RAPPELS DE COURS

Chapitre 1 - Espaces vectoriels et applications linéaires

Dans toute la suite, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

1.1 Définition d'un espace vectoriel

- Un $\mathbb{K}$-espace vectoriel est un triplet $(E, +, .)$, où E est un ensemble muni :
  - d'une loi de composition interne sur E, notée $+$, telle que $(E,+)$ soit un groupe abélien, c'est-à-dire :
    (i) la loi $+$ est une application de $E \times E$ dans E (loi interne) :
    $\forall\,(x,y) \in E^2,\ x+y \in E$;
    (ii) la loi $+$ est associative : $\forall\,(x,y,z) \in E^3,\ (x+y)+z = x+(y+z)$;
    (iii) la loi $+$ est commutative : $\forall\,(x,y) \in E^2,\ x+y = y+x$;
    (iv) la loi $+$ possède un élément neutre, noté 0_E : $\forall\, x \in E,\ x+0_E = x$;
    (v) tout élément x de E possède un symétrique pour la loi $+$, noté $-x$: $x+(-x) = 0_E$.
  - d'une loi de composition externe sur $\mathbb{K}$, notée $.$, c'est-à-dire une application $\begin{cases} \mathbb{K} \times E & \longrightarrow & E \\ (\lambda,x) & \longmapsto & \lambda.x \end{cases}$ vérifiant les propriétés suivantes :

$$\forall\,(\lambda,\mu) \in \mathbb{K}^2,\ \forall\,(x,y) \in E^2,\ \begin{cases} \text{(i)} & 1_{\mathbb{K}}.x = x\,; \\ \text{(ii)} & \lambda.(x+y) = \lambda.x + \lambda.y\,; \\ \text{(iii)} & (\lambda+\mu).x = \lambda.x + \mu.x\,; \\ \text{(iv)} & \lambda.(\mu.x) = (\lambda\mu).x\,. \end{cases}$$

- **Exemples**

  Il est important de connaître un certain nombre d'espaces vectoriels classiques.
  - $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel.
  - $\mathbb{C}$ est un $\mathbb{R}$-espace vectoriel pour les lois usuelles.
  - Si D est un ensemble quelconque et E un $\mathbb{K}$-espace vectoriel, l'ensemble $\mathcal{A}(D,E)$ des applications de D dans E peut être muni d'une structure de $\mathbb{K}$-espace vectoriel pour les lois $+$ et $.$ définies par :
    $(f,g) \mapsto f+g$ avec : $\forall\, x \in D,\ (f+g)(x) = f(x)+g(x)$;
    $(\lambda,f) \mapsto \lambda.f$ avec : $\forall\, x \in D,\ (\lambda.f)(x) = \lambda.f(x)$.
    En particulier, l'ensemble $E^{\mathbb{N}}$ des suites à valeurs dans E est un $\mathbb{K}$-espace vectoriel.
  - $\mathbb{K}[X]$, ensemble des polynômes à coefficients dans $\mathbb{K}$, est un $\mathbb{K}$-espace vectoriel.
  - $\mathcal{M}_{p,q}(\mathbb{K})$, ensemble des matrices à p lignes et q colonnes à coefficients dans $\mathbb{K}$, est un $\mathbb{K}$-espace vectoriel.

1.2 Espace vectoriel produit

Soient $E_1, E_2, \ldots, E_p$ p $\mathbb{K}$-espaces vectoriels.
L'ensemble produit $E = E_1 \times E_2 \times \ldots \times E_p$ est un $\mathbb{K}$-espace vectoriel pour les lois définies par :

$$\begin{aligned} (x_1,x_2,\ldots,x_p) + (y_1,y_2,\ldots,y_p) &= (x_1+y_1, x_2+y_2, \ldots, x_p+y_p)\,; \\ \lambda.(x_1,x_2,\ldots,x_p) &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_p)\,. \end{aligned}$$

Muni de ces lois, E s'appelle l'espace vectoriel produit des E_i. Le vecteur nul de E est le p-uplet $(0_{E_1},\ldots,0_{E_p})$.

Exemple : $\mathbb{K}^p$ est un $\mathbb{K}$-espace vectoriel pour les lois définies ci-dessus.

1.3 Sous-espace vectoriel

- Soit $(E, +, .)$ un $\mathbb{K}$-espace vectoriel. Par définition, une *partie* F de E est un sous-espace vectoriel de E si $(F, +, .)$ est encore un $\mathbb{K}$-espace vectoriel.

- **Caractérisations d'un sous-espace vectoriel**
  F est un sous-espace vectoriel de E si et seulement si F est une partie non vide de E stable par combinaisons linéaires, ce qui équivaut à :

$$F \subset E, \quad F \neq \emptyset \quad \text{et} \quad \forall \lambda \in \mathbb{K}, \forall (x,y) \in F^2 \ : \ x+y \in F \text{ et } \lambda x \in F$$

  ou encore :

$$F \subset E, \quad F \neq \emptyset \quad \text{et} \quad \forall \lambda \in \mathbb{K}, \forall (x,y) \in F^2 \ : \ \lambda x + y \ \in \ F\,.$$

  ✓ Il ne faut surtout pas oublier de vérifier que F est bien inclus dans E. Dans certains cas, il s'agit de la partie la plus délicate de la vérification.

  ✓ Pour montrer $F \neq \emptyset$, on vérifie en général que $0_E \in F$.

- **Exemples**
  Il existe un certain nombre de sous-espaces vectoriels classiques (donc d'espaces vectoriels) à connaître.
  - $\{0_E\}$ et E sont des sous-espaces vectoriels de E (dits triviaux).
  - Si $n \in \mathbb{N}$, $\mathbb{K}_n[X]$, ensemble des polynômes de $\mathbb{K}[X]$ de degré inférieur ou égal à n, est un sous-espace vectoriel de $\mathbb{K}[X]$.

    ✓ Attention : l'ensemble des polynômes de degré *exactement* n n'est PAS un sous-espace vectoriel de $\mathbb{K}[X]$; en effet, il ne contient même pas le polynôme nul !
  - Si $P \in \mathbb{K}[X]$, l'ensemble $\mathbb{K}[X] \cdot P$ des multiples de P est un sous-espace vectoriel de $\mathbb{K}[X]$.
  - Si I est un intervalle de $\mathbb{R}$, l'ensemble $\mathscr{C}(I,\mathbb{R})$ des fonctions continues sur I à valeurs réelles est un sous-espace vectoriel de $\mathcal{A}(I,\mathbb{R})$.

1.4 Intersection de sous-espaces vectoriels

Si $(E_i)_{i\in I}$ est une famille quelconque de sous-espaces vectoriels de E, leur intersection $\bigcap_{i\in I} E_i$ est encore un sous-espace vectoriel de E.

✓ La *réunion* de sous-espaces vectoriels de E n'est pas, en général, un sous-espace vectoriel de E.
Plus précisément, on peut démontrer que si A et B sont deux sous-espaces vectoriels d'un espace vectoriel E, $A \cup B$ est encore un sous-espace vectoriel de E si et seulement si $A \subset B$ ou $B \subset A$.

1.5 Sous-espace vectoriel engendré, combinaisons linéaires

- Soit X une partie quelconque d'un $\mathbb{K}$-espace vectoriel E.
  L'intersection de tous les sous-espaces vectoriels de E qui contiennent X est encore un sous-espace vectoriel de E ; c'est en fait le plus petit sous-espace vectoriel de E contenant X (au sens de l'inclusion) ; on l'appelle sous-espace vectoriel engendré par X, et on le note $\mathrm{Vect}(X)$.
- Si $X = \emptyset$, $\mathrm{Vect}(\emptyset) = \{0_E\}$. Sinon, $\mathrm{Vect}(X)$ est exactement l'ensemble des combinaisons linéaires des éléments de X, c'est-à-dire les éléments de la forme :

$$\sum_{i=1}^{n} \lambda_i x_i \quad \text{avec } n \in \mathbb{N}^* \text{ et } \lambda_i \in \mathbb{K},\ x_i \in X \text{ pour tout } i\,.$$

  ✓ Par définition, une combinaison linéaire est toujours une somme *finie*.

  ✓ Le résultat précédent montre, en particulier, que l'ensemble des combinaisons linéaires des vecteurs d'une partie de E est un sous-espace vectoriel de E. Cela peut être utile pour abréger certaines démonstrations.

1.6 Famille génératrice

Une famille $(x_i)_{i\in I}$ d'un $\mathbb{K}$-espace vectoriel E est dite génératrice de E si le sous-espace vectoriel qu'elle engendre est égal à l'espace E tout entier :

$$E = \mathrm{Vect}\big((x_i)_{i\in I}\big)\,,$$

c'est-à-dire si tout vecteur de E est combinaison linéaire des x_i .

Toute famille contenant une famille génératrice de E est encore génératrice de E.

1.7 Dépendance et indépendance linéaire

- Une famille finie $(x_i)_{1\leqslant i\leqslant n}$ d'un $\mathbb{K}$-espace vectoriel E est dite libre s'il n'existe pas de relation de dépendance linéaire non triviale entre ces vecteurs, c'est-à-dire si, pour toute famille $(\lambda_i)_{1\leqslant i\leqslant n}$ d'éléments de $\mathbb{K}$:

$$\sum_{i=1}^{n}\lambda_i x_i = 0_E \implies \forall\, i\in[\![1\,;n]\!],\ \lambda_i = 0_{\mathbb{K}}$$

  (on dit aussi que les x_i sont linéairement indépendants).
- Une famille quelconque $(x_i)_{i\in I}$ d'un $\mathbb{K}$-espace vectoriel E est dite libre si *toutes* ses sous-familles finies sont libres.
  Exemple : dans $\mathbb{K}[X]$, toute famille de polynômes non nuls de degrés distincts est libre.
- Une famille finie $(x_i)_{1\leqslant i\leqslant n}$ d'un $\mathbb{K}$-espace vectoriel E est dite liée si elle n'est pas libre, c'est-à-dire s'il existe une famille $(\lambda_i)_{1\leqslant i\leqslant n}$ d'éléments de $\mathbb{K}$, *non tous nuls*, telle que $\sum_{i=1}^{n}\lambda_i x_i = 0_E$ (on dit aussi que les x_i sont linéairement dépendants).
- **Propriétés**
  - Une famille réduite à un élément $\{x\}$ est libre si et seulement si $x\neq 0_E$.
  - Toute famille contenant 0_E est liée.
  - Toute sous-famille d'une famille libre est libre. Toute famille contenant une famille liée est liée.
  - Une famille $(x_i)_{i\in I}$ est liée si et seulement si il existe $j\in I$ tel que x_j soit combinaison linéaire des $(x_i)_{i\in I, i\neq j}$.
    Une famille $(x_i)_{i\in I}$ est libre si aucun des x_i n'est combinaison linéaire des autres.

1.8 Bases

Une famille $\mathscr{B}$ d'un $\mathbb{K}$-espace vectoriel E s'appelle une base de E si elle est à la fois libre et génératrice de E.

Cela équivaut à dire que tout vecteur x de E s'écrit de manière unique comme combinaison linéaire des éléments de $\mathscr{B}$. Les coefficients de cette combinaison linéaire sont les coordonnées de x dans la base $\mathscr{B}$.

1.9 Somme de sous espaces vectoriels

Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de n sous-espaces vectoriels de E ($n\in\mathbb{N}^*$) .

On appelle somme des E_i, et on note $\sum_{i=1}^{n}E_i$, l'ensemble des vecteurs de la forme $\sum_{i=1}^{n}x_i$, où $x_i\in E_i$ pour tout $i\in[\![1\,;n]\!]$.

C'est aussi le sous-espace vectoriel de E engendré par la réunion des E_i, c'est-à-dire le plus petit sous-espace vectoriel de E qui contient tous les E_i :

$$\sum_{i=1}^{n}E_i = \mathrm{Vect}\Big(\bigcup_{1\leqslant i\leqslant n}E_i\Big).$$

> ✓ Les deux caractérisations précédentes de la somme de sous-espaces sont utiles et doivent toutes deux être sues.
> Cependant, on prendra soin de ne pas confondre la *réunion*, qui est une notion ensembliste, et la *somme* de sous-espaces vectoriels.

Exemple : si $x_1,\ldots,x_n$ sont n vecteurs de E, $\sum_{i=1}^{n}\mathbb{K}.x_i = \mathrm{Vect}(x_1,\ldots,x_n)$.

1.10 Somme directe de sous-espaces vectoriels

- Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de n sous-espaces vectoriels de E.
  On dit que la somme $\sum_{i=1}^{n}E_i$ est directe lorsque pour tout $x\in\sum_{i=1}^{n}E_i$ l'écriture $x=\sum_{i=1}^{n}x_i$ avec $x_i\in E_i$ pour tout i est *unique*. On note alors : $\sum_{i=1}^{n}E_i = \bigoplus_{i=1}^{n}E_i$.

Exemple : si $x_1, \dots, x_n$ sont n vecteurs *non nuls* de E, dire que la somme $\sum_{i=1}^{n} \mathbb{K}.x_i$ est directe équivaut à dire que la famille $(x_1, \dots, x_n)$ est libre.

- **Caractérisations d'une somme directe**
  - Pour que la somme $\sum_{1 \leqslant i \leqslant n} E_i$ soit directe, il faut et il suffit que pour toute famille $(x_i)_{1 \leqslant i \leqslant n}$ avec $x_i \in E_i$ pour tout i on ait :
  $$\sum_{i=1}^{n} x_i = 0 \implies \forall\, i \in [\![1\,;n]\!],\ x_i = 0\,.$$
  ✓ Il s'agit de la caractérisation la plus simple à utiliser pour montrer que n sous-espaces vectoriels sont en somme directe dès que $n \geqslant 3$.
  - Pour que la somme $\sum_{i=1}^{n} E_i$ soit directe, il faut et il suffit qu'il existe $j \in [\![1\,;n]\!]$ tel que $\sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} E_i$ soit directe et que $\sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} E_i$ et E_j soient en somme directe (et dans ce cas, cette propriété est vraie pour tout $j \in [\![1\,;n]\!]$).
  ✓ Cette caractérisation est moins utilisée, mais elle peut être pratique dans des démonstrations par récurrence.
- **Somme directe de deux sous-espaces**
  Dans le cas de *deux* sous-espaces vectoriels (et dans ce cas seulement), on a une caractérisation plus simple. En effet, pour que la somme de deux sous-espaces vectoriels E_1 et E_2 de E soit directe, il faut et il suffit que $E_1 \cap E_2 = \{0_E\}$.
- **Sous-espaces supplémentaires**
  Les sous-espaces vectoriels E_i $(1 \leqslant i \leqslant n)$ de E sont dits supplémentaires si $E = \bigoplus_{i=1}^{n} E_i$. Cela équivaut à dire que tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = \sum_{i=1}^{n} x_i$ avec $x_i \in E_i$ pour tout i.
  Exemple : si $x_1, \dots, x_n$ sont n vecteurs *non nuls* de E, dire que $E = \bigoplus_{i=1}^{n} \mathbb{K}.x_i$ équivaut à dire que la famille $(x_1, \dots, x_n)$ est une base de E.
- **Base adaptée à une décomposition en somme directe**
  - Soit $(E_i)_{1 \leqslant i \leqslant n}$ une famille de sous-espaces vectoriels de E telle que $E = \bigoplus_{i=1}^{n} E_i$.
  Si, pour tout $i \in [\![1\,;n]\!]$, $\mathscr{B}_i$ est une base de E_i, alors les $\mathscr{B}_i$ sont deux à deux disjointes et $\mathscr{B} = \bigcup_{i=1}^{n} \mathscr{B}_i$ est une base de E.
  - Soit $\mathscr{B}$ une base de E, et $(\mathscr{B}_i)_{1 \leqslant i \leqslant n}$ une partition de $\mathscr{B}$.
  Si, pour tout $i \in [\![1\,;n]\!]$, $E_i = \mathrm{Vect}(\mathscr{B}_i)$, alors $E = \bigoplus_{i=1}^{n} E_i$.

1.11 Espaces vectoriels de dimension finie

- On dit qu'un $\mathbb{K}$-espace vectoriel est de dimension finie s'il possède une famille génératrice finie.
- **Théorème de la base incomplète**
  Soient L une famille libre et G une famille génératrice de E, espace vectoriel de dimension finie.
  Alors il existe une base $\mathscr{B}$ de E telle que : $L \subset \mathscr{B} \subset L \cup G$ (*autrement dit, on peut compléter L à l'aide de vecteurs de G pour obtenir une base*).
  On en déduit que tout espace vectoriel de dimension finie E possède (au moins) une base (dans le cas où $E = \{0\}$, on peut convenir qu'une base de E est $\emptyset$).
- **Dimension**
  Toutes les bases d'un espace vectoriel E de dimension finie ont le même nombre d'éléments, appelé la dimension de E, notée $\dim_{\mathbb{K}} E$ ou plus simplement $\dim E$.

- **Caractérisation des bases**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension n.
  - Toute famille libre de E a au plus n éléments ; et si elle en a exactement n, c'est une base de E.
  - Toute famille génératrice de E a au moins n éléments ; et si elle en a exactement n, c'est une base de E.

  Exemple : toute famille de polynômes de degrés échelonnés de 0 à n est une base de $\mathbb{K}_n[X]$.
- **Dimension d'un espace vectoriel produit**
  Soient $E_1, E_2, \ldots, E_p$ p $\mathbb{K}$-espaces vectoriels de dimensions finies.
  Alors l'espace vectoriel produit $E = E_1 \times E_2 \times \ldots \times E_p$ est de dimension finie et $\dim E = \sum_{i=1}^{p} \dim(E_i)$.

1.12 Dimension d'un sous-espace vectoriel

- Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soit F un sous-espace vectoriel de E.
  - F est de dimension finie, et $\dim F \leqslant \dim E$.
  - De plus, si $\dim E = \dim F$, alors $F = E$.
- **Rang d'une famille**
  Soit $(x_i)_{i\in I}$ une famille d'éléments d'un $\mathbb{K}$-espace vectoriel E.
  Si le sous-espace vectoriel qu'elle engendre est de dimension finie, on définit le rang de la famille $(x_i)_{i\in I}$ par :
  $$\mathrm{rg}\big((x_i)_{i\in I}\big) = \dim\big(\mathrm{Vect}((x_i)_{i\in I})\big).$$
- **Propriétés**
  - Si E est de dimension finie, alors :
    $$\mathrm{rg}\big((x_i)_{i\in I}\big) \ \leqslant \ \dim(E)\,.$$
  - Si $(x_1,x_2,\ldots,x_p)$ est une famille finie, alors : $\mathrm{rg}(x_1,x_2,\ldots,x_p) \leqslant p\,.$

1.13 Dimension de la somme de deux sous-espaces

- **Dimension d'un supplémentaire**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.
  - Tout sous-espace vectoriel de E possède (au moins) un supplémentaire.
  - Si $E = F \oplus G$, $\dim F + \dim G = \dim E\,.$
- **Formule de Grassmann**
  Soit E un $\mathbb{K}$-espace vectoriel, et soient F et G deux sous-espaces vectoriels de dimensions finies de E. Alors $F + G$ est de dimension finie et :
  $$\dim(F+G) = \dim F + \dim G - \dim(F \cap G)\,.$$
- **Caractérisation de deux sous-espaces vectoriels supplémentaires**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soient F et G deux sous-espaces vectoriels de E.
  Pour que F et G soient supplémentaires, il faut et il suffit que :
  $$F + G = E \quad \text{et} \quad \dim E = \dim F + \dim G$$
  ou bien que :
  $$F \cap G = \{0_E\} \quad \text{et} \quad \dim E = \dim F + \dim G\,.$$

1.14 Dimension de la somme de n sous-espaces

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $(E_i)_{1\leqslant i\leqslant n}$ une famille de sous-espaces vectoriels de E.

- **Dimension de la somme**

  On a l'inégalité : $\dim\left(\sum_{i=1}^{n} E_i\right) \leqslant \sum_{i=1}^{n} \dim E_i$, et il y a égalité si et seulement si la somme est directe.

- **Caractérisation de sous-espaces vectoriels supplémentaires**

  Pour que les E_i soient supplémentaires, il faut et il suffit que :

$$\text{la somme } \sum_{i=1}^{n} E_i \text{ soit directe et que } \dim E = \sum_{i=1}^{n} \dim(E_i)$$

  ou bien que :

$$E = \sum_{i=1}^{n} E_i \quad \text{et} \quad \dim E = \sum_{i=1}^{n} \dim(E_i)\,.$$

1.15 Applications linéaires

- **Définitions**
  - Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $u\colon E \to F$ est dite linéaire si :

$$\forall\,\lambda \in \mathbb{K},\ \forall\,(x,y) \in E^2,\ u(x+y) = u(x) + u(y) \quad \text{et} \quad u(\lambda x) = \lambda u(x)\,.$$

    On peut remplacer cette définition par :

$$\forall\,\lambda \in \mathbb{K},\ \forall\,(x,y) \in E^2,\ u(\lambda x + y) = \lambda u(x) + u(y)\,.$$

    Si u est linéaire, on aura plus généralement, pour toute famille de scalaires $(\lambda_i)_{1\leqslant i\leqslant n}$ et toute famille $(x_i)_{1\leqslant i\leqslant n}$ de vecteurs de E :

$$u\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i u(x_i)\,.$$

  - Un isomorphisme de E sur F est une application linéaire bijective de E sur F.
  - Un endomorphisme de E est une application linéaire de E dans lui-même.
  - Un automorphisme de E est un endomorphisme de E bijectif.
  - On note $\mathscr{L}(E,F)$ l'ensemble des applications linéaires de E dans F, $\mathscr{L}(E)$ l'ensemble des endomorphismes de E et $\mathrm{GL}(E)$ celui des automorphismes de E.

- **Exemples**

  De même qu'il faut connaître certains espaces vectoriels de référence, il faut aussi connaître certaines applications linéaires classiques. Citons-en quelques-unes.
  - L'application nulle $\begin{array}{ccc} E & \longrightarrow & F \\ x & \longmapsto & 0_F \end{array}$.
  - L'application identique $\mathrm{Id}_E\colon \begin{array}{ccc} E & \longrightarrow & E \\ x & \longmapsto & x \end{array}$.
  - L'homothétie de rapport $\lambda \in \mathbb{K}$: $h_\lambda\colon \begin{array}{ccc} E & \longrightarrow & E \\ x & \longmapsto & \lambda.x \end{array}$ $(h_\lambda = \lambda.\mathrm{Id}_E)$.
  - L'application $P \mapsto P'$ de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$.
  - L'application $f \mapsto f(a)$ de $\mathcal{A}(D,E)$ dans E, avec $a \in D$.
  - L'application $f \mapsto \int_a^b f(t)\,\mathrm{d}t$ de $\mathscr{C}([a\,;b],\mathbb{R})$ dans $\mathbb{R}$.
  - Il y a aussi les projections et les symétries, que nous verrons un peu plus loin.

1.16 Opérations sur les applications linéaires

- Soient E et F deux $\mathbb{K}$-espaces vectoriels. Alors $\mathscr{L}(E,F)$ est un sous-espace vectoriel de $\mathcal{A}(E,F)$.
  Cela revient à dire que si u et v sont deux applications linéaires de E dans F et si λ est un scalaire, l'application $\lambda u + v$ est linéaire.
  De plus, lorsque E et F sont de dimensions finies, on a :
  $$\dim \mathscr{L}(E,F) = \dim E \times \dim F\,.$$
- Soient E, F et G trois $\mathbb{K}$-espaces vectoriels.
  Si $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$, alors $v \circ u \in \mathscr{L}(E,G)$.
- Si u est un isomorphisme de E sur F, l'application réciproque u^{-1} est encore linéaire (ce sera donc un isomorphisme de F sur E).

1.17 Image et image réciproque d'un sous-espace, noyau

- Soient E et F deux $\mathbb{K}$-espaces vectoriels, et $u \in \mathscr{L}(E,F)$.
  - Si E' est un sous-espace vectoriel de E, son image $u(E')$ par u est un sous-espace vectoriel de F.
  - Si F' est un sous-espace vectoriel de F, son image réciproque $u^{-1}(F')$ par u est un sous-espace vectoriel de E.
- Soit $u \in \mathscr{L}(E,F)$. L'ensemble image de E par u est un sous-espace vectoriel de F, appelé image de u, et noté $\operatorname{Im} u$. Ainsi :
  $$\operatorname{Im} u = \{u(x) \mid x \in E\}\,.$$
- Soit $u \in \mathscr{L}(E,F)$. L'image réciproque de $\{0_F\}$ par u est un sous-espace vectoriel de E, appelé noyau de u, et noté $\operatorname{Ker} u$. Ainsi :
  $$\operatorname{Ker} u = u^{-1}(\{0_F\}) = \{x \in E \mid u(x) = 0_F\}\,.$$
- **Théorème**
  Soit $u \in \mathscr{L}(E,F)$. Alors :
  $$u \text{ surjective } \iff \operatorname{Im} u = F \quad ; \quad u \text{ injective } \iff \operatorname{Ker} u = \{0_E\}\,.$$
- **Deux propriétés simples mais importantes**
  Soient E, F et G trois $\mathbb{K}$-espaces vectoriels, $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$. Alors :
  - $\operatorname{Im}(v \circ u) \subset \operatorname{Im} v$ et $\operatorname{Ker}(v \circ u) \supset \operatorname{Ker} u$.
  - $v \circ u = 0 \iff \operatorname{Im} u \subset \operatorname{Ker} v$.
- **Théorème d'isomorphisme**
  Soit $u \in \mathscr{L}(E,F)$. La restriction de u à tout supplémentaire de $\operatorname{Ker} u$ dans E est un isomorphisme de ce supplémentaire sur $\operatorname{Im} u$.

1.18 Détermination d'une application linéaire

- Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E, telle que $E = \bigoplus_{i=1}^{n} E_i$, et soit F un $\mathbb{K}$-espace vectoriel.
  Pour tout $i \in [\![1\,;n]\!]$, soit u_i une application linéaire de E_i dans F.
  Alors il existe une et une seule application linéaire $u \in \mathscr{L}(E,F)$ telle que pour tout $i \in [\![1\,;n]\!]$, $u_i = u|_{E_i}$.
  (*autrement dit, une application linéaire est entièrement déterminée par ses restrictions à des sous-espaces vectoriels supplémentaires*).
- Soient E et F deux $\mathbb{K}$-espaces vectoriels, $(e_i)_{i\in I}$ une base de E et $(b_i)_{i\in I}$ une famille de vecteurs de F.
  Alors il existe une et une seule application linéaire $u \in \mathscr{L}(E,F)$ telle que $u(e_i) = b_i$ pour tout $i \in I$.
  (*autrement dit, une application linéaire est entièrement déterminée par les images des vecteurs d'une base*).
- Soient E et F deux $\mathbb{K}$-espaces vectoriels et $u \in \mathscr{L}(E,F)$.
  u est injective (respectivement surjective, respectivement bijective) si et seulement si l'image d'*une* base par u est une famille libre dans F (respectivement génératrice de F, respectivement est une base de F) et, lorsque cela est le cas, cela vaut pour *toutes* les bases.
- Il est important de retenir que, si $u \in \mathscr{L}(E,F)$ et si $(e_i)_{i\in I}$ est une base de E, alors l'image de u est le sous-espace vectoriel de F engendré par les vecteurs $u(e_i)$.

  ✓ Il s'agit certainement de la façon la plus simple de déterminer l'image d'une application linéaire.

1.19 Rang d'une application linéaire

- Soient E et F deux $\mathbb{K}$-espaces vectoriels et $u \in \mathscr{L}(E,F)$. Si $\operatorname{Im} u$ est de dimension finie, sa dimension s'appelle le rang de u, noté $\operatorname{rg} u$.
  Si $(e_i)_{i\in I}$ est une base de E, le rang de u est aussi celui de la famille $\big(u(e_i)\big)_{i\in I}$.
- **Théorème du rang**
  Soient E et F deux $\mathbb{K}$-espaces vectoriels, avec E de dimension finie, et $u \in \mathscr{L}(E,F)$.
  Alors $\operatorname{Im} u$ est de dimension finie et :
  $$\dim(\operatorname{Im} u) + \dim(\operatorname{Ker} u) \;=\; \dim E\,.$$
- **Conséquences**
  u désigne ici une application linéaire de E dans F.
  - Si E est de dimension finie :
    $$\operatorname{rg} u \leqslant \dim E \quad \text{et} \quad \operatorname{rg} u = \dim E \iff u \text{ injective}\,.$$
  - Si F est de dimension finie :
    $$\operatorname{rg} u \leqslant \dim F \quad \text{et} \quad \operatorname{rg} u = \dim F \iff u \text{ surjective}\,.$$
  - Si E et F sont de dimensions finies et **si dimE = dimF**, on a :
    $$u \text{ injective} \iff u \text{ surjective} \iff u \text{ bijective.}$$
  - Soient E et F de dimensions finies, avec $\dim E = \dim F$. Soient $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,E)$ tels que $v \circ u = \mathrm{Id}_E$ (ou $u \circ v = \mathrm{Id}_F$). Alors u et v sont des isomorphismes, réciproques l'un de l'autre.

    ✓ Attention : les deux dernières propriétés ci-dessus ne sont plus vraies lorsque E et F ne sont pas de *même* dimension, ou bien lorsque l'un d'eux n'est pas de dimension finie.
- **Rang de la composée**
  Soient E, F et G trois $\mathbb{K}$-espaces vectoriels, E et F étant de dimensions finies.
  Soient $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$. Alors on a :
  - $\operatorname{rg}(v \circ u) \;\leqslant\; \operatorname{rg} v$ et, si u est surjective, il y a égalité.
  - $\operatorname{rg}(v \circ u) \;\leqslant\; \operatorname{rg} u$ et, si v est injective, il y a égalité.

  En particulier, il y a invariance du rang par la composition avec un isomorphisme.
  - Si u est bijective, $\operatorname{rg}(v \circ u) \;=\; \operatorname{rg} v$.
  - Si v est bijective, $\operatorname{rg}(v \circ u) \;=\; \operatorname{rg} u$.

1.20 Formes linéaires, hyperplans

- **Hyperplan**
  Soit E un $\mathbb{K}$-espace vectoriel. Un hyperplan H de E est un sous-espace vectoriel de E qui admet pour supplémentaire une droite vectorielle D : $E = D \oplus H$.
  Si H est un hyperplan, alors, pour toute droite vectorielle D' non incluse dans H on a $E = D' \oplus H$.
  Lorsque E est de dimension finie $n \geqslant 1$, un sous-espace vectoriel H de E est un hyperplan si et seulement si $\dim H = n - 1$.
- **Lien avec les formes linéaires**
  On rappelle qu'une forme linéaire sur un $\mathbb{K}$-espace vectoriel E est une application linéaire de E dans $\mathbb{K}$.
  Le théorème suivant est très utile pour caractériser les hyperplans.
  - Un sous-espace vectoriel H de E est un hyperplan si et seulement si il existe une forme linéaire φ sur E, non nulle, telle que $H = \operatorname{Ker} \varphi$.
  - Si φ et ψ sont deux formes linéaires non nulles sur E telles que $\operatorname{Ker} \varphi = \operatorname{Ker} \psi$, alors il existe $\lambda \in \mathbb{K}$ tel que $\psi = \lambda\varphi$.

  Pour tout hyperplan H, il existe donc une forme linéaire non nulle φ telle que $H = \{x \in E\,, \varphi(x) = 0\}$. La relation $\varphi(x) = 0$ s'appelle une équation de l'hyperplan H.

- **Expression analytique**

  Lorsque E est de dimension finie n, muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$, toute forme linéaire φ sur E est de la forme :

  $$x = \sum_{i=1}^{n} x_i e_i \longmapsto \sum_{i=1}^{n} a_i x_i$$

  où les a_i sont des scalaires ($a_i = \varphi(e_i)$).

  Il en résulte que tout hyperplan possède une équation de la forme :

  $$\sum_{i=1}^{n} a_i x_i = 0\,.$$

- **Intersection d'hyperplans**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

  Si $(\varphi_1, \ldots, \varphi_p)$ est une famille de p formes linéaires *linéairement indépendantes* sur E ($p \in \mathbb{N}^*$), le sous-espace vectoriel :

  $$F = \bigcap_{i=1}^{p} \operatorname{Ker} \varphi_i = \{x \in E \mid \forall\, i \in [\![1\,;p]\!],\ \varphi_i(x) = 0\}$$

  est un sous-espace vectoriel de E de dimension $n - p$.

  L'ensemble des p équations $\varphi_i(x) = 0$ ($1 \leqslant i \leqslant p$) s'appelle un système d'équations de F.

1.21 Projections, projecteurs, symétries

- **Définitions géométriques**

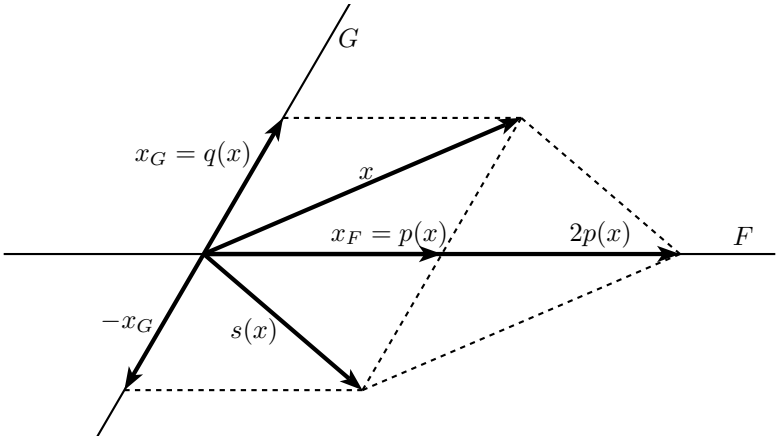
  Soit E un $\mathbb{K}$-espace vectoriel, et F, G deux sous-espaces vectoriels supplémentaires de E : $E = F \oplus G$.

  Tout vecteur $x \in E$ s'écrit donc de manière unique sous la forme : $x = x_F + x_G$ avec $(x_F, x_G) \in F \times G$.

  La projection sur F parallèlement à G est l'application $p : \begin{cases} E & \longrightarrow & E \\ x & \longmapsto & x_F \end{cases}$.

  La symétrie par rapport à F et parallèlement à G est l'application

  $s : \begin{cases} E & \longrightarrow & E \\ x & \longmapsto & x_F - x_G \end{cases}$.



- **Propriétés**
  - $s = 2p - \mathrm{Id}_E$.
  - s et p sont deux endomorphismes de E.
  - $p^2 = p \circ p = p$.
  - $s^2 = s \circ s = \mathrm{Id}_E$ (s est donc bijectif, et $s^{-1} = s$).
  - $\operatorname{Ker}(p - \mathrm{Id}_E) = \operatorname{Im} p = F$ et $\operatorname{Ker} p = G$.
  - $\operatorname{Ker}(s - \mathrm{Id}_E) = F$ et $\operatorname{Ker}(s + \mathrm{Id}_E) = G$.
  - Si p est la projection sur F parallèlement à G, et q est la projection sur G parallèlement à F, $p + q = \mathrm{Id}_E$ (p et q sont dits associés).

- **Projecteurs**

  On appelle projecteur de E tout endomorphisme p de E tel que $p^2 = p$.

  Si p est un projecteur, alors :

  $$E = \operatorname{Im} p \oplus \operatorname{Ker} p \quad ; \quad \operatorname{Im} p = \operatorname{Ker}(p - \operatorname{Id}_E)$$
  et p est la projection sur $\operatorname{Im} p$ parallèlement à $\operatorname{Ker} p$.

  ✓ Pour un projecteur p, il est vraiment important de retenir la propriété $\operatorname{Im} p = \operatorname{Ker}(p - \operatorname{Id}_E)$, qui exprime que les vecteurs de l'image de p sont exactement les vecteurs invariants par p.

  ✓ Il est également utile de retenir que la décomposition de tout vecteur $x \in E$ selon les sous-espaces vectoriels supplémentaires $\operatorname{Im} p$ et $\operatorname{Ker} p$ est :

  $$x = \underbrace{p(x)}_{\in \operatorname{Im} p} + \underbrace{x - p(x)}_{\in \operatorname{Ker} p}.$$

- **Endomorphismes involutifs**

  Si s est un endomorphisme involutif de E ($s^2 = \operatorname{Id}_E$), alors :

  $$E = \operatorname{Ker}(s - \operatorname{Id}_E) \oplus \operatorname{Ker}(s + \operatorname{Id}_E)$$
  et s est la symétrie par rapport à $\operatorname{Ker}(s - \operatorname{Id}_E)$ parallèlement à $\operatorname{Ker}(s + \operatorname{Id}_E)$.

  ✓ Il peut être utile de retenir la décomposition suivante :

  $$\forall\, x \in E,\ x = \underbrace{\frac{x + s(x)}{2}}_{\in \operatorname{Ker}(s - \operatorname{Id}_E)} + \underbrace{\frac{x - s(x)}{2}}_{\in \operatorname{Ker}(s + \operatorname{Id}_E)}.$$

- **Projecteurs associés à une décomposition en somme directe**

  Soit $(E_i)_{1 \leqslant i \leqslant n}$ une famille de sous-espaces vectoriels de E telle que $E = \bigoplus_{i=1}^{n} E_i$.

  Pour tout $x \in E$, il existe donc une unique famille $(x_i)_{1 \leqslant i \leqslant n} \in \prod_{i=1}^{n} E_i$ telle que : $x = \sum_{i=1}^{n} x_i$.

  Considérons, pour tout $i \in [\![1\,;n]\!]$, l'application $p_i : \begin{cases} E & \to & E \\ x & \mapsto & x_i \end{cases}$.

  On a alors :

  – $\forall\, i \in [\![1\,;n]\!]$, $p_i \circ p_i = p_i$.

  – $\forall\, i \in [\![1\,;n]\!]$, p_i est la projection sur E_i parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq i}}^{n} E_j$.

  – $\sum_{i=1}^{n} p_i = \operatorname{Id}_E$.

  – $\forall\, (i,j) \in [\![1\,;n]\!]^2,\ i \neq j \implies p_i \circ p_j = 0_{\mathscr{L}(E)}$.

  On dit que $(p_i)_{1 \leqslant i \leqslant n}$ est la famille de projecteurs canoniquement associée à la décomposition de E en somme directe $E = \bigoplus_{i=1}^{n} E_i$.

Chapitre 2 - Calcul matriciel

Dans toute la suite, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et p, q et n désignent des entiers naturels non nuls. On note :

- $\mathcal{M}_{p,q}(\mathbb{K})$ l'ensemble des matrices de type (p,q) à coefficients dans $\mathbb{K}$.
- $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées d'ordre n.

2.1 Opérations sur les matrices

- **Structure de $\mathbb{K}$-espace vectoriel**
  Soient $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B = (b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ deux matrices de même type (p,q), et soit $\lambda \in \mathbb{K}$ un scalaire.
  La somme des matrices A et B est définie par :
  $$A + B = (a_{ij} + b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}},$$
  et la multiplication externe de A par λ est définie par :
  $$\lambda.A = (\lambda a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}.$$
  Muni de ces lois, $(\mathcal{M}_{p,q}(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel.
  L'élément neutre pour l'addition dans $\mathcal{M}_{p,q}(\mathbb{K})$ est la matrice nulle (dont tous les termes sont nuls), que l'on note $0_{p,q}$, ou 0 s'il n'y a pas d'ambiguïté.
- **Produit de matrices**
  Soient $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{jk})_{\substack{1\leqslant j\leqslant p\\1\leqslant k\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ deux matrices.
  On définit le produit matriciel $AB = (c_{ik})_{\substack{1\leqslant i\leqslant n\\1\leqslant k\leqslant q}} \in \mathcal{M}_{n,q}(\mathbb{K})$ par :
  $$\forall\,(i,k) \in [\![1\,;n]\!] \times [\![1\,;q]\!],\ c_{ik} = \sum_{j=1}^{p} a_{ij}b_{jk}.$$
  Pour tout $k \in [\![1\,;q]\!]$, la k^{e} colonne de AB est égale à AC_k, où C_k désigne la k^{e} colonne de B. De même, pour tout $i \in [\![1\,;n]\!]$, la i^{e} ligne de AB est égale à L_iB, où L_i désigne la i^{e} ligne de A.

  ✓ Le produit matriciel AB n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B.
- **Propriétés**
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{q,r}(\mathbb{K}),\ \forall\, C \in \mathcal{M}_{r,s}(\mathbb{K}),\ A(BC) = (AB)C$.
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B_1,\, B_2 \in \mathcal{M}_{q,r}(\mathbb{K}),\ A(B_1 + B_2) = AB_1 + AB_2$.
  - $\forall\, A_1,\, A_2 \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{q,r}(\mathbb{K}),\ (A_1 + A_2)B = A_1B + A_2B$.
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{q,r}(\mathbb{K}),\ \forall\, \lambda \in \mathbb{K},\ (\lambda A)B = A(\lambda B) = \lambda(AB)$.
- **Base canonique de $\mathcal{M}_{p,q}(\mathbb{K})$**
  - Pour tout $(k,\ell) \in [\![1\,;p]\!] \times [\![1\,;q]\!]$, on note $E_{k\ell}$ la matrice de type (p,q) dont tous les termes sont nuls, sauf celui d'indice (k,ℓ) qui vaut 1. Autrement dit,
    $$E_{k\ell} = (\delta_{ki}\delta_{\ell j})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K}).$$
  - La famille des matrices $(E_{k\ell})_{(k,\ell)\in[\![1;p]\!]\times[\![1;q]\!]}$ forme une base de $\mathcal{M}_{p,q}(\mathbb{K})$, appelée base canonique.
  - $\dim(\mathcal{M}_{p,q}(\mathbb{K})) = pq$.
  - Pour toute matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$, $A = \sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} a_{ij}E_{ij}$, c'est-à-dire que les coordonnées de A dans la base canonique sont les a_{ij}.
- Soient $(E_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}$, $(E'_{jk})_{\substack{1\leqslant j\leqslant q\\1\leqslant k\leqslant r}}$ et $(E''_{ik})_{\substack{1\leqslant i\leqslant p\\1\leqslant k\leqslant r}}$ les bases canoniques respectives de $\mathcal{M}_{p,q}(\mathbb{K})$, de $\mathcal{M}_{q,r}(\mathbb{K})$ et de $\mathcal{M}_{p,r}(\mathbb{K})$. Alors :
  $$\forall\,(i,j) \in [\![1\,;p]\!] \times [\![1\,;q]\!],\ \forall\,(k,\ell) \in [\![1\,;q]\!] \times [\![1\,;r]\!],\ E_{ij}E'_{k\ell} = \delta_{jk}E''_{i\ell}.$$

2.2 Opérations par blocs

- Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. On appelle bloc de A toute matrice $(a_{ij})_{\substack{i\in I\\ j\in J}}$, où I et J sont respectivement des parties de $[\![1\,;p]\!]$ et de $[\![1\,;q]\!]$ formées d'entiers *consécutifs* (lorsque les indices ne sont pas consécutifs on parle de matrice *extraite*).
- **Combinaison linéaire par blocs**
  Soient A, $B \in \mathcal{M}_{p,q}(\mathbb{K})$ décomposées en blocs avec le même découpage :

$$A = \underbrace{\overbrace{\begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{\ell 1} & \dots & A_{\ell m} \end{bmatrix}}^{q}}_{\begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix}}\Bigg\uparrow \begin{matrix} \\ p \\ \\ \end{matrix} \quad \text{et} \quad B = \underbrace{\overbrace{\begin{bmatrix} B_{11} & \dots & B_{1m} \\ \vdots & & \vdots \\ B_{\ell 1} & \dots & B_{\ell m} \end{bmatrix}}^{q}}_{\begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix}}\Bigg\uparrow \begin{matrix} \\ p\,. \\ \\ \end{matrix}$$

  Alors, pour tout $\lambda \in \mathbb{K}$, la matrice $\lambda A + B$ s'écrit, par blocs :

$$\lambda A + B = \underbrace{\overbrace{\begin{bmatrix} \lambda A_{11} + B_{11} & \dots & \lambda A_{1m} + B_{1m} \\ \vdots & & \vdots \\ \lambda A_{\ell 1} + B_{\ell 1} & \dots & \lambda A_{\ell m} + B_{\ell m} \end{bmatrix}}^{q}}_{\begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix}}\Bigg\uparrow \begin{matrix} \\ p\,. \\ \\ \end{matrix}$$

- **Produit par blocs**
  Soient $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,r}(\mathbb{K})$, décomposées en blocs comme suit :

$$A = \underbrace{\overbrace{\begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{\ell 1} & \dots & A_{\ell m} \end{bmatrix}}^{q}}_{\begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix}}\Bigg\uparrow \begin{matrix} \\ p \\ \\ \end{matrix} \quad \text{et} \quad B = \underbrace{\overbrace{\begin{bmatrix} B_{11} & \dots & B_{1n} \\ \vdots & & \vdots \\ B_{m1} & \dots & B_{mn} \end{bmatrix}}^{r}}_{\begin{matrix} \updownarrow q_1 \\ \\ \updownarrow q_m \end{matrix}}\Bigg\uparrow \begin{matrix} \\ q\,, \\ \\ \end{matrix}$$

  et soit $C = AB \in \mathcal{M}_{p,r}(\mathbb{K})$, décomposée en blocs :

$$C = \underbrace{\overbrace{\begin{bmatrix} C_{11} & \dots & C_{1n} \\ \vdots & & \vdots \\ C_{\ell 1} & \dots & C_{\ell n} \end{bmatrix}}^{r}}_{\begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix}}\Bigg\uparrow \begin{matrix} \\ p\,. \\ \\ \end{matrix}$$

  Alors : $\forall\,(i,k) \in [\![1\,;\ell]\!] \times [\![1\,;n]\!]$, $C_{ik} = \displaystyle\sum_{j=1}^{m} A_{ij}B_{jk}$.

2.3 Transposition

Soit $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$.

On appelle transposée de A la matrice $A^\top = (a'_{ji})_{\substack{1\leqslant j\leqslant q\\ 1\leqslant i\leqslant p}} \in \mathcal{M}_{q,p}(\mathbb{K})$ définie par :

$$\forall\,(i,j) \in [\![1\,;p]\!] \times [\![1\,;q]\!],\ a'_{ji} = a_{ij}\,.$$

- **Propriétés**
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K})$, $({}^\top A^\top) = A$.
  - $\forall\, A,\, B \in \mathcal{M}_{p,q}(\mathbb{K})$, $({}^\top A + B) = A^\top + B^\top$.
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K})$, $\forall\,\lambda \in \mathbb{K}$, $({}^\top \lambda A) = \lambda A^\top$.
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K})$, $\forall\, B \in \mathcal{M}_{q,r}(\mathbb{K})$, $({}^\top AB) = B^\top\, A^\top$.
  - L'application $\left\{\begin{array}{ccc} \mathcal{M}_{p,q}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{q,p}(\mathbb{K}) \\ A & \longmapsto & A^\top \end{array}\right.$ est un isomorphisme d'espaces vectoriels.

- **Transposition par blocs**
  Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$, décomposée en blocs :

$$A = \underbrace{\begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{\ell 1} & \dots & A_{\ell m} \end{bmatrix}}_{q} \begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix} \Big\uparrow \Big\downarrow p\,.$$

  Alors la matrice transposée de A s'écrit, par blocs :

$$A^\top = \underbrace{\begin{bmatrix} A_{11}^\top & \dots & A_{\ell 1}^\top \\ \vdots & & \vdots \\ A_{1m}^\top & \dots & A_{\ell m}^\top \end{bmatrix}}_{p} \begin{matrix} \updownarrow q_1 \\ \\ \updownarrow q_m \end{matrix} \Big\uparrow \Big\downarrow q\,.$$

2.4 Structure de l'ensemble des matrices carrées

- On sait que $(\mathcal{M}_n(\mathbb{K}),+,.)$ est un $\mathbb{K}$-espace vectoriel de dimension n^2. De plus, dans $\mathcal{M}_n(\mathbb{K})$, la multiplication matricielle $\times$ est une loi de composition interne.
- **Propriétés de la multiplication interne**
  - Dans $\mathcal{M}_n(\mathbb{K})$, $\times$ est une loi associative, possède un élément neutre I_n, et est distributive à droite et à gauche par rapport à l'addition.
    La matrice I_n est la matrice carrée d'ordre n dont les termes diagonaux valent 1, et dont les termes non diagonaux sont nuls. On l'appelle matrice identité, ou matrice unité.
  - Si $n \geqslant 2$, alors $\times$ n'est pas commutative : il existe des matrices A et B de $\mathcal{M}_n(\mathbb{K})$ telles que $AB \neq BA$. Dans le cas où $AB = BA$, on dit que les matrices A et B commutent.
  - Si $n \geqslant 2$, alors il existe des matrices A et B non nulles de $\mathcal{M}_n(\mathbb{K})$ telles que $AB = 0_n$.
    Exemple : si $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, alors $A^2 = A \times A = 0_2$.

    ✓ Cela implique que les éléments de $\mathcal{M}_n(\mathbb{K})$ ne sont pas réguliers en général, c'est-à-dire qu'une égalité de la forme $AB = AC$ n'implique pas toujours $B = C$.
- **Matrices inversibles**
  Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est dite inversible s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = \mathrm{I}_n$.
  La matrice B est alors unique et est notée A^{-1} ; on l'appelle l'inverse de A.
  On note $\mathrm{GL}_n(\mathbb{K})$ l'ensemble des matrices inversibles d'ordre n, appelé groupe linéaire d'ordre n.
  Propriétés
  - $\forall\, A \in \mathrm{GL}_n(\mathbb{K})$, $A^{-1} \in \mathrm{GL}_n(\mathbb{K})$ et $\left(A^{-1}\right)^{-1} = A$.
  - $\forall\, A,\, B \in \mathrm{GL}_n(\mathbb{K})$, $AB \in \mathrm{GL}_n(\mathbb{K})$ et $(AB)^{-1} = B^{-1}A^{-1}$.
  - $\forall\, A \in \mathrm{GL}_n(\mathbb{K})$, $A^\top \in \mathrm{GL}_n(\mathbb{K})$ et $(A^\top)^{-1} = {}^\top\left(A^{-1}\right)$.

    ✓ $\mathrm{GL}_n(\mathbb{K})$ n'est pas stable par $+$: la somme de deux matrices inversibles n'est pas nécessairement inversible. Par exemple, I_n et $-\mathrm{I}_n$ sont inversibles, mais pas leur somme, qui est la matrice nulle.

2.5 Matrices carrées remarquables

- **Matrices triangulaires**
  . Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. Alors A est dite triangulaire supérieure si :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ i > j \implies a_{ij} = 0\,,$$

  et A est dite triangulaire inférieure si :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ i < j \implies a_{ij} = 0\,.$$

  On note respectivement $\mathcal{T}_n^+(\mathbb{K})$ et $\mathcal{T}_n^-(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures et l'ensemble des matrices triangulaires inférieures d'ordre n.

. **Propriétés**

– $\mathcal{T}_n^+(\mathbb{K})$ et $\mathcal{T}_n^-(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$.

– L'application $\left\{\begin{array}{ccc}\mathcal{T}_n^+(\mathbb{K}) & \longrightarrow & \mathcal{T}_n^-(\mathbb{K}) \\ A & \longmapsto & A^\top\end{array}\right.$ est un isomorphisme d'espaces vectoriels.

– $\dim(\mathcal{T}_n^+(\mathbb{K})) = \dim(\mathcal{T}_n^-(\mathbb{K})) = \dfrac{n(n+1)}{2}$.

– Soient $A = (a_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$ et $B = (b_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$ deux matrices triangulaires supérieures. Alors $C = AB \in \mathcal{T}_n^+(\mathbb{K})$ et admet pour termes diagonaux $c_{11},\ldots,c_{nn}$ avec pour tout $i \in [\![1\,;n]\!]$, $c_{ii} = a_{ii}b_{ii}$.

– Soit $A = (a_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$. Alors : A inversible $\iff \forall\, i \in [\![1\,;n]\!]$, $a_{ii} \neq 0$.

De plus, dans ce cas, A^{-1} est triangulaire supérieure et ses termes diagonaux sont les scalaires $\dfrac{1}{a_{ii}}$.

– Soit $A = (a_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$. Alors : A nilpotente $\iff \forall\, i \in [\![1\,;n]\!]$, $a_{ii} = 0$.

- **Matrices diagonales**

. Une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est dite diagonale si :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ i \neq j \implies a_{ij} = 0\,.$$

On note $\mathcal{D}_n(\mathbb{K}) = \mathcal{T}_n^+(\mathbb{K}) \cap \mathcal{T}_n^-(\mathbb{K})$ l'ensemble des matrices diagonales d'ordre n. C'est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ et $\dim(\mathcal{D}_n(\mathbb{K})) = n$.

Pour tout $(\lambda_1,\ldots,\lambda_n) \in \mathbb{K}^n$, on note $\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ la matrice diagonale $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ telle que : $\forall\, i \in [\![1\,;n]\!]$, $a_{ii} = \lambda_i$.

. **Propriétés**

– Soient $A = \mathrm{diag}(a_{11},\ldots,a_{nn}) \in \mathcal{D}_n(\mathbb{K})$ et $B = \mathrm{diag}(b_{11},\ldots,b_{nn}) \in \mathcal{D}_n(\mathbb{K})$ deux matrices diagonales. Alors $AB = \mathrm{diag}(a_{11}b_{11},\ldots,a_{nn}b_{nn}) \in \mathcal{D}_n(\mathbb{K})$.

– Soit $A = \mathrm{diag}(a_{11},\ldots,a_{nn}) \in \mathcal{D}_n(\mathbb{K})$. Alors :

$$A \text{ inversible} \iff \forall\, i \in [\![1\,;n]\!],\ a_{ii} \neq 0\,.$$

De plus, dans ce cas, $A^{-1} = \mathrm{diag}\left(\dfrac{1}{a_{11}},\ldots,\dfrac{1}{a_{nn}}\right)$.

– Soit $D \in \mathcal{D}_n(\mathbb{K})$ une matrice diagonale dont les termes diagonaux sont *deux à deux distincts*. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors :

$$AD = DA \iff A \text{ diagonale}\,.$$

✓ Autrement dit, si une matrice commute avec une matrice diagonale *à éléments diagonaux distincts*, elle est diagonale. Ce résultat, très utile, ne figure pas au programme officiel et est démontré en [**??**].

- **Matrices scalaires**

. Une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est dite scalaire s'il existe $\lambda \in \mathbb{K}$ tel que $A = \mathrm{diag}(\lambda,\ldots,\lambda) = \lambda \mathrm{I}_n$.

. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors :

$$A \text{ scalaire} \iff \forall\, M \in \mathcal{M}_n(\mathbb{K}),\ AM = MA\,.$$

✓ Autrement dit, les matrices carrées qui commutent avec toutes les autres sont les matrices scalaires. Ce résultat classique ne figure pas au programme officiel et est démontré en [**??**].

- **Matrices symétriques, matrices antisymétriques**

. Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée.
A est dite symétrique si $A^\top = A$, et antisymétrique si $A^\top = -A$.

✓ Les éléments diagonaux d'une matrice antisymétrique sont nécessairement nuls (c'est une conséquence de la définition).

On note respectivement $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices symétriques et l'ensemble des matrices antisymétriques d'ordre n.

. **Propriétés**

– $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$.

– $\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$.

– $\dim(\mathcal{S}_n(\mathbb{K})) = \dfrac{n(n+1)}{2}$ et $\dim(\mathcal{A}_n(\mathbb{K})) = \dfrac{n(n-1)}{2}$.

2.6 Matrice d'une application linéaire

- **Matrice d'une famille de vecteurs**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension n, rapporté à une base $\mathscr{B} = (e_1, \ldots, e_n)$ et soient $x_1, \ldots, x_p$ des vecteurs de E.

  Pour tout $j \in [\![1\,;p]\!]$, le vecteur x_j s'écrit dans la base $\mathscr{B}$ sous la forme $x_j = \sum_{i=1}^{n} a_{ij} e_i$ avec $a_{ij} \in \mathbb{K}$.

  La matrice $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} \in \mathcal{M}_{n,p}(\mathbb{K})$ s'appelle la matrice de la famille $(x_1, \ldots, x_p)$ dans la base $\mathscr{B}$.

  Il s'agit donc de la matrice obtenue en écrivant dans chaque colonne les coordonnées dans $\mathscr{B}$ des vecteurs de la famille.

- **Matrice d'une application linéaire**

  Soient :

  – E un $\mathbb{K}$-espace vectoriel de dimension q, rapporté à une base $\mathscr{B}_E = (e_1, \ldots, e_q)$;

  – F un $\mathbb{K}$-espace vectoriel de dimension p, rapporté à une base $\mathscr{B}_F = (e'_1, \ldots, e'_p)$;

  – $u \in \mathscr{L}(E,F)$.

  On appelle matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$ la matrice dans $\mathscr{B}_F$ du système de vecteurs $(u(e_1), \ldots, u(e_q))$. On la note généralement $M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$.

  Ainsi : $M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) = (a_{ij}) \in \mathcal{M}_{p,q}(\mathbb{K})$, où a_{ij} désigne la i^{e} coordonnée de $u(e_j)$ dans la base $\mathscr{B}_F$ soit :

  $$\forall\, j \in [\![1\,;q]\!]\ ,\ u(e_j) = \sum_{i=1}^{p} a_{ij} e'_i\,,$$

  ce que l'on peut visualiser par :

  $$M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & \ldots & u(e_q) \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1q} \\ a_{21} & a_{22} & \ldots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pq} \end{pmatrix} \end{array} \begin{array}{l} \rightarrow e'_1 \\ \rightarrow e'_2 \\ \vdots \\ \rightarrow e'_p \end{array}$$

  L'application $\left\{ \begin{array}{rcl} \mathscr{L}(E,F) & \longrightarrow & \mathcal{M}_{p,q}(\mathbb{K}) \\ u & \longmapsto & M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) \end{array} \right.$ est linéaire, c'est-à-dire :

  $$\forall\,(u,v) \in \mathscr{L}(E,F)^2,\ \forall\,\lambda \in \mathbb{K},\ M_{\mathscr{B}_E}^{\mathscr{B}_F}(\lambda u + v) = \lambda . M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) + M_{\mathscr{B}_E}^{\mathscr{B}_F}(v)\,.$$

  De plus, cette application est bijective (donc c'est un isomorphisme entre $\mathscr{L}(E,F)$ et $\mathcal{M}_{p,q}(\mathbb{K})$).

  En particulier, pour toute matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$, il existe une unique application linéaire $a \in \mathscr{L}(\mathbb{K}^q, \mathbb{K}^p)$ telle que la matrice de a dans les bases canoniques de $\mathbb{K}^q$ et $\mathbb{K}^p$ soit égale à A. L'application a ainsi définie est appelée l'application linéaire canoniquement associée à A.

  > ✓ Il est important de remarquer que si u est une application linéaire d'un espace de dimension q vers un espace de dimension p alors $M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ est une matrice de type (p,q).

- **Matrice d'un endomorphisme**

  Dans le cas où u est un endomorphisme d'un espace vectoriel E de dimension n muni d'une base $\mathscr{B}$, on appelle matrice de u dans la base $\mathscr{B}$ la matrice $M_{\mathscr{B}}^{\mathscr{B}}(u) \in \mathcal{M}_n(\mathbb{K})$, et on la note simplement $M_{\mathscr{B}}(u)$.

  Pour toute base $\mathscr{B}$ de E, on a $M_{\mathscr{B}}(\mathrm{Id}_E) = \mathrm{I}_n$. Plus généralement, si u est une homothétie de E de rapport λ, $M_{\mathscr{B}}(u) = \lambda \mathrm{I}_n$: c'est une matrice scalaire.

- **Expression analytique d'une application linéaire**

  Soient :

  – E un $\mathbb{K}$-espace vectoriel de dimension q, rapporté à une base $\mathscr{B}_E = (e_1, \ldots, e_q)$;

  – F un $\mathbb{K}$-espace vectoriel de dimension p, rapporté à une base $\mathscr{B}_F = (e'_1, \ldots, e'_p)$;

  – $u \in \mathscr{L}(E,F)$;

  – $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) = (a_{ij}) \in \mathcal{M}_{p,q}(\mathbb{K})$ la matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$.

Soit $x = \sum_{j=1}^{q} x_j e_j \in E$, et soit $y = u(x) = \sum_{i=1}^{p} y_i e'_i \in F$ son image par u. Alors :

$$\forall\, i \in [\![1\,;p]\!],\ y_i = \sum_{j=1}^{q} a_{ij} x_j \quad \text{(expression analytique de } u \text{ dans } \mathscr{B}_E \text{ et } \mathscr{B}_F\text{)}.$$

Cela s'écrit matriciellement : $\boxed{Y = AX}$, où :

- $X = \begin{pmatrix} x_1 & \dots & x_q \end{pmatrix}^\top$ est la matrice colonne des coordonnées de x dans $\mathscr{B}_E$;
- $Y = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^\top$ est la matrice colonne des coordonnées de y dans $\mathscr{B}_F$.

- **Matrice d'une composée d'applications linéaires**
  - Soient E, F et G des $\mathbb{K}$-espaces vectoriels de bases respectives $\mathscr{B}_E$, $\mathscr{B}_F$ et $\mathscr{B}_G$. Soient $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$ deux applications linéaires. Alors :
  $$M_{\mathscr{B}_E}^{\mathscr{B}_G}(v \circ u) = M_{\mathscr{B}_F}^{\mathscr{B}_G}(v) \times M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)\,.$$
  - Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathscr{B}$, et soit $u \in \mathscr{L}(E)$. On note $A = M_{\mathscr{B}}(u)$. Alors :
  $$u \in \mathrm{GL}(E) \iff A \in \mathrm{GL}_n(\mathbb{K})\,.$$
  De plus, dans ce cas, $A^{-1} = M_{\mathscr{B}}(u^{-1})$.

2.7 Formules de changement de base

- **Matrices de passage**
  - Soit E un $\mathbb{K}$-espace vectoriel de dimension n, et $\mathscr{B} = (e_1, \dots, e_n)$ et $\mathscr{B}' = (e'_1, \dots, e'_n)$ deux bases de E. On appelle matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ la matrice du système $(e'_1, \dots, e'_n)$ dans la base $\mathscr{B}$. On la note généralement $P_{\mathscr{B}}^{\mathscr{B}'}$.

    ✓ Ainsi, pour former la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, on écrit dans les colonnes les coordonnées dans $\mathscr{B}$ des vecteurs de $\mathscr{B}'$.
  - **Propriétés**
    - Si $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$ sont trois bases de E, alors $P_{\mathscr{B}}^{\mathscr{B}''} = P_{\mathscr{B}}^{\mathscr{B}'} \times P_{\mathscr{B}'}^{\mathscr{B}''}$.
    - $P_{\mathscr{B}}^{\mathscr{B}'}$ est inversible et $\left(P_{\mathscr{B}}^{\mathscr{B}'}\right)^{-1} = P_{\mathscr{B}'}^{\mathscr{B}}$.

- **Formule de changement de base pour un vecteur**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension n, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, et P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.
  Soit x un vecteur de E, X la matrice colonne de ses coordonnées dans $\mathscr{B}$ et X' celle de ses coordonnées dans $\mathscr{B}'$.
  On a alors la relation : $\boxed{X = PX'}$.

- **Formule de changement de bases pour une application linéaire**
  Soient :
  - E un $\mathbb{K}$-espace vectoriel de dimension q, muni de deux bases $\mathscr{B}_E$ et $\mathscr{B}'_E$;
  - F un $\mathbb{K}$-espace vectoriel de dimension p, muni de deux bases $\mathscr{B}_F$ et $\mathscr{B}'_F$;
  - $P \in \mathrm{GL}_q(\mathbb{K})$ la matrice de passage de $\mathscr{B}_E$ à $\mathscr{B}'_E$;
  - $Q \in \mathrm{GL}_p(\mathbb{K})$ la matrice de passage de $\mathscr{B}_F$ à $\mathscr{B}'_F$;
  - $u \in \mathscr{L}(E,F)$, $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ la matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$, $A' = M_{\mathscr{B}'_E}^{\mathscr{B}'_F}(u)$ celle dans les bases $\mathscr{B}'_E$ et $\mathscr{B}'_F$ (A, $A' \in \mathcal{M}_{p,q}(\mathbb{K})$).

  On a alors la relation : $\boxed{A' = Q^{-1}AP}$.

- **Formule de changement de base pour un endomorphisme**
  Soient :
  - E un $\mathbb{K}$-espace vectoriel de dimension n, muni de deux bases $\mathscr{B}_E$ et $\mathscr{B}'_E$;
  - $P \in \mathrm{GL}_n(\mathbb{K})$ la matrice de passage de $\mathscr{B}_E$ à $\mathscr{B}'_E$;

– $u \in \mathscr{L}(E)$, $A = M_{\mathscr{B}_E}(u)$ sa matrice dans la base $\mathscr{B}$ et $A' = M_{\mathscr{B}'_E}(u)$ celle dans la base $\mathscr{B}'$ (A, $A' \in \mathcal{M}_n(\mathbb{K})$).

On a alors la relation : $\boxed{A' = P^{-1}AP}$.

- **Matrices semblables**

  Soient A et A' deux matrices carrées de $\mathcal{M}_n(\mathbb{K})$. On dit que A et A' sont semblables s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A' = P^{-1}AP$.

  ✓ Dire que A et A' sont semblables signifie donc que ce sont les matrices dans deux bases d'un même endomorphisme u d'un espace vectoriel de dimension n. Cette interprétation est importante à retenir.

2.8 Rang d'une matrice

Le rang d'une matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$ est, par définition, le rang de ses vecteurs colonnes (éléments de $\mathcal{M}_{p,1}(\mathbb{K})$, que l'on peut assimiler à des éléments de $\mathbb{K}^p$). On le note $\operatorname{rg} A$. D'autres interprétations sont possibles :

– si $A = M_{\mathscr{B}_F}(x_1, \ldots, x_q)$ est la matrice d'une famille de vecteurs $(x_1, \ldots, x_q)$ de F relativement à une base $\mathscr{B}_F$, alors :

$$\operatorname{rg} A = \operatorname{rg}(x_1, \ldots, x_q) = \dim\big(\operatorname{Vect}(x_1, \ldots, x_q)\big);$$

– si $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ est la matrice d'une application linéaire $u \in \mathscr{L}(E,F)$ relativement à des bases $\mathscr{B}_E$ et $\mathscr{B}_F$, alors $\operatorname{rg} A = \operatorname{rg} u$.

- Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors on a les résultats suivants.
  – $\operatorname{rg} A \leqslant \min(p,q)$.
  – $\operatorname{rg} A = \operatorname{rg}(A^\top)$.
  – Si $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ est la matrice d'une application linéaire $u \in \mathscr{L}(E,F)$ dans des bases $\mathscr{B}_E$ et $\mathscr{B}_F$, alors :

$$\operatorname{rg} A = p \iff u \text{ surjective} \quad ; \quad \operatorname{rg} A = q \iff u \text{ injective}.$$

- **Caractérisation de l'inversibilité d'une matrice**

  Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors les propriétés suivantes sont équivalentes :

  (i) A est inversible,

  (ii) A est inversible à droite ($\exists\, B \in \mathcal{M}_n(\mathbb{K}) \mid AB = \mathrm{I}_n$),

  (iii) A est inversible à gauche ($\exists\, B \in \mathcal{M}_n(\mathbb{K}) \mid BA = \mathrm{I}_n$),

  (iv) $\operatorname{rg} A = n$.

- **Rang d'un produit**

  Soient $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,r}(\mathbb{K})$. Alors on a les résultats suivants.
  – $\operatorname{rg}(AB) \leqslant \min(\operatorname{rg} A, \operatorname{rg} B)$.
  – Si $p = q$ et $A \in \mathrm{GL}_p(\mathbb{K})$, alors $\operatorname{rg}(AB) = \operatorname{rg}(B)$.
  – Si $q = r$ et $B \in \mathrm{GL}_q(\mathbb{K})$, alors $\operatorname{rg}(AB) = \operatorname{rg}(A)$.

- **Caractérisation des matrices de rang r**

  Le rang d'une matrice non nulle $A \in \mathcal{M}_{p,q}(\mathbb{K})$ est le plus grand entier $r \geqslant 1$ tel que l'on puisse extraire de A une matrice carrée inversible d'ordre r.

2.9 Matrices par blocs et sous-espaces stables

- Soit E un $\mathbb{K}$-espace vectoriel, et u un endomorphisme de E.

  Un sous-espace vectoriel F de E est dit stable par u si $u(F) \subset F$, c'est-à-dire si :

$$\forall\, x \in F,\ u(x) \in F\,.$$

  On peut alors définir l'endomorphisme u_F induit par u sur F par :

$$\forall\, x \in F,\ u_F(x) = u(x)\,.$$

  ✓ Il ne faut pas confondre l'endomorphisme induit u_F, qui n'est défini que si F est stable par u, avec la restriction $u|_F$ qui est une application linéaire de F dans E.

- Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$, soient E_1 et E_2 deux sous-espaces vectoriels *supplémentaires* de E, et soit $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ une base adaptée à la décomposition $E = E_1 \oplus E_2$.
  Soit $p = \dim E_1$, u un endomorphisme de E et $A = M_{\mathscr{B}}(u)$ sa matrice dans $\mathscr{B}$.
  A peut s'écrire par blocs sous la forme :
$$A = \begin{array}{c} \begin{array}{cc} u(\mathscr{B}_1) & u(\mathscr{B}_2) \\ \downarrow & \downarrow \end{array} \\ \left[\begin{array}{cc} A_1 & B_2 \\ B_1 & A_2 \end{array}\right] \begin{array}{l} \to \mathscr{B}_1 \\ \to \mathscr{B}_2 \end{array} \\ \begin{array}{cc} \longleftrightarrow & \longleftrightarrow \\ p & n-p \end{array} \end{array}.$$
  On a alors :
$$E_1 \text{ est stable par } u \iff B_1 = 0 \quad ; \quad E_2 \text{ est stable par } u \iff B_2 = 0\,.$$
- Si E_1 est stable par u, alors A est de la forme : $\begin{bmatrix} A_1 & B_2 \\ 0 & A_2 \end{bmatrix}$.
  Une telle matrice est dite triangulaire supérieure par blocs.
  Dans ce cas, $A_1 = M_{\mathscr{B}_1}(u_1)$ où u_1 est l'endomorphisme induit par u sur E_1.
- Si E_1 et E_2 sont stables par u, alors A est de la forme $\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$.
  Une telle matrice est dite diagonale par blocs.
  Dans ce cas, $A_1 = M_{\mathscr{B}_1}(u_1)$ et $A_2 = M_{\mathscr{B}_1}(u_2)$, où u_1 et u_2 sont les endomorphismes induits par u respectivement sur E_1 et E_2.
- **Matrices triangulaires par blocs**
  - L'ensemble des matrices de la forme $\begin{array}{c} \left[\begin{array}{cc} A_1 & B_2 \\ 0 & A_2 \end{array}\right] \begin{array}{l} \updownarrow p \\ \updownarrow n-p \end{array} \\ \begin{array}{cc} \longleftrightarrow & \longleftrightarrow \\ p & n-p \end{array} \end{array}$
    est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$, stable par multiplication.
  - $A = \begin{bmatrix} A_1 & B_2 \\ 0 & A_2 \end{bmatrix}$ est inversible si et seulement si A_1 *et* A_2 sont inversibles.
    De plus, dans ce cas, A^{-1} est de la forme : $A^{-1} = \begin{bmatrix} A_1^{-1} & B_2' \\ 0 & A_2^{-1} \end{bmatrix}$.
- **Matrices diagonales par blocs**
  $A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ est inversible si et seulement si A_1 *et* A_2 sont inversibles.
  De plus, dans ce cas, $A^{-1} = \begin{bmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{bmatrix}$.

2.10 Trace d'une matrice carrée

- On appelle trace d'une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ le scalaire :
$$\operatorname{tr} A \;=\; \sum_{i=1}^{n} a_{ii} \quad \text{(somme des éléments diagonaux)}\,.$$
  - L'application $\operatorname{tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est une forme linéaire.
  - $\forall\, A \in \mathcal{M}_{n,p}(\mathbb{K})$, $\forall\, B \in \mathcal{M}_{p,n}(\mathbb{K})$, $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
- **Trace d'un endomorphisme**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$.
  Pour toute base $\mathscr{B}$ de E, le scalaire $\operatorname{tr}(M_{\mathscr{B}}(u))$ ne dépend pas de la base $\mathscr{B}$ choisie. Ce scalaire s'appelle la trace de l'endomorphisme u, notée $\operatorname{tr} u$.
  - L'application $\operatorname{tr} : \mathscr{L}(E) \to \mathbb{K}$ est une forme linéaire.
  - $\forall\, u,\, v \in \mathscr{L}(E),\ \operatorname{tr}(v \circ u) = \operatorname{tr}(u \circ v)$.
- **Trace d'un projecteur**
  Si p est un projecteur de E, alors $\operatorname{tr} p = \operatorname{rg} p$.

2.11 Opérations élémentaires sur les lignes d'une matrice

- Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Pour tout $i \in [\![1\,;p]\!]$, on note L_i la i^{e} ligne de A. On appelle opération élémentaire sur les lignes de A l'une des opérations suivantes :
  - échange des lignes L_i et L_j, notée $L_i \leftrightarrow L_j$;
  - ajout à la ligne L_i de la ligne L_j multipliée par un scalaire $\lambda \in \mathbb{K}$, où $i \neq j$, notée $L_i \leftarrow L_i + \lambda L_j$;
  - multiplication de la ligne L_i par un scalaire $\lambda \in \mathbb{K}^*$, notée $L_i \leftarrow \lambda L_i$.

  On dit que deux matrices A et A' de $\mathcal{M}_{p,q}(\mathbb{K})$ sont équivalentes par lignes s'il est possible de passer de A à A' par une suite finie d'opérations élémentaires sur les lignes. On le note $A \underset{L}{\sim} A'$.

  La relation $\underset{L}{\sim}$ est une relation d'équivalence sur $\mathcal{M}_{p,q}(\mathbb{K})$.

  > De manière analogue, on définit les notions d'opérations élémentaires sur les colonnes d'une matrice, et on note $A \underset{C}{\sim} A'$ le fait que les matrices A et A' soient équivalentes par colonnes. Alors $A \underset{L}{\sim} A' \iff A^\top \underset{C}{\sim} A'^\top$.

- **Matrices élémentaires**
  Soit $(E_{ij})_{1 \leqslant i,j \leqslant p}$ la base canonique de $\mathcal{M}_p(\mathbb{K})$, et soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$.
  - Si A' est la matrice obtenue à partir de A par l'opération :
    $$L_i \leftarrow L_i + \lambda L_j \ (i \neq j)\,,$$
    alors $A' = P_1 A$, où $P_1 = I_p + \lambda E_{ij}$.
  - Si A' est la matrice obtenue à partir de A par l'opération :
    $$L_i \leftarrow \lambda L_i \ (\lambda \in \mathbb{K}^*)\,,$$
    alors $A' = P_2 A$, où $P_2 = I_p + (\lambda - 1)E_{ii}$.
  - Si A' est la matrice obtenue à partir de A par l'opération :
    $$L_i \leftrightarrow L_j\,,$$
    alors $A' = P_3 A$, où $P_3 = I_p - E_{ii} - E_{jj} + E_{ij} + E_{ji}$.
  - Les matrices P_1, P_2 et P_3 sont appelées matrices élémentaires. Ce sont des matrices carrées inversibles.

    > ✓ Les hypothèses $i \neq j$ (pour l'opération $L_i \leftarrow L_i + \lambda L_j$) et $\lambda \neq 0$ (pour l'opération $L_i \leftarrow \lambda L_i$) sont *indispensables* : elles garantissent, en effet, que les matrices P_1 et P_2 sont inversibles.

    > ✓ On obtient des résultats similaires pour les opérations élémentaires sur les colonnes : il faut alors multiplier A *à droite* par une matrice élémentaire $Q \in \mathrm{GL}_q(\mathbb{K})$.

- **Matrices échelonnées**
  - Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Pour tout $i \in [\![1\,;p]\!]$, on note :
    $$m_i = \begin{cases} \min\{j \in [\![1\,;q]\!] \mid a_{ij} \neq 0\} & \text{si } L_i \neq 0, \\ q+i & \text{si } L_i = 0. \end{cases}$$
    Si la ligne L_i est non nulle, m_i désigne donc le numéro de colonne du premier terme non nul de L_i en partant de la gauche. Le terme a_{ij} correspondant s'appelle un pivot.
  - On dit que A est échelonnée par lignes si la suite finie $(m_i)_{1 \leqslant i \leqslant p}$ est strictement croissante.
    Ainsi, dans une matrice échelonnée par lignes, la position des pivots croît strictement avec i jusqu'aux lignes nulles éventuelles, et si une ligne est nulle, alors toutes les lignes suivantes sont nulles.
  - Une matrice A échelonnée par lignes est dite échelonnée réduite par lignes si A est nulle ou si tous les pivots de A valent 1 et sont les seuls termes non nuls de leur colonne.
  - On définit de manière analogue les notions de matrice échelonnée par colonnes et de matrice échelonnée réduite par colonnes.

- **Théorème de Gauss-Jordan**
  Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$.
  - Il existe une unique matrice $R \in \mathcal{M}_{p,q}(\mathbb{K})$ échelonnée réduite par lignes telle que $A \underset{L}{\sim} R$. De plus, il existe $U \in \mathrm{GL}_p(\mathbb{K})$, produit de matrices élémentaires, telle que $UA = R$.
  - Il existe une unique matrice $R' \in \mathcal{M}_{p,q}(\mathbb{K})$ échelonnée réduite par colonnes telle que $A \underset{C}{\sim} R'$. De plus, il existe $V \in \mathrm{GL}_q(\mathbb{K})$, produit de matrices élémentaires, telle que $AV = R'$.

2.12 Applications de la méthode du pivot

- **Calcul du rang d'une matrice**

$$\forall\, A,\, A' \in \mathcal{M}_{p,q}(\mathbb{K}),\, A \underset{L}{\sim} A' \implies \operatorname{rg} A = \operatorname{rg} A'\,.$$

Pour déterminer le rang de A, il suffit d'appliquer la méthode du pivot pour obtenir une matrice A' échelonnée par lignes. Le rang de A est alors égal au nombre de pivots de A'.

- **Caractérisation des matrices de rang r**

Pour tout $r \in [\![1\,;\min(p,q)]\!]$, on note $J_{p,q,r} = \begin{bmatrix} I_r & 0_{r,q-r} \\ 0_{p-r,r} & 0_{p-r,q-r} \end{bmatrix} \in \mathcal{M}_{p,q}(\mathbb{K})$.

Alors, pour toute matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$:

$$\operatorname{rg} A = r \iff \exists\,(U,V) \in \mathrm{GL}_p(\mathbb{K}) \times \mathrm{GL}_q(\mathbb{K}) \text{ telles que } UAV = J_{p,q,r}\,.$$

- **Calcul de l'inverse d'une matrice carrée** (s'il existe)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors il existe $P_1, P_2, \ldots, P_m \in \mathrm{GL}_n(\mathbb{K})$ (matrices élémentaires) telles que : $P_m P_{m-1} \ldots P_1 A = R$, où R est l'unique matrice échelonnée réduite par lignes obtenue par l'algorithme du pivot. De plus,

$$A \text{ inversible} \iff R = \mathrm{I}_n\,.$$

Dans ce cas, $A^{-1} = P_m P_{m-1} \ldots P_1 \mathrm{I}_n$, ce qui signifie que A^{-1} se déduit de I_n par la *même* suite d'opérations sur les lignes que celle utilisée pour déduire I_n de A.

2.13 Systèmes linéaires

- Un système linéaire de p équations à q inconnues est un système de la forme :

$$(S) \quad \begin{cases} a_{11}x_1 + \ldots + a_{1q}x_q &= b_1 \\ a_{21}x_1 + \ldots + a_{2q}x_q &= b_2 \\ \vdots & \vdots \\ a_{p1}x_1 + \ldots + a_{pq}x_q &= b_p \end{cases}$$

où les coefficients $(a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}}$ et $(b_i)_{1 \leqslant i \leqslant p}$ sont donnés dans $\mathbb{K}$. $x_1, \ldots, x_q$, éléments de $\mathbb{K}$, sont les inconnues.

Le système est dit homogène si $b_1 = \ldots = b_p = 0$.

- **Interprétations**
  - La matrice $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ s'appelle la matrice du système ; le rang de A s'appelle le rang du système.

    Soit alors $B = \begin{pmatrix} b_1 & \ldots & b_p \end{pmatrix}^\top \in \mathcal{M}_{p,1}(\mathbb{K})$ et $X = \begin{pmatrix} x_1 & \ldots & x_q \end{pmatrix}^\top \in \mathcal{M}_{q,1}(\mathbb{K})$; le système (S) s'écrit alors matriciellement sous la forme

$$AX = B\,.$$

  - Soient E et F deux $\mathbb{K}$-espaces vectoriels tels que $\dim E = q$ et $\dim F = p$. Soient alors $\mathscr{B}_E$ et $\mathscr{B}_F$ deux bases respectivement de E et de F, et $u \in \mathscr{L}(E,F)$ telle que $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$.

    Le système (S) peut alors s'écrire $u(x) = b$, avec $x \in E$ de coordonnées $(x_1, \ldots, x_q)$ dans $\mathscr{B}_E$ et $b \in F$ de coordonnées $(b_1, \ldots, b_p)$ dans $\mathscr{B}_F$.

    On notera que le rang du système est aussi égal au rang de l'application linéaire u.

- On note $\mathscr{S}$ l'ensemble des solutions de (S). On dit que (S) est compatible si $\mathscr{S} \neq \emptyset$, et incompatible sinon.

On note $\operatorname{Ker} A = \{X \in \mathcal{M}_{q,1}(\mathbb{K}) \mid AX = 0\}$ l'ensemble des solutions du système homogène (S_H) associé à (S), et on note $\operatorname{Im} A = \operatorname{Vect}(C_1, \ldots, C_q)$ le sous-espace vectoriel de $\mathcal{M}_{p,1}(\mathbb{K})$ engendré par les colonnes de A.

  - (S) compatible $\iff B \in \operatorname{Im} A$.
  - Si (S) est compatible, alors il existe $X_0 \in \mathcal{M}_{p,q}(\mathbb{K})$ tel que $AX_0 = B$; X_0 s'appelle une solution particulière du système. L'ensemble des solutions du système est alors :

$$\mathscr{S} = X_0 + \operatorname{Ker} A = \{X_0 + X_H \mid X_H \in \operatorname{Ker} A\}\,.$$

- Deux systèmes équivalents par lignes ont même ensemble de solutions. Pour résoudre (S), il suffit d'appliquer l'algorithme de Gauss-Jordan à la matrice complète du système $\begin{bmatrix} A & B \end{bmatrix}$ pour obtenir un système (S') dont la matrice associée R est échelonnée réduite par lignes.

  On appelle alors inconnues principales les inconnues correspondant aux pivots de R, et on appelle inconnues secondaires les autres. Si (S) est compatible, alors le système (S') permet d'exprimer les inconnues principales en fonction des inconnues secondaires.

  > ✓ Le choix des inconnues principales et secondaires est souvent arbitraire. Cependant leur nombre est toujours le même : si r est le rang du système, il y a r inconnues principales et $q - r = \dim(\operatorname{Ker} A)$ inconnues secondaires.

- Un système linéaire (S) est dit de Cramer si sa matrice associée est inversible. Tout système de Cramer admet une solution unique.

- **Caractérisation des matrices inversibles**

  Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors les propositions suivantes sont équivalentes :

  (i) A est inversible,

  (ii) Le système homogène $AX = 0$ n'admet que 0 comme solution,

  (iii) Pour tout $B \in \mathcal{M}_{n,1}(\mathbb{K})$, le système $AX = B$ admet une unique solution,

  (iv) Pour tout $B \in \mathcal{M}_{n,1}(\mathbb{K})$, le système $AX = B$ admet au moins une solution.

2.14 Déterminants

- **Applications n–linéaires**

  Soit $n \in \mathbb{N}^*$. Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $f\colon E^n \to F$ est dite n–linéaire si, pour tout $j \in [\![1\,;n]\!]$ et pour toute famille $(a_i)_{i\in[\![1;n]\!]\setminus\{j\}}$ composée de $(n-1)$ vecteurs de E, l'application partielle :

  $$f_j : \begin{cases} E & \longrightarrow & F \\ x & \longmapsto & f(a_1,\ldots,a_{j-1},x,a_{j+1},\ldots,a_n) \end{cases}$$

  est une application linéaire sur E.

  Lorsque $F = \mathbb{K}$, on parle de *forme* n-linéaire.

- **Applications n-linéaires antisymétriques**

  Soit f une application n-linéaire de E^n dans F, $n \geqslant 2$.

  - f est dite symétrique si pour tout couple d'indices $(i,j) \in [\![1\,;n]\!]^2$ avec $i < j$ et pour tout $(x_i)_{1\leqslant i\leqslant n} \in E^n$,
    $$f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = f(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n)$$
  - f est dite antisymétrique si pour tout couple d'indices $(i,j) \in [\![1\,;n]\!]^2$ avec $i < j$ et pour tout $(x_i)_{1\leqslant i\leqslant n} \in E^n$,
    $$f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = -f(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n)$$
  - Si f est une application n-linéaire antisymétrique, alors pour toute famille $(x_i)_{1\leqslant i\leqslant n} \in E^n$, on a
    $$f(x_1,\ldots,x_n) = 0 \text{ dès qu'il existe deux indices } i \neq j \text{ tels que } x_i = x_j.$$

- **Déterminant d'une matrice carrée**

  Soit $f\colon \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ une application.

  En identifiant toute matrice $M \in \mathcal{M}_n(\mathbb{K})$ au n-uplet $(C_1,\ldots,C_n)$ composé de ses n colonnes, f peut être considérée comme une application de $(\mathcal{M}_{n,1}(\mathbb{K}))^n$ à valeurs dans $\mathbb{K}$. On dit alors que f est linéaire par rapport aux colonnes si f est une application n–linéaire sur $\mathcal{M}_{n,1}(\mathbb{K})$.

  Théorème

  Il existe une unique application $f\colon \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ telle que :

  (i) f est linéaire par rapport aux colonnes ;

  (ii) f est antisymétrique par rapport aux colonnes ;

  (iii) $f(\mathrm{I}_n) = 1$.

  Cette application est appelée application déterminant et est notée $A \mapsto \det A$.

- **Propriétés**

  – Si A admet une colonne nulle ou deux colonnes égales, alors $\det A = 0$.

- $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ \forall\, \lambda \in \mathbb{K},\ \det(\lambda A) = \lambda^n \det A$.
- $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ A \in \mathrm{GL}_n(\mathbb{K}) \iff \det A \neq 0$.
- $\forall\, A,\ B \in \mathcal{M}_n(\mathbb{K}),\ \det(AB) = \det A \times \det B = \det(BA)$.
- $\forall\, A \in \mathrm{GL}_n(\mathbb{K}),\ \det(A^{-1}) = \dfrac{1}{\det A}$.
- Deux matrices semblables ont même déterminant.
- $\forall\, A \in \mathcal{M}_n(\mathbb{K})$, $\det(A^\top) = \det A$.

- **Calcul du déterminant par opérations élémentaires**
  - Le déterminant d'une matrice triangulaire est égal au produit de ses termes diagonaux.
  - Le déterminant d'une matrice est multiplié par -1 lors de l'échange de deux de ses colonnes ou de deux de ses lignes.
  - Pour tout scalaire λ, le déterminant d'une matrice est multiplié par λ lorsqu'une colonne ou une ligne est multipliée par λ.
  - Le déterminant d'une matrice reste inchangé en ajoutant à l'une de ses colonnes une combinaison linéaire des *autres* colonnes ou en ajoutant à l'une de ses lignes une combinaison linéaire des *autres* lignes.

- **Calcul d'un déterminant par blocs**

  Soit $A \in \mathcal{M}_n(\mathbb{K})$, écrite par blocs sous la forme $A = \begin{bmatrix} A_1 & B_2 \\ 0 & A_2 \end{bmatrix}$, où A_1 et A_2 sont deux blocs *carrés*. Alors $\det A = \det A_1 \times \det A_2$.

- **Développement suivant une rangée**
  - Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$. Pour tout $(i,j) \in [\![1\,;n]\!]^2$, on note Δ_{ij} le déterminant obtenu en supprimant la i-ème ligne et la j-ème colonne de A.
    On dit alors que Δ_{ij} est le mineur d'indice (i,j) de la matrice A.
  - Développement de $\det A$ selon la j-ème colonne :
  $$\forall\, j \in [\![1\,;n]\!],\ \det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}\,.$$
  - Développement de $\det A$ selon la i-ème ligne :
  $$\forall\, i \in [\![1\,;n]\!],\ \det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}\,.$$
  - Le terme $(-1)^{i+j}\Delta_{ij}$ s'appelle le cofacteur d'indice (i,j) de la matrice A.

- **Déterminant d'une famille de vecteurs dans une base**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathscr{B}$.
  Soit $\mathcal{F} = (x_i)_{1 \leqslant i \leqslant n} \in E^n$ une famille de n vecteurs de E, et $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ la matrice de cette famille dans la base $\mathscr{B}$ c'est-à-dire :
  $$\forall\, j \in [\![1\,;n]\!],\ x_j = \sum_{i=1}^{n} a_{ij} e_i\,.$$
  On appelle déterminant de la famille $\mathcal{F}$ dans la base $\mathscr{B}$ le déterminant de la matrice A. Il est noté $\det_{\mathscr{B}}(\mathcal{F})$.

  ✓ Le déterminant d'une famille de vecteurs dépend de la base $\mathscr{B}$ choisie.

  ✓ Le déterminant n'est défini que pour une famille de vecteurs dont le cardinal est égal à la dimension de E.

  Avec les notations précédentes : $\mathcal{F}$ base de $E \iff \det_{\mathscr{B}}(\mathcal{F}) \neq 0$.

- **Déterminant d'un endomorphisme**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathscr{B}$.
  Pour tout endomorphisme u de E, le scalaire $\det\big(M_{\mathscr{B}}(u)\big)$ est indépendant de la base $\mathscr{B}$ choisie. On appelle ce scalaire le déterminant de l'endomorphisme u, noté $\det u$.

  Propriétés
  - $\det \mathrm{Id}_E = 1$.
  - $\forall\, u \in \mathscr{L}(E),\ \forall\, \lambda \in \mathbb{K},\ \det(\lambda u) = \lambda^n \det u$.

– $\forall\, u \in \mathscr{L}(E),\ u \in \mathrm{GL}(E) \iff \det u \neq 0$.

– $\forall\, u,\, v \in \mathscr{L}(E),\ \det(u \circ v) = \det u \times \det v$.

– $\forall\, u \in \mathrm{GL}(E),\ \det(u^{-1}) = \dfrac{1}{\det u}$.

– Pour tout endomorphisme u de E et pour toute famille $(x_i)_{1 \leqslant i \leqslant n} \in E^n$:

$$\det_{\mathscr{B}}\big(u(x_1),\ldots,u(x_n)\big) \;=\; \det u \;\times\; \det_{\mathscr{B}}(x_1,\ldots,x_n)\,.$$

- **Déterminant de Vandermonde**

  On appelle déterminant de Vandermonde un déterminant de la forme :

$$V(a_1,a_2,\ldots,a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-1} \end{vmatrix},$$

  où $(a_1,\ldots,a_n) \in \mathbb{K}^n$.

  On sait calculer ce déterminant :

$$V(a_1,a_2,\ldots,a_n) = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i)\,.$$

  Il en résulte que :

$$V(a_1,a_2,\ldots,a_n) \neq 0 \iff \text{les } a_i \text{ sont deux à deux distincts}\,.$$

Chapitre 3 - Réduction des endomorphismes

Réduire un endomorphisme $u \in \mathscr{L}(E)$ consiste, dans le cas le plus général, à trouver une famille $(E_i)_{1\leqslant i\leqslant p}$ de sous-espaces vectoriels de E, non réduits à $\{0\}$, stables par u, tels que $E = \bigoplus_{i=1}^{p} E_i$.

En dimension finie, cela revient à trouver une base $\mathscr{B}$ de E où la matrice de u est diagonale par blocs :

$$M_{\mathscr{B}}(u) = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \boxed{A_2} & \\ & \ddots & \\ 0 & & \boxed{A_p} \end{pmatrix}$$

chaque bloc A_i étant la matrice dans une base de E_i de l'endomorphisme u_i induit par u sur E_i (cf. [2.9]).

Le cas le plus simple est celui où les E_i sont de dimension 1, ce qui revient à chercher les droites vectorielles stables par u.

3.1 Éléments propres d'un endomorphisme

- Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E.
  - On dit que $\lambda \in \mathbb{K}$ est une valeur propre de u s'il existe un vecteur $x \in E$, $\underline{x \neq 0_E}$, tel que $u(x) = \lambda x$.
  - On dit que $x \in E$ est un vecteur propre de u si $x \neq 0_E$ et s'il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$.
    On dit alors que x est *un* vecteur propre associé à la valeur propre λ, ou que λ est *la* valeur propre associé au vecteur propre x.
    Cela équivaut à dire que la droite vectorielle $\mathbb{K}.x$ est stable par u.
  - On appelle spectre de u, noté $\mathrm{Sp}_{\mathbb{K}}(u)$ ou plus simplement $\mathrm{Sp}(u)$, l'ensemble des valeurs propres de u (dans $\mathbb{K}$).
- Il est *très important* de bien connaître et comprendre les équivalences suivantes :

$$\begin{aligned} \lambda \in \mathrm{Sp}(u) &\iff \exists\, x \neq 0 \text{ tel que } u(x) = \lambda x \\ &\iff \exists\, x \neq 0 \text{ tel que } (u - \lambda \mathrm{Id}_E)(x) = 0 \\ &\iff \mathrm{Ker}(u - \lambda \mathrm{Id}_E) \neq \{0\} \\ &\iff (u - \lambda \mathrm{Id}_E) \quad \textit{non} \quad \textit{injective.} \end{aligned}$$

  En particulier : 0 est valeur propre de $u \iff u$ non injectif.
- Si $\lambda \in \mathrm{Sp}(u)$, on appelle sous-espace propre de u associé à la valeur propre λ le sous-espace vectoriel :

$$E_\lambda(u) = \mathrm{Ker}(u - \lambda \mathrm{Id}_E)\,.$$

  Ainsi, $E_\lambda(u) = \{x \in E \mid u(x) = \lambda x\}$ est constitué des vecteurs propres de u associés à la valeur propre λ et du vecteur nul.

  ✓ On rappelle que le vecteur nul n'est *pas* un vecteur propre !
- La propriété suivante, qui découle de la définition, est très souvent utilisée.
  Le sous-espace propre associé à la valeur propre 0 est le noyau de u.
  On a aussi une propriété importante concernant les valeurs propres non nulles.
  Les sous-espaces propres de u associés à une valeur propre non nulle sont inclus dans $\mathrm{Im}\, u$.
  (En effet, si x est un vecteur propre associé à la valeur propre $\lambda \neq 0$ on a $u(x) = \lambda x$ donc $x = u\left(\frac{x}{\lambda}\right) \in \mathrm{Im}\, u$).

3.2 Propriétés des sous-espaces propres

u désigne toujours ici un endomorphisme d'un espace vectoriel E.

- Si λ est une valeur propre de u, $E_\lambda(u)$ est stable par u, et l'endomorphisme induit par u sur $E_\lambda(u)$ est l'homothétie de rapport λ.

- Si $v \in \mathscr{L}(E)$ *commute* avec u, $E_\lambda(u)$ est stable par v.
- Soit $p \in \mathbb{N}^*$ et $(\lambda_i)_{1\leqslant i\leqslant p}$ une famille de valeurs propres de u *deux à deux distinctes*.
  Alors les sous-espaces propres $E_{\lambda_i}(u)$ pour $1 \leqslant i \leqslant p$ sont en somme directe.
- Soit $(\lambda_i)_{1\leqslant i\leqslant p}$ une famille de valeurs propres de u *deux à deux distinctes* et, pour tout $i \in [\![1\,;p]\!]$, x_i un vecteur propre associé à la valeur propre λ_i.
  Alors la famille $(x_i)_{1\leqslant i\leqslant p}$ est *libre*.
- On en déduit immédiatement que, dans un espace vectoriel de dimension n, un endomorphisme possède au plus n valeurs propres distinctes.

3.3 Éléments propres d'une matrice

- Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$.
  - On dit que $\lambda \in \mathbb{K}$ est une valeur propre de A s'il existe un vecteur colonne $V \in \mathcal{M}_{n,1}(\mathbb{K})$, $\underline{V \neq 0}$, tel que $AV = \lambda V$.
  - On dit alors que le vecteur colonne $V \in \mathcal{M}_{n,1}(\mathbb{K})$ est un vecteur propre de A associé à la valeur propre λ.
  - On appelle spectre de A, noté $\mathrm{Sp}_{\mathbb{K}}(A)$ ou plus simplement $\mathrm{Sp}(A)$, l'ensemble des valeurs propres de A (dans $\mathbb{K}$).
  - Si $\lambda \in \mathrm{Sp}(A)$, on appelle sous-espace propre de A associé à λ l'ensemble $E_\lambda(A) = \{V \in \mathcal{M}_{n,1}(\mathbb{K}) \mid AV = \lambda V\}$.
    C'est évidemment un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbb{K})$.

    ✓ Lorsque A est à coefficients réels, il est important de distinguer ses valeurs propres réelles et ses valeurs propres complexes, c'est-à-dire les ensembles $\mathrm{Sp}_{\mathbb{R}}(A)$ et $\mathrm{Sp}_{\mathbb{C}}(A)$.
    Dans le premier cas, les vecteurs propres seront des vecteurs colonnes à chercher dans $\mathcal{M}_{n,1}(\mathbb{R})$ et dans le second cas dans $\mathcal{M}_{n,1}(\mathbb{C})$.

    De plus, si λ est une valeur propre complexe de la matrice A à coefficients réels, alors $\bar{\lambda}$ est encore une valeur propre de A et

$$V \in E_\lambda(A) \iff \overline{V} \in E_{\bar{\lambda}}(A)\,.$$

- **Lien entre éléments propres d'une matrice et d'un endomorphisme**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^*$ et $\mathscr{B}$ une base de E.
  Soit u un endomorphisme de E et A la matrice de u dans la base $\mathscr{B}$.
  - $\lambda \in \mathbb{K}$ est valeur propre de $A \iff \lambda$ est valeur propre de u.
  - Soit $x \in E$ et V la matrice colonne de ses coordonnées dans $\mathscr{B}$.

$$x \in E_\lambda(u) \iff V \in E_\lambda(A)\,.$$

3.4 Polynôme caractéristique

E désigne ici un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^*$.

- Soit $u \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$. Par définition d'une valeur propre on a :

$$\lambda \in \mathrm{Sp}(u) \iff \det(\lambda \mathrm{Id}_E - u) = 0\,.$$

  Soit $\mathscr{B}$ une base de E, $A = (a_{ij})$ la matrice de u dans $\mathscr{B}$. Alors :

$$\det(\lambda \mathrm{Id}_E - u) = \det(\lambda \mathrm{I}_n - A) = \begin{vmatrix} \lambda - a_{11} & \dots & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & \dots & \ddots & \vdots \\ -a_{n1} & \dots & \dots & \lambda - a_{nn} \end{vmatrix}.$$

  D'après la formule de développement du déterminant, il apparaît clairement que $\det(\lambda \mathrm{Id}_E - u)$ est une fonction polynôme en λ, de degré n.
  On appelle polynôme caractéristique de u le polynôme, noté χ_u, défini par :

$$\forall\, \lambda \in \mathbb{K},\ \chi_u(\lambda) = \det(\lambda \mathrm{Id}_E - u) \quad \text{ou} \quad \chi_u = \det(X \mathrm{Id}_E - u)\,.$$

  Si $\mathscr{B}$ est une base quelconque de E, et si $A = M_{\mathscr{B}}(u)$, on a aussi :

$$\forall\, \lambda \in \mathbb{K},\ \chi_u(\lambda) = \det(\lambda \mathrm{I}_n - A) \quad \text{ou} \quad \chi_u = \det(X \mathrm{I}_n - A)\,.$$

  Le polynôme $\chi_A = \det(X \mathrm{I}_n - A)$ s'appelle naturellement le polynôme caractéristique de la matrice A.

- Les valeurs propres d'un endomorphisme de E (ou d'une matrice carrée) sont exactement les racines dans $\mathbb{K}$ de son polynôme caractéristique.

  ✓ Comme il a déjà été dit, lorsque A est une matrice à coefficients réels, il convient de distinguer ses valeurs propres dans $\mathbb{R}$ ou dans $\mathbb{C}$, c'est-à-dire les racines de χ_A dans $\mathbb{R}$ ou dans $\mathbb{C}$.

- Deux matrices semblables ont même polynôme caractéristique (donc mêmes valeurs propres).
- Le polynôme caractéristique d'une matrice est le même que celui de sa transposée. Donc pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, $\operatorname{Sp} A = \operatorname{Sp} A^\top$.
- Le polynôme caractéristique de u est un polynôme unitaire de degré n. Plus précisément, il s'écrit :
$$\chi_u = X^n - (\operatorname{tr} u)X^{n-1} + \cdots + (-1)^n \det u\,.$$
- Si le polynôme caractéristique χ_u de u est *scindé* (ce qui est toujours le cas si $\mathbb{K} = \mathbb{C}$), on peut écrire :
$$\chi_u = \prod_{i=1}^{n}(X - \lambda_i)\,,$$
où les λ_i sont les valeurs propres, distinctes ou non, de u. À l'aide des relations coefficients-racines on obtient alors les relations (*très importantes*) :
$$\operatorname{tr} u = \lambda_1 + \cdots + \lambda_n \quad ; \quad \det u = \lambda_1\lambda_2\cdots\lambda_n\,.$$
✓ On prendra soin cependant de n'utiliser ces relations que lorsque le polynôme caractéristique de u est *scindé*. Ainsi, dans le cas d'une matrice à coefficients réels, il faut considérer toutes les valeurs propres dans $\mathbb{C}$.
- **Ordre de multiplicité d'une valeur propre**
  Soit $u \in \mathscr{L}(E)$ et soit $\lambda \in \operatorname{Sp}_{\mathbb{K}}(u)$.
  On dit que λ est valeur propre d'ordre de multiplicité m_λ de u si λ est une racine d'ordre m_λ du polynôme caractéristique χ_u de u.
  On a alors :
$$1 \leqslant \dim E_\lambda(u) \leqslant m_\lambda\,.$$

3.5 Endomorphismes ou matrices diagonalisables

- Un endomorphisme u d'un $\mathbb{K}$-espace vectoriel E de dimension finie est dit diagonalisable s'il existe une base de E formée de vecteurs propres de u.
  Cela équivaut à dire qu'il existe une base $\mathscr{B}$ de E dans laquelle la matrice de u est diagonale ; $\mathscr{B}$ est formée de vecteurs propres de u, et les valeurs propres de u (éventuellement confondues) sont les éléments de la diagonale de $M_{\mathscr{B}}(u)$.
- **Diagonalisabilité et sous-espaces propres**
  Un endomorphisme u est diagonalisable si et seulement si il vérifie l'une des propriétés (équivalentes) suivantes :
  (i) E est somme (directe) des sous-espaces propres de u : $E = \bigoplus_{\lambda \in \operatorname{Sp}(u)} E_\lambda(u)$,
  (ii) $\dim E = \sum_{\lambda \in \operatorname{Sp}(u)} \dim\left(E_\lambda(u)\right)$,
  (iii) χ_u est scindé dans $\mathbb{K}[X]$ et, pour tout $\lambda \in \operatorname{Sp}(u)$, on a : $m_\lambda = \dim E_\lambda(u)$ (en notant m_λ l'ordre de multiplicité de la valeur propre λ).

  Cas particulier : si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension n qui possède n valeurs propres distinctes, alors u est diagonalisable (et, dans ce cas, tous ses sous-espaces propres sont de dimension 1).

  ✓ Attention ! il ne s'agit là que d'une condition *suffisante*, et non nécessaire, de diagonalisabilité. Considérer par exemple $u = \operatorname{Id}_E$: u est diagonalisable, mais u admet une seule valeur propre, 1.

3.6 Sous-espaces stables

- Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie E.
  Soit F un sous-espace vectoriel de E *stable* par u, et $v \in \mathscr{L}(F)$ l'endomorphisme de F induit par u.
  Alors le polynôme caractéristique de v divise le polynôme caractéristique de u. En particulier on a :
$$\operatorname{Sp}(v) \subset \operatorname{Sp}(u) \quad \text{et} \quad \forall\,\lambda \in \operatorname{Sp}(v),\ E_\lambda(v) = E_\lambda(u) \cap F\,.$$

- **Diagonalisabilité et sous-espaces stables**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathscr{L}(E)$.
  - Si F est un sous-espace vectoriel de E *stable* par u, et si $v \in \mathscr{L}(F)$ est l'endomorphisme de F induit par u, alors :
    $$u \text{ diagonalisable} \implies v \text{ diagonalisable}.$$
  - *Lorsque u est diagonalisable*, un sous-espace vectoriel F de E est stable par u si et seulement si F admet une base formée de vecteurs propres de u.

    ✓ Ce résultat tombe en défaut lorsque u n'est pas diagonalisable : par exemple, si u est une rotation d'un espace euclidien de dimension 3, différente de $\pm\mathrm{Id}_E$, le plan orthogonal à l'axe de cette rotation est stable par u mais ne contient aucun vecteur propre de u.

3.7 Polynômes d'endomorphismes

- Soit E un $\mathbb{K}$-espace vectoriel, et $u \in \mathscr{L}(E)$.
  À tout polynôme $P = \sum_{k=0}^{d^\circ(P)} a_k X^k$ de $\mathbb{K}[X]$ on peut associer le polynôme d'endomorphisme : $P(u) = \sum_{k=0}^{d^\circ(P)} a_k u^k$ (où u^k désigne, comme d'habitude, l'endomorphisme défini par : $u^0 = \mathrm{Id}_E$ et, pour tout $k \geqslant 1$: $u^k = u \circ u^{k-1}$).
  $P(u)$ est encore un endomorphisme de E.

  ✓ Si P est le polynôme constant égal à 1, $P(u) = \mathrm{Id}_E$.

  ✓ Si x est un vecteur de E, l'écriture $P(u)(x)$ a un sens (c'est l'image de x par l'endomorphisme $P(u)$) mais l'écriture $P(u(x))$ ne veut rien dire !
- Si P, $Q \in \mathbb{R}[X]$, $\lambda \in \mathbb{R}$ et $u \in \mathscr{L}(E)$:
  - $(P+Q)(u) = P(u) + Q(u)$;
  - $(\lambda P)(u) = \lambda P(u)$;
  - $(PQ)(u) = P(u) \circ Q(u)$;
  - les endomorphismes $P(u)$ et $Q(u)$ commutent.
- Soit E un $\mathbb{K}$-espace vectoriel, et u, $v \in \mathscr{L}(E)$, tels que $u \circ v = v \circ u$.
  Alors pour tout $P \in \mathbb{R}[X]$, $\mathrm{Im}(P(u))$ et $\mathrm{Ker}(P(u))$ sont stables par v.
- Soit F un sous-espace vectoriel de E, stable par $u \in \mathscr{L}(E)$ et soit $P \in \mathbb{K}[X]$.
  Alors F est stable par $P(u)$.
- Si λ est une valeur propre de u et si $P \in \mathbb{K}[X]$, alors $P(\lambda)$ est une valeur propre de $P(u)$.

3.8 Polynômes annulateurs

- Soit $u \in \mathscr{L}(E)$. Un polynôme $P \in \mathbb{K}[X]$ s'appelle un polynôme annulateur de u si $P(u) = 0$ (endomorphisme nul).
  Si P est annulateur de u, tout multiple de P est encore annulateur de u.
- Si E est un $\mathbb{K}$-espace vectoriel *de dimension finie*, tout endomorphisme $u \in \mathscr{L}(E)$ possède un polynôme annulateur non nul.

  ✓ Lorsque E n'est pas de dimension finie, un endomorphisme peut ne pas posséder de polynôme annulateur autre que le polynôme nul (considérer par exemple l'endomorphisme u de $\mathbb{K}[X]$ qui à tout polynôme P associe son polynôme dérivé P'.)
- **Deux exemples importants d'utilisation d'un polynôme annulateur**
  Soit $u \in \mathscr{L}(E)$, possédant un polynôme annulateur (non nul) P.

. **Calcul de l'inverse**

Si P s'écrit $P = \sum_{k=0}^{n} a_k X^k$ avec $\underline{a_0 \neq 0}$, la relation $P(u) = 0$ s'écrit aussi : $u \circ \left(\sum_{k=1}^{n} a_i u^{i-1} \right) = -a_0 \mathrm{Id}_E$, ce qui prouve que u est inversible et permet d'exprimer u^{-1} à l'aide des puissances de u.

. **Calcul des puissances**

Pour tout $n \in \mathbb{N}$, la division euclidienne de X^n par P s'écrit : $X^n = PQ_n + R_n$ avec $\deg(R_n) < \deg(P)$.

On a alors : $u^n = P(u) \circ Q_n(u) + R_n(u) = R_n(u)$, ce qui permet d'exprimer u^n à l'aide des premières puissances de u.

3.9 Polynômes de matrices

- Soit $M \in \mathcal{M}_n(\mathbb{K})$. À tout polynôme $P = \sum_{k=0}^{d^\circ(P)} a_k X^k$ de $\mathbb{K}[X]$, on peut associer le polynôme de matrice $P(M) = \sum_{k=0}^{d^\circ(P)} a_k M^k$ (avec $M^0 = I_n$).

  Il est clair que, si $M = M_{\mathscr{B}}(u)$, où u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension n rapporté à une base $\mathscr{B}$, on aura $P(M) = M_{\mathscr{B}}\big(P(u)\big)$.

  On dira qu'un polynôme P est annulateur de M si $P(M) = 0$. Cela équivaut à dire que P est annulateur de u.
- Si deux matrices carrées A et A' sont semblables et si P est un polynôme annulateur de A, P est aussi un polynôme annulateur de A'.
- Si $A \in \mathcal{M}_n(\mathbb{K})$ et si P est un polynôme annulateur de A, P est aussi un polynôme annulateur de $A^\top$.

3.10 Polynômes annulateurs et réduction

E désigne ici un $\mathbb{K}$-espace vectoriel de dimension finie non nulle.

- Soit $u \in \mathscr{L}(E)$ et P un polynôme annulateur non nul de u. Alors toute valeur propre (dans $\mathbb{K}$) de u est une racine de P (dans $\mathbb{K}$).

  Autrement dit : $\mathrm{Sp}(u)$ est *inclus* dans l'ensemble des racines de P.

  ✓ Attention : si P est un polynôme annulateur quelconque de u, toutes ses racines ne sont pas nécessairement valeurs propres de u !
- Si $u \in \mathscr{L}(E)$ possède un polynôme annulateur (non nul) scindé dans $\mathbb{K}[X]$ et n'ayant que des racines simples, alors u est diagonalisable.
- Si u est diagonalisable, le polynôme $\prod_{\lambda \in \mathrm{Sp}(u)} (X - \lambda)$ est annulateur de u.
- Un endomorphisme u d'un $\mathbb{K}$-espace vectoriel de dimension finie est diagonalisable si et seulement si il possède un polynôme annulateur scindé dans $\mathbb{K}[X]$ et à racines simples.
- **Théorème de Cayley-Hamilton**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et u un endomorphisme de E. Alors : $\chi_u(u) = 0_{\mathscr{L}(E)}$.

  Autrement dit : le polynôme caractéristique de u est aussi un polynôme annulateur de u.

3.11 Trigonalisation

E désigne ici un $\mathbb{K}$-espace vectoriel de dimension finie non nulle.

- Un endomorphisme $u \in \mathscr{L}(E)$ est dit trigonalisable s'il existe une base $\mathscr{B}$ de E telle que $M_{\mathscr{B}}(u)$ soit triangulaire supérieure.

  Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est dite trigonalisable si l'endomorphisme a de $\mathbb{K}^n$ qui lui est canoniquement associé est trigonalisable.

  Cela revient à dire que A est semblable à une matrice triangulaire supérieure, c'est-à-dire qu'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $T \in \mathcal{T}_n^+(\mathbb{K})$ telles que : $T = P^{-1}AP$ (ou $A = PTP^{-1}$).

- Un endomorphisme $u \in \mathscr{L}(E)$ est trigonalisable si et seulement si son polynôme caractéristique est scindé dans $\mathbb{K}[X]$.
  En particulier, si $\mathbb{K} = \mathbb{C}$, tout endomorphisme de E est trigonalisable, et toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable.

  ✓ Si $u \in \mathscr{L}(E)$ est trigonalisable et si $\mathscr{B}$ est une base de E dans laquelle la matrice T de u est triangulaire supérieure, alors les éléments diagonaux de T sont exactement les valeurs propres de u (chacune étant comptée avec son ordre de multiplicité).

3.12 Endomorphismes nilpotents

- Un endomorphisme u d'un espace vectoriel E est dit nilpotent s'il existe un entier naturel n tel que $u^n = 0$.
  On appelle alors indice de nilpotence de u le plus petit entier p tel que $u^p = 0$.
- Si u est un endomorphisme nilpotent d'indice p, tout polynôme X^k avec $k \geqslant p$ est annulateur de u.
- Si E est de dimension finie, et si u est nilpotent d'indice p, alors $p \leqslant \dim E$.
- Un endomorphisme u d'un $\mathbb{C}$-espace vectoriel de dimension finie est nilpotent si et seulement si sa seule valeur propre est 0.

  ✓ Ce résultat tombe en défaut si E est un $\mathbb{R}$-espace vectoriel ; par exemple, la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ admet pour seule valeur propre *réelle* 0 mais n'est pas nilpotente.
- Si u est un endomorphisme d'un espace vectoriel de dimension n, u est nilpotent si et seulement si son polynôme caractéristique est égal à X^n.
- Si u est un endomorphisme d'un espace vectoriel de dimension finie, u est nilpotent si et seulement si il existe une base de E dans laquelle la matrice de u est triangulaire supérieure à éléments diagonaux nuls.

Chapitre 4 - Espaces préhilbertiens réels

4.1 Produit scalaire

- Une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel E est une application φ de $E\times E$ dans $\mathbb{R}$ linéaire par rapport à chacune des deux variables, c'est-à-dire :
  - pour tout $y\in E$, l'application $x\mapsto\varphi(x,y)$ est une forme linéaire sur E ;
  - pour tout $x\in E$, l'application $y\mapsto\varphi(x,y)$ est une forme linéaire sur E.

  La forme bilinéaire $\varphi\colon E^2\to\mathbb{R}$ est dite symétrique si :
  $$\forall\,(x,y)\in E^2,\ \varphi(x,y)=\varphi(y,x)\,.$$
  Si φ est une application de E^2 dans $\mathbb{R}$ symétrique et linéaire par rapport à l'une des variables, alors φ est bilinéaire.
- On appelle produit scalaire sur un $\mathbb{R}$-espace vectoriel E une forme bilinéaire symétrique φ qui est, de plus :
  - positive, c'est-à-dire : $\forall\,x\in E,\ \varphi(x,x)\geqslant 0$;
  - et définie, c'est-à-dire : $\forall\,x\in E,\ \varphi(x,x)=0\implies x=0_E$.

  Dire que φ est définie positive peut aussi s'écrire directement :
  $$\forall\,x\in E\setminus\{0\}\,,\ \varphi(x,x)>0\,.$$
  Si φ est un produit scalaire sur un $\mathbb{R}$-espace vectoriel E, le couple (E,φ) s'appelle un espace préhilbertien réel.
  Le produit scalaire de deux vecteurs x et y se note en général : $(x\,|\,y)$ ou $\langle x\,|\,y\rangle$.
- **Exemples**
  - Dans $\mathbb{R}^n$, pour $x=(x_1,\dots,x_n)$ et $y=(y_1,\dots,y_n)$, on peut poser :
    $$\langle x\,|\,y\rangle=x_1y_1+\cdots+x_ny_n\,.$$
    Il s'agit du produit scalaire canonique sur $\mathbb{R}^n$.
    En notant $X\in\mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne $\begin{pmatrix}x_1&\dots&x_n\end{pmatrix}^\top$ et $Y\in\mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne $\begin{pmatrix}y_1&\dots&y_n\end{pmatrix}^\top$, on a aussi : $\langle x\,|\,y\rangle=X^\top Y$.
    ✓ Dans l'écriture ci-dessus, on assimile la matrice $X^\top Y$, de type (1,1), à un nombre réel.
  - Dans $\mathcal{M}_n(\mathbb{R})$, si $A=(a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ et $B=(b_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$, on peut poser :
    $$\langle A\,|\,B\rangle=\operatorname{tr}(A^\top B)=\sum_{1\leqslant i,j\leqslant n}a_{ij}b_{ij}\,.$$
    Il s'agit du produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$.
  - Dans l'ensemble $\mathscr{C}([a\,;b],\mathbb{R})$ des fonctions continues sur le segment $[a\,;b]$, on peut poser :
    $$\langle f\,|\,g\rangle=\int_a^b f(t)g(t)\,\mathrm{d}t$$
    (on remarquera que ce n'est pas un produit scalaire sur l'espace vectoriel $\mathscr{CM}([a\,;b]\mathbb{R})$ des fonctions continues par morceaux sur $[a\,;b]$.)

4.2 Norme associée à un produit scalaire

- Si E est un espace préhilbertien réel, où le produit scalaire est noté $\langle\,\cdot\,|\,\cdot\,\rangle$, on appelle norme (euclidienne) associée l'application :
  $$N\colon x\in E\mapsto\|x\|=\sqrt{\langle x\,|\,x\rangle}\,.$$
  (cela a bien un sens car $\langle x\,|\,x\rangle\geqslant 0$).
- **Exemples**

- Dans $\mathbb{R}^n$, muni du produit scalaire canonique, si $x = (x_1, \ldots, x_n)$ on a : $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.
- Dans $\mathcal{M}_n(\mathbb{R})$, muni du produit scalaire canonique, si $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ on a $\|A\| = \sqrt{\sum_{1 \leqslant i,j \leqslant n} a_{i,j}^2}$.
- Dans l'ensemble $\mathscr{C}([a\,;b],\mathbb{R})$ des fonctions continues sur le segment $[a\,;b]$ muni du produit scalaire vu ci-dessus, on aura : $\|f\| = \sqrt{\int_a^b f(t)^2 \,\mathrm{d}t}$.

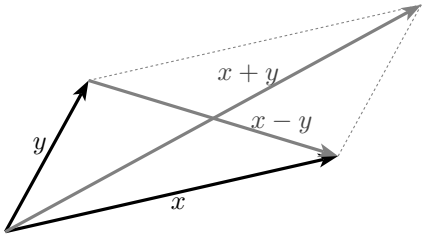
- **Propriétés**
  - $\forall\, x \in E\,, \|x\| = 0_{\mathbb{R}} \iff x = 0_E$.
  - $\forall\, x \in E, \forall\, \lambda \in \mathbb{R},\ \|\lambda x\| = |\lambda|\, \|x\|$.
  - Un vecteur x est dit unitaire si $\|x\| = 1$.

    Si x est un vecteur non nul, alors $\dfrac{x}{\|x\|}$ est unitaire.
  - $\forall\, (x,y) \in E^2,\ \|x+y\|^2 = \|x\|^2 + 2\,\langle x\,|\,y\rangle + \|y\|^2$.
  - $\forall\, (x,y) \in E^2,\ \langle x+y\,|\,x-y\rangle = \|x\|^2 - \|y\|^2$.
  - *Identités de polarisation*

$$\forall\, (x,y) \in E^2,\quad \begin{cases} \langle x\,|\,y\rangle &= \dfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \\ \langle x\,|\,y\rangle &= \dfrac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right) \end{cases}.$$

  - *Identité du parallélogramme*

$$\forall\, (x,y) \in E^2,\ \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$



- **Inégalité de Cauchy-Schwarz**
  Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :

$$|\langle x\,|\,y\rangle| \leqslant \|x\|\,.\,\|y\|$$

  et il y a égalité si et seulement si le système $\{x,y\}$ est lié.
- **Inégalité de Minkowski (ou inégalité triangulaire)**
  Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :

$$\|x+y\| \leqslant \|x\| + \|y\|$$

  et il y a égalité si et seulement si la famille $\{x,y\}$ est positivement liée (c'est-à-dire qu'il existe un réel λ *positif* tel que $x = \lambda y$ ou $y = \lambda x$).
  On en déduit :

$$\forall\, x,\, y \in E,\ \big|\,\|x\| - \|y\|\big| \leqslant \|x-y\|\ .$$

4.3 Orthogonalité

Dans toute la suite, E désigne un espace préhilbertien réel, où le produit scalaire est noté $\langle\,\cdot\,|\,\cdot\,\rangle$.

- **Vecteurs orthogonaux**
  On dit que deux vecteurs x et y de E sont orthogonaux, et on note $x \perp y$, lorsque $\langle x\,|\,y\rangle = 0$.
  La relation $\langle x\,|\,y\rangle = 0$ est équivalente à $\langle y\,|\,x\rangle = 0$. La relation d'orthogonalité est donc symétrique.

- **Orthogonal d'une partie**
  Soit A une partie non vide d'un espace préhilbertien E.
  Un vecteur de E est dit orthogonal à A s'il est orthogonal à tout vecteur de A.
  L'ensemble des vecteurs orthogonaux à A s'appelle l'orthogonal de A, noté $A^\perp$. Ainsi :
  $$A^\perp = \{x \in E \mid \forall\, a \in A,\ \langle x \,|\, a\rangle = 0\}\,.$$
- Si x est un vecteur non nul de E, alors $\{x\}^\perp$ est un hyperplan de E.
  Un supplémentaire de $\{x\}^\perp$ est la droite vectorielle $\mathbb{K}.x$.
- Si A est une partie non vide de E, alors $A^\perp$ est un sous-espace vectoriel de E.
- **Propriétés de l'orthogonal**
  - $E^\perp = \{0\}$ et $\{0\}^\perp = E$.
  - $\forall\,(A,B) \in \mathscr{P}(E)^2,\ A \subset B \implies B^\perp \subset A^\perp$.
  - $\forall\, A \in \mathscr{P}(E),\ A^\perp = \mathrm{Vect}(A)^\perp$.
    Il en résulte que, si F est un sous-espace vectoriel de E et si $(f_i)_{i\in I}$ est une base de F, un vecteur x appartient à $F^\perp$ si et seulement si pour tout $i \in I$ on a $\langle x \,|\, f_i\rangle = 0$.
  - $\forall\, A \in \mathscr{P}(E),\ A \subset \left(A^\perp\right)^\perp$.
  - Si F et G sont des sous-espaces vectoriels de E : $(F+G)^\perp = F^\perp \cap G^\perp$.

    ✓ Si F est un sous-espace vectoriel de E, on n'a pas nécessairement $F = \left(F^\perp\right)^\perp$. Un contre-exemple se trouve en [**??**].
- Soient F et G deux sous-espaces vectoriels d'un espace préhilbertien E.
  On dit que F et G sont des sous-espaces vectoriels orthogonaux, et on note $F \perp G$ lorsque :
  $$\forall\,(x,y) \in F \times G, \quad \langle x \,|\, y\rangle = 0\,,$$
  autrement dit lorsque tout vecteur de F est orthogonal à tout vecteur de G.
  Cela équivaut à : $F \subset G^\perp$ ou $G \subset F^\perp$.
  Remarque : dans $\mathbb{R}^3$, on pourra ainsi parler d'une droite et d'un plan orthogonaux, mais deux plans P_1 et P_2 ne peuvent pas être orthogonaux.
  Cependant, si les droites $P_1^\perp$ et $P_2^\perp$ sont orthogonales, les plans P_1 et P_2 sont dits perpendiculaires.
- Si F et G sont deux sous-espaces vectoriels orthogonaux de E, alors $F \cap G = \{0\}$.
  Plus généralement, des sous-espaces vectoriels orthogonaux deux à deux sont en somme directe.

4.4 Familles orthogonales, orthonormales

- Une famille $(x_i)_{i\in I}$ de vecteurs d'un espace préhilbertien réel E est dite orthogonale si : $\forall\,(i,j) \in I^2,\ i \neq j \implies \langle x_i \,|\, x_j\rangle = 0$
  Elle est dite orthonormale si, de plus, $\|x_i\| = 1$ pour tout $i \in I$, c'est-à-dire si :
  $$\forall\,(i,j) \in I^2,\ \langle x_i \,|\, x_j\rangle = \delta_{ij}\,.$$
  Si $(x_i)_{i\in I}$ est une famille orthogonale de vecteurs *non nuls*, alors la famille des vecteurs $\left(\dfrac{x_i}{\|x_i\|}\right)_{i\in I}$ est orthonormale.
- **Relation de Pythagore**
  Si $(x_1,\ldots,x_p)$ est une famille orthogonale d'un espace préhilbertien réel, alors :
  $$\left\|\sum_{i=1}^{p} x_i\right\|^2 = \sum_{i=1}^{p} \|x_i\|^2\,.$$
  Cas particulier : le théorème de Pythagore
  Soit E un espace préhilbertien réel, et x, y deux vecteurs de E. Alors :
  $$x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2\,.$$
- Si $(x_i)_{i\in I}$ est une famille orthogonale de vecteurs *non nuls*, alors cette famille est libre.
  En particulier, toute famille orthonormale est libre.

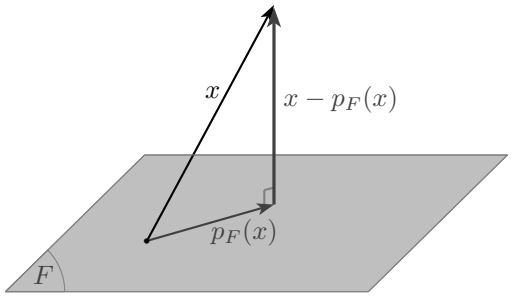
4.5 Bases orthonormales

- Un espace vectoriel euclidien est un espace préhilbertien réel de dimension finie.
- Si E est un espace euclidien, on appelle base orthonormale de E toute base de E qui est aussi une famille orthonormale.
  En vertu de la propriété ci-dessus, pour qu'une famille $(x_1,\dots,x_n)$ de vecteurs d'un espace euclidien forme une base orthonormale, il faut et il suffit que l'on ait simultanément :
  - la famille est orthonormale, c'est-à-dire : $\forall\,(i,j)\in[\![1\,;n]\!]^2,\ \langle x_i\,|\,x_j\rangle=\delta_{i,j}$,
  - et $n=\dim E$.
- Tout espace euclidien (non réduit à $\{0\}$) possède une base orthonormale.
- **Exemples**
  - Dans $\mathbb{R}^n$ muni du produit scalaire canonique, la base canonique est une base orthonormale.
  - Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, la base canonique $(E_{ij})_{1\leqslant i,j\leqslant n}$ est une base orthonormale.
- **Expression analytique du produit scalaire dans une base orthonormale**
  - Soit E un espace euclidien et $\mathscr{B}=(e_1,\dots,e_n)$ une base orthonormale de E.
    Les coordonnées $(x_1,\dots,x_n)$ d'un vecteur x dans cette base sont données par :
    $$\forall\, i\in[\![1\,;n]\!],\ x_i=\langle e_i\,|\,x\rangle\ .$$
    Ainsi : $\forall\, x\in E,\ x=\displaystyle\sum_{i=1}^{n}\langle e_i\,|\,x\rangle\, e_i$.
  - Soient x et y deux vecteurs de E :
    $$x=\sum_{i=1}^{n}x_ie_i\quad\text{ et }y=\sum_{i=1}^{n}y_ie_i\quad\text{avec }(x_1,\dots,x_n)\text{ et }(y_1,\dots,y_n)\in\mathbb{K}^n\,.$$
    En notant $X=\begin{pmatrix}x_1&\dots&x_n\end{pmatrix}^\top$ et $Y=\begin{pmatrix}y_1&\dots&y_n\end{pmatrix}^\top$ les matrices colonnes formées des coordonnées dans $\mathscr{B}$ des vecteurs x et y, on a alors (en assimilant comme d'habitude une matrice de type (1,1) à un nombre réel) :
    $$\langle x\,|\,y\rangle=\sum_{i=1}^{n}x_iy_i=X^\top Y=Y^\top X\quad\text{ et }\quad\|x\|=\sqrt{\sum_{i=1}^{n}x_i^2}\quad,\quad\|x\|^2=X^\top X\,.$$

    ✓ Il est important de noter que les formules données ci-dessus ne sont valables que si la base choisie est *orthonormale*.

4.6 Projection orthogonale sur un sous-espace de dimension finie

- Soit F un sous-espace vectoriel *de dimension finie* d'un espace préhilbertien E. Alors :
  $$E=F\oplus F^\perp\qquad\text{et}\qquad\left(F^\perp\right)^\perp=F\,.$$
  On en déduit que si F et G sont deux sous-espaces vectoriels d'un espace préhilbertien de dimension finie E, alors :
  $$(F\cap G)^\perp\ =\ F^\perp+G^\perp\,.$$
- $F^\perp$ s'appelle le supplémentaire orthogonal de F.
  On peut alors considérer la projection p_F sur F parallèlement à $F^\perp$; on l'appelle projection orthogonale sur F.



- **Expression analytique**
  Si $(e_1,\dots,e_p)$ est une base *orthonormale* de F, on a la formule :

$$\forall\, x \in E,\ p_F(x) = \sum_{i=1}^{p} \langle e_i \,|\, x\rangle\, e_i\,.$$

  Si $(e_1,\dots,e_p)$ est seulement une base *orthogonale* de F, on a la formule :

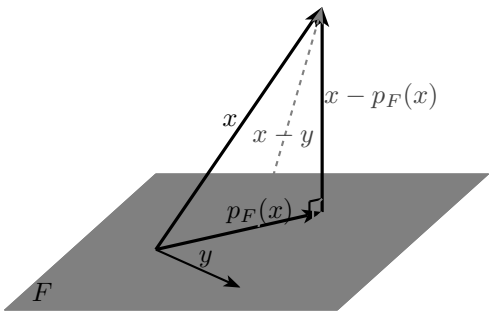
$$\forall\, x \in E,\ p_F(x) = \sum_{i=1}^{p} \frac{\langle e_i \,|\, x\rangle}{\|e_i\|^2}\, e_i\,.$$

  En particulier, le projeté orthogonal sur la droite D de base $a \neq 0$ est :

$$p_D(x) = \frac{\langle a \,|\, x\rangle}{\|a\|^2}\, a\,.$$

- **Lien avec la distance**
  Pour tout x appartenant à E, le vecteur $p_F(x)$ est l'unique vecteur de F réalisant la distance de x à F, c'est-à-dire :

$$\|x - p_F(x)\| = \inf_{y\in F} \|x - y\| = \mathrm{d}(x,F)\,.$$



  D'après le théorème de Pythagore, on a :

$$d(x,F)^2 = \|x\|^2 - \|p_F(x)\|^2\,.$$

  En particulier, si l'on connaît une base *orthonormale* $(e_1,\dots,e_p)$ de F, on aura la formule, pour tout $x \in E$:

$$d(x,F)^2 = \|x\|^2 - \sum_{i=1}^{p} \langle e_i \,|\, x\rangle^2.$$

  Si la base $(e_1,\dots,e_p)$ est seulement *orthogonale*, on aura

$$d(x,F)^2 = \|x\|^2 - \sum_{i=1}^{p} \frac{\langle e_i \,|\, x\rangle^2}{\|e_i\|^2}\,.$$

  Des formules précédentes on déduit immédiatement **l'inégalité de Bessel**.
  Si $(e_1,\dots,e_n)$ est une famille *orthonormale* d'un espace préhilbertien E, pour tout $x \in E$ on a :

$$\sum_{i=1}^{n} \langle e_i \,|\, x\rangle^2 \leqslant \|x\|^2\,.$$

4.7 Procédé d'orthonormalisation de Gram-Schmidt

- Soit E un espace préhilbertien réel.
  On se donne dans E une famille *libre* de vecteurs $(a_1,a_2,\ldots,a_n,\ldots)$, finie ou non.
  Il existe alors une et une seule famille *orthonormale* $(e_1,e_2,\ldots,e_n,\ldots)$ telle que :
  i) pour tout $k \geqslant 1$, $\text{Vect}(e_1,\ldots,e_k) = \text{Vect}(a_1,\ldots,a_k)$;
  ii) pour tout $k \geqslant 1$, $\langle e_k \,|\, a_k\rangle > 0$.
  Dans ce théorème, la condition ii) ne sert qu'à assurer l'unicité des e_k.
- Le procédé de construction est tout aussi important à retenir que le théorème. Les vecteurs e_k se calculent par récurrence de la façon suivante :
  – on pose $e_1 = \dfrac{a_1}{\|a_1\|}$;
  – si l'on a calculé $e_1,\ldots,e_k$ alors on considère le vecteur :
  $$b_{k+1} = a_{k+1} - \sum_{i=1}^{k} \langle e_i \,|\, a_{k+1}\rangle \, e_i \,,$$
  puis l'on pose $e_{k+1} = \dfrac{b_{k+1}}{\|b_{k+1}\|}$.
  Pour retenir facilement ces formules, il suffit de remarquer que le vecteur $\sum_{i=1}^{k} \langle e_i \,|\, a_{k+1}\rangle \, e_i$ n'est autre que le projeté orthogonal $p(a_{k+1)}$ de a_{k+1} sur le sous-espace vectoriel $\text{Vect}(e_1,\ldots,e_k) = \text{Vect}(a_1,\ldots,a_k)$, d'après la formule qui a été rappelée précédemment. Le vecteur $a_{k+1} - p(a_{k+1})$ appartient donc au supplémentaire orthogonal de $\text{Vect}(a_1,\ldots,a_k)$ dans $\text{Vect}(a_1,\ldots,a_k,a_{k+1})$.
  Dans de nombreux cas, trouver une base orthogonale (et non orthonormale) suffit. On trouve alors une telle base $(e'_1,\ldots,e'_n)$ comme suit :
  – on pose $e'_1 = a_1$;
  – si l'on a calculé $e'_1,\ldots,e'_k$ alors on pose :
  $$e'_{k+1} = a_{k+1} - \sum_{i=1}^{k} \frac{\langle e'_i \,|\, a_{k+1}\rangle}{\|e'_i\|^2} e'_i \,.$$
- **Interprétation matricielle**
  Supposons maintenant E de dimension finie n, et soit $\mathscr{B} = (a_1,\ldots,a_n)$ une base de E. Le procédé ci-dessus a permis de construire une base orthonormale $\mathscr{B}' = (e_1,\ldots,e_n)$.
  On peut alors vérifier que la matrice de passage de la base $\mathscr{B}$ à son orthonormalisée par le procédé de Schmidt $\mathscr{B}'$ est une matrice triangulaire supérieure à éléments diagonaux strictement positifs.

4.8 Caractérisation des projecteurs orthogonaux et des symétries orthogonales

E dédigne toujours un espace préhilbertien réel.

- **Caractérisation des projecteurs orthogonaux parmi les projecteurs**
  Soit p un *projecteur* de E ($p \in \mathscr{L}(E)$ et $p \circ p = p$). Les trois propriétés suivantes sont équivalentes :
  (i) p est un projecteur orthogonal (c'est-à-dire $\text{Im}\, p \perp \text{Ker}\, p$),
  (ii) pour tout $x \in E$, $\|p(x)\| \leqslant \|x\|$,
  (iii) pour tous x, $y \in E$, $\langle p(x) \,|\, y\rangle = \langle x \,|\, p(y)\rangle$.

  ✓ La propriété **(ii)** implique que p est une application linéaire continue (cf. chapitre sur les espaces vectoriels normés).
- **Caractérisation des symétries orthogonales parmi les symétries**
  Soit p une *symétrie* de E ($s \in \mathscr{L}(E)$ et $s \circ s = \text{Id}_E$). Les trois propriétés suivantes sont équivalentes :
  (i) s est une symétrie orthogonale,
  (ii) pour tout $x \in E$, $\|s(x)\| = \|x\|$,
  (iii) pour tous x, $y \in E$, $\langle s(x) \,|\, y\rangle = \langle x \,|\, s(y)\rangle$.

4.9 Représentation des formes linéaires d'un espace euclidien

- Soit φ une forme linéaire sur un espace euclidien E.
  Alors il existe un et un seul vecteur $a \in E$ tel que
  $$\forall\, x \in E,\ \varphi(x) = \langle a \,|\, x \rangle\ .$$

- Si H est un hyperplan de E, on sait qu'il existe une forme linéaire non nulle φ telle que $H = \operatorname{Ker} \varphi$. D'après le théorème précédent, il existe donc $a \in E$ tel que $\forall\, x \in E,\ \varphi(x) = \langle a \,|\, x \rangle$.
  Ainsi :
  $$H = \{x \in E \mid \langle a \,|\, x \rangle = 0\}\ .$$
  $\mathbb{R}.a$ est la droite orthogonale à H ; ses vecteurs sont appelés les vecteurs *normaux* à H.

- Si $\mathscr{B} = (e_1, \ldots, e_n)$ est une base orthonormale de E et si φ est une forme linéaire sur E, pour tout $x = \sum_{i=1}^{n} x_i e_i$ dans E on a
  $$\varphi(x) = \sum_{i=1}^{n} x_i \varphi(e_i) = \sum_{i=1}^{n} a_i x_i = \langle a \,|\, x \rangle$$
  où a est le vecteur de coordonnées $a_i = \varphi(e_i)$ dans $\mathscr{B}$. Il s'agit de l'expression analytique de la forme linéaire φ.

Chapitre 5 - Endomorphismes des espaces euclidiens

Dans tout ce chapitre, E désigne un espace euclidien, c'est-à-dire un espace préhilbertien réel de dimension finie, où le produit scalaire est noté $\langle \cdot | \cdot \rangle$.

5.1 Isométries vectorielles

- Une isométrie est un endomorphisme u de E tel que :

$$\forall\,(x,y) \in E^2,\ \langle u(x) \,|\, u(y)\rangle = \langle x \,|\, y\rangle\ .$$

  On dit que u conserve le produit scalaire.

- Une isométrie u peut aussi être caractérisée par l'une des propriétés équivalentes suivantes :

  (i) $\forall\, x \in E,\ \|u(x)\| = \|x\|$ (on dit que u conserve la norme),

  (ii) u transforme une/toute base orthonormale en une base orthonormale.

  La propriété (ii) implique que, si u est une isométrie, u est aussi un endomorphisme bijectif (pour cette raison, une isométrie est aussi appelée automorphisme orthogonal).

- **Exemple :** une symétrie orthogonale par rapport à un sous-espace vectoriel F de E est une isométrie.

- **Sous-espaces stables**
  Soit u une isométrie vectorielle de E.
  Si F est un sous-espace de E stable par u, alors $F^\perp$ est aussi stable par u.

- **Matrices orthogonales**
  - Une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est dite orthogonale si c'est la matrice d'une isométrie *dans une base orthonormale*, ou encore si c'est la matrice de passage d'une base orthonormale à une base orthonormale.
    Cela équivaut à dire que $A^\top A = AA^\top = I_n$.
    Pour reconnaître rapidement si une matrice est orthogonale, il est important de se souvenir de la propriété suivante :

    > une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si ses vecteurs colonnes (ou ses vecteurs lignes), considérés comme éléments de $\mathbb{R}^n$, forment une famille orthonormale de $\mathbb{R}^n$ pour le produit scalaire canonique.

  - L'ensemble $\mathcal{O}_n(\mathbb{R})$ des matrices carrées orthogonales d'ordre n est un sous-groupe du groupe $(\mathrm{GL}_n(\mathbb{R}),\times)$ des matrices inversibles.
  - Si A est une matrice orthogonale, toute valeur propre λ de A, réelle ou complexe, est telle que $|\lambda| = 1$.
    En particulier, les seules valeurs propres possibles d'une isométrie sont ± 1.

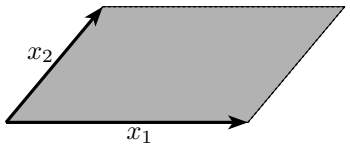
- **Groupe orthogonal**
  - L'ensemble $\mathcal{O}(E)$ des isométries de E est un groupe pour la loi $\circ$ (c'est un sous-groupe de $(\mathrm{GL}(E),\circ)$, appelé groupe orthogonal).
  - Si $u \in \mathcal{O}(E)$, alors $|\det u| = 1$.
    L'ensemble des isométries de déterminant égal à $+1$ est un sous-groupe de $\mathcal{O}(E)$, appelé groupe spécial orthogonal et noté $\mathcal{O}^+(E)$; ses éléments sont appelés des rotations ou isométries positives.
    Une symétrie orthogonale par rapport à un hyperplan de H s'appelle une réflexion ; c'est une isométrie de déterminant -1.
    Une isométrie de déterminant -1 s'appelle aussi une isométrie négative ; leur ensemble est noté $\mathcal{O}^-(E)$.

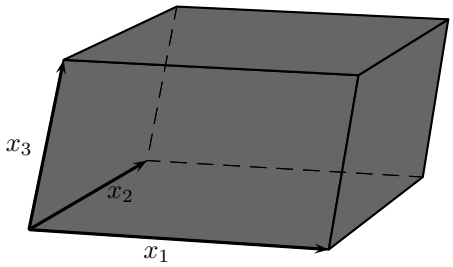
5.2 Produit vectoriel

- **Produit mixte**
  Soit E un espace vectoriel euclidien orienté de dimension n, et $(x_1,\dots,x_n)$ une famille de vecteurs de E.
  Alors le déterminant de cette famille dans *toute* base orthonormale directe est le même ; il s'appelle le produit mixte de $(x_1,\dots,x_n)$ et est noté $[x_1,\dots,x_n]$:

$$[x_1,\dots,x_n] = \det_{\mathscr{B}}(x_1,\dots,x_n) \quad \text{où } \mathscr{B} \text{ est une base orthonormale directe.}$$

Interprétation géométrique
En dimension 2, la valeur absolue du produit mixte de x_1 et x_2 représente l'aire du parallélogramme construit sur x_1 et x_2.
En dimension 3, la valeur absolue du produit mixte de trois vecteurs x_1, x_2 et x_3 représente le volume du parallélépipède construit sur ces trois vecteurs.





- **Produit vectoriel**
  Soit E_3 un espace vectoriel euclidien orienté de dimension 3, et u, v deux vecteurs de E_3. Alors il existe un et un seul vecteur $a \in E_3$ tel que :
  $$\forall\, w \in E_3,\ [u,v,w] = \langle a \,|\, w \rangle\,.$$
  a s'appelle le produit vectoriel de u et v et est noté $u \wedge v$. Il est donc caractérisé par :
  $$\forall\, (u,v,w) \in E_3^3,\ [u,v,w] = \langle u \wedge v \,|\, w \rangle\,.$$
- **Propriétés**
  - Soient u et v deux vecteurs de E_3.
    - $u \wedge v = 0_E \iff$ la famille (u,v) est liée.
    - Le vecteur $u \wedge v$ appartient à $\mathrm{Vect}(u,v)^\perp$. De plus, si la famille (u,v) est libre, la famille $(u,v,u \wedge v)$ est une base directe.
  - L'application $\left\{ \begin{array}{ccc} E_3 \times E_3 & \longrightarrow & E_3 \\ (u,v) & \longmapsto & u \wedge v \end{array} \right.$ est bilinéaire antisymétrique.
  - Si u et v ont pour coordonnées respectives (x,y,z) et (x',y',z') dans une base orthonormale directe $\mathscr{B}$, les coordonnées dans cette même base du produit vectoriel $u \wedge v$ sont :
    $$\begin{vmatrix} y & y' \\ z & z' \end{vmatrix},\ -\begin{vmatrix} x & x' \\ z & z' \end{vmatrix} \text{ et } \begin{vmatrix} x & x' \\ y & y' \end{vmatrix}.$$
  - *Formule du double produit vectoriel*
    $$\forall\, (u,v,w) \in E_3^3,\ u \wedge (v \wedge w) = \langle u \,|\, w \rangle\, v - \langle u \,|\, v \rangle\, w\,.$$

5.3 Classification des isométries en dimension 2

On note ici E un plan vectoriel euclidien, muni d'une base orthonormale $\mathscr{B}$, et on supposera E orienté par le choix de cette base.
Soit u une isométrie de E. Il y a deux cas possibles.

- $\det u = +1$
  u est alors une rotation. Il existe un réel $\theta \in\,]-\pi\,;\pi]$, appelé angle de la rotation, tel que la matrice de u dans *toute* base orthonormale *directe* soit de la forme :
  $$M_{\mathscr{B}}(u) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$
  Sa matrice dans une base orthonormale *indirecte* s'obtient en changeant θ en $-\theta$.
- $\det u = -1$
  Il existe alors un réel θ tel que
  $$M_{\mathscr{B}}(u) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$
  mais ici, θ dépend de la base orthonormale $\mathscr{B}$ choisie.
  Il est alors facile de vérifier, en cherchant les vecteurs propres de u pour les valeurs propres ± 1, que u est une symétrie orthogonale par rapport à une droite D (réflexion). Cette droite est engendrée par le vecteur de coordonnées $\left(\cos\dfrac{\theta}{2}, \sin\dfrac{\theta}{2}\right)$ dans la base $\mathscr{B}$.

5.4 Classification des isométries en dimension 3

On note ici E un espace vectoriel euclidien orienté de dimension 3.
Soit u une isométrie de E. Il y a deux cas possibles.

- $\det u = +1$
  u est une rotation. On cherche alors l'ensemble des vecteurs invariants par u, c'est-à-dire le sous-espace propre $E_1(u)$ associé à la valeur propre 1. On montre qu'il n'y a que deux cas possibles :
  - soit $\dim E_1(u) = 3$: u est alors l'application identique de E.
  - soit $\dim E_1(u) = 1$: l'ensemble des vecteurs invariants de u est ici une droite D, appelée *axe de la rotation*. Le plan $P = D^\perp$ est alors stable par u ; en considérant un vecteur de base unitaire e_1 de l'axe D et une base orthonormale (e_2,e_3) de P telle que la base $\mathscr{B} = (e_1,e_2,e_3)$ soit directe, la matrice de u dans $\mathscr{B}$ est de la forme :
  $$M_{\mathscr{B}}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$
  θ est l'angle de la rotation ; il vérifie la relation : $\operatorname{tr} u = 1 + 2\cos\theta$ (on rappelle que la trace d'un endomorphisme ne dépend pas de la base dans laquelle on écrit sa matrice). Cette relation très simple permet de déterminer facilement $\pm\theta$. Pour déterminer la valeur exacte de θ, dont le signe dépend de l'orientation, nous vous invitons à vous reporter à l'exemple traité en [**??**] page **??**.
- $\det u = -1$
  On montre alors qu'il existe une base orthonormale $\mathscr{B} = (e_1,e_2,e_3)$ (que l'on peut choisir directe) telle que la matrice de u dans cette base soit de la forme :
  $$M_{\mathscr{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$
  - si $\theta \equiv 0 \pmod{2\pi}$, u est la symétrie orthogonale par rapport au plan P de base (e_2,e_3). Ce plan est le sous-espace propre $E_1(u)$.
  - si $\theta \equiv \pi \pmod{2\pi}$, on a alors simplement $u = -\mathrm{Id}_E$.
  - sinon, comme on peut le vérifier facilement à l'aide de la représentation matricielle, u est la composée commutative de la réflexion par rapport à P et de la rotation d'axe $\mathbb{R}e_1$ et d'angle θ ; cet axe est alors l'ensemble des vecteurs x tels que $u(x) = -x$, c'est-à-dire le sous-espace propre $E_{-1}(u)$, et l'ensemble des vecteurs invariants de u est réduit à $\{0\}$. L'angle θ de la rotation vérifie ici la relation : $\operatorname{tr} u = -1 + 2\cos\theta$.
  u s'appelle l'antirotation d'axe $\mathbb{R}e_1$ et d'angle θ.

  ✓ Vous remarquerez que la simple détermination des sous-espaces propres de u et le calcul de sa trace permettent de caractériser u presque entièrement (au signe près de θ dans le cas d'une rotation ou d'une antirotation).

Les isométries sont en fait caractérisées par la dimension du sous-espace des vecteurs invariants, comme le montrent les tableaux récapitulatifs suivants :

- si u est une isométrie en dimension 2

$\dim E_1(u)$	Nature de u	Type d'isométrie
0	rotation	+
1	réflexion d'axe $E_1(u)$	−
2	Id_E	+

- si u est une isométrie en dimension 3

$\dim E_1(u)$	Nature de u	Type d'isométrie
0	antirotation d'axe $E_{-1}(u)$	−
1	rotation d'axe $E_1(u)$	+
2	réflexion par rapport au plan $E_1(u)$	−
3	Id_E	+

5.5 Endomorphismes symétriques

- Un endomorphisme u d'un espace vectoriel euclidien E est dit symétrique si

$$\forall\,(x,y)\in E^2,\ \langle u(x)\,|\,y\rangle = \langle x\,|\,u(y)\rangle\,.$$

  L'ensemble $\mathscr{S}(E)$ des endomorphismes symétriques de E est un sous-espace vectoriel de $\mathscr{L}(E)$.

- Un endomorphisme est symétrique si et seulement si sa matrice dans une/toute base *orthonormale* est symétrique.

- Un projecteur est un endomorphisme symétrique si et seulement si c'est un projecteur orthogonal.

- Une symétrie est un endomorphisme symétrique si et seulement si c'est une symétrie orthogonale.

- Soit u un endomorphisme symétrique de E.
  Si F est un sous-espace vectoriel de E stable par u, alors $F^\perp$ est aussi stable par u.

- **Théorème spectral**
  Soit E un espace vectoriel euclidien, et u un endomorphisme symétrique de E.

  (i) Le polynôme caractéristique de u est scindé dans $\mathbb{R}[X]$.

  (ii) Les sous-espaces propres de u sont orthogonaux deux à deux.

  (iii) Il existe une base orthonormale de E formée de vecteurs propres de u (ce que l'on traduit en disant brièvement : l'endomorphisme u est diagonalisable dans une base orthonormale).

  Puisque la matrice d'un endomorphisme symétrique dans une base orthonormale est une matrice symétrique, on en déduit le théorème suivant.
  Si S est une matrice symétrique réelle, il existe une matrice diagonale D et une matrice orthogonale P telles que :

$$S \;=\; PDP^{-1} \;=\; PD^tP$$

  (on dit aussi que S est orthodiagonalisable).

Chapitre 6 - Espaces vectoriels normés

Dans tout ce chapitre, les espaces vectoriels considérés sont des espaces vectoriels sur $\mathbb{K}$, où $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
Si $\lambda \in \mathbb{K}$, l'écriture $|\lambda|$ désigne la *valeur absolue* de λ si $\mathbb{K} = \mathbb{R}$, et le *module* de λ si $\mathbb{K} = \mathbb{C}$.

6.1 Espaces vectoriels normés

- **Normes sur un espace vectoriel**
  Soit E un $\mathbb{K}$-espace vectoriel. On appelle norme sur E une application $N : E \longrightarrow \mathbb{R}$ telle que :

  (i) $\forall\, x \in E,\ N(x) \geqslant 0$ (*positivité*) ;

  (ii) $\forall\, x \in E,\ N(x) = 0 \implies x = 0$ (*séparation*) ;

  (iii) $\forall\, x \in E,\ \forall\, \lambda \in \mathbb{K},\ N(\lambda x) = |\lambda| N(x)$ (*homogénéité*) ;

  (iv) $\forall\, (x,y) \in E^2,\ N(x+y) \leqslant N(x) + N(y)$ (*inégalité triangulaire*).

  On dit alors que le couple (E,N) est un espace vectoriel normé. On écrit souvent $N(x) = \|x\|$.
  Un vecteur x d'un espace vectoriel normé $(E, \|\cdot\|)$ est dit unitaire si $\|x\| = 1$. Pour tout vecteur $x \in E$ non nul, $\dfrac{x}{\|x\|}$ est unitaire.
- **Inégalité triangulaire**
  Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Alors :
  – $\forall\, (x,y) \in E^2,\ \big|\, \|x\| - \|y\| \,\big| \leqslant \|x+y\| \leqslant \|x\| + \|y\|$.
  – $\forall\, (x_1,\ldots,x_n) \in E^n,\ \forall\, (\lambda_1,\ldots,\lambda_n) \in \mathbb{K}^n,\ \left\| \sum_{i=1}^{n} \lambda_i x_i \right\| \leqslant \sum_{i=1}^{n} |\lambda_i|\, \|x_i\|$.
- **Exemples**
  . La valeur absolue dans $\mathbb{R}$ et le module dans $\mathbb{C}$ sont des normes.
  . Soit $(E, \langle\cdot|\cdot\rangle)$ un espace vectoriel préhilbertien réel. Alors la norme euclidienne associée au produit scalaire $\langle\cdot|\cdot\rangle$, définie par $\|x\| = \sqrt{\langle x\,|\,x\rangle}$, est une norme sur E.
  . Soit E un espace vectoriel de dimension finie n, muni d'une base $\mathscr{B}$.
  Pour tout vecteur x de coordonnées $(x_1,\ldots,x_n) \in \mathbb{K}^n$ dans $\mathscr{B}$, on note :

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \ , \quad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \quad \text{et} \quad \|x\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i| \ .$$

  Alors $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont trois normes sur E.
  . Soient $a < b$ deux réels. Pour toute fonction $f \in \mathscr{C}([a\,;b],\mathbb{K})$, on note :

$$\|f\|_1 = \int_a^b |f(t)|\,\mathrm{d}t \,, \quad \|f\|_2 = \sqrt{\int_a^b |f(t)|^2\,\mathrm{d}t} \quad \text{et} \quad \|f\|_\infty = \sup_{t \in [a;b]} |f(t)| \ .$$

  Alors $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont trois normes sur $\mathscr{C}([a\,;b],\mathbb{K})$.

  ✓ L'existence de $\|f\|_\infty$ est assurée car f est continue sur un segment donc elle est bornée.
- **Distance associée à une norme**
  Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Pour tout $(x,y) \in E^2$, on appelle distance de x à y le réel $d(x,y) = \|x-y\|$.
  Propriétés
  – $\forall\, (x,y) \in E^2,\ d(x,y) \geqslant 0$.
  – $\forall\, (x,y) \in E^2,\ d(x,y) = d(y,x)$.
  – $\forall\, (x,y) \in E^2,\ d(x,y) = 0 \iff x = y$.
  – $\forall\, (x,y,z) \in E^3,\ \big|d(x,y) - d(y,z)\big| \leqslant d(x,z) \leqslant d(x,y) + d(y,z)$.

6.2 Boules

- Soit $(E, \|\cdot\|)$ un espace vectoriel normé, soit a un vecteur de E et soit r un réel positif. On appelle boule *ouverte* de centre a et de rayon r l'ensemble :

$$\mathcal{B}(a,r) = \{x \in E \mid d(a,x) < r\} = \{x \in E \mid \|a - x\| < r\}\,.$$

  On appelle boule *fermée* de centre a et de rayon r l'ensemble :

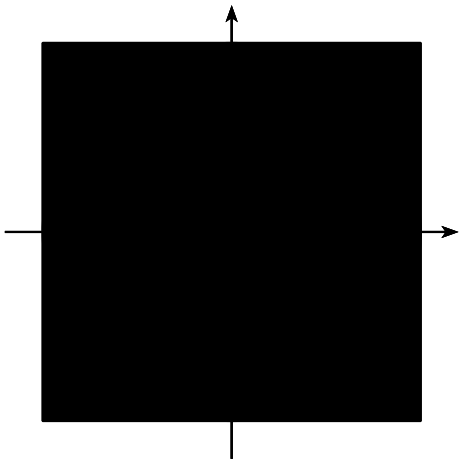
$$\mathcal{B}_f(a,r) = \{x \in E \mid d(a,x) \leqslant r\} = \{x \in E \mid \|a - x\| \leqslant r\}\,.$$

  On appelle sphère de centre a et de rayon r l'ensemble :

$$\mathcal{S}(a,r) = \{x \in E \mid d(a,x) = r\} = \{x \in E \mid \|a - x\| = r\}\,.$$

  Dans le cas où $a = 0_E$ et $r = 1$, on parle de boule unité et de sphère unité.

- **Exemples**
  - Dans $\mathbb{R}$ muni de la valeur absolue $|\cdot|$:
    $\mathcal{B}(a,r) =]a - r\,;a + r[$, $\mathcal{B}_f(a,r) = [a - r\,;a + r]$ et $\mathcal{S}(a,r) = \{a - r\,, a + r\}$.
  - La figure ci-après représente les boules unités fermées pour les trois normes usuelles de $\mathbb{R}^2$.
    - Pour la norme $\|\cdot\|_1$, $B_1 = \big\{(x,y) \in \mathbb{R}^2 \;\big|\; |x| + |y| \leqslant 1\big\}$.
    - Pour la norme $\|\cdot\|_2$, $B_2 = \big\{(x,y) \in \mathbb{R}^2 \;\big|\; x^2 + y^2 \leqslant 1\big\}$.
    - Pour la norme $\|\cdot\|_\infty$, $B_\infty = \big\{(x,y) \in \mathbb{R}^2 \;\big|\; |x| \leqslant 1 \text{ et } |y| \leqslant 1\big\}$.



- **Parties convexes**
  Soit E un $\mathbb{K}$-espace vectoriel. Pour tout $(x,y) \in E^2$, on appelle segment $[x\,;y]$ l'ensemble :

$$[x\,;y] = \{tx + (1-t)y \mid t \in [0\,;1]\}\,.$$

  Une partie A d'un espace vectoriel est dite convexe si, pour tout $(x,y) \in A^2$, le segment $[x\,;y]$ est inclus dans A.
  Dans un espace vectoriel normé $(E, \|\cdot\|)$, toute boule (ouverte ou fermée) est une partie convexe de E.

- **Parties bornées**
  Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Une partie A de E est dite bornée si :

$$\exists\, M \in \mathbb{R},\, \forall\, x \in A,\ \|x\| \leqslant M\,.$$

  Autrement dit, A est bornée s'il existe un réel positif M tel que $A \subset \mathcal{B}_f(0,M)$.

6.3 Équivalence des normes en dimension finie

- On dit que deux normes N_1 et N_2 sur E sont équivalentes s'il existe deux réels strictement positifs α et β tels que :

$$\forall\, x \in E,\ \alpha N_1(x) \leqslant N_2(x) \leqslant \beta N_1(x)\,,$$

  ce qui équivaut à dire que les rapports $\dfrac{N_1}{N_2}$ et $\dfrac{N_2}{N_1}$ sont bornés sur $E \setminus \{0\}$.

- Soit E un $\mathbb{K}$-espace vectoriel de *dimension finie*.
  Toutes les normes sur E sont équivalentes.

  ✓ Ce résultat ne subsiste pas en dimension infinie.
  Par exemple, dans $E = \mathscr{C}([a\,;b],\mathbb{R})$, les normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ ne sont pas équivalentes.

- **Exemple**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension n, muni d'une base $\mathscr{B}$. On note $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ les normes usuelles associées. Alors :

$$\forall\, x \in E,\ \|x\|_\infty \leqslant \|x\|_1 \leqslant n \cdot \|x\|_\infty \quad \text{et} \quad \|x\|_\infty \leqslant \|x\|_2 \leqslant \sqrt{n} \cdot \|x\|_\infty\ .$$

  Cela démontre que les trois normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont équivalentes.

6.4 Suites d'un espace vectoriel normé

Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

- On note $E^{\mathbb{N}}$ l'ensemble des suites définies sur $\mathbb{N}$ à valeurs dans E, c'est-à-dire l'ensemble des applications de la forme $u : \begin{cases} \mathbb{N} & \longrightarrow & E \\ n & \longmapsto & u_n \end{cases}$.
  Pour toutes suites u, $v \in E^{\mathbb{N}}$ et pour tout scalaire $\lambda \in \mathbb{K}$, on note :

$$u + v : n \longmapsto u_n + v_n \quad \text{et} \quad \lambda.u : n \longmapsto \lambda u_n\ .$$

  Muni de ces lois, l'ensemble $(E^{\mathbb{N}}, +\,, .)$ est un $\mathbb{K}$-espace vectoriel.

- **Suites bornées**
  Soit $u \in E^{\mathbb{N}}$ une suite. On dit que u est bornée si :

$$\exists\, M \in \mathbb{R},\ \forall\, n \in \mathbb{N},\ \|u_n\| \leqslant M\ .$$

  Ainsi, u est bornée si l'ensemble $\{u_n \mid n \in \mathbb{N}\}$ est une partie bornée de E.
  L'ensemble noté $\ell^\infty(E)$ des suites bornées d'éléments de E est un sous-espace vectoriel de $E^{\mathbb{N}}$ qui peut être muni d'une structure d'espace vectoriel normé pour la norme $\|\cdot\|_\infty$ définie par :

$$\forall\, u \in \ell^\infty(E),\ \|u\|_\infty = \sup_{n\in\mathbb{N}} \|u_n\|\ .$$

- **Suites convergentes**
  Soit $u \in E^{\mathbb{N}}$ une suite. On dit que u est convergente s'il existe $\ell \in E$ tel que :

$$\forall\, \varepsilon > 0,\ \exists\, n_0 \in \mathbb{N} \text{ tel que } \forall\, n \in \mathbb{N},\ n \geqslant n_0 \implies \|u_n - \ell\| < \varepsilon\ .$$

  Cela signifie que, pour toute boule $\mathcal{B}$ de centre ℓ et de rayon ε strictement positif (arbitrairement petit), la suite u est à valeurs dans $\mathcal{B}$ à partir d'un certain rang.
  Si u est convergente, alors le vecteur ℓ introduit dans la définition est unique. On dit alors que u converge vers ℓ, ou que ℓ est la limite de la suite u, et l'on note :

$$\ell = \lim_{n\to+\infty} u_n\ .$$

  Une suite qui n'est pas convergente est dite divergente.

- **Propriétés des suites convergentes**
  – Dire que $\ell = \lim\limits_{n\to+\infty} u_n$ signifie aussi que la suite réelle $\big(\|u_n - \ell\|\big)$ converge vers 0 quand $n \to +\infty$.

– Toute suite convergente est bornée.

✓ Une suite bornée n'est pas forcément convergente !

– Soient (u_n) et (v_n) deux suites d'éléments de E convergeant respectivement vers ℓ et ℓ'. Alors la suite $(u_n + v_n)$ converge vers $\ell + \ell'$.

– Soit (λ_n) une suite d'éléments de $\mathbb{K}$ convergeant vers $\lambda \in \mathbb{K}$, et (u_n) une suite d'éléments de E convergeant vers $\ell \in E$. Alors, la suite $(\lambda_n u_n)$ converge vers $\lambda\ell$.

Cela entraîne que l'ensemble des suites convergentes à valeurs dans E est un sous-espace vectoriel de $E^{\mathbb{N}}$. De plus, l'application qui à une suite convergente associe sa limite est linéaire.

- **Suites extraites**

  Soit $u \in E^{\mathbb{N}}$. On appelle suite extraite de u toute suite de la forme $\left(u_{\varphi(n)}\right)$, où $\varphi\colon \mathbb{N} \to \mathbb{N}$ est une application strictement croissante.

  Si $u \in E^{\mathbb{N}}$ est une suite convergente de limite $\ell \in E$, toute suite extraite de u converge vers ℓ.

- **Cas de la dimension finie**

  Soit E un $\mathbb{K}$-espace vectoriel de *dimension finie*, et soient N_1 et N_2 deux normes sur E. Soit $u \in E^{\mathbb{N}}$, et $\ell \in E$. N_1 et N_2 étant équivalentes, on a :

  – u bornée pour $N_1 \iff u$ bornée pour N_2.

  – u converge vers ℓ pour $N_1 \iff u$ converge vers ℓ pour N_2.

  ✓ Ce résultat peut tomber en défaut si E n'est pas de dimension finie.

  Exemple : dans $E = \mathscr{C}([0\,;1],\mathbb{R})$ muni des normes usuelles $\|\cdot\|_1$ et $\|\cdot\|_\infty$,

  – la suite $f_n : t \mapsto nt^n$ est bornée pour $\|\cdot\|_1$, mais pas pour $\|\cdot\|_\infty$;

  – la suite $g_n : t \mapsto t^n$ converge vers 0 pour $\|\cdot\|_1$, mais diverge pour $\|\cdot\|_\infty$ (plus précisément, cette suite converge simplement mais pas uniformément, comme nous le verrons dans le chapitre consacré aux suites de fonctions).

- Soit $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie p, muni d'une base $\mathscr{B} = (e_1, \dots, e_p)$. Soit $u \in E^{\mathbb{N}}$, et soit $\ell \in E$. On note $(\ell_1, \dots, \ell_p) \in \mathbb{K}^p$ les coordonnées de ℓ dans la base $\mathscr{B}$ et, pour tout $n \in \mathbb{N}$, on note $\left(u_n^{(1)}, \dots, u_n^{(p)}\right) \in \mathbb{K}^p$ les coordonnées de u_n dans la base $\mathscr{B}$, c'est-à-dire :

$$\ell = \sum_{i=1}^{p} \ell_i e_i \qquad \text{et} \qquad \forall\, n \in \mathbb{N},\ u_n = \sum_{i=1}^{p} u_n^{(i)} e_i\,.$$

  Les suites scalaires $(u_n^{(i)})$ ainsi définies sont appelées les suites coordonnées de u dans la base $\mathscr{B}$. Alors :

$$\lim_{n\to+\infty} u_n = \ell \iff \forall\, i \in [\![1\,;p]\!],\ \lim_{n\to+\infty} u_n^{(i)} = \ell_i\,.$$

  Par conséquent, , l'étude de la convergence d'une suite sur un espace vectoriel normé de dimension finie peut se ramener à l'étude de la convergence de ses suites coordonnées dans une base.

6.5 Topologie d'un espace vectoriel normé

Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

- **Points intérieurs à une partie, intérieur**

  Soit A une partie non vide de E. On dit que $a \in A$ est intérieur à A s'il existe une boule ouverte non vide de centre a qui est incluse dans A.

  L'ensemble des points intérieurs à A s'appelle l'intérieur de A, et se note $\overset{\circ}{A}$. Ainsi :

$$\overset{\circ}{A} = \{a \in A \mid \exists\, r > 0,\ \mathcal{B}(a,r) \subset A\}\,.$$

  Exemple : soient $a \in E$ et $r > 0$. Alors :

$$\overset{\circ}{\mathcal{B}(a,r)} = \mathcal{B}(a,r) \quad , \quad \overset{\circ}{\mathcal{B}_f(a,r)} = \mathcal{B}(a,r) \quad \text{et} \quad \overset{\circ}{\mathcal{S}(a,r)} = \emptyset\,.$$

- **Points adhérents à une partie, adhérence**

  Soit A une partie non vide de E. On dit que $x \in E$ est un point adhérent à A si toute boule ouverte non vide de centre x contient au moins un élément de A.

  L'ensemble des points adhérents à A s'appelle l'adhérence de A et se note $\overline{A}$. Ainsi :

  $$x \in \overline{A} \iff \forall\, r > 0,\, \mathcal{B}(x,r) \cap A \neq \emptyset \iff \forall\, r > 0,\, \exists\, a \in A,\, d(x,a) < r\,.$$

  Exemple : soient $a \in E$ et $r > 0$. Alors :

  $$\overline{\mathcal{B}(a,r)} = \mathcal{B}_f(a,r) \quad , \quad \overline{\mathcal{B}_f(a,r)} = \mathcal{B}_f(a,r) \quad \text{et} \quad \overline{\mathcal{S}(a,r)} = \mathcal{S}(a,r)\ .$$

- **Caractérisation séquentielle des points adhérents à une partie**

  Soit $A \subset E$, et soit $x \in E$. Alors x est un point adhérent à A si et seulement si il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers x.

  L'adhérence de A est donc l'ensemble des limites des suites convergentes d'éléments de A.

- **Partie dense**

  Soit A une partie d'un espace vectoriel normé E. On dit que A est dense dans E si $\overline{A} = E$.

  Cela équivaut à dire que tout élément de E est limite d'une suite d'éléments de A, ou encore que toute boule de E rencontre A.

  Exemple : $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans $\mathbb{R}$.

- **Frontière**

  Soit $A \subset E$. On appelle frontière de A l'ensemble $\partial A = \overline{A} \setminus \overset{\circ}{A}$. La frontière de A est donc l'ensemble des points de E adhérents à A et non intérieurs à A.

  On peut aussi définir la frontière de A par : $\partial A = \overline{A} \cap \overline{\complement_E^A}$, où $\complement_E^A$ désigne le complémentaire de A dans E. La frontière de A est donc l'ensemble des points adhérents à la fois à A et au complémentaire de A.

- **Ouverts**

  Soit A une partie de E. On dit que A est un ouvert de E (ou une partie ouverte de E) si tous les éléments de A sont intérieurs à A, c'est-à-dire :

  $$\forall\, a \in A,\, \exists\, r > 0,\, \mathcal{B}(a,r) \subset A\,.$$

  Propriétés

  - A ouvert $\iff \overset{\circ}{A} = A$.
  - $\emptyset$ et E sont des ouverts.
  - Une boule ouverte est un ouvert.
  - La réunion d'une famille quelconque d'ouverts est un ouvert.
  - L'intersection d'une famille *finie* d'ouverts est un ouvert.

- **Voisinages**

  Soit E un espace vectoriel normé, et $a \in E$. On dit qu'une partie V de E est un voisinage de a s'il existe $r > 0$ tel que $\mathcal{B}(a,r) \subset V$.

  On note $\mathcal{V}(a)$ l'ensemble de voisinages de a.

  Une partie A de E est un ouvert si et seulement si A est voisinage de tous ses points.

  Dans le cas particulier $E = \mathbb{R}$, on définit les voisinages de $+\infty$ (respectivement $-\infty$) comme les parties de $\mathbb{R}$ contenant un intervalle de la forme $]a\,;+\infty[$ (respectivement $]-\infty\,;a[$).

- **Fermés**

  Soit $A \subset E$. On dit que A est un fermé de E (ou une partie fermée de E) si le complémentaire $\complement_E^A$ de A dans E est un ouvert.

  Propriétés

  - A fermé $\iff \overline{A} = A$.
  - $\emptyset$ et E sont des fermés.
  - Une boule fermée, une sphère, sont des fermés.
  - L'intersection d'une famille quelconque de fermés est un fermé.
  - La réunion d'une famille finie de fermés est un fermé.

- **Caractérisation séquentielle des fermés**
  A est un fermé si et seulement si toute suite d'éléments de A qui converge admet pour limite un élément de A.
- **Cas de la dimension finie**
  Toutes les notions topologiques définies ci-dessus dépendent a priori du choix initial de la norme $\|\cdot\|$ sur E. Si E est de *dimension finie*, l'équivalence des normes assure que ces définitions sont indépendantes de la norme choisie.

6.6 Limites

$(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont ici deux espaces vectoriels normés, et A désigne une partie non vide de E.

- Soit f une application de A dans F, et soit $a \in \overline{A}$ un point adhérent à A.
  On dit que f admet une limite en a s'il existe $\ell \in F$ tel que :
  $$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in A,\ \|x-a\|_E \leqslant \alpha \implies \|f(x)-\ell)\|_F \leqslant \varepsilon$$
  ou bien, en termes de voisinages :
  $$\forall V \in \mathcal{V}(\ell),\ \exists U \in \mathcal{V}(a) \text{ tel que } f(U \cap A) \subset V\,.$$
  L'intérêt de la définition faisant intervenir des voisinages est qu'elle s'applique aussi au cas où $a = \pm\infty$ lorsque $E = \mathbb{R}$ et au cas $\ell = \pm\infty$ lorsque $F = \mathbb{R}$.
  Si f admet une limite en a, alors le vecteur ℓ introduit dans la définition est unique. On dit alors que ℓ est la limite de f en a, et on note :
  $$\ell = \lim_{x \to a} f(x)\,.$$
- **Caractérisation séquentielle de la limite**
  Soit f une application de A dans F, soit $a \in \overline{A}$ et soit $\ell \in F$.
  Les propriétés suivantes sont équivalentes :
  (i) $\lim\limits_{x \to a} f(x) = \ell$;
  (ii) pour toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers a, la suite $\big(f(x_n)\big)_{n\in\mathbb{N}}$ converge vers ℓ.
- **Opérations algébriques sur les limites**
  Soit f et g deux applications de A dans F et φ une application de A dans $\mathbb{K}$. Soit $a \in \overline{A}$.
  – Si $\lim\limits_{x\to a} f(x) = \ell$ et $\lim\limits_{x\to a} g(x) = \ell'$ existent, alors $\lim\limits_{x\to a}(f+g)(x)$ existe et est égale à $\ell + \ell'$.
  – Si $\lim\limits_{x\to a} f(x) = \ell$ et $\lim\limits_{x\to a} \varphi(x) = \lambda (\in \mathbb{K})$ existent, alors $\lim\limits_{x\to a}(\varphi.f)(x)$ existe et est égale à $\lambda.\ell$.
- **Limite d'une composée**
  Soient E, F et G trois espaces vectoriels normés, $f\colon A \subset E \to F$ et
  $g\colon B \subset F \to G$ deux applications.
  On suppose $f(A) \subset B$, et $\lim\limits_{x\to a} f(x) = b$.
  Alors $b \in \overline{B}$, et si $\lim\limits_{y\to b} g(y) = \ell \in G$, alors $\lim\limits_{x\to a}(g\circ f)(x) = \ell$.
- **Cas de la dimension finie**
  On suppose ici que f est une application définie sur une partie A d'un espace vectoriel normé E, à valeurs dans un espace vectoriel normé F de *dimension finie*. Soit $\mathscr{B} = (e_1, \ldots, e_p)$ une base de F.
  On note $(f_1, \ldots, f_p)$ les applications coordonnées de f dans la base $\mathscr{B}$, c'est-à-dire :
  $$\forall x \in A,\ f(x) = \sum_{i=1}^{p} f_i(x) e_i \quad \text{avec} \quad f_i : A \to \mathbb{K}\,.$$
  Soit $a \in \overline{A}$. Pour que f admette une limite $\ell = \sum\limits_{i=1}^{p} \ell_i e_i \in F$ quand x tend vers a, il faut et il suffit que, pour tout $i \in [\![1\,;p]\!]$, f_i admette une limite dans $\mathbb{K}$ quand x tend vers a, et on a alors : $\lim\limits_{x\to a} f_i(x) = \ell_i$.

6.7 Continuité

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, et D une partie non vide de E.

- Soit f une application de D dans F, et soit a *appartenant à* D. On dit que f est continue en a si f admet une limite en a (qui vaut alors nécessairement $f(a)$), c'est-à-dire :

$$\forall\,\varepsilon > 0,\ \exists\,\alpha > 0 \text{ tel que } \forall\, x \in D,\ \|x-a\|_E \leqslant \alpha \implies \|f(x)-f(a)\|_F \leqslant \varepsilon\,.$$

  On dit que f est continue sur D si f est continue en tout point de D. On note $\mathscr{C}(D,F)$ l'ensemble des applications continues sur D à valeurs dans F.

- Les notions de limite et de continuité sont inchangées si l'on remplace les normes par des normes équivalentes. Lorsque E et F sont de dimensions finies, ces notions sont donc indépendantes des normes choisies.

  En revanche, si deux normes ne sont pas équivalentes, une application peut être continue au sens d'une norme et pas de l'autre.

- **Caractérisation séquentielle de la continuité**

  Soit f une application de D dans F, et $a \in D$.

  f est continue en a si et seulement si pour toute suite (x_n) d'éléments de D qui converge vers a, la suite $\big(f(x_n)\big)$ converge vers $f(a)$.

- **Opérations algébriques**

  Soient f et g deux applications de D dans F et φ une application de D dans $\mathbb{K}$. Soit $a \in D$.

  – Si f et g sont continues en a, alors $f+g$ est continue en a.

  – Si f et φ sont continues en a, l'application $\varphi.f$ est continue en a.

  En particulier, $\mathscr{C}(D,F)$ est un sous-espace vectoriel de $\mathcal{A}(D,F)$.

- **Composée d'applications continues**

  Soient E, F et G trois espaces vectoriels normés, D une partie de E et Δ une partie de F.

  Soit f une application de D dans F telle que $f(D) \subset \Delta$, et g une application de Δ dans G.

  – Si f est continue en $a \in A$ et si g est continue en $f(a)$, $g \circ f$ est continue en a.

  – Si $f \in \mathscr{C}(D,\Delta)$ et $g \in \mathscr{C}(\Delta,G)$, alors $g \circ f \in \mathscr{C}(D,G)$.

- **Fonctions lipschitziennes**

  Soit f une application de D dans F. On dit que f est lipschitzienne sur D s'il existe un réel k (positif) tel que :

$$\forall\,(x,y) \in E^2,\ \|f(x)-f(y)\|_F \leqslant k\,\|x-y\|_E\,.$$

  On dit alors que f est lipschitzienne de rapport k.

  Si f est lipschitzienne sur D, alors f est continue sur D.

  > ✓ La réciproque est fausse ! Par exemple, la fonction exponentielle est continue sur $\mathbb{R}$, sans être lipschitzienne sur $\mathbb{R}$.

- **Image réciproque d'un ouvert ou d'un fermé**

  Soient E et F deux espaces vectoriels normés, et f une application *continue* de E dans F.

  – L'image réciproque par f de tout ouvert de F est un ouvert de E.

  – L'image réciproque par f de tout fermé de F est un fermé de E.

  Exemples

  ▪ Soient f, $g\colon \mathbb{R} \longrightarrow \mathbb{R}$, continues. Alors :

$$\{x \in \mathbb{R} \mid f(x) < g(x)\} \text{ est un ouvert de } \mathbb{R}\,.$$
$$\{x \in \mathbb{R} \mid f(x) \leqslant g(x)\} \text{ est un fermé de } \mathbb{R}\,.$$

  (En effet, le premier ensemble est l'image réciproque par l'application continue $f-g$ de l'ouvert $]-\infty\,;0[$ et le second est l'image réciproque du fermé $]-\infty\,;0]$).

  ▪ Soit $f \in \mathscr{C}(\mathbb{R},\mathbb{R})$ et soient

$$\Gamma = \big\{(x,y) \in \mathbb{R}^2 \;\big|\; y = f(x)\big\} \text{ le graphe de } f\,,$$
$$\mathscr{E} = \big\{(x,y) \in \mathbb{R}^2 \;\big|\; y > f(x)\big\} \text{ l'épigraphe strict de } f\,.$$

  Alors Γ est une partie fermée de $\mathbb{R}^2$ et $\mathscr{E}$ est une partie ouverte de $\mathbb{R}^2$.

(En effet, il suffit ici de considérer l'application continue $(x,y) \mapsto y - f(x)$ de $\mathbb{R}^2$ dans $\mathbb{R}$. Γ est l'image réciproque du fermé $\{0\}$, et $\mathscr{E}$ celle de l'ouvert $]0\,;+\infty[$).

- L'ensemble $\mathrm{GL}_n(\mathbb{K})$ des matrices carrées inversibles d'ordre n est un ouvert de $\mathcal{M}_n(\mathbb{K})$.

  (En effet, c'est l'image réciproque de l'ouvert $\mathbb{K} \setminus \{0\}$ par l'application « déterminant », qui est continue de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$).

- **Cas de la dimension finie**

  On suppose ici que F est de *dimension finie*. On munit F d'une base $\mathscr{B} = (e_1, \dots, e_p)$. Soit $f \in \mathcal{A}(D,F)$, et $(f_1, \dots, f_p)$ les applications coordonnées de f dans la base $\mathscr{B}$. Alors :

  f est continue en $a \in D$ si et seulement si pour tout $i \in [\![1\,;n]\!]$ f_i est continue en a.

- **Image d'une partie fermée bornée**

  Soit K une partie *fermée* et *bornée* d'un espace vectoriel normé *de dimension finie*, et f une application *continue* de K dans un espace vectoriel normé F.

  Alors $f(K)$ est une partie fermée bornée de F.

  Cela entraîne que si f est continue sur K, alors f est bornée et $\|f\|$ admet sur K un maximum et un minimum.

- **Caractérisation des applications linéaires continues**

  Soient E et F deux espaces vectoriels normés, et $u \in \mathscr{L}(E,F)$. Les propriétés suivantes sont équivalentes :

  (i) u est continue,

  (ii) il existe un réel $k \geqslant 0$ tel que pour tout $x \in E$, $\|u(x)\|_F \leqslant k \|x\|_E$,

  (iii) u est lipschitzienne.

- **Caractérisation des applications multilinéaires continues**

  La définition d'une application p-linéaire a été donnée en [**2.14**].

  Soient $(E_i)_{1\leqslant i\leqslant p}$ p espaces vectoriels normés, muni chacun d'une norme notée $\|\cdot\|_i$.

  On munit l'espace vectoriel produit $E = \prod_{i=1}^{p} E_i$ de la norme $\|\cdot\|_\infty$ définie par :

  $$\forall\, x = (x_1, \dots, x_p) \in E, \quad \|x\|_\infty = \max_{1\leqslant i\leqslant p} \|x_i\|_i \ .$$

  Soit F un espace vectoriel normé, et $u : E \longrightarrow F$ une application p-linéaire. Alors, les propriétés suivantes sont équivalentes :

  (i) u est continue,

  (ii) il existe un réel $k \geqslant 0$ tel que, pour tout $x = (x_1, \dots, x_p) \in E$, on ait

  $$\|u(x_1, \dots, x_p)\|_F \leqslant k \|x_1\|_1 \cdot \|x_2\|_2 \cdots \|x_p\|_p \ .$$

- **Exemples fondamentaux d'applications continues**

  - Si E est de *dimension finie*, alors toute application linéaire $u : E \to F$ est continue sur E.
  - Si $E_1, \dots, E_p$ sont des espaces vectoriels normés de dimensions finies, toute application p-linéaire sur $E_1 \times \cdots \times E_p$ est continue.
  - Toute application polynomiale $\varphi\colon \mathbb{K}^n \to \mathbb{K}$ est continue.
  - L'application $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est continue.

Chapitre 7 - Suites numériques

Les suites numériques ont été étudiées en première année, et nous avons vu dans le chapitre précédent les principales définitions et résultats concernant plus généralement les suites à valeurs vectorielles.
Cependant il nous a semblé important d'introduire un bref chapitre de révisions à ce sujet compte tenu de l'importance qu'ont les suites numériques dans tout le reste du programme. De plus, il existe des résultats spécifiques aux suites à valeurs réelles, liés à la relation d'ordre sur $\mathbb{R}$.

7.1 Limite d'une suite réelle

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles.

- On dit que la suite u admet le réel ℓ pour limite si :
$$\forall\,\varepsilon>0,\ \exists\, n_0\in\mathbb{N},\ \forall\, n\geqslant n_0,\ |u_n-\ell|<\varepsilon\,.$$
- On dit que la suite u tend vers $+\infty$ (respectivement $-\infty$) si :
$$\forall\, A>0,\ \exists\, n_0\in\mathbb{N},\ \forall\, n\geqslant n_0,\ u_n>A \text{ (respectivement } u_n<-A)\,.$$

✓ Dire qu'une suite $(u_n)_{n\in\mathbb{N}}$ est convergente signifie que $\lim\limits_{n\to+\infty} u_n=\ell$ où ℓ est un réel fini. Une suite qui tend vers $\pm\infty$ est divergente.

7.2 Composition des limites

Soit f une fonction numérique et $(u_n)_{n\in\mathbb{N}}$ une suite réelle telle que u_n appartienne au domaine de définition de f au moins à partir d'un certain rang.

– Si $\lim\limits_{n\to+\infty} u_n=a$ et si $\lim\limits_{x\to a} f(x)=\ell$, alors $\lim\limits_{n\to+\infty} f(u_n)=\ell$.

– Si $\lim\limits_{n\to+\infty} u_n=a$ et si f est continue en a, alors $\lim\limits_{n\to+\infty} f(u_n)=f(a)$.

7.3 Propriétés des suites réelles liées à l'ordre

- **Prolongement des inégalités (cas de la convergence)**
Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles convergeant respectivement vers ℓ et ℓ'.
  – Si $\ell<\ell'$, il existe un entier n_0 tel que $\forall\, n\in\mathbb{N},\ n\geqslant n_0 \implies u_n<v_n$.
  – S'il existe un entier n_0 tel que $\forall\, n\in\mathbb{N},\ n\geqslant n_0 \implies u_n\leqslant v_n$, alors $\ell\leqslant\ell'$.

  ✓ Pour appliquer cette propriété, il faut préalablement démontrer *l'existence* des limites.

  ✓ Si pour tout $n\geqslant n_0$ on a $u_n<v_n$, on peut quand même avoir $\ell=\ell'$. Le passage à la limite affaiblit la relation d'ordre.
- **Prolongement des inégalités (cas de la divergence)**
Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles telles que :
$$\exists\, n_0\in\mathbb{N} \text{ tel que } \forall\, n\in\mathbb{N},\ n\geqslant n_0 \implies u_n\leqslant v_n\,.$$
Alors on a :
$$\lim_{n\to+\infty} u_n=+\infty \implies \lim_{n\to+\infty} v_n=+\infty$$
et
$$\lim_{n\to+\infty} v_n=-\infty \implies \lim_{n\to+\infty} u_n=-\infty\,.$$
- **Théorème d'encadrement**
Soit $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ trois suites réelles, telles que :
$$\exists\, n_0\in\mathbb{N} \text{ tel que } \forall\, n\in\mathbb{N},\ n\geqslant n_0 \implies u_n\leqslant v_n\leqslant w_n$$
et
$$\lim_{n\to+\infty} u_n=\lim_{n\to+\infty} w_n=\ell\,.$$
Alors la suite $(v_n)_{n\in\mathbb{N}}$ est convergente, et $\lim\limits_{n\to+\infty} v_n=\ell$.

- **Théorème de la limite monotone**
  Soit (u_n) une suite de réels croissante (respectivement décroissante), au moins à partir d'un certain rang.
  Alors (u_n) est convergente si et seulement si elle est majorée (respectivement minorée).
  (On peut préciser : si (u_n) est croissante non majorée, alors $\lim\limits_{n\to+\infty} u_n = +\infty$).

7.4 Suites adjacentes

- **Définition**
  Soit$(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles. On dit que ces deux suites sont adjacentes si :

$$\begin{cases} (u_n) \text{ est croissante} \\ (v_n) \text{ est décroissante} \\ \lim\limits_{n\to+\infty}(v_n - u_n) = 0 \end{cases}.$$

- **Théorème**
  Deux suites adjacentes sont convergentes, vers la même limite.
- On a de plus : $\forall\,(p,q) \in \mathbb{N}^2$, $u_p \leqslant \ell \leqslant v_q$, avec inégalité stricte si les deux suites sont strictement monotones.
  Cette remarque est importante car elle permet d'obtenir facilement un encadrement de ℓ.

7.5 Suites arithmético-géométriques

On considère ici l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ vérifiant une relation de récurrence de la forme :

$$\forall\, n \in \mathbb{N},\ u_{n+1} = au_n + b$$

avec a, $b \in \mathbb{K}$.

- Si $a = 1$, la suite est tout simplement arithmétique de raison b et l'on a alors $u_n = u_0 + nb$ pour tout entier n.
- Si $a \neq 1$, la suite constante égale à $\ell = \dfrac{b}{1-a}$ vérifie bien la relation de récurrence (il suffit de résoudre l'équation $\ell = a\ell + b$). Si l'on pose alors $v_n = u_n - \ell$ pour tout n, on a :

$$v_{n+1} = u_{n+1} - \ell = (au_n + b) - (a\ell + b) = a(u_n - \ell) = av_n\,,$$

  de sorte que la suite $(v_n)_{n\in\mathbb{N}}$ est géométrique de raison a. On aura donc $v_n = a^n v_0$ pour tout n, d'où l'on tire finalement :

$$\forall\, n \in \mathbb{N},\ u_n = v_n + \ell = a^n\left(u_0 - \frac{b}{1-a}\right) + \frac{b}{1-a}\cdot$$

  ✓ Il ne faut bien sûr pas retenir cette dernière formule par cœur mais seulement la méthode ; les calculs sont alors faciles à refaire.

  ✓ Il faut savoir adapter les résultats dans le cas où la suite n'est définie qu'à partir d'un rang n_0.

  Pour cela, il suffit de se rappeler que les formules deviennent :

  - $u_n = u_{n_0} + (n - n_0)b$ pour une suite arithmétique de raison b ;
  - $v_n = a^{n-n_0}v_{n_0}$ pour une suite géométrique de raison a.

7.6 Suites récurrentes linéaires d'ordre 2

On considère ici l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ qui vérifient une relation de la forme :

$$\forall\, n \in \mathbb{N},\ au_{n+2} + bu_{n+1} + cu_n = 0 \qquad (L)$$

avec a, b, $c \in \mathbb{K}$, $a \neq 0$.
On sait alors que l'ensemble $\mathscr{S}$ des solutions est un $\mathbb{K}$-espace vectoriel de dimension 2 (la démonstration est rappelée en [**??**] page **??**).
On peut alors chercher des éléments particuliers de $\mathscr{S}$: une suite géométrique $n \mapsto r^n$ avec $r \in \mathbb{K}^*$ appartient à $\mathscr{S}$ si et seulement si

$$\forall\, n \in \mathbb{N},\ ar^{n+2} + br^{n+1} + cr^n = 0 \quad \text{soit} \quad ar^2 + br + c = 0\,.$$

L'équation (E_c) : $ar^2 + br + c = 0$ s'appelle l'équation caractéristique de la récurrence.
On étudie alors ses solutions. Il faut distinguer deux cas, selon que le corps de base est $\mathbb{R}$ ou $\mathbb{C}$. En notant Δ le discriminant de l'équation caractéristique, on a les résultats suivants.

- Si $\mathbb{K} = \mathbb{C}$
  - Si $\Delta \neq 0$, (E_c) possède deux racines distinctes r_1 et r_2 ; dans ce cas, les suites $n \mapsto r_1^n$ et $n \mapsto r_2^n$ sont deux solutions linéairement indépendantes de (L). L'ensemble des solutions de (L) est donc l'ensemble des suites de la forme :
  $$u_n = \lambda_1 r_1^n + \lambda_2 r_2^n \ , \ (\lambda_1,\lambda_2) \in \mathbb{C}^2 \, .$$
  - Si $\Delta = 0$, (E_c) possède une racine double r ; dans ce cas, les suites $n \mapsto r^n$ et $n \mapsto nr^n$ sont deux solutions linéairement indépendantes de (L). L'ensemble des solutions de (L) est alors l'ensemble des suites de la forme :
  $$u_n = (\lambda n + \mu) r^n \ , \ (\lambda,\mu) \in \mathbb{C}^2 \, .$$
- Si $\mathbb{K} = \mathbb{R}$
  - Si $\Delta \geqslant 0$, on a les mêmes résultats que ci-dessus (mais avec ici λ_1, λ_2, λ et μ réels).
  - Si $\Delta < 0$, l'équation (E_c) admet deux racines complexes non réelles conjuguées $re^{\pm i\theta}$, et alors les suites $n \mapsto r^n \cos n\theta$ et $n \mapsto r^n \sin n\theta$ sont deux solutions linéairement indépendantes de (L). L'ensemble des solutions de (L) est ici l'ensemble des suites de la forme :
  $$u_n = (\lambda \cos n\theta + \mu \sin n\theta) r^n \ , \ (\lambda,\mu) \in \mathbb{R}^2 \, .$$

✓ Les suites arithmético-géométriques ou les suites récurrentes linéaires peuvent apparaître dans plusieurs types d'exercices (calculs de déterminants, calculs de probabilités...). Il est donc important d'apprendre les méthodes rappelées ci-dessus.

Chapitre 8 - Séries numériques

8.1 Définition de la convergence d'une série

Soit $(u_n)_{n\in\mathbb{N}}$ une suite de nombres complexes. On appelle série de terme général u_n et l'on note $\sum_{n\in\mathbb{N}} u_n$ la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ définie par :

$$\forall\, n \in \mathbb{N},\ S_n = \sum_{k=0}^{n} u_k\,.$$

On dit alors que la série $\sum u_n$ converge si la suite $(S_n)_{n\in\mathbb{N}}$ converge.

Dans ce cas, on appelle somme de la série $\sum u_n$ la limite de la suite $(S_n)_{n\in\mathbb{N}}$; cette somme est notée $\sum_{n=0}^{+\infty} u_n\,.$

✓ La nature (convergence ou divergence) de la série n'est pas modifiée si l'on part d'un entier $n_0 \neq 0$ (mais, en cas de convergence, la valeur de la somme peut être modifiée).

✓ On fera particulièrement attention aux notations : l'écriture $\sum u_n$ ou $\sum_{n\in\mathbb{N}} u_n$ est une simple abréviation pour dire « la série de terme général u_n », et ne doit pas être utilisée dans des calculs ; l'écriture $\sum_{n=0}^{N} u_n$ désigne une « vraie » somme, et la notation $\sum_{n=0}^{+\infty} u_n$, *qui ne doit être utilisée que si la série converge*, désigne la limite de la suite des sommes partielles.

8.2 Condition nécessaire de convergence d'une série

Si la série $\sum u_n$ converge alors $\lim\limits_{n\to+\infty} u_n = 0$.

✓ Cela équivaut au résultat essentiel suivant : une série dont le terme général ne tend pas vers 0 est divergente ; on parle de divergence grossière.

Mais attention : on ne peut rien dire a priori d'une série dont le terme général tend vers 0. Il faut d'ailleurs retenir l'exemple de la série harmonique $\sum_{n\geqslant 1} \frac{1}{n}$, dont le terme général tend vers zéro mais qui est divergente.

8.3 Opérations algébriques sur les séries

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites et λ un scalaire.

- Si les séries $\sum u_n$ et $\sum v_n$ convergent alors la série $\sum(\lambda u_n + v_n)$ converge et :

$$\sum_{n=0}^{+\infty} (\lambda u_n + v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n\,.$$

- Si la série $\sum u_n$ converge et si la série $\sum v_n$ diverge, alors la série $\sum(u_n + v_n)$ diverge.

  ✓ On ne peut rien dire en général de la somme de deux séries divergentes.

- Soit $\sum u_n$ une série de nombres complexes. La série $\sum u_n$ converge si et seulement si les deux séries $\sum \mathcal{R}e\,(u_n)$ et $\sum \mathcal{I}m\,(u_n)$ convergent et dans ce cas :

$$\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \mathcal{R}e\,(u_n) + \mathrm{i} \sum_{n=0}^{+\infty} \mathcal{I}m\,(u_n)\,.$$

  ✓ Cependant, pour étudier une série $\sum u_n$ à valeurs dans $\mathbb{C}$, on préfèrera dans la plupart des cas démontrer l'absolue convergence (voir ci-après) en considérant la série des modules $\sum |u_n|$.

8.4 Convergence absolue

Soit (u_n) une suite de nombres complexes. La série $\sum u_n$ est dite absolument convergente si la série de nombres réels positifs $\sum |u_n|$ est convergente.

Si une série de nombres complexes est absolument convergente, elle est convergente.

✓ La réciproque de ce théorème est fausse ; il faut connaître le contre-exemple classique : la série harmonique alternée $\sum_{n\geqslant 1} \frac{(-1)^n}{n}$ est convergente alors que la série harmonique $\sum_{n\geqslant 1} \frac{1}{n}$ est divergente.

✓ Ce résultat montre l'importance des résultats spécifiques sur les séries à termes positifs qui suivent.

8.5 Règle de comparaison n° 1 pour les séries à termes positifs

Une série à termes réels positifs est convergente si et seulement si la suite de ses sommes partielles est majorée.

8.6 Règle de comparaison n° 2 pour les séries à termes positifs

Soit (u_n) et (v_n) deux suites de nombres réels telles que :

$$0 \leqslant u_n \leqslant v_n \quad \text{à partir d'un certain rang.}$$

– Si la série $\sum v_n$ converge, la série $\sum u_n$ converge.
– Si la série $\sum u_n$ diverge, la série $\sum v_n$ diverge.

8.7 Règle de comparaison n° 3 pour les séries à termes positifs

Soit (u_n) une suite à valeurs réelles et (v_n) une suite de nombres réels *positifs*.
On suppose qu'il existe un réel $\lambda \neq 0$ tel que : $u_n \underset{+\infty}{\sim} \lambda v_n$.

Alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

✓ Il est important de se rappeler que cette règle ne s'applique qu'à des séries de signe constant (au moins à partir d'un certain rang).

8.8 Règle de comparaison n° 4

La règle suivante permet de conclure lorsque l'on connaît « l'ordre de grandeur » du terme général d'une série de nombres complexes. Elle est donc particulièrement efficace.
Soit (u_n) une suite *à valeurs dans* $\mathbb{C}$ et (v_n) une suite *de nombres réels positifs*.
On suppose que : $u_n \underset{+\infty}{=} O(v_n)$ (ou que $u_n \underset{+\infty}{=} o(v_n)$).
Alors, si la série $\sum v_n$ converge, la série $\sum u_n$ converge (absolument).

✓ On fera très attention là encore lors de l'utilisation de cette règle : la série $\sum u_n$ est à termes complexes quelconques, mais il est indispensable que la série $\sum v_n$ à laquelle on la compare soit, elle, à *termes réels positifs*.

8.9 Comparaison à une intégrale

Soit f une fonction continue par morceaux sur un intervalle de la forme $[n_0\,;+\infty[$ ($n_0 \in \mathbb{N}$), *à valeurs réelles positives* et *décroissante*. Alors :

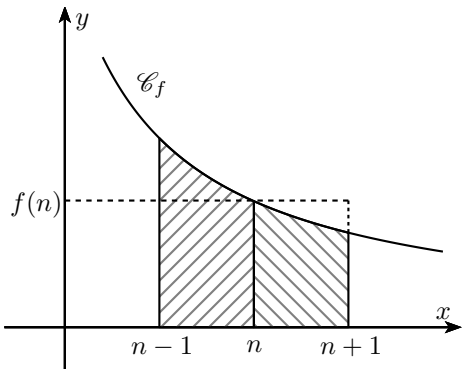
$$\text{la série } \sum_{n\geqslant n_0} f(n) \text{ converge} \iff \int_{n_0}^{+\infty} f \text{ existe}\,.$$

✓ Il est important de retenir non seulement ce théorème, mais aussi et surtout sa démonstration. En effet, la méthode consistant à encadrer le terme général $f(n)$ par des intégrales de f sur des segments permet d'obtenir plus généralement des encadrements des sommes partielles ou du reste de la série lorsqu'elle converge.

Il faut donc retenir (ou savoir retrouver) l'encadrement suivant, valable pour tout $n \geqslant n_0 + 1$:

$$\int_n^{n+1} f(t)\,\mathrm{d}t \leqslant f(n) \leqslant \int_{n-1}^{n} f(t)\,\mathrm{d}t\,.$$

Il est illustré par la figure ci-contre.



8.10 Règle de d'Alembert

Soit $\sum u_n$ une série à termes réels strictement positifs, telle que $\lim\limits_{n\to+\infty}\left(\dfrac{u_{n+1}}{u_n}\right) = \ell$ existe dans $\overline{\mathbb{R}}$ ($\ell \in [0\,;+\infty]$).

– Si $\ell < 1$, la série $\sum u_n$ converge.

– Si $\ell > 1$, la série $\sum u_n$ diverge.

✓ Lorsque $\ell = 1$, on ne peut rien dire a priori . Exemples : $\sum \frac{1}{n}$ et $\sum \frac{1}{n^2}$.

8.11 Séries de référence

- **Séries géométriques**
  Soit q un nombre complexe. La série $\sum q^n$ est convergente si et seulement si $|q| < 1$. Et dans ce cas on a :

$$\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}\,, \quad \text{et plus généralement :} \quad \sum_{n=n_0}^{+\infty} q^n = \frac{q^{n_0}}{1-q}\,.$$

- **Séries de Riemann**
  La série $\displaystyle\sum_{n\in\mathbb{N}^*} \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

- **La série exponentielle**
  Soit z un nombre complexe. Alors la série $\sum \frac{z^n}{n!}$ est absolument convergente. Sa somme s'appelle l'exponentielle de z :

$$\mathrm{e}^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}\,.$$

✓ Lorsque z est un nombre réel, l'inégalité de Taylor-Lagrange permet de démontrer que cette définition coïncide avec celle de la fonction exponentielle que vous connaissiez déjà.

8.12 Produit de Cauchy

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites à valeurs complexes.

On appelle série produit de Cauchy des séries de terme général u_n d'une part et de terme général v_n d'autre part, la série de terme général w_n avec :

$$\forall\, n \in \mathbb{N},\ w_n = \sum_{k=0}^{n} u_k v_{n-k} = \sum_{\substack{p+q=n \\ p,q\in\mathbb{N}}} u_p v_q\,.$$

Si $\sum u_n$ et $\sum v_n$ sont absolument convergentes, alors $\sum w_n$ est absolument convergente, et de plus :

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n\right)\left(\sum_{n=0}^{+\infty} v_n\right).$$

Ce théorème permet de démontrer :

$$\forall\, z,\, z' \in \mathbb{C},\ \mathrm{e}^{z+z'} = \mathrm{e}^z\,\mathrm{e}^{z'}\,.$$

8.13 Séries alternées

Une série à termes réels $\sum u_n$ est dite alternée si la suite $\big((-1)^n u_n\big)$ est de signe constant.

- **Critère spécial des séries alternées, ou critère de Leibniz**
  Soit $\sum u_n$ une série alternée. On suppose que :
  - la suite $\big(|u_n|\big)$ est décroissante ;
  - $\lim\limits_{n\to+\infty} u_n = 0$.

  Alors la série $\sum u_n$ converge.

  ✓ Il est important de noter que ce critère ne donne qu'une condition *suffisante* de convergence pour une série alternée. Par exemple, la série de terme général $u_n = \dfrac{(-1)^n}{n+(-1)^n}$ $(n \geqslant 2)$ est convergente, (on le montre en effectuant un développement limité), cependant elle ne vérifie pas les hypothèses du critère de Leibniz (la suite $(|u_n|)$ n'étant pas décroissante).

- **Résultats complémentaires importants**

  Soit $\sum\limits_{n\in\mathbb{N}} u_n$ une série alternée *qui vérifie les hypothèses du critère spécial*, et soit $S = \sum\limits_{n=0}^{+\infty} u_n$ sa somme. On a les résultats suivants.

  (i) S est comprise entre deux sommes partielles d'indices consécutifs.

  (ii) S est du signe de u_0, et $|S| \leqslant |u_0|$.

  (iii) Si l'on note $R_n = \sum\limits_{k=n+1}^{+\infty} u_k$ le reste d'ordre n, alors R_n est du signe de u_{n+1} et $|R_n| \leqslant |u_{n+1}|$.

8.14 Séries télescopiques

Il s'agit de séries de la forme $\sum v_n$, où $v_n = u_n - u_{n-1}$ $(n \geqslant 1)$.

Les sommes partielles sont alors : $V_n = \sum\limits_{k=1}^{n} v_k = u_n - u_0$ donc on peut énoncer le théorème suivant :

La *série* $\sum\limits_{n\in\mathbb{N}^*} (u_n - u_{n-1})$ converge si et seulement si la *suite* (u_n) converge et, dans ce cas : $\sum\limits_{k=1}^{+\infty} (u_k - u_{k-1}) = \lim\limits_{n\to+\infty} u_n - u_0$.

✓ Cette propriété, qui peut sembler simpliste, est en réalité très importante : elle permet non seulement de calculer la somme de certaines séries, mais aussi, dans certains cas, d'étudier la convergence d'une suite en la ramenant à celle d'une série, donc en utilisant les nombreux outils supplémentaires disponibles sur les séries.

Chapitre 9 - Suites et séries de fonctions

9.1 Convergence d'une suite de fonctions

- Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications définies sur un intervalle I, à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$.
  On dit que cette suite converge simplement sur I si et seulement si :
  $$\text{pour tout } x\in I,\ \lim_{n\to+\infty} f_n(x) \text{ existe (dans } \mathbb{K}).$$
  Dans ce cas, on peut définir une application $f\colon I\longrightarrow\mathbb{K}$ par :
  $$\forall\, x\in I,\ f(x)=\lim_{n\to+\infty} f_n(x)\,.$$
  f s'appelle la limite simple de la suite $(f_n)_{n\in\mathbb{N}}$.
- On dit que la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur I vers f si, à partir d'un certain rang, la suite (f_n-f) est bornée sur I et si
  $$\lim_{n\to+\infty} \|f_n-f\|_\infty^I=0\,,$$
  où $\|f_n-f\|_\infty^I$ désigne $\sup\limits_{x\in I}|f_n(x)-f(x)|$ (norme de la convergence uniforme).

  ✓ Lorsque l'on étudie la convergence uniforme, il est indispensable de préciser sur quelle partie elle a lieu ; la phrase « $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f » ne veut rien dire si l'on ne précise pas « sur I ».

9.2 Régularité de la fonction limite

- **Théorème de la double limite**
  Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications de I dans $\mathbb{K}$, qui converge *uniformément* sur I vers une application $f\colon I\to\mathbb{K}$.
  Soit a une extrémité de I (éventuellement $\pm\infty$). On suppose que, pour tout entier n (au moins à partir d'un certain rang), la limite $\lim\limits_{\substack{x\to a\\ x\in I}} f_n(x)=\ell_n$ existe.
  Alors la suite $(\ell_n)_{n\in\mathbb{N}}$ converge vers un élément $\ell\in\mathbb{K}$, et de plus : $\lim\limits_{\substack{x\to a\\ x\in I}} f(x)=\ell$.
  (c'est-à-dire , en abrégé : $\lim\limits_{x\to a}\lim\limits_{n\to+\infty} f_n(x)=\lim\limits_{n\to+\infty}\lim\limits_{x\to a} f_n(x)$).
- **Continuité**
  Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications de I dans $\mathbb{K}$, qui converge vers une application $f\colon I\to\mathbb{K}$, *la convergence étant uniforme sur tout segment inclus dans I.*
  Si les f_n sont continues sur I (au moins à partir d'un certain rang), alors f est continue sur I.

  ✓ S'il y a convergence uniforme sur I, il y aura bien sûr convergence uniforme sur tout segment inclus dans I ; l'hypothèse du théorème est donc surtout importante lorsque I n'est pas un segment, car pour majorer $|f_n-f|$ il est souvent plus commode de le faire sur un segment.

  ✓ Ce théorème peut dans certains cas servir à démontrer qu'il n'y a pas convergence uniforme, en utilisant un raisonnement par l'absurde.
- **Dérivabilité de la limite d'une suite de fonctions de classe $\mathscr{C}^1$**
  Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications de I dans $\mathbb{K}$. On suppose que :
  - les f_n sont de classe $\mathscr{C}^1$ sur I ;
  - la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur I vers une fonction f ;
  - la suite des dérivées $(f_n')_{n\in\mathbb{N}}$ converge vers une application $g\colon I\to\mathbb{K}$, *la convergence étant uniforme sur tout segment inclus dans I.*

  Alors, la fonction f est de classe $\mathscr{C}^1$ sur I, la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur tout segment inclus dans I, et, pour tout $x\in I$, $f'(x)=g(x)$ (soit, en abrégé, $(\lim f_n)'=\lim f_n'$).
- **Dérivabilité de la limite d'une suite de fonctions de classe $\mathscr{C}^k$**
  Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de classe $\mathscr{C}^k$ ($k\in\mathbb{N}^*$) sur un intervalle I de $\mathbb{R}$, à valeurs dans $\mathbb{K}$. On suppose que :

– pour tout $j \in [\![0\,;k-1]\!]$, la suite de fonctions $(f_n^{(j)})_{n\in\mathbb{N}}$ converge simplement sur I ;

– la suite de fonctions $(f_n^{(k)})_{n\in\mathbb{N}}$ converge simplement sur I vers une fonction g, *la convergence étant uniforme sur tout segment inclus dans* I.

Alors, la fonction $f = \lim\limits_{n\to+\infty} f_n$ est de classe $\mathscr{C}^k$ sur I, on a $f^{(k)} = g$ et pour $j \in [\![0\,;k]\!]$, chaque suite $(f_n^{(j)})_{n\in\mathbb{N}}$ converge uniformément vers $f^{(j)}$ sur tout segment inclus dans I.

- **Intégration sur un segment**
  Soit (f_n) une suite d'applications d'un segment $[a\,;b]$ dans $\mathbb{K}$, qui converge uniformément sur $[a\,;b]$ vers une application $f\colon [a\,;b]\to\mathbb{K}$. Alors :

$$\lim_{n\to+\infty}\int_a^b f_n(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t\,.$$

9.3 Convergence d'une série de fonctions

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'applications d'un intervalle I dans $\mathbb{K}$.
On peut alors considérer la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ définies par :

$$\forall\, x\in I,\ S_n(x) = \sum_{k=0}^{n} u_k(x)\,.$$

Étudier la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$, c'est étudier la suite de fonctions $(S_n)_{n\in\mathbb{N}}$.

- **Convergence simple**
  On dit que la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ converge simplement sur I s'il existe une application $S\colon I\to\mathbb{K}$ telle que la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ converge simplement sur I vers S.
  Cela signifie donc que, pour tout $x \in I$, la série $\sum\limits_{n\in\mathbb{N}} u_n(x)$, à valeurs dans $\mathbb{K}$, converge et que $S(x) = \sum\limits_{n=0}^{+\infty} u_n(x)$.
  S s'appelle alors la somme de la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$. On définit également le reste d'ordre n : $R_n = S - S_n = \sum\limits_{k=n+1}^{+\infty} u_k$. La suite de fonctions $(R_n)_{n\in\mathbb{N}}$ converge simplement sur I vers la fonction nulle.
- **Convergence uniforme**
  On dit que la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ converge uniformément sur I s'il existe une application $S\colon I\to\mathbb{K}$ telle que la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ converge uniformément sur I vers S.
  Cela équivaut à dire que la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ converge simplement sur I et que la suite des restes $(R_n)_{n\in\mathbb{N}}$ converge *uniformément* sur I vers la fonction nulle.

  ✓ La convergence uniforme sur I de la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ implique nécessairement la convergence uniforme de la suite (u_n) vers la fonction nulle sur I (puisque $u_n = R_{n-1} - R_n$).
  En raisonnant par l'absurde, cette remarque peut être utilisée pour montrer qu'une série de fonctions ne converge pas uniformément sur I : il suffit pour cela de trouver une suite (x_n) d'éléments de I telle que la suite $\big(u_n(x_n)\big)$ ne converge pas vers 0.
- **Convergence normale**
  On dit que la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ converge normalement sur I si :

  – les fonctions u_n sont bornées sur I (au moins à partir d'un certain rang),

  – et la série numérique $\sum\limits_{n\geqslant 0} \|u_n\|_\infty^I$ est convergente (en notant comme d'habitude : $\|u_n\|_\infty^I = \sup\limits_{x\in I} |u_n(x)|$).

  Si la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ est *normalement convergente* sur I, alors :

  (i) Pour tout $x\in I$, la série $\sum\limits_{n\in\mathbb{N}} u_n(x)$ est absolument convergente dans $\mathbb{K}$.

  (ii) La série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ est uniformément convergente sur I.

9.4 Convergence uniforme d'une série de fonctions et limite

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur I, à valeurs dans $\mathbb{K}$, et soit $a \in \overline{I}$ (éventuellement $a = \pm\infty$).

On suppose que, pour tout entier n, la limite $\lim\limits_{\substack{x\to a\\ x\in I}} u_n(x) = \ell_n$ existe, et que la série $\sum_{n\in\mathbb{N}} u_n$ est *uniformément convergente sur* I. Notons $S = \sum_{n=0}^{+\infty} u_n$. Alors :

$$\text{la série } \sum_{n\in\mathbb{N}} \ell_n \text{ converge et } \lim_{\substack{x\to a\\ x\in I}} S(x) = \sum_{n=0}^{+\infty} \ell_n$$

(soit, en abrégé : $\lim\limits_a \left(\sum_{n=0}^{+\infty} u_n\right) = \sum_{n=0}^{+\infty} \lim\limits_a u_n$).

Ce résultat porte le nom de *théorème de la double limite*.

9.5 Convergence uniforme d'une série de fonctions et continuité

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur un intervalle I, à valeurs dans $\mathbb{K}$.

Si les u_n sont continues sur I et si la série *converge uniformément sur tout segment inclus dans* I, alors sa somme S est continue sur I.

> ✓ Plutôt que de démontrer la convergence uniforme sur *tout* segment de I, on peut se contenter de montrer la convergence uniforme sur tous les intervalles d'une famille (J_α) telle que $\bigcup\limits_\alpha J_\alpha = I$.

9.6 Convergence d'une série de fonctions et dérivation

- Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur un intervalle I, à valeurs dans $\mathbb{K}$.

  On suppose que :
  - les u_n sont de classe $\mathscr{C}^1$ sur I ;
  - la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge simplement sur I ; on notera S sa somme ;
  - la série de fonctions $\sum_{n\in\mathbb{N}} u'_n$ converge simplement sur I, *la convergence étant uniforme sur tout segment inclus dans* I.

  Alors :
  - la fonction S est de classe $\mathscr{C}^1$ sur I ;
  - la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge uniformément sur tout segment inclus dans I ;
  - pour tout $x \in I$, on a : $S'(x) = \sum_{n=0}^{+\infty} u'_n(x)$.
- Comme pour les suites de fonctions, on a également un énoncé pour démontrer que la somme d'une série de fonctions est de classe $\mathscr{C}^k$.

  Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur un intervalle I, à valeurs dans $\mathbb{K}$. On suppose que :
  - les u_n sont de classe $\mathscr{C}^k$ sur I ;
  - chaque série de fonctions $\sum_{n\in\mathbb{N}} u_n^{(j)}$ pour $j \in [\![0\,;k-1]\!]$ converge simplement sur I ;
  - la série de fonctions $\sum_{n\in\mathbb{N}} u_n^{(k)}$ converge simplement sur I, *la convergence étant uniforme sur tout segment inclus dans* I.

  Alors la fonction somme S est de classe $\mathscr{C}^k$ sur I, chaque série $\sum u_n^{(j)}$ avec $j \in [\![0\,;k]\!]$ converge uniformément vers $S^{(j)}$ sur tout segment de I et :

$$\forall\, j \in [\![0\,;k]\!],\ \forall\, x \in I,\ S^{(j)}(x) = \sum_{n=0}^{+\infty} u_n^{(j)}(x)\,.$$

✓ Là encore, plutôt que de démontrer la convergence uniforme sur *tout* segment de I, on peut se contenter de montrer la convergence uniforme sur tous les intervalles d'une famille (J_α) telle que $\bigcup_\alpha J_\alpha = I$.

9.7 Convergence d'une série de fonctions et intégration sur un segment

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur *un segment* $[a\,;b] \subset \mathbb{R}$, à valeurs dans $\mathbb{K}$.
On suppose que les u_n sont continues sur $[a\,;b]$, et que la série $\sum_{n\in\mathbb{N}} u_n$ converge *uniformément sur* $[a\,;b]$. Notons $S = \sum_{n=0}^{+\infty} u_n$.

Alors S est continue sur $[a\,;b]$, la série $\sum_{n\in\mathbb{N}} \int_a^b u_n(t)\,\mathrm{d}t$ converge, et

$$\int_a^b S(t)\,\mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^b u_n(t)\,\mathrm{d}t\,.$$

Remarque : le théorème s'applique également lorsque les u_n sont seulement *continues par morceaux*, mais il faut alors vérifier la continuité par morceaux de S, celle-ci n'étant plus assurée par la convergence uniforme.

Chapitre 10 - Séries entières

10.1 Rayon de convergence d'une série entière

- Une série entière est une série de fonctions de la forme :

$$\sum_{n\in\mathbb{N}} a_n z^n\,,$$

où z est la variable complexe et $(a_n)_{n\in\mathbb{N}}$ une suite de nombres complexes.

- **Lemme d'Abel**
On suppose qu'il existe un complexe non nul z_0 tel que la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ soit bornée.
  - Pour tout $z\in\mathbb{C}$: $a_n z^n \underset{n\to+\infty}{=} O\left(\left|\dfrac{z}{z_0}\right|^n\right)$.
  - Pour tout complexe z tel que $|z|<|z_0|$, la série numérique $\displaystyle\sum_{n\in\mathbb{N}} a_n z^n$ est absolument convergente.

- **Définition du rayon de convergence**
Soit $\sum\limits_{n\in\mathbb{N}} a_n z^n$ une série entière. L'ensemble des réels positifs r tels que la suite $(a_n r^n)_{n\in\mathbb{N}}$ soit bornée est un intervalle contenant 0.
On appelle alors rayon de convergence R de cette série entière la borne supérieure (dans $\overline{\mathbb{R}}$) de cet intervalle :

$$R=\sup\{r\in\mathbb{R}_+ \mid (a_n r^n) \text{ bornée}\} \quad (R\in\mathbb{R}_+\cup\{+\infty\})\,.$$

- **Disque de convergence**
Soit $\sum\limits_{n\in\mathbb{N}} a_n z^n$ une série entière de rayon de convergence R.
  - Si $R>0$, alors pour tout z tel que $|z|<R$, la série $\sum\limits_{n\in\mathbb{N}} a_n z^n$ est absolument convergente.
  - Si $R<+\infty$ alors pour tout z tel que $|z|>R$, la série $\sum\limits_{n\in\mathbb{N}} a_n z^n$ est (grossièrement) divergente.
  - Si $R=0$, la série ne converge que pour $z=0$.
  - Si $R=+\infty$, la série est convergente (absolument) pour tout $z\in\mathbb{C}$.

  La boule ouverte $\mathcal{B}(0,R)$ dans $\mathbb{C}$ s'appelle le disque de convergence.

  ✓ On ne peut rien dire dans le cas général du comportement de la série entière sur le cercle d'incertitude $\{z\in\mathbb{C} \mid |z|=R\}$.

- **Autres caractérisations du rayon de convergence**
Le rayon de convergence R d'une série entière $\sum\limits_{n\in\mathbb{N}} a_n z^n$ peut être défini par l'une des propriétés suivantes, toutes équivalentes :
  - $R=\sup\left\{|z|\,,\ z\in\mathbb{C}\ \middle|\ \text{la série } \sum\limits_{n\in\mathbb{N}} a_n z^n \text{ est absolument convergente}\right\}$;
  - $R=\sup\left\{|z|\,,\ z\in\mathbb{C}\ \middle|\ \text{la série } \sum\limits_{n\in\mathbb{N}} a_n z^n \text{ est convergente}\right\}$;
  - $R=\sup\left\{|z|\,,\ z\in\mathbb{C}\ \middle|\ \lim\limits_{n\to+\infty} a_n z^n=0\right\}$;
  - $R=\sup\{|z|\,,\ z\in\mathbb{C} \mid \text{la suite } (a_n z^n) \text{ est bornée}\}$ (définition).

- **Utilisation des règles de comparaison**
Soient $\sum\limits_{n\in\mathbb{N}} a_n z^n$ et $\sum\limits_{n\in\mathbb{N}} b_n z^n$ deux séries entières de rayons de convergences respectifs R_a et R_b.
  - Si $a_n \underset{n\to+\infty}{\sim} b_n$, alors $R_a=R_b$.
  - Si $|a_n|\leqslant|b_n|$ (au moins à partir d'un certain rang), alors $R_a\geqslant R_b$.
  - Si $a_n=O(b_n)$ (ou $a_n=o(b_n)$), alors $R_a\geqslant R_b$.

10.2 Opérations sur les séries entières

- **Somme de deux séries entières**
  Soient $\sum_{n\in\mathbb{N}} a_n z^n$ et $\sum_{n\in\mathbb{N}} b_n z^n$ deux séries entières de rayons de convergences respectifs R_a et R_b.
  - Si $R_a \neq R_b$, la série entière $\sum_{n\in\mathbb{N}} (a_n+b_n)z^n$ a pour rayon de convergence $R = \min(R_a,R_b)$.
  - Si $R_a = R_b$, la série entière $\sum_{n\in\mathbb{N}} (a_n+b_n)z^n$ a un rayon de convergence $R \geqslant R_a$.
  - Dans les deux cas on a, pour tout z tel que $|z| < \min(R_a,R_b)$:
  $$\sum_{n=0}^{+\infty}(a_n+b_n)z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n.$$
- **Produit de Cauchy de deux séries entières**
  Soient $\sum_{n\in\mathbb{N}} a_n z^n$ et $\sum_{n\in\mathbb{N}} b_n z^n$ deux séries entières de rayons de convergences respectifs R_a et R_b.
  Posons, pour tout $n \in \mathbb{N}$: $c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{\substack{p,q\in\mathbb{N}\\ p+q=n}} a_p b_q$.
  - Le rayon de convergence R_c de la série entière $\sum_{n\in\mathbb{N}} c_n z^n$ est tel que $R_c \geqslant \min(R_a,R_b)$.
  - De plus, pour tout complexe z tel que $|z| < \min(R_a,R_b)$, on a :
  $$\sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n\right)\left(\sum_{n=0}^{+\infty} b_n z^n\right).$$

10.3 Convergence normale d'une série entière

La série entière $\sum_{n\in\mathbb{N}} a_n x^n$ de la variable réelle x, de rayon de convergence $R > 0$, est normalement convergente sur tout *segment* inclus dans l'intervalle ouvert de convergence $]-R\,;R[$.

✓ Il n'y a pas nécessairement convergence normale ni même convergence uniforme de la série entière sur *tout* l'intervalle de convergence. Le contre-exemple qui suit est classique, et doit être su.

La série entière de la variable réelle $\sum_{n\in\mathbb{N}^*} (-1)^{n-1}\dfrac{x^n}{n}$ *converge simplement vers* $x \mapsto \ln(1+x)$ *sur* $]-1\,;1]$.
Il n'y a pas convergence normale sur $]-1\,;1[$ *(car* $\|u_n\|_\infty^{]-1;1[} = \dfrac{1}{n}$*).*
Il n'y a pas convergence uniforme sur $]-1\,;1[$ *(car le reste* $\sum_{k=n+1}^{+\infty} (-1)^{k-1}\dfrac{x^k}{k}$ *n'est pas borné au voisinage de* -1*).*
Il y a cependant convergence uniforme sur tout segment $[a\,;1]$ *avec* $-1 < a < 1$ *(cela se montre en utilisant la majoration du reste d'une série alternée).*

10.4 Continuité de la somme d'une série entière

Soit $\sum_{n\in\mathbb{N}} a_n x^n$ une série entière de rayon de convergence $R > 0$ ($x \in \mathbb{R}$).

La fonction somme $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est continue sur l'intervalle ouvert de convergence $]-R\,;R[$.

Conséquence

Avec les mêmes notations, f admet au voisinage de 0 un développement limité à tout ordre $p \in \mathbb{N}^*$, obtenu par troncature de la série entière :
$$f(x) = \sum_{n=0}^{p} a_n x^n + O\big(x^{p+1}\big).$$

10.5 Dérivation

- **Séries dérivée et primitive**
  Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière.
  - On appelle série dérivée de cette série la série entière :
  $$\sum_{n\geqslant 1} n a_n z^{n-1} = \sum_{n\geqslant 0} (n+1) a_{n+1} z^n \,.$$
  - On appelle série primitive de cette série la série entière :
  $$\sum_{n\geqslant 0} \frac{a_n}{n+1} z^{n+1} = \sum_{n\geqslant 1} \frac{a_{n-1}}{n} z^n \,.$$
- Une série entière, sa série dérivée et sa série primitive ont même rayon de convergence.
- **Dérivation d'une série entière de la variable réelle**
  Soit (a_n) une suite d'éléments de $\mathbb{C}$, et $\sum_{n\in\mathbb{N}} a_n x^n$ une série entière de la variable réelle x, de rayon de convergence $R > 0$. Posons, pour $x \in]-R\,;R[$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.
  Alors f est de classe $\mathscr{C}^1$ sur $]-R\,;R[$ et, pour tout $x \in]-R\,;R[$:
  $$f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n.$$
  Conséquence
  Avec les mêmes notations, la fonction f est de classe $\mathscr{C}^\infty$ sur sur $]-R\,;R[$ et, pour tout entier naturel k et pour tout $x \in]-R\,;R[$:
  $$f^{(k)}(x) = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} a_n x^{n-k} = \sum_{n=0}^{+\infty} \frac{(n+k)!}{n!} a_{n+k} x^n.$$
  Avec les mêmes notations on a aussi :
  $$\forall\, k \in \mathbb{N},\ a_k = \frac{f^{(k)}(0)}{k!}\,.$$
- **Unicité du développement en série entière**
  Soient deux séries entières $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ et $g(x) = \sum_{n=0}^{+\infty} b_n x^n$, de rayons de convergence non nuls.
  Si f et g coïncident dans un voisinage de 0, on a $a_n = b_n$ pour tout entier n.

10.6 Intégration d'une série entière de la variable réelle

Soit (a_n) une suite d'éléments de $\mathbb{C}$, et $\sum_{n\in\mathbb{N}} a_n x^n$ une série entière de la variable réelle x, de rayon de convergence $R > 0$. Posons, pour $x \in]-R\,;R[$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.

- Pour tout segment $[a\,;b] \subset]-R\,;R[$, on a : $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \sum_{n=0}^{+\infty} \int_a^b (a_n x^n)\,\mathrm{d}x.$
- En particulier, pour tout $x \in]-R\,;R[$, on a : $\displaystyle\int_0^x f(t)\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{a_n x^{n+1}}{n+1}\,.$

10.7 Fonction développable en série entière

On dit qu'une fonction $f\colon \mathbb{C} \to \mathbb{C}$ est développable en série entière dans un voisinage V de 0 s'il existe une série entière $\sum_{n\in\mathbb{N}} a_n z^n$ de rayon de convergence $R > 0$ telle que, pour tout $z \in V$, $f(z) = \sum_{n=0}^{+\infty} a_n z^n$.
D'après les résultats précédents, on a donc :

– f est de classe $\mathscr{C}^\infty$ sur $]-R\,;R[$;

– $\forall\, n \in \mathbb{N},\ a_n = \dfrac{f^{(n)}(0)}{n!}$.

Ainsi, la série entière $\sum\limits_{n\in\mathbb{N}} a_n x^n$ n'est autre que la série de Taylor de f en 0.

✓ Cette propriété n'admet pas de réciproque.

– Il existe des fonctions de classe $\mathscr{C}^\infty$ au voisinage de 0 dont la série de Taylor en 0 a un rayon de convergence nul.

– Il existe des fonctions de classe $\mathscr{C}^\infty$ au voisinage de 0 dont la série de Taylor en 0 a un rayon de convergence non nul, mais dont la somme ne coïncide pas avec f (sauf en 0).

10.8 Développements en série entière usuels

Les développements en série entière des fonctions usuelles sont rappelés page suivante. Quelques remarques concernant ce tableau.

– Les développements en série entière de $\sqrt{1+x}$ et de $\dfrac{1}{\sqrt{1+x}}$ n'y figurent pas, car ils s'obtiennent directement à partir du développement de $(1+x)^\alpha$ pour $\alpha = \pm\dfrac{1}{2}$.

Cependant, il faut savoir écrire ces développements sous forme synthétique. On trouve :

$$\forall\, x \in\,]-1\,;1[,\ \sqrt{1+x} = 1 + \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{(2n-2)!}{2^{2n-1} n!(n-1)!} x^n\,,$$

et

$$\forall\, x \in\,]-1\,;1[,\ \frac{1}{\sqrt{1+x}} = 1 + \sum_{n=1}^{+\infty} (-1)^{n} \frac{(2n)!}{2^{2n} (n!)^2} x^n.$$

– De même, le développements en série entière de Arcsin s'obtient simplement en intégrant le développement de $\dfrac{1}{\sqrt{1-x^2}}$.

Il faut savoir le retrouver ; on obtient :

$$\forall\, x]-1\,;1[,\ \operatorname{Arcsin}(x) = \sum_{n=0}^{+\infty} \frac{(2n)!}{2^{2n}(n!)^2} \frac{x^{2n+1}}{2n+1}\,.$$

Développements en série entière usuels

Fonction	Développement en série entière	Domaine
e^x	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^n}{n!}$	$\mathbb{R}$
$\mathrm{sh}(x)$	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}$	$\mathbb{R}$
$\mathrm{ch}(x)$	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}$	$\mathbb{R}$
$\sin(x)$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$	$\mathbb{R}$
$\cos(x)$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$	$\mathbb{R}$
$\dfrac{1}{1-x}$	$\displaystyle\sum_{n=0}^{+\infty} x^n$	$]-1\,;1[$
$\dfrac{1}{1+x}$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n x^n$	$]-1\,;1[$
$\ln(1+x)$	$\displaystyle\sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}$	$]-1\,;1]$
$\dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right)$	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{2n+1}$	$]-1\,;1[$
$\mathrm{Arctan}(x)$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$	$[-1\,;1]$
$(1+x)^\alpha$	$\displaystyle 1+\sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}\,x^n$	$]-1\,;1[$

Chapitre 11 - Fonctions vectorielles d'une variable réelle

On considère dans ce chapitre des applications définies sur un intervalle I de $\mathbb{R}$ (non réduit à un point), à valeurs dans un espace vectoriel normé E.
Les résultats énoncés généralisent ceux obtenus en première année pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Tous les espaces vectoriels normés intervenant dans ce chapitre seront supposés *de dimension finie*.

11.1 Définitions

- Soit f une application de I dans E, et $t_0 \in I$.
  On dit que f est dérivable en t_0 si $\lim\limits_{\substack{t\to t_0\\ t\neq t_0}} \dfrac{1}{t-t_0}\big(f(t)-f(t_0)\big)$ existe dans E.
  Dans ce cas, cette limite se note $f'(t_0)$ ou $\dfrac{\mathrm{d}f}{\mathrm{d}t}(t_0)$ et s'appelle le vecteur dérivé de f en t_0.

- **Développement limité**
  Dire que f est dérivable en t_0 peut aussi s'écrire :
  $$\exists\, \ell \in E \text{ tel que } f(t) = f(t_0) + (t-t_0).\ell + o(t-t_0)$$
  où $o(t-t_0)$ désigne une fonction vectorielle de la forme $(t-t_0)\varepsilon(t)$ avec $\varepsilon\colon I \to E$ telle que $\lim\limits_{t\to t_0} \varepsilon(t) = 0$.
  Cette dernière expression s'appelle un développement limité de f au voisinage de t_0.
  Dans le cas où f est dérivable en t_0, ce développement limité peut aussi s'écrire sous la forme :
  $$f(t_0+h) = f(t_0) + hf'(t_0) + o(h)\,.$$

- Si f est dérivable en t_0, alors f est continue en t_0.

  ✓ La réciproque de cette propriété est fausse. Il suffit par exemple de considérer l'application $f : x \mapsto |x|$ de $\mathbb{R}$ dans $\mathbb{R}$.

  f est évidemment continue sur $\mathbb{R}$ (elle est lipschitzienne de rapport 1), mais elle n'est pas dérivable en 0 (car $f'_g(0) = -1$ et $f'_d(0) = 1$).

- **Dérivation par coordonnées**
  On suppose E de dimension n, rapporté à une base $(e_1,\dots,e_n)$.
  Soit f une application de I dans E. Pour tout $t \in I$, on peut écrire :
  $$f(t) = \sum_{i=1}^{n} f_i(t)e_i$$
  où les $f_i\colon I \to \mathbb{K}$ sont les applications coordonnées de f.
  Alors, f est dérivable en t_0 si et seulement si, pour tout $i \in [\![1\,;n]\!]$, f_i est dérivable en t_0, et on a alors :
  $$f'(t_0) = \sum_{i=1}^{n} f'_i(t_0)e_i\,.$$

  Exemples
  - Une fonction $f\colon I \longrightarrow \mathbb{C}$ est dérivable en $t_0 \in I$ si et seulement si les fonctions $\mathcal{R}e\,(f)$ et $\mathcal{I}m\,(f)$ le sont, et on a alors $f'(t_0) = (\mathcal{R}e\,(f))'(t_0) + \mathrm{i}(\mathcal{I}m\,(f))'(t_0)$.
  - Une application $A\colon \begin{cases} I & \longrightarrow & \mathcal{M}_{p,q}(\mathbb{K}) \\ t & \longmapsto & (a_{i,j}(t))_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}} \end{cases}$ est dérivable en t_0 si et seulement si pour tout $(i,j) \in [\![1\,;p]\!] \times [\![1\,;q]\!]$ les fonctions coefficients $t \mapsto a_{i,j}(t)$ le sont, et dans ce cas la matrice $A'(t_0)$ est la matrice dont les coefficients sont les $a'_{i,j}(t_0)$.

- **Dérivabilité sur un intervalle**
  - Une application $f\colon I \longrightarrow E$ est dite dérivable sur I si pour tout $t_0 \in I$, f est dérivable en t_0.
    Si tel est le cas, on peut définir l'application dérivée de f, notée f', qui à tout $t \in I$ associe le vecteur $f'(t) \in E$.
    L'ensemble des applications dérivables sur I à valeurs dans E sera noté $\mathscr{D}(I,E)$.

- Une application $f\colon I \longrightarrow E$ est dite de classe $\mathscr{C}^1$ sur I si f est dérivable sur I et si f' est en plus continue sur I.

  On notera $\mathscr{C}^1(I,E)$ l'ensemble des applications de classe $\mathscr{C}^1$ sur I à valeurs dans E.

  ✓ Il existe des fonctions dérivables qui ne sont pas de classe $\mathscr{C}^1$! Autrement dit, l'inclusion $\mathscr{C}^1(I,E) \subset \mathscr{D}(I,E)$ est stricte.

11.2 Opérations sur les dérivées

- Soient f et g deux applications de I dans E, dérivables en t_0.
  Alors l'application $f+g$ est dérivable en t_0, et $(f+g)'(t_0) = f'(t_0) + g'(t_0)$.
- Soit f une application de I dans E, et λ une application de I dans $\mathbb{K}$, dérivables en t_0.
  Alors l'application $\lambda.f$ est dérivable en t_0 et $(\lambda.f)'(t_0) = \lambda'(t_0)f(t_0) + \lambda(t_0)f'(t_0)$.

  En particulier, si $\alpha \in \mathbb{K}$ est constant, αf est dérivable en t_0 et $(\alpha f)'(t_0) = \alpha f'(t_0)$.
  Il en résulte que l'ensemble $\mathscr{D}(I,E)$ est un sous-espace vectoriel de $\mathscr{C}^0(I,E)$, et l'application $f \mapsto f'$ est linéaire de $\mathscr{D}(I,E)$ dans $\mathcal{A}(I,E)$.
- Soient E, F deux espaces vectoriels normés, f une application d'un intervalle I de $\mathbb{R}$ à valeurs dans E, et u une application linéaire de E dans F.
  Si f est dérivable en t_0, $u \circ f$ l'est aussi et : $(u \circ f)'(t_0) = u[f'(t_0)]$.
- Soient E, F, G trois espaces vectoriels normés, et $B\colon E \times F \longrightarrow G$ une application bilinéaire.
  Soient $f\colon I \longrightarrow E$ et $g\colon I \longrightarrow F$ deux applications dérivables en $t_0 \in I$.
  Alors l'application $\varphi\colon \begin{cases} I & \longrightarrow & G \\ t & \longmapsto & B\big(f(t),g(t)\big) \end{cases}$ est dérivable en t_0 et

  $$\varphi'(t_0) = B\big(f'(t_0),g(t_0)\big) + B\big(f(t_0),g'(t_0)\big).$$

  Exemple : soit E un espace préhilbertien réel, muni d'un produit scalaire noté $\langle \cdot | \cdot \rangle$.
  Si f et g sont deux applications définies sur I, à valeurs dans E et dérivables, alors l'application $\varphi\colon \begin{cases} I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \langle f(t) \,|\, g(t) \rangle \end{cases}$ est dérivable et :

  $$\langle f \,|\, g \rangle' = \langle f' \,|\, g \rangle + \langle f \,|\, g' \rangle .$$

  On en déduit ensuite que, si $f\colon I \longrightarrow E$ est dérivable en t_0 et si $f(t_0) \neq 0$, l'application $t \mapsto \|f(t)\|$ est dérivable en t_0, et :

  $$\|f\|'(t_0) = \frac{\langle f(t_0) \,|\, f'(t_0) \rangle}{\|f(t_0)\|} .$$
- Le théorème précédent peut se généraliser.
  Soient $E_1, \ldots, E_n$ et F des espaces vectoriels normés, et $g\colon E_1 \times E_2 \times \cdots \times E_n \longrightarrow F$ une application n-linéaire.
  Pour tout $i \in [\![1\,;n]\!]$, soit $f_i\colon I \longrightarrow E_i$ une application dérivable sur I.
  Alors l'application $\varphi\colon \begin{cases} I & \longrightarrow & F \\ t & \longmapsto & g\big(f_1(t),f_2(t),\ldots,f_n(t)\big) \end{cases}$ est dérivable sur I et, pour tout $t \in I$:

  $$\varphi'(t) = \sum_{i=1}^{n} g\big(f_1(t),\ldots,f_{i-1}(t),f_i'(t),f_{i+1}(t),\ldots,f_n(t)\big).$$

  Exemples

  - Si $f_1, \ldots, f_n$ sont n applications dérivables sur I et à valeurs dans $\mathbb{K}$, leur produit $g\colon t \mapsto f_1(t)f_2(t)\cdots f_n(t)$ est dérivable sur I et pour tout $t \in I$:

    $$g'(t) = f_1'(t)f_2(t)\cdots f_n(t) + f_1(t)f_2'(t)f_3(t)\cdots f_n(t) + \cdots + f_1(t)f_2(t)\cdots f_{n-1}(t)f_n'(t) .$$

  - On sait que l'application $\det\colon \begin{cases} \mathcal{M}_n(\mathbb{K}) & \longrightarrow & K \\ M & \longmapsto & \det M \end{cases}$ est une application n-linéaire des colonnes de la matrice M.

Soit alors $M : \left\{ \begin{array}{ccc} I & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ t & \longmapsto & M(t) \end{array} \right.$ une application dérivable sur I.

Si l'on note $C_1(t), \ldots, C_n(t)$ les colonnes de la matrice $M(t)$, alors les applications $t \mapsto C_i(t)$ sont dérivables sur I (car leurs fonctions coordonnées le sont). Par suite, l'application $\varphi : t \mapsto \det(M(t))$ est dérivable sur I et :

$$\varphi'(t) = \sum_{j=1}^{n} \det\big(C_1(t), \ldots, C_{j-1}(t), C_j'(t), C_{j+1}(t), \ldots, C_n(t)\big).$$

- **Fonction composée**

  Soit $\varphi : I \longrightarrow J$ et $f : J \longrightarrow E$, où I et J sont des intervalles de $\mathbb{R}$ et E un espace vectoriel normé.
  Soit $t_0 \in I$. Si φ est dérivable en t_0 et si f est dérivable en $\varphi(t_0)$, alors $f \circ \varphi$ est dérivable en t_0 et :

$$(f \circ \varphi)'(t_0) = \varphi'(t_0) f'[\varphi(t_0)] .$$

11.3 Dérivées successives

- Soit f une application de I dans E. On peut définir par récurrence, si elles existent, les dérivées successives de f de la façon suivante :
  - on pose $f^{(0)} = f$ (dérivée d'ordre zéro) ;
  - pour $k \in \mathbb{N}^*$, on dira que f est k fois dérivable sur I si $f^{(k-1)}$ est dérivable sur I et on pose alors $f^{(k)} = \big(f^{(k-1)}\big)'$ (dérivée d'ordre k).

  On notera $\mathscr{D}^k(I,E)$ l'ensemble des applications k fois dérivables sur I à valeurs dans E.

  Pour $k \in \mathbb{N}^*$, on dira que f est de classe $\mathscr{C}^k$ sur I si f est k fois dérivable sur I et si $f^{(k)}$ est continue sur I ; on note $\mathscr{C}^k(I,E)$ l'ensemble des applications de classe $\mathscr{C}^k$ sur I à valeurs dans E.
  Enfin on dira que f est de classe $\mathscr{C}^\infty$ sur I si elle est de classe $\mathscr{C}^k$ pour tout $k \in \mathbb{N}$, et on notera $\mathscr{C}^\infty(I,E)$ l'ensemble des applications de classe $\mathscr{C}^\infty$ sur I à valeurs dans E.
- $\mathscr{D}^k(I,E)$, $\mathscr{C}^k(I,E)$ et $\mathscr{C}^\infty(I,E)$ sont des sous-espaces vectoriels de $\mathcal{A}(I,E)$.
  Si $f, g \in \mathscr{D}^k(I,E)$ et si $\lambda, \mu \in \mathbb{K}$, on a : $(\lambda f + \mu g)^{(k)} = \lambda f^{(k)} + \mu g^{(k)}$.
- On suppose E de dimension n, rapporté à une base $(e_1, \ldots, e_n)$.

  Soit f une application de I dans E. Pour tout $t \in I$, on peut écrire $f(t) = \sum_{i=1}^{n} f_i(t) e_i$ (les $f_i : I \to \mathbb{K}$ sont les applications coordonnées).
  Alors, f est k-fois dérivable (respectivement de classe $\mathscr{C}^k$) sur I si et seulement si, pour tout $i \in [\![1\,;n]\!]$, f_i est k fois dérivable (respectivement de classe $\mathscr{C}^k$) sur I, et on a alors :

$$\forall t \in I,\ f^{(k)}(t) = \sum_{i=1}^{n} {f_i}^{(k)}(t) e_i .$$

- **Formule de Leibniz**

  Soient E, F, G trois espaces vectoriels normés, et $B : E \times F \longrightarrow G$ une application bilinéaire.
  Soient $f : I \longrightarrow E$ et $g : I \longrightarrow F$ deux applications n fois dérivables (respectivement de classe $\mathscr{C}^n$) sur I .
  Alors l'application $\varphi : \left\{ \begin{array}{ccc} I & \longrightarrow & G \\ t & \longmapsto & B\big(f(t), g(t)\big) \end{array} \right.$ est n fois dérivable (respectivement de classe $\mathscr{C}^n$) sur I, et l'on a :

$$\forall t \in I,\ \varphi^{(n)}(t) = \sum_{k=0}^{n} \binom{n}{k} B\big(f^{(k)}(t), g^{(n-k)}(t)\big) .$$

- Soit $\varphi : I \longrightarrow J$ et $f : J \longrightarrow E$, où I et J sont des intervalles de $\mathbb{R}$ et E un espace vectoriel normé.
  Si $\varphi \in \mathscr{C}^n(I,J)$ et $f \in \mathscr{C}^n(J,E)$, alors $f \circ \varphi \in \mathscr{C}^n(I,E)$.

11.4 Rappels du cours de 1ère année

On vient ici d'étudier brièvement la dérivabilité d'applications définies sur un intervalle I de $\mathbb{R}$, *à valeurs dans un espace vectoriel normé*.
Dans le cas d'applications définies sur un intervalle I de $\mathbb{R}$, *à valeurs dans* $\mathbb{R}$ (une telle application s'appelle une fonction numérique de la variable réelle), il y a des résultats supplémentaires vus en première année et qu'il est important de rappeler.

- **Théorème de la bijection**
  Soit $f\colon I \longrightarrow \mathbb{R}$ une application *continue* et *strictement monotone* sur un intervalle I de $\mathbb{R}$.
  Alors $J = f(I)$ est un intervalle de $\mathbb{R}$ et $f^{-1} : J \longrightarrow I$ est continue, strictement monotone, de même sens de variation que f.

- **Dérivée de la fonction réciproque**
  Soit $f\colon I \longrightarrow J$ une bijection continue de I sur J.
  Si f est dérivable en $a \in I$ et si $f'(a) \neq 0$, alors f^{-1} est dérivable en $b = f(a)$ et on a :
  $$\big(f^{-1}\big)'(b) = \frac{1}{f'(a)} = \frac{1}{f' \circ f^{-1}(b)}\,.$$

- **Caractérisation des $\mathscr{C}^k$-difféomorphismes**
  Soit $k \in \mathbb{N}^* \cup \{\infty\}$. Soit f une application de classe $\mathscr{C}^k$ d'un intervalle I sur l'intervalle $J = f(I)$, telle que f' *ne s'annule pas sur* I.
  Alors f est bijective de I sur J, et f^{-1} est de classe $\mathscr{C}^k$ sur J.

  ✓ Une telle application s'appelle un $\mathscr{C}^k$-difféomorphisme de I sur J.

- Soit f une fonction numérique définie sur un intervalle I, et admettant en un point t_0 *intérieur à* I un extrémum relatif.
  Si f est dérivable en t_0, alors $f'(t_0) = 0$.

  ✓ Le fait que t_0 est un point intérieur est indispensable ! Par exemple, la fonction $f : \left\{ \begin{array}{ccc} [0\,;1] & \longrightarrow & [0\,;1] \\ t & \longmapsto & t \end{array} \right.$ admet un minimum en 0 et un maximum en 1, mais sa dérivée ne s'annule jamais !

- **Théorème de Rolle**
  Soit f une fonction numérique continue sur l'intervalle fermé $[a\,;b]$ $(a < b)$, dérivable sur l'intervalle ouvert $]a\,;b[$, et telle que $f(a) = f(b)$.
  Alors, il existe $c \in]a\,;b[$ tel que $f'(c) = 0$.

- **Théorème des accroissements finis**
  Soit f une fonction numérique continue sur l'intervalle fermé $[a\,;b]$ $(a < b)$ et dérivable sur l'intervalle ouvert $]a\,;b[$.
  Alors, il existe c appartenant à $]a\,;b[$ tel que : $f(b) - f(a) = (b-a)f'(c)$.
  Interprétation géométrique
  Si f est une fonction numérique continue sur $[a\,;b]$ et dérivable sur $]a\,;b[$, il existe (au moins) un point du graphe de f où la tangente est parallèle à la corde qui joint les points A de coordonnées $\big(a,f(a)\big)$ et B de coordonnées $\big(b,f(b)\big)$.

- **Théorème de prolongement de la dérivée**
  Soit $f\colon [a\,;b] \to \mathbb{R}$. On suppose que :
  - f est continue sur $[a\,;b]$;
  - f est dérivable sur $]a\,;b]$;
  - $\lim\limits_{x \to a^+} f'(x) = \ell \in \overline{\mathbb{R}}$.

  Alors : $\lim\limits_{x \to a^+} \dfrac{f(x) - f(a)}{x - a} = \ell\,.$
  Lorsque ℓ est finie, on en déduit que f est dérivable en a^+ et que $f'_d(a) = \ell$.
  Lorsque $\ell = \pm\infty$, la courbe de f admet au point de coordonnées $\big(a,f(a)\big)$ une tangente verticale, et f n'est pas dérivable ena^+.

Chapitre 12 - Intégration sur un segment

Ce chapitre étend les notions étudiées en 1re année aux fonctions continues par morceaux.
Les applications considérées ici sont définies sur un *segment* $[a\,;b]$ de $\mathbb{R}$ ($a<b$) et à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$.

12.1 Intégrale d'une fonction en escalier sur un segment

- **Définitions**
  Une application $f:[a\,;b]\to\mathbb{K}$ est dite en escalier sur $[a\,;b]$ s'il existe une subdivision $a=x_0<x_1<\ldots<x_{n-1}<x_n=b$ de l'intervalle $[a\,;b]$ telle que la restriction de f à chaque intervalle $]x_{i-1}\,;x_i[$ pour $1\leqslant i\leqslant n$ soit constante. Une telle subdivision est dite adaptée à f.
  Si I est un intervalle quelconque de $\mathbb{R}$, f est dite en escalier sur I si elle l'est sur tout segment inclus dans I.
  L'ensemble $\mathscr{E}sc(I,\mathbb{K})$ des fonctions en escalier de I dans K est un sous-espace vectoriel de $\mathcal{A}(I,\mathbb{K})$.
- **Intégrale d'une fonction en escalier**
  Soit $f\in\mathscr{E}sc([a\,;b],\mathbb{K})$ et $\sigma=(x_0,\ldots,x_n)$ une subdivision adaptée à f, c'est-à-dire :
  $$\forall\, i\in[\![1\,;n]\!],\quad \forall\, t\in\,]x_{i-1}\,;x_i[,\quad f(t)=c_i$$
  où les c_i sont des éléments de $\mathbb{K}$. On pose :
  $$I_\sigma(f)=\sum_{i=1}^{n}(x_i-x_{i-1})c_i\,.$$
  Alors le nombre $I_{\sigma(f)}$ est indépendant de la subdivision σ *adaptée* à f choisie.
  On l'appelle intégrale de f sur $[a\,;b]$. Il est noté $\displaystyle\int_{[a;b]}f$ ou $\displaystyle\int_a^b f$ ou $\displaystyle\int_a^b f(t)\,\mathrm{d}t$.

12.2 Intégrale d'une fonction continue par morceaux sur un segment

- **Définitions**
  Soit $[a\,;b]$ ($a<b$) un segment de $\mathbb{R}$, et f une application de $[a\,;b]$ dans $\mathbb{K}$.
  On dit que f est continue (respectivement de classe $\mathscr{C}^1$) par morceaux sur $[a\,;b]$ s'il existe une subdivision $a=x_0<x_1<\ldots<x_{n-1}<x_n=b$ de l'intervalle telle que la restriction de f à chaque intervalle $]x_{i-1},x_i[$ pour $1\leqslant i\leqslant n$ se prolonge en une application continue (respectivement de classe $\mathscr{C}^1$) sur $[x_{i-1},x_i]$. Une telle subdivision est dite adaptée à f.
  On note $\mathscr{CM}(I,\mathbb{K})$ l'ensemble des applications continues par morceaux de I dans $\mathbb{K}$. C'est un sous-espace vectoriel de de $\mathcal{A}(I,\mathbb{K})$.
- **Approximation d'une fonction continue par morceaux**
  Soit f une application continue par morceaux sur $[a\,;b]$ à valeurs dans $\mathbb{K}$.
  Il existe une suite $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions en escalier sur $[a\,;b]$ qui converge uniformément vers f sur $[a\,;b]$.
- **Intégrale d'une fonction continue par morceaux**
  Soit $f\in\mathscr{CM}([a\,;b],\mathbb{K})$. Pour toute suite $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions en escalier qui converge uniformément vers f sur $[a\,;b]$, la suite $\left(\displaystyle\int_a^b\varphi_n\right)_{n\in\mathbb{N}}$ converge dans $\mathbb{K}$.
  De plus, sa limite ne dépend que de f, et pas de la suite $(\varphi_n)_{n\in\mathbb{N}}$ choisie. Cette limite s'appelle l'intégrale de f sur $[a\,;b]$, et est notée $\displaystyle\int_a^b f$ ou $\displaystyle\int_{[a;b]}f$.

12.3 Propriétés de l'intégrale

- **Linéarité**
  L'application $I:\left\{\begin{array}{ccc}\mathscr{CM}([a\,;b],\mathbb{K}) & \longrightarrow & \mathbb{K}\\ f & \longmapsto & \displaystyle\int_{[a;b]}f\end{array}\right.$ est linéaire.

- **Invariance**

  Soient f, $g \in \mathscr{CM}([a\,;b],\mathbb{K})$. Si f et g ne diffèrent qu'en un nombre fini de points, alors $\displaystyle\int_a^b f = \int_a^b g\,.$

- **Cas des fonctions à valeurs dans $\mathbb{C}$**
  Si $f \in \mathscr{CM}([a\,;b],\mathbb{C})$, on peut écrire : $\forall\, t \in [a\,;b]$, $f(t) = f_1(t) + \mathrm{i} f_2(t)$, $f_1 = \mathcal{R}e\,(f)$ et $f_2 = \mathcal{I}m\,(f)$ étant à valeurs dans $\mathbb{R}$.
  Alors f_1 et f_2 sont continues par morceaux sur $[a\,;b]$ et on a :

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b f_1(t)\,\mathrm{d}t + \mathrm{i}\int_a^b f_2(t)\,\mathrm{d}t\,.$$

- Soit $f \in \mathscr{CM}([a\,;b],\mathbb{K})$. Alors $|f|$ est continue par morceaux sur $[a\,;b]$ et :

$$\left|\int_a^b f\right| \leqslant \int_a^b |f|$$

  (où $|\ |$ désigne la valeur absolue si $\mathbb{K} = \mathbb{R}$ et le module si $\mathbb{K} = \mathbb{C}$).

- **Positivité**

  – Si f est une application continue par morceaux sur $[a\,;b]$, *à valeurs réelles positives*, alors $\displaystyle\int_a^b f \geqslant 0$.

  – Si f, $g \in \mathscr{CM}([a\,;b],\mathbb{R})$ sont telles que $f \leqslant g$ sur $[a\,;b]$, alors $\displaystyle\int_a^b f \leqslant \int_a^b g$.

  – Si f est une fonction *continue* sur $[a\,;b]$, avec $a < b$, *à valeurs réelles positives*, telles que $\displaystyle\int_a^b f = 0$, alors $f = 0$ sur $[a\,;b]$.

- **Inégalité de la moyenne**
  Si $f \in \mathscr{CM}([a\,;b],\mathbb{K})$, on a :

$$\left|\int_a^b f\right| \leqslant (b-a)\,\|f\|_\infty$$

  (où $\|f\|_\infty$ désigne comme d'habitude le réel $\sup\{|f(t)|\,,\ t \in [a\,;b]\}$).

- **Relation de Chasles**
  Soit $f \in \mathscr{CM}([a\,;b],\mathbb{K})$ et $c \in\,]a\,;b[\,$.
  Alors $f\big|_{[a;c]}$ et $f\big|_{[c;b]}$ sont continues par morceaux et l'on a :

$$\int_a^b f = \int_a^c f + \int_c^b f\,.$$

- **Inégalité de Cauchy-Schwarz**
  Pour f, $g \in \mathscr{CM}([a\,;b],\mathbb{R})$ on a :

$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right| \leqslant \sqrt{\int_a^b f(t)^2\,\mathrm{d}t}\cdot\sqrt{\int_a^b g(t)^2\,\mathrm{d}t}\,.$$

  De plus, lorsque f et g sont *continues*, il y a égalité si et seulement si la famille (f,g) est liée.

12.4 Sommes de Riemann

- **Définition**
  Soit $\sigma = (x_0,\ldots,x_n)$ une subdivision de $[a\,;b]$.
  On appelle pas de cette subdivision le réel :

$$\max\{x_{i+1} - x_i,\ i \in [\![0\,;n-1]\!]\}\,.$$

  On appelle famille subordonnée à σ toute famille $c = (c_i)_{1\leqslant i\leqslant n}$ telle que $c_i \in [x_{i-1}\,;x_i]$ pour tout $i \in [\![1\,;n]\!]$.

  Si f est une fonction continue par morceaux sur $[a\,;b]$, à valeurs dans $\mathbb{K}$, le scalaire $S(f,\sigma,c) = \displaystyle\sum_{i=1}^{n}(x_i - x_{i-1})f(c_i)$ s'appelle une somme de Riemann relative à f, σ, c.

- **Théorème**

  Soit f une fonction continue par morceaux sur $[a\,;b]$ à valeurs dans $\mathbb{K}$.

  Alors, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, pour toute subdivision σ de $[a\,;b]$ de pas inférieur à α et toute suite subordonnée c, on ait :

$$\left| S(f,\sigma,c) - \int_a^b f \right| < \epsilon$$

  (autrement dit, les sommes de Riemann associées à f tendent vers $\int_a^b f$ lorsque le pas de la subdivision tend vers 0).

- **Cas particulier**

  On utilise le plus souvent le théorème précédent dans le cas où σ est une subdivision régulière de pas $\dfrac{b-a}{n}$, c'est-à-dire $x_i = a + i\dfrac{b-a}{n}\cdot$ On a dans ce cas :

$$S(f,\sigma,c) = \frac{b-a}{n}\sum_{i=1}^{n} f(c_i) \quad (c_i \in [x_{i-1}\,;x_i])\,.$$

  Dire que le pas de la subdivision tend vers 0 signifie que n tend vers $+\infty$. On a donc :

$$\lim_{n\to+\infty} \frac{b-a}{n}\sum_{i=1}^{n} f(c_i) = \int_a^b f\,.$$

  On peut également choisir $c_i = x_i$ ($c_i = x_{i-1}$ ou $c_i = \dfrac{x_i + x_{i-1}}{2}$ sont aussi des choix possibles). Dans ce cas, la somme de Riemann s'écrit :

$$S(f,\sigma,c) = \frac{b-a}{n}\sum_{i=1}^{n} f\left(a + i\frac{b-a}{n}\right).$$

12.5 Primitives

- **Définitions**
  - Soit f une application *continue* sur un intervalle $I \subset \mathbb{R}$, à valeurs dans $\mathbb{K}$.
    On appelle primitive de f toute application $F\colon I \to \mathbb{K}$, de classe $\mathscr{C}^1$, telle que $F' = f$.
  - Soit f une application *continue par morceaux* sur un intervalle $I \subset \mathbb{R}$, à valeurs dans $\mathbb{K}$.
    On appelle alors primitive de f une application $F\colon I \to \mathbb{K}$ telle que :
    - F est continue et de classe $\mathscr{C}^1$ par morceaux sur I ;
    - et $F'(t) = f(t)$ en tout point t de I où f est continue.
- Si $f \in \mathscr{CM}(I,\mathbb{K})$, deux primitives quelconques de f diffèrent d'une constante.
- **Théorème fondamental**

  Si $f \in \mathscr{CM}(I,\mathbb{K})$, et si $a \in I$, l'application $F_a\colon \left\{\begin{array}{ccc} I & \longrightarrow & \mathbb{K} \\ x & \longmapsto & \displaystyle\int_a^x f(t)\,\mathrm{d}t \end{array}\right.$ est une primitive de f.

  C'est l'unique primitive de f qui s'annule en a.
- **Conséquences**
  - Si f est continue par morceaux sur $[a\,;b]$ et si F est une primitive de f, on a :

$$\int_a^b f = F(b) - F(a) \quad \text{ce que l'on écrit : } \int_a^b f = \big[F(t)\big]_a^b\,.$$

  - Si f est *continue* et *de classe $\mathscr{C}^1$ par morceaux* sur $[a\,;b]$, on a :

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

  (avec, ici, un abus de notation, f' n'étant définie que sur $[a\,;b]$ privé d'un nombre fini de points.)

- **Généralisation : intégrale fonction de ses bornes**
  Soit f une application continue de I dans $\mathbb{K}$, et u, v deux applications de classe $\mathscr{C}^1$ définies sur un intervalle J, à valeurs dans I.
  Alors, la fonction $F : \begin{cases} J & \longrightarrow & \mathbb{K} \\ x & \longmapsto & \displaystyle\int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t \end{cases}$ est de classe $\mathscr{C}^1$ sur J et

$$\forall\, x \in J,\ F'(x) = v'(x) f\big[v(x)\big] - u'(x) f\big[u(x)\big]\,.$$

12.6 Intégration par parties

Soient f et g des applications *continues* et *de classe $\mathscr{C}^1$ par morceaux* sur I, à valeurs dans $\mathbb{K}$. Alors :

$$\forall\,(a,b) \in I^2,\ \int_a^b f'g = \big[f(t)g(t)\big]_a^b - \int_a^b fg'\,.$$

12.7 Changement de variable

- *Cas d'une fonction continue*
  Soit $f \in \mathscr{C}^0(I,\mathbb{K})$ et $\varphi \in \mathscr{C}^1([a\,;b],\mathbb{R})$ telle que $\varphi([a\,;b]) \subset I$. Alors :

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,\mathrm{d}t = \int_a^b \varphi'(u).f[\varphi(u)]\,\mathrm{d}u\,.$$

- *Extension aux fonctions continues par morceaux*
  Soit $f \in \mathscr{CM}(I,\mathbb{K})$ et $\varphi \in \mathscr{C}^1([a\,;b],\mathbb{R})$, strictement monotone, telle que $\varphi([a\,;b]) \subset I$. Alors :

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,\mathrm{d}t = \int_a^b \varphi'(u).f[\varphi(u)]\,\mathrm{d}u\,.$$

12.8 Formules de Taylor

- **Formule de Taylor avec reste intégrale**
  Soit $f : I \longrightarrow \mathbb{K}$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N})$ sur I et de classe $\mathscr{C}^{n+1}$ par morceaux sur I. Alors, pour tous a, $b \in I$:

$$f(b) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t\,.$$

- **Inégalité de Taylor-Lagrange**
  Soit $f : I \longrightarrow \mathbb{K}$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N})$ sur I et de classe $\mathscr{C}^{n+1}$ par morceaux sur I. Alors, pour tous a, $b \in I$:

$$\left| f(b) - \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) \right| \leqslant \frac{|b-a|^{n+1}}{(n+1)!} \left\| f^{(n+1)} \right\|_\infty^{[a;b]}\,.$$

- **Formule de Taylor-Young**
  Soit $f : I \longrightarrow \mathbb{K}$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N})$ sur I. Alors, pour tous a, $x \in I$:

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + o\big((x-a)^n\big)\,.$$

> ✓ Parmi ces trois formules, celle avec reste intégrale est la plus précise. C'est donc celle à privilégier, mais dans beaucoup de cas l'inégalité de Taylor-Lagrange, plus simple, peut suffire.
> Quant à la formule de Taylor-Young, elle ne doit servir QUE pour les développements limités : c'est une formule locale, alors que les deux premières sont des formules globales.

12.9 Primitives usuelles

Dans le tableau page suivante on rappelle quelques primitives de fonctions usuelles.

Les domaines de définition ne sont pas mentionnés.

Le réel a qui intervient est supposé non nul.

Fonction	Primitive	Fonction	Primitive
$(ax+b)^\alpha$ $(\alpha \neq -1)$	$\dfrac{(ax+b)^{\alpha+1}}{a(\alpha+1)}$	$\mathrm{sh}(ax+b)$	$\dfrac{1}{a}\,\mathrm{ch}(ax+b)$
$\dfrac{1}{ax+b}$	$\dfrac{1}{a}\ln\lvert ax+b\rvert$	$\mathrm{ch}(ax+b)$	$\dfrac{1}{a}\,\mathrm{sh}(ax+b)$
e^{ax+b}	$\dfrac{1}{a}\,\mathrm{e}^{ax+b}$	$\dfrac{1}{\mathrm{sh}\,x}$	$\ln\left\lvert \mathrm{th}\dfrac{x}{2}\right\rvert$
$\ln x$	$x\ln x - x$	$\dfrac{1}{\mathrm{ch}\,x}$	$2\,\mathrm{Arctan}(\mathrm{e}^x)$
$\sin(ax+b)$	$-\dfrac{1}{a}\cos(ax+b)$	$\dfrac{1}{\mathrm{sh}^2\,x}$	$\coth x$
$\cos(ax+b)$	$\dfrac{1}{a}\sin(ax+b)$	$\dfrac{1}{\mathrm{ch}^2\,x}$	$\mathrm{th}\,x$
$\dfrac{1}{\sin x}$	$\ln\left\lvert\tan\dfrac{x}{2}\right\rvert$	$\mathrm{th}\,x$	$\ln(\mathrm{ch}\,x)$
$\dfrac{1}{\cos x}$	$\ln\left\lvert\tan\left(\dfrac{x}{2}+\dfrac{\pi}{4}\right)\right\rvert$	$\dfrac{1}{x^2+a^2}$	$\dfrac{1}{a}\,\mathrm{Arctan}\left(\dfrac{x}{a}\right)$
$\dfrac{1}{\sin^2 x}$	$-\cot x$	$\dfrac{1}{x^2-a^2}$	$\dfrac{1}{2a}\ln\left\lvert\dfrac{x-a}{x+a}\right\rvert$
$\dfrac{1}{\cos^2 x}$	$\tan x$	$\dfrac{1}{\sqrt{x^2+a^2}}$	$\ln(x+\sqrt{x^2+a^2})$
$\tan x$	$-\ln\lvert\cos x\rvert$	$\dfrac{1}{\sqrt{x^2-a^2}}$	$\ln\lvert x+\sqrt{x^2-a^2}\rvert$
$\tan^2 x$	$\tan x - x$	$\dfrac{1}{\sqrt{a^2-x^2}}$	$\mathrm{Arcsin}\left(\dfrac{x}{\lvert a\rvert}\right)$

On retiendra qu'une primitive de $\dfrac{1}{x^2-a^2}$ (qui intervient fréquemment) se retrouve facilement en écrivant :

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

Chapitre 13 - Intégration sur un intervalle quelconque

On considère dans ce chapitre des applications à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

13.1 Définitions

- Soit f une fonction continue par morceaux sur un intervalle $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$.

  On dit que l'intégrale généralisée (ou impropre) $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si

  $$\lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t \quad \text{existe.}$$

  S'il en est ainsi, cette intégrale est notée :

  $$\int_a^b f(t)\,\mathrm{d}t \quad \text{ou} \quad \int_a^b f \quad \text{ou} \quad \int_{[a;b[} f\,.$$

  Dans le cas contraire on dit que l'intégrale généralisée $\int_a^b f$ diverge.

- **Intégrale faussement impropre**
  Soit f continue par morceaux sur $[a\,;b[$ avec a et b *finis*, à valeurs dans $\mathbb{K}$.

  Si f admet une limite *finie* en b^-, alors l'intégrale $\displaystyle\int_{[a;b[} f$ converge.

  De plus, si l'on note $\tilde{f}$ le prolongement de f par continuité en b, on a :

  $$\int_{[a;b[} f = \int_a^b \tilde{f}\,.$$

  ✓ Attention à ne pas utiliser ce résultat dans le cas où $b = +\infty$!
  Une fonction définie sur un intervalle de la forme $[a\,;+\infty[$ peut avoir une limite en $+\infty$ sans que son intégrale converge. Il existe même des fonctions dont l'intégrale converge et qui n'ont pas de limite en $+\infty$.
  Ainsi, contrairement aux séries, il n'y a aucun lien entre limite en $+\infty$ et existence de l'intégrale dans le cas général.

- Soit f une fonction continue par morceaux sur $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$, et soit $c \in]a\,;b[$. L'intégrale généralisée $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si et seulement si l'intégrale $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ converge et, dans ce cas :

  $$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t\,.$$

  ✓ Ainsi, pour une fonction continue par morceaux sur $[a\,;b[$, l'existence de l'intégrale est une notion locale au voisinage de b.

- Soit f une fonction continue par morceaux sur un intervalle $]a\,;b]$ avec $-\infty \leqslant a < b < +\infty$, à valeurs dans $\mathbb{K}$.

  On dit que l'intégrale généralisée $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si

  $$\lim_{x\to a^+} \int_x^b f(t)\,\mathrm{d}t \quad \text{existe.}$$

  S'il en est ainsi, cette intégrale est notée :

  $$\int_a^b f(t)\,\mathrm{d}t \quad \text{ou} \quad \int_a^b f \quad \text{ou} \quad \int_{]a;b]} f\,.$$

  Dans le cas contraire on dit que l'intégrale $\displaystyle\int_a^b f$ diverge.

- Soit f une application continue par morceaux sur un intervalle $]a\,;b[$ avec $-\infty \leqslant a < b \leqslant +\infty$ à valeurs dans $\mathbb{K}$. Les propriétés suivantes sont équivalentes :

  (i) il existe $c \in]a\,;b[$ tel que les intégrales $\displaystyle\int_a^c f(t)\,\mathrm{d}t$ et $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ convergent,

(ii) pour tout $c \in]a\,;b[$, les intégrales $\displaystyle\int_a^c f(t)\,\mathrm{d}t$ et $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ convergent.

Si tel est le cas, la somme des deux intégrales généralisées $\displaystyle\int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$ ne dépend pas de c. On la note :

$$\int_a^b f(t)\,\mathrm{d}t \quad \text{ou} \quad \int_a^b f \quad \text{ou} \quad \int_{]a;b[} f\,,$$

et on dit alors que l'intégrale généralisée $\displaystyle\int_a^b f$ converge.

On dira donc que l'intégrale $\displaystyle\int_a^b f$ diverge s'il existe $c \in]a\,;b[$ tel que *l'une au moins* des intégrales $\displaystyle\int_a^c f(t)\,\mathrm{d}t$ ou $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ diverge.

13.2 Propriétés

Les propriétés qui suivent sont énoncées dans le cas d'une fonction continue par morceaux sur un intervalle de la forme $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$. On obtient bien sûr des résultats analogues pour les deux autres cas d'intégrale généralisée.

- Soit f et g deux fonctions continues par morceaux sur $[a\,;b[$ à valeurs dans $\mathbb{K}$ et λ, μ deux scalaires.

  On suppose que les intégrales $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ et $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ convergent.

  Alors l'intégrale généralisée $\displaystyle\int_a^b (\lambda f + \mu g)(t)\,\mathrm{d}t$ converge et :

  $$\int_a^b (\lambda f + \mu g)(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t\,.$$

  En d'autres termes : le sous-ensemble de $\mathscr{CM}([a\,;b[,\mathbb{K})$ formé des fonctions dont l'intégrale converge est un sous-espace vectoriel de $\mathscr{CM}([a\,;b[,\mathbb{K})$, et sur cet espace, l'application $f \mapsto \displaystyle\int_a^b f$ est une forme linéaire.

  ✓ Il résulte immédiatement de la proposition précédente que, si f a une intégrale divergente sur $[a\,;b[$ et g une intégrale convergente, alors l'intégrale de $f+g$ sera divergente.

  On ne peut cependant rien dire a priori de la somme de deux intégrales divergentes.

- **Positivité**
  - Soit f une fonction continue par morceaux sur $[a\,;b[$, *à valeurs réelles*.

    Si $f \geqslant 0$ sur $[a\,;b[$ et si l'intégrale de f est convergente, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t \geqslant 0$.
  - Soient f et g deux fonctions continues par morceaux sur $[a\,;b[$, à valeurs réelles.

    Si $f(t) \leqslant g(t)$ pour tout $t \in [a\,;b[$ et si les intégrales de f et de g convergent, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t \leqslant \int_a^b g(t)\,\mathrm{d}t$.
  - Soit f une fonction *continue* sur $[a\,;b[$, *à valeurs réelles positives*, telle que l'intégrale de f sur $[a\,;b[$ est convergente.

    Alors : $\displaystyle\int_a^b f(t)\,\mathrm{d}t = 0 \implies f = 0$ sur $[a\,;b[$.

- **Cas d'une fonction à valeurs dans $\mathbb{C}$**

  Soit f une fonction continue par morceaux sur $[a\,;b[$, à valeurs dans $\mathbb{C}$.

  Pour que l'intégrale de f sur $[a\,;b[$ soit convergente, il faut et il suffit que les intégrales de $\mathcal{R}e\,(f)$ et de $\mathcal{I}m\,(f)$ le soient, et, dans ce cas on a

  $$\int_a^b f(t)\,\mathrm{d}t = \int_a^b \mathcal{R}e\,(f(t))\,\mathrm{d}t + \mathrm{i}\int_a^b \mathcal{I}m\,(f(t))\,\mathrm{d}t\,.$$

13.3 Règles de convergence pour les fonctions à valeurs positives

Les propriétés qui suivent sont énoncées dans le cas d'une fonction continue par morceaux sur un intervalle de la forme $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$. On obtient des résultats analogues pour une fonction continue par morceaux sur un intervalle de la forme $]a\,;b]$ avec $-\infty \leqslant a < b < +\infty$.

> ✓ Cependant, dans le cas d'une fonction définie sur un intervalle ouvert $]a\,;b[$, il faut faire deux études séparées, l'une au voisinage de a et l'autre au voisinage de b.

- Soit f continue par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles *positives*.
  Soit l'application $F\colon x \in [a\,;b[\longmapsto \displaystyle\int_a^x f(t)\,\mathrm{d}t$.
  Alors l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si et seulement si F est majorée et, dans ce cas, $\displaystyle\int_a^b f(t)\,\mathrm{d}t = \lim_{x\to b^-} F(x)$.
  Dans le cas contraire, $\displaystyle\lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t = +\infty$.
- Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles.
  On suppose qu'il existe $c \in [a\,;b[$ tel que :
  $$\forall\, t \in [c\,;b[\,,\ 0 \leqslant f(t) \leqslant g(t)\,.$$
  – Si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge.
  – Si $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ diverge, alors $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ diverge.
- Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles, *positives* au voisinage de b.
  – Si $f \underset{b^-}{=} O(g)$ et si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge.
  – Si $f \underset{b^-}{=} o(g)$ et si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge.
- Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles.
  On suppose g positive au voisinage de b. S'il existe une constante réelle $k \neq 0$ telle que $f \underset{b^-}{\sim} kg$, alors les intégrales de f et g sur $[a\,;b[$ sont de même nature.

13.4 Intégrales de référence

Pour pouvoir utiliser les critères de comparaison précédents, il faut connaître un certain nombre d'intégrales de référence.

- **Exponentielle et logarithme**
  . $\displaystyle\int_0^1 \ln t\,\mathrm{d}t$ converge (et $\displaystyle\int_0^1 \ln t\,\mathrm{d}t = -1$).
  . $\displaystyle\int_0^{+\infty} \mathrm{e}^{-at}\,\mathrm{d}t$ converge si et seulement si $a > 0$ (et dans ce cas $\displaystyle\int_0^{+\infty} \mathrm{e}^{-at}\,\mathrm{d}t = \frac{1}{a}$).
- **Fonctions de Riemann**
  Il s'agit des fonctions $t \mapsto \dfrac{1}{t^\alpha}$ pour $t > 0$ et $\alpha \in \mathbb{R}$.
  . $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ converge si et seulement si $\alpha > 1$.
  . $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t^\alpha}$ converge si et seulement si $\alpha < 1$.
  . Et plus généralement, si a et b sont des réels *finis* tels que $a < b$, alors :
  $$\int_a^b \frac{\mathrm{d}t}{(t-a)^\alpha} \quad \text{converge si et seulement si } \alpha < 1\,.$$

13.5 Changement de variable dans une intégrale généralisée

Soit f une fonction continue par morceaux sur un intervalle $]\alpha\,;\beta[$ à valeurs réelles ou complexes, et φ une bijection strictement monotone d'un intervalle $]a\,;b[$ sur $]\alpha\,;\beta[$, de classe $\mathscr{C}^1$ sur $]a\,;b[$.

Alors l'intégrale $\displaystyle\int_\alpha^\beta f$ est convergente si et seulement si l'intégrale $\displaystyle\int_a^b (f\circ\varphi)\varphi'$ l'est, et, dans ce cas :

$$\int_\alpha^\beta f(t)\,\mathrm{d}t = \int_a^b f\circ\varphi\,(u)\varphi'(u)\,\mathrm{d}u\,.$$

13.6 Intégration par parties

Soient f et g deux fonctions continues et de classe $\mathscr{C}^1$ par morceaux sur un intervalle $[a\,;b[$. Si $\lim\limits_{t\to b^-} f(t)g(t)$ existe , alors les intégrales $\displaystyle\int_a^b f'g$ et $\displaystyle\int_a^b fg'$ sont de même nature, et, lorsqu'elles convergent, on a :

$$\int_a^b f'(t)g(t)\,\mathrm{d}t = \lim_{t\to b^-} f(t)g(t) - f(a)g(a) - \int_a^b f(t)g'(t)\,\mathrm{d}t\,.$$

ce que l'on écrit encore $\displaystyle\int_a^b f'(t)g(t)\,\mathrm{d}t = \big[f(t)g(t)\big]_a^b - \int_a^b f(t)g'(t)\,\mathrm{d}t$.

Ce théorème s'adapte sans difficulté au cas d'un intervalle de la forme $]a\,;b]$ (il faut alors que $\lim\limits_{t\to a^+} f(t)g(t)$ existe) ou au cas d'un intervalle de la forme $]a\,;b[$ (et il faut alors que les *deux* limites $\lim\limits_{t\to b^-} f(t)g(t)$ et $\lim\limits_{t\to a^+} f(t)g(t)$ existent).

> ✓ Pour appliquer ce théorème il est *indispensable* de vérifier proprement l'existence de la limite. En cas de doute, on reviendra à la formule sur un segment, et à la fin des calculs on fera un (ou des) passage(s) à la limite (que l'on justifiera bien sûr !).

13.7 Intégrales absolument convergentes

Lorsqu'une fonction n'est pas à valeurs réelles positives (ou de signe constant), les théorèmes de comparaison vus plus haut ne s'appliquent pas. On peut cependant s'y ramener dans certains cas grâce à la définition et au théorème suivants.

- **Définition**

  Soit f une fonction continue par morceaux sur un intervalle $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$.

  On dit que l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente si l'intégrale $\displaystyle\int_a^b |f(t)|\,\mathrm{d}t$ est convergente.

  Lorsque l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente, on dit aussi que f est intégrable sur $[a\,;b[$.

  On a bien sûr une définition analogue dans le cas d'une fonction définie sur un intervalle de la forme $]a\,;b]$ ou $]a\,;b[$.

  Dans la définition ci-dessus, $|f(t)|$ désigne la valeur absolue de $f(t)$ si f est à valeurs réelles, et son module si f est à valeurs dans $\mathbb{C}$.

- **Théorème**

  Soit f une fonction continue par morceaux sur un intervalle $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$.

  Si l'intégrale $\displaystyle\int_a^b f$ est absolument convergente, elle est convergente, et alors :

  $$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}t\,.$$

  > ✓ La réciproque du théorème précédent est *fausse*.
  >
  > Il existe en effet des intégrales qui sont convergentes sans être absolument convergentes (une telle intégrale est dite semi-convergente). Exemple : $\displaystyle\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t$ est convergente mais pas absolument convergente.

- Le théorème précédent permet d'utiliser certains critères de comparaison même lorsqu'on ne connaît pas le signe de la fonction ; on a en particulier le résultat suivant.

  Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$.
  On suppose (seulement) g à valeurs réelles *positives* au voisinage de b.

  Si $f \underset{b^-}{=} O(g)$ ou si $f \underset{b^-}{=} o(g)$ et si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente.

13.8 Suite de fonctions intégrables

Contrairement à l'intégration sur un segment (cf. [9.2]), même la convergence uniforme d'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ vers une fonction f ne suffit pas, lorsque l'intervalle I est quelconque, à assurer que :

$$\lim_{n\to+\infty} \int_I f_n = \int_I f\,.$$

Il nous faut un résultat plus puissant ; c'est le **théorème de convergence dominée de Lebesgue**.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues par morceaux sur un intervalle I, à valeurs dans $\mathbb{K}$. On suppose que :

– la suite de fonctions f_n converge simplement sur I vers une fonction f continue par morceaux sur I ;

– il existe une fonction φ continue par morceaux sur I, à valeurs réelles positives, et intégrable sur I, telle que : $\forall\, n \in \mathbb{N}$, $|f_n| \leqslant \varphi$ sur I (*hypothèse de domination*).

Alors les f_n sont intégrables sur I, f est intégrable sur I, et

$$\int_I f \;=\; \lim_{n\to+\infty} \int_I f_n$$

(*l'existence de cette limite est assurée par le théorème*).

13.9 Série de fonctions intégrables

Deux théorèmes sont disponibles pour calculer l'intégrale de la somme d'une série de fonctions, en plus de celui déjà vu en [9.7] dans le cas où il y a convergence uniforme sur un segment.

- **Théorème d'intégration terme à terme**

  Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions continues par morceaux sur un intervalle I, à valeurs dans $\mathbb{C}$. On suppose que :

  – $\forall\, n \in \mathbb{N}$, u_n est intégrable sur I ;

  – la série de fonctions $\sum u_n$ converge simplement sur I, et sa fonction somme s est continue par morceaux sur I ;

  – la série $\sum \int_I |u_n|$ converge.

  Alors $s = \displaystyle\sum_{n=0}^{+\infty} u_n$ est intégrable sur I, et $\displaystyle\int_I s \;=\; \sum_{n=0}^{+\infty} \int_I u_n$ (*cette série étant absolument convergente*).

- **Application du théorème de convergence dominée**

  Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions continues par morceaux sur un intervalle I, à valeurs dans $\mathbb{C}$. On notera, pour $n \in \mathbb{N}$, $s_n = \displaystyle\sum_{k=0}^{n} u_k$. On suppose que :

  – la série de fonctions $\sum u_n$ converge simplement sur I, et sa fonction somme $s = \displaystyle\sum_{n=0}^{+\infty} u_n$ est continue par morceaux sur I ;

  – il existe une fonction φ continue par morceaux sur I, à valeurs réelles positives, et intégrable sur I, telle que : $\forall\, n \in \mathbb{N}$, $|s_n| \leqslant \varphi$ sur I (*hypothèse de domination*).

  Alors toutes les fonctions u_n, ainsi que s, sont intégrables sur I, et $\displaystyle\int_I s = \sum_{n=0}^{+\infty} \int_I u_n$ (*la convergence de cette série étant assurée par le théorème*).

13.10 Intégrales dépendant d'un paramètre

- **Continuité**

  Soit f une fonction définie sur $I \times J$, où I et J sont des intervalles réels, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On suppose que :

  – $\forall\, x \in I$, $t \mapsto f(x,t)$ est continue par morceaux sur J ;

  – $\forall\, t \in J$, la fonction $x \mapsto f(x,t)$ est continue sur I ;

  – il existe une fonction φ, continue par morceaux sur J, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* J telle que :

  $$\forall\,(x,t) \in I \times J,\ \big|f(x,t)\big| \leqslant \varphi(t)\ (\textit{hypothèse de domination}).$$

  Alors la fonction $g\colon x \mapsto \displaystyle\int_J f(x,t)\,\mathrm{d}t$ est définie et continue sur I.

  Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in K \times J$, pour tout segment K inclus dans I.

  ✓ La remarque précédente est très importante, et simplifie souvent beaucoup la recherche d'une fonction dominante.

- **Dérivabilité**

  Soit f une fonction définie sur $I \times J$, où I et J sont des intervalles réels, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On suppose que :

  – $\forall\, x \in I$, $t \mapsto f(x,t)$ est continue par morceaux et intégrable sur J ;

  – $\forall\, t \in J$, la fonction $x \mapsto f(x,t)$ est dérivable sur I ;

  – pour tout $x \in I$, la fonction $t \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue par morceaux sur J ;

  – pour tout $t \in J$, la fonction $x \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue sur I ;

  – il existe une fonction φ, continue par morceaux sur J, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* J telle que :

  $$\forall\,(x,t) \in I \times J,\ \left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant \varphi(t)\ (\textit{hypothèse de domination}).$$

  Alors la fonction $g\colon x \mapsto \displaystyle\int_J f(x,t)\,\mathrm{d}t$ est de classe $\mathscr{C}^1$ sur I, et :

  $$\forall\, x \in I,\ g'(x) = \int_J \frac{\partial f}{\partial x}(x,t)\,\mathrm{d}t \quad \textit{(formule de Leibniz)}.$$

  Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in K \times J$, pour tout segment K inclus dans I.

  ✓ Si l'on vous demande de prouver que la fonction g est de classe $\mathscr{C}^1$, appliquez directement ce dernier théorème sans passer par le premier !

- **Généralisation**

  Le résultat suivant se démontre aisément par récurrence à partir du précédent.

  Soit f une fonction définie sur $I \times J$, où I et J sont des intervalles réels, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $n \in \mathbb{N}^*$.

  On suppose que :

  – $f, \dfrac{\partial f}{\partial x}, \ldots, \dfrac{\partial^n f}{\partial x^n}$ existent, sont continues par rapport à x sur I et continues par morceaux par rapport à t sur J ;

  – $\forall\, x \in I$, les fonctions $t \mapsto f(x,t)$, $t \mapsto \dfrac{\partial f}{\partial x}(x,t)$, $\ldots$, $t \mapsto \dfrac{\partial^{n-1} f}{\partial x^{n-1}}(x,t)$ sont intégrables sur J ;

  – il existe une fonction φ, continue par morceaux sur J, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* J telle que :

  $$\forall\,(x,t) \in I \times J,\ \left|\frac{\partial^n f}{\partial x^n}(x,t)\right| \leqslant \varphi(t)\ (\textit{hypothèse de domination}).$$

Alors la fonction $g\colon x \mapsto \displaystyle\int_J f(x,t)\,\mathrm{d}t$ est de classe $\mathscr{C}^n$ sur I, et :

$$\forall\, k \in [\![1\,;n]\!],\ \forall\, x \in I,\ g^{(k)}(x) = \int_J \frac{\partial^k f}{\partial x^k}(x,t)\,\mathrm{d}t\,.$$

Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in K \times J$, pour tout segment K de I.

✓ Si l'on vous demande de montrer que g est de classe $\mathscr{C}^n$ voire $\mathscr{C}^\infty$, il faut utiliser directement ce dernier théorème.

Chapitre 14 - Espaces probabilisés

Avant d'aborder les probabilités, il nous a paru indispensable de faire quelques rappels de 1re année concernant les dénombrements et les coefficients binomiaux.
Nous mentionnons aussi quelques résultats sur les séries doubles, qui sont admis et qui sont indispensables pour tous les calculs de probabilités dans des espaces dénombrables.

14.1 Dénombrement

- **p-listes**
  Soient n et p deux entiers strictement positifs et E un ensemble à n éléments.
  Une p-liste d'éléments de E est un p-uplet $(x_1,\dots,x_p)$ d'éléments de E.

  ✓ Dans une p-liste, les éléments ne sont pas nécessairement distincts, mais le rang des éléments dans la p-liste est important.

  Le nombre de p-listes d'un ensemble E à n éléments est $\mathrm{Card}(E^p) = n^p$.
- **p-listes d'éléments distincts de E, ou arrangements**
  Soient n et p deux entiers tels que $0 < p \leqslant n$ et E un ensemble à n éléments.
  Une p-liste d'éléments distincts de E ou p-arrangement est une p-liste $(x_1,\dots,x_p)$ d'éléments de E telle que $x_i \neq x_j$ pour $i \neq j$.

  ✓ Dans un p-arrangement, les éléments sont tous distincts, et le rang des éléments dans la liste est important.

  Si $\mathrm{Card}\, E = n$, le nombre de p-listes d'éléments distincts de E est le nombre noté A_n^p égal à :
  $$A_n^p = n(n-1)\cdots(n-p+1) = \frac{n!}{(n-p)!}\cdot$$
- **Permutations**
  Une permutation d'un ensemble E de cardinal n est une n-liste d'éléments distincts de E.
  Le nombre de permutations d'un ensemble E de cardinal n est : $A_n^n = n!$.
- **Parties d'un ensemble**
  Soit E un ensemble de cardinal n ($n \in \mathbb{N}$) et $\mathcal{P}(E)$ l'ensemble des parties de E. Alors : $\mathrm{Card}\,\mathcal{P}(E) = 2^n$.
- **Parties de E à p éléments, combinaisons**
  Soit $p \in \mathbb{N}$. On appelle combinaison à p éléments d'un ensemble E toute partie de E contenant p éléments.

  ✓ Dans une combinaison, les éléments sont tous distincts et leur rang est sans importance.

  Si $\mathrm{Card}\, E = n$, le nombre de combinaisons à p éléments de E ($p \leqslant n$) se note $\binom{n}{p}$ et l'on a :
  $$\binom{n}{p} = \frac{1}{p!}A_n^p = \frac{n!}{p!(n-p)!}\cdot$$
  On posera $\binom{0}{0} = 1$.
  Si $n \in \mathbb{N}$ et $p \in \mathbb{Z}$ sont tels que $p < 0$ ou $p > n$, on posera $\binom{n}{p} = 0$.
- **Propriétés des coefficients binomiaux**
  1. Pour $n \in \mathbb{N}$, $\binom{n}{0} = \binom{n}{n} = 1$; $\binom{n}{1} = n$; $\binom{n}{2} = \frac{n(n-1)}{2}\cdot$
  2. Pour $n \in \mathbb{N}$ et $k \in \mathbb{Z}$, $\binom{n}{k} = \binom{n}{n-k}\cdot$
  3. Formule sans nom :
     $$\forall\, n \in \mathbb{N},\ \forall\, k \in \mathbb{Z},\ k\binom{n}{k} = n\binom{n-1}{k-1}\cdot$$
  4. Formule de Pascal :
     $$\forall\, n \in \mathbb{N},\ \forall\, k \in \mathbb{Z},\ \binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}\cdot$$

5. Généralisation de la formule de Pascal :

$$\text{pour } 0 \leqslant p \leqslant n,\ \sum_{k=p}^{n} \binom{k}{p} = \binom{n+1}{p+1}.$$

6. Formule de Vandermonde :

$$\forall\,(n,m,k) \in \mathbb{N}^3,\ \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}.$$

7. Cas particulier : $\displaystyle\sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}$.

8. Formule du binôme :

$$\forall\, n \in \mathbb{N},\ \forall\,(a,b) \in \mathbb{C}^2,\quad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k\,.$$

- **Applications de E_p dans F_n**
  Soient p et n deux entiers strictement positifs, E_p un ensemble de cardinal p dont les éléments sont notés $(e_1,\ldots,e_p)$ et F_n un ensemble de cardinal n.
  Soit $\mathcal{A}(E_p,F_n)$ l'ensemble des applications de E_p dans F_n.
  À tout élément f de $\mathcal{A}(E_p,F_n)$, on peut faire correspondre la p-liste d'éléments de F_n : $(f(e_1),f(e_2),\ldots,f(e_p))$.
  Réciproquement, à toute p-liste $(y_1,\ldots,y_p)$ d'éléments de F_n correspond une et une seule application de E_p dans F_n définie par : $f(e_i) = y_i$ pour tout i.
  On en déduit les résultats suivants.
  - Le nombre d'applications de E_p dans F_n est égal à : n^p.
  - Le nombre d'injections de E_p dans F_n $(n \geqslant p)$ est égal à : A_n^p.
  - Le nombre de bijections de E_n dans F_n $(n = p)$ est égal à : $A_n^n = n!$.

14.2 Ensembles dénombrables

- Deux ensembles E et F sont dits équipotents s'il existe une bijection de E sur F. C'est une relation d'équivalence.
  Dans le cas où E et F sont finis, ils sont équipotents si et seulement si ils ont même cardinal.
- Un ensemble est dit dénombrable s'il est équipotent à $\mathbb{N}$.
  E est dénombrable si et seulement si on peut le décrire sous la forme

$$E = \{x_n,\ n \in \mathbb{N}\}\ .$$

  Toute partie infinie d'un ensemble dénombrable est dénombrable.
- Un ensemble E est dit au plus dénombrable s'il est équipotent à une partie de $\mathbb{N}$. Cela équivaut à dire qu'il est soit fini, soit dénombrable.
  E est au plus dénombrable si et seulement si on peut le décrire sous la forme

$$E = \{x_i,\ i \in I\} \qquad \text{où} \qquad I \text{ est une partie de } \mathbb{N}\,.$$

- **Ensembles usuels**
  - $\mathbb{Z}$ est dénombrable.
  - $\mathbb{N} \times \mathbb{N}$ est dénombrable.
  - Plus généralement, si E et F sont dénombrables, leur produit $E \times F$ l'est aussi.
  - $\mathbb{Q}$ est dénombrable.
  - Une réunion finie ou dénombrable d'ensembles dénombrables est un ensemble dénombrable.
  - L'ensemble $\mathcal{P}(\mathbb{N})$ des parties de $\mathbb{N}$ n'est pas dénombrable.
  - L'ensemble $\{0,1\}^{\mathbb{N}}$ des suites à valeurs dans $\{0,1\}$ n'est pas dénombrable.
  - $\mathbb{R}$ n'est pas dénombrable.

14.3 Séries doubles

- On considère ici une suite double $(u_{p,q})_{(p,q)\in\mathbb{N}^2}$ de nombres complexes.
  Le théorème suivant est très utilisé dans les calculs de probabilités ; il s'agit du **théorème de Fubini**.

  On suppose que :

  – pour tout q fixé dans $\mathbb{N}$, la série $\sum_{p\in\mathbb{N}} u_{p,q}$ est absolument convergente ;

  – la série de terme général $S_q = \sum_{p=0}^{+\infty} |u_{p,q}|$ est convergente.

  Dans ces conditions :

  – pour tout p fixé dans $\mathbb{N}$, la série $\sum_{q\in\mathbb{N}} u_{p,q}$ est absolument convergente ;

  – la série de terme général $S'_p = \sum_{q=0}^{+\infty} |u_{p,q}|$ est convergente ;

  – on a les égalités :

$$\sum_{p=0}^{+\infty} S'_p = \sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty}|u_{p,q}|\right) = \sum_{q=0}^{+\infty} S_q = \sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty}|u_{p,q}|\right)$$
$$\text{et} \quad \sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty} u_{p,q}\right) = \sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty} u_{p,q}\right).$$

  Si ces condition sont vérifiées, la somme ci-dessus est simplement notée $\sum_{p,q\in\mathbb{N}} u_{p,q}$, et on dit que la famille $(u_{p,q})_{(p,q)\in\mathbb{N}^2}$ est sommable.

- Un cas particulier important et fréquent d'application du théorème de Fubini est le suivant.
  Soient $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites de nombres complexes.

  Pour que la famille double $(a_p b_q)_{(p,q)\in\mathbb{N}^2}$ soit sommable il faut et il suffit que les séries $\sum_{p=0}^{+\infty} a_p$ et $\sum_{q=0}^{+\infty} b_q$ soient absolument convergentes, et on a alors :

$$\sum_{p,q\in\mathbb{N}} a_p b_q = \left(\sum_{p=0}^{+\infty} a_p\right)\left(\sum_{q=0}^{+\infty} b_q\right).$$

14.4 Espace probabilisable

- Soient Ω un ensemble appelé univers et $\mathcal{A}$ une partie de $\mathcal{P}(\Omega)$, c'est-à-dire un ensemble de parties de Ω.
  On dit que $\mathcal{A}$ est une tribu ou une σ-algèbre de parties de Ω si :

  – $\Omega \in \mathcal{A}$;

  – si $A \in \mathcal{A}$ alors $\overline{A} \in \mathcal{A}$ ($\overline{A}$ désigne le complémentaire de A dans Ω) ;

  – pour toute partie I de $\mathbb{N}$, et toute famille $(A_i)_{i\in I}$ d'éléments de $\mathcal{A}$, $\bigcup_{i\in I} A_i \in \mathcal{A}$.

  Les éléments de $\mathcal{A}$ sont appelés des évènements, et le couple $(\Omega,\mathcal{A})$ s'appelle un espace probabilisable.

- **Propriétés**
  Soient Ω un ensemble et $\mathcal{A}$ une tribu de partie de Ω.

  – $\emptyset \in \mathcal{A}$.

  – Si A et B sont deux évènements de $\mathcal{A}$, alors $A \cup B$, $A \cap B$ et $A \setminus B$ sont dans $\mathcal{A}$.

  – Si I est une partie de $\mathbb{N}$ et si pour tout $i \in I$, $A_i \in \mathcal{A}$ alors $\bigcap_{i\in I} A_i \in \mathcal{A}$.

- **Un peu de vocabulaire**

Terminologie probabiliste	Terminologie ensembliste
évènement certain	Ω
évènement impossible	$\emptyset$
évènement élémentaire ω	singleton $\{\omega\}$
évènement contraire de A : $\bar{A}$	complémentaire de A dans Ω
évènement A ou B	$A \cup B$
évènement A et B	$A \cap B$
les évènements A et B sont incompatibles	$A \cap B = \emptyset$
A implique B	$A \subset B$
système complet d'évènements	partition

14.5 Espace probabilisé

- Soit $(\Omega,\mathcal{A})$ un espace probabilisable. On appelle probabilité sur $(\Omega,\mathcal{A})$ toute application P de $\mathcal{A}$ dans $[0\,;1]$ vérifiant :

  (i) $P(\Omega) = 1$;

  (ii) si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements deux à deux incompatibles alors :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) \qquad (\sigma\text{-additivité})$$

  (cette écriture sous-entend que la série converge).

  Le triplet $(\Omega,\mathcal{A},P)$ est alors appelé espace probabilisé.

- **Germe de probabilité**

  Soit Ω un ensemble au plus dénombrable, et $\{p_\omega \mid \omega \in \Omega\}$ une famille de réels positifs telle que $\sum_{\omega\in\Omega} p_\omega = 1$.

  Alors il existe une unique probabilité P sur $\big(\Omega,\mathcal{P}(\Omega)\big)$ telle que :

$$\forall\, \omega \in \Omega,\ P(\{\omega\}) = p_\omega\,.$$

  Cette probabilité est alors définie par :

$$\forall\, A \in \mathcal{P}(\Omega),\ P(A) = \sum_{\omega\in A} p_\omega\,.$$

  ✓ Ce théorème permet donc de généraliser au cas où Ω est dénombrable ce que vous avez vu en 1re année dans le cas fini : pour définir une probabilité, il suffit de la définir pour les évènements élémentaires, sous réserve bien sûr que toutes ces probabilités soient des réels positifs et que leur somme soit une série convergente de somme 1.

- **Propriétés**

  Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et soient A et B deux évènements.

  - $P(\emptyset) = 0$.
  - $P(\overline{A}) = 1 - P(A)$.
  - Si A et B sont incompatibles, $P(A \cup B) = P(A) + P(B)$.
  - Plus généralement : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
  - $P(A \setminus B) = P(A) - P(A \cap B)$.
  - Si $A \subset B$ alors $P(A) \leqslant P(B)$ (*croissance*).
  - Si $(A_i)_{i\in I}$ (I partie de $\mathbb{N}$) est un système complet d'évènements, $\sum_{i\in I} P(A_i) = 1$.

- Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé.

– Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'évènements de $\mathcal{A}$ ($A_n \subset A_{n+1}$) alors la suite $\left(P(A_n)\right)_{n\in\mathbb{N}}$ converge et :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n) \quad (\textit{continuité croissante}).$$

– Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'évènements de $\mathcal{A}$ ($A_{n+1} \subset A_n$) alors la suite $\left(P(A_n)\right)_{n\in\mathbb{N}}$ converge et :

$$P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n) \quad (\textit{continuité décroissante}).$$

- Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements de $\mathcal{A}$ alors :

$$P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \lim_{n\to+\infty} P\left(\bigcup_{k=0}^{n} A_k\right).$$

$$P\left(\bigcap_{k=0}^{+\infty} A_k\right) = \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right).$$

- Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements de $\mathcal{A}$ alors :

– Pour tout $n \in \mathbb{N}$, $P\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \sum_{n=0}^{n} P(A_n)$ (*sous-additivité finie*).

– $P\left(\bigcup_{k=0}^{+\infty} A_k\right) \leqslant \sum_{n=0}^{+\infty} P(A_n)$ (*sous-additivité*), cette dernière somme pouvant être $+\infty$.

- Un évènement $A \in \mathcal{A}$ est dit quasi-impossible si $P(A) = 0$ et quasi-certain si $P(A) = 1$.

  Un système quasi-complet d'évènements est une famille $(A_i)_{i\in I}$ ($I \subset \mathbb{N}$) d'éléments de $\mathcal{A}$, non vides, deux à deux incompatibles, telle que $P\left(\bigcup_{i\in I} A_i\right) = 1$.

14.6 Probabilité conditionnelle

- Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et B un évènement de probabilité non nulle.

  Pour tout évènement A, on pose $P_B(A) = \dfrac{P(A\cap B)}{P(B)}$.

  Alors P_B est une probabilité sur $(\Omega,\mathcal{A})$ appelée probabilité conditionnelle relative à B ou encore probabilité sachant B. $P_B(A)$ est aussi noté $P(A \mid B)$.

  Si $P(A) \neq 0$, on peut aussi écrire $P(B \mid A) = \dfrac{P(A\cap B)}{P(A)}$ et donc :

$$P(A) \times P(B \mid A) = P(B) \times P(A \mid B).$$

  Comme P_B est une probabilité, toutes les propriétés vues en [**14.5**] peuvent lui être appliquées (par exemple : $P\left(\overline{A} \mid B\right) = 1 - P(A \mid B)$).

- **Formule des probabilités composées**

  Directement d'après la définition, on a pour A, $B \in \mathcal{A}$:

$$P(A\cap B) = P(B)P(A \mid B).$$

  Plus généralement, si $(A_i)_{1\leqslant i\leqslant n}$ est une famille d'évènements tels que $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \neq 0$, on a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)\ldots P_{A_1\cap A_2\cap\cdots\cap A_{n-1}}(A_n),$$

  ce qui s'écrit aussi :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)\,P(A_3|A_1\cap A_2)\ldots P(A_n|A_1\cap A_2\cap\cdots\cap A_{n-1}).$$

- **Formule des probabilités totales**
  Soit $(A_i)_{i\in I}$ un système quasi-complet d'évènements de probabilités non nulles. Alors pour tout évènement B on a :
  $$P(B) = \sum_{i\in I} P(A_i \cap B) = \sum_{i\in I} P(B \mid A_i)\, P(A_i)$$
  (lorsque I est infini dénombrable, cela sous-entend que la série écrite converge).

  On peut interpréter cette formule en disant que B est un effet provenant de différentes causes A_i.
  Un cas particulier important de la formule des probabilités totales est celui où l'on prend comme système complet d'évènements un système de la forme $\left\{A, \overline{A}\right\}$. La formule s'écrit alors :
  $$P(B) = P(B \mid A)\, P(A) + P\left(B \mid \overline{A}\right) P(\overline{A})\,.$$

- **Formule de Bayes**
  On a vu que, si A et B sont deux évènements de probabilités non nulles, on a :
  $$P(A \mid B) = \frac{P(A) \times P(B \mid A)}{P(B)} \qquad (*)$$
  Si maintenant on considère un système complet ou quasi-complet d'évènements $(A_i)_{i\in I}$ (avec toujours I partie de $\mathbb{N}$), et si A est l'un de ces évènements, soit $A = A_j$ pour un $j \in I$, la formule des probabilités totales appliquée à l'évènement B donne la **formule de Bayes**.

  > Soient $(A_i)_{i\in I}$ un système quasi-complet d'évènements de probabilités non nulles et B un évènement de probabilité non nulle. On a :
  > $$P(A_j \mid B) = \frac{P(A_j) P(B \mid A_j)}{\sum_{i\in I} P(A_i) P(B \mid A_i)}\,.$$

  Cette formule s'appelle aussi formule de probabilités des causes.

  > ✓ Cette formule est affreuse ! Il vaut mieux retenir la formule $(*)$, beaucoup plus simple et qui est une conséquence directe de la définition d'une probabilité conditionnelle, puis obtenir la valeur de $P(B)$ par la formule des probabilités totales.

14.7 Indépendance d'évènements

- Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. On dit que deux évènements A et B sont indépendants pour la probabilité P si :
  $$P(A \cap B) = P(A) \times P(B)\,.$$
  On a donc alors $P(B \mid A) = P(B)$ et $P(A \mid B) = P(A)$.
  Si A et B sont deux évènements indépendants alors $\overline{A}$ et B, A et $\overline{B}$ et $\overline{A}$ et $\overline{B}$ le sont aussi.

  > ✓ Ne pas confondre « évènements indépendants » et « évènements incompatibles » ! D'ailleurs, il est facile de vérifier que deux évènements incompatibles sont indépendants si et seulement si l'un d'entre eux est quasi-impossible.
  > De plus, la notion d'indépendance dépend de la probabilité P alors que la notion d'incompatibilité est purement ensembliste.

- Soient $(A_i)_{i\in I}$ une famille d'évènements (avec $I \subset \mathbb{N}$, fini ou non).
  - On dit que les évènements sont deux à deux indépendants pour la probabilité P si pour tout $i \neq j$ A_i et A_j sont indépendants.
  - On dit que les évènements sont mutuellement indépendants (ou tout simplement indépendants) pour la probabilité P si pour *tout* ensemble fini d'indices $J \subset I$, $P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j)$.

  > ✓ Il est clair que si des évènements sont mutuellement indépendants, ils le sont deux à deux, mais la réciproque est fausse.

- Soient $(A_i)_{i\in I}$ une famille d'évènements indépendants (avec $I \subset \mathbb{N}$, fini ou non).
  Soit $(B_i)_{i\in I}$ une famille d'évènements telle que :
  $$\forall\, i \in I,\ B_i = A_i \quad \text{ou} \quad B_i = \overline{A_i}\,.$$
  Alors les B_i sont indépendants.

Chapitre 15 - Variables aléatoires discrètes

15.1 Variables aléatoires discrètes

- **Définitions**
  - Soit $(\Omega,\mathcal{A})$, un espace probabilisable. On appelle variable aléatoire discrète définie sur $(\Omega,\mathcal{A})$ à valeurs dans un ensemble E toute application X de Ω dans E telle que :
    - son image $X(\Omega)$ est au plus dénombrable (on écrira donc $X(\Omega) = \{x_i,\, i \in I\}$ avec $I \subset \mathbb{N}$) ;
    - pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$ appartient à $\mathcal{A}$ (c'est-à-dire est un évènement).

      Cet ensemble se note, en abrégé : $(X = x)$.

    On a alors, plus généralement, le résultat suivant :

    pour toute partie $A \subset E$, l'ensemble $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$ est un évènement de $\mathcal{A}$, que l'on note en abrégé : $(X \in A)$.
  - Si $X(\Omega)$ est un ensemble fini, on dit que X est une variable aléatoire finie.
  - Si $X(\Omega)$ est une partie de $\mathbb{R}$, on dit que X est une variable aléatoire réelle.
  - Si X est constante, on dit que c'est une variable aléatoire certaine.
- **Fonction d'une variable aléatoire**

  Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$ à valeurs dans un ensemble E et f une fonction définie sur $X(\Omega)$ à valeurs dans un ensemble F.

  Alors $f \circ X$ est une variable aléatoire discrète sur $(\Omega,\mathcal{A},P)$ (on la notera simplement $f(X)$).
- **Loi d'une variable aléatoire**

  Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé, et X une variable aléatoire discrète à valeurs dans un ensemble E.

  L'application $f : \begin{cases} X(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P(X = x) \end{cases}$ s'appelle la loi de probabilité de X.

  ✓ Pour répondre à la question « déterminer la loi de X », il faut commencer par donner clairement $X(\Omega)$. Ensuite, pour chaque élément x_i de cet ensemble $X(\Omega)$ il faut donner $P(X = x_i)$.

  Le résultat suivant est extrêmement important et permet en particulier de vérifier la cohérence des résultats obtenus.

  Soit X une variable aléatoire discrète. Si $X(\Omega) = \{x_i,\, i \in I\}$ alors la famille d'évènements $(X = x_i)_{i\in I}$ est un système complet d'évènements. En particulier on a $\displaystyle\sum_{i\in I} P(X = x_i) = 1$.

  De plus, puisque $(X = x_i)_{i\in I}$ est un système complet d'évènements, on peut appliquer la formule des probabilités totales pour n'importe quel évènement A :

$$P(A) = \sum_{i\in I} P(A \cap (X = x_i)) = \sum_{i\in I} P\left(A \mid X = x_i\right) P(X = x_i)\,.$$

- **Germe d'une variable aléatoire**

  Soit $(\Omega,\mathcal{A})$ un espace probabilisable, $\{x_i,\, i \in I\}$ un ensemble au plus dénombrable, et $(p_i)_{i\in I}$ une famille de réels positifs telle que $\displaystyle\sum_{i\in I} p_i = 1$.

  Alors il existe une probabilité P sur $(\Omega,\mathcal{A})$ et une variable aléatoire discrète X sur $(\Omega,\mathcal{A},P)$, à valeurs dans $\{x_i,\, i \in I\}$, telles que :

$$\forall\, i \in I,\ \ P(X = x_i) = p_i\,.$$

  ✓ L'intérêt de ce résultat est qu'il permet d'étudier une variable aléatoire définie par sa loi, sans avoir à préciser l'espace probabilisé sur lequel on travaille.
- **Fonction de répartition d'une variable aléatoire réelle**

  Soit X une variable aléatoire discrète réelle. On appelle fonction de répartition de X l'application $F : \mathbb{R} \to \mathbb{R}$ définie par :

$$F(x) = P(X \leqslant x)\,.$$

  Propriétés

  - F est une fonction en escalier ;

– $\forall\, x \in \mathbb{R},\ F(x) \in [0\,;1]$;

– F est croissante ;

– $\lim\limits_{x\to-\infty} F(x) = 0$ et $\lim\limits_{x\to+\infty} F(x) = 1$;

– F est continue à droite en tout point ;

– si a et b sont des réels tels que $a < b$, $P(a < X \leqslant b) = F(b) - F(a)$;

– en notant $X(\Omega) = \{x_i,\ i \in I\}$, les x_i étant rangés par ordre croissant, pour tout $i \in I$ tel que $i-1 \in I$ (on a donc $x_{i-1} < x_i$), on a :
$$P(X = x_i) = F(x_i) - F(x_{i-1})\,.$$

15.2 Lois discrètes finies usuelles

- **Loi uniforme**
On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi uniforme sur $[\![1\,;n]\!]$, et l'on note $X \hookrightarrow \mathcal{U}([\![1\,;n]\!])$ si :
$$\text{(i) } X(\Omega)) = [\![1\,;n]\!] \qquad \text{(ii) } \forall\, x \in X(\Omega),\ P(X = x) = \frac{1}{n}\cdot$$
Situation : tirage au hasard d'un entier parmi $[\![1\,;n]\!]$ (ou une boule dans une urne…), tous les tirages étant équiprobables.

- **Loi de Bernoulli**
Soit $p \in [0\,;1]$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi de Bernoulli de paramètre p, et l'on note $X \hookrightarrow \mathcal{B}(p)$ si :
$$\text{(i) } X(\Omega)) = \{0,1\} \qquad \text{(ii) } P(X = 1) = p \ \text{ et } \ P(X = 0) = 1 - p\,.$$
Les cas $p = 0$ et $p = 1$ n'étant pas intéressants, on considère en général $p \in\,]0\,;1[$.
Situation : expérience de type succès-échec (pile ou face, tirage d'une boule noire ou blanche…).

- **Loi binomiale**
Soit n un entier non nul et p un réel dans $[0\,;1]$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi binomiale de paramètres n et p, et l'on note $X \hookrightarrow \mathcal{B}(n,p)$ si :
$$\text{(i) } X(\Omega)) = [\![0\,;n]\!] \qquad \text{(ii) } \forall\, k \in [\![0\,;n]\!],\ P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}\,.$$
On note traditionnellement $q = 1 - p$ de sorte que $P(X = k) = \binom{n}{k} p^k q^{n-k}$.
Situation : expérience de type succès-échec que l'on effectue n fois dans les mêmes conditions ; X désigne le nombre de fois où l'expérience a amené un succès (lancers de pièces, tirage avec remise de boules dans une urne…).

15.3 Lois discrètes infinies usuelles

- **Loi géométrique**
Soit $p \in\,]0\,;1[$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi géométrique de paramètre p, et l'on note $X \hookrightarrow \mathcal{G}(p)$ si :
$$\text{(i) } X(\Omega)) = \mathbb{N}^* \qquad \text{(ii) } \forall\, n \in \mathbb{N}^*,\ P(X = n) = (1-p)^{n-1} p\,.$$
Il est facile de vérifier que $\sum\limits_{n=1}^{+\infty} P(X = n)$ est bien égal à 1 (série géométrique).
Situation : expérience de type succès-échec que l'on répète dans des conditions identiques ; X désigne le nombre d'épreuves effectuées jusqu'à ce que l'on obtienne un succès pour la première fois.
Une propriété importante de la loi géométrique est d'être *sans mémoire*. En effet, les épreuves étant indépendantes, la loi de probabilité du nombre d'épreuves à répéter jusqu'à l'obtention d'un premier succès est la même quel que soit le nombre d'échecs obtenus auparavant.
Mathématiquement, cela se traduit par :
$$\forall\,(n,m) \in \mathbb{N}^{*2},\ P\,(X = n + m \mid X > m) = P(X = n)\,,$$

ou, puisque $P(X > n) = \sum_{k=n+1}^{+\infty} P(X = k)$ par :

$$\forall\,(n,m) \in \mathbb{N}^{*2},\ P\,(X > n+m \mid X > m) = P(X > n)\,.$$

La loi géométrique est la seule loi de probabilité discrète qui possède cette propriété.

> ✓ Pour une loi géométrique, l'évènement $(X > n)$ lorsque $n \in \mathbb{N}^*$ représente le fait de n'avoir que des échecs lors des n premières épreuves ; on a donc $P(X > n) = (1-p)^n$ (et cela reste vrai pour $n = 0$).
>
> Il est facile de vérifier que l'on a bien : $(1-p)^n = \sum_{k=n+1}^{+\infty} P(X = k)$.

- **Loi de Poisson**

  Soit λ un réel strictement positif. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi de Poisson de paramètre λ, et l'on note $X \hookrightarrow \mathcal{P}(\lambda)$ si :

  $$\text{(i) } X(\Omega)) = \mathbb{N} \qquad \text{(ii) } \forall\, n \in \mathbb{N},\ P(X = n) = \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!}\,.$$

  Situation : la loi de Poisson sert à estimer le nombre de réalisations d'un évènement dans un intervalle de temps donné (arrivée de clients au guichet d'une banque en une heure, arrivée de malades aux urgences d'un hôpital en une nuit…). Le paramètre λ correspond au nombre moyen d'évènements survenus.

  Ce n'est bien sûr qu'une approximation, qui est justifiée par le résultat suivant.

- **Approximation d'une loi binomiale par une loi de Poisson**

  Si (X_n) est une suite de variables aléatoires telle que X_n suit une loi binomiale $\mathcal{B}(n,p_n)$ avec $\lim\limits_{n\to+\infty} np_n = \lambda > 0$, alors la suite (X_n) converge simplement vers une variable aléatoire X qui suit une loi de Poisson de paramètre λ :

  $$\forall\, k \in \mathbb{N},\ \lim_{n\to+\infty} P(X_n = k) = \lim_{n\to+\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!} = P(X = k)\,.$$

15.4 Couples de variables aléatoires discrètes

- **Loi d'un couple de variables aléatoires**

  Soit $(\Omega,\mathcal{A},P)$, un espace probabilisé et E un ensemble. Soient X et Y deux variables aléatoires discrètes sur $(\Omega,\mathcal{A})$, à valeurs dans E.

  - On appelle loi conjointe du couple de variables aléatoires discrètes (X,Y) la donnée de l'ensemble des couples $\big((x_i,y_j),p_{i,j}\big)$ où

    $$X(\Omega) = \{x_i,\ i \in I\}\,,\ Y(\Omega) = \{y_j,\ j \in J\} \quad (I,\ J \text{ parties de } \mathbb{N})$$

    $$\text{et } p_{i,j} = P\big((X,Y) = (x_i,y_j)\big) = P\left(X = x_i \cap Y = y_j\right)\,.$$

  - Les lois de X et Y s'appellent alors les lois marginales du couple (X,Y).

  - Pour tout indice $j \in J$ tel que $P(Y = y_j)$ soit non nul, on appelle loi conditionnelle de X sachant $(Y = y_j)$ l'application de $X(\Omega)$ dans $[0\,;1]$ définie par la relation :

    $$x_i \longmapsto P_{(Y=y_j)}(X = x_i) = P\,(X = x_i \mid Y = y_j) = \frac{P\left(X = x_i \cap Y = y_j\right)}{P(Y = y_j)}\,.$$

  - Pour tout indice $i \in I$ tel que $P(X = x_i)$ soit non nul, on appelle loi conditionnelle de Y sachant $(X = x_i)$ l'application de $Y(\Omega)$ dans $[0\,;1]$ définie par la relation :

    $$y_j \longmapsto P_{(X=x_i)}(Y = y_j) = P\,(Y = y_j \mid X = x_i) = \frac{P\left(X = x_i \cap Y = y_j\right)}{P(X = x_i)}\,.$$

- **Lien entre lois conjointes et lois marginales**

  Il est important de bien savoir jongler entre loi du couple, loi marginale et loi conditionnelle. Il s'agit en fait uniquement d'appliquer la définition des probabilités conditionnelles ou alors d'appliquer la formule des probabilités totales. On a donc les relations suivantes.

- Lois marginales à partir de la loi du couple :

$$P(X=x_i)=\sum_{j\in J}P\left(X=x_i\cap Y=y_j\right)$$
$$P(Y=y_j)=\sum_{i\in I}P\left(X=x_i\cap Y=y_j\right)$$

- Lois marginales à partir des lois conditionnelles :

$$P(X=x_i)=\sum_{j\in J}P(Y=y_j)P_{(Y=y_j)}(X=x_i)$$
$$P(Y=y_j)=\sum_{i\in I}P(X=x_i)P_{(X=x_i)}(Y=y_j)$$

- Loi du couple à partir des lois conditionnelles :

$$P\left(X=x_i\cap Y=y_j\right)=P(X=x_i)P_{(X=x_i)}(Y=y_j)$$
$$P\left(X=x_i\cap Y=y_j\right)=P(Y=y_j)P_{(Y=y_j)}(X=x_i)$$

- **Indépendance de variables aléatoires**

- Soit (X,Y) un couple de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$.
  On dit que les variables X et Y sont indépendantes si :

$$\forall\, i\in I,\ \forall\, j\in J,\ P\left(X=x_i\cap Y=y_j\right)=P(X=x_i)P(Y=y_j)\,.$$

  Cela équivaut à dire que la loi conjointe est le produit des lois marginales.

- X et Y sont indépendantes si et seulement si :

$$\forall\, A\subset X(\Omega),\ \forall\, B\subset Y(\Omega),\ P\left((X\in A)\cap(Y\in B)\right)=P(X\in A)\times P(Y\in B)\,.$$

- Soient $X_1,\ldots,X_n$ n variables aléatoires discrètes. On dit que les variables $X_1,\ldots,X_n$ sont mutuellement indépendantes (ou tout simplement indépendantes) lorsque pour tout $(x_1,\ldots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)$:

$$P\big((X_1=x_1)\cap\cdots\cap(X_n=x_n)\big)=P(X_1=x_1)\times\cdots\times P(X_n=x_n)\,.$$

- **Fonction de deux variables aléatoires indépendantes**

- Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, indépendantes. Soient f et g deux applications définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement.
  Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Le résultat du théorème précédent s'étend à plus de deux variables aléatoires.

- Soient $X_1,\ldots,X_n$ n variables aléatoires discrètes mutuellement indépendantes et soit $p\in[\![2\,;n-1]\!]$. Alors toute variable aléatoire fonction des variables $X_1,\ldots,X_p$ est indépendante de toute variable aléatoire fonction des variables $X_{p+1},\ldots,X_n$.

- **Loi d'une fonction de deux variables aléatoires**

Soit X et Y deux variables aléatoires discrètes et g une fonction quelconque définie sur $X(\Omega)\times Y(\Omega)$. Alors $Z=g(X,Y)$ est une variable aléatoire discrète.

- **Somme de variables aléatoires**

- Soient $(\Omega,\mathcal{A},P)$, un espace probabilisé et (X,Y) un couple de variables aléatoires discrètes sur Ω, à valeurs dans $\mathbb{R}$. Notons $X(\Omega)=\{x_i,\ i\in I\}$ et $Y(\Omega)=\{y_j,\ j\in J\}$ où I, J sont deux parties de $\mathbb{N}$.

  Soit $Z=X+Y$. On sait déjà que Z est une variable aléatoire réelle discrète. Pour tout $z\in Z(\Omega)$, l'évènement $(Z=z)$ est la réunion disjointe et au plus dénombrable des évènements $(X=x_i\cap Y=y_j)$ pour tous les couples $(i,j)\in I\times J$ tels que $x_i+y_j=z$. On aura donc :

$$P(Z=z)=\sum_{(i,j)\ \text{tq}\ x_i+y_j=z}P\left(X=x_i\cap Y=y_j\right)\,.$$

  Dans le cas où les variables X et Y sont indépendantes on aura donc aussi :

$$P(Z=z)=\sum_{(i,j)\ \text{tq}\ x_i+y_j=z}P(X=x_i)(Y=y_j)\,.$$

- Si $X_1,\dots,X_n$ sont n variables aléatoires mutuellement indépendantes suivant la loi de Bernoulli de paramètre p, la variable aléatoire $X = X_1 + X_2 + \cdots + X_n$ suit la loi binomiale $\mathcal{B}(n,p)$.
- Si $X_1,\dots,X_k$ sont k variables aléatoires mutuellement indépendantes suivant les lois binomiales de paramètres respectifs $(n_1,p),\dots,(n_k,p)$ alors la variable aléatoire $X_1 + \cdots + X_k$ suit la loi binomiale de paramètre $\left(\sum_{i=1}^{k} n_i, p\right)$.
- Si X et Y sont deux variables aléatoires indépendantes suivant les lois de Poisson de paramètres respectifs λ et μ alors $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$.

15.5 Espérance d'une variable aléatoire discrète

Les variables aléatoires considérées dans ce paragraphe sont à valeurs réelles.

- **Définition**
  Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$.
  - Si $X(\Omega)$ est fini et si $(x_i,p_i)_{1\leqslant i\leqslant n}$ est la loi de X, alors l'espérance de X est :
  $$E(X) = \sum_{1\leqslant i\leqslant n} p_i x_i = \sum_{i=1}^{n} x_i P(X = x_i)\,.$$
  - Si $X(\Omega)$ est infini et si $(x_n,p_n)_{n\in\mathbb{N}}$ est la loi de X, alors on dira que X admet une espérance si la série de terme général $p_n x_n$ est *absolument* convergente. Dans ce cas l'espérance de X est :
  $$E(X) = \sum_{n=0}^{+\infty} p_n x_n = \sum_{n=0}^{+\infty} x_n P(X = x_n)\,.$$

  Remarques
  - Une variable aléatoire discrète finie a toujours une espérance ce qui n'est pas le cas pour une variable aléatoire discrète infinie.
  - L'absolue convergence est nécessaire car l'ordre dans lequel les termes sont additionnés ne doit pas intervenir (en effet, lors d'une expérience aléatoire, les résultats peuvent apparaître dans n'importe quel ordre), et cette condition est liée à l'absolue convergence de la série.
- **Cas d'une variable aléatoire à valeurs entières**
  Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$.
  X admet une espérance si et seulement si la série de terme général $P(X \geqslant k)$ converge, et dans ce cas :
  $$E(X) = \sum_{k=1}^{+\infty} P(X \geqslant k) = \sum_{k=0}^{+\infty} P(X > k)\,.$$
- **Positivité de l'espérance**
  Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$. Si $X \geqslant 0$ (c'est-à-dire si $X(\omega) \geqslant 0$ pour tout $\omega \in \Omega$) et si son espérance existe, alors $E(X) \geqslant 0$.

  De plus, si $X \geqslant 0$ et si $E(X) = 0$, alors $P(X = 0) = 1$ (on dit que X est quasi certaine).
- **Théorème de transfert**
  Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans un ensemble E. Notons $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$.
  Soit f une application de $X(\Omega)$ dans $\mathbb{R}$. La variable aléatoire réelle $f(X)$ admet une espérance si et seulement si la série $\sum_{n\in\mathbb{N}} f(x_n)P(X = x_n)$ est absolument convergente, et dans ce cas :
  $$E(f(X)) = \sum_{n\in N} f(x_n)P(X = x_n)\,.$$

  L'intérêt de ce théorème est de pouvoir calculer l'espérance de $f(X)$ sans avoir besoin de calculer sa loi, seulement en connaissant la loi de X.

- **Linéarité de l'espérance**
  Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, et soit $\lambda \in \mathbb{R}$.
  Si les espérances de X et de Y existent, alors l'espérance de $\lambda X + Y$ existe et

$$E(\lambda X + Y) = \lambda E(X) + E(Y)\,.$$

  Soit X une variable aléatoire discrète dont l'espérance existe. Puisque $E(1) = 1$ il résulte du théorème précédent que la variable aléatoire $X - E(X)$ est d'espérance nulle. On l'appelle variable aléatoire centrée associée à X.

- **Croissance de l'espérance**
  - Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, dont l'espérance existe.
    Si $X \leqslant Y$ sur Ω, alors $E(X) \leqslant E(Y)$.
  - Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$.
    On suppose que l'espérance de Y existe, et que $|X| \leqslant Y$ sur Ω.
    Alors l'espérance de X existe.
    Cas particulier : si X est une variable aléatoire discrète bornée, son espérance existe.

- **Produit de deux variables aléatoires indépendantes**
  Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, dont l'espérance existe.
  Si X et Y sont indépendantes, la variable aléatoire $Z = XY$ possède une espérance finie et l'on a :

$$E(XY) = E(X)E(Y)\,.$$

- **Inégalité de Markov**
  Soit X une variable aléatoire discrète réelle positive, dont l'espérance existe.
  Alors, pour tout réel $a > 0$:

$$P(X \geqslant a) \leqslant \frac{E(X)}{a}\,.$$

15.6 Moments d'une variable aléatoire réelle discrète

Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et X une variable aléatoire réelle discrète et soit $r \geqslant 1$ un entier.

– Si $X(\Omega)$ est fini et si $(x_i,p_i)_{1\leqslant i\leqslant n}$ est la loi de X, alors le moment d'ordre r de X est :

$$m_r(X) = E(X^r) = \sum_{i=1}^{n} p_i x_i^r\,.$$

– Si $X(\Omega)$ est infini et si $(x_n,p_n)_{n\in\mathbb{N}}$ est la loi de X, alors on dira que X admet un moment d'ordre r si la série de terme général $p_n x_n^r$ est absolument convergente. Dans ce cas le moment d'ordre r de X est :

$$m_r(X) = E(X^r) = \sum_{n=0}^{+\infty} p_n x_n^r = \sum_{n=0}^{+\infty} x_n^r P(X = x_n)\,.$$

Théorème
Si X admet un moment d'ordre $r \geqslant 2$, elle admet des moments d'ordre s pour tout $s \in [\![1\,;r]\!]$.

15.7 Variance

- Soit X une variable aléatoire réelle dont le moment d'ordre 2, c'est-à-dire $E(X^2)$, existe.
  – La variable aléatoire $X - E(X)$ admet aussi un moment d'ordre 2, appelé variance de X. Ainsi :

$$V(X) = E\left(\big(X - E(X)\big)^2\right).$$

  – On a la relation :

$$V(X) = E(X^2) - E(X)^2 \quad (\textit{formule de Kœnig-Huygens}).$$

  On appelle alors écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$.

- Soit X une variable aléatoire réelle dont la variance $V(X)$ existe.
  Si a et b sont deux réels, la variable $aX+b$ admet aussi une variance et :
  $$V(aX+b)=a^2V(X)\,.$$
- **Inégalité de Bienaymé-Tchebychev**
  Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Alors, pour tout réel $\varepsilon>0$:
  $$P\big(|X-E(X)|\geqslant\varepsilon\big)\leqslant\frac{V(X)}{\varepsilon^2}\,.$$
  Cette inégalité montre que la variance permet de mesurer la dispersion de la variable X autour de sa moyenne.
  Conséquence
  Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Notons $m=E(X)$.
  Si $V(X)=0$ alors $P(X=m)=1$ (X est quasi certaine).

15.8 Moments des lois usuelles

- **Loi uniforme**
  Soit $X\hookrightarrow\mathcal{U}(\llbracket1\,;n\rrbracket$. Alors :
  $$E(X)=\frac{n+1}{2}\quad\text{et}\quad V(X)=\frac{n^2-1}{12}\,.$$
- **Loi binomiale**
  Soit $X\hookrightarrow\mathcal{B}(n,p)$. Alors :
  $$E(X)=np\quad\text{et}\quad V(X)=npq\,.$$
- **Loi géométrique**
  Soit $X\hookrightarrow\mathcal{G}(p)$. Alors :
  $$E(X)=\frac{1}{p}\quad\text{et}\quad V(X)=\frac{q}{p^2}\,.$$
- **Loi de Poisson**
  Soit $X\hookrightarrow\mathcal{P}(\lambda)$. Alors :
  $$E(X)=\lambda\quad\text{et}\quad V(X)=\lambda\,.$$

15.9 Covariance

- Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, ayant chacune un moment d'ordre 2.
  Alors l'espérance de XY existe et
  $$E(XY)^2\leqslant E(X^2)E(Y^2)\quad(\textit{inégalité de Cauchy-Schwarz}).$$
- Si X et Y sont deux variables aléatoires réelles admettant un moment d'ordre 2, alors la variable aléatoire $X+Y$ admet aussi un moment d'ordre 2.
- Si X et Y sont deux variables aléatoires réelles admettant chacune un moment d'ordre 2, on appelle covariance de X et Y le réel :
  $$\operatorname{cov}(X,Y)=E\big[(X-E(X))(Y-E(Y))\big]$$
  (l'existence est assurée par les propriétés précédentes).
  Un calcul simple montre que l'on a aussi : $\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$.
  On a alors la relation :
  $$V(X+Y)=V(X)+V(Y)+2\operatorname{cov}(X,Y)\,.$$
- **Propriétés**
  Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé.
  - L'ensemble $\mathcal{V}(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω est un sous-espace vectoriel de $\mathcal{A}(\Omega,\mathbb{R})$.
  - L'ensemble $\mathcal{V}^1(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω admettant une espérance est un sous-espace vectoriel de $\mathcal{V}(\Omega,\mathbb{R})$.

– L'ensemble $\mathcal{V}^2(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω admettant un moment d'ordre 2 est un sous-espace vectoriel de $\mathcal{V}^1(\Omega,\mathbb{R})$.

– L'application $(X,Y) \mapsto \text{cov}\,(X,Y)$ est une forme bilinéaire symétrique positive (mais non définie) sur l'espace vectoriel $\mathcal{V}^2(\Omega,\mathbb{R})$, et pour tout $X \in \mathcal{V}^2(\Omega,\mathbb{R})$, $\text{cov}\,(X,X) = V(X)$.

– Si X et Y appartiennent à $\mathcal{V}^2(\Omega,\mathbb{R})$, on a l'inégalité de Cauchy-Schwarz :

$$|\text{cov}\,(X,Y)| \leqslant \sigma(X)\sigma(Y)\,.$$

- Si X et Y sont deux variables aléatoires réelles admettant chacune un moment d'ordre 2 et **indépendantes**, alors leur covariance est nulle.
  Il en résulte que l'on a alors $V(X+Y) = V(X) + V(Y)$.

  ✓ La réciproque de cette propriété est fausse : deux variables aléatoires peuvent avoir une covariance nulle sans être indépendantes.

- **Généralisation**
  Si $X_1,\ldots,X_n$ sont n variables aléatoires réelles possédant toutes un moment d'ordre 2, en utilisant la bilinéarité de la covariance on obtient :

$$V(X_1+\cdots+X_n) = V(X_1)+\cdots+V(X_n) + 2\sum_{1\leqslant i<j\leqslant n} \text{cov}\,(X_i,X_j)\,.$$

  Il en résulte que si $X_1,\ldots,X_n$ sont n variables aléatoires réelles possédant toutes un moment d'ordre 2 et *indépendantes deux à deux* :

$$V(X_1+\cdots+X_n) = \sum_{i=1}^{n} V(X_i)\,.$$

- Soient X et Y deux variables aléatoires réelles dont la variance existe et est non nulle. On appelle coefficient de corrélation linéaire de X et Y le nombre réel :

$$\rho(X,Y) = \frac{\text{cov}\,(X,Y)}{\sigma(X)\sigma(Y)}\,.$$

  Propriétés
  Soient X et Y deux variables aléatoires réelles dont la variance existe et est non nulle.

  – $|\rho(X,Y)| \leqslant 1$;

  – $|\rho(X,Y)| = 1$ si et seulement si Y est une fonction quasi affine de X, c'est-à-dire s'il existe a et b réels tels que l'évènement $(Y = aX+b)$ soit quasi certain.

15.10 Variables aléatoires à valeurs dans $\mathbb{N}$

- Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$.
  La fonction génératrice de X est la série entière :

$$G_X(t) = \sum_{n=0}^{+\infty} P(X=n)t^n\,.$$

- La fonction génératrice d'une variable aléatoire à valeurs entières a un rayon de convergence R_X supérieur ou égal à 1.
- G_X est de classe $\mathscr{C}^\infty$ sur $]-R_X\,;R_X[$ et :

$$\forall\, n\in\mathbb{N},\ P(X=n) = \frac{G_X^{(n)}(0)}{n!}\,.$$

- Pour tout $t \in]-R_X\,;R_X[$, $G_X(t) = E(t^X)$.
- **Fonctions génératrices des lois usuelles**

  . Si X suit la loi binomiale de paramètre (n,p), sa fonction génératrice est la fonction polynomiale définie par :

$$\forall\, t\in\mathbb{R},\ G_X(t) = (pt+q)^n.$$

  . Si X suit la loi géométrique de paramètre p, le rayon de convergence de sa série génératrice est $\frac{1}{q}$ et :

$$\forall\, t \in \left]-\tfrac{1}{q}\,;\tfrac{1}{q}\right[,\ G_X(t) = \frac{pt}{1-qt}\,.$$

- Si X suit la loi de Poisson de paramètre λ, sa série génératrice est de rayon de convergence $+\infty$ et :

$$\forall\, t \in \mathbb{R},\ G_X(t) = \mathrm{e}^{\lambda(t-1)}.$$

- **Fonction génératrice de la somme de deux variables aléatoires indépendantes**

  Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$, G_X et G_Y leurs fonctions génératrices respectives, et R_X, R_Y leurs rayons de convergence.

  Si X et Y sont indépendantes on a :

$$\text{pour tout } t \text{ tel que } |t| < \min(R_X,R_Y),\ G_{X+Y}(t) = G_X(t)G_Y(t)\,.$$

- **Fonction génératrice et moments**

  - Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice. On suppose que $R_X > 1$, ce qui implique que G_X est de classe $\mathscr{C}^\infty$ au voisinage de 1. Alors :

$$\forall\, k \in \mathbb{N}^*,\ G_X^{(k)}(1) = E\big(X(X-1)\dots(X-k+1)\big)\,.$$

    En particulier :

$$E(X) = G'_X(1)\,,\ E(X(X-1)) = G''_X(1) \text{ d'où } E(X^2) = G'_X(1) + G''_X(1)\,.$$

  - Réciproquement, soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice.

    L'espérance de X existe si et seulement si G_X est dérivable à gauche en 1 et la variance de X existe si et seulement si G_X est deux fois dérivable à gauche en 1.

    Dans ce cas on aura : $E(X) = G'_X(1^-)$ et $E(X^2) = G'_X(1^-) + G''_X(1^-)$.

15.11 Loi faible des grands nombres

Soit $(X_i)_{i\in N}$ une suite de variables aléatoires réelles définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$ indépendantes, de même loi, possédant une espérance $m = E(X)$ et une variance $\sigma^2 = V(X)$. On pose :

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n}\,.$$

Alors on a :

$$\forall\, \varepsilon > 0,\ P\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}\,.$$

Par conséquent :

$$\forall\, \varepsilon > 0,\ \lim_{n\to+\infty} P\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) = 0\,.$$

Chapitre 16 - Équations différentielles linéaires

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, I est un intervalle de $\mathbb{R}$ et E est un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

16.1 Équations différentielles linéaires, systèmes différentiels

- Une équation différentielle (vectorielle) linéaire du premier ordre est une équation de la forme :

$$x' = ax + b \qquad (L)$$

où a est une application continue de I dans $\mathscr{L}(E)$ et b une application continue de I dans E.
Une solution de cette équation est une application x dérivable de I dans E telle que :

$$\forall\, t \in I,\ x'(t) = a(t)\big(x(t)\big) + b(t)\,.$$

- Soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de E, et :
  - $X(t) = \begin{pmatrix} x_1(t) & \dots & x_n(t) \end{pmatrix}^\top$ la matrice colonne formée des coordonnées du vecteur $x(t)$ dans $\mathscr{B}$, où les x_i sont des applications dérivables de I dans $\mathbb{K}$;
  - $B(t) = \begin{pmatrix} b_1(t) & \dots & b_n(t) \end{pmatrix}^\top$ la matrice colonne formée des coordonnées du vecteur $b(t)$ dans $\mathscr{B}$, où les b_i sont des applications continues de I dans $\mathbb{K}$;
  - $A(t) = (a_{ij}(t))_{1 \leqslant i,j \leqslant n}$ la matrice dans $\mathscr{B}$ de l'application linéaire $a(t)$, où les a_{ij} sont des applications continues de I dans $\mathbb{K}$.

  L'équation $x' = ax + b$ peut alors s'écrire matriciellement sous la forme :

$$X' = AX + B \quad \text{ou :} \quad \forall\, t \in I\ ,\ X'(t) = A(t)X(t) + B(t) \qquad (L)$$

  On obtient alors ce que l'on appelle un système différentiel :

$$\left\{ \begin{array}{rcl} x_1'(t) & = & a_{11}(t)x_1(t) + \cdots + a_{1n}(t)x_n(t) + b_1(t) \\ x_2'(t) & = & a_{21}(t)x_1(t) + \cdots + a_{2n}(t)x_n(t) + b_2(t) \\ \vdots & & \vdots \\ x_n'(t) & = & a_{n1}(t)x_1(t) + \cdots + a_{nn}(t)x_n(t) + b_n(t) \end{array} \right.$$

16.2 Théorème de Cauchy-Lipschitz linéaire

- Résoudre le problème de Cauchy pour une équation différentielle $x' = a\,x + b$ sur I consiste à étudier l'existence et l'unicité d'une solution vérifiant une condition initiale du type $x(t_0) = x_0$ avec $t_0 \in I$ et $x_0 \in E$ donnés.
- **Théorème de Cauchy-Lipschitz linéaire**
  Soit E un espace vectoriel de dimension finie, I un intervalle de $\mathbb{R}$, a une application continue de I dans $\mathscr{L}(E)$ et b une application continue de I dans E.
  Pour tout $(t_0, x_0) \in I \times E$, il existe une et une seule solution sur I de l'équation différentielle $x' = a\,x + b$ vérifiant la condition initiale $x(t_0) = x_0$.

  ✓ La continuité des applications a et b est indispensable.

  Par exemple, l'équation différentielle $x' = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$ n'a pas de solution sur $\mathbb{R}$ (en effet, une telle solution ne pourrait être dérivable en 0).
- Matriciellement, ce théorème s'énonce ainsi.
  Soient $A\colon I \to \mathcal{M}_n(\mathbb{K})$ et $B\colon I \to \mathcal{M}_{n,1}(\mathbb{K})$ des applications continues.
  Soit $t_0 \in I$, et soit $X_0 \in \mathcal{M}_{n,1}(\mathbb{K})$.
  Alors il existe une unique application $X\colon I \to \mathcal{M}_{n,1}(\mathbb{K})$ telle que :

$$\forall\, t \in I,\ X'(t) = A(t)X(t) + B(t) \quad \text{ et } \quad X(t_0) = X_0\,.$$

16.3 Structure de l'ensemble des solutions

- L'équation (H) : $x' = a\,x$ s'appelle l'équation homogène associée à l'équation (L) : $x' = a\,x + b$.
  L'ensemble $\mathscr{S}_H$ des solutions de l'équation différentielle homogène (H) est un sous-espace vectoriel de l'espace vectoriel $\mathscr{C}^1(I,E)$ des fonctions de classe $\mathscr{C}^1$ de I dans E.

  De plus, cet espace vectoriel est de dimension finie, et $\dim \mathscr{S}_H = \dim E$.

  ✓ Ce résultat n'est valable que sur un *intervalle*.

  Exemple : l'espace des solutions sur $\mathbb{R}^*$ de l'équation $x' = 0$ est de dimension 2, puisque les solutions sont les applications telles que $x(t) = C_1$ sur $]-\infty\,;0[$ et $x(t) = C_2$ sur $]0\,;+\infty[$.
- Si x_1 est une solution particulière de l'équation (L) : $x' = a\,x + b$ (il en existe, d'après le théorème de Cauchy-Lipschitz), alors l'ensemble des solutions de l'équation (L) est exactement l'ensemble des fonctions somme de x_1 et d'une solution quelconque de l'équation homogène (H) :

$$\mathscr{S}_L = x_1 + \mathscr{S}_H\,.$$

- **Principe de superposition des solutions**
  Soient $b_1 : I \to E$ et $b_2 : I \to E$ deux applications continues, et λ_1, λ_2 dans $\mathbb{K}$.
  Soient x_1 et x_2 deux solutions respectives des équations différentielles

$$(L_1) : \quad x' = ax + b_1 \qquad \text{et} \qquad (L_2) : \quad x' = ax + b_2\,.$$

  Alors $\lambda_1 x_1 + \lambda_2 x_2$ est solution de l'équation différentielle :

$$(L) : \quad x' = ax + \lambda_1 b_1 + \lambda_2 b_2\,.$$

16.4 Systèmes différentiels linéaires homogènes à coefficients constants

On s'intéresse ici à des systèmes différentiels de la forme : $(H) \quad X'(t) = AX(t)$, où $A \in \mathcal{M}_n(\mathbb{K})$ ne dépend pas de t.

- On suppose que A admet une valeur propre λ.
  Si V est un vecteur propre associé à cette valeur propre, alors la fonction $t \mapsto \mathrm{e}^{\lambda t}V$ est solution sur $\mathbb{R}$ de l'équation homogène $X' = AX$.
- On suppose ici que A est diagonalisable. Il existe donc une base $(V_1,\dots,V_n)$ de $\mathbb{K}^n$ formée de vecteurs propres de A, pour les valeurs propres $(\lambda_1,\dots,\lambda_n)$.
  Alors les fonctions $X_i : t \mapsto \mathrm{e}^{\lambda_i t}V_i$, pour $i \in [\![1\,;n]\!]$, forment une base de l'espace vectoriel $\mathscr{S}_H$ des solutions de l'équation homogène $X' = AX$.
  Toute solution de l'équation homogène est donc de la forme :

$$t \longmapsto \sum_{i=1}^{n} \alpha_i \mathrm{e}^{\lambda_i t}V_i \quad \text{avec } (\alpha_1,\dots,\alpha_n) \in \mathbb{K}^n\,.$$

16.5 Équations différentielles linéaires scalaires du premier ordre

Il s'agit d'équations différentielles de la forme :

$$\alpha(t)x' + \beta(t)x = \gamma(t)\,,$$

où les fonctions α, β et γ sont continues sur un certain intervalle I, à valeurs dans $\mathbb{K}$. Les solutions x sont à chercher parmi les applications de I dans $\mathbb{K}$.

- La *première* chose à faire est d'examiner si la fonction α peut ou non s'annuler sur I. Si oui, on cherche les sous-intervalles de I où α ne s'annule pas, et on résoudra l'équation sur *chacun* de ces intervalles (et il y aura des constantes d'intégration différentes sur chaque intervalle).
- Un tel intervalle J étant choisi, on peut alors diviser l'équation par $\alpha(t)$ et on se ramène ainsi à une équation différentielle sous forme résolue :

$$x' = a(t)x + b(t) \quad (L)$$

  On peut remarquer que les solutions d'une telle équation sont nécessairement de classe $\mathscr{C}^1$ sur J.

- **Intégration de l'équation homogène** (H) : $x' = a(t)x$
  Les solutions de l'équation différentielle $x' = a(t)x$ sont les fonctions de la forme :
  $$t \longmapsto C\exp\left(\int_{t_0}^{t} a(u)\,\mathrm{d}u\right),$$
  où C est une constante dans $\mathbb{K}$ et où $t \mapsto \displaystyle\int_{t_0}^{t} a(u)\,\mathrm{d}u$ est une primitive de la fonction a.

  ✓ On retrouve ainsi le théorème de Cauchy-Lipschitz : l'ensemble des solutions de (H) sur J est ici un $\mathbb{K}$-espace vectoriel de dimension 1.

- **Intégration de l'équation avec second membre** (L) : $x' = a(t)x + b(t)$
  L'ensemble des solutions de l'équation complète (L) est l'ensemble des fonctions de la forme $x = x_L + y$, où y est solution de l'équation homogène (H) et où x_L est une solution particulière de (L).
  Si l'on connaît une solution particulière x_1 de (L), la résolution est terminée.
  Sinon, on peut rechercher une telle solution par la *méthode de variation de la constante*, expliquée ci-dessous.
  Soit y une solution non nulle de (H), de la forme $y(t) = \exp\left(\displaystyle\int_{t_0}^{t} a(u)\,\mathrm{d}u\right)$. Une telle solution ne peut s'annuler, et on peut donc chercher les solutions de (L) sous la forme $x(t) = C(t)y(t)$ avec C de classe $\mathscr{C}^1$.
  Un calcul simple montre alors que (L) est équivalente à l'équation :
  $$C'(t)y(t) = b(t) \quad \text{soit} \quad C'(t) = \frac{b(t)}{y(t)}.$$
  Cela permet d'obtenir $C(t)$ par une simple primitivation, puis d'en déduire la solution $x(t)$.

  ✓ Il est bon de retenir les formules ci-dessus, en particulier de se rappeler que dans l'équation transformée apparaît seulement $C'(t)$ et non plus $C(t)$. Il est en fait inutile de refaire le calcul de la dérivée de $t \mapsto C(t)y(t)$ lorsqu'on connaît ces formules.

- **Cas d'une équation à coefficients constants**
  Il s'agit ici d'une équation de la forme $x' - ax = b(t)$, où $a \in \mathbb{K}$ est une constante. Dans ce cas, l'ensemble des solutions de l'équation homogène $x' - ax = 0$ est l'ensemble des fonctions de la forme $t \mapsto C\mathrm{e}^{at}$.
  La méthode de variation de la constante s'applique bien sûr pour la résolution de l'équation complète, mais il y a un cas particulier à retenir.
  Si le second membre est de la forme $b(t) = \mathrm{e}^{mt}P(t)$ où $m \in \mathbb{C}$ et où P est un polynôme, on peut rechercher (avec des coefficients à déterminer) une solution particulière de l'équation $x' - ax = \mathrm{e}^{mt}P(t)$ sous la forme :
  - $\mathrm{e}^{mt}Q(t)$ avec Q polynôme de même degré que P si $m \neq a$,
  - $\mathrm{e}^{mt}tQ(t)$ avec Q polynôme de même degré que P si $m = a$.

- Pour résoudre complètement l'équation initiale $\alpha(t)x' + \beta(t)x = \gamma(t)$ sur tout l'intervalle I, il faut ensuite (éventuellement) raccorder les solutions trouvées précédemment sur les intervalles où la fonction α ne s'annule pas.

16.6 Équations différentielles linéaires scalaires du second ordre

Il s'agit d'équations différentielles de la forme :
$$\alpha(t)x'' + \beta(t)x' + \gamma(t)x = \delta(t),$$
où les fonctions α, β, γ et δ sont continues sur un certain intervalle I, à valeurs dans $\mathbb{K}$.

- Comme pour les équations du premier ordre, il faut résoudre une telle équation sur un intervalle $J \subset I$ où la fonction α ne s'annule pas.
  Un tel intervalle J étant choisi, on peut alors diviser l'équation par $\alpha(t)$ et on se ramène ainsi à une équation différentielle sous forme résolue :
  $$x'' = a(t)x' + b(t)x + c(t) \quad (L)$$
  où a, b et c sont des applications continues de J dans $\mathbb{K}$.
  On peut remarquer que les solutions x d'une telle équation sont nécessairement de classe $\mathscr{C}^2$ sur J.

- L'ensemble $\mathscr{S}_H$ des solutions de l'équation homogène est un espace vectoriel de dimension 2 ; ainsi, si l'on connaît deux solutions x_1 et x_2 de (H), linéairement indépendantes (c'est-à-dire non proportionnelles), on aura :
  $$\mathscr{S}_H = \left\{\lambda_1 x_1 + \lambda_2 x_2 \ ,\ (\lambda_1,\lambda_2) \in \mathbb{K}^2\right\}.$$

- L'ensemble $\mathscr{S}_L$ des solutions de (L) sur J est un espace affine de dimension 2 : $\mathscr{S}_L = x_L + \mathscr{S}_H$, où x_L est une solution particulière de (L).
- Le théorème de Cauchy-Lipschitz linéaire s'énonce ici : pour tout $t_0 \in J$ et tout $(x_0,x_0') \in \mathbb{K}^2$, il existe une solution et une seule solution de (L) sur J vérifiant les conditions initiales

$$x(t_0) = x_0 \quad \text{et} \quad x'(t_0) = x_0' \,.$$

- **Cas où l'on connaît une solution de l'équation homogène**
  Dans ce cas, on peut, comme pour les équations du premier ordre, appliquer la méthode de variation de la constante (encore appelée *méthode de Lagrange*). En voici le principe.
  Considérons l'équation différentielle (pas forcément écrite sous forme résolue) :

$$\alpha(t)x'' + \beta(t)x' + \gamma(t)x = \delta(t) \quad (L)\,.$$

  Supposons connue une solution x_1 de l'équation *homogène*.
  Sur un intervalle où la fonction x_1 ne s'annule pas, on peut chercher une solution quelconque de l'équation (L) sous la forme : $x(t) = y(t)x_1(t)$ (en effet, puisque x_1 ne s'annule pas, la fonction $y = \dfrac{x}{x_1}$ est bien définie et deux fois dérivable).
  Quelques calculs montrent alors que x est solution de (L) si et seulement si :

$$\alpha(t)x_1(t)y''(t) + (2\alpha(t)x_1'(t) + \beta(t)x_1(t))\,y'(t) = \delta(t)\,.$$

  C'est une équation scalaire du premier ordre pour la fonction y'. Il ne reste « plus qu'à » la résoudre pour trouver y', d'où l'on déduit y puis x.

16.7 Équations différentielles du 2e ordre à coefficients constants

Soit (L) une équation différentielle de la forme :

$$x'' + ax' + bx = c(t) \quad \text{avec } a,\, b \in \mathbb{K} \text{ et } c \text{ continue sur } I \text{ à valeurs dans } \mathbb{K}$$

(où $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$).

L'équation homogène associée est $(H) : x'' + ax' + bx = 0$.

- **Résolution de l'équation homogène**
  On cherche des solutions de cette équation de la forme $t \mapsto \mathrm{e}^{rt}$. Un calcul immédiat montre que cette fonction est solution de (H) si et seulement si $r^2 + ar + b = 0$.
  À l'équation (H), on associe donc son *équation caractéristique :*

$$(E_c) : r^2 + ar + b = 0,$$

  et on étudie ses solutions. Il faut distinguer deux cas, selon que le corps de base est $\mathbb{R}$ ou $\mathbb{C}$; on notera Δ le discriminant de l'équation caractéristique.
  - Si $\mathbb{K} = \mathbb{C}$
    - Si $\Delta \neq 0$, (E_c) possède deux racines distinctes r_1 et r_2 ; dans ce cas, les fonctions $x_1 : t \mapsto \mathrm{e}^{r_1 t}$ et $x_2 : t \mapsto \mathrm{e}^{r_2 t}$ sont deux solutions linéairement indépendantes de (H), et $\mathscr{S}_H$ est l'ensemble des applications de la forme :

$$t \mapsto \lambda_1 \mathrm{e}^{r_1 t} + \lambda_2 \mathrm{e}^{r_2 t} \quad , \quad (\lambda_1,\lambda_2) \in \mathbb{C}^2\,.$$

    - Si $\Delta = 0$, (E_c) possède une racine double r ; dans ce cas, les fonctions $x_1 : t \mapsto \mathrm{e}^{rt}$ et $x_2 : t \mapsto t\mathrm{e}^{rt}$ sont deux solutions linéairement indépendantes de (H), et $\mathscr{S}_H$ est l'ensemble des applications de la forme :

$$t \mapsto (\lambda_1 t + \lambda_2)\mathrm{e}^{rt} \quad , \quad (\lambda_1,\lambda_2) \in \mathbb{C}^2\,.$$

  - Si $\mathbb{K} = \mathbb{R}$
    - Si $\Delta \geqslant 0$, on a les mêmes résultats que ci-dessus, mais avec $(\lambda_1,\lambda_2) \in \mathbb{R}^2$.
    - Si $\Delta < 0$, l'équation (E_c) possède deux racines complexes non réelles conjuguées $\alpha \pm \mathrm{i}\beta$; alors les fonctions $x_1 : t \mapsto \mathrm{e}^{\alpha t}\cos\beta t$ et $x_2 : t \mapsto \mathrm{e}^{\alpha t}\sin\beta t$ sont deux solutions linéairement indépendantes de (H), et $\mathscr{S}_H$ est l'ensemble des applications de la forme :

$$t \mapsto (\lambda_1 \cos\beta t + \lambda_2 \sin\beta t)\mathrm{e}^{\alpha t} \quad , \quad (\lambda_1,\lambda_2) \in \mathbb{R}^2\,.$$

- **Résolution de l'équation avec second membre**

  Il y a bien sûr toujours la méthode de variation de la constante, puisque l'on vient de trouver deux solutions de l'équation homogène. Il y a cependant un cas particulier à connaître.

  Si le second membre est de la forme $\mathrm{e}^{mt}P(t)$, où P est un polynôme et $m \in \mathbb{C}$, on cherchera une solution particulière de (L) sous la forme :

$$x(t) = t^k \mathrm{e}^{mt} Q(t) \text{ où } \begin{cases} k = 0 & \text{si } m \text{ n'est pas racine de } (E_c) \\ k = 1 & \text{si } m \text{ est racine simple de } (E_c) \\ k = 2 & \text{si } m \text{ est racine double de } (E_c) \end{cases}$$

  avec Q polynôme de même degré que P.

Chapitre 17 - Calcul différentiel

Dans tout le chapitre, n, p et q désignent des entiers naturels non nuls.
On étudie ici des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}^n$; ces espaces vectoriels étant de dimensions finies, toutes les normes y sont équivalentes. On notera donc $\|\ \|$ une norme quelconque sur $\mathbb{R}^p$ ou $\mathbb{R}^n$.
Les applications considérées seront toujours définies sur un *ouvert* de $\mathbb{R}^p$; cela assure que, si f est définie en $a \in U$, alors elle reste définie dans tout un voisinage de a.

17.1 Rappels

- **Applications coordonnées**
  Si f est une application de $U \subset \mathbb{R}^p$ dans $\mathbb{R}^n$, on a :
  $$\forall\, x \in U,\ f(x) = \big(f_1(x), \dots, f_n(x)\big)$$
  où les f_i sont des applications de U dans $\mathbb{R}$: ce sont les applications coordonnées de f.
  L'utilisation de ces applications coordonnées permet de ramener l'étude des fonction de $\mathbb{R}^p$ dans $\mathbb{R}^n$ à celle des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$.
- **Applications partielles**
  Soit $f\colon U \subset \mathbb{R}^p \to \mathbb{R}^n$, et $a = (a_1, \dots, a_p) \in U$. Pour tout $j \in [\![1\,;p]\!]$, la j-ème application partielle de f en a est l'application :
  $$f_{a,j} : x \mapsto f(a_1, \dots, a_{j-1}, x, a_{j+1}, \dots, a_p)\,.$$
  Elle est définie sur un ouvert de $\mathbb{R}$ contenant a_j et à valeurs dans $\mathbb{R}^n$.
- **Continuité**
  Une application $f\colon U \subset \mathbb{R}^p \to \mathbb{R}^n$ d'un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$ est dite continue en $a \in U$ si :
  $$\forall\, \varepsilon > 0,\ \exists\, \alpha > 0 \text{ tel que } \forall\, x \in U,\ \|x - a\| < \alpha \implies \|f(x) - f(a)\| < \varepsilon$$
  ou, plus simplement :
  $$\lim_{x \to a} f(x) = f(a)\,.$$
- Une combinaison linéaire, une composée, un produit ou un quotient (pour des applications à valeurs dans $\mathbb{R}$) d'applications continues est une application continue là où elle est définie.
- Une application $f\colon U \subset \mathbb{R}^p \to \mathbb{R}^n$ est continue si et seulement si ses applications coordonnées le sont.
- Si $f\colon U \subset \mathbb{R}^p \to \mathbb{R}^n$ est continue en $a \in U$, alors toutes ses applications partielles $f_{a,j}$ en ce point sont continues.

  ✓ La réciproque de cette propriété est fausse : la continuité des seules applications partielles ne suffit pas à assurer la continuité en un point.

17.2 Dérivées partielles

- **Dérivée selon un vecteur**
  Soit f une application définie sur un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$, et a un point de U. Soit $\overrightarrow{v}$ un vecteur non nul de $\mathbb{R}^p$, et t un réel tel que $a + t\overrightarrow{v} \in U$.
  On dit que f admet une dérivée selon le vecteur $\overrightarrow{v}$ au point a si la fonction $t \mapsto f(a + t\overrightarrow{v})$ est dérivable en 0.
  Le vecteur dérivé correspondant est appelé dérivée de f suivant le vecteur $\overrightarrow{v}$, et est noté $\dfrac{\partial f}{\partial \overrightarrow{v}}(a)$ ou $D_{\overrightarrow{v}} f(a)$.
  Ainsi :
  $$\frac{\partial f}{\partial \overrightarrow{v}}(a) = D_{\overrightarrow{v}} f(a) = \lim_{t \to 0} \frac{f(a + t\overrightarrow{v}) - f(a)}{t}\,.$$
- **Dérivées partielles**
  Soit f une application définie sur un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$, et a un point de U.
  Soit $(\overrightarrow{e_1}, \dots, \overrightarrow{e_p})$ la base canonique de $\mathbb{R}^p$.
  On appelle dérivées partielles (premières) de f ses dérivées suivant les vecteurs $\overrightarrow{e_1}, \dots, \overrightarrow{e_p}$ (si elles existent).

La j-ème dérivée partielle en a est notée $\partial_j f(a)$ ou $\dfrac{\partial f}{\partial x_j}(a)$. Ainsi :

$$\forall\, j \in [\![1\,;p]\!],\ \frac{\partial f}{\partial x_j}(a) = D_{\overrightarrow{e_j}} f(a) = \lim_{t\to 0} \frac{f(a+t\overrightarrow{e_j}) - f(a)}{t}\,.$$

C'est donc aussi la dérivée en a_j de la j-ème application partielle $f_{a,j}$:

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t\to 0} \frac{f(a_1,\dots,a_{j-1},a_j+t,a_{j+1},\dots,a_p) - f(a_1,\dots,a_p)}{t}\,.$$

17.3 Fonctions de classe $\mathscr{C}^1$

- Soient U un ouvert de $\mathbb{R}^p$ et f une application de U dans $\mathbb{R}^n$.
  On dit que f est de classe $\mathscr{C}^1$ sur U si f admet des dérivées partielles par rapport à chacune des variables en tout point de U, et si ces dérivées partielles sont continues sur U.
- **Théorème**
  Si f est de classe $\mathscr{C}^1$ sur U, alors f admet en tout point $a \in U$ le développement limité d'ordre 1 :

  $$f(a+h) = f(a) + \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a) + o(\|h\|)\,.$$

  Conséquence : si f est de classe $\mathscr{C}^1$ sur U, elle y est continue.
- **Différentielle**
  - Soit f une application de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^p$, à valeurs dans $\mathbb{R}^n$. Pour tout $a \in U$, on appelle différentielle de f en a l'application linéaire notée $\mathrm{d}f_a$, de $\mathbb{R}^p$ dans $\mathbb{R}^n$, définie par :

    $$\forall\, h = (h_1,\dots,h_p) \in \mathbb{R}^p,\ \mathrm{d}f_a(h) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a)\,.$$

    Avec cette notation, le développement limité de f en a s'écrit :

    $$f(a+h) = f(a) + \mathrm{d}f_a(h) + o(\|h\|)\,.$$

  - Si f est de classe $\mathscr{C}^1$ sur U, alors pour tout $a \in U$ et pour tout vecteur $\overrightarrow{v}$ de E, f admet une dérivée en a selon $\overrightarrow{v}$ et :

    $$D_{\overrightarrow{v}} f(a) = \mathrm{d}f_a(\overrightarrow{v})\,.$$

    ✓ La réciproque est fausse : une fonction peut admettre une dérivée en un point selon n'importe quel vecteur sans être même continue en ce point.
- **Opérations sur les fonctions de classe $\mathscr{C}^1$**
  - Si f et g sont de classe $\mathscr{C}^1$ de U dans $\mathbb{R}^n$, alors pour tout λ réel, $\lambda f + g$ est de classe $\mathscr{C}^1$ sur U.
  - Si f et g sont de classe $\mathscr{C}^1$ de U dans $\mathbb{R}$, leur produit fg est de classe $\mathscr{C}^1$ sur U.
  - Si f et g sont de classe $\mathscr{C}^1$ de U dans $\mathbb{R}$, leur quotient $\dfrac{f}{g}$ est de classe $\mathscr{C}^1$ sur $U \setminus \{x \in U \mid g(x) = 0\}$ (qui est bien un ouvert).
  - Si f et g sont de classe $\mathscr{C}^1$ de U dans $\mathbb{R}^n$, si B est une application bilinéaire de $(\mathbb{R}^n)^2$ dans $\mathbb{R}^n$, alors $B(f,g)$ est de classe $\mathscr{C}^1$ sur U.
- **Gradient**
  Soit f une application de classe $\mathscr{C}^1$ d'un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$. Soit $a \in U$.
  On appelle gradient de f en a , noté $\overrightarrow{\mathrm{grad}} f(a)$ ou $\nabla f(a)$ le vecteur de $\mathbb{R}^p$ dont les coordonnées dans la base canonique sont :

  $$\left(\frac{\partial f}{\partial x_1}(a),\dots,\frac{\partial f}{\partial x_p}(a)\right).$$

  Lorsqu'on munit $\mathbb{R}^p$ de son produit scalaire canonique, le gradient est donc caractérisé par la relation :

  $$\forall\, h \in \mathbb{R}^p,\ D_h f(a) = \mathrm{d}f_a(h) = \langle \nabla f(a) \,|\, h\rangle\,.$$

  Compte tenu de l'inégalité de Cauchy-Schwarz, cette relation montre que le vecteur gradient (lorsqu'il est non nul) donne la direction selon laquelle la dérivée de f est maximale.

  ✓ $\mathrm{d}f_a$ étant ici une forme linéaire, on peut faire le lien entre le gradient de f et le théorème de représentation des formes linéaires dans un espace euclidien vu en [4.9].

- **Extrema**
  Soit f une application de classe $\mathscr{C}^1$ d'un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$. Soit $a \in U$.
  . On dit que f admet un minimum (respectivement un maximum) global en a si :
  $$\forall\, x \in U,\ f(x) \geqslant f(a) \quad (\text{respectivement } f(x) \leqslant f(a))\,.$$
  . On dit que f admet un minimum (respectivement un maximum) local en a si :
  $$\exists\, V \in \mathcal{V}(a) \ \text{ tel que } \ \forall\, x \in V,\ f(x) \geqslant f(a) \quad (\text{respectivement } f(x) \leqslant f(a))\,.$$
  . Si f admet un extremum local en a, alors $\nabla(f)(a) = 0$ (c'est-à-dire que pour tout $j \in [\![1\,;p]\!]$, $\dfrac{\partial f}{\partial x_j}(a) = 0$).

  ✓ La réciproque de ce théorème est fausse : le gradient de f peut s'annuler en a (a s'appelle alors un point critique de f) sans qu'il s'agisse d'un extremum.

17.4 Composée de fonctions de classe $\mathscr{C}^1$

- **Différentielle d'une composée**
  Soient U un ouvert de $\mathbb{R}^p$ et V un ouvert de $\mathbb{R}^q$. Soit f de classe $\mathscr{C}^1$ de U dans $\mathbb{R}^q$ et g de classe $\mathscr{C}^1$ de V dans $\mathbb{R}^n$, telles que $f(U) \subset V$.
  Alors $g \circ f$ est de classe $\mathscr{C}^1$ de U dans $\mathbb{R}^n$ et l'on a :
  $$\forall\, a \in U,\ \mathrm{d}(g \circ f)_a = \mathrm{d}g_{f(a)} \circ \mathrm{d}f_a$$
  (on rappelle que $\mathrm{d}f_a \in \mathscr{L}(\mathbb{R}^p,\mathbb{R}^q)$, $\mathrm{d}g_{f(a)} \in \mathscr{L}(\mathbb{R}^q,\mathbb{R}^n)$ et $\mathrm{d}(g \circ f)_a \in \mathscr{L}(\mathbb{R}^p,\mathbb{R}^n)$).
  En notant $h = g \circ f$, $(h_1,\dots,h_n)$ ses applications coordonnées, $(g_1,\dots,g_n)$ celles de g et $(f_1,\dots,f_q)$ celles de f, puis en notant $(x_1,\dots,x_p)$ les coordonnées d'un élément de $\mathbb{R}^p$ et $(y_1,\dots,y_q)$ celles d'un élément de $\mathbb{R}^q$ (cela correspond à $y = f(x)$), la formule précédente se traduit par :
  $$\forall\, i \in [\![1\,;n]\!],\ \forall\, j \in [\![1\,;p]\!],\ \frac{\partial h_i}{\partial x_j}(a) = \sum_{k=1}^{q} \frac{\partial g_i}{\partial y_k}(f(a))\frac{\partial f_k}{\partial x_j}(a)\,.$$
- **Règle de la chaîne**
  Soit f une application de classe $\mathscr{C}^1$ d'un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$.
  Soient $(\varphi_1,\dots,\varphi_p)$ p fonctions de classe $\mathscr{C}^1$ définies sur un intervalle I de $\mathbb{R}$ telles que :
  $$\forall\, t \in I,\ \big(\varphi_1(t),\dots,\varphi_p(t)\big) \in U\,.$$
  Alors la fonction $g\colon t \mapsto f\big(\varphi_1(t),\dots,\varphi_p(t)\big)$ est de classe $\mathscr{C}^1$ sur I et l'on a :
  $$\forall\, t \in I,\ g'(t) = \sum_{j=1}^{p} \varphi_j'(t)\frac{\partial f}{\partial x_j}\big(\varphi_1(t),\dots,\varphi_p(t)\big)\,.$$
- **Corollaire : caractérisation des fonctions constantes**
  Une application f de classe $\mathscr{C}^1$ sur un ouvert *convexe* U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$ est constante sur U si et seulement si sa différentielle $\mathrm{d}f$ est nulle en tout point de U.
  Cela équivaut à dire que toutes les dérivées partielles de f sont nulles sur U.

17.5 Dérivées partielles d'ordre deux

- Soit f une fonction définie sur un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$, et $(i,j) \in [\![1\,;p]\!]^2$
  On suppose que f possède une j^{e} dérivée partielle $\partial_j f$ ou $\dfrac{\partial f}{\partial x_j}$. Cette dernière peut elle-même posséder une i^{e} dérivée partielle qui sera alors notée :
  $$\partial^2_{i,j} f = \partial_i(\partial_j f) \quad \text{ou} \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right),$$
  et appelée *dérivée partielle seconde* de f par rapport aux variables x_i,x_j.
- On dit que $f\colon U \to \mathbb{R}^n$ est de classe $\mathscr{C}^2$ sur U si toutes ses dérivées partielles secondes existent et sont continues sur U.
- **Théorème de Schwarz**
  Si $f \in \mathscr{C}^2(U,\mathbb{R}^n)$ où U est un ouvert de $\mathbb{R}^p$, on a :
  $$\forall\,(i,j) \in [\![1\,;p]\!]^2,\ \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}\,.$$

17.6 Application aux courbes planes

- Une courbe du plan définie par une équation cartésienne est un ensemble de points (x,y) du plan vérifiant une équation de la forme :
$$f(x,y) = 0$$
où f est une application de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.
- Un point $M(x_0,y_0)$ de la courbe $\mathcal{C}$ d'équation $f(x,y) = 0$ est dit régulier si ce n'est pas un point critique de f, c'est-à-dire si :
$$\nabla f(x_0,y_0) \neq 0\,.$$
- Si M_0 est un point régulier de $\mathcal{C}$, la courbe admet en ce point une tangente ; c'est la droite passant par M_0 et de vecteur normal $\nabla f(x_0,y_0)$, c'est-à-dire d'équation :
$$\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) = 0\,.$$

17.7 Application aux surfaces

- Une surface définie par une équation cartésienne est un ensemble de points (x,y,z) de l'espace vérifiant une équation de la forme :
$$f(x,y,z) = 0$$
où f est une application de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}$.
- Un point $M(x_0,y_0,z_0)$ de la surface $\mathcal{S}$ d'équation $f(x,y,z) = 0$ est dit régulier si ce n'est pas un point critique de f, c'est-à-dire si :
$$\nabla f(x_0,y_0,z_0) \neq 0\,.$$
- Si M_0 est un point régulier de $\mathcal{S}$, la surface admet en ce point un plan tangent ; c'est le plan passant par M_0 et de vecteur normal $\nabla f(x_0,y_0,z_0)$, c'est-à-dire d'équation :
$$\frac{\partial f}{\partial x}(x_0,y_0,z_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0,z_0)(y-y_0) + \frac{\partial f}{\partial z}(x_0,y_0,z_0)(z-z_0) = 0\,.$$
La droite passant par M_0 et de vecteur directeur $\nabla f(x_0,y_0,z_0)$ s'appelle la normale à $\mathcal{S}$ en M_0.
- On appelle courbe tracée sur la surface $\mathcal{S}$ d'équation $f(x,y,z) = 0$ tout arc paramétré (I,γ) où I est un intervalle de $\mathbb{R}$ et où $\gamma\colon I \to \mathbb{R}^3$ est une application de classe $\mathscr{C}^1$ à valeurs dans U telle que :
$$\forall\, t \in I,\ f(\gamma(t)) = 0$$
c'est-à-dire que pour tout t le point $M(t)$ de l'arc γ appartient à $\mathcal{S}$.
- En dérivant la relation précédente à l'aide de la règle de la chaîne on obtient :
$$\forall\, t \in I,\ \langle \nabla f(\gamma(t)) \,|\, \gamma'(t) \rangle = 0$$
donc en tout point régulier (de l'arc et de la surface), la tangente à l'arc est orthogonale au gradient et par suite appartient au plan tangent à la surface.