque, noter encore l'intégrale sur  $J$  d'une application  $\vec{f}$ , élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ ,  $\int_J \vec{f}$ .

### 3.3 Propriétés de l'intégrale

Nous allons donner les principales propriétés de l'intégrale des fonctions continues par morceaux à valeurs vectorielles. Pour l'essentiel, ce sont les propriétés de l'intégrale des fonctions en escalier qui « passent par densité » aux applications continues par morceaux. Il conviendra de revoir dans le cours de sup. les propriétés spécifiques de l'intégrale des applications à valeurs réelles, ainsi que les sommes de Riemann.

**Théorème 3.3.1.** — LINÉARITÉ DE L'INTÉGRALE —

L'application,  $\mathcal{C}_m^0(J, \mathbf{F}) \rightarrow \mathbf{F}$ ;  $\vec{f} \mapsto \int_J \vec{f}$  est linéaire.

*Preuve du théorème 3.3.1.* — Soient  $\vec{f}$  et  $\vec{g}$  des éléments de  $\mathcal{C}_m^0(J, \mathbf{F})$  et  $\lambda$  et  $\mu$  des éléments de  $\mathbf{K}$ . Montrons que  $\int_J \lambda \vec{f} + \mu \vec{g} = \lambda \int_J \vec{f} + \mu \int_J \vec{g}$ .

Soient  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  une suite d'éléments de  $\mathcal{E}(J, \mathbf{F})$  qui converge uniformément vers  $\vec{f}$  et  $(\vec{\psi}_n)_{n \in \mathbf{N}}$  une suite d'éléments de  $\mathcal{E}(J, \mathbf{F})$  qui converge uniformément vers  $\vec{g}$ . Autrement dit  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  converge vers  $\vec{f}$  et  $(\vec{\psi}_n)_{n \in \mathbf{N}}$  converge vers  $\vec{g}$  dans  $(\mathcal{B}(J, \mathbf{F}), \|\cdot\|_\infty)$ , donc  $(\lambda \vec{\varphi}_n + \mu \vec{\psi}_n)_{n \in \mathbf{N}}$  converge dans cet espace vectoriel normé vers  $\lambda \vec{f} + \mu \vec{g}$ , donc converge uniformément vers  $\lambda \vec{f} + \mu \vec{g}$ . Donc

$$\int_J \lambda \vec{\varphi}_n + \mu \vec{\psi}_n \xrightarrow{n \rightarrow +\infty} \int_J \lambda \vec{f} + \mu \vec{g} \quad (\text{VI.8})$$

Or, d'après la linéarité de l'intégrale des applications en escalier (cf. 3.1.2.), pour tout naturel  $n$ ,

$$\int_J \lambda \vec{\varphi}_n + \mu \vec{\psi}_n = \lambda \int_J \vec{\varphi}_n + \mu \int_J \vec{\psi}_n,$$

et, par définition de l'intégrale  $\int_J \vec{\varphi}_n \xrightarrow{n \rightarrow +\infty} \int_J \vec{f}$  et  $\int_J \vec{\psi}_n \xrightarrow{n \rightarrow +\infty} \int_J \vec{g}$ , donc

$$\int_J \lambda \vec{\varphi}_n + \mu \vec{\psi}_n \xrightarrow{n \rightarrow +\infty} \lambda \int_J \vec{f} + \mu \int_J \vec{g}. \quad (\text{VI.9})$$

D'après (VI.8–VI.9), par unicité de la limite,

$$\int_J \lambda \vec{f} + \mu \vec{g} = \lambda \int_J \vec{f} + \mu \int_J \vec{g}.$$

D'où le résultat. □

**Théorème 3.3.2.** — INTÉGRATION DE LA NORME —

Soit  $\vec{f}$  un élément de  $\mathcal{C}_m^0(J, \mathbf{F})$  et  $\|\cdot\|$  une norme sur  $\mathbf{F}$ . Alors  $\left\| \int_J \vec{f} \right\|$  est élément de  $\mathcal{C}_m^0(J, \mathbf{F})$  et

$$\left\| \int_J \vec{f} \right\| \leq \int_J \|\vec{f}\|.$$

*Preuve du théorème 3.3.2.* — Que  $\left\| \int_J \vec{f} \right\|$  soit élément de  $\mathcal{C}_m^0(J, \mathbf{F})$ , est laissé en exercice.

Soit  $(\vec{\varphi}_n)_{n \in \mathbf{N}}$  une suite d'éléments de  $\mathcal{E}(J, \mathbf{F})$  qui converge uniformément vers  $\vec{f}$ .

D'une part, pour tout naturel  $n$ ,  $\|\vec{\varphi}_n\|$  est élément de  $\mathcal{E}(J, \mathbf{R})$ . En effet si  $(a_i)_{i=0, \dots, m}$  est une subdivision de  $J$  subordonnée à  $\vec{\varphi}_n$ , alors pour  $i = 0, \dots, m-1$ ,  $\vec{\varphi}_n$ , et donc  $\|\vec{\varphi}_n\|$  sont constantes sur  $]a_i, a_{i+1}[$ .

D'autre part on montre facilement que  $(\|\vec{\varphi}_n\|)_{n \in \mathbf{N}}$  converge uniformément vers  $\|\vec{f}\|$ , en effet pour tout naturel  $n$  et tout élément  $x$  de  $J$ ,

$$\left| \|\vec{f}(x)\| - \|\vec{\varphi}_n(x)\| \right| \leq \|\vec{\varphi}_n(x) - \vec{f}(x)\|.$$

Donc, par définition de l'intégrale,

$$\int_J \|\vec{\varphi}_n\| \xrightarrow{n \rightarrow +\infty} \int_J \|\vec{f}\|. \quad (\text{VI.10})$$

et

$$\int_J \vec{\varphi}_n \xrightarrow{n \rightarrow +\infty} \int_J \vec{f}. \quad (\text{VI.11})$$

Or, d'après 2.1.3., pour tout naturel  $n$ ,

$$\left\| \int_J \vec{\varphi}_n \right\| \leq \int_J \|\vec{\varphi}_n\|. \quad (\text{VI.12})$$

Par passage à la limite dans l'inégalité (VI.12), en utilisant (VI.10–VI.11) et la continuité de la norme, il vient :

$$\left\| \int_J \vec{f} \right\| \leq \int_J \|\vec{f}\|.$$

D'où le résultat. □

**Théorème 3.3.3.** — EFFET D'UNE APPLICATION LINÉAIRE —

Soit  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$  un espace vectoriel normé, soit  $\vec{\ell}$  une application linéaire de  $\mathbf{F}$  dans  $\mathbf{G}$ , soit  $\vec{f}$  un élément de  $\mathcal{C}_m^0(J, \mathbf{F})$  et soit  $\|\cdot\|$  une norme sur  $\mathbf{F}$ . Alors

- i)  $\vec{\ell} \circ \vec{f} \in \mathcal{C}_m^0(J, \mathbf{G})$ .
- ii)  $\int_J \vec{\ell} \circ \vec{f} = \vec{\ell} \left( \int_J \vec{f} \right)$ .

La preuve de ce résultat est laissée en exercice. Elle repose évidemment sur le résultat analogue pour les applications en escalier 2.1.4.

Le théorème 3.3.3. admet les corollaires importants suivants. Leurs preuves sont laissées en exercice, elles sont proches de celles des corollaires 2.1.10 et 2.1.11.

**Corollaire 3.3.4.** — Soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_d)$  une base de  $\mathbf{F}$  notée  $\mathcal{B}$ .  $\vec{f}$  désigne une application de  $J$  dans  $\mathbf{F}$  et l'on note, pour  $i = 1, 2, \dots, d$ , la  $i^e$  application composante de  $\vec{f}$  dans la base  $\mathcal{B}$ .

$\vec{f}$  est continue par morceaux, si et seulement si, pour tout élément  $i$  de  $\{1, 2, \dots, d\}$ ,  $f_i$  est continue par morceaux. De plus, si c'est le cas,

$$\int_J \vec{f} = \sum_{i=1}^d \int_J f_i \cdot \vec{e}_i.$$

Autrement dit la  $i^e$  composante de l'intégrale de  $\vec{f}$  dans  $\mathcal{B}$  est l'intégrale de la  $i^e$  composante de  $\vec{f}$  dans  $\mathcal{B}$ .

Ce résultat est important, il permet de généraliser des résultats connus pour l'intégrale d'applications à valeurs réelles, à l'intégrale d'applications vectorielles, en raisonnant sur les composantes dans une base de ces dernières.

**Corollaire 3.3.5.** — Soit  $f$  une application de  $J$  dans  $\mathbf{C}$ .  $f$  est continue par morceaux, si et seulement si  $\operatorname{Re} f$  et  $\operatorname{Im} f$  le sont. Si tel est le cas, alors,

$$\int_J f = \int_J \operatorname{Re} f + i \int_J \operatorname{Im} f.$$

**Corollaire 3.3.6.** — Soit  $f$  une application de  $J$  dans  $\mathbf{C}$ .  $f$  est continue par morceaux, si et seulement si  $\bar{f}$  l'est. Si tel est le cas, alors,

$$\int_J \bar{f} = \overline{\int_J f}.$$

Définissons maintenant l'intégrale d'une application entre deux points et introduisons une nouvelle notation.

**Définition 3.3.7.** — Soit  $I$  un intervalle non vide quelconque, soit  $\vec{f} : I \rightarrow \mathbf{F}$ ;  $x \mapsto \vec{f}(x)$  une application continue par morceaux et soient  $a$  et  $b$  des éléments de  $I$ . L'intégrale de  $\vec{f}$  entre  $a$  et  $b$ , notée  $\int_a^b \vec{f}(x) dx$ , est définie par :

- $\int_a^b \vec{f}(x) dx = \int_{[a,b]} \vec{f}|_{[a,b]},$  pour  $a < b$ ,

1. Notons que  $\vec{\ell}(\mathbf{F})$  est, comme  $\mathbf{F}$ , de dimension finie.

- $\int_a^b \vec{f}(x)dx = \vec{0}_{\mathbf{F}}$ , pour  $a = b$ ,
- $\int_a^b \vec{f}(x)dx = - \int_{[b,a]} \vec{f}(x)dx$ , pour  $b < a$ .

Avec ces notations on démontre comme en sup. le résultat suivant :

**Définition 3.3.8.** — LOI DE CHASLES —

Soit  $I$  un intervalle non vide quelconque, soit  $\vec{f} : I \rightarrow \mathbf{F}$ ;  $x \mapsto \vec{f}(x)$  une application continue par morceaux et soient  $a, b$  et  $c$  des éléments de  $I$ . Alors,

$$\int_a^b \vec{f}(x)dx = \int_a^c \vec{f}(x)dx + \int_c^b \vec{f}(x)dx.$$

Pour finir ce paragraphe donnons, avec les notation du 3.3.7., un formulaire qui se déduit immédiatement des résultats précédents.

**Formulaire 3.3.9.** — Soit  $I$  un intervalle non vide quelconque. Soient  $\vec{f} : I \rightarrow \mathbf{F}$ ;  $x \mapsto \vec{f}(x)$  et  $\vec{g} : I \rightarrow \mathbf{F}$ ;  $x \mapsto \vec{g}(x)$  des applications continues par morceaux. On notera pour  $i = 1, 2, \dots, d$ ,  $f_i$  la  $i^e$  application composante de  $\vec{f}$  dans une base  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_d)$  de  $\mathbf{F}$ , notée  $\mathcal{B}$ . Soient enfin  $(\mathbf{G}, \|\cdot\|_{\mathbf{G}})$  un espace vectoriel normé et  $\vec{\ell}$  une application linéaire de  $\mathbf{F}$  dans  $\mathbf{G}$ . Alors,

$$i) \int_a^b \lambda \cdot \vec{f}(x) + \mu \cdot \vec{g}(x)dx = \lambda \cdot \int_a^b \vec{f}(x)dx + \mu \int_a^b \vec{g}(x)dx,$$

$$ii) \left\| \int_a^b \vec{f}(x)dx \right\| \leq \left| \int_a^b \|\vec{f}(x)\| dx \right|,^1$$

$$iii) \vec{\ell} \left( \int_a^b \vec{f}(x)dx \right) = \int_a^b \vec{\ell} \circ \vec{f}(x)dx,$$

$$iv) \int_a^b \vec{f}(x)dx = \sum_{i=1}^d \int_a^b f_i(x)dx \cdot \vec{e}_i,$$

v) Si  $f$  est une application de  $I$  dans  $\mathbf{C}$  continue par morceaux, alors,

$$\int_a^b f(x)dx = \int_a^b \operatorname{Re} f(x)dx + i \int_a^b \operatorname{Im} f(x)dx \text{ et } \int_a^b \overline{f(x)}dx = \overline{\int_a^b f(x)dx}.$$

### 3.4 Primitives et intégration

Nous allons rapidement montrer qu'en un certain sens, l'intégration est le « contraire » de la dérivation ». En fait nous ne ferons qu'énoncer des résultats connus de longue date pour l'intégrale d'applications numériques. On laissera le lecteur prouver la plupart de ces résultats. Pour ce faire il pourra, soit raisonner sur les composantes dans une base des applications considérées, auxquelles il appliquera les résultats de sup. (cf. 3.3.4.), soit, ce qui est plus élégant, reproduire presque mot à mot les preuves de sup. pour prouver dans l'ordre dans lequel ils se présentent, les résultats qui suivent.

Dans tout ce paragraphe  $I$  désigne un intervalle d'intérieur non vide quelconque. Soit  $\vec{f}$  une application de  $I$  dans  $\mathbf{F}$  continue par morceaux et  $a$  un point de  $I$ . Nous nous proposons d'étudier la régularité de l'application suivante :

$$\vec{\Phi}_a : I \rightarrow \mathbf{F}; x \mapsto \int_a^x \vec{f}(t)dt.$$

Rappelons avant de commencer que  $\vec{f}$  a notamment une limite à droite et à gauche en tout point intérieur à  $I$ , de plus si l'extrémité inférieure de  $I$  est un point  $\alpha$  de  $\mathbf{R}$ ,  $\vec{f}$  admet en  $\alpha$  une limite à droite, si l'extrémité supérieure de  $I$  est un point  $\beta$  de  $\mathbf{R}$ ,  $\vec{f}$  admet en  $\beta$  une limite à gauche. Nous pouvons maintenant énoncer le lemme technique suivant sur lequel repose l'essentiel de notre étude.

**Lemme 3.4.1.** — l'application  $\vec{\Phi}_a$  admet en tout point  $x$  intérieur à  $I$  une dérivée à droite et à gauche et plus précisément,

$$(\vec{\Phi}_a)'_d(x) = \vec{f}(x_{+0}), (\vec{\Phi}_a)'_g(x) = \vec{f}(x_{-0}).$$

1. La valeur absolue est ici nécessaire, car on peut avoir  $a > b$ .

De plus, si l'extrémité inférieure de  $I$  est un point  $\alpha$  de  $I$ ,  $\vec{\Phi}_a$  admet en  $\alpha$  une dérivée à droite et

$$(\vec{\Phi}_a)'_d(\alpha) = \vec{f}(\alpha_{+0}),$$

si l'extrémité supérieure de  $I$  est un point  $\beta$  de  $I$ ,  $\vec{\Phi}_a$  admet en  $\beta$  une dérivée à gauche et

$$(\vec{\Phi}_a)'_g(\beta) = \vec{f}(\beta_{-0}).$$

*Preuve du lemme 3.4.1. —*

On en déduit les deux corollaires suivants

**Corollaire 3.4.2.** — *L'application  $\vec{\Phi}_a$  est continue.*

**Corollaire 3.4.3.** — *Si  $\vec{f}$  est continue alors l'application*

$$\vec{\Phi}_a : I \rightarrow \mathbf{F}; x \mapsto \int_a^x \vec{f}(t)dt,$$

*est de classe  $\mathcal{C}^1$ , et*

$$\vec{\Phi}'_a : I \rightarrow \mathbf{F}; x \mapsto \vec{f}(x).$$

**Définition 3.4.4.** — *On appelle primitive d'une application  $\vec{f}$  de  $I$  dans  $\mathbf{F}$  continue, toute application  $\vec{\Phi}$  de  $I$  dans  $\mathbf{F}$  dérivable, telle que  $\vec{\Phi}' = \vec{f}$ .*

Le corollaire 3.4.3. nous apprend que toute application continue sur un intervalle admet au moins une primitive. Plus précisément :

**Proposition 3.4.5.** — *Soit  $\vec{f}$  une application de l'intervalle  $I$  à valeurs dans  $\mathbf{F}$  continue. Alors,*

*i) Deux primitives de  $\vec{f}$  diffèrent d'une constante.*

*ii) Pour tout élément  $a$  de  $I$ , il existe une unique primitive de  $\vec{f}$  nulle en  $a$ , c'est :  $\vec{\Phi}_a : I \rightarrow \mathbf{F}; x \mapsto \int_a^x \vec{f}(t)dt$ .*

Les primitives, comme dans le cas particulier des fonctions à valeurs réelles, servent à calculer des intégrales :

**Proposition 3.4.6.** — *Soit  $\vec{f}$  une application de l'intervalle  $I$  à valeurs dans  $\mathbf{F}$  continue. Soient  $a$  et  $b$  des points de  $I$ , et  $\vec{\Phi}$  une primitive de  $\vec{f}$ . Alors,*

$$\int_a^b \vec{f}(t)dt = \vec{\Phi}(b) - \vec{\Phi}(a).$$

**Exercice 3.4.7.** — Soit  $\vec{f}$  une application de  $\mathbf{R}$  dans  $\mathbf{F}$  continue. On considère l'application

$$\vec{G} : \mathbf{R} \rightarrow \mathbf{F}; x \mapsto \int_{2x}^{x^2} \vec{f}(t)dt.$$

Montrer que  $\vec{G}$  est de classe  $\mathcal{C}^1$  et donner sa dérivée.

*Solution de l'exercice 3.4.7. —*

L'intégration par parties se généralise au cas vectoriel en remplaçant le produit par une forme bilinéaire.

**Proposition 3.4.8.** — INTÉGRATION PAR PARTIES —

Soient  $\mathbf{F}_1$  et  $\mathbf{F}_2$  des espaces vectoriels de dimensions finies, soit  $(\mathbf{G}, \|\cdot\|)$  un espace vectoriel normé, soit  $\vec{B}$  une application bilinéaire de  $\mathbf{F}_1 \times \mathbf{F}_2$  dans  $\mathbf{G}$ , soient  $\vec{f}_1$  et  $\vec{f}_2$  des applications définies sur  $I$ , à valeurs respectivement dans  $\mathbf{F}_1$  et  $\mathbf{F}_2$ , de classe  $\mathcal{C}^1$ . Alors

$$\int_a^b \vec{B}(\vec{f}_1(t), \vec{f}_2(t)) dt = [\vec{B}(\vec{f}_1(t), \vec{f}_2(t))]_a^b - \int_a^b \vec{B}(\vec{f}_1'(t), \vec{f}_2(t)) dt,$$

$$\int_a^b \vec{B}(\vec{f}_1(t), \vec{f}_2'(t)) dt = [\vec{B}(\vec{f}_1(t), \vec{f}_2(t))]_a^b - \int_a^b \vec{B}(\vec{f}_1(t), \vec{f}_2'(t)) dt.$$

Rearquons que  $t \mapsto \vec{B}(\vec{f}_1(t), \vec{f}_2(t))$ ,  $t \mapsto \vec{B}(\vec{f}_1'(t), \vec{f}_2(t))$  prennent en fait leurs valeurs dans l'espace vectoriel de dimension finie  $\vec{B}(\mathbf{F}_1 \times \mathbf{F}_2)$ , l'étude de leurs intégrales ne sort donc pas du cadre de ce cours.

*Preuve de 3.4.8.* — C'est une conséquence immédiate de 2.1.14. et de 3.4.6. □

Un cas particulier classique est celui où l'application  $f_1$  est à valeurs dans  $\mathbf{K}$ ,  $\vec{f}_2$  dans le  $\mathbf{K}$ -e.v.  $\mathbf{F}$  et  $B$  le produit d'un vecteur par un scalaire :

$$B : \mathbf{K} \times \mathbf{F}; (\lambda, \vec{x}) \mapsto \lambda \cdot \vec{x}.$$

La formule s'écrit alors :

$$\int_a^b f_1'(t) \cdot \vec{f}_2(t) dt = [f_1(t) \cdot \vec{f}_2(t)]_a^b - \int_a^b f_1(t) \cdot \vec{f}_2'(t) dt,$$

$$\int_a^b f_1(t) \cdot \vec{f}_2'(t) dt = [f_1(t) \cdot \vec{f}_2(t)]_a^b - \int_a^b f_1'(t) \cdot \vec{f}_2(t) dt.$$

**Exercice** — Soit  $([a, b], \mathbf{F})$  un arc paramétré, de classe  $\mathcal{C}^2$  d'un plan euclidien  $\mathbf{E}_2$ , muni de sa structure affine canonique. On suppose que  $\mathbf{F}(a) = \mathbf{F}(b)$  et que  $\vec{\mathbf{F}}'(a) = \vec{\mathbf{F}}'(b)$ . On note  $\langle \cdot | \cdot \rangle$  le produit scalaire et  $\|\cdot\|$  la norme associée. Soit  $\mathbf{O}$  un point de  $\mathbf{E}_2$  non situé sur le support de l'arc. On suppose que pour tout  $t \in [a, b]$ ,  $\mathbf{F}''(t) = \frac{\vec{\mathbf{O}\mathbf{F}}(t)}{\|\vec{\mathbf{O}\mathbf{F}}(t)\|^2}$ . (Cet arc modélise un mouvement à accélération centrale de centre  $\mathbf{O}$ ). Evaluer

$$\frac{1}{2} \int_a^b \|\vec{\mathbf{F}}'(t)\|^2 dt.$$

Qu'en concluez vous ?

*Solution* —



### 3.5 Inégalité des accroissements finis, formules de Taylor

Utilisons le lien entre primitive et intégrale pour donner des estimations globales d'une application sur un segment

Commençons par l'inégalité des accroissements finis.

**Proposition 3.5.1.** — INÉGALITÉ DES ACCROISSEMENTS FINIS —

Soit  $\vec{f}$  une application de l'intervalle  $I$  à valeurs dans  $\mathbf{F}$  de classe  $\mathcal{C}^1$ .  $\|\cdot\|$  désigne une norme quelconque sur  $\mathbf{F}$ , et l'on suppose qu'il existe un réel  $k$  tel que pour tout élément  $x$  de  $I$ ,  $\|\vec{f}'(x)\| \leq k$ . Alors, pour tout  $x$  et tout  $y$  points de  $I$ ,

$$\|\vec{f}(x) - \vec{f}(y)\| \leq k|x - y|.$$

*Preuve de la proposition 3.5.1.* — Soient  $x$  et  $y$  des points de  $I$ . L'application  $\vec{f}$  étant de classe  $\mathcal{C}^1$ ,  $\vec{f}'$  est continue et  $\vec{f}$  en est une primitive. Donc d'après 2.4.5.,

$$\vec{f}(x) - \vec{f}(y) = \int_y^x \vec{f}'(t) dt.$$

Donc d'après 3.3.9. point ii),

$$\|\vec{f}(x) - \vec{f}(y)\| \leq \left| \int_y^x \|\vec{f}'(t)\| dt \right|.$$

Donc par positivité de l'intégrale d'une fonction à valeurs réelles (cf. cours de sup.)

$$\|\vec{f}(x) - \vec{f}(y)\| \leq \left| \int_y^x k dt \right| = k|x - y|.$$

D'où le résultat. □

Passons à la formule de Taylor avec reste intégrale. Les composantes d'une application  $\vec{f}$  supposée de classe  $\mathbf{C}^{n+1}$  étant de classe  $\mathbf{C}^{n+1}$ , la formule de Taylor avec reste intégrale appliquée à chacune d'elles donne le résultat qui suit. Toutefois il est plus élégant de donner de ce résultat une preuve directe et cette démarche présente aussi l'intérêt de revisiter la démonstration donnée dans le cas d'une fonction réelle, dont elle reprend la démarche.

**Proposition 3.5.2.** — FORMULE DE TAYLOR AVEC RESTE INTÉGRALE —

Soient  $n$  un entier naturel et  $\vec{f}$  une application d'un intervalle  $I$ , non réduit à un point à valeurs dans  $\mathbf{F}$  de classe  $\mathcal{C}^{n+1}$ , pour tout couple  $(a, b)$  de points de  $I$ ,

$$\vec{f}(b) = \sum_{k=0}^n \frac{(b-a)^k}{k!} \cdot \vec{f}^{(k)}(a) + R_n,$$

$$\text{où } R_n = \int_a^b \frac{(b-t)^n}{n!} \cdot \vec{f}^{(n+1)}(t) dt.$$

$R_n$  s'appelle le reste d'ordre  $n$ , il donne le nom à la formule.

*Preuve de la proposition 3.5.2. —*

La formule de Taylor est une formule globale qui donne l'expression de  $f$  en un point en fonction des valeurs des dérivées successive de  $f$  en  $a$  et d'un reste. Le reste n'est pas explicite, il est donné par une intégrale. La formule suivante donne une majoration du reste ce qui permet de mesurer l'écart entre  $f(b)$  est la partie polynômiale de la formule.

**Proposition 3.5.3.** — INÉGALITÉ DE TAYLOR-LAGRANGE —

Soient  $n$  un entier naturel et  $\vec{f}$  une application d'un segment  $[a, b]$  (resp.  $[b, a]$ ) non réduit à un point, à valeurs dans  $\mathbf{F}$  de classe  $\mathcal{C}^{n+1}$ . Soit  $M$  un réel tel que pour tout élément  $t$  de  $[a, b]$ , (resp.  $[b, a]$ ),

$$\|\vec{f}^{(n+1)}(t)\| \leq M,$$

où  $\|\cdot\|$  désigne une norme quelconque sur  $\mathbf{F}$ . Alors :

$$\left\| \vec{f}(b) - \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) \right\| \leq M \frac{|b-a|^{n+1}}{(n+1)!}.$$

**Remarque** —

1. La continuité de  $\vec{f}^{(n+1)}$  sur le segment  $[a, b]$  (resp.  $[b, a]$ ) assure l'existence du réel  $M$  que l'on peut en particulier prendre égal à  $\sup_{t \in [a, b], (\text{resp. } [b, a])} \vec{f}^{(n+1)}(t)$ .
2. Pour  $n = 0$ , on retrouve.....

*Preuve de la proposition 3.5.3. —*

Bien que l'inégalité de Taylor-Lagrange, soit de nature globale elle peut donner aussi la formule locale de Taylor-Young 2.2.9., sous les hypothèses fortes du programme.

Soit  $\vec{f}$  une application de  $I$  dans  $F$  de classe  $\boxed{\mathcal{C}^{n+1}}$ , où  $n$  est entier naturel. et soit  $a$  un point de  $I$ . Traitons le cas où  $a$  est intérieur à  $I$  (on traite de même le cas où  $a$  est extrémité de  $I$ ).

Choisissons un réel  $\eta > 0$  tel que  $[a - \eta, a + \eta]$  soit inclus dans  $I$ ...

Finalement

$$\boxed{\vec{f}(x) = \sum_{k=0}^n \frac{(x-a)^k}{k!} \cdot \vec{f}^{(k)}(a) + \vec{O}_{\substack{x \rightarrow a \\ x \in I}} \left( (x-a)^{n+1} \right)}$$

La formule de Taylor est souvent énoncée avec un  $\vec{O}_{\substack{x \rightarrow a \\ x \in I}} \left( (x-a)^n \right)$  ce qui est moins bien, c'est sûrement un vestige du temps pas si lointain où les taupins connaissaient la vraie formule de Taylor avec comme seule hypothèse l'existence en  $a$  d'une dérivée  $n^e$ , et où l'on ne pouvait pas faire mieux...

### 3.6 Intégrale de la limite uniforme d'une suite d'applications

Donnons brutalement le seul résultat de ce paragraphe.

**Proposition 3.6.1.** — *Soit  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications d'un segment  $[a, b]$  à valeurs dans  $\mathbf{K}$  (ou dans  $\mathbf{F}$ ) continues. On suppose que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément vers une application  $f$ . Alors*

1.  $f$  est continue.
2. La suite  $\left( \int_a^b f_n(t) dt \right)_{n \in \mathbf{N}}$  converge vers  $\int_a^b f(t) dt$ .

*Preuve de la proposition 3.6.1.* — Le premier point à déjà été prouvé dans la partie 1. Notons que c'est la continuité de  $f$  qui permet ici de parler de son intégrale. passons au second point. Soit  $\varepsilon$  un réel strictement positif.

Nous verrons dans un autre chapitre de nombreux théorèmes d'interversion de limite et d'intégrale, plus puissants que la proposition 2.6.1. Notons que les deux hypothèses suivantes :

- $f$  est définie sur un segment ;
  - La convergence de la suite  $(f_n)_{n \in \mathbf{N}}$  est uniforme ;
- sont cruciales.

Donnons des exemples où l'une puis l'autre de ces deux hypothèses est absente et où la conclusion du théorème ne demeure plus.

- Soit  $(f_n)_{n \in \mathbf{N}^*}$  la suite d'applications de  $[-1, 1]$  dans  $\mathbf{R} \dots$

- Soit  $(g_n)_{n \in \mathbf{N}^*}$  la suite d'applications de  $\mathbf{R}_+$  dans  $\mathbf{R} \dots$



## Chapitre VII

# CALCUL DIFFÉRENTIEL, ÉPISODE 1

Le cours de calcul différentiel est tout à la fois un sujet vaste, nouveau et parmi les plus important de l'année tant par la place qu'il occupe en mathématiques que par le rôle central qu'il joue dans la modélisation des sciences. Nous avons, pour ces raisons choisi de le scinder en deux chapitres. Le premier et présent est un chapitre de découverte, riche en notions nouvelles : dérivée directionnelle, différentielle, applications de classes  $\mathcal{C}^1$ . Il est pauvre en applications significatives, exceptées peut être l'étude des extremums et se concentre sur les outils de calcul. Le second présentera dans quelques temps, ce que pompeusement l'on pourrait appeler la géométrie différentielle, permettant de traiter en exercice de nombreuses applications essentielles, circulation de champs de vecteurs, équations aux dérivées partielles, extrémums liés etc.

Dans tout ce chapitre  $\mathbf{E}$  désignera un espace vectoriel sur le corps des **nombres réels** de dimension finie non nulle  $p$  et  $\mathbf{F}$  un espace vectoriel sur le corps des **nombres réels** de dimension finie non nulle  $n$ . Le plus souvent,  $\mathbf{E}$  et parfois  $\mathbf{F}$  seront munis de leur structure naturelle d'espace affine. Concrètement nous noterons par des minuscules italiques  $m, p, q$  etc., les points de l'espace affine  $\mathbf{E}$ , par des minuscules italiques coiffées de flèches  $\vec{h}, \vec{x}, \vec{k}$  etc., les éléments de l'espace vectoriel  $\mathbf{E}$ . Nous écrirons indifféremment, pour  $p$  et  $m$  points de  $\mathbf{E}$ ,  $\overrightarrow{mp}$  ou  $p - m$  l'unique vecteur qui ajouté à  $m$  donne  $p$ .

Rappelons qu'en dimension finie, toutes les normes sont équivalentes, donc dans l'étude de toutes les notions topologiques nous ne préciserons pas les normes dont on équipe  $\mathbf{E}$  et  $\mathbf{F}$ . Nous désignerons donc par  $\|\cdot\|_{\mathbf{E}}$  une norme sur  $\mathbf{E}$  et par  $\|\cdot\|_{\mathbf{F}}$  une norme sur  $\mathbf{F}$  que nous ne préciserons pas d'avantage. Parfois lorsque  $\mathbf{E}$  sera muni d'une base  $\mathcal{B}$ , pour des raisons de commodité, nous ferons appelle à la norme  $\|\cdot\|_{\mathcal{B},\infty}$ , qui, rappelons le, est définie par  $\|\vec{x}\|_{\mathcal{B},\infty} = \sup_{i=0,\dots,p} (|x_i|)$ , où, pour  $i = 1, 2, \dots, p$ ,  $x_i$  désigne la  $i^{\text{e}}$  coordonnée de  $\vec{x}$  dans  $\mathcal{B}$ .

## RAPPELS ET COMPLÉMENTS SUR LA CONTINUITÉ

### 1.1 Exemples d'étude de continuité

Nous avons vu dans le cours de topologie que la continuité peut s'obtenir par des théorèmes de transfert, ici au contraire nous allons traiter des exemples où les théorèmes de transfert s'avèrent insuffisants.

**Exemple 1.1.1.** — Etudions la continuité de l'application

$$f : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} \frac{x^4 + 3xy^2 - 5y^3}{x^2 + y^2}, & \text{pour } (x, y) \neq (0, 0), \\ 0, & \text{pour } (x, y) = (0, 0). \end{cases}$$

**Exemple 1.1.2.** — Etudions la continuité de l'application

$$g : \mathbf{R}_+ \times \mathbf{R}_+ \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} \frac{x^2 - 3xy^2}{x+y}, & \text{pour } (x, y) \neq (0, 0), \\ 0, & \text{pour } (x, y) = (0, 0). \end{cases}$$

## 1.2 Mise en garde sur la continuité partielle

Revenons sur une mise en garde déjà faite dans le cours de topologie. la continuité en un point d'un e.v.n. des composantes d'une application à valeurs dans un espace vectoriel de dimension finie assure la continuité de l'application en ce point. Hélas! comme nous l'avons dit alors, il n'y a rien de similaire lorsque l'application est définie sur un espace vectoriel de dimension finie. Reprenons le contre-exemple donné alors. Soit l'application

$$f : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} 1, & \text{pour } y = x \text{ et } x \neq 0, \\ 0, & \text{pour } y \neq x, \\ 0, & \text{pour } x = y = 0. \end{cases}$$

Notons  $f_1$  et  $f_2$  les « applications partielles » en  $(0, 0)$  :

$$f_1 : \mathbf{R} \rightarrow \mathbf{R}; x \mapsto f(x, 0),$$

$$f_2 : \mathbf{R} \rightarrow \mathbf{R}; y \mapsto f(0, y),$$

$f_1$  et  $f_2$  sont continues en 0 en effet...

. Par contre  $f$  n'est pas continue en  $(0, 0)$ , en effet.

Nous allons voir qu'il à pire encore. Pour ce faire, définissons une notion importante, celle de continuité directionnelle.

Soit  $\vec{f}$  une application d'une partie non vide  $U$  de  $\mathbf{E}$ , à valeurs dans  $\mathbf{F}$ . Soit  $a$  un point de  $U$  ( $\mathbf{E}$  est vu comme un espace affine) et  $\vec{u}$  un vecteur non nul de  $\mathbf{E}$ . Nous noterons, dans ce cours  $\vec{f}_{a,\vec{u}}$  la fonction de la variable réelle  $t$  définie par :

$$\boxed{f_{a,\vec{u}}(t) = \vec{f}(a + t\vec{u})}$$

Lorsque  $t$  décrit  $\mathbf{R}$ ,  $a + t\vec{u}$  décrit la droite  $D$ , passant par  $a$  et dirigé par  $\vec{u}$ . Les valeurs de  $\vec{f}_{a,\vec{u}}$  sont donc celles que prend  $\vec{f}$  sur  $D \cap U$ , cf. figure ci-dessous.

**Remarque** — La fonction  $\vec{f}_{a,\vec{u}}$  est définie en 0, nous verrons même que si  $U$  est ouvert elle est définie au voisinage de 0.

Cette remarque rend licite la définition suivante.

**Définition 1.2.1.** — On dit que  $\vec{f}$  est continue en  $a$  suivant (ou selon)  $\vec{u}$  si, par définition,  $\vec{f}_{a,\vec{u}}$  est continue en 0.

**Remarque** — Soit  $\lambda$  un réel non nul. on pose  $\vec{v} := \lambda\vec{u}$ .  $\vec{f}_{a,\vec{v}}(t) = \vec{f}_{a,\vec{u}}(\dots)$ . Donc  $\vec{f}_{a,\vec{v}}$  est continue en 0 si et seulement si  $\vec{f}_{a,\vec{u}}$  l'est. Donc  $\vec{f}$  est continue en  $a$  selon  $\vec{v}$  si et seulement si elle l'est suivant  $\vec{u}$ . On peut donc parler de continuité en  $a$  dans la *direction du vecteur*  $\vec{u}$ .

La notion de continuité directionnelle est une notion très faible, puisque la continuité en un point dans toutes les directions ne donne même pas la continuité au point, comme en témoigne le contre exemple édifant suivant :

**Exemple 1.2.2.** — Soit l'application

$$f : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} 1, & \text{pour } y = x^2 \text{ et } x \neq 0, \\ 0, & \text{pour } y \neq x^2, \\ 0, & \text{pour } x = y = 0. \end{cases}$$



**Exercice** — Soit l'application  $;\mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \begin{cases} \frac{xy^2}{x^2+y^4}, & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases}$

1. Montrer que  $f$  est continue dans toutes les directions en  $(0, 0)$ .
2.  $f$  est-elle continue en  $(0, 0)$ .

*Solution* —

## DÉRIVÉES DIRECTIONNELLES, APPLICATIONS DE CLASSE $\mathcal{C}^1$ ET DIFFÉRENTIABLES

Dans toute cette partie  $U$  est un **ouvert** de  $\mathbf{E}$  non vide.

### 2.1 Dérivées directionnelles

Dans tout ce paragraphe on considère une application  $\vec{f}$  de  $U$  dans  $\mathbf{F}$ . On désignera par  $a$  un point de  $U$  et  $\vec{u}$  un vecteur de  $\mathbf{F}$  non nul.

**Proposition 2.1.1.** — *L'application  $\vec{f}_{a, \vec{u}}$  est définie au voisinage de 0, ou, pour être plus précis, il existe un élément  $\varepsilon$  de  $\mathbf{R}_+^*$  tel que  $\vec{f}_{a, \vec{u}}$  soit définie sur  $] -\varepsilon, \varepsilon[$ .*

*Preuve de la proposition 2.1.1.* —

La question de la dérivabilité de  $\vec{f}_{a, \vec{u}}$  en 0 est donc pertinente. On pose la définition suivante :

**Proposition 2.1.2.** — On dit que  $\vec{f}$  est dérivable en  $a$  si, par définition,  $\vec{f}_{a,\vec{u}}$  est dérivable en 0. Si tel est le cas  $\vec{f}'_{a,\vec{u}}(0)$  s'appelle dérivée de  $\vec{f}$  en  $a$  selon (suivant)  $\vec{u}$ . On la note  $D_{\vec{u}}\vec{f}(a)$ .

**Remarque** —

1.  $D_{\vec{u}}\vec{f}(a)$  est un élément de  $\mathbf{F}$ .
2. soit  $\lambda$  un réel non nul. on pose  $\vec{v} := \lambda\vec{u}$ .  $\vec{f}_{a,\vec{v}}(t) = \vec{f}_{a,\vec{u}}(\lambda t)$ . Donc  $\vec{f}$  est dérivable en  $a$  selon  $\vec{v}$  si et seulement si elle l'est suivant  $\vec{u}$ . On peut donc parler de dérivabilité en  $a$  dans la direction du vecteur  $\vec{u}$ . Attention cependant, en général  $D_{\vec{v}}\vec{f}(a) \neq D_{\vec{u}}\vec{f}(a)$ .
3. En revenant à la définition de la dérivée d'une application d'une variable réelle,  $D_{\vec{u}}\vec{f}(a)$  est lorsque elle existe donnée par

$$D_{\vec{u}}\vec{f}(a) = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{\vec{f}(a + t\vec{u}) - \vec{f}(a)}{t} \quad (\text{VII.1})$$

Au risque de décevoir, disons le tout net, la notion de dérivée directionnelle, sans autre hypothèse est une notion sans grand intérêt. Reprenons en effet l'exemple 1.2.2. On a ...

CAS où  $\mathbf{E} = \mathbf{R}^p$

Considérons le cas simple d'une application  $\vec{g}$  de  $\mathbf{R}^p$  dans  $\mathbf{F}$  :

$$\vec{g} : \mathbf{R}^p \rightarrow \mathbf{F}; (x_1, \dots, x_p) \mapsto \vec{g}(x_1, \dots, x_p).$$

Donnons nous un point  $(a_1, \dots, a_p)$  de  $\mathbf{R}^p$ . pour tout élément  $i$  de  $\{1, \dots, p\}$ , on définit l'application  $\vec{g}_i$ , (application partielle)

$$\vec{g}_i : \mathbf{R} \rightarrow \mathbf{F}; x_i \mapsto \vec{g}(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_p),$$

$\vec{g}_i$  se note encore, de façon parlante,  $\vec{g}(a_1, \dots, a_{i-1}, \cdot, a_{i+1}, \dots, a_p)$ . On peut étudier la dérivabilité de  $\vec{g}_i$  en  $a_i$ , si  $\vec{g}_i$  est dérivable en  $a_i$ ,  $\vec{g}'_i(a_i)$  s'appelle  $i^{\text{e}}$  dérivée partielle en  $(a_1, \dots, a_p)$  ou dérivée partielle par rapport à  $x_i$  en  $(a_1, \dots, a_p)$ , on la note :

$$\partial_i \vec{g}(a_1, \dots, a_p) \text{ ou } \frac{\partial \vec{g}}{\partial x_i}(a_1, \dots, a_p).$$

De façon quelque peu caricaturale la  $i^{\text{e}}$  dérivée partielle est la dérivée de  $\vec{g}$  « comme fonction de  $x_i$  ».

Regardons d'un peu plus près, avec les notations précédente, en fait :

$$\vec{g}_i(\dots) = \vec{g}(\dots)$$

Donc quand elle existe  $\partial_i \vec{g}(a_1, \dots, a_p)$  est en faite la dérivée de  $\vec{g}(\dots)$

RETOUR AU CAS GÉNÉRAL

L'espace vectoriel  $\mathbf{E}$  est muni d'une base  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ , notée  $\mathcal{B}$ .

**Définition 2.1.3.**—On appelle, sous réserve évidemment d'existence, pour  $i = 1, \dots, p$ ,  $i^{\text{e}}$  dérivée partielle de  $\vec{f}$  au point  $a$  dans la base  $\mathcal{B}$ , la quantité  $D_{\vec{e}_i}\vec{f}(a)$ . On la note  $\partial_i \vec{f}(a)$

Lorsque  $\mathbf{E} = \mathbf{R}^p$ , on parle souvent de dérivée partielle sans préciser la base, on sous-entend alors qu'il s'agit de la base canonique.

**Notations** — Soit  $i \in \{1, \dots, p\}$ . Notons  $(x_1, x_2, \dots, x_p)$  les coordonnées d'un vecteur générique de  $\mathbf{E}$  dans  $\mathcal{B}$ . En fait  $\partial_i \vec{f}(a)$  est la  $i^{\text{e}}$  dérivée partielle de la fonction définie au voisinage de  $a$  par

$$(x_1, \dots, x_p) \mapsto \vec{f}(x_1\vec{e}_1 + \dots + x_p\vec{e}_p)$$

ce qui alors justifie que l'on note fréquemment la  $i^{\text{e}}$  dérivée partielle de  $\vec{f}$  en  $a$ ,  $\frac{\partial \vec{f}}{\partial x_i}(a)$  plutôt que  $\partial_i \vec{f}(a)$ . En fait le programme nous autorise même à d'identifier abusivement  $\vec{f}$  et  $(x_1, \dots, x_p) \mapsto \vec{f}(x_1\vec{e}_1 + \dots + x_p\vec{e}_p)$ , nous n'userons pas de cette dernière liberté.

Cette notation, fort commune que nous venons de légitimer à l'instant et que des siècles d'histoire ont consacrée est néanmoins dangeureuse. Expliquons; soit l'application

$$f : \mathbf{R}^2 \rightarrow \mathbf{R}; (x_1, x_2) \mapsto x_1 x_2 e^{x_1 - x_2}.$$

Notons  $\vec{e}_1 := (1, 0)$  et  $\vec{e}_2 := (0, 1)$  et posons  $\vec{e}_1' := \vec{e}_1$ ,  $\vec{e}_2' := \vec{e}_1 + \vec{e}_2$ . On a immédiatement que  $(\vec{e}_1', \vec{e}_2')$  est une base de  $\mathbf{R}^2$ . Pour tout élément  $(x_1, x_2)$  de  $\mathbf{R}^2$ , nous noterons  $(x_1', x_2')$  ses coordonnées dans la base  $(\vec{e}_1', \vec{e}_2')$  de sorte que :

$$\begin{cases} x_1 = x_1' + x_2', \\ x_2 = x_2', \end{cases} \quad \begin{cases} x_1' = x_1 - x_2, \\ x_2' = x_2. \end{cases} \quad \text{Exprimons } f \text{ dans les nouvelles coordonnées :}$$

$$\tilde{f} : \mathbf{R}^2 \rightarrow \mathbf{R}; (x_1', x_2') \mapsto f(x_1' + x_2', x_2') = \dots\dots\dots$$

$$\frac{\partial \tilde{f}}{\partial x_1'}(x_1', x_2') =$$

$$\frac{\partial f}{\partial x_1}(x_1, x_2) =$$

$$\frac{\partial \tilde{f}}{\partial x_2'}(x_1', x_2') =$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) =$$

Il convient donc de voir les dérivées partielles comme des dérivées suivant des vecteurs plutôt que « par rapport à des variables », comme le favorise la notation  $\frac{\partial f}{\partial x_i}(a)$ .

**Remarque :** Il existe dans les sciences molles et hélas parfois en mathématiques d'autres usages abusifs et peu rigoureux de la notation  $\frac{\partial f}{\partial x_i}$ , nous nous limiterons au seul au programme :

$$\frac{\partial \text{nom d'une fonction}}{\partial \text{nom d'une variable}} \underbrace{\hspace{1cm}}_{2\text{mm}} (\text{nom d'un point de } U)$$

### Exercices d'application directe

- Existence et calcul des dérivées partielles en tout point  $(x, y)$  de  $\mathbf{R}^2$  des applications

(a)

$$f_1 : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto (\arctan(x + y)) \exp^{xy^2}$$

(b)

$$f_2 : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto \int_x^{y^2} f(t) dt,$$

où  $f$  est une application continue de  $\mathbf{R}$  dans  $\mathbf{R}$

2. Existence et calcul des dérivées partielles en  $(1, 1, 1)$  de l'application

$$f_3 : (\mathbf{R}_+^*)^3 \rightarrow \mathbf{R}; (x, y, z) \mapsto \frac{x^y}{z}.$$

3. Existence et calcul des dérivées partielles en tout point  $(x, y, z)$  de  $\mathbf{R}^3$  de l'application

$$f_4 : \mathbf{R}^3 \rightarrow \mathbf{R}^3; (x, y, z) \mapsto (x, \sin(xy), z + y + x^2).$$

4. Existence et calcul des dérivées partielles en tout point  $(m_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$  de  $\mathcal{M}_n(\mathbf{R})$  des applications suivantes :

(a)

$$\text{tr} : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathbf{R}; M \mapsto \text{tr} M$$

(b)

$$T : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathcal{M}_n(\mathbf{R}); M \mapsto {}^t M$$

5. Soit  $f$  une application de  $\mathbf{R}^2$  dans  $\mathbf{R}$  qui admet en tout point de  $\mathbf{R}^2$  deux dérivées partielles. Existence et calcul des dérivées partielles en tout point  $(x, y)$  de  $\mathbf{R}^2$  de l'application :

$$g : \mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto f(y, x)$$

On suppose  $f$  antisymétrique, que dire de ses dérivées partielles ?

*Solution de l'exercice —*

Revenons sur un point important de la définition 2.1.2. Dans cette définition et même dans toute la partie 2, nous avons supposé que  $\vec{f}$  est définie sur un *ouvert*. Cette précaution est essentielle, on a vu que c'est elle qui assure que l'application  $\vec{f}_{a,\vec{u}}$  est bien définie au voisinage de  $a$ . Dans l'exemple qui va suivre cette nécessité de travailler sur un ouvert va se manifester de façon criante et impérieuse. Prenons pour  $\mathbf{E}$ ,  $\mathbf{R}^2$  et  $U := \{(x_1, x_2) \in \mathbf{R}^2 | x_2 \geq x_1^2 \text{ et } x_2 \leq 2x_1^2\}$ . Ainsi

défini,  $U$  n'est pas ouvert.

Toutefois dans le cas où  $U$  est un rectangle fermé de  $\mathbf{R}^2$  de la forme  $[a, b] \times [c, d]$  (ou un pavé de  $\mathbf{R}^3$ ,  $[a, b] \times [c, d] \times [e, f]$ ), la formule (VII.1) permet encore de définir la notion de dérivée partielle (dans la base canonique). Ainsi, si  $U$  est le rectangle  $[a, b] \times [c, d]$ , sous réserve d'existence

$$\partial_1 \vec{f}(a, c) =$$

Pour terminer donnons une définition utile dans la suite.

**Définition 2.1.4.** — Soit  $\vec{u}$  un vecteur non nul de  $\mathbf{E}$ . On suppose que  $\vec{f}$  admet en tout point  $m$  de  $U$  une dérivée suivant  $\vec{u}$ . On appelle alors (application) dérivée suivant  $\vec{u}$ , l'application notée  $D_{\vec{u}}\vec{f}$  :

$$D_{\vec{u}}\vec{f} : U \rightarrow \mathbf{F}; m \mapsto D_{\vec{u}}\vec{f}(m).$$

En particulier si  $\mathcal{B}$  est une base de  $\mathbf{E}$ ,  $i$  un élément de  $\{1, \dots, p\}$  et si la  $i^{\text{e}}$  dérivée partielle de  $\vec{f}$  dans  $\mathcal{B}$  existe en tout point  $m$  de  $U$ . On appelle alors  $i^{\text{e}}$  (application) dérivée, dans  $\mathcal{B}$ , l'application notée  $D_i\vec{f}$  :

$$\partial_i \vec{f} : U \rightarrow \mathbf{F}; m \mapsto \partial_i \vec{f}(m).$$

Cette application peut aussi se noter  $\frac{\partial \vec{f}}{\partial x_i}$ , lorsque  $(x_1, \dots, x_p)$  désigne les coordonnées d'un élément générique de  $E$  dans  $\mathcal{B}$ .

## 2.2 Formule fondamentale du calcul différentiel

Nous allons aborder maintenant une formule fondamentale du calcul différentiel. Comme son nom l'indique, le calcul différentiel a pour objet de mesurer la différence entre les valeurs prises par une application en des points voisins.

Dans tout ce paragraphe nous considérerons encore une application  $\vec{f}$  de l'ouvert  $U$  dans  $\mathbf{F}$ . Nous désignerons par  $a$  un point de  $U$ . Notre objectif sera de mesurer  $\vec{f}(a + \vec{h}) - \vec{f}(a)$ , lorsque  $\vec{h}$  tend vers  $\vec{0}_{\mathbf{E}}$ . Nous allons commencer par une démarche purement heuristique, pour intuitionner le résultat que nous démontrerons par la suite.

### ÉTUDE HEURISTIQUE

Nous supposons provisoirement, pour simplifier, que  $\mathbf{E}$  est un espace vectoriel de dimension 3 dont  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$  est une base. Evaluons  $\vec{f}(a + \vec{h}) - \vec{f}(a)$ . Dans la base  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ ,  $\vec{h}$  se décompose en  $\vec{h} = h_1\vec{e}_1 + h_2\vec{e}_2 + h_3\vec{e}_3$ .

$$\vec{f}(a + \vec{h}) - \vec{f}(a) = \dots$$

Nous avons donc intuité le résultat suivant

**Proposition 2.2.1.** — Soit  $\vec{f}$  une application de l'ouvert  $U$  à valeurs dans  $\mathbf{F}$  et soit  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  une base de  $\mathbf{E}$ , notée  $\mathcal{B}$ . Nous supposons que :

- $\vec{f}$  admet en tout point  $m$  de  $U$ ,  $p$  dérivées partielles dans la base  $\mathcal{B}$ ,  $\partial_1 \vec{f}(m), \partial_2 \vec{f}(m), \dots, \partial_p \vec{f}(m)$  ;
- Les applications dérivées partielles dans la base  $\mathcal{B}$ ,  $\partial_1 \vec{f}, \partial_2 \vec{f}, \dots, \partial_p \vec{f}$  sont continues sur  $U$ .

Alors, pour tout point  $m$  de  $U$ ,

$$\vec{f}(m + \vec{h}) = \vec{f}(m) + \dots \quad (\text{VII.2})$$

*Preuve de la proposition 2.2.1.* — Soit  $a$  un point de  $U$ . On munit  $\mathbf{E}$  de la norme  $\|\cdot\|_{\mathcal{B},\infty}$ . Soit  $\varepsilon$  un élément de  $\mathbf{R}_+^*$ . On veut montrer qu'il existe ...

$U$  étant ouvert on peut choisir  $r$  élément de  $\mathbf{R}_+^*$  tel que  $B_0(a, r) \subset U$ . Donc pour tout vecteur  $\vec{h}$  de  $\mathbf{E}$  tel que  $\|\vec{h}\|_{\mathcal{B},\infty} < r$ ,  $\vec{f}(a + \vec{h})$  est défini. Prenons un tel vecteur  $\vec{h}$  et notons  $(h_1, h_2, \dots, h_p)$  ses coordonnées dans la base  $\mathcal{B}$ . Posons de plus,  $a_0 := a, a_i := a + \sum_{j=1}^i h_j \vec{e}_j$ , pour  $i = 1, 2, \dots, p$ . Ainsi  $a_p = \dots$

$$\vec{f}(a + \vec{h}) - \vec{f}(a) - \sum_{i=1}^p h_i \partial_i \vec{f}(a) = \dots$$

Notons que la somme précédente a bien un sens, en effet :

Posons  $\eta := \dots$  . Si  $\vec{h}$  a été choisi tel que  $\|\vec{h}\|_{\mathcal{B},\infty} < \eta$  alors

et donc

$$\|\vec{f}(a + \vec{h}) - f(a) - \sum_{i=1}^p h_i \partial_i \vec{f}(a)\|_{\mathbf{F}} \leq \varepsilon \|\vec{h}\|_{\mathcal{B},\infty}.$$

$\varepsilon$  étant quelconque, on a le résultat. □

Définissons maintenant les applications de classe  $\mathcal{C}^1$  sur l'ouvert  $U$ .

**Proposition, définition 2.2.2.**—Avec les notations précédentes, les trois propositions suivantes sont équivalentes

1. Pour tout vecteur  $\vec{u}$  de  $\mathbf{E}$  non nul, l'application dérivée directionnelle  $D_{\vec{u}}\vec{f}$  est définie sur  $U$  et continue.
2. Dans toute base de  $\mathbf{E}$ ,  $\vec{f}$  admet  $p$  applications dérivées partielles continues.
3. Il existe une base de  $\mathbf{E}$ , dans laquelle  $\vec{f}$  admet  $p$  applications dérivées partielles continues.

Si  $\vec{f}$  vérifie l'une de ces propositions (donc les trois) on dit que  $f$  est de classe  $\mathcal{C}^1$ . On dit encore pour des raisons que nous expliquerons plus tard que  $f$  est continûment différentiable.

*Preuve de la proposition 2.2.2.*— De toute évidence si 1. est vraie alors 2. est vraie et si 2. est vraie alors 3. l'est. Reste à prouver le gros morceau : si 3. est vraie alors 1. l'est aussi.

Supposons que 3. soit vraie.

On a au passage prouvé le résultat suivant que bientôt nous retrouverons sous des hypothèses plus faibles : *Pour  $\vec{f}$  de classe  $\mathcal{C}^1$ , et  $\vec{u}$  vecteur non nul de  $\mathbf{E}$ , en désignant par  $u_i$  la  $i^{\text{e}}$  coordonnée de  $\vec{u}$  dans  $\mathbf{E}$*

$$D_{\vec{u}}\vec{f} = \sum_{i=1}^p u_i \partial_i \vec{f}$$

**Remarque 2.2.3.** — Dans le cas où  $\mathbf{E}$  est  $\mathbf{R}$  et  $U$  un intervalle de  $\mathbf{R}$ , on retrouve la notion d'application de classe  $\mathcal{C}^1$  vue en MPSI, en effet dans ce cas la dérivabilité de  $\vec{f}$  en un point  $m$  équivaut à sa dérivabilité en  $m$  suivant le vecteur 1 et dans le cas de dérivabilité  $f'(m) = \partial_1 \vec{f}(m)$ .

On a vu en MPSI que les applications d'une variable réelle de classe  $\mathcal{C}^1$  sont continues, ce résultat demeure dans le cadre de ce chapitre :

**Proposition 2.2.4.** — *Une application  $\vec{f}$  de  $U$  dans  $\mathbf{F}$  de classe  $\mathcal{C}^1$  est continue.*

Rappelons que l'on a vu que, sans hypothèse de continuité des dérivées directionnelles, leur existence n'assure pas la continuité.

*Preuve de la proposition 2.2.4.* —

**Proposition 2.2.5.** — *Soit  $\vec{u}$  un vecteur non nul de  $\mathbf{E}$ . Soient  $\vec{f}$  et  $\vec{g}$  des applications de  $U$  dans  $\mathbf{F}$  admettant en un point  $m$  de  $U$  des dérivées suivant  $\vec{u}$ . Alors pour tout  $\lambda$  et tout  $\mu$  réels,  $\lambda\vec{f} + \mu\vec{g}$  admet une dérivée suivant  $\vec{u}$  en  $m$  et :*

$$D_{\vec{u}}(\lambda\vec{f} + \mu\vec{g})(m) = \lambda D_{\vec{u}}\vec{f}(m) + \mu D_{\vec{u}}\vec{g}(m).$$

La preuve de cette proposition, qui résulte directement de la linéarité de la dérivation d'une fonction d'une variable réelle, est laissée en exercice.

On en déduit immédiatement le corollaire suivant :



**Proposition 2.2.6.** — *L'ensemble des applications de  $U$  dans  $\mathbf{F}$  de classe  $\mathcal{C}^1$  est un sous-espace vectoriel  $\mathcal{F}(U, \mathbf{F})$  et de  $\mathcal{C}^0(U, \mathbf{F})$ . On le note  $\mathcal{C}^1(U, \mathbf{F})$ .*

*De plus pour tout élément non nul  $\vec{u}$  de  $\mathbf{E}$ , l'application*

$$D_{\vec{u}} : \mathcal{C}^1(U, \mathbf{F}) \rightarrow \mathcal{C}^0(U, \mathbf{F}) ; \vec{f} \mapsto D_{\vec{u}} \vec{f}$$

*est linéaire.*

**Proposition 2.2.7.** — *Soit  $(\vec{e}_1', \vec{e}_2', \dots, \vec{e}_n')$  une base de  $\mathbf{F}$  notée  $\mathcal{B}'$ . On désigne par  $f_i$  la  $i^{\text{e}}$  application composante de  $\vec{f}$  dans  $\mathcal{B}'$ .  $\vec{u}$  désignant toujours un vecteur de  $\mathbf{E}$  non nul et  $a$  un point de  $U$ .*

*L'application  $\vec{f}$  est dérivable en  $a$  suivant  $\vec{u}$ , si et seulement si  $f_1, f_2, \dots, f_n$  sont dérivables en  $a$  suivant  $\vec{u}$ . Si tel est le cas, alors :*

$$D_{\vec{u}} \vec{f}(a) = \sum_{i=1}^n D_{\vec{u}} f_i(a) \vec{e}_i'.$$

Nous laissons la preuve de cette proposition en exercice, preuve facile qui repose sur un résultat analogue pour la dérivation des fonctions d'une variable réelle.

Traitions des exemples.

**Exemple niais** — Prenons pour  $\mathbf{E}$ ,  $\mathbf{R}^2$ , pour  $\mathbf{F}$ ,  $\mathbf{R}^3$  pour  $U$ ,  $\mathbf{E}$  entier et

$$\vec{f} : \mathbf{R}^2 \rightarrow \mathbf{R}^3 ; (x, y) \mapsto (xe^{xy}, y \sin x, x \cos y).$$

Montrons que  $\vec{f}$  est de classe  $\mathcal{C}^1$ . Et calculons  $D_{(-1,2)} \vec{f}(3\pi, 1)$ .

**Exemple subtil mais classique** — Soit  $m$  un entier naturel non nul. soit l'application

$$\delta : \mathcal{M}_m(\mathbf{R}) \rightarrow \mathbf{R} ; M \mapsto \text{Det}(M)$$

Montrons que  $\delta$  est de classe  $\mathcal{C}^1$ .

### 2.3 Applications différentiables

On a vu dans le paragraphe précédent que  $\mathcal{C}^1(U, \mathbf{F})$  est un sous-espace vectoriel de  $\mathcal{C}^0(U, \mathbf{F})$ . Par ailleurs dans le cas des applications d'un intervalle  $I$  dans  $\mathbf{F}$ , il existe entre l'espace vectoriel des applications continues et celui des applications continûment dérivables un espace intermédiaire, celui des applications dérivables.

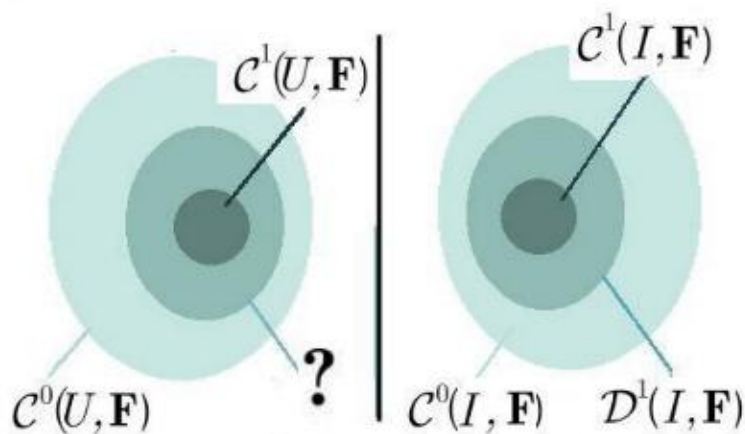


FIGURE VII.1 –

Existe-t-il un espace vectoriel intermédiaire entre  $\mathcal{C}^0(U, \mathbf{F})$  et  $\mathcal{C}^1(U, \mathbf{F})$ , qui serait l'analogue de  $\mathcal{D}^1(I, \mathbf{F})$  pour les applications de  $I$  dans  $\mathbf{F}$  ?

L'espace recherché ne saurait être celui des applications admettant en tout point des dérivées dans toutes les directions, en effet...

Pour continuer remarquons qu'une application dérivable est aussi une application admettant en tout point ...

...

Quel pourrait-être l'analogue de cette propriété, pour une application de  $U$  dans  $\mathbf{F}$ ? Pour le savoir prenons  $\vec{f}$  une application de  $U$  dans  $\mathbf{F}$  de classe  $\mathcal{C}^1$ . On a vu, formule (VII.2) que

$$\vec{f}(m + \vec{h}) = \vec{f}(m) + \dots$$

Remarquons que pour  $\vec{g}$  application de  $I$  dans  $\mathbf{F}$  dérivable en un point  $a$  de  $I$ ,

$$\vec{g}(a + h) = \vec{g}(a) + \dots$$

On voit donc maintenant comment définir l'analogue de dérivable (on dira différentiable) pour une application  $\vec{f}$  de  $U$  dans  $\mathbf{F}$  :

**Définition 2.3.1.** — Soit  $\vec{f}$  application de  $U$  dans  $\mathbf{F}$ . On dit que  $\vec{f}$  est différentiable en  $a$  si, par définition, il existe une application linéaire  $\vec{\ell}$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , telle que :

$$\vec{f}(a + \vec{h}) = \vec{f}(a) + \vec{\ell}(\vec{h}) + o(\|\vec{h}\|_{\mathbf{E}}) ; (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

On dira que  $\vec{f}$  est différentiable si, par définition elle est différentiable en tout point de  $U$ .

Nous avons vu en MPSI l'unicité du développement limité à un ordre donné, on a ici naturellement, avec les notations de 2.3.1. :

**Proposition 2.3.2.** — On suppose qu'il existe des applications linéaires  $\vec{\ell}_1$  et  $\vec{\ell}_2$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , telles que :

$$\vec{f}(a + \vec{h}) = \vec{f}(a) + \vec{\ell}_1(\vec{h}) + o(\|\vec{h}\|_{\mathbf{E}}) \text{ et } \vec{f}(a + \vec{h}) = \vec{f}(a) + \vec{\ell}_2(\vec{h}) + o(\|\vec{h}\|_{\mathbf{E}}) ; (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

Alors  $\vec{\ell}_1 = \vec{\ell}_2$ .

*Preuve de la proposition 2.3.2.* — Soit  $\vec{h}$  un élément de  $\mathbf{E}$ .

On a donc établi :

**Proposition, définition 2.3.3.** — Soit  $\vec{f}$  application de  $U$  dans  $\mathbf{F}$ .  $\vec{f}$  est différentiable en  $a$  point de  $U$  si et seulement si il existe une unique application linéaire  $\vec{\ell}$  de  $\mathbf{E}$  dans  $\mathbf{F}$ , telle que :

$$\vec{f}(a + \vec{h}) = \vec{f}(a) + \vec{\ell}(\vec{h}) + o(\|\vec{h}\|_{\mathbf{E}}) ; (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

L'application  $\vec{\ell}$  s'appelle différentielle de  $\vec{f}$  au point  $a$ . On la note  $d\vec{f}(a)$ . L'image d'un vecteur  $\vec{h}$  de  $\mathbf{E}$  par  $d\vec{f}(a)$  sera noté le plus souvent,  $d\vec{f}(a) \cdot \vec{h}$ .

Si  $\vec{f}$  est différentiable, on note  $d\vec{f}$  l'application de  $U$  dans  $\mathcal{L}(\mathbf{E}, \mathbf{F})$  qui à un point  $m$  de  $U$  associe  $d\vec{f}(m)$ . Cette application s'appelle (application) différentielle de  $\vec{f}$ .

On retiendra :

$$\vec{f}(a + \vec{h}) = \vec{f}(a) + d\vec{f}(a) \cdot \vec{h} + o(\|\vec{h}\|_{\mathbf{E}}) ; (\vec{h} \rightarrow \vec{0}_{\mathbf{E}})$$

$$d\vec{f} \in \dots$$

$$d\vec{f}(a) \in \dots$$

$$d\vec{f}(a) \cdot \vec{h} \in \dots$$

On va maintenant montrer que l'ensemble des applications différentiables joue bien le même rôle dans l'ensemble des applications de  $U$  dans  $\mathbf{F}$  que  $\mathcal{D}^1(I, \mathbf{F})$  pour l'ensemble des applications de  $I$  dans  $\mathbf{F}$ . Nous prouverons successivement que :

- L'ensemble des applications différentiables de  $U$  dans  $\mathbf{F}$  est un sous-espace vectoriel de  $\mathcal{F}(U, \mathbf{F})$  ;
- Il contient  $\mathcal{C}^1(I, \mathbf{F})$  ;
- Il est contenu dans  $\mathcal{C}^0(I, \mathbf{F})$  ;
- pour  $U = I$  on retrouve  $\mathcal{D}^1(I, \mathbf{F})$ .

La premier point est d'une démonstration aisée, nous laissons en exercice le résultat suivant :

**Proposition 2.3.4.** — Soient  $\vec{f}$  et  $\vec{g}$  des applications de  $U$  dans  $\mathbf{F}$  différentiables en un point  $m$  de  $U$ . Alors pour tout  $\lambda$  et tout  $\mu$  réels,  $\lambda\vec{f} + \mu\vec{g}$  est différentiable en  $m$  et :

$$d(\lambda\vec{f} + \mu\vec{g})(m) = \lambda d\vec{f}(m) + \mu d\vec{g}(m).$$

L'ensemble des applications de  $U$  dans  $\mathbf{F}$  différentiables est un sous-espace vectoriel de  $\mathcal{F}(U, \mathbf{F})$ . On le note dans ce cours  $\mathcal{D}^1(U, \mathbf{F})$ . De plus l'application

$$d ; \mathcal{D}^1(U, \mathbf{F}) \rightarrow \dots \quad ; f \mapsto df$$

est linéaire.

On a vu précédemment en utilisant (VII.2), le résultat suivant :

**Proposition 2.3.5.** —  $\mathcal{C}^1(U, \mathbf{F})$  est inclus dans  $\mathcal{D}^1(U, \mathbf{F})$ . De plus pour tout point  $m$  de  $U$  et tout vecteur  $\vec{h}$  de  $\mathbf{E}$ ,

$$d\vec{f}(m) \cdot (\vec{h}) = \dots ,$$

où...

Avec les notations précédemment utilisées :

**Proposition 2.3.6.** —

- Si  $\vec{f}$  est différentiable en un point  $a$  de  $U$ , alors elle est continue en  $a$ .
- $\mathcal{D}^1(U, \mathbf{F})$  est inclus dans  $\mathcal{C}^0(U, \mathbf{F})$ .

*Preuve de la proposition 2.3.6.* — Le second point résulte immédiatement du premier. Prouvons le premier point.

Soit  $\vec{g}$  une application d'un intervalle ouvert (non vide)  $I$  à valeurs dans  $\mathbf{F}$ , soit  $a$  un point de  $I$ .

- Supposons  $\vec{g}$  dérivable en  $a$ .

- Supposons  $\vec{g}$  différentiable en  $a$ .

$\vec{g}$  est donc différentiable en  $b$  si et seulement si elle est dérivable en  $b$ , et si tel est le cas

$\vec{g}'(b) = \quad ; d\vec{g}(b) \cdot h =$
---

L'ensemble des applications de  $I$  dans  $\mathbf{F}$  dérivables se confond avec celui des applications différentiables, ce qui assure la légitimité et l'innocuité de la notation choisie pour ce dernier espace.

Dans le cas d'une application de classe  $\mathcal{C}^1$  on a montré que la différentielle s'exprime au moyen des dérivées partielles. On va voir que ce résultat demeure dans le cas général d'applications différentiables.

**Proposition 2.3.7.** — *Soit  $\vec{f}$  une application de  $U$  dans  $\mathbf{F}$ , différentiable en un point  $a$  de  $U$ , alors  $\vec{f}$  admet en  $a$  des dérivées suivant tout vecteur  $\vec{u}$  de  $\mathbf{E}$  non nul et :*

$$D_{\vec{u}}\vec{f}(a) = d\vec{f}(a) \cdot \vec{u}. \quad (\text{VII.3})$$

En particulier,  $\vec{f}$  admet dans  $\mathcal{B}$ ,  $p$  dérivées partielles en  $a$ , et de plus pour tout vecteur  $\vec{h}$  de  $\mathbf{E}$

$$d\vec{f}(a) \cdot \vec{h} = \sum_{i=1}^p h_i \partial_i \vec{f}(a), \quad (\text{VII.4})$$

$h_i$  désignant, pour  $i = 1, 2, \dots, p$ , la  $i^{\text{e}}$  composante de  $\vec{h}$  dans  $\mathcal{B}$ .

**Remarque —**

1. La réciproque de 2.2.7. est fausse : on a vu qu'il existe des applications dérivables dans toutes les directions, et non continues donc non différentiables (exemple 1.2.2.)
2. Il existe des applications différentiables qui ne sont pas de classe  $\mathcal{C}^1$ , voir feuilles d'exercices.

## NOTATIONS

Nous allons maintenant parler des notations.

- *Premier cas : fonction à valeurs réelles*

$f$  désigne ici une application de  $U$  dans  $\mathbf{R}$  différentiable en un point  $a$  de  $U$ . On vient de voir que pour tout vecteur  $\vec{h}$  de  $\mathbf{E}$  :

$$df(a) \cdot \vec{h} = \sum_{i=1}^p h_i \partial_i f(a),$$

$h_i$  désignant, pour  $i = 1, 2, \dots, p$ , la  $i^{\text{e}}$  composante de  $\vec{h}$  dans  $\mathcal{B}$ . Notons  $(e_1^*, e_2^*, \dots, e_p^*)$  la base duale de la base  $\mathcal{B}$ . La formule précédente s'écrit :

$$df(a) = \sum_{i=1}^p \dots$$

Pour des raisons historiques, si l'on note  $(x_1, x_2, \dots, x_p)$  les coordonnées dans  $\mathcal{B}$  d'un vecteur générique  $\vec{x}$  de  $\mathbf{E}$ , alors on note, pour  $i = 1, 2, \dots, p$ ,  $e_i^*$ ,  $dx_i$  d'où :

$$df(a) = \sum_{i=1}^p \dots$$

Ou encore

$$df(a) = \sum_{i=1}^p \dots$$

Dans le cas où  $f$  est différentiable on a donc

$$df = \sum_{i=1}^p \dots$$

- *Second cas : fonction à valeurs vectorielles*

On revient à présent au cas général d'une application  $\vec{f}$  de  $U$  dans  $\mathbf{F}$ . On la suppose différentiable au point  $a$ .  
On note encore comme précédemment

$$d\vec{f}(a) = \sum_{i=1}^p \partial_i \vec{f}(a) dx_i$$

ou encore

$$d\vec{f}(a) = \sum_{i=1}^p \frac{\partial \vec{f}}{\partial x_i}(a) dx_i$$

Cette notation est abusive, en effet ...

Nous avons vu qu'une application de classe  $\mathcal{C}^1$ , est dite encore continûment différentiable, la proposition suivante donne la raison de cette terminologie.

**Proposition 2.3.8.** — Soit  $(\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n)$  une base de  $\mathbf{F}$  notée  $\mathcal{B}'$ . On désigne par  $f_i$  la  $i^{\text{e}}$  application composante de  $\vec{f}$  dans  $\mathcal{B}'$ .  $\vec{u}$  désignant toujours un vecteur de  $\mathbf{E}$  non nul et  $a$  un point de  $U$ .

L'application  $\vec{f}$  est différentiable en  $a$ , si et seulement si  $f_1, f_2, \dots, f_n$  sont différentiable en  $a$ . Si tel est le cas, alors :

$$d\vec{f}(a) = \sum_{i=1}^p df_i(a) \vec{e}'_i.$$

Nous laissons la preuve (facile) de ce résultat en exercice. Remarquons que ce résultat permet dans l'étude de la différentiabilité de se ramener à des applications à valeur réelles. C'est d'après 2.2.7., aussi le cas de l'étude des dérivées directionnelles.

**Proposition 2.3.9.** — Soit  $f$  une application de  $U$  dans  $\mathbf{F}$ . L'application  $\vec{f}$  est de classe  $\mathcal{C}^1$  si et seulement si elle est différentiable et sa différentielle  $df$  est continue.

*Preuve de la proposition 2.3.8.* —

**Exemples 2.3.10.** —

1. *Exemple élémentaire*

Soit l'application  $\vec{f} : \mathbf{R}^2 \rightarrow \mathbf{R}^3; (x, y) \mapsto (y, xy, ye^x)$ . Montrons que cette application est différentiable et

donnons l'expression de sa différentielle.

2. *Exemple au programme : application constante*

Soit  $\vec{f}$  une application constante de  $U$  dans  $\mathbf{F}$ , alors pour tout élément  $a$  de  $U$  et tout élément  $\vec{h}$  de  $\mathbf{E}$ ,

$$d\vec{f}(a) \cdot h = \dots$$

$$d\vec{f}(a) = \dots$$

$$d\vec{f} = \dots$$

On verra une réciproque faible :

*Si  $U$  est connexe par arcs et si la différentielle de  $\vec{f}$  est nulle alors  $f$  est constante*

3. *Exemple au programme : application linéaire*

Soit  $\vec{f}$  la restriction à  $U$  d'une application linéaire de  $\mathbf{E}$  dans  $\mathbf{F}$ ,  $\vec{\ell}$ . Montrons que cette application est différentiable et donnons l'expression de sa différentielle.

L'application  $\vec{f}$  est même de classe  $\mathcal{C}^1$ .

4. PRODUIT SCALAIRE

$\mathbf{E}$  est muni d'un produit scalaire  $\cdot$  dont la norme associée est notée  $\|\cdot\|$ . L'espace  $\mathbf{E} \times \mathbf{E}$  sera muni comme à l'accoutumée de la norme  $\|\cdot\|$ , définie par  $\|(\vec{x}, \vec{y})\|_\infty = \sup\{\|\vec{x}\|, \|\vec{y}\|\}$ . Montrer que l'application

$$p : \mathbf{E} \times \mathbf{E} \rightarrow \mathbf{R}; (\vec{x}, \vec{y}) \mapsto \langle \vec{x} | \vec{y} \rangle$$



est différentiable, puis de classe  $\mathcal{C}^1$ .

Généralisons

### 5. Application bilinéaire

Un résultat plus général encore fort est au programme et suivra. Soient  $\mathbf{E}_1$  et  $\mathbf{E}_2$  des espaces vectoriels de dimension finie. Pour  $i = 1, 2$   $\|\cdot\|_i$  désigne une norme sur  $\mathbf{E}_i$ . L'espace produit  $\mathbf{E}_1 \times \mathbf{E}_2$  sera muni comme à l'accoutumée de la norme  $\|\cdot\|_\infty$ , définie par  $\|(\vec{x}_1, \vec{x}_2)\|_\infty = \sup\{\|\vec{x}_1\|_1, \|\vec{x}_2\|_2\}$ . Soit  $\vec{b}$  une application bilinéaire de  $\mathbf{E}_1 \times \mathbf{E}_2$  dans  $\mathbf{F}$ . Montrons que  $\vec{b}$  est différentiable, puis de classe  $\mathcal{C}^1$ .

## INTERPRÉTATION DE LA DIFFÉRENTIELLE

Plaçons nous d'abord dans le cadre d'une application  $\vec{f}$  de  $U$  dans  $\mathbf{F}$  différentiable en un point  $a$  de  $U$ .

L'égalité  $f(a + \vec{h}) - \vec{f}(a) = d\vec{f}(a) \cdot \vec{h} + o(\vec{h})$  ( $\vec{h} \rightarrow \vec{0}_{\mathbf{E}}$ ) montre que au première ordre *l'accroissement de l'application* ( $f(a + \vec{h}) - \vec{f}(a)$ ) *est linéaire en l'accroissement  $\vec{h}$  de la variable*.

L'égalité précédente peut aussi s'écrire :

$$f(m) = \vec{f}(a) + d\vec{f}(a) \cdot (\overrightarrow{am}) + o(\|\overrightarrow{am}\|).$$

L'application

$$\mathbf{E} \rightarrow \mathbf{F}; m \mapsto \vec{f}(a) + d\vec{f}(a) \cdot (\overrightarrow{am})$$

est une application...  
quoi.

On l'appelle \_\_\_\_\_ . Voyons pour-

Particularisons notre étude.

— *Premier cas* :  $\mathbf{E} = \mathbf{F} = \mathbf{R}$  et  $U$  est un intervalle de  $\mathbf{R}$ . Rappelons des souvenirs de terminale.

$$f(x) = f(a) + \underbrace{(x-a)f'(a)}_{d f(a) \cdot (x-a)} + o(x-a).$$

Soit  $D$  la droite affine de  $\mathbf{R}^2$  d'équation :

$$D : z = f(a) + (x - a)f'(a).$$

pour tout  $x$  suffisamment voisin de  $a$ ,  $d(x, f(x); D) = o(x - a)$ . Cette propriété, on pourrait le montrer, *caractérise*  $D$  parmi les droites passant par  $(a, f(a))$ , en ce sens  $D$  est la droite passant par  $(a, f(a))$ , la plus proche du graphe de  $f$  on l'appelle tangente au graphe de  $f$  au point  $(a, f(a))$ .

— *Second cas* :  $\mathbf{E} = \mathbf{R}^2$ ,  $\mathbf{F} = \mathbf{R}$ . On prendra pour  $a$  le couple  $(a_1, a_2)$  et  $\|\cdot\|$  est une norme quelconque sur  $\mathbf{R}^2$ .

$$f((x, y)) = f(a_1, a_2) + df(a_1, a_2) \cdot ((x - a_1, y - a_2)) + o(\|(x - a_1, y - a_2)\|),$$

soit

$$f((x, y)) = f(a_1, a_2) + (x - a_1) \frac{\partial f}{\partial x}(a_1, a_2) + (y - a_2) \frac{\partial f}{\partial y}(a_1, a_2) + o(\|(x - a_1, y - a_2)\|).$$

Soit

Pour tout  $(x, y)$  suffisamment voisin de  $(a_1, a_2)$ ,

$$d((x, y), f(x, y); P) = o(\|(x - a_1, y - a_2)\|).$$

Cette propriété, on pourrait le montrer, *caractérise*  $P$  parmi les plans passant par  $(a_1, a_2, f(a_1, a_2))$ , en ce sens  $P$  est le plan passant par  $(a_1, a_2, f(a_1, a_2))$ , la plus proche du graphe de  $f$  on l'appelle plan tangent au graphe de  $f$  au point  $(a_1, a_2, f(a_1, a_2))$ .

On retient : *Equation du plan tangent en  $(a_1, a_2, f(a_1, a_2))$*

$$P : z = f(a_1, a_2) + (x - a_1) \frac{\partial f}{\partial x}(a_1, a_2) + (y - a_2) \frac{\partial f}{\partial y}(a_1, a_2)$$

Nous verrons dans le prochain chapitre de calcul différentiel une autre approche de la notion de plan tangent.

## 2.4 Opérations algébriques sur les applications différentiables

On a déjà vu que la différentiation des applications se comportait bien vis à vis des combinaisons linéaires cf. proposition 2.3.4.

Etudions maintenant la question de la composition d'applications différentiables. Outre les espaces vectoriels sur  $\mathbf{R}$  de dimensions finies, on considère un  $\mathbf{R}$ -espace vectoriel  $\mathbf{G}$  de dimension finie  $q$  sur  $\mathbf{R}$ .

**Proposition 2.4.1.** — Soit toujours  $\vec{f}$  application de  $U$  à valeurs dans  $\mathbf{F}$  et soit  $\vec{g}$  une application d'un ouvert  $V$  non vide de  $\mathbf{F}$  à valeurs dans  $\mathbf{G}$ . On suppose que  $\vec{f}(U) \subset V$ , ainsi dispose-t-on de l'application  $\vec{g} \circ \vec{f}$  notée  $\vec{\phi}$ , de  $U$  dans  $\mathbf{G}$ . On suppose que  $\vec{f}$  est différentiable en un point  $a$  de  $U$  et que  $\vec{g}$  est différentiable au point  $\vec{f}(a)$ . Alors  $\vec{\phi}$  est différentiable en  $a$  et :

$$\underbrace{d(\vec{g} \circ \vec{f})(a)}_{\dots} = \underbrace{d\vec{g}(\vec{f}(a))}_{\dots} \circ \underbrace{d\vec{f}(a)}_{\dots}.$$

*Preuve de la proposition 2.4.1.* — On posera  $b := \vec{f}(a)$  et  $\vec{\varphi} = \vec{g} \circ \vec{f}$ . Comme  $V$  est un ouvert de  $\mathbf{F}$ , il est possible de se donner un élément  $\rho$  de  $\mathbf{R}_+^*$  tel que  $B_o(b, \rho) \subset V$ . Par continuité de  $\vec{f}$  en  $a$  et comme  $U$  est un ouvert de  $\mathbf{E}$ , il est possible de se donner un élément  $r$  de  $\mathbf{R}_+^*$  tel que  $B_o(a, r) \subset U$  et  $\vec{f}(B_o(a, r)) \subset B_o(b, \rho)$ .

La différentiabilité de  $\vec{f}$  en  $a$  assure l'existence d'une application  $\vec{\varepsilon}$  définie sur  $B_o(a, r)$  à valeurs dans  $\mathbf{F}$  de limite nulle en  $\vec{0}_{\mathbf{E}}$  telle que pour tout  $\vec{h} \in B_o(a, r)$  :

$$\vec{f}(a + \vec{h}) = b + d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}). \quad (\text{VII.5})$$

Celle de  $\vec{g}$  en  $b$  assure l'existence d'une application  $\vec{\eta}$  définie sur  $B_o(b, \rho)$  à valeurs dans  $\mathbf{G}$  de limite nulle en  $\vec{0}_{\mathbf{F}}$  telle que pour tout  $\vec{k} \in B_o(b, \rho)$  :

$$\vec{g}(b + \vec{k}) = \vec{g}(b) + d\vec{g}(b) \cdot \vec{k} + \|\vec{k}\|_{\mathbf{F}} \vec{\eta}(\vec{k}). \quad (\text{VII.6})$$

Donc pour tout  $\vec{h} \in B_o(a, r)$ , par linéarité de  $d\vec{g}(b)$  :

$$\begin{aligned} \vec{\varphi}(a + \vec{h}) &= \vec{\varphi}(a) + d\vec{g}(b) \cdot (d\vec{f}(a) \cdot \vec{h}) + \\ &\quad \|\vec{h}\|_{\mathbf{E}} d\vec{g}(b) \cdot (\vec{\varepsilon}(\vec{h})) + \\ &\quad \left\| d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}) \right\|_{\mathbf{F}} \vec{\eta} \left( d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}) \right). \end{aligned} \quad (\text{VII.7})$$

• D'une part  $d\vec{g}(b) \cdot (\vec{\varepsilon}(\vec{h})) \xrightarrow{\vec{h} \rightarrow \vec{0}_{\mathbf{E}}} \vec{0}_{\mathbf{F}}$ , ce par continuité de l'application linéaire  $d\vec{g}(b)$  en  $\vec{0}_{\mathbf{E}}$  ( $\mathbf{E}$  est de dimension finie).  
Donc :

$$\|\vec{h}\|_{\mathbf{E}} d\vec{g}(b) \cdot (\vec{\varepsilon}(\vec{h})) = o(\|\vec{h}\|_{\mathbf{E}}) \quad (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

• D'autre part pour tout vecteur  $\vec{h}$  non nul de  $B_o(a, r)$ ,  $\left\| d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}) \right\|_{\mathbf{F}} = \|\vec{h}\|_{\mathbf{E}} \left\| d\vec{f}(a) \cdot \frac{\vec{h}}{\|\vec{h}\|_{\mathbf{E}}} + \vec{\varepsilon}(\vec{h}) \right\|_{\mathbf{F}}$ . Or  $d\vec{f}(a)$  application continue est bornée sur la sphère unité de  $\mathbf{E}$ , qui est compacte (puisque fermée bornée en dimension finie) et  $\vec{\varepsilon}$ , de limite nulle en  $\vec{0}_{\mathbf{E}}$ , est bornée au voisinage de ce point, donc

$$\left\| d\vec{f}(a) \cdot \frac{\vec{h}}{\|\vec{h}\|_{\mathbf{E}}} + \vec{\varepsilon}(\vec{h}) \right\|_{\mathbf{F}} = \vec{O}(1) \quad (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

Par ailleurs  $\vec{\eta} \left( d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}) \right)$  tend vers  $\vec{0}_{\mathbf{G}}$  lorsque  $\vec{h} \rightarrow \vec{0}_{\mathbf{E}}$ , grâce à la continuité en  $\vec{0}_{\mathbf{F}}$  de  $d\vec{f}(a)$ . Donc :

$$\left\| d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}) \right\|_{\mathbf{F}} \vec{\eta} \left( d\vec{f}(a) \cdot \vec{h} + \|\vec{h}\|_{\mathbf{E}} \vec{\varepsilon}(\vec{h}) \right) = o(\|\vec{h}\|_{\mathbf{E}}) \quad (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

De ces deux points il vient que :

$$\vec{\varphi}(a + \vec{h}) = \vec{\varphi}(a) + d\vec{g}(b) \circ d\vec{f}(a) \cdot \vec{h} + o(\|\vec{h}\|_{\mathbf{E}}) \quad (\vec{h} \rightarrow \vec{0}_{\mathbf{E}}).$$

Compte tenu de la linéarité de  $d\vec{g}(b) \circ d\vec{f}(a)$ , ceci dit que  $\vec{\varphi}$  est différentiable en  $a$  et  $d\vec{\varphi}(a) = d\vec{g}(b) \circ d\vec{f}(a)$ .  $\square$

Examinons ce résultat dans diverses situations particulières.

1. *Premier exemple*

$\mathbf{E} = \mathbf{F} = \mathbf{R}$ ,  $U$  et  $V$  sont des intervalles.

2. *Deuxième exemple*

$\mathbf{E} = \mathbf{R}$ ,  $U$  est un intervalle,  $\mathbf{F}$  et  $\mathbf{G}$  sont quelconques.

$$(\vec{g} \circ \vec{f})'(a) =$$

3. *Troisième exemple (sous-cas du précédent)*

$E = \mathbf{R}$ ,  $U$  est un intervalle,  $\mathbf{F} = \mathbf{R}^n$  et  $\mathbf{G}$  est quelconque.  $g$  différentiable en  $b$ , admet  $n$  dérivées partielles en  $b$  (dans la base canonique), notées  $\partial_1 \vec{g}(b), \partial_2 \vec{g}(b), \dots, \partial_n \vec{g}(b)$ , ou bien  $\frac{\partial \vec{g}}{\partial x_1}(b), \frac{\partial \vec{g}}{\partial x_2}(b), \dots, \frac{\partial \vec{g}}{\partial x_n}(b)$ . On note pour  $i = 1, 2, \dots, n$ ,  $f_i$  la  $i^{\text{e}}$  composante de  $\vec{f}$  dans la base canonique :

$$\vec{f} : U \rightarrow \mathbf{R}^n ; t \mapsto (f_1(t), f_2(t), \dots, f_n(t)).$$

$$\vec{g} \circ \vec{f} : U \rightarrow \mathbf{G} ; t \mapsto \vec{g}(f_1(t), f_2(t), \dots, f_n(t)).$$

$$(\vec{g} \circ \vec{f})'(a) =$$

ou, si l'on préfère

$$(\vec{g} \circ \vec{f})'(a) =$$

**Exercice** —

1. On se place dans le cadre de l'exemple 2.3.10–5 :  $\mathbf{E}_1$  et  $\mathbf{E}_2$  désignent des espaces vectoriels de dimension finie. Pour  $i = 1, 2$   $\vec{f}_i$  désigne une application d'un intervalle ouvert  $U$  dans  $\mathbf{E}_i$  et  $\vec{b}$  une application bilinéaire de  $\mathbf{E}_1 \times \mathbf{E}_2$  dans  $\mathbf{F}$ . On suppose  $\vec{f}_1$  et  $\vec{f}_2$  dérivable en un point  $a$  de  $U$ . Montrer que

$$\vec{\phi} : U \rightarrow \mathbf{G}; t \mapsto \vec{b}(\vec{f}_1(t), \vec{f}_2(t))$$

est dérivable, en utilisant 2.4.1. et donner sa dérivée en  $a$ . Où a-t-on déjà vu ce résultat ?

2.  $\mathbf{E}$  est muni d'un produit scalaire  $\cdot$  dont la norme associée est notée  $\|\cdot\|$ . Montrer que l'application

$$\phi : \mathbf{E} \setminus \{\vec{0}_{\mathbf{E}}\} \rightarrow \mathbf{R}; \vec{x} \mapsto \frac{1}{\|\vec{x}\|}$$

est différentiable, puis de classe  $\mathcal{C}^1$ .

**Solution** —

Donnons un résultat au programme plus fort que celui de la première question de l'exercice.

Par  $\mathbf{E}_1$  et  $\mathbf{E}_2$  on désigne des espaces vectoriels de dimensions finies. Pour  $i = 1, 2$ ,  $\vec{f}_i$  désigne une application de l'ouvert  $U$  de  $\mathbf{E}$  dans  $\mathbf{E}_i$  et  $\vec{b}$  une application bilinéaire de  $\mathbf{E}_1 \times \mathbf{E}_2$  dans  $\mathbf{F}$ . On suppose  $\vec{f}_1$  et  $\vec{f}_2$  différentiables en un point  $a$  de  $U$ . Montrons que

$$\vec{\phi} : U \rightarrow \mathbf{G}; m \mapsto \vec{b}(\vec{f}_1(m), \vec{f}_2(m))$$

est différentiable en  $a$  et donnons sa différentielle.

On a prouvé :

**Proposition 2.4.2.** — Soient  $\mathbf{E}_1$  et  $\mathbf{E}_2$  des espaces vectoriels de dimensions finies, pour  $i = 1, 2$ ,  $\vec{f}_i$  application de l'ouvert  $U$  de  $\mathbf{E}$  dans  $\mathbf{E}_i$ ,  $\vec{b}$  une application bilinéaire de  $\mathbf{E}_1 \times \mathbf{E}_2$  dans  $\mathbf{F}$  et

$$\vec{\phi} : U \rightarrow \mathbf{G}; m \mapsto \vec{b}(\vec{f}_1(m), \vec{f}_2(m)).$$

Alors  $\Phi$  est différentiable en  $a$  et pour tout vecteur  $\vec{h}$  de  $\mathbf{E}$ ,

$$d\vec{\Phi}(a) = \dots$$

**Exercice** — DIFFÉRENTIEL DE L'INVERSION — On munit  $\mathbf{R}^n$  de sa structure euclidienne canonique. Soit  $\omega$  un point de  $\mathbf{R}^n$  et

$$\Phi : \mathbf{R}^n \setminus \{\omega\} \rightarrow \mathbf{R}^n; m \mapsto \omega + \frac{1}{\|\vec{\omega m}\|^2} \vec{\omega m}$$

## 2.5 Matrice Jacobienne

Nous avons dans le paragraphe précédent pu noter que la composition de différentielles est parfois un peu lourd. Pour palier cette lourdeur technique, nous allons introduire un mode de calcul matriciel.

Par  $\vec{f}$  nous continuons à désigner une application de  $U$  dans  $\mathbf{F}$ ,  $\mathbf{E}$  sera toujours muni de la base  $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ , tandis que  $(\vec{e}_1', \vec{e}_2', \dots, \vec{e}_n')$  sera une base de  $\mathbf{F}$  que nous noterons  $\mathcal{B}'$ . Pour  $i = 1, 2, \dots, n$ , la  $i^{\text{e}}$  composante de  $\vec{f}$  dans  $\mathcal{B}'$  sera désignée par  $f_i$ , si elle existe, la  $j^{\text{e}}$  dérivée partielle de  $f_i$  dans  $\mathcal{B}'$  en un point  $m$  de  $U$  sera noté  $\frac{\partial f_i}{\partial x_j}(m)$ .

**Définition 2.5.1.** — On suppose  $\vec{f}$  différentiable en un point  $a$  de  $U$ . On appelle matrice jacobienne de  $\vec{f}$  au point  $a$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}'$ , la matrice, élément de  $\mathcal{M}_{n,p}(\mathbf{R})$ ,

$$\left( \frac{\partial f_i}{\partial x_j}(a) \right)_{\substack{i=1,\dots,n \\ j=1,\dots,p}}$$

On la note  $J_{\vec{f}}(a)$ .

Lorsque  $\mathbf{E} = \mathbf{R}^p$  et  $\mathbf{F} = \mathbf{R}^n$ , on parle de matrice jacobienne de  $f$  au point  $a$ , sans préciser les bases dont on équipe  $\mathbf{R}^p$  et  $\mathbf{R}^n$ , qui sont alors de façon sous-entendue les bases canoniques.

$$J_{\vec{f}}(a) = \begin{pmatrix} f_1 & \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_p}(a) \\ f_2 & \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_p}(a) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_n & \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \dots & \frac{\partial f_n}{\partial x_p}(a) \end{pmatrix}$$

L'utilité de la matrice jacobienne provient du résultat suivant.

**Proposition 2.5.2.** — Avec les notations de la définition précédente,  $J_{\vec{f}}(a)$  est la matrice de  $d\vec{f}(a)$ , dans les bases  $\mathcal{B}$  et  $\mathcal{B}'$  :

$$J_{\vec{f}}(a) = \text{mat}_{\mathcal{B}, \mathcal{B}'}(d\vec{f}(a)).$$

*Preuve de la proposition 2.5.2.* —

### Exemple — COORDONNÉES CYLINDRIQUES —

Soit l'application

$$\vec{\phi}_c : \mathbf{R}^3 \rightarrow \mathbf{R}^3; (r, \theta, \zeta) \mapsto (r \cos \theta, r \sin \theta, \zeta).$$

$$J_{\vec{\phi}_c}(r, \theta, \zeta) =$$

## UTILISATION PRATIQUE DE LA MATRICE JACOBIENNE

Comme dans la proposition 2.4.1., on considère  $\vec{f}$  notre application de  $U$  dans  $\mathbf{F}$  et  $\vec{g}$  une application d'un ouvert  $V$  non vide de  $\mathbf{F}$  à valeurs dans  $\mathbf{G}$ . On suppose que  $\vec{f}(U) \in V$ , ainsi dispose-t-on de l'application  $\vec{g} \circ \vec{f}$  de  $U$  dans  $\mathbf{G}$ . On suppose que  $\vec{f}$  est différentiable en un point  $a$  de  $U$  et que  $\vec{g}$  est différentiable au point  $\vec{f}(a)$ . On a vu dans cette proposition que  $\vec{g} \circ \vec{f}$  est différentiable en  $a$  et :

$$d(\vec{g} \circ \vec{f})(a) = d\vec{g}(\vec{f}(a)) \circ d\vec{f}(a).$$

On désigne par  $\mathcal{B}$ ,  $\mathcal{B}'$  et  $\mathcal{B}''$  des bases respectivement de  $\mathbf{E}$ ,  $\mathbf{F}$  et  $\mathbf{G}$ . On note  $J_{\vec{f}}(a)$  la matrice jacobienne de  $\vec{f}$  au point  $a$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}'$ ,  $J_{\vec{g}}(\vec{f}(a))$  la matrice jacobienne de  $\vec{g}$  au point  $\vec{f}(a)$  dans les bases  $\mathcal{B}'$  et  $\mathcal{B}''$  et  $J_{\vec{g} \circ \vec{f}}(a)$  la matrice jacobienne de  $\vec{g} \circ \vec{f}$  au point  $a$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}''$

L'égalité précédente se traduit matriciellement :

$$\boxed{J_{\vec{g} \circ \vec{f}}(a) = J_{\vec{g}}(\vec{f}(a)) J_{\vec{f}}(a)} \quad (\text{VII.8})$$

Utilisons cette égalité matricielle pour calculer des dérivées partielles.

Notation :

- $\vec{\phi} := \vec{g} \circ \vec{f}$
- Pour  $k = 1, 2, \dots, n$ ,  $f_k$  est la  $k^{\text{e}}$  composante de  $\vec{f}$  dans  $\mathcal{B}'$  ;
- pour  $i = 1, 2, \dots, q$ ,  $g_i$  est la  $i^{\text{e}}$  composante de  $\vec{g}$  dans  $\mathcal{B}''$  ;
- Pour  $i = 1, 2, \dots, q$ ,  $\phi_i$  est la  $i^{\text{e}}$  composante de  $\vec{\phi}$  dans  $\mathcal{B}''$  (notons que  $\phi_i = \dots$
- Pour  $k = 1, 2, \dots, n$ , le symbole  $\frac{\partial}{\partial y_k}$  désigne la dérivation d'une application définie sur  $V$  par rapport au  $k^{\text{e}}$  vecteur de  $\mathcal{B}'$  ; Pour  $j = 1, 2, \dots, p$ , le symbole  $\frac{\partial}{\partial x_j}$  désigne la dérivation d'une application définie sur  $U$  par rapport au  $j^{\text{e}}$  vecteur de  $\mathcal{B}$ .

Soit  $i \in \{1, 2, \dots, q\}$  et  $j \in \{1, 2, \dots, p\}$ , calculons  $\frac{\partial \phi_i}{\partial x_j}(a)$ . D'après (VII.8),

$$\frac{\partial \phi_i}{\partial x_j}(a) =$$

*Cas particulier :*

$\mathbf{F} = \mathbf{R}^n$  et  $\mathcal{B}'$  en est la base canonique. Alors

$$\phi_i : m \mapsto g_i(f_1(m), f_2(m), \dots, f_n(m)),$$

et

$$\frac{\partial \phi_i}{\partial x_j}(a) =$$

Voilà ce que certains nomment « règle de la chaîne », notons que cette formule découle aussi du troisième exemple après 2.4.1. où le rôle de  $t$  est tenu d'une certaine façon par la variable  $x_j$ .

Étudions sur un exemple la matrice jacobienne d'une bijection réciproque.



**Exemple** — COORDONNÉES POLAIRES —

Soit l'application

$$\vec{\phi}_p : \mathbf{R}_+^* \times ]-\pi, \pi[ \rightarrow \dots\dots\dots; (r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

c'est une bijection. On note  $\vec{\psi}_p$  la bijection réciproque  $\vec{\phi}_p$  est de classe  $\mathcal{C}^1$ , ce qui signifie, que l'application obtenue en étendant le domaine d'arrivée de  $\vec{\phi}_p$  à l'espace vectoriel  $\mathbf{R}^2$  l'est, de même dans la suite nous parlerons d'application de classe  $\mathcal{C}^1$  à valeurs dans une partie de  $\mathbf{R}^2$ . en effet.....

$\vec{\phi}_p$  induit une bijection de ..... sur  $\mathbf{R}_+^* \times \mathbf{R}$ , la bijection réciproque est l'application

$$\vec{\psi}_1 : \mathbf{R}_+^* \times \mathbf{R} \rightarrow \dots\dots\dots; (x, y) \mapsto \dots\dots\dots$$

L'application

$\vec{\psi}_1$  est de classe  $\mathcal{C}^1$ ,

On pourrait montrer, par exemple en l'explicitant, chose non totalement évidente, que de même  $\vec{\psi}_p$  est de classe  $\mathcal{C}^1$ .  
Déterminons la matrice jacobienne de  $\vec{\psi}_p$  en un point  $(x, y)$  de .....

## FONCTIONS À VALEURS RÉELLES

Dans toute cette partie,  $U$  est un ouvert non vide de  $\mathbf{E}$  et  $f$  une application de  $U$  dans  $\mathbf{R}$ .

### 3.1 Vecteur gradient

On suppose que  $\mathbf{E}$  est muni d'une structure euclidienne  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$ , on désignera par  $\| \cdot \|$  la norme euclidienne associée à cette structure et l'on se donne une base ORTHONORMÉE  $(\vec{e}_1, \dots, \vec{e}_p)$ , notée  $\mathcal{B}_o$ .

Supposons que  $f$  soit différentiable en un point  $a$  de  $U$ , on a vu que pour tout vecteurs  $\vec{h}$  de  $\mathbf{E}$ ,

$$df(a) \cdot \vec{h} = \sum \dots\dots\dots$$

où pour  $i = 1, \dots, p$ ,  $h_i$  désigne la  $i^{\text{e}}$  coordonnées de  $\vec{h}$  dans  $\mathcal{B}_o$ .

Donc  $df(a) \cdot \vec{h}$  est le produit scalaire de  $\vec{h}$  par le vecteur de  $\mathbf{E}$  dont le  $p$ -uplet de coordonnées dans  $\mathcal{B}_o$  est

$$(\dots\dots\dots)$$

Afin de donner plus de sens à cette constatation et plus de portée à la chose démontrons un résultat général sur la représentation des formes linéaires qui semble au programme.

Observons pour commencer que pour tout vecteur  $\vec{u}$  de  $\mathbf{E}$  l'application  $\langle \vec{u} | \cdot \rangle$ , c'est-à-dire l'application

$$\mathbf{E} \rightarrow \mathbf{R}; \vec{x} \mapsto \langle \vec{u} | \vec{x} \rangle,$$

est une forme linéaire, c'est une conséquence immédiate de.....

Réciproquement nous allons montrer que toute forme linéaire est de cette forme.

soit l'application

$$\Phi : \mathbf{E} \rightarrow \mathbf{E}^*, \vec{u} \mapsto \langle \vec{u} | \cdot \rangle$$

- L'application  $\Phi$  est linéaire...
  
- L'application  $\Phi$  est injective...
  
- L'application  $\Phi$  est un isomorphisme de  $\mathbf{E}$  sur  $\mathbf{E}^*$

On a prouvé :

**Proposition 3.1.1.** — *L'application*

$$\Phi : \mathbf{E} \rightarrow \mathbf{E}^*, \vec{u} \mapsto \langle \vec{u} | \cdot \rangle$$

*est un isomorphisme. Autrement dit pour toute forme linéaire  $\ell$  sur  $\mathbf{E}$  il existe un et un seul vecteur  $\vec{u}$  de  $\mathbf{E}$  tel que :*

$$\ell = \langle \vec{u} | \cdot \rangle.$$

Remarquons que le vecteur qui « représente » une forme linéaire par produit scalaire a été défini sans recourir à une base, il est déterminé uniquement par la structure euclidienne de  $\mathbf{E}$ .

Étudions à présent ses coordonnées dans une base orthonormée.

Soit  $\ell$  une forme linéaire sur notre espace euclidien  $\mathbf{E}$ , notons  $\vec{u} = \Phi^{-1}(\ell)$ .

Notons  $(e_1^*, \dots, e_p^*)$  la base duale de  $\mathcal{B}_o$ , ( $e_i^*$  est pour  $i = 1, \dots, p$ , la  $i^{\text{e}}$  forme coordonnée dans  $\mathcal{B}$ ).

On a le résultat suivant :

**Proposition 5.1.2.** — *Soit une forme linéaire  $\ell$ . Les coordonnées de  $\Phi^{-1}(\ell)$  dans une base orthonormée quelconque et celle de  $\ell$  dans la base duale sont les mêmes*

Revenons à notre problème initial en appliquant le présent résultat à la forme linéaire  $df(a)$ . Nous noterons la base duale de  $\mathcal{B}_o$ ,  $(dx_1, \dots, dx_p)$ , comme, nous l'avons dit, il est de tradition en calcul différentiel, lorsque  $(x_1, \dots, x_n)$  est la notation pour les coordonnées dans  $\mathcal{B}_o$  d'un vecteur générique de  $\mathbf{E}$ . Alors :

**Proposition définition 5.1.3.** — *Si  $f$  est différentiable au point  $a$  de  $U$ , alors il existe un et un seul vecteur  $\vec{u}$  de  $\mathbf{E}$  tel que pour tout élément  $\vec{h}$  de  $\mathbf{E}$ ,*

$$df(a) \cdot \vec{h} = \langle \vec{u} | \vec{h} \rangle.$$

Ce vecteur  $\vec{u}$  est appelé gradient de  $f$  au point  $a$  et se note  $\vec{\nabla} f(a)$ .

Dans  $\mathcal{B}_o$ , base ORTHONORMÉE quelconque, les coordonnées de  $\vec{\nabla} f(a)$  sont les dérivées partielles de  $f$  en  $a$  dans  $\mathcal{B}_o$  :

$$\vec{\nabla} f(a) = \sum_{i=1}^p \partial_i f(a) \vec{e}_i$$

*Preuve de la proposition 5.1.2.* —

Résumons :

Pour tout  $\vec{h} \in \mathbf{E}$ ,

$$df(a) = \langle \vec{\nabla} f(a) | \cdot \rangle$$

$$df(a) \cdot \vec{h} = D_{\vec{h}} f(a) = \langle \vec{\nabla} f(a) | \vec{h} \rangle$$

$$\begin{array}{|l} df(a) = \sum_{i=1}^p \partial_i f(a) dx_i \\ \vec{\nabla} f(a) = \sum_{i=1}^p \partial_i f(a) \vec{e}_i \end{array} \quad \text{ou encore} \quad \begin{array}{|l} df(a) = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) dx_i \\ \vec{\nabla} f(a) = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) \vec{e}_i \end{array}$$

**Exercices d'application** —

1. Soit l'application

$$N : \mathbf{E} \rightarrow \mathbf{R}; \vec{x} \mapsto \|\vec{x}\|^2.$$

Montrer que pour tout  $\vec{x} \in \mathbf{R}$ ,  $N$  est différentiable en  $\vec{x}$  et calculer  $\vec{\nabla} N(x)$

2. Soit  $A \in \mathcal{S}_p(\mathbf{R})$ . On munit  $\mathcal{M}_{p,1}(\mathbf{R})$  du produit scalaire canonique  $\langle \cdot | \cdot \rangle$ . Notons que pour tout couple  $(X, Y)$  d'éléments de  $\mathcal{M}_{p,1}(\mathbf{R})^2$ ,  $\langle X | Y \rangle = {}^t X Y$ . Soit l'application

$$Q : \mathcal{M}_{p,1}(\mathbf{R}) \rightarrow \mathbf{R}; X \mapsto {}^t X A X.$$

Montrer que pour tout  $X \in \mathcal{M}_{p,1}(\mathbf{R})$ ,  $Q$  est différentiable en  $X$  et calculer  $\nabla Q(X)^1$ .

*Solution* —

### Exemple géométrique — GRADIENT EN COORDONNÉES POLAIRES —

Il est utile en physique lorsque l'on étudie des système ayant une symétrie de révolution d'étudier l'expression du gradient en coordonnées polaires. On considère comme espace euclidien  $\mathbf{R}^2$  muni de sa structure euclidienne canonique. Soit  $f$  une application de  $\mathbf{R}^2 - (0, 0)$  dans  $\mathbf{R}$  de classe  $\mathcal{C}^1$ . Considérons l'application  $\tilde{f}$ , « expression de  $f$  en coordonnées polaires »,

$$\tilde{f} : \mathbf{R}_+^* \times \mathbf{R} : (r, \theta) \mapsto f(r \cos \theta, r \sin \theta).$$

Pour tout réel  $\theta$  on pose  $\vec{u}(\theta) := (\cos \theta, \sin \theta)$ ,  $\vec{v}(\theta) := (-\sin \theta, \cos \theta)$ . On note immédiatement que pour tout réel  $\theta$ , la famille  $\mathcal{B}_\theta = (\vec{u}(\theta), \vec{v}(\theta))$  est une... . Notons également que si l'on note  $\vec{u}$  (rep.  $\vec{v}$ ) l'application qui à un réel  $\theta$  associe  $\vec{u}(\theta)$  (res.  $\vec{v}(\theta)$ ), alors ces applications sont dérivables et

$$\vec{u}'(\theta) = \quad \vec{v}'(\theta) =$$

$\mathcal{B}_\theta$  est connu sous le nom de *base mobile*,  $\vec{u}(\theta)$  le vecteur *radial* de  $\mathcal{B}_\theta$ , et  $\vec{v}(\theta)$  le vecteur *orthoradial*.

Soit  $(r, \theta) \in \mathbf{R}_+^* \times \mathbf{R}$ . Nous nous proposons de donner les coordonnées de  $\overrightarrow{\text{grad}}(f)(r \cos \theta, r \sin \theta)$  dans la base  $\mathcal{B}_\theta$ , en

1. On ne met pas ici de flèche sur  $\nabla$ , car les vecteurs de  $\mathcal{M}_{p,1}(\mathbf{R})$  n'en portent pas.

fonction des dérivées partielles de  $\tilde{f}$  au point  $(r, \theta)$ .

D'où la formule :

$$\vec{\nabla}(f)(r \cos \theta, r \sin \theta) = \vec{u}(\theta) + \vec{v}(\theta)$$

#### PREMIÈRE INTERPRÉTATION GÉOMÉTRIQUE DU GRADIENT

Supposons l'application  $f$  différentiable. Soit  $a$  un point de  $U$ , non critique pour  $f$ .

Le caractère différentiable de  $f$  en  $a$  dit :

$$f(a+h) - f(a) = df(a) \cdot h + o(\|\vec{h}\|), (\vec{h} \rightarrow \vec{0}_{\mathbf{E}})$$

$df(a) \cdot h$  représente donc au première ordre près l'augmentation de  $f$  entre  $a$  et  $a + \vec{h}$ . Etudions ce terme. Il vaut encore  $D_{\vec{h}}f(a) = \langle \vec{\nabla}f(a) | \vec{h} \rangle$ . Prenons  $\vec{h}$  unitaire....

Donc  $D_{\vec{h}}f(a)$  est :

- maximum si et seulement si....
- minimum si et seulement si....
- nul si et seulement si....

Intuitivement, d'après le premier point, le vecteur gradient indique...

Examinons le dernier point dans le cas particulier où  $\mathbf{E} = \mathbf{R}^2$ . Soit  $k$  un réel. Notons  $\Lambda_k$  l'ensemble de niveau de  $f$ ,

$\{(x, y) \in U \mid f(x, y) = k\}$ , en général cet ensemble est une courbe...

Nous donnerons une preuve rigoureuse à cette interprétation dans le second chapitre de calcul différentiel.

## 3.2 Extremums locaux

Commençons par une définition

**Définition 3.3.1.** — Soient  $D$  une partie de  $\mathbf{E}$  et  $g$  une application de  $D$  dans  $\mathbf{R}$ , soit enfin  $a$  un point de  $D$ . On dit que  $g$  admet en  $a$  un maximum (resp. minimum) local en  $a$  si, par définition il existe un voisinage  $V$  de  $a$  tel que pour tout  $x \in V \cap D$ ,  $g(x) \leq g(a)$  (resp.  $g(x) \geq g(a)$ ).

On dit que  $g$  admet en  $a$  un extrémum local si, par définition il admet en  $a$  un maximum local ou un minimum local.

On dit que  $g$  admet en  $a$  un maximum (resp. minimum) local strict en  $a$  si, par définition il existe un voisinage  $V$  de  $a$  tel que pour tout élément  $x$  de  $V \cap D$ , distinct de  $a$ ,  $g(x) < g(a)$  (resp.  $g(x) > g(a)$ ).

**Proposition 3.3.2.** — Soit  $f$  une application d'un ouvert  $U$  de  $\mathbf{E}$  à valeurs réelles, de classe  $\mathcal{C}^1$ . Si  $f$  admet en un point  $a$  de  $U$  un extrémum local, alors  $df(a) = 0_{\mathbf{E}^*}$ .

*Preuve de la proposition 3.3.2.* — Supposons que  $f$  admette un extremum en  $a$ . Quitte à changer  $f$  en  $-f$ , ce qui ne change pas l'annulation de la différentielle, on suppose que  $f$  admet en  $a$  un maximum.

Un point  $a$  de  $U$  en lequel s'annule  $df(a)$  est appelé *point critique de  $f$* .

**Mise en garde** — Dans ce théorème deux points sont essentielles :  $f$  est défini sur un ouvert, l'annulation de la différentielle en  $a$  est une condition nécessaire à ce que  $f$  ait en  $a$  un extremum, mais n'est pas une condition suffisante. Illustrons par des contre exemple ces deux points.

- Soit  $G$  une application d'un ouvert  $U$  de  $\mathbf{E}$  à valeurs réelles différentiable. On considère  $D$  une partie de  $U$  non ouverte et  $g$  la restriction de  $G$  à  $D$ .  
L'application  $g$  peut admettre en un point  $a$  de  $U$  sans que  $dG(a)$  soit nulle. Exemple  
— Commençons par le cas simple où  $\mathbf{E}$  est  $\mathbf{R}$ .

— Donnons à présent un exemple dans le cas  $\mathbf{E} = \mathbf{R}^2$ .

- Répétons le, l'annulation de la différentielle en  $a$  n'est pas suffisante pour avoir un extremum local en  $a$ .  
Exemples.
  - Cas où  $\mathbf{E}$  est  $\mathbf{R}$ .
  - Cas où  $\mathbf{E} = \mathbf{R}^2$ .

## PRATIQUE —

La recherche des extremums locaux, ne semble plus un objectif affirmé du programme qui préfère appliquer la proposition précédente à la recherche des extremums globaux. Ceci dit si  $f$  atteint en un point  $a$  sa borne supérieure (resp. inférieure) alors elle admet en  $a$  un maximum (resp. minimum) local, et donc la recherche des extremums globaux passe plus ou moins par l'étude des extremums locaux. C'est parti.

**Exemple 1** — Soit l'application  $F : \Delta \rightarrow \mathbf{R}; z \mapsto |\cos z|$ , où  $\Delta = \{z \in \mathbf{C} \mid |z| \leq 1\}$ . où par définition,  $\cos(z) = \frac{e^{iz} + e^{-iz}}{2}$ .

Pour nous ramener à une fonctions de  $\mathbf{R}^2$  dans  $\mathbf{R}$  on pose

$$f : D \rightarrow \mathbf{R}; (x, y) \mapsto \dots$$

où  $D = \dots$

$F$  admet un extremum local (global) en un point  $z$  de  $\Delta$ , si et seulement si  $f$  admet un extremum local (global) de même nature en .....

Après calcul, pour tout élément  $(x, y)$  de  $D$ ,  $f(x, y) = \dots$

$D$  étant compact et  $f$  continue (en vertu des théorèmes de transfère)  $f$  atteint sa borne supérieure (resp. inférieure) en un point de  $D$ , point en lequel  $f$  admet un maximum (resp. minimum) local. Recherchons les point en lesquels  $f$  admet un extremum local.



- *Etude sur  $\overset{\circ}{D}$*

- *Etude sur la frontière de  $D$*

Concluons :

**Exemple 2** —

Soit  $a$  un réel strictement positif, et  $f$  l'application,

$$g : ]0, +\infty[^2 \rightarrow \mathbf{R}; (x, y) \mapsto \frac{a}{x} + \frac{a}{y} + \frac{xy}{a^2}.$$

Montrer que  $f$  atteint sa borne inférieure en un unique point à déterminer.

A la différence du premier exemple, la fonction  $g$  n'est pas définie sur un compact et *a priori* la borne inférieure peut

ne pas être atteinte (l'existence de la borne inférieure elle, est acquise puisque.....

## DÉRIVÉES D'ORDRES SUPÉRIEURS

### 4.1 Fonctions de classe $\mathcal{C}^k$

Dans toute ce paragraphe  $U$  est un **ouvert** de  $\mathbf{E}$  non vide, et  $\vec{f}$  désigne une application de  $U$  dans  $\mathbf{F}$ .

Commençons par des définitions.

**Définition 4.1.1.** —

- Soient  $\vec{u}$  et  $\vec{v}$  des vecteurs non nuls de  $\mathbf{E}$ . On suppose que l'application dérivée directionnelle  $D_{\vec{u}}\vec{f}$  est définie au voisinage d'un point  $a$  de  $U$ . Alors si  $D_{\vec{v}}(D_{\vec{u}}\vec{f})(a)$  existe on dit que  $\vec{f}$  est deux fois dérivable en  $a$ , selon  $\vec{u}$  puis  $\vec{v}$ .  $D_{\vec{v}}(D_{\vec{u}}\vec{f})(a)$  s'appelle la dérivée d'ordre 2 de  $\vec{f}$  selon  $\vec{u}$  puis  $\vec{v}$  en  $a$ , on la note

$$D_{\vec{v},\vec{u}}^2\vec{f}(a).$$

Lorsque pour tout point  $m$  de  $U$ ,  $D_{\vec{v},\vec{u}}^2\vec{f}(m)$  est définie l'application

$$D_{\vec{v},\vec{u}}^2\vec{f} : U \rightarrow \cdots \quad ; m \mapsto D_{\vec{v},\vec{u}}^2\vec{f}(m)$$

est appelée application dérivée seconde de  $f$  selon  $\vec{u}$  puis  $\vec{v}$ .

- Plus généralement, soient un entier  $k \geq 2$  et  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$  des vecteurs non nuls de  $\mathbf{E}$ . On suppose que l'application dérivée directionnelle  $D_{\vec{u}_{k-1}}(D_{\vec{u}_{k-2}} \cdots (D_{\vec{u}_2}(D_{\vec{u}_1}\vec{f})))$  est définie au voisinage d'un point  $a$  de  $U$ . Alors si  $D_{\vec{u}_k}(D_{\vec{u}_{k-1}}(D_{\vec{u}_{k-2}} \cdots (D_{\vec{u}_2}(D_{\vec{u}_1}\vec{f}))))(a)$  existe on dit que  $\vec{f}$  est  $k$  fois dérivable en  $a$ , selon  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_{k-1}$  puis  $\vec{u}_k$ .  $D_{\vec{u}_k}(D_{\vec{u}_{k-1}}(D_{\vec{u}_{k-2}} \cdots (D_{\vec{u}_2}(D_{\vec{u}_1}\vec{f}))))(a)$  s'appelle la dérivée d'ordre  $k$  de  $\vec{f}$  selon  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$  en  $a$ , on la note

$$D_{\vec{u}_k, \vec{u}_{k-1}, \dots, \vec{u}_2, \vec{u}_1}^k\vec{f}(a).$$

Lorsque pour tout point  $m$  de  $U$ ,  $D_{\vec{u}_k, \vec{u}_{k-1}, \dots, \vec{u}_2, \vec{u}_1}^k\vec{f}(m)$  est définie l'application

$$D_{\vec{u}_k, \vec{u}_{k-1}, \dots, \vec{u}_2, \vec{u}_1}^k\vec{f} : U \rightarrow \cdots \quad ; m \mapsto D_{\vec{u}_k, \vec{u}_{k-1}, \dots, \vec{u}_2, \vec{u}_1}^k\vec{f}(m)$$

est appelée *application dérivée d'ordre  $k$  de  $f$  selon  $\vec{u}_1, \vec{u}_2, \dots$ , puis  $\vec{u}_{k-1}$ .*

Envisageons le cas particulier des dérivées d'ordre  $k$  suivant les vecteurs d'une base.

On suppose que  $\mathbf{E}$  est muni d'une base  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  notée  $\mathcal{B}$ , dans laquelle les coordonnées d'un vecteur générique  $\vec{x}$  seront notées dans l'ordre  $x_1, x_2, \dots, x_p$ .

**Définition 4.1.2.** — Soit un entier  $k \geq 2$   $i_1, i_2, \dots, i_k$  des éléments de  $\{1, 2, \dots, p\}$ . Si  $D_{\vec{e}_{i_k}, \vec{e}_{i_{k-1}} \dots \vec{e}_{i_2}, \vec{e}_{i_1}}^k \vec{f}(a)$  existe, on dit que c'est une dérivée partielle de  $\vec{f}$  au point  $a$  d'ordre  $k$ , on la note

$$\partial_{i_k, i_{k-1} \dots i_2, i_1}^k \vec{f}(a),$$

ou encore

$$\frac{\partial^k \vec{f}}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_2} \partial x_{i_1}} (a).$$

Si pour tout élément  $m$  de  $U$  ;  $\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_2} \partial x_{i_1}} (m)$  existe, l'application

$$\frac{\partial^k \vec{f}}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_2} \partial x_{i_1}} : U \rightarrow \mathbf{F}; m \mapsto \frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_2} \partial x_{i_1}} (m)$$

est dite (application) *dérivée partielle de  $\vec{f}$  d'ordre  $k$ .*

Notons que  $\vec{f}$  admet en un point au plus  $\dots$  dérivées partielles d'ordre  $k$ .

**Exemple** — Ici  $\mathbf{E} = \mathbf{R}^2$ ,  $\mathbf{F} = \mathbf{R}^2$  et

$$\vec{f} : \mathbf{R}^2 \mapsto \mathbf{R}^2; (x, y) \mapsto (x^2 \exp(-x) \cos(y), x^3 \exp(x) \sin(y)).$$

Montrons l'existence et donnons l'expression de l'application  $\frac{\partial^3 f}{\partial x \partial y \partial x}$ .

A présent, montrons l'existence et donnons la valeur de  $D_{(1,2)(3,4)}^2 \vec{f}$ .

**Définition 4.1.3.** — soit un entier naturel  $k \geq 1$ . On dit que  $f$  est de classe  $\mathcal{C}^k$  si, par définition, il existe une base  $\mathcal{B}$  telle que pour tout  $k$ -uplet  $(i_1, i_2, \dots, i_k)$  d'éléments de  $\{1, 2, \dots, p\}$  l'application  $\partial_{i_k, i_{k-1} \dots i_2, i_1}^k \vec{f}$  est définie sur  $U$

et continue. Autrement dit  $f$  est de classe  $\mathcal{C}^k$  si elle admet dans une base,  $p^k$  dérivées partielles d'ordre  $k$  continues.

**Remarques 4.1.4.** —

- Dans le cas  $k = 1$  on retrouve la définition déjà donnée de la classe  $\mathcal{C}^1$  ;
- Si  $\vec{f}$  est de classe  $\mathcal{C}^k$ , alors, en particulier, toute application dérivée partielle d'ordre inférieur ou égal à  $k$  est définie sur  $U$ .

Dans le même ordre d'idées on a :

**Proposition 4.1.5.** — Soit  $k$  un entier naturel non nul. Si  $\vec{f} \in \mathcal{C}^k$  alors  $\vec{f} \in \mathcal{C}^{k-1}$ .

*Preuve de 4.1.5.* —

- Pour  $k = 1$ , nous avons vu que la proposition 4.1.5. est vraie (cf. 2.2.6.).
- Supposons que  $k \geq 2$  et que  $\vec{f}$  soit de classe  $\mathcal{C}^k$ .

Soit  $(i_1, i_2, \dots, i_{k-1})$  un  $k - 1$ -uplet d'éléments de  $\{1, 2, \dots, p\}$ . Comme  $\vec{f}$  est de classe  $\mathcal{C}^k$ , l'application  $\partial_{i_{k-1} \dots i_2, i_1}^{k-1} \vec{f}$  est définie sur  $U$  (cf. remarque 4.1.4.2) et de plus pour  $i = 1, 2, \dots, p$ ,  $\partial_{i, i_{k-1} \dots i_2, i_1}^k \vec{f} = \partial_i \left( \partial_{i_{k-1} \dots i_2, i_1}^{k-1} \vec{f} \right)$ , est définie sur  $U$  et est continue. Donc  $\partial_{i_{k-1} \dots i_2, i_1}^{k-1} \vec{f}$  est de classe  $\mathcal{C}^1$ , donc d'après 2.2.6., continue.

Le  $k - 1$ -uplet étant quelconque, on a montré que  $\vec{f}$  est de classe  $\mathcal{C}^{k-1}$ .

D'où le résultat. □

Donnons un résultat qui généralise 2.2.2.

**Proposition 4.1.6.** — Soit  $k$  un entier naturel non nul. L'application  $f$  est de classe  $\mathcal{C}^k$  si et seulement si pour tout  $k$ -uplet de vecteurs non nuls de  $\mathbf{E}$   $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ , l'application  $D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k \vec{f}$  est définie sur  $U$  et continue.

*Preuve de la proposition 4.1.6.* —

•

- La réciproque du premier point est immédiate.

**Proposition 4.1.6.** — Soient  $k$  un entier naturel non nul,  $\vec{f}$  et  $\vec{g}$  des applications de  $U$  dans  $\mathbf{F}$  et  $\lambda$  et  $\mu$  des réels,  $a$  un point de  $U$ . Pour tout  $k$ -uplet de vecteurs non nuls de  $\mathbf{E}$   $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ , si  $D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k \vec{f}(a)$  et  $D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k \vec{g}(a)$  existent, alors  $D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k (\lambda \vec{f} + \mu \vec{g})(a)$  existe et :

$$D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k (\lambda \vec{f}(a) + \mu \vec{g}) = \lambda D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k \vec{f}(a) + \mu D_{\vec{u}_k, \vec{u}_{k-1} \dots \vec{u}_2, \vec{u}_1}^k \vec{g}(a).$$

La preuve de ce résultat, un peu fastidieuse mais très simple, repose sur 2.2.5. Elle est laissée en exercice.

On en déduit immédiatement le résultat suivant qui généralise 2.2.6. :

**Corollaire 4.1.7.** — Soit  $k$  un entier naturel non nul. L'ensemble des applications de  $U$  dans  $F$  de classe  $\mathcal{C}^k$  est un sous-espace vectoriel de  $\mathcal{F}(U, \mathbf{F})$  et de  $\mathcal{C}^0(U, \mathbf{F})$ . On le note  $\mathcal{C}^k(U, \mathbf{F})$ .

de plus pour tout vecteur  $\vec{u}$  de  $\mathbf{E}$  non nul, l'application

$$D_{\vec{u}} : \mathcal{C}^k(U, \mathbf{F}) \rightarrow \mathcal{C}^{k-1}(U, \mathbf{F}); \vec{f} \mapsto D_{\vec{u}} \vec{f}$$

est linéaire.

Dans le cas où  $\mathbf{F}$  est  $\mathbf{R}$  ou  $\mathbf{C}$ , vu comme un  $\mathbf{R}$ -espace vectoriel de dimension 2, on a un résultat plus fort :

**Proposition 4.1.8.** — Soit  $k$  un entier naturel non nul.  $\mathcal{C}^k(U, \mathbf{R})$  (resp.  $\mathcal{C}^k(U, \mathbf{C})$ ) est une sous-algèbre de  $\mathcal{F}(U, \mathbf{R})$  (resp.  $\mathcal{F}(U, \mathbf{C})$ ).

**Proposition 4.1.9.** — Soit  $k$  un entier naturel non nul. Soit toujours  $\vec{f}$  une application de  $U$  à valeur dans  $\mathbf{F}$  et soit  $\vec{g}$  une application d'un ouvert  $V$  non vide de  $\mathbf{F}$  à valeurs dans un espace vectoriel  $\mathbf{G}$  de dimension finie. On suppose que  $\vec{f}(U) \in V$ , ainsi dispose-t-on de l'application  $\vec{g} \circ \vec{f}$  de  $U$  dans  $\mathbf{G}$ . Alors si  $\vec{f}$  et  $\vec{g}$  sont de classe  $\mathcal{C}^k$ , alors  $\vec{g} \circ \vec{f}$  est de classe  $\mathcal{C}^k$ .

*Preuve de la proposition 4.1.9.* — La preuve dans le cas général est hors programme. Montrons néanmoins le résultat dans le cas simple où  $k = 1$ .

**Proposition 4.1.10.** — Soient  $k$  un entier naturel non nul. Soit  $(\vec{e}_1', \vec{e}_2', \dots, \vec{e}_n')$  une base de  $\mathbf{F}$  notée  $\mathcal{B}'$ . On désigne par  $f_i$  la  $i^{\text{e}}$  application composante de  $\vec{f}$  dans  $\mathcal{B}'$ ,  $\vec{u}$  désignant toujours un vecteur de  $\mathbf{E}$  non nul et  $a$  un point de  $U$ .

L'application  $\vec{f}$  est de classe  $\mathcal{C}^k$ , si et seulement si  $f_1, f_2, \dots, f_n$  sont de classe  $\mathcal{C}^k$  et si c'est le cas, les dérivées directionnelles d'ordre  $k$  de  $\vec{f}$  on comme  $i^{\text{e}}$  composante les dérivées directionnelles d'ordre  $k$  de  $f_i$  selon les mêmes vecteurs, pour  $i = 1, 2, \dots, n$ .

*Preuve de la proposition 4.1.10.* — C'est une conséquence directe de 2.2.7. et du fait que la continuité d'une application équivaut à la continuité de ses applications composantes dans une base.

Regardons maintenant comment prouver concrètement le caractère  $\mathcal{C}^k$  d'une application.

**Exercice** — Soit l'application

$$\vec{f} : \mathbf{R}^2 \rightarrow \mathbf{R}^2 \quad (x, y) \mapsto (\cos(xy) \sin(y^2), y \exp(x)).$$

Montrer que  $\vec{f}$  est pour tout entier  $k$  de classe  $\mathcal{C}^k$ , on dit, comme pour les applications d'une variable réelle, de classe  $\mathcal{C}^\infty$ .

**Exemple** — Montrons que les application polynômiales et les applications rationnelles définies sur un ouvert de  $\mathbf{R}^p$  sont de classe  $\mathcal{C}^\infty$ .

## 4.2 Théorème de Schwarz

On considère encore une application  $\vec{f}$  d'un ouvert  $U$  de  $\mathbf{E}$  à valeurs dans  $\mathbf{F}$ .

Nous admettrons le théorème suivant :

**Proposition 4.2.1.** — THÉORÈME DE SCHWARZ —

*On suppose que  $\vec{f}$  est de classe  $\mathcal{C}^2$ . Pour tout  $\vec{u}$  et tout  $\vec{v}$  vecteurs de  $\mathbf{E}$  non nuls,*

$$D^2_{\vec{u},\vec{v}}\vec{f} = D^2_{\vec{v},\vec{u}}\vec{f}.$$

**Mise en garde !**

Le résultat est faux sous la seule hypothèse d'existence de ces dérivées  $D_{\vec{u},\vec{v}}\vec{f}$  et  $D_{\vec{v},\vec{u}}\vec{f}$  (cf. exercices).

**Corollaire 4.2.2.** — Soit  $\mathcal{B}$  une base de  $\mathbf{E}$ . Si  $\vec{f}$  est de classe  $\mathcal{C}^2$ , pour tout  $i$  et tout  $j$  éléments de  $\{1, \dots, p\}$ , dans la base  $\mathcal{B}$ , on a l'égalité des dérivées partielles :

$$\partial_{i,j}^2 \vec{f} = \partial_{j,i}^2 \vec{f}.$$

Le théorème de Schwarz se généralise comme suit.

**Proposition 4.2.3.** — Soient un entier  $k \geq 2$  et  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$  des vecteurs de  $\mathbf{E}$  quelconques. On suppose que  $\vec{f}$  est de classe  $\mathcal{C}^k$ . Alors pour toute permutation  $\sigma$  de  $\{1, 2, \dots, p\}$ ,

$$D_{\vec{u}_k, \vec{u}_{k-1}, \dots, \vec{u}_2, \vec{u}_1}^k \vec{f} = D_{\vec{u}_{\sigma(k)}, \vec{u}_{\sigma(k-1)}, \dots, \vec{u}_{\sigma(2)}, \vec{u}_{\sigma(1)}}^k \vec{f}.$$

Dit brutalement : l'ordre de dérivation selon  $k$  vecteurs n'a pas d'importance.

En particulier, si  $\mathbf{E}$  est muni d'une base  $\mathcal{B}$ , pour tout  $k$ -uplet  $(i_1, i_2, \dots, i_k)$  éléments de  $\{1, \dots, p\}$ , dans la base  $\mathcal{B}$ , on a l'égalité des dérivées partielles :

$$D_{i_k, i_{k-1}, \dots, i_2, i_1}^k \vec{f} = D_{i_{\sigma(k)}, i_{\sigma(k-1)}, \dots, i_{\sigma(2)}, i_{\sigma(1)}}^k \vec{f}.$$

**Exercice 4.2.4.** —

1. Soit  $\vec{f} : \mathbf{R}^2 \rightarrow \mathbf{F}$ ;  $(x, y) \mapsto f(x, y)$  une application de classe  $\mathcal{C}^2$ . Soient  $(a, b)$  et  $(c, d)$  des vecteurs non nuls de  $\mathbf{R}^2$ . Exprimer  $D_{(a,b),(c,d)}^2$  au moyen des dérivées partielles de  $\vec{f}$ .
2. Soit  $f$  une application de  $\mathbf{R}^2 - (0, 0)$  dans  $\mathbf{R}$  de classe  $\mathcal{C}^2$ . On appelle laplacien de  $f$ , l'application notée  $\Delta(f)$  de  $U$  dans  $\mathbf{R}$ ,  $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$ . Le laplacien intervient de façon cruciale dans beaucoup de domaines des sciences. Considérons l'application  $\tilde{f}$ , « expression de  $f$  en coordonnées polaires »,

$$\tilde{f} : \mathbf{R}_+^* \times \mathbf{R} : (r, \theta) \mapsto f(r \cos \theta, r \sin \theta).$$

Exprimer pour un couple  $(r, \theta)$  de  $\mathbf{R}_+^* \times \mathbf{R}$ ,  $\Delta(f)(r \cos \theta, r \sin \theta)$  en fonction des valeurs des dérivées partielles de  $\tilde{f}$  au point  $(r, \theta)$ .

**Solution de l'exercice 4.2.4.** —

1. ...

Finalement

$$D_{(a,b),(c,d)}^2 = \dots$$



2. ...

À suivre...

## Chapitre VIII

# GROUPES, ANNEAUX, CORPS

## GROUPES

### 1.1 Rappel du cours de première année

Nous allons rappeler l'essentiel du cours de première année. Les preuves des résultats seront souvent omises, au mieux esquissées ; nous laissons au lecteur le soin de s'assurer qu'il s'en souvient. Dans le cas contraire, il est impératif qu'il consulte son cours de sup.

Nous illustrerons ce cours d'exemples divers tirés du programme de MPSI (algèbre linéaire, géométrie...) qui nécessitent des connaissances qui dépassent la théorie des groupes et qui ont été logiquement et chronologiquement introduites après le cours sur les groupes.

**Définition 1..1.1.** — On appelle *groupe* tout couple  $(G, *)$ , où  $G$  est un ensemble,  $*$  une loi interne sur  $G$  et qui jouit des propriétés suivantes :

1. La loi  $*$  est **associative**<sup>1</sup>, c'est-à-dire que pour tout  $x$ , tout  $y$  et tout  $z$  éléments de  $G$ ,

$$(x * y) * z = x * (y * z) ;$$

2. la loi  $*$  possède un **élément neutre**<sup>2</sup>, c'est-à-dire qu'il existe un élément  $e$  de  $G$ , tel que, pour tout élément  $g$  de  $G$ ,

$$e * g = g * e = g ;$$

3. tout élément de  $G$  possède un **symétrique**, c'est-à-dire que, pour tout élément  $g$  de  $G$ , il existe un élément  $x'$  de  $G$  tel que,

$$g * g' = g' * g = e.$$

Si de plus la loi  $*$  est commutative, on dit que le groupe est commutatif ou abélien<sup>3</sup>

**Remarques** —

- On montre que dans un groupe tout élément possède un et un seul symétrique, on peut donc parler « du symétrique » d'un élément ;
- Par abus, lorsque il n'y a pas d'ambiguïté sur la loi dont on munit l'ensemble  $G$ , on parle du « groupe  $G$  » au lieu « du groupe  $(G, *)$  ».

**Notations, terminologie** — Les deux opérations élémentaires que sont l'addition et la multiplication (dans  $\mathbf{R}$  par exemple), inspirent deux types de notations pour les groupes.

- LA NOTATION MULTIPLICATIVE :
  - la loi est notée  $*$ ,  $\star$ ,  $\times$ ,  $\cdot$ ,  $\circ$ , ...
  - on parle du *produit*  $x * y$  de deux éléments  $x$  et  $y$  du groupe  $G$ ,
  - le symétrique d'un élément  $x$  de  $G$  est noté  $x^{-1}$ , on l'appelle parfois inverse de  $x$ ,

1. Cette propriété autorise dans la pratique à ne pas utiliser de parenthèses dans l'écriture d'opérations successives utilisant la loi  $*$ . On peut donc écrire des expressions de la forme :  $x * y * z$ ,  $x * y * z * w$ ,  $x_1 * x_2 * \dots * x_n$  etc.

2. Rappelons l'unicité de l'élément neutre : une loi interne qui possède un élément neutre possède un et un seul élément neutre.

3. Cette terminologie honore le grand mathématicien *Niels Abel* (1802-1829) qui fut l'un des premiers à étudier la structure de certains groupes commutatifs.

— pour tout élément  $k$  de  $\mathbf{Z}$ , et tout élément  $x$  de  $G$ , on définit  $x^k$  (lire «  $x$  puissance  $k$  » par

$$\begin{cases} \underbrace{x * x * \cdots * x}_{k \text{ termes}}, & \text{pour } k > 0, \\ e, \text{ neutre de } G, & \text{pour } k = 0, \\ \underbrace{x^{-1} * x^{-1} * \cdots * x^{-1}}_{-k \text{ termes}}, & \text{pour } k < 0. \end{cases}$$

• LA NOTATION ADDITIVE :

- la loi est notée  $+$ ,
- on parle de la *somme*  $x + y$  de deux éléments  $x$  et  $y$  du groupe  $G$ ,
- le symétrique d'un élément  $x$  de  $G$  est noté  $-x$ , on l'appelle parfois opposé de  $x$ ,
- pour tout élément  $k$  de  $\mathbf{Z}$ , et tout élément  $x$  de  $G$ , on définit  $k \cdot x^1 \ll k \text{ fois } x \gg$  par

$$\begin{cases} \underbrace{x + x + \cdots + x}_{k \text{ termes}}, & \text{pour } k > 0, \\ e, \text{ neutre de } G, & \text{pour } k = 0, \\ \underbrace{(-x) + (-x) + \cdots + (-x)}_{-k \text{ termes}}, & \text{pour } k < 0. \end{cases} \quad (\text{VIII.1})$$

- Comme nous y engage la coutume, nous réserverons la notation additive aux seuls groupes abéliens, la notion multiplicative, elle est employée aussi bien pour des groupes abéliens que pour des groupes non commutatifs.
- Dans la suite nous noterons l'élément neutre d'un groupe  $G$ ,  $e_G$ , ou simplement  $e$  s'il n'y a qu'un seul groupe et donc aucun risque de confusion.

**Exemples** —

- $(\mathbf{Z}, +)$ ,  $(\mathbf{Q}, +)$ ,  $(\mathbf{R}, +)$ ,  $(\mathbf{C}, +)$ ,  $(\mathbf{R}^*, \times)$  et  $(\mathbf{C}^*, \times)$  sont des groupes abéliens.
- Soit  $E$  un ensemble non vide. Notons l'ensemble des bijections de  $E$  sur  $E$ , appelées *permutations* de  $E$ ,  $S(E)$ . L'ensemble  $S(E)$  peut être muni d'une loi interne : la *composition* que l'on note  $\circ$ . L'ensemble  $S(E)$  muni de cette loi,  $(S(E), \circ)$ , est un groupe. En général ce groupe n'est pas abélien.  
Cas particulier : l'ensemble  $E$  est l'ensemble  $\{1, 2, \dots, n\}$ , où  $n$  est un entier supérieur ou égal à 1. On note dans ce cas le groupe  $(S_n, \circ)$  et on l'appelle groupe symétrique d'indice  $n$ .
- Soient  $n$  un entier supérieur ou égal à 1 et  $\mathbf{K}$  un corps (le programme de MPSI impose de se limiter aux cas  $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{K} = \mathbf{C}$ ). L'ensemble des matrices carré d'ordre  $n$ , à coefficients dans  $\mathbf{K}$ , inversibles, muni de la multiplication est un groupe, on le note  $(GL_n(\mathbf{K}), \times)$  (groupe linéaire d'indice  $n$ ). En général il n'est pas commutatif.

**Définition 1.1.2.** — Soit  $(G, *)$  un groupe. On appelle sous-groupe de  $(G, *)$ , tout couple  $(H, \underset{H}{*})$  où :

- $H$  est une partie de  $G$  stable par la loi  $*$ ,
- $\underset{H}{*}$  est la loi induite par  $*$  sur  $H$ ,
- $(H, \underset{H}{*})$  est un groupe.

**Remarque** — Soit  $(H, \underset{H}{*})$  un sous-groupe de  $(G, *)$ . Dans la pratique on ne distingue pas la loi  $*$  et la loi qu'elle induit sur  $H$ , on note donc le sous-groupe  $(H, *)$ . Pire, on considère, par un abus commun, que c'est la partie  $H$  qui est un sous-groupe. on trouve donc des expressions du type :

« Soit  $H$  un sous-groupe de  $(G, *)$ , » voire même, « soit  $H$  un sous-groupe de  $G$  ».

Donnons une caractérisation pratique des sous-groupes.

**Proposition 1.1.3.** — Soit  $(G, *)$  un groupe. Soit  $H$  une partie de  $G$ . Avec les abus précédemment signalés,  $H$  est un sous-groupe de  $(G, *)$  si et seulement si

1. la partie  $H$  est non vide,
2. la partie  $H$  est stable par la loi  $*$ ,
3. la partie  $H$  est stable par passage au symétrique, c'est-à-dire que, pour tout élément  $x$  de  $H$ ,  $x^{-1} \in H$ .

**Remarques** —

- Dans la plupart des cas on prouve que  $H$  est non vide en montrant que l'élément neutre du groupe  $G$  est élément de  $H$ .
- On peut dans le critère précédent remplacer 2 et 3 par : Pour tout  $x$  et tout  $y$  éléments de  $H$ ,  $x * y^{-1} \in H$ .
- Pour prouver qu'un ensemble muni d'une loi est un groupe, dans la pratique, on montre très souvent qu'il s'agit d'un sous-groupe d'un groupe connu.

**Exemples** —

1. Souvent on omet de noter le signe «  $\cdot$  » :  $(kx$  pour  $k \cdot x$ ).

- Soit  $(G, *)$  un groupe.  $\{e\}$ , où  $e$  est l'élément neutre, et  $G$  sont des sous-groupes de  $(G, *)$ . Le premier est dit *sous-groupe trivial*, un sous-groupe de  $(G, *)$  distinct de  $G$  est dit *propre*.
- l'ensemble  $\{-1, 1\}$  est un sous-groupe de  $(\mathbf{R}^*, \times)$ .
- l'ensemble  $U$  des nombres complexes de module 1 est un sous-groupe de  $(\mathbf{C}^*, \times)$ . Pour tout entier  $n$  supérieur ou égal à 1, l'ensemble  $U_n$  des racines complexes  $n^e$  de l'unité est un sous-groupe de  $(\mathbf{C}^*, \times)$ , et aussi de  $(U, \times)$ .
- Soit  $\mathbf{E}$  un espace vectoriel (sur  $\mathbf{R}$  ou  $\mathbf{C}$ ). L'ensemble des bijections linéaires de  $\mathbf{E}$  sur  $\mathbf{E}$  (ou automorphismes) est un sous-groupe de  $(S(\mathbf{E}), \circ)$ . On l'appelle groupe linéaire de  $\mathbf{E}$  et on le note  $GL(\mathbf{E})$ .
- Soit  $\mathcal{E}$  le plan ou l'espace affine. L'ensemble des bijections affines de  $\mathcal{E}$  sur  $\mathcal{E}$  est un sous-groupe de  $(S(\mathcal{E}), \circ)$ . On l'appelle groupe affine de  $\mathcal{E}$  et on le note  $GA(\mathcal{E})$ .
- Soit  $(\mathbf{E}, \cdot)$  un espace euclidien. L'ensemble des automorphismes de  $\mathbf{E}$ , qui conservent le produit scalaire (appelés automorphismes orthogonaux) est un sous-groupe de  $GL(\mathbf{E})$ ; on l'appelle groupe orthogonal et on le note  $O(\mathbf{E})$ .

**Proposition 1.1.4.** — *L'intersection d'une famille, finie ou non, de sous-groupes d'un groupe  $(G, *)$  est un sous-groupe de  $(G, *)$ .*

*Preuve de la proposition 1.1.4.* — Soit  $(H_i)_{i \in I}$  une famille de sous-groupes de  $(G, *)$ .

**Proposition-définition 1.1.5.** — GROUPE PRODUIT

Soient  $(G_1, *)_1$  et  $(G_2, *)_2$  des groupes. La loi notée  $*$  définie sur  $G_1 \times G_2$  par, pour tout  $(x_1, x_2)$  et tout  $(y_1, y_2)$  éléments de  $G_1 \times G_2$ ,

$$(x_1, x_2) * (y_1, y_2) = (x_1 *_1 y_1, x_2 *_2 y_2),$$

est une loi interne. De plus,  $(G_1 \times G_2, *)$  est un groupe, appelé *groupe produit* des groupes  $(G_1, *)_1$  et  $(G_2, *)_2$ .

**Remarque** — On peut généraliser ce résultat à un produit de  $n$  groupes,  $n \geq 2$ .

Nous allons maintenant étudier la notion de morphisme de groupes. Un morphisme de groupes est une application d'un groupe dans un autre qui respecte les lois, plus précisément :

**Définition 1.1.6.** — Soient  $(G_1, *)_1$  et  $(G_2, *)_2$  des groupes. On appelle *morphisme* du groupe  $(G_1, *)_1$  dans le groupe  $(G_2, *)_2$ , toute application  $\varphi$  de  $G_1$  dans  $G_2$  telle que pour tout  $x$  et tout  $y$  éléments de  $G_1$ ,

$$\varphi(x *_1 y) = \varphi(x) *_2 \varphi(y).$$

**Proposition 1.1.7.** — PROPRIÉTÉS DES MORPHISMES DE GROUPE

Soit  $\varphi$  un morphisme d'un groupe  $(G_1, *)_1$  dans une groupe  $(G_2, *)_2$ . Alors,

1. L'image par  $\varphi$  du neutre est le neutre :

$$\varphi(e_{G_1}) = e_{G_2};$$

2. pour tout élément  $x$  de  $G_1$ ,

$$\varphi(x^{-1}) = (\varphi(x))^{-1};$$

3. Si  $H_1$  est un sous-groupe de  $(G_1, *)_1$ , alors l'image par  $\varphi$  de  $H_1$ ,  $\varphi(H_1)$ , est un sous-groupe de  $(G_2, *)_2$  ;

4. Si  $H_2$  est un sous-groupe de  $(G_2, *)_2$ , alors l'image réciproque par  $\varphi$  de  $H_2$ ,  $\varphi^{-1}(H_2)$ , est un sous-groupe de  $(G_1, *)_1$ .

*Preuve de la proposition 1.1.7. —*

1.  $\varphi(e_{G_1} *_1 e_{G_1}) = \dots$

2. Soit  $x$  un élément de  $G_1$ .  $\varphi(x *_1 x^{-1}) = \dots$

3. Soit  $H_1$  un sous-groupe de  $(G_1, *)$ .

Donc  $\varphi(H_1)$  est sous-groupe de  $(G_2, *)$ .

4. Soit  $H_2$  un sous-groupe de  $(G_2, *)$ .

Donc  $\varphi^{-1}(H_2)$  est sous-groupe de  $(G_1, *)$ .

**Remarque** — On a rappelé que  $\{e_{G_2}\}$  est un sous-groupe de  $(G_2, *)$  (sous-groupe trivial), donc, d'après le point 4,  $\varphi^{-1}(\{e_{G_2}\})$  est un sous-groupe de  $(G_1, *)$ ; on l'appelle *noyau* de  $\varphi$  et on le note  $\text{Ker}(\varphi)$ . le noyau de  $\varphi$  sert à caractériser l'injectivité de  $\varphi$  :

**Proposition 1.1.8.** — *Soit  $\varphi$  un morphisme d'un groupe  $(G_1, *)$  dans un groupe  $(G_2, *)$ . Alors  $\varphi$  est injectif si et seulement si  $\text{Ker}(\varphi) = \{e_{G_1}\}$ .*

**Remarque** — Pour prouver qu'un ensemble est un sous-groupe, il est parfois pratique de montrer que c'est le noyau d'un morphisme de groupes.

**Exemples** —

- Soit  $n$  un élément de  $\mathbf{N}^*$ . La signature est un morphisme de  $(S_n, \circ)$  dans  $(\{-1, 1\}, \times)$ ;

$$\varepsilon : S_n \rightarrow \{-1, 1\}; \sigma \mapsto \varepsilon(\sigma).$$

L'ensemble des permutations de signature égale à 1 est un sous-groupe de  $(S_n, \circ)$ , en effet. . .

On appelle ce sous-groupe *groupe alterné d'indice  $n$*  et on le note  $\mathcal{A}_n$ .

On suppose maintenant que  $n \geq 2$ . Montrons que le cardinal de  $\mathcal{A}_n$  est la moitié du cardinal de  $S_n$ , en effet. . .

Désignons par  $\tau_{1,2}$  la transposition « échangeant » 1 et 2 et considérons

$$\Phi : S_n \rightarrow S_n ; \sigma \mapsto \tau_{1,2} \circ \sigma.$$

Donc  $|\mathcal{A}_n| = \frac{1}{2}|S_n| = \frac{n!}{2}$ .

- Soit  $(\mathbf{E}, \cdot)$  un espace euclidien de dimension 2 ou 3. On montre, dans le cours de sup, que tout élément de  $O(\mathbf{E})$  a un déterminant égal à 1 ou  $-1$ . Soit l'application

$$d : O(\mathbf{E}) \rightarrow \{-1, 1\} ; u \mapsto \det(u).$$

l'application  $d$  est un morphisme du groupe  $(O(\mathbf{E}), \circ)$  dans le groupe  $(\{-1, 1\}, \times)$ . Les éléments de  $O(\mathbf{E})$  de déterminant 1 sont appelés rotations. L'ensemble des rotations forme un sous-groupe de  $O(\mathbf{E})$ , en effet...

On l'appelle groupe spécial orthogonal et on le note  $SO(\mathbf{E})$ .

- Soit  $(\mathbf{E}, \cdot)$  un espace euclidien de dimension 2, orienté. Soit  $r$  l'application de  $\mathbf{R}$  dans  $SO(\mathbf{E})$ , qui à un réel  $\theta$  associe la rotation d'angle de mesure  $\theta$ .

$$r : \mathbf{R} \rightarrow SO(\mathbf{E}) ; \theta \mapsto r(\theta),$$

est un morphisme du groupe  $(\mathbf{R}, +)$  dans le groupe  $(SO(\mathbf{E}), \circ)$ . Ce morphisme est surjectif, mais il n'est pas injectif, en effet...

Voici maintenant une notion très importante, celle d'isomorphisme de groupes.

**Définition 1.1.9.** — Soient  $(G_1, *)$  et  $(G_2, *)$  des groupes. On appelle isomorphisme du groupe  $(G_1, *)$  sur le groupe  $(G_2, *)$ , toute application  $\varphi$  de  $G_1$  dans  $G_2$  telle que :

1. l'application  $\varphi$  soit un morphisme,
2. l'application  $\varphi$  soit une bijection de  $G_1$  sur  $G_2$ .

**Proposition 1.1.10.** — Soit  $\varphi$  un isomorphisme du groupe  $(G_1, *)$  sur le groupe  $(G_2, *)$ . Alors  $\varphi^{-1}$ , bijection réciproque de  $\varphi$  est un isomorphisme du groupe  $(G_2, *)$  sur le groupe  $(G_1, *)$ .

Il résulte de la proposition 1.1.10. que l'on peut dire que deux groupes sont isomorphes lorsqu'il existe un isomorphisme de l'un sur l'autre (et donc du second sur le premier). La notion de groupes isomorphes est fondamentale en algèbre. Des groupes  $(G_1, *)$  et  $(G_2, *)$  isomorphes ont même structure, ou si l'on préfère, le groupe  $(G_2, *)$  est le groupe  $(G_1, *)$  dont on aurait, en quelque sorte, « changé » le nom des éléments grâce à l'isomorphisme  $\varphi$  de  $(G_1, *)$  sur  $(G_2, *)$ . En particulier des groupes isomorphes sont simultanément commutatifs ou non, si l'un d'eux est fini alors l'autre aussi et ils ont même cardinal...

**Exemple** —

- Soit  $\mathbf{E}$  un espace vectoriel sur un corps  $\mathbf{K}$ , de dimension finie non nulle  $n$ . Les groupes  $(GL(\mathbf{E}), \circ)$  et  $(GL_n(\mathbf{K}), \times)$  sont isomorphes. En effet, prenons  $\mathcal{B}$  une base de  $\mathbf{E}$ , alors l'application

$$\Phi : GL(\mathbf{E}) \rightarrow GL_n(\mathbf{K}) ; u \mapsto \text{mat}_{\mathcal{B}}(u)$$

est un isomorphisme du groupe  $(GL(\mathbf{E}), \circ)$  sur le groupe  $(GL_n(\mathbf{K}), \times)$  (cf. cours de sup.).

- Soit  $E$  un ensemble de cardinal  $n$  non nul. Nous nous proposons de montrer que les groupes  $(S(E), \circ)$  et  $(S_n, \circ)$  sont isomorphes. Dire que le cardinal de  $E$  vaut  $n$  signifie qu'il existe une bijection  $B$  de  $E$  sur  $\{1, 2, \dots, n\}$ , c'est-à-dire que concrètement, on peut numéroter les éléments de  $E$  de 1 à  $n$ . Nous allons associer à une permutation  $\sigma$  de  $E$  une permutation  $s$  de  $\{1, \dots, n\}$  de façon naturelle : si l'image d'un élément  $a$  de  $E$  par  $\sigma$  est  $b$  nous définirons l'image du numéro de  $a$  par  $s$ , comme étant le numéro de  $b$ . On est dans la vie courante accoutumé

à ce genre de pratique, où par exemple dans une course de chevaux (ou de cyclistes), on dira que le numéro 12 a dépassé le numéro 6, plutôt que de dire que tel cheval (ou cycliste) a dépassé tel autre. Mathématiquement l'application qui à  $\sigma$  associe  $s$  se définit par :

$$\Phi : S(E) \rightarrow S_n ; \sigma \mapsto B \circ \sigma \circ B^{-1}.$$

Remarquons que  $B \circ \sigma \circ B^{-1}$  est bien un élément de  $SS_n$ , puisque...

$$\begin{array}{ccc} E & \xrightarrow{\sigma} & E \\ B \downarrow & & \downarrow B \\ \{1,2,\dots,n\} & \xrightarrow{s} & \{1,2,\dots,n\} \end{array}$$

Reste à prouver que  $\Phi$  est un isomorphisme de  $(S(E), \circ)$  et  $(S_n, \circ)$ .

Donc  $\Phi$  est un isomorphisme de  $(S(E), \circ)$  et  $(S_n, \circ)$ . Donc  $S(E)$  et  $S_n$  ont même cardinal, on peut parler du nombre de permutations d'un ensemble à  $n$  éléments, sans préciser cet ensemble; ce nombre vaut  $\text{card}(S_n)$  soit  $n!$ .

## 1.2 Sous-groupes de $\mathbf{Z}$ , groupe $\mathbf{Z}/n\mathbf{Z}$

Nous commençons par le rappel d'un résultat déjà revu cette année :

**Proposition 1.2.1.** — SOUS-GROUPES DE  $(\mathbf{Z}, +)$  —

Les sous groupe de  $(\mathbf{Z}, +)$  sont exactement les ensembles de la forme  $n\mathbf{Z}$ , avec  $n$  élément de  $\mathbf{N}$ .

*Preuve de la proposition 1.2.1.* — La preuve a été donnée en MPSI et revue en exercice dans le chapitre sur les e.v.n.

Nous allons maintenant introduire un groupe fondamental, le groupe  $\mathbf{Z}/n\mathbf{Z}$ . Ce groupe est d'un usage dans la vie courante presque aussi important que celui de  $\mathbf{Z}$ . Explications. Imaginons par exemple, qu'au mépris des théories scientifiques et théologiques, on représente la succession des jours par l'ensemble  $\mathbf{Z}$ . Lorsque l'on parle de l'emploi du temps du *lundi*, on identifie des jours qui diffèrent d'un multiple de 7, en fait les jours sont regroupés en sept catégories, (lundi, mardi,...,dimanche) qui chacune regroupe des jours qui diffèrent entre eux d'une ou plusieurs semaines, c'est-à-dire d'un multiple de sept jours. Il s'agit là d'une représentation de l'ensemble  $\mathbf{Z}/7\mathbf{Z}$ .

Passons à la définition générale.

Jusque à la fin de ce paragraphe,  $n$  désigne un entier naturel non nul

**Définition 1.2.2** On dit qu'un entier  $x$  est congru modulo  $n$  à un entier  $y$  si  $n$  divise  $y - x$ . On notera alors :  $x \equiv y[n]$ , on trouve aussi les notations  $x = y[n]$  ou  $x \equiv y \bmod n$  ou  $x = y \bmod n$ .

**Exemples** —  $6 \equiv 20[14]$  ;  $0 \equiv 100[2]$  ;  $-6 \equiv 5[11]$ .

**Proposition 1.2.3** — La relation sur  $\mathbf{Z}$ , définie par 1.2.2, (relation de congruence modulo  $n$ ), est une relation d'équivalence.

*Preuve de la proposition 1.2.3.* — Il faut montrer que cette relation est  
Ce qui est assez évident, montrons par exemple qu'elle est...

Maintenant quelques remarques évidentes.

**Remarque** —

1. La symétrie de la relation de congruence autorise à dire que des entiers *sont* congrus modulo  $n$ .
2. Des entiers  $x$  et  $y$  sont congrus modulo  $n$  si et seulement si ils ont même reste dans la division par  $n$ .
3. Dire que des entiers  $x$  et  $y$  sont congrus modulo  $n$  ne signifie rien d'autre que  $y - x$  appartient au sous-groupe de  $\mathbf{Z}$ ,  $n\mathbf{Z}$ .

**Terminologie-notations** — La classe d'équivalence d'un entier  $x$  dans la relation de congruence modulo  $n$  sera notée  $\bar{x}^n$  ou même, s'il n'y a pas risque de confusion  $\bar{x}$ . L'ensemble des classes d'équivalence (ensemble quotient) sera noté  $\mathbf{Z}/n\mathbf{Z}$  (cf. point 3. de la remarque).

Dans ce cours, nous réserverons dans la mesure du possible les lettres de la fin de l'alphabet  $x, y, z, x_i, y'$  etc. pour désigner les éléments de  $\mathbf{Z}$  et celles du début  $a, b, c, a_i, b'$  etc. pour désigner les classes de congruence modulo  $n$ , c'est-à-dire les éléments de  $\mathbf{Z}/n\mathbf{Z}$ .

**Exemples** —

$$\bar{5}^2 = \dots$$

$$\bar{3} = \{\dots - 5, -2, 1, 4, 7, 10, 13, \dots\}.$$

**Proposition 1.2.4.** —  $\mathbf{Z}/n\mathbf{Z}$  est possède exactement  $n$  éléments.



*Preuve de la proposition 1.2.4. —*

Ainsi a-t-on :  $\mathbf{Z}/1\mathbf{Z} = \{\bar{0}\}$  ;  $\mathbf{Z}/2\mathbf{Z} = \{\bar{0}, \bar{1}\}$  ;  $\mathbf{Z}/n\mathbf{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$ .

Nous allons maintenant munir l'ensemble  $\mathbf{Z}/n\mathbf{Z}$  d'une addition, dans le but d'en faire un groupe. Il serait agréable que la somme de deux éléments  $\bar{x}$  et  $\bar{y}$  fût  $\overline{x+y}$ . Toutefois avant de définir ainsi l'addition sur  $\mathbf{Z}/n\mathbf{Z}$ , il convient de s'assurer que l'élément  $\overline{x+y}$  ne dépend que des deux éléments de  $\mathbf{Z}/n\mathbf{Z}$  dont on calcul la somme et non de leurs représentants particuliers  $x$  et  $y$ . Ce qui va être fait grâce au concours du lemme suivant :

**Lemme 1.2.5.** — *Soient  $x, x', y$  et  $y'$  des éléments de  $\mathbf{Z}$ . On suppose que :*

$$x \equiv x' [n] \text{ et } y \equiv y' [n].$$

*Alors*

$$x + y \equiv x' + y' [n].$$

*Preuve du lemme 1.2.5* — Par hypothèse il existe  $k$  et  $h$ , éléments de  $\mathbf{Z}$ , tels que :

$$x' = x + kn, \quad y' = y + hn.$$

Par addition de ces égalités :  $x' + y' = x + y + (k + h)n$  et donc  $x + y \equiv x' + y' [n]$ . □

**Définition 1.3.2.** — *Soit  $a$  et  $b$  des éléments de  $\mathbf{Z}/n\mathbf{Z}$ . On appelle somme de  $a$  par  $b$  noté  $a +_n b$  l'élément de  $\mathbf{Z}/n\mathbf{Z}$ ,  $\overline{x+y}$ , où  $x$  et  $y$  sont des représentants respectifs de  $a$  et  $b$ , quantité, qui d'après le lemme 1.2.5 ne dépend pas du choix des représentants.*

Dans la suite on notera souvent, lorsqu'il n'y a pas d'ambiguïté, la loi  $+_n$  simplement  $+$ .

On retient

$$\boxed{\bar{x} +_n \bar{y} = \overline{x+y}}$$

**Exemple** —

- Dans  $\mathbf{Z}/6\mathbf{Z}$ ,  
 $\bar{4} + \bar{3} = \dots$   
 $\bar{2} + \bar{5} = \dots$
- Dans  $\mathbf{Z}/11\mathbf{Z}$ ,  
 $\bar{4} + \bar{8} = \dots$   
 $\bar{2} + \bar{10} = \dots$

La loi  $+_n$  est une loi interne sur  $\mathbf{Z}/n\mathbf{Z}$  et l'on a le résultat de structure suivant :

**Proposition 1.2.6.** —  $\left(\mathbf{Z}/n\mathbf{Z}, +_n\right)$  est un groupe commutatif.

*Preuve de la proposition 1.2.5. —*

Par définition de l'addition  $+$  sur  $\mathbf{Z}/n\mathbf{Z}$ , on a le résultat suivant :

**Proposition-définition 1.2.7.** — *Soit  $\varphi_n$ , (ou  $\pi_n$ ) l'application de  $\mathbf{Z}$  dans  $\mathbf{Z}/n\mathbf{Z}$  qui à un entier  $x$  associe  $\bar{x}^n$ . Cette application est un*  
*surjectif.*

*On l'appelle projection canonique ou surjection canonique de  $\mathbf{Z}$  sur  $\mathbf{Z}/n\mathbf{Z}$ .*

### 1.3 Groupes engendrés, sous-groupes engendrés

Nous abordons une notion capitale sur les groupes, la notion de partie génératrice, c'est-à-dire de parties d'un groupe, dont les éléments et leurs symétriques composés entre eux donnent tous les éléments du groupe, un peu à la manière des parties génératrices des espaces vectoriels dont les combinaisons linéaires donnent tous les vecteurs de l'espace.

**Proposition-définition 1.3.1.** — *Soit  $(G, *)$  un groupe et  $A$  une partie de  $G$ . L'ensemble des sous-groupes de  $G$  qui contiennent  $A$  possède un plus petit élément pour l'inclusion.*

*Cet élément s'appelle sous-groupe engendré par  $A$ , on le note  $\langle A \rangle$ . Pour un élément  $a$  de  $G$ ,  $\langle \{a\} \rangle$  est souvent noté plus simplement  $\langle a \rangle$  et appelé sous-groupe engendré par  $a$ .*

**Exemples** —  $\langle \emptyset \rangle = \dots$  ;  $\langle G \rangle = \dots$ , dans  $(\mathbf{R}^*, \times)$   $\langle -1 \rangle = \dots$ .

*Preuve de la proposition 1.3.1. —* Nous allons donner deux preuves fort différentes, mais toutes deux instructives.

PREMIÈRE PREUVE :

SECONDE PREUVE :

**Définition 1.3.2.** — *On dit qu'une partie  $A$  d'un groupe  $(G, *)$  engendre  $G$  ou encore est une partie génératrice de  $G$ , si  $\langle A \rangle = G$*

Donnons des exemples de parties génératrices de groupes usuels.

**Exemples** —

1.  $(\mathbf{Z}, +)$ .
2.  $(\mathbf{Z}/n\mathbf{Z}, +)$ , pour  $n$  élément de  $\mathbf{N}^*$ .
3.  $(U_n, \times)$ , pour  $n$  élément de  $\mathbf{N}^*$ .
4.  $(S_n, \circ)$ , pour  $n$  élément de  $\mathbf{N}^*$ .
5.  $(O(\mathbf{R}^2), \circ)$  ( $\mathbf{R}^2$  est muni de sa structure euclidienne canonique).

**Définition 1.3.3.** — *Un groupe est dit monogène si il admet une partie génératrice à un élément. Un groupe monogène fini est dit cyclique*<sup>1</sup>

**Exemples** — On a vu précédemment que  $(\mathbf{Z}, +)$  est monogène et que, pour  $n$  élément de  $\mathbf{N}^*$ ,  $(\mathbf{U}_n, \times)$ , et  $(\mathbf{Z}/n\mathbf{Z}, +)$  sont cycliques.

**Exercices** — Montrer que les groupes  $(\mathbf{U}_n, \times)$ , et  $(\mathbf{Z}/n\mathbf{Z}, +)$  sont isomorphes.

Le résultat de l'exercice précédent résulte en fait d'un résultat plus fort : à isomorphisme près, il n'existe qu'un seul groupe cyclique d'ordre (de cardinal) donné. Plus précisément on a le résultat suivant.

**Proposition 1.3.4.** — *Soit  $(G, +)$  un groupe monogène. Alors :*

- *Si  $G$  est infini, alors  $(G, +)$  est isomorphe à  $(\mathbf{Z}, +)$  ;*
- *Si  $G$  est fini (cyclique), alors  $(G, +)$  est isomorphe à  $(\mathbf{Z}/n\mathbf{Z}, +)$ , où  $n$  désigne l'ordre de  $(G, +)$ .*

*Dit brutalement, à isomorphisme près il existe un et un seul groupe monogène infini,  $\mathbf{Z}$  et pour tout entier  $n \geq 1$ , il existe un et un seul groupe cyclique d'ordre  $n$ ,  $\mathbf{Z}/n\mathbf{Z}$ .*

*Preuve de la proposition 1.3.4* — Soit  $a$  un élément de  $G$  qui l'engendre. D'après la seconde preuve de 1.3.1.,

$$G = \{ \quad \quad \quad \}.$$

D'où l'idée de considérer l'application :  $\varphi : \mathbf{Z} \rightarrow G, k \mapsto \quad$ . On alors le résultat suivant :

**Lemme**<sup>2</sup> — *L'application  $\varphi$  est*

1. Nous verrons dans la suite l'origine de cette terminologie

2. ce lemme est important et au programme

Nous démontrerons ce lemme un peu technique mais sans malice à fin la preuve.

Comme promis donnons la preuve du lemme.

*Preuve du lemme* —

Ce résultat appelle plusieurs remarques.

On peut désormais noter un groupe cyclique d'ordre  $n$ ,  $\{e, a, a^2, a^3, \dots, a^{n-1}\}$ . La structure *cyclique* du groupe apparaît ainsi pleinement, justifiant une terminologie *a priori* étrange.

Un groupe cyclique  $G$ , isomorphe à un groupe  $\mathbf{Z}/n\mathbf{Z}$ , est donc, comme ce dernier groupe commutatif, on peut donc, si on le désire le noter additivement :  $\{ \underset{n \cdot a}{e}, 1 \cdot a, 2 \cdot a, 3 \cdot a, \dots, (n-1) \cdot a \}$ .

**Remarque** — Soit  $G$  un groupe cyclique d'ordre  $n$ , alors pour tout élément  $g$  de  $G$ ,  $g^n = \dots$ , en effet.....

**Définition** — Soient  $(G, *)$  un groupe, et  $g$  un de ses éléments. Si le sous-groupe (monogène) engendré par  $g$  est fini, alors on dit que  $g$  est d'ordre fini et l'on appelle ordre de  $g$  l'ordre de ce sous-groupe. Sinon on dit que  $g$  est d'ordre infini.

Si  $g$  est d'ordre fini on note l'ordre de  $g$ ,  $\omega(g)$  :

$$\omega(g) = |\langle g \rangle|.$$

De l'étude des groupes cycliques et plus particulièrement du noyau de  $\varphi$  il vient le résultat suivant :

**Proposition** — Soient  $x$  un élément d'ordre fini d'un groupe  $(G, *)$  et  $p$  un élément de  $\mathbf{Z}$ . Alors,  $g^p = e_G$  si et seulement si  $\omega(x)$  divise  $p$ .

**Proposition** — Soit  $G$  un groupe fini. L'ordre de tout élément  $g$  de  $G$  divise l'ordre de  $G$ ,

$$\omega(g) \mid |G|.$$

*Preuve de la proposition* —

Seule la démonstration dans les cas où  $G$  est abélien est exigible. Nous nous limiterons à ce cas, mais nous traiterons en exercice d'un résultat plus fort dans le cas général, le théorème de Lagrange qui dit que l'ordre d'un sous-groupe, d'un groupe fini divise l'ordre du groupe.

**Exercices** —

1. Le groupe  $\mathbf{Z}/4\mathbf{Z}$  et le groupe produit  $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$  sont ils isomorphes ?
2. Même question pour les groupes produits  $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$  et  $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ .

*Solution* —

Nous reviendrons sur ce résultat.

# ANNEAUX, CORPS

## 2.1 Rappel du cours de première année

Comme nous l'avons fait pour les groupes nous allons nous livrer à une rétrospective rapide des notions d'anneau et de corps à travers le programme de MPSI. Dans ce panorama le choix des exemples n'obéira à aucune nécessité logique supposant déjà connues des notions à venir.

**Définition 2.1.1.** — On appelle anneau tout triplet  $(A, +, \times)$ , où  $A$  est un ensemble,  $+$  et  $\times$  des lois internes sur  $G$  et qui jouit des propriétés suivantes :

1.  $(A, +)$  est un **groupe abélien**,
2. La loi  $\times$  est **associative**
3. La loi  $\times$  possède un **élément neutre**
4. La loi  $\times$  est **distributive par rapport à l'addition**, c'est-à-dire que, pour tout  $x$ , tout  $y$  et tout  $z$  éléments de  $A$ ,

$$x \times (y + z) = (x \times y) + (x \times z) \text{ et } (y + z) \times x = (y \times x) + (z \times x).$$

Si de plus la loi  $\times$  est commutative, on dit que l'anneau est commutatif.

**Remarque** — Par abus, on parle très souvent de « l'anneau  $A$  » au lieu de « l'anneau  $(A, +, \times)$  », du moins lorsqu'il n'y a pas de doute sur les lois dont on équipe  $A$  pour obtenir une structure d'anneau.

**Terminologie notations** — Soit un anneau  $(A, +, \times)$ .

- On appelle généralement la loi  $+$  *addition* et la loi  $\times$  *multiplication*. L'élément neutre pour la loi  $+$  est appelé *zéro*, celui pour la multiplication *unité*; nous les noterons dans ce cours respectivement  $0_A$  et  $1_A$ . Le symétrique d'un élément  $x$  de  $A$  pour l'addition est appelé opposé (il est noté  $-x$ ), si  $x$  admet un symétrique pour la multiplication on appelle ce dernier *inverse* de  $x$  et on le note  $x^{-1}$ .
- On omet souvent dans les calculs dans un anneau les parenthèses en convenant que la multiplication est prioritaire sur l'addition, ainsi notera-t-on, pour  $x, y, z$  éléments de  $A$  la quantité  $(x \times y) + (x \times z)$ ,

$$x \times y + x \times z.$$

- Signalons enfin que l'on abrège  $+(-x)$  en  $-x$  et que le signe  $\times$  est assez fréquemment omis, ainsi peut-on écrire, pour  $x, y, z$  éléments de  $A$ , la quantité  $(x \times y) + (-z)$ ,

$$xy - z.$$

Donnons maintenant quelques règles de calcul dans un anneau. Pour plus de détails on se reportera au cours de Sup.

RÈGLES DE CALCULS DANS UN ANNEAU —

Soit un anneau  $(A, +, \times)$

1. Pour tout élément  $x$  de  $A$ ,

$$x \times 0_A = 0_A \times x = 0_A \text{ (} 0_A \text{ est absorbant)}.$$

2. Pour tout  $x$  et tout  $y$  éléments de  $A$ ,

$$(-x) \times y = x \times (-y) = -(x \times y),$$

cette dernière quantité peut donc sans ambiguïté se noter  $-x \times y$ , ou même parfois  $-xy$ .

3. Pour tout  $x$  et tout  $y$  élément de  $A$  et tout élément  $k$  de  $\mathbf{Z}$ ,

$$(k \cdot x) \times y = x \times (k \cdot y) = k \cdot (x \times y).$$

On notera donc sans aucun soucis cette quantité  $k \cdot x \times y$  où même  $k \cdot xy$  voire  $kxy$ .

4. Soient  $x$  et  $y$  des éléments de  $A$  qui commutent entre eux<sup>1</sup>, c'est-à-dire que  $x \times y = y \times x$ . Alors, pour tout élément  $n$  de  $\mathbf{N}$ ,

$$(x + y)^n = \binom{n}{0} \cdot x^n \times y^0 + \binom{n}{1} \cdot x^{n-1} \times y^1 + \cdots + \binom{n}{n-1} \cdot x^1 \times y^{n-1} + \binom{n}{n} \cdot x^0 \times y^n,$$

ce que l'on note de façon plus concise,

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} \cdot x^k \times y^{n-k} \text{ (binôme de Newton),}$$

1. Cette condition est primordiale. Dans un anneau non commutatif on devra toujours la vérifier.



avec

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Dans cette formule, on note pour  $a$  élément de  $A$  et  $p$  naturel,  $a^p$  la quantité  $\underbrace{a \times a \cdots \times a}_{p \text{ termes}}$  si  $p$  est non nul et

$1_A$  si  $p$  est nul, cf. notation multiplicative d'un groupe;  $\binom{n}{k}$  est un entier et le signe  $\ll \cdot \gg$  relève donc de la définition (VIII.1).

5. Soient  $x$  et  $y$  des éléments de  $A$  qui commutent entre eux. Alors avec les notations du point précédent, pour tout entier  $n$  supérieur ou égal à 1,

$$x^n - y^n = (x - y) \times (x^{n-1} \times y^0 + x^{n-2} \times y^1 + \cdots + x^1 \times y^{n-2} + x^0 \times y^{n-1}).$$

### Exemples —

- Soit  $\{a\}$  un singleton noté  $A$ . On peut munir  $A$  d'une et d'une seule structure d'anneau. Il existe une et une seule loi interne sur  $A$ , celle qui au couple  $(a, a)$  associe  $a$ . En prenant pour l'addition et pour la multiplication cette loi, c'est-à-dire en posant  $a + a = a$  et  $a \times a = a$ ,  $(A, +, \times)$  est un anneau, comme une vérification immédiate le montre. Notons que pour cet anneau,  $0_A = 1_A = a$ .  
Réciproquement, Si pour un anneau  $(A, +, \times)$ ,  $0_A = 1_A$  alors l'ensemble sous-jacent à l'anneau,  $A$ , est un singleton, comme le prouve 2.1.2.1.
- Un anneau dont l'ensemble sous-jacent est un singleton est dit *trivial* (ou nul puisqu'alors  $A = \{0_A\}$ ).
- $(\mathbf{Z}, +, \times)$ ;  $(\mathbf{Q}, +, \times)$ ;  $(\mathbf{R}, +, \times)$  et  $(\mathbf{C}, +, \times)$  sont des anneaux commutatifs.
- Dans cet exemple et les suivants,  $\mathbf{K}$  désigne  $\mathbf{R}$  ou  $\mathbf{C}$ , on pourrait prendre pour  $\mathbf{K}$  un corps quelconque, mais le programme de MPSI nous limite à ces deux cas. Alors  $(\mathbf{K}[X], +, \times)$  est un anneau commutatif.
- Soit  $\mathbf{E}$  un espace vectoriel sur  $\mathbf{K}$ . Alors  $(\mathcal{L}(\mathbf{E}), +, \circ)$  est un anneau, en général non commutatif.
- Soit  $n$  un entier naturel non nul. Alors  $(\mathcal{M}_n(\mathbf{K}), +, \times)$ , est un anneau en général non commutatif.
- Soit  $D$  un ensemble non vide. l'ensemble  $\mathcal{F}(D, \mathbf{K})$  des applications de  $D$  dans le corps  $\mathbf{K}$ , muni de l'addition et de la multiplication usuelles des applications,  $(\mathcal{F}(D, \mathbf{K}), +, \times)$ , est un anneau commutatif<sup>1</sup>. L'unité de cet anneau est...

On a le résultat suivant, de preuve très facile mais fastidieuse.

### Proposition définition 2.1.2. — — ANNEAU PRODUIT

Soient  $(A_1, +_1, \times_1)$  et  $(A_2, +_2, \times_2)$  des anneaux. La loi notée  $+$  (resp.  $\times$ ) définie sur  $A_1 \times A_2$  par, pour tout  $(x_1, x_2)$  et tout  $(y_1, y_2)$  éléments de  $A_1 \times A_2$ ,

$$(x_1, x_2) + (y_1, y_2) = (x_1 +_1 y_1, x_2 +_2 y_2), \text{ (resp. } (x_1, x_2) \times (y_1, y_2) = (x_1 \times_1 y_1, x_2 \times_2 y_2))$$

est une loi interne. De plus,  $(A_1 \times A_2, +, \times)$  est un anneau, appelé anneau produit des groupes  $(G_1, *_1)$  et  $(G_2, *_2)$ .

**Remarque** — On peut généraliser ce résultat à un produit de  $n$  anneaux,  $n \geq 2$ .

**Définition 2.1.3.** — Soit  $(A, +, \times)$  un anneau. On appelle sous-anneau de  $(A, +, \times)$ , tout triplet  $(B, +_B, \times_B)$  où :

- $B$  est une partie de  $A$  stable par les lois  $+$  et  $\times$ ,
- $1_A$  est élément de  $B$ .
- $+$  et  $\times$  sont les lois induites respectivement par  $+$  et  $\times$  sur  $B$ ,
- $(B, +_B, \times_B)$  est un anneau.

**Remarques** — Soit  $(B, +_B, \times_B)$  un sous-anneau de  $(A, +, \times)$ . Dans la pratique on ne distingue pas les lois  $+$  et  $\times$  et les lois qu'elles induisent sur  $B$ , on note donc le sous-anneau  $(B, +, \times)$ . Pire, on considère, par un abus commun, que c'est la partie  $B$  qui est un sous-anneau. on trouve donc des expressions du type :

$\ll$  Soit  $B$  un sous-anneau de  $(A, +, \times)$ ,  $\gg$  voire même,  $\ll$  soit  $B$  un sous-anneau de  $A \gg$ .

Attention dans la définition 2.1.3. le second point est essentiel, il assure notamment qu'un sous-anneau à même unité que l'anneau.  $A$  peut admettre des parties  $B$  stables par les lois  $+$  et  $\times$ , et telles que  $B$  muni des lois induites soit un anneau sans que pour autant  $B$  soit un sous-anneau de  $A$ . Par exemple on peut prendre pour  $(A, +, \times)$  un anneau non trivial et  $\{0_A\}$  pour  $B$ . Donnons un contre-exemple moins élémentaire. On considère l'anneau  $(\mathcal{M}_4(\mathbf{R}), +, \times)$  et l'ensemble  $B$  des éléments de  $(\mathcal{M}_4(\mathbf{R}), +, \times)$  de la forme

$$\begin{pmatrix} M & O_2 \\ O_2 & O_2 \end{pmatrix},$$

1. Le résultat est encore vrai si l'on remplace  $\mathbf{K}$  par un anneau  $A$  : l'ensemble des applications de  $D$  dans  $A$ , muni de l'addition et de la multiplication usuelles est un anneau, si de plus l'anneau  $A$  est commutatif alors l'anneau  $\mathcal{F}(D, A)$  est commutatif.

où  $M$  est un quelconque élément de  $\mathcal{M}_2(\mathbf{R})$ .

L'ensemble  $B$  est stable pour les lois  $+$  et  $\times$ , muni des lois induites, c'est un anneau. Toutefois  $B$  n'est pas un sous-anneau de  $(\mathcal{M}_4(\mathbf{R}), +, \times)$ , puisque qu'il ne contient pas l'unité de ce dernier, à savoir  $I_4$ . L'unité de l'anneau  $B$  est la matrice  $\begin{pmatrix} I_2 & O_2 \\ O_2 & O_2 \end{pmatrix}$ .

Donnons une caractérisation pratique des sous-anneaux. Soient  $(A, +, \times)$  un anneau et une partie  $B$  de  $A$ . On montre sans mal que  $B$  est un sous-anneau de  $(A, +, \times)$  si et seulement si

- $B$  est un sous-groupe de  $(A, +)$ ,
- $B$  est stable par  $\times$ ,
- $1_A$  est élément de  $B$ .

Donc en utilisant I.1.3,  $B$  est un sous-anneau de  $(A, +, \times)$  si et seulement si

- i.  $B$  est non vide,
  - ii.  $B$  est stable par  $+$ ,
  - iii.  $B$  est stable par passage à l'opposé,
  - iv.  $B$  est stable par  $\times$ ,
  - v.  $1_A$  est élément de  $B$ .
- } (sous groupe)

Puisque la condition v. est plus forte que la condition i., on obtient le résultat suivant.

**Proposition 2.1.4.** — *Soit  $(A, +, \times)$  un anneau. Une partie  $B$  de  $A$  est un sous-anneau de  $(A, +, \times)$  si et seulement si*

1.  $1_A$  est élément de  $B$ ,
2.  $B$  est stable par  $+$ ,
3.  $B$  est stable par passage à l'opposé,
4.  $B$  est stable par  $\times$ .

**Remarques** —

- Comme dans la caractérisation des sous-groupes, on peut remplacer 2 et 3 par :  $B$  est stable par différence.
- Le point 1 est *essentiel* et ne saurait être remplacé par quoi que ce soit. On a besoin impérativement de l'appartenance de  $1_A$  à  $B$  contrairement au cas des sous-groupes où l'on montre, *par exemple*, l'appartenance de  $e$ .

**Exemples** —

- L'ensemble  $\mathbf{Z}$  est un sous-anneau de  $\mathbf{Q}$ , de  $\mathbf{R}$  et de  $\mathbf{C}$ , l'ensemble  $\mathbf{Q}$  est un sous-anneau de  $\mathbf{R}$  et  $\mathbf{C}$ , l'ensemble  $\mathbf{R}$  est un sous-anneau de  $\mathbf{C}$ .
- Soient  $n$  un élément de  $\mathbf{N}$ . On note  $\mathcal{D}_n(\mathbf{K})$  l'ensemble des éléments de  $\mathcal{M}_n(\mathbf{K})$  diagonaux et  $\mathcal{T}_n^+(\mathbf{K})$  l'ensemble des éléments de  $\mathcal{M}_n(\mathbf{K})$  triangulaires supérieures (avec toujours  $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{K} = \mathbf{C}$ ). Alors  $\mathcal{D}_n(\mathbf{K})$  et  $\mathcal{T}_n^+(\mathbf{K})$  sont des sous-anneaux de  $(\mathcal{M}_n(\mathbf{K}), +, \times)$ . En effet...

Nous allons maintenant rappeler la définition d'anneaux particuliers, les anneaux intègres et les corps.

**Définition 2.1.5.** — ANNEAU INTÈGRE

Un anneau  $(A, +, \times)$  est dit *intègre* si par définition,

1. l'anneau  $(A, +, \times)$  est non trivial,
2. l'anneau  $(A, +, \times)$  est commutatif,
3. pour tout  $a$  et tout  $b$  éléments de  $A$ , si  $a \times b = 0_A$  alors  $a = 0_A$  ou  $b = 0_A$ .

Il arrive que l'on exprime le point 3 en disant que  $A$  est *sans diviseur de  $0_A$* , ce qui n'est guère heureux, puisque dans tout anneau tout élément divise le zéro.

**Remarque** — Dans un anneau intègre, pour résoudre une équation on a intérêt à factoriser.

**Exemples, contre-exemples —**

$\mathbf{K}$  désigne le corps  $\mathbf{R}$  ou le corps  $\mathbf{C}$ , ce pour se conformer au programme, on pourrait en fait prendre un corps quelconque.

- $(\mathbf{Z}, +, \times)$ ,  $(\mathbf{Q}, +, \times)$ ,  $(\mathbf{R}, +, \times)$  et  $(\mathbf{C}, +, \times)$  sont des anneaux intègres.
- L'anneau des polynômes à coefficients dans  $\mathbf{K}$ ,  $(\mathbf{K}[X], +, \times)$  est un anneau intègre.
- Soit  $n$  un entier naturel supérieur ou égal à 2,  $(\mathcal{M}_n(\mathbf{K}), +, \times)$  est un anneau non intègre, ce à deux titres : d'une part il n'est pas commutatif, d'autre part il n'est pas sans diviseur de zéro en effet

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}^2 = 0_n.$$

**Définition 2.1.6. — CORPS —**

On appelle corps, tout anneau  $(K, +, \times)$  tel que l'on ait :

1. l'anneau  $(K, +, \times)$  est non trivial,
2. l'anneau  $(K, +, \times)$  est commutatif,
3. tout élément de  $K$  distinct de  $0_K$  est inversible.

**Remarque** — Soit  $(K, +, \times)$  un corps. Alors  $K - \{0_K\}$  est stable par  $\times$  et  $K - \{0_K\}$  muni de la loi induite par  $\times$  est un groupe. L'exercice suivant propose une généralisation de ce résultat.

**Exercice** — Soit  $(A, +, \times)$  un anneau. Notons  $A^*$  l'ensemble des éléments inversibles de cet anneau.  $A^*$  est stable par  $\times$  et  $A^*$  muni de la loi induite par  $\times$  est un groupe.

*Solution* —

**Exemples, contre-exemples —**

- $(\mathbf{Z}, +, \times)$  est un anneau intègre mais pas un corps.
- $(\mathbf{Q}, +, \times)$ ,  $(\mathbf{R}, +, \times)$  et  $(\mathbf{C}, +, \times)$  sont des corps.
- $(\mathbf{K}[X], +, \times)$  est un anneau intègre mais pas un corps.
- $(\mathbf{K}(X), +, \times)$  est un corps.  
( $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{K} = \mathbf{C}$ .)

**Proposition 2.1.7. —** *Tout corps est un anneau intègre.*

*Preuve de la proposition 2.1.8. —* Soit  $(\mathbf{K}, +, \times)$  un corps. Par définition même d'un corps, l'anneau  $(K, +, \times)$  est commutatif et non trivial; reste à montrer qu'il est sans diviseur de zéro.

Soient  $a$  et  $b$  des éléments de  $K$  tels que :

$$a \times b = 0_K. \quad (\text{VIII.2})$$

— 1<sup>e</sup> cas :  $a = 0_K$ .

— 2<sup>nd</sup> cas :  $a \neq 0_K$ . Alors  $a$  est inversible; donc en multipliant (VIII.2) à droite par  $a^{-1}$ , on a :

$$a^{-1} \times a \times b = a^{-1} \times 0_K,$$

donc d'après 2.1.2.1,

$$b = 0_K.$$

Donc au total,  $a$  ou  $b$  est nul. L'anneau  $(K, +, \times)$  est donc intègre.  $\diamond$

### 2.1.9. — SOUS-CORPS —

On appelle sous-corps d'un corps  $(K, +, \times)$  tout sous-anneau de  $(K, +, \times)$  (considéré comme un anneau), qui est un corps.

### 2.1.10. — CARACTÉRISATION DES SOUS-CORPS —

Soient  $(K, +, \times)$  un corps et  $k$  une partie de  $K$ . Avec les abus terminologiques habituels,  $k$  est un sous-corps de  $(K, +, \times)$  si et seulement si

1.  $1_K \in k$ ,
  2.  $k$  est stable par  $+$ ,
  3.  $k$  est stable par passage à l'opposé,
  4.  $k$  est stable par  $\times$ ,
  5.  $k$  est stable par passage à l'inverse, c'est-à-dire que pour tout élément  $x$  de  $k$  non nul,  $x^{-1} \in k$ .
- (sous-anneau)

**Remarques** — On peut remplacer 1 par :  $k$  contient un élément non nul.

Définissons maintenant pour les anneaux, comme nous l'avons fait pour les groupes, une notion de morphisme.

### Définition 2.1.11. — MORPHISMES D'ANNEAUX —

Soit  $(A_1, +_1, \times_1)$  et  $(A_2, +_2, \times_2)$  des anneaux. On appelle morphisme de l'anneau  $(A_1, +_1, \times_1)$  dans l'anneau  $(A_2, +_2, \times_2)$ , toute application  $\varphi$  de  $A_1$  dans  $A_2$  telle que :

1.  $\varphi(1_{A_1}) = 1_{A_2}$ ,
2. pour tout  $x$  et tout  $y$  éléments de  $A_1$ ,  $\varphi(x +_1 y) = \varphi(x) +_2 \varphi(y)$ ,
3. pour tout  $x$  et tout  $y$  éléments de  $A_1$ ,  $\varphi(x \times_1 y) = \varphi(x) \times_2 \varphi(y)$ .

Si de plus  $\varphi$  est une bijection de  $A_1$  sur  $A_2$ , on dit que  $\varphi$  est un isomorphisme de l'anneau  $(A_1, +_1, \times_1)$  sur l'anneau  $(A_2, +_2, \times_2)$ .

Si les anneaux  $(A_1, +_1, \times_1)$  et  $(A_2, +_2, \times_2)$  sont des corps, on parle de morphismes ou d'isomorphismes de corps<sup>1</sup>.

### Proposition 2.1.12. — PROPRIÉTÉS DES MORPHISMES D'ANNEAUX —

Soit  $\varphi$  un morphisme de l'anneau  $(A_1, +_1, \times_1)$  dans l'anneau  $(A_2, +_2, \times_2)$ . Outre les propriétés résultant de ce que  $\varphi$  est un morphisme du groupe  $(A_1, +_1)$  dans le groupe  $(A_2, +_2)$ , on a :

1. Pour tout élément  $x_1$  de  $A_1$  inversible,  $\varphi(x_1^{-1}) = (\varphi(x_1))^{-1}$ .
2. l'image direct d'un sous-anneau de  $(A_1, +_1, \times_1)$  par  $\varphi$  est un sous-anneau de  $(A_2, +_2, \times_2)$ .
3. l'image réciproque d'un sous-anneau de  $(A_2, +_2, \times_2)$  par  $\varphi$  est un sous-anneau de  $(A_1, +_1, \times_1)$ .

De plus si  $(A_1, +_1, \times_1)$  et  $(A_2, +_2, \times_2)$  sont des corps donc  $\varphi$  un morphisme de corps, alors

4. Pour tout élément  $x_1$  de  $A_1$  non nul,  $\varphi(x_1^{-1}) = (\varphi(x_1))^{-1}$ .
5. l'image direct d'un sous-corps de  $(A_1, +_1, \times_1)$  par  $\varphi$  est un sous-corps de  $(A_2, +_2, \times_2)$ .
6. l'image réciproque d'un sous-corps de  $(A_2, +_2, \times_2)$  par  $\varphi$  est un sous-corps de  $(A_1, +_1, \times_1)$ .

**Exemples** — Par  $\mathbf{K}$  on désigne indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$ .

- On désigne par  $D$  un ensemble non vide et par  $a$  un élément de  $D$ . L'application

$$\varphi : \mathcal{F}(D, \mathbf{K}) \rightarrow \mathbf{K}; f \mapsto f(a)$$

est un morphisme de l'anneau  $\mathcal{F}(D, \mathbf{K})$  dans l'anneau  $(\mathbf{K}, +, \times)$ .

- Soient  $(A, +, \times)$  un anneau et  $B$  un sous-anneau de  $(A, +, \times)$ . L'injection canonique de  $B$  dans  $A$ ,

$$B \rightarrow A; x \mapsto x$$

est un morphisme de l'anneau  $B$  dans l'anneau  $A$ .

- Soit  $\mathbf{E}$  un  $\mathbf{K}$ -espace vectoriel de dimension finie  $n$  non nulle. Soit  $\mathcal{B}$  une base de  $\mathbf{E}$ . L'application

$$\text{mat}_{\mathcal{B}} : \mathcal{L}(\mathbf{E}) \rightarrow \mathcal{M}_n(\mathbf{K}); u \mapsto \text{mat}_{\mathcal{B}}(u)$$

est un isomorphisme de l'anneau  $(\mathcal{L}(\mathbf{E}), +, \circ)$  sur l'anneau  $(\mathcal{M}_n(\mathbf{K}), +, \times)$ . On dira, comme pour les groupes que ces deux anneaux sont isomorphes ; ils ont même structure. Une conséquence de l'existence de cet isomorphisme est que pour montrer un résultat portant sur  $\mathcal{L}(\mathbf{E})$ , on peut plutôt choisir de travailler sur  $\mathcal{M}_n(\mathbf{K})$ <sup>2</sup>.

1. Dans le cas de corps si  $\varphi(1_{A_1}) \neq \{0_{A_2}\}$ , la condition 1. est une conséquence de 3. (cf. propriétés des morphismes de groupes).

2. Le programme de MPSI montre même que  $\mathcal{L}(\mathbf{E})$  et  $\mathcal{M}_n(\mathbf{K})$  se munissent de structures d'algèbres et, en tant qu'algèbres sont isomorphes

- La conjugaison,

$$\varphi : \mathbf{C} \rightarrow \mathbf{C}; z \mapsto \bar{z},$$

est un isomorphisme du corps  $\mathbf{C}$  sur lui-même. Un isomorphisme d'un anneau (resp. d'un corps) sur lui-même est appelé *automorphisme* d'anneaux (resp. de corps).

## 2.2 Divisibilité dans les anneaux intègres, retours sur l'arithmétique dans $\mathbf{Z}$ et $\mathbf{K}[X]$

Dans toute ce paragraphe,  $(A, +, \times)$  désigne un anneau *intègre*.

Nous allons pour l'essentiel réviser quelques résultats d'arithmétiques vus en sup., en les éclairant d'un jour nouveau. cette partie ne reprend pas d'une façon exhaustive le cours de MPSI, qui doit être revu à cette occasion.

Commençons par une observation dans l'anneau  $(\mathbf{Z}, +, \times)$  étudié en MPSI. Soient  $a$  et  $b$  des éléments de  $\mathbf{Z}$ ,  $a$  divise  $b$  si par définition il existe un élément  $k$  de  $\mathbf{Z}$  tel que  $b = k \times a$  donc on a immédiatement que  $a$  divise  $b$  si et seulement si  $b\mathbf{Z} \subset a\mathbf{Z}$ . Cette équivalence est intéressante car la relation de divisibilité n'est pas une relation d'ordre sur  $\mathbf{Z}$  (elle n'est pas ), par contre la relation d'inclusion est une relation d'ordre sur  $\mathcal{P}(\mathbf{Z})$ . Exprimer ainsi la divisibilité offre de nombreux avantages comme nous allons le voir. Afin de généraliser cette propriété à des anneaux plus généraux, nous allons introduire une notion importante.

**Définition 2.2.1.** — IDÉAL D'UN ANNEAU INTÈGRE<sup>1</sup> —

On appelle idéal de l'anneau  $(A, +, \times)$ , toute partie  $I$  de  $A$  telle que

1.  $I$  est un sous groupe du groupe  $(A, +)$ .
2. Pour tout élément  $x$  de  $I$  et tout élément  $a$  de  $A$ ,  $a \times x \in I$ , autrement dit  $I$  est stable par multiplication par un élément de l'anneau.

**Exemple 2.2.2.** — Soit  $a$  un élément de  $A$ . On note  $aA$  (ou  $Aa$ ), l'ensemble des éléments  $x$  de  $A$  de la forme  $k \times a$ , où  $k$  est un élément quelconque de  $A$ . L'ensemble  $aA$  est un idéal de  $(A, +, \times)$  (ou plus succinctement de  $A$ ). En effet

**Remarque 2.2.3.** — l'idéal  $aA$  défini dans l'exemple précédent est même le plus petit idéal contenant  $a$ . on dit que c'est l'idéal engendré par  $\{a\}$  ou par  $a$ . On dit encore que  $a$  est un générateur de cet idéal.

**Proposition 2.2.4.** — IDÉAUX DE  $\mathbf{Z}$  —

Les idéaux de  $(\mathbf{Z}, +, \times)$  sont exactement les ensembles de la forme  $n\mathbf{Z}$ , où  $n$  est un élément de  $\mathbf{N}$

*Preuve de la proposition 2.2.4.* —

**Remarque 2.2.5.** — Soit  $n$  un élément de  $\mathbf{N}^*$ .  $n\mathbf{Z}$  admet deux et seulement deux générateurs  $n$  et  $-n$ .

Revenons au point de départ, la relation de divisibilité. On a d'abord comme dans le cas de  $\mathbf{Z}$  la définition suivante :

**Définition 2.2.6.** — Soient  $a$  et  $b$  des éléments de  $A$ . On dit que  $a$  divise  $b$  si, par définition, il existe un élément  $k$  de  $A$  tel que :

$$b = k \times a.$$

<sup>1</sup>. La définition d'un idéal peut être donné dans un anneau commutatif quelconque, mais cette notion n'a vraiment d'intérêt que dans un anneau intègre.

On dit encore que  $a$  est un diviseur de  $a$  ou que  $b$  est un multiple de  $a$ . On note  $a|b$ .

**Proposition 2.2.7.** — Soient  $a$  et  $b$  des éléments de  $A$ .  $a$  divise  $b$  si et seulement si l'idéal  $bA$  est inclus dans  $aA$ .

*Preuve de la proposition 2.2.7.* —

Donnons pour finir un résultat souvent utile pour montrer qu'un ensemble est un idéal.

**Proposition 2.2.8.** — Soient  $(B, +, \times)$  un anneau et  $\varphi$  un morphisme de l'anneau intègre  $(A, +, \times)$  dans l'anneau  $(B, +, \times)$ . Le noyau de  $\varphi$ , c'est-à-dire l'ensemble  $\varphi^{-1}(\{0_B\})$  est un idéal de  $(A, +, \times)$ .

*Preuve de la proposition 2.2.8.* —

## RETOUR SUR LE PPCM ET PGCD

Soient  $a$  et  $b$  des éléments de  $\mathbf{Z}$  on dit qu'un entier  $d$  est plus petit commun diviseur de  $a$  et  $b$  en abrégé PGCD si par définition :

1.  $d|a$  et  $d|b$  ( $d$  est un diviseur commun de  $a$  et  $b$ ).
2. Pour tout élément  $d'$  de  $\mathbf{Z}$ , si  $d'|a$  et  $d'|b$ , alors  $d'|d$  ( $d$  est multiple de tout diviseur commun de  $a$  et  $b$ ).

Lorsque  $a$ ,  $b$  et  $d$ , PGCD de  $a$  et  $b$ , sont positifs non tous deux nuls, on montre que  $d$  est bien le *plus grand diviseur commun* de  $a$  et  $b$  au sens de la relation d'ordre ordinaire sur  $\mathbf{N}$ ,  $\leq$ . Mais il est plus avisé de considérer que, par définition, il est le plus grand diviseur commun au sens de la relation de *divisibilité* qui est une relation d'ordre sur  $\mathbf{N}$ .

Nous allons reformuler cette définition en termes d'idéaux, ce qui nous permettra de montrer l'existence de PGCD et de généraliser à d'autres anneaux. Rappelons qu'en MPSI l'existence du PGCD a été montrée par l'algorithme d'Euclide.

Soient  $a$ ,  $b$  et  $d$  des éléments de  $\mathbf{Z}$ .  $d$  est PGCD de  $a$  et  $b$  si et seulement si :

1.  $a\mathbf{Z} \subset d\mathbf{Z}$  et  $b\mathbf{Z} \subset d\mathbf{Z}$
2. Pour tout élément  $d'$  de  $\mathbf{Z}$ ...

Donc  $d$  est PGCD de  $a$  et  $b$  si et seulement si  $d\mathbf{Z}$  est le plus

*Conclusion : Les PGCD de  $a$  et  $b$  sont exactement les*

Soient  $a$  et  $b$  des éléments de  $\mathbf{Z}$  on dit qu'un entier  $m$  est plus petit commun diviseur de  $a$  et  $b$  en abrégé PPCM si par définition :

1.  $a|m$  et  $b|m$  ( $m$  est un multiple commun de  $a$  et  $b$ ).
2. Pour tout élément  $m'$  de  $\mathbf{Z}$ , si  $a|m'$  et  $b|m'$ , alors  $m|m'$  ( $m$  est diviseur de tout multiple commun de  $a$  et  $b$ ).

Lorsque  $a$ ,  $b$  et  $m$ , PPCM de  $a$  et  $b$ , sont positifs, on montre que  $m$  est bien le *plus petit multiple commun* de  $a$  et  $b$  au sens de la relation d'ordre ordinaire sur  $\mathbf{N}$ ,  $\leq$ . Mais il est plus avisé de considérer que, par définition, il est le plus petit multiple commun au sens de la relation de *divisibilité* qui est une relation d'ordre sur  $\mathbf{N}$ .

Nous allons reformuler cette définition en termes d'idéaux, ce qui nous permettra de montrer l'existence de PPCM et de généraliser à d'autres anneaux.

Soient  $a$ ,  $b$  et  $m$  des éléments de  $\mathbf{Z}$ .  $m$  est PPCM de  $a$  et  $b$  si et seulement si :

1.  $m\mathbf{Z} \subset a\mathbf{Z}$  et  $m\mathbf{Z} \subset b\mathbf{Z}$
2. pour tout élément  $m'$  de  $\mathbf{Z}$ ...

Donc  $m$  est PPCM de  $a$  et  $b$  si et seulement si  $m\mathbf{Z}$  est le plus

*Conclusion : Les PPCM de  $a$  et  $b$  sont exactement les*

La preuve de l'existence de PGCD et de PPCM dans  $\mathbf{Z}$  que nous venons de donner repose sur le fait que les idéaux de  $\mathbf{Z}$  sont monogènes c'est-à-dire de la forme  $a\mathbf{Z}$ , avec  $a \in \mathbf{Z}$ , propriété que  $\mathbf{Z}$  partage avec les anneaux de polynômes comme nous allons maintenant le montrer.

**Proposition 2.2.9.** — Soit  $\mathbf{K}$  un sous corps de  $\mathbf{C}$ . On a vu que  $(\mathbf{K}[X], +, \times)$  est un anneau intègre. Alors

- Tout idéal Im de  $\mathbf{K}[X]$  est monogène, c'est-à-dire de la forme  $P\mathbf{K}[X]$  où  $P$  est un élément de  $\mathbf{K}[X]$ .
- Si  $P$  et  $P'$  sont des générateurs d'un même idéal  $\mathfrak{J}$  de  $\mathbf{K}[X]$ , alors ils sont associés, c'est-à-dire qu'il existe  $k$ , élément non nul de  $\mathbf{K}$ , tel que  $P=kP'$ .
- Tout idéal de  $\mathbf{K}[X]$  non réduit à  $\{0\}$  admet un et un seul générateur unitaire.



*Preuve de la proposition 2.2.9. —*

Donnons une nouvelle définition du PPCM et PGCD.

Dans la suite  $(P, +, \times)$  désigne indifféremment l'anneau  $(\mathbf{Z}, +, \times)$  ou l'anneau  $(\mathbf{K}[X], +, \times)$ ,  $\mathbf{K}$  désignant un sous corps de  $\mathbf{C}$ .

Tout ce qui suit se généralise à tout anneau intègre dont les idéaux sont monogènes (anneau principal), toutefois le programme nous limite aux seuls cas de  $(\mathbf{Z}, +, \times)$  ou de  $(\mathbf{K}[X], +, \times)$ .

**Proposition 2.2.10.** — Soient  $a$  et  $b$  des éléments de  $P$ . alors  $aP \cap bP$  et  $aP + bP$  sont des idéaux de  $(P, +, \times)$ .

*Preuve de la proposition 2.2.10.* —

**Définition 2.2.11.** — Soient  $a$  et  $b$  des éléments de  $P$ .

- On appelle plus grand commun diviseur de  $a$  et  $b$ , en abrégé PGCD, tout générateur de  $aP + bP$ .
- On appelle plus petit commun multiple de  $a$  et  $b$ , en abrégé PPCM, tout générateur de  $aP \cap bP$ .

**Notations** —

- Par convention l'orsque l'on parle DU PGCD de  $a$  et  $b$ , il s'agit lorsque  $aP + bP \neq \{0\}$ , dans le cas de  $\mathbf{Z}$  du générateur strictement positif de  $a\mathbf{Z} + b\mathbf{Z}$ , dans le cas de  $\mathbf{K}[X]$  du générateur unitaire de  $a\mathbf{K}[X] + b\mathbf{K}[X]$ .
- Par convention l'orsque l'on parle DU PPCM de  $a$  et  $b$ , il s'agit lorsque  $aP \cap bP \neq \{0\}$ , dans le cas de  $\mathbf{Z}$  du générateur strictement positif de  $a\mathbf{Z} \cap b\mathbf{Z}$ , dans le cas de  $\mathbf{K}[X]$  du générateur unitaire de  $a\mathbf{K}[X] \cap b\mathbf{K}[X]$ .
- LE PGCD de  $a$  et de  $b$  se note  $\text{PGCD}(a, b)$  ou  $a \wedge b$ .
- LE PPCM de  $a$  et de  $b$  se note  $\text{PPCM}(a, b)$  ou  $a \vee b$ .

Comme le montre l'introduction dans le cas de  $\mathbf{Z}$  on a les deux propositions suivantes.

**Proposition 2.2.12.** — Soient  $a, b$  et  $d$  des éléments de  $P$ . Alors  $d$  est PGCD de  $a$  et  $b$  si et seulement si :

1.  $d|a$  et  $d|b$  ( $d$  est un diviseur commun de  $a$  et  $b$ ).
2. Pour tout élément  $d'$  de  $\mathbf{Z}$ , si  $d'|a$  et  $d'|b$ , alors  $d'|d$  ( $d$  est multiple de tout diviseur commun de  $a$  et  $b$ ).

**Proposition 2.2.13.** — Soient  $a, b$  et  $d$  des éléments de  $P$ . Alors  $d$  est PPCM de  $a$  et  $b$  si et seulement si :

1.  $a|m$  et  $b|m$  ( $m$  est un multiple commun de  $a$  et  $b$ ).
2. Pour tout élément  $m'$  de  $\mathbf{Z}$ , si  $a|m'$  et  $b|m'$ , alors  $m|m'$  ( $m$  est diviseur de tout multiple commun de  $a$  et  $b$ ).

**Remarque** — On peut généraliser ce qui précède.

- On peut définir du PGCD d'une famille finie d'éléments de  $P$ ,  $(a_1, a_2, \dots, a_n)$ , ( $n \geq 2$ ), comme étant les générateurs de  $a_1P + a_2P + \dots + a_nP$ . On montre que ce sont des diviseurs communs de  $a_1, a_2, \dots, a_n$ , multiples de tout autre diviseur commun. La convention qui définit LE PGCD de deux éléments s'étant à un nombre quelconque.
- On peut définir les PPCM d'une famille finie d'éléments de  $P$ ,  $(a_1, a_2, \dots, a_n)$ , ( $n \geq 2$ ), comme étant les générateurs de  $a_1P \cap a_2P \cap \dots \cap a_nP$ . On montre que ce sont des multiples communs de  $a_1, a_2, \dots, a_n$ , diviseurs de tout autre multiple commun. La convention qui définit LE PPCM de deux éléments s'étant à un nombre quelconque.

**Proposition 2.2.14.** — Soient  $a$  et  $b$  des éléments de  $P$ . On dit que  $a$  et  $b$  sont premiers entre eux si, par définition,  $\text{PGCD}(a, b) = 1$

Soient  $a_1, a_2, \dots, a_n$ , ( $n \geq 2$ ), des éléments de  $P$ . On dit qu'ils sont premiers dans leur ensemble si, par définition,  $(a_1, a_2, \dots, a_n)$  admet 1 comme PGCD. On dit qu'ils sont premiers entre eux deux à deux si, par définition, pour tout couple  $(i, j)$  d'éléments distincts de  $\{1, 2, \dots, n\}$ ,  $\text{PGCD}(a_i, a_j) = 1$ .

Il a été vu en MPSI comment calculer le PGCD et le PPCM de deux éléments de  $P$  grâce à une décomposition en facteurs irréductibles de ces deux éléments. Le lecteur est prié de s'y reporter.

Apprécions maintenant la force de cette nouvelle formulation du PGCD, en revisitant des résultats de sup.

**Proposition 2.2.15.** — THÉORÈME DE BEZOUT —

Soient  $a$  et  $b$  des éléments de  $P$ . Alors  $a$  et  $b$  sont premiers entre eux si et seulement si, il existe des éléments  $u$  et  $v$  de  $P$  tels que :

$$au + bv = 1.$$

Pour le calcul pratique des éléments  $u$  et  $v$  on utilise l'algorithme d'Euclide étendu, voir TD info.

L'algorithme d'Euclide, outre son intérêt pratique, montre l'invariance du PGCD par passage à un sur corps.

*Preuve de la proposition 2.2.15.—*

**Proposition 2.2.15.** — THÉORÈME DE GAUSS —

*Soient  $a, b$  et  $c$  des éléments de  $P$ . On suppose que  $a$  est premier avec  $b$ . Si  $a$  divise le produit  $bc$ . Alors  $a$  divise  $c$*

*Preuve de la proposition 2.2.15*

**Proposition 2.2.16.** — *Soient  $a, b_1, b_2, \dots, b_n$  des éléments de  $P$  ( $n \in \mathbb{N}^*$ ). Si pour  $i = 1, 2, \dots, n$ ,  $a$  est premier avec  $b_i$ , alors  $a$  est premier avec le produit  $b_1 b_2 \dots b_n$ .*

*Preuve de la proposition 2.2.16 —*

**Exercice** — Soient  $a$  et  $b$  des éléments de  $P$  premiers entre eux. Soient  $u_0$  et  $v_0$  des éléments de  $P$  tels que :

$$u_0a + v_0b = 1_P,$$

éléments dont l'existence est assurée par le lemme de Bezout.

Déterminer tous les couples  $(u, v)$  d'éléments de  $P$  tels que :

$$ua + vb = 1$$

au moyen de  $(u_0, v_0)$ .

*Solution —*

**Exercice** — ÉQUATION DIOPHANTINNE — Soit  $a, b$  et  $c$  des entiers relatifs. On se propose de résoudre dans  $\mathbf{Z}^2$  l'équation d'inconnue  $(x, y)$

$$ax + by = c. \quad (\text{VIII.3})$$

1. On désigne par  $d$  le PGCD (plus grand commun diviseur) de  $a$  et de  $b$ . Montrer que si l'équation (VIII.3) admet une solution, alors  $d$  divise  $c$ .
2. On suppose dans cette question que  $d$  divise  $c$ , et l'on note  $a', b'$  et  $c'$  les entiers relatifs définis par :

$$a = a'd, \quad b = b'd, \quad c = c'd.$$

Montrer que l'équation (VIII.3) admet au moins une solution notée  $(x_0, y_0)$ .

3. On suppose toujours que  $d$  divise  $c$ . Donner toutes les solutions de l'équation (VIII.3) en fonction de  $(x_0, y_0)$ .

*Solution —*

Nous allons donner dans  $P$  la notion d'élément irréductible, et revoir les théorème de décomposition en produits d'irréductibles.

**Définition 2.2.16.** — Soient  $a$  et  $b$  des éléments de  $P$ . On dit que  $a$  est associé à  $b$  si il existe  $u$  élément inversible de  $P$  tel que  $a = ub$ .

La relation sur  $P$  « être associés à » est une relation d'équivalence, on peut donc employer des expressions du type :  $a$  et  $b$  sont associés.

Cas  $P = \mathbf{Z}$  —

Tout élément  $a$  de  $\mathbf{Z}$  est associé à lui-même et à son opposé et à aucun autre élément.

Cas  $P = \mathbf{K}[X]$  —

Pour tout élément  $A$  de  $\mathbf{K}[X]$ , l'ensemble des polynômes associés à  $A$  est :

$$\{uA, u \in \mathbf{K}^*\}.$$

Donnons une définition de l'irréductibilité valable du reste dans tout anneau intègre.

**Définition 2.2.17.** — *Un élément  $a$  de  $P$  est dit irréductible si :*

1. *L'élément  $a$  n'est pas inversible ;*
2. *Si il existe  $b$  et  $c$  éléments de  $P$  tels que  $a = bc$ , alors  $b$  ou  $c$  est inversible (ou si l'on préfère  $c$  ou  $b$  est associé à  $a$ ).*

Cas  $P = \mathbf{Z}$  —

un entier  $a$  est irréductible si il est différent de 1 et de  $-1$  et si il n'admet pas de diviseur autres que  $\pm 1$  et  $\pm a$ , autrement dit  $a$  est irréductible si et seulement si il admet exactement deux diviseurs positifs. Ainsi 0 n'est pas irréductible, 1 n'est pas irréductible,  $-3$  est irréductible, 12 n'est pas irréductible.

Cas  $P = \mathbf{K}[X]$  —

Un polynôme  $Q$  de  $\mathbf{K}[X]$  est irréductible si et seulement si il est non constant et si ses seuls diviseurs sont les polynômes qui lui sont associés et les éléments de  $\mathbf{K}^*$ .

Dans  $\mathbf{Z}$  on appelle nombre premier tout élément irréductible positif. Ainsi tout irréductible de  $\mathbf{Z}$  est associé à un et un seul nombre premier.

Passons à  $\mathbf{K}[X]$ . Le polynôme nul n'étant pas irréductible, tout polynôme élément de  $\mathbf{K}[X]$ , irréductible, est associé à un et seulement à un polynôme unitaire (nécessairement irréductible). On note dans la suite  $\mathcal{P}$  l'ensemble des éléments de  $\mathbf{K}[X]$  irréductibles et unitaire.

Passons à la décomposition dans  $\mathcal{P}$  en produit d'irréductibles.

Il a été vu l'an passé :

**Proposition 2.2.18.** — *Tout élément de  $\mathbf{Z}$  non nul peut au signe près s'écrire comme le produit de nombres premiers, cette écriture étant unique, à l'ordre près de ses facteurs.*

Nous renvoyons au cours de sup. pour la preuve, ainsi que pour la définition de la valuation  $p$ -adique d'un entier. Nous allons maintenant rappeler l'énoncé similaire pour  $\mathbf{K}[X]$ .

**Proposition 2.2.19.** — DÉCOMPOSITION EN FACTEURS IRREDUCTIBLES —

*Soit  $A$  un élément de  $\mathbf{K}[X]$  non nul. Il existe un unique couple  $(u, \nu)$  constitué d'un élément  $u$  non nul de  $\mathbf{K}$  (un élément inversible de  $\mathbf{K}[X]$ ) et d'une application  $\nu$  de  $\mathcal{P}$  dans  $\mathbf{N}$ , presque nulle, tel que*

$$A = u \prod_{P \in \mathcal{P}} P^{\nu(P)} \quad (\text{VIII.4})$$

Pratiquement si  $A$  est un polynôme non constant, il s'écrit :

$$A = u P_1^{n_1} P_2^{n_2} \dots P_k^{n_k}, \quad (\text{VIII.5})$$

où les  $P_i$ ,  $i = 1, \dots, k$ , étant des éléments distincts de  $\mathcal{P}$  et  $n_1, \dots, n_k$  des entiers supérieurs ou égaux à 1.

Preuve de 2.1.19. — La preuve repose sur le lemme suivant

**Lemme** — *Tout polynôme non constant  $P$  de  $\mathbf{K}[X]$  admet un diviseur irréductible.*

*Preuve du lemme* — Soit  $P$  un élément non constant de  $\mathbf{K}[X]$ . Notons  $D$  l'ensemble des diviseurs de  $P$  non constants....



Nous allons dans les cas particulier où  $\mathbf{K}$  est  $\mathbf{C}$  ou  $\mathbf{R}$  rappeler la forme des irréductibles de  $\mathbf{K}[X]$ .

**Proposition 2.2.20.** — *Les éléments irréductibles de  $\mathbf{C}[X]$  sont les polynômes de degré 1.*

**Proposition 2.2.21.** — *Les éléments irréductibles de  $\mathbf{R}[X]$ .....*

**Exercice** — Décomposer en produit d'irréductibles les éléments de  $\mathbf{R}[X]$  suivants :

1.  $X^4 + 3X^2 - 4$
2.  $X^4 + 1, X^4 + X^2 + 1$

*Solution* —

**Exercice** — Soit  $P$  un élément de  $\mathbf{R}[X]$  tel que pour tout élément  $x$  de  $\mathbf{R}$ ,  $P(x) \geq 0$ . Montrer qu'il existe  $A$  et  $B$  éléments de  $\mathbf{R}[X]$  tels que :

$$P = A^2 + B^2.$$

*Solution* —

**Exercice** — Montrer que  $X^3 + X + 1$  n'admet pas de racine rationnelle. En déduire que ce polynôme est irréductible dans  $\mathbf{Q}[X]$ .



### 2.3 L'anneau $\mathbf{Z}/n\mathbf{Z}$

Dans ce paragraphe  $n$  désignera un entier supérieur ou égal à 1.

Nous allons maintenant offrir à  $\mathbf{Z}/n\mathbf{Z}$  une multiplication dans le but affirmé d'en faire un anneau. Comme dans le cas l'addition, il serait agréable que le produit de deux éléments  $\bar{x}$  et  $\bar{y}$  fût  $\overline{xy}$ . Toutefois avant de définir ainsi le produit sur  $\mathbf{Z}/n\mathbf{Z}$ , il convient de s'assurer que l'élément  $\overline{xy}$  ne dépend que des deux éléments de  $\mathbf{Z}/n\mathbf{Z}$  dont on effectue le produit et non de leurs représentants particuliers  $x$  et  $y$ . Ces préventions vont être écartées par le lemme qui suit.

**Lemme 2.3.1.** — Soient  $x, x', y$  et  $y'$  des éléments de  $\mathbf{Z}$ . On suppose que :

$$x \equiv x' [n] \text{ et } y \equiv y' [n].$$

Alors

$$xy \equiv x'y' [n].$$

*Preuve du lemme 2.3.1* — Par hypothèse il existe  $k$  et  $h$ , éléments de  $\mathbf{Z}$ , tels que :

$$x' = x + kn, \quad y' = y + hn.$$

Par multiplication de ces égalités :  $x'y' = xy + (hkn + hx + ky)n$  et donc  $xy \equiv x'y' [n]$ . □

**Définition 2.3.2.** — Soit  $a$  et  $b$  des éléments de  $\mathbf{Z}/n\mathbf{Z}$ . On appelle produit de  $a$  par  $b$  noté  $a \times_n b$  l'élément de  $\mathbf{Z}/n\mathbf{Z}$ ,  $\overline{x \times_n y}$ , où  $x$  et  $y$  sont des représentants respectifs de  $a$  et  $b$ , cette quantité, d'après le lemme 2.3.1 ne dépend pas du choix des représentants.

Dans la suite on notera souvent, lorsqu'il n'y a pas d'ambiguïté, la loi  $a \times_n b$  simplement  $\times$ .

On retient

$$\boxed{\overline{x \times_n y} = \overline{x \times y}}$$

**Exemple** —

- Dans  $\mathbf{Z}/6\mathbf{Z}$ ,  
 $\bar{4} \times \bar{3} = \dots$   
 $\bar{2} \times \bar{5} = \dots$
- Dans  $\mathbf{Z}/11\mathbf{Z}$ ,  
 $\bar{4} \times \bar{8} = \dots$   
 $\bar{2} \times \bar{10} = \dots$

La loi  $\times_n$  est une loi interne sur  $\mathbf{Z}/n\mathbf{Z}$  et l'on a le résultat de structure suivant :

**Proposition 2.3.3.** —  $\left(\mathbf{Z}/n\mathbf{Z}, +_n, \times_n\right)$  est un anneau commutatif.

*Preuve de la proposition 2.3.3.* — Donnons une preuve succincte.

Par définition de la multiplication  $\times_n$  sur  $\mathbf{Z}/n\mathbf{Z}$ , on a le résultat suivant :

**Proposition 2.3.4.** — *la surjection canonique de  $\mathbf{Z}$  sur  $\mathbf{Z}/n\mathbf{Z}$ ,*

$$\pi_n : \mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z}; x \mapsto \bar{x},$$

*est...*

**Remarque** — Nous verrons qu'en général l'anneau  $(\mathbf{Z}/n\mathbf{Z}, +, \times)$  n'est pas intègre :

**Application** —

La compatibilité des congruences avec l'addition et de la multiplication donne, par exemple, des critères de divisibilité.

Soit  $x$  un élément de  $\mathbf{N}$ , d'écriture décimale  $x = a_n a_{n-1} \dots a_1 a_0$ , c'est-à-dire que

$$x = \sum_{i=0}^n a_i 10^i.$$

#### 1. DIVISIBILITÉ PAR 9

*L'entier  $x$  est divisible par 9 si et seulement si*

## 2. DIVISIBILITÉ PAR 3

*L'entier  $x$  est divisible par 3 si et seulement si*

## 3. DIVISIBILITÉ PAR 5

*L'entier  $x$  est divisible par 5 si et seulement si*

## 4. DIVISIBILITÉ PAR 11

*L'entier  $x$  est divisible par 11 si et seulement si*

## 5. DIVISIBILITÉ PAR 6

*L'entier  $x$  est divisible par 6 si et seulement si*

**Exercice —**

1. Trouver le dernier chiffre dans l'écriture décimale de  $7^{(7^7)}$ .
2. Soient  $a$ ,  $b$  et  $c$  des entiers naturels tels que  $7|(a^3 + b^3 + c^3)$  Montrer que  $7|abc$ .

*Solution* —

GÉNÉRATEURS DE  $(\mathbf{Z}/n\mathbf{Z}, +)$

Nous savons que  $(\mathbf{Z}/n\mathbf{Z}, +)$  est un groupe cyclique. Quels en-sont les générateurs ?

Nous allons commencer par un exemple : on prend  $n = 6$ .

$$\begin{aligned}\langle \bar{0} \rangle &= \{\bar{0}\}. \\ \langle \bar{1} \rangle &= \{ \quad \quad \quad \}, \\ \langle \bar{2} \rangle &= \{ \quad \quad \quad \}, \\ \langle \bar{3} \rangle &= \{ \quad \quad \quad \}, \\ \langle \bar{4} \rangle &= \{ \quad \quad \quad \}, \\ \langle \bar{5} \rangle &= \{ \quad \quad \quad \}.\end{aligned}$$

Conclusion :

Plus généralement

**Proposition 2.3.5.** — *Supposons  $n \geq 1$ . Soit  $p$  un élément de  $\mathbf{Z}$ . Les trois propositions suivantes sont équivalentes*

- i.  $\bar{p}$  est un générateur du groupe  $(\mathbf{Z}/n\mathbf{Z}, +)$ .
- ii.  $\bar{p}$  est un élément inversible de l'anneau  $(\mathbf{Z}/n\mathbf{Z}, +, \times)$ .
- iii.  $p$  et  $n$  sont premiers entre eux.

*Preuve de la proposition 2.3.5. —*

**Définition 2.3.6.** — Pour tout élément  $m$  de  $\mathbf{N}^*$  on note  $\varphi(m)$  le nombre d'éléments de  $\{1, 2, \dots, m\}$  qui sont premiers avec  $m$ . L'application

$$\varphi : \mathbf{N}^* \rightarrow \mathbf{N}; m \mapsto \varphi(m)$$

est appelée indicatrice d'Euler.

La proposition 2.3.5. nous dit donc que pour tout élément  $m$  de  $\mathbf{N}^*$ ,  $\varphi(m)$  est le nombre d'éléments inversibles de  $\mathbf{Z}/m\mathbf{Z}$ , et aussi le nombre de générateurs de  $\mathbf{Z}/m\mathbf{Z}$ .

D'après la proposition précédente on a :

**Proposition 2.3.7.** — Pour tout élément  $m$  de  $\mathbf{N}^*$ ,  $\varphi(m)$  est le nombre d'éléments inversibles de l'anneau du  $(\mathbf{Z}/m\mathbf{Z}, +, \times)$  et le nombre de générateurs du groupe  $(\mathbf{Z}/m\mathbf{Z}, +)$ .

**Proposition 2.3.8.** — On a l'équivalence des trois propositions suivantes.

- i.  $(\mathbf{Z}/n\mathbf{Z}, +, \times)$  est un anneau intègre.
- ii.  $(\mathbf{Z}/n\mathbf{Z}, +, \times)$  est un corps.
- iii.  $n$  est un nombre premier.

*Preuve de la proposition 2.3.8. —*

Nous allons donner une généralisation du petit théorème de Fermat.

**Proposition 2.3.9. — THÉORÈME D'EULER —**

*Soit  $x$  un entier premier avec  $n$ ,*

$$x^{\varphi(n)} \equiv 1 [n].$$

Dans le cas où  $n$  est un nombre premier, la condition  $x$  premier avec  $n$  devient :  $n$  ne divise pas  $x$ . La conclusion est alors :

$$x^{n-1} \equiv 1 [n].$$

On retrouve le petit théorème de Fermat. En effet.....

*Preuve de la proposition 2.3.9. —*

*Première preuve :*

*Seconde preuve :*

Passons au Théorème Chinois. Nous avons vu dans la partie sur les groupes que  $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$  et  $\mathbf{Z}/8\mathbf{Z}$  sont non isomorphes. Par contre nous allons voir que par exemples, les groupes que  $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$  et  $\mathbf{Z}/6\mathbf{Z}$  sont isomorphes, non seulement en tant que groupes mais aussi en tant qu'anneaux.

**Proposition 2.3.10.** — Théorème chinois —

*Soient  $p$  et  $q$  des entiers positifs non nuls et premiers entre eux. Alors l'anneau  $\mathbf{Z}/pq\mathbf{Z}$  et l'anneau produit  $\mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/q\mathbf{Z}$  sont isomorphes.*

**Preuve de 2.3.9**

Pour commencer considérons l'application  $\Phi : \mathbf{Z} \rightarrow \mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/q\mathbf{Z}; x \mapsto (\bar{x}^p, \bar{x}^q)$ , en des termes plus brutaux  $\Phi = (\pi_p, \pi_q)$ . Cette dernière forme nous assure que  $\Phi$  est un .....

La preuve aurait pu aussi reposer sur le résultat classique qui suit.

**Exercice 2.3.11.** — Lemme des restes chinois —

Soient  $p$  et  $q$  des entiers positifs et non nuls, premiers entre eux. Pour tout couple d'entier  $(a, b)$ , le système d'équation d'inconnue entière  $x$ ,

$$\begin{cases} x \equiv a \ [p], \\ x \equiv b \ [q], \end{cases} \quad (\text{VIII.6})$$

admet une solution. De plus si  $x_0$  est une solution de (VIII.6), alors l'ensemble des solutions de (VIII.6) est :  $x_0 + (pq)\mathbf{Z}$ .

*Solution de l'exercice 2.3.11.* —

**Remarque** — Le théorème chinois peut s'étendre par récurrence à des produits quelconque : Soient  $q_1, q_2, \dots, q_k$  des éléments de  $\mathbf{N}^*$  premiers deux à deux. Alors en notant  $m$  leur produit : l'anneau  $\mathbf{Z}/m\mathbf{Z}$  est isomorphe à l'anneau produit  $\mathbf{Z}/q_1\mathbf{Z} \times \mathbf{Z}/q_2\mathbf{Z} \times \dots \times \mathbf{Z}/q_k\mathbf{Z}$ .

Ce résultat peut aussi se prouver à partir d'une version plus générale du lemme des restes chinois que nous proposons en exercice.

**Exercice** — Soit  $m$  un entier supérieur ou égal à 2. Soit  $(a_1, a_2, \dots, a_m)$  un  $m$ -uplet d'entiers strictement positifs premiers entre eux deux à deux, soit  $(x_1, x_2, \dots, x_m)$  une famille d'entiers. Nous nous proposons de donner une méthode de résolution dans  $\mathbf{Z}$  du système d'inconnue  $x$  suivant<sup>1</sup> :

$$x = x_i \ [a_i], \ i = 1, 2, \dots, m. \quad (\text{VIII.7})$$

1. Montrer que pour tout entier  $i$  élément de  $\{1, 2, \dots, m\}$ , le système d'inconnue  $x$  :

$$\begin{cases} x \equiv a_i \ [a_i], \\ x \equiv 0 \ [a_1 \times \dots \times a_{i-1} \times a_{i+1} \times \dots \times a_m], \end{cases} \quad (\text{VIII.8})$$

a des solutions éléments de  $\mathbf{Z}$ . Comment en déterminer une ?

2. Notons pour tout entier  $i$  élément de  $\{1, 2, \dots, m\}$ ,  $s_i$  une solution du système (VIII.8). Donner une solution  $x_0$  du système de départ (VIII.7) en fonction de  $s_1, s_2, \dots, s_m$ . Donner toutes les solutions de (VIII.7).
3. Résoudre dans  $\mathbf{Z}$ ,

$$\begin{cases} x \equiv 5 \ [8], \\ x \equiv 23 \ [27], \\ x \equiv 10 \ [11]. \end{cases}$$

1. Ce problème fut à l'origine étudié par les astologues chinois.  $x_i$  représente la date d'apparition d'une planète  $P_i$  dans une direction donnée,  $a_i$  sa période, on cherche la date  $x$  où les  $m$  planètes  $P_1, P_2, \dots, P_m$  seront alignées.



*Solution* —

Nous nous proposons pour finir de donner l'expression de la valeur l'indicatrice d'Euler en un entier  $n$  au moyen des facteurs premiers de  $n$ .

Ce résultat repose sur la proposition suivante :

**Proposition 2.3.12.** — Soit  $q_1$  et  $q_2$  des éléments de  $\mathbf{N}^*$  premiers entre eux. Alors :

$$\varphi(q_1 q_2) = \varphi(q_1) \varphi(q_2).$$

*Preuve de la proposition 2.3.12.* —

Examinons alors quelques cas particuliers.

- Soit  $p$  un nombre premier Alors :

$$\varphi(p) = \dots\dots\dots$$

En effet

- Soit  $p$  un nombre premier et  $\alpha$  en entier strictement positif. Alors :

$$\varphi(p^\alpha) = \dots\dots\dots$$

En effet

Donc si  $n$  se décompose en produite de nombres premiers :  $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$  où  $p_1, \dots, p_k$  sont des nombres premiers et  $\alpha_1, \dots, \alpha_k$  des entiers strictement positifs, alors...

On a prouvé :

**Proposition 2.3.9.** — *Si  $n$  admet la décomposition en nombres premiers  $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$  où  $p_1, \dots, p_k$  sont des nombres premiers et  $\alpha_1, \dots, \alpha_k$  des entiers strictement positifs, alors*

$$\varphi(n) = n \prod_{i=1}^k \left(1 - \frac{1}{p_i}\right).$$

# ALGÈBRES

## 3.1 Définition

Dans ce paragraphe  $\mathbf{K}$  désigne un corps. Dans la pratique le programme semble se limiter aux seuls cas où  $\mathbf{K}$  est un sous corps de  $\mathbf{C}$

Nous allons maintenant revenir sur la structure d'algèbre déjà entrevue à diverses reprises. C'est la structure la plus complète et la plus naturelle dans laquelle on puisse calculer, une algèbre est à la fois un espace vectoriel et un anneau, avec une condition de compatibilité entre la multiplication de l'anneau et l'opération du corps de base sur l'espace vectoriel. Précisément.

**Définition 3.1.1.** — ALGÈBRE —

On appelle algèbre sur le corps  $\mathbf{K}$  (en abrégé  $\mathbf{K}$ -algèbre), tout quadruplet  $(\mathbf{E}, +, \times, \cdot)$ ,

où  $\mathbf{E}$  est un ensemble,  $+$  et  $\times$  des lois internes sur  $\mathbf{E}$ ,  $\cdot$  une opération de l'ensemble  $\mathbf{K}$  sur  $\mathbf{E}$ , tel que :

1.  $(\mathbf{E}, +, \times)$  soit un anneau.
2.  $(\mathbf{E}, +, \cdot)$  soit un espace vectoriel sur  $\mathbf{K}$ .
3. Pour tout  $\alpha$  élément  $\mathbf{K}$ , tout  $\mathbf{x}$  et tout  $\mathbf{y}$  éléments de  $\mathbf{E}$ .

$$(\alpha \cdot \mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\alpha \cdot \mathbf{y}) = \alpha \cdot (\mathbf{x} \times \mathbf{y}). \quad (\text{VIII.9})$$

Si de plus la loi  $\times$  est commutative on dit que l'algèbre est commutative.

Lorsque pour une algèbre  $(\mathbf{E}, +, \times, \cdot)$ , l'espace vectoriel sous jacent  $(\mathbf{E}, +, \cdot)$  est de dimension finie  $n$ , on dit que l'algèbre est de dimension finie  $n$ .

- Comme pour les autres structures lorsque les lois vont de soi on parle de « l'algèbre  $\mathbf{H}$  ».
- Comme dans les anneaux la loi «  $\times$  » peut être omise et comme dans les espaces vectoriels on ne note pas la loi «  $\cdot$  ». La propriété (VIII.9) fait alors que l'on note en général l'élément de  $\mathbf{K}$  devant les éléments de  $\mathbf{H}$  :  $\alpha \mathbf{x} \mathbf{y}$  pour  $(\alpha \cdot \mathbf{x}) \times \mathbf{y}$

Donnons des exemples qui peuvent être réutilisés sans démonstration.

**Exemples 3.1.2.** —

- Il a été vu en MPSI que  $(\mathbf{K}, +, \times)$  peut être muni naturellement d'une structure de  $\mathbf{K}$ -espace vectoriel, le produit d'un scalaire  $\lambda$  élément de  $\mathbf{K}$  par un vecteur  $k$  de  $\mathbf{K}$ ,  $\lambda \cdot k$  étant alors simplement le produit ordinaire de  $\lambda$  par  $\mathbf{K}$  :

$$\lambda \cdot k = \lambda \times k.$$

$(\mathbf{K}, +, \cdot)$  est alors un espace vectoriel,  $(\mathbf{K}, +, \times)$  un corps donc un anneau commutatif, c'est aussi une  $\mathbf{K}$ -algèbre commutative la propriété (VIII.9) résulte directement de la commutativité et l'associativité de  $\times$ .

- Une situation un peu plus complexe est obtenue lorsque l'on dispose d'un sous corps de  $\mathbf{K}$ .  $\mathbf{K}$  se munit alors d'une structure de -espace vectoriel, l'opération «  $\cdot$  » de l'ensemble sur  $\mathbf{K}$  étant la restriction à  $\mathbf{L} \times \mathbf{K}$  de «  $\times$  », multiplication du corps  $\mathbf{K}$ . On a alors comme dans le point précédent que  $(\mathbf{K}, +, \cdot)$  est une algèbre commutative. Ainsi,  $\mathbf{C}$  peut être vu comme une  $\mathbf{R}$ -algèbre de dimension 2,  $\mathbf{R}$  comme une  $\mathbf{Q}$ -algèbre de dimension non finie.
- En MPSI l'ensemble  $\mathbf{K}[X]$  des polynômes à coefficients dans le corps  $\mathbf{K}$  (c'est-à-dire l'ensemble des suites à valeurs dans  $\mathbf{K}$  à support fini) a été muni d'une addition (l'addition ordinaire des suites) et d'une multiplication un peu particulière (produit de convolution). Il a été vu alors que  $(\mathbf{K}[X], +, \times)$  était un anneau. Par ailleurs,  $\mathbf{K}[X]$  peut, et cela a été vu en sup., être muni d'une structure de  $\mathbf{K}$ -espace vectoriel, l'opération de  $\mathbf{K}$  sur  $\mathbf{K}[X]$  est du reste induite par la multiplication d'une suite à valeur dans  $\mathbf{K}$  par un élément de  $\mathbf{K}$ . On peut cette année se convaincre sans mal que  $(\mathbf{K}[X], +, \times, \cdot)$  est une  $\mathbf{K}$ -algèbre commutative. Cette algèbre n'est pas de dimension finie. Une base en est la famille  $(\delta_p)_{p \in \mathbf{N}}$ , où pour tout  $p \in \mathbf{N}$ ,

$$\delta_p = (\delta_{p,n})_{n \in \mathbf{N}} = (\underbrace{0, 0, 0, \dots, 0}_{p+1}, 1, 0, \dots).$$

Traditionnellement pour des raisons historiques et à cause du lien fort entre polynôme et fonction polynôme, on note  $\delta_1, X$ . Alors la définition de la multiplication conduit pour tout  $p \in \mathbf{N}$ , à :  $X^p = \delta_p$ . Le polynôme  $P = (a_n)_{n \in \mathbf{N}}$  s'écrit donc :

$$P = \sum_{n=0}^{+\infty} a_n X^n,$$

ou encore  $P = \sum_{n=0}^d a_n X^n$ , où  $d$  désigne le degré de  $P$ , quand ce dernier n'est pas nul, c'est-à-dire

$$d = \max\{n \in \mathbf{N} \mid a_n \neq 0\}$$

et où  $d = 0$ , par exemple, quand  $P = 0$ .

- Le cours de MPSI montre également que  $(\mathcal{M}_n(\mathbf{K}), +, \times)$  est un anneau et que  $(\mathcal{M}_n(\mathbf{K}), +, \cdot)$  est un  $\mathbf{K}$ -espace vectoriel. La vérification de la propriété (VIII.9), pour  $(\mathcal{M}_n(\mathbf{K}), +, \times, \cdot)$  est une simple formalité qui fait de  $(\mathcal{M}_n(\mathbf{K}), +, \times, \cdot)$  une  $\mathbf{K}$ -algèbre, en général non commutative. Sa dimension est  $n^2$ .
  - De même le cours de sup. nous livre une autre  $\mathbf{K}$ -algèbre :  $(\mathcal{L}(\mathbf{E}), +, \circ, \cdot)$ , où  $\mathbf{E}$  est un  $\mathbf{K}$  espace vectoriel. Elle est en général non commutative et si  $\mathbf{E}$  est de dimension finie  $n$ , alors elle est de dimension  $n^2$ , puisqu'alors l'espace vectoriel  $(\mathcal{L}(\mathbf{E}), +, \cdot)$  est isomorphe à  $(\mathcal{M}_n(\mathbf{K}), +, \cdot)$  par l'application qui à un endomorphisme de  $\mathbf{E}$  associe sa matrice dans une base  $\mathcal{B}$  donnée.
  - Soit  $D$  un ensemble. L'ensemble des applications de  $D$  dans  $\mathbf{K}$ ,  $\mathcal{F}(D, \mathbf{K})$ , parfois aussi noté  $\mathbf{K}^D$  muni de l'addition et de la multiplication usuelles des applications, est un anneau commutatif. Par ailleurs muni de l'addition et de la multiplication usuelle d'une application par un élément de  $\mathbf{K}$ , c'est un espace vectoriel. On montre immédiatement que  $(\mathcal{F}(D, \mathbf{K}), +, \times, \cdot)$ , est une  $\mathbf{K}$ -algèbre (en général de dimension infinie). Ce résultat s'étend à  $\mathcal{F}(D, \mathbf{E})$ , où  $\mathbf{E}$  est une  $\mathbf{K}$ -algèbre.
- Un cas particulier est l'algèbre des suites à valeurs dans  $\mathbf{K}$ ,  $(\mathcal{F}(\mathbf{N}, \mathbf{K}), +, \times, \cdot)$ .

### 3.2 Sous-algèbre

Comme pour les groupes, anneaux, espaces-vectoriels, on définit une notion de sous-structure pour les algèbres.

**Définition 3.2.1.** — SOUS-ALGÈBRE —

Soit  $(\mathbf{E}, +, \times, \cdot)$  une algèbre sur  $\mathbf{K}$ . On appelle sous-algèbre de  $(\mathbf{E}, +, \times, \cdot)$ , tout quadruplet  $(\mathbf{H}, +_{\mathbf{H}}, \times_{\mathbf{H}}, \cdot_{\mathbf{H}})$  où :

- $\mathbf{H}$  est une partie de  $\mathbf{E}$  stable par  $+$ , par  $\times$  et par  $\cdot$  et contenant l'unité de l'anneau  $(\mathbf{E}, +, \times, \cdot)$ ;
- $+$ ,  $\times$  les lois induites sur  $\mathbf{H}$  respectivement par  $+$  et  $\times$ ,  $\cdot_{\mathbf{H}}$  l'opération de  $\mathbf{K}$  sur  $\mathbf{H}$  induite par  $\cdot$ ;
- $(\mathbf{H}, +_{\mathbf{H}}, \times_{\mathbf{H}}, \cdot_{\mathbf{H}})$  est une algèbre sur  $\mathbf{K}$ .

**Remarque** — Abusivement on dit que  $\mathbf{H}$  est une sous-algèbre.

On a alors le résultat immédiat suivant.

**Proposition 3.2.2.** — Soit  $(\mathbf{E}, +, \times, \cdot)$  un e.v. sur  $\mathbf{K}$ . Une partie  $\mathbf{H}$  de  $\mathbf{E}$  est une sous-algèbre de  $(\mathbf{E}, +, \times, \cdot)$  si et seulement si  $\mathbf{H}$  est un sous-anneau de  $(\mathbf{E}, +, \times)$  et un sous-e.v.  $(\mathbf{E}, +, \cdot)$ .

Les caractérisations des sous-anneaux et des sous-espaces vectoriels conduisent à celle d'une sous-algèbre :

Désignons par  $\mathbf{H}$  une partie d'une algèbre  $(\mathbf{E}, +, \times, \cdot)$ .

Sous-anneau

- $\alpha$ .  $1_{\mathbf{E}} \in \mathbf{H}$ ,
- $\beta$ . La partie  $\mathbf{H}$  est stable par addition,
- $\gamma$ . La partie  $\mathbf{H}$  est stable par passage à l'opposé,
- $\delta$ . La partie  $\mathbf{H}$  est stable par  $\times$  ;

Sous-espace vectoriel

- a.  $\mathbf{H}$  et non vide,
- b. La partie  $\mathbf{H}$  est stable par combinaison linéaire.,

la propriété b. assure  $\beta$ . et  $\gamma$ . la propriété  $\alpha$ . assure a.

Donc :

**Proposition 3.2.3** — CARACTÉRISATION DES SOUS-ALGÈBRES —

Soit  $(\mathbf{E}, +, \times, \cdot)$  un e.v. sur  $\mathbf{K}$ . Soit  $\mathbf{H}$  une partie de  $\mathbf{E}$ .  $\mathbf{H}$  est une sous-algèbre de  $(\mathbf{E}, +, \times, \cdot)$  si et seulement si

1.  $1_{\mathbf{E}} \in \mathbf{H}$ .
2. La partie  $\mathbf{H}$  est stable par combinaison linéaire.
3. La partie  $\mathbf{H}$  est stable par  $\times$ .

**Exemples 3.2.4.** —

- Dans la  $\mathbf{R}$ -algèbre  $\mathbf{C}$  (exemple 3.1.2. deuxième point),  $\mathbf{R}$  est une sous algèbre.

- L'ensemble  $\mathcal{T}_n(\mathbf{K})$  des matrices triangulaires supérieures d'ordre  $n$  à coefficients dans  $\mathbf{K}$  est, on l'a vu un sous anneau de l'anneau  $\mathcal{M}_n(\mathbf{K})$ , c'est aussi un sous-espace vectoriel de l'espace vectoriel  $\mathcal{M}_n(\mathbf{K})$ , (ce de façon immédiate) c'est donc une sous algèbre de l'algèbre  $\mathcal{M}_n(\mathbf{K})$ .
- On a déjà vu dans le cours de topologie que si  $D$  est une partie d'un espace vectoriel normé et  $\mathbf{K}$  est  $\mathbf{R}$  ou  $\mathbf{C}$ , l'ensemble  $\mathcal{C}^0(D, \mathbf{K})$  des applications continues de  $D$  dans  $\mathbf{K}$  est une sous algèbre de l'algèbre  $\mathcal{F}(D, \mathbf{K})$ .

### 3.3 Morphisme d'algèbres

**Définition 3.3.1.** — MORPHISME D'ALGÈBRES —

Soient  $(\mathbf{E}, +, \times, \cdot)$  et  $(\mathbf{F}, +, \times, \cdot)$  des algèbres sur  $\mathbf{K}$ . On appelle *morphisme de l'algèbre*  $(\mathbf{E}, +, \cdot)$  dans l'algèbre  $(\mathbf{F}, +, \cdot)$  toute application  $\mathbf{f}$  de  $\mathbf{E}$  dans  $\mathbf{F}$  qui soit un morphisme de l'anneau  $(\mathbf{E}, +, \times)$  dans l'anneau  $(\mathbf{F}, +, \times)$  et une application linéaire de l'e.v.  $(\mathbf{E}, +, \cdot)$  dans l'e.v.  $(\mathbf{F}, +, \cdot)$ ; c'est-à-dire qui vérifie :

- $\mathbf{f}(1_{\mathbf{E}}) = 1_{\mathbf{F}}$ ,
- Pour tout  $(\mathbf{x}, \mathbf{y}) \in \mathbf{E}^2$ , tout  $(\alpha, \beta) \in \mathbf{K}^2$ ,  $\mathbf{f}(\alpha \cdot \mathbf{x} + \beta \cdot \mathbf{y}) = \alpha \cdot \mathbf{f}(\mathbf{x}) + \beta \cdot \mathbf{f}(\mathbf{y})$ ;
- Pour tout  $(\mathbf{x}, \mathbf{y}) \in \mathbf{E}^2$ ,  $\mathbf{f}(\mathbf{x} \times \mathbf{y}) = \mathbf{f}(\mathbf{x}) \times \mathbf{f}(\mathbf{y})$ .

Un morphisme d'algèbres bijectif est appelé *isomorphisme d'algèbres*, un isomorphisme d'une algèbre sur elle-même est appelé *automorphisme d'algèbre*.

**Exemple 3.3.2.** —

- Dans la  $\mathbf{C}$ -algèbre  $\mathbf{C}$  (exemple 3.1.2. premier point), La conjugaison est un automorphisme d'algèbre.
- Soit  $\mathbf{E}$  un  $\mathbf{K}$  espace vectoriel de dimension  $n$  dont  $\mathcal{B}$  est une base. L'application

$$\mathcal{L}(\mathbf{E}) \rightarrow \mathcal{M}_n(\mathbf{K}); u \mapsto \text{Mat}_{\mathcal{B}}(u)$$

est un isomorphisme de l'algèbre  $\mathcal{L}(\mathbf{E})$  sur l'algèbre  $\mathcal{M}_n(\mathbf{K})$ .

— L'application

$$\mathcal{F}(D, \mathbf{K}) \rightarrow \mathbf{K}; f \mapsto f(d),$$

où  $d$  est un élément de  $D$  est un morphisme de l'algèbre  $\mathcal{F}(D, \mathbf{K})$  dans l'algèbre  $\mathbf{K}$ .

On montrera sans mal les propriétés suivantes que l'on peut déduire de celles analogues pour les morphismes d'anneaux et les applications linéaires.

**Proposition 3.3.3.** — PROPRIÉTÉS DES MORPHISMES D'ALGÈBRES —

Soit  $\mathbf{f}$  un morphisme d'une algèbre  $(\mathbf{E}, +, \times, \cdot)$  dans un algèbre  $(\mathbf{F}, +, \times, \cdot)$ .

1. Si  $\mathbf{H}$  est un sous-algèbre de  $(\mathbf{E}, +, \times, \cdot)$ , alors  $\mathbf{f}(\mathbf{H})$ , en est une de  $(\mathbf{F}, +, \times, \cdot)$ .
2. Si  $\mathbf{H}'$  est un sous-algèbre de  $(\mathbf{F}, +, \times, \cdot)$ , alors  $\mathbf{f}^{-1}(\mathbf{H}')$ , en est une de  $(\mathbf{E}, +, \times, \cdot)$ .

## Chapitre IX

# SÉRIES D'APPLICATIONS $\mathcal{C}^k$ , SÉRIES ENTIÈRES

Nous allons dans la première partie étudier la régularité de la limite d'une suite ou de la somme d'une série. Nous donnerons notamment un résultat qui précise sous quelles hypothèses la somme d'une série d'applications de classe  $\mathcal{C}^k$  est de classe  $\mathcal{C}^k$ . Ce théorème jouera un rôle important dans la seconde partie, où nous étudierons les séries entières, il s'agit de séries d'applications  $\sum u_n$ , où  $u_n(x)$  est de la forme  $a_n x^n$ , elles jouent un rôle essentiel en analyse, et occupent une place de choix aux concours.

## RÉGULARITÉ DES LIMITES DE SUITES ET DES SOMMES DE SÉRIES DE FONCTIONS

Nous allons étudier la régularité de la limite d'une suite d'applications. Dans le premier paragraphe nous rappellerons le résultat qui assure la continuité de la limite d'une suite d'applications continues et le résultat analogue pour la somme d'une série. Le deuxième paragraphe traitera de l'intégration de la limite d'une suite d'applications. Après avoir rappelé le résultat déjà vu dans un chapitre précédent, nous le généraliserons à la primitivation. Ce dernier résultat permettra dans le troisième paragraphe de démontrer un résultat de dérivation d'une limite d'une suite d'applications. Dans ce même paragraphe en généralisant ce résultat, nous étudierons le caractère  $\mathcal{C}^k$  de la limite d'une suite ou de la somme d'une série. Toujours dans le même paragraphe suivra une application aux fonctions exponentielles.

Dans toute cette partie  $I$  désigne un intervalle non vide quelconque. Par  $\mathbf{K}$  nous désignons indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$  et par  $\mathbf{F}$  un  $\mathbf{K}$ -espace vectoriel de dimension finie. Les applications seront, sauf mention contraire, définies sur  $I$  et à valeurs dans  $\mathbf{K}$  ou, à la rigueur, à valeurs dans  $\mathbf{F}$ . Dans les preuves nous traiterons le cas d'applications à valeurs dans  $\mathbf{K}$ , la généralisation  $\mathbf{F}$  étant instantanée.

### 1.1 Suites et séries de fonctions et continuité

Rappelons un théorème sur la continuité de la limite d'une suite d'applications et son corollaire sur les séries d'applications. Le lecteur est prié de se reporter au chapitre où ces résultats furent introduits, pour plus de détails et de commentaires.

**Théorème 1.1.1** — *Soit  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou dans  $\mathbf{F}$ , continues, qui converge uniformément vers une application  $f$ . Alors l'application  $f$  est continue (sur  $I$ ).*

**Corollaire 1.1.2** — *Soit  $(u_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  (ou  $\mathbf{F}$ ), continues. On suppose que la série  $\sum u_n$  converge uniformément. Alors sa somme est continue.*

Nous avons déjà observé sur des exemples que la convergence uniforme sur les segments de  $I$  suffit à prouver la continuité de la limite. Nous allons énoncer ce résultat qui, quoique non explicitement au programme est implicitement utilisé dans le résultat principal du paragraphe suivant.

**Lemme 1.1.3** — *Soit  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou dans  $\mathbf{F}$ , continues, qui converge uniformément sur tout segment de  $I$ . Alors :*

- la suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement sur  $I$  vers une application  $f$ ,
- $f$  est continue sur  $I$ .

On dira que  $(f_n)_{n \in \mathbf{N}}$  converge vers  $f$ , uniformément sur tout segment de  $I$ .

*Preuve du lemme 1.1.3* — Commençons par le premier point. Soit  $t$  un élément de  $I$ . Il existe un segment  $J$  de  $I$  qui contient  $t$ ,  $[t, t]$  par exemple! La suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément sur  $J$ , donc simplement, en particulier la suite  $(f_n(t))_{n \in \mathbf{N}}$  converge. Notons  $f(t)$  sa limite. Le point  $t$  étant quelconque, on dispose donc d'une application  $f$  de  $I$  dans  $\mathbf{K}$  qui est la limite simple sur  $I$  de la suite  $(f_n)_{n \in \mathbf{N}}$ .

Passons au second point. Soit  $x_0$  un élément de  $I$ . Il résulte de ce que tout point de  $I$  admet un voisinage relativement à  $I$  qui est un segment. Précisément, nous noterons  $a$  la borne inférieure de  $I$  et  $b$  sa borne supérieure.

Traisons pour commencer le cas où  $x_0$  n'est pas une borne de l'intervalle  $I$ . Posons :

$$c := \begin{cases} \frac{a + x_0}{2}, & \text{si } a > -\infty, \\ x_0 - 1, & \text{si } a = -\infty, \end{cases} \quad d := \begin{cases} \frac{b + x_0}{2}, & \text{si } b < +\infty, \\ x_0 + 1, & \text{si } b = +\infty. \end{cases}$$

ainsi a-t-on que  $[c, d]$  est un segment de  $I$  et que  $x_0 \in ]c, d[$ . La convergence uniforme de la suite  $(f_n)_{n \in \mathbf{N}}$  sur le segment  $[c, d]$  assure la continuité de  $f|_{[c, d]}$ , donc *a fortiori* celle de  $f|_{]c, d[}$ . L'intervalle  $]c, d[$  étant un ouvert relatif de  $I$ ,  $f$  est continue en tout point de  $]c, d[$ , donc en particulier en  $x_0$ .

Traisons à présent le cas où  $x_0$  est une borne de l'intervalle. Nous prendrons par exemple  $x_0 = a$ , l'autre cas se traite de même. On a donc que  $a$  est réel. Posons :

$$d := \begin{cases} \frac{b + a}{2}, & \text{si } b < +\infty, \\ a + 1, & \text{si } b = +\infty. \end{cases}$$

ainsi a-t-on que  $[a, d]$  est un segment de  $I$ . La convergence uniforme de la suite  $(f_n)_{n \in \mathbf{N}}$  sur le segment  $[c, d]$  assure la continuité de  $f|_{[a, d]}$ , donc *a fortiori* celle de  $f|_{[a, d[}$ . L'intervalle  $[a, d[$  étant un ouvert relatif de  $I$ ,  $f$  est continue en tout point de  $[a, d[$ , donc en particulier en  $a$ .

On a donc montré la continuité de  $f$  en un point quelconque  $x_0$  de  $I$ , donc l'application  $f$  est continue.  $\square$

## 1.2 Suites et séries de fonctions et intégration

Voici pour commencer le rappel d'un résultat déjà rencontré

**Proposition 1.2.1** — Soit une suite  $(f_n)_{n \in \mathbf{N}}$  d'applications d'un segment  $[a, b]$  dans  $\mathbf{K}$  ou dans  $\mathbf{F}$  continues, qui converge uniformément vers une application  $f$ . Alors,

- L'application  $f$  est continue;
- La suite  $\left(\int_a^b f_n(t) dt\right)_{n \in \mathbf{N}}$  converge vers  $\int_a^b f(t) dt$ .

Dit de façon brutale : « L'intégrale de la limite est la limite des intégrales ».

**Remarque** —

- La continuité de la limite  $f$ , qui découle de la proposition 1.1.1, est citée ici pour justifier l'existence de  $\int_a^b f(t) dt$ .
- Dans le chapitre où nous avons énoncé ce résultat nous avons insisté sur le fait que ce théorème est faux pour une suite d'applications qui ne convergerait pas uniformément, ou si l'on remplace le segment  $[a, b]$  par un intervalle quelconque.

On avait également énoncé la traduction de ce résultat pour les séries.

**Proposition 1.2.2** — Soit une suite  $(u_n)_{n \in \mathbf{N}}$  d'applications d'un segment  $[a, b]$  dans  $\mathbf{K}$  ou dans  $\mathbf{F}$  continues. On suppose que la série d'applications  $\sum u_n$  converge uniformément de somme  $S$ . alors,

- L'application  $S$  est continue;
- La série  $\sum \int_a^b u_n(t) dt$  converge de somme  $\int_a^b S(t) dt$ .

Ou, plus brutalement, « La somme de la série des intégrales est l'intégrale de la somme de la série ».

Donnons un corollaire de la proposition 1.2.1, qui est en quelque sorte une généralisation de l'inégalité triangulaire pour la norme  $N_1$ .

**Exercice 1.2.3** — Soit une suite  $(u_n)_{n \in \mathbf{N}}$  d'applications d'un segment  $[a, b]$  dans  $\mathbf{K}$  continues. On suppose que la série d'applications  $\sum u_n$  converge normalement de somme  $S$ . Alors la série  $\sum N_1(u_n)$  converge et :

$$N_1\left(\sum_{n=0}^{+\infty} u_n\right) \leq \sum_{n=0}^{+\infty} N_1(u_n).$$

*Solution de l'exercice 1.2.3 —*

Donnons maintenant un résultat plus fort que celui de la proposition 1.2.1.

**Proposition 1.2.4** — *Soit  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$ , continues qui converge uniformément sur tout segment de  $I$ . D'après le lemme 1.1.3,  $(f_n)_{n \in \mathbf{N}}$  converge simplement sur  $I$  vers une application  $f$  de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$  qui est continue. Soit de plus  $a$  un élément de  $I$ , posons :*

$$F_n : I \rightarrow \mathbf{K}(\mathbf{F}); x \mapsto \int_a^x f_n(t) dt,$$

*pour tout entier naturel  $n$ , et*

$$F : I \rightarrow \mathbf{K}(\mathbf{F}); x \mapsto \int_a^x f(t) dt.$$

*Alors la suite  $(F_n)_{n \in \mathbf{N}}$  converge vers  $F$ , uniformément sur tout segment de  $I$ .*

**Remarque** — Même si l'on suppose que la suite  $(f_n)_{n \in \mathbf{N}}$  converge uniformément sur  $I$  et pas seulement sur les segments de  $I$ , on ne peut pas affirmer que  $(f_n)_{n \in \mathbf{N}}$  converge uniformément sur  $I$ . Donnons un exemple. Prenons  $I = \mathbf{R}$ ,  $a = 0$  et, pour tout entier naturel  $n$ ,

$$f_n : \mathbf{R} \rightarrow \mathbf{R}; x \mapsto \dots$$

*Preuve de la proposition 1.2.4 —*



On peut énoncer un résultat analogue à 1.2.4 pour les séries de fonctions qui s'en déduit en raisonnant sur la suite des sommes partielles de la série.

**Proposition 1.2.5** — *Soit  $(u_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$ , continues. on suppose que la série  $\sum u_n$  converge uniformément sur tout segment de  $I$ . D'après le lemme 1.1.3,  $\sum u_n$  converge simplement sur  $I$  de somme une application  $S$  de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$  qui est continue. Soit de plus  $a$  un élément de  $I$ , posons :*

$$U_n : I \rightarrow \mathbf{K}(\mathbf{F}); t \mapsto \int_a^t u_n(t) dt,$$

*pour tout entier naturel  $n$ , et*

$$F : I \rightarrow \mathbf{K}(\mathbf{F}); t \mapsto \int_a^t S(t) dt.$$

*Alors la série  $\sum U_n$  converge, uniformément sur tout segment de  $I$ , et a pour somme  $F$ .*

*« La série des primitives converge, uniformément sur tout segment de  $I$ , et a pour somme la primitive de la somme. »*

### 1.3. Suites et séries de fonctions et dérivation

Commençons par une mise en garde. la convergence uniforme d'une suite de fonctions  $(f_n)_{n \in \mathbf{N}}$  vers une application  $f$  permet, comme nous l'avons vu, de contrôler l'écart entre  $f_n$  et  $f$ , mais en aucun cas les variations de  $f_n - f$ . C'est pourquoi, lorsque pour tout  $n$ ,  $f_n$  est dérivable, on ne sait pas, *a priori*, si  $f'_n$  converge vers  $f'$  ni même si  $f'$  existe. Donnons des exemples.

**Exemples—**

1. Pour tout élément  $n$  de  $\mathbf{N}^*$ , on définit l'application

$$f_n : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \frac{1}{\sqrt{n}} \sin(nt).$$

2. Soit l'application

$$f : [-1, 1] \rightarrow \mathbf{R}; t \mapsto |t|.$$

L'application  $f$  n'est pas dérivable en 0, elle est pourtant limite uniforme d'une suite d'applications dérivables et même  $\mathcal{C}^\infty$ . En effet ...

Pour contrôler le caractère dérivable de la limite, il est essentiel de contrôler la convergence de la suite des dérivées. C'est l'essence du théorème suivant.

**Proposition 1.3.1 —** THÉORÈME DE DÉRIVATION D'UNE LIMITE—

Soit  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$ . On suppose :

1. Pour tout entier naturel  $n$ , l'application  $f_n$  est de classe  $\mathcal{C}^1$ .
2. La suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers une application  $f$ .

3. La suite  $(f'_n)_{n \in \mathbf{N}}$  converge vers une application  $g$ , uniformément sur tout segment de  $I$ .

Alors :

- i) La suite  $(f_n)_{n \in \mathbf{N}}$  converge vers  $f$ , uniformément sur tout segment de  $I$ .
- ii) L'application  $f$  est de classe  $\mathcal{C}^1$ .
- iii)  $f' = g$ .

*Preuve de la proposition 1.3.1*— Soit  $a$  un point de  $I$ . Pour tout entier naturel  $n$  et tout élément  $t$  de  $I$ ,

$$f_n(t) = f_n(a) + \int_a^t f'_n(s) \mathrm{d}s.$$

Notons, pour tout entier naturel  $n$ ,

$$F_n : I \rightarrow \mathbf{K}; t \mapsto \int_a^t f'_n(s) \mathrm{d}s$$

et

$$F : I \rightarrow \mathbf{K}; t \mapsto \int_a^t g(s) \mathrm{d}s.$$

On va montrer que  $(f_n)_{n \in \mathbf{N}}$  converge, uniformément sur tout segment, vers

$$\tilde{f} : I \rightarrow \mathbf{K}; t \mapsto \dots$$

Ce résultat se généralise :

**Proposition 1.3.2** — Soient  $k$  un élément de  $\mathbf{N}^*$  et  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$ . On suppose :

1. Pour tout entier naturel  $n$ , l'application  $f_n$  est de classe  $\mathcal{C}^k$ .
2. Pour  $i = 0, 1, \dots, k-1$ , La suite  $(f_n^{(i)})_{n \in \mathbf{N}}$  converge simplement vers une application notée  $g_i$ .
3. La suite  $(f_n^{(k)})_{n \in \mathbf{N}}$  converge vers une application notée  $g_k$ , uniformément sur tout segment de  $I$ .

Alors :

- i) Pour  $i = 0, 1, \dots, k$ , La suite  $(f_n^{(i)})_{n \in \mathbf{N}}$  converge vers  $g_i$ , uniformément sur tout segment de  $I$ .
- ii) L'application  $g_0$ , limite de la suite  $(f_n)_{n \in \mathbf{N}}$ , est de classe  $\mathcal{C}^k$ .
- iii) Pour  $i = 0, 1, \dots, k$ ,  $g_0^{(i)} = g_i$ .

Dit brutalement : « La dérivée  $i^{\text{e}}$  de la limite est la limite de la suite des dérivées  $i^{\text{e}}$  ».

*Preuve de la proposition 1.3.2* — La proposition 1.3.2 se déduit par récurrence de la proposition 1.3.1. Nous laissons en exercice la rédaction de cette récurrence.

Nous allons donner une traduction de ce dernier résultat en termes de séries d'applications qui s'obtient en appliquant 1.3.2 à la suite des sommes partielles de la série.

**Proposition 1.3.3** — Soient  $k$  un élément de  $\mathbf{N}^*$  et  $(u_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$  ou  $\mathbf{F}$ . On suppose :

1. Pour tout entier naturel  $n$ , l'application  $u_n$  est de classe  $\mathcal{C}^k$ .
2. Pour  $i = 0, 1, \dots, k-1$ , ...
3. La ...

Alors :

- i) Pour  $i = 0, 1, \dots, k$ , La
  - ii) L'application  $S_0$ , somme de la série  $\sum u_n$ , est de classe  $\mathcal{C}^k$ .
  - iii) Pour  $i = 0, 1, \dots, k$ ,
- Dit brutalement : « La dérivée  $i^{\text{e}}$  de ... »

Donnons un exercice d'application.

**Exercice 1.3.4.** —

Soit la série de fonctions  $\sum_{n \geq 1} u_n$ , où pour tout entier naturel  $n$  non nul,

$$u_n : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto \frac{(-1)^{n+1}}{n} e^{-nx}.$$

1. Montrer que la série de fonctions  $\sum_{n \geq 1} u_n$  converge simplement sur  $\mathbf{R}_+$ . Nous noterons  $f$  l'application de  $\mathbf{R}_+$  dans  $\mathbf{R}$  somme de cette série.
2. Montrer que  $f$  est continue.
3. Montrer que  $f$  est dérivable en tout point de  $\mathbf{R}_+^*$  et donner l'expression de  $f'(x)$ , pour tout élément  $x$  de  $\mathbf{R}_+^*$ .
4. En déduire une expression de  $f$ , sans signe somme, au moyen des applications usuelles.

*Solution de l'exercice 1.3.4. —*

série  $\sum \frac{z^n}{n!}$  converge absolument. La somme s'appelle exponentielle de  $z$ , notée  $\exp(z)$ , on définit ainsi l'application exponentielle sur  $\mathbf{C}$  :

$$\exp : \mathbf{C} \rightarrow \mathbf{C}; z \mapsto \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

On a montré également que pour tout couple  $(z_1, z_2)$  de complexes  $\exp(z_1 + z_2) = \exp(z_1) \exp(z_2)$ , ainsi la restriction de  $\exp$  à  $\mathbf{R}$  nous apparait comme l'application exponentielle déjà connue.

On peut alors généraliser la définition à  $\mathcal{M}_n(\mathbf{K})$ , (l'exponentielle complexe correspondant au cas  $\mathcal{M}_1(\mathbf{C})$ !). Nous avons vu que

pour toute matrice  $M \in \mathcal{M}_n(\mathbf{K})$ , la série  $\sum \frac{M^n}{n!}$  converge absolument. La somme s'appelle exponentielle de  $M$ , notée  $\exp(M)$ , on définit ainsi l'application exponentielle sur  $\mathcal{M}_n(\mathbf{K})$  :

$$\exp : \mathcal{M}_n(\mathbf{K}) \rightarrow \mathcal{M}_n(\mathbf{K}); z \mapsto \sum_{n=0}^{+\infty} \frac{M^n}{n!}.$$

La preuve de la convergence absolue de cette série se faisait en équipant  $\mathcal{M}_n(\mathbf{K})$  d'une norme  $\|\cdot\|$  telle qu'il existe  $C \in \mathbf{K}$  tel que pour tout  $(A, B) \in \mathcal{M}_n(\mathbf{K})$ ,

$$\|AB\| \leq C\|A\|\|B\|. \quad (\text{IX.1})$$

Dans le cas (idéal) où  $C = 1$ , on parle de norme d'algèbre.

Par exemple la norme  $\|\cdot\|_\infty$  définie dans un précédent chapitre par :

$$\|M\|_\infty = \sup_{\substack{i=1,\dots,n \\ j=1,\dots,n}} |m_{i,j}|,$$

pour tout  $M \in \mathcal{M}_n(\mathbf{K})$ , vérifie (IX.1), avec  $C = n$ .

Remarquons que  $\frac{1}{\sqrt{n}}\|\cdot\|_\infty$  est une norme vérifiant (IX.1) pour  $C = 1$ , c'est aussi le cas de la norme

$$\mathcal{M}_n(\mathbf{K}) \rightarrow \mathcal{M}_n(\mathbf{K}); M \mapsto \max \left\{ \sum_{j=1}^n |m_{i,j}|, i = 1, \dots, n \right\}$$

Passons rapidement au cas de l'exponentielle d'un endomorphisme d'un  $\mathbf{K}$ -espace vectoriel  $\mathbf{E}$  de dimension finie  $n$  (non nulle). Une fois  $\mathbf{E}$  muni d'une base  $\mathcal{B}$ , nous disposons de l'isomorphisme d'algèbres (cf. chapitre 7),

$$\mathcal{L}(\mathbf{E}) \rightarrow \mathcal{M}_n(\mathbf{K}); f \mapsto \text{Mat}_{\mathcal{B}}(f).$$

Si  $\|\cdot\|$  est une norme sur  $\mathcal{M}_n(\mathbf{K})$ , on définit alors une norme sur  $\mathcal{L}(\mathbf{E})$  par :

$$\mathcal{L}(\mathbf{E}) \rightarrow \mathbf{R}_+; f \mapsto \|\text{Mat}_{\mathcal{B}}(f)\|.$$

Si  $\|\cdot\|$  est une norme d'algèbre alors par construction la norme que l'on vient de lui associer sur  $\mathcal{L}(\mathbf{E})$  l'est aussi.

Nous disposons donc de normes d'algèbre sur  $\mathcal{L}(\mathbf{E})$ . Nous n'avons besoin de rien de plus pour démontrer, pour tout endomorphisme  $f$  de  $\mathbf{E}$ , la convergence absolue de la série  $\sum \frac{f^n}{n!}$ . La somme s'appelle exponentielle de  $f$ , notée  $\exp(f)$ , on définit ainsi l'application exponentielle sur  $\mathcal{L}(\mathbf{E})$  :

$$\exp : \mathcal{L}(\mathbf{E}) \rightarrow \mathcal{L}(\mathbf{E}); z \mapsto \sum_{n=0}^{+\infty} \frac{f^n}{n!}.$$

On remarquera qu'avec les notations précédentes

$$\text{Mat}_{\mathcal{B}}(\exp(f)) = \exp(\text{Mat}_{\mathcal{B}}(f))$$

en effet

Ceci permet de travailler dans la pratique sur des matrices si l'on souhaite.

Dans la suite  $(A, +, \times, \cdot)$  Désigne indifféremment l'une des 3 algèbres  $\mathbf{K}$ ,  $\mathcal{M}_n(\mathbf{K})$  ou  $\mathcal{L}(\mathbf{E})$ , avec  $\mathbf{E}$ ,  $\mathbf{K}$ -espace vectoriel de dimension finie  $n$ . L'espace vectoriel  $(A, +, \cdot)$  est muni d'une norme d'algèbre quelconque (toutes les normes sont équivalentes).

Soit maintenant  $a$  un élément de  $A$ . On peut en particulier considérer l'application suivante :

$$\Phi_a : \mathbf{R} \rightarrow A ; t \mapsto \exp(t \cdot a)$$

Dans le cas où l'on travaille dans l'algèbre  $\mathbf{R}$ , on sait depuis belle lurette que  $\Phi_a$  est dérivable et que  $\Phi'_a = a\Phi_a$ . Nous allons généraliser ce résultat.

**Proposition 1.3.5.** — *l'application  $\Phi_a$  est de classe  $\mathcal{C}^1$  et :*

$$\Phi'_a = a \times \Phi_a.$$

D'où le corollaire immédiat :

**Corollaire 1.3.6.** — *L'application  $\Phi_a$  est de classe  $\mathcal{C}^\infty$  et pour tout entier naturel  $k$ ,*

$$\Phi_a^{(k)} = \dots$$

**Remarque** — On peut montrer que  $a$  et  $\Phi_a(t)$  commutent pour tout réel  $t$ . En effet  $a$  commute avec toute somme partielle de la série dont la somme est  $\Phi_a(t)$ , la continuité du produit par un élément de  $A$  assure par passage à la limite le résultat.

#### ÉTUDE DES SYSTÈMES DIFFÉRENTIELS LINÉAIRES

L'étude précédente de  $\Phi_a$  conduit à la résolution théorique des systèmes différentiels linéaires.

Soient  $n$  un élément de  $\mathbf{N}^*$  et  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . On se propose d'étudier le système différentiel linéaire

$$X' = MX. \quad (\text{IX.2})$$

On appelle solution (sous-entendu sur  $\mathbf{R}$ ) de (IX.2), toute application  $\Phi$  de  $\mathbf{R}$  dans  $\mathcal{M}_{n,1}(\mathbf{K})$  de classe  $\mathcal{C}^{11}$ , telle que  $\Phi' = M\Phi$ .

Le cours de terminale nous révèle que pour  $m$  réel, les solutions de l'équation

$$x' = mx$$

sont les applications de la forme  $R \rightarrow \mathbf{R}; t \mapsto c \exp(mt)$ , où  $c$  est un réel quelconque. Un résultat analogue s'énonce dans le cas du système (IX.2).

**Proposition 1.3.7.** — Une application  $\Phi$  de  $\mathbf{R}$  dans  $\mathcal{M}_{n,1}(\mathbf{K})$  de classe  $\mathcal{C}^1$  est solution de (IX.2), si et seulement si il existe un élément  $C$  de  $\mathcal{M}_{n,1}(\mathbf{K})$  tel que :

$$\Phi : \mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{K}); t \mapsto \cdot \quad (\text{IX.3})$$

La preuve de ce résultat se calque sur celle du cas réel que l'on vient d'évoquer. Elle est cependant quelque peu plus technique.

*Preuve de la proposition 1.3.7.* —

- HYPOTHÈSE : Il existe  $C$  élément de  $\mathcal{M}_{n,1}(\mathbf{K})$  tel que  $\Phi$  soit de la forme (IX.3).

- HYPOTHÈSE L'application  $\Phi$  est solution de (IX.2).

1. Dérivable, ici reviendrait au même puisque toute application dérivable vérifiant  $\Phi' = M\Phi$  est de classe  $\mathcal{C}^1$ .



**Remarque** — L'ensemble des solutions de (IX.2) est un espace vectoriel de dimension  $\dots$ .

Nous reparlerons de ce résultat dans le cours sur les équations différentielles linéaires.

De 1.3.7. se déduit le corollaire suivant :

**Corollaire 1.3.8.** — *Soit  $X_0$  un élément de  $\mathcal{M}_{n,1}(\mathbf{K})$ . Il existe une et une seule solution de (IX.2) qui prenne la valeur  $X_0$  en 0, c'est*

$$\mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{K}); t \mapsto$$

Ce corollaire admet un corollaire immédiat :

**Corollaire 1.3.9.** — *Soit  $X_0$  un élément de  $\mathcal{M}_{n,1}(\mathbf{K})$  et  $t_0$  un réel. Il existe une et une seule solution de (IX.2) qui prenne la valeur  $X_0$  en  $t_0$ , c'est*

$$\mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{K}); t \mapsto$$

*On dit que cette solution vérifie la condition initiale  $X(t_0) = X_0$ .*

# SÉRIES ENTIÈRES

Nous allons nous intéresser à la possibilité d'écrire une application  $f$  sous la forme

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots,$$

la somme du second membre étant infinie. Au XVIII<sup>e</sup> siècle, beaucoup dont Lagrange pensaient pouvoir mettre toute fonctions régulière sous cette forme. On sait depuis que la chose n'est pas possible et qu'il est des applications, même indéfiniment dérivables, qui ne se laissent pas écrire de la sorte.

Dans un premier temps nous étudierons les séries d'applications de la forme  $\sum a_n x^n$  appelées séries entières. Le premier paragraphe se consacrera plus particulièrement à leur convergence. Le deuxième étudiera la somme et le produit de telles séries. Dans le troisième sera abordé dérivation et la primitivation de leur somme. C'est dans le quatrième paragraphe que nous aborderons enfin la possibilité d'écrire une application comme la somme d'une série entière.

## 2.1 Définition d'une série entière, rayon de convergence

**Définition 2.1.1.** — On appelle série entière de la variable complexe  $z$  (resp. réelle  $t$ ), toute série d'application  $\sum u_n$ , où pour tout entier naturel  $n$ , il existe un complexe  $a_n$  tel que

$$u_n : \mathbf{C} \rightarrow \mathbf{C}; z \mapsto a_n z^n, \\ (\text{resp. } u_n : \mathbf{R} \rightarrow \mathbf{C}; t \mapsto a_n t^n.)$$

On la note traditionnellement et abusivement  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ).

La notation abusive d'une série entière engendre quelques ambiguïtés. Si  $z$  désigne un nombre complexe donné,  $\sum a_n z^n$  peut désigner deux choses :

- La série entière  $\sum a_n z^n$  ;
- La série complexe  $\sum a_n z^n$ .

Pour éviter ce genre de risque on précisera toujours en écrivant « la série entière  $\sum a_n z^n$ , » ou bien « la série complexe  $\sum a_n z^n$ , ». Pour notre part, nous nous efforcerons de surcroît, de noter un complexe donné  $z_0, z_1, Z$  etc., plutôt que  $z$ .

Nous allons montrer que le domaine de convergence d'une série entière a une forme bien particulière. Commençons par un lemme technique, moteur de cette étude.

**Lemme 2.1.2** — LEMME D'ABEL —

Soit  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ). Soit  $z_0$  un complexe (resp.  $t_0$  un réel). On suppose que la suite  $(a_n z_0^n)_{n \in \mathbf{N}}$  (resp.  $(a_n t_0^n)_{n \in \mathbf{N}}$ ) est bornée. Alors pour tout complexe  $z_1$  tel que  $|z_1| < |z_0|$  (resp. tout réel  $t_1$  tel que  $|t_1| < |t_0|$ ) la série complexe  $\sum a_n z_1^n$  (resp. la série réelle  $\sum a_n t_1^n$ ) converge absolument.

*Preuve du lemme 2.1.2* — Donnons la preuve pour une série complexe, le cas réel se traite de même. Soit  $z_1$  un complexe tel que  $|z_1| < |z_0|$ . Pour tout entier naturel  $n$ ,

Donc la série  $\sum a_n z_1^n$  converge absolument. □

**Proposition-définition 2.1.3** — RAYON DE CONVERGENCE —

Soit  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ). Il existe un et un seul élément  $R$  de  $\mathbf{R}_+ \cup \{+\infty\}$  tel que l'on ait, pour tout complexe  $z_0$  (resp. réel  $t_0$ ) :

- i. Si  $|z_0| < R$  (resp.  $|t_0| < R$ ), alors la série complexe  $\sum a_n z_0^n$  (resp. la série réelle  $\sum a_n t_0^n$ ) converge absolument;
- ii. Si  $|z_0| > R$  (resp.  $|t_0| > R$ ), alors la série complexe  $\sum a_n z_0^n$  (resp. la série réelle  $\sum a_n t_0^n$ ) diverge grossièrement.

L'élément  $R$  est appelé Rayon de convergence de la série entière. L'ensemble  $\{z \in \mathbf{C}, |z| < R\}$  (resp.  $\{z \in \mathbf{C}, |t| < R\}$ ) s'appelle disque ouvert de convergence de la série entière  $\sum a_n z^n$  (resp. intervalle ouvert de convergence de la série entière  $\sum a_n z^n$ ).

*Preuve de la proposition 2.1.3* — Qu'il n'existe au plus qu'un élément  $R$  satisfaisant aux conditions [i.] et [ii.] est une évidence. Nous ne montrerons donc que l'existence d'un tel élément. Soit  $A$  l'ensemble des réels  $\rho$  positifs ou nuls tels que  $(a_n \rho^n)_{n \in \mathbf{N}}$  soit bornée. L'ensemble  $A$  est non vide puisque'il contient 0.

- PREMIER CAS :  $A$  est majoré on pose alors  $R := \dots$
- SECOND CAS :  $A$  est non majoré on pose alors  $R := \dots$

**Commentaires** —

Dans le cas où  $\mathbf{R} = 0$ , pour tout  $z$  élément de  $\mathbf{C}^*$ , la série complexe  $\sum a_n z^n$  diverge, bien sûr,  $\sum a_n 0^n$  converge. Dans le cas où  $R = +\infty$ , pour tout élément  $z$  de  $\mathbf{C}$ , la série complexe  $\sum a_n z_0^n$  converge.

Donnons des exemples

Etudions le rayon de convergence de la série entière  $\sum a_n z^n$ , pour les différents exemples suivants.

- EXEMPLE 1  
Pour tout entier naturel  $n$ ,  $a_n = 0$ . On a alors  $R = \dots$
- EXEMPLE 2  
Pour tout entier naturel  $n$ ,  $a_n = \dots$

Donc  $R = 0$ .

- EXEMPLE 3  
Pour tout entier naturel  $n$ ,  $a_n = \frac{1}{n!}$ .

Donc  $R = \dots$

- EXEMPLE 4 :

Pour tout entier naturel  $n$ ,  $a_n = \frac{1}{n^\alpha}$ , où  $\alpha$  est un réel donné.

- EXEMPLE 5

Pour tout entier naturel  $p$ ,  $a_{2p} = 0$  et  $a_{2p+1} = pe^{2p}$ .

Donc  $R = \dots$

- EXEMPLE 6 :

Pour tout entier naturel  $n$ ,  $a_n$  est la  $n^{\text{e}}$  décimale de  $\pi$ .

Donc  $R = \dots$

ÉTUDE AU BORD — Soit  $\sum a_n z^n$  une série entière de la variable complexe  $z$  (le cas complexe se traite de façon similaire). On suppose que son rayon de convergence  $R$  n'est ni nul ni infini. Le disque ouvert de convergence est alors au sens géométrique un disque. Que dire de la série complexe  $\sum a_n z_0^n$ , pour  $z_0$  un nombre complexe de module  $R$ ?

Nous allons voir sur des exemples que l'on ne peut rien dire de général.

Étudions le rayon de convergence de la série entière  $\sum a_n z^n$ , pour les différents exemples suivants.

- Pour tout entier naturel  $n$ ,  $a_n = \frac{1}{n^2}$ , on a vu, dans l'exemple 4, que  $R = 1$ . Soit  $z_0$  un complexe de module 1

Donc la série complexe  $\sum \frac{1}{n^2} z_0^n \dots$

- Pour tout entier naturel  $n$ ,  $a_n = \frac{1}{n}$ , on a vu, dans l'exemple 4, que  $R = 1$ .

**Proposition 2.1.4.**— Soient  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ) de rayon de convergence  $R_a$  et Soit  $\sum b_n z^n$  (resp.  $\sum b_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ) de rayon de convergence  $R_b$ . Alors :

1. Si  $a_n = O(b_n)$  ( $n \rightarrow +\infty$ ), alors  $R_a \geq R_b$ .
2. Si  $a_n \sim b_n$  ( $n \rightarrow +\infty$ ), alors  $R_a = R_b$ .

*Preuve de la proposition 2.1.4.* —

Etudions à présent le mode de convergence d'une série entière.

**Proposition 2.1.5.**— Soit  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ) de rayon de convergence  $R$  non nul. La série entière converge normalement sur tout disque fermé de centre<sup>1</sup> 0, inclus dans le disque ouvert (resp. l'intervalle ouvert) de convergence.

En général il n'y a pas de convergence normale, ni même

1. On a alors immédiatement le résultat pour tout disque fermé de centre quelconque (et même pour tout compact), inclus dans le disque ouvert de convergence.

uniforme sur le disque ouvert de convergence, ce même si  $R = +\infty$ .

*Preuve de la proposition 2.1.5.—*

Donnons des exemples où il n'y a pas de convergence uniforme sur le disque ouvert de convergence.

- EXEMPLE 1

Soit la série entière  $\sum z^n$ .

- EXEMPLE 2

Soit la série entière  $\sum \frac{z^n}{n!}$ .

La convergence n'est pas normale sur  $\mathbf{C}$ . En effet, l'application  $\mathbf{C} \rightarrow \mathbf{C}; z \mapsto \frac{z^n}{n!}$  n'est bornée pour aucun entier naturel  $n$  non nul.

Montrons que la convergence n'est pas d'avantage uniforme. Soit  $n$  un entier naturel, notons  $R_n$  l'application le reste d'ordre  $n$  de la série. Pour tout entier  $p \geq n+1$ , on a :

$$R_n(p) = \sum_{k=n+1}^{+\infty} \frac{p^k}{k!} \geq \frac{p^p}{p!}.$$

Or  $\frac{p^p}{p!} \xrightarrow{p \rightarrow +\infty} +\infty$ , donc  $R_n$  n'est pas borné. Or si la série convergerait uniformément sur  $\mathbf{C}$ , pour  $n$  suffisamment grand  $R_n$  serait borné sur  $\mathbf{C}$  et on aurait  $\|R_m\|_\infty \xrightarrow{m \rightarrow +\infty} 0$ .

Donc la série entière ne converge pas uniformément sur  $\mathbf{C}$ .

**Proposition 2.1.6.**— Soit  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ) de rayon de convergence  $R$  non nul. On note  $D$  le disque ouvert (resp. intervalle ouvert) de convergence. Soit  $S$  l'application somme de la série entière :

$$S : D \rightarrow \mathbf{C}; z \mapsto \sum_{n=0}^{+\infty} a_n z^n,$$

$$(resp. S : D \rightarrow \mathbf{R}; t \mapsto \sum_{n=0}^{+\infty} a_n t^n.)$$

L'application  $S$  est continue sur le disque ouvert (resp. intervalle ouvert) de convergence  $D$ .

**Preuve de la proposition 2.1.6.**—

## 2.2 Sommes et produits de séries entières

Soient  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) et  $\sum b_n z^n$  (resp.  $\sum b_n t^n$ ) des séries entières de la variable complexe  $z$  (resp. réelle  $t$ ) de rayons de convergence respectifs  $R$  et  $R'$ .

**Définition 2.2.1.**—

1. On appelle série entière somme de  $\sum a_n z^n(t^n)$  et de  $\sum b_n z^n(t^n)$  la série entière  $\sum s_n z^n(t^n)$ , où pour tout entier naturel  $n$ ,

$$s_n = a_n + b_n.$$

2. On appelle série entière produit (de Cauchy) de  $\sum a_n z^n(t^n)$  et de  $\sum b_n z^n(t^n)$  la série entière  $\sum p_n z^n(t^n)$ , où pour tout entier naturel  $n$ ,

$$p_n = \sum_{\substack{i=1\dots n \\ j=1\dots n \\ i+j=n}} a_i b_j.$$

**Remarques 2.2.2** —

- On observe facilement que la somme et le produit sont des opérations, sur l'ensemble des séries entières de la variable complexe (resp. réelle) *commutatives*.
- Pour tout entier naturel  $n$ ,  $p_n$  s'écrit également :

$$p_n = \sum_{i=1}^n a_i b_{n-i} = \sum_{j=1}^n a_{n-j} b_j.$$

- On observe une forte analogie entre le produit (de Cauchy) de deux séries entières et le produit de deux polynômes  $P$  et  $Q$  à coefficients dans un corps  $\mathbf{K}$ . En effet si  $P = \sum_{n \in \mathbf{N}} a_n X^n$  et  $Q = \sum_{n \in \mathbf{N}} b_n X^n$ , où  $(a_n)_{n \in \mathbf{N}}$



et  $(b_n)_{n \in \mathbf{N}}$  sont des familles d'éléments de  $\mathbf{K}$  presque nulles, alors  $PQ = \sum_{n \in \mathbf{N}} p_n X^n$  avec  $p_n = \sum_{\substack{i=1 \dots n \\ j=1 \dots n \\ i+j=n}} a_i b_j$ .

Effectuer le produit de deux séries entières  $\sum a_n z^n$  et  $\sum b_n z^n$  revient formellement à « regrouper les termes de degré  $n$  » dans  $\sum a_n z^n \sum b_n z^n$ .

Dans la suite nous noterons :

- $R_s$  Le rayon de convergence de la série somme  $\sum s_n z^n$  ;
- $R_p$  Le rayon de convergence de la série somme  $\sum p_n z^n$ .

On a alors le résultat suivant :

**Proposition 2.2.3** — RAYON DE CONVERGENCE DE LA SOMME ET DU PRODUIT —

1. (a)  $\boxed{R_s \geq \min\{R, R'\}}$  ;

(b) De plus si  $R \neq R'$  alors  $\boxed{R_s = \min\{R, R'\}}$  ;

(c) Soit  $z_0$  un complexe tel que  $|z_0| < \min\{R, R'\}$ , (resp. un élément  $t_0$  de  $\mathbf{R}$  tel que  $|t_0| < \min\{R, R'\}$ ), alors :

$$\begin{aligned} \sum_{n=0}^{+\infty} s_n z_0^n &= \sum_{n=0}^{+\infty} a_n z_0^n + \sum_{n=0}^{+\infty} b_n z_0^n, \\ (\text{resp. } \sum_{n=0}^{+\infty} s_n t_0^n &= \sum_{n=0}^{+\infty} a_n t_0^n + \sum_{n=0}^{+\infty} b_n t_0^n). \end{aligned}$$

2. (a)  $\boxed{R_p \geq \min\{R, R'\}}$  ;

(b) Soit  $z_0$  un complexe tel que  $|z_0| < \min\{R, R'\}$ , (resp. un élément  $t_0$  de  $\mathbf{R}$  tel que  $|t_0| < \min\{R, R'\}$ ), alors :

$$\begin{aligned} \sum_{n=0}^{+\infty} p_n z_0^n &= \sum_{n=0}^{+\infty} a_n z_0^n \times \sum_{n=0}^{+\infty} b_n z_0^n, \\ (\text{resp. } \sum_{n=0}^{+\infty} p_n t_0^n &= \sum_{n=0}^{+\infty} a_n t_0^n \times \sum_{n=0}^{+\infty} b_n t_0^n). \end{aligned}$$

**Remarque** — On peut avoir  $R \neq R'$  et cependant  $R_p > \min(R, R')$ . Voici un exemple. Pour tout entier naturel  $n$ , on pose :

- $a_n = 1$ .
- $b_n = 0$ , pour  $n \geq 2$ ,  $b_0 = 1$ ,  $b_1 = -1$ .

On a immédiatement  $R = \dots$ ,  $R' = \dots$ . Pour tout entier naturel  $n$ ,  $p_n = \dots$ . Donc  $R_p = \dots$ , et donc  $R_p > \min(R, R')$ .

**Preuve de la proposition 2.2.3** —

1. Soit  $z_0$  un complexe tel que  $|z_0| < \min(R, R')$ .

2. Soit  $z_0$  un complexe tel que  $|z_0| < \min(R, R')$ . Pour tout entier naturel  $n$  on pose :  $A_n := a_n z_0^n$ ,  $B_n = b_n z_0^n$  et  $P_n = \sum_{i=0}^n A_i B_{n-i}$ , de sorte que :

$$P_n = \dots$$

## 2.3 Intégration et dérivation des séries entières

Soient  $n$  un entier naturel,  $a_n$  un élément de  $\mathbf{C}$  et  $u_n$  l'application

$$u_n : \mathbf{R} \rightarrow \mathbf{C}; t \mapsto a_n t^n.$$

L'application  $u_n$  est  $\mathcal{C}^1$  et

$$u'_n : \mathbf{R} \rightarrow \mathbf{C}; t \mapsto n a_n t^{n-1}.$$

Par ailleurs

$$U_n : \mathbf{R} \rightarrow \mathbf{C}; t \mapsto \frac{a_n t^n}{n+1}.$$

est une primitive de  $U_n$ .

On va définir par analogie de façon purement formelle les séries entières dérivées et primitives d'une série entière.

**Définition 2.3.1.** — Soit  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ).

1. Sa série entière dérivée est par définition la série entière :

$$\sum_{n \geq 1} n a_n z^{n-1} (t^{n-1}) = \sum_{n \geq 0} (n+1) a_{n+1} z^n (t^n).$$

2. Sa série entière primitive est par définition la série entière :

$$\sum_{n \geq 0} \frac{a_n}{n+1} z^{n+1} (t^{n+1}) = \sum_{n \geq 1} \dots z^n (t^n).$$

**Remarque 2.3.2** —

1. La série dérivée de la série primitive est ...
2. La série primitive de la série dérivée est ...

Les séries dérivée et primitive n'ont qu'un sens formel. Afin de leur donner un sens, nous allons déterminer leur rayon de convergence.

**Proposition 1.3.3.** — Soit  $\sum a_n z^n$  (resp.  $\sum a_n t^n$ ) une série entière de la variable complexe  $z$  (resp. réelle  $t$ ). La série entière  $\sum a_n z^n (t^n)$  sa série entière dérivée et sa série entière primitive ont même rayon de convergence.

*Preuve de la proposition 1.3.3.* — Remarquons qu'il suffit de prouver que la série entière et sa série entière dérivée ont

même rayon compte tenu de ...

**Proposition 2.3.4.** — Soit  $\sum a_n t^n$  une série entière de la variable réelle  $t$  de rayon de convergence  $R$ , non nul. Soit  $S$  l'application somme :

$$S : ]-R, R[ \rightarrow \mathbf{C} \quad t \mapsto \sum_{n=0}^{+\infty} a_n t^n.$$

Soit  $a$  un élément de  $] -R, R[$ . Pour tout élément  $t$  de  $] -R, R[$ ,  $\int_a^t S(\tau) d\tau$  est définie, la série  $\sum_{n \geq 0} a_n \int_a^t \tau^n d\tau$  converge

de somme  $\int_a^t S(\tau) d\tau$  :

$$\sum_{n=0}^{+\infty} a_n \int_a^t \tau^n d\tau = \int_a^t S(\tau) d\tau.$$

*Preuve de la proposition 2.3.4. —*

**Pratique** — Gardons le cadre de la proposition 2.3.4. Soit  $t$  un élément de  $] - R, R[$ . En particulier, pour tout  $t$  élément de  $] - R, R[$ ,  $\int_0^t S(\tau) d\tau$  est la valeur de la valeur de la somme au point  $t$  de la série primitive :

$$\int_0^t S(\tau) d\tau = \sum_{n=0}^{+\infty} \frac{a_n t^{n+1}}{n+1}.$$

Donc si  $F$  désigne une primitive sur  $] - R, R[$  de  $S$ , alors :

$$F : ] - R, R[ \rightarrow \mathbf{C}; t \mapsto F(0) + \sum_{n=0}^{+\infty} \frac{a_n t^{n+1}}{n+1}.$$

Autrement dit :

*La primitive est obtenue en primitivant terme à terme la somme de la série entière, sans oublier d'ajouter la valeur en 0 de la primitive.*

**Proposition 2.3.4.** — *Soit  $\sum a_n t^n$  une série entière de la variable réelle  $t$ , de rayon de convergence  $R$ , non nul. Soit  $S$  l'application somme. Alors :*

1. *L'application  $S$  est élément de  $\mathcal{C}^\infty(] - R, R[, \mathbf{C})$  ;*
2. *L'application  $S'$  est la somme de la série entière dérivée ;*
3. *Plus généralement, pour tout élément  $k$  de  $\mathbf{N}$  et tout élément  $t$  de  $] - R, R[$ ,*

$$S^{(k)}(t) = \sum \dots$$

*Autrement dit, la dérivée  $k^e$  de la somme s'obtient. . .*

*Preuve de la proposition 2.2.5. —*

## 2.4 Développement en séries entière d'une application

Nous revenons à la préoccupation initiale de cette partie. Ecrire une application comme la somme d'une série entière.

**Définition 2.4.1.** — Soit  $f$  une fonction de la variable réelle  $t$ , à valeurs réelles ou complexes définie au voisinage d'un réel  $t_0$ . On dit que  $f$  est développable en série entière, au voisinage de  $t_0$ , si, par définition, il existe une série entière  $\sum a_n t^n$  de rayon de convergence  $R$  non nul, un élément  $r$  de  $]0, R]$ , tels que pour tout élément  $t$  de  $]t_0 - r, t_0 + r[$ ,  $f(t)$  soit défini et donné par :

$$f(t) = \sum_{n=0}^{+\infty} a_n (t - t_0)^n.$$

Cette dernière égalité s'appelle développement de  $f$  en série entière.

**Remarque** — Soit  $f$  une fonction de la variable réelle  $t$ , à valeurs réelles ou complexes définie au voisinage d'un réel  $t_0$ . Soit  $g$  la fonction définie au voisinage de 0 par :  $g(t) = f(t_0 + t)$ . Alors on vérifie immédiatement que  $f$  est développable en séries entière, au voisinage de  $t_0$ , si et seulement si  $g$  est développable en série entière, au voisinage de 0.

C'est pourquoi dans la suite nous nous limiterons aux développements en séries entières au voisinage de 0. Nous parlerons simplement de fonctions développables en séries entière en sous-entendant au voisinage de 0.

### RECHERCHE DE CONDITIONS NÉCESSAIRES

Soit  $f$  une fonction de la variable réelle  $t$ , à valeurs réelles ou complexes définie au voisinage de 0.

Supposons que  $f$  soit développable en série entière. Il existe une série entière  $\sum a_n t^n$  de rayon de convergence  $R$  non nul, un élément  $r$  de  $]0, R]$ , tels que pour tout élément  $t$  de  $] - r, +r[$ , on ait  $f(t)$  défini et donné par :

$$f(t) = \sum_{n=0}^{+\infty} a_n t^n.$$

On déduit de la proposition...

que  $f|_{]-r, r[}$  est de classe  $\mathcal{C}^\infty$ . De plus

Donc, pour tout entier naturel  $k$

$a_k = \text{_____}$

On a donc prouvé le résultat d'unicité suivant

**Proposition 2.4.2.** — UNICITÉ DU DÉVELOPPEMENT EN SÉRIE ENTIÈRE — Soit  $f$  une fonction de la variable réelle  $t$ , à valeurs réelles ou complexes, définie au voisinage de 0, développable en série entière. Alors elle est de classe  $\mathcal{C}^\infty$  au voisinage de 0 et admet un et un seul développement en série entière :

$$f(t) = \sum_{n=0}^{+\infty} \dots$$

En particulier si des séries entières de la variable réelle ou complexe  $x$ ,  $\sum a_n x^n$  et  $\sum b_n x^n$  ont des sommes qui coïncident sur un voisinage de 0 alors les suites  $(a_n)_{n \in \mathbf{N}}$  et  $(b_n)_{n \in \mathbf{N}}$  sont égales.

Résumons, si  $f$  est développable en série entière alors elle est  $\mathcal{C}^\infty$  au voisinage de 0 et la série entière qui intervient dans la définition 2.4.1. est

$$\sum \dots$$

Plus généralement, pour une fonction  $f$ ,  $\mathcal{C}^\infty$  au voisinage de 0, développable ou non en série entière, la série entière précédente s'appelle série de *Mac-Laurin* de  $f$ .

**Remarque** — Il résulte de 2.4.2. que si  $f$  soit développable en série entière et est paire (resp. impaire) alors

**Mise en garde** — Il ne faut pas croire que toute fonction  $f$ ,  $\mathcal{C}^\infty$  au voisinage de 0 est développable en série entière. D'abord rien n'assure que sa série de Mac-Laurin ait un rayon de convergence non nul. Il y a pire ! La série de Mac-Laurin peut avoir un rayon de convergence non nul, sans que pour autant il existe un voisinage de 0 sur lequel la somme de la série de Mac-Laurin et la fonction  $f$  coïncident. Donnons un exemple relevant de cette dernière situation.

**Exemple** — Soit l'application  $f : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \begin{cases} 0, & \text{si } t \leq 0, \\ \exp(-\frac{1}{t^2}), & \text{si } t > 0. \end{cases}$

#### RECHERCHE DE CONDITIONS SUFFISANTES

Il n'existe pas au programme de conditions suffisantes pour qu'une fonction soit développable en série entière. Une méthode parfois employée consiste à utiliser la...

Explication. Soit  $f$  une fonction de la variable réelle  $t$ , à valeurs réelles ou complexe, telle qu'il existe un réel  $a$ , tel que  $f$  soit  $\mathcal{C}^\infty$  sur  $] -a, a[$ . pour tout élément  $t$  de  $] -a, a[$ ,

$$f(t) = \dots$$



On peut aussi, en partant de développements en série entière connus, en générer d'autre grâce à des opérations algébriques simples (somme, produit...) ou par primitivation ou dérivation. Il convient donc de disposer d'un petit nombre de développements en série entière, dits de référence.

#### DÉVELOPPEMENT EN SÉRIE ENTIÈRE DE RÉFÉRENCE

Les développements en série entière encadrés sont au programme.

• *Exponentielle, fonctions trigonométriques*

On a déjà vu que par définition, pour tout complexe  $z$ ,

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

On en déduit, pour tout élément  $a$  de  $\mathbf{C}$ , le développement en série entière

$$\exp(at) = \sum_{n=0}^{+\infty} \frac{(at)^n}{n!}, \text{ pour tout } t \in \mathbf{R}$$

En particulier

$$\exp(t) = \sum_{n=0}^{+\infty} \frac{t^n}{n!}, \text{ pour tout } t \in \mathbf{R}$$

 (IX.4)

et

$$\exp(it) = \sum_{n=0}^{+\infty} \frac{(it)^n}{n!}, \text{ pour tout } t \in \mathbf{R}. \quad (\text{IX.5})$$

On déduit de (IX.4)

$$\operatorname{ch}(t) = \sum_{n=0}^{+\infty} \frac{t^{2n}}{(2n)!}, \text{ pour tout } t \in \mathbf{R}.$$

et

$$\operatorname{sh}(t) = \sum_{n=0}^{+\infty} \frac{t^{2n+1}}{(2n+1)!}, \text{ pour tout } t \in \mathbf{R}.$$

On déduit de (IX.5)

$$\cos(t) = \sum_{n=0}^{+\infty} \frac{(-1)^n t^{2n}}{(2n)!}, \text{ pour tout } t \in \mathbf{R}$$

et

$$\sin(t) = \sum_{n=0}^{+\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!}, \text{ pour tout } t \in \mathbf{R}$$

• *Etude de  $\ln(1+t)$*

On a vu que pour tout complexe  $z$  de module strictement plus inférieur à 1, que la série géométrique  $\sum z^n$  converge et :

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n.$$

On en déduit immédiatement que :

$$\frac{1}{1-t} = \sum_{n=0}^{+\infty} t^n, \text{ pour tout } t \in ]-1, 1[$$

Donc

$$\frac{1}{1+t} = \sum_{n=0}^{+\infty} (-1)^n t^n, \text{ pour tout } t \in ]-1, 1[.$$

On peut en déduire le développement en série entière de  $t \mapsto \ln(1+t)$ , en effet

$$\ln(1+t) = \sum_{n=0}^{+\infty} \frac{(-1)^n t^{n+1}}{n+1}, \text{ pour tout } t \in ]-1, 1[$$

On a de même

$$\ln(1-t) = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1} t^{n+1}}{n+1}, \text{ pour tout } t \in ]-1, 1[$$

• *Etude de  $(1+t)^\alpha$*

Soit  $\alpha$  un réel. Étudions  $t \mapsto (1+t)^\alpha$ .

— Premier cas :  $\alpha$  est un entier naturel

— Second cas :  $\alpha \notin \mathbf{N}$

Soit  $f : ]-1, 1[ \rightarrow \mathbf{R}; t \mapsto (1+t)^\alpha$ .  $f$  est de classe  $\mathcal{C}^\infty$  et pour tout entier naturel  $n$  non nul,

$$f^{(n)} : ]-1, 1[ \rightarrow \mathbf{R}; t \mapsto \dots$$

Soit  $t$  élément de  $] -1, 1[$ . Pour tout élément  $n$  de  $\mathbf{N}^*$ , la formule de Taylor avec reste intégral, à l'ordre  $n$ , s'écrit

$$f(t) = 1 + \sum_{k=1}^n \frac{f^{(k)}(0)}{k!} t^k + R_n(t),$$

où l'on a posé :  $R_n = \int_0^t \frac{f^{(n+1)}(u)}{n!} (t-u)^n du$ .

Commençons par examiner le cas où  $t \in ]-1, 0]$ . Pour tout  $n \in \mathbf{N}^*$ ,

$$|R_n(t)| \leq \frac{|\alpha| |\alpha-1| \dots |\alpha-n|}{n!} \int_t^0 \left( \frac{u-t}{1+u} \right)^n (1+u)^{\alpha-1} du.$$

Étudions l'application  $h : ]-1, 0[ \rightarrow \mathbf{R}; u \mapsto \frac{u-t}{1+u}$ .

Donc dans le cas où  $t \in ]-1, 0]$ ,  $R_n(t) \xrightarrow{n \rightarrow +\infty} 0$ .

On montre de même, dans le cas où  $t \in ]0, 1[$ , que :  $R_n(t) \xrightarrow{n \rightarrow +\infty} 0$ .

Pour tout  $t \in ]-1, 1[$ ,  $(1+t)^\alpha = 1 +$

Posons pour tout  $n \geq 1$ ,  $\binom{\alpha}{n} := \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}$  et  $\binom{\alpha}{0} := 1$ .

On a, pour  $\alpha$  élément de  $\mathbf{R} - \mathbf{N}$

$$(1+t)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} t^n$$

**Remarque** — Pour  $\alpha = -1$ , on retrouve  $\frac{1}{1+t} = \sum_{n=0}^{+\infty} (-1)^n t^n$ .

On peut déduire directement de ce dernier développement, les développements de arcsin, arctan.

Traitons pour commencer le cas au programme de arctan.

$$\arctan(t) = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{2n+1}, \text{ pour tout } t \in ]-1, 1[$$

Passons à  $\arcsin$ .

On verra en exercice d'autres procédés de détermination de développements en série entière, notamment, l'utilisation d'équations différentielles, d'équations algébriques...

## 2.5 Fonctions génératrices

Nous allons donner une application des séries entières aux probabilités.

Dans ce paragraphe, par  $X$  nous désigneront une variable aléatoire à valeurs dans  $\mathbf{N}$ . Nous ne nous intéressons pas à l'espace probabilisé sur lequel elle est défini, reléguant ces considérations dans le chapitre de probabilités, seul nous importe sa loi.

Nous pouvons associé à  $X$  une série entière de la variable réelle  $t$ ,  $\sum P_X(n)t^n$ . Nous allons étudier le rayon de convergence de cette série et voir ses utilisations concrètes dans le calcul des moments de  $X$ , et même dans la détermination de la loi de  $X$  lorsque celle-ci nous est inconnue.

**Proposition-définition 2.5.1.** — *La série entière  $\sum P_X(n)t^n$  a un rayon de convergence supérieur ou égal à 1. La somme de cette série entière s'appelle fonction génératrice de  $X$  on la note  $G_X$ .*

*Preuve de la proposition 2.5.1. —*

**Proposition 2.5.2.** — *La fonction génératrice  $G_X$  est continue sur  $[-1, 1]$ .*

*Preuve de 2.5.2. —*

*Premier cas* : le rayon de convergence  $R$  de  $\sum P_X(n)t^n$  est strictement supérieur à 1. alors...

*Second cas* :  $R = 1$

### Exemples

- $X$  suit une loi de Bernoulli de paramètre  $p$ . alors  $R = +\infty$  et pour tout  $t \in \mathbf{R}$ ,

$$G_X(t) = \dots\dots\dots$$

- $X$  suit une loi binomiale de paramètre  $(n, p)$ . alors  $R = +\infty$  et pour tout  $t \in \mathbf{R}$ ,

$$G_X(t) = \dots\dots\dots$$

- $X$  suit une loi de Poisson de paramètre  $\lambda$ . alors

$$G_X(t) = \dots\dots\dots$$

La fonction génératrice est définie par la loi de  $X$ , il est à noter que l'on peut retrouver la loi de  $X$  à partir de  $G_X$  de façon très simple :

$$P_X(k) = \dots\dots\dots, \text{ pour tout entier } k$$

Nous allons voir comment la fonction génératrice de  $X$  permet d'en déterminer les moments.

**Proposition 2.5.3.** — *La variable aléatoire  $X$  admet une espérance finie si et seulement si  $G_X$  est dérivable en 1 et si c'est le cas :*

$$E(X) = \sum_{n=1}^{+\infty} nP_X(n) = G'_X(1).$$

*Preuve de la proposition 2.5.3.* —

Là encore deux cas. Le premier est celui trivial où  $R$  est strictement supérieur à 1, alors....

Passons au second cas, celui où  $R = 1$ .

- Supposons que  $X$  admette une espérance finie.

• Montrons que si  $X$  est dérivable en 1 alors  $\sum_{n \geq 1} nP_X(n)$  converge, ou plutôt montrons la contraposée de cette implication.

On suppose donc que  $\sum_{n \geq 1} nP_X(n)$  diverge.

D'où par contraposition si  $X$  est dérivable en 1 alors  $\sum_{n \geq 1} nP_X(n)$  converge.

D'où 2.5.3.

**Proposition 2.5.4.** —  $X$  admet une espérance finie et une variance si et seulement si  $G_X$  est deux fois dérivable dérivable en 1 et si tel est le cas :

$$V(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2.$$



*Preuve de la proposition 2.5.4. —*

**Proposition 2.5.5.** — Soient  $X_1$  et  $X_2$  des variables aléatoires définies sur un univers  $\Omega$  à valeurs dans  $\mathbf{N}$  indépendantes. Alors pour tout  $t \in [-1, 1]$

$$G_{X_1+X_2}(t) = G_{X_1}(t) \times G_{X_2}(t).$$

*Preuve de la proposition 2.5.5. —*

**Corollaire 2.5.6.** — Soient  $X_1, X_2, \dots, X_p$  des variables aléatoires définies sur un univers  $\Omega$  à valeurs dans  $\mathbf{N}$  mutuellement indépendantes. Alors pour tout  $t \in [-1, 1]$

$$G_{X_1+X_2+\dots+X_p}(t) = G_{X_1}(t) \times G_{X_2}(t) \times \dots \times G_{X_p}(t).$$

*Preuve du corollaire 2.5.6. —*

### Exercices —

1. Soient deux variable aléatoires  $X_1$  et  $X_2$  définies sur un même espace probabilisé et qui suivent des lois de Poisson de paramètres respectifs  $\lambda_1$  et  $\lambda_2$ . Etudiez la loi de  $X_1 + X_2$ .
2. On suppose que la variable aléatoire  $X$  suit une loi binomiale de paramètre  $(m, p)$  et que  $X = X_1 + X_2$  où  $X_1$  et  $X_2$  sont des variables aléatoires indépendantes à valeurs dans  $\mathbf{N}$ . Déterminer les lois de  $X_1$  et de  $X_2$ .

*Solution —*

Afin de mesurer l'efficacité des fonction génératrices, introduisons une nouvelle loi : la loi géométrique.

### LOI GÉOMÉTRIQUE —

Soit  $p \in ]0, 1[$ . On dit que  $X$  suit une loi géométrique de paramètre  $p$ , et on note  $X \sim \mathcal{G}(p)$  si pour tout entier  $k \geq 1$

$$\boxed{P_X(k) = (1-p)^{k-1}p.}$$

Cette formule définit bien une loi puisque pour tout  $k \in \mathbf{N}^*$ ,  $(1-p)^k p \geq 0$  et

$$\sum_{k=1}^{\infty} (1-p)^{k-1} p = \dots$$

Cette loi intervient de façon « naturelle » dans beaucoup de situations. Considérons par exemple une suite de variables aléatoires indépendantes  $(Y_i)_{i \in \mathbf{N}}$  qui suivent une loi de Bernoulli de même paramètre  $p$ . Cette suite peut modéliser une suite de tirages avec remise dans une urne contenant des boules noires et des boules blanches,  $p$  étant la probabilité

de piocher une boule blanche (le rapport du nombre de boules blanches sur le nombre total de boules, dans le cas de tirage équiprobables). Soit alors  $X$  la variable aléatoire...

$X$  suit une loi géométrique

Le calcul direct de l'espérance  $E(X) = p \sum_{k=1}^{+\infty} k(1-p)^{k-1}$ , n'est pas trivial. Par contre

$$G_X = \sum_{k=1}^{+\infty} \dots$$

Finalement si  $X \sim \mathcal{G}(p)$  alors :  $\boxed{E(X) = \frac{1}{p}}, \quad \boxed{V(X) = \frac{1-p}{p^2}}.$

Pour clore en beauté le présent chapitre donnons une utilisation voisine des séries entières dans un problème de dénombrement.

### Exercice —

Pour tout entier  $n$  supérieur ou égal à 1 on note  $I_n$  le nombre d'involutions de  $\{1, 2, \dots, n\}$ , c'est-à-dire le nombre d'applications  $f$  de  $\{1, 2, \dots, n\}$  dans  $\{1, 2, \dots, n\}$  telles que  $f \circ f = id$ . On se propose de déterminer pour tout entier  $n \geq 1$ ,  $I_n$ . On admet le fait évident, que  $I_n$  est plus généralement le nombre d'involutions d'un quelconque ensemble à  $n$  éléments.

1. (a) Montrer que toute involution est bijective.
- (b) Pour tout entier  $n$  supérieur ou égal à 1, montrer que  $I_n$  est le nombre d'involution de  $\{1, 2, \dots, n+1\}$  qui fixent  $n$ .
- (c) Pour tout entier  $n$  supérieur ou égal à 2 montrer que :

$$I_{n+1} = I_n + nI_{n-1}$$

2. On se propose d'étudier la série entière de la variable réelle  $t$ ,  $\sum_{n \geq 1} \frac{I_n}{n!} t^n$ . Notons que  $\frac{I_n}{n!}$  représente la probabilité pour qu'une permutation de  $\{1, \dots, n\}$  soit une involution lorsque  $S_n$  s'est vu muni de la probabilité uniforme.

- (a) Montrer que le rayon  $R$  de la série entière  $\sum_{n \geq 1} \frac{I_n}{n!} t^n$  n'est pas nul. On note  $S$  sa somme.
- (b) Montrer que  $S$  satisfait à une équation différentielle linéaire.
- (c) En déduire l'expression de  $S$ .
- (d) En déduire l'expression de  $I_n$  pour tout entier  $n \geq 1$ .
- (e) Comment reprendre cette dernière question sans avoir prouvé *a priori* que  $R$  est non nul.



## Chapitre X

# COMPLÉMENT SUR L'INTÉGRALE

Le présent chapitre traite de l'intégrale. La première partie est dévolue à l'étude des intégrales dites généralisées ou impropres c'est-à-dire sur un intervalle quelconque. La seconde étudie le passage à la limite dans des intégrales éventuellement généralisées.

## INTÉGRALE GÉNÉRALISÉE

### 1.1 Définitions, première propriétés

Nous allons définir l'intégrale d'applications continues par morceaux sur un intervalle quelconque. Pour ce faire nous utiliserons l'intégrale sur un segment déjà définie.

**Définition 1.1.1.** — INTÉGRALE IMPROPRE OU GÉNÉRALISÉE — Soit  $f$  une application d'un intervalle  $[a, b[$  où  $a < b \leq +\infty$  (resp.  $]a, b]$  où  $-\infty \leq a < b$ ), à valeurs dans  $\mathbf{K}$ , continue par morceaux. On dit que l'intégrale  $\int_a^b f(t)dt$  converge si par définition,

$$[a, b[ \rightarrow \mathbf{K}; x \mapsto \int_a^x f(t)dt$$

$$( \text{ resp. } ]a, b] \rightarrow \mathbf{K}; x \mapsto \int_x^b f(t)dt )$$

admet une limite lorsque  $x$  tend vers  $b$  (resp. vers  $a$ ).

Si c'est le cas on note  $\int_a^b f(t)dt$  la limite elle-même.

Si l'intégrale ne converge pas on dira sans surprises qu'elle diverge.

Plus généralement, si  $g$  est une application d'un intervalle  $I$  qui contient  $[a, b[$  à valeurs dans  $\mathbf{K}$  et continue par morceaux, alors l'intégrale généralisée  $\int_a^b g|_{[a, b[}(t)dt$  se note plus simplement  $\int_a^b g(t)dt$ .

Nous avons dans la définition précédente envisagé le cas d'une application définie aussi bien sur un intervalle de la forme  $[a, b[$  que sur un de la forme  $]a, b]$ . Dans la suite, afin de ne pas alourdir notre propos, nous énoncerons le plus souvent les résultats pour le premier type d'intégrale, à charge pour le lecteur d'en adapter les énoncés dans le second cas.

Exemples 1.1.2. —

1. L'intégrale  $\int_0^{+\infty} t dt \dots$

2. L'intégrale  $\int_0^{+\infty} e^{-t} dt \dots$

3. L'intégrale  $\int_0^1 \ln t dt$  ....

**Notations** — On peut écrire plus simplement  $\int_a^b f$  l'intégrale généralisée  $\int_a^b f(t)dt$ , on trouve encore une autre notation, plus précise puisqu'elle contient la nature du domaine de définition de l'intégrande,  $\int_{[a,b[} f$ .

A ce propos observons que dans la pratique on ne définit pas une application  $f$  d'un intervalle  $[a, b[$  dans  $\mathbf{K}$ , mais l'on donne une expression de la forme :

$$\int_a^b f(t)dt$$

où  $f(t)$  est une expression, fonction de  $t$ , plus ou moins compliquée utilisant des applications usuelles. Si le domaine de définition de cette fonction contient  $[a, b]$  (qui est alors nécessairement un segment) et que  $[a, b] \rightarrow \mathbf{K}; t \mapsto f(t)$  est continue par morceaux, on a à faire en fait à une intégrale ordinaire d'une application continue sur un segment :  $\int_{[a,b]} f$ . Si par contre l'une des borne de l'intégrale, disons  $b$ , n'est pas dans le domaine de définition de cette fonction qui est cependant définie sur  $[a, b[$ , et que  $[a, b[ \rightarrow \mathbf{K}; t \mapsto f(t)$  est continue par morceaux, c'est que nous avons à faire à une intégrale généralisée. On verra dans un instant la généralisation au cas où la fonction n'est définie ni en  $a$  ni en  $b$ .

Les notation  $\int_a^b f(t)dt$  ou  $\int_a^b f$  qui occultent le domaine de définition de l'intégrande, ne sont pas en fait problématiques. En effet si l'on considère une application  $f$  d'un segment  $[a, b]$  dans  $\mathbf{K}$  continue par morceaux et que l'on considère  $f_1$  sa restriction à  $[a, b[$  alors  $\int_a^b f_1$  converge et  $\int_a^b f_1 = \int_{[a,b]} f$  en effet....

Bref, si l'on considère sans raison une intégrale ordinaire comme une intégrale impropre, cela n'a pas d'importance, il convient néanmoins d'étudier avec soin avant de se lancer dans une étude inutile et peut-être coûteuse, si oui ou non on peut inclure les bornes dans le domaine de définition de l'intégrande.

Dans le même ordre d'idée dans  $\int_a^b f(t)dt$  une des bornes,  $b$  par exemple, peut ne pas appartenir aux domaine de définition de  $f$ , mais l'application  $f$  peut admettre un prolongement par continuité  $\tilde{f}$  à  $[a, b]$ . C'est le cas par exemple lorsque l'on étudie

$$\int_{-\pi}^0 \frac{\sin t}{t} dt$$

où  $[-\pi, 0[ \rightarrow \mathbf{R}; t \mapsto \frac{\sin t}{t}$  se prolonge par continuité à  $[-\pi, 0]$  par la valeur.....en 0.

Alors en vertu de ce qui vient d'être dit l'intégrale converge. Dans la pratique on écrit :

« ....

**Remarque 1.1.3.** — Soit  $f$  une application d'un intervalle  $[a, b[$  où  $a < b \leq +\infty$ , alors pour tout  $a' \in [a, b[$ , l'intégrale impropre  $\int_a^b f(t)dt$  converge si et seulement si  $\int_{a'}^b f(t)dt$  converge (avec les abus de notations mentionnés dans la définition 1.1.1.).

De plus si c'est le cas  $\int_{a'}^b f(t)dt$  tend vers 0 lorsque  $a'$  tend vers  $b$ .

En effet soit  $a' \in [a, b[$ . Pour tout élément  $x$  de  $[a, b[$  la loi de Chasles pour l'intégrale sur un segment affirme :

$$\int_a^x f(t)dt = \int_a^{a'} f(t)dt + \int_{a'}^x f(t)dt$$

Donc  $\int_a^x f(t)dt$  admet une limite lorsque  $x$  tend vers  $a$  si et seulement si  $\int_{a'}^x f(t)dt$  en admet une. □

**Proposition, définition 1.1.4.** — Soit  $f$  une application d'un intervalle  $]a, b[$  où  $-\infty \leq a < b \leq +\infty$  à valeurs dans  $\mathbf{K}$ , continue par morceaux. Si il existe un réel  $c$  élément de  $]a, b[$  tel que  $\int_c^b f$  et  $\int_a^c$  convergent, alors, pour tous élément  $d$  de  $]a, b[$   $\int_a^b f$  et  $\int_a^d$  convergent et si tel est le cas  $\int_a^d f + \int_d^b f = \int_a^c f + \int_c^b f$ .

Si tel est le cas on dit que l'intégrale  $\int_a^b f(t)dt$  converge, et l'on note on note  $\int_a^b f(t)dt$  ou  $\int_a^b f$  ou  $\int_{]a,b[} f$  la quantité  $\int_a^c f + \int_c^b f$ .

**Exemple 1.1.5.** — Etudions la convergence de l'intégrale  $\int_{-\infty}^{+\infty} \frac{1}{1+t^2} dt$ .

**Notation** — Nous disposons donc de la notion d'intégrale généralisée sur un intervalle  $I$  non réduit à un point quelconque. Notons  $a$  et  $b$  les extrémités de  $I$  avec  $a < b$ . Les point  $a$  et  $b$  sont ou ne sont pas éléments de  $I$ , le cas ou  $a$  et  $b$  seraient éléments de  $I$  est sans intérêt puisqu'il nous ramène au chapitre sur les intégrales sur un segment. Soit enfin  $f$  une application de  $I$  dans  $\mathbf{K}$  telle que  $\int_a^b f(t)dt$  converge Soient  $\alpha, \beta$  des éléments de  $I \cup \{a, b\}$ , avec  $\alpha < \beta$ . Alors d'après 1.1.4. l'intégrale  $\int_\alpha^\beta$  converge et l'on pose par convention, comme pour l'intégrale sur un segment :

$$\int_\beta^\alpha f = - \int_\alpha^\beta f, \int_\alpha^\alpha f = 0.$$

On dit aussi que  $\int_\beta^\alpha$  converge.

Avec cette convention la loi de Chasles vue pour l'intégrale sur un segment perdure.

**Proposition 1.1.6.** — LOI DE CHASLES —

Soit un intervalle  $I$  non réduit à un point quelconque. Notons  $a$  et  $b$  les extrémités de  $I$  avec  $a < b$ . Soient  $\alpha, \beta$  et  $\gamma$  des éléments de  $I \cup \{a, b\}$ , avec  $\alpha < \beta$  et enfin  $f$  une application de  $I$  dans  $\mathbf{K}$  telle que  $\int_a^b f(t)dt$  converge. Alors les intégrales  $\int_\alpha^\beta f, \int_\alpha^\gamma f, \int_\beta^\gamma f$  convergent et :

$$\int_\alpha^\beta f = \int_\alpha^\gamma f + \int_\gamma^\beta f.$$

*Preuve de la proposition 1.1.6.* — La preuve est particulièrement simple et longue... à détailler, traitons par exemple le cas où  $\alpha$  et  $\beta$  sont dans  $I$  et  $\gamma = a$ ,  $a$  n'étant pas élément de  $I$

Pour tout élément  $x$ , de  $]a, b[$  la loi de Chasles pour l'intégrale sur un segment nous laisse écrire :

$$\int_\alpha^\beta f = \int_\alpha^x f + \int_x^\beta f.$$

Laissons tendre  $x$  vers  $\gamma$ , c'est-à-dire ici  $a$  et la formule est prouvée. □

**Proposition 1.1.7.** — Soit  $I$  un intervalle quelconque d'extrémité  $a$  et  $b$  avec  $a < b$ . Soient  $f$  et  $g$  des applications de  $I$  dans  $\mathbf{K}$ , continues par morceaux et  $\lambda$  et  $\mu$  des élément de  $\mathbf{K}$ .

Si les intégrales  $\int_a^b f$  et  $\int_a^b g$  convergent, alors  $\int_a^b \lambda f + \mu g$  converge et :

$$\int_a^b \lambda f + \mu g = \lambda \int_a^b f + \mu \int_a^b g.$$

*Preuve de la proposition 1.1.7.* — Le cas où  $I$  est un segment nous ramène évidemment à un résultat connu.

Traitons par exemple le cas où  $I = [a, b[$ .

$$\int_a^x \lambda f + \mu g = \lambda \int_a^x f + \mu \int_a^x g,$$

pour tout élément  $x$  de  $[a, b[$ , grâce à la linéarité de l'intégrale sur un segment. Reste à laisser  $x$  tendre vers  $a$  pour que l'on ait le résultat. □



## 1.2 Intégrale généralisée de fonctions positives, fonction intégrables

Examinons à présent le cas des intégrales généralisées d'applications positives. Nous allons retrouver dans cette étude des points communs avec les séries.

Dans tout ce paragraphe  $\phi$  désigne une application d'un intervalle  $I$  quelconque d'extrémités  $a$  et  $b$  avec  $a < b$ , dans  $\mathbf{R}$ , à valeurs positives.

Le cas où  $I$  est un segment est dépourvu d'intérêt. Dans les démonstrations on traitera le cas où  $I = [a, b[$  avec  $a < b \leq +\infty$ . Le cas d'un intervalle de la forme  $]a, b]$ , se traite de même, celui d'un intervalle de la forme  $]a, b[$ , se ramène par définition au cas des deux types d'intervalles précédentes.

**Proposition 1.2.1.** — *L'intégrale  $\int_a^b \phi$  converge si et seulement si*

$$[a, b[ \rightarrow \mathbf{R}; x \mapsto \int_a^x \phi(t) dt$$

*est majorée.*

*Preuve de la proposition 1.2.1.* —

On notera que si  $\int_a^b \phi$  diverge alors  $\int_a^x \phi \xrightarrow{x \rightarrow +\infty} +\infty$

**Corollaire 1.2.2.** — Soit  $\psi$  une application de  $[a, b[$  dans  $\mathbf{R}$  à valeurs positives telle que pour tout élément  $t$  de  $[a, b[$ ,

$$\phi(t) \leq \psi(t).$$

Si  $\int_a^b \psi$  converge alors  $\int_a^b \phi$  converge.

*Preuve du corollaire 1.2.2.* —

**Proposition 1.2.3.** — POSITIVITÉ DES INTÉGRALES GÉNÉRALISÉES —

1. Si  $\int_a^b \phi$  converge alors

$$\int_a^b \phi \geq 0.$$

2. si  $f$  et  $g$  sont des applications de  $I$  dans  $\mathbf{R}$ , continues par morceaux, telles que pour tout  $t \in [a, b[$ ,

$$f(t) \leq g(t)$$

Si  $\int_a^b f$  et  $\int_a^b g$  convergent alors

$$\int_a^b f \leq \int_a^b g.$$

*Preuve de la proposition 1.2.3. —*

1. Pour tout  $x$  tel que  $a \leq x < b$ , la bien connu positivité de l'intégrale sur un segment nous assure que :  $\int_a^x \phi \geq 0$ .  
Par passage à la limite dans une inégalité large, voilà que  $\int_a^b \phi \geq 0$ .
2. En appliquant à  $g - f$ , application positive le premier point, grâce à la linéarité de l'intégrale généralisée (1.1.7.), on a :  $\int_a^b f \leq \int_a^b g$ .

D'où la proposition.

**Définition 1.2.4.** Soit  $f$  une application de  $I$  dans  $\mathbf{K}$  continue par morceaux, on dit que l'intégrale généralisée  $\int_a^b f$  converge absolument (ou est absolument convergente), si l'intégrale généralisée  $\int_a^b |f|$  converge.

Si l'intégrale  $\int_a^b |f|$  converge, on dit que  $f$  est intégrable sur  $I$ , (ou simplement intégrable s'il a été clairement stipulé que le domaine de définition de  $f$  est  $I$ .)

**Remarque 1.2.5.** — Prenons pour fixer les idées,  $I = [a, b[$ . Par la remarque 1.1.3., on voit que si  $f$  est intégrable sur  $[a, b[$  alors elle l'est sur  $[a', b[$  pour tout  $a' \in [a, b[$ , c'est pourquoi on dit de manière floue et relâchée «  $f$  est intégrable au voisinage de  $b$  ».

Comme pour les séries on a le résultat crucial suivant :

**Proposition 1.2.6.** — Soit  $f$  une application de  $I$  dans  $\mathbf{K}$  continue par morceaux. Si  $\int_a^b f$  converge absolument alors  $\int_a^b f$  converge, et de plus

$$\left| \int_a^b f \right| \leq \int_a^b |f|.$$

**Remarque** — L'intégrale  $\int_a^b f$  peut converger sans converger absolument. On donnera bientôt un exemple d'une telle intégrale.

*Preuve de la proposition 1.2.6. —*

PREMIER CAS :  $f$  est à valeurs réelles

On note  $f_+$  et  $f_-$  les applications :

$$f_+ : I \rightarrow \mathbf{R}; x \mapsto \begin{cases} f(x), & \text{si } f(x) \geq 0, \\ 0, & \text{si } f(x) < 0, \end{cases}$$

$$f_- : I \rightarrow \mathbf{R}; x \mapsto \begin{cases} -f(x), & \text{si } f(x) \leq 0, \\ 0, & \text{si } f(x) > 0, \end{cases}$$

On a immédiatement les propriétés suivantes :

1. Les applications  $f_+$  et  $f_-$  sont à valeurs positives.
2.  $|f| = f_+ + f_-$ .
3.  $f = f_+ - f_-$ .

SECOND CAS :  $f$  est à valeurs complexes

Dans la pratique pour montrer la convergence d'une intégrale on peut commencer par essayer de montrer sa convergence absolue c'est-à-dire l'intégrabilité de l'intégrande. En effet il existe des théorèmes de comparaisons puissants qui permettent de montrer facilement la convergence d'une intégrale d'une fonction positive ; ils constituent le coeur du paragraphe suivant.

### 1.3 Théorèmes de comparaison

Nous allons dans ce paragraphe donner des théorèmes de comparaison pour les intégrales de fonctions positives.

Dans tout ce paragraphe on considère un intervalle  $I = [a, b[$  avec  $a < b \leq +\infty$ . On laisse le soin au lecteur de traduire les résultats dans le cas où  $I$  est un intervalle quelconque d'extrémités  $a$  et  $b$  avec  $a < b$ , (le cas où  $I$  est un segment est dépourvu d'intérêt).

Nous allons Commencer par généraliser la proposition 1.2.2.

**Proposition 1.3.1.** — Soient  $f$  et  $g$  des applications de  $[a, b[$ , avec  $a < b \leq +\infty$  à valeurs dans  $\mathbf{R}$ , continues par morceaux. On suppose qu'il existe un élément  $a'$  de  $[a, b[$  et un réel  $k \geq 0$  tels que pour tout élément  $t$  de  $[a', b[$ ,

$$0 \leq f(t) \leq kg(t).$$

On a alors :

1. Si  $g$  est intégrable alors  $f$  est intégrable.
2. Si  $f$  n'est pas intégrable alors  $g$  n'est pas intégrable.

On énonce parfois l'hypothèse en disant que  $f$  est positive et majorée par  $kg$  au voisinage de  $b$ .

*Preuve de la proposition 1.3.1.* — Le second point est la contraposée du premier. Le premier résulte directement de la

proposition 1.2.2. en effet

Application —

1. Montrer la convergence de l'intégrale  $\int_{-\infty}^{+\infty} \exp(-t^2) dt$

2. Montrer la convergence de l'intégrale  $\int_0^{+\infty} \frac{\sin t}{t} \exp^{-t} dt$

Une reformulation plus concise de la proposition précédente est :

**Proposition 1.3.2.** — Soient  $f$  et  $g$  des applications de  $[a, b[$ , avec  $a < b \leq +\infty$  à valeurs dans  $\mathbf{R}$ , continues par morceaux. On suppose que  $f$  et  $g$  sont positives au voisinage de  $b$  et que

$$f(t) \underset{t \rightarrow b}{=} O(g(t)).$$

Alors si  $g$  est intégrable, alors  $f$  est intégrable.

Cette proposition admet le corollaire évident :

**Corollaire 1.3.3.** — Soient  $f$  et  $g$  des applications de  $[a, b[$ , avec  $a < b \leq +\infty$  à valeurs dans  $\mathbf{R}$ , continues par morceaux. On suppose que  $f$  et  $g$  sont positives au voisinage de  $b$  et que

$$f(t) \underset{t \rightarrow b}{=} o(g(t)).$$

Alors si  $g$  est intégrable, alors  $f$  est intégrable.

Preuve du corollaire 1.3.3. —

**Corollaire 1.3.4.** — Soient  $f$  et  $g$  des applications de  $[a, b[$ , avec  $a < b \leq +\infty$  à valeurs dans  $\mathbf{R}$ , continues par morceaux. On suppose que  $f$  et  $g$  sont positives au voisinage de  $b$  et que

$$f(t) \underset{t \rightarrow b}{\sim} g(t).$$

Alors  $f$  est intégrable si et seulement si  $g$  l'est.

Preuve du corollaire 1.3.4. —

Notons que si l'une des deux application  $f$  et  $g$  est positive au voisinage de  $a$ , alors les deux le sont.

Lorsque l'on étudie l'intégrabilité d'une application positive sans avoir *a priori* d'idée sur sa nature (convergente ou divergente) le dernier corollaire est à privilégier car il permet de conclure dans les deux cas.

#### FONCTION DE RÉFÉRENCES —

Afin de montrer qu'une intégrale est absolument convergente il faut pouvoir comparer la valeur absolue de l'intégrande à des fonctions positives, dites de référence dont on connaît la nature, intégrable ou non.

- Soit  $\alpha$  un réel. On note

$$f_\alpha : [1, +\infty[ \rightarrow \mathbf{R}; t \mapsto \frac{1}{t^\alpha}.$$

Donc

$f_\alpha$  est intégrable si et seulement si  $\alpha > 1$

, autrement dit : 

$\int_1^{+\infty} \frac{1}{t^\alpha} dt$  converge si et seulement si  $\alpha > 1$

- Soient  $\alpha$  un réel, et  $a$  et  $b$  des réels tels que  $a < b$ . On note

$$g_\alpha : ]a, b] \rightarrow \mathbf{R}; t \mapsto \frac{1}{(t-a)^\alpha}.$$

Donc

$g_\alpha$  est intégrable si et seulement si  $\alpha \dots\dots\dots$  , autrement dit :  $\int_a^b \frac{1}{(t-a)^\alpha} dt$  converge si et seulement si  $\alpha \dots\dots\dots$  • De même  $t \mapsto \frac{1}{|t-b|^\alpha}$  est intégrable sur  $[a, b[$  si et seulement si  $\dots\dots\dots$

### Entrainement 1.3.5.

Déterminer, dans les différents cas suivants, si l'application  $f$  est ou n'est pas intégrable.

1.

$$f : \mathbf{R}_+^* \rightarrow \mathbf{R}; t \mapsto \frac{\sin t}{t} e^{-t}.$$

2. Soit  $g$  un élément de  $\mathcal{C}^0([0, 1], \mathbf{R})$ . On pose :

$$f : [0, 1[ \rightarrow \mathbf{R}; t \mapsto \frac{g(t)}{\sqrt{1-t^2}}.$$

3.

$$f : \left] 0, \frac{1}{2} \right] \rightarrow \mathbf{R}; x \mapsto \frac{1}{\sqrt{x} \ln(x)}.$$

4.

$$f : \left] 0, \frac{1}{2} \right] \rightarrow \mathbf{R}; x \mapsto \frac{\ln(x)}{\sqrt{x}}.$$

5.

$$f : [2, +\infty[ \rightarrow \mathbf{R}; t \mapsto \frac{1}{\ln(t)t^2}.$$

6.

$$f : [2, +\infty[ \rightarrow \mathbf{R}; t \mapsto \frac{\ln(t)^6}{t^2}.$$

7.

$$f : [2, +\infty[ \rightarrow \mathbf{R}; t \mapsto \frac{\ln(t)^6}{t}.$$

8.

$$f : [2, +\infty[ \rightarrow \mathbf{R}; t \mapsto \frac{1}{\ln(t)^2 \sqrt{t}}.$$

9.

$$f : [2, +\infty[ \rightarrow \mathbf{R}; t \mapsto \frac{1}{\ln(t)^a t},$$



où  $a$  est un réel positif.

Donnons à présent un exemple très classique que nous croiserons plusieurs fois dans ce chapitre. Il s'agit, comme promis, d'une intégrale convergente mais non absolument convergente.

$$\int_0^{+\infty} \frac{\sin t}{t} dt.$$

- ABSOLUE CONVERGENCE

- *Etude au voisinage de 0*

- *Etude au voisinage de  $+\infty$*

- CONVERGENCE

**Mise en garde** — Nous avons vu dans la définition de la convergence et de l'absolue convergence d'une intégrale, ainsi que dans les théorèmes de comparaison une similitude forte entre les séries et les intégrales impropres. Il faut cependant faire preuve de prudence dans l'analogie qu'on y pourrait déceler entre la convergence de séries et celle d'intégrales impropres. En particulier on peut trouver une application  $f$  de  $[a, +\infty[$ , continue, positive, et intégrable, qui ne tende pas vers 0 en  $+\infty$  : le fait de converger en  $+\infty$  vers 0 est une CNNNS<sup>1</sup> pour que  $f$  soit intégrable. Donnons un exemple

---

1. Condition ni nécessaire ni suffisante

**Exemple** —

Par contre on peut montrer :

**Exercice 1.3.6.** — Soit  $f$  une application de  $[a, +\infty[$  dans  $\mathbf{R}$ , où  $a$  est un réel, continue par morceaux, intégrable. Si  $f$  admet une limite  $\ell$  en  $+\infty$ , alors  $\ell = 0$ .

*Solution* —

#### INTÉGRATION DES RELATIONS DE COMPARAISON

Comme pour les séries à termes positifs il est possible de sommer des relations de comparaison entre suites, il est possible d'intégrer des relations de comparaison entre applications positives. Toujours comme pour les séries on doit distinguer le cas convergent et divergent.

**Proposition 1.3.7.** — Soient  $\phi$  et  $\psi$  des applications de  $[a, b[$  dans  $\mathbf{R}$ , à valeurs positives .

On suppose que  $\phi(t) = O(\psi(t))$  (resp.  $\phi(t) = o(\psi(t))$ )  
 $\quad \quad \quad t \rightarrow b \quad \quad \quad t \rightarrow b$

- PREMIER CAS :  $\psi$  est intégrable.

Alors  $\phi$  est intégrable et :

$$\int_x^b \phi(t) dt = O_{x \rightarrow b} \left( \int_x^b \psi(t) dt \right)$$

$$(\text{ resp. } \int_x^b \phi(t) dt = o_{x \rightarrow b} \left( \int_x^b \psi(t) dt \right) .)$$

- SECOND CAS :  $\phi$  est non intégrable.

Alors  $\psi$  est non intégrable et :

$$\int_a^x \phi(t) dt = O_{x \rightarrow b} \left( \int_a^x \psi(t) dt \right)$$

$$(\text{ resp. } \int_a^x \phi(t) dt = o_{x \rightarrow b} \left( \int_a^x \psi(t) dt \right) .)$$

*Preuve de la proposition 1.3.7. —*

On montre deux cas :

- l'application  $\phi$  est intégrable et  $\phi(t) = O_{t \rightarrow b}(\psi(t))$ .

- l'application  $\phi$  est non intégrable et  $\phi(t) = o_{t \rightarrow b}(\psi(t))$ .

**Corollaire 1.3.8.** — Soient  $f$  une applications de  $[a, b[$  dans  $\mathbf{K}$ , et  $\psi$  une applications de  $[a, b[$  dans  $\mathbf{R}$ , à valeurs positives .

On suppose que  $f(t) = O_{t \rightarrow b}(\psi(t))$  (resp.  $f(t) = o_{t \rightarrow b}(\psi(t))$ ).

- PREMIER CAS :  $\psi$  est intégrable.

Alors  $f$  est intégrable et :

$$\int_x^b f(t) dt = O_{x \rightarrow b} \left( \int_x^b \psi(t) dt \right)$$

$$(\text{ resp. } \int_x^b f(t) dt = o_{x \rightarrow b} \left( \int_x^b \psi(t) dt \right) .)$$

- **SECOND CAS** :  $f$  est non intégrable.

Alors  $\psi$  est non intégrable et :

$$\int_a^x f(t)dt = O_{x \rightarrow b} \left( \int_a^x \psi(t)dt \right)$$

$$( \text{ resp. } \int_a^x f(t)dt = o_{x \rightarrow b} \left( \int_a^x \psi(t)dt \right) . )$$

*Preuve du corollaire 1.3.8. —*

**Proposition 1.3.9.** —

Soient  $\phi$  et  $\psi$  des applications de  $[a, b[$  dans  $\mathbf{R}$ , à valeurs positives .

On suppose que  $\phi(t) \underset{t \rightarrow b}{\sim} (\psi(t))$

- **PREMIER CAS** :  $\phi$  et  $\psi$  sont intégrables.

Alors  $\phi$  est intégrable et :

$$\int_x^b \phi(t)dt \underset{x \rightarrow b}{\sim} \int_x^b \psi(t)dt.$$

- **SECOND CAS** :  $\phi$  et  $\psi$  sont non intégrables.

Alors  $\psi$  est non intégrable et :

$$\int_a^x \phi(t)dt \underset{x \rightarrow b}{\sim} \int_a^x \psi(t)dt.$$

*Preuve de la proposition 1.3.9. —*

**Exercice 1.3.10.** — Déterminer des équivalents simples, lorsque  $x$  tend vers  $+\infty$ , des quantités suivantes :

1.  $\int_x^{+\infty} \frac{e^{-\frac{1}{t}}}{t^c} dt$ , pour  $c$  élément de  $]1, +\infty[$ .

2.  $\int_0^x e^{t^2} dt$ .

3.  $\int_e^x \frac{dt}{\ln t}$ .

*Solution de l'exercice 1.3.10.*

## 1.4 Propriétés des fonctions intégrables

Dans ce paragraphe, une fois encore  $I$  désigne un intervalle quelconque d'extrémités  $a$  et  $b$  avec  $a < b$ , dans  $\mathbf{R}$ , à valeurs positives.

Dans les démonstrations on privilégiera le cas où  $I = [a, b[$  avec  $a < b \leq +\infty$ .

La proposition 1.3.1. nous fournit immédiatement le résultat suivant

**Proposition 1.4.1.** — L'ensemble des applications de  $I$  dans  $\mathbf{K}$ , continues par morceaux intégrables est un sous-espace vectoriel de  $\mathcal{C}_m^0(I, \mathbf{K})$ . Notons le dans ce cours  $\mathcal{I}(I, \mathbf{K})$ . De plus l'application

$$\mathcal{I}(I, \mathbf{K}) \rightarrow \mathbf{K}; f \mapsto \int_I f$$

est linéaire.

**Proposition 1.4.2.** — INÉGALITÉ TRIANGULAIRE — Soient  $f$  et  $g$  des élément de  $\mathcal{I}(I, \mathbf{K})$ . Alors

$$\left| \int_I f + g \right| \leq \int_I |f + g| \leq \int_I |f| + \int_I |g|.$$

*Preuve de la proposition 1.4.2.* —

Examinons le cas particulier des applications continues. L'ensemble des applications de  $I$  dans  $\mathbf{K}$  continues et intégrables est évidemment un sous-espace vectoriel de  $\mathcal{I}(I, \mathbf{K})$ , en effet....

Les applications continues intégrables jouissent de propriétés particulières que nous allons examiner.

Pour une application  $f$  continue de  $I$  sur  $\mathbf{K}$  et  $c$  un point de  $I$ , on a vu dans le cours sur l'intégrale sur un segment que

$$F_c : I \rightarrow \mathbf{K}; x \mapsto \int_c^x f(t)dt$$

est une primitive de  $F$ , c'est même l'unique primitive de  $f$  nulle en  $c$ . Si  $f$  est intégrable, alors on peut généraliser ce résultat en prenant pour  $c$  une des bornes de  $I$  même si celle-ci n'est pas élément de  $I$ .

**Proposition 1.4.3.** — Soit  $f$  une application de  $I$  dans  $\mathbf{K}$  continue et intégrable. Alors les applications

$$F_a : I \rightarrow \mathbf{K}; x \mapsto \int_a^x f(t)dt$$

$$F_b : I \rightarrow \mathbf{K}; x \mapsto \int_b^x f(t)dt$$

sont des primitives de  $f$ .

$F_a(x)$  (resp.  $F_b(x)$ ) tend vers 0 lorsque  $x$  tend vers  $a$  (resp. vers  $b$ ). c'est la seule primitive de  $f$  satisfaisant cette propriété.

*Preuve de la proposition. 1.4.3.*

Soit  $F$  une primitive quelconque de  $f$  que nous fournit la continuité de  $f$ . Par définition et la proposition 1.1.4., pour tout  $x$  élément de  $I$ ,  $F(x) - F(y) = \int_y^x f(t)dt$  admet une limite lorsque  $y$  tend vers  $a$  qui est  $F_a(x)$ . Donc  $F$  admet une limite en  $a$  et pour tout  $x \in I$ ,

$$F_a(x) = F(x) - \lim_a F$$

$F_a$  qui diffère de  $F$  d'une constante est donc une primitive de  $f$ . L'égalité précédente montre que  $F_a(x)$  tend vers 0 lorsque  $x$  tend vers  $a$ .

Toute primitive de  $f$  diffère de  $F_a$  d'une constante donc si une primitive de  $f$  tend vers 0 lorsque  $x$  tend vers  $a$ , c'est  $F_a$ . □

**Proposition 1.4.4.** — Soit  $f$  une application de  $\mathcal{I}(I, \mathbf{R})$ .

1. Si  $f$  est à valeur positive, alors  $\int_I f \geq 0$ .
2. Si  $f$  est continue et à valeur positive, si  $\int_I f = 0$  alors  $f$  est nulle

*Preuve de la proposition 1.4.4.*

**Exercice 1.4.5.** — Soit  $\mathbf{E}$  l'espace vectoriel des application de  $I$  dans  $\mathbf{K}$  continues. Montrer que l'application

$$N_1 : \mathbf{E} \rightarrow \mathbf{R}_+; f \mapsto \int_I |f|$$

est une norme.

*Solution de l'exercice 1.4.5. —*

**Exercice 1.4.6.** — Soit  $\omega$  l'application

$$\mathbf{R} \rightarrow \mathbf{R}; t \mapsto \exp(-t^2)$$

Soit  $\mathbf{H}$  l'ensemble des applications  $f$  de  $\mathbf{R}$  dans  $\mathbf{R}$  continues telles que  $f^2\omega$  soit intégrable.

1. Montrer que si  $f$  et  $g$  sont dans  $\mathbf{H}$ , alors  $fg\omega$  est intégrable.
2. Montrer que  $\mathbf{H}$  est un espace vectoriel qui contient les applications polynômes.
3. Montrer que

$$\mathbf{H}^2 \rightarrow \mathbf{R}; (f, g) \mapsto \int_{\mathbf{R}} fg\omega$$

est un produit scalaire sur  $\mathbf{H}$ . On le note  $\langle \cdot | \cdot \rangle$  et  $\mathbf{N}_2$  la norme associée.

4. Montrer que tout élément  $f$  de  $\mathbf{H}$  polynomial, l'intégrale  $\int_{\mathbf{R}} f\omega$  converge et qu'il existe un réel  $c$  tel que

$$\int_{\mathbf{R}} f\omega \leq cN_2(f).$$

*Solution de l'exercice 1.4.6. —*

## 1.5 Méthode de calcul, intégration par parties, changement de variable

**Proposition 1.5.1.** — *Dans une intégration par parties, il y a plusieurs parties...on écrit donc parties avec un  $S$*



Considérons des applications  $f$  et  $g$  d'un intervalle  $I$  d'extrémités  $a$  et  $b$ , à valeurs dans  $\mathbf{K}$ , de classe  $\mathcal{C}^1$ . On peut appliquer la formule de l'intégration par parties sur tout segment  $[\alpha, \beta]$  inclus dans  $I$ , (si  $a$  ou  $b$  est élément de  $I$  on peut prendre bien sûr  $\alpha = a$  ou  $\beta = b$ ),

$$\int_{\alpha}^{\beta} f'(t)g(t)dt = [f(t)g(t)]_{\alpha}^{\beta} - \int_{\alpha}^{\beta} f(t)g'(t)dt.$$

Observons que si deux des trois termes intervenant dans cette égalité admettent une limite lorsque  $\alpha$  tend vers  $a$  et  $\beta$  vers  $b$ , alors le troisième aussi.

En particulier si le terme entre crochet admet une limite lorsque  $\alpha$  tend vers  $a$  et  $\beta$  vers  $b$  (on dira que  $[f(t)g(t)]_a^b$  converge ou  $[fg]_a^b$  converge) alors les intégrales  $\int_{\alpha}^{\beta} f'(t)g(t)dt$ ,  $\int_{\alpha}^{\beta} f(t)g'(t)dt$  sont de même nature. On a alors si ces intégrales convergent :

$$\boxed{\int_a^b f'(t)g(t)dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)dt.}$$

que l'on peut écrire plus simplement

$$\boxed{\int_a^b f'g = [fg]_a^b - \int_a^b fg'.$$

Les quantités  $[f(t)g(t)]_a^b$  ou  $[fg]_a^b$  désignent alors aussi la limite elle-même de  $[f(t)g(t)]_{\alpha}^{\beta}$ , lorsque  $\alpha$  tend vers  $a$  et  $\beta$  vers  $b$ .

Dans la pratique on a le choix entre faire l'intégration par parties avec des « bornes de sécurité »  $\alpha$  et  $\beta$  et appliquer le théorème.

**Exemple 1.5.2.** — Soit l'application

$$f : \mathbf{R}_+ \rightarrow \mathbf{R}; x \mapsto xe^{-x}.$$

Montrons que  $f$  est intégrable et calculons  $\int_0^{+\infty} f(t)dt$ .

Pour ce qui est du changement de variables dans une intégrale généralisée on peut procéder de même, c'est-à-dire se ramener à un segment en recourant à des bornes de sécurité et appliquer l'un des deux théorèmes de changement de variables de MPS1. Il existe toutefois un théorème au programme de MP de changement de variables dans une intégrale généralisée. Dans des cas concrets, son utilité reste discutable.

Nous allons rappeler les résultats de MPSI, et les appliquer dans le calculs d'intégrales généralisées. Viendra ensuite le nouveau théorème de changement de variables.

**Théorème 1.5.3.** — CHANGEMENT DE VARIABLES —

Soient  $f$  une application d'un intervalle  $J$  dans  $\mathbf{K}$ , continue. Soient  $[\alpha, \beta]$  un segment non réduit à un point et  $\varphi$  un élément de  $\mathcal{C}^1([\alpha, \beta], \mathbf{R})$ , à valeurs dans  $J$ . Alors

$$\int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t)dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(u)du.$$

On dit que l'on a fait le changement de variables  $\ll u = \varphi(t) \gg$ .

Ce théorème sert surtout à passer du premier membre au second membre dans le calcul d'une intégrale. Il s'emploie notamment lorsqu'on recourt aux règles de Bioche. Donnons un exemple

**Exemple 1.5.4.** — Existence et calcul de

$$\int_2^{+\infty} \frac{1}{(\operatorname{sh}^3(t))} dt.$$

Le théorème 1.5.1. admet le corollaire suivant.

**Corollaire 1.5.5.** — CHANGEMENT DE VARIABLES BIJECTIF —

Soit  $\varphi$  une bijection d'un intervalle  $I$ , non réduit à un point, sur un intervalle  $J$ , de classe  $\mathcal{C}^1$ . Soit  $f$  une application

de  $J$  dans  $\mathbf{K}$ , continue. Soient  $a$  et  $b$  des éléments de  $J$ . Alors

$$\int_a^b f(u)du = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t))\varphi'(t)dt.$$

Ce résultat sert à passer du premier membre au second membre. Donnons une application.

**Exemple 1.5.6.** — Existence et calcul de  $\int_0^{+\infty} e^{-\sqrt{u}}du$ .

Terminons par un théorème au programme.

**Théorème 1.5.7.** — CHANGEMENT DE VARIABLES DANS UNE INTÉGRALE GÉNÉRALISÉE—  
étant données une fonction  $f$  continue sur un intervalle  $]a, b[$ , ( $a < b$ ) et une fonction  $\varphi$  de  $]\alpha, \beta[$  dans  $]a, b[$  bijective, strictement croissante et de classe  $\mathcal{C}^1$ , les intégrales

$$\int_a^b f(t)dt \text{ et } \int_\alpha^\beta f(\varphi(u))\varphi'(u)du$$

sont de même nature et égales en cas de convergence.

*Preuve de la proposition 1.5.7.* —

**Remarque** —

1. La proposition est encore vraie pour une application,  $\varphi$  strictement décroissante (en échangeant les bornes d'une des intégrales dans l'égalité).

2. La proposition est encore vraie pour un intervalle semi ouvert  $[a, b[$  ou  $]a, b]$ .
3. Dans la pratique dans un sujet de concours on veillera à rédiger soigneusement le premier changement de variables, les autres pourront être rédigés de manière plus relâchée du reste le programme précise :  
« Les étudiants peuvent appliquer ce résultat sans justification dans des cas de changements de variable simples (fonctions affines, puissances, exponentielle, logarithme). »

## 1.6 Comparaison intégrale/série

On a vu dans le cours sur les séries que pour  $f$  application de  $[n_0, +\infty[$  dans  $\mathbf{R}$ , continue par morceaux, à valeurs positives et décroissante (avec  $n_0$  un entier naturel), la série  $\sum_{n \geq n_0} f(n)$  converge si et seulement si  $\int_{n_0}^x f(t)dt$  tend vers une limite lorsque  $x$  tend vers  $+\infty$ , autrement dit si et seulement si  $f$  est intégrable (cf. 1.1.11). Le lecteur est prié de revoir ce résultat et les applications qui furent alors traitées. Nous allons maintenant, présenter un résultat plus fort.

**Proposition 1.4.1.** — *Soient  $n_0$  un élément de  $\mathbf{N}$  et  $f$  une application de  $[n_0, +\infty[$  dans  $\mathbf{R}$ , continue par morceaux, à valeurs positives et décroissante. Pour tout entier  $n$  supérieur ou égal à  $n_0$ , on pose :*

$$w_n = \int_{n-1}^n f(t)dt - f(n).$$

*Alors la série  $\sum_{n \geq n_0+1} w_n$  converge.*

*Preuve de la proposition 1.4.1. —*

◇

On retrouve le résultat :

**Corollaire 1.4.2** — *Sous les hypothèses de la proposition 1.4.1., L'application  $f$  est intégrable si et seulement si  $\sum_{n \geq n_0} f(n)$  converge.*

*Preuve de la proposition 1.4.2. —*



## INTERVERSION LIMITES INTÉGRALES

Nous allons étudier le problème de passage à la limite « sous le signe intégrale », problème qui formellement s'écrit

$$\lim_{n \rightarrow \infty} \int_I f_n \stackrel{?}{=} \int_I \lim_{n \rightarrow \infty} f_n$$

Et le problème analogue sur les séries d'applications :

$$\sum_{n=0}^{+\infty} \int_I f_n \stackrel{?}{=} \int_I \sum_{n=0}^{+\infty} f_n$$

Rappelons l'état actuel de nos connaissances.

**Rappel —**

- Soit une suite  $(f_n)_{n \in \mathbf{N}}$  d'applications d'un segment  $[a, b]$  dans  $\mathbf{K}$  continues, qui converge uniformément vers une application  $f$ . Alors :
  - L'application  $f$  est continue;
  - La suite  $\left( \int_a^b f_n(t) dt \right)_{n \in \mathbf{N}}$  converge vers  $\int_a^b f(t) dt$ .
- Soit une suite  $(u_n)_{n \in \mathbf{N}}$  d'applications d'un segment  $[a, b]$  dans  $\mathbf{K}$  continues. On suppose que la série d'applications  $\sum u_n$  converge uniformément de somme  $S$ . Alors :
  - L'application  $S$  est continue;
  - La série  $\sum \int_a^b u_n(t) dt$  converge de somme  $\int_a^b S(t) dt$ .

Nous avons insisté sur la nécessité des deux hypothèses que sont la *convergence uniforme* et le fait que les applications soient définies sur un *segment*. Nous allons maintenant relâcher ces hypothèses, bien évidemment, il faudra leur substituer un d'autre types de conditions.

Dans le premier paragraphe nous énoncerons le théorème de convergence dominée qui repose sur une hypothèse, dite de domination, qui consiste en une majoration des applications de la suites. Le deuxième paragraphe donnera un résultat spécifique aux séries qui rappellera les théorèmes de Fubini pour les séries doubles. Le troisième paragraphe étudiera la régularité d'intégrales d'applications qui dépendent d'un paramètre, ce qui s'interprète encore comme un passage à la limite sous le signe intégrale.

## 2.1 Théorème de convergence dominée

Dans tout le paragraphe  $I$  est un intervalle quelconque, non réduit à un point. Les applications sont à valeurs dans le corps  $\mathbf{K}$  qui est indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$

Nous admettrons le théorème suivant :

**Théorème 2.1.1.** — THÉORÈME DE CONVERGENCE DOMINÉE —

Soit  $(f_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$ , continues par morceaux. On suppose que :

1. La suite  $(f_n)_{n \in \mathbf{N}}$  converge simplement vers une application  $f$  de  $I$  dans  $\mathbf{K}$ , continue par morceaux.
2. Il existe une application  $g$  de  $I$  dans  $\mathbf{R}_+$ , continue par morceaux et intégrable, telle que, pour tout entier naturel  $n$ ,  $|f_n| \leq g$ , (hypothèse de domination).

Alors :

- i)  $f$  et pour tout entier naturel  $n$ ,  $f_n$  sont intégrables.
- ii)  $\int_I f_n \xrightarrow{n \rightarrow +\infty} \int_I f$ .

**Remarque** —

1. Le programme nous limite à n'intégrer que des applications continues par morceaux. Comme aucun théorème ne vient assurer qu'une limite simple d'une suite d'applications continues par morceaux est continue par morceaux, dans la proposition 2.1.1. on est obligé de supposer  $f$  continue par morceaux. Cette limitation n'est que technique est imposée par l'absence d'une vraie théorie de l'intégration. Dans un sujet de concours on pourra pardonner l'absence de vérification d'une telle hypothèse surtout au bout d'un grand nombre d'applications du théorème de convergence dominée. Il en va tout autrement de l'hypothèse de domination centrale et incontournable. L'oublier signifie un zéro à la question.
2. On peut remplacer la suite  $(f_n)$  par une famille  $(f_\lambda)_{\lambda \in I}$ , où  $I$  est un intervalle de réels. Ce point de vu sera examiné et démontrer sur des exemples dans l'étude de la continuité et des limites dans l'intégrale à paramètre.

**Exemples** —

1. Montrer que pour tout  $\int_0^{\frac{\pi}{2}} \cos^n(t) dt \xrightarrow{n \rightarrow +\infty} 0$ .

On a déjà montré la chose en utilisant des méthodes élémentaires. Utilisons maintenant le théorème de convergence dominée.

2. Pour tout élément  $n$  de  $\mathbf{N}^*$ , on définit l'application :  $f_n : \mathbf{R}_+ \rightarrow \mathbf{R}; t \mapsto e^{(-t^n)}$ . Montrer pour tout élément  $n$  de  $\mathbf{N}^*$ , l'existence de  $\int_{\mathbf{R}_+} f_n$ . Déterminer la limite éventuelle de la suite  $\left( \int_{\mathbf{R}_+} f_n \right)_{n \in \mathbf{N}^*}$ .

## 2.2 Séries d'intégrales

Dans tout ce paragraphe  $I$  désigne toujours un intervalle quelconque, non réduit à un point. Attaquons nous au problème

$$\sum_{n=0}^{+\infty} \int_I f_n \stackrel{?}{=} \int_I \sum_{n=0}^{+\infty} f_n.$$

On pourrait en théorie appliquer sans mal le théorème de convergence dominée à la suite des sommes partielles, mais dans la pratique la vérification de la condition de domination n'est pas simple. Nous allons donner et admettre un résultat propre aux séries, d'un emploi aisé. Toutefois dans certaines circonstances l'utilisation du théorème de convergence dominée est obligée (voir exercices).

**Théorème 2.2.1.** — *Soit  $(u_n)_{n \in \mathbf{N}}$  une suite d'applications de  $I$  dans  $\mathbf{K}$ , continues par morceaux, telle que la série  $\sum u_n$  converge simplement, de somme  $S$ , supposée continue par morceaux. On suppose :*

1. *Pour tout entier naturel  $n$ ,  $u_n$  est intégrable.*
2. *La série à terme positifs  $\sum \int_I |u_n|$  converge.*

Alors

- i)  *$S$  est intégrable.*
- ii) *La série d'éléments de  $\mathbf{K}$ ,  $\sum \int_I u_n$  converge.*
- iii) 
$$\sum_{n=0}^{+\infty} \int_I u_n = \int_I \underbrace{\sum_{n=0}^{+\infty} u_n}_S$$

**Mise en garde** — Comme dans le théorème de convergence dominée, seules les contraintes du programme nous obligent à supposer la continuité par morceaux de  $S$ .

L'hypothèse 2. est par contre essentielle. En particulier la valeur absolue dont est affublées  $u_n$  ne saurait être impunément omise et remplacé par : la série  $\sum \int_I u_n$  converge. Illustrons par un exemple.

Pour tout entier naturel  $n$  on considère l'application

$$f_n : \mathbf{R}_+ \rightarrow \mathbf{R}; t \mapsto \begin{cases} \sin t, & \text{si } t \in [n2\pi, (n+1)2\pi], \\ 0 & \text{sinon.} \end{cases}$$



Ce théorème est à rapprocher du théorème de Fubini sur les « série doubles »

<ul style="list-style-type: none"> <li>• Soit <math>(u_n)_{n \in \mathbf{N}}</math> une suite d'éléments de <math>\mathcal{C}_m^0(I, \mathbf{K})</math>  <math>u_n : I \rightarrow \mathbf{K}, ; p \mapsto u_n(p)</math>.</li> <li>• Pour tout <math>n \in \mathbf{N}</math>, <math>u_n</math> est intégrable  (i.e. <math>\int_I  u_n(p)  dp</math> converge)</li> <li>• <math>\sum_{n \geq 0} u_n</math> converge simplement  (i.e. <math>\sum_{n \geq 0} u_n(p)</math> converge, pour tout <math>p \in I</math>).</li> <li>★ <math>\sum_{n \geq 0} \int_I  u_n(p)  dp</math> converge.</li> </ul> <p>Alors :</p> <ul style="list-style-type: none"> <li>– <math>\sum_{n \geq 0} \int_I u_n(p) dp</math> converge.</li> <li>– <math>\sum_{n=0}^{+\infty} u_n</math> est intégrable.</li> <li>– <math>\sum_{n=0}^{+\infty} \int_I u_n(p) dp = \int_I \sum_{n=0}^{+\infty} u_n(p) dp</math>.</li> </ul>	<ul style="list-style-type: none"> <li>• Soit <math>(u_{n,p})_{(n,p) \in \mathbf{N}^2}</math> une suite double d'éléments de <math>\mathbf{K}</math>.</li> <li>• Pour tout <math>n \in \mathbf{N}</math>, <math>\sum_{p \geq 0}  u_{n,p} </math> converge.</li> <li>• <math>\sum_{n \geq 0} u_{n,p}</math> converge, pour tout <math>p \in \mathbf{N}</math>.</li> <li>★ <math>\sum_{n \geq 0} \left( \sum_{p=0}^{+\infty}  u_{n,p}  \right)</math> converge.</li> </ul> <p>Alors :</p> <ul style="list-style-type: none"> <li>– <math>\sum_{n \geq 0} \left( \sum_{p=0}^{+\infty} u_{n,p} \right)</math> converge.</li> <li>– <math>\sum_{p \geq 0} \left( \sum_{n=0}^{+\infty} u_{n,p} \right)</math> converge absolument.</li> <li>– <math>\sum_{n=0}^{+\infty} \left( \sum_{p=0}^{+\infty} u_{n,p} \right) = \sum_{p=0}^{+\infty} \left( \sum_{n=0}^{+\infty} u_{n,p} \right)</math>.</li> </ul>
---	--

Ces deux résultats présentent quelques différences, en particulier dans la colonne de droite, l'hypothèse ....., traduction de celles sur l'intégrale, est inutile, Mais dans les deux cas, c'est l'hypothèse ★ qui joue un rôle central, hypothèse dans laquelle il convient surtout de ne pas omettre le module.

**Exercice** — Justifier l'existence de l'intégrale suivante, et donner sa valeur au moyen de la somme d'une série,

$$\int_0^1 \frac{\ln(u)}{u-1} du.$$

*Solution—*

### 2.3 Intégrales dépendant d'un paramètre

Par  $A$  nous désignerons une partie de  $\mathbf{R}^n$  ( $n$  étant un entier naturel non nul), et par  $I$  un intervalle quelconque non vide. Dans tout ce paragraphe nous considérons une application :

$$f : A \times I \rightarrow \mathbf{K}; (\vec{x}, t) \mapsto f(\vec{x}, t).$$

Pour tout élément  $\vec{x}$  de  $A$ ,  $f(\vec{x}, \cdot)$  désigne l'application

$$I \rightarrow \mathbf{K}; t \mapsto f(\vec{x}, t)$$

et, pour tout élément  $t$  de  $I$ ,  $f(\cdot, t)$  désigne l'application

$$A \rightarrow \mathbf{K}; \vec{x} \mapsto f(\vec{x}, t).$$

Par ailleurs, lorsque cela a un sens on notera

$$F : A \rightarrow \mathbf{K}; \vec{x} \mapsto \int_I f(\vec{x}, \cdot).$$

Pour être plus clair on notera plutôt,  $F(\vec{x}) = \int_I f(\vec{x}, t)dt$ . On dit que  $F$  est une intégrale dépendant du paramètre  $\vec{x}$ .

On se propose d'étudier la régularité de  $F$ . Nous commencerons par la continuité, puis par la dérivabilité. Remarquons dès maintenant que la continuité de  $F$  n'est pas un problème simple. En effet même en supposant  $f$  continue (ce qui sera en général insuffisant), La continuité de  $F$  en un point  $\vec{a}$  de  $A$  se traduit par  $\lim_{\vec{x} \rightarrow \vec{a}} F(\vec{x}) = F(\vec{a})$ . Soit

$$\lim_{\vec{x} \rightarrow \vec{a}} \int_I f(\vec{x}, t)dt = \int_I f(\vec{a}, t)dt = \int_I \lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}, t)dt$$

Il s'agit d'un problème d'interversion de passage à la limite sous le signe intégrale. On a rappelé dans le paragraphe précédent combien délicate est ce genre de question.

#### CONTINUITÉ D'UNE INTÉGRALE DÉPENDANT D'UN PARAMÈTRE

Voilà le résultat essentiel.

**Théorème 2.3.1.** —

*On suppose :*

1. Pour tout élément  $\vec{x}$  de  $A$ ,  $f(\vec{x}, \cdot)$  est continue par morceaux.
2. Pour tout élément  $t$  de  $I$ ,  $f(\cdot, t)$  est continue.
3. Il existe une application  $\varphi$  de  $I$  dans  $\mathbf{R}$  à valeurs positives, continue par morceaux intégrable, telle que, pour tout élément  $\vec{x}$  de  $A$  et tout élément  $t$  de  $I$ ,

$$|f(\vec{x}, t)| \leq \varphi(t), \text{ (hypothèse de domination).}$$

Alors l'application  $F$  est bien définie sur  $A$  et continue.

**Remarque —**

1. L'hypothèse 3 de domination assure le caractère intégrable de  $f(\vec{x}, \cdot)$  pour tout élément  $\vec{x}$  de  $A$ , donc la définition de  $F$ . Dans la pratique il est fréquent que l'intégrabilité de  $f(\vec{x}, \cdot)$  soit au préalable connu de sorte que  $F$  soit bien définie et qu'il soit alors pertinent de s'interroger sur sa continuité éventuelle.
2. L'hypothèse 1. est là pour des raisons techniques : dans le programme ne figure que l'intégration des fonctions continues par morceaux. Une théorie plus aboutie de l'intégration permet de s'en affranchir.

*Preuve du théorème 2.3.1. —*

Il est fort souvent difficile de vérifier l'hypothèse de domination, en effet il est parfois impossible de déterminer une application  $\varphi$  qui majore  $|f(\vec{x}, \cdot)|$  pour tout élément  $\vec{x}$  de  $A$ . On va donc affaiblir cette hypothèse :

Supposons 1 et 2 vérifiés. Si pour tout point  $a$  de  $A$  il existe un voisinage  $V$  de  $a$  relativement à  $A$ , une application  $\varphi_V$  de  $I$  dans  $\mathbf{R}$  à valeurs positives, continue par morceaux intégrable, telle que, pour tout élément  $\vec{x}$  de  $V$  et tout élément  $t$  de  $I$ ,

$$|f(\vec{x}, t)| \leq \varphi_V(t).$$

alors  $F$  est définie sur  $A$  et est continue, en effet :

On procède ainsi dans la pratique. Il est un cas particulier explicitement au programme :

**Généralisation 2.3.2.** — *En supposant que  $A$  est un intervalle de  $\mathbf{R}$ , on peut dans le théorème 2.3.1, remplacer l'hypothèse 3, par :*

*3'. Pour tout segment  $S$  de  $A$ , il existe une application  $\varphi_S$  de  $I$  dans  $\mathbf{R}$  à valeurs positives, continue par morceaux intégrable, telle que, pour tout élément  $x$  de  $S$  et tout élément  $t$  de  $I$ ,*

$$|f(x, t)| \leq \varphi_S(t).$$

*La conclusion reste inchangée.*

*Preuve de la généralisation 2.3.2. —*

**Exemple** — TRANSFORMÉE DE LAPLACE —

1. *Cas d'une application intégrable*

Soit  $g$  un élément de  $\mathcal{C}^0(\mathbf{R}_+, \mathbf{R})$ . On suppose de plus  $g$  intégrable. Pour tout élément  $x$  de  $\cdots$ , l'application

$$\mathbf{R}_+ \rightarrow \mathbf{R}; t \mapsto g(t)e^{-xt}$$

est intégrable. En effet,

On dispose donc de l'application

$$F : \cdots \rightarrow \mathbf{R}; x \mapsto \int_0^{+\infty} g(t)e^{-xt} dt.$$

Montrons que  $F$  est continue.

## 2. Cas d'une application bornée

Soit  $g$  un élément de  $\mathcal{C}^0(\mathbf{R}_+, \mathbf{R})$ . On suppose de plus  $g$  est bornée. Pour tout élément  $x$  de  $\cdots$ , l'application

$$\mathbf{R}_+ \rightarrow \mathbf{R}; t \mapsto g(t)e^{-xt}$$

est intégrable. En effet,

On dispose donc de l'application

$$\cdots \rightarrow \mathbf{R}; \quad x \mapsto \int_0^{+\infty} g(t)e^{-xt} dt.$$

Montrons que  $F$  est continue.

## LIMITE D'UNE INTÉGRALE DÉPENDANT D'UN PARAMÈTRE

Il n'existe pas, à proprement parler, de théorème de passage à la limite dans une intégrale dépendant d'un paramètre ; dans le programme, il est seulement fait mention d'une généralisation du théorème de convergence dominée à une famille  $(f_\lambda)_{\lambda \in I}$ , où  $I$  est un intervalle. La continuité est un cas particulier de passage à la limite ; les cas qui ne s'y réduisent pas, peuvent se traiter grâce au théorème de convergence dominée, par une démarche semblable à celle ayant servi à démontrer le théorème de continuité 2.2.1., soit que l'on utilise la généralisation évoquée par le programme, soit qu'on parte du théorème de convergence dominée lui-même, ce qui revient à démontrer la généralisation. Donnons un exemple.

**Exercice** — On reprend l'exemple 1 précédent :  $g$  est un élément de  $\mathcal{C}^0(\mathbf{R}_+, \mathbf{R})$  intégrable. Déterminer la limite éventuelle en  $+\infty$  de

$$\cdots \rightarrow \mathbf{R}; x \mapsto \int_0^{+\infty} g(t)e^{-xt} dt.$$

*Solution* —

**Remarque** — On peut parfois déterminer de manière élémentaire une limite dans une intégrale dépendant d'un paramètre. Reprenons l'exemple 2 précédent :  $g$  est un élément de  $\mathcal{C}^0(\mathbf{R}_+, \mathbf{R})$  bornée. On dispose donc de l'application

$$\cdots \rightarrow \mathbf{R}; x \mapsto \int_0^{+\infty} g(t)e^{-xt} dt.$$

Traisons le cas particulier où l'intervalle  $I$  est un segment. Le programme ne prévoit rien, mais nous allons donner une méthode générale, à reproduire pour traiter ce problème, dans le cas où  $A$  est un intervalle.

**Cas d'un segment** —

On suppose :

- $I$  est un segment  $[a, b]$ .
- L'application  $f : A \times [a, b] \rightarrow \mathbf{K}; (x, t) \mapsto f(x, t)$  est supposée *continue*.

Alors  $F : A \rightarrow \mathbf{K}; \mapsto \int_a^b f(x, t)dt$  est continue.

En effet :

Illustrons par un exemple :



**Exemple** — PRODUIT DE CONVOLUTION PÉRIODIQUE —

Soient  $f$  et  $g$  des applications continues de  $\mathbf{R}$  dans  $\mathbf{R}$  continues,  $2\pi$  périodiques. On définit l'application

$$f * g : \mathbf{R} \rightarrow \mathbf{R}; x \mapsto \int_{-\pi}^{+\pi} f(t)g(x-t)dt.$$

Montrons que  $f * g$  est continue. Soit  $h$  l'application de

$$h : \mathbf{R} \times [-\pi, +\pi] \rightarrow \mathbf{R} \quad (x, t) \mapsto f(t)g(x-t).$$

## DÉRIVATION D'UNE INTÉGRALE DÉPENDANT D'UN PARAMÈTRE

On désigne toujours par  $I$  un intervalle quelconque non vide,  $f$  une application de  $A \times I$  dans  $\mathbf{K}$  mais maintenant  $A$  est un intervalle quelconque d'intérieur non vide.

D'abord, brutalement, le théorème au programme :

**Théorème 2.3.3.** —

On suppose :

1. Pour tout élément  $x$  de  $A$ ,  $f(x, \cdot)$  est continue par morceaux intégrable
2. L'application  $\frac{\partial f}{\partial x}$  est définie sur  $A \times I^1$ .
3.  $\frac{\partial f}{\partial x}$  vérifie les hypothèses de  $f$  dans I.3.1, c'est-à-dire :
  - 3.1. Pour tout élément  $x$  de  $A$ ,  $\frac{\partial f}{\partial x}(x, \cdot)$  est continue par morceaux (intégrable<sup>2</sup>)
  - 3.2. Pour tout élément  $t$  de  $I$ ,  $\frac{\partial f}{\partial x}(\cdot, t)$  est continue.
  - 3.3. Il existe une application  $\psi$  de  $I$  dans  $\mathbf{R}$  à valeurs positives, continue par morceaux intégrable, telle que, pour tout élément  $x$  de  $A$  et tout élément  $t$  de  $I$ ,

$$\left| \frac{\partial f}{\partial x}(x, t) \right| \leq \psi(t), \text{ (hypothèse de domination)}.$$

Alors l'application  $F$  est bien définie sur  $A$ , de classe  $C^1$  et, pour tout élément  $x$  de  $A$ ,

$$F'(x) = \int_I \frac{\partial f}{\partial x}(x, t)dt.$$

1. L'ensemble  $A \times I$  n'est pas nécessairement ouvert, toutefois nous avons, dans le cours de calcul différentiel, généralisé la notion de dérivées partielles, notamment au cas du produit de deux intervalles.)

2. L'intégrabilité est facultative puisqu'assurée par 3.3.

Nous laissons en exercice la preuve de ce résultat qui fait appel, comme celle de 2.3.1., au théorème de convergence dominée.

Comme pour le théorème de continuité, il existe une généralisation au programme.

**Généralisation 2.3.4.** — *On peut dans 2.3.3., remplacer 3.3 par :*

–3.3' *Pour tout segment  $J$  de  $A$  il existe une application  $\psi_J$  de  $I$  dans  $\mathbf{R}$  à valeurs positives, continue par morceaux intégrable, telle que, pour tout élément  $x$  de  $J$  et tout élément  $t$  de  $I$ ,*

$$\left| \frac{\partial f}{\partial x}(x, t) \right| \leq \psi_J(t), .$$

**Exemple** —  $g$  est un élément de  $\mathcal{C}^0(\mathbf{R}_+, \mathbf{R})$  bornée. On dispose, on l'a vu, de l'application

$$\mathbf{R}_+^* \rightarrow \mathbf{R}; \quad x \mapsto \int_0^{+\infty} g(t)e^{-xt} dt.$$

Montrons que  $F$  est de classe  $\mathcal{C}^1$  et donnons l'expression de sa dérivée.

**Théorème 2.3.5** —

*Soit  $k$  un entier naturel supérieur ou égal à 1. On suppose :*

1.  $\frac{\partial^k f}{\partial x^k}$  est définie sur  $A \times I$ . (ce qui exige l'existence de  $\frac{\partial^j f}{\partial x^j}$ , pour  $j = 0, 1, 2, \dots, k$ ).
2. Pour  $j = 0, 1, \dots, k$

*Pour tout élément  $x$  de  $A$ ,  $\frac{\partial^j f}{\partial x^j}(x, \cdot)$  est continue par morceaux intégrable. (ce qui assure au passage l'existence de  $F : A \rightarrow \mathbf{K}, x \mapsto \int_I f(x, t) dt$ ).*

3. Pour tout élément  $t$  de  $I$ ,  $\frac{\partial^k f}{\partial x^k}(\cdot, t)$  est continue.
4. Il existe une application  $\psi$  de  $I$  dans  $\mathbf{R}$  à valeurs positives, continue par morceaux intégrable, telle que, pour tout élément  $x$  de  $A$  et tout élément  $t$  de  $I$ ,

$$\left| \frac{\partial^k f}{\partial x^k}(x, t) \right| \leq \psi(t), \text{ (hypothèse de domination)}.$$

Alors  $F$  est définie sur  $A$ , de classe  $\mathcal{C}^k$ , pour tout élément  $j$  de  $\{0, \dots, k\}$ , pour tout élément  $x$  de  $A$ ,

$$F^{(j)}(x) = \int_I \frac{\partial^j f}{\partial x^j}(x, t) dt.$$

### Cas d'un segment —

Supposons que  $I$  soit un segment et que l'on ait :

- $f$  est continue;
- $\frac{\partial f}{\partial x}$  est définie sur  $A \times I$  et continue.

C'est notamment le cas lorsque  $f$  est de classe  $\mathcal{C}^1$ .

Alors l'application  $F$  de classe  $\mathcal{C}^1$  et, pour tout élément  $x$  de  $A$ ,  $F'(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt$ .

En effet :

**Exercice** — Soit  $x$  un élément de  $[1, +\infty[$ . existence et calcul de

$$I(x) = \int_1^2 \ln(t)t^x dt.$$

*Solution* —

**Exercice** — Soient  $k$  un élément de  $\mathbf{N}^*$ ,  $f$  une application de  $\mathbf{R}$  dans  $\mathbf{R}$  de classe  $\mathcal{C}^k$  et  $a$  un réel. On pose

$$g : \mathbf{R} \rightarrow \mathbf{R}; \begin{cases} x \mapsto \frac{f(x)-f(a)}{x-a}, & \text{pour } x \neq a, \\ f'(a), & \text{pour } x = a. \end{cases}$$

Montrons que l'application  $g$  est de classe  $\mathcal{C}^{k-1}$ .

*Solution* —

#### UN EXEMPLE CLASSIQUE : LA FONCTION GAMMA

La fonction  $\Gamma$  introduite par Leonhard Euler joue un rôle capital dans les mathématiques et les sciences où on la retrouve dans de nombreux domaines. Cette application est un prolongement « naturel » de la fonction factorielle.

Soit  $\Gamma$  la fonction de la variable réel  $x$  définie par :

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

#### ETUDE DU DOMAINE DE DÉFINITION

Étudions les problèmes de définitions, dus à la présence de la borne supérieure,  $+\infty$  et au fait que pour  $x < 1$ , l'intégrande n'est pas définie en 0. Nous ne réservons aucun traitement de faveur au cas où  $x < 1$ , nous considérerons donc, pour tout réel  $x$  l'application :

$$f_x : \mathbf{R}_+^* \rightarrow \mathbf{R}; t \mapsto t^{x-1} e^{-t}.$$

Pour tout réel  $x$ ,  $f_x$  est continue (par morceaux).

- *Intégrabilité au voisinage de  $+\infty$*

- *Intégrabilité au voisinage de 0*

Le domaine de définition de  $\Gamma$  est donc ...

On dispose donc de l'application

$$\Gamma : \cdots \rightarrow \mathbf{R}; x \mapsto \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

RÉGULARITÉ DE  $\Gamma$

Montrons que l'application  $\Gamma$  est de classe  $\mathcal{C}^\infty$  et pour tout élément  $k$  de  $\mathbf{N}$ ,

$$\Gamma^{(k)} : \mathbf{R}_+^* \rightarrow \mathbf{R}; x \mapsto \int_0^{+\infty} (\ln(t))^k t^{x-1} e^{-t} dt.$$

## EQUATION FONCTIONNELLE

Pour tout élément  $x$  de  $\mathbf{R}_+^*$ ,

$$\boxed{\Gamma(x+1) = x\Gamma(x).}$$

Remarquons que  $\Gamma(1) = \int_0^{+\infty} e^{-t} = 1$ . Par récurrence, on en déduit grâce à la proposition 2.4.2., que, pour tout élément  $n$  de  $\mathbf{N}^*$ ,

$$\Gamma(n) = \dots \quad .$$

## Chapitre XI

# RÉDUCTION DES ENDOMORPHISMES, LE POINT DE VUE ALGÈBRIQUE

Dans ce court chapitre,  $n$  désigne un entier naturel non nul,  $\mathbf{E}$  désigne un espace vectoriel sur  $\mathbf{K}$  de dimension finie non nulle  $n$ , (avec  $\mathbf{K}$  un corps, qui dans la pratique sera  $\mathbf{R}$  ou  $\mathbf{C}$ ),  $u$  un endomorphisme de  $\mathbf{E}$  et  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ .

## RAPPELS

Nous avons vu dans le premier chapitre de l'année le résultat suivant.

**Proposition 1.1.1.** — *On a l'équivalence des propriétés suivantes :*

1. *L'endomorphisme  $u$  (resp. la matrice  $M$ ) est diagonalisable.*
2. *la somme des espaces propres de  $u$  (resp.  $M$ ) est égale à  $\mathbf{E}$  (resp.  $\mathcal{M}_{n,1}(\mathbf{K})$ ).*
3. *Le polynôme caractéristique de  $u$  (resp.  $M$ ) est scindé et la multiplicité de chaque valeur propre est égale à la dimension de l'espace propre correspondant.*

Rappelons que la multiplicité d'une valeur propre est . . . .rieur à la dimension du sous-espace propre associé.

Nous étudierons les polynômes d'endomorphismes qui fournissent un critère puissant du caractère diagonalisable. Ce résultat est essentiel dans les exercices d'oraux ardues dont il vient à bout, là où les méthodes développées dans le premier chapitre sur la réduction sont en général impuissantes.

## POLYNÔMES D'ENDOMORPHISMES

### 2.1 Définitions, Premières propriétés

Nous allons « remplacer » dans un polynôme l'indéterminée par l'endomorphisme  $u$ .

Rappelons pour commencer que dans l'algèbre  $(\mathcal{L}(\mathbf{E}), +, \circ, \cdot)$ , comme dans toute algèbre, pour tout élément  $i$  de  $\mathbf{N}$ ,  $u^i$  désigne :

- L'unité de l'algèbre,  $\text{id}_{\mathbf{E}}$ , pour  $i = 0$  ;
- $\underbrace{u \circ \cdots \circ u}_{i \text{ termes}}$  pour  $i$  non nul.

**Définition 2.1.1.** — *Soit  $P$  un élément de  $\mathbf{K}[X]$ , décomposé dans la base canoique de  $\mathbf{K}[X]$  en :*

$$P = \sum_{i=0}^d a_i X^i,$$

$d$  désigne ici le degré de  $P$ .

- On note  $P(u)$  l'endomorphisme  $\sum_{i=0}^d a_i u^i$  ;



— On note  $P(M)$  l'élément de  $\mathcal{M}_n(\mathbf{K})$   $\sum_{i=0}^d a_i M^i$ .

Lorsque  $P$  est nulle  $P(u)$  est l'endomorphisme nul,  $P(M)$  la matrice  $O_n$ , d'après la convention usuelle sur une somme vide.

On dit que l'on a substitué  $u$  (resp.  $M$ ) dans  $P$  à l'indéterminé  $X$  pour obtenir  $P(u)$  (resp.  $P(M)$ ). On dit que  $P(u)$  est un polynôme d'endomorphisme et que  $P(M)$  est un polynôme de matrice.

**Remarque** —  $P(u)$  et  $P(M)$  sont bien définis car  $\mathcal{L}(\mathbf{E})$  et  $\mathcal{M}_n(\mathbf{K})$  sont des  $\mathbf{K}$ -algèbres, donc des structures dans lesquelles il est possible de multiplier entre eux les éléments, de les multiplier par un élément de  $\mathbf{K}$  et de les ajouter. Plus généralement on pourrait définir la substitution, dans un élément de  $\mathbf{K}[X]$ , d'un élément d'une  $\mathbf{K}$ -algèbre.

**Exemples** —

$P$	0	1	$X^3 + 2X + 8$	$X^d - 1$
$P(u)$				
$P(M)$				

L'application qui à un élément de  $\mathcal{L}(\mathbf{E})$  associe sa matrice dans une base étant un morphisme de l'algèbre  $(\mathcal{L}(\mathbf{E}), +, \circ, \cdot)$  dans l'algèbre  $(\mathcal{M}_n(\mathbf{K}), +, \times, \cdot)$ , on a immédiatement la proposition suivante :

**Proposition 2.1.2.** — Soit  $P$  un élément de  $\mathbf{K}[X]$ . Supposons que  $M$  soit la matrice de  $u$  dans une base  $\mathcal{B}$  de  $\mathbf{E}$ . Alors  $P(M)$  est la matrice de  $P(u)$  dans la base  $\mathcal{B}$ .

Cette proposition nous autorise à raisonner indifféremment sur des polynômes d'endomorphismes ou des polynômes de matrices. Nous pourrions donc nous contenter d'énoncer les résultats pour l'endomorphisme  $u$  ou la matrice  $M$ , à charge pour le lecteur de traduire dans le cas non traité.

**Proposition 2.1.3.** — L'application

$$\Phi : \mathbf{K}[X] \rightarrow \mathcal{L}(\mathbf{E}) \text{ ( resp. } \mathcal{M}_n(\mathbf{K}) \text{ )}, ; P \mapsto P(u) \text{ ( resp. } P(M) \text{ )}$$

est un morphisme de l'algèbre  $(\mathbf{K}[X], +, \times, \cdot)$  dans l'algèbre  $(\mathcal{L}(\mathbf{E}), +, \circ, \cdot)$  (resp.  $(\mathcal{M}_n(\mathbf{K}), +, \times, \cdot)$ ). Ce qui concrètement signifie que, pour tout  $(P, Q)$  élément de  $\mathbf{K}[X]^2$  et tout  $(\lambda, \mu)$  élément de  $\mathbf{K}^2$ ,

- $(\lambda \cdot P + \mu \cdot Q)(u) = \lambda \cdot P(u) + \mu \cdot Q(u)$ ,  $(\lambda \cdot P + \mu \cdot Q)(M) = \lambda \cdot P(M) + \mu \cdot Q(M)$ ;
- $(P \times Q)(u) = P(u) \circ Q(u)$ ,  $(P \times Q)(M) = P(M) \times Q(M)$  ;
- Dans le polynôme  $X^0$ , la substitution à l'indéterminé  $X$  de  $u$  donne  $\text{id}_{\mathbf{E}}$ , celle de  $M$ ,  $I_n$ .

Cette proposition permet pratiquement de remplacer dans un polynôme  $P$ , qui peut s'exprimer au moyen de sommes et de produits,  $X$  par  $M$  pour avoir  $P(M)$ .

**Exemple** — Soit  $P$  le polynôme  $X^2(X - 3) + (X - 3)^4(X^3 - 2X)$ . alors :

$$P(M) = \dots$$

$$P(u) = \dots$$

et pour  $\vec{x}$  élément de  $\mathbf{R}$ ,

$$P(u)(\vec{x}) = \dots$$

*Preuve de la proposition 2.1.3.* — La preuve de ce résultat est facile et sans malice. Nous ne montrerons que le deuxième

point, légèrement plus délicat.

Gardons les notations de la proposition 2.1.3. L'image de l'algèbre  $\mathbf{K}[X]$  par le morphisme d'algèbre  $\Phi$  est donc une sous-algèbre de  $\mathcal{L}(u)$ . On la note naturellement  $\mathbf{K}[u]$ . Cette algèbre hérite de la commutativité de  $\mathbf{K}[X]$ . En effet soient  $v_1$  et  $v_2$  des éléments de  $\mathbf{K}[X]$ . Il existe  $P_1$  et  $P_2$  éléments de  $\mathbf{K}[X]$  tels que  $v_1 = P_1(u)$ ,  $v_2 = P_2(u)$ .

$$v_1 \circ v_2 = P_1(u) \circ P_2(u) = (P_1 \times P_2)(u) = (P_2 \times P_1)(u) = P_2(u) \circ P_1(u) = v_2 \circ v_1.$$

Voilà prouvée la proposition suivante :

**Proposition 2.1.4.** — L'ensemble  $\mathbf{K}[u]$  (resp  $\mathbf{K}[M]$ ) est une sous-algèbre de  $\mathcal{L}(\mathbf{E})$ , (resp.  $\mathcal{M}_n(\mathbf{K})$ ) commutative.

**Exercice 2.1.5.** — Prouver que si  $M$  est diagonalisable, alors

$$\chi_M(M) = 0_n.$$

On verra dans la partie 3 que ce résultat demeure même pour une matrice non diagonalisable. Ce résultat porte le nom de *Théorème de Cayley-Hamilton*.

*Solution de l'exercice 2.1.5.* — Il existe une matrice diagonale  $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ , noté  $\Delta$ , un élément  $P$  de  $\text{GL}_n(\mathbf{K})$ , tels que :  $M = P\Delta P^{-1}$ .

Donnons pour finir un résultat sur les valeurs propres d'un polynôme d'endomorphisme.

**Proposition 2.1.6.** — Soient  $P$  un élément de  $\mathbf{K}[X]$  et  $\lambda$  une valeur propre de l'endomorphisme  $u$ . Alors  $P(\lambda)$  est une valeur propre de  $P(u)$ .

*Preuve de la proposition 2.1.6.* — Nous allons donner deux méthodes.

— PREMIÈRE MÉTHODE

— SECONDE MÉTHODE

**Exercice 2.1.7.** — Soit  $M$  un élément de  $\mathcal{M}_{13}(\mathbf{R})$  non nul diagonalisable dans  $\mathcal{M}_{13}(\mathbf{C})^1$ , tel que :  $M^3 = -M$ . Déterminer le spectre complexe de  $M$ , c'est-à-dire le spectre de  $M$ , considérée comme un élément de  $\mathcal{M}_n(\mathbf{C})$ .

*Solution de l'exercice 2.1.7.* —

## 2.2 Polynômes annulateurs, polynôme minimal

Nous allons étudier les polynômes  $P$  tels que  $P(u)$  soit nul, on parle de *polynômes annulateurs* de  $u$ .

**Proposition 2.2.1.** — *L'ensemble  $\mathfrak{I}$  des éléments  $P$  de  $\mathbf{K}[X]$  tels que  $P(u) = 0_{\mathcal{L}(\mathbf{E})}$  (ou  $P(M) = 0_n$ ) est un idéal non nul de  $\mathbf{K}[X]$ .*

*Preuve de la proposition 2.2.1.* —

---

1. On montrera dans le paragraphe suivant que l'hypothèse  $M$  est diagonalisable est en fait inutile.

L'ensemble  $\mathfrak{I}$  est appelé *idéal annulateur* de  $u$  (ou de  $M$ ). Comme, on vient de le voir, cet idéal est non nul, il admet donc un unique générateur unitaire.

**Définition 2.2.2.** — *Le générateur unitaire de l'idéal annulateur de  $u$  (resp.  $M$ ) s'appelle polynôme minimal de  $u$  (resp de  $M$ ). On le notera  $\mu_u$  (resp.  $\mu_M$ ).*

**Exemples** —

- Le polynôme minimal de  $0_n$  est... , celui de  $2021I_n$  est... .
- Soit  $T$  l'endomorphisme de  $\mathcal{M}_n(\mathbf{K})$  qui à une matrice  $A$  associe sa transposée. Son polynôme minimal est ... , en effet
- Soit  $v$  un endomorphisme de  $\mathbf{E}$  nilpotent d'ordre  $p$ . Son polynôme minimal est ... , en effet

**Proposition 2.2.3.** — *Soit  $\mathbf{F}$  un sous espace vectoriel de  $\mathbf{E}$  stable par  $u$ , alors le polynôme minimal de l'endomorphisme  $v$  induit par  $u$  sur  $\mathbf{F}$  divise le polynôme minimal de  $u$  :*

$$\mu_v \mid \mu_u.$$

*Preuve de la proposition 2.2.3.* —

**Exercice 2.2.4.** — On a vu en exercice que pour des matrices  $A$  et  $B$  de  $\mathcal{M}_n(\mathbf{K})$ ,  $\chi_{AB} = \chi_{BA}$ . A-t-on  $\mu_{AB} = \mu_{BA}$  ?

*Solution de l'exercice 2.2.4.* —

**Proposition 2.2.5.** — *Soit  $P$  un polynôme annulateur de  $u$  (resp.  $M$ ). Alors les valeurs propres de  $u$  (resp.  $M$ ) sont racines de  $P$ .*

*Preuve de la proposition 2.2.5.* —

On a le corollaire immédiat

**Corolaire 2.2.6.** — *Les valeurs propres de  $u$  (resp.  $M$ ) sont racines du polynôme minimal.*

Réciproquement

**Proposition 2.2.7.** — *Toute racine du polynôme minimal de  $u$  (resp.  $M$ ) sont des valeurs propres*

*Preuve du théorème 2.2.7.* —

Bien sûr le théorème de Cayley-Hamilton, dont nous avons déjà parlé, prouve ce résultat. Mais nous allons en fournir une preuve élémentaire

**Conclusion 2.2.8.** — Les racines du polynôme minimal sont exactement les valeurs propres.

**Exercice 2.2.9.** — *Déterminer les polynômes minimaux des matrices suivantes :*

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

*Solution de l'exercice 2.2.9.*

Donnons une application des polynômes annulateurs au calcul des puissances d'une matrice.

**Application 2.2.10** — CALCUL DES PUISSANCES D'UNE MATRICE—

Pour calculer les puissances d'une matrice, on a dans le passé calculé les puissances de matrices, en les diagonalisant. C'est une méthode efficace, toutefois il peut arriver que la diagonalisation soit difficile (ordre élevé, présence de paramètre), ou encore, tout simplement, que la matrice ne soit pas diagonalisable. Nous allons maintenant développer une méthode plus élaborée pour calculer les puissances d'une matrice, diagonalisable ou non, qui utilise un polynôme annulateur, le polynôme minimal le plus souvent.

*Calculer pour tout élément  $m$  de  $\mathbf{N}$ , la puissance  $m^e$  des matrices suivantes :*

a)  $\begin{pmatrix} \cos(\theta) & -2\sin(\theta) \\ \frac{\sin(\theta)}{2} & \cos(\theta) \end{pmatrix}$ , où  $\theta$  est un réel non congru à 0 modulo  $\pi$ .

b)  $\begin{pmatrix} -3 & -3 & 2 \\ 1 & 1 & -2 \\ 2 & 4 & -4 \end{pmatrix}.$

*Solution de l'exercice 2.2.10. —*

a) Le polynôme caractéristique de  $M$  est :  $\chi_M = X^2 - 2\cos\theta X + 1$ . Donc,

$$\operatorname{sp}(M) = \{ \quad , \quad \}$$

b) ...

De façon plus profonde on a le résultat suivant :

**Proposition 2.2.11.** — Soit  $d$  le degré du polynôme minimal de  $u$ . Alors  $(u^0, u^1, \dots, u^{d-1})$  est une base de  $\mathbf{K}[u]$ .

*Preuve de la proposition 2.2.11.* —

On a déjà rencontré en exercice et dans un cas particulier le théorème de Cayley-Hamilton<sup>1</sup> :

**Proposition 2.2.12.** — THÉORÈME DE CAYLEY-HAMILTON—

$\chi_u(u) = 0_{\mathcal{L}(\mathbf{E})}$ ,  $\chi_M(M) = 0_n$ .

Autrement dit le polynôme minimal de  $u$  (resp.  $M$ ) divise le polynôme caractéristique de  $u$  (resp.  $M$ ).

La preuve de ce résultat est non exigible. Nous en avons vu une preuve en exercice dans le cas où  $M$  est diagonalisable (cf. 2.1.5.). Nous en donnerons une preuve riche en travaux dirigés. Voici une preuve élémentaire utilisant la trigonalisation dans le cas où  $\mathbf{K} = \mathbf{R}$  ou  $\mathbf{C}$ .

*Preuve de la proposition 2.2.12.* —

Soit  $M$  un élément de  $\mathcal{M}_n(\mathbf{K})$ . Montrons que  $\chi_M(M) = 0_n$ . la matrice  $M$  est *a fortiori* élément de  $\mathcal{M}_n(\mathbf{C})$  et est trigonalisable dans ce dernier espace. Il existe  $Q$  élément  $\mathrm{GL}_n(\mathbf{C})$ ,  $T$  matrice triangulaire supérieure, élément de  $\mathcal{M}_n(\mathbf{C})$ , tels que  $M = QTQ^{-1}$ . La matrice  $T$  est de la forme

$$T = \begin{pmatrix} \lambda_0 & \times & \dots & \dots & \times \\ 0 & \lambda_1 & \times & \dots & \times \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & \times \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

(les complexes  $\lambda_1, \lambda_2, \dots, \lambda_n$  sont les valeurs propres complexes de  $M$ ).

**Application aux matrices nilpotentes** — Soit  $v$  un endomorphisme nilpotent de  $\mathbf{E}$  d'ordre  $p$ . Nous avons vu dans

---

1. **William HAMILTON (1805 – 1865)**— Mathématicien irlandais polyglote et grand amateur de whiskey, il propose la première exposition rigoureuse de  $\mathbf{C}$ , vu comme  $\mathbf{R}^2$  muni d'une addition et d'un produit et définit une structure analogue au corps des nombres complexes de dimension 4 : les *quaternions*. Son apport à la mécanique est de tout premier plan.

**Arthur CAYLEY (1821 – 1895)** — Mathématicien anglais, très productif dans tous les domaines des mathématiques. Il donnera le premier une preuve de ce théorème.

le premier chapitre que :  $p \leq n, \dots$

## POLYNÔMES ANULATEURS ET CARACTÉRISATION DE LA DIAGONALISATION

### 3.1 Lemme des noyaux

Nous allons établir deux critères voisins très puissants pour déterminer si un endomorphisme est diagonalisable. Ces critères permettront de résoudre des exercices théoriques ou complexes. Pour commencer, afin d'établir ces résultats, nous allons exposer un lemme au programme, d'une grande importance.

**Proposition 3.1.1** — LEMME DES NOYAUX —

Soient  $P_1$  et  $P_2$  des éléments de  $\mathbf{K}[X]$ , premiers entre eux. Alors, en posant  $P := P_1 P_2$ ,

- i)  $\text{Ker}(P_1(u))$  et  $\text{Ker}(P_2(u))$  sont des sous-espaces vectoriels de  $E$ , stable par  $u$ .
- ii)  $\text{Ker}(P(u)) = \text{Ker}(P_1(u)) \oplus \text{Ker}(P_2(u))$ .

*Preuve de la proposition 3.1.1.* — La preuve de i) est évidente, en effet, d'après 2.1.4.,  $u$  et  $P_i(u)$  commutent, pour  $u = 1, 2$ . Donc d'après ch. 1, 1.1.2,  $\text{Ker}(P_1(u))$  et  $\text{Ker}(P_2(u))$  sont stables par  $u$ .

Passons à ii). Nous allons successivement montrer que la somme de  $\text{Ker}(P_1(u))$  et  $\text{Ker}(P_2(u))$  est directe, puis que cette somme est incluse dans  $\text{Ker}(P(u))$  et enfin que  $\text{Ker}(P(u))$  est incluse dans la somme de  $\text{Ker}(P_1(u))$  et  $\text{Ker}(P_2(u))$ .

- Montrons que la somme de  $\text{Ker}(P_1(u))$  et  $\text{Ker}(P_2(u))$  est directe.

Donc  $\text{Ker}(P_1(u)) + \text{Ker}(P_2(u)) = \text{Ker}(P_1(u)) \oplus \text{Ker}(P_2(u))$ .

- Montrons que  $\text{Ker}(P_1(u)) + \text{Ker}(P_2(u)) \subset \text{Ker}(P(u))$ .



- Montrons que  $\text{Ker}(P(u)) \subset \text{Ker}(P_1(u)) + \text{Ker}(P_2(u))$ .

D'après ces trois points, on a bien  $\text{Ker}(P(u)) = \text{Ker}(P_1(u)) \oplus \text{Ker}(P_2(u))$ . □

Ce théorème se généralise à un nombre de polynômes supérieur à 2.

**Proposition 3.1.2.** — LEMME DES NOYAUX —

Soient  $m$  un entier naturel supérieur ou égal à 2 et  $P_1, P_2, \dots, P_m$  des éléments de  $\mathbf{K}[X]$ , premiers entre eux deux à deux. On pose  $P = P_1 P_2 \dots P_{m-1} P_m$ . Alors :

i)  $\text{Ker}(P_i(u))$  est un sous-espace vectoriel de  $E$ , stable par  $u$ .

ii)  $\text{Ker}(P(u)) = \bigoplus_{i=1}^m \text{Ker}(P_i(u))$ .

*Preuve de la proposition 3.1.2.* — La preuve de i) est identique à celle donnée pour 3.1.1. Passons à ii)

Une première application du lemme des noyaux est la décomposition de  $\mathbf{E}$  en sous-espaces stables qui suit.

**Proposition 3.1.3** — Supposons que  $u$  possède un polynôme annulateur scindé  $P$ ,

$$P = c \prod_{i=1}^k (X - \alpha_i)^{q_i},$$

où  $\alpha_1, \dots, \alpha_k$  sont des éléments de  $\mathbf{K}$ ,  $c$  de  $\mathbf{K}^*$  et  $q_1, \dots, q_k$  des entiers supérieurs ou égaux à 1.

Alors  $\mathbf{E}$  est la somme directe des sous-espaces  $\text{Ker}((\alpha_i \text{id}_{\mathbf{E}} - u)_i^q)$ ,  $i = 1, \dots, k$ , ces sous-espaces sont stables par  $u$  et  $u$  induit sur chacun des  $\text{Ker}((\alpha_i \text{id}_{\mathbf{E}} - u)_i^q)$  la somme d'une homothétie (de rapport  $\alpha_i$ ) et d'un endomorphisme nilpotent.

*Preuve de la proposition 3.1.3. —*

**Corollaire 3.1.4.** — *supposons que  $M$  admette un polynôme anulateur scindé*

$$P = c \prod_{i=1}^k (X - \alpha_i)^{q_i},$$

*où  $\alpha_1, \dots, \alpha_k$  sont des éléments de  $\mathbf{K}$ ,  $c$  de  $\mathbf{K}^*$  et  $q_1, \dots, q_k$  des entiers supérieurs ou égaux à 1. Alors  $M$  est semblable à une matrice digonale par blocs  $\text{diag}(M_1, \dots, M_k)$  où pour  $i = 1, \dots, k$   $M_i$  est la somme d'une homothétie (de rapport  $\alpha_i$ ),  $\alpha_i I_{d_i}$  et d'une matrice nilpotente  $N_i$*

$$M_i = \alpha_i I_{d_i} + N_i.$$

*Preuve du corollaire 3.1.4. —*

**Remarque** — La proposition 3.1.3. s'applique en particulier lorsque le polynôme caractéristique est scindé....

Une conséquence de cette remarque est la décomposition de Dundford, qui sera vue en exercice.

## 3.2 Critères de diagonalisation

Nous allons présenter un critère du caractère diagonalisable d'une matrice. Résultat très puissant et théorique il est l'outil des situations abstraites et désespérées

**Théorème 3.2.1.** — *L'endomorphisme  $u$  (resp. La matrice  $M$ ) est diagonalisable si et seulement si il existe un polynôme  $P$ , élément de  $\mathbf{K}[X]$ , scindé à racines simples tel que :*

$$P(u) = 0_{\mathcal{L}_{\mathbf{E}}} \text{ resp. } (P(M) = 0_n).$$

Un polynôme  $P$  tel que  $P(u) = 0_{\mathcal{L}(\mathbf{E})}$  (resp.  $P(M) = 0_n$ ) est dit polynôme annulateur pour  $u$  (resp.  $M$ ).

*Preuve du théorème 3.2.1.* — Pour commencer remarquons que l'on peut se contenter de prouver le théorème pour la matrice  $M$ . En effet le résultat étant acquis pour une matrice  $M$ , on l'obtient immédiatement pour  $u$  en prenant pour  $M$  la matrice de  $u$  dans une base, en effet on a vu que  $u$  est diagonalisable si et seulement si  $M$  l'est (cf. ch. 1, I.4.3.) et d'après 2.1.2,  $P(u) = 0_{\mathcal{L}(\mathbf{E})}$  si et seulement si  $P(M) = 0_n$ .

- HYPOTHÈSE :  $M$  est diagonalisable.

Soient  $\lambda_1, \lambda_2, \dots, \lambda_k$  les valeurs propres de  $M$ , deux à deux distinctes, et  $m_k$  l'ordre de multiplicité de la valeur propre  $\lambda_k$ ,  $k = 1, 2, \dots, k$ . La matrice  $M$  étant diagonalisable, il existe  $Q$  élément de  $\mathrm{GL}_n(\mathbf{K})$  tel que  $M = Q\Delta Q^{-1}$ , avec

$$\Delta = \mathrm{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{m_1}, \underbrace{\lambda_2, \dots, \lambda_2}_{m_2}, \dots, \underbrace{\lambda_k, \dots, \lambda_k}_{m_k})$$

- HYPOTHÈSE : il existe un élément  $P$  de  $\mathbf{K}[X]$  scindé à racines simples tel que  $P(M) = 0_n$ .

Le polynôme  $P$  s'écrit  $P = c \prod_{i=1}^m (X - \alpha_i)$ , avec  $\alpha_1, \alpha_2, \dots, \alpha_m$ , éléments de  $\mathbf{K}$  deux à deux distincts et  $c$  un élément de  $\mathbf{K}^*$ .

Remarquons que nous avons en fait montré que si  $m$  est diagonalisable, alors  $\prod_{i=1}^k (X - \lambda_i)$ , noté parfois  $\prod_{\lambda \in \text{sp}(M)} (X - \lambda)$  est un polynôme annulateur pour  $M$ . La réciproque est vraie puisque ce polynôme est scindé à racines simples.

Observons qu'en fait si  $M$  est diagonalisable  $\prod_{\lambda \in \text{sp}(M)} (X - \lambda)$  est .....

En effet

D'où le résultat suivant :

**Proposition 3.2.2.** — *L'endomorphisme  $u$  (resp. la matrice  $M$ ) est diagonalisable si et seulement si*

**Application élémentaire** — On reprend l'exercice 2.1.7. Soit  $M$  un élément de  $\mathcal{M}_n(R)$  tel que :  $M^3 = -M$ . Montrer que  $M$  est diagonalisable dans  $\mathcal{M}_n(\mathbf{C})$ .

Donnons maintenant un exercice dont la résolution est bien difficile sans le théorème 3.2.1.

**Exercice 3.2.3.** — *Soit  $u$  un élément de  $L(\mathbf{E})$ . On considère l'application  $\Psi : \mathcal{L}(\mathbf{E}) \rightarrow \mathcal{L}(\mathbf{E}) ; v \mapsto u \circ v$ . L'application  $\Psi$  est linéaire, donc élément de  $\mathcal{L}(\mathcal{L}(\mathbf{E}))$ . On suppose que  $u$  est diagonalisable. Montrer  $\Psi$  est diagonalisable.*

*Solution de l'exercice 3.2.3. —*

Un autre résultat, celui là au programme, se déduit également du théorème 3.2.1 :

**Proposition 3.2.4.** — *Soit  $\mathbf{F}$  un sous espace vectoriel de  $\mathbf{E}$  stable par  $u$ . Si  $u$  est diagonalisable, alors l'endomorphisme  $v$  induit par  $u$  sur  $\mathbf{F}$  est diagonalisable.*

Ce résultat n'est pas simple à prouver, en utilisant la seule définition de la diagonalisation. En effet il faudrait déterminer une base de vecteurs propres dont une sous-famille soit une base de  $\mathbf{F}$ . Sans le recours, d'une façon ou d'une autre, au théorème 3.1.3, la chose semble fort mal aisée.

*Preuve du théorème 3.2.4. —*

Pour terminer études dans le cas où le polynôme caractéristique est scindé la forme du polynôme caractéristique et minimal.

**Exercice** — Notons  $\lambda_1, \dots, \lambda_k$  les  $k$  valeurs propres deux à deux distinctes de  $u$ , et  $m_1, \dots, m_k$  leurs multiplicités respectives.

Par définition

$$\chi_u = \prod (X - \lambda_i)^{m_i}.$$

D'après 2.2.8. le polynôme minimal est donc scindé de la forme

$$\mu_u = \prod (X - \lambda_i)^{p_i}.$$

où pour  $i = 1, \dots, k$ ,  $0 < p_i \leq m_i$ . D'après 3.1.3 on a également que

$$\mathbf{E} = \bigoplus_{i=1}^k \text{Ker}((\lambda_i I_n - u)^{m_i})$$

et que pour  $i = 1, \dots, k$ , le sous espace  $\text{Ker}(\lambda_i I_n - u)^{m_i}$ , noté  $\mathbf{F}_i$  est stable par  $u$ , qui y induit un endomorphisme  $u_i$ , somme de  $\lambda_i \text{id}_{\mathbf{F}_i}$  et de  $v_i$  endomorphisme nilpotent.

1. Montrer que pour  $i = 1, \dots, k$ ,  $m_i = \dim(\mathbf{F}_i)$ .
2. Montrer pour  $i = 1, \dots, k$ ,  $p_i$  est l'ordre de nilpotence de  $v_i$ .



## Chapitre XII

# ESPACES PRÉHILBERTIENS

Dans ce chapitre nous reprendrons pour l'essentiel les résultats sur le produit scalaire vus en MPSI que nous compléterons.

Dans une première partie nous reprendrons et compléterons les résultats de MPSI sur le produit scalaire. Dans une seconde nous étudierons la notion de suite totale orthonormée, qui est une notion cruciale de l'analyse et des sciences en générales.

### Espaces préhilbertiens réels

Dans toute cette partie  $\mathbf{E}$  désigne un espace vectoriel sur le corps  $\mathbf{R}$  de dimension *a priori* quelconque.

#### 1.1 Définitions, règles de calculs

**Définition 1.1.1.** — On appelle *forme bilinéaire* (resp. *forme bilinéaire symétrique*) sur  $\mathbf{E}$ , toute application bilinéaire (resp. bilinéaire symétrique) de  $\mathbf{E} \times \mathbf{E}$  dans  $\mathbf{R}$ .

**Exemples 1.1.2.** —

1. Soient  $a$  et  $b$  des réels tels que  $a < b$ . Prenons pour  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}_m^0([a, b], \mathbf{R})$ . L'application

$$\mathbf{E} \times \mathbf{E} \rightarrow \mathbf{R}; (f, g) \mapsto \int_a^b f(t)g(t)dt$$

est une forme bilinéaire symétrique sur  $\mathbf{E}$ .

2. Prenons pour  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([0, 1], \mathbf{R})$ . L'application

$$\mathbf{E} \times \mathbf{E} \rightarrow \mathbf{R}; (f, g) \mapsto \int_0^1 \sqrt{1+t^2} f(t)g(t)dt$$

est une forme bilinéaire symétrique sur  $\mathbf{E}$ .

3. Soit  $n$  un élément de  $\mathbf{N}^*$ . L'application

$$\mathcal{M}_n(\mathbf{R}) \times \mathcal{M}_n(\mathbf{R}) \rightarrow \mathbf{R}; (A, B) \mapsto \text{Tr}(AB)$$

est une forme bilinéaire symétrique sur  $\mathbf{E}$ .

**Définition 1.1.3.** — Soit  $\varphi$  une forme bilinéaire symétrique.

- On dit que  $\varphi$ , est *positive* (resp. *négative*) si par définition, pour tout élément  $\vec{x}$  de  $\mathbf{E}$ ,  $\varphi(\vec{x}, \vec{x}) \geq 0$ , (resp.  $\varphi(\vec{x}, \vec{x}) \leq 0$ ).
- On dit que  $\varphi$ , est *définie* si par définition, pour tout élément  $\vec{x}$  de  $\mathbf{E}$ ,  $\varphi(\vec{x}, \vec{x}) = 0$  si et seulement si  $\vec{x} = \vec{0}_E$ .

**Exemples 1.1.4.** — Reprenons les exemples 1.1.2.



1.

2.

3.

**Définition 1.1.5.** — Soit  $\mathbf{E}$  un espace vectoriel sur le corps  $\mathbf{R}$ . on appelle produit scalaire sur  $\mathbf{E}$  toute forme bilinéaire  $\varphi$ , définie sur  $\mathbf{E}$ , symétrique, définie, positive.

**Notations** — Lorsque  $\varphi$  est un produit scalaire sur  $\mathbf{E}$  on note généralement, pour  $\vec{x}$  et  $\vec{y}$  vecteurs de  $\mathbf{E}$ ,  $\varphi(\vec{x}, \vec{y}) = \langle \vec{x} | \vec{y} \rangle$  ou  $(\vec{x} | \vec{y})$  ... On note alors  $\varphi$ ,  $\langle \cdot | \cdot \rangle$  ou  $(\cdot | \cdot)$  ...

**Définition 1.1.6 .** — On appelle espace préhilbertien réel tout espace vectoriel sur  $\mathbf{R}$ , muni d'un produit scalaire, ou pour être plus précis, tout couple  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$ , où  $\mathbf{E}$  est un espace vectoriel sur  $\mathbf{R}$ , et  $\langle \cdot | \cdot \rangle$  un produit scalaire sur  $\mathbf{E}$ . Si  $\mathbf{E}$  est de dimension finie on dit que  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$  est un espace euclidien.

**Exemples 1.1.7.** —

1. L'application  $\varphi$

$$\mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}; ((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) \mapsto \sum_{i=1}^n x_i y_i$$

est un produit scalaire sur  $\mathbf{R}^n$ . On l'appelle produit scalaire canonique sur  $\mathbf{R}^n$ .

2. Prenons pour  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([-1, 1], \mathbf{R})$ . On a vu en sup. que l'application

$$\mathbf{E} \times \mathbf{E} \rightarrow \mathbf{R}; (f, g) \mapsto \int_{-1}^1 f(t)g(t)dt$$

est un produit scalaire sur  $\mathbf{E}$ .

3. Dans le précédent chapitre on a vu l'exemple suivant. soient  $\omega$  l'application

$$\mathbf{R} \rightarrow \mathbf{R}; t \mapsto \exp(-t^2)$$

et  $\mathbf{H}$  l'ensemble des applications  $f$  de  $\mathbf{R}$  dans  $\mathbf{R}$  continues telles que  $f^2\omega$  soit intégrable. On avait alors montré :

— si  $f$  et  $g$  sont dans  $\mathbf{H}$ , alors  $fg\omega$  est intégrable,

—  $\mathbf{H}$  est un espace vectoriel qui contient les applications polynômes,

— enfin

$$\mathbf{H}^2 \rightarrow \mathbf{R}; (f, g) \mapsto \int_{\mathbf{R}} fg\omega$$

est un produit scalaire sur  $\mathbf{H}$ .

Donnons un exemple assez voisin

4. Prenons pour  $\mathbf{E}$  l'ensemble  $\ell^2(\mathbf{N})$ , des suites  $(u_n)_{n \in \mathbf{N}}$ , notée plus simplement  $u$ , à valeurs réelles, telles que la série  $\sum u_n^2$  converge. Montrons que pour tout couple  $(u, v)$  d'éléments de  $\ell^2(\mathbf{N})$ , la série  $\sum u_n v_n$  converge absolument.

On en déduit que  $\ell^2(\mathbf{N})$  est un espace vectoriel. En effet.....

On peut donc définir l'application

$$\varphi : \ell^2(\mathbf{N}) \times \ell^2(\mathbf{N}) \rightarrow \mathbf{R}; (u, v) \mapsto \sum_{n=0}^{+\infty} u_n v_n.$$

Cette application, on l'a vu, est un produit scalaire sur  $\ell^2(\mathbf{N})$ , en effet

Dans tout la suite de ce paragraphe,  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$  désigne un espace préhilbertien.

On a vu en MPSI l'inégalité de Cauchy-Schwarz. Cette inégalité vaut en fait pour une forme bilinéaire positive, précisément on a :

**Proposition 1.1.8.** — INÉGALITÉ DE CAUCHY-SCHWARZ —

Soit  $\varphi$  une forme bilinéaire symétrique positive sur  $E$ . On note  $q$  l'application

$$q : \mathbf{E} \rightarrow \mathbf{E}; \vec{x} \mapsto \varphi(\vec{x}, \vec{x})$$

(forme quadratique associée à  $\varphi$ ).

Soient  $\vec{x}$  et  $\vec{y}$  des éléments de  $\mathbf{E}$ . Alors :

1. On l'inégalité suivante

$$|\varphi(\vec{x}, \vec{y})| \leq \sqrt{q(\vec{x})q(\vec{y})}. \quad (\text{XII.1})$$

2. Si  $\vec{x}$  et  $\vec{y}$  sont colinéaires, alors il y a égalité dans (XII.1).

3. On suppose de plus que  $\varphi$  est définie. Si il y a égalité dans (XII.1), alors  $\vec{x}$  et  $\vec{y}$  sont colinéaires.

Preuve de la proposition 1.1.8. —

1. On considère l'application

$$p : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto q(\vec{x} + t\vec{y}).$$

Pour tout réel  $t$ ,  $\varphi$  étant bilinéaire et symétrique,

$$p(t) = q(\vec{y})t^2 + 2\varphi(\vec{x}, \vec{y})t + q(\vec{x}).$$

Envisageons deux cas.

- PREMIER CAS :  $q(\vec{y}) = 0$ <sup>1</sup>

l'application  $p$  est alors une application polynômiale de degré au plus 1. Or  $q$  étant positive, pour tout réel  $t$ ,  $p(t) \geq 0$ . Donc l'application  $p$  ne saurait être de degré 1 (sans quoi elle prendrait des valeurs strictement négatives). Donc  $d^\circ p = 0$  et donc  $\varphi(\vec{x}, \vec{y}) = 0$ . L'inégalité (XII.1) est alors vérifiée puisque elle dégénère en  $0 \leq 0$ .

- SECOND CAS :  $q(\vec{y}) \neq 0$

l'application  $p$  est alors une application polynômiale de degré 2,  $q$  étant positive, elle ne prend que des valeurs positives ou nulles; donc son discriminant est positif ou nul, soit :

$$\varphi(\vec{x}, \vec{y})^2 - q(\vec{x})q(\vec{y}) \leq 0.$$

D'où l'on déduit immédiatement (XII.1).

2. Supposons que  $\vec{x}$  et  $\vec{y}$  soient colinéaires. Il existe donc un réel  $\lambda$  tel que  $\vec{x} = \lambda\vec{y}$  ou  $\vec{y} = \lambda\vec{x}$ . Traitons par exemple le premier cas.

D'une part

$$|\varphi(\vec{x}, \vec{y})| = |\varphi(\lambda\vec{y}, \vec{y})| = |\lambda| |\varphi(\vec{y}, \vec{y})| = |\lambda| q(\vec{y}),$$

en tenant compte de la positivité de  $q(\vec{y})$ .

D'autre part,

$$\sqrt{q(\vec{x})q(\vec{y})} = \sqrt{q(\lambda\vec{y})q(\vec{y})} = \sqrt{\lambda^2 q(\vec{y})q(\vec{y})} = |\lambda| q(\vec{y}),$$

toujours grâce à  $q(\vec{y}) \geq 0$ . D'où l'égalité dans (XII.1).

3. On prend maintenant  $q$  définie. Supposons qu'il y ait égalité dans (XII.1). Traitons encore deux cas.

- PREMIER CAS :  $q(\vec{y}) = 0$

La forme bilinéaire symétrique  $\varphi$  étant définie, on en déduit que  $\vec{y} = \vec{0}_{\mathbf{E}}$  et donc  $\vec{x}$  et  $\vec{y}$  sont colinéaires.

- SECOND CAS :  $q(\vec{y}) \neq 0$

L'égalité dans (XII.1) donne que le discriminant du polynôme du second degré  $p$  est nul. Il existe donc un (et un seul) réel  $t_0$  tel que  $q(\vec{x} + t_0\vec{y}) = p(t_0) = 0$ . On a toujours grâce au caractère définie de  $\varphi$ ,  $\vec{x} + t_0\vec{y} = \vec{0}_{\mathbf{E}}$ .

Donc  $\vec{x}$  et  $\vec{y}$  sont colinéaires. □

**Application** — On considère la forme bilinéaire symétrique de l'exemple 1.1.2.1. L'inégalité de Cauchy-Schwarz s'écrit : Pour tout  $f$  et tout  $g$  éléments de  $C_m^0([a, b], \mathbf{R})$ ,

$$\left| \int_a^b f(t)g(t)dt \right| \leq \left( \int_a^b f(t)^2 dt \right)^{1/2} \left( \int_a^b g(t)^2 dt \right)^{1/2}.$$

Dans le cas d'un produit scalaire elle prend la forme :

1. Attention!  $q$  n'est pas *a priori* définie, donc on ne peut conclure à la nullité de  $\vec{y}$ .

pour tout  $\vec{x}$  et tout  $\vec{y}$  éléments, de  $\mathbf{E}$ ,

$$|\langle \vec{x} | \vec{y} \rangle| \leq \sqrt{\langle \vec{x} | \vec{x} \rangle} \sqrt{\langle \vec{y} | \vec{y} \rangle},$$

avec égalité si et seulement si  $\vec{x}$  et  $\vec{y}$  sont colinéaires.

Nous allons en déduire un lemme important.

**Lemme 1.1.9.** — INÉGALITÉ DE MINKOWSKI —

Pour tout  $\vec{x}$  et tout  $\vec{y}$  éléments, de  $\mathbf{E}$ ,

$$\sqrt{\langle \vec{x} + \vec{y} | \vec{x} + \vec{y} \rangle} \leq \sqrt{\langle \vec{x} | \vec{x} \rangle} + \sqrt{\langle \vec{y} | \vec{y} \rangle},$$

De plus, il y a égalité dans cette inégalité si et seulement si  $\vec{x}$  et  $\vec{y}$  sont colinéaires de même sens.

Preuve du lemme 1.1.9. —

- Par bilinéarité et symétrie du produit scalaire on a :

$$\langle \vec{x} + \vec{y} | \vec{x} + \vec{y} \rangle = \langle \vec{x} | \vec{x} \rangle + \langle \vec{y} | \vec{y} \rangle + 2\langle \vec{x} | \vec{y} \rangle.$$

Or d'après l'inégalité de Cauchy-Schwarz,

$$|\langle \vec{x} | \vec{y} \rangle| \leq |\langle \vec{x} | \vec{y} \rangle| \leq \sqrt{\langle \vec{x} | \vec{x} \rangle} \sqrt{\langle \vec{y} | \vec{y} \rangle}, \quad (\text{XII.2})$$

donc

$$\langle \vec{x} + \vec{y} | \vec{x} + \vec{y} \rangle \leq \langle \vec{x} | \vec{x} \rangle + \langle \vec{y} | \vec{y} \rangle + 2\sqrt{\langle \vec{x} | \vec{x} \rangle} \sqrt{\langle \vec{y} | \vec{y} \rangle} = \left( \sqrt{\langle \vec{x} | \vec{x} \rangle} + \sqrt{\langle \vec{y} | \vec{y} \rangle} \right)^2$$

La quantité  $\sqrt{\langle \vec{x} | \vec{x} \rangle} + \sqrt{\langle \vec{y} | \vec{y} \rangle}$  étant positive ou nulle, on a l'inégalité promise.

- Traitons de l'égalité. Il y a égalité dans l'inégalité de Minkowski, si et seulement si les deux inégalités dans (XII.2) sont des égalités. Or il y a égalité dans l'inégalité de Cauchy-Schwarz, si et seulement si les vecteurs sont colinéaires, de plus le produit scalaire de deux vecteurs colinéaires est positif si et seulement si ces vecteurs sont de même sens.

Donc, finalement il y a égalité dans cette inégalité si et seulement si  $\vec{x}$  et  $\vec{y}$  sont colinéaires de même sens.  $\square$

**Corollaire-définition 1.1.10.** — L'application  $\mathbf{E} \rightarrow \mathbf{R}; \vec{x} \mapsto \sqrt{\langle \vec{x} | \vec{x} \rangle}$  est une norme sur  $\mathbf{E}$ . On la notera dans la suite  $\|\cdot\|$ . On dit que c'est la norme associée au produit scalaire  $\langle \cdot | \cdot \rangle$ . Une norme associée à un produit scalaire est dite hilbertienne, ou euclidienne surtout si la dimension est finie.

Preuve du corollaire 1.1.10. — La démonstration est immédiate, seule la preuve de l'inégalité triangulaire n'est pas triviale, mais cette dernière est fournie instantanément par l'inégalité de Minkowski.  $\square$

Nous allons maintenant donner un certain nombre de formules fondamentales. On désigne par  $\vec{x}$  et  $\vec{y}$  des éléments de  $\mathbf{E}$ .

INÉGALITÉ DE CAUCHY-SCHWARZ —

$$|\langle \vec{x} | \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|,$$

avec égalité si et seulement si  $\vec{x}$  et  $\vec{y}$  sont colinéaires.

Exprimons le produit scalaire en fonction de la norme. Partons de

$$\langle \vec{x} + \vec{y} | \vec{x} + \vec{y} \rangle = \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\langle \vec{x} | \vec{y} \rangle. \quad (\text{XII.3})$$

D'où

FORMULE DE POLARISATION —

$$\langle \vec{x} | \vec{y} \rangle = \frac{1}{2} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2).$$

On a aussi

$$\langle \vec{x} - \vec{y} | \vec{x} - \vec{y} \rangle = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\langle \vec{x} | \vec{y} \rangle. \quad (\text{XII.4})$$

Par différence de (64) et de (65), on obtient :

$$\langle \vec{x} | \vec{y} \rangle = \frac{1}{4} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2).$$

Enfin, par somme de (64) et de (65), on obtient l'égalité suivante.

$$\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = 2 (\|\vec{x}\|^2 + \|\vec{y}\|^2).$$

Cette égalité s'appelle ainsi car elle admet l'interprétation géométrique suivante :

Cette inégalité joue un rôle capitale dans les mathématiques. C'est elle qui est le moteur de la preuve du théorème de projection sur un convexe fermé (voir exercices). Ce théorème admet de nombreuses applications en analyse, notamment dans l'étude d'équations différentielles et aux dérivées partielles.

L'expression du produit scalaire en fonction de la norme permet de déduire des propriétés du produit scalaire de propriétés similaires pour la norme. Donnons un exemple

**Exercice** — Soit  $\ell$  un endomorphisme de  $\mathbf{E}$ . on suppose que  $\ell$  conserve la norme. Montrer qu'il conserve le produit scalaire.

**Remarque** — Dans un exercice, lorsque l'on demande de montrer qu'une application est une norme et que cette dernière s'écrit avec un radical, il s'agit souvent d'une norme hilbertienne. il est alors plus facile de déterminer le produit scalaire dont elle provient que de prouver directement que c'est une norme. Ceci provient du fait que, dans le cas d'une norme hilbertienne, la preuve de l'inégalité triangulaire utilise deux résultats délicats : les inégalités de Cauchy-Schwarz et de Minkowski.

**Exercice** — Soit  $n$  un entier strictement positif. Montrer que l'application

$$N : \mathcal{M}_n(\mathbf{R}) \rightarrow \mathbf{R}, ; A \mapsto \sqrt{\text{Tr}(^tAA)}$$

est une norme sur  $\mathcal{M}_n(\mathbf{R})$ .

## 1.2 Orthogonalité

Dans le paragraphe 1.2, nous désignerons par  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$  un espace préhilbertien réel et par  $\|\cdot\|$  la norme associée.

**Définition 1.2.1.** — Soient  $\vec{x}$  et  $\vec{y}$  des éléments de  $\mathbf{E}$  on dit que  $\vec{x}$  est orthogonal à  $\vec{y}$ , si par définition  $\langle \vec{x} | \vec{y} \rangle = 0$ . On note  $\vec{x} \perp \vec{y}$ .

**Remarque** — Le caractère symétrique du produit scalaire fait que si  $\vec{x}$  est orthogonal à  $\vec{y}$ , alors  $\vec{y}$  est orthogonal à  $\vec{x}$ , ce qui autorise à dire que *des vecteurs sont orthogonaux*.

**Définition 1.2.2.** — Soit  $(\vec{x}_i)_{i \in I}$  une famille de vecteurs de  $\mathbf{E}$ , ( $I$  est un ensemble quelconque non vide.)

1. On dit que la famille  $(\vec{x}_i)_{i \in I}$  est orthogonale si, par définition, pour tout  $i$  et tout  $j$  éléments distincts de  $I$ ,

$$\vec{x}_i \perp \vec{x}_j.$$

2. On dit que la famille  $(\vec{x}_i)_{i \in I}$  est orthonormale (ou orthonormée) si, par définition, elle est orthogonale et pour tout élément  $i$  de  $I$ ,  $\|\vec{x}_i\| = 1$ .

**Remarque** — La famille  $(\vec{x}_i)_{i \in I}$  est donc orthonormale si et seulement si, pour tout  $i$  et tout  $j$  éléments de  $I$ ,  $\langle \vec{x}_i | \vec{x}_j \rangle = \delta_{i,j}$ .

**Proposition 1.2.3.** — Soit  $(\vec{x}_i)_{i \in I}$  une famille orthogonale de vecteurs de  $\mathbf{E}$ . Si pour tout élément  $i$  de  $I$ ,  $\vec{x}_i \neq \vec{0}$ , alors la famille est libre.

En particulier toute famille orthonormale de vecteurs de  $\mathbf{E}$  est libre.

*Preuve de la proposition 1.2.3. —*

**Remarque** — Etant donnée un vecteur  $\vec{x}$ , non nul de  $\mathbf{E}$ , il existe deux et seulement deux vecteurs colinéaires à  $\vec{x}$  et de norme 1 (ou comme on dit normés),

$$\frac{1}{\|\vec{x}\|}\vec{x} \text{ et } -\frac{1}{\|\vec{x}\|}\vec{x}.$$

**Proposition 1.2.4.** — THÉORÈME DE TCHU-KUNG PYTHAGORE —

1. Soient  $\vec{x}$  et  $\vec{y}$  des éléments de  $\mathbf{E}$ .  $\vec{x} \perp \vec{y}$  si et seulement si  $\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$ .
2. Soit  $(\vec{x}_i)_{i \in \{1, \dots, n\}}$  une famille finie de vecteurs de  $\mathbf{E}$ . Si  $(\vec{x}_i)_{i \in \{1, \dots, n\}}$  est orthogonale alors

$$\left\| \sum_{i \in \{1, \dots, n\}} \vec{x}_i \right\|^2 = \sum_{i \in \{1, \dots, n\}} \|\vec{x}_i\|^2.$$

*Preuve de la proposition 2.2.4. —*

Nous allons maintenant examiner l'existence, en dimension finie, de bases orthonormées. Nous allons démontrer en fait plus que leur existence, à savoir, donner un procédé algorithmique de construction de telles bases. Commençons par la proposition essentielle :

**Proposition 1.2.5.** — PROCÉDÉ D'ORTHONORMALISATION DE GRAM ET SCHMIDT —

Soit  $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$  une famille **libre** de vecteurs de  $\mathbf{E}$ .

- Il existe une et une seule famille orthonormée de vecteurs de  $\mathbf{E}$ ,  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ , telle que :
  - i. Pour tout élément  $k$  de  $\{1, 2, \dots, p\}$ ,  $\text{vec}(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k) = \text{vec}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k)$ .

- ii. Pour tout élément  $k$  de  $\{1, 2, \dots, p\}$ ,  $\langle \vec{x}_k | \vec{e}_k \rangle > 0$ .
- De plus cette famille est donnée par la relation de récurrence suivante :

$$\begin{cases} \vec{e}_1 = \frac{1}{\|\vec{x}_1\|} \vec{x}_1, \\ \text{Pour } k = 1, \dots, p-1, \vec{e}_{k+1} = \frac{1}{\|\vec{y}_{k+1}\|} \vec{y}_{k+1}, \text{ où } \vec{y}_{k+1} = \vec{x}_{k+1} - \sum_{i=1}^k \langle \vec{x}_{k+1} | \vec{e}_i \rangle \vec{e}_i. \end{cases}$$

**Remarque** — Si l'on supprime la condition ii., on n'a plus que l'unicité de la famille au signe près de chacun de ses vecteurs, donc en fait  $2^p$  familles.

*Preuve de la proposition 1.2.5.* — Nous renvoyons au cours de sup. pour cette preuve. Du reste nous traiterons en exercice un résultat voisin dans le cadre de familles non finies.

**Corollaire 1.2.6.** — Si l'on suppose  $\mathbf{E}$  de dimension finie non nulle  $n$ , alors  $\mathbf{E}$  admet des bases orthonormées.

*Preuve de la proposition 1.2.6.* —

**Corollaire 1.2.7.** — THÉORÈME DE LA BASE INCOMPLÈTE ORTHONORMÉE — soit  $(\vec{x}_1, \dots, \vec{x}_p)$  une famille orthonormée de  $\mathbf{E}$  supposé de dimension finie  $n$ . Alors cette famille se complète en une base orthonormée de  $\mathbf{E}$ ,

$$(\vec{x}_1, \dots, \vec{x}_p, \vec{x}_{p+1}, \dots, \vec{x}_n).$$

*Preuve du corollaire 1.2.7.*

#### EXPRESSION DES COORDONNÉES ET DU PRODUIT SCALAIRE DANS UNE BASE

Nous supposons pour établir le formulaire ci-dessous que  $\mathbf{E}$  est de dimension finie non nulle  $n$ . On considère  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base orthonormée de  $\mathbf{E}$ , notée  $\mathcal{B}$ . Soient  $\vec{x}$  et  $\vec{y}$  des éléments de  $\mathbf{E}$ . Pour  $i = 1, 2, \dots, n$  on note  $x_i$  (resp.  $y_i$ ) la  $i^{\text{e}}$  coordonnée de  $\vec{x}$  (resp.  $\vec{y}$ ) dans  $\mathcal{B}$ , et  $X$  (resp.  $Y$ ) le vecteur colonne des coordonnées de  $\vec{x}$  (resp.  $\vec{y}$ ) dans  $\mathcal{B}$ .

le vecteur  $\vec{x}$  se décompose dans  $\mathcal{B}$  en  $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$

Donc

$$\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$$



Calculons  $\langle \vec{x} | \vec{y} \rangle = \dots$

Le produit scalaire s'exprime facilement au moyen des vecteurs colonnes coordonnée

$$\langle \vec{x} | \vec{y} \rangle = \dots$$

$$\langle \vec{x} | \vec{y} \rangle = \sum_{i=1}^n$$

En particulier,

$$\|x\| = \left( \sum_{i=1}^n \right)^{1/2}$$

Enfin, si  $d$  désigne la distance associée à  $\|\cdot\|$ ,

$$d(x, y) = \left( \sum_{i=1}^n \right)^{1/2}$$

**Complément** — On suppose que  $\mathbf{E}$  est de dimension finie non nulle  $n$ . On considère  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  une base quelconque de  $\mathbf{E}$ , notée  $\mathcal{B}$ . Soient  $\vec{x}$  et  $\vec{y}$  des éléments de  $\mathbf{E}$ .  $X$  (resp.  $Y$ ) le vecteur colonne des coordonnées de  $\vec{x}$  (resp.  $\vec{y}$ ) dans  $\mathcal{B}$ ). Prenons par ailleurs  $\mathcal{B}_o$  une base orthonormée de  $\mathbf{E}$  et désignons par  $P$  la matrice de passage de  $\mathcal{B}_o$  à  $\mathcal{B}$

### Exercice 1.2.8. — POLYNÔMES DE LEGENDRE —

1. Montrer que l'application

$$\mathbf{R}[X] \times \mathbf{R}[X] \rightarrow \mathbf{R}; (p, q) \mapsto \int_{-1}^1 p(t)q(t)dt$$

est un produit scalaire sur  $\mathbf{R}[X]$ . Dans la suite  $\mathbf{R}[X]$  sera muni de ce produit que l'on notera  $\langle \cdot | \cdot \rangle$ .

2. Montrer qu'il existe une et une seule famille orthogonale de  $\mathbf{R}[X]$ ,  $(p_n)_{n \in \mathbf{N}}$ , telle que pour tout entier naturel  $n$ ,  $p_n$  soit unitaire de degré  $n$ .

Cette famille s'appelle famille des polynômes orthogonaux de Legendre. Nous étudierons en T.D. (et en détail) les propriétés de différentes familles de polynômes orthogonaux.

Solution de l'exercice 1.2.8. —

1. ...

2. On pourrait montrer en même temps l'unicité et l'existence, mais pour rendre plus fluide la rédaction, nous choisirons de séparer l'existence et l'unicité.

- *Existence*

Supposons avoir construit une famille orthogonale  $(p_0, p_1, \dots, p_m)$  de  $\mathbf{R}[X]$ , telle que  $p_k$  soit unitaire de degré  $k$ , pour  $k = 0, 1, \dots, m$ . La famille  $(p_0, p_1, \dots, p_m)$  est une famille libre de  $\mathbf{R}[X]_m$  (cf. ...) elle en est donc une base (son cardinal est égal à la dimension de  $\mathbf{R}[X]_m$ ). Donc tout élément  $p$  de  $\mathbf{R}[X]$  de degré  $m+1$  unitaire s'écrit

$$p = X^{m+1} + \sum_{k=0}^m \alpha_k p_k,$$

avec  $\alpha_0, \alpha_1, \dots, \alpha_m$  des réels.

Donc la suite  $(p_n)_{n \in \mathbf{N}}$  d'éléments de  $\mathbf{R}[X]$  définie par récurrence par

$$\begin{cases} p_0 = \\ \text{Pour tout } n \in \mathbf{N}, p_{n+1} = \dots \end{cases}$$

répond à la question.

- *Unicité*

Soient  $(p_n)_{n \in \mathbf{N}}$  et  $(\tilde{p}_n)_{n \in \mathbf{N}}$  des familles orthogonale de  $\mathbf{R}[X]$  de polynômes unitaires de degrés échelonnés.

D'où l'unicité.

*On suppose de nouveau que  $\mathbf{E}$  est de dimension quelconque.*

**Définition 1.2.9.** — Soient  $\mathbf{F}$  et  $\mathbf{G}$  des sous-espaces vectoriels de  $\mathbf{E}$ . On dit que  $\mathbf{F}$  et  $\mathbf{G}$  sont orthogonaux si, par définition, tout vecteur de l'un est orthogonal à tout vecteur de l'autre; on note alors :  $\mathbf{F} \perp \mathbf{G}$ .

**Proposition 1.2.10.** — Soient  $\mathbf{F}$  et  $\mathbf{G}$  des sous-espaces vectoriels de  $\mathbf{E}$ . Si  $\mathbf{F} \perp \mathbf{G}$  alors  $\mathbf{F}$  et  $\mathbf{G}$  sont en somme directe. On dit alors que  $\mathbf{F}$  et  $\mathbf{G}$  sont en somme directe orthogonale et l'on note :  $\mathbf{F} + \mathbf{G} = \mathbf{F} \oplus \mathbf{G}$ .

*Preuve de la proposition 2.2.9.* —

**Définition 1.2.11.** — Soient  $\mathbf{F}$  et  $\mathbf{G}$  des sous-espaces vectoriels de  $\mathbf{E}$ . On dit que  $\mathbf{F}$  et  $\mathbf{G}$  sont des supplémentaires orthogonaux si, par définition ils sont supplémentaires et orthogonaux, c'est-à-dire si  $\mathbf{E} = \mathbf{F} \oplus \mathbf{G}$ . Plus généralement on dit qu'une famille  $(\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_p)$  est une famille de supplémentaires orthogonaux, si, par définition,

- c'est une famille de supplémentaires :  $\mathbf{E} = \bigoplus_{i=1}^p \mathbf{F}_i$ ;
- les éléments de la famille sont deux à deux orthogonaux : pour tout couple  $(i, j)$  d'éléments distincts de  $\{1, 2, \dots, p\}$ ,  $\mathbf{F}_i \perp \mathbf{F}_j$ .

On note  $\mathbf{E} = \bigoplus_{i=1, \dots, p} \mathbf{F}_i$

On ne sais rien, pour le moment, quant à l'existence, pour tout sous-espace vectoriel  $\mathbf{F}$  de  $\mathbf{E}$ , d'un supplémentaire orthogonal. Donnons par contre un résultat d'unicité :

**Proposition 1.2.12.** — Soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$ . On suppose qu'il admet  $\mathbf{G}_1$  et  $\mathbf{G}_2$  comme supplémentaires orthogonaux. Alors  $\mathbf{G}_1 = \mathbf{G}_2$ .

*Preuve de la proposition 2.2.12.* —

La précédente proposition permet d'user en toute légitimité d'expressions du type « LE supplémentaire orthogonal », du moins lorsque l'on est assuré de son existence. Il permet aussi de définir les projections orthogonales :

**Définition 1.2.13.** — PROJECTION ORTHOGONALE, SYMÉTRIE ORTHOGONALE —

Soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$ , admettant un supplémentaire orthogonal. On appelle projection orthogonale sur  $\mathbf{F}$  (resp. symétrie orthogonale par rapport à  $\mathbf{F}$ ) la projection sur  $\mathbf{F}$  (resp. symétrie par rapport à  $\mathbf{F}$ ) suivant LE supplémentaire orthogonal de  $\mathbf{F}$ .

Généralisons. Soit  $(\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_p)$  une famille de supplémentaires orthogonaux. Pour  $i = 1, 2, \dots, p$ , on peut parler de la projection orthogonale sur  $\mathbf{F}_i$  (associée à la décomposition de  $\mathbf{E}$ ,  $\mathbf{E} = \bigoplus_{i=1}^p \mathbf{F}_i$ ), en effet  $\mathbf{F}_i$  admet pour supplémentaire orthogonal  $\bigoplus_{\substack{j=1 \\ j \neq i}}^p \mathbf{F}_j$ .

**Définition 1.2.14.** — ORTHOGONAL D'UN SOUS-ESPACE VECTORIEL —

Soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$ . On appelle orthogonal de  $\mathbf{F}$ , l'ensemble des vecteurs  $\vec{y}$  de  $\mathbf{E}$  qui sont orthogonaux à  $\mathbf{F}$ , c'est-à-dire orthogonal à tout vecteur de  $\mathbf{F}$ . On le note  $\mathbf{F}^\perp$ .

**Proposition 1.2.15.** — Soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$ . L'orthogonal de  $\mathbf{F}$  est un sous-espace vectoriel de  $\mathbf{E}$ .

*Preuve de la proposition 1.2.15.* —

L'orthogonalité d'un vecteur à  $\mathbf{F}$  se teste sur une partie génératrice, précisément :

**Proposition 1.2.16.** — Soient  $\mathbf{F}$  un sous-espace vectoriel dont  $(\vec{f}_i)_{i \in I}$  est une famille génératrice. Pour qu'un vecteur  $\vec{y}$  de  $\mathbf{E}$  soit orthogonal à  $\mathbf{F}$ , il faut et il suffit que pour tout élément  $i$  de  $I$ ,  $\vec{y} \perp \vec{f}_i$ .

*Preuve de la proposition 2.2.16.* — La nécessité de cette condition est évidente, passons à sa suffisance. Supposons que pour tout élément  $i$  de  $I$ ,  $\vec{y} \perp \vec{f}_i$ .

**Remarque** — On peut parler de l'orthogonal d'une partie  $A$  de  $\mathbf{E}$ , il s'agit de l'ensemble des vecteurs orthogonaux à tous les éléments de  $A$ , encore noté  $A^\perp$ . La précédente proposition nous dit que c'est aussi l'orthogonal de  $\text{vect}(A)$ . Ainsi Lorsque l'orthogonal  $\{a\}$ , est celui de la droite  $\{\vec{a}\}$ , il se note souvent simplement  $\vec{a}^\perp$ .

**Exercice 1.2.17.** — On reprend les notations de l'exercice 2.2.7.  $(p_n)_{n \in \mathbf{N}}$  est la suite des polynômes orthogonaux de Legendre. Montrer que pour tout entier naturel  $n$ ,  $p_{n+1}$  est orthogonal à  $\mathbf{R}[X]_n$ .

*Solution de l'exercice 1.2.17. —*

Faisons le lien entre orthogonal et supplémentaire orthogonal.

**Proposition 1.2.18.** — Soit  $\mathbf{F}$  un sous-espace vectoriel de  $\mathbf{E}$  on suppose qu'il admet un supplémentaire orthogonal noté  $\mathbf{G}$ . Alors  $\mathbf{G}$  est l'orthogonal de  $\mathbf{F}$ .

*Preuve de la proposition 1.2.18. —*

Récapitulons :

- un sous-espace admet au plus un supplémentaire orthogonal ;
- si un sous-espace admet un supplémentaire orthogonal, alors il s'agit de son orthogonal.

Nous allons dans le paragraphe suivant, sous certaines hypothèses, prouver l'existence d'un supplémentaire orthogonal.

### 1.3 Théorème de projection

Dans le paragraphe 1.3, nous désignerons encore par  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$  un espace préhilbertien réel et par  $\|\cdot\|$  la norme associée. Nous allons montrer qu'un sous-espace vectoriel  $\mathbf{F}$  de dimension finie de  $\mathbf{E}$  admet un supplémentaire orthogonal. Pour commencer nous établirons le résultat dans le cas où  $\mathbf{E}$  lui-même est de dimension finie, reprenant ainsi le cours de MPSI.

CAS PARTICULIER OÙ  $\mathbf{E}$  EST DE DIMENSION FINIE

On suppose que  $\mathbf{E}$  est de dimension finie non nulle  $n$ . Soit  $\mathbf{F}$  un sous-espace vectoriel  $\mathbf{F}$  de  $\mathbf{E}$  de dimension non nulle  $p$ .

Pour commencer montrons que  $\mathbf{E}$  admet une base orthonormée  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$  adaptée à  $\mathbf{F}$ .

Nous noterons dans la suite  $\mathcal{B}_0$  cette base.

Posons  $\mathbf{G} := \dots$ . Montrons que  $\mathbf{G}$  est le supplémentaire orthogonal de  $\mathbf{F}$ .

On voit bien dans ce cas, que comme le dit 1.2.18., que  $\mathbf{G} = \mathbf{F}^\perp$ .

On dispose donc de la projection orthogonale sur  $\mathbf{F}$ , notée  $p_F$ . Donnons en l'expression. Soit  $\vec{x}$  un élément de  $\mathbf{E}$ .

donc

$$p_f(\vec{x}) = \sum_{i=1}^p \dots \vec{e}_i. \quad (\text{XII.5})$$

CAS GÉNÉRAL,  $\mathbf{E}$  EST UN ESPACE PRÉHILBERTIEN DE DIMENSION FINIE OU NON

Donnons le résultat :

**Proposition 1.3.1.** — *Tout sous-espace vectoriel  $\mathbf{F}$  de  $\mathbf{E}$  de dimension finie, admet un supplémentaire orthogonal.*

D'après 1.2.12. et 1.2.18.  $\mathbf{F}$  admet donc un et un seul supplémentaire orthogonal qui est  $\mathbf{F}^\perp$ .

*Preuve de la proposition 1.3.1.* — On exclut dans la suite le cas trivial où  $\mathbf{F} = \{\vec{0}_{\mathbf{E}}\}$ , cas où  $\mathbf{F}$  admet pour supplémentaire

orthogonal... . On dispose donc d'une base orthonormée  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  de  $\mathbf{F}$ , obtenue...

On est donc en mesure de définir la projection orthogonale sur  $\mathbf{F}$ .

**Remarque** — On montrera peut-être en exercice pour les candidats ENS qu'un sous-espace vectoriel  $\mathbf{F}$  de  $\mathbf{E}$ , de dimension quelconque, admet un supplémentaire dans les deux cas suivants :

- $(\mathbf{E}, \langle \cdot | \cdot \rangle)$  un espace de Hilbert et  $\mathbf{F}$  est fermé dans  $(\mathbf{E}, \| \dots \|)$  ;
- $\mathbf{F}$  est, Plus généralement, une partie complète de  $(\mathbf{E}, \| \dots \|)$ .

Il existe des sous-espaces vectoriels qui ne possède pas de supplémentaire orthogonal. Donnons un exemple

**Exemple** —

**Proposition 1.3.2.** — THÉORÈME DE PROJECTION —

Soit  $\mathbf{F}$  un espace-vectoriel de dimension finie de  $\mathbf{E}$ , ( $\mathbf{E}$  est de dimension quelconque). Notons  $p_{\mathbf{F}}$  la projection orthogonale sur  $\mathbf{F}$ . Soient  $\vec{x}$  un élément de  $\mathbf{E}$  et  $\vec{a}$  un élément de  $\mathbf{F}$ . On a alors l'équivalence des trois propositions suivantes :

1.  $\|\vec{x} - \vec{a}\| = d(\vec{x}, \mathbf{F})$ .
2.  $(\vec{x} - \vec{a}) \perp \mathbf{F}$ .
3.  $\vec{a} = p_{\mathbf{F}}(\vec{x})$ .

Ce résultat exprime essentiellement, que la distance de  $\vec{x}$  à  $\mathbf{F}$  est atteinte en un et un seul point, le projeté orthogonal de  $\vec{x}$  sur  $\mathbf{F}$ .

*Preuve de la proposition 1.3.2. —*

La preuve de 1.3.1. a montré que pour tout élément  $\vec{x}$  de  $\mathbf{E}$ , et pour toute base  $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$  orthonormée de  $\mathbf{F}$  (supposé donc ici non nul), la projection de  $\vec{x}$  sur  $\mathbf{F}$  est donnée par la formule (XII.5).

$$p_{\mathbf{F}}(\vec{x}) = \sum_{i=1}^p \langle \vec{x}, \vec{e}_i \rangle \vec{e}_i$$

Nous en déduisons immédiatement l'inégalité dite de Bessel :

$$\sum_{i=1}^p \langle \vec{x}, \vec{e}_i \rangle^2 \leq \|\vec{x}\|^2$$

Tirons de cette formule des cas particuliers intéressants. Soit  $\vec{u}$  un vecteur non nul de  $\mathbf{E}$ . On désigne par  $\Delta$  la droite vectorielle qu'il engendre. Notons respectivement  $p_{\Delta}$  et  $s_{\Delta}$  la projection orthogonale sur  $\Delta$  et la symétrie orthogonale



par rapport à  $\Delta$ .

Donc, pour tout élément  $\vec{x}$  de  $\mathbf{E}$ ,

$p_{\Delta}(\vec{x}) =$	$s_{\Delta}(\vec{x}) =$
-------------------------	-------------------------

Donnons un exercice très classique.

**Exercice 1.3.3.** — *Existence et valeur de*

$$\min_{(a,b) \in \mathbf{R}^2} \int_{-1}^1 |x^2 + ax + b|^2 \, dx.$$

*Solution de l'exercice 2.3.3.* —

Dans le même esprit donnons une application importante du théorème de projection.

**Application 1.3.4.** — DROITE DES MOINDRES CARRÉS —

On s'intéresse au problème suivant : On dispose de  $p$  mesures d'une grandeur  $y$ ,  $(y_1, y_2, \dots, y_p)$ , obtenue pour  $p$  valeurs distinctes d'une variable  $x$ ,  $(x_1, x_2, \dots, x_p)$ . On sait par ailleurs que  $y$  est une fonction affine de  $x$ . On va s'intéresser à la détermination de cette fonction affine. Idéalement il existerait des réels  $a$  et  $b$  tels que pour tout élément  $i$  de  $\{1, \dots, p\}$ ,  $y_i = ax_i + b$ , ou, si l'on préfère tels que les points  $(x_i, y_i)$ ,  $i = 1, \dots, p$ , soient sur la droite d'équation  $y = ax + b$ . Dans la pratique, soit à cause des erreurs de mesure soit parce que la loi n'est qu'approximativement affine il n'en est rien. On va donc chercher  $a$  et  $b$ , pour qu'en un certain sens, la droite d'équation  $y = ax + b$  passe au « plus près » des points  $(x_i, y_i)$ ,  $i = 1, \dots, p$ .

Une possibilité est de minimiser la somme des carrés des écarts entre les  $y_i$  (valeur mesurée) et  $ax_i + b$  (valeur idéale pour la loi  $y = ax + b$ ), précisément nous allons étudier l'existence d'un minimum pour l'ensemble  $\{D(a, b), (a, b) \in \mathbf{R}^2\}$ , où pour tout élément  $(a, b)$  de  $\mathbf{R}^2$ ,

$$D(a, b) = \sum_{i=1}^p (y_i - (ax_i + b))^2.$$

Le choix de la quantité à minimiser est avant tout un choix pragmatique. Certes, dans certains cas, il est possible de donner un sens physique à  $D(a, b)$ , mais le plus souvent cette quantité est choisie parce que, comme nous allons le montrer elle conduit facilement à l'existence d'un et d'un seul minimum pour  $\{D(a, b), (a, b) \in \mathbf{R}^2\}$ , et n'a pas plus de sens que la somme des valeurs absolues ou des puissances 4 des écarts entre les quantités  $y_i$  et  $ax_i + b$ .

Formalisons le problème

## FAMILLES TOTALE ORTHONORMÉES

### 2.1 Un exemple : Les polynômes de Legendre

On désigne par  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([-1, 1], \mathbf{R})$  que l'on munit du produit scalaire  $\langle \cdot | \cdot \rangle$  définie par :

$$\langle f | g \rangle = \int_{-1}^1 f(t)g(t)dt,$$

pour tout  $f$  et tout  $g$  éléments de  $\mathbf{E}$  (cf. 1.1.3-2.). La norme associée sera notée  $\| \cdot \|_2$ . On identifiera ici les fonctions polynomiales de  $[-1, 1]$  dans  $\mathbf{R}$  et les polyômes auxquels elles sont associées.

Par ailleurs,  $(p_n)_{n \in \mathbf{N}}$  désignant la suite des polynômes orthogonaux de Legendre (exercice 1.2.8.), nous poserons, pour tout entier naturel  $n$ ,

$$\tilde{p}_n := \frac{p_n}{\|p_n\|_2}$$

et pour tout élément  $f$  de  $\mathbf{E}$ ,

$$c_n(f) := \langle \tilde{p}_n | f \rangle,$$

coefficient de Fourier-Legendre de  $f$  d'ordre  $n$ .

Notons que  $(\tilde{p}_n)_{n \in \mathbf{N}}$  est une famille orthonormée de  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$ . C'est par ailleurs une base de  $\mathbf{R}[X]$ . Plus précisément,  $(\tilde{p}_0, \dots, \tilde{p}_n)$  est pour tout entier naturel  $n$  une base orthonormée de  $\mathbf{R}[X]_n$ .

Soit  $f$  un élément de  $\mathbf{E}$ .

On désigne, pour tout  $n \in \mathbf{N}$ ; par  $S_n(f)$  le projeté orthogonal de  $f$  sur  $\mathbf{R}[X]_n$ . Montrons que la suite  $(S_n(f))_{n \in \mathbf{N}}$  converge vers  $f$  dans l'espace vectoriel normé  $(\mathbf{E}, \| \cdot \|_2)^1$ .

On a vu dans la partie précédente que  $S_N(f)$  s'exprime au moyen de termes de  $(\tilde{p}_n)_{n \in \mathbf{N}}$  et de  $(c_n(f))_{n \in \mathbf{N}}$ .

$$S_N(f) =$$

$$\text{et } \|S_N(f)\|_2^2 = \sum_{n=0}^N \dots\dots\dots$$

L'inégalité de Bessel dit.....

On en déduit la convergence de la série  $\sum_{n \geq 0} c_n(f)^2 \dots$

1. Cela ne signifie pas que  $(S_n(f))_{n \in \mathbf{N}}$  converge uniformément ni même simplement vers  $f$  !

On peut du reste préciser la somme...

Finalement on obtient l'inégalité dite de Parseval :

$$\sum_{n=0}^{+\infty} c_n(f)^2 =$$

## 2.2 Famille totale orthonormée, définitions, propriétés

Dans le paragraphe précédent, toute les propriété repose sur le fait que  $\mathbf{R}[X]$  est dense dans  $(\mathbf{E}, \|\cdot\|_2)$  conséquence immédiate de la densité de  $\mathbf{R}[X]$  dans  $(\mathbf{E}, \|\cdot\|_\infty)$ , notons que  $\mathbf{R}[X] = \text{vect} \dots$

D'où les définitions qui vont suivre.

$(\mathbf{E}, \langle \cdot | \cdot \rangle)$  désigne comme à l'accoutumé un espace préhilbertien de dimension quelconque, mais ce qui va suivre n'a d'intérêt que si  $\mathbf{E}$  est de dimension infinie. La norme associée au produit scalaire sera notée  $\|\cdot\|$ . Les propriétés topologiques que l'on étudiera dans  $\mathbf{E}$  s'entendent pour l'espace vectoriel normé  $(\mathbf{E}, \|\cdot\|)$ .

**Définition 2.2.1.** — On appelle suite totale de  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$ , toute suite  $(\vec{x}_n)_{n \in \mathbf{N}}$  telle que l'adhérence de l'espace vectoriel qu'elle engendre soit égale à  $\mathbf{E}$  :

$$\overline{\text{vect}((x_n)_{n \in \mathbf{N}})} = \mathbf{E}$$

**Exemple** — En prenant comme dans le 2.1.,  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([-1, 1], \mathbf{R})$  que l'on munit du produit scalaire  $\langle \cdot | \cdot \rangle$  définie par :  $\langle f | g \rangle = \int_{-1}^1 fg$ , La famille  $(X^n)_{n \in \mathbf{N}}$  est totale, en effet...

**Définition 2.2.2.** — On appelle suite totale orthonormée de  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$ , toute suite  $(\vec{e}_n)_{n \in \mathbf{N}}$  totale telle que la famille  $(\vec{e}_n)_{n \in \mathbf{N}}$  soit orthonormée.

**Exemple** — En prenant toujours,  $\mathbf{E}$  l'espace vectoriel  $\mathcal{C}^0([-1, 1], \mathbf{R})$  que l'on munit du produit scalaire  $\langle \cdot | \cdot \rangle$  définie par :  $\langle f | g \rangle = \int_{-1}^1 fg$ , La famille  $(\tilde{p}_n)_{n \in \mathbf{N}}$  est on la vu dans le premier paragraphe, totale orthonormée.

On donnera d'autres exemples de familles totales obtenues grâce aux polynômes orthonormés en exercice ou TD.

Donnons un exemple important qui joue un peu le même rôle pour les familles totales orthonormées, que  $\mathbf{R}^n$  et sa base canonique pour les bases orthonormées d'un espace de dimension  $n$ .

**Exemple 2.2.3.** — Reprenons l'exemple 1.1.7-4 de  $\ell^2(\mathbf{N})$ . Notons toute suite  $(u_n)_{n \in \mathbf{N}}$  réelle plus simplement  $u$  et pour tout  $p \in \mathbf{N}$  on désigne par  $\delta_{p,\cdot}$  la suite

$$(\delta_{p,n})_{n \in \mathbf{N}},$$

où  $\delta_{p,n}$  désigne le symbole de Kronecker associé à  $p$  et  $n$ .<sup>1</sup> Alors  $(\delta_{p,\cdot})_{p \in \mathbf{N}}$  est une suite totale orthonormée, en

1. En algèbre dans l'étude des polynômes  $\delta_{1,\cdot}$ , se note  $X$ , et l'on montre que pour tout  $p \in \mathbf{N}$ ,  $\delta_{p,\cdot} = X^p \dots$

effet.....

Notons que  $\text{vect}((\delta_{p,\cdot})_{p \in \mathbf{N}})$  est .....!!!

**Notations** —

Jusqu'à la fin on considère une suite totale orthonormée de  $(\mathbf{E}, \langle \cdot | \cdot \rangle)$   $(\vec{e}_n)_{n \in \mathbf{N}}$  et  $\vec{x}$  un élément de  $\mathbf{E}$ .

On pose pour tout  $n \in \mathbf{N}$ ,  $c_n(\vec{x}) = \langle \vec{x} | \vec{e}_n \rangle$ ,  $\mathbf{E}_n = \text{vect}(\vec{e}_0, \dots, \vec{e}_n)$ . et  $S_n(\vec{x})$  la somme partielle d'ordre  $n$  de la série d'éléments de  $\mathbf{E}$ ,  $\sum c_n(\vec{x}) \vec{e}_n$

$$S_n(\vec{x}) = \sum_{k=0}^n c_k(\vec{x}) \vec{e}_k$$

Notons que d'après le 1.3.,

**Proposition 2.2.4.** — Pour tout entier naturel  $n$ ,  $S_n(\vec{x})$  est.....

Les résultats vus en 2.1. pour les polynômes de Legendre sont encore vrais :

**Proposition 2.2.5.** —

1. Pour tout entier naturel  $n$ ,  $\|S_n(\vec{x})\|^2 = \sum_{k=0}^n \dots$

2. Pour tout entier naturel  $n$ ,  $\|S_n(\vec{x})\|^2 \leq \|\vec{x}\|^2$ , inégalité de Bessel.

3. La série  $\sum c_n(\vec{x}) \vec{e}_n$  converge de somme  $\vec{x}$  dans l'espace vectoriel normé  $(\mathbf{E}, \|\cdot\|)$ ,

4. La série  $\sum \dots$  converge et

$$\|\vec{x}\|^2 = \sum_{n=0}^{+\infty} \dots, \text{égalité de Parseval.}$$

*Preuve 2.2.5.* — La preuve est identique à celle des résultats similaires dans le cas particulier 2.1.

Revenons sur la formule de Parseval. Dans l'exemple 2.2.3. on a muni  $\ell^2(\mathbf{N})$  d'un produit scalaire et donc d'une norme associée, notons la  $n_2$ . La formule de Parseval nous dit que l'on dispose de l'application

$$\Phi : \mathbf{E} \rightarrow \ell^2(\mathbf{N}); \vec{y} \mapsto (c_n(\vec{y}))_{n \in \mathbf{N}}$$

et qu'elle est une.....

En particulier :

**Proposition 2.2.5.** — FORMULE DE PARSEVAL POLAIRE —  
pour tout couple  $(\vec{x}_1, \vec{x}_2)$  d'éléments de  $\mathbf{E}$ ,

$$\langle \vec{x}_1 | \vec{x}_2 \rangle = \sum_{i=1}^n \dots\dots\dots$$

#### L'EXEMPLE DES SÉRIES DE FOURIER DES FONCTIONS PAIRES

Soit  $\mathcal{P}$  l'espace vectoriel des applications paires continues de  $\mathbf{R}$  dans  $\mathbf{R}$ ,  $2\pi$ -périodiques. On le munit du produit scalaire :

$$\mathcal{P}^2 \rightarrow \mathbf{R}; (f, g) \mapsto \frac{2}{\pi} \int_{[0, \pi]} fg,$$

noté  $\langle \cdot | \cdot \rangle$ .

La norme associée sera notée  $\| \cdot \|_2$ .

La suite  $(\cos(n \cdot))_{n \in \mathbf{N}}$  est, au premier terme près, orthonormée.

Montrons qu'elle est totale.

Soit  $f \in \mathcal{P}$

On note traditionnellement pour tout  $n \in \mathbf{N}$ ,  $a_n(f) = \langle f | \cos(n \cdot) \rangle$ . La série de fonctions

$$\frac{a_0(f)}{2} + \sum_{n \geq 1} a_n(f) \cos(n \cdot)$$

s'appelle série de Fourier de  $f$ .

La formule de Parseval s'écrit :

La série de Fourier de  $f$  converge dans  $\mathcal{P}$ , muni de  $\| \cdot \|_2$ .

Il ne faut pas croire que la série de Fourier converge ne serait-ce que simplement vers  $f$ . Par contre si  $f$  est de classe  $\mathcal{C}^1$  alors la série de Fourier de  $f$  converge normalement vers  $f$ .

Montrons ce résultat en admettant que pour une application impaire  $g$  la série  $\sum_{n \geq 1} b_n(g)^2$ , où  $b_n = \langle f | \sin(n \cdot) \rangle$ , converge. La preuve est similaire quoi que plus technique à celle pour la série  $\sum a_n(g)^2$ , (la suite  $(\sin(n \cdot))_{n \in \mathbf{N}^*}$  est aussi orthonormée totale dans l'espace vectoriel des applications impaires continues de  $\mathbf{R}$  dans  $\mathbf{R}$ ,  $2\pi$ -périodiques.

Le résultat demeure pour  $f$  continue et  $\mathcal{C}^1$  par morceaux, en effet....

**Application** — Donnons la valeur de la somme de la série  $\sum_{n \geq 1} \frac{1}{n^2}$ .

Soit  $g$  l'application de  $\mathbf{R}$  dans  $\mathbf{C}$ ,  $2\pi$ -périodique qui coïncide sur  $[-\pi, +\pi[$  avec l'application valeur absolue (fonction en dents de scie).

Elle est continue. L'application  $g$  est aussi  $\mathcal{C}^1$  par morceaux. En effet la restriction de  $f$  à  $]0, +\pi[$  se prolonge à  $[0, +\pi]$  en l'application identité,  $id$ , qui est de classe  $\mathcal{C}^1$ , et la restriction de  $f$  à  $] - \pi, 0[$  se prolonge à  $[-\pi, 0]$  en  $-id$  qui est aussi de classe  $\mathcal{C}^1$ ....







## Chapitre XIII

# ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES

Dans tout ce chapitre, nous désignerons par  $\mathbf{K}$  indifféremment le corps des nombres complexes ou celui des nombres réels.

## INTRODUCTION

Dans un premier paragraphe, nous donnerons des généralités sur les équations linéaires, équations différentielles ou non. Dans un second nous justifierons l'importance de l'étude des équations différentielles linéaires, que nous situerons dans le cadre plus général des équations différentielles ordinaires.

### 1.1 Equations linéaires

Rappelons des résultats élémentaires propres aux équations linéaires.

Soient  $\mathbf{E}$  et  $\mathbf{F}$  des  $\mathbf{K}$ -espaces vectoriels,  $\ell$  un élément de  $\mathcal{L}(\mathbf{E}, \mathbf{F})$  et  $\vec{b}$  un élément de  $\mathbf{F}$ . On étudie la résolution dans  $\mathbf{E}$  des équations, d'inconnue  $\vec{x}$ , suivantes :

$$\ell(\vec{x}) = \vec{b}. \quad (\text{e})$$

$$\ell(\vec{x}) = \vec{0}_{\mathbf{F}}. \quad (\text{h})$$

Ces équations sont dites *linéaires*. On dit que  $\vec{b}$  est le second membre de la première équation. La seconde, qui en est un cas particulier, est dite *homogène* ou *sans second membre*. On dit encore que (h) est l'*équation homogène associée* à (e).

Ce type d'équations, très général, recouvre des situations variées et diverses. Donnons des exemples déjà rencontrés en MPSI ou en MP.

**Exemples —**

#### 1. SYSTÈMES LINÉAIRES

$\mathbf{E} = \mathcal{M}_{n,1}(\mathbf{K})$ ,  $\mathbf{F} = \mathcal{M}_{p,1}(\mathbf{K})$ , avec  $n$  et  $p$  des entiers strictement positifs. Soit  $M$  un élément de  $\mathcal{M}_{p,n}(\mathbf{K})$ . on prend pour  $\ell$  l'application :

$$\ell : \mathcal{M}_{n,1}(\mathbf{K}) \rightarrow \mathcal{M}_{p,1}(\mathbf{K}); \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto M \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

et pour  $\vec{b}$  un élément  $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$  de  $\mathcal{M}_{p,1}(\mathbf{K})$ . Les équations (e) et (h), avec pour inconnue  $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ , prennent alors respectivement la forme

$$M \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

et

$$M \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Ces équations ont été étudiées en MPSI.

## 2. EQUATIONS AUX DÉRIVÉES PARTIELLES LINÉAIRES

On prend pour  $\mathbf{E}$ , l'espace  $\mathcal{C}^1(\mathbf{R}^2, \mathbf{R})$ , pour  $\mathbf{F}$ ,  $\mathcal{C}^0(\mathbf{R}^2, \mathbf{R})$ , pour  $\ell$ , l'application

$$\ell : \mathcal{C}^1(\mathbf{R}^2, \mathbf{R}) \rightarrow \mathcal{C}^0(\mathbf{R}^2, \mathbf{R}); f \mapsto 3\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial y},$$

où  $x$  désigne abusivement l'application  $\mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto x$ . Enfin on prend pour  $\vec{b}$  l'application  $\mathbf{R}^2 \rightarrow \mathbf{R}; (x, y) \mapsto x^2 + y^2$ . Les équations (e) et (h), avec pour inconnue  $f$  et avec quelques abus classiques de notations, prennent respectivement la forme

$$3\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial y} = 0$$

et

$$3\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial y} = x^2 + y^2.$$

Ces équations ont agréablement égayé le cours de calcul différentiel.

## 3. SUITES À RÉCURRENCE LINÉAIRE D'ORDRE 2

On note  $\mathcal{R}$  (resp.  $\mathcal{H}$ ) l'ensemble des suites réelles  $(u_n)_{n \in \mathbf{N}}$ , telles que, pour tout entier naturel  $n$ ,

$$u_{n+2} - 4u_{n+1} + 4u_n = 2^n + n, \text{ (resp. } u_{n+2} - 4u_{n+1} + 4u_n = 0).$$

Les ensemble  $\mathcal{R}$  (resp.  $\mathcal{H}$ ) est l'ensemble des solutions de l'équation (e) (resp. (h)), avec  $\mathbf{E} =$  ,  $\mathbf{F} =$  ,  $\vec{b} =$  et  $\ell$  l'application :

## 4. EQUATION DIFFÉRENTIEL LINÉAIRE DU PREMIER ORDRE

Soient  $I$  un intervalle d'intérieur non vide,  $f$  et  $g$  des éléments de  $\mathcal{C}^0(I, \mathbf{R})$ . On a étudié, en MPSI, les équations différentielles

$$\frac{dy}{dt} - f(t)y = g(t), \tag{3}$$

$$\frac{dy}{dt} - f(t)y = 0. \tag{4}$$

Les équations (3) et (4) sont respectivement des cas particuliers des équations (e) et (h), avec  $\mathbf{E} =$  ,  $\mathbf{F} =$  ,  $\vec{b} =$  et  $\ell$  l'application :

Revenons dans le cas général à l'étude de (e) et (h).

Remarquons que l'ensemble des solutions de (h), noté  $\mathcal{S}_h$ , est le noyau de  $\ell$ , c'est donc un espace vectoriel :

**Proposition 1.1.1.** — *L'ensemble  $\mathcal{S}_h$ , des solutions de l'équation (h), est un sous-espace vectoriel de  $\mathbf{F}$  (le noyau de  $\ell$ ). Il est donc en particulier non vide.*

Étudions maintenant l'ensemble des solutions de (e), noté  $\mathcal{S}_e$ .

Distinguons deux cas.

- PREMIER CAS :  $\vec{b} \notin \text{Im}(\ell)$

Alors clairement, l'équation linéaire avec second membre (e) n'admet pas de solution.

$$\boxed{\mathcal{S}_e = \emptyset}$$

- SECOND CAS :  $\vec{b} \in \text{Im}(\ell)$

Il existe par définition un élément  $\vec{x}_p$  de  $\mathbf{E}$  tel que :

$$\ell(\vec{x}_p) = \vec{b}. \quad (5)$$

Cet élément est une solution de (e). Soit alors  $\vec{u}$  un élément de  $\mathbf{E}$ . Cet élément est solution de (e) si par définition :

$$\ell(\vec{u}) = \vec{b}. \quad (6)$$

Donc en utilisant (5–6) et la linéarité de  $\ell$ ,  $\vec{u}$  est solution de (e) si et seulement si

$$\ell(\vec{u} - \vec{x}_p) = \vec{0}_{\mathbf{F}}.$$

Autrement dit  $\vec{u}$  est solution de (e) si et seulement si  $\vec{u} - \vec{x}_p$  est solution de (h). Donc :

$$\boxed{\mathcal{S}_e = \vec{x}_p + \mathcal{S}_h}.$$

Au siècle passé on aimait à traduire ce résultat en disant que *la solution générale de (e) est la somme d'une solution particulière de (e) et de la solution générale de (h)*.

Dans un langage moderne nous dirons que  $\mathcal{S}_e$  est un *espace affine dirigé par  $\mathcal{S}_h$* .

Récapitulons :

**Proposition 1.1.2.** — *L'ensemble des solutions de (e) est :*

- soit l'ensemble vide;
- soit un espace affine dirigé par l'espace vectoriel des solutions de (e).

Sans spécifier d'avantage l'équation, il n'est pas possible d'aller plus loin. Il est arrivé effectivement que  $\mathcal{S}_e$  soit vide, comme il a été vu en MPSI dans le cadre de l'exemple 1.

## 1.2 Linéarisation des équations différentielles

Il arrive que des problèmes physiques conduisent à des équations différentielles linéaires, citons l'exemple classique du circuit RLC. Toutefois les équations différentielles linéaires sont rares « dans la nature » ! Elles sont en général le fruit d'une idéalisation, voir d'une simplification audacieuse d'équations non linéaires. Expliquons à ce propos, comment on ramène l'étude d'une équation différentielle d'un certain type (équation autonome), à celle d'une équations différentielle linéaire.

Soient  $n$  un entier strictement positif et  $F$  une application de  $\mathcal{M}_{n,1}(\mathbf{K})$  dans  $\mathcal{M}_{n,1}(\mathbf{K})$ , de classe  $\mathcal{C}^1$ . On considère l'équation différentielle :

$$\frac{dX}{dt} = F(X). \quad (7)$$

Il s'agit là d'un type assez général d'équations, les équations autonomes, qui ont hélas disparues du programme. Donnons un exemple.

**Exemple 1.2.1.** — Prenons  $n = 2$ ,  $k$  un réel strictement positif et

$$F : \mathcal{M}_{2,1}(\mathbf{R}) \rightarrow \mathcal{M}_{2,1}(\mathbf{R}); \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ -\sin(x) - ky \end{pmatrix}.$$

Cette équation modélise, dans de bonnes unités, les variations de la mesure  $x$  de l'angle fait par un pendule simple avec la verticale au cours du temps  $t$ . Le terme  $-ky$  permet de tenir compte de frottements.

Revenons au cas général. supposons qu'il existe un élément  $X_0$  de  $\mathcal{M}_{n,1}(\mathbf{K})$ , tel que

$$F(X_0) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

On dit que  $X_0$  est une position d'équilibre de (7) ; en effet l'application constante

$$\mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{K}) ; t \mapsto X_0$$

est une solution sur  $\mathbf{R}$  de (7). 10 secondes de vérification... On pourrait montrer que c'est la seule.

Etudions la fonction  $F$  au voisinage de  $X_0$ .

$$F(X_0 + H) = F(X_0) + dF(X_0) \cdot H + o(H) = dF(X_0) \cdot H + o(H) \quad (H \rightarrow 0_{\mathcal{M}_{n,1}(\mathbf{K})}).$$

Linéariser l'équation (7), au voisinage de  $X_0$ , consiste à y remplacer  $F$  par son application linéaire tangente en  $X_0$  ; plus précisément après le changement d'inconnues  $H = X - X_0$ , on remplace (7) par :

$$\frac{dH}{dt} = dF(X_0) \cdot H. \quad (8)$$

On a en quelque sorte « négligé » le petit  $o$  dans l'expression de  $F$ .

Une telle équation est linéaire, en effet elle peut se mettre sous la forme (h), avec

$$\ell :$$

Ce type d'équation sera étudiée dans la suite.

**Exemples** — Reprenons l'exemple 1.2.1.

1. Considérons la position d'équilibre  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  (physiquement cette position d'équilibre correspond au pendule laissé sans vitesse initiale la « tête en bas »). L'équation linéarisée au voisinage de ce point est

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x - ky \end{pmatrix}.$$

On reconnaît l'équation classique de l'oscillateur harmonique amorti.

2. Considérons la position d'équilibre  $\begin{pmatrix} \pi \\ 0 \end{pmatrix}$  (physiquement cette position d'équilibre correspond au pendule laissé sans vitesse initiale la « tête en haut »). L'équation linéarisée au voisinage de ce point est

$$\frac{d}{dt} \begin{pmatrix} h \\ y \end{pmatrix} = \begin{pmatrix} y \\ h - ky \end{pmatrix}.$$

Se pose maintenant tout naturellement le rapport entre les solutions de l'équation (7) et celles de (8). Cette question est difficile et dépasse le cadre du programme. Nous nous bornerons à proposer, en exercice, la comparaison d'une équation et de son équation linéarisée. Disons que dans diverse situations le comportement des solutions (8) donne celui des solutions de (7). Dans d'autres ce n'est pas le cas.

Dans les sciences, le recours à la linéarisation est donc fréquente, les équations linéaires étant beaucoup plus simples à étudier. Nous allons donc jusqu'à la fin de ce chapitre traité de divers types d'équations linéaires.

## RAPPELS DE MPSI SUR LES ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES

Nous allons dans cette partie, rappeler brièvement, sans démonstrations les résultats principaux de sup. sur les équations différentielles linéaires. Nous les compléterons un peu et nous les illustrerons par des exemples. Le premier paragraphe est consacré à l'équation linéaire du premier ordre, le second à l'équation linéaire du second ordre à coefficients constants.

## 2.1 Équations linéaires scalaire du premier ordre

On désigne dans tout le paragraphe par  $I$  un intervalle d'intérieur non vide.

**Définition 2.1.1.** — Soient  $f$  et  $g$  des applications de  $I$  dans  $\mathbf{K}$ , continues.

— On appelle solution de l'équation

$$\frac{dy}{dt} = f(t)y + g(t), \quad (9)$$

(équation différentielle linéaires scalaire du premier ordre)

toute application  $\varphi$  d'un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ , à valeurs dans  $\mathbf{K}$ , dérivable, telle que :

$$\varphi' = f\varphi + g.$$

— Soient  $t_0$  un élément de  $I$  et  $y_0$  un élément de  $\mathbf{K}$ . On appelle solution du problème de Cauchy

$$\begin{cases} \frac{dy}{dt} = f(t)y + g(t), \\ y(t_0) = y_0, \end{cases}$$

toute solution  $\varphi$  de (9) définie sur un intervalle  $J$  contenant  $t_0$  et telle que  $\varphi(t_0) = y_0$ . On dit encore qu'une telle solution satisfait à la condition initiale  $y(t_0) = y_0$ .

— Dans le cas particulier où  $g$  est nulle l'équation qui devient

$$\frac{dy}{dt} = f(t)y, \quad (10)$$

est dite homogène ou sans second membre. On dit que l'équation (10) est l'équation homogène associée à (9).

— Les applications  $f$  et  $g$  sont appelées coefficients de l'équation (9);  $g$  est encore appelé second membre de l'équation.

**Remarques 2.1.2.** —

1. Lorsque l'on s'intéresse à des solutions sur un intervalle  $J$  donné, Le couple d'équations ((9),(10)) est un cas particulier du couple d'équations ((e),(h)), avec  $\mathbf{E} =$  ,  $\mathbf{F} =$  ,  $\vec{b} =$  et

$\ell :$

Ceci explique la terminologie choisie.

2. On peut montrer *a priori*, que toute solution de (9) est de classe  $\mathcal{C}^1$ .
3. On notera l'abus de notation, classique et préconisé par le programme, dans ces équations : la variable  $t$  ne figure pas dans l'inconnue mais dans les applications  $f$  et  $g$ . Cet abus n'est tolérable, que dans l'écriture de l'équation, jamais dans les égalités obtenues en y substituant une solution à l'inconnue.

### SOLUTIONS DE L'ÉQUATION HOMOGÈNE

L'ensemble des solutions de (10) sur un intervalle  $J$  fixé, et d'après le 1.1., un espace vectoriel. Nous pouvons ici être plus précis.

**Proposition 2.1.3.** — Soit  $J$  un intervalle inclus dans  $I$ , d'intérieur non vide ( $I$  par exemple).

1. L'ensemble des solutions de (10) définies sur  $J$ , noté  $\mathcal{S}_{10}(J)$ , est la droite vectorielle sur  $\mathbf{K}$  engendrée par

$$h_0 : J \rightarrow \mathbf{K}; t \mapsto \exp(F(t)),$$

où  $F$  est une primitive de  $f$  sur  $J$ .

2. Soient  $t_0$  un élément de  $J$  et  $y_0$  un élément de  $\mathbf{K}$ . Il existe une et une seule solution définie sur  $J$ , solution du problème de Cauchy

$$\begin{cases} \frac{dy}{dt} = f(t)y, \\ y(t_0) = y_0, \end{cases}$$

c'est :

$$J \rightarrow \mathbf{K}; t \mapsto y_0 \exp \left( \int_{t_0}^t f(s) ds \right).$$

## SOLUTIONS DE L'ÉQUATION AVEC SECOND MEMBRE

D'après l'étude du I.1. l'ensemble de solutions de (9), définies sur un intervalle  $J$  donné, est *a priori*, soit l'ensemble vide, soit un espace affine dirigé par  $\mathcal{S}_{10}(J)$ . Nous pouvons montrer que le premier cas ne se produit jamais. Pour montrer l'existence de solutions on utilise une méthode appelée hélas, *méthode de la variation de la constante*, terminologie paradoxale et malheureuse et oximore qui occulte ce dont il s'agit réellement : *la résolution complète de (9) par le changement d'inconnue* «  $Y = h_0 Z$  ». Plutôt que de refaire la théorie, rafraîchissons nous les idées par un exemple. Nous donnerons une rédaction précise mais concise, qui peut s'utiliser telle quelle dans un écrit de concours.

**Exemple 2.1.4.** — On prend  $\mathbf{K} = \mathbf{R}$ . Résolvons sur  $] -\frac{\pi}{2}, \frac{\pi}{2}[$ , l'équation :

$$\frac{dy}{dt} = \tan(t)y + \frac{t}{\cos(t)}. \quad (11)$$

- (Identification de l'équation)

C'est une équation scalaire, linéaire, du premier ordre, à coefficients continus.

- (Etude de l'équation homogène)<sup>1</sup>

Pour tout  $t \in ] -\frac{\pi}{2}, \frac{\pi}{2}[$ ,

$$\int_0^t \tan(s) ds = -\ln |\cos(t)| = \ln \left( \frac{1}{\cos(t)} \right).$$

Le cours de MPSI dit que l'ensemble des solutions sur  $] -\frac{\pi}{2}, \frac{\pi}{2}[$  de l'équation homogène associée à (11) est la droite vectorielle dirigée par

$$h_0 : ] -\frac{\pi}{2}, \frac{\pi}{2}[ \rightarrow \mathbf{R}; t \mapsto \frac{1}{\cos(t)}.$$

- (Résolution de l'équation avec second membre)

Soit  $\varphi$  une application<sup>2</sup> de  $] -\frac{\pi}{2}, \frac{\pi}{2}[$  dans  $\mathbf{R}$ , dérivable. Posons  $\psi := \varphi \cos$ ; de sorte que  $\varphi = \frac{\psi}{\cos}$ . L'application  $\psi$  est dérivable puisque les applications  $\varphi$  et cosinus le sont. l'application  $\varphi$  est solution de (11) si et seulement si, pour tout élément  $t$  de  $] -\frac{\pi}{2}, \frac{\pi}{2}[$ ,

$$\frac{\psi'(t)}{\cos(t)} = \frac{t}{\cos(t)}.$$

Donc  $\varphi$  est solution si et seulement il existe un réel  $c$ , tel que, pour tout élément  $t$  de  $] -\frac{\pi}{2}, \frac{\pi}{2}[$ ,

$$\psi(t) = \frac{t^2}{2} + c,$$

(une telle application est bien dérivable).

Donc l'ensemble des solutions de (11) définies sur  $] -\frac{\pi}{2}, \frac{\pi}{2}[$ , est l'ensemble des applications de la forme

$$\varphi_c : ] -\frac{\pi}{2}, \frac{\pi}{2}[ \rightarrow \mathbf{R}; t \mapsto \underbrace{\frac{t^2}{2\cos(t)}}_A + \underbrace{\frac{c}{\cos(t)}}_B,$$

où  $c$  est un quelconque réel.

(Remarquons que le terme A correspond à une solution particulière de (11),  $(\varphi_0)$ , le terme B à la solution générale de l'équation homogène associée.)

La méthode de résolution, baptisée méthode de la variation de la constante, permet aussi de montrer l'existence et l'unicité de la solution sur un intervalle donné d'un problème de Cauchy. Elle en explicite la forme, mais nous ne la rappellerons pas, d'une part parce que ce n'est pas une formule cruciale, d'autre part parce que nous en donnerons une forme plus générale dans la suite.

Récapitulons tout cela.

**Proposition 2.1.5.** — THÉORÈME DE CAUCHY-LIPSCHITZ LINÉAIRE —

Soit  $J$  un intervalle inclus dans  $I$ , d'intérieur non vide ( $I$  par exemple).

1. L'ensemble des solutions de (9), définies sur  $J$ , noté  $\mathcal{S}_9(J)$ , est, une droite affine dirigé par  $\mathcal{S}_{10}(J)$ .
2. Soient  $t_0$  un élément de  $J$  et  $y_0$  un élément de  $\mathbf{K}$ . Il existe une et une seule solution définie sur  $J$ , solution du problème de Cauchy

$$\begin{cases} \frac{dy}{dt} = f(t)y + g(t), \\ y(t_0) = y_0, \end{cases}$$

1. Cette étude est facultative, seule une solution de l'équation homogène est nécessaire.

2. Une *application* notez bien, et non une solution qui nous obligerait à une étude réciproque.

La méthode de variation des constante consiste en fait à abaisser le degré de l'équation par changement de variables : l'application  $\phi$  est solution de (9) si et seulement la dérivée de l'application  $\psi$  introduite est solution d'une équation différentielle de degré 0. Plus généralement — et on le verra — pour l'équation d'ordre 2 cette méthode permet de ramener la résolution d'une équation différentielle scalaire d'ordre quelconque à une équation d'ordre un de moins, dès lors qu'est connue une solution de l'équation homogène. Plus généralement la même technique peut être employer pour étudier des équations différentielles non linéaires qui ont une partie linéaire, ou, ce qui est un peu la même chose, pour étudier des inéquations différentielles. Donnons des exemples.

**Exercice 2.1.6.** —

1. Soit  $\phi$  un élément de  $\mathcal{C}^1(\mathbf{R}_+, \mathbf{R})$ . On suppose que pour tout réel  $t > 0$ ,

$$\phi'(t) + 2\phi = \frac{1}{1 + \phi^2(t)}. \quad (\text{XIII.1})$$

*Montrer que  $\phi$  est bornée.*

2. — LEMME DE GRONWALL —

*Soient  $t_0$  un réel,  $u_1$  un réel strictement positif et  $f$  une application de  $[t_0, +\infty[$  continue*

*Soit  $\phi_1$  la solution du problème de Cauchy :  $\begin{cases} \frac{dy}{dt} = f(t)y, \\ y(t_0) = u_1. \end{cases}$*

*Soit  $\phi$  une application  $[t_0, +\infty[$  dérivable telle que pour tout  $t \in [t_0, +\infty[$ ,  $\phi'(t) \leq f(t)\phi(t)$  et  $\phi(t_0) \leq u_1$ . Montrer que pour tout  $t \in [t_0, +\infty[$ ,  $\phi(t) \leq \phi_1(t)$ .*

3. Soit  $f$  une application de  $\mathbf{R}$  dans  $\mathbf{R}$ , dérivable. On suppose que l'application  $f' + 2019f$  admet 0 comme limite en  $+\infty$ . Montrer que  $f$  admet 0 comme limite en  $+\infty$ .

*Solution de l'exercice 2.1.6.* —



Signalons qu'il existe dans la nature, une version de la troisième question de cet exercice en termes de suites.

**Exercice** — Soit  $(u_n)_{n \in \mathbf{N}}$  une suite réelle. On suppose que la suite  $(u_{n+1} + \frac{1}{2019}u_n)_{n \in \mathbf{N}}$  converge vers 0. Montrer que  $(u_n)_{n \in \mathbf{N}}$  converge vers 0.

**Remarques** — Il peut arrivé que l'on connaisse une solution particulière de (9), parce qu'elle « saute au yeux », parce que l'énoncé de l'exercice la donne ou la suggère... On peut alors donner directement grâce à I.1.2. l'ensemble des solutions de (9).

#### CAS DES ÉQUATIONS NON RÉSOLUES

Pour finir disons quelques mots des équations linéaires « non résolues ». Soient  $a$ ,  $b$  et  $c$  des applications de  $I$  dans  $\mathbf{K}$  continues. On s'intéresse à la résolutions des équations

$$a(t)\frac{dy}{dt} + b(t)y + c(t) = 0, \quad (12)$$

et

$$a(t)\frac{dy}{dt} + b(t)y = 0; \quad (13)$$

c'est-à-dire à la détermination des applications  $\varphi$  d'un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ , à valeurs dans  $\mathbf{K}$ , dérivables, telles que respectivement  $a\varphi' + b\varphi + c = 0$  et  $a\varphi' + b\varphi = 0$ .

Fixons un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ . On s'intéresse aux solutions de (9) et (10) définies sur  $J$ . Le couple d'équations ((9),(10)) est alors un cas particulier du couple d'équations ((e),(h)), avec  $\mathbf{E} =$ ,  $\mathbf{F} =$ ,  $\vec{b} =$  et

$\ell :$

- PREMIER CAS : l'application  $a$  ne s'annule pas sur  $J$

L'étude des équations (12) et (13) se ramènent évidemment respectivement à celles de

$$\frac{dy}{dt} = -\frac{b(t)}{a(t)}y - \frac{c(t)}{a(t)} = 0, \quad (14)$$

et

$$\frac{dy}{dt} = -\frac{b(t)}{a(t)}y = 0. \quad (15)$$

Précisément, l'ensemble des solutions de (12) (resp.(13)) sur  $J$ , noté  $\mathcal{S}_{12}(J)$  (resp.  $\mathcal{S}_{13}(J)$ ) est l'ensemble des solutions définies sur  $J$  de (14) (resp. 15) En particulier :

$\mathcal{S}_{13}(J)$  est une droite vectorielle;  $\mathcal{S}_{12}(J)$  est une droite affine dirigé par  $\mathcal{S}_{13}(J)$ .

- SECOND CAS : l'application  $a$  s'annule sur  $J$

La situation est différente. D'après le I.1.1.,  $\mathcal{S}_{13}(J)$  est un espace vectoriel. Mais sa dimension n'est pas nécessairement 1, comme nous allons voir.  $\mathcal{S}_{12}(J)$  est, selon I.1.2., soit le vide, soit un espace affine dirigé par  $\mathcal{S}_{13}(J)$ .

On verra sur un exemple que  $\mathcal{S}_{12}(J)$  peut être vide.

Illustrons notre propos d'exemples

**Exemples 2.1.7.** — Nous allons étudier les solutions définie sur  $\mathbf{R}$  de l'équation

$$t\frac{dy}{dt} - ky = 0, \quad (16)$$

où  $k = 2$ ,  $k = 1$ , ou  $k = \frac{1}{2}$ . Nous prendrons ici  $\mathbf{K} = \mathbf{R}$ .

Notons  $I_1 := \mathbf{R}_-^*$  et  $I_2 := \mathbf{R}_+^*$ . Soit  $i \in \{1, 2\}$ .

1. Résolution de (16) sur l'intervalle  $I_i$ <sup>1</sup>

Il s'agit d'une équation différentielle linéaire du premier ordre à coefficients continus sur  $I_i$ , le coefficient de  $\frac{dy}{dt}$  ne s'annule pas, donc d'après le cours de MPSI, l'ensemble des solutions sur  $I_i$  est l'ensemble des applications de la formes :

$$I_i \rightarrow \mathbf{R}; t \mapsto c_i \exp(k \ln(|t|)) = c_i |t|^k,$$

avec  $c_i$  un réel quelconque.

2. Résolution de (16) sur l'intervalle  $\mathbf{R}$

1. Attention on ne parle de solution d'une équation différentielle que sur un *intervalle* .

- (Recherche de condition nécessaire par continuité).

Soit  $\varphi$  une solution de (16) définie sur  $\mathbf{R}^1$ . D'après le point précédent il existe des réels  $c_1$  et  $c_2$  tels que, pour tout réel  $t$ ,  $\varphi(t) = c_1|t|^k$ , pour  $t < 0$ ;  $\varphi(t) = c_2|t|^k$ , pour  $t > 0$ .

Or  $c_1|t|^k \xrightarrow[t \rightarrow 0^-]{} 0$ ;  $c_2|t|^k \xrightarrow[t \rightarrow 0^+]{} 0$ . L'application  $\varphi$  étant continue en 0, on en déduit que  $\varphi(0) = 0$ . Donc nécessairement :

$$\varphi : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \begin{cases} c_1|t|^k, & \text{pour } t < 0, \\ c_2|t|^k, & \text{pour } t \geq 0. \end{cases} \quad (17)$$

- (Etude réciproque).

Soit  $c_1$  et  $c_2$  des réels et  $\varphi_{c_1, c_2}$  l'application donnée par (17). Montrons que  $\varphi_{c_1, c_2}$  est dérivable. La restriction de  $\varphi_{c_1, c_2}$  aux intervalles  $I_1$  et  $I_2$  est dérivable ces intervalles étant *ouverts*,  $\varphi_{c_1, c_2}$  est dérivable en tout point de  $\mathbf{R}^*$ . Etudions la dérivabilité en 0.

— Cas :  $k = 2$

$$\frac{\varphi_{c_1, c_2}(t) - \varphi_{c_1, c_2}(0)}{t} = \frac{c_1 t^2 - 0}{t} \xrightarrow[t \rightarrow 0^-]{} 0.$$

Donc  $\varphi_{c_1, c_2}$  est dérivable à gauche en 0 de dérivée à gauche nulle. De même  $\varphi_{c_1, c_2}$  est-elle dérivable à droite en 0 de dérivée à droite nulle. Donc  $\varphi_{c_1, c_2}$  est dérivable en 0, et  $\varphi'_{c_1, c_2}(0) = 0$ .

Par ailleurs on a vu que pour tout élément  $t$  de  $\mathbf{R}_+^*$ ,  $t\varphi'_{c_1, c_2}(t) = k\varphi_{c_1, c_2}(t)$ ; on vérifie immédiatement que cette égalité est vraie pour  $t = 0$ . Donc  $\varphi'_{c_1, c_2}(t)$  est solution de (16)

Conclusion : l'ensemble  $\mathcal{S}_{16}(\mathbf{R})$  des solutions de (16), définies sur  $\mathbf{R}$ , vaut

$$\mathcal{S}_{16}(\mathbf{R}) = \{\varphi_{c_1, c_2}, c_1 \in \mathbf{R}, c_2 \in \mathbf{R}\}.$$

On observe que pour tout réel  $c_1$  et tout réel  $c_2$ ,  $\varphi_{c_1, c_2} = c_1\varphi_{1,0} + c_2\varphi_{0,1}$ ; donc  $\mathcal{S}_{16} = \text{vect}(\varphi_{1,0}, \varphi_{0,1})$ . Par ailleurs la liberté de la famille  $\mathcal{S}_{16} = \text{vect}(\varphi_{1,0}, \varphi_{0,1})$  se prouve sans mal, donc  $\mathcal{S}_{16}$  est *un espace vectoriel de dimension 2*.

— Cas :  $k = 1$

Une étude élémentaire montre que  $\varphi_{c_1, c_2}$  est dérivable en 0, si et seulement si  $c_2 = -c_1$ . Par ailleurs pour tout élément  $c_1$  de  $\mathbf{R}$ ,  $\varphi_{-c_1, c_1}$  qui vaut  $c_1 \text{id}_{\mathbf{R}}$  est solution de (16).

Conclusion :

$$\mathcal{S}_{16}(\mathbf{R}) = \{c_1 \text{id}_{\mathbf{R}}, c_1 \in \mathbf{R}\}.$$

$\mathcal{S}_{16}(\mathbf{R})$  est donc *la droite vectorielle engendrée par  $\text{id}_{\mathbf{R}}$* .

— Cas :  $k = \frac{1}{2}$

$$\frac{\varphi_{c_1, c_2}(t) - \varphi_{c_1, c_2}(0)}{t} = \frac{c_1(-t)^{1/2} - 0}{t} \xrightarrow[t \rightarrow 0^-]{} \begin{cases} -\infty, & \text{pour } c_1 > 0, \\ 0, & \text{pour } c_1 = 0, \\ +\infty, & \text{pour } c_1 < 0. \end{cases}$$

Donc si  $\varphi_{c_1, c_2}$  est dérivable en 0 alors  $c_1 = 0$  et de même  $c_2 = 0$ , autrement dit  $\varphi_{c_1, c_2} = 0$ . Réciproquement, l'application nulle sur  $\mathbf{R}$ ,  $0_{\mathbf{R}}$ , est évidemment solution sur  $\mathbf{R}$  de (16).

Conclusion :

$$\mathcal{S}_{16}(\mathbf{R}) = \{0_{\mathbf{R}}\},$$

*espace vectoriel de dimension 0.*

**Exemples 2.1.8.** — Nous allons étudier les solutions éventuelles, définies sur  $\mathbf{R}$ , de l'équation

$$t^2 \frac{dy}{dt} - ty - 1 = 0. \quad (18)$$

Nous prendrons ici  $\mathbf{K} = \mathbf{R}$ . Notons  $I_1 := \mathbf{R}_-^*$  et  $I_2 := \mathbf{R}_+^*$ . Soit  $i \in \{1, 2\}$ .

1. Résolution de (18) sur l'intervalle  $I_i$

- (Identification)

Il s'agit d'une équation différentielle linéaire du premier ordre à coefficients continus sur  $I_i$ , le coefficient de  $\frac{dy}{dt}$  ne s'annule pas.

- (Résolution de l'équation homogène associées)

D'après le cours de MPSI, l'ensemble des solutions sur  $I_i$ , de l'équation homogène associée à (18) est la droite engendrée par

$$I_i \rightarrow \mathbf{R}; t \mapsto \exp(\ln(|t|)) = |t|,$$

ou encore, plus simplement, par  $h_0 = \text{id}_{I_i}$

1. Il y en a toujours, l'ensemble des solutions sur  $\mathbf{R}$  est, on l'a dit, un espace vectoriel !

- (*Méthode dite de variation de la constante*)

Soit  $\varphi$  une application de  $I_i$  dans  $\mathbf{R}$  dérivable. posons  $\psi := \frac{\varphi}{h_0}$ . Ainsi :  $\varphi = h_0\psi$ . L'application  $\psi$ , quotient de deux applications dérivables, l'est et  $\varphi$  est solution sur  $I_i$  de (18) si et seulement si  $\psi'(t) = \frac{1}{t^3}$ , pour tout élément  $t$  de  $I_i$ . Donc  $\varphi$  est solution sur  $I_i$  de (18) si et seulement si, il existe un réel  $c_i$  tel que :

$$\varphi(t) = t(c_i \cdots \quad),$$

pour tout élément  $t$  de  $I_i$ .

2. *Résolution de (18) sur  $\mathbf{R}$*

Soit  $\varphi$  une éventuelle<sup>1</sup> solution de (18) sur  $\mathbf{R}$ .

---

1. On n'est pas sûr qu'il en existe.

Donc l'ensemble des solutions de (18), définies sur  $\mathbf{R}$  est donc ...

## 2.2 Equations linéaires scalaires du second ordre à coefficient constant

**Définition 2.2.1.** — Soient  $b, c$  des réels,  $a$  un réel non nul et  $g$  une application de  $I$  dans  $\mathbf{K}$ , continue.

— On appelle solution de l'équation

$$a \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + cy = g(t), \quad (19)$$

(équation différentielle linéaire scalaire du second ordre à coefficients constants)  
toute application  $\varphi$  d'un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ , à valeurs dans  $\mathbf{K}$ , deux fois dérivable, telle que :

$$a\varphi'' + b\varphi' + c\varphi = g.$$

— Soient  $t_0$  un élément de  $I$ ,  $y_0$  et  $y'_0$  des éléments de  $\mathbf{K}$ . On appelle solution du problème de Cauchy

$$\begin{cases} a \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + cy = g(t), \\ y(t_0) = y_0, \\ \frac{dy}{dt}(t_0) = y'_0, \end{cases}$$

toute solution  $\varphi$  de (19) définie sur un intervalle  $J$  contenant  $t_0$  et telle que  $\varphi(t_0) = y_0$  et  $\varphi'(t_0) = y'_0$ . On dit encore qu'une telle solution satisfait à la condition initiale  $y(t_0) = y_0$ ,  $y'(t_0) = y'_0$ .

— Dans le cas particulier où  $g$  est l'application nulle, l'équation qui devient

$$a \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + cy = 0, \quad (20)$$

est dite homogène ou sans second membre. On dit que l'équation (10) est l'équation homogène associée à (9). Dans le cas de l'équation homogène, on peut prendre  $I = \mathbf{R}$ .

— Les réels  $a, b, c$  et parfois l'application  $g$  sont appelés coefficients de l'équation (19); l'application  $g$  est encore appelé second membre.

**Remarque 2.2.2.** —

1. Lorsque l'on s'intéresse à des solutions sur un intervalle  $J$  donné, Le couple d'équations ((19),(20)) est un cas particulier du couple d'équations ((e),(h)), avec  $\mathbf{E} = \mathcal{D}^2(J, \mathbf{R})$ ,  $\mathbf{F} = \mathcal{F}(J, \mathbf{R})$   $\vec{b} = g$  et

$$\ell : J \rightarrow \mathbf{R}; \varphi \mapsto a\varphi'' + b\varphi' + c\varphi.$$

Ceci explique la terminologie choisie.

2. On peut montrer *a priori*, que toute solution de (19) est de classe  $\mathcal{C}^2$ .
3. On notera l'abus de notation, classique et préconisé par le programme, dans ces équations : la variable  $t$  ne figure pas dans l'inconnue mais dans l'application  $g$ . Cet abus n'est tolérable, que dans l'écriture de l'équation, jamais dans les égalités obtenues en y substituant une solution à l'inconnue.
4. On a pris  $a \neq 0$ . Pour  $a = 0$ , l'équation (20) deviendrait une équation scalaire du premier ordre à coefficients constants, étudiée en terminale, (20) serait, elle, un cas particulier de l'équation scalaire du premier ordre étudiée dans le paragraphe précédent.

## SOLUTIONS DE L'ÉQUATION HOMOGÈNE

L'ensemble des solutions de (20) sur un intervalle  $J$  fixé, et d'après le I.1., un espace vectoriel. Nous pouvons ici être plus précis.

**Proposition 2.2.3.** — *Soit  $J$  un intervalle inclus dans  $\mathbf{R}$ , d'intérieur non vide ( $\mathbf{R}$  par exemple). Alors l'ensemble des solutions de (20) définies sur  $J$ , noté  $\mathcal{S}_{20}(J)$ , est un plan vectoriel.*

Le programme de MPSI va plus loin et précise une base de  $\mathcal{S}_{20}(J)$ . Pour cela on introduit l'équation du second degré d'inconnue  $x$ ,

$$ax^2 + bx + c = 0 \quad (21)$$

appelée équation caractéristique de (20). Nous distinguerons plusieurs cas :

- **PREMIER CAS :  $\mathbf{K} = \mathbf{C}$**

Notons  $\lambda_1$  et  $\lambda_2$  les solutions complexes de (21).

— *Premier sous-cas :  $\lambda_1 \neq \lambda_2$*

Alors  $\mathcal{S}_{20}(J)$  admet comme base  $(\varphi_1, \varphi_2)$  avec :

$$\varphi_1 : J \rightarrow \mathbf{C} ; t \mapsto \exp(\lambda_1 t) ; \varphi_2 : J \rightarrow \mathbf{C} ; t \mapsto \exp(\lambda_2 t).$$

Ainsi  $\mathcal{S}_{20}(J)$  est-il l'ensemble des applications de la forme :

$$J \rightarrow \mathbf{C} ; t \mapsto c_1 \exp(\lambda_1 t) + c_2 \exp(\lambda_2 t),$$

avec  $c_1$  et  $c_2$  des complexes quelconques.

— *Second sous-cas :  $\lambda_1 = \lambda_2$*

Alors  $\mathcal{S}_{20}(J)$  admet comme base  $(\varphi_1, \varphi_2)$  avec :

$$\varphi_1 : J \rightarrow \mathbf{C} ; t \mapsto \exp(\lambda_1 t) ; \varphi_2 : J \rightarrow \mathbf{C} ; t \mapsto t \exp(\lambda_1 t).$$

Ainsi  $\mathcal{S}_{20}(J)$  est-il l'ensemble des applications de la forme :

$$J \rightarrow \mathbf{C} ; t \mapsto (c_1 + c_2 t) \exp(\lambda_1 t),$$

avec  $c_1$  et  $c_2$  des complexes quelconques.

- **SECOND CAS :  $\mathbf{K} = \mathbf{R}$**

Notons  $\Delta$  le discriminant de l'équation (21).

— *Premier sous-cas :  $\Delta > 0$*

L'équation (21) admet deux solutions réelles distinctes  $\lambda_1$  et  $\lambda_2$ . Alors  $\mathcal{S}_{20}(J)$  admet comme base  $(\varphi_1, \varphi_2)$  avec :

$$\varphi_1 : J \rightarrow \mathbf{R} ; t \mapsto \exp(\lambda_1 t) ; \varphi_2 : J \rightarrow \mathbf{R} ; t \mapsto \exp(\lambda_2 t).$$

Ainsi  $\mathcal{S}_{20}(J)$  est-il l'ensemble des applications de la forme :

$$J \rightarrow \mathbf{R} ; t \mapsto c_1 \exp(\lambda_1 t) + c_2 \exp(\lambda_2 t),$$

avec  $c_1$  et  $c_2$  des réels quelconques.

— *Deuxième sous-cas :  $\Delta = 0$*

L'équation (21) admet une unique solution complexe,  $\lambda_1$ , qui est réelle. Alors  $\mathcal{S}_{20}(J)$  admet comme base  $(\varphi_1, \varphi_2)$  avec :

$$\varphi_1 : J \rightarrow \mathbf{R} ; t \mapsto \exp(\lambda_1 t) ; \varphi_2 : J \rightarrow \mathbf{R} ; t \mapsto t \exp(\lambda_1 t).$$

Ainsi  $\mathcal{S}_{20}(J)$  est-il l'ensemble des applications de la forme :

$$J \rightarrow \mathbf{R} ; t \mapsto (c_1 + c_2 t) \exp(\lambda_1 t),$$

avec  $c_1$  et  $c_2$  des réels quelconques.

— *Troisième sous-cas :  $\Delta < 0$*

L'équation (21) admet deux solutions complexes, non réelles et conjuguées,  $\lambda_1$  et  $\lambda_2$ . On notera  $\alpha := \operatorname{Re}(\lambda_1)$  et  $\beta := \operatorname{Im}(\lambda_1)$ , ainsi a-t-on :  $\lambda_1 = \alpha + i\beta$ ,  $\lambda_2 = \alpha - i\beta$ . Alors  $\mathcal{S}_{20}(J)$  admet comme base  $(\varphi_1, \varphi_2)$  avec :

$$\varphi_1 : J \rightarrow \mathbf{R} ; t \mapsto \exp(\alpha t) \cos(\beta t) ; \varphi_2 : J \rightarrow \mathbf{R} ; t \mapsto \exp(\alpha t) \sin(\beta t).$$

Ainsi  $\mathcal{S}_{20}(J)$  est-il l'ensemble des applications de la forme :

$$J \rightarrow \mathbf{R} ; t \mapsto \exp(\alpha t) (c_1 \cos(\beta t) + c_2 \sin(\beta t)),$$

avec  $c_1$  et  $c_2$  des réels quelconques.

Chaque solution de (20) est déterminée par deux éléments de  $\mathbf{K}$ , ses coordonnées dans la base  $(\varphi_1, \varphi_2)$ . Il en résulte que les deux égalités d'une condition initiale déterminent entièrement la solution, plus précisément :

**Proposition 2.2.4.** — THÉORÈME DE CAUCHY-LIPSCHITZ LINÉAIRE —

Soient  $J$  un intervalle d'intérieur non vide ( $\mathbf{R}$  par exemple),  $t_0$  un élément de  $J$ ,  $y_0$  et  $y'_0$  des éléments de  $\mathbf{K}$ . Le problème de Cauchy

$$\begin{cases} a \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + cy = 0, \\ y(t_0) = y_0, \\ \frac{dy}{dt}(t_0) = y'_0, \end{cases}$$

admet une et une seule solution définie sur  $J$ .

**Exemples** — Déterminer la solution définie sur  $\mathbf{R}$  des problèmes de Cauchy suivant :

$$\begin{cases} \frac{d^2 y}{dt^2} + \frac{dy}{dt} + y = 0, \\ y(0) = 1, \\ \frac{dy}{dt}(0) = 0. \end{cases} \quad \text{et} \quad \begin{cases} \frac{d^2 y}{dt^2} - y = 0, \\ y(0) = 1, \\ \frac{dy}{dt}(0) = 0. \end{cases}$$

## SOLUTIONS DE L'ÉQUATION AVEC SECOND MEMBRE

L'étude générale de l'équation (20) avec second membre quelconque n'est pas au programme de MPSI. Elle sera donnée, dans un cadre plus général, en fin de chapitre. En revanche, le cours de sup. prévoit l'étude de cette équation lorsque le second membre y prend une forme particulière, à savoir :

$$a \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + cy = e \exp(\lambda t), \quad (22)$$

avec  $\lambda$  et  $e$  des éléments de  $\mathbf{K}$ . Nous étendrons généreusement ces résultats aux équations du type :

$$a \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + cy = p(t) \exp(\lambda t), \quad (23)$$

avec  $\lambda$  un élément de  $\mathbf{K}$  et  $p$  une application polynômiale.

D'après I.1.2., il suffit de déterminer une solution particulière  $\varphi_p$  de (23) pour obtenir l'ensemble des solutions de cette équation sur  $J$ , noté  $\mathcal{S}_{23}(J)$  :

$$\mathcal{S}_{23}(J) = \varphi_p + \mathcal{S}_{20}(J)$$

**Proposition 2.2.5.** —  $J$  un intervalle d'intérieur non vide ( $\mathbf{R}$  par exemple). Il existe une (et même une seule) solution sur  $J$  de (20) la forme :

$$J \rightarrow \mathbf{K}; t \mapsto t^m q(t) \exp(\lambda t),$$

où  $q$  est une application polynômiale de même degré que  $p$ ,  $m$  l'ordre de multiplicité de  $\lambda$  comme solution de l'équation caractéristique (21) ( si  $\lambda$  n'est pas racine de  $ax^2 + bx + c$ , alors  $m = 0$ , si  $\lambda$  est racine simple de  $ax^2 + bx + c$ , alors  $m = 1$ , enfin si  $\lambda$  est racine double, alors  $m = 2$ .)

**Remarque.** — La proposition 2.2.5. s'applique au cas d'une équation différentielle linéaire d'ordre 1 à coefficient constant.

Avant d'illustrer de quelques exemples, rappelons le principe si évident, mais oh combien utile, de superposition des solutions, principe qui permet de se ramener à l'étude de seconds membres « élémentaires ». Nous donnerons un énoncé abstrait dans le cadre de l'équation linéaire la plus générale.

**Proposition 2.2.6.** — SUPERPOSITION DES SOLUTIONS —

Soient  $\mathbf{E}$  et  $\mathbf{F}$  des  $\mathbf{K}$ -espaces vectoriels,  $\ell$  un élément de  $\mathcal{L}(\mathbf{E}, \mathbf{F})$ ,  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n$  des éléments de  $\mathbf{F}$  et  $\alpha_1, \alpha_2, \dots, \alpha_n$  des éléments de  $\mathbf{K}$ . Pour  $i = 1, 2, \dots, n$ , on suppose disposer, d'une solution  $\vec{x}_i$  de l'équation

$$\ell(\vec{x}) = \vec{b}_i.$$

Alors,  $\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n$  est solution de

$$\ell(\vec{x}) = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \dots + \alpha_n \vec{b}_n.$$

Appliquons cela sur le champ !

**Application** —

1. Résolvons sur  $\mathbf{R}$  l'équation différentielle

$$\frac{d^2 y}{dt^2} - 4 \frac{dy}{dt} + 4y = 1515t + 814t \exp(2t). \quad (24)$$

- (Identification de l'équation)  
C'est une équation différentielle linéaire scalaire du second ordre à coefficients constants avec un second membre continue.
- SOLUTION DE L'ÉQUATION HOMOGENE ASSOCIÉE  
l'équation caractéristique,  $x^2 - 4x + 4 = 0$  admet une solution double 2. Les solutions sur  $\mathbf{R}$  de l'équation homogène associée à (24) sont donc les applications de la forme :

$$\mathbf{R} \rightarrow \mathbf{R}; t \mapsto (c_1 + c_2 t) \exp(2t),$$

avec  $c_1$  et  $c_2$  des réels quelconques.

- RECHERCHE D'UNE SOLUTION PARTICULIÈRE DE (24)

— *Solution particulière de*  $\frac{d^2 y}{dt^2} - 4 \frac{dy}{dt} + 4y = t \exp(0t)$ .

Il existe, d'après le cours de MPSI, une unique solution sur  $\mathbf{R}$  de  $\varphi_{p,1}$  de cette équation de la forme :

$$\varphi_{p,1} : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

Après calcul on trouve :

— *Solution particulière de*  $\frac{d^2 y}{dt^2} - 4 \frac{dy}{dt} + 4y = t \exp(2t)$ .

Il existe, d'après le cours de MPSI, une unique solution sur  $\mathbf{R}$  de  $\varphi_{p,2}$  de cette équation de la forme :

$$\varphi_{p,2} : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

Après calcul on trouve :

— D'où une solution particulière de (24) :

$$\varphi_p : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

- CONCLUSION les solutions sur  $\mathbf{R}$  de l'équation (24) sont donc les applications de la forme :

$$\boxed{\mathbf{R} \rightarrow \mathbf{R}; t \mapsto (c_1 + c_2 t) \exp(2t) + \varphi_p(t)},$$

avec  $c_1$  et  $c_2$  des réels quelconques.

2. Résolvons sur  $\mathbf{R}$  l'équation différentielle

$$\frac{d^2 y}{dt^2} + y = \sin(t) + \operatorname{ch}(t). \quad (25)$$

- (Identification de l'équation)  
C'est une équation différentielle linéaire scalaire du second ordre à coefficients constants avec un second membre continue.

- SOLUTION DE L'ÉQUATION HOMOGÈNE ASSOCIÉE

Les solutions sur  $\mathbf{R}$  de l'équation homogène associée à (25) sont donc les applications de la forme :

$$\mathbf{R} \rightarrow \mathbf{R}; t \mapsto c_1 \cos(t) + c_2 \sin(t),$$

avec  $c_1$  et  $c_2$  des réels quelconques<sup>1</sup>.

- RECHERCHE D'UNE SOLUTION PARTICULIÈRE DE (25)

— *Solution particulière de  $\frac{d^2 y}{dt^2} + y = \exp(it)$*

Il existe, d'après le cours de MPSI, une unique solution sur  $\mathbf{R}$ , complexe de  $\varphi_{p,1}$  de cette équation de la forme :

$$\varphi_{p,1} : \mathbf{R} \rightarrow \mathbf{C}; t \mapsto \dots$$

Après calcul on trouve :

— *Solution particulière de  $\frac{d^2 y}{dt^2} + y = \sin(t)$*

$$\varphi_{p,2} : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

— *Solution particulière de  $\frac{d^2 y}{dt^2} + y = \exp(t)$*

Il existe, d'après le cours de MPSI, une unique solution sur  $\mathbf{R}$ , complexe de  $\varphi_{p,3}$  de cette équation de la forme :

$$\varphi_{p,3} : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

Après calcul on trouve :

— *Solution particulière de  $\frac{d^2 y}{dt^2} + y = \operatorname{ch}(t)$*

$$\varphi_{p,4} : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

— D'où une solution particulière de (25) :

$$\varphi_p : \mathbf{R} \rightarrow \mathbf{R}; t \mapsto \dots$$

- CONCLUSION les solutions sur  $\mathbf{R}$  de l'équation (25) sont donc les applications de la forme :

$$\boxed{\mathbf{R} \rightarrow \mathbf{R}; t \mapsto c_1 \cos(t) + c_2 \sin(t) + \dots},$$

avec  $c_1$  et  $c_2$  des réels quelconques.

### Remarques —

1. Les résultats donnés pour les second membres de la forme  $p(t) \exp(\lambda t)$ ,  $p$  polynomial sont aussi valables pour l'équation scalaire du premier ordre à coefficient constant étudiée en terminale. Ils seront encore valables dans le cas de l'équation scalaire à coefficients constants d'ordre quelconque, dont nous dirons quelques mots dans la suite.

2. Le point précédent permet le calcul des primitives du type  $\int p(t) \exp(\lambda t) dt$ , avec  $p$  polynomiale. En effet, il s'agit

de résoudre l'équation différentielle  $\frac{dy}{dt} + 0y = p(t) \exp(\lambda t)$ . Donnons des exemples. Calculons les primitives

$$\int t^2 \exp(3t) dt, \int \sin(t) \exp(3t) dt, \dots$$

1. Pour cette équation (oscillateur harmonique), il serait tout à fait inconvenant que l'on recourût à l'équation caractéristique (cf. cours de terminale).



3. Il existe des résultats analogues pour les suites vérifiant des relations de récurrence linéaires. Déterminons les suite réelles  $(u_n)_{n \in \mathbf{N}}$  telles que pour tout entier naturel  $n$ ,

$$u_{n+2} - 4u_{n+1} + 4u_n = 1515n + 814.n2^n$$

## ÉQUATION LINÉAIRES VECTORIELLES DU PREMIER ORDRE À COEFFICIENT CONSTANT

Nous allons étudier les équations linéaires du premier ordre. Nous nous intéresserons à des équations dont les solutions, contrairement à celles du paragraphe précédent, sont à valeurs vectorielles. Nous nous limiterons dans cette partie au cas où le coefficient est constant, réservant le cas général à la partie suivante. Une bonne partie du travail a déjà été faite dans les chapitres sur les séries et séries de fonctions et le cours sur la réduction des endomorphismes.

Le premier paragraphe donnera des généralités et fixera la terminologie, le deuxième traitera du cas sans second membre, d'un point de vue, d'abord théorique, puis pratique, le dernier étudiera le cas de l'équation avec second membre, généralisant la méthode de variation de la constante, rappelée en II.1.

Dans toute cette partie  $\mathbf{K}$  désigne indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$ ,  $\mathbf{E}$  est un espace vectoriel sur  $\mathbf{K}$  de dimension finie, non nulle  $n$  et  $I$  un intervalle d'intérieur non vide. Enfin  $a$  désigne un élément de  $\mathcal{L}(\mathbf{E})$ .

### 3.1 Définitions, généralités

**Définition 3.1.1.** — Soit  $\vec{b}$  une applications de  $I$  dans  $\mathbf{E}$ , continue.

— On appelle solution de l'équation

$$\frac{d\vec{y}}{dt} = a(\vec{y}) + \vec{b}(t), \quad (26)$$

(équation différentielle linéaires vectorielles du premier ordre à coefficient constant)

toute application  $\vec{\varphi}$  d'un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ , à valeurs dans  $\mathbf{E}$ , dérivable, telle que :

$$\vec{\varphi}' = a \circ \vec{\varphi} + \vec{b}.$$

— Soient  $t_0$  un élément de  $I$  et  $\vec{y}_0$  un élément de  $\mathbf{E}$ . On appelle solution du problème de Cauchy

$$\begin{cases} \frac{d\vec{y}}{dt} = a \cdot \vec{y} + \vec{b}(t), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$$

toute solution  $\vec{\varphi}$  de (26) définie sur un intervalle  $J$  contenant  $t_0$  et telle que  $\vec{\varphi}(t_0) = \vec{y}_0$ . On dit encore qu'une telle solution satisfait à la condition initiale  $\vec{y}(t_0) = \vec{y}_0$ .

— Dans le cas particulier où  $\vec{b}$  est nulle et où donc  $I = \mathbf{R}$ , l'équation (26) qui devient

$$\frac{d\vec{y}}{dt} = a(\vec{y}), \quad (27)$$

est dite homogène ou sans second membre. On dit que l'équation (27) est l'équation homogène associée à (26).

— L'application  $a$  et parfois  $\vec{b}$ , sont appelées coefficients de l'équation (9);  $\vec{b}(t)$  est encore appelée second membre de l'équation.

**Remarques 3.1.2.** —

1. Lorsque l'on s'intéresse à des solutions sur un intervalle  $J$  donné, Le couple d'équations ((26),(27)) est un cas particulier du couple d'équations ((e),(h)), avec pour  $\mathbf{E}$ ,  
pour  $\mathbf{F}$ , , et

$\ell$  :

Ceci explique la terminologie choisie.

2. Si  $b$  est de classe  $\mathcal{C}^\infty$ , alors on peut montrer *a priori*, par récurrence, que toute solution de (26) est de classe  $\mathcal{C}^\infty$  (c'est notamment le cas de l'équation (27)).
3. On notera l'abus de notations, classique et préconisé par le programme, dans ces équations : la variable  $t$  ne figure pas dans l'inconnue mais dans l'application  $\vec{b}$ . Cet abus n'est tolérable, que dans l'écriture de l'équation, jamais dans les égalités obtenues en y substituant une solution à l'inconnue.
4. Il arrive que dans (26), l'image par  $a$  d'un élément  $\vec{x}$  de  $\mathbf{E}$  soit noté par  $a \cdot \vec{x}$  ou  $a\vec{x}$ , cette notation, dans le contexte des équations différentielles, est fréquente, elle renforce l'analogie entre le cas vectoriel et le cas scalaire et est à rapprocher de la notation de la différentielle. Le programme officiel vient d'y renoncer mais ça et là peut trainer quelques vestiges tenaces de cette coutume.

EXPRESSION MATRICIELLE

Soit  $\mathcal{B}$  une base de  $\mathbf{E}$ . Soient  $(a_{i,j})_{\substack{i=1\dots n \\ j=1\dots n}}$  la matrice de  $a$  dans  $\mathcal{B}$ , que l'on notera encore  $A$  et  $(b_1, b_2, \dots, b_n)$ , le vecteur

colonne des coordonnées de  $\vec{b}$  dans  $\mathcal{B}$ , noté plus simplement  $B$ . Soit  $J$  un intervalle inclus dans  $I$ , d'intérieur non vide et  $\vec{\varphi}$  une application de  $J$  dans  $\mathbf{E}$ , dérivable. Notons  $\Phi$  le vecteur colonne des composantes de  $\vec{\varphi}$  dans  $\mathcal{B}$  et pour  $i = 1, 2, \dots, n$ ,  $\varphi_i$ , la  $i^{\text{e}}$  composante de  $\varphi$  dans  $\mathcal{B}$ .

D'après le cours, sur la représentation matricielle des endomorphismes, L'application  $\vec{\varphi}$  est solution de (26) si et seulement si,  $\Phi' = A\Phi + B$ . Si c'est le cas, on dit que  $\Phi$  est solution du *système différentiel linéaire*

$$\frac{dX}{dt} = AX + B(t). \quad (28)$$

En explicitant ce qui précède,  $\vec{\varphi}$  est solution de (26) si et seulement si :

$$\varphi'_i = \sum_{j=1}^n a_{i,j} \varphi_j + b_i, \text{ pour } i = 1, 2, \dots, n.$$

Si c'est le cas on dit que  $(\varphi_1, \varphi_2, \dots, \varphi_n)$  est solution du *système d'équations différentielles linéaires*

$$\begin{cases} \frac{dx_1}{dt} = \sum_{j=1}^n a_{1,j} x_j + b_1(t) \\ \frac{dx_2}{dt} = \sum_{j=1}^n a_{2,j} x_j + b_2(t) \\ \vdots \\ \frac{dx_n}{dt} = \sum_{j=1}^n a_{n,j} x_j + b_n(t) \end{cases} \quad (29)$$

Dans les études théoriques on raisonnera sur l'équation (26). Le système (28), qui est en fait un cas particulier du précédent, où  $\mathbf{E} = \mathcal{M}_{n,1}(\mathbf{K})$ , lui sera préféré dans les exercices pratiques. Nous énoncerons les résultats pour (26), à charge pour le lecteur de les traduire en termes matriciels pour (28).

## 3.2 Étude de l'équation homogène

L'équation (26)

$$\frac{d\vec{y}}{dt} = a(\vec{y})$$

a été étudiée, d'un point de vue théorique dans le chapitre sur les séries de fonctions. Soit  $J$  un intervalle d'intérieur non vide, que l'on peut prendre égal à  $\mathbf{R}$ . On y a vu que l'ensemble des solutions de (26) sur  $J$  est l'ensemble des applications de la forme

$$J \rightarrow \mathbf{E}; t \mapsto \exp(ta)(\vec{c}),$$

où  $\vec{c}$  est un quelconque vecteur de  $\mathbf{E}$ . Ce résultat, rappelons le reposait essentiellement sur le lemme suivant :

**Lemme 3.2.1.** — *L'application  $\chi_a : \mathbf{R} \rightarrow \mathcal{L}(\mathbf{E}) ; t \mapsto \exp(ta)$  est dérivable et  $\chi'_a = a \circ \chi_a$ .*

Comme  $\exp(0a) = \exp(0_{\mathcal{L}(\mathbf{E})}) = \text{id}_{\mathbf{E}}$ , pour tout élément  $\vec{y}_0$  de  $\mathbf{E}$ , le problème de Cauchy

$$\begin{cases} \frac{d\vec{y}}{dt} = a(\vec{y}), \\ \vec{y}(0) = \vec{y}_0, \end{cases}$$

admet une et une seule solution définie sur  $J$  qui est

$$J \rightarrow \mathbf{E} ; t \mapsto \exp(ta)(\vec{y}_0).$$

Plus généralement pour tout élément  $t_0$  de  $J$  et pour tout élément  $\vec{y}_0$  de  $\mathbf{E}$ , le problème de Cauchy

$$\begin{cases} \frac{d\vec{y}}{dt} = a(\vec{y}), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$$

admet une et une seule solution définie sur  $J$  qui est

$$J \rightarrow \mathbf{E} ; t \mapsto \exp(\dots) .$$

En effet

L'ensemble des solutions de (27) sur  $J$ , noté  $\mathcal{S}_{27}(J)$ , est d'après I.1.1., un espace vectoriel. L'existence et l'unicité d'une solution sur  $J$  du problème de Cauchy va permettre, comme nous allons voir, de déterminer la dimension de  $\mathcal{S}_{27}(J)$ .

Soit  $t_0$  un élément de  $J$ . Nous noterons dans la suite, pour tout élément  $\vec{y}_0$ , LA solution définie sur  $J$  du problème de Cauchy  $\begin{cases} \frac{d\vec{y}}{dt} = a(\vec{y}), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$

Considérons alors les applications

$$\mathcal{I} : \mathbf{E} \rightarrow \mathcal{S}_{27}(J) ; \vec{y}_0 \mapsto \varphi_{\vec{y}_0}$$

et

$$\mathcal{J} : \mathcal{S}_{27}(J) \rightarrow \mathbf{E} ; \varphi \mapsto \varphi(t_0).$$

Donc  $\mathcal{I}$  et  $\mathcal{J}$  sont des isomorphismes. Donc,  $\mathcal{S}_{27}(J)$  isomorphe à  $\mathbf{E}$ , est de dimension  $n$ .

Récapitulons

**Proposition 3.2.2.** — *Soit  $J$  un intervalle d'intérieur non vide, par exemple  $\mathbf{R}$ . Notons  $\mathcal{S}_{27}(J)$  l'ensemble des solutions de (27) définie sur  $J$ .*

1.  $\mathcal{S}_{27}(J)$  est l'ensemble des applications de la forme

$$J \rightarrow \mathbf{E}; t \mapsto \exp(ta)(\vec{c}),$$

où  $\vec{c}$  est un quelconque vecteur de  $\mathbf{E}$ .

2. Pour tout élément  $t_0$  de  $J$  et pour tout élément  $\vec{y}_0$  de  $\mathbf{E}$ , Le problème de Cauchy

$$\begin{cases} \frac{d\vec{y}}{dt} = a(\vec{y}), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$$

admet une et une seule solution,  $\varphi_{\vec{y}_0}$  définie sur  $J$  :

$$\varphi_{\vec{y}_0} : J \rightarrow \mathbf{E}; t \mapsto \exp((t - t_0)a)(\vec{y}_0).$$

(THÉORÈME DE CAUCHY-LIPSCHITZ LINÉAIRE)

3. L'ensemble  $\mathcal{S}_{27}(J)$  est un sous-espace vectoriel de l'espace vectoriel des applications de  $J$  dans  $\mathbf{E}$  dérivables (et même de  $\mathcal{C}^\infty(J, \mathbf{E})$ ), de dimension  $n$ . Les applications

$$\mathcal{I} : \mathbf{E} \rightarrow \mathcal{S}_{27}(J); \vec{y}_0 \mapsto \varphi_{\vec{y}_0}$$

et

$$\mathcal{J} : \mathcal{S}_{27}(J) \rightarrow \mathbf{E}; \varphi \mapsto \varphi(t_0),$$

sont des isomorphismes.

**Corollaire 3.2.3.** — *Soit  $b$  un endomorphisme de  $\mathbf{E}$  qui commute avec  $a$ . alors :*

$$\exp(a + b) = \exp(a) \circ \exp(b)$$

*Soient  $A$  et  $B$  des éléments de  $\mathcal{M}_n(\mathbf{K})$  qui entre eux commutent, alors*

$$\exp(A + B) = \exp(A) \exp(B)$$

++++

*Preuve du corollaire 3.2.3.* — Le deuxième point se déduit immédiatement du premier en raisonnant sur les endomorphismes de  $\mathcal{M}_n(k)\mathbf{K}$  canoniquement associés à  $A$  et  $B$ . Montrons le premier.

## RÉSOLUTION PRATIQUE

Le point 1 de la proposition 3.2.2. n'a guère qu'une utilité théorique. Nous nous intéressons maintenant à la résolution concrète du système différentiel

$$\frac{dX}{dt} = AX. \quad (30)$$

La résolution pratique de ce type de système a déjà été abondamment traitée dans le chapitre sur la réduction des endomorphismes. Récapitulons.

Nous nous intéresserons, tant qu'à faire, aux solutions définies sur  $\mathbf{R}$ . Nous noterons  $\mathcal{S}_{30}(\mathbf{R})$  l'ensemble des solutions définie sur  $\mathbf{R}$  de (30). Pour alléger les notations on pose  $\mathbf{E} := \mathcal{C}^1(\mathbf{R}, \mathcal{M}_{n,1}(\mathbf{K}))^1$

### 1. LA MATRICE $A$ EST DIAGONALISABLE DANS $\mathbf{K}$

On diagonalise  $A$  dans  $\mathcal{M}_n(\mathbf{K})$   $M$  admet  $n$  valeurs propres éléments de  $\mathbf{K}$  distinctes ou non  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Il existe une base de  $\mathcal{M}_{n,1}(\mathbf{K})$ , constituée vecteurs propres de  $M$ ,  $(U_1, U_2, \dots, U_n)$ . Notons, pour  $i = 1, 2, \dots, n$ ,  $\lambda_i$  la valeur propre associé à  $U_i$ . Posons enfin  $P = (U_1 U_2 \dots U_n)$ .  $P$  est élément de  $\mathrm{GL}_n(\mathbf{K})$  et

$$M = P\Delta P^{-1},$$

avec  $\Delta := \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ .

Pour tout élément  $\Phi$  de  $\mathbf{E}$ , l'application  $P^{-1}\Phi$  est de classe  $\mathcal{C}^1$ . On dispose donc de l'application

$$\mathcal{I}_1 : \mathbf{E} \rightarrow \mathbf{E}; \Phi \mapsto P^{-1}\Phi.$$

1. on aurait pu poser  $E := \mathcal{D}^1(\mathbf{R}, \mathcal{M}_{n,1}(\mathbf{K}))$ , cela n'a pas d'importance, les solutions de (28) étant même indéfiniment dérivables (cf. remarque 3.1.2)

et de même de l'application :

$$\mathcal{I}_2 : \mathbf{E} \rightarrow \mathbf{E}; \Psi \mapsto P\Psi.$$

Notons que  $\mathcal{I}_2 \circ \mathcal{I}_1 = \text{id}_{\mathbf{E}}$  et  $\mathcal{I}_1 \circ \mathcal{I}_2 = \text{id}_{\mathbf{E}}$ . Donc  $\mathcal{I}_1$  et  $\mathcal{I}_2$  sont des bijections de  $\mathbf{E}$  sur  $\mathbf{E}$  (et même, du reste, des isomorphismes)

Un élément quelconque  $\Phi$  de  $\mathbf{E}$ , est solution de (30) si et seulement si  $\Phi' = M\Phi = P\Delta P^{-1}\Phi$ , soit si et seulement si  $(P^{-1}\Phi)' = \Delta P^{-1}\Phi$ . Donc  $\Phi$  est solution de (30) si et seulement si  $\mathcal{I}_1(\Phi)$  est solution sur  $\mathbf{R}$  du système

$$\frac{dY}{dt} = \Delta Y. \quad (31)$$

Autrement dit,  $\mathcal{I}_1$  induit une bijection<sup>1</sup> de  $\mathcal{S}_{30}(\mathbf{R})$  sur  $\mathcal{S}_{31}(\mathbf{R})$  :

$$\mathcal{S}_{30}(\mathbf{R}) = \mathcal{I}_2(\mathcal{S}_{31}(\mathbf{R})). \quad (32)$$

Or  $\mathcal{S}_{31}(\mathbf{R})$  est l'ensemble des applications de la forme :

$$\mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{K}); t \mapsto \begin{pmatrix} c_1 \exp(\lambda_1 t) \\ c_2 \exp(\lambda_2 t) \\ \vdots \\ c_n \exp(\lambda_n t) \end{pmatrix},$$

où  $c_1, c_2, \dots, c_n$  sont des éléments quelconques de  $\mathbf{K}$ . Pour le voir on peut soit utiliser le premier point de 3.2.2., soit résoudre « ligne par ligne » le système. Donc d'après (32),  $\mathcal{S}_{30}(\mathbf{R})$  est l'ensemble des applications :

$$\Phi_{c_1, c_2, \dots, c_n} : \mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{K}); t \mapsto c_1 \exp(\lambda_1 t)U_1 + c_2 \exp(\lambda_2 t)U_2 + \dots + c_n \exp(\lambda_n t)U_n,$$

où  $c_1, c_2, \dots, c_n$  sont des éléments quelconques de  $\mathbf{K}$ .

Faisons quelques remarques.

Pour  $i = 1, 2, \dots, n$  posons pour alléger  $\Phi_i = \Phi_{0,0,\dots,0, \underset{\rightarrow}{1}, 0, \dots, 0}$ . alors pour tout élément  $(c_1, c_2, \dots, c_n)$  de  $\mathbf{K}^n$ ,

$$\Phi_{c_1, c_2, \dots, c_n} = \sum_{i=1}^n c_i \Phi_i.$$

Donc la famille  $(\Phi_i)_{i=1, \dots, n}$  engendre  $\mathcal{S}_{30}(\mathbf{R})$ . La dimension de  $\mathcal{S}_{30}(\mathbf{R})$  étant  $n$  (cf. 3.2.2. -3),  $(\Phi_i)_{i=1, \dots, n}$  en est une base.

Soient  $Y_0$  un élément de  $\mathcal{M}_{n,1}(\mathbf{K})$  et  $(c_1, c_2, \dots, c_n)$  de  $\mathbf{K}^n$ .  $\Phi_{c_1, c_2, \dots, c_n}(0) = Y_0$  si et seulement si

$$\sum_{i=1}^n c_i U_i = Y_0, \text{ soit encore si et seulement si } (U_1 U_2 \dots U_n) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = P \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = Y_0. \text{ La solution définie sur } \mathbf{R} \text{ de}$$

(30) est donc  $\Phi_{c_1, c_2, \dots, c_n}$  avec  $(c_1, c_2, \dots, c_n) = {}^t(P^{-1}Y_0)$ .

Exemple : Déterminer les solutions définies sur  $\mathbf{R}$  et à valeurs dans  $\mathcal{M}_{2,1}(\mathbf{R})$  de

$$\frac{dX}{dt} = \begin{pmatrix} 6 & 4 \\ 5 & 5 \end{pmatrix} X. \quad (33)$$

## 2. LA MATRICE $A$ N'EST PAS DIAGONALISABLE DANS $\mathbf{K}$

1. et même un isomorphisme

(a) *Premier sous-cas :  $\mathbf{K} = \mathbf{R}$  et  $A$  est diagonalisable dans  $\mathcal{M}_n(\mathbf{C})$*

On commence par déterminer les solutions à valeurs dans  $\mathcal{M}_n(\mathbf{C})$  en utilisant le point 1.

Notons  $\mathcal{S}_{30,\mathbf{C}}(\mathbf{R})$  l'ensemble des solutions de (30) définie sur  $\mathbf{R}$  à valeurs dans  $\mathcal{M}_{n,1}(\mathbf{C})$  et  $\mathcal{S}_{30,\mathbf{R}}(\mathbf{R})$  celles à valeurs  $\mathcal{M}_{n,1}(\mathbf{R})$ . La matrice  $A$  étant à coefficients réels, si un complexe  $\lambda$  est valeur propre, alors  $\bar{\lambda}$  est valeur propre avec la même multiplicité, et si  $U$  est un vecteur propre associé à  $\lambda$ ,  $\bar{U}$  en est un associé à  $\bar{\lambda}$ . Ainsi existe-t-il une base de  $\mathcal{M}_{n,1}(\mathbf{C})$ ,  $(U_1, U_2, \dots, U_p, V_1, \bar{V}_1, \dots, V_q, \bar{V}_q)$ , constituée de vecteurs propres de  $A$  telle que :

— Pour  $i = 1, 2, \dots, p$ ,  $U_i$  est associé à une valeur propre réelle  $\mu_i$  et est élément de  $\mathcal{M}_{n,1}(\mathbf{R})$ .

— Pour  $j = 1, 2, \dots, q$ ,  $V_j$  est associé à une valeur propre complexe  $\lambda_j$  de partie imaginaire strictement positive.

—  $p \in \mathbf{N}$  et  $q \in \mathbf{N}^*$  ( $A$  n'est pas diagonalisable dans  $\mathcal{M}_n(\mathbf{R})$ )

D'après le premier point, on dispose de la base de  $\mathcal{S}_{30,\mathbf{C}}(\mathbf{R})$ ,

$$(\Phi_1, \Phi_2, \dots, \Phi_p, \Psi_1, \bar{\Psi}_1, \Psi_2, \bar{\Psi}_2, \dots, \Psi_q, \bar{\Psi}_q)$$

avec

— Pour  $i = 1, 2, \dots, p$ ,  $\Phi_i : \mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{R})$ ;  $t \mapsto \exp(\mu_i t) U_i$ ;

— Pour  $j = 1, 2, \dots, q$ ,  $\Psi_j : \mathbf{R} \rightarrow \mathcal{M}_{n,1}(\mathbf{C})$ ;  $t \mapsto \exp(\lambda_j t) V_j$ ;

On la notera  $\mathcal{B}_{\mathbf{C}}$ . Posons pour  $j = 1, 2, \dots, p$ ,  $F_j := \frac{\Psi_j + \bar{\Psi}_j}{2}$  et  $G_j := \frac{\Psi_j - \bar{\Psi}_j}{2i}$ . Alors la famille

$$(\Phi_1, \Phi_2, \dots, \Phi_p, F_1, G_1, F_2, G_2, \dots, F_q, G_q)$$

est une base de  $\mathcal{S}_{30,\mathbf{C}}(\mathbf{R})$ , puisque son déterminant dans la base  $\mathcal{B}_{\mathbf{C}}$  vaut

$$\left| \text{diag} \left( I_p, \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix}, \dots, \begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix}}_{q \text{ termes}} \right) \right| = \left( -\frac{2}{4i} \right)^q \neq 0$$

Or pour  $j = 1, 2, \dots, p$ ,  $F_j = \text{Re}(\Psi_j)$  et  $G_j = \text{Im}(\Psi_j)$ . Donc les éléments de la dernière base sont à valeurs dans  $\mathcal{M}_{n,1}(\mathbf{R})$ ,  $(\Phi_1, \Phi_2, \dots, \Phi_p, F_1, G_1, F_2, G_2, \dots, F_q, G_q)$  est donc une famille de  $n$  éléments de  $\mathcal{S}_{30,\mathbf{R}}(\mathbf{R})$ . Cette famille libre dans  $\mathcal{S}_{30,\mathbf{C}}(\mathbf{R})$ , est *a fortiori* libre dans  $\mathcal{S}_{30,\mathbf{R}}(\mathbf{R})$ . Or d'après 3.2.1–3,  $\mathcal{S}_{30,\mathbf{R}}(\mathbf{R})$  est de dimension  $n$ , donc  $(\Phi_1, \Phi_2, \dots, \Phi_p, F_1, G_1, F_2, G_2, \dots, F_q, G_q)$  est une base de  $\mathcal{S}_{30,\mathbf{R}}(\mathbf{R})$ . On en déduit la forme des solutions de (30). Regardons sur un exemple.

Exemple : Déterminer les solutions définies sur  $\mathbf{R}$  et à valeurs dans  $\mathcal{M}_{3,1}(\mathbf{R})$  de

$$\frac{dX}{dt} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\sqrt{3} \\ 0 & \sqrt{3} & 1 \end{pmatrix} X. \quad (34)$$

(b) *Second sous-cas :  $A$  n'est pas diagonalisable dans  $\mathcal{M}_n(\mathbf{C})$*

Alors on trigonalise  $A$ .



- Si  $\mathbf{K} = \mathbf{C}$ , on trigonalise  $A$  dans  $\mathcal{M}_n(\mathbf{C})$  (ce qui est toujours possible).
- Si  $\mathbf{K} = \mathbf{R}$ , on trigonalise  $A$  si possible dans  $\mathcal{M}_n(\mathbf{R})$  et sinon dans  $\mathcal{M}_n(\mathbf{C})$ .

Donnons un exemple.

Exemple : Déterminer les solutions définies sur  $\mathbf{R}$  et à valeurs dans  $\mathcal{M}_{3,1}(\mathbf{R})$  de

$$\frac{dX}{dt} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 0 \end{pmatrix} X. \quad (35)$$

Pour finir, donnons à titre d'application, la résolution d'équations d'ordre trois à coefficients constants, équations qui généralisent les équations d'ordre 2 à coefficients constants, étudiées en 2.1. ; nous dirons un mot des équation scalaire générale d'ordre  $n$  dans la partie suivante.

**Exercice 3.2.4.** — Résoudre sur  $\mathbf{R}$  l'équation différentielle

$$\frac{d^3y}{dt^3} + \frac{d^2y}{dt^2} + \frac{dy}{dt} - 3y = 0, \quad (36)$$

*c'est-à-dire déterminer l'ensemble  $\mathcal{S}_{36}(\mathbf{R})$  des éléments  $\varphi$  de  $\mathcal{C}^3(\mathbf{R}, \mathbf{R})$ , tels que  $\varphi''' + \varphi'' + \varphi' - 3\varphi = 0$ .*

*Solution de l'exercice 3.2.4.* — Observons d'abord que  $\mathcal{S}_{36}$  est un sous-espace vectoriel de  $\mathcal{C}^3(\mathbf{R}, \mathbf{R})$ ; c'est en effet le noyau de l'application linéaire :

Considérons le système

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (37)$$

Soit

$$\mathcal{I}_3 : \mathcal{C}^3(\mathbf{R}, \mathbf{R}) \rightarrow \mathcal{C}^1(\mathbf{R}, \mathcal{M}_{3,1}(\mathbf{R})); \varphi \mapsto \begin{pmatrix} \varphi \\ \varphi' \\ \varphi'' \end{pmatrix}.$$

$\mathcal{I}_3$  est clairement linéaire (linéarité de la dérivation), montrons qu'il induit un isomorphisme de  $\mathcal{S}_{36}(\mathbf{R})$  sur  $\mathcal{S}_{37}(\mathbf{R})$ .

- On a défini le système  $\mathcal{S}_{37}(\mathbf{R})$  de sorte que  $\mathcal{I}_3(\mathcal{S}_{36}(\mathbf{R})) \subset \mathcal{S}_{37}(\mathbf{R})$ . On dispose donc de  $\tilde{\mathcal{I}}_3$ , application linéaire induite par  $\mathcal{I}_3$  de  $\mathcal{S}_{36}(\mathbf{R})$  dans  $\mathcal{S}_{37}(\mathbf{R})$ .
- $\tilde{\mathcal{I}}_3$  est injectif. En effet...

- $\tilde{\mathcal{I}}_3$  est surjectif. En effet...

Déterminons  $\mathcal{S}_{37}(\mathbf{R})$ .

Soit  $M$  la matrice du système (37). C'est une matrice... , l'étude du polynôme caractéristique d'une telle matrice a été fait en TD :

$$\chi_M = (X^3 + X^2 + X - 3) = (1 - X)(X^2 + 2X + 3).$$

Tout cela nous rappelle l'équation caractéristique d'une équation du second ordre à coefficients constants! Le spectre complexe de  $M$  est  $\{1, -1 + i\sqrt{2}, -1 - i\sqrt{2}\}$ . Nous ne pouvons pas diagonaliser  $M$  dans  $\mathcal{M}_3(\mathbf{R})$ . Diagonalisons dans

$\mathcal{M}_3(\mathbf{C})$ .

**Exercice 3.2.5.** — Résoudre sur  $\mathbf{R}$  l'équation différentielle

$$\frac{d^3y}{dt^3} + 3\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + y = 0, \quad (38)$$

*Solution de l'exercice 3.2.5.* — Comme dans l'exercice précédent, observons que  $\mathcal{S}_{38}$  est un sous-espace vectoriel de  $\mathcal{C}^3(\mathbf{R}, \mathbf{R})$  et Considérons le système

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (39)$$

Soit

$$\mathcal{I}_4 : \mathcal{C}^3(\mathbf{R}, \mathbf{R}) \rightarrow \mathcal{C}^1(\mathbf{R}, \mathcal{M}_{3,1}(\mathbf{R})); \varphi \mapsto \begin{pmatrix} \varphi \\ \varphi' \\ \varphi'' \end{pmatrix}.$$

$\mathcal{I}_4$  est clairement linéaire et induit un isomorphisme de  $\mathcal{S}_{36}(\mathbf{R})$  sur  $\mathcal{S}_{37}(\mathbf{R})$ . Déterminons  $\mathcal{S}_{39}(\mathbf{R})$ .

Soit  $M$  la matrice du système (37). C'est une matrice... , l'étude du polynôme caractéristique d'une telle matrice a été fait en TD :

$$\chi_M = X^3 + 3X^2 + 3X + 1 = (1 + X)^3.$$

Le spectre de  $M$  est  $\{-1\}$ .  $M$  n'est pas diagonalisable, car sinon elle serait semblable, donc égale, à  $-I_3$ .  $M$  est en

revanche trigonalisable dans  $\mathcal{M}_3(\mathbf{R})$  car...

Terminons par des exercices sur la stabilité de la solution nulle.

**Exercice 3.2.6.** —

Soit le système différentiel linéaire :

$$\frac{dX}{dt} = MX, \quad (40)$$

où  $M \in \mathcal{M}_n(\mathbf{K})$ . Pour simplifier nous supposons que  $M$  est diagonalisable dans  $\mathcal{M}_n(\mathbf{C})$  et par  $\lambda_1, \dots, \lambda_n$  nous désignerons les  $n$  valeurs propres complexes de  $M$  distinctes ou non.

L'application nulle est solution de (40). On dit que  $O_{n,1}$  est une position d'équilibre du système.

On dit que  $O_{n,1}$  est une position d'équilibre stable si pour tout réel  $\varepsilon > 0$ , il existe un réel  $\eta > 0$  tel que pour toute solution  $\Phi$  de (40), si  $\|\Phi(0)\| < \eta$  alors pour tout réel  $t \geq 0$ ,  $\|\Phi(t)\| < \varepsilon$ , si en plus  $\Phi(t)$  tend vers  $O_{n,1}$  lorsque  $t$  tend vers  $+\infty$ , on dit que  $\Phi$  est asymptotiquement stable.

Dans cette définition,  $\|\cdot\|$  désigne une norme sur  $\mathcal{M}_{n,1}(\mathbf{K})$  que nous ne préciserons pas puisque toutes les normes sont équivalentes et que l'on vérifie sans mal que dans la définition de la stabilité le choix de la norme est donc sans importance.

1. Dans cette question on prend  $\mathbf{K} = \mathbf{C}$ .

- (a) Montrer que si les parties réelles de toutes les valeurs propres de  $M$  sont toutes strictement négatives alors la position d'équilibre du système est asymptotiquement stable.
  - (b) Montrer que si une des parties réelles des valeurs propres de  $M$  est strictement positive alors la position d'équilibre n'est pas stable.
  - (c) Que dire de la position d'équilibre lorsque les parties réelles des valeurs propres sont toutes nulles.
2. Reprendre la question précédente avec  $\mathbf{K} = \mathbf{R}$ .
  3. On suppose dans cette question que  $n = 2$ . Tracer la forme des trajectoires de (40) c'est-à-dire les arcs paramétrés  $(\mathbf{R}, \Phi)$ , où  $\Phi$  est solution de (40), dans les deux cas suivant :
    - (a)  $\lambda_1$  et  $\lambda_2$  sont des réels tels que :  $0 < \lambda_1 < \lambda_2$ .
    - (b)  $\lambda_1$  et  $\lambda_2$  sont des réels tels que :  $\lambda_1 < 0 < \lambda_2$ .

### 3.3 Étude de l'équation avec second membre

Nous nous intéresserons, tant qu'à faire, aux solutions de (26), définies sur  $I$  entier. Nous noterons  $\mathcal{S}_{26}(I)$  l'ensemble des solutions de (26) définies sur  $I$ . *A priori*  $\mathcal{S}_{26}(I)$  est soit le vide soit un espace affine dirigé par  $\mathcal{S}_{27}(I)$ . Nous allons montrer que le premier cas est impossible. La méthode employée pour résoudre (26) sera une méthode de changement d'inconnue qui se calque sur celle utilisée pour l'équation scalaire du premier ordre, on la nomme encore, hélas, *méthode de variation de la constante*.

Rappelons que nous avons vu, dans le cours sur les séries de fonctions, que pour tout réel  $t$ ,  $\exp(ta)$  et  $\exp(-ta)$  sont des éléments de  $\mathcal{L}(E)$  inverses l'un de l'autre, résultat que généralise le corollaire 3.3.3.

Soit  $\vec{\varphi}$  une application de  $I$  dans  $\mathbf{E}$  dérivable. Posons

$$\vec{\psi} : I \rightarrow \mathbf{E}; t \mapsto \exp(-ta)(\vec{\varphi}(t))$$

de sorte que, d'après la propriété rappelée,  $\vec{\varphi}(t) = \exp(ta)(\vec{\psi}(t))$ , pour tout  $t \in I$ . Le lemme 3.2.1. assure que  $\vec{\psi}$  est dérivable et que pour tout  $t \in I$ ,

$$\vec{\varphi}'(t) = (a \circ \exp(ta))(\vec{\psi}(t)) + \exp(ta)(\vec{\psi}'(t)),$$

Donc  $\vec{\varphi}$  est solution sur  $I$  de (26), si et seulement si  $(\exp(\cdot a)(\vec{\psi}))' = \vec{b}$ , soit, si et seulement si

$$\vec{\psi}' = \exp(-\cdot a)(\vec{b}). \quad (41)$$

L'application  $\exp(-\cdot a)(\vec{b})$  étant continue ( $\exp(-\cdot a)$  et  $\vec{b}$  le sont et.....), elle admet des primitives sur  $J$  soit  $\vec{p}_0$  l'une d'elle. D'après (41),  $\vec{\varphi}$  est solution sur  $I$  de (26) si et seulement si il existe un élément  $\vec{k}$  de  $\mathbf{E}^1$  tel que  $\vec{\psi} = \vec{p}_0 + \vec{k}$ .

---

1. Rappelons que le cours sur l'intégrale d'une fonction vectorielle, nous apprend que toutes les primitives de  $\exp(-\cdot a)(\vec{b})$  diffèrent de  $\vec{p}_0$  d'une constante, élément de  $\mathbf{E}$ .

Donc  $\mathcal{S}_{26}(I)$  est l'ensemble des applications de la forme

$$\vec{\varphi}_{\vec{k}} : I \rightarrow \mathbf{E}; t \mapsto \underbrace{\exp(ta)(\vec{p}_0(t))}_A + \underbrace{\exp(ta)(\vec{k})}_B. \quad (42)$$

Remarquons que A est  $\cdot$ , B est  $\cdot$ .

On en déduit immédiatement le théorème suivant

**Proposition 3.3.1.** — THÉORÈME DE CAUCHY-LIPSCHITZ LINÉAIRE —

Soient  $t_0$  un élément de  $I$  et  $\vec{y}_0$  un de  $\mathbf{E}$ . Le problème de Cauchy

$$\begin{cases} \frac{d\vec{y}}{dt} = a(\vec{y}) + \vec{b}(t), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$$

admet une et une seule solution définie sur  $I$ , qui est<sup>1</sup>

$$I \rightarrow \mathbf{E}; t \mapsto \exp\left(\int_{t_0}^t a(\vec{y}_0) \, d\tau\right) (\vec{y}_0) + \int_{t_0}^t \exp\left(\int_{t_0}^{\tau} a(\vec{y}_0) \, d\sigma\right) \vec{b}(\tau) \, d\tau$$

### Exercice

Preuve de la proposition 3.3.1. —

**Remarque 3.3.2.** —

1. Entraînez vous à retrouver très vite cette formule.

1. Ce résultat généralise La proposition 2.1.5., en identifiant l'élément  $a$  de  $\mathbf{K}$  est l'application linéaire  $\mathbf{K} \rightarrow \mathbf{K}; y \mapsto ay$ .
2. La méthode de variation de la constante est essentiellement théorique. Dans la résolution pratique, très souvent, on raisonne sur la forme matricielle (28) et l'on procède à la réduction de la matrice  $M$  comme dans la résolution de l'équation homogène (cf 2.1). On est alors amené à résoudre des équations linéaires du premier ordre scalaires avec second membre. On peut évidemment appliquer à chaque équation ainsi obtenue la méthode de variation de la constante cf. 2.1, ce qui revient en fait à présenter la méthode de variation de la constante pour (28) sous une forme différente. Passons à des exemples.

**Exemple 3.3.3.** — Déterminer les solutions définies sur  $\mathbf{R}$  et à valeurs dans  $\mathcal{M}_{2,1}(\mathbf{R})$  de

$$\frac{dX}{dt} = \begin{pmatrix} 6 & 4 \\ 5 & 5 \end{pmatrix} X + \begin{pmatrix} \exp(t) \\ 0 \end{pmatrix}. \quad (43)$$

**Exemple 3.3.4.** — Déterminer les solutions définies sur  $]-\frac{\pi}{2}, \frac{\pi}{2}[$  et à valeurs dans  $\mathcal{M}_{2,1}(\mathbf{R})$  de

$$\frac{dX}{dt} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} X + \begin{pmatrix} 0 \\ \tan(t) \end{pmatrix}. \quad (44)$$

# ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES VECTORIELLES DU PREMIER ORDRE À COEFFICIENT QUELCONQUE

Nous allons étudier les équations linéaires du premier ordre vectorielles, dans lesquelles, contrairement à celles du 3., le coefficient dépend de la variable  $t$ . Elles généralisent donc à la fois celles du 3., et les équations linéaires scalaires du premier ordre, dont nous avons rappelé l'étude en 2.1.

Le premier paragraphe donnera des généralités terminologiques et énoncera un théorème d'existence et d'unicité de la solution d'un problème de Cauchy, le deuxième traitera du cas sans second membre, le dernier étudiera le cas de l'équation avec second membre, grâce à une adaptation de la méthode de variation de la constante, vue en 3.3.

Dans toute cette partie  $\mathbf{K}$  désigne indifféremment  $\mathbf{R}$  ou  $\mathbf{C}$ ,  $\mathbf{E}$  est un espace vectoriel sur  $\mathbf{K}$  de dimension finie, non nulle  $n$  et  $I$  un intervalle d'intérieur non vide. Enfin  $a$  désigne une application de  $I$  à valeurs dans  $\mathcal{L}(\mathbf{E})$ , continue

$$a : I \rightarrow \mathcal{L}(\mathbf{E}) ; t \mapsto a(t).$$

## 4.1 Définitions, généralités

**Définition 4.1.1.** — Soit  $\vec{b}$  une applications de  $I$  dans  $\mathbf{E}$ , continue.

— On appelle solution de l'équation

$$\frac{d\vec{y}}{dt} = a(t)(\vec{y}) + \vec{b}(t), \quad (45)$$

(équation différentielle linéaires vectorielles du premier ordre)

toute application  $\vec{\varphi}$  d'un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ , à valeurs dans  $\mathbf{E}$ , dérivable, telle que, pour tout élément  $t$  de  $J$  :

$$\vec{\varphi}'(t) = a(t)(\vec{\varphi}(t)) + \vec{b}(t).$$

— Soient  $t_0$  un élément de  $I$  et  $\vec{y}_0$  un élément de  $\mathbf{E}$ . On appelle solution du problème de Cauchy

$$\begin{cases} \frac{d\vec{y}}{dt} = a(t)(\vec{y}) + \vec{b}(t), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$$

toute solution  $\vec{\varphi}$  de (26) définie sur un intervalle  $J$  contenant  $t_0$  et telle que  $\vec{\varphi}(t_0) = \vec{y}_0$ . On dit encore qu'une telle solution satisfait à la condition initiale  $\vec{y}(t_0) = \vec{y}_0$ .

— Dans le cas particulier où  $\vec{b}$  est nulle, l'équation (26) qui devient

$$\frac{d\vec{y}}{dt} = a(t)(\vec{y}), \quad (46)$$

est dite homogène ou sans second membre. On dit que l'équation (46) est l'équation homogène associée à (45).

— L'applications  $a$  et parfois  $\vec{b}$ , sont appelées coefficients de l'équation (9) ;  $\vec{b}(t)$  est encore appelée second membre de l'équation.

**Remarques 4.1.2.** —

1. Lorsque l'on s'intéresse à des solutions sur un intervalle  $J$  donné, Le couple d'équations ((45),(46)) est un cas particulier du couple d'équations ((e),(h)), avec pour  $\mathbf{E}$ , , et

$\ell$  :

Ceci explique la terminologie choisie.

2. On peut montrer, *a priori*, que toute solution de l'équation (45) est en fait une application de classe  $\mathcal{C}^1$
3. On notera l'abus de notations, classique et préconisé par le programme, dans ces équations : la variable  $t$  ne figure pas dans l'inconnue mais dans les applications  $a$  et  $\vec{b}$ . Cet abus n'est tolérable, que dans l'écriture de l'équation, jamais dans les égalités obtenues en y substituant une solution à l'inconnue.
4. Il arrive que l'on désigne l'image par  $a(t)$  d'un élément  $\vec{x}$  de  $\mathbf{E}$  par  $a(t) \cdot \vec{x}$  ou  $a(t) \vec{x}$ , au lieu de  $a(t)(\vec{x})$ , cette notation classique est plus légère, renforce l'analogie entre le cas vectoriel et le cas scalaire, le programme officiel y a renoncé...



## EXPRESSION MATRICIELLE

Soit  $\mathcal{B}$  une base de  $\mathbf{E}$ . Pour tout élément  $t$  de  $I$ , on note  $(a_{i,j}(t))_{i=1\dots n, j=1\dots n}$  la matrice de  $a(t)$  dans  $\mathcal{B}$ , que l'on notera

encore  $A(t)$  et  ${}^t(b_1, b_2, \dots, b_n)$ , le vecteur colonne des composantes de  $\vec{b}$  dans  $\mathcal{B}$ , noté plus simplement  $B$ . Soit  $J$  un intervalle inclus dans  $I$ , d'intérieur non vide et  $\vec{\varphi}$  une application de  $J$  dans  $\mathbf{E}$ , dérivable. Notons  $\Phi$  le vecteur colonne des composantes de  $\vec{\varphi}$  dans  $\mathcal{B}$  et pour  $i = 1, 2, \dots, n$ ,  $\varphi_i$ , la  $i^{\text{e}}$  composante de  $\vec{\varphi}$  dans  $\mathcal{B}$ .

D'après le cours sur la représentation matricielle des endomorphismes, L'application  $\vec{\varphi}$  est solution de (45) si et seulement si, pour tout élément  $t$  de  $J$ ,  $\Phi'(t) = A(t)\Phi(t) + B(t)$ . Si c'est le cas, on dit que  $\Phi$  est solution du *système différentiel linéaire*

$$\frac{dX}{dt} = A(t)X + B(t). \quad (47)$$

En explicitant ce qui précède,  $\vec{\varphi}$  est solution de (45) si et seulement si pour tout élément  $t$  de  $J$ ,

$$\varphi'_i(t) = \sum_{j=1}^n a_{i,j}(t)\varphi_j(t) + b_i(t), \text{ pour } i = 1, 2, \dots, n.$$

Si c'est le cas on dit que  $(\varphi_1, \varphi_2, \dots, \varphi_n)$  est solution du *système d'équations différentielles linéaires*

$$\left\{ \begin{array}{l} \frac{dx_1}{dt} = \sum_{j=1}^n a_{1,j}(t)x_j + b_1(t) \\ \frac{dx_2}{dt} = \sum_{j=1}^n a_{2,j}(t)x_j + b_2(t) \\ \vdots \\ \frac{dx_n}{dt} = \sum_{j=1}^n a_{n,j}(t)x_j + b_n(t) \end{array} \right. \quad (48)$$

Dans les études théoriques on raisonnera sur l'équations (45). Le système (47), qui est en fait un cas particulier du précédent, où  $\mathbf{E} = \mathcal{M}_{n,1}(\mathbf{K})$ , lui sera préféré dans les exercices pratiques. Nous énoncerons les résultats pour (45), à charge pour le lecteur de les traduire en termes matriciels pour (47).

Nous allons ici donner un théorème assurant l'existence et l'unicité de la solution sur un intervalle donné d'un problème de Cauchy.

**Proposition 5.1.3.** — THÉORÈME DE CAUCHY-LIPSCHITZ LINÉAIRE —

Soit  $J$  un intervalle inclus dans  $I$ , d'intérieur non vide. Soient  $t_0$  un élément de  $J$  et  $\vec{y}_0$  un de  $\mathbf{E}$ . Alors le problème de Cauchy

$$\left\{ \begin{array}{l} \frac{d\vec{y}}{dt} = a(t) \cdot \vec{y} + \vec{b}(t), \\ \vec{y}(t_0) = \vec{y}_0, \end{array} \right.$$

admet une et une seule solution définie sur  $J$ .

Nous admettrons la preuve de ce théorème qui est non exigible des élèves est assez technique. Tout au plus allons-nous donner un schéma de démonstration

*Idées générales de la preuve du théorème 4.1.3.* —

La preuve repose sur l'expression intégrale du problème de Cauchy :

Soit  $\vec{\varphi}$  une application de  $J$  dans  $\mathbf{E}$  continue.  $\varphi$  est solution si et seulement si, pour tout élément  $t$  de  $J$ ,

$$\vec{\varphi}(t) = \vec{y}_0 + \int_{t_0}^t a(\tau) \cdot \vec{\varphi}(\tau) + \vec{b}(\tau) d\tau.$$

en effet

La formulation intégrale du problème de Cauchy est au programme.

Introduisons alors l'application linéaire

$$L : \mathcal{C}^0(J, \mathbf{E}) \rightarrow \mathcal{C}^0(J, \mathbf{E}); \vec{\varphi} \mapsto L(\vec{\varphi}),$$

où  $L(\vec{\varphi}) : J \rightarrow \mathbf{E}; t \mapsto \vec{y}_0 + \int_{t_0}^t a(\tau) \cdot \vec{\varphi}(\tau) + \vec{b}(\tau) d\tau$ . Une application  $\vec{\varphi}$  de  $J$  dans  $\mathbf{E}$  continue est donc solution sur  $J$  du problème de Cauchy si et seulement si elle est un *point fixe* de  $L$ .

Soit  $\vec{\varphi}_1$  et  $\vec{\varphi}_2$  des éléments  $\mathcal{C}^0(J, \mathbf{E})$ .

Prenons  $K$  un segment de  $J$

On munit  $\mathbf{E}$  d'une norme  $\|\cdot\|_{\mathbf{E}}$  et  $\mathcal{C}^0(K, \mathbf{E})$  de la norme de la convergence uniforme  $\|\cdot\|_{\infty}$  associée ( $K$  est un segment). On note  $\|\cdot\|$  la norme de  $\mathcal{L}(\mathbf{E})$  subordonnée à  $\|\cdot\|$ , (voir TD) et  $\|a\|_{\infty}$  désigne la quantité  $\sup_{t \in K} (\|a(t)\|)$ . On prouve alors par récurrence que pour tout entier  $n \geq 0$ ,

$$\|L^n(\varphi_2) - L^n(\varphi_1)\|_{\infty} \leq \frac{\|a\|_{\infty}^n |J|^n}{n!} \|\varphi_2 - \varphi_1\|_{\infty}$$

Si  $\varphi_1$  et  $\varphi_2$  sont solution du problème de Cauchy considéré  $L^n(\varphi_i) = \varphi_i$ , pour  $i = 1, 2$  et en laissant tendre  $n$  vers  $+\infty$  dans l'inégalité précédente on obtient que  $\varphi_1$  et  $\varphi_2$  coïncident sur  $K$  segment quelconque de  $J$ . Donc  $\varphi_1 = \varphi_2$ . Voilà pour l'unicité de la solution sur  $J$  du problème de Cauchy.

Pour l'existence, prenons  $\psi \in \mathcal{C}^0(J, \mathbf{E})$  et posons pour tout entier  $n \geq 0$ ,  $\psi^n = L^n(\psi)$ , on vient de montrer que pour le segment  $K$  précédent et la norme de la convergence uniforme associée :

$$\|\psi_{n+1} - \psi_n\|_{\infty} \leq \frac{\|a\|_{\infty}^n |J|^n}{n!} \|\psi_1 - \psi_0\|_{\infty}$$

On obtient la convergence normale sur  $K$  de la série  $\sum \psi_{n+1} - \psi_n$ . Le segment  $K$  étant quelconque  $\sum \psi_{n+1} - \psi_n$  converge normalement sur tout segment de  $J$ . On en déduit par télescopie une limite pour  $\psi_n$  qui est un point fixe de  $L$  (grâce à la convergence normale sur tout segment), donc une solution du problème de Cauchy.

Soit  $J$  un intervalle de  $I$  d'intérieur non vide. Notons  $\mathcal{S}_{45}(J)$  l'ensemble des solutions de (45), définies sur  $J$  et  $\mathcal{S}_{46}(J)$  celui des solutions de (46), définies sur  $J$ .

D'après la remarque VI.1.2.-1,  $\mathcal{S}_{46}(J)$  est un *sous-espace vectoriel* de  $\mathcal{D}^1(J, \mathbf{E})$  (et même, grâce à VI.1.2.-2, de  $\mathcal{C}^1(J, \mathbf{E})$ ).  $\mathcal{S}_{45}(J)$  étant d'après le théorème de Cauchy-Lipschitz linéaire, non vide, c'est un *espace affine* dirigé par  $\mathcal{S}_{46}(J)$ .

Soit  $t_0$  un élément de  $J$ . Nous noterons dans la suite, pour tout élément  $\vec{\varphi}_{\vec{y}_0}$ , la solution définie sur  $J$  du problème de Cauchy  $\begin{cases} \frac{d\vec{y}}{dt} = a(t)(\vec{y}), \\ \vec{y}(t_0) = \vec{y}_0, \end{cases}$

Considérons alors les applications

$$\mathcal{I} : \mathbf{E} \rightarrow \mathcal{S}_{46}(J); \vec{y}_0 \mapsto \varphi_{\vec{y}_0}$$

et

$$\mathcal{J} : \mathcal{S}_{46}(J) \rightarrow \mathbf{E}; \varphi \mapsto \varphi(t_0).$$

Donc  $\mathcal{I}$  et  $\mathcal{J}$  sont des isomorphismes. Donc,  $\mathcal{S}_{46}(J)$  isomorphe à  $\mathbf{E}$ , est de dimension  $n$ .

Récapitulons

**Proposition 4.1.4.** — *Soit  $J$  un intervalle d'intérieur non vide, par exemple  $\mathbf{R}$ .*

1. *L'ensemble  $\mathcal{S}_{46}(J)$  est un sous-espace vectoriel de l'espace vectoriel des applications de  $J$  dans  $\mathbf{E}$  dérivables (et même de  $\mathcal{C}^\infty(J, \mathbf{E})$ ), de dimension  $n$ .*
2. *Les applications*

$$\mathcal{I} : \mathbf{E} \rightarrow \mathcal{S}_{46}(J); \vec{y}_0 \mapsto \vec{\varphi}_{\vec{y}_0}$$

*et*

$$\mathcal{J} : \mathcal{S}_{46}(J) \rightarrow \mathbf{E}; \vec{\varphi} \mapsto \vec{\varphi}(t_0),$$

*sont des isomorphismes.*

3. *L'ensemble  $\mathcal{S}_{45}(J)$  est un sous-espace affine de de l'espace vectoriel des applications de  $J$  dans  $\mathbf{E}$  dérivables, dirigé par  $\mathcal{S}_{46}(J)$ .*

**Application du théorème de Cauchy-Lipschitz Linéaire** —

Soit  $\vec{\varphi}$  une solution de (46) définie sur  $I$ ,  $\vec{\psi}$  une solution de (45) définie sur  $I$ .

1. Soit  $\vec{\varphi}$  est nulle sur  $I$ , soit  $\vec{\varphi}$  ne s'annule pas.

2. On suppose de plus que  $I = \mathbf{R}$ , et que  $a$  et  $b$  sont impaires. Montrons que  $\vec{\psi}$  est paire.

3. On suppose encore que  $I = \mathbf{R}$ , et que  $a$  et  $b$  sont  $2\pi$ -périodiques. Montrer que  $\vec{\psi}$  est  $2\pi$ -périodiques si et

seulement si  $\vec{\psi}(0) = \vec{\psi}(2\pi)$ .

## 4.2 Etude de l'équation homogène

Au risque de décevoir, disons le tout net : à la différence de l'équation linéaire scalaire du premier ordre (cf. 2.1) où de l'équation linéaire vectorielle du premier ordre à coefficient constant, *on ne dispose pas de formules explicites pour la forme des solutions de (46)* (sauf si  $n = 1$ ).

	$\mathbf{E} = \mathbf{K}$	$\dim(\mathbf{E}) > 1$
coefficient constant	$\frac{dy}{dt} = ay; a \in \mathbf{K}$ $\mathbf{R} \rightarrow \mathbf{K}; t \mapsto c \exp(at), c \in \mathbf{K}$	$\frac{d\vec{y}}{dt} = a(\vec{y}); a \in \mathcal{L}(\mathbf{E})$ $\mathbf{R} \rightarrow \mathbf{E}; t \mapsto \exp(ta) \cdot \vec{c}, \vec{c} \in \mathbf{E}$
coefficient non constant	$\frac{dy}{dt} = a(t)y; a \in \mathcal{C}^0(I, \mathbf{K})$ $I \rightarrow \mathbf{K}; t \mapsto c \exp\left(\int a(t)dt\right), c \in \mathbf{K}$	$\frac{d\vec{y}}{dt} = a(t)(\vec{y}); a \in \mathcal{C}^0(I, \mathcal{L}(\mathbf{E}))$

Soit  $J$  un intervalle de  $I$  d'intérieur non vide. On ne sait pas résoudre en général (46), mais on sait que l'ensemble des solutions sur  $J$  est un espace vectoriel de dimension  $n$ . Donc si l'on dispose d'une famille libre d'éléments de  $\mathcal{S}_{46}(J)$ ,  $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ , l'ensemble des solutions de sur  $J$  de (46) est  $\text{vect}(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ , c'est-à-dire l'ensemble des applications de la forme :

$$J \rightarrow \mathbf{E}; t \mapsto c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) + \dots + c_n \vec{\varphi}_n(t),$$

avec  $c_1, c_2, \dots, c_n$  des éléments de  $\mathbf{K}$  quelconques. On voit donc toute l'importance de savoir déterminer si  $n$  solutions sont, ou ne sont pas indépendantes. On dispose du résultat suivant :

**Proposition 4.2.1.** — *Soit  $J$  un intervalle de  $I$  d'intérieur non vide et  $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$  des solutions de (46) sur  $J$ . Soit  $t_0$  un point de  $J$ . Alors  $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$  est une base de  $\mathcal{S}_{46}$  si et seulement si  $(\vec{\varphi}_1(t_0), \vec{\varphi}_2(t_0), \dots, \vec{\varphi}_n(t_0))$  est une base de  $\mathbf{E}$ .*

*Preuve de la proposition 4.2.1. —*

**Remarque** — Si  $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$  ne sont plus supposée être solutions de (46) sur  $J$ , mais seulement des applications de  $J$  dans  $\mathbf{E}$ , alors subsiste uniquement l'implication :

Pour tester la liberté en un point  $t_0$  de  $J$  d'une famille de  $n$  éléments de  $\mathcal{S}_{46}(J)$ , on introduit l'outil suivant

**Définition IV.2.2.** — WRONSKIEN —

*Soient  $J$  un intervalle de  $I$  d'intérieur non vide,  $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$  une famille d'éléments de  $\mathcal{S}_{46}(J)$  sur  $J$ , et  $\mathcal{B}$  une base de  $\mathbf{E}$ . On appelle wronskien dans la base  $\mathcal{B}$  du  $n$ -uplet de solution  $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ , l'application :*

$$W_{\mathcal{B}} : J \rightarrow \mathbf{K}; t \mapsto \det_{\mathcal{B}}(\vec{\varphi}_1(t), \vec{\varphi}_2(t), \dots, \vec{\varphi}_n(t)).$$

*Lorsque  $\mathbf{E}$  est  $\mathcal{M}_{n,1}(\mathbf{K})$ , on ne précise pas en générale la base  $\mathcal{B}$ ; qui par défaut; est la base canonique de  $\mathcal{M}_{n,1}(\mathbf{K})$ .*

Avec les notations de la définition IV.2.2., pour un élément  $t$  de  $I$ ,  $(\vec{\varphi}_1(t), \vec{\varphi}_2(t), \dots, \vec{\varphi}_n(t))$  est liée si et seulement si ...  
 . D'après IV.2.1., se présentent deux possibilités :

- Soit  $W_{\mathcal{B}}$  est l'application nulle sur  $J$  ;
- Soit  $W_{\mathcal{B}}$  ne s'annule pas sur  $J$ .

On peut en récapitulant, en déduire le résultat suivant :

**Proposition 4.2.3.** — *Avec les notations de la définition IV.2.2., on a l'équivalence des propositions suivantes :*

1. Pour tout élément  $t$  de  $J$ ,  $W_{\mathcal{B}}(t) \neq 0$ .
2. Il existe un élément  $t_0$ , de  $J$   $W_{\mathcal{B}}(t_0) \neq 0$ .
3. ...

### 4.3 Étude de l'équation avec second membre

Le théorème 4.1.3. assure que  $\mathcal{S}_{45}(J)$  est non vide. C'est donc un espace affine dirigé par  $\mathcal{S}_{46}(J)$ . Nous allons donner une méthode de résolution de (45) lorsque l'on connaît une base de solutions de (46).

#### CAS PARTICULIER

Pour commencer nous allons revenir quelques instants au cas constant. On considère les équations étudiées en 3.

$$\frac{d\vec{y}}{dt} = a(\vec{y}) + \vec{b}(t), \quad (49)$$

$$\frac{d\vec{y}}{dt} = a(\vec{y}), \quad (50)$$

$a$  étant un vecteur de  $\mathbf{E}$ . Pour simplifier l'écriture, prenons  $I = \mathbf{R}$ . Désignons par  $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$  une base de  $\mathcal{S}_{50}(\mathbf{R})$ . Notons pour  $i = 1, 2, \dots, n$ ,  $\vec{y}_i := \vec{\varphi}_i(0)$ . La solution  $\vec{\varphi}_i$  satisfait à la condition initiale  $\vec{y}(0) = \vec{y}_i$ , d'après 3.2.2, on a donc

$$\vec{\varphi}_i : \mathbf{R} \rightarrow \mathbf{E}; t \mapsto \exp(ta)(\vec{y}_i),$$

pour  $i = 1, 2, \dots, n$ . Notons que  $(\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n)$  est une base de  $\mathbf{E}$  en effet ...

Soit maintenant  $\vec{\varphi}$  une application de  $\mathbf{R}$  dans  $\mathbf{E}$  dérivable. Posons  $\vec{\psi} := \exp(-\cdot a(\vec{\varphi}))$ , de sorte que  $\vec{\varphi} = \exp(\cdot a)(\vec{\psi})$ . La méthode de résolution de (49), vue dans le 3.3 consiste à déterminer  $\vec{\psi}$ , pour que  $\vec{\varphi}$  soit solution de (49). Détaillons cette méthode

#### CAS GÉNÉRAL

On revient à l'étude de l'équation générale (45). On va s'inspirer de ce qui précède pour expliciter la méthode de résolution de (45) dite de « variation des constantes ». Cette terminologie est malheureuse, et l'oxymoron masque la nature de la démarche (45) : résoudre dans une base mobile de solution de (46). Cette méthode suppose donc la connaissance préalable, d'une base de solutions de (46).

Soient  $J$  un intervalle de  $I$  d'intérieur non vide et  $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$  une base de  $\mathcal{S}_{46}(J)$ . Pour tout élément  $t$  de  $J$ ,  $(\vec{\varphi}_1(t), \vec{\varphi}_2(t), \dots, \vec{\varphi}_n(t))$  est une base de  $\mathbf{E}$ , notée  $\mathcal{B}_t$ . Soit alors  $\vec{\varphi}$  une application dérivable de  $J$  dans  $\mathbf{E}$ . Notons pour  $i = 1, 2, \dots, n$  et tout élément  $t$  de  $J$ ,  $\psi_i(t)$  la  $i^{\text{e}}$  composante de  $\vec{\varphi}(t)$  dans  $\mathcal{B}_t$ .

commençons par un lemme technique :

**Lemme** — Soit  $\vec{g}$  une application de  $J$  dans  $\mathbf{E}$  de classe  $\mathcal{C}^k$ , où  $k = 0$  ou  $1$ . Notons pour  $i = 1, 2, \dots, n$  et tout élément  $t$  de  $J$ ,  $\gamma_i(t)$  la  $i^{\text{e}}$  composante de  $\vec{g}$  dans  $\mathcal{B}_t$ . Alors les applications  $\gamma_1, \dots, \gamma_n$  sont de classe  $\mathcal{C}^k$ .

*Preuve du lemme* — Soit  $\mathcal{B}$  une base de  $\mathbf{E}$ , désignons par  $\Gamma$  l'application de  $J$  dans  $\mathcal{M}_{n,1}(\mathbf{K})$ ,  $\Gamma = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$ , par  $p_{i,j}$  la  $i^{\text{e}}$  application composante de  $\vec{\varphi}_j$  dans  $\mathcal{B}$ , et  $g_i$  celle de  $\vec{g}$ , par  $P$  l'application de  $J$  dans  $\mathcal{M}_n(\mathbf{K})$ ,  $P : J \rightarrow \mathcal{M}_n(\mathbf{K})$ ;  $t \mapsto$

$(p_{i,j}(t))_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$  et  $G : J \rightarrow \mathcal{M}_{n,1}(\mathbf{K}) ; t \mapsto \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}$ . Ainsi pour tout élément  $t$  de  $J$  :

$$P(t)\Gamma(t) = G(t).$$

Notons que  $\vec{g}$  étant de classe  $\mathcal{C}^k$  ses composantes dans  $\mathcal{B}$  sont  $\mathcal{C}^k$  aussi, et de même  $\vec{\phi}_1, \dots, \vec{\phi}_n$  étant de classe  $\mathcal{C}^k$ , les fonctions coefficients de  $P$  sont de classe  $\mathcal{C}^k$ .

Notons que pour tout  $t \in J$ ,  $\det(P(t))$  est  $w_{\mathcal{B}}(t)$  valeur en  $t$  du wronskien de  $(\vec{\phi}_1, \dots, \vec{\phi}_n)$ , dans  $\mathcal{B}$ , donc le système d'inconnue  $X$ ,  $P(t)X = G(t)$ , est de Crammer et donc pour  $j = 1, 2, \dots, n$ ,  $\gamma_j(t)$  est donné par les célèbres formules de Crammer, précisément :

$$\gamma_j = \frac{\det \begin{pmatrix} p_{1,1} & \dots & p_{1,j-1} & g_1 & p_{1,j+1} & \dots & p_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ p_{n,1} & \dots & p_{n,j-1} & g_n & p_{n,j+1} & \dots & p_{n,n} \end{pmatrix}}{\det(P)},$$

pour  $j = 1, \dots, n$  et donc  $\gamma_j$  est de classe  $\mathcal{C}^k$ .

Le lemme est prouvé passons à la résolution de (45).

**Exemple** — Déterminer les solutions définies sur  $J$  avec ici  $J = ]-\frac{\pi}{2}, \frac{\pi}{2}[$  de

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \frac{t}{\cos\left(\frac{t^2}{2}\right)} \\ 0 \end{pmatrix} \quad (51)$$





## ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES SCALAIRES

Nous allons traiter des équations linéaires scalaires d'ordre  $n$ . Un premier paragraphe donnera les généralités sur ce typ d'équation et montrera comment elle se ramène à une équation linéaire du premier ordre.

Les paragraphes suivant seront consacrés au cas particulier de l'équation d'ordre 2, qui joue un rôle important dans les sciences, notamment en mécanique.

### 5.1 Équation scalaire d'ordre $n$

On désigne par  $I$  un intervalle d'intérieur non vide et  $n$  un entier naturel non nul.

**Définition 5.1.1.** — Soient  $a_n, a_{n-1}, \dots, a_0$  et  $d$  des applications de  $I$  dans  $\mathbf{K}$ , continues.

— On appelle solution de l'équation

$$a_n(t) \frac{d^n y}{dt^n} + \dots + a_1(t) \frac{dy}{dt} + a_0(t)y = d(t), \quad (52)$$

(équation différentielle linéaires scalaire d'ordre  $n$ )

toute application  $\varphi$  d'un intervalle  $J$  d'intérieur non vide, inclus dans  $I$ , à valeurs dans  $\mathbf{K}$ ,  $n$  fois dérivable, telle que :

$$a_n \varphi^{(n)} + \dots + a_1 \varphi' + a_0 \varphi = d.$$

— Soient  $t_0$  un élément de  $I$ ,  $y_0$  et  $y_1 \dots y_{n-1}$  des éléments de  $\mathbf{K}$ . On appelle solution du problème de Cauchy

$$\begin{cases} a_n(t) \frac{d^n y}{dt^n} + \dots + a_1(t) \frac{dy}{dt} + a_0(t)y = d(t), \\ \frac{d^i y}{dt^i}(t_0) = y_i, \text{ pour } i = 0, 1, \dots, n-1, \end{cases}$$

toute solution  $\varphi$  de (52) définie sur un intervalle  $J$  contenant  $t_0$  et telle que  $\varphi^{(i)}(t_0) = y_i$  pour  $i = 0, 1, \dots, n-1$ .

On dit encore qu'une telle solution satisfait à la condition initiale  $\frac{d^i y}{dt^i}(t_0) = y_i$ , pour  $i = 1, \dots, n$ .

— Dans le cas particulier où  $d$  est nulle l'équation qui devient

$$a_n(t) \frac{d^n y}{dt^n} + \dots + a_1(t) \frac{dy}{dt} + a_0(t)y = 0, \quad (53)$$

est dite homogène ou sans second membre. On dit que l'équation (53) est l'équation homogène associée à (52).

— Les applications  $a_i$ ,  $i = 0, \dots, n$  et  $d$  sont appelées coefficients de l'équation (19);  $d(t)$  est encore appelé second membre de l'équation.

**Remarques 5.1.2.** —

1. Lorsque l'on s'intéresse à des solutions sur un intervalle  $J$  donné, Le couple d'équations ((52),(53)) est un cas particulier du couple d'équations ((e),(h)), avec  $\mathbf{E} = \mathcal{D}^n(J, \mathbf{K})$ ,  $\mathbf{F} = \mathcal{F}(J, \mathbf{K})$ ,  $\vec{b} = d$  et

$$\ell : \mathcal{D}^n(J, \mathbf{K}) \rightarrow \mathcal{F}(J, \mathbf{K}); \varphi \mapsto a_n \varphi^{(n)} + \dots + a_1 \varphi' + a_0 \varphi.$$

Ceci explique la terminologie choisie.

2. On notera l'abus de notation, classique et préconisé par le programme, dans ces équations : la variable  $t$  ne figure pas dans l'inconnue mais dans les applications  $a_i$  et  $d$ . Cet abus n'est tolérable, que dans l'écriture de l'équation, jamais dans les égalités obtenues en y substituant une solution à l'inconnue.

Soit  $J$  un intervalle inclus dans  $I$  d'intérieur non vide. D'après la remarque 5.1.2.–1 l'ensemble des solutions de (53) définie sur  $J$ , noté  $\mathcal{S}_{53}$  est un sous espace vectoriel de  $\mathcal{D}^n(J, \mathbf{K})$ , l'ensemble des solutions de (52) définie sur  $J$ , noté  $\mathcal{S}_{52}$  est lui, soit le vide, soit un sous-espace affine dirigé par  $\mathcal{S}_{53}$ . On ne peut en dire plus. Pour aller plus loin il convient de se placer sur un intervalle sur lequel  $a_n$  ne s'annule pas.

Dans la suite de cette partie, sauf mention contraire,

$J$  désignera un intervalle inclus dans  $I$ , sur lequel  $a_n$  ne s'annule pas

On va ramener l'étude de (52) à celle d'un système linéaire du premier ordre. Soit le système différentiel

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} & & & \\ & & & \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \\ \\ \\ \end{pmatrix} \quad (54)$$

Soient  $y_0, \dots, y_{n-1}$  des éléments de  $\mathbf{K}$ , et  $t_0$  un élément de  $J$ .

— D'après 4.1.3 il existe une et une seule solution sur  $J$ ,  $\begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix}$ , du problème de Cauchy

$$\left\{ \begin{array}{l} (54), \\ \begin{pmatrix} x_1(t_0) \\ \vdots \\ x_n(t_0) \end{pmatrix} = \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix}. \end{array} \right. \quad (55)$$

L'application  $\varphi_1$  est  $n$  fois dérivable, en effet...

De plus  $\varphi_1$  est solution du problème de Cauchy

$$\left\{ \begin{array}{l} (52), \\ \frac{dy^i}{dt^i}(t_0) = y_i, \text{ pour } i = 0, \dots, n-1, \end{array} \right. \quad (56)$$

qui admet donc au moins une solution définie sur  $J$ .

— Réciproquement, Soit  $\varphi$  une solution sur  $J$  du problème de Cauchy (56) Alors  $\begin{pmatrix} \varphi \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix}$  est solution du problème de Cauchy (55) et donc  $\varphi = \varphi_1$ . En effet

En conclusion le problème de Cauchy (56), admet une et une seule solution sur  $J$ . En particulier, l'ensemble des solutions de (52) sur  $J$ , noté  $\mathcal{S}_{52}(J)$ , est non vide.

Envisageons le cas particulier de l'équation sans second membre (53). Le système (54) devient

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} & & & \\ & & & \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (57)$$

Soit  $\varphi$  une application de  $J$  dans  $\mathbf{K}$ , dérivable. Si  $\varphi$  est solution de (53) alors  $\begin{pmatrix} \varphi \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix}$  est solution de (57) sur  $J$ .

Donc l'application

$$\mathcal{K} : \mathcal{D}^n(J, \mathbf{K}) \rightarrow \mathcal{D}^1(J, \mathcal{M}_{n,1}(\mathbf{K})) ; \varphi \mapsto \begin{pmatrix} \varphi \\ \vdots \\ \varphi^{n-1} \end{pmatrix}$$

induit un isomorphisme  $\tilde{\mathcal{K}}$ , de l'espace vectoriel des solutions de (53) définies sur  $J$ , noté  $\mathcal{S}_{53}(J)$ , sur  $\mathcal{S}_{57}(J)$ , espace

vectorel des solutions de (57) sur  $J$ . En effet...

Donc  $\mathcal{S}_{53}(J)$  a comme dimension  $n$ .

Récapitulons :

**Proposition 5.1.3.** —

1. Pour tout élément  $t_0$  de  $J$ , tout  $(y_0, \dots, y_{n-1})$  éléments de  $\mathbf{K}$ , il existe une et une seule solution sur  $J$  du problème de Cauchy (56).

(Théorème de Cauchy-Lipschitz linéaire)

2. L'espace vectoriel des solutions de (53) définies sur  $J$  est un sous-espace vectoriel de  $\mathcal{D}^n(J, \mathbf{K})$  de dimension  $n$ , noté  $\mathcal{S}_{53}(J)$ .

3. L'ensemble des solutions de (52) définies sur  $J$  est un sous-espace affine de  $\mathcal{D}^n(J, \mathbf{K})$  dirigé par  $\mathcal{S}_{53}(J)$ .

**! La proposition 5.1.3. est fausse lorsque  $a_n$  s'annule sur  $J$ .**

## 5.2 Cas particulier de l'équation linéaire scalaire d'ordre 2

Nous allons maintenant nous intéresser à l'équation linéaire scalaire du deuxième ordre. L'ordre 2 est particulièrement important, en effet la définition même du temps en mécanique fait que les équations qu'on y rencontre sont souvent du deuxième d'ordre.

On désigne dans cette partie par  $I$  un intervalle d'intérieur non vide.

DÉFINITIONS, GÉNÉRALITÉS

On désigne par  $I$  un intervalle d'intérieur non vide.

On va considérer les cas particuliers des équations (52) et (59).

$$a(t) \frac{d^2 y}{dt^2} + b(t) \frac{dy}{dt} + c(t)y = d(t), \quad (58)$$

(équation différentielle linéaire scalaire du deuxième ordre)

$$a(t) \frac{d^2 y}{dt^2} + b(t) \frac{dy}{dt} + c(t)y = 0, \quad (59)$$

Dans la suite de cette partie, sauf mention contraire,

$J$  désignera un intervalle inclus dans  $I$ , sur lequel  $a$  ne s'annule pas

On a vu que l'on peut ramener l'étude de (58) à celle d'un système linéaire du premier ordre. Soit le système différentiel

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} & \\ & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \\ \end{pmatrix} \quad (60)$$

Rappelons ce qui vient d'être vu.

Soient  $y_0, y'_0$  des éléments de  $\mathbf{K}$ , et  $t_0$  un élément de  $J$ .

— D'après 4.1.3 il existe une et une seule solution sur  $J$ ,  $\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$ , du problème de Cauchy

$$\begin{cases} (60), \\ \begin{pmatrix} x_1(t_0) \\ x_2(t_0) \end{pmatrix} = \begin{pmatrix} y_0 \\ y'_0 \end{pmatrix}. \end{cases} \quad (61)$$

L'application  $\varphi_1$  est deux fois dérivable, et est solution du problème de Cauchy

$$\begin{cases} (58), \\ y(t_0) = y_0, \frac{dy}{dt}(t_0) = y'_0, \end{cases} \quad (62)$$

qui admet donc au moins une solution définie sur  $J$ .

— Réciproquement, Soit  $\varphi$  une solution sur  $J$  du problème de Cauchy (62) Alors  $\begin{pmatrix} \varphi \\ \varphi' \end{pmatrix}$  est solution du problème de Cauchy (61) et donc  $\varphi = \varphi_1$ .

En conclusion le problème de Cauchy (62), admet une et une seule solution sur  $J$ . En particulier, l'ensemble des solutions de (58) sur  $J$ , noté  $\mathcal{S}_{58}(J)$ , est non vide.

Envisageons le cas particulier de l'équation sans second membre (59). Le système (60) devient

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} & \\ & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (63)$$

Soit  $\varphi$  une application de  $J$  dans  $\mathbf{K}$ , dérivable. Si  $\varphi$  est solution de (59) alors  $\begin{pmatrix} \varphi \\ \varphi' \end{pmatrix}$  est solution de (63) sur  $J$ . Donc l'application

$$\mathcal{K} : \mathcal{D}^2(J, \mathbf{K}) \rightarrow \mathcal{D}^1(J, \mathcal{M}_{2,1}(\mathbf{K})) ; \varphi \mapsto \begin{pmatrix} \varphi \\ \varphi' \end{pmatrix}$$

induit un isomorphisme  $\tilde{\mathcal{K}}$ , de l'espace vectoriel des solutions de (59) définies sur  $J$ , noté  $\mathcal{S}_{59}(J)$ , sur  $\mathcal{S}_{63}(J)$ , espace vectoriel des solutions de (63) sur  $J$ .

Donc  $\mathcal{S}_{59}(J)$  a comme dimension 2.

Récapitulons :

**Proposition 5.2.1.** —

1. Pour tout élément  $t_0$  de  $J$ , tout  $y_0$  et tout  $y'_0$  éléments de  $\mathbf{K}$ , il existe une et une seule solution sur  $J$  du problème de Cauchy (62).

(Théorème de Cauchy-Lipschitz linéaire)

2. L'espace vectoriel des solutions de (59) définies sur  $J$  est un sous-espace vectoriel de  $\mathcal{D}^2(J, \mathbf{K})$  de dimension 2, noté  $\mathcal{S}_{59}(J)$ .
3. L'ensemble des solutions de (58) définies sur  $J$  est un sous-espace affine de  $\mathcal{D}^2(J, \mathbf{K})$  dirigé par  $\mathcal{S}_{59}(J)$ .

**! La proposition 5.1.3. est fausse lorsque  $a$  s'annule sur  $J$ .**

Nous donnerons dans le dernier paragraphe des exemples d'études de solutions sur un intervalle où  $a$  s'annule.

**Remarques 5.2.2.** —

1. D'après la remarque 4.1.2.-2, on peut montrer que toute solution de (58) est *a priori* de classe  $\mathcal{C}^2$ .
2. Comme dans le cas de l'équation vectorielle du premier ordre la solution sur  $J$  d'un problème de Cauchy homogène dépend linéairement de  $(y_0, y'_0)$ . Précisément en notant  $\varphi_{y_0, y'_0}$  la solution sur  $J$  de

$$\begin{cases} (59), \\ y(t_0) = y_0, \frac{dy}{dt}(t_0) = y'_0, \end{cases} ,$$

les applications

$$\mathcal{I} : \mathbf{K}^2 \rightarrow \mathcal{S}_{59}(J) ; (y_0, y'_0) \mapsto \varphi_{y_0, y'_0}$$

et

$$\mathcal{J} : \mathcal{S}_{59}(J) \rightarrow \mathbf{K}^2 ; \varphi \mapsto (\varphi(t_0), \varphi'(t_0)),$$

sont des isomorphismes réciproques l'un de l'autre. On peut prouver ce résultat directement, où utiliser IV.1.4.-2 et l'isomorphisme  $\tilde{\mathcal{K}}$ .

Donnons quelques applications classiques du théorème de Cauchy-Lipschitz linéaire.

**Application du théorème de Cauchy-Lipschitz Linéaire —**

Soit  $\varphi$  une solution de (59) définie sur  $J$ ,  $\vec{\psi}$  une solution de (58) définie sur  $J$ .

1. Soit  $t_0$  un point de  $J$ . On suppose que  $\varphi(t_0) = 0$  et  $\varphi'(t_0) = 0$ . Montrons que  $\varphi$  est nulle.

2. On suppose que  $J = \mathbf{R}$ , et que  $a$  et  $b$   $c$  et  $d$  sont  $2\pi$ -périodiques. Montrer que  $\psi$  est  $2\pi$ -périodiques si et seulement si  $\psi(0) = \psi(2\pi)$  et  $\psi'(0) = \psi'(2\pi)$ .

Nous avons vu que la résolution de (59) se ramène à celle du système linéaire d'ordre 1 (63), or, nous l'avons dit au IV.2, on ne dispose pas d'expression générale de la solution d'un tel système. On ne dispose donc pas d'avantage d'une expression de la solution de (58). On sait simplement que si  $\varphi_1$  et  $\varphi_2$  sont des solutions sur  $J$  de (58), indépendante, alors l'espace vectoriel des solutions sur  $J$  de (58) est le plan engendré par  $(\varphi_1, \varphi_2)$  c'est-à-dire l'ensemble des applications de la forme :

$$\varphi : J \rightarrow \mathbf{K}; t \mapsto c_1\varphi_1 + c_2\varphi_2,$$

où  $c_1$  et  $c_2$  sont des éléments quelconques de  $\mathbf{K}$ .

Il est donc important de savoir si des solutions  $\varphi_1$  et  $\varphi_2$  sont indépendantes, ou comme on dit si  $(\varphi_1, \varphi_2)$  forme un *système fondamental de solutions*. Pour ce faire on a besoin de la définition suivante.

**Définition 5.2.3.** — WRONSKIEN —

Soit  $\varphi_1$  et  $\varphi_2$  des solutions de (59) définies sur  $J$ . On appelle wronskien du couple  $(\varphi_1, \varphi_2)$ , l'application

$$w : J \rightarrow \mathbf{K}; t \mapsto \begin{vmatrix} \varphi_1(t) & \varphi_2(t) \\ \varphi_1'(t) & \varphi_2'(t) \end{vmatrix}$$

L'application  $w$  s'appelle encore wronskien de  $(\varphi_1, \varphi_2)$ .

Avec les notations de la définition on a le résultat :

**Proposition 5.2.4.** — Les trois propositions suivantes sont équivalentes :

1. Le couple  $(\varphi_1, \varphi_2)$  forme un système fondamental de solutions.
2. Il existe un élément  $t_0$  de  $J$  tel que  $w(t_0) \neq 0$ .
3. Pour tout élément  $t$  de  $J$ ,  $w(t) \neq 0$ .

*Preuve de la proposition 5.2.4.* —

Avant d'étudier sur des exemples diverses méthodes de détermination d'un système fondamental de solutions, nous allons étudier l'équation avec second membre.

#### ÉTUDE DE L'ÉQUATION AVEC UN SECOND MEMBRE

Nous considérons toujours un intervalle  $J$  sur lequel  $a$  ne s'annule pas. Alors, comme nous l'avons dit dans le paragraphe précédent,  $\mathcal{S}_{58}$  est espace affine de dimension 2 dirigé par  $\mathcal{S}_{59}$ . Nous allons nous intéresser à la résolution de (58), dans un premier temps en supposant disposer d'un système fondamental de solutions de l'équation homogène, puis dans un second temps en supposant ne disposer que d'une solution de l'équation homogène.

On suppose que  $(\varphi_1, \varphi_2)$  est un système fondamental de solutions de (59) définies sur  $J$ . On a vu que résoudre (58) revient à résoudre le système (60) et que  $\left( \begin{pmatrix} \varphi_1 \\ \varphi_1' \end{pmatrix}, \begin{pmatrix} \varphi_2 \\ \varphi_2' \end{pmatrix} \right)$  est une base de solution du système (63), système homogène associé à (60). On pourrait se contenter de cette remarque, puisque la méthode de résolution de (60), connue sous le

nom de variation des constantes, a été explicitée au IV.4.3. Néanmoins, nous allons transcrire soigneusement<sup>1</sup> cette méthode, dans le cas particulier de l'équation (58), afin de pouvoir la rédiger directement et rapidement.

le wronskien de  $(\varphi_1, \varphi_2)$  qui est celui de  $\left(\begin{pmatrix} \varphi_1 \\ \varphi_1' \end{pmatrix}, \begin{pmatrix} \varphi_2 \\ \varphi_2' \end{pmatrix}\right)$  dans la base canonique de  $\mathcal{M}_{2,1}(\mathbf{K})$ , ne s'annule pas. il en résulte que pour tout élément  $t$  de  $J$ ,  $\left(\begin{pmatrix} \varphi_1(t) \\ \varphi_1'(t) \end{pmatrix}, \begin{pmatrix} \varphi_2(t) \\ \varphi_2'(t) \end{pmatrix}\right)$  est une base de  $\mathcal{M}_{2,1}(\mathbf{K})$ , que nous noterons  $\mathcal{B}_t$ .

Soit alors  $\varphi$  une application de  $J$  dans  $\mathbf{K}$  deux fois dérivable. Notons pour tout élément  $t$  de  $J$   $(\psi_1(t), \psi_2(t))$  le couple des coordonnées de  $\begin{pmatrix} \varphi(t) \\ \varphi'(t) \end{pmatrix}$  dans  $\mathcal{B}_t$ . Nous avons vu que les applications  $\psi_1$  et  $\psi_2$  ainsi définies de  $J$  dans  $\mathbf{K}$  sont dérivables, en conclusion on a le résultat suivant :

*Il existe un unique couple  $(\psi_1, \psi_2)$  d'éléments de  $\mathcal{D}^1(J, \mathbf{K})^2$  tel que :*

$$\varphi = \psi_1 \varphi_1 + \psi_2 \varphi_2, \quad (64)$$

$$\varphi' = \psi_1 \varphi_1' + \psi_2 \varphi_2'. \quad (65)$$

Remarquons tout de suite qu'en dérivant (64) et en retranchant (65) on obtient

$$\psi_1' \varphi_1 + \psi_2' \varphi_2 = 0. \quad (66)$$

1. Il s'agit d'une des questions les plus mal traitées du programme, et la littérature foisonne d'acrobaties invraisemblables et à la limite malhonnêtes,

2. ou de  $\mathcal{C}^1(J, \mathbf{K})$ , cf remarque IV.1.2.-2





**Exemple** — Etudions les solutions définies sur  $I = ]-\frac{\pi}{2}, \frac{\pi}{2}[$ , à valeurs réelles, de l'équation différentielle

$$\frac{d^2y}{dt^2} + y = \frac{1}{\cos(t)} \quad (67)$$

Le coefficient de  $\frac{d^2y}{dt^2}$  ne s'annule pas sur  $I$ . Les applications  $\varphi_1 : I \rightarrow \mathbf{R}; t \mapsto \cos(t)$ ,  $\varphi_2 : I \rightarrow \mathbf{R}; t \mapsto \sin(t)$  sont solutions de (67), elles sont de plus indépendantes puisque  $\begin{vmatrix} \varphi_1(0) & \varphi_2(0) \\ \varphi_1'(0) & \varphi_2'(0) \end{vmatrix} = 1 \neq 0$ .

Soit  $\varphi$  un élément de  $\mathcal{C}^2(I, \mathbf{R})$ . D'après le cours, il existe des éléments de  $\mathcal{C}^1(I, \mathbf{R})$ ,  $\psi_1$  et  $\psi_2$ , (composantes de  $\begin{pmatrix} \varphi \\ \varphi' \end{pmatrix}$ ) dans la « base mobile »  $t \mapsto \left( \begin{pmatrix} \varphi_1 \\ \varphi_1' \end{pmatrix}, \begin{pmatrix} \varphi_2 \\ \varphi_2' \end{pmatrix} \right)$

#### COMPLÉMENT : MÉTHODES PRATIQUES DE RÉOLUTION

Nous n'aurons pas toujours la chance de disposer d'un système fondamental de solutions de l'équation (59). On va donner quelques exemples de méthode de résolution.

- Il se peut que l'on ne dispose que d'une solution  $\varphi_1$  de (59) qui *ne s'annule pas* sur  $J$ .

Soit alors  $\varphi$  un élément de  $\mathcal{D}^2(J, \mathbf{K})$ . La non annulation de  $\varphi_1$ , permet de poser  $\psi_1 := \frac{\varphi}{\varphi_1}$ , de sorte que  $\varphi = \psi_1 \varphi_1$ . On peut alors chercher une condition sur  $\psi$ , nécessaire et suffisante pour que  $\varphi$  soit solution

- On peut chercher un changement de variables  $\ll \tau = \chi(t) \gg$ , avec  $\chi$   $\mathcal{C}^2$ -difféomorphisme, qui ramène l'équation à une équation linéaire du second ordre à coefficients constants ;
- On peut chercher un changement de fonctions inconnues de la forme  $z = \lambda y$ , avec  $\lambda$  une application de classe  $\mathcal{C}^2$  qui ne s'annule pas ;
- Dans le cas où les coefficients de l'équation s'y prêtent (coefficients polynômiaux) on peut chercher successivement :
  - Une solution monôme,
  - Une solution de la forme  $t \mapsto t^\alpha$ ,  $\alpha$  réel,
  - Une solution polynomiale,
  - Une solution somme d'une série entière,
  - Une solution produit de la somme d'une série entière est d'une fonction puissance,  $t \mapsto t^\alpha \sum_{n=0}^{+\infty} a_n t^n$ ,  $\alpha$  réel, (ce qui permet lorsque le coefficient de  $\frac{d^2 y}{dt^2}$  s'annule d'obtenir des solutions qui ne sont pas  $\mathcal{C}^\infty$  au voisinage de 0) ;
- Compter sur l'énoncé ou, à l'oral, sur la gentillesse de l'examineur.

Pour finir, un mot sur la résolution de (58) sur l'intervalle  $I$  dans le cas  $a$  s'annule. On l'a dit,  $\mathcal{S}_{59}(I)$  est un espace vectoriel, mais la dimension n'est pas nécessairement 2. Quant à  $\mathcal{S}_{58}(I)$ , c'est soit le vide, soit un espace affine, les deux cas étant possibles. Dans la pratique, on procède comme pour l'équation scalaire du premier ordre. Supposons que  $a$  s'annule sur  $I$  en  $n$  points. On résout d'abord l'équation sur les intervalles ouverts d'extrémités deux de ces points successifs (il y en a au maximum  $n+1$ ). On s'intéresse ensuite aux solutions sur  $I$  en utilisant la forme qu'une éventuelle solution prend sur chaque intervalle du type précédent. Dans ce cas *a priori*, la dimension de l'espace vectoriel  $\mathcal{S}_{59}(I)$  est comprise au sens large entre  $\dots$

Illustrons tout cela de quelques beaux exemples.

### Exemples 5.2.6. —

1. Etudions les solutions réelles, définies sur  $] -1, +\infty]$ , de l'équation

$$(t+1)\frac{d^2 y}{dt^2} - 2\frac{dy}{dt} - (t-1)y = t \exp(-t). \quad (69)$$

On pourra chercher une solution de l'équation homogène associée, de la forme

$$]-1, +\infty] \rightarrow \mathbf{R}; \quad t \mapsto p(t) \exp(t),$$

avec  $p$  une application polynômiale.

2. Soit  $m$  un réel. Déterminons les solutions réelles définies sur  $\mathbf{R}$  de l'équation

$$(1+t^2)^2 \frac{d^2 y}{dt^2} + 2t(1+t^2) \frac{dy}{dt} + my = 0. \quad (70)$$

On va chercher un changement de variables qui conduit à une équation linéaire à coefficients constants.

3. Déterminons les solutions réelles définies sur  $\mathbf{R}_+^*$  de l'équation

$$t^2 \frac{d^2 y}{dt^2} + 3t \frac{dy}{dt} + y = 0. \quad (71)$$

(équation d'Euler<sup>1</sup>)

---

1. On appelle équation d'Euler d'ordre 2, toute équation  $\alpha t^2 \frac{d^2 y}{dt^2} + \beta t \frac{dy}{dt} + \gamma y = 0$ , où  $\alpha, \beta, \gamma$  sont des réels.



4. Cherchons les solutions réelles définies sur  $\mathbf{R}_+^*$  de l'équation

$$t^2 \frac{d^2 y}{dt^2} + t \frac{dy}{dt} + \left( t^2 - \frac{1}{4} \right) y = 0. \quad (72)$$

On va chercher des solutions de la forme

$$]0, R[ \rightarrow \mathbf{R}; \quad t \mapsto t^\alpha \sum_{n=0}^{+\infty} a_n t^n,$$

avec  $\alpha$  un réel et  $\sum a_n t^n$  une série entière de rayon de convergence  $R > 0$ .

5. Reprenons l'équation (69).

$$(t+1)\frac{d^2y}{dt^2} - 2\frac{dy}{dt} - (t-1)y = t\exp(-t).$$

Nous allons chercher les solutions réelles, définies sur  $\mathbf{R}$  de cette équation, et nous intéresser à l'étude d'un problème de Cauchy associé.