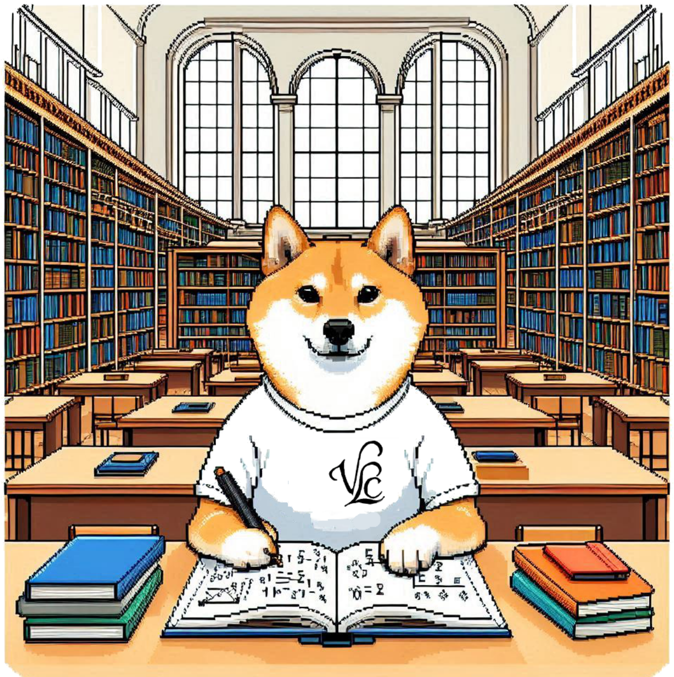
Math PT. 2025 - 2026

Mathématiques & Informatique - F. Gaunard
http://frederic.gaunard.com
Lycée Voltaire, Paris 11e.

Le cours de mathématiques en PT



Ce document synthétise les polycopiés de cours des 20 chapitres présentés lors de l'année scolaire 2025-2026.
Il présente sûrement encore quelques coquilles.
Je tiens à remercier les collègues dont les documents en ligne d'excellente qualité ont fortement influencé ce qui va suivre : François Lemeux, Gilles Rebelles et Jerôme von Buhren.
Je remercie également Philippe Hesse pour le partage de l'intégralité de ses ressources, qui ont elle aussi grandement contribué à l'écriture de ce cours.

Contents

0 Remise en route (Révisions)

Afin de relancer la machine, on propose un premier « chapitre » de rappels de formules qu'il est indispensable de maîtriser. On aura donc lu l'intégralité de ce qui suit **avant** la rentrée.
Il serait naïf de croire qu'il s'agit d'un résumé de ce qu'il faudrait avoir retenu du cours de première année, mais les notions ci-dessous seront utilisées tout au long des autres chapitres.

0.1 Trigonométrie

☞ Le formulaire qui suit doit être connu sur le bout des doigts, ou sinon on doit être capable d'en retrouver très rapidement tous les éléments sur un coin de brouillon.

La plupart des formules qui suivent sont valables pour tous réels x et y, exceptées celles mettant en jeu des tangentes, où il faudra veiller à ce que tous les termes soient définis. Soient $x, y \in \mathbb{R}$.

$$x^2 + y^2 = 1 \iff \exists \theta \in \mathbb{R} \left\{ \begin{array}{lcl} x & = & \cos(\theta) \\ y & = & \sin(\theta) \end{array} \right.$$

$$\forall \theta \in \mathbb{R}, \quad \cos^2(\theta) + \sin^2(\theta) = 1,$$

ce qui donne aussi immédiatement, pour tout $\theta \not\equiv \frac{\pi}{2} [\pi]$, $1 + \tan^2(\theta) = \dfrac{1}{\cos^2(\theta)}$ et

$$\forall \theta \in \mathbb{R}, \quad |\cos(\theta)| \leqslant 1 \quad \text{et} \quad |\sin(\theta)| \leqslant 1.$$

Proposition. 0.1. **Valeurs particulières**

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos(\theta)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\sin(\theta)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\tan(\theta)$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	

Proposition. 0.2. **Propriétés de symétrie du cercle trigonométrique**

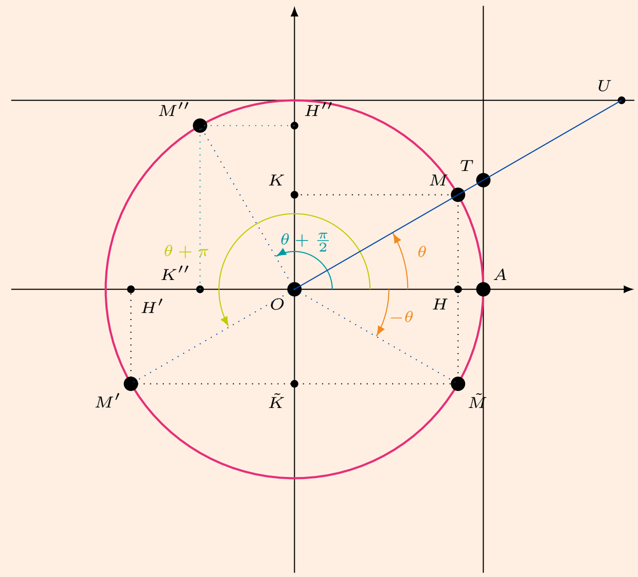
Soit $\theta \in \mathbb{R}$.

$$\cos(\theta+2\pi)=\cos(\theta) \qquad \sin(\theta+2\pi)=\sin(\theta)$$

$$\cos(\theta+\pi)=-\cos(\theta) \qquad \sin(\theta+\pi)=-\sin(\theta)$$

$$\cos\left(\frac{\pi}{2}+\theta\right)=-\sin(\theta) \qquad \sin\left(\frac{\pi}{2}+\theta\right)=\cos(\theta)$$
$$\cos\left(\frac{\pi}{2}-\theta\right)=\sin(\theta) \qquad \sin\left(\frac{\pi}{2}-\theta\right)=\cos(\theta)$$

$$\cos(-\theta)=\cos(\theta) \qquad \sin(-\theta)=-\sin(\theta)$$
$$\tan(-\theta)=-\tan(\theta) \qquad \tan(\theta+\pi)=\tan(\theta)$$



☞ Sur la figure,

$$\cos(\theta)=\overline{OH},\ \sin(\theta)=\overline{OK},\ \tan(\theta)=\overline{AT}$$

$$\tilde{M}=\mathcal{S}_{(OA)}(M),\ M'=\rho_{0,\pi}(M),\ M''=\rho_{0,\frac{\pi}{2}}(M),$$
$$\tilde{K}=\mathcal{S}_{(OA)}(K),\ H'=\rho_{0,\pi}(H)=\mathcal{S}_{(OK)}(H),\ H''=\rho_{0,\frac{\pi}{2}}(H),\ K''=\rho_{0,\frac{\pi}{2}}(K).$$

Proposition. 0.3. **Formules d'addition**

Soit $(x,y)\in\mathbb{R}^2$. Pour tout couple (x,y) tel que les quantités qui suivent sont bien définies (celles faisant intervenir une tangente), on a :

$$\cos(x+y)=\cos(x)\cos(y)-\sin(x)\sin(y) \qquad \cos(x-y)=\cos(x)\cos(y)+\sin(x)\sin(y)$$
$$\sin(x+y)=\sin(x)\cos(y)+\cos(x)\sin(y) \qquad \sin(x-y)=\sin(x)\cos(y)-\cos(x)\sin(y)$$
$$\tan(x+y)=\frac{\tan(x)+\tan(y)}{1-\tan(x)\tan(y)} \qquad \tan(x-y)=\frac{\tan(x)-\tan(y)}{1+\tan(x)\tan(y)}$$

☞ Ces formules se retrouvent facilement en identifiant parties réelles et parties imaginaires dans la relation $e^{i(x+y)}=e^{ix}e^{iy}$.

Corollaire 0.4.

Soient α et β deux réels non nuls. Si l'on pose $r=\sqrt{\alpha^2+\beta^2}$, alors il existe $\varphi\in\mathbb{R}$ tel que
$$\forall\theta\in\mathbb{R},\qquad \alpha\cos(\theta)+\beta\sin(\theta)=r\cos(\theta-\varphi).$$

Exercice 0.1. **Extrait du cahier de vacances**

Résoudre l'équation $1+\sin(x)-\cos(x)=0$.

Proposition. 0.5. **Formules de duplication**

Soit $\theta\in\mathbb{R}$.

$$\cos(2\theta)=\cos^2(\theta)-\sin^2(\theta) \qquad \sin(2\theta)=2\sin(\theta)\cos(\theta)$$
$$\cos(2\theta)=2\cos^2(\theta)-1$$
$$\cos(2\theta)=1-2\sin^2(\theta) \qquad \tan(2\theta)=\frac{2\tan(\theta)}{1-\tan^2(\theta)} \quad \left(\theta\not\equiv\frac{\pi}{2}\ [\pi]\ \text{ et }\ \theta\not\equiv\frac{\pi}{4}\ \left[\frac{\pi}{2}\right]\right)$$

Proposition. 0.6. **Formules de linéarisation**

Soit $(x,y) \in \mathbb{R}^2$.

$$\cos^2(x) = \frac{1+\cos(2x)}{2} \qquad \cos(x)\cos(y) = \frac{1}{2}\left(\cos(x+y)+\cos(x-y)\right)$$
$$\sin^2(x) = \frac{1-\cos(2x)}{2} \qquad \sin(x)\sin(y) = \frac{1}{2}\left(\cos(x-y)-\cos(x+y)\right)$$
$$\tan^2(x) = \frac{1-\cos(2x)}{1+\cos(2x)} \qquad \sin(x)\cos(y) = \frac{1}{2}\left(\sin(x+y)+\sin(x-y)\right)$$

Proposition. 0.7. **Formules de factorisation**

Soit $(a,b) \in \mathbb{R}^2$.

$$\cos(a)+\cos(b) = 2\cos\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right) \qquad \cos(a)-\cos(b) = -2\sin\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)$$
$$\sin(a)+\sin(b) = 2\sin\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right) \qquad \sin(a)-\sin(b) = 2\cos\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)$$
$$\tan(a)+\tan(b) = \frac{\sin(a+b)}{\cos(a)\cos(b)} \qquad \tan(a)-\tan(b) = \frac{\sin(a-b)}{\cos(a)\cos(b)}$$

Exercice 0.2. **Extrait du cahier de vacances**

Résoudre sur $\mathbb{R}$, puis sur l'intervalle $[0; 2\pi]$, l'équation $\cos(x)+\cos(3x) = 0$.

Proposition. 0.8. **Dérivées des fonctions trigonométriques**

Les fonctions cosinus, sinus et tangentes sont continues et dérivables sur leur ensemble de définition. Pour tout réel x, on a :

$$\cos'(x) = \cos\left(x+\frac{\pi}{2}\right) = -\sin(x) \qquad \sin'(x) = \sin\left(x+\frac{\pi}{2}\right) = \cos(x)$$
$$\tan'(x) = \frac{1}{\cos^2(x)} = 1+\tan^2(x) \qquad \left(x \not\equiv \frac{\pi}{2}\ [\pi]\right)$$

☞ On parlera à nouveau des fonctions trigonométriques (et des fonctions trigo inverses) dans le **Chapitre 2**.

0.2 Sommes, Produits

Définitions

Définition 0.1. **Symboles Σ et Π**

Soit $(a_k)_{k\in\mathbb{N}}$ une suite de nombres complexes. Pour tout $n \in \mathbb{N}$, la quantité $\sum_{k=0}^{n} a_k$ est définie récursivement par:

$$\sum_{k=0}^{0} a_k = a_0 \quad \text{et } \forall n \in \mathbb{N},\ \sum_{k=0}^{n+1} a_k = \left(\sum_{k=0}^{n} a_k\right) + a_{n+1}.$$

De manière analogue, la quantité $\prod_{k=0}^{n} a_k$ est définie récursivement par:

$$\prod_{k=0}^{0} a_k = a_0 \quad \text{et } \forall n \in \mathbb{N},\ \prod_{k=0}^{n+1} a_k = \left(\prod_{k=0}^{n} a_k\right) \times a_{n+1}.$$

Remarque 0.1.
Soit $p \in \mathbb{N}$, $I = \{i_0, \ldots, i_p\}$ un ensemble fini d'entiers, et $(a_k)_{k\in\mathbb{N}}$ une suite de nombres complexes. On définit alors la quantité $\sum_{i\in I} a_i$ comme étant la somme $\sum_{k=0}^{p} a_{i_k} = a_{i_0} + \cdots + a_{i_p}$.
Cela permet par exemple de définir $\sum_{k\in\{2p;\ p\in[\![0,m]\!]\}} a_k$, qui est la somme des $m+1$ premiers termes d'indice pair de la suite (a_n), c'est-à-dire $a_0 + a_2 + \cdots + a_{2m}$. En pratique, on préférera la notation $\sum_{k=0}^{m} a_{2k}$.
On a la même chose pour les produits.

☞ Par convention, on pose pour toute suite (a_k):

✗ Si $n < 0$, $\sum_{k=0}^{n} a_k = \sum_{i\in\emptyset} a_i = 0$.
✗ Si $n < 0$, $\prod_{k=0}^{n} a_k = \prod_{i\in\emptyset} a_i = 1$.

Définition 0.2. **Factorielle d'un entier**
Pour $n \in \mathbb{N}$, on appelle **factorielle** de n, et on note $n!$ la quantité

$$n! = \prod_{k=1}^{n} k.$$

En particulier, $0! = 1$.

0.2.1 Quelques formules

Proposition. 0.9. **Somme des (puissances des) premiers entiers**
Pour tout entier n, on a

$$(i)\ \sum_{k=1}^{n} k = \sum_{k=0}^{n} k = \frac{n(n+1)}{2}, \quad (ii)\ \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \quad (iii)\ \sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

Exercice 0.3. **Extrait du cahier de vacances**
Calculer $\sum_{k=1}^{n} k(2k^2-1)$. On vérifiera le résultat obtenu par récurrence

Proposition. 0.10. **Sommes télescopiques**
Soient n un entier et (a_k) une suite de complexes. On a $\sum_{k=0}^{n} (a_{k+1} - a_k) = a_{n+1} - a_0$.

Exercice 0.4.
Calculer, pour tout $n \in \mathbb{N}^*$, $\sum_{k=1}^{n} \ln\left(1 + \frac{1}{k}\right)$.

Proposition. 0.11. **Somme des premiers termes d'une suite géométrique**
Soit x un nombre complexe. Pour tout entier n, on a $\sum_{k=0}^{n} x^k = \begin{cases} \dfrac{1-x^{n+1}}{1-x} & \text{si } x \neq 1 \\ n+1 & \text{sinon} \end{cases}$.

Proposition. 0.12.
Soient $n \in \mathbb{N}^*$ et $z, a, b \in \mathbb{C}$. On a alors les résultats suivants.

$$(i)\ z^n - 1 = (z-1)\sum_{k=0}^{n-1} z^k, \qquad (ii)\ \ a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}.$$

Exercice 0.5.
Démontrer les deux formules de la Proposition précédente.

0.2.2 Coefficients binomiaux

Définition 0.3. **Coefficient binomial**
Soient $p \leqslant n$ des entiers. On appelle **coefficient binomial** p parmi n, et on note $\binom{n}{p}$ la quantité définie par

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}.$$

Proposition. 0.13. **Propriétés algébriques des coefficients binomiaux**
Soient $p \leqslant n$ des entiers.

i. Les coefficients binomiaux vérifient une propriété de **symétrie** : $\binom{n}{p} = \binom{n}{n-p}$.

ii. Ils vérifient la **formule du triangle de Pascal** : $\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}$.

Exercice 0.6. **Extrait du cahier de vacances**
Montrer, par la méthode de votre choix, que, pour tout $p \in \mathbb{N}^*$ et tout $n \geq p+1$, $\displaystyle\sum_{k=p}^{n-1} \binom{k-1}{p-1} = \binom{n-1}{p}$.

Proposition. 0.14. **Formule du binôme de Newton**
Soient $a, b \in \mathbb{C}$, et $n \in \mathbb{N}$. On a alors:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i.$$

☞ Comme on le rappellera au cours du **Chapitre 4**, on peut appliquer cette formule à deux matrices (resp. applications linéaires) **qui commutent**.

Exercice 0.7.
Soit $n \in \mathbb{N}$ et $\theta \in \mathbb{R}$. Calculer $\displaystyle\sum_{k=0}^{n} \binom{n}{k} \cos(k\theta)$, puis $\displaystyle\sum_{k=0}^{n} \cos(k\theta)$ dans le cas où $\theta \not\equiv 0\ [2\pi]$.

0.2.3 Sommation par paquets

Proposition. 0.15. **Extrait du cahier de vacances**
Soit $E \subset \mathbb{C}$ et I des ensembles finis, ainsi qu'une partition $(E_i)_{i \in I}$ de E.

Alors, $\displaystyle\sum_{x \in E} x = \sum_{i \in I} \left(\sum_{x \in E_i} x \right)$.

Par exemple, si $I = [\![1, n]\!]$, on peut écrire $\displaystyle\sum_{x \in E} x = \sum_{x \in E_1} x + \cdots + \sum_{x \in E_n} x$.

Exercice 0.8.

Soit $n \in \mathbb{N}^*$. Calculer $(1+i)^{4n}$, puis en déduire les valeurs de

$$\sum_{k=0}^{2n} (-1)^k \binom{4n}{2k} \quad \text{et de} \quad \sum_{k=0}^{2n-1} (-1)^k \binom{4n}{2k+1}.$$

0.2.4 Sommes doubles

On considère maintenant une suite de complexes $(a_{i,j})$ indexée par **deux indices** i et j. On cherche donc à écrire la somme de ses termes lorsque les deux indices parcourent un certain ensemble.
Ainsi, la *somme double*

$$\sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} a_{i,j}$$

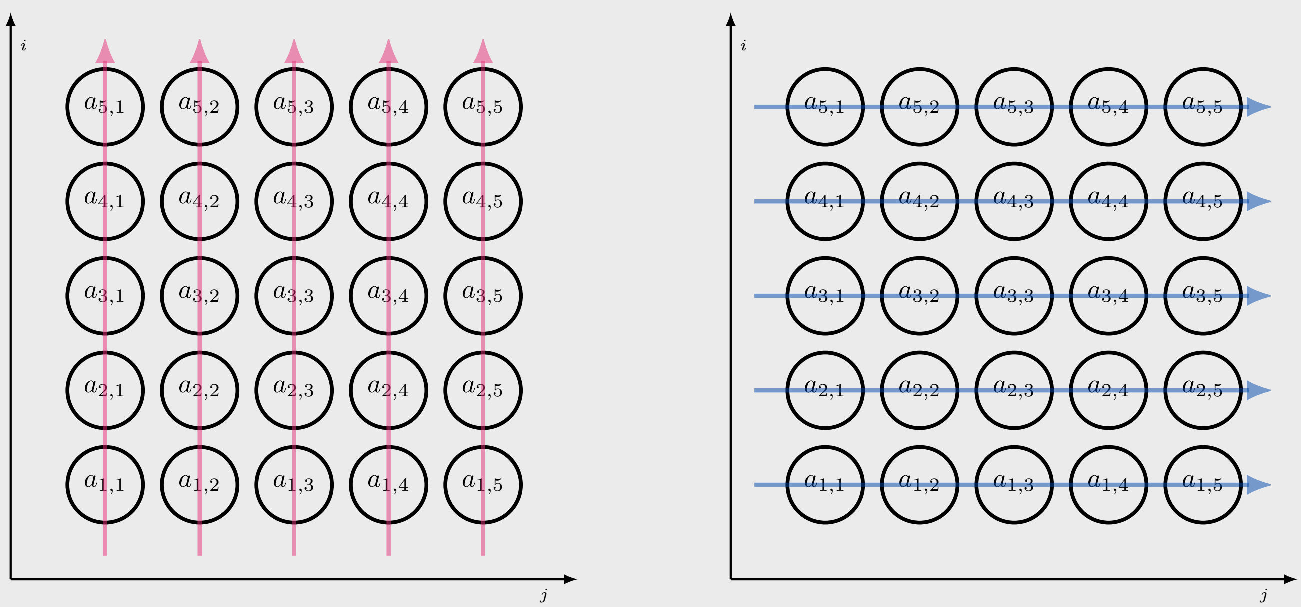
représente la somme de tous les termes $a_{i,j}$ dont le premier indice est compris entre 1 et n et dont le deuxième indice est compris entre 1 et m. Lorsque i et j parcourent le même ensemble $\{1, 2, \cdots, n\}$, la somme précédente s'écrira $\displaystyle\sum_{1 \le i,j \le n} a_{i,j}$.

Méthode 0.1. **Permutation des indices de sommation**

Lorsque les indices ne dépendent pas l'un de l'autre, on peut permuter l'ordre de sommation, ce qui revient à lister les termes sommés dans un autre ordre.

$$\sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} a_{i,j} = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{i,j} \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{i,j} \right).$$

Par exemple, pour $1 \le i, j \le 5$, on peut illustrer cette propriété avec le schéma suivant : $\displaystyle\sum_{j=1}^{5} \sum_{i=1}^{5} a_{i,j} = \sum_{i=1}^{5} \sum_{j=1}^{5} a_{i,j}$.



Exercice 0.9.

Soient $x \in \mathbb{R} \setminus \{1\}$ et $n \in \mathbb{N}^*$.

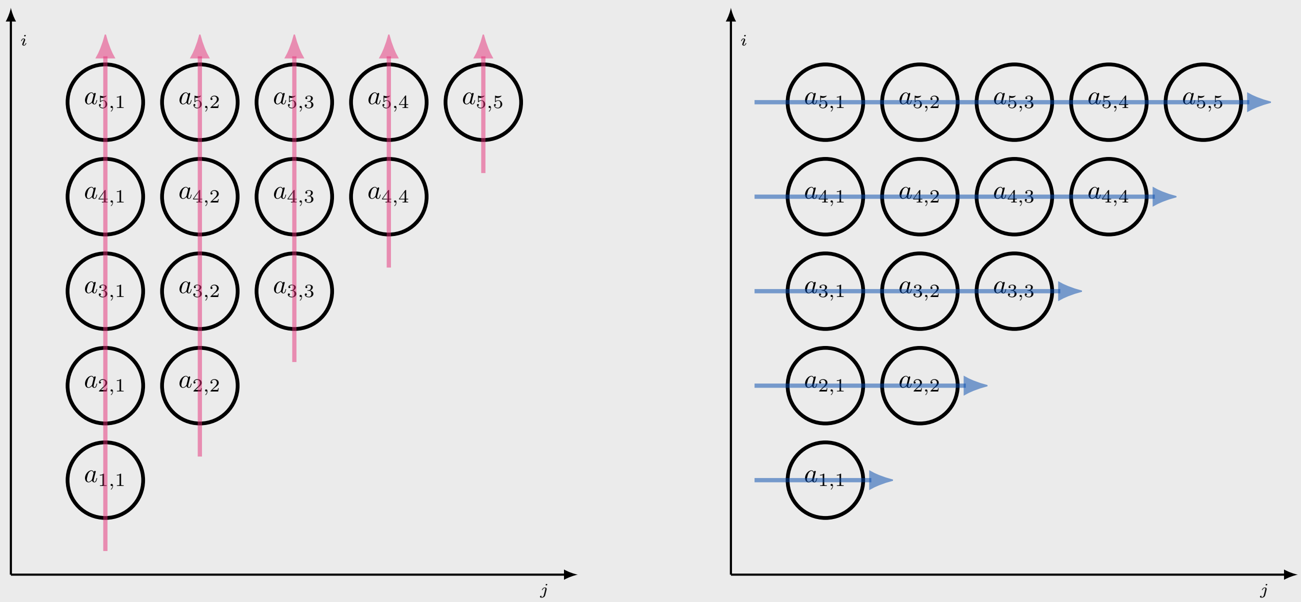
On cherche à calculer de deux manières différentes la somme $\displaystyle\sum_{k=1}^{n} kx^k$.

1. En écrivant $\displaystyle\sum_{k=1}^{n} kx^k$ comme une somme double, répondre à la question posée.
2. Retrouver le résultat à partir de la formule de la s **Proposition 11**.

Exercice 0.10.
Calculer, pour tout $n \in \mathbb{N}^*$, $\displaystyle\sum_{1 \leqslant i,j \leqslant n} \max(i,j)$.

Méthode 0.2. **Permutation des indices de sommation, dépendance du deuxième indice**
Parfois, le deuxième indice varie en fonction du premier et l'ordre de sommation peut avoir une incidence sur la simplification des calculs. On fait donc attention à cette dépendance d'indices lors de la permutation de ceux-ci

$$1 \leq j \leq i \leq n \iff \begin{cases} 1 \leq j \leq n \\ j \leq i \leq n \end{cases} \iff \begin{cases} 1 \leq i \leq n \\ 1 \leq j \leq i \end{cases}$$



$$\sum_{j=1}^{5} \sum_{i=j}^{5} a_{i,j} = \sum_{i=1}^{5} \sum_{j=1}^{i} a_{i,j}.$$

Exemple 0.1.

$$\begin{aligned} \sum_{1 \leq i \leq j \leq n} \frac{i}{j+1} &= \sum_{i=1}^{n} \sum_{j=i}^{n} \frac{i}{j+1} = \sum_{j=1}^{n} \sum_{i=1}^{j} \frac{i}{j+1} = \sum_{j=1}^{n} \frac{1}{j+1} \sum_{i=1}^{j} i \\ &= \sum_{j=1}^{n} \frac{1}{j+1} \times \frac{j(j+1)}{2} = \frac{1}{2} \sum_{j=1}^{n} j = \frac{n(n+1)}{4}. \end{aligned}$$

Exercice 0.11.
Calculer $\displaystyle\sum_{1 \leq i \leq j \leq n} ij$.

0.3 Nombres complexes, Racines de l'unité

Exercice 0.12. **Extrait du cahier de vacances**
On pose $z = \dfrac{1+\sqrt{2}+i}{1+\sqrt{2}-i}$. Calculer$|z|$. Mettre z sous forme algébrique. Calculer z^{2025}.

Exercice 0.13. **Extrait du cahier de vacances**

Pour chacune des deux questions (indépendantes) suivantes, proposer une résolution par le calcul puis une autre résolution géométrique.

1. Déterminer les valeurs de z pour lesquelles z, $\dfrac{1}{z}$ et $1-z$ ont même module.
2. Trouver tous les nombres complexes z vérifiant $|z| = |z-4|$ et $\arg(z) = \arg(z+1+i)$.

Proposition. 0.16. **Formules d'Euler**

Pour tout $\theta \in \mathbb{R}$, on a :

$$\cos(\theta) = \frac{e^{i\theta}+e^{-i\theta}}{2} \quad \text{et} \quad \sin(\theta) = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

Méthode 0.3. **Linéarisation d'une expression trigonométrique**

Soient $(p,q) \in \mathbb{N}^2$ et $x \in \mathbb{R}$. Pour linéariser l'expression $\cos^p(x)\sin^q(x)$, on peut procéder comme suit.

1. On commence par utiliser les formules d'Euler :
$$\cos^p(x)\sin^q(x) = \left(\frac{e^{ix}+e^{-ix}}{2}\right)^p \left(\frac{e^{ix}-e^{-ix}}{2i}\right)^q.$$
2. On développe ensuite les expressions $\left(\dfrac{e^{ix}+e^{-ix}}{2}\right)^p$ et $\left(\dfrac{e^{ix}-e^{-ix}}{2i}\right)^q$ à l'aide de la formule du binôme de Newton, puis on développe le produit de deux facteurs ainsi obtenu.
3. Enfin, il suffit de remarquer que la somme obtenue fait apparaître des couples de termes conjugués : grâce encore une fois à la formule d'Euler on peut ne faire apparaitre que des cosinus et des sinus.

Exercice 0.14. **Extrait cahier de vacances**

Linéariser $\cos^4(x)$.

Proposition. 0.17. **Formule de Moivre**

Pour tout $\theta \in \mathbb{R}$ et tout $n \in \mathbb{N}$, on a : $\cos(n\theta) + i\sin(n\theta) = (\cos(\theta) + i\sin(\theta))^n$.

Méthode 0.4.

La formule de Moivre ci-dessus, combinée la formule du binôme et une identification des parties réelles et imaginaires permet d'exprimer $\cos(n\theta)$ et $\sin(n\theta)$ en fonction des puissances de $\cos(\theta)$ et $\sin(\theta)$)

$$\cos(n\theta) = \sum_{\substack{k=0 \\ k \text{ pair}}}^{n} \binom{n}{k} \cos^{n-k}(\theta)(i\sin(\theta))^k, \qquad \sin(n\theta) = \sum_{\substack{k=0 \\ k \text{ pair}}}^{n} \binom{n}{k} \cos^{n-k}(\theta)(i\sin(\theta))^k.$$

Exercice 0.15.

Exprimer $\cos(3\theta)$ en fonction de $\cos(\theta)$.

0.3.1 Racines de l'unité

Dans cette section, on considère un entier naturel non nul n.

Définition 0.4. **Racines de l'unité**

On appelle **racine n-ième de l'unité** toute solution dans $\mathbb{C}$ de l'équation $z^n = 1$. On note alors $\mathbb{U}_n$ l'ensemble des racines n-ièmes de l'unité :

$$\mathbb{U}_n = \{z \in \mathbb{C} : \ z^n = 1\}.$$

Proposition. 0.18.

On a les résultats suivants :

i. $\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} \ : \ k \in [\![0, n-1]\!] \right\}$.

ii. $\mathbb{U}_n$ contient exactement n éléments.

iii. Posons $\omega = e^{\frac{2i\pi}{n}}$. On peut alors écrire $\mathbb{U}_n = \{\omega^k : \ k \in [\![0, n-1]\!]\}$.

Proposition. 0.19. **Somme des racines $n-$ièmes de l'unité**

Soit $n \in \mathbb{N}^*$ et $\omega = e^{\frac{2i\pi}{n}}$. On a

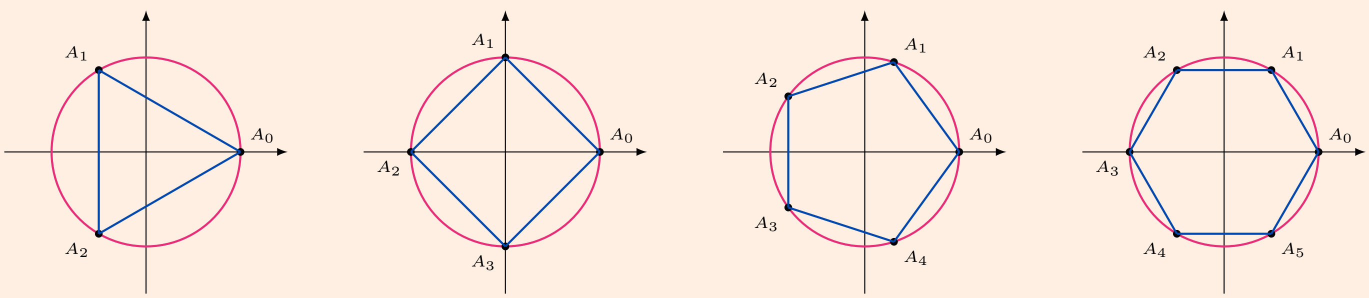
$$\sum_{z \in \mathbb{U}_n} z = \sum_{k=0}^{n-1} \omega^k = \begin{cases} 0 & \text{si } n \neq 1 \\ 1 & \text{si } n = 1 \end{cases}.$$

Exemple 0.2.

On a $\mathbb{U}_1 = \{1\}$, $\mathbb{U}_2 = \{1, -1\}$, $\mathbb{U}_3 = \{1, j, j^2\}$ où $j = e^{\frac{2i\pi}{3}} = \dfrac{-1 + i\sqrt{3}}{2}$.
De plus, on a $j^2 = j^{-1} = \overline{j}$ et $1 + j + j^2 = 0$.

Proposition. 0.20. **Polygones réguliers associés aux racines $n-$ièmes de l'unité**

Pour $n \geqslant 3$, les points dont l'affixe est dans $\mathbb{U}_n$ sont les sommets d'un polygone régulier inscrit dans le cercle trigonométrique.



$\mathbb{U}_3$ $\mathbb{U}_4$ $\mathbb{U}_5$ $\mathbb{U}_6$

Définition 0.5.

Soient $a \in \mathbb{C}$ et $n \in \mathbb{N}^*$. On dit que z est une racine n-ième de a si $z^n = a$.

Proposition. 0.21. **Racines d'un complexe non nul**

Soit $a \in \mathbb{C}^*$. Alors :

i. a admet exactement n racines n-ièmes.
ii. Si b est une racine n-ième quelconque de a, alors l'ensemble des racines n-ièmes de a est :

$$\{z \in \mathbb{C}^* : \ z^n = a\} = \{bz : \ z \in \mathbb{U}_n\} = \left\{be^{\frac{2ik\pi}{n}} : \ k \in [\![0, n-1]\!]\right\}.$$

☞ Pour obtenir toutes les racines n-ièmes d'un complexe $a \in \mathbb{C}^*$, il suffit donc d'en déterminer une, puis de multiplier celle-ci par toutes les racines n-ème de l'unité.

Exercice 0.16. **Extrait du cahier de vacances**

Déterminer les racines cubiques de $-8i$.

0.4 Polynômes

0.4.1 Généralités

Dans toute la suite, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
On note $\mathbb{K}[X]$ est l'ensemble des polynômes à cœfficients dans $\mathbb{K}$.
Pour tout $n \in \mathbb{N}$, $\mathbb{K}_n[X]$ est l'ensemble des polynômes à cœfficients dans $\mathbb{K}$ de degré inférieur (ou égal) à n.

Remarque 0.2.

Dans la suite la lettre X ne désigne pas une variable, mais une *indéterminée*. Il n'aurait ainsi aucun sens d'écrire par exemple une formule commençant par $\forall X \in \mathbb{K}$, ou d'écrire "prenons $X = 1$".

Proposition. 0.22. **Théorème d'identification des cœfficients**

Deux polynômes sont égaux si et seulement si ils ont les mêmes cœfficients.

☞ En particulier, un polynôme est nul si et seulement si tous ses cœfficients sont nuls.

Exercice 0.17. **Formule de Vandermonde**

Soient $n \in \mathbb{N}$ et $P(X) = (X+1)^n$.

1. En calculant de deux manières différentes le polynôme P^2, montrer que :

$$\forall k \in [\![0, 2n]\!], \quad \sum_{i=0}^{k} \binom{n}{i}\binom{n}{k-i} = \binom{2n}{k}.$$

2. En déduire la *formule de Vandermonde*: $\displaystyle\sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}$.

Théorème 0.23. **Division euclidienne sur les polynômes**

Soit $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$. Alors, il existe un unique couple $(Q, R) \in \mathbb{K}[X]^2$ tel que :

$$(i)\ A = BQ + R, \qquad (ii)\ \deg(R) < \deg(B).$$

On dit alors que Q est le **quotient** et R le **reste** de la division euclidienne de A par B.

Exercice 0.18.

Calculer le quotient et le reste de la division euclidienne de $X^4 - 1$ par $X - 1$.

Exercice 0.19.

Calculer $\displaystyle\int_0^1 \frac{x^4 + x - 1}{x^2 + 1} \mathrm{d}x$.

☞ Des rappels d'intégration apparaîtront au **Chapitre 7**.

Définition 0.6.

Soit $A, B \in \mathbb{K}[X]$. On dit que B **divise** A s'il existe $Q \in \mathbb{K}[X]$ tel que $A = BQ$.

☞ Il est clair que B divise A si et seulement si le reste dans la division euclidienne de A par B est nul.

0.4.2 Racines d'un polynôme

Définition 0.7. **Racine d'un polynôme**

Soit $P \in \mathbb{K}[X]$ et $\alpha \in \mathbb{K}$. On dit que α est une **racine** (ou est **un zéro**) de P si $P(\alpha) = 0$.

Théorème 0.24. **Caractérisation des racines d'un polynôme**

Soit $P \in \mathbb{K}[X]$ et $\alpha \in \mathbb{K}$. Les assertions suivantes sont équivalentes.

i. α est une racine de P.
ii. $(X - \alpha)$ divise P.
iii. $\exists Q \in \mathbb{K}[X],\ P(X) = (X - \alpha)Q(X)$.

Définition 0.8. **Ordre d'une racine**

Soit $P \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $k \in \mathbb{N}$. On dit que α est une **racine d'ordre** k de P si $(X-\alpha)^k$ divise P et $(X-\alpha)^{k+1}$ ne divise pas P. L'entier k est alors appelé **l'ordre de multiplicité** de α.

Théorème 0.25. **Caractérisation des racines de multiplicité k d'un polynôme**

Soit $P \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $k \in \mathbb{N}$. Les assertions suivantes sont équivalentes.

i. α est une racine de multiplicité k de P.
ii. $\exists Q \in \mathbb{K}[X],\ P(X) = (X - \alpha)^k Q(X)$ et $Q(\alpha) \neq 0$.
iii. $\forall i \in [\![0, k-1]\!]$, $P^{(i)}(\alpha) = 0$ et $P^{(k)}(\alpha) \neq 0$.

0.4.3 Nombre de racines d'un polynôme

Théorème 0.26.
Soit $P \in \mathbb{K}[X]$ un polynôme de degré $n \in \mathbb{N}$. Alors P admet au plus n racines distinctes.

Corollaire 0.27.
Soit $P \in \mathbb{K}[X]$ un polynôme de degré $n \in \mathbb{N}$. Alors le nombre de racines de P, comptées avec multiplicité, est inférieur ou égal à n.

C'est-à-dire que si $\alpha_1, \dots, \alpha_r$ sont des racines de P de multiplicité respective $m_1, \dots, m_r$, on a $\sum_{k=1}^{r} m_k \leqslant n$.

Corollaire 0.28.
Soit $P \in \mathbb{K}_n[X]$. Si P admet au moins $n+1$ racines distinctes, alors $P = 0$.

☞ On en déduit que *tout polynôme admettant une infinité de racines distinctes est nul.*

Corollaire 0.29.
Soit $P, Q \in \mathbb{K}_n[X]$ et $\alpha_1, \dots, \alpha_{n+1}$ des scalaires deux à deux distincts. Si, pour tout $i \in [\![1, n+1]\!]$, $P(\alpha_i) = Q(\alpha_i)$, alors $P = Q$.

☞ Deux polynômes qui coïncident en un nombre de points strictement supérieur à leur degré coïncident partout.

0.4.4 Polynômes scindés

Définition 0.9. **Polynôme scindé**
Soit $P \in \mathbb{K}[X]$. On dit que P est **scindé sur** $\mathbb{K}$ lorsque P est constant ou lorsque il existe $\lambda \in \mathbb{K}^*$ et $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ (pas nécessairement distinctes) tels que :

$$P(X) = \lambda \prod_{i=1}^{n} (X - \alpha_i).$$

En d'autres termes, un polynôme est scindé s'il peut se décomposer comme produit de polynômes de degré 1.

Exemple 0.3.
$X^2 + 1$ est scindé sur $\mathbb{C}$, mais pas sur $\mathbb{R}$.

Théorème 0.30. **Théorème de d'Alembert-Gauss**
Tout polynôme de $\mathbb{C}[X]$ de degré supérieur ou égal à 1 admet au moins une racine complexe.

Preuve. Résultat admis. □

Théorème 0.31.
Tout polynôme à cœfficients complexes est scindé sur $\mathbb{C}$.

Remarque 0.3.
Tout polynôme à cœfficients complexes peut alors se mettre sous la forme $P(X) = \lambda \prod_{i=1}^{r} (X - \alpha_i)^{m_i}$, où $\{\alpha_1, \dots, \alpha_r\}$ est l'ensemble des racines de P, λ est son cœfficient dominant, et m_i est la multiplicité de α_i en tant que racine de P. Cette écriture est alors unique à l'ordre des facteurs près.

Exercice 0.20. **Extrait du cahier de vacances**
Factoriser $X^5 - 1$ dans $\mathbb{R}[X]$.

Remarque 0.4.
On voit sur cet exemple que tout polynôme à cœfficients réels, à défaut d'être scindé, se décompose comme un produit de polynômes de degré 1 et de polynômes de degré 2 à discriminant strictement négatif.

0.4.5 Relations cœfficients racines

Proposition. 0.32.
Soit $P(X) = aX^2 + bX + c \in \mathbb{K}[X]$, avec $a \neq 0$, et $(\alpha, \beta) \in \mathbb{K}^2$. Les assertions suivantes sont équivalentes.

i. α et β sont les racines de P.
ii. $P(X) = a(X - \alpha)(X - \beta)$.
iii. $\alpha + \beta = \dfrac{-b}{a}$ et $\alpha\beta = \dfrac{c}{a}$.

☞ On retrouve le résultat suivant: a et b sont racines de $X^2 - SX + P$, où $S = a + b$ et $P = ab$.

On a enfin le résultat suivant, pour les polynômes de degré supérieur.

Proposition. 0.33.
Soit $n \in \mathbb{N}^*$, $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{K}[X]$, avec $a_n \neq 0$ un polynôme scindé et $\alpha_1, \ldots, \alpha_n$ les racines de P. Alors :

$$(i)\ \sum_{k=1}^{n} \alpha_k = -\frac{a_{n-1}}{a_n}, \qquad (ii)\ \prod_{k=1}^{n} \alpha_k = (-1)^n \frac{a_0}{a_n}.$$

0.5 Suites numériques

0.5.1 Convergence des suites

Définition 0.10. **Limite d'une suite**

- ✗ On dit qu'une suite (u_n) **diverge vers l'infini**, ce que l'on notera $\lim\limits_{n\to+\infty} u_n = +\infty$, si elle vérifie :

$$\forall A > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ u_n > A.$$

- ✗ On dit qu'une suite (u_n) **diverge vers moins l'infini**, ce que l'on notera $\lim\limits_{n\to+\infty} u_n = -\infty$, si elle vérifie :

$$\forall A < 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ u_n < A$$

- ✗ Soit (u_n) une suite et $\ell \in \mathbb{R}$. On dit que (u_n) **converge vers** ℓ, ce que l'on note $\lim\limits_{n\to+\infty} u_n = \ell$, si elle vérifie

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

Proposition. 0.34. **Prolongement des inégalités**
Soient a, b des réels, et (u_n) une suite convergente telle qu'il existe $n_0 \in \mathbb{N}$ vérifiant: $\forall n \geqslant n_0,\ a < u_n < b$.
Alors,

$$a \leqslant \lim_{n\to+\infty} u_n \leqslant b.$$

Théorème 0.35. **Théorème de la limite monotone**
Soit (u_n) une suite réelle monotone.

- ✗ Si (u_n) est croissante, alors:
  i. si la suite est majorée, elle possède une limite finie ℓ;
  ii. sinon, elle tend vers $+\infty$.

  Et dans le cas où $\lim\limits_{n\to+\infty} u_n = \ell$, on a : $\forall n \in \mathbb{N}, u_n \leqslant \ell$.

- ✗ Si (u_n) est décroissante, alors:
  i. si la suite est minorée elle possède une limite finie ℓ;
  ii. sinon, elle tend vers $-\infty$.

  Et dans le cas où $\lim\limits_{n\to+\infty} u_n = \ell$, on a : $\forall n \in \mathbb{N}, u_n \geqslant \ell$.

☞ Toute suite monotone possède donc une limite, finie ou infinie.

Définition 0.11.
On dit que deux suites (a_n) et (b_n) sont adjacentes, si

- ✗ l'une de ces deux suites est croissante et l'autre est décroissante;
- ✗ la différence des deux termes généraux tend vers 0 : $\lim\limits_{n\to+\infty}(b_n - a_n) = 0$.

Théorème 0.36. **Suites adjacentes**
Soient (a_n) et (b_n) deux suites adjacentes. Alors elles convergent toutes les deux et ont même limite.

Exercice 0.21. **Irrationalité de e et suites adjacentes**

1. Montrer que les suites (u_n) et (v_n) définies, pour tout $n \in \mathbb{N}$, par
$$u_n = \sum_{k=0}^{n} \frac{1}{k!}, \quad \text{et} \quad v_n = \sum_{k=0}^{n} \frac{1}{k!} + \frac{1}{n \cdot n!}$$
sont adjacentes.
2. En déduire que e est irrationnel.

Définition 0.12. **Suite extraite**
Étant donnée une suite $(u_n)_{n\in\mathbb{N}}$, on appelle **suite extraite** de (u_n) ou **sous-suite** de (u_n) toute suite de la forme $\left(u_{\varphi(n)}\right)_{n\in\mathbb{N}}$ où $\varphi : \mathbb{N} \to \mathbb{N}$ est une fonction strictement croissante.

☞ Une sous-suite d'une suite (u_n) est donc une suite obtenue en ne conservant que certains des termes de (u_n), extraits dans l'ordre.

Proposition. 0.37.
Si une suite (u_n) possède une limite, alors toute suite extraite de (u_n) possède la même limite.

Théorème 0.38.

Soit (u_n) une suite, et ℓ qui est soit un réel soit égal à $\pm\infty$. On a alors : $u_n \underset{n\to+\infty}{\longrightarrow} \ell \iff \begin{cases} u_{2n} \underset{n\to+\infty}{\longrightarrow} \ell \\ u_{2n+1} \underset{n\to+\infty}{\longrightarrow} \ell \end{cases}$.

Théorème 0.39. **Encadrement et comparaison**
Soient (u_n), (v_n) et (w_n) trois suites. Alors

i. Si, il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geqslant n_0$, $u_n \leqslant v_n$ et que (u_n) diverge vers $+\infty$, alors (v_n) diverge aussi vers $+\infty$.

ii. Si, il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geqslant n_0$, $u_n \leqslant v_n$ et que (v_n) diverge vers $-\infty$, alors (u_n) diverge aussi vers $-\infty$.

iii. Si, il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geqslant n_0$, $w_n \leqslant u_n \leqslant v_n$ et que (w_n) et (v_n) convergent vers une même limite finie ℓ, alors (u_n) converge également vers ℓ.

Exercice 0.22.
Soit $x \in \mathbb{R}$. Déterminer $\lim\limits_{n\to+\infty} \dfrac{\lfloor nx \rfloor}{n}$.

0.5.2 Suites usuelles

Exercice 0.23.
Dans chacun des cas suivants, déterminer u_n en fonction de n. On en profitera pour rappeler la méthode.

1. $u_0 = 2$ et $\forall n \in \mathbb{N}, u_{n+1} = 3u_n - 2$
2. $u_0 = u_1 = 1$ et $\forall n \in \mathbb{N}, u_{n+2} - 5u_{n+1} + 6u_n = 0$.
3. $u_0 = 0, u_1 = 1$ et $\forall n \in \mathbb{N}, u_{n+2} - 4u_{n+1} + 4u_n = 0$.

0.5.3 Suites récurrentes du type $u_{n+1} = f(u_n)$

Méthode 0.5. **Étude suite récurrente $u_{n+1} = f(u_n)$ - avec f croissante**

Pour étudier une suite $(u_n)_{n\geqslant 0}$ définie par son premier terme u_0 et une relation de récurrence du type $u_{n+1} = f(u_n)$, on suit les étapes ci-contre :

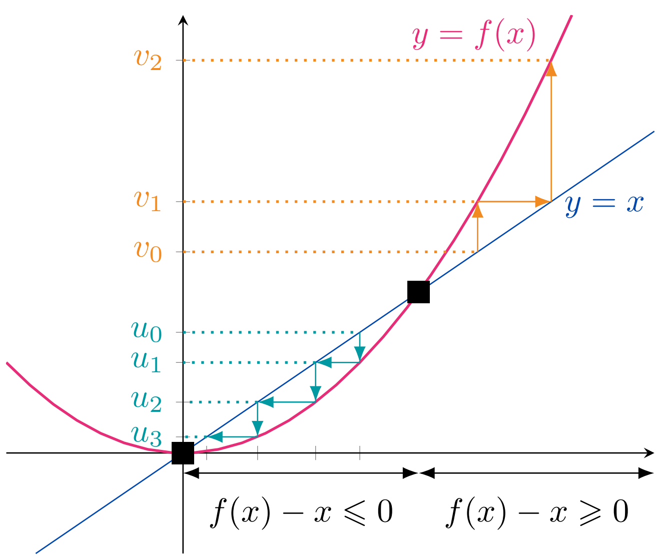
- ✗ On étudie brièvement les variations de f sur son ensemble de définition ;
- ✗ On vérifie que la suite $(u_n)_{n\geqslant 0}$ est bien définie (il n'y a pas de terme u_n qui tombe en dehors du domaine de définition de f pour un certain n, rendant impossible le calcul des termes suivants).
  Pour ce faire, en général on montre une condition plus restrictive, souvent que la suite est bornée (tous les termes sont dans un certain intervalle, stable par f) par récurrence.
- ✗ On étudie les variations de la suite.
  Pour ce faire, si f est croissante, on peut utiliser une récurrence, si l'on connait les deux premiers termes de la suite (pour initialiser la récurrence) - les autres termes seront rangés dans le même ordre.
  Sinon, on étudie le signe de $f(x) - x$.

  Attention, écrire que la suite est croissante car la fonction est croissance est une hérésie. Une fonction croissante peut générer une suite croissante ainsi qu'une suite décroissante (selon la localisation du premier terme).
  Par contre, une fonction croissante génère toujours une suite **monotone**.

- ✗ On utilise le **théorème de convergence monotone** pour justifier de la convergence de la suite vers une certaine limite ℓ (qu'on ne connait pas à ce stade) qui se trouve dans le même intervalle (**fermé**) que les termes de la suite.

- ✗ On utilise la continuité de f au voisinage de ℓ pour obtenir l'équation de point fixe vérifié par ℓ (à savoir $f(\ell) = \ell$) et déterminer la valeur de ℓ.
  Les points fixes de f sont les seules valeurs possibles pour une limite finie (lorsque f est continue).

Exemple 0.4. **Suite récurrente et schéma en escalier**



Sur l'exemple ci-contre,

- ✗ f est (continue et) croissante sur $\mathbb{R}_+$.

- ✗ Les points fixes, donc limites finies éventuelles pour une suite de la forme $u_{n+1} = f(u_n)$ sont $\ell = 0$ ou $\ell = 1$.

- ✗ $f(x) - x \leqslant 0$ sur $[0, 4]$. Si on prend un premier terme $u_0 \in [0; 4]$, tous les termes de la suite sont dans le même intervalle (stable par f) et la suite $(u_n)_{n\geqslant 0}$ générée est **décroissante**. On voit qu'elle converge vers $\ell = 0$.

- ✗ $f(x) - x \geqslant 0$ sur $[4, +\infty[$. Si on prend un premier terme $v_0 > 4$, alors tous les termes de la suite donc dans ce même intervalle (stable par f) et la suite $(v_n)_{n\geqslant 0}$ générée est **croissante**. On montre par l'absurde qu'elle ne peut pas converger (car $v_n \geqslant v_0 > 4$) donc elle diverge vers $+\infty$.

☞ Lorsque f est décroissante, on peut étudier les suites extraites de rangs pairs et impairs qui seront elles monotones. Si elles convergent vers une même limite, (u_n) est convergence vers cette limite.

Exercice 0.24.

Soit f la fonction définie sur $\mathbb{R}_+$ par $f(x) = x\ln(1+x)$. On considère la suite $(u_n)_{n\in\mathbb{N}}$ définie par $u_0 \in]0;+\infty[$ et $u_{n+1} = f(u_n)$ pour tout $n \in \mathbb{N}$.

1. **a.** Montrer que f est $\mathcal{C}^2$ sur $[0;+\infty[$ et calculer, pour tout $x \in [0;+\infty[$, $f'(x)$ et $f''(x)$.
   b. En déduire les variations de f.

2. **a.** Étudier le signe de la fonction $g : x \mapsto f(x) - x$ sur $\mathbb{R}_+$.
   b. Quelles sont les limites possibles de la suite $(u_n)_{n\in\mathbb{N}}$?

3. On suppose dans cette question : $u_0 \in]e-1;+\infty[$.
   a. Montrer que : pour tout $n \in \mathbb{N}$, $e - 1 < u_n \leq u_{n+1}$.
   b. En déduire que la suite $(u_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$.

4. On suppose, dans cette question : $u_0 \in]0;e-1[$.
   Étudier la convergence de $(u_n)_{n\in\mathbb{N}}$.

Exercice 0.25.

Soit (u_n) définie par $u_0 = 1$ et, pour $n \in \mathbb{N}$, $u_{n+1} = f(u_n)$ avec $f(x) = 1 + \dfrac{2}{x}$.

1. Montrer que, pour tout $n \in \mathbb{N}$, u_n est bien défini et que $u_n \in [1;3]$.
2. Calculer u_1, u_2, u_3. La suite (u_n) est-elle monotone?
3. Montrer que (u_{2n}) est croissante et que (u_{2n+1}) est décroissante.
4. Conclure quant à la nature de la suite (u_n).

☞ Dans le cas où la fonction f serait contractante (qu'elle soit croissante ou non), on peut utiliser le résultat suivant. Il donne une information sur la vitesse de convergence (qui est sous-géométrique) et permet notamment l'écriture de script en `Python` pour obtenir une valeur approchée de la limite.

Théorème 0.40. **Suites récurrentes et fonction contractante**

Soient $q \in [0,1[$ et $f : [a,b] \to [a,b]$ dérivable telle que $\forall x \in [a,b],\ |f'(x)| \leqslant q$.
On définit une suite (u_n) en posant $u_0 \in [a,b]$ et $u_{n+1} = f(u_n)$. On a alors les résultats suivants.

i. f possède un unique point fixe ℓ;
ii. $\forall n \in \mathbb{N},\ |u_{n+1} - \ell| \leqslant q|u_n - \ell|$;
iii. $\forall n \in \mathbb{N},\ |u_n - \ell| \leqslant q^n|u_0 - \ell|$;
iv. (u_n) converge vers ℓ.

Exercice 0.26.

On considère l'équation (E) $2x + \sin(x) = 4$.

1. Montrer que (E) possède une unique solution que l'on notera α.
2. Montrer que la suite $(u_n)_{n\geqslant 0}$ définie par $u_0 = 0$ et $u_{n+1} = \dfrac{1}{2}(4 - \sin(u_n))$ converge vers α.
3. Déterminer un entier n_0 tel que u_{n_0} soit une valeur approchée de α à ε près.
4. Écrire une fonction `Python` prenant en argument un réel $\varepsilon > 0$ qui renvoie une valeur approchée de α à ε près.

0.5.4 Comparaison de suites au voisinage de l'infini

Ces notions seront rappelées dans le cadre des fonctions et détaillées dans le **Chapitre 1**.

Définition 0.13. **Quantités équivalentes, négligeables et dominées**

- ✗ Soient (u_n) et (v_n) deux suites réelles. On suppose que la suite (v_n) ne s'annule pas à partir d'un certain rang. On dit que (u_n) est **négligeable** devant (v_n) (en $+\infty$), et on note $u_n \underset{+\infty}{=} o(v_n)$ si et seulement si

$$\lim_{n\to+\infty} \frac{u_n}{v_n} = 0.$$

- ✗ Soient (u_n) et (v_n) deux suites réelles. On suppose que la suite (v_n) ne s'annule pas à partir d'un certain rang. On dit que les quantités u_n et v_n sont **équivalentes** (en $+\infty$), et on note $u_n \underset{+\infty}{\sim} v_n$ si et seulement si

$$\lim_{n\to+\infty} \frac{u_n}{v_n} = 1.$$

- ✗ Soient (u_n) et (v_n) deux suites réelles. On suppose que la suite (v_n) ne s'annule pas à partir d'un certain rang. On dit que (u_n) est **dominée** par (v_n), et on note $u_n \underset{+\infty}{=} O(v_n)$ si et seulement si la suite $\left(\dfrac{u_n}{v_n}\right)$ est bornée.

Exercice 0.27.

Montrer que, si (u_n) est négligeable devant (v_n), alors (u_n) est dominée par (v_n).
Que dire de la réciproque ?

Proposition. 0.41.

Si (u_n) est une suite convergente de limite ℓ **non nulle**, alors on a $u_n \sim \ell$.

On prendra bien garde à ne **jamais** écrire qu'une quantité est équivalente à 0, ce qui n'a aucun sens.

Proposition. 0.42. **Caractérication de l'équivalence par les petits-o**

On a l'équivalence :

$$u_n \sim v_n \Longleftrightarrow u_n = v_n + o(v_n) \qquad (n \to +\infty)$$

☞ La réciproque indique qu'on peut négliger les termes négligeables dans une somme. Ainsi

$$u_n + o(u_n) \sim u_n.$$

Proposition. 0.43. **Équivalents et limites**

Si les suites (u_n) et (v_n) sont équivalentes, alors elles sont de même nature, c'est-à-dire

i. Si $u_n \sim v_n$ et si (v_n) converge vers ℓ, alors (u_n) converge également vers ℓ.
ii. Si $u_n \sim v_n$ et si (v_n) diverge vers $\pm\infty$, alors (u_n) diverge également vers $\pm\infty$.

1 Fonctions d'une variable réelle

Dans tout le chapitre, I désigne un intervalle non trivial de $\mathbb{R}$ (c'est à dire non vide et non réduit à un seul point).
Si les bornes de I sont deux réels a et b, on appelle *adhérence* de I, et on note $\overline{I}$, l'ensemble $\overline{I} = I \cup \{a, b\}$.
Plus généralement, l'adhérence d'une partie A de $\mathbb{R}$ est l'ensemble obtenu en rajoutant à A les points de sa *frontière*.
Par extension, on note $\overline{\mathbb{R}} = [-\infty; +\infty]$.

Le début de ce chapitre propose principalement des rappels sur l'étude locale et globale des fonctions d'une variable réelle à valeurs réelles.
On étend ensuite certains résultats aux fonctions à valeurs vectorielles.

Définition 1.1. **Voisinage d'un point**

Soit $x_0 \in \mathbb{R}$. Alors,

- ✗ tout intervalle de la forme $]x_0 - \delta, x_0 + \delta[$ avec $\delta > 0$ est appelé "un voisinage de x_0";
- ✗ tout intervalle de la forme $]x_0, x_0 + \delta[$ avec $\delta > 0$ est appelé "un voisinage de x_0^+";
- ✗ tout intervalle de la forme $]x_0 - \delta, x_0[$ avec $\delta > 0$ est appelé "un voisinage de x_0^-";
- ✗ tout intervalle de la forme $]A, +\infty[$ avec $A > 0$ est appelé "un voisinage de $+\infty$";
- ✗ tout intervalle de la forme $]-\infty, A[$ avec $A < 0$ est appelé "un voisinage de $-\infty$".

L'ensemble des voisinages de $\alpha \in \overline{\mathbb{R}}$ est noté $V(\alpha)$.

☞ Tout voisinage de $\alpha \in \overline{\mathbb{R}}$ est un intervalle ouvert et non trivial, car on a pris $\delta > 0$.

Définition 1.2.

Soit $\alpha \in \overline{\mathbb{R}}$ et $P(x)$ une propriété dépendant de x. On dit que "$P(x)$ est vraie au voisinage de α" s'il existe un voisinage V de α tel que $P(x)$ soit vraie pour tout $x \in V$.

Exemple 1.1.

- ✗ Dire que $f : \mathbb{R} \to \mathbb{R}$ est positive au voisinage de $+\infty$ c'est dire qu'il existe un voisinage de $+\infty$ dans lequel f est positive, c'est à dire :
$$\exists A > 0, \forall x \in]A, +\infty[,\ f(x) \geqslant 0.$$
- ✗ Soit x_0 un réel. Dire que $f : I \to \mathbb{R}$ est définie au voisinage de x_0 c'est dire que :
$$\exists \delta > 0,]x_0 - \delta, x_0 + \delta[\subset I.$$

1.1 Limites d'une fonction à valeurs réelles

Définition 1.3. **Limite en x_0**

Soit $f : I \to \mathbb{R}$. Si x_0 et ℓ sont des réels:

- ✗ $\lim\limits_{x \to x_0} f(x) = \ell$ équivaut à $\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in I,\ \ (|x - x_0| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon)$.
- ✗ $\lim\limits_{x \to x_0} f(x) = +\infty$ équivaut à $\forall A > 0,\ \exists \delta > 0,\ \forall x \in I,\ \ (|x - x_0| < \delta \Longrightarrow f(x) > A)$.
- ✗ $\lim\limits_{x \to +\infty} f(x) = \ell$ équivaut à $\forall \varepsilon > 0,\ \exists A > 0,\ \forall x \in I,\ \ (x > A \Longrightarrow |f(x) - \ell| < \varepsilon)$.

Cette définition s'applique également si α, β sont de la forme x_0^+ ou x_0^- avec $x_0 \in \mathbb{R}$, ce qui donne par exemple :
$$\lim_{x \to x_0^+} f(x) = \ell \Leftrightarrow \forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in I, (0 < x - x_0 < \delta \Longrightarrow |f(x) - \ell| < \varepsilon).$$

Exercice 1.1.
Soit $f: \mathbb{R} \to \mathbb{R}$ une fonction impaire. On suppose que f admet ℓ pour limite en $+\infty$. Montrer que f admet $-\ell$ pour limite en $-\infty$.

1.1.1 Quelques résultats sur les limites

Théorème 1.1. **Caractérisation séquentielle de la limite**
Soient $f : I \to \mathbb{R}$ une fonction, $\alpha \in \overline{I}$ et $\beta \in \overline{\mathbb{R}}$. Alors $\lim\limits_{x\to\alpha} f(x) = \beta$ si et seulement si pour toute suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de I tendant vers α, on a $\lim\limits_{n\to+\infty} f(u_n) = \beta$.

Exercice 1.2.
Soit $f: \mathbb{R} \to \mathbb{R}$ une fonction continue et périodique (de période T) admettant une limite finie en $+\infty$. Montrer que f est constante.

Théorème 1.2. **Prolongement des inégalités**
Soient $f : I \to \mathbb{R}$, $\alpha \in \overline{I}$, et a, b des réels. On suppose que f possède une limite finie en α et que

$$\forall x \in I,\ a < f(x) < b.$$

On a alors

$$a \leqslant \lim_{x\to\alpha} f(x) \leqslant b.$$

☞ Le passage à la limite prolonge donc tout type d'inégalités en **inégalités larges**.

Lemme 1.3. **Théorème des gendarmes**
Soient $f, g, h : I \to \mathbb{R}$, $\alpha \in \overline{I}$, et ℓ un réel. On suppose encore:

- ✗ $\forall x \in I,\ g(x) \leqslant f(x) \leqslant h(x)$;
- ✗ $\lim\limits_{x\to\alpha} g(x) = \lim\limits_{x\to\alpha} h(x) = \ell$.

Alors (f possède une limite lorsque x tend vers α et) $\lim\limits_{x\to\alpha} f(x) = \ell$.

Exercice 1.3.
Montrer, à l'aide de considérations géométriques, que, pour tout $x \in \left]0, \frac{\pi}{2}\right[$,

$$0 < \sin(x) \leq x \leq \tan(x).$$

En déduire $\lim\limits_{x\to0} \dfrac{\sin(x)}{x} = 1$.

Les théorèmes suivants sont l'expression fonctionnelle du théorème de la limite monotone vu dans le cadre des suites.

Théorème 1.4. **Théorème de la limite monotone**

Soient $a, b \in \overline{\mathbb{R}}$, de sorte que $a < b$, et $f :]a, b[\to \mathbb{R}$ une fonction.
Si la fonction f est croissante sur $]a, b[$, alors elle admet une limite, finie ou infinie, en a et en b. Plus précisément, on a:

i. ✗ si f n'est pas majorée sur $]a, b[$, alors $\lim\limits_{x\to b^-} f(x) = +\infty$.
✗ si f est majorée sur $]a, b[$, alors f possède une limite finie en b^-.

ii. ✗ si f n'est pas minorée sur $]a, b[$, alors $\lim\limits_{x\to a^+} f(x) = -\infty$.
✗ si f est minorée sur $]a, b[$, alors f possède une limite finie en a^+.

Théorème 1.5. **Théorème de la limite monotone, bis**

Soient $a, b \in \overline{\mathbb{R}}$, de sorte que $a < b$, et $f :]a, b[\to \mathbb{R}$ une fonction.
Si la fonction f est décroissante sur $]a, b[$, alors elle admet une limite, finie ou infinie, en a et en b. Plus précisément, on a:

i. ✗ si f n'est pas majorée sur $]a, b[$, alors $\lim\limits_{x\to a^+} f(x) = +\infty$;
✗ si f est majorée sur $]a, b[$, alors f possède une limite finie en a^+.

ii. ✗ si f n'est pas minorée sur $]a, b[$, alors $\lim\limits_{x\to b^-} f(x) = -\infty$;
✗ si f est minorée sur $]a, b[$, alors f possède une limite finie en b^-.

1.1.2 Comparaison locale (au voisinage d'un point)

Définition 1.4. **Négligeabilité**

Soient f et g deux fonctions définies sur un voisinage V de $\alpha \in \overline{\mathbb{R}}$. On suppose que g ne s'annule pas sur V.
On dit que f est **négligeable devant** g au voisinage de α si et seulement si $\lim\limits_{x\to\alpha} \dfrac{f(x)}{g(x)} = 0$.

On note $f(x) \underset{x\to\alpha}{=} o(g(x))$ ou $f(x) = o(g(x)), x \to \alpha$.

Exemple 1.2.

✗ On peut alors écrire les relations suivantes

$$(i)\ x^2 \underset{x\to 0}{=} o(x), \qquad (ii)\ e^x \underset{x\to+\infty}{=} o\left(e^{2x}\right), \qquad (iii)\ \frac{1}{x^2} \underset{x\to+\infty}{=} o\left(\frac{1}{x}\right)$$

✗ Dire que $f(x) \underset{x\to\alpha}{=} o(1)$, c'est dire que f tend vers 0 en α. On a la même conclusion sous l'hypothèse $f(x) \underset{x\to\alpha}{=} o(c)$ où $c \in \mathbb{R}^*$.

Remarque 1.1.

Il faut bien comprendre et garder à l'esprit que la négligeabilité en un point n'entraîne nullement la négligeablité ailleurs. Il suffit de regarder l'exemple (i) ci-dessus

$$x^2 \underset{x\to 0}{=} o(x), \quad \text{mais} \quad x^2 \underset{x\to 2}{\neq} o(x) \quad \text{et} \quad x^2 \underset{x\to+\infty}{\neq} o(x).$$

On comprend l'importance de ne pas simplement écrire $f(x) = o(g(x))$ sans préciser le voisinage concerné!

Théorème 1.6. **Croissances comparées**

Soient r, s des réels.

i. On a
$$r < s \Rightarrow x^r \underset{x \to +\infty}{=} o(x^s) \qquad \text{et} \qquad r < s \Rightarrow x^s \underset{x \to 0}{=} o(x^r).$$
En d'autres termes, "les grandes puissances l'emportent au voisinage de $+\infty$, et les petites au voisinage de 0".

ii. Si $r > 0$, on a
$$x^r \underset{x \to +\infty}{=} o(e^x), \qquad \ln x \underset{x \to +\infty}{=} o(x^r) \qquad \text{et} \qquad \ln x \underset{x \to +\infty}{=} o(e^{rx}).$$
En d'autres termes, "au voisinage de $+\infty$, l'exponentielle l'emporte sur les puissances qui l'emportent sur le logarithme".

Exercice 1.4.

Calculer les limites de $x \mapsto x^x$ aux bornes de son domaine de définition.

Définition 1.5. **Domination**

Soient f et g deux fonctions définies sur un voisinage V de $\alpha \in \overline{\mathbb{R}}$. On suppose que g ne s'annule pas sur V.

On dit que f **est dominée par** g au voisinage de α si et seulement si $x \mapsto \dfrac{f(x)}{g(x)}$ est bornée sur V.

On note $f(x) \underset{x \to \alpha}{=} O(g(x))$ ou $f(x) = O(g(x)),\ x \to \alpha$.

Proposition. 1.7.

Soient f et g deux fonctions définies sur un voisinage de $\alpha \in \overline{\mathbb{R}}$. On a alors:
$$f(x) \underset{x \to \alpha}{=} o(g(x)) \implies f(x) \underset{x \to \alpha}{=} O(g(x)).$$

Définition 1.6. **Quantités équivalentes**

Soient f et g deux fonctions définies sur un voisinage V de $\alpha \in \overline{\mathbb{R}}$. On suppose que g ne s'annule pas sur V.

On dit que $f(x)$ est **équivalent** à $g(x)$ au voisinage de α si et seulement si $\lim\limits_{x \to \alpha} \dfrac{f(x)}{g(x)} = 1$.

On note $f(x) \underset{x \to \alpha}{\sim} g(x)$ ou $f(x) \sim g(x),\ x \to \alpha$.

Exercice 1.5.

Interpréter le résultat de l'**Exercice 3** comme un équivalent au voisinage de 0. Qu'en est-il au voisinage de $+\infty$?

☞ On remarquera qu'il découle immédiatement de la définition que, si $\ell \in \mathbb{R}^*$,
$$\lim_{x \to \alpha} f(x) = \ell \iff f(x) \underset{x \to \alpha}{\sim} \ell.$$

Proposition. 1.8.

Toute fonction polynomiale est équivalente à son monôme de plus haut degré en $\pm\infty$.
Toute fonction polynomiale est équivalente à son monôme de plus bas degré en 0.

☞ Par exemple,
$$2x^7 - 18x^3 + 91 \underset{x \to +\infty}{\sim} 2x^7, \quad \text{ou encore} \quad -12x^3 + 2x^2 + 5x \underset{x \to 0}{\sim} 5x.$$

Proposition. 1.9.
Soient f et g deux fonctions définies sur un voisinage V de $\alpha \in \overline{\mathbb{R}}$.

i. On peut traduire un équivalent à l'aide d'une égalité au voisinage de α :

$$f(x) \underset{x\to\alpha}{\sim} g(x) \Longleftrightarrow f(x) \underset{x\to\alpha}{=} g(x) + o(g(x)).$$

☞ La réciproque indique notamment qu'on peut négliger les termes négligeables dans une somme. Ainsi :

$$f(x) + o(f(x)) \underset{x\to\alpha}{\sim} f(x).$$

ii. Deux quantités équivalentes au voisinage d'un point ont la même limite en ce point. La réciproque est fausse bien entendu.

$$f(x) \underset{x\to\alpha}{\sim} g(x) \Longrightarrow \lim_{x\to\alpha} f(x) = \lim_{x\to\alpha} g(x).$$

Exercice 1.6.
Utiliser des équivalents pour déterminer les limites suivantes.

$$(i) \lim_{x\to+\infty} e^x - x^3 - \ln(x); \qquad (ii) \lim_{x\to+\infty} \frac{x^2 - \ln(x)}{\sqrt{e^x + x^3}}; \qquad (iii) \lim_{x\to+\infty} \frac{\ln(1+x^3)}{\ln(x)}.$$

Remarque 1.2.
Il existe un certain nombre d'erreurs classiques sur les équivalents, auxquelles il faut prendre garde.

i. **On ne peut pas sommer les équivalents**:

$$f(x) \underset{x\to\alpha}{\sim} g(x) \text{ n'implique pas } (f+h)(x) \underset{x\to\alpha}{\sim} (g+h)(x).$$

On a par exemple $-x^3 + x^2 \underset{+\infty}{\sim} -x^3 + x$, et pourtant on n'a pas $x^2 \underset{+\infty}{\sim} x$!

ii. **On ne peut pas élever une équivalence à une puissance dépendant de** x:

$$f(x) \underset{x\to\alpha}{\sim} g(x) \text{ n'implique pas } f(x)^{h(x)} \underset{x\to\alpha}{\sim} g(x)^{h(x)}.$$

On a par exemple $2^{1/x} \underset{+\infty}{\sim} 1$, et pourtant $\left(2^{1/x}\right)^x \underset{+\infty}{\sim} 1^x$ est inexact.

iii. **On ne peut pas composer à gauche des équivalents**:

$$f \sim g \text{ n'implique pas } h \circ f \sim h \circ g.$$

On a par exemple $x^2 + x \underset{+\infty}{\sim} x^2$, mais par contre on n'a pas $e^{x^2+x} \underset{+\infty}{\sim} e^{x^2}$.

iv. On peut toutefois effectuer des produits d'équivalents, ou élever une équivalence à une puissance constante.

Exercice 1.7.
Montrer que : $\exp\left(\left(2 - \dfrac{1}{x}\right)\ln(x)\right) - x^2 \underset{x\to+\infty}{\sim} -x\ln(x)$

1.2 Continuité d'une fonction à valeurs réelles

Définition 1.7. **Continuité en x_0**
Soient $f : I \to \mathbb{R}$ et $x_0 \in I$. On dit que :

✗ f est **continue** en x_0 si $\lim_{x\to x_0} f(x) = f(x_0)$, c'est-à-dire:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in I,\quad (|x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \varepsilon)\,.$$

✗ f est **continue à droite** (resp. **à gauche**) de x_0 si $\lim_{x\to x_0^+} f(x) = f(x_0)$ (resp. $\lim_{x\to x_0^-} f(x) = f(x_0)$).

Exemple 1.3.

La fonction définie par $f(x) = \begin{cases} x, & \text{si } x \geqslant 0 \\ 1, & \text{si } x < 0 \end{cases}$ n'est pas continue en 0.

Théorème 1.10. **Caractérisation séquentielle de la continuité**

Soient $f : I \to \mathbb{R}$ et $x_0 \in I$.
Les assertions suivantes sont équivalentes.

i. f est continue en x_0;
ii. Pour toute suite numérique $(u_n)_{n\geq 0}$ (dont les termes appartiennent à un *voisinage* de x_0 à partir d'un certain rang) qui converge vers x_0, on a que la suite $(f(u_n))_{n\geqslant 0}$ converge vers $f(x_0)$.

Définition 1.8. **Continuité sur un intervalle**

Soient $f : I \to \mathbb{R}$. On dit que f **est continue sur** I si elle est continue en chacun des réels $x_0 \in I$.

Remarque 1.3.

Il a été dit en Terminale que les fonctions continues sont celles que l'on peut "tracer sans lever le stylo". Cette notion est informelle et ne peut donc pas être utilisée dans une démonstration, mais il peut être utile de la garder à l'esprit puisqu'elle permet d'appréhender la majorité des exemples rencontrés au concours.

Exercice 1.8.

Montrer que la fonction f définie sur $\mathbb{R}$ par $f(x) = \begin{cases} x \ln x, & \text{si } x > 0 \\ 0, & \text{sinon} \end{cases}$ est continue sur $\mathbb{R}$.

☞ Il est possible de rencontrer des fonctions qui ne sont continues nulle part (voir par exemple l'**Exercice ??**).

Théorème 1.11. **Théorème des valeurs intermédiaires, $1^{\text{ière}}$ version**

Soient $f : I \to \mathbb{R}$ une fonction continue, et $a \leqslant b$ deux réels de I. Si $f(a)f(b) \leqslant 0$, alors il existe $c \in [a, b]$ tel que $f(c) = 0$.

Exemple 1.4.

Considérons une fonction polynomiale de la forme $f(x) = x^3 + ax^2 + bx + c$.
Puisque $f(x) \to +\infty$, $x \to +\infty$, on en déduit que f est à valeurs positives pour x supérieur à un réel A donné.
De même, comme $f(x) \to -\infty$, $x \to -\infty$, il vient que f est à valeurs négatives pour x inférieur à un réel A' donné.
On déduit du théorème des valeurs intermédiaires (appliqué à f sur $[A', A]$) que f s'annule en au moins un réel $x \in [A', A]$.

☞ On a ainsi démontré que toute fonction polynomiale de degré 3 possède une racine réelle (ce résultat se généralise à tous les polynômes de degré impair).

Exercice 1.9.

Montrer que toute fonction continue $f : [a, b] \longrightarrow [a, b]$ admet au moins un point fixe.

Théorème 1.12. **Théorème des valeurs intermédiaires, 2^{nde} version**

Soit $f : I \to \mathbb{R}$ une fonction continue, et $a \leqslant b$ deux réels de I. Alors pour tout y compris entre $f(a)$ et $f(b)$, il existe $c \in [a, b]$ tel que $f(c) = y$.

☞ Attention, le théorème des valeurs intermédiaires (TVI pour les intimes) ne permet pas de compter le nombre d'antécédents d'une valeurs de l'intervalle image. Pour cela, on utilise plutôt le corollaire suivant.

Théorème 1.13. **Théorème de la bijection monotone**

Soit $f : I \to \mathbb{R}$ une fonction strictement monotone et continue. Alors f réalise une bijection de I sur $f(I)$, sa bijection réciproque $f^{-1} : f(I) \to I$ est continue sur $f(I)$, et elle est de même monotonie que f.

Exercice 1.10.

Pour tout $n \in \mathbb{N}^*$, on considère l'équation

$$(E_n) \qquad e^x - x = n.$$

1. Montrer que, pour tout $n \in \mathbb{N}^*$, (E_n) admet exactement deux solutions réelles α_n et β_n telles que $\alpha_n < 0 < \beta_n$.
2. Montrer que $(\beta_n)_{n \geqslant 1}$ est croissante et déterminer sa limite.

On rappelle qu'un segment est un intervalle fermé et borné de la forme $[a, b]$. Le résultat suivant indique que toute fonction continue sur un segment y possède une valeur maximale et une valeur minimale, ou autrement dit que **toute fonction continue sur un segment est bornée et atteint ses bornes**.

Théorème 1.14. **Théorème des bornes atteintes**

Soient $a \leqslant b$ des réels, et $f : [a, b] \to \mathbb{R}$ une fonction continue. Il existe alors $(\alpha, \beta) \in [a, b]^2$ tels que

$$\forall x \in [a, b],\ f(\alpha) \leqslant f(x) \leqslant f(\beta).$$

Plus précisément,

$$f(\alpha) = \inf_{t \in [a,b]} f(t) = \min_{t \in [a,b]} f(t), \qquad f(\beta) = \sup_{t \in [a,b]} f(t) = \max_{t \in [a,b]} f(t).$$

☞ Ce résultat n'est plus vrai si l'on considère une fonction continue sur un intervalle quelconque. Prenons par exemple la fonction définie par $f(x) = x$ sur l'intervalle $]0, 1]$. On a $\inf_{x \in]0,1]} f(x) = 0$, et pourtant f ne prend jamais la valeur 0.

Exercice 1.11.

Soit f une fonction continue sur $[0, 1]$. Déterminer $\displaystyle\lim_{n \to +\infty} \int_0^1 x^n f(x) \mathrm{d}x$.

1.3 Dérivabilité d'une fonction à valeur réelle

Définition 1.9. **Taux d'accroissement**

Soient $f : I \to \mathbb{R}$ une fonction et $x_0 \in I$. On appelle **taux d'accroissement** de f en x_0, et l'on note Δ_{f,x_0}, la fonction définie pour tout $x \in I \setminus \{x_0\}$ par:

$$\Delta_{f,x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Définition 1.10. **Dérivabilité en x_0**

Soient $f : I \to \mathbb{R}$ une fonction et $x_0 \in I$.

✗ On dit que f est **dérivable en** x_0 si son taux d'accroissement admet une limite finie en x_0; on note alors $f'(x_0)$ cette limite.

✗ On dit que f est **dérivable à droite** (resp. **à gauche**) de x_0 si son taux d'accroissement admet une limite finie en x_0^+ (resp. x_0^-).

Théorème 1.15.

Soient $f : I \to \mathbb{R}$ une fonction et $x_0 \in I$. On a alors

$$f \text{ est dérivable en } x_0 \iff \begin{cases} f \text{ est dérivable à gauche et à droite en } x_0 \\ \textbf{et } \lim_{x \to x_0^-} \Delta_{f,x_0}(x) = \lim_{x \to x_0^+} \Delta_{f,x_0}(x) \end{cases}$$

Proposition. 1.16.

Soient $x_0 \in \mathbb{R}$ et f une fonction définie sur un voisinage V de x_0. Alors

$$f \text{ dérivable en } x_0 \implies f \text{ continue en } x_0.$$

Preuve. On peut écrire (pour $x \neq x_0$ dans un voisinage de x_0)

$$f(x) = f(x_0) + (x - x_0)\Delta_{f,x_0}(x).$$

Il est alors immédiat que si f dérivable en x_0, $\Delta_{f,x_0}(x)$ admet une limite finie lorsque $x \to x_0$ et donc $f(x) \to f(x_0)$, ce qui signifie que f est continue en x_0. □

La réciproque est bien évidemment fausse !

Exercice 1.12.
La fonction f de l'**Exercice 3** est-elle dérivable en 0 ?

Proposition. 1.17.
Soit $f : I \to \mathbb{R}$ une fonction dérivable en x_0. On appelle **tangente à la courbe représentative** de f au point d'abscisse a la droite passant par le point de coordonnées $(x_0, f(x_0))$ et de coefficient directeur $f'(a)$. Cette tangente admet donc pour équation :

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

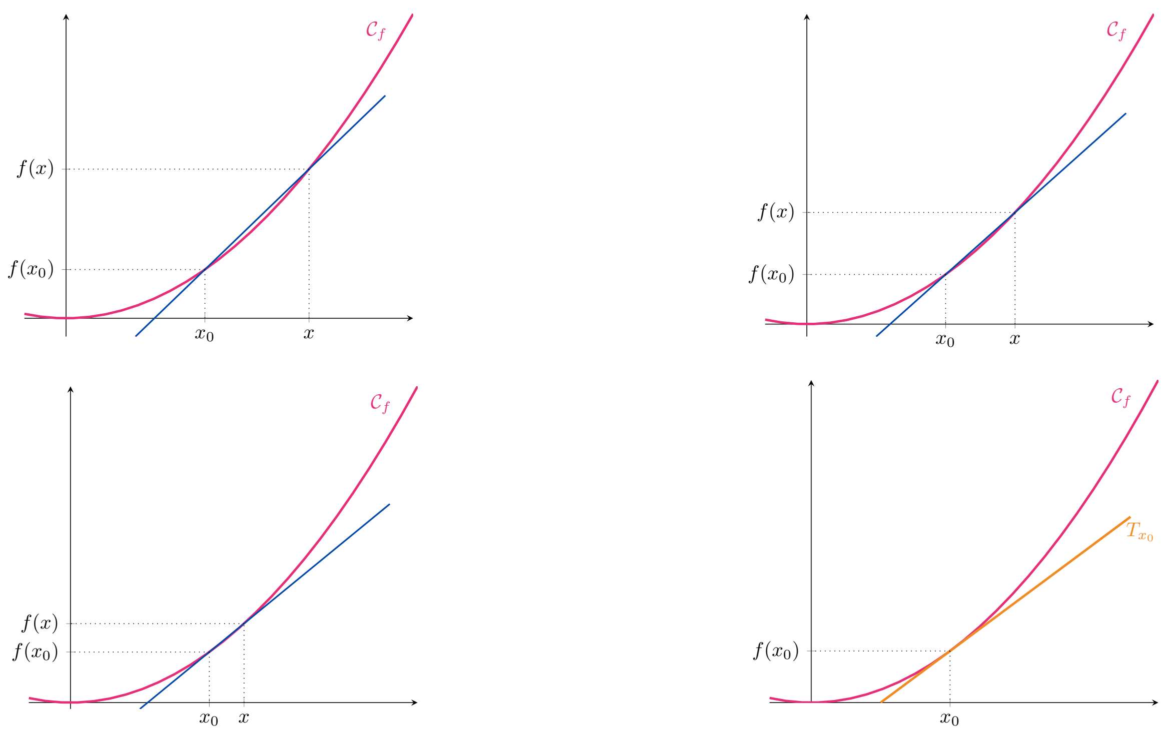


Illustration. La tangente en x_0 comme "limite" des cordes.

Proposition. 1.18. **Dérivée d'une composée**
Soit J un intervalle véritable de $\mathbb{R}$, $f : I \to J$ et $g : J \to \mathbb{R}$ deux fonctions et $x_0 \in I$.
On suppose f dérivable en x_0 et g dérivable en $f(x_0)$. Alors $g \circ f$ est dérivable en x_0, et on a :

$$(g \circ f)'(x_0) = g'(f(x_0)) \times f'(x_0).$$

Proposition. 1.19. **Dérivée de la bijection réciproque**
Soient J un intervalle véritable de $\mathbb{R}$, $f : I \to J$ une fonction bijective, $x_0 \in I$ et $y_0 = f(x_0) \in J$.
On suppose de plus que f est dérivable en x_0. Si $f'(x_0) \neq 0$, alors f^{-1} est dérivable en y_0, de sorte que

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Définition 1.11. **Extremum local**

Soient $f: I \to \mathbb{R}$ et x_0 un point intérieur à I. On dit que f présente un **extremum local** en x_0 s'il existe un voisinage $V \in V((x_0)$ tel que

$$\forall x \in V(x_0), \quad f(x) - f(x_0) \text{ est de signe constant .}$$

Lorsque ce signe est positif (resp. négatif), on parle de **minimum local** (resp. **maximum local**).

Proposition. 1.20. **Extremum local et dérivée : condition nécessaire**

Soient $f: I \to \mathbb{R}$ et x_0 un point intérieur à I où f est dérivable.
Si f admet un extremum (local) en x_0, alors nécessairement $f'(x_0) = 0$.

Preuve. Considérons un voisinage V de x_0 sur lequel $f(x) - f(x_0)$ est constant. Alors, pour $x \in V$, Δ_{f,x_0} est de signe opposé lorsque $x < x_0$ et $x > x_0$. La limite, qui est égale à $f'(x_0)$ par hypothèse, est alors à la fois positive (ou nulle) et négative (ou nulle). Elle est donc nulle. □

1.3.1 Propriétés globales des fonctions dérivables

Dans cette section, les lettres a et b désigneront deux réels tels que $a < b$.

Théorème 1.21. **Théorème de Rolle**

Soit $f : [a, b] \to \mathbb{R}$ une fonction telle que:

i. f est continue sur $[a, b]$;
ii. f est dérivable sur $]a, b[$;
iii. $f(a) = f(b)$.

Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.

Exercice 1.13.

Soit f une fonction dérivable admettant $n \geqslant 2$ racines distinctes. Alors f' admet au moins $n - 1$ racines distinctes.

Théorème 1.22. **Égalité des accroissements finis**

Soit $f : [a, b] \to \mathbb{R}$ une fonction. On suppose que :

i. f est continue sur $[a, b]$;
ii. f est dérivable sur $]a, b[$.

Alors, il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c) \times (b - a)$.

☞ Il n'existe pas de méthode simple et générale permettant de calculer la valeur de c.
☞ Il n'y a pas unicité du point c.

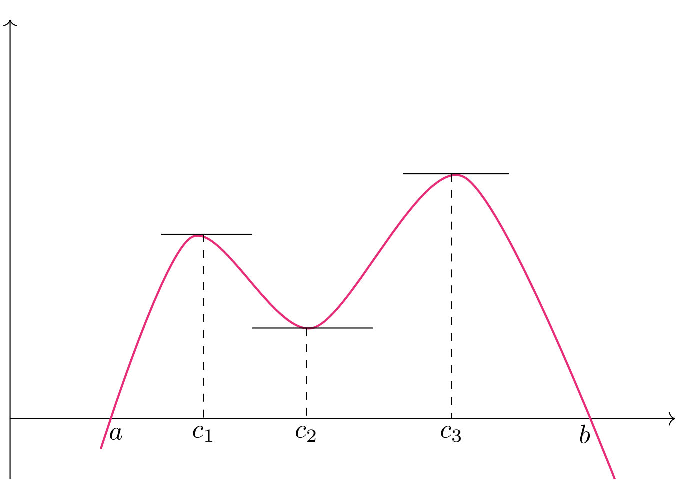


Illustration du Théorème de Rolle.

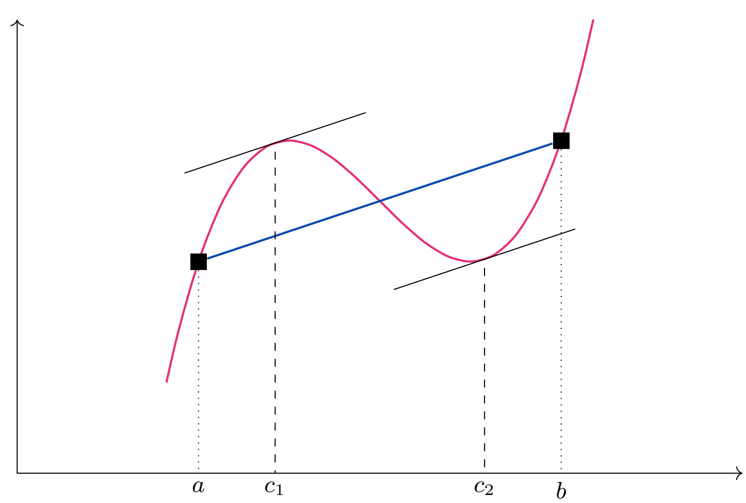


Illustration de l'égalité des accroissements finis.

Théorème 1.23. **Inégalité des accroissements finis**

Soient $f : [a, b] \to \mathbb{R}$ une fonction et m, M deux réels. On suppose que :

i. f est continue sur $[a, b]$;
ii. f est dérivable sur $]a, b[$;
iii. $\forall x \in]a, b[, m \leqslant f'(x) \leqslant M$.

Alors on a:

$$m(b-a) \leqslant f(b) - f(a) \leqslant M(b-a).$$

Le résultat suivant indique que, sous certaines conditions, la connaissance de $\lim_{x \to x_0} f'(x)$ permet de déterminer la valeur de $\lim_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0}$.

Théorème 1.24. **Théorème de la limite de la dérivée**

Soient $f : I \to \mathbb{R}$ une fonction, $x_0 \in I$ et $\ell_2 \in \overline{\mathbb{R}}$. On suppose que :

i. f est continue sur I;
ii. f est dérivable sur $I \setminus \{x_0\}$;
iii. $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} f'(x) = \ell_2$.

On a alors :

1. Si $\ell_2 \in \mathbb{R}$, alors f est dérivable en x_0 de sorte que $f'(x_0) = \ell_2$;
2. Si $\ell_2 = \pm\infty$, alors f n'est pas dérivable en x_0.

Définition 1.12. **Fonction de classe $\mathcal{C}^1$**

Soit $f : I \to \mathbb{R}$. On dit que f est **de classe** $\mathcal{C}^1$ sur I si f est dérivable sur I et f' est continue sur I.

Théorème 1.25. **Théorème de prolongement $\mathcal{C}^1$**

Soient $x_0 \in I$, $f : I \setminus \{x_0\} \to \mathbb{R}$ une fonction et ℓ_1, ℓ_2 des réels. On suppose que :

i. f est de classe $\mathcal{C}^1$ sur $I \setminus \{x_0\}$;
ii. $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} f(x) = \ell_1$;
iii. $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} f'(x) = \ell_2$.

Alors, f se prolonge en une fonction $\mathcal{C}^1$ sur I, et plus précisément:

1. f se prolonge par continuité en x_0 en une fonction $\tilde{f}$;
2. $\tilde{f}$ est de classe $\mathcal{C}^1$ sur I ;
3. $\tilde{f}(x_0) = \ell_1$ et $\tilde{f}'(x_0) = \ell_2$.

Exercice 1.14.

Montrer que la fonction f définie par $f(x) = \begin{cases} x^2, & \text{si } x > 0 \\ x^3, & \text{si } x < 0 \end{cases}$ est prolongeable sur $\mathbb{R}$ en une fonction de classe $\mathcal{C}^1$, et que la fonction g définie par $g(x) = \begin{cases} x^2, & \text{si } x > 0 \\ x, & \text{si } x < 0 \end{cases}$ ne l'est pas.

1.3.2 Dérivées successives

Définition 1.13. **Fonctions de classe $\mathcal{C}^k$**

Soit $f : I \to \mathbb{R}$.

- ✗ On dit que f **est de classe** $\mathcal{C}^0$ sur I si f est continue sur I.
- ✗ On dit que f **est de classe** $\mathcal{C}^1$ sur I si f est dérivable sur I et f' est continue sur I.
- ✗ Soit $n \in \mathbb{N}^*$. On dit que f est de classe $\mathcal{C}^n$ sur I si f est dérivable sur I et f' est de classe $\mathcal{C}^{n-1}$ sur I.
- ✗ On dit que f est de classe $\mathcal{C}^\infty$ sur I si f est de classe $\mathcal{C}^n$, pour tout $n \in \mathbb{N}$.

On note :

- ✗ $\mathcal{C}^n(I)$ l'ensemble des fonctions de classe $\mathcal{C}^n$ sur I et à valeurs réelles.
- ✗ $\mathcal{C}^\infty(I)$ l'ensemble des fonctions de classe $\mathcal{C}^\infty$ sur I et à valeurs réelles.
- ✗ $f^{(k)}$ la **dérivée** k-ème de f. Par convention, on a donc $f^{(0)} = f$.

☞ Il existe des fonctions dérivables en tout point et qui ne sont pas de classe $\mathcal{C}^1$.

Exemple 1.5.

La fonction définie par $f(x) = \begin{cases} x^2 \sin\left(\dfrac{1}{x}\right), & \text{si } x \neq 0 \\ 0, & \text{si } x = 0 \end{cases}$ est dérivable mais pas $\mathcal{C}^1$ en 0.

Théorème 1.26. **Formule de Leibniz**

Soient $n \in \mathbb{N}$, et $f, g : I \to \mathbb{R}$ deux fonctions de classe $\mathcal{C}^n$. Alors $f \times g$ est de classe $\mathcal{C}^n$, de sorte que

$$(f \times g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} \times g^{(n-k)}.$$

Remarque 1.4.

La formule ci-dessus peut sembler, dans son écriture, proche de celle du binôme de Newton. Mais d'une part elle ne fait pas apparaître des puissances mais des dérivées, et d'autre part son sens est très différent, puisqu'elle permet de calculer la dérivée $n-$ième d'un produit (et non de développer la puissance $n-$ième d'une somme).

Exercice 1.15.

Soit $n \in \mathbb{N}^*$. Calculer la dérivée $n-$ième de $f : x \mapsto (x^2+1)e^{2x}$.

Exercice 1.16. **Polynômes de Legendre, Extrait Mathématiques A 2025**

Pour $n \in \mathbb{N}$, on pose

$$U_n(X) = \left(X^2 - 1\right)^n \qquad \text{et} \qquad P_n(X) = U_n^{(n)}(X).$$

On rappelle que si $P \in \mathbb{R}[X]$ et $n \in \mathbb{N}$, on note $P^{(n)}$ la dérivée n-ème du polynôme P. On remarque que $U_0(X) = P_0(X) = 1$.

1. Calculer P_1 et P_2.
2. Déterminer le degré de U_n puis celui de P_n pour $n \in \mathbb{N}$.
3. On fixe $n \in \mathbb{N}$.

a. Vérifier que :

$$\left(X^2 - 1\right) U_n' = 2nXU_n \qquad (E)$$

b. Rappeler la formule de Leibniz exprimant la dérivée p-ème d'un produit de deux polynômes, pour $p \in \mathbb{N}$.

c. En dérivant $n+1$ fois l'égalité (E), montrer que :

$$\varphi\left(P_n\right) = n(n+1)P_n.$$

On pourra utiliser cette égalité dans les parties suivantes.

1.4 Formules de Taylor pour une fonction à valeurs réelles

Les lettres a et b désigneront dans cette section deux réels tels que $a < b$. Nous allons énoncer *les formules de Taylor*, qui généralisent le théorème des accroissement finis aux fonctions plusieurs fois dérivables.
Elles permettent d'écrire toute fonction f de classe $\mathcal{C}^{n+1}$ sur $[a,b]$ sous la forme d'une somme d'un polynôme de degré n et d'un reste :

$$f(b) = f(a) + \frac{f'(a)}{1!}(b-a) + \frac{f''(a)}{2!}(b-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(b-a)^n + R_n.$$

Dans les théorèmes qui suivent, on établit certaines propriétés de ce reste.

Théorème 1.27. **Formule de Taylor avec reste intégral**

Soient $n \in \mathbb{N}$, $f : I \to \mathbb{R}$ et $x_0 \in I$. Si f est de classe $\mathcal{C}^{n+1}$ sur I, alors pour tout $x \in I$:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{f^{(n+1)}(t)}{n!}(x-t)^n \mathrm{d}t.$$

Et si on considère une fonction $f : [a,b] \to \mathbb{R}$ de classe $\mathcal{C}^{n+1}$ on a :

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \int_{a}^{b} \frac{f^{(n+1)}(t)}{n!}(b-t)^n \mathrm{d}t.$$

☞ Cette formule n'est pas exigible au concours. Elle se démontre par récurrence, par intégrations par parties successives. Par contre il faut connaitre les deux suivantes.

Le résultat suivant majore l'erreur commise lors de l'approximation de $f(b)$ par le polynôme de Taylor de f en a.
La preuve découle immédiatement de l'a formule de Taylor avec reste intégral et de l'inégalité triangulaire.

Théorème 1.28. **Inégalité de Taylor-Lagrange**

Soient $n \in \mathbb{N}$ et $f : [a,b] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$. On a :

$$\left| f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k \right| \leqslant \frac{(b-a)^{n+1}}{(n+1)!} \sup_{x\in[a,b]} \left| f^{(n+1)}(x) \right|.$$

Exercice 1.17.

Écrire l'inégalité de Taylor-Lagrange à l'ordre n pour l'exponentielle sur l'intervalle $[0,x]$.

☞ La formule suivante permet d'obtenir les développements limités.

Théorème 1.29. **Formule de Taylor-Young**

Soit $n \in \mathbb{N}$, $x_0 \in \mathbb{R}$, et f une fonction de classe $\mathcal{C}^n$ sur un voisinage V de x_0.
Alors f admet un développement limité à l'ordre n donné par:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o\left((x-x_0)^n\right), \quad x \to x_0.$$

En posant $x = x_0 + h$, il vient:

$$f(x_0+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}h^k + o\left(h^n\right), \quad h \to 0.$$

Remarque 1.5.

La formule de Taylor-Young a l'avantage d'être plus simple que les précédentes, mais l'inconvénient de ne proposer qu'une approximation locale, c'est-à-dire qu'**elle n'est pertinente que lorsque x tend vers** x_0.

Exercice 1.18.

Déterminer un $\mathrm{DL}_3(0)$ de la fonction $x \mapsto \displaystyle\int_0^x e^{-t^2/2}\mathrm{d}t$.

1.5 Fonctions à valeurs réelles usuelles

1.5.1 Les fonctions hyperboliques

Définition 1.14. **Fonctions hyperboliques**

On appelle **cosinus hyperbolique** et **sinus hyperbolique**, et l'on note (respectivement) ch et sh les fonctions définies pour tout réel x par

$$\forall x \in \mathbb{R}, \qquad \operatorname{ch}(x) = \frac{e^x + e^{-x}}{2}, \quad \text{et} \quad \operatorname{sh}(x) = \frac{e^x - e^{-x}}{2}.$$

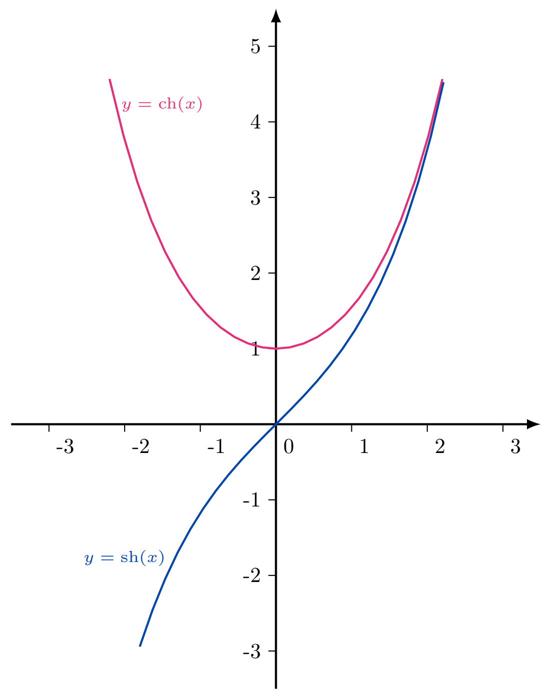
Proposition. 1.30.

On a les résultats suivants.

i. La fonction ch est paire, et sh est impaire;
ii. Pour tout réel x on a $\operatorname{ch}^2(x) - \operatorname{sh}^2(x) = 1$;
iii. Les fonctions ch et sh sont dérivables sur $\mathbb{R}$, de sorte que $\operatorname{ch}' = \operatorname{sh}$ et $\operatorname{sh}' = \operatorname{ch}$.

On peut alors tracer les tableaux de variation de ch et sh, puis en déduire leur courbe représentative.

x	$-\infty$		0		$+\infty$
ch	$+\infty$	↘	1	↗	$+\infty$
sh	$-\infty$	↗	0	↗	$+\infty$



1.5.2 Les fonctions puissances

Définition 1.15.

Soit $a \in \mathbb{R}$. Pour tout $x > 0$, on appelle x puissance a et l'on note x^a la quantité définie par

$$x^a = e^{a \ln x}.$$

✏ Pour effectuer un calcul complexe avec une quantité de la forme u^v, on revient à la définition en écrivant $u^v = e^{v \ln u}$.

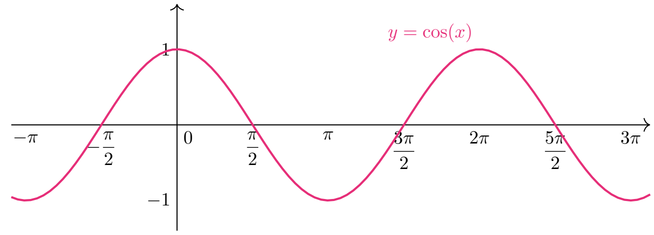
Exercice 1.19.

Calculer la dérivée de la fonction f définie par $f:\ x \mapsto x^x$.

1.5.3 Les fonctions trigonométriques et leurs réciproques

Cosinus et arccosinus.

Proposition. 1.31.
La fonction cosinus est dérivable, et sa dérivée est donnée par $(\cos)' = -\sin$.



On obtient alors le graphe ci-contre en considérant par ailleurs que la fonction cosinus est paire et 2π-périodique.

☞ La fonction cos ne réalise pas une bijection: pour tout réel $y \in [-1, 1]$ l'équation $y = \cos x$ admet une infinité de solutions (le cosinus est périodique). Néanmoins, le cosinus est strictement décroissant sur $[0, \pi]$, ce qui permettra de l'inverser sous réserve de le restreindre à $[0, \pi]$.

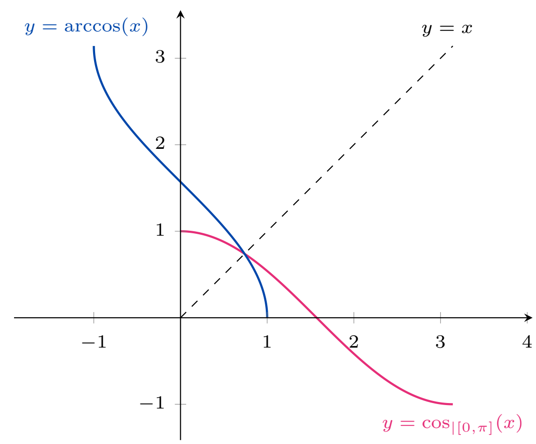
Proposition. 1.32.
La fonction $\cos_{|[0,\pi]}$ réalise une bijection strictement décroissante de $[0, \pi]$ sur $[-1, 1]$.

Définition 1.16. **Arccosinus**
On appelle **arccosinus**, et on note arccos, la bijection réciproque de la fonction $\cos_{|[0,\pi]}$.

☞ On a donc

$$\forall y \in [-1, 1],\ \cos(\arccos\ y) = y \quad \text{et} \quad \forall x \in [0, \pi],\ \arccos(\cos x) = x.$$

En d'autres termes, si $y \in [-1, 1]$, $\arccos(y)$ est l'unique $x \in [0, \pi]$ tel que $\cos x = y$.



Proposition. 1.33.
On a les résultats suivants.

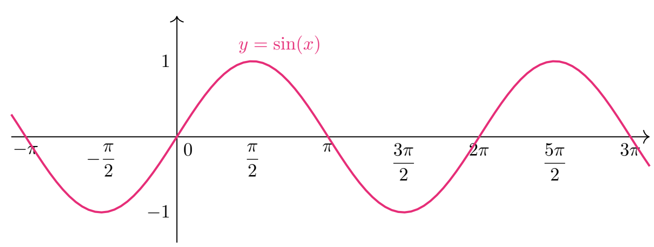
i. La fonction arccos réalise une bijection continue et strictement décroissante de $[-1, 1]$ sur $[0, \pi]$.

ii. La fonction arccos est dérivable sur $]-1, 1[$, de sorte que:

$$\forall x \in]-1, 1[,\ \arccos'(x) = -\frac{1}{\sqrt{1 - x^2}}.$$

Sinus et arcsinus.

Proposition. 1.34.
La fonction sinus est dérivable, et sa dérivée est donnée par $(\sin)' = \cos$.



On obtient alors le graphe ci-contre en considérant par ailleurs que la fonction sinus est impaire et 2π-périodique.

De même on restreint le sinus à un intervalle où il est strictement monotone pour pouvoir l'inverser.

Proposition. 1.35.
La fonction $\sin_{|[-\frac{\pi}{2},\frac{\pi}{2}]}$ réalise une bijection strictement croissante de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ sur $[-1, 1]$.

Définition 1.17. **Arcsinus**
On appelle **arcsinus**, et on note arcsin, l'inverse de la fonction $\sin_{|[-\frac{\pi}{2},\frac{\pi}{2}]}$.

☞ On a donc

$$\forall y \in [-1, 1],\ \sin(\arcsin\ y) = y \quad \text{et} \quad \forall x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],\ \arcsin(\sin x) = x.$$

En d'autres termes, si $y \in [-1, 1]$, $\arcsin(y)$ est l'unique $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ tel que $\sin x = y$.

Exercice 1.20.

Calculer $\arcsin(1)$ puis $\arcsin\left(\frac{1}{2}\right)$.

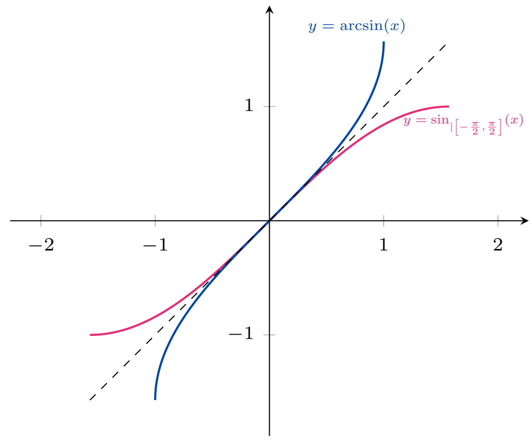
Proposition. 1.36.

On a les résultats suivants.

i. La fonction arcsin réalise une bijection continue et strictement croissante de $[-1,1]$ sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

ii. La fonction arcsin est dérivable sur $]-1,1[$, de sorte que:

$$\forall x \in]-1,1[,\ \arcsin'(x) = \frac{1}{\sqrt{1-x^2}}.$$



Tangente et arctangente.

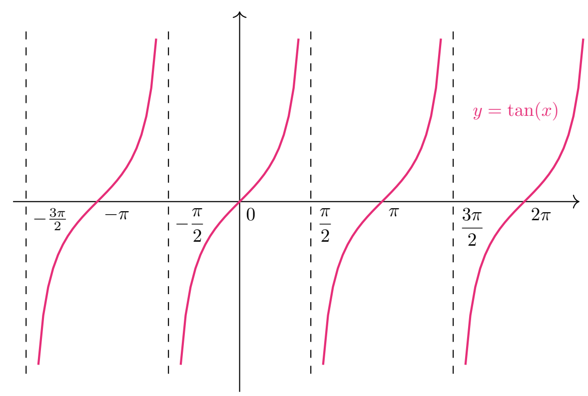
Proposition. 1.37.

La fonction tangente est définie sur

$$\mathcal{D} = \left\{x \in \mathbb{R}; \forall k \in \mathbb{Z}, x \neq \frac{\pi}{2} + k\pi\right\}.$$

Sur $\mathcal{D}$, la fonction tangente est dérivable, impaire et π-périodique, et sa dérivée est donnée par

$$(\tan)' = \frac{1}{\cos^2} = 1 + \tan^2.$$



Proposition. 1.38.

La fonction $\tan_{|]-\frac{\pi}{2},\frac{\pi}{2}[}$ réalise une bijection strictement croissante de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur $\mathbb{R}$.

Définition 1.18. **Arctangente**

On appelle **arctangente**, et on note arctan, l'inverse de la fonction $\tan_{|]-\frac{\pi}{2},\frac{\pi}{2}[}$.

☞ On a donc

$$\forall y \in \mathbb{R},\ \tan(\arctan\ y) = y \quad \text{et} \quad \forall x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[,\ \arctan(\tan x) = x.$$

En d'autres termes, si $y \in \mathbb{R}$, $\arctan(y)$ est l'unique $x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ tel que $\tan(x) = y$.

Proposition. 1.39.

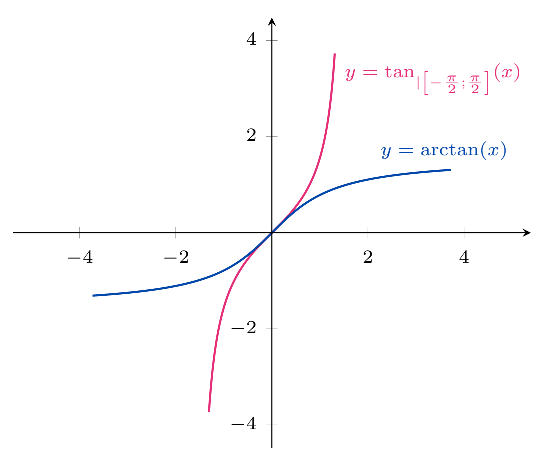
On a les résultats suivants.

i. La fonction arctan réalise une bijection continue et strictement croissante de $\mathbb{R}$ sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.

ii. La fonction arctan est dérivable, de sorte que:

$$\forall x \in \mathbb{R},\ \arctan'(x) = \frac{1}{1+x^2}.$$

iii. $\lim\limits_{x \to +\infty} \arctan(x) = \frac{\pi}{2}$ et $\lim\limits_{x \to -\infty} \arctan(x) = -\frac{\pi}{2}$.



Exercice 1.21.

Montrer que pour tout $x \in \mathbb{R}^*$, on a $\arctan x + \arctan \dfrac{1}{x} = \pm\dfrac{\pi}{2}$.

1.6 Développements limités usuels de fonctions à valeurs réelles

On rappelle quelques développements usuels, à connaître sur le bout des doigts.

Théorème 1.40. **DL usuels en 0**

Soit $\alpha \in \mathbb{R}$. Quel que soit $n \in \mathbb{N}$, on a les développements limités suivants.

$$
\begin{aligned}
e^x &\underset{x\to 0}{=} 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + o(x^n) \underset{x\to 0}{=} \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n)\\
\cos(x) &\underset{x\to 0}{=} 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n})\\
&\underset{x\to 0}{=} \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n})\\
\sin(x) &\underset{x\to 0}{=} x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o\left(x^{2n+1}\right)\\
&\underset{x\to 0}{=} \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o\left(x^{2n+1}\right)\\
\tan(x) &\underset{x\to 0}{=} x + \frac{x^3}{3} + o(x^3)\\
\arctan(x) &\underset{x\to 0}{=} x - \frac{x^3}{3} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+1})\\
\ln(1+x) &\underset{x\to 0}{=} x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n-1} \frac{x^n}{n} + o(x^n) \underset{x\to 0}{=} \sum_{k=1}^{n} (-1)^{k-1} \frac{x^k}{k} + o(x^n)\\
(1+x)^\alpha &\underset{x\to 0}{=} 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n + o(x^n) \underset{x\to 0}{=} \sum_{k=0}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k + o(x^n)\\
\frac{1}{1-x} &\underset{x\to 0}{=} 1 + x + x^2 + \cdots + x^n + o(x^n) \underset{x\to 0}{=} \sum_{k=0}^{n} x^k + o(x^n)\\
\operatorname{ch}(x) &\underset{x\to 0}{=} 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + o(x^{2n}) \underset{x\to 0}{=} \sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} + o(x^{2n})\\
\operatorname{sh}(x) &\underset{x\to 0}{=} x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+1}) \underset{x\to 0}{=} \sum_{k=0}^{n} \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+1})
\end{aligned}
$$

Remarque 1.6.

- ☞ Dans l'énoncé ci-dessus, tous les développements donnés sont à l'ordre n, sauf celui de cos qui est à l'ordre $2n$, ainsi que ceux de sin et arctan qui sont à l'ordre $2n+1$.
- ☞ On peut démontrer aisément que si une fonction paire (resp. impaire) admet un développement limité, alors celui-ci ne fait apparaitre que des termes d'ordre pair (resp. impair).

- ✏ On peut additionner, soustraire, multiplier, composer et intégrer (mais pas toujours dériver. . .) des développements limités.
- ✏ Pour calculer un développement limité d'une fonction $x \mapsto f(x)$ en $x_0 \neq 0$ on effectue un changement de variable, en posant $x = x_0 + h$, avec h qui tend vers 0. On calcule alors un DL(0) de $h \mapsto f(x_0 + h)$, puis on revient au développement recherché en posant $h = x - x_0$.

✏ Pour calculer un développement *asymptotique* d'une fonction $x \mapsto f(x)$ lorsque $x \to +\infty$ on effectue un changement de variable, en posant $t = 1/x$, avec t qui tend vers 0. On calcule alors un DL(0) de $t \mapsto f(1/t)$, puis on revient au développement recherché en revenant à x.

✏ Deux développements limités sont égaux si et seulement si ils ont les mêmes cœfficients.

Exercice 1.22.

Donner un $\mathrm{DL}_4(0)$ de $\ln(\cos(h))$.

Exercice 1.23.

Donner un développement de $\sqrt{1+x^2}$ au voisinage de $+\infty$, à la précision $\dfrac{1}{x}$. Interprétation graphique?

1.7 Fonctions d'une variable réelle à valeurs vectorielles

Dans cette dernière section, n désigne un entier naturel égal à 2 ou 3. On considère alors l'espace vectoriel $\mathbb{R}^n$ muni de sa structure euclidienne, et on fixe un repère orthonormé $\mathcal{R} = (O, e_1, ..., e_n)$.
Le produit scalaire canonique est noté $\langle \cdot | \cdot \rangle$ et la norme associée $\| \cdot \|$.
On note aussi d la distance euclidienne associée :

$$\forall\, x, y \in \mathbb{R}^n, \qquad d(x,\ y) = \|x - y\|.$$

.

1.7.1 Définitions

Définition 1.19.

Une fonction f définie sur I est dite à valeurs vectorielles s'il existe un espace vectoriel E tel que, pour tout $t \in I$, $f(t) \in E$.

☞ Si E est de dimension finie n, on peut fixer une base de celui-ci. Ainsi, on peut faire correspondre à tout fonction $f :\ I \to E$ à valeurs vectorielle une fonction

$$f :\ I \to \mathbb{R}^n$$

telle que, pour tout $t \in I$, $f(t)$ est égal aux coordonnées du vecteurs $f(t)$ dans la base choisie. Cette observation permet de se ramener à l'étude des fonctions à valeurs dans $\mathbb{R}^n$.

Définition 1.20.

Soit $f :\ I \to \mathbb{R}^n$. On appelle **fonctions coordonnées** de f les fonctions $x_1, ..., x_n$ définies sur I et à valeurs dans $\mathbb{R}$ telles que

$$\forall t \in I, \qquad f(t) = (x_1(t), ..., x_n(t)).$$

Exemple 1.6.

Notons $\vec{\mathcal{P}}$ le plan vectoriel euclidien usuel et $(\vec{i},\ \vec{j})$ sa base canonique.
La fonction $f :\ t \in [0, 2\pi[\longmapsto \cos(t)\ \vec{i} + \sin(t)\ \vec{j}$ est une fonction à valeurs vectorielles qui correspond à l'application $f :\ t \mapsto (\cos(t), \sin(t))$, à valeurs dans $\mathbb{R}^2$.

Les fonctions coordonnées sont $x_1 :\ t \mapsto \cos(t)$ et $x_2 : t \mapsto \sin(t)$.

1.7.2 Limites et continuité

Sans surprise, la notion de limite étend celle qu'on connait déjà à l'aide de la distance d donnée par la norme; on dira que $f(t)$ tend vers ℓ si la distance entre les deux vecteurs tend vers 0.

Définition 1.21.

Soient $f: I \to \mathbb{R}^n$ à valeurs vectorielles, ℓ un vecteur de $\mathbb{R}^n$ et $\alpha \in \overline{I}$.
On dit que $f(t)$ tend vers ℓ lorsque t tend vers α, ce qu'on note $\lim_{t\to\alpha} f(t) = \ell$ si

$$\lim_{t\to\alpha} d\left(f(t),\ \ell\right) = 0$$

ce qui équivaut à

$$\lim_{t\to\alpha} \|f(t) - \ell\| = 0.$$

Cette notion de limite se ré-exprime à l'aide des fonctions coordonnées. Plus précisément, on a le résultat suivant.

Proposition. 1.41.

Soient $f: I \to \mathbb{R}^n$, $x_1, ..., x_n$ ses fonctions coordonnées, ℓ un vecteur de $\mathbb{R}^n$ de coordonnées $(\ell_1, ..., \ell_n)$, et $\alpha \in \overline{I}$.
Alors,

$$\lim_{t\to\alpha} f(t) = \ell \iff \forall i \in [\![1, n]\!],\ \lim_{t\to\alpha} x_i(t) = \ell_i.$$

Preuve. Commençons par observer que, pour tout $t \in I$,

$$\|f(t) - \ell\| = \sqrt{\sum_{i=1}^{n}(x_i(t) - \ell_i)^2}.$$

Le sens $\Leftarrow$ est immédiat par continuité de la fonction carrée puis par somme et continuité de la fonction racine sur $\mathbb{R}_+$.
Réciproquement, supposons $\|f(t) - \ell\| \longrightarrow 0$. Alors, pour tout $i \in [\![1, n]\!]$,

$$0 \leq (x_i - \ell_i)^2 \leq \sum_{i=1}^{n}(x_i(t) - \ell_i)^2 \underset{t\to\alpha}{\longrightarrow} 0,$$

par hypothèse et continuité de la fonction carrée.
Le théorème des gendarmes permet de conclure. □

Maintenant qu'on dispose de la notion de limite, on peut définir la notion de continuité pour une fonction à valeurs vectorielles.

Définition 1.22.

Soient $f: \to \mathbb{R}^n$ et $t_0 \in I$. On dit que f est **continue en** t_0 si $\lim_{t\to t_0} f(t) = f(t_0)$.
Si f est continue en tout point $t_0 \in I$, on dit alors que f **est continue sur** I.

La proposition ci-dessous donne immédiatement le résultat suivant qui permet d'affirmer que la continuité d'une fonction à valeurs vectorielles est caractérisée par la continuité de ses fonctions coordonnées.

Proposition. 1.42.

Soient $f: I \to \mathbb{R}^n$, $x_1, ..., x_n$ ses fonctions coordonnées et $t_0 \in \overline{I}$.
Alors,

$$f \text{ est continue en } t_0 \iff \forall i \in [\![1, n]\!],\ x_i \text{ est continue en } t_0.$$

Les théorèmes généraux sur la continuité des fonctions d'une variable réelle s'étendent alors de manière immédiate aux fonctions à valeurs vectorielles.

Proposition. 1.43. **Théorèmes généraux**

La somme, le produit par un scalaire, le produit scalaire, le produit vectoriel (s'il est défini) de fonctions à valeurs vectorielles continues sur I est encore une fonction (à valeurs vectorielles) continue sur I.

1.7.3 Dérivabilité

Définition 1.23.
Soient $f:\ I \to \mathbb{R}^n$ et $t_0 \in I$. On dit que f est **dérivable en** t_0 si
$$\lim_{t\to t_0} \frac{1}{t-t_0}\left(f(t)-f(t_0)\right)$$
existe et est finie (et élément de $\mathbb{R}^n$). Auquel cas, cette limite sera notée
$$f'(t_0) \quad \text{ou} \quad \frac{\mathrm{d}f}{\mathrm{d}t}(t_0).$$
Si f est dérivable en tout point $t_0 \in I$, on dit alors que f **est dérivable sur** I et la fonction f' est appelée **dérivée de** f.

Comme précédemment avec la continuité, on établit une caractérisation de la dérivabilité de f sur ses fonctions coordonnées.

Proposition. 1.44.
Soient $f:\ I \to \mathbb{R}^n$, $x_1, ..., x_n$ ses fonctions coordonnées et $t_0 \in \overline{I}$.
Alors,
$$f \text{ est dérivable en } t_0 \iff \forall i \in [\![1,n]\!],\ x_i \text{ est dérivable en } t_0.$$
Lorsque c'est le cas, on a même
$$\frac{\mathrm{d}f}{\mathrm{d}t}(t_0) = (x_1'(t_0), ..., x_n'(t_0)).$$

☞ **Pour dériver une fonction à valeurs vectorielles, il suffit donc de dériver ses fonctions coordonnées.**

Preuve. Il suffit d'observer que, pour tout $t \in I \setminus \{t_0\}$,
$$\frac{1}{t-t_0}\left(f(t)-f(t_0)\right) = \frac{1}{t-t_0}\left(x_1(t)-x_1(t_0), ..., x_n(t)-x_n(t_0)\right)$$
et d'appliquer le résultat de la **Proposition 1.42.** avec la définition de dérivabilité d'une fonction à valeurs réelles. □

La dérivation coordonnée par coordonnée permet notamment d'obtenir les résultats suivants.

Proposition. 1.45. **Théorèmes généraux**
Soient $f, g:\ I \to \mathbb{R}^n$ des fonctions dérivables sur I, $\lambda\mu \in \mathbb{R}$ et $\varphi: J \to I$, où J est un intervalle (non trivial) de $\mathbb{R}$.
Alors,

i. $\lambda f + \mu g$ est encore dérivable sur I, et $(\lambda f + \mu g)' = \lambda f' + \mu g'$.

ii. $f \circ \varphi$ est dérivable sur I, et $(f\circ\varphi)' = \varphi' \times (f' \circ \varphi)$.

iii. La fonction $t \in I \mapsto \langle f(t) \mid g(t)\rangle$ est dérivable sur I et $\langle f \mid g\rangle' = \langle f' \mid g\rangle + \langle f \mid g'\rangle$.

iv. Si elle a un sens, la fonction $t \in I \mapsto \det(f(t),\ g(t))$ est dérivable sur I et $\det(f, g)' = \det(f',\ g) + \det(f,\ g')$.

v. Si elle a un sens, la fonction $t \in I \mapsto f(t) \wedge g(t)$ est dérivable sur I et $[f \wedge g]' = f' \wedge g + f \wedge g'$.

vi. Si, pour tout $t \in I$, $f(t) \neq 0$, alors $t \mapsto \|f(t)\|$ est dérivable sur I et $\|f\|' = \dfrac{\langle f' \mid f\rangle}{\|f\|}$.

1.7.4 Fonctions de classe $\mathcal{C}^k$ et formule de Taylor-Young

Définition 1.24.
Soient $f:\ I \to \mathbb{R}^n$ et $k \in \mathbb{N}$. On dit que f est de classe $\mathcal{C}^k$ (sur I) si f est dérivable k fois (sur I) et si sa dérivée $k-$ème $f^{(k)}$ est continue (sur I).
L'ensemble des fonctions de classe $\mathcal{C}^k$ sur I à valeurs dans $\mathbb{R}^n$ sera noté $\mathcal{C}^k(I,\, \mathbb{R}^n)$.
Enfin, on dit que f est de classe $\mathcal{C}^\infty$ (sur I) si elle est de classe $\mathcal{C}^k$ (sur I) pour tout entier $k \in \mathbb{N}$.
Lorsque f est de classe $\mathcal{C}^k$ (sur I), on note aussi, pour $t \in I$,

$$\frac{\mathrm{d}^k f}{\mathrm{d}t}(t) = f^{(k)}(t).$$

Naturellement, encore une fois, la fonction f sera de classe $\mathcal{C}^k$ si et seulement si toutes ses fonctions coordonnées le sont, et on obtient l'expression de $f^{(k)}$ à l'aide des dérivées $k-$ème des fonctions coordonnées de f.

$$\forall t \in I, \qquad f^{(k)}(t) = (x_1^{(k)}(t), ..., x_n^{(k)}(t)).$$

Proposition. 1.46. **Formules de Leibniz**
Soient $m \in \mathbb{N}$ et $f, g:\ I \to \mathbb{R}^n$ deux fonctions de classe $\mathcal{C}^m$ sur I. Alors, $\langle f \mid g\rangle$ et $f \wedge g$ sont encore de classe $\mathcal{C}^m$ sur I, et

$$\langle f \mid g\rangle^{(m)} = \sum_{k=0}^{m} \binom{m}{k} \left\langle f^{(k)} \mid g^{(m-k)} \right\rangle \qquad \text{et} \qquad (f \wedge g)^{(m)} = \sum_{k=0}^{m} \binom{m}{k} f^{(k)} \wedge g^{(m-k)}.$$

Si $\varphi : I \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^m$ sur I, alors $t \mapsto \varphi(t)f(t)$ est encore de classe $\mathcal{C}^m$ sur I et

$$(\varphi f)^{(m)} = \sum_{k=0}^{m} \binom{m}{k} \varphi^{(k)} f^{(m-k)}.$$

Preuve. La démonstration est admise mais le résultat se montre par récurrence. □

Théorème 1.47. **Formule du Taylor-Young**
Soient $m \in \mathbb{N}$, f une fonction de classe $\mathcal{C}^m$ sur I et $t_0 \in I$. Alors, pour tout h dans un voisinage de 0, on peut écrire

$$f(t_0 + h) \underset{h\to 0}{=} \sum_{k=0}^{m} \frac{f^{(k)}(t_0)}{k!} h^k + o(h^m),$$

où $o(h^m)$ désigne un vecteur dont toutes les coordonnées sont négligeables devant h^m.

Preuve. Il suffit d'appliquer la formule de Taylor-Young à chaque fonction coordonnées de f. □

☞ **On obtient les développements limités des fonctions à valeurs vectorielles en calculant les développements limités coordonnée par coordonnée.**

Exercice 1.24.
Déterminer le développement limité de la fonction

$$f:\ t \in]-\pi; \pi[\longmapsto (\cos(t), \sin(t)).$$

1.7.5 Interprétation cinématique

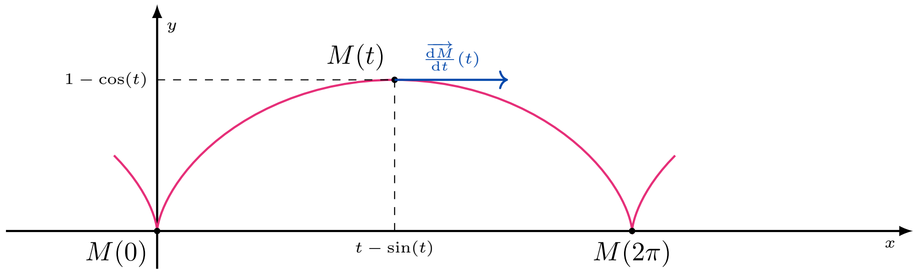
Lorsque l'on étudie le déplacement d'un mobile dans le plan, sa position en fonction du temps est donnée par une fonction $M:\ \mathbb{R}_+ \to \mathbb{R}^2$ qui sera dite à **valeurs ponctuelles**.
Les fonctions coordonnées x et y de M donnent, à l'instant t, les coordonnées du point $M(t)$ où se trouve le mobile.
Le vecteur $M'(t) = (x'(t), y'(t))$ est alors appelé *vecteur vitesse* du mobile à l'instant t; il s'agit d'une fonction à *valeurs vectorielles*.

On notera parfois $\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t)$ la dérivée de M à l'instant t. Le graphe (ou **courbe**) de M représente alors la *trajectoire* du mobile.

La figure ci-dessous représente la trajectoire d'un mobile donnée par la fonction

$$M : t \mapsto (t - \sin(t),\ 1 - \cos(t))\,.$$

Le vecteur vitesse est donné, pour tout $t \in \mathbb{R}_+$ par : $\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) = (1 - \cos(t),\ \sin(t))$.



2 Espaces vectoriels et applications linéaires

Dans tout ce chapitre, la lettre $\mathbb{K}$ désignera l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$.

2.1 Généralités et exemples

Définition 2.1. **Notion de $\mathbb{K}-$espace vectoriel**

On appelle $\mathbb{K}$**-espace vectoriel**, tout triplet $(E, +, \cdot)$ composé d'un ensemble E et de deux applications:

- l'une appelée *loi de composition interne* et notée $+$, définie par $\left\{ \begin{array}{ccc} E \times E & \to & E \\ (u, v) & \mapsto & u + v \end{array} \right.$
- l'autre appelée *loi de composition externe* et notée $\cdot$, définie par $\left\{ \begin{array}{ccc} \mathbb{K} \times E & \to & E \\ (\lambda, u) & \mapsto & \lambda \cdot u \end{array} \right.$

qui vérifient les propriétés suivantes quels que soient les vecteurs $(u, v, w) \in E^3$ et les scalaires $(\lambda, \mu) \in \mathbb{K}^2$:

i. La somme est associative: $(u + v) + w = u + (v + w)$.
ii. La somme est commutative: $u + v = v + u$.
iii. Il existe un vecteur de E, appelé *vecteur nul* et noté 0_E, qui vérifie : $u + 0_E = u$.
iv. Tout vecteur $u \in E$ possède un *opposé* $u' \in E$ qui vérifie : $u + u' = 0_E$.
L'opposé de u sera noté $-u$.

v. Le produit par un scalaire est distributif sur la somme de vecteurs: $\lambda \cdot (u+v) = \lambda \cdot u + \lambda \cdot v$.
vi. Le produit par un scalaire est distributif sur la somme de scalaires: $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$.
vii. Le scalaire $1 \in \mathbb{K}$ est neutre: $1 \cdot u = u$.
viii. Le produit par un scalaire est associatif sur les vecteurs: $(\lambda\mu) \cdot u = \lambda \cdot (\mu \cdot u)$.

☞ Un espace vectoriel est un ensemble muni d'une *structure algébrique* qui permet de faire du calcul et plus particulièrement des *combinaisons linéaires* d'éléments de l'espace.
On présente ci-après quelques espaces vectoriels de référence que nous rencontrerons très souvent.

Exemple 2.1.

- ✗ En géométrie, l'ensemble $\overrightarrow{\mathcal{P}}$ des vecteurs du plan (resp. l'ensemble $\overrightarrow{\mathcal{E}}$ des vecteurs de l'espace) muni des opérations naturelles $+$ et $\cdot$ forme un $\mathbb{R}$-espace vectoriel.
- ✗ $(\mathbb{K}, +, \times)$ est un $\mathbb{K}$-espace vectoriel.
- ✗ $(\{0_{\mathbb{K}}\}, +, \times)$ est un $\mathbb{K}$-espace vectoriel.
- ✗ Étant donné deux entiers $n, p > 0$, l'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ des matrices de taille $n \times p$ à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel, dont le vecteur nul est la matrice nulle de taille $n \times p$.
- ✗ Pour tout $n \in \mathbb{N}^*$, $\mathbb{K}^n$ muni des opérations naturelles est un $\mathbb{K}$-espace vectoriel, dont le vecteur nul est $0_{\mathbb{K}^n} = (0, \ldots, 0)$.
- ✗ L'ensemble $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel dont le vecteur nul est le polynôme nul. Il en est de même pour l'ensemble $\mathbb{K}_n[X]$ des polynômes à coefficients dans $\mathbb{K}$ de degré inférieur ou égal à n (pour $n \in \mathbb{N}$).
- ✗ L'ensemble $\mathbb{K}^{\mathbb{N}}$ des suites de scalaires peut être muni d'une structure de $\mathbb{K}$-espace vectoriel, dont le vecteur nul est la suite dont tous les termes sont nuls.
- ✗ L'ensemble des fonctions définies sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel. Le vecteur nul associé est alors la fonction $\left\{ \begin{array}{ccc} I & \to & \mathbb{R} \\ x & \mapsto & 0 \end{array} \right.$ notée $\underline{0}$.

- ✗ Soit $n \in \mathbb{N}$. L'ensemble des fonctions de classe $\mathcal{C}^n$ sur un intervalle véritable I et à valeurs dans $\mathbb{R}$ (noté $\mathcal{C}^n(I, \mathbb{R})$) est un $\mathbb{R}$-espace vectoriel. Le vecteur nul associé est alors la fonction $\left\{ \begin{array}{ccc} I & \to & \mathbb{R} \\ x & \mapsto & 0 \end{array} \right.$ notée $\underline{0}$.
- ✗ Soit $(E, +, \cdot)$ un $\mathbb{K}$-espace vectoriel et X un ensemble quelconque. Alors E^X, l'ensemble des applications de X dans E peut être muni d'une structure de $\mathbb{K}$-espace vectoriel, dont le vecteur nul est l'application nulle de X dans E.

☞ Par abus de notation et pour alléger celles-ci, on note parfois le vecteur nul 0 sans préciser en indice l'espace vectoriel dont il est l'élément nul, ce qui peut, si on travaille sans trop de discernement prêter parfois à confusion avec le 0 de l'ensemble $\mathbb{K}$. On accordera une attention tout particulière aux objets qu'on manipule.

Dans la suite de ce chapitre, la lettre E désignera (sauf mention contraire) un $\mathbb{K}$-espace vectoriel.

Définition 2.2. **Combinaison linéaire**

Soient $u_1, \dots, u_p$ des vecteurs de E. On dit qu'un vecteur $v \in E$ est une **combinaison linéaire** des vecteurs $u_1, \dots, u_p$ lorsqu'on peut trouver des scalaires $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que

$$v = \lambda_1 u_1 + \cdots + \lambda_p u_p.$$

Exercice 2.1.

Dans chaque cas, le vecteur w est-il combinaison des vecteurs u et v?

1. Dans $\mathbb{R}^3$, $w = (1, 2, 1)$, $u = (2, 1, 2)$ et $v = (1, 1, 2)$.

2. Dans $\mathcal{M}_2(\mathbb{R})$,

$$w = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}, \quad u = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix}, \quad \text{et} \quad v = \begin{pmatrix} 1 & -1 \\ -2 & 0 \end{pmatrix}$$

3. Dans $\mathbb{R}_3[X]$, $w = X^2 + 1$, $u = X^3 - X$ et $v = 2X$.

Remarque 2.1.

☞ Plutôt que dire qu'un vecteur v est combinaison linéaire d'un seul vecteur u, on dira plutôt que u et v sont **colinéaires**.

☞ Le vecteur nul est colinéaire à n'importe quel vecteur (pour tout $u \in E$, $0_E = 0 \cdot u$).

Attention, la notion de colinéarité n'a aucun sens dès lors qu'on parle d'un nombre de vecteurs strictement supérieur à 2.

Définition 2.3.

Soit $(u_1, u_2, ..., u_k)$ une famille de k vecteurs de E. L'ensemble de toutes les combinaisons linéaires des vecteurs de cette famille est noté

$$\mathrm{Vect}(u_1, u_2, ..., u_k) = \{v \in E \;:\; v = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_k u_k,\ \lambda_1, ..., \lambda_k \in \mathbb{K}\}.$$

Exercice 2.2.

Écrire les ensembles suivants sous forme d'un Vect(...).

1. $A = \{(x, y, -2x + y) \;:\; (x, y) \in \mathbb{R}^2\}$;

2. $B = \left\{ (x, y, z) \in \mathbb{R}^3 \;:\; \left\{ \begin{array}{rcl} x + y + z & = & 0 \\ x + 2y & = & 0 \end{array} \right. \right\}$

3. $C = \left\{ P \in \mathbb{C}_3[X] :\ X^2 P' - 2P = 0 \right\}$.

Proposition. 2.1. **Simplification dans un Vect(...)**

Soit $(u_1, u_2, ..., u_k)$ une famille de vecteurs de E.

i. Si $\lambda_1, \dots, \lambda_k$ sont des réels **non nuls**, alors :
$$\mathrm{Vect}(u_1, \cdots, u_k) = \mathrm{Vect}(\lambda_1 u_2, \dots, \lambda_k u_k);$$

ii. Si u_k est combinaison linéaire de $(u_1 \cdots, u_{k-1})$, alors
$$\mathrm{Vect}(u_1, \cdots, u_{k-1}, u_k) = \mathrm{Vect}(u_1, \cdots, u_{k-1});$$

☞ En particulier, si $u_k = 0_E$, alors :
$$\mathrm{Vect}(u_1, \cdots, u_{k-1}, 0_E) = \mathrm{Vect}(u_1, \cdots, u_{k-1}).$$

iii. Si $\lambda_1, \dots, \lambda_k$ sont des réels **avec $\boldsymbol{\lambda_1 \neq 0}$** alors :
$$\mathrm{Vect}(u_1, u_2, ..., u_k) = \mathrm{Vect}\left(\sum_{i=1}^{k} \lambda_i u_i, u_2, \dots, u_k\right).$$

☞ Ainsi, dans un Vect(...), on peut supprimer les vecteurs qui sont combinaisons linéaires des autres, on peut remplacer un vecteur par un multiple non nul de celui-ci, ou par une combinaison linéaire des autres vecteurs dès lors qu'on ne supprime pas la *contribution* du vecteur remplacé.

Exercice 2.3.

Montrer que, dans $\mathbb{K}_1[X]$,
$$\mathrm{Vect}(X-1, X+1, 2X+2) = \mathrm{Vect}(1, X).$$

2.2 Familles finies de vecteurs

Dans cette section toutes les familles de vecteurs considérées sont finies. Cette notion sera généralisée par la suite.

Définition 2.4.

Une sous-famille d'une famille $\mathcal{F} = (u_i)_{i\in I}$ de vecteurs de E indexée par I est une famille de la forme $\mathcal{F}' = (u_i)_{i\in J}$ où $J \subset I$. On dit aussi que $\mathcal{F}$ est une sur-famille de $\mathcal{F}'$.

2.2.1 Notion de famille génératrice (de E)

Définition 2.5. **Famille génératrice de E**

Soit $(u_i)_{1\leqslant i\leqslant p}$ une famille de vecteurs de E. Si tout vecteur $x \in E$ peut s'écrire comme une combinaison linéaire des $(u_i)_{1\leqslant i\leqslant p}$, on dit que la famille $(u_i)_{1\leqslant i\leqslant p}$ est une **famille génératrice de** E, ou que cette famille **génère** ou *engendre* E. En d'autres termes on a:
$$\begin{aligned}(u_i)_{1\leqslant i\leqslant p} \text{ est génératrice de } E \quad &\Longleftrightarrow \quad \forall x \in E,\ \exists(\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p,\ x = \sum_{i=1}^{p} \lambda_i u_i \\ &\Longleftrightarrow \quad E = \mathrm{Vect}(u_1, ..., u_p).\end{aligned}$$

Exercice 2.4.

Montrer que deux vecteurs non nuls et orthogonaux du plan $\overrightarrow{\mathbf{P}}$ forment une famille génératrice de celui-ci.

Proposition. 2.2.

Soit $(u_i)_{1\leqslant i\leqslant p}$ une famille génératrice de E. Alors,

i. La famille obtenue en changeant l'ordre des vecteurs de la famille initiale est génératrice de E;

ii. Pour tout $v \in E$, la famille $(u_1, u_2, ..., u_p, v)$ est encore génératrice de E, autrement dit, toute sur-famille d'une famille génératrice de E est encore génératrice de E.

iii. Si il existe $i \in [\![1; p]\!]$ tel que u_i est combinaison linéaire des u_j $(j \neq i)$, alors la famille $(u_1, ..., u_{i-1}, u_{i+1}, ..., u_p)$ reste génératrice de E.

☞ Le dernier point de la proposition ci-dessus permet de construire une sous-famille génératrice de cardinal *minimal* en "supprimant" les vecteurs de la famille qui sont combinaison linéaire des autres.

2.2.2 Notion de famille libre

Définition 2.6.
Soit $(u_i)_{1\leqslant i\leqslant p}$ une famille de vecteurs de E. On dit que $(u_i)_{1\leqslant i\leqslant p}$ est une **famille libre**, ou que ses vecteurs sont **linéairement indépendants** si

$$\forall(\lambda_1,\dots,\lambda_p)\in\mathbb{K}^p,\ \left[\lambda_1 u_1+\cdots+\lambda_p u_p=0_E \quad\Longrightarrow\quad \lambda_1=\cdots=\lambda_p=0_{\mathbb{K}}\right].$$

En d'autres termes, une famille libre est une famille dont toute combinaison linéaire nulle est triviale : la seule liaison possible entre les vecteurs est celle dont tous les coefficients sont nuls.
Inversement, lorsque des vecteurs ne forment pas une famille libre, on dit qu'ils sont **linéairement dépendants** ou qu'ils forment une **famille liée**.

La notion de famille liée généralise la notion de vecteurs colinéaires dès que le nombre de vecteurs est supérieur à deux.

Exercice 2.5.
Est-ce que les vecteurs $(1,0,1),(1,1,2)$ et $(2,1,2)$ forment une famille libre de $\mathbb{R}^3$?

Exemple 2.2.

- ✗ $\left(x\mapsto 1,\ \cos,\ \cos^2,\ x\mapsto\cos(2x)\right)$ est une famille liée de $\mathbb{R}^{\mathbb{R}}$.
- ✗ Pour tout $\omega\in\mathbb{R}^*$, $(x\mapsto\cos(\omega x),\ x\mapsto\sin(\omega x))$ est une famille libre de $\mathbb{R}^{\mathbb{R}}$.

Proposition. 2.3.
Une famille composée d'un unique vecteur est liée si et seulement si celui-ci est nul.

Proposition. 2.4.
Deux vecteurs sont colinéaires si et seulement si la famille qu'ils forment est liée.

Exemple 2.3.
Les vecteurs $(1,0)$ et $(0,1)$ de $\mathbb{R}^2$ sont visiblement non colinéaires: ils forment une famille libre.

Proposition. 2.5.
On a les résultats suivants :

i. Une famille de vecteurs est liée si et seulement si l'un (au moins) de ses vecteurs est combinaison linéaire des autres.
ii. Soient une famille libre $(u_i)_{1\leqslant i\leqslant p}$ de vecteurs de E, et $u_{p+1}\in E$. Alors la famille $(u_i)_{1\leqslant i\leqslant p+1}$ est liée si et seulement si u_{p+1} est combinaison linéaire de $u_1,\dots,u_p$.

Proposition. 2.6.
On a les résultats suivants.
i. Toute sous-famille d'une famille libre est libre.
ii. Toute sur-famille d'une famille liée est liée.
iii. Si $(u_i)_{1\leqslant i\leqslant p}$ est libre et si $v\notin \mathrm{Vect}((u_i)_{1\leqslant i\leqslant p})$, alors $(u_1,\dots,u_p,v)$ est encore libre.

2.2.3 Notion de base

Définition 2.7.
Soit $(u_i)_{i\in I}$ une famille de vecteurs de E. On dit que $(u_i)_{i\in I}$ est une **base** de E si c'est à la fois une famille libre et génératrice.

Proposition. 2.7. **Décomposition d'un vecteur dans une base**

Soit $(u_i)_{1\leq i\leq n}$ une famille de vecteurs de E. Alors $(u_i)_{1\leqslant i\leqslant n}$ est une base de E **si et seulement si** :

$$\forall x \in E, \exists!\ (\lambda_1,\dots,\lambda_n)\in\mathbb{K}^n,\ x=\sum_{i=1}^{n}\lambda_i u_i.$$

Définition 2.8.

Soit $\mathcal{B}=(u_i)_{1\leqslant i\leqslant n}$ une base de E et $x\in E$. On appelle **coordonnées** de x dans $(u_i)_{1\leqslant i\leqslant n}$ l'**unique** n-uplet $(\lambda_1,\dots,\lambda_n)\in\mathbb{K}^n$ tel que $x=\sum_{i=1}^{n}\lambda_i u_i$.

On note souvent le $n-$uplet de coordonnées de x dans la base $\mathcal{B}$ sous forme d'une matrice colonne :

$$x=\sum_{i=1}^{n}\lambda_i u_i \iff \mathrm{mat}_{\mathcal{B}}(x)=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}\in\mathcal{M}_{n,1}(\mathbb{K}).$$

Les bases sont donc les familles de vecteurs pour lesquelles il existe une notion de *coordonnées*. L'existence de celles-ci correspond au caractère générateur, et leur unicité à la liberté.

Exemple 2.4. **Base canonique de $\mathbb{K}^n$**

i. Soient $E=\mathbb{K}^2$, $e_1=(1,0)$ et $e_2=(0,1)$. Tout vecteur $x=(\lambda,\mu)\in\mathbb{K}^2$ s'écrit $x=\lambda(1,0)+\mu(0,1)=\lambda e_1+\mu e_2$, ce qui prouve que la famille (e_1,e_2) génère $\mathbb{K}^2$. De plus on a pour $(\lambda_1,\lambda_2)\in\mathbb{K}^2$:

$$\lambda_1 e_1+\lambda_2 e_2=0_E \iff (\lambda_1,\lambda_2)=(0,0) \iff \lambda_1=0 \quad\text{et}\quad \lambda_2=0.$$

Cela montre que la famille (e_1,e_2) est une famille libre; c'est donc une base de $\mathbb{K}^2$, appelée la **base canonique** de $\mathbb{K}^2$.

ii. Plus généralement, soient $n\in\mathbb{N}^*$ et $E=\mathbb{K}^n$. Pour tout $i\in[\![1,n]\!]$, on considère le vecteur:

$$e_i=(\underbrace{0,\dots,0}_{i-1\text{ zéros}},1,\underbrace{0,\dots,0}_{n-i\text{ zéros}}).$$

Autrement dit, on pose $e_1=(1,0\dots,0)$, $e_2=(0,1,0\dots,0),\dots,e_n=(0,\dots,0,1)$.
On montre en raisonnant comme ci-dessus que la famille $(e_i)_{1\leqslant i\leqslant n}$ est une famille libre et génératrice de $\mathbb{K}^n$; c'est donc une base de $\mathbb{K}^n$. On l'appelle la **base canonique** de $\mathbb{K}^n$.

☞ Ces bases sont dites **canoniques** dans le sens où elles apparaissent comme des bases *naturelles* de l'espace vectoriel considéré: les coordonnées d'un vecteur dans la base canonique ne sont autres que les *composantes* (ou paramètres) de ce vecteur.
On connait les bases canoniques d'autres espaces vectoriels de référence.

Exemple 2.5. **Bases canoniques - suite**

i. La base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ est la famille $\mathcal{B}=(E_{i,j})_{1\leq i\leq n,1\leq j\leq p}$ constituée de np éléments où, pour tout couple $(i,j)\in[\![1,n]\!]\times[\![1,p]\!]$,

$$\begin{array}{cc} & j\text{-ème colonne} \\ & \downarrow \\ \begin{array}{c} \\ \\ \\ i\text{-ème ligne}\rightarrow \\ \\ \\ \end{array} & \begin{pmatrix} 0 & \dots & 0 & \dots & 0\\ 0 & \dots & 0 & \dots & 0\\ \vdots & & \vdots & & \vdots\\ 0 & \dots & 1 & \dots & 0\\ \vdots & & \vdots & & \vdots\\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} \end{array} = E_{i,j}.$$

ii. La base canonique de $\mathbb{K}_n[X]$, espace vectoriel des polynômes à coefficients dans $\mathbb{K}$, de degré inférieur ou égal à n est la famille $\mathcal{B}=(1,X,X^2,\dots,X^n)$. Elle est constituée de $n+1$ vecteurs.

Exercice 2.6.
Montrer que la famille

$$\mathcal{F} = \left(\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ 1 & 0 \end{pmatrix} \right)$$

forme une base $\mathcal{M}_2(\mathbb{C})$.

2.3 Sous-espaces vectoriels

2.3.1 Définition, caractérisation et exemples

Définition 2.9. **Sous-espace vectoriel**
On appelle **sous-espace vectoriel** de E toute partie F de E telle que :
i. $0_E \in F$;
ii. $\forall (x, y) \in F^2,\ x + y \in F$ (F est stable par la loi $+$);
iii. $\forall x \in F,\ \forall \lambda \in \mathbb{K},\ \lambda \cdot x \in F$ (F est stable par la loi $\cdot$).

Proposition. 2.8.
Soit F un sous-espace vectoriel de E. Alors F, muni des lois $+$ et $\cdot$ restreintes à F est un $\mathbb{K}$-espace vectoriel.

Preuve. Comme F est stable pour $+$ et $\cdot$, celles-ci définissent des lois de composition interne et externe sur F. Il suffit alors de vérifier les huit points de la **Définition 2.1**, ce qui découle du fait que les lois considérées sont des restrictions de lois les vérifiant sur E. □

Proposition. 2.9. **Caractérisation des sous-espaces vectoriels**
Soit F un ensemble. Alors F est un sous-espace vectoriel de E si, et seulement si :

i. $F \subset E$;
ii. $0_E \in F$;
iii. $\forall (x, y) \in F^2,\ \forall (\lambda, \mu) \in \mathbb{K}^2,\ \lambda x + \mu y \in F$ (F est stable par combinaison linéaire).

☞ Cette caractérisation est très pratique, c'est d'ailleurs celle-ci qu'on vérifiera dans la plupart des cas. Il ne sera jamais nécessaire de revenir à la définition d'espace vectoriel. Pour montrer qu'un ensemble est un espace vectoriel, on montrera que c'est un sous-espace vectoriel d'un espace vectoriel de référence.

Exemple 2.6.

- ✗ Soient $\alpha_1, \ldots, \alpha_n$ des scalaires. L'ensemble $\left\{ (x_1, \ldots, x_n) \in \mathbb{K}^n ; \sum_{i=1}^{n} \alpha_i x_i = 0 \right\}$ est un sous-espace vectoriel de $\mathbb{K}^n$.

  Mais si β est un scalaire non nul, alors l'ensemble $\left\{ (x_1, \ldots, x_n) \in \mathbb{K}^n ; \sum_{i=1}^{n} \alpha_i x_i = \beta \right\}$ n'est pas un sous-espace vectoriel de $\mathbb{K}^n$, car il ne contient pas le vecteur nul.

- ✗ $\{0_E\}$ et E sont des sous-espaces vectoriels de E.

- ✗ Soit I un intervalle véritable de $\mathbb{R}$. L'ensemble $\mathbb{C}^0(I)$ des fonctions continues de I vers $\mathbb{R}$ est un sous-$\mathbb{R}$-espace vectoriel de $\mathbb{R}^I$.

- ✗ L'ensemble $\mathbb{K}_n[X]$ des polynômes de degré inférieur ou égal à n et à cœfficients dans $\mathbb{K}$ est un sous-espace vectoriel de $\mathbb{K}[X]$.

Proposition. 2.10. **Sous-espace engendré**

Soient $p \in \mathbb{N}^*$ et $u_1, \ldots, u_p$ des vecteurs de E. On a alors les résultats suivants.

i. $\operatorname{Vect}(u_1, \ldots, u_p)$ est un sous-espace vectoriel de E, appelé le **sous-espace engendré** par $u_1, \ldots, u_p$.

ii. La famille $(u_1, \ldots, u_p)$ est une famille génératrice de $\operatorname{Vect}(u_1, \ldots, u_p)$.

iii. Si la famille $(u_1, \ldots, u_p)$ est libre, alors c'est une base de $\operatorname{Vect}(u_1, \ldots, u_p)$.

Exemple 2.7.

Soit $\omega \in \mathbb{R}^*$, $\underline{0} : \mathbb{R} \to \mathbb{R}$ l'application nulle, et $\mathcal{S} = \{y \in \mathcal{C}^2(\mathbb{R}); y'' + \omega^2 y = \underline{0}\}$. Alors $\mathcal{S}$ est un $\mathbb{R}$-espace vectoriel admettant pour base la famille $(x \mapsto \cos(\omega x),\ x \mapsto \sin(\omega x))$.

Exercice 2.7.

1. Montrer que $F = \{(2x, 3y, x+y, x) \in \mathbb{K}^4 :\ (x, y) \in \mathbb{K}^2\}$ est un espace vectoriel, et en déterminer une base.

2. Montrer que $G = \left\{(x, y, z, t) \in \mathbb{K}^4 : \left\{ \begin{array}{rcl} x - y + z + t & = & 0 \\ x + z & = & 0 \end{array} \right.\right\}$ est un espace vectoriel, et en déterminer une base.

3. Montrer que $H = \{M \in \mathcal{M}_2(\mathbb{K}) :\ M^T = M\}$ est un espace vectoriel, et en déterminer une base.

2.3.2 Intersection et somme de sous-espaces vectoriels

Proposition. 2.11. **Intersection de s-ev**

Si F_1 et F_2 sont des sous-espaces vectoriels de E, alors $F_1 \cap F_2$ est un sous-espace vectoriel de E. De plus, l'intersection d'un nombre quelconque de sous-espaces vectoriels de E est un sous-espace vectoriel de E.

Remarque 2.2.

Il est erroné en revanche de penser que la réunion de deux sous-espaces vectoriels est un sous-espace vectoriel. Il suffit pour s'en convaincre de prendre par exemple deux droites vectorielles de $\mathbb{R}^2$: $\operatorname{Vect}((1,0))$ et $\operatorname{Vect}((0,1))$: la somme d'un élément de la première droite et d'un élément de la seconde n'est pas élément de la réunion qui n'est donc pas stable par somme (ni par combinaison linéaire).

Remarque 2.3.

Soit $n \in \mathbb{N}^*$. On déduit alors directement des **Exemples 2.6** que tout sous-ensemble de $\mathbb{K}^n$ défini par un système d'équations cartésiennes homogènes est un sous-espace vectoriel de $\mathbb{K}^n$.

Définition 2.10.

Soit $p \in \mathbb{N}^*$ et $(F_1, \ldots, F_p)$ une famille de sous-espaces vectoriels de E. On appelle **somme** de $F_1, \ldots, F_p$, et l'on note $F_1 + \cdots + F_p$, l'ensemble défini par:

$$F_1 + \cdots + F_p = \{u_1 + \cdots + u_p;\ \forall i \in [\![1, p]\!],\ u_i \in F_i\}.$$

Proposition. 2.12.

Soit $p \in \mathbb{N}^*$ et $(F_1, \ldots, F_p)$ une famille de sous-espaces vectoriels de E. Alors $F_1 + \cdots + F_p$ est un sous-espace vectoriel de E.

Définition 2.11. **Somme directe**

Soit $p \in \mathbb{N}^*$ et $(F_1, \ldots, F_p)$ une famille de sous-espaces vectoriels de E. On dit que $F_1, \ldots, F_p$ sont en **somme directe** si tout vecteur de $F_1 + \cdots + F_p$ s'écrit **de manière unique** comme somme de vecteurs appartenant aux F_i, c'est-à-dire:

$$F_1 + \cdots + F_p \text{ directe} \Leftrightarrow \forall u \in F_1 + \cdots + F_p, \exists!\,(u_1, \ldots, u_p) \in F_1 \times \cdots \times F_p, u = u_1 + \cdots + u_p.$$

Dans ce cas, la somme de $F_1, \ldots, F_p$ sera notée $F_1 \oplus \cdots \oplus F_p = \bigoplus_{i=1}^{p} F_i$ au lieu de $F_1 + \cdots + F_p$.

Remarque 2.4.

Soit $(e_1, \ldots, e_n)$ une base de E. On a alors $E = \bigoplus_{i=1}^{n} \operatorname{Vect}(e_i)$.

Les sommes directes de deux sous-espaces se caractérisent comme suit.

Proposition. 2.13.

Soient F_1, F_2 deux sous-espaces vectoriels de E. Alors la somme $F_1 + F_2$ est directe si et seulement si $F_1 \cap F_2 = \{0_E\}$.

Dans le cas général, on ne dispose que du résultat suivant.

Proposition. 2.14.

Une somme de sous-espaces vectoriels est directe si et seulement si le vecteur nul se décompose de façon unique dans celle-ci.

On a enfin la notion de couples de sous-espaces supplémentaires.

Définition 2.12. **Sous-espaces supplémentaires**

Soient F_1, F_2 deux sous-espaces vectoriels de E. On dit que F_1 et F_2 sont **supplémentaires** dans E si leur somme est directe et vaut E, c'est-à-dire si $E = F_1 \oplus F_2$. En d'autres termes:

$$E = F_1 \oplus F_2 \iff \forall u \in E, \exists!\,(u_1, u_2) \in F_1 \times F_2,\; u = u_1 + u_2.$$

C'est-à-dire que F_1 et F_2 sont supplémentaires dans E si et seulement si tout vecteur de E s'écrit de manière unique comme somme d'un vecteur de F_1 et d'un vecteur de F_2

Exemple 2.8.

- ✗ Soit P (resp. I) ensemble des fonctions $f : \mathbb{R} \to \mathbb{R}$ paires (resp. impaires). Alors $\mathbb{R}^{\mathbb{R}} = P \oplus I$.
- ✗ Soit (u, v) une base de $\mathbb{K}^2$. On a alors $\operatorname{Vect}(u) \oplus \operatorname{Vect}(v) = \mathbb{K}^2$ donc $\operatorname{Vect}(v)$ est **un** supplémentaire de $\operatorname{Vect}(u)$. Mais comme $(u, u+v)$ est encore une base de $\mathbb{K}^2$, $\operatorname{Vect}(u+v)$ est un autre supplémentaire de $\operatorname{Vect}(u)$. En particulier, il n'y a pas unicité du supplémentaire.
  On utilisera donc un pronom indéfini pour parler d'un supplémentaire.

2.4 Dimension d'un espace vectoriel

2.4.1 Définition et théorème fondamental

Définition 2.13.

On dit qu'un $\mathbb{K}$-espace vectoriel est de **dimension finie** lorsqu'il admet une base de cardinal fini.
Dans le cas contraire, il est dit de **dimension infinie**.

Théorème 2.15. **Dimension d'un ev**

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Alors toutes les bases de E sont finies et ont même nombre de vecteurs. On appelle alors **dimension** de E, et l'on note $\dim(E)$ (ou $\dim_{\mathbb{K}}(E)$), le nombre de vecteurs d'une base de E.

☞ Dès qu'on est en mesure d'expliciter **une** base d'un espace vectoriel, on peut en déduire sa dimension.

Exemple 2.9. **Dimension des espaces de référence**

i. L'espace vectoriel trivial $\{0\}$ est de dimension 0.
ii. $\dim(\mathbb{K}^n) = n$.
iii. $\dim(\mathcal{M}_{n,p}(\mathbb{K})) = np$.
iv. $\dim \mathbb{K}_n[X] = n + 1$.
v. Soit $\omega \in \mathbb{R}^*$. L'ensemble des solutions de $y'' + \omega^2 y = \underline{0}$ est de dimension 2.
vi. $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ et $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.

Exemple 2.10.

i. Soit $u \in E$, tel que $u \neq 0_E$. Alors $\text{Vect}\,(u) = \{\lambda u : \ \lambda \in \mathbb{K}\}$ est appelé la *droite vectorielle* de E engendrée par u ; c'est un espace vectoriel de dimension 1.

ii. Soient u et v deux vecteurs non colinéaires de E. Alors $\text{Vect}(u, v) = \{\lambda u + \mu v : \ (\lambda, \mu) \in \mathbb{K}^2\}$ est appelé le *plan vectoriel* de E engendré par u et v ; c'est un espace vectoriel de dimension 2.

Exercice 2.8.

Montrer que $\text{Vect}(\cos, \sin)$ est un espace vectoriel de dimension 2.

2.4.2 Le théorème de la base incomplète et ses conséquences

Théorème 2.16. **Théorème de la base incomplète**

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, $\mathcal{L}$ une famille libre finie de vecteurs de E et $\mathcal{G}$ une famille génératrice finie de vecteurs de E. Alors E admet une base $\mathcal{B}$ obtenue à partir de $\mathcal{L}$ en ajoutant des vecteurs de $\mathcal{G}$.

Exercice 2.9.

Donner une base de $\mathbb{K}^4$ contenant les vecteurs $(1, 1, 0, -1)$ et $(1, -1, 0, 1)$.

Corollaire 2.17.

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Alors :

i. Toute famille libre finie peut être complétée en une base finie de E.
ii. De toute famille génératrice finie, on peut extraire une sous-famille qui est une base finie de E.

☞ Le point *ii.* ci-dessus est appelé le *théorème de la base extraite* (ou base trop pleine).
Le résultat ci-dessous indique que les bases sont à la fois *les familles libres maximales* et *les familles génératrices minimales*.

Théorème 2.18. **Cardinal des familles et dimension**

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{F}$ une famille finie de vecteurs de E.

✗ *i.* Si $\mathcal{F}$ est libre, alors $\text{card}(\mathcal{F}) \leqslant n$.
ii. Si $\mathcal{F}$ est libre, alors $\mathcal{F}$ est une base si et seulement si $\text{card}(\mathcal{F}) = n$.

✗ *i.* Si $\mathcal{F}$ est génératrice, alors $\text{card}(\mathcal{F}) \geqslant n$.
ii. Si $\mathcal{F}$ est génératrice, alors $\mathcal{F}$ est une base si et seulement si $\text{card}(\mathcal{F}) = n$.

Preuve. Résulte des théorèmes de la base incomplète et de la base extraite. □

Remarque 2.5.

Par contraposition, on déduit directement de $i)$ et de $j)$:

✗ Si $\text{card}(\mathcal{F}) > n$, alors $\mathcal{F}$ est liée.
✗ Si $\text{card}(\mathcal{F}) < n$, alors $\mathcal{F}$ n'est pas génératrice.

Corollaire 2.19.

Soient E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{F}$ une famille finie de vecteurs de E. Si deux des assertions suivantes sont vérifiées :

i. $\mathcal{F}$ est libre;
ii. $\mathcal{F}$ est génératrice;
iii. $\text{card}(\mathcal{F}) = n$;

alors la troisième assertion est également vérifiée, et $\mathcal{F}$ est une base.

☞ Pour montrer qu'une famille de vecteurs forme une base d'une espace vectoriel (dont on connait **déjà** la dimension), on montre souvent qu'elle est libre et on conclut par une phrase rigoureuse expliquant que cela suit car la famille "a le bon nombre de vecteurs".

Exercice 2.10.
Soit $n \in \mathbb{N}^*$. Montrer que la famille $(1, X+1, (X+1)^2, ..., (X+1)^n)$ forme une base de $\mathbb{K}_n[X]$.

2.4.3 Espaces vectoriels de dimension infinie

Les notions de famille libre ou génératrice se généralisent aux familles infinies de vecteurs.

Définition 2.14.
Soit I un ensemble infini, $\mathcal{F} = (u_i)_{i \in I}$ une famille de vecteurs de E. On note $\mathcal{P}_{\mathrm{f}}(I)$ l'ensemble des parties finies de I.

i. On dit que la famille $\mathcal{F}$ est libre si toute sous-famille finie de $\mathcal{F}$ est libre, c'est-à-dire si :

$$\forall J \in \mathcal{P}_{\mathrm{f}}(I),\ \forall (\lambda_j)_{j \in J} \in \mathbb{K}^J,\ \left[\sum_{j \in J} \lambda_j u_j = 0_E \iff \forall j \in J, \lambda_j = 0\right].$$

ii. On dit que la famille $\mathcal{F}$ est génératrice de E si tout vecteur x de E s'écrit comme une combinaison linéaire d'un nombre **fini** d'éléments de la famille $\mathcal{F}$, c'est-à-dire si :

$$\forall x \in E,\ \exists J \in \mathcal{P}_{\mathrm{f}}(I),\ \exists (\lambda_j)_{j \in J} \in \mathbb{K}^J, \quad x = \sum_{j \in J} \lambda_j u_j.$$

iii. On dit que la famille $\mathcal{F}$ est une base de E si elle est à la fois libre et génératrice de E.

Proposition. 2.20.
La famille $(X^k)_{k \geqslant 0}$ est une base de $\mathbb{K}[X]$, appelée la *base canonique* de $\mathbb{K}[X]$.

Corollaire 2.21.
$\mathbb{K}[X]$ est un espace vectoriel de dimension infinie.

Preuve. Si tel n'était pas le cas, alors toute famille libre de $\mathbb{K}[X]$ aurait moins de $n = \dim \mathbb{K}[X]$ éléments, ce qui contredirait le théorème précédent. □

Corollaire 2.22.
Soit I un intervalle véritable de $\mathbb{R}$. Alors $\mathbb{K}^I$ est un $\mathbb{K}$-espace vectoriel de dimension infinie.

La définition suivante et le théorème associé sont souvent utiles pour montrer qu'une famille de polynômes est libre.

Définition 2.15. **Famille de polynômes échelonnée**
Soit $n \in \mathbb{N}$ et $(P_k)_{0 \leqslant k \leqslant n}$ une famille de $\mathbb{K}[X]$. On dit que $(P_k)_{0 \leqslant k \leqslant n}$ est **échelonnée** si

$$-\infty < \deg P_0 < \deg P_1 < \cdots < \deg P_n.$$

Exemple 2.11.

- ✗ La base canonique de $\mathbb{K}[X]$ est échelonnée.
- ✗ La famille $((X+1)^k)_{k \geq 0}$, de l'**Exercice 2.10.**, est échelonnée.

Proposition. 2.23. **Base et famille échelonnée en degré**
Soit $n \in \mathbb{N}$ et $(P_k)_{k \in \mathbb{N}}$ une famille de $\mathbb{K}[X]$.

i. Si $(P_k)_{0 \leqslant k \leqslant n}$ échelonnée, alors $(P_k)_{0 \leqslant k \leqslant n}$ est libre.
ii. Si $\forall k \in [\![0, n]\!], \deg P_k = k$, alors $(P_k)_{0 \leq k \leq n}$ est une base de $\mathbb{K}_n[X]$.
iii. Si $\forall k \in \mathbb{N}$, $\deg P_k = k$, alors $(P_k)_{k \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$.

Remarque 2.6.
Il est erroné de croire que toute base de $\mathbb{K}[X]$ est une famille échelonnée: $X, X+1$ forment par exemple une base de $\mathbb{K}_1[X]$.

2.4.4 Croissance de la dimension

Proposition. 2.24.

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors :

i. F est de dimension finie.
ii. $\dim(F) \leq \dim(E)$.
iii. $\dim(F) = \dim(E) \iff F = E$.

Ce résultat est analogue à un résultat sur le cardinal des ensembles finis.
☞ En pratique, pour montrer que deux sous-espaces vectoriels sont égaux, il suffit de démontrer que l'un est inclus dans l'autre, puis de vérifier qu'ils ont même dimension.

2.4.5 Dimension et somme de sous-espaces

Proposition. 2.25.

[Somme (directe) et dimension] Soient E un $\mathbb{K}$-espace vectoriel de dimension finie, F et G deux sous-espaces vectoriels de E en somme directe. On a alors

$$\dim(F \oplus G) \dim(F) + \dim(G).$$

Plus généralement, on a la formule suivante.

Proposition. 2.26. **Formule de Grassmann**

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie, F et G deux sous-espaces vectoriels de E. On a alors

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

On observe que la formule ci-dessus est analogue à $\operatorname{card}(A \cup B) = \operatorname{card}(A) + \operatorname{card}(B) - \operatorname{card}(A \cap B)$.

Définition 2.16.

Soient $\mathcal{F} = (f_i)_{i \in I}$ et $\mathcal{F}' = ({f_j}')_{j \in J}$ deux familles de vecteurs de E telles que les ensembles d'indices I et J sont disjoints (*i.e.* $I \cap J = \emptyset$). Pour $k \in I \cup J$, on pose :

$$g_k = \begin{cases} f_k, & \text{si } k \in I \\ {f_k}', & \text{si } k \in J. \end{cases}$$

On appelle alors **réunion disjointe** des familles $\mathcal{F}$ et $\mathcal{F}'$, et on note $\mathcal{F} \sqcup \mathcal{F}'$, la famille $(g_k)_{k \in I \cup J}$. Intuitivement, la réunion disjointe de familles correspond à leur *concaténation* (ou juxtaposition).

On définit de même l'union disjointe d'un nombre fini quelconque de familles.

Proposition. 2.27. **Base adaptée à la somme directe**

Soit $p \in \mathbb{N}^*$ et $(F_1, \ldots, F_p)$ une famille de sous-espaces vectoriels de E tels que $E = \displaystyle\bigoplus_{i=1}^{p} F_i$.

Alors toute union disjointe de bases de $F_1, \ldots, F_p$ est une base de E. Cette base est dite **adaptée à la décomposition**

$$E = \bigoplus_{i=1}^{p} F_i.$$

Corollaire 2.28.

Soit $p \in \mathbb{N}^*$ et $(F_1, \ldots, F_p)$ une famille de sous-espaces vectoriels de E tels que $E = \displaystyle\bigoplus_{i=1}^{p} F_i$.

Alors $\dim E = \displaystyle\sum_{i=1}^{p} \dim F_i$.

Preuve. Résulte directement des **Théorèmes 2.27** et **2.25**. □

☞ Dans le cas particulier d'une somme de deux sous-espaces, la réciproque du **Théorème 2.27** est au programme.

Théorème 2.29.

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F_1, F_2 deux sous-espaces vectoriels de E. Alors F_1 et F_2 sont supplémentaires dans E si et seulement si l'union disjointe d'une base de F_1 et d'une base de F_2 est une base de E.

Théorème 2.30.
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel de E. On a alors les résultats suivants.

i. F admet au moins un supplémentaire dans E.
ii. Tous les supplémentaires de F dans E ont la même dimension, égale à $\dim(E) - \dim(F)$.

Proposition. 2.31.
Soient E un $\mathbb{K}$-espace vectoriel de dimension finie, F et G deux sous-espaces vectoriels de E. Si deux des assertions suivantes sont vérifiées

i. $F \cap G = \{0_E\}$
ii. $F + G = E$
iii. $\dim(F) + \dim(G) = \dim E$

alors la troisième assertion est également vérifiée, et alors les sous-espaces F, G sont supplémentaires dans E.

Exercice 2.11.
Vérifier que $F = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ et $G = \mathrm{Vect}(1, 1, 1)$ sont supplémentaires dans $\mathbb{R}^3$.

2.4.6 Rang d'une famille finie de vecteurs

Définition 2.17. **Rang d'une famille finie de vecteurs**
Soit $\mathcal{F}$ une famille finie de vecteurs de E. On appelle **rang** de $\mathcal{F}$, et on note $\mathrm{rg}\,(\mathcal{F})$, la dimension du sous-espace vectoriel $\mathrm{Vect}\mathcal{F}$.
En d'autres termes, on pose
$$\mathrm{rg}\,(\mathcal{F}) = \dim\left(\mathrm{Vect}\mathcal{F}\right).$$

Exercice 2.12.
Soit (u, v) une base de $\mathbb{R}^2$. Calculer le rang de $(u, v, u + v)$.

☞ Le rang d'une famille de vecteurs est la dimension de l'espace que cette famille engendre.
La caractérisation suivante est aussi fréquemment utile que la définition.

Proposition. 2.32.
Le rang d'une famille finie de vecteurs est égal au cardinal de sa plus grande sous-famille libre.

Preuve. Posons $F = \mathrm{Vect}(u_1, \ldots, u_p)$. Puisque la famille $u_1, \ldots, u_p$ génère F, sa plus grande sous-famille libre est une base de F (par le théorème de la base extraite), et le cardinal de cette dernière est donc égal à $\dim(F) = \mathrm{rg}\,(u_1, \ldots, u_p)$. □

☞ Pour calculer le rang d'une famille finie $\mathcal{F} = (u_1, \ldots, u_p)$:

✗ on commence par regarder si elle est libre ou pas.
✗ Sinon, on retire un vecteur qui est combinaison linéaire des autres.
Si la famille obtenue n'est toujours pas libre, on en retire encore un vecteur etc...
Ce faisant, on calcule une base du sous-espace vectoriel $\mathrm{Vect}(u_1, \ldots, u_p)$ en utilisant la méthode de la base extraite.
De manière *duale*, on peut également employer la méthode de la base incomplète et, partant de la famille $\emptyset$, rajouter des vecteurs de $\mathcal{F}$ tant que cela donne encore une famille libre.

2.4.7 Notion d'hyperplan

Définition 2.18. **Hyperplan**
On appelle **hyperplan** de E tout sous-espace vectoriel de E admettant une droite pour supplémentaire.

Proposition. 2.33. **Dimension d'un hyperplan**
Soient E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$, et F un sous-espace vectoriel de E. Alors F est un hyperplan de E si, et seulement si $\dim(F) = n - 1$.

Exemple 2.12.

- ✗ Dans $\mathbb{R}^3$ les hyperplans sont des plans vectoriels, et dans $\mathbb{R}^2$ les hyperplans sont des droites vectorielles.
- ✗ Pour tout $n \in \mathbb{N}^*$, $\mathbb{K}_{n-1}[X]$ est un hyperplan de $\mathbb{K}_n[X]$.
- ✗ $F = \{P \in \mathbb{K}[X];\ P(0) = 0\}$ est un hyperplan de $\mathbb{K}[X]$.

2.5 Notion d'application linéaire

Dans la suite de ce chapitre, les lettres E, F désigneront (sauf mention contraire) des $\mathbb{K}$-espace vectoriels.

2.5.1 Définition, caractérisation et exemples

Définition 2.19. **Application linéaire, morphisme**

Soit $f : E \to F$ une application. On dit que f est une **application linéaire** de E dans F, ou un **morphisme** d'espace vectoriels entre E et F, si elle vérifie les propriétés suivantes.

i. $\forall\,(x, y) \in E^2,\ f\,(x + y) = f\,(x) + f\,(y)\,.$

ii. $\forall x \in E, \forall \lambda \in \mathbb{K},\ f\,(\lambda \cdot x) = \lambda \cdot f\,(x)\,.$

Définition 2.20. **Types de morphismes**

- ✗ On appelle **isomorphisme** de E dans F toute application linéaire bijective $f : E \to F$.
- ✗ On appelle **endomorphisme** de E toute application linéaire $f : E \to E$.
- ✗ On appelle **automorphisme** de E toute application linéaire bijective $f : E \to E$.
- ✗ On appelle **forme linéaire** de E toute application linéaire $f : E \to \mathbb{K}$.
  L'ensemble des formes linéaires sur E est noté E^*.

☞ On note $\mathcal{L}(E, F)$ l'ensemble des morphismes de E dans F, $\mathcal{L}(E)$ l'ensemble des endomorphismes de E et $\mathrm{GL}(E)$ l'ensemble des automorphismes de E.

Proposition. 2.34. **St abilité et endomorphisme induit**

Soit $f \in \mathcal{L}(E, F)$. Un sous-espace vectoriel A de E est dit **stable** par f si et seulement si $f(A) \subset A$.
Dans ce cas, l'endomorphisme $f_A : x \in A \longmapsto f(x) \in A$ est un endomorphisme de A, appelé **endomorphisme induit** par f dans A.

On a la caractérisation suivante des applications linéaires.

Proposition. 2.35.

Soit $f : E \to F$ une application. Alors $f \in \mathcal{L}(E, F)$ si et seulement si

$$\forall (u, v) \in E^2, \forall (\lambda, \mu) \in \mathbb{K}^2,\ f\,(\lambda u + \mu v) = \lambda f\,(u) + \mu f\,(v)\,.$$

Les applications linéaires entre deux espaces vectoriels E et F sont donc les applications dont l'action est compatible avec la structure d'espace vectoriel: une combinaison linéaire est transformée en une combinaison linéaire. On peut ainsi étudier ces applications en utilisant les outils de l'algèbre linéaire.

Exemple 2.13. **Exemple d'applications linéaires**

- ✗ Quel que soit l'espace vectoriel E, l'application identité $\mathrm{id}_E : E \longrightarrow E,\ x \longmapsto x$ et l'application nulle $0_{\mathcal{L}(E)} : E \longrightarrow E,\ x \mapsto 0_E$ sont des endomorphismes de E.

- ✗ Soit $\lambda \in \mathbb{K}^*$. L'application $H_k : E \longrightarrow E,\ x \mapsto \lambda \cdot x$ est un automorphisme de E, appelé l'**homothétie vectorielle** de rapport λ. On la note $\lambda \mathrm{id} E$.

- ✗ L'application de $\mathbb{K}^2$ dans $\mathbb{K}^3$ définie par $\varphi : (x, y) \mapsto (x + y, x - y, y)$ est linéaire.

- ✗ Soit I un intervalle véritable de $\mathbb{R}$. L'application $D : \mathcal{C}^{k+1}(I) \longrightarrow \mathcal{C}^k(I),\ f \longmapsto f'$ est linéaire.

✗ Soient $a < b$ des réels. L'application $\Psi : \mathcal{C}^0([a,b]) \longrightarrow \mathbb{R},\ f \longmapsto \int_a^b f(x)\mathrm{d}x$ est linéaire.

✗ L'application de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$ définie par $\varphi : P \mapsto X^2P - XP'$ est linéaire.

✗ Soit $n, p > 0$ entiers. L'application $\tau : M \mapsto M^T$ est un isomorphisme de $\mathcal{M}_{n,p}(\mathbb{K})$ vers $\mathcal{M}_{p,n}(\mathbb{K})$.

✗ L'application `shift` : $\mathbb{R}^{\mathbb{N}} \longrightarrow \mathbb{R}^{\mathbb{N}},\ (u_n)_{n\in\mathbb{N}} \longmapsto (u_{n+1})_{n\in\mathbb{N}}$ est linéaire.

Proposition. 2.36.
Soit $f : E \to F$ une application linéaire. On a les résultats suivants.

i. $f(0_E) = 0_F$.

ii. Soit $n \in \mathbb{N}^*$. Pour tout $(x_1, \ldots, x_n) \in E^n$ et $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ on a $f\left(\sum_{i=1}^n \lambda_i x_i\right) = \sum_{i=1}^n \lambda_i f(x_i)$.

Remarque 2.7.
L'image du vecteur nul (de l'espace de départ) est toujours le vecteur nul (de l'espace d'arrivée), mais ces deux vecteurs n'ont pas nécessairement la même nature.

Proposition. 2.37.
Soient E, F, G des $\mathbb{K}$-espaces vectoriels, ainsi que $(f_1, f_2) \in \mathcal{L}(E,F)^2$ et $(g_1, g_2) \in \mathcal{L}(F,G)^2$. On a:

$$(g_1 + g_2) \circ f_1 = g_1 \circ f_1 + g_2 \circ f_1 \qquad \text{et} \qquad g_1 \circ (f_1 + f_2) = g_1 \circ f_1 + g_1 \circ f_2.$$

Ce théorème montre que la compostion est distributive par rapport à l'addition. Pour cette raison, la composée de deux morphismes $g \circ f$ est parfois notée gf, à la manière d'un produit; néanmoins, la composition n'est pas commutative, puisque $g \circ f$ et $f \circ g$ sont en général distincts.
Par ailleurs, la notation f^n désigne f composée n fois avec elle-même : $f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ fois}}$.

2.5.2 Noyau et image d'une application linéaire

Définition 2.21. **Image et noyau**
Soit $\varphi : E \longrightarrow F$ une application linéaire.

i. On appelle **noyau** de φ l'ensemble défini par $\mathrm{Ker}(\varphi) = \{x \in E :\ \varphi(x) = 0_F\} = \varphi^{-1}(\{0_F\})$.

ii. On appelle **image** de φ l'ensemble défini par $\mathrm{Im}(\varphi) = \{y \in F : \exists x \in E, y = \varphi(x)\} = \{\varphi(x) : x \in E\} = \varphi(E)$.

Proposition. 2.38.
Soit $\varphi : E \to F$ une application linéaire. Alors :

i. $\mathrm{Ker}(\varphi)$ est un sous-espace vectoriel de E ;
ii. $\mathrm{Im}\varphi$ est un sous-espace vectoriel de F.

☞ Ce résultat est très utile afin de reconnaitre directement des sous-espaces vectoriels. Dès qu'on peut écrire un ensemble comme image ou noyau d'une application linéaire, c'est automatiquement un (sous-) espace vectoriel!

Exercice 2.13.
Soient a, b, c des réels. Montrer que l'ensemble des solutions à valeurs dans $\mathbb{R}$ de l'équation différentielle $ay'' + by' + cy = 0$ est un $\mathbb{R}$-espace vectoriel.

Proposition. 2.39. **Injectivité et Surjectivité**

Soit $\varphi : E \to F$ une application linéaire. On a les résultats suivants:

i. φ est surjective si et seulement si $\operatorname{Im}(\varphi) = F$.
ii. φ est injective si et seulement si $\operatorname{Ker}(\varphi) = \{0_E\}$.

2.5.3 Notion d'équation linéaire

Définition 2.22.

On appelle **équation linéaire** d'inconnue $x \in E$ toute équation de la forme $f(x) = y_0$, où $f : E \to F$ est une application linéaire et $y_0 \in F$.

Exemple 2.14.

- ✗ Toute équation différentielle du type $y' + a(x)y = b(x)$ est une équation linéaire.
- ✗ Tout système linéaire de n équations à p inconnues est une équation linéaire au sens de la définition ci-dessus.

Proposition. 2.40.

Soit $f \in \mathcal{L}(E, F), y_0 \in F$ et $f(x) = y_0$ une équation linéaire. Si x_1 est une solution particulière de cette équation, alors l'ensemble $\mathcal{S}$ de toutes ses solutions vérifie:

$$\mathcal{S} = \{x_1 + u;\ u \in \operatorname{Ker}(f)\}.$$

Exemple 2.15.

- ✗ La solution générale d'une équation différentielle est égale à la somme d'une solution particulière de l'équation, et de la solution générale de l'équation homogène associée.
- ✗ La solution générale d'un système linéaire est égale à la somme d'une solution particulière du système, et de la solution générale du système homogène associé.

2.5.4 Caractérisation des isomorphismes, isomorphisme fondamental

Définition 2.23.

Soit $\mathcal{E} = (e_1, \ldots, e_n)$ une famille finie de vecteurs de E et $f \in \mathcal{L}(E, F)$. La famille $(f(e_1), \ldots, f(e_n))$ est appelée l'**image** de $\mathcal{E}$ par f et est notée $f(\mathcal{E})$.

Proposition. 2.41.

Soient E, F deux $\mathbb{K}$-espaces vectoriels tel que E soit de dimension finie, $\mathcal{E}$ une famille finie de vecteurs de E et $f \in \mathcal{L}(E, F)$. Si $\mathcal{E}$ génère E, alors $f(\mathcal{E})$ génère $\operatorname{Im}(f)$.

Proposition. 2.42.

Soient E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $f \in \mathcal{L}(E, F)$. On a les résultats suivants.

i. f est bijective si, et seulement si, l'image par f de toute base de E est une base de F.
ii. f est bijective si, et seulement si, il existe une base de E dont l'image par f est une base de F.

On en déduit les résultats suivants.

Théorème 2.43. **Isomorphisme fondamental**

Soient E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie. Il existe un isomorphisme entre E et F si et seulement si $\dim(E) = \dim(F)$. En particulier, tout $\mathbb{K}$-espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

Exemple 2.16.

Si l'on dispose d'un isomorphisme entre $\mathbb{K}^n$ et $\mathbb{K}^p$, alors $n = p$. Il n'existe par exemple pas d'isomorphisme $f : \mathbb{K}^3 \to \mathbb{K}^2$.

Exercice 2.14.

Soit $n \in \mathbb{N}^*$. Montrer que l'application $\psi : \mathbb{R}_{n-1}[X] \longrightarrow \mathbb{R}^n$, $\sum_{k=0}^{n-1} a_k X^k \longmapsto (a_0, \ldots, a_{n-1})$ est un isomorphisme.

2.5.5 Rang d'une application linéaire

Définition 2.24. **Rang d'une application linéaire**

Soit $f \in \mathcal{L}(E, F)$. Si $\mathrm{Im}(f)$ est de dimension finie, on appelle **rang** de f, et l'on note $\mathrm{rg}(f)$ la dimension de $\mathrm{Im}(f)$, c'est-à-dire

$$\mathrm{rg}(f) = \dim(\mathrm{Im}(f)).$$

Proposition. 2.44.

Soit $f \in \mathcal{L}(E, F)$, et $(e_1, \ldots, e_p)$ une base de E. On a alors:

i. $\mathrm{rg}(f) = \mathrm{rg}\left(f\left(e_1\right), \ldots, f\left(e_p\right)\right)$.

ii. $\mathrm{rg}(f) \leqslant \min(\dim(E), \dim(F))$.

Théorème 2.45. **Théorème du rang**

Soient E et F des espaces vectoriels tels que E est de dimension finie, et $f \in \mathcal{L}(E, F)$. Alors on a

$$\dim(\mathrm{Ker}(f)) + \mathrm{rg}(f) = \dim(E).$$

Remarque 2.8.

Attention, il est erroné de penser que $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont supplémentaires. Déjà car ce ne sont *a priori* pas des sous-espaces du même espace vectoriel, mais même si f est un endomorphisme ce résultat n'est pas vrai en général.

Théorème 2.46. **Corollaire du Th. du rang pour les endomorphismes**

Soit E un espace vectoriel de **dimension finie**, et $f \in \mathcal{L}(E)$. On a alors les équivalences suivantes:

$$f \text{ injective} \iff f \text{ surjective} \iff f \text{ bijective}.$$

Remarque 2.9.

Ce résultat n'est pas vrai dans le cas d'une application linéaire quelconque. Par contre, il reste valable dans le cas où $f \in \mathcal{L}(E, F)$ avec $\dim(E) = \dim(F)$.

☞ Pour montrer qu'un endomorphisme est un automorphisme, il suffit de vérifier son injectivité ou sa surjectivité, mais il est inutile de vérifier les deux.

2.5.6 Propriétés des hyperplans

Proposition. 2.47. **Caractérisation d'un hyperplan**

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $H \subset E$. Alors H est un hyperplan de E si et seulement si H est le noyau d'une forme linéaire $\varphi : E \to \mathbb{K}$ non nulle.

Corollaire 2.48.

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\mathcal{B}$ une base de celui-ci, et $H \subset E$. Alors H est un hyperplan de E si et seulement si il existe $(a_1, \ldots, a_n) \in \mathbb{K}^n \setminus \{(0, \ldots, 0)\}$ tels que H admette pour équation $a_1 x_1 + \cdots + a_n x_n = 0$ dans la base $\mathcal{B}$.

Proposition. 2.49.

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. L'intersection de p hyperplans de E est un sous-espace vectoriel de dimension supérieure à $n - p$.

Réciproquement, tout sous-espace de dimension $n - p$ est l'intersection de p hyperplans.

Exemple 2.17.
Considérons un système linéaire homogène à p équations et n inconnues. L'ensemble $\mathcal{S}$ des solutions associé est alors un sous-espace vectoriel de $\mathbb{K}^n$, qui s'écrit comme intersection de p hyperplans (chacune des équations du système caractérise un hyperplan): $\mathcal{S}$ est donc de dimension supérieure à $n-p$.

En particulier, l'ensemble des solutions d'un système linéaire homogène à deux équations et trois inconnues est soit une droite, soit un plan soit $\mathbb{K}^3$ tout entier.

2.6 Projecteurs et symétries

2.6.1 Les projecteurs vectoriels

Définition 2.25. **Projecteur**
Soient F_1, F_2 deux sous-espaces vectoriels supplémentaires dans E. On appelle **projecteur sur** F_1 **parallèlement à** F_2 l'application $p : E \to E$ définie $p(u_1+u_2) = u_1$ pour tout $(u_1, u_2) \in F_1 \times F_2$.

Cette définition a bien un sens puisque $E = F_1 \oplus F_2$, donc pour tout $u \in E$, il existe un unique couple $(u_1, u_2) \in F_1 \times F_2$ tel que $u = u_1 + u_2$.

Exercice 2.15.
Soit E un espace vectoriel, et F_1, F_2 deux sous-espaces vectoriels supplémentaires dans E. Soit p (resp. q) le projecteur sur F_1 parallèlement à F_2 (resp. sur F_2 parallèlement à F_1). Montrer que $p+q = \mathrm{id}_E$ et $pq = qp = 0_{\mathcal{L}(E)}$.

Proposition. 2.50.
Soient F_1, F_2 deux sous-espaces vectoriels supplémentaires dans E. On note alors p le projecteur sur F_1 parallèlement à F_2. On a les résultats suivants:

i. $p \in \mathcal{L}(E)$;
ii. $p \circ p = p$;
iii. $F_2 = \mathrm{Ker}(p)$;
iv. $F_1 = \mathrm{Im}(p)$;
v. $\mathrm{Im}(p) = \mathrm{Ker}(p - \mathrm{id}_E)$.

☞ Le point *v.* ci-dessus assure que $x \in F_1$ si et seulement si $p(x) = x$ pour tout $x \in E$.

Réciproquement, on a le résultat suivant.

Proposition. 2.51. **Caractérisation des projecteurs**
Soit $p \in \mathcal{L}(E)$ tel que $p \circ p = p$. Alors $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ sont supplémentaires dans E, et p est le projecteur sur $\mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$.

☞ Les projecteurs sont donc les endomorphismes f caractérisés par la relation $f \circ f = f$.

2.6.2 Les symétries vectorielles

Définition 2.26. **Symétrie vectorielle**
Soient F_1, F_2 deux sous-espaces vectoriels supplémentaires dans E. On appelle **symétrie par rapport à** F_1 **et parallèlement à** F_2 l'application linéaire $s : E \to E$ définie par $s(u_1+u_2) = u_1 - u_2$ pour tout $(u_1, u_2) \in F_1 \times F_2$.

Proposition. 2.52.
Soient F_1, F_2 deux sous-espaces vectoriels supplémentaires dans E. On note p le projecteur sur F_1 parallèlement à F_2 et s la symétrie par rapport à F_1 parallèlement à F_2. On a alors $s = 2p - \mathrm{id}_E$.

Proposition. 2.53.

Soient F_1, F_2 deux sous-espaces vectoriels supplémentaires dans E. On note alors s la symétrie par rapport à F_1 parallèlement à F_2. On a les résultats suivants:

i. $s \in \mathcal{L}(E)$;
ii. $s^2 = \mathrm{id}_E$;
iii. $F_1 = \mathrm{Ker}(s - \mathrm{id}_E)$;
iv. $F_2 = \mathrm{Ker}(s + \mathrm{id}_E)$.

☞ Le point *iii.* ci-dessus assure que $x \in F_1$ si et seulement si $s(x) = x$ pour tout $x \in E$.
Le point *iv.* ci-dessus assure que $x \in F_2$ si et seulement si $s(x) = -x$ pour tout $x \in E$.

Réciproquement, on a le résultat suivant.

Proposition. 2.54. **Caractérisation des symétries vectorielles**

Soit $s \in \mathcal{L}(E)$ tel que $s^2 = \mathrm{id}_E$. Alors $\mathrm{Ker}(s - \mathrm{id}_E)$ et $\mathrm{Ker}(s + \mathrm{id}_E)$ sont supplémentaires dans E, et s est la symétrie sur $\mathrm{Ker}(s - \mathrm{id}_E)$ parallèlement à $\mathrm{Ker}(s + \mathrm{id}_E)$.

☞ Les symétries sont donc les endomorphismes f caractérisés par la relation $f \circ f = \mathrm{id}_E$.

Exemple 2.18.

✗ Soit $n \in \mathbb{N}^*$. L'endomorphisme $\tau : M \mapsto M^T$ de $\mathcal{M}_n(\mathbb{K})$ est la symétrie par rapport à $\mathcal{S}_n(\mathbb{K})$ parallèlement à $\mathcal{A}_n(\mathbb{K})$.
✗ Soit σ la fonction définie sur $\mathbb{R}$ par $\sigma(x) = -x$. L'endomorphisme $\Phi : f \mapsto f \circ \sigma$ de $\mathbb{R}^{\mathbb{R}}$ est la symétrie par rapport à l'ensemble des fonctions paires, parallèlement à l'ensemble des fonctions impaires.

2.6.3 Famille de projecteurs associée à une décomposition en somme directe

Définition 2.27.

Soit $n \in \mathbb{N}^*$ et $(F_1, \dots, F_n)$ une famille de sous-espaces vectoriels de E tels que $E = \bigoplus_{k=1}^{n} F_k$.

Pour tout $i \in [\![1, n]\!]$, on note p_i le projecteur sur F_i parallèlement à $\bigoplus_{\substack{k=1\\k\neq i}}^{n} F_k$.

La famille $(p_i)_{1 \leqslant i \leqslant n}$ est appelée **famille des projecteurs associés à la décomposition** $E = \bigoplus_{k=1}^{n} F_k$.

Exemple 2.19.

Si les F_i sont tous des droites vectorielles, alors ils correspondent aux axes d'un repère, et les p_i sont ainsi les projections sur ces axes.

Proposition. 2.55.

Avec les notations de la **Définition 2.27.**, on a :

i. $p_1 + \dots + p_n = \mathrm{id}_E$;
ii. $p_i \circ p_j = 0_{\mathcal{L}(E)}$ si $i \neq j$ sont des entiers dans $[\![1, n]\!]$.

3 Courbes paramétrées

Dans tout ce chapitre, on considère l'espace euclidien $(\mathbb{R}^2, \|\cdot\|)$ et on note d la distance associée.
Sauf mention explicite du contraire, I désigne un intervalle réel (non trivial) et $\overline{I}$ son adhérence.

3.1 Arcs (ou courbes) paramétré(e)s du plan

3.1.1 Définitions et exemples

Définition 3.1. **Arc ou courbe paramétré(e)**

On appelle **arc (ou courbe) paramétré(e)** (de classe $\mathcal{C}^k$) toute partie γ de $\mathbb{R}^2$ de la forme $\gamma = \{M(t) \mid t \in I\}$, où $M : I \to \mathbb{R}^2$ est une fonction de classe $\mathcal{C}^k$ sur I.

M est appelé **paramétrage** de γ et, réciproquement, γ est appelé **support** (ou trajectoire) de M.

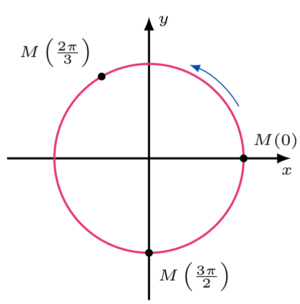
Pour tout $t \in I$, le point $M(t)$ de γ sera appelé le point de paramètre t de γ.

Si $t \mapsto x(t)$ et $t \mapsto y(t)$ désignent les fonctions coordonnées de M, on dira que $\left\{\begin{array}{lcl} x & = & x(t) \\ y & = & y(t) \end{array}\right.$, $t \in I$ est un paramétrage de γ.

Exemple 3.1.

$$\left\{\begin{array}{lcl} x(t) & = & \cos(t) \\ y(t) & = & \sin(t) \end{array}\right. , \ t \in [0, 2\pi[$$

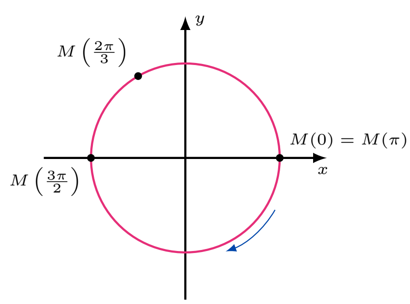
est un paramétrage du cercle unité. Le cercle est parcouru dans le sens direct (inverse des aiguilles d'une montre).



Exemple 3.2.

$$\left\{\begin{array}{lcl} x(t) & = & \cos(-2t) \\ y(t) & = & \sin(-2t) \end{array}\right. , \ t \in [0, 2\pi[$$

est un autre paramétrage du cercle unité. Le cercle est parcouru deux fois dans le sens des aiguilles d'une montre.

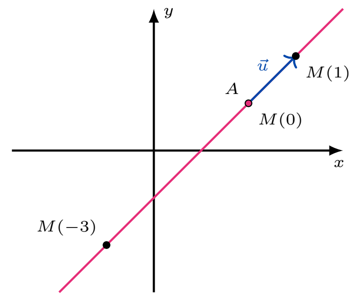


Exemple 3.3.

$$\begin{cases} x(t) &= a+\alpha t \\ y(t) &= b+\beta t \end{cases}, \ t \in [0, 2\pi[$$

est un paramétrage de la droite passant par le point $A(a,b)$ et dirigée par le vecteur $\vec{u}(\alpha,\beta)$ (que l'on note aussi $A+\mathbb{R}\vec{u}$) et d'équation

$$\beta(x-a)-\alpha(y-b)=0.$$

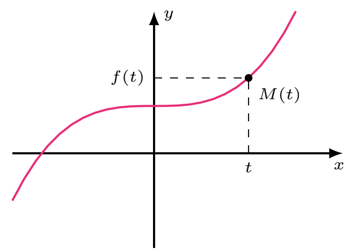


Exemple 3.4.

Soit $f:\ I \to \mathbb{R}$ une fonction réelle. Alors,

$$\begin{cases} x(t) &= t \\ y(t) &= f(t) \end{cases}, \ t \in I$$

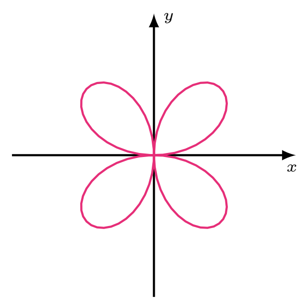
est un paramétrage de la courbe représentative de f.



Exemple 3.5.

$$\begin{cases} x(t) &= 2\sin^2(t)\cos(t) \\ y(t) &= 2\cos^2(t)\sin(t) \end{cases}, \ t \in [0, 2\pi[$$

est un paramétrage du *quadrifolium*, ou trèfle à quatre feuilles.



3.1.2 Points simples, points multiples d'une courbe paramétrée

Définition 3.2. **Point multiple**

On considère une courbe γ paramétrée par $M:\ I \to \mathbb{R}^2$. Un point A de γ est appelé **point multiple** s'il existe (au moins) deux réels (distincts) t_1 et t_2 tels que $M(t_1)=M(t_2)$ et t_2-t_1 n'est pas une période pour M.
On dira plus précisément que $M(t)$ est double s'il correspond à deux valeurs distinctes du paramètres, triple s'il correspond à trois valeurs distinctes du paramètre, etc...
Un point qui n'est pas multiple est dit **simple**.

Exemple 3.6.

Dans l'**Exemple 3.1**, tous les points sont simples.
Dans l'**Exemple 3.5**, le point $O(0,0)$ est multiple; on y passe quatre fois (on dira qu'il est quadruple).

Exercice 3.1.

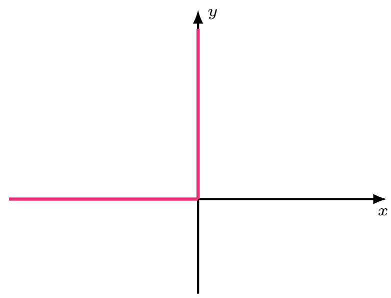
Déterminer les éventuels points doubles de la courbe paramétrée par $\begin{cases} x(t) &= 3t^3+2t^2-t-1 \\ y(t) &= 3t^2+2t+1 \end{cases}, \ t \in \mathbb{R}$.

3.1.3 Tangente en un point

Afin de motiver la définition qui va suivre, observons un exemple de courbe paramétrée. Considérons la courbe γ paramétrée par $M:\ t \mapsto (x(t), y(t))$ où

$$\forall t \in \mathbb{R}, \qquad x(t) = \begin{cases} -t^2, & \text{si } t < 0 \\ 0, & \text{si } t \geq 0 \end{cases} \qquad \text{et} \qquad y(t) = \begin{cases} 0, & \text{si } t < 0 \\ t^2, & \text{si } t \geq 0 \end{cases}$$

dont la représentation est la suivante.



Bien que M soit dérivable (et même de classe $\mathcal{C}^\infty$) en tout point $t \in \mathbb{R}$, on observe que γ n'admet pas de tangente au point $O(0,0)$.

Définition 3.3. **Point régulier, courbe régulière**

Soit γ une courbe paramétrée de classe $\mathcal{C}^1$ paramétrée par $M:\ I \to \mathbb{R}^2$. On dit que le point $M(t_0)$ de paramètre $t_0 \in I$ est un **point régulier** de γ si

$$\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) \neq \vec{0}.$$

Sinon, le point est dit **singulier** (ou stationnaire).
Une courbe paramétrée γ est dite **régulière** si tous ses points sont réguliers.

Si γ est de classe $\mathcal{C}^2$, le point $M(t_0)$ est dit **birégulier** si la famille

$$\left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0), \frac{\mathrm{d}^2\overrightarrow{M}}{\mathrm{d}t^2}(t_0) \right)$$

est libre.

Afin de donner la définition qui va suivre, on rappelle que si A est un point du plan et $\vec{u}$ un vecteur non nul, la droite dirigée par $\vec{u}$ et passant par A est notée $A + \mathbb{R}\vec{u}$.

On dit que la courbe γ paramétrée par $M:\ I \to \mathbb{R}^2$ admet une tangente en $M(t_0)$ si la droite $(M(t_0)M(t))$ admet une *position limite* lorsque t tend vers t_0; auquel cas la droite limite est la tangente.

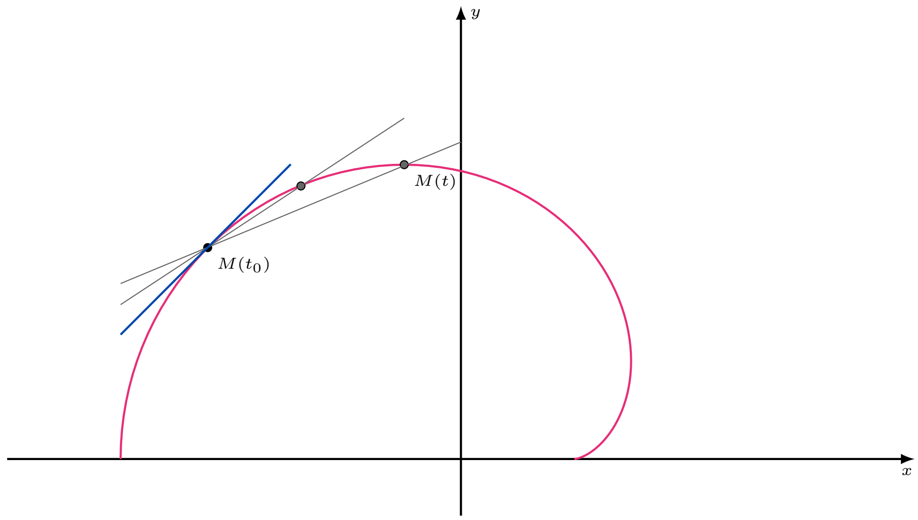
Définition 3.4.

Soient γ une courbe paramétrée par $M:\ I \to \mathbb{R}^2$ et $t_0 \in I$. On note $\vec{d}(t)$ un vecteur unitaire, directeur de la corde $(M(t_0)M(t))$.

i. Si $\vec{d}(t)$ admet une limite finie $\vec{d}_0^+ \in \mathbb{R}^2$ lorsque $t \to t_0^+$, on dit γ admet la droite $M(t_0) + \mathbb{R}\vec{d}_0^+$ comme **tangente à droite** en $M(t_0)$;

ii. Si $\vec{d}(t)$ admet une limite finie $\vec{d}_0^- \in \mathbb{R}^2$ lorsque $t \to t_0^-$, on dit γ admet la droite $M(t_0) + \mathbb{R}\vec{d}_0^-$ comme **tangente à gauche** en $M(t_0)$;

iii. Si les tangentes à droite et à gauche existent et sont une seule et même droite $M(t_0) + \mathbb{R}\vec{d}_0$, celle-ci est appelée **tangente** en $M(t_0)$ à la courbe.



Proposition. 3.1.
Soit γ une courbe de classe $\mathcal{C}^1$ paramétrée par $M: I \to \mathbb{R}^2$. Alors, en tout point régulier $M(t_0)$ de γ. Plus précisément, la droite
$$M(t_0) + \mathbb{R}\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0)$$
est tangente à γ au point $M(t_0)$.

Preuve. Soit $M(t_0)$ un point régulier. Si $t \in I, t \neq t_0$, un vecteur directeur unitaire de la corde $(M(t_0)M(t)$ est donné par
$$\begin{aligned}
\vec{d}(t) &= \frac{1}{\|M(t_0) - M(t)\|}(x(t) - x(t_0), y(t) - y(t_0)) \\
&= \frac{t - t_0}{\sqrt{(x(t) - x(t_0))^2 + (y(t) - y(t_0))^2}}\left(\frac{x(t) - x(t_0))}{t - t_0}, \frac{y(t) - y(t_0)}{t - t_0}\right) \\
&= \frac{\pm 1}{\sqrt{\left(\frac{x(t)-x(t_0)}{t-t_0}\right)^2 + \left(\frac{y(t)-y(t_0)}{t-t_0}\right)^2}}\left(\frac{x(t) - x(t_0))}{t - t_0}, \frac{y(t) - y(t_0)}{t - t_0}\right) \\
&\xrightarrow[t \to t_0]{} \frac{\pm 1}{\left\|\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0)\right\|}(x'(t_0), y'(t_0)) =: \vec{d}_0^{\pm}.
\end{aligned}$$
On aura observé que, comme $M(t_0)$ est supposé régulier, $\left\|\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0)\right\| \neq 0$. Comme les deux droites sont les mêmes :
$$M(t_0) + \mathbb{R}\vec{\mathrm{d}}_0^+ = M(t_0) + \mathbb{R}\vec{\mathrm{d}}_0^-,$$
on a bien le résultat énoncé. □

Exercice 3.2.
Déterminer une équation de la tangente en un point quelconque de la courbe γ paramétrée par
$$M: t \longmapsto \begin{cases} x(t) = \cos(t) \\ y(t) = \sin(t) \end{cases}, \ t \in [0, 2\pi[.$$
Vérifier ensuite qu'elle est perpendiculaire à la droite $(OM(t_0))$, où O désigne l'origine du repère (et le centre du cercle dont M est une paramétrisation).

Exemple 3.7.
Si $f: I \to \mathbb{R}$ est une fonction réelle de classe $\mathcal{C}^1$, la tangente à la courbe représentative de f au point d'abscisse t_0 est également la tangente à la courbe γ paramétrée par $M: t \in I \mapsto (t, f(t))$ au point $M(t_0)$ et a pour vecteur directeur
$$\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) = (1, f'(t_0)).$$
On retrouve bien que $f'(t_0)$ en est le coefficient directeur.

☞ Une courbe paramétrée peut avoir une tangente verticale, contrairement à ce à quoi on est habitué pour des graphes de fonctions d'équation $y = f(x)$.

Exercice 3.3.
Trouver les points où la tangente à la courbe de Lissajous, paramétrée par

$$\begin{cases} x(t) &= \sin(2t) \\ y(t) &= \sin(3t) \end{cases}, \; t \in [-\pi, \pi]$$

est verticale, puis horizontale.

Remarque 3.1.
Si on écrit la formule de Taylor-Young (ou le développement limité), à l'ordre 1, en t_0, pour le paramétrage M d'un courbe γ, on a

$$M(t) = M(t_0) + \frac{\overrightarrow{\mathrm{d}M}}{\mathrm{d}t}(t_0)(t - t_0) + \overrightarrow{o(t - t_0)}.$$

Notamment, la *partie principale* correspond au paramétrage de la tangente à γ en $M(t_0)$.

Remarque 3.2.
La notion de tangente en un point multiple n'a pas toujours de sens, car il peut y avoir différentes tangentes en un tel point, chacune étant associée à une valeur différente du paramètre en lequel le point est atteint.
Il faudra donc, le cas échéant, toujours préciser en quelle valeur du paramètre on calcule la tangente.

3.1.4 Étude locale aux points singuliers

Considérons un point singulier $M(t_0)$ (c'est à dire que $\frac{\overrightarrow{\mathrm{d}M}}{\mathrm{d}t}(t_0) = \vec{0}$).
Comme le vecteur dérivée est nul, il n'est d'aucune utilité pour la recherche d'une tangente ; il parait naturel de chercher à préciser les termes suivants dans le développement limité de M. Notons alors

$$p = \min\left\{k : \frac{\overrightarrow{\mathrm{d}^k M}}{\mathrm{d}t^k}(t_0) \neq \vec{0}\right\}.$$

On peut alors écrire, pour t au voisinage de t_0,

$$M(t) = M(t_0) + \frac{\overrightarrow{\mathrm{d}^p M}}{\mathrm{d}t^p}(t_0)\frac{(t - t_0)^p}{p!} + o\left((t - t_0)^p\right)$$

et, dans ce cas, la courbe paramétrée $t \mapsto M(t_0) + \frac{\overrightarrow{\mathrm{d}^p M}}{\mathrm{d}t^p}(t_0)\frac{(t - t_0)^p}{p!}$ est la (demi-)tangente à γ en $M(t_0)$.
En effet, si p est pair, il s'agit d'une demi-droite et parler de tangente est donc ambigu (on parlera plutôt de *rebroussement*). Il nous faut alors une seconde direction pour préciser ce qui se passe. Le résultat est alors le suivant.

Proposition. 3.2.
Soient $n \geq 2$ un entier, $M : I \to \mathbb{R}^2$ un paramétrage de classe $\mathcal{C}^n$ d'une courbe γ et $t_0 \in I$.
On suppose qu'il existe deux entiers $(p, q) \in [\![1, n]\!]^2$ tels que

$$p = \min\left\{k \geq 2 : \frac{\overrightarrow{\mathrm{d}^k M}}{\mathrm{d}t^k}(t_0) \neq \vec{0}\right\}, \qquad q = \min\left\{k > p : \left(\frac{\overrightarrow{\mathrm{d}^p M}}{\mathrm{d}t^p}(t_0), \; \frac{\overrightarrow{\mathrm{d}^k M}}{\mathrm{d}t^k}(t_0)\right) \text{ est libre }\right\}.$$

Alors, on a les cas de figure suivants (illustrés ci-après) :

i. Si p est impair et que q est pair, on a un **point d'allure ordinaire**;
ii. Si p est impair et que q est impair, on a un **point d'inflexion**;
iii. Si p est pair et que q est impair, on a un **point de rebroussement de première espèce**;
iv. Si p est pair et que q est pair, on a un **point de rebroussement de seconde espèce**.

Preuve. La démonstration repose sur le développement limité obtenu ci-avant ; pour t au voisinage de t_0,

$$M(t) \quad = \quad M(t_0) + \frac{\overrightarrow{\mathrm{d}^p M}}{\mathrm{d}t^p}(t_0)\frac{(t - t_0)^p}{p!} + \ldots + \frac{\overrightarrow{\mathrm{d}^q M}}{\mathrm{d}t^q}(t_0)\frac{(t - t_0)^q}{q!} + (t - t_0)^q\vec{\varepsilon}(t - t_0),$$

où $\vec{\varepsilon}$ est une fonction à valeurs dans $\mathbb{R}^2$, continue au voisinage de t_0 telle que $\lim_{h\to 0} \|\varepsilon(h)\| = 0$.
Pour alléger les notations, notons

$$\vec{v} = \frac{\mathrm{d}^p\overrightarrow{M}}{\mathrm{d}t^p}(t_0), \qquad \vec{w} = \frac{\mathrm{d}^q\overrightarrow{M}}{\mathrm{d}t^q}(t_0).$$

Observons que si l'on note $P \in \mathcal{M}_2(\mathbb{R})$ la matrice de passage de la base canonique (e_1, e_2) de $\mathbb{R}^2$ vers la base $(\vec{v}, \vec{w})$, alors, les coordonnées de $\varepsilon(t - t_0)$ dans cette nouvelle base tendent encore vers 0 lorsque $t \to t_0$:

$$\|P^{-1}\varepsilon(h)\| \leq \|P\|_2\|\varepsilon(h)\| \underset{h\to 0}{\longrightarrow} 0.$$

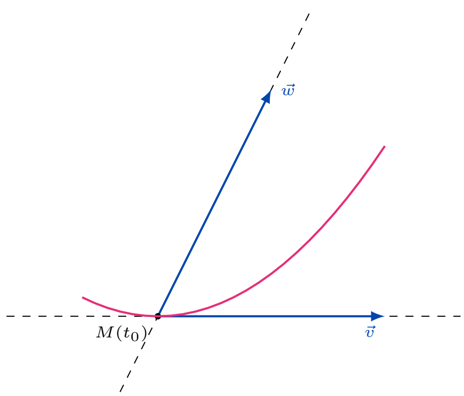
Notons alors $(X(t), Y(t))$ les coordonnées de $M(t)$ dans le repère $\mathcal{R}' = (M(t_0), \vec{v},\ \vec{w})$.
On a donc immédiatement, d'une part,

$$Y(t) = \frac{(t-t_0)^q}{q!} + o((t-t_0)^q).$$

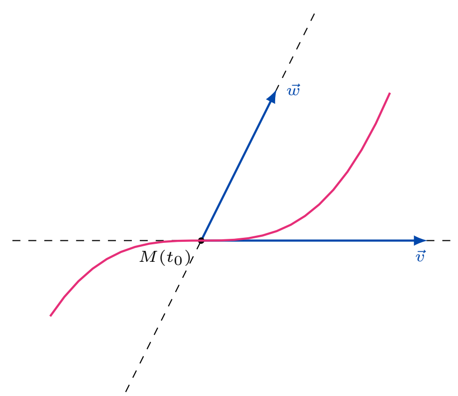
Comme, pour tout $k \in [\![p, q-1]\!]$, $\dfrac{\mathrm{d}^k\overrightarrow{M}}{\mathrm{d}t^k}(t_0)$ est colinéaire à $\dfrac{\mathrm{d}^p\overrightarrow{M}}{\mathrm{d}t^p}(t_0)$ et que, pour $k \geq p$, $(t-t_0)^k = o\left((t-t_0)^p\right)$ (lorsque $t \to t_0$), on a aussi

$$X(t) = \frac{(t-t_0)^p}{p!} + o((t-t_0)^p).$$

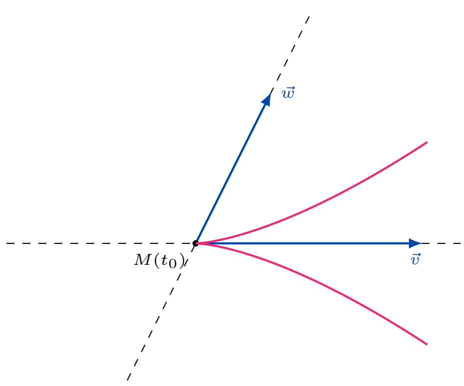
En $M(t_0)$, γ admet alors une tangente de vecteur directeur $\vec{v} = \dfrac{\mathrm{d}^p\overrightarrow{M}}{\mathrm{d}t^p}(t_0)$. La position de la courbe par rapport à cette tangente est donnée par les parités de p et q, ce qu'on représente avec les figures ci-dessous. On a un des quatre cas de figure suivants :



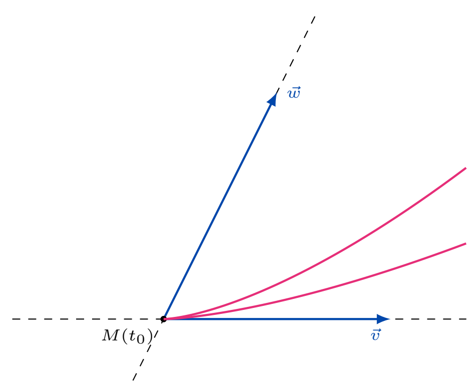
i. p impair, q pair
point d'allure normale



ii. p impair, q impair
point d'inflexion



iii. p pair, q impair
point de rebroussement de première espèce



iv. p pair, q pair
point de rebroussement de seconde espèce

□

Remarque 3.3.

✗ Tout point birégulier est un point d'allure normale (car on a dans ce cas $p=1, q=2$).

✗ Un point d'inflexion est un point où la courbe traverse sa tangente.
Si la courbe est régulière (*i.e.* pas de point singulier), les points d'inflexion sont à chercher parmi ceux pour lesquels

$$\frac{\mathrm{d}^p \overrightarrow{M}}{\mathrm{d}t^p}(t_0), \quad \text{et} \quad w = \frac{\mathrm{d}^q \overrightarrow{M}}{\mathrm{d}t^q}(t_0)$$

sont colinéaires.

✗ Un point de rebroussement est un point où les demi-tangentes à droite et à gauche sont une même demi-droite. Il s'agit toujours d'un point singulier. Mais un point singulier n'est pas toujours un point de rebroussement !

✗ Lorsque cela permet un calcul plus simple, on peut écrire ne développement limité de $M(t)$ pour lire immédiatement les valeurs de p et q.

✗ Si la limite

$$\lim_{t\to t_0} \frac{y(t)-y(t_0)}{x(t)-x(t_0)} \left(= \lim_{t\to t_0} \frac{y'(t)}{x'(t)}\right) = \ell$$

existe et est finie, alors ℓ est le coefficient directeur de la tangente à γ en $M(t_0)$. Si la limite précédente existe et est infinie, la tangente est verticale.

Exercice 3.4.
Étudier le point singulier à l'origine de chacune des courbes ci-dessous :

1. γ_1 paramétrée par $\left\{\begin{array}{lcl} x(t) & = & t^2 \\ y(t) & = & t^5 \end{array}\right.$.

2. γ_2 paramétrée par $\left\{\begin{array}{lcl} x(t) & = & t^5 \\ y(t) & = & t^3 \end{array}\right.$.

3. γ_3 paramétrée par $\left\{\begin{array}{lcl} x(t) & = & t^2 \ln(1+t) \\ y(t) & = & t^2 \left(\exp(t^2)-1\right) \end{array}\right.$.

Exercice 3.5.
Déterminer et étudier les points singuliers de la courbe γ paramétrée par $\left\{\begin{array}{lcl} x(t) & = & t^2-t^3 \\ y(t) & = & t^2+t^3 \end{array}\right.$.

3.1.5 Branches infinies

Définition 3.5. **Branches infinies**
Soient $M : I \to \mathbb{R}^2$ un paramétrage d'une courbe γ et $t_0 \in \overline{I}$. On note $x, y : I \to \mathbb{R}$ les fonctions coordonnées de M. On dit que M a une **branche infinie** en t_0 lorsque

$$\lim_{t\to t_0} M(t) = \lim_{t\to t_0} \|\overrightarrow{OM(t)}\| = \lim_{t\to t_0} \sqrt{x(t)^2+y(t)^2} = +\infty$$

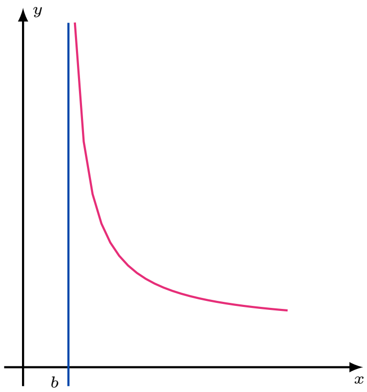
ce qui revient au même de dire qu'au moins l'une des fonction $t \mapsto |x(t)|$ ou $t \mapsto |y(t)|$ tend vers l'infini lorsque $t \to t_0$.

Pour chaque branche infinie, on cherche s'il existe une asymptote, c'est à dire une droite que va venir *épouser* la courbe, c'est à dire avec laquelle la distance tend vers 0 lorsque $t \to t_0$. Plus précisément, on a la définition suivante.

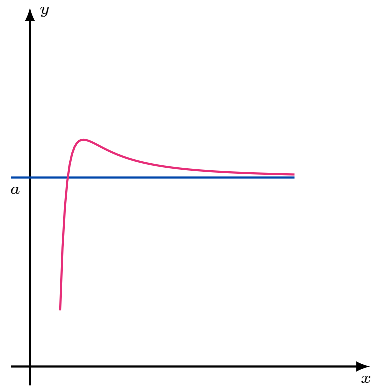
Définition 3.6. **Asymptote**
Soient $t_0 \in \overline{I}$, $M : I \to \mathbb{R}^2$ un paramétrage d'une courbe γ admettant une branche infinie en t_0. On dit que la droite $\mathcal{D}$ est **asymptote** à la courbe γ en t_0 si

$$\lim_{t\to t_0} d\left(M(t), \mathcal{D}\right) = 0.$$

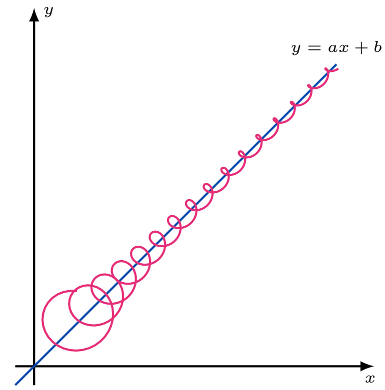
Une asymptote peut être verticale, horizontale ou oblique, comme on peut le voir avec les illustrations ci-dessous.



Asymptote verticale :
$x(t) \underset{t \to t_0}{\longrightarrow} b^+$ et $y(t) \underset{t \to t_0}{\longrightarrow} +\infty$.

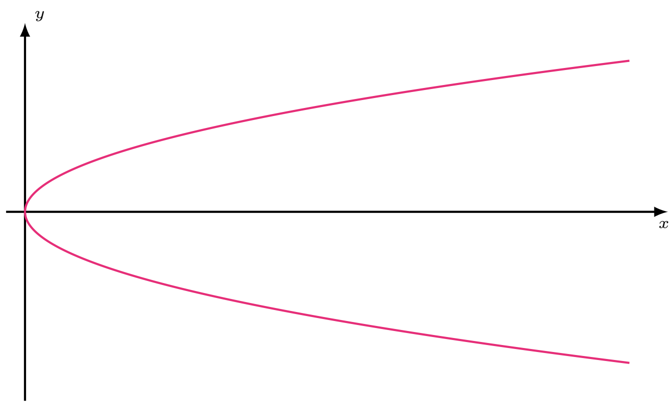


Asymptote horizontale :
$x(t) \underset{t \to t_0}{\longrightarrow} +\infty$ et $y(t) \underset{t \to t_0}{\longrightarrow} a^+$.



Asymptote oblique :
$[y(t) - (ax(t) + b)] \underset{t \to t_0}{\longrightarrow} 0$.

☞ **Attention**, une branche infinie peut ne pas admettre de droite asymptote (comme dans le cas d'une parabole), on parle alors *branche parabolique* de direction une certaine droite.



Branches paraboliques de direction horizontale :
$y(t) = o(x(t)),\ t \to \pm\infty$.

☞ **Pour déterminer la nature de la branche infinie**, en t_0; on forme le quotient $\dfrac{y(t)}{x(t)}$:

- ✗ Si $\dfrac{y(t)}{x(t)} \underset{t \to t_0}{\longrightarrow} \pm\infty$, on dira que γ admet une branche parabolique de direction (verticale) (Oy).
- ✗ Si $\dfrac{y(t)}{x(t)} \underset{t \to t_0}{\longrightarrow} 0$, on dira que γ admet une branche parabolique de direction (horizontale) (Ox).
- ✗ Si $\dfrac{y(t)}{x(t)} \underset{t \to t_0}{\longrightarrow} a \in \mathbb{R}^*$, et que $y(t) - ax(t) \underset{t \to t_0}{\longrightarrow} b \in \mathbb{R}$, on dira que γ admet une asymptote d'équation $y = ax + b$. Dans ce cas, il est nécessaire d'étudier le signe de $y(t) - (ax(t) + b)$ lorsque $t \to t_0$ pour connaître la position relative de la courbe par rapport à son asymptote.
- ✗ Si $\dfrac{y(t)}{x(t)} \underset{t \to t_0}{\longrightarrow} a \in \mathbb{R}^*$, et que $y(t) - ax(t) \underset{t \to t_0}{\longrightarrow} \pm\infty$, on dira que γ admet une branche parabolique de direction la droite d'équation $y = ax$.

Exercice 3.6.

Étudier la branche infinie de la courbe γ paramétrée par $\begin{cases} x(t) &= \dfrac{t^3+1}{t^2+1} \\ y(t) &= \dfrac{t^3+t^2}{t^2+1} \end{cases}$ lorsque $t \to +\infty$.

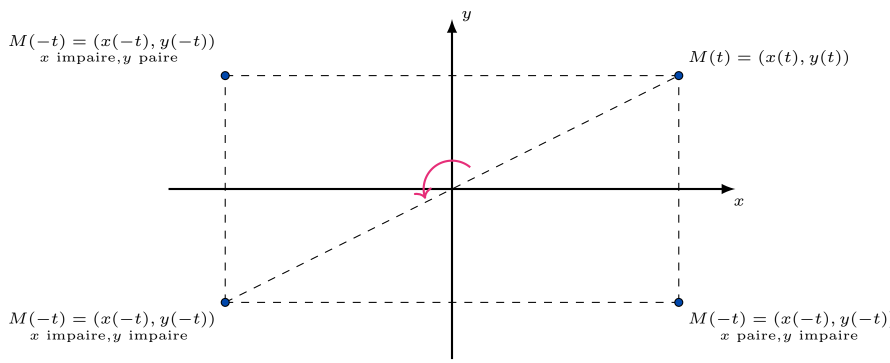
3.1.6 Réduction du domaine d'étude

Considérons une courbe γ paramétrée par $M :\ I \to \mathbb{R}^2$ dont les fonctions coordonnées sont notées x et y.

Il est clair que si x et y sont toutes deux des fonctions $T-$**périodiques** définies sur $I = \mathbb{R}$, il suffit d'étudier x et y sur $[0, T]$ (ou tout autre intervalle d'amplitude T) puis de tracer la courbe paramétrée par

$$\begin{cases} x &= x(t) \\ y &= y(t) \end{cases}, \quad t \in [0; T].$$

D'autres observations et transformations géométriques permettent de réduire le domaine d'étude et d'obtenir γ à partir de l'arc paramétré réduit auquel on fait subir les transformations géométriques susmentionnées.



- ✗ Si x et y sont toutes les deux paires, on restreint le domaine d'étude à $I \cap \mathbb{R}_+$. La courbe ainsi tracée est complète.

- ✗ Si x et y sont toutes les deux impaires, on restreint le domaine d'étude à $I \cap \mathbb{R}_+$. On complète ensuite la courbe ainsi tracée par symétrie de centre O.

- ✗ x est paire et y est impaire, on restreint le domaine d'étude à $I \cap \mathbb{R}_+$. On complète ensuite la courbe ainsi tracée par symétrie d'axe (Ox).

- ✗ x est impaire et y est paire, on restreint le domaine d'étude à $I \cap \mathbb{R}_+$. On complète ensuite la courbe ainsi tracée par symétrie d'axe (Oy).

- ✗ Si le changement de t en $a + b - t$ fait apparaître une transformation géométrique permettant de passer de $M(t)$ à $M(a + b - t)$, on restreint le domaine d'étude de $I = [a, b]$ à $\left[a, \frac{a+b}{2}\right]$, et on complète ensuite la courbe en appliquant cette transformation.

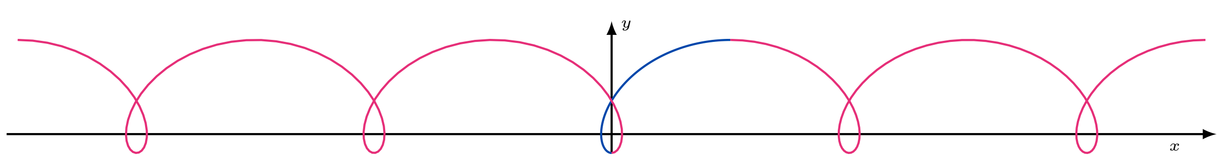
- ✗ Si le changement de t en $1/t$ fait apparaître une transformation géométrique permettant de passer de $M(t)$ à $M(1/t)$, on restreint le domaine d'étude de $I =]0, +\infty[$ à $]0, 1[$, et on complète ensuite la courbe en appliquant cette transformation.

- ✗ Si, pour tout $t \in I = \mathbb{R}$, $M(t + T) = M(t) + \vec{u}$, on restreint de domaine d'étude à $[0, T]$ et on complète la courbe par translations de vecteurs $k \cdot \vec{u}$, pour $k \in \mathbb{Z}$.

Exercice 3.7.

Déterminer un domaine d'étude le plus simple possible de la courbe γ paramétrée par $\begin{cases} x(t) &= t - \dfrac{3}{2}\sin(t) \\ y(t) &= 1 - \dfrac{3}{2}\cos(t) \end{cases}$.

On pourra ensuite expliquer la construction de la courbe sur la figure ci-dessous.



3.1.7 Plan d'étude d'une courbe paramétrée

Pour étudier une courbe paramétrée, en général, on suit le plan d'étude suivant :

1. On détermine l'ensemble de définition puis le domaine d'étude le plus simple possible.

2. On étudie les fonctions coordonnées x et y, on trace le tableau de variations à 5 lignes :

t	
$x'(t)$	
x	
y	
$y'(t)$	

3. On étudie les éventuels points singuliers.

4. On étudie les éventuelles branches infinies et les tangentes aux points qui apparaissent dans le tableau de variations.

5. Réalisation d'un brouillon du tracé. Pour cela :

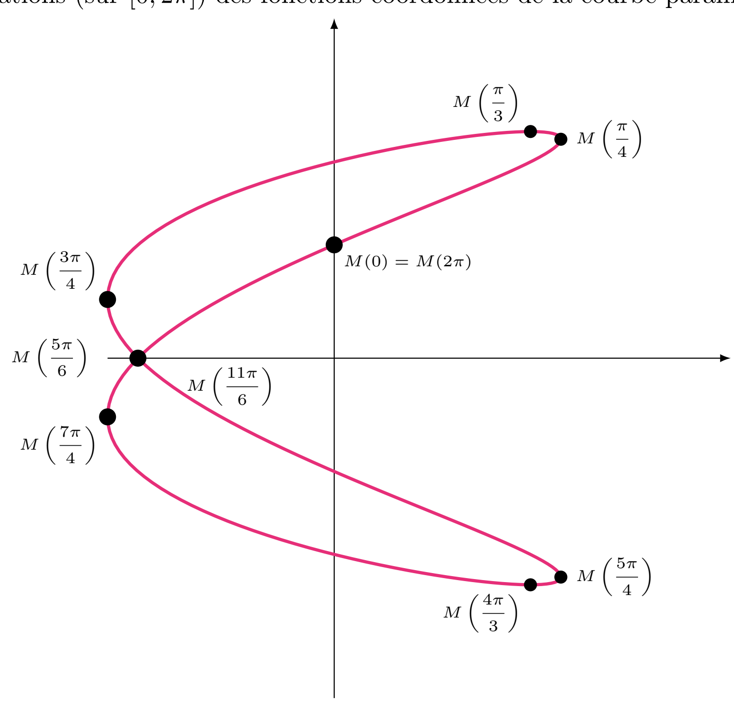
- ✗ Si x croît et y croît, on va vers la droite et vers le haut.
- ✗ Si x croît et y décroît, on va vers la droite et vers le bas.
- ✗ Si x décroît et y croît, on va vers la gauche et vers le haut.
- ✗ Si x décroît et y décroît, on va vers la gauche et vers le bas.

Le brouillon du tracé permet notamment de faire apparaître d'éventuels points multiples, dont on vérifiera l'existence par le calcul.

6. Construction méticuleuse de la courbe, asymptotes et tangentes, en utilisant éventuellement les transformations géométriques qui auraient pu permettre de réduire le domaine d'étude.

Exercice 3.8.

Dresser les tableaux de variations (sur $[0, 2\pi]$) des fonctions coordonnées de la courbe paramétrée représentée ci dessous



3.1.8 Un exemple d'étude complète : le folium de Descartes

On étudie dans cette dernière question la courbe γ paramétrée par

$$\begin{cases} x(t) &= \dfrac{3t}{1+t^3} \\ \\ y(t) &= \dfrac{3t^2}{1+t^3} \end{cases}$$

1. Les fonctions coordonnées sont toutes deux définie sur $\mathbb{R} \setminus \{-1\}$.
Pour $t \in \mathbb{R} \setminus \{-1\}$, on a

$$x\left(\frac{1}{t}\right) = y(t), \qquad y\left(\frac{1}{t}\right) = x(t).$$

Comme l'application $t \mapsto \frac{1}{t}$ est bijective de $[-1, 0[$ sur $]-\infty, -1]$ et de $]0, 1]$ sur $[1, +\infty[$, la restriction de l'étude à l'intervalle $[-1, 1]$ permet d'avoir une courbe complète en terminant le tracé par symétrie par rapport à la première bissectrice d'équation $y = x$.

2. On a, pour tout $t \in]-1, 1]$,

$$x'(t) = \frac{3(1-2t^3)}{(1+t^3)^2}, \qquad y'(t) = \frac{3t(2-t^3)}{(1+t^3)^2},$$

ce qui donne sans difficulté le tableau de variations ci-dessous :

t	-1		0		$\sqrt[3]{\frac{1}{2}}$		1
$x'(t)$	$\Vert$	$+$	$\vdots$	$+$	0	$-$	
x	$\Vert$ $-\infty$	$\nearrow$	0	$\nearrow$	$\sqrt[3]{4}$	$\searrow$	$3/2$
y	$\Vert$ $+\infty$	$\searrow$	0	$\nearrow$	$\sqrt[3]{2}$	$\nearrow$	$3/2$
$y'(t)$	$\Vert$	$-$	0	$+$	$\vdots$	$+$	

3. Il n'y a pas de point singulier.

4. Branches infinies et tangentes.
Lorsque t tend vers -1, on a $x(t) \underset{t\to-1}{\longrightarrow} -\infty, \qquad y(t) \underset{t\to-1}{\longrightarrow} +\infty.$

On a donc une branche infinie en -1 (et on remarque que c'est la seule). Comme $\lim\limits_{t\to-1} \dfrac{y(t)}{x(t)} = -1$, puis que

$$y(t) - (-x(t)) = \frac{3t}{t^2-t+1} \underset{t\to-1}{\longrightarrow} -1,$$

la droite $\mathcal{D}$ d'équation $y = -x - 1$ est asymptote oblique à la courbe. On peut même préciser que,

$$y(t) - (-x(t) - 1) = \frac{(t+1)^2}{t^2-t+1} > 0,$$

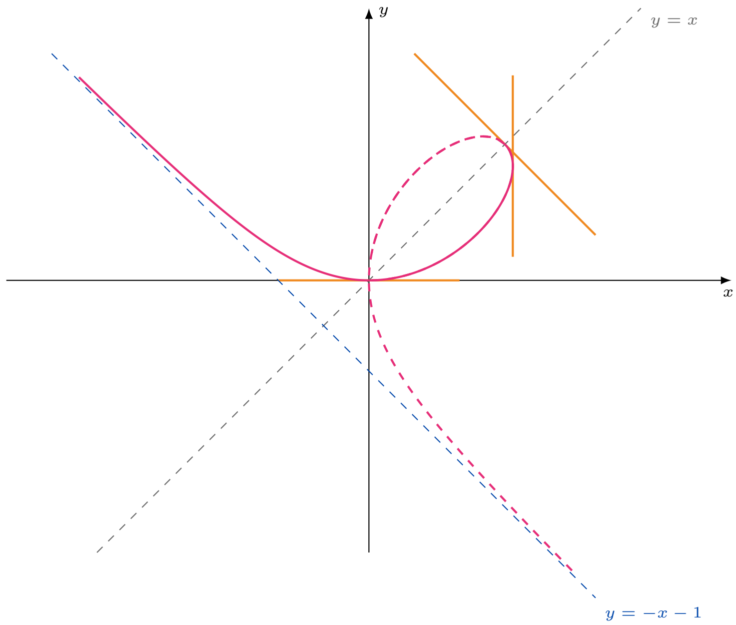
donc γ est au dessus de son asymptote.

En $M(0)$, la tangente à γ a pour coefficient directeur $\dfrac{y'(0)}{x'(0)} = 0$ et on a une tangente horizontale.

En $M\left(\sqrt[3]{\frac{1}{2}}\right)$, $\dfrac{y'(t)}{x'(t)} \underset{t\to\sqrt[3]{\frac{1}{2}}}{\longrightarrow} \infty$, on a donc une tangente verticale au point $M\left(\sqrt[3]{\frac{1}{2}}\right) = (\sqrt[3]{4}, \sqrt[3]{2})$.

En $M(1)$, la tangente à γ a pour coefficient directeur $\dfrac{y'(1)}{x'(1)} = -1$.

5. On peut tracer la courbe. On s'aide de l'asymptote et des tangentes. Puis on complète par symétrie par rapport à la première bissectrice d'équation $y = x$.



Exercice 3.9.

On cherche à représenté la courbe γ paramétrée par $M : t \longmapsto \begin{cases} x(t) &= \tan\left(\frac{t}{3}\right) \\ y(t) &= \sin(t) \end{cases}$.

1. a. Déterminer le domaine de définition D de M.

b. Calculer, pour $t \in D$, $M(t + 6\pi)$. En déduire qu'on peut réduire le domaine d'étude à un domaine D' centré en 0 qu'on explicitera.

c. Calculer, pour $t \in D'$, $M(-t)$. En déduire qu'on peut réduire le domaine d'étude à un ensemble D'' deux fois plus petit que le précédent.

d. Calculer alors, pour $t \in D''$, $M(3\pi - t)$ et en déduire un domaine d'étude le plus simple possible.

2. Dresser le tableau de variations des fonctions coordonnées x et$y$ sur ce domaine.

3. Vérifier qu'il n'y a pas de point singulier.

4. Montrer qu'il n'y a sur ce domaine qu'une seule branche infinie et préciser sa nature.

5. Déterminer les points éventuels où la tangente à la courbe est horizontale, ainsi que l'équation de la tangente au point $M(0)$.

6. Tracer soigneusement la courbe complète.

Exercice 3.10.

Étudier et tracer la courbe γ paramétrée par $\begin{cases} x(t) &= \cos^3(t) \\ y(t) &= \sin^3(t) \end{cases}$.

3.2 Enveloppe d'une famille de droites

Définition 3.7. **Famille de droites**

On appelle *famille de droites* (du plan) toute famille de droites $(\mathcal{D}_t)_{t\in I}$, où I est un intervalle (non trivial), de sorte que $\mathcal{D}_t$ soit paramétrée par

$$(\mathcal{D}_t) : A(t) + \lambda \overrightarrow{u}(t), \quad \lambda \in \mathbb{R},$$

avec A et $\overrightarrow{u}$ des fonctions vectorielles de classe $\mathcal{C}^1$ définies sur I.

Exemple 3.8.

Le cercle trigonométrique est paramétré par $M(t)\left\{\begin{array}{lcl} x & = & \cos(t) \\ y & = & \sin(t)\end{array}\right. \quad , \quad t \in [0, 2\pi].$

Pour tout $t \in [0; 2\pi]$, la tangente au point $M(t)$ est la droite

$$(\mathcal{D}_t) : \ M(t) + \lambda \frac{\overrightarrow{\mathrm{d}M}}{\mathrm{t}}(t), \qquad \lambda \in \mathbb{R}.$$

Comme M est de classe $\mathcal{C}^\infty$, il s'agit bien d'une famille de droites (avec la définition ci-dessus), en l'occurrence de la famille de droites de paramétrisation

$$\forall t \in [0, 2\pi], \qquad (\mathcal{D}_t) : \left\{\begin{array}{lcl} x & = & \cos(t) - \lambda \sin(t) \\ y & = & \sin(t) + \lambda \cos(t)\end{array}\right. , \qquad \lambda \in \mathbb{R}.$$

Plus généralement, on peut, pour toute courbe paramétrée γ régulière de classe $\mathcal{C}^1$ définir la famille de ses tangentes, paramétrée par

$$(\mathcal{D}_t) : \ M(t) + \lambda \frac{\overrightarrow{\mathrm{d}M}}{\mathrm{t}}(t), \qquad \lambda \in \mathbb{R}.$$

On peut alors se poser la question de la réciproque du procédé ; étant donnée une famille de droite

$$(\mathcal{D}_t)_{t \in I} : \ A(t) + \lambda \overrightarrow{u}(t), \qquad \lambda \in \mathbb{R},$$

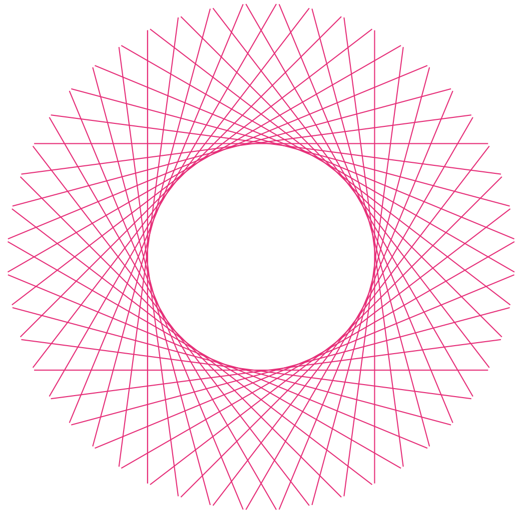
sil existe une (ou plusieurs) courbe(s) du plant dont $(\mathcal{D}_t)_{t\in I}$ est la famille des tangentes.

Définition 3.8. **Enveloppe**

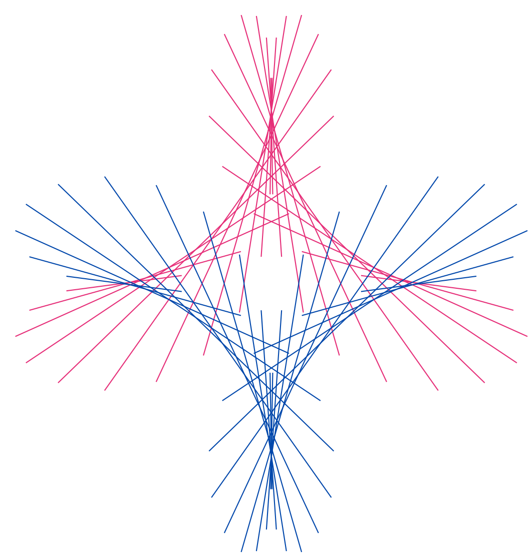
Soit $(\mathcal{D}_t)_{t\in I}$ une famille de droites du plan. On appelle **enveloppe** de $(\mathcal{D}_t)_{t\in I}$ toute courbe γ de classe $\mathcal{C}^1$ dont $(\mathcal{D}_t)_{t\in I}$ est exactement la famille des tangentes.

Exemple 3.9.

On représente ci-dessous deux enveloppes de famille de droites.



L'enveloppe de la famille des tangentes du cercle trigonométrique est le cercle trigonométrique lui-même.



L'astroïde comme enveloppe de la famille de ses tangentes

$$(\mathcal{D}_t) : \begin{pmatrix} \cos^3(t) \\ \sin^3(t) \end{pmatrix} + \lambda \begin{pmatrix} -3\sin(t)\cos^2(t) \\ 3\cos(t)\sin^2(t) \end{pmatrix}, \qquad \lambda \in \mathbb{R}.$$

Théorème 3.3.

Si $\mathcal{D}_t$ est la droite passant par le point $A(t)$ de vecteur directeur $\vec{u}(t)$, si les fonctions $t \mapsto A(t)$ et $t \mapsto \vec{u}(t)$ sont de classe $\mathscr{C}^1$ sur I et si, pour tout $t \in I$, la famille $(\vec{u}'(t), \vec{u}(t))$ est libre (ce qui se caractérise par un déterminant non nul), alors la famille de droites $(\mathcal{D}_t)_{t\in I}$ possède une et une unique enveloppe Γ.

La famille $(\mathcal{D}_t)$: $A(t) + \lambda \overrightarrow{u}(t)$ $(\lambda \in \mathbb{R})$ étant donnée, on cherche une courbe paramétrée γ telle que, pour tout t, $\mathcal{D}_t$ est tangente à γ et, pour tout point M de γ, il existe t tel que $\mathcal{D}_t$ est la tangente à γ en M.

La seconde condition implique alors que, pour tout t, il existe $\lambda = \lambda(t)$ tel que $A(t) + \lambda(t)\overrightarrow{u}(t)$ est un point de γ. On cherche donc un paramétrage de γ sous cette forme.

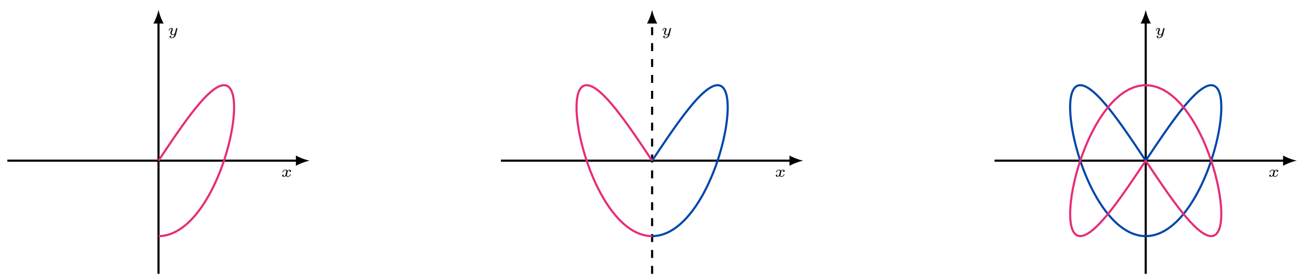
On veut de plus que $\mathcal{D}_t$ soit la tangente à γ en $M(t) = A(t) + \lambda(t)\overrightarrow{u}(t)$ c'est à dire que $\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{dt}}(t)$ et $\overrightarrow{u}(t)$ soient colinéaires, ce qui se traduit par

$$\det\left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{dt}}(t),\ \overrightarrow{u}(t)\right) = 0.$$

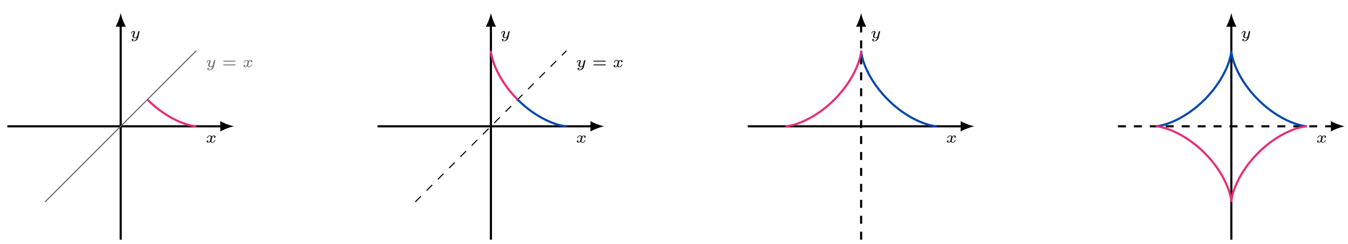
3.3 Quelques tracés

On propose les constructions très sommaires de quelques courbes paramétrées classiques pour illustrer et clore ce chapitre.

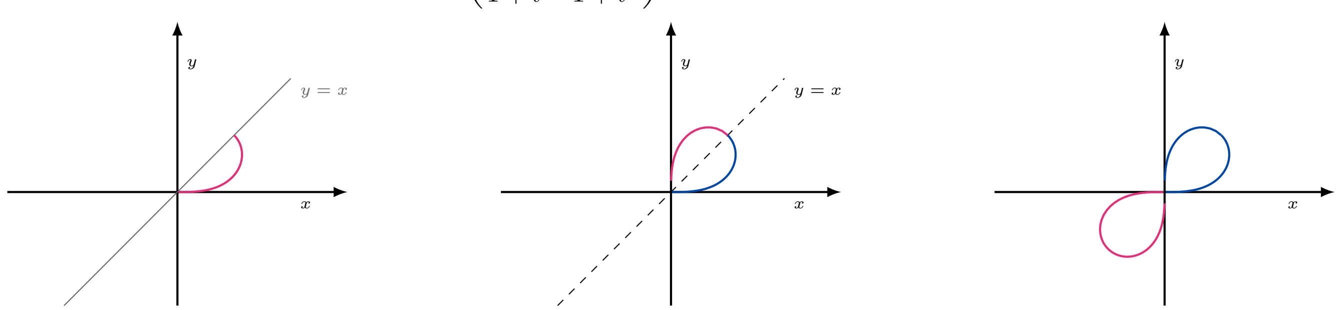
✗ Courbe de Lissajous $\gamma : t \mapsto (\sin(2t), \sin(3t))$.



✗ Astroïde $\gamma : t \mapsto (\cos^3(t), \sin^3(t))$.



✗ Lemniscate de Bernoulli : $\gamma : t \mapsto \left(\dfrac{t}{1+t^4}, \dfrac{t^3}{1+t^4}\right)$.



✗ Cardioïde : $\gamma \mapsto (2\cos(t) - \cos(2t), 2\sin(t) - \sin(2t))$.

4 Représentation matricielle des applications linéaires

Dans tout ce chapitre, la lettre $\mathbb{K}$ désignera l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$. Ce chapitre reprend naturellement toutes les notions déjà (ré-)introduites au cours du **Chapitre 2**.

La première section de ce chapitre propose des rappels sur les matrices et les éléments du calcul matriciel.

4.1 Rappels sur les matrices et le calcul matriciel

Dans cette section, n, p désigneront deux entiers strictement positifs.

4.1.1 Définition et exemples

Définition 4.1.
On appelle **matrice** à n lignes, p colonnes et à coefficients dans $\mathbb{K}$ tout tableau rectangulaire d'éléments de $\mathbb{K}$ possédant n lignes et p colonnes:

$$\begin{array}{cc} & \begin{array}{cccccc} & & & j\text{-ème colonne} & & \\ & & & \downarrow & & \end{array} \\ i\text{-ème ligne} \rightarrow & \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,j} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,j} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i,1} & a_{i,2} & \dots & a_{i,j} & \dots & a_{i,p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,j} & \dots & a_{n,p} \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K}). \end{array}$$

Lorsque l'on emploiera la notation ci-dessus, on dira que A est la matrice de *terme général* $a_{i,j}$.

- ✗ Une matrice à n lignes, p colonnes est aussi appelée une matrice de **taille** $n \times p$.
- ✗ L'ensemble des matrices de taille $n \times p$ à cœfficients dans $\mathbb{K}$ est noté $\mathcal{M}_{n,p}(\mathbb{K})$.
- ✗ Une matrice de taille $n \times n$ est appelée une matrice carrée de taille n ou une matrice carrée d'**ordre** n.
- ✗ L'ensemble $\mathcal{M}_{n,n}(\mathbb{K})$ des matrices carrées d'ordre n est noté $\mathcal{M}_n(\mathbb{K})$.
- ✗ Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Le cœfficient de A situé à l'intersection de la i-ième ligne et de la j-ième colonne sera appelé le cœfficient d'indice (i, j) de A, et sera noté $[A]_{i,j}$.

☞ L'important est de retenir que, par convention, le premier indice correspond au numéro de la ligne, et le second à celui de la colonne.

Définition 4.2. **Matrices particulières**

- ✗ La matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$ dont tous les cœfficients sont nuls est appelée **matrice nulle** de taille $n \times p$, et notée $0_{n,p}$.

- ✗ Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ vérifiant la formule $\forall (i,j) \in [\![1,n]\!]^2, (i \neq j \Rightarrow a_{i,j} = 0)$ est appelée **matrice diagonale**.
Elle est notée $\operatorname{diag}(a_{1,1}, \dots, a_{n,n})$:
$$\operatorname{diag}(a_{1,1}, \dots, a_{n,n}) = \begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & & a_{n,n} \end{pmatrix}.$$
Il s'agit des matrices dont les seuls cœfficients pouvant non nuls sont situés sur la diagonale du tableau. Attention, une matrice diagonale peut avoir des coefficients nuls aussi sur la diagonale. D'ailleurs la *matrice nulle* est une matrice diagonale.

- ✗ La matrice carrée $M \in \mathcal{M}_n(\mathbb{K})$ diagonale dont tous les coefficients diagonaux sont égaux à 1 est appelée **matrice identité** de taille n, et notée I_n :
$$I_n = \operatorname{diag}(1, 1, ..., 1).$$

- ✗ Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ vérifiant la formule $\forall (i,j) \in [\![1,n]\!]^2, (i > j \Rightarrow a_{i,j} = 0)$ est appelée **matrice triangulaire supérieure**.
$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix}.$$

- ✗ Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ vérifiant la formule $\forall (i,j) \in [\![1,n]\!]^2, (i < j \Rightarrow a_{i,j} = 0)$ est appelée **matrice triangulaire inférieure**.
$$\begin{pmatrix} a_{1,1} & 0 & 0 & \dots & 0 \\ a_{2,1} & a_{2,2} & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ a_{n,1} & \dots & & a_{n,n-1} & a_{n,n} \end{pmatrix}.$$
☞ Une matrice diagonale est donc à la fois une matrice triangulaire supérieure **et** inférieure.

- ✗ Soit $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$. La matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$ dont tous les cœfficients sont nuls, excepté le cœfficient d'indices (i,j) qui vaut 1, est notée $E_{i,j}$.
$$E_{i,j} = \begin{pmatrix} 0 & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & & & \vdots & & \vdots \\ \vdots & & & \vdots & & \vdots \\ 0 & \cdots & \cdots & 1 & \cdots & 0 \\ \vdots & & & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

4.1.2 Structure d'espace vectoriel

Étant données deux matrices de même taille, on définit leur somme de manière naturelle, c'est-à-dire cœfficient par cœfficient.

Définition 4.3.
Soient deux matrices $A=(a_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ et $B=(b_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ dans $\mathcal{M}_{n,p}(\mathbb{K})$. On note alors $A+B$ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de terme général $c_{i,j}$ définie par
$$\forall (i,j)\in [\![1,n]\!]\times[\![1,p]\!],\ c_{i,j}=a_{i,j}+b_{i,j}.$$

Si A et B n'ont pas la même taille, alors la somme $A+B$ n'est pas définie.

On peut aussi définir le produit d'une matrice par un scalaire, obtenu en multipliant chacun des termes de la matrice par le scalaire considéré.

Définition 4.4.
Soient $\lambda\in\mathbb{K}$ et $A=(a_{i,j})$ dans $\mathcal{M}_{n,p}(\mathbb{K})$. On note alors $\lambda\cdot A$ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de terme général $c_{i,j}$ défini par
$$\forall (i,j)\in [\![1,n]\!]\times[\![1,p]\!],\ c_{i,j}=\lambda a_{i,j}.$$

Proposition. 4.1.
$(\mathcal{M}_{n,p}(\mathbb{K}),+,\cdot)$ est un $\mathbb{K}$-espace vectoriel.

Proposition. 4.2.
La famille formée des matrices $E_{i,j}$, avec $(i,j)\in [\![1,n]\!]\times[\![1,p]\!]$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$, appelée la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$, de sorte que $\dim(\mathcal{M}_{n,p}(\mathbb{K}))=np$.

4.1.3 Produit de matrices

Contrairement à la somme, le produit de deux matrices n'est pas défini par le produit terme-à-terme de leurs cœfficients. On commence par définir le produit d'une matrice ligne par une matrice colonne.

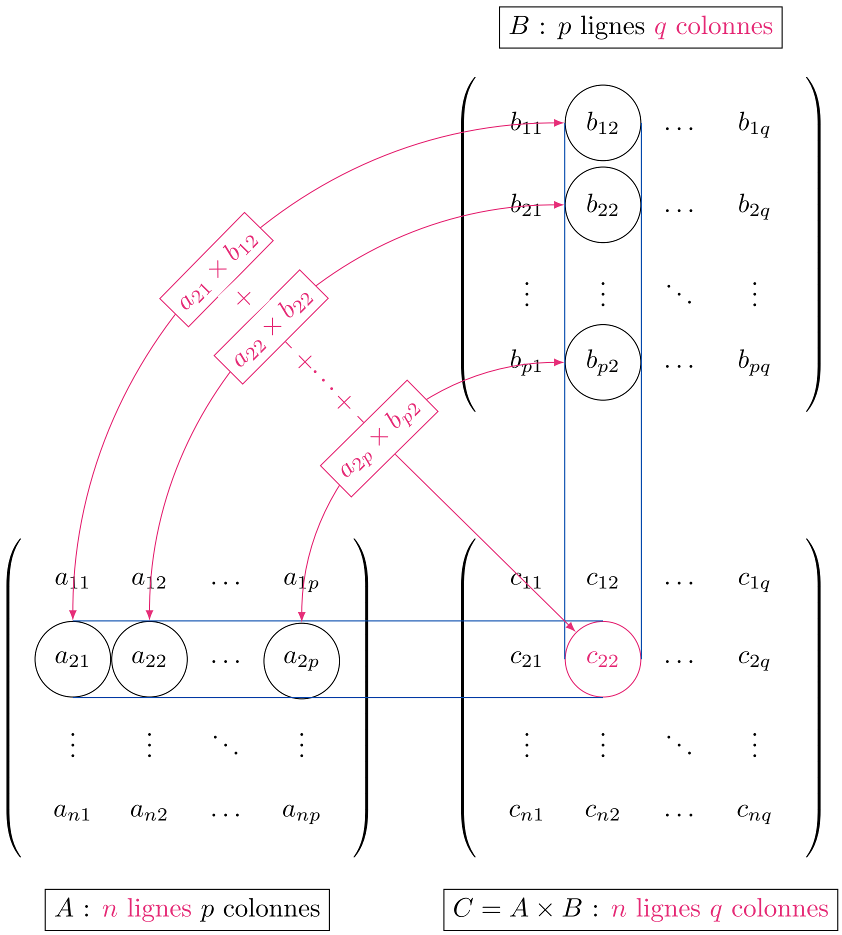
Définition 4.5. **Produit matriciel**
On définit le **produit** d'une matrice A de n lignes et p colonnes avec une matrice B de p lignes et q colonnes comme la matrice de n lignes et q colonnes suivante:
$$\forall A=(a_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq p}}\in\mathcal{M}_{n,p}(\mathbb{K}),\ \forall B=(b_{k,j})_{\substack{1\leq k\leq p\\1\leq j\leq q}}\in\mathcal{M}_{p,q}(\mathbb{K}),\ AB=\left(\sum_{k=1}^{p}a_{i,k}b_{k,j}\right)_{\substack{1\leq i\leq n\\1\leq j\leq q}}\in\mathcal{M}_{n,q}(\mathbb{K}).$$

Remarque 4.1.
On ne peut calculer le produit AB que si le nombre de colonnes de A égale le nombre de lignes de B.

En particulier, il peut être possible de calculer le produit AB mais pas le produit BA! Si les deux matrices sont carrées de même taille, on peut calculer AB et BA mais ces deux produits ne sont (en général) pas égaux.

Exemple 4.1.
Soit $M\in\mathcal{M}_n(\mathbb{K})$. On a $M0_n=0_nM=0_n$ et $I_n\times M=M\times I_n=M$.

✎ Le coefficient (i,j) du produit AB est le produit de la i-ème ligne de A avec la j-ème colonne de B. On peut disposer (**sur son brouillon**) les calculs ainsi:



Exercice 4.1.

Calculer les produits AB et BA avec $A = \begin{pmatrix} 1 & 2 & 4 \\ -2 & -3 & 1 \\ 3 & -5 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} -1 & 0 & -4 \\ 0 & 2 & 0 \\ 5 & -4 & 3 \end{pmatrix}$.

Proposition. 4.3. **Produit de matrices diagonales**

Soient $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n$ des scalaires. On a alors

$$\mathrm{diag}(\lambda_1, \dots, \lambda_n) \times \mathrm{diag}(\mu_1, \dots, \mu_n) = \mathrm{diag}\lambda_1\mu_1, \dots, \lambda_n\mu_n).$$

En d'autres termes, le produit de deux matrices diagonales est la matrice diagonale obtenue en faisant le produit terme-à-terme des cœfficients diagonaux.

Proposition. 4.4. **Distributivité du produit par rapport aux combinaisons linéaires**

Soient A, B, C trois matrices telles que B, C soient de même taille, et λ, μ des éléments de $\mathbb{K}$.

i. Si A est compatible à gauche de B on a $(\lambda \cdot A) \times B = A \times (\lambda \cdot B) = \lambda \cdot (A \times B)$.
ii. Si A est compatible à gauche de B et de C, on a $A \times (\lambda \cdot B + \mu \cdot C) = \lambda \cdot A \times B + \mu \cdot A \times C$.
iii. Si A est compatible à droite de B et de C, on a $(\lambda \cdot B + \mu \cdot C) \times A = \lambda \cdot B \times A + \mu \cdot C \times A$.
iv. Si A est compatible à gauche de B, et B est compatible à gauche de C, on a $A \times (B \times C) = (A \times B) \times C$.

4.1.4 Puissances de matrices

Définition 4.6. **Puissances de matrices**

Soient $k \in \mathbb{N}$ et A une matrice **carrée** de $\mathcal{M}_n(\mathbb{K})$. On appelle **puissance** k-ième de A, et on note A^k, la matrice

$$A^0 = I_n \quad \text{et} \quad A^k = \underbrace{A \times \cdots \times A}_{k \text{ fois}}.$$

Proposition. 4.5. **Puissances de matrices diagonales**

Soient $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n$ des scalaires. On a alors, pour tout entier $k \in \mathbb{N}$,

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n)^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k).$$

En d'autres termes, les puissances d'une matrice diagonale sont les matrices diagonales obtenues élevant à la puissance les cœfficients diagonaux.

Proposition. 4.6. **Formule du binôme**

Soient A et B deux matrices carrées de même taille qui **commutent**. Alors, pour tout entier $n \in \mathbb{N}$,

$$(A+B)^n = \sum_{i=0}^{n} \binom{n}{i} A^i B^{n-i}.$$

Exercice 4.2.

On considère les matrices $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ et $N = A - 2I$.

1. Expliciter N et calculer N^2 puis N^3. En déduire, pour $k \geq 3$ l'expression e N^k.
2. Donner une expression, pour tout $n \in \mathbb{N}$, de A^n en fonction de n de I, de N et N^2.
3. En déduire le tableau matriciel explicite de A^n.

4.1.5 Polynômes de matrices

Définition 4.7.

Soient $P = a_n X^n + ... + a_1 X + a_0 \in \mathbb{K}[X]$ et $A \in \mathcal{M}_p(\mathbb{R})$ une matrice carrée. On définit l'évaluation de P en A comme la matrice

$$P(A) = a_n A^n + ... + a_1 A + a_0 I_p \in \mathcal{M}_p(\mathbb{K}).$$

Lorsque $P(A) = 0_p$, on dit que P est un **polynôme annulateur** de A.

☞ On peut utiliser les propriétés des polynômes pour obtenir des informations sur les matrices (*inverse*, puissances).

Exercice 4.3.

Soient $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$ et $P = X^2 - 3X + 2$.

1. Calculer $P(A)$.
2. Soit $n \geq 3$. Effectuer la division euclidienne de X^n par P.
3. En déduire l'expression de A^n.

4.1.6 Matrices carrées inversibles

Définition 4.8. **Matrice inversible**

Soit $n \in \mathbb{N}^*$ et $M \in \mathcal{M}_n(\mathbb{K})$. On dit que M est **inversible** s'il existe une matrice M^{-1} de taille n telle que

$$M \times M^{-1} = M^{-1} \times M = I_n.$$

La matrice M^{-1} est alors appelée **l'inverse** de M. L'ensemble des matrices inversibles d'ordre n à cœfficients dans $\mathbb{K}$ est noté $\mathrm{GL}_n\mathbb{K}$, et est appelé **groupe linéaire d'ordre n sur** $\mathbb{K}$.

☞ Pour montrer qu'une matrice A est l'inverse d'une matrice M de taille n, il est inutile de prouver les deux formules $AM = I_n$ et $MA = I_n$: il suffit d'en prouver une, l'autre sera alors nécessairement vraie.

Exemple 4.2.

Soit $n \in \mathbb{N}^*$. La matrice I_n est inversible, de sorte que $I_n^{-1} = I_n$.

Proposition. 4.7.

Soient A et B deux matrices inversibles d'ordre n. Alors le produit AB est inversible et l'on a
$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proposition. 4.8.

Soit M une matrice inversible. Alors M^{-1} l'est également, et $\left(M^{-1}\right)^{-1} = M$.

Exercice 4.4.

Soit M une matrice carré d'ordre n et a, b, c des scalaires tels que $aM^2 + bM + cI_n = 0_n$. Montrer que si c est non nul, alors M est inversible.

Proposition. 4.9. **Matrice diagonale inversible**

Soient $n \in \mathbb{N}^*$, ainsi que $a_1, \ldots, a_n$ des scalaires non nuls. Alors la matrice $\operatorname{diag}(a_1, \ldots, a_n)$ est inversible, d'inverse $\operatorname{diag}\left(a_1^{-1}, \ldots, a_n^{-1}\right)$.

Proposition. 4.10. **Matrice inversible et équation matricielle**

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors, on a équivalence

i. A est inversible

ii. $\forall\, Y \in \mathcal{M}_{n,1}(\mathbb{K})$, l'équation matricielle $AX = Y$ d'inconnue $X \in \mathcal{M}_{n,1}(\mathbb{K})$ admet une unique solution.

Auquel cas, on a
$$X = A^{-1}Y.$$

On l'a mentionné ci-avant, pouvoir inverser une matrice A revient à pouvoir trouver une unique solution au système $AX = Y$, ce qui revient à dire que ce système est de Cramer.
La méthode du pivot de Gauss pour la résolution de système permet alors d'obtenir plusieurs résultats.

Proposition. 4.11.

Soit A une matrice carrée de taille n.
Toute matrice obtenue à partir de A par des *opérations sur les lignes* (introduites dans la méthode du pivot de Gauss) a les même propriétés d'inversibilité que A.

Une première conséquence de ce résultat est le corollaire suivant qui caractérise les matrices triangulaires inversibles.

Proposition. 4.12. **Caractérisation de l'inversibilité des matrices triangulaires**

Soit $T \in \mathcal{M}_n(\mathbb{K})$ une matrice triangulaire (inférieure ou supérieure). Alors, T est inversible si et seulement si tous ses coefficients diagonaux sont non nuls.

Ce critère ne marche que si la matrice est triangulaire.
Il n'est pas question de dire qu'une matrice quelconque sans zéro sur la diagonale est inversible !
Par exemple,
$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$
n'est pas inversible (et pourtant n'a aucun zéro sur la diagonale).

4.1.7 Transposition d'une matrice

Définition 4.9. **Transposée**

Soient $n, p \geqslant 1$ des entiers, et $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ une matrice de taille $n \times p$.
On appelle alors **transposée** de A, et on note ${}^tA$ ou $A^\top$, la matrice B de taille $p \times n$ dont le terme général $b_{i,j}$ est donné par $b_{i,j} = a_{j,i}$ pour $(i,j) \in [\![1,p]\!] \times [\![1,n]\!]$.

La transposée d'une matrice M est la matrice dont les lignes sont les colonnes de M ou, ce qui revient au même, la matrice dont les colonnes sont les lignes de M.

Exemple 4.3.

- ✗ La transposée de $\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$ est $\begin{pmatrix}1&4&7\\2&5&8\\3&6&9\end{pmatrix}$, et celle de $\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}$ est $\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}$.
- ✗ La transposée de $\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}$ est $\begin{pmatrix}x_1&\cdots&x_n\end{pmatrix}$ et celle de $\begin{pmatrix}x_1&\cdots&x_p\end{pmatrix}$ est $\begin{pmatrix}x_1\\\vdots\\x_p\end{pmatrix}$.

Remarque 4.2.
Pour calculer la transposée d'une matrice, il suffit donc d'opérer une symétrie sur ces termes par rapport à la diagonale de la matrice, ce qui revient à intervertir les indices de lignes et de colonnes dans les cœfficients de celle-ci.

Proposition. 4.13.
Soient $\lambda \in \mathbb{K}$ et M, N deux matrices. On a

i. $\left(M^\top\right)^\top = M$.
ii. Si M et N ont même taille on a $(\lambda M + \mu N)^\top = \lambda M^\top + \mu N^\top$ (en d'autres termes, $M \mapsto M^\top$ est donc une application linéaire).
iii. Si M est compatible à gauche de N on a $(MN)^\top = N^\top \cdot M^\top$.
iv. Si M est inversible, alors $M^\top$ l'est également, de sorte que $\left(M^\top\right)^{-1} = \left(M^{-1}\right)^\top$.

Définition 4.10. **Matrices symétriques et antisymétriques**

- ✗ Une matrice carrée M est dite **symétrique** si $M^\top = M$. On notera $\mathcal{S}_n(\mathbb{K})$ l'ensemble des matrices symétriques d'ordre n.
- ✗ Une matrice carrée M est dite **antisymétrique** si $M^\top = -M$. On notera $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices antisymétriques d'ordre n.

Remarque 4.3.

- ✗ Une matrice qui n'est pas carrée ne peut pas être symétrique.
- ✗ Les termes diagonaux d'une matrice antisymétrique A sont nuls, puisque l'on a $a_{i,i} = -a_{i,i}$ pour tout i.

Exemple 4.4.

- ✗ Toute matrice symétrique de taille 3 est de la forme $\begin{pmatrix}\alpha&a&b\\a&\beta&c\\b&c&\gamma\end{pmatrix}$, avec $a, b, c, \alpha, \beta, \gamma$ des scalaires.
- ✗ Toute matrice antisymétrique de taille 3 est de la forme $\begin{pmatrix}0&-a&-b\\a&0&-c\\b&c&0\end{pmatrix}$, avec a, b, c des scalaires.

Proposition. 4.14.
$\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$sont des sous-espaces vectoriels supplémentaires de $\mathcal{M}_n(\mathbb{K})$: $\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$.

Exercice 4.5.
Soit $M \in \mathcal{M}_n(\mathbb{K})$. Montrer que $M \times M^\top$ et $M^\top + M$ sont symétriques.

4.1.8 Matrices semblables, matrices diagonalisables

Définition 4.11. **Matrices semblables**

Soient $n \geqslant 1$ un entier et A, B des matrices carrées d'ordre n. On dit que A est **semblable** à B s'il existe une matrice inversible P telle que :

$$A = PBP^{-1}.$$

Proposition. 4.15.

Soient N et P des matrices carrées de même taille, de sorte que P soit inversible. Alors pour tout entier n, on a

$$\left(PNP^{-1}\right)^n = PN^nP^{-1}.$$

☞ Lorsque deux matrices sont semblables, il est aisé de calculer les puissances de l'une à partir des puissances de l'autre. Une matrice A étant donnée, il est donc intéressant de chercher une matrice B **la plus simple possible** qui est semblable à A; typiquement on cherchera une matrice B diagonale (qui présente l'intérêt de posséder des puissances très simples à calculer).

Définition 4.12. **Matrice diagonalisable**

Soient $n \in \mathbb{N}^*$ et M une matrice carrée de taille n. On dit que M **est diagonalisable** s'il existe une matrice diagonale D, ainsi qu'une matrice inversible P de taille n tels que

$$M = PDP^{-1},$$

en d'autres termes, M est diagonalisable si elle est semblable à une matrice diagonale.

☞ L'étude des matrices *diagonalisables* sera faite dans un chapitre ultérieur.

4.1.9 Trace d'une matrice

Définition 4.13. **Trace d'une matrice**

Soit $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ une matrice de taille n. On appelle **trace** de A, et on note $\operatorname{tr}(A)$, la somme de ses termes diagonaux. En d'autres termes:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{i,i}.$$

Proposition. 4.16. **Propriétés de la trace**

Soient A, B deux matrices de taille n, et $\lambda, \mu \in \mathbb{K}$. On a alors les résultats suivants.

i. $\operatorname{tr}(\lambda A + \mu B) = \lambda \operatorname{tr}(A) + \mu \operatorname{tr}(B)$ (en d'autres termes la trace est une forme linéaire sur $\mathcal{M}_n(K)$) ;
ii. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Exercice 4.6.

Montrer que si A et B sont deux matrices semblables de $\mathcal{M}_n(\mathbb{K})$, alors elles ont la même trace.

4.2 Représentation matricielle des applications linéaires en dimension finie

Dans cette section, les lettres E, F, G désigneront des $\mathbb{K}$-espaces vectoriels de dimension **finie**.

4.2.1 Matrice d'un vecteur et d'une famille dans une base

Définition 4.14.

Soient $n = \dim(E)$, et $\mathcal{B}$ une base de E et $u \in E$, dont les coordonnées dans la base $\mathcal{B}$ sont $(u_1, \ldots, u_n)$. On appelle **matrice** de u dans la base $\mathcal{B}$, et l'on note $\operatorname{Mat}_{\mathcal{B}}(u)$ la matrice colonne de taille $(n, 1)$ formée des coordonnées de u dans cette base:

$$\operatorname{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

Exemple 4.5.

Soit u le vecteur de $\mathbb{K}^2$ tel que $u = i + 2j$ où (i, j) est la base canonique de $\mathbb{K}^2$. La matrice de u dans la base (i, j) est $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, celle dans la base $(i + j, i)$ est $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

☞ On peut identifier tout vecteur de $\mathbb{K}^2$ avec sa matrice dans la base canonique.

Exemple 4.6.

Soit le polynôme $P = X^2 - X + 2 \in \mathbb{K}_2[X]$. Sa matrice dans la base $(1, X, X^2)$ est $\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$.

Définition 4.15.

Soient $n = \dim(E)$, et $\mathcal{B}$ une base de E et $(u_1, \ldots, u_p)$ une famille finie de vecteurs de E. On appelle **matrice de la famille** $(u_1, \ldots, u_p)$ dans la base $\mathcal{B}$, et l'on note $\mathrm{Mat}_{\mathcal{B}}(u_1, \ldots, u_p)$ la matrice de taille (n, p) dont la j-ième colonne est la matrice de u_j dans la base $\mathcal{B}$.

4.2.2 Matrice d'une application linéaire

Définition 4.16. **Matrice d'une application linéaire**

Soit $f \in \mathcal{L}(E, F)$, $\mathcal{B} = (e_1, \ldots, e_p)$ une base de E, et $\mathcal{B}' = ({e_1}', \ldots, {e_n}')$ une base de F.
On appelle **matrice de** f dans les bases $\mathcal{B}, \mathcal{B}'$, et l'on note $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$, la matrice dont la j-ième colonne est $\mathrm{Mat}_{\mathcal{B}'}(f(e_j))$, c'est-à-dire la matrice de la famille $(f(e_1), \ldots, f(e_p))$ dans la base $\mathcal{B}'$.

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \begin{array}{c} \begin{array}{ccc} f(e_1) & & f(e_p) \\ \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \end{array} \begin{array}{l} \leftarrow \text{coordonnée selon } e'_1 \\ \\ \leftarrow \text{coordonnée selon } e'_n \end{array} .$$

Remarque 4.4.

À une application linéaire sont donc associées différentes matrices, selon les bases $\mathcal{B}, \mathcal{B}'$ choisies.

☞ Dans le cas où f est un endomorphisme et où l'on choisit $\mathcal{B} = \mathcal{B}'$, la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f)$ est appelée **matrice de** f **dans la base** $\mathcal{B}$. On peut donc donner la définition restreinte suivante.

Définition 4.17. **Matrice d'un endomorphisme**

Soit $f \in \mathcal{L}(E)$ et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E.
On appelle **matrice de** f **dans la base** $\mathcal{B}$, et l'on note $\mathrm{Mat}_{\mathcal{B}}(f)$, la matrice dont la j-ième colonne est $\mathrm{Mat}_{\mathcal{B}}(f(e_j))$, c'est-à-dire la matrice de la famille $(f(e_1), \ldots, f(e_n))$ dans la base $\mathcal{B}$.

☞ Pour calculer la matrice d'un endomorphisme f dans une base $\mathcal{B}$, on commence par calculer les images des vecteurs de $\mathcal{B}$ par f, puis on décompose celles-ci dans la base $\mathcal{B}$.
Si $\mathcal{B}$ est la base canonique de E, on note $\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{Mat}_{\mathrm{can}}(f)$.

Exemple 4.7.

Supposons $\dim(E) = n$. Alors dans **n'importe quelle base** de E:

- ✗ la matrice de id_E est la matrice I_n;
- ✗ la matrice de $0_{\mathcal{L}(E)}$ est la matrice 0_n;
- ✗ si $\lambda \in \mathbb{K}$, la matrice de l'homothétie $h_\lambda : x \mapsto \lambda x$ de rapport λ est la matrice scalaire λI_n.

Exercice 4.7.

Pour tout $P \in \mathbb{K}[X]$ on pose $\Phi(P) = P + X^2 P''$.

1. Montrer $\Phi \in \mathcal{L}(\mathbb{K}_2[X])$.
2. Donner la matrice de Φ dans la base canonique de $\mathbb{K}_2[X]$.

Proposition. 4.17. **Caractérisation d'une application linéaire par une matrice**

Soient $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{B}' = (e_1{}', \dots, e_n{}')$ une base de F. Alors, pour toute matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, il existe une unique application linéaire $f : E \to F$ telle que

$$M = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f).$$

Remarque 4.5.

✗ Une fois que l'on a fixé deux bases $\mathcal{B}, \mathcal{B}'$, chaque matrice correspond donc à une unique application linéaire. Mais dans l'absolu, une matrice correspond à différentes applications linéaires: tout dépend du couple de bases dans lequel on choisit d'interpréter cette matrice!

✗ Étant donnée $M \in \mathcal{M}_{n,p}(\mathbb{K})$, il existe donc une unique application $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ telle que $M = \mathrm{Mat}_{\mathrm{can}}(f)$ (il s'agit de l'application qui envoie le j-ième vecteur de la base canonique de $\mathbb{K}^p$ sur la j-ième colonne de M), appelée l'*application linéaire canoniquement associée à M*.

Proposition. 4.18. **Stabilité et représentation par blocs**

Soient F, G deux sous-espaces supplémentaires de E, de dimensions respectives $p, n-p$ et n, et une base adaptée $\mathcal{B} = (e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ à la décomposition $E = F \oplus G$. Soit encore $f \in \mathcal{L}(E)$.
Alors, F est stable par f si et seulement si la matrice M de f dans la base $\mathcal{B}$ est de la forme:

$$\begin{matrix} \\ \\ \\ n-p \text{ lignes} \left\{ \vphantom{\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}} \right. \end{matrix} \begin{pmatrix} * & \cdots & * & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ * & \cdots & * & . & & . \\ 0 & \cdots & 0 & . & & . \\ & & & \vdots & & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix}$$

$$\underbrace{\hphantom{0 \cdots 0}}_{p \text{ colonnes}} \quad \underbrace{\hphantom{* \cdots *}}_{n-p \text{ colonnes}}$$

ce qu'on peut aussi écrire

$$M = \left(\begin{array}{c|c} A & B \\ \hline 0_{n-p,p} & C \end{array} \right),$$

où $A \in \mathcal{M}_p(\mathbb{K}), B \in M_{p,n-p}(\mathbb{K}), C \in \mathcal{M}_{n-p}(\mathbb{K})$ et $0_{n-p,p}$ est la matrice nulle de $\mathcal{M}_{n-p,p}(\mathbb{K})$.
Une telle matrice est dite **définie par blocs**.

Exercice 4.8.

Montrer que l'endomorphisme canoniquement associé à la matrice $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ possède un plan et une droite stable.

4.2.3 Correspondance entre opérations sur les applications linéaires et calcul matriciel

Opération d'une application linéaire sur un vecteur via les matrices associées

Proposition. 4.19.

Soient $\Phi : E \to F$ une application linéaire, $\mathcal{B}$ une base de E et $\mathcal{B}'$ une base de F. Alors

$$\forall u \in E, \ \mathrm{Mat}_{\mathcal{B}'}(\Phi(u)) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\Phi) \times \mathrm{Mat}_{\mathcal{B}}(u)$$

Remarque 4.6.

- ✗ Dans le cas d'un endomorphisme $\Phi \in \mathcal{L}(E)$, on a, pour toute base $\mathcal{B}$ de E,
$$\forall u \in E, \mathrm{Mat}_{\mathcal{B}}\left(\Phi(u)\right) = \mathrm{Mat}_{\mathcal{B}}(\Phi) \times \mathrm{Mat}_{\mathcal{B}}\left(u\right).$$
- ✗ Soit $u \in E$ et $\Phi \in \mathcal{L}(E,F)$. Si $\mathcal{B}, \mathcal{B}''$ sont deux bases de E, on peut effectuer le produit $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\Phi) \times \mathrm{Mat}_{\mathcal{B}''}(u)$ puisque ces matrices sont compatibles, mais le résultat ne sera *a priori* pas égal à la matrice de $\Phi(u)$, dans quelque base que ce soit.

Exercice 4.9.

Soit $\Phi \in \mathcal{L}(\mathbb{K}_2[X])$ admettant $M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ pour matrice dans la base $(1, 1+X, X^2)$.

1. Donner une expression explicite de $\Phi(P)$ pour $P \in \mathbb{K}_2[X]$, et montrer que Φ est un projecteur.
2. Donner une base de $\mathrm{Ker}(\Phi)$.

Le résultat suivant permet de calculer le noyau et l'image d'une application linéaire à partir d'une matrice représentant celle-ci.

Proposition. 4.20.
Supposons $\dim(F) = n$, et considérons des bases $\mathcal{B}$ de E et $\mathcal{B}'$ de F.
Soit encore $f \in \mathcal{L}(E,F)$, et M la matrice de f dans les bases $\mathcal{B}, \mathcal{B}'$. On a:

i. Pour tout $u \in E$: $f(u) = 0_F$ si et seulement si $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) \times \mathrm{Mat}_{\mathcal{B}}(u) = 0_{n,1}$.
ii. Pour tout $v \in F$: $v \in \mathrm{Im}(f)$ si et seulement si $\exists u \in E,\ \mathrm{Mat}_{\mathcal{B}'}(v) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) \times \mathrm{Mat}_{\mathcal{B}}(u)$.

☞ Soit $f \in \mathcal{L}(E,F)$ et $M \in \mathcal{M}_{n,p}(\mathbb{K})$ la matrice de f dans des bases $\mathcal{B}, \mathcal{B}'$.
On pose alors $\mathrm{Ker}(M) = \{X \in \mathcal{M}_{p,1}(\mathbb{K});\ MX = 0_{n,1}\}$.
On en déduit que pour déterminer une base de $\mathrm{Ker}(f)$, on peut commencer par déterminer une base $U_1, \ldots, U_k$ de $\mathrm{Ker}(M)$: $\mathrm{Ker}(f)$ admettra alors $u_1, \ldots, u_k$ pour base, où chacun des u_i est tel que $U_i = \mathrm{Mat}_{\mathcal{B}}(u_i)$.

Sommes, compositions et inverses d'applications linéaires via les matrices associées

Les résultats suivants indiquent que le calcul algébrique sur les applications linéaires correspond au calcul matriciel: la correspondance entre matrices et applications linéaires présente donc une contrepartie opératoire. Chaque problème d'algèbre linéaire peut ainsi être considéré comme un problème matriciel, ou comme un problème sur les applications linéaires: il peut donc être résolu en utilisant indifféremment les techniques de l'une ou l'autre de ces théories.

Proposition. 4.21.
On munit E, F, G de bases respectives $\mathcal{B}_E, \mathcal{B}_F, \mathcal{B}_G$. Soient encore $\lambda \in \mathbb{K}$, $k \in \mathbb{N}$, $(f,g) \in \mathcal{L}(E,F)^2$, $h \in \mathcal{L}(F,G)$ et $h_2 \in \mathcal{L}(E)$. On a alors les résultats suivants.

i. $\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(\lambda f) = \lambda \cdot \mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(f)$;
ii. $\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(f+g) = \mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(f) + \mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(g)$;
iii. $\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_G}(h \circ f) = \mathrm{Mat}_{\mathcal{B}_F,\mathcal{B}_G}(h) \times \mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(f)$;
iv. $\mathrm{Mat}_{\mathcal{B}_E}(h_2^k) = \left(\mathrm{Mat}_{\mathcal{B}_E}(h_2)\right)^k$.

Exemple 4.8.
Supposons $\dim(E) = n$, et soit $f \in \mathcal{L}(E)$ tel que $f^3 - 2f = 0_{\mathcal{L}(E)}$. Alors si M désigne sa matrice dans une base quelconque, celle-ci vérifie $M^3 - 2M = 0_n$.

Théorème 4.22.
Supposons $\dim(E) = p$, $\dim(F) = n$ et considérons des bases $\mathcal{B}$ de E et $\mathcal{B}'$ de F.
Alors, l'application

$$\begin{array}{ccc} \mathcal{L}(E,F) & \longrightarrow & \mathcal{M}_{n,p}(\mathbb{K}) \\ f & \longmapsto & \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) \end{array}$$

est un isomorphisme d'espaces vectoriels.

Théorème 4.23.
On a $\dim(\mathcal{L}(E,F)) = \dim(E) \cdot \dim(F)$.

Preuve. Soit $p = \dim E$ et $n = \dim F$. On a prouvé que $\mathcal{L}(E,F)$ et $\mathcal{M}_{n,p}(\mathbb{K})$ sont isomorphes, donc ils ont même dimension. D'où le résultat, car $\dim(\mathcal{M}_{n,p}(\mathbb{K})) = n \times p = \dim(E) \times \dim(F)$. ☐

Le résultat suivant prouve que les isomorphismes correspondent aux matrices inversibles.

Proposition. 4.24.
Soient $\mathcal{B}$ une base de E et $\mathcal{B}'$ une base de F. Soit encore $f \in \mathcal{L}(E,F)$.
Alors, f est inversible si et seulement si sa matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ l'est, et dans ce cas on a

$$\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(f^{-1}) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)^{-1}.$$

☞ En particulier dès qu'un endomorphisme est représenté par une matrice inversible, il est bijectif.

4.2.4 Changement de base

Définition 4.18.
Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E.
On appelle **matrice de passage** de $\mathcal{B}$ à $\mathcal{B}'$, et l'on note $P_{\mathcal{B},\mathcal{B}'}$, la matrice définie par

$$P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E).$$

En d'autres termes, la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$: Notant $\mathcal{B} = (e_1, ..., e_n)$ et $\mathcal{B}' = (e'_1, ..., e'_n)$, on forme la matrice de passage comme suit.

$$P_{\mathcal{B},\mathcal{B}'} = \begin{array}{c} \begin{array}{ccc} e'_1 & & e'_n \\ \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix} \end{array} \begin{array}{l} \leftarrow \text{coordonnée selon } e_1 \\ \\ \leftarrow \text{coordonnée selon } e_n \end{array} \quad ,$$

et

$$P_{\mathcal{B}',\mathcal{B}} = \begin{array}{c} \begin{array}{ccc} e_1 & & e_n \\ \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} b_{1,1} & \dots & b_{1,p} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,p} \end{pmatrix} \end{array} \begin{array}{l} \leftarrow \text{coordonnée selon } e'_1 \\ \\ \leftarrow \text{coordonnée selon } e'_n \end{array}$$

où

$$e'_k = \sum_{i=1}^{n} a_{i,k} e_i, \qquad e_j = \sum_{i=1}^{n} b_{i,j} e'_i.$$

Exemple 4.9.

La matrice de passage de la base canonique de $\mathbb{R}^3$ à $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ est $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$.

Proposition. 4.25.
Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E. Pour tout $u \in E$ on a $\mathrm{Mat}_{\mathcal{B}}(u) = P_{\mathcal{B},\mathcal{B}'} \times \mathrm{Mat}_{\mathcal{B}'}(u)$.

☞ La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ permet ainsi de passer des coordonnées dans la base $\mathcal{B}'$ à celles dans la base $\mathcal{B}$ (*i.e.* d'exprimer **les anciennes coordonnées en fonction des nouvelles**).

Remarque 4.7.
De prime abord il semble que le nom de ces matrices ait été mal choisi... mais en pratique la définition est la bonne. On retiendra par exemple les phrases suivantes:

- ✗ la matrice de passage P d'une ancienne base vers une nouvelle est la matrice des vecteurs de la nouvelle base dans l'ancienne (ce qui est somme toute logique: on décrit les nouveaux vecteurs en fonction de ceux déjà connus);
- ✗ mais cette matrice P donne les anciennes coordonnées en fonction des nouvelles (pour obtenir les nouvelles en fonction des anciennes il faudra utiliser P^{-1}, voir la **Proposition 4.26**).

Exemple 4.10.
Soit $f \in \mathrm{Gl}(E)$ et $\mathcal{B}$ une base de E. Alors $\mathrm{Mat}_{\mathcal{B}}(f)$ est la matrice de passage de la base $\mathcal{B}$ à la base $f(\mathcal{B})$. Autrement dit, toute matrice inversible est une matrice de passage.

On démontre que, réciproquement, toute matrice de passage est inversible.

Proposition. 4.26.
Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E. La matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ est inversible et son inverse est la matrice de passage $P_{\mathcal{B}',\mathcal{B}}$.

Exemple 4.11.
La matrice de passage de $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right)$ à la base canonique de $\mathbb{R}^3$ est

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

Si $u \in \mathbb{R}^3$ admet pour coordonnées x, y, z dans la base canonique, alors ses coordonnées dans la base $\mathcal{B}$ sont données par:

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Exercice 4.10.
Déterminer la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

1. $\mathcal{B}$ est la base canonique de $\mathbb{K}_2[X]$ et $\mathcal{B}' = (-X^2 + 2X + 1, X + 1, X^2 + 2)$.
2. $\mathcal{B} = (1, X - 2, X^2 - 3X + 3)$ et $\mathcal{B}'$ est la base canonique de $\mathbb{K}_2[X]$.

Proposition. 4.27. **Formule de changement de base**
Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Pour tout $f \in \mathcal{L}(E)$ on a:

$$\mathrm{Mat}_{\mathcal{B}}(f) = P \times \mathrm{Mat}_{\mathcal{B}'}(f) \times P^{-1}$$

Proposition. 4.28.
Deux matrices sont semblables si et seulement si ce sont les matrices d'un même endomorphisme.

Théorème 4.29. **Trace d'un endomorphisme**
Toutes les matrices associées à un endomorphisme ont la même trace. Celle-ci est appelée **trace** de f et est notée $\mathrm{tr}(f)$.

Exercice 4.11.
Montrer que la trace d'un projecteur est égale à son rang.

4.3 Calcul effectif du rang et de l'inverse d'une matrice

4.3.1 Opérations élémentaires

Définition 4.19.
Soit $A \in \mathcal{M}_{n,p}\mathbb{K}$; on note $L_1, \ldots, L_n$ ses lignes.
On appelle opérations élémentaires sur les lignes de A les transformations suivantes, où $(i,k) \in [\![1,n]\!]^2$ avec $i \neq k$, $\lambda \in \mathbb{K}^*$ et $\alpha \in \mathbb{K}$:

- ✗ Remplacer la ligne L_i par la ligne λL_i, ce qui est noté $L_i \leftarrow \lambda L_i$;
- ✗ Échanger les lignes L_i et L_k, ce qui est noté $L_i \leftrightarrow L_k$;
- ✗ Remplacer la ligne L_i par la ligne $L_i + \alpha L_k$, ce qui est noté $L_i \leftarrow L_i + \alpha L_k$.

Définition 4.20. **Matrices élémentaires**
Soient $(i,k) \in [\![1,n]\!]^2$ tels que $i \neq k$, $\lambda \in \mathbb{K}^*$ et $\alpha \in \mathbb{K}$. On appelle **matrices élémentaires** les matrices de $\mathcal{M}_n\mathbb{K}$ suivantes.

i. On appelle **matrice de dilatation** toute matrice de $\mathcal{M}_n\mathbb{K}$ de la forme : $D_{n,i}(\lambda) = I_n + (\lambda - 1)E_{i,i}$.

$$D_{n,i}(\lambda) = \begin{array}{c} \begin{array}{ccccccc} &&&i&&& \\ &&&\downarrow&&& \end{array} \\ \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array} .$$

ii. On appelle **matrice de transposition** toute matrice de la forme: $P_{n,i,k} = I_n + (E_{i,k} + E_{k,i} - E_{i,i} - E_{k,k})$.

$$P_{n,i,k} = \begin{array}{c} \begin{array}{ccccccccc} &&i&&&&k&& \\ &&\downarrow&&&&\downarrow&& \end{array} \\ \begin{pmatrix} \ddots &&&&&&&& \\ & 1 &&&&&&& \\ && 0 &&&& 1 && \\ &&& 1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& 1 &&& \\ && 1 &&&& 0 && \\ &&&&&&& 1 & \\ &&&&&&&& \ddots \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow k \\ \\ \\ \end{array} .$$

iii. On appelle **matrice de transvection** toute matrice de $\mathcal{M}_n\mathbb{K}$ de la forme : $T_{n,i,k}(\alpha) = I_n + \alpha E_{i,k}$.

$$T_{n,i,k}(\alpha) = \begin{array}{c} \begin{array}{ccccccc} &&&&k&& \\ &&&&\downarrow&& \end{array} \\ \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 && \alpha && \\ &&& \ddots &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \\ \end{array} .$$

Proposition. 4.30.

Soit $A \in \mathcal{M}_{n,p}\mathbb{K}$; on note $L_1, \dots, L_n$ ses lignes.
Soit $(i,k) \in [\![1,n]\!]^2$ tels que $i \neq k$, $\lambda \in \mathbb{K}^*$ et $\alpha \in \mathbb{K}$. On a les résultats suivants.

i. Remplacer la ligne L_i par la ligne λL_i revient à multiplier A par $D_{n,i}(\lambda)$ à gauche.
ii. Échanger les lignes L_i et L_k revient à multiplier A par $P_{n,i,k}$ à gauche.
iii. Remplacer la ligne L_i par la ligne $L_i + \alpha L_k$ revient à multiplier A par $T_{n,i,k}(\alpha)$ à gauche.

Remarque 4.8.

On définit de même les opérations élémentaires sur les colonnes d'une matrice, et on introduit ainsi la notion de matrices **équivalentes par colonnes**. On montre alors qu'*effectuer des opérations élémentaires sur les colonnes d'une matrice revient à multiplier celle-ci par des matrices élémentaires à droite*, de même que ci-dessus.

4.3.2 Rang d'une matrice

Définition 4.21. **Rang d'une matrice**

Soit M une matrice. On appelle rang de M, et on note $\mathrm{rg}(M)$ le rang de la famille formée des vecteurs colonnes de M.

Proposition. 4.31.

Soit M une matrice. Si f est une application linéaire dont M est la matrice (dans un couple de bases quelconques), alors $\mathrm{rg}(M) = \mathrm{rg}(f)$.

Exercice 4.12.

Déterminer sans calcul l'image et le noyau de l'endomorphisme de $\mathbb{K}^3$ dont la matrice dans la base canonique est

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

☞ Pour déterminer l'image et le noyau d'une application linéaire $f \in \mathcal{L}(E,F)$, on peut utiliser la méthode suivante.

- ✗ Donner la matrice M de f dans un couple de bases quelconques.
- ✗ Appliquer l'algorithme de la base incomplète pour déterminer le rang de la famille formée par les vecteurs colonnes de M, et ainsi une base de $\mathrm{Im}(f)$.
- ✗ On déduit alors du calcul précédent $\dim(\mathrm{Ker}(f)) = \dim(E) - \dim(\mathrm{Im}(f))$, et pour déterminer une base du noyau, il suffit de déterminer $\dim(\mathrm{Ker}(f))$ à l'aide de combinaisons linéaires nulles des colonnes de M. En n'oubliant pas de vérifier qu'à ces combinaisons linéaires ne sont pas associés des vecteurs formant une famille liée.

Définition 4.22.

i. Une matrice est dite *échelonnée par lignes* si, après la première ligne, chaque ligne non nulle commence par strictement plus de zéros que la ligne précédente.
ii. On appelle **pivot** le premier cœfficient non nul de chaque ligne non nulle d'une matrice échelonnée par lignes.

☞ Pour démontrer que toute matrice est équivalente par lignes à une matrice échelonnée, il suffit d'utiliser l'algorithme de Gauss-Jordan.

Remarque 4.9.

Soit A une matrice échelonnée par lignes de taille $n \times p$. Soit r le nombre de lignes non nulles de A. Alors $r \leqslant \min(n, p)$. Pour tout $k \in [\![1, r]\!]$, soit j_k l'indice de colonne du pivot de la ligne k. Alors $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant p$, les pivots de A sont les cœfficients $a_{1,j_1}, \ldots, a_{r,j_r}$ et A est de la forme suivante :

$$A = \begin{pmatrix} 0 & \cdots 0 & a_{1,j_1} & & \cdots & & & & & a_{1,p} \\ 0 & & 0 & \cdots & 0 & a_{2,j_2} & \cdots & & & a_{2,p} \\ \vdots & & & & & & & & & \vdots \\ 0 & & & \cdots & & & 0 & a_{r,j_r} & \cdots & a_{r,p} \\ 0 & & & \cdots & & & & 0 & \cdots & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & & & \cdots & & & & & & 0 \end{pmatrix}$$

Proposition. 4.32.

Le rang d'une matrice est égal au nombre de pivots de toute matrice échelonnée qui lui est équivalente par lignes.

☞ Autrement dit, partant d'une matrice dont on cherche à connaître le rang, on fait des opérations sur ses lignes jusqu'à obtenir une matrice échelonnée dont le rang est "lisible".

Remarque 4.10.

On définit de même la notion de système échelonné. Le rang d'un système est alors égal au nombre de pivot de tout système échelonné qui lui est équivalent, via des opérations élémentaires sur les lignes.

Proposition. 4.33.

Le rang d'une matrice est le rang de la famille formée de ses lignes. En d'autres termes, on a $\mathrm{rg}(M) = \mathrm{rg}(M^\top)$ pour toute matrice M.

4.3.3 Calcul de l'inverse d'une matrice inversible

Proposition. 4.34.

Une matrice de taille n est inversible si et seulement si elle est de rang n.

☞ Comme les opérations élémentaires préservent le rang, si la question est uniquement de savoir si une matrice est inversible ou non, on peut donc lui faire subir l'algorithme de Gauss-Jordan jusqu'à obtenir une matrice triangulaire qui aura les mêmes propriétés d'inversibilité que la matrice de départ et pour laquelle un critère est connu.

Exemple 4.12.

La matrice $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}$ est-elle inversible ? On peut raisonner comme ceci.

$$A \underset{L}{\sim} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

donc le rang de ces deux matrices est le même et notamment

$$A \text{ inversible} \quad \Longleftrightarrow \quad \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \text{ est inversible.}$$

Or, cette dernière matrice est triangulaire supérieure sans zéro sur la diagonale; elle est donc inversible et il en est de même pour A.

En plus de permettre de savoir si une matrice est inversible, l'algorithme de Gauss-Jordan permet de calculer l'inverse de la matrice.

☞ On considère une matrice A carrée de taille n ($n \in \mathbb{N}^*$) dont **on sait** qu'elle est inversible. Pour calculer A^{-1}, on applique l'algorithme du pivot de Gauss **total** sur la matrice A avec comme second membre I_n, en le prolongeant jusqu'à obtenir la matrice identité à la place de A. La matrice obtenue à la place de I_n est alors A^{-1}.

En effet, effectuer une opération sur les lignes d'une matrice revient à multiplier celle-ci à gauche par une matrice inversible. Si on appelle $P_1, \ldots, P_k$ les matrices inversibles correspondants aux opérations effectuées pour passer de la matrice M à $I_n n$, on a alors l'égalité matricielle suivante, où $Q = P_k P_{k-1} \cdots P_1 P_0$:

$$QM = I_n.$$

Cela signifie alors que Q est en fait l'inverse de M. Comme on a effectué les mêmes opérations sur la matrice I_n et qu'on a obtenu la matrice N, on a aussi l'égalité :

$$QI_n = N.$$

Autrement dit, $Q = N$, ce qui prouve que N est bien l'inverse de M.

Exemple 4.13. **Pivot simultané**

Reprenons la matrice

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

que nous savons être inversible.
Déterminons A^{-1} à l'aide d'un Pivot de Gauss simultané. On présente les choses comme ceci:

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

On commence donc par faire, comme précédemment, $L_3 \longleftarrow 2L_3 - L1$.

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 2 \end{array}\right)$$

On fait ensuite $L_1 \longleftarrow L_1 - L_3$ et $L_2 \longleftarrow L_2 + L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 2 & 0 & -2 \\ 0 & 2 & 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 0 & 2 \end{array}\right)$$

Enfin, on se ramène à des coefficients diagonaux égaux à 1

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1/2 & 1/2 & 1 \\ 0 & 0 & 1 & -1 & 0 & 2 \end{array}\right)$$

On peut alors **vérifier** que

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ -1/2 & 1/2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ -1/2 & 1/2 & 1 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et conclure que

$$A^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ -1/2 & 1/2 & 1 \\ -1 & 0 & 2 \end{pmatrix}.$$

Exercice 4.13.

Inverser les matrices P et Q ci-dessous à l'aide d'un pivot de Gauss simultané.

$$P = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

5
Des séries et des hommes

Ce chapitre propose des rappels et compléments sur les séries numériques ainsi qu'une application : la définition de l'exponentielle complexe.
Dans toute la suite, $\mathbb{K}$ désignera $\mathbb{R}$ ou $\mathbb{C}$.

5.1 Définitions

Définition 5.1. **Série numérique**

- ✗ On appelle **série numérique** ou **série** toute suite $(S_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K}$ telle que

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

  où $(u_n)_{n\in\mathbb{N}}$ est une suite à valeurs dans $\mathbb{K}$. On note alors cette série $\sum u_n$ ou $\sum\limits_{n\geq 0} u_n$.
- ✗ u_n est appelé le **terme général** de la série ;
- ✗ pour tout entier n, S_n est la **somme partielle** d'ordre n de la série ;
- ✗ une série $\sum u_n$ est dite à termes positifs si son terme général est à valeurs réelles et si, pour tout $n \geq 0$, $u_n \geq 0$.

☞ On rencontrera plus tard la notion de série entière qui désigne une série de *fonctions* et non plus de *nombres* d'où la précision avec le qualificatif *numérique* pour les séries étudiées dans ce chapitre.

Remarque 5.1. **Définition APCR**
Étant donnée une suite $(u_n)_{n\geq n_0}$ définie à partir d'un certain rang n_0 (APCR), on définit de même une série notée $\sum\limits_{n\geq n_0} u_n$, dont les sommes partielles sont définies par $S_n = \sum\limits_{k=n_0}^{n} u_k$ pour tout $n \geq n_0$.

Avec Python.

Si l'on dispose d'un programme `liste_u(n)` qui renvoie **la liste** des termes consécutifs (jusqu'au rang n) de la suite (u_n), on peut obtenir :

- ✗ la somme partielle d'indice n avec la commande `np.sum(liste_u(n))` ;
- ✗ la liste des termes consécutifs de la suite des sommes partielles grâce à la commande `np.cumsum(liste_u(n))`.

Autrement, on peut aussi utiliser une boucle `for` pour calculer S_n :

```
s= liste_u[0]
for k in range(1,n+1) :
    s=s+liste_u[k]
```

Définition 5.2. **Série convergente**

- ✗ Une série $\sum_{n\geq 0} u_n$ est dite **convergente** si la suite $(S_n)_{n\in\mathbb{N}}$ de ses sommes partielles est convergente.

  Auquel cas, la limite de la suite des sommes partielles est appelée **somme de la série** et est notée $\sum_{n=0}^{+\infty} u_n$.

  En d'autres termes on a :
$$\sum_{n=0}^{+\infty} u_n = \lim_{N\to+\infty} \sum_{n=0}^{N} u_n.$$
- ✗ Dans le cas où la suite des sommes partielles diverge, on dit que la série est **divergente**.
- ✗ Une série $\sum_{n\geq 0} u_n$ est dite **absolument convergente** si la série $\sum_{n\geq 0} |u_n|$ est convergente.
- ✗ Si une série $\sum_{n\geq 0} u_n$ converge, on appelle **reste à l'ordre** N de $\sum_{n\geq 0} u_n$, et on note R_N, le réel défini par
$$R_N = \sum_{n=N+1}^{+\infty} u_n = \sum_{n=0}^{+\infty} u_n - \sum_{n=0}^{N} u_n.$$

Remarque 5.2. **Attention, danger !**

- ✗ L'expression $\sum_{n=0}^{+\infty} u_n$ n'a un sens **que si** la série de terme général u_n **converge**, et désigne alors sa limite. On se gardera donc de l'écrire dans le cas où la suite diverge, et **on ne l'écrira qu'après avoir démontré la convergence.**
- ✗ L'expression $\sum_{n\geq 0} u_n$ désigne simplement la série de terme général u_n, et est donc toujours valable.
- ✗ La formule $\sum_{n\geq 0} u_n = \sum_{n=0}^{+\infty} u_n$ n'a donc **aucun sens**, puisque le premier membre est une série, alors que le second est un nombre (si tant est qu'il existe!).
- ✗ Lorsque l'on parlera de la *nature d'une série*, on se réfèrera à son caractère absolument convergeant, convergeant ou divergeant.

Remarque 5.3. **Nature et premier terme sommé**

Soit n_0 un entier **fixé**. Remarquant qu'on peut décomposer (par la relation de Chasles) une somme partielle
$$\sum_{k=0}^{n} u_k = \sum_{k=0}^{n_0-1} u_k + \sum_{k=n_0}^{n} u_k,$$
il est clair que la nature de la série ne dépend par de la somme des n_0 premiers termes, celle-ci étant constante. La valeur du premier terme de la série ne modifie donc pas le caractère convergent ou divergent mais en revanche, cela modifie, en cas de convergence, la valeur de la somme.

Définition 5.3. **Cas des séries de termes complexes**

Dans le cas où $\sum u_n$ est une série à termes complexes, celle-ci converge si et seulement si les deux séries $\sum \mathrm{Re}(u_n)$ et $\sum \mathrm{Im}(u_n)$ convergent, et lorsque c'est le cas on a
$$\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \mathrm{Re}(u_n) + i\sum_{n=0}^{+\infty} \mathrm{Im}(u_n).$$

5.2 Premiers résultats

5.2.1 Propriétés algébriques des séries convergentes

Le théorème suivant montre que l'ensemble des séries convergentes est un $\mathbb{K}$-espace vectoriel.

Théorème 5.1. **Structure d'espace vectoriel**

Soient $\sum_{n\geq 0} u_n$ et $\sum_{n\geq 0} v_n$ deux séries convergentes, et $(\lambda,\mu)\in\mathbb{K}^2$.

Alors, la série $\sum_{n\geq 0}(\lambda u_n+\mu v_n)$ converge, et

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

Remarque 5.4. **Attention, encore...**

Si $\sum_{n\geq 0} u_n$ ou $\sum_{n\geq 0} v_n$ diverge, l'égalité de la proposition précédente n'a aucun sens! Par exemple, bien que l'on ait

$$\forall n\geq 1,\ \frac{1}{n(n+1)}=\frac{1}{n}-\frac{1}{n+1},$$

écrire $\displaystyle\sum_{n=1}^{+\infty}\frac{1}{n(n+1)}=\sum_{n=1}^{+\infty}\frac{1}{n}-\sum_{n=1}^{+\infty}\frac{1}{n+1}$ n'aurait aucun sens puisque les séries de droite divergent (voir le critère de Riemann énoncé ci-après).

5.2.2 Critères élémentaires de convergence ou divergence

Proposition. 5.2.

Soit $\sum_{n\geq 0} u_n$ une série telle qu'il existe un entier $n_0\in\mathbb{N}$ vérifiant, pour tout $n\geq n_0, u_n=0$ (la suite (u_n) est donc nulle APCR).

Alors la série $\sum_{n\geq 0} u_n$ converge, et l'on a $\displaystyle\sum_{n=0}^{+\infty}u_n=\sum_{n=0}^{n_0-1}u_n$.

Le résultat suivant donne une condition **nécessaire (mais pas suffisante!)** de convergence.

Proposition. 5.3. **Une condition nécessaire**

Soit $\sum_{n\geq 0} u_n$ une série convergente. Alors $\lim_{n\to+\infty}u_n=0$.

Définition 5.4. **Divergence grossière**

Si $(u_n)_{n\in\mathbb{N}}$ ne tend pas vers 0, alors la série $\sum_{n\geq 0} u_n$ est divergente et on dit qu'elle diverge grossièrement.

Exercice 5.1.

Soit H_n la somme partielle d'indice n de la *série harmonique* $\sum\frac{1}{k}$.

1. Que vérifie le terme général de cette série?

2. Montrer que, pour tout $n\geq 1$,

$$H_{2n}-H_n\geq\frac{1}{2}.$$

3. En déduire la divergence de la série harmonique.

Remarque 5.5. **Une erreur à ne pas commettre**

L'exercice précédent propose alors un contre-exemple de série qui diverge bien que son terme général tende vers 0.

On prendra donc bien garde de ne **jamais** écrire que la convergence vers 0 du terme général implique celle de la série.

C'est nécessaire, mais il faut en fait savoir si on tend vers 0 *assez vite* de sorte que la quantité sommée est *assez petite*. Comme on le verra ci-après, la recherche d'un équivalent ou d'une relation de négligeabilité (pour une série à termes positifs) permet de déterminer la nature de celle-ci : c'est une information plus précise que la simple limite.

Les outils d'étude des suites s'appliquent aussi aux séries (ce sont des suites de sommes partielles....). Notamment, le théorème de convergence monotone implique le résultat suivant.

Proposition. 5.4.

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

Exercice 5.2.

On considère la série $\sum_{n\geq 1} \frac{1}{n^2}$, dont on note S_n la somme partielle d'indice n.

1. Montrer que la suite (S_n) est croissante.
2. Montrer que, pour tout $n \geq 2$, on a $\frac{1}{n^2} \leq \frac{1}{n-1} - \frac{1}{n}$.
3. En déduire que la série est convergente. On note S sa somme.
4. Déduire également de la Question **2.**, que, pour tout $n \in \mathbb{N}^*$ et tout $p \geq n$, on a

$$\sum_{k=n+1}^{p} \frac{1}{k^2} \leq \frac{1}{n} - \frac{1}{p}.$$

En déduire une majoration du reste.
5. Écrire un programme en `Python` permettant d'obtenir une valeur approchée à 10^{-3} de S.

5.2.3 Séries téléscopiques

Proposition. 5.5. **Convergence des séries téléscopiques**

Soit $(a_n)_{n\geq 0} \in \mathbb{K}^{\mathbb{N}}$. La série $\sum_{n\geq 0} (a_{n+1} - a_n)$ converge si et seulement si la suite (a_n) converge, et dans ce cas :

$$\sum_{n=0}^{+\infty} (a_{n+1} - a_n) = \lim_{n\to+\infty} a_n - a_0.$$

Exercice 5.3.

Montrer que la série $\sum_{n\geq 1} \ln\left(1 + \frac{1}{n}\right)$ est divergente mais pas grossièrement divergente.

Exercice 5.4.

Soient f la fonction définie sur $[0;1]$ par $f : x \mapsto x - x^2$ et (u_n), de premier terme $u_0 \in]0;1[$, vérifiant la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = f(u_n) = u_n - u_n^2.$$

1. Dresser le tableau de variations de f.
2. Montrer que $\forall n \in \mathbb{N},\ u_n \in]0;1[$.
3. Étudier la monotonie puis la convergence de (u_n). Déterminer sa limite.
4. Montrer que la série $\sum u_n^2$ converge et préciser sa somme.

5.2.4 Série géométrique

Proposition. 5.6. **Série géométrique**

Soit $z \in \mathbb{C}$. La série $\sum_{n \geq 0} z^n$ converge si et seulement si $|z| < 1$.
Dans ce cas, il y a convergence absolue, et l'on a :

$$\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}.$$

Si $|z| \geq 1$, cette série diverge grossièrement.

Remarque 5.6. **Développement limité et sommes partielles**

Le résultat précédent permet de retrouver le développement limité de $x \mapsto \dfrac{1}{1-x}$.
Attention, on ne peut pas généraliser ce genre de résultat (ne serait-ce que parce que les développements limités sont des formules locales).
Certaines fonctions ont la propriété d'être *développables* en série (entière), ce qui sera proprement introduit et détaillé dans le **Chapitre 9**.

5.2.5 Comparaison à l'intégrale et séries de Riemann

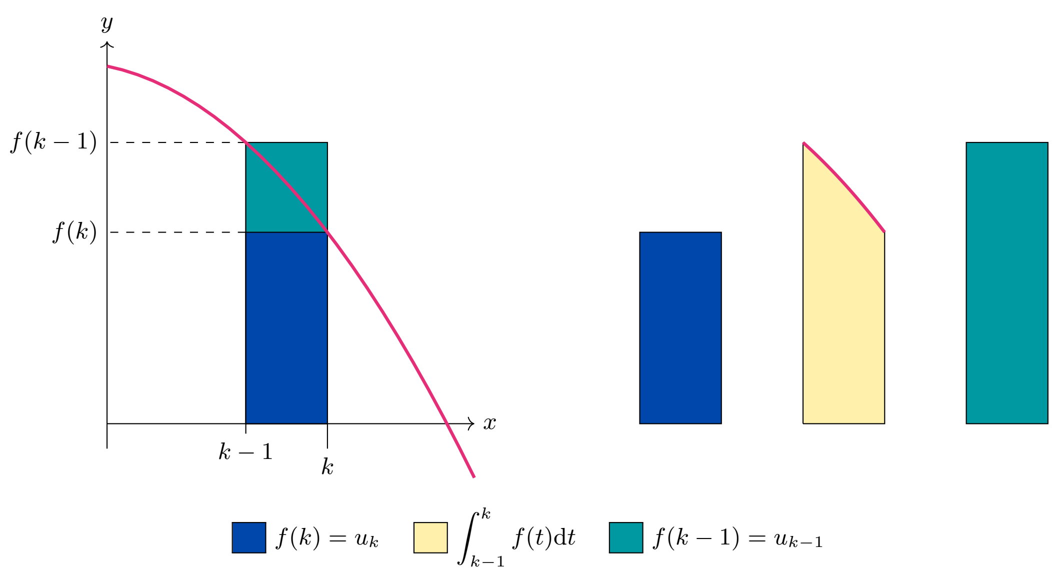
Pour déterminer la nature d'une série à termes positifs, on peut comparer ses sommes partielles à une intégrale, via le lemme suivant. La démonstration de ce résultat est un raisonnement classique à savoir refaire.

Proposition. 5.7. **Comparaison série / intégrale**

Soit $f : [0, +\infty[\to \mathbb{R}$ une fonction continue et décroissante. Alors :

i. Pour tout $n \in \mathbb{N}$, $f(n+1) \leq \displaystyle\int_n^{n+1} f(t)\mathrm{d}t \leq f(n)$;

ii. Pour tout$n \in \mathbb{N}$, $\displaystyle\sum_{k=1}^{n+1} f(k) \leq \int_0^{n+1} f(t)\mathrm{d}t \leq \sum_{k=0}^{n} f(k)$.



Définition 5.5. **Série de Riemann**

On appelle **série de Riemann** la série de terme général $\dfrac{1}{n^\alpha}$, où $\alpha \in \mathbb{R}$.

Théorème 5.8. **Critère de Riemann**

La série de Riemann de terme général $\dfrac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

Exercice 5.5. **Une série de Bertrand**

Montrer que la série de terme général $\dfrac{1}{n \ln n}$ diverge et donner un équivalent de $\displaystyle\sum_{n=2}^{N} \frac{1}{n \ln n}$.

Remarque 5.7.

- ✗ La comparaison à l'intégrale est fréquemment utilisée pour déterminer un équivalent de la somme partielle d'une série divergente (ou du reste d'une série convergente).
- ✗ Cette méthode peut également s'appliquer à des fonctions croissantes.
- ✗ La comparaison série/intégrale ne permet pas (le plus souvent) de calculer la somme d'une série convergente.

5.2.6 Séries alternées

Théorème 5.9. **Critère spécial des séries alternées**

Soit $(u_n)_{n\geq 0}$ une suite décroissante à termes positifs, qui converge vers 0. On note, pour $N \in \mathbb{N}$, la somme partielle $S_N = \displaystyle\sum_{n=0}^{N} (-1)^n u_n$.

On a les résultats suivants.

i. La série $\sum (-1)^n u_n$ converge.

ii. Pour tout entier N on a $S_{2N+1} \leq \displaystyle\sum_{n=0}^{+\infty} (-1)^n u_n \leq S_{2N}$.

iii. Pour tout $n \in \mathbb{N}$, la somme $R_n = \displaystyle\sum_{k=n}^{+\infty} (-1)^k u_k$ de la série a le même signe que son premier terme et est majoré par celui-ci en valeur absolue :

$$\forall n \in \mathbb{N}, \qquad 0 \leq |R_n| = (-1)^n R_n \leq u_n.$$

Exercice 5.6.

1. Déterminer la nature de la série de terme général $\dfrac{(-1)^n}{n^\alpha}$, en fonction de $\alpha \in \mathbb{R}$.

2. Soit $\alpha \in \mathbb{R}$ tel que $\displaystyle\sum_{n\geq 1} \frac{(-1)^n}{n^\alpha}$ converge.

Déterminer un entier n_0 tel que S_{n_0} soit une valeur approchée de la somme de la série à 10^{-3} près.

5.3 Théorèmes de comparaison pour les séries à termes positifs

S'il est en général difficile de calculer la somme d'une série, on dispose de critères de convergence qui permettent de déterminer la nature d'une série, en la comparant à une série de nature connue.

Théorème 5.10. **Critère de comparaison**

Soient $\displaystyle\sum_{n\geq 0} u_n$ et $\displaystyle\sum_{n\geq 0} v_n$ deux séries à termes positifs telles que $0 \leq u_n \leq v_n$ à partir d'un certain rang. Alors,

i. Si $\displaystyle\sum_{n\geq 0} v_n$ converge alors $\displaystyle\sum_{n\geq 0} u_n$ converge, et on a dans ce cas $\displaystyle\sum_{n=0}^{+\infty} u_n \leq \sum_{n=0}^{+\infty} v_n$.

ii. Si $\displaystyle\sum_{n\geq 0} u_n$ diverge alors $\displaystyle\sum_{n\geq 0} v_n$ diverge.

Remarque 5.8. **Comparaison des séries à termes négatifs**

Si on étudie une série dont le terme général u_n est négatif, on peut encore utiliser ce résultat quitte à considérer $-u_n$. Le seul cas où on ne peut pas utiliser ce théorème est le cas où $(u_n)_{n\in\mathbb{N}}$ n'est pas de signe constant.

Exercice 5.7.

1. Montrer que la série $\displaystyle\sum_{n\geq 1}\frac{1}{n(n+1)}$ converge, à l'aide du théorème de comparaison.
2. Calculer la somme de cette série.

Exercice 5.8.

Montrer que la série de terme général $\dfrac{\ln n}{n^2}$ converge.

Théorème 5.11. **Critère d'équivalence**

Soient $\displaystyle\sum_{n\geq 0} u_n$ et $\displaystyle\sum_{n\geq 0} v_n$ deux séries de **signe constant**.

Si, lorsque $n\to+\infty$, $u_n\sim v_n$ alors $\displaystyle\sum_{n\geq 0} u_n$ et $\displaystyle\sum_{n\geq 0} v_n$ sont de même nature.

Remarque 5.9.

Attention, ce théorème ne s'applique pas aux séries à termes réels de signe variable, ni à celles de termes complexes non réels.
On s'en convaincra en regardant l'**Exercice 5.10**.

Exercice 5.9.

Déterminer la nature de la série de terme général $\ln\left(1+\dfrac{1}{2^n}\right)$.

Exercice 5.10. **Une erreur classique**

1. Montrer que la série $\displaystyle\sum\frac{(-1)^n}{\sqrt{n}}$ converge.
2. Montrer que :
$$\frac{(-1)^n}{\sqrt{n}+(-1)^n}=\frac{(-1)^n}{\sqrt{n}}-\frac{1}{n}+op\left(\frac{1}{n}\right).$$
3. En déduire que la série $\displaystyle\sum\frac{(-1)^n}{\sqrt{n}+(-1)^n}$ diverge. Commenter.

5.4 Étude de la convergence absolue

5.4.1 Théorème fondamental et théorème de comparaison

Pour déterminer la nature d'une série, on se contentera le plus souvent de déterminer si elle est absolument convergente, et on sera ainsi ramené à l'étude d'une série à termes positifs. Le théorème suivant permettra d'en déduire la convergence de la série.

Théorème 5.12.

Toute série absolument convergente est convergente.

Remarque 5.10.

La réciproque du résultat ci-dessus est inexacte. Par exemple la série de terme général $\dfrac{(-1)^n}{n}$ converge (elle est alternée), alors que celle de terme général $\dfrac{1}{n}$ diverge.

Exemple 5.1.

La série de terme général $u_n = \dfrac{\sin(n)}{n^2}$ est absolument convergente, et donc convergente.

Le théorème suivant montre que l'ensemble des séries absolument convergentes est un sous-espace vectoriel de l'ensemble des séries convergentes.

Proposition. 5.13.

Soient $\sum_{n\geq 0} u_n$ et $\sum_{n\geq 0} v_n$ deux séries absolument convergentes et $(\lambda,\mu)\in\mathbb{K}^2$. Alors la série $\sum_{n\geq 0}(\lambda u_n+\mu v_n)$ est absolument convergente.

Les différents théorèmes de comparaison s'adaptent à la notion de série absolument convergente.

Théorème 5.14. **Critères de comparaison**

Soient $\sum_{n\geq 0} u_n$ et $\sum_{n\geq 0} v_n$ deux séries.

i. Si, pour tout $n\in\mathbb{N}$ (ou au moins APCR), $|u_n|\leq|v_n|$ et si $\sum_{n\geq 0} v_n$ converge absolument, alors $\sum_{n\geq 0} u_n$ converge absolument.

ii. Si $u_n \underset{n\to+\infty}{=} O(v_n)$ (ou $u_n \underset{n\to+\infty}{=} o(v_n)$) et si $\sum_{n\geq 0} v_n$ converge absolument, alors $\sum_{n\geq 0} u_n$ converge absolument.

iii. Si $u_n \underset{n\to+\infty}{\sim} v_n$ alors la convergence absolue de $\sum_{n\geq 0} v_n$ équivaut à celle de $\sum_{n\geq 0} u_n$.

Exercice 5.11.

Étudier la nature de la série de terme général $(-1)^n ne^{-n}$.

Proposition. 5.15. **Inégalité triangulaire généralisée**

Soient $\sum_{n\geq 0} u_n$, $\sum_{n\geq 0} v_n$ deux séries et $M\geq 0$.

i. Si la série $\sum_{n\geq 0} u_n$ est absolument convergente, alors : $\left|\sum_{n=0}^{+\infty} u_n\right| \leq \sum_{n=0}^{+\infty}|u_n|$.

ii. Si, pour tout $n\in\mathbb{N}$, $|v_n|\leq M$, et si $\sum_{n\geq 0} u_n$ est absolument convergente, alors $\sum_{n\geq 0} v_n u_n$ est absolument convergente, et

$$\sum_{n=0}^{+\infty}|v_n u_n| \leq M\sum_{n=0}^{+\infty}|u_n|.$$

Autrement dit, le produit des termes généraux d'une suite bornée et du terme général d'une série absolument convergente est encore le terme général d'une suite absolument convergente.

5.4.2 Règle de d'Alembert

Le théorème suivant est fréquemment utile. On en réutilisera une version dans le cadre des séries entières. Il est indispensable de le connaître parfaitement.

Théorème 5.16. **Règle de d'Alembert**

Soient $\ell \in \mathbb{R} \cup \{+\infty\}$ et $\sum_{n \geq 0} u_n$ une série à termes non nuls à partir d'un certain rang telle que

$$\lim_{n \to +\infty} \left| \frac{u_{n+1}}{u_n} \right| = \ell.$$

i. Si $\ell < 1$ alors $\sum_{n \geq 0} u_n$ converge absolument.

ii. Si $\ell > 1$ alors $\sum_{n \geq 0} u_n$ diverge grossièrement.

Remarque 5.11.

Si $\ell = 1$, on ne peut pas conclure. En effet, si $u_n = \frac{1}{n^\alpha}$ on a $\lim_{n \to +\infty} \frac{u_{n+1}}{u_n} = 1$, et pourtant $\sum u_n$ est tantôt divergente, tantôt convergente suivant la valeur de α.

Il existe un résultat, nommé critère de **Raabe-Duhamel** (voir **Exercice 5.??**), qui fait intervenir le deuxième terme du développement asymptotique de u_{n+1}/u_n, mais il n'est pas au programme en **PT**.
On pourra aussi regarder l'**Exercice 5.??**, qui propose un critère dans un des cas limite de la règle de d'Alembert.

Exercice 5.12.

Étudier la nature de la série de terme général $n^2 z^n$ où $z \in \mathbb{C}$.

5.4.3 Produit de Cauchy

Définition 5.6. **Produit de Cauchy**

Soient $\sum_{n \geq 0} u_n$, $\sum_{n \geq 0} v_n$ deux séries. On appelle **produit de Cauchy** de $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ la série $\sum_{n \geq 0} w_n$ où

$$w_n = \sum_{k+i=n} u_k v_i = \sum_{k=0}^{n} u_k v_{n-k}.$$

Théorème 5.17. **Produit de Cauchy de deux séries abs. convergentes**

Soient $\sum_{n \geq 0} u_n$, $\sum_{n \geq 0} v_n$ deux séries absolument convergentes.

Alors leur produit de Cauchy $\sum_{n \geq 0} w_n$ converge absolument, et on a

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

5.5 Application : l'exponentielle complexe

Soit $z = x + iy \in \mathbb{C}$ avec $x, y \in \mathbb{R}$. En PTSI, on a défini l'exponentielle complexe de z en posant $e^z = e^x e^{iy} = e^x(\cos y + i \sin y)$. On peut maintenant donner une autre expression de l'exponentielle complexe.

Théorème 5.18. **Série exponentielle**

i. Soit $z \in \mathbb{C}$. La série $\sum_{n \geq 0} \frac{z^n}{n!}$ converge absolument.

ii. Pour tout $z \in \mathbb{C}$ on a $\sum_{n=0}^{+\infty} \frac{z^n}{n!} = e^z$.

Preuve. Soit $x \in \mathbb{R}$. La fonction exponentielle est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ donc l'inégalité de Taylor-Lagrange à l'ordre n en 0 donne :

$$\left|\mathrm{e}^x - \sum_{k=0}^{n} \frac{x^k}{k!}\right| = \left|\mathrm{e}^x - \sum_{k=0}^{n} \frac{\mathrm{e}^0}{k!}(x-0)^k\right| \leqslant \frac{|x-0|^{n+1}}{(n+1)!} \max_{t\in[0,x]} \mathrm{e}^t \leqslant \frac{|x|^{n+1}}{(n+1)!}\mathrm{e}^{|x|} \underset{n\to+\infty}{\longrightarrow} 0$$

puisque $|x|^n = o(n!)$ pour tout réel x.

Soit maintenant $z \in \mathbb{C}$. La série $\sum_{n\geq 0} \frac{z^n}{n!}$ est alors absolument convergente (donc convergente). On note $S(z)$ sa somme.

Si $z \in \mathbb{R}$, on a alors $S(z) = e^z$ par unicité de la limite. Si $z = iy \in i\mathbb{R}$, alors

$$\sum_{k=0}^{n} \frac{z^k}{k!} = \sum_{k=0}^{n} \frac{(iy)^k}{k!} = \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j y^{2j}}{(2j)!} + i \sum_{j=0}^{\lfloor (n-1)/2 \rfloor} \frac{(-1)^j y^{2j+1}}{(2j+1)!}$$

Toujours par Taylor-Lagrange (à l'ordre n) appliqué aux fonctions cos et sin de classe $\mathcal{C}^\infty$, on a (en majorant $|\sin|$ et $|\cos|$ par 1)

$$\left|\cos(y) - \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j y^{2j}}{(2j)!}\right| \leq \frac{|y-0|^{n+1}}{(n+1)!} \underset{n\to+\infty}{\longrightarrow} 0$$

et

$$\left|\sin(y) - \sum_{j=0}^{\lfloor (n-1)/2 \rfloor} \frac{(-1)^j y^{2j+1}}{(2j+1)!}\right| \leq \frac{|y-0|^{n+1}}{(n+1)!} \underset{n\to+\infty}{\longrightarrow} 0.$$

Ainsi, en identifiant parties réelles et imaginaires, on a bien

$$\sum_{k=0}^{+\infty} \frac{z^k}{k!} = \sum_{k=0}^{+\infty} \frac{(iy)^k}{k!} = \cos(y) + i\sin(y) = e^{iy}.$$

Enfin, si $z = x + iy$, alors, les séries $\sum \frac{x^k}{k!}$ et $\sum \frac{(iy)^k}{k!}$ étant toutes deux absolument convergentes, leur produit de Cauchy l'est aussi. Or, le terme général de ce produit de Cauchy vaut

$$\sum_{k=0}^{n} \frac{x^k}{k!} \frac{(iy)^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k (iy)^{n-k} = \frac{(x+iy)^n}{n!} = \frac{z^n}{n!},$$

par la formule du binôme. Ainsi,

$$S(z) = e^x e^{iy} = e^{x+iy} = e^z.$$

$\square$

6 Déterminant

Dans tout ce chapitre, la lettre $\mathbb{K}$ désignera l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$.

6.1 Théorème fondamental, et expression explicite en dimension 2 et 3

Soit $n \in \mathbb{N}^*$. Étant données des matrices $C_1, \ldots, C_n$ dans $\mathcal{M}_{n,1}(\mathbb{K})$, on notera $[C_1, \ldots, C_n]$ la matrice de taille n dont les colonnes sont $C_1, \ldots, C_n$.

Théorème & Définition 6.1. **Application déterminant**

Soit $n \in \mathbb{N}^*$. Il existe une unique application $f : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ vérifiant les propriétés suivantes:

i. f est **n-linéaire**, c'est-à-dire linéaire par rapport à chacune des n colonnes de son argument: si $C_1, \ldots, C_n$ sont dans $\mathcal{M}_{n,1}(\mathbb{K})$ et si $j \in [\![1, n]\!]$, alors $X \mapsto f([C_1, \ldots, C_{j-1}, X, C_{j+1}, \ldots, C_n])$ est dans $\mathcal{L}(\mathcal{M}_{n,1}(\mathbb{K}), \mathbb{K})$;

ii. f est **antisymétrique**, c'est-à-dire que si $A \in \mathcal{M}_n(\mathbb{K})$ et si B est la matrice obtenue à partir de A en échangeant deux de ses colonnes, alors $f(A) = -f(B)$;

iii. $f(I_n) = 1$.

Cette unique application est appelée **déterminant** et est notée $\det$.

Si $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$, son déterminant sera noté $\det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}$.

Preuve. Résultat admis. □

Il résulte de la définition le résultat suivant.

Proposition. 6.1.

Le déterminant d'une matrice ayant deux colonnes égales est nul.

Proposition. 6.2.

Soient $x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3$ des scalaires. On a alors les résultats suivants:

i. $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - y_1x_2.$

ii. $\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_1y_3z_2 - x_3y_2z_1 - x_2y_1z_3$

Remarque 6.1.

- ✗ On dispose de même d'une formule explicite dans le cas général, mais celle-ci est hors-programme.
- ✗ Pour retrouver la formule *ii.* ci-dessus, on utilise la *règle de Sarrus*, qui consiste a recopier sous le tableau les deux premières lignes de celui-ci, puis à sommer le produit des termes rencontrés le long de chaque diagonale, affublé d'un signe moins si la diagonale est croissante.

Exercice 6.1.

Calculer $\begin{vmatrix} 1 & 1 & -2 \\ 2 & -1 & 3 \\ 4 & 1 & -1 \end{vmatrix}$.

6.2 Propriétés élémentaires du déterminant d'une matrice carrée

Proposition. 6.3. **Propriétés du déterminant**

Soient $(A, B) \in \mathcal{M}_n(\mathbb{K})^2$ et $\lambda \in \mathbb{K}$. On a les résultats suivants.

i. $\det(\lambda \cdot A) = \lambda^n \det(A)$;
ii. Le déterminant d'une matrice ayant deux colonnes égales est nul;
iii. Le déterminant d'une matrice dont une colonne est combinaison linéaire des autres est nul;
iv. Le déterminant d'une matrice ayant une colonne nulle est nul;
v. Soient $C_1, \ldots, C_n$ les colonnes de A, et $i \neq j$ dans $[\![1, n]\!]$. Alors si B est la matrice obtenue à partir de A en effectuant l'opération $C_i \leftarrow C_i + \lambda C_j$, on a $\det A = \det B$;
vi. $\det(AB) = \det(A) \cdot \det(B)$;
vii. $\det(A^\top) = \det(A)$.
viii. $A \in \mathrm{GL}_n(\mathbb{K}) \iff \det(A) \neq 0$;
ix. $A \in \mathrm{GL}_n(\mathbb{K}) \Longrightarrow \det\left(A^{-1}\right) = \dfrac{1}{\det(A)}$;

Preuve. Les points *vi.* et *vii.* sont admis. □

☞ En particulier, une qui matrice possède deux colonnes proportionnelles a un déterminant nul.

Remarque 6.2.

✗ Le point *vii.* montre que le déterminant permet de caractériser les matrices inversibles.
✗ Le point *ix.* assure que les lignes et les colonnes d'un déterminant jouent le même rôle. Les propriétés *i.*, *ii.*, *,iii.*, *,iv.* sont donc également vraies si on considère les lignes au lieu des colonnes.
✗ L'opération $C_i \leftarrow \alpha C_i + \lambda C_j$ multiplie la valeur du déterminant par α.
✗ Le point *vi.* permet de voir que le déterminant est *invariant par changement de base*. Si A et B sont semblables, avec $A = PBP^{-1}$, alors

$$\det(A) = \det\left(PBP^{-1}\right) = \det(PB) \times \det\left(P^{-1}\right) = \det\left(P^{-1}\right) \times \det(PB) = \det\left(P^{-1}PB\right) = \det(B).$$

6.3 Développement par rapport à une ligne ou une colonne

Proposition. 6.4. **Développement par rapport à une ligne ou une colonne**

Soit $M \in \mathcal{M}_n(\mathbb{K})$. Pour tout $(i, j) \in [\![1, n]\!]^2$ on note $m_{i,j}$ le terme de M d'indice (i, j), et $M_{i,j}$ la matrice obtenue à partir de M en supprimant la i-ème ligne et la j-ème colonne. On a alors les formules suivantes:

✗ **Développement par rapport à la j-ème colonne**. Pour tout $j \in [\![1, n]\!]$:

$$\det M = \sum_{i=1}^{n} (-1)^{i+j} m_{i,j} \times \det M_{i,j}.$$

✗ **Développement par rapport à la i-ème ligne.** Pour tout $i \in [\![1, n]\!]$:

$$\det M = \sum_{j=1}^{n} (-1)^{i+j} m_{i,j} \times \det M_{i,j}.$$

Preuve. Résultat admis. □

☞ Ces formules permettent de ramener le calcul d'un déterminant de taille n à un déterminant de taille $n - 1$, et en raisonnant par récurrence, on peut ainsi calculer sa valeur.

☞ En dimension 3 les formules de développement par rapport à une colonne sont ainsi:

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = x_1 \begin{vmatrix} y_2 & y_3 \\ z_2 & z_3 \end{vmatrix} - y_1 \begin{vmatrix} x_2 & x_3 \\ z_2 & z_3 \end{vmatrix} + z_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}$$

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = - x_2 \begin{vmatrix} y_1 & y_3 \\ z_1 & z_3 \end{vmatrix} + y_2 \begin{vmatrix} x_1 & x_3 \\ z_1 & z_3 \end{vmatrix} - z_2 \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}$$

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = x_3 \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} - y_3 \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} + z_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$$

☞ En pratique on retrouve ces formules ci-dessus en considérant que chaque terme de la somme est le produit d'un coefficient de la colonne choisie, multiplié par le déterminant de *la matrice obtenue en enlevant la ligne et la colonne de ce coefficient*, et multiplié par un signe, qui dépend de la position du coefficient.
On peut facilement mémoriser ce signe à l'aide du tableau :

$$\begin{bmatrix} + & - & + & \cdots & \cdots \\ - & + & - & \cdots & \cdots \\ \vdots & & & & \vdots \\ \vdots & \cdots & & \cdots & + \end{bmatrix}$$

Remarque 6.3.
Pour simplifier le développement d'un déterminant par rapport à l'une de ses colonnes, on peut d'abord transformer la colonne considérée de sorte qu'elle ne contienne qu'un seul terme non nul, en effectuant des opérations élémentaires du type $C_i \leftarrow C_i + \lambda C_j$, puisque celles-ci ne changent pas la valeur du déterminant.

Corollaire 6.5. **Déterminant d'une matrice triangulaire**
Le déterminant d'une matrice triangulaire supérieure ou inférieure est égal au produit de ses termes diagonaux.

Exercice 6.2. **Déterminant de Vandermonde**
Soient $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n$ des scalaires. Montrer que

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & \cdots & x_2^{n-1} \\ \vdots & & & & & \vdots \\ 1 & x_n & x_n^2 & \cdots & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Ce déterminant est noté $V(x_1, \ldots, x_n)$ et est appelé *déterminant de Vandermonde.*

6.4 Exercices type d'application

S'il est aisé de calculer un déterminant (on peut toujours développer par rapport à une ligne ou une colonne, autant de fois que nécessaire), il est plus difficile d'en obtenir une forme factorisée. On présente ici quelques exercices type illustrant les méthodes courantes.

6.4.1 Calcul en dimension 3

Pour factoriser un déterminant de taille 3, il suffit d'appliquer l'algorithme du pivot de Gauss pour se ramener à une matrice triangulaire (supérieure).

Exercice 6.3.
Soit $B = \begin{pmatrix} 5 & 1 & -1 \\ 2 & 4 & -2 \\ 1 & -1 & 3 \end{pmatrix}$. Déterminer pour quelles valeurs de $\lambda \in \mathbb{K}$ la matrice $B - \lambda I_3$ est inversible.

6.4.2 Cas général

Dans le cas général, on peut développer par rapport à une ligne ou une colonne afin de faire apparaître une relation de récurrence.

Exercice 6.4.

Soient $n \in \mathbb{N}^*$ et $a_0, \ldots, a_n$ des scalaires. Calculer $\begin{vmatrix} a_n & a_{n-1} & \cdots & \cdots & a_1 & a_0 \\ -1 & x & 0 & \cdots & \cdots & 0 \\ 0 & -1 & x & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & -1 & x & 0 \\ 0 & \cdots & & 0 & -1 & x \end{vmatrix}$.

6.5 Déterminant d'une famille ou d'un endomorphisme dans une base donnée

Dans cette section, on appelle E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^*$, muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$.

Définition 6.1. **Déterminant d'une famille de vecteurs dans une base**

Soit $\mathcal{F}_n = (u_1, \ldots, u_n)$ une famille de n vecteurs de E. On appelle **déterminant de $\mathcal{F}_n$ dans la base** $\mathcal{B}$ et on note $\det_{\mathcal{B}}(\mathcal{F}_n)$ le déterminant de la matrice de la famille $\mathcal{F}_n$ dans la base $\mathcal{B}$.

Remarque 6.4.

- ✗ De même que le déterminant d'une matrice rectangulaire n'est pas défini, dans un espace de dimension n on ne définit que le déterminant de n vecteurs.
  La famille dont on calcule le déterminant a donc nécessairement un cardinal égal à la dimension de l'espace dont les vecteurs sont éléments.

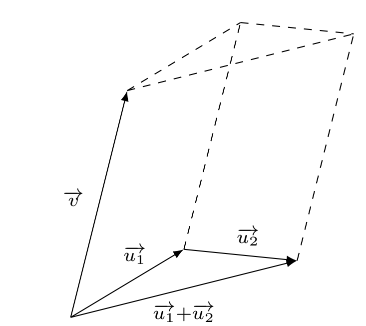
- ✗ Dans le cas où E est égal au plan ou a l'espace usuel (et donc $n = 2$ ou 3), si $\mathcal{B}$ est une base orthonormée directe, cette définition coïncide avec celle du produit mixte. Dans le premier cas on obtient donc l'aire du parallélogramme construit sur les deux vecteurs, et dans le second le volume du parallélépipède construit sur les trois vecteurs.

Proposition. 6.6. **Base et déterminant**

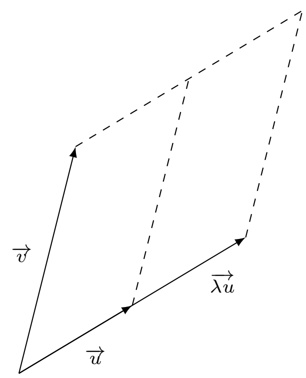
i. Une famille $(u_1, \ldots, u_n)$ de E est liée si et seulement si $\det_{\mathcal{B}}(u_1, \ldots, u_n) = 0$.

ii. Une famille $(u_1, \ldots, u_n)$ de E est une base si et seulement si $\det_{\mathcal{B}}(u_1, \ldots, u_n) \neq 0$.

☞ Géométriquement le déterminant de n vecteurs $u_1, \ldots, u_n$ est égal au volume orienté du parallélépipède construit sur ces n vecteurs, la base choisie correspondant à l'unité de volume. Le déterminant permet donc de définir la notion de volume dans un espace vectoriel de dimension finie quelconque. En effet, le dessin ci-dessous justifie que, dans le cas $n = 2$, l'aire $\mathcal{A}(\overrightarrow{u}, \overrightarrow{v})$ du parallélogramme construit sur $\overrightarrow{u}, \overrightarrow{v}$ est bien linéaire à gauche. On justifierait de même que l'aire est linéaire à droite, et comme elle est orientée elle est clairement alternée.



$\mathcal{A}(\overrightarrow{u_1}+\overrightarrow{u_2}, \overrightarrow{v}) = \mathcal{A}(\overrightarrow{u_1}, \mathrm{Vect} v) + \mathcal{A}(\overrightarrow{u_2}, \overrightarrow{v})$



$\mathcal{A}(\lambda \overrightarrow{u}, \overrightarrow{v}) = \lambda \mathcal{A}(\overrightarrow{u}, \overrightarrow{v})$

Remarque 6.5.
Si E est égal au plan ou a l'espace usuel, on en déduit que le produit mixte d'une famille de vecteurs est nul si et seulement si cette famille est liée, et ce même si la base n'est pas orthonormée directe.

Théorème & Définition 6.2. **Déterminant d'un endomorphisme**
Soit $f \in \mathcal{L}(E)$. Toutes les matrices associées à f ont le même déterminant.
On appelle **déterminant de** f, et on note $\det(f)$, le déterminant de la matrice de f dans n'importe quelle base.

Le résultat suivant montre notamment que le déterminant permet de caractériser les automorphismes.

Proposition. 6.7.
Soient $(f, g) \in \mathcal{L}(E)^2$ et $\lambda \in \mathbb{K}$. On a les résultats suivants.

i. $\det(\lambda \cdot f) = \lambda^n \det(f)$;
ii. $\det(f \circ g) = \det(f) \cdot \det(g)$;
iii. $f \in \mathrm{GL}(E) \iff \det(f) \neq 0$;
iv. $f \in \mathrm{GL}(E) \Longrightarrow \det\left(f^{-1}\right) = \dfrac{1}{\det(f)}$.

Exercice 6.5.
Quelles sont les valeurs possibles du déterminant d'un projecteur? D'une symétrie?

7 Intégration généralisée

Ce chapitre propose, dans une première partie des rappels sur le cours d'intégration de première année, puis étend ensuite la notion d'intégrale à des intervalles non bornés et/ou non fermés.

7.1 Rappels : intégrale d'une fonction continue sur un segment

7.1.1 Définition et propriétés

On peut montrer que tout fonction continue $f : [a, b] \to \mathbb{R}$ peut s'écrire comme limite d'une fonction en escalier. Étant immédiat le calcul de l'aire sous la courbe d'une fonction en escalier, on peut alors définir l'intégrale de f sur $[a, b]$ comme une limite.
En utilisant une subdivision régulière (à pas constant) de $[a, b]$, on obtient le résultat suivant.

Définition 7.1. **Somme de Riemann à l'ordre n**

Soient f une fonction continue sur $[a, b]$ et $n \in \mathbb{N}^*$.

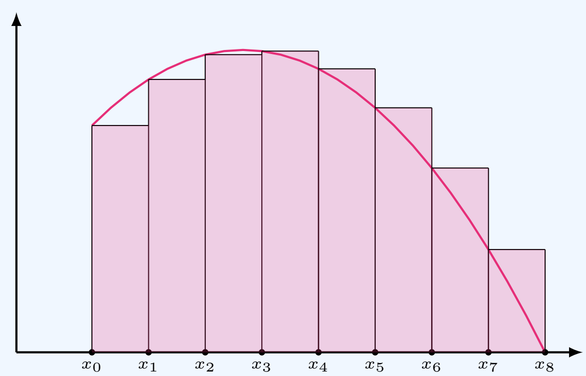
On définit la subdivision $(x_0, ..., x_n)$ régulière de $[a, b]$ (de pas $(b - a)/n$) par :

$$\forall k \in [\![0, n]\!],\ x_k = a + k\frac{b-a}{n}.$$

Puis, la **somme de Riemann à l'ordre** n de f est alors définie par

$$R_n(f) = \frac{b-a}{n}\sum_{k=0}^{n-1} f(x_k).$$

Celle-ci correspond à l'aire algébrique des n rectangles de largeur $(b - a)/n$ et de hauteurs respectives $f(x_0)$, ... $f(x_{n-1})$.



Avec $n = 8$

Théorème & Définition 7.1. **Sommes de Riemann à pas constant**

Soit f une fonction continue sur $[a, b]$. En reprenant les notations de la **Définition 7.1.** ci-dessus, la suite des sommes de Riemann $(R_n(f))_{n\geq 1}$ est convergente et on appelle intégrale de f sur $[a, b]$ sa limite. On note :

$$\lim_{n\to+\infty} \frac{b-a}{n}\sum_{k=0}^{n-1} f(x_k) =: \int_a^b f(t)\mathrm{d}t.$$

☞ Le réel $\mathcal{A} = \displaystyle\int_a^b f(t)\mathrm{d}t$ mesure, en unités d'aire, l'aire algébrique du domaine délimité par la courbe représentative de f, l'axe des abscisses et les droites d'équation $x = a$ et $x = b$.

Il découle de cette définition de l'intégrale les propriétés suivantes.

Proposition. 7.1. **Propriétés de l'intégrale**

Soient $f, g : [a,b] \to \mathbb{R}$ deux fonctions continues. On a les résultats suivants.

i. **Antisymétrie.** $\displaystyle\int_b^a f(t)\mathrm{d}t = -\int_a^b f(t)\mathrm{d}t.$

ii. **Relation de Chasles.** Pour tout $c \in [a,b]$, $\displaystyle\int_a^b f(t)\mathrm{d}t = \int_a^c f(t)\mathrm{d}t + \int_c^b f(t)\mathrm{d}t.$

iii. **Linéarité**. Pour tous $\lambda, \mu \in \mathbb{R}$, $\displaystyle\int_a^b (\lambda f(t) + \mu g(t))\,\mathrm{d}t = \lambda \int_a^b f(t)\mathrm{d}t + \mu \int_a^b g(t)\mathrm{d}t.$

iv. **Positivité.** On a : $\displaystyle[\forall t \in [a,b],\ f(t) \geq 0] \Longrightarrow \int_a^b f(t)\mathrm{d}t \geq 0.$

De plus :

$$\left.\begin{array}{l} \forall t \in [a,b],\ f(t) \geq 0 \\ \\ \displaystyle\int_a^b f(t)\mathrm{d}t = 0 \end{array}\right\} \Longrightarrow [\forall t \in [a,b],\ f(t) = 0]\,.$$

v. **Croissance.** On a : $\displaystyle[\forall t \in [a,b],\ f(t) \leq g(t)] \Longrightarrow \int_a^b f(t)\mathrm{d}t \leq \int_a^b g(t)\mathrm{d}t.$

vi. **Inégalité triangulaire.** On a : $\displaystyle\left|\int_a^b f(t)\mathrm{d}t\right| \leq \int_a^b |f(t)|\mathrm{d}t.$

Remarque 7.1.

Les propriétés *iv.*, *v.* et *vi.* deviennent fausses si jamais $a > b$. Auquel cas, on utilise la propriété *i.* pour obtenir des inégalités dans l'autre sens.

Les propriétés ci-dessous permettent de démontrer le résultat ci-dessous, fondamental comme son nom l'indique.

Théorème 7.2. **Théorème fondamental de l'analyse**

Soient f une fonction continue sur $[a,b]$ et $c \in [a,b]$. Alors, la fonction F définie sur $[a,b]$ par

$$F(x) = \int_c^x f(t)\mathrm{d}t$$

est l'unique primitive de f sur $[a,b]$ qui s'annule en c.
À ce titre, F est de classe $\mathcal{C}^1$ sur $[a,b]$, et $F' = f$.

Corollaire 7.3. **TFA, version 2**

Toute fonction continue sur $[a,b]$ admet des primitives sur ce même intervalle.

Corollaire 7.4.

Soit f une fonction continue sur $[a,b]$ et F une de ses primitives. Alors

$$\int_a^b f(t)\mathrm{d}t = [F(t)]_a^b = F(b) - F(a).$$

Exercice 7.1.

Calculer, en primitivant *à vue*, les intégrales suivantes

$$i.\ \int_0^1 \sqrt{t}\mathrm{d}t, \qquad ii.\ \int_1^2 \frac{x}{x^2+1}\mathrm{d}x, \qquad iii.\ \int_0^{\ln(2)} te^{t^2}\mathrm{d}t, \qquad iv.\ \int_1^x \frac{\mathrm{d}s}{s}, \qquad v.\ \int_0^\pi \sin^2(\theta)\mathrm{d}\theta.$$

Exercice 7.2.

Calculer les limites, lorsque $n \to +\infty$, de

$$i.\ \frac{1}{n\sqrt{n}}\sum_{k=1}^n \sqrt{k}, \quad ii.\ \sum_{k=1}^n \frac{1}{n+k}, \quad iii.\ \sum_{k=1}^n \frac{n}{n^2+k^2}.$$

Exercice 7.3. **Fonction définie par une intégrale**

Soit F la fonction définie sur $]0;1[$ par $F(x) = \displaystyle\int_{x^2}^{x} \frac{\mathrm{d}t}{\ln(t)}$.

1. Montrer que F est de classe $\mathcal{C}^1$ sur $]0,1[$ puis étudier ses variations.
2. Montrer que F peut se prolonger en une fonction continue sur $[0;1]$.

Remarque 7.2.

Si f est une fonction définie sur un intervalle $[a,b]$ et à valeurs dans $\mathbb{C}$, alors on peut écrire, pour tout $t \in [a,b]$, $f(t) = \mathrm{Re}(f)(t) + i\ \mathrm{Im}(f)(t)$ et définir l'intégrale de f sur $[a,b]$ par la formule

$$\int_a^b f(t)\mathrm{d}t = \int_a^b \mathrm{Re}(f)(t)\mathrm{d}t + i\int_a^b \mathrm{Im}(f)(t)\mathrm{d}t.$$

On peut par exemple écrire, pour tout $\lambda \in \mathbb{C}^*$, $\displaystyle\int_a^b e^{\lambda t}\mathrm{d}t = \frac{e^{\lambda b} - e^{\lambda a}}{\lambda}$.

Exercice 7.4.

Calculer $\displaystyle\int_0^{\pi} \cos(\theta)e^{\theta}\mathrm{d}\theta$.

7.1.2 Méthodes pratiques de calcul d'intégrales

⋆ Intégration par parties.

Théorème 7.5. **Intégration par parties**

Soient $u, v : [a,b] \to \mathbb{R}$ deux fonctions de classe $\mathcal{C}^1$ sur $[a,b]$. On a alors

$$\int_a^b u'(t)v(t)\mathrm{t} = \Big[u(t)v(t)\Big]_a^b - \int_a^b u(t)v'(t)\mathrm{d}t.$$

☞ Lorsque l'on utilisera une intégration par parties, on n'oubliera pas de vérifier que les fonctions u, v sont $\mathcal{C}^1$ sur l'intervalle considéré, et **on fera apparaître cette hypothèse sur sa copie**.

Exercice 7.5.

Donner une primitive de $x \mapsto \ln(x)$ puis, avec la même méthode de $x \mapsto \arccos(x)$.

Exercice 7.6.

Pour tout entier n, on pose $I_n = \displaystyle\int_0^1 t^n e^{-t}\mathrm{d}t$.

1. **a.** Montrer que pour tout entier n, on a $0 \leq I_n \leq \dfrac{1}{n+1}$.
   b. En déduire que (I_n) converge, et déterminer sa limite.

2. **a.** Donner une relation entre I_n et I_{n+1}.
   b. En déduire un équivalent de I_n.

☞ Si (I_n) est une suite d'intégrales, une intégration par parties permet souvent de déterminer une relation entre I_n et I_{n+1}, dont on peut ensuite déduire une expression de I_n en fonction de n.

⋆ Intégrales de la forme $\displaystyle\int P(x)e^{ax}\mathrm{d}x$, $\displaystyle\int P(x)\cos(bx)e^{ax}\mathrm{d}x$ et $\displaystyle\int P(x)\sin(bx)e^{ax}\mathrm{d}x$ où P est un polynôme et $a \in \mathbb{C}$

Pour intégrer le produit d'un polynôme et d'une exponentielle, il suffit d'effectuer des intégrations par parties successives, en dérivant le polynôme (jusqu'à obtenir un terme constant, ce qui requiert donc $\deg(P)$ intégrations par parties successives). En identifiant $\cos(ax)$ et $\sin(ax)$ comme les parties réelles et imaginaires de e^{iax}, la même méthode fonctionne pour intégrer les fonctions ci-dessus.

Exercice 7.7.

Calculer les intégrales $\displaystyle\int_0^1 (x^2+x+1)e^{-x}\mathrm{d}x$ et $\displaystyle\int_0^1 x^2\cos(2x)e^{-x}\mathrm{d}x$.

⋆ Changement de variable.

La formule de dérivation d'une fonction composée permet d'établir la formule de changement de variable.

Théorème 7.6. **Formule de changement de variable**

Soient α,β des réels tels que $\alpha<\beta$. On se donne une fonction continue $f:[a,b]\to\mathbb{R}$, et $\varphi:[\alpha,\beta]\to[a,b]$ une fonction de classe $\mathcal{C}^1$ sur $[\alpha,\beta]$. Alors

$$\int_\alpha^\beta f(\varphi(x))\varphi'(x)\mathrm{d}x=\int_{\varphi(\alpha)}^{\varphi(\beta)} f(t)\mathrm{d}t.$$

☞ Il est capital de se rappeler (et d'écrire sur sa copie) que la fonction φ doit être de classe $\mathcal{C}^1$.
Dans la majeure partie des exercices (mais pas toujours), le changement de variable est donné, sous la forme $u=u(t)$. On vérifie donc que u est de classe $\mathcal{C}^1$. On regarde ce que deviennent les bornes et surtout on oublie pas que $\mathrm{d}u=u'(t)\mathrm{d}t$.

Exercice 7.8.

Donner une primitive de $x\mapsto\dfrac{\sin(x)}{1+\cos^2(x)}$ sur $\mathbb{R}$ en posant $u=\cos(x)$.

⋆ Calcul d'intégrales de la forme $\int R\left(e^x\right)\mathrm{d}x$ où R est une fraction rationnelle.

Pour intégrer une fraction rationnelle (*i.e.* un quotient de polynômes) en l'exponentielle, on pose $u=e^x$, ce qui permet de se ramener à une fraction rationnelle en u. Si cette dernière ne s'intègre pas aisément, on utilise la méthode détaillée dans la suite.

Exercice 7.9.

En posant $u=e^x$, calculer l'intégrale $\displaystyle\int_0^1\frac{e^x}{e^{2x}+1}\mathrm{d}x$.

⋆ Calcul d'intégrales de la forme $\int\sin^q(x)\cos^p(x)\mathrm{d}x$ où p,q sont des entiers naturels.

- ✗ Si p ou q est impair on effectue un changement de variable. Plus précisément, si p est impair, on pose $u=\sin(x)$ et si q est impair, on pose $u=\cos(x)$.
- ✗ Si p et q sont pairs, on linéarise le produit $\cos^p(x)\sin^q(x)$ à l'aide des formules d'Euler.

Exercice 7.10.

Calculer les intégrales $\displaystyle\int_0^{\pi/2}\sin^3(x)\cos^2(x)\mathrm{d}x$ et $\displaystyle\int_0^{\pi}\cos^2(x)\mathrm{d}x$.

⋆ Calcul d'intégrales de la forme $\int\frac{P(x)}{x^2+px+q}\mathrm{d}x$ où P est un polynôme de degré supérieur ou égal à 2.

Pour intégrer une fraction rationnelle de la forme $\dfrac{P(x)}{x^2+px+q}$ où P est un polynôme tel que $\deg(P)>1$, on commence par effectuer la division euclidienne de P par x^2+px+q, ce qui permet de se ramener à une fraction rationnelle de la forme $\dfrac{ax+b}{x^2+px+q}$.
On calcule alors les racines du polynôme X^2+pX+q puis on utilise éventuellement le résultat suivant (admis) pour faire apparaître une combinaison d'éléments simples à primitiver.

Théorème 7.7. **Décomposition en éléments simples**

Soient a, b, p, q des réels. On note Δ le discriminant du polynôme du second degré $x^2 + px + q$.

i. Si $\Delta > 0$, le polynôme possède deux racines distinctes c et d, et il existe alors deux réels α, β tels que
$$\forall x \in \mathbb{R} \setminus \{c, d\}, \ \frac{ax+b}{x^2+px+q} = \frac{ax+b}{(x-c)(x-d)} = \frac{\alpha}{x-c} + \frac{\beta}{x-d}.$$

ii. Si $\Delta = 0$, le polynôme possède une racine double c, et il existe alors deux réels α, β tels que
$$\forall x \in \mathbb{R} \setminus \{c\}, \ \frac{ax+b}{x^2+px+q} = \frac{ax+b}{(x-c)^2} = \frac{\alpha}{(x-c)^2} + \frac{\beta}{x-c}$$
.

On distingue les cas pour la poursuite du calcul. En reprenant les notations du théorème ci-dessus,

✗ **1er cas.** Si $\Delta > 0$,
$$\int \frac{ax+b}{x^2+px+q} \mathrm{d}x = \int \frac{\alpha}{x-c} \mathrm{d}x + \int \frac{\beta}{x-d} \mathrm{d}x,$$

où chaque intégrale se calcul à l'aide de logarithmes.

✗ **2ème cas.** Si $\Delta = 0$,
$$\int \frac{ax+b}{x^2+px+q} \mathrm{d}x = \int \frac{\alpha}{(x-c)^2} \mathrm{d}x + \int \frac{\beta}{x-c} \mathrm{d}x,$$

où la première intégrale se calcule en reconnaissant la dérivée d'une inverse et la seconde avec un logarithme.

✗ **3ème cas.** Si $\Delta < 0$, on fait alors apparaître au numérateur la dérivée du dénominateur, en cherchant deux réels α, β tels que
$$\int \frac{ax+b}{x^2+px+q} \mathrm{d}x = \alpha \int \frac{2x+p}{x^2+px+q} \mathrm{d}x + \beta \int \frac{1}{x^2+px+q} \mathrm{d}x.$$

La première intégrale se calcule en observant que $\frac{2x+p}{x^2+px+q}$ est la dérivée de $\ln|x^2+px+q|$.

Pour calculer la seconde, on met le polynôme $x^2 + px + q$ sous forme canonique:
$$\int \frac{1}{x^2+px+q} \mathrm{d}x = \int \frac{1}{\left(x+\frac{p}{2}\right)^2 + q - \frac{p^2}{4}} \mathrm{d}x = \int \frac{1}{\left(x+\frac{p}{2}\right)^2 + \left(\sqrt{q-\frac{p^2}{4}}\right)^2} \mathrm{d}x$$

puisque le discriminant du dénominateur est strictement négatif (on a donc $q - p^2/4 > 0$).

On pose alors $u = x + p/2$ ce qui permet de se ramener à une intégrale de la forme $\displaystyle\int \frac{du}{u^2 + \left(\sqrt{q-\frac{p^2}{4}}\right)^2}$, que l'on peut calculer en utilisant la formule $\displaystyle\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{a} \arctan\left(\frac{x}{a}\right)\right) = \frac{1}{x^2+a^2}$.

Exercice 7.11.

Calculer les intégrales
$$i. \int_0^{1/2} \frac{2x+1}{x^2+x-2} \mathrm{d}x, \quad ii. \int_0^{1/2} \frac{2x+1}{x^2-2x+1} \mathrm{d}x, \quad iii. \int_0^{1} \frac{-3x+2}{x^2-x+1} \mathrm{d}x.$$

7.1.3 Primitives usuelles

Chacune des lignes du tableau ci-dessous se lit comme suit: la fonction f admet (sur son intervalle de définition) F pour primitive, éventuellement sous la condition *cond*. La lettre u désigne une fonction, et c la constante d'intégration.

$f(x)$	$F(x)$	*cond*
e^{zx}	$\frac{1}{z}e^{zx}+c$	$z\in\mathbb{C}^*$
$\dfrac{u'(x)}{u(x)}$	$\ln\lvert u(x)\rvert+c$	
x^α	$\dfrac{1}{\alpha+1}x^{\alpha+1}+c$	$\alpha\neq-1$
$\alpha u'u^{\alpha-1}$	$u^\alpha+c$	$\alpha\neq 0$
$\dfrac{u'}{2\sqrt{u}}$	$\sqrt{u}+c$	
a^x	$\frac{a^x}{\ln a}+c$	$a>0, a\neq 1$
$\cos(ax)$	$\dfrac{1}{a}\sin(ax)+c$	$a\neq 0$
$\sin(ax)$	$-\dfrac{1}{a}\cos(ax)+c$	$a\neq 0$
$\dfrac{1}{\cos^2(x)}=1+\tan^2(x)$	$\tan(x)+c$	
$\dfrac{1}{\sin^2(x)}=1+\cotan^2(x)$	$\cotan(x)+c$	
$\dfrac{1}{a^2+x^2}$	$\dfrac{1}{a}\arctan\left(\dfrac{x}{a}\right)+c$	$a\neq 0$
$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin(x)+c$	
$-\dfrac{1}{\sqrt{1-x^2}}$	$\arccos(x)+c$	

7.2 Intégrale impropre et intégrale convergente

7.2.1 Définitions

Définition 7.2. **Intégrale impropre**

Soient $a,b\in\overline{\mathbb{R}}$ (avec $a<b$) et f une fonction continue sur l'intervalle **semi-ouvert** $[a;b[$ qui peut être borné, ou non). On dit que l'intégrale

$$\int_a^b f(t)\mathrm{d}t$$

est **impropre** en b. Si la fonction est continue sur l'intervalle semi-ouvert $]a;b]$, l'intégrale précédente est alors impropre en a.

Le caractère impropre en a (ou en b) peut avoir plusieurs origines : la fonction f n'est pas définie en a ou la borne a est elle-même infinie ce qui dans les deux cas se représentent graphiquement par une surface "infinie".

Il existe un cas qu'on peut immédiatement traiter.

Proposition. 7.8.

Soit f une fonction continue sur un intervalle borné semi-ouvert $[a;b[$ admettant une limite finie en b (*i.e.* prolongeable par continuité à b). On dit que l'intégrale

$$\int_a^b f(t)\mathrm{d}t$$

est **faussement impropre** (en b) et que c'est une intégrale ***convergente***.

Exercice 7.12.

Que peut-on dire des intégrales suivantes?

$$i.\ \int_0^1 t\ln(t)\mathrm{d}t,\qquad ii.\ \int_0^1 \frac{e^x-1}{x}\mathrm{d}x$$

Le caractère prolongeable par continuité est suffisant à donner un sens à l'intégrale, mais pas nécessaire comme on va le voir avec la définition et les exemples suivants.

Définition 7.3. **Convergence d'une intégrale impropre (d'un seul côté)**

Soit f une fonction continue sur l'intervalle semi-ouvert $[a\,;\,b[$ où $b \in \overline{\mathbb{R}}$.

i. On dit que l'intégrale (impropre à droite, donc en b) $\displaystyle\int_a^b f(t)\mathrm{d}t$ est **convergente** lorsque la fonction $x \mapsto \int_a^x f(t)\mathrm{d}t$ admet une limite finie quand x tend vers b, et dans ce cas, on pose

$$\int_a^b f(t)\mathrm{d}t = \lim_{x\to b}\int_a^x f(t)\mathrm{d}t.$$

ii. Sinon, on dit que l'intégrale **diverge**.

On définit de la même façon la notion d'intégrale convergente lorsque f est continue sur $]a; b]$ et que l'intégrale est impropre en a (à gauche donc) :

$$\int_a^b f(t)\mathrm{d}t = \lim_{x\to a}\int_x^b f(t)\mathrm{d}t.$$

☞ **Attention**, si l'intégrale est impropre des deux côtés, il faut traiter la convergence séparément pour chacune des bornes (voir ci-dessous).

☞ Étudier la nature d'une intégrale signifie déterminer si l'intégrale est convergente ou non.

Exercice 7.13.

Déterminer la nature des intégrales suivantes

$$i.\ \int_0^{+\infty} e^{-\alpha t}\mathrm{d}t, \qquad ii.\ \int_1^{+\infty}\frac{\mathrm{d}t}{t}, \qquad iii.\ \int_e^{+\infty}\frac{\mathrm{d}t}{t\ln(t)}, \qquad iv.\ \int_0^1\frac{\mathrm{d}t}{\sqrt{t}}.$$

Définition 7.4. **Convergence d'une intégrale impropre des deux côtés**

Soit f une fonction continue sur l'intervalle ouvert $]a, b[$ (où $a, b \in \overline{\mathbb{R}}$). Soit $c \in]a, b[$. On dit que l'intégrale $\displaystyle\int_a^b f(t)\mathrm{d}t$ converge si et seulement si **les deux intégrales**

$$\int_a^c f(t)\mathrm{d}t \qquad \textbf{et} \qquad \int_c^b f(t)\mathrm{d}t$$

convergent. Dans ce cas, on a

$$\int_a^b f(t)\mathrm{d}t = \int_a^c f(t)\mathrm{d}t + \int_c^b f(t)\mathrm{d}t.$$

Remarque 7.3.

La relation de Chasles prouve que la définition ci-dessus ne dépend pas du point c choisi.

Exercice 7.14.

Montrer que $\displaystyle\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{1+x^2}$ est convergente et calculer sa valeur.

Remarque 7.4.

On prendra garde au fait que, bien que l'on ait

$$\lim_{A\to+\infty}\int_{-A}^{A} x\mathrm{d}x = \lim_{A\to+\infty} 0 = 0,$$

l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x\mathrm{d}x$ ne converge pas, puisque $\displaystyle\int_0^{+\infty} x\mathrm{d}x$ et $\displaystyle\int_{-\infty}^{0} x\mathrm{d}x$ ne convergent pas!

On ne peut donc pas faire tendre simultanément les deux bornes vers l'infini. On doit vraiment montrer la convergence des deux intégrales impropre séparément.

On peut ensuite généraliser la notion d'intégrale aux fonctions continues sauf en un nombre fini de points.

Définition 7.5. **Intégrale impropre plusieurs fois**

Soient $a, b \in \overline{\mathbb{R}}$ tels que $a \leq b$.
On se donne encore $a_1, \dots, a_n$ dans $\overline{\mathbb{R}}$ tels que $a_1 = a < a_2 < \cdots < a_{n-1} < a_n = b$.
On considère alors une fonction f continue sur $]a, b[$, sauf éventuellement en chacun des points $a_1, \dots, a_n$.
On dit que l'intégrale de f sur $]a, b[$ converge si chacune des intégrales $\displaystyle\int_{a_i}^{a_{i+1}} f(t)\mathrm{d}t$ converge, et lorsque c'est le cas on pose

$$\int_a^b f(t)\mathrm{d}t = \sum_{i=1}^{n-1} \int_{a_i}^{a_{i+1}} f(t)\mathrm{d}t.$$

On peut enfin généraliser cette notion aux fonctions à valeurs complexes.

Définition 7.6. **Intégrale impropre à valeurs complexes**

Soient $a, b \in \overline{\mathbb{R}}$, tels que $a \leq b$. Soit f une fonction à valeurs dans $\mathbb{C}$, continue sur $]a, b[$ sauf éventuellement en un nombre fini de points.
On dit que l'intégrale de f sur $]a, b[$ converge si les intégrales impropres $\displaystyle\int_a^b \mathrm{Re}(f(t))\mathrm{d}t$ et $\displaystyle\int_a^b \mathrm{Im}(f(t))\mathrm{d}t$ convergent, et lorsque c'est le cas on pose

$$\int_a^b f(t)\mathrm{d}t = \int_a^b \mathrm{Re}(f(t))\mathrm{d}t + i\int_a^b \mathrm{Im}(f(t))\mathrm{d}t.$$

7.2.2 Intégrales de référence

Parmi les intégrales impropres dont il est facile de déterminer la nature, certaines ont un rôle central dans la mesure où elles apparaissent souvent et qu'on s'en sert pour justifier de la nature d'autres intégrales par comparaison.

Proposition. 7.9.

Soit $\alpha \in \mathbb{R}$. L'intégrale $\displaystyle\int_0^{+\infty} e^{-\alpha x}\mathrm{d}x$ converge si et seulement si $\alpha > 0$.
Auquel cas,

$$\int_0^{+\infty} e^{-\alpha x}\mathrm{d}x = \frac{1}{\alpha}.$$

Proposition. 7.10. **L'intégrale de Riemann en $+\infty$**

Soit $\alpha \in \mathbb{R}$. L'intégrale $\displaystyle\int_1^{+\infty} \frac{1}{x^\alpha}\mathrm{d}x$ converge si et seulement si $\alpha > 1$.
Auquel cas,

$$\int_1^{+\infty} \frac{1}{x^\alpha}\mathrm{d}x = \frac{1}{\alpha - 1}.$$

Remarque 7.5.

Dans la proposition précédente, le résultat reste vrai si l'on remplace la borne de gauche 1 par n'importe quel réel strictement positif.

Proposition. 7.11. **L'intégrale de Riemann en 0**

Soit $\beta \in \mathbb{R}$. L'intégrale $\displaystyle\int_0^1 \frac{1}{x^\beta}\mathrm{d}x$ converge si et seulement si $\beta < 1$.
Auquel cas,

$$\int_0^1 \frac{1}{x^\beta}\mathrm{d}x = \frac{1}{1 - \beta}.$$

Remarque 7.6.

On observera que les conditions de convergence en 0 et en $+\infty$ étant incompatibles, l'intégrale $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{x^\alpha}$ est divergente pour tout $\alpha \in \mathbb{R}$.

Corollaire 7.12.

Soit $\alpha \in \mathbb{R}$, et $a < b$ des réels. Les intégrales $\displaystyle\int_a^b \frac{1}{(t-a)^\alpha}\mathrm{d}t$ et $\displaystyle\int_a^b \frac{1}{(b-t)^\alpha}\mathrm{d}t$ convergent si et seulement si $\alpha < 1$.

Proposition. 7.13.

L'intégrale $\displaystyle\int_0^1 \ln(x)\mathrm{d}x$ converge et vaut -1.

7.2.3 Propriétés de l'intégrale généralisée

Dans cette section, on fixe $a \in \mathbb{R}$ et $b \in \overline{\mathbb{R}}$ tels que $a \leq b$. Les résultats suivants seront donc énoncés pour des intégrales impropres à droite (sauf mention du contraire), mais les résultats analogues à gauche sont également vrais.
Plus précisément, les résultats qui suivent étendent aux intégrales impropres à droite les propriétés algébriques élémentaires des intégrales rappelés en début de chapitre; bien entendu on dispose de résultats analogues concernant les intégrales impropres à gauche, ou plusieurs fois impropres.

Proposition. 7.14. **Antisymétrie**

Soit une fonction f continue sur $[a,b[$. Alors $\displaystyle\int_b^a f(t)\mathrm{d}t$ converge si et seulement si $\int_a^b f(t)\mathrm{d}t$ converge, et lorsque c'est le cas, on a

$$\int_b^a f(t)\mathrm{d}t = -\int_a^b f(t)\mathrm{d}t.$$

Proposition. 7.15. **Linéarité**

Soient deux fonctions f, g continues sur $[a,b[$.
Si $\displaystyle\int_a^b f(t)\mathrm{d}t$ et $\int_a^b g(t)\mathrm{d}t$ convergent, alors $\displaystyle\int_a^b (\lambda f(t) + \mu g(t))\,\mathrm{d}t$ converge, et on a

$$\int_a^b (\lambda f(t) + \mu g(t))\,\mathrm{d}t = \lambda\int_a^b f(t)\mathrm{d}t + \mu\int_a^b g(t)\mathrm{d}t.$$

☞ Le résultat ci-dessus permet notamment de démontrer que l'ensemble des fonctions admettant une intégrale sur $[a,b[$ est un sous-espace vectoriel de l'ensemble des fonctions définies sur $[a,b[$.

Remarque 7.7.

Avant d'écrire

$$\int_a^b (f(t) + g(t))\,\mathrm{d}t = \int_a^b f(t)\mathrm{d}t + \int_a^b g(t)\mathrm{d}t,$$

on vérifiera qu'au moins deux des trois intégrales considérées convergent; on pourra alors en déduire que la troisième converge, puis que la linéarité s'applique.
Mais attention à ne pas écrire qu'une intégrale convergente est somme de deux intégrales divergentes (voir l'exercice ci-dessous).

Exercice 7.15.

Calculer $\displaystyle\int_1^{+\infty} \frac{1}{t(t+1)}\mathrm{d}t$.

Proposition. 7.16. **Croissance**

Soient deux fonctions f, g continues sur $[a, b[$. On suppose que $\int_a^b f(t)\mathrm{d}t$ et $\int_a^b g(t)\mathrm{d}t$ convergent. On a alors le résultat suivant:

$$\forall t \in [a, b[,\ f(t) \leq g(t) \implies \int_a^b f(t)\mathrm{d}t \leq \int_a^b g(t)\mathrm{d}t.$$

Proposition. 7.17. **Positivité**

Soit $f : [a, b[\to \mathbb{R}$ continue et à valeurs positives telle que $\int_a^b f(t)\mathrm{d}t$ converge. On a alors les résultats suivants.

i. $\int_a^b f(t)\mathrm{d}t \geq 0$.

ii. $\int_a^b f(t)\mathrm{d}t = 0 \iff \forall t \in [a, b[,\ f(t) = 0$.

7.2.4 À propos d'une condition nécessaire de convergence

Proposition. 7.18.

Soit $a \in \mathbb{R}$, et $f : [a, +\infty[\to \mathbb{R}$ une fonction continue telle que f possède une limite ℓ en $+\infty$.
On a alors

$$\int_a^{+\infty} f(t)\mathrm{d}t \text{ converge} \implies \ell = 0.$$

Remarque 7.8.

Attention à bien comprendre cet énoncé ; la conclusion s'applique seulement si on sait que f admet une limite en $+\infty$. En effet, il est possible de trouver une fonction f continue, sans limite en $+\infty$, dont l'intégrale converge.

Exercice 7.16.

On considère la fonction f, affine par morceaux, nulle sur $\mathbb{R}_-$, définie et **continue** sur $\mathbb{R}$ comme suit : pour tout $n \in \mathbb{N}$,

- ✗ $f\left(n + \frac{1}{2}\right) = 1$;
- ✗ sur $[n, n+1]$, f est nulle en dehors de $\left[n + \frac{1}{2} - \frac{1}{n^2}; n + \frac{1}{2} + \frac{1}{n^2}\right]$.

1. **a**. Dessiner l'allure de la courbe de f.
b. Justifier que $f(x)$ n'admet pas de limite lorsque $x \to +\infty$.

2. Calculer, pour tout $n \in \mathbb{N}$, $\int_n^{n+1} f(t)\mathrm{d}t$ puis $\int_0^n f(t)\mathrm{d}t$. En déduire la convergence de $\int_0^{+\infty} f(t)\mathrm{d}t$.

Remarque 7.9.

Naturellement aussi, la réciproque est fausse. La fonction f peut avoir une limite finie nulle en $+\infty$ sans que l'intégrale converge (on aura un contre-exemple en tête à tout moment).
Si par contre la limite existe et n'est pas nulle, on est certain que l'intégrale diverge. On dit même qu'elle **diverge grossièrement**.

7.2.5 Comparaison pour des fonctions de signe constant

Lemme 7.19.

Soit une fonction f continue et positive sur $[a, b[$. Alors $\int_a^b f(t)\mathrm{d}t$ converge si et seulement si $x \mapsto \int_a^x f(t)\mathrm{d}t$ est majorée sur $[a, b[$, et lorsque c'est le cas, on a

$$\forall x \in [a, b],\ \int_a^x f(t)\mathrm{d}t \leq \int_a^b f(t)\mathrm{d}t.$$

Théorème 7.20. **Comparaison des fonctions positives**

Soient deux fonctions f, g continues sur $[a, b[$ tels que

$$\forall t \in [a, b[, \ 0 \leq f(t) \leq g(t)$$

On a alors les résultats suivants.

i. Si $\displaystyle\int_a^b g(t)\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\mathrm{d}t$ converge, et l'on a $\displaystyle\int_a^b f(t)\mathrm{d}t \leq \int_a^b g(t)\mathrm{d}t$.

ii. Si $\displaystyle\int_a^b f(t)\mathrm{d}t$ diverge, alors $\displaystyle\int_a^b g(t)\mathrm{d}t$ diverge.

☞ Quitte à changer f en $-f$, on peut adapter ce résultat au cas des fonctions négatives. **L'essentiel ici est donc que la fonction soit de signe constant**.

Exercice 7.17.

Justifier que l'intégrale $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{x^3+1}$ est convergente.

Théorème 7.21. **Comparaison des équivalents de signe constant**

Soient $f, g : [a, b[\to \mathbb{R}$ deux fonctions continues et de signe constant au voisinage de b.

Si $f \underset{b}{\sim} g$, alors $\displaystyle\int_a^b f(t)\mathrm{d}t$ et $\displaystyle\int_a^b g(t)\mathrm{d}t$ sont de même nature.

Exercice 7.18.

Justifier que l'intégrale $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{\mathrm{ch}(x)}$ est convergente.

Exercice 7.19.

Montrer que $\displaystyle\int_0^{+\infty} \frac{\ln(t)}{1+t^2}\mathrm{d}t$ converge.

7.2.6 Calculs d'intégrales généralisées

⋆ Utilisation d'une primitive

Proposition. 7.22.

Soient $a, b \in \overline{\mathbb{R}}$ tels que $a \leq b$. Soient f une fonction continue sur $]a, b[$, et F une primitive de f sur $]a, b[$.

Si $\displaystyle\int_a^b f(t)\mathrm{d}t$ converge, on a:

$$\int_a^b f(t)\mathrm{d}t = \lim_{x \to b^-} F(x) - \lim_{x \to a^+} F(x).$$

⋆ Intégration par parties

Théorème 7.23. **IPP généralisée**

Soient $a, b \in \overline{\mathbb{R}}$, tels que $a \leq b$, et u, v deux fonctions de classe $\mathcal{C}^1$ sur $]a, b[$.
Si le produit uv admet des limites finies en a^+ et b^-, alors:

i. Les deux intégrales $\displaystyle\int_a^b u'(t)v(t)\mathrm{d}t$ et $\displaystyle\int_a^b u(t)v'(t)\mathrm{d}t$ sont de même nature;

ii. $\displaystyle\int_a^b u'(t)v(t)\mathrm{d}t = \Big[u(t)v(t)\Big]_a^b - \int_a^b u(t)v'(t)\mathrm{d}t$ lorsque les intégrales convergent.

Exercice 7.20.

Calculer $\displaystyle\int_0^{+\infty} te^{-t}\mathrm{d}t$.

⋆ Changement de variable

Théorème 7.24. **Changement de variable généralisé**

Soient $a, b \in \overline{\mathbb{R}}$ tels que $a \leq b$. Soient φ une bijection strictement monotone de classe $\mathcal{C}^1$ sur $]a, b[$, et f une fonction continue sur l'intervalle $]\alpha, \beta[$ où $\alpha = \lim\limits_{x \to a^+} \varphi(x)$ et $\beta = \lim\limits_{x \to b^-} \varphi(x)$.

Alors, les intégrales $\displaystyle\int_\alpha^\beta f(t)\mathrm{d}t$ et $\displaystyle\int_a^b f(\varphi(u))\varphi'(u)\mathrm{d}u$ sont de même nature, et si elles convergent alors elles sont égales.

☞ Lorsque l'on posera $t = \varphi(u)$ dans $\displaystyle\int_a^b f(\varphi(u))\varphi'(u)\mathrm{d}u$, on n'oubliera pas de vérifier que φ est de classe $\mathcal{C}^1$ et **strictement monotone** sur $]a, b[$.

Exercice 7.21.

Calculer $\displaystyle\int_0^{+\infty} \frac{1}{\sqrt{t}(1+t)}\mathrm{d}t$ en posant $u = \sqrt{t}$.

7.3 Fonctions intégrables

Sauf mention explicite du contraire, I dans cette section désigne un intervalle ouvert de $\mathbb{R}$.

7.3.1 Définition

Définition 7.7. **Fonction intégrable**

Soit f une fonction continue sur I, à valeurs dans $\mathbb{K}$.
On dit que f est **intégrable** sur I si $\displaystyle\int_I |f(t)|\mathrm{d}t$ est convergente. On note $L^1(I)$ l'ensemble des fonctions intégrables sur I.

Lorsque f est intégrable sur I, on dit aussi que l'intégrale $\displaystyle\int_I f(t)\mathrm{d}t$ est **absolument convergente.**

Remarque 7.10.

- ✗ Si f est de signe constant, alors cette notion équivaut à la convergence de l'intégrale impropre $\displaystyle\int_I f(t)\mathrm{d}t$.
- ✗ Si f est continue sur $I = [a, b]$ segment, alors f est automatiquement intégrable: on a une intégrale ordinaire.

7.3.2 Propriétés des fonctions intégrables

Théorème 7.25.

Si f est intégrable sur I, alors $\displaystyle\int_I f(t)\mathrm{d}t$ converge.

☞ En d'autres termes, la convergence absolue entraîne la convergence "simple".

Remarque 7.11.

- ✗ Ce théorème donne une condition **suffisante** de convergence d'une intégrale impropre: pour prouver que $\displaystyle\int_a^b f(t)\mathrm{d}t$ converge, il suffit de prouver que $\displaystyle\int_a^b |f(t)|\mathrm{d}t$ converge.
- ✗ La réciproque du théorème ci-dessus est fausse (voir l'**Exercice 7.23.**).
  Une telle intégrale est dite *semi-convergente.*

Exercice 7.22.

Montrer que $\displaystyle\int_1^{+\infty} \frac{\cos x}{x^2}\mathrm{d}x$ converge.

Exercice 7.23. **Intégrale semi-convergente**

Montrer que $\displaystyle\int_0^{+\infty} \frac{\sin t}{t} \mathrm{d}t$ converge mais que $t \mapsto \dfrac{\sin t}{t}$ n'est pas intégrable sur $[0, +\infty[$.

Théorème 7.26. **Structure de s-ev**

Si f, g sont intégrables sur I et si λ, μ sont des scalaires, alors $\lambda f + \mu g$ est intégrable sur I de sorte que:

$$\int_I (\lambda f(t) + \mu g(t))\,\mathrm{d}t = \lambda \int_I f(t)\mathrm{d}t + \mu \int_I g(t)\mathrm{d}t.$$

☞ L'ensemble $L^1(I)$ est donc un sous-espace vectoriel de l'ensemble des fonctions définies sur celui-ci.

Théorème 7.27. **Inégalité triangulaire**

Si f est intégrable sur I, alors $\displaystyle\left|\int_I f(t)\mathrm{d}t\right| \leq \int_I |f(t)|\,\mathrm{d}t.$

Théorème 7.28. **Intégrale nulle**

Si f est continue et intégrable sur I, alors

$$\int_I |f(t)|\mathrm{d}t = 0 \Longrightarrow (\forall t \in I,\ f(t) = 0)\,.$$

7.3.3 Théorèmes de comparaison pour les fonctions intégrables

Théorème 7.29. **Comparaison des équivalents**

Soient f, g deux fonctions continues sur $I = [a, b[$ telles que $f \underset{b}{\sim} g$. Alors f est intégrable sur $[a, b]$ si et seulement si g est intégrable sur $[a, b]$.

Théorème 7.30. **Comparaison des o et des O à droite**

Soient f, g deux fonctions continues sur $I = [a, b[$. Si g est intégrable sur $[a, b]$ et si $f \underset{b}{=} o(g)$ (ou si $f \underset{b}{=} O(g)$), alors f est intégrable sur $[a, b]$.

☞ On a des résultats analogues si f, g sont continues sur $I =]a, b]$ et si les relations de comparaisons ci-dessus ont lieu à gauche (en a).

Exercice 7.24.

Montrer que $\displaystyle\int_1^{+\infty} t\sin(t)\ln(t)e^{-t}\mathrm{d}t$ converge.

Corollaire 7.31. **Test de Riemann à l'infini**

Soit f continue sur $[a, +\infty[$ avec $a > 0$. S'il existe $\alpha > 1$ tel que

$$\lim_{t\to+\infty} t^\alpha |f(t)| = 0$$

alors f est intégrable sur $[a, +\infty[$.

7.4 Permutation somme et intégrale

La permutation de limites et une opération toujours délicate et qu'on ne peut pas systématiquement faire. On conclut alors ce chapitre avec un théorème d'interversion dont les hypothèses doivent être rigoureusement vérifiées avant application.

Théorème 7.32. **Permutation somme et intégrale**

Soit $(f_n)_{n\geq 0}$ une suite de fonctions intégrables sur un même intervalle I telles que, pour tout $t \in I$, la série $\sum_{n\geq 0} f_n(t)$ est convergente. On note alors

$$f : t \in I \longmapsto \sum_{n=0}^{+\infty} f_n(t).$$

On suppose que :

i. f est continue sur I, sauf éventuellement aux bords de I ;

ii. La série $\sum_{n\geq 0} \int_I |f_n(t)| \mathrm{d}t$ est convergente.

Alors, f est intégrable sur I et

$$\int_I f(t)\mathrm{d}t = \sum_{n=0}^{+\infty} \int_I f_n(t)\mathrm{d}t.$$

Exercice 7.25.

Montrer que

$$\int_0^1 \frac{\ln(t)}{t-1}\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{1}{(n+1)^2}.$$

8 Réduction des endomorphismes

Dans tout ce chapitre, la lettre $\mathbb{K}$ désignera l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$.

Dans différentes situation rencontrées préalablement, on a déjà pu observer qu'un endomorphisme pouvait être représenté, dans des bases différentes, par des matrices différentes et dans certains cas des matrices plus "simples" ou plus "pratiques" pour le calcul, notamment par des matrices diagonales.

En effet, si A et D représentent un même endomorphisme f (défini sur un espace E de dimension n) dans deux bases différentes, notées respectivement $\mathcal{B}_1$ et $\mathcal{B}_2$, alors les matrices A et D sont semblables, le lien entre les deux matrices fait intervenir la matrice de passage $P = P_{\mathcal{B}_1,\mathcal{B}_2}$

$$A = PDP^{-1}$$

ce qui implique notamment que, pour tout $k \in \mathbb{N}$ (et même $k \in \mathbb{Z}$ si f est un automorphisme ou de manière équivalence si A est inversible)

$$A^k = PD^kP^{-1}$$

et si D est diagonale, ses puissances se calculent très facilement

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} \Longrightarrow D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n^k \end{pmatrix}.$$

L'objectif de ce chapitre est, partant d'une matrice carrée A de taille n, de trouver, **si possible**, une base de $\mathcal{M}_{n,1}(\mathbb{K})$ (et donc une matrice de passage P) de sorte que $P^{-1}AP$ soit une **matrice diagonale** ou triangulaire.

8.1 Endomorphismes et matrices diagonalisables

8.1.1 Exemples introductifs

Lors de l'introduction dans les **Chapitre 2 et 4** des matrices de projection et de symétrie, on a vu qu'on pouvait trouver des bases dans lesquelles les matrices de ces endomorphismes étaient diagonales.

Plus précisément, supposons que E est de dimension $n \in \mathbb{N}^*$. Soit F_1 un sous-espace de E de dimension k, F_2 un supplémentaire de F_1 dans E et p le projecteur sur F_1 parallèlement à F_2.
Alors, en prenant une base $\mathcal{B} = (e_1, \dots, e_k, e_{k+1}, \dots, e_n)$ adaptée à la décomposition $E = F_1 \oplus F_2$, on obtient :

$$\mathrm{Mat}_{\mathcal{B}}(p) = \begin{pmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & 1 & & & & & \\ \vdots & & \ddots & & & & \vdots \\ & & & 1 & & & \\ \vdots & & & & 0 & & \vdots \\ & & & & & \ddots & \\ 0 & 0 & \cdots & & & & 0 \end{pmatrix} = \left(\begin{array}{c|c} I_k & 0_{k,n-k} \\ \hline 0_{n-k,k} & 0_{n-k} \end{array}\right)$$

qui est bien diagonale. En effet, on a:

- ✗ $p(e_i) = e_i$ si $i \in [\![1, k]\!]$, car dans ce cas $e_i \in F_1$;
- ✗ $p(e_i) = 0_E$ si $i \in [\![k+1, n]\!]$, car dans ce cas $e_i \in F_2$.

Exercice 8.1.
Justifier que toute symétrie admet une matrice diagonale, dans une base bien choisie.

Remarque 8.1.
Il existe des endomorphismes qui ne peuvent pas être représentés par une matrice diagonale (voir l'**Exercice 8.2**).

8.1.2 Définition

Définition 8.1. **Endomorphismes et matrices diagonalisables**
Supposons que E est de dimension $n \in \mathbb{N}^*$.

- ✗ Un endomorphisme f est dit **diagonalisable** s'il existe une base $\mathcal{B}$ de E dans laquelle la matrice de f est diagonale.
- ✗ Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **diagonalisable** s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ et une matrice diagonale $D \in \mathcal{M}_n(\mathbb{K})$ telles que $A = PDP^{-1}$, ou en d'autres termes si elle est semblable à une matrice diagonale.

Théorème 8.1.
Un endomorphisme est diagonalisable si et seulement si l'une de ses matrices est diagonalisable. Auquel cas, toutes ses matrices sont diagonalisables.

8.2 Éléments propres d'un endomorphisme

L'observation suivante est fondamentale.

Si f possède, dans une base $\mathcal{B} = (e_1, \dots, e_n)$, une matrice diagonale de la forme

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

alors on a :

$$\forall k \in [\![1, n]\!],\ f(e_k) = \lambda_k e_k$$

c'est-à-dire que $\mathcal{B}$ est une base de E formée de vecteurs colinéaires à leur image par f. Et cette condition est clairement nécessaire et suffisante. Elle motive les définitions suivantes.

8.2.1 Vecteurs propres, valeurs propres

Définition 8.2. **Valeur propre, vecteur propre**
Soit $\lambda \in \mathbb{K}$.

- ✗ λ est appelé une **valeur propre** de f s'il existe un vecteur **non nul** $x \in E$ tel que $f(x) = \lambda x$.
- ✗ On appelle **vecteur propre de** f **associé à** λ tout vecteur **non nul** $x \in E$ tel que $f(x) = \lambda x$.
- ✗ L'ensemble des valeurs propres de f est appelé **spectre** de f et est noté $\mathrm{Sp}(f)$.
- ✗ Si $\lambda \in \mathrm{Sp}(f)$, on note $E_\lambda(f)$ (ou E_λ si cela ne prête pas à confusion) l'ensemble $\mathrm{Ker}(\lambda \mathrm{Id}_E - f)$: c'est l'ensemble des vecteurs propres associés à λ auquel on a ajouté 0_E.

Remarque 8.2.

- ✗ La relation $f(x) = \lambda x$ est appelée **équation aux éléments propres**.
- ✗ Un vecteur propre n'est jamais nul.
- ✗ Lorsque l'on est en dimension finie, cette définition s'étend naturellement aux matrices carrés: on parle de vecteurs propres ou de valeurs propres associés à une matrice carrée. Par exemple, si $M \in \mathcal{M}_n(\mathbb{K})$, alors $\lambda \in \mathbb{K}$ est une valeur propre de M s'il existe $X \in \mathcal{M}_{n,1}(\mathbb{K})$ non nul tel que $MX = \lambda X$, et X est alors un vecteur propre de M associé à λ.

Théorème 8.2.
Soit $\lambda \in \mathbb{K}$. On a les équivalences suivantes :

$$\begin{aligned} \lambda \in \mathrm{Sp}(f) &\iff \exists x \in E \setminus \{0\},\ f(x) = \lambda x \\ &\iff \mathrm{Ker}(\lambda \mathrm{Id}_E - f) \neq \{0_E\} \\ &\iff \lambda \mathrm{Id}_E - f \text{ n'est pas injective} \end{aligned}$$

Et si E est de dimension finie on peut ajouter

$$\begin{aligned} \lambda \in \mathrm{Sp}(f) &\iff \lambda \mathrm{Id}_E - f \text{ n'est pas bijective} \\ &\iff \det(\lambda \mathrm{Id}_E - f) = 0 \end{aligned}$$

Corollaire 8.3.
Soit $\lambda \in \mathbb{K}$. On a les résultats suivants :

i. $E_\lambda = \mathrm{Ker}(\lambda \mathrm{Id}_E - f)$, et donc E_λ est un sous-espace vectoriel de E.
ii. Le scalaire 0 est une valeur propre associée à f si et seulement si f n'est pas injectif.

8.2.2 Notion de polynôme caractéristique et conséquences en dimension finie

On suppose dans cette section que E est de dimension $n \in \mathbb{N}^*$.

Théorème 8.4.
L'application $x \in \mathbb{K} \longmapsto \det(x \cdot \mathrm{Id}_E - f)$ est polynomiale de degré n, et son cœfficient dominant vaut 1.

Preuve. Résultat admis. □

Définition 8.3. **Polynôme caractéristique**
On appelle **polynôme caractéristique** de f, et on note $\chi_f(X)$, le polynôme défini par

$$\chi_f(X) = \det(X \cdot \mathrm{Id}_E - f).$$

Remarque 8.3.

- ✗ Le **Théorème 8.4** prouve que χ_f est bien un polynôme, qu'il est unitaire, et que son degré est égal à la dimension de l'espace.
- ✗ On définit de même le polynôme caractéristique d'une matrice $M \in \mathcal{M}_n(\mathbb{K})$ en posant $\chi_M(X) = \det(X \cdot I_n - M)$, qui est égal au polynôme caractéristique de tout endomorphisme associé.
- ✗ Le polynôme caractéristique ne dépend pas de la base dans laquelle on se place pour le calculer (car si deux matrices sont semblables elles ont le même polynôme caractéristique).

Théorème 8.5.
Pour tout $\lambda \in \mathbb{K}$ on a:
$$\lambda \in \mathrm{Sp}(f) \iff \chi_f(\lambda) = 0.$$
En d'autres termes, les valeurs propres de f sont les racines de son polynôme caractéristique χ_f.

Preuve. Il suffit de reprendre les équivalences vues précédemment:
$$\begin{aligned}\lambda \in \mathrm{Sp}(f) &\iff \mathrm{Ker}(f - \lambda \mathrm{Id}_E) \neq \{0_E\}\\ &\iff f - \lambda \mathrm{Id}_E \text{ n'est pas injective}\\ &\iff f - \lambda \mathrm{Id}_E \text{ n'est pas bijective}\\ &\iff \det(f - \lambda \mathrm{Id}_E) = 0 \iff \det(\lambda \mathrm{Id}_E - f) = 0\end{aligned}$$

□

Exercice 8.2.
Soit $f \in \mathcal{L}(\mathbb{R}^2)$ dont la matrice dans une base donnée est $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Déterminer $\mathrm{Sp}(f)$. En déduire que f n'est pas diagonalisable.

Proposition. 8.6. **Valeurs propres d'une matrice triangulaire**
Les valeurs propres d'une matrice triangulaire supérieure
$$T = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & * \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$
sont ses termes diagonaux $\lambda_1, \ldots, \lambda_n$.

Remarque 8.4.

- ✗ Ce résultat est en particulier vrai dans le cas des matrices diagonales.
- ✗ Ce résultat est encore vrai pour une matrice triangulaire inférieure (et la preuve est la même).

On peut enfin déduire de la notion de polynômes caractéristique des conséquences sur le nombre de valeurs propres d'un endomorphisme en dimension finie.

Théorème 8.7.
Toute matrice carrée à cœfficients complexes possède au moins une valeur propre complexe.

Preuve. Résulte du théorème de d'Alembert-Gauss. □

☞ On en déduit que toute matrice carrée à coefficients réels possède au moins une valeur propre complexe, mais celle-ci n'est **pas forcément réelle**.

Définition 8.4. **Multiplicité d'une valeur propre**
Soit $\lambda \in \mathrm{Sp}(f)$. On appelle **ordre de multiplicité de la valeur propre** λ de f, et on note $m(\lambda)$, l'ordre de multiplicité de λ en tant que racine du polynôme χ_f.

Corollaire 8.8. **Nombre de valeurs propres**

i. Si E est un $\mathbb{R}$-espace vectoriel, alors f possède au plus $n = \dim E$ valeurs propres **distinctes**.

ii. Si E est un $\mathbb{C}$-espace vectoriel, alors f possède exactement n valeurs propres **comptées avec leur ordre de multiplicité**.

Preuve. Résulte des **Théorèmes 8.4** et **8.5**. □

8.2.3 Propriétés des éléments propres

☞ Les résultats qui suivent sont vrais en dimension quelconque, éventuellement infinie.

Proposition. 8.9.
Soit $x \in E$ un vecteur non nul. Alors x est un vecteur propre de f si et seulement si $\mathrm{Vect}(x)$ est stable par f.

Proposition. 8.10.
Soit $\lambda \in \mathrm{Sp}(f)$. Alors E_λ est stable par f.

Proposition. 8.11.
Soit $\lambda \in \mathrm{Sp}(f)$. Si $\lambda \neq 0$ alors $E_\lambda \subset \mathrm{Im} f$, et de plus $E_0 = \mathrm{Ker}(f)$.

Le résultat suivant possède de nombreuses conséquences.

Proposition. 8.12. **Principe de concaténation**
Soit $p \in \mathbb{N}^*$. Si $u_1, \ldots, u_p$ sont des vecteurs propres associés à des valeurs propres $\lambda_1, \ldots, \lambda_p$ deux-à-deux distinctes de f, alors la famille $u_1, \ldots, u_p$ est libre.

Corollaire 8.13. **Principe de concaténation, version 2**

i. Toute somme finie de sous-espaces propres associés à des valeurs propres deux-à-deux distinctes est directe.
ii. Toute union disjointe de familles libres provenant de sous-espaces propres deux-à-deux distincts est libre.

Exercice 8.3.

1. Déterminer le spectre de $D: \begin{array}{ccc} \mathcal{C}^\infty(\mathbb{R}) & \to & \mathcal{C}^\infty(\mathbb{R}) \\ f & \mapsto & f' \end{array}$.

2. Pour tout réel λ on pose $f_\lambda : x \mapsto e^{\lambda x}$. En déduire que la famille $(f_\lambda)_{\lambda \in \mathbb{R}}$ est libre.

☞ Les résultats qui suivent ne sont vrais que si on suppose que $\dim(E) = n \in \mathbb{N}^*$.

Théorème 8.14.
Soit $\lambda \in \mathrm{Sp}(f)$ de multiplicité $m(\lambda)$. Alors:

$$1 \leq \dim \mathrm{Ker}(\lambda \mathrm{Id}_E - f) \leq m(\lambda).$$

Corollaire 8.15.
Soit $\lambda \in \mathrm{Sp}(f)$. Si $m(\lambda) = 1$, alors $\dim E_\lambda = \dim \mathrm{Ker}(\lambda \mathrm{Id}_E - f) = 1$.

Preuve. Résulte directement du théorème précédent. □

8.3 Quelques critères de réduction

On suppose dans cette section que E est de dimension $n \in \mathbb{N}^*$.

8.3.1 Critères de diagonalisabilité

On a déjà le résultat élémentaire suivant; il indique notamment que les endomorphismes diagonalisables sont ceux possédant *suffisament de vecteurs propres pour former une base.*

Théorème 8.16.

f est diagonalisable si et seulement si il existe une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E formée de vecteurs propres de f. Lorsque c'est le cas, si on appelle λ_i la valeur propre associée à e_i, on a

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Exercice 8.4.

Soit $n \geq 2$. Est-ce que l'endomorphisme $\Phi : \begin{array}{ccc} \mathbb{K}_n[X] & \to & \mathbb{K}_n[X] \\ P & \mapsto & P'' \end{array}$ est diagonalisable?

Rappel 8.1.

Un polynôme P est **scindé** s'il peut s'écrire comme un produit de polynômes de degré 1, c'est-à-dire s'il peut s'écrire

$$P = \alpha \prod_{k=1}^{p} (X - \lambda_k),$$

avec $p \in \mathbb{N}$, et $\alpha, \lambda_1, \ldots, \lambda_p$ des scalaires.

Ainsi, tout polynôme est scindé dans $\mathbb{C}[X]$, mais pas nécessairement dans $\mathbb{R}[X]$.

Exemple 8.1.

$(X-1)^2 X$ est scindé dans $\mathbb{R}[X]$, $(X^2+1)X = (X-i)(X+i)X$ est scindé dans $\mathbb{C}[X]$ mais pas dans $\mathbb{R}[X]$.

On peut déjà en déduire une condition suffisante de diagonalisabilité.

Théorème 8.17. **CS de diagonalisabilité**

Si f possède $n = \dim E$ valeurs propres distinctes, alors f est diagonalisable, et tous ses sous-espaces propres sont des droites.

En d'autres termes, si χ_f est scindé à racines simples, alors f est diagonalisable, et tous ses sous-espaces propres sont des droites.

Remarque 8.5.

Attention! Il est erroné de croire que si un endomorphisme est diagonalisable, alors il possède n valeurs propres distinctes! Un contre-exemple est donné par la matrice identité de taille $n \geq 2$, qui est diagonale et donc diagonalisable, bien qu'elle possède une seule valeur propre.

Exemple 8.2.

La matrice $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est diagonalisable dans $\mathcal{M}_2(\mathbb{C})$, alors qu'elle ne l'était pas dans $\mathcal{M}_2(\mathbb{R})$.

Exercice 8.5.

Déterminer sans calcul si $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix}$ et $C = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 0 & 3 \end{pmatrix}$ sont diagonalisables.

On peut enfin donner des conditions nécessaires et suffisantes de diagonalisabilité.

Théorème 8.18. **CNS de diagonalisabilité**

Les propriétés suivantes sont équivalentes.

i. f est diagonalisable;

ii. $E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} E_\lambda$;

iii. $\dim(E) = \sum_{\lambda \in \mathrm{Sp}(f)} \dim(E_\lambda)$;

iv. χ_f est scindé, et $\dim E_\lambda = m(\lambda)$ pour tout $\lambda \in \mathrm{Sp}(f)$.

☞ Pour diagonaliser un endomorphisme f de matrice A dans une base $\mathcal{B}$, on procède le plus souvent comme suit.

- ✗ On calcule son polynôme caractéristique $\det(X \cdot I_n - A)$, si possible sous forme factorisée, et on détermine ses racines. S'il n'est pas scindé, f n'est pas diagonalisable.

- ✗ Si χ_f est scindé, on détermine pour chaque valeur propre λ_k son sous-espace propre $\mathrm{Ker}(\lambda_k \cdot I_n - A)$ en résolvant le système linéaire

$$(\lambda_k \cdot I_n - A) \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

  Si, pour l'une au moins des valeurs propres la dimension du sous-espace propre est strictement inférieure à l'ordre de multiplicité, alors f n'est pas diagonalisable.

- ✗ Enfin, si f est diagonalisable, de valeurs propres $\lambda_1, \ldots, \lambda_p$ de multiplicités respectives $m_1, \ldots, m_p$ (telles que $\forall k \in [\![1, p]\!],\ m_k = \dim E_{\lambda_k}$) alors on a :

$$A = P \left(\begin{array}{c|c|c|c} \lambda_1 I_{m_1} & \mathbb{0}_{m_1,m_2} & \cdots & \mathbb{0}_{m_1,m_p} \\ \hline \mathbb{0}_{m_2,m_1} & \lambda_2 I_{m_2} & & \vdots \\ \hline \vdots & & \ddots & \vdots \\ \hline \mathbb{0}_{m_p,m_1} & \cdots & \cdots & \lambda_p I_{m_p} \end{array}\right) P^{-1},$$

  où P est la matrice de passage de la base $\mathcal{B}$ vers la base

$$\mathcal{B}' = (\underbrace{u_{1,1}, \ldots, u_{1,m_1}}_{\text{base de } E_{\lambda_1}}, \underbrace{u_{2,1}, \ldots, u_{2,m_2}}_{\text{base de } E_{\lambda_2}}, \ldots, \underbrace{u_{p,1}, \ldots, u_{p,m_p}}_{\text{base de } E_{\lambda_p}}).$$

  Enfin, on calcule P^{-1} à partir de P en cas de besoin (parfois seule la matrice diagonale est demandée).

Remarque 8.6.

L'ordre dans lequel apparaissent les valeurs propres sur la diagonale est l'ordre dans lequel on *concatène* les bases des sous-espaces propres associés.

Exercice 8.6.

Déterminer si chacune des matrices suivantes est diagonalisable, et si c'est le cas donner une matrice inversible P et une matrice diagonale D telle que cette matrice s'écrive PDP^{-1}.

$$i.\ A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad ii.\ B = \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}.$$

8.3.2 Critère de trigonalisabilité

Lorsqu'un endomorphisme n'est pas diagonalisable, on peut chercher à le *trigonaliser*.

Définition 8.5. **Endomorphismes et matrices trigonalisables**

Supposons que E est de dimension $n \in \mathbb{N}^*$.

- ✗ Un endomorphisme f est dit **trigonalisable** s'il existe une base $\mathcal{B}$ de E dans laquelle la matrice de f est triangulaire supérieure.
- ✗ Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **trigonalisable** s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ et une matrice triangulaire supérieure $T \in \mathcal{M}_n(\mathbb{K})$ telles que $A = PTP^{-1}$, en d'autres termes si A est semblable à une matrice triangulaire supérieure.

Théorème 8.19.

Un endomorphisme est trigonalisable si et seulement si l'une de ses matrices est trigonalisable.
Auquel cas, toutes ses matrices sont trigonalisables.

Théorème 8.20. **CNS de trigonalisabilité**

Un endomorphisme est trigonalisable si et seulement si son polynôme caractéristique est scindé.

Preuve. La condition nécessaire est immédiate. La condition suffisante est admise. □

Corollaire 8.21.

Toute matrice est trigonalisable dans $\mathcal{M}_n(\mathbb{C})$.

Aucune méthode pratique de trigonalisation n'est au programme. Mais on étudiera attentivement l'exemple suivant (concrètement il faut savoir trigonaliser une matrice de taille 3 sans aide).

Exercice 8.7.

Trigonaliser la matrice $B = \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$.

8.3.3 Somme et produit des valeurs propres

On déduit de ce qui précède que la trace (*resp.* le déterminant) d'une matrice est égal à la somme (*resp.* le produit) de ses valeurs propres.
On peut aussi préciser certains coefficients de son polynôme caractéristique.

Proposition. 8.22.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $\lambda_1, \ldots, \lambda_n$ ses valeurs propres (énumérées avec multiplicité). On a les résultats suivants :

i. $\mathrm{Tr}(A) = \displaystyle\sum_{k=1}^{n} \lambda_k$;

ii. $\det(A) = \displaystyle\prod_{k=1}^{n} \lambda_k$;

iii. $\chi_A(X) = X^n - \mathrm{Tr}(A)X^{n-1} + \ldots + (-1)^n \det(A)$.

Remarque 8.7.

Ce résultat reste vrai si $A \in \mathcal{M}_n(\mathbb{R})$, mais il faut garder à l'esprit que ce sont toutes les valeurs propres complexes de A qui sont prises en compte.

8.4 Premières applications de la réduction

8.4.1 Puissances d'une matrice

☞ Pour calculer les puissances d'une matrice diagonalisable A, on écrit celle-ci sous la forme

$$A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} P^{-1}$$

avec $\lambda_1, \ldots, \lambda_n$ ses valeurs propres (pas nécessairement deux-à-deux distinctes), et on utilise alors la relation

$$A^m = P \begin{pmatrix} \lambda_1^m & 0 & \cdots & 0 \\ 0 & \lambda_2^m & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^m \end{pmatrix} P^{-1}$$

valable pour tout $m \in \mathbb{N}$.

☞ Pour calculer les puissances d'une matrice trigonalisable A, on écrit celle-ci sous la forme

$$A = P \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & * \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} P^{-1}$$

avec $\lambda_1, \ldots, \lambda_n$ ses valeurs propres. On remarque alors que

$$T = \underbrace{\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}}_{D} + \underbrace{\begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}}_{N}$$

avec N qui est *nilpotente* (c'est-à-dire que ses puissances sont nulles à partir d'un certain rang).
Dans le cas où D et N commutent, on peut alors utiliser la formule du binôme de Newton pour en déduire les puissances de A.

Exercice 8.8.

Calculer les puissances de $B = \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$.

8.4.2 Suites récurrentes linéaires d'ordre p à cœfficients constants

⋆ **Cas** $p = 2$

On considère ici une suite à valeurs dans $\mathbb{K}$ vérifiant une relation de récurrence de la forme $\forall n \in \mathbb{N},\ u_{n+2} + bu_{n+1} + cu_n = 0$, où $(b, c) \in \mathbb{K} \times \mathbb{K}^*$, les premiers termes u_0, u_1 étant donnés. On cherche à exprimer u_n en fonction de n, u_0, u_1.

Soit $n \in \mathbb{N}$. On pose $U_n = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$, ce qui donne $U_{n+1} = \underbrace{\begin{pmatrix} -b & -c \\ 1 & 0 \end{pmatrix}}_{A} U_n$, d'où $U_n = A^n U_0$..

Le problème se ramène donc au calcul de A^n.

Proposition. 8.23.
Soit $(b,c) \in \mathbb{K} \times \mathbb{K}^*$ et F l'ensemble des suites $(u_n) \in \mathbb{K}^{\mathbb{N}}$ vérifiant la relation $\forall n \in \mathbb{N},\ u_{n+2} + bu_{n+1} + cu_n = 0$.

i. F est un espace vectoriel de dimension 2.

ii. Si l'équation caractéristique admet deux racines distinctes α, β, alors
$$F = \{(u_n) \in \mathbb{K}^{\mathbb{N}};\ \exists(\lambda,\mu) \in \mathbb{K}^2, \forall n \in \mathbb{N},\ u_n = \lambda\alpha^n + \mu\beta^n\}.$$

iii. Si l'équation caractéristique admet une unique racine réelle $\alpha \neq 0$, alors il existe des scalaires λ, μ tels que
$$F = \{(u_n) \in \mathbb{K}^{\mathbb{N}};\ \exists(\lambda,\mu) \in \mathbb{K}^2, \forall n \in \mathbb{N},\ u_n = (\lambda n + \mu)\alpha^n\}.$$

Remarque 8.8.
Si $\mathbb{K} = \mathbb{R}$ et si l'équation caractéristique admet deux racines complexes conjuguées non réelles $re^{\pm i\theta}$, alors on peut démontrer qu'il existe deux réels λ, μ tels que
$$F = \{(u_n) \in \mathbb{R}^{\mathbb{N}};\ \exists(\lambda,\mu) \in \mathbb{R}^2, \forall n \in \mathbb{N},\ u_n = r^n(\lambda\cos(n\theta) + \mu\sin(n\theta))\}.$$

⋆ Cas général

Soit $p \in \mathbb{N}$ et $a_0, \ldots, a_{p-1}$ des scalaires. On considère maintenant une suite à valeurs dans $\mathbb{K}$ vérifiant une relation de récurrence de la forme
$$\forall n \in \mathbb{N},\ u_{n+p} = a_{p-1}u_{n+p-1} + a_{p-2}u_{n+p-2} + \cdots + a_0u_n,$$
les premiers termes $u_0, u_1, \ldots, u_{p-1}$ étant donnés. On cherche à exprimer u_n en fonction de $n, u_0, u_1, \ldots, u_{p-1}$.

Soit $n \in \mathbb{N}$. On pose
$$U_n = \begin{pmatrix} u_{n+p-1} \\ u_{n+p-2} \\ \vdots \\ u_n \end{pmatrix} \text{ et } A = \begin{pmatrix} a_{p-1} & a_{p-2} & \cdots & \cdots & a_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$
On a ainsi
$$\begin{pmatrix} u_{n+p} \\ u_{n+p-1} \\ \vdots \\ u_{n+1} \end{pmatrix} = \begin{pmatrix} a_{p-1} & a_{p-2} & \cdots & \cdots & a_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} u_{n+p-1} \\ u_{n+p-2} \\ \vdots \\ u_n \end{pmatrix} \iff U_{n+1} = AU_n.$$
D'où $U_n = A^nU_0$ en raisonnant par récurrence. Le problème se ramène donc encore au calcul de A^n.
On ne présente ici que le cas où A admet p valeurs propres distinctes $\lambda_1, \ldots, \lambda_p$ (les autres cas sont hors-programmes).

Dans ce cas A est diagonalisable, et il existe donc une matrice de passage P tel que pour tout entier n:
$$U_n = P\begin{pmatrix} \lambda_1^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p^n \end{pmatrix} P^{-1}U_0$$
ce qui prouve que dans ce cas (u_n) est combinaison linéaire des p suites géométriques $(\lambda_1^n), (\lambda_2^n), \ldots, (\lambda_p^n)$. On peut notamment en déduire que l'ensemble des solutions est un espace vectoriel de dimension p (ce qui reste vrai même si A ne possède pas n valeurs propres distinctes).

Remarque 8.9.
On démontre aisément (voir le cours sur le déterminant) que $\chi_A(X) = X^p - a_{p-1}X^{p-1} - \cdots - a_1X - a_0$: on retrouve dans ce cas la notion d'équation caractéristique, dont les racines permettent d'exprimer le terme général de la suite considérée.

9 Probabilités générales

Dans ce chapitre nous allons reprendre la théorie des probabilités vue en première année, non seulement dans le cadre des univers finis mais aussi d'univers *dénombrables infinis.* La nécessité de cette extension provient des situations concrètes que l'on souhaite modéliser. Prenons par exemple le jeu suivant : on lance une pièce de monnaie jusqu'à obtenir Pile. Pour modéliser cette situation, on est obligé d'utiliser un univers infini, puisque le nombre de lancers que l'on devra effectuer n'est *a priori* pas borné.

9.1 Rappels : dénombrement dans les ensembles finis ou dénombrables

9.1.1 Cardinal d'un ensemble fini

Proposition. 9.1.
Deux ensembles finis ont le même cardinal si et seulement si ils sont équipotents, c'est à dire et seulement si il existe une bijection entre eux.

Proposition. 9.2.
Soient E, F deux ensembles finis, et $A \subset E$. On a les résultats suivants.

i. Si $E \cap F = \emptyset$, alors $\text{Card}\,(E \cup F) = \text{Card}(E) + \text{Card}(F)$.
ii. Si $\overline{A}$ est le complémentaire de A dans E, alors $\text{Card}(A) = \text{Card}(E) - \text{Card}(\overline{A})$
iii. Dans tous les cas on a $\text{Card}\,(E \cup F) = \text{Card}(E) + \text{Card}(F) - \text{Card}\,(E \cap F)$.

Proposition. 9.3.
Soient E, F des ensembles finis. Alors $\text{Card}\,(E \times F) = \text{Card}(E) \times \text{Card}(F)$.

Proposition. 9.4.
Soient E et F deux ensembles finis. Alors $\text{Card}\left(F^E\right) = \text{Card}(F)^{\text{Card}(E)}$.

Proposition. 9.5.
Soit E un ensemble fini. Alors l'ensemble $\mathcal{P}(E)$ est également fini, de cardinal $2^{\text{Card}(E)}$.

9.1.2 Dénombrement dans les ensembles finis

Listes ou tirages avec remise

Imaginons que chaque ensemble est une urne, et que ses éléments sont autant de boules dans celle-ci. Construire un ensemble revenant à prélever des éléments dans d'autres ensembles déjà existants, il y a une correspondance entre le dénombrement des ensembles mathématiques et le nombre d'issues associées aux tirages dans les urnes. Les résultats qui suivent seront démontrés en utilisant ce paradigme.

Soit E un ensemble fini. On appelle liste d'éléments de E tout p-uplet $(e_1, \dots, e_p)$ d'éléments (pas nécessairement distincts...) de E, où $p \in \mathbb{N}$.

Paradigme. Soit E un ensemble à n éléments. Il y a autant de manières de construire une liste de p éléments de E que de prélever **avec remise** p .

Proposition. 9.6. **Nombre de tirages avec remise**

Il y a n^p manières de tirer avec remise p boules dans une urne en contenant n.

Preuve. On a n choix au premier tirage, puis n choix au second (car on a remis la première boule dans l'urne), puis n au troisième,... et enfin n choix au p-ième. Il y a donc n^p configurations possibles au total (la **Proposition 9.3** assure que ces nombres de possibilités *se multiplient*). □

Exemple 9.1.

Redémontrons ainsi que le nombre d'applications $f : [\![1,p]\!] \to [\![1,n]\!]$ est égal à n^p.
Construire une application $f : [\![1,p]\!] \to [\![1,n]\!]$ c'est choisir les valeurs de $f(1),\dots,f(p)$ dans l'ensemble $[\![1,n]\!]$. Il y a n choix pour $f(1)$, n choix pour $f(2)$,... et toujours n choix pour $f(p)$. Ce qui donne au final n^p choix.

Arrangements ou tirages sans remise

Définition 9.1. **Arrangement**

Soit E un ensemble fini, et $1 \leq p \leq \text{Card}(E)$ un entier. On appelle **arrangement** de p éléments de E tout p-uplet $(e_1,\dots,e_p)$ d'éléments **distincts** de E.

☞ Deux éléments d'un arrangement sont toujours distincts: c'est ce qui les différencie des listes.

Paradigme. Soit E un ensemble à n éléments. Il y a autant de manières de construire un arrangement de $p \leq n$ éléments de E que de prélever (les unes après les autres et) **sans remise** p boules dans une urne en contenant n.

Définition 9.2.

Soient $n \geq p \geq 1$ des entiers. On note A_n^p le nombre d'arrangements de p éléments d'un ensemble à n éléments, c'est-à-dire le nombre de manières de tirer **sans remise** p boules dans une urne en contenant n.

Proposition. 9.7. **Nombre de tirages sans remise**

Soient $n \geq p \geq 1$. On a $A_n^p = \dfrac{n!}{(n-p)!}$.

Preuve. Soit une urne contenant n boules; dénombrons le nombre de manières d'y choisir p boules sans remise. On a n choix au premier tirage, puis $n-1$ choix au second, puis $n-2$ choix au troisième,... et $n-(p-1)$ choix au p-ième. Il y a donc $n(n-1)\cdots(n-p+1) = \dfrac{n!}{(n-p)!}$ configurations possibles au total. □

Exemple 9.2. **Injections de $[\![1,p]\!]$ dans $[\![1,n]\!]$**

Soient $p \leq n$ des entiers. Le nombre d'injections $f : [\![1,p]\!] \to [\![1,n]\!]$ est égal à A_n^p.

Construire une application $f : [\![1,p]\!] \to [\![1,n]\!]$ c'est choisir les valeurs de $f(1),\dots,f(p)$ dans l'ensemble $[\![1,n]\!]$. Pour que f soit injective, il faut et il suffit que ces valeurs soient deux à deux distinctes. Cela prouve que construire une injection $f : [\![1,p]\!] \to [\![1,n]\!]$ équivaut à choisir un arrangement de p éléments sur les n (le premier élément choisi sera $f(1)$, le second sera $f(2)$,...), d'où le résultat.

Proposition. 9.8.

Soient E et F deux ensembles finis tels que $\text{Card}E = p$ et $\text{Card}F = n$ avec $p \leq n$. Alors le nombre d'injections $f : E \to F$ est égal à A_n^p.

Application des arrangements aux permutations d'un ensemble

Définition 9.3. **Permutations d'un ensemble fini**

Soit E un ensemble fini. On appelle **permutation** de E toute bijection de E dans E, et on note $\mathfrak{S}(E)$ l'ensemble des permutations de E.
Soit $n \in \mathbb{N}^*$. On note $\mathfrak{S}_n$ l'ensemble des permutations de $[\![1,n]\!]$.

Proposition. 9.9.

Soit E un ensemble fini de cardinal n. Alors $\text{Card}\left(\mathfrak{S}(E)\right) = n!$.

Preuve. On sait que si $f : E \to E$, alors f est bijective si et seulement si elle est injective. Or il y a A_n^n injections de E dans E en vertu de la **Proposition 9.2**, ce qui donne le résultat. □

Corollaire 9.10.
Soient E, F deux ensembles finis tels que $\mathrm{Card}E = \mathrm{Card}F = n$. Alors le nombre de bijections $f : E \to F$ est égal à $n!$.

Proposition. 9.11.
Il existe $n!$ manières d'ordonner les éléments d'un ensemble de cardinal n.

Preuve. Ordonner les éléments de E revient à attribuer à chacun d'eux un unique numéro, compris entre 1 et n. Le nombre recherché est donc le nombre de bijections entre E et $[\![1, n]\!]$, soit $n!$. □

Exemple 9.3.
Le nombre de mots que l'on peut écrire avec les lettres P, I, Z, Z, A n'est pas égal à $5!$. En effet, si l'on prend un mot écrit avec ces 5 lettres et qu'on permute les deux Z, on obtient encore le même mot: il n'y a donc que $\frac{5!}{2!}$ manières d'arranger ces lettres entre elles.

Exercice 9.1.
Combien y a-t-il d'anagrammes du mot $AVADAKEDAVRA$?

Combinaisons ou tirages simultanés

Définition 9.4. **Combinaison**
Soit E un ensemble fini, et $1 \leq p \leq \mathrm{Card}(E)$ un entier. On appelle **combinaison** de p éléments parmi E tout sous-ensemble de E de cardinal p (c'est-à-dire tout ensemble de p éléments issus de E).

Exemple 9.4.
Si $E = \{a, b, c\}$, alors $\{a, b\}$ est une combinaison de deux éléments parmi E, et (a, b) est arrangement de deux éléments de E.

Remarque 9.1.
Il existe dans les uplets (arrangements et listes) une notion d'ordre, absente des ensembles:

- ✗ lorsque l'on représente un ensemble, on ne se préoccupe pas de l'ordre dans lequel on énumère ses éléments. Par exemple, les deux ensembles $\{a, b\}$ et $\{b, a\}$ sont égaux;
- ✗ par contre, on considère que les deux uplets (a, b) et (b, a) sont distincts: l'ordre dans lequel apparaissent a et b importe.

La différence entre une combinaison et un arrangement (ou une liste) est donc que les éléments d'un arrangement (ou d'une liste) sont ordonnés, contrairement à ceux d'une combinaison. Ils sont en outre deux à deux distincts, contrairement à ceux d'une liste.

Paradigme. Soit E un ensemble à n éléments, et $p \leq n$. Il y a autant de manières de construire une partie à p éléments de E que de prélever simultanément (on parle de "poignée") p boules dans une urne en contenant n.

Proposition. 9.12.
Soient $1 \leq p \leq n$ des entiers. Le nombre de parties à p éléments d'un ensemble à n éléments est $\displaystyle\binom{n}{p} = \frac{n!}{p!(n-p)!}$.

Preuve. Soit U une urne contenant n boules. Commençons par choisir p boules sans remise: il y a

$$A_n^p = n(n-1)\cdots(n-p+1) = \frac{n!}{(n-p)!}$$

choix possibles.
Mais tous ces tirages ne correspondent pas à des ensembles de boules différents. Il reste donc à dénombrer, étant donné un ensemble à p éléments, combien de tirages différents lui correspondent : ce nombre est égal à $p!$, puisqu'il s'agit du nombre

de manières d'ordonner p éléments (c'est-à-dire de choisir leur ordre d'apparition). Finalement, le nombre recherché est donc égal à $\dfrac{n!}{(n-p)!}$ divisé par $p!$, d'où le résultat. $\square$

Exemple 9.5.

Il y a $\dbinom{n}{p}$ applications strictement croissantes de $[\![1,p]\!]$ dans $[\![1,n]\!]$.

En effet, on commence par choisir l'image de f, c'est-à-dire un ensemble I de cardinal p dans $[\![1,n]\!]$: il y a pour cela $\dbinom{n}{p}$ possibilités.

Il suffit alors d'observer qu'il n'y a rien d'autre à choisir: puisque f est croissante, le plus petit élément de I sera nécessairement $f(1)$, le second plus petit sera nécessairement $f(2)$,..., et le plus grand sera nécessairement $f(p)$. D'où le résultat.

Exercice 9.2.

Un maître d'école distribue des bons points à ses 24 élèves, sous la forme d'images plastifiées.

1. On suppose que le maître d'école distribue 12 images deux à deux différentes. Combien y a-t-il de distributions possibles
   a. s'il donne au plus une image à chaque élève?
   b. si chaque élève peut recevoir un nombre quelconque d'images?
2. Reprendre la Question **1.a.** dans le cas où les images sont toutes identiques.

9.1.3 Ensembles dénombrables

Définition 9.5. **Ensemble (fini ou dénombrable)**

- ✗ Un ensemble E est dit **dénombrable** s'il est équipotent à $\mathbb{N}$, c'est-à-dire s'il existe une bijection entre E et $\mathbb{N}$.
- ✗ Un ensemble E est dit **fini ou dénombrable** s'il est équipotent à une partie de $\mathbb{N}$, c'est-à-dire s'il existe une bijection entre E et une partie de $\mathbb{N}$.

Exemple 9.6.

$\{e,\pi\}$, $\mathbb{N}^*$, $\mathbb{Z}$ et $\left\{\dfrac{1}{n} : n \in \mathbb{N}^*\right\}$ sont finis ou dénombrables.

Proposition. 9.13.

Un ensemble E non vide est fini ou dénombrable si et seulement si on peut l'écrire sous la forme $E=\{x_i;\ i\in I\}$, où $I\subset\mathbb{N}$ et où les x_i sont distincts.

☞ On peut démontrer que si E est fini ou dénombrable, alors:

- ✗ soit il est de cardinal fini, et alors il est équipotent à une partie finie de $\mathbb{N}$;
- ✗ soit il est de cardinal infini, et alors il est équipotent à $\mathbb{N}$.

Exemple 9.7.

Tous les ensembles ne sont pas dénombrables : $\mathbb{R}$ n'est pas dénombrable et $[0,1]$ non plus.

Proposition. 9.14.

Le produit cartésien de deux ensembles finis ou dénombrables est fini ou dénombrable.

Exemple 9.8.

$\mathbb{N}^p$ est dénombrable pour tout $p\in\mathbb{N}^*$.

On peut démontrer que l'union d'un nombre dénombrable d'ensembles dénombrables reste dénombrable.

Proposition. 9.15.
$\mathbb{Q}$ est dénombrable.

Méthodologie du dénombrement

On conclue ce chapitre par un "point méthode" pour savoir comment dénombrer les possibilités offertes par une situation donnée, à l'aide des outils précédemment introduits.

✎ **On décompose la situation:**

- ✗ Si la situation à étudier peut se décrire en plusieurs cas **disjoints** (à l'aide par exemple d'un "ou bien" ou "soit"), alors le dénombrement peut s'obtenir en **ajoutant** les dénombrement de chacun des cas.
- ✗ Si la situation à étudier peut être décrite par plusieurs **étapes successives** ("puis", "et", ...) alors le dénombrement peut s'obtenir en **multipliant** les dénombrements de chacune des étapes.
- ✗ Il peut être parfois beaucoup plus simple de dénombrer l'ensemble complémentaire.

✎ **On dénombre chaque composante de la situation décomposée:**

- ✗ Si l'ordre intervient et si les répétitions sont autorisées, il y a n^d façons de choisir d fois successives un élément parmi n.
- ✗ Si l'ordre intervient mais que les répétitions ne sont pas autorisées, il y a A_n^d façons de ranger d éléments parmi n.
- ✗ Si l'ordre n'intervient pas et que le répétitions ne sont pas autorisées, il y a $\binom{n}{d}$ façons de choisir d éléments parmi n.

Exercice 9.3.
On tire successivement, avec remise, 5 boules dans une urne qui en contient 3 noires et 4 blanches. Combien y a-t-il de tirages:
1. en tout?
2. comportant 3 noires et 2 blanches?
3. comportant au plus une noire?
4. comportant au moins deux noires?

9.2 Espaces probabilisables et probabilisés

9.2.1 Univers, tribus, évènements

Définition 9.6.
Dans une situation réelle donnée (aussi appelée **épreuve** ou **expérience aléatoire**) on appelle **univers** et l'on note Ω un ensemble fini ou dénombrable dont les éléments représentent toutes les évolutions possibles de l'épreuve considérée. Les éléments $\omega \in \Omega$ sont appelés des **issues**.

L'énoncé d'un exercice de probabilité (c'est à dire d'une expérience aléatoire) ne précise pas (toujours) l'univers Ω explicitement. C'est donc parfois à celui ou celle qui en fait l'étude de *décrire* cet ensemble d'une manière rigoureuse et commode afin de travailler dessus. Cette étape s'appelle la **modélisation** de l'expérience.

Exemple 9.9.
Regardons les univers avec lesquels on peut modéliser les expériences suivantes.

- ✗ **Expérience 1.** On lance un dé cubique à 6 faces et on lit le numéro inscrit sur la face supérieure. Les résultats possibles sont alors $\Omega = \{1, 2, 3, 4, 5, 6\}$ qui est un ensemble de cardinal 6.
- ✗ **Expérience 2.** On lance deux dés discernables (un rouge et un bleu par exemple) et on note les résultats obtenus. On voit alors qu'ici $\Omega = \{(i, j) :\ 1 \leq i \leq 6,\ 1 \leq j \leq 6\} = [\![1; 6]\!]^2$, qui est un ensemble à 36 éléments.

✗ **Expérience 3.** Cette fois les deux dés sont identiques, mais on note encore les résultats obtenus. Comme on ne peut plus faire la différence entre les deux dés, les deux résultats de l'expérience précédent $(1;3)$ et $(3;1)$ sont en particulier dans cette nouvelle expérience l'expression de la même issue $\{1;3\}$. Ainsi, on a

$$\Omega = \{\{i,j\} : \ 1 \leq i \leq j \leq 6\}.$$

De plus,

$$\mathrm{Card}(\Omega) = \sum_{1\leq i\leq j\leq 6} 1 = \sum_{j=1}^{6}\sum_{i=1}^{j} 1 = \sum_{j=1}^{6} j = \frac{6\times 7}{2} = 21.$$

☞ En pratique, on ne détaillera pas l'univers Ω dans la plupart des exercices: la modélisation sera implicite.

La définition suivante étend les notions de réunion et d'intersection à un nombre infini dénombrable d'ensembles.

Définition 9.7. **Intersection et union dénombrable**

Soient Ω un univers et $(A_i)_{i\in I}$ une famille finie ou dénombrable de parties de Ω.

La réunion $\bigcup\limits_{i\in I} A_i$ est définie par

$$\omega \in \bigcup_{i\in I} A_i \iff \exists i \in I, \omega \in A_i$$

De même, l'intersection $\bigcap\limits_{i\in I} A_i$ est définie par

$$\omega \in \bigcap_{i\in I} A_i \iff \forall i \in I, \omega \in A_i.$$

Dans un univers infini non dénombrable, toutes les parties ne peuvent pas être considérées comme des évènements (il existe des parties trop *compliquées* pour que, dans certains cas, leur probabilité soit définie...). On a ainsi besoin de définir la notion de **tribu** qui sera *a posteriori* l'ensemble des évènements.

Définition 9.8. **Notion de tribu, espace probabilisable**

Soit Ω un univers.

✗ On appelle **tribu** (ou σ-algèbre) tout sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$ vérifiant les propriétés suivantes:

i. $\Omega \in \mathcal{T}$;
ii. si $A \in \mathcal{T}$, alors $\overline{A} \in \mathcal{T}$;
iii. pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{T}$, on a $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{T}$.

✗ Si $\mathcal{T}$ est une tribu sur Ω, on dit que $(\Omega, \mathcal{T})$ est un espace probabilisable.

Exemple 9.10.

Si Ω est un univers, $\{\emptyset, \Omega\}$ et $\mathcal{P}(\Omega)$ sont des tribus.

☞ De même que l'univers, la tribu utilisée n'est en général pas précisée dans l'énoncé.

Définition 9.9. **Évènements**

Soit $(\Omega, \mathcal{T})$ un espace probabilisable.

✗ On appelle **évènement** de $(\Omega, \mathcal{T})$ toute partie A de Ω telle que $A \in \mathcal{T}$.
✗ On appelle **évènement**[a] **élémentaire** de $(\Omega, \mathcal{T})$ toute partie de Ω de la forme $\{\omega_i\}$ où $\omega_i \in \Omega$.

[a] sous réserve qu'une telle partie soit un évènement... cela dépend de la tribu considérée.

Proposition. 9.16.
Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On a les résultats suivants.

i. $\emptyset \in \mathcal{T}$;
ii. si $A, B \in \mathcal{T}$ alors $A \cup B,\ A \cap B$ et $A \setminus B \in \mathcal{T}$;
iii. pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{T}$, on a $\displaystyle\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{T}$.

On va donc décrire les *évènements* qui nous intéressent à l'aide des opérations de la théorie des ensembles.

Exemple 9.11.

Le concierge d'un immeuble possède à son trousseau un exemplaire de chacune des clés des appartements de l'immeuble, y compris celle de son domicile. En rentrant d'une soirée arrosée, il ne sait discerner laquelle ouvre sa porte d'entrée. N'ayant pas tous ses moyens, il remet dans le trousseau la clé après l'avoir testée, même si celle-ci n'est clairement pas la bonne. On note S l'évènement correspondant à la découverte de la bonne clé et à l'ouverture de la porte.

S'il est facile de comprendre l'évènement S et de le formuler "oralement", son expression permettant le calcul de la probabilité correspondante n'est pas si simple. Pour décrire précisément la structure de l'évènement, on a besoin d'introduire les évènements (pour $n \geq 1$ entier) A_n : "Le concierge réussit enfin à ouvrir la porte après n tentatives" et B_n : "la $n-$ième clé testée est la bonne". On peut alors écrire les choses comme suit

$$S = \bigcup_{n\in\mathbb{N}^\star} A_n = \bigcup_{n\in\mathbb{N}^\star} \left(\left(\bigcap_{k=1}^{n-1} \overline{B}_k \right) \cap B_n \right).$$

Définition 9.10. **Le langage des probabilités**
Le tableau suivant indique comment traduire en théorie des probabilités le vocabulaire élémentaire de la théorie des ensembles.

Vocabulaire des ensembles	Vocabulaire des probabilités	Notation
la tribu	l'ensemble des évènements	$\mathcal{T}$
la partie pleine de Ω	l'évènement certain	Ω
l'ensemble vide	l'évènement impossible	$\emptyset$
le complémentaire de la partie A dans Ω	l'évènement contraire de A	$\overline{A}$
l'intersection des parties A et B	la conjonction de A et B	$A \cap B$
la réunion des parties A et B	la disjonction de A et B	$A \cup B$
les parties A et B sont disjointes	A et B sont incompatibles	$A \cap B = \emptyset$
la partie A est incluse dans la partie B	la réalisation de A implique celle de B	$A \subset B$

Définition 9.11. **Système complet d'évènements**
Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $(A_i)_{i\in I}$ une famille d'évènements. On dit que $(A_i)_{i\in I}$ est un **système complet d'évènements** (s.c.e) pour Ω si :

i. $\Omega = \displaystyle\bigcup_{i\in I} A_i$;
ii) $\forall (i,j) \in I^2,\ i \neq j \Longrightarrow A_i \cap A_j = \emptyset$.

☞ Étant donnée une expérience, une et une seule partie d'un système complet se réalise.

Exemple 9.12.

✗ Si A est un évènement, alors le système $A, \overline{A}$ est complet.

- ✗ Si $\Omega = \{\omega_i;\ i \in I\}$, et si les ω_i sont deux à deux distincts, alors le système $(\{\omega_i\})_{i\in I}$ constitué de tous les évènements élémentaires est complet.

9.2.2 Probabilité sur un espace probabilisable

Définition 9.12. **Probabilité et espace probabilisé**

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On appelle **mesure de probabilité** sur $(\Omega, \mathcal{T})$ ou plus simplement **probabilité** sur $(\Omega, \mathcal{T})$ toute application $\mathbf{P} : \mathcal{T} \to [0, 1]$ vérifiant les propriétés suivantes.

i. **Probabilité de l'évènement certain.** $\mathbf{P}(\Omega) = 1$.

ii. **Axiome de σ-additivité.** Quelle que soit la suite $(A_n)_{n\in\mathbb{N}}$ d'évènements (deux à deux) incompatibles, la série $\sum_{n\geq 0} \mathbf{P}(A_n)$ converge et sa somme vaut $\sum_{n=0}^{+\infty} \mathbf{P}(A_n) = \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right)$.

Le triplet $(\Omega, \mathcal{T}, \mathbf{P})$ est alors appelé un **espace probabilisé.**

Jusqu'à la fin de ce chapitre, on fixe un espace probabilisé $(\Omega, \mathcal{T}, \mathbf{P})$.

Remarque 9.2.

Seuls les évènements possèdent une probabilité, pas les parties quelconques de Ω. La probabilité *mesure* les parties qui sont *mesurables*.

Exercice 9.4.

On considère une urne contenant $p \geq 1$ boules noires et une boule blanche. On effectue une infinité de tirages avec remise dans celle-ci.
Pour $n \geq 1$, on dit que l'évènement X_n est réalisé si et seulement si on obtient pour la première fois la boule blanche au n-ième tirage.
Donner la probabilité de X_n, puis la probabilité d'obtenir la boule blanche après un nombre fini de tirage.

Définition 9.13. **Evènement négligeable ou quasi-certain**

- ✗ Un évènement est dit **négligeable** si sa probabilité vaut 0.
- ✗ Un évènement est dit **presque sûr** ou **quasi-certain** si sa probabilité vaut 1.

Remarque 9.3.

Attention !
Il ne faut pas confondre l'*évènement impossible* (qui correspond à l'ensemble $\emptyset$) avec un évènement négligeable, et l'*évènement certain* (qui est Ω) avec un évènement presque certain.

Par exemple, dans l'exercice précédent:

- ✗ l'évènement *obtenir la boule blanche en temps fini* est presque certain mais pas certain;
- ✗ l'évènement *ne jamais obtenir la boule blanche* est négligeable mais pas impossible.

Cas d'un univers fini

Si Ω est un univers fini, c'est toujours la tribu $\mathbf{P}(\Omega)$ qui est utilisée. Ainsi toutes les parties de Ω sont des événements, et possèdent donc une probabilité.

Construire une probabilité sur un univers fini $\Omega = \{\omega_1, \ldots, \omega_n\}$ revient à associer une probabilité p_i à chaque issue ω_i.
Le réel $p_i \in [0, 1]$ sera la probabilité que l'issue ω_i se réalise, et comme Ω décrit l'ensemble des possibilités on a $\sum_{i=1}^{n} p_i = 1$.

On peut alors définir la probabilité d'un évènement A comme étant la somme des probabilités des issues le constituant, ce qui revient à poser

$$\mathbf{P}(A) = \sum_{\{i \in [\![1,n]\!];\ \omega_i \in A\}} p_i.$$

Le cas particulier suivant est le plus fréquent.

Théorème & Définition 9.1. **Probabilité uniforme sur un univers fini.**

On appelle **probabilité uniforme** sur un univers fini $\Omega = \{\omega_1, \ldots, \omega_n\}$ la probabilité $\mathbf{P}$ définie par :

$$\forall i \in [\![1,n]\!],\ \mathbf{P}(\{\omega_i\}) = \frac{1}{n}.$$

La probabilité d'un évènement est alors donnée par le rapport

$$\frac{\text{nombre de cas favorables}}{\text{nombre total de cas}}.$$

En d'autres termes, on a $\mathbf{P}(A) = \dfrac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$ pour tout évènement A.

Remarque 9.4.

Par abus de langage, on emploie parfois la locution *au hasard* pour indiquer que l'univers est muni de la probabilité uniforme.

☞ En général, le lancer d'un dé, d'une pièce de monnaie, ou le choix d'une boule dans une urne sont effectués de manière équiprobable.

Exercice 9.5.

On lance trois fois de suite un dé à six faces. Calculer la probabilité d'obtenir trois fois le chiffre 1.

Remarque 9.5.

Il n'existe pas de probabilité uniforme sur un univers infini dénombrable
Supposons en effet avoir un univers infini dénombrable $\Omega = \{\omega_n;\ n \in \mathbb{N}\}$ muni d'une probabilité uniforme P. À chacun des évènements élémentaires $\{\omega_n\}$ est alors associée la même probabilité $p \in [0,1]$, ce qui permet d'écrire

$$\mathbf{P}(\Omega) = \sum_{n=0}^{+\infty} \mathbf{P}(\{\omega_n\}) = \sum_{n=0}^{+\infty} p = p \sum_{n=0}^{+\infty} 1 = \begin{cases} +\infty & \text{si } p \neq 0 \\ 0 & \text{sinon} \end{cases}.$$

Dans les deux cas cela contredit la propriété $\mathbf{P}(\Omega) = 1$, et termine ainsi la démonstration.

Exercice 9.6.

On tire 4 cartes dans un jeu en contenant 32. Quelle est la probabilité d'avoir exactement un as? Au plus un as? Au moins un as?

9.2.3 Probabilités et opérations ensemblistes

Étant donné une famille dénombrable d'évènements $(A_i)_{i \in I}$, le symbole $\displaystyle\sum_{i \in I} \mathbf{P}(A_i)$ désignera la somme de la série de terme général $(\mathbf{P}(A_i))_{i \in I}$, sous réserve de convergence.

Proposition. 9.17. **σ-additivité finie ou dénombrable**

Soit $(A_i)_{i \in I}$ une famille d'évènements incompatibles. On a $\displaystyle\sum_{i \in I} \mathbf{P}(A_i) = \mathbf{P}\left(\bigcup_{i \in I} A_i\right)$.

Proposition. 9.18.

Soit $(A_i)_{i \in I}$ un système complet. On a alors $\displaystyle\sum_{i \in I} \mathbf{P}(A_i) = 1$.

Remarque 9.6. **Système quasi-complet d'évènements**

On dit qu'une famille $(A_i)_{i\in I}$ d'évènements est un système quasi-complet pour Ω si :

i. $\mathbf{P}\left(\bigcup_{i\in I} A_i\right) = 1$;

ii. $\forall (i,j) \in I^2,\ i \neq j \Longrightarrow A_i \cap A_j = \emptyset$.

Tous les théorèmes vus dans la suite (en particulier la formule des probabilités totales) utilisant un système complet sont encore vrais avec un système quasi-complet.

On peut redémontrer tous les résultats vus dans le cours de première année, puisque ceux-ci se prouvaient à partir de la formule de σ-additivité finie et de la relation $\mathbf{P}(\Omega) = 1$. On y renvoie pour les preuves.

Proposition. 9.19.

Soient A, B deux évènements. On a alors les résultats suivants.

i. $\mathbf{P}\left(\overline{A}\right) = 1 - \mathbf{P}(A)$.

ii. $\mathbf{P}(\emptyset) = 0$.

iii. Si $A \subset B$, alors $\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A)$.

iv. Si $A \subset B$, alors $\mathbf{P}(A) \leq \mathbf{P}(B)$ (toute probabilité est une fonction croissante).

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

Enfin, la formule des probabilités totales reste vraie dans le cas où le système complet considéré est infini dénombrable.

Théorème 9.20. **Formule des probabilités totales, première version**

Soit B un évènement et $(A_i)_{i\in I}$ un système complet. Alors $\sum_{i\in I} \mathbf{P}(B \cap A_i)$ converge, et on a

$$\mathbf{P}(B) = \sum_{i\in I} \mathbf{P}(B \cap A_i).$$

Le théorème suivant permet enfin de calculer la probabilité d'une réunion dénombrable d'évènements pas nécessairement incompatibles, dans deux cas particuliers.

Définition 9.14. **Monotonie au sens de l'inclusion**

Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'évènements.

✗ $(A_n)_{n\in\mathbb{N}}$ est dite **croissante** (au sens de l'inclusion) si $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$.

✗ $(A_n)_{n\in\mathbb{N}}$ est dite **décroissante** (au sens de l'inclusion) si $A_n \supset A_{n+1}$ pour tout $n \in \mathbb{N}$.

Théorème 9.21.

Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'évènements. On a alors les résultats suivants.

i. **Continuité croissante.** Si $(A_n)_{n\in\mathbb{N}}$ est croissante, alors la suite $(\mathbf{P}(A_n))_{n\in\mathbb{N}}$ converge et

$$\lim_{n\to+\infty} \mathbf{P}(A_n) = \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right).$$

ii. **Continuité décroissante.** Si $(A_n)_{n\in\mathbb{N}}$ est décroissante, alors la suite $(\mathbf{P}(A_n))_{n\in\mathbb{N}}$ converge et

$$\lim_{n\to+\infty} \mathbf{P}(A_n) = \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

iii. **Sous-additivité.** Si la série $\sum_{n\geq 0} \mathbf{P}(A_n)$ converge, alors on a $\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbf{P}(A_n)$.

Corollaire 9.22.
Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'évènements. On a alors les résultats suivants.

i. $\displaystyle\lim_{N\to+\infty}\mathbf{P}\left(\bigcup_{n=0}^{N}A_n\right)=\mathbf{P}\left(\bigcup_{n=0}^{+\infty}A_n\right)$.

ii. $\displaystyle\lim_{N\to+\infty}\mathbf{P}\left(\bigcap_{n=0}^{N}A_n\right)=\mathbf{P}\left(\bigcap_{n=0}^{+\infty}A_n\right)$.

Exercice 9.7.
On joue une infinité de fois au loto. Montrer que presque sûrement, on finit par gagner.

9.3 Indépendance et conditionnement

Définition 9.15. **Indépendance**
Soient A, B deux évènements. On dit que A et B sont **indépendants** si

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

L'indépendance de deux évènements correspond le plus souvent à un choix effectué lors de la modélisation, et n'est alors pas démontrable.
Par exemple, les évènements élémentaires correspondants aux lancers successifs d'une même pièce de monnaie...
Cela correspond parfois à la décision de négliger certains paramètres, afin de simplifier le modèle. Mais dans certains cas, l'indépendance est une conséquence de la modélisation, et c'est le calcul qui permet de la vérifier.

Remarque 9.7.

Attention !
L'indépendance et l'incompatibilité sont deux notions différentes et on fera bien attention à le pas les confondre. Par exemple, dans le cas où $\mathbf{P}(A)\in]0,1[$, les évènements A et Ω sont indépendants mais pas incompatibles, alors que les évènements A et $\overline{A}$ sont incompatibles mais certainement pas indépendants.

Proposition. 9.23.
On a alors les résultats suivants.

i. Tout évènement est indépendant de l'évènement impossible $\emptyset$ et de l'évènement certain Ω.
ii. Deux évènements A et B sont indépendants si et seulement si A et $\overline{B}$ le sont.
iii. Soient B un évènement, et $A_1,\dots,A_n$ des évènements deux à deux incompatibles, tels que chacun d'entre eux soit indépendant de B. Alors B est indépendant de $A_1\cup\cdots\cup A_n$.

Il nous reste encore à définir la notion d'indépendance pour une famille finie d'évènements $A_1,...,A_m$. Il faut déjà comprendre que l'indépendance deux à deux de ceux-ci ne correspond pas à l'indépendance globale de la famille: on peut imaginer que A_1 ne soit conséquence d'aucun des autres A_i, mais que pourtant la conjonction des $(A_i)_{i>1}$ implique A_1.

Définition 9.16. **Indépendance mutuelle**
Soit $(A_i)_{1\leq i\leq m}$ une famille finie d'évènements.

i. $A_1,\dots,A_m$ sont dits **deux à deux indépendants** si, quels que soient $i,j\in[\![1,m]\!]$:

$$i\neq j\Longrightarrow\mathbf{P}\left(A_i\cap A_j\right)=\mathbf{P}(A_i)\times\mathbf{P}(A_j).$$

ii. $A_1,\dots,A_m$ sont dits **mutuellement indépendants** si, quel que soit $k\in[\![1,m]\!]$ et les entiers distincts $i_1,\dots,i_k$ appartenant à $[\![1,m]\!]$:

$$\mathbf{P}\left(A_{i_1}\cap\cdots\cap A_{i_k}\right)=\mathbf{P}(A_{i_1})\times\cdots\times\mathbf{P}(A_{i_k}).$$

Remarque 9.8.
On obtient par exemple que trois évènements A, B, C sont mutuellement indépendants si: $\mathbf{P}(A\cap B)=\mathbf{P}(A)\mathbf{P}(B)$, $\mathbf{P}(A\cap C)=\mathbf{P}(A)\mathbf{P}(C)$, $\mathbf{P}(B\cap C)=\mathbf{P}(B)\mathbf{P}(C)$ **et** $\mathbf{P}(A\cap B\cap C)=\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$.

L'exercice ci-dessous montre que l'indépendance mutuelle est une notion plus forte que l'indépendance deux à deux.

Exercice 9.8.

On lance deux fois de suite une pièce équilibrée. On considère les évènements A "le premier lancer donne *Pile*", B "le second lancer donne *Pile* et C "les deux lancers donnent le même résultat".
Montrer que ces trois évènements sont deux à deux indépendants, mais pas mutuellement indépendants.

Théorème 9.24.

Soit une famille finie d'évènements $(A_i)_{i\in I}$. On définit une famille $(B_i)_{i\in I}$ en posant $B_i = A_i$ ou $B_i = \overline{A_i}$ pour tout $i \in I$. Alors $(A_i)_{i\in I}$ est mutuellement indépendante si et seulement si $(B_i)_{i\in I}$ l'est.

☞ Par abus de langage, on dira parfois des évènements d'une famille qu'ils sont *indépendants* au lieu de mutuellement indépendants. Mais dans le cas de l'indépendance deux-à-deux, il faudra toujours préciser pour éviter les confusions.

9.3.1 Probabilités conditionnelles

Définition 9.17. **Probabilité conditionnelle**

Soit A un évènement non négligeable. Pour tout évènement B, on appelle **probabilité conditionnelle de B sachant A** le réel noté $\mathbf{P}\left(B|A\right)$ ou $\mathbf{P}_A(B)$ défini par

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

Proposition. 9.25.

Soit $(\Omega, \mathcal{T}, \mathbf{P})$ un espace probabilisé, et A un évènement non négligeable. L'application $\mathbf{P}_A : \mathcal{T} \to [0,1], B \mapsto \mathbf{P}_A(B)$ est une probabilité sur $(\Omega, \mathcal{T})$.

Remarque 9.9.

Attention !
Écrire (ou même penser) que $B|A$ (ou parler de l'évènement B sachant A) est un évènement n'a aucun sens. L'écriture $\mathbf{P}(B|A)$ est une notation, notamment pratique lorsque l'évènement A est lourd à écrire en indice de $\mathbf{P}$ ce qui rendrait peu lisible la quantité considérée.

En pratique, on est souvent amené à considérer des *processus aléatoires*, c'est-à-dire des successions complexes d'étapes aléatoires plus simples (ces processus sont alors représentés par des arbres). Pour modéliser globalement un tel phénomène, on peut utiliser la notion de probabilité conditionnelle afin d'étudier l'interaction entre deux expériences successives.

Proposition. 9.26.

Soient A, B deux évènements non négligeables. Alors $\mathbf{P}(A\cap B) = \mathbf{P}_B(A)\mathbf{P}(B) = \mathbf{P}_A(B)\mathbf{P}(A)$.

Preuve. Résulte trivialement de la définition de la probabilité conditionnelle. □

Proposition. 9.27.

Soient A, B deux évènements non négligeables. Alors les assertions suivantes sont équivalentes.

i. A et B sont indépendants.
ii. $\mathbf{P}_B(A) = \mathbf{P}(A)$.
iii. $\mathbf{P}_A(B) = \mathbf{P}(B)$.

9.3.2 Formule des probabilités totales

On peut donner la seconde expression de la formule des probabilités totales. Celle-ci est fondamentale.

Théorème 9.28. **Formule des probabilités totales, seconde version**

Soit $(A_i)_{i\in I}$ un système complet d'évènements non négligeables, et B un évènement. Alors $\sum_{i\in I} \mathbf{P}_{A_i}(B)\mathbf{P}(A_i)$ converge, et on a

$$\mathbf{P}(B) = \sum_{i\in I} \mathbf{P}_{A_i}(B)\mathbf{P}(A_i).$$

Preuve. Se déduit directement de la première version de la formule des probabilités totales, couplée avec la définition d'une probabilité conditionnelle. □

Remarque 9.10.

Dans le cas d'un système complet quelconque $(A_i)_{i\in I}$, on pose $\mathbf{P}(B|A_i)\mathbf{P}(A_i) = 0$ dès que $\mathbf{P}(A_i) = 0$, pour tout évènement B. On vérifie alors aisément que la formule ci-dessus reste valable avec cette convention, ce qui en pratique permet de l'appliquer sans avoir à vérifier l'hypothèse que les évènements du système complet sont non négligeables.

Exercice 9.9.

On dispose de deux pièces de monnaie: l'une d'entre elles est équilibrée, et l'autre possède deux côtés face. On choisit au hasard l'une des deux pièces, puis on la jette. Quelle est la probabilité d'obtenir *Pile*?

Exercice 9.10.

On lance une pièce de monnaie équilibrée jusqu'à apparition du premier *Pile*. On sait qu'un *Pile* apparait presque sûrement. Si le premier *Pile* apparaît avec le $k-$ième lancer, on remplit une urne de k boules numérotées de 1 à k, et on pioche alors une boule dans cette urne. Pour $j \in \mathbb{N}^*$, quelle est la probabilité de piocher une boule avec le numéro j ?

9.3.3 Formule de Bayes

Il est des situations dans lesquelles on peut avoir besoin d'inverser le conditionnement. Si on lance deux pièces dont une est truquée (et dont on connait les paramètres), il est facile de connaitre la probabilité d'obtenir telle ou telle face sachant la pièce qu'on lance.
Mais on pourrait vouloir connaître la probabilité d'avoir lancé une pièce truquée sachant ce qu'on a obtenu comme face.

Théorème 9.29. **Formule de Bayes ou formule de probabilité des causes**

Soient A, B deux évènements non négligeables. On a

$$\mathbf{P}_B(A) = \frac{\mathbf{P}_A(B)\mathbf{P}(A)}{\mathbf{P}(B)}.$$

Preuve. Trivial. □

Exemple 9.13.

La formule de Bayes est d'une grande importance pratique. Pour déterminer par l'exemple l'efficacité d'un vaccin, il faut déterminer la probabilité $\mathbf{P}\left(\text{être infecté}|\text{ être vacciné}\right)$. Mais on ne peut *a priori* pas déterminer cette probabilité expérimentalement (à moins d'inoculer volontairement la maladie à des personnes vaccinées!).
Par contre on peut aisément déterminer une valeur approchée de la probabilité $\mathbf{P}\left(\text{être vacciné}|\text{ être infecté}\right)$: il suffit pour cela, après avoir vacciné la population, d'effectuer un sondage dans un hôpital, et de déterminer parmi les patients infectés la proportion de ceux qui avaient été vaccinés. On déduit alors de la formule de Bayes une valeur approchée de $\mathbf{P}\left(\text{être infecté}|\text{ être vacciné}\right)$.

Théorème 9.30. **Formule de Bayes ou formule de probabilité des causes, seconde version**

Soit $(A_i)_{i\in I}$ un système complet d'évènements non négligeables, et B un évènement non négligeable. Alors pour tout $k \in I$, on a

$$\mathbf{P}_B(A_k) = \frac{P_{A_k}(B)P(A_k)}{\sum_{i\in I} \mathbf{P}_{A_i}(B)\mathbf{P}(A_i)}.$$

Preuve. Se déduit de la formule précédente, en utilisant la formule des probabilités totales pour calculer $\mathbf{P}(B)$. □

Exercice 9.11. **Le paradoxe des maladies rares**

Vous êtes directeur de cabinet du ministre de la santé. Une maladie est présente dans la population avec la proportion d'une personne malade sur 10000. Un responsable d'un grand laboratoire pharmaceutique vient vous vanter son nouveau test de dépistage: si une personne est malade, le test est positif à 99%. Si une personne n'est pas malade, le test est positif à 0,1%.

Que penser de ce test?

Exercice 9.12.

On reprend l'expérience de l'**Exercice 9. 9**. On suppose que le lancer a donné face ; quelle est la probabilité que ce soit la pièce équilibrée qui a été lancée?

9.3.4 Formule des probabilités composées

Les probabilités conditionnelles permettent de calculer la probabilité d'une suite d'évènements décroissants, connaissant la probabilité de chacun sachant les précédents.

Théorème 9.31. **Formules des probabilités composées**

Soient $n \geq 2$ évènements $A_1, \ldots, A_n$ non négligeables.

i. Si $(A_i)_{1\leq i\leq n}$ est décroissante, on a

$$\begin{aligned}\mathbf{P}(A_n) &= \left(\prod_{i=1}^{n-1} \mathbf{P}(A_{i+1}|A_i)\right) \times \mathbf{P}(A_1)\\ &= \mathbf{P}(A_n|A_{n-1})\mathbf{P}(A_{n-1}|A_{n-2})\cdots\mathbf{P}(A_2|A_1)\mathbf{P}(A_1)\end{aligned}$$

ii. Dans le cas général, on a

$$\begin{aligned}\mathbf{P}(A_1\cap\cdots\cap A_n) = \left(\prod_{i=1}^{n-1}\mathbf{P}\left(\bigcap_{k=1}^{i+1}A_k\middle|\bigcap_{k=1}^{i}A_k\right)\right)\times\mathbf{P}(A_1) &= \left(\prod_{i=1}^{n-1}\mathbf{P}\left(A_{i+1}\middle|\bigcap_{k=1}^{i}A_k\right)\right)\times\mathbf{P}(A_1)\\ =&\mathbf{P}(A_1)\mathbf{P}_{A_1}(A_2)\mathbf{P}_{A_1\cap A_2}(A_3)\times\ldots\times\mathbf{P}\left(A_n\middle|\bigcap_{k=1}^{n-1}A_k\right)\end{aligned}$$

☞ On utilise cette formule (notamment celle du *ii.* pour calculer la probabilité d'une intersection d'évènements non indépendants.

Exercice 9.13.

On considère une urne contenant une boule blanche et deux boules noires. On y effectue des tirages selon le protocole suivant: si l'on tire la boule blanche, la partie s'arrête; si l'on tire une boule noire, on remet celle-ci dans l'urne ainsi qu'une seconde boule noire (avant le tirage suivant).

1. Soit N_k l'évènement "obtenir une boule noire au k-ième tirage". Les évènements N_k, N_{k+1} sont-ils indépendants?
2. Soit $k \geq 2$. Montrer que $(N_i)_{1\leq i\leq k}$ est décroissante. En déduire $\mathbf{P}(N_k)$.
3. Quelle est la probabilité que le jeu ne s'arrête pas?

Exercice 9.14.

Une galette des rois est découpées en N parts; une seule contient la fève et les $N-1$ autres rien du tout que du beurre et du gras. On distribue les N parts successivement aux N convives qui font la queue pour êtres servi.e.s.
Montrer qu'une fois positionnées dans la queue, toutes les personnes gourmandes de ce groupe ont la même probabilité d'obtenir la fève.

10 Séries entières

Dans le **Chapitre 5**, on a rappelé qu'il était possible de *développer en série* la fonction exponentielle, et même de donner un sens à l'exponentielle complexe

$$\forall z \in \mathbb{C}, \quad e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

Cette série est un cas particulier de **série entière** : une série $\sum a_n z^n$ dont le terme général fait intervenir des puissances d'un paramètre complexe z.
Les séries entières jouent un rôle fondamental dans diverses branches des mathématiques ; elles permettent de représenter de nombreuses fonctions, de résoudre des équations différentielles, de caractériser les lois de variables aléatoires... On introduit dans ce chapitre les éléments qui permettent de les étudier.

Comme d'habitude, dans tout ce chapitre, la lettre $\mathbb{K}$ désignera l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$.

10.1 Notion de série entière, de rayon de convergence

10.1.1 Définition, exemples

Comme rappelé ci-avant, on peut écrire l'exponentielle de tout nombre complexe comme la somme d'une certaine série. Mais, pour d'autres séries dépendant d'un paramètre complexe, la convergence n'a pas lieu pour tout $z \in \mathbb{C}$.

Exercice 10.1.

1. Déterminer l'ensemble $\mathcal{D}$ des nombres réels $x \in \mathbb{R}$ tels que la série $\sum x^n$ converge.
2. Même question avec les série $\sum e^n x^n$, $\sum \cos(n) x^n$ et $\sum \frac{x^n}{n}$.
3. Déterminer l'ensemble $\mathcal{D}$ des complexes $z \in \mathbb{C}$ tels que la série $\sum z^n$ converge.
4. Même question avec la série $\sum e^n z^n$.

Définition 10.1. **Série entière**

On appelle **série entière** de la variable réelle ou complexe z toute série de la forme $\sum a_n z^n$ avec $(a_n) \in \mathbb{K}^{\mathbb{N}}$.
Les termes de la suite (a_n) sont appelés les **coefficients** de la série.

La fonction $z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ est alors appelée **fonction somme** ou plus simplement **somme** de la série.

On appelle **domaine de convergence** de la série le domaine de définition de la fonction $z \mapsto \sum_{n=0}^{+\infty} a_n z^n$, c'est à dire l'ensemble des valeurs de z pour lesquelles la série converge.

Exemple 10.1. **Série géométrique**

La série géométrique $\sum z^n$ est une série entière de la variable complexe z, de domaine de convergence $\{z \in \mathbb{C};\ |z| < 1\}$ (ou $]-1, 1[$ si on la considère comme une série de la variable réelle).

Sa somme vaut $\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$.

Remarque 10.1.

Bien que $z \mapsto \dfrac{1}{1-z}$ soit définie sur $\mathbb{C} \setminus \{1\}$, elle ne coïncide avec $z \mapsto \displaystyle\sum_{n=0}^{+\infty} z^n$ que sur $\{z \in \mathbb{C};\ |z| < 1\}$.

10.1.2 Rayon de convergence

On s'intéresse dans un premier temps au domaine de convergence des séries entières.

Proposition. 10.1. **Lemme d'Abel**

Soit $\sum a_n z^n$ une série entière, et $z_0 \in \mathbb{C}$. Si la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ est bornée, alors pour tout $z \in \mathbb{C}$ tel que $|z| < |z_0|$, la série $\sum a_n z^n$ converge absolument.

Preuve. Notons que $z_0 = 0$ n'est pas un cas intéressant à traiter : la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ est stationnaire en 0 donc bornée et il n'existe pas de nombre complexe $z \in \mathbb{C}$ tel que $|z| < z_0 = 0$.
Notons maintenant $M \in \mathbb{R}_+$tel que pour tout $z \in \mathbb{C}, |a_n z_0^n| \leqslant M$ avec $z_0 \neq 0$.
Pour tout $n \in \mathbb{N}$ et $|z| < |z_0|$:

$$|a_n z^n| = \left| a_n z_0^n \left(\frac{z}{z_0}\right)^n \right| \leqslant M \left|\frac{z}{z_0}\right|^n.$$

La série $\sum a_n z^n$ est donc absolument convergente par comparaison avec la série géométrique $\sum \left(\dfrac{z}{z_0}\right)^n$ absolument convergente car $|z| < |z_0| \iff \left|\dfrac{z}{z_0}\right| < 1$. □

Rappel 10.2.

Soit $A \subset \mathbb{R}$. On appelle *borne supérieure de* A l'élément α de $\overline{\mathbb{R}}$ défini comme suit:

- ✗ α est le plus petit élément de l'ensemble des majorants de A, si A est non vide et majoré.
- ✗ $\alpha = +\infty$ si A est non vide et non majoré.
- ✗ $\alpha = -\infty$ si A est vide.

La borne supérieure α de A est notée $\alpha = \sup A$.

Définition 10.2. **Rayon de convergence**

On appelle **rayon de convergence** de la série entière $\sum a_n z^n$ l'élément R de $\overline{\mathbb{R}}$ défini par:

$$R = \sup\{r \geq 0\ ;\ (a_n r^n)_{n\in\mathbb{N}} \text{ est bornée}\}.$$

Exemple 10.2. **Quelques rayons de convergence**

Avec les remarques précédentes,

- ✗ La série géométrique $\sum z^n$ admet $R = 1$ pour rayon de convergence.
- ✗ La série exponentielle $\sum \dfrac{z^n}{n!}$ admet $R = +\infty$ pour rayon de convergence.
- ✗ La série $\sum n! z^n$ admet $R = 0$ pour rayon de convergence.

Théorème 10.2.

Soit $\sum a_n z^n$ une série entière, et R son rayon de convergence. On a les résultats suivants:

i. si $|z| < R$, alors $\sum a_n z^n$ converge absolument;
ii. si $|z| > R$, alors $\sum a_n z^n$ diverge grossièrement;
iii. si $|z| = R$, alors on ne peut rien dire en général.

Preuve.

i. Supposons $|z| < R$. On peut trouver $r > 0$ tel que $|z| < r < R$. La suite $(a_n r^n)_{n\in\mathbb{N}}$ est donc bornée. Puisque $|z| < r$, la série $\sum a_n z^n$ est absolument convergente par le lemme d'Abel.

ii. Supposons $|z| > R$.
La suite $(a_n z^n)_{n\in\mathbb{N}}$ est donc non bornée et ne converge donc pas vers 0. La série $\sum a_n z^n$ est donc grossièrement divergente.

□

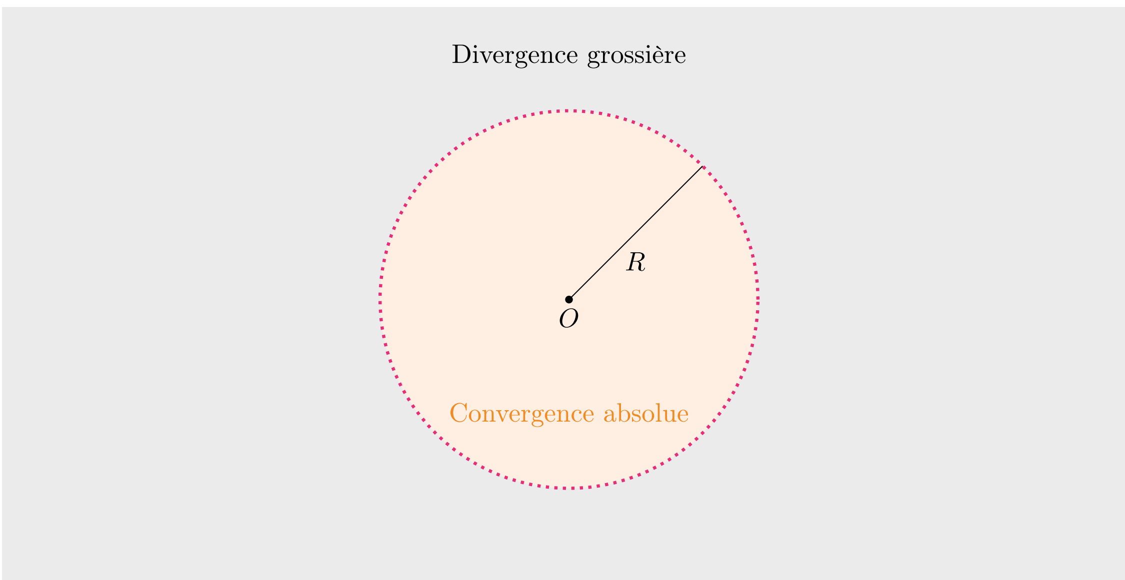
Remarque 10.2.

Soit $\sum a_n z^n$ une série entière, et R son rayon de convergence.

- ✗ Si on considère que la variable z est réelle, on en déduit que la série converge sur $]-R, R[$ (appelé *intervalle ouvert de convergence*) mais la convergence en $z = -R$ et $z = R$ nécessite une étude supplémentaire: le domaine de convergence peut être $]-R, R[, [-R, R[,]-R, R]$ ou $[-R, R]$.

- ✗ Si on considère que la variable z est complexe, alors la série converge sur $\{z \in \mathbb{C};\ |z| < R\}$ (appelé *disque ouvert de convergence*), mais la convergence dans le cas $|z| = R$ nécessite une étude supplémentaire.

☞ On retiendra qu'on ne peut rien dire à l'avance sur la convergence sur le bord du *disque de convergence*.
Sur la figure ci-dessous, le cercle en pointillés est une zone d'incertitude pour la convergence de la série entière de rayon de convergence R.



10.1.3 Calcul pratique du rayon de convergence

Méthode 10.6. **Utilisation du lemme d'Abel**

Soit une série entière $\sum a_n z^n$ de rayon de convergence R.

Si on trouve $z_0 \in \mathbb{C}$ tel que $(a_n z_0^n)_{n\in\mathbb{N}}$ est bornée (par exemple lorsque $\sum a_n z_0^n$ converge), alors $R \geq |z_0|$.

Inversement, si on trouve $z_0 \in \mathbb{C}$ tel que $\sum a_n z_0^n$ diverge, alors $R \leq |z_0|$.

Exemple 10.3.

Considérons la série entière $\sum \frac{z^n}{n}$.

La série $\sum \frac{(-1)^n}{n}$ converge donc $R \geqslant 1$. La série $\sum \frac{1}{n}$ diverge donc $R \leqslant 1$.
Ainsi, $R = 1$.

Exercice 10.2.

Soit $\sum a_n z^n$ une série entière qui converge pour $z = \sqrt{2}i$ et qui diverge pour $z = 1 + i$. Déterminer son rayon de convergence.

Méthode 10.7. **Utilisation de la règle de d'Alembert**

Pour déterminer le rayon de convergence d'une série entière $\sum a_n z^n$ on peut utiliser la règle de d'Alembert, dans le cas où a_n est non nul à partir d'un certain rang $n_0 \in \mathbb{N}$.

On pose $u_n(z) = a_n z^n$ pour tout $n \geq n_0$ et $z \in \mathbb{C}^*$. Alors $\left|\frac{u_{n+1}(z)}{u_n(z)}\right| = \left|\frac{a_{n+1}}{a_n} z\right|$, et dans le cas où $\left|\frac{a_{n+1}}{a_n}\right|$ possède une limite ℓ:

- ✗ si $\ell = 0$, alors $\forall z \in \mathbb{C}$, $\left|\frac{u_{n+1}(z)}{u_n(z)}\right| \underset{n\to+\infty}{\longrightarrow} 0$ et donc la série converge: le rayon de convergence est $R = +\infty$;
- ✗ si $\ell = +\infty$, alors $\forall z \in \mathbb{C}^*$, $\left|\frac{u_{n+1}(z)}{u_n(z)}\right| \underset{n\to+\infty}{\longrightarrow} +\infty$ et donc la série diverge: le rayon de convergence est $R = 0$;
- ✗ si $\ell \in \mathbb{R}_+^*$, alors $\forall z \in \mathbb{C}$, $\left|\frac{u_{n+1}(z)}{u_n(z)}\right| \underset{n\to+\infty}{\longrightarrow} \ell|z|$ et donc si $\ell|z| < 1$ la série converge, et si $\ell|z| > 1$ la série diverge: on a ainsi $R = \frac{1}{\ell}$.

Remarque 10.3.

On pourra retenir que, si $\sum a_n z^n$ est une série entière dont les cœfficients sont **non nuls** telle que

$$\left|\frac{a_{n+1}}{a_n}\right| \underset{n\to+\infty}{\longrightarrow} \ell,$$

alors son rayon de convergence est $R = \frac{1}{\ell}$.

Mais **attention** le théorème, ainsi formulé, ne s'adapte pas au cas des *séries lacunaires* (voir l'**Exercice 10.4**).

Exemple 10.4.

$\sum_{n\geq 2} \frac{z^n}{\ln n}$ admet $R = 1$ pour rayon de convergence.

Exercice 10.3.

Reprendre les série de l'**Exemple 10.2** en utilisant la règle de d'Alembert.

En pratique, le même raisonnement s'adapte à certains types de *séries lacunaires*, c'est-à-dire de séries possédant une infinité de termes nuls.

Exercice 10.4.

Déterminer le rayon de convergence de la série entière $\sum \frac{x^{2n+1}}{n3^n}$.

Utilisation d'un encadrement ou d'un équivalent

Proposition. 10.3.

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, de rayon de convergence respectif R_a et R_b. Si on a $|a_n| \leq |b_n|$ à partir d'un certain rang, alors $R_a \geq R_b$.

Preuve. On a bien sûr pour tout $n \in \mathbb{N}$ et $z \in \mathbb{C}$, $|a_n z^n| \leqslant |b_n z^n|$. Ainsi,

$$\begin{aligned} |z| < R_b &\Longrightarrow \sum |b_n z_n| \text{ cv} \\ &\Longrightarrow \sum |a_n z^n| \text{ cv par comparaison de séries à termes positifs} \\ &\Longrightarrow \sum a_n z^n \text{cv} \\ &\Longrightarrow |z| \leqslant R_a \end{aligned}$$

On en déduit que $R_b \leqslant R_a$. □

Exemple 10.5.

La série entière $\sum \cos(n) z^n$ a pour rayon de convergence $R = 1$. En effet, on a, pour tout entier $n \in \mathbb{N}$, $|\cos(n)| \leqslant 1$ et la série $\sum z^n$ a pour rayon de convergence $R' = 1$ donc $R \geqslant R' = 1$.
Mais, la série $\sum \cos(n) 1^n$ diverge grossièrement donc $R \leqslant 1$.

Exercice 10.5.

Déterminer le rayon de convergence R de la série entière $\sum e^{\cos(n)} z^n$.

Proposition. 10.4.

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, de rayon de convergence respectif R_a et R_b. Si on a $|a_n| \underset{n \to +\infty}{\sim} |b_n|$, alors $R_a = R_b$.

Preuve. Pour tout $z \in \mathbb{C}$, les séries $\sum |a_n z^n|$ et $\sum |b_n z^n|$ sont alors de même nature d'après le critère d'équivalence des séries à termes positifs. Par définition du rayon de convergence, ceux-ci sont égaux. □

Exercice 10.6.

Déterminer le rayon de convergence de $\sum \arctan(n) z^n$.

Proposition. 10.5.

Les séries entières $\sum a_n z^n$ et $\sum n a_n z^n$ possèdent le même rayon de convergence.

Preuve. On peut faire une preuve directe (c'est un bel exercice); on se contentera ici de l'interpréter comme conséquence du **Théorème 10.12**. □

Exemple 10.6.

La série entière $\sum \frac{z^n}{n}$ a le même rayon de convergence que la série entière $\sum n \frac{z^n}{n} = \sum z^n$ c'est-à-dire $R = 1$.

Exercice 10.7.

Déterminer le rayon de convergence de la série entière $\sum \frac{(-1)^n}{n(n-1)} z^n$.

10.2 Propriétés de la somme d'une série entière

10.2.1 Opérations sur les séries entières

Proposition. 10.6. **Multiplication par un scalaire et rayon de convergence**

Soit $\sum a_n z^n$ une série entière, R son rayon de convergence, et $\lambda \in \mathbb{K}^*$. Alors $\sum \lambda a_n z^n$ admet également R pour rayon de convergence.

De plus, on a $\sum_{n=0}^{+\infty} \lambda a_n z^n = \lambda \sum_{n=0}^{+\infty} a_n z^n$ pour tout complexe z tel que la série converge.

Proposition. 10.7. **Somme et rayon de convergence**

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, de rayon de convergence respectif R_a et R_b.
Alors, le rayon de convergence R_{a+b} de la série $\sum (a_n + b_n) z^n$ est tel que $R_{a+b} \geq \min(R_a, R_b)$.
De plus, si on suppose encore $R_a \neq R_b$, alors on a $R_{a+b} = \min(R_a, R_b)$.
Enfin, on a $\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n$ pour tout complexe z tel que la série converge.

Proposition. 10.8. **Produit de Cauchy et rayon de convergence**

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, de rayon de convergence respectif R_a et R_b.
Alors, le rayon de convergence R_c de leur produit de Cauchy, $\sum c_n z^n$ avec $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, est tel que $R_c \geq \min(R_a, R_b)$.
De plus, on a $\sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n\right) \cdot \left(\sum_{n=0}^{+\infty} b_n z^n\right)$ pour tout complexe z tel que $|z| < \min(R_a, R_b)$.

10.2.2 Propriétés de la somme vue comme une fonction de la variable réelle

Dans cette section, on ne s'intéresse qu'aux séries entières de la variable réelle. Les résultats suivants sont **admis**, conformément au programme officiel.

Théorème 10.9. **Continuité**

Soit $\sum a_n z^n$ une série entière, de rayon de convergence R. Alors, la fonction $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est continue sur l'intervalle $]-R, R[$.
De plus, si f est définie en $\pm R$, alors elle est continue en ces points.

Remarque 10.4.

On pourra retenir que $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est continue sur tout son intervalle de définition.

Théorème 10.10. **Dérivation terme à terme**

Soit $\sum a_n z^n$ une série entière, de rayon de convergence R. Alors, la fonction $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est dérivable sur l'intervalle $]-R, R[$, de sorte que:

$$\forall x \in]-R, R[, \; f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}.$$

En d'autres termes, les séries entières se *dérivent terme à terme* sur leur intervalle ouvert de convergence.

Remarque 10.5.

- ✗ Dans la série $\sum a_n x^n$, le terme $a_0 x^0$ est constant, et disparait donc lorsque l'on dérive.
- ✗ **Attention**, la fonction $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ n'est pas en général dérivable sur tout son intervalle de définition (contrairement à la continuité), car elle peut ne pas être dérivable au bord même en cas de convergence.

Exercice 10.8.

Déterminer le rayon de convergence, et calculer la somme de la série $\sum n x^n$ où $x \in \mathbb{R}$.

Corollaire 10.11. **Classe $\mathcal{C}^\infty$ de la somme**

Soit $\sum a_n z^n$ une série entière, de rayon de convergence R. Alors la fonction $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est de classe $\mathcal{C}^\infty$ sur l'intervalle $]-R, R[$.

Théorème 10.12. **Intégration terme à terme**

Soit $\sum a_n x^n$ une série entière de rayon de convergence $R > 0$.

On note $f : x \longmapsto \sum_{n=0}^{+\infty} a_n x^n$ sa fonction somme et F une primitive de la fonction $f \in \mathcal{C}^0(]-R;R[)$.

La série entière $\sum \frac{a_n}{n+1} x^{n+1}$ est une série entière de rayon de convergence R et

$$\forall x \in]-R, R[, F(x) = F(0) + \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}.$$

Autrement dit, pour pour tout $x \in]-R;R[$,

$$F(x) - F(0) = \int_0^x f(t)\mathrm{d}t = \int_0^x \sum_{n=0}^{+\infty} a_n t^n \mathrm{d}t = \sum_{n=0}^{+\infty} \int_0^x a_n t^n \mathrm{d}t = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

Remarque 10.6.
Il n'y a pas au programme de théorème permettant d'intégrer ou de dériver une série entière en $x = \pm R$.

Remarque 10.7.
Le **Théorème 10.12** ne doit pas être confondu avec le **Théorème 7.32** de permutation somme et intégrale du **Chapitre 7**.

Exercice 10.9.

1. Montrer $\forall x \in]-1, 1[,\ \arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$, et justifier que le rayon de convergence de cette série entière est 1.
2. En déduire $\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}$.

10.3 Développement en série entière d'une fonction

10.3.1 Définition et résultats fondamentaux

Définition 10.3. **Fonction DSE**

Soit $\alpha \in \overline{\mathbb{R}_+^*}$ et $I =]-\alpha, \alpha[$ un intervalle ouvert centré en 0. Une fonction f définie sur I est dite **développable en série entière** sur I s'il existe une série entière $\sum a_n x^n$ convergente sur I telle que

$$\forall x \in I,\ f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Remarque 10.8.

- ✗ Si f est développable en série entière sur un intervalle I, alors f est nécessairement $\mathcal{C}^\infty$ sur I (mais cette condition n'est pas suffisante).
- ✗ Si f est développable en série entière sur un intervalle $I =]-\alpha, \alpha[$, alors le rayon de convergence R de la série vérifie $R \geq \alpha$.
- ✗ I n'est pas nécessairement le domaine de définition de f ; en d'autres termes, f peut n'être développable en série entière **que sur une partie** de son domaine de définition (voir les exemples ci-dessous).

Exemple 10.7. **Quelques fonctions DSE**

On a prouvé précédemment les résultats suivants:

- ✗ $x \mapsto \dfrac{1}{1-x}$ est développable en série entière sur $]-1,1[$: $\forall x \in]-1,1[,\ \dfrac{1}{1-x} = \displaystyle\sum_{n=0}^{+\infty} x^n.$
- ✗ $x \mapsto e^x$ est développable en série entière sur $\mathbb{R}$: $\forall x \in \mathbb{R},\ e^x = \displaystyle\sum_{n=0}^{+\infty} \frac{x^n}{n!}.$
- ✗ $x \mapsto \arctan x$ est développable en série entière sur $]-1,1[$: $\forall x \in]-1,1[,\ \arctan x = \displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}.$

Théorème 10.13. **Condition nécessaire d'une fonction DSE**

Soit f une fonction développable en série entière sur un intervalle $I =]-\alpha, \alpha[$ avec $\alpha \in \overline{\mathbb{R}_+^*}$, de sorte que

$$\forall x \in I,\ f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

On a alors:

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

Ce théorème dit qu'en cas où il est possible de développer en série entière, les coefficients sont nécessairement ceux de la formule de Taylor-Young (c'est à dire ceux du développement limité).

Preuve. Supposons que f soit développable en série entière sur $I :=]-\alpha, \alpha[$. C'est à dire, supposons, qu'il existe $\sum a_n x^n$ une série entière de rayon de convergence R telle que $\alpha \leqslant R$ et vérifiant

$$\forall x \in]-\alpha; \alpha[, f(x) = \sum_{k=0}^{n} a_k x^k.$$

Alors f est de classe $\mathcal{C}^\infty$ sur $]-\alpha; \alpha[$ par le corollaire du **Théorème 10.10** de dérivation terme à terme.
De plus, on obtient par récurrence

$$\forall n \in \mathbb{N}, \forall x \in]-\alpha; \alpha[, f^{(p)}(x) = \sum_{n=p}^{+\infty} n(n-1)\dots(n-p+1) a_n x^{n-p}.$$

En effet, le caractère héréditaire se montre avec les arguments suivants :

$$\begin{aligned} f^{(p+1)}(x) = \left(f^{(p)}\right)'(x) = \left(\sum_{n=p}^{+\infty} n(n-1)\dots(n-p+1)a_n x^{n-p}\right)' &= \sum_{n=p}^{+\infty} n(n-1)\dots(n-p+1)(n-p)a_n x^{n-p-1} \\ &= \sum_{n=p+1}^{+\infty} n(n-1)\dots(n-(p+1)-1)a_n x^{n-(p+1)}. \end{aligned}$$

En particulier pour tout $p \in \mathbb{N}$, $f^{(p)}(0) = p!a_p$: le développement en série entière de f au voisinage de 0 coïncide bien avec la série de Taylor en 0. □

Remarque 10.9.

- ✗ On en déduit que si une fonction f est dérivable en série entière sur un intervalle I, son développement est nécessairement celui de Taylor en 0:

$$\forall x \in I,\ f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Toutefois cette égalité n'est pas nécessairement vérifiée, et même si c'est le cas elle n'est pas nécessairement valable sur tout le domaine de définition (voir l'exemple de $x \mapsto \dfrac{1}{1-x}$ ci-dessus).

- ✗ Il existe des fonctions $\mathcal{C}^\infty$, dont la série de Taylor $\sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$ converge sur $\mathbb{R}$, bien que l'égalité

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$$

ne soit vérifiée **pour aucun** réel x.

Remarque 10.10. **Ne pas confondre**

Il ne faut pas confondre la notion de développement en série entière et celle de développement limité:

- ✗ un développement limité est une notion locale, alors qu'un développement en série entière est une notion globale;
- ✗ un développement limité ne fait intervenir qu'une somme finie, alors qu'un développement en série entière fait intervenir la somme d'une série.

Toutefois, une fonction développable en série entière admet nécessairement un développement limité à tout ordre en 0, puisque l'on démontre aisément pour tout $n \in \mathbb{N}$:

$$f(x) = \sum_{k=0}^{+\infty} \frac{f^{(k)}(0)}{k!} x^k \implies f(x) \underset{x\to 0}{=} \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k + o(x^n).$$

Exemple 10.8.

On peut voir (et on renvoie au **Devoir Surveillé n°2**, Automne 2024) que la fonction f définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R}, \qquad f(x) = \begin{cases} \exp\left(-\dfrac{1}{x^2}\right), & \text{si } x \neq 0 \\ 0, & \text{si } x = 0 \end{cases}$$

était de classe $\mathcal{C}^\infty$ au voisinage de 0 et admettait un développement limité nul à tout ordre.
f ne peut donc pas être développable en série entière car elle n'est identiquement nulle sur aucun intervalle de la forme $]-\alpha, \alpha[$.

Corollaire 10.14. **Identification des coefficients sur les séries entières**

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, et $r > 0$ tels que:

$$\forall x \in]-r, r[, \ \sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} b_n x^n.$$

Alors on a $\forall n \in \mathbb{N}, \ a_n = b_n$.

Proposition. 10.15. **Parité et coefficient du DSE**

Soit f une fonction développable en série entière sur un intervalle I, de sorte que $\forall x \in I, \ f(x) = \sum_{n=0}^{+\infty} a_n x^n$. On a alors:

i. si f est paire, alors : $\forall n \in \mathbb{N}, \ a_{2n+1} = 0$;
ii. si f est impaire, alors : $\forall n \in \mathbb{N}, \ a_{2n} = 0$.

Preuve. Soit f développable en série entière sur $]-r; r[$.
Si f est paire, alors, pour tout $x \in]-r; r[$, on a :

$$\begin{aligned} f(-x) = f(x) &\Longleftrightarrow \sum_{n=0}^{+\infty} a_n(-x)^n = \sum_{n=0}^{+\infty} a_n x^n \\ &\Longrightarrow \sum_{p=0}^{+\infty} a_{2p+1} x^{2p+1} = -\sum_{p=0}^{+\infty} a_{2p+1} x^{2p+1} \\ &\Longrightarrow \sum_{p=0}^{+\infty} a_{2p+1} x^{2p+1} = 0 \Longrightarrow \forall p \in \mathbb{N}, a_{2p+1} = 0 \end{aligned}$$

par unicité du développement en série entière de la fonction nulle.

De manière analogue, le développement en série entière d'une fonction impaire ne contient que des termes de degré impair. □

10.3.2 Méthodes pratiques de développement

Par une formule de Taylor

Proposition. 10.16.
Pour tout $x \in \mathbb{R}$, on a

$$\cos(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad \text{et} \quad \sin(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

Exercice 10.10.
Montrer que la fonction définie par $f(x) = \dfrac{\cos x - 1}{x^2}$ se prolonge en une fonction $\mathcal{C}^\infty$ sur $\mathbb{R}$.

Par opérations élémentaires (combinaisons linéaires ou produit)

Proposition. 10.17.
Pour tout $x \in \mathbb{R}$, on a

$$\operatorname{ch}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!} \quad \text{et} \quad \operatorname{sh}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}.$$

Exercice 10.11.
Déterminer un développement en série entière de $x \mapsto \dfrac{1}{(x-1)(x-2)}$ de deux façons différentes.

Par intégration

Proposition. 10.18.

Pour tout $x \in]-1,1[$, on a $\ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} x^n$. De plus le rayon de convergence de cette série est $R = 1$.

Exercice 10.12.

Calculer $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n2^n}$ et $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}$.

Proposition. 10.19.

Pour tout $\forall x \in]-1,1[$, on a $\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$. De plus le rayon de convergence de cette série est 1.

Preuve. Voir l'**Exercice 10.19**. □

Par équation différentielle

Théorème 10.20.

Soit $\alpha \in \mathbb{R} \setminus \mathbb{N}$. La fonction $f_\alpha : x \mapsto (1+x)^\alpha$ est l'unique solution du problème de Cauchy suivant:

$$\begin{cases} (1+x)y' - \alpha y = 0 \\ y(0) = 1 \end{cases}.$$

On en déduit $\forall x \in]-1,1[,\ (1+x)^\alpha = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$, et le rayon de convergence de cette série est 1.

Preuve. Notons $(\mathcal{P})$ le problème de Cauchy susmentionné.
Il est clair que la fonction f_α est solution de $(\mathcal{P})$. Déterminons alors les solutions de $(\mathcal{P})$ développables en série entière, par Analyse-Synthèse.

✗ **Analyse**. Supposons qu'il existe y une fonction DSE sur $]-r;r[$ (avec $r \geq 0$) solution de (E) avec $y(0) = 1$.
Soit $x \in]-r;r[$. On a $y(x) = \sum_{n=0}^{+\infty} a_n x^n$. Par le théorème de dérivation terme à terme, on obtient $y'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}$.
Ainsi, y est solution de $(\mathcal{P})$ si et seulement si :

$$\begin{aligned}
(1+x)\sum_{n=1}^{+\infty} n a_n x^{n-1} = \alpha \sum_{n=0}^{+\infty} a_n x^n &\iff \sum_{n=1}^{+\infty} n a_n x^{n-1} + \sum_{n=1}^{+\infty} n a_n x^n = \sum_{n=0}^{+\infty} \alpha a_n x^n. \\
&\iff \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n + \sum_{n=1}^{+\infty} n a_n x^n = \sum_{n=0}^{+\infty} \alpha a_n x^n. \\
&\iff a_1 + \sum_{n=1}^{+\infty} \left((n+1)a_{n+1} + n a_n\right) x^n = \alpha a_0 + \sum_{n=1}^{+\infty} \alpha a_n x^n.
\end{aligned}$$

Par unicité du développement en série entière, y est solution de $(\mathcal{P})$ si et seulement si

$$\begin{cases} a_1 &= \alpha a_0 \\ a_{n+1} &= \dfrac{\alpha - n}{n+1} a_n, \qquad n \in \mathbb{N}^* \end{cases} \iff \forall n \in \mathbb{N},\ a_{n+1} = \frac{\alpha - n}{n+1} a_n.$$

On obtient alors par une récurrence (que l'on laisse en exercice) que :$\forall n \in \mathbb{N}^*, a_n = \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} a_0$.

Ainsi, pour tout $x \in]-r;r[$: $y(x) = a_0 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} x^n$.

Puisque $y(0) = 1$, on a $a_0 = 1$.

✗ **Synthèse**. On pose $y(x) = \sum_{n=0}^{+\infty} a_n x^n$ avec

$$a_0 = 1 \quad \text{et} \quad \forall n \in \mathbb{N}^*,\ a_n = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}.$$

On détermine maintenant le rayon de convergence R de la série entière $\sum a_n x^n$. On utilise la règle de d'Alembert : pour tout $x \neq 0$,

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{\alpha(\alpha-1)\dots(\alpha-n+1)(\alpha-n)}{(n+1)!}\frac{n!}{\alpha(\alpha-1)\dots(\alpha-n+1)}\right| = \left|\frac{\alpha-n}{n+1}\right| \underset{n\to+\infty}{\longrightarrow} 1.$$

On en déduit que $R = 1$. On a donc montré que y est développable en série entière sur $]-1;1[$, $y(0) = 1$ et y est solution de $(\mathcal{P})$ d'après les équivalences écrites dans la partie Analyse.

Les fonctions f_α et $x \mapsto \sum a_n x^n$ sont toutes deux solution de $(\mathcal{P})$. Elles coïncident par unicité de la solution au problème de Cauchy. f_α est donc développable en série entière sur $]-1;1[$ et pour tout $x \in]-1;1[$:

$$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n.$$

□

Remarque 10.11.

✗ Si $\alpha \in \mathbb{N}$, il suffit d'utiliser le binôme de Newton pour retrouver la formule ci-dessus, qui devient alors une somme finie (et donc son rayon de convergence est $+\infty$).

✗ La formule de Taylor permet également d'obtenir le développement de $x \mapsto (1+x)^\alpha$.

10.3.3 Développements en série usuels

Soit $\alpha \in \mathbb{R} \setminus \mathbb{N}$. Les développements suivants doivent être sus.

Fonction	DSE	Rayon de convergence
$\frac{1}{1-x}$	$\sum_{n=0}^{+\infty} x^n$	$R = 1$
$\ln(1+x)$	$\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} x^n$	$R = 1$
e^x	$\sum_{n=0}^{+\infty} \frac{x^n}{n!}$	$R = +\infty$
$\operatorname{ch}(x)$	$\sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}$	$R = +\infty$
$\operatorname{sh}(x)$	$\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}$	$R = +\infty$
$(1+x)^\alpha$	$\sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$	$R = 1$
$\cos(x)$	$\sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n}$	$R = +\infty$
$\sin(x)$	$\sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$	$R = +\infty$
$\arctan(x)$	$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$	$R = 1$

11 Espaces préhilbertiens réels

Dans ce chapitre, E désigne un $\mathbb{R}$-espace vectoriel. On généralise à E la notion de produit scalaire vue dans le cas du plan et de l'espace euclidien.

11.1 Notion de produit scalaire et de norme sur un espace vectoriel

11.1.1 Définitions, exemples

Définition 11.1. **Produit scalaire**

On appelle **produit scalaire** sur E toute application $\varphi : E^2 \to \mathbb{R}$ possédant les propriétés suivantes quels que soient les vecteurs u, u', v, v' de E, et les réels λ, λ'.

i. Linéarité à gauche: $\varphi(\lambda u + \lambda' u', v) = \lambda\varphi(u, v) + \lambda'\varphi(u', v)$.
ii. Linéarité à droite: $\varphi(u, \lambda v + \lambda' v') = \lambda\varphi(u, v) + \lambda'\varphi(u, v')$.
iii. Symétrie: $\varphi(u, v) = \varphi(v, u)$.
iv. φ est définie positive: $\varphi(u, u) \geq 0$, et $\varphi(u, u) = 0 \iff u = 0_E$.

On dit également qu'un produit scalaire est une **forme bilinéaire**, symétrique et définie positive.

Si φ est un produit scalaire, la quantité $\varphi(x, y)$ sera le plus souvent notée $x \cdot y$, $(x|y)$, $< x, y >$ ou (x, y).

Remarque 11.1.

- ✗ Plutôt que de montrer que φ vérifie *i.* et *ii.*, on commence par montrer que φ vérifie `i.` (linéarité à gauche) puis qu'elle est symétrique (c'est à dire qu'elle vérifie *iii.*). Elle vérifie alors nécessairement *ii.*

- ✗ Il est possible de définir plusieurs produits scalaires différents sur un même espace vectoriel.

Exemple 11.1.

i. Soient $n \in \mathbb{N}^*$ et $u = (x_1, \ldots, x_n)$, $v = (y_1, \ldots, y_n)$ dans $\mathbb{R}^n$.
Alors la relation
$$(u|v) = \sum_{k=1}^{n} x_k y_k$$
définit un produit scalaire sur $\mathbb{R}^n$, appelé **produit scalaire canonique** (ou usuel) sur $\mathbb{R}^n$.
Si $n = 2$ (resp. 3) on retrouve ainsi le produit scalaire qui avait été défini sur $\mathbb{R}^2$ (resp. $\mathbb{R}^3$).

ii. Soient $a < b$ des réels. Si $f, g \in \mathcal{C}^0([a, b])$ la relation
$$(f|g) = \int_a^b f(t)g(t)\mathrm{d}t$$
définit un produit scalaire sur $\mathcal{C}^0([a, b])$.
iii. La relation ci-dessus ne définit pas un produit scalaire sur $\mathcal{C}^0(\mathbb{R})$.

iv. Soient $a_0 < a_1 < \cdots < a_n$ des réels. Si $P, Q \in \mathbb{R}_n[X]$ la relation
$$(P|Q) = \sum_{k=0}^{n} P(a_k)Q(a_k)$$
définit un produit scalaire sur $\mathbb{R}_n[X]$.

v. La relation ci-dessus ne définit pas un produit scalaire sur $\mathbb{R}[X]$.

vi. Soient A, B deux matrices de $\mathcal{M}_n(\mathbb{R})$. La relation

$$\langle A, B\rangle = \operatorname{Tr}\left(A^\top B\right)$$

définit un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$.

Définition 11.2. **Espace euclidien, espace pré-hilbertien**

✗ On appelle **espace vectoriel préhilbertien réel** tout $\mathbb{R}$-espace vectoriel muni d'un produit scalaire.

✗ On appelle **espace vectoriel euclidien** tout $\mathbb{R}$-espace vectoriel **de dimension finie** muni d'un produit scalaire.

Définition 11.3. **Norme associée à un produit scalaire**

Soit E un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire $(.|.)$. On appelle **norme** préhilbertienne associée à $(.|.)$, et on note $\|\cdot\|$, l'application définie par

$$\begin{array}{cccc} \|\cdot\| & E & \longrightarrow & E \\ & x & \longmapsto & \sqrt{(x|x)} \end{array}$$

Dans la suite, E désigne un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire $(.|.)$, et $\|\cdot\|$ désigne la norme associée.

11.1.2 Propriétés fondamentales

Proposition. 11.1.

Pour tout vecteur u de E, on a $(u|0_E) = 0$.

Proposition. 11.2. **Propriétés de la norme préhilbertienne**

On a les propriétés suivantes, quel que soit le vecteur u de E et le réel λ.

i. $\|u\| \geq 0$.

ii. $\|u\| = 0 \iff u = 0_E$.

iii. $\|\lambda u\| = |\lambda| \cdot \|u\|$.

Proposition. 11.3. **Propriétés liant la norme au produit scalaire**

On a alors les propriétés suivantes, quels que soient les vecteurs u, v de E et le réel λ.

i. **Formule d'Al-Kashi.** $\|u+v\|^2 = \|u\|^2 + 2\,(u|v) + \|v\|^2$.

ii. **Identité du parallélogramme.** $\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$.

iii. **Formules de polarisation.** $(u|v) = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right)$ et $(u|v) = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right)$.

Théorème 11.4. **Inégalité de Cauchy-Schwarz**

Soient deux vecteurs u, v de E. On a : $|(u|v)| \leq \|u\| \times \|v\|$.
De plus, cette inégalité est une égalité si et seulement si u et v sont colinéaires.

Preuve. L'idée est de considérer la fonction polynomiale réelle de degré 2 $f:\ \lambda \longmapsto \|x+\lambda y\|^2$. Le discriminant de celle-ci est alors négatif ou nul. □

Remarque 11.2.

L'inégalité de Cauchy-Schwarz permet, dans un espace vectoriel muni d'un produit scalaire, de définir l'angle géométrique (c'est-à-dire compris entre 0 et π) entre deux vecteurs u, v non nuls en posant:

$$\widehat{(u,v)} = \arccos\left(\frac{(u|v)}{\|u\| \cdot \|v\|}\right).$$

Exemple 11.2.
L'inégalité de Cauchy-Schwarz donne
$$\left|\int_0^1 f(t)g(t)\mathrm{d}t\right| \leq \sqrt{\int_0^1 f(t)^2\mathrm{d}t} \times \sqrt{\int_0^1 g(t)^2\mathrm{d}t}$$
pour toutes fonctions f, g continues sur $[0,1]$. Et on a égalité si et seulement si f, g sont colinéaires.

Exercice 11.1.

Soient $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$. Montrer $\left(\sum_{k=1}^n x_k y_k\right)^2 \leq \left(\sum_{k=1}^n x_k^2\right) \times \left(\sum_{k=1}^n y_k^2\right)$.

Proposition. 11.5. **Inégalités triangulaires**

On a les propriétés suivantes, quels que soient les vecteurs u, v de E.

i. **Inégalité triangulaire (ou de Minkowski).** $\|u+v\| \leq \|u\| + \|v\|$.
De plus, cette inégalité est une égalité si et seulement si u et v sont colinéaires et de même sens.

ii. **Inégalité triangulaire inversée.** $|\ \|u\| - \|v\|\ | \leq \|u - v\|$.

11.1.3 Notion de norme

Partant des propriétés précédentes, on peut généraliser la notion de norme.

Définition 11.4.
On appelle **norme** sur E toute application $N : E \to \mathbb{R}_+$ vérifiant les propriétés suivantes quels que soient $(u, v) \in E^2$ et $\lambda \in \mathbb{R}$.

i. $N(u) \geq 0$, et $N(u) = 0 \iff u = 0_E$.
ii. $N(\lambda u) = |\lambda| \cdot N(u)$.
iii. $N(u+v) \leq N(u) + N(v)$.

L'inégalité triangulaire inversée se déduit alors aisément des propriétés ci-dessus en raisonnant comme précédemment.

Exemple 11.3.

- ✗ La valeur absolue est une norme sur $\mathbb{R}$.
- ✗ La norme préhilbertienne associée à un produit scalaire est bien une norme au sens de la définition ci-dessus.

Définition 11.5. **Vecteur normé**
Soit N une norme sur E. On dit qu'un vecteur $x \in E$ est unitaire ou **normé** si $N(x) = 1$.

Exemple 11.4.
Si $x \neq 0_E$, les vecteurs $\dfrac{1}{N(x)}x$ et $-\dfrac{1}{N(x)}x$ sont normés. Ce sont les seuls vecteurs normés colinéaires à x.

Étant donnée une norme, on peut définir une notion de distance.

Définition 11.6. **Distance associée à une norme**
Soit N une norme sur E. On définit une application $d : E^2 \to \mathbb{R}_+$, appelée **distance associée** à N, en posant $d(x, y) = N(x - y)$.

Proposition. 11.6. **Propriétés de la distance**

On a alors les propriétés suivantes, quels que soient les vecteurs u, v, w de E.

i. $d(u,v) = 0 \iff u = v$.
ii. $d(u,v) = d(v,u)$.
iii. $d(u,w) \leq d(u,v) + d(v,w)$.

11.2 Familles de vecteurs et orthogonalité

11.2.1 Définitions et propriétés élémentaires

Définition 11.7. **Vecteurs orthogonaux**

Deux vecteurs x, y de E sont dit **orthogonaux** si $(x|y) = 0$. Cette propriété sera notée $x \perp y$.

Théorème 11.7. **Théorème de Pythagore**

Soit $(x,y) \in E^2$. On a alors:

$$x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Preuve. Ce fameux théorème se déduit immédiatement et sans difficulté de la formule d'Al-Kashi. □

Définition 11.8. **Famille orthogonale \ orthonormale**

Soit $(e_i)_{i\in I}$ une famille de vecteurs de E.

✗ Une famille $(e_i)_{i\in I}$ est dite **orthogonale** si ses vecteurs sont deux-à-deux orthogonaux, c'est-à-dire si $\forall (i,j) \in I^2,\ i \neq j \Rightarrow (e_i|e_j) = 0$.

✗ La famille $(e_i)_{i\in I}$ est dite **orthonormale** (ou orthonormée) si ses vecteurs sont deux-à-deux orthogonaux et normés, c'est-à-dire si $\forall (i,j) \in I^2,\ i \neq j \Rightarrow (e_i|e_j) = 0$ et $\forall i \in I,\ \|e_i\| = 1$.

Exemple 11.5.

La base canonique de $\mathbb{R}^n$ est orthonormée pour le produit scalaire usuel, pour tout $n \in \mathbb{N}^*$.

Théorème 11.8. **Formule de Pythagore**

Soit $n \in \mathbb{N}^*$ et $(u_i)_{1\leq i\leq n}$ une famille orthogonale de E. On a alors:

$$\left\| \sum_{i=1}^{n} u_i \right\|^2 = \sum_{i=1}^{n} \|u_i\|^2.$$

Proposition. 11.9.

Toute famille orthogonale de vecteurs non nuls est libre (qu'elle soit finie ou pas).

☞ On en déduit directement que toute famille orthonormée est libre.

Exercice 11.2.

Dans cet exercice, on se place dans l'espace préhilbertien $\mathcal{C}^0([0,2\pi])$ muni du produit scalaire défini par

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\mathrm{d}x.$$

On considère la famille $\mathcal{F} = (x \mapsto a_p \cos(px))_{p\in\mathbb{N}}$.

1. Déterminer pour quelle valeur de $(a_p)_{p\in\mathbb{N}}$ la famille $\mathcal{F}$ est orthonormée.
2. En déduire que $\mathcal{F}$ est libre.

11.2.2 Propriétés élémentaires des bases orthonormées en dimension finie

Les coordonnées d'un vecteur dans une base orthonormée se calculent aisément: il suffit de projeter celui-ci sur les vecteurs de base.

Théorème 11.10.

Soit E un **espace vectoriel euclidien** et $(e_i)_{1\leq i\leq n}$ une base orthonormée de E. Alors pour tout vecteur $x \in E$ on a

$$x = \sum_{i=1}^{n} (x|e_i)e_i.$$

De même, le produit scalaire de deux vecteurs se calcule aisément dans les bases orthonormées: quitte à passer aux matrices de ces vecteurs, tout se passe comme si on était dans $\mathbb{R}^n$ muni du produit scalaire canonique.

Théorème 11.11.

Soit E un **espace vectoriel euclidien**, $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une base orthonormée de E, et $x = \sum_{i=1}^{n} x_i e_i$, $y = \sum_{i=1}^{n} y_i e_i$ deux vecteurs de E. On a alors les résultats suivants.

i. $(x|y) = \sum_{i=1}^{n} x_i y_i$ et $\|x\|^2 = \sum_{i=1}^{n} x_i^2$.

ii. En d'autres termes, si X (resp. Y) est la matrice de x (resp. y) dans la base $\mathcal{B}$, on a $(x|y) = X^\top \times Y$ et $\|x\|^2 = X^\top \times X$.

11.2.3 Construction de bases orthonormées: le procédé d'orthogonalisation de Gram-Schmidt

La preuve du théorème suivant correspond à la mise en œuvre de l'*algorithme d'orthogonalisation de Gram-Schmidt*.

Théorème 11.12. **Procédé d'orthogonalisation de Gram-Schmidt**

Soit $n \in \mathbb{N}^*$ et $(e_i)_{1\leq i\leq n}$ une famille libre de vecteurs de E. Il existe alors une unique famille $(v_i)_{1\leq i\leq n}$ de vecteurs de E telle que:

i. La famille $(v_i)_{1\leq i\leq n}$ est orthonormée;

ii. $\forall k \in [\![1, n]\!]$, $\operatorname{Vect}(e_1, \dots, e_k) = \operatorname{Vect}(v_1, \dots, v_k)$;

iii. $\forall k \in [\![1, n]\!]$, $(e_k|v_k) > 0$.

Remarque 11.3.

- ✗ Si l'on supprime la condition *iii.*, on perd l'unicité.
- ✗ Le principe de la démonstration doit impérativement être connu, puisqu'à chaque fois que l'on utilise concrètement ce résultat, on le redémontre.
- ✗ En pratique, on commence par construire une famille orthogonale, et on ne la norme qu'à la fin (cela simplifie les calculs).
- ✗ Dès que le cardinal de la famille libre est supérieur à 4, le recours à un outil informatique de calcul est précieux. Pour implémenter cet algorithme, le plus commode est de normer directement les vecteurs et de mémoriser les formules de récurrence suivantes:

$$v_1 = \frac{1}{\|e_1\|}e_1, \qquad \forall k \in [\![2, n]\!],\ w_k = e_k - \sum_{i=1}^{k-1} (e_k|v_i)v_i \quad \text{et} \quad v_k = \frac{1}{\|w_k\|}w_k.$$

Exercice 11.3.

Appliquer l'algorithme de Gram-Schmidt à la famille $(1,0,0), (1,1,0), (1,1,1)$ de $\mathbb{R}^3$ muni du produit scalaire canonique.

Exercice 11.4. **Polynômes de Legendre**

On muni $\mathbb{R}_2[X]$ du produit scalaire défini par $(P|Q) = \int_{-1}^{1} P(t)Q(t)\mathrm{d}t$. Appliquer l'algorithme de Gram-Schmidt à la base canonique de $\mathbb{R}_2[X]$.

Théorème 11.13. **Théorème de la base incomplète orthonormée**
Soit E un **espace vectoriel euclidien**. Toute famille orthonormée de E se complète en une base orthonormée de E.

Corollaire 11.14.
Tout **espace vectoriel euclidien** admet une base orthonormée.

11.3 Orthogonal d'un sous espace vectoriel

Définition 11.9. **Orthogonal**
Soit $A \subset E$.

- ✗ On appelle **orthogonal** de A, et l'on note $A^\perp$, l'ensemble des vecteurs $x \in E$ qui sont orthogonaux à tous les vecteurs de A. En d'autres termes on a
$$A^\perp = \{x \in E;\ \forall a \in A,\ (x|a) = 0\}.$$
- ✗ On dit qu'un vecteur $x \in E$ est **orthogonal** à A, si $x \in A^\perp$.

Proposition. 11.15.
Soit $A \subset E$. On a les résultats suivants.

i. $E^\perp = \{0_E\}$ et $\{0_E\}^\perp = E$.
ii. $A^\perp$ est un sous-espace vectoriel de E.
iii. Si A est sous-espace vectoriel de E, $A \cap A^\perp = \{0_E\}$.
iv. $A \subset {A^\perp}^\perp$.

Remarque 11.4.
Le point *i.* établit que 0_E est le seul vecteur orthogonal à tout l'espace. C'est par ailleurs le seul vecteur orthogonal à lui-même.

Théorème 11.16. **Supplémentaire orthogonal**
Soit F un sous-espace vectoriel **de dimension finie** de E. Alors $E = F \oplus F^\perp$. En d'autres termes, $F^\perp$ est un supplémentaire de F dans E, appelé **le supplémentaire orthogonal** de F..

Remarque 11.5.
En particulier, dans le cas où E est de dimension finie, on en déduit $\dim(F^\perp) = \dim E - \dim F$.

Proposition. 11.17.
Soit E un **espace vectoriel euclidien**, et F un sous-espace vectoriel de E. On a alors ${F^\perp}^\perp = F$.

Proposition. 11.18.
Soit E un **espace vectoriel euclidien**, et H un hyperplan de E. Alors il existe un vecteur non nul $v_N \in E$ tel que
$$\forall u \in E,\ u \in H \iff (u|v_N) = 0.$$

Remarque 11.6.

- ✗ Un tel vecteur v_N est appelé un **vecteur normal** à H.
- ✗ H est entièrement déterminé par la donnée d'un tel vecteur.
- ✗ Si on se place dans une base orthonormée $\mathcal{B}$ de E, et si on appelle $(a_1, \ldots, a_n)$ les coordonnées de v_N dans celle-ci, on obtient que H admet $a_1x_1 + \cdots + a_nx_n = 0$ pour équation cartésienne dans $\mathcal{B}$. On retrouve ainsi, dans le cas particulier des espaces euclidiens, un résultat connu.
- ✗ H admet exactement deux vecteurs normaux et normés, qui sont opposés. L'équation obtenue ci-dessus est alors appelée une *équation normale* de H.

11.4 Projections et symétries orthogonales

Dans cette section, la lettre F désigne un sous-espace vectoriel **de dimension finie** d'un espace préhilbertien quelconque E.

Définition 11.10. **Projcteur orthogonal**

Soit F un sous-espace vectoriel de dimension finie de E. On appelle **projecteur orthogonal sur** F, et on note p_F, le projecteur sur F parallèlement à $F^\perp$.

Remarque 11.7.

En tant que projecteur, p_F est linéaire et vérifie $p_F^2 = p_F$, $F^\perp = \mathrm{Ker}(p_F)$, $F = \mathrm{Im}(p_F) = \mathrm{Ker}(p_F - \mathrm{Id}_E)$.

Proposition. 11.19.

Soit F un sous-espace vectoriel de dimension finie de E et p_F le projecteur orthogonal sur F. Pour tout $u \in E$, on a alors $u - p_F(u) \in F^\perp$.

☞ Pour tout $x \in E$, on peut écrire $x = p_F(x) + x - p_F(x)$, et il suit, par **Pythagore**, que

$$||x||^2 = ||p_F(x)||^2 + ||x - p_F(x)||^2.$$

Exercice 11.5.

On considère $E = \mathbb{R}^3$ muni de sa structure euclidienne canonique. On introduit l'endomorphisme p dont la matrice dans la base canonique de E est donnée par

$$A = \frac{1}{6}\begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{pmatrix}$$

Montrer que p est une projection orthogonale sur un plan dont on précisera l'équation.

Théorème 11.20.

Soit F un sous-espace vectoriel de dimension finie de E, et $(e_1, \ldots, e_p)$ une base orthonormée de F. On a alors

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

Exercice 11.6.

Utiliser cette formule pour retrouver la formule de récurrence nécessaire à l'implémentation de l'algorithme de Gram-Schmidt.

Exercice 11.7.

Soit $E = \mathbb{R}^4$ muni de son produit scalaire canonique et on note $\mathcal{B} = (e_1, e_2, e_3, e_4)$ sa base canonique. On considère le sous-espace vectoriel

$$F = \left\{ u = (x, y, z, t) \in \mathbb{R}^4 : \left\{ \begin{array}{rcl} x + y & = & 0 \\ z + t & = & 0 \end{array} \right. \right\}$$

1. Déterminer une base orthonormale de F.
2. Déterminer la matrice M de la projection orthogonale sur F dans $\mathcal{B}$.

Exercice 11.8.

On considère $E = \mathbb{R}_n[X]$ muni du produit scalaire

$$\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t}\mathrm{d}t$$

1. Vérifier que $\langle \cdot, \cdot \rangle$ est bien un produit scalaire sur E.
2. Préciser la valeur de $\langle X^k, X^j \rangle$, pour $1 \leq k, j \leq n$.
3. Déterminer $p_F(X^2)$, où $F = \mathbb{R}_1[X]$. (*On pourra procéder de deux façons différentes.*)

Théorème 11.21.
Soit F un sous-espace vectoriel de dimension finie de E et $x \in E$. Alors $p_F(x)$ minimise la distance entre x et un vecteur de F, et c'est l'unique vecteur dans ce cas.

On en déduit la définition suivante, qui possède de nombreuses applications.

Définition 11.11. **Distance d'un vecteur à un sous-espace**
Soit F un sous-espace vectoriel de dimension finie de E et $x \in E$. On appelle **distance de** x à F, et on note $\mathrm{d}(x, F)$ le réel défini par

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

Théorème 11.22. **Inégalité de Bessel**
Soit F un sous-espace vectoriel de dimension finie de E, et $(e_1, \ldots, e_p)$ une base orthonormée de F. Pour tout $x \in E$ on a alors $\|p_F(x)\| \leq \|x\|$, qui s'écrit encore

$$\sum_{i=1}^{p} (x|e_i)^2 \leq \|x\|^2.$$

Exercice 11.9.
Soit $p \in \mathcal{L}(E)$ une projection. Le but de l'exercice est de montrer la propriété suivante:

$$p \text{ projection orthogonale } \iff \forall x \in E,\ ||p(x)|| \leq ||x||.$$

1. Montrer le sens $\Rightarrow$.
2. On veut montrer la réciproque. On suppose donc, que pour tout $x \in E$, on a $||p(x)|| \leq ||x||$. Soient $(y, z) \in \mathrm{Im}(p) \times \mathrm{Ker}(p)$.
On introduit la fonction $\varphi :\ t \mapsto ||y + tz||^2$.
  a. Justifier que φ est dérivable sur $\mathbb{R}$ et expliciter $\varphi'(t)$.
  b. Montrer que, pour tout $t \in \mathbb{R}$, $\varphi(t) \geq \varphi(0)$. En déduire que $\varphi'(0) = 0$.
  c. Conclure.

Exercice 11.10.
En revenant à l'**Exercice 11.5**, déterminer la distance de $(1, 1, 1)$ au plan sur lequel on projette.

Exercice 11.11.
Déterminer $\displaystyle\min_{\alpha \in \mathbb{R}} \int_0^1 (t^2 - \alpha)^2 \mathrm{d}t$.

On peut de façon analogue définir la notion de symétrie orthogonale.

Définition 11.12. **Symétrie orthogonale**
Soit F un sous-espace vectoriel de dimension finie de E. On appelle **symétrie orthogonale par rapport à** F, et l'on note s_F, la symétrie par rapport à F parallèlement à $F^\perp$.

Remarque 11.8.

- ✗ En tant que symétrie, s_F est linéaire et vérifie $s_F^2 = \mathrm{Id}_E$, $F^\perp = \mathrm{Ker}(s_F + \mathrm{Id}_E)$, $F = \mathrm{Ker}(s_F - \mathrm{Id}E)$. On a encore $s_F = 2p_F - \mathrm{Id}$, où p_F est le projecteur orthogonal sur F.
- ✗ On en déduit notamment que si $u \in E$, alors $u - s_F(u) \in F^\perp$ et $u + s_F(u) \in F$.
- ✗ Dans le cas où E est de dimension finie et $F = H$ est un hyperplan de E, on dit que s_H est la *réflexion* d'hyperplan H.

On a enfin le résultat suivant.

Théorème 11.23.

Soit E un **espace vectoriel euclidien**, et F un sous-espace vectoriel de E. Les matrices de p_F et s_F dans une base orthonormée sont des matrices symétriques.

Preuve. Notons $A = (a_{i,j})$ la matrice de p_F, projection orthogonale sur F, dans une b.o.n $(u_1, ..., u_n)$ de (E). Alors, on peut voir que

$$(\forall i, j, \qquad a_{i,j} = \langle p_F(u_j), u_i \rangle.)$$

Observons alors que, pour tout $x \in E$, $x - p_F(x) \in F^\perp$ et donc, comme, pour tout $(y \in E,\ p_F(y) \in F)$, pour tous $x, y \in E$, $\langle x - p_F(x), p_F(y) \rangle = 0$, ce qui donne, par linéarité à gauche, puis en reproduisant le raisonnement $\langle x, p_F(y) \rangle = \langle p_F(x), p_F(y) \rangle = \langle y, p_F(x) \rangle$. En prenant $(x = u_i)$ et $(y = u_j)$, on a bien $a_{u,j} = a_{j,i}$ et donc A est symétrique. Il est découle que la matrice de s_F aussi. $\square$

Remarque 11.9.

Attention, la matrice d'une symétrie n'est pas nécessairement symétrique (il en va de même des projecteurs). Par exemple,

$$S = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$$

vérifie $S^2 = I_2$ donc c'est la matrice d'une symétrie, bien qu'elle ne soit pas symétrique.

12 Équations différentielles

Dans ce chapitre, nous allons apprendre à résoudre de nouvelles équations différentielles (c'est à dire davantage d'équations différentielles que celles qu'on sait déjà résoudre grâce au cours de première année, notamment grâce au développement en série entière).
On considère un intervalle non trivial I de $\mathbb{R}$, et $\mathbb{K}$ désignera soit $\mathbb{R}$ soit $\mathbb{C}$.

12.1 Résultats généraux

12.1.1 Vocabulaire et structure de l'ensemble des solutions

Définition 12.1.
Soit $n \in \mathbb{N}^*$.

✗ On appelle **équation différentielle linéaire d'ordre** n toute équation de la forme

$$(E) \quad a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = b,$$

où
- $a_0, \ldots, a_n$ et b sont des fonctions continues sur I et à valeurs dans $\mathbb{K}$;
- y est une fonction inconnue.

✗ Dans le cas où la fonction b est nulle, on dit que l'équation est **homogène** (ou sans second membre).

✗ On appelle **équation homogène associée** à E l'équation obtenue en remplaçant b par 0.
Dans le cas où la fonction a_n est constante et vaut 1, l'équation est dite **normalisée** ou **résolue** en $y^{(n)}$.

✗ Soit $J \subset I$ un intervalle véritable. On dit qu'une fonction y est **solution** de (E) sur J si y est n fois dérivable sur J, de sorte que $\forall x \in J,\ a_n(x)y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \cdots + a_1(x)y'(x) + a_0(x)y(x) = b(x)$.

✗ Résoudre une équation différentielle, c'est déterminer l'ensemble de ses solutions sur I.

☞ Par abus de notation, toute équation différentielle de la forme $a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = b$ sera notée $a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = b(x)$, pour marquer le fait que $a_0, \ldots, a_n, b$ sont *a priori* des fonctions et non des constantes.

Théorème 12.1. **Structure de l'ensemble des solutions**
Soit $\mathcal{S}$ l'ensemble des solutions sur I de $a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = b(x)$, et $\mathcal{S}_0$ l'ensemble des solutions sur I de l'équation homogène associée. On a alors les résultats suivants.

i. $\mathcal{S}_0$ est un sous-$\mathbb{K}$-espace vectoriel de $\mathbb{K}^I$.
ii. Si (E) possède au moins une solution y_0, alors $\mathcal{S} = \{y_0 + y_h;\ y_h \in \mathcal{S}_0\}$.

Remarque 12.1.

- ✗ Une équation différentielle homogène possède une infinité de solution, mais on peut en pratique donner sa solution en fonction d'une ou plusieurs constantes, dont la valeur est associée aux conditions initiales (voir la suite). Lorsque celles-ci ne sont pas fixées, et que les constantes sont quelconques, on parle de *la solution générale* de l'équation.
  Par exemple, $y' = ay$ admet $y : x \mapsto \lambda e^{ax}$ pour solution générale.

- ✗ Le point *ii.* prouve que la solution générale d'une équation (E) est somme de la solution générale de l'équation homogène associée (E_0) et d'une solution *particulière* de (E). Il peut être retenue sous la forme suivante: $y_{\text{gén},E} = y_{\text{gén},E_0} + y_{\text{part},E}$.

12.1.2 Théorème de Cauchy-Lipschitz

Théorème 12.2. **Cauchy-Lipschitz**

On considère une équation différentielle de la forme $a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = b(x)$, avec $a_0, \ldots, a_n, b$ continues sur I, et on suppose que a_n ne s'annule pas. Pour tout $x_0 \in I$ et $(\alpha_0, \ldots, \alpha_{n-1}) \in \mathbb{K}^n$, l'équation possède une unique solution y sur I qui vérifie

$$y(x_0) = \alpha_0,\ y'(x_0) = \alpha_1,\ \ldots,\ y^{(n-1)}(x_0) = \alpha_{n-1}.$$

Preuve. Résultat admis. □

Remarque 12.2.

- ✗ La condition a_n ne s'annule pas sur I est **nécessaire**: sans elle le résultat n'est plus vrai (voir les problèmes de recollement de solutions dans la suite). Dans ce cas, on peut alors diviser l'équation par a_n pour obtenir une équation normalisée.

- ✗ La donnée d'une équation différentielle d'ordre n et de conditions initiales associées est appelée un **problème de Cauchy**. Mais il faut bien noter la forme des conditions initiales: leur nombre est égal à l'ordre de l'équation, soit n, et elles portent sur les valeurs de $y, \ldots, y^{(n-1)}$ en un **même point** x_0.

Définition 12.2.

Étant donnée une équation différentielle (E), on appelle **courbe intégrale associée** à (E) toute courbe représentative d'une solution de (E).

Exemple 12.1.

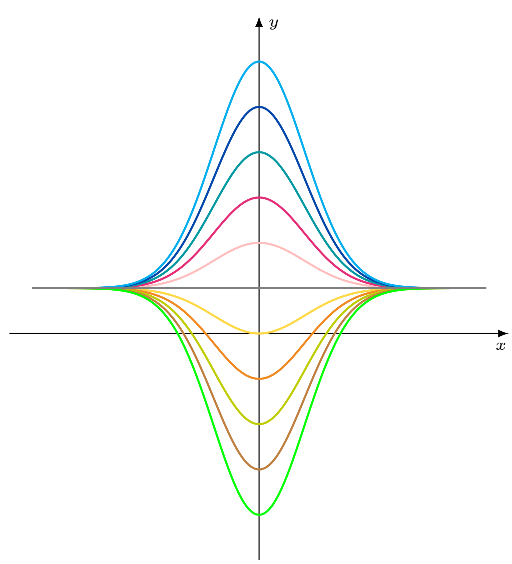
Le théorème précédent montre que l'ensemble des courbes intégrales d'une équation différentielle forme une partition de $I \times \mathbb{K}$. Dans le cas où $\mathbb{K} = I = \mathbb{R}$, on obtient ainsi une partition du plan $\mathbb{R}^2$: en d'autres termes, il passe en chaque point du plan une courbe intégrale et une seule.

La figure ci-contre donne le tracé des courbes représentatives de la solution y_k du problème de Cauchy

$$\begin{cases} y' + xy = x \\ y(0) = k \end{cases}$$

pour $k \in [\![-5, 5]\!]$.
On verra qu'on peut résoudre ce problème de Cauchy et que

$$y_k : x \longmapsto (k-1)e^{-x^2/2} + 1.$$



12.1.3 Principe de superposition des solutions

Théorème 12.3. **Principe de superposition**

Soient $b_1, b_2 : I \to \mathbb{K}$ des fonctions continues, et $\lambda \in \mathbb{K}$. Soient encore y_1 (resp. y_2) une solution de l'équation différentielle $a_n(x)y^{(n)} + \cdots + a_0(x)y = b_1(x)$ (resp. $a_n(x)y^{(n)} + \cdots + a_0(x)y = b_2(x)$). On a alors les résultats suivants.

i. λy_1 est une solution de $a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = \lambda b_1(x)$.

ii. $y_1 + y_2$ est une solution de $a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = b_1(x) + b_2(x)$.

12.2 Équations linéaires d'ordre un

Cette section présente des résultats vus en première année. On explique dans un premier temps comment résoudre une équation normalisée, puis comment traiter le cas général.

12.2.1 Méthode de résolution des équations normalisées

Cas des équations homogènes.

On commence par le cas d'une équation normalisée homogène.

Proposition. 12.4.

Soit $a : I \to \mathbb{K}$ une fonction continue. La solution générale de $y' + a(x)y = 0$ est $y : x \mapsto \lambda e^{-A(x)}$ où $\lambda \in \mathbb{K}$, et où A est une primitive de a sur I.
En d'autres termes, l'ensemble des solutions de $y' + a(x)y = 0$ sur I est de dimension 1, et c'est la droite vectorielle dirigée par $y : x \mapsto e^{-A(x)}$, où A est une primitive de a sur I.

Remarque 12.3.

Dans le cadre d'une équation de la forme $y' + a(x)y = 0$, le théorème de Cauchy-Lipschitz assure que la seule solution s'annulant est la fonction nulle, et donc qu'aucune autre solution ne change de signe. Cela permet de retrouver le résultat précédent en raisonnant comme suit:

$$\begin{aligned} y' + a(x)y = 0 &\iff \frac{y'}{y} = -a(x) \iff \ln|y| = -A(x) + c \\ &\iff |y| = e^{-A(x)+c} \iff y = \underbrace{\pm e^c}_{\lambda} e^{-A(x)}. \end{aligned}$$

Remarque 12.4.

On choisira la fonction $x \mapsto A(x)$ la plus simple possible. En pratique on commence par calculer une primitive quelconque de $x \mapsto a(x)$ (en calculant une intégrale), puis on choisit pour A la primitive obtenue en supprimant tous les termes constants présents dans celle calculée précédemment, étant entendu que la primitive d'une fonction est définie à une constante près.

Cas des équations avec second membre.

Pour pouvoir résoudre une équation normalisée quelconque, il suffit de savoir en déterminer une solution particulière (on pourra alors appliquer le **Théorème 12.1**). Le plus souvent, on applique la méthode de **variation de la constante**.

Proposition. 12.5. **Variation de la constante**

Soient $a, b : I \to \mathbb{K}$ des fonctions continues. Alors l'équation différentielle $y' + a(x)y = b(x)$ admet sur I une solution particulière de la forme $y_1 : x \mapsto \lambda(x)e^{-A(x)}$, où A est une primitive de a (et où λ est une primitive de $x \mapsto b(x)e^{A(x)}$).

Méthode 12.8. **Variation de la constante**

On recherchera une solution particulière de $y' + a(x)y = b(x)$ sous la forme $y_1(x) = \lambda(x)e^{-A(x)}$, où λ est une fonction dérivable sur I; il suffira alors d'injecter y_1 dans l'équation pour déterminer λ. On aura toujours

$$\lambda'(x) = b(x)e^{A(x)}$$

qu'il faudra donc primitiver.

Exercice 12.1.

Résoudre sur $\mathbb{R}_+^*$ l'équation différentielle $y' + \dfrac{x}{x^2+1}y = \dfrac{1}{(1+\sqrt{x})\sqrt{x^2+1}}$.

On peut encore citer le résultat suivant, qui n'est valable que dans le cas des équations à cœfficients constants, mais qui est parfois plus rapide d'utilisation que la méthode de variation de la constante.

Proposition. 12.6. **Forme des solutions particulières**

Soient $(a, b, \alpha) \in \mathbb{K}^3$, $\omega \in \mathbb{R}$, et P un polynôme à cœfficients dans $\mathbb{K}$.

i. Si $a \neq 0$, l'équation $y' + ay = b$ admet une solution particulière constante, définie par $y_1(x) = \frac{b}{a}$.

ii. L'équation $y' + ay = P(x)e^{\alpha x}$ admet une solution particulière définie par

- ✗ $y_1(x) = Q(x)e^{\alpha x}$ si $\alpha \neq -a$, où Q est un polynôme de même degré que P;
- ✗ $y_1(x) = xQ(x)e^{\alpha x}$ si $\alpha = -a$, où Q est un polynôme de même degré que P.

iii. Les équations $y' + ay = P(x)\cos\omega x$ et $y' + ay = P(x)\sin\omega x$ admettent une solution particulière définie par $y_1(x) = Q(x)\cos\omega x + R(x)\sin\omega x$, où Q, R sont des polynômes de degré inférieur à celui de P.

Exercice 12.2.

Déterminer l'ensemble des solutions des équations différentielles suivantes :

i. $y' + 2y = 3$ *ii.* $y' - y = t^2 + 1$ *iii.* $y' + y = te^t$

Méthode 12.9. **Plan d'étude**

Pour résoudre une équation différentielle linéaire d'ordre un normalisée, on peut procéder comme suit.

1. *Résolution de l'équation homogène.* On obtient la solution générale $y_h : x \mapsto \lambda e^{-A(x)}$.
2. *Recherche d'une solution particulière y_1 de l'équation non homogène.* S'il n'en existe pas d'évidente, on utilise la méthode de variation de la constante pour chercher une solution particulière de la forme $y_1 : x \mapsto \lambda(x)e^{-A(x)}$.
3. **Conclusion.** On écrit que la solution générale de l'équation considérée est $y_h + y_1$. Et si l'énoncé fournit une condition initiale $y(x_0) = \alpha_0$, on détermine l'unique valeur de la constante qui permet de la réaliser en résolvant $y(x_0) = \alpha_0$.

Exercice 12.3.

Déterminer les solutions de l'équation différentielle $y' - 3y = t\ln(t)e^{3t}$.

12.2.2 Cas général: recollement des solutions

Considérons une équation différentielle (E) : $a_1(x)y' + a_0(x)y = b(x)$, à résoudre sur un intervalle $I =]\alpha, \beta[$. Dans le cas où a_1 ne s'annule pas, il suffit de diviser l'équation par $a_1(x)$, puis de résoudre l'équation *normalisée*

$$(E') : \ y' + \frac{a_0(x)}{a_1(x)}y = \frac{b(x)}{a_1(x)}$$

en utilisant les méthodes précédentes.

Mais dans le cas où par exemple a_1 s'annule en un unique réel $x_0 \in]\alpha, \beta[$, il faut procéder différemment: on commence par résoudre (E') sur $]\alpha, x_0[$ et sur $]x_0, \beta[$, puis on essaye de recoller les deux solutions générales ainsi obtenues de façon à obtenir une solution sur $]\alpha, \beta[$ de (E). Considérons pour commencer l'exemple suivant.

Méthode 12.10. **Principe de recollement**

Considérons une équation différentielle (E) : $a_1(x)y' + a_0(x)y = b(x)$, à résoudre sur un intervalle $I =]\alpha, \beta[$ où a_1 s'annule en un unique réel $x_0 \in]\alpha, \beta[$.
Soit alors (E') l'équation

$$y' + \frac{a_0(x)}{a_1(x)}y = \frac{b(x)}{a_1(x)};$$

on appelle $y_{1,\lambda}$ la solution générale de (E') sur $]\alpha, x_0[$ (qui dépend donc d'un paramètre λ), et $y_{2,\mu}$ la solution générale de (E') sur $]x_0, \beta[$ (qui dépend d'un paramètre μ).

- ✗ **Analyse**: soit y une solution de (E) sur I; il existe alors deux scalaires λ, μ tels que

$$(\star)\ y(x) = \begin{cases} y_{1,\lambda}(x) \text{ si } x \in]\alpha, x_0[\\ y_{2,\mu}(x) \text{ si } x \in]x_0, \beta[\end{cases}$$

  Et toujours parce que y est solution de (E) on a:

  i. y est continue et dérivable en x_0;
  ii. y vérifie (E) en x_0.

- ✗ **Synthèse**: réciproquement, on démontre aisément que toute fonction y définie par $(\star)$ et vérifiant *i.* et *ii.* est une solution de (E) sur I.

Pour résoudre (E) sur I il suffit donc de résoudre (E') sur $]\alpha, x_0[$ et $]x_0, \beta[$ (c'est-à-dire déterminer $y_{1,\lambda}$ et $y_{2,\mu}$), puis de déterminer les valeurs de λ, μ pour lesquelles *i.* et *ii.* sont vérifiées.

Exercice 12.4.

Résoudre sur $\mathbb{R}$ l'équation $xy' - \frac{1}{2}y = x$.

Remarque 12.5.

Cette méthode se généralise au cas où a_1 s'annule plusieurs fois sur I : il faut alors effectuer un recollement en chaque point où a_1 s'annule.

12.3 Équations linéaires d'ordre deux

12.3.1 Cas des équations à cœfficients constants

Théorème 12.7. **Résolution des équations homogènes dans $\mathbb{C}$**

Soient a, b, c des complexes tels que $a \neq 0$, $ay'' + by' + cy = 0$ une équation différentielle homogène et $ar^2 + br + c = 0$ l'équation caractéristique associée. On appelle $\mathcal{S}_0(\mathbb{C})$ l'ensemble de ses solutions qui sont à valeurs dans $\mathbb{C}$. On a alors les résultats suivants.

i. Si l'équation caractéristique possède deux racines distinctes r_1 et r_2, alors $\mathcal{S}_0(\mathbb{C})$ est l'ensemble des fonctions de la forme $y : x \mapsto \lambda e^{r_1 x} + \mu e^{r_2 x}$ avec $(\lambda, \mu) \in \mathbb{C}^2$.

ii. Si l'équation caractéristique possède une racine double r, alors $\mathcal{S}_0(\mathbb{C})$ est l'ensemble des fonctions de la forme $y : x \mapsto (\lambda + \mu x)e^{rx}$ avec $(\lambda, \mu) \in \mathbb{C}^2$.

Preuve. On se ramène à une équation différentielle matricielle d'ordre un, qui se résout par réduction matricielle. □

Théorème 12.8. **Résolution des équations homogènes dans $\mathbb{R}$**

Soient a, b, c des réels tels que $a \neq 0$, $ay'' + by' + cy = 0$ une équation différentielle homogène et $ar^2 + br + c = 0$ l'équation caractéristique associée.
On appelle $\mathcal{S}_0(\mathbb{R})$ l'ensemble de ses solutions à valeurs réelles. On a alors les résultats suivants.

i. Si l'équation caractéristique possède deux racines réelles distinctes r_1 et r_2, alors $\mathcal{S}_0(\mathbb{R})$ est l'ensemble des fonctions de la forme $y : x \mapsto \lambda e^{r_1 x} + \mu e^{r_2 x}$ avec $(\lambda, \mu) \in \mathbb{R}^2$.

ii. Si l'équation caractéristique possède deux racines complexes conjuguées distinctes $r_1 = \alpha + i\beta$ et $r_2 = \alpha - i\beta$, alors $\mathcal{S}_0(\mathbb{R})$ est l'ensemble des fonctions de la forme

$$y : x \mapsto e^{\alpha x}(\lambda \cos \beta x + \mu \sin \beta x)$$

avec $(\lambda, \mu) \in \mathbb{R}^2$.
Dans ce cas, les solutions sont également les fonctions de la forme $y : x \mapsto K e^{\alpha x} \cos(\beta x + \varphi)$, avec K, φ réels.

iii. Si l'équation caractéristique possède une racine double r, alors $\mathcal{S}_0(\mathbb{R})$ est l'ensemble des fonctions de la forme $y : x \mapsto (\lambda + \mu x)e^{rx}$ avec $(\lambda, \mu) \in \mathbb{R}^2$.

On a enfin le résultat suivant, qui permet de donner une solution particulière dans un cas particulier.

Proposition. 12.9. **Solution particulière dans le cas exponentielle/polynôme**

Soient a, b, c, z des complexes, et P un polynôme. L'équation différentielle $ay'' + by' + cy = e^{zx}P(x)$ admet une solution particulière de la forme

$$y_1(x) = e^{zx} x^{m(z)} Q(x),$$

où Q est un polynôme de même degré que P, et où $m(z)$ est la multiplicité de z en tant que racine de l'équation caractéristique.

Remarque 12.6.

Ce résultat peut également être utilisé dans le cas des équations du premier ordre: il suffit en effet de prendre $a = 0$, on retrouve alors le **Théorème 12.6**.

Exercice 12.5.

Résoudre l'équation différentielle $y'' - 2y' + y = \cos x$.

12.3.2 Dimension de l'espace des solutions dans les conditions de Cauchy

Théorème 12.10.

Soient a, b, c des fonctions continues sur I, telles que a ne s'annule pas. Alors l'ensemble des solutions de l'équation différentielle $a(x)y'' + b(x)y' + c(x)y = 0$ est un sous-espace vectoriel de $\mathbb{K}^I$ de dimension 2.

Preuve. Le théorème de Cauchy-Lipschitz induit un isomorphisme de l'ensemble des solutions vers $\mathbb{K}^2$. $\square$

Théorème 12.11.

Soient a, b, c, f des fonctions continues sur I, telles que a ne s'annule pas. Alors l'ensemble des solutions de l'équation différentielle $a(x)y'' + b(x)y' + c(x)y = f(x)$ n'est pas vide si on suppose que l'équation homogène associée possède une solution y_1 qui ne s'annule pas sur I.

Preuve. Sous l'hypothèse que y_1 ne s'annule pas on peut faire une variation de la constante: en injectant $y = \lambda y_1$ on voit que λ' vérifie une équation du premier ordre normalisable et dépend donc linéairement d'une constante, par suite $y = \lambda y_1$ dépend linéairement de deux constantes (on retrouve le résultat précédent). $\square$

La *preuve* ci-dessus donne la méthode suivante.

Méthode 12.11.

La méthode de variation de la constante correspond à un changement de fonction inconnue qui permet de ramener la résolution d'une équation d'ordre 2 à celle d'une équation d'ordre 1, partant d'une solution particulière ne s'annulant pas de l'équation homogène d'ordre 2. Pour cette raison, la variation de la constante s'appelle aussi la *méthode d'abaissement de l'ordre*.

12.4 Exemples de résolution lorsque les cœfficients ne sont pas constants

12.4.1 Recherche d'une solution particulière puis variation de la constante

Exercice 12.6.
Résoudre l'équation $x^2y'' - 3xy' + 4y = x^3$ en commençant par déterminer une solution polynomiale de l'équation homogène associée.

12.4.2 Recherche d'une solution développable en série entière

Exemple 12.2.
On considère l'équation différentielle $(\mathcal{E}) : \left(t^2+t\right) y'' + (3t+1)y' + y = 0$.

On cherche une solution $y(t) = \displaystyle\sum_{n=0}^{+\infty} a_n t^n$ développable en série entière au voisinage de 0.

- ✗ **Analyse.** On suppose qu'il existe une telle solution $y(t) = \displaystyle\sum_{n=0}^{+\infty} a_n t^n$ développable en série entière au voisinage de 0 ont on note R le rayon de convergence.
  Pour tout $t \in]-R; R[$ en dérivant terme à terme

$$
\begin{aligned}
&\left(t^2+t\right)\sum_{n=2}^{+\infty} n(n-1)a_n t^{n-2} + (3t+1)\sum_{n=1}^{+\infty} n a_n t^{n-1} + \sum_{n=0}^{+\infty} a_n t^n = 0 \\
\iff & \sum_{n=2}^{+\infty} n(n-1)a_n t^n + \sum_{n=2}^{+\infty} n(n-1)a_n t^{n-1} + 3\sum_{n=1}^{+\infty} n a_n t^n + \sum_{n=1}^{+\infty} n a_n t^{n-1} + \sum_{n=0}^{+\infty} a_n t^n = 0 \\
\iff & \sum_{n=2}^{+\infty} n(n-1)a_n t^n + \sum_{n=1}^{+\infty} (n+1)n a_{n+1} t^n + 3\sum_{n=1}^{+\infty} n a_n t^n + \sum_{n=0}^{+\infty} (n+1)a_{n+1} t^n + \sum_{n=0}^{+\infty} a_n t^n = 0 \\
\iff & \sum_{n=0}^{+\infty} n(n-1)a_n t^n + \sum_{n=0}^{+\infty} (n+1)n a_{n+1} t^n + 3\sum_{n=0}^{+\infty} n a_n t^n + \sum_{n=0}^{+\infty} (n+1)a_{n+1} t^n + \sum_{n=0}^{+\infty} a_n t^n = 0
\end{aligned}
$$

  Par unicité du développement en série entière, on obtient pour tout $n \in \mathbb{N}$,

$$
\begin{aligned}
& n(n-1)a_n + (n+1)n a_{n+1} + 3n a_n + (n+1)a_{n+1} + a_n = 0 \\
\iff & a_n\left(n^2 - n + 3n + 1\right) + a_{n+1}\left(n^2 + n + n + 1\right) = 0 \\
\iff & a_n(n+1)^2 + a_{n+1}(n+1)^2 = 0
\end{aligned}
$$

  On obtient $\forall n \in \mathbb{N}, a_{n+1} = -a_n = (-1)^n a_0$.

- ✗ **Synthèse.** On pose $y(t) = \displaystyle\sum_{n=0}^{+\infty} a_0(-1)^n t^n$.
  Cette série entière a pour rayon de convergence $R = 1$. De plus, y vérifie bien sur $]-1;1[$ l'équation $(\mathcal{E})$ d'après les équivalences écrites dans la partie analyse.

- ✗ **Conclusion.** On obtient ici une fonction usuelle :

$$\forall t \in]-1;1[,\ y_p(t) = \frac{a_0}{1-(-t)} = \frac{a_0}{1+t}$$

  solution particulière de $(\mathcal{E})$.

Remarque 12.7.
Lorsque l'on cherche une solution développable en série entière d'une équation différentielle, on n'oubliera pas de déterminer le rayon de convergence de celle-ci (qui résulte en général directement de la relation de récurrence trouvée entre les coefficients, via la règle de d'Alembert).

Exercice 12.7.

Résoudre le problème de Cauchy $\begin{cases} y'' + xy' + y = 0 \\ y(0) = 1 \\ y'(0) = 0 \end{cases}$ en cherchant la solution sous forme de fonction développable en série entière.

12.4.3 Changement de variable

Méthode 12.12.

On considère l'équation différentielle $(E) : a(t)y'' + b(t)y'(t) + c(t)y = d(t)$ sur I.
On suppose qu'il existe un changement de variable $\varphi : J \to I, x \longmapsto \varphi(x) = t$ de classe $\mathcal{C}^2$ sur J et tel que φ^{-1} est de classe $\mathcal{C}^2$ sur I.
On pose $z(x) = y(\varphi(x)) = y(t)$. On dérive deux fois la fonction z par rapport à x.
On remplace t par $\varphi(x)$ dans (E).
On obtient une équation d'inconnue z et on retrouve la solution générale

$$y(t) = z \circ \varphi^{-1}(t).$$

Exemple 12.3.

Résolvons, sur $]-1;1[$, l'équation différentielle $(E) :\ \left(1 - t^2\right) y'' - ty' + y = 0$.

On va poser $t = \cos(x)$.

On le sait bien : la fonction $\varphi : x \mapsto \cos(x)$ est continue strictement décroissante sur $[0;\pi]$ donc réalise une bijection de $[0;\pi]$ sur $[-1;1]$. De plus, φ est de classe $\mathcal{C}^2$ sur $[0;\pi]$.
La bijection réciproque $\varphi^{-1} : t \mapsto \arccos(t) = x$ est de classe $\mathcal{C}^2$ sur $]-1;1[$.
On pose pour $x \in]0;\pi[$: $z(x) = y(\varphi(x)) = y(\cos(x)) = y(t)$.
Pour tout $x \in]0;\pi[$:

- ✗ $z'(x) = -\sin(x)y'(\cos x)$
- ✗ $z''(x) = \sin^2(x)y''(\cos x) - \cos(x)y'(\cos x)$

La fonction $t \mapsto y(t)$ est solution sur $]-1;1[$ de (E) si et seulement si pour tout $x \in]0;\pi[$:

$$\underbrace{\left(1 - \cos^2(x)\right) y''(\cos(x)) - \cos(x)y'(\cos(x))}_{=z''(x)} + \underbrace{y(\cos x)}_{=z(x)} = 0.$$

Ainsi, z est solution sur $]0;\pi[$ de l'équation $z'' + z = 0$.
L'équation caractéristique $X^2 + 1 = 0$ a pour racine $X = \pm i = 0 \pm 1 \times i$.
Il existe donc $(A, B) \in \mathbb{R}^2$ tel que $z(x) = e^{0x}(A\cos(1 \times x) + B\sin(1 \times x)) = A\cos(x) + B\sin(x)$.
La solution générale de (E) sur $]-1;1[$ est alors :

$$y(t) = z \circ \varphi^{-1}(t) = A\cos(\arccos(t)) + B\sin(\underbrace{\arccos(t)}_{\in]0;\pi[}) = At + B\sqrt{1 - t^2}, \qquad (A, B) \in \mathbb{R}^2.$$

Exercice 12.8.

Résoudre l'équation $x^2y'' + xy' - 4y = 4x^2$ sur $\mathbb{R}_+^*$ en posant $x = e^t$.

13 Isométries d'un espace euclidien

Dans tout le chapitre, E désigne un espace vectoriel euclidien de dimension $n \geq 2$ dont le produit scalaire sera noté $(\cdot|\cdot)$.
On notera sans distinction (mais avec discernement) $\langle \cdot, \cdot \rangle$ le produit scalaire canonique sur $\mathbb{R}^n$ et sur $\mathcal{M}_{n,1}(\mathbb{R})$.
Si $X, Y \in \mathcal{M}_{n,1}(\mathbb{R})$, on rappelle que $\langle X, Y \rangle = X^\top \cdot Y$.
La norme associé à tout produit scalaire sera toujours notée $\|\cdot\|$.

13.1 Matrices orthogonales

13.1.1 Généralités

Définition 13.1. **Matrices orthogonales, groupe orthogonal**

Soit $M \in \mathcal{M}_n(\mathbb{R})$. On dit que M est une matrice **orthogonale** d'ordre n si :

$$M \times M^\top = M^\top \times M = I_n.$$

On notera $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales d'ordre n.
$\mathcal{O}_n(\mathbb{R})$ est appelé le **groupe orthogonal d'ordre n**.

Proposition. 13.1. **Propriétés immédiates**

On a les résultats suivants.

i. Si $M \in \mathcal{O}_n(\mathbb{R})$, alors $\det(M) = \pm 1$.
ii. $M \in \mathcal{O}_n(\mathbb{R})$,, alors M est inversible et $M^{-1} = M^\top$.
iii. $M \in \mathcal{O}_n(\mathbb{R})$ si et seulement si $M \times M^\top = I_n$ si et seulement si $M^\top \times M = I_n$
iv. $I_n \in \mathcal{O}_n(\mathbb{R})$.
v. $\mathcal{O}_n(\mathbb{R})$ est stable par produit matriciel.
vi. Si $M \in \mathcal{O}_n(\mathbb{R})$, alors $M^{-1} \in \mathcal{O}_n(\mathbb{R})$ (ou encore $\mathcal{O}_n(\mathbb{R})$ est stable par passage à l'inverse).

Remarque 13.1.

Les propriétés *iv.*, *v.* et *vi.* montrent que $\mathcal{O}_n(\mathbb{R})$ possède une *structure de groupe* mais cette notion n'est pas au programme en **PT**.

Définition 13.2. **Groupe spécial orthogonal**

On introduit et note $\mathcal{SO}_n(\mathbb{R}) = \{M \in \mathcal{O}_n(\mathbb{R}) : \det(M) = 1\}$, et cet ensemble est appelé le **groupe spécial orthogonal**.

Proposition. 13.2.

Soit $M \in \mathcal{M}_n(\mathbb{R})$. On note $C_1, \ldots, C_n$ les vecteurs colonnes de M et $L_1, \ldots, L_n$ ses vecteurs lignes. Les assertions suivantes sont équivalentes.

i. $M \in \mathcal{O}_n(R)$.
ii. $C_1, \ldots, C_n$ est une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$;
iii. $L_1, \ldots, L_n$ est une base orthonormée de $\mathcal{M}_{1,n}(\mathbb{R})$.

Preuve. Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R})$.

$$M \in \mathcal{O}_n(\mathbb{R}) \iff M^\top M = I_n \quad \iff \quad \forall i, j \in [\![1, n]\!], \quad \left(M^\top M\right)_{i,j} = \delta_{i,j}$$

$$\iff \forall i, j \in [\![1, n]\!], \quad \sum_{k=1}^{n} m_{k,i} m_{k,j} = C_i^\top C_j = \delta_{i,j}$$

Le même calcul appliqué à $M^\top$ donne le résultat pour les lignes de M. $\square$

Exemple 13.1.

Les matrices $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ sont orthogonales.

Le résultat suivant montre que les matrices orthogonales sont les matrices de passage entre deux bases orthonormées.

Corollaire 13.3.
Soient $\mathcal{B}$ une base orthonormée de E, $\mathcal{B}'$ une base de E et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.
Alors, $\mathcal{B}'$ est orthonormée si et seulement si P est une matrice orthogonale.

Méthode 13.13.
Si P est une matrice de changement de bases orthonormées, alors son inverse est $P^{-1} = P^\top$ et donc il s'obtient sans calcul.

13.1.2 Orientation d'un espace euclidien

Définition 13.3. **Orientation de E**

i. Deux bases orthonormées $\mathcal{B}, \mathcal{B}'$ de E sont dites de **même orientation** (ou de même sens) si la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est de déterminant 1.

ii. Deux bases orthonormées $\mathcal{B}, \mathcal{B}'$ de E sont dites d' **orientation contraire** (ou de sens contraire) si la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est de déterminant -1.

iii. **Orienter** E, c'est choisir une base orthonormée $\mathcal{B}$ de E comme référence. On dira alors que toute base de même sens est **directe**, et que toute base de sens contraire est **indirecte**.

Remarque 13.2.
Dans le plan, on choisit le plus souvent le sens trigonométrique, et dans l'espace le sens *des trois doigts de la main droite*. Mais ce n'est pas une obligation.

Exemple 13.2.
On oriente l'espace euclidien $\mathbb{R}^2$ par la base canonique (e_1, e_2).
Les bases (e_2, e_1) et $(-e_1, e_2)$ sont indirectes.
Les bases $(e_2, -e_1)$ et $(-e_1, -e_2)$ sont directes. La base $\left(\frac{1}{\sqrt{2}}(e_1 + e_2),\ \frac{1}{\sqrt{2}}(-e_1 + e_2)\right)$ est directe.

Théorème 13.4.
Soit H un hyperplan de E $(h_1, \dots, h_{n-1})$ une base orthonormée de H et u un vecteur unitaire dirigeant $H^\perp$.
Alors, $(h_1, \dots, h_{n-1}, u)$ et $(h_1, \dots, h_{n-1}, -u)$ sont deux bases orthonormées de E de sens contraires.

Remarque 13.3.
On en déduit que pour définir un angle orienté dans un hyperplan, il faut d'abord orienter l'hyperplan en choisissant un vecteur normal.

Méthode 13.14. **Orientation de l'espace euclidien $\mathbb{R}^3$**

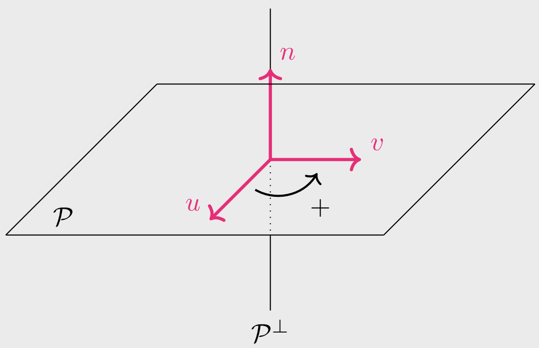
- ✗ Orienter une droite de $\mathbb{R}^3$ consiste à choisir un vecteur unitaire (de norme 1) de cette droite.
- ✗ Orienter un plan P de $\mathbb{R}^3$ consiste à choisir une base orthonormée de $\mathcal{P}$.

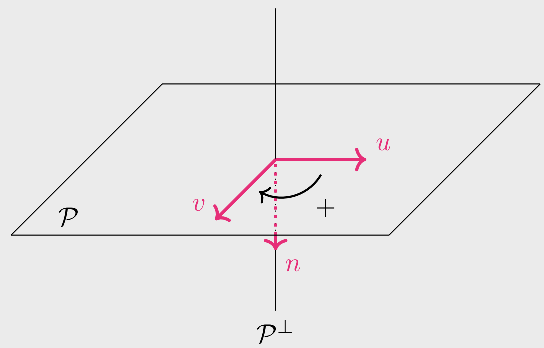
De manière équivalente il suffit de choisir un vecteur unitaire n normal à P (et une orientation de $\mathbb{R}^3$).
En effet ce vecteur induit une orientation sur le plan $\mathcal{P} = \mathrm{Vect}(n)^{\perp}$. Pour toutes bases orthonormées $(u, v), (u', v')$ de P, on a en effet:

$$\det_{(u,v)}(u', v') = \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \begin{vmatrix} \alpha & \beta & 0 \\ \gamma & \delta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \det_{(u,v,n)}(u', v', n)$$

Par conséquent les bases $(u, v), (u', v')$ de $\mathcal{P}$ ont la même orientation si et seulement si les bases $(u, v, n), (u', v', n)$ de $\mathbb{R}^3$ ont la même orientation. On décrète que (u, v) est une base directe de $\mathcal{P}$ si (u, v, n) est une base directe de $\mathbb{R}^3$.

Modifier le choix du vecteur unitaire normal à $\mathcal{P}$ modifie l'orientation de $\mathcal{P}$.





Définition 13.4. **Produit mixte**

Soit E un espace euclidien orienté par le choix d'une base orthonormée directe $\mathcal{B} = (e_1, \ldots, e_n)$. Soient $x_1, \ldots, x_n$ des vecteurs de E. On appelle **produit mixte** de ces n vecteurs le nombre :

$$[x_1, \ldots, x_n] = \det_{\mathcal{B}}(x_1, \ldots, x_n).$$

Il est indépendant de la base orthonormée directe choisie.

On vérifie que le produit mixte est en effet indépendant de la b.o.n directe choisie.
Soient $\mathcal{B}'$ est une autre base orthonormée directe de E et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, qui vérifie donc $\det(P) = 1$. On a alors

$$\det_{\mathcal{B}}(x_1, ..., x_n) = \det(P)\det_{\mathcal{B}'} = \det_{\mathcal{B}'}.$$

Remarque 13.4.

- ✗ On a mentionné dans le **Chapitre 6** que le produit mixte $[x, y]$ dans $\mathbb{R}^2$peut être interprété comme comme l'aire algébrique du parallélogramme engendré par les vecteurs (x, y).
- ✗ De même, on a interprété le produit mixte $[x, y, z]$ dans $\mathbb{R}^3$ comme le volume algébrique du parallélépipède engendré par les vecteurs (x, y, z).

13.1.3 Produit vectoriel en dimension 3

Dans tout ce paragraphe on considère un espace euclidien E de dimension 3.
Soit $\mathcal{B} = (e_1, e_2, e_3)$ une base orthonormée directe de E.

Définition 13.5. **Produit vectoriel**

Soient $x, y \in E$ de coordonnées $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$.
On définit le **produit vectoriel** $x \wedge y$ des vecteurs x, y le vecteur de coordonnées dans $\mathcal{B}$:

$$\begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

Proposition. 13.5. **Propriétés du produit vectoriel et lien avec le produit mixte**

On a les résultats suivants :

i. $x \wedge y$ est orthogonal au plan $\mathrm{Vect}(x, y)$: $(x \wedge y \mid x) = 0$ et $(x \wedge y \mid y) = 0$.
ii. $x \wedge y = 0_E$ si et seulement si (x, y) est liée.
iii. $(x \wedge y \mid z) = \det(x, y, z) = [x, y, z]$.

Remarque 13.5.

Dans $\mathbb{R}^3$, on peut de manière équivalente définir le produit vectoriel $x \wedge y$ comme l'unique vecteur directement orthogonal au plan $\mathrm{Vect}(x, y)$ de norme $\|x\| \cdot \|y\| \cdot |\sin(x, y)|$.
On retrouve alors géométriquement l'interprétation du produit produit mixte $[x, y, z] = (x \wedge y \mid z)$ comme le volume orienté du parallélépipède construit sur (x, y, z).

Théorème 13.6. **Base orthonormée directe de E**

Soit (x, y) une famille orthonormale de E. Alors $(x, y, x \wedge y)$ est une base orthonormée directe de E.

13.2 Théorème spectral : réduction des matrices symétriques réelles

On rappelle que $\mathcal{S}_n(\mathbb{R})$ désigne le sous-espace vectoriel des matrices symétriques de $\mathcal{M}_n(\mathbb{R})$.

Proposition. 13.7.

Soient $A \in \mathcal{S}_n(\mathbb{R})$, et $\lambda \neq \mu$ deux valeurs propres de A. Alors les sous-espaces propres E_λ et E_μ associés sont orthogonaux.

Preuve. Soient $\lambda \neq \mu$ deux valeurs propres de A et $(X, Y) \in E_\lambda \times E_\mu$. Alors

$$\lambda\mu\langle X, Y\rangle = \langle \lambda AX, AY\rangle = (AX)^\top AY = X^\top A^\top AY = X^\top A^2 Y = \mu^2\langle X, Y\rangle.$$

Par symétrie, on obtient

$$\lambda\mu\langle X, Y\rangle = \mu^2\langle X, Y\rangle = \lambda^2\langle X, Y\rangle.$$

Comme λ et μ sont distinctes, elles ne peuvent être toutes deux nulles. Supposons, sans perte de généralité que $\mu \neq 0$.

- ✗ Si $\lambda = 0$ alors $0 = \lambda\mu\langle X, Y\rangle = \mu^2\langle X, Y\rangle$ donc $\langle X, Y\rangle = 0$ et X et Y sont bien orthogonaux.
- ✗ Si $\lambda \neq 0$ alors

$$0 = \lambda\mu\langle X, Y\rangle = \mu^2\langle X, Y\rangle \iff 0 = \mu(\lambda - \mu)\langle X, Y\rangle$$

ce qui implique, car $\lambda \neq \mu$, que $\langle X, Y\rangle = 0$ et X et Y sont encore orthogonaux.

□

Théorème 13.8. **Théorème spectral**

Toute matrice symétrique réelle est diagonalisable dans une base orthonormée.

Preuve. **Résultat admis**.

Néanmoins, on ne résiste pas à l'envie de proposer une preuve dans le cas particulier et simple où $n = 2$. Pour le plaisir. Soit $A \in \mathcal{S}_2(\mathbb{R})$ telle que

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

Observons que A est toujours diagonalisable. En effet, son polynôme caractéristique vaut

$$\chi_A(\lambda) = \lambda^2 - (a + c)\lambda + \left(ac - b^2\right).$$

Ce dernier a pour discriminant

$$\Delta = (a+c)^2 - 4\left(ac - b^2\right) = (a-c)^2 + 4b^2 \geq 0.$$

Ainsi, A admet toujours au moins une valeur propre réelle. Distinguons deux cas.

✗ Si $\Delta = (a-c)^2 + 4b^2 = 0$. Ceci signifie que $a = c$ et $b = 0$, et donc que A est en fait la matrice aI_2 qui est déjà diagonale, on peut évidemment l'écrire comme $A = I_2 \cdot A \cdot I_2^\top$.

✗ Si $\Delta > 0$. Dans ce cas, χ_A possède deux racines distinctes et donc A admet deux valeurs propres distinctes. Elle est donc diagonalisable.
Les deux sous-espaces propres sont de dimension 1. Il suffit de normaliser un vecteur propre (non nul donc) associé à chaque valeur propre. Par le résultat qui précède, la famille obtenue par concaténation des deux vecteurs normalisés est orthogonale (et libre) et forme donc une base orthonormale de vecteurs propres de A.

Cette preuve ne s'étend hélas pas au cas $n > 2$ où cela devient plus compliqué. □

Remarque 13.6. **Réciproque**

Observons que la réciproque est vraie et facile à montrer ; si $A \in \mathcal{M}_n(\mathbb{R})$ est diagonalisable dans une base orthonormée, il existe une matrice orthogonale P et une matrice diagonale D telles que $A = PDP^\top$. Mais alors, D étant bien entendu symétrique en tant que matrice diagonale, on a

$$A^\top = \left(PDP^\top\right)^\top = \left(P^\top\right)^\top D^\top P^\top = PDP^\top = A.$$

Remarque 13.7. **À retenir**

Si A est une matrice symétrique réelle:

- ✗ toutes les valeurs propres de A sont réelles ;
- ✗ il existe une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ formée de vecteurs propres de A ;
- ✗ il existe donc une matrice diagonale D et une matrice orthogonale P telle que $A = PDP^{-1} = PDP^\top$.

On dit que la matrice A est *orthogonalement diagonalisable*.

Méthode 13.15. **Diagonalisation en B.O.N**

En pratique, pour diagonaliser une matrice symétrique en base orthonormée, on procède comme suit :

i. on dit que la matrice est symétrique **réelle** donc elle est diagonalisable;
ii. on cherche ses valeurs propres ;
iii. On détermine une B.O.N $\mathcal{B}_i$ de chaque sous-espace propre (éventuellement à l'aide du procédé de Gram-Schmidt)
iv. La base $\mathcal{B}$ obtenue par concaténation des bases $\mathcal{B}_i$ est une B.O.N de $\mathcal{M}_{n,1}(\mathbb{R})$.
v. La matrice de passage P de la base canonique vers cette base est orthogonale et on a

$$P^{-1}AP = P^\top AP = \mathrm{diag}(\lambda_1, ..., \lambda_1, ..., \lambda_\ell, ..., \lambda_\ell),$$

où chaque λ_i est répété un nombre de fois correspondant à sa multiplicité.

Exercice 13.1.

Diagonaliser en base orthonormée la matrice $M = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$.

13.3 Isométries vectorielles

Définition 13.6. **Isométrie vectorielle**

Soit $f \in \mathcal{L}(E)$. On dit que f est une **isométrie vectorielle** si :

$$\forall x \in E,\ \|f(x)\| = \|x\|,$$

c'est à dire si f préserve la norme.
On notera $\mathcal{O}(E)$ l'ensemble des isométries vectorielles sur E.

Exemple 13.3.

- ✗ Id_E est une isométrie.
- ✗ Une projection orthogonale est une isométrie si et seulement si c'est l'identité.

Exercice 13.2.
Dans chaque cas, justifier que f est une isométrie de $\mathbb{R}^2$ et déterminer son spectre :

i. $f = -\mathrm{id}_{\mathbb{R}^2}$; *ii.* $f(x, y) = (-y, x)$; *iii.* $f(x, y) = (y, x)$.

Proposition. 13.9.
Soit $f \in \mathcal{L}(E)$. Alors, $f \in \mathcal{O}(E)$ si et seulement si f préserve le produit scalaire, c'est à dire si et seulement si

$$\forall x, y \in E,\ (f(x)|\, f(y)) = (x|\, y).$$

Preuve. Soit $f \in \mathcal{L}(E)$.

- ✗ Supposons que f préserve le produit scalaire. Soit $x \in E$. $\|f(x)\|^2 = (f(x)|\, f(x)) = (x|\, x) = \|x\|^2$ donc f préserve la norme et $f \in \mathcal{O}(E)$.

- ✗ Supposons que $f \in \mathcal{O}(E)$. Soient $x, y \in E$. On utilise l'identité de polarisation :

$$\begin{aligned}(f(x)|\, f(y)) &= \frac{1}{4}\left(\|f(x)+f(y)\|^2 - \|f(x)-f(y)\|^2\right) = \frac{1}{4}\left(\|f(x+y)\|^2 - \|f(x-y)\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = (x|\, y)\end{aligned}$$

  f préserve donc le produit scalaire.

□

Théorème 13.10.
Une symétrie est une isométrie si et seulement si elle est orthogonale.

Preuve. Soit $s \in \mathcal{L}(E)$ une symétrie.

- ✗ Si s est orthogonale, alors $\mathrm{Ker}(s - \mathrm{id})$ et $\mathrm{Ker}(s + \mathrm{id})$ sont orthogonaux. Soit alors $x \in E$ rel que $x = y + z$ avec $(y, z) \in \mathrm{Ker}(s - \mathrm{id}) \times \mathrm{Ker}(s + \mathrm{id})$ et $y \perp z$. Comme $s(x) = y - z$, on a, par Pythagore

$$\|s(x)\|^2 = \|y - z\|^2 = \|y\|^2 + \|-z\|^2 = \|y\|^2 + \|z\|^2 = \|y + z\|^2 = \|x\|^2$$

  et donc $s \in \mathcal{O}(E)$.
- ✗ Supposons que $s \in \mathcal{O}(E)$. Montrons que $\mathrm{Ker}(s - \mathrm{id})$ et $\mathrm{Ker}(s + \mathrm{id})$ sont orthogonaux. Soit $(y, z) \in \mathrm{Ker}(s - \mathrm{id}) \times \mathrm{Ker}(s + \mathrm{id})$. Alors, comme s préserve le produit scalaire,

$$(y|\, z) = (s(y)|\, s(z)) = (y|\, -z) = -(y|\, z)$$

  donc $(y|\, z) = 0$ et s est bien orthogonale.

□

Théorème 13.11. **Structure de groupe**

i. $\mathrm{Id}_E \in \mathcal{O}(E)$.
ii. Si $f \in \mathcal{O}(E)$ alors f est bijective et $f^{-1} \in \mathcal{O}(E)$.
iii. $\mathcal{O}(E)$ est stable pour la composition d'applications.

Le résultat suivant caractérise les isométries vectorielles, et relie cette notion à celle de matrice orthogonale.

Proposition. 13.12.

Soit $f \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes.

i. f est une isométrie ;
ii. f conserve le produit scalaire ;
iii. f transforme toute base orthonormée de E en une base orthonormée ;
iv. f transforme au moins une base orthonormée de E en une base orthonormée;
v. la matrice de f dans au moins une base orthonormée est orthogonale;
vi. la matrice de f dans toute base orthonormée est orthogonale.

Remarque 13.8.

Ces derniers résultats justifient l'existence d'une autre terminologie ; on appelle aussi les isométries vectorielles des **automorphismes orthogonaux.**

Méthode 13.16.

En pratique, pour montrer qu'une application linéaire est une isométrie, on calcule sa matrice M dans une base orthonormée et on montre que M est orthogonale, en vérifiant que ses vecteurs colonnes forment une famille orthonormée.

Proposition. 13.13. **Déterminant d'une isométrie**

Si $f \in \mathcal{O}(E)$ alors $\det(f) = \pm 1$.

Remarque 13.9.

Attention ! La réciproque du résultat ci-dessus est fausse (il suffit par exemple de considérer une symétrie non orthogonale).

Définition 13.7. **Groupe spécial orthogonal**

On introduit et on note $\mathcal{SO}(E) = \{f \in \mathcal{O}(E) :\ \det(f) = 1\}$. Cet ensemble est appelé le **groupe spécial orthogonal**, il est aussi parfois noté $\mathcal{O}^+(E)$.
Les isométries de $\mathcal{SO}(E)$ sont dites **directes** (ou positives), celles de déterminant -1 sont dites **indirectes** (ou négatives).

Théorème 13.14.

i. Une isométrie est directe si et seulement si elle transforme les bases orthonormées directes en bases orthonormées directes (on dit que f **conserve l'orientation**).
ii. La composée de deux isométries directes (*resp.* indirectes) est une isométrie directe.
iii. Si f est une isométrie, alors f et f^{-1} ont même sens.

Exemple 13.4.

On voit aisément sur un dessin que, dans le plan, les rotations sont des isométries directes et les symétries orthogonales sont des isométries indirectes.

Théorème 13.15. **Éléments propres d'une isométrie**

Soit f une isométrie de E.

i. $\mathrm{Sp}_{\mathbb{R}}(f) \subset \{-1, 1\}$.
ii. Si -1 et 1 sont des valeurs propres de f, alors E_1 et E_{-1} sont orthogonaux.

Preuve. Soit $\lambda \in \mathrm{Sp}_{\mathbb{R}}(f)$. Alors il existe $x \neq 0$ tel que $f(x) = \lambda x$, mais alors $\|x\| = \|f(x)\| = \|\lambda x\| = |\lambda| \cdot \|x\|$ donc $|\lambda| = 1$. On a bien $\mathrm{Sp}_{\mathbb{R}}(f) \subset \{-1, 1\}$.
Si 1 et -1 sont toutes deux valeurs propres de f, considérons deux vecteurs propres x et y respectivement associés à 1 et -1. Alors, $(x|\,y) = (f(x)|\,f(y)) = (x|\,-y) = -(x|y)$ et donc $(x|\,y) = 0$ et les deux sous-espaces propres sont bien orthogonaux. □

Théorème 13.16.
Soit f une isométrie, et F un sous-espace vectoriel de E stable par f. Alors $F^{\perp}$ est également stable par f.

Preuve. Supposons F stable par $f \in \mathcal{O}(E)$. Observons que $f(F) = F$ par inclusion évidente et égalité des dimensions. Soit $y \in F^{\perp}$. Pour tout $z \in F$, il existe $x \in F$ tel que $z = f(x)$ et donc, $(f(y)|\, z) = (f(y)|\, f(x)) = (y|\, x) = 0$.
Ainsi, $f(y) \in F^{\perp}$. □

Corollaire 13.17.
Soit f une isométrie de E.
Si $\lambda \in \mathrm{Sp}(f)$, alors ${E_\lambda}^{\perp}$ est stable par f, et l'endomorphisme induit par f sur ${E_\lambda}^{\perp}$ est encore une isométrie.

13.4 Isométries en dimension 2

Dans cette section, on suppose que $n = 2$, notamment E et $\mathbb{R}^2$ sont alors isomorphes.
Sauf mention explicite du contraire, on considère $\mathcal{B} = (e_1, e_2)$ une base orthonormée de $\mathbb{R}^2$.

On cherche alors à **identifier** et **interpréter géométriquement** les éléments de $\mathcal{O}(\mathbb{R}^2)$ et $\mathcal{SO}(\mathbb{R}^2)$.

Pour ce faire, commençons par observer les choses suivantes.
Soient $f \in \mathcal{O}\left(\mathbb{R}^2\right)$ et $M = \mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Tout d'abord

$$\begin{aligned} M^\top M = I_2 \quad & \Longleftrightarrow \quad \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} = I_2 \\ & \Longleftrightarrow \quad \begin{cases} a^2+c^2 & = & 1 \\ b^2+d^2 & = & 1 \\ ab+cd & = & 0 \end{cases} \end{aligned}$$

Or,

$$a^2 + c^2 = 1 \Longleftrightarrow \exists \theta \in \mathbb{R} : \quad \begin{cases} a = \cos\theta \\ c = \sin\theta \end{cases}$$

et

$$b^2 + d^2 = 1 \Longleftrightarrow \exists \varphi \in \mathbb{R} : \quad \begin{cases} b = \cos\varphi \\ d = \sin\varphi \end{cases} .$$

Il suit qu'on peut d'ores et déjà écrire $M = \begin{pmatrix} \cos\theta & \cos\varphi \\ \sin\theta & \sin\varphi \end{pmatrix}$.
De plus, les formules usuelles de trigonométrie donnent $ab + cd = \cos\theta\cos\varphi + \sin\theta\sin\varphi = \cos(\theta - \varphi)$.
Donc $ab + cd = 0 \Longleftrightarrow \theta - \varphi = \frac{\pi}{2} \quad [\pi] \Longleftrightarrow \varphi = \theta + \varepsilon\frac{\pi}{2} \quad [2\pi] \quad$ avec $\quad \varepsilon = \pm 1$.

$$\cos\varphi = \cos\left(\theta + \varepsilon\frac{\pi}{2}\right) = -\varepsilon\sin\theta; \quad \sin\varphi = \sin\left(\theta + \varepsilon\frac{\pi}{2}\right) = \varepsilon\cos\theta$$

D'où

$$M \in \mathcal{O}_2(\mathbb{R}) \iff \exists \theta \in \mathbb{R} \quad \text{tel que} \quad M = \begin{pmatrix} \cos\theta & -\varepsilon\sin\theta \\ \sin\theta & \varepsilon\cos\theta \end{pmatrix} \quad \text{avec} \quad \varepsilon = \det(M) \in \{-1, 1\}.$$

Théorème 13.18. **Isométries du plan**

$$M \in \mathcal{O}_2(\mathbb{R}) \iff \exists \theta \in \mathbb{R} \quad \text{tel que} \quad M = \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix}.$$

Plus précisément,

i. $M \in \mathcal{SO}_2(\mathbb{R}) \iff \det(M) = 1 \iff \exists \theta \in \mathbb{R} \quad \text{tel que} \quad M = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$

Auquel cas, $\chi_M(\lambda) = \lambda^2 - 2\cos(\theta) + 1$ et :

✗ Si $\theta \neq 0[\pi]$, on a $\mathrm{Sp}_{\mathbb{R}}(M) = \emptyset$.
✗ Si $\theta = 0[\pi]$, on on a $M = \pm I_2$.

ii. $\det(M) = -1 \iff \exists \theta \in \mathbb{R} \quad \text{tel que} \quad M = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$

Auquel cas, $M \in \mathcal{S}_2(\mathbb{R})$ est symétrique réelle, donc diagonalisable dans une base orthonormée directe avec $\mathrm{Sp}_{\mathbb{R}}(M) = \{-1, 1\}$. Il existe donc $P \in \mathcal{SO}_2(\mathbb{R})$ telle que

$$M = P \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} P^\top.$$

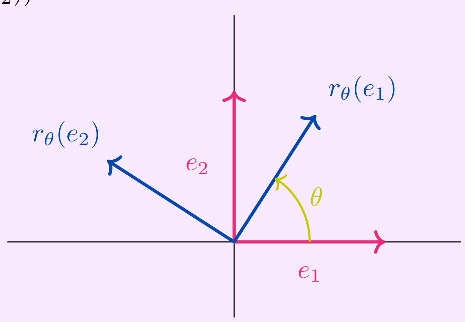
Définition 13.8. **Rotation d'angle θ**

Soit $\theta \in \mathbb{R}$. On appelle **rotation de E d'angle θ**, l'endomorphisme de E, noté r_θ, dont la matrice dans la base orthonormée directe $\mathcal{B}$ est

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Remarque 13.10. **Angle orienté**

Si $v = r_\theta(u)$, on dit que la mesure de l'angle orienté (u, v) est θ (modulo 2π).
Notamment $\theta = (e_1, r_\theta(e_1)) = (e_2, r_\theta(e_2))$.



Proposition. 13.19. **Propriétés des rotations du plan**

Soient $\theta, \theta' \in \mathbb{R}$. Alors, on a :

i. $r_\theta = r_{\theta'} \iff \theta = \theta'[2\pi]$
ii. $r_\theta \circ r_{\theta'} = r_{\theta'} \circ r_\theta = r_{\theta+\theta'}$;
iii. r_θ est bijective, de bijection réciproque $r_\theta^{-1} = r_{-\theta}$.

Remarque 13.11.

La matrice d'un changement de B.O.N directe étant celle d'une rotation d'angle θ', un changement de B.O.N. directe ne change pas la matrice d'une rotation :

$$P^\top M P = R_{-\theta'} R_\theta R_{\theta'} = R_{-\theta'+\theta+\theta'} = M.$$

Définition 13.9. **Réflexion du plan**

Soit $\mathcal{B} = (e_1, e_2)$ une base orthonormée directe du plan. Soit $\mathcal{D}$ une droite (vectorielle) de $\mathbb{R}^2$ (c'est à dire un hyperplan de $\mathbb{R}^2$).
On rappelle **réflexion par rapport à** $\mathcal{D}$ la symétrie orthogonale par rapport à $\mathcal{D}$, que l'on note $S_{\mathcal{D}}$.

Plus précisément, si $\theta \in \mathbb{R}$ est tel que $D = \operatorname{Vect}\left(\cos\left(\frac{\theta}{2}\right) e_1 + \sin\left(\frac{\theta}{2}\right) e_2\right)$, alors

$$\operatorname{Mat}_{\mathcal{B}}(s_{\mathcal{D}}) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

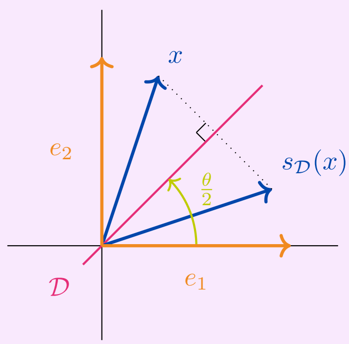
C'est une isométrie indirecte.

Remarque 13.12.

Soit $s_{\mathcal{D}}$ la réflexion dont la matrice est $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ dans la b.o.n directe $\mathcal{B} = (e_1, e_2)$.

En notant $u = \cos\left(\dfrac{\theta}{2}\right) e_1 + \sin\left(\dfrac{\theta}{2}\right) e_2$ et en formant une base $\mathcal{B}' = (u, v)$ orthonormée directe, la matrice de la réflexion d'axe $\mathcal{D} = \operatorname{Vect}(u)$ dans la base $\mathcal{B}'$ est

$$\operatorname{Mat}_{\mathcal{B}'}(s_{\mathcal{D}}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$



Attention ! Contrairement au cas des rotations, la matrice d'une réflexion dépend de la b.o.n directe dans laquelle on la calcule.

Théorème 13.20. **Classification des isométrie du plan**

On peut établir le tableau de synthèse suivant.

Nature de l'isométrie	Déterminant	Spectre réel	s.e. propres	Matrice dans une b.o.n directe
Id_E	1	$\{1\}$	$E_1 = \mathbb{R}^2$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$-\operatorname{Id}_E$	1	$\{-1\}$	$E_{-1} = \mathbb{R}^2$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
r_θ : Rotation d'angle θ ($\theta \neq 0, \theta \neq \pi$)	1	$\emptyset$	/	$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$
$s_{\mathcal{D}}$: Réflexion d'axe $\mathcal{D} = \operatorname{Vect}(u)$	-1	$\{-1, 1\}$	$E_1 = \operatorname{Vect}(u)$, $E_{-1} = E_1^{\perp}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Proposition. 13.21. **Composée de deux isométries d'un plan euclidien**

On a les résultats suivants.

i. la composée de deux rotations d'un plan euclidien est une rotation;
ii. la composée de deux réflexions d'un plan euclidien est une rotation;
iii. la composée d'une rotation et d'une réflexion d'un plan euclidien est une réflexion.

Exercice 13.3.

Montrer que la composée d'une rotation et d'une réflexion est une réflexion.

Méthode 13.17. **Nature d'une isométrie du plan**

Pour déterminer la nature et les caractéristiques d'une isométrie f, on commence par calculer la dimension de $E_1(f) = \mathrm{Ker}(f - \mathrm{Id}_E)$:

- ✗ Si $\dim \mathrm{Ker}(f - \mathrm{Id}_E) = 2$, alors $f = \mathrm{Id}_E$.
- ✗ Si $\dim \mathrm{Ker}(f - \mathrm{Id}_E) = 1$, alors f est la réflexion par rapport à $\mathcal{D} = \mathrm{Ker}(f - \mathrm{Id}_E)$.
- ✗ Si $\dim \mathrm{Ker}(f - \mathrm{Id}_E) = 0$, alors f est une rotation. Pour déterminer l'angle de cette rotation, il suffit de lire celui-ci sur la matrice de f dans une base orthonormée (ou d'utiliser la relation $\mathrm{Tr}(f) = 2\cos\theta$).

Exercice 13.4.

Dans chaque cas, identifier $f \in \mathcal{L}\left(\mathbb{R}^2\right)$ canoniquement associé à M :

$$i.\quad M = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix}, \qquad ii.\quad M = \frac{1}{5}\begin{pmatrix} 4 & -3 \\ -3 & -4 \end{pmatrix}.$$

13.5 Isométries en dimension 3

Dans cette section, on suppose que $n = 3$, notamment E et $\mathbb{R}^3$ sont alors isomorphes.
Sauf mention explicite du contraire, on considère $\mathcal{B} = (e_1, e_2, e_3)$ une base orthonormale de $\mathbb{R}^3$.

On cherche alors à **identifier** et **interpréter géométriquement** les éléments de $\mathcal{O}(\mathbb{R}^3)$ et $\mathcal{SO}(\mathbb{R}^3)$.

On procède de la même manière que dans la section précédente.
Soient $f \in \mathcal{O}\left(\mathbb{R}^3\right)$ (avec $f \notin \{\mathrm{Id}_E, -\mathrm{Id}_E\}$). Comme le polynôme caractéristique de f est degré 3, il admet au moins une racine réelle qui vaut 1 ou -1 (d'après le **Théorème 13.15**).
Le sous-espace propre associe $E_{\pm 1}$ est stable par f , ainsi que sont orthogonal, et l'endomorphisme $\tilde{f}$ induit par f sur $E_{\pm 1}^{\perp}$ est encore une isométrie vectorielle.
On peut montrer que, nécessairement, E_1 ou E_{-1} est de dimension 1 et donc $\tilde{f}$ (sur l'orthogonal) est une isométrie plane classifiée dans la section précédente (d'après le **Théorème 13.17**).
Il existe donc, par concaténation de bases orthonormées de $E_{\pm 1}$ et $E_{\pm 1}^{\perp}$, une base orthonormée de $\mathbb{R}^3$ dans laquelle la matrice de f est

$$M = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad \text{ou } M = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & \sin\theta & -\cos\theta \end{pmatrix}, \quad \text{avec } \theta \in \mathbb{R}.$$

Définition 13.10. **Rotation d'axe orienté et d'angle θ**

Soit $u \neq 0$, ainsi que deux vecteurs v, w non colinéaires et orthogonaux à u, et $\theta \in \mathbb{R}$. On pose $\mathcal{D} = \mathrm{Vect}(u)$ et $\mathcal{P} = \mathcal{D}^{\perp} = \mathrm{Vect}(v, w)$. On appelle **rotation d'axe D orienté par u et d'angle θ** l'unique application linéaire r définie par:

- ✗ $r_{|\mathcal{D}} = \mathrm{Id}_{\mathcal{D}}$;
- ✗ $r_{|\mathcal{P}}$ est la rotation d'angle θ dans le plan $\mathcal{P}$ orienté par u.

Proposition. 13.22.

Toute rotation E_3 est une isométrie directe.

Soit u un vecteur unitaire, $\mathcal{D} = \mathrm{Vect}(u)$, $\theta \in \mathbb{R}$ et $\mathcal{B} = (u, v, w)$ une base orthonormée directe. Alors la matrice de la rotation d'axe $\mathcal{D}$ orienté par u et d'angle θ dans la base $\mathcal{B}$ est :

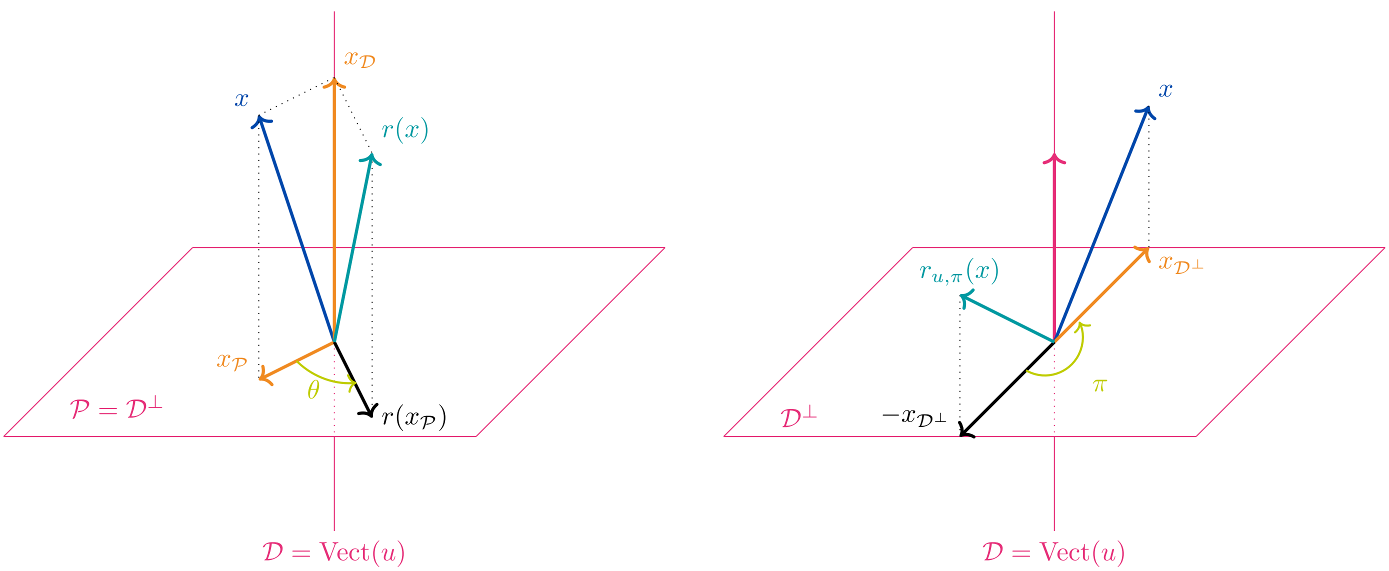
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Remarque 13.13.

- ✗ La rotation d'axe D orienté par u et d'angle θ est souvent notée $r_{u,\theta}$.
- ✗ Si on change u en $-u$, alors l'orientation de $\mathcal{P}$ change; ainsi on a, $r_{u,-\theta} = r_{-u,\theta}$.
- ✗ Contrairement au cas du plan euclidien, la matrice d'une rotation dépend désormais de la base dans laquelle on se place. La base $\mathcal{B}$ utilisée ci-dessus est dite *adaptée* à la rotation, en ce sens qu'elle lui confère une matrice diagonale par blocs.
- ✗ En prenant $\theta = \pi$ on obtient la matrice de la symétrie orthogonale par rapport à $\mathcal{D}$, aussi appelée **demi-tour** ou **retournement**.

Illustration 13.1. **Rotation et retournement**

Les figures ci-dessous représentent respectivement une rotation d'axe $\mathcal{D}$ orienté par u et d'angle θ et un retournement d'axe $\mathcal{D}$ orienté par u.



Méthode 13.18. **Identifier une isométrie directe de l'espace**

Soit f l'endomorphisme dont la matrice dans une b.o.n directe $\mathcal{B}$ est M.

- ✗ On vérifie que $M^\top M = I_3$, c'est-à-dire $M \in \mathcal{O}_3(\mathbb{R})$.
  f est bien une isométrie vectorielle.
- ✗ On vérifie que $\det(f) = 1$; f est une isométrie positive, c'est donc une rotation d'axe dirigée par u et d'angle θ.
- ✗ On détermine l'axe $\mathcal{D} = \mathrm{Vect}(u)$ de la rotation en résolvant $MX = X$.
- ✗ L'angle de la rotation vérifie
  $$\mathrm{Tr}(M) = 1 + 2\cos(\theta)$$
  donc
  $$\theta = \pm \arccos\left(\frac{\mathrm{Tr}(M) - 1}{2}\right).$$
  On trouve le signe de θ en observant que :
  $$\sin(\theta) = \det_{\mathcal{B}}(u, x, f(x)),$$
  où on choisit $x \in \mathcal{D}^\perp$ unitaire quelconque.
  En effet, notant $y = u \wedge x$, (x, y) est une b.o.n directe de $D^\perp$ et (u, x, y) est une b.o.n directe de $\mathbb{R}^3$, donc on obtient
  $$\det_{\mathcal{B}}(u, x, f(x)) = [u, x, f(x)] = (u \wedge x \mid f(x)) = (y \mid \cos\theta x + \sin\theta y) = \|y\|^2 \sin\theta = \sin\theta.$$

Exemple 13.5.

Déterminons les caractéristiques géométriques de l'endomorphisme de $\mathbb{R}^3$ canoniquement associé à :

$$M = \frac{1}{4}\begin{pmatrix} 3 & 1 & \sqrt{6} \\ 1 & 3 & -\sqrt{6} \\ -\sqrt{6} & \sqrt{6} & 2 \end{pmatrix}$$

- ✗ $M^\top M = I_3$ et $\det(M) = 1$. C'est donc une rotation.
- ✗ $MX = X \iff X \in \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right)$. On pose alors $u = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. L'axe de la rotation est $\mathcal{D} = \mathrm{Vect}(u)$.
- ✗ $\mathrm{Tr}(M) = 2$ donc $\cos\theta = \pm\frac{\pi}{3}$ et on montre que $\sin\theta \geqslant 0$. Donc $\theta = \frac{\pi}{3}$.

Exercice 13.5.

Reconnaitre l'endomorphisme dont la matrice dans la base canonique de $\mathbb{R}^3$ est $M = \dfrac{1}{3}\begin{pmatrix} 2 & -1 & -2 \\ 2 & 2 & 1 \\ 1 & -2 & 2 \end{pmatrix}$.

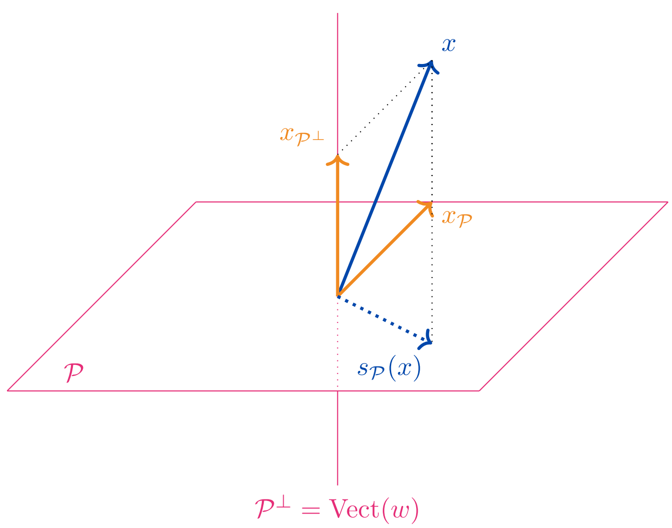
On précisera ses éléments caractéristiques.

Définition 13.11. **Réflexion**

On appelle **réflexion** de $\mathbb{R}^3$ toute symétrie orthogonale par rapport à un hyperplan de $\mathbb{R}^3$ (c'est-à-dire un plan). Si $\mathcal{P}$ est un plan vectoriel, on notera $s_\mathcal{P}$ la réflexion par rapport à $\mathcal{P}$.

Illustration 13.2. **Réflexion**

La figure ci-dessous représente la réflexion par rapport au plan $\mathcal{P}$.



Proposition. 13.23.

Soit $\mathcal{P}$ un plan vectoriel. Alors :

i. Si (u, v) est une base de $\mathcal{P}$, et $w \neq 0$ est orthogonal à $\mathcal{P}$, alors

$$\mathrm{Mat}_{(u,v,w)}(s_\mathcal{P}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

ii. $s_\mathcal{P}$ est une isométrie indirecte.

iii. La matrice d'une réflexion dans une base orthonormée est symétrique.

Remarque 13.14.

La base (u, v, w) utilisée ci-dessus est dite *adaptée* à la réflexion: c'est une base formée de vecteurs propres de $s_\mathcal{P}$. En raisonnant de même on prouve que toute symétrie est diagonalisable.

Proposition. 13.24. **Composée d'une symétrie par une rotation**

Soit u un vecteur unitaire, $\mathcal{D} = \text{Vect}(u)$, $\theta \in \mathbb{R}$ et $\mathcal{B} = (u, v, w)$ une base orthonormée directe.
Soit r la rotation d'axe $\mathcal{D}$ orienté par u et d'angle θ, et s la réflexion par rapport à $\mathcal{D}^\perp = \text{Vect}(v, w)$.

i. $r \circ s$ est une isométrie indirecte;

ii. la matrice de $r \circ s$ dans la base $\mathcal{B}$ est $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$;

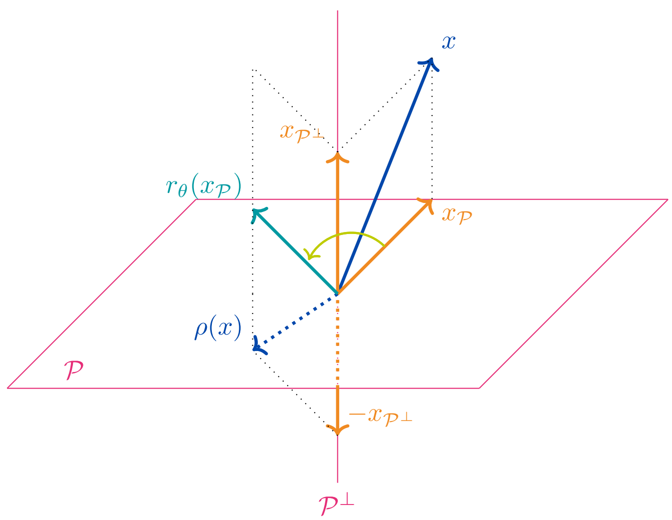
iii. r et s commutent.

Définition 13.12. **Rotation-Réflexion**

On appelle **rotation-réflexion** la composée commutative d'une réflexion s par rapport à un plan $\mathcal{P}$ et d'une rotation r d'axe $\mathcal{D} = \mathcal{P}^\perp$.

Illustration 13.3. **Rotation-Réflexion**

La figure ci-dessous représente la rotation-réflexion ρ composé de la réflexion s par rapport au plan $\mathcal{P}$ et de la rotation r_θ d'axe $\mathcal{D} = \mathcal{P}^\perp$ et d'angle θ.



Théorème 13.25. **Isométrie de $\mathbb{R}^3$**

Toute isométrie de $\mathbb{R}^3$ est soit :

i. une rotation;
ii. une réflexion;
iii. une rotation-réflexion.

Remarque 13.15.

- ✗ Il y a deux types de symétries parmi les isométries de $\mathbb{R}^3$: les réflexions (=symétrie par rapport à un plan) et les demi-tours (=symétrie par rapport à une droite=rotation d'angle π).
- ✗ Les seules matrices orthogonales de taille 3 à être symétriques sont les matrices de symétrie.

Corollaire 13.26.

Soit $f \in \mathcal{SO}(\mathbb{R}^3)$. Alors f est une rotation.

Théorème 13.27. **Classification des isométries de $\mathbb{R}^3$**

On peut établir le tableau de synthèse suivant.

Nature de l'isométrie	Déterminant	Spectre	Matrice dans une certaine b.o.n.
Id_E	1	$\{1\}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Demi-tour	1	$\{1, -1\}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$
Rotation d'angle $\theta(\neq 0, \pi)$	1	$\{1\}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$
$-\mathrm{Id}_E$	-1	$\{-1\}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$
Réflexion	-1	$\{1, -1\}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Rotation-réflexion	-1	$\{-1\}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$

Méthode 13.19.

Pour déterminer la nature et les caractéristiques d'une isométrie f de E_3, on commence par caculer $\dim E_1(f) = \dim \mathrm{Ker}(f - \mathrm{Id})$:

- ✗ si $\dim \mathrm{Ker}(f - \mathrm{Id}) = 3$, alors $f = \mathrm{Id}$;
- ✗ si $\dim \mathrm{Ker}(f - \mathrm{Id}) = 2$, alors f est la réflexion par rapport à $\mathcal{P} = \mathrm{Ker}(f - \mathrm{Id})$;
- ✗ si $\dim \mathrm{Ker}(f - \mathrm{Id}) = 1$, alors f est une rotation d'axe $\mathcal{D} = \mathrm{Ker}(f - \mathrm{Id})$;
- ✗ si $\dim \mathrm{Ker}(f - \mathrm{Id}) = 0$, alors f la composée commutative d'une réflexion et d'une rotation.

Exercice 13.6.

Reconnaitre l' endomorphisme dont la matrices dans la base canonique de $\mathbb{R}^3$ est $M = \dfrac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}$.

On précisera ses éléments caractéristiques.

Méthode 13.20.

Soit $f \in \mathcal{O}(\mathbb{R}^3)$ admettant M pour matrice dans une base orthonormée $\mathcal{B}$. Si $\det M = -1$, alors f est soit une réflexion (ce sera le cas si et seulement si la matrice est symétrique), soit la composée commutative d'une réflexion et d'une rotation d'axe orthogonal, et pour déterminer les éléments caractéristiques de f dans le second cas on procède comme suit:

- ✗ l'axe $\mathcal{D}$ de la rotation est $E_{-1}(M)$, soit l'ensemble des solutions de l'équation $MX = -X$;
- ✗ le plan de la réflexion est $\mathcal{P} = \mathcal{D}^\perp$;
- ✗ l'angle θ de f est donné (au signe près) par $\mathrm{Tr}(M) = -1 + 2\cos\theta$ (si $\theta \equiv 0[2\pi]$ la rotation est triviale, et donc f est une réflexion);
- ✗ θ est du signe de $\det_{\mathcal{B}}(u, x, f(x))$ pour n'importe quel vecteur x unitaire et orthogonal à u, u étant le vecteur choisi pour diriger et orienter l'axe de la rotation.

14 Fonctions de plusieurs variables

14.1 Éléments de topologie

Soit $n = 2$ ou 3. On considère dans cette section l'espace vectoriel $\mathbb{R}^n$ muni d'un repère orthonormé direct $\mathcal{R} = (O, e_1, e_2, ..., e_n)$, et on considère la structure euclidienne usuelle associée.

Les fonctions que l'on va considérer par la suite seront définie sur (une partie de) $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^n$. Afin de pouvoir généraliser les notions usuelles de régularité (et de pouvoir définir la notion de limite), nous avons besoin d'introduire quelques rudiments de *topologie* sur $\mathbb{R}^n$.

La complémentaire d'une partie A de $\mathbb{R}^n$ sera noté ${}^cA$.

Définition 14.1. **Distance associé à la norme euclidienne**

On appelle **distance associée à la norme euclidienne** sur $\mathbb{R}^n$ l'application définie par

$$\begin{array}{rccc} \mathrm{d}: & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}_+ \\ & (x,y) & \longmapsto & \|x-y\|. \end{array}$$

La définition ci-dessous généralise la notion d'intervalle centré en un point.

Définition 14.2. **Boule ouverte, boule fermée**

Soit $a \in \mathbb{R}^n$ et $r > 0$. On appelle **boule ouverte** (resp. **fermée**) de centre a et de rayon r l'ensemble défini par $B_o(a,r) = \{x \in \mathbb{R}^n;\ \mathrm{d}(a,x) < r\}$ (resp. $B_f(a,r) = \{x \in \mathbb{R}^n;\ \mathrm{d}(a,x) \leq r\}$).

Remarque 14.1.

Si $n = 2$ on parle plutôt de **disque**, et si $n = 1$ une boule est un intervalle (fermé ou ouvert).

Définition 14.3. **Intérieur, adhérence, frontière**

Soit A une partie non vide de $\mathbb{R}^n$ et a un point de $\mathbb{R}^n$.

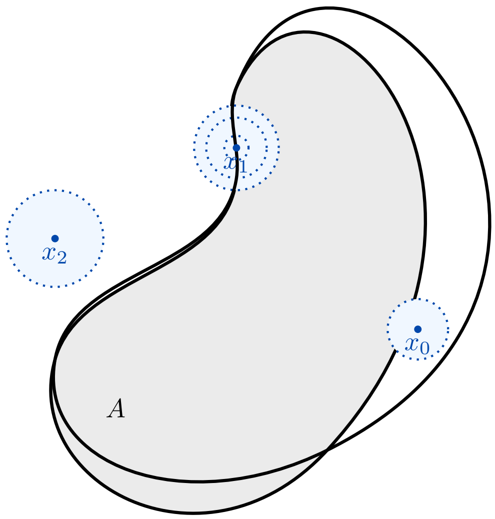
i. a est un **point intérieur** à A s'il existe une boule ouverte de centre a contenue dans A.
On dit alors que A est un *voisinage* de a.
ii. a est un **point extérieur** à A s'il existe une boule ouverte de centre a contenue dans ${}^cA$.
iii. a est un **point adhérent** à A si toute boule ouverte de centre a rencontre A.
iv. a est un point de la **frontière** (ou du **bord**) de A si toute boule ouverte de centre a rencontre A et ${}^cA$.

L'ensemble des points intérieurs à A est appelé l'**intérieur** de A et est noté $\mathring{A}$. L'ensemble des points adhérents à A est appelé l'**adhérence** de A et est noté $\overline{A}$.
La frontière de A est parfois notée $\mathrm{Fr}(A)$ ou ∂A et on a $\partial A = \overline{A} \setminus \mathring{A}$.

Remarque 14.2.

- ✗ a est extérieur à A si et seulement si a est intérieur à ${}^cA$.
- ✗ Tout point de A est adhérent à A. Par contre, il est possible qu'aucun point ne soit à l'intérieur de A même si $A \neq \emptyset$ (par exemple dans le cas d'un cercle).
- ✗ Tout point de la frontière de A est adhérent à A et à ${}^cA$. Mais il est possible que A ne contienne pas certains points de sa frontière (par exemple une boule ouverte n'en contient aucun).

Exemple 14.1. **Point intérieur, Point extérieur : illustration**



Sur la figure ci-contre :

- ✗ x_0 est un point intérieur à A : on peut trouver un disque ouvert centré en x_0 intégralement contenu dans A;
- ✗ x_2 est un point extérieur à A (ou intérieur à ${}^cA$) : on peut trouver un disque ouvert centré en x_2 intégralement contenu dans ${}^cA$;
- ✗ x_1 est sur le frontière de A : toute disque ouvert centré en x_1 rencontre A et ${}^cA$.

Définition 14.4. **Partie ouverte, partie fermée**

Soit A une partie de $\mathbb{R}^n$.

i. A est **ouverte** si et seulement si tout point de A lui est intérieur.
ii. A est **fermée** si et seulement si ${}^cA$ est ouvert.

Exemple 14.2.

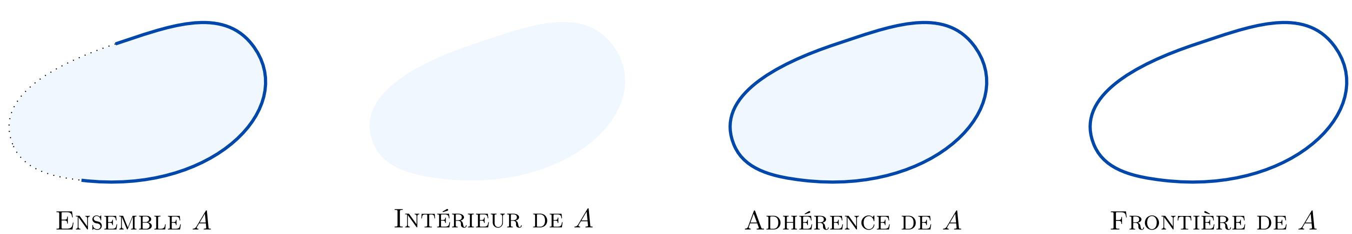
Les boules ouvertes sont des ouverts, les boules fermées sont des fermées.

Remarque 14.3.

- ✗ Une partie de $\mathbb{R}^n$ est ouverte si et seulement si elle ne contient aucun point de sa frontière.
- ✗ Une partie de $\mathbb{R}^n$ est fermée si et seulement si elle contient tous les points de sa frontière.

Exemple 14.3. **Adhérence, intérieur, frontière : illustration**

Sur la figure ci-dessous, la partie A n'est ni ouverte, ni fermée ; en effet, elle contient des éléments de sa frontière.



Définition 14.5. **Partie bornée**

Une partie A de $\mathbb{R}^n$ est **bornée** si et seulement si il existe $r > 0$ tel que $\forall a \in A, \|a\| \leq r$.

Remarque 14.4. **Autre formulation du caractère borné**

Une partie A de $\mathbb{R}^n$ est bornée si et seulement si il existe une boule de centre O contenant A.

Exercice 14.1.

Représenter dans le plan les parties suivantes, avec leur frontière. Sont-elles ouvertes, fermées, bornées?

i. $A = \left\{(x,y) \in \mathbb{R}^2;\ -1 < x \leq 1\right\}$.

ii. $B = \left\{(x,y) \in \mathbb{R}^2;\ \max(|x|,|y|) \leq 1\right\}$.

iii. $C = \left\{(x,y) \in \mathbb{R}^2;\ 0 < x^2 + y^2 < 1\right\}$.

iv. $D = \left\{(x,y) \in \mathbb{R}^2;\ x > 0\right\}$.

14.2 Fonctions de deux variables à valeurs réelles

On propose d'introduire les notions en commençant par le cas des fonctions de deux variables, ce qui sera généralisé par la suite dans une autre section.

14.2.1 Définitions, Représentations graphiques

Définition 14.6. **Fonction de deux variables à valeurs réelles**

On appelle **fonction de deux variables à valeurs réelles** toute fonction $f : \begin{array}{ccc} \mathcal{D} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array}$ où $\mathcal{D} \subset \mathbb{R}^2$.

Exemple 14.4.

$f : (x,y) \mapsto \dfrac{1}{x-y} e^{x^2+y^2}$ est une fonction de deux variables définie sur $\{(x,y) \in \mathbb{R}^2;\ x \neq y\}$.

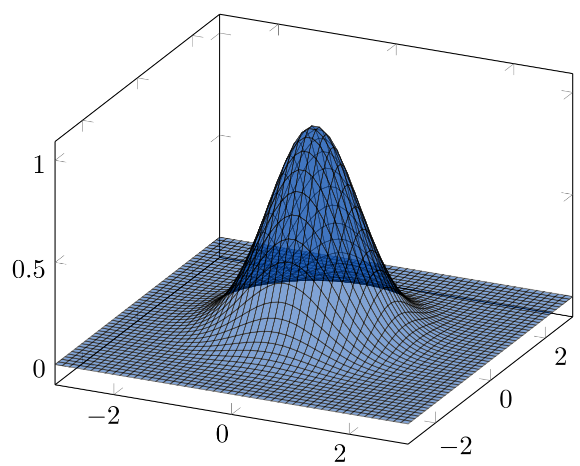
Définition 14.7. **Surface représentative**

On considère un repère orthonormé direct (O, e_1, e_2, e_3) de l'espace, et une fonction de deux variables f définie sur $\mathcal{D} \subset \mathbb{R}^2$.

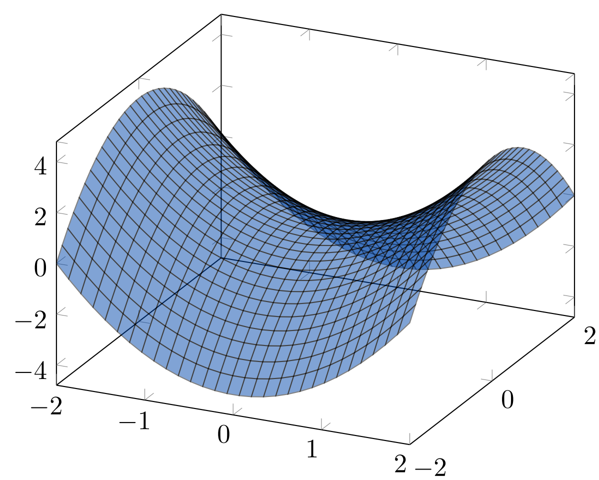
On appelle **surface représentative** de f la *surface* d'équation $z = f(x,y)$, c'est-à-dire l'ensemble des points de coordonnées $(x, y, f(x,y))$ où $(x,y) \in \mathcal{D}$.

Exemple 14.5.

On représente ci-dessous les surfaces représentatives de deux fonctions.



$z = \exp\left(-x^2 - y^2\right), \quad (x,y) \in [-3,3]^2$



$z = x^2 - y^2, \quad (x,y) \in [-2,2]^2$

Définition 14.8. **Lignes de niveau**

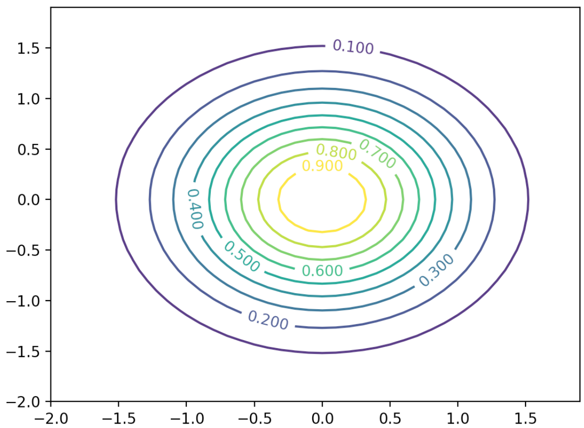
Soient $f : \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables et $k \in \mathbb{R}$. On appelle **ligne de niveau** k de f la courbe

$$\{(x,y) \in \mathcal{U} :\ f(x,y) = k\}$$

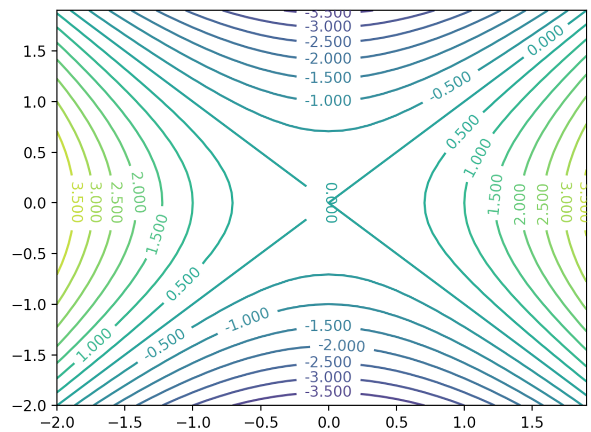
Remarque 14.5.
Si f est la fonction qui à la longitude et la latitude associe l'altitude, alors les lignes de niveau représentent les points qui sont à la même altitude : si on se promène sur une ligne de niveau, on ne monte pas et on ne descend pas : on reste au même niveau.

Exemple 14.6.

On représente quelques niveaux pour les fonctions de l'**Exemple 14.5**.



$e^{-x^2-y^2} = k, \quad k = 0.1, 0.2, ..., 1,\ (x, y) \in [-2, 2]^2.$



$x^2 - y^2 = k, \quad k = -4, -3.5, -3, ..., 4,\ (x, y) \in [-2, 2]^2.$

14.2.2 Continuité

Définition 14.9. **Limite finie en un point**
Soit $\mathcal{U} \subset \mathbb{R}^2$ un ouvert, $f : \mathcal{U} \to \mathbb{R}$, et $(a, b) \in \overline{\mathcal{U}}$.
On dit que f **possède une limite** $\ell \in \mathbb{R}$ en (a, b), ce que l'on note $\lim\limits_{(x,y)\to(a,b)} f(x, y) = \ell$, si :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall (x, y) \in \mathcal{U}, ||(x, y) - (a, b)|| < \delta \Longrightarrow |f(x, y) - \ell| < \varepsilon.$$

Définition 14.10. **Continuité sur un intervalle**
Soit $\mathcal{U} \subset \mathbb{R}^2$ un ouvert, et $f : \mathcal{U} \to \mathbb{R}$. On dit que f est **continue** sur $\mathcal{U}$ si

$$\forall (a, b) \in \mathcal{U}, \quad \lim_{(x,y)\to(a,b)} f(x, y) = f(a, b).$$

Exemple 14.7.
Les projecteurs $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont continues sur $\mathbb{R}^2$.

Théorème 14.1.
Toute combinaison linéaire, tout produit ou quotient (dont le dénominateur ne s'annule pas) de fonctions de deux variables continues est continue.

Exemple 14.8.
La fonction $f(x, y) = y^4 - 3y^3x + yx^2 + xy$ est continue sur $\mathbb{R}^2$. Plus généralement, est continue toute fonction f telle que $f(x, y)$ est une expression polynomiale de x et de y.

Remarque 14.6.
Dans le plan, il existe **une infinité de manières** pour (x, y) de tendre vers (a, b), donc la définition ci-dessus est très contraignante. S'il existe deux chemins vers (a, b) donnant à $f(x, y)$ des limites différentes, alors f n'est pas continue en (a, b).

Exercice 14.2.

Pour $i \in [\![1,2]\!]$, on note $f_i : \mathbb{R}^2 \to \mathbb{R}$ définie par $f_i(x,y) = \begin{cases} \dfrac{x^i y}{x^2+y^2}, & \text{si } (x,y) \neq (0,0) \\ 0, & \text{si } (x,y) = (0,0) \end{cases}$.

Étudier, pour $i \in [\![1,2]\!]$, la continuité de f_i en $(0,0)$.
Commenter.

Théorème 14.2.

Si $f : \mathbb{R}^2 \to \mathbb{R}$ est continue et si $\mathcal{D}$ est une partie fermée et bornée de $\mathbb{R}^2$, alors $f(\mathcal{D})$ est un intervalle fermé et borné.

Remarque 14.7.

Ce résultat prolonge le théorème des bornes atteintes : toute fonction continue sur un segment est bornée et atteint ses bornes.
Il est important dans le cadre de la recherche des *extrema* d'une fonction.

Théorème 14.3. **Images réciproques d'ouverts et fermés par une fonction continue**

Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^2$, et $f : \mathcal{U} \to \mathbb{R}$ une fonction continue. Alors l'ensemble $\{(x,y) \in \mathcal{U};\ f(x,y) > 0\}$ est un ouvert, et les ensembles $\{(x,y) \in \mathcal{U};\ f(x,y) \geq 0\}$, $\{(x,y) \in \mathcal{U};\ f(x,y) = 0\}$ sont des fermés.

14.2.3 Dérivées en un point

Dérivées partielles

Définition 14.11.

Soit f une fonction définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, $(x_0,y_0) \in \mathcal{U}$. On appelle fonctions partielles de f en (x_0,y_0) les deux applications suivantes.

$$f(\cdot,y_0) : x \mapsto f(x,y_0) \qquad \text{et} \qquad f(x_0,\cdot) : y \mapsto f(x_0,y).$$

Remarque 14.8.

Si S est la surface représentative de f (d'équation $z = f(x,y)$) alors la courbe représentative de $f(\cdot,y_0)$ (resp. $f(x_0,\cdot)$) est l'intersection de S avec le plan d'équation $y = y_0$ (resp. $x = x_0$).

Remarque 14.9.

- ✗ Le fait que f soit définie sur un ouvert implique que la fonction partielle $f(\,\cdot\,,y_0)$ est définie sur un voisinage de x_0 (et que $f(x_0,\cdot)$ est définie sur un voisinage de y_0).
- ✗ Si les deux applications $f(x_0,\cdot)$ et $f(\cdot,y_0)$ sont continues, on dit que f est *continue par rapport* à x et y en (x_0,y_0).
  Néanmoins, cette propriété n'implique pas la continuité (conjointe) de f en (x_0,y_0).

Définition 14.12. **Dérivée partielle**

Soit f une fonction définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$ et $(x_0,y_0) \in \mathcal{U}$.

i. Si $\displaystyle\lim_{x \to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}$ existe, alors on dit que f admet une **dérivée partielle par rapport à la première variable** en (x_0,y_0), et l'on note

$$\frac{\partial f}{\partial x}(x_0,y_0) = \lim_{x \to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}.$$

ii. Si $\displaystyle\lim_{y \to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}$ existe, alors on dit que f admet une **dérivée partielle par rapport à la seconde variable** en (x_0,y_0), et l'on note

$$\frac{\partial f}{\partial y}(x_0,y_0) = \lim_{y \to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}.$$

Remarque 14.10.

- ✗ Dire que f admet une dérivée partielle par rapport à la première variable (resp. seconde) en (x_0, y_0), c'est dire que la fonction $f(\cdot, y_0)$ (resp. $f(x_0, \cdot)$) est dérivable en x_0 (resp. y_0). Dans les deux cas, on *gèle* l'une des deux variables avant de dériver par rapport à l'autre.

- ✗ Géométriquement, ce procédé s'interprète comme la recherche de deux droites tangentes à la surface représentative de f en (x_0, y_0), la première dans le plan $y = y_0$ et la seconde dans le plan $x = x_0$ (la donnée de ces deux droites permet alors de définir le plan tangent en (x_0, y_0)).

Méthode 14.21.
Pour calculer $\dfrac{\partial f}{\partial x}(x, y)$, on dérive l'expression $f(x, y)$, en considérant que y est une constante par rapport à x.
De même, pour calculer $\dfrac{\partial f}{\partial y}(x, y)$, on dérive l'expression $f(x, y)$ en considérant que x est une constante par rapport à y.

Exercice 14.3.
Calculer les dérivées partielles de la fonction définie sur $\mathbb{R}^2$ par $f(x, y) = e^x(\cos(x) - \cos(y))$.

Exercice 14.4.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction admettant une dérivée partielle par rapport à sa première variable.
L'égalité $\dfrac{\partial f}{\partial x}(x, x) = \dfrac{\mathrm{d}}{\mathrm{d}x}(f(x, x))$ est-elle vraie?

Remarque 14.11.
Dans la notation $\dfrac{\partial f}{\partial x}(x, y)$, le x situé au dénominateur ne représente pas une variable (contrairement à celui entre parenthèses), mais indique par rapport à quelle *place* (la première ou la seconde) on a dérivé f! Pour éviter les confusion, on pourra préférer les notations $\dfrac{\partial f}{\partial x_1}(x, y)$ ou $\partial_1 f(x, y)$.

Théorème 14.4.
Toute combinaison linéaire, tout produit ou quotient (dont le dénominateur ne s'annule pas) de fonctions admettant des dérivées partielles en un point admet encore des dérivées partielles en ce point.

Remarque 14.12.
Les projecteurs
$$(x, y) \mapsto x, \qquad (x, y) \mapsto y$$
admettent des dérivées partielles en tout point $(x_0, y_0) \in \mathbb{R}^2$.
Il découle alors du Théorème ci-dessus que toutes les fonctions polynomiales admettent des dérivées partielles en tout point $(x_0, y_0) \in \mathbb{R}^2$.

Définition 14.13. **Gradient**
Soit f une fonction définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, et $(x_0, y_0) \in \mathcal{U}$ un point où f admet des dérivées partielles. On appelle alors gradient de f en (x_0, y_0), et l'on note $\nabla f(x_0, y_0)$ ou $(\overrightarrow{\text{grad}}\, f)(x_0, y_0)$, le vecteur défini par
$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0),\ \frac{\partial f}{\partial y}(x_0, y_0)\right).$$

Exercice 14.5.
Soit f la fonction définie sur $\mathbb{R}^2$ par $f(x, y) = ye^{-x^2-y}$.
1. Montrer que f admet des dérivées partielles d'ordre 1 et les calculer.
2. Déterminer le gradient de la fonction f au point $(1, 2)$.

Dérivées selon un vecteur quelconque

Définition 14.14.
Soient f une fonction définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, un point $a = (x_0, y_0) \in \mathcal{U}$ et $\vec{u}$ un vecteur de $\mathbb{R}^2$. On dit que f est **dérivable en** a **selon le vecteur** $\vec{u}$ si la fonction $g: \ t \mapsto f(a + t\vec{u})$ est dérivable en 0.
Lorsque c'est le cas, la dérivée $g'(0)$ est notée $D_{\vec{u}} f(a)$ et est appelée dérivée de f en a selon le vecteur $\vec{u}$.

Remarque 14.13.
On voit immédiatement que les dérivées partielles de f en a sont les dérivées de f en a selon les vecteurs de la base canonique de $\mathbb{R}^2$.
Il est donc clair que si f est dérivable suivant tout vecteur en a, alors f admet des dérivées partielles par rapport à chacune de ses variables en a.

L'exemple de fonction ci-dessous est très important ; il propose une fonction admettant des dérivées partielles en un point sans être dérivable suivant tout vecteur en ce point.
Attention donc aux conclusions hâtives.

Exercice 14.6.

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = \begin{cases} 1, & \text{si } x \neq 0 \text{ et } y \neq 0, \\ 0, & \text{si } x = 0 \text{ ou } y = 0 \end{cases}$.

1. Montrer que f admet des dérivées partielles en $(0,0)$ et calculer celles-ci.
2. Est-ce que f est dérivable en $(0,0)$ selon le vecteur $\vec{u} = (1,1)$?

14.2.4 Fonctions de classe $\mathcal{C}^1$

Définition 14.15. **Fonctions de classe $\mathcal{C}^1$ sur $\mathcal{U} \subset \mathbb{R}^2$**
Une fonction f définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$ est dite de classe $\mathcal{C}^1$ si elles possèdent des dérivées partielles en tout point de $\mathcal{U}$, et si celles-ci sont continues.
L'ensemble des fonctions de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^2$ sera noté $\mathcal{C}^1(\mathcal{U}, \mathbb{R})$.

Remarque 14.14.
Le fait que les fonctions partielles de f soient $\mathcal{C}^1$ n'implique pas que f l'est (toujours car la continuité par rapport à x et y séparément n'implique pas la continuité par rapport à (x,y)).

Remarque 14.15.
Toute expression polynomiale en x et en y définit une fonction de classe $\mathcal{C}^1$.

Théorème 14.5.
Toute combinaison linéaire, tout produit ou quotient (dont le dénominateur ne s'annule pas) de fonctions de classe $\mathcal{C}^1$ est encore de classe $\mathcal{C}^1$.

Théorème 14.6. **Composition à gauche**
Soient $f: \mathcal{U} \subset \mathbb{R}^2$ une fonction de classe $\mathcal{C}^1$ à valeurs dans $I \subset \mathbb{R}$ et $\varphi : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Alors, $\varphi \circ f$ est encore de classe $\mathcal{C}^1$ (sur $\mathcal{U}$) et, pour tout $(x,y) \in \mathcal{U}$, on a

$$\frac{\partial}{\partial x}(\varphi \circ f)(x,y) = \varphi'(f(x,y)) \frac{\partial f}{\partial x}(x,y), \qquad \text{et} \qquad \frac{\partial}{\partial y}(\varphi \circ f)(x,y) = \varphi'(f(x,y)) \frac{\partial f}{\partial y}(x,y).$$

Exercice 14.7.
Justifier que $f = (x,y) \mapsto \ln(x^2 + y^2 + 1) - x$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et calculer ses dérivées partielles d'ordre 1 en tout point de $\mathbb{R}^2$.

Théorème 14.7. **Formule de Taylor-Young, ordre 1**

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$ et $(x_0, y_0) \in \mathcal{U}$. On a alors lorsque $(h,k) \to (0,0)$:

$$\begin{aligned} f(x_0+h, y_0+k) &= f(x_0,y_0) + \nabla f(x_0,y_0) \cdot \begin{pmatrix} h \\ k \end{pmatrix} + o\left(\|(h,k)\|\right) \\ &= f(x_0,y_0) + h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0) + o\left(\|(h,k)\|\right) \end{aligned}$$

Exemple 14.9.

Soit $f\,(x,y) \in \mathbb{R}^2 \mapsto ye^x + e^{2y} + x^2$. Il est clair que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. On peut alors écrire le développement limité de f à l'ordre 1 en $(0,0)$ grâce à la formule de Taylor-Young à l'ordre 1. Pour (x,y) proche de $(0,0)$,

$$\begin{aligned} f(x,y) &= f(0,0) + \nabla(f)(0,0) \cdot \begin{pmatrix} x \\ y \end{pmatrix} + o\left(\|(x,y)\|\right) \\ &= 1 + \begin{pmatrix} 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + o\left(\|(x,y)\|\right) \\ &= 1 + 3y + o\left(\|(x,y)\|\right) \end{aligned}$$

Remarque 14.16. **(hyperPlan tangent)**

La surface d'équation $z = f(x,y)$ admet alors un **plan tangent** en tout point (x_0, y_0, z_0) tel que $z_0 = f\,(x_0, y_0)$. Celui-ci a pour équation :

$$z = f\,(x_0,y_0) + \frac{\partial f}{\partial x}\,(x_0,y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}\,(x_0,y_0) \cdot (y - y_0)\,.$$

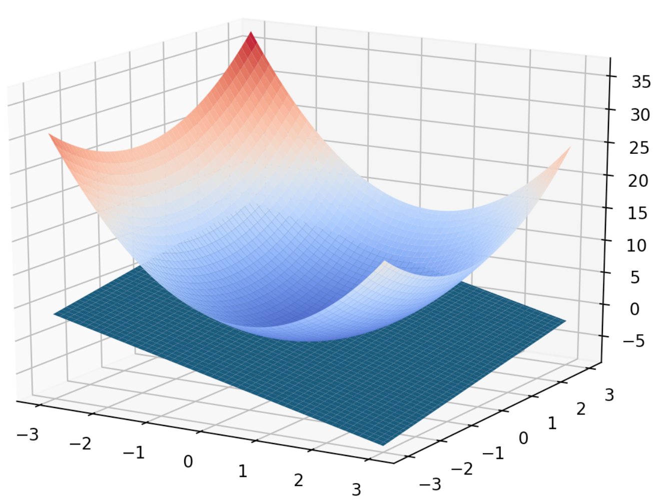
Cette équation est de la forme $a\,(x-x_0) + b\,(y-y_0) + c\,(z-z_0) = 0$ avec $a = \frac{\partial f}{\partial x}\,(x_0,y_0)\,, b = \frac{\partial f}{\partial y}\,(x_0,y_0)$ et $c = -1$. Le vecteur $\left(\frac{\partial f}{\partial x}\,(x_0,y_0)\,, \frac{\partial f}{\partial y}\,(x_0,y_0)\,, -1\right)$ est normal au plan tangent.

Exemple 14.10.

Le code `Python` téléchargeable **ici** permet de visualiser un exemple; on a représenté le plan tangent à la surface représentative de $f:\ (x,y) \mapsto 2(x-1/2)^2 + (y+1/2)^2$ en 0.

Ce plan affine a pour équation

$$z = f(0) + \langle \nabla f(0), (x,y) \rangle = \frac{3}{4} - 2x + y$$



Remarque 14.17.

Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$ et $a \in \mathcal{U}$. Soit $\vec{u} = (u_1, u_2)$ un vecteur de $\mathbb{R}^2$. Pour t assez proche de 0, comme $\mathcal{U}$ est ouvert, le point $a + t\vec{u}$ est encore dans $\mathcal{U}$. On peut donc définir la fonction $g : t \mapsto f(a + t\vec{u})$. Cette fonction est alors de classe $\mathcal{C}^1$ (sur le voisinage de 0 sur lequel elle est définie) - ce qui découle de la règle de la chaîne qu'on précise après - et on a :

$$g'(t) = \langle \nabla f(a + t\vec{u}), \vec{u} \rangle = \frac{\partial f}{\partial x}(a + t\vec{u})u_1 + \frac{\partial f}{\partial y}(a + t\vec{u})u_2.$$

En particulier, comme l'on sait que les dérivées partielles de f sont continues sur $\mathcal{U}$, on peut faire tendre $t \to 0$; on obtient $D_{\vec{u}} f(a) = \lim_{t\to 0} g'(t) = \langle \nabla f(a), \vec{u} \rangle$.

14.2.5 Formules de dérivation des composées (ou règle de la chaîne)

On déduit du résultat précédent la formule de **dérivation des composées** ci-dessous.

Théorème 14.8. **Règle de la chaîne**

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, I un intervalle non trivial de $\mathbb{R}$ et $\varphi : t \in I \mapsto (x(t), y(t)) \in \mathcal{U}$ une fonction vectorielle de classe $\mathcal{C}^1$.
Alors la fonction définie par $f \circ \varphi : \ t \in I \mapsto f\,(x(t), y(t))$ est de classe $\mathcal{C}^1$ sur I, de sorte que pour tout $t \in I$:

$$\begin{aligned}
(f \circ \varphi)'(t) &= \frac{\partial f}{\partial x}\Big(x(t), y(t)\Big) \times x'(t) + \frac{\partial f}{\partial y}\Big(x(t), y(t)\Big) \times y'(t) \\
&= \nabla f\Big(x(t), y(t)\Big) \cdot \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}
\end{aligned}$$

En d'autres termes, on a $(f \circ \varphi)' = (\nabla f \circ \phi) \cdot \varphi'$.

Preuve. φ est de classe $\mathcal{C}^1$ sur I, donc x et y sont dérivables et pour tout $t \in I$:

$$\begin{aligned}
(f \circ \varphi)(t+h) - (f \circ \varphi)(t) &= f(x(t+h), y(t+h)) - f(x(t), y(t)) \\
&\underset{h\to 0}{=} f(x(t) + \underbrace{hx'(t) + \mathrm{o}(h)}_{=h'}, y(t) + \underbrace{hy'(t) + \mathrm{o}(h)}_{=k'}) - f(x(t), y(t))
\end{aligned}$$

Or $f\,(x(t) + h', y(t) + k') \underset{h\to 0}{=} f(x(t), y(t)) + \frac{\partial f}{\partial x}(x(t), y(t)) \cdot h' + \frac{\partial f}{\partial y}(x(t), y(t)) \cdot k' + \mathrm{o}\left(\sqrt{h'^2 + k'^2}\right)$.
D'où,

$$(f \circ \varphi)(t+h) - (f \circ \varphi)(t) \underset{h\to 0}{=} \left[x'(t)\frac{\partial f}{\partial x}(x(t), y(t)) + y'(t)\frac{\partial f}{\partial y}(x(t), y(t))\right] \cdot h + \mathrm{o}(h).$$

□

Remarque 14.18.

Appliquer la règle de la chaîne revient à dériver f le long de la courbe paramétrée par $M(t) = (x(t), y(t))$.

Exercice 14.8.

Soit $f(x, y) = x^2 + y^2$. Calculer la dérivée de $f(\cos t, \sin t)$ de trois façons différentes.

Théorème 14.9. **Dérivée selon un vecteur et gradient**

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, un point $a \in \mathcal{U}$ et un vecteur $v \in \mathbb{R}^2$. Alors f est dérivable en a selon le vecteur v, et

$$D_v f(a) = \langle \nabla f(a), v \rangle\,.$$

Théorème 14.10. **Règle de la chaîne, généralisation**

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, x et y deux fonctions de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{V}$ de $\mathbb{R}^2$, telles que (x, y) soit à valeurs dans $\mathcal{U}$.

Alors la fonction définie par $\Phi : (u, v) \in \mathcal{V} \mapsto f\Big(x(u, v), y(u, v)\Big)$ est de classe $\mathcal{C}^1$ sur $\mathcal{V}$ de sorte que pour tout $(u, v) \in \mathcal{V}$:

$$\begin{aligned} \frac{\partial \Phi}{\partial u}(u, v) &= \frac{\partial f}{\partial x}\Big(x(u, v), y(u, v)\Big) \times \frac{\partial x}{\partial u}(u, v) + \frac{\partial f}{\partial y}\Big(x(u, v), y(u, v)\Big) \times \frac{\partial y}{\partial u}(u, v) \\ &= \nabla f\Big(x(u, v), y(u, v)\Big) \cdot \begin{pmatrix} \frac{\partial x}{\partial u}(u, v) \\ \frac{\partial y}{\partial u}(u, v) \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \frac{\partial \Phi}{\partial v}(u, v) &= \frac{\partial f}{\partial x}\Big(x(u, v), y(u, v)\Big) \times \frac{\partial x}{\partial v}(u, v) + \frac{\partial f}{\partial y}\Big(x(u, v), y(u, v)\Big) \times \frac{\partial y}{\partial v}(u, v) \\ &= \nabla f\Big(x(u, v), y(u, v)\Big) \cdot \begin{pmatrix} \frac{\partial x}{\partial v}(u, v) \\ \frac{\partial y}{\partial v}(u, v) \end{pmatrix} \end{aligned}$$

Preuve. Il suffit d'appliquer le résultat précédent aux applications partielles $\Phi(\cdot, v)$ et $\Phi(u, \cdot)$. □

Remarque 14.19.

On déduit des deux résultats précédents que toute fonction de deux variables obtenue par composition de fonctions $\mathcal{C}^1$ (d'une ou deux variables) est elle-même de classe $\mathcal{C}^1$.

Exercice 14.9.

Soit $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$. On pose $g : (r, \theta) \in \mathbb{R}_+^* \times]0, 2\pi[\mapsto f(r\cos\theta, r\sin\theta)$. Justifier que g est de classe $\mathcal{C}^1$, et calculer ses dérivées partielles.

14.2.6 Fonctions de classe $\mathcal{C}^2$ et théorème de Schwarz

Définition 14.16. **Dérivées partielles d'ordre 2**

Soit f une fonction définie sur $\mathbb{R}^2$ et admettant des dérivées partielles d'ordre 1. Si les fonctions dérivées partielles

$$(x, y) \mapsto \frac{\partial f}{\partial x}(x, y) \qquad \text{et} \qquad (x, y) \mapsto \frac{\partial f}{\partial y}(x, y)$$

admettent également des dérivées partielles d'ordre 1, on dit que f admet des dérivées partielles d'ordre 2. On note alors

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Définition 14.17. **Fonctions de classe $\mathcal{C}^2$**

Soit f une fonction définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$.
f est dite de classe $\mathcal{C}^2$ sur $\mathcal{U}$ si les (quatre) dérivées partielles d'ordre 2 existent en tout point de $\mathcal{U}$ et sont continues.

Théorème 14.11. **Théorème de Schwarz**

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$. Pour tout $(x_0, y_0) \in \mathcal{U}$ on a

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

Preuve. Résultat admis. □

Exercice 14.10.

Soit f définie par $f(x,y) = \arctan\left(\dfrac{x}{y}\right) + \ln(x^2+y^2)$ sur $]0,+\infty[^2$. Calculer $\dfrac{\partial^2 f}{\partial x \partial y}(x,y)$ et $\dfrac{\partial^2 f}{\partial y \partial x}(x,y)$.

Définition 14.18. **Matrice Hessienne**

Soit f une fonction définie sur $\mathcal{U} \subset \mathbb{R}^2$ admettant des dérivées partielles d'ordre 2 en un point $a \in \mathcal{U}$. La **matrice hessienne** de f au point a est la matrice $H_f(a) \in \mathcal{M}_2(\mathbb{R})$ suivante :

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x \partial y}(a) \\ \frac{\partial^2 f}{\partial y \partial x}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}.$$

La matrice hessienne est aussi parfois notée $\nabla^2 f(a)$.

Remarque 14.20.

Il découle du théorème de Schwarz que, si f est de classe $\mathcal{C}^2$, alors sa matrice hessienne est symétrique réelle et donc **diagonalisable**.

Théorème 14.12. **Formule de Taylor-Young, ordre 2**

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$ et $(x_0,y_0) \in \mathcal{U}$. On a alors lorsque $(h,k) \to (0,0)$:

$$\begin{aligned} f(x_0+h, y_0+k) &= f(x_0,y_0) + \nabla f(x_0,y_0) \cdot \begin{pmatrix} h \\ k \end{pmatrix} + \frac{1}{2} \begin{pmatrix} h & k \end{pmatrix} H_f(x_0,y_0) \begin{pmatrix} h \\ k \end{pmatrix} + o\left(\|(h,k)\|^2\right) \\ &= f(x_0,y_0) + h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0) + \\ &\qquad \frac{1}{2}\left(h^2 \frac{\partial^2 f}{\partial x^2}(x_0,y_0) + 2hk\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right) + o\left(\|(h,k)\|^2\right) \end{aligned}$$

14.3 Généralisation aux fonctions vectorielles de p variables

Dans la suite, on fixe des entiers $n, p \in [\![1,3]\!]$, et on s'intéresse aux fonctions $f : \mathbb{R}^p \to \mathbb{R}^n$. Tous les résultats de cette partie seront admis.
On notera sans distinction $\|\cdot\|$ la norme euclidienne sur $\mathbb{R}^n$ et sur $\mathbb{R}^p$ ainsi que d la distance euclidienne : $d(u,v) = \|u-v\|$.

14.3.1 Définitions

Définition 14.19. **Vocabulaire**

Soit $f : \mathbb{R}^p \to \mathbb{R}^n$.

- ✗ Si $n = 1$, f est à valeurs numériques : pour tout $(x_1,\ldots,x_p) \in \mathbb{R}^p$, $f(x_1,\ldots,x_p)$ est un réel.
- ✗ Si $n > 1$, f est à valeurs vectorielles: pour tout $(x_1,\ldots,x_p) \in \mathbb{R}^p$, $f(x_1,\ldots,x_p)$ est un vecteur de $\mathbb{R}^n$.
- ✗ Si $n > 1$, on peut écrire pour tout $(x_1,\ldots,x_p) \in \mathbb{R}^p$, $f(x_1,\ldots,x_p) = \Big(f_1(x_1,\ldots,x_p),\ldots,f_n(x_1,\ldots,x_p)\Big)$ où les $f_i : \mathbb{R}^p \to \mathbb{R}$ sont appelées les **fonctions coordonnées** de f.

Exemple 14.11.

$f : (x,y,z) \mapsto (\cos(x+y+z), x-z)$ est une fonction de $\mathbb{R}^3$ dans $\mathbb{R}^2$.

14.3.2 Continuité et dérivées partielles en un point

Définition 14.20. **Limite**

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}^n$, et $a = (a_1,\ldots,a_p) \in \overline{\mathcal{U}}$. On dit que f possède une limite $\ell \in \mathbb{R}^n$ en a, ce que l'on note $\lim\limits_{u \to a} f(u) = \ell$, si la formule ci-dessous est vérifiée:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall u \in \mathcal{U}, (\|u-a\| < \delta \Longrightarrow \|f(u)-\ell\| < \varepsilon).$$

Théorème 14.13. **Limite et fonctions coordonnées**

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}^n$, et $f_1, \dots, f_n$ ses fonctions coordonnées. Soit encore $a = (a_1, \dots, a_p) \in \overline{\mathcal{U}}$ et $\ell = (\ell_1, \dots, \ell_n) \in \mathbb{R}^n$. On a alors:

$$\lim_{u \to a} f(u) = \ell \iff \forall i \in [\![1, n]\!],\ \lim_{u \to a} f_i(u) = \ell_i.$$

Définition 14.21.

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}^n$. f est dite continue sur $\mathcal{U}$ si, pour tout $a \in \mathcal{U}$, on a $\lim\limits_{u \to a} f(u) = f(a)$.

Théorème 14.14. **Continuité et fonctions coordonnées**

f est continue si et seulement si ses fonctions coordonnées le sont.

Théorème 14.15.

Toute combinaison linéaire, ou (lorsque cela a un sens) tout produit scalaire, produit vectoriel, déterminant ou quotient (dont le dénominateur ne s'annule pas) de fonctions de plusieurs variables continues est continue.

Théorème 14.16.

Si $f : \mathbb{R}^p \to \mathbb{R}^n$ est continue et si Ω est une partie fermée et bornée de $\mathbb{R}^p$, alors $f(\Omega)$ est une partie fermée et bornée de $\mathbb{R}^n$.

Théorème 14.17. **Image réciproque par une fonction continue**

Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^p$, et $f : \mathcal{U} \to \mathbb{R}$ une fonction continue. Alors l'ensemble $\{x \in \mathcal{U};\ f(x) > 0\}$ est un ouvert, et les ensembles $\{x \in \mathcal{U};\ f(x) \geq 0\}$, $\{x \in \mathcal{U};\ f(x) = 0\}$ sont des fermés.

On peut alors généraliser la notion de dérivée partielle.

Définition 14.22.

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}^n$, $a = (a_1, \dots, a_p) \in \mathcal{U}$ et $i \in [\![1, n]\!]$.

Si $\lim\limits_{x \to a_i} \dfrac{f(a_1, \dots, x, \dots, a_p) - f(a_1, \dots, a_i, \dots, a_p)}{x - a_i}$ existe, alors on dit que f admet une dérivée partielle par rapport à la i-ème variable en $a = (a_1, \dots, a_p)$, et l'on note

$$\frac{\partial f}{\partial x_i}(a) = \lim_{x \to a_i} \frac{f(a_1, \dots, x, \dots, a_p) - f(a_1, \dots, a_i, \dots, a_p)}{x - a_i}.$$

Remarque 14.21.

On peut définir la dérivée d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}^n$ selon un vecteur v quelconque, exactement comme dans le cas des fonctions d'une variable.

14.3.3 Fonctions de classe $\mathcal{C}^1$

Définition 14.23. **Classe $\mathcal{C}^1$**

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, et $f : \mathcal{U} \to \mathbb{R}^n$. f est dite de classe $\mathcal{C}^1$ si elles possèdent des dérivées partielles en tout point de $\mathcal{U}$, et si celles-ci sont continues.

Théorème 14.18.

Toute combinaison linéaire, ou (lorsque cela a un sens) tout produit scalaire, produit vectoriel, déterminant ou quotient de fonctions de classe $\mathcal{C}^1$ est encore de classe $\mathcal{C}^1$.

Remarque 14.22.

On peut démontrer qu'une fonction à valeurs vectorielles est $\mathcal{C}^1$ si et seulement si ses fonctions coordonnées le sont.

14.3.4 Extension des résultats pour $p \geq 3$ et $n = 1$

Définition 14.24. **Gradient**

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}$, et $a \in \mathcal{U}$ un point où f admet des dérivées partielles. On appelle alors **gradient** de f en a, et l'on note $\nabla f(a)$, le vecteur défini par

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_p}(a)\right).$$

Théorème 14.19. **Taylor-Young à l'ordre 1**

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et $a \in \mathcal{U}$. On a alors, si $h = (h_1, \ldots, h_p) \in \mathbb{R}^p$:

$$\begin{aligned} f(a+h) &\underset{h \to 0_{\mathbb{R}^p}}{=} f(a) + \langle \nabla f(a), h \rangle + o(\|h\|) \\ &= f(a) + \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a) + o(\|h\|) \end{aligned}$$

Théorème 14.20.

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$.
Soit $m \in \mathbb{N}^*$, $\mathcal{V} \subset \mathbb{R}^m$ un ouvert, $g : \mathcal{V} \to \mathcal{U}$ une fonction de classe $\mathcal{C}^1$ de fonctions coordonnées $g_1, \ldots, g_p$.
Alors $f \circ g$ est de classe $\mathcal{C}^1$ sur $\mathcal{V}$, de sorte que pour tout $i \in [\![1, m]\!]$:

$$\begin{aligned} \forall x = (x_1, \ldots, x_m) \in \mathcal{V},\ \frac{\partial}{\partial x_i} f \circ g(x) &= \frac{\partial f}{\partial x_1}(g(x)) \times \frac{\partial g_1}{\partial x_i}(x) + \cdots + \frac{\partial f}{\partial x_p}(g(x)) \times \frac{\partial g_p}{\partial x_i}(x) \\ &= \sum_{k=1}^p \frac{\partial f}{\partial x_k}(g(x)) \times \frac{\partial g_k}{\partial x_i}(x) \end{aligned}$$

Méthode 14.22.

Si l'on souhaite dériver une composée de la forme $f \circ g$ avec f à valeurs vectorielles, alors il suffit d'appliquer la formule ci-dessus à chacune des fonctions coordonnées de f.

Remarque 14.23. **Matrice Hessienne et Taylor-Young à l'ordre 2**

Soit $\mathcal{U} \subset \mathbb{R}^p$ un ouvert, $f : \mathcal{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ et $a \in \mathcal{U}$.
On peut, comme dans le cas où $p = 2$ définir la matrice Hessienne de f, toujours notée $H_f(a) = \left(\frac{\partial^2}{\partial x_i \partial x_j} f(a)\right)_{1 \leq i,j \leq p}$.
On a alors, si $h = (h_1, \ldots, h_p) \in \mathbb{R}^p$:

$$\begin{aligned} f(a+h) &\underset{h \to 0_{\mathbb{R}^p}}{=} f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2}\left[h \cdot H_f(a) \cdot h^\top\right] + o\left(\|h\|^2\right) \\ &= f(a) + \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2} \sum_{1 \leq i,j \leq p} h_i h_j \frac{\partial^2}{\partial x_i \partial x_j} f(a) + o\left(\|h\|^2\right) \end{aligned}$$

14.4 Extrema d'une fonction numérique de deux variables

Dans toute la suite on considère des fonctions numériques à deux variables (ainsi on revient aux objets étudiés en début de chapitre, $p = 2$ et $n = 1$).

14.4.1 Extrema locaux et globaux, points cols

Définition 14.25. **Extrema locaux**

Soit f une fonction de deux variables définie sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, et $(x_0, y_0) \in U$.

i. On dit que f admet un **minimum (resp. maximum) global** en (x_0, y_0) si $\forall (x, y) \in U,\ f(x, y) \geq f(x_0, y_0)$ (resp. $f(x, y) \leq f(x_0, y_0)$).

ii. On dit que f admet un **minimum (resp. maximum) local** en (x_0, y_0) s'il existe une boule ouverte B de centre (x_0, y_0) incluse dans U telle que $\forall (x, y) \in B,\ f(x, y) \geq f(x_0, y_0)$ (resp. $f(x, y) \leq f(x_0, y_0)$).

Remarque 14.24.
Lorsque f est définie sur un ouvert, tout extremum global est un extremum local. Mais ce résultat est faux si $\mathcal{U}$ n'est pas ouvert, car f peut dans ce cas avoir un extremum global sur la frontière de $\mathcal{U}$.

Le résultat suivant est analogue à un résultat vu dans le cas des fonctions d'une variable réelle.

Théorème 14.21. **Condition nécessaire d'existence d'un extremum local**
Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$. Si f admet un extremum local en $(x_0, y_0) \in \mathcal{U}$ alors

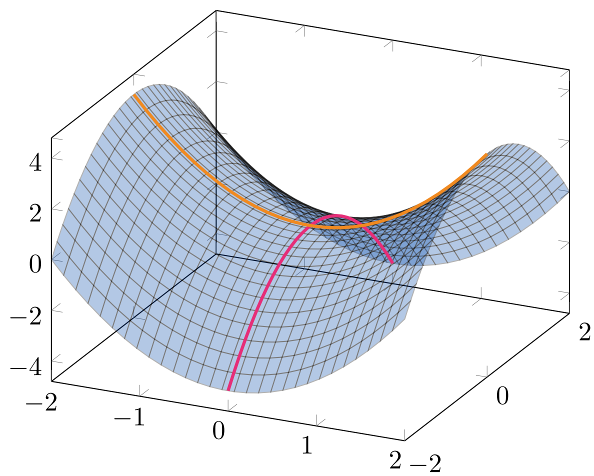
$$\nabla f(x_0, y_0) = 0.$$

En d'autres termes, tout extremum local d'une fonction $\mathcal{C}^1$ en est un **point critique** (ou stationnaire, ou singulier).

Preuve. Si f admet un extremum local en (x_0, y_0), alors la fonction partielle $x \mapsto f(x, y_0)$ admet un extremum local en x_0. Celle-ci étant définie sur un intervalle ouvert, on en déduit que sa dérivée s'annule en x_0, soit $\dfrac{\partial f}{\partial x}(x_0, y_0) = 0$. On démontre de même la seconde égalité. □

Exemple 14.12. **Selle de cheval ou *Pringles***
La figure ci-dessous représente le paraboloïde hyperbolique d'équation $z = x^2 - y^2$ (communément appelé *Pringles*). Le point $(0, 0)$ est un maximum local pour la fonction $f(0, \cdot) : y \mapsto -y^2$ (représentée en rose) et un minimum local pour la fonction $f(\cdot, 0) : x \mapsto x^2$ (représentée en orange).

Il n'est donc pas un extremum local pour f; on parle dans ce cas de **point col** ou de **point selle**, la surface représentative évoquant la forme d'une selle de cheval.



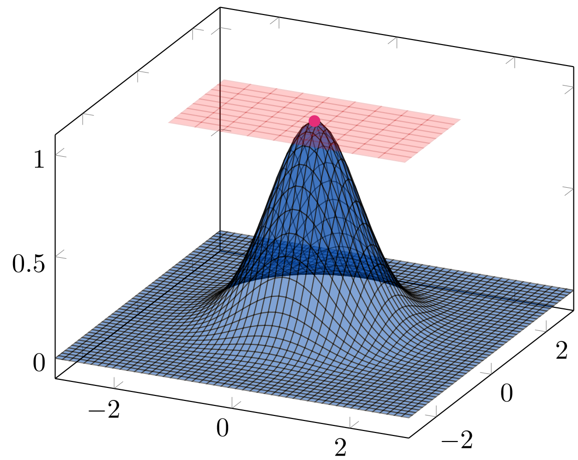
Définition 14.26. **Point col**
Soit f une fonction de deux variables de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$, et $(x_0, y_0) \in U$. On dit que f admet un **point col** en (x_0, y_0) si :

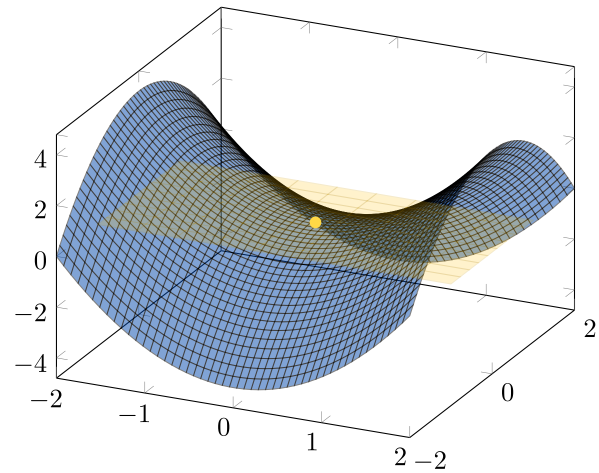
i. $\nabla f(x_0, y_0) = 0$ (c'est à dire si (x_0, y_0) est un point critique) ;
ii. dans toute boule de centre (x_0, y_0) incluse dans $\mathcal{U}$, il existe des points u, u' tels que $f(u) > f(x_0, y_0)$ et $f(u') < f(x_0, y_0)$.

Remarque 14.25.
Soit S une surface d'équation $z = f(x, y)$.

- ✗ Un extremum de f est un point critique au voisinage duquel S reste au-dessus de son plan tangent (si c'est un minimum), ou en-dessous (si c'est un maximum) de celui-ci.
- ✗ Un point col de f est un point critique en lequel le plan tangent est traversé par S.

Exemple 14.13.
On illustre la remarque ci-dessus avec les surfaces déjà rencontrées ci-avant.





Méthode 14.23.
Soit f une fonction définie sur une partie de $\mathbb{R}^2$, et (x_0, y_0) un point critique de celle-ci. Pour déterminer la nature de ce point, il faut étudier le signe de $\Delta(x,y) = f(x,y) - f(x_0,y_0)$ au voisinage de (x_0, y_0) :

- ✗ s'il existe une boule centrée en (x_0, y_0) où Δ est toujours positif, alors (x_0, y_0) est un minimum local;
- ✗ s'il existe une boule centrée en (x_0, y_0) où Δ est toujours négatif, alors (x_0, y_0) est un maximum local;
- ✗ si Δ change de signe sur toute boule centrée en (x_0, y_0), alors (x_0, y_0) est un point col.

☞ On peut montrer qu'on a un point col en un point critique a en étudiant f dans une direction, c'est à dire en étudiant le signe de la fonction $t \mapsto f(a+t\vec{u}) - f(a)$ et en montrant par exemple que cette fonction s'annule en changeant de signe en 0.

Exercice 14.11.
Montrer que la fonction définie sur $\mathbb{R}^2$ par $f(x,y) = x^2y - xy^2$ n'admet pas d'extremum (local ou global).

14.4.2 Condition suffisante locale de second ordre

Le développement limité (formule de Taylor-Young) à l'ordre 2 permet de donner un critère à la présence d'un extremum. En effet, si a est un point critique de f, pour h dans un voisinage de 0, on a

$$f(a+h) - f(a) = \frac{1}{2}\left(h \cdot H_f(a) \cdot h^\top\right) + o\left(\|h\|^2\right).$$

Comme $H_f(a)$ est symétrique réelle (d'après le **Théorème 14.11 de Schwarz**), elle est diagonalisable (dans une base orthonormée). Notant P la matrice (orthogonale) de passage de la base canonique de $\mathbb{R}^2$ à cette base et λ_1, λ_2 les valeurs propres de $H_f(a)$ (comptées avec multiplicité), on aura

$$f(a+h) - f(a) = \lambda_1 k_1^2 + \lambda_2 k_2^2 + o\left(\|k\|\right), \qquad \text{où } \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = P^\top \begin{pmatrix} h_1 \\ h_2 \end{pmatrix},$$

avec $\|k\| = \|h\|$. Des valeurs propres de même signe non nul permettent alors de conclure quant à la nature (locale) du point critique.

Théorème 14.22. **Condition suffisante locale de second ordre, version 1**

Soient f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U}$ et $a \in \mathcal{U}$ un point critique de f.

i. Si $\mathrm{Sp}(H_f(a)) \subset \mathbb{R}_+^*$, alors f présente un minimum local en a.

ii. Si $\mathrm{Sp}(H_f(a)) \subset \mathbb{R}_-^*$, alors f présente un maximum local en a.

iii. Si $\mathrm{Sp}(H_f(a)) \cap \mathbb{R}_+^* \neq \emptyset$ **et** $\mathrm{Sp}(H_f(a)) \cap \mathbb{R}_-^* \neq \emptyset$, alors f présente un point selle en a.

iv. Si $0 \in H_f(a)$ (autrement dit si $H_f(a)$ n'est pas inversible), on ne peut pas conclure.

On ne peut pas conclure si $0 \in \mathrm{Sp}(H_f(a))$ car, en prenant $\vec{u} \in \mathrm{Ker}((H_f(a)))$, on a $f(a+t\vec{u}) - f(a) = o(t^2)$ et il n'est pas possible de savoir si la surface est au dessous ou au dessus de la valeur au point en se déplaçant dans la direction de $\vec{u}$.

Si la formulation précédente est la plus naturelle, on lui préfèrera (en accord avec le programme en **PT**), la version ci-dessous.

Théorème 14.23. **Condition suffisante locale de second ordre, version 2**

Soient f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U}$ et $a \in \mathcal{U}$ un point critique de f.

i. Si $\det(H_f(a)) > 0$, alors f présente un extremum local en a.
Dans ce cas :
- ✗ Si $\mathrm{Tr}(H_f(a)) > 0$, alors f présente un minimum local en a.
- ✗ Si $\mathrm{Tr}(H_f(a)) < 0$, alors f présente un maximum local en a.

ii. Si $\det(H_f(a) < 0$, alors f présente un point col en a.

iii. Si $\det(H_f(a) = 0$ on ne peut pas conclure.

Exercice 14.12.

Soit $f(x,y) = 3xy - x^3 - y^3$. Justifier que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ puis, déterminer, si ils existent, les extremums locaux de f et préciser leurs natures.

Exercice 14.13.

Soit la fonction f définie par : $f(x,y) = x\left((\ln(x))^2 + y^2\right)$.

1. Préciser le domaine de définition de f et le représenter. Justifier qu'il s'agit d'un ouvert de $\mathbb{R}^2$.
2. Montrer que f est de classe $\mathcal{C}^2$ sur ce domaine.
3. Montrer que f admet un seul extremum local. Préciser sa nature et sa valeur.
4. Cet extremum est-il global ?

Remarque 14.26. **Que faire lorsque $0 \in \mathrm{Sp}(H_f(a)$)**

Dans le cas où 0 est valeur propre de la hessienne au point critique, on se laissera souvent guider par l'énoncé.
Une idée naturelle reste d'étudier (le signe) de la fonction $t \mapsto f(a+t\vec{u}) - f(a)$ où $\vec{u} \in \mathrm{Ker}((H_f(a))$ et de comparer avec le signe de l'autre valeur propre.
Un développement limité de g à un ordre supérieur ou égal à 3 permet parfois de conclure...

Exercice 14.14.
Déterminer la nature de l'unique point critique a de f définie sur $\mathbb{R}^2$ par

$$f(x,y) = (x-y)^2 + e^{x+y} - 1 - (x+y) - \frac{1}{2}(x+y)^2.$$

On commencera par montrer que $H_f(a)$ n'est pas inversible et on obtiendra un développement limité à l'ordre 3 en 0 de la fonction $t \mapsto f(a+t\vec{u}) - f(a)$, où $\vec{u}$ est choisi dans le noyau de la hessienne.

14.4.3 Recherche d'*extrema* sur un fermé borné

Théorème 14.24. **Extrema globaux sur un fermé borné**
Soient $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue et $K \subset \mathbb{R}^2$ une partie fermée et bornée de $\mathbb{R}^2$. Alors, f admet des *extrema* globaux sur K.

Méthode 14.24.
Soient K un fermé borné de $\mathbb{R}^n$ et f une fonction continue sur K et de classe $\mathcal{C}^2$ sur $\mathring{K}$. On **sait** donc (par théorème) que f admet minimum et maximum sur K.

- ✗ On recherche les points critiques de f sur $\mathring{K}$ car ce sont les seuls points intérieurs à K où f peut présenter un extremum.
- ✗ On étudie les *extrema* de f sur le bord ∂K.
- ✗ Si f admet des *extrema* sur $\mathring{K}$, on les compare avec ceux sur ∂K. Si f n'admet aucun extremum sur $\mathring{K}$, c'est nécessairement sur ∂K.

Exercice 14.15.
Rechercher les *extrema* globaux de $f : (x,y) \mapsto x^2 - y^2$ sur le disque fermé de centre O et de rayon 1.

Exercice 14.16.
Soit T est l'ensemble des couples $(x,y) \in \mathbb{R}^2$ solutions du système d'inéquations

$$x \geq \frac{1}{4}, \quad y \geq \frac{1}{4}, \quad x+y \leq \frac{3}{4}.$$

Soit f la fonction définie sur T par

$$f(x,y) = \frac{1}{x} + \frac{1}{y} - \frac{2}{x+y}.$$

1. Représenter, sur un même graphique, les domaines T et $\mathring{T}$. Justifier qu'ils sont tous les deux bornés.

2. **a**. Justifier que f est de classe $\mathcal{C}^2$ sur $\mathring{T}$ et continue sur T.

b. Déterminer les dérivées partielles d'ordre 1 sur $\mathring{T}$ de la fonction f.
c. Montrer que f n'admet pas de point critique sur $\mathring{T}$.

3. Montrer alors que f admet un maximum et un minimum sur T, atteint sur ∂T et dont on précisera les valeurs.

14.5 Exemples de résolution d'équations aux dérivées partielles

On ne propose pas de résultats ni de méthode générale de résolution d'équations aux dérivées partielles (EDP).
On proposera uniquement des exemples, à travers différents exercices. On pourra toutefois observer que les résultats suivants sont immédiats. Si $\mathcal{U}$ désigne l'ouvert $I \times J$ où I et J sont deux intervalles ouverts et f une solution des équations aux dérivées partielles proposées.

- ✗ $\forall (x,y) \in \mathcal{U}, \quad \frac{\partial f}{\partial x}(x,y) = 0 : \quad \forall (x,y) \in \mathcal{U}, \quad f(x,y) = \varphi(y) \quad$ avec $\quad \varphi : J \to \mathbb{R}$
- ✗ $\forall (x,y) \in \mathcal{U}, \quad \frac{\partial f}{\partial x}(x,y) = g(x,y) : \quad \forall (x,y) \in \mathcal{U}, \quad f(x,y) = \int g(x,y)\mathrm{d}x + \varphi(y) \quad$ avec $\quad \varphi : J \to \mathbb{R}$
- ✗ $\forall (x,y) \in \mathcal{U}, \quad \frac{\partial^2 f}{\partial x^2}(x,y) = 0 : \quad \forall (x,y) \in \mathcal{U}, \quad f(x,y) = \varphi(y)x + \psi(y) \quad$ avec $\quad \varphi, \psi : J \to \mathbb{R}$.

Exercice 14.17.

Le but de cet exercice est de déterminer toutes les fonctions f de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ telles que

$$\forall (x,y) \in \mathbb{R}^2,\ \frac{\partial f}{\partial x}(x,y) + \frac{\partial f}{\partial y}(x,y) = 1.$$

1. Montrer que $\Psi : (x,y) \mapsto (x+y, x-y)$ réalise une bijection de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ sur un ensemble à préciser, et expliciter Ψ^{-1}.

2. Déterminer les solutions f du problème posé qui sont de la forme $f : (x,y) \mapsto g(x+y, x-y)$, où $g \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$.

3. Conclure.

On dit que l'on a résolu l'équation $\dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial y} = 1$ *en posant* $(u,v) = (x+y, x-y)$.

Exercice 14.18.

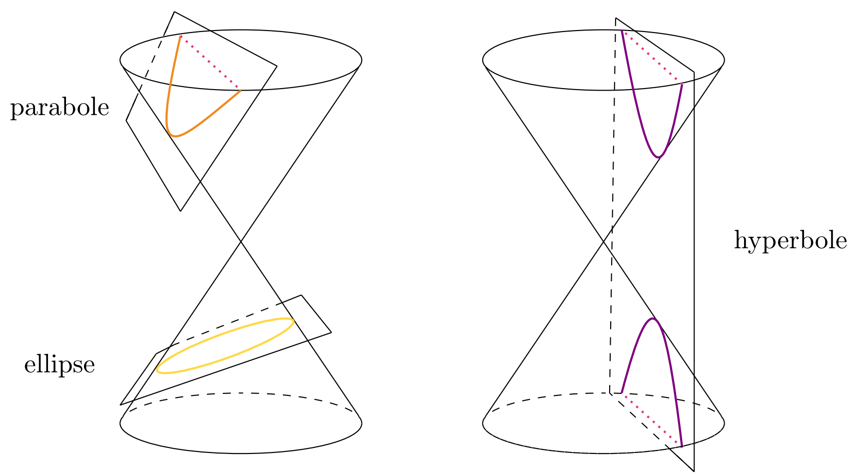
Résoudre l'équation $x\dfrac{\partial f}{\partial y} - y\dfrac{\partial f}{\partial x} = f$ sur $\mathcal{D} = \{(x,y) \in \mathbb{R}^2;\ x > 0\}$ en passant en coordonnées polaires, c'est-à-dire en posant $(x,y) = (r\cos\theta, r\sin\theta)$.

15 Coniques et courbes implicites du plan

Introduction

Les coniques sont des courbes planes étudiées depuis l'antiquité, possédant de nombreuses applications (par exemple en astronomie et en optique).
On peut définir de plusieurs façons les coniques. On peut par exemple montrer que ce sont les courbes du plan obtenues en réalisant l'intersection d'un cône de révolution et d'un plan (excepté qu'on obtient ainsi quelques cas dégénérés qu'on ne rencontre pas avec la définition *monofocale*, qui est la définition que l'on choisit).

Illustration 15.4. **Les coniques comme section d'un cône de révolution par un plan**



On rappelle quelques définitions et résultats. On se place dans un repère orthonormé $\mathcal{R} = (O, \overrightarrow{\imath}, \overrightarrow{\jmath})$ du plan.

Définition 15.1. **Distance d'un point à une partie non vide**

Soit $\mathcal{U}$ une partie non vide du plan et A un point. On note $\mathrm{d}(A,\mathcal{U})$ la borne inférieure des distances entre A et un point de $\mathcal{U}$:

$$\mathrm{d}(A,\mathcal{U}) = \inf\{AM : M \in \mathcal{U}\}.$$

Il existe un unique point H de $(\mathcal{D})$ tel que $\mathrm{d}(A,(\mathcal{D})) = AH$; ce point H est appelé le projeté orthogonal de A sur $(\mathcal{D})$.

Proposition. 15.1. **Calcul d'une distance**

Soient $(\mathcal{D})$ une droite de vecteur directeur u, d'équation cartésienne $\alpha x + \beta y + \gamma = 0$, n un vecteur normal à $(\mathcal{D})$ et $A(x_A, y_A)$ un point. Soit B un point de $(\mathcal{D})$. On a :

$$\mathrm{d}(A,(\mathcal{D})) = \frac{\left|\det\left(u, \overrightarrow{BA}\right)\right|}{\|u\|} = \frac{\left|\langle n, \overrightarrow{BA}\rangle\right|}{\|n\|} = \frac{|\alpha x_A + \beta y_A + \gamma|}{\sqrt{\alpha^2+\beta^2}}.$$

15.1 Définition des coniques par foyer, directrice et excentricité

Définition 15.2. **Coniques, foyer, directrice et excentricité**

Soit $e > 0$, F un point et $(\mathcal{D})$ une droite du plan tels que $F \notin (\mathcal{D})$.
L'ensemble $(\mathcal{C})$ des points M du plan tels que

$$\mathrm{d}(M, F) = e \times \mathrm{d}(M, (\mathcal{D}))$$

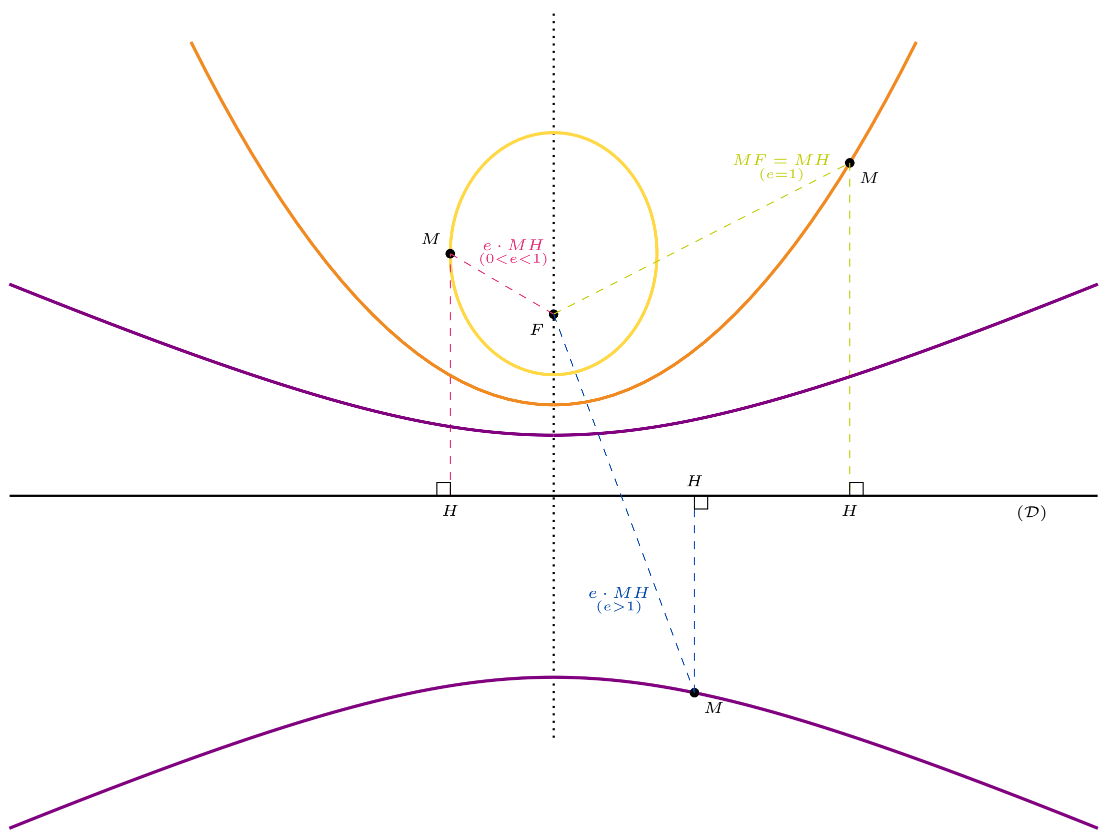
est appelé la **conique** de **foyer** F, de **directrice** $(\mathcal{D})$ et d'**excentricité** e.

Cette conique est appelée:

- ✗ une **parabole** si $e = 1$;
- ✗ une **ellipse** si $0 < e < 1$;
- ✗ une **hyperbole** si $e > 1$.

Illustration 15.5. **Trois coniques définies par un même foyer et une même droite directrice**

On représente sur la figure ci-dessous trois coniques : une ellipse ($e = 1/2$), une parabole et une hyperbole ($e = 2$), ayant le même foyer F et la même droite directrice $\mathcal{D}$.
H est le projeté orthogonal de M sur $\mathcal{D}$.



Remarque 15.1.

- ✗ Si H est le projeté orthogonal de M sur $(\mathcal{D})$, on a $\mathrm{d}(M, F) = e \cdot \mathrm{d}(M, \mathcal{D}) \iff MF = eMH$.
- ✗ Les coniques sont clairement les lignes de niveau de l'application $M \mapsto \dfrac{\mathrm{d}(M, F)}{\mathrm{d}(M, \mathcal{D})}$.

Exercice 15.1.

On se place dans un repère orthonormé $\mathcal{R} = (O, \overrightarrow{\imath}, \overrightarrow{\jmath})$ du plan.
Donner l'équation cartésienne de la parabole de directrice $(\mathcal{D}) : 2x + y = 1$ et de foyer $F(1, 1)$.

15.1.1 L'axe focal

Définition 15.3. **Axe focal**

Soit $(\mathcal{C})$ une conique de foyer F, de directrice $(\mathcal{D})$ et d'excentricité e.
La perpendiculaire à $(\mathcal{D})$ passant par F est appelée l'**axe focal** de la conique.

Théorème 15.2.

L'axe focal d'une conique en est un axe de symétrie.

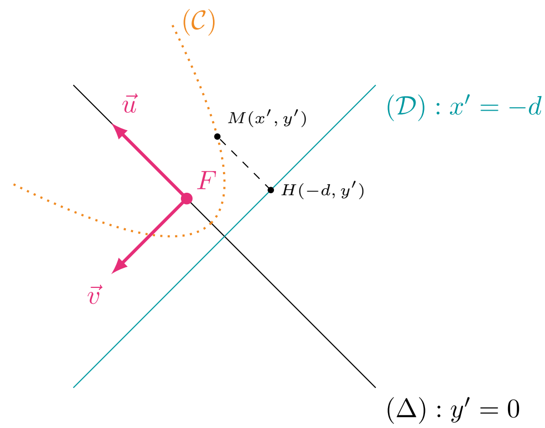
Preuve. Soit $(\mathcal{C})$ une conique de foyer F, de directrice $(\mathcal{D})$ et d'axe focal (Δ). Pour $M \in (\mathcal{C})$, note $M' = \mathcal{S}_{(\Delta)}(M)$ le symétrique de M par rapport à (Δ). Alors

$$\mathrm{d}(M', F) = \mathrm{d}(M, F) \qquad \text{et} \qquad \mathrm{d}(M', (\mathcal{D})) = \mathrm{d}(M, (\mathcal{D}))$$

donc $M' \in (\mathcal{C})$. $\square$

On peut alors définir un repère orthonormé **direct** $\mathcal{R}_1 = (F, \vec{(u)}, \vec{(v)})$ centré en F avec $\vec{u}$ vecteur unitaire directeur de l'axe focal (Δ).
On note $d = \mathrm{d}(F, (\mathcal{D}))$.
Notant (x', y') les coordonnées d'un point dans $\mathcal{R}_1$:

- ✗ (Δ) a pour équation $y' = 0$;
- ✗ $(\mathcal{D})$ a pour équation $x' = -d$;
- ✗ Le projeté orthogonal de $M(x', y')$ sur $(\mathcal{D})$ a pour coordonnées $(-d, y')$.
- ✗ $\mathrm{d}(M, (\mathcal{D})) = \mathrm{d}(M, F) \iff x'^2 + y'^2 = e^2(x' + d)^2$.



Définition 15.4. **Sommets d'une conique**

On appelle **sommets** d'une conique ses points d'intersection avec son axe focal.

Théorème 15.3.

Soit $(\mathcal{C})$ une conique de foyer F, de directrice $(\mathcal{D})$, d'excentricité e et d'axe focal (Δ).
Soit K l'intersection de $(\mathcal{D})$ et de (Δ). On a les résultats suivants.

i. Les paraboles possèdent un unique sommet, qui est le milieu de $[FK]$.
ii. Les ellipses possèdent exactement deux sommets, situés de part et d'autre du foyer mais du même côté de la directrice.
iii. Les hyperboles possèdent exactement deux sommets, situés du même côté par rapport au foyer, mais de part et d'autre de la directrice.

Preuve. On se place dans le repère $\mathcal{R}_1$ introduit ci-avant. Un point $M(x', y')$ de la conique appartient à l'axe focal si et seulement si

$$\begin{cases} x'^2 + y'^2 &= e^2(x'+d)^2 \\ y' &= 0 \end{cases} \iff \begin{cases} x'^2 &= e^2(x'+d)^2 \\ y' &= 0 \end{cases} \iff \begin{cases} (1-e^2)x'^2 - 2de^2x' - e^2d^2 &= 0 \\ y' &= 0 \end{cases}$$

- ✗ Pour $e = 1$ (une parabole), on a alors $-2de^2x' - e^2d^2 = 0$ si et seulement si $x' = -\frac{d}{2}$. Ainsi, une parabole possède un unique sommet de coordonnées $(-\frac{d}{2}, 0)$ dans $\mathcal{R}_1$: c'est bien le milieu de $[FK]$;
- ✗ Pour $e > 0$ avec $e \neq 1$ (ellipse ou hyperbole), on a

$$(1-e^2)x'^2 - 2de^2x' - e^2d^2 = 0 \iff x' = \frac{-de}{1+e} \quad \text{ou} \quad x' = \frac{de}{1-e}.$$

On trouve bien les coordonnées de deux sommets S et S'. Plus précisément :

- ✓ Pour une ellipse, comme $0 < e < 1$, S et S' sont du même côté de $(\mathcal{D})$, et S est entre K et F;
- ✓ Pour une hyperbole, comme $e > 1$, S et S' sont des deux côtés de $(\mathcal{D})$, et S est entre K et F.

$\square$

15.2 Réduction des équations algébriques des coniques

Dans cette section, nous allons donner l'**équation réduite** de chacune des coniques: il s'agit d'une équation cartésienne dans un repère bien choisi (de sorte que l'équation soit particulièrement simple). Cela permettra de tracer aisément la conique.

Dans chaque cas, on se place dans le repère réduit qui **n'est pas le même** que le repère $\mathcal{R}_1$ précédemment explicité.

15.2.1 Paraboles

Définition 15.5. **Repère réduit de la parabole**

Soit $(\mathcal{C})$ une parabole de directrice $(\mathcal{D})$ et de foyer F. Soient (Δ) son axe focal, et K l'intersection de $(\mathcal{D})$ et de (Δ). On appelle **repère réduit associé à la parabole** $(\mathcal{C})$, le repère orthonormé direct $\mathcal{R}_0 = (S, \overrightarrow{u}, \overrightarrow{v})$ où :

- ✗ S est le sommet de la parabole ;
- ✗ $\overrightarrow{u} = \dfrac{\overrightarrow{SF}}{\|\overrightarrow{SF}\|}$.

On pose encore $p = KF = \mathrm{d}(F, (\mathcal{D}))$, qui est appelé le **paramètre de la parabole**.

Théorème 15.4. **Équation réduite d'une parabole**

Dans son repère réduit $\mathcal{R}_0$, une parabole $(\mathcal{C})$ admet pour équation

$$y^2 = 2px.$$

Cette équation est appelée l'**équation réduite** de la parabole.

Théorème 15.5. **Paramétrage**

Un paramétrage d'une parabole dans son repère réduit est $\begin{cases} x(t) = \dfrac{t^2}{2p} \\ y(t) = t \end{cases} \quad t \in \mathbb{R}.$

Illustration 15.6. **Équation réduite d'une parabole**

On représente la parabole $(\mathcal{C})$ dans son repère associé $\mathcal{R}_0 = (S, \overrightarrow{u}, \overrightarrow{v})$.
Dans ce repère, $(\mathcal{C})$ admet pour équation (réduite)

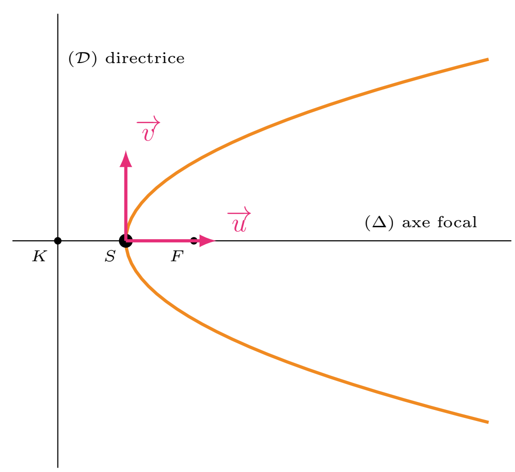
$$y^2 = 2px.$$

- ✗ Le sommet S de $(\mathcal{C})$ est l'origine du repère $\mathcal{R}_0$.
- ✗ Le foyer F de $(\mathcal{C})$ a pour coordonnées $\left(\frac{p}{2}; 0\right)$.
- ✗ La directrice a pour équation cartésienne :

$$(\mathcal{D}) : x = -\frac{p}{2}.$$

- ✗ Une paramétrisation $f : \mathbb{R} \to \mathbb{R}^2$ de $\mathcal{P}$ est donnée par

$$f(t) = \begin{pmatrix} \dfrac{t^2}{2p} \\ t \end{pmatrix}.$$



Remarque 15.2.

Si l'on interverti $\overrightarrow{u}$ et $\overrightarrow{v}$ dans le repère, on obtient l'équation $x^2 = 2py$.

Remarque 15.3.

Un autre paramétrage possible est $\begin{cases} x(t) = 2pt^2 \\ y(t) = 2pt \end{cases} \quad t \in \mathbb{R}.$

Dans son repère réduit, une parabole $(\mathcal{C})$ est la réunion des courbes représentatives des fonctions $x \mapsto \sqrt{2px}$ et $x \mapsto -\sqrt{2px}$, ce qui permet d'en déduire le tracé directement (on pourrait également étudier la représentation paramétrique).

Exercice 15.2.

Déterminer une équation réduite de la parabole de foyer F dont les coordonnées sont $F(1,1)$ dans $\mathcal{R} = (O, \overrightarrow{\imath}, \overrightarrow{\jmath})$ et de directrice $(\mathcal{D})$ d'équation cartésienne $(\mathcal{D}) : x - y = 1$ dans $\mathcal{R}$.
Donner les coordonnées dans $\mathcal{R}$ du sommet de la parabole.

15.2.2 Ellipses

D'après la discussion qui précède (et qui fait suite à la définition d'axe focal), on peut réécrire l'équation de la conique dans $\mathcal{R}_1 = (F, \vec{u}, \vec{v})$. Observant que $0 < e < 1$:

$$\begin{aligned}
x'^2 + y'^2 = e^2 (x' + d)^2 &\iff x'^2 (1 - e^2) - 2dx'e^2 + y'^2 = e^2 d^2 \iff (1 - e^2) \left[x'^2 - \frac{2de^2}{1 - e^2} x' \right] + y'^2 = e^2 d^2 \\
&\iff (1 - e^2) \left[\left(x' - \frac{de^2}{1 - e^2} \right)^2 - \frac{d^2 e^4}{(1 - e^2)^2} \right] + y'^2 = e^2 d^2 \\
&\iff (1 - e^2) \left(x' - \frac{de^2}{1 - e^2} \right)^2 + y'^2 = \frac{e^2 d^2 (1 - e^2) + d^2 e^4}{1 - e^2} \\
&\iff (1 - e^2) \left(x' - \frac{de^2}{1 - e^2} \right)^2 + y'^2 = \frac{e^2 d^2}{1 - e^2} \\
&\iff \frac{\left(x' - \frac{de^2}{1 - e^2} \right)^2}{\left(\frac{ed}{1 - e^2} \right)^2} + \frac{y'^2}{\left(\frac{ed}{\sqrt{1 - e^2}} \right)^2} = 1.
\end{aligned}$$

En posant

$$\begin{cases} \tilde{x} & = & x' - \dfrac{de^2}{1 - e^2} \\ \tilde{y} & = & y' \end{cases} \quad \text{et} \quad \begin{cases} a & = & \dfrac{ed}{1 - e^2} \\ b & = & \dfrac{ed}{\sqrt{1 - e^2}} \end{cases},$$

et Ω le point de coordonnées $\left(\dfrac{de^2}{1 - e^2}, 0 \right)$ dans $\mathcal{R}_1$, on peut définir un nouveau repère $\mathcal{R}_0 = (\Omega, \vec{u}, \vec{v})$ dans lequel l'ellipse a pour équation

$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} = 1.$$

On observe qu'il découle des calculs précédents que Ω est le milieu de $[SS']$ où S et S' sont les deux sommets de l'ellipse obtenus précédemment.

Définition 15.6. **Repère réduit de l'ellipse**

Soit $(\mathcal{E})$ une ellipse de directrice $(\mathcal{D})$, de foyer F et d'excentricité $e \in]0, 1[$. Soient (Δ) son axe focal, et K l'intersection de $(\mathcal{D})$ et de (Δ).
Soient S, S' ses sommets (avec $S \in [KF]$) et Ω le milieu de $[SS']$. On appelle **repère réduit associé à l'ellipse** $(\mathcal{E})$, le repère orthonormé direct $\mathcal{R}_0 = (\Omega, \overrightarrow{u}, \overrightarrow{v})$ où : $\vec{u}$ est défini par

$$\vec{u} = \frac{-\overrightarrow{\Omega F}}{\|\overrightarrow{\Omega F}\|}.$$

Le point Ω est appelé **centre** de l'ellipse.

Théorème 15.6. **Équation réduite d'une ellipse**

Dans son repère réduit $\mathcal{R}_0$, une ellipse $\mathcal{E}$ admet pour équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

avec $a = \Omega S = \Omega S'$. Notant $c = \Omega F = ea$, on a b qui vérifie $b > 0$ et $a^2 = b^2 + c^2$.
Le nombre a est appelé le **demi-grand axe** et le nombre b le **demi-petit axe**.

De plus les quatre points S, S', T et T' de coordonnées respectives (dans $\mathcal{R}_0$) $(-a, 0), (a, 0), (0, b), (0, -b)$, qui sont les points d'intersection de l'ellipse avec des deux axes de symétrie, sont appelés **sommets de l'ellipse**.

L'excentricité de l'ellipse vérifie $e = \sqrt{\dfrac{a^2 - b^2}{a^2}}$.

Le choix du repère réduit est fait de telle sorte que $a > b$, ce qui justifie la terminologie demi-**grand** axe et demi-**petit** axe.
On pourrait définir un autre repère (parfois également appelé *réduit*) dans lequel $b > a$.

Théorème 15.7. **Paramétrage**

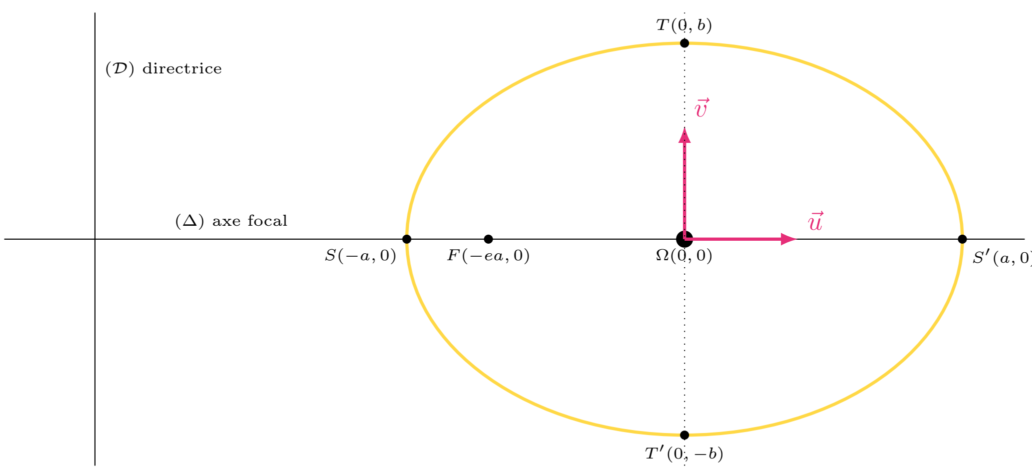
Un paramétrage d'une ellipse $(\mathcal{E})$ d'excentricité e dans son repère réduit est $\begin{cases} x(t) = a\cos(t) \\ y(t) = b\sin(t) \end{cases} \quad t \in \mathbb{R}.$

L'étude de ce paramétrage permet de tracer l'ellipse.
Attention, dans ce paramétrage, t ne désigne pas l'angle polaire du point $M(t) \in (\mathcal{E})$.

Illustration 15.7. **Équation réduite d'une ellipse**

On représente l'ellipse $(\mathcal{E})$ d'excentricité e dans son repère réduit associé $\mathcal{R}_0 = (\Omega, \vec{u}, \vec{v})$.
Dans ce repère, $(\mathcal{E})$ admet pour équation (réduite)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$



- ✗ Le milieu Ω de $[SS']$ est l'origine du repère $\mathcal{R}_0$ et le centre de l'ellipse.
- ✗ Le foyer F de $(\mathcal{E})$ a pour coordonnées $(-ae; 0)$.
- ✗ La directrice a pour équation cartésienne :

$$(\mathcal{D}) : x = -\frac{a^2}{\sqrt{a^2 - b^2}}$$

- ✗ Une paramétrisation $f : \mathbb{R} \to \mathbb{R}^2$ de $\mathcal{E}$ est donnée par

$$f(t) = \begin{pmatrix} a\cos(t) \\ \\ b\sin(t) \end{pmatrix}.$$

Exercice 15.3.
Déterminer une équation réduite de l'ellipse de foyer F dont les coordonnées sont $F(1,1)$ dans $\mathcal{R} = (O, \vec{i}, \vec{j})$, de directrice $(\mathcal{D})$ d'équation cartésienne $(\mathcal{D}) : x - y = 1$ dans $\mathcal{R}$ et d'excentricité $e = \frac{1}{2}$. Donner les coordonnées dans $\mathcal{R}$ du centre et des sommets de l'ellipse.

15.3 Hyperboles

On reprend les calculs précédent dont la conclusion va différer de celle pour l'ellipse, car $e > 1$. Toujours dans le repère $\mathcal{R}_1 = (F, \vec{u}, \vec{v})$,

$$x'^2 + y'^2 = e^2 \left(x' + d\right)^2 \iff \left(1 - e^2\right)\left(x' - \frac{de^2}{1-e^2}\right)^2 + y'^2 = \frac{e^2d^2}{1-e^2}$$
$$\iff \frac{\left(x' + \frac{de^2}{e^2-1}\right)^2}{\left(\frac{ed}{e^2-1}\right)^2} - \frac{y'^2}{\left(\frac{ed}{\sqrt{e^2-1}}\right)^2} = 1.$$

En posant

$$\begin{cases} \tilde{x} &= x' + \dfrac{de^2}{e^2-1} \\ \tilde{y} &= y' \end{cases} \quad \text{et} \quad \begin{cases} a &= \dfrac{ed}{e^2-1} \\ b &= \dfrac{ed}{\sqrt{e^2-1}} \end{cases},$$

et Ω le point de coordonnées $\left(-\dfrac{de^2}{e^2-1}, 0\right)$ dans $\mathcal{R}_1$, on peut définir un nouveau repère $\mathcal{R}_0 = (\Omega, \vec{u}, \vec{v})$ dans lequel l'ellipse a pour équation

$$\frac{\tilde{x}^2}{a^2} - \frac{\tilde{y}^2}{b^2} = 1.$$

Le point Ω est encore le milieu du segment $[SS']$ où S et S' sont les deux sommets de l'hyperbole.

Définition 15.7. **Repère réduit d'une hyperbole**
Soit $(\mathcal{H})$ une hyperbole de directrice $(\mathcal{D})$, de foyer F et d'excentricité e.
Soient S, S' ses sommets, et Ω le milieu de $[SS']$.

On appelle **repère réduit associé à l'hyperbole** $\mathcal{H}$, le repère orthonormé direct $\mathcal{R}_0 = (\Omega, \vec{u}, \vec{v})$ où $\vec{u} = \dfrac{\vec{\Omega F}}{\|\vec{\Omega F}\|}$.

Les axes d'équation (dans $\mathcal{R}_0$) $x = 0$ et $y = 0$ sont des axes de symétrie pour l'hyperbole et le point d'intersection, Ω, s'appelle le **centre** de l'hyperbole.

Théorème 15.8. **Équation réduite d'une hyperbole**
Dans son repère réduit $\mathcal{R}_0$, une hyperbole $\mathcal{H}$ admet pour équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

avec $a = \Omega S$, $c = \Omega F = ea$ et $b > 0$ tel que $a^2 + b^2 = c^2$.

L'excentricité de l'ellipse vérifie $e = \sqrt{\dfrac{a^2+b^2}{a^2}}$.

Théorème 15.9. **Paramétrage de l'hyperbole**
Un paramétrage d'une hyperbole dans son repère $\mathcal{R}_0$ est

$$\begin{cases} x(t) = \varepsilon a \operatorname{ch}(t) \\ y(t) = b \operatorname{sh}(t) \end{cases} \quad (\varepsilon, t) \in \{-1, 1\} \times \mathbb{R}$$

Théorème 15.10. **Asymptotes de l'hyperbole**
Dans le repère réduit $\mathcal{R}_0 = (\Omega, \vec{u}, \vec{v})$ les asymptotes de l'hyperbole ont pour équation $y = \pm\dfrac{b}{a}x$.

Preuve. Lorsque $t \to +\infty$, on a :

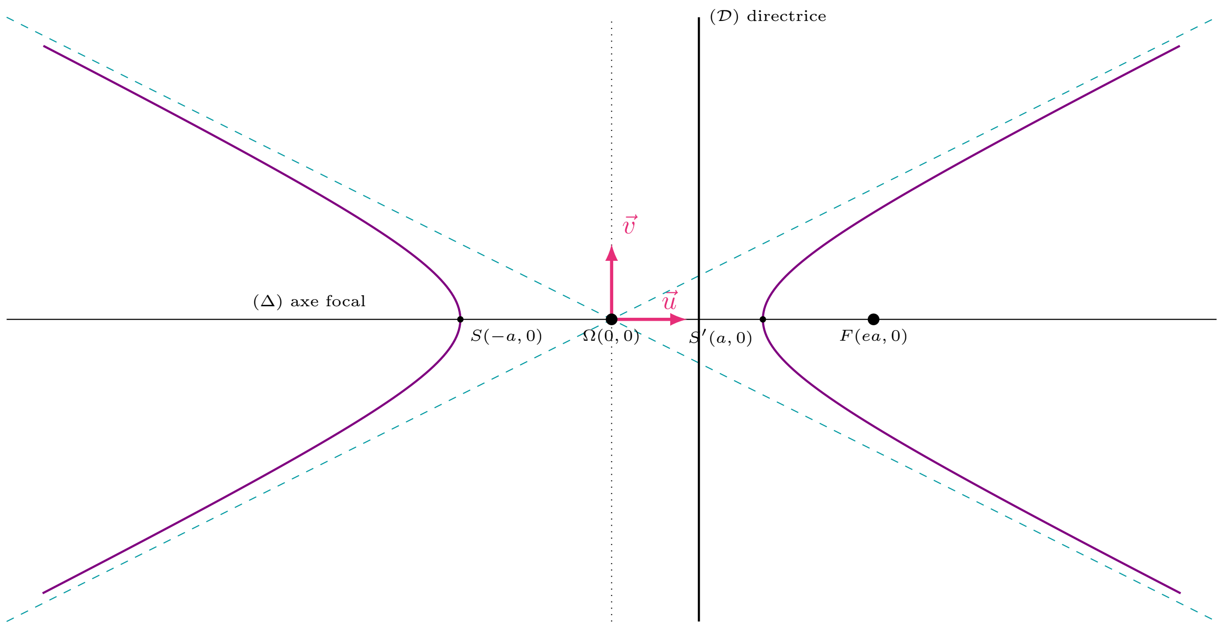
$$\frac{b\mathrm{sh}(t)}{a\mathrm{ch}(t)} \to \frac{b}{a}, \quad \text{et} \quad b\mathrm{sh}(t) - \frac{b}{a} \times a\mathrm{ch}(t) = -be^{-t} \to 0.$$

On utilise alors les propriétés de symétrie de l'hyperbole pour obtenir la second asymptote. □

Illustration 15.8. **Équation réduite d'une hyperbole**

On représente l'hyperbole $(\mathcal{H})$ d'excentricité e dans son repère réduit associé $\mathcal{R}_0 = (\Omega, \overrightarrow{u}, \overrightarrow{v})$.
Dans ce repère, $(\mathcal{H})$ admet pour équation (réduite)

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$



- ✗ Le milieu Ω de $[SS']$ est l'origine du repère $\mathcal{R}_0$ et le centre de l'hyperbole.
- ✗ Le foyer F de $(\mathcal{E})$ a pour coordonnées $(ae; 0)$.
- ✗ La directrice a pour équation cartésienne :

$$x = \frac{a^2}{\sqrt{a^2+b^2}}.$$

- ✗ Les sommets S et S' ont pour coordonnées $(-a, 0)$ et $(a, 0)$.
- ✗ Les asymptotes ont pour équation

$$y = \pm \frac{b}{a} x.$$

- ✗ Une paramétrisation $f : \mathbb{R} \to \mathbb{R}^2$ de chaque branche de $\mathcal{E}$ est donnée par

$$f(t) = \begin{pmatrix} \pm a\mathrm{ch}(t) \\ \\ b\mathrm{sh}(t) \end{pmatrix}.$$

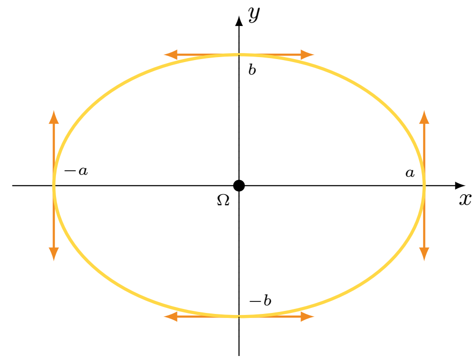
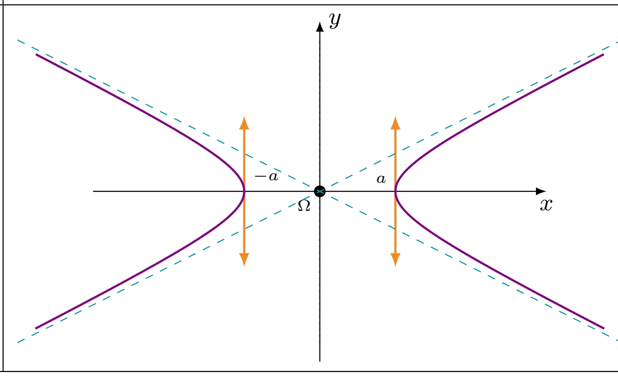
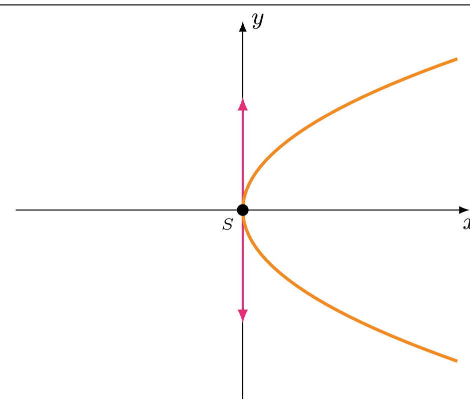
Exercice 15.4.

Déterminer une équation réduite de l'hyperbole de foyer F dont les coordonnées sont $F(1,1)$ dans $\mathcal{R} = (O, \vec{i}, \vec{j})$, de directrice $(\mathcal{D})$ d'équation cartésienne $(\mathcal{D}) : x - y = 1$ dans $\mathcal{R}$ et d'excentricité $e = 2$. Donner les coordonnées dans $\mathcal{R}$, du centre et des sommets de l'hyperbole.

Exercice 15.5.

On dit qu'une hyperbole est équilatère si ses deux asymptotes sont perpendiculaires. Quelle est l'excentricité d'une hyperbole équilatère?

Tableau de synthèse

Nature	Équation réduite	Paramétrage	Représentation
Ellipse	$\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$	$\begin{cases} x(t) &= a\cos(t) \\ y(t) &= b\sin(t)\end{cases}$	
Hyperbole	$\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$	$\begin{cases} x(t) &= \pm a\mathrm{ch}(t) \\ y(t) &= b\mathrm{sh}(t)\end{cases}$	
Parabole	$y^2=2px$	$\begin{cases} x(t) &= \dfrac{t^2}{2p} \\ y(t) &= t\end{cases}$	

15.4 Courbes implicites du plan

Dans la fin de ce chapitre, nous allons établir quelques propriétés élémentaires des courbes du plan définies par une équation cartésienne.
Puis nous démontrerons que les coniques sont les courbes planes dont l'équation est polynomiale de degré 2 (hormis quelques cas dégénérés).

15.4.1 Définition et résultats élémentaires

Définition 15.8. **Courbe implicite du plan**

i. On appelle **courbe implicite du plan** toute une courbe (Γ) admettant dans un repère une équation de la forme $F(x,y)=0$, où F est une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

ii. Dans le cas où F est une fonction polynomiale de degré $k\in\mathbb{N}^*$, on dit que (Γ) est une **courbe algébrique** de degré k.

iii. Γ est dite **régulière** en un point (x,y) si $\nabla F(x,y)\neq 0$.
Une courbe implicite régulière en tout point est dite régulière.

Le résultat suivant sera généralisé et discuté en géométrie de l'espace dans le **Chapitre 18**.
En particulier, une courbe (implicite) régulière admet donc une tangente en tout point.

Théorème 15.11.
Toute courbe implicite régulière possède localement (c'est-à-dire au voisinage de chacun de ses points) un paramétrage régulier $\begin{cases} x = x(t) \\ y = y(t) \end{cases}$ de classe $\mathcal{C}^1$.

Preuve. Résultat admis. □

Théorème 15.12.
Soit (Γ) une courbe implicite régulière d'équation $F(x, y) = 0$. Alors en tout point (x_0, y_0) de (Γ), le gradient $\nabla F(x_0, y_0)$ est normal à la tangente.

Preuve. Soit (x_0, y_0) un point de (Γ) fixé. D'après le théorème précédent, au voisinage de (x_0, y_0), (Γ) admet un paramétrage régulier

$$M : t \mapsto \begin{cases} x = x(t) \\ y = y(t) \end{cases}, \qquad \text{avec } M(t_0) = (x_0, y_0).$$

La tangente à (Γ) en $M(t_0)$ est dirigée par

$$\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) = \begin{pmatrix} x'(t_0) \\ y'(t_0) \end{pmatrix}.$$

Posons $f : t \mapsto F(x(t), y(t))$. Comme f est nulle, sa dérivée l'est aussi ; notamment en $t = t_0$, on a, par la règle de la chaîne,

$$0 = f'(t_0) = \frac{\partial F}{\partial x}(x(t_0), y(t_0))x'(t_0) + \frac{\partial F}{\partial y}(x(t_0), y(t_0))y'(t_0)$$

ce qui se réécrit

$$\left\langle \nabla F(x(t_0), y(t_0)) \middle| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) \right\rangle = 0,$$

ce qu'on cherchait à obtenir. □

Exercice 15.6.
Déterminer l'équation de la tangente au point $(1, 1)$ à l'ellipse d'équation $4x^2 + y^2 = 5$.

Corollaire 15.13.
Soit $F : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. En tout point régulier d'une ligne de niveau de F, le gradient est orthogonal au vecteur vitesse.

Preuve. Il suffit de reprendre la preuve précédente, f reste constante. □

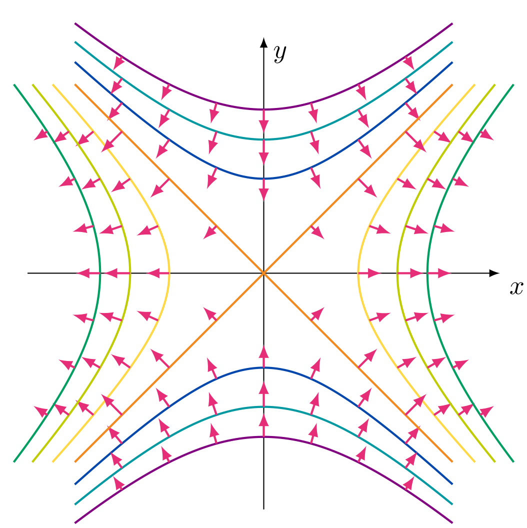
Exemple 15.1.

On retiendra que *lorsqu'il est non nul, le gradient de F est orthogonal aux lignes de niveau et orienté dans le sens des valeurs croissantes de F.*

On pourrait même démontrer que le gradient indique la direction de plus grande pente de la surface représentative de F, c'est-à-dire la direction dans laquelle F augmente le plus vite.

On représente ci-contre quelques lignes de niveaux k et les gradients de la fonction

$$g : (x, y) \in \mathbb{R}^2 \longmapsto x^2 - y^2, \qquad k \in [\![-3, 3]\!].$$



Exercice 15.7.
Tracer quelques lignes de niveaux et placer quelques gradients de la fonction
$$g:\ (x,y)\in\mathbb{R}_+^*\times\mathbb{R}\longmapsto\frac{y}{x}.$$

15.4.2 Le cas particulier des courbes algébriques de degré deux

Définition 15.9. **Courbe algébrique de degré 2**
On appelle *courbe algébrique de degré deux* toute courbe (Γ) du plan admettant dans un repère une équation polynomiale de degré 2 de la forme
$$\overbrace{\underbrace{ax^2+2bxy+cy^2}_{q(x,y)}+\underbrace{dx+ey}_{\ell(x,y)}+f}^{F(x,y)}=0$$
où a,b,c,d,e,f sont des réels, tels que $(a,b,c)\neq(0,0,0)$.

- ✗ $q(x,y)=ax^2+2bxy+cy^2$ est appelé *la partie quadratique* de l'équation $F(x,y)=0$.
- ✗ $\ell(x,y)=dx+ey$ est appelé *la partie linéaire* de l'équation $F(x,y)=0$.

À cette équation est associée la matrice symétrique réelle $A=\begin{pmatrix}a&b\\b&c\end{pmatrix}$, appelée *matrice de la partie quadratique.*

Remarque 15.4.

- ✗ En posant $X=\begin{pmatrix}x\\y\end{pmatrix}$ on voit que l'équation $ax^2+2bxy+cy^2+dx+ey+f=0$ s'écrit matriciellement
$$X^\top AX+\begin{pmatrix}d&e\end{pmatrix}X+f=0$$
où A est la matrice de la partie quadratique.
- ✗ A étant symétrique réelle, elle est donc diagonalisable dans une base orthonormée (d'après le théorème spectral). Cela permettra dans la suite de simplifier l'équation de la courbe en changeant de repère.

Exemple 15.2.
Toutes les courbes suivantes sont des courbes algébriques de degré 2.

i. Les cercles, puisqu'ils possèdent pour équation $(x-\alpha)^2+(y-\beta)^2=R^2$.
ii. Les réunions de deux droites, puisqu'elles possèdent pour équation cartésienne $(ax+by+c)(a'x+b'y+c')=0$.
iii. Les paraboles, d'équation réduite $y^2=2px$.
iv. Les ellipses, d'équation réduite $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$.
v. Les hyperboles, d'équation réduite $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$.

Théorème 15.14. **Intersection d'une courbe algébrique de degré deux et d'une droite**
Soit $(\mathcal{D})$ une droite et (Γ) une courbe algébrique de degré deux, telles que $(\mathcal{D})$ ne soit pas incluse dans (Γ). Alors, l'intersection de $(\mathcal{D})$ et de (Γ) contient au plus deux points.

Preuve. Il suffit d'injecter une paramétrisation de la droite dans l'équation de la courbe. Le problème revient donc à chercher les racines d'un polynôme de degré 2 : il y en a au plus deux. □

15.5 Réduction des équations des courbes algébriques de degré 2

Jusqu'à la fin de ce chapitre on considère une courbe algébrique de degré deux (Γ), d'équation
$$F(x,y)=ax^2+2bxy+cy^2+dx+ey+f=0,$$

et on appelle $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ la matrice symétrique associée.

Nous allons expliquer dans ce qui va suivre comment réduire l'équation d'une courbe algébrique de degré 2 pour montrer que, hormis quelques cas dits *dégénérés*, chacune de ces courbes est une conique de l'un des trois types suivants: parabole, ellipse ou hyperbole.

Ces étapes sont présentées ci-après dans le détail, elles reviennent à changer de repère successivement jusqu'à obtenir une équation réduite.

15.5.1 Première étape : diagonalisation de A dans une b.o.n.d et réduction de la partie quadratique

Dans le cas où A est déjà diagonale, on a alors un premier cas très simple.
En effet, l'équation $ax^2 + cy^2 + dx + ey + f = 0$ a pour membre de gauche la somme de deux polynômes de degré 2 dont la mise sous forme canonique est aisée et permet après un changement de repère (par translation) d'obtenir une équation réduite qui entraîne l'identification de la conique.

Exercice 15.8.
Montrer que les équations suivantes sont celles de coniques dont on précisera la nature. Représenter ensuite celles-ci graphiquement.
i. $(\Gamma) : \ x^2 + 4y^2 - 2x - 16y + 8 = 0.$
ii. $(\Gamma') : \ 4y^2 + 2x + 16y + 14 = 0.$

On s'inspire de ce premier cas simple pour déduire le cas général.
La matrice $A \in \mathcal{S}_2(\mathbb{R})$ est symétrique réelle et donc A est diagonalisable dans une base orthonormée directe de vecteurs propres (on note $\mathrm{Sp}(A) = \{\lambda, \mu\}$) :

$$\exists P \in SO_2(\mathbb{R}) : A = PDP^\top \text{ avec } D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Ainsi :

$$X^\top AX = X^\top PDP^\top X = \left(P^\top X\right)^\top D\left(P^\top X\right) = (X')^\top DX' \text{ avec } X' = P^\top X.$$

On pose $X' = P^\top X = \begin{pmatrix} x' \\ y' \end{pmatrix}$ et l'équation de (Γ) devient alors

$$\lambda x'^2 + \mu y'^2 + \delta x' + \zeta y' + f = 0.$$

La matrice $P \in SO_2(\mathbb{R})$ s'interprète également comme matrice de passage entre deux b.o.n.d de $\mathbb{R}^2$. On a effectué une rotation sur le repère initial $\mathcal{R} = (O, \vec{i}, \vec{j})$.
On a obtenu un nouveau repère $\mathcal{R}' = \left(O, \vec{i}', \vec{j}'\right)$ dans lequel l'expression de la conique est plus simple.

Théorème 15.15.
Soit $k \in \mathbb{R}^*$. L'équation $xy = k$ définit une hyperbole dont les asymptotes sont les axes du repère.

Exercice 15.9.
Démontrer le théorème ci-dessus.

15.5.2 Deuxième étape : mise sous forme canonique et réduction de la partie linéaire

Puisque $(a, b, c) \neq (0, 0, 0)$ la matrice A est non nulle.
Ses valeurs propres ne sont pas toutes deux nulles : $(\lambda, \mu) \neq (0, 0)$. Distinguons alors trois cas.

✗ Supposons $\lambda \neq 0$ et $\mu \neq 0$. L'équation $(\mathcal{C}')$ s'écrit alors :

$$\lambda\left(x' + \frac{\delta}{2\lambda}\right)^2 + \mu\left(y' + \frac{\zeta}{2\mu}\right)^2 + \left(f - \frac{\delta^2}{4\lambda} - \frac{\zeta^2}{4\mu}\right) = 0.$$

On pose $X'' = \begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} x' + \frac{\delta}{2\lambda} \\ y' + \frac{\zeta}{2\mu} \end{pmatrix} = X' + \begin{pmatrix} \frac{\delta}{2\lambda} \\ \frac{\zeta}{2\mu} \end{pmatrix}$ et on obtient l'équation réduite :

$$\lambda x'^2 + \mu y''^2 = g.$$

On a introduit un **nouveau repère** $\mathcal{R}'' = \left(\Omega, \vec{i}'', \vec{j}''\right)$ par translation du vecteur de coordonnées $\begin{pmatrix} \frac{\delta}{2\lambda} \\ \frac{\zeta}{2\mu} \end{pmatrix}$ dans $\mathcal{R}' = \left(O, \vec{i}', \vec{j}'\right)$.

Le centre Ω de ce nouveau repère a pour coordonnées $\left(-\dfrac{\delta}{2\lambda}, -\dfrac{\xi}{2\mu}\right)$ dans $\mathcal{R}'$.
Observons que Ω est alors **centre de symétrie** pour la courbe.
Pour retrouver ses coordonnées dans le repère initial $\mathcal{R}$, on utilise les formules de changement de bases :

$$0 = X'' = X' + \begin{pmatrix} \frac{\delta}{2\lambda} \\ \frac{\zeta}{2\mu} \end{pmatrix} = P^\top X + \begin{pmatrix} \frac{\delta}{2\lambda} \\ \frac{\zeta}{2\mu} \end{pmatrix} \text{ donc } X_\Omega = P \begin{pmatrix} -\frac{\delta}{2\lambda} \\ -\frac{\zeta}{2\mu} \end{pmatrix} = PX'_\Omega.$$

✗ Supposons $\lambda \neq 0$ et $\mu = 0$. On obtient alors de manière analogue une équation réduite :

$$\lambda x''^2 + \zeta y'' = g.$$

✗ Supposons $\mu \neq 0$ et $\lambda = 0$. On obtient alors de manière analogue une équation réduite :

$$\mu y''^2 + \delta x'' = g.$$

15.5.3 Troisième étape : identification

On a obtenu après rotation et translation du repère initial les équations réduites suivantes :

Valeurs Propres de A	$\lambda, \mu \neq 0$	$\lambda \neq 0, \mu = 0$	$\lambda = 0, \mu \neq 0$
Équation réduite dans $\mathcal{R}''$	$\lambda x^2 + \mu y^2 = g$	$\lambda x^2 + \zeta y = g$	$\mu y^2 + \delta x = g$

i. $\boldsymbol{\lambda > 0}$ et $\boldsymbol{\mu > 0}$, les trois situations suivantes se présentent:

✗ Si $g > 0$ alors l'équation réduite peut s'écrire :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

La courbe obtenue est appelée **ellipse**.

✗ Si $g = 0$ alors le seul point solution de l'équation réduite est $(0, 0)$. La conique est réduite au centre du repère.

✗ Si $g < 0$ alors l'équation réduite n'a pas de solution.

Le cas $\boldsymbol{\lambda < 0}$ et $\boldsymbol{\mu < 0}$ est similaire.

ii. Si $\boldsymbol{\lambda > 0}$ et $\boldsymbol{\mu < 0}$, deux situations peuvent se présenter :

✗ Si $g \neq 0$ alors l'équation réduite peut s'écrire :

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \pm 1.$$

La conique obtenue est appelée **hyperbole.**

✗ Si $g = 0$ alors l'équation réduite peut s'écrire

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0 \Longleftrightarrow \left(\frac{x}{a} - \frac{y}{b}\right)\left(\frac{x}{a} + \frac{y}{b}\right) = 0.$$

La conique obtenue est la réunion de deux droites sécantes.

Le cas $\boldsymbol{\lambda < 0}$ et $\boldsymbol{\mu > 0}$ est analogue.

iii. Si $\boldsymbol{\lambda \neq 0}$ et $\boldsymbol{\mu = 0}$. Alors:

✗ Si $\zeta \neq 0$ alors l'équation réduite peut s'écrire $x^2 = 2py$. La conique obtenue est appelée **parabole.**

✗ Si $\zeta = 0$ alors l'équation réduite devient $\lambda x^2 = g$. Suivant le signe de $r = \frac{g}{\lambda}$ on obtient :

✓ la réunion de deux droites parallèles ($r > 0$; d'équations $x = \pm\sqrt{r}$)
✓ une droite ($r = 0$; d'équation $x = 0$)
✓ ou bien l'ensemble vide ($r < 0$).

iv. Si $\boldsymbol{\lambda} = \mathbf{0}$ et $\boldsymbol{\mu} \neq \mathbf{0}$. Alors:

✗ Si $\delta \neq 0$ alors l'équation réduite peut s'écrire : $y^2 = 2px$. La conique obtenue est une **parabole.**

✗ Si $\delta = 0$ alors l'équation réduite devient $\mu y^2 = g$. La nature dépend du signe de $r = \frac{g}{\mu}$.
✓ la réunion de deux droites parallèles ($r > 0$; d'équations $y = \pm\sqrt{r}$)
✓ une droite ($r = 0$; d'équation $y = 0$)
✓ ou bien l'ensemble vide ($r < 0$).

15.6 Classification des courbes algébriques de degré 2

Théorème 15.16. **Classification des courbes algébriques de degré 2**

On considère la courbe algébrique (Γ) de degré 2 d'équation $F(x,y) = ax^2 + 2bxy + cy^2 + dx + ey + f = 0$ avec $(a,b,c) \neq (0,0,0)$.

On note $\Delta = ac - b^2 = \lambda\mu = \det(A)$ où λ, μ sont les valeurs propres de $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

i. Si $\Delta > 0$ alors (Γ) est une ellipse, un point ou l'ensemble vide.
ii. Si $\Delta = 0$ alors (Γ) est une parabole, une droite, la réunion de deux droites parallèles ou l'ensemble vide.
iii. Si $\Delta < 0$ alors (Γ) est une hyperbole ou la réunion de deux droites sécantes.

Si la courbe obtenue est une ellipse, une parabole ou une hyperbole, on parle de **conique propre**. Dans tous les autres cas, on parle de **conique dégénérée**.

Remarque 15.5.
Une conique dégénérée peut avoir une infinité de centres de symétrie (s'il s'agit de deux droites parallèles par exemple), ce qui n'est pas possible avec la définition de conique par foyer et droite directrice.

Exercice 15.10.
Déterminer la nature et les éléments caractéristiques de la conique (Γ) : $x^2 + y^2 + 2xy + 2x + 4y + 3 = 0$.

16 Intégrales à paramètres

16.1 Introduction : problèmes d'interversion

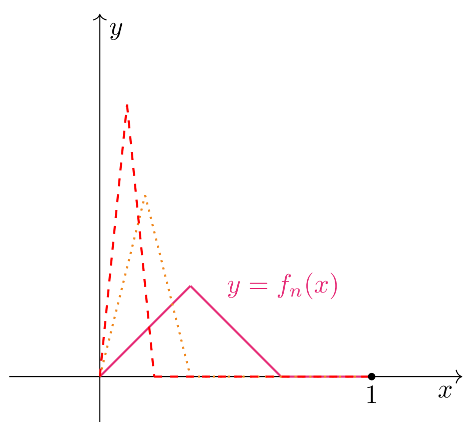
Ce chapitre, intitulé **Intégrales à paramètres** aurait pu tout aussi bien être intitulé **Interversion de limites et d'intégrales**.

En effet, l'étude de la régularité d'une fonction est une étude de limite.
On a déjà à plusieurs reprises rencontré des quantités exprimées sous forme d'intégrales qui dépendaient d'un (ou de plusieurs) paramètres : un entier (suite d'intégrales), un réel (fonction définie par une intégrale) ou des deux (comme le reste intégral dans la formule de Taylor).
Il est naturel de vouloir comprendre l'évolution de ces quantités quand le paramètre varie, voire d'en déterminer une éventuelle limite.

Hélas, on ne peut pas tout simplement intervertir les symboles $\displaystyle\int$ et lim sans s'assurer que des conditions, que l'on présentera ci-après, sont satisfaites.

En effet, il suffit de considérer cet exemple pour s'en convaincre. Soit (f_n) la suite de fonctions (affines par morceaux donc) continues sur $[0,1]$ définie par

$$\forall n \in \mathbb{N}^*, \quad \forall x \in [0,1], \qquad f_n(x) = \begin{cases} n^2 x, & \text{si } 0 \leq x \leq \frac{1}{n} \\ 2n - n^2 x, & \text{si } \frac{1}{n} < x \leq \frac{2}{n} \\ 0, & \text{si } x > \frac{2}{n} \end{cases}$$



Il est alors clair que, pour tout $n \in \mathbb{N}^*$, $\displaystyle\int_0^1 f_n(x)\mathrm{d}x = 1$.
Néanmoins, si $x \in]0,1]$ est fixé, il existe N tel que, si $n \geq N$, $x > \frac{2}{n}$ et donc $f_n(x) = 0$. Ainsi, $\displaystyle\lim_{n\to+\infty} f_n(x) = 0$.
Ceci reste vrai si $x = 0$ car $f_n(0) = 0$. On observe donc ici que

$$1 = \lim_{n\to+\infty} \int_0^1 f_n(x)\mathrm{d}x \neq \int_0^1 \left(\lim_{n\to+\infty} f_n(x)\right)\mathrm{d}x = 0.$$

Il est donc important de prendre quelques précautions.

On peut d'ailleurs rappeler quelques résultats d'interversion déjà établis dans les chapitres précédents, qu'il ne faudra surtout pas confondre avec ce qui suit.

Théorème 7.32. **Permutation somme et intégrale**

Soit $(f_n)_{n\geq 0}$ une suite de fonctions intégrables sur un même intervalle I telles que, pour tout $t \in I$, la série $\sum\limits_{n\geq 0} f_n(t)$ est convergente. On note alors

$$f : t \in I \longmapsto \sum_{n=0}^{+\infty} f_n(t).$$

On suppose que :

i. f est continue sur I, sauf éventuellement aux bords de I ;

ii. La série $\sum\limits_{n\geq 0} \int_I |f_n(t)|\mathrm{d}t$ est convergente.

Alors, f est intégrable sur I et

$$\int_I f(t)\mathrm{d}t = \sum_{n=0}^{+\infty} \int_I f_n(t)\mathrm{d}t.$$

Et le cas particulier pour l'intégration de la somme d'une série entière :

Théorème 10.12. **Intégration terme à terme**

Soit $\sum a_n x^n$ une série entière de rayon de convergence $R > 0$.

On note $f : x \longmapsto \sum\limits_{n=0}^{+\infty} a_n x^n$ sa fonction somme et F une primitive de la fonction $f \in \mathcal{C}^0(] - R; R[)$.

La série entière $\sum \dfrac{a_n}{n+1} x^{n+1}$ est une série entière de rayon de convergence R et

$$\forall x \in] - R, R[, F(x) = F(0) + \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}.$$

Autrement dit, pour pour tout $x \in] - R; R[$,

$$F(x) - F(0) = \int_0^x f(t)\mathrm{d}t = \int_0^x \sum_{n=0}^{+\infty} a_n t^n \mathrm{d}t = \sum_{n=0}^{+\infty} \int_0^x a_n t^n \mathrm{d}t = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

Dans le **Chapitre 14**, lors du petit aperçu sur la résolution d'équations aux dérivées partielles, on a pu rencontrer la situation suivante :

$$\frac{\partial}{\partial x} f(x, y) = g(x, y) \Longrightarrow f(x, y) = \int g(x, y)\mathrm{d}x + \varphi(y),$$

où f et g étaient deux fonctions de deux variables (à valeurs réelles) définies et $\mathcal{C}^1$ sur un domaine du type $I \times J$, et φ une fonction quelconque, de classe $\mathcal{C}^1$ définie sur J.
Apparaît alors un exemple d'intégrale dont l'*intégrande* ne dépend pas que de la variable d'intégration. Il est possible qu'on ne sache pas explicitement calculer cette intégrale.
On donne dans ce chapitre quelques éléments d'étude de ces fonctions.

Dans tout le chapitre, les lettres I et J désignent des intervalles quelconques de $\mathbb{R}$ (non nécessairement bornés), de sorte que I soit non trivial (non vide et non réduit à un point) et J non vide.
Notre objectif est, pour g une fonction définie sur $I \times J$ et à valeurs dans $\mathbb{R}$, d'étudier la régularité de la fonction f définie par une intégrale dite à paramètre :

$$\forall x \in I, \quad f(x) = \int_J g(x, t)\mathrm{d}t.$$

16.2 Définitions, rappels

On *rappelle* le résultat ci-dessous, déjà présenté sous une autre forme dans le **Chapitre 7** d'intégration et qui découle du Théorème fondamental de l'analyse.

Théorème 16.1.

On considère deux fonctions u, v dérivables sur I et à valeurs dans J, ainsi qu'une fonction h continue sur J et à valeurs dans $\mathbb{R}$.
Alors la fonction $f : I \to \mathbb{R}$ définie par :

$$\forall x \in I,\ f(x) = \int_{u(x)}^{v(x)} h(t)\mathrm{d}t$$

est dérivable sur I, de sorte que

$$\forall x \in I,\ f'(x) = v'(x) \times h\Big(v(x)\Big) - u'(x) \times h\Big(u(x)\Big).$$

Dans ce chapitre on s'intéresse à des fonctions définies par une intégrale où la variable se trouve dans l'*intégrande* (et non dans les bornes comme ci-dessus).

Définition 16.1. **Intégrale à paramètre**

On appelle **intégrale à paramètre** toute fonction définie par une relation de la forme

$$\forall x \in I,\ f(x) = \int_J g(x,t)\mathrm{d}t,$$

où $g : I \times J \to \mathbb{R}$ est une fonction de deux variables.

Exemple 16.1. **Transformée de Laplace**

Si $f : \mathbb{R}_+ \to \mathbb{R}$, sa *transformée de Laplace* est définie par $\widehat{f}(x) = \displaystyle\int_0^{+\infty} e^{-xt} f(t)\mathrm{d}t$ (sous réserve d'existence), c'est donc une intégrale à paramètre.

Méthode 16.25.

L'étude d'une intégrale à paramètre commence toujours par l'étude de son domaine de définition:

- ✗ si J est bornée on vérifiera que, pour tout $x \in I$, la fonction d'une variable $t \mapsto g(x,t)$ est continue sur J (on aura alors affaire à une intégrale ordinaire);
- ✗ sinon, l'intégrale est impropre, et on devra alors prouver sa convergence par les méthodes usuelles (par exemple en utilisant le théorème de comparaison).

Exercice 16.1.

Montrer que $f : x \mapsto \displaystyle\int_0^{+\infty} \cos(xt)e^{-t^2}\mathrm{d}t$ est définie pour tout $x \in \mathbb{R}$.

Exercice 16.2.

Étudier le domaine de définition de $f :\ x \mapsto= \displaystyle\int_0^1 \frac{e^{2t}}{t+x}\mathrm{d}t$, où $x \in \mathbb{R}_+$.

Pour étudier la continuité ou la dérivabilité des intégrales à paramètre, nous aurons besoin de l'hypothèse suivante.

Définition 16.2. **Hypothèse de domination**

Soit $g : I \times J \to \mathbb{R}$ une fonction de deux variables. On dit que g vérifie l'**hypothèse de domination** sur $I \times J$ si et seulement si il existe une fonction $\varphi :\ J \to \mathbb{R}$ (de la seule variable d'intégration et donc **indépendante de** l'autre), positive et intégrable sur J telle que:

$$\forall (x,t) \in I \times J,\ |g(x,t)| \leq \varphi(t).$$

Exemple 16.2.

$g : (x,t) \mapsto \cos(xt)e^{-t^2}$ vérifie l'hypothèse de domination sur $\mathbb{R} \times \mathbb{R}_+$. En effet,

- ✗ $\forall (x,t) \in \mathbb{R} \times \mathbb{R}_+, \quad , \left|\cos(xt)e^{-t^2}\right| \leq e^{-t^2}$;
- ✗ $\forall t \in \mathbb{R}_+, \quad e^{-t^2} \geq 0$;

✗ $\int_0^{+\infty} e^{-t^2} \mathrm{d}t$ est une intégrale convergente (autrement dit $t \mapsto e^{-t^2}$ est intégrable sur $\mathbb{R}_+$.

Cette hypothèse de domination assure alors que $f(x) = \int_0^{+\infty} g(x,t)\mathrm{d}t$ est définie sur $\mathbb{R}$, en vertu du théorème de comparaison.

16.3 Continuité, dérivabilité d'une intégrale à paramètre

Commençons par motiver ce qui va suivre par une observation.
L'application $f : x \mapsto \int_0^{+\infty} x\mathrm{e}^{-xt}\, \mathrm{d}t$ est définie sur $\mathbb{R}_+$ (intégrale de référence, pour $x > 0$).

Pour tout $t \geqslant 0$, la fonction $x \mapsto x\mathrm{e}^{-xt}$ est continue sur $\mathbb{R}_+$ et pourtant, f **n'est pas continue** puisque un calcul direct montre que $f(0) = 0$ et $f(x) = 1 \to 1$ lorsque $x \to 0, x > 0$.

Ainsi, la continuité, à $t \in J$ fixé, de $x \mapsto g(x,t)$ ne suffit pas à garantir la continuité de f. Ici,

$$\lim_{x\to 0} \int_0^{+\infty} x\mathrm{e}^{-xt}\, \mathrm{d}t \neq \int_0^{+\infty} \left(\lim_{x\to 0} x\mathrm{e}^{-xt}\right) \mathrm{d}t.$$

16.3.1 Continuité d'une intégrale à paramètre

Théorème 16.2. **Continuité sous le signe $\int$**

Soit $g : I \times J \to \mathbb{R}$ une fonction de deux variables telle que:

i. **Caractère $\mathcal{C}^0$ - étude « en x »**

✗ Pour tout $t \in J$, la fonction $x \mapsto g(x,t)$ est continue sur I;

ii. **Intégrabilité (domination) - étude « en t »**

✗ Pour tout $x \in I$, la fonction $t \mapsto g(x,t)$ est continue sur J;
✗ g vérifie l'hypothèse de domination sur $I \times J$.

Alors, $f : x \mapsto \int_J g(x,t)\mathrm{d}t$ définit une fonction continue sur I.

Preuve. Résultat admis. □

Remarque 16.1.

✗ Si g est continue sur $I \times J$, alors elle l'est nécessairement par rapport à chacune de ses deux variables. **Mais la réciproque est fausse**.
L'hypothèse ici faite sur les fonctions d'une variable est donc plus faible.

✗ L'intégrabilité de $t \mapsto g(x,t)$ sur J résulte de l'hypothèse de domination sur g, en vertu du théorème de comparaison, il n'est donc pas nécessaire de la prouver explicitement si on a vérifié l'hypothèse.

✗ Le programme officiel précise qu'aux concours, il n'est pas nécessaire de justifier l'hypothèse relative à la continuité « par rapport à t » mais on doit quand même citer que la propriété est satisfaite.

Remarque 16.2. **Machine à intervertir $\lim$ et $\int$**

Le théorème de continuité sous le signe $\int$ doit être interprété comme une *machine à intervertir* les symboles lim et $\int$. En effet, si les hypothèses sont satisfaites, pour $x_0 \in I$,

$$\begin{aligned}\lim_{x\to x_0}\int_J g(x,t)\mathrm{d}t &= \lim_{x\to x_0} f(x) = f(x_0) \qquad \text{car } f \text{ continue en } x_0\\ &= \int_J g(x_0,t)\mathrm{d}t = \int_J \left(\lim_{x\to x_0} g(x,t)\right)\mathrm{d}t \qquad \text{car, pour tout } t\in J,\ x\mapsto g(x,t) \text{ continue en } x_0\end{aligned}$$

Exercice 16.3.

Montrer que $f : x \mapsto \displaystyle\int_0^{+\infty} \cos(xt)e^{-t^2}\mathrm{d}t$ est continue sur $\mathbb{R}$.

Méthode 16.26. **Domination locale**

La continuité étant une notion locale, il suffit pour prouver la continuité de f sur I, de prouver que f est continue en chaque point de l'intervalle. Il suffit donc d'appliquer le théorème précédent au voisinage de chaque point, et ainsi il n'est pas nécessaire de vérifier l'hypothèse de domination sur I tout entier. Lorsque l'on raisonne de la sorte, on dit que l'on vérifie **l'hypothèse de domination locale**. On peut remplacer l'hypothèse de domination sur $I \times J$ par

✗ Pour tout $[\alpha,\beta] \subset I$, g vérifie l'hypothèse de domination sur $[\alpha,\beta]\times J$.

Attention cependant, il faudra systématiquement rédiger l'intégralité de l'argumentation en écrivant par exemple, pour $I =]a,b[$,

$$]a,b[\;=\; \bigcup_{a<\alpha<\beta<b} [\alpha,\beta],$$

donc, pour tout $x \in]a,b[$, il existe $[\alpha,\beta] \subset]a,b[$ tels que $x \in [\alpha,\beta]$. Or, f est continue sur $[\alpha,\beta]$ (car ce sera sur cet intervalle que l'on appliquera le théorème de continuité) donc f est continue en x. Ceci étant vrai pour tout $x \in]a,b[$, f est continue sur $]a,b[$.

Exercice 16.4.

Montrer que $f : x \mapsto \displaystyle\int_0^1 \frac{e^{2t}}{t+x}\mathrm{d}t$ est continue sur $\mathbb{R}_+^*$.

Exercice 16.5. **Fonction Gamma d'Euler**

1. Soit $x > 0$. Montrer que la fonction $t \mapsto t^{x-1}e^{-t}$ est intégrable sur $]0,+\infty[$.

On pose alors, pour tout $x \in]0,+\infty[$, $\Gamma(x) = \displaystyle\int_0^{+\infty} t^{x-1}e^{-t}\mathrm{d}t$.

2. Montrer que Γ est continue sur $]0,+\infty[$.
3. Pour tout $x > 0$, exprimer $\Gamma(x+1)$ en fonction de $\Gamma(x)$.
4. Que vaut, pour $n \in \mathbb{N}^*$, $\Gamma(n)$?
5. Montrer que, pour tout $x > 0$, $\Gamma(x) > 0$.

Le résultat suivant n'est pas au programme, il faut donc **le redémontrer systématiquement pour l'utiliser.**
On rappelle qu'un segment désigne un intervalle **fermé et borné** de $\mathbb{R}$.

Corollaire 16.3. **Cas particulier sur un segment, HP**

Soit $g : I \times J \to \mathbb{R}$ une fonction de deux variables telle que:

i. $(x,t) \mapsto g(x,t)$ est continue sur $I \times J$ (**en tant que fonction de deux variables**);
ii. $J = [a,b]$ est un segment.

Alors, $f : x \mapsto \displaystyle\int_J g(x,t)\mathrm{d}t$ définit une fonction continue sur I.

Preuve. On vérifie qu'on est dans le cadre d'application du théorème en vérifiant les hypothèses :

i. **Caractère $\mathcal{C}^0$ - étude « en x »**

✗ g continue sur $I \times [a,b]$ donc, pour tout $t \in [a,b]$, $x \mapsto g(x,t)$ est continue.

ii. **Intégrabilité - étude « en t »**

✗ g continue sur $I \times [a,b]$ donc, pour tout $x \in I$, $t \mapsto g(x,t)$ est continue.
✗ L'hypothèse de domination est vérifiée sur tout segment $[c,d]$ inclus dans I. En effet, $[c,d] \times [a,b]$ étant un ensemble fermé et borné de $\mathbb{R}^2$, la continuité de g justifie l'existence de $M_{c,d} \in \mathbb{R}_+$ telle que :

$$\forall (x,t) \in [c,d] \times [a,b], \quad |f(x,t)| \leqslant M_{c,d} \rightsquigarrow \text{ indépendante de } x \text{ positive et intégrable sur } [a,b]$$

On peut donc appliquer le théorème de continuité sous le signe $\displaystyle\int$. □

Méthode 16.27.

Si l'intégrale n'est pas impropre, on peut donc se passer de l'hypothèse de domination pour prouver la continuité de f, sous réserve que g soit continue sur $I \times J$ (et plus seulement par rapport à chacune de ses deux variables x et t). Mais il faut redémontrer qu'on peut le faire.

Exercice 16.6.

Utiliser le **Corollaire 16.3** pour prouver que $f : x \mapsto \displaystyle\int_0^1 \frac{e^{2t}}{t+x}\mathrm{d}t$ est continue sur $\mathbb{R}_+^*$.

16.3.2 Dérivabilité d'une intégrale à paramètre

Théorème 16.4. **Dérivation sous le signe $\displaystyle\int$**

Soit $g : I \times J \to \mathbb{R}$ une fonction de deux variables telle que:

i. **Caractère $\mathcal{C}^1$ - étude « en x »**

✗ Pour tout $t \in J$, la fonction $x \mapsto g(x,t)$ est $\mathcal{C}^1$ sur I;

ii. **Intégrabilité (domination) - étude « en t »**

✗ Pour tout $x \in I$, la fonction $t \mapsto g(x,t)$ est intégrable sur J;
✗ Pour tout $x \in I$, $t \mapsto \dfrac{\partial}{\partial x} g(x,t)$ est $\mathcal{C}^0$ sur J ;
✗ $\dfrac{\partial}{\partial x} g$ vérifie l'hypothèse de domination sur $I \times J$.

Alors $f : x \mapsto \displaystyle\int_J g(x,t)\mathrm{d}t$ définit une fonction $\mathcal{C}^1$ sur I, de sorte que

$$\forall x \in I, \ f'(x) = \int_J \frac{\partial}{\partial x} g(x,t)\mathrm{d}t.$$

Preuve. Résultat admis. □

Exercice 16.7.

Montrer que $f : x \mapsto \int_0^{+\infty} \cos(xt)e^{-t^2}\mathrm{d}t$ est $\mathcal{C}^1$ sur $\mathbb{R}$. En déduire une expression explicite de f.

Méthode 16.28. **Domination locale**

La dérivabilité est une notion locale, tout comme la continuité. On peut donc utiliser ce théorème en vérifiant **l'hypothèse de domination locale** en tout point de I. C'est à dire qu'on vérifie l'hypothèse de domination pour $\frac{\partial}{\partial x}g$ sur tout domaine du type $[\alpha, \beta] \times J$, où $[\alpha, \beta] \subset I$.

Exercice 16.8. **Gamma, le retour**

On reprend les notations de l'**Exercice 16.5**.

Montrer que Γ est de classe $\mathcal{C}^1$ sur $]0, +\infty[$ et exprimer, pour $x > 0$, $\Gamma'(x)$ sous forme d'intégrale.

Comme précédemment, le résultat suivant n'est pas au programme, mais en exercice on peut le redémontrer (et on doit, le cas échéant, le faire) facilement pour l'utiliser.

Corollaire 16.5. **Cas particulier pour un segmet, HP**

Soit $g : I \times J \to \mathbb{R}$ une fonction de deux variables telle que:

i. $(x, t) \mapsto g(x, t)$ est $\mathcal{C}^1$ sur $I \times J$;
ii. $J = [a, b]$ est un segment.

Alors, $f : x \mapsto \int_J g(x, t)\mathrm{d}t$ définit une fonction $\mathcal{C}^1$ sur I, de sorte que

$$\forall x \in I,\ f'(x) = \int_J \frac{\partial}{\partial x} g(x, t)\mathrm{d}t.$$

Méthode 16.29.

Si l'intégrale n'est pas impropre, on peut donc se passer de l'hypothèse de domination pour prouver le caractère $\mathcal{C}^1$ de f, sous réserve que g soit $\mathcal{C}^1$ sur $I \times J$ et après avoir redémontré le corollaire ci-dessus.

Exercice 16.9. **Calcul de l'intégrale de Gauss**

Soit $f : x \mapsto \int_0^1 \frac{e^{-x^2(1+t^2)}}{1+t^2}\mathrm{d}t$.

1. Montrer que f est $\mathcal{C}^1$ sur $\mathbb{R}$.

2. Calculer $\lim_{x \to +\infty} f(x)$.

3. En déduire que : $\int_0^{+\infty} e^{-t^2}\mathrm{d}t = \frac{\sqrt{\pi}}{2}$.

17 Variables aléatoires discrètes

Dans ce chapitre nous étudions les applications définies sur un espace probabilisé dénombrable $(\Omega, \mathcal{T}, \mathbf{P})$, c'est-à-dire des quantités dont la valeur dépend de l'issue d'une expérience aléatoire. Contrairement a ce qui avait été fait en première année, on ne se limite donc plus au cas des variables aléatoires dont l'ensemble des valeurs possibles est fini.

De même que dans le chapitre sur les probabilités, on étudiera des phénomènes aléatoires sans décrire précisément l'espace probabilisé associé: la donnée *a priori* des probabilités des évènements considérés suffit, celles-ci étant données sous forme de loi d'une variable aléatoire.

Dans tout le chapitre, I désigne (sauf mention du contraire) une partie de $\mathbb{N}$, finie ou non. Tout ensemble de la forme $\{x_i : \ i \in I\}$ est alors fini ou dénombrable.
De même, toute famille de la forme $(x_i)_{i\in I}$ est finie ou dénombrable.

On admet qu'il existe une tribu, que l'on notera $\mathcal{B}$, sur $\mathbb{R}$ contenant tous les intervalles, appelée *tribu des boréliens*.

17.1 Notion de variable aléatoire discrète

17.1.1 Définition et premières propriétés

Définition 17.1. **Variable aléatoire discrète**

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble, et une application $X : \Omega \to E$. On dit que X est une **variable aléatoire discrète** si, pour tout $x \in E$, on a $X^{-1}(\{x\}) \in \mathcal{T}$ (*i.e.* si $X^{-1}(\{x\})$ est un évènement).
L'image $X(\Omega)$ de X sera appelée l'**univers image** de X ou l'ensemble des valeurs possibles de X.

Remarque 17.1. **Variable aléatoire discrète ou finie**

Soit X une variable aléatoire discrète définie sur un espace probabilisable $(\Omega, \mathcal{T})$.

- ✗ Si $X(\Omega)$ est un ensemble fini $\{x_1, \ldots, x_n\}$, on dit que X est une **variable aléatoire (discrète) finie**.
  Si l'univers Ω est fini, alors $X(\Omega)$ l'est également.
- ✗ Si $X(\Omega)$ est un ensemble infini dénombrable $\{x_i |\ i \in I\}$, on dit que X est une variable aléatoire **discrète infinie dénombrable**.

Exemple 17.1.

Une urne contient des boules numérotées de 1 à n.

- ✗ On pioche une boule au hasard dans cette urne et on note X la variable aléatoire qui prend la valeur du numéro de la boule piochée. On a alors $X(\Omega) = [\![1, n]\!]$ et X est une variable aléatoire finie.
- ✗ On considère la variable aléatoire Y qui prend la valeur 1 si et seulement si la boule piochée a un numéro pair et 0 sinon. Dans ce cas, $Y(\Omega) = \{0, 1\}$ et Y est encore une variable aléatoire finie.
- ✗ On pioche maintenant successivement et avec remise jusqu'à obtenir la boule numérotée 1 et on note Z la variable aléatoire qui prend pour valeur le nombre de pioches effectuées. On peut avoir la boule 1 dès la première pioche, mais on peut piocher arbitrairement longtemps une autre boule donc ici $Z(\Omega) = \mathbb{N}^*$ et cette fois Z est une variable aléatoire infinie dénombrable.

Définition 17.2. **Notations des évènements**

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, $X : \Omega \to E$ une variable aléatoire, et $k \in E$.

- ✗ L'évènement $X^{-1}(\{k\}) = \{\omega \in \Omega,\ X(\omega) = k\}$ sera noté $[X = k]$.
- ✗ Si $E \subset \mathbb{R}$, on pose encore:
  - l'évènement $X^{-1}(]-\infty, k]) = \{\omega \in \Omega,\ X(\omega) \leq k\}$ sera noté $[X \leq k]$;
  - l'évènement $X^{-1}(]-\infty, k[) = \{\omega \in \Omega,\ X(\omega) < k\}$ sera noté $[X < k]$.

  On définit de même les évènements $[X \geq k]$ et $[X > k]$.

Proposition. 17.1. **Système complet canonique associé à une variable aléatoire**

Soient $(\Omega, \mathcal{T})$ un espace probabilisable, et $X : \Omega \to E$ une variable aléatoire discrète dont l'univers image est $X(\Omega) = \{x_i |\ i \in I\}$. Alors la famille $(X = x_i)_{i \in I}$ est un système complet d'évènements pour Ω.

Cette dernière proposition est importante, car elle permet d'appliquer la formule des probabilités totales en *découpant* l'univers selon les valeurs que prend une variable aléatoire.

Exercice 17.1.

Une urne contient des boules, indiscernables au toucher, numérotées de 1 à n. On pioche successivement et avec remise des boules dans cette urne. On introduit la variable aléatoire X qui prend pour valeur le rang du premier tirage qui permet d'obtenir une boule avec un numéro supérieur ou égal au numéro de la boule piochée au premier au tirage.

1. Que vaut $X(\Omega)$?

2. Déterminer, pour tout $k \in X(\Omega)$, la probabilité $P([X = k])$.

Jusqu'à la fin de ce chapitre, on fixe un espace probabilisé $(\Omega, \mathcal{T}, \mathbf{P})$. Sauf mention du contraire, toutes les variables aléatoires considérées seront discrètes, définies sur $(\Omega, \mathcal{T})$ et à valeurs dans $\mathbb{R}$.

Remarque 17.2.

Par σ-additivité, on déduit alors, de la proposition précédente, $\mathbf{P}\left(\bigcup_{i \in I} [X = x_i]\right) = \sum_{i \in I} \mathbf{P}(X = x_i) = 1$.

Grâce à la probabilité donnée sur Ω, on peut définir la *loi* d'une variable aléatoire X, qui consiste à associer à chaque valeur que X peut prendre la probabilité que cette valeur soit prise.

Théorème & Définition 17.1. **Loi d'une variable aléatoire**

Soient X une variable aléatoire telle que $X(\Omega) = \{x_i :\ i \in I\}$ et $\mathcal{B} \subset \mathcal{P}(\mathbb{R})$ la tribu des boréliens. Alors, l'application définie par

$$\begin{array}{rccc} \mathbf{P}_X : & \mathcal{B} & \longrightarrow & [0,1] \\ & A & \longmapsto & \mathbf{P}(X \in A) \end{array}$$

est une probabilité sur $(\mathbb{R}, \mathcal{B})$, appelée **loi** de X.

Théorème & Définition 17.2. **Loi et distribution de probabilité**

Soit $\{x_i : i \in I\}$ un ensemble dénombrable de réels.
Une application

$$\begin{array}{ccc} \{x_i :\ i \in I\} & \to & \mathbb{R} \\ x_i & \mapsto & p_i \end{array}$$

est une **distribution de probabilité** si :

$$\forall i \in I,\ p_i \in [0,1] \qquad \text{et} \qquad \sum_{i \in I} p_i = 1.$$

On peut alors lui associer, de manière unique, une loi de probabilité sur $(\mathbb{R}, \mathcal{B})$ en posant

$$\begin{array}{rccc} \mathbf{P} : & \mathcal{B} & \longrightarrow & [0,1] \\ & A & \longmapsto & \displaystyle\sum_{i:\ x_i \in A} p_i \end{array}$$

Plus précisément, il existe une variable aléatoire X telle que $\mathbf{P}_X = \mathbf{P}$, c'est à dire telle que, pour tout $i \in I$, $\mathbf{P}([X = x_i]) = p_i$.

Méthode 17.30. **Déterminer la loi d'une v.a. discrète**

L'application $\mathbf{P}_X$ associe donc à un évènement la probabilité que X y prenne ses valeurs. Si bien entendu la définition de la loi est celle ci-dessus, ce n'est pas exactement cela qu'on va manipuler, mais plutôt la *distribution* des probabilités $x \in X(\Omega) \mapsto \mathbf{P}(X = x)$.

En effet, dans le cas d'un univers Ω discret (fini ou infini), la loi d'une variable aléatoire sera entièrement déterminée (comme le précise le **Théorème 17.2** ci-dessus) par la donnée de l'application

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

Ainsi, pour déterminer la loi d'une variable aléatoire, on cherche, pour chacun des éléments de son univers image, la probabilité que la variable aléatoire prenne cette valeur là.

La loi sera la principale information dont on disposera sur une variable aléatoire, l'univers Ω étant le plus souvent implicite et non précisé.

On peut représenter une loi par un histogramme ou un diagramme en bâtons: la hauteur du bâton d'abscisse x_n est alors la probabilité que X prenne la valeur x_n.
Si $x_n \notin X(\Omega)$, alors $\mathbf{P}(X = x_n) = 0$.

Remarque 17.3. **Variables aléatoires de même loi**

Il est erroné de penser que deux variables aléatoires qui possèdent la même loi sont égales.
Si on lance deux dés équilibrés, ils ne vont pas forcément tomber sur la même face et pourtant, les variables aléatoires qui prennent pour valeur le résultat de chacune des deux faces suivent bien la même loi (elles prennent les valeurs avec la même probabilité $1/6$).

Lorsque deux variables aléatoires X et Y *suivent* la même loi, c'est à dire lorsque $\mathbf{P}_X = \mathbf{P}_Y$ (ou plus simplement qu'elles ont la même distribution de probabilité) on écrira

$$X \sim Y.$$

Ce théorème permet en particulier de définir une variable aléatoire directement par sa distribution de probabilités, sans définir formellement l'univers ni les évènements associés.

Exercice 17.2.

On considère la suite (u_k) définie par

$$\forall k \geq 2, \qquad u_k = \frac{1}{k-1} - \frac{1}{k}.$$

Montrer qu'il existe une variable aléatoire X telle que, pour tout $k \geq 2$, $\mathbf{P}(X = k) = u_k$.

Définition 17.3. **Fonction de répartition**

Soit X une variable aléatoire. On appelle **fonction de répartition** de X la fonction F_X définie sur $\mathbb{R}$ par :

$$\forall t \in \mathbb{R}, \qquad F_X(t) = \mathbf{P}(X \leq t).$$

Proposition. 17.2. **Calcul de la fonction de répartition**

Soit X une variable aléatoire. Alors,

$$\forall t \in \mathbb{R}, \qquad F_X(t) = \sum_{\substack{x \in X(\Omega) \\ x \leq t}} \mathbf{P}(X = x).$$

Remarque 17.4.

On observe que, pour tout $k \in \mathbb{N}$, $[X = k] \cup [X \leq k-1] = [X \leq k]$ et que l'union est disjointe. Il suit que

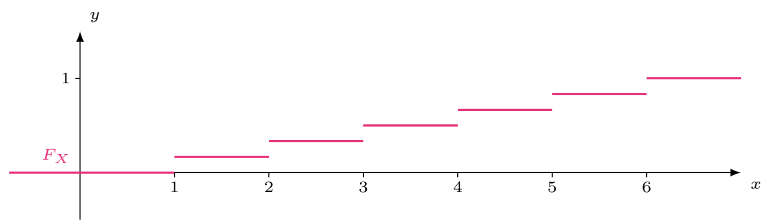
$$\forall k \in \mathbb{N}, \qquad F_X(k) - F_X(k-1) = \mathbf{P}(X = k).$$

☞ On se sert souvent de la fonction de répartition pour calculer la loi du max (ou du min, en passant aux évènements contraires) de plusieurs variables aléatoires *indépendantes*.

Exemple 17.2. **Fonction de répartition d'une loi uniforme**

Soit X la variable aléatoire qui prend pour valeur le résultat de lancer d'un dé équilibré à six faces. Alors

$$\forall t \in \mathbb{R}, \quad F_X(t) = \begin{cases} 0, & \text{si } t < 1 \\ \dfrac{\lfloor t \rfloor}{6}, & \text{si } t \leq 6 \\ 1, & \text{si } t > 6 \end{cases}$$



Proposition. 17.3.
Deux variables aléatoires X et Y suivent la même loi si et seulement si elles ont la même fonction de répartition.

17.2 Lois usuelles discrètes

17.2.1 Loi certaine

Une variable aléatoire sera dite certaine si elle prend toujours la même valeur avec une probabilité 1.

Définition 17.4. **Variable aléatoire certaine**
Soit X une variable aléatoire. On dit que X est **certaine** s'il existe a tel que $\mathbf{P}(X = a) = 1$.

17.2.2 Loi uniforme

La loi uniforme est celle des variables aléatoires **finies** qui prennent leurs différentes valeurs de manière équiprobable.

Définition 17.5. **Loi uniforme**
Soit X une variable aléatoire finie telle que $X(\Omega) = \{x_1, \ldots, x_n\}$. On dit que X suit une **loi uniforme** sur $\{x_1, \ldots, x_n\}$, ce qui sera noté $X \sim \mathcal{U}(\{x_1, \ldots, x_n\})$, si

$$\forall i \in [\![1, n]\!], \ \mathbf{P}(X = x_i) = \frac{1}{n}.$$

Remarque 17.5.
Dans le cas où $X(\Omega) = \mathbb{N}$, dire que X suit une loi uniforme est absurde (car il n'existe pas de probabilité uniforme sur un ensemble infini).

Exemple 17.3.
Une variable aléatoire donnant le résultat du lancer d'un dé équilibré suit la loi $\mathcal{U}[\![1, 6]\!]$.

En `Python` .
La commande `randint(a,b)` (du package `numpy.random`) renvoie une simulation de la loi $\mathcal{U}([\![a, b-1]\!])$.

17.2.3 Loi de Bernoulli

La loi de Bernoulli intervient à chaque fois qu'il est question d'expérience aléatoire à deux alternatives possibles (souvent appelées succès et échec).

Définition 17.6. **Loi de Bernoulli**
Soit $p \in [0, 1]$, et X une variable aléatoire telle que $X(\Omega) = \{0, 1\}$. On dit que X suit une **loi de Bernoulli** de paramètre p, ce qui sera noté $X \sim \mathcal{B}(1, p)$ ou simplement $X \sim \mathcal{B}(p)$, si

$$\mathbf{P}(X = 1) = p \qquad \text{et} \qquad \mathbf{P}(X = 0) = 1 - p.$$

Exemple 17.4.
On considère la variable aléatoire qui donne le résultat du lancer d'une pièce de monnaie truquée. Celle-ci suit alors la loi $\mathcal{B}(1, p)$, pour un certain réel p, égal à la probabilité d'obtenir *Pile* (ou *Face*).

Un autre exemple très utilisé de loi de Bernoulli est celui des variables indicatrices (voir **Exercice 17.??**).

17.2.4 Loi binomiale

Théorème & Définition 17.3. **Loi Binomiale**

Soit $n \in \mathbb{N}^*$, $p \in [0,1]$, et X une variable aléatoire telle que $X(\Omega) = [\![0,n]\!]$. On dit que X suit une **loi binomiale de paramètres** n, p, ce qui sera noté $X \sim \mathcal{B}(n,p)$, si :

$$\forall k \in [\![0,n]\!],\ \mathbf{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Preuve. La formule du binôme garantit que la formule ci-dessus définit bien une distribution de probabilités. □

Proposition. 17.4. **Nombre de succès**

Lorsqu'on répète n fois, de façon indépendante, la même expérience de Bernoulli de paramètre p (disons que p est la probabilité d'un succès), alors la variable aléatoire qui prend la valeur du *nombre de succès* suit la loi $\mathcal{B}(n,p)$.

Exemple 17.5.

On considère une urne contenant des boules, parmi lesquelles il y a une proportion $p \in [0,1]$ de boules blanches, et une proportion $1-p$ de boules noires. On effectue n tirages avec remise dans cette urne, et on appelle X la variable aléatoire qui prend pour valeur le nombre de boules blanches obtenues.
Alors $X \sim \mathcal{B}(n,p)$.

Remarque 17.6.

Une loi de Bernoulli de paramètre p est une loi binomiale de paramètres $1, p$, d'où la notation de la loi de Bernoulli.

En `Python`.

La commande `binomial(n,p)` (du package `numpy.random`) renvoie une simulation de la loi $\mathcal{B}(n,p)$.

17.2.5 Loi géométrique

Théorème & Définition 17.4. **Loi géométrique**

Soit $p \in [0,1]$. On dit que X suit **la loi géométrique** de paramètre p, ce qui se note $X \sim \mathcal{G}(p)$, si $X(\Omega) = \mathbb{N}^*$ et si

$$\forall k \in \mathbb{N}^*, \quad \mathbf{P}(X=k) = p(1-p)^{k-1}.$$

Exercice 17.3.

Démontrer que la formule ci-dessus définit bien une loi de probabilité.

Proposition. 17.5. **Temps d'attente du premier succès**

Lorsqu'on répète, de façon indépendante, la même expérience de Bernoulli de paramètre p (disons que p est la probabilité d'un succès) jusqu'à l'obtention d'un succès, alors la variable aléatoire qui prend la valeur du *nombre d'itérations* (ou du rang du premier succès) suit la loi $\mathcal{G}(p)$.

Exemple 17.6.

On lance une infinité de fois un dé équilibré et on note X la variable aléatoire qui prend la valeur du numéro du lancer où apparait, pour la première fois, la face numérotée "6". Alors $X \sim \mathcal{G}(1/6)$.

Exercice 17.4.

Soit $X \sim \mathcal{G}(p)$. Déterminer, pour tout $k \in \mathbb{N}$, $\mathbf{P}(X > k)$.

Exercice 17.5. **Coïncidence de lois géométriques indépendantes**

On considère deux lois géométriques de même paramètre p supposées *indépendantes* (on renvoie à la **Définition 17.10**). Déterminer $\mathbf{P}(X = Y)$.

Exercice 17.6. **Minimum de deux lois géométriques indépendantes**

On considère deux lois géométriques de même paramètre p supposées *indépendantes*.
Déterminer la loi de $Z = \min(X, Y)$.

La loi géométrique permet aussi de caractériser les lois discrètes *sans mémoire*. On renvoie à l'**Exercice 17.??**).

En `Python` **.**

La commande `geometric(p)` (du package `numpy.random`) renvoie une simulation de la loi $\mathcal{G}(p)$.

17.2.6 Loi de Poisson

Théorème & Définition 17.5. **Loi de Poisson**

Soit $\lambda > 0$. On dit que X suit **loi de Poisson** de paramètre λ, ce qui se note $X \sim \mathcal{P}(\lambda)$, si $X(\Omega) = \mathbb{N}$ et si

$$\forall k \in \mathbb{N}, \quad \mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Exercice 17.7.

Démontrer que la formule ci-dessus définit bien une loi de probabilité.

Remarque 17.7.

La loi de Poisson est aussi appelée *loi des évènements rares*, parce que $P(X = k)$ décroît très vite vers 0 lorsque k tend vers $+\infty$. Contrairement aux lois binomiale ou géométrique, on ne peut pas associer une loi de Poisson à la description d'une expérience aléatoire usuelle.
On s'en sert (et c'est dans ce cas une hypothèse faite *a priori*) pour modéliser des évènements associés à des prises de valeurs arbitrairement grandes mais dont la probabilité décroît très vite : nombre de voiture passant à un péage, nombre d'appels dans un standard téléphonique, apparition d'une mutation génétique chez un sujet...

Exercice 17.8. **Stabilité par somme des lois de Poisson**

Soient X et Y deux variables aléatoires *indépendantes* suivant toutes deux des lois de Poisson de paramètres respectifs λ et μ.
Montrer que

$$X + Y \hookrightarrow \mathcal{P}(\lambda + \mu).$$

17.3 Loi conditionnelle

Définition 17.7. **Loi conditionnelle**

Soit X une variable aléatoire telle que $X(\Omega) = \{x_i : \ i \in I\}$. Soient encore A un évènement non négligeable, et $\mathbf{P}_A$ la probabilité conditionnelle sachant A.
On appelle **loi conditionnelle de X sachant** A la loi de X relativement à l'espace probabilisé $(\Omega, \mathcal{T}, \mathbf{P}_A)$, c'est-à-dire celle associée à la distribution de probabilités

$$\begin{array}{ccc} \{x_i : \ i \in I\} & \to & [0, 1] \\ x_i & \mapsto & \mathbf{P}_A(X = x_i) \end{array}$$

Exercice 17.9.

On lance une pièce donnant pile avec probabilité $p \in]0; 1[$. On note X le rang d'apparition du premier pile et si $[X = n]$, on note Y le nombre de faces obtenus lors de n lancers supplémentaires.

1. Déterminer la loi conditionnelle de Y sachant $[X = n]$.
2. Préciser $Y(\Omega)$ puis montrer que pour tout $k \in \mathbb{N}$,

$$\mathbf{P}(Y = k) = \left(\frac{q}{p}\right)^{k-1} \sum_{n=k}^{+\infty} \binom{n}{k} (pq)^n.$$

Exercice 17.10. **Binomiale conditionnée par une Poisson**

Un auto-stoppeur attend au péage d'une autoroute pendant une certaine période. On suppose que le nombre de véhicules franchisant le péage pendant cette période est modélisé par une variable aléatoire suivant la loi de Poisson de paramètre $\lambda > 0$. A chaque fois qu'un véhicule franchit le péage, il lance une pièce truquée qui donne pile avec la probabilité p. On pose $q = 1 - p$.
On note N, X et Y respectivement le nombre de lancers, le nombre de piles obtenus et le nombre de face obtenus.

1. Soit $n \in \mathbb{N}$. Déterminer la loi conditionnelle de X sachant $[N = n]$.
2. Montrer que X suit une loi de Poisson de paramètre λp.
3. Sans calcul, que peut-on dire de la loi de Y ?

17.4 Loi d'une composée $\varphi(X)$

Soit X une variable aléatoire réelle, et une fonction φ à valeurs réelles définie sur $X(\Omega)$. Alors l'application $\varphi \circ X : \Omega \to \mathbb{R}$ sera notée $\varphi(X)$.

Remarque 17.8.
On exprime ici le fait que, considérant une variable aléatoire donnée par sa loi, on peut *oublier* l'espace probabilisé sous-jacent et considérer X comme une variable! Mais cette notation n'est valable que dans le cadre des variables aléatoires: la composée $f \circ g$ de deux applications ne se note pas $f(g)$ en général!

Nous allons montrer que, si X est une variable aléatoire, on peut exprimer la loi de $\varphi(X)$ en fonction de celle de X. On commence par donner un exemple représentatif.

Exercice 17.11.
Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{N}$ et φ la fonction définie par $\varphi(x) = \cos(\pi x)$. Déterminer la loi de $\varphi(X)$.

Plus généralement, on a le résultat ci-dessous, basé sur le principe suivant: $\varphi(X) = y$ si et seulement si X prend une des valeurs x_i que φ envoie sur y.

Proposition. 17.6.
Soit X une variable aléatoire, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction définie sur $X(\Omega)$.
Si $\varphi(X)$ est une variable aléatoire, alors son univers image est $\varphi(X(\Omega))$, et sa loi est donnée par :

$$\forall y \in Y(\Omega), \quad \mathbf{P}(Y = y) = \mathbf{P}(\varphi(X) = y) = \sum_{x \in X(\Omega)\colon \varphi(x) = y} \mathbf{P}(X = x).$$

Exemple 17.7.
Si $X(\Omega) = \{-1, 0, 1\}$, alors $\mathbf{P}(X^2 = 0) = \mathbf{P}(X = 0)$ et $\mathbf{P}(X^2 = 1) = \mathbf{P}(X = -1) + \mathbf{P}(X = 1)$.

Théorème 17.7.
Soient X, Y deux variables aléatoires, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction définie sur $X(\Omega)$. Si $X \sim Y$, alors $\varphi(X) \sim \varphi(Y)$.

17.5 Couples de variables aléatoires. Loi conjointe

17.5.1 Généralités

On admet qu'il existe une tribu, notée $\mathcal{B}_2$, sur $\mathbb{R}^2$, qui contient toutes les *boules ouvertes*. Elle contient en particulier tous les *rectangles* (produits cartésiens d'intervalles).

Définition 17.8.
On appelle **couple de variables aléatoires** toute fonction $(X, Y) : \Omega \to \mathbb{R}^2$, où X, Y sont deux variables aléatoires réelles.
On appelle alors **loi (conjointe)** du couple (X, Y) l'application

$$\begin{array}{ccc} \mathcal{B}_2 & \to & [0,1] \\ A & \mapsto & \mathbf{P}((X,Y) \in A) \end{array}$$

Remarque 17.9.
Un couple de variable aléatoire discrètes est en fait une variable aléatoire à valeurs dans $\mathbb{R}^2$.

La probabilité $\mathbf{P}(X = x_i \cap Y = y_j)$ sera notée $\mathbf{P}(X = x_i, Y = y_j)$.

Méthode 17.31. **Déterminer la oi d'un couple**

- ✗ Tout comme pour la loi d'une (seule) variable aléatoire, on détermine la loi d'un couple aléatoire en déterminant l'ensemble de ses valeurs possibles $(X, Y)(\Omega)$, puis on calcule $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in (X, Y)(\Omega)$.

  La loi conjointe du couple (X, Y) est déterminée par l'application :
  $$\begin{array}{rccl} \varphi_{X,Y}: & X(\Omega) \times Y(\Omega) & \to & [0;1] \\ & (i,j) & \mapsto & \mathbf{P}\left(X = i,\ Y = j\right). \end{array}$$
- ✗ On observera que $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$ mais qu'il n'y a pas nécessairement égalité.
- ✗ On ne cherchera (donc) pas systématiquement à déterminer l'ensemble des valeurs possibles de (X, Y): on pourra se contenter de donner $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

Exercice 17.12.
Une urne contient 3 boules numérotées de 1 à 3. On tire successivement et sans remise deux boules dans cette urne et on note X le numéro de la première boule et Y le numéro de la deuxième boule.

Déterminer la loi conjointe de (X, Y).

Exercice 17.13.
On lance une pièce qui, à chaque lancer, donne *Pile* avec probabilité p et *Face* avec probabilité $q = 1 - p$ (avec $p \in]0; 1[$). On note X le rang d'apparition du premier *Pile*, Y le rang d'apparition du premier *Face* et Z le rang d'apparition du second *Pile*.
1. Déterminer la loi conjointe de (X, Z).
2. Déterminer la loi conjointe de (X, Y).

Proposition. 17.8.
On a :
$$\sum_{i \in X(\Omega)} \sum_{j \in Y(\Omega)} \mathbf{P}\left([X = i] \cap [Y = j]\right) = \sum_{j \in Y(\Omega)} \sum_{i \in X(\Omega)} \mathbf{P}\left([X = i] \cap [Y = j]\right) = 1.$$

Dans la suite de cette section, J désigne une partie de $\mathbb{N}$.

Définition 17.9. **Lois marginales**
Soit (X, Y) un couple de variables aléatoires tel que $X(\Omega) = \{x_i : i \in I\}$ et $Y(\Omega) = \{y_j : j \in J\}$. On appelle **lois marginales** du couple (X, Y) les lois des variables aléatoires X et Y.

Méthode 17.32. **Déterminer les lois marginales**

Déterminer les lois marginales du coupe (X, Y), c'est déterminer les applications

$$\begin{array}{rcl} X(\Omega) & \to & \mathbb{R} \\ x_i & \mapsto & \mathbf{P}(X = x_i) \end{array} \quad \text{et} \quad \begin{array}{rcl} Y(\Omega) & \to & \mathbb{R} \\ y_j & \mapsto & \mathbf{P}(Y = y_j) \end{array}$$

Pour ce faire, on utilise la **formule des probabilités totales**. Plus précisément,

✗ On obtient la loi de X grâce à la formule des probabilités totales avec le s.c.e $\{[Y = y_j] : j \in J\}$.

$$\begin{aligned} \mathbf{P}([X = x_i]) &= \sum_{j \in J} \mathbf{P}(X = x_i,\ Y = y_j) \\ &= \sum_{j \in J} \mathbf{P}([Y = y_j]) P_{[Y=y_j]}([X = x_i]) \end{aligned}$$

✗ On obtient la loi de Y grâce à la formule des probabilités totales avec le s.c.e $\{[X = x_i] : i \in I\}$.

$$\begin{aligned} \mathbf{P}([Y = y_j]) &= \sum_{i \in I} \mathbf{P}(X = x_i,\ Y = y_j) \\ &= \sum_{i \in I} \mathbf{P}([X = x_i]) P_{[X=x_i]}([Y = y_j]) \end{aligned}$$

Si la loi du couple est résumée dans un tableau à double entrée, les lois marginales s'obtiennent en sommant les éléments de chaque ligne et de chaque colonne.

Remarque 17.10.

Si on peut déduire les lois marginales à partir de la loi conjointe, on ne peut pas déduire la loi conjointe d'un couple en fonction des lois marginales: il y a en effet dans un couple une troisième information, qui consiste en l'interaction entre les deux variables le composant.
En d'autres termes : X_1 peut avoir la même loi que X_2 et Y_1 avoir la même loi que Y_2 sans que les couples (X_1, Y_1) et (X_2, Y_2) aient la même loi.

17.5.2 Notion d'indépendance

Définition 17.10. **Variables aléatoires indépendantes**

Soient X, Y deux variables aléatoires. On dit que X et Y sont **indépendantes**, et on note $X \perp\!\!\!\perp Y$ si, quel que soit $(A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega))$, les évènements $[X \in A]$ et $[Y \in B]$ sont indépendants.

Remarque 17.11.

✗ Dans le cas de deux variables aléatoires indépendantes X et Y, la donnée de la loi du couple (X, Y) est donc équivalente à la donnée des lois de X et de Y.

✗ L'indépendance de deux variables aléatoires ne présage en rien de leur loi. En particulier, deux variables aléatoires peuvent être indépendantes et avoir même loi.

En pratique, c'est toujours la caractérisation suivante que l'on utilise.

Proposition. 17.9. **Caractérisation de l'indépendance de deux v.a.d**

Soient X, Y deux variables aléatoires. Alors X et Y sont indépendantes si et seulement si, quel que soit $(x, y) \in X(\Omega) \times Y(\Omega)$, les évènements $[X = x]$ et $[Y = y]$ sont indépendants. Soit:

$$X \perp\!\!\!\perp Y \iff \forall (x, y) \in X(\Omega) \times Y(\Omega),\ \mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x)\mathbf{P}(Y = y).$$

La notion d'indépendance s'étend à un nombre quelconque de variables aléatoires, en considérant des évènements mutuellement indépendants.

Définition 17.11.

Soit $(X_i)_{1 \le i \le m}$ une famille finie de variables aléatoires. Celles-ci sont dites **mutuellement indépendantes** si, quel que soit $k \in [\![1, m]\!]$, les entiers distincts $i_1, \ldots, i_k$ appartenant à $[\![1, m]\!]$, et $(A_1, \ldots, A_k) \in \prod_{j=1}^{k} \mathcal{P}\left(X_{i_j}(\Omega)\right)$ on a:

$$\mathbf{P}(X_{i_1} \in A_1, \ldots, X_{i_k} \in A_k) = \mathbf{P}(X_{i_1} \in A_1) \times \cdots \times \mathbf{P}(X_{i_k} \in A_k).$$

Remarque 17.12.

- ✗ $X_1, \ldots, X_m$ sont mutuellement indépendantes si et seulement si quel que soit $(A_1, \ldots, A_m) \in \prod_{j=1}^{m} \mathcal{P}(X_j(\Omega))$, les évènements $X_1 \in A_1, \ldots, X_m \in A_m$ sont mutuellement indépendants.
- ✗ Cette définition s'étend aux familles infinies dénombrables, qui sont dites mutuellement indépendantes si toute sous-famille finie est mutuellement indépendante.
- ✗ Une suite (X_i) de variables aléatoires (mutuellement) indépendantes de même loi est appelée une suite de *variables aléatoires indépendantes identiquement distribuées*, abrégé en *v.a.i.i.d.*

Exemple 17.8.

On modélise le plus souvent une suite de n expériences aléatoires indépendantes (comme par exemple n lancers de dés) par une suite (X_k) de n v.a.i.i.d (donnant par exemple le résultat de chacun des n lancers).

Proposition. 17.10.

Soit $(X_i)_{1 \leq i \leq m}$ une famille finie de variables aléatoires. Celles-ci sont mutuellement indépendantes si et seulement si, quel que soit $k \in [\![1, m]\!]$, les entiers distincts $i_1, \ldots, i_k$ appartenant à $[\![1, m]\!]$, et $(x_1, \ldots, x_k) \in X_{i_1}(\Omega) \times \cdots \times X_{i_k}(\Omega)$ on a:

$$\mathbf{P}(X_{i_1} = x_1, \ldots, X_{i_k} = x_k) = \mathbf{P}(X_{i_1} = x_1) \times \cdots \times \mathbf{P}(X_{i_k} = x_k).$$

Théorème 17.11. **Lemme des coalitions**

Soient $1 \leq m < n$ des entiers, ainsi que des variables aléatoires $X_1, \ldots, X_n$. On considère encore des fonctions $f_1, \ldots, f_n$ définies sur $\mathbb{R}$, ainsi que deux fonctions de plusieurs variables $F : \mathbb{R}^m \to \mathbb{R}$ et $G : \mathbb{R}^{n-m} \to \mathbb{R}$. Si $X_1, \ldots, X_n$ sont mutuellement indépendantes, alors les trois suites de variables aléatoires suivantes le sont également:

i. $X_1, \ldots, X_m$;
ii. $f_1(X_1), \ldots, f_n(X_n)$;
iii. $F(X_1, \ldots, X_m), G(X_{m+1}, \ldots, X_n)$.

Remarque 17.13.

Le deuxième point ci-dessus implique par exemple que, si $X \perp\!\!\!\perp Y$ alors $f(X) \perp\!\!\!\perp g(Y)$, quelles que soient les fonctions f et g. Intuitivement cela signifie que, quels que soient les procédés déterministes employés (ici des fonctions de la variable réelle), on ne peut créer une dépendance là où il n'y en avait pas.

Exemple 17.9.

- ✗ Si X, Y, Z sont mutuellement indépendantes, alors $X + Y$ et Z sont indépendantes.
- ✗ Si X_1, X_2, X_3, X_4, X_5 sont 5 variables aléatoires discrètes discrètes mutuellement indépendantes alors les variables $X_1 - X_2 + 2X_4^2$ et $X_3 - X_5^2$ sont indépendantes.
- ✗ Si X^2 et Y^2 ne sont pas indépendantes, alors X et Y ne le sont pas non plus.

17.5.3 Loi d'une somme de deux variables aléatoires

Proposition. 17.12. **Loi de la somme**

Soient X, Y deux variables aléatoires telle que $X(\Omega) = \{x_i | : i \in I\}$ et $Y(\Omega) = \{y_j : j \in J\}$. Alors la loi de $Z = X + Y$ est donnée par

$$Z(\Omega) = \{x_i + y_j \;\; i \in I, j \in J\}$$

et

$$\forall z \in Z(\Omega),\ \mathbf{P}(X + Y = z) = \sum_{(i,j):\, x_i + y_j = z} \mathbf{P}((X = x_i) \cap (Y = y_j))$$

Remarque 17.14.
Si X et Y prennent leurs valeurs dans $\mathbb{N}$, on obtient:

$$\mathbf{P}(X+Y=n)=\sum_{i+j=n}\mathbf{P}(X=i,Y=j)=\sum_{i=0}^{n}\mathbf{P}(X=i,Y=n-i).$$

C'est le plus souvent cette formule que l'on rencontre en exercice.

Exercice 17.14.
Soient X et Y deux variables indépendantes suivant toutes deux la loi géométrique $\mathcal{G}(p)$.
1. Déterminer, en utilisant un s.c.e associé à X, $\mathbf{P}(X=Y)$.
2. Montrer, en utilisant un s.c.e associé à X que pour tout $n\geq 2$: $\mathbf{P}\left(X+Y=n\right)=(n-1)p^2q^{n-2}$.

Théorème 17.13. **Stabilité par somme de certaines lois usuelles**
Soient $X_1,X_2,\ldots,X_m$ des variables aléatoires mutuellement indépendantes. On a les résultats suivants.

i. Si $X_1,\ldots,X_m\sim\mathcal{B}(1,p)$, alors $X_1+\cdots+X_m\sim\mathcal{B}(m,p)$.
ii. Si $X_1\sim\mathcal{B}(n_1,p)$ et $X_2\sim\mathcal{B}(n_2,p)$, alors $X_1+X_2\sim\mathcal{B}(n_1+n_2,p)$.
iii. Si $X_1\sim\mathcal{P}(\lambda_1),\ldots,X_m\sim\mathcal{P}(\lambda_m)$, alors $X_1+\cdots+X_m\sim\mathcal{P}(\lambda_1+...+\lambda_m)$.

17.6 Espérance et variance

17.6.1 Variables aléatoires admettant une espérance

Définition 17.12. **Espérance d'une variable aléatoire discrète**
Soit X une variable aléatoire telle que $X(\Omega)=\{x_i:\quad i\in I\}$. On dit que X **admet une espérance** si la série $\sum x_i\mathbf{P}(X=x_i)$ converge **absolument**.
Lorsque c'est le cas, sa somme est appelée espérance de Xs et notée $\mathbf{E}(X)$.

Remarque 17.15.
L'hypothèse de convergence absolue permet de démontrer que la somme $\displaystyle\sum_{i\in I}x_i\mathbf{P}(X=x_i)$ ne dépend pas de l'ordre dans lequel on somme les termes.

Exemple 17.10.
Toute variable aléatoire définie sur un espace probabilisé fini admet une espérance, et celle-ci est une somme finie.

Exercice 17.15.
Montrer qu'il existe une variable aléatoire X telle que $X(\Omega)=\mathbb{N}^*$ et $\mathbf{P}(X=n)=\dfrac{1}{n(n+1)}$ pour tout $n\in\mathbb{N}^*$.
Montrer ensuite que X n'admet pas d'espérance.

Théorème 17.14. **Espérance des lois usuelles**
Soit X une variable aléatoire.
i. Si X suit une loi certaine, elle prend une unique valeur a, et alors $\mathbf{E}(X)=a$.
ii. Si $X\sim\mathcal{U}\left(\llbracket 1,n\rrbracket\right)$, alors $\mathbf{E}(X)=\dfrac{n+1}{2}$.
iii. Si $X\sim\mathcal{B}(1,p)$, alors $\mathbf{E}(X)=p$.
iv. Si $X\sim\mathcal{B}(n,p)$, alors $\mathbf{E}(X)=np$.
v. Si $X\sim\mathcal{G}(p)$, alors X admet une espérance et $\mathbf{E}(X)=\dfrac{1}{p}$.
vi. Si $X\sim\mathcal{P}(\lambda)$, alors X admet une espérance et $\mathbf{E}(X)=\lambda$.

Proposition. 17.15. **Propriétés de l'espérance**

Soient $\lambda, \mu \in \mathbb{R}$ et X, Y deux variables aléatoires **admettant** une espérance. On a les résultats suivants.

i. **Linéarité.** $\lambda X + \mu Y$ admet une espérance, et on a $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$.
ii. **Positivité.** Si X est à valeurs positives, on a $\mathbf{E}(X) \geq 0$. Et sous cette hypothèses, on a $\mathbf{E}(X) = 0$ si et seulement si $\mathbf{P}(X = 0) = 1$.
iii. **Croissance.** Si $X(\omega) \geq Y(\omega)$ pour tout $\omega \in \Omega$, alors on a $\mathbf{E}(X) \geq \mathbf{E}(Y)$.

Proposition. 17.16. **Domination**

Soient X, Y deux variables aléatoires. Si Y admet une espérance, et si $|X| \leq |Y|$, alors X admet une espérance et $\mathbf{E}(|X|) \leq \mathbf{E}(|Y|)$.

Proposition. 17.17.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ et admettant une espérance. On a alors

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

Définition 17.13. **Variable aléatoire centrée**

Une variable aléatoire X est dite centrée si l'on a $\mathbf{E}(X) = 0$.

Si la loi d'une composée ne s'exprime pas aisément, il n'en est pas de même de son espérance, comme le montre le résultat suivant, extrêmement utile.

Théorème 17.18. **Théorème de transfert**

Soit X une variable aléatoire telle que $X(\Omega) = \{x_i : \ i \in I\}$, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction définie sur $X(\Omega)$. Si $\varphi(X)$ est une variable aléatoire, alors on a les résultats suivants:

i. $\varphi(X)$ admet une espérance si et seulement si $\sum \varphi(x_i)\mathbf{P}(X = x_i)$ converge absolument;
ii. lorsque l'une des deux conditions ci-dessus est vérifiée on a $\mathbf{E}(\varphi(X)) = \sum_{i \in I} \varphi(x_i)\mathbf{P}(X = x_i)$.

Exercice 17.16.

Soient X une loi de Poisson de paramètre $\lambda > 0$ et $Y = 1/(X+1)$. Montrer, à l'aide du théorème de transfert que Y admet une espérance et la déterminer.

Exercice 17.17.

Soit $X \sim \mathcal{B}(n, p)$. Le but est de calculer la probabilité $a = P(A)$ où A est l'évènement réalisé si et seulement si X prend une valeur paire.

1. On introduit $Y = (-1)^X$. Quelle est la loi de Y? Expliciter son espérance en fonction de a.
2. Calculer, à l'aide du théorème de transfert, l'espérance de Y en fonction de n et de p. Conclure.

Remarque 17.16.

✗ Ce résultat montre que pour connaitre l'espérance de $\varphi(X)$, il n'est pas nécessaire de connaitre sa loi: celle de X suffit.
✗ On utilise ce résultat pour calculer $\mathbf{E}(X^2)$ sans déterminer au préalable la loi de X^2.
✗ Le théorème de transfert se généralise aux couples, et aux n-uplets de variables aléatoires.

Théorème 17.19.

Soient X, Y deux variables aléatoires réelles indépendantes admettant une espérance. Alors XY admet également une espérance, et on a $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

17.6.2 Espérance d'un produit

Remarque 17.17.
Soit (X,Y) un couple de deux variables aléatoires discrètes. Alors, le théorème de transfert permet d'écrire, sous réserve de convergence absolue,
$$\mathbf{E}(XY) = \sum_{i\in X(\Omega)} \sum_{j\in Y(\Omega)} ij\mathbf{P}([X=i]\cap[Y=j]).$$

Proposition. 17.20.
Si X et Y sont **indépendantes** et admettent une espérance, alors XY admet une espérance, et on a
$$\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y).$$

On peut utiliser le théorème précédent pour montrer que deux variables aléatoires discrètes **ne sont pas** indépendantes en vérifiant que $\mathbf{E}(XY) \neq \mathbf{E}(X)\mathbf{E}(Y)$.

Exercice 17.18.
Soient X_1, X_2, et X_3, indépendantes, suivant des lois de Bernoulli de paramètre p. Montrer que les variables aléatoires $Y_1 = X_1X_2$ et $Y_2 = X_2X_3$ ne sont pas indépendantes.

17.7 Variables aléatoires admettant une variance

17.7.1 Notion de moment d'ordre quelconque

Définition 17.14. **Moment d'ordre r**
Soit X une variable aléatoire et $r \in \mathbb{N}^*$. On dit que X **admet un moment d'ordre** r si la variable aléatoire X^r admet une espérance. Lorsque c'est le cas on définit son moment d'ordre r, noté $m_r(X)$, par $m_r(X) = \mathbf{E}(X^r)$.

Remarque 17.18.
Si X admet un moment d'ordre $r \in \mathbb{N}^*$, le théorème de transfert donne $m_r(X) = \sum_{x\in X(\Omega)} x^r\mathbf{P}(X=x)$.

Proposition. 17.21.
Soit X une variable aléatoire admettant un moment d'ordre 2. Alors X et $(X-\mathbf{E}(X))^2$ admettent une espérance.

Preuve. Le premier point se déduit de $2|x| \leq x^2+1$. □

Remarque 17.19.
Plus généralement, si X possède un moment d'ordre $r \in \mathbb{N}^*$, alors X possède aussi un moment d'ordre $s \in \mathbb{N}^*$ pour tout $s \leq r$.

17.7.2 Variance

Définition 17.15.
Soit X une variable aléatoire. Si X admet un moment d'ordre 2, alors $\mathbf{E}((X-\mathbf{E}(X))^2)$ est appelée **variance** de X, et est notée $\mathbf{V}(X)$.
On définit encore **l'écart-type** de X, noté $\sigma(X)$, en posant $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

Remarque 17.20.

- ✗ L'espérance d'une variable aléatoire donne sa valeur moyenne. Pour obtenir une description plus précise de la loi d'une variable aléatoire, on cherche à donner une quantité synthétisant la dispersion de celle-ci autour de sa moyenne: on utilise les notions de variance et d'écart-type.

- ✗ Plaçons-nous dans le cadre d'une expérimentation où la grandeur X mesure plusieurs fois de suite un même phénomène naturel. Même si X mesure un phénomène déterministe, on doit le considérer comme étant une variable aléatoire, ne serait-ce que parce que chaque mesure est entachée d'erreurs elles-même aléatoires. Il alors est naturel d'approximer la valeur théorique de X par la moyenne des valeurs mesurées, c'est-à-dire par son espérance $\mathbf{E}(X)$. L'écart-type est alors une approximation de *l'erreur moyenne de mesure par rapport à la valeur théorique*.

- ✗ Si X est exprimé dans une certaine unité u, alors $\mathbf{V}(X)$ est exprimé en u^2. C'est pour cette raison qu'en sciences physiques on considère plutôt l'écart type σ, qui est lui exprimé en u. En mathématiques on préfère la variance, qui possède des propriétés algébriques plus simples et plus naturelles.

☞ En pratique, on calcule souvent la variance à l'aide du théorème de transfert et de la formule suivante.

Proposition. 17.22. **Formule de Kœnig-Huygens**
Soit X une variable aléatoire admettant un moment d'ordre 2. On a alors $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

Proposition. 17.23. **Variance d'une transformation affine**
Soit X une variable aléatoire admettant un moment d'ordre 2. Alors quels que soient les réels a et b, $aX + b$ admet une variance, et $\mathbf{V}(aX + b) = a^2\mathbf{V}(X)$.

Définition 17.16. **Variable aléatoire réduite**
Une variable aléatoire X admettant un moment d'ordre 2 est dite **réduite** si $\mathbf{V}(X) = 1$.

Proposition. 17.24.
Soit X une variable aléatoire admettant un moment d'ordre 2 dont la variance est non nulle.
Alors la variable aléatoire $X^* = \dfrac{1}{\sigma(X)}\,(X - \mathbf{E}(X))$ est centrée et réduite et s'appelle la variable aléatoire centrée-réduite associée à X.

Proposition. 17.25.
Soit X une variable aléatoire admettant un moment d'ordre 2. On a les résultats suivants.

i. $\mathbf{V}(X) \geq 0$;
ii. $\mathbf{V}(X) = 0 \iff P(X = \mathbf{E}(X)) = 1$.

En d'autres termes, la variance est toujours positive, et s'annule si et seulement si la variable aléatoire est certaine.

Preuve. Il suffit d'appliquer la positivité de l'espérance à la variable aléatoire $(X - \mathbf{E}(X))^2$. □

Théorème 17.26. **Variance des lois usuelles**
Soit X une variable aléatoire. On a alors les résultats suivants.

i. Si X suit une loi certaine, alors $\mathbf{V}(X) = 0$.
ii. Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$, alors $\mathbf{V}(X) = \dfrac{n^2 - 1}{12}$.
iii. Si $X \sim \mathcal{B}(1, p)$, alors $\mathbf{V}(X) = p(1 - p)$.
iv. Si $X \sim \mathcal{B}(n, p)$, alors $\mathbf{V}(X) = np(1 - p)$.
v. Si $X \sim \mathcal{G}(p)$, alors X admet une variance et $\mathbf{V}(X) = \dfrac{1 - p}{p^2}$.
vi. Si $X \sim \mathcal{P}(\lambda)$, alors X admet une variance et $\mathbf{V}(X) = \lambda$.

Lois usuelles : synthèse

Loi de X	$X(\Omega)$	$\mathbf{P}(X=k)$	$\mathbf{E}(X)$	$\mathbf{V}(X)$
$\mathcal{U}(\llbracket 1,n\rrbracket)$	$\llbracket 1,n\rrbracket$	$\dfrac{1}{n}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$
$\mathcal{B}(1,p)$	$\{0,1\}$	$\begin{cases} p & \text{si } k=1 \\ 1-p & \text{si } k=0\end{cases}$	p	$p(1-p)$
$\mathcal{B}(n,p)$	$\llbracket 0,n\rrbracket$	$\binom{n}{k} p^k(1-p)^{n-k}$	np	$np(1-p)$
$\mathcal{G}(p)$	$\mathbb{N}^*$	$p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
$\mathcal{P}(\lambda)$	$\mathbb{N}$	$e^{-\lambda}\dfrac{\lambda^k}{k!}$	λ	λ

17.8 Covariance de deux variables aléatoires

La covariance se présente au départ comme un outil introduit pour calculer la variance d'une somme de variables aléatoires. On verra ensuite qu'elle possède une utilité propre.

Proposition. 17.27.
Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Alors $[X-\mathbf{E}(X)]\times[Y-\mathbf{E}(Y)]$ admet une espérance, et l'on a

$$\mathbf{E}\Big([X-\mathbf{E}(X)]\times[Y-\mathbf{E}(Y)]\Big)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y).$$

Définition 17.17. **Covariance**
Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. On appelle **covariance** de X et de Y la quantité définie par

$$\operatorname{cov}(X,Y)=\mathbf{E}\Big([X-\mathbf{E}(X)]\times[Y-\mathbf{E}(Y)]\Big)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y).$$

Intuitivement, la covariance de X et de Y est une mesure relative (ou couplée) des dispersions de X et de Y autour de leurs moyennes. Elle se calcule grâce au théorème de transfert.

Remarque 17.21.
On déduit directement de la définition la formule $\mathbf{V}(X)=\operatorname{cov}(X,X)$.

Théorème 17.28.
Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Si X, Y sont indépendantes, alors on a $\operatorname{cov}(X,Y)=0$.

Remarque 17.22.
Attention! La réciproque du résultat ci-dessus est fausse.

Définition 17.18.
Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Si $\operatorname{cov}(X,Y)=0$, on dit que X et Y dont **décorelées**.

Proposition. 17.29. **Propriétés élémentaires de la covariance**

Soient X, X', Y, Y' des variables aléatoires admettant un moment d'ordre 2, et $\lambda, \mu \in \mathbb{R}$. On a les résultats suivants.

i. **Variance d'une somme**: $X+Y$ admet une variance et l'on a $\mathbf{V}(X+Y) = \mathbf{V}(X) + \mathbf{V}(Y) + 2\mathrm{cov}(X,Y)$.
ii. **Symétrie de la covariance** : $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$.
iii. **Bilinéarité de la covariance**: $\mathrm{cov}(\lambda X + \mu X', Y) = \lambda \cdot \mathrm{cov}(X,Y) + \mu \cdot \mathrm{cov}(X',Y)$ et $\mathrm{cov}(X, \lambda Y + \mu Y') = \lambda \cdot \mathrm{cov}(X,Y) + \mu \cdot \mathrm{cov}(X,Y')$.

Remarque 17.23.

- ✗ On se place sur l'espace vectoriel des variables aléatoires à valeurs réelles définies sur un espace probabilisé donné. La covariance **ressemble** à un produit scalaire (elle est bilinéaire, symétrique et positive), et l'écart-type à la norme associée. Mais la covariance n'est pas définie positive, car $V(X) = 0$ n'implique pas $X = 0$.
  Par analogie, on peut toutefois considérer que la covariance permet de mesurer le *degré d'indépendance linéaire* existant entre deux variables aléatoires.

- ✗ Dans le cas d'une étude statistique, la variance d'une variable aléatoire X est l'erreur moyenne commise lors des mesures de la grandeur associée. La formule $\mathrm{cov}(X,Y) = \frac{1}{2}(V(X+Y) - V(X) - V(Y))$ indique que $\mathrm{cov}(X,Y)$ donne la variation des erreurs commises lorsque l'on passe de (X,Y) à $X+Y$. Celles-ci pouvant s'ajouter ou se compenser, la covariance est donc une donnée précieuse.

Exercice 17.19.

Soient X, Y, Z trois variables aléatoires mutuellement indépendantes. Montrer que $U = X+Y$ et $V = X+Z$ sont indépendantes si et seulement si X est certaine.

Corollaire 17.30. **Variance d'une somme de v.a.d indépendantes**

Soient $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes, admettant un moment d'ordre 2. On a alors:

$$\mathbf{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbf{V}(X_k).$$

Exercice 17.20.

Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes admettant toutes une même variance σ^2. Déterminer la variance de la variable aléatoire

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Remarque 17.24.

- ✗ Dans l'analogie décrite ci-dessus, ce théorème correspond au théorème de Pythagore.
- ✗ Ce résultat permet de retrouver facilement la variance d'une loi binomiale.

On a enfin le résultat suivant, qui provient de l'inégalité de Cauchy-Schwarz.

Proposition. 17.31.

Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Alors XY admet une espérance, et l'on a

$$\mathbf{E}(XY)^2 \leq \mathbf{E}\left(X^2\right)\mathbf{E}\left(Y^2\right) \iff |\mathbf{E}(XY)| \leq \sigma(X)\sigma(Y).$$

Preuve. Résultat admis (il suffit de constater que $\phi : (X,Y) \mapsto \mathbf{E}(XY)$ est bilinéaire, symétrique et positive, et que cela suffit pour faire la preuve de l'inégalité de Cauchy-Schwarz). □

17.9 Série génératrice d'une variable aléatoire entière

Dans cette section, toutes les variables aléatoires considérées seront définies sur $(\Omega, \mathcal{T}, P)$, et supposées *entières* (c'est-à-dire à valeurs dans $\mathbb{N}$).

17.9.1 Définitions et résultats élémentaires

Définition 17.19. **Série génératrice**

Soit X une variable aléatoire entière. On appelle **série génératrice** de X (ou fonction génératrice de X), et on note G_X la fonction qui à t associe la somme de la série entière $\sum_{n\geq 0} \mathbf{P}(X=n)t^n$.

Proposition. 17.32.

Soit X une variable aléatoire entière. Le rayon de convergence de sa série génératrice G_X est supérieur à 1, et on a

$$\forall t \in [-1,1],\ G_X(t) = \sum_{n=0}^{+\infty} \mathbf{P}(X=n)t^n.$$

Le théorème de transfert permet alors immédiatement d'énoncer le résultat suivant.

Corollaire 17.33.

Soit X une variable aléatoire entière. On a $\forall t \in [-1,1],\ G_X(t) = \mathbf{E}\left(t^X\right)$.

Théorème 17.34.

Deux variables aléatoires entières ont même loi si et seulement si elles ont même série génératrice.

Théorème 17.35.

Soit X une variable aléatoire entière, de série génératrice G_X. Alors X admet une espérance si et seulement si G_X est dérivable en 1, et lorsque c'est le cas $\mathbf{E}(X) = G_X'(1)$.

Citons le résultat suivant, qui n'est pas au programme, mais il faut être capable de retrouver la formule donnant la variance en fonction de la série génératrice.

Théorème 17.36.

Soit X une variable aléatoire entière, de série génératrice G_X. Alors X admet un moment d'ordre 2 si et seulement si G_X est deux fois dérivable en 1, et lorsque c'est le cas $\mathbf{V}(X) = G_X''(1) + G_X'(1) - \left(G_X'(1)\right)^2$.

Théorème 17.37. **Somme de v.a. indépendantes et séries génératrices**

Soit X, Y deux variables aléatoires entières. On note R_X (resp. R_Y) le rayon de convergence de G_X (resp. G_Y). Si X, Y sont indépendantes, alors G_{X+Y} a un rayon de convergence supérieur à $R = \min(R_X, R_Y)$, et on a $\forall t \in]-R, R[,\ G_{X+Y}(t) = G_X(t) \cdot G_Y(t)$.

17.9.2 Cas des lois usuelles

Théorème 17.38. **Séries génératrices des lois usuelles**

Soit $n \in \mathbb{N}^*$ et $p \in]0,1[$.

i. Si $X \sim \mathcal{U}(⟦1,n⟧)$, alors $G_X(t) = \dfrac{1}{n}\sum_{k=1}^{n} t^k$ et le rayon de convergence est infini.
ii. Si $X \sim \mathcal{B}(p)$, alors $G_X(t) = 1-p+pt$ et le rayon de convergence est infini.
iii. Si $X \sim \mathcal{B}(n,p)$, alors $G_X(t) = (1-p+pt)^n$ et le rayon de convergence est infini.
iv. Si $X \sim \mathcal{G}(p)$, alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$ et le rayon de convergence est $\dfrac{1}{1-p}$.
v. Si $X \sim \mathcal{P}(\lambda)$, alors $G_X(t) = e^{\lambda(t-1)}$ et le rayon de convergence est infini.

Exercice 17.21.

Retrouver, à l'aide du **Théorème 17.37** le résultat sur la loi d'une somme de variables aléatoires indépendantes suivant une loi de Poisson.

Exercice 17.22.
On lance un dé deux fois de suite. X_1 et X_2 sont les numéros des faces obtenues lors, respectivement, du premier et du second lancer. $S = X_1 + X_2$ est la somme de ces deux numéros.
1. Trouver la fonction génératrice de S.
2. Est-il possible de truquer le dé pour que S suive une loi uniforme sur $\{2, \ldots, 12\}$?

17.10 Quelques résultats asymptotiques

17.10.1 Approximation de la loi binomiale par une loi de Poisson

Le résultat suivant n'est pas au programme, mais permet de justifier que la loi de Poisson est appelée loi des évènements rares.

Proposition. 17.39. **Convergence en loi de la binomiale vers une Poisson**
Soit $(p_n)_{n\in\mathbb{N}}$ une suite de réels compris dans $]0,1[$ et $\lambda > 0$ tels que $\lim\limits_{n\to+\infty} np_n = \lambda$.
Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires telles que $X_n \sim \mathcal{B}(n, p_n)$ pour tout $n \in \mathbb{N}$, alors

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} \mathbf{P}(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

On dit que la suite (X_n) **converge en loi** vers une loi de Poisson. On note

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} Z, \quad Z \sim \mathcal{P}(\lambda).$$

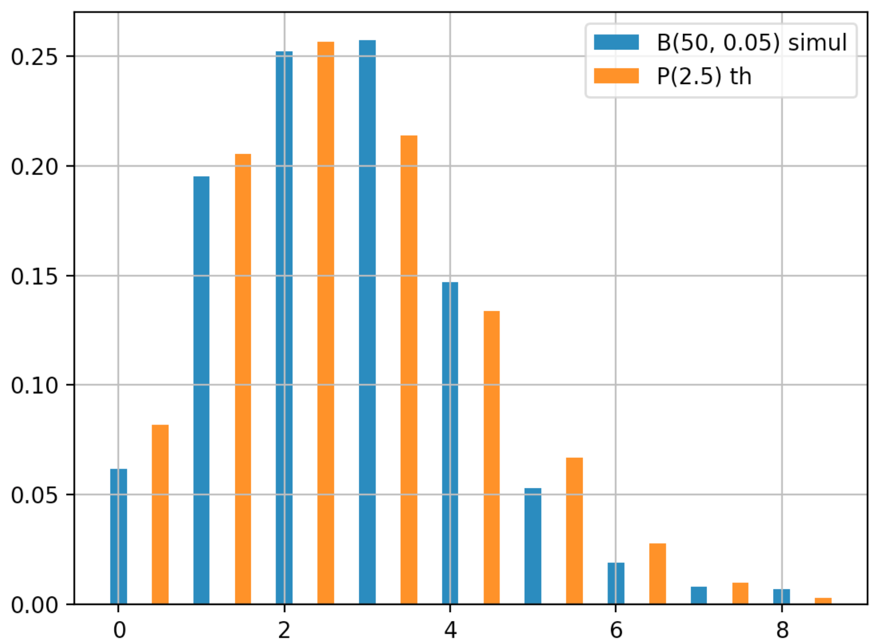
Remarque 17.25.

- ✗ L'intérêt de ce théorème réside dans le fait que le calcul numérique d'une loi de Poisson est plus simple que celui d'une loi binomiale.
- ✗ On peut par exemple utiliser ce résultat si $p_n = \dfrac{\lambda}{n}$.
- ✗ Dans une situation concrète, si n est grand et p est petit, on peut utiliser ce résultat si $np = \lambda$ reste petit par rapport à n. En pratique le résultat est exploitable si $n > 20, p \leq \dfrac{1}{10}$ et $np \leq 5$.
- ✗ Soit une urne contenant une proportion $p \in]0,1[$ de boules blanches. Si p est proche de 0 et n est assez grand, le théorème ci-dessus laisse penser que le nombre de boules blanches obtenues après n tirages suit approximativement une loi de Poisson de paramètre $\lambda = np$. C'est pour cette raison que la loi de Poisson est appelée la *loi des évènements rares*.
- ✗ On pourra retenir le résultat sous la forme suivante : $p_n \underset{n\to+\infty}{\sim} \dfrac{\lambda}{n} \Longrightarrow \mathcal{B}(n, p_n) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{P}(\lambda)$.

On illustre ce résultat avec `Python`.

```
import numpy as np
import numpy.random as rd
import matplotlib.pyplot as plt

X=[rd.binomial(50, 0.05) for k in range(1000)] #echantillon taille 1000
m=max(X)
freq=np.zeros(m+1)
for k in range(len(X)):
    freq[X[k]]+=1
freq=freq/1000

def distr_poisson_th(lam,n):
    y=np.zeros(n+1)
```

```
    y[0]=np.exp(-lam)
    for k in range(1, n+1):
        y[k]=y[k-1]*lam/k
    return y

N=[k for k in range(m+1)]
Z=distr_poisson_th(2.5, m)

plt.grid()
plt.bar(N, freq, width=0.2, label='B(50, 0.05) simul')
plt.bar([k+0.5 for k in range(m+1)], Z, width=0.2, label='P(2.5) th')
plt.legend()
plt.show()
```



En bleu, les hauteur des bâtons correspondent aux *fréquences* d'apparition de chaque valeur lors de la simulation de 1000 variables aléatoires de loi $\mathcal{B}(50, 0.05)$.
En orange, les hauteurs *théoriques* de la distribution d'une loi de Poisson de paramètre 2.5.

17.10.2 Loi (faible) des grands nombres

Proposition. 17.40. **Inégalité de Markov**

Soit X une variable aléatoire à valeurs dans $\mathbb{R}_+$ admettant une espérance. Alors :

$$\forall a > 0,\ \mathbf{P}(X \geq a) \leq \frac{\mathbf{E}(X)}{a}.$$

Proposition. 17.41. **Inégalité de Bienaymé-Tchebychev**

Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors :

$$\forall \varepsilon > 0,\ \mathbf{P}\left(|X - \mathbf{E}(X)| \geq \varepsilon\right) \leq \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

Preuve. Il suffit d'appliquer l'inégalité de Markov à la variable aléatoire $(X - \mathbf{E}(X))^2$. □

Remarque 17.26.

- ✗ Cette inégalité majore la probabilité que X s'éloigne d'au moins ε de sa moyenne.
- ✗ Cette probabilité est d'autant plus petite que ε est grand : plus on s'éloigne de la moyenne moins on a de chance de trouver X.
- ✗ Cette probabilité est d'autant plus petite que $\mathbf{V}(X)$ est petit : une variance faible correspond à une variable aléatoire resserrée autour de sa moyenne.
- ✗ Enfin, dans le cadre d'une expérimentation où l'on mesure plusieurs fois de suite un même phénomène naturel X, cette inégalité majore la probabilité que l'on ait commis une erreur de mesure supérieure à un seuil ε donné (on considère en général un seuil critique).

Définition 17.20. **Moyenne empirique**

Soit $(X_n)_{n\in\mathbb{N}}$ une famille de variables aléatoires indépendantes suivant la même loi. Pour tout $n \in \mathbb{N}$, on appelle **moyenne empirique** de $X_1, \ldots, X_n$ la variable aléatoire $\overline{X}_n$ définie par

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Exemple 17.11.

On lance n fois une pièce de monnaie qui renvoie *Pile* avec probabilité p. Le résultat du k-ième lancer correspond à une variable aléatoire X_k, qui prend la valeur 1 si l'on a obtenu *Pile* et 0 sinon. Les X_k sont indépendantes et suivent la même loi. Dans ce cas, $\overline{X}_n$ est le nombre moyen de *Pile* obtenus. La *loi faible des grands nombres* établira que ce nombre moyen tend vers p lorsque n tend vers $+\infty$, ce qui permet d'ailleurs d'estimer p par observations si jamais le paramètre est inconnu.

Théorème 17.42. **Loi (faible) des grands nombres**

Soit $(X_n)_{n\in\mathbb{N}}$ une famille de variables aléatoires indépendantes suivant la même loi, admettant une espérance μ et une variance σ^2.
Alors, pour tout $n \in \mathbb{N}$ et tout réel $\varepsilon > 0$ on a

$$\mathbf{P}\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \quad \text{et} \quad \lim_{n\to+\infty} \mathbf{P}\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) = 0.$$

Remarque 17.27.

- ✗ Ce résultat établit que, pour tout $\varepsilon > 0$, il devient improbable que $\overline{X}_n \notin]\mu - \varepsilon, \mu + \varepsilon[$ lorsque $n \to +\infty$ (mais pas que cela est impossible). On a de plus un majorant de cette probabilité.
- ✗ Ce résultat est d'une grande importance théorique et pratique. Il indique par exemple qu'en mesurant n fois un phénomène physique, on obtient une moyenne empirique $\overline{X}_n$ qui tend vers la moyenne théorique μ lorsque $n \to +\infty$.
- ✗ Ce résultat permet également d'estimer la valeur moyenne d'une variable dans une population à l'aide d'un sondage.

Exemple 17.12.

Soit $\varepsilon = \dfrac{1}{10}$, et supposons $V(X_1) \leq 1$. Pour $n > 10^4$, on a

$$\mathbf{P}\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{1}{100}.$$

On dit que

$$\left]\overline{X}_n - \frac{1}{10}, \overline{X}_n + \frac{1}{10}\right[$$

est un *intervalle de confiance de μ au seuil de confiance de* 99% *(ou au seuil de risque de* 1%*).*
$\overline{X}_n$ est appelé un **estimateur** de μ, dont l'écart à μ sera donc inférieur à $\dfrac{1}{10}$ dans 99% des cas.

18 Courbes et Surfaces de l'espace

Dans tout ce chapitre, on se place dans l'espace affine euclidien orienté usuel, muni d'un repère orthonormé direct $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$. On peut ainsi identifier cet espace à $\mathbb{R}^3$ muni de sa base canonique.

18.1 Courbes paramétrées de l'espace

18.1.1 Généralités

Définition 18.1. **Courbe paramétrée de l'espace**

Soit $k \in \mathbb{N}$. On appelle **courbe ou arc paramétré(e)** de $\mathbb{R}^3$ de classe $\mathcal{C}^k$ toute application (à valeurs vectorielles)

$$M : t \mapsto \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

définie sur un intervalle non trivial I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}^3$, où x, y, z sont de classe $\mathcal{C}^k$.
L'ensemble $\Gamma = \{M(t);\ t \in I\}$ est appelé **support** (ou trajectoire) de M. On dit aussi que la courbe Γ est paramétrée par M.

Remarque 18.1.

Tout comme ce qu'on avait vu dans le plan, une courbe paramétrée ne se résume pas à son support. Le support est un objet géométrique, tandis que la courbe paramétrée M donne une information supplémentaire : elle indique la manière dont le support est parcouru.
On peut décrire le même objet géométrique de plusieurs manières. Par exemple, le cercle trigonométrique du plan (xOy) est paramétré par les fonctions vectorielles $M_1 : t \mapsto (\cos(t), \sin(t), 0)$ et $M_2 : t \mapsto (\cos(2t), \sin(2t), 0)$.

Exemple 18.1. **Une courbe paramétrée de l'espace : l'hélice circulaire**

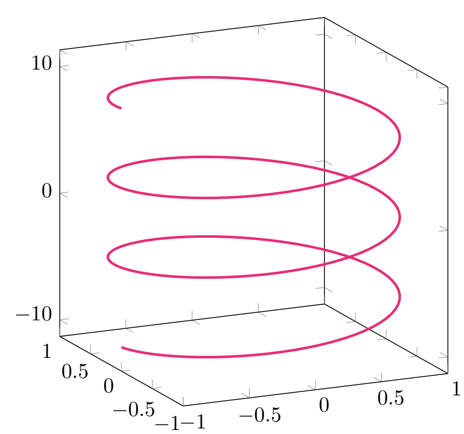
On représente ci-contre le support de la courbe paramétrée par

$$M : t \mapsto \begin{cases} \sin(t) \\ \cos(t) \\ t \end{cases}$$

Il s'agit d'une courbe régulière. On observe de plus qu'en tout point t_0,

$$\left\langle \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) \middle| \ \vec{k} \right\rangle = 1.$$

On en déduit que la tangente forme un angle constante avec l'axe (c'est d'ailleurs la définition de l'hélice).



18.1.2 Étude locale en un point régulier

Dans cette section, on considère I un intervalle non trivial de $\mathbb{R}$, et Γ une courbe paramétrée par $M : I \to \mathbb{R}^3$ de classe $\mathcal{C}^1$.

La tangente en un point d'une courbe paramétrée de l'espace est définie comme dans le cas du plan: il s'agit de la limite des cordes à la courbe passant par le point considéré (sous réserve d'existence).

Définition 18.2. **Point régulier, courbe régulière**

Soit Γ une courbe de $\mathbb{R}^3$ paramétrée par M.

On dit que le point $M(t_0)$ de paramètre $t_0 \in I$ est un **point régulier** de Γ si $\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) \neq \vec{0}$.

Dans le cas contraire, le point sera dit **stationnaire** ou *singulier*.
Une courbe paramétrée est dite **régulière** si tous ses points sont réguliers.

Théorème 18.1.

En tout point régulier $M(t_0)$ où $t_0 \in I$, la courbe Γ possède une tangente, et celle-ci est dirigée par $\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0)$.

Preuve. La preuve est analogue à celle faite dans le cas du plan, à laquelle on renvoie. □

Remarque 18.2.

- ✗ La tangente d'une courbe paramétrée en un point est un objet géométrique : elle ne dépend pas de la paramétrisation de la courbe. En effet, si $\varphi : J \to I$ est une fonction bijective de classe $\mathcal{C}^1$ définie sur un intervalle J de $\mathbb{R}$, alors les courbes paramétrées par $M_1 : I \to \mathbb{R}^3$ et par $M_2 = M_1 \circ \varphi : J \to \mathbb{R}^3$ ont le même support. De plus, si on note $s_0 = \varphi^{-1}(t_0)$, alors on a
$$\frac{\mathrm{d}\overrightarrow{M_2}}{\mathrm{d}t}(s_0) = \varphi'(s_0)\frac{\mathrm{d}\overrightarrow{M_1}}{\mathrm{d}t}(\varphi(s_0)) = \varphi'(s_0)\frac{\mathrm{d}\overrightarrow{M_1}}{\mathrm{d}t}(t_0)$$
donc les vecteurs $\dfrac{\mathrm{d}\overrightarrow{M_2}}{\mathrm{d}t}(s_0)$ et $\dfrac{\mathrm{d}\overrightarrow{M_1}}{\mathrm{d}t}(t_0)$ sont colinéaires.
- ✗ Si $M(t_0)$ est un point régulier de la courbe paramétrée par M, on en déduit une représentation paramétrique de la tangente $\mathcal{T}$ en $M(t_0)$ avec les équivalences
$$A(x,y,z) \in \mathcal{T} \Leftrightarrow \overrightarrow{M(t_0)A} \in \operatorname{Vect}\left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0)\right) \Leftrightarrow \exists t \in \mathbb{R}, \begin{cases} x = x(t_0) + tx'(t_0) \\ y = y(t_0) + ty'(t_0) \\ z = z(t_0) + tz'(t_0) \end{cases}$$
- ✗ Si $M(t_0)$ est un point régulier de la courbe paramétrée par M, on en déduit une équation cartésienne de la tangente $\mathcal{T}$ en $M(t_0)$ avec les équivalences
$$A(x,y,z) \in \mathcal{T} \quad \Leftrightarrow \quad \overrightarrow{M(t_0)A} \in \operatorname{Vect}\left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0)\right) \quad \Leftrightarrow \quad \overrightarrow{M(t_0)A} \wedge \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t_0) = 0_{\mathbb{R}^3}$$
On en déduit un système de trois équations cartésiennes décrivant la droite $\mathcal{T}$. Ces trois équations sont toujours liées, ce qui permet d'en éliminer une.

Exercice 18.1.

On considère la courbe Γ paramétrée par $M(t) = (t, \sin(t), \cos(t))$, $t \in \mathbb{R}$.

1. Montrer que Γ est régulière de classe $\mathcal{C}^1$.

2. Soient $t_0 \in \mathbb{R}$ et $\mathcal{T}$ la tangente à Γ en $M(t_0)$.
 a. Déterminer une représentation paramétrique de $\mathcal{T}$.
 b. Déterminer une équation cartésienne de $\mathcal{T}$.

Définition 18.3. **Courbe plane, courbe gauche**

Une courbe de l'espace est dite **plane** si elle est tracée sur un plan. Elle est dite **gauche** sinon.

Exercice 18.2.

Soit Γ la courbe de représentation paramétrique $\begin{cases} x(t) &= 5\cos(t) \\ y(t) &= 5\sin(t) \\ z(t) &= t \end{cases}, t \in \mathbb{R}$.

1. Montrer que la courbe Γ est régulière et gauche.

2. Déterminer une représentation paramétrique de la tangente à Γ en $M(\pi)$.
3. Dessiner les projections orthogonales de Γ sur les plans (Oxy) et (Oxz).

Définition 18.4. **Plan normal**

Si $t_0 \in I$ est tel qu'il existe une tangente à Γ en $M(t_0)$, alors on appelle **plan normal** à Γ en $M(t_0)$ le plan orthogonal à la tangente à Γ en $M(t_0)$ et passant par $M(t_0)$.

18.2 Surfaces paramétrées de l'espace

18.2.1 Généralités

Définition 18.5. **Surface ou nappe paramétrée**

Soit $k \in \mathbb{N}$. On appelle **surface** (ou **nappe**) **paramétrée** de $\mathbb{R}^3$ de classe $\mathcal{C}^k$ toute application $M : (u,v) \mapsto \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{pmatrix}$ définie sur un ouvert $U \subset \mathbb{R}^2$ et à valeurs dans $\mathbb{R}^3$, où x, y, z sont de classe $\mathcal{C}^k$.
L'ensemble $\mathcal{S} = \{M(u,v);\ (u,v) \in U\}$ est appelé **support** de M. On dit aussi que $\mathcal{S}$ est paramétrée par M.

Dans la suite, on ne considérera que des nappes de classe $\mathcal{C}^1$.

Remarque 18.3.

Une surface paramétrée ne se résume pas à son support. Le support est un objet géométrique, tandis que la surface paramétrée M donne une information supplémentaire : elle indique la manière dont le support est parcouru.
On peut décrire le même objet géométrique de plusieurs manières. Par exemple, les fonctions vectorielles

$$M_1 : (u,v) \mapsto (\cos(u), \sin(u), v) \quad \text{et} \quad M_2 : (u,v) \mapsto (\sin(u), \cos(u), v).$$

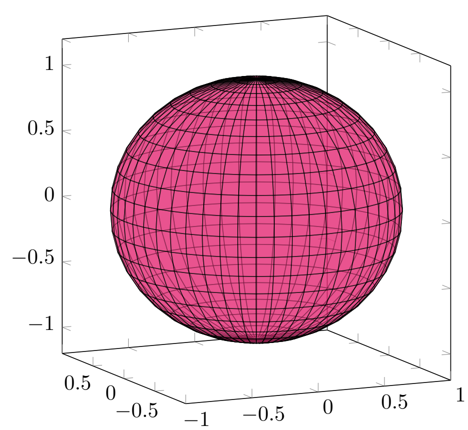
ont comme support le même cylindre.

Exemple 18.2. **Une surface paramétrée : la sphère**

On représente ci-contre le support de la surface paramétrée par

$$M : (u,v) \mapsto \begin{cases} \cos(u)\cos(v) \\ \sin(u)\cos(v) \\ \sin(v) \end{cases}$$

C'est une surface paramétrée de classe $\mathcal{C}^1$.



Exemple 18.3. **Un plan comme surface paramétrée**

Tout plan de l'espace est une surface paramétrée.
Ainsi, le plan passant par le point $A(x_0, y_0, z_0)$ et dirigé par $\vec{U}(a,b,c)$ et $\vec{V}(a',b',c')$ est paramétré par

$$M(u,v) = A + u\vec{U} + v\vec{V}$$

soit :

$$M : (u,v) \mapsto \begin{pmatrix} x_0 + ua + va' \\ y_0 + ub + vb' \\ z_0 + uc + vc' \end{pmatrix}.$$

Définition 18.6. **Courbes coordonnées**

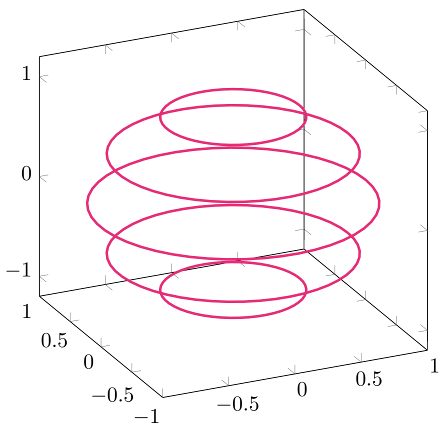
Soit $M : U \to \mathbb{R}^3$ une surface paramétrée de l'espace. On appelle **courbe coordonnée** de la surface paramétrée par M toute courbe paramétrée par une fonction vectorielle de la forme $u \mapsto M(u, v_0)$ ou $v \mapsto M(u_0, v)$ avec $(u_0, v_0) \in U$.

Remarque 18.4.

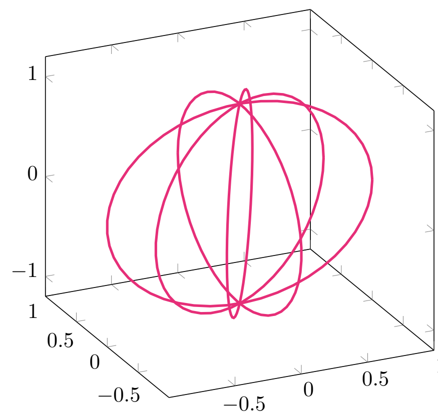
- ✗ Toutes les courbes coordonnées sont *incluses* dans la surface paramétrée : chaque point de leur support est un point de la surface.
- ✗ Ce sont les courbes utilisées par l'ordinateur pour *tracer* une surface paramétrée.
- ✗ Plus généralement, si $u : I \to \mathbb{R}$ et $v : I \to \mathbb{R}$ sont des fonctions de classe $\mathcal{C}^1$ sur un intervalle I, alors la fonction vectorielle $t \mapsto M(u(t), v(t))$ est une courbe paramétrée de l'espace *tracée sur la surface* paramétrée par f.

Exemple 18.4. **Courbes coordonnées de la sphère**

On considère à nouveau la sphère paramétrée par $M : (u, v) \mapsto \begin{cases} \cos(u)\cos(v) \\ \sin(u)\cos(v) \\ \sin(v) \end{cases}$.



Quelques courbes coordonnées paramétrées par $u \mapsto M(u, v_0)$



Quelques courbes coordonnées paramétrées par $v \mapsto M(u_0, v)$

Exercice 18.3.

Soit $\mathcal{S}$ la surface paramétrée par $\begin{cases} x &= v \\ y &= u^2 + 2v \\ z &= uv \end{cases}, (u, v) \in \mathbb{R}^2$.

1. Déterminer les courbes coordonnées de $\mathcal{S}$ passant par son point de paramètre $(-2, -2)$.
2. Déterminer et tracer l'intersection de $\mathcal{S}$ avec les plans (Oxy) et (Oxz).

18.2.2 Étude locale en un point régulier

Dans cette section, on considère un ouvert $U \subset \mathbb{R}^2$, et $\mathcal{S}$ une surface paramétrée par $M : (u, v) \mapsto \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix}$ de classe $\mathcal{C}^1$ sur U.

Définition 18.7. **Point régulier d'une surface paramétrée**

Soit $(u_0, v_0) \in U$. Le point $M(u_0, v_0)$ de $\mathcal{S}$ est dit **régulier** si la famille $\left(\frac{\partial M}{\partial u}(u_0, v_0), \frac{\partial M}{\partial v}(u_0, v_0)\right)$ est libre.
Dans le cas contraire, le point sera dit **stationnaire** ou **singulier**.
Une surface paramétrée est dite de paramétrage **régulier** si tous ses points sont réguliers.

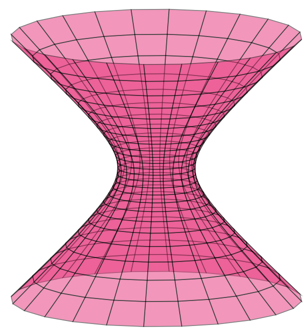
Remarque 18.5.
On déduit des propriétés du produit vectoriel qu'un point $M(u_0, v_0)$ est régulier si et seulement si

$$\frac{\partial M}{\partial u}(u_0, v_0) \wedge \frac{\partial M}{\partial v}(u_0, v_0) \neq \vec{0}.$$

Exemple 18.5. **Une surface paramétrée régulière : l'hyperboloïde à une nappe**

On représente ci-contre le support de la surface paramétrée par

$$M : (u, v) \mapsto \begin{cases} \cos(u)\mathrm{ch}(v) \\ \sin(u)\mathrm{ch}(v) \\ \mathrm{sh}(v) \end{cases}$$



C'est une surface paramétrée régulière de classe $\mathcal{C}^1$. On verra dans une section ultérieure qu'il s'agit d'une *surface réglée*.

En effet,

$$\frac{\partial M}{\partial u}(u, v) = \begin{pmatrix} -\sin(u)\mathrm{ch}(v) \\ \cos(u)\mathrm{ch}(v) \\ 0 \end{pmatrix}, \quad \frac{\partial M}{\partial v}(u, v) = \begin{pmatrix} \cos(u)\mathrm{sh}(v) \\ \sin(u)\mathrm{sh}(v) \\ \mathrm{ch}(v) \end{pmatrix}$$

et donc

$$\frac{\partial M}{\partial u}(u, v) \wedge \frac{\partial M}{\partial v}(u, v) = \begin{pmatrix} \cos(u)\mathrm{ch}^2(v) \\ \sin(u)\mathrm{ch}^2(v) \\ \mathrm{sh}(v)\mathrm{ch}(v) \end{pmatrix} \neq \vec{0}.$$

Définition 18.8. **Plan tangent**

Soit $(u_0, v_0) \in U$ tel que $M(u_0, v_0)$ est un point régulier de $\mathcal{S}$.
On appelle **plan tangent** à $\mathcal{S}$ en $M(u_0, v_0)$ plan passant par $M(u_0, v_0)$ et dirigé par $\frac{\partial M}{\partial u}(u_0, v_0)$ et $\frac{\partial M}{\partial v}(u_0, v_0)$.

Méthode 18.33. **Équation cartésienne d'un plan tangent**

Pour déterminer une équation cartésienne du plan tangent $\mathcal{P}$ en un point régulier $M(u_0, v_0)$, on peut utiliser la relation

$$P \in \mathcal{P} \iff \det\left(\overrightarrow{MP}, \frac{\partial M}{\partial u}(u_0, v_0), \frac{\partial M}{\partial v}(u_0, v_0)\right) = 0.$$

Exemple 18.6. **Plan tangent (en un point régulier à la sphère)**

En reprenant l'exemple de la sphère paramétrée par $M : (u, v) \mapsto \begin{cases} \cos(u)\cos(v) \\ \sin(u)\cos(v) \\ \sin(v) \end{cases}$, on a que $M\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \left(\frac{1}{2}, \frac{1}{2}, \frac{\sqrt{2}}{2}\right)$ est un point régulier de la surface.

Par définition, le plan tangent $\mathcal{P}$ à la surface en $M\left(\frac{\pi}{4}, \frac{\pi}{4}\right)$ est dirigé par les vecteurs

$$\frac{\partial M}{\partial u}\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{et} \quad \frac{\partial M}{\partial v}\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ \sqrt{2} \end{pmatrix}$$

On en déduit qu'une représentation paramétrique de $\mathcal{P}$ est

$$\begin{cases} x = \dfrac{1}{2} - \dfrac{u}{2} - \dfrac{v}{2} \\ y = \dfrac{1}{2} + \dfrac{u}{2} - \dfrac{v}{2} \\ z = \dfrac{\sqrt{2}}{2} + \dfrac{\sqrt{2}}{2} v \end{cases}$$

D'autre part, une équation cartésienne de $\mathcal{P}$ est

$$\begin{vmatrix} x - 1/2 & -1/2 & -1/2 \\ y - 1/2 & 1/2 & -1/2 \\ z - \sqrt{2}/2 & 0 & \sqrt{2}/2 \end{vmatrix} = 0 \quad \Leftrightarrow \quad \sqrt{2}x + \sqrt{2}y + 2z = 2\sqrt{2}.$$

Exercice 18.4.

Montrer que le cylindre paramétré par $(u, v) \mapsto \begin{cases} x = R\cos u \\ y = R\sin u \\ z = v \end{cases}$ est régulier, puis déterminer une équation de son plan tangent au point de coordonnées (x_0, y_0, z_0).

Définition 18.9. **Vecteur normal, droite normale à une surface en un point**

Soit $(u_0, v_0) \in U$ tel que $M(u_0, v_0)$ est un point régulier de $\mathcal{S}$.

- ✗ On appelle **vecteur normal** à la surface paramétrée en $M(u_0, v_0)$ tout vecteur normal au plan tangent à la surface en $M(u_0, v_0)$.
- ✗ On appelle **normale** (ou droite normale) à $\mathcal{S}$ en $M(u_0, v_0)$ la droite orthogonale au plan tangent à $\mathcal{S}$ en $M(u_0, v_0)$ et passant par ce point.

Remarque 18.6.

La droite normale en un point régulier $M(u_0, v_0)$ de la surface paramétrée par M est dirigé par le vecteur normal :

$$\vec{n} = \frac{\partial M}{\partial u}(u_0, v_0) \wedge \frac{\partial M}{\partial v}(u_0, v_0).$$

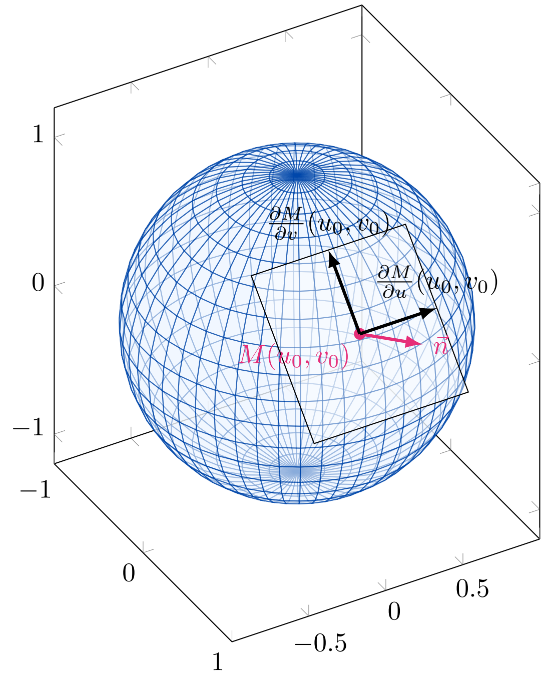
Exemple 18.7.

En reprenant notre exemple fil rouge de la sphère, on a que, pour tout $(u_0, v_0) \in \mathbb{R}^2$ avec $\cos(v_0) \neq 0$, (*les pôles ne sont pas des points réguliers pour ce paramétrage*) la droite normale à la surface paramétrée par M au point $M(u_0, v_0)$ est dirigé par

$$\vec{n} = \frac{\partial M}{\partial u}(u_0, v_0) \wedge \frac{\partial M}{\partial v}(u_0, v_0) = \begin{pmatrix} -\sin(u_0)\cos(v_0) \\ \cos(u_0)\cos(v_0) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} -\cos(u_0)\sin(v_0) \\ -\sin(u_0)\sin(v_0) \\ \cos(v_0) \end{pmatrix} = \cos(v_0) \begin{pmatrix} \cos(u_0)\cos(v_0) \\ \sin(u_0)\cos(v_0) \\ \sin(v_0) \end{pmatrix}.$$

Une représentation paramétrique de la droite normale au point $M(u_0, v_0)$ est

$$\begin{cases} x = \cos(u_0)\cos(v_0) + \cos(u_0)\cos(v_0)\, t \\ y = \sin(u_0)\cos(v_0) + \sin(u_0)\cos(v_0)\, t \\ z = \sin(v_0) + \sin(v_0)\, t \end{cases}$$



18.3 Équations cartésiennes

Dans cette section, $\mathcal{U}$ désigne un ouvert de $\mathbb{R}^3$.

18.3.1 Surfaces implicites de l'espace : généralités

Définition 18.10. **Surface définie par une équation cartésienne**

Soit $F : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ sur $\mathcal{U}$.
On appelle **surface d'équation (cartésienne)** $F(x, y, z) = 0$ l'ensemble de points

$$\mathcal{S} = \{M(x, y, z) : (x, y, z) \in \mathcal{U} \text{ et } F(x, y, z) = 0\}.$$

Remarque 18.7.
Cette autre définition de surface est très générale. On observe par exemple qu'avec cette définition, l'ensemble

$$\{M(x, y, z) \in \mathbb{R}^3 : \ y^2 + z^2 = 0\}$$

est une surface alors qu'il s'agit... de la droite (Ox). On se contentera quand même de cette définition qui a le mérite d'être simple.

Exemple 18.8. **Équation cartésienne de la sphère**

La sphère de centre $(x_0, y_0, z_0) \in \mathbb{R}^3$ et de rayon $R > 0$ est la surface d'équation (cartésienne)

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 - R^2 = 0.$$

Remarque 18.8.

- ✗ Tout plan de l'espace admet une équation cartésienne. En particulier, le plan (Oxy) admet $z = 0$ pour équation.
- ✗ Une même surface peut être décrite par plusieurs équations cartésiennes. Par exemple, le plan d'équation $x + z = 0$ admet aussi $(x + z)^2 = 0$ comme équation cartésienne.

Exercice 18.5.
Soit $\mathcal{C}$ le cylindre de révolution d'axe (Oz) et de rayon $R > 0$. Déterminer une équation cartésienne de $\mathcal{C}$.

Remarque 18.9.

- ✗ Si une surface possède une équation cartésienne de la forme $z = f(x, y)$, on dit que cette dernière est **explicite**. La surface est alors la représentation graphique dans $\mathbb{R}^3$ de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$.
  On obtient alors une paramétrisation de la surface en posant:
  $$\begin{cases} x = u \\ y = v \\ z = f(u, v) \end{cases}.$$
- ✗ La remarque ci-dessus s'adapte au cas des équations de la forme $y = f(x, z)$ ou $x = f(y, z)$ (l'essentiel est que l'on ait une coordonnée en fonction des deux autres). Par opposition, une équation cartésienne de la forme $F(x, y, z) = 0$ sera dite **implicite**.
- ✗ Il n'est pas toujours évident d'obtenir une paramétrisation d'une surface définie par une équation cartésienne implicite (et inversement).

18.3.2 Équation du plan tangent (en un point régulier)

Définition 18.11. **Point régulier d'une surface définie par une équation cartésienne**
Soit $\mathcal{S}$ une surface d'équation $F(x, y, z) = 0$.
Un point $M_0\,(x_0, y_0, z_0)$ de $\mathcal{S}$ est dit **régulier** si $\nabla F\,(M_0) \neq \vec{0}$.
Une surface dont tous les points sont réguliers est dite **régulière**.

On **admettra** dans la suite que si une surface définie par équation cartésienne $F(x, y, z) = 0$ (avec F de classe $\mathcal{C}^1$) est régulière, alors elle admet, en tout point, un paramétrage local régulier de classe $\mathcal{C}^1$.
On avait déjà admis un résultat analogue dans le chapitre précédent (**Théorème 15.11**).
Plus précisément, si au voisinage de $M(x_0, y_0, z_0)$, la surface $\mathcal{S}$ est régulière, alors $\mathcal{S}$ est le support d'une nappe paramétrée régulière (*i.e.* de paramétrage régulier) au voisinage du point (x_0, y_0, z_0) ; il existe alors un voisinage ouvert $\mathcal{U}_0 \subset \mathcal{U}$ de (x_0, y_0, z_0) et une nappe paramétrée $M : \mathcal{V} \subset \mathbb{R}^2 \to \mathbb{R}^3$ tels que
$$\forall (x, y, z) \in \mathcal{U}_0, \qquad [(x, y, z) \in \mathcal{S} \iff \exists (u, v) \in \mathcal{V},\ (x, y, z) = M(u, v)]\,.$$
Il découle de cette observation qu'une surface admet un plan tangent en tout point régulier, et une surface régulière en tout point.

La notion de point régulier ne dépend alors pas du paramétrage.

Théorème 18.2.
Soit F une fonction de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^3$. Soit $\mathcal{S}$ une surface d'équation cartésienne $F(x, y, z) = 0$.
Si (x_0, y_0, z_0) est un point régulier de $\mathcal{S}$, alors $\mathcal{S}$ admet un plan tangent $\mathcal{P}$ en (x_0, y_0, z_0) de vecteur normal $\nabla F(x_0, y_0, z_0)$ et d'équation cartésienne:
$$\frac{\partial F}{\partial x}(x_0, y_0, z_0) \cdot (x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0) \cdot (y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0) \cdot (z - z_0) = 0.$$

Preuve. La remarque qui précède l'énoncé du théorème garantit l'existence du plan tangent.
Soit $M :\ (u, v) \in \mathcal{U}_0 \mapsto (x(u, v), y(u, v), z(u, v))$ un paramétrage local régulier de $\mathcal{S}$ au voisinage de $(x_0, y_0, z_0) = M(u_0, v_0)$.
Alors, en appliquant la règle de la chaîne à la relation
$$\forall (u, v) \in \mathcal{U}, \quad 0 = F(M(u, v)) = F(x(u, v), y(u, v), z(u, v)),$$
on obtient
$$\forall (u, v) \in \mathcal{U}, \qquad 0 = \left\langle \nabla F(M(u, v)) | \frac{\partial M}{\partial u}(u, v) \right\rangle \quad \text{ et } 0 = \left\langle \nabla F(M(u, v)) | \frac{\partial M}{\partial v}(u, v) \right\rangle$$
Donc $\nabla F(M(u_0, v_0))$ est bien orthogonal au plan tangent $\mathcal{P}$. De plus, soit $A(x, y, z)$, on a
$$\begin{aligned} A \in \mathcal{P} \quad &\iff \quad \left\langle \overrightarrow{M(u_0, v_0)A} | \nabla F(x_0, y_0, y_0) \right\rangle = 0 \\ &\iff \quad (x - x_0) \cdot \frac{\partial F}{\partial x}(x_0, y_0, z_0) + (y - y_0) \cdot \frac{\partial F}{\partial y}(x_0, y_0, z_0) + (z - z_0) \cdot \frac{\partial F}{\partial z}(x_0, y_0, z_0) = 0. \end{aligned}$$
□

Remarque 18.10.
En d'autres termes, la normale à $\mathcal{S}$ en (x_0, y_0, z_0) est dirigée par $\nabla F(x_0, y_0, z_0)$.
Comme annoncé dans le chapitre précédent, ce résultat est une généralisation du (**Théorème 15.12**).

Exercice 18.6.
Déterminer le plan tangent en chacun des points de la sphère centrée en O et de rayon 1, et retrouver la normale en chaque point.

Dans le cas particulier des surfaces définies par une équation cartésienne explicite, on a le résultat suivant.

Théorème 18.3.
Soit $\mathcal{S}$ une surface d'équation cartésienne $z = f(x, y)$, avec f de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{V}$ de $\mathbb{R}^2$. Alors $\mathcal{S}$ est régulière.

Preuve. Une équation cartésienne de $\mathcal{S}$ est alors $F(x, y, z) = 0$ avec $F(x, y, z) = f(x, y) - z$ qui est de classe $\mathcal{C}^1$ sur $\mathcal{V} \times \mathbb{R}$ avec

$$\nabla F(x, y, z) = \begin{pmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \\ -1 \end{pmatrix} \neq \vec{0}.$$

□

Remarque 18.11.
Le résultat reste vrai si l'équation est de la forme $x = f(y, z)$ ou $y = f(x, z)$.

Définition 18.12. **Surface de niveau**
Soient $g : \mathcal{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{R}$. On appelle **surface de niveau** λ l'ensemble des points $(x, y, z) \in \mathcal{U}$ vérifiant

$$g(x, y, z) = \lambda.$$

Le résultat suivant est une généralisation du **Corollaire 15.13** énoncé au cours du chapitre précédent.

Corollaire 18.4.
Soit $g : \mathcal{U} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. En tout point régulier d'une surface de niveau de g, le gradient ∇g est orthogonal (au plan tangent) et orienté dans le sens des valeurs croissantes de g.

18.3.3 Courbe tracée sur une surface

Définition 18.13. **Courbe tracée sur une surface**
Soit $\mathcal{S}$ la surface d'équation $F(x, y, z) = 0$.
Une courbe Γ paramétrée par $t \mapsto \begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}$, $t \in I$, est dite **tracée sur la surface** $\mathcal{S}$ si

$$\forall t \in I, \quad F(x(t), y(t), z(t)) = 0.$$

Remarque 18.12.
Si la surface $\mathcal{S}$ est définie par un paramétrage $(u, v) \mapsto \begin{cases} x = x(u, v) \\ y = y(u, v) \\ z = z(u, v) \end{cases}$, $(u, v) \in \mathcal{U}$, une courbe Γ tracée sur $\mathcal{S}$ a une représentation paramétrique de la forme :

$$t \mapsto \begin{cases} x = x(u(t), v(t)) \\ y = y(u(t), v(t)) \\ z = z(u(t), v(t)) \end{cases}, t \in I,$$

où $\quad t \longmapsto (u(t), v(t))$ est $\mathcal{C}^1$ sur I et à valeurs dans $\mathcal{U}$.

Exemple 18.9. **Courbe paramétrée tracée sur une surface (paramétrée)**

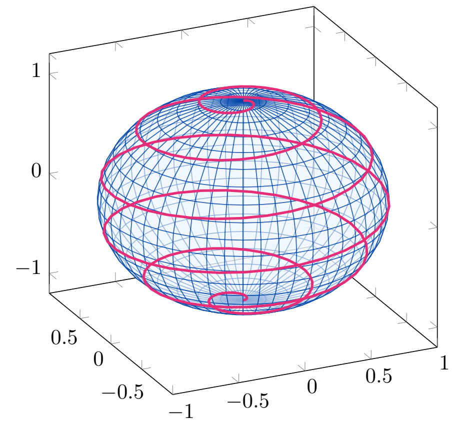
Un exemple de courbe gauche tracée sur une surface.
On représente ci-contre la courbe paramétrée par

$$t \mapsto M(12t, t),$$

où

$$M : (u,v) \mapsto \begin{cases} \cos(u)\cos(v) \\ \sin(u)\cos(v) \\ \sin(v) \end{cases}$$

est un paramétrage de la sphère déjà présenté dans les exemples précédents.



Exercice 18.7.
Soit $\mathcal{S}$ la surface d'équation $z^2 - xy - x + y - 4z + 4 = 0$, $\mathcal{P}$ le plan d'équation $x + y + z = 2$ et $\mathcal{C} = \mathcal{S} \cap \mathcal{P}$.

1. Déterminer une base orthonormée directe $(\vec{u}, \vec{v}, \vec{w})$ de $\mathbb{R}^3$ telle que $\vec{w}$ soit normal à $\mathcal{P}$.
2. Soit $A(1, -1, 2)$. Déterminer des équations cartésiennes de $\mathcal{C}$ dans le repère $(A; \vec{u}, \vec{v}, \vec{w})$.
3. En déduire la nature de $\mathcal{C}$ et ses caractéristiques géométriques.
4. Déterminer une représentation paramétrique de $\mathcal{C}$ dans $(O; \vec{i}, \vec{j}, \vec{k})$.

Remarque 18.13. **Une autre définition de plan tangent**
Le plan tangent en un point régulier d'une surface peut alors être défini comme la réunion des tangentes aux courbes régulières tracées sur la surface passant par ce point.
En effet, si $\gamma : t \mapsto M(u(t), v(t))$ est une courbe paramétrée régulière tracée sur la surface et passant par le point régulier $M(u_0, v_0)$ lorsque $t = t_0$, alors la tangente à la courbe en $M(u_0, v_0)$ est incluse dans le plan tangent à la surface en $M(u_0, v_0)$. En effet, il suffit d'appliquer la règle de la chaine pour obtenir

$$\frac{\mathrm{d}\overrightarrow{\gamma}}{\mathrm{d}t}(t_0) = u'(t_0)\frac{\partial M}{\partial u}(u_0, v_0) + v'(t_0)\frac{\partial M}{\partial v}(u_0, v_0)$$

alors la tangente à la courbe en $M(u_0, v_0)$ est incluse dans le plan tangent à la surface en $M(u_0, v_0)$.

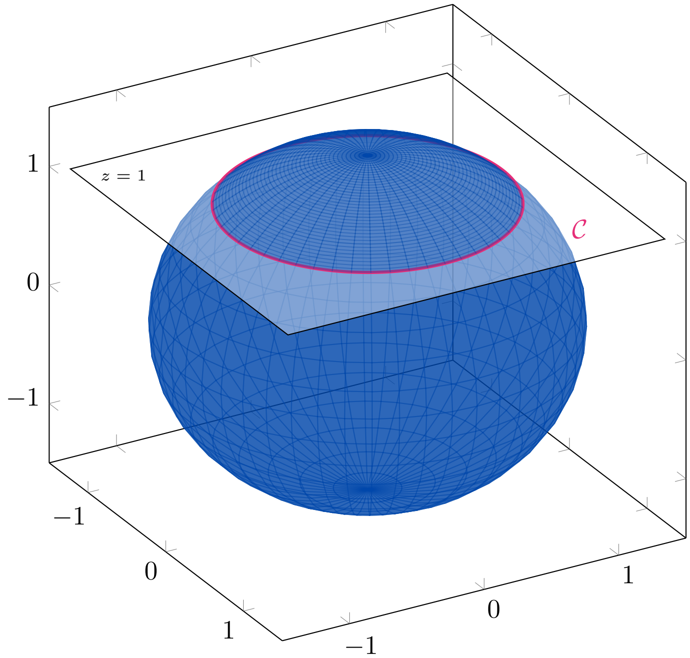
18.3.4 Système d'équations définissant une courbe de l'espace, intersection de surfaces

Définition 18.14. **Courbe implicite de l'espace**
Soient F_1 et F_2 deux fonctions de classe $\mathcal{C}^1$ définies sur l'ouvert $\mathcal{U} \subset \mathbb{R}^3$.
On appelle courbe (implicite) d'équation $\begin{cases} F_1(x,y,z) &= 0 \\ F_2(x,y,z) &= 0 \end{cases}$ l'intersection des surfaces $\mathcal{S}_1$ et $\mathcal{S}_2$ d'équations respectives $F_1(x,y,z) = 0$ et $F_2(x,y,z) = 0$, c'est à dire l'ensemble des points $(x,y,z) \in \mathcal{U}$ vérifiant $F_1(x,y,z) = F_2(x,y,z) = 0$.

Remarque 18.14.
Comme précédemment pour les surfaces, cette définition est très (trop?) générale, mais elle a le mérite d'être simple et on s'en contentera.
Par exemple, l'intersection de deux sphères dans l'espace peut être une sphère, un cercle, un point ou l'ensemble vide.

Exemple 18.10.
Le cercle $\mathcal{C}$ de centre $(0,0,1)$ de rayon 1 dans le plan d'équation $z = 1$ est l'intersection de la sphère de centre $(0,0,0)$ et de rayon $\sqrt{2}$ avec le plan d'équation $z = 1$, donc une équation de $\mathcal{C}$ est

$$\begin{cases} x^2 + y^2 + z^2 = 2 \\ z = 1 \end{cases}$$



18.3.5 Tangente en un points régulier d'une courbe implicite

Définition 18.15. **Point régulier**

Soit $\mathcal{C}$ la courbe de l'espace d'équation $F_1(x,y,z) = F_2(x,y,z) = 0$.
On dit qu'un point $(x_0, y_0, z_0) \in \mathcal{C}$ est **régulier** si les vecteurs $\nabla F_1\,(x_0, y_0, z_0)$ et $\nabla F_2\,(x_0, y_0, z_0)$ ne sont pas colinéaires.
Dans le cas contraire, on dit que le point $(x_0, y_0, z_0) \in \mathcal{C}$ est **singulier**.

On admet à nouveau que, que si $(x_0, y_0, z_0) \in \mathcal{C}$ est un point régulier de la courbe, alors $\mathcal{C}$ est le support d'une courbe paramétrée régulière au voisinage du point (x_0, y_0, z_0).
Plus précisément, il existe un voisinage ouvert $\mathcal{U}_0 \subset U$ de (x_0, y_0, z_0) et une courbe paramétrée régulière $\gamma : I \to \mathbb{R}^3$ telle que

$$\forall (x,y,z) \in \mathcal{U}_0, \quad [(x,y,z) \in \mathcal{C} \iff \exists t \in I, \quad (x,y,z) = \gamma(t)] \,.$$

En particulier et d'après les parties précédentes, on en déduit que la courbe $\mathcal{C}$ admet une tangente au point (x_0, y_0, z_0).

Théorème 18.5. **Tangente en un point régulier d'une courbe implicite**

Soit $\mathcal{C}$ la courbe de l'espace d'équation $F_1(x,y,z) = F_2(x,y,z) = 0$.
En tout point régulier $(x_0, y_0, z_0) \in \mathcal{C}$, la courbe $\mathcal{C}$ admet une tangente qui est l'intersection du plan tangent de la surface d'équation $F_1(x,y,z) = 0$ et du plan tangent de la surface d'équation $F_2(x,y,z) = 0$.

Preuve. La remarque qui précède l'énoncé du théorème garantit l'existence de la tangente.
Soit $\gamma :\ t \in I \mapsto (x(t), y(t), z(t))$ un paramétrage local régulier de $\mathcal{C}$ au voisinage de $(x_0, y_0, z_0) = \gamma(t_0)$. Alors, en appliquant la règle de la chaîne (puis en évaluant en $t = t_0$) à la relation

$$\forall t \in I, \quad 0 = F_1(\gamma(t)) = F_2(\gamma(t)),$$

on obtient

$$\left\langle \nabla F_1(\gamma(t_0) | \frac{\mathrm{d}\overrightarrow{\gamma}}{\mathrm{d}t}(t_0) \right\rangle = \left\langle \nabla F_2(\gamma(t_0) | \frac{\mathrm{d}\overrightarrow{\gamma}}{\mathrm{d}t}(t_0) \right\rangle = 0.$$

Or, $\nabla F_1(x_0, y_0, z_0)$ est normal au plan tangent $\mathcal{P}_1$ en (x_0, y_0, z_0) à la surface d'équation $F_1(x,y,z) = 0$ d'équation donc $\dfrac{\mathrm{d}\overrightarrow{\gamma}}{\mathrm{d}t}(t_0) \in \mathcal{P}_1$.
De même, $\dfrac{\mathrm{d}\overrightarrow{\gamma}}{\mathrm{d}t}(t_0) \in \mathcal{P}_2$. La tangente à $\mathcal{C}$ en (x_0, y_0, z_0) étant dirigée par $\dfrac{\mathrm{d}\overrightarrow{\gamma}}{\mathrm{d}t}(t_0)$, on a bien inclusion de celle-ci dans l'intersection des deux plans tangents.
La régularité du point $(x_0, y_0, z_0) \in \mathcal{C}$ assure que les deux plans considérés ne sont pas parallèles, donc leur intersection est bien une droite et on a égalité. $\square$

Méthode 18.34.

Si le point $M\left(x_0,y_0,z_0\right)$ est un point régulier de la courbe d'équation $F_1(x,y,z)=F_2(x,y,z)=0$, on en déduit une équation cartésienne de la tangente $\mathcal{T}$ en $\left(x_0,y_0,z_0\right)$ avec les équivalences

$$\begin{aligned} A(x,y,z)\in\mathcal{T} &\Leftrightarrow \left\{\begin{array}{l} \left\langle\overrightarrow{MA}\mid\nabla F_1\left(x_0,y_0,z_0\right)\right\rangle=0 \\ \left\langle\overrightarrow{MA}\mid\nabla F_2\left(x_0,y_0,z_0\right)\right\rangle=0 \end{array}\right. \\ &\Leftrightarrow \left\{\begin{array}{l} \left(x-x_0\right)\frac{\partial F_1}{\partial x}\left(x_0,y_0,z_0\right)+\left(y-y_0\right)\frac{\partial F_1}{\partial y}\left(x_0,y_0,z_0\right)+\left(z-z_0\right)\frac{\partial F_1}{\partial z}\left(x_0,y_0,z_0\right)=0 \\ \left(x-x_0\right)\frac{\partial F_2}{\partial x}\left(x_0,y_0,z_0\right)+\left(y-y_0\right)\frac{\partial F_2}{\partial y}\left(x_0,y_0,z_0\right)+\left(z-z_0\right)\frac{\partial F_2}{\partial z}\left(x_0,y_0,z_0\right)=0. \end{array}\right. \end{aligned}$$

En particulier, la tangente en un point régulier de la courbe $\left(x_0,y_0,z_0\right)\in\mathcal{C}$ est dirigée par le vecteur

$$\nabla F_1\left(x_0,y_0,z_0\right)\wedge\nabla F_2\left(x_0,y_0,z_0\right).$$

Exercice 18.8.

Soit Γ la courbe d'équations $\left\{\begin{array}{l} x^3+y^3+z^3=8 \\ xy+yz+xz+1=0 \end{array}\right.$.

1. Vérifier que le point $A(1,2,-1)$ est un point régulier de la courbe Γ.
2. Déterminer un paramétrage de la tangente à Γ en A.
3. Calculer les coordonnées des points d'intersection de Γ avec le plan (Oxy).

Définition 18.16. **Section plane**

On appelle **section plane** d'une surface $\mathcal{S}$ l'intersection entre $\mathcal{S}$ et un plan.

Remarque 18.15.

Si $\mathcal{P}$ est un plan de l'espace, alors le plan tangent de $\mathcal{P}$ en tout point $\left(x_0,y_0,z_0\right)\in\mathcal{P}$ est le plan $\mathcal{P}$ lui-même.
Comme la tangente en un point à une courbe obtenue comme intersection de surface est l'intersection des plans tangents des surfaces (en ce même point), la tangente en un point à une section plane $\mathcal{P}\cap\mathcal{S}$ est l'intersection du plan tangent en ce point à $\mathcal{S}$ et du plan réalisant la section.

Définition 18.17. **Lignes de niveau**

Les sections planes d'une surface $\mathcal{S}$ avec un plan d'équation $z=z_0$ avec $z_0\in\mathbb{R}$ s'appelle les **lignes de niveaux** de la surface $\mathcal{S}$.

Exemple 18.11.

Dans l'**Exemple 18.4**, la première figure correspond aux sections planes de la surface $\mathcal{S}$ d'équation $x^2+y^2+z^2=1$ avec les plans d'équation $z=z_0$. Ce sont les lignes de niveau z_0 de $\mathcal{S}$.

La seconde figure représente des sections planes de $\mathcal{S}$.

Proposition. 18.6.

Si $\mathcal{C}$ est une courbe tracée sur une surface $\mathcal{S}$ et si M est un point régulier à la fois de $\mathcal{S}$ et de $\mathcal{C}$, alors la tangente en M à $\mathcal{C}$ est incluse dans le plan tangent en M à $\mathcal{S}$.

18.4 Surfaces réglées

Dans cette section, on désigne par «surface» soit une surface paramétrée par une fonction vectorielle, soit une surface définie par une équation cartésienne.

Définition 18.18. **Surface réglée**

Une surface $\mathscr{S}$ de l'espace est dite **réglée** s'il existe une famille de droites $\left(\mathscr{D}_u\right)_{u\in I}$ où I est un intervalle de $\mathbb{R}$ telle que $\mathscr{S}=\bigcup_{u\in I}\mathscr{D}_u$.

Définition 18.19. **Génératrices d'une surface réglée**

Les droites contenues dans une surface réglée $\mathscr{S}$ sont appelées **génératrices** de la surface $\mathscr{S}$.

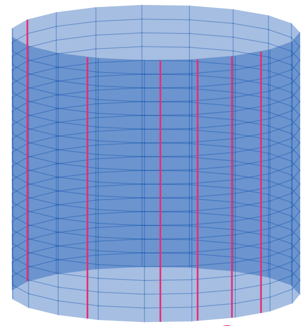
Exemple 18.12.

i. Les plans de l'espace sont des surfaces réglées.

ii. Le cylindre de révolution d'axe (Oz) et de rayon $R > 0$ est une surface réglée.

On déduit immédiatement de la paramétrisation usuelle du cylindre de révolution que ses génératrices sont les droites $\mathcal{D}_u$ $(u \in \mathbb{R})$ paramétrées par

$$\mathcal{D}_u : v \mapsto \begin{pmatrix} \cos(u) \\ \sin(u) \\ 0 \end{pmatrix} + v \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$



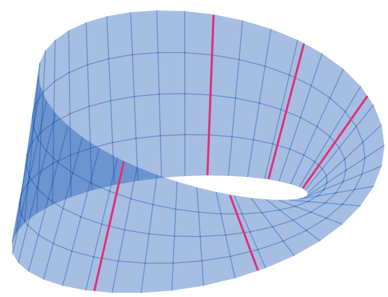
iii. Le *Ruban de Möbius* est une surface réglée.

Avec le paramétrage

$$M : (u, v) \in]0, 2\pi] \times [-1, 1] \mapsto \begin{cases} \left(1 + \frac{v}{2} \cos\left(\frac{u}{2}\right)\right) \cos(u) \\ \left(1 + \frac{v}{2} \cos\left(\frac{u}{2}\right)\right) \sin(u) \\ \frac{v}{2} \sin\left(\frac{u}{2}\right) \end{cases},$$

on peut en effet écrire

$$M(u, v) = \begin{pmatrix} \cos(u) \\ \sin(u) \\ 0 \end{pmatrix} + v \begin{pmatrix} \frac{1}{2} \cos\left(\frac{u}{2}\right) \cos(u) \\ \frac{1}{2} \cos\left(\frac{u}{2}\right) \sin(u) \\ \frac{1}{2} \sin\left(\frac{u}{2}\right) \end{pmatrix} = A(u) + v\vec{d}(u)$$



Théorème 18.7.
Une surface $\mathcal{S}$ est réglée si et seulement si elle admet un paramétrage de ses points de la forme:

$$M(u, v) = A(u) + v \times \vec{d}(u)$$

où $(u, v) \in I \times \mathbb{R}$ avec I un intervalle de $\mathbb{R}$, et $A, \vec{d}$ deux fonctions à valeurs dans $\mathbb{R}^3$ telles que $\vec{d}$ ne s'annule pas.

Remarque 18.16.
Avec les notations de l'énoncé ci-dessus, on a $\mathcal{S} = \bigcup_{u \in I} \mathcal{D}_u$ où la droite $\mathcal{D}_u$ est définie comme passant par le point $A(u)$ est dirigée par le vecteur $\vec{d}(u)$.

Exemple 18.13. **Le *Pringles* comme surface réglée**

Le paraboloïde hyperbolique («*Pringles*») d'équation $x^2 - y^2$ est une surface réglée.
En effet, on peut montrer que cette surface correspond à celle engendrée par la famille de droites $(\mathcal{D}_u)_{u \in \mathbb{R}}$ où $\mathcal{D}_u$ est la droite passant par $A_u(u, 0, u^2)$ et dirigée par $\vec{d}(1, 1, 2u)$ et paramétrée par

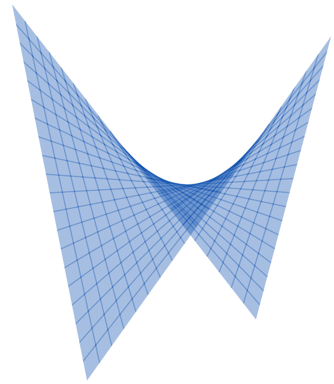
$$\mathcal{D}_u : v \in \mathbb{R} \mapsto \begin{pmatrix} u + v \\ v \\ u^2 + 2uv \end{pmatrix}.$$

Ainsi

$$(x,y,z) \in \bigcup_{u\in\mathbb{R}} \mathcal{D}_u \iff \exists (u,v)\in\mathbb{R}^2,\quad \begin{cases} x &= u+v \\ y &= v \\ z &= u^2+2uv \end{cases}$$

$$\Rightarrow \begin{cases} u &= x-y \\ v &= y \\ z &= (x-y)^2+2(x-y)y = x^2-y^2 \end{cases}$$

$$\Rightarrow (x,y,z)\in\mathcal{S}.$$



Ainsi, la surface engendrée par la famille de droites $(\mathcal{D}_u)_{u\in\mathbb{R}}$ est incluse dans $\mathcal{S}$.
Réciproquement, si $(x,y,z)\in\mathcal{S}$ vérifie $z=x^2-y^2$ alors, en posant $u=x-y$ et $v=y$, on obtient

$$\begin{pmatrix} x\\ y\\ z\end{pmatrix} = \begin{pmatrix} x\\ y\\ x^2-y^2\end{pmatrix} = \begin{pmatrix} u+v\\ v\\ (u+v)^2-v^2\end{pmatrix} = \begin{pmatrix} u+v\\ v\\ u^2+2uv\end{pmatrix} \in \mathcal{D}_u.$$

Exercice 18.9.

Montrer que l'hélicoïde droit paramétré par $\begin{cases} x = v\cos u \\ y = v\sin u \\ z = u\end{cases}$ où $(u,v)\in\mathbb{R}^2$ est une surface réglée.

Proposition. 18.8.
Soit $\mathscr{S}$ une surface réglée de l'espace. Le plan tangent en un point régulier de $\mathscr{S}$ contient les génératrices de $\mathscr{S}$ passant par ce point.

18.5 Surfaces de révolution

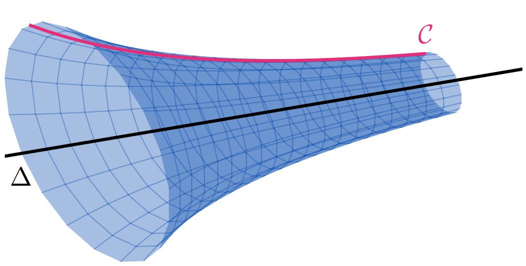
Définition 18.20. **Surface de révolution**

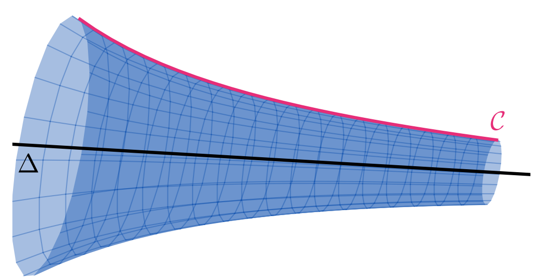
Une surface $\mathscr{S}$ de l'espace est dite **de révolution** si elle s'obtient par la rotation d'une courbe $\mathscr{C}$ autour d'une droite Δ.

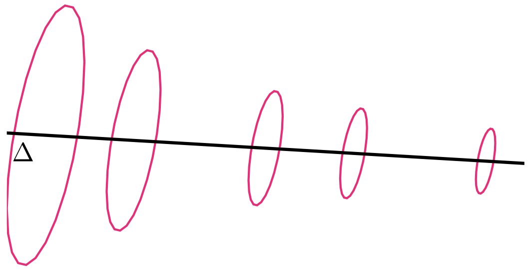
i. La droite Δ s'appelle un **axe de révolution** de la surface $\mathscr{S}$.
ii. On appelle **plan méridien** de $\mathcal{S}$ tout plan contenant Δ.
iii. On appelle **parallèle** de $\mathscr{S}$ les cercles d'axe Δ rencontrant $\mathscr{S}$.
iv. On appelle **méridienne** de $\mathscr{S}$ l'intersection entre un plan méridien et la surface $\mathscr{S}$.

Exemple 18.14. **Une surface de révolution**

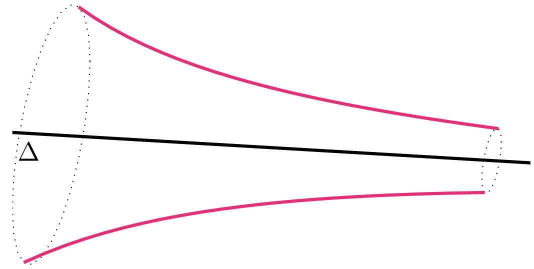
Les deux figures ci-dessous représentent (avec deux angles de vue différents) la même surface de révolution $\mathcal{S}$ obtenue par rotation de la courbe $\mathcal{C}$ autour de l'axe de révolution Δ.







Quelques parallèles de la surface de révolution $\mathcal{S}$ autour de Δ.



Une méridienne de la surface de révolution $\mathcal{S}$ autour de Δ.

Remarque 18.17.

Il est immédiat que si $\mathcal{C}$ est une droite, la surface de révolution $\mathcal{S}$ obtenue par rotation de $\mathcal{C}$ autour de Δ est réglée. On retrouve qu'un cylindre de révolution est une surface réglée, tout comme un cône de révolution.

Remarque 18.18.

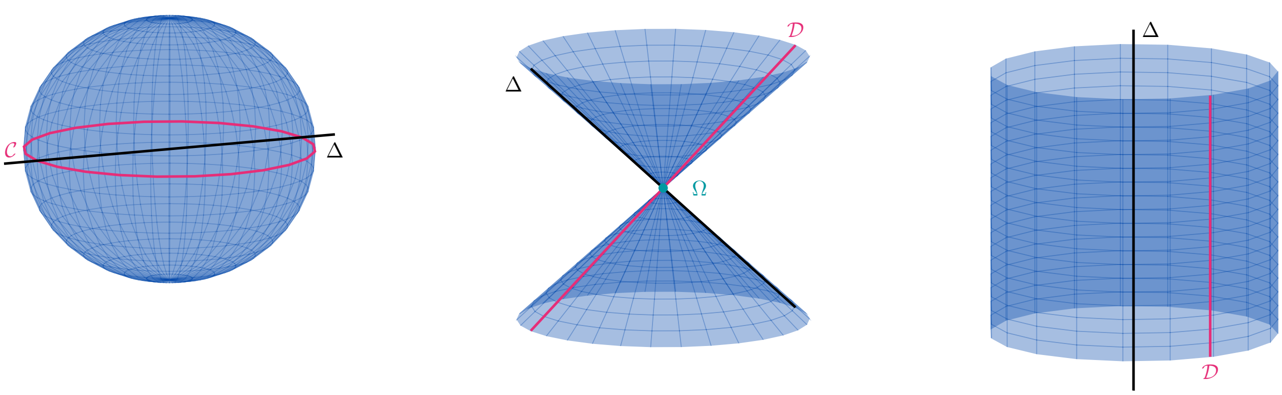
- ✗ Une surface de révolution d'axe Δ est invariante par les rotations d'axe Δ.
- ✗ Une surface de révolution est la réunion de ses méridiennes, qui se déduisent les unes des autres par rotation autour de Δ.
- ✗ Une surface de révolution est la réunion de ses parallèles, qui sont toutes des réunions de cercles.

Exemple 18.15. **Quelques surfaces de révolutions usuelles**

i. La surface de révolution engendrée par la rotation d'un cercle $\mathcal{C}$ autour de n'importe lequel de ses diamètres est une *sphère*, possédant le même rayon et le même centre que le cercle.

ii. La surface de révolution engendrée par la rotation d'une droite $\mathcal{D}$ autour d'un axe Δ sécant en un point Ω est un *cône de révolution*. Ses méridiennes sont les génératrices du cône, et elles forment avec Δ un angle constant (appelé demi-angle au sommet).

iii. La surface de révolution engendrée par la rotation d'une droite $\mathcal{D}$ autour d'un axe Δ parallèle est un *cylindre de révolution*. Ses méridiennes sont les génératrices du cylindre, et ses parallèles sont des cercles de même rayon R que le cylindre, avec $R = \mathrm{d}(\Delta, \mathcal{D})$.



Dans la suite (et la fin de ce chapitre), on considère une surface de révolution obtenue en faisant tourner une courbe $\mathscr{C}$ autour d'un axe Δ.

Dans le cas où Δ est l'un des axes du repères, il est facile d'obtenir une paramétrisation de la surface $\mathscr{S}$. Si $\gamma : I \to \mathbb{R}^3$ est une paramétrisation de la courbe $\mathscr{C}$, alors une paramétrisation de $\mathscr{S}$ est donnée par

$$M(x,y,z) \in \mathscr{S} \quad \Leftrightarrow \quad \exists (t,\theta) \in I \times \mathbb{R}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_\Delta(\theta) \times \gamma(t)$$

où $R_\Delta(\theta)$ est la matrice de la rotation d'axe Δ et d'angle θ.
On rappelle les expressions des matrices de rotation atour d'un des axes du repère ; plus généralement, on renvoie au cours du **Chapitre 15**.

Axe des abscisses (Ox)	**Axe des ordonnées** (Oy)	**Axe des cotes** (Oz)
$R_\Delta(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$	$R_\Delta(\theta) = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$	$R_\Delta(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Exemple 18.16. **Paramétrisation du tore**

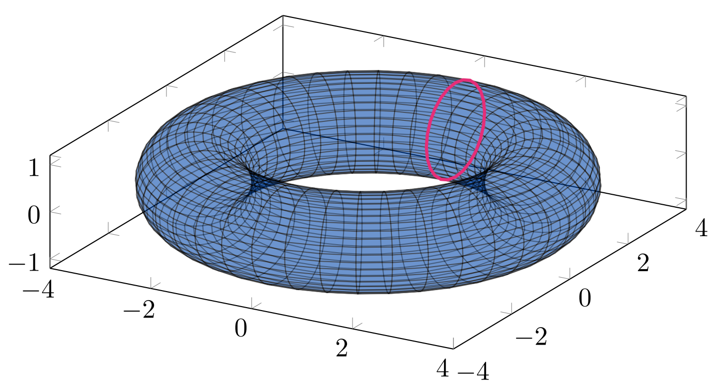
Dans le plan d'équation $x = 0$, on considère le cercle $\mathscr{C}$ de centre $(0,3,0)$ et de rayon 1.
Déterminons une paramétrisation de la surface de révolution $\mathscr{S}$ obtenue en faisant tourner le cercle $\mathscr{C}$ autour de l'axe des cotes (Oz).
La courbe $\mathscr{C}$ est paramétrée par la fonction vectorielle $\gamma : \mathbb{R} \to \mathbb{R}$ définie par

$$\forall t \in \mathbb{R}, \quad \gamma(t) = \begin{pmatrix} 0 \\ 3 + \cos(t) \\ \sin(t) \end{pmatrix}$$

On en déduit qu'une représentation paramétrique de $\mathscr{S}$ est

$$M(x,y,z) \in \mathscr{S} \Leftrightarrow \exists (t,\theta) \in \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 3+\cos(t) \\ \sin(t) \end{pmatrix} = \begin{pmatrix} -\sin(\theta)(3+\cos(t)) \\ \cos(\theta)(3+\cos(t)) \\ \sin(t) \end{pmatrix}.$$

En utilisant cette paramétrisation, on peut tracer cette surface (que l'on appelle un *tore*).



Méthode 18.35. **Équation cartésienne d'une surface de révolution, par les parallèles**

On souhaite obtenir une équation cartésienne de la surface de révolution $\mathcal{S}$ obtenue par rotation de $\mathcal{C}$ (paramétrée par $\gamma : t \in I \mapsto (x(t), y(t), z(t))$) autour de Δ.

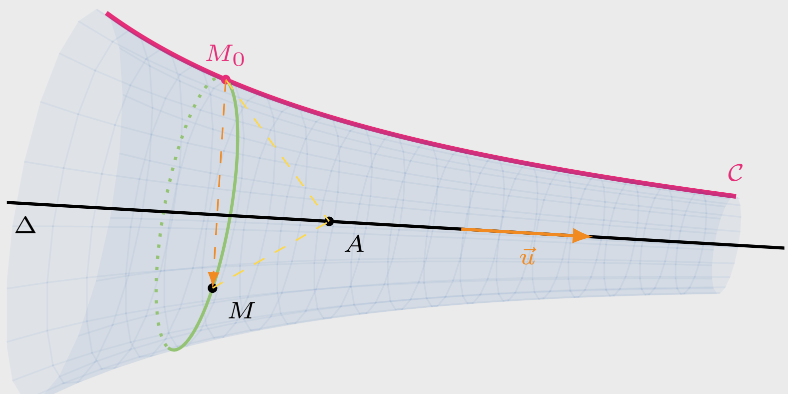
Soient $\vec{u}\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ un vecteur de directeur de Δ et $A(x_A, y_A, z_A) \in \Delta$ arbitraires.

M est un point de $\mathcal{S}$ si et seulement si il existe un parallèle de $\mathcal{S}$ sur lequel se trouve M, c'est à dire un cercle d'axe Δ, passant par un point $M_0 \in \mathcal{C}$, et dans un plan orthogonal à Δ. On peut écrire la chose comme tel :

$$\begin{aligned} M(x,y,z) \in \mathcal{S} \quad &\iff \quad \exists M_0 \in \mathcal{C}, \left\{ \begin{array}{ccc} \overrightarrow{M_0 M} & \perp & \vec{u} \\ AM & = & AM_0 \end{array} \right. \\ &\iff \quad \exists t \in I, \left\{ \begin{array}{ccc} a(x - x(t)) + b(y - y(t)) + c(y - y(t)) & = & 0 \\ (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2 & = & (x_A - x(t))^2 + (y_A - y(t))^2 + (z_A - z(t))^2 \end{array} \right. \end{aligned}$$

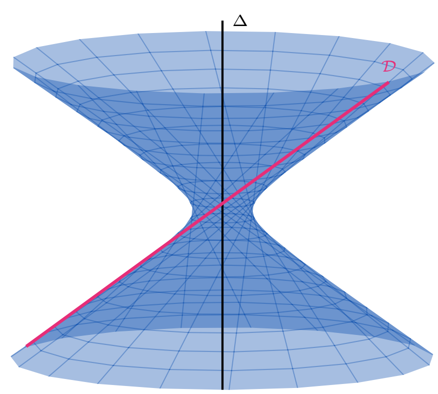
La première équation du système permet d'éliminer t dans la deuxième et d'obtenir ainsi une équation cartésienne de $\mathcal{S}$.

Il est pratique de choisir pour A le point O quand ce dernier est bien un point de l'axe de révolution.



Exercice 18.10.

Déterminer une équation de la surface de révolution obtenue par rotation de la droite $\mathcal{D}$ paramétrée par $t \in \mathbb{R} \mapsto (1+t, t, t)$ autour de l'axe (Oz).



Exercice 18.11.

Soit $\mathcal{S}$ la surface engendrée par la rotation autour de l'axe (Oz) de la courbe paramétrée par $\begin{cases} x = t \\ y = t^2 \\ z = t^2 \end{cases}$ où $t \in \mathbb{R}$.

1. Donner une paramétrisation de $\mathcal{S}$.
2. Obtenir une équation cartésienne de $\mathcal{S}$. Quelle méthode semble la plus pratique ?
3. Déterminer la nature et les éléments caractéristiques de la méridienne de $\mathcal{S}$ associée au plan d'équation $x = 0$.

19 Étude métrique des courbes

Dans ce dernier chapitre, on revient sur l'étude des courbes (paramétrées) planes introduites dans le **Chapitre 3**.
On y introduit la notion d'*abscisse curviligne*, qui permet de calculer la «longueur» d'une courbe du plan, ainsi que les notions de *centre et rayon de courbure*, qui permettent de définir le *cercle de courbure* en un point d'une courbe (le cercle dont la forme «épouse» localement la courbe au mieux).

Sauf mention du contraire, on considère tout au long du chapitre une courbe Γ paramétrée par

$$M : t \in I \longmapsto M(t) = (x(t), y(t)) \in \mathbb{R}^2.$$

19.1 Abscisse curviligne et longueur d'arc

19.1.1 Longueur d'un arc paramétré

Définition 19.1. **Longueur d'un arc régulier**

Soit $(a, b) \in I^2$ avec $a \leqslant b$. La **longueur** d'un arc paramétré par M du point $M(a)$ au point $M(b)$ est la quantité

$$\ell_\Gamma(a, b) = \int_a^b \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\| \mathrm{d}t = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\ \mathrm{d}t.$$

Proposition. 19.1.

La longueur d'un arc paramétré entre deux de ses points est une notion géométrique : elle ne dépend pas de la paramétrisation M de l'arc.

Preuve. Si $\varphi : J \to I$ est une fonction bijective de classe $\mathcal{C}^1$ définie sur un intervalle J de $\mathbb{R}$, alors les courbes paramétrées $M : I \to \mathbb{R}^2$ et par $\tilde{M} = M \circ \varphi : J \to \mathbb{R}^2$ ont le même support.
Supposons que la fonction φ est strictement croissante - donc $\varphi'(u) > 0$ - (le calcul est analogue si φ est strictement décroissante, il suffit de permuter les bornes ce qui est compensé avec le signe $-$ de $\varphi'(u)$), on obtient avec la formule du changement de variable que

$$\begin{aligned} \ell_M(a, b) &= \int_a^b \|M'(t)\|\, \mathrm{d}t \underset{t=\varphi(u)}{=} \int_c^d \|M'(\varphi(u))\|\, \varphi'(u)\mathrm{d}u = \int_c^d \|M'(\varphi(u))\varphi'(u)\|\, \mathrm{d}u \\ &= \int_c^d \left\|\tilde{M}'(u)\right\| \mathrm{d}u = \ell_{\tilde{M}}(c, d) \end{aligned}$$

□

Remarque 19.1.

Si M est la loi horaire d'un mouvement, le nombre $\left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}s} \right\|$ est la vitesse de l'objet à l'instant t.

La longueur de l'arc est alors le produit de la vitesse moyenne par la durée du mouvement entre $M(a)$ et $M(b)$: c'est bien la longueur au sens usuel du terme.

Exemple 19.1. **Longueur d'un segment**

Soient A, B deux points du plan et $M(t) = A + t\overrightarrow{AB},\ t \in [0, 1]$ un paramétrage du segment $\Gamma = [AB]$. On a : $\ell_\Gamma(0, 1) = \left\|\overrightarrow{AB}\right\|$.

Exemple 19.2. **Périmètre du cercle**

La longueur d'arc de la courbe Γ paramétré par la fonction $M : t \mapsto (\cos(t), \sin(t))$ du point $M(0) = (1,0)$ au point $M(2\pi) = (1,0)$ est égale à

$$\ell_\Gamma(0, 2\pi) = \int_0^{2\pi} \sqrt{(-\sin(t))^2 + (\cos(t))^2}\ \mathrm{d}t = \int_0^{2\pi} 1\ \mathrm{d}t = 2\pi.$$

On retrouve le périmètre du cercle trigonométrique.

Remarque 19.2.

Attention ! La quantité $\ell_\Gamma(a, b)$ est appelée **longueur d'arc** et non longueur de la courbe. En effet, il s'agit de la longueur parcourue par le point $M(t)$ entre les instants $t = a$ et $t = b$. En particulier, on peut parcourir plusieurs fois la courbe avant de s'arrêter.
La longueur mesure alors la **totalité du trajet parcouru** et pas simplement la distance sur la courbe entre le point de départ et le point d'arrivée.
Par exemple, la longueur d'arc de la courbe Γ paramétrée par $M : t \mapsto (\cos(t), \sin(t))$ du point $M(0) = (1,0)$ au point $M(4\pi) = (1,0)$ est égale à

$$\ell_\Gamma(0, 4\pi) = \int_0^{4\pi} \sqrt{(-\sin(t))^2 + (\cos(t))^2}\ \mathrm{d}t = \int_0^{4\pi} 1\ \mathrm{d}t = 4\pi$$

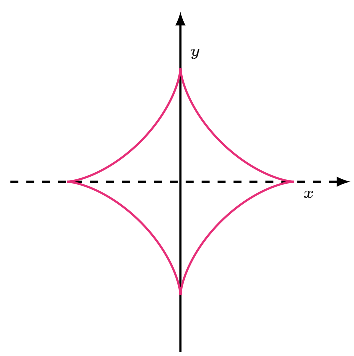
tandis que le périmètre du cercle est seulement de 2π. Cette différence provient du fait que lorsque la variable t décrit l'intervalle $[0, 4\pi]$, le point $M(t)$ effectue deux fois le tour du cercle trigonométrique.

Exercice 19.1. **Longueur d'arc de l'astroïde**

Calculer la longueur d'arc du support l'astroïde paramétrée par

$$\left\{ \begin{array}{rcl} x(t) & = & a\cos^3(t) \\ y(t) & = & a\sin^3(t) \end{array} \right. , \qquad t \in [0, 2\pi]$$

avec $a > 0$.



19.1.2 Abscisse curviligne

Définition 19.2. **Abscisse curviligne**

On appelle **abscisse curviligne** pour Γ toute application $s \in \mathcal{C}^1(I)$ telle que

$$\forall t \in I,\ s'(t) = \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\|.$$

Remarque 19.3.

Le théorème fondamental de l'analyse affirme que la fonction

$$t \mapsto \sqrt{\left\langle \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \mid \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\rangle} = \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\|$$

admet une infinité primitives (de classe $\mathcal{C}^1$) et donc Γ admet une infinité d'abscisses curvilignes, deux d'entre elles étant égales à une constante près.
On se donne alors un point de référence $t_0 \in I$ et on parle alors d'abscisse curviligne d'origine t_0, comme énoncé plus bas.

Exercice 19.2.
On considère la courbe représentative $\mathcal{C}$ de la fonction ch qui admet pour paramétrage $M : t \longmapsto (t, \mathrm{ch}(t))$.
Calculer l'abscisse curviligne d'origine 0 pour $\mathcal{C}$.

Remarque 19.4.
Si $s : I \to \mathbb{R}$ est une abscisse curviligne pour Γ, et s'il existe $t_0 \in I$ tel que $s(t_0) = 0$, alors

$$\forall t \in I,\ s(t) = \int_{t_0}^{t} \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\| \mathrm{d}t = \ell_\Gamma(t_0, t),$$

et le réel t_0 est appelé l'**origine** de s sur Γ.
La formule $s(t) = \ell_\Gamma(t_0, t)$ donne l'étymologie de l'abscisse curviligne: c'est une abscisse que l'on lirait sur la courbe et non plus sur les axes de coordonnées. Le signe indique si l'on se trouve «avant» ou «après» t_0.

Exemple 19.3.
L'abscisse curviligne d'origine 0 de la courbe Γ paramétrée par $M :\ t \mapsto (\cos(t), \sin(t))$ est la fonction

$$s : t \mapsto \int_0^t \sqrt{(-\sin(u))^2 + (\cos(u))^2}\ \mathrm{d}u = \int_0^t 1\ \mathrm{d}u = t.$$

Par définition, l'abscisse curviligne s d'origine t_0 est une fonction de de classe $\mathcal{C}^1$. De plus, si, comme ici, la courbe paramétrée par M est **régulière**, on a

$$\forall t \in I, \quad s'(t) = \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\| > 0.$$

Par le théorème de la bijection monotone (et son extension pour le caractère $\mathcal{C}^1$ de la bijection réciproque), on en déduit que s réalise une bijection de I sur l'intervalle $J = s(I)$ et que son application réciproque $s^{-1} : J \to I$ est de classe $\mathcal{C}^1$ sur J.
L'application $\tilde{M} = M \circ s^{-1} : J \to \mathbb{R}^2$ est un nouveau paramétrage de la courbe.

Remarque 19.5. **Paramétrisation par longueur d'arc pour une courbe régulière**
Considérons une paramétrisation M d'une courbe régulière Γ.

- ✗ En pratique, en posant $s = \varphi(t)$, on considère que toute fonction $h : I \to \mathbb{R}$ de la variable t est aussi une fonction de la variable s via $t = \varphi^{-1}(s)$. En particulier, la formule de dérivation d'une composée donne

$$\frac{\mathrm{d}h}{\mathrm{d}s} = \frac{\mathrm{d}h}{\mathrm{d}t} \times \frac{\mathrm{d}t}{\mathrm{d}s} = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^{-1} \times \frac{\mathrm{d}h}{\mathrm{d}t} = \frac{1}{\|f'(t)\|} \times \frac{\mathrm{d}h}{\mathrm{d}t}$$

- ✗ En tout point, le vecteur $\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}s}$ est alors de norme 1. En effet, d'après la formule précédente, on a

$$\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}s} = \frac{1}{\|M'(t)\|} \times \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t} = \frac{M'(t)}{\|M'(t)\|} \quad \Longrightarrow \quad \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}s} \right\| = 1.$$

  On obtient alors un parcours de Γ *à vitesse constante*.

19.2 Repère de Frenet

19.2.1 Repère de Frenet

Définition 19.3. **Reprère de Frenet associé à un point régulier**

Soit $M(t)$ un point **régulier** de Γ.
On appelle **repère de Frenet** le la courbe paramétrée au point $M(t)$ le repère **orthonormé direct** $(M(t), \vec{T}(t), \vec{N}(t))$ où $\vec{T}(t)$ est le vecteur tangent unitaire de Γ en $M(t)$ défini par

$$\vec{T}(t) = \frac{1}{\left\|\frac{\mathrm{d}\vec{M}}{\mathrm{d}t}(t)\right\|} \frac{\mathrm{d}\vec{M}}{\mathrm{d}t}(t) = \frac{1}{\sqrt{x'(t)^2 + y'(t)^2}} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}.$$

Il découle de cette définition que

$$\vec{N}(t) = \frac{1}{\sqrt{x'(t)^2 + y'(t)^2}} \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix}.$$

Remarque 19.6.

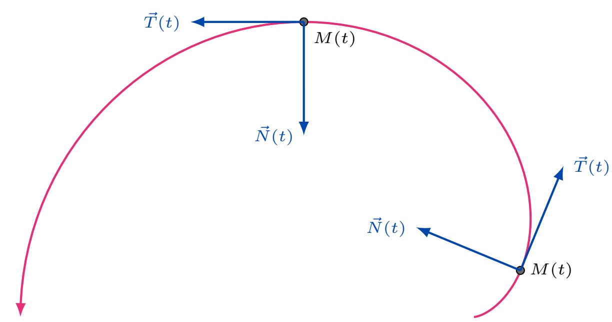
De manière abusive, on pourra écrire $\vec{T}(t) = \dfrac{\mathrm{d}\vec{M}}{\mathrm{d}s}$.
Cette relation a le mérite de souligner que si la courbe est paramétrée par l'abscisse curviligne, alors il suffit de dériver le vecteur position pour obtenir le vecteur tangente unitaire.

Exemple 19.4. **Repère de Frenet et cardioïde**

On considère la (demi-)cardioïde paramétrée par

$$\left\{ \begin{array}{rcl} x(t) & = & 2\cos(t) - \cos(2t) \\ y(t) & = & 2\sin(t) - \sin(2t) \end{array} \right. , \qquad t \in [0, \pi].$$

Soit $M(t)$, $t \neq 0$, un point **régulier**.



Observant que

$$(\sin(2t) - \sin(t))^2 + (\cos(t) - \cos(2t))^2 = 4\sin^2\left(\frac{t}{2}\right),$$

on obtient que

$$\overrightarrow{T(t)} = \frac{1}{2\sin\left(\frac{t}{2}\right)} \begin{pmatrix} \sin(2t) - \sin(t) \\ \cos(t) - \cos(2t) \end{pmatrix},$$

et

$$\overrightarrow{N(t)} = \frac{1}{2\sin\left(\frac{t}{2}\right)} \begin{pmatrix} \cos(2t) - \cos(t) \\ \sin(2t) - \sin(t) \end{pmatrix}.$$

Exercice 19.3.

Déterminer le repère de Frenet en tout point régulier de la parabole $\begin{cases} x = 2pt^2 \\ y = 2pt \end{cases}$ où $p > 0$ est fixé.

Exercice 19.4.
Déterminer le repère de Frenet en tout point régulier de l'astroïde.

19.3 Courbure et dévelopée d'un arc

19.3.1 Courbure et détermination angulaire

Dans cette section, on suppose que le paramétrage M de Γ est de classe $\mathcal{C}^2$.

Par définition du repère de Frenet, on a la relation

$$\forall t \in I, \quad \langle \overrightarrow{T}(t) \mid \overrightarrow{T}(t) \rangle = \|\overrightarrow{T}(t)\|^2 = 1.$$

Comme on a supposé que M est $\mathcal{C}^2$, la fonction vectorielle $\vec{T}: \ t \mapsto \overrightarrow{T}(t)$ est dérivable. En dérivant la relation précédente, on obtient que

$$\forall t \in I, \quad \left\langle \frac{\mathrm{d}\vec{T}}{\mathrm{d}t}(t) \middle| \ \vec{T}(t) \right\rangle = 0$$

Finalement, en utilisant la relation $\frac{\mathrm{d}\vec{T}}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s} \times \frac{\mathrm{d}\vec{T}}{\mathrm{d}t}$, on aboutit à

$$\forall t \in I, \quad \left\langle \frac{\mathrm{d}\vec{T}}{\mathrm{d}s}(t) \middle| \ \vec{T}(t) \right\rangle = 0.$$

On en déduit que les vecteurs $\frac{\mathrm{d}\vec{T}}{\mathrm{d}s}(t)$ et $\vec{N}(t)$ sont colinéaires pour tout $t \in I$, ce qui nous permet de considérer la définition suivante.

Définition 19.4. **Courbure**

i. Soit $M(t)$ un point régulier de Γ.

On appelle **courbure** en $M(t)$ à Γ le réel $\gamma(t)$ défini par la relation $\frac{\mathrm{d}\vec{T}}{\mathrm{d}s}(t) = \gamma(t)N(t)$.

ii. Si Γ est régulière (et donc M est un paramétrage régulier), on appelle **courbure** de M la fonction $\gamma: \ I \to \mathbb{R}$ définie par

$$\forall t \in I, \quad \frac{\mathrm{d}\vec{T}}{\mathrm{d}s}(t) = \gamma(t)N(t)$$

Remarque 19.7.
La courbure est une notion géométrique ; elle ne dépend pas de la paramétrisation de la courbe.

Exemple 19.5. **Courbure d'une droite**

Considérons une droite $\mathcal{D}$ passant par (x_0, y_0) et dirigée par $\vec{u}\begin{pmatrix} a \\ b \end{pmatrix}$ et donc paramétrée par

$$M: t \in \mathbb{R} \longmapsto (x_0 + at, y_0 + bt).$$

Pour tout $t \in \mathbb{R}$, un calcul immédiat donne

$$\overrightarrow{T}(t) = \frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix} a \\ b \end{pmatrix}, \qquad \overrightarrow{N}(t) = \frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix} -b \\ a \end{pmatrix}, \quad \text{et} \quad \frac{\mathrm{d}s}{\mathrm{d}t}(t) = \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\| = \sqrt{a^2+b^2}.$$

Il suit que

$$\frac{\mathrm{d}\vec{T}}{\mathrm{d}s}(t) = \frac{\mathrm{d}t}{\mathrm{d}s} \times \frac{\mathrm{d}\vec{T}}{\mathrm{d}t}(t) = \frac{1}{\sqrt{a^2+b^2}} \times \vec{0} = \vec{0} = 0 \times \overrightarrow{N}(t)$$

et la courbure de la droite est nulle.

Exemple 19.6. **Courbure d'un cercle**

Considérons maintenant le cercle $\mathcal{C}$ de centre $(x_0, y_0) \in \mathbb{R}^2$ et de rayon $R > 0$, paramétré par

$$M: t \mapsto (x_0 + R\cos(t),\ y_0 + R\sin(t)).$$

Pour tout $t \in \mathbb{R}$, on a par un calcul direct que

$$\overrightarrow{T}(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}, \quad \overrightarrow{N}(t) = \begin{pmatrix} -\cos(t) \\ -\sin(t) \end{pmatrix} \quad \text{et} \quad \frac{\mathrm{d}s}{\mathrm{d}t}(t) = \left\| \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\| = R.$$

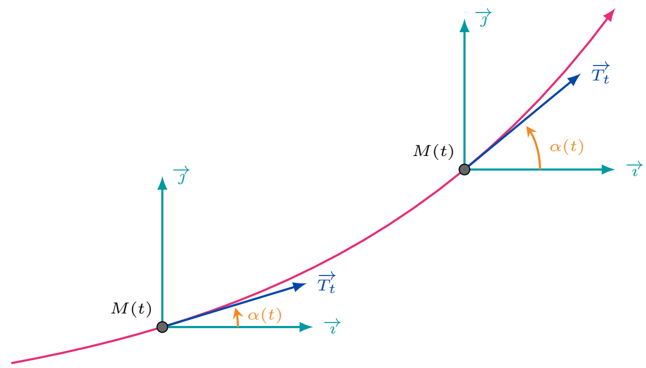
Il suit que

$$\frac{\mathrm{d}\vec{T}}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s} \times \frac{\mathrm{d}\vec{T}}{\mathrm{d}t} = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^{-1} \times \frac{\mathrm{d}\vec{T}}{\mathrm{d}t} = \frac{1}{R} \times \overrightarrow{N}(t)$$

donc la courbure γ est constante de valeur $1/R$.

Continuons nos observations. Soit $t \in \mathbb{R}$ tel quel $M(t)$ est régulier. Le vecteur $\overrightarrow{T}(t)$ est unitaire ; il existe alors $\alpha(t)$ tel que

$$\overrightarrow{T}(t) = \cos(\alpha(t))\vec{\imath} + \sin(\alpha(t))\vec{\jmath}.$$



Théorème & Définition 19.1. **Théorème de relèvement et détermination angulaire**

Soit M un paramétrage de Γ de classe $\mathcal{C}^1$, tel que tout point $M(t)$ soit régulier.
Il existe alors une application $\alpha : I \to \mathbb{R}$ de classe $\mathcal{C}^1$ telle que:

$$\forall t \in I,\ \overrightarrow{T}(t) = \cos(\alpha(t))\vec{\imath} + \sin(\alpha(t))\vec{\jmath}.$$

La fonction α est appelée **détermination angulaire** associée au repère de Frenet.

Preuve. Résultat admis. □

Théorème 19.2. **Formules de Frenet**

Si $\alpha : I \to \mathbb{R}$ est une application de classe $\mathcal{C}^1$ telle que

$$\forall t \in I, \quad \overrightarrow{T}(t) = \begin{pmatrix} \cos(\alpha(t)) \\ \sin(\alpha(t)) \end{pmatrix},$$

alors, on a les relations, notant $\vec{N} : t \mapsto \vec{N}(t)$,

$$\frac{\mathrm{d}\overrightarrow{T}}{\mathrm{d}s}(t) = \gamma(t)\overrightarrow{N}(t), \quad \frac{\mathrm{d}\overrightarrow{N}}{\mathrm{d}s}(t) = -\gamma(t)\overrightarrow{T}(t), \quad \frac{\mathrm{d}\alpha}{\mathrm{d}s}(t) = \gamma(t).$$

Remarque 19.8.

La troisième formule de Frenet nous permet d'interpréter géométriquement la courbure:

- ✗ La courbure est positive si et seulement si la fonction α est croissante. Dans ce cas, la courbe «tourne à gauche (selon l'orientation du plan euclidien)».
- ✗ La courbure est négative si et seulement si la fonction α est décroissante. Dans ce cas, la courbe «tourne à droite (selon l'orientation du plan euclidien)».
- ✗ Plus la courbure est grande en valeur absolue, plus la variation de α est rapide : la courbe «prend un virage serré».
  A contrario, plus la courbure est petite en valeur absolue, plus la variation de α est lente : la courbe «prend un virage ample».

Exemple 19.7.

Reprenons l'**Exemple 19.6** du cercle.
Pour tout $t \in [0, 2\pi]$,

$$\overrightarrow{T}(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} = \begin{pmatrix} \cos\left(t + \frac{\pi}{2}\right) \\ \sin\left(t + \frac{\pi}{2}\right) \end{pmatrix},$$

ainsi $\alpha : t \mapsto t + \dfrac{\pi}{2}$ est une détermination angulaire associée au repère de Frenet. On observe que

$$\frac{\mathrm{d}\alpha}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s} \times \frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{1}{R} \times \alpha'(t) = \frac{1}{R},$$

qui est bien la valeur qu'on avait trouvé pour la courbure en $M(t)$.

Remarque 19.9.

Calculer une courbure nécessite donc de dériver un quotient dont le dénominateur est une racine, ce qui peut s'avérer très lourd... On peut parfois contourner le problème en observant et en raisonnant de la manière suivante. Attention, au cas où on choisirait cette approche, **il est nécessaire de répéter en intégralité** les arguments et étapes ci-dessous. Posons

$$\vec{T}(t) = \varphi(t)\vec{u}(t), \qquad \text{avec} \quad \varphi(t) = \frac{1}{s'(t)} = \frac{1}{\sqrt{x'^2(t) + y'^2(t)}} \quad \text{et} \quad \vec{u}(t) = M'(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}.$$

Comme $\vec{u}$ et $\vec{T}$ sont colinéaires, il est clair que $\vec{u} \perp \vec{N}$ et donc $\langle \vec{u}, \vec{N} \rangle = 0$. Il suit que

$$\begin{aligned} \langle \vec{T}\,'(t),\ \vec{N}(t) \rangle &= \left\langle \varphi'(t)\vec{u}(t) + \varphi(t)\vec{u}\,'(t),\ \vec{N}(t) \right\rangle \\ &= \left\langle \varphi(t)\vec{u}\,'(t),\ \vec{N}(t) \right\rangle = \varphi(t) \left\langle \vec{u}\,'(t),\ \vec{N}(t) \right\rangle. \end{aligned}$$

Mais, par la première formule de Frenet, $\vec{T}\,'(t) = \dfrac{\mathrm{d}T}{\mathrm{d}s} \times \dfrac{\mathrm{d}s}{\mathrm{d}t} = \gamma(t)s'(t)\vec{N}(t)$, ce qui donne (car $\vec{N}$ est de norme 1),

$$s'(t)\gamma(t) = s'(t)\gamma(t)\|\vec{N}(t)\| = \langle \vec{T}\,'(t),\ \vec{N}(t) \rangle = \varphi(t) \left\langle \vec{u}\,'(t),\ \vec{N}(t) \right\rangle$$

ou encore (car $\varphi = 1/s'$) et $\vec{N}(t) = \varphi\vec{v}(t)$ avec $\vec{v}(t) = \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix}$,

$$\gamma(t) = \frac{\varphi(t)}{s'(t)} \left\langle \vec{u}\,'(t),\ \vec{N}(t) \right\rangle = \frac{1}{s'(t)^3} \langle \vec{u}\,'(t), \vec{v}(t) \rangle = \frac{-x''(t)y'(t) - y''(t)x'(t)}{\sqrt{x'^2(t) + y'^2(t)}^3}.$$

On peut refaire en intégralité ce calcul mais on ne peut pas citer cette formule directement.

Exercice 19.5. **Extrait Oral Math I 2023**

On considère la courbe :

$$\Gamma : \begin{cases} x(t) &= 3t - 3t^3 \\ y(t) &= 3t^3 \end{cases}.$$

1. Tracer rapidement la courbe Γ
2. Déterminer le repère de Frénet en un point de Γ.
3. Déterminer le rayon de courbure en un point de Γ.

19.3.2 Cercle osculateur et développée d'un arc

On commence par rappeler la définition suivante.

Définition 19.5. **Point birégulier**

Un point $M(t)$ de Γ est dit **birégulier** si la famille $\left(\dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t),\ \dfrac{\mathrm{d}^2\overrightarrow{M}}{\mathrm{d}t^2}(t) \right)$ est libre.

Γ est birégulière si tout point de Γ est birégulier.

En admettant la formule obtenue ci-avant, qu'on rappelle **hors-programme**,

$$\gamma(t) = \frac{1}{\left\| \dfrac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) \right\|^3} \det \left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t),\ \frac{\mathrm{d}^2\overrightarrow{M}}{\mathrm{d}t^2}(t) \right),$$

on voit que $M(t)$ est birégulier si et seulement si $\gamma(t) \neq 0$.

Définition 19.6. **Cercle, centre et rayon de courbure**

Soit $M(t)$ un point birégulier de la courbe.

i. On appelle **rayon de courbure** au point $M(t)$ le nombre $R(t) = \dfrac{1}{\gamma(t)}$.

ii. On appelle **centre de courbure** au point $M(t)$ le point $C(t)$ défini par la relation $C(t) = M(t) + R(t)\overrightarrow{N}(t)$.

iii. On appelle **cercle de courbure** (ou cercle osculateur) au point $M(t)$ le cercle de centre $C(t)$ et de rayon $|R(t)|$.

Définition 19.7. **Dévelopée d'un arc birégulier**

On appelle **développée** d'un arc **birégulier** l'ensemble de ses centres de courbures.

Exemple 19.8. **Chaînette**

Considérons la courbe de la fonction $f :\ t \longmapsto \mathrm{ch}(t)$ dont une paramétrisation birégulière est naturellement

$$M :\ t \longmapsto (t, \mathrm{ch}(t)).$$

On a immédiatement, pour tout $t \in \mathbb{R}$,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = s'(t) = \left\| \frac{\mathrm{d}\vec{M}}{\mathrm{d}t}(t) \right\| = \sqrt{1 + \mathrm{sh}^2(t)} = \mathrm{ch}(t).$$

Il suit que

$$\overrightarrow{T}(t) = \begin{pmatrix} \dfrac{1}{\mathrm{ch}(t)} \\ \\ \dfrac{\mathrm{sh}(t)}{\mathrm{ch}(t)} \end{pmatrix}, \quad \overrightarrow{N}(t) = \begin{pmatrix} -\dfrac{\mathrm{sh}(t)}{\mathrm{ch}(t)} \\ \\ \dfrac{1}{\mathrm{ch}(t)} \end{pmatrix}, \quad \text{puis} \quad \frac{\mathrm{d}\vec{T}}{\mathrm{d}t}(t) = \begin{pmatrix} -\dfrac{\mathrm{sh}(t)}{\mathrm{ch}^2(t)} \\ \\ \dfrac{1}{\mathrm{ch}^2(t)} \end{pmatrix} = \frac{\mathrm{d}t}{\mathrm{d}s} \times \overrightarrow{N}(t)$$

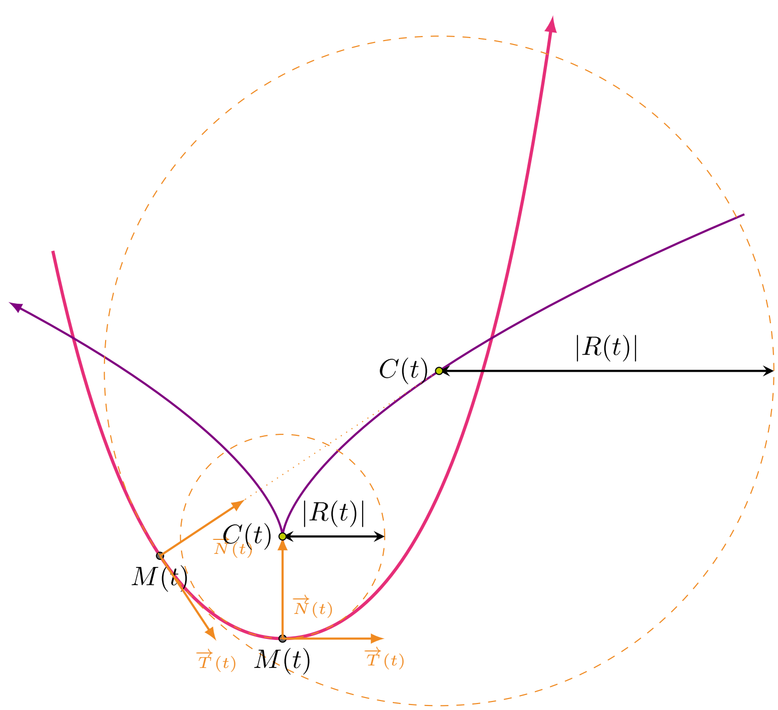
ce qui implique que

$$\frac{\mathrm{d}\vec{T}}{\mathrm{d}s}(t) = \frac{\mathrm{d}t}{\mathrm{d}s} \times \frac{\mathrm{d}\vec{T}}{\mathrm{d}t}(t) = \left(\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 \times \overrightarrow{N}(t) = \gamma(t)\overrightarrow{N}(t).$$

Finalement, on obtient :

$$R(t) = \frac{1}{\gamma(t)} = \mathrm{ch}^2(t), \qquad C(t) = (t - \mathrm{ch}(t)\mathrm{sh}(t), 2\mathrm{ch}(t))\,.$$

Sa développée est la courbe paramétrée par $t \mapsto C(t)$, représentée en violet sur la figure ci-dessous.



Exercice 19.6.

On considère l'ellipse paramétrée par

$$t \in [-\pi, \pi] \longmapsto (2\cos(t), \sin(t)).$$

1. Vérifier qu'il s'agit d'un paramétrage birégulier.

2. Montrer que, pour tout $t \in [-\pi, \pi]$, on a $\gamma(t) = \dfrac{2}{\left(3\sin^2(t)+1\right)^{3/2}}$.

3. En déduire que la développée de l'ellipse est une astroïde dont on précisera un paramétrage.

Théorème 19.3.

La développée d'un arc birégulier est l'enveloppe de la famille de ses normales.

Preuve. Il s'agit de montrer que l'enveloppe de la famille $(\mathcal{D}_t)_{t\in I}$ est bien la courbe paramétrée par

$$C: \ t \in I \longmapsto M(t) + R(t)\overrightarrow{N}(t).$$

Soit $t \in I$ fixé. La droite $\mathcal{D}_t$ normale à la courbe Γ en $M(t)$ est elle-même paramétrée par

$$\mathcal{D}_t : M(t) + \lambda\overrightarrow{N}(t), \quad \lambda \in \mathbb{R}.$$

On cherche alors un paramétrage de l'enveloppe $\mathcal{E}$ de la forme

$$\mu : t \in I \longmapsto M(t) + \lambda(t)\overrightarrow{N}(t),$$

où λ est une fonction de classe $\mathcal{C}^2$ à déterminer. Il suffit finalement de montrer que $\lambda(t) = R(t)$.
La tangente à $\mathcal{E}$ en $\mu(t)$ est dirigée par

$$\frac{\mathrm{d}\overrightarrow{\mu}}{\mathrm{d}t}(t) = \frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t) + \lambda'(t)\overrightarrow{N}(t) + \lambda(t) \times \frac{\mathrm{d}\overrightarrow{N}}{\mathrm{d}t}(t).$$

La tangente ci-dessus et la droite $\mathcal{D}_t$ sont alors confondues si et seulement si les vecteurs $\dfrac{\mathrm{d}\overrightarrow{\mu}}{\mathrm{d}t}(t)$ et $\overrightarrow{N_t}$ sont colinéaires, ce qui est équivalent à ce que leur déterminant soit nul. Mais, par propriétés du déterminant, puis par les formules de Frenet,

$$\begin{aligned}
\det\left(\frac{\mathrm{d}\overrightarrow{\mu}}{\mathrm{d}t}(t),\, \overrightarrow{N}(t)\right) &= \det\left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t}(t),\, \overrightarrow{N}(t)\right) + \lambda(t)\det\left(\frac{\mathrm{d}\overrightarrow{N}}{\mathrm{d}t}(t),\, \overrightarrow{N}(t)\right) \\
&= \frac{\mathrm{d}s}{\mathrm{d}t} \times \det\left(\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}s}(t),\, \overrightarrow{N}(t)\right) + \frac{\mathrm{d}s}{\mathrm{d}t} \times \lambda(t)\det\left(\frac{\mathrm{d}\overrightarrow{N}}{\mathrm{d}s}(t),\, \overrightarrow{N}(t)\right) \\
&= \frac{\mathrm{d}s}{\mathrm{d}t} \times \det\left(\overrightarrow{T}(t),\, \overrightarrow{N}(t)\right) + \frac{\mathrm{d}s}{\mathrm{d}t} \times \lambda(t)\det\left(-\gamma(t)\overrightarrow{T}(t),\, \overrightarrow{N}(t)\right)
\end{aligned}$$

Or, la base $\left(\overrightarrow{T}(t), \overrightarrow{N}(t)\right)$ est directe, ce qui veut dire que $\det\left(\overrightarrow{T}(t), \overrightarrow{N}(t)\right) = 1$. Il suit que

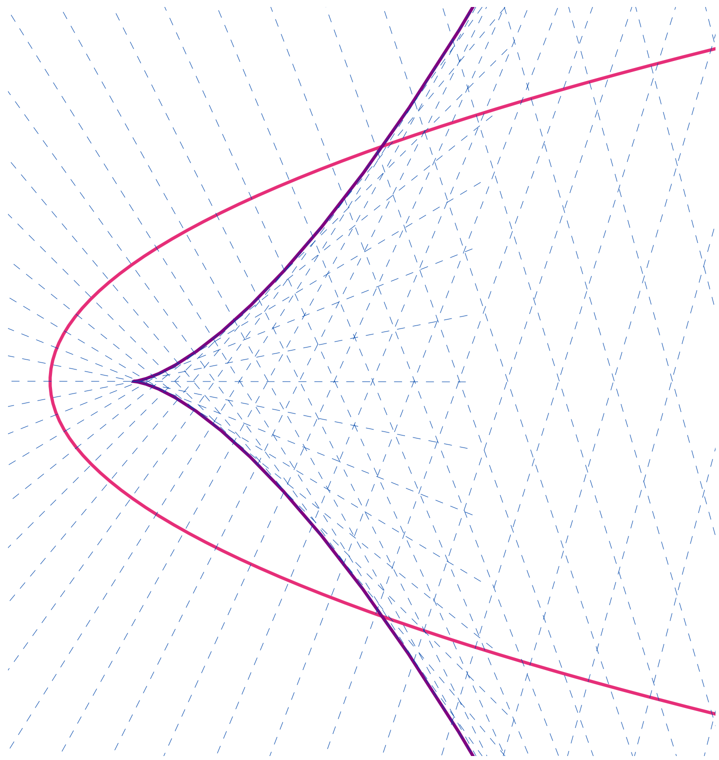
$$\frac{\mathrm{d}\overrightarrow{\mu}}{\mathrm{d}t}(t) \text{ et } \overrightarrow{N}(t) \text{ colinéaires} \iff 0 = \frac{\mathrm{d}s}{\mathrm{d}t}\left(1 - \gamma(t)\lambda(t)\right) \iff \lambda(t) = \frac{1}{\gamma(t)} = R(t),$$

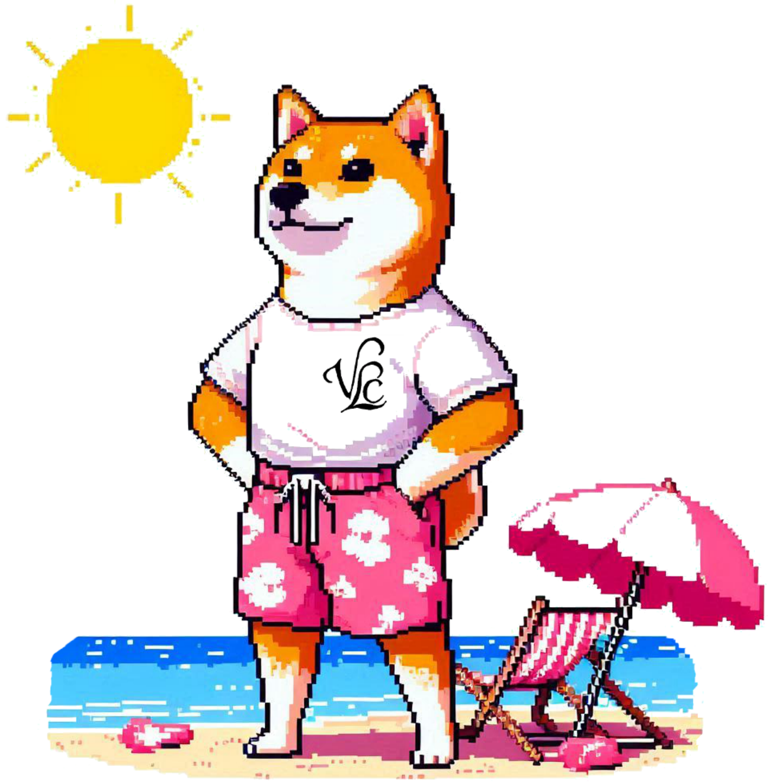
ce qu'on voulait et conclut la démonstration. □

Exercice 19.7.

Déterminer la développée de la parabole paramétrée par $\begin{cases} x = 2pt^2 \\ y = 2pt \end{cases}$ où $p > 0$ est fixé.

Indication : On déterminera l'enveloppe de la famille des normales de la parabole.




VLC