

MATHÉMATIQUES TSI2

SOMMAIRE

1 Révision : Autour des suites et des fonctions réelles	1
Résumé de cours	3
Méthodes	24
2 Révision : nombres complexes et polynômes	29
Résumé de cours	31
Méthodes	40
Vrai/Faux	48
3 Révision : algèbre linéaire	49
Résumé de cours	51
Méthodes	65
4 Révision : Probabilités finies	77
Résumé de cours	79
5 Variables aléatoires discrètes finies	89
Résumé de cours	91
6 Déterminants	103
Résumé de cours	105
Méthodes	110
7 Réduction des endomorphismes	115
Résumé de cours	117
Méthodes	121
Énoncé des exercices	129
Indications	131
Corrigé des exercices	132
8 Intégration sur un intervalle	137
Résumé de cours	139
Méthodes	147
9 Séries numériques	161
Résumé de cours	163
Méthodes	167
10 Power series	179
Résumé de cours	181
Méthodes	186

11 Linear differential equation and differential system	191
Résumé de cours	193
Méthodes	197
Énoncé des exercices	209
Indications	210
Corrigé des exercices	211
12 Pre-Hilbert and euclidean spaces	213
Résumé de cours	215
Méthodes	220
Énoncé des exercices	228
Indications	229
Corrigé des exercices	230
13 Fourier series	231
Résumé de cours	233
Énoncé des exercices	236
Indications	237
Corrigé des exercices	238
14 Functions of several real variables	239
Résumé de cours	241
Méthodes	248
Énoncé des exercices	257
Indications	258
Corrigé des exercices	259
15 Probability of countable universes	261
Résumé de cours	263
Méthodes	267
Énoncé des exercices	272
Indications	273
Corrigé des exercices	274
16 Discrete Random Variables	275
Résumé de cours	277
Méthodes	282
Énoncé des exercices	289
Indications	290
Corrigé des exercices	291
17 Isometries in euclidian space	293
Résumé de cours	295
Méthodes	299
18 Vector functions and parameterized curves	307
Résumé de cours	309
Méthodes	314
Énoncé des exercices	322
Indications	322
Corrigé des vrai/faux	324
Corrigé des exercices	325

Chapitre **1**

**Révision : Autour des suites et
des fonctions réelles**

■ les incontournables

- À propos des suites réelles :
 - ▶ savoir étudier leur convergence par les méthodes au programme, notamment avec des encadrements ou le théorème de la limite monotone ;
 - ▶ savoir la définition d'une limite finie avec ϵ ;
 - ▶ montrer la divergence d'une suite en utilisant des suites extraites ;
 - ▶ savoir montrer que deux suites sont adjacentes ;
 - ▶ savoir trouver le terme général, la limite et la somme partielle des termes d'une suite arithmétique ou géométrique ;
 - ▶ savoir trouver le terme général, la limite d'une suite arithmético-géométrique ;
 - ▶ savoir trouver le terme général, la limite d'une suite récurrente linéaire d'ordre 2.
- À propos des fonctions réelles :
 - ▶ savoir trouver la parité ou la périodicité d'une fonction ;
 - ▶ connaître les formules trigonométriques classiques et notamment savoir traiter les égalités entre \sin , \cos et \tan ;
 - ▶ connaître les encadrements entre x et sa partie entière $\lfloor x \rfloor$;
 - ▶ savoir manier et connaître les propriétés des fonctions exponentielles, logarithmes et puissances ;
 - ▶ savoir faire fonction la définition d'une limite dans des cas simples ;
 - ▶ savoir trouver s'il y a prolongement par continuité ou non ;
 - ▶ savoir montrer qu'une fonction n'a pas de limite en utilisant des suites ;
 - ▶ connaître les équivalents classiques en 0 ;
 - ▶ connaître, savoir énoncer et utiliser le théorème des valeurs intermédiaires (T.V.I) ;
 - ▶ savoir expliquer pourquoi une fonction admet une fonction réciproque en utilisant le théorème du cours ;
 - ▶ savoir tracer et connaître les propriétés principales des fonctions arcsin, arccos et arctan ;
- À propos de la dérivabilité des fonctions réelles :
 - ▶ connaître les dérivées de fonctions classiques ;
 - ▶ savoir trouver la dérivée d'une fonction composée ;
 - ▶ connaître l'énoncé et savoir appliquer le théorème de Rolle et le théorème des accroissements finis (T.A.F) ;
 - ▶ savoir manier le théorème de la limite de la dérivée ;
 - ▶ savoir montrer qu'une fonction est de classe C^1 ou non ;
 - ▶ connaître la formule de Gottfried Leibniz.
- À propos des développements limités :
 - ▶ connaître les développements limités de fonctions usuelles en 0 ;
 - ▶ savoir trouver un développement limité à partir des développements limités usuels ;
 - ▶ savoir utiliser un développement limité pour calculer une limite ou un équivalent ;
 - ▶ savoir utiliser un développement limité pour obtenir la position d'une courbe par rapport à une tangente ou une asymptote oblique ;

Résumé de cours

La première partie de ce chapitre est consacré aux rappels sur les suites. C'est une bonne révision avant le chapitre sur les séries numériques. Il faut savoir qu'une série est tout d'abord ... une suite ! La suite de ce chapitre est une révision du programme de première année TSI1 sur les fonctions usuelles, la dérivabilité et les développements limités. Ces bases sont essentielles en particulier pour le chapitre sur l'intégration sur un intervalle, les courbes paramétrées et les fonctions de deux variables. Il faut revenir de temps en temps revoir les techniques et les notions développées ici et cela tout le long de l'année, quand vous en sentirez le besoin !

■ Nombres réels

On munit \mathbf{R} de l'addition et de la multiplication. Ainsi pour tout $(a, b, c) \in \mathbf{R}^3$,

$$a + b \in \mathbf{R}, a - b \in \mathbf{R}, a + b = b + a, (a + b) + c = a + (b + c)$$

$$a \times b \in \mathbf{R}, \frac{a}{b} \in \mathbf{R} \text{ si } b \neq 0, a \times b = b \times a, a \times (b \times c) = (a \times b) \times c$$

De plus, $a(-b) = (-a)b = -ab$ et $(a + b)c = ac + bc = c(a + b)$.

Le réel 0 vérifie $a + 0 = 0 + a = a$ et $a \times 0 = 0 \times a = 0$. De même, $1 \times a = a \times 1 = a$.

On munit \mathbf{R} d'une relation \leq qui vérifie :

$$\forall (a, b) \in \mathbf{R}^2, \forall c \in \mathbf{R}, a \leq b \Rightarrow a + c \leq b + c \text{ et } \forall c \geq 0, a \leq b \Rightarrow ac \leq bc.$$

Remarque : on définit la notation $<$ par $a < b \Leftrightarrow (a \leq b \text{ et } a \neq b)$ la notation \geq par $a \geq b \Leftrightarrow b \leq a$ et la notation $>$ par $a > b \Leftrightarrow b < a$.

On munit aussi \mathbf{R} d'une **valeur absolue**, application de \mathbf{R} dans \mathbf{R}_+ qui à x associe $|x|$ tel que :

$$|x| = \begin{cases} x & , \quad x \geq 0 \\ -x & , \quad x \leq 0 \end{cases}.$$

Proposition 1.1. — On a : $\forall a \in \mathbf{R}, (|a| = 0 \Leftrightarrow a = 0).$

— On a : $\forall (a, b) \in \mathbf{R}^2, |ab| = |a||b|, |a + b| \leq |a| + |b|, ||a| - |b|| \leq |a - b|.$

Intervalles de \mathbf{R} . On rappelle que $-\infty$ et $+\infty$ désignent l'infini négatif et l'infini positif.

Ils ne sont pas considérés comme des réels mais sont comparables avec les réels.

Pour tout $x \in \mathbf{R}$, $x > -\infty$ et $x < +\infty$.

Soit $(a, b) \in \mathbf{R}^2$, on définit l'intervalle fermé $[a, b]$ par $[a, b] = \{x \in \mathbf{R}, a \leq x \leq b\}$, l'intervalle ouvert $]a, b[$ par $]a, b[= \{x \in \mathbf{R}, a < x < b\}$, l'intervalle semi-ouvert $[a, b[$ par $[a, b[= \{x \in \mathbf{R}, a \leq x < b\}$, l'intervalle semi-ouvert $]a, b]$ par $]a, b] = \{x \in \mathbf{R}, a < x \leq b\}$.

Soit, de même, $(a, b) \in \mathbf{R}^2$, on définit l'intervalle fermé $[a, +\infty[$ par $[a, +\infty[= \{x \in \mathbf{R}, a \leq x\}$, l'intervalle ouvert $]a, +\infty[$ par $]a, +\infty[= \{x \in \mathbf{R}, a < x\}$, l'intervalle fermé $]-\infty, a]$ par $]-\infty, a] = \{x \in \mathbf{R}, x \leq a\}$, l'intervalle ouvert $]-\infty, a[$ par $]-\infty, a[= \{x \in \mathbf{R}, x < a\}$.

On peut remarquer que \mathbf{R} est en fait $]-\infty, +\infty[$.

Mise en œuvre : exercice 01, exercice 02.

Définition : Un sous-ensemble A de \mathbf{R} est dit majoré ou borné supérieurement (resp. minoré ou borné inférieurement) s'il possède un majorant : $\exists M \in \mathbf{R}, \forall x \in A, x \leq M$ (resp. minorant : $\exists m \in \mathbf{R}, \forall x \in A, x \geq m$). Un ensemble à la fois majoré et minoré est dit **borné**.

Remarque : Attention, un majorant ou un minorant de A ne sont pas dans A en général.

Proposition 1.2.— A est borné si, et seulement si, $\exists K > 0, \forall x \in A, |x| \leq K$.

Définition : On appelle **partie entière** du réel x le plus grand entier inférieur ou égal à x . On note $\lfloor x \rfloor$ un tel entier dans la suite.

Proposition 1.3.—

$$\forall x \in \mathbf{R}, \lfloor x \rfloor \leq x < \lfloor x \rfloor + 1 \text{ et } x - 1 < \lfloor x \rfloor \leq x.$$

Mise en œuvre : exercice 03.

■ Suites réelles

Définition : On appelle **suite d'éléments de \mathbf{R}** (ou suite réelle) toute application d'une partie A de \mathbf{N} de la forme $\{n \in \mathbf{N}, n \geq N\}$, où N est fixé et à valeurs dans \mathbf{R} .

Remarques : 1. Bien que le programme officiel n'indique rien, on note souvent $\mathbf{R}^{\mathbf{N}}$ l'ensemble des suites réelles.

2. Si $u : \mathbf{N} \rightarrow \mathbf{R}$ est une suite, l'image d'un entier n par u se note u_n plutôt que $u(n)$. Ainsi, u_n est le terme d'indice n (appelé aussi $n^{\text{ème}}$ terme de la suite). Cette suite pourra être notée $(u_n)_{n \geq N}$ et parfois (u_n) s'il n'y a pas d'ambiguïté sur son premier terme.

Définition : On dit que la suite $u = (u_n)_{n \in \mathbf{N}}$ est constante si et seulement s'il existe $K \in \mathbf{R}$ tel que pour tout $n \in \mathbf{N}$, $u_n = K$.

Opérations sur les suites. Soient u et v deux suites quelconques que l'on suppose données à partir du même rang et $\alpha \in \mathbf{R}$, on définit :

$$u + v : n \mapsto u_n + v_n, uv : n \mapsto u_n v_n, \alpha u : n \mapsto \alpha u_n.$$

La première suite est la somme des suites u et v , la seconde est le produit des suites u et v et la troisième est le produit externe de la suite u et du réel α .

Définition : Majoration et minoration de suite.

- La suite $(u_n)_{n \in \mathbf{N}}$ est majorée par $M \in \mathbf{R}$ si et seulement si : $\forall n \in \mathbf{N}, u_n \leq M$.
- La suite $(u_n)_{n \in \mathbf{N}}$ est minorée par $m \in \mathbf{R}$ si et seulement si : $\forall n \in \mathbf{N}, u_n \geq m$.
- La suite $(u_n)_{n \in \mathbf{N}}$ est bornée si et seulement si elle est majorée et minorée.

On peut remarquer qu'une suite $(u_n)_{n \in \mathbf{N}}$ est bornée si et seulement si :

$$\exists A > 0, \forall n \in \mathbf{N}, |u_n| \leq A.$$

Proposition 1.4.— Toute combinaison linéaire ou tout produit de suites bornées est une suite bornée.

Définition : Monotonie et suites.

- ▶ u est **croissante** (resp. *strictement croissante*) si et seulement si : $\forall n \in \mathbf{N}, u_n \leq u_{n+1}$ (resp. $u_n < u_{n+1}$).
- ▶ u est **décroissante** (resp. *strictement décroissante*) si et seulement si : $\forall n \in \mathbf{N}, u_n \geq u_{n+1}$ (resp. $u_n > u_{n+1}$).
- ▶ u est une suite **monotone** (resp. *strictement monotone*) si et seulement si u est soit croissante, soit décroissante (resp. soit strictement croissante, soit strictement décroissante).

Définition : Suites extraites. On dit que la suite v est une suite extraite de la suite u si et seulement s'il existe une application g strictement croissante de \mathbf{N} dans \mathbf{N} telle que : $\forall n \in \mathbf{N}, v_n = u_{g(n)}$.

Définition : Limite finie d'une suite. On dit que la suite u est **convergente** dans \mathbf{R} si et seulement s'il existe $l \in \mathbf{R}$ tel que : $\forall \epsilon > 0, \exists n_0 \in \mathbf{N}, \forall n \geq n_0, |u_n - l| \leq \epsilon$.

Dans le cas où u est convergente, le réel l s'appelle la **limite de la suite** u et on dit aussi que u est convergente vers l . Il est équivalent d'écrire que u est convergente de limite l et d'écrire que $u - l$ est convergente de limite 0. Une suite qui ne converge pas est dite **divergente**.

Notation : si u est convergente de limite l , on écrit : $\lim_{n \rightarrow +\infty} u_n = l$.

Proposition 1.5.— ▶ La limite d'une suite, si elle existe, est unique.

- ▶ Toute suite convergente est bornée dans \mathbf{R} .
- ▶ Si u est une suite convergente vers l alors toute suite extraite de u converge et de plus nécessairement vers l .

Remarque : Attention, des suites extraites de (u_n) peuvent converger sans que (u_n) converge. Penser à $u_n = (-1)^n$. Cette suite diverge et pourtant les deux suites extraites (u_{2n}) et (u_{2n+1}) convergent.

Mise en œuvre : exercice 04, exercice 05.

Proposition 1.6.— Opérations sur les suites finies—.

Soient u et v deux suites convergentes respectivement vers l et l' , alors :

- ▶ la suite $u + v$ est convergente de limite $l + l'$;
- ▶ pour tout $\lambda \in \mathbf{R}$, la suite λu est convergente de limite λl ;
- ▶ la suite uv est convergente de limite ll' ;
- ▶ si $l' \neq 0$, la suite $\frac{u}{v}$ est convergente de limite $\frac{l}{l'}$.

Définition : Limite infinie d'une suite.

- ▶ On dit que la suite u tend vers $+\infty$ si et seulement si : $\forall A \in \mathbf{R}, \exists N \in \mathbf{N}, \forall n \geq N, u_n \geq A$.
- ▶ On dit que la suite u tend vers $-\infty$ si et seulement si : $\forall A \in \mathbf{R}, \exists N \in \mathbf{N}, \forall n \geq N, u_n \leq A$.

On remarque qu'une suite tendant vers $\pm\infty$ est un cas particulier de suite divergente. On notera $\lim_{n \rightarrow +\infty} u_n = +\infty$ toute suite (u_n) qui tend vers $+\infty$. Il existe des suites divergentes qui ne tendent pas vers $\pm\infty$, par exemple $u_n = (-1)^n$.

Opérations sur les limites infinies.

- ▶ $\lim_{n \rightarrow +\infty} u_n = l$ et $\lim_{n \rightarrow +\infty} v_n = +\infty$ alors $\lim_{n \rightarrow +\infty} (u_n + v_n) = +\infty$.
- ▶ Si $\lim_{n \rightarrow +\infty} u_n = l$ et $\lim_{n \rightarrow +\infty} v_n = -\infty$ alors $\lim_{n \rightarrow +\infty} (u_n + v_n) = -\infty$.
- ▶ Si $\lim_{n \rightarrow +\infty} u_n = +\infty$ et $\lim_{n \rightarrow +\infty} v_n = +\infty$ alors $\lim_{n \rightarrow +\infty} (u_n + v_n) = +\infty$.
- ▶ Si $\lim_{n \rightarrow +\infty} u_n = -\infty$ et $\lim_{n \rightarrow +\infty} v_n = -\infty$ alors $\lim_{n \rightarrow +\infty} (u_n + v_n) = -\infty$.

- Si $\lim_{n \rightarrow +\infty} u_n = l > 0$ et $\lim_{n \rightarrow +\infty} v_n = +\infty$ alors $\lim_{n \rightarrow +\infty} u_n v_n = +\infty$.
- Si $\lim_{n \rightarrow +\infty} u_n = l < 0$ et $\lim_{n \rightarrow +\infty} v_n = +\infty$ alors $\lim_{n \rightarrow +\infty} u_n v_n = -\infty$.
- Si $\lim_{n \rightarrow +\infty} u_n = +\infty$ et $\lim_{n \rightarrow +\infty} v_n = +\infty$ alors $\lim_{n \rightarrow +\infty} u_n v_n = +\infty$.
- Si $\lim_{n \rightarrow +\infty} u_n = l$ et $\lim_{n \rightarrow +\infty} v_n = +\infty$ alors $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0$.
- Si $\lim_{n \rightarrow +\infty} u_n = l$ et $\lim_{n \rightarrow +\infty} v_n = -\infty$ alors $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0$.

Attention, il y a des cas où on ne peut pas conclure, par exemple pour les formes indéterminées.

Définition : Relations de comparaison.

On considère deux suites $u = (u_n)$ et $v = (v_n)$ telles que v_n ne s'annule pas à partir d'un certain rang N .

- On dit que la suite u est **infinitement petite** ou **négligeable** devant la suite v quand n tend vers $+\infty$ si et seulement si $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0$. On note alors $u = o(v)$.
- On dit que la suite u est **dominée** par la suite v quand n tend vers $+\infty$ si et seulement si la suite $\left(\frac{u_n}{v_n}\right)_{n \geq N}$ est bornée. On note alors $u = O(v)$.
- On dit que la suite u est **équivalente** à la suite v quand n tend vers $+\infty$ si et seulement si $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1$. On note alors $u \sim v$.

Proposition 1.7.— On suppose que u, v, w et z sont quatre suites dont les termes ne s'annulent pas à partir d'un certain rang.

Si $u \sim v$ et $w \sim z$ alors $u.w \sim v.z$, $\frac{u}{w} \sim \frac{v}{z}$ et si $\alpha > 0$, $u^\alpha \sim v^\alpha$.

Proposition 1.8.— $u_n \sim v_n \Leftrightarrow u_n - v_n = o(v_n)$.

Proposition 1.9.— Si $(\alpha, \beta, \gamma) \in (\mathbf{R}^*)^3$, $e^{\gamma n} = o(n^\alpha)$ et $n^\alpha = o((\ln n)^\beta)$.

Proposition 1.10.— Relations de comparaisons classiques pour des suites usuelles quand n tend vers $+\infty$

$$\ln \left(1 + \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right), \quad \sin \left(\frac{1}{n}\right) = \frac{1}{n} - \frac{1}{6n^3} + o\left(\frac{1}{n^3}\right)$$

$$\cos \left(\frac{1}{n}\right) = 1 - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right), \quad \exp \left(\frac{1}{n}\right) = 1 + \frac{1}{n} + \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right),$$

$$\frac{1}{1 + \frac{1}{n}} = 1 - \frac{1}{n} + \frac{1}{n^2} + o\left(\frac{1}{n^2}\right), \quad \left(1 + \frac{1}{n}\right)^\alpha = 1 + \frac{\alpha}{n} + \frac{\alpha(\alpha-1)}{2n^2} + o\left(\frac{1}{n^2}\right), \quad \text{avec } \alpha \in \mathbf{Q}.$$

Mise en œuvre : exercice 06.

Proposition 1.11.— Comparaison à une suite dont on connaît la limite—.

- Soient (u_n) et (α_n) deux suites, la suite (α_n) étant convergente vers 0 et soit $N \in \mathbf{N}$ tel que: $\forall n \geq N, |u_n| \leq \alpha_n$ alors $\lim_{n \rightarrow +\infty} u_n = 0$.
- Soient (u_n) et (v_n) deux suites avec $\lim_{n \rightarrow +\infty} v_n = +\infty$ et $N \in \mathbf{N}$ tel que: $\forall n \geq N, v_n \leq u_n$ alors $\lim_{n \rightarrow +\infty} u_n = +\infty$.
- Soient (u_n) , (v_n) et (w_n) trois suites, on suppose que (v_n) et (w_n) sont convergentes de limite l et qu'il existe $N \in \mathbf{N}$ tel que : $\forall n \geq N, v_n \leq u_n \leq w_n$, alors la suite (u_n) est convergente de limite l .

Le dernier résultat a plusieurs noms : le théorème de l'encadrement, le théorème des Gendarmes, le théorème de l'étau ou le théorème du sandwich. À vous de choisir !

Proposition 1.12.— Théorème de la limite monotone —. Toute suite croissante majorée est convergente et toute suite décroissante minorée est convergente.

Donc toute suite monotone bornée est convergente.

On a ainsi plusieurs méthodes pour étudier la convergence d'une suite.

Mise en œuvre : exercice 07, exercice 08.

Définition : On dit que deux suites $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ sont **adjacentes** si et seulement si $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ sont monotones, de sens de variation contraire et que $\lim_{n \rightarrow +\infty} (u_n - v_n) = 0$.

Proposition 1.13.— Théorème des suites adjacentes—.

Deux suites adjacentes sont convergentes dans \mathbf{R} et, de plus, leur limite est commune.

Mise en œuvre : exercice 09, exercice 10.

Cas des suites définies par une relation de récurrence.

• Suites définies par une récurrence linéaire d'ordre 1.

Définition : Soit $r \in \mathbf{R}$. On appelle **suite arithmétique de raison r** toute suite définie par $u_0 \in \mathbf{R}$ et la relation : $\forall n \geq 0, u_{n+1} = u_n + r$.

Application : déterminer u_n en fonction de u_0 , de n et de r . En déduire la somme partielle d'une suite arithmétique $S_n = \sum_{k=0}^n u_k$.

Définition : Soit $q \in \mathbf{R}$. On appelle **suite géométrique de raison q** toute suite définie par $u_0 \in \mathbf{R}$ et la relation : $\forall n \geq 0, u_{n+1} = qu_n$.

Application : déterminer u_n en fonction de u_0 , de n et de q . On rappelle que $\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}$ pour $q \neq 1$.

En déduire la somme partielle d'une suite géométrique $S_n = \sum_{k=0}^n u_k$.

Mise en œuvre : exercice 11, exercice 12, exercice 13, exercice 14.

Définition : Soit $(a, b) \in \mathbf{R}^2$. On appelle **suite arithmético-géométrique** toute suite définie par $u_0 \in \mathbf{R}$ et la relation : $\forall n \geq 0, u_{n+1} = au_n + b$.

Application : déterminer u_n en fonction de a , de b de n et de u_0 .

• **Suites définies par une récurrence linéaire d'ordre 2**

On considère ici les suites définies par la récurrence $u_{n+2} = au_{n+1} + bu_n$, $(a, b) \in \mathbf{R}^2$.

L'approche consiste à déterminer les suites du type $(q^n)_{n \in \mathbf{N}}$ qui vérifient la récurrence. En remplaçant dans cette égalité, on obtient :

$$\forall n \in \mathbf{N}, q^{n+2} = aq^{n+1} + bq^n$$

L'égalité précédente se transforme en divisant par q^n en l'égalité (1) :

$$q^2 = qa + b$$

On considère donc l'équation du second degré (2) : $X^2 - aX - b = 0$.

On écrit d'après (1) que la suite $(q^n)_{n \in \mathbf{N}}$ vérifie la récurrence si et seulement si q vérifie (2).

i. Premier cas : $a^2 + 4b > 0$

(2) possède alors deux solutions q_1 et q_2 réelles et distinctes. Montrons que toute suite qui vérifie la récurrence est de la forme $(\alpha q_1^n + \beta q_2^n)_{n \in \mathbf{N}}$, α et β étant des réels. Il suffit de remarquer que toute suite $(u_n)_{n \in \mathbf{N}}$ qui vérifie la récurrence est parfaitement définie si on connaît ses deux premiers termes u_0 et u_1 . L'égalité, qui doit être valable pour tout $n \in \mathbf{N}$, $u_n = \alpha q_1^n + \beta q_2^n$ fournit une et une seule détermination de (α, β) par le système :

$$\begin{cases} \alpha + \beta &= u_0 \\ \alpha q_1 + \beta q_2 &= u_1 \end{cases}$$

que l'on résout sans difficulté. $\alpha = \frac{-q_2 u_0 + u_1}{q_1 - q_2}$ et $\beta = \frac{u_1 - q_1 u_0}{q_2 - q_1}$.

ii. Deuxième cas : $a^2 + 4b < 0$.

(1) a alors deux solutions q_1 et q_2 complexes conjuguées. On les écrit :

$$q_1 = r e^{i\theta}, \quad q_2 = r e^{-i\theta}$$

Cette partie sera à reprendre quand on aura fait le chapitre 3 c'est-à-dire les révisions sur les complexes.

En remarquant que $q_1^n = r^n (\cos n\theta + i \sin n\theta)$, en utilisant le même procédé que dans le premier cas, on montre que les suites réelles qui vérifient la récurrence sont de la forme :

$$\forall n \in \mathbf{N}, u_n = r^n (\alpha \cos n\theta + \beta \sin n\theta)$$

avec $(\alpha, \beta) \in \mathbf{R}^2$ couple déterminé par la donnée du couple (u_0, u_1) . On trouve :

$$\alpha = u_0 \text{ et } \beta = \frac{u_1}{r \sin \theta} - u_0 \frac{\cos \theta}{\sin \theta}$$

iii. Troisième cas : $a^2 + 4b = 0$.

L'équation (1) a une seule solution réelle et double q . On sait alors que la suite $(q^n)_{n \in \mathbf{N}}$ vérifie la récurrence. On remarque de plus que la suite $(n q^n)_{n \in \mathbf{N}}$ vérifie aussi la récurrence. On montre par le choix de u_0 et de u_1 que toute suite qui vérifie la récurrence s'écrit sous la forme $u_n = q^n (\alpha + \beta n)$, où $(\alpha, \beta) \in \mathbf{R}^2$ est parfaitement déterminé. Ici $\alpha = u_0$ et $\beta = \frac{u_1 - u_0 q}{q}$.

Mise en œuvre : exercice 15, exercice 16, exercice 17.

Remarque : On reprendra dans le chapitre sur la diagonalisation (chapitre 5) les suites définies par la récurrence $u_{n+2} = au_{n+1} + bu_n$.

■ Généralités sur les fonctions à variable réelle et à valeurs dans \mathbf{R}

On note (bien que le programme officiel n'impose rien) $\mathcal{F}(I, \mathbf{R})$ l'ensemble des fonctions définies sur $I \subset \mathbf{R}$ et à valeurs dans \mathbf{R} . Une fonction de I dans \mathbf{R} est donc un cas particulier d'application de E dans F . Ici E et F sont des parties de \mathbf{R} .

Graphes d'une fonction : on appelle graphe d'une fonction f de I dans \mathbf{R} , le sous ensemble G de \mathbf{R}^2 défini par :

$$G = \{(x, y) \in \mathbf{R}^2, x \in I, y = f(x)\}.$$

Le plan étant rapporté à un repère cartésien (O, \vec{i}, \vec{j}) , on appelle courbe représentative de la fonction f l'ensemble (Γ) des points $M(x, y)$ du plan tels que $x \in I$ et $y = f(x)$. On dit aussi que l'équation cartésienne de la courbe (Γ) est $y = f(x)$.

Structure de l'ensemble $\mathcal{F}(I, \mathbf{R})$.

Soient f et g deux éléments de $\mathcal{F}(I, \mathbf{R})$ et $a \in \mathbf{R}$. La somme de f et g est définie par la fonction $f + g$ telle que :

$$\forall x \in I, (f + g)(x) = f(x) + g(x)$$

Le produit de f et g , noté fg (au lieu de $f \times g$) est défini par :

$$\forall x \in I, (fg)(x) = f(x)g(x)$$

Le produit de $a \in \mathbf{R}$ et de f , noté af (au lieu de $a.f$) est défini par

$$\forall x \in I, (af)(x) = af(x)$$

Si $f \in \mathcal{F}(I, \mathbf{R})$, $g \in \mathcal{F}(I, \mathbf{R})$, pour tout $(\alpha, \beta) \in \mathbf{R}^2$, $\alpha f + \beta g \in \mathcal{F}(I, \mathbf{R})$.

Définition : $f \in \mathcal{F}(I, \mathbf{R})$ est **bornée** sur I si et seulement si : $\exists A \geq 0, \forall x \in I, |f(x)| \leq A$.

Exemple : $x \mapsto |x|$, $x \mapsto \lfloor x \rfloor$, $x \mapsto \ln x$, $x \mapsto \exp(x) = e^x$, $x \mapsto x^\alpha$ avec $x > 0$ et $\alpha \in \mathbb{Q}$, $x \mapsto a^x = e^{x \ln a}$ (avec $a > 0$), $x \mapsto \log_a(x) = \frac{\ln x}{\ln a}$ (avec $a > 0$ et $a \neq 1$) sont des fonctions définies sur \mathbf{R} . Rappeler les propriétés de $\ln(x)$, de $\exp(x)$, de x^α et de a^x .

Mise en œuvre : exercice 18, exercice 19, exercice 20.

Proposition 1.14.— Tout produit et toute combinaison linéaire de fonctions bornées sur I est une fonction bornée sur I .

Définition : Soit $f \in \mathcal{F}(I, \mathbf{R})$.

- (i) f est **majorée** sur I si et seulement si
 $\exists M \in \mathbf{R}, \forall x \in I, f(x) \leq M$
- (ii) f est **minorée** sur I si et seulement si
 $\exists m \in \mathbf{R}, \forall x \in I, m \leq f(x)$

Il est clair que f est bornée sur I si et seulement si f est majorée et minorée.

Composition des fonctions. Soit f une fonction définie sur I et g définie sur $f(I)$, la fonction composée $g \circ f$ est la fonction $x \mapsto g(f(x))$. On retrouve les propriétés de la composition des applications. La composition est dite **associative**, c'est-à-dire que si f, g, h sont trois fonctions à valeurs dans \mathbf{R} définies sur les domaines adéquats, $(h \circ g) \circ f = h \circ (g \circ f)$.

On rappelle que la composition n'est pas **commutative**, c'est-à-dire qu'en général, $f \circ g \neq g \circ f$.

Exemple : Calculer $f \circ g$ et $g \circ f$ pour $f : x \mapsto \sqrt{x}$ et $g : x \mapsto x^2$. Conclure.

Définition : Soit $f \in \mathcal{F}(I, \mathbf{R})$, on dit que f admet un **maximum relatif** (resp. un **minimum relatif**) en $x_0 \in I$ s'il existe $\alpha > 0$ tel que

$$\forall x \in I, (|x - x_0| < \alpha \Rightarrow f(x) \leq f(x_0))$$

$$(resp. \forall x \in I, (|x - x_0| < \alpha \Rightarrow f(x) \geq f(x_0)))$$

Définition : Soit $f \in \mathcal{F}(I, \mathbf{R})$, on dit que f admet un **maximum absolu** (resp. un **minimum absolu**) en $x_0 \in I$ si et seulement si :

$$\forall x \in I, f(x) \leq f(x_0)$$

$$(resp. \forall x \in I, f(x) \geq f(x_0))$$

On appelle **extremum** de f tout maximum ou tout minimum de f .

Définition : Soit f une fonction de I dans \mathbf{R} .

- (i) f est **croissante** sur I si et seulement si
 $\forall (x, x') \in I^2, (x \leq x' \Rightarrow f(x) \leq f(x'))$
- (ii) f est **décroissante** sur I si et seulement si
 $\forall (x, x') \in I^2, (x \leq x' \Rightarrow f(x) \geq f(x'))$
- (iii) f est **strictement croissante** sur I si et seulement si
 $\forall (x, x') \in I^2, (x < x' \Rightarrow f(x) < f(x'))$
- (iv) f est **strictement décroissante** sur I si et seulement si
 $\forall (x, x') \in I^2, (x < x' \Rightarrow f(x) > f(x'))$

Une application de $\mathcal{F}(I, \mathbf{R})$ est **monotone** (resp. **strictement monotone**) sur I si et seulement si elle est soit croissante sur I , soit décroissante sur I (resp. soit strictement croissante sur I , soit strictement décroissante sur I).

Exemple : Montrer que $x \mapsto 3x^2 + 2021$ est strictement croissante sur $I = [0, +\infty[$.

Définition : Une fonction f est **paire** sur I si et seulement :

$$x \in I \Leftrightarrow -x \in I \text{ et } \forall x \in I, f(x) = f(-x).$$

Interprétation géométrique. En supposant être dans le plan rapporté à (O, \vec{i}, \vec{j}) , la courbe (Γ) , représentative de f , admet Oy comme axe de symétrie et pour étudier f , il suffit de l'envisager sur $E = I \cap [0, +\infty[$.

Définition : Une fonction f est **impaire** sur I si et seulement :

$$x \in I \Leftrightarrow -x \in I \text{ et } \forall x \in I, f(x) = -f(-x).$$

Interprétation géométrique. Ici la courbe (Γ) admet l'origine O comme centre de symétrie et pour étudier f , il suffit de l'envisager encore sur $E = I \cap [0, +\infty[$.

Définition : Une fonction f est **périodique** sur I s'il existe un nombre T strictement positif tel que l'on ait (1) : $x \in I \Leftrightarrow x + T \in I$ et $\forall x \in I, f(x + T) = f(x)$.

Le plus petit réel T qui vérifie (1) est appelé **période** de la fonction f . On dit aussi que f est **T -périodique**.

Interprétation géométrique. Pour étudier une fonction f de période T , il suffit de l'envisager sur $E = I \cap [\alpha, \alpha + T]$ avec $\alpha \in \mathbf{R}$. Si (\mathcal{C}) est la courbe représentative de la restriction de f à E , la courbe (Γ) représentative de f s'obtient en complétant (\mathcal{C}) par les arcs de courbes qui s'en déduisent par les translations de vecteur $k\vec{V}$ avec $k \in \mathbf{Z}^*$ et $\vec{V}(T, 0)$.

Application : tracer \sin , \cos et \tan sur une période.

■ Formulaire de trigonométrie

Exercice Commencer par faire un tableau donnant les valeurs de $\sin \theta$, de $\cos \theta$ et $\tan \theta$ pour $\theta \in \{0, \pi/6, \pi/4, \pi/3, \pi/2, 2\pi/3, 3\pi/4, \pi, 3\pi/2, 2\pi\}$.

Proposition 1.15.— Formules d'égalités —.

$$\cos u = \cos v \Leftrightarrow \begin{cases} u = v + 2k\pi \\ \text{ou} \\ u = -v + 2k\pi \\ k \in \mathbf{Z} \end{cases}, \sin u = \sin v \Leftrightarrow \begin{cases} u = v + 2k\pi \\ \text{ou} \\ u = \pi - v + 2k\pi \\ k \in \mathbf{Z} \end{cases}, \tan u = \tan v \Leftrightarrow \begin{cases} u = v + k\pi \\ k \in \mathbf{Z} \end{cases}.$$

Dans le cas où $\cos u = \sin v$, on peut utiliser $\sin v = \cos\left(\frac{\pi}{2} - v\right)$ ou $\cos v = \sin\left(\frac{\pi}{2} - v\right)$.

Proposition 1.16.— Formules d'addition et de duplication —.

$$\begin{array}{ll} \blacksquare \cos(a+b) = \cos a \cos b - \sin a \sin b & \blacksquare \cos 2a = \cos^2 a - \sin^2 a \\ \blacksquare \sin(a+b) = \sin a \cos b + \cos a \sin b & \blacksquare \sin 2a = 2 \sin a \cos a \\ \blacksquare \tan(a+b) = \frac{\tan a + \tan b}{1 - \tan a \tan b} & \blacksquare \tan 2a = \frac{2 \tan a}{1 - \tan^2 a} \end{array}$$

Proposition 1.17.— Produits en somme (linéarisation) —.

$$\begin{array}{ll} \blacksquare \cos a \cos b = \frac{1}{2} [\cos(a+b) + \cos(a-b)] \\ \blacksquare \sin a \sin b = \frac{1}{2} [\cos(a-b) - \cos(a+b)] \\ \blacksquare \sin a \cos b = \frac{1}{2} [\sin(a+b) + \sin(a-b)] \end{array}$$

En particulier, lorsque $a = b$: $\cos^2 a = \frac{1}{2} [1 + \cos 2a]$, $\sin^2 a = \frac{1}{2} [1 - \cos 2a]$.

Proposition 1.18.— Transformations de sommes en produits

$$\begin{array}{ll} \blacksquare \cos p + \cos q = 2 \cos \frac{p+q}{2} \cos \frac{p-q}{2} & \blacksquare \sin p + \sin q = 2 \cos \frac{p+q}{2} \sin \frac{p-q}{2} \\ \blacksquare \cos p - \cos q = -2 \sin \frac{p+q}{2} \sin \frac{p-q}{2} & \blacksquare \sin p - \sin q = 2 \sin \frac{p+q}{2} \cos \frac{p-q}{2} \end{array}$$

Mise en œuvre : exercice 21, exercice 22, exercice 23.

Remarque : Pour tout $(a, b) \in \mathbf{R}^2$, avec $|a| + |b| \neq 0$, $a \cos t + b \sin t$ s'écrit :

$$\sqrt{a^2 + b^2} \left(\frac{a}{\sqrt{a^2 + b^2}} \cos t + \frac{b}{\sqrt{a^2 + b^2}} \sin t \right),$$

et si on pose $\cos \phi = \frac{a}{\sqrt{a^2 + b^2}}$, $\sin \phi = \frac{b}{\sqrt{a^2 + b^2}}$, ce qui est toujours possible, on a :

$$a \cos t + b \sin t = A \cos(t - \phi), \text{ avec } A = \sqrt{a^2 + b^2}.$$

A est appelé **l'amplitude** et ϕ est la **phase**.

Mise en œuvre : exercice 24.

■ Étude locale et limite d'une fonction en un point

Vocabulaire : une propriété portant sur une fonction f définie sur I est vraie au voisinage de a signifie que cette propriété est vraie pour tout élément situé sur l'intersection de I avec un intervalle ouvert centré en a si $a \in \mathbf{R}$ ou avec un intervalle de la forme $[A, +\infty[$ si $a = +\infty$ ou enfin avec un intervalle de la forme $]-\infty, A]$ si $a = -\infty$.

Définition : Étant donnés des réels a et l , on dit que f admet l pour limite en a si pour tout réel $\epsilon > 0$, il existe $\delta > 0$ tel que l'on ait l'implication :

$$\forall x \in I, (|x - a| \leq \delta \Rightarrow |f(x) - l| \leq \epsilon).$$

Le nombre l est alors unique et on le note $\lim_{x \rightarrow a} f(x)$.

Mise en œuvre : exercice 25.

Si f admet l pour limite en a , on dit que f tend vers l quand x tend vers a .

Limite à gauche, limite à droite. Elles sont définies en a par restriction de f à $I \cap]-\infty, a[$ pour la limite à gauche en a et par restriction de f à $I \cap]a, +\infty[$ pour la limite à droite en a . Par conséquent, dans le cas où f admet l pour limite en a à gauche, on peut écrire :

$$\forall \epsilon > 0, \exists \delta > 0, \forall x \in I, (0 < a - x \leq \delta \Rightarrow |f(x) - l| \leq \epsilon).$$

et dans le cas où f admet l pour limite en a à droite, on peut écrire :

$$\forall \epsilon > 0, \exists \delta > 0, \forall x \in I, (0 < x - a \leq \delta \Rightarrow |f(x) - l| \leq \epsilon).$$

On note alors respectivement $\lim_{x \rightarrow a^-} f(x) = l$ et $\lim_{x \rightarrow a^+} f(x) = l$.

Définition : limite finie d'une fonction en l'infini. Étant donné un réel l , on dit que f admet l pour limite quand x tend vers $+\infty$ (respectivement $-\infty$) si pour tout $\epsilon > 0$, il existe $A > 0$ tel que :

$$\forall x \in I, (x > A \text{ resp. } x \leq -A) \Rightarrow |f(x) - l| \leq \epsilon.$$

On note $\lim_{x \rightarrow +\infty} f(x) = l$ ou $\lim_{x \rightarrow -\infty} f(x) = l$.

Définition : limite infinie d'une fonction en un point fini. Étant donné un réel a , on dit que f tend vers $+\infty$ (respectivement $-\infty$) quand x tend vers a si pour tout $A > 0$, il existe $\delta > 0$ tel que : $\forall x \in I, (|x - a| < \delta \Rightarrow f(x) > A \text{ (resp } f(x) < -A))$.

Définition : limite infinie d'une fonction en l'infini.

- On dit que f tend vers $+\infty$ (resp. vers $-\infty$) quand x tend vers $+\infty$ si pour tout $A > 0$, il existe $B > 0$ tel que : $\forall x \in I, (x > B \Rightarrow f(x) > A \text{ (resp. } f(x) < -A))$.
 - On dit que f tend vers $+\infty$ (resp. vers $-\infty$) quand x tend vers $-\infty$ si pour tout $A > 0$, il existe $B > 0$ tel que : $\forall x \in I, (x < -B \Rightarrow f(x) > A \text{ (resp. } f(x) < -A))$.
- On note $\lim_{x \rightarrow +\infty} f(x) = +\infty$ dans le premier cas et $\lim_{x \rightarrow +\infty} f(x) = -\infty$ dans le second.

Proposition 1.19.— Soient $a \in \mathbf{R}$ et $l \in \mathbf{R}$, on a les équivalences :

$$\lim_{x \rightarrow a} f(x) = l \Leftrightarrow \lim_{x \rightarrow a} (f(x) - l) = 0 \Leftrightarrow \lim_{h \rightarrow 0} (f(a + h)) = l.$$

Dans la définition de la limite finie, on a parlé de « l'unique réel l tel que » et nous allons justifier maintenant cette unicité.

Proposition 1.20.— **Unicité de la limite** —. Si f admet une limite finie dans un voisinage de a , cette limite est unique.

Proposition 1.21.— Toute fonction admettant une limite finie en un point est bornée dans un voisinage de ce point.

Proposition 1.22.— **Théorème d'encadrement** —.

- Soit f et g deux fonctions admettant deux limites finies l et l' en un point, si $f \leq g$ au voisinage de ce point, on a l'inégalité $l \leq l'$.
- Soient f, g et h trois fonctions telles que $f \leq g \leq h$ au voisinage d'un point en lequel les limites de f et de h sont égales à l . La limite en ce point de g existe et vaut l .

Mise en œuvre : exercice 26.

Proposition 1.23.— **Limite d'une fonction composée** —. Soient f une fonction définie sur I et g une fonction définie sur J tel que $f(I) \subset J$, si f tend vers l quand x tend vers a (fini ou non) et si g tend vers m quand x tend vers l , alors $g \circ f$ tend vers m quand x tend vers a .

► Toute combinaison linéaire de fonctions de $\mathcal{F}(I, \mathbf{R})$ admettant une limite finie en a , est une fonction de $\mathcal{F}(I, \mathbf{R})$ admettant une limite finie en a . Si f et g sont définies sur I et tendant respectivement vers l et m quand x tend vers a alors pour tout $(\alpha, \beta) \in \mathbf{R}^2, \alpha f + \beta g$ tend vers $\alpha l + \beta m$ quand x tend vers a .

► Tout produit de fonctions de $\mathcal{F}(I, \mathbf{R})$ admettant une limite finie en a , est une fonction de $\mathcal{F}(I, \mathbf{R})$ admettant une limite finie en a . Si f et g sont définies sur I et tendant respectivement vers l et m quand x tend vers a alors fg tend vers lm quand x tend vers a .

► Soient f et g deux fonctions définies sur I tendant respectivement vers l et $m \neq 0$ quand x tend vers a . La fonction $\frac{f}{g}$ tend vers $\frac{l}{m}$ quand x tend vers a .

Cas des limites infinies. Les quantités a et l étant finis ou non, nous avons les implications :

$$\lim_{x \rightarrow a} f(x) = +\infty \text{ et } \lim_{x \rightarrow a} g(x) = l \Rightarrow \lim_{x \rightarrow a} (f(x) + g(x)) = +\infty.$$

$$\lim_{x \rightarrow a} f(x) = +\infty \text{ et } \lim_{x \rightarrow a} g(x) = l > 0 \Rightarrow \lim_{x \rightarrow a} f(x)g(x) = +\infty.$$

(Si l est infini, $l > 0$ correspond à $l = +\infty$.)

$$\lim_{x \rightarrow a} f(x) = +\infty \text{ et } \lim_{x \rightarrow a} g(x) = l \Rightarrow \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = +\infty.$$

■ Continuité en un point

Définition : On dit que l'application f est **continue** en a si et seulement si f est définie en a et si $\lim_{x \rightarrow a} f(x) = f(a)$.

Définition : Soit $a \in I$ et f définie sur I .

- L'application f est continue à droite au point a si et seulement si f admet $f(a)$ pour limite à droite au point a .
- L'application f est continue à gauche au point a si et seulement si f admet $f(a)$ pour limite à gauche au point a .

f est dite **discontinue** en a si f n'est pas continue en a .

Proposition 1.24.— f est continue en a si et seulement si pour tout $(u_n) \in \mathbf{R}^{\mathbf{N}}$ convergente vers a alors $(f(u_n))$ converge vers $f(a)$.

On utilise souvent la contraposée de ce résultat pour montrer la non continuité en un point.

Prolongement par continuité

Soit $a \in I$ et f définie sur $I \setminus \{a\}$, admettant une limite finie l en a , on appelle prolongement de f par continuité au point a l'application g définie sur I par : $g : I \rightarrow \mathbf{R}, x \mapsto \begin{cases} f(x) & , \quad x \in I \setminus \{a\} \\ l & , \quad x = a \end{cases}$.

Mise en œuvre : exercice 27.

On peut transporter les opérations sur les limites finies en un point finie dans le cadre de la continuité en un point.

- Si f et g sont continues en a alors pour tout $(\alpha, \beta) \in \mathbf{R}^2$, $\alpha f + \beta g$, fg et $\frac{f}{g}$ (si g ne s'annule pas dans un voisinage de a) sont continues en a .

Application : cas des suites $u_{n+1} = f(u_n)$, où f est continue et croissante sur I stable par f

On remarque déjà que si (u_n) est une suite convergente de limite l , la suite $(f(u_n))$ est convergente de limite $f(l)$ et par conséquent si la suite (u_n) vérifie la récurrence $u_{n+1} = f(u_n)$, nécessairement $l = f(l)$. On admet donc et on utilisera ce résultat.

Remarque. On a l'implication : $\{x \in \mathbf{R}, f(x) = x\}$ est vide $\Rightarrow (u_n)$ est divergente dans \mathbf{R} .

Supposons maintenant f définie sur I intervalle réel stable c'est-à-dire que $f(I) \subset I$.

De plus, on suppose $u_0 \in I$ et $\forall n \geq 0, u_{n+1} = f(u_n)$. On énonce :

Théorème : Si f est croissante sur I alors (u_n) est croissante si $u_1 > u_0$, (u_n) est décroissante si $u_1 < u_0$, (u_n) est constante si $u_0 = u_1$.

Bilan : dans la pratique, on étudie f en décomposant par intervalles stables. On considère I stable sur lequel f est croissante puis on détermine un minorant ou un majorant de (u_n) qui est en général une solution de $f(x) = x$.

Exemple Étudier la convergence de (u_n) définie par : $\forall n \in \mathbf{N}^*, u_n = \ln(u_{n-1} + 3)$ d'abord avec $u_0 = 0$ puis avec $u_0 = 5$. On pourra commencer par dessiner $x \mapsto \ln(x + 3)$.

Remarque : Que se passe-t-il si f est décroissante? D'abord il faut que vous soyez guidé par l'énoncé et l'idée est de prendre les suites définies pour tout $n \in \mathbf{N}$ par $v_n = u_{2n}$ et $w_n = u_{2n+1}$. En effet, $v_{n+1} = u_{2n+2} = (f \circ f)(v_n)$ et $f \circ f$ est croissante et on peut y appliquer le théorème précédent.

■ Domination, négligeabilité et équivalences de fonctions en un point

Définition : Domination au voisinage d'un point. Étant donné un point a (appartenant à I ou à l'extrémité de I) et une fonction φ à valeurs réelles et ne s'annulant pas sur I privé de a , on dit que f définie sur I et à valeurs réelles est **dominée** par φ dans un voisinage de a si, et seulement si, le rapport $\frac{f}{\varphi}$ est borné au voisinage de a . On note alors $f = O(\varphi)$.

Définition : Étant donné un point a (appartenant à I ou à l'extrémité de I) et une fonction φ à valeurs réelles et ne s'annulant pas sur I privé de a , on dit que f définie sur I et à valeurs réelles est **négligeable** devant φ dans un voisinage de a si et seulement si : $\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = 0$.

On note alors $f = o(\varphi)$.

Définition : Soient deux fonctions f et g définies sur I à valeurs réelles et ne s'annulant pas sur $I \setminus \{a\}$. On dit que f est **équivalente** à g au voisinage de a si, et seulement si, $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$. On note alors $f \sim g$.

Proposition 1.26.— Rappels des équivalents classiques en $x = 0$

$$\ln(1+x) \sim x, \exp(x) - 1 \sim \frac{x^2}{2}, (1+x)^\alpha \sim 1 + \alpha x, \sin x \sim x, 1 - \cos x \sim \frac{x^2}{2}, \tan x \sim x$$

Mise en œuvre : exercice 28.

■ Continuité sur un intervalle I

Définition : f est **continue** sur I si et seulement si f est continue en tout point de I (le cas échéant continue à droite ou à gauche aux extrémités de l'intervalle).

L'ensemble des applications continues sur I se note $\mathcal{C}^0(I, \mathbf{R})$.

Il résulte de ce qui précède les **opérations qui conservent la continuité sur un intervalle**.

- Toute combinaison linéaire d'applications continues sur I est une application continue sur I .
- Tout produit d'application continue sur I est une application continue sur I .
- Si f est continue sur I et ne s'annule pas sur I alors $\frac{1}{f}$ est continue sur I .
- Soient I et J deux intervalles, $f : I \rightarrow J$, $g : J \rightarrow \mathbf{R}$. Si f et g sont continues respectivement sur I et J , $g \circ f$ est continue sur I .

Proposition 1.27.— Théorème des valeurs intermédiaires —. Soit f continue sur un intervalle I , on suppose que a et b sont deux points de I tels que $f(a)f(b) \leq 0$. Alors il existe $c \in [a, b]$ tel que $f(c) = 0$.

Corollaire 1.28.— Autre forme du théorème des valeurs intermédiaires —. Soit f continue sur $[a, b]$. Toute valeur comprise entre $f(a)$ et $f(b)$ est atteinte par la fonction f en un point de $[a, b]$.

Mise en œuvre : exercice 29.

Corollaire 1.29.— L'image d'un intervalle par une application continue est un intervalle.

Corollaire 1.30.— Une fonction continue sur un segment est bornée et atteint ses bornes.

Proposition 1.31.— Fonction réciproque et continuité —. Soit $f : I \rightarrow \mathbf{R}$, définie, continue et strictement monotone sur I alors l'application f est bijective et sa fonction réciproque f^{-1} est définie, continue et strictement monotone (de même sens de monotonie) sur $f(I)$.

Warnung : ne pas confondre f^{-1} fonction réciproque de f et $\frac{1}{f}$, fonction inverse de f .

Interprétation géométrique : on peut tracer le graphe d'une fonction réciproque f^{-1} en effectuant la symétrie orthogonale par rapport à $y = x$ du graphe de f .

Application : les fonctions inverses des fonctions trigonométriques

- La restriction de \sin à $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ admet une fonction réciproque \arcsin définie, continue et strictement croissante sur $[-1, 1]$. Dessiner la portion de \sin pertinente et \arcsin sur le même dessin.
- La restriction de \cos à $[0, \pi]$ admet une fonction réciproque \arccos définie, continue et strictement décroissante sur $[-1, 1]$. Dessiner la portion de \cos pertinente et \arccos sur le même dessin.
- La restriction de \tan à $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ admet une fonction réciproque \arctan définie, continue et strictement croissante sur \mathbf{R} . Dessiner la portion de \tan pertinente et \arctan sur le même dessin.

Proposition 1.32.— Pour tout $x \in [-1, 1]$, $\sin(\arcsin x) = x$ et $\cos(\arccos x) = x$.
 Pour tout $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $\arcsin(\sin x) = x$ et pour tout $x \in [0, \pi]$, $\arccos(\cos x) = x$.
 Pour tout $x \in \mathbf{R}$, $\tan(\arctan x) = x$ et pour tout $x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, $\arctan(\tan x) = x$.

Mise en œuvre : exercice 30, exercice 31, exercice 32.

■ Dérivabilité en un point d'une fonction réelle

Dans tout ce paragraphe, f désigne une fonction définie dans un voisinage de $a \in \mathbf{R}$ et à valeurs dans \mathbf{R} .

Définition : On dit que f est **dérivable** en a lorsque le rapport $\frac{f(x) - f(a)}{x - a}$, défini pour $x \neq a$, admet une limite finie quand x tend vers a .

Lorsque cette limite existe, on l'appelle **nombre dérivé** de f en a . On le note $f'(a)$ (parfois aussi $D(f)(a)$). On remarque que : $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = f'(a)$.

Exemple : Montrer que $\frac{\sin x}{x} = 1$ en remarquant que la dérivée de $\sin x$ en tout x est $\cos x$.

Mise en œuvre : exercice 33.

Proposition 1.33.— f est dérivable en a si et seulement s'il existe une application ϵ définie dans un voisinage de a et un réel α vérifiant pour tout x dans un voisinage de a ,

$$\lim_{x \rightarrow a} \epsilon(x) = 0 \text{ et } f(x) = f(a) + (x - a)(\alpha + \epsilon(x))$$

Dans le cas où cette égalité existe, $\alpha = f'(a)$.

Ainsi, f est dérivable en a si, et seulement si, f possède un développement limité à l'ordre 1 au voisinage de a .

Corollaire 1.34.— Si f est dérivable en a alors f est continue en a .

Interprétation géométrique. Si f est dérivable en a , la tangente à la courbe représentative de f dans un voisinage de a a pour équation : $y - f(a) = f'(a)(x - a)$.

Exemple : Écrire une équation de la tangente à $x \mapsto e^x$ en $x = 0$.

f est **dérivable à droite** (resp. **à gauche**) en a si et seulement si l'application $x \mapsto \frac{f(x) - f(a)}{x - a}$ définie dans un voisinage de a privé de $\{a\}$ admet une limite finie **à droite** (resp. **à gauche**) en a .

Interprétation géométrique. Si f est dérivable à droite (resp. à gauche) en a , la courbe représentative de f possède en a une demi-tangente à droite (resp. à gauche).

On a un résultat similaire à la continuité :

Proposition 1.35.— Soit f définie dans un voisinage de a alors f est dérivable en a si et seulement si f est simultanément dérivable à droite et à gauche en a et si l'on a $f'_d(a) = f'_g(a)$.

Opérations et dérivabilité en un point.

Soit f et g deux fonctions dérivables en a , on a alors :

► $f + g$ est dérivable en a et $(f + g)'(a) = f'(a) + g'(a)$.

► fg est dérivable en a et $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

► Pour tout $\lambda \in \mathbf{R}$, λf est dérivable en a et $(\lambda f)'(a) = \lambda f'(a)$.

► Si $f(a) \neq 0$, alors $\frac{1}{f}$ est dérivable en a et $\left(\frac{1}{f}\right)'(a) = -\frac{f'(a)}{(f(a))^2}$.

► Si $g(a) \neq 0$, alors $\frac{f}{g}$ est dérivable en a et $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{(g(a))^2}$.

► Si f est une fonction dérivable en a et g une fonction dérivable en $f(a)$ alors $g \circ f$ est dérivable en a et $(g \circ f)'(a) = f'(a)g'(f(a))$.

Proposition 1.36.— Si u est une fonction dérivable en tout point x ,

$$(e^{u(x)})' = u'(x)e^{u(x)}, (\cos(u(x)))' = -u'(x)\sin(u(x)), (\sin(u(x)))' = u'(x)\cos(u(x))$$

$$(\ln(|u(x)|))' = \frac{u'(x)}{u(x)}, (u^\alpha(x))' = \alpha u'(x)u^{\alpha-1}(x), (\tan(u(x)))' = u'(x)(1 + \tan^2(u(x))).$$

■ Dérivée sur un intervalle

Dans toute cette partie, I désigne un intervalle de \mathbf{R} et les fonctions sont à valeurs dans \mathbf{R} .

Définition : On dit que f est **dérivable** sur I si, et seulement si, f est dérivable en tout point de I . Si f est dérivable sur I alors la fonction $I \rightarrow \mathbf{R}, x \mapsto f'(x)$, où $f'(x)$ est le nombre dérivé de f en x , est appelée **fonction dérivée** de f sur I . Elle est notée f' (ou encore $\frac{df}{dx}$).

Remarque : Attention dans le cas où I est un intervalle au moins fermé d'un côté. Prenons le cas où, par exemple $I = [a, +\infty[$, alors pour que f soit dérivable sur I , il faut en particulier que f soit dérivable en a ce qui impliquera de démontrer seulement la dérivabilité à droite en a car un voisinage de a sera de la forme $[a, a + \beta[$, avec $\beta > 0$.

Proposition 1.37.— Soit f et g deux fonctions dérivables sur I , alors toutes les formules de dérivation des fonctions $f + g$, fg , λf avec $\lambda \in \mathbf{R}$ et f/g dans le cas où g ne s'annule pas sur I , données plus haut pour le point $a \in \mathbf{R}$ sont valables pour tout $x \in I$.

Si f est une fonction dérivable sur I et g une fonction dérivable sur $f(I)$ alors $g \circ f$ est dérivable sur I et $(g \circ f)' = f' \times (g' \circ f)$.

Proposition 1.38.— Soit f une fonction dérivable sur I et admettant un extremum en $a \in I$. Alors : $f'(a) = 0$.

Remarque : La réciproque de la proposition précédente est fausse. La dérivée peut s'annuler en un point sans qu'il y ait obligatoirement un extremum en ce point.

Proposition 1.39.— Soit f une fonction strictement monotone de I sur $J = f(I)$ et dérivable sur I . Si f' ne s'annule pas sur I alors f^{-1} est dérivable sur J et on a l'égalité : $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$.

On peut en déduire les formules suivantes.

Proposition 1.40.— Formules classiques

- La fonction Arcsin est dérivable sur $] - 1, 1[$ et pour tout $x \in] - 1, 1[$, $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$.
- La fonction arccos est dérivable sur $] - 1, 1[$ et pour tout $x \in] - 1, 1[$, $(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$.
- La fonction arctan est dérivable sur \mathbf{R} et pour tout $x \in \mathbf{R}$, $(\arctan x)' = \frac{1}{1+x^2}$.
- Pour tout $x \in [-1, 1]$, $\arcsin x + \arccos x = \frac{\pi}{2}$.
- Pour tout $x \in \mathbf{R}^*$, $\operatorname{Arctan} x + \operatorname{Arctan} \frac{1}{x} = \epsilon \frac{\pi}{2}$ avec $\epsilon = 1$ si $x > 0$ et $\epsilon = -1$ si $x < 0$.

Remarque : Si $f'(a) = 0$ alors f^{-1} n'est pas dérivable en $f(a)$ et la courbe représentative de f^{-1} admet une demi-tangente verticale au point d'abscisse a .

Mise en œuvre : exercice 34.

Théorème 1.41.— Théorème de Rolle.— Soit f une fonction continue sur $[a, b]$, dérivable sur $]a, b[$ et telle que $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que : $f'(c) = 0$.

Sous les hypothèses du théorème de Rolle, on peut affirmer que la fonction f admet au moins une tangente horizontale dans l'intervalle $]a, b[$.

Théorème 1.42.— Théorème des accroissements finis —. Soit f une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que : $f'(c) = \frac{f(b) - f(a)}{b - a}$.

Corollaire 1.43.— Inégalité des accroissements finis.— Soit f une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$. On suppose que f' est bornée sur $]a, b[$ c'est-à-dire qu'il existe deux réels m et M tels que pour tout $x \in]a, b[$, on a $m \leq f'(x) \leq M$. Alors : $m(b - a) \leq f(b) - f(a) \leq M(b - a)$.

Il découle de l'inégalité des accroissements finis que si f est dérivable sur un intervalle I et s'il existe $M \in \mathbf{R}_+$ tel que pour tout $t \in I$, $|f'(t)| \leq M$ alors :

$$\forall (x, y) \in I^2, |f(x) - f(y)| \leq M|x - y|.$$

Mise en œuvre : exercice 35, exercice 36, exercice 37.

Corollaire 1.44.— Soit f continue sur $[a, b]$ et dérivable sur $]a, b[$. On a les équivalences :

- f est constante sur $[a, b] \Leftrightarrow f' = 0$ sur $]a, b[$.
- f est croissante sur $[a, b] \Leftrightarrow f' \geq 0$ sur $]a, b[$.
- f est décroissante sur $[a, b] \Leftrightarrow f' \leq 0$ sur $]a, b[$.

Définition : Soit f définie sur I , on dit que f est **de classe \mathcal{C}^1** sur I si f est dérivable sur I et si f' est continue sur I .

Mise en œuvre : exercice 38.

Définition : p étant un entier supérieur ou égal à 2, soit f définie sur I , on dit que f est p fois dérivable sur I si et seulement si f' est $(p-1)$ fois dérivable sur I et donc si la dérivée $(p-1)^{\text{ème}}$ de f est dérivable sur I . On note $f^{(p)}$ cette dérivée $p^{\text{ème}}$ de f .

On a : $f^{(p)} = [f^{(p-1)}]'$ et par convention $f^{(0)} = f$.

On dit que f est **de classe \mathcal{C}^p** sur I lorsque f est p fois dérivable sur I et $f^{(p)}$ est continue sur I .

On dit que f est **de classe \mathcal{C}^∞** sur I lorsque pour tout $k \in \mathbf{N}^*$, f est de classe \mathcal{C}^k sur I .

Remarque : Ne pas confondre l'ensemble $\mathcal{C}^1(I, \mathbf{R})$ des fonctions dont la dérivée première est continue sur I et l'ensemble $\mathcal{C}(I, \mathbf{R})$ des fonctions seulement continues sur I .

On peut remarquer, par ailleurs que si f et g sont n fois dérivables sur I et λ est un réel alors $f+g$ et λf sont n fois dérivables sur I et : $(f+g)^{(n)} = f^{(n)} + g^{(n)}$, $(\lambda f)^{(n)} = \lambda f^{(n)}$.

De même, si f est n fois dérivable sur I et si g est n fois dérivable sur $f(I)$ alors $g \circ f$ est n fois dérivable sur I .

Pour le produit de deux fonctions, on a une formule intéressante : la formule de Leibniz.

Proposition 1.45.— Formule de Gottfried Leibniz —. Soit n un entier et soit f et g deux fonctions de classe \mathcal{C}^n sur I . Alors le produit fg est de classe \mathcal{C}^n sur I et on a l'égalité :

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

Mise en œuvre : exercice 39.

■ Développements limités d'une fonction au voisinage d'un point

Définition : Si $n \in \mathbf{N}$, $a \in I$, on dit que f , définie sur I ou éventuellement I privé de a , admet un **développement limité à l'ordre n** au voisinage de a si et seulement s'il existe un polynôme P de degré inférieur ou égal à n tel que : $f(x) - P(x-a) = o_a((x-a)^n)$ dans un voisinage de a . On note souvent $DL_n(a)$ ce développement limité, même si rien « d'officiel » n'est écrit au programme de la classe de TSI.

Plus précisément en décomposant P , f admet un développement limité à l'ordre n au voisinage de a si et seulement s'il existe $(a_0, \dots, a_n) \in \mathbf{R}^{n+1}$,

$$f(x) = \sum_{k=0}^n a_k (x-a)^k + o_a((x-a)^n).$$

Le polynôme $P(x-a)$ s'appelle **partie régulière du développement limité** de f à l'ordre n , la différence $f(x) - P(x-a)$ s'appelle **reste du développement limité** de f à l'ordre n .

Dans le cas où $a = 0$, on écrit plus simplement que f admet un développement limité à l'ordre n au voisinage de 0 si, et seulement s'il existe $(a_0, \dots, a_n) \in \mathbf{R}^{n+1}$,

$$f(x) = \sum_{k=0}^n a_k x^k + o(x^n).$$

Cette dernière forme est très importante dans la pratique. Allez voir le tableau plus loin.

Proposition 1.46.— Opérations algébriques et développements limité —. Soit f et g deux fonctions de I dans \mathbf{R} admettant chacune un développement limité d'ordre n au voisinage de $a \in I$ de parties régulières respectives P_n et Q_n , alors :

- Pour tout $(\lambda, \mu) \in \mathbf{R}$, $\lambda f + \mu g$ admet un développement limité au voisinage de a à l'ordre n de partie régulière $\lambda P_n + \mu Q_n$.
- La fonction fg admet un développement limité au voisinage de a à l'ordre n de partie régulière $P_n Q_n$, tronqué de tous les monômes de degré supérieur strictement à n .
- Dans le cas où $f(a) \neq 0$, la fonction $1/f$ admet un développement limité à l'ordre n au voisinage de a .

Proposition 1.47.— Composition et développements limités —. Ici $0 \in I$ et soit f une fonction de I dans \mathbf{R} et g une fonction de J dans \mathbf{R} telles que $f(I) \subset J$. On suppose de plus $f(0) = 0$ et que chacune de ces fonctions admette un développement limité d'ordre n au voisinage de 0 de parties régulières respectives P_n et Q_n , alors $g \circ f$ admet aussi un développement limité d'ordre n au voisinage de 0 dont la partie régulière est $P_n \circ Q_n$, tronqué de tous les monômes de degré supérieur strictement à n .

Warning ! ne pas confondre $P_n Q_n$, produit de deux polynômes et $P_n \circ Q_n$, composé de ces mêmes polynômes !

Proposition 1.48.— Intégration et développement limité —.

Soit f de I dans \mathbf{R} , continue et admettant un développement limité d'ordre n au voisinage de 0 donné par : $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$.

Alors toute primitive F de f sur I possède un développement limité d'ordre $n+1$ au voisinage de 0 donné par :

$$F(x) = F(0) + \sum_{k=0}^n \frac{a_k}{k+1} x^{k+1} + o(x^{n+1}).$$

Proposition 1.49.— Dérivation et développement limité —. Ici $n \geq 1$, soit f de I dans \mathbf{R} , dérivable sur I et admettant un développement limité d'ordre n au voisinage de 0 donné par la formule : $f(x) = \sum_{k=0}^n a_k x^k + o(x^n)$.

Alors le développement limité de sa dérivée f' d'ordre $n-1$ au voisinage de 0, s'il existe est donné par :

$$f'(x) = \sum_{k=1}^n k a_k x^{k-1} + o(x^{n-1}).$$

Proposition 1.50.— Unicité d'un développement limité —. Si f admet un développement limité à l'ordre n au voisinage de a , ce développement limité a une partie régulière unique.

Supposons par exemple que f soit paire et admette un développement limité à l'ordre n au voisinage de 0 que l'on écrit $f(x) = P(x) + o(x^n)$. On en déduit :

$$f(-x) = P(-x) + o((-1)^n x^n) = P(-x) + o(x^n).$$

L'unicité de la partie régulière implique que pour tout x dans un voisinage de 0, $P(x) = P(-x)$.

Corollaire 1.51.— Soit f une fonction admettant un développement limité au voisinage de 0, si f est paire (resp. impaire), la partie régulière de ce développement est un polynôme ne possédant que des monômes de degré pair (resp. de degré impair).

Nous avons vu le problème de l'unicité d'un développement limité en un point, abordons maintenant le problème de l'existence de ce développement limité.

Théorème-Définition 1.52.— Soit f de classe \mathcal{C}^n sur un intervalle I , alors pour tout $a \in I$,

$$\forall x \in I, f(x) = \sum_{k=0}^n \frac{1}{k!} (x-a)^k f^{(k)}(a) + o_a(x-a)^n.$$

Cette formule porte le nom de **formule de Taylor avec reste de Young**.

On peut l'écrire aussi sous la forme :

$$\forall x \in I, f(x) = \sum_{k=0}^n \frac{1}{k!} (x-a)^k f^{(k)}(a) + (x-a)^n \epsilon(x),$$

où ϵ est une fonction définie dans un voisinage de a telle que $\lim_{x \rightarrow a} \epsilon(x) = 0$.

Théorème 1.53.— Développements limités au voisinage de 0 de fonctions usuelles —.

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^k}{k!} + \dots + \frac{x^n}{n!} + o(x^n)$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots + x^k + \dots + x^n + o(x^n)$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^k \frac{x^{2k}}{(2k)!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1})$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2})$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{k+1} \frac{x^k}{k} + \dots + (-1)^{n+1} \frac{x^n}{n} + o(x^n)$$

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + (-1)^k \frac{x^{2k+1}}{2k+1} + \dots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2})$$

$$\tan(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + o(x^6), \quad \frac{1}{\tan(x)} = \frac{1}{x} - \frac{1}{3}x - \frac{1}{45}x^3 + o(x^4)$$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \dots + \frac{(-1)^{n-1}(2n-2)!}{2^{2n-1}(n-1)!n!}x^n + o(x^n)$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \dots + \frac{(-1)^n(2n)!}{2^{2n}(n!)^2}x^n + o(x^n)$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n + o(x^n), \quad \alpha \in \mathbf{R}$$

Mise en œuvre : exercice 40, exercice 41, exercice 42, exercice 43.

Méthodes

■ Autour de la formule de Taylor-Young

□ **Méthode 1.1.—** Comment établir le $DL_n(a)$ d'une fonction f en utilisant la formule de Taylor-Young

- En calculant les valeurs $f^{(k)}(a)$ pour k variant de 0 à n .
- Puis en écrivant l'égalité de Taylor-Young à l'ordre n en a .

Exemple : Déterminer le $DL_2(0)$ de $u : x \mapsto \arctan(1+x)$.

On calcule

$$u'(x) = \frac{1}{2+2x+x^2}, \quad u''(x) = -\frac{2x+2}{(2+2x+x^2)^2}$$

et donc $u(0) = \frac{\pi}{4}$, $u'(0) = \frac{1}{2}$, $u''(0) = -\frac{2}{(2)^2} = -\frac{1}{2}$.

Il reste à appliquer la formule de Taylor-Young à l'ordre 2 :

$$u(x) = u(0) + \frac{x}{1!}u'(0) + \frac{x^2}{2!}u''(0) + o(x^2) \Rightarrow \arctan(1+x) = \frac{\pi}{4} + \frac{x}{2} - \frac{x^2}{4} + o(x^2)$$

Remarque : pour exprimer un développement limité, on utilise volontiers la formule de Taylor-Young dans les cas suivants :

- ▶ l'ordre est petit (1 ou 2) ;
- ▶ la fonction n'est pas issue directement par opérations usuelles de fonctions dont les développements limités sont connus ;
- ▶ dans des exercices théoriques où f n'est pas donnée explicitement à part quelques données du genre parité par exemple.

□ **Méthode 1.2.—** Comment déterminer les dérivées successives d'une fonction f en un point a à l'aide de la formule de Taylor-Young

- On dispose du développement limité de la fonction f à l'ordre n au voisinage de a .
- On identifie pour tout k de 0 à n les quantités $f^{(k)}(a)/k!$ et a_k qui est le coefficient devant $(x-a)^k$ dans le développement limité de f .

■ Détermination et utilisation d'un développement limité

Dans toute cette partie, f et g désignent des fonctions de classe \mathcal{C}^∞ sur un voisinage de 0 et pour tout $n \in \mathbf{N}$, on note respectivement P_n et Q_n leurs parties régulières d'ordre n dans les développements limités de ces fonctions au voisinage de 0. On pose :

$$v_f = \min\{n \in \mathbf{N}, P_n \neq 0\} \text{ et } v_g = \min\{n \in \mathbf{N}, Q_n \neq 0\}.$$

Ainsi si $n \geq v_f$, $f(x) = a_{v_f}x^{v_f} + \dots + a_nx^n + o(x^n)$ et si $n \geq v_g$, $g(x) = b_{v_g}x^{v_g} + \dots + b_nx^n + o(x^n)$.

❑ Méthode 1.3.— Comment calculer le $DL_n(0)$ d'une fonction h qui peut s'écrire sous forme d'une combinaison linéaire du type $h = \alpha f + \beta g$, où $(\alpha, \beta) \in \mathbb{R}^2$

On peut successivement :

[1] déterminer les développements limités à l'ordre n au voisinage de 0 de f et de g ;
si la partie régulière de f est P_n et celle de g est Q_n alors la partie régulière du développement limité de h est $\alpha P_n + \beta Q_n$.

❑ Méthode 1.4.— Comment calculer le $DL_n(0)$ d'une fonction qui s'écrit sous la forme d'un produit de deux fonctions fg

Il suffit de :

[1] calculer le développement limité de f à l'ordre $n - v_g$;
calculer le développement limité de g à l'ordre $n - v_f$;
faire le produit des deux développements limités et garder les termes de degré au plus n .
On peut étendre à un produit supérieur la méthode.

Exemple : Calculons le $DL_8(0)$ de $x \mapsto \sin^6 x$.

On commence par un développement limité de $\sin x$, mais attention l'ordre 3 suffit. En effet, comme la partie principale commence par un x , l'élévation à la puissance 6 donne au minimum des termes en x^6 , et un développement limité de $\sin x$ à l'ordre 3 suffira donc. On écrit

$$\begin{aligned}\sin x &= x - \frac{x^3}{6} + o(x^3) \Rightarrow \\ \sin^6 x &= \left(x - \frac{x^3}{6} + o(x^3)\right)^6 = x^6 \left(1 - \frac{x^2}{6} + o(x^2)\right)^6 = x^6 - x^8 + o(x^2)\end{aligned}$$

Remarque : ceci est général, quand on doit faire un produit un peu lourd, seuls quelques termes seront le plus souvent nécessaires. Cela donne espoir ! Ici, on a eu un développement du type $(a + b)^n$, il peut arriver d'avoir à développer $(a + b + c)^n$ pour certaines valeurs de n ,

$$\begin{aligned}(a + b + c)^2 &= a^2 + b^2 + c^2 + 2ab + 2ac + 2bc \\ (a + b + c)^3 &= a^3 + b^3 + c^3 + 3a^2b + 3ab^2 + 3ac^2 + 3a^2c + 3bc^2 + 3b^2c + 6abc\end{aligned}$$

Là encore, souvent, seuls deux ou trois termes sont utiles si on ne pousse pas trop le développement limité.

❑ Méthode 1.5.— Comment déterminer le $DL_n(0)$ d'une fonction de type $1/f(x)$

[1] On détermine le développement limité de f à un ordre suffisamment élevé. On met en facteur $1/x^{v_f}$. On remarque que $1/f$ est le produit de $1/x^{v_f}$ et de $1/(1 + u)$, où u est un polynôme en x tel que $u(0) = 0$. On fait alors un développement limité de $1/(1 + u)$ au voisinage de 0, à un ordre suffisant. On fait le produit du dernier développement limité avec $1/x^{v_f}$.

Exemple : Déterminer le $DL_3(0)$ de $x \mapsto \frac{1}{\sin(x)}$.

On doit faire le $DL_5(0)$ de \sin et un développement limité de $u \mapsto \frac{1}{1+u}$. Ici $v_f = 1$. On écrit : $\frac{1}{\sin x} = \frac{1}{v_f} \times \frac{1}{1+u}$ avec ici $u = -\frac{x^2}{6} + \frac{x^4}{120}$. La quantité u a une valuation de 2 et un $DL_2(0)$ suffit pour $\frac{1}{1+u}$.

$$\frac{1}{1+u} = 1 - u + u^2 + o(u^2)$$

et donc cela donne, après calculs : $\frac{1}{\sin x} = \frac{1}{x} + \frac{x}{6} + \frac{7}{360}x^3 + o(x^3)$.

Remarque : pour déterminer le $DL_n(0)$ de $\frac{f}{g}$, on fait le produit du $DL_n(0)$ de $\frac{1}{f}$ par le $DL_n(0)$ de g en ne gardant que les termes de degré inférieur ou égal à n .

Passons au cas de **la composition de fonctions** : soit f une fonction de I dans \mathbf{R} et g une fonction de J dans \mathbf{R} telles que $f(I) \subset J$. On suppose de plus $f(0) = 0$.

❑ Méthode 1.6.— Comment déterminer le $DL_n(0)$ d'une fonction de type $f \circ g$

[1] On peut calculer le $DL_n(0)$ de g et déterminer v_g . Puis calculer la partie entière e de $n/v_g + 1$ puis le développement limité d'ordre e au voisinage de 0 de f . Enfin faire alors la composée de la partie régulière du développement limité de f que l'on a obtenu et celui de g et garder les termes de degré au plus n .

Exemple : Déterminer le $DL_7(0)$ de $x \mapsto \sin(1 - \cos x)$.

Pour cela, on écrit

$$g(x) = 1 - \cos x = \frac{x^2}{2} - \frac{x^4}{24} + \frac{x^6}{720} + o(x^7)$$

et $v_g = 2$. Ici la partie entière de $\left(\frac{n}{v_g}\right) + 1$ est 4 et $f(x) = \sin x = x - \frac{x^3}{6} + o(x^4)$. Alors :

$$\begin{aligned} (f \circ g)(x) &= \frac{x^2}{2} - \frac{x^4}{24} + \frac{x^6}{720} - \frac{1}{6} \left(\frac{x^2}{2} - \frac{x^4}{24} + \frac{x^6}{720} \right)^3 + o(x^7) \\ &= \frac{x^2}{2} - \frac{x^4}{24} + \frac{x^6}{720} - \frac{x^6}{48} + o(x^7) = \frac{x^2}{2} - \frac{x^4}{24} - \frac{7x^6}{360} + o(x^7) \end{aligned}$$

❑ **Méthode 1.7.— Comment calculer un développement limité un peu costaud**

On doit :

- se ramener en 0 si le développement limité demandé n'est pas en ce point;
- décomposer le calcul en plusieurs étapes élémentaires;
- attribuer à chacune des étapes un ordre de développement convenable;
- remarquer dans chaque étape des moyens d'avoir les développements limités voulus le plus rapidement possible, par exemple en utilisant une intégration, une dérivation ou d'autres opérations usuelles sur des développements limités génériques.

■ **Calcul de limites ou d'équivalents à l'aide des développements limités**

On reprend les notations de la partie précédente et on se place ici dans le cas où f et g s'annulent en 0 et l'idée est d'étudier le quotient f/g .

❑ **Méthode 1.8.— Comment calculer $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)}$**

On commence par déterminer v_f et v_g en effectuant un développement limité de f et de g à un ordre suffisamment grand :

- ▶ si $v_f > v_g$, la limite cherchée est nulle;
- ▶ si $v_f = v_g$, la limite cherchée est le rapport a_{v_f}/b_{v_g} ;
- ▶ si $v_f < v_g$, le quotient admet des limites infinies à gauche ou à droite de 0.

Remarque : si l'on demande $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$ en un point a où $f(a) = g(a) = 0$, on peut se ramener au cas précédent en posant $X = x - a$ et de même, si l'on demande $\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)}$ en posant $X = 1/x$. On peut étendre à des calculs de limites du type $\lim_{x \rightarrow a} (g \circ f)(x)$, où a est fini ou non. Souvent un développement limité de $f \circ g$ permet ici de lever la forme indéterminée 1^∞ .

❑ **Méthode 1.9.— Comment trouver un équivalent de f/g en 0**

- On commence par déterminer v_f et v_g en effectuant un développement limité de f et de g à un ordre suffisamment grand.
- Un équivalent de f/g en 0 est alors : $\frac{a_{v_f}}{b_{v_g}} x^{v_f - v_g}$.

Exemple : Montrer que dans un voisinage de 0, $\arctan(1+x) - \frac{\pi}{4} \sim \frac{x}{2}$.

On pose ici $f(x) = \arctan(1+x) - \frac{\pi}{4}$ et $g(x) = \frac{x}{2}$. Dans la **méthode 1.1**, on a établi

$$\arctan(1+x) = \frac{\pi}{4} + \frac{x}{2} + o(x)$$

Ainsi, $f(x) = \frac{x}{2} + o(x)$ et $g(x) = \frac{x}{2} = \frac{x}{2} + o(x)$. Donc $v_f = v_g = 1$ et de plus $a_{v_f} = b_{v_g} = \frac{1}{2}$.

Le rapport f/g tend vers 1 quand x tend vers 0 et on a le résultat.

■ Utilisation des développements limités pour l'étude des graphes de fonctions

□ **Méthode 1.10.— Comment connaître une équation d'une tangente et la position de la courbe par rapport à cette tangente en $M(x_0)$**

- On peut déterminer un développement limité de f au voisinage de x_0 à un ordre suffisamment élevé pour que la partie régulière ait au moins un monôme de degré supérieur ou égal à 2 non nul.
- On conclut alors selon la parité du degré de ce monôme.

Exemple : Soit $f : x \mapsto \sin^6 x$. Préciser la position relative de la courbe représentative et de la tangente en son point d'abscisse 0.

On rappelle que le développement limité de f est au voisinage de 0, $f(x) = x^6 - x^8 + o(x^8)$. La partie régulière du développement limité de f n'a pas de termes en x , cela signifie tout simplement que la tangente est $y = 0$. Le premier terme qui suit celui en x (ici $0 \times x$) est x^6 . Donc $k = 6$ est pair, la courbe est localement au-dessus de sa tangente.

□ **Méthode 1.11.— Comment déterminer une asymptote oblique et la position de la branche infinie par rapport à cette asymptote**

On tente de déterminer un développement limité généralisé au voisinage de l'infini de $f(x)$ à un certain ordre $k \geq 2$: $f(x) = a_0x + a_1 + \frac{a_k}{x^{k-1}} + o(\frac{1}{x^{k-1}})$ avec $a_k \neq 0$. La courbe admet alors la droite $y = a_0x + a_1$ pour asymptote et la position de la branche infinie par rapport à l'asymptote est donnée par le signe de $\frac{a_k}{x^{k-1}}$.

Chapitre 2

Révision : nombres complexes et polynômes

■ les incontournables

- savoir manipuler les écritures algébrique et exponentielle des nombres complexes
 - ▶ pour simplifier une expression complexe ;
 - ▶ pour déterminer si un nombre complexe est réel.
- utiliser les nombres complexes en trigonométrie,
 - ▶ pour linéariser une expression trigonométrique ;
 - ▶ pour établir une formule trigonométrique.
- savoir résoudre une équation polynomiale, notamment
 - ▶ déterminer les racines $n^{\text{ièmes}}$ d'un nombre complexe.
 - ▶ calculer les racines carrées d'un nombre complexe, présenté sous forme algébrique ou exponentielle ;
 - ▶ résoudre les équations polynomiales de degré 2.
- savoir résoudre quelques problèmes géométriques dans le plan en passant dans \mathbf{C} et notamment
 - ▶ savoir traduire l'alignement et l'orthogonalité au moyen d'affixes ;
 - ▶ savoir reconnaître et utiliser les transformations $z \mapsto z + b$;
 - ▶ savoir reconnaître et utiliser les transformations $z \mapsto e^{i\theta} z$;
 - ▶ savoir reconnaître et utiliser les transformations $z \mapsto kz$, où $k \in \mathbf{R}$.
- ▶ Faire une division euclidienne sur deux polynômes de degrés pas trop élevés.
- ▶ Décomposer en facteurs irréductibles dans $\mathbf{C}[X]$.
- ▶ Décomposer en facteurs irréductibles dans $\mathbf{R}[X]$ en utilisant les couples de racines conjuguées.

■ et plus si affinités

- Résoudre des équations algébriques de degré supérieur à 2, en se ramenant à une équation plus simple.
- Résoudre des équations du type $a \cos t + b \sin t = c$, en passant dans \mathbf{C} ou par transformation.

Résumé de cours

Ce chapitre est une révision du programme de première année TSI1 sur les nombres complexes, les bases essentielles de trigonométrie et les polynômes. Ces notions sont utiles et reviennent constamment en algèbre et en analyse au cours de l'année de TSI2. Je vous conseille de revenir de temps en temps revoir les techniques et les notions développées ici et cela tout le long de l'année, quand vous en sentirez le besoin ! E adesso avanti !

■ Notation algébrique des nombres complexes

Présentation de \mathbb{C}

Définition : On appelle **nombre complexe** toute quantité de la forme $a + ib$, où $(a, b) \in \mathbb{R}^2$ et où i est un nombre complexe tel que $i^2 = -1$.

a est la **partie réelle** de z et b est la **partie imaginaire** et on note $a = \Re(z)$ et $b = \Im(z)$.

Vocabulaire : Si la partie réelle de z est nulle, on dit que z est **imaginaire pur**.

Théorème 2.1.— Unicité de l'écriture d'un nombre complexe en notation algébrique —.

Pour tout couple $(z, z') \in \mathbb{C}^2$ de nombres complexes,

$$z = z' \iff \begin{cases} \Re z = \Re z' \\ \Im z = \Im z' \end{cases}$$

On note \mathbb{C}^* l'ensemble des nombres complexes non nuls.

Conjugué et module d'un nombre complexe

Définition : Le **conjugué** du nombre complexe $z = a + ib$, où $(a, b) \in \mathbb{R}^2$ est $\bar{z} = a - ib$.

Le conjugué vérifie les différentes propriétés suivantes.

Proposition 2.2.— Soit $(z, z') \in \mathbb{C}^2$ un couple de nombres complexes. Alors :

- $\overline{z + z'} = \bar{z} + \bar{z}'$;
- $\overline{\bar{z}} = z$;
- $\Re(z) = \frac{1}{2}(z + \bar{z})$;
- si $z' \neq 0$, $\overline{z/z'} = \bar{z}/\bar{z}'$;
- $\overline{z \cdot z'} = \bar{z} \bar{z}'$;
- $\Im(z) = \frac{1}{2i}(z - \bar{z})$.

Corollaire 2.3.— Caractérisation des nombres réels, imaginaires purs —. Soit $z \in \mathbb{C}$ un nombre complexe. Alors :

- z est réel $\iff \Im(z) = 0 \iff z = \bar{z}$;
- z est imaginaire pur $\iff \Re(z) = 0 \iff z = -\bar{z}$.

Définition : Le **module** du nombre complexe $z = a + ib$, où $(a, b) \in \mathbb{R}^2$ est le réel positif ou nul défini par $|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$.

Remarque : soit $z \in \mathbb{C}$, on a l'encadrement $\max\{|\Re z|, |\Im z|\} \leq |z| \leq |\Re z| + |\Im z|$.

Proposition 2.4.— Propriétés du module —. Pour tout couple (z, z') de nombres complexes,

$$\begin{array}{ll} \blacksquare |z \cdot z'| = |z| |z'| ; & \blacksquare |z/z'| = |z|/|z'| ; \\ \blacksquare |z + z'| \leq |z| + |z'| ; & \blacksquare |z - z'| \geq ||z| - |z'||. \end{array}$$

Remarque : $|z + z'| = |z| + |z'|$ si et seulement s'il existe un réel $\lambda > 0$ tel que $z' = \lambda z$.

Mise en œuvre : exercice 01, exercice 02, exercice 03.

Le **plan complexe** \mathcal{P} est le plan muni d'un repère orthonormal direct $\mathcal{R} = (O, \vec{i}, \vec{j})$. À tout nombre complexe $z = x + iy$, où $(x, y) \in \mathbf{R}^2$, on associe le point M de \mathcal{P} tel que $\overrightarrow{OM} = x\vec{i} + y\vec{j}$. On dit que M est l'**image du complexe** z et que z est l'**affixe du point** M . On peut associer aussi à z le vecteur $\vec{u} = x\vec{i} + y\vec{j}$. On dit que z est l'**affixe du vecteur** \vec{u} .

Si $A(a)$ et $B(b)$, le module $|b - a|$ correspond à la distance AB .

Ainsi, \overrightarrow{AB} a pour affixe $b - a$ et $\|\overrightarrow{AB}\| = |b - a|$.

■ Ensemble \mathbf{U} des nombres complexes de module 1

Définition : Soit $\theta \in \mathbf{R}$, on appelle **exponentielle imaginaire d'angle** θ , et on note $e^{i\theta}$ le complexe $e^{i\theta} = \cos(\theta) + i \sin(\theta)$.

Proposition 2.5.— Représentation des nombres complexes de module 1 —. Pour tout nombre complexe $z \in \mathbf{U}$, il existe $\theta \in \mathbf{R}$, unique à 2π -près, tel que $z = e^{i\theta}$.

Théorème 2.6.— Règles de calcul pour l'exponentielle imaginaire —. Soit $(\theta, \theta') \in \mathbf{R}^2$, alors :

$$\begin{array}{ll} \blacksquare e^{i0} = 1 ; & \blacksquare e^{-i\theta} = 1/e^{i\theta} = \overline{e^{i\theta}} ; \\ \blacksquare e^{i(\theta+\theta')} = e^{i\theta} \times e^{i\theta'} ; & \blacksquare e^{i(\theta-\theta')} = e^{i\theta}/e^{i\theta'}. \end{array}$$

Formules d'Euler et Moivre

Théorème 2.7.— Pour tout réel $\theta \in \mathbf{R}$ et tout entier relatif $n \in \mathbf{Z}$,

$$\begin{array}{ll} \blacksquare \text{Euler : } \cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} ; & \\ \blacksquare \text{Moivre : } (e^{i\theta})^n = e^{in\theta}, \text{ soit } (\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta). & \end{array}$$

Applications à la trigonométrie

Lemme 2.8.— Factorisation d'une somme d'exponentielles —. Soit $(\theta_1, \theta_2) \in \mathbf{R}^2$, alors

$$\blacksquare e^{i\theta_1} + e^{i\theta_2} = 2 \cos\left(\frac{\theta_1 - \theta_2}{2}\right) e^{i\frac{\theta_1 + \theta_2}{2}} \quad \blacksquare e^{i\theta_1} - e^{i\theta_2} = 2i \sin\left(\frac{\theta_1 - \theta_2}{2}\right) e^{i\frac{\theta_1 + \theta_2}{2}}.$$

On déduit de ces propriétés, les formules de trigonométrie rappelées dans le cours 01.

Mise en œuvre : exercice 04, exercice 05.

■ Notation exponentielle des nombres complexes

Proposition 2.9.— Soit $z \in \mathbf{C}^*$ un nombre complexe non nul. Il existe un couple de réels $(\rho, \theta) \in \mathbf{R}_+^* \times \mathbf{R}$ tel que $z = \rho e^{i\theta} = \rho(\cos \theta + i \sin \theta)$. Cette écriture est appelée **forme exponentielle ou trigonométrique** de z .

Définition : Si $z \in \mathbf{C}^*$ s'écrit $z = \rho e^{i\theta}$, alors nécessairement $\rho = |z|$. On appelle **un argument** de z , et on note $\text{Arg}(z)$ tout nombre réel tel que $z = |z|e^{i \text{Arg}(z)}$.

Interprétation : soit M l'image dans le plan complexe d'un complexe non nul $z = \rho e^{i\theta}$. Alors $\rho = |z|$ est la longueur du vecteur \overrightarrow{OM} et θ est une mesure modulo 2π de l'angle orienté $(\vec{i}, \overrightarrow{OM})$. Il n'y a donc pas unicité de l'écriture exponentielle.

Théorème 2.10.— **Défaut d'unicité de l'écriture en notation exponentielle** —. Pour tout couple $(z, z') \in \mathbf{C}^* \times \mathbf{C}^*$ de nombres complexes non nuls :

$$z = z' \iff \begin{cases} |z| &= |z'| \\ \text{Arg}(z) &\equiv \text{Arg}(z') [2\pi] \end{cases}$$

Notation : dans l'énoncé ci-dessus, on a noté $\theta_1 \equiv \theta_2 (2\pi)$ la relation $\exists k \in \mathbf{Z}, \theta_2 = \theta_1 + 2k\pi$.

Proposition 2.11.— **Propriétés des arguments** —. Soit $(z, z') \in \mathbf{C}^* \times \mathbf{C}^*$ et $n \in \mathbf{Z}$. Alors

- $\text{Arg}(z.z') \equiv \text{Arg}(z) + \text{Arg}(z') (2\pi)$;
- $\text{Arg}(z/z') \equiv \text{Arg}(z) - \text{Arg}(z') (2\pi)$;
- $\text{Arg}(\bar{z}) \equiv -\text{Arg}(z) (2\pi)$;
- $\text{Arg}(z^n) \equiv n \text{Arg}(z) (2\pi)$.

Mise en œuvre : exercice 06, exercice 07, exercice 08.

Fonction exponentielle complexe

Définition : Soit $z = x + iy$ en notation algébrique. On définit l'**exponentielle** de z par :

$$\exp(z) = e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y).$$

Les règles de calcul pour les fonctions exponentielles réelle et imaginaire pure, s'étendent à la fonction exponentielle complexe. On a notamment : $\forall (z, z') \in \mathbf{C}^2, e^z e^{z'} = e^{z+z'}$. De plus, $\forall (z, z') \in \mathbf{C}^2, e^z = e^{z'}$ si et seulement si $z - z' \in 2i\pi\mathbf{Z}$.

Mise en œuvre : exercice 09.

■ Racines $n^{\text{èmes}}$ d'un complexe

Définition : On appelle **racine $n^{\text{ème}}$ de l'unité** tout complexe z vérifiant $z^n = 1$. L'ensemble des racines $n^{\text{èmes}}$ de l'unité est noté \mathbf{U}_n .

Théorème 2.12.— Soit $n \in \mathbf{N}, n \geq 1$. Notons pour $k \in \mathbf{Z}, z_k = \exp\left(\frac{2ik\pi}{n}\right)$. Alors

$$\mathbf{U}_n = \{z_k; k \in \mathbf{Z}\} = \{z_0, z_1, \dots, z_{n-1}\}$$

Exemples : $\mathbf{U}_1 = \{1\}$, $\mathbf{U}_2 = \{-1, 1\}$, $\mathbf{U}_3 = \{1, j, j^2\}$, $\mathbf{U}_4 = \{1, i, -1, -i\}$, où $j = e^{i\frac{2\pi}{3}}$.

Proposition 2.13.— Racines $n^{\text{èmes}}$ d'un complexe non nul quelconque —. Pour tout nombre complexe $\omega \in \mathbf{C}^*$, il existe exactement n complexes z vérifiant $z^n = \omega$.
Si on pose $\omega = \rho e^{i\theta}$, avec $(\rho, \theta) \in \mathbf{R}_+^* \times \mathbf{R}$, il s'agit des complexes définis par :

$$\forall k \in \llbracket 0, n-1 \rrbracket, z_k(\omega) = \rho^{\frac{1}{n}} e^{i\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right)}$$

Proposition 2.14.— Si $z \in \mathbf{U}_n \setminus \{1\}$. Alors $1 + z + z^2 + \dots + z^{n-1} = 0$.

Mise en œuvre : exercice 10 à exercice 16.

■ Applications des complexes aux problèmes de géométrie

Il s'agit surtout de problèmes d'alignement, d'orthogonalité ou de rotation selon un angle donné.

Proposition 2.15.— Soient trois points $A(a)$, $B(b)$ et $C(c)$ tous distincts. On rappelle que l'abscisse de \overrightarrow{AB} est $b - a$ et la norme de \overrightarrow{AB} est $|b - a|$.

- ▶ Trois points tous distincts $A(a)$, $B(b)$ et $C(c)$ sont alignés si et seulement si $\frac{c-a}{b-a} \in \mathbf{R}$.
- ▶ Soient trois points tous distincts $A(a)$, $B(b)$ et $C(c)$. Les vecteurs \overrightarrow{AB} et \overrightarrow{AC} sont orthogonaux si et seulement si $\frac{c-a}{b-a} \in i\mathbf{R}$.
- ▶ Le triangle (ABC) est rectangle en A dans le sens direct si et seulement si les vecteurs \overrightarrow{AB} et \overrightarrow{AC} sont issus l'un de l'autre par une rotation d'angle $\pi/2$.

En termes d'abscisses, c'est :

$$\exists \lambda \in \mathbf{R}_+^*, c - a = i\lambda(b - a) \Leftrightarrow \frac{c-a}{b-a} = -\frac{\bar{c}-\bar{a}}{\bar{b}-\bar{a}}.$$

Si, de plus le triangle est isocèle en A , on a $\lambda = 1$.

- ▶ Le triangle (ABC) est équilatéral dans le sens direct si et seulement si $AB = AC$ et si les vecteurs \overrightarrow{AB} et \overrightarrow{AC} font un angle de $\pi/3$, c'est-à-dire qu'ils sont issus l'un de l'autre par une rotation d'angle $\pi/3$.

En termes d'abscisses, on a : $c - a = e^{i\pi/3}(b - a)$.

- ▶ n points $M_1(z_1), M_2(z_2), \dots, M_n(z_n)$ sont sur le même cercle de centre O si et seulement si $|z_1| = |z_2| = \dots = |z_n|$.

Mise en œuvre : exercice 17, exercice 18, exercice 19.

■ Suites complexes

De façon analogue à $\mathbf{R}^{\mathbf{N}}$ qui est, rappelons le, ensemble des suites réelles, notons $\mathbf{C}^{\mathbf{N}}$ l'ensemble des suites à valeurs complexes que l'on munit des lois internes :

$$\forall ((u_n), (v_n)) \in (\mathbf{C}^{\mathbf{N}})^2, \forall n \in \mathbf{N}, (u+v)_n = u_n + v_n \text{ et } (uv)_n = u_n v_n,$$

et de la loi externe :

$$\forall (u_n) \in \mathbf{C}^{\mathbf{N}}, \forall \alpha \in \mathbf{C}, \forall n \in \mathbf{N}, (\alpha u)_n = \alpha u_n.$$

Définition : Soit $(u_n) \in \mathbf{C}^{\mathbf{N}}$, on définit :

$$\forall n \in \mathbf{N}, (\Re u)_n = \Re(u_n) \text{ et } (\Im u)_n = \Im(u_n),$$

qui sont appelés respectivement **la suite partie réelle** de la suite complexe (u_n) , **la suite partie imaginaire** de la suite complexe (u_n) .

Théorème 2.16.— Théorème de caractérisation : (u_n) est convergente quand n tend vers $+\infty$ si et seulement si les deux suites $\Re u$ et $\Im u$ sont simultanément convergentes. Dans ce cas, on a

$$\lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} (\Re u)_n + i \lim_{n \rightarrow +\infty} (\Im u)_n.$$

On peut étudier les suites complexes qui vérifient des récurrences linéaires d'ordre 1 (suites arithmétiques, géométriques, arithmético-géométriques) et d'ordre 2, et dans ce dernier cas seuls les cas où le discriminant de l'équation caractéristique est nul ou non est à distinguer.

Mise en œuvre : exercice 20.

■ Fonctions à variable réelle et à valeurs complexes

De façon analogue à $\mathcal{F}(I, \mathbf{R})$ qui est, rappelons le, ensemble des fonctions de I partie non vide de \mathbf{R} dans \mathbf{R} , notons $\mathcal{F}(I, \mathbf{C})$ l'ensemble des applications de I partie non vide de \mathbf{R} dans \mathbf{C} et munissons le des lois internes :

$$\forall (f, g) \in (\mathcal{F}(I, \mathbf{C}))^2, \forall x \in I, (f + g)(x) = f(x) + g(x) \text{ et } (fg)(x) = f(x)g(x)$$

et de la loi externe : $\forall f \in \mathcal{F}(I, \mathbf{C}), \forall \alpha \in \mathbf{C}, (\alpha f)(x) = \alpha f(x).$

Définition : Soit $f : I \rightarrow \mathbf{C}$, on définit :

$$\begin{aligned} \Re f : I &\rightarrow \mathbf{R}, x \mapsto \Re(f(x)), \quad \Im f : I \rightarrow \mathbf{R}, x \mapsto \Im(f(x)), \\ |f| : I &\rightarrow \mathbf{R}, x \mapsto |f(x)|, \quad \bar{f} : I \rightarrow \mathbf{C}, x \mapsto \overline{f(x)} \end{aligned}$$

qui sont appelés respectivement **la fonction partie réelle** de f , **la fonction partie imaginaire** de f , **la fonction module** de f et **la fonction conjuguée** de f .

Remarques : 1) $f \in \mathcal{F}(I, \mathbf{C})$ est bornée sur I si et seulement s'il existe $A \in \mathbf{R}$ tel que

$$\forall x \in \mathbf{C}, |f(x)| \leq A.$$

2) L'ensemble des fonctions bornées est stable par produit et combinaisons linéaires.

Limite et continuité en un point

Définition : f admet une limite au point a si et seulement s'il existe l tel que pour tout $\epsilon > 0$, il existe $\alpha > 0$ tel que

$$\forall x \in D_f, (|x - a| < \alpha \Rightarrow |f(x) - l| < \epsilon)$$

Dans ce cas, on dit que l est la limite de f au point a et on écrit $\lim_{x \rightarrow a} f(x) = l$.

Mise en œuvre : exercice 21.

Théorème 2.17.— Théorème de caractérisation : f admet une limite en a si et seulement si $\Re f$ et $\Im f$ admettent une limite au point a .

Dans ce cas, on a

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} (\Re f)(x) + i \lim_{x \rightarrow a} (\Im f)(x).$$

L'ensemble des applications de $\mathcal{F}(I, \mathbf{C})$ admettant une limite en a est stable par produit et par passage aux combinaisons linéaires.

Définition : Continuité au point a

f continue en a si et seulement si f admet $f(a)$ pour limite en a .

On peut écrire de façon équivalente l'implication

$$\forall \epsilon > 0, \exists \alpha > 0, |x - a| < \alpha \Rightarrow |f(x) - f(a)| < \epsilon.$$

Théorème 2.18.— $f \in \mathcal{F}(I, \mathbf{C})$ est continue en a si et seulement si les applications $\Re(f)$ et $\Im(f)$ sont continues en a .

Définition : Continuité sur un intervalle

La fonction f est continue sur I si et seulement si f est continue en tout point de I (à droite ou à gauche aux extrémités de l'intervalle). On note $\mathcal{C}^0(I, \mathbf{C})$ l'ensemble des fonctions continues sur I à valeurs dans \mathbf{C} . Alors $\mathcal{C}^0(I, \mathbf{C})$ a les mêmes stabilités que $\mathcal{C}^0(I, \mathbf{R})$.

■ Polynômes de $\mathbf{K}[X]$

Ici $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$.

Définition : Si $(a_k)_{k \in \mathbf{N}}$ est une suite d'éléments de \mathbf{K} , nulle à partir d'un certain rang, on appelle **polynôme à une indéterminée** de degré p , l'expression :

$$P = a_0 + a_1X + a_2X^2 + \dots + a_pX^p = \sum_{k=0}^p a_kX^k,$$

où p est le plus grand des entiers k tels que $a_k \neq 0$. Les scalaires a_0, \dots, a_p sont les **coefficients** de P . Chacune des quantités a_kX^k est appelée **mônôme**.

L'entier p est le **degré** de P et se note $\deg P$.

Si $\deg P = 0$, P est appelé **polynôme constant**.

Si tous les termes a_k sont nuls, P est le **polynôme nul**, noté 0 . Par convention, $\deg 0 = -\infty$.

Si $p = \deg P$ et si $a_p = 1$, on dit que P est **unitaire**.

On note $\mathbf{K}[X]$ l'ensemble des polynômes à une indéterminée, à coefficients dans \mathbf{K} et pour tout $n \in \mathbf{N}$, $\mathbf{K}_n[X]$ est l'ensemble des polynômes de $\mathbf{K}[X]$ dont le degré est inférieur ou égal à n .

Opérations dans $\mathbf{K}[X]$.

► Posons $P = \sum_{k=0}^p a_kX^k$ et $Q = \sum_{k=0}^p b_kX^k$. On suppose pour simplifier l'écriture que P et Q ont le même nombre de monômes, quitte à ajouter autant de fois que nécessaire 0 .

Alors $P + Q = \sum_{k=0}^p (a_k + b_k)X^k$ et pour tout $\alpha \in \mathbf{K}$, $\alpha P = \sum_{k=0}^p \alpha a_k X^k$.

► Pour le produit de P et Q , on utilise la distributivité de la multiplication par rapport à l'addition. Par exemple :

$$X^k \times (a_i X^i + a_j X^j) = a_i X^{i+k} + a_j X^{j+k}.$$

On définit alors $P^0 = 1$, $P^1 = P$, puis pour tout $n \in \mathbf{N}^*$, $P^n = P^{n-1} \cdot P$.

► Pour la composée de P par Q , on écrit : $P \circ Q = \sum_{k=0}^p a_k Q^k$.

Proposition 2.19.— Soient P et Q deux polynômes de $\mathbf{K}[X]$. On démontre que :

- $\deg(P + Q) \leq \sup(\deg P, \deg Q)$ (et égalité si $\deg P \neq \deg Q$).
- Si $\lambda \in \mathbf{K}^*$, $\deg(\lambda P) = \deg P$.
- $\deg(PQ) = \deg P + \deg Q$.
- $\forall n \in \mathbf{N}$, $\deg(P^n) = n \deg P$.
- $\deg(P \circ Q) = \deg P \deg Q$.

Mise en œuvre : exercice 22, exercice 23.

Définition : Soient A et B deux polynômes de $\mathbf{K}[X]$, on dit que A est un multiple de B dans $\mathbf{K}[X]$ ou que B est un diviseur de A dans $\mathbf{K}[X]$ si et seulement s'il existe $Q \in \mathbf{K}[X]$ tel que $A = BQ$.

Proposition 2.20.— **Division euclidienne dans $\mathbf{K}[X]$** —. Étant donnés A et B deux polynômes de $\mathbf{K}[X]$ avec $B \neq 0$, il existe un couple unique de polynômes (Q, R) de $(\mathbf{K}[X])^2$ tels que

$$A = BQ + R, \deg R < \deg B.$$

Les polynômes Q et R ainsi obtenus sont respectivement appelés quotient et reste dans la division euclidienne de A par B .

Corollaire 2.21.— Soient A et B deux polynômes de $\mathbf{K}[X]$ avec $B \neq 0$. B divise A dans $\mathbf{K}[X]$ si et seulement si le reste de la division euclidienne de A par B est nul.

Définition : À tout polynôme $P = a_0 + a_1 X + \dots + a_n X^n$, on associe une application de \mathbf{K} dans \mathbf{K} définie par : $x \mapsto a_0 + a_1 x + \dots + a_n x^n$, appelée fonction polynomiale associée à P .

On confond polynôme et fonction polynomiale associée. Il en découle la définition suivante qui s'appuie non pas sur le polynôme mais sur sa fonction polynomiale associée.

Définition : Tout $x \in \mathbf{K}$ tel que $P(x) = 0$ est appelé **racine** ou **zéro** de P .

Proposition 2.22.— Soit $P \in \mathbf{K}[X]$ et $a \in \mathbf{K}$. On a les assertions :

- Le reste de la division euclidienne de P par $X - a$ est $P(a)$.
- $X - a$ divise P si et seulement si $P(a) = 0$.

Mise en œuvre : exercice 24, exercice 25.

Corollaire 2.23.— Si n est un entier et $P \in \mathbf{K}_n[X]$ admet au moins $n + 1$ racines deux à deux distinctes dans \mathbf{K} alors P est nul.

Définition : Soient $P \in \mathbf{K}[X]$ un polynôme non constant et a une racine de P , on appelle **multipllicité** de a pour P le plus grand entier $\alpha \geq 1$ tel que $(X - a)^\alpha$ divise P .
On dit aussi que a est une racine de P d'ordre α de P .

Définition : Soit $p \in \mathbf{N}^*$. On appelle **polynôme dérivé** de $P = \sum_{k=0}^p a_k X^k$ le polynôme

$$P' = \sum_{k=0}^{p-1} k a_k X^{k-1}. \text{ Si } P \text{ est un polynôme constant alors } P' \text{ est le polynôme nul.}$$

On peut remarquer que cette définition est « logique » en pensant à la dérivation de la fonction polynomiale associée.

On pose $P^{(0)} = P$. Pour tout $n \in \mathbf{N}^*$, $P^{(n)} = [P^{(n-1)}]'$.

Proposition 2.24.— Formule de Taylor pour un polynôme —.

Si P est un polynôme de $\mathbf{K}[X]$ et si $a \in \mathbf{K}$, alors $P = \sum_{n \geq 0} \frac{P^{(n)}(a)}{n!} (X - a)^n$.

Attention, comme P a un degré fini, $P^{(n)}(a)$ est nul à partir d'une certaine valeur de n et la somme n'est jamais infinie.

Il découle de la formule de Taylor pour un polynôme la caractérisation suivante :

Proposition 2.25.— Caractérisation par les valeurs des dérivées successives —. Soit $P \in \mathbf{K}[X]$, non nul, $a \in \mathbf{K}$ et $\alpha \in \mathbf{N}^*$. Alors a est racine d'ordre α de P si et seulement si :

$$P(a) = P'(a) = \dots = P^{(\alpha-1)}(a) = 0 \text{ et } P^{(\alpha)}(a) \neq 0.$$

Mise en œuvre : exercice 26.

Application : pour montrer que P est divisible par $(X - a)^\alpha$, il suffit de montrer que a est une racine d'ordre α de P .

Mise en œuvre : exercice 27.

Proposition 2.26.— Formule de Leibniz dans $\mathbf{K}[X]$ —. Si P et Q sont deux polynômes de $\mathbf{K}[X]$ alors : $\forall n \in \mathbf{N}$, $[PQ]^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n-k)} = \sum_{k=0}^n \binom{n}{k} P^{(n-k)} Q^{(k)}$.

Définition : $P \in \mathbf{K}[X]$ est **scindé** sur \mathbf{K} si et seulement s'il s'écrit comme un produit de polynômes tous de degré 1.

Dans le cas où P est un polynôme scindé sur \mathbf{K} , il existe des relations entre les coefficients et les racines. Au programme de TSI2, il faut connaître la formule donnant la somme des racines et celle donnant le produit des racines.

Proposition 2.27.— Soit $P \in \mathbf{K}[X]$ scindé avec $p = \deg P \geq 1$, alors il existe un p -uplet $(\alpha_1, \dots, \alpha_p) \in \mathbf{K}^p$ et $a_p \in \mathbf{K}^*$ tel que :

$$P = a_p \prod_{k=1}^p (X - \alpha_k) = \sum_{k=0}^p a_k X^k.$$

Les scalaires $\alpha_1, \dots, \alpha_k$ sont les racines de \mathbf{K} et, de plus, on a les formules :

$$\sum_{k=1}^p \alpha_k = -\frac{a_{p-1}}{a_p} \text{ et } \prod_{k=1}^p \alpha_k = (-1)^p \frac{a_0}{a_p}.$$

Attention, les racines $\alpha_1, \dots, \alpha_p$ ne sont pas nécessairement distinctes.

Définition : Soit $P \in \mathbf{K}[X]$ alors P est **irréductible** dans $\mathbf{K}[X]$ si et seulement si $\deg P \geq 1$ et si $P = AB$ avec A et B deux polynômes de $\mathbf{K}[X]$ alors nécessairement A ou B est un polynôme constant.

Proposition 2.28.— Si $P \in \mathbf{K}[X]$ est irréductible dans $\mathbf{K}[X]$ avec $\deg P > 1$ alors P n'a pas de racine dans \mathbf{K} .

Corollaire 2.29.— **Théorème de d'Alembert-Gauss.**—

Tout polynôme de $\mathbf{C}[X]$ de degré supérieur ou égal à 1 est scindé dans \mathbf{C} , c'est-à-dire que pour tout $P \in \mathbf{C}[X]$ non constant, il existe a complexe non nul, p un entier non nul, $(x_1, \dots, x_p) \in \mathbf{C}^p$, $(r_1, \dots, r_p) \in \mathbf{N}^p$ tels que : $P = a \prod_{k=1}^p (X - x_k)^{r_k}$.

Proposition 2.30.— Soit $P \in \mathbf{R}[X]$ un polynôme non constant de coefficient dominant a . Il existe alors un couple unique (A, B) de polynômes unitaires, où A est scindé sur \mathbf{R} ou constant et B sans racine réelle tel que : $P = aAB$.

Proposition 2.31.— Les polynômes irréductibles de $\mathbf{R}[X]$ sont :

- ▶ soit les polynômes de degré 1.
- ▶ soit les polynômes de degré 2 à discriminant strictement négatif.

Remarque : Si z est un complexe non réel, on peut utiliser la formule :

$$(X - z)(X - \bar{z}) = X^2 - 2\Re(z)X + |z|^2,$$

pour à partir de la décomposition d'un polynôme à coefficients réels dans $\mathbf{C}[X]$ trouver aussi celle en facteurs irréductibles dans $\mathbf{R}[X]$.

En effet, les polynômes $X - z$ et $X - \bar{z}$ sont irréductibles dans $\mathbf{C}[X]$ et $X^2 - 2\Re(z)X + |z|^2$ l'est dans $\mathbf{R}[X]$ car son discriminant est strictement négatif.

Mise en œuvre : exercice 28, exercice 29.

Méthodes

■ Étude d'une expression complexe

□ Méthode 2.1.— Comment montrer qu'un complexe z est réel

- ▶ On peut (s'il n'est pas nul) montrer ou écrire que son argument est un multiple de π .
- ▶ On peut aussi montrer ou écrire qu'il est égal à son conjugué.
- ▶ On peut aussi montrer que sa partie imaginaire est nulle.

Exemple : donnons deux exemples qui développent deux cheminements différents.

▶ Déterminons les valeurs de $n \in \mathbf{N}$ pour lesquelles le complexe $z_n = (1+i)^n$ soit réel. Comme z_n est sous forme d'une puissance n -ième, le mieux est de passer à la forme trigonométrique de $1+i$. On écrit $1+i = \sqrt{2}e^{i\frac{\pi}{4}}$ et donc comme z_n est évidemment non nul, $\text{Arg } z_n = n\text{Arg}(1+i) = n\frac{\pi}{4}$ doit être un multiple de π c'est-à-dire que n doit être un multiple de 4.

▶ Soit $z \in \mathbf{C} - \{-1\}$ et $Z = \frac{z-1}{z+1}$, on veut déterminer z de telle manière que Z soit réel. Pour cela, on écrit que Z est réel si et seulement si $Z = \bar{Z}$, relation qui s'écrit, de manière équivalente par $\frac{z-1}{z+1} = \frac{\bar{z}-1}{\bar{z}+1}$ c'est-à-dire : $z\bar{z} - \bar{z} + z - 1 = z\bar{z} - z + \bar{z} - 1 \Leftrightarrow \bar{z} = z$. Et on en déduit que Z est réel si et seulement si z est réel et différent de -1 .

□ Méthode 2.2.— Comment montrer ou caractériser qu'un complexe z est imaginaire pur

- ▶ On peut montrer ou écrire que son argument est de la forme $\pi/2 + k\pi$, $k \in \mathbf{Z}$.
- ▶ On peut aussi montrer ou écrire qu'il est opposé à son conjugué.
- ▶ On peut aussi montrer que sa partie réelle est nulle.

Exemple : soit $z \in \mathbf{C} - \{-1\}$ et reprenons $Z = \frac{z-1}{z+1}$, on veut déterminer maintenant z de telle manière que Z soit imaginaire pur. Pour cela, on écrit que Z est imaginaire pur si et seulement si $Z = -\bar{Z}$, relation qui s'écrit, de manière équivalente par $\frac{z-1}{z+1} = -\frac{\bar{z}-1}{\bar{z}+1}$. C'est-à-dire : $z\bar{z} - \bar{z} + z - 1 = -z\bar{z} + z - \bar{z} + 1 \Leftrightarrow z\bar{z} = 1$.

Et on en déduit que Z est imaginaire pur si et seulement si z est élément de \mathbf{U} qui doit être bien entendu différent de -1 .

□ Méthode 2.3.— Comment simplifier un complexe z écrit sous forme d'une puissance de complexes, du type Z^n , où $n \in \mathbf{N}$ et Z non nul

Une méthode est d'écrire Z sous forme trigonométrique $Z = \rho e^{i\theta}$ et dans ce cas, on écrit, de façon immédiate

$$z = \rho^n e^{in\theta}$$

Exemple : on peut repartir de l'exemple précédent de la **méthode 2.1** en écrivant immédiatement

$$z = (1 + i)^n = (\sqrt{2})^n e^{in\pi/4}.$$

Remarquons, au passage, que l'idée qui viendrait à certains d'utiliser la **formule du binôme de Newton** pour développer $(1+i)^n$, dans l'espoir de simplifier cette expression, est à sortir rapidement de leur esprit. Ici ce n'est absolument pas indiqué voire contre-indiqué. Par contre la **formule du binôme de Newton** peut aider à calculer certaines sommes.

Ne résistons pas au plaisir de le faire, vous aurez ainsi une méthode gratuite en plus ! Par exemple, comme :

$$(1 + i)^n = (\sqrt{2})^n e^{in\pi/4} = \sum_{k=0}^n \binom{n}{k} i^k,$$

en prenant séparément la partie réelle et la partie imaginaire, on a :

$$(\sqrt{2})^n \cos\left(n\frac{\pi}{4}\right) = \sum_{0 \leq 2k \leq n} (-1)^k \binom{n}{2k}, \quad (\sqrt{2})^n \sin\left(n\frac{\pi}{4}\right) = \sum_{0 \leq 2k+1 \leq n} (-1)^k \binom{n}{2k+1}.$$

❏ Méthode 2.4.— Comment simplifier dans certains cas une expression complexe z écrite sous forme d'une somme de deux ou trois termes

- Si z est une somme ou une différence de complexes conjugués, on remarque alors que

$$z = Z + \overline{Z} = 2\Re(Z) \text{ ou } z = Z - \overline{Z} = 2i\Im(Z).$$

- Si z est une somme de complexes de module 1, on écrit alors $((\alpha, \beta) \in \mathbf{R}^2)$,

$$z = e^{i\alpha} + e^{i\beta} = e^{i\frac{\alpha+\beta}{2}} \left(e^{i\frac{\alpha-\beta}{2}} + e^{i\frac{-\alpha+\beta}{2}} \right) = 2e^{i\frac{\alpha+\beta}{2}} \cos\left(\frac{\alpha-\beta}{2}\right)$$

Exemple : si θ est fixé dans $[0, \pi]$, on considère $z = 1 + \cos \theta + i \sin \theta$ et on écrit successivement :

$$z = 1 + e^{i\theta} = e^{i\frac{\theta}{2}} \left(e^{i\frac{-\theta}{2}} + e^{i\frac{\theta}{2}} \right) = 2e^{i\frac{\theta}{2}} \cos\left(\frac{\theta}{2}\right)$$

On remarque, en passant, que comme $\theta/2 \in [0, \pi/2]$, la forme obtenue de z est la forme trigonométrique (si z est non nul!).

❏ Méthode 2.5.— Comment simplifier une expression complexe z écrite sous forme d'un quotient

- On peut par exemple écrire sous forme trigonométrique le numérateur et le dénominateur de z et utiliser les règles sur le module et l'argument d'un quotient.
- On peut aussi multiplier à la fois le numérateur et le dénominateur par la quantité conjuguée du dénominateur.
- On peut combiner les deux méthodes précédentes.

Exemple : simplifions le complexe $z = \frac{1 + \cos a + i \sin a}{\sqrt{1 + \sin 2a} + i\sqrt{1 - \sin 2a}}$ avec $0 < a < \pi/4$. Notons $num(z)$ et $denom(z)$ le numérateur et le dénominateur du complexe z . On a, en utilisant le développement précédent, $num(z) = 1 + \cos a + i \sin a = 2e^{i\frac{a}{2}} \cos\left(\frac{a}{2}\right)$. Pour le dénominateur, on va utiliser une méthode qui reviendra souvent. On vous conseille de la retenir. Dès qu'il y a une expression radicale, ayez à l'esprit que si l'on a un carré parfait sous ce radical, le monde dans lequel on évolue devient alors le meilleur des mondes (sans être parfait !). Ainsi comme $1 = \cos^2 a + \sin^2 a$ et comme $\sin 2a = 2 \sin a \cos a$, (inutile d'insister sur la connaissance parfaite des formules trigonométriques ?),

$$\begin{aligned} denom(z) &= \sqrt{1 + \sin 2a} + i\sqrt{1 - \sin 2a} \\ &= \sqrt{\cos^2 a + \sin^2 a + 2 \sin a \cos a} + i\sqrt{\cos^2 a + \sin^2 a - 2 \sin a \cos a} \\ &= \sqrt{(\cos a + \sin a)^2} + i\sqrt{(\cos a - \sin a)^2} = |\cos a + \sin a| + i|\cos a - \sin a| \\ &= \cos a + \sin a + i(\cos a - \sin a) = \sqrt{2} \left(\cos\left(\frac{\pi}{4} - a\right) + i \sin\left(\frac{\pi}{4} - a\right) \right) = \sqrt{2} e^{i\frac{\pi}{4} - a} \end{aligned}$$

Au passage, bien retenir cette méthode de simplification d'une somme ou d'une différence d'un cosinus et sinus du même angle car cela reviendra encore (en particulier pour certains calculs de primitives). Ainsi, $z = \frac{2e^{i\frac{a}{2}} \cos\left(\frac{a}{2}\right)}{\sqrt{2}e^{i\frac{\pi}{4} - a}}$ et cela donne $z = \sqrt{2} \cos\left(\frac{a}{2}\right) e^{i\left(\frac{3a}{2} - \frac{\pi}{4}\right)}$.

On remarque encore une fois que nous avons là l'écriture trigonométrique de z .

■ Utilisation des nombres complexes pour simplifier des expressions réelles

□ Méthode 2.6.— Comment linéariser $(\cos^p x)(\sin^q x)$, où $(p, q) \in \mathbf{N}^2$ et $x \in \mathbf{R}$

Ceci revient à écrire cette expression E , comme combinaison linéaire de termes de la forme $\cos kx$ et $\sin lx$, avec $(k, l) \in \mathbf{N}^2$.

- On remplace $\cos x$ par $\frac{e^{ix} + e^{-ix}}{2}$ et $\sin x$ par $\frac{e^{ix} - e^{-ix}}{2i}$.
- Puis on développe $\left(\frac{e^{ix} + e^{-ix}}{2}\right)^p$ et $\left(\frac{e^{ix} - e^{-ix}}{2i}\right)^q$.
- Puis l'expression E étant réelle, son développement est formé d'une combinaison linéaire de termes de la forme $e^{ikx} + e^{-ikx} = 2 \cos kx$ ou $i(e^{ilx} - e^{-ilx}) = 2 \sin lx$, avec $(k, l) \in \mathbf{N}^2$.

Exemple : linéarisons l'expression $\cos^3 x$. On écrit $\cos^3 x = \left(\frac{e^{ix} + e^{-ix}}{2}\right)^3$. Développons à l'aide de l'identité remarquable $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.

Ceux qui ne connaissent pas cette forme doivent l'apprendre maintenant. Elle est issue de la formule du binôme de Newton pour ceux qui l'ont vue ou alors développer $(a + b)^2(a + b)$. Alors $\cos^3 x$

$$= \frac{1}{8} (e^{i3x} + e^{-i3x} + 3e^{ix} + 3e^{-ix}) = \frac{1}{8} (2 \cos 3x + 6 \cos x) = \frac{1}{4} (\cos 3x + 3 \cos x).$$

Donnons aussi $\sin^3 x = \frac{1}{4} (-\sin 3x + 3 \sin x)$ (à retrouver par vous-même!).

❏ **Méthode 2.7.—** Comment développer $\cos(px)$ ou $\sin(px)$, où p est un entier naturel et $x \in \mathbf{R}$ en une combinaison linéaire de termes de la forme $(\cos^k x)(\sin^l x)$, où $(k, l) \in \mathbf{N}^2$

- On remarque que l'on a les égalités $\cos(px) = \Re((\cos x + i \sin x)^p)$ et $\sin(px) = \Im((\cos x + i \sin x)^p)$.
- Puis on développe $(\cos x + i \sin x)^p$.
- Enfin, il reste à récupérer soit la partie réelle, soit la partie imaginaire selon que l'on veuille $\cos(px)$ ou $\sin(px)$.

Le développement de $(\cos x + i \sin x)^p$ se fait :

- soit peu à peu à partir de $(\cos x + i \sin x)^2$;
- soit avec la formule du binôme de Newton (pour ceux qui la connaissent).

Remarque : on démontre (c'est un exercice que l'on pourra faire au moment du chapitre sur les polynômes) que l'expression $\cos(px)$ (pour tout $p \in \mathbf{N}$) peut toujours s'exprimer sous la forme d'un polynôme de variable $\cos x$. Attention, ce résultat n'est pas toujours vrai pour l'expression $\sin(px)$ qui ne peut pas s'exprimer sous la forme d'un polynôme de variable $\sin x$ dans le cas où p est pair. Penser par exemple à $\sin 2x = 2 \cos x \sin x$.

Exemple : écrivons $\sin(5x)$ en fonction de différentes puissances de $\sin x$. On écrit

$$\begin{aligned} \sin(5x) &= \Im((\cos x + i \sin x)^5) \\ &= \Im(\cos^5 x + 5i \cos^4 x \sin x - 10 \cos^3 x \sin^2 x - 10i \cos^2 x \sin^3 x + 5 \cos x \sin^4 x + i \sin^5 x) \\ &= 5 \cos^4 x \sin x - 10 \cos^2 x \sin^3 x + \sin^5 x \end{aligned}$$

et en utilisant $\cos^2 x = 1 - \sin^2 x$ puis en développant $\cos^4 x = (1 - \sin^2 x)^2$, on a :

$$\sin(5x) = 5 \sin x - 20 \sin^3 x + 16 \sin^5 x$$

L'expression est bien ici une fonction polynomiale en $\sin x$.

❏ **Méthode 2.8.—** Comment simplifier certaines sommes de réels en utilisant les nombres complexes

L'idée est d'utiliser un complexe Z_n (qui s'écrit lui aussi sous forme de somme) dont la partie réelle ou la partie imaginaire est la somme de départ à simplifier. Il s'agit de simplifier Z_n et de récupérer alors $\Re(Z_n)$ ou $\Im(Z_n)$.

Donnons les formes les plus classiques, n étant un entier supérieur ou égal à 1, et $q \in \mathbf{C}$:

- si $q \neq 1$ et si $Z_n = \sum_{k=0}^{n-1} q^k$ alors $Z_n = \frac{1 - q^n}{1 - q}$.
- si $Z_n = \sum_{k=0}^n \binom{n}{k} q^k$ alors $Z_n = (1 + q)^n$.

La première formule est bien entendu la formule de la somme partielle d'une suite géométrique et la deuxième est la formule du binôme de Newton.

Exemple : calculons pour tout réel $x \in \mathbf{R}$, la somme $S_n = \sum_{k=0}^n \cos(kx)$. Si x est un multiple de 2π , alors S_n est la somme de $(n+1)$ fois le réel 1 et vaut donc $n+1$.

Si $x \neq 2k\pi$, posons $Z_n = \sum_{k=0}^n e^{ikx} = \frac{1 - e^{i(n+1)x}}{1 - e^{ix}} = \frac{e^{i\frac{(n+1)x}{2}} - 2i \sin\left(\frac{(n+1)x}{2}\right)}{e^{i\frac{x}{2}} - 2i \sin\left(\frac{x}{2}\right)}$

et en suivant la méthode, on récupère la partie réelle : $S_n = \cos\left(\frac{nx}{2}\right) \frac{\sin\left(\frac{(n+1)x}{2}\right)}{\sin\left(\frac{x}{2}\right)}$.

Ce développement est à savoir et il fait partie du Top 10 de ce qu'il faut connaître. Lorsqu'on pose cet exercice (ou un proche) aux oraux de deuxième année et que le candidat est bloqué, cela fait très mauvais effet, vous en conviendrez !

Nous vous laissons le soin de montrer si $x \neq 2k\pi$, $\sum_{k=0}^n \sin(kx) = \sin\left(\frac{nx}{2}\right) \frac{\sin\left(\frac{(n+1)x}{2}\right)}{\sin\left(\frac{x}{2}\right)}$.

■ Résolution d'équations algébriques dans \mathbf{C}

□ **Méthode 2.9.—** Comment résoudre l'équation $z^2 = \omega$, où z et ω sont complexes

Soit l'équation $z^2 = \omega$, d'inconnue $z \in \mathbf{C}$, pour la résoudre :

- on peut écrire le complexe ω sous **forme trigonométrique**, (en restreignant au cas qui reste général où ω n'est pas nul), on a alors $\omega = \rho e^{i\theta}$ avec $\rho \in \mathbf{R}_+^*$ et $\theta \in [0, 2\pi[$. Les solutions de l'équation sont $\sqrt{\rho} e^{i\theta/2}$ et $-\sqrt{\rho} e^{i\theta/2}$. On les appelle **les racines carrées** de ω ;
- on peut écrire aussi le complexe ω sous **forme algébrique** c'est-à-dire $\omega = a + ib$ avec $(a, b) \in \mathbf{R}^2$. On cherche z sous la forme $z = x + iy$ avec $(x, y) \in \mathbf{R}^2$.

La première méthode est à privilégier si la forme trigonométrique de ω est (relativement) simple !

Exemple : cherchons les racines carrées de $1 + i$ en employant la méthode trigonométrique. On écrit $1 + i = \sqrt{2} e^{i\frac{\pi}{4}}$ et les racines carrées de $1 + i$ sont $2^{1/4} e^{i\frac{\pi}{8}}$ et $-2^{1/4} e^{i\frac{\pi}{8}}$ c'est-à-dire

$$2^{1/4} \left(\cos\left(\frac{\pi}{8}\right) + i \sin\left(\frac{\pi}{8}\right) \right) \text{ et } 2^{1/4} \left(\cos\left(\frac{\pi}{8}\right) - i \sin\left(\frac{\pi}{8}\right) \right)$$

Détaillons maintenant la deuxième méthode, partons des implications

$$(x + iy)^2 = a + ib \Rightarrow |x + iy|^2 = |a + ib| \Rightarrow x^2 + y^2 = \sqrt{a^2 + b^2}$$

puis de l'implication : $(x + iy)^2 = a + ib \Rightarrow x^2 - y^2 + i2xy = a + ib$. Il reste à récupérer la partie réelle et la partie imaginaire de la dernière égalité obtenue et on obtient le système

$$\begin{cases} x^2 - y^2 = a \\ 2xy = b \\ x^2 + y^2 = \sqrt{a^2 + b^2} \end{cases} \Leftrightarrow \begin{cases} x^2 = a/2 + \sqrt{a^2 + b^2}/2 \\ xy \text{ et } b \text{ ont même signe} \\ y^2 = -a/2 + \sqrt{a^2 + b^2}/2 \end{cases}$$

On en déduit les deux solutions $z_1 = \sqrt{\frac{a}{2} + \frac{1}{2}\sqrt{a^2 + b^2}} + i\epsilon \sqrt{-\frac{a}{2} + \frac{1}{2}\sqrt{a^2 + b^2}}$ et $-z_1$ en posant $\epsilon = 1$ si $b \geq 0$ et $\epsilon = -1$ si $b < 0$.

Se dire que dans la pratique, il n'y a guère que ces deux méthodes. La deuxième a l'avantage de donner un résultat explicite qui peut d'ailleurs ne pas être très beau à voir si sous les racines, rien ne s'arrange! Notez au passage qu'il arrive de résoudre les équations du type $z^4 = a$, avec $a \in \mathbf{C}$ en cherchant les racines carrées de a puis les racines carrées de ces dernières et tout cela avec la méthode algébrique. Cela fait un bon exercice de colle en perspective!

Exemple : cherchons les racines carrées de $1 + i$ en employant la méthode algébrique. Si on pose $(x + iy)^2 = 1 + i$, avec $(x, y) \in \mathbf{R}^2$, on écrit

$$\begin{cases} x^2 - y^2 = 1 \\ 2xy = 1 \\ x^2 + y^2 = \sqrt{2} \end{cases} \Leftrightarrow \begin{cases} x^2 = 1/2 + \sqrt{2}/2 \\ xy \text{ et } 1 \text{ ont même signe} \\ y^2 = -1/2 + \sqrt{2}/2 \end{cases}$$

On obtient les deux solutions : $z_1 = \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{2}} + i\sqrt{-\frac{1}{2} + \frac{1}{2}\sqrt{2}}$ et $-z_1$. On compare avec l'autre méthode et on en déduit, sachant que $\cos\left(\frac{\pi}{8}\right) > 0$, les relations

$$2^{1/4} \cos\left(\frac{\pi}{8}\right) = \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{2}} \text{ et } 2^{1/4} \sin\left(\frac{\pi}{8}\right) = \sqrt{-\frac{1}{2} + \frac{1}{2}\sqrt{2}}$$

ce qui donne les formules : $\cos\left(\frac{\pi}{8}\right) = \sqrt{\frac{1}{2} + \frac{1}{4}\sqrt{2}}$ et $\sin\left(\frac{\pi}{8}\right) = \sqrt{\frac{1}{2} - \frac{1}{4}\sqrt{2}}$. Voilà une application des nombres complexes à la trigonométrie (très classique, donc à retenir!).

□ Méthode 2.10.— Comment résoudre $(E) : az^2 + bz + c = 0$, $(a, b, c) \in \mathbf{C}^* \times \mathbf{C}^2$

On pose $\Delta = b^2 - 4ac$:

- si $\Delta = 0$, l'équation (E) a pour unique solution $z = -\frac{b}{2a}$;
- si $\Delta \neq 0$, l'équation (E) a deux solutions distinctes $z = \frac{-b + \delta}{2a}$ et $z = \frac{-b - \delta}{2a}$, où δ est une racine carrée de Δ .

Exemple : Soit $(E) : 2z^2 - (9i+1)z - 7 + 11i = 0$. On calcule $\Delta = (1+9i)^2 - 8(-7+11i) = -24 - 70i$. Puis l'on détermine les racines carrées de Δ en employant la méthode algébrique, c'est-à-dire en cherchant des réels x et y tels que $-24 - 70i = (x + iy)^2$.

On obtient le système : $\begin{cases} x^2 - y^2 = -24 \\ 2xy = -70 \\ x^2 + y^2 = \sqrt{24^2 + 70^2} \end{cases} \Leftrightarrow \begin{cases} x^2 = 25 \\ xy = -35 \\ y^2 = 49 \end{cases}$. Les deux racines carrées de

Δ sont $5 - 7i$ et $-5 + 7i$ et les racines de (E) sont $4i - 1$ et $\frac{1}{2}(3 + i)$.

□ **Méthode 2.11.—** Comment résoudre une équation polynomiale (E_n) du type $az^{2n} + bz^n + c = 0$, avec $(a, b, c) \in \mathbf{C}^* \times \mathbf{C}^2$ et n un entier supérieur ou égal à 2

On pose $Z = z^n$ et on résout l'équation (E) du second degré $aZ^2 + bZ + c = 0$. Puis :

- si $\Delta = b^2 - 4ac \neq 0$, alors (E) a deux solutions distinctes z_1 et z_2 . On résout ensuite les deux équations $z^n = z_1$ et $z^n = z_2$. On obtient alors $2n$ solutions;
- si $\Delta = b^2 - 4ac = 0$, alors (E) a une solution z_1 . On résout ensuite l'équation $z^n = z_1$ et on obtient alors n solutions doubles.

■ Résolution d'équations ou d'inéquations trigonométriques

□ **Méthode 2.12.—** Comment résoudre une équation ou une inéquation trigonométrique

On utilise les formules trigonométriques (**proposition 1.15**, **proposition 1.16**, **proposition 1.17** et **proposition 1.18**) et le cercle trigonométrique. Souvent, on se ramène à une équation polynomiale d'inconnue $\cos x$, $\sin x$ ou $\tan x$.

Exemple : L'équation $\cos(2x) + \cos x = 0$ peut se résoudre de deux manières, par exemple en posant $\cos 2x = -\cos x = \cos(x + \pi)$ et donc $2x = \pm(x + \pi) + 2k\pi$, avec $k \in \mathbf{Z}$, donc $x = (2k + 1)\pi$ ou $x = \frac{\pi}{3} + 2k\frac{\pi}{3}$. Une autre façon est de remarquer que $\cos 2x = 2\cos^2 x - 1$ et donc $\cos x$ est solution de l'équation $2X^2 + X - 1 = 0$, c'est-à-dire $X = -1$ ou $X = \frac{1}{2}$.
On retrouve les valeurs précédentes.

□ **Méthode 2.13.—** Comment résoudre $\alpha \cos x + \beta \sin x = \gamma$, où $(\alpha, \beta, \gamma) \in \mathbf{R}^3$

- On peut poser $z = e^{ix}$ et donc $\cos x = \frac{1}{2}(z + z^{-1})$ et $\sin x = \frac{1}{2i}(z - z^{-1})$. L'équation (E) est alors équivalente à une équation polynomiale du second degré d'inconnue z .
- On peut aussi écrire $\alpha \cos x + \beta \sin x$ sous la forme $A \cos(x - \phi)$.

■ Plan complexe

□ Méthode 2.14.— Comment utiliser les affixes pour résoudre des problèmes de géométrie plane. Soient trois points $A(a)$, $B(b)$ et $C(c)$ tous distincts. On rappelle que l'affixe de \overrightarrow{AB} est $b - a$ et la norme de \overrightarrow{AB} est $|b - a|$.

- ▶ Trois points tous distincts $A(a)$, $B(b)$ et $C(c)$ sont alignés si et seulement si $\frac{c - a}{b - a} \in \mathbf{R}$.
- ▶ Soient trois points tous distincts $A(a)$, $B(b)$ et $C(c)$. Les vecteurs \overrightarrow{AB} et \overrightarrow{AC} sont orthogonaux si et seulement si $\frac{c - a}{b - a} \in i\mathbf{R}$.
- ▶ Le triangle (ABC) est rectangle en A dans le sens direct si et seulement si les vecteurs \overrightarrow{AB} et \overrightarrow{AC} sont issus l'un de l'autre par une rotation d'angle $\pi/2$.
En termes d'affixes, c'est :

$$\exists \lambda \in \mathbf{R}_+^*, c - a = i\lambda(b - a) \Leftrightarrow \frac{c - a}{b - a} = -\frac{\bar{c} - \bar{a}}{\bar{b} - \bar{a}}.$$

Si, de plus le triangle est isocèle en A , on a $\lambda = 1$.

- ▶ Le triangle (ABC) est équilatéral dans le sens direct si et seulement si $AB = AC$ et si les vecteurs \overrightarrow{AB} et \overrightarrow{AC} font un angle de $\pi/3$, c'est-à-dire qu'ils sont issus l'un de l'autre par une rotation d'angle $\pi/3$.
En termes d'affixes, on a : $c - a = e^{i\pi/3}(b - a)$.
- ▶ n points $M_1(z_1), M_2(z_2), \dots, M_n(z_n)$ sont sur le même cercle de centre O si et seulement si $|z_1| = |z_2| = \dots = |z_n|$.

Exemple : Déterminons les nombres complexes z tels que les images des nombres complexes z, z^2 et z^3 forment un triangle rectangle. La difficulté c'est que nous ne savons pas en quel point le triangle est rectangle. De plus, on évite les cas où deux (au moins) des trois points sont confondus. Nous prendrons donc : $z \notin \{-1, 0, 1\}$. Selon le sommet en lequel le triangle est rectangle, il existe un réel non nul λ tel que :

$$z - z^3 = \lambda i(z - z^2) \text{ ou } z^3 - z^2 = \lambda i(z - z^2) \text{ ou } z^3 - z = \lambda i(z^3 - z^2).$$

Les conditions s'écrivent alors :

$$1 + z \in i\mathbf{R} \text{ ou } z \in i\mathbf{R} \text{ ou } \frac{1 + z}{z} \in i\mathbf{R}.$$

La première condition correspond à la verticale d'abscisse -1 , la seconde correspond à la verticale d'abscisse 0 . Pour la troisième, posons $z = x + iy$, où $(x, y) \in \mathbf{R}^2$. On a :

$$\frac{1 + z}{z} \in i\mathbf{R} \Leftrightarrow (1 + z)\bar{z} \in i\mathbf{R}.$$

Cela donne $x^2 + y^2 + a = 0$. C'est le cercle de centre $(-1/2, 0)$ et de rayon $1/2$.

Le lieu cherché est donc la réunion des deux droites et du cercle, privé des points communs au cercle et aux droites.

Vrai/Faux

	Vrai	Faux
1. i est égal à sa partie imaginaire.	<input type="checkbox"/>	<input type="checkbox"/>
2. On peut très bien avoir l'égalité $ z + z' = z + z' $ sans que z et z' soient obligatoirement liés par un réel.	<input type="checkbox"/>	<input type="checkbox"/>
3. Deux complexes dont la somme et le produit sont réels, sont réels.	<input type="checkbox"/>	<input type="checkbox"/>
4. $\forall (a, b) \in \mathbf{C}^2, (a + ib = 0 \Rightarrow a = b = 0)$.	<input type="checkbox"/>	<input type="checkbox"/>
5. Le nombre complexe 0 est de module et d'argument nuls.	<input type="checkbox"/>	<input type="checkbox"/>
6. Deux complexes de même module dont la différence des arguments est 2π sont égaux.	<input type="checkbox"/>	<input type="checkbox"/>
7. Si $z \in \mathbf{C}^*$, alors $-z$ a pour argument $\pi + \arg(z)$ et z^n avec $n \in \mathbf{N}$ a pour argument $n + \arg(z)$.	<input type="checkbox"/>	<input type="checkbox"/>
8. Si $z \in \mathbf{C}^*$, il admet deux racines carrées opposées $\sqrt{ z }e^{i(\arg(z))/2}$ et $-\sqrt{ z }e^{i(\arg(z))/2}$	<input type="checkbox"/>	<input type="checkbox"/>
9. Soit z un nombre complexe de module égal à 1. Alors, il existe un entier $n \in \mathbf{N}^*$ tel que le nombre complexe z soit une racine n -ième de l'unité.	<input type="checkbox"/>	<input type="checkbox"/>
10. Les racines carrées du complexe i sont \sqrt{i} et $-\sqrt{i}$.	<input type="checkbox"/>	<input type="checkbox"/>
11. Les racines cubiques de l'unité sont 1, $j = e^{i\pi/3}$ et $\bar{j} = e^{-i\pi/3}$.	<input type="checkbox"/>	<input type="checkbox"/>
12. Si $n \in \mathbf{N}^*$, les n racines $n^{\text{ièmes}}$ de z complexe non nul, s'obtiennent en multipliant une racine $n^{\text{ième}}$ particulière de z par les n racines $n^{\text{ièmes}}$ de 1.	<input type="checkbox"/>	<input type="checkbox"/>
13. Si l'équation $ax^2 + bx + c = 0$ d'inconnue x et de constantes a, b et c complexes, admet pour solution le complexe z , alors \bar{z} est aussi solution.	<input type="checkbox"/>	<input type="checkbox"/>
14. On a l'implication : $e^z = -1 \Rightarrow z = i\pi$.	<input type="checkbox"/>	<input type="checkbox"/>

Chapitre 3

Révision : algèbre linéaire

■ Les incontournables

- ▶ montrer qu'une famille est libre, génératrice dans un espace vectoriel de dimension quelconque ;
- ▶ justifier qu'une application est linéaire ;
- ▶ bien faire le lien entre une application linéaire et sa matrice ;
- ▶ approfondir les notions de somme et de somme directe de sous-espaces vectoriels, ainsi que de sous-espaces vectoriels supplémentaires, et donc de projecteur ;
- ▶ déterminer le rang d'une famille de vecteurs ou d'une application linéaire ;
- ▶ utiliser le théorème du rang ;
- ▶ transformer une matrice en changeant de base ;
- ▶ savoir déterminer les solutions d'un système par différentes méthodes ;
- ▶ savoir inverser une matrice par différentes méthodes.

■ Et plus si affinités

- ▶ établir que deux sous-espaces vectoriels sont supplémentaires dans un espace vectoriel de dimension finie ou non ;
- ▶ calculer la trace d'un endomorphisme d'un \mathbf{K} -espace vectoriel de dimension n ;
- ▶ savoir calculer A^n par différentes méthodes.

Résumé de cours

\mathbf{K} désigne indifféremment l'un des ensembles \mathbf{R} ou \mathbf{C} .

■ Matrices de $\mathcal{M}_{n,p}(\mathbf{K})$ (rappel de TSI1)

Définition : Soit $(n, p) \in \mathbf{N}^* \times \mathbf{N}^*$, on appelle **matrice de n lignes et p colonnes** toute famille $(a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ d'éléments de \mathbf{K} . On la représente sous la forme d'un tableau rectangulaire :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}$$

Dans la pratique, on assimilera la famille $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ et la représentation sous forme de tableau ci-dessus. On note $\mathcal{M}_{n,p}(\mathbf{K})$ l'ensemble des matrices à n lignes et p colonnes.

- Si $p = 1$, (resp $n = 1$) une matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ est appelée une **matrice colonne** (resp. **ligne**).
- Si $n = p$, une matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ est appelée une **matrice carrée d'ordre n** et on note alors $\mathcal{M}_n(\mathbf{K})$ l'ensemble des matrices carrées d'ordre n .
- Dans la suite, pour tout i de 1 à n , L_i désigne la $i^{\text{ème}}$ ligne et pour tout j de 1 à p , C_j désigne la $j^{\text{ème}}$ colonne d'une matrice donnée.

Définition : Soit $(n, p) \in \mathbf{N}^* \times \mathbf{N}^*$ et $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ une matrice de $\mathcal{M}_{n,p}(\mathbf{K})$, on dit que :

- A est **triangulaire supérieure** (resp. **inférieure**) si : $i > j \Rightarrow a_{i,j} = 0$ (resp. $i < j \Rightarrow a_{i,j} = 0$).
- A est **diagonale** si : $i \neq j \Rightarrow a_{i,j} = 0$.
- Pour toute matrice $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ de $\mathcal{M}_{n,p}(\mathbf{K})$, on appelle **diagonale** de A la suite ordonnée de réels $a_{1,1}, \dots, a_{k,k}$, où k est le plus petit des deux entiers n et p .
- On note I_n la matrice carrée d'ordre n diagonale et n'ayant que des 1 sur la diagonale. Ainsi,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ et } I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Nous noterons $O_{n,p}$ la matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ n'ayant que des 0 pour coefficients. Elle se note parfois plus simplement O si la taille des matrices est fixée.

Remarque : On note souvent $\text{diag}(a_{1,1}, \dots, a_{n,n})$ une matrice diagonale carrée d'ordre n ayant pour diagonale principale $a_{1,1}, \dots, a_{n,n}$.

Définition : somme de deux matrices et multiplication par un scalaire—.

Soit $(n, p) \in \mathbf{N}^* \times \mathbf{N}^*$ et soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ et $B = (b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ deux matrices de $\mathcal{M}_{n,p}(\mathbf{K})$, on appelle **somme des matrices A et B** la matrice $A + B$ de $\mathcal{M}_{n,p}(\mathbf{K})$ définie par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket, \quad (A + B)_{i,j} = a_{i,j} + b_{i,j}$$

Si $\lambda \in \mathbf{K}$, on appelle produit de la matrice A et de λ la matrice λA de $\mathcal{M}_{n,p}(\mathbf{K})$ définie par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket, \quad (\lambda A)_{i,j} = \lambda a_{i,j}.$$

Définition : Transposition d'une matrice—.

Soit $(n, p) \in \mathbf{N}^* \times \mathbf{N}^*$ et soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ une matrice de $\mathcal{M}_{n,p}(\mathbf{K})$, on appelle transposée de la matrice A , la matrice A^T (notée aussi tA) de $\mathcal{M}_{p,n}(\mathbf{K})$ définie par :

$$\forall (i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, n \rrbracket, \quad (A^T)_{i,j} = a_{j,i}.$$

Définition : Produit de deux matrices—.

Soit $(n, p, q) \in \mathbf{N}^* \times \mathbf{N}^* \times \mathbf{N}^*$ et $(A, B) \in \mathcal{M}_{n,p}(\mathbf{K}) \times \mathcal{M}_{p,q}(\mathbf{K})$, on définit $A \times B \in \mathcal{M}_{n,q}(\mathbf{K})$ par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, q \rrbracket, \quad (A \times B)_{i,j} = \sum_{k=1}^p a_{i,k} \times b_{k,j}$$

$A \times B$ s'écrira plus simplement AB . Attention à cet abus que nous ferons nous aussi.

Mise en œuvre : exercice 01.

Proposition 3.1.— Écriture matricielle des systèmes d'équations linéaires —. Considérons le système de n équations linéaires à p inconnues (S) ci-dessous. On note $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$ la matrice des coefficients du premier membre, $B = (b_i) \in \mathcal{M}_{n,1}(\mathbf{K})$ la matrice colonne *second membre* et $X = (x_i) \in \mathcal{M}_{p,1}(\mathbf{K})$ la matrice colonne des inconnues. On a alors l'écriture matricielle :

$$(S) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p = b_2 \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = b_n \end{cases} \iff A \times X = B$$

■ Opérations élémentaires sur les matrices

Si i et j sont deux entiers distincts dans $\llbracket 1, p \rrbracket$, on a les opérations sur les lignes :

$$L_i \leftarrow L_i + \lambda L_j, \quad L_i \leftarrow \lambda L_i, \quad L_i \leftrightarrow L_j.$$

Application : résolution d'un système linéaire

On « concatène » les matrices A et B en $(A|B)$.

En utilisant les opérations élémentaires sur les lignes appliquées à la matrice concaténée, on aboutit à une matrice triangulaire supérieure qui est dite équivalente en lignes à la première et en remontant, on en déduit les inconnues du système associé. Les dernières lignes de la matrice terminale peuvent être les relations de compatibilité du système. **Mise en œuvre : exercice 02, exercice 03.**

Si i et j sont deux entiers distincts dans $\llbracket 1, p \rrbracket$, on a les opérations sur les colonnes :

$$C_i \leftarrow C_i + \lambda C_j, \quad C_i \leftarrow \lambda C_i, \quad C_i \leftrightarrow C_j.$$

Attention, ces opérations sont verbotes pour résoudre un système car sinon on modifie les inconnues. Par contre, elles peuvent être utiles pour montrer qu'une famille est libre, génératrice ou trouver le rang d'une matrice (voir plus loin).

■ Matrices carrées

On désigne par n un entier naturel non nul.

Déjà, on peut remarquer que le produit de deux matrices carrées de même ordre existe toujours. Un exercice classique est alors de trouver toutes les matrices $M \in \mathcal{M}_n(\mathbf{K})$ qui commutent avec une matrice A donnée dans $\mathcal{M}_n(\mathbf{K})$, c'est-à-dire les matrices $M \in \mathcal{M}_n(\mathbf{K})$ qui vérifient : $MA = AM$.

Définition : Puissances successives d'une matrice—. Soit $A \in \mathcal{M}_n(\mathbf{K})$ une matrice carrée d'ordre n . On définit les **puissances successives** de A par récurrence :

$$\left\{ \begin{array}{l} \bullet A^0 = I_n \\ \bullet \forall k \in \mathbf{N}, A^{k+1} = A \times A^k \end{array} \right.$$

Proposition 3.2.— Formule du binôme de Newton pour deux matrices commutantes—.

Soit $A, B \in \mathcal{M}_n(\mathbf{K})$ telles que $A \times B = B \times A$. Alors :

$$(A + B)^p = \sum_{k=0}^p \binom{p}{k} A^k \times B^{p-k}.$$

Pour le calcul des puissances successives d'une matrice carrée, voir la **méthode 3.10**.

Mise en œuvre : exercice 04, exercice 05.

Transposée d'une matrice carrée

Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_n(\mathbf{K})$. La transposée de A , notée A^T , est la matrice (\tilde{a}_{ij}) de $\mathcal{M}_n(\mathbf{K})$ définie par $\tilde{a}_{ij} = a_{ji}$.

Proposition 3.3.— Pour toutes matrices A, B de $\mathcal{M}_n(\mathbf{K})$, on a :

$$(A + B)^T = A^T + B^T, (\lambda A)^T = \lambda A^T, \lambda \in \mathbf{K}, (A^T)^T = A$$

Proposition 3.4.— Soit A une matrice de $\mathcal{M}_n(\mathbf{K})$ et B une matrice de $\mathcal{M}_n(\mathbf{K})$, on a :

$$(AB)^T = (B^T)(A^T)$$

Définition : Une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est dite **symétrique** si elle est égale à sa transposée, c'est-à-dire si elle vérifie : $A = A^T$.

Notation : On note $\mathcal{S}_n(\mathbf{K})$ l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbf{K})$.

Définition : Une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est dite *antisymétrique* si elle est égale à l'opposée de sa transposée, c'est-à-dire si elle vérifie : $A = -A^T$.

Notation : On note $\mathcal{A}_n(\mathbf{K})$ l'ensemble des matrices antisymétriques de $\mathcal{M}_n(\mathbf{K})$.

Remarque : Au programme officiel de TSI2, pour l'ensemble des matrices symétriques, on impose la notation $\mathcal{S}_n(\mathbf{K})$ et pour l'ensemble des matrices antisymétriques, on impose $\mathcal{A}_n(\mathbf{K})$.

Mise en œuvre : exercice 06.

■ Matrices carrées inversibles

Définition : Une matrice carrée $A \in \mathcal{M}_n(\mathbf{K})$ est dite *inversible* s'il existe une matrice carrée $B \in \mathcal{M}_n(\mathbf{K})$ telle que $A \times B = I_n$ et $B \times A = I_n$.

Si tel est le cas, B est unique : on l'appelle l'*inverse* de A , et on note $B = A^{-1}$.

Théorème-Définition 3.5.— On note $GL_n(\mathbf{K})$ l'ensemble des matrices carrées d'ordre n inversibles. Il est appelé **groupe linéaire**. Tout produit de matrices de $GL_n(\mathbf{K})$ est une matrice de $GL_n(\mathbf{K})$.

Théorème 3.6.— **Propriétés des matrices inversibles** —. Soit $A, B \in \mathcal{M}_n(\mathbf{K})$, alors :

- Si A est inversible, alors A^{-1} est inversible et $(A^{-1})^{-1} = A$.
- Si A est inversible alors A^T est inversible et $(A^T)^{-1} = (A^{-1})^T$.
- Si A et B sont inversibles, alors $A \times B$ est inversible et $(A \times B)^{-1} = B^{-1} \times A^{-1}$

Théorème 3.7.— **Lien fondamental avec les systèmes linéaires** —. Supposons que la matrice A appartienne à $\mathcal{M}_n(\mathbf{K})$ et la matrice B appartienne à $\mathcal{M}_{n,1}(\mathbf{K})$. On note (S) le système $A \times X = B$. On a alors l'équivalence des propriétés suivantes :

- A est inversible (et appartient donc à $GL_n(\mathbf{K})$.)
- A est équivalent en lignes à I_n . On note donc $A \sim_L I_n$.
- Le système homogène associé à (S) , $AX = 0$ n'admet que la solution nulle.
- Pour tout B , le système (S) admet une unique solution $X = A^{-1} \times B$.
- Pour tout B , le système (S) admet au moins une solution

Remarque : comme conséquences, on obtient les critères d'inversibilité suivants :

- **Inversibilité des matrices triangulaires :** soit $T \in \mathcal{M}_n(\mathbf{K})$ une matrice triangulaire. Alors T est inversible si et seulement si ses coefficients diagonaux sont **tous non nuls**.
- **Inversibilité d'une matrice d'ordre 2 :** soit $A \in \mathcal{M}_2(\mathbf{K})$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Alors A est inversible si et seulement si $a \times d - b \times c \neq 0$, et dans ce cas, $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Théorème 3.8.— Inversibilité à gauche ou à droite —. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors

A est inversible	<i>si et seulement si</i>	il existe $B \in \mathcal{M}_n(\mathbf{K})$ telle que $A \times B = I_n$
	<i>si et seulement si</i>	il existe $B \in \mathcal{M}_n(\mathbf{K})$ telle que $B \times A = I_n$.

Pour le calcul de l'inverse d'une matrice carrée par résolution d'un système linéaire et par la méthode du pivot de Gauss-Jordan, voir les méthodes.

Mise en œuvre : exercice 07, exercice 08, exercice 09, exercice 10.

■ Espaces vectoriels, sous-espaces vectoriels

Définition : Un **espace vectoriel** est un ensemble E muni d'éléments appelés **vecteurs**. On munit E d'une opération entre vecteurs que l'on note généralement $+$ dont le résultat est un vecteur et d'une opération entre les éléments d'un ensemble \mathbf{K} (qui est \mathbf{R} ou \mathbf{C}) et les vecteurs de E que l'on note $.$ généralement et dont le résultat est un élément de E .

Les éléments de \mathbf{K} sont appelés des **scalaires**.

Si $\mathbf{K} = \mathbf{R}$, on dit que E est un espace vectoriel réel et si $\mathbf{K} = \mathbf{C}$, on dit que E est un espace vectoriel complexe. De façon plus générale, E est dit un espace vectoriel sur \mathbf{K} ou aussi un \mathbf{K} -espace vectoriel.

Nous noterons avec une flèche les éléments de E (par exemple \vec{x}) et normalement c'est-à-dire sans flèche les éléments de \mathbf{K} .

Attention, en général un vecteur n'est pas un scalaire et réciproquement (sauf si $E = \mathbf{K}$ ce qui est possible).

Les opérations $+$ et $.$ ont les propriétés suivantes :

$$\forall (\vec{x}, \vec{y}, \vec{z}) \in E^3, \vec{x} + \vec{y} = \vec{y} + \vec{x}, \vec{x} + \vec{0}_E = \vec{x}, (-\vec{x}) + \vec{x} = \vec{0}_E, \vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}.$$

$$\forall (\vec{x}, \vec{y}) \in E, \forall (a, b) \in \mathbf{K}^2, a.(b.\vec{x}) = (ab).\vec{x}, (a + b).\vec{x} = a.\vec{x} + b.\vec{x}, a.(\vec{x} + \vec{y}) = a.\vec{x} + b.\vec{y}, 1.\vec{x} = \vec{x}.$$

Ici $\vec{0}_E$ est appelé le **vecteur nul** de E .

Exemples : Donnons quelques exemples génériques à connaître.

- L'ensemble \mathbf{R} muni de $+$ et de \times est un espace vectoriel réel.
- L'ensemble \mathbf{C} muni de $+$ et de \times est un espace vectoriel complexe.
- L'ensemble \mathbf{K}^2 des couples de scalaires (x, y) , où $x \in \mathbf{K}$ et $y \in \mathbf{K}$ est un espace vectoriel sur \mathbf{K} . Avec les lois :

$$\forall (x, y, x', y', a) \in \mathbf{K}^5, (x, y) + (x', y') = (x + x', y + y'), a.(x, y) = (ax, ay).$$

Ainsi le vecteur nul $\vec{0}_{\mathbf{R}^2} = (0, 0)$.

- De façon plus générale, l'ensemble \mathbf{K}^n des n -uplets de scalaires (x_1, \dots, x_n) , où $x_i \in \mathbf{K}$ pour tout entier $i \in \llbracket 1, n \rrbracket$ est un espace vectoriel sur \mathbf{K} .
- L'ensemble $\mathbf{K}[X]$ des polynômes à coefficients dans \mathbf{K} est un espace vectoriel sur \mathbf{K} .
- L'ensemble $\mathcal{C}^0(I, \mathbf{K})$ des fonctions continues sur I et à valeurs dans \mathbf{K} est un espace vectoriel sur \mathbf{K} .

- L'ensemble $\mathcal{B}(I, \mathbf{K})$ des fonctions définies sur I et bornées sur I et à valeurs dans \mathbf{K} est un espace vectoriel sur \mathbf{K} .
- L'ensemble $\mathcal{C}^1(I, \mathbf{K})$ des fonctions de classe \mathcal{C}^1 sur I et à valeurs dans \mathbf{K} est un espace vectoriel sur \mathbf{K} .
- L'ensemble $\mathcal{C}^\infty(I, \mathbf{K})$ des fonctions de classe \mathcal{C}^∞ sur I et à valeurs dans \mathbf{K} est un espace vectoriel sur \mathbf{K} .
- L'ensemble $\mathbf{C}^{\mathbf{N}}$ des suites à valeurs complexes est un espace vectoriel complexe.
- L'ensemble $\mathcal{M}_{n,p}(\mathbf{K})$ des matrices à coefficients dans \mathbf{K} est un espace vectoriel sur \mathbf{K} (avec les lois classiques d'addition et de multiplication par un élément de \mathbf{K} d'une matrice).

Dans la pratique, démontrer qu'un ensemble est un espace vectoriel

Définition : Définition d'un sous-espace vectoriel Soit E un espace vectoriel sur \mathbf{K} . Alors le sous-ensemble F de E est un sous-espace vectoriel de E si et seulement si l'on a les trois assertions à la fois :

- F est non vide.
- $\forall a \in \mathbf{K}, \forall \vec{x} \in F, a \cdot \vec{x} \in F$.
- $\forall (\vec{x}, \vec{y}) \in F^2, \vec{x} + \vec{y} \in F$.

On peut collapser les deux dernières assertions en une seule :

$\forall a \in \mathbf{K}, \forall (\vec{x}, \vec{y}) \in F^2, a \cdot \vec{x} + \vec{y} \in F$.

Remarque : $\vec{0}_E$ appartient toujours à tout sous-espace vectoriel de E . Ainsi, pour montrer que F n'est pas vide, il suffit de vérifier que $\vec{0}_E \in F$. Ainsi, si $\vec{0}_E \notin F$, alors F n'est pas un sous-espace vectoriel de E .

Exemples : • $F = \{(x, y) \in \mathbf{R}^2, x + 3y = 0\}$ est un sous-espace vectoriel de \mathbf{R}^2 , espace vectoriel sur \mathbf{R} .

• $F = \{(x, y) \in \mathbf{R}^2, x + 3y = 2021\}$ n'est pas un sous-espace vectoriel de \mathbf{R}^2 , espace vectoriel sur \mathbf{R} car $(0, 0) \notin F$.

• Le sous-ensemble des suites convergentes est un sous-espace vectoriel de $\mathbf{R}^{\mathbf{N}}$.

• Le sous-ensemble des fonctions s'annulant en 0 est un sous-espace vectoriel de $\mathcal{F}(\mathbf{R}, \mathbf{R})$, espace vectoriel des fonctions de \mathbf{R} dans \mathbf{R} .

• Le sous-ensemble des complexes imaginaires purs est un sous-espace vectoriel de \mathbf{C} espace vectoriel sur \mathbf{R} .

• Le sous-ensemble $\mathbf{R}_3[X]$ des polynômes de degré au plus 3 est un sous-espace vectoriel de $\mathbf{R}[X]$, espace vectoriel sur \mathbf{R} .

• Le sous-ensemble des complexes de module 1 n'est pas un sous-espace vectoriel de \mathbf{C} espace vectoriel sur \mathbf{R} .

Mise en œuvre : exercice 11, exercice 12, exercice 13, exercice 14.

Proposition 3.9.— $\mathcal{S}_n(\mathbf{K})$ et $\mathcal{A}_n(\mathbf{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbf{K})$.

Proposition 3.10.— **Somme de deux sous-espace vectoriels**

Dans E espace vectoriel sur \mathbf{K} , on considère F et G deux sous-espaces vectoriels de E .

Alors $F + G = \{\vec{x} + \vec{y}, \vec{x} \in F \text{ et } \vec{y} \in G\}$.

Par ailleurs $F + G$ est un sous-espace vectoriel de E .

Définition : Somme directe de deux sous-espaces vectoriels. Sous-espaces vectoriels supplémentaires

Soient F et G deux sous-espaces vectoriels de E , espace vectoriel sur \mathbf{K} .

- On dit que la somme $F + G$ est **directe** si et seulement si $F \cap G = \{0_E\}$.
On note alors $F \oplus G$.
- On dit que F et G sont **supplémentaires** dans E si et seulement si $F \oplus G = E$.

Exemple : Dans le plan \mathbf{R}^2 les deux axes Ox et Oy sont des sous-espaces vectoriels de \mathbf{R}^2 et $Ox \oplus Oy = \mathbf{R}^2$ car tout vecteur de \mathbf{R}^2 s'écrit comme somme d'un vecteur de Ox et d'un vecteur de Oy .

Proposition 3.11.— Les deux sous-espaces vectoriels $\mathcal{S}_n(\mathbf{K})$ et $\mathcal{A}_n(\mathbf{K})$ sont des sous-espaces vectoriels supplémentaires dans $\mathcal{M}_n(\mathbf{K})$.

Mise en œuvre : exercice 15, exercice 16, exercice 17.

Remarque : Nous allons rappeler plus loin des méthodes pour montrer la supplémentarité de deux sous-espaces vectoriels en utilisant la notion de base ou en utilisant la notion de dimension et même aborder la notion de somme de plus de deux sous-espaces vectoriels.

■ Familles libres, familles génératrices, bases

Définition : Combinaison linéaire —. On appelle combinaison linéaire d'un nombre fini de vecteurs $\vec{u}_1, \dots, \vec{u}_p$ de E toute somme $\sum_{i=1}^p \lambda_i \vec{u}_i$, où $\lambda_1, \dots, \lambda_p$ sont des scalaires (c'est-à-dire des éléments de \mathbf{K}), appelés coefficients de la combinaison linéaire.

Proposition 3.12.— L'ensemble des combinaisons linéaires d'une famille finie $X = (\vec{u}_1, \dots, \vec{u}_p)$ est un sous-espace vectoriel de E , appelé **sous-espace vectoriel engendré** par X et noté $\text{Vect}(X)$.

Définition : Famille libre, famille liée —.

► $(\vec{u}_i)_{i \in [1, n]}$, famille **finie** de vecteurs de E , est **libre** (on dit aussi que ces vecteurs sont indépendants) si et seulement si pour toute famille $(\lambda_i)_{i \in [1, n]} \in \mathbf{K}^n$

$$\sum_{i=1}^n \lambda_i \vec{u}_i = 0_E \Rightarrow \forall i \in [1, n], \lambda_i = 0.$$

► On peut étendre à $(\vec{u}_i)_{i \in I}$, famille **dénombrable** de vecteurs de E . Elle est **libre** si et seulement si toutes ses sous-familles finies sont libres.

► Si une famille n'est pas libre, on dit qu'elle est **liée**.

Remarques : ► Toute famille contenue dans une famille libre est libre.

► Toute famille contenant une famille liée est liée. (C'est la contraposition de l'implication précédente.) En particulier, toute famille contenant le vecteur nul est liée.

► Une famille de deux vecteurs est liée si et seulement si ces deux vecteurs sont colinéaires.

► Toute famille contenant le vecteur nul est liée.

Proposition 3.13.— Famille liée et combinaison linéaire —. On suppose I fini. Une famille $(\vec{u}_i)_{i \in I}$ est liée si et seulement si l'un au moins des \vec{u}_i est combinaison linéaire des autres.

Mise en œuvre : exercice 18.

Définition : Famille génératrice —. Une famille X de E est une **famille génératrice de l'espace vectoriel** E si et seulement si tout vecteur de E est une combinaison linéaire finie de vecteurs de E . Dans le cas où X est une famille finie, on a alors : $\text{Vect } X = E$.

Définition : Base —. Une famille de vecteurs du \mathbf{K} -espace vectoriel E est une **base** de E si et seulement si elle est une famille libre et génératrice.

Exemples : ▶ $(X^i)_{i \in \mathbf{N}}$ est une base de $\mathbf{K}[X]$ et $(X^i)_{i \in [0, n]}$ est une base de $\mathbf{K}_n[X]$.

▶ $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ est une base de \mathbf{R}^3 .

Ce sont les bases les plus simples de chacun de ces espaces vectoriels : on dit que ce sont leurs **bases canoniques** respectives.

Théorème 3.14.— Pour tout $(k, l) \in [1, n] \times [1, p]$, notons $E_{k,l}$ la matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ dont tous les coefficients valent 0 sauf celui à l'intersection de la ligne k et de la colonne l qui vaut 1. La famille de matrices $\{E_{k,l}\}$ est une base de $\mathcal{M}_{n,p}(\mathbf{K})$ appelée **base canonique** de $\mathcal{M}_{n,p}(\mathbf{K})$.

Mise en œuvre : exercice 19, exercice 20, exercice 21.

■ Dimension

Définition : Espace vectoriel de dimension finie —. Un espace vectoriel est **de dimension finie** si et seulement s'il admet une famille génératrice finie.

Proposition 3.15.— Théorème d'existence d'une base et de la base incomplète —.

- ▶ Tout \mathbf{K} -espace vectoriel E de dimension finie admet une base finie.
- ▶ On peut extraire de toute famille génératrice finie de E une base de E .
- ▶ Toute famille libre de E peut être complétée en une base de E .

Lemme 3.16.— Dans un espace vectoriel de dimension finie, une famille libre ne peut avoir plus d'éléments qu'une famille génératrice.

Théorème 3.17.— Théorème de la dimension —.

Toutes les bases d'un espace vectoriel E de dimension finie ont le même nombre d'éléments, appelé **dimension de** E et noté $\dim E$.

Par convention, la dimension de $\{0_E\}$ est 0.

Proposition 3.18.— Si $\dim E = n$ et si \mathcal{B} est une famille de n vecteurs de E , alors il y a équivalence entre :

- (1) \mathcal{B} est une base de E (2) \mathcal{B} est une famille libre (3) \mathcal{B} est une famille génératrice de E .

Mise en œuvre : exercice 22, exercice 23, exercice 24, exercice 25.

Proposition 3.19.— Tout sous-espace vectoriel F d'un espace vectoriel de dimension finie E est de dimension finie et $\dim F \leq \dim E$. De plus, si $\dim F = \dim E$, alors $F = E$.

Théorème 3.20.— $(\mathcal{M}_{n,p}(\mathbf{K}), +, \cdot)$ est un espace vectoriel sur \mathbf{K} de dimension np .

Remarque : On a : $\dim \mathcal{S}_n(\mathbf{K}) = \frac{n(n+1)}{2}$ et $\dim \mathcal{A}_n(\mathbf{K}) = \frac{n(n-1)}{2}$.

Mise en œuvre : exercice 26.

Définition : Base adaptée à un sous-espace vectoriel —. Soit F un sous-espace vectoriel de dimension p du \mathbf{K} -espace vectoriel de dimension finie E , une base de E est **adaptée à F** lorsque ses p premiers vecteurs forment une base de F .

Proposition 3.21.— **Coordonnées d'un vecteur dans une base** —. Soit $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ une base de E . Pour tout vecteur \vec{x} de E , il existe un n -uplet unique (x_1, x_2, \dots, x_n) de \mathbf{K}^n tel que :

$$\vec{x} = \sum_{i=1}^n x_i \vec{e}_i.$$

(x_1, x_2, \dots, x_n) est le *n -uplet des coordonnées du vecteur x dans la base \mathcal{B}* .

Mise en œuvre : exercice 27, exercice 28.

Exemple : Les coordonnées d'un polynôme de degré inférieur ou égal à n dans la base canonique de $\mathbf{K}_n[X]$ sont ses coefficients.

Définition : Rang d'une famille de vecteurs —. Le *rang d'une famille finie de vecteurs* est la dimension du sous-espace vectoriel qu'elle engendre.

Notation : Le rang de la famille \mathcal{F} est noté $\text{Rg}(\mathcal{F})$.

Mise en œuvre : exercice 29.

■ Somme et somme directe d'une famille finie de sous-espaces vectoriels

Définition : Somme —. Soit $(E_i)_{i \in [1, n]}$ une famille finie de sous-espaces vectoriels de E , la *somme de ces sous-espaces vectoriels* est le sous-espace vectoriel, noté $\sum_{i=1}^n E_i$, engendré par la réunion des E_i :

$$\sum_{i=1}^n E_i = \text{Vect} \left(\bigcup_{i=1}^n E_i \right).$$

L'appellation « somme de sous-espaces vectoriels » est justifiée par la caractérisation suivante.

Proposition 3.22.— $\sum_{i=1}^n E_i$ est l'ensemble des sommes de vecteurs des E_i , autrement dit :

$$\sum_{i=1}^n E_i = \left\{ \sum_{i=1}^n \vec{u}_i, \text{ où, } \forall i \in \llbracket 1, n \rrbracket, \vec{u}_i \in E_i \right\}.$$

Définition : Somme directe —. La somme des sous-espaces vectoriels de la famille $(E_i)_{i \in \llbracket 1, n \rrbracket}$ est directe si et seulement si tout \vec{u} de $\sum_{i=1}^n E_i$ se décompose **de manière unique** sous la forme :

$$\vec{u} = \sum_{i=1}^n \vec{u}_i, \text{ où } (\vec{u}_i)_{i \in \llbracket 1, n \rrbracket} \in E_1 \times E_2 \times \cdots \times E_n.$$

Dans ce cas, on note la somme de ces sous-espaces vectoriels : $\bigoplus_{i=1}^n E_i$.

Proposition 3.23.— **Caractérisation par l'unicité de la décomposition du vecteur nul**— La somme des sous-espaces vectoriels de la famille $(E_i)_{i \in \llbracket 1, n \rrbracket}$ est directe si et seulement si :

$$\forall (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n) \in E_1 \times E_2 \times \cdots \times E_n, \left(\sum_{i=1}^n \vec{u}_i = 0_E \Rightarrow \forall i \in \llbracket 1, n \rrbracket, \vec{u}_i = 0_{E_i} \right).$$

Théorème-Définition 3.24.— **Base adaptée à une décomposition en somme directe** —.

Soit E un \mathbf{K} -espace vectoriel de dimension finie et $(E_i)_{i \in \llbracket 1, n \rrbracket}$ une famille de n sous-espaces vectoriels de E , de bases respectives \mathcal{B}_i , $1 \leq i \leq n$.

Si la somme des $(E_i)_{i \in \llbracket 1, n \rrbracket}$ est directe, alors $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_n)$ est une base de $\bigoplus_{i=1}^n E_i$.

Si, de plus, $E = \bigoplus_{i=1}^n E_i$ alors $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_n)$ est une base de E dite **base adaptée à la décomposition** $E = \bigoplus_{i=1}^n E_i$.

Corollaire 3.25.— **Dimension de la somme de deux sous-espaces vectoriels** —.

Soient F et G deux sous-espaces vectoriels d'un espace vectoriel de dimension finie E :

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

Mise en œuvre : exercice 30, exercice 31, exercice 32.

■ Application linéaire, lien avec les matrices (rappel de TS11)

On désigne par E et F deux \mathbf{K} -espaces vectoriels.

Définition : Une application $u : E \rightarrow F$ est dite linéaire si elle vérifie :

$$\forall (\vec{x}, \vec{y}) \in E^2, \forall \lambda \in \mathbf{K}, \quad u(\vec{x} + \lambda \vec{y}) = u(\vec{x}) + \lambda u(\vec{y})$$

Proposition 3.26.— L'application trace définie par : $\mathcal{M}_n(\mathbf{K}) \rightarrow \mathbf{K}, A \mapsto \text{Tr}(A)$ est linéaire.

Notation : On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F .

$\mathcal{L}(E, F)$ est un \mathbf{K} -espace vectoriel. Lorsque E et F sont de dimension finie, on a :

$$\dim \mathcal{L}(E, F) = \dim(E) \times \dim(F)$$

- ▷ Une application linéaire de E dans E est appelée endomorphisme de E .
- ▷ Une application linéaire de E dans F bijective est appelée isomorphisme.
- ▷ On dit que deux espaces vectoriels E et F sont isomorphes s'il existe un isomorphisme de E vers F .
- ▷ Deux espaces vectoriels de dimension finie sont isomorphes si et seulement s'ils ont la même dimension. En particulier tout \mathbf{K} -espace vectoriel de dimension n est isomorphe à \mathbf{K}^n .

Mise en œuvre : exercice 33, exercice 34.

Définition d'une application linéaire par image d'une base

Théorème 3.27.— Soit $(\vec{e}_i)_{i \in I}$ une base de E et $(\vec{f}_i)_{i \in I}$ une famille de vecteurs appartenant à un espace vectoriel F .

- ▶ Il existe une unique application linéaire u de E vers F telle que :

$$\forall i \in I, \quad u(\vec{e}_i) = \vec{f}_i.$$

- ▶ Pour que u soit injective, il faut et il suffit que la famille $(\vec{f}_i)_{i \in I}$ soit libre.
- ▶ Pour que u soit surjective, il faut et il suffit que la famille $(\vec{f}_i)_{i \in I}$ engendre F .
- ▶ Pour que u soit bijective, il faut et il suffit que la famille $(\vec{f}_i)_{i \in I}$ soit une base de F .

Une application linéaire u de E vers F est entièrement déterminée par son action sur une base de E .

▷ Supposons que E et F aient pour dimensions respectives p et n .

Soit $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_p)$ une base de E et $\mathcal{C} = (\vec{f}_1, \dots, \vec{f}_n)$ une base de F . Une application linéaire u de E vers F est entièrement déterminée par la donnée des vecteurs $u(\vec{e}_j)$, qui, eux-mêmes, sont déterminés par leurs coordonnées dans la base \mathcal{C} .

Définition : La matrice de u par rapport aux bases \mathcal{B} et \mathcal{C} et on note $\mathcal{M}_{\mathcal{B}, \mathcal{C}}(u)$, la matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ dont la $j^{\text{ième}}$ colonne est formée des coordonnées du vecteur $u(\vec{e}_j)$ dans la base \mathcal{C} .

Remarque : Si l'on note $\mathcal{M}_{\mathcal{B}, \mathcal{C}}(u) = (a_{ij})$, on a pour tout $j \in \llbracket 1, p \rrbracket$: $u(\vec{e}_j) = \sum_{i=1}^n a_{ij} \vec{f}_i$.

Mise en œuvre : exercice 35, exercice 36.

Noyau, image d'une application linéaire

Définition : Soit $u \in \mathcal{L}(E, F)$. On appelle noyau de u l'ensemble noté $\ker(u)$, défini par

$$\ker(u) = \{\vec{x} \in E \mid u(\vec{x}) = 0_F\}$$

Proposition 3.28.— Soit $u \in \mathcal{L}(E, F)$.

Le noyau de u est un sous-espace vectoriel de E .

L'application u est injective si et seulement si son noyau est réduit à $\{0_E\}$.

Définition : Soit $u \in \mathcal{L}(E, F)$. On appelle image de u l'ensemble noté $\text{Im}(u)$ défini par :

$$\text{Im}(u) = u(E) = \{u(\vec{x}) \mid \vec{x} \in E\} = \{y \in F \mid \exists \vec{x} \in E, y = u(\vec{x})\}$$

Proposition 3.29.— Soit $u \in \mathcal{L}(E, F)$. L'image de u est sous-espace vectoriel de F .

Mise en œuvre : exercice 37, exercice 38, exercice 39, exercice 40, exercice 41, exercice 42.

Notion de rang

Définition : Le rang d'une application linéaire u de E vers F est la dimension de son image. Il est noté $\text{Rg}(u)$.

Si E est un \mathbf{K} -espace vectoriel de dimension n et (e_1, \dots, e_n) est une base de E , le rang de $u \in \mathcal{L}(E, F)$ est le rang de la famille de vecteurs $(u(e_1), \dots, u(e_n))$.

Définition : Le rang d'une matrice $A \in \mathcal{M}_{n,p}(\mathbf{K})$ est le rang de ses p vecteurs colonnes dans \mathbf{K}^n . Il est noté $\text{Rg}(A)$.

Le rang de A est un entier vérifiant $\text{Rg}(A) \leq \min(n, p)$. Il s'agit du nombre maximal de colonnes de A linéairement indépendantes dans \mathbf{K}^n .

▷ Le rang d'une matrice est le rang de toute application linéaire qu'elle représente.

▷ Le rang d'une matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ est également le rang de ses n vecteurs lignes dans \mathbf{K}^p . Autrement dit, une matrice et sa transposée ont même rang.

Remarque : On peut utiliser les opérations élémentaires sur les lignes et colonnes pour simplifier une matrice et en déterminer son rang qui lui n'est pas modifié par ces opérations. L'idée est d'aboutir à une matrice triangulaire supérieure ou inférieure. On obtient alors son rang sans souci.

Mise en œuvre : exercice 43.

Théorème 3.30 (théorème du rang).— Soit E un \mathbf{K} -espace vectoriel de dimension finie, F un \mathbf{K} -espace vectoriel et $u \in \mathcal{L}(E, F)$. Alors $\text{Im}(u)$ est de dimension finie et :

$$\dim E = \dim(\ker u) + \text{Rg}(u)$$

Mise en œuvre : exercice 44, exercice 45.

Corollaire 3.31.— Soit E et F deux \mathbf{K} -espaces vectoriels de dimension finie et $u \in \mathcal{L}(E, F)$.

On suppose que E et F ont même dimension : $\dim E = \dim F$.

Les propositions suivantes sont alors équivalentes :

$$u \text{ est bijective} \Leftrightarrow u \text{ est injective} \Leftrightarrow u \text{ est surjective}$$

Changement de bases en dimension finie

Définition : Soit E un \mathbf{K} -espace vectoriel de dimension n , $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$ deux bases de E . On appelle *matrice de passage de la base \mathcal{B} à la base \mathcal{B}'* , la matrice de passage dont la $j^{\text{ème}}$ colonne est formée des coordonnées du vecteur e'_j par rapport à la base (e_1, \dots, e_n) .

Notons P la matrice de passage entre les bases \mathcal{B} et \mathcal{B}' . P est la matrice par rapport aux bases \mathcal{B}' et \mathcal{B} (attention à l'ordre!), de l'application linéaire Id_E .

Soit Q la matrice de passage de la base \mathcal{B}' à la base \mathcal{B} ; Q est la matrice de Id_E par rapport aux bases \mathcal{B} et \mathcal{B}' . On sait alors que QP est la matrice de $\text{Id}_E \circ \text{Id}_E = \text{Id}_E$ par rapport aux bases \mathcal{B}' et \mathcal{B}' ; donc $QP = I_n$. La matrice P est inversible et $P^{-1} = Q$.

Théorème 3.32.— Soit E un \mathbf{K} -espace vectoriel de dimension n , u un endomorphisme de E , \mathcal{B} et \mathcal{B}' deux bases de E . Si l'on note A et A' les matrices de u respectivement dans les bases \mathcal{B} et \mathcal{B}' :

$$A' = P^{-1}AP$$

égalité dans laquelle P désigne la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' .

Mise en œuvre : exercice 46.

Remarque : Supposons que A soit la matrice d'un endomorphisme d'un \mathbf{K} -espace vectoriel E par rapport à une base. Dire que A' est semblable à A signifie que A' est la matrice du même endomorphisme par rapport à une autre base de E .

Définition : On appelle *trace d'un endomorphisme u de E* la trace d'une matrice représentant u dans une base de E . On la note $\text{Tr}(u)$.

Cas particuliers importants : projections et symétries

Définition : Soient F et G deux sous-espaces vectoriels supplémentaires de E . Alors tout vecteur \vec{x} de E s'écrit de façon unique sous la forme $\vec{x} = \vec{x}_F + \vec{x}_G$, où $\vec{x}_F \in F$ et $\vec{x}_G \in G$. L'application qui à \vec{x} associe \vec{x}_F est la *projection sur F dans la direction de G* et l'application qui à \vec{x} associe \vec{x}_G est la *projection sur G dans la direction de F* .

Proposition 3.33.— Les deux applications projections précédentes sont des endomorphismes de E , leur image est le sous-espace vectoriel sur lequel on projette et leur noyau est leur direction.

Dans une base adaptée à la décomposition en somme directe, la matrice d'une projection est diagonale avec des 1 et des 0 sur la diagonale.

Comment définir analytiquement une projection connaissant les deux sous-espaces vectoriels F et G ?

Mise en œuvre : exercice 47.

Définition : Un endomorphisme p de E est un *projecteur* s'il vérifie $p^2 = p$.

Pour un projecteur p de E , on a : $\text{Im}(p) = \text{Ker}(p - \text{Id}_E)$.

Théorème 3.34.— Soit p un projecteur de E . Les sous-espaces vectoriels $\text{Ker}(p)$ et $\text{Im}(p)$ sont supplémentaires dans E : $E = \text{Ker}(p) \oplus \text{Im}(p)$. Alors p est la projection sur $\text{Im}(p)$ dans la direction de $\text{Ker}(p)$.

Comment à partir d'une matrice reconnaître une projection et ses caractéristiques

Mise en œuvre : exercice 48.

Définition : Soient F et G deux sous-espaces vectoriels supplémentaires de E . L'application qui à \vec{x} associe $\vec{x}_F - \vec{x}_G$ est la symétrie vectorielle par rapport à F dans la direction de G .

Proposition 3.35.— Si P est la projection sur F dans la direction de G et s la symétrie vectorielle par rapport à F dans la direction de G alors : $s = 2P - \text{Id}_E$.

Dans une base adaptée à la décomposition en somme directe, la matrice d'une symétrie est diagonale avec des 1 et des -1 sur la diagonale.

Comment définir analytiquement une symétrie connaissant les deux sous-espaces vectoriels F et G ?

Mise en œuvre : exercice 49.

Définition : Un endomorphisme s de E est une symétrie s'il vérifie $s^2 = \text{Id}_E$.

Proposition 3.36.— Si $s^2 = \text{Id}_E$, alors s est la symétrie vectorielle par rapport à $\text{Ker}(s - \text{Id}_E)$ dans la direction de $\text{Ker}(s + \text{Id}_E)$.

Comment à partir d'une matrice reconnaître une symétrie vectorielle et ses caractéristiques

Mise en œuvre : exercice 50.

Toute symétrie de E est un endomorphisme bijectif, égal à son inverse.

■ Sous-espace stable

Définition : On dit qu'un sous-espace vectoriel F de E est stable par un endomorphisme u de E si $u(F)$ est contenu dans F , c'est-à-dire si et seulement si : $u(F) \subset F$.

Théorème 3.37.— Soit E un \mathbf{K} -espace vectoriel de dimension n , F un sous-espace vectoriel de dimension p et u un endomorphisme de E . Les propositions suivantes sont équivalentes :

- ▶ F est stable par u ;
- ▶ pour tout supplémentaire G de F dans E , la matrice de u dans toute base \mathcal{B} adaptée à la décomposition $E = F \oplus G$ est de la forme :

$$\mathcal{M}(u) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

où $A \in \mathcal{M}_p(\mathbf{K})$, $B \in \mathcal{M}_{p,n-p}(\mathbf{K})$ et $C \in \mathcal{M}_{n-p}(\mathbf{K})$.

Méthodes

■ Produit de deux matrices

Avant d'effectuer le produit de deux matrices A et B , dans cet ordre, il faut d'abord vérifier que le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B . Cela donnera un produit AB qui a pour nombre de lignes celui de la matrice A et pour nombre de colonnes celui de la matrice B . Avant de faire le produit, on doit donc avoir la taille de AB ce qui permettra de nous guider.

□ Méthode 3.1.— Comment effectuer le produit de A et B dans cet ordre

- On écrit de gauche à droite le contenu de la matrice A puis celui de la matrice B et enfin la place du produit AB .
- On calcule ensuite chaque coefficient d'indice (i, j) de AB , par exemple pour avoir son coefficient à l'intersection de la ligne i et colonne j , on encadre la ligne L_i de A et la colonne C_j de B puis on effectue la somme des produits deux à deux des coefficients de L_i et de C_j , de gauche vers la droite pour L_i et de haut en bas pour C_j .

$\forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, p\} :$

$$(AB)_{i,j} = \left(\begin{array}{ccc} a_{i,1} & \cdot & a_{i,k} & \cdot & a_{i,n} \end{array} \right) \left(\begin{array}{c} b_{1,j} \\ \cdot \\ b_{k,j} \\ \cdot \\ b_{n,j} \end{array} \right)$$

ce qui donne bien

$$(AB)_{i,j} = a_{i,1}b_{1,j} + \dots + a_{i,k}b_{k,j} + \dots + a_{i,n}b_{n,j}$$

Exemple : 1) Calculons le produit de $A = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}$ par $B = \begin{pmatrix} 1 & -i \\ 1-i & 1+i \end{pmatrix}$. Comme A et B sont des matrices carrées d'ordre 2 à coefficients dans \mathbf{C} , il en est de même de leur produit. Par exemple, le coefficient c_{11} de AB est :

$$c_{11} = i \times 1 - 1 \times (1 - i) = 2i - 1.$$

On trouve :

$$A \times B = AB = \begin{pmatrix} -1+2i & -i \\ 2+i & -1 \end{pmatrix}.$$

De même, on peut calculer :

$$B \times A = BA = \begin{pmatrix} 0 & 0 \\ 2(1+i) & 2(-1+i) \end{pmatrix} = 2 \begin{pmatrix} 0 & 0 \\ 1+i & -1+i \end{pmatrix}.$$

On remarque au passage que $AB \neq BA$, ce qui est généralement le cas.

2) Calculons maintenant le produit matriciel :

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

On remarque que ce produit étant le produit d'une matrice de $\mathcal{M}_{1,2}(\mathbf{R})$ par une matrice de $\mathcal{M}_{2,3}(\mathbf{R})$ puis par une matrice de $\mathcal{M}_{3,3}(\mathbf{R})$ puis par une matrice de $\mathcal{M}_{3,2}(\mathbf{R})$ puis enfin par une matrice de $\mathcal{M}_{2,1}(\mathbf{R})$, on obtient une matrice à une ligne et une colonne. Pour la calculer, d'après l'associativité du produit, on peut calculer par exemple le produit des deux premières, notons le P (c'est une matrice à une ligne et trois colonnes) et parallèlement faire le produit des deux suivantes que l'on note Q (c'est une matrice à trois lignes et deux colonnes). Puis on fait le produit PQ qui donne une matrice à une ligne et deux colonnes et enfin on fait le produit du résultat obtenu avec la dernière matrice. Calculons :

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

On écrit en appliquant la règle de produit rappelée plus haut,

$$P_{1,1} = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} = 2, \quad P_{1,2} = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 3$$

pour les deux premiers coefficients et enfin pour le troisième coefficient

$$P_{1,3} = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} = -3$$

Ensuite, on calcule la matrice Q de la même façon, on trouve

$$Q = \begin{pmatrix} -2 & 3 \\ 3 & -5 \\ 2 & 2 \end{pmatrix}$$

Puis il nous faut calculer le produit PQ , cela donnera (le faire),

$$PQ = \begin{pmatrix} -1 & -15 \end{pmatrix}$$

et enfin le produit de PQ par $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ qui fait (-44) .

Remarque : dans le cas de matrices carrées, on peut remarquer (et on retiendra que) :

- le produit de deux matrices triangulaires supérieures (resp. inférieures) est une matrice triangulaire supérieure (resp. inférieure) ;
- le produit de matrices diagonales est une matrice diagonale.

Mise en œuvre : exercice ??, exercice ??.

■ Comment montrer qu'une famille de vecteurs est libre

- **Méthode 3.2.** — ► Si la famille est finie, on prouve qu'une combinaison linéaire des vecteurs de cette famille est nulle si, et seulement si, ses coefficients sont tous nuls.
- Si la famille n'est pas finie, on prouve que toute sous-famille finie extraite de cette famille est libre. On effectue alors souvent une récurrence.
 - On peut aussi montrer que cette famille est l'image d'une famille libre par une application linéaire injective.
 - Si l'on connaît les composantes des vecteurs dans une base donnée, on peut faire appel au pivot de Gauss (voir le chapitre consacré aux matrices).

Exemple : Pour $n \in \mathbf{N}$, soit $\vec{f}_n : x \mapsto \cos^n(x)$. Montrer que la famille $(\vec{f}_n)_{n \in \mathbf{N}}$ est une famille libre du \mathbf{R} -espace vectoriel $\mathbb{F}(\mathbf{R}, \mathbf{R})$ des fonctions de \mathbf{R} dans \mathbf{R} .

Soit I une partie finie contenue dans \mathbf{N} , montrons que la famille $(\vec{f}_i)_{i \in I}$ est libre. Soit $(\lambda_i)_{i \in I}$ une famille de réels telle que :

$$\sum_{i \in I} \lambda_i \vec{f}_i = 0 \quad \Longleftrightarrow \quad (1) \forall x \in \mathbf{R}, \quad \sum_{i \in I} \lambda_i \cos^i(x) = 0.$$

Notons P le polynôme de $\mathbf{R}[X]$ défini par $P = \sum_{i \in I} \lambda_i X^i$. De l'égalité (1), on déduit que pour tout réel $x : P(\cos(x)) = 0$. On en déduit que P est le polynôme nul car il admet une infinité de racines. Tous les λ_i sont nuls et la famille $(\vec{f}_i)_{i \in I}$ est libre.

■ Comment montrer qu'une famille \mathcal{B} est une base du \mathbf{K} -espace vectoriel E

□ **Méthode 3.3.** — ► On utilise la définition : on montre que \mathcal{B} est libre et génératrice dans E .

- Si $\dim(E) = n$ et si \mathcal{B} est de cardinal n , il suffit de montrer que \mathcal{B} est libre (respectivement génératrice).
- On peut montrer que \mathcal{B} est une base adaptée à une décomposition en somme directe de E .
- On peut aussi montrer que \mathcal{B} est l'image d'une base d'un espace vectoriel F par un isomorphisme de F dans E .

Exemple : Soit E le sous-espace vectoriel de $\mathcal{M}_n(\mathbf{K})$ constitué des matrices diagonales.

1. Montrer que $f : \mathbf{R}^n \rightarrow E, (x_1, \dots, x_n) \mapsto \text{diag}(x_1, \dots, x_n)$ est un isomorphisme.

2. Soit alors A la matrice de E définie par : $A = \text{diag}(1, 2, \dots, n)$. Montrer que la famille $(I_n, A, A^2, \dots, A^{n-1})$ est une base de E .

1. L'application : $f : \mathbf{R}^n \rightarrow E, (x_1, \dots, x_n) \mapsto \text{diag}(x_1, \dots, x_n)$ est un isomorphisme ; en effet elle est clairement linéaire car :

$$f((x_1, \dots, x_n) + \lambda(y_1, \dots, y_n)) = \text{diag}(x_1 + \lambda y_1, \dots, x_n + \lambda y_n) = f((x_1, \dots, x_n)) + \lambda f((y_1, \dots, y_n))$$

pour tout $\lambda \in \mathbf{K}$. De plus, f est surjective par construction.

Enfin, si $(x_1, \dots, x_n) \in \ker f$, $x_1 = \dots = x_n = 0$. f est bien injective donc est un isomorphisme.

2. Ainsi $\dim(E) = \dim(\mathbf{R}^n) = n$. Puisque la famille $(I_n, A, A^2, \dots, A^{n-1})$ est de cardinal n , il suffit de justifier qu'elle est libre pour établir qu'elle forme une base de E . Soit $\lambda_0, \dots, \lambda_{n-1}$ des réels tels que :

$$\lambda_0 I_n + \lambda_1 A + \dots + \lambda_{n-1} A^{n-1} = 0. \quad (*)$$

Introduisons le polynôme P de $\mathbf{R}_{n-1}[X]$ en posant : $P = \lambda_0 + \lambda_1 X + \dots + \lambda_{n-1} X^{n-1}$. De la relation (*), on déduit que pour tout entier $k \in \llbracket 1, n \rrbracket$: $P(k) = 0$. Le polynôme P possède alors n racines distinctes ; étant de degré inférieur ou égal à $n-1$, il est nul. Tous les λ_i sont donc nuls et la famille (I_n, A, \dots, A^{n-1}) est une base de E .

■ Comment montrer que la somme de p sous-espaces vectoriels est directe

□ **Méthode 3.4.**— Soit F_1, \dots, F_p une famille de p sous-espaces vectoriels d'un \mathbf{K} -espace vectoriel E .

► Si $p = 2$, on montre que $F_1 \cap F_2 = \{0_E\}$ pour justifier que F_1 et F_2 sont en somme directe.

(Attention, ce résultat n'est plus suffisant si $p \geq 3$.)

► Si $p \geq 3$, on utilise la proposition du cours. Pour cela on montre que :

$$\left. \begin{array}{l} \vec{f}_1 + \vec{f}_2 + \dots + \vec{f}_p = 0 \\ (\vec{f}_1, \vec{f}_2, \dots, \vec{f}_p) \in F_1 \times F_2 \times \dots \times F_p \end{array} \right\} \implies \forall i \in \llbracket 1, p \rrbracket, \vec{f}_i = 0.$$

Exemple : Soit Q un polynôme non nul de $\mathbf{K}[X]$ et $p \in \mathbf{N}^*$. Pour tout entier $k \in \llbracket 1, p \rrbracket$, on pose $F_k = \text{Vect}(X^k Q)$. Montrer que les sous-espaces vectoriels F_1, \dots, F_p sont en somme directe.

Soit R_1, \dots, R_p des polynômes appartenant respectivement à F_1, \dots, F_p tels que :

$$R_1 + R_2 + \dots + R_p = 0.$$

Pour tout entier $i \in \llbracket 1, p \rrbracket$, il existe un scalaire λ_i tel que $R_i = \lambda_i X^i Q$. Puisque $Q \neq 0$, on a :

$$\lambda_1 X Q + \lambda_2 X^2 Q + \dots + \lambda_p X^p Q = 0 \implies \lambda_1 X + \lambda_2 X^2 + \dots + \lambda_p X^p = 0.$$

Puisque la famille $(X^k)_{k \in \llbracket 1, p \rrbracket}$ est libre, tous les λ_i sont nuls. La somme des F_i est donc directe.

■ Comment montrer que $E = F \oplus G$

□ **Méthode 3.5.** — ► On choisit un vecteur \vec{x} quelconque de E et on montre qu'il s'écrit de manière unique sous la forme $\vec{x} = \vec{f} + \vec{g}$ avec $(\vec{f}, \vec{g}) \in F \times G$. Pour cela, on peut procéder par *analyse-synthèse*.

- *Analyse* – On suppose qu'un tel couple existe et on cherche à exprimer \vec{f} et \vec{g} uniquement en fonction de \vec{x} .
- *Synthèse* – On vérifie que les vecteurs \vec{f} et \vec{g} trouvés précédemment conviennent ; c'est-à-dire que $(\vec{f}, \vec{g}) \in F \times G$ et $\vec{f} + \vec{g} = \vec{x}$.
- On peut aussi montrer que la réunion d'une base de F et d'une base de G est une base de E .
- Si E est de dimension finie, on peut utiliser le **théorème 3.25**. Dans la plupart des cas, on essaiera de montrer que $F \cap G = 0_E$ puis que $\dim(E) = \dim(F) + \dim(G)$.

Exemple : Soit E un \mathbf{K} -espace vectoriel, u un endomorphisme de E vérifiant $u + u^3 = 0$. Montrer que $\ker(u)$ et $\ker(u^2 + u)$ sont supplémentaires dans E .

Soit $\vec{x} \in E$.

► *Analyse* – On suppose qu'il existe un couple $(\vec{y}, \vec{z}) \in \ker(u) \times \ker(u^2 + u)$ tel que $\vec{x} = \vec{y} + \vec{z}$. Puisque $u(\vec{y}) = 0$, on a $u^2(\vec{x}) = u^2(\vec{z})$. Or $\vec{z} \in \ker(u^2 + u)$, ainsi $u^2(\vec{z}) = -\vec{z}$. On vient de prouver que nécessairement $\vec{z} = -u^2(\vec{x})$ et $\vec{y} = \vec{x} + u^2(\vec{x})$.

► *Synthèse* – Posons $\vec{y} = \vec{x} + u^2(\vec{x})$ et $\vec{z} = -u^2(\vec{x})$.

On a bien $\vec{y} + \vec{z} = \vec{x}$.

Puisque $u + u^3 = 0$, alors $u(\vec{y}) = 0$ et $\vec{y} \in \ker(u)$.

On a : $(u^2 + u)(\vec{z}) = -(u^4 + u^2)(\vec{x}) = -u(u^3(\vec{x}) + u(\vec{x})) = 0$. Puisque $u^3 + u = 0$, on a bien $\vec{z} \in \ker(u^2 + u)$.

Tout vecteur de E s'écrit de manière unique comme somme d'un vecteur de $\ker(u)$ et d'un vecteur de $\ker(u^2 + u)$: $E = \ker(u) \oplus \ker(u^2 + u)$.

■ Comment montrer que F est stable par u

□ **Méthode 3.6.** — ► On montre que l'image par u d'un vecteur quelconque de F est encore un vecteur de F .

- Si l'on sait que $F = \text{Vect}(\vec{f}_1, \dots, \vec{f}_n)$, il suffit de montrer que pour tout $i \in \llbracket 1, n \rrbracket$, $u(\vec{f}_i)$ est un vecteur de F .

Exemple : Soit $n \in \mathbf{N}^*$. Montrer que $\mathbf{R}_n[X]$ est stable par l'application linéaire Δ définie sur $\mathbf{R}[X]$ par :

$$\forall P \in \mathbf{R}[X], \quad \Delta(P) = (X^2 + 1)P' - nXP.$$

Puisque $(1, X, \dots, X^n)$ est une base de $\mathbf{R}_n[X]$, pour justifier que $\mathbf{R}_n[X]$ est stable par Δ il suffit de vérifier que pour tout entier $k \in \llbracket 0, n \rrbracket$, on a : $\Delta(X^k) \in \mathbf{R}_n[X]$.

Soit $k \in \llbracket 0, n \rrbracket$, on a : $\Delta(X^k) = (k-n)X^{k+1} + kX^{k-1}$.

▷ Si $k < n$, on a $\deg(\Delta(X^k)) = k+1 \leq n$ et par conséquent $\Delta(X^k) \in \mathbf{R}_n[X]$.

▷ Pour $k = n$, on a $\Delta(X^n) = nX^{n-1} \in \mathbf{R}_n[X]$.

Ainsi, $\mathbf{R}_n[X]$ est stable par Δ .

■ Sachant que $E = F \oplus G$, comment déterminer la projection d'un vecteur \vec{x} sur F parallèlement à G (rappel de TSI1)

□ **Méthode 3.7.**— On décompose le vecteur \vec{x} suivant F et G , c'est-à-dire on détermine l'unique couple $(\vec{f}, \vec{g}) \in F \times G$ tel que $\vec{x} = \vec{f} + \vec{g}$. La projection de \vec{x} sur F parallèlement à G est alors le vecteur \vec{f} .

Exemple : Soit $F = \{(x, y, z) \in \mathbf{R}^3 \mid x + y + z = 0\}$ et $G = \{(x, y, z) \in \mathbf{R}^3 \mid x = -y = z\}$.

1. Justifier que F et G sont supplémentaires dans \mathbf{R}^3 .
2. Soit $(x, y, z) \in \mathbf{R}^3$, déterminer la projection de (x, y, z) sur F parallèlement à G .
3. En déduire la matrice dans la base canonique de \mathbf{R}^3 de la projection sur F parallèlement à G , puis de la projection sur G parallèlement à F .

1. F est un hyperplan de \mathbf{R}^3 , il est de dimension 2. De plus $G = (1, -1, 1)$, et ainsi on a $\dim(F) + \dim(G) = 3$. Puisque $F \cap G = \{0\}$, on a bien : $\mathbf{R}^3 = F \oplus G$.

2. Soit $\vec{X} = (x, y, z) \in \mathbf{R}^3$, on sait qu'il existe un unique couple $(\vec{f}, \vec{g}) \in F \times G$ tel que : $\vec{X} = \vec{f} + \vec{g}$. En posant $\vec{f} = (x_1, y_1, z_1)$ et $\vec{g} = (x_2, y_2, z_2)$, on a :

$$\begin{cases} x = x_1 + x_2 \\ y = y_1 + y_2 \\ z = z_1 + z_2 \\ x_1 + y_1 + z_1 = 0 \\ x_2 = -y_2 \\ x_2 = z_2 \end{cases} \iff \begin{cases} x_2 = x + y + z \\ z_2 = x + y + z \\ y_2 = -x - y - z \\ x_1 = x - x_2 = -y - z \\ y_1 = y - y_2 = x + 2y + z \\ z_1 = z - z_2 = -x - y \end{cases}$$

La projection de (x, y, z) sur F est le vecteur $\vec{f} = (-y - z, x + 2y + z, -x - y)$.

3. Notons P la matrice dans la base canonique de la projection sur F parallèlement à G . On vient de prouver que pour tout $(x, y, z) \in \mathbf{R}^3$, on a :

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

Cette égalité est vraie pour tout $(x, y, z) \in \mathbf{R}^3$ donc

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

Notons Q la matrice de la projection sur G parallèlement à F . Puisque $P + Q = I_3$, on a :

$$Q = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

■ Comment déterminer si une matrice carrée est inversible

□ **Méthode 3.8.** — ► Une **première piste** est de démontrer qu'elle possède un inverse à gauche ou à droite.

► Une **deuxième piste** est d'utiliser le fait que A est la matrice d'un endomorphisme ϕ dans la base canonique. Il suffit alors de trouver ϕ^{-1} et d'exprimer alors sa matrice dans la base canonique. C'est A^{-1} . Cette méthode donne explicitement l'inverse.

► Une **troisième piste** est de résoudre un système :

- si $AX = 0$ a une solution $X \neq 0$, A n'est pas inversible;
- si $AX = B$ n'a pas de solution ou alors n'est pas unique, alors A n'est pas inversible;
- si $AX = B$ a une solution unique, alors A est inversible. De plus l'unique solution X_0 du système vérifie $X_0 = A^{-1}B$ et cela permet de déterminer A^{-1} .

► Une **quatrième piste** est de transformer la matrice en utilisant des opérations élémentaires sur les lignes ou sur les colonnes et aboutir à une matrice triangulaire supérieure ou inférieure (on peut arriver à I_n mais ce n'est pas nécessaire) et conclure quant au rang de la matrice de départ qui est le même que celui de la matrice d'arrivée. On retiendra en particulier que si l'on peut transformer A par opérations élémentaires en une matrice triangulaire supérieure à termes diagonaux tous non nuls, son rang est n et la famille de vecteurs associée à cette matrice est une base de \mathbf{K}^n .

► Une **cinquième piste** est d'utiliser la **méthode de Gauss-Jordan** :

- on écrit d'abord la matrice à p lignes et $2p$ colonnes schématisée par $c|cAI_p$, le symbole $|$ étant juste une barrière virtuelle qui permet de distinguer les deux parties importantes de la matrice;
- on transforme cette matrice par opérations élémentaires **sur les lignes** (sans se soucier de la barrière virtuelle) pour aboutir à la matrice schématisée par $c|cI_pA^{-1}$.

► Une **sixième piste** est de partir de : $a_0I_p + a_1A + \dots + a_nA^n = O_p$, avec $a_0 \neq 0$ et $n \in \mathbf{N}^*$. Ce type d'égalité est fournie par l'énoncé. Alors :

$$A \times \left(-\frac{a_1}{a_0}I_p - \dots - \frac{a_n}{a_0}A^{n-1} \right) = I_p \text{ et } \left(-\frac{a_1}{a_0}I_p - \dots - \frac{a_n}{a_0}A^{n-1} \right) \times A = I_p,$$

ce qui prouve que A est inversible et que $A^{-1} = \left(-\frac{a_1}{a_0}I_p - \dots - \frac{a_n}{a_0}A^{n-1} \right)$.

► Une **septième piste** est de calculer le **déterminant** de la matrice A et on sait que A est inversible si et seulement si $\det(A)$ est non nul. Cette piste est évidemment à mettre à l'écart si vous n'avez pas encore vu les déterminants.

Remarque : pour privilégier telle ou telle piste, tout dépend en fait de ce que l'on veut. Si l'on cherche l'expression de A^{-1} , on pourra par exemple résoudre complètement $AX = B$,

chercher B tel que $AB = I_n$ ou encore utiliser une relation du type $a_0 I_p + a_1 A + \dots + a_n A^n = O_p$, si elle est fournie !

Si l'on cherche simplement à montrer que A n'est pas inversible, on peut par exemple chercher $X \neq 0$ tel que $AX = 0$.

Exemple : Soit B une matrice de $\mathcal{M}_n(\mathbf{K})$, nilpotente d'ordre p , c'est-à-dire telle que :

$$B^p = 0 \text{ et } B^{p-1} \neq 0.$$

Montrer que $I_n - B$ est inversible et trouver cet inverse.

Ici, posons $C = I_n + B + \dots + B^{p-1}$.

Par analogie avec les réels et les sommes de suites géométriques : $(I_n - B)C = I_n - B^p = I_n$.

Donc C est l'inverse de $I_n - B$. On a montré au passage que $I_n - B$ est inversible (car il possède un inverse!). Retenez cet exemple car il est très classique.

■ Comment modifier la matrice d'une application linéaire en changeant les bases

Soit $\mathcal{B} = \{e_1, \dots, e_p\}$ une base de E , $\mathcal{C} = \{f_1, \dots, f_n\}$ une base de F et ϕ une application linéaire de E dans F avec $\dim E = p$ et $\dim F = n$. Considérons deux nouvelles bases, $\mathcal{B}' = \{e'_1, \dots, e'_p\}$ de E et $\mathcal{C}' = \{f'_1, \dots, f'_n\}$ de F . On part de la matrice $M_{\mathcal{B}, \mathcal{C}}(\phi)$ de ϕ par rapport aux bases \mathcal{B} et \mathcal{C} et le but du jeu c'est de trouver la matrice $M_{\mathcal{B}', \mathcal{C}'}(\phi)$ qui est la matrice de ϕ , l'espace vectoriel E étant rapporté à la base \mathcal{B}' et l'espace vectoriel F étant rapporté à la base \mathcal{C}' .

■ Méthode 3.9.— [1] On peut déterminer les matrices de passage de la base \mathcal{B} à la base \mathcal{B}' et de la base \mathcal{C} à la base \mathcal{C}' . On rappelle que leurs colonnes sont les images des vecteurs de la première base, exprimés dans la seconde. On note $P_{\mathcal{B}}^{\mathcal{B}'}$ et $P_{\mathcal{C}}^{\mathcal{C}'}$ les deux matrices de passage. On calcule ensuite $(P_{\mathcal{C}}^{\mathcal{C}'})^{-1}$ par une méthode d'inversion de matrice. On calcule enfin le produit $(P_{\mathcal{C}}^{\mathcal{C}'})^{-1} M_{\mathcal{B}, \mathcal{C}}(\phi) P_{\mathcal{B}}^{\mathcal{B}'}$.

Cette méthode inclut le cas particulier important où ϕ est un endomorphisme (dans ce cas, on prend en général : $\mathcal{B} = \mathcal{C}$ et $\mathcal{B}' = \mathcal{C}'$).

Exemple : soit la matrice $M_{\mathcal{B}}(\phi) = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Cette matrice représente l'endomorphisme ϕ ,

l'espace vectoriel \mathbf{R}^3 étant rapporté à sa base canonique $\mathcal{B} = \{e_1, e_2, e_3\}$.

Nous allons maintenant considérer la nouvelle base $\mathcal{B}' = \{e_1 + e_2 + e_3, -e_1 + e_2, -e_1 + e_3\}$.

On calcule : $P_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ et $(P_{\mathcal{B}}^{\mathcal{B}'})^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$.

On a : $M_{\mathcal{B}'}(\phi) = (P_{\mathcal{B}}^{\mathcal{B}'})^{-1} M_{\mathcal{B}}(\phi) P_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

Remarque : il existe une méthode parfois plus rapide. On part de la remarque que pour tout j variant de 1 à p , la colonne C_j de $M_{\mathcal{B},\mathcal{C}}(\phi)$ représente l'image de e_j exprimé dans \mathcal{C} et que la colonne C_j de $M_{\mathcal{B}',\mathcal{C}'}(\phi)$ représente l'image de e'_j exprimé dans \mathcal{C}' . Il suffit donc de savoir exprimer directement l'image de e'_j qui est une combinaison linéaire des images des vecteurs e_1, \dots, e_p dans \mathcal{C}' . Il faut faire cela bien entendu à partir des images de vecteurs e_j dans \mathcal{C} . Si cela est possible, on évite bien des calculs !

■ Comment calculer une puissance A^p d'une matrice donnée $A \in \mathcal{M}_n(\mathbf{K})$

(Ici pas de connaissance particulière de diagonalisation : on complète la méthode dans le chapitre sur la réduction des matrices.)

□ Méthode 3.10. — ► Une **première piste** est le calcul des premières puissances de A puis de conjecturer l'expression de A^p en fonction de p et enfin démontrer cette formule par récurrence.

► Une **deuxième piste** est l'utilisation de la **formule du binôme de Newton**, on peut procéder de la façon suivante :

- décomposer A sous forme $A = D + N$, où D est diagonale et N ayant des puissances simples (par exemple nilpotente);
- vérifier que D et N commutent;
- appliquer la formule du **binôme de Newton**.

► Une **troisième piste** est d'utiliser une relation du type $a_0 I_p + a_1 A + \dots + a_n A^n = O_p$, avec $n \leq p$. Dans ce cas, l'on pose $P = a_0 + a_1 X + \dots + a_n X^n$.

- On détermine le reste R de la division euclidienne de X^p par P ;
- En ce cas, on a alors $A^p = R(A)$.

► Une **quatrième piste** est de se souvenir que A représente la matrice d'un endomorphisme ϕ par rapport à la base canonique, notée \mathcal{B} de \mathbf{K}^n . Il suffit alors de trouver une base dans laquelle la matrice B de ϕ par rapport à une nouvelle base \mathcal{B}' de \mathbf{K}^n soit assez simple pour que l'on puisse en déduire B^p sans problème. On utilise alors l'implication : $A = P_{\mathcal{B}}^{\mathcal{B}'} B (P_{\mathcal{B}}^{\mathcal{B}'})^{-1} \Rightarrow A^p = P_{\mathcal{B}}^{\mathcal{B}'} B^p (P_{\mathcal{B}}^{\mathcal{B}'})^{-1}$ et le produit donne A^p .

On utilisera abondamment cette piste dans le chapitre sur la diagonalisation.

Exemple : Soit $J = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Calculer J^p , $p \in \mathbb{N}$.

Rapidement, on a : $J^2 = 3J$. Par récurrence, on montre rapidement que : $\forall p \geq 1, J^p = 3^{p-1}J$.

Exemple : Soit $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Calculer A^p par plusieurs méthodes.

1. Illustrons la deuxième piste de la **méthode 3.10**. On remarque que $A = J + I_3$, où J est la matrice précédente (quel heureux hasard!).

$$A^p = \sum_{k=1}^p \binom{p}{k} 3^{k-1} J + \binom{p}{0} I_3 = \frac{1}{3} [4^p - 1] J + I_3.$$

Le lecteur développe A^p .

2. Illustrons la troisième piste de la **méthode 3.10**. On calcule $A^2 = \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix}$ et on remarque que $A^2 = 5A - 4I_3$. Posons $P = X^2 - 5X + 4$. On écrit alors :

$$X^p = (X^2 - 5X + 4)Q + a_p X + b_p,$$

où a_p et b_p sont deux réels à calculer. Comme $X^2 - 5X + 4$ a pour racines 1 et 4, on a $1 = a_p + b_p$ et $4^p = 4a_p + b_p$ ce qui donne $a_p = (4^p - 1)/3$ et $b_p = 4(1 - 4^{p-1})/3$. On conclut que

$$A^p = \frac{1}{3}(4^p - 1)A + \frac{4}{3}(1 - 4^{p-1})I_3 = \frac{1}{3} \begin{pmatrix} 4^p + 2 & 4^p - 1 & 4^p - 1 \\ 4^p - 1 & 4^p + 2 & 4^p - 1 \\ 4^p - 1 & 4^p - 1 & 4^p + 2 \end{pmatrix}$$

Remarque : dans le cas où le polynôme P trouvé est de degré au plus 2, on peut aussi partir de $A^2 = \alpha_2 A + \beta_2 I_3$ et remarquer par récurrence que $A^p = \alpha_p A + \beta_p I_3$. Cette relation est vraie pour $p = 0$ en posant $\alpha_0 = 0$ et $\beta_0 = 1$ et pour $p = 1$ en posant $\alpha_1 = 1$ et $\beta_1 = 0$. On en déduit alors A^p en exprimant α_p et β_p en utilisant des relations de récurrence d'ordre 2. Ainsi, si l'on reprend :

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

on a $\alpha_2 = 5$ et $\beta_2 = -4$. On peut écrire $A^{p+1} = A^p A = (\alpha_p A + \beta_p I_3)A = \alpha_p A^2 + \beta_p A$

$$= \alpha_p(5A - 4I_3) + \beta_p A = (5\alpha_p + \beta_p)A - 4\alpha_p I_3 = \alpha_{p+1}A + \beta_{p+1}I_3$$

Il reste à identifier $5\alpha_p + \beta_p = \alpha_{p+1}$ et $-4\alpha_p = \beta_{p+1}$. Cela donne $\alpha_{p+2} - 5\alpha_{p+1} + 4\alpha_p = 0$ et en utilisant le cours sur les récurrences linéaires d'ordre 2 des suites (si vous l'avez fait), on remarque que α_p est une combinaison linéaire de 1 et de 4^p . En utilisant le fait que $\alpha_0 = 0$ et $\alpha_1 = 1$, on a $\alpha_p = (4^p - 1)/3$ et on retrouve ensuite β_p .

3. Illustrons la quatrième piste de la **méthode 3.10**.

La matrice A représente l'endomorphisme ϕ , l'espace vectoriel \mathbf{R}^3 étant rapporté à sa base canonique $\mathcal{B} = \{e_1, e_2, e_3\}$. Nous allons maintenant considérer la nouvelle base :

$$\mathcal{B}' = \{e_1 + e_2 + e_3, -e_1 + e_2, -e_1 + e_3\}$$

La matrice de passage et son inverse ont été calculés plus haut. Appelons B la matrice de ϕ par rapport à la base \mathcal{B}' . Elle a été calculée elle aussi plus haut. On écrit pour tout p entier,

$$A = PBP^{-1} \Rightarrow A^p = PB^pP^{-1}.$$

On retrouve alors :

$$B^p = \begin{pmatrix} 4^p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow A^p = \frac{1}{3} \begin{pmatrix} 4^p + 2 & 4^p - 1 & 4^p - 1 \\ 4^p - 1 & 4^p + 2 & 4^p - 1 \\ 4^p - 1 & 4^p - 1 & 4^p + 2 \end{pmatrix}$$

■ Comment calculer la trace d'un endomorphisme u de E

□ Méthode 3.11.— On choisit une base de E et on détermine la matrice de u dans cette base. La trace de u est alors la somme des coefficients diagonaux de cette matrice.

Exemple : Soit $A \in \mathcal{M}_n(\mathbf{K})$ une matrice non nulle. Déterminer la trace de l'endomorphisme u de $\mathcal{M}_n(\mathbf{K})$ défini par :

$$\forall M \in \mathcal{M}_n(\mathbf{K}), \quad u(M) = \text{Tr}(M)A.$$

D'après le théorème de la base incomplète, il existe $n^2 - 1$ matrices, E_2, \dots, E_{n^2} , telles que la famille $\mathcal{B} = (A, E_2, \dots, E_{n^2})$ soit une base de $\mathcal{M}_n(\mathbf{K})$. Dans cette base, la matrice de u est de la forme suivante (seule la première ligne est non nulle) :

$$X = \begin{pmatrix} \text{Tr}(A) & \text{Tr}(E_2) & \cdots & \text{Tr}(E_{n^2}) \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

On en déduit que $\text{Tr}(u) = \text{Tr}(A)$.

■ Comment appliquer le théorème du rang à une matrice $A \in \mathcal{M}_{n,p}(\mathbf{K})$

□ Méthode 3.12.— On applique ce théorème à l'application linéaire $u \in \mathcal{L}(\mathbf{K}^p, \mathbf{K}^n)$ canoniquement associée à la matrice A , ce qui donne :

$$p = \dim(\ker(A)) + \text{Rg}(A)$$

où p est le nombre de **colonnes** de la matrice A .

Exemple : Soit a et b deux réels et $A = \begin{pmatrix} 1 & 1 \\ a & b \\ a^2 & b^2 \end{pmatrix}$. Déterminer la dimension de $\ker(A)$.

▷ Si $a = b$, la matrice A est de rang 1. Le théorème du rang nous donne alors :

$$2 = \text{Rg}(A) + \dim(\ker(A)) \quad \text{soit} \quad \dim(\ker(A)) = 1.$$

▷ Si $a \neq b$, la matrice A est de rang 2 et par conséquent $\dim(\ker(A)) = 0$.

Chapitre 4

Révision : Probabilités finies

■ Objectifs

■ Les incontournables :

- ▶ identifier et modéliser l'expérience aléatoire ;
- ▶ traduire ensemblistement l'énoncé ;
- ▷ utiliser les propriétés classiques de probabilités, notamment reconnaître une probabilité conditionnelle ;
- ▷ mettre en fonction les formules des probabilités totales, des probabilités composées et de Bayes et savoir laquelle utiliser ;
- ▷ faire fonctionner l'indépendance de deux événements, l'indépendance mutuelle d'une famille finie d'événements.

■ Et plus si affinités ...

- ▶ Étendre la formule des probabilités totales à une suite d'événements deux à deux incompatibles dont la somme des probabilités vaut 1 ;
- ▶ utiliser la formule des probabilités totales pour obtenir une relation de récurrence entre les termes d'une suite de probabilités.

Résumé de cours

■ Rappels sur les dénombrements

On rappelle que $0! = 1$ et $n! = n \times (n-1) \times \dots \times 1$ pour n entier non nul.

Proposition 4.1.— Nombre d'applications entre deux ensembles finis —. Soit E et F deux ensembles finis de cardinaux respectifs n et p , alors l'ensemble des applications de E dans F est de cardinal p^n , c'est-à-dire $\text{Card } F^{\text{Card } E}$.

Proposition 4.2.— Nombre de parties d'un ensemble fini —. Soit E un ensemble fini de cardinal n alors l'ensemble $\mathcal{P}(E)$ des parties de E est de cardinal 2^n , c'est-à-dire $2^{\text{Card } E}$.

Définition : On appelle **p -liste** (ou **p -uplet**) de **E** ensemble fini tout élément de E^p de la forme (x_1, x_2, \dots, x_p) , où x_1, \dots, x_p sont des éléments de E .

Remarque : L'ordre des éléments compte et il peut y avoir des répétitions.

Proposition 4.3.— Le nombre de p -listes de E de cardinal n est n^p .

Proposition 4.4.— La nombre de p -listes **d'éléments distincts** de E est $\frac{n!}{(n-p)!}$.

Proposition 4.5.— Nombre d'injections entre deux ensembles finis —. Le nombre d'injections d'un ensemble de cardinal p dans un ensemble de cardinal n est $\frac{n!}{(n-p)!}$.

Définition : Soit E un ensemble fini de cardinal n , on appelle **p -combinaison** de E toute partie de E à p éléments.

Proposition 4.6.— Le nombre de **p -combinaison** de E de cardinal n est $\binom{n}{p} = \frac{n!}{(n-p)!p!}$.

On rappelle la formule du binôme de Newton :

$$(x+y)^n = \sum_{p=0}^n \binom{n}{p} x^p y^{n-p} = \sum_{p=0}^n \binom{n}{p} x^{n-p} y^p$$

On a les formules :

$$\forall p \in \llbracket 1, n \rrbracket, \binom{n}{n-p} = \binom{n}{p} \text{ et } \forall p \in \llbracket 1, n-1 \rrbracket, \binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$$

Mise en œuvre : exercice 01, exercice 02, exercice 03, exercice 04.

■ Le langage des probabilités

Définition : Une *expérience aléatoire* est une expérience dont on ne peut prédire avec certitude le résultat. L'étude d'une expérience aléatoire commence par la description des résultats possibles, appelés *éventualités*. L'ensemble des résultats possibles est appelé **univers des possibles**. On le note en général Ω .

Exemples : 1. Si nous lançons une pièce, nous avons deux résultats possibles : Pile ou Face. L'expérience aléatoire est « Lancer d'une pièce ». Donc $\Omega = \{Pile, Face\}$.

2. Si nous lançons un dé cubique dont les faces sont numérotées de 1 à 6, nous avons six résultats possibles. L'expérience aléatoire est « Lancer d'un dé ».

Donc $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Remarque : Attention, pour certains types d'expérience, on peut avoir plusieurs univers possibles. Il faut donc que l'énoncé soit le plus précis possible. Sinon, pour résoudre le problème, il vous faudra faire des hypothèses supplémentaires. Par exemple, quand nous tirons une carte d'un jeu de 52 cartes, nous pouvons nous intéresser au rang de la carte dans le jeu et définir l'univers comme l'ensemble des entiers de 1 à 52. On peut s'intéresser aussi à la couleur de la carte obtenue et définir l'univers comme étant $\{Pique, Coeur, Carreau, Trèfle\}$.

Définition : Un *événement aléatoire* est un événement qui peut se produire ou non, suivant le résultat de l'expérience aléatoire. On le représente par l'ensemble des éventualités qui le réalisent. Il s'agit donc d'une partie A de Ω , on écrit : $A \in \mathcal{P}(\Omega)$, où $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω . On dira que **l'événement A est réalisé** si le résultat ω de cette expérience est élément de A . Enfin, si $A = \{\omega\}$, donc si A n'a qu'un élément, on dit que A est un **événement élémentaire**.

Exemples : 1. Pour l'expérience aléatoire « Lancer d'une pièce », $\Omega = \{Pile, Face\}$.

On sait que $\mathcal{P}(\Omega) = \{\emptyset, \{Pile\}, \{Face\}, \Omega\}$.

Si A est l'événement : « obtenir Pile », alors $A = \{Pile\}$.

2. Pour l'expérience aléatoire « Lancer d'un dé », $\Omega = \{1, 2, 3, 4, 5, 6\}$. Dans ce cas, $\mathcal{P}(\Omega)$ est un ensemble qui possède $6! = 720$ éléments. On ne va pas l'écrire en extension ! Par contre si A est l'événement : « faire un résultat pair » alors $A = \{2, 4, 6\}$.

Si A est l'événement : « faire un résultat supérieur à 7 » alors $A = \emptyset$.

Ω est appelé **l'événement certain** et \emptyset est appelé **l'événement impossible**.

L'identification entre événements aléatoires et parties de Ω permet d'utiliser les opérations ensemblistes classiques :

- La **disjonction ou réunion** de A et B est $A \cup B$.

On rappelle que $\omega \in A \cup B$ si et seulement si $\omega \in A$ ou $\omega \in B$. Attention le « ou » n'est pas exclusif.

- La **conjonction ou intersection** de A et B est $A \cap B$.

On rappelle que $\omega \in A \cap B$ si et seulement si $\omega \in A$ et $\omega \in B$.

- La **négation ou complémentarité** de A est \bar{A} .

On rappelle que $\omega \in \bar{A}$ si et seulement si $\omega \in \Omega$ et $\omega \notin A$.

- La **différence** de A par B est $A - B = A \setminus B = A \cap \bar{B}$.

On rappelle que $\omega \in A \setminus B$ si et seulement si $\omega \in A$ et $\omega \notin B$.

- **Inclusion** de A dans B est $A \subset B$.

On remarque que cela peut aussi se traduire par $A \Rightarrow B$.

Enfin, ces opérations interagissent entre-elles de la manière suivante. Si $(A, B, C) \in (\mathcal{P}(\Omega))^3$,

Commutativité : $A \cap B = B \cap A$, $A \cup B = B \cup A$.

Associativité : $A \cap (B \cap C) = (A \cap B) \cap C$, $A \cup (B \cup C) = (A \cup B) \cup C$.

Distributivité : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Lois de Morgan : $\overline{A \cap B} = \overline{A} \cup \overline{B}$, $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

On peut remarquer que l'associativité permet de définir les parties $\bigcup_{i \in I} A_i$ et $\bigcap_{i \in I} A_i$, où I est un ensemble fini (en général des entiers).

Définition : Deux événements A et B sont dits **incompatibles** lorsque $A \cap B = \emptyset$.

Définition : Une famille finie $(A_i)_{i \in I}$ d'événements forme un **système complet d'événements** si pour tout $(i, j) \in I^2$ avec $i \neq j$, $A_i \cap A_j = \emptyset$ et si $\bigcup_{i \in I} A_i = \Omega$.

Exemples : 1) Un **système complet d'événements** classique est $\{A, \overline{A}\}$ en posant $A \in \mathcal{P}(\Omega)$ fixé. En effet, $A \cap \overline{A} = \emptyset$ et $A \cup \overline{A} = \Omega$.

2) Un autre **système complet d'événements** classique est $\{A \cap B, A \cap \overline{B}, \overline{A} \cap B, \overline{A} \cap \overline{B}\}$ en posant $(A, B) \in \mathcal{P}(\Omega)^2$ fixé.

En effet, $(A \cap B) \cup (A \cap \overline{B}) \cup (\overline{A} \cap B) \cup (\overline{A} \cap \overline{B}) = \Omega$ et ces quatre ensembles sont bien disjoints deux à deux.

Mise en œuvre : Beaucoup d'exercices utilisent les opérations sur les parties d'un ensemble. On les fait plus loin.

■ Espace probabilisé fini

Dans toute la suite, Ω désigne un **ensemble fini**.

Définition : On appelle **probabilité** sur Ω toute application $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$ qui vérifie les deux propriétés suivantes :

- $P(\Omega) = 1$.
- Si A et B sont incompatibles alors $P(A \cup B) = P(A) + P(B)$.

Vocabulaire : Le triplet $(\Omega, \mathcal{P}(\Omega), P)$ est appelé **espace probabilisé fini**. On dit que si $A \in \mathcal{P}(\Omega)$, alors $P(A)$ est la probabilité de A .

On remarque que les deux assertions qui définissent une probabilité suffisent pour fonctionner. Voir le théorème suivant.

Théorème 4.7.— Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini, et si A et B sont des événements alors :

$$P(\emptyset) = 0, P(\overline{A}) = 1 - P(A), P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

$$P(A \setminus B) = P(A) - P(A \cap B), A \subset B \Rightarrow P(A) \leq P(B).$$

Théorème 4.8.— **Formule d'additivité finie.** Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini et $(A_i)_{i \in \llbracket 1, n \rrbracket}$ une famille d'événements deux à deux incompatibles alors :

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

En particulier, si $\Omega = \{\omega_1, \dots, \omega_n\}$, une probabilité sur Ω est entièrement définie si l'on connaît les probabilités des événements élémentaires $\{\omega_i\}$ pour tout $i \in \llbracket 1, n \rrbracket$.

En effet, si $A = \{\omega_i\}_{i \in I}$, où $I \subset \llbracket 1, n \rrbracket$ alors :

$$P(A) = \sum_{i \in I} P(\{\omega_i\}).$$

On en déduit la proposition suivante.

Proposition 4.9.— Soit $\Omega = \{\omega_1, \dots, \omega_n\}$, un ensemble fini de cardinal $n \in \mathbb{N}^*$ et (p_1, \dots, p_n) un n -uplet de nombres réels **positifs**. Pour qu'il existe une probabilité $P : \Omega \rightarrow [0, 1]$ telle que pour tout $i \in \llbracket 1, n \rrbracket$, $P(\{\omega_i\}) = p_i$, il faut et il suffit que :

$$p_1 + p_2 + \dots + p_n = 1.$$

On peut remarquer que comme $p_i \geq 0$ pour tout $i \in \llbracket 1, n \rrbracket$ et $p_1 + \dots + p_n = 1$ alors pour tout $i \in \llbracket 1, n \rrbracket$, $p_i \in [0, 1]$.

Remarque : Souvent dans un exercice, on demande de calculer les probabilités des événements élémentaires, et avant de continuer dans l'exercice, il faut vérifier que toutes ces probabilités sont bien entre 0 et 1 et que leur somme soit égale à 1.

On lance un dé cubique truqué tel que la probabilité de faire 6 est $1/2$ et celle de faire i est $1/10$ pour tout i variant de 1 à 5. Alors si l'on pose $p_i = P(\{i\})$ pour tout $i \in \llbracket 1, 6 \rrbracket$,

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{2} = 1.$$

C'est cohérent.

Puis si A est l'événement : « faire un résultat pair », alors $P(A) = p_2 + p_4 + p_6 = \frac{7}{10}$.

Mise en œuvre : exercice 05, exercice 06, exercice 07.

Cas particulier important : la probabilité uniforme.

Imaginons le lancer d'une pièce non truquée alors $P(\{Pile\}) = P(\{Face\}) = \frac{1}{2}$.

Imaginons le lancer d'un dé cubique non truqué alors $\forall i \in \llbracket 1, 6 \rrbracket, P(\{i\}) = \frac{1}{6}$.

Dans tous ces cas courants, on peut calculer la probabilité de n'importe quel événement par le rapport de deux nombres.

Théorème 4.10.— Soit $A \in \mathcal{P}(\Omega)$ une partie de Ω et on suppose que tous les événements élémentaires ont la même probabilité $\frac{1}{\text{Card}(\Omega)}$. Alors :

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}.$$

On peut interpréter $\text{Card}(A)$ comme le *nombre de cas favorables* et $\text{Card}(\Omega)$ comme le *nombre de cas possibles*.

Mise en œuvre : exercice 08, exercice 09

■ Conditionnement

Proposition 4.11.— Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini et B est un événement tel que $P(B) > 0$ (on dit que B est *non négligeable* alors l'application P_B définie sur $\mathcal{P}(\Omega)$ par l'égalité :

$$\forall A \in \mathcal{P}(\Omega), P_B(A) = \frac{P(A \cap B)}{P(B)}$$

est une probabilité sur Ω .

On a bien : $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = 1$, et si A et C sont deux événements incompatibles,

$$P_B(A \cup C) = \frac{P((A \cup C) \cap B)}{P(B)} = \frac{P((A \cap B) \cup (C \cap B))}{P(B)} = \frac{P(A \cap B)}{P(B)} + \frac{P(C \cap B)}{P(B)}.$$

Et on a bien la seconde assertion de la définition d'une probabilité :

$$P_B(A \cup C) = P_B(A) + P_B(C).$$

(L'avant-dernière égalité est justifiée par le fait que $A \cap B$ et $C \cap B$ sont encore deux événements incompatibles.)

Remarque : On a alors aussi, pour tout $A \in \mathcal{P}(\Omega)$, $P_B(\overline{A}) = 1 - P_B(A)$.

Définition : Si A et B sont deux événements tels que $P(B) > 0$, la *probabilité conditionnelle de A sachant B* est :

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

Notation : on note aussi couramment cette probabilité $P(A|B)$ (à ne pas confondre avec $P(A \setminus B)$).

Remarque : Attention, une erreur classique est de confondre A sachant B et $A \cap B$. Par exemple, dans le lancer d'un dé cubique non pipé, si A est : « obtenir un chiffre pair » et B est : « obtenir un nombre inférieur strictement à 4 » alors la probabilité que l'on obtienne à la fois un chiffre pair et un nombre inférieur strictement à 4 est $P(A \cap B) = 1/6$ car $A \cap B = \{2\}$. Et la probabilité que l'on obtienne un chiffre pair sachant que l'on a obtenu un nombre inférieur strictement à 4 est (car $B = \{1, 2, 3\}$),

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$

Mise en œuvre : exercice 11.

Proposition 4.12. — Formules des probabilités composées —.

► $\forall (A, B) \in (\mathcal{P}(\Omega))^2$,
si $P(B) > 0$, $P(A \cap B) = P(B)P_B(A)$ et si $P(A) > 0$, $P(A \cap B) = P(A)P_A(B)$.
► Si, pour $n \geq 2$, (A_1, A_2, \dots, A_n) est une famille d'événements de l'univers Ω telle que l'on ait :
 $P\left(\bigcap_{i=1}^{n-1} A_i\right) \neq \emptyset$, alors :

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \dots P_{\bigcap_{i=1}^{n-1} A_i}(A_n).$$

Remarques : 1. La probabilité de réalisation simultanée de deux événements est le produit de la probabilité de réalisation de l'un d'entre eux par la probabilité de réalisation de l'autre sachant que le premier est réalisé.

2. Lorsque A_n est le dernier événement d'une suite décroissante $A_n \subset A_{n-1} \subset \dots \subset A_1$, la formule précédente donne un moyen souvent efficace pour calculer $P(A_n)$:

$$P(A_n) = P(A_1)P_{A_1}(A_2)P_{A_2}(A_3) \dots P_{A_{n-1}}(A_n).$$

Exemple : Trois urnes contiennent des boules blanches et noires. L'urne U_1 contient 2 boules blanches et 3 boules noires, l'urne U_2 contient 4 boules blanches et 2 boules noires, l'urne U_3 contient 6 boules blanches et 1 boule noire. Toutes les boules ont la même probabilité d'être piochées. On effectue trois tirages successifs selon le protocole suivant :

- on tire une boule de U_1 , on note sa couleur, on met cette boule dans U_2 ;
- on tire une boule de U_2 , on note sa couleur, on met cette boule dans U_3 ;
- on tire une boule de U_3 et on note sa couleur.

Calculons la probabilité pour que les trois boules dont on a noté la couleur soient de la même couleur.

Notons B_i (respectivement N_i) l'événement « la $i^{\text{ème}}$ boule tirée et notée est blanche (respectivement noire) ». On cherche donc la probabilité de :

$$(N_1 \cap N_2 \cap N_3) \cup (B_1 \cap B_2 \cap B_3).$$

On remarque que les événements $(N_1 \cap N_2 \cap N_3)$ et $(B_1 \cap B_2 \cap B_3)$ sont incompatibles. Donc :

$$P((N_1 \cap N_2 \cap N_3) \cup (B_1 \cap B_2 \cap B_3)) = P(N_1 \cap N_2 \cap N_3) + P(B_1 \cap B_2 \cap B_3).$$

D'après la formule des probabilités composées,

$$P(N_1 \cap N_2 \cap N_3) = P(N_1) \times P_{N_1}(N_2) \times P_{N_1 \cap N_2}(N_3) = \frac{3}{5} \times \frac{3}{7} \times \frac{2}{8} = \frac{9}{140}.$$

De même,

$$P(B_1 \cap B_2 \cap B_3) = P(B_1) \times P_{B_1}(B_2) \times P_{B_1 \cap B_2}(B_3) = \frac{2}{5} \times \frac{5}{7} \times \frac{7}{8} = \frac{35}{140}.$$

$$\text{Ainsi, } P((N_1 \cap N_2 \cap N_3) \cup (B_1 \cap B_2 \cap B_3)) = \frac{9}{140} + \frac{35}{140} = \frac{44}{140} = \frac{11}{35}.$$

Mise en œuvre : exercice 10

Proposition 4.13.— Formules des probabilités totales —. Si $(A_i)_{i \in \llbracket 1, n \rrbracket}$ est un système complet d'événements de l'univers Ω , alors :

$$\text{pour tout événement } B \text{ de } \mathcal{P}(\Omega), P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P_{A_i}(B)P(A_i),$$

avec la convention $P_{A_i}(B)P_{A_i}(B) = 0$ lorsque $P(A_i) = 0$.

Remarque : Un cas fréquent d'application de ce théorème est celui où le système complet d'événements est (A, \bar{A}) . On a alors, pour tout B de $\mathcal{P}(\Omega)$,

$$P(B) = P_A(B)P(A) + P_{\bar{A}}(B)P(\bar{A}).$$

Exemple : Donnons un exemple d'utilisation de la formule des probabilités totales avec un système complet d'événements de quatre éléments.

Quatre urnes U_1, U_2, U_3 et U_4 contiennent des boules blanches et noires. U_1 contient 4 boules blanches et 1 boule noire, U_2 contient 3 boules blanches et 2 boules noires, U_3 contient 2 boules blanches et 3 boules noires, U_4 contient 1 boule blanche et 4 boules noires. Pour une urne donnée, chaque boule a la même probabilité d'être tirée. On effectue un tirage selon le protocole suivant :

- On choisit une des quatre urnes. On suppose que la probabilité de choisir l'urne U_i est $\frac{i}{10}$. (Donc il n'y a pas équiprobabilité du choix de l'urne.)
- On pioche au hasard une boule dans l'urne qui a été choisie.

On veut déterminer la probabilité d'obtenir une boule blanche.

Notons pour tout $i \in \llbracket 1, 4 \rrbracket$, l'événement U_i : « le tirage de la boule se fait dans l'urne U_i ».

Notons B l'événement : « la boule tirée est blanche ».

Il est clair que (U_1, U_2, U_3, U_4) forme un système complet d'événements. D'après la formule des probabilités totales,

$$P(B) = P(U_1)P_{U_1}(B) + P(U_2)P_{U_2}(B) + P(U_3)P_{U_3}(B) + P(U_4)P_{U_4}(B).$$

$$\text{On trouve : } P(B) = \frac{1}{50} (4 + 6 + 6 + 1) = \frac{17}{50}.$$

Proposition 4.14. — Formules de Bayes (ou probabilité des causes) —.

► Si A et B sont deux événements de probabilités non nulles, alors :

$$P_B(A) = \frac{P(A)P_A(B)}{P(B)}.$$

► Si $(A_j)_{j \in \llbracket 1, n \rrbracket}$ est un système complet d'événements de probabilités non nulles et si B est un événement de probabilité non nulle, alors :

$$\forall j \in \llbracket 1, n \rrbracket, P_B(A_j) = \frac{P(A_j)P_{A_j}(B)}{\sum_{i=1}^n P(A_i)P_{A_i}(B)}.$$

Exemple : Une forêt se compose de trois types d'arbres : 30% sont des chênes, 50% sont des peupliers et 20% des hêtres. Suite à une tempête, une maladie se déclare et touche 10% des chênes, 4% des peupliers et 25% des hêtres. On étudie un arbre.

On note M l'événement : « l'arbre étudié est malade » puis C l'événement « l'arbre étudié est un chêne », Q l'événement « l'arbre étudié est un peuplier » et H l'événement « l'arbre étudié est un hêtre ».

1. Calculons dans l'ordre $P_C(M)$ (probabilité que l'arbre étudié soit malade sachant que c'est un chêne), $P_H(M)$ (probabilité que l'arbre étudié soit malade sachant que c'est un hêtre), $P_Q(M)$ (probabilité que l'arbre étudié soit malade sachant que c'est un peuplier).

Pour commencer, d'après l'énoncé, $P(C) = 0.3$, $P(Q) = 0.5$ et $P(H) = 0.2$.

Toujours d'après l'énoncé, $P_C(M) = 0.1$, $P_Q(M) = 0.04$, $P_H(M) = 0.25$.

2. Calculons $P_M(C)$ qui est la probabilité que l'arbre étudié soit un chêne sachant qu'il est malade.

On applique la formule de Thomas Bayes :

$$P_M(C) = \frac{P_C(M)P(C)}{P(M)} = \frac{P_C(M)P(C)}{P_C(M)P(C) + P_H(M)P(H) + P_Q(M)P(Q)}.$$

Cela donne :

$$P_M(C) = \frac{0.1 \times 0.3}{0.1 \times 0.3 + 0.04 \times 0.5 + 0.25 \times 0.2} = \frac{0.1 \times 0.3}{0.1} = 0.3.$$

On peut aussi sur notre lancée calculer $P_M(H) = 0.5$ et $P_M(Q) = 0.2$.

Mise en œuvre : exercice 12, exercice 13, exercice 14

■ Indépendance

Définition : Deux événements A et B sont **indépendants** lorsque : $P(A \cap B) = P(A)P(B)$.

Remarques : 1. Dans ce cas, si, par exemple, $P(B) > 0$, alors $P_B(A) = \frac{P(A \cap B)}{P(B)} = P(A)$, ce qui signifie que la réalisation de B n'influe pas sur la probabilité de réalisation de A . Réciproquement, si $P(B) > 0$ et si $P_B(A) = P(A)$, alors $P(A \cap B) = P(A)P(B)$ et A et B sont indépendants.

2. Ne pas confondre indépendance et incompatibilité : deux événements incompatibles sont en fait généralement dépendants, puisque la réalisation de l'un interdit la réalisation de l'autre.

3. On montre facilement (c'est l'exercice 15) que A et B sont indépendants si et seulement si A et \overline{B} ou \overline{A} et B ou \overline{A} et \overline{B} sont indépendants.

Définition : Les événements d'une famille finie $(A_i)_{1 \leq i \leq n}$ d'événements sont :

► **deux à deux indépendants** lorsque :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i \neq j \Rightarrow P(A_i \cap A_j) = P(A_i) P(A_j)) ;$$

► **indépendants** lorsque, pour toute partie J non vide de $\llbracket 1, n \rrbracket$:

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i).$$

Comment repérer l'indépendance dans un énoncé ?

Parfois, l'indépendance est une hypothèse écrite de l'énoncé donc dans ce cas, on applique directement la formule. Parfois, elle est sous-entendu par l'expérience décrite. Les exemples classiques sont des lancers successifs d'une pièce (truquée ou non). La pièce n'a pas de mémoire. Donc si A_i est l'événement : « le $i^{\text{ème}}$ lancer donne Pile » alors tous les événements A_1, \dots, A_n sont indépendants. Il en de même si l'on lance successivement toujours un dé et de même si l'on tire des boules avec remise dans une urne de composition connue etc.

Exemple : On effectue une suite de lancers d'une pièce (parfaitement équilibrée) à Pile ou Face. On s'intéresse à la probabilité que Face apparaisse pour la première fois au quatrième lancer. On note P_k l'événement : « On a obtenu Pile au $k^{\text{ème}}$ lancer » et F_k l'événement : « On a obtenu Face au $k^{\text{ème}}$ lancer ». L'événement A qui est : « On a obtenu Face pour la première fois au quatrième lancer » peut se traduire par

$$A = P_1 \cap P_2 \cap P_3 \cap F_4.$$

Les résultats obtenus aux différents lancers étant indépendants, il s'ensuit que :

$$P(A) = P(P_1) \times P(P_2) \times P(P_3) \times P(F_4) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}.$$

Remarque : Des événements indépendants sont deux à deux indépendants, mais la réciproque est fausse lorsque leur nombre est supérieur ou égal à 3.

En effet, considérons l'expérience aléatoire qui consiste à lancer un dé deux fois et les événements A , B , C définis respectivement par « le premier chiffre est pair, » « le deuxième chiffre est impair, » « la somme des chiffres est paire ».

On a : $P(A) = P(B) = P(C) = \frac{1}{2}$, $P(A \cap B) = P(B \cap C) = P(C \cap A) = \frac{9}{36} = \frac{1}{4}$ et ces événements sont deux à deux indépendants, mais $P(A \cap B \cap C) = P(\emptyset) \neq P(A)P(B)P(C)$ et ils ne sont pas indépendants.

Mise en œuvre : exercice 15, exercice 16

Chapitre 5

Variables aléatoires discrètes finies

■ Objectifs

■ Les incontournables :

- ▶ savoir déterminer la loi de probabilité d'une v.a.r discrète.
- ▶ savoir calculer l'espérance la variance d'une v.a.r et leurs propriétés;
- ▶ étudier les couples de variables aléatoires discrètes, en particulier indépendantes et calculer leurs lois marginales;
- ▶ connaître les lois discrètes usuelles : loi de Bernoulli, loi Binomiale, loi uniforme et leur espérance et variance;
- ▶ connaître et savoir utiliser l'inégalité de Bienaymé-Tchebycheff;
- ▶ savoir utiliser le théorème de transfert.

■ Et plus si affinités ...

- ▶ avoir quelques idées sur les suites de variables aléatoires mutuellement indépendantes notamment sur la somme de n variables aléatoires mutuellement indépendantes ;
- ▶ savoir déterminer la loi de probabilité d'une somme de deux v.a.r indépendantes;
- ▶ savoir déterminer la loi du minimum ou du maximum de deux v.a.r.

Résumé de cours

Ici $(\Omega, \mathcal{P}(\Omega), P)$ est un espace probabilisé fini.

■ Variable aléatoire discrète et loi de probabilité associée

Définition : Une variable aléatoire discrète X (notée v.a.r ou v.a.r.d) sur $(\Omega, \mathcal{P}(\Omega))$ est une application définie sur Ω dont l'image, $X(\Omega)$, est une partie finie de \mathbb{R} .

L'ensemble des valeurs prises par X sera noté $X(\Omega) = \{x_1, x_2, \dots, x_n\}$.

Remarques : 1) Une variable aléatoire n'est pas plus une variable qu'elle n'est aléatoire, puisqu'il s'agit d'une application fixée entre deux ensembles eux-mêmes fixés. Ce vocable prend son origine dans l'histoire des probabilités et on a tout intérêt à le considérer comme le néologisme insécable « variable aléatoire ».

2) L'ensemble d'arrivée est souvent une partie de \mathbb{N} .

Exemples : Prenons l'exemple de lancer de dés.

Si l'expérience est le lancer d'un dé cubique, si l'on note X le numéro de la face obtenue, X est une v.a.r.d et $X(\Omega) = \llbracket 1, 6 \rrbracket$.

Si l'expérience est le lancer de deux dés cubiques, si l'on note X les numéros obtenus, X est une v.a.r.d et $X(\Omega) = \llbracket 1, 6 \rrbracket^2$.

Si l'expérience est le lancer d'un dé cubique, si l'on note X le nombre minimal de lancers pour obtenir un premier 6, X est une v.a.r.d et $X(\Omega) = \mathbb{N}^*$. On note que ce dernier exemple sort des clous ici car $X(\Omega)$ est de cardinal infini. Les v.a.r.d à univers infini dénombrable sont étudiés plus tard.

Notations :

- Si $x \in X(\Omega)$, l'image réciproque par X de x est $X^{-1}(x) = \{\omega \in \Omega \mid X(\omega) = x\}$. On la note abusivement $\{X = x\}$ ou $(X = x)$. La définition ci-dessus impose que $X^{-1}(x) \in \mathcal{P}(\Omega)$.
- Si U est une partie de $X(\Omega)$, l'image réciproque de U par X est

$$X^{-1}(U) = \{\omega \in \Omega \mid X(\omega) \in U\} = \bigcup_{x \in U} (X = x).$$

On la note $\{X \in U\}$ ou $(X \in U)$.

- Si $x \in X(\Omega)$, $P(X = x)$ est la probabilité de $(X \in \{x\}) = \{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{P}(\Omega)$.
- Si U est une partie de $X(\Omega)$, on : $P(X \in U) = \sum_{x \in U} P(X = x)$. Il suffit donc de connaître tous

les $P(X = x)$ pour $x \in X(\Omega)$ pour pouvoir calculer, au moins de façon théorique, la probabilité de $(X \in U)$, pour n'importe quelle partie U de $X(\Omega)$. D'où la définition suivante :

Définition : Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ telle que $X(\Omega) = \{x_k\}_{k \in \llbracket 1, n \rrbracket}$.

La loi de probabilité de X est la suite $(p_k)_{k \in \llbracket 1, n \rrbracket}$, où, pour tout $k \in \llbracket 1, n \rrbracket$, $p_k = P(X = x_k)$.

Remarque : Alors : $\forall k \in \llbracket 1, n \rrbracket$, $p_k = P(X = x_k) \in [0, 1]$ et $\sum_{k=1}^n p_k = \sum_{k=1}^n P(X = x_k) = 1$.

Proposition 5.1.— Si une variable aléatoire discrète sur $(\Omega, \mathcal{P}(\Omega))$, X , prend ses valeurs dans $\{x_k; k \in \llbracket 1, n \rrbracket\}$, les x_k étant distincts, et si $(p_k)_{k \in \llbracket 1, n \rrbracket}$ est une suite de réels positifs vérifiant $\sum_{k=1}^n p_k = 1$, alors il existe une probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ telle que $\forall k \in \llbracket 1, n \rrbracket, P(X = x_k) = p_k$.

Exemple : Terminons ce paragraphe par un exemple de v.a.r et de loi de probabilité associée. On lance deux dés usuels non pipés et X est la v.a.r égale au plus grand des deux résultats obtenus. Déterminons la loi de X .

On a $X(\Omega) = \llbracket 1, 6 \rrbracket$ avec $\Omega = \{(i, j), i \in \llbracket 1, 6 \rrbracket, j \in \llbracket 1, 6 \rrbracket\}$.

On va pour tout $k \in \llbracket 1, 6 \rrbracket$, dénombrer le nombre de cas favorables qui est $\text{Card}(X = k)$ et diviser par le nombre de cas possibles qui est 36.

Pour commencer, l'événement $(X = 1)$ est $X^{-1}(\{1\}) = \{(1, 1)$. On a : $P(X = 1) = \frac{1}{36}$.

De même, l'événement $(X = 2)$ est $X^{-1}(\{2\}) = \{(1, 2), (2, 1), (2, 2)$. On a : $P(X = 2) = \frac{3}{36}$.

Puis, l'événement $(X = 3)$ est $X^{-1}(\{3\}) = \{(1, 3), (2, 3), (3, 3), (3, 2), (3, 1)$. On a : $P(X = 3) = \frac{5}{36}$.

Puis, l'événement $(X = 4)$ est $X^{-1}(\{4\})$ et $\text{Card } X^{-1}(\{4\}) = 7$. On a : $P(X = 4) = \frac{7}{36}$.

Puis, l'événement $(X = 5)$ est $X^{-1}(\{5\})$ et $\text{Card } X^{-1}(\{5\}) = 9$. On a : $P(X = 5) = \frac{9}{36}$.

Enfin, l'événement $(X = 6)$ est $X^{-1}(\{6\})$ et $\text{Card } X^{-1}(\{6\}) = 11$. On a : $P(X = 6) = \frac{11}{36}$.

On peut vérifier au passage que :

$$\sum_{k=1}^6 P(X = k) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{7}{36} + \frac{9}{36} + \frac{11}{36} = 1.$$

On a bien une loi de probabilité. Par exemple si A est l'événement « le plus grand résultat obtenu est inférieur ou égal à 2 » (on écrit $A = (X \leq 2)$) alors $P(A) = P(X = 1) + P(X = 2) = \frac{4}{36} = \frac{1}{9}$.

Transformée d'une v.a.r.d

Définition : Soit X une v.a.r.d sur $(\Omega, \mathcal{P}(\Omega), P)$ et $f : X(\Omega) \rightarrow \mathbf{R}$, une fonction définie sur $X(\Omega)$. L'application composée $Y : \Omega \rightarrow \mathbf{R}, \omega \mapsto f(X(\omega))$ est une v.a.r.d notée $Y = f(X)$.

La transformée la plus utilisée de X est X^2 . Utile pour la définition de la variance. Voir plus loin.

Exemple : Si X est une variable aléatoire à valeurs dans $\llbracket 1, 5 \rrbracket$ telle que $P(X = k) = \frac{1}{5}$ pour tout $k \in \llbracket 1, 5 \rrbracket$, déterminons la loi de $Y = (X - 3)^2$.

On commence par remarquer que $(X - 3)(\Omega) = \llbracket -2, 2 \rrbracket$ et $Y(\Omega) = \{0, 1, 4\}$.

L'événement $(Y = 0)$ est aussi $(X = 3)$ et donc $P(Y = 0) = \frac{1}{5}$.

L'événement $(Y = 1)$ est aussi $(X = 2) \cup (X = 4)$ et donc $P(Y = 1) = \frac{2}{5}$.

Enfin, l'événement $(Y = 4)$ est aussi $(X = 1) \cup (X = 5)$ et donc $P(Y = 4) = \frac{2}{5}$.

Mise en œuvre : l'exercice 01 et l'exercice 02.

■ Espérance et variance d'une variable aléatoire discrète finie

Définition : L'espérance de la variable aléatoire **réelle** discrète X à valeurs dans $\{x_1, x_2, \dots, x_n\}$ est le réel

$$E(X) = \sum_{k=1}^n x_k P(X = x_k).$$

Exemple : Calculons l'espérance de X la variable aléatoire à valeurs dans $\llbracket 1, 5 \rrbracket$ telle que :

$P(X = k) = \frac{1}{5}$ pour tout $k \in \llbracket 1, 5 \rrbracket$.

On a : $E(X) = \sum_{k=1}^5 k \frac{1}{5} = \frac{1}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5} + \frac{5}{5} = \frac{15}{5} = 3$.

Remarques : 1. L'espérance peut toujours être considérée comme une moyenne pondérée par les probabilités, en effet :

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

2. La variable aléatoire X est **centrée** si $E(X) = 0$.

Proposition 5.2. — Théorème de transfert — Si X est une variable aléatoire et f une application à valeurs réelles définie sur l'image $X(\Omega) = \{x_k, k \in \llbracket 1, n \rrbracket\}$ de X , alors

$$E(f(X)) = \sum_{k=1}^n P(X = x_k) f(x_k).$$

Exemple : Reprenons l'exemple de X une variable aléatoire à valeurs dans $\llbracket 1, 5 \rrbracket$ telle que :

$P(X = k) = \frac{1}{5}$ pour tout $k \in \llbracket 1, 5 \rrbracket$, calculons $E((X - 3)^2)$ avec la formule de transfert. On a :

$$E((X - 3)^2) = \sum_{k=1}^5 P(X = k) f(k),$$

où $f(k) = (k - 3)^2$. On a : $E((X - 3)^2) = \frac{1}{5} ((1 - 3)^2 + (2 - 3)^2 + (3 - 3)^2 + (4 - 3)^2 + (5 - 3)^2)$.

On trouve $E((X - 3)^2) = \frac{10}{5} = 2$.

On peut remarquer que connaissant la loi de $Y = (X - 3)^2$, le calcul direct avec la définition de l'espérance donne $E(Y) = 0 \times \frac{1}{5} + 1 \times \frac{2}{5} + 4 \times \frac{2}{5} = 2$. C'est le même résultat.

Proposition 5.3.— Soit X une v.a.r et $a \in \mathbb{R}$: $E(aX) = aE(X)$, $E(a) = a$, $E(X+a) = E(X) + a$.

La preuve est simple, à faire.

Théorème-Définition 5.4.— **Formule de Huyghens-Koenig** —. Si X est une v.a.r.d, on appelle *variance de X* le réel

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

L'écart-type de X est le réel noté $\sigma(X)$ défini par : $\sigma(X) = \sqrt{V(X)}$.

On pourra prouver l'égalité entre les deux formes plus loin.

Remarque : Ainsi plus la variance est grande, plus les x_k sont éloignés de la moyenne $E(X)$. Plus la variance est proche de 0, plus les x_k sont ramassés autour de $E(X)$. On va essayer de mieux quantifier tout cela un peu plus bas avec l'inégalité de Bienaymé-Tchebychev.

Exemple : Calculons la variance de X la variable aléatoire à valeurs dans $\llbracket 1, 5 \rrbracket$ telle que :

$$P(X = k) = \frac{1}{5} \text{ pour tout } k \in \llbracket 1, 5 \rrbracket.$$

$$\text{On a : } E(X^2) = \sum_{k=1}^5 k^2 \frac{1}{5} = \frac{1}{5} + \frac{4}{5} + \frac{9}{5} + \frac{16}{5} + \frac{25}{5} = \frac{55}{5} = 11.$$

$$\text{Comme } E(X) = 3, V(X) = E(X^2) - E^2(X) = 11 - 9 = 2.$$

Proposition 5.5.—

$$\forall (a, b) \in \mathbb{R}^2, V(aX + b) = a^2 V(X).$$

Remarque : Une variable aléatoire admettant une variance est *centrée réduite* lorsque $E(X) = 0$ et $V(X) = 1$. Par exemple, si $V(X) \neq 0$, $(X - E(X))/\sigma(X)$ est centrée réduite. Pour démontrer ce résultat, on utilise les deux dernières propositions.

Mise en œuvre : exercice 03.

■ Exemples usuels de lois discrètes finies

1. La loi certaine.

Ici $X(\Omega) = \{a\}$.

On dit que X suit la **loi certaine** de valeur $\{a\}$ si et seulement si :

$$P(X = a) = 1.$$

Et on a rapidement : $E(X) = a$, $V(X) = 0$. (À interpréter!)

2. La loi de Bernoulli.

Ici $X(\Omega) = \llbracket 0, 1 \rrbracket$ et $p \in]0, 1[$.

On dit que X suit la **loi de Bernoulli de paramètre p** si et seulement si :

$$P(X = 0) = q = 1 - p \text{ et } P(X = 1) = p.$$

Notation : $X \sim \mathcal{B}(p)$. Et on a rapidement : $E(X) = p$, $V(X) = pq$.

Remarque : C'est la loi du succès. Par exemple le lancer d'une pièce équilibrée. Soit X la v.a.r qui donne 0 si l'on a Face et 1 si l'on a Pile. Alors $X \hookrightarrow \mathcal{B}(\frac{1}{2})$. On peut remarquer que si la pièce est truquée, on a encore une loi de Bernoulli mais son paramètre n'est plus $1/2$.

3. La loi Binomiale.

C'est la loi du nombre de succès en n tentatives. Comme pour la loi de Bernoulli, on considère une expérience aléatoire e qui a deux issues : le succès S ou l'échec E .

On suppose que la probabilité que S se réalise est p . Et donc E se réalise avec une probabilité $q = 1 - p$. On répète n fois dans les mêmes conditions et de manière indépendantes les unes des autres, l'expérience e . On note X la v.a.r.d égale au nombre de succès obtenus en n tentatives.

Finalement, $X(\Omega) = \llbracket 0, n \rrbracket$ et

$$\forall k \in \llbracket 0, n \rrbracket, P(X = k) = \binom{n}{k} p^k q^{n-k}. \text{ On note } X \sim \mathcal{B}(n, p)$$

Et on montre que : $E(X) = np$, $V(X) = npq$.

Remarque : Une loi binomiale peut s'interpréter comme une répétition de lois de Bernoulli indépendantes. Nous le verrons plus loin.

Exemple : 1. On effectue n lancers indépendants les uns des autres d'une pièce de monnaie pour laquelle la probabilité d'obtenir Pile est p .

On note X le nombre de Piles obtenus.

2. Soit une urne \mathcal{U} ayant une proportion p de boules blanches et $q = 1 - p$ de boules noires. On prélève successivement n boules de \mathcal{U} en notant la couleur de la boule tirée et en la remettant dans l'urne. Si X est la v.a.r.d égale au nombre de boules blanches tirées au total, alors X suit une loi binomiale $\mathcal{B}(n, p)$.

3. On note X la v.a.r.d égale au nombre de rois obtenus lorsqu'on tire successivement 5 cartes avec remise dans un jeu de 32 cartes. Précisons la loi de X .

On vérifie rapidement que X suit une loi binomiale $\mathcal{B}(5, \frac{1}{8})$.

En effet, si k est le nombre de rois obtenus en n tirages, les rois sont apparus k fois et il y a $\binom{5}{k}$ possibilités d'apparition. Un de ces événements se produit avec une probabilité de $\left(\frac{1}{8}\right)^k \times \left(\frac{7}{8}\right)^{5-k}$ car un roi a une chance sur huit d'apparaître.

Donc : $P(X = k) = \binom{5}{k} \left(\frac{1}{8}\right)^k \times \left(\frac{7}{8}\right)^{5-k}$.

4. La loi uniforme sur $\llbracket 1, n \rrbracket$.

Ici $X(\Omega) = \llbracket 1, n \rrbracket$.

On dit que X suit la **loi uniforme** sur $\llbracket 1, n \rrbracket$ si et seulement si :

$$\forall k \in \llbracket 1, n \rrbracket, P(X = k) = \frac{1}{n}.$$

Notation : $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$.

Et : $E(X) = \frac{n+1}{2}$, $V(X) = \frac{n^2-1}{12}$.

Mise en œuvre : exercice 04, exercice 05, exercice 06 et exercice 07.

■ Couples de variables aléatoires discrètes

Définition : Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega))$, (X, Y) , application de Ω dans $X(\Omega) \times Y(\Omega)$, est un **couple de variables aléatoires discrètes** sur $(\Omega, \mathcal{P}(\Omega))$.

Un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega))$ est aussi une variable aléatoire discrète sur $(\Omega, \mathcal{P}(\Omega))$ à valeurs dans $X(\Omega) \times Y(\Omega)$, puisque le produit cartésien de deux ensembles dénombrables est dénombrable. Pour tout élément (x_i, y_j) de $X(\Omega) \times Y(\Omega)$, l'événement $((X, Y) = (x_i, y_j))$ désigne $\{\omega \in \Omega \mid X(\omega) = x_i \text{ et } Y(\omega) = y_j\}$ et se note souvent $(X = x_i, Y = y_j)$.

Définition : La loi de probabilité du couple (X, Y) de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega), P)$, considérée comme une variable aléatoire discrète, est la **loi conjointe** de (X, Y) ; elle est déterminée par la donnée, pour tout élément (x_i, y_j) de $X(\Omega) \times Y(\Omega)$, de $P(X = x_i, Y = y_j)$. Les lois de X et Y sont les **lois marginales** de (X, Y) .

Exemple : Une entreprise emploie 10 salariés, 3 ingénieurs, 2 secrétaires et 5 techniciens supérieurs. La direction désigne au hasard trois personnes pour participer à la commission "hygiène et sécurité". X désigne la variable aléatoire donnant le nombre de secrétaires et Y la variable aléatoire donnant le nombre d'ingénieurs participant à cette commission.

Ω est l'ensemble des parties à 3 éléments de l'ensemble des salariés; son cardinal est $\binom{10}{3} = 120$.

Le couple de variables aléatoires (X, Y) est à valeurs dans $\llbracket 0, 2 \rrbracket \times \llbracket 0, 3 \rrbracket$.

Pour décrire la loi conjointe de (X, Y) , on a recours à un tableau à double entrée permettant de noter la probabilité de chaque événement élémentaire.

$$\forall (i, j) \in \llbracket 0, 2 \rrbracket \times \llbracket 0, 3 \rrbracket, P((X, Y) = (i, j)) = \frac{\binom{2}{i} \binom{3}{j} \binom{5}{3-i-j}}{120}.$$

	X=0	X=1	X=2	Loi de Y
Y=0	10/120	20/120	5/120	35/120
Y=1	30/120	30/120	3/120	63/120
Y=2	15/120	6/120	0	21/120
Y=3	1/120	0	0	1/120
Loi de X	56/120	56/120	8/120	1

Les lois marginales de X et de Y sont explicitées dans la dernière ligne pour X et la dernière colonne pour Y

Remarque : Si la loi de (X, Y) est connue, les lois marginales le sont.

Ainsi : $\forall x \in X(\Omega)$,

$$P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y).$$

$$\text{De même : } \forall y \in Y(\Omega), P(Y = y) = \sum_{x \in X(\Omega)} P(X = x, Y = y).$$

En revanche, réciproquement, la connaissance des lois marginales de X et Y ne suffit pas pour définir la loi conjointe de (X, Y) , ce que met en évidence l'exemple ci-dessous.

Exemple : Considérons deux couples de variables aléatoires (X, Y) et (X', Y') prenant tous les deux les quatre valeurs $(0, 0)$, $(0, 1)$, $(1, 0)$ et $(1, 1)$.

Supposons que (X, Y) suive la loi uniforme :

$P((X, Y) = (0, 0)) = P((X, Y) = (1, 0)) = P((X, Y) = (0, 1)) = P((X, Y) = (1, 1)) = \frac{1}{4}$,
alors les lois marginales de X et Y sont : $P(X = 0) = P(X = 1) = \frac{1}{2}$ et $P(Y = 0) = P(Y = 1) = \frac{1}{2}$.

Supposons que (X', Y') suive la loi définie par : $P((X', Y') = (0, 0)) = \frac{1}{8}$,
 $P((X', Y') = (1, 0)) = \frac{3}{8}$, $P((X', Y') = (0, 1)) = \frac{3}{8}$, $P((X', Y') = (1, 1)) = \frac{1}{8}$, alors les lois marginales de X' et Y' sont : $P(X' = 0) = P(X' = 1) = \frac{1}{2}$ et $P(Y' = 0) = P(Y' = 1) = \frac{1}{2}$.

Les lois marginales sont identiques, alors que les lois conjointes correspondantes ne le sont pas.

Définition : Soit x un élément de $X(\Omega)$ tel que $P(X = x) > 0$, la **loi conditionnelle de Y sachant que $(X = x)$** est la probabilité P_x sur $Y(\Omega)$ définie par :

$$\forall y_j \in Y(\Omega), P_x(y_j) = P_{(X=x)}(Y = y_j) = \frac{P(X = x, Y = y_j)}{P(X = x)}.$$

Mise en œuvre : exercice 08, exercice 09, exercice 10.

■ Variables aléatoires discrètes indépendantes

Définition : X et Y étant deux variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$, (X, Y) est un **couple de variables aléatoires indépendantes** si, et seulement si,

$$\forall (x_i, y_j) \in X(\Omega) \times Y(\Omega), P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j).$$

Exemple : On lance quatre fois une pièce de monnaie. On note X le nombre de piles obtenus lors des deux premiers lancers et Y le nombre de piles obtenus lors des deux derniers lancers.

Intuitivement, les valeurs de Y ne sont pas « influencées » par les résultats de X : les variables X et Y semblent indépendantes. Vérifions-le par le calcul.

	X=0	X=1	X=2	Loi de Y
Y=0	1/16	1/8	1/16	1/4
Y=1	1/8	1/4	1/8	1/2
Y=2	1/16	1/8	1/16	1/4
Loi de X	1/4	1/2	1/4	1

La probabilité de chaque événement élémentaire de $X(\Omega) \times Y(\Omega)$ est égale au produit des probabilités marginales correspondantes : (X, Y) est un couple de variables aléatoires indépendantes.

Exemple : On lance **trois** fois une pièce de monnaie. On note X le nombre de piles obtenus lors des deux premiers lancers et Y le nombre de piles obtenus lors des deux derniers lancers.

Intuitivement, les variables X et Y ne semblent pas indépendantes. Vérifions-le par le calcul.

La loi conjointe se résume dans le tableau suivant :

	X=0	X=1	X=2	Loi de Y
Y=0	1/8	1/8	0	1/4
Y=1	1/8	1/4	1/8	1/2
Y=2	0	1/8	1/8	1/4
Loi de Y	1/4	1/2	1/4	1

$P((X, Y) = (0, 0)) \neq P(X = 0)P(Y = 0)$ ou $P((X, Y) = (0, 2)) \neq P(X = 0)P(Y = 2)$: (X, Y) est un couple de variables aléatoires non indépendantes.

Mise en œuvre : exercice 11.

Proposition 5.6.— Si X et Y sont indépendantes, alors, pour toute partie A de $X(\Omega)$ et toute partie B de $Y(\Omega)$, on a :

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

Proposition 5.7.— Si X et Y sont deux variables aléatoires indépendantes alors, pour toutes fonctions f et g , les variables $f(X)$ et $g(Y)$ sont indépendantes.

Définition : Soit $(X_i)_{i \in \llbracket 1, n \rrbracket}$ une famille finie de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega), P)$, ces variables aléatoires sont **indépendantes** si, et seulement si,

$$\forall (x_1, x_2, \dots, x_n) \in \prod_{i \in \llbracket 1, n \rrbracket} X_i(\Omega), P((X_i)_{i \in \llbracket 1, n \rrbracket} = (x_i)_{i \in \llbracket 1, n \rrbracket}) = \prod_{i=1}^n P(X_i = x_i).$$

Proposition 5.8.— Si $(X_i)_{i \in \llbracket 1, n \rrbracket}$ est une famille finie de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega), P)$ indépendantes, alors quel que soit $(A_i)_{i \in \llbracket 1, n \rrbracket} \in \prod_{i=1}^n \mathcal{P}(X_i(\Omega))$, les événements $(X_i \in A_i)_{i \in \llbracket 1, n \rrbracket}$ sont indépendants.

Remarque : Si $(X_i)_{i \in \llbracket 1, n \rrbracket}$ est une famille finie de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega), P)$ indépendantes, alors toutes ses sous-familles sont des familles finies de variables aléatoires discrètes indépendantes.

Définition : Soit $(X_n)_{n \in \mathbf{N}}$ une suite de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega), P)$, ces variables aléatoires sont :

- ▷ **deux à deux indépendantes** si, et seulement si, $\forall (i, j) \in \mathbf{N}^2, (i \neq j \Rightarrow X_i \text{ et } X_j \text{ indépendantes})$.
- ▷ **indépendantes** si, et seulement si, pour toute partie finie I de \mathbf{N} , $(X_i)_{i \in I}$ est une famille finie de variables aléatoires discrètes indépendantes.

Remarques : 1. Il suffit de vérifier que, pour tout $n \in \mathbf{N}$, les variables X_0, X_1, \dots, X_n sont indépendantes. En effet, si c'est le cas, toute sous-famille finie de $(X_n)_{n \in \mathbf{N}}$ est contenue dans une sous-famille de la forme (X_0, X_1, \dots, X_n) , donc est une famille de variables aléatoires discrètes indépendantes, d'après la remarque ci-dessus.

2. Une suite de variables aléatoires discrètes indépendantes est un modèle pour décrire une succession d'épreuves dont les résultats sont indépendants. En particulier, la répétition indépendante d'une même épreuve est modélisée par une suite de variables aléatoires indépendantes suivant la même loi.

Par exemple, un jeu de pile ou face infini est modélisé par $(X_n)_{n \in \mathbf{N}}$, suite de variables aléatoires indépendantes suivant chacune la loi de Bernoulli $\mathcal{B}(\frac{1}{2})$.

Extension

Si X_1, \dots, X_n sont n variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega))$, (X_1, \dots, X_n) , application de Ω dans $X_1(\Omega) \times \dots \times X_n(\Omega)$, est un **n -uplet de variables aléatoires discrètes** sur $(\Omega, \mathcal{P}(\Omega))$.

Un n -uplet de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega))$ est aussi une variable aléatoire discrète sur $(\Omega, \mathcal{P}(\Omega))$ à valeurs dans $X_1(\Omega) \times \dots \times X_n(\Omega)$, puisque le produit cartésien de n ensembles dénombrables est dénombrable. Pour tout élément (x_1, \dots, x_n) de $X_1(\Omega) \times \dots \times X_n(\Omega)$, l'événement $((X_1, \dots, X_n) = (x_1, \dots, x_n))$ (ie $(X_1 = x_1, \dots, X_n = x_n)$) désigne :

$$\{\omega \in \Omega \mid \forall i \in \llbracket 1, n \rrbracket, X_i(\omega) = x_i\}.$$

Définition : La loi de probabilité du n -uplet (X_1, \dots, X_n) de variables aléatoires discrètes sur $(\Omega, \mathcal{P}(\Omega), P)$, considérée comme une variable aléatoire discrète, est la **loi conjointe de** (X_1, \dots, X_n) ; elle est déterminée par la donnée, pour tout élément (x_1, \dots, x_n) de $X_1(\Omega) \times \dots \times X_n(\Omega)$, de $P(X_1 = x_1, \dots, X_n = x_n)$.

Les lois de X_1, \dots, X_n sont les **lois marginales de** (X_1, \dots, X_n) .

■ Sommes et produits de variables aléatoires discrètes

Le cas le plus courant est la somme de 2 variables aléatoires discrètes.

Considérons X et Y deux v.a.r.d définis sur le même univers fini Ω .

La première idée qui vient à l'esprit est de calculer $P(X+Y = k)$ « à la main » pour $k \in (X+Y)(\Omega)$.

$$P(X+Y = k) = \sum_{j \in X(\Omega)} P((X = j) \cap (Y = k - j)).$$

Dans le cas très fréquent, où X et Y sont à valeurs dans \mathbb{N} ,

$$P(X+Y = k) = \sum_{j=0}^n P((X = j) \cap (Y = k - j)).$$

Une question est alors primordiale. Les v.a.r.d dont on fait la somme, sont-elles indépendantes ou non ?

Si l'on peut répondre à cette question de manière positive, alors :

$$P(X+Y = k) = \sum_{j \in X(\Omega)} P(X = j) \times P(Y = k - j).$$

Et si X et Y sont à valeurs dans \mathbb{N} ,

$$P(X+Y = k) = \sum_{j=0}^n P(X = j) \times P(Y = k - j).$$

Exemple : X et Y étant deux variables aléatoires indépendantes suivant des lois binomiales $\mathcal{B}(m, p)$ et $\mathcal{B}(n, p)$ alors $X+Y$ suit une loi binomiale $\mathcal{B}(m+n, p)$.

Mise en œuvre : exercice 12.

Remarque : Si X et Y ne sont pas indépendantes alors on ne peut pas en général expliciter facilement X et Y . Il faut alors se ramener à des formules du genre :

$$P((X = j) \cap (Y = k - j)) = P_{(X=j)}(Y = k - j) P(X = j).$$

Par contre, on peut obtenir $E(X+Y)$ facilement (voir en dessous).

Proposition 5.9. — Linéarité, positivité et croissance —. Soit X et Y deux variables aléatoires réelles discrètes d'espérances finies.

- $\forall \lambda \in \mathbf{R}, E(\lambda X + Y) = \lambda E(X) + E(Y)$: l'espérance est linéaire.
- Si $X \geq 0$, c'est-à-dire si X ne prend que des valeurs positives, on a : $E(X) \geq 0$.
- En conséquence, si : $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$, alors $E(X) \leq E(Y)$: l'espérance est croissante.

Pour la variance $V(X + Y)$, on verra un peu plus loin. Pour cela, il faut introduire **le produit de deux variables aléatoires discrètes**. Ici X et Y sont encore définis sur le même univers Ω fini, et si $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_p\}$, alors $XY(\Omega)$ est constitué de tous les produits distincts $x_i y_j$ avec $i \in \llbracket 1, n \rrbracket$ et $j \in \llbracket 1, p \rrbracket$.

Exemple : Considérons X et Y deux v.a.r.d qui suivent une loi de Bernoulli de paramètre respectif p_1 et p_2 . Déterminons la loi de XY .

Déjà, comme $X(\Omega) = \{0, 1\} = Y(\Omega)$, $XY(\Omega) = \{0, 1\}$. Donc XY suit encore une loi de Bernoulli. De plus son paramètre est : $p = P((X = 1) \cap (Y = 1))$.

Dans le cas où X et Y sont indépendantes, $p = p_1 p_2$.

Proposition 5.10.— Espérance du produit de deux variables aléatoires indépendantes —.

Si X et Y sont deux variables aléatoires discrètes finies avec $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_p\}$, alors :

$$E(XY) = \sum_{i=1}^n \sum_{j=1}^p x_i y_j P(X = x_i, Y = y_j).$$

Si de plus X et Y sont indépendantes alors : $E(XY) = E(X)E(Y)$.

Remarque : La réciproque est fautive en générale, comme le montre l'exemple suivant : considérons le couple de variables aléatoires (X, Y) dont la loi conjointe est donnée par le tableau suivant :

	X=1	X=-1	Loi de Y
Y=-1	1/8	1/8	1/4
Y=0	3/8	1/8	1/2
Y=1	1/8	1/8	1/4
Loi de X	5/8	3/8	1

La loi de la variable aléatoire XY est donc : $P(XY = 1) = \frac{1}{4}$, $P(XY = -1) = \frac{1}{4}$, $P(XY = 0) = \frac{1}{2}$. D'où : $E(XY) = 0$.

Comme $E(X) = \frac{1}{4}$ et $E(Y) = 0$, on a bien : $E(XY) = E(X)E(Y)$.

Pourtant, $P(X = 1, Y = 1) = \frac{1}{8} \neq P(X = 1)P(Y = 1) = \frac{5}{32}$, donc X et Y ne sont pas indépendantes.

Définition : Si X et Y sont deux variables discrètes réelles, leur **covariance** est :

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

Proposition 5.11.— $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.

Remarque : Si X et Y sont deux variables aléatoires discrètes indépendantes, alors $\text{Cov}(X, Y) = 0$, mais la réciproque est fautive.

Mise en œuvre : exercice 13, exercice 14.

Remarque : Pour tout couple de variables aléatoires discrètes (X, Y) définis sur le même univers, on peut définir une multitude de nouvelles variables aléatoires discrètes de la forme $\phi(X, Y)$. En dehors de $X + Y$ et de XY , on peut rencontrer $X - Y$, $\min(X, Y)$, $\max(X, Y)$ etc.

Mise en œuvre : exercice 12, exercice 15, exercice 16, exercice 17.

Extension à la somme de n variables aléatoires réelles discrètes.

On peut construire $X_1 + X_2 + \dots + X_n$, où pour tout $k \in \llbracket 1, n \rrbracket$, X_k est définie sur l'univers fini Ω .

Proposition 5.12.— Si X_1, \dots, X_n suivent la même loi de Bernoulli de paramètre p et sont mutuellement indépendantes alors $X_1 + X_2 + \dots + X_n$ suit la loi binomiale $\mathcal{B}(n, p)$.

Ce résultat permet de montrer qu'une variable aléatoire X suit une loi binomiale en mettant X sous la forme d'une somme de variables aléatoires de Bernoulli de même paramètre et mutuellement indépendantes.

Proposition 5.13.— **Espérance d'une somme finie de variables aléatoires** —.

Soit $(X_i)_{1 \leq i \leq n}$ une famille finie de variables aléatoires réelles discrètes définies sur Ω :

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i).$$

Ainsi si $X_k \sim \mathcal{B}(p)$ pour tout $k \in \llbracket 1, n \rrbracket$, alors on sait que $E(X_k) = p$ et comme on a $E(X_1 + \dots + X_n) = np$, on retrouve l'espérance de la loi $\mathcal{B}(n, p)$.

Proposition 5.14.— **Variance d'une somme de deux variables aléatoires discrètes**—.

Soient X et Y deux variables aléatoires réelles discrètes définies sur Ω :

$$V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y).$$

Si de plus, X et Y sont **indépendantes**, $V(X + Y) = V(X) + V(Y)$

Remarque : Sachez que la formule précédente s'étend. Dans un concours, vous pourriez avoir à l'utiliser (à condition qu'on vous l'admette). **La variance de la somme de n variables aléatoires indépendantes deux à deux est la somme des variances de ces n variables aléatoires.** Ainsi si $X_k \sim \mathcal{B}(p)$ pour tout $k \in \llbracket 1, n \rrbracket$, et en supposant X_1, \dots, X_n indépendantes donc deux à deux indépendantes, alors on sait que $V(X_k) = pq$ et donc $V(X_1 + \dots + X_n) = npq$. On retrouve la variance de la loi $\mathcal{B}(n, p)$.

Mise en œuvre : exercice 20.

■ Inégalité de Bienaymé-Tchebychev

Corollaire 5.15.— Inégalité de Bienaymé-Tchebychev —.

Soit X une variable aléatoire réelle discrète, alors :

$$\forall \varepsilon > 0, P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}.$$

Remarques : 1. L'inégalité de Bienaymé-Tchebychev permet de comprendre ce que mesure la variance : pour $\varepsilon > 0$ fixé, la probabilité que l'écart entre X et $E(X)$ soit supérieur à ε est d'autant plus petite que $V(X)$ est faible. La variance donne donc une indication de la dispersion de X autour de son espérance, c'est-à-dire sa plus ou moins forte tendance à s'écarter de sa moyenne. L'écart-type, qui mesure aussi la dispersion de X , présente l'intérêt de s'exprimer dans la même unité que les valeurs prises par la variable aléatoire X .

2. On passe souvent à l'événement contraire : $\forall \varepsilon > 0, P(|X - E(X)| < \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}$.

Exemple : Illustrons l'inégalité de Bienaymé-Tchebychev en utilisant la loi binomiale.

On réalise 400 fois la même expérience, dont la probabilité de succès est $0.8 = \frac{4}{5}$.

On suppose que les 400 expériences sont indépendantes.

Soit X le nombre de succès obtenus.

Quelle loi suit X ? Calculer $E(X)$ puis donner un minorant de $P(300 < X < 340)$.

Déjà, on remarque que : $X \hookrightarrow \mathcal{B}\left(400, \frac{4}{5}\right) \Rightarrow E(X) = 320$.

$$(300 < X < 340) = (|X - E(X)| < 20)$$

est l'événement contraire de $(|X - E(X)| \geq 20)$.

D'après l'inégalité de Bienaymé-Tchebychev,

$$P(|X - E(X)| \geq 20) \leq \frac{V(X)}{20^2} = \frac{64}{400}.$$

Donc : $P(300 < X < 340) = 1 - P(|X - E(X)| \geq 20) \geq \frac{21}{25}$.

Remarque : De l'inégalité de Bienaymé-Tchebychev, on aboutit directement à la loi faible des grands nombres qui est la lecture de l'inégalité de Bienaymé-Tchebychev appliquée non pas à une v.a.r X mais à une somme de n variables aléatoires.

Mise en œuvre : exercice 18 et exercice 19.

Chapitre 6

Déterminants

■ Objectifs

■ Les incontournables :

- ▶ Savoir calculer un déterminant :
 - ▷ en utilisant les transformations élémentaires du pivot de Gauss ;
 - ▷ par développement par rapport à une ligne ou une colonne ;
 - ▷ grâce à des relations de récurrence.
- ▶ Savoir calculer le déterminant d'une matrice triangulaire par blocs.
- ▶ Connaître et utiliser la notion de déterminant d'un endomorphisme.
- ▶ Connaître les propriétés de l'application déterminant de $\mathcal{M}_n(\mathbf{K})$ dans \mathbf{K} .

■ Et plus si affinités ...

- ▶ Savoir calculer, reconnaître et utiliser les déterminants de Vandermonde.

Résumé de cours

■ Rappel : déterminant de deux vecteurs du plan

Définition : Dans le plan vectoriel \mathbf{R}^2 rapporté à une base orthonormée directe de deux vecteurs (\vec{i}, \vec{j}) , pour tous vecteurs **non nuls** \vec{u}_1 et \vec{u}_2 , on appelle **déterminant** des vecteurs \vec{u}_1 et \vec{u}_2 le nombre :

$$\text{Det}(\vec{u}_1, \vec{u}_2) = [\vec{u}_1, \vec{u}_2] = \|\vec{u}_1\| \cdot \|\vec{u}_2\| \cdot \sin \theta,$$

où θ est l'angle orienté (\vec{u}_1, \vec{u}_2) .

Dans le cas où l'un au moins des deux vecteurs \vec{u}_1 ou \vec{u}_2 est nul, $\text{Det}(\vec{u}_1, \vec{u}_2) = 0$.

Interprétation géométrique :

$\text{Det}(\vec{u}_1, \vec{u}_2)$ est l'aire algébrique du parallélogramme construit sur \vec{u}_1 et \vec{u}_2 .

À faire avec un dessin !

Il n'est parfois pas évident de calculer $\sin \theta$. On a la formule suivante (qui se démontre en passant par le plan complexe mais admise pour nous).

Proposition 6.1.— Dans toute base orthonormée directe du plan, si $\vec{u}_1(x_1, y_1)$ et $\vec{u}_2(x_2, y_2)$ alors

$$\text{Det}(\vec{u}_1, \vec{u}_2) = x_1 y_2 - x_2 y_1.$$

On note par ailleurs cette dernière quantité aussi $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$.

Remarques : 1) On peut étendre la formule au déterminant de deux vecteurs de \mathbf{C}^2 exprimés dans la base canonique de \mathbf{C}^2 .

2) On remarque rapidement que $\text{Det}(\vec{u}, \vec{v}) = -\text{Det}(\vec{v}, \vec{u})$.

3) On remarque aussi que $\text{Det}(\vec{u}, a \cdot \vec{u}) = 0$.

4) On remarque aussi que $\text{Det}(\vec{u} + a \cdot \vec{u}', \vec{v}) = \text{Det}(\vec{u}, \vec{v}) + a \text{Det}(\vec{u}', \vec{v})$.

On dit que le déterminant de deux vecteurs dans le plan est une **forme bilinéaire alternée**.

Proposition 6.2.— On a : $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$.

On dit qu'un déterminant est invariant par passage à la transposition.

■ Rappel : déterminant (ou produit mixte) de trois vecteurs de l'espace

Définition : Dans l'espace rapporté à une base orthonormée directe de trois vecteurs $(\vec{i}, \vec{j}, \vec{k})$, pour tous vecteurs \vec{u}_1, \vec{u}_2 et \vec{u}_3 , on appelle **déterminant** des trois vecteurs \vec{u}_1, \vec{u}_2 et \vec{u}_3 le nombre :

$$\text{Det}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = [\vec{u}_1, \vec{u}_2, \vec{u}_3] = (\vec{u}_1 \wedge \vec{u}_2) \cdot \vec{u}_3,$$

où \cdot désigne le produit scalaire entre deux vecteurs de l'espace et \wedge désigne le produit vectoriel de deux vecteurs de l'espace.

Remarque : On remarque que dans le cas où l'un au moins des trois vecteurs \vec{u}_1, \vec{u}_2 ou \vec{u}_3 est nul, $\text{Det}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = 0$.

Interprétation géométrique :

$\text{Det}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ peut s'interpréter comme le volume algébrique du parallélépipède construit sur \vec{u}_1 , \vec{u}_2 et \vec{u}_3 .

À faire avec un dessin !

Comme pour le déterminant de deux vecteurs dans le plan, on a une formule utilisant les composantes des trois vecteurs qui donne $\text{Det}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$. Cette formule utilise bien entendu celle qui donne le produit scalaire et le produit vectoriel.

Proposition 6.3.— Dans toute base orthonormée directe de l'espace, si $\vec{u}_1(x_1, y_1, z_1)$, $\vec{u}_2(x_2, y_2, z_2)$ et $\vec{u}_3(x_3, y_3, z_3)$ alors

$$\text{Det}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} x_3 + \begin{vmatrix} z_1 & z_2 \\ x_1 & x_2 \end{vmatrix} y_3 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} z_3.$$

On note par ailleurs cette dernière quantité aussi $\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$

Remarques : 1) On peut étendre la formule au déterminant de trois vecteurs de \mathbf{C}^3 exprimés dans la base canonique de \mathbf{C}^3 .

2) On remarque rapidement que si l'on permute deux des trois vecteurs \vec{u} , \vec{v} et \vec{w} , $\text{Det}(\vec{u}, \vec{v}, \vec{w})$ se transforme en son opposé.

3) On peut remarquer aussi que si deux des trois vecteurs sont colinéaires, le déterminant est nul.

4) On peut remarquer aussi que l'application Det est linéaire par rapport à chacune de ses trois variables.

Par exemple, $\text{Det}(\vec{u} + a\vec{u}', \vec{v}, \vec{w}) = \text{Det}(\vec{u}, \vec{v}, \vec{w}) + a\text{Det}(\vec{u}', \vec{v}, \vec{w})$.

Idem pour les deux autres variables.

On dit que le déterminant de trois vecteurs dans le plan est une **forme trilinéaire alternée**.

Exemple : Calculer

$$\Delta = \begin{vmatrix} a & -a & b \\ a & -a & 2b \\ -b & 2a & b \end{vmatrix}$$

de deux manières, directement avec la proposition du cours et aussi avec la règle de Pierre Sarrus que l'on rappelle pour l'occasion.

Mise en œuvre : exercice ??.

Proposition 6.4.— On a l'égalité :

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

On dit qu'un déterminant d'ordre 3 est invariant par passage à la transposition.

■ Déterminant d'une matrice carrée d'ordre n et d'une famille de n vecteurs de \mathbf{K}^n

Définition : Déterminant d'une famille de n vecteurs de \mathbf{K}^n u d'une matrice carrée de $\mathcal{M}_n(\mathbf{K})$

Il existe une unique application $\text{Det} : \mathcal{M}_n(\mathbf{K}) \rightarrow \mathbf{K}$ vérifiant les trois propriétés suivantes :

- Det est **linéaire par rapport à chacune des colonnes de sa variable**, c'est-à-dire que si C_1, \dots, C_n sont des matrices colonnes, pour tout $j \in \llbracket 1, n \rrbracket$, l'application $X \mapsto \text{Det}(C_1, C_2, \dots, C_{j-1}, X, C_{j+1}, \dots, C_n)$ est linéaire de $\mathcal{M}_{n,1}(\mathbf{K})$ vers \mathbf{K} .
- Det est **antisymétrique**, c'est-à-dire que si dans A , on effectue l'opération élémentaire $C_i \leftrightarrow C_j$ avec $i \neq j$, et si l'on note B la matrice obtenue alors $\text{Det}(A) = -\text{Det}(B)$.
- Enfin, $\text{Det}(I_n) = 1$.

On dit que Det est le **déterminant** de A et on note $\text{Det}A$ sa valeur.

Notation

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} \Rightarrow \text{Det}(A) = \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{vmatrix}.$$

Toute matrice carrée d'ordre n est aussi la matrice d'une famille de n vecteurs de \mathbf{R}^n , les n colonnes de cette matrice étant les composantes de ces vecteurs dans la base canonique de \mathbf{R}^n . On en déduit alors la définition logique suivante.

Définition : Soit E , espace vectoriel sur \mathbf{K} , $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E et $\mathcal{F} = \{\vec{u}_1, \dots, \vec{u}_n\}$ une famille de n vecteurs de E . On appelle **déterminant de la famille \mathcal{F} selon \mathcal{B}** le déterminant de la matrice carrée d'ordre n telle que pour tout j variant de 1 à n , sa colonne C_j est constitué des composantes de \vec{u}_j dans \mathcal{B} . On le note $\text{Det}_{\mathcal{B}}(\mathcal{F})$.

Remarques : 1) On remarque que cette définition généralise bien les deux paragraphes précédents pour $n = 2$ et $n = 3$.

2) Si la base \mathcal{B} est fixée pour toute (la base canonique par exemple), on note $\text{Det}_{\mathcal{B}}(\mathcal{F})$ plus simplement $\text{Det}(\mathcal{F})$ et donc on retrouve $\text{Det}(\vec{u}, \vec{v})$ en dimension 2 et $\text{Det}(\vec{u}, \vec{v}, \vec{w})$ en dimension 3.

3) **Warning :** la matrice doit être carrée ou le nombre de vecteurs doit être aussi la dimension pour faire leur déterminant. Ainsi le déterminant de deux vecteurs de l'espace n'existe pas !

Comment calculer un déterminant d'ordre n ?

Si $n = 2$ ou $n = 3$, on a des formules qui permettent de le faire.

Dans le cas général, on peut utiliser plusieurs règles.

Commençons par le développement selon une rangée.

Proposition 6.5.— Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_n(\mathbf{K})$ alors :

$$\forall j \in \llbracket 1, n \rrbracket, \text{Det } A = \sum_{i=1}^n (-1)^{i+j} \Delta_{ij} a_{ij}$$

ou

$$\forall i \in \llbracket 1, n \rrbracket, \text{Det } A = \sum_{j=1}^n (-1)^{i+j} \Delta_{ij} a_{ij},$$

où $\Delta_{i,j}$ est le déterminant d'ordre $n - 1$ issu de $\text{Det } A$, en y enlevant la ligne L_i et colonne C_j .

Dans la première formule, j est fixé arbitrairement : on dit que le déterminant de A est obtenu **en développant suivant la $j^{\text{ème}}$ colonne** ; dans la deuxième formule, i est fixé arbitrairement : on dit que le déterminant de A est obtenu **en développant suivant la $i^{\text{ème}}$ ligne**.

Exemples : 1) Reprendre le calcul de $\Delta = \begin{vmatrix} a & -a & b \\ a & -a & 2b \\ -b & 2a & b \end{vmatrix}$ par le développement selon la colonne

C_3 et aussi selon la ligne L_3 .

2) Écrire un déterminant d'ordre 4 quelconque comme combinaison linéaire de 4 déterminants d'ordre 3.

Corollaire 6.6.— Déterminant d'une matrice triangulaire —.

Le déterminant d'une matrice triangulaire est égal au produit des éléments de sa diagonale.

On a aussi les règles suivantes issues de la définition d'un déterminant (multilinéarité et antisymétrie) :

- ▶ l'opération élémentaire $C_i \leftrightarrow C_j$, où $i \neq j$, transforme le déterminant en son opposé ;
- ▶ l'opération élémentaire $C_i \leftarrow kC_i$, multiplie le déterminant par k ;
- ▶ l'opération élémentaire $C_i \leftarrow C_i + kC_j$, où $i \neq j$, laisse le déterminant invariant.

Ces dernières règles peuvent se combiner avec le développement selon une rangée.

Exemple : On reprend $\Delta = \begin{vmatrix} a & -a & b \\ a & -a & 2b \\ -b & 2a & b \end{vmatrix}$. Faire les opérations $C_2 \leftarrow C_2 + C_1$ et $C_3 \leftarrow C_3 + C_1$.

Puis développer selon la colonne C_2 . Retrouver la valeur de Δ .

Par ailleurs, on admet qu'un déterminant est invariant par transposition et on a donc aussi les règles :

- ▶ l'opération élémentaire $L_i \leftrightarrow L_j$, où $i \neq j$, transforme le déterminant en son opposé ;
- ▶ l'opération élémentaire $L_i \leftarrow kL_i$, multiplie le déterminant par k ;
- ▶ l'opération élémentaire $L_i \leftarrow L_i + kL_j$, où $i \neq j$, laisse le déterminant invariant.

Par ailleurs on peut faire ces opérations en utilisant plusieurs rangées, par exemple

$$C_1 \leftarrow C_1 + 3C_2 + C_3$$

ne modifie pas la valeur du déterminant.

Mise en œuvre : exercice ?? à exercice ??.

Donnons en synthèse les principales propriétés du déterminant d'une matrice carrée d'ordre n .

Proposition 6.7.— ▶ $\forall (A, B) \in (\mathcal{M}_n(\mathbf{K}))^2$, $\text{Det}(A.B) = \text{Det } A \times \text{Det } B$.

▶ $\forall A \in \mathcal{M}_n(\mathbf{K})$, $\forall \lambda \in \mathbf{K}$, $\text{Det}(\lambda A) = \lambda^n \text{Det } A$.

▶ $\forall A \in \mathcal{M}_n(\mathbf{K})$, $(A \in \text{GL}_n(\mathbf{K}) \Leftrightarrow \text{Det}(A) \neq 0)$.

▶ $\forall A \in \text{GL}_n(\mathbf{K})$, $\text{Det}(A^{-1}) = \frac{1}{\text{Det } A}$.

▶ Les déterminants de deux matrices semblables sont égaux.

▶ $\text{Det}(A) = \text{Det}(A^T)$.

Mise en œuvre : exercice ??, exercice ??.

Proposition 6.8.— Caractérisation d'une base —. E étant un \mathbf{K} -espace vectoriel de dimension n , une famille de n vecteurs est une base de E si et seulement si son déterminant dans une base quelconque de E est non nul.

Mise en œuvre : exercice ??.

■ Déterminant d'un endomorphisme

Définition : Soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E , $\phi \in \mathcal{L}(E)$. On appelle **déterminant de l'endomorphisme** ϕ le déterminant de la matrice carrée A d'ordre n représentant ϕ dans la base \mathcal{B} . On le note $\text{Det } \phi$. Comme le déterminant de A et d'une matrice semblable à A sont égaux, $\text{Det } \phi$ ne dépend pas de la base choisie.

Proposition 6.9.— Propriétés du déterminant d'un endomorphisme —. Soit $(\varphi, \psi) \in \mathcal{L}(E)^2$,

1. $\text{Det } (\psi \circ \varphi) = \text{Det } \psi \times \text{Det } \varphi = \text{Det } \varphi \times \text{Det } \psi$.
2. φ est un automorphisme si et seulement si $\text{Det } \varphi \neq 0$.

Dans ce cas : $\text{Det } (\varphi^{-1}) = \frac{1}{\text{Det } \varphi}$.

Mise en œuvre : exercice ??, exercice ??, exercice ??.

Méthodes

■ La pratique du calcul d'un déterminant

Il faut, bien entendu, connaître les formules de développement du déterminant d'une matrice par rapport à une ligne ou une colonne : **proposition 6.5**. Néanmoins, celles-ci conduisent régulièrement à des calculs inextricables (sauf pour la **méthode 6.3**). Les deux premières méthodes décrites ici sont donc plutôt des stratégies de contournement de ces formules.

□ Méthode 6.1.— Comment utiliser des opérations élémentaires pour se ramener au déterminant d'une matrice triangulaire

On met en œuvre une procédure de Gauss classique, avec quelques précautions :

- ▶ l'opération élémentaire $L_i \leftrightarrow L_j$, où $i \neq j$, transforme le déterminant en son opposé ;
- ▶ l'opération élémentaire $L_i \leftarrow kL_i$, multiplie le déterminant par k ;
- ▶ l'opération élémentaire $L_i \leftarrow L_i + kL_j$, où $i \neq j$, laisse le déterminant invariant.

Exemples : 1. Calculer $D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 9 & 6 \\ 7 & 8 & 9 \end{vmatrix}$.

$$L_2 \leftarrow L_2 - 4L_1 \text{ et } L_3 \leftarrow L_3 - 7L_1 \text{ donnent : } D = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & -6 \\ 0 & -6 & -12 \end{vmatrix}.$$

$$L_3 \leftarrow L_3 + 6L_2 \text{ donne : } D = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & -6 \\ 0 & 0 & -48 \end{vmatrix} = -48.$$

2. Calculer $D = \begin{vmatrix} a_1 & a_1 & a_1 & \cdots & a_1 \\ a_1 & a_2 & a_2 & \cdots & a_2 \\ a_1 & a_2 & a_3 & \cdots & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_n \end{vmatrix}$.

Les opérations $L_i \leftarrow L_i - L_{i-1}$, pour tout $i \in \llbracket 2, n \rrbracket$, en partant de $i = n$, donnent :

$$D = \begin{vmatrix} a_1 & a_1 & a_1 & \cdots & a_1 \\ 0 & a_2 - a_1 & a_2 - a_1 & \cdots & a_2 - a_1 \\ 0 & 0 & a_3 - a_2 & \cdots & a_3 - a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n - a_{n-1} \end{vmatrix} = a_1 \prod_{i=2}^n (a_i - a_{i-1}).$$

Ces manipulations peuvent être interrompues à tout moment et le calcul achevé par un développement par rapport à une ligne (ou une colonne).

□ Méthode 6.2.— Comment mettre en évidence un facteur commun aux éléments d'une colonne (ligne)

Il s'agit souvent d'une manipulation préparatoire à la **méthode 6.1**.

- Si, en ajoutant à une colonne une ou plusieurs des autres colonnes, on obtient une colonne formée d'éléments identiques et de zéros, on factorise le déterminant par ce facteur commun en faisant ainsi apparaître une colonne de 1 et de 0.
- Des opérations élémentaires sur les lignes permettent alors de faire apparaître un seul 1 et des 0 dans cette colonne par rapport à laquelle on va développer.

Exemples : 1. Calculer $D = \begin{vmatrix} 0 & a & b & c \\ a & 0 & b & c \\ a & b & 0 & c \\ a & b & c & 0 \end{vmatrix}$.

Les sommes des éléments de chaque ligne sont égales : on ajoute donc à la première colonne les trois autres.

$$D = \begin{vmatrix} a+b+c & a & b & c \\ a+b+c & 0 & b & c \\ a+b+c & b & 0 & c \\ a+b+c & b & c & 0 \end{vmatrix} = (a+b+c) \begin{vmatrix} 1 & a & b & c \\ 1 & 0 & b & c \\ 1 & b & 0 & c \\ 1 & b & c & 0 \end{vmatrix}.$$

En retranchant la première ligne aux trois autres, puis en développant par rapport à la première colonne :

$$D = (a+b+c) \begin{vmatrix} 1 & a & b & c \\ 0 & -a & 0 & 0 \\ 0 & b-a & -b & 0 \\ 0 & b-a & c-b & -c \end{vmatrix} = (a+b+c) \begin{vmatrix} -a & 0 & 0 \\ b-a & -b & 0 \\ b-a & c-b & -c \end{vmatrix} = -abc(a+b+c).$$

2. Calculer $D = \begin{vmatrix} a & b & c \\ b & a & b \\ c & -c & a \end{vmatrix}$.

$$C_1 \leftarrow C_1 + C_2 \text{ donne : } D = \begin{vmatrix} a+b & b & c \\ a+b & a & b \\ 0 & -c & a \end{vmatrix} = (a+b) \begin{vmatrix} 1 & b & c \\ 1 & a & b \\ 0 & -c & a \end{vmatrix}.$$

$L_2 \leftarrow L_2 - L_1$, puis le développement par rapport à la première colonne, donnent :

$$D = (a+b) \begin{vmatrix} 1 & b & c \\ 0 & a-b & b-c \\ 0 & -c & a \end{vmatrix},$$

c'est-à-dire :

$$D = (a+b)(a^2 - ab + bc - c^2) = (a+b)(a-c)(a-b+c).$$

❏ Méthode 6.3.— Comment développer suivant une ligne (colonne) pour mettre en évidence une relation de récurrence

Le champ d'application privilégié de cette méthode est le calcul de déterminants dont l'ordre n est quelconque.

- On développe par rapport à une ligne (ou une colonne) la plus creuse possible pour essayer de trouver une relation de récurrence simple.
- On résout la récurrence, ce qui nécessite de connaître quelques méthodes usuelles.

Exemple : Calculer $D_n = \begin{vmatrix} 0 & 0 & \cdots & \cdots & a_n \\ 0 & 0 & \cdots & a_{n-1} & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & a_2 & 0 & \cdots & 0 \\ a_1 & 0 & \cdots & \cdots & 0 \end{vmatrix}$.

Si l'un des a_j , $j \in \llbracket 1, n \rrbracket$ est nul, la matrice est de rang strictement inférieur à n , et $D_n = 0$. Dans le cas contraire (aucun des a_j n'est nul), développons D_n par rapport à la première ligne :

$$D_n = (-1)^{n+1} a_n \begin{vmatrix} 0 & 0 & \cdots & \cdots & a_{n-1} \\ 0 & 0 & \cdots & a_{n-2} & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & a_2 & 0 & \cdots & 0 \\ a_1 & 0 & \cdots & \cdots & 0 \end{vmatrix}.$$

Et donc :

$$D_n = (-1)^{n+1} a_n D_{n-1} = (-1)^{n-1} a_n D_{n-1}.$$

De cette relation de récurrence, on déduit :

$$\prod_{k=2}^n D_k = \prod_{k=2}^n (-1)^{k-1} \cdot \prod_{k=2}^n a_k \cdot \prod_{k=2}^n D_{k-1}.$$

Il en résulte que, par récurrence sur $n \in \mathbf{N}^*$, $\prod_{k=2}^n D_k$ est non nul, donc *a fortiori* D_n est non nul.

On trouve alors par simplification

$$D_n = D_1 \cdot \prod_{k=2}^n (-1)^{k-1} \cdot \prod_{k=2}^n a_k = D_1 \cdot (-1)^{\sum_{k=2}^n (k-1)} \prod_{k=2}^n a_k.$$

Après examen des valeurs de $D_1 = a_1$ et $D_2 = -a_1 a_2$, on aboutit à :

$$D_n = (-1)^{\frac{n(n-1)}{2}} \prod_{k=1}^n a_k.$$

■ Formule $\text{Det}(A \times B) = \text{Det } A \times \text{Det } B$

Exemple : Soit a, b et c trois complexes, $M = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ et $J = \begin{pmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{pmatrix}$, où j est la racine cubique de l'unité de partie imaginaire strictement positive.

1. Déterminer directement une matrice diagonale D telle que : $MJ = JD$.

2. En déduire une factorisation du déterminant de M en trois facteurs.

1.
$$MJ = \begin{pmatrix} a+b+c & a+bj+cj^2 & a+bj^2+cj \\ a+b+c & c+aj+bj^2 & c+aj^2+bj \\ a+b+c & b+cj+aj^2 & b+cj^2+aj \end{pmatrix}.$$

Si $D = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$, $JD = \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha & \beta j & \gamma j^2 \\ \alpha & \beta j^2 & \gamma j \end{pmatrix}.$

Comme $j^3 = 1$ et $j^4 = j$, $MJ = JD \Leftrightarrow \begin{cases} \alpha = a+b+c \\ \beta = a+bj+cj^2 \\ \gamma = a+bj^2+cj \end{cases}$

2. $MJ = JD \Rightarrow \text{Det } M \times \text{Det } J = \text{Det } J \times \text{Det } D$. Or, $\text{Det } J = (1-j)^3$, donc :

$$\text{Det } M = \text{Det } D = (a+b+c)(a+bj+cj^2)(a+bj^2+cj).$$

Chapitre 7

Réduction des endomorphismes

■ Objectifs

■ Les incontournables

- ▶ trouver les valeurs propres, les vecteurs propres d'un endomorphisme, d'une matrice carrée ;
- ▶ diagonaliser une matrice d'ordre 2 et 3 ;
- ▶ diagnostiquer le caractère diagonalisable ou trigonalisable d'un endomorphisme, d'une matrice carrée ;
- ▶ calculer les puissances successives de matrices diagonalisables ;
- ▶ savoir interpréter matriciellement une récurrence linéaire d'ordre 2 d'une suite de \mathbf{K}^n et utiliser la réduction pour déterminer u_n en fonction de n .

■ Et plus si affinités

- ▶ savoir trigonaliser une matrice carrée d'ordre 2 ou d'ordre 3 s'il y a au moins deux valeurs propres distinctes ;
- ▶ calculer les puissances successives de matrices trigonalisées ;
- ▶ savoir interpréter matriciellement une récurrence linéaire d'ordre supérieur à 2 d'une suite de \mathbf{K}^n et utiliser la réduction pour déterminer u_n en fonction de n .

Résumé de cours

Dans ce chapitre, E est un \mathbf{K} -espace vectoriel de dimension finie, \mathbf{K} étant \mathbf{R} ou \mathbf{C} et $u \in \mathcal{L}(E)$, c'est-à-dire que u est un endomorphisme de E

■ Valeurs et vecteurs propres d'un endomorphisme

Définition : Le scalaire $\lambda \in \mathbf{K}$ est une **valeur propre** de l'endomorphisme u s'il existe un vecteur $\vec{x} \neq 0_E$ tel que $u(\vec{x}) = \lambda\vec{x}$. Un tel vecteur \vec{x} est appelé **vecteur propre** associé à λ .

Le **spectre** de u dans \mathbf{K} , noté $\text{Sp}_{\mathbf{K}}(u)$ (ou parfois simplement $\text{Sp}(u)$, ou encore Sp si l'on a à faire à un seul endomorphisme) est l'ensemble de ses valeurs propres.

Définition : Si λ est une valeur propre de u , le **sous-espace propre** associé à λ est :

$$E_{\lambda}(u) = \ker(u - \lambda \text{Id}_E).$$

Attention : $\vec{0}_E$ n'est pas un vecteur propre et pourtant $\vec{0}_E \in E_{\lambda}(u)$.

Mise en œuvre : exercice 7.1.

Théorème 7.1.— Toute famille de vecteurs propres associés à des valeurs propres deux à deux distinctes est libre.

La somme d'une famille de sous-espaces propres associés à des valeurs propres deux à deux distinctes est directe.

■ Valeurs et vecteurs propres d'une matrice carrée

Définition : Soit M une matrice de $\mathcal{M}_n(\mathbf{K})$ et u l'endomorphisme de \mathbf{K}^n qui lui est canoniquement associé.

Alors les éléments propres de M sont ceux de u , plus précisément :

- ▶ les **valeurs propres** de M sont celles de u ;
- ▶ les **vecteurs propres** de M sont les vecteurs-colonnes canoniquement associés aux vecteurs propres de u , c'est-à-dire vérifiant $M.X = \lambda X$;
- ▶ les **sous-espaces propres** de M sont les sous-espaces de $\mathcal{M}_{n,1}(\mathbf{K})$ canoniquement associés à ceux de u ;
- ▶ enfin, le **spectre** $\text{Sp}_{\mathbf{K}}(M)$ de M est l'ensemble des valeurs propres de M .

Théorème 7.2.— Si $M \in \mathcal{M}_n(\mathbf{R})$, alors $\text{Sp}_{\mathbf{R}}(M) \subset \text{Sp}_{\mathbf{C}}(M)$; autrement dit, une valeur propre réelle d'une matrice réelle est également valeur propre complexe de cette matrice, considérée comme élément de $\mathcal{M}_n(\mathbf{C})$.

Mise en œuvre : exercice 7.2.

■ Polynôme caractéristique

Théorème 7.3.— Le scalaire λ est valeur propre de l'endomorphisme u si et seulement si $\lambda \text{Id}_E - u$ n'est pas inversible.

Définition : Le *polynôme caractéristique* de l'endomorphisme u est $x \mapsto \chi_u(x) = \text{Det}(x \text{Id}_E - u)$.

Définition : Le *polynôme caractéristique* de la matrice $A \in \mathcal{M}_n(\mathbf{K})$ est $x \mapsto \chi_A(x) = \text{Det}(xI_n - A)$.

Exemples : Écrire le polynôme caractéristique de $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ et de $B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 1 & 0 \end{pmatrix}$.

Que remarque t-on ?

Proposition 7.4.— Le degré du polynôme caractéristique est égal à n , dimension de l'espace E et son coefficient dominant est 1.

Proposition 7.5.— Deux matrices semblables ont le même polynôme caractéristique.

Preuve : À faire.

La conséquence est que le polynôme caractéristique d'un endomorphisme u est celui de sa matrice dans n'importe quelle base.

Théorème 7.6.— Les valeurs propres de u ou de A sont les racines de son polynôme caractéristique.

Définition : L'*ordre de multiplicité* d'une valeur propre de u est son ordre de multiplicité en tant que racine du polynôme caractéristique de u .

Théorème 7.7.— Soit λ une valeur propre de u et m_λ la multiplicité de λ dans le polynôme caractéristique. Alors :

$$1 \leq \dim E_\lambda(u) \leq m_\lambda.$$

Exemple : On pose $A = \begin{pmatrix} -1 & 2 & -1 \\ -3 & 4 & -3 \\ -4 & 4 & -4 \end{pmatrix}$. Déterminer son polynôme caractéristique et l'ordre de multiplicité de ses racines. Si $E_\lambda(A)$ est un de ses sous-espaces propres, quelle est obligatoirement sa dimension ?

■ Endomorphismes diagonalisables

Définition : L'endomorphisme u est dit *diagonalisable* s'il existe une base dans laquelle sa matrice est diagonale.

Remarque : on donne ici trois conditions nécessaires et suffisantes (*CNS*) et une condition suffisante (*CS*) de diagonalisabilité. Il est *primordial* de distinguer les deux types de condition, la condition suffisante étant plus facile à manier, mais pas forcément vérifiée...

Théorème 7.8.— CNS de diagonalisabilité 1 —. u est diagonalisable si et seulement s'il existe une base de E formée de vecteurs propres de u , c'est-à-dire s'il existe une base de E dans laquelle la matrice de u est diagonale.

Théorème 7.9.— CNS de diagonalisabilité 2 —. L'endomorphisme u est diagonalisable si et seulement si la somme des dimensions des sous-espaces propres de u est la dimension de E .

Théorème 7.10.— CNS de diagonalisabilité 3 —. L'endomorphisme u est diagonalisable si et seulement si le polynôme caractéristique de u est scindé et si pour chaque valeur propre λ , la dimension de l'espace propre associé est égale à la multiplicité de λ dans le polynôme caractéristique.

Corollaire 7.11.— CS de diagonalisabilité —. Si le polynôme caractéristique de u est scindé à racines simples, **alors** u est diagonalisable, et ses sous-espaces propres sont des droites vectorielles.

Remarque : dans les exercices, on parle parfois de matrice diagonalisable et non d'endomorphisme diagonalisable.

Définition : Une matrice est **diagonalisable** si elle est semblable à une matrice diagonale.

Remarques : 1. M est diagonalisable si et seulement si elle s'écrit sous la forme :

$$M = P.D.P^{-1},$$

(où D est diagonale) les éléments diagonaux étant les valeurs propres, répétées suivant leur ordre de multiplicité, et P est la matrice de passage de la base canonique vers une base de vecteurs propres (disposés dans le même ordre que les valeurs propres associées sur la diagonale de D).

Dans ce contexte, on notera les sous-espaces propres plutôt $E_\lambda(M)$ que $E_\lambda(\phi)$.

2. Attention à un abus de langage courant. À la place d'écrire qu'une matrice est diagonalisable dans $\mathcal{M}_n(\mathbf{R})$ ou dans $\mathcal{M}_n(\mathbf{C})$, on écrit parfois qu'elle est diagonalisable dans \mathbf{R} ou dans \mathbf{C} . La première écriture est rigoureuse mais malheureusement, la seconde est utilisée. Donc on pourra l'utiliser nous aussi mais en ayant conscience de l'abus fait.

Mise en œuvre : exercice 7.3 à exercice 7.7, exercice 7.9.

■ Endomorphismes trigonalisables

Définition : L'endomorphisme u est **trigonalisable** s'il existe une base de E dans laquelle la matrice de u est triangulaire supérieure.

Théorème 7.12.— L'endomorphisme u est trigonalisable si et seulement si son polynôme caractéristique est scindé.

Remarque : Si $\mathbf{K} = \mathbf{C}$, tout endomorphisme est trigonalisable.

Remarque : Une matrice est **trigonalisable** si elle est semblable à une matrice triangulaire.

Il faut noter qu'aucune méthode de trigonalisation effective n'est au programme mais l'étudiant doit être capable de s'en sortir s'il est guidé.

Exemple : Soit $A = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \\ -4 & 2 & 3 \end{pmatrix}$.

1. Montrer que $\chi_A(X) = (X - 3)(X - 1)^2$.

2. Montrer que les deux sous-espaces propres sont $E_3 = \text{Vect}(\vec{u}_1)$ et $E_1 = \text{Vect}(\vec{u}_2)$, où $\vec{u}_1 = (0, 0, 1)$ et $\vec{u}_2 = (1, 1, 1)$. A est-elle diagonalisable? trigonalisable?

3. On pose $P = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$. Montrer qu'il existe a et b tels que $T = \begin{pmatrix} 3 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = P^{-1}AP$.

Déterminer a et b en utilisant $PT = AP$. Conclure.

Mise en œuvre : exercice 7.10, exercice 7.11.

Proposition 7.13.— Dans le cas où le polynôme caractéristique de u est scindé dans \mathbf{K} (c'est-à-dire donc dans le cas où u est trigonalisable dans \mathbf{K}) et si l'on note λ_i ses racines (éléments de \mathbf{K}) d'ordre de multiplicité α_i , pour i variant de 1 à p , alors :

$$\text{Det } u = \prod_{i=1}^p \lambda_i^{\alpha_i} \text{ et } \text{Tr}(u) = \sum_{i=1}^p \alpha_i \lambda_i.$$

Remarques : 1. La proposition 7.13 est donc toujours valable si $\mathbf{K} = \mathbf{C}$.

2. Si ϕ est diagonalisable, il est trigonalisable et la proposition 7.13 reste valable.

3. Cette proposition est utile. En effet, si par exemple $n = 3$ et que l'on connaisse une valeur propre, on fait l'hypothèse que le polynôme caractéristique est scindé (ou alors on se place dans \mathbf{C}) et on peut obtenir les deux autres valeurs propres grâce à la trace et au déterminant. Si l'on connaît deux valeurs propres, la trace suffit.

4. Enfin, faire attention à ce que l'on dit. Si vous dites : « la trace est la somme des valeurs propres », c'est imprécis. Il faut dire : « la trace est la somme des valeurs propres répétées selon leur ordre de multiplicité ».

Mise en œuvre : exercice 7.8.

Quels sont les principales applications de la diagonalisation ?

Certaines applications sont vues plus tard (systèmes différentiels), d'autres dépassent le cadre du cours de Math (matrice d'inertie, matrice de Lorentz etc.). Dans ce chapitre, nous en verrons trois.

1. Le calcul de A^n (voir l'exercice 7.12). L'idée est de remarquer que si l'on a $A = PDP^{-1}$ alors par récurrence, on a pour tout $n \in \mathbf{N}$, $A^n = PD^nP^{-1}$. Et D^n se calcule plus facilement que A^n directement (car D est diagonale ou triangulaire).

2. Le calcul du terme général d'une suite récurrente linéaire (voir l'exercice 7.13).

3. La résolution d'équation matricielle (voir l'exercice 7.14).

Bien entendu, il faut que tous ces problèmes se ramènent à une matrice diagonalisable ou au moins trigonalisable, ce qui est le cas dans les trois exercices cités.

Méthodes

■ Comment déterminer les valeurs propres ou (et) les vecteurs propres de $u \in \mathcal{L}(E)$, de matrice canoniquement associée A

- ❑ **Méthode 7.1.**— ▶ On peut calculer le polynôme caractéristique de A (voir la **méthode 7.2**). Ses racines sont les valeurs propres de a et de u .
- ▶ On peut poser $u(\vec{x}) = \lambda \vec{x}$ ou $AX = \lambda X$ et chercher des conditions vérifiées par λ ou (et) par \vec{x} .
 - ▶ Sous l'hypothèse où u est diagonalisable, $\text{Tr}(u)$ et $\text{Det } u$ donnent respectivement la somme et le produit des valeurs propres. Cela peut être utile pour les trouver.
 - ▶ Le rang de u (qui est celui de A) peut être aussi utile. En effet, si $\text{Rg}(u) = p$ alors 0 est valeur propre de multiplicité au moins $\dim E - p$.

Exemple : Si $n \geq 3$, $A \in \mathcal{M}_n(\mathbf{C})$ avec $A^3 + A^2 + A = 0$. Valeurs propres possibles de A ?

$AX = \lambda X$ avec $X \neq 0$ donne immédiatement :

$$\lambda^3 + \lambda^2 + \lambda = 0.$$

Donc $\lambda \in \{0, j, j^2\}$, $j = e^{i2\pi/3}$.

Par contre, rien ne permet d'affirmer que ces trois valeurs soient effectivement toutes des valeurs propres. Une connaissance plus précise de A est nécessaire pour chercher alors $\ker(A - \lambda I_n)$.

Remarque : Dans certains exercices, on doit caractériser le fait que λ n'est pas valeur propre. Si l'on pose $\dim E = n$, on écrit alors souvent que $A - \lambda I_n$ est inversible ou alors que $A - \lambda I_n$ est injective.

■ Comment calculer un polynôme caractéristique

❑ **Méthode 7.2.**— Soit A une matrice carrée d'ordre $n \geq 1$, $\chi_A(X)$ est un déterminant. On peut chercher à l'obtenir sous forme factorisée (car ce qui nous intéresse en général sont ses racines) et il vaut mieux tenter de commencer par des opérations élémentaires adéquates (par exemple éviter d'utiliser la formule de Sarrus dans le cas $n = 3$).

Le terme dominant de $\chi_A(X)$ est X^n et le suivant (on le donne juste pour vous mais ce n'est pas explicitement à savoir !) est $-\text{Tr}(A)X^{n-1}$. Ceci permet une vérification rapide du calcul qui dépiste une bonne partie des erreurs de calcul.

Exemple : Soit $A = \begin{pmatrix} -1 & 2 & -1 \\ -3 & 4 & -3 \\ -4 & 4 & -4 \end{pmatrix}$. Calculer $\chi_A(X)$.

$$\chi_A(X) = \begin{vmatrix} X+1 & -2 & 1 \\ 3 & X-4 & 3 \\ 4 & -4 & 4+X \end{vmatrix} = (X-1) \begin{vmatrix} 1 & -2 & 1 \\ 1 & X-4 & 3 \\ 0 & -4 & 4+X \end{vmatrix}$$

en faisant : $C_1 \leftarrow C_1 + C_2$. Puis $L_2 \leftarrow L_2 - L_1$ donne :

$$\chi_A(X) = (X-1) \begin{vmatrix} 1 & -2 & 1 \\ 0 & X-2 & 2 \\ 0 & -4 & 4+X \end{vmatrix} = (X-1) \begin{vmatrix} X-2 & 2 \\ -4 & 4+X \end{vmatrix},$$

ce qui donne : $\chi_A(X) = (X-1)(X^2 + 2X) = X(X-1)(X+2)$.

■ Comment diagonaliser une matrice carrée d'ordre n .

❑ **Méthode 7.3.** — ► La méthode normale est de calculer le polynôme caractéristique χ_A de la matrice A . Supposons pour la suite que χ_A soit scindé (quitte à se placer dans **C**). On détermine ses racines $\lambda_1, \dots, \lambda_n$, pas nécessairement distinctes et qui sont les valeurs propres de A . Pour chaque valeur propre **multiple**, on calcule la dimension du sous espace propre associé. Si la somme des dimensions des sous-espaces propres est égale à n , la matrice est diagonalisable. On prend alors une base de chaque sous-espace propre ; en les réunissant, on obtient une base qui diagonalise A .

► Parfois, on peut calculer rapidement le rang de A et on sait alors que 0 est valeur propre et que son sous espace propre associé a pour dimension $n - \text{Rg } A$. Il reste à trouver les autres valeurs propres (utiliser la trace par exemple) .

Mise en œuvre : de l'exercice ?? à l'exercice ??.

Exemples : 1. Réduire la matrice $A = \begin{pmatrix} -1 & 2 & -1 \\ -3 & 4 & -3 \\ -4 & 4 & -4 \end{pmatrix}$.

D'après le calcul fait ci-dessus (exemple **1.** de la **méthode 7.2**), $\chi_A(X) = (X-1)X(X+2)$.

Les valeurs propres sont donc $-2, 0$ et 1 .

Comme χ_A est scindé avec des valeurs propres simples, A est diagonalisable, et de plus, les sous-espaces propres correspondants sont des droites vectorielles.

Pour trouver $E_0(A)$, résolvons le système linéaire d'écriture matricielle $A.X = 0_{\mathcal{M}_3(\mathbf{R})}$, en n'oubliant pas que ses solutions forment une droite vectorielle :

$$\begin{cases} -x + 2y - z = 0 \\ -3x + 4y - 3z = 0 \\ -4x + 4y - 4z = 0 \end{cases} \Leftrightarrow \begin{cases} -x + 2y - z = 0 \\ -x + y - z = 0 \end{cases} \Leftrightarrow \begin{cases} y = 0 \\ x = -z \end{cases}$$

$E_0(A) = \text{Vect } \vec{u}_0$ avec $\vec{u}_0 = (-1, 0, 1)$. Le vecteur $\vec{x} \in E_{-2}(A)$ si et seulement si le vecteur colonne X qui le représente dans la base canonique vérifie $A.X = -2.X$ soit $(A + 2I_3)X = 0_{\mathcal{M}_3(\mathbf{R})}$.

On trouve $E_{-2}(A) = \text{Vect } \vec{u}_{-2}$ avec $\vec{u}_{-2} = (0, 1, 2)$.

De même, pour déterminer $E_1(A)$, on se ramène à la résolution du système $(A - I_3)X = 0_{\mathcal{M}_3(\mathbf{R})}$.

On trouve $E_1(A) = \text{Vect } \vec{u}_1$ avec $\vec{u}_1 = (1, 1, 0)$.


En prenant $P = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$, on obtient $P^{-1}.A.P = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

2. Diagonaliser la matrice $J \in \mathcal{M}_4(\mathbf{R})$ ayant tous ses coefficients égaux à 1.

Immédiatement, $J = 1$ et donc 0 est une valeur propre de J avec $\dim E_0 = \dim \ker J = 3$. Comme $\text{Tr} J = 4$, 4 est la dernière valeur propre. Pour le vérifier, on résout le système $JX = 4X$. On trouve la droite vectorielle de base $\{(1, 1, 1, 1)\}$. Donc J est bien diagonalisable dans \mathbf{R} .

Remarquons qu'ici, se lancer dans $\chi_J(X)$ aurait été maladroit.

■ Comment montrer qu'un endomorphisme n'est pas diagonalisable sans calcul du polynôme caractéristique

 **Méthode 7.4.**— Si A est la matrice associée, on part de $AX = \lambda X$, où $X \neq 0$, et on tente de trouver une contradiction à partir des λ possibles.

Exemple : On appelle **nilpotent** un endomorphisme non nul tel qu'il existe un entier $p \in \mathbf{N}^*$, tel que $f^p = 0_{\mathcal{L}(E)}$. Montrer qu'un tel endomorphisme n'est jamais diagonalisable.

Si A est la matrice associée à f dans la base canonique de E , posons $X \neq 0$ et cherchons λ , une éventuelle valeur propre (de vecteur propre X) :

$$AX = \lambda X \Rightarrow A^p X = \lambda^p X = 0 \Rightarrow \lambda^p = 0 \Rightarrow \lambda = 0.$$

Si f était diagonalisable, il serait l'endomorphisme nul car sa matrice dans une base de vecteurs propres serait... la matrice nulle ! Ce n'est pas le cas.

Remarque : il existe une autre condition de diagonalisabilité très courante : toute matrice qui est symétrique réelle dans $\mathcal{M}_n(\mathbf{R})$ est diagonalisable dans \mathbf{R} . C'est le cas ainsi de la matrice J dans un exemple précédent. C'est dans le chapitre sur les endomorphismes symétriques.

Dans le programme officiel, il faut savoir reconnaître le cas où une matrice est trigonalisable dans \mathbf{K} . Il faut et il suffit que le polynôme caractéristique de la matrice soit scindé dans \mathbf{K} . Après, pour « trigonaliser » il faut être aidé par l'énoncé car aucune méthode de trigonalisation n'est au programme. Maintenant, donnons une méthode (cela peut vous être utile) dans le cas (très courant) de trigonalisation d'une matrice carrée d'ordre 3 ayant au moins deux valeurs propres distinctes.

■ Comment prouver qu'une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est trigonalisable

□ **Méthode 7.5.**— On applique le théorème de caractérisation de la trigonalisation.

- ▶ on calcule le polynôme caractéristique de A ;
- ▶ on vérifie qu'il est scindé sur \mathbf{K} .

Exemple : Soit $a \in \mathbf{R}$ et $A = \begin{pmatrix} 3 & a \\ 2 & 1 \end{pmatrix}$. Déterminer la valeur de a telle que A soit uniquement trigonalisable sur \mathbf{R} , non diagonalisable (sur \mathbf{R} et sur \mathbf{C}).

On a $\chi_A(X) = X^2 - 4X + 3 - 2a = (X - 2)^2 - (1 + 2a)$. Plusieurs cas se présentent suivant les valeurs de $1 + 2a$:

- ▷ Si $1 + 2a > 0$, χ_A est scindé simple sur \mathbf{R} et A est diagonalisable sur \mathbf{R} .
- ▷ Si $1 + 2a < 0$, χ_A n'est pas scindé sur \mathbf{R} donc A n'est pas trigonalisable sur \mathbf{R} . Cependant puisque χ_A est scindé simple sur \mathbf{C} , A est diagonalisable sur \mathbf{C} .
- ▷ Si $1 + 2a = 0$, χ_A est scindé sur \mathbf{R} et 2 est l'unique valeur propre de A . Si A était diagonalisable, il existerait une matrice inversible P telle que : $P^{-1}AP = 2I_2 \iff A = 2I_2$.

Puisque A est différente de $2I_2$, la matrice A n'est pas diagonalisable. Puisque χ_A est scindé sur \mathbf{R} , elle est trigonalisable sur \mathbf{R} .

■ Comment trigonaliser une matrice $A \in \mathcal{M}_3(\mathbf{K})$ sachant que $\chi_A(X) = (X - \alpha)^2(X - \beta)$, où $\alpha \neq \beta$. On suppose aussi que $\dim E_\alpha(A) = 1$.

□ **Méthode 7.6.**— ▶ On sait déjà que A est trigonalisable mais pas diagonalisable dans \mathbf{K} . On détermine une base de $E_\alpha(A)$, que l'on note (\vec{u}_1) puis une base de $E_\beta(A)$, que l'on note (\vec{u}_2) .

- ▶ On complète par un vecteur \vec{u}_3 (en général un des vecteurs de la base canonique) telle que $\mathcal{B} = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$ soit une base de \mathbf{R}^3 . On introduit une matrice P ayant pour colonnes les trois vecteurs de la base \mathcal{B} dans l'ordre. Il existe alors $(a, b) \in \mathbf{K}^2$

$$\text{tel que : } P^{-1}AP = \begin{pmatrix} \alpha & 0 & a \\ 0 & \beta & b \\ 0 & 0 & \alpha \end{pmatrix}.$$

- ▶ Appelons T la matrice triangulaire supérieure $P^{-1}AP$.
On détermine les coefficients a et b en utilisant $AP = PT$ et plus particulièrement en égalant les troisièmes colonnes de ces matrices.
On peut aussi exprimer $f(\vec{u}_3) - \alpha\vec{u}_3$ en fonction de \vec{u}_1 et de \vec{u}_2 .

Exemple : Trigonaliser $A = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \\ -4 & 2 & 3 \end{pmatrix}$.

Rapidement, $\chi_A(X) = (X - 3)(X - 1)^2$. Puis $E_3 = \text{Vect}(\vec{u}_1)$ et $E_1 = \text{Vect}(\vec{u}_2)$, où $\vec{u}_1 = (0, 0, 1)$ et $\vec{u}_2 = (1, 1, 1)$. Il reste à compléter avec $\vec{u}_3(1, 0, 0)$ par exemple. La matrice $P = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$

est inversible. On cherche a et b tels que $T = \begin{pmatrix} 3 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = P^{-1}AP$. On écrit :

$$PT = AP \Rightarrow \begin{cases} b+1 = 2 \\ b = 1 \\ a+b = -4 \end{cases} \Rightarrow b = 1 \text{ et } a = -5.$$

Finalement dans $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, la matrice de l'endomorphisme canoniquement associé à A est :

$$T = \begin{pmatrix} 3 & 0 & -5 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

■ Comment calculer la puissance $p^{\text{ème}}$ d'une matrice A par réduction de cette matrice

□ Méthode 7.7.— En diagonalisant (ou à défaut en trigonalisant) la matrice A . On détermine alors R . On écrit : $A = P.R.P^{-1}$ et on établit la formule : $\forall p \in \mathbf{N}^*, A^p = P.R^p.P^{-1}$. Ceci complète la quatrième piste de la **méthode 3.10**. Cette formule s'étend aux valeurs négatives ou nulle de p si A est inversible. On rappelle au passage que si R est diagonale alors on a l'égalité : $\forall p \in \mathbf{N}^*, \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)^p = \text{diag}(\lambda_1^p, \lambda_2^p, \dots, \lambda_n^p)$.

Exemple : Calculer A^p , où $A = \begin{pmatrix} -1 & 2 & -1 \\ -3 & 4 & -3 \\ -4 & 4 & -4 \end{pmatrix}$.

Cette matrice a déjà été réduite dans la **méthode 7.3**, $A = PRP^{-1}$, où les héros sont les matrices :

$$R = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, P = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}, P^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ -2 & 2 & -1 \\ -1 & 2 & -1 \end{pmatrix}.$$

$$\text{Et : } \forall p \in \mathbf{N}^*, A^p = PR^pP^{-1} = \begin{pmatrix} -1 & 2 & -1 \\ -1 + (-2)^p & 2 - (-2)^p & -1 + (-2)^p \\ 2(-2)^p & -2(-2)^p & 2(-2)^p \end{pmatrix}.$$

Pour $p = 0$, la formule proposée pour R^n donne un autre résultat que I_3 , donc celle de A^p n'est pas applicable au cas $p = 0$ (et encore moins pour $p < 0$ puisque A n'est pas inversible!).

■ Comment résoudre l'équation $X^p = A$, où A est diagonalisable

Ce type de problématique se rencontre dans les sujets TSI de temps en temps donc autant en parler.

□ **Méthode 7.8.**— On commence par diagonaliser A et on détermine la matrice de passage P telle que $A = PDP^{-1}$. Puis on pose $Y = P^{-1}XP$ et on remarquera que :

$$Y^p = D \Leftrightarrow X^p = A.$$

On résout ensuite l'équation $Y^p = D$. (En effet, D étant diagonale, $Y^p = D$ se résout plus simplement que $X^p = A$.)

On en déduit Y puis X .

Exemple : Soit $A = \begin{pmatrix} 11 & -5 & -5 \\ -5 & 3 & 3 \\ -5 & 3 & 3 \end{pmatrix}$. On désire résoudre l'équation $X^2 = A$ dans $\mathcal{M}_3(\mathbb{R})$.

1. Réduire A (on notera $P \in \text{GL}_3(\mathbb{R})$ et D diagonale telles que $A = PDP^{-1}$).
2. Pour $X \in \mathcal{M}_3(\mathbb{R})$, on pose $C = P^{-1}XP$, montrer que $X^2 = A \Leftrightarrow C^2 = D$.
3. Résoudre l'équation $C^2 = D$ d'inconnue $C \in \mathcal{M}_3(\mathbb{R})$ (montrer que C et D commutent puis que C est diagonale) et en déduire les solutions de $X^2 = A$.

1. On remarque que A a deux lignes égales et donc son rang est au plus 2. Ainsi, 0 est valeur propre au moins simple. Comme il nous en manque certainement deux, on calcule quand même $\chi_A(X) = X(X-1)(X-16)$. On sait que A est diagonalisable (car elle possède trois valeurs propres distinctes). Puis $E_0(A) = \text{Vect}(\{(0, -1, 1)\})$, $E_1(A) = \text{Vect}(\{(1, 1, 1)\})$, $E_{16}(A) = \text{Vect}(\{(-2, 1, 1)\})$.

Ce qui donne : $P = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ et $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix}$.

2. On a sans souci : $X^2 = A \Leftrightarrow P^{-1}X^2P = P^{-1}AP \Leftrightarrow C^2 = D$.

3. C'est très classique ! On a évidemment $CD = C^3 = DC$. Posons $C = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$.

L'égalité $CD = DC$ donne : $b = c = d = f = g = h = 0$.

C est bien diagonale. Enfin, $C^2 = D$ donne : $a^2 = 0$, $b^2 = 1$ et $c^2 = 16$. Il reste quatre matrices C possibles. Comme $X = PCP^{-1}$, les solutions de $X^2 = A$ sont :

$$P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} P^{-1}, P \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} P^{-1}, P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix} P^{-1}, P \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix} P^{-1}.$$

■ **Comment trouver, à l'aide d'une méthode matricielle, les suites $(u_n) \in \mathbf{K}^{\mathbf{N}}$ vérifiant une relation de récurrence linéaire d'ordre 2**

□ **Méthode 7.9.**— Soit $(u_n) \in \mathbf{K}^{\mathbf{N}}$, définie par la donnée de u_0 et de u_1 , et la relation de récurrence : $(R) \quad \forall n \in \mathbf{N}, \quad u_{n+2} = a u_{n+1} + b u_n$, (a, b) sont deux éléments fixés de \mathbf{K} .

[1] On pose $U_n = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}$, alors la relation (R) devient $U_{n+1} = A U_n$ avec $A = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix}$.

Par récurrence sur $n \in \mathbf{N}$, on aboutit à $U_n = A^n U_0$.

On applique la méthode précédente (calcul de la puissance $p^{\text{ème}}$ d'une matrice).

Exemple : *Expliciter la suite définie par $u_0 = u_1 = 1$ et $\forall n \in \mathbf{N}, u_{n+2} = u_{n+1} + u_n$.*

En posant $U_n = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}$, alors la relation de récurrence devient $U_{n+1} = A U_n$ avec $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

D'où : $\chi_A(X) = X^2 - X - 1$, donc A a pour valeurs propres $\alpha = \frac{1 + \sqrt{5}}{2}$ et $\beta = \frac{1 - \sqrt{5}}{2}$.

On trouve $A = P \cdot \text{diag}(\alpha, \beta) \cdot P^{-1}$, avec :

$$P = \begin{pmatrix} 1 & 1 \\ \alpha & \beta \end{pmatrix} \text{ et } P^{-1} = \frac{1}{\beta - \alpha} \begin{pmatrix} \beta & -1 \\ -\alpha & 1 \end{pmatrix}.$$

Alors, pour tout $n \in \mathbf{N}$:

$$A^n = P \cdot \text{diag}(\alpha^n, \beta^n) \cdot P^{-1}.$$

Il reste à courageusement faire le calcul.

On obtient :

$$A^n = \frac{1}{\alpha - \beta} \begin{pmatrix} -\alpha^n \beta + \alpha \beta^n & \alpha^n - \beta^n \\ -\alpha \beta (\alpha^n - \beta^n) & \alpha^{1+n} - \beta^{1+n} \end{pmatrix}.$$

Et donc :

$$U_n = A^n \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\alpha - \beta} \begin{pmatrix} -\alpha^n(-1 + \beta) + (\alpha - 1)\beta^n \\ -\alpha^{1+n}(-1 + \beta) + (\alpha - 1)\beta^{1+n} \end{pmatrix}.$$

Comme $\alpha - \beta = \sqrt{5}$, et que $\alpha + \beta = 1$, on trouve finalement : $u_n = \frac{1}{\sqrt{5}} (\alpha^{n+1} - \beta^{n+1})$.

Remarque : Au programme officiel, on n'en voudra pas à l'étudiant (moyen) s'il ne sait pas traiter des récurrences d'ordre supérieur. La méthode peut pourtant se généraliser sans souci. Ainsi, par exemple, pour l'ordre $p = 3$, la matrice associée est d'ordre 3. Imaginons (une imagination très probable!) que dans un problème de concours en étant guidé, cela devient du domaine du possible. Toujours en travaillant à la marge du programme officiel, on pourrait imaginer que la matrice A n'est pas diagonalisable mais seulement trigonalisable. Ce cas est adapté pour obtenir les suites de $\mathbf{C}^{\mathbf{N}}$ qui vérifient une récurrence linéaire. Tout le problème est alors le calcul de A^n , ou plutôt T^n , si T est une matrice triangulaire semblable à A . La méthode reste fondamentalement la même. Et là aussi, si l'on est guidé, tout devient possible : *Sky is the limit !*

■ **Comment déterminer les termes généraux de plusieurs suites définies par une récurrence matricielle, cette matrice étant diagonalisable**

Ce genre d'exercice se rencontre aussi en TSI et à ne pas confondre avec la problématique de la **méthode 7.9**.

□ **Méthode 7.10.**— Par exemple, dans le cas de trois suites u, v et w , définies par :

$$X_{n+1} = \begin{pmatrix} u_{n+1} \\ v_{n+1} \\ w_{n+1} \end{pmatrix} = A \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} = AX_n,$$

on a : $X_n = A^n X_0$. Il reste à diagonaliser A , trouver P la matrice de passage vers la base de vecteurs propres. Puis on pose : $Y_n = P^{-1}X_n$. On calcule Y_n par la relation $Y_n = D^n Y_0$ et on revient à X_n par la relation $PY_n = X_n$.

Remarquer que l'on a besoin d'inverser P que si les conditions initiales sont explicitées.

Exemple : Soient (u_n) , (v_n) et (w_n) définies par $u_0 = -2, v_0 = 1, w_0 = 5$ et :

$$\forall n \in \mathbf{N}, \begin{cases} u_{n+1} = 4u_n - 3v_n - 3w_n \\ v_{n+1} = 3u_n - 2v_n - 3w_n \\ w_{n+1} = 3u_n - 3v_n - 2w_n \end{cases}.$$

Déterminer u_n, v_n et w_n en fonction de n .

On pose $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$ et $X_{n+1} = AX_n$. On a rapidement : $A = \begin{pmatrix} 4 & -3 & -3 \\ 3 & -2 & -3 \\ 3 & -3 & -2 \end{pmatrix}$. Et :

$$\chi_A(X) = X^3 - 3X + 2 = (X - 1)^2(X + 2).$$

Les valeurs propres sont -2 et 1 . On montre rapidement que $\dim E_{-2} = 1$ et $\dim E_1 = 2$. Donc comme $\dim E_{-2} + \dim E_1 = 3$, A est diagonalisable.

En prenant par exemple $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ et $D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, on a l'égalité :

$A = PDP^{-1}$. Donc : $A^n = PD^nP^{-1}$. Posons $Y_n = P^{-1}X_n$.

$$Y_{n+1} = DY_n \Rightarrow Y_n = D^n Y_0 \Rightarrow X_n = PY_n = PD^n Y_0.$$

Posons $(Y_0)^T = (\alpha \beta \gamma)$, alors l'égalité $PY_0 = X_0$ donne $\alpha = 8, \beta = -3$ et $\gamma = -7$. Finalement :

$$\begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 8 \\ -3 \\ -7 \end{pmatrix} = \begin{pmatrix} 8(-2)^n - 10 \\ 8(-2)^n - 7 \\ 8(-2)^n - 3 \end{pmatrix}.$$

Énoncé des exercices

■ Valeurs et vecteurs propres d'un endomorphisme ou d'une matrice

Exercice 7.1 : Soit $f \in \mathcal{L}(E)$ avec E de dimension finie tel que $f^3 - 5f^2 + 6f = 0$. Déterminer les valeurs propres possibles de f .

Exercice 7.2 : Déterminer l'ensemble des matrices $A \in \mathcal{M}_2(\mathbf{R})$ telles que $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ en soit un vecteur propre associé à la valeur propre -2 .

■ Conditions de diagonalisation

Exercice 7.3 : Réduire, lorsque cela est possible, dans $\mathcal{M}_2(\mathbf{R})$:

$$A_1 = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}, A_4 = \begin{pmatrix} m+1 & 1 \\ 2 & m \end{pmatrix}.$$

Exercice 7.4 : Réduire, lorsque cela est possible, dans $\mathcal{M}_2(\mathbf{C})$:

$$A_1 = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}, A_2 = \begin{pmatrix} -1 & 2i \\ -2i & 2 \end{pmatrix}, A_3 = \begin{pmatrix} -1+i & -1 \\ 1 & 1+i \end{pmatrix}, A_4 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Exercice 7.5 : On considère $A = \begin{pmatrix} 13 & -5 & -2 \\ -2 & 7 & -8 \\ -5 & 4 & 7 \end{pmatrix}$.

1. Calculer $(A - 9I_3)^3$. En déduire que A n'est pas diagonalisable.
2. Notons $f \in \mathcal{L}(\mathbf{R}^3)$ canoniquement associé à A et $g = f - 9\text{Id}_E$, montrer que $g^2 \neq 0$. Si \vec{u}_3 est un vecteur tel que $g^2(\vec{u}_3) \neq 0$, on pose $\vec{u}_2 = g(\vec{u}_3)$ et $\vec{u}_1 = g(\vec{u}_2)$. Montrer que $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est une base de \mathbf{R}^3 . Que peut-on dire de la matrice de f dans cette base?

D'après Concours Commun INP (ancien CCP)

Exercice 7.6 : Soit $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ 4 & 8 & -12 \end{pmatrix}$.

1. Calculer le rang de A . En déduire sans calcul le polynôme caractéristique de A .
2. Donner les éléments propres de A . A est-elle diagonalisable?

Exercice 7.7 : Déterminer $(a, b, c) \in \mathbf{R}^3$ pour que $A = \begin{pmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}$ soit diagonalisable.

D'après Concours Communs Mines-Telecom et Concours ENSEA, filière PSI

Exercice 7.8 : Pour $M \in \mathcal{M}_n(\mathbf{C})$, on pose $f(M) = M^T$.

Montrer que f est un endomorphisme et déterminer son spectre et ses sous-espaces propres.

L'endomorphisme f est-il diagonalisable ?

Déterminer alors $\text{Det}(f)$ et $\text{Tr}(f)$.

Exercice 7.9 : On considère l'application f définie sur $\mathbf{R}_2[X]$ par $f(P) = (2X + 1)P - (X^2 - 1)P'$.

1. Vérifier que f est un endomorphisme (en particulier que $f(P) \in \mathbf{R}_2[X]$.)

2. Déterminer sa matrice A dans sa base canonique et réduire A .

3. En déduire une base de $\mathbf{R}_2[X]$ formée de vecteurs propres de f et résoudre $f(P) = 1 + X^2$ en utilisant cette base.

■ Trigonalisation

Exercice 7.10 : Montrer que $A = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est trigonalisable mais non diagonalisable puis donner une base qui trigonalise A .

Exercice 7.11 : $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$ et $M_{\mathcal{B}}(\phi) = \begin{pmatrix} 7 & 3 & 2 & 5 \\ 10 & 7 & 4 & 9 \\ 9 & 6 & 4 & 8 \\ -15 & -9 & -5 & -12 \end{pmatrix}$.

1. $M_{\mathcal{B}}(\phi)$ est-elle diagonalisable dans \mathbf{R} ?

2. On note \vec{V}_1 un vecteur propre associé à la valeur propre la plus petite de $M_{\mathcal{B}}(\phi)$ et on note \vec{V}_2 un vecteur propre associé à sa plus grande valeur propre. On fixera \vec{V}_1 de telle manière que sa première composante soit 3 et on fixera \vec{V}_2 de telle manière que sa dernière composante soit -1 . Montrer que $\mathcal{B}' = \{\vec{V}_1, \vec{V}_2, \vec{e}_3, \vec{e}_4\}$ est une base de \mathbf{R}^4 .

3. Déterminer la matrice de passage P de \mathcal{B} à la base \mathcal{B}' . Calculer P^{-1} .

4. Déterminer $P^{-1}M_{\mathcal{B}}(\phi)P$ et l'écrire en blocs : $\begin{pmatrix} D_1 & C_1 \\ O_2 & B_1 \end{pmatrix}$.

5. Déterminer les valeurs propres de B_1 . On notera $\vec{w}_1(a, b)$ un vecteur propre de B_1 associé à la plus petite des valeurs propres de B_1 et de même $\vec{w}_2(c, d)$ un vecteur propre de B_1 associé à la plus grande valeur propre de B_1 . Déterminer ces vecteurs.

6. Soit $\vec{V}_3 = (0, 0, a, b)$ et $\vec{V}_4 = (0, 0, c, d)$, montrer que $\mathcal{B}'' = \{\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4\}$ est une base de \mathbf{R}^4 .

7. Déterminer la matrice $M_{\mathcal{B}''}(\phi)$. Que remarque t-on ?

■ Application de la diagonalisation pour calculer des puissances successives

Exercice 7.12 : On considère la matrice $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$.

1. A est-elle diagonalisable ?

2. Montrer que A est semblable à $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. En déduire A^n pour tout $n \in \mathbf{N}$.

Exercice 7.13 : Soit (u_n) la suite définie par $u_0 = 0$, $u_1 = -1$, $u_2 = 3$ et la relation :

$$\forall n \in \mathbf{N}, u_{n+3} = -2u_n + u_{n+1} + 2u_{n+2}.$$

1. On pose $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$. Déterminer une matrice A telle que $X_{n+1} = AX_n$.
2. Réduire A et en déduire X_n puis u_n en fonction de n .

Exercice 7.14 : Soit $A = \begin{pmatrix} -1 & 3 & 2 \\ -2 & 4 & 2 \\ -1 & 3 & 2 \end{pmatrix}$.

1. Diagonaliser A . On appelle D la matrice diagonale obtenue et P telle que : $D = P^{-1}AP$.
2. Déterminer $Y \in \mathcal{M}_3(\mathbf{R})$ telles que $Y^2 = D$. En déduire $X \in \mathcal{M}_3(\mathbf{R})$ telles que $X^2 = A$.

Indications

Ex. 7.1

On posera $f(\vec{x}) = \lambda \vec{x}$.

Ex. 7.2

Une mise en bouche, une !

Ex. 7.6

1. Comparer les colonnes.

Ex. 7.7

On commencera par mettre $\chi_A(X)$ sous forme factorisée.

Ex. 7.10

Au programme, il est dit que c'est bien de donner une indication. J'en donne deux. La première est d'aller voir la **méthode ??** pour se donner une idée. La seconde (à ne lire que si vraiment l'on est bloqué !) est de prendre une base dont les deux premiers vecteurs sont une base de $E_1(A)$ et le dernier par exemple $(0, 0, 1)$.

Ex. 7.12

2. Ici deux méthodes, on peut par exemple considérer f l'endomorphisme associé à A dans la base canonique (e_1, e_2, e_3) , puis déterminer une base $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ dans laquelle la matrice de f est A . On peut s'appuyer sur le fait que u_1 et u_2 sont forcément dans $E_1(A)$. Une autre méthode est de poser une matrice P (de passage) telle que ses deux premières colonnes soient \vec{u}_1 et \vec{u}_2 précédents et la troisième colonne de la forme (a, b, c) . Enfin, on a : $AP = PB$. À vous de continuer.

Ex. 7.13

Allez voir la **méthode ??**

Ex. 7.14


C'est la **méthode ??**. Ainsi si $Y^2 = D$ alors Y et D commutent !

Corrigé des exercices

Exercice 7.1

A FAIRE

Exercice 7.2

 On a une infinité de solutions.

On pose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. La relation $A \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -6 \\ 2 \end{pmatrix}$ donne pour solution $A = \begin{pmatrix} a & -6-3a \\ c & -2-3c \end{pmatrix}$, où $(a, c) \in \mathbf{R}^2$.

Exercice 7.3

A FAIRE


Exercice 7.4

- $\chi_{A_1}(X) = X^2 - 4X + 5$ et $\text{Sp}(A_1) = \{2+i, 2-i\}$, A_1 est diagonalisable.
- $\chi_{A_2}(X) = X^2 - X - 6$ et $\text{Sp}(A_2) = \{-2, 3\}$, A_2 est diagonalisable.
- $\chi_{A_3}(X) = X^2 - 2iX - 1$ et $\text{Sp}(A_3) = \{i\}$, A_3 n'est pas diagonalisable.
- $\chi_{A_4}(X) = X^2 - 2X + 2$ et $\text{Sp}(A_4) = \{1+i, 1-i\}$, A_4 est diagonalisable.


Exercice 7.5

1. Soit f l'endomorphisme canoniquement associé à A et $g = f - 9\text{Id}_E$, alors si B est la matrice de g dans la base canonique, on a :

$$B = A - 9I_3 = \begin{pmatrix} 4 & -5 & -2 \\ -2 & -2 & -8 \\ -5 & 4 & -2 \end{pmatrix}, B^2 = 9 \begin{pmatrix} 4 & -2 & 4 \\ 4 & -2 & 4 \\ -2 & 1 & -2 \end{pmatrix}, B^3 = 0.$$

 Retenir ce raisonnement, il revient souvent !

$(X-9)^3$ est un polynôme annulateur de A , donc 9 est la seule valeur propre possible de A . A n'est donc pas diagonalisable, sinon elle serait semblable, puis égale, à $9I_3$, ce qui n'est pas le cas.

 g est nilpotente d'ordre 3, on extrait une base de \mathbf{R}^3 à partir d'images successives d'un vecteur par g , cette idée a été vue dans un chapitre précédent.

2. Choisissons $\vec{u}_3 \notin \text{Ker}(g^2)$, posons $\vec{u}_2 = g(\vec{u}_3)$ puis $\vec{u}_1 = g(\vec{u}_2)$.

Montrons que $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est une base de \mathbf{R}^3 .

Supposons (1) $\alpha\vec{u}_1 + \beta\vec{u}_2 + \gamma\vec{u}_3 = 0_{\mathbf{R}^3}$ soit $\alpha g^2(\vec{u}_3) + \beta g(\vec{u}_3) + \gamma\vec{u}_3 = 0$.

Composons par g^2 : comme $g^3 = 0_{\mathcal{L}(\mathbf{R}^3)}$, on obtient $\gamma g^2(\vec{u}_3) = 0$.


Or $\vec{u}_3 \notin \text{Ker}(g^2)$ donc $\gamma = 0$.

Composons alors (1) par g ; on obtient de même $\beta = 0$ puis on a alors $\alpha = 0$.

$(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est une base dans laquelle la matrice de g est la matrice triangulaire

$$\theta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \text{ celle de } f \text{ la matrice triangulaire : } T = \begin{pmatrix} 9 & 1 & 0 \\ 0 & 9 & 1 \\ 0 & 0 & 9 \end{pmatrix}.$$

Exercice 7.6

 On remarque que la deuxième colonne est le double de la première et que la troisième colonne est l'opposée du triple de la première. Dit ainsi, on dirait presque une devinette !

1. La matrice est de rang 1.

On en déduit que le noyau, c'est-à-dire le sous-espace propre relatif à 0, est de dimension 2 ; par ailleurs, le polynôme caractéristique est de degré 3 et multiple de X^2 , donc il est scindé dans \mathbf{R} .

La somme des valeurs propres, en tenant compte de leur ordre de multiplicité, étant égale à la trace, on en déduit que -7 est la dernière valeur propre.

Le polynôme caractéristique de A est donc $\chi_A(X) = X^2(X+7)$.

2. Le noyau est le sous-espace propre associé à la valeur propre 0, et est le plan d'équation $x + 2y - 3z = 0$ (on obtient trois fois la même équation).

Le sous-espace propre associé à -7 est $\vec{V} = (1, 2, 4)$.

Le sous-espace propre associé à la valeur propre double étant de dimension 2, A est diagonalisable.

Exercice 7.7

Sans souci, on arrive à $\chi_A(X) = (X - 1)^2(X - c)$.

► Si $c \neq 1$, $E_c(A)$ est un sous-espace propre de dimension 1. Il reste à déterminer la dimension de $E_1(A)$. On aboutit à $z = 0$, $y = y$ et $x + ay = x$. Donc si $a = 0$, $E_1(A)$ est le plan d'équation $z = 0$. Dans ce cas $\dim E_c(A) + \dim E_1(A) = 3$ et A est diagonalisable. Si $a \neq 0$, on a $y = z = 0$ et $E_1(A)$ est une droite. Dans ce cas, A n'est pas diagonalisable.

► Si $c = 1$, 1 est valeur propre triple et A n'est diagonalisable que si $A = I_3$. Or ce n'est pas le cas. Donc A n'est pas diagonalisable.

🔗 Finalement, A est diagonalisable ssi $c \neq 1$ et $a = 0$.

Exercice 7.8

Exercice 7.9

- 1.
- 2.
- 3.

Exercice 7.10

$$\chi_A(X) = \begin{vmatrix} X & 0 & -1 \\ -2 & X-1 & 0 \\ 0 & 0 & X-1 \end{vmatrix} = X \begin{vmatrix} -X & 1 \\ 2 & 1-X \end{vmatrix},$$

c'est-à-dire : $\chi_A(X) = X(X^2 - 2X + 1) = X(X - 1)^2$. Puis :

$$(x, y, z) \in E_0(A) \iff \begin{cases} z = 0 \\ 2x + y = 0 \\ z = 0 \end{cases} \Rightarrow E_0(A) = \text{Vect}(u_1), \vec{u}_1 = (1, -2, 0).$$

$$(x, y, z) \in E_1(A) \iff \begin{cases} z = 0 \\ 2x = 0 \end{cases} \Rightarrow E_1(A) = \text{Vect}(\vec{u}_2) \text{ avec } \vec{u}_2 = (0, 1, 0).$$

🔗 Le sous-espace propre associé à $\lambda = 1$, racine double de χ_A , étant de dimension 1, la matrice n'est pas diagonalisable. On est en plein de la méthode ??.

Complétons la famille libre (\vec{u}_1, \vec{u}_2) en une base $\mathcal{B} = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$ par exemple avec le vecteur $\vec{u}_3 = (0, 0, 1)$. Notons f l'endomorphisme de matrice A dans la base canonique de \mathbf{R}^3 . On a : $f(\vec{u}_3) - \vec{u}_3 = \vec{u}_1 + 2\vec{u}_2$ donc la matrice de f

dans la base \mathcal{B} est : $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$.

Exercice 7.11

1. $\chi_A(X) = X^4 - 6X^3 + 13X^2 - 12X + 4 = (X - 1)^2(X - 2)^2$.

Et : $E_2(A) = \text{Vect}((0, 1, 1, -1)) = \text{Vect}(\vec{V}_2)$, $E_1(A) = \text{Vect}((3, 2, 3, -6)) = \text{Vect}(\vec{V}_1)$.

Comme la somme des dimensions des sous-espaces propres est strictement inférieur à $\dim \mathbf{R}^4$, A n'est pas diagonalisable.

🔗 On peut calculer $\chi_A(X)$ par développement selon une rangée.

☞ P étant triangulaire inférieure, il en est de même de P^{-1} .

☞ Une matrice 4×4 est une matrice tout terrain, donc on est rassuré !

2. On voit par exemple : $\text{Det}(\vec{V}_1, \vec{V}_2, \vec{e}_3, \vec{e}_4) \neq 0$.

$$3. P = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 \\ 6 & -1 & 0 & 1 \end{pmatrix} \Rightarrow P^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 3 & 0 & 0 \\ -1 & -3 & 3 & 0 \\ 4 & 3 & 0 & 3 \end{pmatrix}.$$

$$4. \text{Après un calcul (in)surmontable : } A' = \frac{1}{3} \begin{pmatrix} 3 & 0 & 2 & 5 \\ 0 & 6 & 8 & 17 \\ 0 & 0 & -2 & -8 \\ 0 & 0 & 5 & 11 \end{pmatrix}.$$

☞ L'idée fondamentale ici est de partir d'une matrice triangulaire supérieure par blocs que l'on transforme en matrice triangulaire supérieure tout court en diagonalisant (ou trigonalisant) un des blocs (qui représente la réduction de ϕ à un sous-espace vectoriel stable).

☞ On part de $\vec{V}_3 = -8\vec{e}_3 + 5\vec{e}_4$ et $\vec{V}_4 = \vec{e}_3 - \vec{e}_4$.

☞ On remarque qu'elle est triangulaire supérieure.

5. On part de $B_1 = \frac{1}{3} \begin{pmatrix} -2 & -8 \\ 5 & 11 \end{pmatrix}$. Le spectre de B_1 est $\{1, 2\}$.

On peut poser $\vec{w}_1(-8, 5)$ et $\vec{w}_2(1, -1)$ après calculs.

6. Encore une fois, on peut calculer $\text{Det}(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4) \neq 0$.

7. On peut écrire la matrice de ϕ dans la nouvelle base en inversant une nouvelle matrice de passage et en procédant comme à la quatrième question. Puis faire des produits par blocs.

On peut aussi (et c'est ce que nous allons faire) calculer les images de $\vec{V}_1, \vec{V}_2, \vec{V}_3$ et \vec{V}_4 par rapport à eux mêmes. On a déjà : $\phi(\vec{V}_1) = \vec{V}_1$ et $\phi(\vec{V}_2) = 2\vec{V}_2$. Puis $\phi(\vec{V}_3) = -8\phi(\vec{e}_3) + 5\phi(\vec{e}_4)$. On part de $3\phi(\vec{V}_3) = -8\phi(3\vec{e}_3) + 5\phi(3\vec{e}_4) = -8(2\vec{V}_1 + 8\vec{V}_2 - 2\vec{e}_3 + 5\vec{e}_4) + 5(5\vec{V}_1 + 17\vec{V}_2 - 8\vec{e}_3 + 11\vec{e}_4) = 9\vec{V}_1 + 21\vec{V}_2 + 3\vec{V}_3$, en utilisant $3\vec{e}_3 = -\vec{V}_3 - 5\vec{V}_4$ et $3\vec{e}_4 = -\vec{V}_3 - 8\vec{V}_4$. Ainsi : $\phi(\vec{V}_3) = 3\vec{V}_1 + 7\vec{V}_2 + \vec{V}_3$. De même, on trouve, après calculs, $\phi(\vec{V}_4) = -\vec{V}_1 - 3\vec{V}_2 + 2\vec{V}_4$.

La matrice de ϕ dans notre nouvelle base est :
$$\begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & 2 & 7 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Exercice 7.12

1. Par un calcul facile (développer selon la première rangée), on obtient :

$$\chi_A = (X - 1)^3.$$

L'unique valeur propre est donc 1 ; l'espace propre E_1 associé a pour équation $y + z = 0$. Il est de dimension 2.

Comme $\dim E_1(A) = 2 < 3 = m_1$, la matrice A est non diagonalisable.

☞ Remarquer l'analyse avec la méthode ?? dans le cas de deux valeurs propres distinctes. Ici, il n'y en a qu'une mais la matrice triangulaire finale est donnée, ouf !

☞ On peut aussi écrire $B = I_3 + J$ et utiliser le binôme de Newton. Allez voir revoir la méthode concernée !

2. On peut prendre pour base de $E_1(A)$: $u_1(1, 0, 0), u_2(0, -1, 1)$.

Soit $P = \begin{pmatrix} 1 & 0 & a \\ 0 & -1 & b \\ 0 & 1 & c \end{pmatrix}$, $A = P.B.P^{-1} \Rightarrow AP = PB$.

Un calcul rapide donne : $b + c = 1$. On peut prendre : $a = c = 0, b = 1$.

Les matrices A et B sont donc semblables.

On peut montrer par récurrence sur $n \in \mathbf{N}$ que $B^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$.

Alors, pour $n \in \mathbf{N}$, $A^n = P.B^n.P^{-1}$; le calcul donne successivement :

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \text{ et } A^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - n & -n \\ 0 & n & 1 + n \end{pmatrix}.$$


Exercice 7.13


1. On a : $X_{n+1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} X_n$.

2. On trouve : $\chi_A(X) = (X - 2)(X - 1)(X + 1)$ et donc le spectre est $\{-1, 1, 2\}$. On trouve : $P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \\ 1 & 1 & 4 \end{pmatrix}$ et $D = \text{diag}(-1, 1, 2)$. On pose

$X_n = PY_n$, où $Y_n = D^n Y_0$. Il reste à calculer Y_0 par la formule $P^{-1}X_0 = Y_0$. On trouve :

$$Y_0 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \Rightarrow X_n = PD^n \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} (-1)^n - 2 + 2^n \\ (-1)^{n+1} - 2 + 2^{n+1} \\ (-1)^n - 2 + 2^{n+2} \end{pmatrix}.$$

 A est diagonalisable dans \mathbf{R} car ses trois valeurs propres sont réelles et distinctes.

 On en déduit : $u_n = (-1)^n - 2 + 2^n$.

Exercice 7.14

1. $\chi_A(X) = X(X - 1)(X - 4)$, $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, $P = \begin{pmatrix} -1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

2. Posons $Y = (y_{i,j}) \in \mathcal{M}_3(\mathbf{R})$. Si $Y^2 = D$ alors $YD = Y^3 = DY$.

$$YD = DY \Rightarrow [i \neq j \Rightarrow y_{i,j} = 0].$$


La matrice Y est donc diagonale et $Y^2 = D$ permet de conclure :


$$y_{1,1} = 0, y_{2,2} = \pm 1, y_{3,3} = \pm 2.$$

Il reste à écrire que les matrices X telles que $X^2 = A$.

$Y^2 = D \Leftrightarrow PY^2P^{-1} = A$. En posant $X = PYP^{-1}$, on a $X^2 = A$.

Les solutions sont les quatre matrices : $P \begin{pmatrix} 0 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 2 \end{pmatrix} P^{-1}$.

 Les calculs sont classiques, laissés au lecteur.

 Il y a quatre matrices Y possibles.

Chapitre 8

Intégration sur un intervalle

■ Objectifs

■ Les incontournables :

- ▶ connaître les primitives usuelles.
- ▶ Savoir ce que signifie « fonction continue par morceaux » sur un intervalle borné ou non.
- ▶ Savoir prouver la nature d'une intégrale généralisée.
- ▶ Savoir calculer une intégrale définie ou une intégrale généralisée convergente directement ou avec un changement de variable ou une intégration par parties en prenant les précautions d'usage.
- ▶ Connaître les théorèmes de comparaison pour l'intégrabilité des fonctions.
- ▶ Connaître sans confusion le vocabulaire un peu foisonnant (intégrales généralisées, convergentes, absolument convergentes, fonctions intégrables sur un intervalle).

■ Et plus si affinités ...

- ▶ Connaître et savoir exploiter la formule de Taylor avec reste intégral.

Résumé de cours

Dans ce chapitre, les fonctions considérées sont définies sur un intervalle I de \mathbf{R} , non vide et non réduit à un point, et à valeurs dans \mathbf{R} ou \mathbf{C} .

■ Révision : intégrale d'une fonction continue sur le segment $I = [a, b]$.

Sommes de Riemann d'une fonction continue

Définition : Soit $f \in \mathcal{C}([a, b], \mathbf{R})$. Les **sommes de Riemann** $(I_n(f))_{n \in \mathbf{N}^*}$ et $(J_n(f))_{n \in \mathbf{N}^*}$ de f sont les suites définies par : $\forall n \in \mathbf{N}^*$, $I_n(f) = \frac{b-a}{n} \sum_{k=0}^{n-1} f(a_k)$, $J_n(f) = \frac{b-a}{n} \sum_{k=1}^n f(a_k)$, où on a noté pour tout $k \in \llbracket 0, n \rrbracket$, $a_k = a + k \cdot \frac{b-a}{n}$.

Théorème 8.1.— Sommes de Riemann —. Soit $f \in \mathcal{C}([a, b], \mathbf{R})$. Les sommes de Riemann de f $(I_n(f))_{n \in \mathbf{N}^*}$ et $(J_n(f))_{n \in \mathbf{N}^*}$ sont des suites convergentes et :

$$\lim_{n \rightarrow \infty} I_n(f) = J_n(f) = \int_a^b f(t) dt.$$

De plus, si f est croissante (resp. décroissante) sur $[a, b]$, $I_n(f) \leq \int_a^b f(t) dt \leq J_n(f)$ (resp. $J_n(f) \leq \int_a^b f(t) dt \leq I_n(f)$).

$\int_a^b f(t) dt$ se définit donc comme la limite d'une somme de Riemann.

Remarque : Comme $I = [a, b]$, on peut parfois noter $\int_a^b f(t) dt$ plus simplement : $\int_I f$.

Mise en œuvre : exercice 01, exercice 02 (question 2).

Proposition 8.2.— Linéarité de l'intégrale —.

$$\forall (f, g) \in [\mathcal{C}([a, b], \mathbf{R})]^2, \forall (\lambda, \mu) \in \mathbf{K}^2, \int_a^b (\lambda f(t) + \mu g(t)) dt = \lambda \int_a^b f(t) dt + \mu \int_a^b g(t) dt.$$

Proposition 8.3.— Pour tout $f \in \mathcal{C}([a, b], \mathbf{R})$,

$$f \geq 0 \Rightarrow \int_a^b f(t) dt \geq 0, \quad \left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

Proposition 8.4.— On a aussi la relation de Chasles, pour toute fonction $f \in \mathcal{C}([a, b], \mathbf{R})$ et pour tout réel $c \in [a, b]$:

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt.$$

Mise en œuvre : exercice 02.

Proposition 8.5.— Si $f \in \mathcal{C}([a, b], \mathbf{R})$ et en posant $m = \inf_{x \in [a, b]} f(x)$, $M = \sup_{t \in [a, b]} f(t)$, on peut alors en déduire la double inégalité :

$$m(b-a) \leq \int_a^b f(t) dt \leq M(b-a).$$

Proposition 8.6.— Pour tout $(f, g) \in [\mathcal{C}([a, b], \mathbf{R})]^2$,

$$\forall t \in [a, b], f(t) \geq g(t) \Rightarrow \int_a^b f(t) dt \geq \int_a^b g(t) dt.$$

Intégrale fonction d'une extrémité de l'intervalle

Soit $f \in \mathcal{C}([a, b], \mathbf{R})$, à tout $x \in [a, b]$, on associe $F(x) = \int_a^x f(t) dt$.

Proposition 8.7.— La fonction $F : x \mapsto \int_a^x f(t) dt$ est de classe \mathcal{C}^1 sur $[a, b]$. On a : $\forall x \in [a, b]$, $F'(x) = f(x)$. F est donc une primitive de f . C'est l'unique primitive de f sur $[a, b]$ qui s'annule en $x = a$.

Mise en œuvre : exercice 03.

Théorème fondamental du calcul intégral

Si les propriétés de l'intégrale des fonctions continues permettent de montrer l'existence de primitives, inversement, la connaissance d'une primitive permet de calculer l'intégrale d'une fonction continue.

Théorème 8.8.— Théorème fondamental du calcul intégral —.

Soit $f \in \mathcal{C}([a, b], \mathbf{R})$ et F une primitive *quelconque* de f sur I , alors

$$\int_a^b f(t) dt = \left[F(t) \right]_a^b = F(b) - F(a).$$

En particulier, une fonction f de classe \mathcal{C}^1 est une primitive de sa dérivée.

Corollaire 8.9.— Soit $f \in \mathcal{C}^1(I, \mathbf{R})$, $c \in I$, alors

$$\forall x \in I, \quad f(x) - f(c) = \int_c^x f'(t) dt.$$

Calculs d'intégrales et de primitives

Théorème 8.10.— Intégration par parties —. Soit $(u, v) \in \mathcal{C}^1([a, b], \mathbf{R})^2$, alors

$$\int_a^b u'(t) \times v(t) dt = \left[u(t) \times v(t) \right]_a^b - \int_a^b u(t) \times v'(t) dt.$$

Remarque : comme les primitives d'une fonction continue f sont de la forme $F(x) = A + \int_a^x f(t) dt$, ces formules de calcul d'intégrales définies, s'appliquent aussi au calcul de primitives. Par exemple, si u et v sont de classe \mathcal{C}^1 sur un intervalle I , alors (à une constante additive près) :

$$\int u'(x)v(x) dx = u(x)v(x) - \int u(x)v'(x) dx.$$

Quelques primitives usuelles

Fonction	Primitive	Fonction	Primitive
$(x-b)^\alpha$	$\frac{(x-b)^{\alpha+1}}{\alpha+1} + C$	$\frac{1}{\sin^2 x}$	$-\frac{1}{\tan x} + C$
e^{cx}	$\frac{1}{c}e^{cx} + C$	$\frac{1}{\sqrt{x^2-1}}$	$\ln \left x + \sqrt{x^2-1} \right + C$
$\sin ax$	$-\frac{1}{a} \cos ax + C$	$\frac{1}{\sqrt{1+x^2}}$	$\ln \left x + \sqrt{x^2+1} \right + C$
$\cos ax$	$\frac{1}{a} \sin ax + C$	$\frac{1}{\sqrt{a^2-x^2}}$	$\frac{a}{ a } \operatorname{Arcsin} \frac{x}{a} + C$
$\frac{1}{\cos^2 x}$	$\tan x + C$	$\frac{1}{1-x^2}$	$\frac{1}{2} \ln \left \frac{1+x}{1-x} \right + C$
$\frac{1}{x}$	$\ln x + C$	$\frac{1}{a^2+x^2}$	$\frac{1}{a} \operatorname{Arctan} \frac{x}{a} + C$
$\tan x$	$-\ln \cos x + C$	$\ln x$	$x \ln x - x + C$

Dans ce tableau, b est un réel, a est un réel non nul, α est un rationnel différent de -1 , $c \in \mathbf{R}^*$ ou $c \in \mathbf{C}^*$ et C est une constante réelle (ou complexe).

Théorème 8.11.— Changement de variable —. Soit I, J deux segments de \mathbf{R} , $\varphi : J \rightarrow I$ une fonction strictement monotone de classe \mathcal{C}^1 , $f : I \rightarrow \mathbf{R}$ une fonction continue et $(\alpha, \beta) \in J^2$. Alors :

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(t) dt = \int_{\alpha}^{\beta} f(\varphi(u)) \varphi'(u) du.$$

Pour (ré)apprendre à utiliser cette formule, allez à la **méthode 8.1** et à la **méthode 8.2**.

Il est temps de calculer plein de primitives et d'intégrales connaissant les différentes méthodes. Au travail !

Mise en œuvre : exercice 04 à exercice 09.

Théorème 8.12.— Formule de Taylor avec reste intégral —. Soit f de classe \mathcal{C}^{n+1} sur un segment I , alors pour tout $(a, x) \in I^2$,

$$f(x) = \sum_{k=0}^n \frac{1}{k!} (x-a)^k f^{(k)}(a) + \int_a^x \frac{1}{n!} (x-t)^n f^{(n+1)}(t) dt.$$

Une conséquence de la formule de Taylor avec reste intégral est la formule de Taylor-Young, que vous connaissez, j'espère !

Mise en œuvre : exercice 10.

Proposition 8.13.— Extension aux fonctions continues par morceaux et à valeurs dans \mathbf{C}
Si $f \in \mathcal{C}_m([a, b], \mathbf{C})$ alors on peut définir l'intégrale de f sur $[a, b]$:

$$\int_a^b f(t) dt = \int_a^b \Re(f)(t) dt + i \int_a^b \Im(f)(t) dt.$$

Remarque : L'intégrale sur $[a, b]$ d'une fonction à valeurs dans \mathbf{C} vérifie l'inégalité de la moyenne et la relation de Chasles. Si f est continue sur $[a, b]$, elle vérifie la formule de changement de variable et si f et g sont de classe \mathcal{C}^1 sur $[a, b]$, la formule d'intégration par parties.

Dans toute la suite, sauf précision éventuelle, I est de façon indifférenciée soit $[a, b[$ avec b réel ou $+\infty$, soit $I =]a, b]$ avec $a = -\infty$ ou a réel, soit $I = \mathbf{R}$.

■ Intégrales sur un intervalle

Définition : Convergence d'une intégrale sur un intervalle semi-ouvert —.

Si f est une application continue sur $[a, b[$ (b réel ou $+\infty$), l'intégrale (appelée souvent **intégrale généralisée**) $\int_a^b f(t) dt$ est **convergente** lorsque $\int_a^x f(t) dt$ a une limite finie ℓ quand x tend vers b , $x \in [a, b[$.

Dans ce cas, on pose :

$$\int_a^b f(t) dt = \lim_{x \rightarrow b} \int_a^x f(t) dt = \ell$$

Dans le cas contraire, l'intégrale est dite **divergente**.

La définition s'étend naturellement aux intervalles de la forme $]a, b]$ (a réel ou $-\infty$).

Remarque : Si f est une fonction continue sur le segment $[a, b]$, alors cette définition est cohérente avec celle de l'intégrale définie $\int_a^b f(t) dt$.

Mise en œuvre : exercice 11.

Proposition 8.14.— Comparaison par la relation inférieur ou égal —.

On suppose que f et g sont continues sur $I = [a, b[$ ou $I =]a, b]$ et sont à valeurs positives sur I .

- Si : $\forall x \in I, f(x) \leq g(x)$ alors l'intégrale $\int_a^b f(t) dt$ converge si $\int_a^b g(t) dt$ converge.
- Si $f \leq g$ seulement au voisinage de b et si $I = [a, b[$, on a encore le résultat.
- Si $f \leq g$ seulement au voisinage de a et si $I =]a, b]$, on a encore le résultat.

Proposition 8.15.— Intégrales de Riemann —.

L'intégrale de Riemann $\int_1^{+\infty} \frac{dt}{t^\alpha}$ est convergente si et seulement si $\alpha > 1$. Dans ce cas :

$$\int_1^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{\alpha - 1}.$$

L'intégrale de Riemann $\int_0^1 \frac{dt}{t^\alpha}$ est convergente si et seulement si $\alpha < 1$. Dans ce cas :

$$\int_0^1 \frac{dt}{t^\alpha} = \frac{1}{1 - \alpha}.$$

Remarque : L'intégrale de Riemann $\int_0^{+\infty} \frac{dt}{t^\alpha}$ n'est convergente pour aucune valeur de α .

Proposition 8.16.— Intégrales de la fonction logarithme népérien —.

$\int_0^1 \ln t dt$ est convergente et vaut -1 et $\int_1^{+\infty} \ln t dt$ est divergente.

Proposition 8.17.— Intégrales des fonctions exponentielles —.

L'intégrale $\int_0^{+\infty} \exp(-\alpha t) dt$ est convergente si et seulement si $\alpha > 0$. Dans ce cas, elle vaut $\frac{1}{\alpha}$.

Proposition 8.18.— Extension de la formule de Chasles —.

L'intégrale généralisée $\int_a^b f(t) dt$ est convergente si et seulement s'il existe c de $[a, b[$ tel que $\int_c^b f(t) dt$ est convergente. Dans ce cas, $\int_c^b f(t) dt$ est convergente pour tout $c \in [a, b[$, et

$$\forall c \in [a, b[, \int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$$

Proposition 8.19.— Extension de la formule de changement de variable —.

Étant données une fonction f **continue sur** $[a, b[$ et une fonction ϕ **strictement croissante** et de classe \mathcal{C}^1 sur $[\alpha, \beta]$, les intégrales $\int_a^b f(t) dt$ et $\int_\alpha^\beta f(\phi(u))\phi'(u) du$ avec

$$a = \lim_{t \rightarrow \alpha} \phi(u) \text{ et } b = \lim_{t \rightarrow \beta} \phi(u)$$

sont de même nature et égales en cas de convergence.

On adapte sans souci au cas où ϕ est **strictement décroissante** ou au cas où f est **continue sur** $]a, b]$.

Remarque : Quant à la formule d'intégration par parties, dans la pratique, on l'applique non pas sur $[a, b[$ ou sur $]a, b]$ mais sur $[\epsilon, X]$ et l'on fait tendre ϵ vers a et (ou) X vers b , selon que l'intégrale est doublement généralisée (voir plus loin) ou non.

Mise en œuvre : exercice 12, exercice 13.

Définition : Intégrale doublement généralisée —. Si f est continue sur $]a, b[$, $\int_a^b f(t) dt$ est **convergente** si et seulement s'il existe un réel c de $]a, b[$ pour lequel $\int_a^c f(t) dt$ et $\int_c^b f(t) dt$ sont convergentes. Dans ce cas, on pose :

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt.$$

Dans le cas contraire, l'intégrale est dite **divergente**.

Exemple : $\int_{-1}^1 \frac{1}{\sqrt{1-t^2}} dt$ est doublement généralisée. En effet, la fonction $t \mapsto \frac{1}{\sqrt{1-t^2}}$ est continue sur $] -1, 1[$, mais non prolongeable aux bornes de cet intervalle. De plus, elle est convergente, puisque

$$\int_0^x \frac{1}{\sqrt{1-t^2}} dt = \arcsin x \text{ et } \int_{-x}^0 \frac{1}{\sqrt{1-t^2}} dt = \arcsin x$$

admettent toutes deux une limite quand x tend vers 1 par valeurs inférieures.

Remarque : d'après la **proposition 8.18**, s'il existe un tel réel c , alors tous les réels de $]a, b[$ conviennent.

Il arrive que l'on définisse une intégrale doublement généralisée en additionnant plusieurs intégrales (simplement) généralisées sur des intervalles contigus.

Proposition 8.20.— Cas d'une fonction complexe —. Si f est une fonction à valeurs complexes continue sur un intervalle I , l'intégrale généralisée $\int_{\inf I}^{\sup I} f(t) dt$ est convergente si et seulement si

$$\int_{\inf I}^{\sup I} \Re f(t) dt \text{ et } \int_{\inf I}^{\sup I} \Im f(t) dt$$

sont convergentes. Dans ce cas :

$$\int_{\inf I}^{\sup I} f(t) dt = \int_{\inf I}^{\sup I} \Re f(t) dt + i \int_{\inf I}^{\sup I} \Im f(t) dt$$

Proposition 8.21.— Linéarité —. Ici $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} . L'ensemble des applications f continues sur un intervalle I telles que $\int_{\inf I}^{\sup I} f(t) dt$ converge est un sous-espace vectoriel de $\mathcal{C}([a, b], \mathbf{K})$. De plus, si f et g sont dans cet ensemble et si λ et μ sont deux scalaires :

$$\int_{\inf I}^{\sup I} (\lambda f + \mu g)(t) dt = \lambda \int_{\inf I}^{\sup I} f(t) dt + \mu \int_{\inf I}^{\sup I} g(t) dt.$$

■ Intégrabilité d'une fonction continue sur I

Définition : Soit f une application à valeurs dans \mathbf{R} ou \mathbf{C} , continue sur un intervalle I . Si

$$\int_{\inf I}^{\sup I} |f(t)| dt$$

est convergente, on dit que l'intégrale $\int_{\inf I}^{\sup I} f(t) dt$ est **absolument intégrable**.

Définition : Une fonction f est dite **intégrable** sur I si elle est continue sur I et son intégrale sur I est absolument convergente.

Exemple : $t \mapsto \frac{\exp(it)}{t^2} dt$ est intégrable sur $[1, +\infty[$ car :

$$\left| \frac{\exp(it)}{t^2} \right| = \frac{1}{t^2}$$

et l'intégrale $\int_1^{+\infty} \frac{1}{t^2} dt$ est convergente car c'est une intégrale de Riemann convergente (voir la fondamentale **proposition 8.15**).

Proposition 8.22.— Si f est intégrable sur I alors $\int_I f(t) dt$ existe.

A t-on la réciproque ?

Mise en œuvre : exercice 14.

Proposition 8.23.— **Intégrale nulle d'une fonction continue, positive et intégrable** —.

Soit f une fonction continue et intégrable sur un intervalle I .

$$\int_I |f| = 0 \text{ si et seulement si } f \text{ est nulle.}$$

Proposition 8.24.— **Linéarité de $f \mapsto \int_I f(t) dt$** —.

On suppose que f et g sont continues et intégrables sur I et $(\lambda, \mu) \in \mathbf{K}^2$ alors :

$$\int_I (\lambda f(t) + \mu g(t)) dt = \lambda \int_I f(t) dt + \mu \int_I g(t) dt.$$

Proposition 8.25.— **Espace vectoriel $L_1(I)$** —. L'ensemble des fonctions, à valeurs dans \mathbf{K} , continues et intégrables sur un intervalle I est un \mathbf{K} -espace vectoriel, (noté généralement $L_1(I)$).

Proposition 8.26.— **Inégalité de la moyenne** —.

f étant une application intégrable sur un intervalle I :

$$\left| \int_I f \right| \leq \int_I |f|.$$

Proposition 8.27.— **Comparaison de fonctions et intégrabilité** —.

On suppose que f et g sont continues sur $I = [a, b]$, où b est éventuellement $+\infty$.

- Si $|f| \leq |g|$, alors l'intégrabilité de g sur I implique celle de f .
- Si $f(t) = O(g(t))$, quand $t \rightarrow b$ alors l'intégrabilité de g sur I implique celle de f .

Remarque : On peut dans le dernier item remplacer O par o .

Proposition 8.28.— **Comparaison par la relation d'équivalence** —. On suppose que f et g sont continues sur $I = [a, b]$, où b est éventuellement $+\infty$.

Si $f(t) \sim g(t)$ quand $t \rightarrow b$, alors l'intégrabilité de f sur I est équivalente à celle de g .

par la suite, on fera référence à ces résultats en invoquant les « théorèmes de comparaison ».

Remarque : Dans le cas où $I =]a, b]$, où a vaut éventuellement $-\infty$, on adapte la **proposition 8.27** et la **proposition 8.28** et on compare f et g quand $t \rightarrow a$.

Mise en œuvre : exercice 15, exercice 16, exercice 17.

Méthodes

■ Révision : méthodes de calcul de primitives ou d'intégrales sur un segment

On reprend les notations de la formule de changement de variable, c'est la **proposition 8.11**.

□ Méthode 8.1.— Comment et dans quelle circonstance on utilise la formule de changement de variable de la droite vers la gauche

Quand on reconnaît sous une intégrale I la forme $f(\phi(u))\phi'(u)$. Dans ce cas, on pose $t = \phi(u)$, puis on calcule $\phi(\alpha) = a$ et $\phi(\beta) = b$. On vérifie que ϕ est bien de classe \mathcal{C}^1 sur l'intervalle d'extrémités α et β . Puis on remplace $f(\phi(u))$ par $f(t)$ et $\phi'(u)du$ par dt et il reste à intégrer.

Exemple : Calculer l'intégrale

$$I = \int_0^{\frac{\pi}{2}} \sin^2 u \cos u \, du.$$

Faisons fonctionner la formule de la droite vers la gauche. On pose $\phi(u) = \sin u$ et $f(t) = t^2$. Ici $\alpha = 0$ et $\beta = \pi/2$, et $a = 0$ et $b = 1$. On a donc, en utilisant la formule de changement de variable,

$$I = \int_0^1 t^2 \, dt = \frac{1}{3}.$$

Remarque : de cette méthode de changement de variable, on peut établir quelques formules d'intégration rapides, à utiliser sans modération :

$$\begin{aligned} \int u(x)u'(x) \, dx &= \frac{1}{2}u^2(x), \quad \int u^\alpha(x)u'(x) \, dx = \frac{u^{\alpha+1}}{\alpha+1} \text{ avec } \alpha \neq -1, \quad \int \frac{u'(x)}{u(x)} \, dx = \ln |u(x)|, \\ \int \frac{u'(x)}{2\sqrt{u(x)}} \, dx &= \sqrt{u(x)}, \quad \int \exp(u(x))u'(x) \, dx = \exp(u(x)), \quad \int u'(v(x))v'(x) \, dx = u(v(x)), \\ \int \sin(u(x))u'(x) \, dx &= -\cos(u(x)), \quad \int \cos(u(x))u'(x) \, dx = \sin(u(x)) \end{aligned}$$

Ici, les fonctions u et v sont de classe \mathcal{C}^1 sur leur intervalle de définition et toutes ces formules sont valables à une constante additive près.

□ Méthode 8.2.— Comment et dans quelle circonstance on utilise la formule de changement de variable de la gauche vers la droite

Pour rendre l'intégrande plus facile à intégrer et le plus souvent en utilisant un changement de variable intuitif ou donné. Dans $I = \int_a^b f(t) \, dt$, on remplace t par $\phi(u)$ et dt par $\phi'(u) \, du$ puis on cherche deux valeurs α et β telles que $\phi(\alpha) = a$ et $\phi(\beta) = b$ et on vérifie que ϕ est de classe \mathcal{C}^1 sur l'intervalle d'extrémités α et β . Enfin, on écrit alors

$$I = \int_\alpha^\beta f(\phi(u))\phi'(u) \, du$$

Exemple : Calculer $J = \int_{-1}^{1/2} \sqrt{1-t^2} dt$ en posons $t = \sin u$.

Donc, ici $\phi(u) = \sin u$ et $dt = \cos u du$. Par ailleurs, on peut écrire :
 $\alpha = \arcsin(a) = \arcsin(-1) = -\pi/2$ et $\beta = \arcsin(b) = \arcsin(1/2) = \pi/6$.

Il reste : $J = \int_{-\pi/2}^{\pi/6} |\cos u| \cos u du = \int_{-\pi/2}^{\pi/6} \cos^2 u du$

Pour la suite de l'intégration, il reste à linéariser $\cos^2 u = \frac{1}{2}(\cos 2u + 1)$.

On a : $J = \int_{-\pi/2}^{\pi/6} \frac{1}{2}(\cos 2u + 1) du = \left[\frac{\sin 2u}{4} + \frac{u}{2} \right]_{-\pi/2}^{\pi/6} = \frac{\sqrt{3}}{8} + \frac{\pi}{3}$.

□ Méthode 8.3.— Comment et dans quelle circonstance utiliser la formule de l'intégration par parties

On remarque que la quantité $f(t)$ s'écrit $u'(t)v(t)$, où u et v sont deux fonctions dont leurs dérivées sont continues sur un intervalle I et on remarque aussi que la quantité $u(t)v'(t)$ s'intègre plus facilement que $f(t)$. On aura déterminé une primitive u de u' . Puis, on écrit (égalité valable à une constante additive près) :

$$\int u'(t)v(t) dt = u(t)v(t) - \int u(t)v'(t) dt.$$

Il reste à intégrer $u(t)v'(t)$.

La difficulté de la méthode, c'est de choisir les bonnes fonctions u et v . En effet, la quantité $u'(t)v(t)$ n'est pas censée s'intégrer directement (sinon on n'appliquerait pas cette formule) mais par contre, la quantité $u(t)v'(t)$, elle, doit s'intégrer (ou du moins être « plus intégrable »).

Exemples : 1. Déterminer une primitive de $t \mapsto \ln t$ sur $]0, +\infty[$.

On pose $u'(t) = 1$ et $v(t) = \ln t$. Les fonctions u et v sont bien de classe \mathcal{C}^1 sur $]0, +\infty[$ et :
 $\int 1 \times \ln t dt = t \ln t - \int t \frac{1}{t} dt$.

On choisit t pour primitive de 1, car en général, le recours à une constante d'intégration particulière n'est pas utile. Il reste donc : $\int \ln t dt = t \ln t - t + K$, où $K \in \mathbf{R}$.

2. En admettant qu'une primitive de $t \mapsto \frac{1}{\sqrt{1+t^2}}$ est $t \mapsto \ln(t + \sqrt{t^2 + 1})$, déterminer une primitive de $t \mapsto \sqrt{t^2 + 1}$.

Pour cela, on effectue une intégration par parties en posant $u'(t) = 1$ et $v(t) = \sqrt{1+t^2}$. On a :
 $F = \int 1 \times \sqrt{1+t^2} dt = t\sqrt{1+t^2} - \int \frac{t^2}{\sqrt{1+t^2}} dt$, à une constante additive près. On écrit (astuce

à retenir pour d'autres exos!) $\frac{t^2}{\sqrt{1+t^2}} = \frac{t^2 + 1 - 1}{\sqrt{1+t^2}} = \sqrt{1+t^2} - \frac{1}{\sqrt{1+t^2}}$,

on en déduit : $F = t\sqrt{1+t^2} - F + \ln(t + \sqrt{t^2 + 1}) \Rightarrow F = \frac{1}{2} \left[t\sqrt{1+t^2} + \ln(t + \sqrt{t^2 + 1}) \right]$.

Remarque : Parfois, il faut faire plusieurs intégrations par parties successives !

❑ **Méthode 8.4.— Comment intégrer une fonction rationnelle du type**

$$t \mapsto \frac{1}{t^2 + \alpha t + \beta}$$

On peut écrire $t^2 + \alpha t + \beta = \left(t + \frac{\alpha}{2}\right)^2 + \frac{4\beta - \alpha^2}{4}$. Puis on effectue le changement de variable $x = t + \frac{\alpha}{2}$. Enfin, selon le signe de $-\alpha^2 + 4\beta$, on pose d^2 ou $-d^2$ cette dernière quantité et on se ramène au tableau des primitives classiques.

Exemple : Calculer $F : x \mapsto \int_0^x \frac{1}{1+t+t^2} dt$.

Comme on peut écrire : $1+t+t^2 = \left(t + \frac{1}{2}\right)^2 + \frac{3}{4}$, on pose alors le changement de variable $u = t + \frac{1}{2}$ et on écrit $dt = du$. Ce qui donne : $F : x \mapsto \int_{\frac{1}{2}}^{x+\frac{1}{2}} \frac{1}{u^2 + \left(\frac{\sqrt{3}}{2}\right)^2} du$.

Il reste à utiliser le tableau des primitives usuelles,

$$F(x) = \left[\frac{2}{\sqrt{3}} \arctan \frac{2u}{\sqrt{3}} \right]_{\frac{1}{2}}^{x+\frac{1}{2}} = \frac{2}{\sqrt{3}} \arctan \frac{2x+1}{\sqrt{3}} - \frac{2}{\sqrt{3}} \arctan \frac{1}{\sqrt{3}}$$

Or, $\arctan \frac{1}{\sqrt{3}} = \frac{\pi}{6}$ et il reste : $F(x) = \frac{2}{\sqrt{3}} \arctan \frac{2x+1}{\sqrt{3}} - \frac{\pi}{3\sqrt{3}}$. On a fini !

Remarque : on peut étendre aux fonctions rationnelles f du type $t \mapsto \frac{at+b}{t^2+\alpha t+\beta}$, avec $(a, b, \alpha, \beta) \in \mathbb{R}^4$. Pour cela, on peut suivre le cheminement suivant :

- décomposer la fonction $f(t)$ sous la forme $\frac{\frac{a}{2}(2t+\alpha)}{t^2+\alpha t+\beta} + \frac{b-\frac{a\alpha}{2}}{t^2+\alpha t+\beta}$;
- déterminer alors une primitive de $t \mapsto \frac{\frac{a}{2}(2t+\alpha)}{t^2+\alpha t+\beta}$ de forme $\frac{u'}{u}$;
- déterminer une primitive de $t \mapsto \frac{b-\frac{a\alpha}{2}}{t^2+\alpha t+\beta}$ avec la méthode précédente.

❏ Méthode 8.5.— Comment intégrer une fonction rationnelle $f(t)$ de variable t
On décompose en éléments simples la fraction rationnelle associée dans $\mathbf{R}(X)$, c'est-à-dire que l'on écrit $f(t)$ sous la forme d'une combinaison linéaire de fonctions du type

$$t \mapsto P(t), t \mapsto \frac{1}{(t-a)^m}, t \mapsto \frac{at+b}{(t^2+\alpha t+\beta)^n},$$

où P est un polynôme en t , $(m, n) \in (\mathbf{N}^*)^2$, $(a, b, \alpha, \beta) \in \mathbf{R}^4$ avec $\alpha^2 - 4\beta < 0$.
On détermine une primitive de chacun des éléments simples en utilisant le tableau des primitives usuelles et la méthode précédente.

Remarque : la théorie de la décomposition en éléments simples n'est absolument pas au programme de TSI. Ceci dit, si l'on vous guide, par exemple en donnant la forme générale de la décomposition, vous devez y arriver. Un grand nombre d'exercices dans vos concours démarrent ainsi. Par ailleurs, en SI, vous pratiquez aussi des décompositions en éléments simples.

Exemple : Trouver une primitive de la fonction $t \mapsto \frac{1}{1-t^3}$ sur un intervalle ne contenant pas 1.
On commencera par remarquer que $1-t^3 = (1-t)(1+t+t^2)$ et qu'il existe a, b et c trois réels (que l'on déterminera) tels que : $\frac{1}{1-t^3} = \frac{a}{1-t} + \frac{bt+c}{1+t+t^2}$.

On trouve les coefficients a, b et c soit en réduisant au même dénominateur et en identifiant, soit en utilisant quelques astuces de calculs de coefficients. Par exemple, si l'on multiplie la fraction par $(1-t)$ et que l'on applique pour $t=1$, on obtient le coefficient $a = 1/3$. Si on multiplie la fraction par t et que l'on fait tendre t vers $+\infty$, on obtient $-a+b=0$ et donc $b = 1/3$. Enfin, une identification avec $t=0$ donne $c = 2/3$. Dans tous les cas, : $\frac{1}{1-t^3} = \frac{1}{3} \left(\frac{1}{1-t} + \frac{t+2}{1+t+t^2} \right)$.

Il reste à intégrer $t \mapsto \frac{1}{3(1-t)}$ ce qui est immédiat et à intégrer $t \mapsto \frac{t+2}{3(1+t+t^2)}$, or cette dernière primitive a été calculé un peu plus haut : $\frac{1}{3} \left(\sqrt{3} \arctan \frac{2t+1}{\sqrt{3}} + \frac{1}{2} \ln(t^2+t+1) \right)$.

Finalement : $\int \frac{1}{1-t^3} dt = \frac{1}{3} \left(-\ln|1-t| + \sqrt{3} \arctan \frac{2t+1}{\sqrt{3}} + \frac{1}{2} \ln(t^2+t+1) \right)$.

Remarque : On peut procéder parfois sans connaître la théorie de la décomposition. C'est le cas si $f(t)$ est un rapport du type $\frac{P(t)}{(t+a)^n}$, où P est un polynôme. On cherche à faire apparaître $t+a$ dans P .

Exemple : Déterminer une primitive de $f : t \mapsto \frac{t^3}{(1+t)^2}$.

On a : $f(t) = \frac{t^3+1-1}{(1+t)^2} = \frac{(t+1)(t^2-t+1)}{(1+t)^2} - \frac{1}{(1+t)^2}$. De plus, $t^2-t+1 = (t-2)(t+1)+3$.

Il reste : $\int f(t) dt = \int \left[t-2 + \frac{3}{1+t} - \frac{1}{(1+t)^2} \right] dt = \frac{t^2}{2} - 2t + 3 \ln(1+t) + \frac{1}{1+t} + K$, où $K \in \mathbf{R}$.

❑ **Méthode 8.6.— Comment calculer des intégrales du type $I_n = \int_a^b f_n(t) dt$**

Une méthode est de tenter une récurrence. On commence par calculer les premiers éléments de (I_n) , par exemple I_0 et I_1 . Puis à l'aide d'une intégration par parties, on trouve une relation entre I_n et un ou plusieurs de ses prédécesseurs I_{n-1} , I_{n-2} et on en déduit I_n en fonction de n .

Exemple : Calculer pour tout $n \in \mathbf{N}$, l'intégrale $I_n = \int_0^{\frac{\pi}{2}} \sin^n x dx$.

Au départ : $I_0 = \frac{\pi}{2}$ et $I_1 = 1$. Écrivons pour $n \geq 2$, $I_n = \int_0^{\frac{\pi}{2}} \sin^{n-1} x \sin x dx$. Par une intégration par parties, si l'on pose $u = \sin^{n-1} x$ et $v' = \sin x$,

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

En utilisant $\cos^2 x = 1 - \sin^2 x$, $I_n = (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x dx - (n-1) \int_0^{\frac{\pi}{2}} \sin^n x dx$.

On en déduit l'implication : $I_n = (n-1)(I_{n-2} - I_n) \Rightarrow nI_n = (n-1)I_{n-2}$.

$$I_{2p} = \frac{(2p-1)(2p-3)\dots 3 \times 1}{2p(2p-2)(2p-4)\dots \times 4 \times 2} \frac{\pi}{2} \text{ et } I_{2p+1} = \frac{2p(2p-2)\dots 4 \times 2}{(2p+1)(2p-1)(2p-3)\dots \times 5 \times 3}.$$

❑ **Méthode 8.7.— Comment encadrer une fonction entre deux polynômes à l'aide de la formule de Taylor avec reste intégral**

On écrit la formule de Taylor avec reste intégral jusqu'à l'ordre nécessaire pour faire apparaître un des polynômes en tant que partie régulière dans cette formule et on montre que le reste intégral est toujours de même signe.

Exemple : Prouver sans étude de fonction que pour tout $x \in [-\pi, \pi]$, $\cos x \geq 1 - \frac{x^2}{2}$.

On applique la formule de Taylor avec reste intégral à l'ordre 2, $\cos x = 1 - \frac{x^2}{2} + \int_0^x \frac{(x-t)^2}{2} \sin t dt$.

Si $x < 0$, \sin est négative et comme on intègre de 0 à x , $\int_0^x \frac{(x-t)^2}{2} \sin t dt$ est positive et si $x > 0$,

\sin est positive et comme on intègre de 0 à x , $\int_0^x \frac{(x-t)^2}{2} \sin t dt$ est encore positive.

❑ **Méthode 8.8.— Comment calculer la limite d'une suite de somme de Riemann**

On peut commencer par factoriser la somme par $(b-a)/n$ puis trouver la fonction f continue et le segment $[a, b]$ sur lequel on travaille. Enfin on explicite u_n en tant que somme de Riemann et on en déduit que $\lim_{n \rightarrow +\infty} u_n = \int_a^b f(t) dt$.

Exemple : Déterminer la limite de $S_n = \frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{n+n}$ quand $n \rightarrow +\infty$.

Commençons par « arranger » la somme S_n .

Elle peut s'écrire : $S_n = \frac{1}{n} \left(\frac{1}{1+1/n} + \dots + \frac{1}{1+n/n} \right) = \frac{1}{n} \left(f\left(\frac{1}{n}\right) + \dots + f(1) \right)$, en posant $f(x) = \frac{1}{1+x}$. C'est donc une somme de Riemann et on a : $\lim_{n \rightarrow +\infty} S_n = \int_0^1 \frac{dx}{1+x} = \ln 2$.

■ Étude de la convergence d'une intégrale sur un intervalle

□ Méthode 8.9.— Comment étudier la convergence d'une intégrale sur un intervalle à l'aide de la définition du cours

L'application f étant continue par morceaux sur $[a, b[$ (b réel ou $+\infty$), considérons l'intégrale $J = \int_a^b f(t) dt$ et F une primitive de f .

- Pour un x quelconque de $[a, b[$, on calcule : $\int_a^x f(t) dt = F(x) - F(a)$.
- Si $F(x) - F(a)$ admet une limite finie quand x tend vers b , J est convergente et égale à cette limite; sinon elle est divergente.

Exemples : 1. Étudier la nature de $I_1 = \int_0^1 \frac{dt}{\sqrt{t}}$, $I_2 = \int_1^{+\infty} \frac{dt}{\sqrt{t}}$ et $I_3 = \int_0^{+\infty} \cos t dt$.

► I_1 est convergente. En effet, $t \mapsto \frac{1}{\sqrt{t}}$ est continue sur $]0, 1]$ et, pour tout $\varepsilon \in]0, 1]$, on peut écrire : $\int_\varepsilon^1 \frac{dt}{\sqrt{t}} = [2\sqrt{t}]_\varepsilon^1 = 2 - 2\sqrt{\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} 2$. De plus : $\int_0^1 \frac{dt}{\sqrt{t}} = 2$.

► I_2 n'est pas convergente. $t \mapsto \frac{1}{\sqrt{t}}$ est bien continue sur $[1, +\infty[$ mais, pour tout $x \in [1, +\infty[$, on peut écrire : $\int_1^x \frac{dt}{\sqrt{t}} = [2\sqrt{t}]_1^x = 2\sqrt{x} - 2 \xrightarrow{x \rightarrow +\infty} +\infty$ et $\int_1^{+\infty} \frac{dt}{\sqrt{t}}$ est divergente.

► I_3 est divergente. $t \mapsto \cos t$ est bien continue sur $[0, +\infty[$, mais, pour tout $x \in [0, +\infty[$, on peut écrire : $\int_0^x \cos t dt = [\sin t]_0^x = \sin x$ et $x \mapsto \sin x$ n'a pas de limite en $+\infty$.

2. Convergence et, éventuellement, calcul de $A = \int_0^1 \frac{dx}{\sqrt{1-x^2}}$ et $B = \int_{-\infty}^{+\infty} \frac{dt}{1+t^2}$.

► $x \mapsto \frac{1}{\sqrt{1-x^2}}$ est continue sur $[0, 1[$; elle est positive sur \mathbf{R} .

On écrit alors sans souci : $\forall X \in [0, 1[, \int_0^X \frac{dx}{\sqrt{1-x^2}} = [\operatorname{Arcsin} x]_0^X = \operatorname{Arcsin} X \xrightarrow{X \rightarrow 1} \frac{\pi}{2}$.

Conclusion : $A = \int_0^1 \frac{dx}{\sqrt{1-x^2}}$ est convergente et vaut $\frac{\pi}{2}$.

► $t \mapsto \frac{1}{1+t^2}$ est continue sur \mathbf{R} ; elle est positive sur \mathbf{R} .

Pour tout $x \geq 0$, $\int_0^x \frac{dt}{1+t^2} = [\text{Arctan } t]_0^x = \text{Arctan } x \xrightarrow{x \rightarrow +\infty} \frac{\pi}{2}$, donc $\int_0^{+\infty} \frac{dt}{1+t^2} = \frac{\pi}{2}$.

Par parité, on obtient que $\int_{-\infty}^0 \frac{dt}{1+t^2}$ est convergente et vaut aussi $\frac{\pi}{2}$.

Conclusion : $\int_{-\infty}^{+\infty} \frac{dt}{1+t^2}$ est convergente et vaut π .

□ Méthode 8.10.— Comment montrer la convergence d'une intégrale ou l'intégrabilité d'une fonction à l'aide des théorèmes de comparaison

L'application f étant continue sur $[a, b[$ (b réel ou $+\infty$), considérons l'intégrale $\int_a^b f(t) dt$.

[1] Les théorèmes de comparaisons sont établis pour des fonctions **de même signes** ou pour leurs valeurs absolues. Si f est à valeurs réelles, il est donc indispensable de se poser la question du signe de f sur $[a, b[$. Si celui-ci n'est pas constant (au moins autour de la valeur b), on est bloqué pour l'étude de la nature de $\int_a^b f(t) dt$. Par contre, on peut

étudier la nature de $\int_a^b |f(t)| dt$.

Le but du jeu est, ensuite, de comparer f ou $|f|$ à une fonction dont l'intégrale sur $[c, b[$, $c \in [a, b[$, est une intégrale de référence (nature établie en cours). On prendra $c = a$ si l'intégrale de référence est définie en a et on prendra $c > a$ sinon.

Cette comparaison pourra s'effectuer par majoration-minoration ou par équivalence en utilisant la **proposition 8.14**, la **proposition 8.27** et la **proposition 8.28**.

Exemple : Étudier la convergence des intégrales :

$$1. \int_1^{+\infty} \frac{dt}{\sqrt{1+t^4}} \quad 2. \int_1^{+\infty} \frac{dt}{\sqrt{t^4-1}} \quad 3. \int_0^{+\infty} e^{-x^2} dx$$

1. Remarquons que l'intégrande est continue sur $[1, +\infty[$ et positive ce qui autorise l'usage des théorèmes de comparaison. Seule la borne supérieure $+\infty$ pose un problème.

On écrit : $\sqrt{1+t^4} \geq t^2$ donc : $\frac{1}{\sqrt{1+t^4}} \leq \frac{1}{t^2}$, or $\int_1^{+\infty} \frac{dt}{t^2}$ converge donc $\int_1^{+\infty} \frac{dt}{\sqrt{1+t^4}}$ converge.

2. L'intégrande est continue sur $]1, +\infty[$, mais non prolongeable en 1.

Il est donc préférable de scinder l'intégrale en une somme de deux intégrales généralisées.

On écrit : $\int_1^2 \frac{dt}{\sqrt{t^4-1}}$ (problème en 1) et $\int_2^{+\infty} \frac{dt}{\sqrt{t^4-1}}$ (problème en $+\infty$).

Comme la fonction à intégrer est positive, on peut utiliser les théorèmes de comparaison.

On écrit : $\sqrt{t^4-1} \sim \sqrt{t^4+1}$ en $+\infty$ et $\int_2^{+\infty} \frac{dt}{\sqrt{t^4+1}}$ converge donc $\int_2^{+\infty} \frac{dt}{\sqrt{t^4-1}}$ converge.

$\sqrt{t^4-1} = \sqrt{t^2-1}\sqrt{t^2+1} \sim \sqrt{2}\sqrt{2}\sqrt{t-1}$ en 1, et $\int_1^2 \frac{dt}{\sqrt{t-1}}$ converge (car $t \mapsto \frac{1}{\sqrt{t-1}}$ a pour

primitive $t \mapsto 2\sqrt{t-1}$ et $[2\sqrt{t-1}]_1^2$ existe) donc $\int_1^2 \frac{dt}{\sqrt{t^4-1}}$ converge.

Finalement l'intégrale proposée converge.

3. La fonction $x \mapsto e^{-x^2}$ est continue sur \mathbf{R} donc seule la borne infinie pose un problème. Par ailleurs, cette fonction est positive, ce qui autorise l'usage des théorèmes de comparaison. $e^x e^{-x^2} = e^{x-x^2}$ est majorée sur \mathbf{R}_+ (car inférieure à 1 sur $[1, +\infty[$, et continue sur le segment $[0, 1]$) donc $\exists M \in \mathbf{R}_+, 0 \leq e^{-x^2} \leq M e^{-x}$. Comme $\int_{\mathbf{R}_+} e^{-x} dx$ converge, $\int_{\mathbf{R}_+} e^{-x^2} dx$ converge aussi.

■ Calcul d'une intégrale (généralisée) sur un intervalle

□ Méthode 8.11.— Comment pratiquer un changement de variable dans une intégrale sur un intervalle

Si l'on veut pratiquer le changement de variable défini par l'application φ dans l'intégrale généralisée $\int_a^b f(t) dt$, il faut alors :

[1] s'assurer que φ est une bijection strictement monotone de classe \mathcal{C}^1 d'un intervalle $]\alpha, \beta[$ dans $]a, b[$;

établir la convergence de $\int_\alpha^\beta (f \circ \varphi)(u) \varphi'(u) du$ ou de $\int_a^b f(t) dt$;

si φ est strictement croissante, ces deux intégrales sont égales, sinon, elles sont opposées. Dans la plupart des cas, les changements de variables seront suggérés par l'énoncé. Néanmoins, il arrivera, de temps à autres, qu'on espère une certaine autonomie de l'étudiant(e) pour des changements de variables « naturels ».

Exemples : 1. Montrer que $f : x \mapsto \frac{\sin x}{x^2}$ est intégrable sur $[1, +\infty[$, puis que :

$$\int_1^{+\infty} \frac{\sin x}{x^2} dx = \int_0^1 \sin \frac{1}{x} dx.$$

f est continue sur $[1, +\infty[$. Puis :

$$\forall x \in [1, +\infty[, |f(x)| \leq \frac{1}{x^2}$$

et $x \mapsto \frac{1}{x^2}$ est intégrable sur $[1, +\infty[$, donc f est intégrable sur $[1, +\infty[$.

$\varphi : x \mapsto \frac{1}{x}$ est une bijection \mathcal{C}^1 strictement décroissante de $[1, +\infty[$ dans $]0, 1]$.

On a donc un changement de variable valide pour cette intégrale généralisée.

Puis : $\forall x \in [1, +\infty[, |\varphi'(x)| = \frac{1}{x^2}$, donc :

$$\int_1^{+\infty} \frac{\sin x}{x^2} dx = - \int_1^{+\infty} \sin \left(\frac{1}{\varphi(x)} \right) \varphi'(x) dx = \int_0^1 \sin \frac{1}{t} dt.$$

2. Pratiquer le changement de variable $x = \sin t$ sur l'intégrale $\int_{-1}^1 \frac{dx}{(2-x^2)\sqrt{1-x^2}}$.

En déduire sa convergence. Pratiquer le changement de variable $u = \tan t$ sur l'intégrale obtenue et en déduire la valeur de l'intégrale initiale.

▷ $\varphi : t \mapsto \sin t$ est une bijection strictement croissante et de classe \mathcal{C}^1 de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ dans $] -1, 1[$: c'est donc un changement de variable valide pour l'intégrale proposée. Comme :

$$\forall t \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[, \varphi'(t) = \cos t = \sqrt{1 - \sin^2 t},$$

l'intégrale $\int_{-1}^1 \frac{dx}{(2-x^2)\sqrt{1-x^2}}$ et l'intégrale

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\varphi'(t) dt}{(2-\varphi^2(t))\sqrt{1-\varphi^2(t)}} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\sqrt{1-\sin^2 t} dt}{(2-\sin^2 t)\sqrt{1-\sin^2 t}} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{dt}{(1+\cos^2 t)}$$

sont de même nature.

Or, cette dernière est une intégrale définie, donc l'intégrale initiale converge et :

$$\int_{-1}^1 \frac{dx}{(2-x^2)\sqrt{1-x^2}} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{dt}{(1+\cos^2 t)}.$$

▷ $\psi : t \mapsto \tan t$ est une bijection strictement croissante et de classe \mathcal{C}^1 de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ dans \mathbf{R} ; il s'agit d'un changement de variable valide pour l'intégrale précédente.

Comme, $\forall t \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[, \psi'(t) = \frac{1}{\cos^2 t}$ et $\cos^2 t = \frac{1}{1+\tan^2 t}$, on peut écrire :

$$\begin{aligned} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{dt}{(1+\cos^2 t)} &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\cos^2 t}{(1+\cos^2 t)} \frac{1}{\cos^2 t} dt \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{(1+\psi^2(t)) \left(1 + \frac{1}{1+\psi^2(t)}\right)} |\psi'(t)| dt \\ &= \int_{-\infty}^{+\infty} \frac{du}{2+u^2} \\ &= \lim_{X \rightarrow +\infty} \left[\frac{1}{\sqrt{2}} \arctan \frac{u}{\sqrt{2}} \right]_{-X}^{+X} = \frac{\pi}{\sqrt{2}}. \end{aligned}$$

■ **Méthode 8.12.— Comment pratiquer une intégration par parties sur une intégrale sur un intervalle**

Considérons $\int_a^b f(t)g'(t) dt$ avec $I = [a, b[$.

- On vérifie que f et g sont de classe \mathcal{C}^1 sur $[a, b[$. L'existence des limites du produit fg aux bornes de l'intervalle assure que les intégrales de fg' et $f'g$ sont de même nature. On peut appliquer alors directement (dans les cas simples) :

$$\int_a^b f(t)g'(t) dt = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t) dt.$$

- Souvent, par précaution, (et c'est ce que nous ferons dans les exemples), il vaut mieux choisir X dans $[a, b[$, puis pratiquer l'intégration par parties sur $[a, X]$. Et chercher ensuite la limite lorsque X tend vers b de l'expression obtenue.
- Dans le cas où $I =]a, b[$, c'est-à-dire dans le cas où $\int_a^b f(t)g'(t) dt$ est doublement généralisée, on fera l'intégration par parties sur $[\epsilon, X]$ avec $a < \epsilon < X < b$ et on cherche ensuite la limite lorsque ϵ tend vers a et X tend vers b dans l'expression obtenue.

Exemples : 1. Montrer que l'intégrale $\int_0^1 \ln t dt$ est convergente et vaut -1 puis que l'intégrale $\int_1^{+\infty} \ln t dt$ est divergente.

Vous allez chercher la **proposition ??**. On écrit (où $0 < u < v$) :

$$\int_u^v 1 \times \ln t dt = - \int_u^v t \times \frac{1}{t} dt + [t \ln t]_u^v.$$

On commence par faire tendre u vers 0 et v vers 1. La quantité $t \ln t$ tend vers 0 dans les deux cas. Comme $-\int_0^1 t \times \frac{1}{t} dt = -1$, on en déduit bien que $\int_0^1 \ln t dt$ est convergente et vaut -1 .

On fait tendre u vers 1 et v vers $+\infty$. Or : $-\int_u^v t \times \frac{1}{t} dt + [t \ln t]_u^v = v - u + v \ln v - u \ln u$ tend vers $+\infty$. Donc : $\int_1^{+\infty} \ln t dt$ est divergente.

2. Pour $n \in \mathbf{N}$, convergence et calcul de $I_n = \int_0^{+\infty} t^n e^{-t} dt$.

► **Positivité et continuité de l'intégrande.**

$t \mapsto t^n e^{-t}$ est continue, donc continue par morceaux, et positive sur \mathbf{R}_+ : on peut déjà dire que l'intégrale est seulement généralisée en $+\infty$.

► **Étude de la convergence.**

À défaut de trouver un équivalent plus simple de $t^n e^{-t}$ en $+\infty$, on peut en rechercher un majorant. D'après le théorème de croissance comparée, pour tout entier $n \in \mathbf{N}$, $\lim_{n \rightarrow +\infty} t^{n+2} e^{-t} = 0$ donc $t \mapsto t^{n+2} e^{-t}$ est bornée sur un intervalle $[A, +\infty[$ ($A \in \mathbf{R}_+^*$), et elle est par ailleurs bornée sur $[0, A]$. Alors : $\exists M \in \mathbf{R}_+, \forall t \in \mathbf{R}, 0 \leq t^{n+2} e^{-t} \leq M$, donc, pour tout $t \in \mathbf{R}_+^*$: $0 \leq t^n e^{-t} \leq \frac{M}{t^2}$.

On conclut en utilisant la convergence de $\int_1^{+\infty} \frac{dt}{t^2}$.

► **Intégration par parties sur une intégrale ordinaire.**

Effectuons sur l'intégrale $\int_0^X t^n e^{-t} dt$ ($X \in \mathbf{R}_+$) une intégration par parties en posant :

$$\begin{cases} u'(t) = e^{-t} \\ v(t) = t^n \end{cases} \quad \text{et} \quad \begin{cases} u(t) = -e^{-t} \\ v'(t) = nt^{n-1} \end{cases}$$

Alors : $\int_0^X u'(t)v(t) dt = [-t^n e^{-t}]_0^X + \int_0^X nt^{n-1} e^{-t} dt.$

Faisons tendre X vers $+\infty$; le crochet tend vers 0 (croissances comparées) et :

$$I_n = nI_{n-1}.$$

► **Fin du calcul.**

À partir de $I_0 = 1$ et de $I_n = nI_{n-1}$, on obtient par récurrence sur $n \in \mathbf{N}$: $I_n = n!$

En effet, cette relation est vérifiée pour $n = 0$ (c'est une intégrale de référence) et si elle est vraie au rang $n \in \mathbf{N}$ alors $I_n = n!$ donc $I_{n+1} = (n+1)I_n = (n+1)!$.

$$I_n = \int_0^{+\infty} t^n e^{-t} dt \text{ converge et vaut } n!.$$

3. Convergence et calcul de $\int_0^1 \frac{\ln t}{\sqrt{1-t}} dt$.

► **Positivité et continuité de l'intégrande.**

La fonction $t \mapsto \frac{\ln t}{\sqrt{1-t}}$ est continue et négative sur $]0, 1[$.

► **Étude de la convergence.** On a :

$$\frac{\ln t}{\sqrt{1-t}} \underset{0}{\sim} \ln t \text{ et } \frac{\ln t}{\sqrt{1-t}} = \frac{\ln(1+(t-1))}{\sqrt{1-t}} \underset{1}{\sim} -\sqrt{1-t}.$$

Comme $\int_0^{\frac{1}{2}} \ln t dt$ et $\int_{\frac{1}{2}}^1 \sqrt{1-t} dt$ convergent, l'intégrale proposée converge.

► **L'intégration par parties sur l'intégrale généralisée échoue.**

Si l'on applique directement la formule d'intégration par parties en posant :

$$\begin{cases} u'(t) = \frac{1}{\sqrt{1-t}} \\ v(t) = \ln t \end{cases} \quad \text{et} \quad \begin{cases} u(t) = -2\sqrt{1-t} \\ v'(t) = \frac{1}{t} \end{cases},$$

on obtient :

$$\int_0^1 \frac{\ln t}{\sqrt{1-t}} dt = [-2\sqrt{1-t} \ln t]_0^1 + 2 \int_0^1 \frac{\sqrt{1-t}}{t} dt.$$

Les deux termes du membre de droite ne sont pas des réels, car $\lim_{t \rightarrow 0} -2\sqrt{1-t} \ln t = +\infty$

et $\frac{\sqrt{1-t}}{t} \underset{0}{\sim} \frac{1}{t}$ donc $\int_0^1 \frac{\sqrt{1-t}}{t} dt$ n'est pas convergente.

► **Intégration par parties sur une intégrale ordinaire.**

Soit ε et x dans $]0, 1[$ tels que : $0 < \varepsilon < x < 1$:

$$\int_{\varepsilon}^x \frac{\ln t}{\sqrt{1-t}} dt = \left[-2\sqrt{1-t} \ln t \right]_{\varepsilon}^x + 2 \int_{\varepsilon}^x \frac{\sqrt{1-t}}{t} dt.$$

► **Fin du calcul.**

Pour cette dernière intégrale, pratiquons le changement de variable (vous ne l'auriez jamais trouvé tout seul ? Rassurez-vous, on vous le donnera sûrement ! Ce que l'on attend de vous, c'est que vous sachiez l'utiliser) : $u = \sqrt{1-t}$.

Alors : $du = -\frac{dt}{2\sqrt{1-t}}$, $t = \varepsilon \Rightarrow u = \sqrt{1-\varepsilon}$ et $t = x \Rightarrow u = \sqrt{1-x}$, d'où :

$$\begin{aligned} \int_{\varepsilon}^x \frac{\sqrt{1-t}}{t} dt &= - \int_{\sqrt{1-\varepsilon}}^{\sqrt{1-x}} \frac{2u^2}{1-u^2} du \\ &= \int_{\sqrt{1-\varepsilon}}^{\sqrt{1-x}} \left(2 - \frac{2}{1-u^2} \right) du \\ &= \int_{\sqrt{1-\varepsilon}}^{\sqrt{1-x}} \left(2 - \frac{1}{1+u} - \frac{1}{1-u} \right) du \\ &= \left[2u + \ln \frac{1-u}{1+u} \right]_{\sqrt{1-\varepsilon}}^{\sqrt{1-x}} \\ &= 2\sqrt{1-x} - 2\sqrt{1-\varepsilon} + \ln \frac{1-\sqrt{1-x}}{1+\sqrt{1-x}} - \ln \frac{1-\sqrt{1-\varepsilon}}{1+\sqrt{1-\varepsilon}} \end{aligned}$$

Il n'y a pas de problème à faire tendre x vers 1. On obtient alors :

$$\int_{\varepsilon}^1 \frac{\ln t}{\sqrt{1-t}} dt = 2\sqrt{1-\varepsilon} \ln \varepsilon + 2 \left(-2\sqrt{1-\varepsilon} + \ln \frac{1+\sqrt{1-\varepsilon}}{1-\sqrt{1-\varepsilon}} \right).$$

En multipliant numérateur et dénominateur par l'expression conjuguée du dénominateur :

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

D'où : $\ln \frac{1+\sqrt{1-\varepsilon}}{1-\sqrt{1-\varepsilon}} = 2 \ln(1+\sqrt{1-\varepsilon}) - \ln \varepsilon$ et

$$\begin{aligned} \int_{\varepsilon}^1 \frac{\ln t}{\sqrt{1-t}} dt &= 2\sqrt{1-\varepsilon} \ln \varepsilon - 4\sqrt{1-\varepsilon} + 4 \ln(1+\sqrt{1-\varepsilon}) - 2 \ln \varepsilon \\ &= 2(\sqrt{1-\varepsilon} - 1) \ln \varepsilon - 4\sqrt{1-\varepsilon} + 4 \ln(1+\sqrt{1-\varepsilon}) \end{aligned}$$

Comme $(\sqrt{1-\varepsilon} - 1) \ln \varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0$ (car $\sqrt{1-\varepsilon} - 1 \sim -\frac{\varepsilon}{2}$), on écrit :

$$\int_0^1 \frac{\ln t}{\sqrt{1-t}} dt = -4 + 4 \ln 2.$$

Remarque : sachez que l'on peut utiliser l'intégration par parties pour montrer la convergence d'une intégrale généralisée.

□ **Méthode 8.13.— Comment étudier une intégrale généralisée $J = \int_a^b f(t) dt$ dans le cas où f est à valeurs dans \mathbf{C}**

- On peut étudier la nature des intégrales généralisées :

$$K = \int_a^b \Re(f)(t) dt \text{ et } L = \int_a^b \Im(f)(t) dt.$$

Si ces deux intégrales existent, $J = K + iL$.

- On peut étudier (on commencera par calculer $|f(t)|$) si f est intégrable sur le domaine d'intégration. Si c'est le cas, J existe.

Exemple : Étudier la nature de $\int_0^{+\infty} \frac{\exp(it) - 1}{t^{3/2}} dt$.

La fonction à intégrer est continue sur \mathbf{R}_+^* . Par ailleurs, cette fonction est à valeurs dans \mathbf{C} , ce qui n'autorise pas l'usage direct des théorèmes de comparaison.

- En 0, on écrit (en sortant de votre inépuisable culture : $e^{it} = 1 - 2\sin^2(t/2) + i\sin t$) :

$$\frac{e^{it} - 1}{t^{3/2}} = -\frac{2\sin^2 t/2}{t^{3/2}} + i\frac{\sin t}{t^{3/2}} = -\frac{\sqrt{t}}{2} + \frac{i}{\sqrt{t}} + o(1),$$

donc $t \mapsto \frac{e^{it} - 1}{t^{3/2}}$ n'est pas prolongeable en 0.

La partie imaginaire de $\frac{e^{it} - 1}{t^{3/2}}$ est de signe constant et équivalente à $\frac{1}{\sqrt{t}}$; or $\int_0^1 \frac{1}{\sqrt{t}} dt$ converge,

donc $\int_0^1 \Im \left(\frac{e^{it} - 1}{t^{3/2}} \right) dt$ converge, ainsi que $\int_0^1 \Re \left(\frac{e^{it} - 1}{t^{3/2}} \right) dt$.

- En $+\infty$, on a la majoration :

$$\left| \frac{e^{it} - 1}{t^{3/2}} \right| \leq \frac{2}{t^{3/2}}.$$

Or $\int_1^{+\infty} \frac{2}{t^{3/2}} dt$ converge, d'où la convergence absolue de $\int_1^{+\infty} \frac{e^{it} - 1}{t^{3/2}} dt$.

Finalement : $\int_0^{+\infty} \frac{e^{it} - 1}{t^{3/2}} dt$ est convergente.

Conclusion : on voit que l'on a utilisé ici les deux pistes, la première pour l'étude en 0 et la seconde pour l'étude en $+\infty$. Par contre, on n'a pas pu calculer l'intégrale !

■ Fonction définie par une intégrale de bornes variables

□ **Méthode 8.14.—** Comment étudier $F : x \mapsto \int_{\varphi(x)}^{\psi(x)} f(t) dt$, où φ et ψ sont \mathcal{C}^1 sur un intervalle I à valeurs dans un intervalle J sur lequel f est continue

- Comme f est continue et si $a \in J$, alors $G : x \mapsto \int_a^x f(t) dt$, primitive sur J s'annulant en a de f , est de classe \mathcal{C}^1 sur J et $\forall x \in J, G'(x) = f(x)$.
- $\forall x \in I, F(x) = \int_a^{\psi(x)} f(t) dt - \int_a^{\varphi(x)} f(t) dt = G \circ \psi(x) - G \circ \varphi(x)$.
- On conclut au caractère \mathcal{C}^1 de F puis on obtient sa dérivée.
- Les problèmes de limite sont traités soit à l'aide d'inégalités obtenues par la croissance de l'intégrale sur un segment (s'assurer que les bornes sont "dans le bon ordre"), soit en étudiant la nature d'une intégrale généralisée.

Exemple : Soit $f : x \mapsto \int_x^{2x} \frac{dt}{\ln t}$. Ensemble de définition, variations et limite en $+\infty$.

[1] La fonction $t \mapsto \frac{1}{\ln t}$ est définie sur $]0, 1[\cup]1, +\infty[$, donc il est nécessaire et suffisant, pour que $f(x)$ existe, que $[x, 2x] \subset]0, 1[\cup]1, +\infty[$, donc $D =]0, \frac{1}{2}[\cup]1, +\infty[$.

$\varphi : x \mapsto \int_{\frac{1}{2}}^x \frac{dt}{\ln t}$ est la primitive sur $]0, 1[$ s'annulant en $\frac{1}{2}$ de $t \mapsto \frac{1}{\ln t}$: elle est donc dérivable sur $]0, 1[$. Puis $\forall x \in]0, \frac{1}{2}[$, $f(x) = \varphi(2x) - \varphi(x)$. f est donc dérivable sur $]0, \frac{1}{2}[$.

On écrit $\forall x \in]0, \frac{1}{2}[$, $f'(x) = 2\varphi'(2x) - \varphi'(x) = \frac{2}{\ln(2x)} - \frac{1}{\ln x} = \frac{\ln x - \ln 2}{\ln x(\ln 2 + \ln x)}$.

$\psi : x \mapsto \int_2^x \frac{dt}{\ln t}$ est la primitive sur $]1, +\infty[$ s'annulant en 2 de $t \mapsto \frac{1}{\ln t}$: elle est donc dérivable sur $]1, +\infty[$. Puis $\forall x \in]1, +\infty[$, $f(x) = \psi(2x) - \psi(x)$. f est bien dérivable sur $]1, +\infty[$.

On écrit $\forall x \in]1, +\infty[$, $f'(x) = 2\psi'(2x) - \psi'(x) = \frac{2}{\ln(2x)} - \frac{1}{\ln x} = \frac{\ln x - \ln 2}{\ln x(\ln 2 + \ln x)}$.

Sur $]0, \frac{1}{2}[$, $\ln x(\ln 2 + \ln x) > 0$, donc $f'(x) < 0$.

Sur $]1, +\infty[$, $\ln x(\ln 2 + \ln x) > 0$, donc $f'(x) < 0$, si $x \in]1, 2[$ et $f'(x) > 0$, si $x > 2$.

Conclusion : f est décroissante sur $]0, \frac{1}{2}[$ et sur $]1, 2[$ et croissante sur $[2, +\infty[$.

Si $x > 1$, $x \leq t \leq 2x \Rightarrow 0 < \ln x \leq \ln t \leq \ln 2x \Rightarrow \frac{1}{\ln 2x} \leq \frac{1}{\ln t} \leq \frac{1}{\ln x} \Rightarrow \frac{x}{\ln 2x} \leq f(x) \leq \frac{x}{\ln x}$.

On en déduit $\lim_{x \rightarrow +\infty} \frac{x}{\ln 2x} = +\infty$, donc $\lim_{x \rightarrow +\infty} f(x) = +\infty$. Attention, ici la non-intégrabilité de

$t \mapsto \frac{1}{\ln t}$ sur $[2, +\infty[$ ne suffit pas pour conclure car on aboutit à une forme indéterminée.

Remarque : Dans le cas $x \mapsto \int_x^{+\infty} f(t) dt$, on prouve d'abord l'intégrabilité sur un intervalle du type $]a, +\infty[$. $\int_a^{+\infty} f(t) dt$ est alors constant.

Chapitre 9

Séries numériques

■ Objectifs

■ Les incontournables :

- ▶ Savoir montrer qu'une série converge (ou diverge), en utilisant tout ce qui est possible d'utiliser, en particulier la règle de d'Alembert.
- ▶ Savoir appliquer à *bon escient* le critère spécial des séries alternées.
- ▶ Tirer parti de la comparaison entre série et intégrale généralisée.
- ▶ Utiliser le principe de télescopage pour montrer la convergence d'une série et trouver la valeur de sa somme.

■ Et plus si affinités ...

- ▶ Savoir approcher ou encadrer un réel à l'aide de sommes de séries.
- ▶ Trouver un équivalent du reste R_n d'une série numérique.

Résumé de cours

Vous pouvez commencer par réviser $K^{\mathbf{N}}$, l'ensemble des suites de \mathbf{N} dans K , dans le chapitre 1.

■ Séries à termes dans K

Définition : Série associée à une suite —. Soit $u = (u_n)_{n \in \mathbf{N}}$ une suite de $K^{\mathbf{N}}$. On pose, pour tout entier naturel n , $S_n = \sum_{k=0}^n u_k$. La suite $(S_n)_{n \in \mathbf{N}}$ est la **série** associée à la suite $u = (u_n)_{n \in \mathbf{N}}$.

Cette série est notée $\sum_{k=0}^n u_k$, pour la distinguer de la suite $u = (u_n)_{n \in \mathbf{N}}$; on dit que u_n est son **terme général**. $S_n = \sum_{k=0}^n u_k$ est sa **somme partielle d'indice n** .

Remarque : chaque suite définit une série unique et réciproquement, puisque la connaissance des sommes partielles S_n équivaut à celle des termes généraux : $S_0 = u_0$ et $u_n = S_n - S_{n-1}$ si $n \in \mathbf{N}^*$.

Définition : Série convergente —. La série de terme général u_n , $\sum u_n$, est **convergente** si et seulement si la suite $(S_n)_{n \in \mathbf{N}}$ de ses sommes partielles est convergente dans K .

La limite de la suite $(S_n)_{n \in \mathbf{N}}$ est alors la **somme de la série** : on la note $\sum_{n=0}^{+\infty} u_n$.

Si la suite des sommes partielles diverge, alors la série est dite **divergente**.

Remarque : ne pas confondre $\sum u_n$, qui désigne la série elle-même, et $\sum_{n=0}^{+\infty} u_n$ qui désigne la somme de cette série, c'est-à-dire la limite de la suite de ses sommes partielles. Il y a la même différence entre ces notations qu'entre $(u_n)_{n \in \mathbf{N}}$ et $\lim_{n \rightarrow +\infty} u_n$.

Proposition 9.1. — Cas d'une série de nombres complexes —. Une série de nombres complexes $\sum z_n$ est convergente dans \mathbf{C} si, et seulement si, la série $\sum \Re(z_n)$ et la série $\sum \Im(z_n)$ sont convergentes dans \mathbf{R} . Dans ce cas :

$$\sum_{n=0}^{+\infty} z_n = \sum_{n=0}^{+\infty} \Re(z_n) + i \sum_{n=0}^{+\infty} \Im(z_n).$$

Proposition 9.2. — Espace vectoriel des séries convergentes —. L'ensemble des séries convergentes de K , muni de l'addition des suites et de la multiplication des suites par un scalaire, est un K -espace vectoriel.

Remarque : On suppose $\sum u_n$ et $\sum v_n$ des séries convergentes et $\lambda \in \mathbf{K}$. Alors :

► $\sum (u_n + v_n)$ est convergente et $\sum_{n=0}^{+\infty} (u_n + v_n) = \sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n$.

► $\sum (\lambda u_n)$ est convergente et $\sum_{n=0}^{+\infty} (\lambda u_n) = \lambda \sum_{n=0}^{+\infty} u_n$.

En conséquence, si $\sum u_n$ converge et si $\sum v_n$ diverge, alors $\sum (u_n + v_n)$ diverge. (Si elle convergait, alors $\sum v_n$, différence de deux séries convergentes, convergerait.)

Définition : Reste d'une série convergente —. Si la série $\sum u_n$ converge et a pour somme S ,

$$R_n = S - \sum_{k=0}^n u_k = S - S_n \text{ est le } \textbf{reste d'ordre } n \text{ de la série convergente } \sum u_n.$$

On le note $R_n = \sum_{k=n+1}^{+\infty} u_k$.

Proposition 9.3.— Si la série $\sum u_n$ converge, alors la suite (R_n) de ses restes converge vers 0.

Mise en œuvre : exercice 01, exercice 02.

Proposition 9.4.— **Condition nécessaire de convergence** —. Si la série $\sum u_n$ converge, alors son terme général u_n converge vers 0.

ATTENTION, ce n'est pas une condition suffisante : une série dont le terme général tend vers 0 n'est pas nécessairement convergente. La série harmonique en est un exemple.

En revanche, par contraposition, une série dont le terme général ne tend pas vers 0 est divergente : on dit qu'elle est ***grossièrement divergente***.

Définition : Séries alternées —. Une série réelle $\sum u_n$ est dite ***alternée*** si, pour tout entier naturel n , u_n et u_{n+1} sont de signes contraires.

Proposition 9.5.— **Théorème spécial des séries alternées** —. Une série alternée dont la valeur absolue du terme général décroît et tend vers 0 est convergente.

Dans ce cas, la somme S est du signe de u_0 et elle est comprise entre deux sommes partielles consécutives quelconques.

De plus, pour tout n , le reste d'ordre n , R_n , est du signe de u_{n+1} et $|R_n| \leq |u_{n+1}|$.

Mise en œuvre : exercice 03, exercice 04.

■ Séries de nombres réels positifs

Proposition 9.6.— Convergence par majoration de la suite des sommes partielles —. Une série à termes positifs $\sum u_n$ converge si et seulement si la suite $(S_n)_{n \in \mathbf{N}}$ de ses sommes partielles est majorée. De plus, dans ce cas, $\sum_{n=0}^{+\infty} u_n = \sup_{n \in \mathbf{N}} S_n$.

Proposition 9.7.— Séries de référence —.

- **Séries géométriques —.** Si $q \in \mathbf{R}_+$, alors la série $\sum q^n$ de terme général q^n converge si, et seulement si, $q \in [0, 1[$.
- **Séries de Riemann —.** Si $\alpha \in \mathbf{R}$, alors la série $\sum \frac{1}{n^\alpha}$ converge si, et seulement si, $\alpha > 1$.

Preuve : À développer.

Dans le développement précédent, on a comparé une série et une intégrale, on peut généraliser.

Proposition 9.8.— Comparaison série-intégrale —. Si f est une fonction continue par morceaux, décroissante et à valeurs réelles positives sur $[0, +\infty[$, alors la série $\sum f(n)$ converge si et seulement si f est intégrable sur $[0, +\infty[$.

D'autres méthodes permettent de savoir si une série à termes positifs converge ou non.

Proposition 9.9.— Comparaison par majoration-minoration —. Soit (u_n) et (v_n) deux suites de nombres réels positifs telles que, à partir d'un certain rang n_0 , on ait :

$$\forall n \geq n_0, 0 \leq u_n \leq v_n.$$

Si $\sum v_n$ converge, alors $\sum u_n$ converge; si $\sum u_n$ diverge, alors $\sum v_n$ diverge.

Proposition 9.10.— Comparaison par domination —. Soit (u_n) et (v_n) deux suites de nombres réels positifs telles que : $u_n = O(v_n)$ ou $u_n = o(v_n)$.

$$\text{Si } \sum v_n \text{ converge, alors } \sum u_n \text{ converge.}$$

Corollaire 9.11.— Comparaison par équivalence —. Si (u_n) et (v_n) sont deux suites de nombres réels positifs telles que : $u_n \sim v_n$, alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature, c'est-à-dire qu'elles sont simultanément convergentes ou simultanément divergentes.

Proposition 9.12.— Règle de d'Alembert —. Soit $\sum u_n$ une série à termes strictement positifs :
 Si : $\exists k \in]0, 1[$ et $\exists n_0 \in \mathbf{N}$ tels que : $\forall n \in \mathbf{N}, (n \geq n_0 \Rightarrow \frac{u_{n+1}}{u_n} \leq k)$, alors $\sum u_n$ est convergente.
 Si : $\exists n_0 \in \mathbf{N}$ tel que : $\forall n \in \mathbf{N}, (n \geq n_0 \Rightarrow \frac{u_{n+1}}{u_n} > 1)$, alors $\sum u_n$ est divergente.
 Si le rapport $\frac{u_{n+1}}{u_n}$ a une limite l (finie ou infinie) quand n tend vers $+\infty$, on a :
 ► Si $l < 1$, la série $\sum u_n$ converge.
 ► Si $l > 1$, la série $\sum u_n$ diverge grossièrement.
 ► Si $l = 1$, on ne peut pas conclure sur la nature de la série $\sum u_n$ par cette méthode.

Warnung : cette règle de d'Alembert n'est pas la règle universelle. Essayez avec les séries de Riemann.

Mise en œuvre : exercice 05 à exercice 10.

■ Séries absolument convergentes

Définition : Séries absolument convergentes —. La série numérique $\sum u_n$ est **absolument convergente** ou **ACV** lorsque la série $\sum |u_n|$ est convergente.

Proposition 9.13.— Toute série absolument convergente est convergente.

$$\text{En outre, } \left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|.$$

Proposition 9.14.— Série géométrique $\sum z^n$ —. La série géométrique $\sum z^n$, où $z \in \mathbf{C}$, est convergente si, et seulement si, $|z| < 1$; dans ce cas, elle est même absolument convergente et

$$\forall z \in \mathbf{C} \text{ tel que } |z| < 1, \sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$$

Théorème-Définition 9.15.— Série exponentielle —. Pour tout nombre complexe z , la série $\sum \frac{z^n}{n!}$ est absolument convergente. On définit la fonction \exp sur \mathbf{C} en posant :

$$\forall z \in \mathbf{C}, \exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

Mise en œuvre : exercice 11 à exercice 14.

Méthodes

Ici, on notera la somme partielle d'ordre n de la série $\sum u_n$, la quantité $S_n = \sum_{k=0}^n u_k$ et si cette série converge, on notera S sa somme.

■ Séries à termes toujours positifs (ou à termes toujours négatifs)

Dans tout le paragraphe, les séries considérées sont à termes positifs (on peut étendre bien entendu aux séries à termes négatifs).

□ Méthode 9.1.— Comment déterminer la nature de $\sum u_n$

- ▶ Si u_n ne tend pas vers 0, $\sum u_n$ diverge grossièrement.
- ▶ Si $u_n \sim_{+\infty} \frac{A}{n^\alpha}$ alors si $\alpha > 1$, $\sum u_n$ converge et si $\alpha \leq 1$, $\sum u_n$ diverge.
On pourra penser à la formule de Stirling si dans u_n apparaît $n!$
- ▶ Si $u_n \sim_{+\infty} Kq^n$ alors si $|q| < 1$, $\sum u_n$ converge et si $|q| \geq 1$, $\sum u_n$ diverge.
- ▶ Si l'on peut majorer u_n par v_n , (respectivement si $u_n = O(v_n)$) telle que $\sum v_n$ soit convergente alors $\sum u_n$ converge.
- ▶ Si l'on peut minorer u_n par $v_n \geq 0$ (respectivement si $v_n = O(u_n)$) telle que $\sum v_n$ soit divergente alors $\sum u_n$ est divergente.
- ▶ Si $\lim_{n \rightarrow +\infty} n^\alpha u_n = 0$ avec $\alpha > 1$, alors $u_n = o\left(\frac{1}{n^\alpha}\right)$ et $\sum u_n$ converge.
- ▶ Si $\lim_{n \rightarrow +\infty} n^\alpha u_n = +\infty$ avec $\alpha \leq 1$, alors $u_n > \frac{1}{n^\alpha}$ et $\sum u_n$ diverge.
- ▶ Si l'on fait un développement limité de u_n en $1/n$ et si $u_n = v_n + o\left(\frac{1}{n^\alpha}\right)$, où $\alpha > 1$, alors si $\sum v_n$ diverge, $\sum u_n$ diverge et si $\sum v_n$ converge, $\sum u_n$ converge.
- ▶ On peut aussi montrer : $\exists A \in \mathbf{R}, \forall n, S_n \leq A$. Alors $\sum u_n$ converge et $S \leq A$.
- ▶ On peut aussi (parfois) calculer la limite S directement. Si cette limite est finie, $\sum u_n$ converge. Allez voir la **méthode 9.7** pour calculer S .
- ▶ On peut utiliser la règle de d'Alembert. (Lorsque le terme général u_n s'exprime sous forme de produit ou de quotient et comporte des factorielles ou des puissances)
Voir **méthode 9.2** pour la façon de l'utiliser.
- ▶ On peut comparer avec une intégrale généralisée si $u_n = f(n)$, où f décroît, au moins à partir d'un certain rang de n . Voir **méthode 9.3**.

Exemples : 1. Étudier la nature de la série de terme général $\frac{1}{n \cos^2 n}$.

2. Étudier la nature de la série de terme général $\frac{n^2}{2^n + (-1)^n}$.

3. Étudier la nature de la série de terme général $e - \left(1 + \frac{1}{n}\right)^n$.

4. Déterminer la nature de la série de terme général $u_n = \binom{2n}{n} \frac{1}{4^n}$.

1. Pour tout $n \geq 1$, $\frac{1}{n \cos^2 n} \geq \frac{1}{n}$ et $\sum_{n \geq 1} \frac{1}{n}$ diverge, donc $\sum_{n \geq 1} \frac{1}{n \cos^2 n}$ diverge.

2. $0 \leq \frac{n^2}{2^n + (-1)^n} \sim \frac{n^2}{2^n} = o\left(\frac{1}{n^2}\right)$ car $\frac{n^4}{2^n} \rightarrow 0$; puisque $\sum_{n \geq 1} \frac{1}{n^2}$ converge, il en est de même pour la série numérique (à termes positifs) $\sum_{n \geq 1} \frac{n^2}{2^n + (-1)^n}$.

3. $\left(1 + \frac{1}{n}\right)^n = \exp\left(n \ln\left(1 + \frac{1}{n}\right)\right)$, et par ailleurs un petit développement limité à l'ordre 2 donne : $n \ln\left(1 + \frac{1}{n}\right) = n\left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\right) = 1 - \frac{1}{2n} + o\left(\frac{1}{n}\right)$ et donc : $\left(1 + \frac{1}{n}\right)^n = e \cdot \exp\left(-\frac{1}{2n} + o\left(\frac{1}{n}\right)\right)$, ce qui garantit que $v_n = e - \left(1 + \frac{1}{n}\right)^n$ tend vers 0 et est équivalent à $-\frac{1}{2n}$ (en utilisant $e^u - 1 \sim u$ quand u tend vers 0).

On en déduit que v_n est négatif pour n assez grand, et donc la série de terme général v_n est de même nature que la série harmonique ; $\sum v_n$ diverge.

4. En usant de James Stirling, $u_n = \frac{(2n)!}{(n!)^2} \frac{1}{4^n} \sim \frac{(2n)^{2n} e^{-2n} \sqrt{4\pi n}}{4^n n^{2n} e^{-2n} 2\pi n} = \frac{1}{\sqrt{\pi n}}$ et donc la série $\sum u_n$ diverge par comparaison avec une série de Riemann.

❑ Méthode 9.2.— Comment utiliser la règle de d'Alembert

- On s'assure que les u_n sont non nuls, au moins à partir d'un certain rang.
- On cherche la limite ℓ , si elle existe, de $\frac{u_{n+1}}{u_n}$:
 - ▷ si $\ell < 1$, $\sum u_n$ converge (absolument) ;
 - ▷ si $\ell > 1$, $\sum u_n$ diverge (grossièrement) ;
 - ▷ si $\ell = 1$, le critère ne permet pas de conclure.

Exemple : Soit $\alpha \in \mathbf{R}_+$, déterminer suivant la valeur du réel α la nature de la série de terme général $u_n = \binom{2n}{n} \alpha^n$.

$$\frac{u_{n+1}}{u_n} = \frac{\binom{2n+2}{n+1} \alpha^{n+1}}{\binom{2n}{n} \alpha^n} = \frac{(n!)^2 (2n+2)!}{(2n)! ((n+1)!)^2} \frac{\alpha^{n+1}}{\alpha^n} = \frac{(2n+2)(2n+1)}{(n+1)^2} \alpha = 2\alpha \frac{2n+1}{n+1} \rightarrow 4\alpha$$

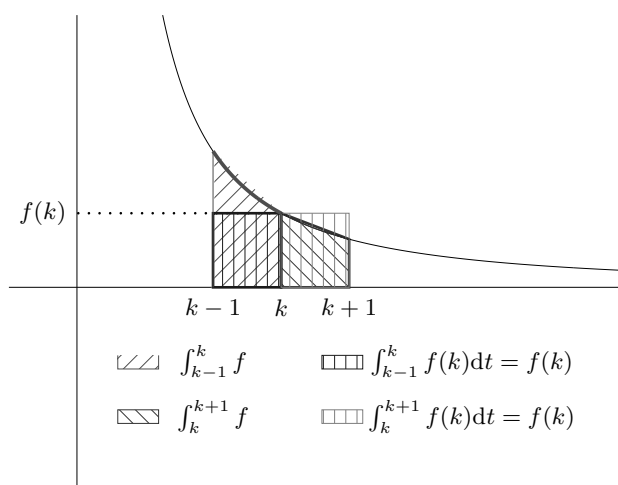
Il en résulte que si $\alpha < \frac{1}{4}$, alors $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} < 1$, donc la série $\sum u_n$ converge.

Si $\alpha > \frac{1}{4}$, alors $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} > 1$, donc la série $\sum u_n$ diverge.

Si $\alpha = \frac{1}{4}$, on ne peut pas conclure. (Voir à ce sujet l'exemple 4 qui suit la **méthode 9.1.**)

□ Méthode 9.3.— Comment encadrer des sommes partielles par la comparaison série-intégrale et en déduire la nature d'une série numérique

f est une fonction continue, positive et décroissante sur \mathbf{R}^+ . Pour déterminer la nature de la série $\sum f(n)$, on peut utiliser directement le théorème du cours permettant de se ramener à l'intégrabilité de f sur \mathbf{R}^+ . Mais, pour encadrer une somme partielle ou un reste de cette série, il est plus efficace de construire et d'interpréter la figure :



- $\int_{k-1}^k f(t) dt$ et $\int_k^{k+1} f(t) dt$ s'interprètent respectivement comme les aires des portions de plan "sous la courbe" :

$\{(x, y), k-1 \leq x \leq k, 0 \leq y \leq f(x)\}$ et $\{(x, y), k \leq x \leq k+1, 0 \leq y \leq f(x)\}$.
 $f(k)$ s'interprète indifféremment comme l'aire du rectangle $[k-1, k] \times [0, f(k)]$ ou l'aire du rectangle $[k, k+1] \times [0, f(k)]$.

- On a l'encadrement : $\int_k^{k+1} f(t) dt \leq f(k) \leq \int_{k-1}^k f(t) dt$.
 ► On somme membre à membre les encadrements précédents, en rajoutant, au besoin, $f(0)$ pour obtenir :

$$f(0) + \int_1^{n+1} f(t) dt \leq \sum_{k=0}^n f(k) \leq f(0) + \int_0^n f(t) dt$$

- On calcule les intégrales pour obtenir un encadrement.

Exemple : Étudier la nature de la série de terme général $u_n = \frac{1}{n \ln n (\ln(\ln n))^2}$.

La fonction $f : t \mapsto \frac{1}{t \ln t (\ln(\ln t))^2}$ est continue, décroissante et positive sur $]e, +\infty[$, donc la série

$\sum_{n \geq 3} \frac{1}{n \ln n (\ln(\ln n))^2}$ est convergente si et seulement si f est intégrable sur $[3, +\infty[$.

$\varphi : \begin{cases} [3, +\infty[& \rightarrow [\ln(\ln 3), +\infty[\\ t & \mapsto \ln(\ln t) \end{cases}$ est une bijection de classe \mathcal{C}^1 , donc c'est un changement de

variable valide pour une intégrale impropre sur $[3, +\infty[$. De plus : $\varphi'(t) = \frac{1}{t \ln t}$.

$$\int_3^{+\infty} \frac{dt}{t \ln t (\ln(\ln t))^2} = \int_{\ln 3}^{+\infty} \frac{du}{u^2} = \left[-\frac{1}{u} \right]_{\ln 3}^{+\infty} = \frac{1}{\ln 3}.$$

On a obtenu ainsi l'intégrabilité de f sur $[3, +\infty[$, donc la convergence de $\sum_{n \geq 3} \frac{1}{n \ln n (\ln(\ln n))^2}$.

■ Séries à termes quelconques

Dans tout le paragraphe, les séries considérées sont à termes dans \mathbf{R} ou dans \mathbf{C}

□ Méthode 9.4.— Comment étudier la nature d'une série alternée $\sum u_n$

On rappelle qu'ici $u_n = \varepsilon(-1)^n |u_n|$, où $\varepsilon = \pm 1$.

- Si $|u_n|$ ne tend pas vers 0, la série diverge grossièrement.
- Si la suite $(|u_n|)_{n \in \mathbf{N}}$ décroît et converge vers 0, on applique le critère des séries alternées. La série $\sum u_n$ converge. En bonus, on peut majorer $|S - S_n|$ par $|u_{n+1}|$. On dit que l'erreur commise en remplaçant S par S_n est inférieure à la valeur absolue du premier terme négligé.
- Si la suite $|u_n|$ tend vers 0 sans monotonie (ou si l'on ne peut pas le prouver), on fait un *D.L* de u_n à partir de la variable $1/n$. On se rapproche ici de la huitième piste de la **méthode 9.1**.

Exemples : 1. Justifier la convergence de $\sum \frac{(-1)^n}{2n+1}$ (de somme S). Déterminer $n_0 \in \mathbf{N}$ tel que $|S_n - S| \leq 10^{-2}$ et en déduire un encadrement de S de longueur 10^{-2} .

Il s'agit clairement d'une série alternée dont la valeur absolue du terme général, $\frac{1}{2n+1}$, décroît et tend vers 0 : cette série converge, d'après le critère spécial des séries alternées.

$S - S_n$ est le reste R_n d'ordre n de cette série convergente.

L'une des conclusions du critère spécial des séries alternées est que : $\forall n \in \mathbf{N}, |R_n| \leq |u_{n+1}|$.

Ici, $|S_n - S| \leq \frac{1}{2n+3}$.

Pour avoir $|S_n - S| \leq 10^{-2}$, il est suffisant d'avoir : $\frac{1}{2n+3} \leq \frac{1}{100}$, donc $n \geq 49$.

Finalement 49, ou tout entier supérieur ou égal à 49, est un entier n_0 possible.

De plus, S , comprise entre deux sommes partielles consécutives quelconques, est minorée par les sommes partielles d'indices impairs, d'où : $S_{49} \leq S \leq S_{49} + 10^{-2}$.

2. Étudier la convergence de $\sum_{n \geq 2} \frac{(-1)^n}{n + (-1)^n}$.

Posons $u_n = \frac{(-1)^n}{n + (-1)^n}$. On ne peut pas appliquer le critère des séries alternées.

En effet, si $p \in \mathbf{N}^*$, $|u_{2p}| = \frac{1}{2p+1}$ et $|u_{2p+1}| = \frac{1}{2p}$. On a :

$$|u_{2p}| < |u_{2p+1}|$$

donc $|u_n|$ ne décroît pas (même si sa limite est 0).

Un développement limité de u_n donne :

$$u_n = \frac{(-1)^n}{n} \left[\frac{1}{1 + \frac{(-1)^n}{n}} \right] = \frac{(-1)^n}{n} - \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)$$

Si l'on pose $v_n = \frac{(-1)^n}{n}$, on remarque que la série $\sum v_n$ suit le critère des séries alternées et est donc convergente. Si l'on pose $w_n = -\frac{1}{n^2}$, cette série est à termes tous négatifs et $\sum (-w_n)$ est à termes positifs et est convergente (voir la **proposition ??**). Enfin $o\left(\frac{1}{n^2}\right)$ représente le terme général d'une série absolument convergente donc convergente. Par somme, la série $\sum u_n$ est convergente.

□ Méthode 9.5.— Comment montrer qu'une série $\sum u_n$ est absolument convergente

On montre la convergence de la série à termes positifs $\sum |u_n|$, en utilisant la **méthode 9.1**, la **méthode 9.2** ou la **méthode 9.3**.

Exemple : Nature de la série $\sum_{n \geq 1} \frac{\sin n^2}{n^2}$.

Le terme général $u_n = \frac{\sin n^2}{n^2}$ n'est pas de signe constant à partir d'un certain rang.

$\forall n \geq 1$, $\left| \frac{\sin n^2}{n^2} \right| \leq \frac{1}{n^2}$ et $\sum_{n \geq 0} \frac{1}{n^2}$ est convergente, donc $\sum_{n \geq 1} \left| \frac{\sin n^2}{n^2} \right|$ est convergente.

$\sum_{n \geq 1} \frac{\sin n^2}{n^2}$ est absolument convergente, donc convergente.

□ **Méthode 9.6.— Comment étudier la nature (et trouver la somme éventuelle) d'une série $\sum u_n$ à termes complexes**

- On peut calculer $|u_n|$. Si cette quantité ne tend pas vers 0, il en est de même de u_n . Il y a divergence grossière. Si $|u_n| \sim v_n$, où v_n est le terme général d'une série convergente, alors $\sum u_n$ est absolument convergente donc convergente.
- On peut étudier les séries réelles $\sum \Re(u_n)$ et $\sum \Im u_n$. Si elles convergent simultanément, il en est de même de $\sum u_n$. Si l'une (au moins) des deux diverge, il en est de même de $\sum u_n$. Quant à la somme, en cas de convergence, sa partie réelle est la somme de $\sum \Re(u_n)$ et sa partie imaginaire est la somme de $\sum \Im(u_n)$.

Exemples : 1. Étudier la nature de la série $\sum_{n \geq 1} \frac{n(2+i)^n}{2^n}$.

On a :

$$|u_n| = \left| \frac{n(2+i)^n}{2^n} \right| = \frac{n}{2^n} |2+i|^n = n \left(\frac{\sqrt{5}}{2} \right)^n.$$

Cette quantité tend vers $+\infty$ et donc la série $\sum_{n \geq 1} \frac{n(2+i)^n}{2^n}$ diverge grossièrement.

2. Étudier la nature de la série $\sum_{n \geq 1} \frac{1}{\sqrt{n}+i}$.

On écrit :

$$u_n = \frac{1}{\sqrt{n}+i} = \frac{\sqrt{n}-i}{n+1} = \frac{\sqrt{n}}{n+1} - i \frac{1}{n+1}.$$

Les séries $\sum_{n \geq 1} \Re(u_n)$ et $\sum_{n \geq 1} \Im(u_n)$ divergent et il en est de même donc de $\sum_{n \geq 1} \frac{1}{\sqrt{n}+i}$.

Outre les méthodes utilisant les résultats relatifs aux séries entières que nous verrons dans un chapitre ultérieur, donnons les principaux moyens de déterminer la somme d'une série convergente.

□ **Méthode 9.7.— Comment déterminer la somme d'une série convergente $\sum u_n$**

- L'utilisation des séries géométriques quand u_n est de la forme Aq^n , où $-1 < q < 1$.
- L'utilisation de : $\forall z \in \mathbf{C}, \exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$.
- Le **principe du télescopage** quand le terme général u_n s'écrit facilement sous la forme $f(n) - f(n+1)$, où f est une application de \mathbf{R}^+ dans \mathbf{K} :

- On calcule la somme partielle d'ordre n par "télescopage" :

$$\begin{aligned} S_n &= \sum_{k=0}^n (f(k) - f(k+1)) = \sum_{k=0}^n f(k) - \sum_{k=0}^n f(k+1) \\ &= \sum_{k=0}^n f(k) - \sum_{k=1}^{n+1} f(k) = f(0) - f(n+1). \end{aligned}$$

- On détermine, si elle existe, la limite de f en $+\infty$ et on conclut sur la nature de la série et la valeur de sa somme éventuelle.
- Il existe des formes de télescopage plus complexes, par exemple :
 $\sum_{n=p}^q (f(n+1) - f(n-1))$ ou $\sum_{n=p}^q (f(n+1) - 2f(n) + f(n-1))$;
dans ce cas, on peut décomposer linéairement ces sommes partielles, puis décaler les indices pour faire apparaître des sommes communes qui se simplifient, ne laissant que quelques termes résiduels.
- Il existe des pistes plus originales. On peut calculer seulement S_{2p} ou seulement S_{2p+1} . Et même utiliser des sommes de Riemann !

Exemples : 1. Étudier la convergence de $\sum_{n \geq 0} \frac{5 \times 2^{n+1}}{3^{n-1}}$. Calculer alors sa somme.

Posons $u_n = \frac{5 \times 2^{n+1}}{3^{n-1}}$. On constate que : $\forall n \in \mathbf{N}, \frac{u_{n+1}}{u_n} = \frac{2}{3}$, donc $\sum u_n$ est une série géométrique de premier terme $u_0 = 30$ et de raison $\frac{2}{3} \in]-1, 1[$.

Cette série est donc convergente et sa somme est $30 \times \frac{1}{1 - \frac{2}{3}} = 90$.

2. Soit $\theta \in \mathbf{R}$, montrer la convergence de $\sum_{n \geq 0} \frac{\sin(n\theta)}{2^n}$, et calculer sa somme.

La série $\sum_{n \geq 0} \frac{e^{in\theta}}{2^n}$ converge absolument (voir **méthode 9.1**) car $\left| \frac{e^{in\theta}}{2^n} \right| = \frac{1}{2^n}$ et que $\sum_{n=0}^{+\infty} \frac{1}{2^n}$ converge, donc sa partie imaginaire converge ;

$$S = \sum_{n=0}^{+\infty} \frac{e^{in\theta}}{2^n} = \sum_{n=0}^{+\infty} \left(\frac{e^{i\theta}}{2} \right)^n = \frac{1}{1 - \frac{e^{i\theta}}{2}} = \frac{2}{2 - e^{i\theta}} = \frac{2}{2 - \cos \theta - i \sin \theta} = \frac{4 - 2 \cos \theta + 2i \sin \theta}{4 - 4 \cos \theta + \cos^2 \theta + \sin^2 \theta}.$$

Donc on en déduit : $\sum_{n=0}^{+\infty} \frac{\sin(n\theta)}{2^n} = \Im S = \frac{2 \sin \theta}{5 - 4 \cos \theta}$.

3. Convergence et somme de $\sum_{n \geq 1} \frac{1}{n(n+1)}$.

$$\forall n \geq 1, \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1} \text{ et } S_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \dots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}.$$

$$\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = \lim_{n \rightarrow +\infty} S_n = 1.$$

4. Convergence et somme de $\sum_{n \geq 0} \ln \cos \frac{a}{2^n}, \quad a \in \left]0, \frac{\pi}{2}\right[$.

Remarquons que, pour tout $k \in \mathbf{N}$, $\sin \frac{a}{2^{k-1}} = \sin \left(2 \frac{a}{2^k}\right) = 2 \cos \frac{a}{2^k} \sin \frac{a}{2^k} > 0$.

Et donc : $\ln \left(\cos \frac{a}{2^k}\right) = \ln \left(\frac{\sin \frac{a}{2^{k-1}}}{2 \sin \frac{a}{2^k}}\right) = \ln \left(2^{k-1} \sin \frac{a}{2^{k-1}}\right) - \ln \left(2^k \sin \frac{a}{2^k}\right)$, soit :

$$\sum_{k=0}^n \ln \left(\cos \frac{a}{2^k}\right) = \ln \left(\frac{1}{2} \sin(2a)\right) - \ln \left(2^n \sin \frac{a}{2^n}\right).$$

Or, $\sin \frac{a}{2^n} \sim \frac{a}{2^n}$ donc $2^n \sin \frac{a}{2^n}$ tend vers a en $+\infty$; $\sum_{n=0}^{+\infty} \ln \cos \frac{a}{2^n} = \ln \frac{\sin(2a)}{2a}$.

5. Convergence et somme de $\sum_{n \geq 2} \frac{2n-1}{n^3-n}$. Indication : $\frac{2n-1}{n^3-n} = \frac{1}{2(n-1)} + \frac{1}{n} - \frac{3}{2(n+1)}$.

$$\begin{aligned} \sum_{n=2}^m \frac{2n-1}{n^3-n} &= \sum_{n=2}^m \frac{1}{2(n-1)} + \sum_{n=2}^m \frac{1}{n} - \sum_{n=2}^m \frac{3}{2(n+1)} \\ &= \sum_{n'=1}^{m-1} \frac{1}{2n'} + \sum_{n=2}^m \frac{1}{n} - \sum_{n'=3}^{m+1} \frac{3}{2n} \\ &= \frac{1}{2} - \frac{1}{2m} + \left(\frac{1}{2} + 1 - \frac{3}{2}\right) \left(\sum_{n=2}^m \frac{1}{n}\right) - \frac{3}{2(m+1)} + \frac{3}{4} \\ &= \frac{5}{4} - \frac{1}{2m} - \frac{3}{2(m+1)} \text{ qui tend vers } \frac{5}{4} \text{ quand } m \text{ tend vers } +\infty. \end{aligned}$$

6. On pose : $\forall n \geq 1, u_n = \frac{(-1)^n}{n}$. La série $\sum_{n \geq 1} u_n$ est-elle absolument convergente ?

On pose alors : $\forall p \in \mathbf{N}^*, S_{2p} = \sum_{k=1}^{2p} u_k$, montrer : $S_{2p} = -\frac{1}{p} \sum_{k=1}^p \frac{1}{1 + \frac{k}{p}}$.

En déduire $\sum_{n=1}^{+\infty} u_n$ et retrouver la nature de $\sum_{n \geq 1} u_n$.

► Comme $|u_n| = \frac{1}{n}$, il est clair que $\sum_{n \geq 1} u_n$ n'est pas absolument convergente.

► On écrit : $S_{2p} = -1 + \frac{1}{2} - \frac{1}{3} + \dots - \frac{1}{2j+1} + \frac{1}{2j} + \dots - \frac{1}{2p+1} + \frac{1}{2p}$.

Or, $-\frac{1}{p} \sum_{k=1}^p \frac{1}{1+\frac{k}{p}} = -\sum_{k=1}^p \frac{1}{p+k} = -\frac{1}{p+1} - \frac{1}{p+2} - \dots - \frac{1}{2p-1} - \frac{1}{2p}.$

On écrit alors : $S_{2p} = -1 - \frac{1}{3} - \dots - \frac{1}{2j+1} - \dots - \frac{1}{2p-1} + \frac{1}{2} \left[1 + \dots + \frac{1}{j} + \dots + \frac{1}{p} \right].$

Ou encore : $S_{2p} = -1 - \frac{1}{2} - \dots - \frac{1}{2p-1} - \frac{1}{2p} + \frac{1}{2} + \dots + \frac{1}{2p} + \frac{1}{2} \left[1 + \dots + \frac{1}{p} \right].$

Ainsi, $S_{2p} = \left[-1 - \frac{1}{2} - \dots - \frac{1}{2p-1} - \frac{1}{2p} \right] + \left[1 + \dots + \frac{1}{p} \right]$, ce qui donne bien la forme indiquée.

► Si l'on pose $f(t) = \frac{1}{1+t}$, on reconnaît en $-\frac{1}{p} \sum_{k=1}^p \frac{1}{1+\frac{k}{p}}$ une somme de Riemann.

Et : $\lim_{p \rightarrow +\infty} S_{2p} = -\int_0^1 f(t) dt = -\ln 2.$ Comme $S_{2p+1} = S_{2p} - \frac{1}{2p+1}$, $\lim_{p \rightarrow +\infty} S_{2p+1} = -\ln 2.$

Donc : $\sum_{n=1}^{+\infty} u_n = -\ln 2$ et la série $\sum_{n \geq 1} u_n$ converge.

❏ Méthode 9.8.— Comment déterminer un équivalent du reste d'une somme partielle de la série $\sum u_n$

► Dans le cas d'une série convergente dont on connaît la somme S et l'expression S_n pour tout n , alors $R_n = S - S_n$ s'écrit sous la forme $g(n)$. On effectue un développement limité en $1/n$.

► Si $u_n = f(n)$, où f est continue, positive et décroissante sur \mathbf{R}^+ , on utilise :

$$\int_{n+1}^{+\infty} f(t) dt \leq \sum_{k=n+1}^{+\infty} f(k) \leq \int_n^{+\infty} f(t) dt.$$

Cette double inégalité peut permettre d'obtenir un équivalent du reste d'ordre n .

Exemples : 1. Soit la série de terme général : $u_n = \ln \left(1 - \frac{1}{n^2} \right)$ définie pour $n \geq 2$.

Montrer que u_n peut s'écrire sous la forme $u_n = \phi(n) - \phi(n-1)$.

En déduire la somme de la série. Montrer que $R_n = \sum_{k=n+1}^{+\infty} u_k \sim \frac{-1}{n}.$

On commence par écrire : $u_n = \ln \left(\frac{n^2-1}{n^2} \right) = \ln \left(\frac{n-1}{n} \right) + \ln \left(\frac{n+1}{n} \right).$

Si l'on pose $\phi(n) = \ln(n+1) - \ln n$, on a bien : $u_n = \phi(n) - \phi(n-1).$

Par télescopage, on en déduit la somme de la série :

$$S_n = \sum_{k=2}^n [\phi(k) - \phi(k-1)] = \phi(n) - \phi(1),$$

c'est-à-dire : $S_n = \phi(n) - \ln 2.$ Donc : $\lim_{n \rightarrow +\infty} \phi(n) = 1 \Rightarrow \sum_{n=2}^{+\infty} u_n = -\ln 2.$

Posons $R_n = \sum_{k=n+1}^{+\infty} u_k$. Cette quantité existe et tend vers 0 quand n tend vers $+\infty$ car la série associée est convergente. De plus : $R_n = S - S_n = -\ln 2 - (\phi(n) - \ln 2) = -\phi(n)$.

Or, quand n tend vers $+\infty$, $\phi(n) = \ln\left(1 + \frac{1}{n}\right) \sim \frac{1}{n} \Rightarrow R_n \sim \frac{-1}{n}$.

2. a. Soient deux séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ à termes positifs telles que $\sum_{n \geq 0} v_n$ soit convergente.

Montrer, au voisinage de $+\infty$, que : $\left[u_n \sim v_n \Rightarrow \sum_{p=n}^{+\infty} u_p \sim \sum_{p=n}^{+\infty} v_p \right]$.

Indication : on utilisera le fait que : $u_n \sim v_n \Leftrightarrow u_n - v_n = o(v_n)$.

b. Soit $\alpha > 1$ et $u_n \sim_{+\infty} \frac{1}{n^\alpha}$ alors montrer que : $\sum_{p=n}^{+\infty} u_p \sim_{+\infty} \frac{1}{(\alpha-1)n^{\alpha-1}}$.

a. Comme $u_n \sim v_n$, $u_n - v_n = o(v_n)$ et donc :

$$\forall \epsilon > 0, \exists N_0 \in \mathbf{N}, \forall n \geq N_0, -\epsilon v_n \leq u_n - v_n \leq \epsilon v_n.$$

Posons $S_n = \sum_{k=0}^n u_k$ et $T_n = \sum_{k=0}^n v_k$. Pour tout $n \geq N_0$ et tout $p \in \mathbf{N}$,

$$-\epsilon(T_{n+p} - T_n) \leq S_{n+p} - S_n \leq \epsilon(T_{n+p} - T_n).$$

On fait tendre p vers $+\infty$. En posant $A_n = \sum_{p=n}^{+\infty} u_p$ et $Z_n = \sum_{p=n}^{+\infty} v_p$ alors :

$$\epsilon Z_n \leq A_n - Z_n \leq \epsilon Z_n.$$

(En effet, les suites S_n et T_n étant convergentes, il en est de même de $(T_{n+p} - T_n)$ et de $(S_{n+p} - S_n)$. Finalement, $A_n - Z_n = o(Z_n)$ et on a bien : $A_n \sim Z_n$.)

b. Soit $\alpha > 1$ et partons maintenant de l'intégration (à faire) :

$$\int_n^{n+1} \frac{dt}{t^\alpha} = \frac{1}{\alpha-1} \left(\frac{1}{n^{\alpha-1}} - \frac{1}{(n+1)^{\alpha-1}} \right)$$

Par ailleurs :

$$\frac{1}{(n+1)^\alpha} < \int_n^{n+1} \frac{dt}{t^\alpha} < \frac{1}{n^\alpha} \Rightarrow \int_n^{n+1} \frac{dt}{t^\alpha} \sim \frac{1}{n^\alpha}.$$

Donc si $u_n \sim \frac{1}{n^\alpha}$ alors $u_n \sim \int_n^{n+1} \frac{dt}{t^\alpha}$ par transitivité des équivalences.

Comme la série de terme général $\frac{1}{n^\alpha}$ est convergente, en utilisant le résultat précédent :

$$\sum_{p=n}^{+\infty} u_p \sim \int_n^{+\infty} \frac{dt}{t^\alpha}$$

Or $\int_n^{+\infty} \frac{dt}{t^\alpha}$ vaut $\frac{1}{(\alpha-1)n^{\alpha-1}}$. On en déduit le résultat.

3. On considère la série $\sum_{n \geq 2} \frac{1}{n(\ln n)^\alpha}$, où $\alpha > 1$.

a. Montrer sa convergence.

b. Encadrer son reste R_n d'ordre n . En déduire un équivalent de R_n .

a. La fonction $f : x \mapsto \frac{1}{x(\ln x)^\alpha}$ est décroissante et positive sur $[2, +\infty[$ donc la série :

$$\sum_{n \geq 2} \frac{1}{n(\ln n)^\alpha},$$

de terme général $f(n)$, converge si et seulement si f est intégrable sur $[2, +\infty[$. Or, $\forall X \geq 2$,

$$\int_2^X \frac{dx}{x(\ln x)^\alpha} = \left[\frac{(\ln x)^{-\alpha+1}}{-\alpha+1} \right]_2^X = \frac{1}{(1-\alpha)(\ln X)^{\alpha-1}} - \frac{1}{(1-\alpha)(\ln 2)^{\alpha-1}} \xrightarrow{X \rightarrow +\infty} \frac{1}{(\alpha-1)(\ln 2)^{\alpha-1}}.$$

En conclusion, f est intégrable sur $[2, +\infty[$, donc $\sum_{n \geq 2} \frac{1}{n(\ln n)^\alpha}$ converge.

b. En exploitant la figure décrite dans la méthode :

$$\forall k \geq 3, \int_k^{k+1} f(t) dt \leq \frac{1}{k(\ln k)^\alpha} \leq \int_{k-1}^k f(t) dt.$$

En sommant de $n+1$ à N , puis en faisant tendre N vers $+\infty$, on obtient successivement, par la relation de Chasles :

$$\int_{n+1}^{N+1} f(t) dt \leq \sum_{k=n+1}^N \frac{1}{k(\ln k)^\alpha} \leq \int_n^N f(t) dt.$$

$$\int_{n+1}^{+\infty} f(t) dt \leq R_n \leq \int_n^{+\infty} f(t) dt.$$

D'après le calcul fait à la première question, en remplaçant 2 successivement par $n+1$ et n :

$$\frac{1}{(\alpha-1)(\ln(n+1))^{\alpha-1}} \leq R_n \leq \frac{1}{(\alpha-1)(\ln n)^{\alpha-1}}.$$

De cet encadrement, il est facile de déduire :

$$R_n \sim \frac{1}{(\alpha-1)(\ln n)^{\alpha-1}}.$$

Chapitre 10

Power series

■ Objectifs

■ Les incontournables :

- ▶ Savoir déterminer et interpréter le rayon de convergence d'une série entière.
- ▶ Connaître le mode de convergence d'une série entière sur l'intervalle ou le disque de convergence.
- ▶ Savoir manipuler les séries entières (combinaisons linéaires, dérivation, intégration, décalage d'indices).
- ▶ Connaître les développements classiques en série entière.
- ▶ Utiliser les savoir-faire ci-dessus pour
 - ▷ calculer une somme de série entière;
 - ▷ développer en série entière une fonction définie par une formule analytique.

■ Et plus si affinités ...

- ▶ Exprimer certaines intégrales « non calculables » comme des sommes de séries numériques.
- ▶ Savoir calculer les sommes de certaines séries numériques.

Résumé de cours

■ Définition, convergence

Définition : Série entière —.

Une **série entière** de la variable réelle x (respectivement complexe z) est une série de la forme

$$\sum_{n \geq 0} a_n x^n \text{ (respectivement } \sum_{n \geq 0} a_n z^n),$$

où $(a_n)_{n \in \mathbf{N}}$ est une suite à termes réels (respectivement complexes), avec la convention $x^0 = 1$ (respectivement $z^0 = 1$) (même si $x = 0$ ou $z = 0$).

Dans la suite, on énonce les résultats dans \mathbf{C} (qu'on adapte au cas réel sans souci).

Théorème 10.1.— Lemme d'Abel —.

Soit $z_0 \in \mathbf{C}^*$, si la suite $(a_n z_0^n)_{n \in \mathbf{N}}$ est bornée alors, pour tout nombre complexe z tel que $|z| < |z_0|$, la série $\sum_{n \geq 0} a_n z^n$ est absolument convergente.

Proposition 10.2.— Pour $\sum_{n \geq 0} a_n z^n$, donnée, il existe R unique dans $\mathbf{R}_+ \cup \{+\infty\}$ tel que :

1. $|z| < R \Rightarrow \sum_{n \geq 0} a_n z^n$ converge absolument
2. $|z| > R \Rightarrow \sum_{n \geq 0} a_n z^n$ diverge grossièrement

Plus précisément : $R = \sup \left\{ |z|, (|a_n z^n|)_{n \geq 0} \text{ est bornée} \right\} = \inf \left\{ |z|, \sum_{n \geq 0} a_n z^n \text{ diverge} \right\}$

si $R = 0$, la propriété $|z| < R$ est fausse pour tout élément de \mathbf{C} , et de même, si $R = +\infty$, la propriété $|z| > R$ est fausse pour tout élément de \mathbf{C} , et les implications ci-dessus sont vraies, en vertu du principe de logique : « si p est fausse, $p \Rightarrow q$ est vraie quelle que soit q ».

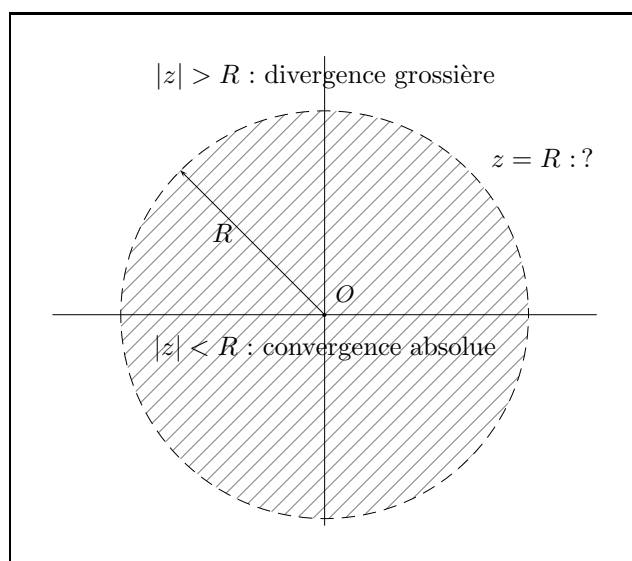
Définition : Rayon et disque de convergence —. Le réel R défini en **proposition 10.2** est appelé **rayon de convergence** de la série entière $\sum_{n \geq 0} a_n z^n$.

Le disque (ouvert) $D_R = \mathcal{B}(0, R)$ est appelé **disque de convergence** de cette série entière : si $R = 0$, alors $D_R = \emptyset$, et si $R = \infty$, alors $D_R = \mathbf{C}$.

Proposition 10.3.— D'après la **proposition 10.1**, toute série entière converge absolument en tout point de son disque de convergence, et diverge grossièrement en tout point extérieur au disque de convergence (c'est-à-dire en tout point z tel que $|z| > R$.) Aucune propriété générale ne précise le comportement de la série sur le **cercle d'incertitude** de centre O et de rayon R .

Remarque : Notez les cas particuliers importants :

si $R = +\infty$: f est définie sur \mathbf{C} tout entier et si $R = 0$: f n'est définie qu'en $z = 0$.



Proposition 10.4.— Règle de d'Alembert pour les séries entières non lacunaires —.

Soit $\sum_{n \geq 0} a_n z^n$ une série entière **non lacunaire**, c'est-à-dire dont les coefficients a_n sont tous non nuls à partir d'un certain rang et de rayon de convergence R . Alors :

$$\text{Si } \lim_{n \rightarrow +\infty} \frac{|a_{n+1}|}{|a_n|} = \ell \in [0, +\infty] \Rightarrow \begin{cases} \text{si } \ell = 0 & , \quad R = +\infty \\ \text{si } \ell \in]0, +\infty[& , \quad R = \frac{1}{\ell} \\ \text{si } \ell = +\infty & , \quad R = 0 \end{cases} .$$

Remarques : 1. La série entière $\sum_{n \geq 0} \frac{1}{n+1} z^{2n+1}$ est lacunaire, en effet, $a_{2n} = 0$ et $a_{2n+1} \neq 0$.

On ne peut donc pas appliquer la **proposition 10.4**. Voir la **méthode 10.1**.

2. La série $\sum_{n \geq 1} n^\alpha x^n$ a pour rayon de convergence 1 pour tout réel α .

En effet, en posant $a_n = n^\alpha \neq 0$ pour tout $n \in \mathbf{N}^*$ et en appliquant la **proposition 10.4**, le rapport $\frac{|a_{n+1}|}{|a_n|} = \left(1 + \frac{1}{n}\right)^\alpha$ tend vers 1 quand n tend vers $+\infty$ (ou vaut 1 si $\alpha = 0$).

Mise en œuvre : exercice 01 à exercice 04

■ Opérations sur les séries entières

Dans tout le paragraphe, R_a (respectivement R_b) est le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$ (respectivement $\sum_{n \geq 0} b_n z^n$). Leurs sommes respectives sont f et g .

Proposition 10.5.—

(i) Si l'on a : $|a_n| \leq |b_n|$ alors : $R_a \geq R_b$. (ii) Si l'on a : $|a_n| \sim |b_n|$ alors : $R_a = R_b$.

Remarque : Dans le cas (i), on peut remplacer par $a_n = O(b_n)$ ou $a_n = o(b_n)$.

Proposition 10.6.— Combinaison linéaire —. Si λ et μ sont deux complexes *non nuls* alors la série entière de somme $\lambda f + \mu g$ a un rayon de convergence $R \geq \inf(R_a, R_b)$, avec égalité si $R_a \neq R_b$.

■ Séries entières d'une variable réelle

Soit $\sum_{n \geq 0} a_n x^n$ une série entière de rayon de convergence R *non nul*.

Sur $] -R, R[$, $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est \mathcal{C}^∞ (donc en particulier continue) et on peut dériver et intégrer terme à terme.

Proposition 10.7.— ▶ $f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n$.

▶ $f^{(p)}(x) = \sum_{n=p}^{+\infty} n(n-1) \cdots (n-p+1) a_n x^{n-p} = \sum_{m=0}^{+\infty} (m+p)(m+p-1) \cdots (m+1) a_{m+p} x^m$.

▶ $F(x) = \int_0^x f(t) dt = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$.

Les séries entières de sommes f' , $f^{(p)}$ et F ont le même rayon de convergence que f .

Remarque : En particulier, pour tout $[\alpha, \beta] \subset] -R, R[$, $\int_\alpha^\beta \sum_{n=0}^{+\infty} a_n t^n dt = \sum_{n=0}^{+\infty} \int_\alpha^\beta a_n t^n dt$.

Corollaire 10.8.— Principe d'identification —. Soit $\sum_{n \geq 0} a_n x^n$ et $\sum_{n \geq 0} b_n x^n$ deux séries entières de rayon de convergence non nul dont les sommes coïncident sur un voisinage de 0, alors :

$$\forall n \in \mathbf{N}, a_n = b_n \text{ (et bien sûr, les deux séries ont même rayon).}$$

En particulier, le développement en série entière d'une fonction paire (respectivement impaire) ne contient que des termes de degré pair (respectivement impair).

Définition : Fonction développable en série entière —.

Une fonction f définie sur $I \subset \mathbf{R}$ est **développable en série entière au voisinage de 0** sur $] -r, r[\subset I$ (avec $r > 0$) si et seulement s'il existe une série entière $\sum_n a_n x^n$ de rayon de convergence r telle que :

$$\forall x \in] -r, r[, f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Corollaire 10.9.— Si f est développable en série entière au voisinage de 0 alors f est de classe \mathcal{C}^∞ au voisinage de 0.

Théorème 10.10.— Si f admet un développement en série entière au voisinage de 0 sur $] - r, r[$ alors ce développement est unique.

Proposition 10.11.— Si $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ est la somme d'une série entière de rayon de convergence R non nul, alors $a_n = \frac{f^{(n)}(0)}{n!}$ pour $n \in \mathbf{N}$.

■ Développements classiques en série entière d'une fonction à variable réelle

(1)	$\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n$	$R = 1$
(2)	$\ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n$	$R = 1$
(3)	$e^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$	$R = \infty$
(4)	$\cos x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n}$	$R = \infty$
(5)	$\sin x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$	$R = \infty$
(6)	$\operatorname{ch} x = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^{2n}$	$R = \infty$
(7)	$\operatorname{sh} x = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!} x^{2n+1}$	$R = \infty$
(8)	$\arctan x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$	$R = 1$

Si $\alpha \in \mathbf{R}$, en posant $\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} = \prod_{k=1}^n \frac{\alpha-k+1}{k}$ et $\binom{\alpha}{0} = 1$,

$$(9) \quad (1+x)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n \quad \text{avec} \quad R = \begin{cases} 1 & \text{si } \alpha \notin \mathbf{N} \\ +\infty & \text{si } \alpha \in \mathbf{N} \end{cases}$$

Remarque : Le développement en série entière de $(1+x)^{-1/2}$ permet de déterminer par substitution ceux de $\frac{1}{\sqrt{1-x^2}}$ et de $\frac{1}{\sqrt{1+x^2}}$, puis par intégration terme à terme celui de $\arcsin x$ (et de $\arccos x$).

■ **Extension admise : développement en série entière de l'exponentielle à variable complexe**

$$e^z = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n \text{ avec } R = \infty$$

Remarque : Il y a deux types d'exercices principaux.

Développer une fonction connue en série entière. **Mise en œuvre : exercice 05 à exercice 09**

Déterminer la somme d'une série entière. **Mise en œuvre : exercice 10 à exercice 17.**

On exclut ici la méthode utilisant les équations différentielles linéaires car ce sera dans le prochain chapitre.

Méthodes

■ Rayon de convergence d'une série entière

□ Méthode 10.1.— Comment déterminer un rayon de convergence.

- ▶ On peut utiliser la règle de D'Alembert pour les séries numériques.
- ▶ On peut encadrer le rayon de convergence, en utilisant :
 - ▶ la convergence ou la divergence de la série en un point z_0 de module connu, puis la **proposition 10.2** ;
 - ▶ une comparaison du coefficient a_n avec celui d'une série de rayon de convergence connu et en utilisant la **proposition 10.5**.

Exemples : 1. Déterminer le rayon de la série entière $\sum_{n \geq 0} \binom{2n}{n} z^{2n+1}$.

Posons $u_n(z) = \binom{2n}{n} z^{2n+1}$. Alors : $\left| \frac{u_{n+1}(z)}{u_n(z)} \right| = \frac{2(n+1)!}{((n+1)!)^2} \cdot \frac{((n)!)^2}{(2n)!} |z|^2$.

Ainsi : $\lim_{n \rightarrow +\infty} \left| \frac{u_{n+1}(z)}{u_n(z)} \right| = \lim_{n \rightarrow +\infty} \frac{(2n+1)(2n+2)}{(n+1)^2} |z|^2 = 4|z|^2$.

Alors, si $0 < |z_0| < \frac{1}{2}$, $\lim_{n \rightarrow +\infty} \left| \frac{u_{n+1}(z_0)}{u_n(z_0)} \right| = 4|z_0|^2 < 1$ et $\sum_{n \geq 0} \binom{2n}{n} z_0^{2n+1}$ converge. Donc $R \geq \frac{1}{2}$.

Si $|z_0| > \frac{1}{2}$, $\lim_{n \rightarrow +\infty} \left| \frac{u_{n+1}(z_0)}{u_n(z_0)} \right| = 4|z_0|^2 > 1$ et $\sum_{n \geq 0} \binom{2n}{n} z_0^{2n+1}$ diverge. Donc $R \leq \frac{1}{2} \Rightarrow R = \frac{1}{2}$.

2. On considère la série entière $\sum p_n z^n$, où p_n est la $n^{\text{ème}}$ décimale du nombre π : $(p_n)_{n \in \mathbf{N}} = (3, 1, 4, 1, 5, \dots)$; déterminer son rayon de convergence R .

On considère la fonction g définie par la somme de la série entière : $g(z) = \sum_{n=0}^{+\infty} p_n z^n$.

Bien sûr, $g(0.1) = \sum_{n=0}^{+\infty} p_n 10^{-n} = \pi$ (suite décimale illimitée) donc $R \geq 0.1$.

Par ailleurs, comme $0 \leq p_n \leq 9$ par définition, $\sum_{n \geq 0} p_n z_0^n$ converge absolument si $|z_0| < 1$

(d'après les théorèmes de comparaison des séries à termes positifs), donc $R \geq \sup_{|z| < 1} |z| = 1$.

Par ailleurs, $\sum_{n \geq 0} p_n$ diverge grossièrement : la suite (p_n) , à valeurs entières, ne peut tendre vers 0

que si elle stationne en 0, ce qui impliquerait que π soit décimal. On en conclut que $R \leq 1$.

Finalement, $R = 1$.

■ Somme d'une série entière

□ Méthode 10.2.— Comment déterminer la somme d'une série entière

Soit la série entière $\sum_{n \geq p} a_n x^n$ de somme $S(x)$, où $x \in \mathbf{R}$ ou $x \in \mathbf{C}$.

- ▶ On commence par déterminer, si c'est possible, son rayon de convergence R . Sinon on admet que $R > 0$.
- ▶ On peut calculer $S(x)$ en se ramenant à des combinaisons linéaires de sommes :
$$\sum_{n \geq p} x^n, \sum_{n \geq p} nx^n, \sum_{n \geq p} n(n-1)x^n, \sum_{n \geq p} \frac{1}{n}x^n, \sum_{n \geq p} \frac{1}{n!}x^n, \text{ etc.}$$
On conclut en utilisant les opérations de dérivation, intégration, multiplication par x^p , changement d'indice, changement de variable $y = x^2$, etc.
- ▶ On peut aussi passer par $S'(x)$ ou $S''(x)$, s'ils sont plus simples à calculer.
- ▶ On peut "pousser" la piste précédente en aboutissant à une équation différentielle linéaire d'inconnue S et de valeur initiale $S(0)$. On se limite à des équations d'ordre un dans ce chapitre (pour l'ordre deux, allez au chapitre consacré aux équations différentielles).
- ▶ On peut aussi partir d'une relation entre différents termes consécutifs de la suite (a_n) puis multiplier par x^n et sommer ensuite. On trouve alors une égalité où intervient $S(x)$. On en déduit cette quantité.

Exemple : Déterminer la valeur de $f(x) = \sum_{n=3}^{+\infty} \frac{x^n}{n(n-1)(n-2)}$, si $x \in]-1, 1[$.

On pourra rechercher des réels α, β et γ tels que $\frac{1}{n(n-1)(n-2)} = \frac{\alpha}{n} + \frac{\beta}{n-1} + \frac{\gamma}{n-2}$.

La série entière $\sum_{n \geq 3} \frac{x^n}{n(n-1)(n-2)}$ a pour rayon 1. On travaille sur $]-1, 1[$.

On réduit : $\frac{\alpha}{n} + \frac{\beta}{n-1} + \frac{\gamma}{n-2} = \frac{2\alpha + n(-3\alpha - 2\beta - \gamma) + n^2(\alpha + \beta + \gamma)}{n(n-1)(n-2)}$ puis on identifie et on trouve $\alpha = \frac{1}{2}$, $\beta = -1$ et $\gamma = \frac{1}{2}$. Alors, pour tout $x \in]-1, 1[$,

$$\begin{aligned} f(x) &= \sum_{n=3}^{+\infty} \frac{x^n}{2n} - \sum_{n=3}^{+\infty} \frac{x^n}{n-1} + \sum_{n=3}^{+\infty} \frac{x^n}{2(n-2)} \\ &= \sum_{n=1}^{+\infty} \frac{x^n}{2n} - \frac{x}{2} - \frac{x^2}{4} - x \sum_{n=3}^{+\infty} \frac{x^{n-1}}{n-1} + x^2 \sum_{n=3}^{+\infty} \frac{x^{n-2}}{2(n-2)} \\ &= \sum_{n=1}^{+\infty} \frac{x^n}{2n} - \frac{x}{2} - \frac{x^2}{4} - x \sum_{m=2}^{+\infty} \frac{x^m}{m} + x^2 \sum_{p=1}^{+\infty} \frac{x^p}{2p} \\ &= \frac{-1}{2} \ln(1-x) - \frac{x}{2} - \frac{x^2}{4} - x(-\ln(1-x) - x) - \frac{x^2}{2} \ln(1-x) \\ &= \frac{1}{4}(-2x + 3x^2) - \frac{1}{2}(x-1)^2 \ln(1-x). \end{aligned}$$

On peut aussi intégrer deux fois $f''(x)$. En particulier, pour tout $x \in]-1, 1[$, on a :

$$f''(x) = \sum_{n=3}^{+\infty} \frac{x^{n-2}}{n-2} = -\ln(1-x) \Rightarrow f'(x) = x + \ln(1-x) - x \ln(1-x).$$

Exemple : Pour $n \geq 1$, on pose $a_n = \sum_{k=1}^n \frac{1}{k}$.

1. Déterminer le rayon R de $\sum_{n \geq 1} a_n x^n$. Pour $x \in]-R, R[$, on pose : $f(x) = \sum_{n=1}^{+\infty} a_n x^n$.
2. Déterminer $a_n - a_{n-1}$ pour $n \geq 2$ et en déduire une expression simple de $f(x)$.

1. On pose $u_n = a_n x^n$. Alors : $\lim_{n \rightarrow +\infty} \left| \frac{u_{n+1}}{u_n} \right| = \left| \frac{\sum_{k=1}^n \frac{1}{k} + \frac{1}{n+1}}{\sum_{k=1}^n \frac{1}{k}} x \right| = |x|$ car $\lim_{n \rightarrow +\infty} a_n = +\infty$.

On peut en déduire que si $|x| < 1$, alors $\sum_{n \geq 1} a_n x^n$ est absolument convergente et si $|x| > 1$, la série

diverge (même grossièrement). Le rayon de convergence est $R = 1$.

2. Il est clair que $k \geq 2 \Rightarrow a_k - a_{k-1} = \frac{1}{k}$.

On multiplie par x^k de chaque côté dans cette égalité, puis on somme jusqu'à $n \geq 2$:

$$\sum_{k=2}^n a_k x^k - \sum_{k=2}^n a_{k-1} x^k = \sum_{k=2}^n \frac{x^k}{k} = \sum_{k=1}^n a_k x^k - a_1 x - \sum_{k=1}^n a_k x^{k+1}.$$

On aboutit alors (en faisant tendre n vers $+\infty$) à : $f(x) - a_1 x - \sum_{k=1}^{+\infty} a_k x^{k+1} = -\ln(1-x) - x$.

Comme $a_1 = 1$, il reste (pour $|x| < 1$) : $f(x) - x f(x) = -\ln(1-x) \Rightarrow f(x) = -\frac{\ln(1-x)}{1-x}$.

Exemple : Calculer $S(x) = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$ (sans le formulaire du cours!).

On montre rapidement que cette série entière est de rayon de convergence $R = +\infty$ et que de plus $S'(x) = S(x)$ pour tout $x \in \mathbf{R}$. Comme S est solution de l'équation différentielle $y' - y = 0$, il existe $K \in \mathbf{R}$ tel que $S(x) = K e^x$ pour tout $x \in \mathbf{R}$. Comme $S(0) = 1$, $S(x) = e^x$.

■ Développement en série entière

□ Méthode 10.3.— Comment montrer qu'une fonction est développable en série entière

Une combinaison linéaire ou un produit de fonctions développables en série entière sont développables en série entière. La dérivée ou une primitive d'une fonction développable en série entière l'est aussi sur le même intervalle.

Exemple : Montrer que $x \mapsto e^x \int_0^x \frac{dt}{1+t}$ est développable en série entière.

$x \mapsto e^x$ est développable en série entière sur \mathbf{R} . $x \mapsto \frac{1}{1+x}$ est développable en série entière sur $] -1, 1[$ donc sa primitive $x \mapsto \int_0^x \frac{dt}{1+t}$ est développable en série entière sur le même intervalle. Le produit $x \mapsto e^x \int_0^x \frac{dt}{1+t}$ est donc développable en série entière sur $] -1, 1[$.

□ **Méthode 10.4.— Comment obtenir un développement en série entière d'une certaine fonction**

- Certaines fonctions (en particulier celles dont l'expression contient des logarithmes ou des arctangentes) ont une dérivée simple, dont on peut déterminer un développement en série entière.
Quand on l'a obtenu, il ne reste plus qu'à intégrer terme à terme (et ajuster comme il faut la constante d'intégration).
- On peut "pousser" la piste précédente en déterminant une équation différentielle vérifiée par la fonction $f(x)$ puis en posant $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ et en supposant que $R > 0$. On détermine des relations entre des termes consécutifs de la suite (a_n) . La donnée de $f(0)$ et éventuellement de $f'(0)$ donne l'initialisation de la suite. Tout ceci permet d'avoir a_n en fonction de n et le rayon de convergence. Cette piste est bien développée et expliquée dans la **méthode 11.9**, qui se trouve dans le chapitre sur les équations différentielles. Nous nous limitons par contre dans ce chapitre à des équations différentielles linéaires d'ordre 1.
- On peut développer la fonction en combinaison linéaire d'autres fonctions dont on connaît les développements en série entière (on peut pour cela faire une décomposition en éléments simples ou une linéarisation, on peut utiliser aussi des formules du genre $\ln(PQ) = \ln P + \ln Q$).
- On peut utiliser la série de Taylor de f , à condition d'avoir prouvé que le reste de Taylor-Maclaurin tend vers 0. Voir la **proposition ??**.

Exemple : Déterminer les développements en série entière de :

1. $x \mapsto \arctan x$. 2. $x \mapsto \sin^2 x$. 3. $x \mapsto (1+x)^a$, où $a \in \mathbf{R} \setminus [\mathbf{N} \cup \{-1\}]$.

1. On sait que pour $|u| < 1$, $\frac{1}{1+u} = \sum_{n=0}^{+\infty} (-1)^n u^n$. On pose $u = x^2$. Alors :

$$\frac{1}{1+x^2} = \sum_{n=0}^{+\infty} (-1)^n x^{2n} \Rightarrow \arctan x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1} + K.$$

Puis $K = 0$ car $\arctan 0 = 0$. Le rayon de convergence est $R = 1$.

2. On a, en linéarisant : $\sin^2 x = \frac{1 - \cos 2x}{2} = x^2 - \frac{1}{3}x^4 + \dots + (-1)^{n+1} 2^{2n-1} \frac{1}{(2n)!} x^{2n} + \dots$, le rayon de convergence étant $R = +\infty$.

Remarque : Utiliser le développement en série entière de sin puis un produit de Cauchy aurait été ici plus que maladroit !

3. On remarque que $f : x \mapsto (1+x)^a$ est solution de : $(1+x)y' = ay(x)$ avec $y(0) = 1$. D'après le cours de première année, f est la seule solution de cette équation différentielle.

Écrivons $f(x) = \sum_{n=0}^{+\infty} a_n x^n$, de rayon de convergence R que nous calculerons par la suite. On

remplace dans l'équation différentielle $f(x)$ par cette somme et $f'(x)$ par la somme $\sum_{n=1}^{+\infty} n a_n x^{n-1}$.

On obtient le développement en série entière de $(1+x)f'(x) - af(x)$ qui est de la forme $\sum_{n=0}^{+\infty} c_n x^n$.

Or cette somme est nulle et par unicité du développement en série entière, $c_n = 0$ pour tout $n \in \mathbf{N}$. Le développement de $(1+x)f'(x) - af(x) = 0$ donne :

$$c_0 = a_1 - aa_0, \quad c_1 = 2a_2 + a_1(1-a), \quad \dots, \quad c_n = (n+1)a_{n+1} + a_n(n-a)$$

et avec la remarque précédente, on obtient alors :

$$a_1 = aa_0, \quad 2a_2 = (a-1)a_1, \quad \dots, \quad (n+1)a_{n+1} = (a-n)a_n$$

et en multipliant membre à membre ces différentes égalités, on obtient :

$$a_0 a_1 \dots a_n a(a-1) \dots (a-n) = 1 \times 2 \times \dots \times (n+1) a_1 a_2 \dots a_{n+1},$$

ce qui donne en simplifiant par les a_i tous non nuls ce qui se voit par récurrence immédiate :

$$(n+1)! a_{n+1} = a_0 a(a-1) \dots (a-n) \Rightarrow a_{n+1} = \frac{a(a-1) \dots (a-n)}{(n+1)!}$$

car $a_0 = 1$. Enfin, comme $\left| \frac{a_{n+1}}{a_n} \right|$ tend vers 1 quand n tend vers $+\infty$, $R = 1$.

On en déduit finalement : $\forall x \in]-1, 1[, (1+x)^a = 1 + \sum_{n=1}^{+\infty} \frac{a(a-1) \dots (a-(n-1))}{n!} x^n$.

□ Méthode 10.5.— Comment montrer qu'une fonction est de classe \mathcal{C}^k (ou \mathcal{C}^∞) sur un intervalle du type $] -r, r[$ (ou sur \mathbf{R})

Il suffit de prouver que f possède un D.S.E. autour de 0 de la forme $\sum_{n \geq 0} a_n x^n$.

On détermine son rayon de convergence r et on peut en déduire d'après le cours que f est \mathcal{C}^∞ sur $] -r, r[$. (Si $r = +\infty$, $f \in \mathcal{C}^\infty(\mathbf{R})$.)

Chapitre 11

Linear differential equation and differential system

■ Objectifs

■ Les incontournables :

- ▶ Connaître les théorèmes d'existence et d'unicité des solutions (Cauchy linéaire).
- ▶ Connaître la structure des solutions d'une équation différentielle linéaire d'ordre 1 ou d'ordre 2 ou un système linéaire d'ordre 1, c'est-à-dire que ces solutions sont constituées de la somme d'une solution particulière et des solutions quelconques de l'équation ou du système homogène associé.
- ▶ Savoir résoudre une équation différentielle linéaire du premier ordre à coefficients variables, avec la méthode de variation de la constante.
- ▶ Savoir résoudre une équation différentielle linéaire d'ordre 2 à coefficients constants, le second membre étant de la forme $\lambda \cos(\omega t)$, ou $\lambda \sin(\omega t)$ ou λe^{mt} , λ étant une constante.

■ Et plus si affinités ...

- ▶ Savoir résoudre une équation différentielle linéaire du premier ordre à coefficients variables, en tenant compte des intervalles de résolution.
- ▶ Savoir trouver les solutions polynomiales d'une équation différentielle linéaire du second ordre à coefficients variables.
- ▶ Savoir trouver les solutions d'une équation différentielle linéaire du second ordre à coefficients variables quand on connaît une solution de l'équation sans second membre associée.
- ▶ Savoir résoudre une équation différentielle linéaire d'ordre 2 à coefficients constants, le second membre étant de la forme $P(t) \cos(\omega t)$, ou $P(t) \sin(\omega t)$ ou $P(t)e^{mt}$, P étant un polynôme.
- ▶ Savoir faire le lien entre une équation différentielle linéaire du second ordre et un système différentiel.
- ▶ Savoir résoudre un système différentiel dans le cas où A est trigonalisable (avec $n \geq 3$).
- ▶ Savoir résoudre un système différentiel du type $X'(t) = AX(t) + B(t)$, dans le cas où on donne la forme d'une solution particulière $X_p(t)$ du système complet.
- ▶ Savoir résoudre une équation différentielle linéaire homogène à coefficients constants d'ordre au moins 3 en se ramenant à un système linéaire d'ordre 1.

Résumé de cours

On désigne par I un intervalle (non vide) de \mathbf{R} et par \mathbf{K} l'ensemble \mathbf{R} ou \mathbf{C} .

■ Équations différentielles linéaires d'ordre 1

Définition : Si a et f sont deux applications continues de I dans \mathbf{K} , une **équation différentielle linéaire du premier ordre** est une relation entre une fonction y et sa dérivée y' définie par :

$$(E) \quad y'(t) + a(t)y(t) = f(t)$$

Une fonction $y \in \mathcal{C}^1(I)$ à valeurs dans \mathbf{K} qui vérifie (E) est appelée **solution** de (E) . Enfin, on appelle **courbe intégrale** le graphe d'une solution y de (E) sur I .

Notation : on note $\mathcal{S}_E(I)$ l'ensemble des solutions de l'équation (E) .

Définition : on appelle **équation homogène** associée à (E) l'équation différentielle linéaire (E_H) définie par :

$$(E_H) : y'(t) + a(t)y(t) = 0$$

Proposition 11.1.— On a : $\dim \mathcal{S}_{E_H}(I) = 1$.

Plus précisément, l'ensemble des solutions de E_H s'écrit sous la forme :

$$\mathcal{S}_{E_H}(I) = \left\{ I \rightarrow \mathbf{K} : t \mapsto \lambda \exp \left(- \int a(t) dt \right), \lambda \in \mathbf{K} \right\}$$

Proposition 11.2.— Si y_p est une solution particulière de l'équation (E) , alors :

$$\mathcal{S}_E(I) = y_p + \mathcal{S}_{E_H}(I)$$

Pour résoudre (E) , on résout (E_H) et on détermine une solution particulière y_p de (E) . Soit on la cherche directement soit on utilise la **méthode de variation de la constante**. Voir pour développer ces deux pistes, la **méthode 11.2**, la **méthode 11.3** et la **méthode 11.4**.

Théorème-Définition 11.3.— **Théorème de Cauchy linéaire.**

Soit a et f deux applications continues de I dans \mathbf{K} et (E) l'équation différentielle $y' + ay = f$, le problème de Cauchy associé au couple (t_0, y_0) de $I \times \mathbf{K}$ est la recherche des solutions de (E) vérifiant la condition initiale (ou condition de Cauchy) $y(t_0) = y_0$.

Le problème de Cauchy admet une unique solution.

Une seule courbe intégrale définie sur I passe par le point (t_0, y_0) de $I \times \mathbf{K}$.

Mise en œuvre : exercice 01 à exercice 07.

■ Systèmes différentiels linéaires d'ordre 1

On désigne par n un entier naturel non nul. On identifie \mathbf{K}^n à l'espace vectoriel des matrices colonnes, $\mathcal{M}_{n,1}(\mathbf{K})$.

Définition : On appelle *système différentiel linéaire d'ordre 1 à coefficients constants* sur I tout système du type :

$$X'(t) = AX(t) \quad (S)$$

d'inconnue $X : I \rightarrow \mathbf{K}^n$ de classe \mathcal{C}^1 , où $A \in \mathcal{M}_n(\mathbf{K})$.

Existence de solutions, problème de Cauchy

Définition : Une *solution sur l'intervalle I* du système (S) est une fonction $X : I \rightarrow \mathbf{K}^n$ dérivable sur I telle que : $\forall t \in I, X'(t) = AX(t)$.

Remarque : Une solution de (S) sur I est de classe \mathcal{C}^1 sur I .

Proposition 11.4.— Théorème de Cauchy linéaire.—

Pour tout $(t_0, X_0) \in I \times \mathbf{K}^n$, il existe une unique application $X : I \rightarrow \mathbf{K}^n, t \mapsto X(t)$ de classe \mathcal{C}^1 vérifiant le système (S) et la condition initiale $X(t_0) = X_0$.

Notons $\mathcal{S}(I)$ l'ensemble des solutions $X : I \rightarrow \mathbf{K}^n$ du système différentiel $(S) : X'(t) = AX(t)$.

Proposition 11.5.— $\mathcal{S}(I)$ est un sous-espace vectoriel de $\mathcal{C}^1(I, \mathbf{K}^n)$, isomorphe à $\mathcal{M}_{n,1}(\mathbf{K})$ et donc : $\dim \mathcal{S}(I) = n$.

Mise en œuvre : exercice 08 à exercice 11.

■ Équations différentielles linéaires scalaires d'ordre 2

Ici les fonctions a, b et f sont continues sur I et à valeurs dans \mathbf{K}

Définition : On appelle *équation différentielle linéaire d'ordre 2* une équation du type :

$$y''(t) + a(t)y'(t) + b(t)y(t) = f(t) \quad (E)$$

la fonction inconnue étant $y : I \rightarrow \mathbf{K}$.

Remarque : On notera $y''(t) + a(t)y'(t) + b(t)y(t) = f(t)$ l'équation (E) si l'on souhaite préciser le nom de la variable libre. Il est souvent pratique de confondre y et $y(t)$.

Vocabulaire : ► Les fonctions a et b sont appelées *coefficients* de (E) et la fonction f le *second membre* de (E) .

► On dit que l'équation (E) est à coefficients constants si les fonctions a et b sont constantes. On dit qu'elle est *homogène* ou (communément dit) sans *second membre* si $f = 0$.

► On appelle *équation homogène associée* à (E) l'équation différentielle :

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0 \quad (E_H)$$

L'équation (E) s'appelle alors l'équation *complète*.

Proposition 11.6.— Lien entre équation différentielle linéaire d'ordre 2 homogène et système différentiel d'ordre 1. Ici, a et b sont constants.

L'équation : $y''(t) + ay'(t) + by(t) = 0$ est équivalente au système associé :

$$\begin{cases} z(t) = y'(t) \\ z'(t) + az(t) + by(t) = 0 \end{cases}$$

qui s'écrit $X'(t) = AX$ avec $X(t) = \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$, $A(t) = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}$.

Remarque : Attention, on peut généraliser, c'est-à-dire que toute équation différentielle linéaire du second ordre à coefficients non constants et à second membre non nul, peut se mettre sous la forme d'un système différentiel d'ordre 1 du type $X'(t) = A(t)X(t) + B(t)$, mais un tel système différentiel est hors programme en 2TSI.

Structure de l'ensemble des solutions

Définition à connaître : Une *solution de l'équation (E) sur l'intervalle I* est une fonction φ deux fois dérivable sur I telle que pour tout $t \in I$,

$$\varphi''(t) + a(t)\varphi'(t) + b(t)\varphi(t) = f(t).$$

Notation : $\mathcal{C}^2(I)$ est l'ensemble des fonctions dont la dérivée seconde est continue sur I .

Proposition 11.7.— L'ensemble $\mathcal{S}_{E_H}(I)$ des solutions sur I de l'équation homogène (E_H) est un sous-espace vectoriel de $\mathcal{C}^2(I)$.

L'ensemble $\mathcal{S}_E(I)$ des solutions sur I de l'équation (E) est soit vide, soit l'ensemble constitué de la somme d'une solution particulière de (E) et de toutes les solutions de $\mathcal{S}_{E_H}(I)$.

Remarques : ► Si y_p est une solution de (E) sur I , on a donc : $\mathcal{S}_E(I) = y_p + \mathcal{S}_H(I)$.

► On dit que la solution générale sur I de l'équation complète (E) est la somme d'une solution particulière y_p et de la solution générale de l'équation homogène (E_H), associée à (E).

Existence de solutions, problème de Cauchy

Proposition 11.8.— Théorème de Cauchy linéaire.— Soit $t_0 \in I$ et $(\alpha, \beta) \in \mathbf{K}^2$.

Il existe une unique solution au système :
$$\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = f(t) \\ y(t_0) = \alpha \\ y'(t_0) = \beta \end{cases} \quad (E)$$

Proposition 11.9.— L'espace vectoriel $\mathcal{S}_{E_H}(I)$ est de dimension 2.

Remarque : La recherche de la solution générale d'une équation homogène à coefficients constants, bien qu'au programme de première année, est rappelée dans la **méthode 11.5** (si $\mathbf{K} = \mathbf{C}$) et dans la **méthode 11.6** (si $\mathbf{K} = \mathbf{R}$).

Mise en œuvre : exercice 12 à exercice 17.

■ Extension : équations différentielles linéaires d'ordre n homogènes et à coefficients constants

On suppose $n \geq 2$.

Étant donné une fonction y de classe \mathcal{C}^n sur I et a_0, \dots, a_{n-1} des éléments de \mathbf{K} ($\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$), on pose :

$$X(t) = \begin{pmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(n-1)}(t) \end{pmatrix} \text{ et } A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

Alors, on a le résultat suivant :

Théorème 11.10.— Les propositions suivantes sont équivalentes :

- y est solution sur I de :

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \dots + a_0y(t) = 0.$$

- X est solution sur I du système différentiel $X'(t) = AX(t)$.

Comportement asymptotique des solutions

Théorème 11.11.— Soit A une matrice de $\mathcal{M}_n(\mathbf{K})$ diagonalisable et X une solution du système différentiel $X'(t) = AX(t)$. Si toutes les valeurs propres de A ont une partie réelle strictement négative alors :

$$\lim_{n \rightarrow +\infty} X(t) = 0.$$

Méthodes

■ Équations différentielles linéaires d'ordre 1

□ **Méthode 11.1.— Comment résoudre $(E_H) : y'(t) + a(t)y(t) = 0$ sur I**

[1] On vérifie d'abord que $t \mapsto a(t)$ est une application continue sur I ;

On détermine (si possible) une primitive notée G de l'application a sur I ;

Et enfin : $\mathcal{S}_{E_H}(I) = \{I \rightarrow \mathbf{K} : t \mapsto \lambda \exp(-G(t)), \lambda \in \mathbf{K}\}$.

□ **Méthode 11.2.— Comment déterminer une solution particulière de l'équation $(E) : y'(t) + a y(t) = f(t)$ sur I dans le cas où a est constant**

- Si $f(t) = f_1(t) + \dots + f_n(t)$, où chaque fonction f_i est un produit d'un polynôme et d'une exponentielle ou le produit d'un polynôme et d'une fonction sin ou cos, on doit chercher une solution particulière y_{p_1} de l'équation différentielle notée $(E_1) : y'(t) + a y(t) = f_1(t)$ et ainsi de suite jusqu'à une solution particulière y_{p_n} de l'équation différentielle notée $(E_n) : y'(t) + a y(t) = f_n(t)$.

Pour cela, on utilise le **tableau ci-après**. $y_{p_1} + \dots + y_{p_n}$ est une solution particulière de (E) en utilisant le **principe de superposition**.

- Si f n'a pas la forme précédente, on utilise la méthode de variation de la constante (voir **méthode 11.3**).
- On peut aussi chercher une solution développable en série entière comme pour une équation différentielle linéaire d'ordre 2 (voir **méthode 11.9**).

Remarque : Si a n'est pas constant, le tableau n'est plus valable et la méthode de variation de la constante est en général obligatoire (voir **méthode 11.3**).

Forme de $f(t)$	Forme de y_p solution particulière de (E)
P_n	Q_n
$P_n(t)e^{kt}$ avec $k \neq -a$	$Q_n(t)e^{kt}$
$P_n(t)e^{kt}$ avec $k = -a$	$tQ_n(t)e^{kt}$
$P_n(t) \cos(kt)$ ou $P_n(t) \sin(kt)$	$Q_n(t) \cos(kt) + Z_n(t) \sin(kt)$

(Dans tout le tableau, P_n , Q_n et Z_n désignent trois polynômes de degré n et $k \in \mathbf{R}$).

Exemple : Déterminer l'unique solution de $y'(t) + 3y(t) = e^t + te^{-3t} + \cos t + 2$, $y(0) = 0$

On a $\mathcal{S}_{E_H}(\mathbf{R}) = \{t \mapsto Ke^{-3t}, K \in \mathbf{R}\}$. Puis, on pose $f_1(t) = e^t$. On cherche une solution y_1 de $(E_1) : y'(t) + 3y(t) = f_1(t)$ de la forme $y_1(t) = Ae^t$. En identifiant, $A = \frac{1}{4}$. Puis, $f_2(t) = te^{-3t}$.

On cherche une solution y_2 de $(E_2) : y'(t) + 3y(t) = f_2(t)$ de la forme $y_2(t) = t(Bt + C)e^{-3t}$.

En identifiant, $B = \frac{1}{2}$ et $C = 0$. Puis on pose ensuite $f_3(t) = \cos t$.
 On cherche une solution y_3 de $(E_3) : y'(t) + 3y(t) = f_3(t)$ de la forme $y_1(t) = D \cos t + E \sin t$.
 En identifiant, on trouve $D = \frac{3}{10}$ et $E = \frac{1}{10}$. Enfin, on pose $f_4(t) = 2$.
 On cherche une solution y_4 de $(E_4) : y'(t) + 3y(t) = f_4(t)$ de la forme $y_4(t) = L$. Et : $L = \frac{2}{3}$.
 Ainsi, sans la condition initiale, $y(t) = \frac{1}{4}e^t + \frac{t^2}{2}e^{-3t} + \frac{3}{10} \cos t + \frac{1}{10} \sin t + \frac{2}{3} + Ke^{-3t}$.
 La condition $y(0) = 0$ donne $K = -\frac{73}{60}$.

□ Méthode 11.3.— Comment résoudre $(E) : y'(t) + a(t)y(t) = f(t)$ sur I avec la méthode dite de variation de la constante (ici a peut varier)

[1] On résout d'abord l'équation homogène associée (E_H) , pour cela allez voir la **méthode 11.1**. On notera y_h une solution particulière non nulle de l'équation (E_H) que l'on choisit. Cette solution peut s'écrire $y_h(t) = \exp(-G(t))$;

On cherche les solutions y de (E) sous la forme $y : t \mapsto \lambda(t)y_h(t)$ où λ est une fonction de $\mathcal{C}^1(I)$ en identifiant cette forme de y dans (E) . On trouvera une égalité du type $\lambda'(t) = h(t)$, où h est une fonction continue sur I , que l'on intégrera, ce qui donne : $\lambda(t) = H(t) + L$, où $L \in \mathbf{R}$ et H est une primitive de h ;

Et alors en posant $y_p(t) = H(t)y_h(t)$:

$$\mathcal{S}_E(I) = \{I \rightarrow \mathbf{R} : t \mapsto L \exp(-G(t)) + y_p(t), L \in \mathbf{R}\}$$

On reconnaît au passage que $H(t)y_p(t)$ est une solution particulière de (E) .

Exemples : 1. Résoudre $(E) : y'(t) - 2ty(t) = e^t e^{t^2}$ définie sur \mathbf{R} .

Comme l'équation (E_H) a déjà été traitée plus haut, on choisit $y_h : t \mapsto e^{t^2}$ pour solution particulière de l'équation homogène (E_H) associée à l'équation (E) . On cherche donc les solutions de (E) de la forme $y : t \mapsto \lambda(t)y_h(t)$, où λ est une fonction de $\mathcal{C}^1(\mathbf{R})$. y_h est solution de (E_H) , donc : $y'_h(t) - 2ty_h(t) = 0$ et en remplaçant dans (E) , on peut écrire :

$$(\lambda y_h)'(t) - 2t\lambda(t)y_h(t) = \lambda(t)y'_h(t) + \lambda'(t)y_h(t) - 2t\lambda(t)y_h(t) = \lambda'(t)y_h(t).$$

$$\text{De façon successive, } \lambda y_h \in \mathcal{S}_E(\mathbf{R}) \Leftrightarrow \begin{cases} \lambda'(t)y_h(t) = e^t e^{t^2} \\ t \in \mathbf{R} \end{cases} \Leftrightarrow \begin{cases} \lambda'(t) = e^t \\ t \in \mathbf{R} \end{cases} \Leftrightarrow \begin{cases} \lambda(t) = e^t + L \\ t \in \mathbf{R} \end{cases}.$$

D'où les solutions de (E) sont de la forme : $t \mapsto e^t e^{t^2} + L e^{t^2}$, où $L \in \mathbf{R}$.

2. Résoudre $(E) : t(t-1)y'(t) + y(t) = \ln t$ avec $I =]1, +\infty[$. (Posé à ESM Saint-Cyr)

L'équation homogène (E_H) se ramène à : $y'(t) = -\frac{y(t)}{t(t-1)}$. Comme $\frac{-1}{t(t-1)} = \frac{1}{t} - \frac{1}{t-1}$, on obtient : $y(t) = \frac{Ct}{t-1}$, où $C \in \mathbf{R}$. Il reste à chercher les solutions y de (E) sous la forme $y(t) = \frac{\lambda(t)t}{t-1}$.

En remplaçant dans (E) , $\lambda'(t) = \frac{\ln t}{t^2} \Rightarrow \lambda(t) = -\frac{\ln t}{t} + \int \frac{dt}{t^2} = -\frac{\ln t}{t} - \frac{1}{t} + L$, où $L \in \mathbf{R}$.

Ainsi l'ensemble des solutions est : $\mathcal{S}_E(I) = \left\{ t \mapsto \frac{Lt - \ln t - 1}{t-1}, L \in \mathbf{R} \right\}$.

■ Systèmes différentiels linéaires d'ordre 1

Dans cette partie, soit le système différentiel (S) : $X'(t) = AX(t)$.

□ Méthode 11.4.— Comment résoudre (S)

Plusieurs cas se présentent suivant la nature de la matrice A .

- A est une matrice **diagonale**.

L'équation (S) se ramène au système :

$$\begin{cases} x'_1(t) = \lambda_1 x_1(t) + b_1(t) \\ \vdots \\ x'_n(t) = \lambda_n x_n(t) + b_n(t) \end{cases}$$

On résout séparément ces n équations différentielles linéaires.

- A est une matrice **triangulaire** (supérieure pour fixer les idées)

L'équation (S) se ramène alors au système :

$$\begin{cases} x'_1(t) = \lambda_1 x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1(t) \\ x'_2(t) = \lambda_2 x_2 + \cdots + a_{2n}x_n + b_2(t) \\ \vdots \\ x'_{n-1}(t) = \lambda_{n-1}x_{n-1} + a_{n-1,n}x_n + b_{n-1}(t) \\ x'_n(t) = \lambda_n x_n + b_n(t) \end{cases}$$

On résout ce système par remontée, en commençant par la dernière équation qui permet de déterminer x_n . On détermine alors successivement x_{n-1}, \dots, x_2, x_1 .

- **Cas général.** On diagonalise ou à défaut on trigonalise la matrice A (en se plaçant sur le corps \mathbf{C} , on rappelle que la trigonalisation est toujours possible).

On détermine une matrice inversible P (qui est indépendante de t) telle que la matrice $\Delta = P^{-1}AP$ soit diagonale ou triangulaire.

On note $Y(t)$ les coordonnées d'un vecteur exprimées dans la base définie par P . On rappelle la relation de changement de base $X(t) = PY(t)$. L'équation (S) devient :

$$\begin{aligned} [X'(t) = AX(t)] &\Rightarrow [X'(t) = (P\Delta P^{-1})X(t)] \Rightarrow \\ [P^{-1}X'(t) = \Delta(P^{-1}X(t))] &\Rightarrow [Y'(t) = \Delta Y(t)]. \end{aligned}$$

On est ramené à un des cas précédents. On détermine la solution générale Y .

On revient aux inconnues primaires en utilisant $X(t) = PY(t)$.

Exemple : Résoudre le système différentiel (S) suivant : $\begin{cases} x'_1(t) = x_1(t) + 4x_2(t) \\ x'_2(t) = x_1(t) + x_2(t) \end{cases}$

D'après Concours Commun Mines-Telecom

En posant $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$, (S) s'écrit $X'(t) = AX(t)$.

On remarque que le polynôme caractéristique de A est $\chi_A = (X-3)(X+1)$. La matrice A admet deux valeurs propres réelles distinctes, elle est donc diagonalisable sur \mathbf{R} .

Il existe une matrice inversible P telle que $\Delta = P^{-1}AP = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$.

Après calculs, la matrice : $P = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}$ convient.

On pose $X = PY$, l'équation devient : $Y' = \Delta Y$.

D'où le système :
$$\begin{cases} y_1'(t) = 3y_1(t) & (L_1) \\ y_2'(t) = -y_2(t) & (L_2) \end{cases}.$$

Ces deux équations sont indépendantes l'une de l'autre. On les résout successivement.

La solution générale de (L_1) est : $y_1(t) = \lambda_1 e^{3t}$ $\lambda_1 \in \mathbf{R}$.

De même, la solution générale de (L_2) est : $y_2(t) = \lambda_2 e^{-t}$ $\lambda_2 \in \mathbf{R}$.

La solution générale du système (L) est alors : $X = PY = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$

Après avoir déroulé les calculs :
$$\begin{cases} x_1(t) = 2\lambda_1 e^{3t} - 2\lambda_2 e^{-t} \\ x_2(t) = \lambda_1 e^{3t} + \lambda_2 e^{-t} \end{cases}$$

Remarque : On peut ajouter une condition initiale du type $X_0(t_0) = (x_0, z_0)$.

On en déduit alors λ_1 et λ_2 . Par ailleurs, on remarque que le calcul de P^{-1} est **inutile** pour trouver l'espace vectoriel des solutions.

■ Équations différentielles linéaires à coefficients constants du second ordre

Soit $(a, b) \in \mathbf{K}^2$ et $(E) : y''(t) + ay'(t) + by(t) = f(t)$, d'équation homogène associée (E_H) . L'équation (1) : $x^2 + ax + b = 0$ est l'**équation caractéristique** associée à (E_H) .

□ **Méthode 11.5.**— **Comment résoudre (E_H) (avec $(a, b) \in \mathbf{C}^2$) dans \mathbf{C} —.**

► On peut utiliser le cours de première année. On résout (1).

Si $a^2 - 4b \neq 0$, (1) admet deux racines complexes r_1 et r_2 distinctes et :

$$\mathcal{S}_{E_H}(\mathbf{R}) = \left\{ \mathbf{R} \rightarrow \mathbf{C}, t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}, (\lambda, \mu) \in \mathbf{C}^2 \right\}.$$

Si $a^2 - 4b = 0$, (1) admet une racine complexe double r et :

$$\mathcal{S}_{E_H}(\mathbf{R}) = \left\{ \mathbf{R} \rightarrow \mathbf{C}, t \mapsto (\lambda + \mu t) e^{rt}, (\lambda, \mu) \in \mathbf{C}^2 \right\}.$$

► On peut aussi remarquer que (E_H) est associée (voir la **proposition 11.6**) à

$$X'(t) = AX(t), \text{ où } A(t) = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \text{ et } X(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}.$$

Ici $\chi_A(x) = x^2 + ax + b$. On retrouve les deux cas $a^2 - 4b \neq 0$ ou $a^2 - 4b = 0$.

Dans le premier cas, A est semblable à la matrice diagonale $\begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix}$. On

continue en utilisant la **méthode 11.4**. Dans le deuxième cas, A est semblable à la matrice $\begin{pmatrix} r_1 & c \\ 0 & r_2 \end{pmatrix}$, où $c \in \mathbf{C}$. On continue toujours en utilisant la **méthode 11.4**.

On retrouve dans les deux cas la forme de $y(t)$ souhaitée.

Exemple : Déterminer tous les $y : \mathbf{R} \rightarrow \mathbf{C}$ tels que : $y''(t) - (2 + 4i)y'(t) + (-3 + 4i)y(t) = 0$.

1. En utilisant l'équation caractéristique. 2. En utilisant le système différentiel associé.

1. L'équation caractéristique est :

$$x^2 - (2 + 4i)x + (-3 + 4i) = 0.$$

Son discriminant est 0 et on a une racine double : $r_1 = 2i + 1$.

Finalement,

$$\mathcal{S}_E(\mathbf{R}) = \{\mathbf{R} \rightarrow \mathbf{C} : t \mapsto (\lambda + \mu t)e^{2it+t}, (\lambda, \mu) \in \mathbf{C}^2\}.$$

2. Le système différentiel associé a pour matrice : $A = \begin{pmatrix} 0 & 1 \\ 3 - 4i & 2 + 4i \end{pmatrix}$.

On trouve rapidement :

$$\chi_A(x) = \text{Det}(xI_2 - A) = x^2 - (2 + 4i)x + (-3 + 4i) = (x - 2i - 1)^2.$$

Le seul sous-espace propre est $E_{1+2i}(A) = \text{Vect}(\vec{u})$, où $\vec{u}(1, 1 + 2i)$.

Posons par exemple : $P = \begin{pmatrix} 1 & 0 \\ 1 + 2i & 1 \end{pmatrix}$. Alors :

$$\Delta = P^{-1}AP = \begin{pmatrix} 1 + 2i & 1 \\ 0 & 1 + 2i \end{pmatrix}.$$

Si l'on pose $X(t) = \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$ et $Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$, $X'(t) = AX(t) \Leftrightarrow Y'(t) = \Delta Y(t)$.

L'équation $y_2'(t) = (1 + 2i)y_2(t)$ donne : $y_2(t) = \alpha e^{(2i+1)t}$, avec $\alpha \in \mathbf{C}$.

Il reste l'équation : $y_1'(t) = (1 + 2i)y_1(t) + y_2(t) = (1 + 2i)y_1(t) + \alpha e^{(2i+1)t}$. Un petit coup de la **méthode 11.2** et $y_1(t) = (\alpha t + \beta)e^{(2i+1)t}$, où $\beta \in \mathbf{C}$. Il reste à utiliser $X(t) = PY(t)$, pour en déduire $y(t)$. On peut remarquer qu'ici, ce choix de la seconde colonne de P donne : $y(t) = y_1(t)$, ce qui facilite les calculs. On aboutit ainsi encore à une combinaison linéaire de $e^{(2i+1)t}$ et $te^{(2i+1)t}$. Encore une fois, $\mathcal{S}_E(\mathbf{R}) = \{\mathbf{R} \rightarrow \mathbf{C} : t \mapsto (\lambda + \mu t)e^{2it+t}, (\lambda, \mu) \in \mathbf{C}^2\}$.

□ **Méthode 11.6.— Comment résoudre (E_H) (avec $(a, b) \in \mathbf{R}^2$) dans \mathbf{R} —.**

■ On peut se ramener à la **méthode 11.5** et utiliser la formule de Moivre dans le cas où le discriminant de (1) est négatif pour ne conserver que des solutions réelles.

■ On peut aussi utiliser directement :

► si $a^2 - 4b < 0$, (1) admet deux racines $r = \alpha \pm i\beta$ distinctes et conjuguées et :

$$\mathcal{S}_{E_H}(\mathbf{R}) = \{\mathbf{R} \rightarrow \mathbf{R}, t \mapsto e^{\alpha t}(\lambda \cos(\beta t) + \mu \sin(\beta t)), (\lambda, \mu) \in \mathbf{R}^2\}.$$

► si $a^2 - 4b = 0$, (1) admet une racine réelle double r et :

$$\mathcal{S}_{E_H}(\mathbf{R}) = \{\mathbf{R} \rightarrow \mathbf{R}, t \mapsto (\lambda + \mu t)e^{rt}, (\lambda, \mu) \in \mathbf{R}^2\}.$$

► si $a^2 - 4b > 0$, (1) admet deux racines réelles r_1 et r_2 distinctes et :

$$\mathcal{S}_{E_H}(\mathbf{R}) = \{\mathbf{R} \rightarrow \mathbf{R}, t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}, (\lambda, \mu) \in \mathbf{R}^2\}.$$

❑ **Méthode 11.7.— Comment résoudre l'équation complète (E)**

- On résout l'équation homogène associée (E_H) en utilisant la **méthode 11.6**.
- On cherche une solution particulière y_p de (E), soit fournie par l'énoncé, soit "devinée" par l'étudiant. Pour ce dernier cas, il faut que $f(t)$ soit de la forme $f(t) = f_1(t) + \dots + f_n(t)$, où chaque fonction f_i est un polynôme, une exponentielle ou des fonctions trigonométriques usuelles. On doit chercher une solution particulière y_{p_1} de l'équation différentielle (E_1) : $y''(t) + a y'(t) + b y(t) = f_1(t)$ et ainsi de suite jusqu'à une solution particulière y_{p_n} de l'équation différentielle notée (E_n) : $y''(t) + a y'(t) + b y(t) = f_n(t)$. Pour cela, on utilise le **tableau ci-après**. Enfin, on utilise la remarque que $y_{p_1} + \dots + y_{p_n}$ est une solution particulière de (E) en utilisant le **principe de superposition**.

$$\mathcal{S}_E(I) = \mathcal{S}_{E_H}(I) + y_{p_1}(t) + \dots + y_{p_n}(t).$$

Parfois, si $f(t)$ est moins classique, on peut chercher une solution particulière développable en série entière. On est en plein dans la **méthode 11.9**.

- On peut aussi utiliser la **proposition 11.6** et la **méthode 11.4**, dans le cas où $B(t)$ est non nul. Mais c'est généralement plus long ! On résout alors l'équation complète sans passer par celle de l'équation homogène auparavant.

Ici P_n et Q_n sont des **polynômes réels de degré n** , $(K, M, N, c, d) \in \mathbf{R}^5$, $k \in \mathbf{C}$.

Forme du second membre $f(t)$	Forme de y_1 solution particulière de (E)
$P_n(t)$ avec $b \neq 0$	$Q_n(t)$
$P_n(t)$ avec $b = 0, a \neq 0$	$t Q_n(t)$
$P_n(t)$ avec $a = b = 0$	$t^2 Q_n(t)$
$K e^{kt}$ avec k non racine de (1)	$M e^{kt}$
$K e^{kt}$ avec k racine simple de (1)	$M t e^{kt}$
$K e^{kt}$ avec k racine double de (1)	$M t^2 e^{kt}$
$K \cos(dt)$ avec id non racine de (1)	$M \cos(dt) + N \sin(dt)$
$K \sin(dt)$ avec id non racine de (1)	$M \cos(dt) + N \sin(dt)$
$K \cos(dt)$ avec id racine de (1)	$t M \cos(dt) + t N \sin(dt)$
$K \sin(dt)$ avec id racine de (1)	$t M \cos(dt) + t N \sin(dt)$

Exemples : 1. Résoudre l'équation (E) : $y''(t) + 2y'(t) + y(t) = t + 1 + e^t + e^{-t}$ de conditions initiales $y(0) = y'(0) = 0$ avec le tableau.

L'équation caractéristique $x^2 + 2x + 1$ a pour racine double -1 .

Donc : $\mathcal{S}_{E_H}(\mathbf{R}) = \{t \mapsto K_1 t e^{-t} + K_2 e^{-t}, (K_1, K_2) \in \mathbf{R}^2\}$. Attaquons l'équation complète.

On pose : $f_1(t) = t + 1$, $f_2(t) = e^t$, $f_3(t) = e^{-t}$.

On cherche une solution particulière y_{p_i} de chaque (E_i) : $y''(t) + 2y'(t) + y(t) = f_i(t)$.

$y_{p_1}(t)$ est de la forme $a_1 t + a_2$, on trouve $a_1 = 1$ et $a_2 = -1$. Puis $y_{p_2}(t)$ est de la forme $b e^t$.

On trouve $b = \frac{1}{4}$. Puis $y_{p_3}(t)$ est de la forme $t^2 c_1 e^{-t}$ car -1 est racine double. On trouve $c_1 = \frac{1}{2}$.

Il reste à exploiter les conditions initiales qui donnent $K_1 = -\frac{1}{2}$ et $K_2 = \frac{3}{4}$.

Ainsi : $\mathcal{S}_E(\mathbf{R}) = \left\{ t \mapsto \left(\frac{3}{4} - \frac{t}{2} + \frac{t^2}{2} \right) e^{-t} + \frac{1}{4} e^t + t - 1 \right\}.$

2. Résoudre (avec le tableau) $y''(t) + \omega y(t) = 2 \cos t$, où $\omega \in \mathbf{R}^*$ avec $I = \mathbf{R}$.

L'équation homogène (E_H) : $y''(t) + \omega y(t) = 0$ a pour équation caractéristique (1) : $x^2 + \omega = 0$.

► Si $\omega > 0$, (1) donne $\{i\sqrt{\omega}, -i\sqrt{\omega}\}$. Et : $\mathcal{S}_{E_H}(\mathbf{R}) = \{t \mapsto C_1 \sin(t\sqrt{\omega}) + C_2 \cos(t\sqrt{\omega})\}$.

Si $\omega \neq 1$, i n'est pas solution de (1) et on cherche une solution particulière de la forme : $y_1(t) = a \cos t + b \sin t$. On dérive deux fois et après identification, on obtient $a = \frac{2}{\omega - 1}$ et $b = 0$.

Finalement : $\mathcal{S}_E(\mathbf{R}) = \left\{ t \mapsto C_1 \sin(t\sqrt{\omega}) + C_2 \cos(t\sqrt{\omega}) + \frac{2}{\omega - 1} \cos t \right\}.$

Si $\omega = 1$, i est solution de (1) et on cherche une solution particulière de la forme : $y_1(t) = at \cos t + bt \sin t$. On trouve $a = 0$ et $b = 1$. Finalement : $\mathcal{S}_E(\mathbf{R}) = \{t \mapsto C_1 \sin t + C_2 \cos t + t \sin t\}.$

► Si $\omega < 0$, (1) a pour solution $\{-\sqrt{-\omega}, \sqrt{-\omega}\}$. Déjà, $\mathcal{S}_{E_H}(\mathbf{R}) = \{t \mapsto C_1 e^{-\sqrt{-\omega}t} + C_2 e^{\sqrt{-\omega}t}\}.$

Pour la suite, la solution particulière se trouve comme dans le cas $\omega \neq 1 > 0$.

On trouve : $\mathcal{S}_E(\mathbf{R}) = \left\{ t \mapsto C_1 e^{-\sqrt{-\omega}t} + C_2 e^{\sqrt{-\omega}t} + \frac{2}{\omega - 1} \cos t \right\}.$

■ Équations différentielles linéaires à coefficients non constants du second ordre

On comprend que les méthodes qui vont suivre restent valables pour le cas où les coefficients sont constants.

Soit a, b et c trois fonctions continues sur I , la fonction a ne s'annulant pas sur I .

On note : $(E) : a(t)y''(t) + b(t)y'(t) + c(t)y(t) = f(t)$, d'équation homogène associée (E_H).

L'ensemble des solutions de (E) est noté $\mathcal{S}_E(I)$ et celles de (E_H) est noté $\mathcal{S}_{E_H}(I)$.

□ Méthode 11.8.— Comment déterminer $\mathcal{S}_E(I)$ en effectuant un changement de variable dans (E)

On suppose que l'on dispose d'une fonction φ bijective définie d'un intervalle J sur l'intervalle I . On suppose que φ est de classe \mathcal{C}^2 sur J et φ^{-1} est de classe \mathcal{C}^2 sur I . On introduit alors une nouvelle fonction z définie sur J par : $z(x) = y(\varphi(x))$.

On dérive deux fois la fonction z et on écrit l'équation (E) sous la forme :

$$a(\varphi(x))y''(\varphi(x)) + b(\varphi(x))y'(\varphi(x)) + c(\varphi(x))y(\varphi(x)) = f(\varphi(x))$$

On essaye alors d'obtenir une nouvelle équation en z , (généralement à coefficients constants pour que la méthode ait une efficacité). On résout cette équation, on obtient donc l'expression de $z(x)$ puis on détermine la solution générale de (E) en utilisant φ^{-1} .

Exemple : Résoudre sur $] -1, 1[$, $(E) : (1 - t^2)y''(t) - ty'(t) + y(t) = 0$ en posant $t = \cos x$.

$$z(x) = y(\cos x) \Rightarrow z'(x) = -(\sin x)y'(\cos x) \Rightarrow z''(x) = -(\cos x)y'(\cos x) + (1 - \cos^2 x)y''(\cos x).$$

$$y \in \mathcal{S}_{(E)}(] -1, 1[) \Leftrightarrow \forall x \in]0, \pi[, \quad \underbrace{(1 - \cos^2 x)y''(\cos x) - \cos(x)y'(\cos x)}_{z''(x)} + \underbrace{y(\cos x)}_{z(x)} = 0.$$

z est alors solution sur $]0, \pi[$ de l'équation différentielle : $z'' + z = 0$ dont la solution générale est $z : x \mapsto \lambda \cos(x) + \mu \sin(x)$. Et : $\mathcal{S}_E(] -1, 1[) = \{t \mapsto \lambda t + \mu \sqrt{1 - t^2}, (\lambda, \mu) \in \mathbf{R}^2\}.$

□ **Méthode 11.9.—** Comment déterminer une solution $y(t) = \sum_{n=0}^{+\infty} a_n t^n$ de (E)

développable en série entière (à ne lire qu'après avoir fait le chapitre 10)

Traisons uniquement le cas d'une équation du second ordre. On suppose que $0 \in I$ et que f est développable en série entière (de rayon de convergence $R > 0$) et que les fonctions a , b et c sont polynomiales en t . La méthode peut s'étendre au cas où a s'annule en 0.

Alors : $y'(t) = \sum_{n=1}^{+\infty} n a_n t^{n-1}$ et $y''(t) = \sum_{n=2}^{+\infty} n(n-1) a_n t^{n-2}$

- 1. On injecte ces deux égalités dans l'équation (E).
- 2. En faisant éventuellement des changements d'indice, on écrit le premier membre de (E) sous la forme regroupée $\sum_{n=0}^{+\infty} b_n t^n$.
- 3. On détermine le développement en série entière de $f(t)$.
- 4. On utilise l'unicité du développement en série entière pour déterminer une relation de récurrence entre les termes de la suite (a_n) .
- 5. On détermine l'expression de a_n .
- 6. On calcule enfin le rayon de convergence de la série entière $\sum a_n t^n$.

Exemple : Appliquer cette méthode pour déterminer une solution particulière de l'équation différentielle (E) : $y'' + ty' + y = 0$.

On cherche une solution de (E) sous la forme $y(t) = \sum_{n=0}^{+\infty} a_n t^n$.

On dérive deux fois la fonction y et on injecte le tout dans l'équation (E), il vient :

$$a_0 + 2a_2 + \sum_{n=1}^{+\infty} (a_n + n a_n + (n+2)(n+1)a_{n+2}) t^n = 0.$$

Par unicité du développement en série entière, $\forall n \geq 0, a_n + (n+2)a_{n+2} = 0$.

La relation précédente est une relation entre deux termes de même parité : on peut donc exprimer tous les termes d'indices pairs en fonction de a_0 et les termes d'indices impairs en fonction de a_1 . Soit $p \geq 0$, on a :

$$\begin{array}{ll} a_{2p} = \frac{-1}{2p} a_{2p-2} & a_{2p+1} = \frac{-1}{2p+1} a_{2p-1} \\ a_{2p-2} = \frac{-1}{2(p-1)} a_{2p-4} & a_{2p-1} = \frac{-1}{2p-1} a_{2p-3} \\ \vdots & \vdots \\ a_2 = \frac{-1}{2} a_0 & a_3 = \frac{-1}{3} a_1 \end{array}$$

On en déduit (par multiplication terme à terme) que :

$$\forall p \geq 0 \quad a_{2p} = \frac{(-1)^p}{2^p p!} a_0 \quad \text{et} \quad a_{2p+1} = \frac{(-1)^p 2^p p!}{(2p+1)!} a_1.$$

On cherche une solution particulière de (E) , on est libre du choix de (a_0, a_1) , la série entière correspondante doit avoir cependant un rayon de convergence non nul.

► On choisit alors $a_0 = 1$ et $a_1 = 0$. La série obtenue est alors : $\sum_{p \geq 0} \frac{(-1)^p}{2^p p!} t^{2p}$.

On reconnaît la définition de l'exponentielle :

$$\forall t \in \mathbf{R}, \quad \sum_{p=0}^{+\infty} \frac{(-1)^p}{2^p p!} t^{2p} = e^{-t^2/2}.$$

En conclusion, la fonction $t \mapsto e^{-t^2/2}$ est une solution particulière de (E) sur \mathbf{R} .

► On choisit $a_0 = 0$ et $a_1 = 1$. La série obtenue est alors : $\sum_{p \geq 0} \frac{(-1)^p 2^p p!}{(2p+1)!} t^{2p+1}$.

Déterminons son rayon de convergence. Pour tout $t \neq 0$, on a :

$$\left| \frac{a_{2p+1} t^{2p+1}}{a_{2p-1} t^{2p-1}} \right| = \frac{t^2}{2p+1} \xrightarrow{p \rightarrow +\infty} 0.$$

D'après la règle de D'Alembert la série de terme général $a_{2p+1} t^{2p+1}$ est convergente pour tout $t \in \mathbf{R}$. Le rayon de convergence de la série entière $\sum \frac{(-1)^p 2^p p!}{(2p+1)!} t^{2p+1}$ est donc infini.

La fonction $\phi : t \mapsto \sum_{p=0}^{+\infty} \frac{(-1)^p 2^p p!}{(2p+1)!} t^{2p+1}$ est une solution de (E) sur \mathbf{R} .

Remarque : $t \mapsto e^{-\frac{t^2}{2}}$ et ϕ sont deux solutions de (E) sur \mathbf{R} linéairement indépendantes. En effet, l'une est paire l'autre est impaire. On sait que $\mathcal{S}_E(\mathbf{R})$ est un \mathbf{R} -espace vectoriel de dimension 2 (**théorème 11.9**). Ainsi :

$$\mathcal{S}_E(\mathbf{R}) = \left\{ t \mapsto \lambda e^{-t^2/2} + \mu \phi(t), (\lambda, \mu) \in \mathbf{R}^2 \right\}.$$

Dernière méthode pour ceux qui en veulent plus...

Remarque : Dans le programme officiel, il est question de savoir « se débrouiller » sur des équations différentielles linéaires du type $(E) : y''(t) + a(t)y'(t) + b(t)y(t) = f(t)$, où a et b sont continues sur I . On sait que l'ensemble des solutions $\mathcal{S}_E(I)$ est alors un espace vectoriel de dimension 2. Or, ces équations peuvent provenir d'équations plus générales du type

$$(E) : a(t)y''(t) + b(t)y'(t) + c(t)y(t) = f(t).$$

Bien entendu, l'idée est de diviser par $a(t)$ et de se ramener au cas précédent (dans le domaine I , où a ne s'annule pas). Un exercice classique au programme est de chercher une solution développable en série entière (même si $0 \notin I$). On pourrait donc dans certains sujets de concours avoir à étendre les solutions de (E) dans le cas où $a(t)$ de (E) s'annule. On vous invite alors à parcourir les exemples qui suivent.

❏ Méthode 11.10.— Comment chercher d'éventuelles solutions d'une équation différentielle linéaire d'ordre un ou deux sur un domaine de \mathbf{R} où les conditions d'application des résultats de cours ne sont pas vérifiées

Attention le prolongement doit être dérivable (respectivement deux fois dérivable) si (E) est d'ordre 1 (respectivement d'ordre 2). Ici l'équation se présente ainsi :

$$(E) : a(t)y'(t) + b(t)y(t) = f(t) \text{ ou } (E) : a(t)y''(t) + b(t)y'(t) + c(t)y(t) = f(t).$$

On commence à écrire les intervalles I_1, \dots, I_k sur lesquels la fonction a ne s'annule pas. On sait alors, d'après le cours, que sur chacun de ces intervalles, l'ensemble des solutions de l'équation homogène associée est un espace vectoriel de dimension 1 ou 2 selon l'ordre de (E) .

- On peut commencer par résoudre (E) sur chacun des intervalles I_1, \dots, I_k .
- On peut par analyse-synthèse, étudier l'existence d'éventuelles solutions sur un ensemble plus vaste.
- On peut étudier l'existence d'une solution en $t_0 \notin I_1 \cup \dots \cup I_k$ en effectuant un *D.L* au voisinage de t_0 des solutions trouvées sur les deux intervalles "approchant" t_0 . L'idée est de montrer que ce prolongement est dérivable en t_0 (si (E) est d'ordre 1) et deux fois dérivable en t_0 (si (E) est d'ordre 2).
- On peut aussi dans le cas où $0 \notin I_1 \cup \dots \cup I_k$, chercher une solution développable en série entière autour de 0 pour ajouter 0 au domaine de validité (Voir encore la **méthode 11.9**).
- On peut utiliser une base de $\mathcal{S}_E(I_j)$ et de $\mathcal{S}_E(I_k)$ pour "raccorder" des solutions entre les deux intervalles consécutifs I_j et I_k .

Exemples : 1. Soit l'équation différentielle $(E) : t^2 y'(t) + y(t) = 0$.

Résoudre (E) sur \mathbf{R}^* , sur \mathbf{R}_-^* et sur \mathbf{R}_+^* puis déterminer les solutions sur \mathbf{R} .

L'équation différentielle est équivalente à : $y'(t) + \frac{1}{t^2}y(t) = 0$, ce qui est une équation différentielle du type $y'(t) + a(t)y(t) = 0$ avec a application continue sur \mathbf{R}^* , donc en appliquant les résultats de cours, on obtient sur \mathbf{R}_-^* et sur \mathbf{R}_+^* les solutions suivantes :

$$\mathcal{S}_E(\mathbf{R}_-^*) = \{\mathbf{R}_-^* \rightarrow \mathbf{R} : t \mapsto \lambda e^{\frac{1}{t}}, \lambda \in \mathbf{R}\}; \mathcal{S}_E(\mathbf{R}_+^*) = \{\mathbf{R}_+^* \rightarrow \mathbf{R} : t \mapsto \mu e^{\frac{1}{t}}, \mu \in \mathbf{R}\}$$

Solutions sur \mathbf{R} : on procède par analyse et synthèse.

Analyse : on suppose que l'équation (E) admet des solutions sur \mathbf{R} et soit f l'une d'entre elles. f est solution de (E) sur \mathbf{R} , donc $f|_{\mathbf{R}_-^*}$ appartient à $\mathcal{S}_E(\mathbf{R}_-^*)$ et il existe λ réel tel que pour tout x réel strictement négatif, $f(x) = \lambda e^{\frac{1}{x}}$. De même, il existe μ réel tel que pour tout x réel strictement positif, $f(x) = \mu e^{\frac{1}{x}}$. De plus f est continue en 0 donc $\lim_{x \rightarrow 0} f(x)$ existe et vaut $f(0)$, ce qui impose $\mu = 0$ et $f(0) = 0$. On en déduit donc que si f est solution de (E) sur \mathbf{R} , alors il existe λ tel que :

$$f : t \mapsto f(t) = \begin{cases} \lambda e^{\frac{1}{t}} & \text{si } t < 0 \\ 0 & \text{si } t \geq 0 \end{cases}$$

Synthèse : on vérifie aisément que les fonctions ainsi obtenues sont dérivables et solutions de

l'équation (E) sur \mathbf{R} . Conclusion :

$$\mathcal{S}_E(\mathbf{R}) = \left[f : \mathbf{R} \rightarrow \mathbf{R} : t \mapsto f(t) = \begin{cases} \lambda e^{\frac{1}{t}} & \text{si } t < 0 \\ 0 & \text{si } t \text{ est positif ou nul} \end{cases}, \lambda \in \mathbf{R} \right].$$

2. Résoudre (E) : $t(t-1)y'(t) + y(t) = \ln t$ sur $]0, 1[$ et sur $]1, +\infty[$ puis sur \mathbf{R}_+^* . (ESM Saint-Cyr)

On reconnaît le dernier exemple de la **méthode 11.3**.

On sait que les solutions sont $y : t \mapsto \begin{cases} \frac{C_1 t - \ln t - 1}{t-1} & \text{si } t \in]0, 1[\\ \frac{C_2 t - \ln t - 1}{t-1} & \text{si } t \in]1, +\infty[\end{cases}$, où C_1 et C_2 sont

deux constantes réelles. Pour déterminer une solution qui « contienne » la valeur $t = 1$, il reste à effectuer un développement limité autour de $t = 1$.

On pose $t = 1 + h$ et on applique : $\ln(1+h) = h - \frac{h^2}{2} + o(h^2)$.

Si $h > 0$, $y(1+h) = \frac{C_2 - 1}{h} + C_2 - 1 + \frac{h}{2} + o(h)$. Cette quantité a une limite finie quand h tend vers 0^+ si, et seulement si, $C_2 = 1$. Alors $y(1) = 0$ et y est alors dérivable à droite en 1 avec $y'_d(1) = \frac{1}{2}$. De même, en supposant $h < 0$ et en faisant tendre h vers 0^- , on aboutit à : $C_1 = 1$ et y est dérivable à gauche en 1 avec $y'_g(1) = \frac{1}{2}$. Conclusion :

$$\mathcal{S}_E(\mathbf{R}_+^*) = \left[t \mapsto \begin{cases} 1 - \frac{\ln t}{t-1} & \text{si } t \neq 1 \\ 0 & \text{si } t = 1 \end{cases} \right].$$

3. Soit (E) : $t^2 y''(t) + t y'(t) - (t^2 + t + 1)y(t) = 0$.

a. Déterminer les solutions de (E) développables en série entière autour de 0.

Que dire du rayon de convergence de cette série ? Déterminer explicitement ces solutions.

b. Chercher une solution particulière de la forme $t \mapsto t^r e^{\alpha t}$, où $r \in \mathbf{Z}$ et $\alpha \in \mathbf{R}$

c. Déterminer l'ensemble des solutions de (E) sur \mathbf{R} . (Issu de Mines-Ponts, filière PSI)

a. Partons de $y(t) = \sum_{n=0}^{+\infty} a_n t^n$ et supposons que le rayon de convergence de cette série soit strictement positif. Alors :

$$-(t^2 + t + 1)y(t) = -\sum_{k=0}^{+\infty} a_k t^k - \sum_{k=1}^{+\infty} a_{k-1} t^k - \sum_{k=2}^{+\infty} a_{k-2} t^k.$$

Et, de même :

$$t y'(t) = \sum_{k=1}^{+\infty} k a_k t^k \text{ et } t^2 y''(t) = \sum_{k=2}^{+\infty} k(k-1) a_k t^k.$$

On remplace dans (E) et on trouve $a_0 = 0$ et :

$$\forall k \geq 2, a_{k-2} + a_{k-1} - (k-1)(k+1)a_k = 0.$$

On calcule maintenant les premiers termes : $a_2 = -\frac{1}{3}a_1$, $a_3 = \frac{1}{3 \times 4}a_1$, $a_4 = -\frac{1}{3 \times 4 \times 5}a_1$.

On prouve alors par récurrence (le faire) : $\forall n \geq 1$, $a_n = \frac{2 \times (-1)^{n+1}}{(n+1)!}a_1$.

En conclusion, les fonctions $y_\mu : t \mapsto \sum_{k=1}^{+\infty} \frac{\mu(-1)^{k+1}}{(k+1)!}t^k$, où $\mu \in \mathbf{R}$ sont les solutions de (E) développables en série entière autour de 0.

On peut prouver rapidement que ces développements en série entière ont un rayon de convergence infini. C'est très semblable à l'exemple développé après la **méthode 11.9**.

Calculons maintenant y_1 (la solution avec $\mu = 1$).

$$\forall t \neq 0, y_1(t) = \frac{1}{t} \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{(k+1)!} t^{k+1} = \frac{1}{t} \sum_{k=2}^{+\infty} \frac{(-1)^k}{k!} t^k = \frac{1}{t} (e^{-t} - 1 + t).$$

b. Posons $y : t \mapsto t^r e^{\alpha t}$ (avec $t \neq 0$). Alors :

$$y'(t) = (rt^{r-1} + \alpha t^r) e^{\alpha t} \text{ et } y''(t) = (r(r-1)t^{r-2} + 2\alpha r t^{r-1} + \alpha^2 t^r) e^{\alpha t}.$$

(E) devient :

$$(\alpha^2 - 1)t^{r+2} + (2\alpha r + \alpha - 1)t^{r+1} + (r^2 - 1)t^r = 0.$$

Comme ceci est valable quel que soit $t \neq 0$, $\alpha^2 = 1$, $\alpha(2r+1) = 0$, $r^2 = 1$.

L'unique possibilité est $\alpha = r = -1$. On obtient une solution $z_1 : t \mapsto \frac{1}{t}e^{-t}$ sur $I = \mathbf{R}_-^* \cup \mathbf{R}_+^*$.

c. Sur \mathbf{R}_+^* ou de même sur \mathbf{R}_-^* , on a deux solutions non proportionnelles (l'une est continue en 0 et l'autre non).

Prenons $y_1 : t \mapsto \frac{1}{t}(e^{-t} - 1 + t)$ et $z_1 : t \mapsto \frac{1}{t}e^{-t}$.

D'après le cours, on sait que $S_E(\mathbf{R}_+^*)$ et $S_E(\mathbf{R}_-^*)$ sont des espaces vectoriels de dimension 2. Une base de chacun d'eux est $\{y_1, z_1\}$. Les solutions sur \mathbf{R}_+^* sont de la forme $Ay_1 + Bz_1$ et les solutions sur \mathbf{R}_-^* sont de la forme $Cy_1 + Dz_1$. Pour trouver les solutions définies sur \mathbf{R} , on étudie le raccordement de ces combinaisons linéaires quand t tend vers 0. On obtient : $B = D = 0$ et $A = C$. Il reste les fonctions Ay_1 , c'est-à-dire les fonctions y_μ , où $\mu \in \mathbf{R}$. Les seules solutions sur \mathbf{R} sont finalement les solutions développables en série entière autour de 0 (ce qui est logique!)

Énoncé des exercices

■ Équations différentielles linéaires d'ordre 1

Exercice 11.1 :

Exercice 11.2 :

Exercice 11.3 :

Exercice 11.4 :

Exercice 11.5 :

■ Équations différentielles non linéaires d'ordre 1

Exercice 11.6 :

Exercice 11.7 :

■ Systèmes différentiels linéaires d'ordre 1

Exercice 11.8 :

Exercice 11.9 :

Exercice 11.10 :

Exercice 11.11 :

Exercice 11.12 :

■ Équations différentielles linéaires d'ordre 2

Exercice 11.13 :

Exercice 11.14 :

Exercice 11.15 :

Exercice 11.16 :

Exercice 11.17 :

Indications

- ___ Ex. 11.1 _____
- ___ Ex. 11.2 _____
- ___ Ex. 11.3 _____
- ___ Ex. 11.4 _____
- ___ Ex. 11.5 _____
- ___ Ex. 11.6 _____
- ___ Ex. 11.7 _____
- ___ Ex. 11.8 _____
- ___ Ex. 11.9 _____
- ___ Ex. 11.10 _____
- ___ Ex. 11.11 _____
- ___ Ex. 11.12 _____
- ___ Ex. 11.13 _____
- ___ Ex. 11.14 _____
- ___ Ex. 11.15 _____
- ___ Ex. 11.16 _____
- ___ Ex. 11.17 _____

Corrigé des exercices

____ **Exercice 11.1** _____
ZUM MACHEN

____ **Exercice 11.2** _____
ZUM MACHEN

____ **Exercice 11.3** _____
ZUM MACHEN

____ **Exercice 11.4** _____
ZUM MACHEN

____ **Exercice 11.5** _____
ZUM MACHEN

____ **Exercice 11.6** _____
ZUM MACHEN

____ **Exercice 11.7** _____
ZUM MACHEN

____ **Exercice 11.8** _____
ZUM MACHEN

____ **Exercice 11.9** _____
ZUM MACHEN

____ **Exercice 11.10** _____
ZUM MACHEN

____ **Exercice 11.11** _____
ZUM MACHEN

____ **Exercice 11.12** _____
ZUM MACHEN

____ **Exercice 11.13** _____
ZUM MACHEN

____ **Exercice 11.14** _____
ZUM MACHEN

____ **Exercice 11.15** _____
ZUM MACHEN

____ **Exercice 11.16** _____
ZUM MACHEN

____ **Exercice 11.17** _____
ZUM MACHEN

Chapitre 12

Pre-Hilbert and euclidean spaces

■ Objectifs

■ Les incontournables :

- ▶ Savoir manipuler les identités remarquables sur les normes (développement de $\|\vec{u} + \vec{v}\|^2$, identité de polarisation).
- ▶ Connaître les applications géométriques du produit scalaire dans le cas usuel du plan ou de l'espace. En particulier, savoir en déduire des équations de plans ou (et) de droites.
- ▶ Faire le lien entre un produit scalaire et sa norme associée.
- ▶ Interpréter le cas d'égalité dans l'inégalité triangulaire.
- ▶ Savoir reconnaître une situation où il faut mettre en œuvre l'inégalité (ou le cas d'égalité) de Cauchy-Schwarz.
- ▶ Savoir
 - ▷ orthogonaliser une base avec le procédé de Gram-Schmidt ;
 - ▷ projeter un vecteur sur un sous-espace de dimension finie ;
 - ▷ déterminer et interpréter la distance d'un vecteur à un espace.

■ Et plus si affinités ...

- ▶ Connaître les règles de calculs sur l'orthogonal d'un espace.
- ▶ Savoir interpréter un problème de minimisation en introduisant un produit scalaire.

Résumé de cours

On considère ici un espace vectoriel E sur \mathbf{R} de dimension finie ou non.

■ Définition d'un espace préhilbertien réel ou d'un espace euclidien

Définition : On appelle **produit scalaire** sur E toute application :

$$\begin{aligned} \varphi : E \times E &\rightarrow \mathbf{R} \\ (\vec{x}, \vec{y}) &\mapsto \langle \vec{x}, \vec{y} \rangle \end{aligned}$$

possédant les propriétés suivantes :

► **linéarité à droite :** pour tous réels β_1, β_2 et pour tous vecteurs \vec{x}, \vec{u}_1 et \vec{u}_2 :

$$\langle \vec{x}, \beta_1 \vec{u}_1 + \beta_2 \vec{u}_2 \rangle = \beta_1 \langle \vec{x}, \vec{u}_1 \rangle + \beta_2 \langle \vec{x}, \vec{u}_2 \rangle ;$$

- **symétrie :** $\forall (\vec{x}, \vec{y}) \in E^2, \langle \vec{y}, \vec{x} \rangle = \langle \vec{x}, \vec{y} \rangle$;
- **positive :** pour tout $\vec{x} \in E, \langle \vec{x}, \vec{x} \rangle \geq 0$;
- **définie :** pour tout $\vec{x} \in E, \langle \vec{x}, \vec{x} \rangle = 0 \Rightarrow \vec{x} = 0$.

Remarque : Grâce à la symétrie, on a automatiquement la **linéarité à gauche** (qui peut être choisie à la place de la linéarité à droite dans la définition).

$$\forall (\alpha_1, \alpha_2) \in \mathbf{R}^2 \text{ et } \forall (\vec{u}_1, \vec{u}_2, \vec{y}) \in E^3 : \langle \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2, \vec{y} \rangle = \alpha_1 \langle \vec{u}_1, \vec{y} \rangle + \alpha_2 \langle \vec{u}_2, \vec{y} \rangle .$$

Vocabulaire : Un espace vectoriel réel E de dimension infinie (respectivement finie) muni d'un produit scalaire est appelé **espace préhilbertien réel**, (respectivement **espace euclidien**).

Exemples : Montrer que les applications suivantes sont des produits scalaires sur E :

- $E = \mathbf{R}^n, \langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i$;
- $E = \mathcal{M}_n(\mathbf{R}), \langle A, B \rangle = \text{Tr}({}^tAB) = \text{Tr}(A^T B)$; (à prouver pour $n = 2$)
- $E = \mathcal{C}^0([a, b], \mathbf{R}), \langle f, g \rangle = \int_a^b f(t)g(t) dt$.

Mise en œuvre : exercice 01.

Proposition 12.1.— Inégalité de Cauchy-Schwarz —. Si \langle , \rangle est un produit scalaire de E :

$$\forall (\vec{x}, \vec{y}) \in E^2, |\langle \vec{x}, \vec{y} \rangle| \leq \sqrt{\langle \vec{x}, \vec{x} \rangle} \sqrt{\langle \vec{y}, \vec{y} \rangle} .$$

De plus, \vec{x} et \vec{y} sont colinéaires si et seulement si l'on a l'égalité.

Mise en œuvre : exercice 04.

Corollaire 12.2.— Soit (E, \langle , \rangle) un espace préhilbertien réel ou un espace euclidien. L'application N de E dans \mathbf{R}_+ définie par : $\forall \vec{x} \in E, N(\vec{x}) = \sqrt{\langle \vec{x}, \vec{x} \rangle}$ vérifie :

$$\begin{aligned} \forall \vec{x} \in E, \forall \lambda \in \mathbf{R}, \quad N(\lambda \vec{x}) &= |\lambda| N(\vec{x}) ; \\ \forall \vec{x} \in E, \quad N(\vec{x}) &= 0 \quad \Rightarrow \quad \vec{x} = 0 ; \\ \forall (\vec{x}, \vec{y}) \in E^2, \quad N(\vec{x} + \vec{y}) &\leq N(\vec{x}) + N(\vec{y}) . \end{aligned}$$

Corollaire 12.3.— Norme associée à un produit scalaire—.

Soit $(E, \langle \cdot, \cdot \rangle)$ espace préhilbertien réel ou euclidien. L'application N définie ci-dessus est une norme sur E , appelée *norme euclidienne associée au produit scalaire* $\langle \cdot, \cdot \rangle$.

Notation : Nous adopterons dans la suite du chapitre la notation usuelle : $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$.

Remarque : On peut réécrire les propriétés vues ci-dessus avec les notations usuelles :

► **Inégalité de Cauchy-Schwarz**

$$\forall (\vec{x}, \vec{y}) \in E^2, |\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|$$

► **Identité de polarisation**

$$\forall (\vec{x}, \vec{y}) \in E^2, \langle \vec{x}, \vec{y} \rangle = \frac{1}{4} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2) = \frac{1}{2} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2)$$

► **Propriétés de la norme**

$$\begin{array}{ll} \forall \vec{x} \in E, \forall \lambda \in \mathbf{R}, & \|\lambda \vec{x}\| = |\lambda| \|\vec{x}\|; \\ \forall \vec{x} \in E, & \|\vec{x}\| = 0 \iff \vec{x} = \vec{0} \\ \forall (\vec{x}, \vec{y}) \in E^2, & \|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \end{array}$$

Proposition 12.4.— Identité du parallélogramme —.

$$2 (\|\vec{x}\|^2 + \|\vec{y}\|^2) = (\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2).$$

■ Orthogonalité

Dans ce paragraphe, on munit E du produit scalaire $\langle \cdot, \cdot \rangle$. La norme associée est notée $\|\cdot\|$.

Définition : Un vecteur est dit *unitaire* si sa norme est égale à 1.

Définition : Deux vecteurs de E sont *orthogonaux* lorsque leur produit scalaire est nul.

Remarque : Le vecteur nul est orthogonal à tout vecteur.

Définition : **Sous-espaces vectoriels orthogonaux —.** Deux sous-espaces vectoriels F et G de E sont *orthogonaux* lorsque :

$$\forall (\vec{x}, \vec{y}) \in F \times G, \langle \vec{x}, \vec{y} \rangle = 0.$$

On note alors : $F \perp G$.

Proposition 12.5.— Si F et G sont orthogonaux, $F \cap G = \{0_E\}$.

La somme de deux sous-espaces vectoriels orthogonaux est directe.

Définition : **Orthogonal d'un s.e.v —.** Soit X un sous-espace vectoriel de E . On appelle *orthogonal* de X et on note X^\perp l'ensemble des vecteurs de E qui sont orthogonaux à tout vecteur de X : $X^\perp = \{\vec{y} \in E \text{ tel que } \forall \vec{x} \in X, \langle \vec{x}, \vec{y} \rangle = 0\}$.

Proposition 12.6.— Sous-espace vectoriel orthogonal d'un s.e.v —.

L'orthogonal X^\perp de X , sous-espace vectoriel de E , est lui-même un sous-espace vectoriel de E .

Proposition 12.7.— Si \mathcal{B} est une base de X , sous-espace vectoriel de E , un vecteur est orthogonal à X si et seulement s'il est orthogonal à tout vecteur de \mathcal{B} .

Mise en œuvre : exercice 08.

Théorème-Définition 12.8.— **Somme directe orthogonale** —. Si F_1, F_2, \dots, F_p sont p sous-espaces vectoriels de E deux à deux orthogonaux, alors leur somme est une somme directe. On dit que la somme de ces sous-espaces vectoriels est une **somme directe orthogonale**.

Définition : Familles orthogonales, orthonormées —. Une famille de vecteurs indexée sur une famille I , $(\vec{u}_i)_{i \in I}$, est **orthogonale** lorsque ses éléments sont orthogonaux deux à deux :

$$\forall (i, j) \in I^2, (i \neq j \Rightarrow \langle \vec{u}_i, \vec{u}_j \rangle = 0).$$

Si, de plus, tous ses éléments sont des vecteurs unitaires, alors la famille est **orthonormée**.

Si $(\vec{u}_i)_{i \in I}$ est une famille orthonormée : $\forall (i, j) \in I^2, \langle \vec{u}_i, \vec{u}_j \rangle = \delta_i^j$, où δ_i^j est le **symbole de Kronecker**, qui vaut 0, si $i \neq j$ et 1 si $i = j$.

Proposition 12.9.— **Liberté d'une famille orthogonale sans vecteur nul** —.

Une famille *orthogonale* finie, dont aucun des vecteurs n'est nul, est libre.

En particulier, toute famille *orthonormée* finie est libre.

Proposition 12.10.— **Relation de Pythagore** —.

Si $(\vec{u}_i)_{1 \leq i \leq n}$ est une famille orthogonale finie, alors : $\left\| \sum_{i=1}^n \vec{u}_i \right\|^2 = \sum_{i=1}^n \|\vec{u}_i\|^2$.

Proposition 12.11.— Soit F un sous-espace vectoriel de E , alors :

- ▶ F et F^\perp sont des sous-espaces vectoriels orthogonaux.
- ▶ Si G est un sous-espace vectoriel orthogonal à F alors $G \subset F^\perp$.
- ▶ Si F et F^\perp sont supplémentaires, F^\perp s'appelle le **supplémentaire orthogonal** de F et par définition, il est unique.

Proposition 12.12.— **Existence et unicité du supplémentaire orthogonal** —.

Tout sous-espace vectoriel de **dimension finie** de E admet un supplémentaire orthogonal.

Corollaire 12.13.— En dimension finie, si F s.e.v de E : $\dim F + \dim F^\perp = \dim E$.

Définition : Projection orthogonale —. Soit F un sous-espace vectoriel de E admettant un supplémentaire orthogonal. La **projection orthogonale** p_F **sur** F est la projection sur F parallèlement à son supplémentaire orthogonal F^\perp . Alors : $\text{Ker } p_F = (\text{Im } p_F)^\perp$ et $(\text{Ker } p_F)^\perp = \text{Im } p_F$.

Proposition 12.14.— **Propriétés d'une projection orthogonale** —.

Si p_F est la projection orthogonale sur F , pour tout $x \in E$,

$$\vec{x} - p_F(\vec{x}) \in F^\perp, \quad \|\vec{x}\|^2 = \|p_F(\vec{x})\|^2 + \|\vec{x} - p_F(\vec{x})\|^2, \quad \|p_F(\vec{x})\| \leq \|\vec{x}\|.$$

Mise en œuvre : exercice 05.

■ Orthonormalisation. Cas général

Proposition 12.15.— Orthogonalisation/Orthonormalisation de Gram-Schmidt —.

Soit F un sous-espace de dimension finie de E et $(\vec{u}_i)_{1 \leq i \leq n}$ une base de F .

On pose $\vec{v}_1 = \vec{u}_1$ et pour k variant de 1 à $n - 1$:

$$\vec{v}_{k+1} = \vec{u}_{k+1} - \sum_{i=1}^k \frac{\langle \vec{u}_{k+1}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$$

Alors la famille $(\vec{v}_i)_{1 \leq i \leq n}$ est une base *orthogonale* de F adaptée aux espaces $U_p = \text{Vect}((\vec{u}_i)_{1 \leq i \leq p})$, c'est-à-dire que $\text{Vect}((\vec{v}_i)_{1 \leq i \leq p}) = \text{Vect}((\vec{u}_i)_{1 \leq i \leq p})$.

Pour obtenir une base *orthonormale* de F , il suffit de poser ensuite : $\vec{w}_i = \frac{1}{\|\vec{v}_i\|} \vec{v}_i$, pour $i \in \llbracket 1, n \rrbracket$.

Exemple : On pourra voir la **méthode 12.8** dans le cas classique $n = 3$. On munit $\mathbf{R}_2[X]$ du produit scalaire : $\langle P, Q \rangle = P(1)Q(1) + P'(1)Q'(1) + P''(1)Q''(1)$. Déterminer une base orthonormée de $\mathbf{R}_2[X]$ pour ce produit scalaire.

Ici $F = E = \mathbf{R}_2[X]$. Un espace vectoriel est un sous-espace vectoriel de lui-même.

Il suffit maintenant de partir de la base canonique $\{1, X, X^2\}$ et de lui appliquer le procédé d'orthonormalisation de Gram-Schmidt. C'est à vous !

Mise en œuvre : exercice 02 et exercice 03.

Remarque : On peut se poser la question. E possède-t-il toujours une base orthonormale ?

La réponse est oui si E est de dimension finie. C'est l'objet du prochain paragraphe.

Proposition 12.16.— Projection orthogonale sur un sous-espace vectoriel de dimension finie

Soit F un sous-espace vectoriel de dimension finie de E , $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ une base *orthogonale* de F et p_F la projection orthogonale sur F , \vec{x} étant un vecteur quelconque de E :

$$p_F(\vec{x}) = \sum_{i=1}^p \frac{\langle \vec{e}_i, \vec{x} \rangle}{\|\vec{e}_i\|^2} \vec{e}_i$$

Corollaire 12.17.— Dans le cas où la base est *orthonormale*, cette formule devient :

$$p_F(x) = \sum_{i=1}^p \langle \vec{e}_i, \vec{x} \rangle \vec{e}_i.$$

Remarque : Ces deux formules permettent de rendre opérationnel l'algorithme de Gram-Schmidt, décrit ci-dessous.

Mise en œuvre : exercice 06 et exercice 07.

Proposition 12.18.— Inégalité de Bessel —. Soit $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ une famille orthonormale finie de l'espace préhilbertien ou euclidien $(E, \langle \cdot, \cdot \rangle)$ et $F = \text{Vect}(\vec{e}_1, \dots, \vec{e}_p)$.

$$\forall \vec{x} \in E, \|p_F(\vec{x})\|^2 = \sum_{i=1}^p |\langle \vec{e}_i, \vec{x} \rangle|^2 \leq \|\vec{x}\|^2.$$

■ Orthonormalisation dans le cas où E est un espace euclidien

On suppose ici que E est de dimension finie.

Proposition 12.19.— Théorème d'orthonormalisation de Gram-Schmidt —.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien de dimension n et $\mathcal{B} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ une base de E .

Il existe une base orthonormale $\mathcal{B}' = (\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n)$ de E telle que :

$$\forall i \in \llbracket 1, n \rrbracket, \vec{w}_i \in \text{Vect}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_i).$$

Remarque : on utilise dans la pratique l'algorithme d'orthonormalisation de Gram-Schmidt de la **proposition 12.15** et en particulier la **méthode 12.8** qui permet de « faire fonctionner » l'algorithme pour n pas trop grand.

Corollaire 12.20.— Existence d'une base orthonormale en dimension finie —.

Tout espace vectoriel euclidien admet au moins une base orthonormale.

Toute famille orthonormale d'un espace vectoriel euclidien peut être complétée en une base orthonormale.

Proposition 12.21.— Expression du produit scalaire et de la matrice d'un endomorphisme, E étant rapporté à une de ses bases orthonormales—.

Si E est rapporté à une de ses bases orthonormales $\mathcal{B} = (\vec{w}_i)_{1 \leq i \leq n}$ et si l'on pose $\vec{x}(x_1, \dots, x_n)$ et $\vec{y}(y_1, \dots, y_n)$ alors : $\forall k \in \llbracket 1, n \rrbracket, \langle \vec{x}, \vec{w}_k \rangle = x_k$ et on a les assertions :

$$(i) \|\vec{x}\| = \sqrt{\sum_{i=1}^n x_i^2}, (ii) \langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i, (iii) d(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Puis, si $X = \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix}$ sont les matrices colonnes associées, $\langle \vec{x}, \vec{y} \rangle = X^T Y$.

■ Distance d'un vecteur à un sous-espace vectoriel dans un espace euclidien

Corollaire 12.22.— Distance et projection orthogonale —. Pour tout vecteur \vec{x} de E , il existe un unique vecteur \vec{z} de F tel que : $d(\vec{x}, F) = \min \{\|\vec{x} - \vec{y}\|, \vec{y} \in F\} = d(\vec{x}, \vec{z})$. Ce vecteur est $p_F(\vec{x})$.

Mise en œuvre : exercice 09 à exercice 11.

Méthodes

■ Produits scalaires

□ Méthode 12.1.— Comment montrer que \langle , \rangle est un produit scalaire

- On commence par vérifier (à moins que cela ne soit évident) que $(\vec{x}, \vec{y}) \mapsto \langle \vec{x}, \vec{y} \rangle$ est une application de E^2 dans \mathbf{R} .
- On vérifie ensuite que \langle , \rangle est symétrique : $\forall (\vec{x}, \vec{y}) \in E^2, \langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$.
- On vérifie la linéarité (mais d'un seul côté ! En effet, la symétrie montrée en amont affirme alors la linéarité de l'autre côté). Par exemple, on montre que $\vec{y} \mapsto \langle \vec{x}, \vec{y} \rangle$ est linéaire pour tout \vec{x} fixé.
- On vérifie que \langle , \rangle est positive, c'est-à-dire que $\langle \vec{x}, \vec{x} \rangle \geq 0$ pour tout $\vec{x} \in E$.
- Il reste à vérifier que \langle , \rangle est définie, c'est-à-dire que :

$$\forall \vec{x} \in E, \langle \vec{x}, \vec{x} \rangle = 0 \Rightarrow \vec{x} = \vec{0}.$$

Remarque : On dit qu'un produit scalaire est une forme bilinéaire symétrique définie positive. Ce vocabulaire n'est plus dans le programme officiel de 2TSI. Mais, culture, quand tu nous tiens...

Exemple : Montrer : $(\mathbf{R}_3[X])^2 \rightarrow \mathbf{R}, (P, Q) \mapsto \sum_{k=0}^{2014} P(k)Q(k)$ est un produit scalaire sur $\mathbf{R}_3[X]$.

- $(P, Q) \mapsto \langle P, Q \rangle$ est une application de $(\mathbf{R}_3[X])^2$ dans \mathbf{R} d'après l'énoncé. Facile!
- \langle , \rangle est bien symétrique car pour tout $(P, Q) \in (\mathbf{R}_3[X])^2$:

$$\langle P, Q \rangle = \sum_{k=0}^{2014} P(k)Q(k) = \sum_{k=0}^{2014} Q(k)P(k) = \langle Q, P \rangle.$$

- Fixons P et soit $\phi_P : Q \mapsto \sum_{k=0}^{2014} P(k)Q(k)$. Mission : $\phi_P \in \mathcal{L}(\mathbf{R}_3[X])$? Alors :

$$\forall (Q_1, Q_2) \in (\mathbf{R}_3[X])^2, \forall a \in \mathbf{R}, \phi_P(aQ_1 + Q_2) = \sum_{k=0}^{2014} P(k)[aQ_1(k) + Q_2(k)] = a\phi_P(Q_1) + \phi_P(Q_2).$$

- \langle , \rangle est positive car $\langle P, P \rangle = \sum_{k=0}^{2014} P(k)^2 \geq 0$ pour tout $P \in \mathbf{R}_3[X]$.
- $\langle P, P \rangle = \sum_{k=0}^{2014} P(k)^2 = 0 \Rightarrow [P \in \mathbf{R}_3[X] \text{ admet 2015 racines distinctes}] \Rightarrow P = 0$.

Passons à un cas particulier de produit scalaire assez courant !

□ **Méthode 12.2.—** Comment montrer que $(P, Q) \mapsto \langle P, Q \rangle = \int_a^b P(t)Q(t)w(t)dt$ définit un produit scalaire sur $\mathbf{R}[X]$

$]a, b[$ est un intervalle de \mathbf{R} , w une fonction strictement positive, continue et intégrable sur $]a, b[$.

- Vérifier si besoin est que l'intégrale $\langle P, Q \rangle$ est convergente (dans le cas où elle est généralisée bien sûr !)
- La symétrie et la bilinéarité résultent des propriétés de la multiplication dans \mathbf{R} et de la linéarité de l'intégrale.
- Le caractère positif résulte de la positivité de l'intégrale ($b > a$) et de celui des fonctions w et P^2 .
- Le caractère défini est dû :
 - ▷ au fait que $w.P^2$ est positive et continue sur $]a, b[$, alors

$$\int_a^b P^2 w = 0 \Rightarrow P^2 w = 0 \Rightarrow P = 0 \text{ sur }]a, b[;$$

- ▷ au fait que P est un polynôme : il ne peut s'annuler sur un intervalle $]a, b[$ que s'il est nul car il aurait une infinité de racines.

Exemple : Montrer que $(P, Q) \mapsto \langle P, Q \rangle = \int_{-1}^1 \frac{P(t)Q(t)}{\sqrt{1-t^2}} dt$ définit un produit scalaire sur $\mathbf{R}[X]$.

- **Existence de l'intégrale :** si $(P, Q) \in \mathbf{R}[X]$ alors $\frac{|P(t)Q(t)|}{\sqrt{1-t^2}} = O\left(\frac{1}{\sqrt{1-t^2}}\right)$ au $V(\pm 1)$.

Par ailleurs, $\int_{-1}^1 \frac{dt}{\sqrt{1-t^2}}$ converge donc $\langle P, Q \rangle$ converge absolument.

- **Symétrie et bilinéarité :** elles résultent des propriétés de linéarité de l'intégrale.

- **Positivité :** si $P \in \mathbf{R}[X]$ alors $\langle P, P \rangle = \int_{-1}^1 \frac{P(t)^2}{\sqrt{1-t^2}} dt \geq 0$.

- **Caractère défini :** comme P et $t \mapsto \frac{1}{\sqrt{1-t^2}}$ sont positives et continues sur $] -1, 1[$, il en va de même pour leur produit et $\langle P, P \rangle = 0 \Rightarrow \forall t \in] -1, 1[, P(t) = 0$.
Comme un polynôme qui s'annule sur un intervalle vaut 0, $\langle P, P \rangle = 0 \Rightarrow P = 0$.

Remarque : Pour plus de précisions quant aux notions d'intégrale convergente et d'intégrabilité, on se reportera avec profit au chapitre concernant les intégrales sur un intervalle.

Remarque : On peut remplacer dans la **méthode 12.2** les polynômes P et Q par des fonctions f et g continues sur $[a, b]$. Le processus est identique.

□ Méthode 12.3.— Quand et comment utiliser l'inégalité de Cauchy-Schwarz

Nous rencontrons des espaces préhilbertiens ou euclidiens dont les éléments sont des matrices, des applications, des n -uplets de réels ou de complexes : on peut donc mettre en œuvre l'inégalité de Cauchy-Schwarz dans des situations très variées.

- Dans l'espace préhilbertien réel ou euclidien $(E, \langle \cdot, \cdot \rangle)$, l'inégalité de Cauchy-Schwarz s'écrit de deux façons possibles :

$$\forall (\vec{u}, \vec{v}) \in E^2, |\langle \vec{u} | \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\| \Leftrightarrow |\langle \vec{u} | \vec{v} \rangle|^2 \leq \|\vec{u}\|^2 \|\vec{v}\|^2.$$

- Si l'on reconnaît cette structure dans une inégalité à démontrer, il importe, d'abord, de définir soigneusement le produit scalaire que l'on va utiliser.
- Il faut, ensuite, choisir tout aussi soigneusement les deux vecteurs pour lesquels on va écrire l'inégalité de Cauchy-Schwarz afin d'aboutir à l'inégalité attendue.
- L'égalité a lieu si et seulement si les deux vecteurs sont colinéaires.

Exemples :

1. a. Montrer : $\left\{ \begin{array}{ccc} (\mathbf{R}^3)^2 & \longrightarrow & \mathbf{R} \\ ((x, y, z), (x', y', z')) & \mapsto & 2xx' + yy' + 5zz' \end{array} \right.$ est un produit scalaire sur \mathbf{R}^3 .

b. x, y et z sont trois réels tels que $2x^2 + y^2 + 5z^2 \leq 1$. Montrer que : $(x + y + z)^2 \leq \frac{17}{10}$.

1. a. Nous laissons au lecteur le soin de « s'éclater » à développer les différentes assertions qui confèrent à $\langle (x, y, z), (x', y', z') \rangle$ le titre tant envié de produit scalaire sur \mathbf{R}^3 .

b. Appliquons l'inégalité de Cauchy-Schwarz pour le produit scalaire ci-dessus aux vecteurs (x, y, z) et $\left(\frac{1}{2}, 1, \frac{1}{5}\right)$, on obtient :

$$(x + y + z)^2 \leq \|(x, y, z)\|^2 \left\| \left(\frac{1}{2}, 1, \frac{1}{5}\right) \right\|^2 = \frac{17}{10}.$$

Attention aux réflexes idiots, $\|(x, y, z)\|^2$ vaut ici $2x^2 + y^2 + 5z^2$ et non $x^2 + y^2 + z^2$.

2. f étant une fonction continue sur $[0, +\infty[$ et a et b étant deux réels tels que $0 \leq a < b$, montrer que : $\frac{b^3 - a^3}{3} \int_a^b f^2(t) dt \geq \left(\int_a^b t f(t) dt \right)^2$. Étudier le cas d'égalité.

Appliquons l'inégalité de Cauchy-Schwarz pour le produit scalaire canonique sur $\mathcal{C}^0([a, b], \mathbf{R})$ aux fonctions $t \mapsto t$ et f :

$$\left(\int_a^b t f(t) dt \right)^2 \leq \left(\int_a^b t^2 dt \right) \left(\int_a^b f^2(t) dt \right),$$

d'où le résultat.

Il y a égalité si et seulement si f et $t \mapsto t$ sont colinéaires, donc si et seulement s'il existe un réel k tel que, pour $t \in [a, b]$, $f(t) = kt$.

■ Orthogonalité

□ **Méthode 12.4.—** Comment montrer l'orthogonalité de deux vecteurs ou de deux sous-espaces vectoriels de E

Dans $(E, \langle \cdot, \cdot \rangle)$:

- Soit $(\vec{u}, \vec{v}) \in E^2$, alors : $\vec{u} \perp \vec{v} \Leftrightarrow \langle \vec{u}, \vec{v} \rangle = 0$.
- Si F et G sont deux sous-espaces vectoriels de E ,

$$F \perp G \Leftrightarrow (\forall (\vec{u}, \vec{v}) \in F \times G, \langle \vec{u}, \vec{v} \rangle = 0).$$

Exemple : Montrer que les fonctions $t \mapsto \sin nt$ et $t \mapsto \cos pt$, où n et p sont des entiers naturels non nuls, sont des vecteurs orthogonaux de $E = \mathcal{C}^0([0, 2\pi], \mathbf{R})$ pour le produit scalaire :

$$(f, g) \mapsto \langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)g(t) dt.$$

Le produit de ces deux fonctions est $\langle \sin nt, \cos pt \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin nt \cos pt dt$.

$t \mapsto \sin nt \cos pt$ est aussi une fonction impaire, donc cette intégrale est nulle : les fonctions proposées sont des vecteurs orthogonaux de E pour ce produit scalaire.

Remarque : Attention, cette notion d'orthogonalité dépend du produit scalaire choisi !

□ **Méthode 12.5.—** Comment montrer que deux sous-espaces vectoriels sont supplémentaires orthogonaux

On est dans $(E, \langle \cdot, \cdot \rangle)$:

- On commence par montrer l'orthogonalité en utilisant la **méthode 12.4** puisque deux sous-espaces vectoriels orthogonaux ont leur intersection réduite à $\{0_E\}$.
Il suffit, ensuite, de montrer que E est égal à leur somme ou, en dimension finie, que la dimension de E est la somme de leurs dimensions.
- Si un sous-espace vectoriel F de E admet un supplémentaire orthogonal (et c'est le cas s'il est de **dimension finie**), celui-ci est unique et c'est F^\perp .

$$F^\perp = \{x \in E / \forall y \in F, \langle x, y \rangle = 0\}.$$

Exemple : $\mathcal{M}_n(\mathbf{R})$ est muni du produit scalaire canonique : $\langle A | B \rangle = \text{Tr}(A^T B)$, $\mathcal{S}_n(\mathbf{R})$ est le sous-espace vectoriel des matrices symétriques et $\mathcal{A}_n(\mathbf{R})$ est le sous-espace vectoriel des matrices antisymétriques. Montrer que $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont supplémentaires orthogonaux.

$\forall (A, B) \in \mathcal{S}_n(\mathbf{R}) \times \mathcal{A}_n(\mathbf{R}), \langle A, B \rangle = \text{Tr}(A^T B) = \text{Tr}(AB)$; or, une matrice et sa transposée ont la même trace et $(AB)^T = B^T A^T = -BA$, donc $\text{Tr}(AB) = -\text{Tr}(BA)$.

D'autre part, il est bien connu que : $\text{Tr}(AB) = \text{Tr}(BA)$, donc : $\langle A, B \rangle = \text{Tr}(AB) = 0$.

$\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont orthogonaux, donc leur somme est directe orthogonale.

$\forall A \in \mathcal{M}_n(\mathbf{R}), A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$. Comme $\frac{1}{2}(A + A^T)$ est symétrique et $\frac{1}{2}(A - A^T)$ antisymétrique, $\mathcal{M}_n(\mathbf{R}) = \mathcal{S}_n(\mathbf{R}) + \mathcal{A}_n(\mathbf{R})$.

$\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont supplémentaires orthogonaux dans $\mathcal{M}_n(\mathbf{R})$.

□ Méthode 12.6.— Comment montrer qu'une application est un projecteur orthogonal dans E

Soit p un endomorphisme de $(E, \langle \cdot, \cdot \rangle)$.

- On s'assure d'abord que p est un projecteur (*i.e* $p \circ p = p$).
- Si l'on connaît $\text{Mat}_{\mathcal{B}}(p)$ et si \mathcal{B} est orthonormale pour $\langle \cdot, \cdot \rangle$, il suffit de vérifier que cette matrice est symétrique. Alors p est un projecteur orthogonal.
- On peut montrer sinon que : $\forall x \in E, \|p(\vec{x})\| \leq \|\vec{x}\|$. (Voir **exercice ??**.)
- On peut montrer aussi que $\text{Ker } p$ et $\text{Im } p$ sont orthogonaux avec la **méthode 12.4**.

□ Méthode 12.7.— Comment déterminer le projeté orthogonal d'un vecteur sur un sous-espace vectoriel

F est un sous-espace vectoriel de $(E, \langle \cdot, \cdot \rangle)$ admettant un supplémentaire orthogonal et \vec{x} est un vecteur de E . Savoir calculer pratiquement le projeté orthogonal de \vec{x} sur F est utile pour déterminer la matrice d'une projection orthogonale et pour résoudre certains problèmes de minimisation de distance.

- Si l'on dispose d'une base orthogonale $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ du sous-espace vectoriel F de dimension finie, on emploie la formule : $p_F(\vec{x}) = \sum_{i=1}^p \frac{\langle \vec{e}_i, \vec{x} \rangle}{\|\vec{e}_i\|^2} \vec{e}_i$.
- Si ce n'est pas le cas, on peut avoir recours à une orthogonalisation de Gram-Schmidt, allez voir la **méthode 12.8**.
- On peut également utiliser la caractérisation :

$$\vec{y} = p_F(\vec{x}) \Leftrightarrow \vec{y} \in F \text{ et } \vec{x} - \vec{y} \in F^\perp.$$
On traduira l'appartenance à F^\perp par l'orthogonalité à chaque vecteur d'une base de F .
- Évidemment, si l'on a : $\vec{x} = \vec{x}_F + \vec{x}_{F^\perp}$, avec $(\vec{x}_F, \vec{x}_{F^\perp}) \in F \times F^\perp$, alors :
 $p_F(\vec{x}) = \vec{x}_F$.

Exemple : E étant un préhilbertien de dimension finie, déterminer l'image d'un vecteur \vec{x} de E par la projection orthogonale sur une droite vectorielle, puis par la projection orthogonale sur son hyperplan supplémentaire orthogonal.

Considérons la droite $D = \mathbf{K}\vec{u}$, où \vec{u} est un vecteur non nul, $H = D^\perp$ et (p, q) le couple de projecteurs orthogonaux associés à la décomposition $E = D \oplus H$.

Tout vecteur \vec{x} de E se décompose de manière unique en : $\vec{x} = k\vec{u} + \vec{y}$, $k \in \mathbf{K}$, $\vec{y} \in H$.

$$\langle \vec{u}, \vec{x} \rangle = k \|\vec{u}\|^2, \text{ d'où } p(\vec{x}) = k \vec{u} = \frac{\langle \vec{u}, \vec{x} \rangle}{\|\vec{u}\|^2} \vec{u}.$$

q est la projection orthogonale sur H , donc $q = \text{Id}_E - p$ et : $\forall \vec{x} \in E, q(\vec{x}) = \vec{x} - p(\vec{x}) = \vec{x} - \frac{\langle \vec{u}, \vec{x} \rangle}{\|\vec{u}\|^2} \vec{u}$.

Remarque : Ainsi, dans le cas particulier de la projection orthogonale \vec{x}_F d'un vecteur \vec{x} sur un hyperplan F de E , il est plus simple en général de déterminer le projeté orthogonal \vec{x}_{F^\perp} sur la droite vectorielle F^\perp et d'utiliser : $\vec{x}_F = \vec{x} - \vec{x}_{F^\perp}$.

□ Méthode 12.8.— Comment déterminer une base orthonormée d'un sous-espace vectoriel F de $(E, \langle \cdot, \cdot \rangle)$ en appliquant le procédé d'orthonormalisation de Gram-Schmidt

On suppose ici $\dim F = 3$ et notons $\mathcal{B} = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$ une base de F .

► On pose $\vec{v}_1 = \vec{u}_1$ puis $\vec{v}_2 = \vec{u}_2 + a_1 \vec{v}_1$, où $a_1 \in \mathbf{R}$ est déterminé par la condition : $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$. Cela donne : $a_1 = -\frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2}$ puis on explicite \vec{v}_2 .

► On pose ensuite $\vec{v}_3 = \vec{u}_3 + b_1 \vec{v}_1 + b_2 \vec{v}_2$, où $(b_1, b_2) \in \mathbf{R}^2$ est déterminé par les conditions : $\langle \vec{v}_1, \vec{v}_3 \rangle = \langle \vec{v}_2, \vec{v}_3 \rangle = 0$.

On trouve : $b_1 = -\frac{\langle \vec{v}_1, \vec{u}_3 \rangle}{\|\vec{v}_1\|^2}$ et $b_2 = -\frac{\langle \vec{v}_2, \vec{u}_3 \rangle}{\|\vec{v}_2\|^2}$ puis on explicite \vec{v}_3 .

► À ce niveau, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ est une base orthogonale de F .

La famille $\left(\frac{\vec{v}_1}{\|\vec{v}_1\|}, \frac{\vec{v}_2}{\|\vec{v}_2\|}, \frac{\vec{v}_3}{\|\vec{v}_3\|} \right)$ est une base orthonormée de F .

Remarque : on peut utiliser directement les formules de la **proposition 12.15** (qui fournissent a_1, b_1 et b_2) mais on vous conseille de les retrouver en utilisant $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ pour $i \neq j$.

Exemple : On munit $\mathbf{R}_2[X]$ du produit scalaire : $\langle P, Q \rangle = P(1)Q(1) + P'(1)Q'(1) + P''(1)Q''(1)$. Déterminer une base orthonormée de $\mathbf{R}_2[X]$ pour ce produit scalaire.

Ici $F = E = \mathbf{R}_2[X]$. Un espace vectoriel est un sous-espace vectoriel de lui-même.

Il suffit maintenant de partir de la base canonique $\{1, X, X^2\}$ et de lui appliquer le procédé d'orthonormalisation de Gram-Schmidt.

On cherche trois polynômes V_1, V_2 et V_3 tels que :
$$\begin{cases} V_1 &= 1 \\ V_2 &= X + a.1 \\ V_3 &= X^2 + b.1 + cV_2 \end{cases}.$$

On utilise $\langle V_1, V_2 \rangle = \langle V_1, V_3 \rangle = \langle V_2, V_3 \rangle = 0$. On trouve $a = -1$, puis $b = -1$ et $c = -2$.

On en déduit :
$$\begin{cases} V_1 &= 1 \\ V_2 &= X - 1 \\ V_3 &= (X - 1)^2 \end{cases} \text{ puis : } \langle V_1, V_1 \rangle = \langle V_2, V_2 \rangle = 1 \text{ et } \langle V_3, V_3 \rangle = 4.$$

Une base orthonormale de $\mathbf{R}_2[X]$ pour $\langle \cdot, \cdot \rangle$ est : $\{W_1 = 1, W_2 = X - 1, W_3 = \frac{1}{2}(X - 1)^2\}$.

On a fini mais remarquons que l'on aurait pu partir de n'importe quel base de $\mathbf{R}_2[X]$ et pas nécessairement de la base canonique. Ainsi, partir de la base $\{1, X - 1, (X - 1)^2\}$ aurait été intelligent car cette base est orthogonale pour $\langle \cdot, \cdot \rangle$!

Remarque : On peut étendre la **méthode 12.8** dans le cas où $\dim F = n \geq 4$. Mais les calculs sont de plus en plus techniques, à moins de démarrer avec une base « quasi » orthogonale pour

\langle, \rangle . (Voir à ce sujet l'exemple précédent.) Tout cela dépend donc de l'énoncé de l'exercice ou du problème. Maintenant, s'il s'agit d'une application de Python, c'est la machine qui fait les calculs !

■ Les problèmes classiques dans un espace préhilbertien ou euclidien

□ **Méthode 12.9.**— Comment déterminer la distance d'un vecteur à un sous-espace vectoriel F de E , ce sous-espace vectoriel ayant pour base orthonormale $(\vec{e}_1, \dots, \vec{e}_p)$

On utilise le projeté $p_F(\vec{x})$ d'un vecteur \vec{x} sur un espace F pour calculer $d(\vec{x}, F)$:

► $d(\vec{x}, F)^2 = \|\vec{x} - p_F(\vec{x})\|^2 = \|\vec{x}\|^2 - \|p_F(\vec{x})\|^2$ d'après le théorème de Pythagore.

► Si $p(\vec{x})$ est obtenu grâce à la formule $p_F(x) = \sum_{k=1}^p \langle \vec{e}_k, \vec{x} \rangle \vec{e}_k$, alors :

$$\|p_F(\vec{x})\|^2 = \sum_{k=1}^p |\langle \vec{e}_k, \vec{x} \rangle|^2, \text{ toujours par application du théorème de Pythagore.}$$

Exemple : On munit encore $\mathbf{R}_2[X]$ de : $\langle P, Q \rangle = P(1)Q(1) + P'(1)Q'(1) + P''(1)Q''(1)$. Calculer la distance de X^2 à $\mathbf{R}_1[X]$.

Appelons d la distance de X^2 à $\mathbf{R}_1[X]$ et Z la projection orthogonale de X^2 sur $\mathbf{R}_1[X]$.

Il est clair que (d'après l'exemple qui suit la **méthode 12.8**) que $\{W_1, W_2\}$ est une base orthonormale de $\mathbf{R}_1[X]$ pour \langle, \rangle (Justifiez le vous-même!).

Alors : $Z = \langle X^2, 1 \rangle 1 + \langle X^2, X - 1 \rangle (X - 1) = 1 + 2(X - 1) = 2X - 1$.

La distance cherchée est : $d = \|X^2 - Z\| = \|(X - 1)^2\| = 2$.

□ **Méthode 12.10.**— Comment résoudre certains problèmes de minimisation dans un préhilbertien

On considère une fonction φ de \mathbf{R}^p ($p \in \mathbf{N}^*$) dans \mathbf{R}_+ à minimiser, c'est-à-dire que l'on cherche sa valeur minimale (et éventuellement, les antécédents de cette valeur minimale).

Si l'on peut trouver un espace préhilbertien réel E , un vecteur \vec{v} de E et un sous-espace F de E de dimension p et de base $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ tels que :

$$\forall (x_i)_{i \in [1, p]} \in \mathbf{R}^p, \quad \varphi(x_1, x_2, \dots, x_p) = \left\| \vec{v} - \sum_{i=1}^p x_i \vec{u}_i \right\|^2,$$

alors le minimum de φ est $(d(\vec{v}, F))^2$; il est atteint lorsque $\vec{u} = \sum_{i=1}^p x_i \vec{u}_i$ est le projeté orthogonal de \vec{v} sur F . Il est souvent pratique de choisir $(\vec{u}_1, \dots, \vec{u}_p)$ orthonormée.

Exemples : 1. Dans $\mathbf{R}_2[X]$ muni de $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$, on pose $F = \mathbf{R}_1[X]$.

a. Déterminer une base orthonormée de F puis la projection orthogonale p sur F .

b. Montrer que $d(t^2, F) = \|t^2 - p(t^2)\|$ et déterminer cette distance.

c. En déduire $\inf_{(a,b) \in \mathbb{R}^2} \int_0^1 (t^2 - a - bt)^2 dt$.

a. On détermine une base orthonormée de $F = \mathbf{R}_1[X]$ en partant par exemple de $\{1, t\}$ et en utilisant le procédé d'orthonormalisation de Gram-Schmidt. Après calcul : $\{1, \sqrt{3}(2t-1)\}$. Ainsi, si P est un polynôme quelconque de $\mathbf{R}_2[X]$,

$$p(P(t)) = \langle P, 1 \rangle 1 + 3 \langle P, 2t-1 \rangle (2t-1) = \int_0^1 P(t) dt + 3 \left[\int_0^1 P(t)(2t-1) dt \right] (2t-1).$$

b. L'égalité $d(t^2, F) = \|t^2 - p(t^2)\|$ découle du cours. On commence par calculer :

$$p(t^2) = \int_0^1 t^2 dt + 3 \left[\int_0^1 (2t^3 - t^2) dt \right] (2t-1) = t - \frac{1}{6}.$$

Il reste à faire maintenant :

$$\|t^2 - p(t^2)\| = \sqrt{\int_0^1 \left(t^2 - t + \frac{1}{6} \right)^2 dt} = \frac{1}{\sqrt{180}} = \frac{\sqrt{5}}{30}.$$

c. Grâce à la **méthode 12.10**, on peut écrire : $\inf_{(a,b) \in \mathbb{R}^2} \int_0^1 (t^2 - a - bt)^2 dt = d(t^2, F) = \frac{\sqrt{5}}{30}$.

2. On reprend l'exemple de la **méthode 12.5**. Soit $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$. Expliquer pourquoi, si $M = (m_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbf{R})$, $\min_{M \in \mathcal{S}_n(\mathbf{R})} \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - m_{ij})^2$ existe et vaut :

$$\frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - a_{ji})^2.$$

On a vu dans l'exemple de la **méthode 12.5** que $\mathcal{S}_n(\mathbf{R})$ et $\mathcal{A}_n(\mathbf{R})$ sont supplémentaires orthogonaux

dans $\mathcal{M}_n(\mathbf{R})$ pour le produit scalaire canonique $\langle A, B \rangle = \text{Tr}(A^T B) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ji} b_{ji} \right)$.

$$\langle A, A \rangle = \text{Tr}(A^T A) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2, \text{ donc } \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - m_{ij})^2 = \|A - M\|^2.$$

On recherche donc le carré de la plus petite distance entre la matrice A et les matrices symétriques, donc le carré de la distance $d(A, \mathcal{S}_n(\mathbf{R}))$: c'est le carré de la distance entre A et son projeté orthogonal sur $\mathcal{S}_n(\mathbf{R})$ qui est, d'après le même exemple, $\frac{1}{2} (A + A^T)$.

Or : $A - \frac{1}{2} (A + A^T) = \frac{1}{2} (A - A^T) = (\frac{1}{2} (a_{ij} - a_{ji}))_{1 \leq i, j \leq n}$. On a donc bien :

$$\min_{M \in \mathcal{S}_n(\mathbf{R})} \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - m_{ij})^2 = \left\| \frac{1}{2} (A - A^T) \right\|^2 = \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - a_{ji})^2.$$

Énoncé des exercices

■ ???

Exercice 12.1 :

Exercice 12.2 :

Exercice 12.3 :

Exercice 12.4 :

Exercice 12.5 :

■ ???

Exercice 12.6 :

Exercice 12.7 :

■ ???

Exercice 12.8 :

Exercice 12.9 :

Exercice 12.10 :

Exercice 12.11 :

Exercice 12.12 :

■ ???

Exercice 12.13 :

Exercice 12.14 :

Exercice 12.15 :

Exercice 12.16 :

Exercice 12.17 :

Indications

- ___ Ex. 12.1 _____
- ___ Ex. 12.2 _____
- ___ Ex. 12.3 _____
- ___ Ex. 12.4 _____
- ___ Ex. 12.5 _____
- ___ Ex. 12.6 _____
- ___ Ex. 12.7 _____
- ___ Ex. 12.8 _____
- ___ Ex. 12.9 _____
- ___ Ex. 12.10 _____
- ___ Ex. 12.11 _____
- ___ Ex. 12.12 _____
- ___ Ex. 12.13 _____
- ___ Ex. 12.14 _____
- ___ Ex. 12.15 _____
- ___ Ex. 12.16 _____
- ___ Ex. 12.17 _____

Corrigé des exercices

____ **Exercice 12.1** _____
ZUM MACHEN

____ **Exercice 12.2** _____
ZUM MACHEN

____ **Exercice 12.3** _____
ZUM MACHEN

____ **Exercice 12.4** _____
ZUM MACHEN

____ **Exercice 12.5** _____
ZUM MACHEN

____ **Exercice 12.6** _____
ZUM MACHEN

____ **Exercice 12.7** _____
ZUM MACHEN

____ **Exercice 12.8** _____
ZUM MACHEN

____ **Exercice 12.9** _____
ZUM MACHEN

____ **Exercice 12.10** _____
ZUM MACHEN

____ **Exercice 12.11** _____
ZUM MACHEN

____ **Exercice 12.12** _____
ZUM MACHEN

____ **Exercice 12.13** _____
ZUM MACHEN

____ **Exercice 12.14** _____
ZUM MACHEN

____ **Exercice 12.15** _____
ZUM MACHEN

____ **Exercice 12.16** _____
ZUM MACHEN

____ **Exercice 12.17** _____
ZUM MACHEN

Chapitre 13

Fourier series

■ Objectifs

■ Les incontournables :

- ▶ Savoir calculer les coefficients de Fourier, en tenant compte de l'éventuelle parité.
- ▶ Utiliser le théorème de Dirichlet pour établir la convergence et déterminer la somme d'une série de Fourier.
- ▶ À partir du théorème de Dirichlet ou de la formule de Parseval, calculer une somme de série.

■ Et plus si affinités ...

- ▶ Interpréter dans une structure préhilbertienne les résultats sur les séries de Fourier.

Résumé de cours

Dans tout le chapitre, T désigne un réel strictement positif.

■ Fonctions continues et de classe \mathcal{C}^1 par morceaux

Ici a et b sont deux réels avec $a < b$.

Définition : Une fonction f définie sur le segment $[a, b]$, à valeurs dans \mathbf{R} est dite **continue par morceaux sur $[a, b]$** s'il existe une subdivision $a_0 = a < a_1 < \dots < a_n = b$ telle que la restriction de f à chaque sous-intervalle $]a_i, a_{i+1}[$ soit prolongeable en une fonction continue sur $[a_i, a_{i+1}]$.

Définition : Une fonction T -périodique est dite **continue par morceaux sur \mathbf{R}** si elle est continue par morceaux sur un intervalle de la forme $[a, a + T]$.

Théorème 13.1.— L'ensemble des fonctions à valeurs réelles, T -périodiques et continues par morceaux sur \mathbf{R} est un espace vectoriel sur \mathbf{R} .

Théorème 13.2.— Soit f une fonction à valeurs réelles, T -périodique et continue par morceaux sur \mathbf{R} .

Alors pour tout réel $\alpha \in \mathbf{R}$,
$$\int_{\alpha}^{\alpha+T} f(t) dt = \int_0^T f(t) dt.$$

Définition : Une fonction f définie sur le segment $[a, b]$, à valeurs dans \mathbf{R} est dite **de classe \mathcal{C}^1 par morceaux sur $[a, b]$** s'il existe une subdivision $a_0 = a < a_1 < \dots < a_n = b$ telle que la restriction de f à chaque sous-intervalle $]a_i, a_{i+1}[$ soit prolongeable en une fonction de classe \mathcal{C}^1 sur $[a_i, a_{i+1}]$.

Définition : Une fonction T -périodique est dite **de classe \mathcal{C}^1 par morceaux sur \mathbf{R}** si elle est de classe \mathcal{C}^1 par morceaux sur un intervalle de la forme $[a, a + T]$.

■ Coefficients et série de Fourier

Définition : Étant donné f , une fonction à valeurs réelles, T -périodique et continue par morceaux sur \mathbf{R} , on appelle **coefficients de Fourier** de f , les réels définis par (où $\omega = \frac{2\pi}{T}$, $\alpha \in \mathbf{R}$ et $n \in \mathbf{N}^*$) :

$$\begin{aligned} a_0(f) &= \frac{1}{T} \int_{\alpha}^{\alpha+T} f(t) dt \\ a_n(f) &= \frac{2}{T} \int_{\alpha}^{\alpha+T} f(t) \cos(n\omega t) dt . \\ b_n(f) &= \frac{2}{T} \int_{\alpha}^{\alpha+T} f(t) \sin(n\omega t) dt \end{aligned}$$

Définition : Pour tout réel t , la série :

$$a_0(f) + \sum_{n \geq 1} [a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)]$$

s'appelle la **série de Fourier** de la fonction f en t .

Remarque : Dans la pratique, selon l'intervalle sur lequel f est définie explicitement, on prend des valeurs particulières de α . Ainsi si f est explicite sur $[0, T]$, on prend $\alpha = 0$ et si f est explicite sur $[-T/2, T/2]$, on prend $\alpha = -T/2$. On peut aussi tenir compte de ce qui suit.

Proposition 13.3.— Si f est T -périodique, **continue par morceaux sur \mathbf{R}** et **paire**, alors pour $n \in \mathbf{N}^*$:

$$a_0(f) = \frac{2}{T} \int_0^{T/2} f(t) dt, \quad a_n(f) = \frac{4}{T} \int_0^{T/2} f(t) \cos(n\omega t) dt, \quad b_n(f) = 0.$$

Proposition 13.4.— Si f est T -périodique, **continue par morceaux sur \mathbf{R}** et **impaire**, alors pour $n \in \mathbf{N}^*$:

$$a_0(f) = 0, \quad a_n(f) = 0, \quad b_n(f) = \frac{4}{T} \int_0^{T/2} f(t) \sin(n\omega t) dt.$$

On note pour tout $N \in \mathbf{N}^*$:

$$\forall t \in \mathbf{R}, S_N(f)(t) = a_0(f) + \sum_{n=1}^N [a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)].$$

C'est la $N^{\text{ème}}$ **somme partielle de la série de Fourier** de f .

■ Interprétation géométrique

On note ici $\mathcal{C}_T(\mathbf{R})$ l'espace vectoriel des fonctions à valeurs réelles, continues sur \mathbf{R} et T -périodiques.

Théorème 13.5.— L'application

$$\Psi : \mathcal{C}_T(\mathbf{R}) \times \mathcal{C}_T(\mathbf{R}) \rightarrow \mathbf{R}, (f, g) \mapsto \frac{1}{T} \int_0^T f(t) g(t) dt$$

est un **produit scalaire** sur $\mathcal{C}_T(\mathbf{R})$.

La norme associée à ce produit scalaire est pour tout $f \in \mathcal{C}_T(\mathbf{R})$,

$$\|f\|^2 = \frac{1}{T} \int_0^T f^2(t) dt.$$

Proposition 13.6.— (i) Dans l'espace préhilbertien $\mathcal{C}_T(\mathbf{R})$, muni de Ψ , les fonctions $t \mapsto \cos(n\omega t)$ et $t \mapsto \sin(n\omega t)$ pour tout $n \in \mathbf{N}^*$, forment une famille orthogonale.
(ii) Soit $N \in \mathbf{N}^*$ et $f \in \mathcal{C}_T(\mathbf{R})$. La $N^{\text{ème}}$ somme partielle de la série de Fourier de f , $S_N(f)$, est la projection orthogonale de f sur l'espace vectoriel engendré par les fonctions $t \mapsto \cos(n\omega t)$ pour n décrivant $\llbracket 0, N \rrbracket$ et $t \mapsto \sin(n\omega t)$ pour n décrivant $\llbracket 1, N \rrbracket$.

■ Théorèmes de convergence

Théorème 13.7.— Formule de Parseval.

Soit f une fonction réelle, T -périodique et **continue par morceaux sur \mathbf{R}** .

Les séries $\sum_{n \geq 0} a_n(f)^2$ et $\sum_{n \geq 1} b_n(f)^2$ sont convergentes et on a :

$$\frac{1}{T} \int_0^T f^2(t) dt = a_0(f)^2 + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n(f)^2 + b_n(f)^2).$$

Théorème 13.8.— Théorème de Dirichlet.

Soit f une fonction réelle, T -périodique et **de classe \mathcal{C}^1 par morceaux sur \mathbf{R}** .

Alors, pour tout $t \in \mathbf{R}$, la série de Fourier de f au point t converge et on a :

$$a_0(f) + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) = \frac{1}{2} \lim_{h \rightarrow 0} (f(t+h) + f(t-h)).$$

Corollaire 13.9.— Théorème de Dirichlet pour une fonction continue.

Soit f une fonction réelle, T -périodique, **continue sur \mathbf{R}** et **de classe \mathcal{C}^1 par morceaux sur \mathbf{R}** .

Alors, pour tout $t \in \mathbf{R}$, la série de Fourier de f au point t converge et on a :

$$a_0(f) + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) = f(t).$$

Énoncé des exercices

■ ???

Exercice 13.1 :

Exercice 13.2 :

Exercice 13.3 :

Exercice 13.4 :

Exercice 13.5 :

■ ???

Exercice 13.6 :

Exercice 13.7 :

■ ???

Exercice 13.8 :

Exercice 13.9 :

Exercice 13.10 :

Exercice 13.11 :

Exercice 13.12 :

■ ???

Exercice 13.13 :

Exercice 13.14 :

Exercice 13.15 :

Exercice 13.16 :

Exercice 13.17 :

Indications

- ___ Ex. 13.1 _____
- ___ Ex. 13.2 _____
- ___ Ex. 13.3 _____
- ___ Ex. 13.4 _____
- ___ Ex. 13.5 _____
- ___ Ex. 13.6 _____
- ___ Ex. 13.7 _____
- ___ Ex. 13.8 _____
- ___ Ex. 13.9 _____
- ___ Ex. 13.10 _____
- ___ Ex. 13.11 _____
- ___ Ex. 13.12 _____
- ___ Ex. 13.13 _____
- ___ Ex. 13.14 _____
- ___ Ex. 13.15 _____
- ___ Ex. 13.16 _____
- ___ Ex. 13.17 _____

Corrigé des exercices

____ **Exercice 13.1** _____
ZUM MACHEN

____ **Exercice 13.2** _____
ZUM MACHEN

____ **Exercice 13.3** _____
ZUM MACHEN

____ **Exercice 13.4** _____
ZUM MACHEN

____ **Exercice 13.5** _____
ZUM MACHEN

____ **Exercice 13.6** _____
ZUM MACHEN

____ **Exercice 13.7** _____
ZUM MACHEN

____ **Exercice 13.8** _____
ZUM MACHEN

____ **Exercice 13.9** _____
ZUM MACHEN

____ **Exercice 13.10** _____
ZUM MACHEN

____ **Exercice 13.11** _____
ZUM MACHEN

____ **Exercice 13.12** _____
ZUM MACHEN

____ **Exercice 13.13** _____
ZUM MACHEN

____ **Exercice 13.14** _____
ZUM MACHEN

____ **Exercice 13.15** _____
ZUM MACHEN

____ **Exercice 13.16** _____
ZUM MACHEN

____ **Exercice 13.17** _____
ZUM MACHEN

Chapitre 14

Functions of several real variables

■ Objectifs

■ Les incontournables :

- ▶ étudier les notions de base : dérivées partielles, applications de classe \mathcal{C}^1 , différentielle, gradient, points critiques, dérivées partielles d'ordre 2 ;
- ▶ savoir calculer les dérivées et les dérivées partielles dans des cas très simples de compositions de fonctions \mathcal{C}^1 ;
- ▶ savoir appliquer les notions de base à l'étude de courbes du plan et de surfaces de l'espace ;
- ▶ chercher les extremums globaux sur un fermé borné d'une fonction numérique de classe \mathcal{C}^1 ;
- ▶ savoir montrer qu'une fonction est de classe \mathcal{C}^2 sur \mathbf{R}^2 ;
- ▶ énoncer le théorème de Schwarz ;
- ▶ être capable d'utiliser un changement de variables affine ou le passage en coordonnées polaires pour étudier des exemples d'équations aux dérivées partielles du premier ordre ;
- ▶ étudier quelques exemples d'équations aux dérivées partielles du second ordre.

■ Et plus si affinités ...

- ▶ connaître les propriétés du gradient ;
- ▶ étudier des courbes tracées sur une surface ;
- ▶ étudier les tangentes aux courbes régulières de classe \mathcal{C}^1 tracées sur une surface.
- ▶ savoir faire des changements de variables un peu « exotiques » sur une équation aux dérivées partielles ;

Résumé de cours

\mathbf{R}^p est muni de sa norme euclidienne usuelle, notée $\|\cdot\|$. On note par ailleurs $d(\vec{a}, \vec{x}) = \|\vec{a} - \vec{x}\|$.
On se limite en pratique au cas $p \leq 3$.

■ Boules, parties bornées et convexes

Définition : Boules ouvertes et fermées —. \vec{a} étant un élément de \mathbf{R}^p et r un réel positif, on appelle **boule fermée** (respectivement **boule ouverte**) de centre \vec{a} et de rayon r , et on note $\mathcal{B}_f(\vec{a}, r)$ (respectivement $\mathcal{B}_o(\vec{a}, r)$), l'ensemble des éléments \vec{x} de E tels que $d(\vec{x}, \vec{a}) \leq r$ (respectivement $d(\vec{x}, \vec{a}) < r$).

Remarque : L'ensemble $\mathcal{B}_f(\vec{a}, r) \setminus \mathcal{B}_o(\vec{a}, r) = \{\vec{x} \in E / d(\vec{x}, \vec{a}) = r\}$ est appelé la **sphère** de centre \vec{a} et de rayon r .

Exemple : Dans $(\mathbf{R}, |\cdot|)$, $\mathcal{B}_o(\vec{a}, r) =]\vec{a} - r, \vec{a} + r[$ et $\mathcal{B}_f(\vec{a}, r) = [\vec{a} - r, \vec{a} + r]$.

Définition : Partie bornée —. Une partie de \mathbf{R}^p est **bornée** si, et seulement si, elle est incluse dans une boule de centre $0_{\mathbf{R}^p}$.

■ Topologie de \mathbf{R}^p

Point intérieur —. Un élément \vec{x} de \mathbf{R}^p est un **point intérieur à la partie A** si, et seulement s'il existe une boule ouverte de centre \vec{x} incluse dans A , c'est-à-dire :

$$\exists r > 0, \quad \mathcal{B}_o(\vec{x}, r) \subset A.$$

On appelle **intérieur de A** (que l'on note dans cet ouvrage par commodité $\overset{\circ}{A}$) l'ensemble des points intérieurs de A .

Définition : Partie ouverte —.

Une partie A de \mathbf{R}^p est un **ouvert** (ou une **partie ouverte**) de \mathbf{R}^p si, et seulement si,
 $\forall \vec{x} \in A, \exists r > 0$ tel que $\mathcal{B}_o(\vec{x}, r) \subset A$.

Proposition 14.1.— Une boule ouverte de \mathbf{R}^p est un ouvert de \mathbf{R}^p .

Remarque : On démontre en exercice que la réunion d'une famille quelconque d'ouverts est un ouvert et que l'intersection d'une famille finie d'ouverts est un ouvert. Ces deux propriétés pourront vous être utiles à condition de les justifier bien entendu car hors programme en TSI2 dans le cours !

Définition : Point adhérent —. Un élément \vec{x} de E est **adhérent à la partie A** si, et seulement si, toute boule ouverte de centre \vec{x} a au moins un point commun avec A , c'est-à-dire :

$$\forall r > 0, \quad \mathcal{B}_o(\vec{x}, r) \cap A \neq \emptyset.$$

On appelle **adhérence de A** (que l'on note dans cet ouvrage par commodité $\text{adh}(A)$) l'ensemble des points adhérents de A .

Définition : Partie fermée —. Une partie A de E est un **fermé** (ou une **partie fermée**) de \mathbf{R}^p si, et seulement si, son complémentaire dans E est un ouvert de \mathbf{R}^p .

Proposition 14.2.— Une boule fermée et une sphère de \mathbf{R}^p sont des parties fermées de \mathbf{R}^p .

Remarque : La réunion d'une famille finie de fermés est un fermé et l'intersection d'une famille quelconque de fermés est un fermé. Encore une fois, ces deux propriétés utiles doivent être prouvables par vos soins en exercice car elles ne sont pas citées dans le cours.

Proposition 14.3.— Partie fermée et points adhérents —. Une partie est fermée si, et seulement si, elle contient tous ses points adhérents.

Définition : On appelle *frontière* de A l'ensemble $\text{adh}(A) \setminus \overset{\circ}{A}$. On le note souvent $\text{Fr}(A)$ ou δA .

■ Fonctions d'une partie de \mathbf{R}^p dans \mathbf{R}

Proposition 14.4.— Limite en un point —. Soit f une application d'une partie A de \mathbf{R}^p à valeurs dans \mathbf{R} et \vec{a} un élément de \mathbf{R}^p adhérent à A .

Étant donné un élément b de \mathbf{R} , on dit que f *admet b comme limite en \vec{a}* si :

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tel que : } \forall \vec{x} \in A, \|\vec{x} - \vec{a}\| \leq \delta \Rightarrow |f(\vec{x}) - b| \leq \varepsilon.$$

Le réel b est alors unique et on le note $b = \lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x})$ ou $\lim_{\vec{a}} f$.

Définition : Continuité, prolongement par continuité —. Lorsque \vec{a} appartient à A , si f admet une limite b en \vec{a} , alors $b = f(\vec{a})$ et f est dite *continue en \vec{a}* .

Si \vec{a} n'appartient pas à A et si f admet une limite b en \vec{a} , alors on dit que f est *prolongeable par continuité en \vec{a}* .

Proposition 14.5.— Limite d'une combinaison linéaire —.

Soit f et g deux applications de A dans \mathbf{R} admettant pour limites respectives b et c au point adhérent \vec{a} de A . Pour tout couple (α, β) de \mathbf{K}^2 , $\lim_{\vec{x} \rightarrow \vec{a}} (\alpha f + \beta g)(\vec{x}) = \alpha b + \beta c$.

Proposition 14.6.— Limite d'un produit —.

Soit f une application de A dans \mathbf{R} admettant b pour limite au point adhérent \vec{a} de A et g une application de A dans \mathbf{R} ayant α pour limite en \vec{a} . Le produit $g.f$ admet pour limite αb en \vec{a} .

Proposition 14.7.— Limite de l'inverse d'une fonction scalaire, d'un quotient —.

Étant donné une fonction g de A dans \mathbf{R} , de limite β en \vec{a} , avec $\beta \neq 0$, alors :

► $\exists r > 0, \forall \vec{x} \in \mathcal{B}_o(\vec{a}, r) \cap A, g(\vec{x}) \neq 0$.

► En posant $\mathcal{B}' = \mathcal{B}_o(\vec{a}, r) \cap A$, la fonction $\frac{1}{g}$ est définie sur \mathcal{B}' ; \vec{a} est un point adhérent de

\mathcal{B}' et $\frac{1}{g}$ admet pour limite $\frac{1}{\beta}b$ en \vec{a} .

■ Continuité des fonctions de \mathbf{R}^p dans \mathbf{R}

Définition : Fonction continue sur une partie A —. Une application définie sur une partie A de \mathbf{R}^p est *continue sur A* si, et seulement si, elle est continue en tout point de A .

Proposition 14.8.— Théorèmes généraux —. Utiles pour montrer la continuité d'une fonction :

- Si f est une application continue sur une partie A de \mathbf{R}^p et B une partie de \mathbf{R}^p incluse dans A , alors la restriction de f à B , $f|_B$, est continue sur B .
- Si f est une application continue sur une partie A de \mathbf{R}^p et g une application continue sur une partie B de \mathbf{R}^p contenant $f(A)$, alors $g \circ f$ est continue sur A .
- Si f et g sont deux applications continues sur une partie A de \mathbf{R}^p , si λ et μ sont deux scalaires quelconques, alors $\lambda f + \mu g$ est continue sur A .
- Si f et g sont deux applications continues sur une partie A de \mathbf{R}^p à valeurs numériques, alors $f.g$ est continue sur A .

Proposition 14.9.— Une fonction continue sur une partie fermée et bornée de \mathbf{R}^p et à valeurs dans \mathbf{R} , est bornée et atteint ses bornes.

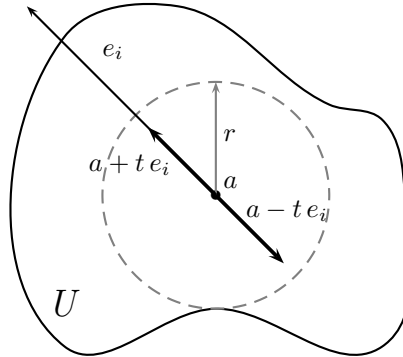
■ Fonctions de classe \mathcal{C}^1

Les fonctions considérées ici sont définies sur une partie ouverte U de \mathbf{R}^p et à valeurs dans \mathbf{R} .

Soit $(\vec{e}_1, \dots, \vec{e}_p)$ la base canonique de \mathbf{R}^p et $a = (a_1, \dots, a_p)$ un point de la partie ouverte U de \mathbf{R}^p . Comme U est un ouvert :

$$\exists r > 0 \text{ tel que } B(a, r) \subset U.$$

Pour $i \in \llbracket 1, p \rrbracket$, $\forall t \in]-r, r[$, $a + t\vec{e}_i \in U$, puisque : $\|(a + t\vec{e}_i) - a\| = |t| \|\vec{e}_i\| = |t| < r$.



On pose alors, pour $t \in]-r, r[$, $\varphi_i(t) = f(a + t\vec{e}_i)$.

Définition : Dérivées partielles d'ordre 1 en un point —. Si φ_i , fonction réelle de la variable réelle t , est dérivable en 0, on dit que f **admet une dérivée partielle d'ordre 1 en a par rapport à la $i^{\text{ème}}$ variable**.

Cette dérivée partielle est notée $\partial_i f(a)$ ou $\frac{\partial f}{\partial x_i}(a)$.

On a donc : $\partial_i f(a) = \varphi_i'(0)$ ou $\partial_i f(a) = \lim_{t \rightarrow 0} \frac{f(a + t\vec{e}_i) - f(a)}{t}$.

Plus précisément, dans le cas d'une fonction de deux variables :

$$\begin{aligned}\partial_1 f(a_1, a_2) &= \frac{\partial f}{\partial x}(a_1, a_2) = \lim_{t \rightarrow 0} \frac{f(a_1 + t, a_2) - f(a_1, a_2)}{t} \\ \partial_2 f(a_1, a_2) &= \frac{\partial f}{\partial y}(a_1, a_2) = \lim_{t \rightarrow 0} \frac{f(a_1, a_2 + t) - f(a_1, a_2)}{t}\end{aligned}$$

Définition : Fonctions dérivées partielles —. Si une fonction réelle f définie sur un ouvert U de \mathbf{R}^p admet une dérivée partielle à l'ordre 1 par rapport à la $i^{\text{ème}}$ coordonnée en tout point de U , on définit sur U la $i^{\text{ème}}$ **fonction dérivée partielle à l'ordre 1** :

$$\frac{\partial f}{\partial x_i} : \begin{cases} U \longrightarrow \mathbf{R} \\ x = (x_1, \dots, x_p) \mapsto \partial_i f(x) = \frac{\partial f}{\partial x_i}(x) \end{cases}$$

Dans la pratique, pour obtenir la fonction dérivée partielle par rapport à l'une des variables, on dérive f en considérant les autres variables comme des constantes.

Définition : Fonction de classe \mathcal{C}^1 —. Une fonction définie sur un ouvert U de \mathbf{R}^p est **de classe \mathcal{C}^1 sur U** si, et seulement si, ses dérivées partielles d'ordre 1 existent et sont continues sur U .

Exemple : Les deux (ou trois) applications coordonnées $\theta_i : (x_1, x_2, x_3) \mapsto x_i$ sont de classe \mathcal{C}^1 sur $\mathbf{R}^p = \mathbf{R}^2$ ou \mathbf{R}^3 , avec :

$$\forall (i, j) \in \llbracket 1, p \rrbracket, \partial_j \theta_i : (x_1, x_2, x_3) \mapsto \delta_{i,j} = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases} \quad (\text{symbole de Kronecker}).$$

Proposition 14.10. — Opérations sur les fonctions de classe \mathcal{C}^1 —. Si U est un ouvert de $\mathbf{R}^p = \mathbf{R}^2$ ou \mathbf{R}^3 et si f et g sont deux applications réelles de classe \mathcal{C}^1 sur U , alors :

- $f + g$ est de classe \mathcal{C}^1 sur U et, $\forall i \in \llbracket 1, p \rrbracket$, $\frac{\partial(f+g)}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i}$.
- $\forall \lambda \in \mathbf{R}$, λf est de classe \mathcal{C}^1 sur U et, $\forall i \in \llbracket 1, p \rrbracket$, $\frac{\partial(\lambda f)}{\partial x_i} = \lambda \frac{\partial f}{\partial x_i}$.
- fg est de classe \mathcal{C}^1 sur U et, $\forall i \in \llbracket 1, p \rrbracket$, $\frac{\partial(fg)}{\partial x_i} = \frac{\partial f}{\partial x_i} g + f \frac{\partial g}{\partial x_i}$.
- Si g ne s'annule pas sur U , $\frac{f}{g}$ est de classe \mathcal{C}^1 sur U et, $\forall i \in \llbracket 1, p \rrbracket$, $\frac{\partial\left(\frac{f}{g}\right)}{\partial x_i} = \frac{\frac{\partial f}{\partial x_i} g - f \frac{\partial g}{\partial x_i}}{g^2}$.
- Si φ est une fonction numérique de classe \mathcal{C}^1 sur un intervalle I tel que $f(U) \subset I$, alors $\varphi \circ f$ est \mathcal{C}^1 sur U et, $\forall i \in \llbracket 1, p \rrbracket$, $\frac{\partial(\varphi \circ f)}{\partial x_i} = \varphi' \circ f \times \frac{\partial f}{\partial x_i}$.

Exemple : Les fonctions polynômes de deux (ou trois) variables $(x_1, x_2, x_3) \mapsto \sum_{k=1}^m a_i x_1^{n_1} x_2^{n_2} x_3^{n_3}$

sont des applications de classe \mathcal{C}^1 sur \mathbf{R}^p . Si le dénominateur ne s'annule pas sur un ouvert U de \mathbf{R}^p , les fonctions quotients de polynômes de deux (ou trois) variables sont des applications de classe \mathcal{C}^1 sur U .

Théorème-Définition 14.11. — Développement limité à l'ordre 1 —. Si f est une fonction de classe \mathcal{C}^1 sur l'ouvert U de \mathbf{R}^2 et si $a = (a_1, a_2)$ est un point quelconque de U , alors :

$$\forall \vec{x} = (x_1, x_2) \in U, f(\vec{x}) = f(a) + (x_1 - a_1) \frac{\partial f}{\partial x_1}(a) + (x_2 - a_2) \frac{\partial f}{\partial x_2}(a) + \|\vec{x} - a\| \varepsilon(\vec{x}), \text{ avec } \lim_{\vec{x} \rightarrow a} \varepsilon(\vec{x}) = 0.$$

On dit que f admet en a un développement limité à l'ordre 1.

Proposition 14.12.— Une fonction de classe \mathcal{C}^1 sur U est continue sur U .

■ Règle de la chaîne

Proposition 14.13.— **Règle de la chaîne** —. Soit f une application de classe \mathcal{C}^1 sur une partie ouverte U de \mathbf{R}^p et x_1, \dots, x_p deux (ou trois) fonctions de classe \mathcal{C}^1 sur un intervalle I telles :

$$\forall t \in I, (x_1(t), \dots, x_p(t)) \in U.$$

$g : t \mapsto f(x_1(t), \dots, x_p(t))$ est de classe \mathcal{C}^1 sur I et

$$\forall t \in I, g'(t) = \frac{\partial f}{\partial x_1} x'_1(t) + \dots + \frac{\partial f}{\partial x_p} x'_p(t).$$

(les dérivées partielles étant évaluées en $(x_1(t), \dots, x_p(t))$.)

Interprétation géométrique. Si $\Gamma : I \rightarrow \mathbf{R}^p$, $t \mapsto (x_1(t), \dots, x_p(t))$ est un arc de classe \mathcal{C}^1 , alors le vecteur $t \mapsto (x'_1(t), \dots, x'_p(t))$ est la dérivée le long de Γ .

Corollaire 14.14.— **Dérivées partielles de $(u, v) \mapsto f(x(u, v), y(u, v))$** —. Soit f une fonction réelle de classe \mathcal{C}^1 sur un ouvert U de \mathbf{R}^2 et x et y deux fonctions de classe \mathcal{C}^1 sur un ouvert V de \mathbf{R}^2 telles que : $\forall (u, v) \in V, (x(u, v), y(u, v)) \in U$.

$g : (u, v) \mapsto f(x(u, v), y(u, v))$ est une fonction réelle de classe \mathcal{C}^1 sur V , avec :

$$\forall (u, v) \in V, \begin{cases} \frac{\partial g}{\partial u}(u, v) = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} \\ \frac{\partial g}{\partial v}(u, v) = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} \end{cases}.$$

(Les dérivées partielles de f sont évaluées en $(x(u, v), y(u, v))$, celles de x et y en (u, v) .)

Exemple : le cas particulier des coordonnées polaires.

$x : (r, \theta) \mapsto r \cos \theta$ et $y : (r, \theta) \mapsto r \sin \theta$ sont des fonctions réelles de classe \mathcal{C}^1 sur \mathbf{R}^2 ; si f est fonction réelle de classe \mathcal{C}^1 sur \mathbf{R}^2 , alors $g : (r, \theta) \mapsto f(r \cos \theta, r \sin \theta)$ est de classe \mathcal{C}^1 sur \mathbf{R}^2 , avec : $\forall (r, \theta) \in \mathbf{R}^2$,

$$\begin{cases} \frac{\partial g}{\partial r}(r, \theta) = \cos \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + \sin \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \\ \frac{\partial g}{\partial \theta}(r, \theta) = -r \sin \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + r \cos \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \end{cases}.$$

■ Gradient

Définition : Gradient de f en un point —. Soit f une application réelle de classe \mathcal{C}^1 sur une partie ouverte U de $\mathbf{R}^p = \mathbf{R}^2$ ou \mathbf{R}^3 et a un point de U , le **gradient de f au point a** est le vecteur

$$\nabla f(a) = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a) \left[\frac{\partial f}{\partial z}(a) \right] \right).$$

■ Applications géométriques

Définition : Soit U une partie ouverte de \mathbf{R}^2 , f une fonction réelle de classe \mathcal{C}^1 sur U et λ un réel, le système $\begin{cases} (x, y) \in U \\ f(x, y) = \lambda \end{cases}$ définit une courbe plane de \mathbf{R}^2 : **la ligne de niveau λ de f** .

Proposition 14.15.— Soit Γ la ligne de niveau 0 de f . Si $f(x_0, y_0) = 0$ et si $\nabla f(x_0, y_0) \neq (0, 0)$, alors le point $M_0(x_0, y_0)$ est un point régulier de Γ .

La tangente à Γ en $M_0(x_0, y_0)$ est la droite passant par M_0 et de vecteur normal $\nabla f(x_0, y_0)$, c'est-à-dire la droite d'équation :

$$(x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Corollaire 14.16.— En un point régulier $M_0(x_0, y_0)$ de la ligne de niveau Γ_λ de f , $\nabla f(x_0, y_0)$ est orthogonal à Γ_λ et orienté dans le sens des valeurs croissantes de f .

Définition : Soit U un ouvert de \mathbf{R}^3 et f une fonction réelle de classe \mathcal{C}^1 sur U , le système $\begin{cases} (x, y, z) \in U \\ f(x, y, z) = 0 \end{cases}$ définit une **surface de \mathbf{R}^3** .

Si $f(x_0, y_0, z_0) = 0$ et si $\nabla f(x_0, y_0, z_0) \neq (0, 0, 0)$, alors le point $M_0(x_0, y_0, z_0)$ est **un point régulier** de cette surface.

Si $\varphi : t \mapsto (x(t), y(t), z(t))$ est une application de classe \mathcal{C}^1 d'un intervalle I de \mathbf{R} , à valeurs dans U , telle que : $\forall t \in I, f(x(t), y(t), z(t)) = 0$, alors (I, φ) est **une courbe tracée sur cette surface**.

Exemple : Courbes coordonnées d'une surface d'équation $z = g(x, y)$: ce sont les courbes planes déterminées par l'intersection de cette surface avec l'un des plans d'équation $x = x_0$ ou $y = y_0$ ou $z = z_0$.

Par exemple, pour $x = 0$, $(\mathbf{R}, t \mapsto (0, t, g(0, t)))$ est une courbe coordonnée tracée sur cette surface.

Définition : Le plan tangent à la surface définie par $f(x, y, z) = 0$ en son point régulier $M_0(x_0, y_0, z_0)$ est le plan passant par M_0 et de vecteur normal $\nabla f(x_0, y_0, z_0)$, c'est-à-dire le plan d'équation :

$$(x - x_0) \frac{\partial f}{\partial x}(x_0, y_0, z_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0, z_0) + (z - z_0) \frac{\partial f}{\partial z}(x_0, y_0, z_0) = 0.$$

Remarque : La tangente en un point à une courbe régulière tracée sur une surface régulière est incluse dans le plan tangent en ce point à cette surface.

■ Dérivées partielles d'ordre 2

Définition : Dérivées partielles d'ordre 2 —. Soit f une fonction réelle de deux ou trois variables définie sur un ouvert U de \mathbf{R}^2 ou \mathbf{R}^3 , si la $i^{\text{ème}}$ dérivée partielle à l'ordre 1 de f admet une dérivée partielle à l'ordre 1 par rapport à la $j^{\text{ème}}$ coordonnée au point a de U , alors cette dérivée, notée

$\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$, est la dérivée partielle d'ordre 2 de f en a .

Si elle existe en tout a de U , on définit sur U la fonction dérivée partielle d'ordre 2 : $\frac{\partial^2 f}{\partial x_j \partial x_i}$.

Remarques : 1. L'ordre des variables au dénominateur est naturel si l'on veut bien se souvenir que $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$.

2. Si l'on dérive deux fois par rapport à la même variable, on note : $\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right) = \frac{\partial^2 f}{\partial x_i^2}$.

Définition : Fonction de plusieurs variables de classe \mathcal{C}^2 —. L'application f est **de classe \mathcal{C}^2 sur U** lorsqu'elle est de classe \mathcal{C}^1 sur U et lorsque ses fonctions dérivées partielles sont de classe \mathcal{C}^1 sur U .

Théorème 14.17.— Théorème d'Hermann Schwarz —.

Si f est une fonction de \mathbf{R}^p dans \mathbf{R} , de classe \mathcal{C}^2 sur une partie ouverte U , alors, en tout point a de U :

$$\forall (i, j) \in \llbracket 1, p \rrbracket^2, \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

■ **Extremums d'une fonction de \mathbf{R}^p dans \mathbf{R}**

Extremum global, extremum local : soit f une fonction réelle définie sur une partie Δ de \mathbf{R}^p .

▷ f admet un **maximum global** (respectivement un **minimum global**) au point a de Δ si :

$$\forall x \in \Delta, \quad f(x) \leq f(a) \text{ (respectivement } f(x) \geq f(a)).$$

▷ La fonction f admet un **maximum local** (respectivement un **minimum local**) au point a de U si :

$$\exists r > 0 \text{ tel que : } \forall x \in B(a, r) \cap U, \quad f(x) \leq f(a) \text{ (respectivement } f(x) \geq f(a)).$$

Le maximum (respectivement minimum) local ou global est le réel $f(a)$.

▷ La fonction f admet un **extremum local** au point a de U si f admet un minimum local ou un maximum local au point a . En l'absence de précision, lorsqu'on parlera d'extremum, il s'agira d'extremum local.

Définition : Point critique —. Un point en lequel le gradient de f s'annule est un **point critique** de f .

Théorème 14.18.— Condition nécessaire d'existence d'un extremum local sur un ouvert —.

Si une fonction réelle de classe \mathcal{C}^1 sur un ouvert U atteint un extremum local en un point a de U , alors a est un point critique.

Remarque : Sur un **ouvert**, les points en lesquels une fonction de classe \mathcal{C}^1 est susceptible d'atteindre un extremum seront à rechercher parmi ses points critiques sur cet ouvert.

Méthodes

■ Dérivées partielles d'ordre 1 et fonction de classe \mathcal{C}^1 sur un ouvert $U \subset \mathbf{R}^p$

□ **Méthode 14.1.—** Comment montrer qu'une application f définie sur un ouvert U de $\mathbf{R}^p = \mathbf{R}^2$ ou \mathbf{R}^3 est de classe \mathcal{C}^1 sur U

On utilisera en priorité les théorèmes généraux (propositions 14.10 et 14.13). Pour les rares points présentant une difficulté, on pourra faire une étude locale en revenant aux définitions. En résumé, on montre l'existence de $\frac{\partial f}{\partial x_i}$ pour tout $i \in \llbracket 1, p \rrbracket$ puis on montre la continuité de chacune de ces fonctions en appliquant la **méthode ??**.

Exemples : 1. Montrer que $f : (x, y) \mapsto \arctan\left(\frac{y}{x}\right)$ est de classe \mathcal{C}^1 sur $\mathbf{R}^2 \setminus \{(0, y), y \in \mathbf{R}\}$.

$(x, y) \mapsto \frac{y}{x}$ est \mathcal{C}^1 sur $\mathbf{R}^2 \setminus \{(0, y), y \in \mathbf{R}\}$ comme quotient de fonctions polynômes des deux variables x et y , celui du dénominateur ne s'annulant pas. $t \mapsto \arctan t$ est \mathcal{C}^1 sur \mathbf{R} , donc, par composition, f est de classe \mathcal{C}^1 sur $\mathbf{R}^2 \setminus \{(0, y), y \in \mathbf{R}\}$.

$$\forall (x, y) \in \mathbf{R}^2 \setminus \{(0, y), y \in \mathbf{R}\}, \frac{\partial f}{\partial x}(x, y) = \frac{-\frac{y}{x^2}}{1 + \frac{y^2}{x^2}} = \frac{-y}{x^2 + y^2} \text{ et } \frac{\partial f}{\partial y}(x, y) = \frac{\frac{1}{x}}{1 + \frac{y^2}{x^2}} = \frac{x}{x^2 + y^2}.$$

2. Montrer $f : \begin{cases} \mathbf{R}^2 & \rightarrow \\ (x, y) & \mapsto \begin{cases} \frac{2xy}{x^2+y^2}, \text{ si } (x, y) \neq (0, 0) \\ 0, \text{ si } (x, y) = (0, 0) \end{cases} \end{cases}$ est \mathcal{C}^1 sur $\mathbf{R}^2 \setminus \{(0, 0)\}$, que f admet

des dérivées partielles par rapport à x et par rapport à y en $(0, 0)$, mais n'est pas \mathcal{C}^1 en $(0, 0)$.

f , quotient de fonctions polynômes des deux variables x et y dont le dénominateur ne s'annule pas, est \mathcal{C}^1 sur $\mathbf{R}^2 \setminus \{(0, 0)\}$. Puis : $\forall t \neq 0, \frac{f(0+t, 0) - f(0, 0)}{t} = 0$, donc $\lim_{t \rightarrow 0} \frac{f(0+t, 0) - f(0, 0)}{t} = 0$:

f est dérivable par rapport à x en $(0, 0)$ et $\frac{\partial f}{\partial x}(0, 0) = 0$.

De même, f est dérivable par rapport à y en $(0, 0)$ et $\frac{\partial f}{\partial y}(0, 0) = 0$.

$\forall a \neq 0, f(a, a) = 1$, donc, si elle existe, la limite de f en $(0, 0)$ ne peut être égale à $0 = f(0, 0)$: f n'est pas continue en $(0, 0)$, donc ne peut être de classe \mathcal{C}^1 en ce point (cf **proposition 14.12**). Bien sûr, il est aussi possible de calculer l'une ou l'autre des dérivées partielles de f sur $\mathbf{R}^2 \setminus \{(0, 0)\}$ et de montrer qu'elle n'est pas continue en $(0, 0)$.

□ **Méthode 14.2.— Quand utiliser la règle de la chaîne**

- Dans le cas de $g : t \mapsto f(x_1(t), \dots, x_p(t))$: on vérifie préalablement que f est \mathcal{C}^1 sur un ouvert U de \mathbf{R}^p et que les fonctions d'une seule variable x_1, x_2, \dots, x_p sont \mathcal{C}^1 sur un intervalle I , avec, $\forall t \in I, (x_1(t), x_2(t), \dots, x_p(t)) \in U$.

g est alors de classe \mathcal{C}^1 sur I et, $\forall t \in I, g'(t) = \sum_{i=1}^p x'_i(t) \frac{\partial f}{\partial x_i}(x_1(t), \dots, x_p(t))$.

- Dans le cas de $h : (u, v) \mapsto f(x(u, v), y(u, v))$: on vérifie que f est \mathcal{C}^1 sur un ouvert U de \mathbf{R}^2 et que x et y sont deux fonctions de classe \mathcal{C}^1 sur un ouvert V de \mathbf{R}^2 , avec :

$\forall (u, v) \in V, (x(u, v), y(u, v)) \in U$, h est alors classe \mathcal{C}^1 sur V et, avec la règle de la chaîne, on obtient $\frac{\partial h}{\partial u}$ en fixant v et $\frac{\partial h}{\partial v}$ en fixant u .

Exemples : 1. Si f est une fonction réelle de classe \mathcal{C}^1 sur un ouvert U de \mathbf{R}^2 contenant au moins le cercle unité, montrer que la fonction $\varphi : t \mapsto f(\cos t, \sin t)$ est de classe \mathcal{C}^1 sur \mathbf{R} et calculer sa dérivée.

$x : t \mapsto \cos t$ et $y : t \mapsto \sin t$ sont de classe \mathcal{C}^1 sur \mathbf{R} et, pour tout réel t , $(\cos t, \sin t)$ appartient au cercle unité, donc à l'ouvert U . f étant de classe \mathcal{C}^1 sur un ouvert U , φ est \mathcal{C}^1 sur \mathbf{R} et la règle de la chaîne donne :

$$\forall t \in \mathbf{R}, \varphi'(t) = -\sin t \frac{\partial f}{\partial x}(\cos t, \sin t) + \cos t \frac{\partial f}{\partial y}(\cos t, \sin t).$$

2. Soient $f : (x, y) \mapsto x^3 + y$, $a : (u, v) \mapsto \cos(uv)$ et $b : (u, v) \mapsto u^2v$. Montrer que $g : (u, v) \mapsto f(a(u, v), b(u, v))$ est de classe \mathcal{C}^1 sur \mathbf{R}^2 et calculer ses dérivées partielles.

Déjà, f , a et b étant de classe \mathcal{C}^1 sur \mathbf{R}^2 , g est de classe \mathcal{C}^1 sur \mathbf{R}^2 .

Calculer $\frac{\partial g}{\partial u}(u, v)$ revient à calculer la dérivée en u de $g_1 : t \mapsto f(a(t, v), b(t, v))$, où v est supposée constante. Si $a_1 : t \mapsto a(t, v)$ et si $b_1 : t \mapsto b(t, v)$,

$$a'_1(u) = \frac{\partial a}{\partial u}(u, v) \text{ et } b'_1(u) = \frac{\partial b}{\partial u}(u, v).$$

En utilisant la règle de la chaîne, on peut retrouver la formule du cours (si, fait improbable, on l'avait oubliée) :

$$\frac{\partial g}{\partial u}(u, v) = g'_1(u) = \frac{\partial f}{\partial x}(a(u, v), b(u, v)) \frac{\partial a}{\partial u}(u, v) + \frac{\partial f}{\partial y}(a(u, v), b(u, v)) \frac{\partial b}{\partial u}(u, v).$$

Ici : $\forall (u, v) \in \mathbf{R}^2, \frac{\partial g}{\partial u}(u, v) = 3 \cos^2(uv) [-v \sin(uv)] + 1.2uv = 2uv - 3v \cos^2(uv) \sin(uv)$.

Et : $\frac{\partial g}{\partial v}(u, v) = 3 \cos^2(uv) [-u \sin(uv)] + 1.u^2 = u^2 - 3u \cos^2(uv) \sin(uv)$.

■ Gradient et applications

□ Méthode 14.3.— Comment trouver une équation d'une tangente à une courbe plane définie par $f(x, y) = 0$

Soit U un ouvert de \mathbf{R}^2 , f doit être une application de classe \mathcal{C}^1 sur U et soit Γ la courbe définie par $(x, y) \in U$ et $f(x, y) = 0$.

Si (x_0, y_0) est un couple de U tel que $f(x_0, y_0) = 0$ et si $\nabla f(x_0, y_0) \neq 0_{\mathbf{R}^2}$, alors $M_0(x_0, y_0)$ est un point régulier de Γ , $\nabla f(x_0, y_0)$ est un vecteur normal à Γ en M_0 et une équation cartésienne de la tangente à Γ en M_0 est :

$$(x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Exemple : Équation de la tangente en $M_0 = (x_0, y_0)$ à $\Gamma : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

La fonction $f : \begin{cases} \mathbf{R}^2 & \rightarrow & \mathbf{R} \\ (x, y) & \mapsto & \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \end{cases}$ est \mathcal{C}^1 sur $U = \mathbf{R}^2$ et :

$$\forall (x_0, y_0) \in \mathbf{R}^2 \setminus \{(0, 0)\}, \nabla f(x_0, y_0) = \left(\frac{2x_0}{a^2}, \frac{2y_0}{b^2} \right) \neq 0_{\mathbf{R}^2}.$$

Tout point $M_0 = (x_0, y_0)$ de Γ est régulier et l'équation cherchée est :

$$(x - x_0) \cdot \frac{2x_0}{a^2} + (y - y_0) \cdot \frac{2y_0}{b^2} = 0 \Leftrightarrow \frac{2x_0x}{a^2} + \frac{2y_0y}{b^2} - \frac{2x_0^2}{a^2} - \frac{2y_0^2}{b^2} = 0 \Leftrightarrow \frac{x_0x}{a^2} + \frac{y_0y}{b^2} = 1.$$

□ Méthode 14.4.— Comment déterminer l'équation cartésienne d'un plan tangent à une surface

Soit U un ouvert de \mathbf{R}^3 , f une application réelle de classe \mathcal{C}^1 sur U et S la surface de \mathbf{R}^3 définie par $(x, y, z) \in U$ et $f(x, y, z) = 0$.

Si (x_0, y_0, z_0) est un triplet de U tel que $f(x_0, y_0, z_0) = 0$ et si $\nabla f(x_0, y_0, z_0) \neq 0_{\mathbf{R}^3}$, alors $M_0(x_0, y_0, z_0)$ est un point régulier de S .

Le plan tangent à S en M_0 est le plan passant par M_0 et normal au vecteur $\nabla f(x_0, y_0, z_0)$.

Une équation cartésienne de ce plan est :

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0)(z - z_0) = 0.$$

Exemple : Déterminer les plans tangents à la surface S d'équation cartésienne $x^2 + y^2 + 2z^2 = 1$ et orthogonaux à la droite $D : x = \frac{y}{3} = -\frac{z}{2}$.

On remarque, tout d'abord, que le système $x = \frac{y}{3} = -\frac{z}{2}$ définit bien une droite vectorielle de \mathbf{R}^3 : c'est l'ensemble des (x, y, z) qui sont colinéaires à $(1, 3, -2)$, donc $\text{Vect}((1, 3, -2))$.

Un vecteur directeur de D est donc $e_1 + 3e_2 - 2e_3$.

Soit $A(a, b, c)$ un point de S et Π_A le plan tangent à S en A :

$$\begin{aligned} M(x, y, z) \in \Pi_A &\Leftrightarrow (\nabla f(A) | \overrightarrow{AM}) = 0 \\ &\Leftrightarrow 2a(x - a) + 2b(y - b) + 4c(z - c) = 0 \\ &\Leftrightarrow ax + by + 2cz = 1 \end{aligned}$$

Π_A est orthogonal à D si, et seulement si, $(a, b, 2c)$ et $(1, 3, -2)$ sont colinéaires, donc si, et seulement si, il existe $\lambda \in \mathbf{R}$ tel que :
$$\begin{cases} a = \lambda \\ b = 3\lambda \\ 2c = -2\lambda \end{cases}.$$

L'hypothèse $A \in S$ impose $a^2 + b^2 + 2c^2 = 1$, ce qui donne $\lambda^2 = \frac{1}{12}$.

Il existe deux plans tangents à S et orthogonaux à D . Ils ont pour équations cartésiennes :

$$\lambda x + 3\lambda y - 2\lambda z = 1 \quad \text{avec} \quad \lambda = \pm \frac{1}{2\sqrt{3}}.$$

❑ Méthode 14.5.— Comment trouver les extremums globaux d'une fonction f de classe \mathcal{C}^1 sur une partie fermée bornée

Soit U une partie de \mathbf{R}^2 fermée bornée et $f : U \subset \mathbf{R}^2 \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 sur U .

- Une fonction réelle continue est bornée et atteint ses bornes sur un fermé borné : on a ainsi l'existence des extremums globaux pour f de classe \mathcal{C}^1 sur U .
- On commence par chercher les extremums locaux sur l'intérieur de U , puis on s'occupera des extremums sur sa frontière.
- 1. Sur l'intérieur de U , qui est un ouvert, on détermine les points critiques de f en résolvant le système : $\nabla f(x, y) = (0, 0)$.
- 2. Si (a, b) est l'un de ces points critiques, on calcule $f(a + h, b + k) - f(a, b)$ et on cherche à en déterminer le signe pour (h, k) voisin de $(0, 0)$.
- 3. Si $f(a + h, b + k) - f(a, b)$ est de signe constant pour (h, k) dans un certain disque de centre $(0, 0)$, alors f atteint en (a, b) un extremum qui vaut $f(a, b)$; sinon, f ne présente pas d'extremum en (a, b) et (pour briller dans les salons) sachez que ce point est alors appelé un **point selle** ou un **point col** pour f .
- Il reste ensuite à étudier f sur la frontière de U , ce qui se ramène souvent à l'étude de plusieurs fonctions d'une seule variable sur différents intervalles.

Exemple : Soit $f : (x, y) \mapsto x^2 + y^2 - 2x - 4y$.

Trouver les extremums globaux de f sur le rectangle $[0, 3] \times [1, 5] = U$.

f est de classe \mathcal{C}^1 , donc continue, sur le fermé borné U : elle admet sur U un maximum global et un minimum global. Chacun d'eux est atteint soit en un point critique de l'intérieur $\overset{\circ}{U}$ de U , soit en un point de la frontière de U .

Recherche des extremums locaux sur $\overset{\circ}{U}$: $\nabla f(x, y) = (0, 0) \Leftrightarrow (2x - 2, 2y - 4) = (0, 0) :$
 $(1, 2)$ est le seul point critique de f sur $\overset{\circ}{U}$.

Pour tout (h, k) tel que $(1 + h, 2 + k) \in U$,

$f(1 + h, 2 + k) - f(1, 2) = (1 + h)^2 + (2 + k)^2 - 2(1 + h) - 4(2 + k) - (-5) = h^2 + k^2 \geq 0$, donc f atteint en $(1, 2)$ un minimum global sur U qui vaut -5 .

Il reste à trouver le maximum global qui est nécessairement sur la frontière de U , puisqu'il n'y a pas d'autre point critique dans $\overset{\circ}{U}$.

Recherche du maximum sur la frontière de U :

▷ Sur le segment $\{(x, 1), x \in [0, 3]\}$, $f(x, 1) = x^2 - 2x - 3$ et, sur $[0, 3]$, $x \mapsto x^2 - 2x - 3$ a 0 pour maximum en 3.

▷ Sur le segment $\{(x, 5), x \in [0, 3]\}$, $f(x, 5) = x^2 - 2x + 5$ et, sur $[0, 3]$, $x \mapsto x^2 - 2x + 5$ a 8 pour maximum en 3.

▷ Sur le segment $\{(0, y), y \in [1, 5]\}$, $f(0, y) = y^2 - 4y$ et, sur $[1, 5]$, $y \mapsto y^2 - 4y$ a 5 pour maximum en 5.

▷ Sur le segment $\{(3, y), y \in [1, 5]\}$, $f(3, y) = y^2 - 4y + 3$ et, sur $[1, 5]$, $y \mapsto y^2 - 4y + 3$ a 8 pour maximum en 5.

f atteint en $(3, 5)$ un maximum global sur U qui vaut 8.

■ Fonctions de classe \mathcal{C}^2 sur un ouvert U de \mathbf{R}^p

□ Méthode 14.6.— Comment montrer qu'une fonction n'est pas de classe \mathcal{C}^2

Il suffit d'utiliser la contraposée du théorème de Schwarz.

Exemple : Soit $f : (x, y) \mapsto \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & \text{si } (x, y) \neq (0, 0) \\ 0, & \text{si } (x, y) = (0, 0) \end{cases}$. Montrer que f est deux fois

dérivable par rapport à x et à y sur \mathbf{R}^2 tout entier, mais n'est pas de classe \mathcal{C}^2 sur \mathbf{R}^2 .

Sur $\mathbf{R}^2 \setminus \{(0, 0)\}$, f est \mathcal{C}^2 , comme quotient de fonctions polynômes, donc de fonctions \mathcal{C}^2 sur \mathbf{R}^2 , celle du dénominateur ne s'annulant pas sur $\mathbf{R}^2 \setminus \{(0, 0)\}$.

Comme $|f(x, y)| \leq |xy|$, il est clair que $f(x, y) \xrightarrow{(x, y) \rightarrow (0, 0)} 0$, donc f est continue en $(0, 0)$, et, par

suite, sur \mathbf{R}^2 . Les dérivées partielles en $(0, 0)$ sont nulles, car $t \mapsto f(t, 0)$ et $t \mapsto f(0, t)$ sont toutes les deux égales à la fonction nulle.

Soit $y \neq 0$, $\lim_{t \rightarrow 0} \frac{f(t, y) - f(0, y)}{t} = \lim_{t \rightarrow 0} \frac{y(t^2 - y^2)}{t^2 + y^2} = -y$, donc : $\frac{\partial f}{\partial x}(0, y) = -y$.

Soit $x \neq 0$, $\lim_{t \rightarrow 0} \frac{f(x, t) - f(x, 0)}{t} = \lim_{t \rightarrow 0} \frac{x(x^2 - t^2)}{x^2 + t^2} = x$, donc : $\frac{\partial f}{\partial y}(x, 0) = x$.

On a l'implication : $\forall t \neq 0, \frac{\frac{\partial f}{\partial y}(t, 0) - \frac{\partial f}{\partial y}(0, 0)}{t} = 1 \Rightarrow \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$.

Et, de même : $\forall t \neq 0, \frac{\frac{\partial f}{\partial x}(0, t) - \frac{\partial f}{\partial x}(0, 0)}{t} = -1 \Rightarrow \frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$.

Conclusion : f est deux fois dérivable par rapport à x et à y en $(0, 0)$, donc sur \mathbf{R}^2 tout entier, mais d'après le théorème de Schwarz, f n'est pas de classe \mathcal{C}^2 en $(0, 0)$.

■ Équations aux dérivées partielles du premier et du second ordre

On ne traite ici que le cas des fonctions de deux variables.

► Il s'agit d'abord d'équations aux dérivées partielles du premier ordre du type :

$$(E) \quad \alpha(x, y) \frac{\partial f}{\partial x}(x, y) + \beta(x, y) \frac{\partial f}{\partial y}(x, y) + \gamma(x, y) f(x, y) = \delta(x, y),$$

où α, β, γ et δ sont continues sur U et f est la fonction inconnue. La plupart du temps, il faut appliquer un changement de variables pour résoudre ce type d'équation (voir la **méthode 14.8** générale). Au programme officiel de TSI2, il faut que vous soyez au top dans le cas de changement par transformation affine, c'est-à-dire où l'on pose $u = ax + by$, $v = cx + dy$ (voir la **méthode 14.9**) et dans le cas d'un changement en coordonnées polaires (c'est la **méthode 14.10**). Maintenant, vous pouvez vous préparer éventuellement à des changements « un peu plus costauds », en étant guidé et dans les meilleurs concours.

► Il s'agit aussi d'équations aux dérivées partielles du second ordre du type :

$$(E) \quad \alpha(x, y) \frac{\partial^2 f}{\partial x^2}(x, y) + \beta(x, y) \frac{\partial^2 f}{\partial x \partial y}(x, y) + \gamma(x, y) \frac{\partial^2 f}{\partial y^2}(x, y) + \delta(x, y) f(x, y) = \epsilon(x, y),$$

où $\alpha, \beta, \gamma, \delta$ et ϵ sont continues sur U et f est la fonction inconnue. (Parfois dans (E), peuvent apparaître des termes du premier ordre en plus mais la méthode à utiliser ne changera pas!) On applique alors généralement là aussi un changement de variables et les remarques plus haut sur l'ordre un restent valables. Les calculs sont juste plus long. Attention ici comme la solution f est de classe \mathcal{C}^2 *a priori*, le théorème de Schwarz est valable et les deux dérivées croisées se confondent.

Remarque : On peut parfois trouver les solutions d'une équation aux dérivées partielles directement sans changement de variables (quand c'est simple!).

□ **Méthode 14.7.—** Comment résoudre une équation du type (1) : $\frac{\partial^2 f}{\partial x \partial y} = g$ ou

(2) : $\frac{\partial^2 f}{\partial x^2} = g$ ou (3) : $\frac{\partial^2 f}{\partial y^2} = g$, où g est une fonction définie sur \mathbf{R}^2

► Pour (1), on fixe par exemple x et on intègre par rapport à y .

On obtient $\frac{\partial f}{\partial x} = G(x, y) + A(x)$, où G est obtenu en primitivant g par rapport à y et A est une fonction de classe \mathcal{C}^1 sur \mathbf{R} . Il reste à primitiver $x \mapsto G(x, y) + A(x)$ (donc par rapport à x). On vous laisse terminer!

► Pour (2) et pour (3), il faut primitiver deux fois par rapport à une seule des variables. Quand g est fonction de f , on a une équation différentielle classique (de variable x ou de variable y). On la résout (les constantes dépendent de la variable non utilisée).

Exemple : Trouver les applications f de \mathbf{R}^2 dans \mathbf{R} , de classe \mathcal{C}^2 telles que : $\frac{\partial^2 f}{\partial x \partial y} = 0$ sur \mathbf{R}^2 .

Fixons y quelconque. La fonction $\Psi_y : x \mapsto \frac{\partial f}{\partial y}(x, y)$ a une dérivée première nulle et il existe α telle que : $\forall x \in \mathbf{R}, \Psi_y(x) = \alpha(y)$. Comme $y \mapsto \frac{\partial f}{\partial y}(0, y)$ est de classe \mathcal{C}^1 , $\alpha \in \mathcal{C}^1(\mathbf{R})$.

Fixons x maintenant. L'intégration de $\frac{\partial f}{\partial y}$ donne : $f(x, y) = A(y) + B(x)$, où A est une fonction de classe \mathcal{C}^2 sur \mathbf{R} . De plus, $B(x) = f(x, 0) - A(0)$ et donc B est aussi de classe \mathcal{C}^2 sur \mathbf{R} . La réciproque est immédiate.

□ Méthode 14.8.— Comment résoudre par changement de variables une équation aux dérivées partielles (E) d'inconnue f définie sur $U \subset \mathbf{R}^2$ et à valeurs dans \mathbf{R}

On utilise un changement de variable défini par $u = \phi_1(x, y)$ et $v = \phi_2(x, y)$. On suppose que f cherchée est de classe \mathcal{C}^1 (cas premier ordre) ou de classe \mathcal{C}^2 (cas second ordre, donc Schwarz est utilisable) et que ϕ_1 et ϕ_2 sont de classe \mathcal{C}^1 (cas premier ordre) ou de classe \mathcal{C}^2 (cas second ordre) sur U .

[1] On détermine $\Delta = \{(\phi_1(x, y), \phi_2(x, y)), \text{ où } (x, y) \in U\}$. On calcule les dérivées partielles premières (et aussi secondes si (E) est d'ordre 2) de f en fonction de celles de la nouvelle inconnue g définie pour tout $(x, y) \in U$ par : $g(u, v) = f(\phi_1(x, y), \phi_2(x, y))$. Pour cela, on applique la **proposition 14.14**. On remplace dans (E) les dérivées partielles de f par leur expression en fonction de celles de g . On obtient une nouvelle équation aux dérivées partielles (F). On résout alors (F) sur Δ . On en déduit les fonctions f solutions de (E) sur U .

□ Méthode 14.9.— Comment résoudre en utilisant un changement de variable par transformation affine (E) : $\lambda \frac{\partial f}{\partial x}(x, y) + \mu \frac{\partial f}{\partial y}(x, y) = \delta(x, y)$, où $(\lambda, \mu) \in \mathbf{R}^2$

On note $h : \mathbf{R}^2 \rightarrow \mathbf{R}^2$, $(x, y) \mapsto (ax + by, cx + dy)$ et g telle que : $g \circ h = f$.

On impose à h d'être bijective. En effet, $h \in \mathcal{L}(\mathbf{R}^2)$ et de matrice : $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Comme son déterminant est $ad - bc$, on en déduit que h est bijective si, et seulement si, $ad - bc \neq 0$, ce que l'on suppose dans la suite. On peut donc écrire que : $g = f \circ h^{-1}$ car h^{-1} est parfaitement définie. Il en est alors de même de g .

On conclut : $ad - bc \neq 0 \Rightarrow \exists g, \forall (x, y) \in \mathbf{R}^2, f(x, y) = g(ax + by, cx + dy)$.

Remarquons ensuite que h et donc h^{-1} sont de classe \mathcal{C}^1 sur \mathbf{R}^2 .

Posons $u = ax + by$ et $v = cx + dy$. On a les formules :

$$\begin{cases} \frac{\partial f}{\partial x} = \frac{\partial g}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial g}{\partial v} \frac{\partial v}{\partial x} \\ \frac{\partial f}{\partial y} = \frac{\partial g}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial g}{\partial v} \frac{\partial v}{\partial y} \end{cases} \Rightarrow \begin{cases} \frac{\partial f}{\partial x} = a \frac{\partial g}{\partial u} + c \frac{\partial g}{\partial v} \\ \frac{\partial f}{\partial y} = b \frac{\partial g}{\partial u} + d \frac{\partial g}{\partial v} \end{cases}.$$

Il reste à terminer en choisissant a, b, c et d tels que l'on ait $\frac{\partial g}{\partial u} = 0$ ou $\frac{\partial g}{\partial v} = 0$.

Exemple : Résoudre (E) : $\forall (x, y) \in \mathbf{R}^2, \frac{\partial f}{\partial x}(x, y) + 2 \frac{\partial f}{\partial y}(x, y) = x$.

On remplace $\frac{\partial f}{\partial x}$ par $a \frac{\partial g}{\partial u} + c \frac{\partial g}{\partial v}$ et $\frac{\partial f}{\partial y}$ par $b \frac{\partial g}{\partial u} + d \frac{\partial g}{\partial v}$. On obtient : $(a + 2b) \frac{\partial g}{\partial u} + (c + 2d) \frac{\partial g}{\partial v} = 0$.

Il suffit de prendre (par exemple) $a = 1, b = 0, c = -2$ et $d = 1$. On utilise donc le changement de variables défini par $(u = x, v = -2x + y)$. u et v sont exprimés en fonction de x et y par des fonctions de classe \mathcal{C}^1 sur \mathbf{R}^2 et $h : \mathbf{R}^2 \rightarrow \mathbf{R}^2, (x, y) \mapsto (x, -2x + y)$ est une bijection (application linéaire de déterminant non nul).

f est solution de (E) si, et seulement si, $\forall (u, v) \in \mathbf{R}^2, \frac{\partial g}{\partial u} = u$.

On en déduit que : $\forall (u, v) \in \mathbf{R}^2, g(u, v) = \frac{u^2}{2} + A(v)$ où A est une fonction de classe \mathcal{C}^1 sur \mathbf{R} .

La solution générale de l'équation (E) est alors la fonction de classe \mathcal{C}^1 sur \mathbf{R}^2 :

$$\forall (x, y) \in \mathbf{R}^2, f(x, y) = \frac{x^2}{2} + A(-2x + y)$$

Remarque : On peut étendre à des équations du type : $\lambda \frac{\partial^2 f}{\partial x^2} + \mu \frac{\partial^2 f}{\partial x \partial y} + \nu \frac{\partial^2 f}{\partial y^2} = 0$.

Exemple : Ici $c \in \mathbf{R}_+^*$ fixé. Déterminer $f : \mathbf{R}^2 \mapsto \mathbf{R}$ de classe \mathcal{C}^2 vérifiant l'équation (E) dite des cordes vibrantes : $\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial y^2} = 0$ à l'aide du changement $u = x - cy, v = x + cy$.

On pose $g(u, v) = f(x, y)$. Comme $x = \frac{1}{2}(u + v)$ et $y = \frac{1}{2c}(v - u)$, la fonction $h : (x, y) \mapsto (u, v)$ est bijective et de classe \mathcal{C}^2 . La composée $g = f \circ h^{-1}$ est bijective et est de classe \mathcal{C}^2 .

On a rapidement :

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial v} \text{ et } \frac{\partial f}{\partial y} = -c \frac{\partial g}{\partial u} + c \frac{\partial g}{\partial v}.$$

On applique encore le même théorème du cours et le théorème de Schwarz pour aboutir :

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial u}{\partial x} \frac{\partial}{\partial u} \left(\frac{\partial g}{\partial u} + \frac{\partial g}{\partial v} \right) + \frac{\partial v}{\partial x} \frac{\partial}{\partial v} \left(\frac{\partial g}{\partial u} + \frac{\partial g}{\partial v} \right) = \frac{\partial^2 g}{\partial u^2} + 2 \frac{\partial^2 g}{\partial u \partial v} + \frac{\partial^2 g}{\partial v^2}.$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial u}{\partial y} \frac{\partial}{\partial u} \left(-c \frac{\partial g}{\partial u} + c \frac{\partial g}{\partial v} \right) + \frac{\partial v}{\partial y} \frac{\partial}{\partial v} \left(-c \frac{\partial g}{\partial u} + c \frac{\partial g}{\partial v} \right) = c^2 \frac{\partial^2 g}{\partial u^2} - 2c^2 \frac{\partial^2 g}{\partial u \partial v} + c^2 \frac{\partial^2 g}{\partial v^2}.$$

Il reste à remplacer et on obtient :

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial y^2} = 0 \Rightarrow 4 \frac{\partial^2 g}{\partial u \partial v} = 0 \Rightarrow \frac{\partial^2 g}{\partial u \partial v} = 0.$$

On montre (voir l'exemple qui suit la **méthode 14.7**), qu'il existe alors deux fonctions A et B de classe \mathcal{C}^2 sur \mathbf{R} telles que $g(u, v) = A(u) + B(v)$.

Finalement : $\forall (x, y) \in \mathbf{R}^2, f(x, y) = A(x + cy) + B(x - cy)$.

❑ Méthode 14.10.— Comment résoudre une équation aux dérivées partielles (E) en utilisant le changement de variables en coordonnées polaires

On suppose (E) d'inconnue $f : (x, y) \mapsto f(x, y)$ définie sur $U \subset \mathbf{R}^2$.

[1] On pose donc $x = r \cos \theta$, $y = r \sin \theta$. On commence par trouver le domaine Δ , où évolue (r, θ) quand (x, y) parcourt U . On pose alors $g(r, \theta) = f(x, y)$.

Attention au point O qui ne possède pas de coordonnées polaires. On supposera généralement $r > 0$ dans les calculs et on étendra le résultat en $(0, 0)$ par continuité. Dans le cas où (E) possède le groupement $-y \frac{\partial f}{\partial x} +$

$x \frac{\partial f}{\partial y}$ ou $x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y}$, on peut directement remplacer la première quantité par

$\frac{\partial g}{\partial \theta}$ ou la seconde par $r \frac{\partial g}{\partial r}$, ce qui permet d'aller plus vite. D'ailleurs la présence de tels groupements justifie de passer en coordonnées polaires ! Sinon

$$\begin{cases} \frac{\partial g}{\partial r}(r, \theta) &= \cos \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + \sin \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \\ \frac{\partial g}{\partial \theta}(r, \theta) &= -r \sin \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + r \cos \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \end{cases}$$

donne par inversion ($r > 0$) :

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) &= \cos \theta \frac{\partial g}{\partial r}(r, \theta) - \frac{1}{r} \sin \theta \frac{\partial g}{\partial \theta}(r, \theta) \\ \frac{\partial f}{\partial y}(x, y) &= \sin \theta \frac{\partial g}{\partial r}(r, \theta) + \frac{1}{r} \cos \theta \frac{\partial g}{\partial \theta}(r, \theta) \end{cases}.$$

Puis on remplace dans E. Il reste à trouver g et à remplacer r et θ par x et y . Pour r , on utilise $r = \sqrt{x^2 + y^2}$ en restreignant $r > 0$. Pour θ , cela dépend de l'intervalle retenu. Par exemple, $\theta = \arctan \frac{y}{x}$ en restreignant $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ ou

alors $\theta = 2 \arctan \frac{y}{x + \sqrt{x^2 + y^2}}$ en restreignant $\theta \in]-\pi, \pi[$.

Exemple : Posons $U = \{(x, y) \in \mathbf{R}^2, x > 0\} = \mathbf{R}_+^* \times \mathbf{R}$.

En usant d'un changement en coordonnées polaires, résoudre :

$$(F) : \forall (x, y) \in U, -y \frac{\partial f}{\partial x}(x, y) + x \frac{\partial f}{\partial y}(x, y) = \sqrt{x^2 + y^2}.$$

x et y sont donc exprimés en fonction de r et θ par des fonctions de classe \mathcal{C}^1 sur la partie ouverte $V =]0, +\infty[\times]-\frac{\pi}{2}, \frac{\pi}{2}[$ et l'application φ définie sur $V =]0, +\infty[\times]-\frac{\pi}{2}, \frac{\pi}{2}[$ par l'égalité $\varphi(r, \theta) = (r \cos \theta, r \sin \theta)$ est une bijection de V dans U . On définit g en posant : $g = f \circ \varphi$.

On va exprimer les dérivées partielles de g en fonction de celles de f en utilisant la **méthode 14.10**.

La fonction f est solution de l'équation (F) si, et seulement si,

$$\forall (r, \theta) \in V, -r \sin \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + r \cos \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) = r \iff \frac{\partial g}{\partial \theta}(r, \theta) = r.$$

On en déduit : $\forall (r, \theta) \in V, g(r, \theta) = r\theta + A(r)$, où A est une fonction de classe \mathcal{C}^1 sur \mathbf{R}_+^* .

La solution générale de l'équation (F) est alors :

$$\forall (x, y) \in U, f(x, y) = \sqrt{x^2 + y^2} \arctan \left(\frac{y}{x} \right) + A \left(\sqrt{x^2 + y^2} \right).$$

Énoncé des exercices

■ ???

Exercice 14.1 :

Exercice 14.2 :

Exercice 14.3 :

Exercice 14.4 :

Exercice 14.5 :

■ ???

Exercice 14.6 :

Exercice 14.7 :

■ ???

Exercice 14.8 :

Exercice 14.9 :

Exercice 14.10 :

Exercice 14.11 :

Exercice 14.12 :

■ ???

Exercice 14.13 :

Exercice 14.14 :

Exercice 14.15 :

Exercice 14.16 :

Exercice 14.17 :

Indications

- ___ Ex. 14.1 _____
- ___ Ex. 14.2 _____
- ___ Ex. 14.3 _____
- ___ Ex. 14.4 _____
- ___ Ex. 14.5 _____
- ___ Ex. 14.6 _____
- ___ Ex. 14.7 _____
- ___ Ex. 14.8 _____
- ___ Ex. 14.9 _____
- ___ Ex. 14.10 _____
- ___ Ex. 14.11 _____
- ___ Ex. 14.12 _____
- ___ Ex. 14.13 _____
- ___ Ex. 14.14 _____
- ___ Ex. 14.15 _____
- ___ Ex. 14.16 _____
- ___ Ex. 14.17 _____

Corrigé des exercices

____ **Exercice 14.1** _____
ZUM MACHEN

____ **Exercice 14.2** _____
ZUM MACHEN

____ **Exercice 14.3** _____
ZUM MACHEN

____ **Exercice 14.4** _____
ZUM MACHEN

____ **Exercice 14.5** _____
ZUM MACHEN

____ **Exercice 14.6** _____
ZUM MACHEN

____ **Exercice 14.7** _____
ZUM MACHEN

____ **Exercice 14.8** _____
ZUM MACHEN

____ **Exercice 14.9** _____
ZUM MACHEN

____ **Exercice 14.10** _____
ZUM MACHEN

____ **Exercice 14.11** _____
ZUM MACHEN

____ **Exercice 14.12** _____
ZUM MACHEN

____ **Exercice 14.13** _____
ZUM MACHEN

____ **Exercice 14.14** _____
ZUM MACHEN

____ **Exercice 14.15** _____
ZUM MACHEN

____ **Exercice 14.16** _____
ZUM MACHEN

____ **Exercice 14.17** _____
ZUM MACHEN

Chapitre 15

Probability of countable universes

■ Objectifs

■ Les incontournables :

- ▶ Savoir reconnaître un ensemble dénombrable;
- ▶ étendre les compétences acquises en 1TSI sur les espaces probabilisés finis aux espaces probabilisés dénombrables :
 - ▷ le conditionnement et ses grandes formules : probabilités composées, probabilités totales, formule de Bayes;
 - ▷ indépendance de deux événements, indépendance mutuelle d'une famille finie d'événements.

■ Et plus si affinités ...

- ▶ Étendre la formule des probabilités totales à une suite d'événements deux à deux incompatibles dont la somme des probabilités vaut 1 ;
- ▶ être capable de produire un exemple où l'indépendance deux à deux n'entraîne pas l'indépendance mutuelle.

Résumé de cours

■ Ensembles dénombrables

Définition : Un ensemble A est **fini** si et seulement s'il existe un entier naturel N tel qu'il existe une bijection de A dans $\llbracket 1, N \rrbracket$. Dans ce cas, N est le cardinal de A : $\text{Card } A = N$.

Un ensemble A est **dénombrable** si et seulement s'il existe une bijection de A dans \mathbf{N} . Son cardinal est alors infini : $\text{Card } A = +\infty$.

Un ensemble sera **au plus dénombrable** si et seulement s'il est fini ou dénombrable.

Remarque : Un ensemble est dénombrable si et seulement s'il peut s'écrire en extension sous la forme $\{x_n, n \in \mathbf{N}\}$, autrement dit si l'on peut numéroter tous ses éléments à l'aide d'entiers naturels.

1. Un ensemble fini peut s'écrire en extension $\{x_n, n \in \llbracket 1, N \rrbracket\}$.

2. Si A est dénombrable et si φ est une bijection de A dans \mathbf{N} , alors chaque élément a de A peut se noter x_n , avec $n = \varphi(a)$.

Proposition 15.1.— \mathbf{N} et \mathbf{Z} sont des ensembles dénombrables.

■ Espace probabilisé

Définition : Une expérience aléatoire est une expérience dont on peut prédire avec certitude le résultat. L'étude d'une expérience aléatoire commence par la description des résultats possibles, appelés éventualités. L'ensemble des éventualités est appelé univers des possibles ou, plus simplement, univers. On le note généralement Ω .

Jusqu'à présent, on se limitait à un univers de cardinal fini. Désormais, il est dénombrable (donc peut être infini).

Exemple : Par exemple, on jette une pièce et on s'intéresse au rang de la première appartition de « pile ». Ici $\Omega = \mathbf{N}^*$.

Rappelons que $\mathcal{P}(\Omega)$ désigne l'ensemble des parties d'un ensemble Ω .

Définition : Un événement A est un ensemble d'éventualités. Il s'agit donc d'une partie de Ω . À la suite d'une expérience aléatoire, on dit que l'événement A est réalisé si et seulement si l'éventualité ω résultant de cette expérience est élément de A .

Vocabulaire : L'événement $\bar{A} = \Omega \setminus A$ est l'événement contraire de l'événement A , Ω est l'événement certain, \emptyset est l'événement impossible.

Lien avec les opérations ensemblistes

- L'événement (A ou B), appelé **disjonction** de A et B , est modélisé par $A \cup B$.
- L'événement (A et B), appelé **conjonction** de A et B , est modélisé par $A \cap B$.
- Le fait que la réalisation de l'événement A entraîne celle de B ($A \Rightarrow B$) se traduit par $A \subset B$.
- $\bigcup_{n=0}^{+\infty} A_n$ est l'événement qui se réalise si au moins un des A_n se réalise.

- $\bigcap_{n=0}^{+\infty} A_n$ est l'événement qui se réalise si tous les A_n se réalisent simultanément.

Événements incompatibles, système complet d'événements

- Deux événements sont **incompatibles** lorsque leur intersection est l'événement impossible.
- Une famille finie ou dénombrable $(A_i)_{i \in J}$ d'événements forme un système complet d'événements si les A_i sont deux à deux incompatibles et recouvrent Ω :

$$\bigcap_{i \in J} A_i = \Omega \text{ et pour tout couple } (i, j) \in J^2, \text{ avec } i \neq j, A_i \cap A_j = \emptyset.$$

On a les lois de Morgan pour tout $n \in \mathbb{N}^*$, $\overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \overline{A_i}$ et $\overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \overline{A_i}$.

De même, on a : $\overline{\bigcup_{n=0}^{+\infty} B_n} = \bigcap_{n=0}^{+\infty} \overline{B_n}$ et que : $\overline{\bigcap_{n=0}^{+\infty} B_n} = \bigcup_{n=0}^{+\infty} \overline{B_n}$. La preuve est dans l'exercice 01.

Mise en œuvre : exercice 01 et exercice 02.

Définition : On appelle **probabilité sur** Ω toute application P de $\mathcal{P}(\Omega)$ dans $[0, 1]$ vérifiant :

- $P(\Omega) = 1$;
- (σ -additivité) Pour toute suite $(A_n)_{n \in \mathbf{N}}$ d'événements deux à deux incompatibles de \mathcal{A} ,

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n).$$

Le couple (Ω, P) est alors un **espace probabilisé**.

Remarques : 1. Si la suite d'événements deux à deux incompatibles (A_0, A_1, \dots, A_n) est finie, la propriété est encore vraie.

Il suffit de considérer cette suite comme infinie, en posant : $\forall j > n, A_j = \emptyset$.

2. Une façon de construire une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ sera de considérer une suite $(p_n)_{n \in \mathbf{N}}$ de nombres positifs telle que la série $\sum p_n$ soit convergente et de somme 1.

Si $\Omega = \{x_n, n \in \mathbf{N}\}$ alors on posera alors :

$$\forall n \in \mathbf{N}, P(\{x_n\}) = p_n.$$

On définit ensuite la probabilité d'un événement A par : $P(A) = \sum_{x \in A} P(\{x\})$, la somme considérée pouvant être la somme d'une série.

Proposition 15.2.— Propriétés principales des probabilités—.

1. Si $A \in \mathcal{P}(\Omega)$, $P(\overline{A}) = 1 - P(A)$. En conséquence, $P(\emptyset) = 0$.
2. *Monotonie* : Si $A \subset B$: $P(A) \leq P(B)$.
3. Si $(A, B) \in (\mathcal{P}(\Omega))^2$, alors : $P(A) + P(B) = P(A \cup B) + P(A \cap B)$.

Mise en œuvre : exercice 03 à exercice 05.

■ Conditionnement et indépendance

Soit (Ω, P) un espace probabilisé.

Proposition 15.3.— Si B est un événement tel que $P(B) > 0$, alors l'application P_B définie sur $\mathcal{P}(\Omega)$ par l'égalité :

$$\forall A \in \mathcal{P}(\Omega), P_B(A) = \frac{P(A \cap B)}{P(B)}$$

est une probabilité sur Ω .

Définition : Si A et B sont deux événements tels que $P(B) > 0$, la **probabilité conditionnelle de A sachant B** est :

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

Notation : on note aussi couramment cette probabilité $P(A|B)$.

Proposition 15.4.— **Formules des probabilités composées** —.

► $\forall (A, B) \in (\mathcal{P}(\Omega))^2$,

si $P(B) > 0$, $P(A \cap B) = P(B)P_B(A)$ et si $P(A) > 0$, $P(A \cap B) = P(A)P_A(B)$.

► Si, pour $n \geq 2$, (A_1, A_2, \dots, A_n) est une famille d'événements de l'univers Ω telle que l'on ait :

$P\left(\bigcap_{i=1}^{n-1} A_i\right) \neq \emptyset$, alors :

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \dots P_{\bigcap_{i=1}^{n-1} A_i}(A_n).$$

Remarques : 1. La probabilité de réalisation simultanée de deux événements est le produit de la probabilité de réalisation de l'un d'entre eux par la probabilité de réalisation de l'autre sachant que le premier est réalisé.

2. Lorsque A_n est le dernier événement d'une suite décroissante $A_n \subset A_{n-1} \subset \dots \subset A_2 \subset A_1$, la formule précédente donne un moyen souvent efficace pour calculer $P(A_n)$:

$$P(A_n) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \dots P_{A_{n-1}}(A_n).$$

Définition : Un **système complet dénombrable d'événements** est une famille $(A_n)_{n \in \mathbf{N}}$ d'événements deux à deux incompatibles et dont l'union est égale à Ω (c'est-à-dire qu'il y a toujours exactement l'un des A_n qui se réalise).

Proposition 15.5.— **Formules des probabilités totales** —. Si $(A_n)_{n \in \mathbf{N}}$ est un système complet dénombrable d'événements de l'univers Ω , alors :

$$\text{pour tout événement } B \text{ de } \mathcal{P}(\Omega), P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n),$$

avec la convention $P_{A_n}(B)P(A_n) = 0$ lorsque $P(A_n) = 0$.

Remarques : 1. Ce résultat est valable pour un système complet fini d'événements et un cas fréquent d'application de ce théorème est celui où le système complet d'événements est (A, \bar{A}) .

2. La formule reste valable dans le cas d'une suite $(A_n)_{n \in \mathbf{N}}$ d'événements deux à deux incompatibles tels que $\sum_{n=0}^{+\infty} P(A_n) = 1$.

Proposition 15.6.— Formules de Bayes (ou probabilité des causes) —.

► Si A et B sont deux événements de probabilités non nulles, alors :

$$P_B(A) = \frac{P(A)P_A(B)}{P(B)}.$$

► Si $(A_n)_{n \in \mathbf{N}}$ est un système complet dénombrable d'événements de probabilités non nulles et si B est un événement de probabilité non nulle, alors :

$$\forall j \in \mathbf{N}, P_B(A_j) = \frac{P(A_j)P_{A_j}(B)}{\sum_{n=0}^{+\infty} P(A_n)P_{A_n}(B)}.$$

Définition : Deux événements A et B sont **indépendants** lorsque : $P(A \cap B) = P(A)P(B)$.

Remarques : 1. Dans ce cas, si, par exemple, $P(B) > 0$, alors $P_B(A) = \frac{P(A \cap B)}{P(B)} = P(A)$, ce qui signifie que la réalisation de B n'influe pas sur la probabilité de réalisation de A . Réciproquement, si $P(B) > 0$ et si $P_B(A) = P(A)$, alors $P(A \cap B) = P(A)P(B)$ et A et B sont indépendants.

2. Ne pas confondre indépendance et incompatibilité : deux événements incompatibles sont en fait généralement dépendants, puisque la réalisation de l'un interdit la réalisation de l'autre.

3. On montre facilement (faites-le !) que A et B sont indépendants si et seulement si A et \bar{B} ou \bar{A} et B ou \bar{A} et \bar{B} sont indépendants.

Définition : Les événements d'une famille finie $(A_i)_{1 \leq i \leq n}$ sont :

- **deux à deux indépendants** lorsque : $\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i \neq j \Rightarrow P(A_i \cap A_j) = P(A_i)P(A_j))$;
- **mutuellement indépendants** lorsque, pour toute partie J non vide de $\llbracket 1, n \rrbracket$:

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i).$$

Remarque : Des événements mutuellement indépendants sont deux à deux indépendants, mais la réciproque est fausse lorsque leur nombre est supérieur ou égal à 3.

Mise en œuvre : exercice 06 à exercice 11.

Méthodes

■ Espaces probabilisés

□ Méthode 15.1.— Comment implanter une probabilité sur un univers dénombrable

Soit $\Omega = \{x_n, n \in \mathbf{N}\}$ un ensemble infini dénombrable.

Pour construire une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, on admet qu'il suffit :

- de rechercher une suite $(p_n)_{n \in \mathbf{N}}$ de nombres positifs telle que la série :

$$\sum p_n$$

soit convergente et de somme 1 ;

- de poser alors :

$$\forall n \in \mathbf{N}, P(\{x_n\}) = p_n$$

et

$$\forall A \in \mathcal{P}(\Omega), P(A) = \sum_{x \in A} P(\{x\}).$$

Exemple : On pose :

$$\forall n \in \mathbf{N}, P(n) = \frac{1}{2^{n+1}}.$$

Montrer que l'on définit ainsi une probabilité sur $(\mathbf{N}, \mathcal{P}(\mathbf{N}))$ et calculer la probabilité de l'événement :

$$B = \{n \in \mathbf{N}, n \geq 10\}.$$

Il s'agit de vérifier :

$$\sum_{n=0}^{+\infty} P(n) = 1.$$

Ici, la série géométrique de premier terme $\frac{1}{2}$ et de raison $\frac{1}{2}$ est convergente et a pour somme

$$\frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1 : P \text{ définit bien une probabilité sur } (\mathbf{N}, \mathcal{P}(\mathbf{N})) :$$

$$P(B) = \sum_{n=10}^{+\infty} P(n) = \frac{\left(\frac{1}{2}\right)^{11}}{1 - \frac{1}{2}} = \frac{1}{2^{10}}.$$

■ Conditionnement et indépendance

□ Méthode 15.2.— Comment utiliser la formule de la probabilité conditionnelle et la formule des probabilités composées

- Il faut d'abord être capable de faire la différence entre une probabilité conditionnée par un événement et la probabilité d'une intersection d'événements, d'où une analyse fine de l'énoncé.
- Ensuite, on met en œuvre l'une des formules équivalentes :

$$\text{si } P(B) > 0, P_B(A) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A \cap B) = P(B)P_B(A).$$

- On peut être amené à utiliser la formule généralisée des probabilités composées :

$$\text{si } P\left(\bigcap_{i=1}^{n-1} A_i\right) \neq \emptyset, \text{ alors } P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \dots P_{\bigcap_{i=1}^{n-1} A_i}(A_n).$$

Exemple : Un rat se trouve dans un labyrinthe face à quatre portes dont une seule conduit à la sortie. Chaque fois qu'il choisit une mauvaise porte, le rat reçoit une légère décharge électrique et revient à son point de départ. On s'intéresse au nombre d'essais utilisés pour trouver la bonne porte. On envisage successivement trois hypothèses.

a. Le rat a une mémoire parfaite. À chaque nouvel essai, il évite toutes les mauvaises portes choisies précédemment et choisit au hasard parmi les restantes.

b. Le rat a une mémoire immédiate. À chaque nouvel essai, il évite la mauvaise porte de l'essai précédent et choisit au hasard parmi les trois autres.

c. Le rat n'a pas de mémoire. Il choisit à chaque essai de façon équiprobable l'une des portes. Pour chacune de ces hypothèses, définir l'univers et calculer la probabilité de chaque événement élémentaire.

Dans le chapitre suivant, la rédaction de la solution d'un tel exercice sera simplifiée par l'usage du vocabulaire des variables aléatoires.

Pour l'instant, si on note A_k l'événement : « au k -ième essai, le rat trouve la bonne porte », l'événement B_n : « le rat utilise n essais pour trouver la bonne porte » s'écrit :

$$B_n = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_{n-1}} \cap A_n.$$

Avec une mémoire parfaite : l'univers est $\{1, 2, 3, 4\}$.

$$P(B_1) = \frac{1}{4}; P(B_2) = P(\overline{A_1}) P_{\overline{A_1}}(A_2) = \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}; \text{ il faut ensuite utiliser la formule généralisée}$$

$$\text{des probabilités composées : } P(B_3) = P(\overline{A_1}) P_{\overline{A_1}}(\overline{A_2}) P_{\overline{A_1} \cap \overline{A_2}}(A_3) = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{4} \text{ et } P(B_4) = \frac{1}{4}.$$

Avec une mémoire immédiate : l'univers est alors \mathbf{N}^* . Avec les mêmes notations que ci-dessus :

$$P(B_n) = P(\overline{A_1}) P_{\overline{A_1}}(\overline{A_2}) \dots P_{\overline{A_{n-2}}}(\overline{A_{n-1}}) P_{\overline{A_{n-1}}}(A_n) = \frac{3}{4} \left(\frac{2}{3}\right)^{n-2} \frac{1}{3}.$$

Sans mémoire : l'univers est encore \mathbf{N}^* . L'absence de mémoire du rat vaut hypothèse d'indépendance mutuelle des différents événements, d'où :

$$P(B_n) = P(\overline{A_1}) P(\overline{A_2}) \dots P(\overline{A_{n-1}}) P(A_n) = \left(\frac{3}{4}\right)^{n-1} \frac{1}{4}.$$

❏ Méthode 15.3.— Comment utiliser la formule des probabilités totales

Pour calculer la probabilité d'un événement B avec la formule des probabilités totales, il faut :

- ▶ Mettre en évidence un système $(A_n)_{n \in \mathbf{N}}$ complet d'événements, fini ou dénombrable, de l'univers ; on le reconnaît par le fait que ces événements sont exclusifs et que tout événement élémentaire est réalisé pour un seul des A_n ;
- ▶ Disposer ensuite des probabilités de B conditionnées par les A_n : $P_{A_n}(B)$, et des probabilités des A_n ;
- ▶ Enfin, mettre en œuvre la formule :
$$P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n) ;$$
- ▶ Une occasion fréquente d'utilisation de cette formule se trouve dans la formule de Bayes.

Exemple : On choisit un nombre N au hasard entre 1 et 4, puis un nombre k au hasard entre 1 et N . Quelle est la probabilité de B_j : « le dernier nombre obtenu est j », $j \in \{1, 2, 3, 4\}$?

L'univers est l'ensemble $\Omega = \{(N, k), 1 \leq k \leq N \leq 4\}$.

Notons A_N l'événement : « le premier nombre obtenu est N », (A_1, A_2, A_3, A_4) forme un système complet d'événements de l'univers.

$$\triangleright P(B_1) = \sum_{N=1}^4 P_{A_N}(B_1)P(A_N) = 1 \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} + \frac{1}{3} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{25}{48}.$$

$$\triangleright P(B_2) = \sum_{N=1}^4 P_{A_N}(B_2)P(A_N) = 0 \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} + \frac{1}{3} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{13}{48}.$$

$$\triangleright P(B_3) = \sum_{N=1}^4 P_{A_N}(B_3)P(A_N) = 0 \times \frac{1}{4} + 0 \times \frac{1}{4} + \frac{1}{3} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{7}{48}.$$

$$\triangleright \text{Enfin, } P(B_4) = P_{A_4}(B_4)P(A_4) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16} = \frac{3}{48}.$$

Comment retrouver ce dernier résultat autrement ?

❏ Méthode 15.4.— Quand et comment utiliser la formule de Bayes

Il s'agit de calculer $P_B(A)$ à partir d'un énoncé donnant les moyens d'accéder à $P(A)$, $P_A(B)$ et $P(B)$.

- ▶ On utilise alors la formule : $P_B(A) = \frac{P_A(B)P(A)}{P(B)} ;$
- ▶ Si $A = A_j$ est l'un des événements d'un système complet d'événements $(A_n)_{n \in \mathbf{N}}$, on met en œuvre la formule :
$$P_B(A_j) = \frac{P(A_j)P_{A_j}(B)}{\sum_{n=0}^{+\infty} P(A_n)P_{A_n}(B)}.$$

Exemple : Un test sanguin a une probabilité de 0,95 de détecter un certain virus lorsque celui-ci est effectivement présent. Il donne néanmoins un faux résultat positif pour 1% des personnes non

infectées. Si 0,5% de la population est porteuse du virus, quelle est la probabilité qu'une personne ait le virus sachant qu'elle a un test positif ?

Notons V : « la personne testée a le virus » et T : « la personne testée a un test positif ».
On cherche $P_T(V)$. Or on sait que : $P(V) = 0,005$, $P_V(T) = 0,95$, $P_{\overline{V}}(T) = 0,01$.

$$\begin{aligned} P_T(V) &= \frac{P(T \cap V)}{P(T)} = \frac{P_V(T)P(V)}{P_V(T)P(V) + P_{\overline{V}}(T)P(\overline{V})} \\ &= \frac{0,95 \times 0,005}{0,95 \times 0,005 + 0,01 \times 0,995} = 0,323 \end{aligned}$$

Le test n'est pas fiable : si la personne présente un test positif, la probabilité qu'elle ne soit pas porteuse du virus est deux fois plus élevée que celle qu'elle le soit !



Méthode 15.5.— Quand et comment utiliser ou mettre en évidence l'indépendance de deux événements

Deux événements A et B sont indépendants si :

- ▶ $P(A \cap B) = P(A)P(B)$;
- ▶ Avec $P(B) > 0$, $P_B(A) = P(A)$; avec $P(A) > 0$, $P_A(B) = P(B)$;
- ▶ Si l'énoncé le suggère et on a alors droit aux formules précédentes.

Exemple : Montrer qu'un événement presque impossible (de probabilité 0) est indépendant de tout autre événement.

Soit A un événement presque impossible et B un événement quelconque, par la croissance d'une probabilité : puisque $A \cap B \subset A$, $P(A \cap B) \leq P(A)$, donc $P(A \cap B) = 0$ et $P(A \cap B) = P(A)P(B)$.

Un événement presque impossible est indépendant de tout autre événement.

Remarque : l'indépendance de deux événements n'est pas intrinsèque à ces événements, mais dépend de la probabilité choisie, comme le montre l'exemple suivant :

On lance deux fois une pièce. A est l'événement « obtenir pile au premier lancer » et B est l'événement « obtenir deux fois le même résultat. » L'univers est $\Omega = \{PP, PF, FP, FF\}$.

▷ **Si la pièce est non faussée :** $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$ et $P(A \cap B) = \frac{1}{4} = P(A)P(B)$. A et B sont indépendants pour la probabilité uniforme.

▷ **Si la pièce est faussée** de sorte que la probabilité d'obtenir pile à un lancer soit $\frac{3}{5}$.

$$P(A) = \frac{3}{5}, P(B) = \left(\frac{2}{5}\right)^2 + \left(\frac{3}{5}\right)^2 = \frac{13}{25} \text{ et } P(A \cap B) = \left(\frac{3}{5}\right)^2 = \frac{9}{25} \neq P(A)P(B).$$

Donc A et B ne sont pas indépendants pour cette probabilité.

□ Méthode 15.6.— Quand et comment utiliser ou mettre en évidence l'indépendance mutuelle d'une famille d'événements

Les événements de la famille finie $(A_i)_{1 \leq i \leq n}$ sont mutuellement indépendants si :

- Pour toute partie J non vide de $\llbracket 1, n \rrbracket$: $P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$;
- Si l'énoncé considère une suite d'épreuves indépendantes, alors toute suite $(A_i)_{i \geq 1}$ telle que la réalisation de chaque A_i est déterminée uniquement par le résultat de la i -ème épreuve est une suite indépendante d'événements.

Exemple : On suppose que les événements A_1, \dots, A_5 sont mutuellement indépendants. Donner une expression plus simple des probabilités :

$$P_{A_3}(A_1), P_{A_2 \cap A_4}(A_3), P_{A_2}(A_1 \cap A_3), P_{A_3 \cap A_4 \cap A_5}(A_1 \cap A_2).$$

▷ A_1 et A_3 sont indépendants (l'indépendance mutuelle entraîne l'indépendance deux à deux)
Et donc : $P_{A_3}(A_1) = P(A_1)$.

En utilisant l'associativité de l'intersection :

$$\triangleright P_{A_2 \cap A_4}(A_3) = \frac{P(A_3 \cap A_2 \cap A_4)}{P(A_2 \cap A_4)} = \frac{P(A_3)P(A_2)P(A_4)}{P(A_2)P(A_4)} = P(A_3).$$

$$\triangleright P_{A_2}(A_1 \cap A_3) = \frac{P(A_2 \cap A_1 \cap A_3)}{P(A_2)} = \frac{P(A_2)P(A_1)P(A_3)}{P(A_2)} = P(A_1)P(A_3).$$

$$\triangleright P_{A_3 \cap A_4 \cap A_5}(A_1 \cap A_2) = \frac{P(A_3 \cap A_4 \cap A_5 \cap A_1 \cap A_2)}{P(A_3 \cap A_4 \cap A_5)} = \frac{\prod_{i=1}^5 P(A_i)}{P(A_3)P(A_4)P(A_5)} = P(A_1)P(A_2).$$

On pourrait établir des formules plus générales, laissées au lecteur, comme l'on dit !

Énoncé des exercices

■ ???

Exercice 15.1 :

Exercice 15.2 :

Exercice 15.3 :

Exercice 15.4 :

Exercice 15.5 :

■ ???

Exercice 15.6 :

Exercice 15.7 :

■ ???

Exercice 15.8 :

Exercice 15.9 :

Exercice 15.10 :

Exercice 15.11 :

Exercice 15.12 :

■ ???

Exercice 15.13 :

Exercice 15.14 :

Exercice 15.15 :

Exercice 15.16 :

Exercice 15.17 :

Indications

- ___ Ex. 15.1 _____
- ___ Ex. 15.2 _____
- ___ Ex. 15.3 _____
- ___ Ex. 15.4 _____
- ___ Ex. 15.5 _____
- ___ Ex. 15.6 _____
- ___ Ex. 15.7 _____
- ___ Ex. 15.8 _____
- ___ Ex. 15.9 _____
- ___ Ex. 15.10 _____
- ___ Ex. 15.11 _____
- ___ Ex. 15.12 _____
- ___ Ex. 15.13 _____
- ___ Ex. 15.14 _____
- ___ Ex. 15.15 _____
- ___ Ex. 15.16 _____
- ___ Ex. 15.17 _____

Corrigé des exercices

- ____ **Exercice 15.1** _____
ZUM MACHEN
- ____ **Exercice 15.2** _____
ZUM MACHEN
- ____ **Exercice 15.3** _____
ZUM MACHEN
- ____ **Exercice 15.4** _____
ZUM MACHEN
- ____ **Exercice 15.5** _____
ZUM MACHEN
- ____ **Exercice 15.6** _____
ZUM MACHEN
- ____ **Exercice 15.7** _____
ZUM MACHEN
- ____ **Exercice 15.8** _____
ZUM MACHEN
- ____ **Exercice 15.9** _____
ZUM MACHEN
- ____ **Exercice 15.10** _____
ZUM MACHEN
- ____ **Exercice 15.11** _____
ZUM MACHEN
- ____ **Exercice 15.12** _____
ZUM MACHEN
- ____ **Exercice 15.13** _____
ZUM MACHEN
- ____ **Exercice 15.14** _____
ZUM MACHEN
- ____ **Exercice 15.15** _____
ZUM MACHEN
- ____ **Exercice 15.16** _____
ZUM MACHEN
- ____ **Exercice 15.17** _____
ZUM MACHEN

Chapitre 16

Discrete Random Variables

■ Objectifs

■ Les incontournables :

- ▶ consolider les connaissances sur les variables aléatoires finies et explorer la notion de variable aléatoire à image dénombrable ;
- ▶ étendre aux variables aléatoires discrètes les concepts d'espérance et de variance ;
- ▶ connaître l'inégalité de Bienaymé-Tchebychev ;
- ▶ étudier les couples de variables aléatoires discrètes, en particulier indépendantes ;
- ▶ appliquer ces nouveaux concepts aux lois discrètes usuelles : loi géométrique et loi de Poisson ;
- ▶ étudier l'approximation de la loi binomiale par la loi de Poisson.

■ Et plus si affinités ...

- ▶ avoir quelques idées sur les suites de variables aléatoires mutuellement indépendantes ;
- ▶ utiliser les inégalités sur les couples de variables aléatoires : encadrement du coefficient de corrélation linéaire ou inégalité de Cauchy-Schwarz ;

Résumé de cours

■ Variable aléatoire discrète et loi de probabilité associée

Définition : Une *variable aléatoire discrète* X est une application définie sur Ω (univers fini ou dénombrable) et à valeurs réelles. On peut préciser que son image $X(\Omega)$ est finie ou dénombrable (au plus dénombrable).

Notations :

- Si $x \in X(\Omega)$, l'image réciproque par X de x est $X^{-1}(x) = \{\omega \in \Omega \mid X(\omega) = x\}$. On la note abusivement $\{X = x\}$ ou $(X = x)$.
- Si U est une partie de $X(\Omega)$, l'image réciproque de U par X est

$$X^{-1}(U) = \{\omega \in \Omega \mid X(\omega) \in U\} = \bigcup_{x \in U} (X = x).$$

On la note $\{X \in U\}$ ou $(X \in U)$.

- Si $x \in X(\Omega)$, $P(X = x)$ est la probabilité de $(X \in \{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$.
- Si U est une partie de $X(\Omega)$, on : $P(X \in U) = \sum_{x \in U} P(X = x)$ (cette somme étant éventuellement

celle d'une série). Il suffit donc de connaître tous les $P(X = x)$ pour $x \in X(\Omega)$ pour pouvoir calculer, au moins de façon théorique, la probabilité de $(X \in U)$, pour n'importe quelle partie U de $X(\Omega)$. D'où la définition suivante :

Définition : Soit X une variable aléatoire discrète telle que $X(\Omega) = \{x_n\}_{n \in I}$, où $I \subset \mathbf{N}$.

La loi de probabilité de X est la suite $(p_n)_{n \in I}$, où, pour tout $n \in I$, $p_n = P(X = x_n)$.

Remarque : Alors : $\forall n \in I, p_n = P(X = x_n) \in [0, 1]$ et $\sum_{n \in I} p_n = \sum_{n \in I} P(X = x_n) = 1$.

Proposition 16.1.— Si une variable aléatoire discrète X prend ses valeurs dans $\{x_n; n \in I\}$, les x_n étant distincts, et si $(p_n)_{n \in I}$ est une suite de réels positifs vérifiant $\sum_{n \in I} p_n = 1$, alors il existe une probabilité P sur (Ω, \mathcal{A}) telle que, pour tout $n \in I$, $P(X = x_n) = p_n$.

Définition : La *fonction de répartition* F_X d'une variable aléatoire discrète *réelle* X sur (Ω, \mathcal{A}) est la fonction définie sur \mathbf{R} par :

$$\forall x \in \mathbf{R}, F_X(x) = P(X \leq x).$$

Proposition 16.2.— La fonction de répartition F_X d'une variable aléatoire discrète réelle X est une fonction positive, croissante, de limites 0 en $-\infty$ et 1 en $+\infty$.

Définition : Soit ϕ une application réelle définie sur un intervalle I contenant $X(\Omega)$, alors $\phi \circ X = \phi(X)$ est une variable aléatoire réelle sur Ω , appelée **image de X par ϕ** .

■ Espérance d'une variable aléatoire discrète

Définition : La variable aléatoire *réelle* discrète X à valeurs dans un ensemble dénombrable $\{x_n; n \geq 0\}$ est dite **d'espérance finie** si la série $\sum x_n P(X = x_n)$ est absolument convergente ; si tel est le cas, on appelle **espérance de X** le réel

$$E(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n).$$

Remarques : 1. Si $X(\Omega)$ est fini, l'espérance de X est toujours définie.

Si $X(\Omega) = \{x_n; n \geq 0\}$ est dénombrable, mais contenu dans \mathbf{R}^+ ou dans \mathbf{R}^- , l'absolue convergence équivaut à la convergence, donc la vérification se fait en même temps que le calcul de l'espérance. Le cas critique est celui de $X(\Omega)$ dénombrable avec une infinité de valeurs positives et une infinité de valeurs négatives.

2. L'espérance peut toujours être considérée comme une moyenne pondérée, étendue à une famille éventuellement dénombrable de valeurs.

3. On admet que la somme $\sum_{n=0}^{+\infty} x_n P(X = x_n)$ ne dépend pas de l'ordre d'énumération.

4. Une variable aléatoire discrète bornée est d'espérance finie.

5. La variable aléatoire X est **centrée** si elle est d'espérance finie avec $E(X) = 0$.

Proposition 16.3.— Théorème du transfert —. Si X est une variable aléatoire et f une application à valeurs réelles définie sur l'image $X(\Omega) = \{x_n, n \in \mathbf{N}\}$ de X , alors $f(X)$ est d'espérance finie si et seulement si la série $\sum_{n \geq 0} P(X = x_n) f(x_n)$ converge absolument. Dans ce cas, on a :

$$E(f(X)) = \sum_{n=0}^{+\infty} P(X = x_n) f(x_n).$$

Proposition 16.4.— Linéarité, positivité et croissance —. Soient X et Y deux variables aléatoires réelles discrètes.

- $\forall \lambda \in \mathbf{R}, E(\lambda X + Y) = \lambda E(X) + E(Y)$: l'espérance est linéaire.
- Si $X \geq 0$, c'est-à-dire si X ne prend que des valeurs positives, on a : $E(X) \geq 0$.

Conséquence : si X est une variable aléatoire discrète d'espérance finie, la variable aléatoire $X - E(X)$ est centrée.

■ Variance, écart-type d'une variable aléatoire discrète réelle

Théorème-Définition 16.5.— Formule de Huyghens-Koenig —. Si la variable aléatoire X^2 est d'espérance finie, alors X est elle-même d'espérance finie.

Si X^2 est d'espérance finie, on appelle **variance de X** le réel

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

L'écart-type de X est le réel noté $\sigma(X)$ défini par : $\sigma(X) = \sqrt{V(X)}$.

Remarque : La variance est donc indifféremment « la moyenne des carrés des écarts à la moyenne » ou « la moyenne des carrés moins le carré de la moyenne ».

Proposition 16.6.—

$$\forall (a, b) \in \mathbf{R}^2, V(aX + b) = a^2 V(X).$$

Remarque : Une variable aléatoire admettant une variance est *centrée réduite* lorsque $E(X) = 0$ et $V(X) = 1$. Par exemple, si $V(X) \neq 0$, $\frac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Corollaire 16.7.— **Inégalité de Bienaymé-Tchebychev** —.

Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie, alors :

$$\forall \varepsilon > 0, P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}.$$

Remarques : 1. L'inégalité de Bienaymé-Tchebychev permet de comprendre ce que mesure la variance : pour $\varepsilon > 0$ fixé, la probabilité que l'écart entre X et $E(X)$ soit supérieur à ε est d'autant plus petite que $V(X)$ est faible. La variance donne donc une indication de la dispersion de X autour de son espérance, c'est-à-dire sa plus ou moins forte tendance à s'écarter de sa moyenne. L'écart-type, qui mesure aussi la dispersion de X , présente l'intérêt de s'exprimer dans la même unité que les valeurs prises par la variable aléatoire X .

2. On passe souvent à l'événement contraire :

$$\forall \varepsilon > 0, P(|X - E(X)| < \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}.$$

■ Lois usuelles

Loi géométrique

Définition : Soit $p \in]0, 1[$, la variable aléatoire X suit la loi géométrique de paramètre p si et seulement si :

$$P(X = 0) = 0 \text{ et } \forall k \in \mathbf{N}^*, P(X = k) = p(1 - p)^{k-1}.$$

Notation : $X \hookrightarrow \mathcal{G}(p)$.

Remarque : La loi géométrique peut être interprétée comme rang du premier succès dans une suite illimitée d'épreuves de Bernoulli indépendantes et de même paramètre p .

Proposition 16.8.— Si $X \hookrightarrow \mathcal{G}(p)$, alors son espérance et sa variance sont :

$$E(X) = \frac{1}{p}, V(X) = \frac{1 - p}{p^2}.$$

Loi de Poisson

Définition : Une variable aléatoire suit la loi de Poisson de paramètre λ lorsque :

$$\forall n \in \mathbf{N}, P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}.$$

Remarque : On définit bien ainsi la loi de probabilité d'une variable aléatoire, puisque, pour tout $n \in \mathbf{N}$, $\frac{e^{-\lambda} \lambda^n}{n!} \geq 0$ et :

$$\sum_{n=0}^{+\infty} \frac{e^{-\lambda} \lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1.$$

Notation : $X \hookrightarrow \mathcal{P}(\lambda)$.

Proposition 16.9.— Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors son espérance et sa variance sont :

$$E(X) = V(X) = \lambda.$$

Résultat asymptotique important

Proposition 16.10.— Si, pour tout n , $X_n \hookrightarrow \mathcal{B}(n, p_n)$ et si $\lim_{n \rightarrow +\infty} np_n = \lambda$, alors, pour tout $k \in \mathbf{N}$, on a :

$$\lim_{n \rightarrow +\infty} P(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Remarque : Dans la pratique, lorsque $X_n \hookrightarrow \mathcal{B}(n, p_n)$, le calcul de $P(X_n = k)$ devient difficile lorsque n est grand. On cherche alors une approximation en supposant que X_n suit une loi de Poisson plutôt qu'une loi binomiale. De façon théorique, il faut vérifier que np_n a une limite finie λ ; en pratique, on se contentera des conditions : n est grand ($n \geq 50$) et p_n petit ($p_n \leq 0,1$) et $np_n \leq 15$ (np_n sera alors une approximation de λ).

■ Bonus en marge du programme : fonctions génératrices

Soit X une variable aléatoire à valeurs dans \mathbf{N} .

Proposition 16.11.— Le rayon de convergence R_X de la série entière $\sum P(X = n) t^n$ est au moins égal à 1.

Définition : La *fonction génératrice (ou série génératrice)* d'une variable aléatoire X à valeurs dans \mathbf{N} est définie par :

$$G_X(t) = E(t^X) = \sum_{n=0}^{+\infty} P(X = n) t^n.$$

Proposition 16.12.— La loi d'une variable aléatoire X à valeurs dans \mathbf{N} est caractérisée par sa fonction génératrice G_X .

Proposition 16.13.— La variable aléatoire X admet une espérance $E(X)$ si et seulement si G_X est dérivable en 1 et, si tel est le cas, $E(X) = G'_X(1)$.

Bonus en marge du programme : fonctions génératrices des lois usuelles

Proposition 16.14.— ► **Loi de Bernoulli de paramètre** $p \in]0, 1[$: si $X \hookrightarrow \mathcal{B}(p)$,

pour tout $t \in \mathbb{R}$, $G_X(t) = 1 - p + pt$;

► **loi binomiale de paramètres** $n \geq 2$ et $p \in]0, 1[$: si $X \hookrightarrow \mathcal{B}(n, p)$,

pour tout $t \in \mathbb{R}$, $G_X(t) = (1 - p + pt)^n$;

► **loi géométrique de paramètre** $p \in]0, 1[$: si $X \hookrightarrow \mathcal{G}(p)$,

pour tout $t \in [-1/(1 - p), 1/(1 - p)[$, $G_X(t) = \frac{pt}{1 - (1 - p)t}$;

loi de Poisson de paramètre $\lambda > 0$: si $X \hookrightarrow \mathcal{P}(\lambda)$,

pour tout $t \in \mathbb{R}$, $G_X(t) = e^{\lambda(t-1)}$.

Méthodes

■ Lois de probabilité d'une variable aléatoire discrète

□ Méthode 16.1.— Comment déterminer la loi de probabilité d'une variable aléatoire discrète

Soit X une variable aléatoire discrète sur un espace probabilisé (Ω, \mathcal{A}, P) et

$X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$ son image.

Déterminer la loi de probabilité de X , c'est préciser $P(\{x_n\})$, pour tout $n \in \mathbf{N}$.

Pour cela, on peut :

- se livrer à un calcul de probabilités en utilisant le matériel du chapitre précédent ;
- déterminer d'abord la fonction de répartition F_X de X pour en déduire la loi de probabilité ; par exemple, si $X(\Omega) = \mathbf{N}$, $\forall k \in \mathbf{N}$, $P(X = k) = F_X(k) - F_X(k-1)$;
- reconnaître une loi usuelle décrite en TSI1 ou en TSI2.

Exemples : 1. *Que signifie qu'une variable aléatoire X suit la loi de Bernoulli de paramètre $p \in]0, 1[$? Que signifie qu'une variable aléatoire X suit la loi binomiale de paramètres $n \geq 2$ et $p \in]0, 1[$?*

Une variable aléatoire suit la loi de Bernoulli de paramètre $p \in]0, 1[$ si $X(\Omega) = \{0, 1\}$, avec $P(X = 1) = p$ (et donc $P(X = 0) = 1 - p$). On note : $X \hookrightarrow \mathcal{B}(p)$.

Une variable aléatoire suit la loi binomiale de paramètres $n \geq 2$ et $p \in]0, 1[$ si elle est la somme de n variables de Bernoulli mutuellement indépendantes suivant la même loi $\mathcal{B}(p)$.

On note : $X \hookrightarrow \mathcal{B}(n, p)$.

2. *On lance un dé parfait et on recommence indéfiniment. Soit X la variable aléatoire prenant pour valeur le rang d'apparition du premier 6. Déterminer la loi de X .*

X est bien une variable aléatoire discrète puisque $X(\Omega) = \mathbf{N}^*$ (on peut remarquer que $\Omega = \{1, 2, 3, 4, 5, 6\}^{\mathbf{N}^*}$, ensemble des suites d'éléments de $\{1, 2, 3, 4, 5, 6\}$, n'est pas, lui, dénombrable). Pour tout $n \geq 1$, notons A_n l'événement « le $n^{\text{ème}}$ lancer donne un 6 ». La réalisation ou la non réalisation de A_n ne dépend que du $n^{\text{ème}}$ lancer et les lancers sont indépendants les uns des autres, donc les A_n sont des événements mutuellement indépendants.

► $(X = 1) = A_1$ et $P(X = 1) = \frac{1}{6}$.

► $(X = 2) = \overline{A_1} \cap A_2$, donc $P(X = 2) = P(\overline{A_1})P(A_2) = \frac{5}{6} \times \frac{1}{6}$.

► Plus généralement, pour tout $k \in \mathbf{N}^*$, $(X = k) = \overline{A_1} \cap \dots \cap \overline{A_{k-1}} \cap A_k$, donc :

$$P(X = k) = P(\overline{A_1}) \dots P(\overline{A_{k-1}})P(A_k) = \left(\frac{5}{6}\right)^{k-1} \times \frac{1}{6}.$$

Remarque : dans ce genre de situation, on peut reconnaître directement que X suit la loi géométrique de paramètre $\frac{1}{6}$: $X \hookrightarrow \mathcal{G}(\frac{1}{6})$.

□ Méthode 16.2.— Comment déterminer l'espérance et la variance d'une variable aléatoire discrète

- Au delà des formules de définition, il est indispensable de connaître certaines sommes (finies ou sommes de séries convergentes).
 - ▷ Dans le cas fini, citons $\sum_{k=1}^n k = \frac{n(n+1)}{2}$, $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$,
 $\sum_{k=1}^n k^3 = \frac{n^2(n+1)^2}{4}$, $\sum_{k=0}^n \binom{n}{k}$ et $\sum_{k=1}^n k \binom{n}{k}$ (voir **exercice ??**) ;
 - ▷ dans le cas infini, il faut connaître les sommes des séries entières usuelles.
- On peut aussi avoir à utiliser certaines propriétés vues en cours :
 - ▷ pour l'espérance, la linéarité, la formule du transfert ou la formule $E(XY) = E(X)(Y)$, lorsque X et Y sont indépendantes ;
 - ▷ pour la variance, $V(aX + b) = a^2V(X)$.
- Il faut également connaître l'espérance et la variance de chaque loi usuelle : Bernoulli, binomiale, géométrique, Poisson.

Exemple : Une urne contient $n \geq 3$ boules indiscernables au toucher, deux sont blanches et les autres sont noires. On tire une à une, et sans remise, les n boules de l'urne. X est la variable aléatoire égale au rang de la première boule blanche tirée. Calculer $E(X)$ et $V(X)$.

Les boules étant tirées successivement et sans remise, Ω est l'ensemble des permutations des n boules, donc $\text{Card } \Omega = n!$. Il y a deux boules blanches, donc $X(\Omega) = \llbracket 1, n-1 \rrbracket$.

Si N_k désigne l'événement : « c'est une boule noire qui apparaît au k -ième tirage », alors $(X = k) = N_1 \cap N_2 \cap \dots \cap N_{k-1} \cap \overline{N_k}$. La formule des probabilités composées donne :

$$\begin{aligned} P(X = k) &= P(N_1)P_{N_1}(N_2) \dots P_{N_1 \cap \dots \cap N_{k-2}}(N_{k-1})P_{N_1 \cap \dots \cap N_{k-1}}(\overline{N_k}) \\ &= \frac{n-2}{n} \cdot \frac{n-3}{n-1} \dots \frac{n-k}{n-k+2} \cdot \frac{2}{n-k+1} = \frac{2(n-k)}{n(n-1)} \end{aligned}$$

$$\begin{aligned} E(X) &= \sum_{k=1}^{n-1} kP(X = k) = \frac{2}{n(n-1)} \sum_{k=1}^{n-1} k(n-k) = \frac{2}{(n-1)} \sum_{k=1}^{n-1} k - \frac{2}{n(n-1)} \sum_{k=1}^{n-1} k^2 \\ &= \frac{2}{(n-1)} \times \frac{n(n-1)}{2} - \frac{2}{n(n-1)} \times \frac{n(n-1)(2n-1)}{6} = n - \frac{2n-1}{3} \text{ et } E(X) = \frac{n+1}{3}. \end{aligned}$$

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 = \sum_{k=1}^{n-1} k^2 P(X = k) - \frac{(n+1)^2}{9} = \frac{2}{n(n-1)} \sum_{k=1}^{n-1} k^2(n-k) - \frac{(n+1)^2}{9} \\ &= \frac{2}{n-1} \sum_{k=1}^{n-1} k^2 - \frac{2}{n(n-1)} \sum_{k=1}^{n-1} k^3 - \frac{(n+1)^2}{9} = \frac{n(2n-1)}{3} - \frac{n(n-1)}{2} - \frac{(n+1)^2}{9} \end{aligned}$$

$$\text{Donc : } V(X) = \frac{(n+1)(n-2)}{18}.$$

Exemple : Une urne contient 3 boules blanches et 2 noires. on effectue, dans cette urne, des tirages successifs avec remise de chaque boule après tirage.

X est la variable aléatoire qui prend pour valeur le nombre de tirages nécessaires à l'obtention de la première boule blanche. Calculer l'espérance et la variance de X .

$X(\Omega) = \mathbf{N}^*$. X donne le rang du premier succès dans une suite illimitée d'épreuves de Bernoulli indépendantes de même paramètre $\frac{3}{5}$: elle suit donc la loi géométrique de paramètre $\frac{3}{5}$.

$$X \hookrightarrow \mathcal{G}\left(\frac{3}{5}\right), \text{ donc } E(X) = \frac{1}{\frac{3}{5}} = \frac{5}{3} \text{ et } V(X) = \frac{1 - \frac{3}{5}}{\left(\frac{3}{5}\right)^2} = \frac{10}{9}.$$

■ Méthode 16.3.— Comment étudier un couple de variables aléatoires discrètes

(X, Y) étant un couple de variables aléatoires discrètes sur Ω :

- Il faut commencer par établir la loi conjointe et les lois marginales. Dans le cas infini, il ne sera plus possible de rassembler les résultats dans un tableau ;
- Leur éventuelle indépendance sera vérifiée par :
 $\forall (x, y) \in X(\Omega) \times Y(\Omega), P(X = x, Y = y) = P(X = x)P(Y = y),$
ou par l'indépendance, $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, de $(X = x)$ et $(Y = y)$.
- On peut avoir à chercher la loi conditionnelle de X sachant que $(Y = y)$ (ou le contraire) ;
- Pour calculer leur covariance et leur coefficient de corrélation, on utilisera :
 $\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$
et $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$

Exemple : On effectue une suite de lancers avec une pièce non équilibrée pour laquelle la probabilité d'obtenir pile vaut $\frac{3}{4}$ (et celle d'obtenir face vaut donc $\frac{1}{4}$). On note X la longueur de la première chaîne obtenue, c'est-à-dire le nombre de tirages initiaux donnant le même résultat que le premier tirage. On note Y la longueur de la deuxième chaîne. Ainsi, si les premiers tirages donnent PPPFFP (peu importe la suite), on aura $X = 4$ et $Y = 2$. Déterminer la loi conjointe et les lois marginales du couple (X, Y) , puis la loi conditionnelle de X sachant $(Y = j)$. X et Y sont-elles indépendantes ?

La loi marginale de X se calcule assez aisément : $X(\Omega) = \mathbf{N}^*$ et on aura $X = k$ si la suite de lancers commence par k pile suivi d'une face ou par k face suivis d'un pile, cas incompatibles qui donnent $P(X = k) = \left(\frac{3}{4}\right)^k \frac{1}{4} + \left(\frac{1}{4}\right)^k \frac{3}{4} = \frac{3^k + 3}{4^{k+1}}.$

Pour déterminer la loi de Y , le plus simple est de passer par la loi conjointe du couple : on aura $(X, Y) = (i, j)$ si on débute par i pile, puis j face et à nouveau un pile, ou bien i face, j pile et une face, soit une probabilité de $P(X = i, Y = j) = \left(\frac{3}{4}\right)^i \left(\frac{1}{4}\right)^j \frac{3}{4} + \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^j \frac{1}{4} = \frac{3^{i+1} + 3^j}{4^{i+j+1}}.$

On utilise ensuite la formule des probabilités totales pour obtenir la loi de Y , avec le système complet d'événements $(X = i)_{i \geq 1}$.

$$\begin{aligned}
P(Y = j) &= \sum_{i=1}^{+\infty} P(X = i)P_{(X=i)}(Y = j) = \sum_{i=1}^{+\infty} P(X = i, Y = j) \\
&= \sum_{i=1}^{+\infty} \frac{3^{i+1} + 3^j}{4^{i+j+1}} = \frac{1}{4^j} \sum_{i=1}^{+\infty} \left(\frac{3}{4}\right)^{i+1} + \frac{3^j}{4^j} \sum_{i=1}^{+\infty} \left(\frac{1}{4}\right)^{i+1} \\
&= \frac{1}{4^j} \left(\frac{3}{4}\right)^2 \frac{1}{1 - \frac{3}{4}} + \frac{3^j}{4^j} \frac{1}{4^2} \frac{1}{1 - \frac{1}{4}} = \frac{9 + 3^{j-1}}{4^{j+1}}
\end{aligned}$$

On constate que la loi de Y n'est pas la même que celle de X , ce qui ne sautait pas aux yeux.

La loi conditionnelle de X sachant $(Y = j)$:

$$P_{(Y=j)}(X = i) = \frac{P((X = i) \cap (Y = j))}{P(Y = j)} = \frac{3^{i+1} + 3^j}{4^{i+j+1}} \cdot \frac{4^{j+1}}{9 + 3^{j-1}} = \frac{3^{i+1} + 3^j}{4^i(9 + 3^{j-1})}.$$

La loi marginale de X étant différente de la loi conditionnelle de X sachant $(Y = j)$, X et Y ne sont pas indépendantes.

□ Méthode 16.4.— Comment étudier une variable aléatoire définie comme une fonction d'une ou de plusieurs v.a.r.d dans les cas les plus courants

- Pour étudier $Y = \phi(X)$, on remarque que $Y(\Omega) = \phi(X(\Omega))$ et pour tout $k \in Y(\Omega)$, l'événement $Y = k$ est égal à $X = \phi^{-1}(\{k\})$. On en déduit la loi de Y , c'est-à-dire toutes les probabilités $P(Y = k)$. Pour $E(Y)$, on utilise la **formule du transfert**.

- Pour étudier $X + Y$, où X et Y sont deux v.a.r définis sur Ω , on calcule $P(X + Y = k)$ "à la main" pour $k \in (X + Y)(\Omega)$.

$$\text{On part de la formule : } P(X + Y = k) = \sum_{j \in X(\Omega)} P((X = j) \cap (Y = k - j)).$$

Dans le cas très fréquent, où X et Y sont à valeurs dans \mathbb{N} ,

$$P(X + Y = k) = \sum_{j=0}^n P((X = j) \cap (Y = k - j)).$$

Si les v.a.r dont on fait la somme, sont indépendantes :

$$P(X + Y = k) = \sum_{j \in X(\Omega)} P(X = j) \times P(Y = k - j).$$

$$\text{Et si } X \text{ et } Y \text{ sont à valeurs dans } \mathbb{N}, P(X + Y = k) = \sum_{j=0}^n P(X = j) \times P(Y = k - j).$$

- Pour étudier $Y = \min(X_1, \dots, X_n)$ ou $Z = \max(X_1, \dots, X_n)$, on calcule $P(Y = i)$, où $i \in Y(\Omega)$: $P(Y \geq i) = P((X_1 \geq i) \cap \dots \cap (X_n \geq i))$.
Puis on en déduit : $P(Y = i) = P(Y \geq i) - P(Y \geq i + 1)$.
De même, pour calculer $P(Z = i)$, où $i \in Z(\Omega)$, on calcule :
 $P(Z \leq i) = P((X_1 \leq i) \cap \dots \cap (X_n \leq i))$.
Puis on en déduit : $P(Z = i) = P(Z \leq i) - P(Z \leq i - 1)$.

Exemple : X et Y étant deux variables aléatoires indépendantes suivant des lois binomiales $\mathcal{B}(m, p)$ et $\mathcal{B}(n, p)$, montrer que $X + Y$ suit une loi binomiale $\mathcal{B}(m + n, p)$. (on utilisera que

$$\sum_{i=0}^k \binom{m}{i} \binom{n}{k-i} = \binom{m+n}{k}.$$

$(X = k)_{k \in \llbracket 0, m \rrbracket}$ forment un système complet d'événements, donc, d'après la loi des probabilités totales :

$$\forall k \in \llbracket 0, m+n \rrbracket, P(X + Y = k) = \sum_{i=0}^k P(X = i)P_{(X=i)}(Y = k - i).$$

Du fait de l'indépendance, $\forall k \in \llbracket 0, m+n \rrbracket, P(X + Y = k) = \sum_{i=0}^k P(X = i)P(Y = k - i).$

$$\begin{aligned} P(X + Y = k) &= \sum_{i=0}^k \binom{m}{i} p^i (1-p)^{m-i} \binom{n}{k-i} p^{k-i} (1-p)^{n-k+i} \\ &= \sum_{i=0}^k \binom{m}{i} \binom{n}{k-i} p^k (1-p)^{m+n-k} \\ &= p^k (1-p)^{m+n-k} \sum_{i=0}^k \binom{m}{i} \binom{n}{k-i} = \binom{m+n}{k} p^k (1-p)^{m+n-k} \end{aligned}$$

$X + Y$ suit bien la loi binomiale $\mathcal{B}(m + n, p)$.

Exemple : Soit X_1, \dots, X_k , k v.a.r.d de même loi uniforme sur $\llbracket 1, n \rrbracket$ et supposées mutuellement indépendantes. On pose $Z = \max(X_1, \dots, X_k)$.

Déterminer la loi de Z et calculer son espérance pour $k = 2$ et $k = 3$.

$\forall i \leq n, P(Z \leq i) = P[(X_1 \leq i) \cap \dots \cap (X_k \leq i)] = P(X_1 \leq i)^k = \left(\frac{i}{n}\right)^k$ car les variables X_1, \dots, X_k sont mutuellement indépendantes.

Et : $\forall i \in \llbracket 2, n \rrbracket, P(Z = i) = P(Z \leq i) - P(Z \leq i-1) = \frac{i^k - (i-1)^k}{n^k}.$

De plus, $P(Z = 1) = \left(\frac{1}{n}\right)^k.$

Écrivons $E(Z)$ dans le cas général sous forme d'une somme :

$$E(Z) = 1 \times P(Z = 1) + \sum_{i=2}^n \frac{i^{k+1} - i \times (i-1)^k}{n^k}.$$

Dans le cas $k = 2$, on a donc :

$$E(Z) = \left(\frac{1}{n}\right)^2 + \sum_{i=2}^n \frac{i^3 - i \times (i-1)^2}{n^2} = \sum_{i=1}^n \frac{2i^2 - i}{n^2}.$$

On rappelle : $\sum_{i=1}^n i = \frac{n(n+1)}{2}, \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}.$

On en déduit : $E(Z) = \frac{(n+1)(4n-1)}{6n}.$

Enfin si $k = 3$: $E(Z) = \left(\frac{1}{n}\right)^3 + \sum_{i=2}^n \frac{i^4 - i \times (i-1)^3}{n^3} = \sum_{i=1}^n \frac{i + 3i^3 - 3i^2}{n^3}$, ce qui nous donne en rappelant la formule $\sum_{i=1}^n i^3 = \frac{n^2(n+1)^2}{4}$: $E(Z) = \frac{(n+1)(3n-1)}{4n}$.

■ Approximations

□ **Méthode 16.5.— Comment utiliser l'inégalité de Bienaymé-Tchebychev**
Le mieux est de voir l'exemple.

Exemple : On réalise 400 fois la même expérience, dont la probabilité de succès est $0,8 = \frac{4}{5}$. On suppose que les 400 expériences sont indépendantes. Soit X le nombre de succès obtenus. Calculer $E(X)$ puis donner un minorant de $P(300 < X < 340)$.

X suit la loi binomiale $\mathcal{B}\left(400, \frac{4}{5}\right)$, donc $E(X) = 320$.

$$(300 < X < 340) = (|X - E(X)| < 20)$$

est l'événement contraire de $(|X - E(X)| \geq 20)$.

D'après l'inégalité de Bienaymé-Tchebychev, $P(|X - E(X)| \geq 20) \leq \frac{V(X)}{20^2} = \frac{64}{400}$, donc :

$$P(300 < X < 340) = 1 - P(|X - E(X)| \geq 20) \geq \frac{21}{25}.$$

Exemple : On utilise un dé équilibré. Cherchons le nombre de lancers qu'il faut effectuer pour pouvoir affirmer avec un risque d'erreur inférieur à 0,05, que la fréquence d'apparition du numéro 1 au cours de ces n lancers sera dans l'intervalle $\left] \frac{1}{6} - \frac{1}{100}, \frac{1}{6} + \frac{1}{100} \right[$?

X , variable aléatoire donnant le nombre d'as à l'issue de ces n lancers, suit la loi binomiale $\mathcal{B}\left(n, \frac{1}{6}\right)$. La fréquence d'apparition du numéro 1 au cours de ces n lancers est $\frac{X}{n}$.

$$\left(\frac{X}{n} \in \left] \frac{1}{6} - \frac{1}{100}, \frac{1}{6} + \frac{1}{100} \right[\right) = \left(\left| X - \frac{n}{6} \right| < \frac{n}{100} \right)$$

qui est l'événement contraire de $\left(\left| X - \frac{n}{6} \right| \geq \frac{n}{100} \right)$. D'après l'inégalité de Bienaymé-Tchebychev,

$$P\left(\left| X - \frac{n}{6} \right| \geq \frac{n}{100} \right) \leq \frac{V(X)}{\frac{n^2}{10^4}} = \frac{10^4 \frac{5n}{36}}{n^2}.$$

Puis, on utilise la deuxième remarque après la **proposition 16.7**.

$$P\left(\left|X - \frac{n}{6}\right| < \frac{n}{100}\right) > 1 - \frac{10^4 \frac{5n}{36}}{n^2} = 1 - \frac{5 \times 10^4}{36n}.$$

Dire que le risque d'erreur est inférieur à 0,05 (5 pour cent), c'est dire que la probabilité de l'événement $\left(\frac{X}{n} \in \left[\frac{1}{6} - \frac{1}{100}, \frac{1}{6} + \frac{1}{100}\right]\right)$ est supérieure à 0,95. Pour que ce soit le cas, il suffit que $1 - \frac{5 \times 10^4}{36n^2} \geq \frac{95}{100}$, ce qui donne $n \geq \frac{10^6}{36}$, c'est-à-dire $n \geq 167$.

□ Méthode 16.6.— Quand approcher une loi binomiale par une loi de Poisson

Une loi de Poisson étant la loi des événements rares, une loi binomiale sera d'autant mieux approchée par une telle loi que le paramètre p sera petit et que n sera grand. Dans ce cas, on approchera le paramètre λ de la loi de Poisson par la valeur de np , à condition qu'elle ne soit pas trop grande (à cause de la condition théorique : $\lim_{n \rightarrow +\infty} np_n = \lambda$.)

Exemple : Soit une variable aléatoire $X \hookrightarrow \mathcal{B}(300, 0,01)$. Est-il légitime de l'approcher par une loi de Poisson ? Quel est alors le paramètre de cette loi ? Donner une approximation de $P(X \geq 3)$.

$n = 300 > 50$ et $np = 3 < 15$, donc on peut approcher la loi de X par $\mathcal{P}(3)$.

$$P(X \geq 3) \approx \sum_{k=3}^{+\infty} e^{-3} \frac{3^k}{k!}. \text{ Or, } \sum_{k=3}^{+\infty} e^{-3} \frac{3^k}{k!} = e^{-3} \left(e^3 - 1 - 3 - \frac{9}{2} \right) \approx 0,577.$$

Énoncé des exercices

■ ???

Exercice 16.1 :

Exercice 16.2 :

Exercice 16.3 :

Exercice 16.4 :

Exercice 16.5 :

■ ???

Exercice 16.6 :

Exercice 16.7 :

■ ???

Exercice 16.8 :

Exercice 16.9 :

Exercice 16.10 :

Exercice 16.11 :

Exercice 16.12 :

■ ???

Exercice 16.13 :

Exercice 16.14 :

Exercice 16.15 :

Exercice 16.16 :

Exercice 16.17 :

Indications

- ___ Ex. 16.1 _____
- ___ Ex. 16.2 _____
- ___ Ex. 16.3 _____
- ___ Ex. 16.4 _____
- ___ Ex. 16.5 _____
- ___ Ex. 16.6 _____
- ___ Ex. 16.7 _____
- ___ Ex. 16.8 _____
- ___ Ex. 16.9 _____
- ___ Ex. 16.10 _____
- ___ Ex. 16.11 _____
- ___ Ex. 16.12 _____
- ___ Ex. 16.13 _____
- ___ Ex. 16.14 _____
- ___ Ex. 16.15 _____
- ___ Ex. 16.16 _____
- ___ Ex. 16.17 _____

Corrigé des exercices

- ___ Exercice 16.1 _____
- ___ Exercice 16.2 _____
- ___ Exercice 16.3 _____
- ___ Exercice 16.4 _____
- ___ Exercice 16.5 _____
- ___ Exercice 16.6 _____
- ___ Exercice 16.7 _____
- ___ Exercice 16.8 _____
- ___ Exercice 16.9 _____
- ___ Exercice 16.10 _____
- ___ Exercice 16.11 _____
- ___ Exercice 16.12 _____
- ___ Exercice 16.13 _____
- ___ Exercice 16.14 _____
- ___ Exercice 16.15 _____
- ___ Exercice 16.16 _____
- ___ Exercice 16.17 _____

Chapitre 17

Isometries in euclidian space

■ Objectifs

■ Les incontournables :

- ▶ Se familiariser avec les isométries vectorielles (ou automorphismes orthogonaux) d'un espace euclidien.
- ▶ Savoir caractériser les matrices orthogonales.
- ▶ Connaître le théorème fondamental sur la diagonalisabilité des endomorphismes et des matrices symétriques.
- ▶ Savoir classer les endomorphismes orthogonaux de \mathbf{R}^2 .
- ▶ Savoir étudier une rotation dans \mathbf{R}^3 et trouver en particulier son axe et son angle.
- ▶ Connaître analytiquement et géométriquement le produit vectoriel.
- ▶ Connaître le spectre d'une isométrie vectorielle et préciser les éléments ϕ de $O(\mathbf{R}^3)$ dans le cas où ϕ est diagonalisable dans \mathbf{R} .

■ Et plus si affinités ...

- ▶ Savoir écrire analytiquement une symétrie orthogonale par rapport au *s.e.v* F , à partir d'un système d'équations de F ou d'une base de F dans le cas $n \leq 3$.
- ▶ Savoir classer les isométries vectorielles de \mathbf{R}^3 .
- ▶ Savoir faire un changement intelligent de base orthonormée dans \mathbf{R}^3 pour par exemple écrire analytiquement une rotation dans \mathbf{R}^3 .

Résumé de cours

$(E, \langle \cdot, \cdot \rangle)$ est un espace euclidien de dimension finie n de norme associée $\| \cdot \|$.

■ Isométries vectorielles d'un espace euclidien

Définition : Isométries vectorielles (ou automorphismes orthogonaux) —. Un endomorphisme u de E est une **isométrie vectorielle** lorsqu'il conserve le produit scalaire, c'est-à-dire lorsque :

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

Proposition 17.1. — Conservation de la norme —. Un endomorphisme u de E est une isométrie vectorielle si et seulement s'il conserve la norme, c'est-à-dire si et seulement si :

$$\forall \vec{x} \in E, \|u(\vec{x})\| = \|\vec{x}\|.$$

Proposition 17.2. — Bijektivité —. Une isométrie vectorielle est un automorphisme de E .

Proposition 17.3. — Soit u un endomorphisme de E , les propositions suivantes sont équivalentes :

- (i) u est une isométrie vectorielle de E ;
- (ii) u est un endomorphisme de E transformant toute base orthonormée de E en une base orthonormée de E ;
- (iii) il existe une base orthonormée de E que l'endomorphisme u transforme en une base orthonormée de E ;

Définition : Groupe orthogonal —.

L'ensemble $O(E)$ des automorphismes orthogonaux de E est appelé **groupe orthogonal** de E .

Proposition 17.4. — Stabilité des s.e.v —.

Si F est un s.e.v de E , stable par $\phi \in O(E)$, alors F^\perp est stable par ϕ .

■ Matrices orthogonales

Définition : Matrice orthogonale —. Une matrice de $\mathcal{M}_n(\mathbf{R})$ est **orthogonale** lorsque l'endomorphisme de \mathbf{R}^n qui lui est canoniquement associé est une isométrie vectorielle.

Notation : L'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbf{R})$ est noté $O(n)$ ou $O_n(\mathbf{R})$.

Théorème-Définition 17.5. — Matrices orthogonales —.

Soit M une matrice de $\mathcal{M}_n(\mathbf{R})$, les propositions suivantes sont équivalentes :

- (i) M est une matrice orthogonale ;
- (ii) les colonnes de M forment une base orthonormée de \mathbf{R}^n ;
- (iii) M est inversible et $M^{-1} = M^T$;
- (iv) $M^T M = I_n$ (ou $M M^T = I_n$) ;
- (v) les lignes de M forment une base orthonormée de \mathbf{R}^n .

Remarques : 1. M est orthogonale si et seulement si M^{-1} l'est d'une part, et si et seulement si M^T l'est d'autre part, ce qui justifie l'équivalence (ii) \iff (v).

2. L'inverse de $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$, $\theta \in \mathbf{R}$, qui est orthogonale, est $\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$.

3. Écrire $M^T M = I_n$ est en effet équivalent à écrire $M M^T = I_n$ car alors : $M^{-1} = M^T$.

Proposition 17.6.— La matrice de passage entre deux bases orthonormées est orthogonale.

Théorème-Définition 17.7.— **Caractérisation matricielle d'une isométrie vectorielle** —.

Les propositions suivantes sont équivalentes :

- (i) u est une isométrie vectorielle ;
- (ii) dans toute base orthonormée, la matrice associée à u est orthogonale ;
- (iii) il existe une base orthonormée dans laquelle la matrice associée à u est orthogonale.

Proposition 17.8.— **Déterminant d'une matrice orthogonale** —.

Le déterminant d'une matrice orthogonale (respectivement d'une isométrie vectorielle) vaut ± 1 .

Définition : Groupe spécial orthogonal d'ordre n —. L'ensemble des matrices orthogonales de déterminant égal à 1 est appelé **groupe spécial orthogonal d'ordre n** et noté $\mathrm{SO}(n)$ ou $\mathrm{SO}_n(\mathbf{R})$.

Proposition 17.9.— **Valeurs propres réelles d'une matrice orthogonale** —.

Les valeurs propres réelles d'une matrice de $\mathrm{O}(n)$ appartiennent à $\{-1, 1\}$.

■ Espace euclidien orienté de dimension 2 ou 3.

Théorème-Définition 17.10.— **Orientation-Produit mixte** —. Ici $n \in \{2, 3\}$. Orienter \mathbf{R}^n , c'est choisir une base orthonormée $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ de \mathbf{R}^n . Soit $\mathcal{B}' = \{\vec{e}'_1, \dots, \vec{e}'_n\}$ une autre base orthonormée de \mathbf{R}^n . On dit qu'elle est **directe** si $\mathrm{Det}_{\mathcal{B}}(\mathcal{B}') = 1$ et **indirecte** si $\mathrm{Det}_{\mathcal{B}}(\mathcal{B}') = -1$.

Si \vec{u}, \vec{v} sont deux vecteurs de \mathbf{R}^2 (respectivement \vec{u}, \vec{v} et \vec{w} sont trois vecteurs de \mathbf{R}^3), le réel $\mathrm{Det}_{\mathcal{B}}(\vec{u}, \vec{v})$, noté $[\vec{u}, \vec{v}]$ (respectivement $\mathrm{Det}_{\mathcal{B}}(\vec{u}, \vec{v}, \vec{w})$, noté $[\vec{u}, \vec{v}, \vec{w}]$) est indépendant de la base orthonormée directe \mathcal{B} dans laquelle on le calcule. Ce nombre est appelé **produit mixte** des deux vecteurs \vec{u} et \vec{v} (respectivement des trois vecteurs $\vec{u}, \vec{v}, \vec{w}$).

Dans le plan \mathbf{R}^2 , la base orthonormée directe de référence est telle que le second vecteur est obtenu « en tournant » d'un angle droit le premier vecteur dans le sens inverse des aiguilles d'une montre (appelé aussi sens trigonométrique).

Remarque : Ainsi, $P_{\mathcal{B}}^{\mathcal{B}'} \in \mathrm{SO}_n(\mathbf{R})$ si \mathcal{B}' est directe et $P_{\mathcal{B}}^{\mathcal{B}'} \in \mathrm{O}_n(\mathbf{R}) \setminus \mathrm{SO}_n(\mathbf{R})$ si \mathcal{B}' est indirecte.

Application : orientation dans le plan et l'espace d'une droite ou d'un plan

- 1. On oriente une droite D de \mathbf{R}^2 ou \mathbf{R}^3 par la donnée d'un vecteur \vec{n} de norme 1, base de D .
- 2. Soit $\{\vec{u}_1, \vec{u}_2\}$ une base orthonormée d'un plan P de \mathbf{R}^3 . Orienter P c'est choisir un vecteur \vec{u}_3 orthogonal à P tel que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ soit directe.

Théorème-Définition 17.11.— **Produit vectoriel** —. Si \vec{u} et \vec{v} sont deux vecteurs de \mathbf{R}^3 , il existe un unique vecteur \vec{w} de \mathbf{R}^3 tel que :

$$\forall \vec{X} \in \mathbf{R}^3, [\vec{u}, \vec{v}, \vec{X}] = \langle \vec{w}, \vec{X} \rangle.$$

Ce vecteur \vec{w} est appelé **produit vectoriel de \vec{u} par \vec{v}** et noté $\vec{u} \wedge \vec{v}$.

Proposition 17.12.— Propriétés du produit vectoriel —.

1. L'application $(\vec{u}, \vec{v}) \mapsto \vec{u} \wedge \vec{v}$ est bilinéaire alternée, c'est-à-dire qu'elle est linéaire par rapport à chacune des variables (l'autre étant fixée) et que $\vec{v} \wedge \vec{u} = -\vec{u} \wedge \vec{v}$.
2. $\vec{u} \wedge \vec{v}$ est orthogonal à \vec{u} et à \vec{v} . De plus : $\vec{u} \wedge \vec{v} = 0_{\mathbf{R}^3} \Leftrightarrow \vec{u}$ et \vec{v} colinéaires.
3. Si \vec{u} et \vec{v} sont non colinéaires, $(\vec{u}, \vec{v}, \vec{u} \wedge \vec{v})$ est une base directe de \mathbf{R}^3 , c'est-à-dire : $[\vec{u}, \vec{v}, \vec{u} \wedge \vec{v}] > 0$.

Proposition 17.13.— Expression analytique du produit vectoriel —.

Si, dans une base orthonormée directe de \mathbf{R}^3 , \vec{u} et \vec{v} ont respectivement pour coordonnées (x, y, z) et (x', y', z') alors $\vec{u} \wedge \vec{v}$ a, dans cette base, pour coordonnées : $(yz' - y'z, zx' - xz', xy' - x'y)$.

Définition : Angle de deux vecteurs de \mathbf{R}^3 —. Soient \vec{u} et \vec{v} deux vecteurs non nuls de \mathbf{R}^3 , la mesure de l'angle des vecteurs \vec{u} et \vec{v} est l'unique réel θ de l'intervalle $[0, \pi]$ tel que : $\frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \|\vec{v}\|} = \cos \theta$. (Il s'agit d'un angle « géométrique ».)

Proposition 17.14.— Définition géométrique du produit vectoriel —. Soient \vec{u} et \vec{v} deux vecteurs non liés et P le plan engendré par \vec{u} et \vec{v} . Soit \vec{w} un vecteur unitaire orientant P . Alors : $\vec{u} \wedge \vec{v} = \|\vec{u}\| \times \|\vec{v}\| (\sin \theta) \vec{w}$. Le réel θ désigne une mesure de l'angle orienté de \vec{u} et \vec{v} selon \vec{w} dans le plan P .

Proposition 17.15.— Interprétation géométrique comme aire ou volume —.

1. $\|\vec{u} \wedge \vec{v}\|$ est l'aire du parallélogramme construit sur \vec{u} et \vec{v} .
2. $|\langle \vec{u}, \vec{v}, \vec{w} \rangle|$ est le volume du parallélépipède construit sur \vec{u}, \vec{v} et \vec{w} .

■ Cas des symétries orthogonales en dimension 2 ou 3

Définition : Soit p_F la projection orthogonale sur le s.e.v F de E . Alors $s_F = 2p_F - Id_E$ est la symétrie orthogonale par rapport à F .

Remarque : une symétrie orthogonale est un élément de $O(E)$. Les symétries orthogonales permettent de classer $O(E)$. C'est à connaître si $\dim E = 2$ et peut être utile si $\dim E = 3$ pour ceux qui veulent approfondir. Dans le cas où $\dim E > 3$, cela est réservé à quelques sujets de concours.

Corollaire 17.16.— Caractérisation matricielle d'une symétrie orthogonale —.

Soit ϕ un endomorphisme de E représenté par une matrice M dans une base orthonormée de E . ϕ est une symétrie orthogonale si et seulement si $M^2 = I_n$ et $M^T = M$.

■ Isométries vectorielles d'un plan euclidien**Proposition 17.17.— Matrices des automorphismes orthogonaux de \mathbf{R}^2 —.**

La matrice, dans une base orthonormée quelconque, d'un automorphisme orthogonal de \mathbf{R}^2 est de l'une des deux formes suivantes : $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ ou $\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$, $\theta \in \mathbf{R}$.

Donc : $SO(2) = \left\{ R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \theta \in \mathbf{R} \right\}$.

Définition : Rotations vectorielles planes —.

Les éléments de $\text{SO}(\mathbf{R}^2)$ sont appelés **rotations vectorielles planes**.

Pour toute rotation vectorielle plane r , il existe un réel θ , unique à un multiple entier de 2π près, tel que, quelle que soit la base \mathcal{B} orthonormée directe de \mathbf{R}^2 , $M_{\mathcal{B}}(r) = R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. θ est une mesure de l'angle de cette rotation, qui est, alors, la rotation d'angle θ .

Remarque : Le produit des rotations d'angles θ et θ' est la rotation d'angle $\theta + \theta'$.

Proposition 17.18.— Si r est la rotation d'angle θ , alors, pour tout vecteur unitaire \vec{a} de \mathbf{R}^2 :
 $\cos \theta = \langle \vec{a}, r(\vec{a}) \rangle$ et $\sin \theta = [\vec{a}, r(\vec{a})]$.
 (le déterminant étant calculé dans une base orthonormée directe).

Proposition 17.19.— Automorphismes orthogonaux négatifs du plan —. Les automorphismes orthogonaux du plan qui ne sont pas dans $\text{SO}(\mathbf{R}^2)$ sont des symétries orthogonales.

■ Isométries d'un espace euclidien de dimension trois

Proposition 17.20.— Matrices des automorphismes orthogonaux de $\text{SO}(\mathbf{R}^3)$ —.
 $f \in \text{SO}(\mathbf{R}^3)$ si et seulement s'il existe une base orthonormée directe $(\vec{u}, \vec{v}, \vec{w})$ de \mathbf{R}^3 dans laquelle la matrice de f est : $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$.
 f est la rotation de \mathbf{R}^3 d'axe $D = \text{Vect}(\vec{u})$ et d'angle θ , que l'on note $r(D, \theta)$.

Le choix de \vec{u} oriente le plan \vec{u}^\perp . f est aussi bien $r_{(D, \theta)}$ que $r_{(\text{Vect}(-u), -\theta)}$.

Remarque : La classification complète des éléments de $\text{O}(\mathbf{R}^3)$ est plus complexe que dans $\text{O}(\mathbf{R}^2)$. En dehors des rotations (qui incluent les symétries orthogonales par rapport à une droite) qui sont des isométries positives, on a pour isométries négatives les symétries orthogonales par rapport à des plans et aussi les composées de ces symétries et des rotations qui ne laissent aucun vecteur invariant (à part le vecteur nul).

■ Réduction des matrices symétriques réelles dans un espace euclidien

Soit $n \geq 1$. On suppose E espace vectoriel euclidien de dimension n , muni de son produit scalaire euclidien canonique.

Proposition 17.21.— Diagonalisabilité des matrices symétriques réelles—.
 Si $A \in \mathcal{S}_n(\mathbf{R})$ alors les sous-espaces propres de A sont deux à deux orthogonaux.

Théorème 17.22.— Théorème spectral—.

Toute matrice symétrique réelle est diagonalisable au moyen d'une matrice orthogonale. C'est-à-dire que si $A \in \mathcal{S}_n(\mathbf{R})$, il existe une matrice $P \in \text{O}_n(\mathbf{R})$ et une matrice diagonale D telle que :

$$D = P^{-1}AP = P^TAP.$$

Méthodes

■ Isométries vectorielles et matrices orthogonales

□ Méthode 17.1.— Comment montrer qu'une matrice est orthogonale

Soit $A \in \mathcal{M}_n(\mathbf{R})$, $A \in O(n)$ lorsque :

- ▶ Ses vecteurs colonnes (ou ses vecteurs lignes) forment une base orthonormée de \mathbf{R}^n .
- ▶ $A^T A = I_n$.
- ▶ Elle est la matrice dans une base orthonormée d'un endomorphisme orthogonal.

Soit $A \in O(n)$, A appartient à $SO(n)$ lorsque :

- ▶ Son déterminant est égal à 1.
- ▶ La base orthonormée de \mathbf{R}^n formée par ses vecteurs colonnes (ou ses vecteurs lignes) est directe.
- ▶ Dans le cas particulier où $n = 3$, le produit vectoriel de ses deux premières colonnes (lignes) est égal à la troisième.

Exemple : Soit $A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$, avec a, b et c réels non tous nuls.

1. Montrer que $I_3 + A$ est inversible. Que dire de $I_3 - A$?
2. Montrer que $(I_3 - A)(I_3 + A)^{-1} \in SO(3)$.

$$1. \quad \text{Det}(I_3 + A) = \begin{vmatrix} 1 & -c & b \\ c & 1 & -a \\ -b & a & 1 \end{vmatrix} = 1 + a^2 + b^2 + c^2 \neq 0.$$

$I_3 + A$ est inversible, $I_3 - A$ aussi puisque $\text{Det}(I_3 - A) = 1 + a^2 + b^2 + c^2$.

$$2. \quad \text{Det}[(I_3 - A)(I_3 + A)^{-1}] = \text{Det}(I_3 - A) \frac{1}{\text{Det}(I_3 + A)} = 1.$$

▶ Remarquons que $(I_3 + A)^T = I_3 - A$, que $(I_3 - A)(I_3 + A) = I_3 - A^2 = (I_3 + A)(I_3 - A)$ et que, si M est inversible, M^T l'est aussi, avec $(M^T)^{-1} = (M^{-1})^T$.

$$\begin{aligned} ((I_3 - A)(I_3 + A)^{-1})((I_3 - A)^T(I_3 + A)^{-1}) &= (I_3 - A)(I_3 + A)^{-1}(I_3 - A)^{-1}(I_3 + A) \\ &= (I_3 - A)((I_3 - A)(I_3 + A))^{-1}(I_3 + A) \\ &= (I_3 - A)(I_3 - A^2)^{-1}(I_3 + A) \\ &= (I_3 - A)((I_3 + A)(I_3 - A))^{-1}(I_3 + A) \\ &= (I_3 - A)(I_3 - A)^{-1}(I_3 + A)^{-1}(I_3 + A) = I_3 \end{aligned}$$

$$(I_3 - A)(I_3 + A)^{-1} \in SO(3).$$

□ **Méthode 17.2.—** Comment montrer que $\phi \in \mathcal{L}(E)$ est une isométrie vectorielle ($E, \langle \cdot, \cdot \rangle$) est un espace euclidien. Pour montrer que $\phi \in \mathcal{O}(E)$, on peut :

- ▶ Revenir à la définition en montrant qu'il conserve le produit scalaire :
 $\forall (\vec{x}, \vec{y}) \in E^2, \langle \phi(\vec{x}), \phi(\vec{y}) \rangle = \langle \vec{x}, \vec{y} \rangle$, ou la norme : $\forall \vec{x} \in E, \|\phi(\vec{x})\| = \|\vec{x}\|$.
- ▶ Montrer qu'il transforme une base orthonormée en une base orthonormée.
- ▶ Vérifier que sa matrice dans une base orthonormée donnée est orthogonale.

Exemple : Soit \vec{u} un vecteur unitaire de \mathbf{R}^n . Montrer que l'application linéaire de \mathbf{R}^n dans \mathbf{R}^n définie par $f(\vec{x}) = \vec{x} - 2 \langle \vec{u}, \vec{x} \rangle \vec{u}$ est une isométrie vectorielle de \mathbf{R}^n .

$$\begin{aligned} \forall (\vec{x}, \vec{y}) \in (\mathbf{R}^n)^2, \langle f(\vec{x}), f(\vec{y}) \rangle &= \langle \vec{x} - 2 \langle \vec{u}, \vec{y} \rangle \vec{u}, \vec{y} - 2 \langle \vec{u}, \vec{x} \rangle \vec{u} \rangle \\ &= \langle \vec{x}, \vec{y} \rangle - 2 \langle \vec{u}, \vec{y} \rangle \langle \vec{u}, \vec{x} \rangle - 2 \langle \vec{u}, \vec{x} \rangle \langle \vec{u}, \vec{y} \rangle + 4 \langle \vec{u}, \vec{x} \rangle \langle \vec{u}, \vec{y} \rangle \|\vec{u}\|^2 = \langle \vec{x}, \vec{y} \rangle. \end{aligned}$$

□ **Méthode 17.3.—** Comment identifier un endomorphisme ϕ , à partir de sa matrice canoniquement associée dans un espace euclidien de dimension $n = 3$.

Soit $M \in \mathcal{M}_3(\mathbf{R})$ et ϕ l'endomorphisme de \mathbf{R}^n canoniquement associé.

On commence par rechercher si M est orthogonale (en faisant par exemple $M^T M$).

- ▶ Si c'est le cas ($M^T M = I_n$), on peut calculer $\text{Det } M = \pm 1$.
 1. Si $\text{Det } M = 1$ alors $M \in \text{SO}_3(\mathbf{R})$. Si, de plus, $M = M^T$, alors M est une symétrie orthogonale. M est alors nécessairement la matrice d'une symétrie orthogonale par rapport à une droite. La recherche des vecteurs invariants donne une équation de cette droite.
 Si $M \neq M^T$, on continue alors avec la **méthode 17.6**.
 2. Si $\text{Det } M = -1$, si encore $M = M^T$, alors ϕ est une symétrie orthogonale par rapport à $F = \{\vec{u} \in E, \phi(\vec{u}) = \vec{u}\}$. Obligatoirement : $\dim F = 2$, et alors ϕ est une symétrie orthogonale par rapport à un plan.
 Sinon, si $M^T \neq M$, ϕ est une composée d'une symétrie orthogonale par rapport à un plan et d'une rotation. C'est une antirotation. L'étude est alors guidée car on sort du programme officiel.
- ▶ Si ce n'est pas le cas ($M^T M \neq I_n$), on peut tenter le calcul de M^2 .
 1. Si $M^2 = M$, ϕ est une projection. On sait alors que si $M = M^T$, la projection est orthogonale. Tout ceci est précisé dans la **méthode 12.6**.
 2. Si $M^2 = I_n$, ϕ est une symétrie vectorielle (mais non orthogonale).
 Allez voir la **méthode ??**.
 3. Dans les autres cas, espérez de l'énoncé ou d'une aide de l'examinateur qui ne saurait tarder !

Remarque : Cet algorithme a des « itinéraires bis ». En effet :

1. Si les colonnes de M sont C_1, C_2 et C_3 alors montrer que $C_1 \wedge C_2 = \pm C_3$, dit que l'on a à faire à une isométrie et si plus précisément, $C_1 \wedge C_2 = C_3$, on est devant une belle rotation.

2. Si M est une matrice de symétrie orthogonale, elle est semblable à une matrice du type $\text{Diag}(\pm 1, \pm 1, \pm 1)$, et $\text{Tr}(M)$ donne des renseignements sur le nombre de 1 et -1 . On en déduit le *s.e.v* F . Voir l'exemple ci-dessous.

Exemple : Déterminer la nature et les éléments caractéristiques de l'endomorphisme f de \mathbf{R}^3

canoniquement associé à $S = \frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$.

- On vérifie que les colonnes de S forment une base orthonormée de \mathbf{R}^3 , donc $S \in O(3)$.
- On constate par ailleurs que S est symétrique (son nom n'est pas un piège...), donc S est une matrice de symétrie orthogonale; elle est donc diagonalisable et non diagonale, de spectre inclus dans $\{-1, 1\}$, donc elle est forcément semblable à une des matrices $R = \text{diag}(-1, -1, 1)$ ou $-R = \text{diag}(1, 1, -1)$. Puisque $\text{Tr} S = 1$, par élimination, S est semblable à $-R$, S est une symétrie orthogonale par rapport à un plan.
- Reste à trouver le plan de cette symétrie qui est le sous-espace propre de S associé à la valeur propre 1 : $E_1(S)$. Comme $S - I_3$ est de rang 1, de lignes proportionnelles à $(1, -1, 1)$, f est la symétrie orthogonale par rapport au plan d'équation $x - y + z = 0$.

Remarque : Dans le cas $n = 2$, la **méthode 17.3** est de l'artillerie trop lourde. En effet, la **proposition 17.17** permet de voir immédiatement sur la matrice de ϕ (qui est donnée dans la base canonique de \mathbf{R}^2 qui est une base orthonormée pour le produit scalaire canonique) s'il s'agit d'une matrice de $O(2)$ et même de $\text{SO}(2)$.

□ **Méthode 17.4.—** Comment calculer le produit vectoriel $\vec{u} \wedge \vec{v}$ dans \mathbf{R}^3 .

- On peut calculer ses coordonnées avec la **proposition 17.13**.
- On peut le représenter géométriquement (après avoir vérifié que \vec{u} et \vec{v} ne sont pas colinéaires car alors $\vec{u} \wedge \vec{v} = \vec{0}$). Étant donné le plan $\Pi = \text{Vect}(\vec{u}, \vec{v})$, on place \vec{u} puis \vec{v} dans le sens trigonométrique sur Π . Puis, $\vec{u} \wedge \vec{v}$ est un vecteur orthogonal à Π pointant du plan vers notre oeil lorsqu'on regarde le plan Π "par dessus". Sa norme est donnée par la **proposition 17.14**.
- On peut appliquer aussi la bilinéarité ou l'antisymétrie de \wedge . Allez voir la **proposition 17.12**. Si $\{\vec{i}, \vec{j}, \vec{k}\}$ est une base orthonormée directe de l'espace et si \vec{u} et \vec{v} sont exprimés dans cette base, on peut utiliser : $\vec{i} \wedge \vec{j} = \vec{k}$, $\vec{j} \wedge \vec{k} = \vec{i}$, $\vec{k} \wedge \vec{i} = \vec{j}$.

Exemple : Soit $\{\vec{i}, \vec{j}, \vec{k}\}$ la base canonique orthonormée directe et les vecteurs :

$$\vec{u} = (\vec{i} \wedge \vec{j}) \wedge \vec{j}, \vec{v} = (\vec{j} \wedge \vec{k}) \wedge \vec{k}, \vec{w} = (\vec{k} \wedge \vec{i}) \wedge \vec{i}.$$

Calculer le vecteur $\vec{z} = (\vec{u} \wedge \vec{v}) \wedge \vec{w}$.

On commence par simplifier $\vec{u} = \vec{k} \wedge \vec{j} = -\vec{i}$, $\vec{v} = \vec{i} \wedge \vec{k} = -\vec{j}$ et $\vec{w} = \vec{j} \wedge \vec{i} = -\vec{k}$.

Il reste : $\vec{z} = [(-\vec{i}) \wedge (-\vec{j})] \wedge (-\vec{k}) = \vec{k} \wedge (-\vec{k}) = \vec{0}$.

❑ **Méthode 17.5.— Dans quelles circonstances user du produit vectoriel dans \mathbf{R}^3 .**

- Le produit vectoriel sert à construire une base orthonormée directe. Si \vec{u} et \vec{v} sont de norme 1 et orthogonaux alors $\vec{u} \wedge \vec{v}$ est l'unique vecteur de l'espace tel que $\{\vec{u}, \vec{v}, \vec{u} \wedge \vec{v}\}$ soit une base orthonormée directe de l'espace.
- $\|\vec{u} \wedge \vec{v}\|$ est l'aire du parallélogramme construit sur \vec{u} et \vec{v} .
- Si C_1, C_2 et C_3 sont les trois colonnes d'une matrice A de $\mathcal{M}_3(\mathbf{R})$, si $C_1 \wedge C_2 = C_3$ alors $A \in \text{SO}_3(\mathbf{R})$ et si $C_1 \wedge C_2 = -C_3$ alors $A \in \text{O}_3(\mathbf{R}) \setminus \text{SO}_3(\mathbf{R})$.
- Le calcul de $\text{Det}(\vec{u}, \vec{v}, \vec{w}) = [\vec{u}, \vec{v}, \vec{w}]$ est aussi celui de $(\vec{u} \wedge \vec{v}) \cdot \vec{w}$.
- Soit dans \mathcal{B} orthonormée directe, une droite (D) d'équation
$$\begin{cases} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \end{cases}$$
. Les vecteurs $\vec{n}_1(a_1, b_1, c_1)$ et $\vec{n}_2(a_2, b_2, c_2)$ ne sont pas colinéaires (sinon D ne serait pas une droite). Leur produit vectoriel $\vec{n}_1 \wedge \vec{n}_2$ est une base de (D) .
- Le produit vectoriel permet de déterminer le signe de $\sin \theta$, où θ est l'angle d'une rotation de l'espace. Voir la **méthode 17.6**.

Remarque : Les applications du produit vectoriel sont innombrables dans les autres disciplines scientifiques. Donnons en deux. En SI, le produit vectoriel est utilisé pour définir les torseurs et en physique, il est utilisé entre-autre pour le calcul d'un champ magnétostatique.

❑ **Méthode 17.6.— Comment déterminer les éléments caractéristiques de la rotation de \mathbf{R}^3 canoniquement associée à une matrice $M \in \text{SO}(3)$**

On note ϕ l'endomorphisme de \mathbf{R}^3 canoniquement associé à M .

- Chercher les vecteurs invariants qui doivent former une droite vectorielle. On choisit pour vecteur \vec{I} un vecteur directeur unitaire de cette droite.
- Choisir \vec{J} unitaire et orthogonal à \vec{I} . On peut alors calculer $\vec{I} \wedge \vec{J} = \vec{K}$. $\mathcal{B} = (\vec{I}, \vec{J}, \vec{K})$ est une base orthonormée directe, relativement à laquelle la matrice de

$$\phi \text{ est de la forme "réduite" : } \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

- À l'aide de la matrice initiale, on calcule les coordonnées de $\phi(\vec{J})$ dans la base canonique.

Le produit scalaire de deux vecteurs étant indépendant de la base orthonormée directe dans laquelle on le calcule :

$$\langle \phi(\vec{J}), \vec{J} \rangle = \cos \theta \text{ et } \langle \phi(\vec{J}), (\vec{I} \wedge \vec{J}) \rangle = [\vec{I}, \vec{J}, \phi(\vec{J})] = \sin \theta.$$

La formule $1 + 2 \cos \theta = \text{Tr}(M)$ permet de vérifier le calcul de $\cos \theta$. Si $\sin \theta \geq 0$, ϕ est la rotation d'axe $\text{Vect}(\vec{I})$ orienté par \vec{I} et d'angle θ ; si $\sin \theta < 0$, ϕ est la rotation d'axe $\text{Vect}(\vec{I})$ orienté par \vec{I} et d'angle $-\theta$.

Remarque : On peut « court-circuiter » certaines étapes. Le calcul de \vec{K} n'est pas indispensable. Il nous faut \vec{I} par $MX = X$. Puis $\text{Tr}(M) = 1 + 2 \cos \theta$ donne directement la valeur de $\cos \theta$.

On choisit alors un vecteur \vec{u} orthogonal à \vec{I} (pas nécessairement unitaire) et $\text{Det}(\vec{I}, \vec{u}, \phi(\vec{u}))$ donne le signe de $\sin \theta$, ce qui nous suffit pour conclure.

Exemples : Déterminer la nature et les éléments caractéristiques de l'endomorphisme ϕ de \mathbf{R}^3 canoniquement associé à $A = \frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$.

► $A = \frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$ est une matrice orthogonale puisque ses vecteurs colonnes forment une

base orthonormée de \mathbf{R}^3 . De plus, le produit vectoriel de ses deux premiers vecteurs colonnes est égal à son troisième vecteur colonne, donc cette base orthonormée est directe et $A \in \text{SO}(3)$.

► L'axe de ϕ est la droite vectorielle de ses vecteurs invariants :

$$\phi((x, y, z)) = (x, y, z) \Leftrightarrow A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} x + 2y - 2z = 3x \\ -2x + 2y + z = 3y \\ 2x + y + 2z = 3z \end{cases}$$

$$\Leftrightarrow \begin{cases} -x + 2y - 2z = 0 \\ -2x - y + z = 0 \end{cases}$$

$$L_2 \leftarrow L_2 - 2L_1 \Leftrightarrow \begin{cases} -x + 2y - 2z = 0 \\ -5y + 5z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = z \end{cases}$$

L'axe de la rotation vectorielle ϕ est $\Delta = \text{Vect}(0, 1, 1)$.

► Posons $\vec{I} = \frac{1}{\sqrt{2}}(0, 1, 1)$; choisissons pour \vec{J} un vecteur unitaire orthogonal à \vec{I} , par exemple, $\vec{J} = (1, 0, 0)$ puis posons : $\vec{K} = \vec{I} \wedge \vec{J} = \frac{1}{\sqrt{2}}(0, 1, -1)$. $(\vec{I}, \vec{J}, \vec{K})$ est une base orthonormée directe de

\mathbf{R}^3 relativement à laquelle la matrice de ϕ est la matrice réduite : $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$.

Puisque : $\phi(\vec{J}) = \cos \theta \vec{J} + \sin \theta \vec{K}$ donc $\cos \theta = \langle \phi(\vec{J}) | \vec{J} \rangle$ et $\sin \theta = \langle \phi(\vec{J}) | \vec{K} \rangle$.

$$\frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \Rightarrow \phi(\vec{J}) = \frac{1}{3}(1, -2, 2).$$

$$\begin{cases} \cos \theta = \left\langle \frac{1}{3}(1, -2, 2) | (1, 0, 0) \right\rangle = \frac{1}{3} \\ \sin \theta = \left\langle \frac{1}{3}(1, -2, 2) | \frac{1}{\sqrt{2}}(0, 1, -1) \right\rangle = \frac{-4}{3\sqrt{2}} \end{cases},$$

donc $\theta = -\text{Arccos} \frac{1}{3}$. En conclusion, ϕ est la rotation d'axe Δ orienté par \vec{I} et d'angle $-\text{Arccos} \frac{1}{3}$.

Remarque : la relation $1 + 2 \cos \theta = \text{Tr} A$ donne aussi $\cos \theta = \frac{1}{3}$.

□ **Méthode 17.7.—** Comment calculer l'image $s_F(\vec{x})$ de $\vec{x} \in \mathbf{R}^n$ par une symétrie orthogonale par rapport à un sous-espace vectoriel F de \mathbf{R}^n si $n \leq 3$

- On peut utiliser la formule du projeté sur un espace de dimension finie, après avoir déterminé une base orthogonale $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ de F : $p_F(\vec{x}) = \sum_{k=1}^p \frac{\langle \vec{u}_k, \vec{x} \rangle}{\langle \vec{u}_k, \vec{u}_k \rangle} \vec{u}_k$.

Puis on utilise $\vec{x} + s_F(\vec{x}) = 2p_F(\vec{x})$.

- Si F est un plan (cas où $n = 3$), il peut être avantageux de passer par la projection orthogonale sur F^\perp , et par la formule $s_F = -s_{F^\perp}$.
- On peut aussi faire un changement de base orthonormée vers une base adaptée à F , mais les calculs sont en général fastidieux.

Exemple : Déterminer la matrice de la réflexion s de \mathbf{R}^3 par rapport au plan P d'équation cartésienne $x + y + z = 0$.

Un vecteur normal à P est $\vec{n} = (1, 1, 1)$. Le projeté orthogonal de $\vec{v} = (x, y, z)$ sur $\text{Vect}(\vec{n})$ est

$$\vec{h} = \frac{\langle \vec{v}, \vec{n} \rangle}{\|\vec{n}\|^2} \vec{n} = \frac{x + y + z}{3} \vec{n}.$$

Le symétrique $s(\vec{v})$ de \vec{v} par rapport à P vérifie $s(\vec{v}) = \vec{v} - 2\vec{h}$, donc

$$s(\vec{v}) = \frac{1}{3} (x - 2y - 2z, -2x + y - 2z, -2x - 2y + z)$$

et la matrice de s est : $S = \frac{1}{3} \begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$.

□ **Méthode 17.8.—** Comment déterminer la matrice (dans $\mathcal{B} = \{\vec{i}, \vec{j}, \vec{k}\}$ orthonormée directe) d'une rotation vectorielle de \mathbf{R}^3 définie par ses éléments géométriques

Soit r la rotation d'axe $\text{Vect}(\vec{a})$, orienté par \vec{a} , et d'angle θ . Pour obtenir la matrice de r dans la base canonique \mathcal{B} , on peut :

- Poser $\vec{I} = \frac{1}{\|\vec{a}\|} \vec{a}$, choisir \vec{J} unitaire et orthogonal à \vec{a} , puis poser $\vec{K} = \vec{I} \wedge \vec{J}$.

La famille $\mathcal{B}' = (\vec{I}, \vec{J}, \vec{K})$ est une base orthonormée directe, relativement à laquelle

la matrice de r est de la forme "réduite" $R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$.

- Si $P_{\mathcal{B}}^{\mathcal{B}'}$ est la matrice de passage de \mathcal{B} à \mathcal{B}' , $P_{\mathcal{B}}^{\mathcal{B}'}$ est orthogonale. Donc : $(P_{\mathcal{B}}^{\mathcal{B}'})^{-1} = (P_{\mathcal{B}}^{\mathcal{B}'})^T$. Ainsi $M = P_{\mathcal{B}}^{\mathcal{B}'} R (P_{\mathcal{B}}^{\mathcal{B}'})^T$ est la matrice cherchée.

Exemple : Déterminer la matrice dans la base canonique \mathcal{B} de la rotation r d'angle $\frac{\pi}{3}$ autour de l'axe dirigé et orienté par $\vec{a} = (1, 1, 1)$.

Posons $\vec{I} = \frac{1}{\sqrt{3}}(1, 1, 1)$, $\vec{J} = \frac{1}{\sqrt{2}}(1, 0, -1)$, puis $\vec{K} = \vec{I} \wedge \vec{J} = \frac{1}{\sqrt{6}}(-1, 2, -1)$. $\mathcal{B}' = (\vec{I}, \vec{J}, \vec{K})$ est une base orthonormée directe, relativement à laquelle la matrice de r est de la forme « réduite »

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \frac{\pi}{3} & -\sin \frac{\pi}{3} \\ 0 & \sin \frac{\pi}{3} & \cos \frac{\pi}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \text{ et } P_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix}.$$

$$\begin{aligned} M = P_{\mathcal{B}'}^{\mathcal{B}} R (P_{\mathcal{B}'}^{\mathcal{B}})^T &= \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix}. \end{aligned}$$

Chapitre 18

Vector functions and parameterized curves

■ Objectifs

■ Les incontournables :

- ▶ Savoir montrer qu'une fonction vectorielle est (ou n'est pas) dérivable, de classe \mathcal{C}^1 ou \mathcal{C}^k .
- ▶ Savoir faire un développement limité d'ordre 1 autour d'un point d'une fonction vectorielle si cette fonction est dérivable en ce point.
- ▶ Savoir dériver des applications bilinéaires, des applications composées avec des applications bilinéaires.
- ▶ Savoir étudier une courbe paramétrée plane, en commençant par la diminution de l'intervalle d'étude.
- ▶ Savoir déterminer les tangentes éventuelles en un point régulier d'une courbe paramétrée plane.
- ▶ Savoir étudier la position d'une courbe paramétrée par rapport à une asymptote oblique.
- ▶ Savoir calculer la longueur d'une courbe.

■ Et plus si affinités ...

- ▶ Savoir déterminer les tangentes éventuelles et la position du support de l'arc par rapport à cette tangente en un point non régulier d'une courbe paramétrée plane.
- ▶ Savoir dériver un produit vectoriel ou un déterminant.
- ▶ Faire le lien entre les notions mathématiques d'arc et de courbe, et leurs pendants cinématiques.
- ▶ Savoir appliquer la formule de Leibniz pour les fonctions vectorielles.

Résumé de cours

Ici $n \geq 1$. Les fonctions étudiées dans ce chapitre sont des applications d'un intervalle I de \mathbf{R} , non vide et non réduit à un point, à valeurs dans \mathbf{R}^n . On les appelle des **fonctions vectorielles** (dans le cas particulier $n = 1$, on parle alors aussi de **fonctions numériques**). Les résultats du chapitre *Espaces vectoriels normés de dimension finie* sur les limites et la continuité des fonctions de I dans \mathbf{R}^n s'appliquent évidemment ici.

f et g désignent dans tout le résumé de cours, sauf précision, des fonctions de I dans \mathbf{R}^n et $a \in I$.

■ Dérivée en un point, fonctions de classe \mathcal{C}^1

Définition : Dérivabilité en un point — f est **dérivable en a** si, et seulement si, l'application définie sur $I \setminus \{a\}$ par :

$$t \mapsto \frac{1}{t-a} (f(t) - f(a))$$

admet un vecteur \vec{A} de \mathbf{R}^n pour limite en a .

\vec{A} est la **dérivée de f au point a** et ce vecteur est noté $f'(a)$, $Df(a)$, ou encore $\frac{df}{dt}(a)$.

On définit de même les dérivées à droite et à gauche de f en a par les limites à droite et à gauche, si elles existent, de la même application. Elles sont notées $f'_d(a)$ ou $Df(a^+)$ ou encore $\frac{df}{dt}(a^+)$. (Notations analogues à gauche).

Remarque : Soit $a \in I$ qui n'est pas une de ses extrémités (un point **intérieur** à I), f est dérivable en a si, et seulement si, f est dérivable à droite en a , dérivable à gauche en a et $f'_d(a) = f'_g(a)$.

Proposition 18.1.— Caractérisation par la dérivabilité des applications composantes —.

Soit $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ une base de \mathbf{R}^n et $(f_i)_{i \in [1, n]}$ les applications composantes de f relativement à cette base : pour tout $t \in I$, $f(t) = \sum_{i=1}^n f_i(t) \vec{e}_i$.

f est dérivable en a si, et seulement si, ses applications composantes sont dérivables en a .

Dans ce cas, on a : $f'(a) = \sum_{i=1}^n f'_i(a) \vec{e}_i$.

Proposition 18.2.— Dérivabilité et existence d'un développement limité à l'ordre 1 —.

f est dérivable en a si, et seulement si, il existe un élément \vec{A} de \mathbf{R}^n et une fonction vectorielle $\vec{\varepsilon}$ tels que l'on ait, au voisinage de a :

$$f(t) = f(a) + (t-a)\vec{A} + (t-a)\vec{\varepsilon}(t),$$

où $\lim_{t \rightarrow a} \vec{\varepsilon}(t) = 0$. Dans ce cas, $f'(a) = \vec{A}$.

Corollaire 18.3.— Si f est dérivable en a , alors f est continue en a .

Remarque : La réciproque est fautive, comme le montre l'exemple de la fonction $x \rightarrow |x|$ (ici $n = 1$) qui est continue en 0 sans être dérivable en ce point.

Interprétation cinématique : Si l'on interprète le paramètre t comme le temps, le vocabulaire change. $\Gamma = f(I)$ est la *trajectoire* du point mobile $f(t)$. La fonction f est la *loi horaire du mouvement*.

$f'(a)$ est le vecteur vitesse du point mobile à l'instant a .

Définition : **Dérivabilité sur un intervalle, application dérivée** —. Une fonction f est **dérivable sur un intervalle** lorsqu'elle est dérivable en tout point de cet intervalle.

L'application associant à chaque point de cet intervalle la dérivée de la fonction en ce point est l'**application dérivée** ou **dérivée** de cette fonction. On la note : f' ou Df ou $\frac{df}{dt}$.

De plus, f est **de classe** \mathcal{C}^1 sur I lorsqu'elle est dérivable sur I et lorsque f' est continue sur I . L'ensemble de ces applications est noté $\mathcal{C}^1(I, \mathbf{R}^n)$.

Remarque : La dérivabilité et le caractère \mathcal{C}^1 , au même titre que la continuité, sont des **propriétés locales** d'une application, dans le sens où elles sont d'abord définies en un point, pour être ensuite étendues à un intervalle en disant qu'elles sont vraies en tout point de cet intervalle.

Proposition 18.4.— Caractérisation par le caractère \mathcal{C}^1 des applications composantes —. Soit $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ une base de \mathbf{R}^n et $(f_i)_{i \in \llbracket 1, n \rrbracket}$ les applications composantes de f relativement à cette base. f est une application de classe \mathcal{C}^1 de I dans \mathbf{R}^n si, et seulement si, chaque application composante de f est une application de classe \mathcal{C}^1 de I dans \mathbf{R} . Dans ce cas, on a : $\forall t \in I, f'(t) = \sum_{i=1}^n f'_i(t) \vec{e}_i$

On sait qu'un repère orthonormé du plan étant fixé, on peut « assimiler » \mathbf{R}^2 et \mathbf{C} .

Corollaire 18.5.— Cas particulier d'une fonction à valeurs dans \mathbf{C} —. Soit $f \in \mathcal{F}(I, \mathbf{C})$. Les propositions suivantes sont équivalentes :

- f est de classe \mathcal{C}^1 sur I ;
- $\Re f$ et $\Im f$ sont de classe \mathcal{C}^1 sur I ;
- \overline{f} est de classe \mathcal{C}^1 sur I .

Dans ces conditions, $f' = (\Re f)' + i(\Im f)'$ et $(\overline{f})' = \overline{f'}$.

Proposition 18.6.— Linéarité de la dérivation —. Si f et g sont dérivables sur un intervalle I et si λ et μ sont deux éléments quelconques de \mathbf{R} , alors la fonction $\lambda f + \mu g$ est dérivable sur I et $(\lambda f + \mu g)' = \lambda f' + \mu g'$.

Proposition 18.7.— Espace vectoriel $\mathcal{C}^1(I, \mathbf{R}^n)$ —. $\mathcal{C}^1(I, \mathbf{R}^n)$ est un espace vectoriel sur \mathbf{R} .

Proposition 18.8.— Si I et J sont deux intervalles de \mathbf{R} , ϕ une application dérivable de J dans \mathbf{R} , et f une fonction vectorielle dérivable de I dans \mathbf{R}^n alors l'application $f \circ \phi$ est dérivable sur J et : $\forall t \in J, (f \circ \phi)'(t) = \phi'(t) \times (f' \circ \phi)(t)$.

■ Applications linéaires, applications bilinéaires et dérivabilité

Proposition 18.9.— Composée d'une application linéaire et d'une application dérivable —.

Soit $p \in \mathbf{N}^*$ et L une application linéaire de \mathbf{R}^n dans \mathbf{R}^p .

- Si f est dérivable en a , alors l'application $L \circ f$ est dérivable en a et $(L \circ f)'(a) = (L \circ f')(a)$.
- Si f est de classe \mathcal{C}^1 sur I , alors l'application $L \circ f$ l'est aussi et $(L \circ f)' = L \circ f'$ sur I .

Proposition 18.10.— Composée d'une application bilinéaire et d'applications dérivables —.

Soit $(p, q) \in [\mathbf{N}^*]^2$ et B une application bilinéaire de $\mathbf{R}^p \times \mathbf{R}^q$ dans \mathbf{R}^n , f une fonction vectorielle de I dans \mathbf{R}^p et g une fonction vectorielle de I dans \mathbf{R}^q . On définit l'application $B(f, g)$ en posant : pour tout t de I , $B(f, g)(t) = B(f(t), g(t))$.

- Si f et g sont dérivables en a , alors l'application $B(f, g)$ est dérivable en a et

$$B(f, g)'(a) = B(f'(a), g(a)) + B(f(a), g'(a)).$$

- Si f et g sont de classe \mathcal{C}^1 sur I , alors l'application $B(f, g)$ l'est aussi et

$$B(f, g)' = B(f', g) + B(f, g').$$

Applications

1. Dérivée du produit de deux fonctions numériques dérivables

Si f et g sont deux fonctions dérivables de I dans \mathbf{R} , puisque $\begin{cases} \mathbf{R} \times \mathbf{R} & \rightarrow \mathbf{R} \\ (x, y) & \mapsto xy \end{cases}$ est bilinéaire, on obtient : fg est dérivable sur I et $(fg)' = f'g + fg'$.

2. Dérivée du produit d'une fonction numérique dérivable par une fonction vectorielle dérivable

Si φ est une fonction dérivable sur I à valeurs dans \mathbf{R} et f une fonction dérivable sur I à valeurs dans \mathbf{R}^n , \mathbf{K} -espace vectoriel normé de dimension finie, puisque $\begin{cases} \mathbf{R} \times \mathbf{R}^n & \rightarrow \mathbf{R}^n \\ (\lambda, x) & \mapsto \lambda x \end{cases}$ est bilinéaire, on obtient : φf est dérivable sur I et $(\varphi f)' = \varphi' f + \varphi f'$.

3. Dérivée du produit ou du carré scalaire de fonctions vectorielles à valeurs dans \mathbf{R}^n

Si f et g sont deux fonctions dérivables sur I à valeurs dans \mathbf{R}^n puisque

$$\begin{cases} \mathbf{R}^n \times \mathbf{R}^n & \rightarrow \mathbf{R} \\ (\vec{x}, \vec{y}) & \mapsto \langle \vec{x}, \vec{y} \rangle \end{cases} \text{ est bilinéaire, on obtient que } \langle f, g \rangle \text{ est dérivable et}$$

$$[\langle f, g \rangle]' = \langle f', g \rangle + \langle f, g' \rangle.$$

De plus, $\|f\|^2 = \langle f, f \rangle$ donc $(\|f\|^2)' = 2\langle f, f' \rangle$.

Corollaire 18.11.— f est de norme constante (déplacement sur une sphère) si et seulement si f et f' sont orthogonaux.

4. Dérivée du déterminant de fonctions vectorielles à valeurs dans \mathbf{R}^2

Si f et g sont dérivables sur I à valeurs dans \mathbf{R}^2 alors puisque $\begin{cases} \mathbf{R}^2 \times \mathbf{R}^2 & \rightarrow \mathbf{R} \\ (\vec{x}, \vec{y}) & \mapsto \text{Det}(\vec{x}, \vec{y}) \end{cases}$ est bilinéaire, $\text{Det}(f, g)$ est dérivable et $[\text{Det}(f, g)]' = \text{Det}(f', g) + \text{Det}(f, g')$.

■ Fonctions de classe \mathcal{C}^k

Définition : Fonctions de classe \mathcal{C}^k —. Soit k un entier naturel. f est dite de classe :

- \mathcal{C}^0 sur I si elle est continue sur I ,
- \mathcal{C}^{k+1} sur I si elle est dérivable sur I et si sa dérivée f' est de classe \mathcal{C}^k sur I ,
- \mathcal{C}^∞ sur I si elle est de classe \mathcal{C}^k sur I , pour tout entier naturel k .

Notation : $\mathcal{C}^k(I, \mathbf{R}^n)$ désigne l'ensemble des fonctions de classe \mathcal{C}^k sur I , à valeurs dans \mathbf{R}^n .

Il est clair que : $\forall k \in \mathbf{N}, \mathcal{C}^k(I, \mathbf{R}^n) \supset \mathcal{C}^{k+1}(I, \mathbf{R}^n) \supset \mathcal{C}^\infty(I, \mathbf{R}^n)$.

Définition : Dérivées successives —. Si $f \in \mathcal{C}^k(I, \mathbf{R}^n)$, avec $k \geq 2$, la dérivée $k^{\text{ème}}$ de f sur I est l'application $f^{(k)}$ définie par : $f^{(0)} = f$ et $\forall i \in \llbracket 1, k \rrbracket, f^{(i)} = (f^{(i-1)})'$.

Proposition 18.12.— Espace vectoriel $\mathcal{C}^k(I, \mathbf{R}^n)$ —.

Pour $0 \leq k \leq +\infty$, $(\mathcal{C}^k(I, \mathbf{R}^n), +, \cdot)$ est un \mathbf{R} -espace vectoriel.

Proposition 18.13.— Caractérisation par le caractère \mathcal{C}^k des applications composantes —.

Soit $b = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ une base de \mathbf{R}^n et $(f_i)_{i \in \llbracket 1, n \rrbracket}$ les applications composantes de f relativement à cette base : pour tout $t \in I$, $f(t) = \sum_{i=1}^n f_i(t) \vec{e}_i$.

f est une application de classe \mathcal{C}^k de I dans \mathbf{R}^n si, et seulement si, chaque application composante de f est une application de classe \mathcal{C}^k de I dans \mathbf{R} . Dans ce cas, on a :

$$\forall p \in \llbracket 1, k \rrbracket, \forall t \in I, f^{(p)}(t) = \sum_{i=1}^n f_i^{(p)}(t) \vec{e}_i.$$

Proposition 18.14.— Composition d'applications de classe \mathcal{C}^k —.

Soient I et J deux intervalles de \mathbf{R} , φ une fonction de classe \mathcal{C}^k de J dans I et f une application de classe \mathcal{C}^k de I dans \mathbf{R}^n alors l'application $f \circ \varphi$ est de classe \mathcal{C}^k sur J .

Proposition 18.15.— Prolongement en un point du caractère \mathcal{C}^k —. Soit $f \in \mathcal{F}([a, b], \mathbf{R}^n)$ une fonction vectorielle. Si f est continue sur $[a, b]$, de classe \mathcal{C}^k sur $]a, b]$ et telle que :

$$\forall p \in \llbracket 1, k \rrbracket, \lim_{t \rightarrow a} f^{(p)}(t) = \vec{l}_p \in \mathbf{R}^n$$

alors f est de classe \mathcal{C}^k sur $[a, b]$ et $\forall p \in \llbracket 1, k \rrbracket, f^{(p)}(a) = \vec{l}_p$.

■ Arcs paramétrés de classe \mathcal{C}^k . Étude locale

L'objectif et l'application principale est l'étude de courbes dans le plan euclidien orienté \mathbf{R}^2 . Ceci dit, les notions de ce paragraphe sont valables dans le cas où $n \geq 3$. Sauf mention contraire, si $n = 2$, on le suppose muni du repère orthonormal $(O; \vec{i}, \vec{j})$, et si $n = 3$, on se place dans le repère orthonormal $(O; \vec{i}, \vec{j}, \vec{k})$. Si $n = 3$, on parle de courbe gauche ; si $n = 2$, on parle de courbe plane. Néanmoins, une courbe définie dans \mathbf{R}^3 peut être incluse dans un plan et donc devenir plane !

Définition : Si $k \in \mathbf{N}^* \cup \{+\infty\}$, un **arc paramétré** Γ de classe \mathcal{C}^k est la donnée d'un intervalle I de \mathbf{R} et d'une fonction vectorielle $\vec{\gamma} : I \rightarrow \mathbf{R}^n$ de classe \mathcal{C}^k .

On appelle **courbe associée à l'arc** Γ l'image de l'application $\vec{\gamma}$, c'est-à-dire l'ensemble des points M de \mathbf{R}^n pour lesquels il existe $t \in I$ tel que $\vec{OM}(t) = \vec{\gamma}(t)$.

Donc si $n = 2$, $\vec{OM}(t) = x(t)\vec{i} + y(t)\vec{j}$, où x et y sont des fonctions de classe \mathcal{C}^k et si $n = 3$, alors $\vec{OM}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$, où x , y et z sont des fonctions de classe \mathcal{C}^k .

Par exemple, les cercles et les droites se paramètrent très bien, le faire !

Définition : On considère un arc paramétré Γ , défini par un intervalle I et une fonction vectorielle $\vec{\gamma} : I \rightarrow \mathbf{R}^n$ de classe \mathcal{C}^k et un élément t_0 de I .

Un point de paramètre $t_0 \in I$ est dit **régulier** si $\vec{\gamma}'(t_0) \neq \vec{0}$.

En un tel point, la droite dirigée par $\vec{\gamma}'(t_0)$ et passant par le point $M(t_0)$ s'appelle la **tangente** à l'arc Γ au point de paramètre $M(t_0)$.

Définition : Un point de paramètre t_0 est dit **singulier** ou **stationnaire** si $\vec{\gamma}'(t_0) = \vec{0}$.

Voir la **méthode 18.4** pour l'étude locale autour d'un point non régulier.

■ Étude des branches infinies d'une courbe plane

Définition : Soit $t_0 \in \overline{\mathbf{R}}$ une extrémité de l'intervalle I . On dit que l'arc Γ présente en t_0 une **branche infinie** si $\lim_{t \rightarrow t_0} \|\vec{OM}(t)\| = +\infty$. On pose $\vec{OM}(t) = f(t)\vec{i} + y(t)\vec{j}$.

- Si $\lim_{t \rightarrow t_0} x(t) = \alpha$ et $\lim_{t \rightarrow t_0} y(t) = \pm\infty$, on dit que la droite $x = \alpha$ est **asymptote** à la courbe.
- Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ et $\lim_{t \rightarrow t_0} y(t) = \beta$, la droite $y = \beta$ est **asymptote** à la courbe.
- Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ et $\lim_{t \rightarrow t_0} y(t) = \pm\infty$, on forme le rapport $f(t) = \frac{y(t)}{x(t)}$.
- ▷ Si $\lim_{t \rightarrow t_0} f(t)$ n'existe pas, Γ admet une **branche infinie**.
- ▷ Si $\lim_{t \rightarrow t_0} f(t) = \pm\infty$, Γ présente une **branche parabolique de direction** (Oy) .
- ▷ Si $\lim_{t \rightarrow t_0} f(t) = 0$, Γ présente une **branche parabolique de direction** (Ox) .
- ▷ Si $\lim_{t \rightarrow t_0} f(t) = a \neq 0$, on forme alors la différence $g(t) = y(t) - ax(t)$.
- Si $\lim_{t \rightarrow t_0} g(t)$ n'existe pas, Γ présente une **branche asymptotique de pente** a .
- Si $\lim_{t \rightarrow t_0} g(t) = \pm\infty$, Γ présente une **branche parabolique de pente** a .
- Si $\lim_{t \rightarrow t_0} g(t) = b$, la droite d'équation $y = ax + b$ est **asymptote** à l'arc Γ .

■ Longueur d'un arc paramétré de classe \mathcal{C}^1

Soit un arc de classe \mathcal{C}^1 paramétré par :

$$\begin{array}{ccc} \vec{\gamma} : I & \longrightarrow & \mathbf{R}^2 \\ t & \longmapsto & \vec{OM}(t) = x(t)\vec{i} + y(t)\vec{j} \end{array} \quad \text{ou} \quad \begin{array}{ccc} \vec{\gamma} : I & \longrightarrow & \mathbf{R}^3 \\ t & \longmapsto & \vec{OM}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k} \end{array}$$

Définition : Soit $(t_1, t_2) \in I^2$. La **longueur** de l'arc $M(t_1)M(t_2)$ que décrit le point $M(t)$ lorsque t parcourt le segment $[t_1, t_2]$ est le réel : $\int_{t_1}^{t_2} \|\vec{\gamma}'(t)\| dt$.

Méthodes

■ Dérivabilité des fonctions vectorielles

Commencer par revoir la méthode ?? et la méthode ??, vues dans le paragraphe consacré aux fonctions vectorielles du chapitre sur les espaces vectoriels normés de dimension finie.

□ **Méthode 18.1.**— Comment montrer la dérivabilité et calculer la dérivée d'une fonction vectorielle f de I dans \mathbf{R}^n

- Si l'on connaît ou l'on "soupçonne" $\vec{l} = f'(t_0)$, on peut utiliser la définition de la dérivée, où $\|\cdot\|$ est une norme bien choisie sur \mathbf{R}^n :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall t \in I \cap]t - \eta, t + \eta[, \left\| \frac{f(t) - f(t_0)}{t - t_0} - \vec{l} \right\| < \varepsilon.$$

- On peut montrer la dérivabilité de chaque fonction composante $t \mapsto x_i(t)$ si l'on pose $f(t) = (x_1(t), \dots, x_n(t))$. Et alors : $f'(t) = (x'_1(t), \dots, x'_n(t))$. Ainsi, f est dérivable en $t_0 \in I$ si, et seulement si, chaque fonction composante l'est en t_0 .
- On peut aussi utiliser un développement limité d'ordre 1 en s'aidant de la **proposition 18.2**. Voir les exemples de la **méthode 18.2**, où on l'applique.
- Si l'expression de f est "costaud" on peut utiliser la **proposition 18.10**.

Exemples : 1. Soit $f :]-1, 1[\rightarrow \mathbf{R}^2, t \mapsto \begin{cases} (0, 0) & \text{si } -1 < t \leq 0 \\ (t^2 \sin \frac{1}{t}, t^2 \cos \frac{1}{t}) & \text{si } 0 < t < 1 \end{cases}$.

a. Montrer que f est dérivable sur $] - 1, 1[$. f est-elle deux fois dérivable ?

b. Montrer que $f(] - 1, 1[)$ est la réunion de $(0, 0)$ et d'une partie bornée de \mathbf{R}^2 située en dehors de la boule $\mathcal{B}(O, 1)$.

a. Comme $t \mapsto t^2 \sin \frac{1}{t}$ et $t \mapsto t^2 \cos \frac{1}{t}$ sont dérivables sur $] - 1, 1[\setminus \{0\}$, il en est de même de f . On peut même ajouter que, les fonctions composantes étant de classe \mathcal{C}^∞ sur $] - 1, 1[\setminus \{0\}$, il en est de même de f . Maintenant, si $t \neq 0$, $\frac{f(t) - f(0)}{t - 0} = \begin{pmatrix} t \sin \frac{1}{t} \\ t \cos \frac{1}{t} \end{pmatrix}$. Comme chaque composante tend vers 0 quand t tend vers 0, f est dérivable en 0 avec $f'(0) = (0, 0)$. On peut remarquer que par contre $\lim_{t \rightarrow 0} \frac{f'(t) - f'(0)}{t - 0}$ n'existe pas ! Donc f n'est pas deux fois dérivable en 0.

b. $\forall t \in] - 1, 1[\setminus \{0\}, \|f'(t)\|^2 = (2t \sin \frac{1}{t} - \cos \frac{1}{t})^2 + (2t \cos \frac{1}{t} + \sin \frac{1}{t})^2 = 4t^2 + 1 \in]1, 5[$.

2. Soit $f : [a, b] \rightarrow \mathbf{R}^n$ continue sur $[a, b]$ et dérivable sur $]a, b[$. Montrer que f est constante sur $[a, b]$ si et seulement si : $\forall t \in]a, b[, f'(t) = \vec{0}$.

Les n applications composantes $t \mapsto x_i(t)$ de f sont continues sur $[a, b]$ et dérivables sur $]a, b[$. f est constante si, et seulement si, chacune des fonctions composantes est constante. D'après le cours de première année sur la dérivation des fonctions numériques, cela équivaut à écrire que les n fonctions composantes ont toutes leur dérivée nulle sur $]a, b[$. On peut conclure.

❏ Méthode 18.2.— Comment exprimer et utiliser la dérivée d'une application bilinéaire ou d'un déterminant de fonctions vectorielles

- Dans le cas d'une application bilinéaire B de $\mathbf{R}^p \times \mathbf{R}^q$ dans \mathbf{R}^n , on applique tout simplement la formule de la **proposition 18.10**. La première difficulté est parfois d'identifier l'application bilinéaire.
- Dans le cas particulier d'un produit scalaire de deux fonctions vectorielles f et g dérivables de I dans \mathbf{R}^n : l'application $t \mapsto \langle f(t), g(t) \rangle$ est dérivable sur I et on peut appliquer sur I : $[\langle f, g \rangle]' = \langle f', g \rangle + \langle f, g' \rangle$.
- Dans le cas particulier d'un déterminant de deux fonctions vectorielles f et g dérivables de I dans \mathbf{R}^2 (ce qui découle du précédent par composition) : l'application $t \mapsto \text{Det}_{\mathcal{B}}(f(t), g(t))$ est dérivable sur I et on encore peut appliquer sur I la formule de la **proposition 18.10** : $[\text{Det}_{\mathcal{B}}(f, g)]' = \text{Det}_{\mathcal{B}}(f', g) + \text{Det}_{\mathcal{B}}(f, g')$.
- On peut aussi utiliser l'outil des développements limités. Ainsi par exemple si f est une fonction de classe \mathcal{C}^1 de $I \subset \mathbf{R}$ dans \mathbf{R}^p , et si $f(t) = (x_1(t), \dots, x_p(t))$ pour tout $t \in I$ alors : $x_i(t+h) = x_i(t) + hx'_i(h) + h\epsilon_{x_i}(h)$ pour tout $i \in \llbracket 1, p \rrbracket$, où $\lim_{h \rightarrow 0} \epsilon_{x_i}(h) = 0$. Et : $f(t+h) = f(t) + hf'(h) + h\vec{\epsilon}(h)$, où $\lim_{h \rightarrow 0} \vec{\epsilon}(h) = \vec{0}$, en posant le vecteur $\vec{\epsilon} = (\epsilon_{x_1}, \dots, \epsilon_{x_n})$. C'est l'application directe de la **proposition 13.2**. On généralisera cette formule à la **méthode 18.4** si $f \in \mathcal{C}^k(I, \mathbf{R}^2)$.

Exemple : Montrer que, \mathbf{R}^3 étant orienté par \mathcal{B} et f et g étant deux fonctions vectorielles, dérivables de I dans \mathbf{R}^3 alors $f \wedge g$ est dérivable sur I et on a sur I : $[f \wedge g]' = f' \wedge g + f \wedge g'$.

- On peut utiliser le fait que $t \mapsto f(t) \wedge g(t)$ est bilinéaire et donc utiliser directement la formule de la **proposition 18.10**.
- On peut aussi faire une preuve directe en utilisant des développements limités d'ordre 1 :

$$\forall t \in I, \forall h, f(t+h) = f(t) + hf'(t) + h\vec{\epsilon}_1(h) \text{ et } g(t+h) = g(t) + hg'(t) + h\vec{\epsilon}_2(h),$$

où $\vec{\epsilon}_1$ et $\vec{\epsilon}_2$ sont deux fonctions vectorielles définies dans un voisinage de 0 et tendant toutes les deux vers 0 quand $h \rightarrow 0$.

$f(t+h) \wedge g(t+h)$ se développe en la somme de $f(t) \wedge g(t)$, de $h[f'(t) \wedge g(t) + f(t) \wedge g'(t)]$, de $h[f(t) \wedge \vec{\epsilon}_2(h) + \vec{\epsilon}_1(h) \wedge g_1(t)]$ et $h^2[f'(t) \wedge g'(t) + \vec{\epsilon}_1(t) \wedge \vec{\epsilon}_2(t)]$.

Et les produits vectoriels : $f(t) \wedge \vec{\epsilon}_2(h) + \vec{\epsilon}_1(h) \wedge g_1(t)$ et $h(f'(t) \wedge g'(t) + \vec{\epsilon}_1(t) \wedge \vec{\epsilon}_2(t))$ tendent vers 0 quand h tend vers 0, (pour t fixé). On utilise pour montrer cela les résultats du **chapitre 7**. Ceci fait :

$$f(t+h) \wedge g(t+h) = f(t) \wedge g(t) + h[f'(t) \wedge g(t) + f(t) \wedge g'(t)] + h\vec{\epsilon}(h),$$

où $\lim_{h \rightarrow 0} \vec{\epsilon}(h) = 0$.

D'après la **proposition 18.2**, la dérivée de $t \mapsto f(t) \wedge g(t)$ est $t \mapsto f'(t) \wedge g(t) + f(t) \wedge g'(t)$.

2. a. Soit n fonctions vectorielles f_1, \dots, f_n , dérivables de I dans \mathbf{R}^n .

Montrer que l'application $t \mapsto \text{Det}_{\mathcal{B}}(f_1(t), \dots, f_n(t))$ est dérivable sur I .

Montrer alors que sur I : $[\text{Det}_{\mathcal{B}}(f_1(t), \dots, f_n(t))]' = \sum_{j=1}^n \text{Det}_{\mathcal{B}}(f_1(t), \dots, f'_j(t), \dots, f_n(t))$.

■ Étude et construction d'un arc plan

□ **Méthode 18.3.** — Comment réduire l'intervalle I de $\Gamma : \begin{cases} x = x(t) \\ y = y(t) \end{cases}, t \in I.$

- Si les fonctions x et y sont T -**périodiques**, il suffit d'étudier Γ sur un intervalle d'amplitude T , du type $[\alpha, \alpha + T[\cap I$ où $\alpha \in \mathbf{R}$ pour avoir toute la courbe.
- Si les fonctions x et y sont T -**antipériodiques**, c'est-à-dire :

$$\begin{cases} x(t+T) = -x(t) \\ y(t+T) = -y(t) \end{cases},$$

on étudie sur $[\alpha, \alpha + T[\cap I$ puis on fait une **symétrie de centre O** pour avoir toute la courbe.

- Si les fonctions x ou y sont T -**pseudopériodiques**, c'est-à-dire, par exemple,

$$\exists K, \begin{cases} x(t+T) = x(t) + K \\ y(t+T) = y(t) \end{cases},$$

on étudie sur $[\alpha, \alpha + T[\cap I$ puis le reste de la courbe se déduit par une succession de **translations de vecteurs $K\vec{i}$** .

Dans la suite de la méthode, on suppose que I est centré en $a \in \mathbf{R}$.

- Si l'on a : $\begin{cases} x(2a-t) = x(t) \\ y(2a-t) = -y(t) \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **symétrie orthogonale par rapport à (Ox)** .
- Si l'on a : $\begin{cases} x(2a-t) = -x(t) \\ y(2a-t) = y(t) \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **symétrie orthogonale par rapport à (Oy)** .
- Si l'on a : $\begin{cases} x(2a-t) = -x(t) \\ y(2a-t) = -y(t) \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **symétrie par rapport à O** .
- Si l'on a : $\begin{cases} x(2a-t) = y(t) \\ y(2a-t) = x(t) \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **symétrie orthogonale par rapport à $y = x$** .
- Si l'on a : $\begin{cases} x(2a-t) = x(t) + \alpha \\ y(2a-t) = y(t) + \beta \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **translation de vecteur $\alpha\vec{i} + \beta\vec{j}$** .
- Si l'on a : $\begin{cases} x(2a-t) = -x(t) + \alpha \\ y(2a-t) = y(t) \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **symétrie orthogonale par rapport à $x = \frac{\alpha}{2}$** .
- Si l'on a : $\begin{cases} x(2a-t) = x(t) \\ y(2a-t) = -y(t) + \beta \end{cases}$ alors on étudie Γ pour $t \in [a, +\infty[\cap I$ puis le reste de la courbe se déduit par **symétrie orthogonale par rapport à $y = \frac{\beta}{2}$** .

Le cas le plus fréquent est celui où $a = 0$ (alors I est centré en $t = 0$).

- Remarques :** 1. On a le même type de résultat en remplaçant $a - t$ par $\frac{1}{t}$.
 2. Le cas $a = 0$ correspond à x et y paires ou impaires.
 3. Quelquefois (on peut le voir dans les exemples développés), on a à composer plusieurs réductions. Il faut alors faire attention à la réduction de l'intervalle d'étude à chaque étape.

Exemples : Déterminer l'intervalle minimal d'étude et les symétries éventuelles des arcs :

$$\begin{array}{lll}
 1. \Gamma_1 : \begin{cases} x(t) = \frac{t}{1+t^3} \\ y(t) = \frac{t^2}{1+t^3} \end{cases} & 2. \Gamma_2 : \begin{cases} x(\theta) = \theta - \sin \theta \\ y(\theta) = \cos \theta - 1 \end{cases} & 3. \Gamma_3 : \begin{cases} x(t) = \cos^3 t + \sin t \\ y(t) = \sin^3 t + \cos t \end{cases} \\
 4. \Gamma_4 : \begin{cases} x(t) = \cos^3 t \\ y(t) = \sin^3 t \end{cases} & 5. \Gamma_5 : \begin{cases} x(t) = \sin 2t \\ y(t) = \sin 3t \end{cases} & 6. \Gamma_6 : \begin{cases} x(t) = \cos t \\ y(t) = \sin\left(\frac{t}{3}\right) \end{cases}
 \end{array}$$

1. Ici $I = \mathbf{R} \setminus \{-1\}$. Aucune période ou de parité n'apparaît. On constate que $x\left(\frac{1}{t}\right) = y(t)$ et que $y\left(\frac{1}{t}\right) = x(t)$. On en déduit que la partie de Γ_1 décrite lorsque t décrit $I_1 =]0, 1[$ [respectivement $I_2 =]-1, 0[$] est symétrique par rapport à la première bissectrice (droite d'équation $y = x$) de la partie correspondant à $J_1 =]1, +\infty[$ (respectivement $J_2 =]-\infty, -1[$).

Bilan : l'intervalle minimal d'étude est $I_1 \cup I_2$, on trace la partie du support de Γ correspondante à $I_1 \cup I_2$ et on effectue la symétrie orthogonale par rapport à $y = x$ pour avoir le support complet.

2. Ici pour tromper l'ennemi, la variable est θ et non plus t .

On a : $\begin{cases} x(\theta + 2\pi) = x(\theta) + 2\pi \\ y(\theta + 2\pi) = y(\theta) \end{cases}$, donc on étudie sur un intervalle du type $[\alpha, \alpha + 2\pi]$ puis le

reste de la courbe se déduit par une succession de translations de vecteurs $2\pi\vec{i}$.

Comme x est impaire et y est paire, on prend $\alpha = -\pi$, ce qui permet de considérer $[-\pi, \pi]$. Ainsi on réduit l'intervalle d'étude à $[0, \pi]$.

Bilan : l'intervalle d'étude minimal est $J = [0, \pi]$, on trace la partie du support de Γ_2 correspondante à J puis on effectue une symétrie orthogonale par rapport à (Oy) et enfin une succession de translations de vecteurs $2\pi\vec{i}$ pour avoir le support complet.

Remarque : Le lecteur un peu doué verra que $\begin{cases} x(2\pi - \theta) = 2\pi - x(\theta) \\ y(2\pi - \theta) = y(\theta) \end{cases}$, et il pourra faire le chemin suivant : étude sur $J = [0, \pi]$ puis on trace la partie du support de Γ_2 correspondante. Ensuite, on effectue la symétrie orthogonale d'axe $x = \pi$. Il reste à faire la succession de translations de vecteurs $2\pi\vec{i}$. Il n'y a donc pas unicité des chemins possibles pour obtenir tout l'arc.

3. Le domaine de définition est \mathbf{R} . Puis 2π est une période commune à x et à y . On étudie donc Γ_3 sur $[\alpha, \alpha + 2\pi]$, où α sera choisi par les autres symétries. Puis, π est une antipériode pour x et y , et on étudie donc Γ_3 sur $[\alpha, \alpha + \pi]$, et on complètera par une symétrie centrale de centre O . De plus, $\begin{cases} x\left(\frac{\pi}{2} - t\right) = y(t) \\ y\left(\frac{\pi}{2} - t\right) = x(t) \end{cases}$, et donc on centre $[\alpha, \alpha + \pi]$ en $\frac{\pi}{4}$ et on ramène l'étude à $J = \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$.

Bilan : on étudie puis on trace la partie de Γ_3 correspondante à J puis on effectue une symétrie orthogonale par rapport à $y = x$, puis une symétrie de centre O . On obtient alors Γ_3 complet.

4. Comme x et y sont 2π -périodiques, il suffit d'étudier l'arc sur un intervalle d'amplitude 2π .

On remarque ensuite que π est une antipériode pour x et y , on étudie sur $[\alpha, \alpha + \pi]$ puis on effectue une symétrie par rapport à O . Puis on exploite le fait que x est paire et y est impaire et donc on

réduit à $[0, \frac{\pi}{2}]$ puis on effectue une symétrie par rapport à (Ox) . Enfin : $\begin{cases} x(\frac{\pi}{2} - t) = y(t) \\ y(\frac{\pi}{2} - t) = x(t) \end{cases}$ et on réduit encore à $[0, \pi/4]$ puis on effectue une symétrie par rapport à $y = x$.

Bilan : on étudie puis on trace la partie de Γ_4 correspondante à $J = [0, \pi/4]$ puis on fait une symétrie orthogonale par rapport à $y = x$ puis une symétrie orthogonale par rapport à Ox et enfin une symétrie par rapport à O , ce qui fournit enfin le support de Γ_4 complet.

5. x et y sont définies dans \mathbf{R} . L'arc Γ_5 fait partie des courbes de Lissajous (quand vous tracerez le support, vous penserez au tracé qu'on a avec un oscilloscope !) On commence par remarquer que 2π est une période commune à x et à y . On obtient donc toute la courbe Γ_5 quand t décrit $[\alpha, \alpha + 2\pi]$, où α va être précisé. Comme x et y sont impaires, on prend $\alpha = -\pi$ et on se ramène à $[0, \pi]$. Il reste

à faire une symétrie de centre O pour obtenir l'arc sur $[-\pi, \pi]$. Enfin, $\begin{cases} x(\pi - t) = -x(t) \\ y(\pi - t) = y(t) \end{cases}$ et donc l'intervalle minimal d'étude est $[0, \pi/2]$. Il reste une symétrie d'axe Oy pour revenir à $[0, \pi]$.

Bilan : on étudie puis on trace la partie de l'arc Γ_5 correspondant à $J = [0, \pi/2]$ puis on effectue successivement une symétrie orthogonale par rapport à Oy puis une symétrie centrale de centre O . On a alors l'arc Γ_5 complet.

6. La plus petite période commune de x et y est 6π : l'arc est entièrement décrit lorsque t décrit un intervalle d'amplitude 6π , $[-3\pi, 3\pi]$ est choisi, vu la suite. x est paire et y impaire : $M(-t)$ est le symétrique de $M(t)$ par rapport à l'axe des abscisses. Il suffit de tracer l'arc pour t décrivant $[0, 3\pi]$, puis de symétriser la portion obtenue par rapport à l'axe des abscisses pour obtenir la totalité de l'arc. Puis, $\forall t \in [0, 3\pi]$, $x(3\pi - t) = -x(t)$ et $y(3\pi - t) = y(t)$: $M(3\pi - t)$ est le symétrique de $M(t)$ par rapport à l'axe des ordonnées. Il suffit de tracer l'arc pour t décrivant $[0, \frac{3\pi}{2}]$, puis de symétriser la portion obtenue par rapport à l'axe des ordonnées pour obtenir la portion de l'arc obtenu lorsque t décrit $[0, 3\pi]$.

Bilan : on étudie puis on trace la partie de Γ_6 correspondant à $J = [0, \frac{3\pi}{2}]$ puis on fait successivement la symétrie orthogonale par rapport à (Oy) puis la symétrie orthogonale par rapport à (Ox) . On a alors le tracé de Γ_6 complet.

□ Méthode 18.4.— Comment déterminer la tangente en un point $t_0 \in I$ du support d'un arc paramétré plan $\Gamma = (I, f)$ de classe \mathcal{C}^1 , où $f(t) = (x(t), y(t))$ et connaître la position de ce support par rapport à sa tangente D

- Si $x'(t_0) \neq 0$ et $y'(t_0) \neq 0$, $M(t_0)$ est régulier et la tangente est portée par $f'(t_0)$.
- Si $x'(t_0) = 0$ et $y'(t_0) \neq 0$, $M(t_0)$ est régulier et la tangente est verticale.
- Si $x'(t_0) \neq 0$ et $y'(t_0) = 0$, $M(t_0)$ est régulier et la tangente est horizontale.
- Si $x'(t_0) = 0$ et $y'(t_0) = 0$, $M(t_0)$ n'est pas régulier. La pente de la tangente est $\lim_{t \rightarrow t_0} \frac{y'(t)}{x'(t)}$ si cette limite existe. On peut alors en déduire une équation de D .

Par ailleurs, les symétries du support de l'arc permettent d'obtenir généralement le comportement de la courbe par rapport à sa tangente. S'il faut aller plus loin, on effectue un développement limité de x et de y , au voisinage de t_0 , à un ordre suffisamment élevé (en général 3 ou 4), ce qui permet d'avoir l'allure autour de $M(t_0)$ et la position de Γ par rapport à D . Pour cela, on peut se placer dans un repère local de centre $M(t_0)$. Voir les exemples des courbes C_3 à C_5 qui suivent.

Remarques : 1. Dans le cas d'un point régulier, un développement limité de la fonction vectorielle f au voisinage de t_0 peut donner l'expression de $f'(t_0)$. De plus, on peut obtenir une équation (de la forme $y = ax + b$) de cette tangente D connaissant $M(t_0) \in D$ et sa direction $f'(t_0)$.

2. On se retrouve obligatoirement dans une des quatre situations suivantes.

Planche 1. Point ordinaire, la courbe Γ reste du même côté par rapport à la tangente.

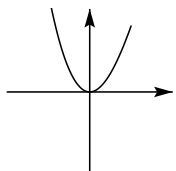


Planche 2. Point d'inflexion, la courbe Γ traverse la tangente en O .

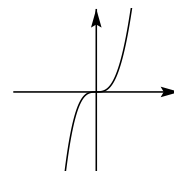


Planche 3. Point de rebroussement de première espèce, Γ est de part et d'autre de la tangente.

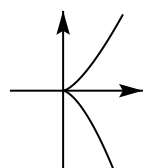
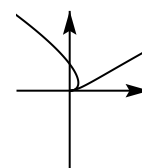


Planche 4. Point de rebroussement de deuxième espèce, Γ reste du même côté de la tangente.



(Pour simplifier, la tangente est portée par l'axe horizontal, le lecteur effectuera la rotation nécessaire dans la pratique de l'arc étudié.) Les cas de la planche 1 et de la planche 2 correspondent aux points réguliers et la planche 1 est de loin la plus fréquente. Le cas de la planche 3 se rencontre souvent dès que $M(t = t_0)$ n'est pas régulier. Les symétries de l'arc permettent en général de le détecter. Le cas de la planche 4 est plus marginal et c'est souvent l'étude complète de l'arc ou une étude locale soignée qui permet de le différencier de la planche 3.

Exemples : Étudier la tangente à l'arc paramétré plan au point t_0 pour les cas suivants :

$$1. C_1 : \begin{cases} x(t) = \frac{1}{\sin t} \\ y(t) = \frac{1}{\cos t} \\ t_0 = \frac{\pi}{4} \end{cases} \quad 2. C_2 : \begin{cases} x(t) = \sin^3 t \\ y(t) = \cos t - \cos^4 t \\ t_0 = 0. \end{cases} \quad 3. C_3 : \begin{cases} x(t) = 2t + t^2 \\ y(t) = 2t - \frac{1}{t^2} \\ t_0 = -1 \end{cases},$$

$$4. C_4 : \begin{cases} x(t) = e^{t-1} - t \\ y(t) = t^3 - 3t + 2 \\ t_0 = 1 \end{cases} \quad 5. C_5 : \begin{cases} x(t) = 3(\sin t - t) \\ y(t) = t^3 + t^5 \\ t_0 = 0. \end{cases}.$$

❑ **Méthode 18.5.—** Comment connaître la position d'une courbe par rapport à son asymptote oblique (quand t tend vers t_0 appartenant à l'adhérence de I).

- ▶ Il faut au préalable que $|x(t)|$ et $|y(t)|$ tendent simultanément vers $+\infty$ quand $t \rightarrow t_0$. Sinon, voir le résumé de cours.
- ▶ On détermine $\lim_{t \rightarrow t_0} \frac{y(t)}{x(t)}$. Si cette limite est nulle ou infini, on est en présence d'une branche parabolique. Si cette limite est $l \neq 0$, on continue.
- ▶ On détermine $\lim_{t \rightarrow t_0} y(t) - lx(t)$. Si cette limite est infinie, on a une direction asymptotique ($y = lx$). Si cette limite est p finie, on continue.
- ▶ La droite $\Delta : y = lx + p$ est alors asymptote oblique au support de l'arc Γ . On étudie alors le signe de $y(t) - lx(t) - p$.

Exemple : Étudier les branches infinies de l'arc paramétré $\mathcal{C} : x(t) = \frac{t}{t^2 - 1}, y(t) = \frac{t^2}{t - 1}$.

▶ Lorsque $t \rightarrow \pm\infty$, $|y(t)|$ tend vers $+\infty$ et $|x(t)|$ tend vers 0. Donc la droite d'équation $x = 0$ pour asymptote verticale.

▶ Lorsque t tend vers -1 , $|x(t)|$ tend vers $+\infty$ et $y(t)$ tend vers $-\frac{1}{2}$. La courbe \mathcal{C} admet la droite d'équation $y = -\frac{1}{2}$, comme asymptote horizontale.

▶ Lorsque t tend vers 1, $x(t)$ et $y(t)$ tendent simultanément vers ∞ .

On écrit alors : $\frac{y(t)}{x(t)} = \frac{t^2}{t-1} \times \frac{t^2-1}{t} = t(t+1)$, quantité qui tend vers 2 quand t tend vers 1.

Puis : $y(t) - 2x(t) = \frac{t(t+2)}{t-1}$, quantité qui tend vers $\frac{3}{2}$ quand t tend vers 1. Donc Γ admet la droite $\Delta : y = 2x + 3/2$ pour asymptote oblique. Pour la position de Δ par rapport à \mathcal{C} , on écrit :

$$y(t) - 2x(t) - \frac{3}{2} = \frac{t(t+2)}{t-1} - \frac{3}{2} = \frac{(t-1)(2t+3)}{2(t-1)}.$$

\mathcal{C} est donc en dessous de son asymptote quand t tend vers 1 par valeurs inférieures et au dessus quand t tend vers 1 par valeurs supérieures.

❑ **Méthode 18.6.—** Comment étudier puis tracer le support (dans un plan) d'un arc paramétré

- ▶ On commence par réduire au maximum l'intervalle d'étude. Voir la **méthode 18.3**.
- ▶ On calcule les dérivées x' et y' et on dresse le tableau de variation de x et de y l'un en dessous de l'autre.
- ▶ On cherche les tangentes horizontales ou verticales éventuelles. On étudie la pente de la tangente en un point non régulier avec la **méthode 18.4**.
- ▶ On étudie les branches infinies et notamment on recherche les équations des asymptotes (s'il y en a). Voir la **méthode 18.5**.
- ▶ On peut chercher les points doubles, c'est-à-dire des valeurs $t_1 \neq t_2$ telles que $M(t_1) = M(t_2)$. Les symétries suffisent parfois pour les trouver. Sinon, on les détermine. Ceci n'est en général possible que si x et y sont des fonctions trigonométriques simples ou des fonctions polynomiales de degré assez bas.
- ▶ On trace le support sur l'intervalle réduit (en dessinant les asymptotes et les tangentes particulières) puis on effectue les symétries (en respectant le bon ordre).

Exemple : Étudier et tracer la courbe paramétrée C_6 : $\begin{cases} x(t) = \cos t \\ y(t) = \sin(\frac{t}{3}) \end{cases}, t \in \mathbf{R}.$

► x et y sont définis sur \mathbf{R} et on a déjà développé les réductions éventuelles de l'étude. C'est le dernier exemple qui suit la **méthode 18.3**. On a vu qu'il suffit de tracer l'arc pour t décrivant $[0, \frac{3\pi}{2}]$, puis de symétriser la portion obtenue par rapport à l'axe des ordonnées pour obtenir la portion de l'arc obtenu lorsque t décrit $[0, 3\pi]$.

► On dresse le tableau de variation sur $J = [0, \frac{3\pi}{2}]$.

Les fonctions x et y sont dérivables sur J .

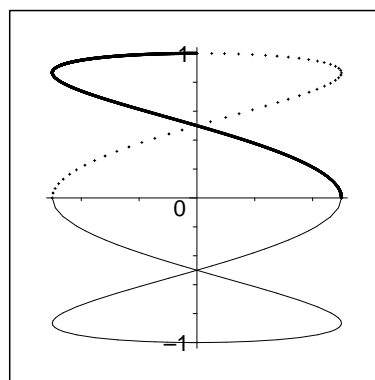
Alors : $\forall t \in J, x'(t) = -\sin t$ et $y'(t) = \frac{1}{3} \cos(\frac{t}{3})$, d'où le tableau de variation conjoint de x et y à droite.

► Comme x' s'annule en $t = 0$ et $t = \pi$, C_6 présente une tangente verticale en $M(0) = (1, 0)$ et $M(\pi) = (-1, \frac{\sqrt{3}}{2})$. Comme y' s'annule en $t = \frac{3\pi}{2}$, C_6 présente une tangente horizontale en $M(\frac{3\pi}{2}) = (0, 1)$.

t	0	π	$\frac{3\pi}{2}$		
$x'(t)$	0	-	0	+	1
$x(t)$	1	\searrow	-1	\nearrow	0
$y(t)$	0	\nearrow			1
$y'(t)$	$\frac{1}{3}$	+			0

► Ici pas de branches infinies. La courbe est contenue dans un carré.

► La partie de l'arc obtenue lorsque t décrit $[0, \frac{3\pi}{2}]$ figure en trait fort. Pour la tracer, on part de $M(0)$ avec une tangente verticale. Comme x décroît et y croît, on va vers la gauche tout en montant vers $M(\pi)$. On arrive en ce point avec une tangente verticale. Puis x et y croissent. On va vers la droite tout en montant jusqu'à $M(\frac{3\pi}{2})$, où la tangente est horizontale. Par symétrie par rapport à l'axe des ordonnées, on obtient, en pointillé, la partie de l'arc obtenue lorsque t décrit $[\frac{3\pi}{2}, 3\pi]$; par symétrie par rapport à l'axe des abscisses, on obtient, en trait plein, la partie de l'arc obtenue lorsque t décrit $[-3\pi, 0]$. Au total, on bien la totalité de l'arc qui est une **courbe de Lissajous**.



► Il y a des points doubles mais les symétries les fournissent immédiatement car en ces points, on traverse les axes. Il s'agit d'abord du point $M(\pi/2) = M(3\pi/2) = (0, 1/2)$ et ensuite (par symétrie) du point $M(-\pi/2) = M(-3\pi/2) = (0, -1/2)$.

□ Méthode 18.7.— Comment déterminer la longueur d'une courbe

Soit l'arc de classe \mathcal{C}^1 paramétré par : $\begin{matrix} \vec{\gamma} : I & \longrightarrow & \mathbf{R}^n \\ t & \longmapsto & \overrightarrow{OM(t)} = x_1(t)\vec{e}_1 + \dots + x_n(t)\vec{e}_n \end{matrix}$

La longueur entre $M(t_1)$ et $M(t_2)$ est donnée par : $\int_{t_1}^{t_2} \sqrt{\sum_{k=1}^n (x'_k)^2(t)} dt.$

La longueur totale est donnée par : $\int_I \sqrt{\sum_{k=1}^n (x'_k)^2(t)} dt.$

Il est conseillé de faire l'étude et le tracé de l'arc avant pour éventuellement utiliser les symétries pour réduire les bornes d'intégration.

Énoncé des exercices

■ ???

Exercice 18.1 :

Exercice 18.2 :

Exercice 18.3 :

Exercice 18.4 :

Exercice 18.5 :

■ ???

Exercice 18.6 :

Exercice 18.7 :

■ ???

Exercice 18.8 :

Exercice 18.9 :

Exercice 18.10 :

Exercice 18.11 :

Exercice 18.12 :

■ ???

Exercice 18.13 :

Indications

____ Ex. 18.1 _____

____ Ex. 18.2 _____

____ Ex. 18.3 _____

____ Ex. 18.4 _____

____ Ex. 18.5 _____

____ Ex. 18.6 _____

- ____ **Ex. 18.7** _____
- ____ **Ex. 18.8** _____
- ____ **Ex. 18.9** _____
- ____ **Ex. 18.10** _____
- ____ **Ex. 18.11** _____
- ____ **Ex. 18.12** _____
- ____ **Ex. 18.13** _____

Corrigé des vrai/faux

1	2	3	4	5	6	7	8	9	10	11	12	13	14
F	F	F	F	F	V	F	V	F	F	F	V	F	F

1. Attention, une partie imaginaire est un réel!
3. Soit S la somme de deux complexes z_1 et z_2 et P son produit. Les deux complexes sont solutions de l'équation $x^2 - Sx + P = 0$, on peut le voir en développant $(x - z_1)(x - z_2)$. On a le résultat. Il suffit de considérer l'équation $x^2 + x + 1 = 0$ dont les racines sont j et j^2 . On a $S = -1$ et $P = 1$. Pourtant j et j^2 ne sont pas réels.
4. Prendre $a = 1$ et $b = i$. L'assertion est donc fausse. Pourtant si on impose à a et b d'être réels, elle serait vraie!
5. 0 n'a pas d'argument, devons-nous l'écrire? Ben oui, on l'a écrit!
6. Dès que la différence des arguments est un multiple de 2π , si leurs modules sont les mêmes, ils sont égaux. L'assertion est vraie.
9. Par exemple, $e^{i\pi^2} \in \mathbf{U}$ et il n'existe pas $(k, n) \in \mathbf{Z}^2$ tel que $e^{i2\pi/n} = e^{i\pi^2}$.
10. Ne jamais écrire \sqrt{i} . L'assertion n'a donc pas de sens. Elle est considérée comme fausse.
11. Attention, $j = e^{i2\pi/3}$.
13. C'est faux sauf si a, b et c sont réels.
14. Il y a une infinité de solutions! En effet, si on pose $z = x + iy$ avec $(x, y) \in \mathbf{R}^2$, on peut écrire $e^{x+iy} = e^x e^{iy} = -1 = e^{i\pi}$ et cela donne $x = 0$ et $y = \pi + 2k\pi$. Les solutions de l'équation sont tous les complexes de la forme $z = i(\pi + 2k\pi)$, $k \in \mathbf{Z}$.

■ Quelques pièges à éviter

❏ Erreurs classiques

- Oublier d'écrire (ou de vérifier) que z doit être non nul avant de considérer $\text{Arg}(z)$.
- Prendre la partie imaginaire d'un complexe en gardant i en facteur.
- Écrire $|z| = z\bar{z}$ à la place de $|z|^2 = z\bar{z}$.
- Oublier que $|z + z'| = |z| + |z'|$ est équivalent à l'existence de $\lambda > 0$ tel que $z' = \lambda z$.
- Donner l'impression de découvrir $j = e^{\frac{2i\pi}{3}}$ pour la première fois et ne pas avoir le réflexe d'écrire que $\bar{j} = j^2$ et $1 + j + j^2 = 0$.
- Écrire \sqrt{z} pour désigner la racine carrée de z , ce qui ne veut strictement rien dire car il y a en général deux racines carrées distinctes.
- Oublier, quand on calcule les racines $n^{\text{ièmes}}$ d'un complexe non nul, d'ajouter $2k\pi$ à l'argument de ce dernier avant de diviser par n .
- Oublier que si les trois quantités a, b, c sont réelles, l'équation $az^2 + bz + c = 0$ dans le cas où $\Delta < 0$, admet deux solutions complexes conjuguées.
- Dire par contre que l'équation $az^2 + bz + c = 0$ admet deux solutions complexes conjuguées dans le cas où l'une au moins des trois quantités a, b, c n'est pas réelle.

Corrigé des exercices

- ___ Exercice 18.1 _____
- ___ Exercice 18.2 _____
- ___ Exercice 18.3 _____
- ___ Exercice 18.4 _____
- ___ Exercice 18.5 _____
- ___ Exercice 18.6 _____
- ___ Exercice 18.7 _____
- ___ Exercice 18.8 _____
- ___ Exercice 18.9 _____
- ___ Exercice 18.10 _____
- ___ Exercice 18.11 _____
- ___ Exercice 18.12 _____
- ___ Exercice 18.13 _____