

Mathématiques pour le traitement du signal

C. Jauberthie - E. Montseny

v1 - 2015/2016

Table des matières

1	Séries	7
1.1	Rappels sur les suites numériques	7
1.2	Séries numériques	10
1.2.1	Définitions et premiers résultats	10
1.2.2	Séries à termes positifs	12
1.2.3	Séries à termes de signe quelconque	16
1.3	Séries entières	18
1.4	Séries de Fourier	23
1.4.1	Définitions	23
1.4.2	Propriétés et théorème important	29
2	Intégration	31
3	Transformation de Fourier et peigne de Dirac	33
4	Résolution d'équations différentielles linéaires	35
4.1	Introduction	35
4.2	Résolution des équations différentielles linéaires	37
4.2.1	Méthodologie générale	37
4.2.2	EDO linéaires d'ordre 1	37
4.2.3	EDO linéaires à coefficients constants d'ordre 2	41
5	Algèbre linéaire	47
5.1	Espaces Vectoriels	47
5.1.1	Définitions	47
5.1.2	Sous-espaces vectoriels	49
5.1.3	Familles libres et génératrices	49

5.1.4	Bases, dimension	51
5.2	Applications linéaires	52
5.2.1	Rappels sur les applications	52
5.2.2	Applications linéaires	53
5.3	Matrices	55
5.3.1	Des applications linéaires aux matrices	55
5.3.2	Définitions et notations	55
5.3.3	Structure de $\mathcal{M}_{m,n}(\mathbb{K})$	57
5.3.4	Propriétés et opérations sur les matrices	59
5.3.5	Matrices de changement de base	63
5.4	Réduction d'endomorphismes	64
5.4.1	Valeurs propres, vecteurs propres et sous-espaces propres	65
5.4.2	Diagonalisation de matrices	66
6	Probabilités et statistiques	69

Introduction et références

Le but de ce cours est d'introduire de manière la plus pragmatique possible, sans pour autant sacrifier la rigueur, quelques-unes des nombreuses notions d'analyse et d'algèbre indispensables pour le traitement du signal et l'ingénierie en général.

Les notions abordées sont vastes et nécessitent chacune un module à part entière pour être correctement approfondies ; ainsi, elles sont étudiées dans ce cours dans leur forme la plus simple afin de donner les bases nécessaires au lecteur désireux d'en savoir plus et ainsi lui permettre de se référer à des ouvrages spécialisés ; quelques ouvrages de référence, réputés pour leur qualité pédagogique et leur rigueur, sont donnés ci-dessous, à titre informatif.

Quelques ouvrages pour s'entraîner :

- *Mathématiques, Jean-Marie Monier, édition Dunod* (il en existe plein pour à peu près toutes les classes préparatoires !)
- *Les maths en tête, tome d'Analyse ou tome d'Algèbre, Xavier Gourdon, édition Ellipses* (une référence, mais un peu plus difficile d'accès).
- *Analyse numérique et équations différentielles, Jean-Pierre Demailly, EDP Sciences.*
- Tout autre ouvrage de mathématiques pour licence ou classes préparatoires.
- Tous les petits exemples et exercices éparpillés dans ce poly !

Chapitre 1

Séries

Pré-requis indispensables :

- Notions sur les suites numériques.
- Notions classiques sur les fonctions d'une variable réelle et calculs d'intégrales simples.
- Calculs de dérivées et primitives usuelles.
- Calculs de limites et règles de comparaison classiques (exp/ln/pol - fonctions trigo).
- Notions d'équivalence (\sim) de fonctions en un point.
- Développements limités usuels.
- Techniques basiques de décompositions en éléments simples.

1.1 Rappels sur les suites numériques

On se cantonne dans ce cours aux suites numériques réelles, mais les résultats s'étendent sans difficulté au cas des suites complexes.

Définition 1.1.1

Une **suite numérique** est une application à valeur réelle (ou complexe) dont l'argument est un nombre entier, soit $u : \mathbb{N} \mapsto \mathbb{R}$. Par commodité, on note $u = (u_n)_{n \in \mathbb{N}}$ ou $u = (u_n)_{n \geq 0}$ un tel objet.

Les suites peuvent être définies de manière explicite (expression de u_n en fonction de n), ou alors par récurrence (expression de u_{n+1} en fonction de u_n, u_{n-1} etc. et de conditions initiales). Dans ce cours, les suites servant essentiellement pour l'étude de séries, on manipulera principalement des suites explicitement définies, mais il est important de savoir manipuler les relations de récurrence !

Exemples : Voici quelques exemples de suites numériques classique.

- Suite arithmétique (définie par récurrence) : $v_{n+1} = v_n + a \forall n \geq 1, v(0) = 0$.
- Suite géométrique (définition explicite) : $u_n = a^n$ pour tout $n \in \mathbb{N}$, avec $a > 0$.
- $w_n = \frac{2^n}{n!} \forall n \geq 1$.

Naïvement, on peut noter qu'une suite est une fonction dont l'argument n est défini sur un ensemble dénombrable (\mathbb{N} ou un sous-ensemble de \mathbb{N} le plus souvent, ou encore \mathbb{Z}) au lieu d'être une variable x définie sur un continuum de valeurs (\mathbb{R}, \mathbb{C} etc.)

Ainsi, on peut additionner et multiplier des suites (lorsqu'elles ont le même ensemble de définition évidemment) comme on peut le faire pour les fonctions, et un grand nombre de notions élémentaires sur les fonctions se transpose sans difficultés aux suites ; en particulier, les notions de limite (à l'infini), de borne, de croissance, nous seront utiles.

Définition 1.1.2

Soit une suite $u = (u_n)_{n \in \mathbb{N}}$.

- u est dite **bornée** si $\exists M \in \mathbb{R}$ tel que $\forall n \in \mathbb{N}, |u_n| \leq M$.
- On dit que la suite u **converge vers** l si la limite de u à l'infini est égale à l , i.e. $\forall \varepsilon \geq 0$, il existe un indice $N \in \mathbb{N}$ à partir duquel on est assuré que :

$$|u_n - l| \leq \varepsilon, \quad \forall n \geq N. \quad (1.1)$$

Le nombre l est également appelé **limite de la suite** u ; elle est unique (très simple à démontrer, à faire pour s'amuser).

- u est dite **majorée** (resp. **minorée**) si $\exists M \in \mathbb{R}$ (resp. $m \in \mathbb{R}$) tel que $u_n \leq M$ (resp. $u_n \geq m$) pour tout n . On notera que u est majorée et minorée équivaut à u bornée.
- u est dite **croissante** si $u_{n+1} \geq u_n$ pour tout $n \in \mathbb{N}$. La définition d'une suite décroissante est alors évidente.

Important : La notion de "propriété valable à **partir d'un certain rang**", implicitement introduite dans la définition précédente, est essentielle dans tout ce cours sur les suites et séries numériques ; il convient de bien se familiariser avec cette notion qui est à la base de nombreux résultats et raisonnements.

Bien sûr, toutes les règles usuelles de calcul de limites dans le cadre des fonctions de la variable réelle (comparaison polynôme-exponentielle-logarithme, fractions rationnelles, etc.) restent valables dans le cadre des suites.

Exemples : Reprenons les suites de l'exemple précédent :

- Quelle que soit la valeur de a , on se doute bien qu'une telle suite va tendre vers $+\infty$. Pour le montrer, il faut exprimer v_n en fonction de n en exploitant la relation de récurrence : $v_n = v_{n-1} + a = v_{n-2} + 2a = \dots = an$.
- La convergence de u_n dépend de la valeur de a ; si $a < 1$, la suite converge vers 0, si $a > 1$ est elle évidemment divergente (sa limite est $+\infty$). Si $a = 1$, sa limite vaut 1.
- Pour w_n le résultat est moins immédiat, car on est a priori en présence d'une forme indéterminée. Soit $m > 2$ un nombre entier. A partir d'un certain rang $n > m$, on peut écrire :

$$\frac{2^n}{n!} = \frac{2^m}{m!} \frac{2^{n-m}}{n!/m!} = \frac{2^m}{m!} \frac{2^{n-m}}{n(n-1)\dots \underbrace{(m+1)}_{=n-(n-m+1)}} \leq \frac{2^m}{m!} \frac{2^{n-m}}{\underbrace{m^{n-m}}_{\text{Cte} \rightarrow 0 \text{ car } m > 2}}$$

Une autre démonstration plus élégante utilise la notion d'équivalence à l'infini (cf. après la définition 1.1.5) et la formule de Stirling.

Exercices : Calculer les limites des suites :

$$- a_n = \frac{3n^2+5n+2}{5n^3-5n^2} \rightarrow 0$$

$$- c_n = \frac{n}{\ln n} \rightarrow +\infty$$

$$- b_n = \frac{3n+2}{2n-3} \rightarrow \frac{3}{2}$$

$$- d_n = \frac{(-1)^n}{2^n} \rightarrow 0$$

Attention : une suite (de même qu'une fonction) peut ne pas avoir de limite !

Exemple : La suite définie par $u_n = n \cos\left(\frac{n\pi}{2}\right)$ n'a pas de limite (tracer ses valeurs dans le plan en s'aidant du cercle trigo).

Proposition 1.1.3 (*Stabilité*)

Si u et v sont deux suites convergentes vers a et b respectivement, alors $u + v$ et uv convergent vers $a + b$ et ab respectivement ; le résultat reste valable pour $\frac{u}{v}$ si $b \neq 0$.
Enfin, λu converge vers λa .

On donne ci-après un résultat de convergence utile pour l'étude des séries.

Proposition 1.1.4

Toute suite croissante majorée est convergente.

Démonstration : Évidente avec un raisonnement par l'absurde : une suite croissante non convergente a nécessairement pour limite $+\infty$, ce qui est contradictoire avec l'hypothèse de suite majorée. ■

Définition 1.1.5 (*Équivalence de deux suites*)

Deux suites numériques $u := (u_n)_{n \geq 0}$ et $v := (v_n)_{n \geq 0}$ sont dites **équivalentes** (lorsque $n \rightarrow +\infty$) si et seulement si $u - v$ négligeable devant v , i.e. s'il existe une suite $(\varepsilon_n)_{n \geq 0}$ avec $\lim_{n \rightarrow +\infty} \varepsilon_n = 0$ telle qu'à partir d'un certain rang $u_n = v_n + \varepsilon_n v_n$ ¹. On le note $u \underset{\infty}{\sim} v$.

Lorsque v_n ne s'annule pas à partir d'un certain rang, cette définition peut également s'écrire plus simplement :

$$u \underset{\infty}{\sim} v \Leftrightarrow \lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1. \tag{1.2}$$

Remarque importante : On notera que la relation \sim est conservée par le produit ($u_1 \sim u_2$ et $v_1 \sim v_2 \Rightarrow u_1 v_1 \sim u_2 v_2$) et par l'élévation à n'importe quelle puissance $\alpha \in \mathbb{R}$ ($u \sim v \Rightarrow u^\alpha \sim v^\alpha$, donc en particulier pour la notion d'inverse avec $\alpha = -1$) mais n'est pas compatible avec l'addition² et encore moins avec la composition d'équivalents !

1. En d'autres termes, u_n est égal à v_n plus éventuellement quelque chose de négligeable devant v_n , ce qui est plus faible que d'imposer $u - v \rightarrow 0$, qui est suffisant mais non nécessaire.

Il n'y a pas de règle générale pour la composition. Par exemple $x + 1 \underset{\infty}{\sim} x$, et pourtant $e^{x+1} = e^x e^1 \underset{\infty}{\sim} e^x$; en réalité, il faut s'assurer au préalable que la différence entre la fonction et son équivalent (i.e. ce que l'on décide de négliger) tende bien vers 0! Pour la composition par le log, c'est plus logique, il faut s'assurer que l'équivalent ne s'approche pas de la valeur 1 au voisinage du point.

La notion d'équivalence entre suites sera essentielle dans l'étude des séries numériques. Il est indispensable de savoir manipuler les calculs de limites et les règles de développements limités usuels (qui seront utilisés en $\frac{1}{n}$ qui est voisin de 0) afin d'établir simplement ces équivalences.

Exemples : — On peut reprendre l'exemple $w_n = \frac{2^n}{n!} \forall n \geq 1$ déjà traité, et utiliser la formule de Stirling : $n! \underset{\infty}{\sim} \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$. Ainsi, on a après calculs :

$$w_n \underset{\infty}{\sim} \frac{1}{\sqrt{2\pi}} e^{-n \ln n} \underset{n \rightarrow +\infty}{\rightarrow} 0.$$

— Soit la suite définie par $z_n = \frac{1}{\sqrt{n}} \ln\left(1 + \frac{1}{\sqrt{n}}\right) > 0$. Un DL de $\ln(1+x)$ avec $x = \frac{1}{\sqrt{n}}$ (proche de 0 lorsque $n \rightarrow +\infty$) permet d'écrire : $\ln\left(1 + \frac{1}{\sqrt{n}}\right) \underset{\infty}{\sim} \frac{1}{\sqrt{n}}$, et par suite $z_n \underset{\infty}{\sim} \frac{1}{n}$, donc la suite $(z_n)_{n \in \mathbb{N}}$ converge vers 0.

1.2 Séries numériques

1.2.1 Définitions et premiers résultats

Les séries numériques sont aux suites ce que les intégrales sont aux fonctions : une somme ; elles constituent un premier pas vers un objet plus général : les séries de fonctions. Les applications sont nombreuses, tant de manière directe (séries de Fourier, résolutions d'EDP, etc.) qu'indirectes (calculs d'intégrales complexes par exemple).

Définition 1.2.1 (*Série numérique*)

Soit $(u_k)_{k \in \mathbb{N}}$ une suite numérique. Le nombre

$$S_n = \sum_{k=0}^n u_k \tag{1.3}$$

est appelée **somme partielle d'indice n** (de la suite (u_k)). En considérant cette somme pour tout n , on obtient une suite (la suite des sommes partielles), notée $(S_n)_{n \in \mathbb{N}}$, que l'on appelle généralement la **série de terme général** u_k . Elle est fréquemment notée de manière synthétique $\sum_{k \in \mathbb{N}} u_k$ ou encore $\sum u_k$ lorsqu'aucune confusion n'est à craindre.

2. Considérer par exemple $-\frac{1}{n} + \frac{1}{n^3}$ et $\frac{1}{n} + \frac{1}{n^2}$ à l'infini : leur somme est égale à $\frac{1}{n^2} + \frac{1}{n^3} \underset{\infty}{\sim} \frac{1}{n^2}$, alors que la somme de leurs équivalents est égale à 0.

Remarque : Attention de ne pas confondre la **suite** $(u_n)_{n \in \mathbb{N}}$ et la **série** $\sum_{n \in \mathbb{N}} u_n$! Ceci est d'autant plus facile que l'on a vite tendance à mélanger les indices k et n (**muets !**) dans les expressions synthétiques des séries ; il faut apprendre à jongler avec les notations et ne pas s'attacher aux noms des indices utilisés.

Exemples : — Soit la suite $u_k = k$. Alors $S_n = 0 + 1 + 2 + 3 + \dots + n$; c'est une somme arithmétique qui vaut : $S_n = \frac{n(n+1)}{2}$ (résultat bien connu depuis le lycée).
 — Soit la suite $u_k = a^k$ avec $a \in \mathbb{R} \setminus \{0, 1\}$. Alors $S_n = 1 + a + a^2 + \dots + a^n$; la somme partielle s'exprime : $S_n = \frac{1-a^{n+1}}{1-a}$ (résultat bien connu également !).

Définition 1.2.2 (*Convergence d'une série numérique*)

La série $\sum_{k \in \mathbb{N}} u_k$ est dite **convergente** si et seulement si la suite des sommes partielles $(S_n)_{n \in \mathbb{N}}$ converge (en tant que suite donc) vers une limite finie S , appelée **somme de la série** ; elle peut se noter $S = \sum_{k=0}^{+\infty} u_k$ (à ne pas confondre avec $\sum_{k \in \mathbb{N}} u_k$ qui représente la série en tant qu'objet mathématique !).

Si au contraire $\lim_{n \rightarrow \infty} S_n = \pm\infty$, on dit que la série est **divergente**.

Exemples : Reprenons les exemples précédents.

- $u_k = k$; la série $\sum_{k \in \mathbb{N}} u_k$ est évidemment divergente (somme infinie de termes positifs de plus en plus grands...).
- $u_k = a^k$; d'après l'expression de S_n et l'étude de sa limite $\lim_{n \rightarrow \infty} S_n$, on peut dire que $\sum_{k \in \mathbb{N}} u_k$ est une série convergente si et seulement si $0 \leq |a| < 1$ et que sa somme vaut $\frac{1}{1-a}$. Si $|a| > 1$, la série est divergente.
 N.B. Si $a = 1$, alors $S_n = n + 1 \xrightarrow{n \rightarrow +\infty} +\infty$ donc la série $\sum u_k$ diverge également.

Remarque importante : Pour des raisons évidentes, **la nature d'une série est inchangée lorsque l'on ajoute ou retranche un nombre fini de termes** (un nombre fini de termes ajoutés ou retranchés ne saurait porter une somme à une valeur infinie...) ; couplée à la notion de propriété "à partir d'un certain rang" évoquée pour les suites, elle est très utile dans le cadre de l'étude de la nature d'une série !

Proposition 1.2.3 (*Stabilité*)

Soient $\sum_{n \in \mathbb{N}} u_n$ et $\sum_{n \in \mathbb{N}} v_n$ convergentes respectivement vers a et b . Alors $\sum_{n \in \mathbb{N}} u_n + \sum_{n \in \mathbb{N}} v_n$ et $\sum_{n \in \mathbb{N}} \lambda u_n$ ($\lambda \in \mathbb{R}$) convergent respectivement vers $a + b$ et λa (l'ensemble des séries numériques convergentes est un espace vectoriel).

La proposition qui suit est essentielle pour l'étude de la nature des séries numériques. Elle doit constituer le premier réflexe de tout étudiant avant de se lancer dans les calculs.

Proposition 1.2.4 (*Condition nécessaire de convergence*)

Si la série $\sum_{n \in \mathbb{N}} u_n$ converge, alors nécessairement $u_n \xrightarrow[n \rightarrow +\infty]{} 0$. Dans le cas contraire, la série est dite **grossièrement divergente**.

Attention : la réciproque de ce théorème est évidemment fautive (il suffit de considérer $\sum_{n \in \mathbb{N}} \frac{1}{n}$ qui diverge et pourtant $u_n \xrightarrow[n \rightarrow +\infty]{} 0$).

Exercice : Soit la série de terme général $u_n = \frac{1}{n(n+1)}$.

- La série est-elle grossièrement divergente ?
- Montrer que la série $\sum_{n \geq 0} u_n$ converge en calculant sa somme (une décomposition en éléments simple pourra aider...).

Sol : on a $u_n = \frac{1}{n} - \frac{1}{n+1}$, donc les termes s'annulent deux à deux dans la somme partielle et on a $S_n = 1 + \frac{1}{n+1} \xrightarrow[n \rightarrow +\infty]{} 1$; la somme de la série est donc égale à 1, et elle est évidemment, de fait, convergente.

A quelques exceptions académiques près, calculer la somme d'une série de manière analytique relève du miracle (de même que le calcul d'intégrales et de primitives générales est impossible). Cependant, il n'est pas toujours nécessaire de calculer leur somme : connaître leur nature nous suffit. De nombreux résultats existent pour nous aider dans cette tâche ; c'est l'objet des sections suivantes.

Quoi qu'il en soit, l'étude de la nature d'une série $\sum u_n$ se fera selon des étapes séquentielles bien déterminées :

- Si $u_n \not\rightarrow 0$ alors la série diverge grossière d'après la prop. 1.2.4.
- Sinon, si u_n est de signe constant (à partir d'un certain rang), on cherche à appliquer les règles de la section 1.2.2.
- Sinon, on cherche à établir la **convergence absolue**, cf. la prop-def 1.2.13.
- Sinon, si u_n est de signe alterné alors on peut essayer le thm 1.2.15.
- Sinon, utiliser les résultats sur les séries de signe quelconque (cf. thm 1.2.16).

1.2.2 Séries à termes positifs

Avant toute chose, il convient de fixer la terminologie.

Définition 1.2.5 (*Séries à termes positifs*)

Une série $\sum_{n \in \mathbb{N}} u_n$ est dite **à termes positifs** si $u_n \geq 0$ à partir d'un certain rang.

Nous avons déjà évoqué le fait que la nature d'une série ne saurait être remise en cause par l'ajout ou le retrait d'un nombre fini de termes ; c'est ce qui motive la définition précédente.

De plus, il est évident que la notion de série à termes positifs englobe finalement les séries à termes de signe constant (à partir d'un certain rang toujours) dans la mesure où le

cas des séries à termes négatifs se traite en considérant $v_n = -u_n \geq 0$ pour l'étudier. Quoi qu'il en soit, il conviendra de **toujours vérifier au préalable que le terme général u_n entre bien dans ce cadre avant d'appliquer une des méthodes données dans cette section !**

Le premier résultat ci-dessous est directement déduit du théorème 1.1.4 sur les suites.

Théorème 1.2.6

Soit $\sum_{n \in \mathbb{N}} u_n$ une série à termes positifs. Alors la suite des sommes partielles $(S_n)_{n \in \mathbb{N}}$ est croissante, et on a $\sum_{n \in \mathbb{N}} u_n$ convergente si et seulement si (S_n) est majorée.

Démonstration : On a $S_{n+1} = S_n + u_{n+1} \geq S_n$, donc (S_n) est bien une suite croissante ; ainsi, d'après la proposition 1.1.4, si elle est majorée alors elle est convergente, i.e. $\sum_{n \in \mathbb{N}} u_n$ convergente. ■

Exemple : La série $\sum_{n \in \mathbb{N}^*} \frac{1}{n^2}$ est convergente car $\frac{1}{k^2} \leq \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$, d'où S_n majorée par

$$S_n = \sum_{k=1}^n \frac{1}{k^2} \leq 1 + \sum_{k=2}^n \left(\frac{1}{k-1} - \frac{1}{k} \right) = 2 - \frac{1}{n} \leq 2.$$

Attention, cela ne signifie en aucun cas que la somme de la série vaut 2 !

On donne ci-après quelques techniques classiques pour la détermination de la nature (i.e. dire si elle est convergente ou divergente sans nécessairement avoir à calculer sa somme) d'une série numérique à termes positifs. Il est indispensable de tous les connaître dans la mesure où chacun d'entre eux peut servir ou se révéler peu utile selon l'expression de u_n .

Un certain nombre des théorèmes qui seront donnés par la suite nécessite de disposer de séries "type" dont on connaît la nature, dans la mesure où ils sont basés sur des comparaisons, équivalences, etc. On donne donc **ci-après un résultat qu'il est essentiel de connaître** pour l'étude des séries numériques.

Proposition 1.2.7 (Séries de Riemann et de Bertrand)

La série (dite de Riemann) de la forme :

$$\sum_{n \in \mathbb{N}^*} \frac{1}{n^\alpha} \tag{1.4}$$

converge **si et seulement si $\alpha > 1$** .

Un résultat plus général est donné par la série de Bertrand de la forme :

$$\sum_{n \in \mathbb{N}^*} \frac{1}{n^\alpha (\ln n)^\beta} \tag{1.5}$$

qui converge si et seulement si $\alpha > 1$ (quelque soit β) ou si $\alpha = 1$ et $\beta > 1$.

Démonstration : Plusieurs preuves possibles ; notamment une assez simple, basée sur le théorème de comparaison avec une intégrale). ■

On dispose désormais d'une large classe de séries sur laquelle il sera possible de s'appuyer pour utiliser les théorèmes qui suivent.

On introduit ci-après deux règles simples permettant de déterminer la nature d'une série numérique en étudiant son terme général. L'une (d'Alembert) est particulièrement utile lorsque l'on est en présence de fonctions puissance ou de factorielles dans le terme général u_n , l'autre (Cauchy) lorsque l'on est en présence de puissances nièmes dans le terme général u_n .

Théorème 1.2.8 (Règle de d'Alembert)

Soit $\sum_{n \in \mathbb{N}} u_n$ une série à termes positifs telle que $\lim_{n \rightarrow +\infty} \frac{|u_{n+1}|}{|u_n|} = l$. Alors :

- Si $l < 1$, la série $\sum_{n \in \mathbb{N}} u_n$ converge.
- Si $l > 1$, la série $\sum_{n \in \mathbb{N}} u_n$ diverge.
- Si $l = 1$, on ne peut conclure par cette méthode.

Exemple : Soit la série $\sum_{n \in \mathbb{N}^*} \frac{a^n}{n!}$ avec $a > 0$. Alors cette série converge d'après le thm. 1.2.8 car $\frac{u_{n+1}}{u_n} = \frac{a}{n+1} \xrightarrow{n \rightarrow +\infty} 0 < 1$.

Théorème 1.2.9 (Règle de Cauchy)

Soit $\sum_{n \in \mathbb{N}} u_n$ une série à termes positifs telle que $\lim_{n \rightarrow +\infty} \sqrt[n]{|u_n|} = l$. Alors :

- Si $l < 1$, la série $\sum_{n \in \mathbb{N}} u_n$ converge.
- Si $l > 1$, la série $\sum_{n \in \mathbb{N}} u_n$ diverge.
- Si $l = 1$, on ne peut conclure par cette méthode.

Exemple : Soit la série de terme général $u_n = \left(\frac{n}{1+2n}\right)^n$. On a $\sqrt[n]{u_n} = \frac{n}{1+2n} \xrightarrow{n \rightarrow +\infty} \frac{1}{2} < 1$ donc la série $\sum_{n \in \mathbb{N}} u_n$ converge d'après le thm. 1.2.9.

Dans les lignes qui suivent, on introduit toute une série de théorèmes très utiles, permettant de déterminer la nature d'une série en la comparant à une autre série dont on connaîtrait la nature (à ce titre, le théorème 1.2.7 est essentiel).

Théorème 1.2.10 (Comparaison)

Soient $\sum_{n \in \mathbb{N}} u_n$ et $\sum_{n \in \mathbb{N}} v_n$ deux séries à termes positifs telles que $u_n \leq v_n$ à partir d'un certain rang. Alors :

$$\left[\begin{array}{l} - \sum_{n \in \mathbb{N}} v_n \text{ convergente} \Rightarrow \sum_{n \in \mathbb{N}} u_n \text{ convergente.} \\ - \sum_{n \in \mathbb{N}} u_n \text{ divergente} \Rightarrow \sum_{n \in \mathbb{N}} v_n \text{ divergente.} \end{array} \right.$$

Démonstration : Ce résultat est particulièrement intuitif (d'accord ?). Pour le démontrer proprement, il suffit de se référer au théorème 1.2.6. Si v_n est convergente, alors la suite des sommes partielles $(Sv_n)_{n \in \mathbb{N}}$ est une suite majorée ; notons M un majorant. Comme $u_n \leq v_n$ pour tout n ³, il est clair que $Su_n \leq Sv_n \leq M$, donc $(Su_n)_{n \in \mathbb{N}}$ est une suite majorée, et par suite $\sum_{n \in \mathbb{N}} u_n$ est une série convergente toujours d'après le théorème 1.2.6.

On démontre le second point de manière similaire ($(Sv_n)_{n \in \mathbb{N}}$ n'est pas majoré). ■

Il est important de noter que ce théorème peut s'utiliser dans les deux sens, tant pour montrer qu'une série converge que pour montrer qu'elle diverge.

Exemples : — La série $\sum_{n \geq 1} \frac{1}{n^2 + \sqrt{n}}$ converge car $\frac{1}{n^2 + \sqrt{n}} \leq \frac{1}{n^2}$ avec $\sum_{n \geq 1} \frac{1}{n^2}$ converge (Riemann).
 — La série $\sum_{n \geq 1} \frac{\ln n}{n}$ diverge car $\frac{\ln n}{n} \geq \frac{1}{n}$ pour $n \geq 1$ et $\sum_{n \geq 1} \frac{1}{n}$ diverge d'après le thm 1.2.10.

Théorème 1.2.11 (Équivalence)

Soient $\sum_{n \in \mathbb{N}} u_n$ et $\sum_{n \in \mathbb{N}} v_n$ deux séries à termes positifs telles que $u_n \sim v_n$ lorsque $n \rightarrow +\infty$. Alors $\sum_{n \in \mathbb{N}} u_n$ et $\sum_{n \in \mathbb{N}} v_n$ sont **de même nature** (i.e. convergent ou divergent toutes deux).

Exemple : — Soit série déjà traitée plus haut : $\sum_{n \geq 1} u_n$ avec $u_n = \frac{1}{n^2 + \sqrt{n}}$.

On sait que $n^2 + \sqrt{n} \underset{\infty}{\sim} n^2$ (car \sqrt{n} évidemment négligeable devant n^2 à l'infini), donc $u_n \underset{\infty}{\sim} \frac{1}{n^2}$; comme $\sum_{n \geq 1} \frac{1}{n^2}$ converge (Riemann), on en déduit que $\sum_{n \geq 1} u_n$ converge d'après le thm 1.2.11.

— Considérons maintenant la série de terme général $u_n = \frac{1}{\sqrt{n}} \ln \left(1 + \frac{1}{\sqrt{n}} \right) > 0$. On a vu dans la section 1.1 que $u_n \underset{\infty}{\sim} \frac{1}{n}$. Or $\sum_{n \geq 1} \frac{1}{n}$ diverge (Riemann toujours !), donc il en va de même de $\sum_{n \geq 1} u_n$ d'après le thm 1.2.11.

Enfin, cette dernière règle peut s'avérer très utile pour établir la nature d'une série en passant par un calcul d'intégrale (ou du moins que l'on peut se prononcer sur sa nature).

Théorème 1.2.12 (Comparaison avec une intégrale)

Soit une fonction $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ positive et décroissante. Alors, l'intégrale (impropre) $\int_1^{+\infty} f(x) dx$ et la série $\sum_{n \geq 1} f(n)$ sont de même nature.

3. après avoir au besoin "placé" l'indice 0 au rang n_0 à partir duquel c'est vrai pour tout n ; on rappelle que l'ajout/retrait d'un nombre fini de termes ne change pas la nature d'une série

Exemple : On va utiliser ce théorème pour établir les conditions de convergence des séries de Riemann données en début de cette section. Soit la fonction $f : x \mapsto \frac{1}{x^\alpha}$ avec $\alpha > 1$ définie sur $[1, \infty[$; elle est positive et décroissante, donc en vertu le théorème 1.2.12, la nature de l'intégrale $\int_1^{+\infty} f(x)dx$ déterminera la nature de la série de Riemann associée $\sum_{n \geq 1} \frac{1}{n^\alpha}$. Or $\forall X \geq 1$:

$$\int_1^X \frac{1}{n^\alpha} dx = \frac{1}{-\alpha + 1} [x^{-\alpha+1}]_1^X = \frac{1}{-\alpha + 1} (X^{-\alpha+1} - 1). \quad (1.6)$$

La convergence de cette intégrale lorsque $X \rightarrow +\infty$ est évidemment fonction de la valeur de α : elle converge (vers $\frac{1}{\alpha-1}$) si $\alpha > 1$ et diverge sinon.

Remarque importante : Il est **essentiel** de noter que **tous** les résultats donnés dans cette section peuvent s'utiliser en cascade ! En effet, on peut par exemple déterminer dans un premier temps un équivalent v_n à notre terme général u_n , puis, afin de déterminer la nature de $\sum v_n$, utiliser un autre résultat tel qu'une comparaison ou encore Cauchy/d'Alembert ; une fois la nature de $\sum v_n$ établie, on peut alors conclure quant à la nature de $\sum u_n$.

1.2.3 Séries à termes de signe quelconque

Lorsque le terme général u_n n'est pas de signe constant, l'étude de sa nature est moins simple, mais on dispose malgré tout de nombreux résultats. On ne présentera que quelques résultats dans ce cours, le lecteur désireux d'en apprendre plus trouvera son bonheur dans tout livre traitant des séries numériques.

La première chose à étudier lorsque l'on est confronté à une série à termes de signe quelconque est d'étudier sa convergence absolue.

Proposition-Définition 1.2.13 (*Convergence absolue*)

Une série $\sum_{n \in \mathbb{N}} u_n$ est dite **absolument convergente** lorsque la série $\sum_{n \in \mathbb{N}} |u_n|$ est convergente. On peut alors montrer que **toute série absolument convergente est convergente**.

Si une série est convergente mais n'est pas absolument convergente⁴, elle est dite **semi-convergente**.

Exemple : $\sum_{n \in \mathbb{N}} \frac{\cos(n)}{n^2+1}$ n'est pas une série à terme général de signe constant. Cependant, on peut montrer qu'elle est absolument convergente (et donc convergente) par comparaison avec une série de Riemann : $|\frac{\cos(n)}{n^2+1}| \leq \frac{1}{n^2}$, avec $\sum_{n \in \mathbb{N}} \frac{1}{n^2}$ converge.

4. Par exemple, la série $\sum_{n \in \mathbb{N}^*} \frac{(-1)^n}{n}$.

Si la convergence absolue ne donne rien, d'autres résultats existent, notamment pour les séries dites alternées (i.e. dont le terme général u_n change de signe à chaque coup.

Définition 1.2.14 (*Série alternée*)

On appelle **série alternée** toute série de la forme $\sum_{n \in \mathbb{N}} (-1)^n a_n$ avec a_n de signe constant pour tout n .

Théorème 1.2.15

Soit $\sum_{n \in \mathbb{N}} (-1)^n a_n$ une série alternée. Si la suite $(a_n)_{n \in \mathbb{N}}$ est décroissante et tend vers 0, alors la série est convergente, et on a les encadrements et majorations suivants :

$$\dots \leq S_{2n-1} \leq S_{2n+1} \leq \dots \leq S \leq \dots \leq S_{2n+2} \leq S_{2n} \leq \dots \quad (1.7)$$

$$R_n \leq a_{n+1}. \quad (1.8)$$

On comprend aisément que la convergence de ce genre de série soit notablement moins contraignante que la convergence absolue dans la mesure où les termes successifs ont la possibilité de se "compenser" (puisque'ils ont un signe opposé) et amener la série à converger alors qu'elle divergerait en valeur absolue.

Exemple : En utilisant ce théorème, on montre sans difficulté que la série $\sum_{n \in \mathbb{N}^*} \frac{(-1)^n}{n}$ est convergente

($a_n = \frac{1}{n}$) alors que l'on sait que la série $\sum_{n \in \mathbb{N}^*} \frac{1}{n}$ est divergente.

Pour les séries de signe quelconque qui n'entrerait pas dans ce cadre, les résultats sont moins nombreux. On en citera un, le plus connu, basé sur la transformation d'Abel, qui est aux séries ce que l'intégration par partie est aux intégrales. Le principal résultat de cette transformation est le suivant.

Théorème 1.2.16 (*Règle d'Abel*)

Soit une série de la forme $\sum_{n \in \mathbb{N}} u_n v_n$ avec :

(i) $(u_n)_{n \in \mathbb{N}}$ une suite décroissante vers 0,

(ii) $\exists M > 0$ tel que $|\sum_{k=0}^n v_n| \leq M$ pour tout n .

Alors la série $\sum_{n \in \mathbb{N}} u_n v_n$ est convergente.

Remarque : On notera qu'en prenant $v_n = (-1)^n$, on retrouve le théorème des séries alternées, la condition (ii) étant trivialement vérifiée ($M = 1$).

1.3 Séries entières

Les séries entières sont un cas particuliers simple de séries de fonctions où les fonctions f_n sont des monômes de la forme $a_n x^n$. Leur structure particulière fait que leur étude est, *in fine*, assez similaires à celle des séries numériques.

On les retrouve dans de nombreux domaines d'application : développements en séries entières, séries de Taylor, résolutions d'équations différentielles, analyse complexe, etc.

Définition 1.3.1 (*Série entière*)

On appelle **série entière** une série (de fonctions) de la forme $\sum_{n \in \mathbb{N}} a_n x^n$ avec x une variable réelle et $(a_n)_{n \in \mathbb{N}}$ une suite numérique.

Remarque : Même si on se bornera dans ce cours à l'étude des séries entières réelles, le cadre s'étend sans difficulté aux séries entières complexes $\sum a_n z^n$ où $z \in \mathbb{C}$. Bien que non indispensable pour nous, c'est un cadre important puisque sous-jacent aux séries de Fourier (cf. section 1.4)

Exemples : — Un exemple bien connu : $\sum_{n \in \mathbb{N}} \frac{x^n}{n!}$, c'est-à-dire $a_n = \frac{1}{n!} \forall n$. On sait depuis nos petites classes que la valeur de cette somme n'est autre que e^x , pour tout $x \in \mathbb{R}$.

— $\sum_{n \in \mathbb{N}} \sin(n) x^n$ est un autre exemple de série entière.

De même que pour les séries numériques et les séries de fonctions, la question de la **convergence** des séries entières se pose ; la notion centrale pour les séries est celle de **rayon de convergence**, et elle est assez simple à appréhender car assez intuitive (contrairement aux convergences propres aux séries de fonctions telles que la convergence uniforme ou normale).

En effet, $\sum x^n$ est une série géométrique dont le résultat de convergence est bien connu (cf. section 1.2.1) : il faut et suffit que $|x| < 1$. Une série entière n'est pas exactement une série géométrique (du fait de la présence de a_n), mais on se doute bien que comme a_n est multiplié par x^n dont la croissance est géométrique, il va falloir qu'asymptotiquement les termes a_n compensent l'éventuelle croissance géométrique des x^n pour que la série puisse converger.

Vu autrement - car la variable est x (les a_n eux sont fixés!) - une telle série ne pourra converger que d'une part si les a_n ont une croissance asymptotiquement au plus géométrique (sinon le terme x^n , géométrique lui, sera dominé par leur croissance), et d'autre part si $|x|$ est choisi sur une certaine plage de valeurs (éventuellement infinie) qui sera telle que le produit $a_n x^n$ soit sommable, i.e. que la série $\sum a_n x^n$ converge. C'est sur cela que s'appuie la notion de **rayon de convergence** d'une série entière.

Proposition-Définition 1.3.2 (*Rayon de convergence*)

Soit une série entière $\sum_{n \in \mathbb{N}} a_n x^n$. On appelle **rayon de convergence** le nombre (fini ou

infini) :

$$R = \sup \{r \geq 0 \text{ tel que la suite } (a_n r^n)_{n \in \mathbb{N}} \text{ est bornée}\}^5. \quad (1.9)$$

On a alors les résultats suivants :

- (i) Pour tout $|x| < R$, la série (numérique) $\sum a_n x^n$ converge.
- (ii) Pour tout $|x| > R$, la série diverge.
Pour $|x| = R$, elle peut converger ou diverger selon les cas.
- (iii) Enfin, la série (de fonctions) $\sum a_n x^n$ converge normalement (et donc uniformément) sur tout intervalle de la forme $[0, r]$ avec $r < R$ (ce qui implique évidemment (i)).

Exemple : Il est bien connu que la série géométrique $\sum x^n$ converge si et seulement si $x < 1$ (et sa somme est alors égale à $\frac{1}{1-x}$). Dans le cas contraire, elle diverge. Son rayon de convergence est donc $R = 1$, et on notera que sur son rayon de convergence (i.e. pour $x = 1$, la série diverge).

Ces résultats de convergence, en particulier (iii), sont essentiels, car ils garantissent un certain nombre de résultats bien pratiques, qui sont la conséquence de résultats propres aux séries de fonctions uniformément convergentes : continuité de la somme, dérivation de la somme, permutation $\int - \sum$ etc. que l'on verra dans le théorème 1.3.4. Dans un premier temps, on propose ci-dessous deux règles simples, directement issues des règles de d'Alembert et de Cauchy pour les séries numériques, permettant le calcul pratique du rayon de convergence d'une série entière.

Proposition 1.3.3 (*Calcul pratique du rayon de convergence*)

Soit une série entière $\sum a_n x^n$. Alors, le rayon de convergence R de cette série entière est égal à $R = \frac{1}{\lambda}$ avec λ donné par l'une des deux méthodes suivantes (lorsque ces limites existent !) :

- **Règle de d'Alembert** : $\lambda = \lim_{n \rightarrow +\infty} \left| \frac{a_{n+1}}{a_n} \right|$,
- **Règle de Cauchy** : $\lambda = \lim_{n \rightarrow +\infty} \sqrt[n]{|a_n|}$.

Démonstration : Triviale en utilisant ces critères sur la série numérique de terme général $a_n x^n$. ■

Exemples : Avec ces résultats, on montre facilement que :

- La série $\sum \frac{x^n}{n!}$ a un RC égal à $+\infty$.
- La série $\sum n^\alpha x^n$ a un RC égal à 1. (avec $\alpha \in \mathbb{R}$).
- La série $\sum n! x^n$ a un RC égal à 0. (On pouvait s'en douter ; d'accord ?)

Lorsque le rayon de convergence est déterminé, on sait qu'une série entière y converge (en tant que série de fonction) normalement, et donc uniformément. Ce résultat garantit un ensemble de résultats très pratiques dont on en donne deux ci-dessous.

5. i.e. tel qu'il existe $M \geq 0$ tel que pour tout n on ait $|a_n r^n| \leq M$.

Théorème 1.3.4 (Dérivation et primitive d'une série entière)

Soit une série entière $\sum a_n x^n$ de rayon de convergence R . Alors :

- (i) La fonction $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ (fonction somme de la série entière) est une fonction continue et même dérivable sur $] -R, R[$, et sa dérivée est une série entière de même rayon de convergence R et donnée, pour tout $x \in] -R, R[$, par :

$$f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} \text{ (ce qui en prenant } x = 0 \text{ implique : } a_1 = f'(0)\text{)}. \quad (1.10)$$

et son rayon de convergence est R .

- (ii) Ce résultat s'étend par récurrence à toutes les dérivées de f , qui est de classe \mathcal{C}^∞ , et on a pour tout $x \in] -R, R[$, par :

$$f^{(m)}(x) = \sum_{n=m}^{+\infty} n(n-1) \dots (n-m+1) a_n x^{n-m} \quad (1.11)$$

ce qui implique, en prenant $x = 0$:

$$\forall n, a_n = \frac{f^{(n)}(0)}{n!} \text{ et donc pour tout } x \in] -R, R[, f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n \quad (1.12)$$

- (iii) Une primitive de la série entière f est une série entière de même rayon de convergence R et donnée, pour tout $x \in] -R, R[$, par :

$$\sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}. \quad (1.13)$$

Remarque : Il est inutile de retenir par cœur les formules du théorème précédent : il suffit pour les retrouver de savoir intégrer et dériver les monômes x^n !

Exemple : D'après l'exemple précédent, on sait que $\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$, $\forall x \in] -1, 1[$. En utilisant le résultat (1.13) du dernier théorème, on peut affirmer que $\ln(1+x)$ admet le DSE suivant, de même rayon de convergence :

$$\ln(1-x) = \sum_{n=0}^{+\infty} \frac{x^n}{n}.$$

Exercice important : Dédurre de l'exemple précédent le DSE de $\frac{1}{1+x}$ et de $\ln(1+x)$.

La relation (1.12) du Théorème 1.3.4 donne à s'interroger sur les conditions sous lesquelles

une fonction f donnée serait **développable en série entière**, c'est-à-dire serait égale, sur un certain domaine de convergence, à une série entière, qui prendrait à la forme (1.12).

Sans entrer dans les détails, il existe certaines conditions pour lesquelles on est assuré de l'existence d'un tel développement (une condition nécessaire est évidemment que f soit de classe \mathcal{C}^∞ ; d'accord ?). Son calcul se ferait alors en utilisant les dérivées successives de f , en vertu de l'expression (1.12).

Il est **indispensable** de connaître l'expression du développement en série entière (DSE) usuels donnés ci-après et de savoir les manipuler en utilisant les résultats de dérivation et intégration du Théorème 1.3.4, permettant ainsi d'avoir accès à une large classe de DSE classiques.

Développements en série entières de fonctions usuelles

$$\forall x \in \mathbb{R}, \quad e^x = 1 + x + \frac{x^2}{2} + \dots = \sum_{n=0}^{+\infty} \frac{x^n}{n!},$$

$$\forall x \in \mathbb{R}, \quad \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots = \sum_{p=0}^{+\infty} (-1)^p \frac{x^{2p+1}}{(2p+1)!},$$

$$\forall x \in \mathbb{R}, \quad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots = \sum_{p=0}^{+\infty} (-1)^p \frac{x^{2p}}{(2p)!},$$

$$\forall x \in \mathbb{R}, \quad \operatorname{sh} x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots = \sum_{p=0}^{+\infty} \frac{x^{2p+1}}{(2p+1)!},$$

$$\forall x \in \mathbb{R}, \quad \operatorname{ch} x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots = \sum_{p=0}^{+\infty} \frac{x^{2p}}{(2p)!},$$

$$\forall x \in]-1, 1[, \forall \alpha \in \mathbb{R}, \quad (1+x)^\alpha = 1 + \frac{\alpha x}{1!} + \frac{\alpha(\alpha-1)x^2}{2!} + \dots = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)x^n}{n!}$$

De cette dernière ($\alpha = -1$) on peut tirer les classiques classique DSE :

$$\forall x \in]-1, 1[, \quad \frac{1}{1-x} = 1 + x + x^2 + \dots = \sum_{n=0}^{+\infty} x^n$$

$$\forall x \in]-1, 1[, \quad \frac{1}{1+x} = 1 - x + x^2 - \dots = \sum_{n=0}^{+\infty} (-1)^n x^n,$$

et par intégration :

$$\forall x \in]-1, 1[, \quad \ln(1-x) = -x - \frac{x^2}{2} + \frac{x^3}{3} - \dots = -\sum_{n=1}^{+\infty} \frac{x^n}{n}$$

$$\forall x \in]-1, 1[, \quad \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}.$$

Le lecteur désireux de s'entraîner pourra utiliser les résultats précédents pour déterminer les DSE (classiques) de $\sqrt{1+x}$, $\frac{1}{\sqrt{1+x}}$, $\frac{1}{1+x^2}$ et $\arctan x$.

Exercice : Donner le développement en série entières de la fonction $\frac{1}{(x+1)^2}$.

1.4 Séries de Fourier

1.4.1 Définitions

Les séries de Fourier sont un outil de choix pour l'étude des signaux **périodiques**. Le résultat essentiel de cette section est que tout signal périodique (assez régulier) peut être écrit comme étant la somme de signaux élémentaires sinusoïdaux.

Définition 1.4.1 (*Coefficients de Fourier*)

Soit f une fonction T -périodique, à valeur réelle ou complexe. On appelle **coefficients de Fourier exponentiels de f** la suite de nombres complexes définie par :

$$c_n = \frac{1}{T} \int_a^{a+T} f(t) e^{-\frac{2in\pi t}{T}} dt, \forall n \in \mathbb{Z} \quad (1.14)$$

Si la fonction f est à valeurs **réelles**, on utilise souvent les **coefficients de Fourier trigonométriques de f** :

$$a_0 = \frac{1}{T} \int_a^{a+T} f(t) dt, \quad (1.15)$$

$$a_n = \frac{2}{T} \int_a^{a+T} f(t) \cos\left(\frac{2\pi nt}{T}\right) dt, \forall n \in \mathbb{N}^*, \quad (1.16)$$

$$b_n = \frac{2}{T} \int_a^{a+T} f(t) \sin\left(\frac{2\pi nt}{T}\right) dt, \forall n \in \mathbb{N}^*.$$

Remarque Importante : Dans certains ouvrages, le coefficient a_0 est défini selon la même formule que les coefficients a_n (avec un coefficient $\frac{2}{T}$ donc). Il s'agit purement d'une convention n'ayant aucune conséquence autre que celle de se souvenir de remplacer a_0 par $2a_0$ dans toutes les formules l'impliquant (Dirichlet, Parseval etc.). L'utilisation de l'une ou l'autre des définitions dépend surtout des us et coutumes des différentes branches de la physique et de l'ingénierie.

De même, il est fréquent de trouver ces formules exprimées en fonction de la pulsation fondamentale $\omega_0 := \frac{2\pi}{T}$; là encore, il suffit de remplacer.

Les coefficients c_n sont souvent appelés **spectre de f** . Ces coefficients étant complexes, on les représente habituellement sur deux graphiques, l'un dédié au module, et l'autre à la phase :

On reparlera plus en détail de cette notion de spectre fréquentiel d'un signal avec la transformation de Fourier.

Remarque : Si les coefficients a_n et b_n sont souvent utilisés lorsqu'on manipule des fonctions réelles, les coefficients c_n présentent l'intérêt, en plus d'être plus synthétiques, de faire l'analogie avec la transformée de Fourier (continue) \hat{f} d'une fonction non périodique f (cf. chapitre dédié à la transformation de Fourier).

Exemple : — Soit la fonction 2π -périodique h définie par (cf. figure 1.1) :

$$h(t) = \begin{cases} 1 & \text{si } t \in [0, \pi] \\ 0 & \text{si } t \in [\pi, 2\pi] \end{cases}$$

Ses coefficients de Fourier sont donnés par :

$$\begin{aligned} a_0 &= 1/2, \quad a_n = 0 \quad \forall n \in \mathbb{N}^* \\ b_n &= 0 \text{ si } n \text{ pair et } \frac{2}{n\pi} \text{ si } n \text{ est impair.} \end{aligned}$$

— Soit la fonction L -périodique g définie par (cf. figure 1.2) :

$$g(t) = \begin{cases} \frac{2}{L}t & \text{si } t \in [0, \frac{L}{2}] \\ 2 - \frac{2}{L}t & \text{si } t \in [\frac{L}{2}, L] \end{cases} \quad (1.17)$$

Ses coefficients de Fourier sont donnés par :

$$\begin{aligned} a_0 &= 1/2, \quad a_n = 0 \text{ si } n \text{ pair et } \frac{-4}{n^2\pi^2} \text{ si } n \text{ est impair.} \\ b_n &= 0 \quad \forall n \in \mathbb{N} \text{ (} f \text{ est paire).} \end{aligned} \quad (1.18)$$

Définition 1.4.2 (*Synthèse*)

On appelle **développement en série de Fourier (DSF) de f** (on parle aussi de **synthèse de Fourier**) la série :

$$\sum_{n \in \mathbb{Z}} c_n e^{\frac{2i\pi nt}{T}}, \quad (1.19)$$

ou, dans le cas des coefficients trigonométriques :

$$a_0 + \sum_{n \geq 1} a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right). \quad (1.20)$$

On a alors le résultat fondamental suivant.

Théorème 1.4.3 (*Théorème de Dirichlet*)

Soit f une fonction T -périodique, de classe \mathcal{C}^1 par morceaux. Alors, le DSF de f converge simplement vers $\frac{f(t^-) + f(t^+)}{2}$ (demi-somme entre les limites à gauche et à droite), soit :

$$\forall t, \frac{f(t^-) + f(t^+)}{2} = \sum_{n \in \mathbb{Z}} c_n e^{\frac{2i\pi nt}{T}}, \quad (1.21)$$

ou encore, en utilisant les coefficients trigonométriques :

$$\forall t, \frac{f(t^-) + f(t^+)}{2} = a_0 + \sum_{n \geq 1} a_n \cos\left(\frac{2\pi n t}{T}\right) + b_n \sin\left(\frac{2\pi n t}{T}\right). \quad (1.22)$$

Concrètement, ça signifie que en tout point de continuité de la fonction f , le DSF de f vaut $f(t)$. Si f est discontinue en t , son DSF est égal à $\frac{f(t^-) + f(t^+)}{2}$.

Ainsi, toute fonction périodique f est entièrement caractérisée par un ensemble dénombrable de coefficients $\{a_n\}_{n \in \mathbb{N}}$ et $\{b_n\}_{n \in \mathbb{N}^*}$: on peut recomposer la fonction f par la seule connaissance de ces coefficients, qui traduisent la contribution de la $n^{\text{ème}}$ harmonique (fréquence $\frac{n}{T}$) dans le signal f ⁶. Dit autrement, toute fonction périodique f (suffisamment régulière) est une **superposition** (pondérée) **de signaux monochromatiques** aux fréquences multiples de la fréquence fondamentale $\frac{1}{T}$ (fréquence de f).

Cette propriété est à la base du développement de l'analyse harmonique. Outre une représentation intuitive des signaux, elle permet de simplifier de nombreux problèmes en utilisant une telle décomposition, de telle sorte que l'analyse de Fourier se retrouve dans à peu près toutes les sciences de l'ingénieur : mécanique, acoustique, traitement du signal, analyse des systèmes linéaires, etc.

En particulier, on peut noter l'importance de ce développement dans l'optique de la transmission de ces signaux dans des systèmes linéaires invariants (ex : filtrage). En effet, ceux-ci ayant pour fonctions propres les signaux monochromatiques :

Ainsi, sous réserve de connaître $F(n\omega_0) \forall n \in \mathbb{Z}$, le développement en série de Fourier de l'entrée permet de déduire aisément le développement en série de Fourier de la sortie :

6. Concernant les coefficients c_n , leur interprétation est identique mais ceux-ci nécessitent de manipuler des fréquences négatives (puisque $k \in \mathbb{Z}$) qui n'ont pas d'interprétation physique parlante mais qui sont nécessaires au formalisme mathématique des séries de Fourier.

Cette remarque sur $F(n\omega_0)$, $n \in \mathbb{Z}$ mènera par la suite à la notion de réponse fréquentielle $F(\omega)$, $\omega \in \mathbb{R}$ (Fourier) et de fonction de transfert $F(p)$, $p \in \mathbb{C}$ (Laplace) du système.

Remarque : La formule (1.21) pour une fonction continue s'écrit :

$$\forall t, f(t) = \sum_{n \in \mathbb{Z}} c_n e^{\frac{2i\pi nt}{T}}.$$

Cette écriture ressemble étrangement à une décomposition de la fonction f sur une base (de fonctions); la notion de base est effectivement sous-jacente au développement en séries de Fourier.

Exemple : Reprenons l'exemple de la fonction triangle g traitée précédemment. Cette fonction est de classe \mathcal{C}^1 par morceaux et est continue. Ainsi, elle s'identifie en tout point à son développement en série de Fourier :

$$\forall t \in \mathbb{R}, g(t) = \frac{1}{2} + \sum_{p \geq 0} \frac{-4}{(2p+1)^2 \pi^2} \cos\left(\frac{2\pi(2p+1)t}{L}\right).$$

De plus, on peut (par exemple) aisément en déduire l'égalité (loin d'être triviale) suivante, en prenant $t = \frac{L}{2\pi}$:

$$\sum_{p \geq 0} \frac{-4}{(2p+1)^2 \pi^2} \cos(2p+1) = \frac{1}{\pi} - \frac{1}{2}.$$

On donne ci-après un exemple de synthèse de Fourier pour les deux signaux g et h dont on a calculé les coefficients de Fourier précédemment.

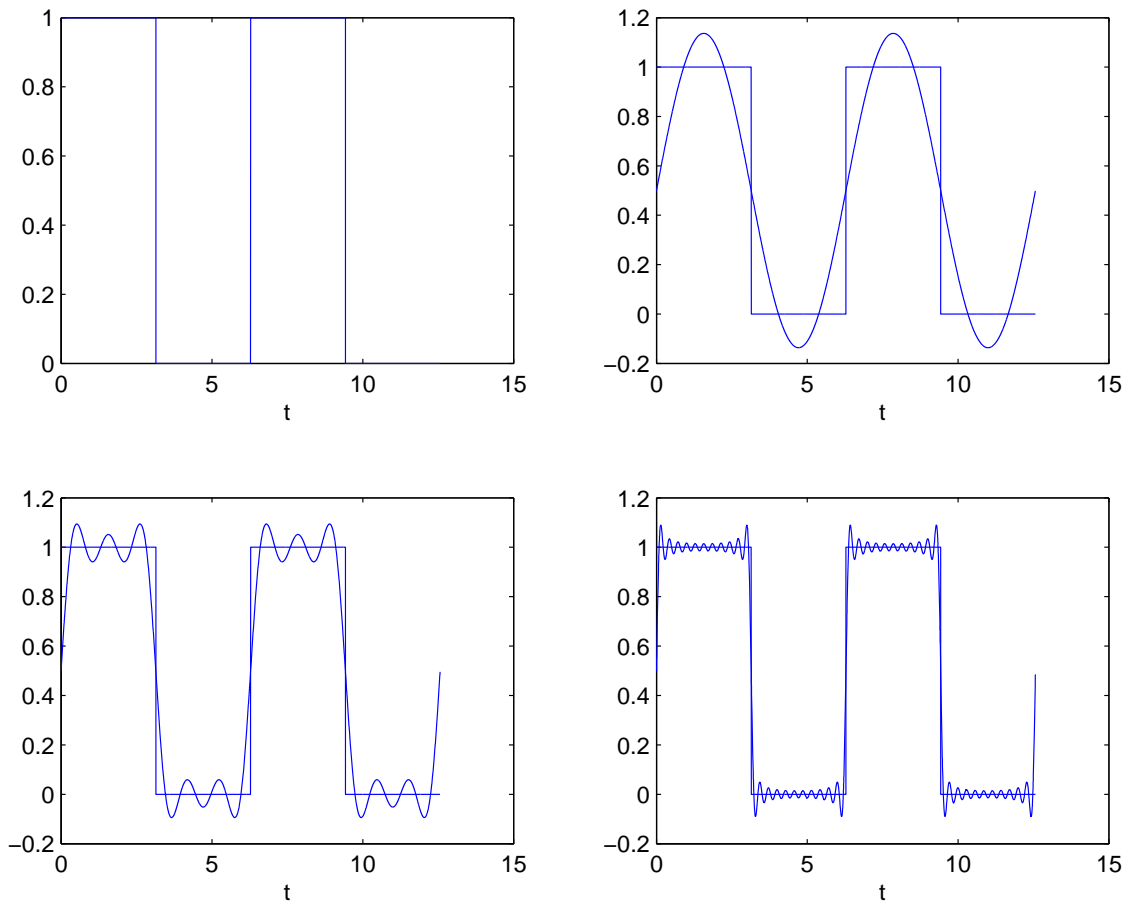


FIGURE 1.1 – Synthèse de Fourier de la fonction h jusqu'à la première, 5^e et 21^e harmonique.

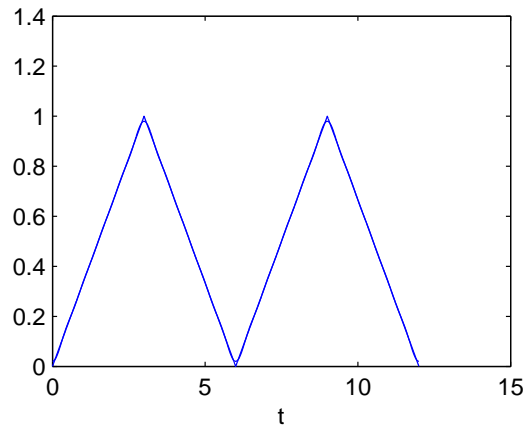
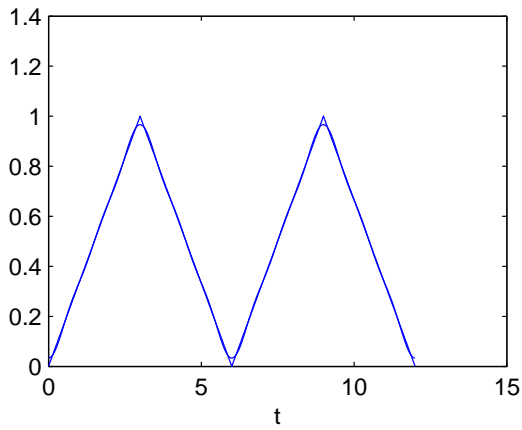
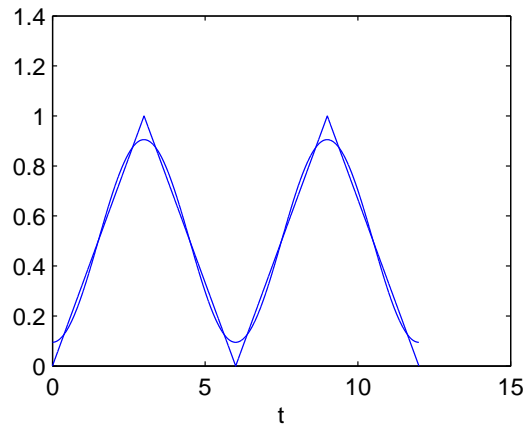
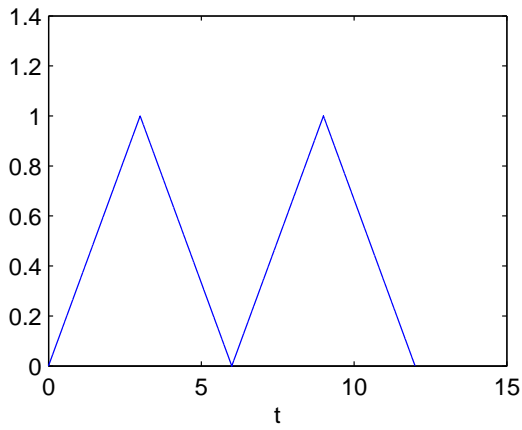


FIGURE 1.2 – Synthèse de Fourier de la fonction g ($L = 6$) jusqu'à la première, 5^e et 9^e harmonique.

1.4.2 Propriétés et théorème important

Notons tout d'abord que, dans le cas où f est une fonction à valeurs réelles, on a les relations suivantes entre les coefficients exponentiels et trigonométriques :

$$c_0 = a_0, \forall n \in \mathbb{N}^*, c_n = \frac{a_n - ib_n}{2}, c_{-n} = \frac{a_n + ib_n}{2} = \overline{c_n}$$

et réciproquement :

$$a_0 = c_0 ; \forall n \in \mathbb{N}^*, a_n = c_n + c_{-n}, b_n = i(c_n - c_{-n}).$$

On a la propriété :

Proposition 1.4.4

Si f est impaire, alors $a_n = 0 \forall n \in \mathbb{N}$ et $b_n = \frac{4}{T} \int_0^{T/2} f(t) \sin\left(\frac{2\pi nt}{T}\right) dt, \forall n \in \mathbb{N}^*$.

Si f est paire, alors $b_n = 0 \forall n \in \mathbb{N}^*$ et $a_n = \frac{4}{T} \int_0^{T/2} f(t) \cos\left(\frac{2\pi nt}{T}\right) dt, \forall n \in \mathbb{N}$ (et bien sûr $a_0 = \frac{2}{T} \int_0^{T/2} f(t) dt$).

Le théorème qui suit est fondamental : il donne une relation directe entre l'énergie d'un signal périodique et ses coefficients de Fourier.

Théorème 1.4.5 (Formule de Parseval)

Soit $f \in L^2([0, T])$ une fonction T -périodique. Alors :

$$\frac{1}{T} \int_a^{a+T} |f(t)|^2 dt = \sum_{n \in \mathbb{Z}} |c_n|^2. \quad (1.23)$$

Si on utilise les coefficients trigonométriques (fonction à valeurs réelles) :

$$\frac{1}{T} \int_a^{a+T} |f(t)|^2 dt = a_0^2 + \frac{1}{2} \sum_{n \geq 1} (|a_n|^2 + |b_n|^2).$$

Remarque : Comme souvent en mathématiques, ce théorème peut être employé de plusieurs manières. On peut utiliser cette relation pour donner l'expression de l'énergie d'un signal à partir de ses coefficients de Fourier, mais aussi par exemple pour calculer explicitement la valeur d'une série numérique "compliquée" (comme on l'a fait en utilisant le théorème de Dirichlet dans un exemple précédent).

Chapitre 2

Intégration

cf. cours de C. Jauberthie

Chapitre 3

Transformation de Fourier et peigne de Dirac

cf. cours de C. Jauberthie

Chapitre 4

Résolution d'équations différentielles linéaires

Pré-requis indispensables :

- Fonctions usuelles et propriétés.
- Savoir calculer des dérivées et primitives usuelles.

4.1 Introduction

Les équations différentielles (ordinaires, ou EDO, pour les distinguer des équations aux dérivées partielles, plus générales) sont des équations régissant l'évolution de phénomènes dynamiques ; de ce fait, il est évident qu'elles sont omniprésentes dans tous les domaines de la physique et il est indispensable de savoir les aborder avec méthode.

Ne seront abordées dans ce cours que les équations linéaires linéaires d'ordre 1 et 2 à coefficients constants, sachant que la majorité des résultats et techniques qui seront présentés s'étendent sans difficultés aux EDO linéaires d'ordre n . En dehors de ce cadre (coefficients non constants et ordre supérieur ou égal à 2, et a fortiori non linéaire), il existe d'innombrables résultats, cas particuliers, astuces, etc. qui nécessitent la lectures d'ouvrages dédiés.

Notons enfin que, comme pour les intégrales, les séries, etc., le nombre d'équations différentielles que l'on peut résoudre analytiquement est dans l'absolu très maigre... bien souvent, il faut faire appel à des méthodes de résolution numérique (Runge-Kutta etc.).

Définition 4.1.1 (*Équation différentielle*)

Une **équation différentielle** est une équation dont la solution est une fonction y , et impliquant les dérivées successives de cette dernière :

$$y^{(n)}(t) = F(t, y(t), y'(t), \dots, y^{(n-1)}(t)), \quad \forall t > t_0, \quad (4.1)$$

avec $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, et n un entier non nul, appelé **ordre** de l'EDO (c'est le plus haut degré de dérivation de y présent dans l'équation).

Une solution $y(t)$ d'une telle EDO devra évidemment être n fois dérivable sur Ω et vérifier la relation (4.1) à chaque instant t .

Remarques : — Il existe une autre forme de (4.1), un peu plus générale, dite non résolue :

$$G(t, y(t), y'(t), \dots, y^{(n-1)}(t), y^{(n)}(t)) = 0, \forall t > t_0.$$

— Il est fréquent de rencontrer des systèmes d'équations différentielles, c'est-à-dire un ensemble d'équations différentielles scalaires couplées les unes aux autres ; elles sont alors de la forme (4.1) mais avec F à valeurs dans \mathbb{R}^p .

Si l'on adjoint à l'équation (4.1) des conditions initiales, c'est-à-dire des valeurs imposées à l'instant¹ $t = 0$ pour $y, y', \dots, y^{(n-1)}$ (position initiale, vitesse initiale, etc. par exemple), et que l'on cherche les solutions de (4.1) satisfaisant ces conditions initiales, on parle de **problème de Cauchy**.

On peut montrer que, sous des hypothèses peu contraignantes sur la fonction F (continue et localement Lipschitzienne), on dispose d'un résultat fondamental, le théorème de **Cauchy-Lipschitz**, assurant l'existence et l'unicité d'une solution $y(t)$ au problème de Cauchy.

Dans le cadre des équations différentielles linéaires à coefficients continus dans lequel on se placera dans ce cours (coefficients constants en particulier), ce théorème est immédiatement vérifié et nous assure l'existence et l'unicité de solutions aux équations que l'on étudiera.

Définition 4.1.2 (Équation différentielle linéaire)

Une équation différentielle est dite **linéaire** lorsque F est une fonction linéaire de $y, y', \dots, y^{(n-1)}$; l'équation prend alors la forme :

$$\begin{cases} a_n(t)y^{(n)}(t) + \dots + a_1(t)y'(t) + a_0(t)y(t) = u(t), \forall t > t_0, \\ y(0) = y_0, y'(0) = y_1, \dots, y^{(p-1)}(0) = y_{p-1} \text{ (C.I.)}, \end{cases} \quad (4.2)$$

où les $a_i : I \rightarrow \mathbb{R}$ et $u : I \rightarrow \mathbb{R}$ (appelé **second membre** ou **terme source**) sont des fonctions continues.

Lorsque les $a_i(t)$ sont des constantes, l'équation est dite **à coefficients constants**.

Alors on a le résultat suivant, dont la démonstration est admise.

Théorème 4.1.3 (Cauchy-Lipschitz)

Le problème de Cauchy (4.2) admet une unique solution $y(t)$ définie sur tout I .

Exemples : — Équation (non linéaire) de l'évolution de l'angle θ d'un pendule simple de longueur l , sans viscosité :

$$\begin{cases} \ddot{\theta}(t) + \frac{g}{l} \sin(\theta(t)) = 0, \\ \theta(0) = \frac{\pi}{4}, \dot{\theta}(0) = 0. \end{cases}$$

1. Il s'agit bien sûr d'une convention : toute autre instant initial t_0 est possible.

— Équation (linéaire) de la décharge d'un condensateur dans un circuit RC :

$$\begin{cases} u(t) + RC u'(t) = 0, \\ u(0) = u_0. \end{cases}$$

4.2 Résolution des équations différentielles linéaires

4.2.1 Méthodologie générale

On abordera dans ce cours la méthodologie pour les EDO linéaires à coefficients constants, principalement d'ordre 1 et 2 (même si le cas général sera brièvement traité pour l'ordre 1, et des indications seront données pour l'ordre n à coefficient constant).

Résoudre ces équations différentielles se déroulera toujours en 3 étapes fondamentales :

- (i) Résolution de l'**équation homogène** (sans second membre) \rightarrow polynôme caractéristique, racines \rightarrow solution générale $y_h(t)$ de l'équation homogène.
- (ii) Détermination d'une solution particulière $y_p(t)$ de l'équation **avec second membre** (intuition, règles de second membre polynomial/exponentiel, ou méthode de la variation de la constante).
- (iii) Alors on a la solution globale : $y(t) = y_h(t) + y_p(t)$, et on identifie les constantes présentes dans $y_h(t)$ en utilisant les conditions initiales.

4.2.2 EDO linéaires d'ordre 1

Le cas des EDO linéaires d'ordre 1 sera traité dans le cas général (i.e. coefficients non constants) fait de sa simplicité. Une EDO d'ordre linéaire d'ordre 1 est de la forme² :

$$\begin{cases} y'(t) = a(t)y(t) + u(t), \\ y(0) = y_0. \end{cases} \quad (4.3)$$

où $a(t)$ et $u(t)$ sont des fonctions continues³ et $y_0 \in \mathbb{R}$ (condition initiale).

(i) Résolution de l'équation homogène

On considère donc dans un premier temps l'équation sans second membre :

$$y'(t) = a(t)y(t). \quad (4.4)$$

2. Ou plus précisément le problème de Cauchy qui y est associé.

3. Cette hypothèse sur $u(t)$ peut en réalité être affaiblie.

En divisant par ⁴ $y(t)$, on peut écrire : $\frac{y'(t)}{y(t)} = -a(t)$, ce qui donne après intégration $\ln |y(t)| = \int_0^t a(s)ds + K$ et par suite $|y(t)| = B e^{A(t)}$, où $A(t)$ est une primitive de la fonction $a(t)$.

On a donc démontré la proposition suivante (quitte à poser $C = -B$ si $B < 0$) :

Proposition 4.2.1 (Solutions de l'équation homogène)

Les solutions de l'équation homogène sont de la forme :

$$y_h(t) = C e^{A(t)}, \quad C \in \mathbb{R}, \quad (4.5)$$

où $A(t)$ est une primitive de $a(t)$.

Notons que dans le cas où $a(t) = a$ constante, on a le résultat classique :

$$y_h(t) = C e^{at}, \quad C \in \mathbb{R}, \quad (4.6)$$

résultat qui peut également s'obtenir par la méthode du polynôme caractéristique (cf. section 4.2.3).

Remarque : Un résultat essentiel bien que non approfondi dans ce cours est que l'ensemble des solutions de l'équation homogène (4.4), muni des lois classiques d'addition et de multiplication par un scalaire, est un **espace vectoriel de dimension 1**. On constate alors dans l'expression (4.5) que $e^{A(t)}$ est une base de cet espace vectoriel. Ce résultat, fondamental, s'étend aux EDO d'ordre n à coefficients constants (on est alors en présence d'un e.v. de dimension n , engendré là-encore par une base de fonctions exponentielles).

(ii) Solution particulière de l'équation avec second membre

On considère cette fois l'équation complète :

$$y'(t) = a(t)y(t) + u(t), \quad (4.7)$$

et on en cherche une solution particulière, sans se soucier des conditions initiales ⁵.

C'est l'étape la plus délicate de la résolution. Elle peut se faire à l'intuition dans des cas simples, soit par une méthode systématique appelée variation de la constante ; enfin, quelques règles existent également pour le cas spécifique des équations à coefficients constants.

4. Cette opération est licite en vertu du théorème de Cauchy-Lipschitz, qui garantit que $y(t) \neq 0 \forall t$ pour toute solution autre que la solution identiquement nulle. En effet, supposons qu'il existe une solution non identiquement nulle, qui s'annule en un instant t_0 ; le problème de Cauchy sur $[t_0, +\infty[: y'(t) = a(t)y(t), y(t_0) = 0$ admet une **unique** solution (par Cauchy-Lipschitz), qui ne peut que la solution identiquement nulle puisque celle-ci en est également solution.

5. Ce point est important : on lit parfois qu'on considère l'EDO avec second membre et conditions initiales nulles ; ce n'est pas une nécessité : il faut retenir que la valeur prise en $t = 0$ par la solution particulière importe peu puisqu'elles sera corrigée par la dernière étape de résolution

► **Méthode de la variation de la constante**

Elle présente l'avantage d'être très simple (sur le principe du moins) et d'être tout le temps valable. Son inconvénient étant de nécessiter à son tour la résolution d'une équation différentielle, qui peut s'avérer impossible analytiquement.

Le nom de cette méthode est explicite : elle consiste à rechercher une solution particulière sous la même forme que celle de y_h mais en considérant que la constante C est cette fois une fonction du temps :

$$y_p(t) = C(t) e^{A(t)},$$

l'objectif étant alors de déterminer une fonction $C(t)$. Pour cela, on injecte cette solution dans l'équation (4.7), et on fait apparaître, après simplification (à faire pour s'entraîner!), une équation (elle aussi différentielle!) vérifiée par $C(t)$:

$$C'(t) = u(t) e^{-A(t)}. \tag{4.8}$$

Après résolution de cette nouvelle équation différentielle (dont il suffit de trouver **une** solution $C_p(t)$, peu importe laquelle, le plus simple sera le mieux!), on a entièrement déterminé notre solution particulière qui s'exprimera $y_p(t) = C_p(t) e^{-A(t)}$.

► **Cas à coefficients constants : second membre polynôme-exponentiel**

Si $a(t) = a$, il existe des règles bien pratiques pour la recherche de solution particulière dans le cas où le second membre $u(t)$ possède une forme bien particulière. On donne ci-après un des principaux résultats.

Proposition 4.2.2

Si $u(t) = Q(t)e^{rt}$ où r est un réel et Q un polynôme à coefficients réels, alors (4.7) admet une solution de la forme

$$\begin{aligned} y_p(t) &= P(t)e^{rt}, \text{ si } r \neq a, \\ y_p(t) &= tP(t)e^{rt}, \text{ si } r = a, \end{aligned}$$

où P est un polynôme de même degré que $Q(t)$.

Démonstration : La fonction $t \mapsto e^{rt}$ étant continûment dérivable et non nulle sur I , on peut rechercher une solution de (4.7) sous la forme : $\hat{y}(t) = u(t)e^{rt}$. On a :

$$y_p'(t) = a y_p(t) + u(t) \iff u'(t)e^{rt} + u(t)re^{rt} = a u(t)e^{rt} + Q(t)e^{rt} \iff u'(t) + u(t)(r-a) = Q(t).$$

- si $r = a$, alors u est une primitive de Q c'est à dire un polynôme de degré $q + 1$,

- si $r \neq a$, alors si u polynôme de degré p , par identification des coefficients on a $p = q$. ■

Ainsi, ce résultat renseigne sur la structure de la solution particulière. Il reste évidemment à **poser** $y_p(t)$ de cette forme, **l'injecter dans l'équation (4.7)** et résoudre pour pleinement la **déterminer** !

Exemple : $y_p(t) = -(t^2 + 2t + 3)e^t$ est une solution particulière de $y'(t) = 2y(t) + (t^2 + 1)e^t$.

► **Principe de superposition**

Pour finir, signalons une propriété fondamentale des EDO linéaires, très pratique pour les calculs de solutions particulières.

Proposition 4.2.3 (Principe de superposition)

Soient y_{p_u} et y_{p_v} deux solutions particulières respectives d'une EDO du premier ordre $y' = ay + u$ lorsque le second membre $u(t)$ est respectivement $u(t) = f(t)$ et $u(t) = g(t)$. Alors, $y_{p_u} + y_{p_v}$ est une solution particulière de cette même équation différentielle avec pour second membre $u(t) = f(t) + g(t)$.

En pratique, cela signifie que pour trouver une solution particulière à une EDO dont le second membre est une somme de termes simples, il suffit de trouver une solution particulière pour chacun des termes source, la solution particulière globale étant alors la somme de ces solutions particulières élémentaires.

Exemple : On cherche une solution particulière à l'EDO suivante :

$$y'(t) = ay(t) + Q(t) \cos(\omega t). \tag{4.9}$$

Comme $\cos(\omega t) = \frac{e^{i\omega t} + e^{-i\omega t}}{2}$, on sait alors qu'il nous suffira de trouver une solution particulière pour chacune des équations suivantes (avec $R(t)$ qui est simplement $Q(t)/2$) :

$$y'(t) = ay(t) + R(t)e^{i\omega t} \text{ et } y'(t) = ay(t) + R(t)e^{-i\omega t},$$

chose qui sera aisée en utilisant la proposition 4.2.2 ; en additionnant ces deux solutions particulières, on a alors une solution particulière de (4.9) d'après le principe de superposition.

Second membre polynomial-sinusoïdal : Du fait du principe de superposition, on peut aisément déduire de la proposition 4.2.2 une règle concernant les second membres sinusoidaux. Il est en effet aisé de montrer que si $u(t) = Q_1(t) \cos(\omega t + \varphi) + Q_2(t) \sin(\omega t + \varphi)$, alors on pourra chercher une solution particulière sous la forme $y_p(t) = P_1(t) \cos(\omega t + \varphi) + P_2(t) \sin(\omega t + \varphi)$.

(iii) Solution globale

On sait alors que l'unique solution du problème (4.3) est donnée par la somme de de "la" solution générale de l'équation homogène et de la solution particulière précédemment déterminées :

$$y(t) = y_h(t) + y_p(t). \tag{4.10}$$

Cette dernière étape (très simple) consiste simplement à déterminer la valeur de la constante C présente dans y_h en utilisant la condition initiale en posant $y_h(0) + y_p(0) = y_0$, qui devient $C + y_p(0) = y_0$, soit $C = y_0 - y_p(0)$.

Exercice d'application : Soit l'équation différentielle définie sur $[1, +\infty[$:

$$ty'(t) = 2y(t) + t^3; \quad y(1) = 3.$$

(i) Solutions générales de l'équation homogène données par :

$$y_h(t) = Ce^{2\ln|t|} = Ct^2, C \in \mathbb{R} \quad (4.11)$$

(ii) Détermination d'une solution particulière par variation de la constante : on cherche y_p sous la forme $y_p(t) = C(t)t^2$, et donc $y_p'(t) = C'(t)t^2 + 2tC(t)$; en reportant dans l'EDO (4.11), on obtient :

$$tC'(t)t^3 + 2t^2C(t) = 2C(t)t^2 + t^3 \Leftrightarrow C'(t) = 1.$$

On peut donc prendre $C(t) = t$, ce qui donne la solution particulière :

$$y_p(t) = t^3.$$

(iii) Les solutions de (4.11) sur $[1, +\infty[$ ont donc pour expression $y(t) = C(t)t^2 + t^3$, $C \in \mathbb{R}$, et la solution telle que $y(1) = 3$ est donc obtenue pour $C = 2$:

$$\boxed{y(t) = 2t^2 + t^3.}$$

Cas vectoriel

Dans le cas vectoriel $Y'(t) = A(t)Y(t)$, où $A(t)$ est une matrice $n \times n$, il n'existe pas de méthode générale permettant le calcul analytique des solutions $Y_h(t)$ de l'équation homogène. De nombreux résultats existent cependant.

En particulier, si la matrice $A(t)$ est constante, il est possible d'étendre le résultat scalaire en définissant la notion bien pratique d'exponentielle de matrice (on en reparlera dans le chapitre d'algèbre linéaire); les résultats sont alors similaires à ceux établis dans le cas scalaire (à coefficients constants).

4.2.3 EDO linéaires à coefficients constants d'ordre 2

On s'intéresse aux équations à coefficients constants car mis à part les EDO d'ordre 1 (inclu), il n'y a pas de méthode systématique permettant de résoudre les EDO à coefficients variables.

Une EDO du second ordre à coefficients constants est de la forme :

$$\begin{cases} y''(t) = ay'(t) + by(t) + u(t), \\ y(0) = y_0, y'(0) = y_1, \end{cases} \quad (4.12)$$

où I est un intervalle de \mathbb{R} et $a, b \in \mathbb{R}$, le second membre $u : I \rightarrow \mathbb{R}$ est une fonction continue, et y_0, y_1 les conditions initiales sur y .

La méthodologie introduite dans la section 4.2.1 reste évidemment valable.

(i) Résolution de l'équation homogène

On dispose du résultat fondamental suivant :

Théorème 4.2.4

Soient r_1 et r_2 les deux racines de l'équation caractéristique de (4.12) :

$$r^2 - ar - b = 0, \quad (4.13)$$

et Δ son discriminant.

Alors, les solutions de l'équation homogène sont de la forme :

— si $\Delta > 0$ (r_1 et r_2 racines distinctes réelles) :

$$y(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t}, \quad C_1, C_2 \in \mathbb{R},$$

— si $\Delta = 0$ (racine réelle double $r_1 = r_2 = \alpha$) :

$$y(t) = (C_1 t + C_2) e^{\alpha t}, \quad C_1, C_2 \in \mathbb{R},$$

— si $\Delta < 0$ (racines complexes conjuguées : $r_1 = \bar{r}_2 = \alpha + i\omega$) :

$$y(t) = C_1 e^{r_1 t} + C_2 e^{\bar{r}_1 t}, \quad C_1, C_2 \in \mathbb{R},$$

ce qui peut être écrit, de manière équivalente :

$$y(t) = (C_1 \sin(\omega t) + C_2 \cos(\omega t)) e^{\alpha t}, \quad C_1, C_2 \in \mathbb{R}.$$

Démonstration : La fonction $t \mapsto e^{r_1 t}$ étant deux fois dérivable et non nulle sur I , on peut chercher les solutions sur I de l'équation homogène sous la forme $y(t) = u(t)e^{r_1 t}$. On a alors :

$$y'(t) = (u'(t) + r_1 u(t)) e^{r_1 t} \text{ et } y''(t) = (u''(t) + 2r_1 u'(t) + r_1^2 u(t)) e^{r_1 t},$$

d'où :

$$\begin{aligned} & y''(t) = ay'(t) + by(t) \\ \iff & (u''(t) + 2r_1 u'(t) + r_1^2 u(t)) e^{r_1 t} = [b(u'(t) + r_1 u(t)) + au(t)] e^{r_1 t} \\ \iff & [u''(t) + (2r_1 - b)u'(t) + (r_1^2 - br_1 - a)u(t)] e^{r_1 t} = 0, \end{aligned}$$

soit, puisque r_1 racine de (4.13) et $e^{r_1 t} > 0$:

$$u''(t) + (2r_1 - b)u'(t) = 0.$$

On pose $\Delta = b^2 + a$ et on distingue alors 3 cas différents :

— si $\Delta > 0$ alors les racines de (4.13) sont $\frac{b \pm \sqrt{\Delta}}{2}$ donc $r_1 \neq \frac{b}{2} \implies 2r_1 - b \neq 0$. En posant $z = u'$, on se ramène donc à l'équation du premier ordre suivante :

$$z'(t) + (2r_1 - b)z(t) = 0,$$

dont les solutions sont données par $z(t) = Ce^{-(2r_1-b)t}$, $C \in \mathbb{R}$. Les solutions u sont donc données par les primitives de z c'est à dire par :

$$u(t) = -\frac{1}{(2r_1-b)}Ce^{-(2r_1-b)t} + C_1, C, C_1 \in \mathbb{R}.$$

On a alors :

$$y(t) = \left(-\frac{1}{(2r_1-b)}Ce^{-(2r_1-b)t} + C_1\right)e^{r_1t} = -\frac{1}{(2r_1-b)}Ce^{(b-r_1)t} + C_1e^{r_1t},$$

soit, en posant $C_2 = -\frac{1}{(2r_1-b)}C$ et puisque $r_2 = b - r_1$:

$$y(t) = C_2e^{r_2t} + C_1e^{r_1t}.$$

- si $\Delta = 0$ alors $r_1 = r_2 = \frac{b}{2}$, d'où $u''(t) = 0$ et donc $u(t) = C_1t + C_2$, $C_1, C_2 \in \mathbb{R}$.
- si $\Delta < 0$ alors $\bar{r}_1 = r_2$. La démarche est la même que dans le cas $\Delta > 0$, la seule différence étant que, comme r_1 et r_2 sont complexes, C_1 et C_2 le sont aussi. Comme l'on cherche des solutions à valeurs réelles, on prendra $\bar{C}_1 = C_2$. Ainsi les solutions s'écriront :

$$y(t) = \bar{C}_1e^{\bar{r}_1t} + C_1e^{r_1t} = 2\operatorname{Re}(C_1e^{r_1t}) \in \mathbb{R},$$

c'est à dire :

$$\begin{aligned} y(t) &= 2\operatorname{Re}\left((\operatorname{Re}(C_1) + i\operatorname{Im}(C_1))e^{(\operatorname{Re}(r_1)+i\operatorname{Im}(r_1))t}\right) \\ &= 2[\operatorname{Re}(C_1)\cos(\operatorname{Im}(r_1)t) - \operatorname{Im}(C_1)\sin(\operatorname{Im}(r_1)t)]e^{\operatorname{Re}(r_1)t} \\ &= [\mu_2\cos(\operatorname{Im}(r_1)t) + \mu_1\sin(\operatorname{Im}(r_1)t)]e^{\operatorname{Re}(r_1)t} \end{aligned}$$

avec $\mu_1 = -2\operatorname{Im}(C_1)$ et $\mu_2 = 2\operatorname{Re}(C_1)$. ■

Remarque : Comme établi dans la preuve ci-dessus, les solutions de l'équation homogène forment un sous espace vectoriel de dimension 2 de l'espace vectoriel $\mathcal{C}^2(I, \mathbb{R})$ des fonctions deux fois continûment dérivables sur I à valeur dans \mathbb{R} .

(ii) Solution particulière de l'équation avec second membre

Comme dans le cas des systèmes du premier ordre, cette étape est la plus délicate. On cherche une solution particulière $y_p(t)$ à l'équation différentielle avec second membre sans se soucier de la condition initiale.

Les techniques dont on dispose sont les mêmes :

► Méthode de la variation de la constante

La démarche est similaire à celle effectuée dans le cadre des systèmes du premier ordre, et considérer l'expression de la solution de l'équation homogène y_h en considérant que les constantes sont des fonctions de t et vérifiant :

$$\begin{cases} y_p(t) = C_1(t)v_1(t) + C_2(t)v_2(t), \\ y_p'(t) = C_1(t)v_1'(t) + C_2(t)v_2'(t). \end{cases}$$

et injecter cette solution particulière dans l'équation avec second membre. On obtient alors le système :

$$\begin{cases} C_1'(t)v_1(t) + C_2'(t)v_2(t) = 0, \\ C_1'(t)v_1'(t) + C_2'(t)v_2'(t) = b(t). \end{cases}$$

La résolution de ce système (différentiel) permet de déterminer $C_1(t)$ et $C_2(t)$ et ainsi déterminer la solution particulière recherchée.

► **Second membre polynôme-exponentiel**

De même que dans le cadre des équations d'ordre 1, il existe des résultats bien pratiques pour la détermination d'un second membre lorsque le second membre de l'équation est polynomial et/ou exponentiel.

Proposition 4.2.5

Si $u(t) = Q(t)e^{rt}$ où r est un réel et Q un polynôme à coefficients réels de degré q , l'EDO avec second membre admet une solution de la forme :

$$\begin{aligned} y_p(t) &= P(t)e^{rt}, \text{ si } r \text{ n'est pas racine de l'équation caractéristique (4.13),} \\ y_p(t) &= tP(t)e^{rt}, \text{ si } r \text{ est racine simple de l'équation caractéristique (4.13),} \\ y_p(t) &= t^2P(t)e^{rt}, \text{ si } r \text{ est racine double de l'équation caractéristique (4.13),} \end{aligned}$$

où $P(t)$ est un polynôme de même degré que $Q(t)$.

Démonstration : Similaire à celle de la proposition 4.2.2. ■

Ici encore, ce résultat renseigne sur la structure de la solution particulière. Il reste évidemment à **poser** $y_p(t)$ de cette forme, **l'injecter dans l'équation** et résoudre pour pleinement la **déterminer**.

► **Principe de superposition**

Celui reste valable (il est valable quel que soit l'ordre de l'EDO, c'est une propriété de la linéarité de l'équation) : pour trouver la solution particulière, il suffit de déterminer une solution particulière pour chaque terme qui compose le second membre, et additionner le tout.

Second membre polynomial-sinusoïdal : Ici encore, on peut utiliser les résultats ci-dessous pour traiter les second membre de type sinusoïdal, en faisant toutefois attention aux différents cas possible (fonction des racines de l'éq caractéristique) de la proposition 4.2.5. Un exemple sera traité en TD.

(iii) Solution globale

On sait alors que l'unique solution du problème (4.12) est donnée par la somme de de "la" solution générale de l'équation homogène et de la solution particulière précédemment

déterminées :

$$y(t) = y_h(t) + y_p(t). \quad (4.14)$$

Cette dernière étape, toujours très simple, consiste à déterminer la valeur des constantes C_1 et C_2 présentes dans y_h en utilisant les conditions initiales $y(0) = y_0$ et $y'(0) = y_1$.

Chapitre 5

Algèbre linéaire

Pré-requis indispensables :

Dans ce cours, \mathbb{K} désigne un corps commutatif, typiquement \mathbb{R} ou \mathbb{C} muni des lois usuelles d'addition et de multiplication. On utilisera souvent l'abréviation e.v. pour espace vectoriel. Bien que beaucoup de notions introduites dans ce chapitre soient généralisables, nous travaillerons essentiellement sur des espaces vectoriels de dimension finie.

5.1 Espaces Vectoriels

5.1.1 Définitions

Soient E un ensemble et \mathbb{K} un corps commutatif. On note $+$ (resp. \cdot) une loi interne sur E (resp. loi externe), c'est-à-dire opérant sur deux éléments de E (resp. sur un élément de E et un élément de \mathbb{K}) :

$$\begin{aligned} + : E \times E &\rightarrow E & \cdot : \mathbb{K} \times E &\rightarrow E \\ (x, y) &\mapsto x + y & (x, \lambda) &\mapsto \lambda \cdot x \end{aligned}$$

Définition 5.1.1

On dit que $(E, +, \cdot)$ est un **espace vectoriel sur \mathbb{K}** (ou \mathbb{K} -e.v) si :

- $(E, +)$ est un groupe commutatif (ou abélien) :
 - la loi $+$ est commutative sur E : $\forall x, y \in E, x + y = y + x$
 - la loi $+$ est associative : $\forall x, y, z \in E, x + (y + z) = (x + y) + z$
 - il existe un élément neutre dans E , noté 0_E , vérifiant : $\forall x \in E, x + 0_E = x$
 - tout élément x de E admet un symétrique pour $+$, noté $(-x)$, tel que : $x + (-x) = 0_E$
(moyen mnémotechnique pour les axiomes du groupe commutatif : CANS)

- La loi externe \cdot vérifie :
 - $\forall \lambda, \mu \in \mathbb{K}, \forall x \in E, (\lambda + \mu).x = \lambda.x + \mu.x,$
 - $\forall \lambda \in \mathbb{K}, \forall x, y \in E, \lambda.(x + y) = \lambda.x + \lambda.y$
 - $\forall \lambda, \mu \in \mathbb{K}, (\lambda\mu).x = \lambda.(\mu.x)$
 - $\forall x \in E, 1_{\mathbb{K}}.x = x$

Terminologie : Les éléments de \mathbb{K} sont communément appelés des **scalaires** et ceux de E des **vecteurs**.

Les espaces vectoriels sont donc des ensembles d'objets de même nature (communément appelés *vecteurs* dans le cas classique $E = \mathbb{R}^N$) possédant certaines propriétés et munis d'une loi d'addition et de multiplication par un scalaire.

Attention : Dans ce chapitre, on veillera à toujours garder à l'esprit la nature des objets que l'on manipule. Ne pas confondre la loi sur le corps \mathbb{K} et la loi sur l'espace vectoriel E , qui agissent sur des objets de nature différente. Par exemple, lorsqu'on écrit $(\lambda\mu).x$, $\lambda\mu$ est un produit entre deux éléments du corps \mathbb{K} , alors que le \cdot désigne le produit entre un élément de \mathbb{K} (ici $\lambda\mu$) et un élément x de E . La notation $+$ est quant à elle utilisée pour \mathbb{K} et E .

Exemple : 1. \mathbb{R}^n est un \mathbb{R} -e.v s'il est muni des lois $+$ et \cdot définies pour tous $x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$ par :

$$x + y := (x_1 + y_1, \dots, x_n + y_n)$$

$$\lambda.x := (\lambda x_1, \dots, \lambda x_n)$$

2. $\mathcal{C}^0(I, \mathbb{R}) = \{\text{fonctions continues de } I \subset \mathbb{R} \text{ dans } \mathbb{R}\}$ est un \mathbb{R} -e.v s'il est muni des lois $+$ et \cdot définies pour tous $f, g \in E, \lambda \in \mathbb{R}$ par :

$$\forall x \in I, (f + g)(x) := f(x) + g(x) \text{ et } \forall x \in I, (\lambda.f)(x) := \lambda f(x)$$

Proposition 5.1.2

Soit E un \mathbb{K} -e.v. Pour tout $\lambda \in \mathbb{K}$ et $x \in E$, on a :

$$\lambda.x = 0_E \Leftrightarrow (\lambda = 0_{\mathbb{K}} \text{ ou } x = 0_E)$$

Démonstration : $(\Leftarrow) \lambda.0_E = \lambda.(0_E + 0_E) \Leftrightarrow \lambda.0_E = \lambda.0_E + \lambda.0_E$ donc $\underbrace{\lambda.0_E + (-\lambda.0_E)}_{0_E} =$

$$\lambda.0_E + \underbrace{\lambda.0_E + (-\lambda.0_E)}_{0_E} \text{ d'où } 0_E = \lambda.0_E + 0_E \text{ et donc } 0_E = \lambda.0_E$$

(\Rightarrow) Soient $\lambda \in \mathbb{K}$ et $x \in E$ tels que $\lambda.x = 0_E$. Si $\lambda \neq 0_{\mathbb{K}}$, alors $x = (\lambda^{-1}\lambda).x = \lambda^{-1}.(\lambda.x) = \lambda^{-1}.0_E = 0_E$. ■

5.1.2 Sous-espaces vectoriels

Définition 5.1.3

Soit un \mathbb{K} -e.v E et une partie **non vide** $F \subset E$. On dit que F est un **sous-espace vectoriel** (s.e.v) de E si :

1. F est stable pour $+$: $\forall x, y \in F$, on a $x + y \in F$
2. F est stable pour \cdot : $\forall \lambda \in \mathbb{K}, \forall x \in F$, on a $\lambda.x \in F$.

Remarque : Ces conditions peuvent se résumer à : F stable par combinaison linéaire, i.e. $\forall \lambda \in \mathbb{K}, \forall x, y \in F, \lambda.x + y \in F$.

Exemple : 1. $\mathbb{R} \times \{0\}$ est un s.e.v de \mathbb{R}^2 . Plus généralement, tout ensemble de la forme $\{(x, \lambda.x), \lambda \in \mathbb{K}\}$ est un s.e.v de \mathbb{R}^2 .
2. $\mathcal{C}^1(\mathcal{I}, \mathbb{R}) = \{\text{fonctions continuellement différentiables de } \mathcal{I} \subset \mathbb{R} \text{ dans } \mathbb{R}\}$ est un s.e.v de $\mathcal{C}^0(\mathcal{I}, \mathbb{R})$.

La proposition suivante est utile pour établir qu'un ensemble est un espace vectoriel.

Proposition 5.1.4

Tout sous-espace vectoriel est un espace vectoriel.

Démonstration : $(F, +)$ groupe commutatif :

- la commutativité et l'associativité de $+$ dans F découle des propriétés de $+$ dans E .
- neutre : comme F stable pour $+$, on a $(-1).x + x = ((-1) + 1).x = 0_E \in F$, donc 0_E neutre pour $+$ dans F .

► symétrique : $\forall x \in F, (-1).x \in F$ est le symétrique de x pour $+$, noté $-x$.

Les axiomes sur la loi \cdot se vérifient grâce aux propriétés de \cdot sur E . ■

5.1.3 Familles libres et génératrices

Définition 5.1.5

Soit E un \mathbb{K} -e.v et x_1, \dots, x_p des éléments de E . On appelle **combinaison linéaire de** x_1, \dots, x_p tout élément de la forme :

$$\lambda_1.x_1 + \dots + \lambda_p.x_p, \text{ où } \lambda_i \in \mathbb{K} \text{ pour tout } i = 1 : p.$$

On pourra noter cette combinaison : $\sum_{i=1}^p \lambda_i.x_i$.

Exemple : $E = \mathbb{R}^4$, soient x_1, x_2, x_3 des vecteurs de \mathbb{R}^4 . Alors $y = 3.x_1 - 2.x_3$ est une combinaison linéaire de x_1, x_2, x_3

Définition 5.1.6

On dit que la famille (x_1, \dots, x_p) d'éléments de E est **libre** si :

$$\forall \lambda_1, \dots, \lambda_p \text{ dans } \mathbb{K}, \sum_{i=1}^p \lambda_i \cdot x_i = 0_E \Rightarrow \lambda_i = 0_{\mathbb{K}} \text{ pour tout } i = 1 : p.$$

Dans le cas contraire, on dit que la famille est **liée** :

$$\exists \lambda_1, \dots, \lambda_p \text{ dans } \mathbb{K} \text{ non tous nuls tels que } \sum_{i=1}^p \lambda_i \cdot x_i = 0_E.$$

Lorsqu'une famille est libre, on dit aussi que les vecteurs qui la constituent sont **linéairement indépendants** ; cela signifie qu'aucun vecteur de la famille ne peut s'exprimer comme combinaison linéaire des autres. Cette notion généralise la notion de vecteurs colinéaires dans \mathbb{R}^2 .

Exemples : 1. $E = \mathbb{R}^2$. Les vecteurs $x_1 = (1, 2)$ et $x_2 = (-3, -6)$ sont liés car $3 \cdot x_1 + x_2 = 0$ (ils sont colinéaires). En revanche, soit $x_3 = (-1, 3)$; la famille (x_1, x_3) est libre car $\forall \lambda_1, \lambda_3$ dans \mathbb{K} :

$$\lambda_1 \cdot x_1 + \lambda_3 \cdot x_3 = 0_{\mathbb{R}^2} \Leftrightarrow \begin{cases} \lambda_1 - \lambda_3 = 0 \\ 2\lambda_1 + 3\lambda_3 = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_3 = 0. \end{cases}$$

2. $E = \mathbb{R}^3$. Les vecteurs $x_1 = (1, 2, -1)$, $x_2 = (-1, 0, 4)$ et $x_3 = (2, 2, -5)$ ne forment pas une famille libre (i.e. elle est liée) car $x_3 = x_1 - x_2$, soit encore $x_1 - x_2 - x_3 = 0$.

Proposition-Définition 5.1.7

Soit E un \mathbb{K} -e.v et A une partie non vide de E . L'ensemble des combinaisons linéaires des éléments de A est un s.e.v de E , appelé **sous-espace vectoriel engendré par A** , et noté $\text{Vect}(A)$. C'est le plus petit s.e.v contenant A .

Si de plus $\text{Vect}(A) = E$, on dit que A **engendre** E , ou encore que la famille A est **génératrice**.

Démonstration : ► Montrons que $\text{Vect}(A)$ est un s.e.v de E . Tout d'abord $\text{Vect}(A)$ est non vide puisqu'il contient évidemment 0_E (qui est combinaison linéaire de n'importe quels $a_i \in A$ avec des coefficients nuls). Ensuite, il est clair que $\text{Vect}(A) \subset E$ puisque ses éléments sont des combinaisons linéaires d'éléments de A , donc de E , et donc appartient à E puisque c'est un espace vectoriel. Enfin, montrons que $\text{Vect}(A)$ est stable par combinaison linéaire. Soient $\lambda \in \mathbb{K}$ et $x, y \in E$. Alors :

$$\lambda \cdot x + y = \lambda \cdot \sum_{i=1}^{r_1} \lambda_i \cdot a_i + \sum_{j=1}^{r_2} \mu_j \cdot a_j = \sum_{i=1}^{r_1} (\lambda \lambda_i) \cdot a_i + \sum_{j=1}^{r_2} \mu_j \cdot a_j.$$

Il est clair que cette expression peut se mettre sous la forme d'une combinaison linéaire

d'éléments de A :

$$\lambda.x + y = \sum_{i=1}^{r_3} \nu_i.a_i,$$

donc $\lambda.x + y \in Vect(A)$.

► Soit H un s.e.v tel que $A \subset H$. On a $Vect(A) \subset H$. En effet, soit $x \in Vect(A)$, alors x est une combinaison linéaire d'éléments de A , donc de H , et par stabilité de H , $x \in H$. Ainsi, tout s.e.v contenant A contient aussi $Vect(A)$, ce qui montre que $Vect(A)$ est le plus petit sous-espace vectoriel contenant A . ■

5.1.4 Bases, dimension

Proposition-Définition 5.1.8

Une famille $\mathcal{E} = (e_1, \dots, e_n)$ d'éléments de E est une **base** si elle est libre et génératrice, ce qui est équivalent à l'assertion suivante :

$$\forall x \in E, \exists! (x_1, \dots, x_n) \in \mathbb{K}^n \text{ tel que } x = \sum_{i=1}^n x_i.e_i. \quad (5.1)$$

Dit autrement : tout élément x se décompose de manière unique sur la base \mathcal{E} . Les scalaires $\{x_i\}_{i=1:n}$ sont appelés **coordonnées de x dans la base \mathcal{E}** .

Démonstration : On va montrer l'équivalence entre la définition d'une base et l'assertion (5.1).

(\Rightarrow) \mathcal{E} base de E , donc \mathcal{E} génératrice : tout élément x de E s'écrit comme combinaison linéaire de la famille \mathcal{E} : $x = \sum_{i=1}^n x_i.e_i$. Cette décomposition est unique. En effet, si $x =$

$\sum_{i=1}^n x'_i.e_i$, alors $0_E = \sum_{i=1}^n (x_i - x'_i).e_i$, ce qui entraîne $x_i = x'_i$ puisque \mathcal{E} est libre.

(\Leftarrow) Si $\forall x \in E, \exists! (x_1, \dots, x_n) \in \mathbb{K}^n$ tel que $x = \sum_{i=1}^n x_i.e_i$, il est évident que la famille \mathcal{E} est génératrice. Elle est également libre puisque $0_E = \sum_{i=1}^n 0.e_i$ est, par unicité, la seule combinaison linéaire qui soit nulle. ■

Le notion de base est essentielle en algèbre linéaire. Elle permet d'établir que tout élément d'un espace vectoriel est entièrement caractérisé par un nombre dénombrable (fini dans notre cas) de coefficients.

Exemple : $\{e_i\}_{i=1:n}$ avec $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$, le 1 étant à la i ème position, est une base de \mathbb{R}^n , appelée base canonique.

Remarque importante : Il convient de bien faire la distinction entre un **vecteur** et ses coordonnées, surtout dans \mathbb{R}^n . Un vecteur de \mathbb{R}^n est un n-uplet de nombres réels, il existe indépendamment d'une base (c'est simplement un ensemble de n nombres). Ses coordonnées sont des coefficients qui correspondent à sa décomposition sur une base donnée (elles dépendent donc

de la base). Ainsi, lorsqu'on écrit $(2, 1, 2)$ sans précision supplémentaire, il s'agit du vecteur $(2, 1, 2) \in \mathbb{R}^n$. La représentation d'un vecteur par ses coordonnées se faisant avec la même notation (n-uplet), il faudra impérativement associer une base à n-uplet de coordonnées (on remarque que l'écriture d'un vecteur coïncide avec ses coordonnées dans la base canonique).

Définition 5.1.9

Un espace vectoriel E est de **dimension finie** s'il existe une famille finie génératrice de E .

Théorème-Définition 5.1.10 (*admis*)

Tout espace de dimension finie admet au moins une base finie. Le nombre d'éléments d'une base de E , identique pour toutes les bases, est appelé **dimension** de E .

Exemple : $E = \mathbb{R}^2$. La famille $((1, 1), (1, 0))$ est une base de E (le montrer!). Ainsi, E est un e.v de dimension 2.

Les résultats qui suivent sont utiles en pratique.

Proposition 5.1.11 (*admise*)

Soit E un espace de dimension n finie. Alors

- Toute famille libre a au plus n éléments
- Toute famille génératrice a au moins n éléments.

Ainsi, toute famille libre de n éléments est une base. De même, toute famille génératrice de n éléments est une base.

- Soit F un s.e.v de E . Si $\dim(F) = \dim(E)$, alors $F = E$.

5.2 Applications linéaires

5.2.1 Rappels sur les applications

Soient X, Y deux ensembles, et $f : X \rightarrow Y$ une application. L'élément $y \in Y$ tel que $y = f(x)$ est l'**image** de x par f . Un élément $x \in X$ tel que $f(x) = y$ est un **antécédent** de y par f .

Définition 5.2.1

L'application f est dite :

- **injective** si $\forall x_1, x_2, f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ (unicité de l'antécédent s'il existe)
- **surjective** si $\forall y \in Y, \exists x \in X$ tel que $y = f(x)$ (existence d'au moins un antécédent)
- **bijective** si elle est injective et surjective (existence et unicité de l'antécédent)

Définition 5.2.2

Soient E, F, G trois espaces vectoriels, et soient les applications $f : E \rightarrow F, g : E \rightarrow F, h : F \rightarrow G$. On définit les opérations suivantes :

- la somme de f et $g : \forall x \in E, (f + g)(x) := f(x) + g(x)$
- la composition de h par $f : \forall x \in E, (g \circ f)(x) := h(f(x))$.

5.2.2 Applications linéaires

Définition 5.2.3

Soient E, F deux e.v. Une application $f : E \rightarrow F$ est **linéaire** (ou est un morphisme d'e.v) si :

- $\forall x, y \in E, f(x + y) = f(x) + f(y)$
- $\forall \lambda \in \mathbb{K}, \forall x \in E, f(\lambda.x) = \lambda.f(x)$

Si de plus $E = F$, on dit que f est un **endomorphisme**.

- Remarque :**
1. Les deux conditions de linéarité peuvent se résumer en une seule : $\forall x, y \in E, \forall \lambda \in \mathbb{K}, f(\lambda.x + y) = \lambda.f(x) + f(y)$
 2. Pour toute application linéaire f , on a $f(0_E) = 0_F$ car $f(0_E) = f(x - x) = f(x) - f(x) = 0_F$.

Attention : La notation \cdot est utilisé à la fois pour la loi externe sur E et sur F .

Notation : L'ensemble des applications linéaires (resp. des endomorphismes) de E dans F est noté $\mathcal{L}(E, F)$ (resp. $\mathcal{L}(E)$). Ce sont tous deux des espaces vectoriels.

- Exemple :**
1. $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$
 $(x, y, z) \mapsto (x + y, x - y) \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$
 2. $\delta : \mathcal{C}^0(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}$
 $f \mapsto f(0) \in \mathcal{L}(\mathcal{C}^0(\mathbb{R}, \mathbb{R}), \mathbb{R})$

Définition 5.2.4

Soient E, F deux e.v, et $f \in \mathcal{L}(E, F)$. On appelle **noyau de f** l'ensemble noté $\ker(f)$ défini par :

$$\ker(f) := f^{-1}(\{0_F\}) = \{x \in E ; f(x) = 0_F\} \subset E.$$

On appelle **image de f** l'ensemble noté $\text{Im } f$ défini par :

$$\text{Im } f := f(E) = \{y \in F ; \exists x \in E \text{ tel que } y = f(x)\} \subset F.$$

- Exemple :** Soit $f : (x, y) \in \mathbb{R}^2 \mapsto x + y \in \mathbb{R}$. On a $\ker(f) = \{(x, y) \in \mathbb{R}^2 / x + y = 0\} = \{(x, y) \in \mathbb{R}^2 / x = -y\} = \{\lambda.(-1, 1), \lambda \in \mathbb{R}\}$, donc $\ker(f) = \text{Vect}((-1, 1))$.
 $\text{Im } f = \mathbb{R}$ car $\forall z \in \mathbb{R}, \exists (x, y) \in \mathbb{R}^2$ tel que $x + y = z$ (il en existe une infinité en fait).

Proposition 5.2.5

Soient E, F deux e.v, et $f \in \mathcal{L}(E, F)$. On a les assertions suivantes :

1. $\ker(f)$ et $\text{Im } f$ sont respectivement des s.e.v de E et F
2. f injective $\Leftrightarrow \ker(f) = \{0_E\}$
3. f surjective $\Leftrightarrow \text{Im } f = F$.

Démonstration : 1. Le fait que $\ker(f)$ et $\text{Im } f$ sont des s.e.v est immédiat.

2. (\Rightarrow) f injective. Soit $x \in \ker(f)$. On a alors $f(x) = 0_F = f(0_E)$. L'injectivité de f entraîne $x = 0_E$, donc $\ker(f) = \{0_E\}$.
 (\Leftarrow) $\ker(f) = \{0_E\}$. On a $f(x_1) = f(x_2) \Leftrightarrow f(x_1) - f(x_2) = 0_F \Leftrightarrow f(x_1 - x_2) = 0_F$, et donc $x_1 - x_2 \in \ker(f)$, soit $x_1 = x_2$ puisque $\ker(f) = \{0_E\}$.
3. f surjective $\Leftrightarrow f(E) = F \Leftrightarrow \text{Im } f = F$. ■

Ce résultat est utile en pratique pour établir qu'un ensemble est un espace vectoriel : il suffit de l'exprimer comme étant le noyau d'une certaine application linéaire.

Exemple : 1. L'application $f : (x, y) \in \mathbb{R}^2 \mapsto x + y$ est surjective car $\text{Im } f = \mathbb{R}$, mais n'est pas injective car $\ker(f) \neq \{0_E\}$ (cf. exemple précédent)

2. L'ensemble $E_1 = (x, y, z) / x + 2y - z = 0$ est un espace vectoriel. En effet, soit $f : (x, y, z) \in \mathbb{R}^3 \mapsto x + 2y - z$. Alors, $E_1 = \ker(f)$, donc E est un s.e.v de \mathbb{R}^3 .

Le théorème qui suit est important dans la mesure où il relie la dimension de E et la dimension des noyau et image d'une application linéaire sur E ; il peut ainsi servir à établir certaines propriétés sur les applications (injective, bijective), les familles (liées, libres) etc.

Théorème 5.2.6 (Théorème du rang)

Soient E, F deux e.v de dimension finie, et $f \in \mathcal{L}(E, F)$. Alors

$$\dim(E) = \dim(\ker(f)) + \text{rg}(f),$$

où $\text{rg}(f) := \dim(\text{Im } f)$ est appelé **rang de f** .

Remarque : On montre facilement que le rang de f est égal au rang de la famille de vecteurs (de F) constituée des images des vecteurs de la base de E : $\text{rg}(f) = \text{rg}(f(e_1), \dots, f(e_n))$, soit encore le nombre de vecteurs libres parmi cette famille.

Exemple : Soit $f : (x, y) \in \mathbb{R}^2 \mapsto x + y \in \mathbb{R}$. Comme vu précédemment, $\ker(f) = \text{Vect}((-1, 1))$, donc $\dim(\ker(f))=1$. On en déduit d'après le théorème du rang que $\dim(\text{Im } f) = \text{rg}(f) = \dim(\mathbb{R}^2) - 1 = 1$. On en déduit que $\text{Im } f = \mathbb{R}$ (inclu et de même dimension) et donc que f est surjective.

5.3 Matrices

5.3.1 Des applications linéaires aux matrices

Soient E, F deux e.v, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $\mathcal{B}' = (e'_1, \dots, e'_m)$ base de F , et $f \in \mathcal{L}(E, F)$. Soit $x \in E$. Alors x admet une (unique) décomposition sur \mathcal{B} :

$$f(x) = f\left(\sum_{j=1}^n x_j \cdot e_j\right) = \sum_{j=1}^n x_j \cdot f(e_j) \quad (\text{avec } x_j \in \mathbb{K}).$$

Ainsi, on constate que la seule connaissance des images des vecteurs de base e_j par f suffit à caractériser l'application f . Comme $f(e_j) \in F$, on peut décomposer chaque $f(e_j)$ sur \mathcal{B}' :

$$f(x) = \sum_{j=1}^n x_j \cdot \sum_{i=1}^m a_{ij} \cdot e'_i.$$

L'application f est donc entièrement caractérisée par les coefficients $\{a_{ij}\}_{i=1:m, j=1:n}$ (scalaires de \mathbb{K}), que l'on regroupe dans un objet appelé **matrice** :

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Un intérêt des matrices est qu'elles permettent une manipulation simplifiée et intuitive des applications linéaires. Ainsi, les applications de $\mathcal{L}(E, F)$ seront avantageusement assimilées à leurs représentations matricielles. Le but des paragraphes qui suivent est de définir un ensemble de règles sur les matrices correspondant aux règles sur $\mathcal{L}(E, F)$.

5.3.2 Définitions et notations

Définition 5.3.1

Une **matrice** est un ensemble (fini) de coefficients $\{a_{ij}\}_{i=1:m, j=1:n}$ de \mathbb{K} , généralement représentés par un tableau :

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Si $m = 1$ (resp. $n = 1$) on parle de **matrice ligne** (resp. **colonne**). On utilise souvent les termes de **vecteur ligne** et **vecteur colonne**.

Si $m = n$, on dit que la matrice est **carrée**.

Notations : — On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices à éléments dans \mathbb{K} de m lignes et n colonnes, et plus simplement $\mathcal{M}_n(\mathbb{K})$ pour les matrices carrées de taille $n \times n$.

- Un élément x de E sera représenté par ses composantes dans la base \mathcal{B} considérée pour E , la matrice colonne correspondante étant notée $[x]_{\mathcal{B}}$ (ou x si aucune confusion n'est à craindre) :

$$[x]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (\text{avec } x = \sum_{j=1}^n x_j \cdot e_j)$$

- On note M_{ij} l'élément i, j de M ; pour une matrice colonne u , on note u_i la i -ème composante de u .

Définition 5.3.2

Soient E, F deux e.v, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $\mathcal{B}' = (e'_1, \dots, e'_m)$ une base de F . Une application $f \in \mathcal{L}(E, F)$ sera désormais représentée par sa matrice $(a_{ij})_{i=1:m, j=1:n}$, où a_{ij} est la i -ième composante de $f(e_j)$ dans la base \mathcal{B}' , soit :

$$[f]_{\mathcal{B}'}^{\mathcal{B}} := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \text{où } \forall j = 1 : n, f(e_j) = \sum_{i=1}^m a_{ij} \cdot e'_i$$

Dit autrement, chaque colonne de $[f]_{\mathcal{B}'}^{\mathcal{B}}$ est l'image de e_j par f exprimée dans la base \mathcal{B}' . La matrice $[f]_{\mathcal{B}'}^{\mathcal{B}}$ est appelée **matrice représentative de f relativement aux bases \mathcal{B} et \mathcal{B}'** .

Remarque (importante) : La matrice représentative d'une application linéaire **dépend des bases** considérées sur E et F !

Exemple : Soit $f : (x, y) \in \mathbb{R}^2 \mapsto (x + y, x - y, y) \in \mathbb{R}^3$. La matrice de f relativement aux bases canoniques de \mathbb{R}^2 et \mathbb{R}^3 est donnée par :

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix}$$

$f(e_1) \quad f(e_2)$

5.3.3 Structure de $\mathcal{M}_{m,n}(\mathbb{K})$

Proposition-Définition 5.3.3

On définit sur $\mathcal{M}_{m,n}(\mathbb{K})$ une loi interne $+$ et une loi externe \cdot qui lui confèrent la structure de \mathbb{K} -e.v :

$$+ : \mathcal{M}_{m,n}(\mathbb{K}) \times \mathcal{M}_{m,n}(\mathbb{K}) \rightarrow \mathcal{M}_{m,n}(\mathbb{K}) \quad \text{où } (A+B)_{ij} := a_{ij} + b_{ij}$$

$$(A, B) \mapsto A + B$$

$$\cdot : \mathbb{K} \times \mathcal{M}_{m,n}(\mathbb{K}) \rightarrow \mathcal{M}_{m,n}(\mathbb{K}) \quad \text{où } (\lambda.A)_{i,j} := \lambda a_{ij}$$

$$(\lambda, A) \mapsto \lambda.A$$

Démonstration : $(\mathcal{M}_{m,n}(\mathbb{K}), +)$ groupe commutatif :

► $\forall A, B, C \in \mathcal{M}_{m,n}(\mathbb{K}), A + B = B + A$ car $a_{ij} + b_{ij} = b_{ij} + a_{ij}$; pour les mêmes raisons, $(A + B) + C = A + (B + C)$

► neutre : la matrice nulle $0_{m,n}$ est neutre pour $+$

► symétrique : $\forall A \in \mathcal{M}_{m,n}(\mathbb{K})$, il existe un symétrique $(-A)$ défini par $(-A)_{ij} = -a_{ij}$.

De plus, pour tout $\lambda, \mu \in \mathbb{K}$ et $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$, la loi \cdot vérifie :

► $\lambda.(A+B) = \lambda.A + \lambda.B$ car $(\lambda.(A+B))_{ij} = \lambda(A+B)_{ij} = \lambda(a_{ij} + b_{ij}) = \lambda a_{ij} + \lambda b_{ij} = (\lambda.A)_{ij} + (\lambda.B)_{ij}$

► $(\lambda+\mu).A = \lambda.A + \mu.A$ car $((\lambda+\mu).A)_{ij} = (\lambda+\mu)a_{ij} = \lambda a_{ij} + \mu a_{ij} = (\lambda.A)_{ij} + (\mu.A)_{ij} = (\lambda.A + \mu.A)_{ij}$ ■

Ces lois sur les matrices correspondent aux opérations que l'on peut effectuer sur les applications de $\mathcal{L}(E, F)$. Ainsi, la somme de deux applications f, g de $\mathcal{L}(E, F)$ s'effectue en additionnant leurs matrices respectives : $[f + g]_{\mathcal{B}}^{\mathcal{B}'} = [f]_{\mathcal{B}}^{\mathcal{B}'} + [g]_{\mathcal{B}}^{\mathcal{B}'}$. De même, la multiplication par un scalaire λ se traduit par une multiplication de la matrice par λ : $[\lambda.f]_{\mathcal{B}}^{\mathcal{B}'} = \lambda.[f]_{\mathcal{B}}^{\mathcal{B}'}$.

On définit ci-après une troisième loi sur les matrices, à savoir un produit interne entre matrices.

Définition 5.3.4

(multiplication de matrices) Soient $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$. On définit la matrice produit de A par B , noté $A \times B$ ou plus simplement AB , comme suit :

$$(AB)_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

La matrice résultante est de taille $m \times p$.

Remarque : 1. Ce produit n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B .

2. Ce produit n'est en général **pas** commutatif.

Exemple : Soient $A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -3 \\ 0 & 1 \\ -1 & 0 \end{pmatrix}$. Alors $AB = \begin{pmatrix} 3 & -3 \\ 1 & -5 \end{pmatrix}$. On re-

marque que le produit BA est quant à lui impossible à cause des dimensions qui ne concordent pas (nombre de colonnes de B différents du nombre de lignes de A).

De même que précédemment, il s'avère que cette loi \times sur les matrices correspond à une opération sur les applications qu'elles représentent. C'est l'objet de la proposition suivante.

Proposition 5.3.5

Soient E, F, G trois espaces vectoriels de dimensions m, n, p finies, et de bases respectives $\mathcal{E} = (e_1, \dots, e_m)$, $\mathcal{F} = (e'_1, \dots, e'_n)$, $\mathcal{G} = (e''_1, \dots, e''_p)$. Soient enfin $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. La composition de g par f se traduit par un produit de leurs matrices :

$$[g \circ f]_{\mathcal{E}}^{\mathcal{G}} = [g]_{\mathcal{F}}^{\mathcal{G}} [f]_{\mathcal{E}}^{\mathcal{F}}.$$

Démonstration : On note $A = [f]_{\mathcal{E}}^{\mathcal{F}}$, $B = [g]_{\mathcal{F}}^{\mathcal{G}}$ et $C = [g \circ f]_{\mathcal{E}}^{\mathcal{G}}$. On va montrer que $C = BA$.

Par définition de C , on a pour tout $j = 1 : m$,

$$(g \circ f)(e_j) = \sum_{i=1}^p c_{ij} e''_i.$$

Par ailleurs :

$$\begin{aligned} (g \circ f)(e_j) &= g(f(e_j)) = g\left(\sum_{k=1}^n a_{kj} e'_k\right) \text{ par définition de } A \\ &= \sum_{k=1}^n a_{kj} g(e'_k) = \sum_{k=1}^n a_{kj} \sum_{i=1}^p b_{ik} e''_i = \sum_{i=1}^p \left(\sum_{k=1}^n b_{ik} a_{kj}\right) e''_i, \end{aligned}$$

et, puisque la décomposition sur une base est unique :

$$\forall i = 1 : m, j = 1 : p, c_{ij} = \sum_{k=1}^n b_{ik} a_{kj} = (BA)_{ij},$$

d'où $C = BA$. ■

Enfin, l'objet de la proposition qui suit est d'établir que le résultat d'une application linéaire f appliquée sur un élément x (i.e $f(x)$) n'est autre que le produit (matriciel) entre la matrice représentative de f et la matrice colonne des coordonnées de x dans la base considérée.

Proposition 5.3.6

Soient E, F deux espaces vectoriels de dimensions m, n finies, et de base respective $\mathcal{B} = (e_1, \dots, e_m)$, $\mathcal{B}' = (e'_1, \dots, e'_n)$. Soit $f \in \mathcal{L}(E, F)$. Alors, pour tout $x \in E$, les composantes de $f(x)$ dans la base \mathcal{B}' s'expriment :

$$[f(x)]_{\mathcal{B}'} = [f]_{\mathcal{B}}^{\mathcal{B}'} [x]_{\mathcal{B}}.$$

Démonstration : On note $A = [f]_{\mathcal{B}}^{\mathcal{B}'}$, $X = [x]_{\mathcal{B}}$, $Y = [f(x)]_{\mathcal{B}'}$. Pour tout $x \in E$, on a par définition de Y :

$$f(x) = \sum_{i=1}^n y_i e'_i.$$

Par ailleurs :

$$\begin{aligned} f(x) &= f\left(\sum_{j=1}^m x_j e_j\right) = \sum_{j=1}^m x_j f(e_j) = \sum_{j=1}^m x_j \sum_{i=1}^n a_{ij} e'_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij} x_j\right) e'_i = \sum_{i=1}^n (AX)_i e'_i \end{aligned}$$

d'où :

$$\forall i = 1 : n, y_i = (AX)_i.$$

Ainsi, $Y = AX$. ■

On peut constater que cette représentation matricielle permet de visualiser le caractère linéaire de la fonction f , qui s'écrit comme le produit (matriciel) entre une constante (matricielle) A et la "variable" x , généralisant ainsi la représentation classique de \mathbb{R} dans \mathbb{R} : $x \mapsto ax$.

5.3.4 Propriétés et opérations sur les matrices

Proposition-Définition 5.3.7 (*rang*)

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$. On appelle **rang de A** le rang de la famille constituée des colonnes de A , qui est aussi égal au rang des lignes de A .

Notons que les notions de noyau et d'image sont les mêmes que pour les applications linéaires :

$$\begin{aligned} \ker(A) &= \{X \in \mathbb{K}^n / AX = 0_{\mathbb{K}^m}\} \\ \text{Im } A &= \{AX, X \in \mathbb{K}^n\}. \end{aligned}$$

En particulier, le rang d'une matrice est égal au rang de l'application linéaire qu'elle représente :

$$\text{rg}(A) = \text{rg}(f).$$

Exemple : Soit la matrice A suivante :

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 2 \\ 0 & 2 & -2 \\ 1 & 1 & 0 \end{pmatrix}.$$

Alors $\text{rg}(A) = 2$. En effet, en notant C_1, C_2, C_3 les colonnes de A , on a $\text{rg}(A) = \text{rg}((C_1, C_2, C_3))$. Or (C_1, C_2, C_3) est liée car $C_3 = C_1 - C_2$, donc $\text{rg}(A) \leq 2$. En revanche, (C_1, C_2) est libre, donc $\text{rg}(A) = 2$, et la matrice (ou l'application qu'elle représente) n'est pas surjective. De plus, le théorème du rang indique que $\dim(\ker(A)) = 3 - 2 = 1$, l'application n'est donc pas injective car $\ker(A) \neq \{0_{\mathbb{R}^3}\}$. Cherchons une base de $\ker(A)$.

$$X = (x, y, z) \in \ker(A) \iff AX = 0_{\mathbb{R}^4} \iff \begin{cases} x + z = 0 \\ x - y + 2z = 0 \\ 2y - 2z = 0 \\ x + y = 0. \end{cases} \iff \begin{cases} y = -x \\ z = -x \end{cases}$$

Donc $X = (x, -x, -x) = x(1, -1, -1)$. Une base de $\ker(A)$ est $(1, -1, -1)$.

Définition 5.3.8 (transposée)

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$. On appelle **matrice transposée de A** la matrice notée $A^T \in \mathcal{M}_{n,m}(\mathbb{K})$ (parfois notée tA) définie par :

$$\forall i = 1 : n, \forall j = 1 : m, (A^T)_{ij} = A_{ji}.$$

Si $A = A^T$, on dit que A est **symétrique**.

Dit autrement : les lignes de A^T sont constituées des colonnes de A .

Exemple : Soit $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}$, alors $A^T = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$.

Définition 5.3.9

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

— On dit que A est **inversible** s'il existe une matrice $A' \in \mathcal{M}_n(\mathbb{K})$ telle que :

$$AA' = A'A = I_n.$$

Si A est inversible, on note A^{-1} son inverse.

— On appelle **trace de A** le scalaire :

$$\text{tr}(A) := \sum_{i=1}^n a_{ii} \quad (\text{somme des éléments diagonaux})$$

Exemple : Soit $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$. Alors $\text{tr}(A) = 0 + 2 = 2$. De plus, A est inversible. En effet, soit

$$A' = \begin{pmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{pmatrix}. \text{ Alors :}$$

$$AA' = I_2 \Leftrightarrow \begin{cases} a'_{21} = 1 \\ a'_{22} = 0 \\ a'_{11} + 2a'_{21} = 0 \\ a'_{12} + 2a'_{22} = 1 \end{cases} \Leftrightarrow A' = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}.$$

On a donc bien $A'A = AA' = I_2$ donc $A^{-1} = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}$.

On donne dans la proposition qui suit quelques propriétés utiles lors des calculs.

Proposition 5.3.10

1. $\forall A \in \mathcal{M}_{m,n}(\mathbb{K}), (A^T)^T = A$
2. $\forall A \in \mathcal{M}_{m,n}(\mathbb{K}), \forall B \in \mathcal{M}_{n,p}(\mathbb{K}), (AB)^T = B^T A^T$
3. $\forall A, B \in \mathcal{M}_{m,n}(\mathbb{K}), \text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$
4. $\forall A \in \mathcal{M}_{m,n}(\mathbb{K}), \forall B \in \mathcal{M}_{n,m}(\mathbb{K}), \text{tr}(AB) = \text{tr}(BA)$
5. Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, on a A^T inversible et $(A^T)^{-1} = (A^{-1})^T$
6. Si $A, B \in \mathcal{M}_n(\mathbb{K})$ inversibles, on a AB inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

Démonstration : 1. évident

$$2. ((AB)^T)_{ij} = (AB)_{ji} = \sum_k a_{jk} b_{ki} = \sum_k b_{ki} a_{jk} = (B^T A^T)_{ij}$$

3. évident

$$4. \text{tr}(AB) = \sum_i (AB)_{ii} = \sum_i \sum_j a_{ij} b_{ji} = \sum_j \sum_i b_{ji} a_{ij} = \sum_j (BA)_{jj} = \text{tr}(BA)$$

5. Il est évident que A' est inversible. De plus, $A^T (A^{-1})^T = (A^{-1}A)^T$ d'après 2, d'où $A^T (A^{-1})^T = I_n$, de même pour $(A^{-1})^T A^T$.

6. $AB (AB)^{-1} = ABB^{-1}A^{-1} = AI_n A^{-1} = I_n$, de même pour $(AB)^{-1} AB$, d'où le résultat. ■

Définition 5.3.11 (déterminant)

Soit $A \in \mathcal{M}_2(\mathbb{K})$. On définit le déterminant de la matrice A , noté $|A|$, par :

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Pour $A \in \mathcal{M}_3(\mathbb{K})$, on définit alors :

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Enfin, pour $A \in \mathcal{M}_n(\mathbb{K})$, on définit le déterminant développé selon la ligne i :

$$|A| = \sum_{j=1}^n (-1)^{i+j} a_{ij} |M_{ij}|$$

où M_{ij} est la (sous-)matrice A à laquelle on a "enlevé" la ligne i et la colonne j (les déterminants $|M_{ij}|$ s'appellent les **mineurs** de A).

Remarque importante : Cette formule est valable quelle que soit la ligne de développement i choisie. De plus, le développement du déterminant peut se faire, de la même manière, selon une colonne j en inversant les rôles de i et j , le résultat étant évidemment le même.

Exemple :
$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 5 \\ 2 & 1 & 1 \end{vmatrix} = (-1)^{1+1} 1 \begin{vmatrix} 0 & 5 \\ 1 & 1 \end{vmatrix} + (-1)^{1+2} 2 \begin{vmatrix} 0 & 5 \\ 2 & 1 \end{vmatrix} + (-1)^{1+3} 3 \begin{vmatrix} 0 & 0 \\ 2 & 1 \end{vmatrix} = -15.$$

(développement par rapport à la première ligne). Il est bien sûr plus judicieux de faire le développement selon la seconde ligne : $|A| = (-1)^{2+3} 5 \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -15.$

On a également les propriétés suivantes, intéressantes pour les calculs de déterminants, utiles notamment pour la réduction d'endomorphismes.

Proposition 5.3.12 (Propriétés du déterminant)

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

1. Permuter deux lignes ou deux colonnes change le signe du déterminant.
2. Si une ligne ou une colonne de A possède un facteur commun α , ce nombre peut-être mis en facteur du déterminant.
3. Toute opération de la forme :

$$L_i \leftarrow L_i + \sum_{\substack{k=1 \\ k \neq i}}^n L_k$$

laisse le déterminant inchangé. Cette dernière propriété est essentielle pour faire apparaître des 0 dans le déterminant (méthode de pivot), simplifiant ainsi notablement les calculs, en particuliers pour les calculs de valeurs propres.

Exemples :

1.
$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = - \begin{vmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{vmatrix}.$$
2.
$$\begin{vmatrix} 1 & 2 & 3 \\ -2 & 8 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 & 3 \\ -1 & 4 & 3 \\ 7 & 8 & 9 \end{vmatrix}$$
3.
$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & -1 & -2 \\ 2 & 1 & 1 \end{vmatrix} \stackrel{L_3 \leftarrow L_3 - L_2}{=} \begin{vmatrix} 1 & 2 & 1 \\ 2 & -1 & -2 \\ 0 & 2 & 3 \end{vmatrix} \stackrel{L_2 \leftarrow L_2 - 2L_1}{=} \begin{vmatrix} 1 & 2 & 1 \\ 0 & -5 & -4 \\ 0 & 2 & 3 \end{vmatrix} = -7.$$

5.3.5 Matrices de changement de base

Comme on l'a vu, l'écriture des vecteurs et matrices dépendent des bases dans lesquelles ils sont exprimés. En pratique, on est souvent amené à se placer dans des bases non canoniques pour simplifier certains problèmes. Les matrices de passage permettent de ramener ces opérations de changement de base à une multiplication matricielle.

Proposition-Définition 5.3.13

Soient \mathcal{B} et \mathcal{B}' deux bases d'un e.v E . On appelle **matrice de passage de la base \mathcal{B} à la base \mathcal{B}'** la matrice dont les colonnes sont les vecteurs de la base \mathcal{B}' exprimés dans la base \mathcal{B} :

$$P_{\mathcal{B},\mathcal{B}'} = [e'_1 | \dots | e'_n]_{\mathcal{B}}.$$

Ainsi, pour exprimer dans la base \mathcal{B} un vecteur $x \in E$ exprimé en base \mathcal{B}' , il suffit d'effectuer le produit :

$$[x]_{\mathcal{B}} = P_{\mathcal{B},\mathcal{B}'} [x]_{\mathcal{B}'}$$

On a de plus la propriété :

$$P_{\mathcal{B}',\mathcal{B}} = (P_{\mathcal{B},\mathcal{B}'})^{-1}.$$

De même, pour les applications linéaires, on peut envisager d'effectuer un changement de base sur E et/ou F .

Proposition 5.3.14

Soient $f \in \mathcal{L}(E, F)$, $\mathcal{E}, \mathcal{E}'$ deux bases de E et $\mathcal{F}, \mathcal{F}'$ deux bases de F . On note $A = [f]_{\mathcal{E}}^{\mathcal{F}}$, $A' = [f]_{\mathcal{E}'}^{\mathcal{F}'}$, P la matrice de passage de la base \mathcal{E} à \mathcal{E}' et Q la matrice de passage de la base \mathcal{F} à \mathcal{F}' . On a alors la relation suivante :

$$A' = Q^{-1}AP. \tag{5.2}$$

Remarques : 1. En d'autres termes : la matrice représentative de f dans les nouvelles bases $\mathcal{E}', \mathcal{F}'$ s'obtient en pré et post multipliant la matrice de f dans les anciennes bases par des matrices de passage. Bien entendu, la relation (5.2) peut-être écrite de diverses manières ;

l'essentiel est d'en retenir une et de s'y tenir, notamment concernant l'expression des matrices de changement de base en jeu qui sont souvent source d'erreur et de confusion.

2. Ce genre d'opération est très utilisé pour la réduction d'endomorphismes (diagonalisation et triangularisation de matrices, etc.).

Exemple : Soit $f \in \mathcal{L}(\mathbb{R}^2)$ (endomorphisme) et \mathcal{E} la base canonique de \mathbb{R}^2 , avec $[f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 13 & -5 \\ 11 & -3 \end{pmatrix}$.

Soient deux vecteurs $u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ de E . On note $\mathcal{B} = (u_1, u_2)$ une nouvelle base de E . Exprimons la matrice de f dans la nouvelle base \mathcal{B} . Calculons la matrice de passage :

$$P_{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{matrix} e_1 \\ e_2 \end{matrix} \begin{matrix} u_1 \\ u_2 \end{matrix}.$$

On doit maintenant calculer son inverse. Pour cela, on exprime la relation entre les vecteurs u_1, u_2 et la base canonique e_1, e_2 :

$$\begin{cases} u_1 = e_1 + e_2 \\ u_2 = -e_1 + e_2, \end{cases}$$

et on inverse cette relation en résolvant le système pour exprimer e_1, e_2 en fonction de u_1, u_2 :

$$\begin{cases} e_1 = u_1 - e_2 \\ e_2 = u_2 + e_1 = u_2 + u_1 - e_2 \end{cases} \iff \begin{cases} e_1 = \frac{1}{2}(u_1 - u_2) \\ e_2 = \frac{1}{2}(u_2 + u_1). \end{cases}$$

Ainsi :

$$P_{\mathcal{E},\mathcal{B}}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{matrix} u_1 \\ u_2 \end{matrix} \begin{matrix} e_1 \\ e_2 \end{matrix}, \text{ et } [f]_{\mathcal{B}}^{\mathcal{B}} = P_{\mathcal{E},\mathcal{B}}^{-1} [f]_{\mathcal{E}}^{\mathcal{E}} P_{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 8 & -16 \\ 0 & 2 \end{pmatrix}$$

On peut constater que ce changement de base a "transformé" la matrice en une matrice dite triangulaire (qui n'a que des 0 en dessous de la diagonale). Cette opération est un cas particulier de réduction d'endomorphisme.

5.4 Réduction d'endomorphismes

Rappel : Un endomorphisme $f \in \mathcal{L}(E)$ est une application linéaire de E dans E . La matrice représentative d'un endomorphisme est donc une matrice carrée de taille n .

Dans cette section on s'intéressera principalement à introduire les notions clés à la **diagonalisation d'endomorphismes**, c'est-à-dire à la diagonalisation de leur matrice représentative (lorsque cela est possible) par changement de bases. D'autres réductions moins exigeantes existent mais ne sont pas présentées ici car plus techniques ; cependant, elles sont basées sur les mêmes notions, et le lecteur désireux d'en apprendre plus pourra se référer à n'importe quel ouvrage d'algèbre linéaire pour en savoir plus.

5.4.1 Valeurs propres, vecteurs propres et sous-espaces propres

La notion de valeur propre (ou vecteur propre, ou direction propre) est essentielle, on la rencontre dans toutes les disciplines scientifiques sous une forme ou une autre. Bien qu'appliquée aux matrices ici, la terminologie s'applique bien entendu de la même manière aux endomorphismes que ces matrices représentent.

Définition 5.4.1

On appelle **valeur propre - vecteur propre** tout couple $(\lambda, x) \in \mathbb{K} \times E \setminus \{0_E\}$ tel que :

$$Ax = \lambda x.$$

L'ensemble des valeurs propres d'une matrice est appelé **spectre** de la matrice et est noté $Sp(A)$.

Remarque : Selon la nature du vecteur propre (vecteur de \mathbb{R}^n , fonction etc.), on pourra parler de direction propre ou encore fonction propre.

Soit λ une valeur propre de A ; si x est un vecteur propre associé à λ , il est solution de :

$$Ax = \lambda x \Leftrightarrow (A - \lambda I_n)x = 0_E \Leftrightarrow x \in \ker(A - \lambda I_n).$$

Déterminer l'ensemble des vecteurs propres associés à une valeur propre revient donc à déterminer $\ker(A - \lambda I_n)$ (c'est-à-dire en déterminer une base).

Définition 5.4.2

On appelle **sous-espace propre** associé à la valeur propre λ le s.e.v noté E_λ :

$$E_\lambda := \ker(A - \lambda I_n).$$

Ces notions sont à la base du processus de diagonalisation (ou autre réduction en général) d'une matrice.

Exemple : Soit la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$. Sachant que 1 est valeur propre de cette matrice, déterminons le sous-espace propre E_1 :

$$\begin{aligned} v \in E_1 &\Leftrightarrow Av = 1.v \Leftrightarrow (A - 1I_n)v = 0_{\mathbb{R}^2} \Leftrightarrow \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Leftrightarrow \begin{cases} 0 = 0 \\ v_1 + v_2 = 0 \end{cases} \Leftrightarrow v_2 = -v_1. \end{aligned}$$

Ainsi, $v \in E_1 \Leftrightarrow v = \begin{pmatrix} v_1 \\ -v_1 \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. On a donc déterminé une base de E_1 , à savoir le vecteur $(1, -1)$, E_1 est de dimension 1. On pourrait en faire de même avec la valeur propre 2 et déterminer E_2 (on trouve $(0, 1)$ comme vecteur de base).

5.4.2 Diagonalisation de matrices

La diagonalisation d'une matrice consiste à déterminer un changement de base telle que la matrice soit diagonale dans cette base. Lorsqu'elle est possible, cette opération est intéressante car elle simplifie considérablement les calculs faisant intervenir la matrice en question. Si cela n'est pas possible, d'autres réduction existent (triangularisation, réduction de Jordan etc.).

Les grandes étapes d'un processus de diagonalisation sont :

1. Détermination des valeurs propres de la matrice

La plupart du temps, on utilise le théorème 5.4.3.

2. Détermination des sous-espaces propres associés à chaque valeur propre

Pour chaque valeur propre λ , on détermine une base du sous-espace E_λ en posant $(A - \lambda I_n) x = 0_E$.

3. La matrice est-elle diagonalisable ?

Après chaque calcul de sous-espace propre de l'étape précédente, on se réfère le théorème 5.4.4 pour savoir celui-ci est mis en défaut ou pas en s'assurant que la dimension de l'espace propre est bien égale à la multiplicité de la valeur propre associée. Si à l'issue du processus il ne l'a jamais été, alors la matrice est diagonalisable.

4. Diagonalisation de A .

Si A est diagonalisable, alors on sait que la matrice est diagonale lorsqu'elle est exprimée dans la base constituée de ses vecteurs propres, ou, plus précisément, constituée des vecteurs de base des différents sous-espaces propres E_{λ_i} (on peut montrer qu'ils forment une base de E). On note v_1, \dots, v_n ces vecteurs de base, P la matrice de passage de la base canonique à la base des vecteurs propres $P = [v_1 | \dots | v_n]$, et $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Alors, on a :

$$D = P^{-1}AP, \text{ ou encore } A = PDP^{-1}.$$

Ainsi, l'application linéaire représentée par A dans la base canonique est représentée par une matrice diagonale dans la base des vecteurs propres. En pratique comme en théorie, la notion de diagonalisation de matrice (et de manière générale de réduction d'endomorphisme) est très importante, car elle permet de transformer un problème faisant intervenir la matrice A en un problème simplifié équivalent faisant intervenir la matrice diagonale D , en utilisant la relation ci-dessus. Du fait de sa structure intéressante, on pourra résoudre des problèmes de manière plus aisée, et "repasser" à la solution du problème original en utilisant la matrice P .

Voici les théorèmes fondamentaux utilisés dans les précédents points.

Théorème-Définition 5.4.3

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On appelle **polynôme caractéristique** de A , noté p_A , le polynôme défini par :

$$p_A(\lambda) = \det(A - \lambda I_n).$$

On a alors le résultat fondamental suivant : les valeurs propres de la matrice A sont données par les racines du polynôme caractéristique (c'est-à-dire qu'elles sont solution de $p_A(\lambda) = 0$).

Exemple : Soit la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$. Dans un précédent exemple, il a été affirmé que 1 et 2 sont valeurs propres. On va ici le montrer en calculant son polynôme caractéristique :

$$p_A(\lambda) = \det(A - \lambda I_n) = \begin{vmatrix} 1 - \lambda & 0 \\ 1 & 2 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda).$$

Les valeurs propres de A étant les racines de p_A , on en déduit immédiatement que celles-ci sont 1 et 2.

Théorème 5.4.4

Une matrice A est diagonalisable si et seulement si :

1. Son polynôme caractéristique est sous la forme : $p_A(\lambda) = K \prod_{i=1}^r (\lambda - \lambda_i)^{m_i}$, c'est-à-dire qu'il est le produit de polynômes de degré un élevés à une puissance entière m_i (on dit qu'il est **scindé**).
2. La **multiplicité** m_i de la valeur propre λ_i est égale à la dimension du sous-espace propre associé E_{λ_i} .

En particulier, si $p_A(\lambda)$ n'admet que des racines simples, la matrice est diagonalisable.

Exemple : $p_A(\lambda) = (\lambda - 1)^2(\lambda + 5)$ est scindé ; $p_A(\lambda) = (\lambda^2 + 3\lambda - 1)(\lambda + 5)$ n'est pas scindé dans \mathbb{R} (mais il l'est dans \mathbb{C} !) : on voit ici l'importance de ne pas perdre de vue le corps \mathbb{K} sur lequel on travaille.

Exemple : Soit la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$. On déroule les différentes étapes :

1. Déjà fait lors d'un exemple précédent : les valeurs propres sont 1 et 2. Ces valeurs propres sont de multiplicité 1 (car $p_A(\lambda) = (1 - \lambda)^1 (2 - \lambda)^1$).
2. Les sous-espaces E_1 et E_2 ont été déterminés dans un précédent exemple : $E_1 = \text{Vect}((1, -1))$ et $E_2 = \text{Vect}((0, 1))$.
3. La dimension des sous-espaces propres est égale à la multiplicité des valeurs propres (égales à 1), le polynôme caractéristique est scindé dans \mathbb{R} , donc la matrice A est diagonalisable.
4. La matrice de passage s'exprime $P = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, et on a :

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = P^{-1}AP.$$

Chapitre 6

Probabilités et statistiques

cf. cours de C. Jauberthie