Harmonisation des Connaissances en Mathématiques

U.E. de 3me année
de la Licence Physique et Ingénieries,
parcours IUP GSI

et

1re année de Master PTR

Polycopié du cours
2005

François Dumas

PLAN ET PROGRESSION

1. Première série: Révision et mise au point sur certaines notions élémentaires

1. Nombres complexes

1. Forme algébrique: partie réelle, partie imaginaire, opérations sur les complexes, module, conjugaison, interprétation géométrique. – 2. Forme trigonométrique, forme exponentielle: argument, notation $e^{i\theta}$, formule de Moivre, formule d'Euler, applications, exponentielle complexe.

2. Equations et systèmes d'équations

1. Equations algébriques à une inconnue dans $\mathbb{R}$ et dans $\mathbb{C}$: racines n-ièmes dans $\mathbb{C}$, méthode générale de résolution des équations de degré 2, solutions réelles, solutions complexes. – 2. Equations linéaires à 2 ou 3 inconnues: systèmes linéaires, méthodes de résolution, pivot, déterminant 2×2, interprétation géométrique.

3. Développements limités

1. Notion de développement limité: définition, formule de Taylor-Young, réduction au cas des développements limités au voisinage de 0, développements limités généralisés. – 2. Méthodes de calcul: somme, produit, division, substitution, dérivation et primitivisation, exemples.

4. Intégrales et primitives

1. Intégrale d'une fonction continue par morceaux sur un intervalle fermé borné: principe de construction et interprétation géométrique, principales propriétés, liens avec les primitives, cas des fonctions à valeurs complexes. 2. Méthodes de calcul: primitives usuelles, changement de variables, intégration par parties, exemples.

5. Equations différentielles linéaires

1. Equations différentielles linéaires du premier ordre: solution générale de l'équation homogène associée, solutions particulières de l'équation avec second membre, variation de la constante, exemples. – 2. Equations différentielles linéaires du deuxième ordre à coefficients constants: solution générale de l'équation homogène associée, équation caractéristique, exemples de recherche de solutions particulières de l'équation avec second membre.

2. Deuxième série: Rappels et compléments d'analyse

6. Intégrales généralisées

1. Notion d'intégrale généralisée: divers types de problèmes, intégrale convergente ou divergente, méthodes de calculs, exemple de Riemann. – 2. Critères de convergence: critères de majoration et d'équivalence pour les fonctions réelles positives, exemple de la fonction eulérienne Γ, convergence absolue, exemples divers.

7. Séries numériques

1. Notion de série: sommes partielles, terme général, série convergente ou divergente, exemples des séries géométriques. – 2. Cas des séries à termes réels positifs: critères de majoration et d'équivalence, exemple des séries de Riemann, règles de d'Alembert et de Cauchy, exemples. – 3. Cas des séries à termes réels ou complexes quelconques: convergence absolue, séries réelles alternées, exemples.

8. Séries entières

1. Notion de série entière: rayon de convergence, disque de convergence, méthodes de calculs, exemples, série exponentielle, série entière géométrique. – 2. Fonctions d'une variable réelle définie par la somme d'une série entière: intervalle de convergence, continuité et dérivabilité, application à certaines équations différentielles, exemples. – 3. Développement en série entière d'une fonction d'une variable réelle: règles de calculs et exemples classiques.

9. Séries de Fourier

1. Coefficients de Fourier: forme réelle, forme complexe, série de Fourier, exemples. – 2. Convergences de la série de Fourier: théorème de Dirichlet, formule de Plancherel-Parseval, exemples et applications.

10. Fonctions de plusieurs variables

1. Fonctions scalaires de deux variables: continuité, dérivées partielles, formules de compositions, cas des coordonnées polaires, dérivées partielles d'ordre supérieur, extension aux fonctions scalaires de trois variables, coordonnées cylindriques, coordonnées sphériques. – 2. Fonctions vectorielles de deux ou trois variables: fonctions coordonnées, matrice jacobienne, exemples et applications, opérateurs classiques de l'analyse vectorielle.

.../...

3. Troisième série: Rappels et compléments d'algèbre linéaire, et applications

11. Vecteurs

1. Sous-espaces vectoriels de $\mathbb{R}^n$: opérations sur les vecteurs, combinaisons linéaires, notion de sous-espace vectoriel, intersection, somme directe et sous-espaces supplémentaires. – 2. Familles génératrices, familles libres, bases et dimension: bases et somme directe, théorème de la base incomplète.

12. Matrices

1. Calcul matriciel: notion de matrice, somme, produit, cas des matrices carrées, inversibilité, matrices semblables, matrices diagonales ou triangulaires, trace. – 2. Matrices et applications linéaires: notion d'endomorphisme, matrices d'un endomorphisme, application au calcul de l'inverse d'une matrice carrée inversible, noyau d'un endomorphisme.

13. Changements de bases

1. Les principes du changement de bases: matrice d'une famille finie de vecteurs dans une base, matrice de passage, première formule de changement de bases, seconde formule de changement de bases. – 2. Exemples d'applications de changements de bases: calcul des puissances n-ièmes d'une matrice et applications à des problèmes de suites, résolution de systèmes différentiels linéaires à coefficients constants, exemples.

14. Déterminants

1. Définition et propriétés du déterminant: développements, opérations sur les lignes, transposition et opérations sur les colonnes, exemples, multiplicativité, inversibilité. 2 – Applications des déterminants: calcul de l'inverse d'une matrice carrée, déterminant d'un endomorphisme, caractérisation des bases, formules de Cramer pour les systèmes d'équations linéaires.

15. Diagonalisation

1. Valeurs propres et polynôme caractéristique: valeurs propres, vecteurs propres, sous-espaces propres, polynôme caractéristique, multiplicité algébrique et multiplicité géométrique d'une valeur propre, exemples. – 2. Diagonalisation: notion d'endomorphisme et de matrice diagonalisables, conditions de diagonalisabilité, exemples, application à la résolution des systèmes différentiels linéaires, exemples, quelques mots sur la trigonalisation.

Leçon 1

Nombres complexes

1. Forme algébrique

1.1 Forme algébrique d'un nombre complexe.

On note $\mathbb{C}$ l'ensemble des nombres complexes. Tout nombre complexe z s'écrit de façon unique:

$$\boxed{z = x + iy \quad \text{avec } x, y \in \mathbb{R}.}$$

Le réel x s'appelle la partie réelle de z; le réel y s'appelle la partie imaginaire de z; on note parfois:

$$\boxed{x = \Re(z), \quad y = \Im(z).}$$

Deux nombres complexes sont égaux s'ils ont même partie réelle et même partie imaginaire:

$$\boxed{\text{si } x, x', y, y' \in \mathbb{R}, \quad x + iy = x' + iy' \ \Leftrightarrow\ x = x' \text{ et } y = y'}$$

Un nombre complexe est réel si et seulement si sa partie imaginaire est nulle. Un nombre complexe dont la partie réelle est nulle s'appelle un imaginaire pur; leur ensemble est noté $i\mathbb{R}$.

$$\boxed{\text{si } \ x, x', y, y' \in \mathbb{R}, \quad alors \quad [x + iy \in \mathbb{R} \ \Leftrightarrow\ y = 0], \quad \text{et} \quad [x + iy \in i\mathbb{R} \ \Leftrightarrow\ x = 0].}$$

1.2 Opérations sur les nombres complexes.

• Addition: la somme de deux nombres complexes s'obtient en additionnant leurs parties réelles et leurs parties imaginaires:

$$\boxed{(x + iy) + (x' + iy') = (x + x') + i(y + y').}$$

Propriétés. Cette addition a toutes les bonnes propriétés usuelles: commutativité $[z + z' = z' + z]$, associativité $[z + (z' + z'') = (z + z') + z'']$. Le nombre complexe $0 = 0 + i0$ est neutre pour cette addition $[0 + z = z + 0 = z$ pour tout $z \in \mathbb{C}]$. L'opposé d'un nombre complexe $z \in \mathbb{C}$ est $-z = (-x) + i(-y)$; c'est le nombre complexe tel que $z + (-z) = (-z) + z = 0$.

• Multiplication: le produit de deux nombres complexes est défini de façon à prolonger le produit des réels et à partir de la formule fondamentale:

$$\boxed{i^2 = i \times i = -1.}$$

Le produit de deux nombres complexes quelconques est donné par la formule:

$$\boxed{(x + iy)(x' + iy') = (xx' - yy') + i(xy' + x'y).}$$

Propriétés. Cette multiplication a toutes les bonnes propriétés usuelles: commutativité $[zz' = z'z]$, associativité $[z(z'z'') = (zz')z'']$. Le nombre complexe $1 = 1 + i0$ est neutre pour cette addition $[1z = z1 = z$ pour tout $z \in \mathbb{C}]$. La multiplication est distributive sur l'addition $[z(z' + z'') = zz' + zz'']$.

• Inverse: tout nombre complexe z non-nul admet un inverse noté z^{-1} ou $\frac{1}{z}$; c'est par définition l'unique nombre complexe tel que $zz^{-1} = z^{-1}z = 1$; sa partie réelle et sa partie imaginaire sont données par la formule fondamentale:

$$\boxed{\frac{1}{x + iy} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.}$$

En effet. Remarquons d'abord que le fait que $z = x + iy \neq 0$ équivaut à dire que x et y ne sont pas simultanément nuls, donc $x^2 + y^2 \neq 0$. On pose alors $x' = \frac{x}{x^2+y^2}$ et $y' = -\frac{y}{x^2+y^2}$. On calcule:

$$xx' - yy' = x\tfrac{x}{x^2+y^2} - y\tfrac{-y}{x^2+y^2} = 1 \text{ et } xy' + x'y = x\tfrac{-y}{x^2+y^2} + \tfrac{x}{x^2+y^2}y = 0,$$

ce qui prouve que $(x + iy)(x' + iy') = 1$ et montre le résultat voulu.

1.3 Conjugaison et module.

• Conjugué d'un nombre complexe. On appelle conjugué d'un nombre complexe z le nombre complexe noté $\overline{z}$ qui a la même partie réelle que z et une partie imaginaire opposée:

$$\boxed{\text{si } z = x + iy \ \text{ avec } x, y \in \mathbb{R}, \ \text{ alors } \ \overline{z} = x - iy.}$$

Les propriétés suivantes sont immédiates:

$\overline{\overline{z}} = z$	$\overline{z+z'} = \overline{z} + \overline{z}'$ et $\overline{-z} = -\overline{z}$
$z - \overline{z} = 2i\Im(z)$ et $z + \overline{z} = 2\Re(z)$	$\overline{zz'} = \overline{z}.\overline{z}'$
$\overline{z} = z \Leftrightarrow z \in \mathbb{R}$ et $\overline{z} = -z \Leftrightarrow z \in i\mathbb{R}$	si $z \neq 0$, $\overline{z^{-1}} = (\overline{z})^{-1}$

• Module d'un nombre complexe. Soit $z = x + iy$ un nombre complexe quelconque, avec $x, y \in \mathbb{R}$. Le produit de z par son conjugué est toujours un réel positif:

$$z\overline{z} = (x + iy)(x - iy) = x^2 + y^2 \geq 0.$$

On appelle module de z, noté $|z|$, la racine carrée de ce réel:

$$\boxed{\text{si } z = x + iy \ \text{ avec } x, y \in \mathbb{R}, \quad \text{alors } \ |z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2} \ \in \mathbb{R}_+.}$$

On résume dans le tableau ci-dessous les principales propriétés du module.

$\lvert z\rvert = 0 \Leftrightarrow z = 0$	$\lvert zz'\rvert = \lvert z\rvert.\lvert z'\rvert$	$\lvert z + z'\rvert \leq \lvert z\rvert + \lvert z'\rvert$
$\lvert\overline{z}\rvert = \lvert z\rvert$	$\lvert z^{-1}\rvert = \lvert z\rvert^{-1}$ pour $z \neq 0$	$\lvert\Re(z)\rvert \leq \lvert z\rvert$ avec égalité pour $z \in \mathbb{R}$
$\lvert -z\rvert = \lvert z\rvert$		$\lvert\Im(z)\rvert \leq \lvert z\rvert$ avec égalité pour $z \in i\mathbb{R}$

1.4 Interprétation géométrique.

Dans le plan (affine euclidien) rapporté à un repère orthonormé, tout point M est déterminé par ses deux coordonnées réelles: son abscisse x et son ordonnée y. Le nombre complexe $z = x+iy$ est alors appelé l'*affixe* du point M. Réciproquement, il est clair qu'à tout nombre complexe $z = x + iy$ avec $x, y \in \mathbb{R}$, on associe l'unique point M de coordonnées (x, y) dans le repère choisi. On a ainsi une correspondance bijective entre les points du plan et les nombres complexes.

Il est clair que, si M d'affixe z, alors

- le point M' d'affixe $-z$ est le symétrique de M par rapport à l'origine O du repère.
- le point M_1 d'affixe $\overline{z}$ est le symétrique de M par rapport à l'axe des abscisses du repère.
- le point M_2 d'affixe $-\overline{z}$ est le symétrique de M par rapport à l'axe des ordonnées du repère.

Figure 1

De plus, par définition même du module d'une part et de la distance dans le plan d'autre part, on a:

$$\boxed{\begin{array}{l}\text{Pour tout point } M \text{ d'affixe } z, \text{ le réel positif } |z| = \sqrt{x^2 + y^2} = OM \\ \text{est la distance entre le point } M \text{ et l'origine } O \text{ du repère.}\end{array}}$$

Il en résulte que, si A et B sont deux points d'affixes respectives z_A et z_B, alors

- $|z_B - z_A| = AB = OD$

  où D est le point d'affixe $z_B - z_A$.
- $|z_A + z_B| = OC$

  où C est le point défini par la somme de vecteurs (règle du parallélogramme): $\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OC}$.

Figure 2

2.1 Argument d'un nombre complexe non-nul.

Soit $z = x + iy$ un nombre complexe quelconque, avec $x, y \in \mathbb{R}$. Notons ρ son module:

$$\rho = |z| = \sqrt{\overline{z}z} = \sqrt{x^2 + y^2}.$$

Si $\rho = 0$, alors $z = 0$, et réciproquement.
On supposera désormais que z est non-nul, donc $\rho > 0$.
Introduisons alors le nombre complexe $z_1 = \frac{1}{\rho}z$. Il s'écrit sous forme algébrique $z_1 = x_1 + iy_1$ avec:

$$x_1 = \frac{x}{\rho} = \frac{x}{\sqrt{x^2+y^2}} \quad \text{et} \quad y_1 = \frac{y}{\rho} = \frac{y}{\sqrt{x^2+y^2}}.$$

On a $|z_1| = |\frac{1}{\rho}z| = \frac{1}{|\rho|}|z| = \frac{1}{\rho}|z| = 1$. Ainsi les réels x_1 et y_1 vérifient $x_1^2 + y_1^2 = 1$. Il existe donc un réel θ tel que $x_1 = \cos\theta$ et $y_1 = \sin\theta$. Ainsi on a montré que:

$$\boxed{\text{tout } z \in \mathbb{C} \text{ non-nul s'écrit:} \quad z = \rho(\cos\theta + i\sin\theta), \quad \text{avec } \rho = |z| > 0, \ \text{ et } \ \theta \in \mathbb{R}.}$$

L'expression ci-dessus s'appelle la *forme trigonométrique* du nombre complexe non-nul z.

Un réel θ satisfaisant cette égalité s'appelle un *argument* du nombre complexe non-nul z.
On note $\theta = \arg z$ (modulo 2π, voir ci-dessous).

• Remarques.

1. Par définition, θ est une mesure en radians de l'angle orienté des vecteurs $(\overrightarrow{OI}, \overrightarrow{OM})$ où M est le point du plan d'affixe z et I le point de l'axe des abscisses d'affixe 1.

2. *Attention*: θ n'est pas unique;

   tout réel $\theta + 2k\pi$ avec $k \in \mathbb{Z}$ est un autre argument de z. On appelle parfois argument principal de z, noté $\operatorname{Arg} z$, celui des arguments de z qui appartient à $[-\pi, \pi[$.

3. On a modulo 2π les égalités:

   $\arg\overline{z} = -\arg z$, $\arg(-z) = \arg z + \pi$, $\arg(-\overline{z}) = \pi - \arg z$.

Figure 3

• Propriété fondamentale:
si θ est un argument de z et θ' un argument de z', alors $\theta + \theta'$ est un argument du produit zz'.
Comme on sait déjà que $|zz'| = |z||z'|$, on peut donc résumer en:

$$\boxed{\rho(\cos\theta + i\sin\theta) \times \rho'(\cos\theta' + i\sin\theta') = \rho\rho'\,[\cos(\theta + \theta') + i\sin(\theta + \theta')].}$$

Preuve. Soient $z = \rho(\cos\theta + i\sin\theta)$ et $z' = \rho'(\cos\theta' + i\sin\theta')$. On sait que $|zz'| = \rho\rho'$.
Posons $z_1 = \frac{1}{\rho}z = \cos\theta + i\sin\theta$ et $z_1' = \frac{1}{\rho'}z' = \cos\theta' + i\sin\theta'$, de sorte que $zz' = \rho\rho' z_1 z_1'$. On calcule alors:

$$\begin{aligned} z_1 z_1' &= (\cos\theta + i\sin\theta)(\cos\theta' + i\sin\theta') = (\cos\theta\cos\theta' - \sin\theta\sin\theta') + i(\cos\theta\sin\theta' + \sin\theta\cos\theta') \\ &= \cos(\theta + \theta') + i\sin(\theta + \theta'), \quad \text{ce qui montre le résultat voulu.} \end{aligned}$$

Une conséquence immédiate mais très importante pour les applications est la formule suivante:

• Formule de Moivre: pour tout $n \in \mathbb{N}$ et tout $\theta \in \mathbb{R}$, on a: $\boxed{(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta).}$

2.2 Notation exponentielle.

On introduit la notation: pour tout $\theta \in \mathbb{R}$

$$\boxed{\exp(i\theta) = \cos\theta + i\sin\theta}, \quad \text{ou encore} \quad \boxed{\mathrm{e}^{i\,\theta} = \cos\theta + i\sin\theta}.$$

Il résulte alors immédiatement des résultats du 2.1 ci-dessus que l'on a, pour tout $\theta \in \mathbb{R}$:

$$\boxed{\begin{aligned} |\,\mathrm{e}^{i\,\theta}\,| &= 1, \\ \overline{\mathrm{e}^{i\,\theta}} &= \mathrm{e}^{-i\,\theta}, \end{aligned}} \qquad \begin{array}{l} \text{d'où l'on tire} \\ \text{les formules d'Euler:} \end{array} \qquad \boxed{\begin{aligned} \cos\theta &= \tfrac{1}{2}(\mathrm{e}^{i\,\theta} + \mathrm{e}^{-i\,\theta}), \\ \sin\theta &= \tfrac{1}{2i}(\mathrm{e}^{i\,\theta} - \mathrm{e}^{-i\,\theta}). \end{aligned}}$$

Comme conséquence de la propriété fondamentale du 2.1, on obtient:

$$\boxed{\begin{array}{l}\text{pour tous } \theta, \theta' \in \mathbb{R},\\ \quad e^{i\theta}\, e^{i\theta'} = e^{i(\theta+\theta')}\end{array}}, \text{ en particulier } \boxed{\begin{array}{c}\text{pour tous } \theta \in \mathbb{R},\ n \in \mathbb{N}\\ (e^{i\theta})^n = e^{in\theta}\end{array}} \text{ et } \boxed{\begin{array}{c}\text{pour tout } \theta \in \mathbb{R},\\ \frac{1}{e^{i\theta}} = e^{-i\theta}\end{array}}.$$

Avec cette notation:

$$\boxed{\text{tout } z \in \mathbb{C} \text{ non-nul s'écrit:} \quad z = \rho\, e^{i\theta}, \quad \text{avec } \rho = |z| > 0, \text{ et } \theta \in \mathbb{R}.}$$

Cette expression est appelée la *forme exponentielle* du nombre complexe non-nul z. Elle est particulièrement bien adaptée aux calculs multiplicatifs puisque, dans un produit, les modules se multiplient alors que les arguments s'ajoutent:

$$\boxed{\begin{array}{l}\text{pour tous } \theta, \theta' \in \mathbb{R},\ \rho, \rho' \in \mathbb{R}_+^*\\ \quad \rho\, e^{i\theta} \,.\, \rho'\, e^{i\theta'} = (\rho\rho')\, e^{i(\theta+\theta')}\end{array}}, \text{ et en particulier } \boxed{\begin{array}{c}\text{pour tous } \theta \in \mathbb{R},\ \rho \in \mathbb{R}_+^*,\ n \in \mathbb{N}\\ (\rho\, e^{i\theta})^n = \rho^n\, e^{in\theta}\end{array}}.$$

2.3 Quelques applications.

Les applications des résultats ci-dessus sur les nombres complexes sont innombrables. Sans insister ici sur les applications de nature purement algébriques ou géométriques (très importantes mais pas prioritaires pour nous), donnons deux techniques utiles en analyse.

• Linéarisation des puissances de sin ou cos. C'est une méthode utilisé pour le calcul de certaines primitives de fonctions trigonométriques. Il s'agit d'exprimer $\cos^n x$ ou $\sin^n x$ comme une somme de $\cos(px)$ ou de $\sin(px)$ avec $p \leq n$. Pour cela, on applique directement les formules d'Euler, et la formule du binôme.

Exemple: $\cos^3 x = \frac{1}{2^3}(e^{ix} + e^{-ix})^3 = \frac{1}{2^3}\left[(e^{ix})^3 + 3(e^{ix})^2(e^{-ix}) + 3(e^{ix})(e^{-ix})^2 + (e^{-ix})^3\right]$

$= \frac{1}{2^3}\left[e^{3ix} + e^{-3ix} + 3(e^{ix} + e^{-ix})\right] = \frac{1}{4}\cos 3x + \frac{3}{4}\cos x$

• Problème réciproque. Il s'agit d'exprimer $\cos(nx)$ ou $\sin(nx)$ comme un polynôme en $\cos x$ ou $\sin x$. Pour cela, on applique directement la formule de Moivre, et la formule du binôme.

Exemple: On part du développement:

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

d'où: $\cos(3x) = \Re(\cos x + i\sin x)^3 = \cos^3 x - 3\cos x\sin^2 x$ et $\sin(3x) = \Im(\cos x + i\sin x)^3 = 3\cos^2 x\sin x - \sin^3 x$.

2.4 Fonction exponentielle complexe.

• *La fonction* $t \mapsto e^{it}$. D'après ce que l'on a vu en 2.1 et 2.2, on peut considérer la fonction:

$$\begin{array}{rl} f: & \mathbb{R} \to \mathbb{C}\\ & t \mapsto e^{it} = \cos t + i\sin t\end{array}$$

Sur le plan algébrique, rappelons ses deux propriétés fondamentales:

$$\boxed{|\,e^{it}\,| = 1 \quad \text{pour tout } t \in \mathbb{R}.} \qquad \text{et} \qquad \boxed{e^{it}\, e^{it'} = e^{i(t+t')} \quad \text{pour tous } t, t' \in \mathbb{R}.}$$

Sur le plan de l'analyse, il suffit de rappeler que les fonctions cos et sin sont dérivables sur $\mathbb{R}$ et vérifient $\cos' = -\sin$ et $\sin' = \cos$ pour conclure que:

$$\boxed{f \text{ est dérivable sur } \mathbb{R} \text{ et vérifie } (e^{it})' = i\, e^{it}.}$$

• *La fonction* $z \mapsto e^z$. Soit $z \in \mathbb{C}$; notons $z = x + iy$ avec $x, y \in \mathbb{R}$. Comme on vient de le voir, on peut définir $e^{iy} = \cos y + i\sin y \in \mathbb{C}$. Par ailleurs, la fonction exponentielle réelle d'une variable réelle définit $e^x \in \mathbb{R}$. Le produit du nombre complexe e^{iy} par le nombre réel e^x sera noté e^z. On définit ainsi:

$$\begin{array}{rl} e: & \mathbb{C} \to \mathbb{C}\\ & z = x + iy \mapsto e^z = e^x(e^{iy}) = e^x(\cos y + i\sin y)\end{array}$$

C'est un simple exercice de vérifier à partir de cette définition que:

$$\boxed{e^{z+z'} = e^z\, e^{z'} \quad \text{pour tous } z, z' \in \mathbb{C}}.$$

On verra plus loin (lors de l'étude des séries entières) d'autres façons de définir cette même fonction exponentielle complexe.

Equations et systèmes d'équations

1. Equations algébriques à une inconnue dans $\mathbb{R}$ et dans $\mathbb{C}$

1.1 Racines n-ièmes dans $\mathbb{R}$.

Soit A un nombre réel fixé quelconque. Soit n un entier ≥ 1. Une racine n-ième de A dans $\mathbb{R}$ est un réel x solution de l'équation $x^n = A$. Il est clair que, si $A = 0$, la seule solution est $x = 0$. On prend donc maintenant $A \neq 0$. Rappelons les résultats bien connus suivants:

1. Pour $n = 2$, tout réel $A > 0$ admet 2 racines carrées opposées dans $\mathbb{R}$ (celle des deux qui est positive est notée $\sqrt{A}$, l'autre est donc $-\sqrt{A}$), et tout réel $A < 0$ n'admet aucune racine carrée dans $\mathbb{R}$.
2. Pour $n = 3$, tout réel A admet une unique racine cubique dans $\mathbb{R}$.
3. Pour tout $n \geq 1$, tout réel $A > 0$ admet dans $\mathbb{R}$ une unique racine n-ième *positive* $A^{1/n} = \exp(\frac{1}{n} \ln A)$.

1.2 Racines n-ièmes dans $\mathbb{C}$.

• Prenons maintenant A un nombre complexe fixé quelconque. Soit n un entier ≥ 1. Une racine n-ième de A dans $\mathbb{C}$ est un complexe z solution de l'équation $z^n = A$. Si $A = 0$, la seule solution est $z = 0$. On prend donc maintenant $A \neq 0$. On a le résultat suivant:

Théorème. *Tout nombre complexe non-nul A admet exactement n racines n-ièmes distinctes dans $\mathbb{C}$. Plus précisément, si A est donné sous forme trigonométrique:*

$$A = r\,\mathrm{e}^{i\alpha} = r(\cos\alpha + i\sin\alpha), \qquad \textit{avec}\ \ r = |A| \in \mathbb{R}_+^* \ \ \textit{et}\ \ \alpha = \arg A \in \mathbb{R},$$

les n racines n-ièmes $z_0, z_1, z_2, \ldots, z_{n-1}$ de A sont données sous forme trigonométrique par:

$$z_k = r^{1/n}\,\mathrm{e}^{i\frac{\alpha+2k\pi}{n}} = r^{1/n}(\cos\tfrac{\alpha+2k\pi}{n} + i\sin\tfrac{\alpha+2k\pi}{n}), \qquad \textit{pour}\ k = 0, 1, 2, \ldots, n-1.$$

Remarque 1. Sans écrire en détails la preuve de ce théorème, observons que c'est une application simple et directe de la formule de Moivre. En effet, si l'on cherche $z = \rho\,\mathrm{e}^{i\theta}$ tel que $z^n = A$, l'égalité des modules implique $\rho^n = r$ et l'égalité des arguments implique $\alpha = n\theta$ modulo 2π.

Remarque 2. Précisons bien que $r^{1/n}$ désigne la racine n-ième réelle positive du réel strictement positif r.

Remarque 3. On pourrait penser résoudre le problème sans avoir recours à la forme trigonométrique, c'est-à-dire écrire A sous forme algébrique $A = a+ib$ avec $a, b \in \mathbb{R}$ et chercher $z = x+iy$ avec $x, y \in \mathbb{R}$ tel que $(x+iy)^n = a+ib$ en développant et identifiant les parties réelles et imaginaires. On pourra le faire à titre d'exercice pour $n = 2$ (forme algébrique des racines carrées), mais le procédé devient vite impossible pour les plus grandes valeurs de n.

• Un cas particulier important dans la pratique est celui où $A = 1$. Il s'agit donc de calculer les n nombres complexes distincts solutions de l'équation $z^n = 1$. On les appelle les *racines n-ièmes de l'unité*. On déduit du théorème précédent:

Proposition.

(i) *Pour tout entier $n \geq 0$, les racines n-ièmes de l'unité sont les n nombres complexes:*

$$\boxed{\omega_0 = 1, \quad \omega_1 = \mathrm{e}^{i\frac{2\pi}{n}}, \quad \omega_2 = \mathrm{e}^{i\frac{4\pi}{n}}, \quad \omega_3 = \mathrm{e}^{i\frac{6\pi}{n}}, \quad \ldots, \quad \omega_{n-1} = \mathrm{e}^{i\frac{2(n-1)\pi}{n}}.}$$

(ii) *Pour tout $k = 0, 1, 2, \ldots, n-1$, on a:* $\omega_k = \omega_1^k$.

(iii) *On a:* $\omega_0 + \omega_1 + \omega_2 + \cdots + \omega_{n-1} = 0$.

(iv) *Les n points du plan complexe qui ont pour affixe les racines n-ièmes de l'unité sont situés sur le cercle unité U (cercle de centre O de rayon 1), et sont les sommets du polygone régulier à n côtés inscrit dans U centré en O et passant par le point de l'axe réel d'abscisse 1.*

• Exemples.

Pour $n=2$, on a: $\omega_0 = 1,$
$\omega_1 = e^{i\frac{2\pi}{2}} = e^{i\pi} = -1.$

Les racines carrés de 1 sont 1 et -1;
elles vérifient $1+(-1)=0$.

Pour $n=3$, on a: $\omega_0 = 1,$
$\omega_1 = e^{i\frac{2\pi}{3}} = \cos(\frac{2\pi}{3}) + i\sin(\frac{2\pi}{3}) = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$
$\omega_2 = \omega_1^2 = e^{i\frac{4\pi}{3}} = \cos(\frac{4\pi}{3}) + i\sin(\frac{4\pi}{3}) = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$

On note $\boxed{j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}}$, donc $\boxed{j^2 = \overline{j} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}}$.

Les racines cubiques de 1 sont 1, j et j^2;
elles vérifient $\boxed{1+j+j^2=0}$.

Pour $n=4$, on a: $\omega_0 = 1,$
$\omega_1 = e^{i\frac{2\pi}{4}} = \cos(\frac{\pi}{2}) + i\sin(\frac{\pi}{2}) = i$
$\omega_2 = \omega_1^2 = e^{i\frac{4\pi}{4}} = \cos\pi + i\sin\pi = -1.$
$\omega_3 = \omega_1^3 = e^{i\frac{6\pi}{4}} = \cos(\frac{3\pi}{2}) + i\sin(\frac{3\pi}{2}) = -i.$

Les racines quatrièmes de 1 sont 1, i, -1 et $-i$;
elles vérifient $1+i+(-1)+(-i)=0$.

Pour $n=6$, on a: $\omega_0 = 1,$
$\omega_1 = e^{i\frac{2\pi}{6}} = \cos(\frac{\pi}{3}) + i\sin(\frac{\pi}{3}) = \frac{1}{2} + i\frac{\sqrt{3}}{2} = -j^2$
$\omega_2 = \omega_1^2 = e^{i\frac{4\pi}{6}} = \cos(\frac{2\pi}{3}) + i\sin(\frac{2\pi}{3}) = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = j.$
$\omega_3 = e^{i\frac{6\pi}{6}} = \cos\pi + i\sin\pi = -1$
$\omega_4 = e^{i\frac{8\pi}{6}} = \cos(\frac{4\pi}{3}) + i\sin(\frac{4\pi}{3}) = -\frac{1}{2} - i\frac{\sqrt{3}}{2} = j^2$
$\omega_5 = e^{i\frac{10\pi}{6}} = \cos(\frac{5\pi}{3}) + i\sin(\frac{5\pi}{3}) = \frac{1}{2} - i\frac{\sqrt{3}}{2} = -j.$

Les racines sixièmes de 1 sont $1, -j^2, j, -1, j^2, -j$;
elles vérifient $1 - j^2 + j - 1 + j^2 - j = 0$.

$n=2$
$n=3$
$n=4$
$n=6$

1.2 Equations de degré 2.

• Considérons une équation de degré 2 à coefficients réels:

$$\boxed{ax^2+bx+c=0, \quad \text{avec } a,b,c\in\mathbb{R},\ a\neq 0.} \qquad (1)$$

On introduit son discriminant: $\boxed{\Delta = b^2-4ac \ \in \mathbb{R}.}$

Rappelons le résultat bien connu suivant, qui donne dans tous les cas les solutions réelles de (1).

$$\begin{cases} \textit{si } \Delta>0, & \textit{alors (1) a deux solutions distinctes dans } \mathbb{R}\textit{, données par: } x_1 = \frac{-b-\sqrt{\Delta}}{2a} \textit{ et } x_2 = \frac{-b+\sqrt{\Delta}}{2a}, \\ \textit{si } \Delta=0, & \textit{alors (1) a une unique solution dans } \mathbb{R}\textit{, donnée par: } x_0 = -\frac{b}{2a}, \\ \textit{si } \Delta<0, & \textit{alors (1) n'a aucune solution dans } \mathbb{R}. \end{cases}$$

Revenons sur le cas où $\Delta<0$. Le réel Δ n'a pas de racine carrée dans $\mathbb{R}$, mais il a deux racines carrées distinctes dans $\mathbb{C}$. En effet, puisque $-\Delta>0$, on peut considérer $\sqrt{-\Delta}$, et les deux nombres complexes $\delta = i\sqrt{-\Delta}$ et $-\delta = -i\sqrt{-\Delta}$ qui vérifient $\delta^2 = (-\delta)^2 = (i\sqrt{-\Delta})^2 = (-i\sqrt{-\Delta})^2 = (-1)\times(-\Delta) = \Delta$. En particulier, on a $b^2-\delta^2 = 4ac$. Il en résulte que, suivant le même calcul que dans le cas $\Delta>0$, les nombres complexes $z_1 = \frac{1}{2a}(-b-\delta)$ et $z_2 = \frac{1}{2a}(-b+\delta)$ vérifient:

$$a(x-z_1)(x-z_2) = ax^2 - a(z_1+z_2)x + az_1z_2 = ax^2 - a\frac{-2b}{2a}x + a\frac{b^2-\delta^2}{4a^2} = ax^2+bx+c.$$

Donc z_1 et z_2 sont solutions de (1). De plus, ils sont conjugués car: $z_1 = \frac{-b}{2a} - i\frac{\sqrt{-\Delta}}{2a}$ et $z_2 = \frac{-b}{2a} + i\frac{\sqrt{-\Delta}}{2a}$.

Conclusion: *si* $\Delta<0$, *alors l'équation à coefficients réels* (1) *a deux solutions distinctes dans* $\mathbb{C}$, *qui sont les deux nombres complexes conjugués:*

$$z_1 = \frac{-b-i\sqrt{-\Delta}}{2a} \quad \textit{et} \quad z_2 = \frac{-b+i\sqrt{-\Delta}}{2a}$$

• Considérons maintenant une équation de degré 2 à coefficients complexes:

$$\boxed{az^2 + bz + c = 0, \quad \text{avec } a, b, c \in \mathbb{C},\ a \neq 0,} \qquad (2)$$

et son discriminant:

$$\boxed{\Delta = b^2 - 4ac \ \in \mathbb{C}.}$$

Comme Δ est a priori un nombre complexe, ça n'a aucun sens de parler de son signe. Mais dans $\mathbb{C}$ tout nombre complexe a deux racines carrées opposées (voir ci-dessus), et on peut donc consdérer les deux nombres complexes δ et $-\delta$ tels que $\delta^2 = \Delta$. Des calculs en tout point similaires à ceux faits ci-dessus dans le cas particulier où les coefficients de l'équation sont réels montrent alors que:

> *L'équation à coefficients complexes* (2) *a deux solutions dans* $\mathbb{C}$, *qui sont:*
> $$z_1 = \frac{-b-\delta}{2a} \quad \text{et} \quad z_2 = \frac{-b+\delta}{2a}$$
> *où* $\pm\delta$ *sont les deux racines carrées complexes de* Δ. *Les solutions* z_1 *et* z_2 *sont confondues et égales à* $-\frac{b}{2a}$ *lorsque* $\Delta = 0$, *et distinctes si* $\Delta \neq 0$.

2. Rudiments pratiques sur les systèmes d'équations linéaires

2.1 Equations linéaires à deux inconnues.

Fixons trois réels a, b, c.
On suppose que a et b ne sont pas tous les deux nuls (l'un au moins est non-nul), c'est-à-dire $(a, b) \neq (0, 0)$.
Formons l'équation linéaire à deux inconnues:

$$\boxed{ax + by = c} \qquad (1)$$

Une solution réelle de l'équation (1) est un couple de réels (x_0, y_0) pour lequel la relation $ax_0 + by_0 = c$ est effectivement vérifiée. Résoudre l'équation dans $\mathbb{R}$ consiste à déterminer l'ensemble de toutes ses solutions réelles. L'ensemble $\mathcal{S}$ des solutions de (1) est donc un sous-ensemble de l'ensemble $\mathbb{R}^2$ des couples de réels.

• Résolution algébrique de (1). Deux cas peuvent se présenter:

1. Si $b = 0$, alors on a forcément $a \neq 0$, et (1) équivaut à $x = \frac{c}{a}$. Ainsi une solution est un couple dont le premier terme vaut la constante $\frac{c}{a}$ et dont le second terme est quelconque dans $\mathbb{R}$. On note:
$$\mathcal{S} = \{ \ (\tfrac{c}{a}\, ,\, y) \ ; \ y \in \mathbb{R} \ \}.$$
2. Si $b \neq 0$, alors $y = \frac{1}{b}(c - ax) = -\frac{a}{b}x + \frac{c}{b}$. Ainsi une solution est un couple dont le second terme s'exprime en fonction du premier suivant la relation précédente. On note:
$$\mathcal{S} = \{ \ (x\, ,\, -\tfrac{a}{b}x + \tfrac{c}{b}) \ ; \ x \in \mathbb{R} \ \}.$$

Dans les deux cas, l'ensemble $\mathcal{S}$ est non-vide, et il est infini, dépendant d'un paramètre réel.

• Interprétation géométrique de (1).

Dans le plan (affine, euclidien) muni d'un repère orthonormé, on peut identifier tout point M au couple (x, y) de ses coordonnées relativement à ce repère. Comme tout sous-ensemble de $\mathbb{R}^2$, l'ensemble $\mathcal{S}$ s'identifie à un sous-ensemble de points du plan. Il est bien connu qu'il s'agit ici d'une droite, dont on rappelle ci-dessous sans démonstration la description:

1. Si $b = 0$, alors $\mathcal{S}$ est la droite verticale passant par le point de l'axe horizontal d'abscisse $\frac{c}{a}$.
2. Si $b \neq 0$, alors $\mathcal{S}$ est la droite de pente $-\frac{a}{b}$ coupant l'axe vertical au point d'ordonnée $\frac{c}{b}$.

cas $b = 0$

cas $b \neq 0$, avec $a \neq 0$ ou $a = 0$

Remarque. On a dès le départ exclu le cas $a = b = 0$. Dans ce cas en effet, la résolution de (1) est triviale:

ou $c = 0$, alors l'équation $0x + 0y = 0$ est trivialement vérifiée par tout couple de réels. Donc $\mathcal{S} = \mathbb{R}^2$.

ou $c \neq 0$, alors l'équation $0x + 0y = c \neq 0$ n'admet aucune solution. On dit que $\mathcal{S}$ est vide, ce que l'on note $\mathcal{S} = \emptyset$.

2.2 Systèmes d'équations linéaires à deux inconnues.

Fixons six réels a, b, c, a', b', c'. Formons le système d'équations linéaires à deux inconnues:

$$(\Sigma) \qquad \boxed{\begin{cases} ax + by = c & (1) \\ a'x + b'y = c' & (2) \end{cases}}$$

Une solution réelle du système (Σ) est un couple de réels (x_0, y_0) pour lequel les deux relations $ax_0 + by_0 = c$ et $a'x_0 + b'y_0 = c'$ sont simultanément vérifiées, c'est-à-dire qui est solution à la fois de (1) et de (2). L'ensemble $\mathcal{S}$ des solutions de (Σ) est donc le sous-ensemble de $\mathbb{R}^2$ formé des couples de réels qui appartiennent à la fois à l'ensemble $\mathcal{S}_1$ des solutions de (1) et à l'ensemble $\mathcal{S}_2$ des solutions de (2). On dit que $\mathcal{S}$ est l'intersection de $\mathcal{S}_1$ et $\mathcal{S}_2$, et l'on note $\mathcal{S} = \mathcal{S}_1 \cap \mathcal{S}_2$.

• Les cas dégénérés.

→ Si $(a, b) = (0, 0)$ et $c \neq 0$, alors (1) n'a pas de solution, et donc le système (Σ) n'en a pas non plus; dans ce cas $\mathcal{S} = \emptyset$.

→ Si $(a, b) = (0, 0)$ et $c = 0$, alors tout couple de réels est solution de (1), donc les solutions du système (Σ) sont celles de la seule équation (2); on est ramené à l'étude de 2.1 et $\mathcal{S} = \mathcal{S}_2$ peut être soit $\emptyset$ (lorsque $a' = b' = 0$ et $c' \neq 0$), soit $\mathbb{R}^2$ tout entier (lorsque $a' = b' = c' = 0$), soit un ensemble infini à un paramètre correspondant à une droite.

→ Si $(a, b) \neq (0, 0)$ mais si $(a', b') = (0, 0)$, idem en échangeant les rôles des équations (1) et (2).

On pourra donc supposer dans la suite que $(a, b) \neq (0, 0)$ et $(a', b') \neq (0, 0)$.

• Méthode géométrique.

Sous réserve que chacun des deux couples $(a; b)$ et (a', b') soit différent de $(0, 0)$, l'ensemble $\mathcal{S}_1$ des solutions de (1) correspond à une droite D_1 du plan, l'ensemble $\mathcal{S}_2$ des solutions de (2) correspond à une droite D_2, et l'ensemble $\mathcal{S}$ des solutions de (Σ) est l'intersection des droites D_1 et D_2. Trois cas sont possibles:

si $ab' \neq a'b$, les pentes des droites sont distinctes, donc elles ne sont pas parallèles; leur intersection est un unique point M_0, et ses coordonnées forment l'unique solution du système (Σ).

si $ab' = a'b$, les droites sont parallèles;

ou $b'c = bc'$, dans ce cas les droites D_1 et D_2 sont confondues, et donc $\mathcal{S} = \mathcal{S}_1 = \mathcal{S}_2$ est l'ensemble infini à un paramètre correspondant à cette droite.

ou $b'c \neq bc'$, dans ce cas les droites D_1 et D_2 sont strictement parallèles, donc elles n'ont aucun point commun, et $\mathcal{S}$ est alors vide.

premier cas	deuxième cas	troisième cas

• MÉTHODE ALGÉBRIQUE.

Il résulte en particulier de ce qui précède que:

$$\boxed{\text{le système } (\Sigma) \text{ a une unique solution si et seulement si } \; a'b - ab' \neq 0.}$$

On appelle *déterminant* du système (Σ) le réel $ab' - a'b$. On le note: $\boxed{\left|\begin{smallmatrix} a & b \\ a' & b' \end{smallmatrix}\right| = ab' - a'b}$.

Un système dont le déterminant est non-nul s'appelle un *système de Cramer*; il a donc une unique solution. Lorsque le déterminant est nul, le système peut avoir pour ensemble de solutions: soit $\mathbb{R}^2$ tout entier, soit un sous-ensemble infini de $\mathbb{R}^2$ dépendant linéairement d'un paramètre (une droite), soit l'ensemble vide (aucune solution).

La question est de disposer de méthode permettant de déterminer explicitement l'ensemble des solutions.

Un premier procédé (élémentaire) est la méthode dite *par substitution*.

> Il consiste à exprimer par exemple y en fonction de x dans l'une des équations, on reporte cette expression dans l'autre équation qui ne fait alors intervenir que x, on résout en x, et on en déduit y en revenant à l'expression de départ. Pour évidente qu'elle soit, cette méthode a deux inconvénients: (1) elle est assez efficace pour des systèmes avec des coefficients numériques mais se prête mal aux raisonnements avec discussion suivant divers cas pour les systèmes dépendant de paramètres, (2) elle devient plus difficile à mettre en œuvre pour des systèmes de plus de deux équations avec plus de deux inconnues.

C'est pourquoi on préferera souvent la méthode dite *méthode de Gauss* (ou *méthode du pivot*, ou encore *méthode d'échelonnement par combinaisons linéaires*), qui a l'avantage de s'appliquer dans un cadre tout à fait général. Nous la décrirons au paragraphe suivant. Contentons-nous ici de donner un exemple:

EXEMPLE. Fixons un paramètre réel $\alpha \neq 0$ et considérons le système $(\Sigma)\begin{cases} x + \alpha y = 1 & (1) \\ \alpha x + 4y = 2 & (2) \end{cases}$.

Comme $\alpha \neq 0$, on obtient un système équivalent en multipliant (1) par α: $\begin{cases} \alpha x + \alpha^2 y = \alpha & (1') \\ \alpha x + 4y = 2 & (2) \end{cases}$.

On remplace ensuite (2) par la différence (2) − (1): $\begin{cases} \alpha x + \alpha^2 y = \alpha & (1') \\ (4-\alpha^2)y = 2 - \alpha & (2') \end{cases}$.

Trois cas apparaissent alors, suivant que $4 - \alpha^2$ est nul ou non:

- Si $\alpha \neq 2$ et $\alpha \neq -2$, alors on tire de (2') que $y = \frac{2-\alpha}{4-\alpha^2} = \frac{1}{\alpha+2}$. On reportant dans (1), il vient que $x = 1 - \alpha y = 1 - \frac{\alpha}{\alpha+2} = \frac{2}{\alpha+2}$. Dans ce cas, le système a une unique solution qui est le couple $(\frac{2}{\alpha+2}, \frac{1}{\alpha+2})$.

  L'ensemble $\mathcal{S}$ des solutions est alors noté: $\boxed{\mathcal{S} = \{(\frac{2}{\alpha+2}, \frac{1}{\alpha+2})\}}$.
- Si $\alpha = -2$, alors on réécrit (Σ): $\begin{cases} x - 2y = 1 \\ -2x + 4y = 2 \end{cases}$ ou encore en multipliant (2) par $-\frac{1}{2}$, $\begin{cases} x - 2y = 1 \\ x - 2y = -1 \end{cases}$.

  Il est clair que les deux équations sont incompatibles. Le système n'a pas de solution. On note $\boxed{\mathcal{S} = \emptyset}$.
- Si $\alpha = 2$, alors on réécrit (Σ): $\begin{cases} x + 2y = 1 \\ 2x + 4y = 2 \end{cases}$ ou encore en multipliant (2) par $\frac{1}{2}$, $\begin{cases} x + 2y = 1 \\ x + 2y = 1 \end{cases}$. Il est clair que les deux équations sont équivalentes. On a $\boxed{\mathcal{S} = \{\ (1-2y, y)\ ;\ y \in \mathbb{R}\}}$.

On remarquera que le réel $4 - \alpha^2$ qui apparaît naturellement dans cette méthode de résolution n'est autre que le déterminant du système; il est donc normal que sa nullité ou non soit le point crucial de la discussion.

• REMARQUE. On considère de même des systèmes de 3 (ou plus) équations à deux inconnues correspondant à l'intersection de 3 droites (ou plus...) dans le plan, avec tous les cas de positions relatives de ces droites.

2.3 Systèmes d'équations linéaires à trois inconnues ou plus.

Un système linéaire de trois équations à trois inconnues dans $\mathbb{R}$ est un système de la forme:

$$(\Sigma) \begin{cases} ax + by + cz = d & (1) \\ a'x + b'y + c'z = d' & (2) \\ a''x + b''y + c''z = d'' & (3) \end{cases},$$

où les coefficients $a, a', a'', b, b', b'', c, c', c'', d, d', d''$ sont fixés dans $\mathbb{R}$. Une solution de (Σ) est donc un triplet $(x_0, y_0, z_0) \in \mathbb{R}^3$ pour lequel ces trois relations sont simultanément vérifiées. L'ensemble des solutions $\mathcal{S}$ que l'on cherche à déterminer lorsque l'on résout le système est donc un sous-ensemble de $\mathbb{R}^3$.

Remarque. Si aucune des 3 équations n'est dégénérée, un tel système s'interprète géométriquement comme l'intersection de 3 plans dans l'espace, qui peut être vide, ou réduite à un point (le système a alors une seule solution, qui est un triplet de réels), ou égale à une droite (l'ensemble des solutions du système dépend alors linéairement d'un paramètre), ou égale à un plan (l'ensemble des solutions dépend alors linéairement de deux paramètres).

Lorsque le système a une seule solution, on dit que c'est un système *de Cramer*. Comme dans le cas précédent, cela se produit si et seulement si le *déterminant* du système est non-nul. Mais dans le cas d'un système 3×3, le déterminant est plus difficile à définir que dans le cas 2×2. On verra cela plus tard dans un autre chapitre.

• Méthode de Gauss.

Elle repose sur les trois règles fondamentales suivantes, que l'on montre en toute généralité, pour tout système linéaire de m équations à n inconnues:

(R1) *si l'on permute deux équations du système, on obtient un système équivalent (c'est-à-dire qui a exactement le même ensemble de solutions);*

(R2) *si l'on multiplie les deux membres d'une des équations du système par un même réel* non-nul*, on obtient un système équivalent;*

(R3) *Si l'on ajoute à l'une des équations du système le produit d'une autre équation par un réel, on obtient un système équivalent.*

Elles permettent de se ramener à un système (équivalent à celui de départ) qui est *échelonné*, c'est-à-dire tel que le nombre d'inconnues apparaissant dans chaque équation diminue strictement d'équation en équation. La dernière donne alors l'une des inconnues, dont on reporte la valeur dans les équations précédentes, et l'on remonte le système dans l'autre sens pour trouver de proche en proche toutes les inconnues.

• Exemples numériques.

$$(\Sigma_1)\ \begin{cases} x+y+z=1 \\ x+2y-z=-1 \\ 2x+y+3z=1 \end{cases} \xrightarrow[(3)-2\times(1)]{(2)-(1)} \begin{cases} x+y+z=1 \\ \quad y-2z=-2 \\ \ -y+z=-1 \end{cases} \xrightarrow{(3)+(2)} \begin{cases} x+y+z=1 \\ \quad y-2z=-2 \\ \qquad -z=-3 \end{cases} \longrightarrow \begin{cases} z=3 \\ y=2z-2=4 \\ x=-y-z+1=-6 \end{cases}$$

Le système (Σ_1) a une unique solution, qui est le triplet $(-6,4,3)$. On note $\mathcal{S}_1 = \{\ (-6,4,3)\ \}$.

$$(\Sigma_2)\ \begin{cases} x+y+z=1 \\ x+2y-z=-1 \\ 2x+y+4z=4 \end{cases} \xrightarrow[(3)-2\times(1)]{(2)-(1)} \begin{cases} x+y+z=1 \\ \quad y-2z=-2 \\ \ -y+2z=2 \end{cases} \longrightarrow \begin{cases} x+y+z=1 \\ \quad y-2z=-2 \end{cases} \longrightarrow \begin{cases} y=2z-2 \\ x=-y-z+1=-3z+3 \end{cases}$$

Les solutions de (Σ_2) sont donc tous les triplets de la forme $(-3z+3, 2z-2, z)$ avec z quelconque. Le système a donc une infinité de solutions dépendant linéairement d'un paramètre; on note $\mathcal{S}_2 = \{\ (-3z+3, 2z-2, z)\ ;\ z \in \mathbb{R}\ \}$.

$$(\Sigma_3)\ \begin{cases} x+y+z=1 \\ x+2y-z=-1 \\ 2x+y+4z=3 \end{cases} \xrightarrow[(3)-2\times(1)]{(2)-(1)} \begin{cases} x+y+z=1 \\ \quad y-2z=-2 \\ \ -y+2z=1 \end{cases} \longrightarrow \begin{cases} x+y+z=1 \\ \quad y-2z=-2 \\ \quad y-2z=-1 \end{cases} \longrightarrow \text{les deux dernières équations sont incompatibles}$$

Le système (Σ_3) n'a donc aucune solution; on note $\mathcal{S}_3 = \emptyset$.

• Exemple avec paramètres. Soient α et β fixés dans $\mathbb{R}$. On considère dans $\mathbb{R}^3$ le système linéaire:

$$(\Sigma)\ \left\{ \begin{array}{ccccccc} x & + & \beta y & + & \alpha z & = & 0 \quad (1) \\ x & + & y & + & \alpha z & = & 0 \quad (2) \\ x & + & y & + & z & = & 0 \quad (3) \end{array} \right. \xrightarrow[(2)-(3)]{(1)-(2)} (\Sigma')\ \left\{ \begin{array}{ccc} (\beta-1)y & = & 0 \quad (1') \\ (\alpha-1)z & = & 0 \quad (2') \\ x+y+z & = & 0 \quad (3) \end{array} \right.$$

On raisonne donc en distinguant les quatre cas suivants:

1er cas: $\alpha \neq 1$ et $\beta \neq 1$

$(1')$ a pour seule solution $y=0$ et $(2')$ a pour seule solution $z=0$. Alors (3) devient $x=0$. L'ensemble des solutions de (Σ) est donc l'ensemble à un seul élément $\{(0,0,0)\}$.

2ème cas: $\alpha = 1$ et $\beta \neq 1$

$(1')$ a pour seule solution $y=0$ et tout $z \in \mathbb{R}$ est solution de $(2')$. Alors (3) devient $x+z=0$. L'ensemble des solutions de (Σ) est le sous-ensemble $D = \{(-z,0,z)\,;\, z \in \mathbb{R}\}$ de $\mathbb{R}^3$; il y a une infinité de solutions dépendant linéairement d'un paramètre.

3ème cas: $\alpha \neq 1$ et $\beta = 1$

$(2')$ a pour seule solution $z=0$ et tout $y \in \mathbb{R}$ est solution de $(1')$. Alors (3) devient $x+y=0$. L'ensemble des solutions de (Σ) est le sous-ensemble $D' = \{(-y,y,0); y \in \mathbb{R}\}$ de $\mathbb{R}^3$; il y a une infinité de solutions dépendant linéairement d'un paramètre.

4ème cas: $\alpha = 1$ et $\beta = 1$

Tout $y \in \mathbb{R}$ est solution de $(1')$ et tout $z \in \mathbb{R}$ est solution de $(2')$. Le système équivaut à la seule équation $x+y+z=0$. L'ensemble des solutions de (Σ) est le sous-ensemble $P = \{(-y-z,y,z)\,;\, y,z \in \mathbb{R}\}$ de $\mathbb{R}^3$; il y a une infinité de solutions dépendant linéairement de deux paramètres.

Développements limités

1. Notion de développement limité

1.1 Définition et premières propriétés.

On se donne un réel x_0 fixé.
On considère un voisinage de x_0, c'est-à-dire ici un intervalle $]x_0-\alpha, x_0+\alpha[$, avec $\alpha>0$.
Soit f une fonction réelle d'une variable réelle définie sur $]x_0-\alpha, x_0+\alpha[$, sauf éventuellement au point x_0.

On dit que f admet un *développement limité à l'ordre* n (où $n\in\mathbb{N}$ fixé) au voisinage de x_0 lorsqu'il existe un réel $\beta<\alpha$, des réels $a_0, a_1, \dots, a_n$, et une fonction $\varepsilon :]x_0-\beta, x_0+\beta[\to \mathbb{R}$ tels que:

$$\begin{cases} f(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + (x-x_0)^n\varepsilon(x) & \text{pour tout } x \in]x_0-\beta, x_0+\beta[\,, \\ \text{avec } \lim\limits_{x\to x_0} \varepsilon(x) = 0 \end{cases}$$

Le polynôme $a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n$ s'appelle la *partie régulière* du D.L.,
le terme $(x-x_0)^n\varepsilon(x)$ s'appelle le *reste.*

On peut démontrer les propriétés suivantes:

1. Si f admet un D.L. à l'ordre n au voisinage de x_0, alors il est unique.
2. Si f admet un D.L. à l'ordre n au voisinage de x_0, alors f admet un D.L. au voisinage de x_0 à tout ordre $m\leq n$, dont la partie régulière s'obtient en prenant les $m+1$ premiers termes du D.L. à l'ordre n.
3. Si f admet un D.L. à l'ordre $n\geq 0$ au voisinage de x_0, alors $\lim\limits_{x\to x_0} f(x) = a_0$.

Exemple: Considérons la fonction $f(x) = \frac{1}{1-x}$ sur l'intervalle $I = \left]-\frac{1}{2}, +\frac{1}{2}\right[$. Pour tout $n\in\mathbb{N}$, on sait que $1+x+x^2+x^3+\cdots+x^n = \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x} - x^n\frac{x}{1-x}$. Posons $\varepsilon(x) = \frac{x}{1-x}$ pour tout $x\in I$. On a donc:

$$\frac{1}{1-x} = 1+x+x^2+x^3+\cdots+x^n+x^n\varepsilon(x), \quad \text{avec } \lim_{x\to 0}\varepsilon(x) = 0.$$

1.2 Liens avec la formule de Taylor-Young

Théorème. *Si f est de classe C^{n-1} au voisinage de x_0 et si $f^{(n)}(x_0)$ existe, alors f admet un D.L. à l'ordre n au voisinage de x_0, qui est donné par la formule de Taylor-Young:*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \tfrac{1}{2}f''(x_0)(x-x_0)^2 + \cdots + \tfrac{1}{n!}f^{(n)}(x_0)(x-x_0)^n + (x-x_0)^n\varepsilon(x).$$

Remarque 1. Ce théorème (que nous rappelons ici sans démonstration) donne une condition suffisante pour l'existence d'un D.L., mais elle n'est pas nécessaire. C'est-à-dire qu'on peut avoir l'existence d'un D.L. sans que les hypothèses de la formule de Taylor-Young soient satisfaites.

> Contre-exemple: soit f la fonction définie par $f(x) = x+x^2+x^3\cos\frac{1}{x}$ si $x\neq 0$, et par $f(0)=0$.
>
> Posons $\varepsilon(x) = x\cos\frac{1}{x}$. On a $\lim\limits_{x\to 0}\epsilon(x) = 0$ puisque $|\varepsilon(x)|\leq |x|$. Donc, par construction même, la fonction f admet $f(x) = x+x^2+x^2\varepsilon(x)$ comme D.L. à l'ordre 2 au voisinage de 0. Et pourtant, f est bien continue et dérivable sur $\mathbb{R}$, mais on peut aisément vérifier que $f''(0)$ n'existe pas [car $\frac{1}{x}(f'(x)-f'(0)) = 2+3x\cos\frac{1}{x}+\sin\frac{1}{x}$ n'admet pas de limite quand $x\to 0$].

Remarque 2. Ce théorème peut permettre de calculer effectivement un D.L.

> Exemple: soit f la fonction définie par $f(x) = \sqrt{2+2\,\mathrm{e}^x}$ pour tout $x\in\mathbb{R}$. Elle est de classe C^∞ sur $\mathbb{R}$ donc le théorème ci-dessus s'applique. Déduisons-en par exemple un D.L. de f à l'ordre 3 au voisinage de 0. Pour cela, on calcule les dérivées successives:
>
> $$f'(x) = (2+2\,\mathrm{e}^x)^{-1/2}\,\mathrm{e}^x, \quad f''(x) = (2+2\,\mathrm{e}^x)^{-3/2}(\mathrm{e}^{2x}+2\,\mathrm{e}^x), \quad f'''(x) = (2+2\,\mathrm{e}^x)^{-5/2}(\mathrm{e}^{3x}+2\,\mathrm{e}^{2x}+4\,\mathrm{e}^x).$$
>
> On en tire en particulier les valeurs: $f(0)=2,\ f'(0)=\frac{1}{2},\ f''(0)=\frac{3}{8},\ f'''(0)=\frac{7}{32}$.
>
> En les reportant dans la formule de Taylor-Young, on en déduit le D.L. de $f(x)$ à l'ordre 2 au voisinage de 0:
>
> $$\sqrt{2+2\,\mathrm{e}^x} = 2+\tfrac{1}{2}x+\tfrac{3}{16}x^2+\tfrac{7}{192}x^3+x^3\varepsilon(x).$$

On verra plus loin d'autres méthodes (en général plus efficaces et rapides) de calculs explicites de D.L.

1.3 Réduction au cas des D.L. au voisinage de 0.

Reprenons les données et notations du paragraphe 1.1. Il est clair que dire que f admet un D.L. à l'ordre n au voisinage de x_0 équivaut à dire que la fonction g définie par $g(x) = f(x+x_0)$ admet un D.L. à l'ordre n au voisinage de 0, ce dernier étant donné par $g(x) = a_0+a_1x+a_2x^2+\cdots+a_nx^n+x^n\varepsilon'(x)$, avec $\varepsilon'(x) = \varepsilon(x_0+x)$.

En clair: *l'étude d'un D.L. peut toujours se ramener à celle d'un D.L. au voisinage de 0.*

C'est pourquoi, classiquement, on se limite pour les développements techniques qui vont suivre, à considérer des D.L. au voisinage de 0. Répétons ce que deviennent dans ce cas la définition 1.1 et le théorème 1.2:

• On dit qu'une fonction f définie au voisinage de 0 admet un D.L. à l'ordre n au voisinage de 0 lorsqu'il existe un intervalle $]-\beta,+\beta[$, des réels $a_0, a_1, \dots, a_n$, et une fonction $\varepsilon :]-\beta,+\beta[\to \mathbb{R}$ tels que:

$$\boxed{\begin{cases} f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + x^n\varepsilon(x) & \text{pour tout } x \in]-\beta,+\beta[\,, \\ \text{avec } \lim\limits_{x\to 0} \varepsilon(x) = 0 \end{cases}}$$

• Si f est de classe C^{n-1} au voisinage de 0 et si $f^{(n)}(0)$ existe, alors f admet un D.L. à l'ordre n au voisinage de 0, qui est donné par la formule de Taylor-Young:

$$\boxed{f(x) = f(0) + f'(0)x + \tfrac{1}{2}f''(0)x^2 + \cdots + \tfrac{1}{n!}f^{(n)}(0)x^n + x^n\varepsilon(x).}$$

Exemple fondamental. Un exemple d'application du théorème ci-dessus est celui de la fonction $f(x) = \mathrm{e}^x$. Elle est de classe C^∞ sur $\mathbb{R}$ et vérifie $\mathrm{e}'(x) = \mathrm{e}(x)$ pour tout $x \in \mathbb{R}$, de sorte que toutes les dérivées successives $\mathrm{e}^{(n)}(x)$ sont aussi égales à e^x. Il en résulte que $\mathrm{e}^{(n)}(0) = \cdots = \mathrm{e}''(0) = \mathrm{e}'(0) = \mathrm{e}(0) = 1$, et l'application de la formule de Taylor-Young donne un D.L. de e^x à tout ordre n au voisinage de 0:

$$\mathrm{e}^x = 1 + x + \tfrac{1}{2!}x^2 + \tfrac{1}{3!}x^3 + \cdots + \tfrac{1}{n!}x^n + x^n\varepsilon(x)$$

Remarque. Si une fonction paire f admet un D.L. au voisinage de 0, alors la partie régulière de ce D.L. n'admet que des puissances de x d'exposant pair (c'est-à-dire $a_i = 0$ si i impair). Idem si f impaire.

1.4 Généralisations des développements limités.

• Si f est définie seulement à droite , on peut parler de D.L. à droite: la définition est la même, mais la fonction ε est définie a priori seulement à droite de 0, et vérifie $\lim\limits_{x\to 0, x>0} \varepsilon(x) = 0$. Idem à gauche.

• Soit f une fonction définie sur un intervalle de la forme $]\alpha,+\infty[$. On dit que f admet un *développement limité (généralisé) au voisinage de* $+\infty$ lorsque la fonction $x \mapsto f(\frac{1}{x})$ admet un D.L. au voisinage de 0 à droite. Idem au voisinage de $-\infty$.

Exemple. Considérons la fonction $f(x) = \frac{x}{x-1}$ sur $]1,+\infty[$. On a $f(x) = \frac{1}{1-\frac{1}{x}}$. Lorsque x tend vers $+\infty$, l'inverse $y = \frac{1}{x}$ tend vers 0 à droite. Or on a vu en 1.1 que la fonction $y \mapsto \frac{1}{1-y}$ admet pour D.L. à un ordre n quelconque au voisinage de 0: $\frac{1}{1-y} = 1 + y + y^2 + y^3 + \cdots + y^n + y^n\varepsilon(y)$. On en déduit que:

$$\tfrac{x}{1-x} = 1 + \tfrac{1}{x} + \tfrac{1}{x^2} + \tfrac{1}{x^3} + \cdots + \tfrac{1}{x^n} + \tfrac{1}{x^n}\varepsilon(\tfrac{1}{x}), \quad \text{avec } \lim_{x\to+\infty} \varepsilon(\tfrac{1}{x}) = 0$$

est un D.L. (généralisé) de la fonction f à l'ordre n au voisinage de $+\infty$.

• Même si f n'admet pas de D.L. au voisinage de 0, il peut exister une fonction g équivalente à f au voisinage de 0 telle que le quotient $\frac{f}{g}$ admette un D.L. au voisinage de 0, permettant d'avoir une expression de la forme:

$$f(x) = g(x)[a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + x^n\varepsilon(x)] \quad \text{avec } \lim_{x\to 0} \varepsilon(x) = 0$$

Exemple. $\frac{1}{x-x^2} = \frac{1}{x}[1 + x + x^2 + x^3 + \cdots + x^n + x^n\varepsilon(x)]$ avec $\lim\limits_{x\to 0} \varepsilon(x) = 0$

2. Méthodes de calculs

Le principe général que l'on adopte dans la pratique est de connaître quelques D.L. classiques de fonctions de références (déterminés par exemple grâce au théorème 1.2), puis de combiner ceux-ci grâce aux règles rappelées ci-dessous (sans démonstration).

2.1 Quelques D.L. classiques

Au voisinage de 0, pour tout entier $n \geq 1$, et avec $\lim\limits_{x\to 0} \varepsilon(x) = 0$, on a le D.L. fondamental:

$$\boxed{\mathrm{e}^x = 1 + x + \tfrac{1}{2!}x^2 + \tfrac{1}{3!}x^3 + \cdots + \tfrac{1}{n!}x^n + x^n\varepsilon(x)}$$

d'où l'on tire:

$$\operatorname{ch} x = 1 + \tfrac{1}{2!}x^2 + \tfrac{1}{4!}x^4 + \cdots + \tfrac{1}{(2n)!}x^{2n} + x^{2n+1}\varepsilon(x)$$

$$\operatorname{sh} x = x + \tfrac{1}{3!}x^3 + \tfrac{1}{5!}x^5 + \cdots + \tfrac{1}{(2n+1)!}x^{2n+1} + x^{2n+2}\varepsilon(x)$$

$$\cos x = 1 - \tfrac{1}{2!}x^2 + \tfrac{1}{4!}x^4 + \cdots + (-1)^n\tfrac{1}{(2n)!}x^{2n} + x^{2n+1}\varepsilon(x)$$

$$\sin x = x - \tfrac{1}{3!}x^3 + \tfrac{1}{5!}x^5 + \cdots + (-1)^n\tfrac{1}{(2n+1)!}x^{2n+1} + x^{2n+2}\varepsilon(x)$$

De même à partir du D.L. fondamental:

$$\boxed{(1+x)^\alpha = 1 + \alpha x + \tfrac{\alpha(\alpha-1)}{2!}x^2 + \tfrac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots + \tfrac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)}{n!}x^n + x^n\varepsilon(x)}$$

on obtient:

$$\tfrac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + x^n\varepsilon(x)$$

$$\tfrac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + x^n\varepsilon(x)$$

$$\tfrac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n} + x^{2n+1}n\varepsilon(x)$$

$$\tfrac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 + \cdots + x^{2n} + x^{2n+1}n\varepsilon(x)$$

$$\ln(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \cdots + (-1)^n\tfrac{1}{n}x^n + x^n\varepsilon(x)$$

$$\arctan x = x - \tfrac{1}{3}x^3 + \tfrac{1}{5}x^5 - \tfrac{1}{7}x^7 + \cdots + (-1)^n\tfrac{1}{2n+1}x^{2n+1} + x^{2n+2}\varepsilon(x)$$

$$\operatorname{argth} x = x + \tfrac{1}{3}x^3 + \tfrac{1}{5}x^5 + \tfrac{1}{7}x^7 + \cdots + \tfrac{1}{2n+1}x^{2n+1} + x^{2n+2}\varepsilon(x)$$

Enfin, sans avoir de forme simple pour un D.L. à l'ordre n quelconque, on a au voisinage de 0 et à l'ordre 8:

$$\tan x = x + \tfrac{1}{3}x^3 + \tfrac{2}{15}x^5 + \tfrac{17}{315}x^7 + x^8\varepsilon(x)$$

$$\operatorname{th} x = x - \tfrac{1}{3}x^3 + \tfrac{2}{15}x^5 - \tfrac{17}{315}x^7 + x^8\varepsilon(x)$$

2.2 Somme et produit de D.L.

Proposition. *Si f et g admettent toutes les deux un D.L. à l'ordre n au voisinage de 0, alors:*

(i) *la fonction $f+g$ admet un D.L. à l'ordre n au voisinage de 0 dont la partie régulière est la somme des parties régulières des D.L. de f et g.*

(ii) *la fonction fg admet un D.L. à l'ordre n au voisinage de 0 dont la partie régulière est la somme des termes de degré $\leq n$ dans le produit des parties régulières des D.L. de f et g.*

Exemple. Calculons un D.L. à l'ordre 5 au voisinage de 0 de $f(x) = \sin x(\operatorname{ch} x - \frac{1}{1+x})$

On a d'abord: $\operatorname{ch} x - \frac{1}{1+x} = (1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4) - (1 - x + x^2 - x^3 + x^4 - x^5) + x^5\varepsilon(x)$
$= x - \frac{1}{2}x^2 + x^3 - \frac{23}{24}x^4 + x^5 + x^5\varepsilon(x)$.

Puis: $f(x) = (x - \frac{1}{6}x^3 + \frac{1}{120}x^5 + x^6\varepsilon(x))(x - \frac{1}{2}x^2 + x^3 - \frac{23}{24}x^4 + x^5 + x^5\varepsilon(x))$
$= x^2 - \frac{1}{2}x^3 + (-\frac{1}{6} + 1)x^4 + (-\frac{23}{24} + \frac{1}{12})x^5 + x^5\varepsilon(x)$
$= x^2 - \frac{1}{2}x^3 + \frac{5}{6}x^4 - \frac{7}{8}x^5 + x^5\varepsilon(x)$.

2.3 Composition de D.L.

Proposition. *Si f et g admettent toutes les deux un D.L. à l'ordre n au voisinage de 0, et si $f(0)=0$, alors la fonction $g\circ f$ composée de g par f admet un D.L. à l'ordre n au voisinage de 0, dont la partie régulière s'obtient en substituant la partie régulière du D.L. de f dans la partie régulière du D.L. de g, et en ne conservant que les termes de degré $\leq n$.*

Exemple. Calculons un D.L. à l'ordre 4 au voisinage de 0 de $h(x)=\mathrm{e}^{\sin x}$.

On a $h=g\circ f$ où $\begin{cases} f(x)=\sin x = x-\frac{1}{6}x^3+x^4\varepsilon(x) \text{ vérifie bien } f(0)=0,\\ g(x)=\mathrm{e}^x=1+x+\frac{1}{2}x^2+\frac{1}{6}x^3+\frac{1}{24}x^4+x^4\varepsilon(x).\end{cases}$

$$\begin{aligned}\text{donc: } h(x) &= 1+(x-\tfrac{1}{6}x^3)+\tfrac{1}{2}(x-\tfrac{1}{6}x^3)^2+\tfrac{1}{6}(x-\tfrac{1}{6}x^3)^3+\tfrac{1}{24}(x-\tfrac{1}{6}x^3)^4+x^4\varepsilon(x)\\ &= 1+(x-\tfrac{1}{6}x^3)+\tfrac{1}{2}(x^2-\tfrac{2}{6}x^4+\tfrac{1}{36}x^6)+\tfrac{1}{6}(x^3-\tfrac{3}{6}x^5+\cdots)+\tfrac{1}{24}(x^4+\cdots)+x^4\varepsilon(x)\\ &= 1+x+\tfrac{1}{2}x^2-\tfrac{1}{8}x^4+x^4\varepsilon(x).\end{aligned}$$

2.4 Intégration de D.L.

Proposition. *Si f est continue sur un intervalle I contenant 0 et admet un D.L. à l'ordre n au voisinage de 0, alors toute primitive F de f sur I admet un D.L. à l'ordre $n+1$ au voisinage de 0, dont la partie régulière s'obtient en intégrant terme à terme la partie régulière du D.L. de f.*

Exemple. Calculons un D.L. à l'ordre 7 au voisinage de 0 de $F(x)=\arcsin x$.

On a $F'(x)=f(x)=\frac{1}{\sqrt{1-x^2}}=(1-x^2)^{-1/2}$. On en obtient un D.L. avec la formule donnant un D.L. pour $(1+x)^\alpha$ en prenant $\alpha=-\frac{1}{2}$ et en remplaçant x par $-x^2$. On obtient:

$$\begin{aligned} f(x) &= 1+(-\tfrac{1}{2})(-x^2)+\tfrac{1}{2}(-\tfrac{1}{2})(-\tfrac{3}{2})(-x^2)^2+\tfrac{1}{6}(-\tfrac{1}{2})(-\tfrac{3}{2})(-\tfrac{5}{2})(-x^2)^3+x^6\varepsilon(x)\\ &= 1+\tfrac{1}{2}x^2+\tfrac{3}{8}x^4+\tfrac{5}{16}x^6+x^6\varepsilon(x).\end{aligned}$$

On intègre terme à terme pour conclure: $F(x)=x+\frac{1}{6}x^3+\frac{3}{40}x^5+\frac{5}{112}x^7+x^7\varepsilon(x)+K$,

avec K constante réelle. Mais comme on sait que $\arcsin 0=0$, on a forcément $K=0$. On conclut donc finalement: $\arcsin x = x+\frac{1}{6}x^3+\frac{3}{40}x^5+\frac{5}{112}x^7+x^7\varepsilon(x)$.

2.5 Quotient de D.L.

Proposition. *Si f et g admettent toutes les deux un D.L. à l'ordre n au voisinage de 0, et si $g(0)\neq 0$, alors la fonction f/g quotient de f par g admet un D.L. à l'ordre n au voisinage de 0, dont la partie régulière est le quotient à l'ordre n dans la division suivant les puissances croissantes de x de la partie régulière du D.L. de f par la partie régulière du D.L. de g.*

Exemple. Calculons un D.L. à l'ordre 5 au voisinage de 0 de $h(x)=\frac{1-\cos(2x)}{x\,\mathrm{sh}\,x}$.

D'une part: $1-\cos(2x)=\frac{1}{2!}(2x)^2-\frac{1}{4!}(2x)^4+\frac{1}{6!}(2x)^6+x^7\varepsilon(x)=x^2\big(2-\frac{2}{3}x^2+\frac{4}{45}x^4+x^5\varepsilon(x)\big)$,

d'autre part: $x\,\mathrm{sh}\,x = x\big(x+\frac{1}{3!}x^3+\frac{1}{5!}x^5+x^6\varepsilon(x)\big)=x^2\big(1+\frac{1}{6}x^2+\frac{1}{120}x^4+x^5\varepsilon(x)\big)$.

Donc $h(x)=\frac{f(x)}{g(x)}$, où l'on a posé: $\begin{cases} f(x)=\frac{1-\cos(2x)}{x^2}=2-\frac{2}{3}x^2+\frac{4}{45}x^4+x^5\varepsilon(x),\\ g(x)=\frac{\mathrm{sh}\,x}{x}=1+\frac{1}{6}x^2+\frac{1}{120}x^4+x^5\varepsilon(x).\end{cases}$

On effectue la division suivant les puissances croissantes de x de la partie régulière de $f(x)$ par la partie régulière de $g(x)$:

$$\begin{array}{r|l} 2-\frac{2}{3}x^2+\frac{4}{45}x^4 & 1+\frac{1}{6}x^2+\frac{1}{120}x^4 \\ \cline{2-2} \underline{2+\frac{1}{3}x^2+\frac{1}{60}x^4} & 2-x^2+\frac{43}{180}x^4 \\ -x^2+\frac{13}{180}x^4 & \\ \underline{-x^2-\frac{1}{6}x^4} & \\ \frac{43}{180}x^4 & \end{array}$$

Comme h est paire et que donc seuls des termes d'exposant pairs apparaissent dans le D.L., on conclut qu'un D.L. de $h(x)$ à l'ordre 5 au voisinage de 0 est:

$$h(x)=2-x^2+\frac{43}{180}x^4+x^5\varepsilon(x).$$

Remarque pratique. Effectuer une telle division est souvent fastidieux; aussi est-il nécessaire, avant de se lancer dans ce type de calcul, de s'assurer qu'une autre méthode (par exemple une composition avec $\frac{1}{1\pm x}$) ne permet pas d'arriver plus simplement au résultat.

Exemple. Calculons un D.L. à l'ordre 4 au voisinage de 0 de $f(x) = \frac{x}{\sin x}$.

On pense a priori à un quotient (il est tout à fait possible de faire la division suivant les puissances croissantes de x pour parvenir au résultat), mais on peut aussi procéder de la manière suivante.

On a: $\sin x = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 + x^5\varepsilon(x)$. Posons: $u = \frac{1}{x}(\sin x - x) = -\frac{1}{6}x^2 + \frac{1}{120}x^4 + x^4\varepsilon(x)$.

Il est clair que u tend vers 0 quand x tend vers 0. De plus, par définition de u, on a:

$$f(x) = \frac{x}{\sin x} = \frac{x}{xu+x} = \frac{1}{1+u}.$$

Or $\frac{1}{1+u} = 1 - u + u^2 + u^2\varepsilon(u)$, et on peut substituer (en appliquant la proposition de 2.3):

$$\begin{aligned} f(x) &= 1 - (-\tfrac{1}{6}x^2 + \tfrac{1}{120}x^4) + (-\tfrac{1}{6}x^2 + \tfrac{1}{120}x^4)^2 + x^4\varepsilon(x) \\ &= 1 + \tfrac{1}{6}x^2 + (-\tfrac{1}{120} + \tfrac{1}{36})x^4 + x^4\varepsilon(x) \\ &= 1 + \tfrac{1}{6}x^2 + \tfrac{7}{360}x^4 + x^4\varepsilon(x). \end{aligned}$$

2.6 Applications des D.L.

Un D.L. permet de donner une approximation d'une fonction par un polynôme, ce qui est précieux dans bien des situations: calculs de limites, étude locale d'une courbe, étude des branches infinies, position d'une courbe par rapport à ses asymptotes,... Sans entrer dans les détails, donnons quelques exemples de calculs de limite où les D.L. permettent de "lever une indétermination".

Exemple. Calculons $\lim\limits_{x\to 0} \frac{1-\cos x}{\tan^2 x}$.

A priori, on a une forme indéterminée car le numérateur et le dénominateur tendent tous les deux vers 0. Mais au voisinage de 0, on a $\cos x = 1 - \frac{1}{2}x^2 + x^3\varepsilon(x)$ et $\tan x = x + \frac{1}{3}x^3 + x^4\varepsilon'(x)$. Donc:

$$\frac{1-\cos x}{\tan^2 x} = \frac{\frac{1}{2}x^2 + x^3\varepsilon(x)}{x^2 + x^3\varepsilon'(x)} = \frac{\frac{1}{2} + x\varepsilon(x)}{1 + x\varepsilon'(x)},$$

et comme $\lim\limits_{x\to 0}\varepsilon(x) = \lim\limits_{x\to 0}\varepsilon'(x) = 0$, on conclut que $\lim\limits_{x\to 0} \frac{1-\cos x}{\tan^2 x} = \frac{1}{2}$.

Exemple. Calculons $\lim\limits_{x\to +\infty} \left[x - x^2\ln(1+\frac{1}{x})\right]$.

A priori, on a une forme indéterminée car chacun des deux termes de la différence tend vers $+\infty$. Mais au voisinage de $+\infty$, on a:

$$\ln(1+\tfrac{1}{x}) = \tfrac{1}{x} - \tfrac{1}{2x^2} + \tfrac{1}{3x^3} + \tfrac{1}{x^3}\varepsilon(\tfrac{1}{x}), \quad \text{avec } \lim_{x\to+\infty}\varepsilon(\tfrac{1}{x}) = 0,$$

car $\ln(1+u) = u - \frac{1}{2}u^2 + \frac{1}{3}u^3 + u^3\varepsilon(u)$ pour $u = \frac{1}{x}$ au voisinage de 0. Donc:

$$x^2\ln(1+\tfrac{1}{x}) = x - \tfrac{1}{2} + \tfrac{1}{3x} + \tfrac{1}{x}\varepsilon(\tfrac{1}{x}), \quad \text{avec } \lim_{x\to+\infty}\varepsilon(\tfrac{1}{x}) = 0,$$

d'où l'on tire que $\lim\limits_{x\to+\infty}\left[x - x^2\ln(1+\frac{1}{x})\right] = \lim\limits_{x\to+\infty}\left[\frac{1}{2} - \frac{1}{3x} - \frac{1}{x}\varepsilon(\frac{1}{x})\right] = \frac{1}{2}$

Notons pour finir que plusieurs limites "classiques" (et à connaître !) découle immédiatement de calculs avec les D.L. C'est le cas entre autres de:

$$\lim_{x\to 0}\frac{\sin x}{x} = 1, \qquad \lim_{x\to 0}\frac{\cos x - 1}{x} = 0, \qquad \lim_{x\to 0}\frac{\mathrm{e}^x - 1}{x} = 1$$

Leçon 4

Intégrales et primitives

1. Intégrale d'une fonction continue par morceaux sur un intervalle fermé borné

1.1 Principe de la définition de l'intégrale

On fixe un intervalle fermé borné $I = [a, b]$ de $\mathbb{R}$, avec $a \leq b$ deux réels.
Une *subdivision* de I est une famille finie $(x_0, x_1, x_2, \dots, x_n)$ de réels de I tels que:

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Première étape: intégrale d'une fonction en escalier. Une fonction $f : I \to \mathbb{R}$ est dite *en escalier* sur I s'il existe une subdivision $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ de I telle que f soit constante sur chacun des intervalles ouverts $]x_{i-1}, x_i[$, pour tout $1 \leq i \leq n$.

Si l'on appelle α_i la valeur constante prise par f sur $]x_{i-1}, x_i[$, on définit *l'intégrale* $\int_I f$ comme le réel:

$$\int_I f = \sum_{i=1}^{n} (x_i - x_{i-1})\alpha_i.$$

Figure 1

Par définition, ce réel $\int_I f$ mesure l'aire limitée par la courbe représentative de f et l'axe des abscisses, en comptant positivement ce qui est au dessus de l'axe, et négativement ce qui est en-dessous.

Seconde étape: intégrale d'une fonction continue par morceaux. Une fonction $f : I \to \mathbb{R}$ est dite *continue par morcaux* sur I s'il existe une subdivision $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ de I telle que f soit continue sur chacun des intervalles ouverts $]x_{i-1}, x_i[$, pour tout $1 \leq i \leq n$, et telle que f admette une limite à gauche et une limite à droite (par forcément égales) en chacun des points de la subdivision.

Figure 2

Il est clair que les fonctions continues sur I et les fonctions en escalier sur I sont des exemples de fonctions continues par morceaux sur I.

On montre (et nous admettons) que toute fonction f continue par morceaux sur I peut être "approchée uniformément" (notion que nous ne pouvons pas détailler ici) par une suite (f_n) de fonctions en escalier, et que la limite de la suite de réels $(\int_I f_n)$ est alors convergente. C'est cette limite, dont on montre qu'elle ne dépend pas de la suite (f_n) choisie, que l'on appelle *l'intégrale* de f sur I.

On la note $\int_I f$, ou encore $\int_a^b f(t)\,dt$.

Figure 3

Remarque 1: le réel $\int_a^b f(t)\,dt$ mesure l'aire limitée par la courbe représentative de f et l'axe des abscisses, en comptant positivement ce qui est au dessus de l'axe, et négativement ce qui est en-dessous.

Remarque 2: le réel $\int_a^b f(t)\,dt$ ne dépend pas des valeurs prises par f aux éventuels points de discontinuité (qui sont de toute façon en nombre fini).

1.2 Principales propriétés.

(i) Linéarité: si f et g sont deux fonctions continues par morceaux sur $I = [a, b]$, et si λ est un réel fixé, alors on a:

$$\int_a^b (f+g)(t)\,dt = \int_a^b f(t)\,dt + \int_a^b g(t)\,dt \quad \text{et} \quad \int_a^b \lambda f(t)\,dt = \lambda(\int_a^b f(t)\,dt).$$

(ii) Egalité presque partout: si f et g sont deux fonctions continues par morceaux sur $I = [a, b]$ telles que $f(t) = g(t)$ pour tout $t \in I$ sauf éventuellement un nombre fini, alors $\int_a^b f(t)\,dt = \int_a^b g(t)\,dt$.
En d'autres termes, on ne change pas la valeur de l'intégrale si on modifie la valeur prise par f en un nombre fini de points de I.

(iii) Relation de Chasles: si f est une fonction continue par morceaux sur $[a, b]$ et sur $[b, c]$, alors f est continue par morceaux sur $[a, c]$ et l'on a:

$$\int_a^b f(t)\,dt + \int_b^c f(t)\,dt = \int_a^c f(t)\,dt.$$

(iv) Positivité: si f est une fonction continue par morceaux sur $I = [a, b]$, et positive sur I (sauf éventuellement en un nombre fini de points), alors $\int_a^b f(t)\,dt \geq 0$.

Il en résulte bien sûr que, si $f(t) \geq g(t)$, alors $\int_a^b f(t)\,dt \geq \int_a^b g(t)\,dt$

1.3 Lien avec les primitives.

D'après les propriétés (ii) et (iii) ci-dessus, toute intégrale d'une fonction continue par morceaux sur I se ramène à une somme d'intégrales de fonctions continues sur des sous-intervalles de I.

En effet: soit f continue par morceaux sur $I = [a, b]$. Soit $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ une subdivision de I telle que f soit continue sur chaque intervalle ouvert $]x_{i-1}, x_i[$ et admette une limite à gauche et une limite à droite en chacun des x_i. Pour tout $1 \leq i \leq n$, on définit une fonction $f_i : [x_{i-1}, x_i] \to \mathbb{R}$ en posant:

$$f_i(t) = f(t) \text{ si } t \in]x_{i-1}, x_i[, \qquad f_i(x_{i-1}) = \lim_{t \to x_{i-1}, t > x_{i-1}} f(t) \quad \text{et} \quad f_i(x_i) = \lim_{t \to x_i, t < x_i} f(t).$$

Par construction, f_i est continue sur $[x_{i-1}, x_i]$ et coïncide avec f sur $]x_{i-1}, x_i[$. Donc $\int_{x_{i-1}}^{x_i} f_i(t)\,dt = \int_{x_{i-1}}^{x_i} f(t)\,dt$ d'après la propriété d'égalité presque partout. Et donc, d'après la relation de Chasles:

$$\int_a^b f(t)\,dt = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(t)\,dt = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f_i(t)\,dt.$$

Il suffit donc de savoir calculer l'intégrale d'une fonction continue. Sous cette hypothèse, la notion d'intégrale est intimement liée à celle de primitive. Tout repose sur le résultat suivant (que nous admettons).

Théorème fondamental. *Soit $f : [a, b] \to \mathbb{R}$ continue sur $[a, b]$.*

(i) *En notant* $\boxed{F_a(x) = \int_a^x f(t)\,dt}$ *pour tout $x \in [a, b]$, l'application $F_a : [a, b] \to \mathbb{R}$ est une primitive de f, ce qui signifie par définition que:*

$$F_a \text{ est dérivable sur } [a, b], \text{ et } \boxed{F_a'(x) = f(x)} \text{ pour tout } x \in [a, b].$$

(ii) *Les primitives F de f sur $[a, b]$ sont toutes les fonctions F de la forme $F_a + \lambda$, avec $\lambda \in \mathbb{R}$.*

(iii) *La fonction F_a est la primitive de f sur $[a, b]$ qui s'annule en a.*

(iv) *Pour toute primitive F de f sur $[a, b]$, on a* $\boxed{\int_a^b f(t)\,dt = F(b) - F(a)}$, *que l'on note: $[F(x)]_a^b$.*

(v) *En particulier, $\int_b^a f(t)\,dt = -\int_a^b f(t)\,dt$.*

On verra à la section suivante comment on calcule techniquement les primitives, et donc les intégrales.

1.4 Cas des fonctions à valeurs complexes.

On prend toujours $I = [a, b]$, mais on suppose maintenant que $f : I \to \mathbb{C}$ est à valeurs complexes. On définit classiquement la fonction *partie réelle* $f_1 : I \to \mathbb{R}$ et la fonction *partie imaginaire* $f_2 : I \to \mathbb{R}$ de f par:

$$f(t) = f_1(t) + i f_2(t) \text{ pour tout } t \in I.$$

On peut montrer aisément que f est continue par morceaux sur I si et seulement si f_1 et f_2 le sont, de sorte que l'on définit naturellement:

$$\int_a^b f(t)\,dt = \int_a^b f_1(t)\,dt + i \int_a^b f_2(t)\,dt.$$

En outre, si f est continue par morceaux sur I, il en est de même de la fonction $|f|$, (qui désigne le module de f, et donc la valeur absolue dans le cas particulier où f est à valeurs réelles), et l'on a:

$$\boxed{\left| \int_a^b f(t)\,dt \right| \leq \int_a^b |f(t)|\,dt}.$$

Au niveau élémentaire, le calcul des intégrales via celui des primitives repose sur deux arguments:

- on connaît certaines primitives de fonctions usuelles ou de référence (parce qu'on sait reconnaître ces fonctions comme des dérivées d'autres fonctions connues);
- on se ramène à ces primitives de référence en les combinant par linéarité, par intégration par parties, et par changement de variables (voir ci-dessous).

2.1 Quelques primitives usuelles.

On rappelle dans le premier tableau ci-dessous les quelques primitives les plus fondamentales (à connaître absolument). La fonction F est, pour chaque ligne, UNE primitive de la fonction f donnée. On n'oubliera pas les constantes additives pour obtenir les autres primitives.

données	$f(x)$	$F(x)$	intervalle de validité
$n \in \mathbb{N}$	x^n	$\frac{1}{n+1}x^{n+1}$	$\mathbb{R}$
$\alpha \in \mathbb{R}, \alpha \neq -1$	x^α	$\frac{1}{\alpha+1}x^{\alpha+1}$	$\mathbb{R}_+^*$
	$\frac{1}{x}$	$\ln x$	$\mathbb{R}_+^*$
	e^x	e^x	$\mathbb{R}$
	$\cos x$	$\sin x$	$\mathbb{R}$
	$\sin x$	$-\cos x$	$\mathbb{R}$
	$\operatorname{ch} x$	$\operatorname{sh} x$	$\mathbb{R}$
	$\operatorname{sh} x$	$\operatorname{ch} x$	$\mathbb{R}$

D'autres primitives usuelles sont données dans le second tableau ci-dessous.

$f(x)$	$F(x)$	intervalle de validité
$\frac{1}{\cos^2 x} = 1 + \tan^2 x$	$\tan x$	$\left](2p-1)\frac{\pi}{2}, (2p+1)\frac{\pi}{2}\right[$, où $p \in \mathbb{Z}$
$\frac{1}{\sin^2 x} = 1 + \operatorname{cotan}^2 x$	$-\operatorname{cotan} x$	$]p\pi, (p+1)\pi[$, où $p \in \mathbb{Z}$
$\frac{1}{1+x^2}$	$\arctan x$	$\mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$]-1, 1[$
$\frac{1}{\sqrt{x^2+1}}$	$\ln(x + \sqrt{x^2+1}) = \operatorname{argsh} x$	$\mathbb{R}$
$\frac{1}{\sqrt{x^2-1}}$	$\ln\lvert x + \sqrt{x^2-1}\rvert = \operatorname{argch} x$	$]1, +\infty[$
$\frac{1}{1-x^2}$	$\frac{1}{2}\ln\lvert\frac{1+x}{1-x}\rvert = \operatorname{argth} x$	$]-1, 1[$

EXEMPLES.

$$\int_0^1 \frac{1}{1+t^2}\,dt = [\arctan x]_0^1 = \arctan 1 - \arctan 0 = \frac{\pi}{4}.$$

$$\int_2^3 (5t^2 + 2t - 1)\,dt = [\frac{5}{3}x^3 + x^2 - x]_2^3 = 51 - \frac{46}{3} = \frac{107}{3}.$$

$$\int_0^{\pi/2} \sin x = [-\cos x]_0^{\pi/2} = 1.$$

2.3 Changement de variables.

THÉORÈME.
Pour toute fonction $\varphi : [a,b] \to \mathbb{R}$ de classe C^1 et toute fonction continue $f : [\varphi(a), \varphi(b)] \to \mathbb{R}$, on a:

$$\boxed{\int_a^b f[\varphi(t)]\,\varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(u)\,du}.$$

EXEMPLE 1. Calculons $\displaystyle\int_{-3/2}^{\sqrt{5}} \frac{2t}{\sqrt{t^2+4}}\,dt$.

On pose $u = \varphi(t) = t^2+4$, donc $du = \varphi'(t)\,dt = 2t\,dt$. Par ailleurs, $\varphi(-\frac{3}{2}) = \frac{25}{4}$ et $\varphi(\sqrt{5}) = 9$.

D'où: $\displaystyle\int_{-3/2}^{\sqrt{5}} \frac{2t}{\sqrt{t^2+4}}\,dt = \int_{25/4}^{9} \frac{1}{\sqrt{u}}\,du = \int_{25/4}^{9} u^{-1/2}\,du = [2\,x^{1/2}]_{25/4}^{9} = [2\sqrt{x}]_{25/4}^{9} = 1.$

EXEMPLE 2. Calculons $\displaystyle\int_0^1 \sqrt{1-t^2}\,dt$.

On pose $u = \varphi(t) = \arcsin t$, c'est-à-dire $t = \sin u$, donc $dt = \cos u\,du$. Par ailleurs, $\varphi(0) = 0$ et $\varphi(1) = \frac{\pi}{2}$.

D'où: $\displaystyle\int_0^1 \sqrt{1-t^2}\,dt = \int_0^{\pi/2} \sqrt{1-\sin^2 u}\,\cos u\,du = \int_0^{\pi/2} \cos^2 u\,du.$

Pour achever le calcul, on est alors conduit à linéariser le $\cos^2$, c'est-à-dire à utiliser la formule de trigonométrie $\cos 2a = 2\cos^2 a - 1$ pour écrire:

$\displaystyle\int_0^{\pi/2} \cos^2 u\,du = \frac{1}{2}\int_0^{\pi/2} (1+\cos 2u)\,du = \frac{1}{2}[x + \frac{1}{2}\sin 2x]_0^{\pi/2} = \frac{\pi}{4}.$

2.4 Intégration par parties.

THÉORÈME. *Soient u et v deux fonctions $[a,b] \to \mathbb{R}$ continues sur $[a,b]$ et de classe C^1 par morceaux sur $[a,b]$; alors on a:*

$$\boxed{\int_a^b u'(t)v(t)\,dt = \big[u(t)v(t)\big]_a^b - \int_a^b u(t)v'(t)\,dt}.$$

EXEMPLE 1. Calculons $\displaystyle\int_0^\pi t\cos t\,dt$. On pose $u = t$ donc $u' = 1$, et $dv = \cos t\,dt$ donc $v = \sin t$.

On obtient alors $\displaystyle\int_0^\pi t\cos t\,dt = \big[t\sin t\big]_0^\pi - \int_0^\pi \sin t\,dt = 0 - \big[-\cos t\big]_0^\pi = -2.$

EXEMPLE 2. Calculons $\displaystyle\int_1^{\mathrm{e}} \ln t\,dt$. On pose $u = \ln t$ donc $u' = \frac{1}{t}\,dt$, et $dv = dt$ donc $v = t$.

On obtient alors $\displaystyle\int_1^{\mathrm{e}} \ln t\,dt = \big[t\ln t\big]_1^{\mathrm{e}} - \int_1^{\mathrm{e}} dt = \mathrm{e} - (\mathrm{e}-1) = 1.$

EXEMPLE 3. Calculons $X = \displaystyle\int_0^{\pi/2} \mathrm{e}^{-2t}\cos t\,dt$.

On pose $u = \mathrm{e}^{-2t}$ donc $u' = -2\,\mathrm{e}^{-2t}\,dt$, et $dv = \cos t\,dt$ donc $v = \sin t$.

On obtient alors $X = \big[\mathrm{e}^{-2t}\sin t\big]_0^{\pi/2} + 2\displaystyle\int_0^{\pi/2} \mathrm{e}^{-2t}\sin t\,dt = \mathrm{e}^{-\pi} + 2\int_0^{\pi/2} \mathrm{e}^{-2t}\sin t\,dt.$

On réapplique la formule avec: $u = \mathrm{e}^{-2t}$ donc $u' = -2\,\mathrm{e}^{-2t}\,dt$, et $dv = \sin t\,dt$ donc $v = -\cos t$.

Donc: $\displaystyle\int_0^{\pi/2} \mathrm{e}^{-2t}\sin t\,dt = \big[-\mathrm{e}^{-2t}\cos t\big]_0^{\pi/2} - 2\int_0^{\pi/2} \mathrm{e}^{-2t}\cos t\,dt = 1 - 2X.$

On a ainsi: $X = \mathrm{e}^{-\pi} + 2(1-2X)$ d'où $X = \frac{1}{5}(\mathrm{e}^{-\pi}+2)$.

Equations différentielles linéaires

1. EQUATIONS DIFFÉRENTIELLES LINÉAIRES DU PREMIER ORDRE

1.1 Données et terminologie.

• Une *équation différentielle linéaire du premier ordre* est une équation différentielle que l'on peut mettre sous la forme:

$$\boxed{y'(x) + a(x)y(x) = b(x),} \qquad (1)$$

où a et b sont deux fonctions fixées d'une variable réelle x définies sur un intervalle I, et y est une fonction de x dérivable sur I. La fonction $b(x)$ est appelée le *second membre* de l'équation. Une fonction y de x dérivable sur I qui satisfait effectivement la relation (1) pour tout $x \in I$ est appelée une *solution* de (1) sur I. *Résoudre* (1) sur I consiste à déterminer toutes ses solutions.

• Par tradition, on dit que l'on a obtenu *la solution générale* de (1) lorsqu'on a déterminé toutes ses solutions. Si on connaît une fonction y_0 qui est solution de (1), on dit que y_0 est *une solution particulière* de (1).

• On appelle équation différentielle *homogène* (ou *sans second membre*) associée à (1) l'équation différentielle:

$$\boxed{y'(x) + a(x)y(x) = 0,} \qquad (2)$$

Par opposition, l'équation (1) de départ est appelée équation *complète* ou *avec second membre.* L'intérêt de l'introduction de (2) résulte de l'observation suivante:

OBSERVATION FONDAMENTALE: *la solution générale de (1) s'obtient en ajoutant une solution particulière de (1) à la solution générale de l'équation homogène (2).*

EN EFFET, soit y_0 une solution particulière de (1) sur un intervalle I. On a donc $y_0'(x) + a(x)y_0(x) = b(x)$ pour tout $x \in I$. Donc pour toute fonction y dérivable sur I, on a:

$$y'(x) + a(x)y(x) = b(x) \;\Leftrightarrow\; y'(x) + a(x)y(x) = y_0'(x) + a(x)y_0(x) \;\Leftrightarrow\; (y - y_0)'(x) + a(x)(y - y_0)(x) = 0,$$

ce qui prouve que y est solution de (1) si et seulement si $y - y_0$ est solution de (2).

1.2 Solution générale de l'équation homogène.

PROPOSITION:*la solution générale de l'équation homogène* (2) *est donnée par:*

$$y(x) = C\,\mathrm{e}^{-A(x)},$$

où $A(x)$ est une primitive quelconque de $a(x)$ et C une constante quelconque décrivant $\mathbb{R}$.

EN EFFET, sur tout intervalle I où $y(x)$ ne s'annule pas, (2) équivaut à: $\frac{y'(x)}{y(x)} = -a(x)$.

Donc la solution générale de (2) sur I vérifie $\ln|y(x)| = -A(x) + k$ avec $A(x)$ primitive de $a(x)$ et k constante quelconque dans $\mathbb{R}$. Ou encore $|y(x)| = \mathrm{e}^k\,\mathrm{e}^{-A(x)}$. Comme par hypothèse $y(x)$ ne change pas de signe sur I, on obtient le résultat voulu avec $C = \mathrm{e}^k$ si $y(x) \geq 0$ et $C = -\mathrm{e}^k$ si $y(x) \leq 0$.

EXEMPLES.

La solution générale sur $I =]0, +\infty[$ de $y' - \frac{1}{x}y = 0$ est $y(x) = C\,\mathrm{e}^{\ln x} = Cx$, avec $C \in \mathbb{R}$ quelconque.

La solution générale sur $I =]0, +\infty[$de $y' + \frac{1}{x}y = 0$ est $y(x) = C\,\mathrm{e}^{-\ln x} = \frac{C}{x}$, avec $C \in \mathbb{R}$ quelconque.

La solution générale sur $I =]0, +\infty[$de $y' - \frac{x+1}{x}y = 0$ est $y(x) = C\,\mathrm{e}^{x+\ln x} = Cx\,\mathrm{e}^x$, avec $C \in \mathbb{R}$ quelconque.

1.3 Solution particulière de l'équation avec second membre.

• *Première méthode:* détermination directe d'une solution plus ou moins évidente.

EXEMPLE. Résoudre sur $I = \mathbb{R}$ l'équation (1) suivante: $y' - 2y = 3$. L'équation homogène associée est $y' - 2y = 0$, dont la solution générale est $C\,\mathrm{e}^{2x}$, avec $C \in \mathbb{R}$. De plus, il est clair que la fonction constante $y = -\frac{3}{2}$ est une solution particulière de (1). On conclut que la solution générale de (1) est $y(x) = C\,\mathrm{e}^{2x} - \frac{3}{2}$, avec $C \in \mathbb{R}$.

• *Seconde méthode:* procédé systématique dit de la “variation de la constante” (encore un héritage terminologique de l'histoire !). En s'appuyant sur la forme de la solution générale de (2), la méthode consiste à rechercher une solution particulière de (1) qui soit de la forme:

$$y(x) = z(x)\,\mathrm{e}^{-A(x)},$$

avec toujours $A(x)$ une primitive de $a(x)$, mais z qui est cette fois une fonction de x. On a:

$$y'(x) = z'(x)\,\mathrm{e}^{-A(x)} - z(x)a(x)\,\mathrm{e}^{-A(x)}.$$

Dire qu'une telle $y(x)$ est solution de (1) équivaut donc à:

$$z'(x)\,\mathrm{e}^{-A(x)} - z(x)a(x)\,\mathrm{e}^{-A(x)} + a(x)z(x)\,\mathrm{e}^{-A(x)} = b(x).$$

Le second et le troisième termes se simplifient (c'est le but de la méthode !) et l'on obtient $z'(x) = b(x)\,\mathrm{e}^{A(x)}$. On détermine alors $z(x)$ par calcul d'une primitive de $b(x)\,\mathrm{e}^{A(x)}$, et on reporte cette fonction $z(x)$ trouvée dans $y(x) = z(x)\,\mathrm{e}^{-A(x)}$ pour obtenir une solution particulière de (1).

Exemple. Résoudre sur $I =]0, +\infty[$ l'équation (1) suivante: $y' - \frac{1}{x}y = \frac{1}{x}\,\mathrm{e}^{-\frac{1}{x}}$.

L'équation homogène associée est $y' - \frac{1}{x}y = 0$, dont la solution générale est Cx, avec $C \in \mathbb{R}$. On cherche alors une solution particulière de (1) sous la forme $y(x) = z(x)x$. Elle vérifie $y'(x) = z'(x)x + z(x)$, donc est solution de (1) si et seulement si $z'(x)x + z(x) - \frac{1}{x}z(x)x = \frac{1}{x}\,\mathrm{e}^{-\frac{1}{x}}$. On en tire $z'(x) = \frac{1}{x^2}\,\mathrm{e}^{-\frac{1}{x}}$, dont une primitive est $z(x) = \mathrm{e}^{-\frac{1}{x}}$. On conclut qu'une solution particulière de (1) est $y(x) = x\,\mathrm{e}^{-\frac{1}{x}}$, et donc la solution générale de (1) est $y(x) = Cx + x\,\mathrm{e}^{-\frac{1}{x}}$, avec $C \in \mathbb{R}$.

2. Equations différentielles linéaires du deuxième ordre à coefficients constants

On se limite ici au cas particulier où les fonctions coefficients des dérivées successives de la fonction inconnue y sont des constantes.

2.1 Données et terminologie.

• Une *équation différentielle linéaire du deuxième ordre* à coefficients constants est une équation différentielle que l'on peut mettre sous la forme:

$$\boxed{ay''(x) + by'(x) + cy(x) = d(x),} \qquad (1)$$

où a, b, c sont des réels fixés (des constantes, pas des fonctions) tels que $a \neq 0$, le second membre $d(x)$ est une fonction de x définie sur un intervalle I, et y est une fonction de x deux fois dérivable sur I. Comme dans le cas précédent, on parle de solution générale et de solution particulière de l'équation.

• On appelle équation différentielle *homogène* (ou *sans second membre*) associée à (1) l'équation différentielle:

$$\boxed{ay''(x) + by'(x) + cy(x) = 0,} \qquad (2)$$

On a encore (la vérification est analogue à celle faite en 1.1):

Observation fondamentale: *la solution générale de (1) s'obtient en ajoutant une solution particulière de (1) à la solution générale de l'équation homogène (2).*

2.2 Solution générale de l'équation homogène.

Définition: les données étant celles de 2.1, on appelle *équation caractéristique* associée à (1) et à (2) l'équation algébrique de degré 2 à coefficients réels:

$$\boxed{ar^2 + br + c = 0.} \qquad \text{(E)}$$

On introduit classiquement le discriminant de (E):

$$\boxed{\Delta = b^2 - 4ac.}$$

Rappelons que l'on sait résoudre une équation algébrique comme (E):

si $\Delta > 0$, alors (E) admet deux racines réelles distinctes $r_1 = \frac{1}{2a}(-b - \sqrt{\Delta})$ et $r_2 = \frac{1}{2a}(-b + \sqrt{\Delta})$,

si $\Delta = 0$, alors (E) admet une racine réelle double $r_0 = -\frac{b}{2a}$,

si $\Delta < 0$, alors (E) admet deux racines complexes distinctes $r_1 = \frac{1}{2a}(-b - i\sqrt{-\Delta})$ et $r_2 = \frac{1}{2a}(-b + i\sqrt{-\Delta})$, et puisque a, b, c sont réels, elles sont conjuguées, c'est-à-dire de la forme $r_1 = \alpha + i\beta$ et $r_2 = \alpha - i\beta$ avec $\alpha, \beta \in \mathbb{R}$.

L'existence de solutions de (E) dans $\mathbb{R}$ ou $\mathbb{C}$ suivant le signe de Δ détermine alors les solutions de l'équation différentielle homogène (2). C'est l'objet du théorème suivant, que l'on rappelle ici sans démonstration.

Théorème. *Avec les données et hypothèses ci-dessus, trois cas sont possibles:*

(i) *Si* $\Delta > 0$, *alors* (E) *admet deux racines réelles distinctes* r_1 *et* r_2, *et la solution générale de* (2) *est alors:*

$$y(x) = C_1\,\mathrm{e}^{r_1 x} + C_2\,\mathrm{e}^{r_2 x}, \quad \text{avec } C_1, C_2 \in \mathbb{R}.$$

(ii) *Si* $\Delta = 0$, *alors* (E) *admet une racine réelle double* r_0, *et la solution générale de* (2) *est alors:*

$$y(x) = (C_1 x + C_2)\,\mathrm{e}^{r_0 x}, \quad \text{avec } C_1, C_2 \in \mathbb{R}.$$

(iii) *Si* $\Delta < 0$, *alors* (E) *admet deux racines complexes conjuguées* $\alpha + i\beta$ *et* $\alpha - i\beta$, *et la solution générale de* (2) *est alors:*

$$y(x) = \mathrm{e}^{\alpha x}[C_1\cos(\beta x) + C_2\sin(\beta x)], \quad \text{avec } C_1, C_2 \in \mathbb{R}.$$

Remarque: en physique, on préfère souvent donner la solution générale dans le cas (iii) sous la forme:

$$y(x) = A\,\mathrm{e}^{\alpha x}\cos(\beta x + \varphi), \quad \text{avec } A \in \mathbb{R}_+, \varphi \in \mathbb{R}.$$

En effet, comme $\cos(\beta x + \varphi) = \cos(\beta x)\cos\varphi - \sin(\beta x)\sin\varphi$, on passe d'une expression à l'autre de la solution y par les formules:

$$\begin{cases} C_1 = A\cos\varphi \\ C_2 = -A\sin\varphi \end{cases} \qquad \begin{cases} A = \sqrt{C_1^2 + C_2^2} \\ \cos\varphi = \frac{C_1}{\sqrt{C_1^2+C_2^2}} \quad \text{et} \quad \sin\varphi = \frac{-C_2}{\sqrt{C_1^2+C_2^2}} \end{cases}$$

Exemples.

• Considérons l'équation différentielle homogène (2) suivante: $y'' - y = 0$. L'équation caractéristique associée est $r^2 - 1 = 0$, qui admet les deux racines réelles distinctes 1 et -1. La solution générale de (2) est donc $y(x) = C_1\,\mathrm{e}^x + C_2\,\mathrm{e}^{-x}$, avec $C_1, C_2 \in \mathbb{R}$ quelconques.

• Considérons l'équation différentielle homogène (2) suivante: $y'' - 2y' + y = 0$. L'équation caractéristique associée est $r^2 - 2r + 1 = 0$, qui admet une racine réelle double 1. La solution générale de (2) est donc $y(x) = (C_1 x + C_2)\,\mathrm{e}^x$, avec $C_1, C_2 \in \mathbb{R}$ quelconques.

• Considérons l'équation différentielle homogène (2) suivante: $y'' + y = 0$. L'équation caractéristique associée est $r^2 + 1 = 0$, qui admet les deux racines complexes conjuguées i et $-i$. Avec $\alpha = 0$ et $\beta = 1$, on conclut que la solution générale de (2) est donc $y(x) = C_1\cos x + C_2\sin x$, avec $C_1, C_2 \in \mathbb{R}$ quelconques. Elle s'exprime aussi sous la forme $y(x) = A\cos(x + \varphi)$, avec $A \in \mathbb{R}_+$ et $\varphi \in \mathbb{R}$ quelconques.

2.3 Solution particulière de l'équation avec second membre.

Sans donner de méthode générale, on indique ci-dessous quelques cas particuliers fréquents de seconds membres.

• *Si le second membre $d(x)$ est un polynôme*, on cherche une solution particulière $y_0(x)$ de (1) qui soit un polynôme.

> Exemple. Considérons l'équation (1) suivante: $y'' - 2y' + y = x$. On cherche y_0 sous la forme $y_0(x) = \lambda x + \mu$; on a alors $y_0''(x) - 2y_0'(x) + y_0(x) = 0 - 2\lambda + \lambda x + \mu$, de sorte que y_0 est solution de (1) si et seulement si $\lambda = 1$ et $\mu - 2\lambda = 0$. D'où $y_0(x) = x + 2$. Comme on avait déjà déterminé au 2.2 la solution générale de l'équation homogène associée, on conclut que la solution générale de (1) est $y(x) = (C_1 x + C_2)\,\mathrm{e}^x + x + 2$, avec $C_1, C_2 \in \mathbb{R}$ quelconques.

On fera preuve de bon sens dans le choix du degré m du polynôme $y_0(x)$ que l'on cherche. Ce degré m dépend bien sûr du degré n du second membre $d(x)$, mais aussi des coefficients a, b, c du premier membre. Il est facile de voir que l'on peut prendre $m = n$ si $c \neq 0$, $m = n + 1$ si $c = 0$ et $b \neq 0$, et $m = n + 2$ si $c = b = 0$ et $a \neq 0$.

• *Si le second membre $d(x)$ est le produit d'un polynôme* par une exponentielle $e^{\alpha x}$, on cherche une solution particulière $y_0(x)$ de (1) qui soit de la même forme.

> EXEMPLE. Considérons l'équation (1) suivante: $y'' - y = e^x$. On cherche y_0 sous la forme $y_0(x) = p(x)\,e^x$ avec p fonction polynomiale; on a alors:
> $$y_0'(x) = p'(x)\,e^x + p(x)\,e^x \quad \text{et} \quad y_0''(x) = p''(x)\,e^x + 2p'(x)\,e^x + p(x)\,e^x,$$
> donc $y_0''(x) - y_0(x) = [p''(x) + 2p'(x)]\,e^x$, de sorte que y_0 est solution de (1) si et seulement si $p''(x) + 2p'(x) = 1$. On peut trouver un polynôme p solution de $p''(x) + 2p'(x) = 1$ qui soit de degré 1: il suffit de prendre $p(x) = \frac{1}{2}x$. Une solution particulière de (1) est donc $y_0(x) = \frac{1}{2}x\,e^x$. Comme on avait déjà déterminé au 2.2 la solution générale de l'équation homogène associée, on conclut que la solution générale de (1) est $y(x) = (\frac{x}{2} + C_1)\,e^x + C_2\,e^{-x}$, avec $C_1, C_2 \in \mathbb{R}$ quelconques.

• *Si le second membre $d(x)$ est une somme $d_1(x) + d_2(x)$*, on cherche une solution particulière $y_0(x)$ de (1) qui soit la somme d'une solution particulière $y_1(x)$ de l'équation $ay'' + by' + c = d_1(x)$ et d'une solution particulière $y_2(x)$ de l'équation $ay'' + by' + c = d_2(x)$. Ceci est utile par exemple si d_1 et d_2 sont de l'un des deux types considérés précédemment.

> EXEMPLE. Considérons l'équation (1) suivante: $y'' - 3y' + 2y = x^2 + x^3\,\mathrm{ch}\,x$.
>
> L'équation homogène associée (2) est $y'' - 3y' + 2y = 0$.
>
> Son équation caractéristique $r^2 - 3r + 2 = 0$ admettant les deux racines réelles distinctes 1 et 2, la solution générale de (2) est $y(x) = C_1\,e^x + C_2\,e^{2x}$ avec $C_1, C_2 \in \mathbb{R}$.
>
> On casse ensuite le second membre $d(x) = x^2 + x^3\,\mathrm{ch}\,x$ en une somme $d(x) = x^2 + \frac{1}{2}x^3\,e^x + \frac{1}{2}x^3\,e^{-x}$.
>
> On cherche une solution particulière y_1 de l'équation $y'' - 3y' + 2y = x^2$ sous forme polynomiale; après calculs, on obtient:
> $$y_1(x) = \frac{1}{2}x^2 + \frac{3}{2}x + \frac{7}{4}.$$
> On cherche une solution particulière y_2 de l'équation $y'' - 3y' + 2y = \frac{1}{2}x^3\,e^x$ sous forme $p(x)\,e^x$ avec $p(x)$ polynomiale. Il vient $p'' - p' = \frac{1}{2}x^3$, soit $p' - p = \frac{1}{8}x^4$, dont une solution polynomiale est $p(x) = -(\frac{1}{8}x^4 + \frac{1}{2}x^3 + \frac{3}{2}x^2 + 3x + 3)$. On obtient donc:
> $$y_2(x) = -(\tfrac{1}{8}x^4 + \tfrac{1}{2}x^3 + \tfrac{3}{2}x^2 + 3x + 3)\,e^x.$$
> On cherche une solution particulière y_3 de l'équation $y'' - 3y' + 2y = \frac{1}{2}x^3\,e^{-x}$ sous forme $q(x)\,e^{-x}$ avec $q(x)$ polynomiale. Il vient $q'' - 5q' + 6q = \frac{1}{2}x^3$, dont une solution polynomiale est $q(x) = \frac{1}{12}x^3 + \frac{5}{24}x^2 + \frac{19}{72}x + \frac{65}{432}$. On obtient donc:
> $$y_3(x) = (\tfrac{1}{12}x^3 + \tfrac{5}{24}x^2 + \tfrac{19}{72}x + \tfrac{65}{432})\,e^{-x}.$$
> On en déduit que la solution générale de (1) est $y(x) = C_1\,e^x + C_2\,e^{2x} + y_1(x) + y_2(x) + y_3(x)$, avec $C_1, C_2 \in \mathbb{R}$ quelconques.

Intégrales généralisées

1. Notion d'intégrale généralisée

On suppose connue la notion d'intégrale d'une fonction continue par morceaux $f : [a,b] \to \mathbb{R}$ sur un intervalle fermé borné (un segment) $[a,b]$, avec $a, b \in \mathbb{R}$, $a \leq b$. On la note $\int_a^b f(t)\,dt$. On cherche à étendre cette notion aux cas où f est définie sur un intervalle infini ou ouvert.

1.1 Premier type de problème.

On considère $f : [a,b[\to \mathbb{R}$ continue par morceaux, avec $a, b \in \mathbb{R}$, $a \leq b$. A priori, f n'est pas définie en b.
Si $\lim\limits_{t\to b, t<b} f(t)$ existe dans $\mathbb{R}$, (appelons alors ℓ cette limite), on prolonge f par continuité à gauche en b en posant $f(b) = \ell$, de façon à obtenir une fonction continue par morceaux sur $[a,b]$, pour laquelle l'intégrale est bien définie. Dans ce cas, il n'y a pas de problème.
La vraie question se pose si $\lim\limits_{t\to b, t<b} f(t) = \pm\infty$ ou n'existe pas. Dans ce cas, le principe est le suivant:

Principe: *on prend $a \leq x < b$, de sorte que f est continue par morceaux sur $[a,x]$; on considère $\int_a^x f(t)\,dt$; puis on fait tendre x vers b à gauche.*

Ceci conduit à la définition suivante:

Définition. *Avec les données ci-dessus, on dit que l'intégrale $\int_a^b f(t)\,dt$ converge lorsque $\lim\limits_{x\to b, x<b} \int_a^x f(t)\,dt$ existe dans $\mathbb{R}$; dans ce cas, on a: $\int_a^b f(t)\,dt = \lim\limits_{x\to b, x<b} \int_a^x f(t)\,dt$.*
Dans le cas contraire, on dit que $\int_a^b f(t)\,dt$ diverge.

Exemple 1. Soit $f : [0,1[\to \mathbb{R}$ définie par $f(t) = \frac{1}{\sqrt{1-t^2}}$.
On a $\lim\limits_{t\to 1, t<1} f(t) = +\infty$.
Pour tout $x \in [0,1[$, on calcule $\int_0^x f(t)\,dt = \big[\arcsin t\big]_0^x = \arcsin x$.
Comme $\lim\limits_{x\to 1} \arcsin x = \frac{\pi}{2}$, on conclut que
$\int_0^1 \frac{dt}{\sqrt{1-t^2}}$ converge et vaut $\frac{\pi}{2}$.

Figure 1

Exemple 2 Soit $f : [0,1[\to \mathbb{R}$ définie par $f(t) = \frac{1}{1-t}$.
On a $\lim\limits_{t\to 1, t<1} f(t) = +\infty$.
Pour $x \in [0,1[$, on calcule $\int_0^x f(t)\,dt = -\big[\ln(1-t)\big]_0^x = -\ln(1-x)$.
Comme $\lim\limits_{x\to 1, x\leq 1} \ln(1-x) = -\infty$, on conclut que
$\int_0^1 \frac{dt}{1-t}$ diverge.

Figure 2

Remarque. On a bien sûr une notion analogue pour $f :]a,b] \to \mathbb{R}$ avec problème à droite en a; on prend $a < x \leq b$, on considère $\int_x^b f(t)\,dt$; puis on fait tendre x vers a à droite.

1.2 Second type de problème.

On considère $f : [a,+\infty[\to \mathbb{R}$ continue par morceaux, avec $a \in \mathbb{R}$. L'intervalle sur lequel on veut intégrer n'est donc pas borné. Le principe est le même que dans le cas précédent:

Principe: *on prend $x > a$, de sorte que f est continue par morceaux sur $[a,x]$; on considère $\int_a^x f(t)\,dt$; puis on fait tendre x vers $+\infty$.*

Ceci conduit à la définition suivante:

Définition. *Avec les données ci-dessus, on dit que l'intégrale* $\int_a^{+\infty} f(t)\,dt$ *converge lorsque* $\lim_{x\to+\infty} \int_a^x f(t)\,dt$ *existe dans* $\mathbb{R}$*; dans ce cas, on pose* $\int_a^{+\infty} f(t)\,dt = \lim_{x\to+\infty} \int_a^x f(t)\,dt$.
Dans le cas contraire, on dit que $\int_a^{+\infty} f(t)\,dt$ *diverge.*

• Exemple 1. Soit $f : [1,+\infty[\to \mathbb{R}$ définie par $f(t) = \frac{1}{t}$.
Pour tout $x \in [1,+\infty[$, on calcule $\int_1^x f(t)\,dt = \big[\ln t\big]_1^x = \ln x$.
Comme $\lim_{x\to+\infty} \ln x = +\infty$, on conclut que
$\int_1^{+\infty} \frac{dt}{t}$ diverge.

Figure 3

• Exemple 2. Soit $f : [1,+\infty[\to \mathbb{R}$ définie par $f(t) = \frac{1}{t^2}$.
Pour tout $x \in [1,+\infty[$, on calcule $\int_1^x f(t)\,dt = \big[-\frac{1}{t}\big]_1^x = 1 - \frac{1}{x}$.
Comme $\lim_{x\to+\infty} 1 - \frac{1}{x} = 1$, on conclut que
$\int_1^{+\infty} \frac{dt}{t^2}$ converge et vaut 1.

Figure 4

• Exemple 3. Soit $f : [1,+\infty[\to \mathbb{R}$ définie par $f(t) = \cos t$.
Pour tout $x \in [1,+\infty[$, on a $\int_1^x f(t)\,dt = \big[\sin t\big]_1^x = \sin x - \sin 1$.
Comme $\lim_{x\to+\infty} \sin x$ n'existe pas dans $\mathbb{R}$, on conclut que
$\int_1^{+\infty} \cos t\,dt$ diverge.

Figure 5

Attention ! *Erreur fréquente et grave:* la condition $\lim_{t\to+\infty} f(t) = 0$ n'est pas suffisante pour assurer la convergence de l'intégrale $\int_a^{+\infty} f(t)\,dt$. (Voir l'exemple 1 ci-dessus).

Elle n'est d'ailleurs pas non plus nécessaire (on verra plus loin que $\int_1^{+\infty} \cos t^2\,dt$ converge bien que $f(t) = \cos t^2$ n'admette pas de limite pour $t \to +\infty$). En revanche on peut montrer que si $f(t)$ admet une limite finie ℓ pour $t \to +\infty$, et si on sait que $\int_a^{+\infty} f(t)\,dt$ est convergente, alors nécessairement $\ell = 0$.

Remarque. On a bien sûr une notion analogue pour un intervalle du type $f :]-\infty, b] \to \mathbb{R}$; on prend $x < b$, on considère $\int_x^b f(t)\,dt$, puis on fait tendre x vers $-\infty$.

1.3 Problèmes mixtes.

On considère $f :]a,b[\to \mathbb{R}$ continue par morceaux, avec $a \in \mathbb{R}$ ou $a = -\infty$, et $b \in \mathbb{R}$ ou $b = +\infty$. On a donc un problème éventuellement aux deux bornes de l'intégrale. Le principe consiste alors à "casser" l'intégrale en une somme de deux intégrales (par la relation de Chasles) et à étudier séparément les deux morceaux. Plus précisément, on a la définition suivante:

Définition. *Avec les données ci-dessus, considérons un réel c tel que* $a < c < b$.
On dit que l'intégrale $\int_a^b f(t)\,dt$ *converge lorsque* $\int_a^c f(t)\,dt$ *et* $\int_c^b f(t)\,dt$ *convergent toutes les deux; dans ce cas, on pose* $\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt$.
Dans le cas contraire, on dit que $\int_a^b f(t)\,dt$ *diverge.*

Exemple 1. Soit $f :]-1,1[\to \mathbb{R}$ définie par $f(t) = \frac{1}{\sqrt{1-t^2}}$. On a $\lim_{t\to1, t<1} f(t) = \lim_{t\to-1, t>-1} f(t) = +\infty$.

On a donc un problème en -1 et un problème en 1. On étudie donc séparément $\int_0^1 \frac{dt}{\sqrt{1-t^2}}$ et $\int_{-1}^0 \frac{dt}{\sqrt{1-t^2}}$.
(On a coupé en 0, on aurait pu choisir n'importe quel autre réel $-1 < c < 1$).
On a vu plus haut que $\int_0^1 \frac{dt}{\sqrt{1-t^2}}$ converge et vaut $\frac{\pi}{2}$. On montre de même que:

$$\int_{-1}^0 \frac{dt}{\sqrt{1-t^2}} = \lim_{x\to-1, x>-1} \int_x^0 \frac{dt}{\sqrt{1-t^2}} = \lim_{x\to-1, x>-1} \big[\arcsin t\big]_x^0 = \lim_{x\to-1, x>-1} (-\arcsin x) = \frac{\pi}{2}.$$

On conclut que $\int_{-1}^1 \frac{dt}{\sqrt{1-t^2}}$ converge et vaut: $\int_{-1}^1 \frac{dt}{\sqrt{1-t^2}} = \int_{-1}^0 \frac{dt}{\sqrt{1-t^2}} + \int_0^1 \frac{dt}{\sqrt{1-t^2}} = \frac{\pi}{2} + \frac{\pi}{2} = \pi$.

Exemple 2. Soit $f :]0,+\infty[\to \mathbb{R}$ définie par $f(t) = \frac{1}{t^2}$. On a un problème en $+\infty$ (intervalle non borné) et un problème en 0 (car $\lim_{t\to 0, t>0} \frac{1}{t^2} = +\infty$). On étudie donc séparément $\int_0^1 \frac{dt}{t^2}$ et $\int_1^{+\infty} \frac{dt}{t^2}$.
(On a coupé en 1, on aurait pu choisir n'importe quel autre réel $0 < c$).
On a vu plus haut que $\int_1^{+\infty} \frac{dt}{t^2}$ converge et vaut 1. On considère par ailleurs:

$$\lim_{x\to 0, x>0} \int_x^1 \frac{dt}{t^2} = \lim_{x\to 0, x>0} \left[-\frac{1}{t}\right]_x^0 = \lim_{x\to 0, x>0} \left(\frac{1}{x} - 1\right) = +\infty.$$

Ainsi $\int_0^1 \frac{dt}{t^2}$ diverge, ce qui, bien que l'autre morceau $\int_1^{+\infty} \frac{dt}{t^2}$ converge, suffit pour conclure que $\int_0^{+\infty} \frac{dt}{t^2}$ diverge.

Attention ! *Erreur fréquente et grave:* la condition $\lim_{x\to+\infty} \int_{-x}^{x} f(t)\,dt$ existe dans $\mathbb{R}$ n'est pas suffisante pour assurer la convergence de l'intégrale $\int_{-\infty}^{+\infty} f(t)\,dt$.

En effet, $\int_{-\infty}^{+\infty} f(t)\,dt$ signifie que les deux intégrales $\int_{-\infty}^{0} f(t)\,dt$ et $\int_0^{+\infty} f(t)\,dt$ convergent, c'est-à-dire que $\lim_{x\to-\infty} \int_x^0 f(t)\,dt$ existe dans $\mathbb{R}$ et $\lim_{x\to+\infty} \int_0^x f(t)\,dt$ existe dans $\mathbb{R}$, ce que n'assure pas l'existence de la seule limite $\lim_{x\to+\infty} \int_{-x}^{x} f(t)\,dt$.

Prenons par exemple $\int_{-\infty}^{+\infty} t\,dt$; elle diverge car $\lim_{x\to+\infty} \int_0^x t\,dt = +\infty$ et $\lim_{y\to-\infty} \int_y^0 t\,dt = -\infty$. Et pourtant $\lim_{x\to+\infty} \int_{-x}^{x} t\,dt = \lim_{x\to+\infty} \left[\frac{t^2}{2}\right]_{-x}^{x} = 0$ par parité de x^2.

1.4 Méthodes de calcul.

• Avant tout, bien repérer la ou les bornes de l'intervalle où il y a un problème. Casser éventuellement en une somme d'intégrales avec problème à la borne de droite (du type $\int_a^b f(t)\,dt$ avec $b = +\infty$ ou $b \in \mathbb{R}$ tel que f n'a pas de limite finie en b à gauche) et d'intégrales avec problème à la borne de gauche (du type $\int_a^b f(t)\,dt$ avec $a = -\infty$ ou $a \in \mathbb{R}$ tel que f n'a pas de limite finie en a à droite).
• Pour le premier type, calculer lorsque c'est possible $\int_a^x f(t)\,dt$ pour un réel x tel que $a < x < b$. Pour cela, on peut utiliser le calcul des primitives, la formule d'intégration par parties, la formule de changement de variables,... Puis, dans le résultat (qui dépend de x) faire tendre x vers b à gauche: si la limite existe dans $\mathbb{R}$, l'intégrale $\int_a^b f(t)\,dt$ converge et est égale la valeur de cette limite; sinon l'intégrale diverge.
• Procéder de même pour le second type, et synthétiser éventuellement les résultats obtenus aux deux bornes comme expliqué ci-dessus en 1.3.

Un exemple fondamental (Riemann). *Pour tout réel α, on a :*

(i) *l'intégrale $\int_1^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$;*

(ii) *l'intégrale $\int_0^1 \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha < 1$.*

Preuve. Pour le point (i), introduisons $x > 1$ et calculons $I(x) = \int_1^x \frac{dt}{t^\alpha} = \int_1^x t^{-\alpha}\,dt$. Si $\alpha = 1$, on a $I(x) = \left[\ln t\right]_1^x = \ln x$, qui tend vers $+\infty$ lorsque x tend vers $+\infty$. Si $\alpha \neq 1$, on a $I(x) = \left[\frac{t^{1-\alpha}}{1-\alpha}\right]_1^x = \frac{1}{1-\alpha}(x^{1-\alpha} - 1)$. Quand x tend vers $+\infty$, cette expression tend vers $+\infty$ si $1-\alpha > 0$, et vers $\frac{1}{\alpha-1}$ si $1-\alpha < 0$. Ce qui prouve le résultat voulu.
Pour le point (ii), introduisons $0 < x < 1$ et calculons $J(x) = \int_x^1 \frac{dt}{t^\alpha} = \int_x^1 t^{-\alpha}\,dt$. Si $\alpha = 1$, on a $J(x) = \left[\ln t\right]_x^1 = -\ln x$, qui tend vers $+\infty$ lorsque x tend vers 0. Si $\alpha \neq 1$, on a $J(x) = \left[\frac{t^{1-\alpha}}{1-\alpha}\right]_x^1 = \frac{1}{\alpha-1}(x^{1-\alpha} - 1)$. Quand x tend vers 0, cette expression tend vers $+\infty$ si $1-\alpha < 0$, et vers $\frac{1}{1-\alpha}$ si $1-\alpha > 0$. Ce qui prouve le résultat voulu. □

Remarque (Linearité de l'intégrale et application aux fonctions à valeurs complexes).
Considérons pour fixer les idées un intervalle $[a,b[$, où a est un réel, et b est soit un réel supérieur à a, soit $+\infty$; (on ferait de même avec un intervalle $]b,a]$, avec a réel et b réel $< a$ ou $-\infty$).

Soient f et g deux applications $[a,b[\to \mathbb{R}$ continues par morceaux. Si les intégrales $\int_a^b f(t)\,dt$ et $\int_a^b g(t)\,dt$ convergent, alors $\int_a^b (f+g)(t)\,dt$ converge et vaut naturellement $\int_a^b f(t)\,dt + \int_a^b g(t)\,dt$. De même $\int_a^b (\lambda f)(t)\,dt$ converge et vaut $\lambda \int_a^b f(t)\,dt$ pour tout réel λ.

Ceci permet en particulier d'étendre tout ce qui précède au cas des fonctions à valeurs complexes.

Si $f : [a, b[\to \mathbb{C}$ continue par morceaux, les fonctions partie réelle $f_1 : [a, b[\to \mathbb{R}$ et partie imaginaire $f_2 : [a, b[\to \mathbb{R}$, définies par $f(t) = f_1(t) + i f_2(t)$ pour tout $t \in [a, b[$, sont continues par morceaux sur $[a, b[$, et l'intégrale $\int_a^b f(t)\,dt$ converge si et seulement si $\int_a^b f_1(t)\,dt$ et $\int_a^b f_2(t)\,dt$ convergent. On a alors $\int_a^b f(t)\,dt = \int_a^b f_1(t)\,dt + i \int_a^b f_2(t)\,dt$.

2. Critères de convergence

Il est très rare dans les faits que l'on sache calculer explicitement une intégrale généralisée. Le plus souvent, on est déjà bien content quand on sait déterminer sa *nature*, c'est-à-dire démontrer qu'elle est convergente ou divergente. On dispose pour cela de divers critères. On rappelle ci-dessous (sans démonstration) les plus usuels. Les deux premiers portent sur les fonctions réelles positives, (attention ils peuvent être faux sinon). On les énonce pour des intégrales généralisées avec problème de convergence à la borne de droite; il y a bien sûr des énoncés analogues pour des intégrales généralisées avec problème de convergence à la borne de gauche.

2.1 Critère de majoration pour les fonctions réelles positives.

Proposition. *Soient f et g deux applications $[a, b[\to \mathbb{R}$ continues par morceaux, où a est un réel, et b est soit un réel supérieur à a, soit $+\infty$. On suppose que, pour t assez grand, on a $0 \leq f(t) \leq g(t)$.*
Si l'intégrale $\int_a^b g(t)\,dt$ converge, alors l'intégrale $\int_a^b f(t)\,dt$ converge.

Remarques.
1. Il en résulte évidemment que, si l'intégrale $\int_a^b f(t)\,dt$ diverge, alors l'intégrale $\int_a^b g(t)\,dt$ diverge.
2. Rappelons que l'expression "pour t assez grand, on a $0 \leq f(t) \leq g(t)$" signifie qu'il existe un réel $c \in [a, b[$ tel que, pour tout $t \in [c, b[$, on a $0 \leq f(t) \leq g(t)$.
3. Si $f(t) \leq g(t)$ pour tout $t \in [a, b[$ et que les deux intégrales convergent, on a: $0 \leq \int_a^b f(t)\,dt \leq \int_a^b g(t)\,dt$.

Corollaire. *Soit $f : [a, +\infty[\to \mathbb{R}$ positive et continue par morceaux, avec $a \in \mathbb{R}+^*$.*

(i) *S'il existe $\alpha > 1$ tel que $\lim\limits_{t\to+\infty} t^\alpha f(t) = 0$, alors $\int_a^{+\infty} f(t)\,dt$ converge.*

(ii) *S'il existe $\alpha \leq 1$ tel que $\lim\limits_{t\to+\infty} t^\alpha f(t) = +\infty$, alors $\int_a^{+\infty} f(t)\,dt$ diverge.*

En effet, si $\lim\limits_{t\to+\infty} t^\alpha f(t) = 0$, on a pour t assez grand $f(t) \leq \frac{1}{t^\alpha}$; comme $\alpha > 1$, on sait d'aprs le point (i) de 1.4 que $\int_a^b \frac{1}{t^\alpha}\,dt$ converge, ce qui prouve (i) grâce à la proposition ci-dessus. Le point (ii) se montre de même.

2.2 Critère d'équivalence pour les fonctions réelles positives.

Proposition. *Soient f et g deux applications $[a, b[\to \mathbb{R}$ continues par morceaux, où a est un réel, et b est soit un réel supérieur à a, soit $+\infty$. On suppose que, pour t assez grand, on a $f(t) \geq 0$ et $g(t) \geq 0$. On suppose aussi que f est équivalente à g au voisinage de b.*
Alors les intégrales $\int_a^b f(t)\,dt$ et $\int_a^b g(t)\,dt$ sont de même nature (c'est-à-dire toutes les deux convergentes ou toutes les deux divergentes).

Remarque. Rappelons que f est équivalente à g au voisinage de b signifie par définition que:
il existe un réel $c \in [a, b[$ tel que, pour tout $t \in [c, b[$, on a $f(t) = g(t)(1 + \varepsilon(t))$ avec $\lim\limits_{t\to b} \varepsilon(t) = 0$.
On note alors $f \sim_b g$.

2.3 Un exemple d'application important: la fonction eulérienne Γ.

Proposition.

(i) *Pour tout réel $x > 0$, l'intégrale $\Gamma(x) = \int_0^{+\infty} \mathrm{e}^{-t} t^{x-1}\,dt$ est convergente.*

(ii) *Pour tout réel $x > 0$, on a $\Gamma(x + 1) = x\Gamma(x)$.*

(iii) *Pour tout entier $n > 0$, on a $\Gamma(n) = (n - 1)! = 1 \times 2 \times 3 \times \cdots \times (n - 1)$.*

Démonstration. Pour montrer (i), on fixe $x > 0$. Notons pour simplifier $f(t) = \mathrm{e}^{-t}t^{x-1}$, de sorte que $\Gamma(x) = \int_0^{+\infty} f(t)\,dt$. Il y a un problème en $+\infty$, et un autre en 0 car $x-1$ peut être strictement négatif et t^{x-1} tend alors vers l'infini pour t tendant vers 0. On étudie donc séparément $\int_0^1 f(t)\,dt$ et $\int_1^{+\infty} f(t)\,dt$.
Etude du problème en 0 : $f(t)$ est positive pour tout $t \in]0,1]$ et $f(t)$ est équivalent à t^{x-1} au voisinage de 0. Or $\int_0^1 t^{x-1}\,dt = \int_0^1 \frac{dt}{t^{1-x}}$ est convergente par application du point (ii) de l'exemple fondamental 1.4, car $1 - x < 1$ par hypothèse. On applique donc la proposition 2.2 pour conclure que $\int_0^1 f(t)\,dt$ converge.
Etude du problème en $+\infty$: On sait que $\lim\limits_{t\to+\infty} \mathrm{e}^{-t}t^{x+1} = 0$, c'est--dire $\lim\limits_{t\to+\infty} t^2 f(t) = 0$. On applique donc le corollaire de la proposition 2.1 pour conclure que $\int_1^{+\infty} f(t)\,dt$ converge.
Bilan : les deux intégrales $\int_0^1 f(t)\,dt$ et $\int_1^{+\infty} f(t)\,dt$ sont convergentes donc, d'après 1.3, $\Gamma(x) = \int_0^{+\infty} \mathrm{e}^{-t}t^{x-1}\,dt$ est convergente.

Pour montrer (ii), on fait une intégration par parties en posant $u'(t) = \mathrm{e}^{-t}$, $u(t) = -\mathrm{e}^{-t}$, $v(t) = t^x$, $v'(t) = xt^{x-1}$. Donc:

$$\begin{aligned}\Gamma(x+1) &= \int_0^{+\infty} \mathrm{e}^{-t}t^x\,dt = \int_0^{+\infty} u'(t)v(t)\,dt \\ &= \big[u(t)v(t)\big]_0^{+\infty} - \int_0^{+\infty} u(t)v'(t)\,dt = \big[-\mathrm{e}^{-t}t^x\big]_0^{+\infty} + x\int_0^{+\infty} \mathrm{e}^{-t}t^{x-1}\,dt,\end{aligned}$$

et comme le premier terme est nul, on conclut que $\Gamma(x+1) = x\Gamma(x)$.
Pour montrer (iii), on utilise le point (ii) pour vérifier que $\Gamma(n) = (n-1)\Gamma(n-1)$, d'où par récurrence $\Gamma(n) = (n-1)\times(n-2)\times(n-3)\times\cdots 3\times 2\times 1\times\Gamma(1)$. On calcule enfin $\Gamma(1) = \int_0^{+\infty} \mathrm{e}^{-t}\,dt = \big[-\mathrm{e}^{-t}\big]_0^{+\infty} = 1$, ce qui montre le résultat voulu. □

2.4 Convergence absolue.

Considérons pour fixer les idées un intervalle $[a,b[$, où a est un réel, et b est soit un réel supérieur à a, soit $+\infty$; (on ferait de même avec un intervalle $]b,a]$, avec a réel et b réel $< a$ ou $-\infty$). Soit $f : [a,b[\to \mathbb{R}$ ou $\mathbb{C}$ continue par morceaux. On peut montrer qu'alors l'application $|f| : [a,b[\to \mathbb{R}^+$ est aussi continue par morceaux. (Rappelons que $|f(t)|$ désigne la valeur absolue ou le module de $f(t)$). La définition suivante est donc fondée.

Définition. *Avec les données ci-dessus, on dit que l'intégrale $\int_a^b f(t)\,dt$ est absolument convergente lorsque l'intégrale $\int_a^b |f(t)|\,dt$ est convergente.*

Exemple. L'intégrale $\int_1^{+\infty} \frac{\mathrm{e}^{it}}{t^2}\,dt$ est absolument convergente.
En effet, puisque $\mathrm{e}^{it} = \cos t + i\sin t$, on a $|\mathrm{e}^{it}| = 1$, de sorte que la fonction à valeurs complexes $f(t) = \frac{\mathrm{e}^{it}}{t^2}$ vérifie $|f(t)| = \frac{1}{t^2}$. Or on sait (voir 1.4) que $\int_1^{+\infty} \frac{1}{t^2}\,dt$ converge; ce qui prouve le résultat voulu.
L'intérêt de la notion d'intégrale absolument convergente est qu'elle fournit une condition suffisante de convergence de l'intégrale, en se ramenant au cas d'une fonction réelle positive, pour laquelle on peut appliquer les critères 2.1 et 2.2. C'est ce que montre le théorème (que nous rappelons sans preuve) suivant.

Théorème. *Avec les données ci-dessus, si l'intégrale $\int_a^b f(t)\,dt$ est absolument convergente, alors elle est convergente, et l'on a:* $\left|\int_a^b f(t)\,dt\right| \le \int_a^b |f(t)|\,dt$.

Attention ! *Erreur fréquente et grave:* le théorème affirme que la convergence absolue entraîne la convergence, mais la réciproque est fausse. En d'autres termes, si $\int_a^b |f(t)|\,dt$ converge, alors $\int_a^b f(t)\,dt$ converge, mais on peut avoir $\int_a^b f(t)\,dt$ qui converge bien que $\int_a^b |f(t)|\,dt$ diverge.

Considérons par exemple l'intégrale $I = \int_1^{+\infty} \frac{\mathrm{e}^{it}}{t}\,dt$.

Elle n'est pas absolument convergente car $\int_1^{+\infty} \left|\frac{\mathrm{e}^{it}}{t}\right| dt = \int_1^{+\infty} \frac{1}{t}\,dt$ qui diverge comme on l'a vu en 1.4.

Montrons que cependant I converge. Pour cela prenons $x > 1$ et posons $I_x = \int_1^x \frac{\mathrm{e}^{it}}{t}\,dt$. Une intégration par parties donne $I_x = \big[\frac{\mathrm{e}^{it}}{it}\big]_1^x + \int_1^x \frac{\mathrm{e}^{it}}{it^2}\,dt = \frac{\mathrm{e}^{ix}}{ix} - \frac{\mathrm{e}^{i}}{i} + \frac{1}{i}\int_1^x \frac{\mathrm{e}^{it}}{t^2}\,dt$. D'une part $\frac{\mathrm{e}^{ix}}{ix}$ tend vers 0 quand x tend vers $+\infty$, car son module est $\frac{1}{x}$. D'autre part $\int_1^x \frac{\mathrm{e}^{it}}{t^2}\,dt$ tend vers une limite finie quand x tend vers $+\infty$, car on a montré ci-dessus que $\int_1^{+\infty} \frac{\mathrm{e}^{it}}{t^2}\,dt$ converge. On conclut que I_x tend vers une limite finie quand x tend vers $+\infty$.

Si $\int_a^b |f(t)|\,dt$ diverge, le théorème ne permet absolument pas de conclure quoi que ce soit sur la convergence de $\int_a^b f(t)\,dt$, qui peut soit converger soit diverger.

2.5 Un exemple d'utilisation des différents résultats.

On avait affirmé à la fin du 1.2 que l'intégrale $\int_1^{+\infty}(\cos t^2)\,dt$ converge. Montrons-le.

Prenons pour cela un réel $x > 1$ et considérons $J_x = \int_1^x(\cos t^2)\,dt$.

On fait le changement de variables $u = t^2$, d'où $du = 2t\,dt$ et donc $dt = \frac{du}{2\sqrt{u}}$. On obtient $J_x = \int_1^{x^2} \frac{\cos u}{2\sqrt{u}}\,du$.

Pour tout $y > 1$, posons $I_y = \int_1^y \frac{\cos u}{\sqrt{u}}\,du$. Une intégration par parties donne $I_y = \left[\frac{\sin u}{\sqrt{u}}\right]_1^y + \frac{1}{2}\int_1^y \frac{\sin u}{u^{3/2}}\,du$.

Le premier terme est sans problème: $\lim\limits_{y\to+\infty}(\frac{\sin y}{\sqrt{y}} - \frac{\sin 1}{\sqrt{1}}) = -\sin 1$.

La limite du second se ramène à l'étude de la convergence de l'intégrale $\int_1^y \frac{\sin u}{u^{3/2}}\,du$.

> On a la majoration: $\left|\frac{\sin u}{u^{3/2}}\right| \leq \frac{1}{u^{3/2}}$; et on sait d'après le point (i) de l'exemple fondamental de 1.4 que $\int_1^{+\infty} \frac{du}{u^{3/2}}\,du$ converge. Donc l'application du critère de majoration 2.1 prouve que $\int_1^{+\infty}\left|\frac{\sin u}{u^{3/2}}\right|\,du$ converge. En d'autres termes, l'intégrale $\int_1^y \frac{\sin u}{u^{3/2}}\,du$ est absolument convergente, et donc convergente d'après le théorème du 2.4. Ceci prouve que $\int_1^y \frac{\sin u}{u^{3/2}}\,du$ tend vers une limite finie (notons-la L) dans $\mathbb{R}$ lorsque y tend vers $+\infty$.

On en déduit que $\lim\limits_{y\to+\infty} I_y = L - \sin 1$, donc $\lim\limits_{x\to+\infty} J_x = \frac{1}{2}\lim\limits_{y\to+\infty} I_y$ existe dans $\mathbb{R}$.

On conclut que $\int_1^{+\infty}(\cos t^2)\,dt$ converge; (bien que la fonction $t \mapsto \cos t^2$ n'admette pas de limite quand $t \to +\infty$).

2.6 Un exemple d'application en physique.

Dans un circuit R, L, C en série (résistance, self-inductance, capacité), avec une tension aux bornes constante V, une charge initiale q_0 du condensateur, et un courant initial nul, le courant à l'instant t est d'intensité $i(t) = \frac{CV-q_0}{LC\omega}\mathrm{e}^{-\frac{R}{2L}t}\sin\omega t$, où l'on a posé $\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$ en supposant $R < 2\sqrt{\frac{L}{C}}$. Montrer que l'énergie produite sous forme de chaleur dans la résistance R pendant un intervalle de temps $[0, +\infty[$ a une valeur finie.

Solution. Cette énergie est donnée par: $W = \int_0^{+\infty} Ri^2(t)\,dt = R\left(\frac{CV-q_0}{LC\omega}\right)^2\int_0^{+\infty}\mathrm{e}^{-\frac{R}{L}t}\sin^2\omega t\,dt$.

En posant $\alpha = \frac{R}{L} > 0$, on est ramené à étudier l'intégrale $J = \int_0^{+\infty}\mathrm{e}^{-\alpha t}\sin^2\omega t\,dt$.
On a la majoration: $0 \leq \mathrm{e}^{-\alpha t}\sin^2\omega t \leq \mathrm{e}^{-\alpha t}$.
Or $\int_0^{+\infty}\mathrm{e}^{-\alpha t}\,dt$ converge, car $\int_0^x \mathrm{e}^{-\alpha t}\,dt = \left[\frac{\mathrm{e}^{-\alpha t}}{-\alpha}\right]_0^x = \frac{1-\mathrm{e}^{-\alpha x}}{\alpha}$ tend vers une limite finie $\frac{1}{\alpha}$ quand $x \to +\infty$.
Donc par application du critère de majoration 2.1, on conclut que J converge, et donc W aussi.

> Remarque. On peut calculer la valeur explicite de J et donc de $W = R\left(\frac{CV-q_0}{LC\omega}\right)^2 J$.
>
> Pour cela, on rappelle d'abord que $\sin^2\omega t = \frac{1}{2}(1 - \cos 2\omega t)$, de sorte que:
>
> $J = \frac{1}{2}\int_0^{+\infty}\mathrm{e}^{-\alpha t}(1-\cos 2\omega t)\,dt = \frac{1}{2}\left[\frac{\mathrm{e}^{-\alpha t}}{-\alpha}\right]_0^{+\infty} - \frac{1}{2}\int_0^{+\infty}\mathrm{e}^{-\alpha t}\cos 2\omega t\,dt = \frac{1}{2\alpha} - \frac{1}{2}K$, avec $K = \int_0^{+\infty}\mathrm{e}^{-\alpha t}\cos 2\omega t\,dt$.
>
> Une première intégration par parties donne $K = \left[\mathrm{e}^{-\alpha t}\frac{\sin 2\omega t}{2\omega}\right]_0^{+\infty} + \int_0^{+\infty}\alpha\mathrm{e}^{-\alpha t}\frac{\sin 2\omega t}{2\omega}\,dt = 0 + \frac{\alpha}{2\omega}H$, avec $H = \int_0^{+\infty}\mathrm{e}^{-\alpha t}\sin 2\omega t\,dt$.
>
> Une seconde intégration par parties donne $H = \left[-\mathrm{e}^{-\alpha t}\frac{\cos 2\omega t}{2\omega}\right]_0^{+\infty} - \int_0^{+\infty}\alpha\mathrm{e}^{-\alpha t}\frac{\cos 2\omega t}{2\omega}\,dt = \frac{1}{2\omega} - \frac{\alpha}{2\omega}K$.
>
> Ainsi $H = \frac{1}{2\omega} - \frac{\alpha}{2\omega}K = \frac{1}{2\omega} - \frac{\alpha}{2\omega}\frac{\alpha}{2\omega}H$, d'où $H = \frac{2\omega}{\alpha^2+4\omega^2}$, puis $K = \frac{\alpha}{\alpha^2+4\omega^2}$.
>
> On en tire $J = \frac{1}{2\alpha} - \frac{1}{2}K = \frac{1}{2}\left(\frac{1}{\alpha} - \frac{\alpha}{\alpha^2+4\omega^2}\right) = \frac{2\omega^2}{\alpha(\alpha^2+4\omega^2)}$.
>
> Comme $\omega^2 = \frac{1}{LC} - \frac{R^2}{4L^2}$ et $\alpha = \frac{R}{L}$, on a $\alpha^2 + 4\omega^2 = \frac{4}{LC}$, donc:
>
> $W = R\left(\frac{CV-q_0}{LC\omega}\right)^2 J = R\left(\frac{CV-q_0}{LC\omega}\right)^2\frac{2\omega^2}{\alpha(\alpha^2+4\omega^2)}$, et finalement $W = \frac{(CV-q_0)^2}{2C}$.
>
> En particulier, si $q_0 = 0$, l'énergie produite par la charge du condensateur est $W = \frac{1}{2}CV^2$.

Séries numériques

1. Notion de série

1.1 Exemples préliminaires

Exemple 1. Soit $(u_n)_{n\geq 1}$ la suite de nombres réels définie par $u_n = \frac{1}{n(n+1)}$ pour tout $n \geq 1$. On définit à partir de cette suite une nouvelle suite $(S_N)_{N\geq 1}$ de la façon suivante:
$S_1 = u_1 = \frac{1}{2}$, $S_2 = u_1+u_2 = \frac{1}{2}+\frac{1}{6} = \frac{2}{3}$, $S_3 = u_1+u_2+u_3 = \frac{2}{3}+\frac{1}{12} = \frac{3}{4}$, $S_4 = u_1+u_2+u_3+u_4 = \frac{3}{4}+\frac{1}{20} = \frac{4}{5}$,
et plus généralement $S_N = u_1 + u_2 + \cdots + u_N = \sum\limits_{n=1}^{N} u_n$ pour tout entier $N \geq 1$.
Comme $u_n = \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, on peut calculer $S_N = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{N} - \frac{1}{N+1} = 1 - \frac{1}{N+1}$.
Il en résulte que la suite (S_N) converge, et tend vers 1 quand N tend vers $+\infty$.
On note $\sum\limits_{n=1}^{+\infty} u_n = \lim\limits_{N\to+\infty} S_N$, c'est-à-dire que $\sum\limits_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1$.
On dit dans ce cas que la *série* $\sum u_n$ est convergente, et que sa *somme* est 1.

Exemple 2. Soit $(u_n)_{n\geq 1}$ la suite de nombres réels définie par $u_n = \frac{n+1}{n}$ pour tout $n \geq 1$. On définit à partir de cette suite la nouvelle suite $(S_N)_{N\geq 1}$ en posant:
$S_N = u_1 + u_2 + \cdots + u_N = \sum\limits_{n=1}^{N} u_n$ pour tout entier $N \geq 1$.
Comme $u_n = 1 + \frac{1}{n}$, on peut calculer $S_N = 1 + \frac{1}{2} + 1 + \frac{1}{3} + 1 + \frac{1}{4} + \cdots + 1 + \frac{1}{N} = N + (\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{N})$.
En particulier, $S_N \geq N$ pour tout N, ce qui implique que $\lim\limits_{N\to+\infty} S_N = +\infty$, c'est-à-dire que $\sum\limits_{n=1}^{+\infty} \frac{n+1}{n} = +\infty$.
On dit dans ce cas que la *série* $\sum u_n$ est divergente.

Exemple 3. Soit $(u_n)_{n\geq 1}$ la suite de nombres réels définie par $u_n = \frac{1}{n}$ pour tout $n \geq 1$. On définit à partir de cette suite la nouvelle suite $(S_N)_{N\geq 1}$ en posant:
$S_N = u_1 + u_2 + \cdots + u_N = \sum\limits_{n=1}^{N} u_n$ pour tout entier $N \geq 1$.
Considérons $|S_{2N} - S_N| = S_{2N} - S_N = \frac{1}{N+1} + \frac{1}{N+2} + \frac{1}{N+3} + \cdots + \frac{1}{2N}$.
Chacun des N termes de cette somme est $\geq \frac{1}{2N}$, donc $|S_{2N} - S_N| \geq N \times \frac{1}{2N} = \frac{1}{2}$ pour tout $N \geq 1$.
Cette condition permet de montrer que la suite (S_N) ne peut pas converger. En effet, si elle admettait une limite finie ℓ, on aurait pour tout $N \geq 1$ la double inégalité: $\frac{1}{2} \leq |S_{2N} - S_N| \leq |S_{2N} - \ell| + |S_N - \ell|$ avec chacun des deux termes du membre de droite qui tend vers 0 quand N tend vers $+\infty$, d'où une contradiction.
On conclut que dans ce cas la *série* $\sum u_n$ est divergente.

1.2 Définitions

Soit $(u_n)_{n\geq 1}$ une suite de nombres réels (ou complexes). On appelle *suite des sommes partielles* associée à $(u_n)_{n\geq 1}$ la suite $(S_N)_{N\geq 1}$ de nombres réels (ou complexes) définie par:

$$S_N = u_1 + u_2 + \cdots + u_N = \sum_{n=1}^{N} u_n, \quad \text{pour tout entier } N \geq 1.$$

Quand on étudie la suite $(S_N)_{N\geq 1}$, on dit que l'on étudie la *série* $\sum\limits_{n\geq 1} u_n$, appelée série de *terme général* u_n.
La question principale que l'on considère est celle de la convergence ou non de la suite (S_N). Lorsque la suite $(S_N)_{N\geq 1}$ est convergente, on dit que la série $\sum\limits_{n\geq 1} u_n$ est *convergente*, et la limite de la suite $(S_N)_{N\geq 1}$ est appelée la *somme* de la série $\sum\limits_{n\geq 1} u_n$. C'est un nombre (réel ou complexe); on le note: $\sum\limits_{n=1}^{+\infty} u_n = \lim\limits_{N\to+\infty} S_N$.

Dans le cas contraire (si la suite (S_N) diverge vers $\pm\infty$ ou n'admet pas de limite), on dit que la série $\sum\limits_{n\geq 1} u_n$ est *divergente*. Etudier la *nature* d'une série consiste à déterminer si elle est convergente ou divergente.

Remarque 1. On a donné les définitions ci-dessus pour une série de premier indice $n = 1$. Mais la suite (u_n) peut a priori démarrer à u_0, ou u_1, ou u_2, ou plus. Ceci est sans importance pour déterminer la nature de la série de terme général u_n, car il est clair que les premiers termes sont sans influence sur le fait que la suite des sommes partielles converge ou non. En revanche, on fera attention que, dans le cas où la série converge, la valeur de la somme dépend des termes à partir desquels on effectue la sommation.

Remarque 2 (fondamentale). Pour qu'une série $\sum u_n$ converge, il est nécessaire (mais non suffisant) que la suite (u_n) converge vers 0. En d'autres termes, pour qu'une série converge, il est nécessaire (mais non suffisant) que son terme général tende vers 0.

> En effet, si la série $\sum u_n$ converge, la suite $(S_N - S_{N-1})$ converge vers 0 (comme différence de deux suites convergeant vers une même limite). Or par définition des sommes partielles, $S_N - S_{N-1} = u_N$.

Ainsi, si la suite (u_n) ne converge pas vers 0, il n'y a pas de question se poser: on est sûr que la série $\sum u_n$ est divergente. La question de la convergence d'une série ne se pose donc que si le terme général tend vers 0.

Attention ! *Erreur fréquente et grave:* la condition $\lim\limits_{n\to+\infty} u_n = 0$ n'est pas suffisante pour assurer la convergence de la série $\sum u_n$ (elle est nécessaire, voir ci-dessus).
Par exemple, on a vu à l'exemple 3 du 1.1 que la série $\sum \frac{1}{n}$ est divergente; et pourtant on a bien $\lim\limits_{n\to+\infty} \frac{1}{n} = 0$.
Cette série divergente $\sum \frac{1}{n}$ est un exemple classique à connaître, appelé la *série harmonique*.

1.3 L'exemple des séries géométriques.

Fixons un nombre x réel ou complexe. Posons $u_n = x^n$ pour tout entier $n \geq 0$. La série $\sum\limits_{n\geq 0} u_n = \sum\limits_{n\geq 0} x^n$ ainsi déterminée s'appelle la *série géométrique de raison x*.

Proposition. *La série $\sum\limits_{n\geq 0} x^n$ converge si et seulement si $|x| < 1$, et alors sa somme est $\sum\limits_{n=0}^{+\infty} x^n = \frac{1}{1-x}$.*

> *Preuve.* Si $|x| \geq 1$, le terme général x^n ne tend pas vers 0, donc la série diverge. Si $|x| < 1$, alors $\lim\limits_{N\to+\infty} x^{N+1} = 0$ de sorte que $S_N = \sum\limits_{n=0}^{N} x^n = \frac{1-x^{N+1}}{1-x}$ vérifie $\lim\limits_{N\to+\infty} S_N = \frac{1}{1-x}$. □

1.4 Linéarité.

Soient $\sum u_n$ et $\sum v_n$ deux séries. Leur somme est la série $\sum(u_n + v_n)$ de terme général $u_n + v_n$. Le produit de la série $\sum u_n$ par un nombre fixé (réel ou complexe) λ est la série $\sum \lambda u_n$ de terme général λu_n.
On peut montrer facilement que: si les deux séries $\sum u_n$ et $\sum v_n$ convergent, alors la série somme $\sum(u_n+v_n)$ converge, ainsi que la série produit $\sum \lambda u_n$ par tout nombre λ, et l'on a:

$$\sum_{n=1}^{+\infty}(u_n + v_n) = \sum_{n=1}^{+\infty} u_n + \sum_{n=1}^{+\infty} v_n \quad \text{et} \quad \sum_{n=1}^{+\infty}(\lambda u_n) = \lambda \sum_{n=1}^{+\infty} u_n.$$

En particulier, une série $\sum u_n$ de nombres complexes converge si et seulement si les deux séries de nombres réels $\sum x_n$ et $\sum y_n$ définies par $u_n = x_n + iy_n$ pour tout entier n convergent, et l'on a dans ce cas:

$$\sum_{n=1}^{+\infty} u_n = \sum_{n=1}^{+\infty} x_n + i \sum_{n=1}^{+\infty} y_n.$$

2. Cas des séries à termes réels positifs

Il est très rare dans les faits que l'on sache calculer explicitement la somme d'une série, comme on l'a fait à l'exemple 1 de 1.1, ou à l'exemple 1.3. Le plus souvent, on est déjà bien content quand on sait déterminer sa *nature*, c'est-à-dire démontrer qu'elle est convergente ou divergente. Dans le cas des séries à termes réels positifs, on dispose pour cela de divers critères. On rappelle ci-dessous (sans démonstration) les plus usuels.

2.1 Critères de majoration et d'équivalence pour les séries à termes réels positifs.

Proposition 1. *Soit $\sum u_n$ et $\sum v_n$ deux séries de nombres réels telles que $0 \leq u_n \leq v_n$ pour tout n supérieur à un certain rang p.*

(i) *Si $\sum v_n$ converge, alors $\sum u_n$ converge aussi, et $0 \leq \sum_{n=p}^{+\infty} u_n \leq \sum_{n=p}^{+\infty} v_n$.*

(ii) *Si $\sum u_n$ diverge, alors $\sum v_n$ diverge aussi.*

Proposition 2. *Soit $\sum u_n$ et $\sum v_n$ deux séries de nombres réels telles que $u_n \geq 0$ et $v_n \geq 0$ pour tout n supérieur à certain rang p. Si les suites (u_n) et (v_n) sont équivalentes au voisinage de l'infini, alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature (soit toutes les deux convergentes, soit toutes les deux divergentes).*

Rappelons que (u_n) et (v_n) équivalentes au voisinage de l'infini signifie qu'à partir d'un certain rang, on a $u_n = v_n(1+\varepsilon_n)$, avec (ε_n) une suite de réels convergeant vers 0.

2.2 Application: exemple fondamental des séries de Riemann

On appelle série de Riemann toute série de nombres réels positifs de la forme $\sum_{n\geq 1} \frac{1}{n^\alpha}$, où α est un nombre réel fixé.

Théorème. *La série de Riemann $\sum_{n\geq 1} \frac{1}{n^\alpha}$ est convergente si et seulement si $\alpha > 1$.*

Preuve. Si $\alpha \leq 1$, on a $n^\alpha \leq n$, donc $\frac{1}{n^\alpha} \geq \frac{1}{n} > 0$. Comme la série harmonique $\sum \frac{1}{n}$ diverge (voir exemple 3 de 1.1), on applique le point (ii) la proposition 1 de 2.1 pour conclure que $\sum \frac{1}{n^\alpha}$ diverge dans ce cas.

On supposera donc maintenant que $\alpha > 1$; on pose $\beta = \alpha - 1 > 0$. On définit $a_n = \frac{1}{n^\beta}$ et $b_n = a_n - a_{n+1}$ pour tout $n \geq 1$. Comme $\beta > 0$, on a $a_{n+1} < a_n$ donc $b_n > 0$. On raisonne en deux étapes.

1. La série $\sum b_n$ converge. En effet, la somme partielle $B_N = b_1 + b_2 + \cdots + b_N$ vaut $B_N = a_1 - a_2 + a_2 - a_3 + \cdots + a_N - a_{N+1} = a_1 - a_{N+1} = 1 - \frac{1}{(N+1)^\beta}$, qui tend vers 1 quand $N \to +\infty$ puisque $\beta > 0$.
2. $\frac{1}{n^{\beta+1}}$ est équivalent à $\frac{1}{\beta} b_n$ au voisinage de l'infini. En effet, on a: $b_n = \frac{1}{n^\beta} - \frac{1}{(n+1)^\beta} = \frac{1}{n^\beta}[1 - (\frac{n}{n+1})^\beta] = \frac{1}{n^\beta}[1 - (1+\frac{1}{n})^{-\beta}]$; mais $(1+\frac{1}{n})^{-\beta} = 1 - \beta\frac{1}{n} + \frac{1}{n}\varepsilon_n$ avec $\lim_{n\to+\infty} \varepsilon_n = 0$. Donc $b_n = \frac{\beta}{n^{\beta+1}}[1 - \frac{1}{\beta}\varepsilon_n]$.

Ainsi $\frac{1}{n^\alpha} = \frac{1}{n^{\beta+1}}$ est équivalent à $\beta^{-1} b_n$; comme la série de terme général b_n converge, il en est de même de la série de terme général $\beta^{-1} b_n$ (voir 1.4), et l'on conclut avec la proposition 2 de 2.1 que la série $\sum \frac{1}{n^\alpha}$ converge. □

Conséquence pratique. Une des méthodes usuelles pour déterminer la nature d'une série à termes réels positifs consiste à chercher à la comparer à une série de référence dont on connaît la nature (série de Riemann ou série géométrique entre autres), et à appliquer les propositions de 2.1 ci-dessus.

Exemple 1. La série $\sum \frac{e^{-n}}{n^2}$ est convergente, car $0 \leq \frac{e^{-n}}{n^2} \leq \frac{1}{n^2}$ pour tout $n \geq 1$, et la série de Riemann $\sum \frac{1}{n^2}$ est convergente puisque $2 > 1$; (on applique le point (i) de la proposition 1 de 2.1).

Exemple 2. La série $\sum \frac{\ln n}{\sqrt{n}}$ est divergente, car $\frac{\ln n}{\sqrt{n}} \geq \frac{1}{\sqrt{n}} \geq 0$ pour tout $n \geq 3$, et la série de Riemann $\sum \frac{1}{n^{1/2}}$ est divergente puisque $\frac{1}{2} < 1$; (on applique le point (ii) de la proposition 1 de 2.1).

Exemple 3. La série $\sum \frac{5n+1}{n^3+2n+1}$ est convergente, car $\frac{5n+1}{n^3+2n+1}$ est positif et équivalent à $\frac{5}{n^2}$ au voisinage de l'infini, et la série de Riemann $\sum \frac{1}{n^2}$ converge puisque $2 > 1$; (on applique la proposition 2 de 2.1).

Avec la proposition 1 de 2.1 et le théorme ci-dessus, on déduit (par le même raisonnement qu'au corollaire du 2.1 de la leçon sur les intégrales généralisées) le corollaire suivant, dit *rgle de comparaison avec les séries de Riemann.*

Corollaire. *Soit $\sum u_n$ une série termes réels positifs.*

(i) *S'il existe $\alpha > 1$ tel que $\lim_{n\to+\infty} n^\alpha u_n = 0$, alors la série $\sum u_n$ converge.*

(ii) *S'il existe $\alpha \leq 1$ tel que $\lim_{n\to+\infty} n^\alpha u_n = +\infty$, alors la série $\sum u_n$ diverge.*

2.3 Règle de d'Alembert

Proposition. *Soit $\sum u_n$ une série réelle telle que $u_n > 0$ à partir d'un certain rang. On suppose que la suite $(\frac{u_{n+1}}{u_n})$ admet une limite finie $\ell \in \mathbb{R}_+$ pour n tendant vers $+\infty$.*

(i) *Si $\ell < 1$, alors la série $\sum u_n$ est convergente.*

(ii) *Si $\ell > 1$, alors la série $\sum u_n$ est divergente.*

Exemple. Considérons la série $\sum u_n$ pour $u_n = \frac{n!}{n^n}$. On a bien $u_n > 0$ pour tout $n \geq 1$ et $\frac{u_{n+1}}{u_n} = (\frac{n}{n+1})^n = (1+\frac{1}{n})^{-n} = \exp(-n\ln(1+\frac{1}{n}))$ qui tend vers $\mathrm{e}^{-1} = \frac{1}{\mathrm{e}} < 1$ quand $n \to +\infty$. On conclut que $\sum u_n$ converge.

Remarque. Lorsque $\ell = 1$, on ne peut pas conclure; la série $\sum u_n$ peut converger ou diverger (prendre par exemple d'une part $u_n = \frac{1}{n}$, et d'autre part $u_n = \frac{1}{n^2}$).

2.4 Règle de Cauchy

Proposition. *Soit $\sum u_n$ une série réelle telle que $u_n > 0$ à partir d'un certain rang. On suppose que la suite $((u_n)^{1/n})$ admet une limite finie $\ell \in \mathbb{R}_+$ pour n tendant vers $+\infty$.*

(i) *Si $\ell < 1$, alors la série $\sum u_n$ est convergente.*

(ii) *Si $\ell > 1$, alors la série $\sum u_n$ est divergente.*

Exemple. Considérons la série $\sum u_n$ pour $u_n = (\frac{n+1}{2n-1})^n$. On a bien $u_n > 0$ pour tout $n \geq 1$ et $(u_n)^{1/n} = \frac{n+1}{2n-1}$ qui tend vers $\frac{1}{2} < 1$ quand $n \to +\infty$. On conclut que $\sum u_n$ converge.

Dans la pratique, les rgles de d'Alembert et de Cauchy ne s'appliquent qu' des situations o la forme du terme général u_n incite directement prendre un quotient de deux termes consécutifs (séries entires, présence de factorielles,...) ou une racine n-ime (présence d'une puissance n-ime,...). Pour le problme de déterminer la nature d'une série termes positifs, l'outil de base est constitué par les critres de 2.1. La rgle de comparaison avec les séries de Riemann qui en découle (corollaire de 2.2) est un argument classique et d'utilisation fréquente, bien connaître.

Attention ! Les propositions précédentes ne s'appliquent qu'à des séries à termes réels positifs. Les résultats correspondants peuvent être faux pour des séries à termes complexes, ou réels de signe non constant. Pour de telles séries, on rappelle dans la dernière partie (sans démonstration) deux résultats importants.

3. Cas des séries à termes complexes ou réels quelconques.

3.1 Convergence absolue.

Considérons une série à termes réels ou complexes $\sum u_n$. En notant $|u_n|$ la valeur absolue (si u_n réel) ou le module (si u_n est complexe) de u_n, on définit une nouvelle série $\sum |u_n|$ qui est, par définition, une série à termes réels positifs (puisque l'on a toujours $|u_n| \in \mathbb{R}_+$). On peut donc appliquer à cette série les résultats de la section 2 précédente. L'intérêt est que le théorème suivant permet dans certains cas favorables d'en déduire la convergence de la série de départ $\sum u_n$.

Définition. *Avec les données ci-dessus, on dit que la série $\sum u_n$ est absolument convergente lorsque la série $\sum |u_n|$ est convergente.*

Théorème. *Si la série $\sum\limits_{n\geq 1} u_n$ est absolument convergente, alors elle est convergente, et l'on a:*

$$\left|\sum_{n=1}^{+\infty} u_n\right| \leq \sum_{n=1}^{+\infty} |u_n|.$$

Exemple. La série $\sum \frac{\mathrm{e}^{in}}{n^2}\, dt$ est absolument convergente, donc convergente.

En effet, puisque $\mathrm{e}^{in} = \cos n + i \sin n$, on a $|\mathrm{e}^{in}| = 1$, de sorte que la suite complexe $u_n = \frac{\mathrm{e}^{in}}{n^2}$ vérifie $|u_n| = \frac{1}{n^2}$. Or on sait (voir 2.2) que la série $\sum \frac{1}{n^2}$ converge; ce qui prouve le résultat voulu.

Attention ! *Erreur fréquente et grave:* le théorème affirme que la convergence absolue entraîne la convergence, mais la réciproque est fausse. En d'autres termes, si $\sum |u_n|$ converge, alors $\sum u_n$ converge, mais on peut avoir $\sum u_n$ qui converge bien que $\sum |u_n|$ diverge.

Considérons par exemple la série dite *harmonique alternée* $\sum \frac{(-1)^n}{n}$. Elle n'est pas absolument convergente car $\sum |\frac{(-1)^n}{n}|$ est la série harmonique $\sum \frac{1}{n}$ dont on sait qu'elle diverge (voir 1.1 exemple 3, ou théorème de 2.2).

Montrons que cependant la série $\sum\limits_{n\geq 1} \frac{(-1)^n}{n}$ converge. Pour cela, on pourait appliquer directement la proposition du 3.2 ci-dessous. On peut donner aussi la preuve directe suivante (qui a l'avantage de donner aussi la valeur de la somme de la série). On part de l'identité, vraie pour tout $x \in \mathbb{R}$ différent de -1:

$$1 - x + x^2 - x^3 + \cdots + (-1)^N x^N = \sum_{i=0}^{N} (-x)^i = \frac{1-(-x)^{N+1}}{1+x} = \frac{1}{1+x} + (-1)^N \frac{x^{N+1}}{1+x}.$$

Si l'on intègre terme à terme les deux membres de cette égalité entre 0 et 1, on obtient:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^N}{N+1} = \ln 2 + (-1)^N \int_0^1 \frac{x^{N+1}}{1+x}\, dx.$$

Or $0 \leq \int_0^1 \frac{x^{N+1}}{1+x}\, dx \leq \int_0^1 x^{N+1}\, dx = \frac{1}{N+2}$ qui tend vers 0 quand $N \to +\infty$, et donc $\lim\limits_{N\to+\infty} \int_0^1 \frac{x^{N+1}}{1+x}\, dx = 0$.

On conclut que $\lim\limits_{N\to+\infty} \sum\limits_{n=1}^{N+1} \frac{(-1)^{n-1}}{n} = \ln 2$, ce qui prouve que la série $\sum\limits_{n\geq 1} \frac{(-1)^n}{n}$ converge et que sa somme vaut $-\ln 2$.

Si la série $\sum |u_n|$ diverge, le théorème ne permet pas de conclure quoi que ce soit sur la nature de la série $\sum u_n$, qui peut soit converger soit diverger.

Terminons en donnant un résultat pratique sur certains types de séries réelles, dites alternées, dont le signe change de terme en terme, et dont l'exemple de la série harmonique alternée traité ci-dessus est un cas particulier.

3.2 Séries (réelles) alternées.

Définition. Une série réelle $\sum u_n$ est dite *alternée* lorsque son terme général est de la forme $u_n = (-1)^n x_n$ avec $x_n \in \mathbb{R}_+$, ou de la forme $u_n = (-1)^{n+1} x_n$ avec $x_n \in \mathbb{R}_+$.

Proposition. *Soit (x_n) une suite de nombre réels positifs. Si la suite (x_n) est décroissante et converge vers 0, alors la série alternée $\sum (-1)^n x_n$ est convergente.*

Exemple (séries de Riemann alternées).

On considère la série $\sum \frac{(-1)^n}{n^\alpha}$, où α est un réel fixé.

- Si $\alpha \leq 0$, elle est divergente; (en effet, son terme général ne tend pas vers 0).
- Si $\alpha > 1$, elle est absolument convergente donc convergente (d'après 2.2 et 3.1).
- Si $0 < \alpha \leq 1$, elle est convergente (d'après la dernière proposition ci-dessus), mais non absolument convergente (d'après 2.2).

Séries entières

1. Notion de série entière

1.1 Définition d'une série entière.

Soit $(a_n)_{n\geq 0}$ une suite de nombres complexes fixée. Pour tout $z \in \mathbb{C}$, on considère la série:

$$\sum_{n\geq 0} a_n z^n.$$

Une série de ce type s'appelle une *série entière*. Pour chaque valeur de z, la série $\sum_{n\geq 0} a_n z^n$ est une série numérique de nombres complexes, pour laquelle se pose donc la question de savoir si elle convergente ou pas, et si oui de calculer éventuellement sa somme:

$$S(z) = \sum_{n=0}^{+\infty} a_n z^n = \lim_{N\to+\infty} (a_0 + a_1 z + a_2 z^2 + \cdots + a_N z^N) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n + \cdots.$$

Notons E le domaine de convergence de la série entière $\sum a_n z^n$, c'est-à-dire l'ensemble de tous les $z \in \mathbb{C}$ pour lesquels elle est convergente, (il est clair que E contient au moins 0). On peut considérer l'application $S : E \to \mathbb{C}$ qui à tout $z \in E$ associe la somme de la série $S(z) = \sum_{n=0}^{+\infty} a_n z^n$, et étudier les propriétés de cette application. Le premier résultat fondamental est le théorème ci-dessous (que l'on rappelle sans démonstration), qui donne des précisions sur la nature de l'ensemble E.

1.2 Rayon de convergence d'une série entière.

Théorème. *Soit $\sum_{n\geq 0} a_n z^n$ une série entière. Il existe un unique élément R qui est, soit un réel positif, soit $+\infty$, tel que l'on ait:*

(i) *Pour tout $z \in \mathbb{C}$ tel que $|z| < R$, la série $\sum a_n z^n$ est absolument convergente, et donc convergente.*

(ii) *Pour tout $z \in \mathbb{C}$ tel que $|z| > R$, la suite $(a_n z^n)$ est non-bornée, et donc la série $\sum a_n z^n$ est divergente.*

Définition. L'élément R ainsi déterminé est appelé le *rayon de convergence* de la série entière.

Remarque 1. Dire que $R = 0$ signifie que la série entière $\sum a_n z^n$ ne converge pour aucune valeur non-nulle de z, c'est-à-dire que son domaine de convergence E est réduit à $\{0\}$. Au contraire, dire que $R = +\infty$ signifie qu'elle converge (et même absolument) pour tout $z \in \mathbb{C}$ tout entier, c'est-à-dire que son domaine de convergence E est $\mathbb{C}$ tout entier.

Remarque 2. Supposons que $0 < R < +\infty$. Considérons dans le plan complexe le disque ouvert $D(0, R) = \{z \in \mathbb{C}\,;\, |z| < R\}$ de centre 0 et de rayon R. Par définition de R, la série $\sum a_n z^n$ converge absolument en tout point de $D(0, R)$, et diverge en tout point extérieur au disque fermé $\overline{D}(0, R) = \{z \in \mathbb{C}\,;\, |z| \leq R\}$.

En revanche le théorème ne dit rien sur ce qui se passe sur le cercle $C(0, R) = \{z \in \mathbb{C}\,;\, |z| = R\}$ de centre 0 et de rayon R. En fait, on verra sur des exemples que ce cercle peut contenir à la fois des points $z \in \mathbb{C}$ pour lesquels la série $\sum a_n z^n$ converge et des points $z \in \mathbb{C}$ pour lesquels elle diverge. Pour cette raison, ce cercle est parfois appelé cercle d'incertitude de la série entière.

Figure 1

1.3 Méthodes de calcul du rayon de convergence.

Exemple 1 (*en utilisant la définition du rayon de convergence*).

Considérons la série entière $\sum \mathrm{e}^{-\sqrt{n}} z^n$. Cherchons son rayon de convergence R.

On a $|\mathrm{e}^{-\sqrt{n}} z^n| = \exp(-\sqrt{n} + n \ln |z|) = \exp[\sqrt{n}(\sqrt{n} \ln |z| - 1)]$.

Supposons $|z| > 1$; donc $\ln |z| > 0$, ce qui implique que $|\mathrm{e}^{-\sqrt{n}} z^n|$ tend vers $+\infty$ quand $n \to +\infty$. On a donc nécessairement $\sum \mathrm{e}^{-\sqrt{n}} z^n$ divergente. Ainsi la série $\sum \mathrm{e}^{-\sqrt{n}} z^n$ diverge pour tout z tel que $|z| > 1$, ce qui prouve d'après le point (i) du théorème 1.2 que $|z| \geq R$. Donc $R \leq 1$.

Supposons $|z| < 1$; donc $\ln|z| < 0$, ce qui implique que $|\,\mathrm{e}^{-\sqrt{n}}\, z^n|$ tend vers 0 quand $n \to +\infty$. En particulier, la suite $(\mathrm{e}^{-\sqrt{n}}\, z^n)$ est bornée. Ce qui implique d'après le point (ii) du théorème 1.2 que $|z| \leq R$. Donc $1 \leq R$. On conclut finalement que $R = 1$.

EXEMPLE 2 (*en utilisant la définition du rayon de convergence*).

Considérons la série entière $\sum(\sin n)z^n$. Cherchons son rayon de convergence R.

On a $|(\sin n)z^n| \leq |z|^n$. Si $|z| < 1$, la série géométrique de raison z^n est convergente. Le critère de majoration sur les séries à termes positifs assure donc que la série de terme général $|(\sin n)z^n|$ est convergente. Ainsi la série de terme général $(\sin n)z^n$ est absolument convergente pour tout z tel que $|z| < 1$. Ceci prouve que $1 \leq R$. Prenons maintenant $z = 1$. La série $\sum \sin n$ diverge car son terme général ne tend pas vers 0. Ainsi, on a trouvé le point $z = 1$ en lequel la série entière $\sum \sin n z^n$ est divergente, ce qui suffit à prouver que $R \leq 1$. On conclut finalement que $R = 1$.

On peut déduire aussi de la règle de d'Alembert sur les séries numériques la méthode de calcul suivante:

PROPOSITION. *Soit $\sum a_n z^n$ une série entière. On suppose que $a_n \neq 0$ à partir d'un certain rang et que:*

$$\lim_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right| = \ell, \quad \text{avec } \ell \in \mathbb{R}_+ \text{ ou } \ell = +\infty.$$

Alors le rayon de convergence de la série $\sum a_n z^n$ est $R = \frac{1}{\ell}$, (avec les conventions $\frac{1}{0} = +\infty$ et $\frac{1}{+\infty} = 0$).

La série entière ci-dessous fournit un exemple d'application de cette proposition.

1.4 Série exponentielle.

Considérons la série entière $\sum \frac{1}{n!} z^n$. Le rapport $\left|\frac{a_{n+1}}{a_n}\right| = \frac{n!}{(n+1)!} = \frac{1}{n+1}$ tend vers la limite $\ell = 0$ quand $n \to +\infty$. Donc, d'après la proposition de 1.3, le rayon de convergence R est $+\infty$. Cette série entière est appelée série exponentielle, notée e^z ou $\exp z$. On retiendra que:

$$\boxed{\mathrm{e}^z = \exp z = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \cdots \quad \text{absolument convergente pour tout } z \in \mathbb{C}}$$

REMARQUE. - D'une façon générale, le produit de deux séries numériques $U = \sum_{n\geq 0} u_n$ et $V = \sum_{n\geq 0} v_n$ est défini comme la série $W = UV = \sum_{n\geq 0} w_n$ de terme général $w_n = u_0 v_n + u_1 v_{n-1} + \cdots + u_n v_0 = \sum_{p+q=n} u_p v_q$, ce qui traduit le calcul naturel:

$$(u_0 + u_1 + u_2 + \cdots)(v_0 + v_1 + v_2 + \cdots) = \underbrace{(u_0 v_0)}_{w_0} + \underbrace{(u_0 v_1 + u_1 v_0)}_{w_1} + \underbrace{(u_0 v_2 + u_1 v_1 + u_2 v_0)}_{w_2} + \cdots$$

On peut montrer que si chacune des deux séries U et V est absolument convergente, alors la série W aussi, et que sa somme est le produit de la somme de U et de la somme de V. D'après le point (i) du théorème 1.2, ceci s'applique donc en particulier au produit de deux séries entières sur leurs disques de convergence.

Prenons z et z' deux nombres complexes. Posons $u_n = \frac{1}{n!} z^n$ et $v_n = \frac{1}{n!} z'^n$. Les séries $U = \exp z = \sum_{n\geq 0} u_n$ et $V = \exp z' = \sum_{n\geq 0} v_n$ sont absolument convergentes. La série produit W est de terme général:

$$w_n = \sum_{p=0}^{n} u_p v_{n-p} = \sum_{p=0}^{n} \frac{1}{p!} z^p \frac{1}{(n-p)!} z'^{n-p} = \frac{1}{n!} \sum_{p=0}^{n} \frac{n!}{p!\,(n-p)!} z^p z'^{n-p} = \frac{1}{n!}(z+z')^n.$$

Ainsi, la série produit W n'est autre que $\exp(z + z')$. On a donc montré, et on retiendra, que:

$$\boxed{(\exp z)(\exp z') = \exp(z + z') \text{ pour tous } z, z' \in \mathbb{C}.}$$

1.5 Série entière géométrique.

Soit $a \in \mathbb{C}^*$. Considérons la série entière $\sum a^n z^n$. C'est la série géométrique $\sum(az)^n$ de raison az, dont on sait qu'elle est convergente si et seulement si $|az| < 1$, et que sa somme est alors $\frac{1}{1-az}$. On en déduit que: *le rayon de convergence de la série entière $\sum a^n z^n$ est $\frac{1}{|a|}$.*

Les cas $a = 1$ et $a = -1$ sont particulièrement utiles dans la pratique. On retiendra que:

$$\boxed{\begin{array}{rl} 1 + z + z^2 + z^3 + \cdots = \displaystyle\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z} & \text{absolument convergente pour tout } z \in \mathbb{C} \text{ tel que } |z| < 1 \\ 1 - z + z^2 - z^3 + \cdots = \displaystyle\sum_{n=0}^{+\infty} (-1)^n z^n = \frac{1}{1+z} & \text{absolument convergente pour tout } z \in \mathbb{C} \text{ tel que } |z| < 1 \end{array}}$$

2.1 Intervalle de convergence.

On considère toujours une suite $(a_n)_{n\geq 0}$ de nombres complexes, mais on se limite ici à considérer la série entière $\sum a_n z^n$ pour des valeurs réelles de z. Dans ce cas, il est habituel de noter x au lieu de z la variable. On note donc cette série:

$$\textstyle\sum_{n\geq 0} a_n x^n.$$

Si l'on appelle R le rayon de convergence de la série entière complexe $\sum a_n z^n$, l'intersection avec $\mathbb{R}$ du disque ouvert $D(0,R)$ du plan complexe est l'intervalle ouvert $]-R,R[$. Le théorème 2.1 implique donc :

- si $R=+\infty$, la série $\sum a_n x^n$ est absolument convergente pour tout $x\in\mathbb{R}=]-\infty,+\infty[$.
- si $0<R<+\infty$, la série $\sum a_n x^n$ est absolument convergente pour $x\in]-R,R[$, et divergente pour $x\in]-\infty,-R[$ et pour $x\in]R,+\infty[$.

Il en résulte que, pour tout $x\in]-R,R[$, (avec R réel strictement positif ou $+\infty$), la somme de la série

$$S(x)=\sum_{n=0}^{+\infty} a_n x^n = a_0+a_1x+a_2x^2+\cdots+a_nx^n+\cdots.$$

définit une application:

$$S:]-R,R[\longrightarrow\mathbb{C}.$$

On rappelle ci-dessous (sans démonstration) certaines de ses propriétés en tant que fonction d'une variable réelle. Notons auparavant que, dans le cas particulier (fréquent pour beaucoup d'applications pratiques) où l'on suppose que les coefficients a_n sont eux-mêmes tous des réels, la fonction S est à valeurs réelles.

2.2 Continuité et dérivabilité.

Théorème. *Avec les données et notations ci-dessus, la fonction S est de classe C^∞ (c'est-à-dire continue et indéfiniment dérivable) sur l'intervalle $]-R,R[$. En outre, on a pour tout $x\in]-R,R[$:*

$$S(x)=\sum_{n=0}^{+\infty} a_nx^n = a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+\cdots+a_nx^n+\cdots$$
$$S'(x)=\sum_{n=1}^{+\infty} na_nx^{n-1} = a_1+2a_2x+3a_3x^2+4a_4x^3+\cdots+na_nx^{n-1}+\cdots$$
$$S''(x)=\sum_{n=2}^{+\infty} n(n-1)a_nx^{n-2} = 2a_2+6a_2x+12a_3x^2+\cdots+n(n-1)a_nx^{n-2}+\cdots$$

etc...

Remarque: on peut montrer qu'une série entire et sa série entire dérivée ont même rayon de convergence.

Exemple 1. Reprenons l'exemple de la série exponentielle traité en 1.4.

Appliquons lui ce qui précède, avec $R=+\infty$, en notant encore exp ou plus simplement e la fonction S. Cette application définie par: $\mathrm{e}^x=\exp x=\sum_{n=0}^{+\infty}\frac{1}{n!}x^n=1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\cdots$, est de classe C^∞ sur $]-\infty,+\infty[$, et sa dérivée est: $(\mathrm{e}^x)'=0+1+2\frac{x}{2}+3\frac{x^2}{6}+4\frac{x^3}{24}+\cdots$. En d'autres termes, $(\mathrm{e}^x)'=\sum_{n=1}^{+\infty} n\frac{1}{n!}x^{n-1}=\mathrm{e}^x$ pour tout $x\in]-\infty,+\infty[$.

On a montré que: $\mathrm{e}^x=(\mathrm{e}^x)'=(\mathrm{e}^x)''=\cdots$ pour tout $x\in\mathbb{R}$.

Exemple 2. Reprenons l'exemple de la série $\sum x^n$ traité en 1.5.

On a $R=1$. La fonction S définie par: $\frac{1}{1-x}=\sum_{n=0}^{+\infty}x^n=1+x+x^2+x^3+x^4+\cdots$ est de classe C^∞ sur $]-1,1[$, et sa dérivée est: $(\frac{1}{1-x})'=0+1+2x+3x^2+4x^3+\cdots=\sum_{n=1}^{+\infty}nx^{n-1}$ pour tout $x\in]-1,1[$. Par ailleurs, le produit (voir 1.4) de la série $\sum x^n$ par elle-même donne:

$$(1+x+x^2+x^3+x^4+\cdots)(1+x+x^2+x^3+x^4+\cdots)=1+2x+3x^2+4x^3+5x^4+\cdots.$$

On a montré que $(\frac{1}{1-x})'=\frac{1}{(1-x)^2}$.

2.3 Application à la résolution de certaines équations différentielles

On cherche à déterminer, pour une équation différentielle donnée, s'il existe des solutions qui sont la somme d'une série entière sur un intervalle $]-R, R[$ de $\mathbb{R}$.

Exemple. Considérons l'équation différentielle $4xy''(x) + 2y'(x) - y(x) = 0$.
On cherche une solution $y(x)$ qui soit la somme d'une série entière $\sum a_n x^n$.
On a $y(x) = \sum_{n=0}^{+\infty} a_n x^n$, $y'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}$, $y''(x) = \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2}$.

$$\begin{aligned}\text{Donc } 4xy''(x) + 2y'(x) - y(x) &= \sum_{n=2}^{+\infty} 4n(n-1)a_n x^{n-1} + \sum_{n=1}^{+\infty} 2n a_n x^{n-1} - \sum_{n=0}^{+\infty} a_n x^n \\ &= \sum_{m=1}^{+\infty} 4(m+1)m a_{m+1} x^m + \sum_{m=0}^{+\infty} 2(m+1) a_{m+1} x^m - \sum_{m=0}^{+\infty} a_m x^m \\ &= (2a_1 - a_0) + \sum_{m=1}^{+\infty} [4(m+1)m a_{m+1} + 2(m+1)a_{m+1} - a_m] x^m.\end{aligned}$$

Il en résulte que $y(x)$ est solution de l'équation différentielle de départ si et seulement si:

$$a_1 = \tfrac{1}{2} a_0 \quad \text{et} \quad a_{n+1} = \tfrac{1}{(2n+1)(2n+2)} a_n \text{ pour tout } n \geq 1.$$

Imposons de plus la condition intiale $y(0) = 1$. Cela signifie que $a_0 = 1$. D'où l'on tire successivement:

$$a_1 = \tfrac{1}{2},\ \ a_2 = \tfrac{1}{3\times 4} \times \tfrac{1}{2} = \tfrac{1}{4!},\ a_3 = \tfrac{1}{5\times 6} \times \tfrac{1}{4!} = \tfrac{1}{6!},\ \text{ et par récurrence } a_n = \tfrac{1}{(2n)!}.$$

On conclut que $y(x) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^n$, dont le rayon de convergence est clairement $+\infty$ par application directe de la proposition de 1.3.

Remarque. – On définit naturellement à partir de la série exponentielle les deux séries entières:

$$\mathrm{ch} z = \tfrac{1}{2}(\mathrm{e}^z + \mathrm{e}^{-z}) = \sum_{n=0}^{+\infty} \tfrac{1}{(2n)!} z^{2n} \quad \text{et} \quad \cos z = \tfrac{1}{2}(\mathrm{e}^{iz} + \mathrm{e}^{-iz}) = \sum_{n=0}^{+\infty} \tfrac{(-1)^n}{(2n)!} z^{2n},$$

dont le rayon de convergence est $+\infty$. Donc, pour tout $x \in \mathbb{R}$, la solution $y(x)$ trouvée s'exprime comme:

$$y(x) = \sum_{n=0}^{+\infty} \tfrac{1}{(2n)!} x^n = \begin{cases} \sum_{n=0}^{+\infty} \frac{1}{(2n)!} (\sqrt{x})^{2n} = \mathrm{ch}\sqrt{x} & \text{si } x \geq 0, \\ \sum_{n=0}^{+\infty} \frac{1}{(2n)!} (-(\sqrt{-x})^2)^n = \cos\sqrt{-x} & \text{si } x \leq 0. \end{cases}$$

Cette dernière remarque nous introduit à la question qui fait l'objet de la partie 3: savoir si, réciproquement à ce qu'on vient de faire dans la partie 2, une fonction donnée est ou non la somme d'une série entière sur un certain intervalle.

3. Développement en série entière d'une fonction d'une variable réelle.

3.1 Fonction développable en série entière.

Définition. Soit f une fonction d'une variable réelle x définie sur un intervalle I de $\mathbb{R}$ contenant 0. On dit que f est développable en série entière centrée en 0 lorsqu'il existe un intervalle $]-\alpha, \alpha[$ centré en 0 inclus dans I, et une série entière $\sum a_n z^n$ de rayon de convergence $\geq \alpha$, tels que:

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n \quad \text{pour tout } x \in]-\alpha, \alpha[.$$

D'après le théorème 2.2, la fonction f est de classe C^∞ sur $]-\alpha, \alpha[$, et l'on a:

$$a_0 = f(0), \quad a_1 = f'(0), \quad a_2 = \tfrac{1}{2} f''(0), \quad \ldots, a_n = \tfrac{1}{n!} f^{(n)}(0), \quad \ldots$$

ce qui prouve en particulier l'unicité du développement en série entière (s'il existe).

Convention terminologique. Dans toute la suite, "développable en série entière" signifiera "développable en série entière centrée en 0".

Pour tout $a \in \mathbb{R}$, une fonction f définie sur un intervalle I de $\mathbb{R}$ contenant a est dite développable en série entière centrée en a lorsque la fonction $x \mapsto f(x+a)$ est développable en série entière centrée en 0; c'est pourquoi on se ramène à l'étude du cas $a = 0$

Remarque. Une condition nécessaire et suffisante pour l'existence d'un développement en série entière est donnée par la formule de Taylor avec reste intégral.

Cette formule établit que, pour $f :]-\alpha,\alpha[\to \mathbb{C}$ de classe C^∞, on a pour tout entier $N \geq 1$ et tout $x \in]-\alpha,\alpha[$:

$$f(x) = f(0) + f'(0)x + \tfrac{1}{2!}f''(0)x^2 + \cdots + \tfrac{1}{N!}f^{(N)}(0)x^N + R_N(x) = \sum_{n=0}^{N} \tfrac{1}{n!}f^{(n)}(0)x^n + R_N(x),$$

où le reste $R_N(x)$ peut s'exprimer sous la forme $R_N(x) = \int_0^x \frac{1}{N!}(x-t)^N f^{(N+1)}(t)\,dt$.

On montre alors que f est développable en série entière si et seulement s'il existe $0 < \beta \leq \alpha$ tel que $\lim\limits_{N\to+\infty} R_N(x) = 0$ pour tout $x \in]-\beta,\beta[$.

Méthodes de calcul. Un raisonnement "théorique" (celui de la remarque précédente, ou d'autres que nous ne détaillons pas ici) permettant d'établir l'existence d'un développement en série entière pour certaines fonctions de références (exponentielle, puissances,...), on se ramène à ces fonctions de références en utilisant les régles suivantes:

1. la somme de deux fonctions f et g développables en série entière est développable en série entière, et le développement de $f+g$ s'obtient en faisant la somme des deux développements;
2. idem pour le produit fg;
3. la dérivée d'une fonction f développable en série entière est développable en série entière, et le développement de f' s'obtient en dérivant terme à terme le développement de f;
4. idem pour une primitive.

3.2 Première série d'exemples.

Le résultat central est que la fonction exponentielle est développable en série entière, avec un rayon de convergence infini:

$$\boxed{\mathrm{e}^x = \exp x = \sum_{n=0}^{+\infty} \tfrac{1}{n!}x^n = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \tfrac{1}{24}x^4 + \cdots \qquad \text{pour tout } x \in]-\infty,+\infty[}$$

Les fonctions hyperboliques étant définies par $\mathrm{ch}x = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x})$ et $\mathrm{sh}x = \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x})$, les résultats sur les sommes de fonctions développables en séries entières donnent:

$$\boxed{\begin{array}{ll} \mathrm{ch}x = \sum_{n=0}^{+\infty} \frac{1}{(2n)!}x^{2n} = 1 + \frac{1}{2}x^2 + \frac{1}{24}x^4 + \frac{1}{720}x^6 + \cdots & \text{pour tout } x \in]-\infty,+\infty[\\ \mathrm{sh}x = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!}x^{2n+1} = x + \frac{1}{6}x^3 + \frac{1}{120}x^5 + \frac{1}{5040}x^7 + \cdots & \text{pour tout } x \in]-\infty,+\infty[\end{array}}$$

De même pour les fonctions trigonométriques $\cos x = \frac{1}{2}(\mathrm{e}^{ix} + \mathrm{e}^{-ix})$ et $\sin x = \frac{1}{2i}(\mathrm{e}^{ix} - \mathrm{e}^{-ix})$:

$$\boxed{\begin{array}{ll} \cos x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!}x^{2n} = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \cdots & \text{pour tout } x \in]-\infty,+\infty[\\ \sin x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1} = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \cdots & \text{pour tout } x \in]-\infty,+\infty[\end{array}}$$

3.3 Seconde série d'exemples.

Un autre résultat central est que la fonction puissance $(1+x)^\alpha$ est développable en série entière pour tout $\alpha \in \mathbb{R}$, avec un rayon de convergence égal à 1 (infini dans le cas particulier où $\alpha \in \mathbb{N}$):

$$\boxed{(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \tfrac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)}{n!}x^n \qquad \text{pour tout } x \in]-1,1[}$$

Pour $\alpha = -1$, on retrouve $\boxed{\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n \qquad \text{pour tout } x \in]-1,1[}$,

et donc en changeant x et $-x$, la formule $\boxed{\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \qquad \text{pour tout } x \in]-1,1[}$.

En prenant leurs primitives s'annulant en 0, il vient

$$\boxed{\ln(1+x) = \sum_{n=1}^{+\infty} \tfrac{(-1)^{n+1}}{n}x^n \quad \text{et} \quad -\ln(1-x) = \sum_{n=1}^{+\infty} \tfrac{1}{n}x^n \qquad \text{pour tout } x \in]-1,1[},$$

dont la demie-somme conduit à $\boxed{\frac{1}{2}\ln(\frac{1+x}{1-x}) = \sum_{n=0}^{+\infty} \frac{1}{2n+1}x^{2n+1} \quad \text{pour tout } x \in]-1,1[}$.

En remplaçant x par x^2 dans $\frac{1}{1+x} = \sum_{n=0}^{+\infty}(-1)^n x^n$ et en prenant la primitive s'annulant en 0, on obtient:

$$\boxed{\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1}x^{2n+1} \quad \text{pour tout } x \in]-1,1[}.$$

Reprenons maintenant la formule de départ avec $\alpha = -\frac{1}{2}$; il vient:

$$\boxed{\frac{1}{\sqrt{1+x}} = 1 + \sum_{n=1}^{+\infty}(-1)^n \frac{1\times3\times5\times\cdots\times(2n-1)}{2\times4\times6\times\cdots\times2n}x^n \quad \text{pour tout } x \in]-1,1[}.$$

En remplaçant x par $-x^2$ en prenant la primitive s'annulant en 0, on obtient:

$$\boxed{\arcsin x = x + \sum_{n=1}^{+\infty} \frac{1\times3\times5\times\cdots\times(2n-1)}{2\times4\times6\times\cdots\times2n}\,\frac{1}{2n+1}x^{2n+1} \quad \text{pour tout } x \in]-1,1[}.$$

Séries de Fourier

1. Coefficients de Fourier

1.1 Forme réelle (ou trigonométrique) des coefficients.

Soit f une fonction $\mathbb{R} \to \mathbb{C}$ que l'on suppose continue par morceaux, et 2π-périodique (ce qui signifie que $f(x+2\pi) = f(x)$ pour tout $x \in \mathbb{R}$.

Définition. Pour tout $n \in \mathbb{N}$, on définit les *coefficients de Fourier réels* de f par:

$$\boxed{a_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos(nt)\,dt} \qquad \text{et} \qquad \boxed{b_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(nt)\,dt} \; .$$

Remarques.
1. On a toujours $b_0(f) = 0$.
2. Dans le cas particulier où f est paire, on a $b_n(f) = 0$ et $a_n(f) = \frac{2}{\pi}\int_0^{\pi} f(t)\cos(nt)\,dt$, pour tout $n \in \mathbb{N}$.
3. Dans le cas particulier où f est impaire, on a $a_n(f) = 0$ et $b_n(f) = \frac{2}{\pi}\int_0^{\pi} f(t)\sin(nt)\,dt$, pour tout $n \in \mathbb{N}$.

1.2 Forme complexe (ou exponentielle) des coefficients.

Soit f une fonction $\mathbb{R} \to \mathbb{C}$ continue par morceaux et 2π-périodique.

Définition. Pour tout $n \in \mathbb{Z}$, on définit le *coefficient de Fourier complexe* de f par:

$$\boxed{c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\,\mathrm{e}^{-int}\,dt} \; .$$

Attention ! Les coefficients $a_n(f)$ et $b_n(f)$ sont définis pour tout entier n positif, alors que les coefficients $c_n(f)$ sont définis pour tout n entier positif ou négatif.

Puisque $\mathrm{e}^{int} = \cos(nt) + i\sin(nt)$ et $\mathrm{e}^{-int} = \cos(nt) - i\sin(nt)$, on déduit immédiatement les relations:

$$\begin{cases} c_n = \frac{1}{2}(a_n - ib_n) \\ c_{-n} = \frac{1}{2}(a_n + ib_n), \end{cases} \qquad \begin{cases} a_n = c_n + c_{-n} \\ b_n = i(c_n - c_{-n}), \end{cases} \qquad \text{d'où} \; \begin{cases} c_n = c_{-n} = \frac{1}{2}a_n & \text{si } f \text{ est paire,} \\ c_n = -c_{-n} = \frac{1}{2i}b_n & \text{si } f \text{ est impaire.} \end{cases}$$

Remarques calculatoires.

1. Les coefficients de Fourier sont linéaires. Cela signifie que, si f et g sont deux fonctions continues par morceaux et 2π-périodiques sur $\mathbb{R}$, on a $a_n(f+g) = a_n(f) + a_n(g)$, $b_n(f+g) = b_n(f) + b_n(g)$, $c_n(f+g) = c_n(f) + c_n(g)$; et si $\lambda \in \mathbb{C}$, $a_n(\lambda.f) = \lambda.a_n(f)$, $b_n(\lambda.f) = \lambda.b_n(f)$, $c_n(\lambda.f) = \lambda.c_n(f)$.
2. En désignant par $\overline{f}$ la conjuguée de f, on a: $a_n(\overline{f}) = \overline{a_n(f)}$, $b_n(\overline{f}) = \overline{b_n(f)}$, $c_n(\overline{f}) = \overline{c_{-n}(f)}$.
3. Si f est continue sur $\mathbb{R}$, C^1 par morceaux, et 2π-périodique, alors f' peut être prolongée en une fonction continue par morceaux 2π-périodique, dont on peut considérer les coefficients de Fourier. On a alors: $c_n(f') = in\,c_n(f)$ pour tout $n \in \mathbb{Z}$.

   En effet, $c_n(f') = \frac{1}{2\pi}\int_0^{2\pi} f'(t)\,\mathrm{e}^{-int}\,dt$. Une intégration par parties donne:
   $c_n(f') = \frac{1}{2\pi}[f(t)\,\mathrm{e}^{-int}]_0^{2\pi} - \frac{1}{2\pi}\int_0^{2\pi} f(t)(-in)\,\mathrm{e}^{-int}\,dt = \frac{in}{2\pi}\int_0^{2\pi} f(t)\,\mathrm{e}^{-int}\,dt = in\,c_n(f)$.

1.3 Série de Fourier.

Soit f une fonction $\mathbb{R} \to \mathbb{C}$ continue par morceaux et 2π-périodique.

Définition. Pour tout $x \in \mathbb{R}$, on appelle *série de Fourier* de f en x la série:

$$\boxed{\mathrm{SF}_f(x) = \sum_{n\in\mathbb{Z}} c_n(f)\,\mathrm{e}^{inx} = \tfrac{1}{2}a_0(f) + \sum_{n\geq 1}\big(a_n(f)\cos(nx) + b_n(f)\sin(nx)\big)}$$

1.4 Exemples.

Exemple 1. (Fonction en créneaux)

Soit $f : \mathbb{R} \to \mathbb{R}$ impaire, 2π-périodique, telle que $f(t) = 1$ si $0 < t < \pi$ et $f(0) = f(\pi) = 0$.
Il est clair que f est continue par morceaux sur $\mathbb{R}$.

Comme f est impaire, $a_n(f) = 0$ pour tout $n \in \mathbb{N}$,

et $b_n(f) = \frac{2}{\pi}\int_0^\pi f(t)\sin(nt)\,dt = \frac{2}{\pi}\int_0^\pi \sin(nt)\,dt$

$= \frac{2}{\pi}\left[-\frac{1}{n}\cos(nt)\right]_0^\pi = \frac{2}{n\pi}(1-(-1)^n).$

Figure 1

Ainsi $b_n(f) = \begin{cases} 0 & \text{si } n \text{ pair} \\ \frac{4}{n\pi} & \text{si } n \text{ impair} \end{cases}$. On conclut que:

$$\mathrm{SF}_f(x) = \sum_{n\geq 1} b_n(f)\sin(nx) = \sum_{p\geq 0} \frac{4}{(2p+1)\pi}\sin(2p+1)x.$$

Exemple 2. (Fonction en dents de scie)

Soit $g : \mathbb{R} \to \mathbb{R}$ paire, 2π-périodique, telle que $g(t) = t$ si $0 \leq t \leq \pi$.
Il est clair que g est continue sur $\mathbb{R}$.

Comme g est paire, on a $b_n(g) = 0$ pour tout $n \in \mathbb{N}$,

et $a_n(g) = \frac{2}{\pi}\int_0^\pi g(t)\cos(nt)\,dt = \frac{2}{\pi}\int_0^\pi t\cos(nt)\,dt$.

Si $n = 0$, $a_0(g) = \frac{2}{\pi}\int_0^\pi t\,dt = \pi$.

Figure 2

Si $n \geq 1$, on intègre par parties:

$$a_n(g) = \frac{2}{\pi}\Big(\ \left[\frac{1}{n}t\sin(nt)\right]_0^\pi - \int_0^\pi \frac{1}{n}\sin(nt)\,dt\ \Big) = 0 - \frac{2}{n\pi}\int_0^\pi \sin(nt)\,dt = \frac{2}{n\pi}\left[\frac{1}{n}\cos(nt)\right]_0^\pi = \frac{2}{n^2\pi}\big((-1)^n - 1\big).$$

Ainsi $a_n(g) = \begin{cases} 0 & \text{si } n \text{ pair, } n \neq 0 \\ -\frac{4}{n^2\pi} & \text{si } n \text{ impair} \end{cases}$. On conclut que:

$$\mathrm{SF}_g(x) = \frac{1}{2}a_0(g) + \sum_{n\geq 1} a_n(g)\cos(nx) = \frac{\pi}{2} - \sum_{p\geq 0}\frac{4}{(2p+1)^2\pi}\cos(2p+1)x.$$

Exemple 3. (Avec coefficients complexes)

Soit $h : \mathbb{R} \to \mathbb{R}$ paire, 2π-périodique, telle que $h(t) = \mathrm{e}^{-t}$ si $0 \leq t \leq \pi$.
Il est clair que h est continue sur $\mathbb{R}$ (et coïncide avec l'exponentielle sur $[-\pi, 0]$).

Figure 3

$$c_n(h) = \frac{1}{2\pi}\int_0^{2\pi} h(t)\,\mathrm{e}^{-int}\ dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} h(t)\,\mathrm{e}^{-int}\ dt$$

$$= \frac{1}{2\pi}\Big(\ \int_{-\pi}^0 \mathrm{e}^t\,\mathrm{e}^{-int}\ dt + \int_0^\pi \mathrm{e}^{-t}\,\mathrm{e}^{-int}\ dt\ \Big)$$

$$= \frac{1}{2\pi}\Big(\ \left[\frac{1}{1-in}\mathrm{e}^{(1-in)t}\right]_{-\pi}^0 + \left[\frac{1}{-1-in}\mathrm{e}^{(-1-in)t}\right]_0^\pi\ \Big) = \frac{1}{2\pi}\Big(\ \frac{1}{1-in} - \frac{1}{1-in}\mathrm{e}^{-\pi}\,\mathrm{e}^{in\pi} + \frac{1}{1+in} - \frac{1}{1+in}\mathrm{e}^{-\pi}\,\mathrm{e}^{-in\pi}\ \Big).$$

$$= \frac{1}{2\pi}\Big(\ \frac{2}{1+n^2} - \mathrm{e}^{-\pi}(-1)^n\big(\frac{1}{1-in} + \frac{1}{1+in}\big)\Big) = \frac{1}{(1+n^2)\pi}(1 + (-1)^{n+1}\,\mathrm{e}^{-\pi}).$$

On conclut que:

$$\mathrm{SF}_h(x) = \sum_{n\in\mathbb{Z}} c_n(h)\,\mathrm{e}^{inx} = \sum_{n\in\mathbb{Z}}\frac{1}{(1+n^2)\pi}(1 + (-1)^{n+1}\,\mathrm{e}^{-\pi})\,\mathrm{e}^{inx}.$$

Remarquons que, comme h est paire, on a $b_n(h) = 0$ et $a_n(h) = 2c_n(h) = \frac{2}{(1+n^2)\pi}(1 + (-1)^{n+1}\,\mathrm{e}^{-\pi})$, d'où l'on déduit une autre écriture équivalente de cette série de Fourier:

$$\mathrm{SF}_h(x) = \frac{1}{2}a_0(h) + \sum_{n\geq 1} a_n(h)\cos(nx) = \frac{1-\mathrm{e}^{-\pi}}{\pi} + \sum_{n\geq 1}\frac{2}{\pi(1+n^2)}(1 + (-1)^{n+1}\,\mathrm{e}^{-\pi})\cos(nx).$$

Soit $f : \mathbb{R} \to \mathbb{C}$ continue par morceaux et 2π-périodique. Pour tout $x \in \mathbb{R}$, on a défini la série de Fourier $\mathrm{SF}_f(x)$. La première question qui se pose est de savoir si cette série est convergente, pour certaines valeurs de x, voire pour tout $x \in \mathbb{R}$. Si c'est le cas, elle définit une nouvelle application $\mathrm{SF}_f : x \mapsto \mathrm{SF}_f(x)$ de $\mathbb{R}$ dans $\mathbb{C}$, et la seconde question qui se pose est celle des liens entre cette application SF et l'application f de départ. On va voir que, dans les cas les plus favorables, elle est égale à f. On rappelle ci-dessous (sans démonstration) deux résultats utiles relativement à ces questions.

2.1 Théorème de Dirichlet

Soit $f : \mathbb{R} \to \mathbb{C}$ continue par morceaux et 2π-périodique.
Pour tout $x \in \mathbb{R}$, la fonction f admet donc une limite à droite en x, que l'on note parfois $f(x^+)$, et une limite à gauche en x, notée $f(x^-)$.

Dans le cas où x est un point de continuité de la fonction f, on a par définition de la continuité: $f(x^+) = f(x^-) = f(x)$.

En général, on considère la moyenne des limites à gauche et à droite, appelée parfois régularisée de f, définie par: $\widetilde{f}(x) = \frac{1}{2}(f(x^+) + f(x^-))$.

Si x est un point de continuité, alors on a $\widetilde{f}(x) = f(x)$.

Figure 4

Théorème. *Soit $f : \mathbb{R} \to \mathbb{C}$ continue par morceaux et 2π-périodique. On suppose de plus que f est de classe C^1 par morceaux sur $\mathbb{R}$. Alors:*

(i) *pour tout $x \in \mathbb{R}$, la série de Fourier de f en x est convergente, et l'on a: $\mathrm{SF}_f(x) = \frac{1}{2}(f(x^+) + f(x^-))$;*

(ii) *en particulier, en tout $x \in \mathbb{R}$ où f est continue, on a: $\mathrm{SF}_f(x) = f(x)$.*

Remarque. Sous les hypothèses de ce théorème, la série de fonctions SF_f converge donc simplement sur $\mathbb{R}$. On peut montrer que, si l'on suppose de plus que f est continue sur $\mathbb{R}$, alors la série de fonctions SF_f converge normalement (et donc uniformément) sur $\mathbb{R}$.

Exemple 1. Reprenons l'exemple 1 de 1.4. Le théorème de Dirichlet s'applique.

En tout $x \in \mathbb{R}$, (multiple ou non de π), on a: $f(x) = \sum_{p=0}^{\infty} \frac{4}{(2p+1)\pi} \sin(2p+1)x$.

Application: si l'on choisit $x = \frac{\pi}{2}$, on obtient $1 = \sum_{p\geq 0} \frac{4}{(2p+1)\pi} \sin\frac{(2p+1)\pi}{2} = \sum_{p\geq 0} \frac{4}{(2p+1)\pi}(-1)^p$.

On en déduit que: $\boxed{\sum_{p=0}^{+\infty} \frac{(-1)^p}{2p+1} = \frac{\pi}{4}}$

Exemple 2. Reprenons l'exemple 2 de 1.4. Le théorème de Dirichlet s'applique.

En tout $x \in \mathbb{R}$, on a: $g(x) = \frac{\pi}{2} - \sum_{p=0}^{+\infty} \frac{4}{(2p+1)^2\pi} \cos(2p+1)x$.

Application: si l'on choisit $x = 0$, on obtient $0 = \frac{\pi}{2} - \sum_{p=0}^{+\infty} \frac{4}{(2p+1)^2\pi}$. On en déduit que: $\boxed{\sum_{p=0}^{+\infty} \frac{1}{(2p+1)^2} = \frac{\pi^2}{8}}$

2.2 Formules de Plancherel-Parseval.

Théorème. *Soit $f : \mathbb{R} \to \mathbb{C}$ continue par morceaux et 2π-périodique. On suppose de plus que f vérifie $f(x) = \frac{1}{2}(f(x^+) + f(x^-))$ pour tout $x \in \mathbb{R}$. Alors:*

(i) *la série de réels positifs $\sum(|c_n(f)|^2 + |c_{-n}(f)|^2)$ est convergente, et l'on a:*

$$\boxed{|c_0(f)|^2 + \sum_{n=1}^{+\infty}(|c_n(f)|^2 + |c_{-n}(f)|^2) = \frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\, dt} \;;$$

(ii) *lorsque f est à valeurs réelles, la série de réels positifs $\sum(a_n(f)^2 + b_n(f)^2)$ est convergente, et l'on a:*

$$\boxed{\frac{a_0(f)^2}{4} + \frac{1}{2}\sum_{n=1}^{+\infty}(a_n(f)^2 + b_n(f)^2) = \frac{1}{2\pi}\int_0^{2\pi} f(t)^2\, dt} \;.$$

Exemple 1. Reprenons l'exemple 1 de 1.4. On a $a_n(f) = 0$, $b_{2p}(f) = 0$ et $b_{2p+1}(f) = \frac{4}{(2p+1)\pi}$.
La formule (ii) du théorème précédent donne donc $\frac{1}{2}\sum\limits_{p=0}^{+\infty}(\frac{4}{(2p+1)\pi})^2 = \frac{1}{2\pi}\int_0^{2\pi} f(t)^2\,dt = \frac{1}{2\pi}\int_0^{2\pi} dt = 1$,
d'où l'on déduit que $\boxed{\displaystyle\sum_{p=0}^{+\infty}\frac{1}{(2p+1)^2} = \frac{\pi^2}{8}}$.

Remarque. On en tire que: $\sum\limits_{n=1}^{+\infty}\frac{1}{n^2} = \sum\limits_{p=1}^{+\infty}\frac{1}{(2p)^2} + \sum\limits_{p=0}^{+\infty}\frac{1}{(2p+1)^2} = \frac{1}{4}\sum\limits_{p=1}^{+\infty}\frac{1}{p^2} + \frac{\pi^2}{8}$,
d'où $(1-\frac{1}{4})\sum\limits_{n=1}^{+\infty}\frac{1}{n^2} = \frac{\pi^2}{8}$, et finalement $\boxed{\displaystyle\sum_{n=1}^{+\infty}\frac{1}{n^2} = \frac{\pi^2}{6}}$.

De même: $\sum\limits_{n=1}^{+\infty}\frac{(-1)^n}{n^2} = \sum\limits_{p=1}^{+\infty}\frac{1}{(2p)^2} - \sum\limits_{p=0}^{+\infty}\frac{1}{(2p+1)^2} = \frac{1}{4}\sum\limits_{p=1}^{+\infty}\frac{1}{p^2} - \frac{\pi^2}{8} = \frac{1}{4}\frac{\pi^2}{6} - \frac{\pi^2}{8}$, d'où $\boxed{\displaystyle\sum_{n=1}^{+\infty}\frac{(-1)^n}{n^2} = -\frac{\pi^2}{12}}$.

Exemple 2. Reprenons l'exemple 2 de 1.4. On a:

$$b_n(g) = 0,\ a_0(g) = \pi,\ a_{2p}(g) = 0 \text{ si } p \geq 1,\ a_{2p+1}(g) = -\frac{4}{(2p+1]^2\pi}.$$

La formule (ii) du théorème précédent donne donc:

$$\frac{\pi^2}{4} + \frac{1}{2}\sum_{p=0}^{+\infty}(\frac{4}{(2p+1)^2\pi})^2 = \frac{1}{2\pi}\int_0^{2\pi} g(t)^2\,dt = \frac{1}{\pi}\int_0^{\pi} t^2\,dt = \frac{\pi^2}{3},$$

d'où l'on déduit que $\frac{8}{\pi^2}\sum\limits_{p=0}^{+\infty}\frac{1}{(2p+1)^4} = \frac{\pi^2}{3} - \frac{\pi^2}{4}$, et finalement $\boxed{\displaystyle\sum_{p=0}^{+\infty}\frac{1}{(2p+1)^4} = \frac{\pi^4}{96}}$.

Remarque. On en tire que: $\sum\limits_{n=1}^{+\infty}\frac{1}{n^4} = \sum\limits_{p=1}^{+\infty}\frac{1}{(2p)^4} + \sum\limits_{p=0}^{+\infty}\frac{1}{(2p+1)^4} = \frac{1}{16}\sum\limits_{p=1}^{+\infty}\frac{1}{p^4} + \frac{\pi^4}{96}$,
d'où $(1-\frac{1}{16})\sum\limits_{n=1}^{+\infty}\frac{1}{n^4} = \frac{\pi^4}{96}$, et finalement $\boxed{\displaystyle\sum_{n=1}^{+\infty}\frac{1}{n^4} = \frac{\pi^4}{90}}$.

Fonctions de plusieurs variables

On procède en deux étapes: d'abord l'étude des fonctions scalaires de plusieurs variables réelles (deux ou trois variables dans la pratique), c'est-à-dire des fonctions $\mathbb{R}^2 \to \mathbb{R}$ ou $\mathbb{R}^3 \to \mathbb{R}$; puis celle des fonctions vectorielles, c'est-à-dire des fonctions de $\mathbb{R}^2$ ou $\mathbb{R}^3$ à valeurs dans $\mathbb{R}^2$ ou $\mathbb{R}^3$.

1. Fonctions scalaires de plusieurs variables

1.1 Fonction scalaire de deux variables réelles.

• Définition: une fonction scalaire de deux variables réelles est une application:

$$\boxed{\begin{array}{rrcl} f: & U & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & f(x,y) \end{array}},$$

où U est une partie de $\mathbb{R}^2$. On a une notion naturelle de *domaine de définition*, qui est la partie de $\mathbb{R}^2$ formée des couples (x,y) pour lesquels l'expression $f(x,y)$ est bien définie.

Exemple: $(x,y) \mapsto \frac{xy}{x^2+y^2-1}$ est définie sur $U = \mathbb{R}^2 - C$, où $C = \{\ (x,y) \in \mathbb{R}^2\ ;\ x^2+y^2 = 1\ \}$ est le cercle unité.

Exemple: $(x,y) \mapsto \frac{1}{(x-y)}$ est définie sur $U = \mathbb{R}^2 - D$, où $D = \{\ (x,y) \in \mathbb{R}^2\ ;\ x = y\}$ est la première bissectrice.

• Applications partielles. Soit $f : U \to \mathbb{R}$ avec U partie de $\mathbb{R}^2$. Fixons un élément $(a,b) \in U$. On appelle première application partielle associée à f au point (a,b), notée $f_1 = f(\,.\,,b)$ la fonction d'une variable réelle x obtenue en fixant b dans l'expression de $f(x,y)$. On définit de même $f_2 = f(a,\,.\,)$ qui est une fonction de la seule variable y.

$$f_1 = f(\,.\,,b):\ x \mapsto f(x,b), \qquad f_2 = f(a,\,.\,):\ y \mapsto f(a,y).$$

L'application f_1 est définie sur la partie de $\mathbb{R}$ formée des $x \in \mathbb{R}$ tels que $(x,b) \in U$, et f_2 sur la partie de $\mathbb{R}$ formée des $y \in \mathbb{R}$ tels que $(a,y) \in U$. Le rôle de ces applications est très important comme on va le voir dans la suite.

• Opérations sur les fonctions à deux variables. Si $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ sont deux fonctions définies sur une même partie U de $\mathbb{R}^2$, on définit la somme $f+g$ et le produit fg, ainsi que le quotient $\frac{f}{g}$ si de plus g ne s'annule pas sur U:

$$\begin{array}{rrcl} f+g: & U & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & f(x,y)+g(x,y) \end{array}, \quad \begin{array}{rrcl} fg: & U & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & f(x,y)g(x,y) \end{array}, \quad \begin{array}{rrcl} \frac{f}{g}: & U & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & \frac{f(x,y)}{g(x,y)} \end{array}$$

• Projections. Les projections sont par définition les deux applications:

$$\begin{array}{rrcl} p_1: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & x \end{array}, \qquad \begin{array}{rrcl} p_2: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & y \end{array}.$$

A partir des projections, on construit par somme et produit les *fonctions polynomiales en deux variables*, qui sont les fonctions (définies sur $\mathbb{R}^2$) du type $f(x,y) = \sum\limits_{i=0}^{n}\sum\limits_{j=0}^{m} \alpha_{i,j}x^iy^j$, où $\alpha_{i,j} \in \mathbb{R}$ pour tous i,j.

1.2 Continuité.

• Distance dans $\mathbb{R}^2$.

Soit $X = (x,y)$ et $A = (a,b)$ deux éléments quelconques de $\mathbb{R}^2$, la *distance* (euclidienne) entre X et A est par définition:

$$||X - A|| = \sqrt{(x-a)^2 + (y-b)^2}.$$

Pour tout $A \in \mathbb{R}^2$ et tout réel $\varepsilon > 0$, on appelle *boule ouverte* (ou *disque ouvert*) de centre A et de rayon ε l'ensemble:

$$B(A,\varepsilon) = \{\ X \in \mathbb{R}^2\ ;\ ||X - A|| < \varepsilon\ \}.$$

Par définition, on dit qu'un sous-ensemble U de $\mathbb{R}^2$ est un *ouvert* de $\mathbb{R}^2$ lorsque, pour tout $A \in U$, il existe une boule ouverte de centre A qui est entièrement incluse dans U.

Par exemple, une boule ouverte est un ouvert; un demi-plan ouvert est ouvert (prendre $\{ (x,y) \in \mathbb{R}^2 \ ; \ x > 0 \}$ et faire un dessin). En revanche une droite, ou un demi-plan fermé ne sont pas des ouverts (considérer par exemple $\{ (x,y) \in \mathbb{R}^2 \ ; \ x \geq 0 \}$, prendre un point sur le bord et faire un dessin).

un ouvert de $\mathbb{R}^2$

• Définition de la continuité. On définit la continuité de la même façon que dans le cas d'une fonction numérique d'une variable réelle, en remplaçant simplement la valeur absolue par la norme. Une fonction $f : U \to \mathbb{R}$ avec U partie de $\mathbb{R}^2$ est dite *continue en un point* (a,b) de U lorsque: $\lim\limits_{(x,y)\to(a,b)} f(x,y) = f(a,b)$, ce qui signifie que:

$$\forall\, \varepsilon > 0,\ \exists\, \eta > 0,\ \forall\, (x,y) \in U,\ ||(x,y) - (a,b)|| < \eta \ \Rightarrow\ ||f(x,y) - f(a,b)|| < \varepsilon.$$

On dit que f est continue sur U lorsqu'elle continue en tout point de U.
Exemple: les fonctions polynomiales sont continues sur $\mathbb{R}^2$.

• Propriétés.

(i) *Si f et g sont continues sur une même partie U de $\mathbb{R}^2$, alors $f+g$ et fg sont continues sur U.*
Si de plus g ne s'annule pas sur U, alors $\frac{f}{g}$ est continue sur U.

(ii) *Si f est continue sur une partie U de $\mathbb{R}^2$, alors la fonction $|f|$ est aussi continue sur U.*

(iii) *Si f est continue sur une partie U de $\mathbb{R}^2$, et si h est une fonction d'une variable réelle continue sur un intervalle I de $\mathbb{R}$ contenant l'image $f(U)$ de U par f, alors $h \circ f$ est continue sur U.*

• Dans la pratique, ces 3 propriétés, combinées avec les résultats connus sur la continuité des fonctions d'une variable, permettent d'établir la continuité de la plupart des fonctions usuellement rencontrées.

Exemple. Considérons $f(x,y) = \frac{\sin xy}{x-y} + \mathrm{e}^{x+y}$, définie sur $U = \mathbb{R}^2 - \Delta$, où Δ est la droite $\{(x,x) \ ; \ x \in \mathbb{R}\}$.

On décompose en $f = f_1 + f_2$, où $f_1(x,y) = \frac{\sin xy}{x-y}$ et $f_2(x,y) = \mathrm{e}^{x+y}$.

Il est clair d'après (iii) que f_2 est continue sur $\mathbb{R}^2$ comme composée de la fonction polynomiale de deux variables $(x,y) \mapsto x+y$ et de la fonction d'une variable $t \mapsto \mathrm{e}^t$.

Toujours d'après (iii), la fonction $(x,y) \mapsto \sin xy$ est continue sur $\mathbb{R}^2$ comme composée de la fonction polynomiale de deux variables $(x,y) \mapsto xy$ et de la fonction d'une variable $t \mapsto \sin t$. En prenant son quotient par la fonction $(x,y) \to x-y$ continue sur $\mathbb{R}^2$ et non-nulle sur U, on déduit du second point de (i) que f_1 est continue sur U.

Ainsi f_1 et f_2 sont continues sur U, donc $f = f_1 + f_2$ aussi.

• Attention ! Notons pour mémoire (sans insister ni donner plus détails ici) qu'il ne suffit pas a priori que les deux applications partielles $x \to f(x,b)$ et $y \to f(a,y)$ associées à f au point (a,b) soient continues en a et b respectivement pour que f soit continue en (a,b).

1.3 Dérivées partielles du premier ordre.

• Définitions.

Fixons $f : U \to \mathbb{R}$ une fonction de deux variables définie sur un ouvert U de $\mathbb{R}^2$. Les applications partielles $f_1 = f(\,.,b)$ et $f_2 = f(a,\,.)$ sont définies au moins sur un voisinage de a ou de b respectivement.

On dit que f admet une dérivée partielle par rapport à la première variable au point (a,b) si l'application partielle f_1 est dérivable en a.

On note alors: $\frac{\partial f}{\partial x}(a,b) = f_1'(a)$.

On dit que f admet une dérivée partielle par rapport à la seconde variable au point (a,b) si l'application partielle f_2 est dérivable en b.

On note alors: $\frac{\partial f}{\partial y}(a,b) = f_2'(b)$.

Si les dérivées partielles par rapport à chaque variable existent en tout point $(a,b) \in U$, on définit deux nouvelles applications de deux variables:

$$\begin{array}{rcl} \frac{\partial f}{\partial x} : \quad U & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \frac{\partial f}{\partial x}(x,y) \end{array} , \qquad \begin{array}{rcl} \frac{\partial f}{\partial y} : \quad U & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \frac{\partial f}{\partial y}(x,y) \end{array} ,$$

appelées les *fonctions dérivées partielles d'ordre 1* de f. On les calcule simplement en considérant l'une des variables fixée et en dérivant par rapport à l'autre.

Exemple. Si $f(x,y) = x^2 \sin y - y\,\mathrm{e}^{3x}$, on a $\frac{\partial f}{\partial x}(x,y) = 2x \sin y - 3y\,\mathrm{e}^{3x}$ et $\frac{\partial f}{\partial y}(x,y) = x^2 \cos y - \mathrm{e}^{3x}$.

On dit que f est *de classe C^1 sur U* lorsque les deux fonctions dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont définies en tout point de U, et sont continues sur U. On démontre que:

$$f \text{ de classe } C^1 \text{ sur } U \;\Rightarrow\; f \text{ continue sur } U,$$

et les propriétés pratiques (i), (ii), (iii) données en 1.2 sur la continuité restent vraies si l'on remplace dans les énoncés "continue" par "de classe C^1".

• Première formule de composition (ou de changement de variables).

On se donne une fonction de deux variables $f : U \to \mathbb{R}$ de classe C^1 sur un ouvert U de $\mathbb{R}^2$, et deux fonctions d'une variable $u : I \to \mathbb{R}$ et $v : I \to \mathbb{R}$ de classe C^1 sur un intervalle ouvert de $\mathbb{R}$. On suppose que $(u(t), v(t)) \in U$ pour tout $t \in I$, ce qui permet de considérer par composition la fonction d'une variable:

$$\begin{array}{rcl} F : I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & f(u(t), v(t)) \end{array}.$$

On peut démontrer qu'alors F est de classe C^1 sur I, et que sa fonction dérivée $F' = \frac{dF}{dt}$ vérifie:

$$\text{pour tout } a \in I, \quad F'(a) = \frac{\partial f}{\partial u}(u(a), v(a)).\, u'(a) + \frac{\partial f}{\partial v}(u(a), v(a)).\, v'(a).$$

On synthétise ce résultat sous la forme de la formule suivante:

$$\text{Pour } \left\{ \begin{array}{rcl} F : I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & f(u(t), v(t)) \end{array} \right. \text{ de classe } C^1, \text{ on a: } \boxed{\frac{dF}{dt} = \frac{\partial f}{\partial u}\frac{du}{dt} + \frac{\partial f}{\partial v}\frac{dv}{dt}}$$

Exemple. Soit $F(t) = (\mathrm{e}^t + t)(t^2 - 3t)$ de classe C^1 sur $I = \mathbb{R}$. On introduit:

$$\begin{array}{rcl} f : U = \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (u,v) & \longmapsto & uv \end{array} , \qquad \begin{array}{rcl} u : R & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \mathrm{e}^t + t \end{array} , \qquad \begin{array}{rcl} v : R & \longrightarrow & \mathbb{R} \\ t & \longmapsto & t^2 - 3t \end{array} ,$$

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

$$\begin{aligned} \frac{dF}{dt}(a) &= \frac{\partial f}{\partial u}(u(a), v(a)).\, (\mathrm{e}^a + 1) + \frac{\partial f}{\partial v}(u(a), v(a)).\, (2a - 3) \\ &= v(a).\, (\mathrm{e}^a + 1) + u(a).\, (2a - 3) = (a^2 - 3a)(\mathrm{e}^a + 1) + (\mathrm{e}^a + a)(2a - 3), \end{aligned}$$

qui n'est autre dans cet exemple que la formule usuelle de dérivation d'un produit $(uv)' = u'v + uv'$.

• Seconde formule de composition (ou de changement de variables).

On se donne une fonction de deux variables $f : U \to \mathbb{R}$ de classe C^1 sur un ouvert U de $\mathbb{R}^2$, et deux fonctions de deux variables $u : V \to \mathbb{R}$ et $v : V \to \mathbb{R}$ de classe C^1 sur un ouvert V de $\mathbb{R}^2$. On suppose que $(u(x,y), v(x,y)) \in U$ pour tout $(x,y) \in V$, ce qui permet de considérer par composition la fonction de deux variables:

$$\begin{array}{rcl} F : \quad V & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(u(x,y), v(x,y)) \end{array}.$$

On peut démontrer qu'alors F est de classe C^1 sur V, et que ses fonctions dérivées partielles vérifient:

$$\text{pour tout } (a,b) \in V, \quad \left\{ \begin{array}{l} \frac{\partial F}{\partial x}(a,b) = \frac{\partial f}{\partial u}(u(a,b), v(a,b)).\, \frac{\partial u}{\partial x}(a,b) + \frac{\partial f}{\partial v}(u(a,b), v(a,b)).\, \frac{\partial v}{\partial x}(a,b), \\ \\ \frac{\partial F}{\partial y}(a,b) = \frac{\partial f}{\partial u}(u(a,b), v(a,b)).\, \frac{\partial u}{\partial y}(a,b) + \frac{\partial f}{\partial v}(u(a,b), v(a,b)).\, \frac{\partial v}{\partial y}(a,b). \end{array} \right.$$

On synthétise ce résultat sous la forme de la formule suivante:

$$\text{Pour } \left\{ \begin{array}{rcl} F : \quad V & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(u(x,y), v(x,y)) \end{array} \right. \text{ de classe } C^1, \text{ on a: } \boxed{\begin{array}{l} \frac{\partial F}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} \\ \\ \frac{\partial F}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} \end{array}}$$

• Exemple: passage en coordonnées polaires.

Si f est une fonction de deux variables (x, y),
et si l'on note $F(\rho, \theta) = f(\rho\cos\theta, \rho\sin\theta)$, alors:

$$\begin{cases} \frac{\partial F}{\partial \rho}(\rho,\theta) = \frac{\partial f}{\partial x}(\rho\cos\theta,\rho\sin\theta).\, \frac{\partial x}{\partial \rho}(\rho,\theta) + \frac{\partial f}{\partial y}(\rho\cos\theta,\rho\sin\theta).\, \frac{\partial y}{\partial \rho}(\rho,\theta) \\ \frac{\partial F}{\partial \theta}(\rho,\theta) = \frac{\partial f}{\partial x}(\rho\cos\theta,\rho\sin\theta).\, \frac{\partial x}{\partial \theta}(\rho,\theta) + \frac{\partial f}{\partial y}(\rho\cos\theta,\rho\sin\theta).\, \frac{\partial y}{\partial \theta}(\rho,\theta) \end{cases}$$

c'est-à-dire:

$$\begin{cases} \frac{\partial F}{\partial \rho}(\rho,\theta) = \cos\theta.\frac{\partial f}{\partial x}(\rho\cos\theta,\rho\sin\theta) + \sin\theta.\frac{\partial f}{\partial y}(\rho\cos\theta,\rho\sin\theta) \\ \frac{\partial F}{\partial \theta}(\rho,\theta) = \rho(-\sin\theta).\frac{\partial f}{\partial x}(\rho\cos\theta,\rho\sin\theta) + \rho\cos\theta.\frac{\partial f}{\partial y}(\rho\cos\theta,\rho\sin\theta) \end{cases}$$

coordonnées polaires
$x = \rho\cos\theta$,
$y = \rho\sin\theta$

On synthétise ces relations en utilisant la notation matricielle:

$$\boxed{\begin{pmatrix} \frac{\partial F}{\partial \rho} \\ \\ \frac{\partial F}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \\ -\rho\sin\theta & \rho\cos\theta \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \\ \frac{\partial f}{\partial y} \end{pmatrix}}$$

1.4 Dérivées partielles d'ordre supérieur

• Dérivées partielles d'ordre 2.

Soit $f : (x, y) \mapsto f(x, y)$ une fonction de deux variables $U \to \mathbb{R}$ de classe C^1 sur un ouvert U de $\mathbb{R}^2$. Si les fonctions de deux variables $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ admettent elle-même des dérivées partielles par rapport à chacune des variables, on peut considérer les quatre fonctions dérivées partielles d'ordre 2 de f:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

Si de plus ces quatre fonctions sont continues sur U, on dit que f est de classe C^2 sur U.

Exemple. Reprenons l'exemple de 1.3: $f(x, y) = x^2 \sin y - y\, \mathrm{e}^{3x}$.

Donc $\frac{\partial f}{\partial x}(x, y) = 2x\sin y - 3y\,\mathrm{e}^{3x}$ et $\frac{\partial f}{\partial y}(x, y) = x^2\cos y - \mathrm{e}^{3x}$.

D'où $\frac{\partial^2 f}{\partial x^2}(x, y) = 2\sin y - 9y\,\mathrm{e}^{3x}$, $\frac{\partial^2 f}{\partial y^2}(x, y) = -x^2\sin y$ et $\frac{\partial^2 f}{\partial y \partial x}(x, y) = 2x\cos y - 3\,\mathrm{e}^{3x} = \frac{\partial^2 f}{\partial x \partial y}(x, y)$.

L'égalité des deux dernières dérivées secondes dans l'exemple précédent n'est pas fortuite; c'est un cas d'application du théorème suivant.

Théorème de Schwarz. Si f est de classe C^2 sur un ouvert U de $\mathbb{R}^2$, on a: $\boxed{\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{\partial^2 f}{\partial x \partial y}}$

• Dérivées partielles d'ordre n.

En considérant les dérivées partielles par rapport à chacune des variables des dérivées partielles d'ordre 2, on définit de même les dérivées partielles de f d'ordre 3. En itérant, on définit les dérivées partielles de f d'ordre n quelconque, et la notion de fonction de classe C^n.

1.5 Cas des fonctions de trois variables

Tout ce qu'on a fait ci-dessus s'adapte *mutatis mutandis* à des fonctions de trois variables $(x, y, z) \mapsto \mathbb{R}$ d'un ouvert U de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}$. Il y a bien sûr trois dérivées partielles d'ordre 1 à considérer:

$$\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}, \quad \text{et} \quad \frac{\partial f}{\partial z},$$

ainsi que neuf dérivées partielles d'ordre 2.

Exemple. Soit $f(x, y, z) = \dfrac{x^2 \sin y}{z}$, de classe C^2 sur $U = \mathbb{R}^3 - P$, où P est le plan d'équation $z = 0$.

On calcule d'abord: $\dfrac{\partial f}{\partial x} = \dfrac{2x\sin y}{z}$, $\dfrac{\partial f}{\partial y} = \dfrac{x^2\cos y}{z}$, $\dfrac{\partial f}{\partial z} = -\dfrac{x^2\sin y}{z^2}$,

Puis: $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{2\sin y}{z}$, $\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{2x\cos y}{z} = \dfrac{\partial^2 f}{\partial x \partial y}$, $\dfrac{\partial^2 f}{\partial z \partial x} = -\dfrac{2x\sin y}{z^2} = \dfrac{\partial^2 f}{\partial x \partial z}$,

$\dfrac{\partial^2 f}{\partial y^2} = -\dfrac{x^2\sin y}{z}$, $\dfrac{\partial^2 f}{\partial z \partial y} = -\dfrac{x^2\cos y}{z^2} = \dfrac{\partial^2 f}{\partial y \partial z}$, $\dfrac{\partial^2 f}{\partial z^2} = \dfrac{2x^2\sin y}{z^3}$.

La (seconde) formule de composition devient dans le cas de trois variables:

Pour une fonction de trois variables:

$$\begin{cases} F: & V \longrightarrow \mathbb{R} \\ & (x,y,z) \longmapsto f(u(x,y,z), v(x,y,z), w(x,y,z)) \end{cases}$$

de classe C^1 sur un ouvert V de $\mathbb{R}^3$, on a:

$$\boxed{\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} \\ \frac{\partial F}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial y} \\ \frac{\partial F}{\partial z} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial z} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial z} \end{aligned}}$$

• Exemple: passage en coordonnées cylindriques.

Si f est une fonction de trois variables (x, y, z),
et si l'on note $F(r, \theta, z) = f(r\cos\theta, r\sin\theta, z)$, alors:

$$\begin{cases} \dfrac{\partial F}{\partial r} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial r} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial r} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial r} \\[2ex] \dfrac{\partial F}{\partial \theta} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \theta} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \theta} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial \theta} \\[2ex] \dfrac{\partial F}{\partial z} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial z} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial z} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial z} \end{cases}$$

coordonnées cylindriques
$x = r\cos\theta$,
$y = r\sin\theta$,
$z = z$

On synthétise en utilisant la notation matricielle:

$$\boxed{\begin{pmatrix} \frac{\partial F}{\partial r} \\ \frac{\partial F}{\partial \theta} \\ \frac{\partial F}{\partial z} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -r\sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}}$$

• Exemple: passage en coordonnées sphériques.

Si f est une fonction de trois variables (x, y, z), et si
$F(\rho, \theta, \varphi) = f(\rho\cos\theta\cos\varphi, \rho\sin\theta\cos\varphi, \rho\sin\varphi)$, alors:

$$\begin{cases} \dfrac{\partial F}{\partial \rho} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \rho} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \rho} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial \rho} \\[2ex] \dfrac{\partial F}{\partial \theta} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \theta} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \theta} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial \theta} \\[2ex] \dfrac{\partial F}{\partial \varphi} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \varphi} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \varphi} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial \varphi} \end{cases}$$

coordonnées sphériques
$x = \rho\cos\theta\cos\varphi$,
$y = \rho\sin\theta\cos\varphi$,
$z = \rho\sin\varphi$

On synthétise en utilisant la notation matricielle:

$$\boxed{\begin{pmatrix} \frac{\partial F}{\partial r} \\ \frac{\partial F}{\partial \theta} \\ \frac{\partial F}{\partial z} \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\varphi & \sin\theta\cos\varphi & \sin\varphi \\ -\rho\sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & 0 \\ -\rho\cos\theta\sin\varphi & -\rho\sin\theta\sin\varphi & \rho\cos\varphi \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}}$$

2.1 Notion de fonction vectorielle.

• Fonction vectorielle d'une variable réelle. C'est une application définie sur une partie I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}^2$ ou $\mathbb{R}^3$ (ou plus généralement dans $\mathbb{R}^n$ pour $n \geq 2$). Pour la distinguer d'une fonction à valeurs dans $\mathbb{R}$, on note souvent $\vec{f}$ avec une flèche de vecteur une telle fonction. Il est clair que la donnée d'une telle fonction à valeurs dans $\mathbb{R}^2$ (respectivement dans $\mathbb{R}^3$) équivaut à la donnée de deux (respectivement trois) fonctions $I \to \mathbb{R}$ appelées les *fonctions coordonnées* de $\vec{f}$; en résumé:

$$\boxed{\begin{array}{rrcl} \vec{f}: & I & \longrightarrow & \mathbb{R}^2 \\ & x & \longmapsto & \big(f_1(x), f_2(x)\big) \end{array} \quad \text{avec } I \text{ partie de } \mathbb{R}, \text{ et } \begin{array}{l} f_1 : I \to \mathbb{R}, \\ f_2 : I \to \mathbb{R}. \end{array}}$$

ou

$$\boxed{\begin{array}{rrcl} \vec{f}: & I & \longrightarrow & \mathbb{R}^3 \\ & x & \longmapsto & \big(f_1(x), f_2(x), f_3(x)\big) \end{array} \quad \text{avec } I \text{ partie de } \mathbb{R}, \text{ et } \begin{array}{l} f_1 : I \to \mathbb{R}, \\ f_2 : I \to \mathbb{R}, \\ f_3 : I \to \mathbb{R}. \end{array}}$$

Une telle fonction $\vec{f}$ est continue (en un point, ou sur tout I) lorsque chacune des fonctions coordonnées l'est. Elle est dérivable en un point $a \in I$ lorsque chacune des fonctions coordonnées l'est, et on note alors:

$$\vec{f}'(a) = \big(f_1'(a), f_2'(a)\big).$$

• Fonction vectorielle de plusieurs variables réelles. On considère de même des fonctions définies sur une partie U de $\mathbb{R}^2$ ou $\mathbb{R}^3$, et à valeurs dans $\mathbb{R}^2$ ou $\mathbb{R}^3$. Chacune des fonctions coordonnées est alors une fonction scalaire de plusieurs variables du type de celles étudiées dans la première partie ci-dessus. Bien que rien n'empêche a priori de considérer des fonctions $\mathbb{R}^2 \to \mathbb{R}^3$ ou $\mathbb{R}^3 \to \mathbb{R}^2$, on s'intéressera surtout dans la pratique aux cas:

$$\boxed{\begin{array}{rrcl} \vec{f}: & U & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longmapsto & \big(f_1(x,y), f_2(x,y)\big) \end{array} \quad \text{avec } U \text{ partie de } \mathbb{R}^2, \text{ et } \begin{array}{l} f_1 : U \to \mathbb{R}, \\ f_2 : U \to \mathbb{R}. \end{array}}$$

ou

$$\boxed{\begin{array}{rrcl} \vec{f}: & U & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longmapsto & \big(f_1(x,y,z), f_2(x,y,z), f_3(x,y,z)\big) \end{array} \quad \text{avec } U \text{ partie de } \mathbb{R}^3, \text{ et } \begin{array}{l} f_1 : U \to \mathbb{R}, \\ f_2 : U \to \mathbb{R}, \\ f_3 : U \to \mathbb{R}. \end{array}}$$

Là-encore, $\vec{f}$ est continue (en un point, ou sur tout U) lorsque chacune des fonctions coordonnées l'est (au sens où on l'a vu ci-dessus en 1.2). De même, $\vec{f}$ admet une dérivée partielle par rapport à l'une des variables lorsque chacune des fonctions coordonnées admet une dérivée partielle par rapport à cette variable. On note (par exemple en deux variables):

$$\frac{\partial \vec{f}}{\partial x}(a_1,a_2) = \Big(\frac{\partial f_1}{\partial x}(a_1,a_2), \frac{\partial f_2}{\partial x}(a_1,a_2)\Big) \quad \text{et} \quad \frac{\partial \vec{f}}{\partial y}(a_1,a_2) = \Big(\frac{\partial f_1}{\partial y}(a_1,a_2), \frac{\partial f_2}{\partial y}(a_1,a_2)\Big).$$

Enfin, $\vec{f}$ est de classe C^1 sur U lorsque chacune des fonctions coordonnées l'est (au sens où on l'a vu ci-dessus en 1.3).

2.2 Matrice Jacobienne.

• En deux variables. Soit U un ouvert de $\mathbb{R}^2$, et $\vec{f} : U \to \mathbb{R}^2$ de fonctions coordonnées $f_1 : U \to \mathbb{R}$ et $f_2 : U \to \mathbb{R}$. On suppose que $\vec{f}$ est de classe C^1 sur U. On appelle alors *matrice jacobienne* de $\vec{f}$ en un point $a = (a_1, a_2)$ de U la matrice:

$$\boxed{\operatorname{Jac} \vec{f}(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x}(a_1,a_2) & \dfrac{\partial f_1}{\partial y}(a_1,a_2) \\ \dfrac{\partial f_2}{\partial x}(a_1,a_2) & \dfrac{\partial f_2}{\partial y}(a_1,a_2) \end{pmatrix}}$$

Exemple. Soit $\vec{f} : (x,y) \mapsto (\sin(xy), \mathrm{e}^{xy^2})$ de classe C^1 sur $U = \mathbb{R}^2$. Alors $\operatorname{Jac} \vec{f} = \begin{pmatrix} y\cos(xy) & x\cos(xy) \\ y^2\,\mathrm{e}^{xy^2} & 2yx\,\mathrm{e}^{xy^2} \end{pmatrix}$.

Passage en coordonnées polaires. Soit $\vec{f} : (\rho,\theta) \mapsto (\rho\cos\theta, \rho\sin\theta)$. Alors $\operatorname{Jac} \vec{f} = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$.
On retrouve (après transposition) la matrice rencontrée au dernier exemple de 1.3.

• En trois variables. Soit U un ouvert de $\mathbb{R}^3$, et $\vec{f}: U \to \mathbb{R}^3$ de fonctions coordonnées $f_1 : U \to \mathbb{R}$, $f_2 : U \to \mathbb{R}$ et $f_3 : U \to \mathbb{R}$. On suppose que $\vec{f}$ est de classe C^1 sur U. On appelle alors *matrice jacobienne* de $\vec{f}$ en un point $a = (a_1, a_2, a_3)$ de U la matrice:

$$\boxed{\operatorname{Jac}\vec{f}(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(a_1,a_2,a_3) & \frac{\partial f_1}{\partial y}(a_1,a_2,a_3), & \frac{\partial f_1}{\partial z}(a_1,a_2,a_3) \\ \frac{\partial f_2}{\partial x}(a_1,a_2,a_3) & \frac{\partial f_2}{\partial y}(a_1,a_2,a_3), & \frac{\partial f_2}{\partial z}(a_1,a_2,a_3) \\ \frac{\partial f_3}{\partial x}(a_1,a_2,a_3) & \frac{\partial f_3}{\partial y}(a_1,a_2,a_3), & \frac{\partial f_3}{\partial z}(a_1,a_2,a_3) \end{pmatrix}}$$

Exemple. Soit $\vec{f}: (x,y,z) \mapsto (xyz, x^2+y^2+z^2, xy+yz+zx)$. Alors $\operatorname{Jac}\vec{f} = \begin{pmatrix} yz & xz & xy \\ 2x & 2y & 2z \\ y+z & x+z & x+y \end{pmatrix}$.

Passage en coordonnées cylindriques et sphériques.

Soit $\vec{f}: (r,\theta,z) \mapsto (r\cos\theta, r\sin\theta, z)$ et $\vec{g}: (\rho,\theta,\varphi) = (\rho\cos\theta\cos\varphi, \rho\sin\theta\cos\varphi, \rho\sin\varphi)$,

Alors $\operatorname{Jac}\vec{f} = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $\operatorname{Jac}\vec{g} = \begin{pmatrix} \cos\theta\cos\varphi & -\rho\sin\theta\cos\varphi & -\rho\cos\theta\sin\varphi \\ \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\varphi & 0 & \rho\cos\varphi \end{pmatrix}$.

On retrouve (après transposition) les matrices rencontrées aux exemples de 1.5.

2.3 Formule de changement de variables.

Soit U un ouvert de $\mathbb{R}^2$. Soit $\vec{f}: U \to \mathbb{R}^2$ une fonction de classe C^1 sur U. Notons ici u et v ses fonctions coordonnées, qui sont donc des fonctions $U \to \mathbb{R}$ de classe C^1 sur U. Soit V un ouvert de $\mathbb{R}^2$ tel que $\vec{f}(U) \subset V$ et $\vec{g}: V \to \mathbb{R}^2$ une fonction de classe C^1 sur V. Notons g_1 et g_2 ses fonctions coordonnées. On peut considérer la composée $\vec{F} = \vec{g} \circ \vec{f}$.

$$\vec{F}: (x,y) \xrightarrow{\ \vec{f}\ } \big(u(x,y), v(x,y)\big) \xrightarrow{\ \vec{g}\ } \Big(\underbrace{g_1\big(u(x,y),v(x,y)\big)}_{F_1(x,y)}, \underbrace{g_2\big(u(x,y),v(x,y)\big)}_{F_2(x,y)}\Big).$$

En appliquant la seconde formule vue en 1.3 à chacune des fonctions coordonnées F_1 et F_2, on obtient:

$$\frac{\partial F_1}{\partial x} = \frac{\partial g_1}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g_1}{\partial v}\frac{\partial v}{\partial x}, \qquad \frac{\partial F_1}{\partial y} = \frac{\partial g_1}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial g_1}{\partial v}\frac{\partial v}{\partial y},$$

$$\frac{\partial F_2}{\partial x} = \frac{\partial g_2}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g_2}{\partial v}\frac{\partial v}{\partial x}, \qquad \frac{\partial F_2}{\partial y} = \frac{\partial g_2}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial g_2}{\partial v}\frac{\partial v}{\partial y},$$

que l'on écrit globalement sous forme matricielle:

$$\begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{pmatrix} \times \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}.$$

Cela signifie plus explicitement que l'on a en tout point $A = (a,b)$ de U:

$$\begin{pmatrix} \frac{\partial F_1}{\partial x}(a,b) & \frac{\partial F_1}{\partial y}(a,b) \\ \frac{\partial F_2}{\partial x}(a,b) & \frac{\partial F_2}{\partial y}(a,b) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial u}\big(u(a,b),v(a,b)\big) & \frac{\partial g_1}{\partial v}\big(u(a,b),v(a,b)\big) \\ \frac{\partial g_2}{\partial u}\big(u(a,b),v(a,b)\big) & \frac{\partial g_2}{\partial v}\big(u(a,b),v(a,b)\big) \end{pmatrix} \times \begin{pmatrix} \frac{\partial u}{\partial x}(a,b) & \frac{\partial u}{\partial y}(a,b) \\ \frac{\partial v}{\partial x}(a,b) & \frac{\partial v}{\partial y}(a,b) \end{pmatrix}.$$

On aboutit donc à la formule “compacte”:

$$\boxed{\operatorname{Jac}(\vec{g}\circ\vec{f})(A) = \operatorname{Jac}\vec{g}\,(\vec{f}(A)) \times \operatorname{Jac}\vec{f}(A).}$$

Cette formule reste vraie pour des fonctions de $\mathbb{R}^3$ dans $\mathbb{R}^3$, et plus généralement de $\mathbb{R}^n$ dans $\mathbb{R}^n$.

2.4 Opérateurs classiques de l'analyse vectorielle.

• Définitions. On se place dans $\mathbb{R}^3$. En physique, on emploie souvent le terme de *champ scalaire* pour désigner une fonction $f : U \to \mathbb{R}$, et de *champ vectoriel* pour désigner une fonction $f : U \to \mathbb{R}^3$, où U est un ouvert de $\mathbb{R}^3$.

1. Si $f : U \to \mathbb{R}$ est un champ scalaire de classe C^1, le *gradient* de f est le champ vectoriel $\overrightarrow{\mathrm{grad}}\, f : U \to \mathbb{R}^3$ défini par:
$$\overrightarrow{\mathrm{grad}}\, f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}).$$
2. Si $f : U \to \mathbb{R}$ est un champ scalaire de classe C^2, le *laplacien* de f est le champ scalaire $\Delta f : U \to \mathbb{R}$ défini par:
$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$
3. Si $\vec{F} : U \to \mathbb{R}^3$ est un champ vectoriel de classe C^1, et si l'on note P, Q, R les champs scalaires $U \to \mathbb{R}$ de classe C^1 qui sont les fonctions coordonnées de $\vec{F}$, *la divergence* de $\vec{F}$ est le champ scalaire $\mathrm{div}\, \vec{F} : U \to \mathbb{R}$ défini par:
$$\mathrm{div}\, \vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$
4. Avec les mêmes hypothèses et notations que ci-dessus, le *rotationnel* de $\vec{F}$ est le champ vectoriel $\overrightarrow{\mathrm{rot}}\, \vec{F} : U \to \mathbb{R}^3$ défini par:
$$\overrightarrow{\mathrm{rot}}\, \vec{F} = (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \;,\; \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \;,\; \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}).$$

On pourra vérifier à titre d'exercice qu'il résulte immédiatement des formules de Schwarz pour des champs de classe C^2 que:

$$\boxed{\overrightarrow{\mathrm{rot}}(\overrightarrow{\mathrm{grad}}\, f) = \vec{0}} \qquad \text{et} \qquad \boxed{\mathrm{div}(\overrightarrow{\mathrm{rot}}\, \vec{F}) = 0}$$

Plusieurs problèmes classiques d'équations aux dérivées partielles relatifs à ces opérateurs seront vus dans le cours de licence en liaison avec l'intégration des formes différentielles. Citons les deux suivants, pour lesquels on a les théorèmes (admis) ci-dessous.

• Champ scalaire de gradient nul.

Sous certaines bonnes hypothèses topologiques sur l'ouvert U de $\mathbb{R}^3$ (il suffit qu'il soit connexe), on a pour tout champ scalaire $f : U \to \mathbb{R}$ de classe C^1:

$$\boxed{\overrightarrow{\mathrm{grad}}\, f = \vec{0}} \iff \boxed{\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0} \iff \boxed{f \text{ est constant sur } U}$$

• Champ vectoriel de rotationnel nul.

D'une façon générale, pour tout champ scalaire $f : U \to \mathbb{R}$ de classe C^1, on a toujours $\overrightarrow{\mathrm{rot}}\,\overrightarrow{\mathrm{grad}}\, f = \vec{0}$ (voir formules ci-dessus).

Réciproquement, si $\vec{F}$ est un champ vectoriel donné tel que $\overrightarrow{\mathrm{rot}}\, \vec{F} = \vec{0}$, existe-t-il nécessairement un champ scalaire f tel que $\vec{F} = \overrightarrow{\mathrm{grad}}\, f$? Si c'est le cas, on dit que $\vec{F}$ *dérive d'un potentiel.*

Sous certaines bonnes hypothèses topologiques sur l'ouvert U de $\mathbb{R}^3$ (il suffit qu'il soit étoilé), on a une réponse positive à cette question:

$$\boxed{\overrightarrow{\mathrm{rot}}\, \vec{F} = \vec{0}} \iff \boxed{\frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z} \;,\; \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x} \;,\; \frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}} \iff \boxed{\vec{F} \text{ dérive d'un potentiel}}$$

Vecteurs

1. Sous-espaces vectoriels.

1.1 Combinaisons linéaires.

• Vecteurs de $\mathbb{R}^n$. On fixe un entier $n \geq 1$ (dans la pratique, ce sera souvent 2 ou 3). On considère l'ensemble $\mathbb{R}^n$ des n-uplets de réels. Un tel n-uplet est appelé un *vecteur* de $\mathbb{R}^n$ et se note avec des parenthèses:

$$u = (x_1, x_2, \dots, x_n) \ \text{ avec } x_i \in \mathbb{R} \text{ pour tout } 1 \leq i \leq n.$$

Notations Lorsque $n = 2$, un vecteur $u = (x_1, x_2)$ est un couple; on préfère alors souvent ne pas utiliser d'indice et noter $u = (x, y)$. Lorsque $n = 3$, un vecteur $u = (x_1, x_2, x_3)$ est un triplet, et l'on préfère noter $u = (x, y, z)$. Lorsque $n = 1$, on a $\mathbb{R}^1 = \mathbb{R}$, un vecteur est donc dans ce cas simplement un réel $u = x_1$.

Attention ! La notion de vecteur fait intervenir l'ordre dans lequel sont pris les nombres x_i, c'est-à-dire que:

$$[(x_1, x_2, \dots, x_n) = (y_1, y_2, \dots, y_n) \text{ dans } \mathbb{R}^n] \Leftrightarrow [x_i = y_i \text{ dans } \mathbb{R} \text{ pour tout } 1 \leq i \leq n].$$

Par exemple $(x_1, x_2) \neq (x_2, x_1)$ dans $\mathbb{R}^2$ dès lors que $x_1 \neq x_2$ dans $\mathbb{R}$.

• Addition des vecteurs. Elle est définie de façon naturelle à partir de l'addition des nombres réels, en additionnant terme à terme, c'est-à-dire:

$$\text{si } u = (x_1, x_2, \dots, x_n) \text{ et } v = (y_1, y_2, \dots, y_n), \ \text{ alors } \ u + v = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n).$$

Par exemple dans $\mathbb{R}^3$: $(1, -2, 0) + (-1, 4, 3) = (0, 2, 3)$.

On appelle vecteur nul de $\mathbb{R}^n$ le vecteur $\vec{0} = (0, 0, \dots, 0)$. Pour tout vecteur $u = (x_1, x_2, \dots, x_n)$, on appelle opposé de u le vecteur $-u = (-x_1, -x_2, \dots, -x_n)$

• Produit d'un vecteur par un nombre réel. Il est défini de la façon suivante:

$$\text{si } u = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \text{ et } \alpha \in \mathbb{R}, \text{ on pose } \alpha.u = (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \in \mathbb{R}^n.$$

Par exemple dans $\mathbb{R}^3$: $2.(-1, 4, 3) = (-2, 8, 6)$ et $-\frac{1}{3}(6, 0, -5) = (-2, 0, \frac{5}{3})$.

• Propriétés de ces opérations. Elles se déduisent immédiatement des propriétés de l'addition et de la multiplication dans $\mathbb{R}$. En particulier, pour tous $u, v \in \mathbb{R}^n$ et tous $\alpha, \beta \in \mathbb{R}$:

$$\begin{array}{ll} u + (v + w) = (u + v) + w, & (\alpha + \beta).u = \alpha.u + \beta.u, \\ u + \vec{0} = \vec{0} + u = u, & \alpha.(u + v) = \alpha.u + \alpha.v, \\ u + (-u) = (-u) + u = \vec{0}, & \alpha.(\beta.u) = (\alpha\beta).u, \\ u + v = v + u, & 1.u = u. \end{array}$$

Ces huit propriétés fondamentales font de l'ensemble $\mathbb{R}^n$ ce qu'on appelle un *espace vectoriel* sur $\mathbb{R}$.
Elles entraînent en outre comme conséquences:

$$(-\alpha).u = \alpha.(-u) = -\alpha.u, \qquad (-1).u = -u, \qquad \alpha.(u - v) = \alpha.u - \alpha.v, \qquad (\alpha - \beta).u = \alpha.u - \beta.u,$$

ainsi que la propriété suivante, qui joue un rôle important dans beaucoup de raisonnements:

$$(\alpha.u = \vec{0}) \text{ si et seulement si } (\alpha = 0 \text{ dans } \mathbb{R}, \text{ ou } u = \vec{0} \text{ dans } \mathbb{R}^n).$$

• Combinaisons linéaires. Il s'agit simplement de combiner naturellement les deux lois + et . de $\mathbb{R}^n$.

Si u et v sont deux vecteurs de $\mathbb{R}^n$, on appelle combinaison linéaire de u et v tout vecteur de la forme $\alpha.u + \beta.v$, où α et β sont des scalaires.

On peut considérer de même des combinaisons linéaires de trois vecteurs: $\alpha.u + \beta.v + \gamma.w$, ou plus...
Par exemple dans $\mathbb{R}^3$, à partir des vecteurs $u = (1, -2, 3)$ et $v = (0, 5, -1)$, on peut fabriquer la combinaison linéaire $w = 2.u + 3.v = (2, -4, 6) + (0, 15, -3) = (2, 11, 3)$.
Pour calculer plus commodément, on dispose parfois les vecteurs en colonnes. Par exemple dans $\mathbb{R}^4$:

$$\begin{pmatrix} 2 \\ -5 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ -2 \\ 1 \end{pmatrix}, \qquad 3. \begin{pmatrix} 1 \\ 0 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -6 \\ 9 \end{pmatrix}, \qquad - \begin{pmatrix} 1 \\ 8 \\ -1 \\ 0 \end{pmatrix} + \tfrac{1}{2}. \begin{pmatrix} 4 \\ 0 \\ -2 \\ 3 \end{pmatrix} + 3. \begin{pmatrix} 1 \\ 5 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 3 \\ -\frac{9}{2} \end{pmatrix}.$$

Dans la suite, on utilisera l'abréviation c.l. pour combinaison linéaire.

1.2 Sous-espaces vectoriels de $\mathbb{R}^n$

• Exemples préliminaires. Plaçons-nous dans $\mathbb{R}^2$.

1. Considérons d'abord le sous-ensemble F de $\mathbb{R}^2$ constitué des vecteurs (x, y) tels que $x = y$. Si $u = (x, x)$ et $v = (x', x')$ sont deux vecteurs quelconques de F, leur somme $u + v$ reste dans F, puisque $u + v = (x + x', x + x')$. On traduit cette propriété en disant que F est stable pour l'addition. De même, quels que soient un vecteur $u = (x, x)$ de F et un scalaire $\alpha \in \mathbb{R}$, le vecteur $\alpha.u = (\alpha x, \alpha x)$ reste dans F. On traduit cette propriété en disant que F est stable aussi pour le produit externe.
2. Considérons maintenant le sous-ensemble G de $\mathbb{R}^2$ constituée des vecteurs (x, y) tels que $x = 1$. Si $u = (1, y)$ et $v = (1, y')$ sont deux vecteurs quelconques de G, leur somme est le vecteur $u+v = (2, y+y')$, qui n'est plus un vecteur de G. Donc G n'est pas stable pour l'addition de $\mathbb{R}^n$. De même, le produit externe d'un vecteur $u = (1, y)$ par un scalaire $\alpha \in \mathbb{R}$ est $\alpha.u = (\alpha, \alpha y)$ qui n'est plus un élément de G dès lors que le réel α est distinct de 1. Donc G n'est pas non plus stable pour le produit externe.
3. A noter qu'un sous-ensemble peut être stable pour l'addition mais pas pour le produit externe; c'est la cas par exemple du sous-ensemble des vecteurs (x, y) tels que $x \geq 0$.

• Notion de sous-espace vectoriel

Définition. Un sous-ensemble F non-vide de $\mathbb{R}^n$ est appelé un *sous-espace vectoriel* de $\mathbb{R}^n$ lorsque F est *stable* à la fois pour l'addition et pour le produit externe, c'est-à-dire:

$$\boxed{F \text{ est un sous-espace vectoriel de } \mathbb{R}^n} \iff \boxed{\begin{array}{l} \text{pour tous } u, v \in F, \text{ on a: } u + v \in F \\ \text{pour tout } u \in F \text{ et tout } \alpha \in \mathbb{R}, \text{ on a: } \quad \alpha.u \in F \end{array}}$$

On utilisera dans la suite l'abréviation ss-e.v. pour sous-espace vectoriel.

Il est clair qu'être stable à la fois pour l'addition et pour le produit externe signifie être stable par combinaison linéaire:

$$\boxed{F \text{ est un ss-e.v. de } \mathbb{R}^n} \iff \boxed{\begin{array}{c} \text{pour tous } u, v \in F, \\ \text{et tous } \alpha, \beta \in \mathbb{R}, \text{ on a:} \\ \alpha.u + \beta.v \in F \end{array}} \iff \boxed{\begin{array}{l} \text{toute c.l. d'un nombre} \\ \text{quelconque de vecteurs} \\ \text{de } F \text{ reste dans } F \end{array}}$$

• Remarques.

1. Un ss-e.v. F de $\mathbb{R}^n$ contient nécessairement le vecteur nul (en effet, F contient au moins un vecteur u, et donc aussi la c.l. $u - u = \vec{0}$). A contrario, dès lors qu'un sous-ensemble ne contient pas le vecteur nul, on est sûr que ce n'est pas un ss-e.v. !
2. Il est clair que $\{\vec{0}\}$ et $\mathbb{R}^n$ lui-même sont toujours des ss-e.v. de $\mathbb{R}^n$.
3. Comment s'y prendre pour vérifier qu'un sous-ensemble F de $\mathbb{R}^n$ est un ss-e.v. :

   – d'abord, s'assurer que F est non-vide, et le plus simple pour cela est de vérifier que le vecteur nul $\vec{0}$ appartient bien à F.

   – puis montrer que: ($u + v \in F$ pour tous $u, v \in F$) et ($\alpha.u \in F$ pour tous $u \in F$ et $\alpha \in \mathbb{R}$); si les calculs sont simples, on peut vérifier en une seule fois que: ($\alpha.u + \beta.v \in F$ pour tous u, v, α, β).

• Exemples.

Dans $\mathbb{R}^3$, les sous-ensembles $P = \{(x, 0, z)\,;\, x \in \mathbb{R}, z \in \mathbb{R}\}$ et $D = \{(x, 0, x)\,;\, x \in \mathbb{R}\}$ sont des ss-e.v. de $\mathbb{R}^3$. En revanche, $A = \{(x, 1, x)\,;\, x \in \mathbb{R}\}$ n'est pas un ss-e.v. car la somme de deux vecteurs $(x, 1, x)$ et $(x', 1, x')$ est $(x + x', 2, x + x')$ qui n'appartient pas à A. [On pourrait aussi remarquer que $\vec{0} = (0, 0, 0) \notin A$, ce qui suffit à prouver que A n'est pas un ss-e.v. de $\mathbb{R}^3$].

1.3 Intersection de sous-espaces vectoriels

Rappelons d'abord que, si F et G sont deux sous-ensembles quelconques de $\mathbb{R}^n$, on appelle *intersection* de F et G, notée $F \cap G$, le sous-ensemble de $\mathbb{R}^n$ formé des vecteurs qui appartiennent à la fois à F et à G. Le résultat suivant est alors à peu près évident:

Proposition. *Si F et G sont deux ss-e.v. de $\mathbb{R}^n$, alors $F \cap G$ est un ss-e.v. de $\mathbb{R}^n$.*

Exemple. Dans $E = \mathbb{R}^3$, $P = \{(x, 0, z)\,;\, x \in \mathbb{R}, z \in \mathbb{R}\}$, $Q = \{(x, y, 0)\,;\, x \in \mathbb{R}, y \in \mathbb{R}\}$ ont pour intersection $P \cap Q = \{(x, 0, 0)\,;\, x \in \mathbb{R}\}$.

1.4 Somme directe et sous-espaces vectoriels supplémentaires

• Cas de deux sous-espaces

Définition. Soient F et G deux ss-e.v. de $\mathbb{R}^n$. On dit qu'ils sont *supplémentaires* dans $\mathbb{R}^n$ lorsque tout vecteur de $\mathbb{R}^n$ se décompose *de façon unique* en la somme d'un vecteur de F et d'un vecteur de G.
On dit aussi dans ce cas que $\mathbb{R}^n$ est *somme directe* de F et G et on note: $\mathbb{R}^n = F \oplus G$. Ainsi

$$\boxed{\mathbb{R}^n = F \oplus G} \iff \boxed{\text{Pour tout } u \in \mathbb{R}^n, \text{ il existe } v \in F \text{ et } w \in G \text{ uniques tels que } u = v + w.}$$

Remarquons que, si $\mathbb{R}^n = F \oplus G$, alors on a forcément $F \cap G = \{\vec{0}\}$.

Exemple. Dans $E = \mathbb{R}^3$, $P = \{(x, 0, z)\,;\, x \in \mathbb{R}, z \in \mathbb{R}\}$ et $S = \{(0, y, 0)\,;\, y \in \mathbb{R}\}$ sont des ss-e.v. supplémentaires. On note: $\mathbb{R}^3 = P \oplus S$.

• Cas d'un nombre quelconque de sous-espaces

Définition. Soient $F_1, F_2, \ldots, F_p$ des ss-e.v. de $\mathbb{R}^n$. On dit qu'ils sont *supplémentaires* dans $\mathbb{R}^n$ lorsque tout vecteur de $\mathbb{R}^n$ se décompose *de façon unique* en la somme d'un vecteur de F_1, plus un vecteur de F_2, ..., plus un vecteur de F_p.

$$\boxed{\mathbb{R}^n = F_1 \oplus F_2 \oplus \cdots \oplus F_p} \iff \boxed{\begin{array}{c}\text{Pour tout } u \in \mathbb{R}^n, \text{ il existe } v_1 \in F_1, v_2 \in F_2, \ldots, v_p \in F_p \\ \text{uniques tels que: } u = v_1 + v_2 + \cdots + v_p.\end{array}}$$

2. Famille génératrice, famille libre, base et dimension

2.1 Familles génératrices d'un sous-espace vectoriel.

• Exemple préliminaire.

Dans $\mathbb{R}^3$, considérons les deux vecteurs $u = (0, 1, 2)$ et $v = (-1, 3, 1)$. On peut à partir de ces deux vecteurs en fabriquer “beaucoup” d'autres en utilisant l'addition et le produit externe, c'est-à-dire en prenant toutes les combinaisons linéaires $\alpha.u + \beta.v$ où α et β décrivent $\mathbb{R}$.

Peut-on obtenir ainsi tous les vecteurs de $\mathbb{R}^3$? La réponse est non. En effet, le vecteur $w = (1, 1, 1)$ par exemple ne peut jamais s'écrire $\alpha.u + \beta.v$. Il suffit pour s'en convaincre de remarquer que l'on aurait sinon $(1, 1, 1) = (-\beta, \alpha + 3\beta, 2\alpha + \beta)$, d'où $\beta = -1$, $\alpha - 3 = 1$ et $2\alpha - 1 = 1$, qui n'a pas de solution dans $\mathbb{R}$. On traduit cette propriété en disant que la famille de vecteurs $\mathcal{X} = (u, v)$ n'engendre pas $\mathbb{R}^3$ (ou n'est pas une famille génératrice de $\mathbb{R}^3$).

Que peut-on dire alors de l'ensemble des vecteurs de $\mathbb{R}^3$ qui sont de la forme $\alpha.u + \beta.v$? Cet ensemble F est égal à $\{(-\beta, \alpha + 3\beta, 2\alpha + \beta)\,;\, \alpha, \beta \in \mathbb{R}\}$, et il est facile de montrer que c'est un ss-e.v. de $\mathbb{R}^3$. On dit que la famille de vecteurs $\mathcal{X} = (u, v)$ engendre F (ou est une famille génératrice de F).

• Définition.

Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille finie de p vecteurs de $\mathbb{R}^n$. Soit F un ss-e.v. de $\mathbb{R}^n$.
On dit que $\mathcal{X}$ est une *famille génératrice de* F, (ou que $\mathcal{X}$ *engendre* F), lorsque tout vecteur de F s'écrit comme une c.l. de $u_1, u_2, \ldots, u_p$.

> En particulier, pour $F = \mathbb{R}^n$, une famille génératrice de l'espace $\mathbb{R}^n$ tout entier est une famille $\mathcal{X}$ de vecteurs de $\mathbb{R}^n$ telle que n'importe quel vecteur de $\mathbb{R}^n$ puisse s'exprimer comme une c.l. des vecteurs de la famille $\mathcal{X}$.

• Proposition et définition.

Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille finie de p vecteurs de $\mathbb{R}^n$. L'ensemble de toutes les c.l. que l'on peut former avec les vecteurs de $\mathcal{X}$ est un ss-e.v. de $\mathbb{R}^n$. On l'appelle le ss-e.v. *engendré* par $\mathcal{X}$. On le note $\operatorname{Vect} \mathcal{X}$.

> Remarque. Il est clair par définition que $\mathcal{X}$ est une famille génératrice de $\operatorname{Vect} \mathcal{X}$.

• En résumé, pour toute famille $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ de vecteurs de $\mathbb{R}^n$ et tout ss-e.v. F de $\mathbb{R}^n$:

$$\boxed{F = \operatorname{Vect} \mathcal{X}} \iff \boxed{\begin{array}{c}\mathcal{X} = (u_1, u_2, \ldots, u_p) \text{ est une} \\ \text{famille génératrice de } F\end{array}} \iff \boxed{\begin{array}{c}\text{pour tout } v \in F, \text{ il existe } \alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R} \\ \text{tels que: } v = \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p\end{array}}$$

2.2 Familles libres, familles liées.

• Exemples préliminaires

Dans $\mathbb{R}^3$, considérons les vecteurs $u = (0,1,2)$, $v = (-1,3,1)$ et $s = (-2,7,4)$. Il est clair que $u + 2.v + (-1).s = (0,0,0)$. On a donc une combinaison linéaire de ces 3 vecteurs qui est nulle, sans que les 3 coefficients (à savoir $1, 2, -1$) soient simultanément nuls. On traduit cette propriété en disant que la famille $\mathcal{X} = (u, v, s)$ est liée. Cela permet d'exprimer l'un des vecteurs comme une combinaison linéaire des deux autres, par exemple $s = u + 2.v$. Géométriquement, cela signifie que le vecteur s est dans le plan déterminé par les vecteurs u et v.

Prenons maintenant $u = (0,1,2)$, $v = (-1,3,1)$ et $t = (0,1,0)$. Supposons qu'une combinaison linéaire $\alpha.u + \beta.v + \gamma.t$ soit nulle. On aurait donc: $(-\beta, \alpha + 3\beta + \gamma, 2\alpha + \beta) = (0,0,0)$. On en tire d'abord $\beta = 0$, puis $\gamma = 0$ et $\alpha = 0$. La seule combinaison linéaire de ces 3 vecteurs qui soit nulle est celle dont les coefficients valent tous les 3 zéro. On traduit cette propriété en disant que la famille $\mathcal{X}' = (u, v, t)$ est libre. Géométriquement, cela signifie que les vecteurs u, v, t ne sont pas situés dans un même plan de l'espace.

> Commentaire. Soit $\mathcal{X} = (u_1, u_2, \dots, u_p)$ une famille finie de p vecteurs de $\mathbb{R}^n$. La question est de savoir s'il peut exister une c.l. de ces vecteurs qui soit nulle. Bien sr, si l'on prend tous les α_i égaux à 0, on a trivialement $0.u_1 + 0.u_2 + \cdots + 0.u_p = \vec{0}$. On ne retient donc pour la suite que les cas non-triviaux, c'est-à-dire où les α_i sont *non tous nuls*. Rappelons que dire que les scalaires $\alpha_1, \alpha_2, \dots, \alpha_p$ sont *non tous nuls* signifie que *l'un au moins* des α_i est non-nul (ce qui n'exclut pas que certains des α_i puissent être nuls; mais pas tous...). Ceci équivaut aussi à dire que $(\alpha_1, \alpha_2, \dots, \alpha_p) \neq (0, 0, \dots, 0)$.

• Définition 1. On dit que la famille $\mathcal{X} = (u_1, u_2, \dots, u_p)$ est *liée* lorsqu'il existe une c.l. des vecteurs de $\mathcal{X}$, à coefficients non tous nuls, qui est nulle dans $\mathbb{R}^n$.

$$\boxed{\begin{array}{c}\mathcal{X} = (u_1, u_2, \dots, u_p) \text{ est}\\ \text{une famille liée}\end{array}} \iff \boxed{\begin{array}{c}\text{il existe } \alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{R} \text{ NON TOUS NULS}\\ \text{tels que: } \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p = \vec{0}\end{array}}$$

Une telle relation s'appelle une *relation de dépendance linéaire* entre les vecteurs de $\mathcal{X}$.

• Définition 2. On dit que la famille $\mathcal{X} = (u_1, u_2, \dots, u_p)$ est *libre* lorsqu'elle n'est pas liée.
Ceci signifie donc qu'il n'existe pas de relation de dépendance linéaire à coefficients non tous nuls entre les vecteurs de $\mathcal{X}$, c'est-à-dire que la seule c.l. nulle des vecteurs de $\mathcal{X}$ est celle où tous les coefficients sont nuls, ce qui se traduit encore par:

$$\boxed{\begin{array}{c}\mathcal{X} = (u_1, u_2, \dots, u_p) \text{ est}\\ \text{une famille libre}\end{array}} \iff \boxed{\begin{array}{c}\text{si } \alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{R} \text{ vérifient: } \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p = \vec{0},\\ \text{alors nécessairement } \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0\end{array}}$$

On dit aussi dans ce cas que les vecteurs de $\mathcal{X}$ sont *linéairement indépendants.*

On regroupe dans la proposition suivante quelques propriétés utiles des familles libres ou liées.

• Proposition. *Les familles finies de vecteurs de $\mathbb{R}^n$ vérifient les propriétés suivantes:*

(i) *Toute sous-famille d'une famille libre est libre.*

(ii) *Toute famille qui admet une sous-famille liée est liée.*

(iii) *Toute famille dont l'un des vecteurs est $\vec{0}$ est liée.*

(iv) *Une famille formée d'un seul vecteur est liée si et seulement si ce vecteur est nul (donc est libre si et seulement si ce vecteur est non-nul).*

(v) *Une famille est liée si et seulement s'il existe un de ses vecteurs qui est combinaison linéaire des autres.*

(vi) *Une famille de deux vecteurs est liée si et seulement si l'un des deux est le produit de l'autre par un réel.*

2.3 Bases d'un sous-espace vectoriel.

• Définition. Une famille de vecteurs $\mathcal{X} = (u_1, u_2, \dots, u_p)$ est une base d'un ss-e.v. F de $\mathbb{R}^n$ lorsqu'elle est à la fois une famille libre et une famille génératrice de F.

On a la caractérisation fondamentale suivante:

• Théorème.

$$\boxed{\begin{array}{c}\mathcal{X} = (u_1, u_2, \dots, u_p) \text{ est une}\\ \text{base de } F\end{array}} \iff \boxed{\begin{array}{c}\text{pour tout } v \in F, \text{ il existe } \alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{R} \text{ UNIQUES}\\ \text{tels que: } v = \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p\end{array}}$$

Les réels uniques $\alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{R}$ sont appelés les *composantes* du vecteur v dans la base $\mathcal{X}$.

• EXEMPLES.

EXEMPLE 1. Dans $\mathbb{R}^3$, considérons $F = \{(x, y-x, 2x+3y)\,;\, x \in \mathbb{R},\ y \in \mathbb{R}\}$. Tout vecteur de F s'écrit:

$$(x, y-x, 2x+3y) = x.(1,-1,2) + y.(0,1,3),$$

ce qui montre que la famille $\mathcal{X} = (u, v)$ formée par les deux vecteurs $u = (1,-1,2)$ et $v = (0,1,3)$ est une famille génératrice de F. Il est clair par ailleurs qu'elle est libre (par exemple d'après le point (vi) de la proposition du 2.2). Donc $\mathcal{X}$ est une base de F.

EXEMPLE 2. Toujours dans $\mathbb{R}^3$, prenons pour F l'espace $\mathbb{R}^3$ lui-même, et considérons les trois vecteurs $u = (2,0,0)$, $v = (0,1,0)$ et $w = (1,1,1)$. Tout vecteur $(x,y,z) \in \mathbb{R}^3$ s'écrit:

$$(x,y,z) = \tfrac{1}{2}(x-z).u + (y-z).v + z.w,$$

ce qui montre que la famille $\mathcal{X} = (u, v, w)$ est une famille génératrice de $\mathbb{R}^3$. Pour montrer qu'elle est aussi libre, donnons-nous trois réels α, β, γ vérifiant $\alpha.u + \beta.v + \gamma.w = \vec{0}$. On a donc: $(2\alpha+\gamma, \beta+\gamma, \gamma) = (0,0,0)$, d'où l'on tire successivement que $\gamma = 0$, donc $\beta = 0$, donc $\alpha = 0$. Ceci prouve que $\mathcal{X}$ est libre. On conclut finalement que $\mathcal{X}$ est une base de $\mathbb{R}^3$.

• BASE CANONIQUE.

La base canonique de $\mathbb{R}^2$ est $\mathcal{B} = (e_1, e_2)$ avec $e_1 = (1,0)$ et $e_2 = (0,1)$.

La base canonique de $\mathbb{R}^3$ est $\mathcal{B} = (e_1, e_2, e_3)$ avec $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ et $e_3 = (0,0,1)$.

Et plus généralement, pour tout $n \geq 1$, la base canonique de $\mathbb{R}^n$ est $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ où:

$$e_1 = (1,0,0,0,\ldots,0), \quad e_2 = (0,1,0,0,\ldots,0), \quad e_3 = (0,0,1,\ldots,0), \quad \ldots, \quad e_n = (0,0,\ldots,0,1).$$

REMARQUES.

1. Le fait que $(e_1, e_2, \ldots, e_n)$ est effectivement une famille libre et génératrice de $\mathbb{R}^n$ est évident par définition.
2. La base canonique est caractérisée par le fait que, pour tout vecteur $v = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, les composantes de v dans la base canonique sont précisément $(x_1, x_2, \ldots, x_n)$.
3. La base canonique est celle que l'on utilise classiquement en géométrie ou en physique pour repérer un vecteur par ses composantes.

2.4 Dimension d'un sous-espace vectoriel.

Tout repose sur le résultat suivant:

• THÉORÈME ET DÉFINITION. *Soit F un ss-e.v. de $\mathbb{R}^n$. On a les propriétés suivantes:*

(1) *Toutes les bases de F sont formées du même nombre de vecteurs, que l'on appelle la dimension de F, et que l'on note* $\dim F$.

(2) *On a toujours:* $0 \leq \dim F \leq n = \dim \mathbb{R}^n$, *avec de plus* $\begin{cases} \dim F = n \text{ si et seulement si } F = \mathbb{R}^n, \\ \dim F = 0 \text{ si et seulement si } F = \{\vec{0}\}. \end{cases}$

(3) *Plus généralement, si H est un ss-e.v. de $\mathbb{R}^n$ tel que $F \subseteq H$, on a $\dim F \leq \dim H$, avec de plus $\dim F = \dim H$ si et seulement si $F = H$.*

EN PARTICULIER:

- Dans $\mathbb{R}^2$, les ss-e.v. sont: $\mathbb{R}^2$ lui-même (qui est le seul ss-e.v. de dimension 2), tous les ss-e.v. de dimension 1 (qu'on appelle des droites vectorielles), et le ss-e.v. nul $\{\vec{0}\}$ qui est le seul ss-e.v. de dimension 0.
- Dans $\mathbb{R}^3$, les ss-e.v. sont: $\mathbb{R}^3$ lui-même (qui est le seul ss-e.v. de dimension 3), tous les ss-e.v. de dimension 2 (qu'on appelle des plans vectoriels), tous les ss-e.v. de dimension 1 (les droites vectorielles), et le ss-e.v. nul $\{\vec{0}\}$ qui est le seul ss-e.v. de dimension 0.

Le résultat suivant donne une caractérisation des bases en terme de *familles libres maximales*, ou encore de *familles génératrices minimales*.

• PROPOSITION 1. *Soit F un ss-e.v. de $\mathbb{R}^n$. Notons m sa dimension.*

(i) *Une famille libre de vecteurs de F est formée de au plus m vecteurs (et par voie de conséquence, une famille de plus de m vecteurs est nécessairement liée).*

(ii) *Une famille génératrice de F est formée de au moins m vecteurs.*

COMMENTAIRE.

- Dans $\mathbb{R}^3$, une base est formée de 3 vecteurs. Si une famille est formée de plus de 3 vecteurs, elle ne peut pas être une base car elle ne peut pas être libre. Si une famille est formée de moins de 3 vecteurs, elle ne peut pas être une base car elle ne peut pas être génératrice.

- Mais attention ! Dans $\mathbb{R}^3$, une famille de 3 vecteurs n'est pas nécessairement une base. On a vu au premier exemple préliminaire de 2.2 qu'une famille de 3 vecteurs de $\mathbb{R}^3$ peut très bien ne pas être libre.
- Toujours dans $\mathbb{R}^3$, considérons une famille de 3 vecteurs. Pour qu'elle soit une base de $\mathbb{R}^3$, il faut a priori qu'elle soit libre et génératrice. Le résultat suivant montre qu'en fait il suffit de vérifier l'une des deux propriétés, l'autre étant alors automatiquement satisfaite.

• PROPOSITION 2 (*Un résultat très utile dans la pratique.*)
Soit F un ss-e.v. de $\mathbb{R}^n$. Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille de vecteurs de F.

<u>On suppose que le nombre p des vecteurs de F est exactement égal à la dimension de F.</u>

Alors, dans ce cas:

$$\boxed{\begin{array}{c}\mathcal{X} \text{ est une famille} \\ \text{génératrice de } F\end{array}} \iff \boxed{\mathcal{X} \text{ est une base de } F} \iff \boxed{\begin{array}{c}\mathcal{X} \text{ est une famille libre} \\ \text{de vecteurs de } F\end{array}}$$

PAR EXEMPLE, dans $\mathbb{R}^3$, pour montrer qu'une famille de 3 vecteurs est une base de $\mathbb{R}^3$, il suffit de montrer qu'elle est libre ou qu'elle est génératrice; si elle vérifie l'une des deux propriétés, elle vérifie nécessairement les deux.

2.5 Bases et somme directe

Les résultats suivants sont très utiles pour les questions de diagonalisation que l'on verra plus loin.

• PROPOSITION 1 (CAS DE DEUX SOUS-ESPACES).
Soient F et G deux ss-e.v. de $\mathbb{R}^n$. Notons $p = \dim F$ et $q = \dim G$. Les assertions suivantes sont équivalentes.

(i) $\mathbb{R}^n = F \oplus G$.

(ii) *Pour toute base $\mathcal{X}$ de F et toute base $\mathcal{Y}$ de G, la famille formée en adjoignant les p vecteurs de $\mathcal{X}$ et les q vecteurs de $\mathcal{Y}$ est une base de $\mathbb{R}^n$.*

(iii) *Il existe une base $\mathcal{X}$ de F et une base $\mathcal{Y}$ de G telles que la famille formée en adjoignant les p vecteurs de $\mathcal{X}$ et les q vecteurs de $\mathcal{Y}$ soit une base de $\mathbb{R}^n$.*

COROLLAIRE 1. *Si F et G sont deux ss-e.v. de $\mathbb{R}^n$ tels que $\mathbb{R}^n = F \oplus G$, alors:* $\dim F + \dim G = n$.

• PROPOSITION 2 (CAS D'UN NOMBRE QUELCONQUE DE SOUS-ESPACES).
Soient $F_1, F_2, \ldots, F_p$ des ss-e.v. de $\mathbb{R}^n$. Les assertions suivantes sont équivalentes.

(i) $\mathbb{R}^n = F_1 \oplus F_2 \oplus \cdots \oplus F_p$.

(ii) *Une base de $\mathbb{R}^n$ s'obtient en adjoignant entre elles une base de chacun des ss-e.v. F_i.*

COROLLAIRE 2. *Si $\mathbb{R}^n = F_1 \oplus F_2 \oplus \cdots \oplus F_p$, alors* $\dim F_1 + \dim F_2 + \cdots + \dim F_p = n$.

2.6 Théorème dit de la base incomplète

C'est un résultat théorique qui intervient comme argument dans de nombreux raisonnements.

• REMARQUE PRÉLIMINAIRE. Considérons dans $\mathbb{R}^n$ une famille libre $\mathcal{X}$, formée de p vecteurs.
- Comme $\mathcal{X}$ est libre, on sait d'après le point (i) de la proposition 1 de 2.4 que l'on ne peut pas avoir $p > n$.
- Si $p = n$, on sait d'après la proposition 2 de 2.4 que $\mathcal{X}$ est une base.
- Le cas restant $p < n$ fait l'objet du théorème suivant:

• THÉORÈME. *Considérons dans $\mathbb{R}^n$ une famille $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ de p vecteurs de $\mathbb{R}^n$ avec $p < n$. Si $\mathcal{X}$ est libre, alors il existe dans $\mathbb{R}^n$ des vecteurs $v_{p+1}, v_{p+2}, \ldots, v_n$ tels que la famille $(u_1, \ldots, u_p, v_{p+1}, \ldots, v_n)$ soit une base de $\mathbb{R}^n$.*

COMMENTAIRE.

1. Ainsi, toute famille libre est une "base en puissance" au sens que, si elle est "trop petite" pour être une base, on peut toujours lui ajouter des vecteurs pour construire une base.
2. Attention: cette façon de compléter n'est pas unique.
3. Attention aussi: cet énoncé n'a de sens que parce que l'on part d'une famille $\mathcal{X}$ qui est libre (il est trivialement faux sinon d'après le point (ii) de la proposition de 2.2).

• COROLLAIRE. *Tout ss-e.v. F de $\mathbb{R}^n$ admet (au moins) un supplémentaire dans $\mathbb{R}^n$.*

COMMENTAIRE. Là encore, il convient d'insister sur le fait qu'il n'y a pas du tout unicité. C'est une erreur souvent rencontrée que de parler "du" supplémentaire d'un ss-e.v. dans $\mathbb{R}^n$ alors qu'il faut parler "d'un" supplémentaire.

Matrices

1. Calcul matriciel.

1.1 Notion de matrice

• Définition. Soient n et p deux entiers ≥ 1. Une *matrice* à p *lignes* et n *colonnes* à coefficients dans $\mathbb{R}$ est une application: $A : \{1, 2, \dots, p\} \times \{1, 2, \dots, n\} \to \mathbb{R}$ qui, à tout couple (i, j), associe un nombre $a_{i,j}$ de $\mathbb{R}$. On la note sous la forme d'un tableau à p lignes et n colonnes:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{p,1} & a_{p,2} & a_{p,3} & \dots & a_{p,n-1} & a_{p,n} \end{pmatrix},$$

ou encore de façon générique:

$$A = (a_{i,j})_{1 \leq i \leq p,\, 1 \leq j \leq n}.$$

Le premier indice est l'indice de ligne, le second l'indice de colonne.
Les termes $a_{i,i}$ sont appelés les termes diagonaux de A.

• Notation et terminologie.
On note $\mathcal{M}_{p,n}(\mathbb{R})$ l'ensemble des matrices à p lignes et n colonnes à coefficients dans $\mathbb{R}$.
Pour $p = 1$, une matrice de $\mathcal{M}_{1,n}(\mathbb{R})$ s'appelle une *matrice ligne* (c'est un n-uplet).
Pour $n = 1$, une matrice de $\mathcal{M}_{p,1}(\mathbb{R})$ s'appelle une *matrice colonne* (c'est un p-uplet).
Pour $p = n$, une matrice de $\mathcal{M}_{n,n}(\mathbb{R})$ s'appelle une *matrice carrée d'ordre* n.

Exemples.
$$\begin{pmatrix} \pi & -1 \\ 0 & \sqrt{2} \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R}), \quad \begin{pmatrix} 2 & \sqrt{5} & -7 \\ 1 & -2 & -2 \\ 3 & 0 & 9 \\ 0 & 1 & \frac{1}{4} \end{pmatrix} \in \mathcal{M}_{4,3}(\mathbb{R}), \quad \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \in \mathcal{M}_{3,3}(\mathbb{R}), \quad \begin{pmatrix} 5 \\ \frac{1}{2} \\ 0 \\ -1 \end{pmatrix} \in \mathcal{M}_{4,1}(\mathbb{R}).$$

1.2 Opérations sur les matrices.

• Addition des matrices et produit externe d'une matrice par un réel.

Ces deux opérations sont définies de façon naturelle et ne posent pas de difficultés: on peut additionner terme à terme (ie. coefficient par coefficient) deux matrices ayant le même nombre de lignes et le même nombre de colonnes; on peut multiplier une matrice A par un scalaire quelconque α en multipliant par α chaque coefficient de A. On peut donc faire des combinaisons linéaires de matrices de même format.

Exemple.
$$3.\begin{pmatrix} 5 & 1 & 2 \\ 8 & 9 & 3 \\ 1 & -6 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 15 & 3 & 6 \\ 24 & 27 & 9 \\ 3 & -18 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 14 & 3 & 6 \\ 24 & 27 & 8 \\ 3 & -19 & 3 \end{pmatrix},$$

On appelle *matrice nulle* dans $\mathcal{M}_{p,n}(\mathbb{R})$ la matrice $O_{p,n}$ de $\mathcal{M}_{p,n}(\mathbb{R})$ dont tous les coefficients sont nuls. C'est bien sûr l'élément neutre de l'addition:

$$A + O_{p,n} = O_{p,n} + A = A \text{ pour toute } A \in \mathcal{M}_{p,n}(\mathbb{R}).$$

Les règles de calcul suivantes sont évidentes, pour toutes matrices A, B, C dans $\mathcal{M}_{p,n}(\mathbb{R})$ et tous $\alpha, \beta \in \mathbb{R}$:

$$\begin{array}{lll} A + B = B + A & A + (B + C) = (A + B) + C & A - B = -(B - A) \\ (\alpha + \beta).A = \alpha.A + \beta.A & \alpha.(A + B) = \alpha.A + \alpha.B & \alpha.(\beta.A) = (\alpha\beta).A \\ 1.A = A & (-\alpha).A = \alpha.(-A) = -\alpha.A & \alpha.(A - B) = \alpha.A - \alpha.B \end{array}.$$

• Produit matriciel

Sa définition est plus complexe et peut sembler a priori artificielle. Son sens s'éclaircira plus loin en relation avec la composition des applications linéaires.

Définition. Remarquons avant tout que le produit d'une matrice A par une matrice B n'est défini que lorsque le nombre de colonnes de A est égal au nombre de lignes de B. Le coefficient de AB sur la i-ième ligne et la j-ième colonne s'obtient en faisant la somme des n produits terme à terme de la i-ième ligne de A par la j-ième colonne de B.

$$\underbrace{\begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \boxed{a_{i,1}} & \boxed{a_{i,2}} & \dots & \boxed{a_{i,n}} \\ * & * & \dots & * \\ * & * & \dots & * \end{pmatrix}}_{A \text{ a } p \text{ lignes et } n \text{ colonnes}} \times \underbrace{\begin{pmatrix} * & * & \boxed{b_{1,j}} & * & * \\ * & * & \boxed{b_{2,j}} & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & \boxed{b_{n,j}} & * & * \end{pmatrix}}_{B \text{ a } n \text{ lignes et } q \text{ colonnes}} = \underbrace{\begin{pmatrix} * & * & \vdots & * & * \\ * & * & \vdots & * & * \\ \dots & \dots & \boxed{c_{i,j}} & \dots & \dots \\ * & * & \vdots & * & * \\ * & * & \vdots & * & * \end{pmatrix}}_{C = AB \text{ a } p \text{ lignes et } q \text{ colonnes}}.$$

Exemples.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{pmatrix} = \begin{pmatrix} \alpha+2\alpha'+3\alpha'' & \beta+2\beta'+3\beta'' & \gamma+2\gamma'+3\gamma'' \\ 4\alpha+5\alpha'+6\alpha'' & 4\beta+5\beta'+6\beta'' & 4\gamma+5\gamma'+6\gamma'' \end{pmatrix};$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{pmatrix}; \qquad \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & a+c \\ d & e & d+f \\ g & h & g+i \end{pmatrix};$$

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}\begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} a+4b \\ c+4d \\ e+4f \end{pmatrix}; \qquad \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} a+2c+3e & b+2d+3f \end{pmatrix}.$$

Attention ! Le produit AB peut exister sans que le produit BA soit défini; et même si les deux existent, ils ne sont en général pas égaux (cf. les exemples 2 et 3 précédents).

Dans la pratique, on aura besoin de faire le produit:
- 1) d'une matrice carrée d'ordre n par une autre matrice carrée de même ordre n,
- 2) d'une matrice carrée d'ordre n par une matrice colonne à n lignes.

Détaillons donc (en prenant $n = 3$ pour fixer les idées) les formules du produit dans ces cas:

$$\boxed{\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax+by+cz \\ a'x+b'y+c'z \\ a''x+b''y+c''z \end{pmatrix}}$$

$$\boxed{\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}\begin{pmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{pmatrix} = \begin{pmatrix} a\alpha+b\alpha'+c\alpha'' & a\beta+b\beta'+c\beta'' & a\gamma+b\gamma'+c\gamma'' \\ a'\alpha+b'\alpha'+c'\alpha'' & a'\beta+b'\beta'+c'\beta'' & a'\gamma+b'\gamma'+c'\gamma'' \\ a''\alpha+b''\alpha'+c''\alpha'' & a''\beta+b''\beta'+c''\beta'' & a''\gamma+b''\gamma'+c''\gamma'' \end{pmatrix}}$$

1.3 Algèbre des matrices carrées

• Produit des matrices carrées. Soit n un entier ≥ 1 fixé. On note $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{R}$. D'après la définition du 1.2 ci-dessus, pour toutes matrices carrées A et B d'ordre n, les produits AB et BA sont tous les deux définis; donc le produit matriciel définit un produit interne (une multiplication) dans $\mathcal{M}_n(\mathbb{R})$.

Les règles de calcul suivantes sont évidentes, pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{R})$ et tout $\alpha \in \mathbb{R}$:

$$A(BC) = (AB)C, \quad A(B+C) = AB + AC \text{ et } (A+B)C = AC + BC, \quad (\alpha.A)B = A(\alpha.B) = \alpha.(AB).$$

Attention ! erreur grave à ne pas commettre: on a vu qu'en général, $AB \neq BA$.

• Matrice identité.

On note I_n et on appelle matrice identité d'ordre n la matrice dont les coefficients sont égaux à 1 sur la diagonale et à 0 partout ailleurs.

$$\boxed{I_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}}$$

Il est immédiat que I_n est neutre pour le produit matriciel dans $\mathcal{M}_n(\mathbb{R})$, c'est-à-dire que l'on a pour tout $A \in \mathcal{M}_n(\mathbb{R})$:

$$\boxed{A \times I_n = I_n \times A = A}$$

• Matrice nulle.

On note O_n et on appelle matrice nulle d'ordre n la matrice dont tous les coefficients sont égaux à 0 .

$$\boxed{O_n = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}}$$

Il est immédiat que O_n est neutre pour l'addition dans $\mathcal{M}_n(\mathbb{R})$, et absorbant pour le produit matriciel, c'est-à-dire que l'on a pour tout $A \in \mathcal{M}_n(\mathbb{R})$:

$$\boxed{A + O_n = O_n + A = A} \quad \text{et} \quad \boxed{A \times O_n = O_n \times A = O_n}$$

Attention ! Erreur grave à ne pas commettre. Le produit de deux matrices dans $\mathcal{M}_n(\mathbb{R})$ peut être nul sans qu'aucune des deux ne soit nulle. Par exemple:

$$\begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{et pourtant } \begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \ \begin{pmatrix} 0 & 0 \\ 4 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

• La question de l'inversibilité.

Une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est *inversible* lorsqu'il existe dans $\mathcal{M}_n(\mathbb{R})$ une matrice, notée A^{-1}, vérifiant:

$$\boxed{A \times A^{-1} = A^{-1} \times A = I_n}.$$

- Lorsqu'il existe, un tel inverse est nécessairement unique; on dira donc que A^{-1} est *la* matrice inverse de A.
- Il suffit que $AA^{-1} = I_n$ ou que $A^{-1}A = I_n$ pour conclure que A est inversible d'inverse A^{-1}.

Exemples.

La matrice $A = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ est inversible, d'inverse $A^{-1} = \begin{pmatrix} 3 & -2 & 1 \\ -5 & 4 & -2 \\ 1 & -1 & 1 \end{pmatrix}$. Il suffit de calculer le produit de ces deux matrices et de vérifier qu'il vaut I_3.

En revanche, la matrice $B = \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix}$ n'est pas inversible dans $\mathcal{M}_2(\mathbb{R})$. En effet, sinon il existerait une matrice $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$ telle que $\begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. En identifiant la première colonne dans chaque membre, on aurait $3x + 6z = 1$ et $2x + 4z = 0$, ce qui est incompatible.

On verra plus tard des méthodes concrètes d'une part pour déterminer si une matrice carrée donnée est ou non inversible, d'autre part pour calculer son inverse lorsqu'il existe. Contentons-nous ici de signaler les 3 propriétés suivantes:

- La matrice identité I_n est inversible, et $I_n^{-1} = I_n$.
- Si A est inversible, alors A^{-1} est inversible, et $(A^{-1})^{-1} = A$.
- Si A et B sont inversibles, alors AB est inversible, et l'on a:

$$\boxed{(AB)^{-1} = B^{-1}A^{-1}.} \qquad \longleftarrow \textit{Attention à l'ordre !}$$

1.5 Quelques compléments sur les matrices carrées

• Matrices semblables

Deux matrices carrées A et B d'ordre n sont dites *semblables* lorsqu'il existe une matrice carrée P d'ordre n inversible vérifiant $A = PBP^{-1}$.

Ce n'est pour l'instant qu'une définition, mais on verra que cette notion prend tout son sens et tout son intérêt lors de l'étude des changements de bases.

Remarque. L'observation suivante, bien qu'élémentaire, aura plus loin d'importantes applications:

$$\text{si } A = PBP^{-1}, \text{ alors } \ A^m = PB^mP^{-1} \text{ pour tout entier } m \geq 1.$$

En effet: $A^m = (PBP^{-1})^m = (PBP^{-1})(PBP^{-1})(PBP^{-1})\dots(PBP^{-1})$
$= PB(P^{-1}P)B(P^{-1}P)B(P^{-1}\dots P)BP^{-1} = PBI_nBI_nB\dots BP^{-1} = PB^mP^{-1}$.

• Matrices diagonales, matrices triangulaires.

Une matrice A carrée d'ordre n est dite *diagonale* lorsque tous ses coefficients non diagonaux sont nuls. Elle est dite *triangulaire supérieure* lorsque tous les coefficients situés sous la diagonale sont nuls.

Par exemple, $A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ est diagonale et $B = \begin{pmatrix} 3 & 4 & -7 & 9 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ est triangulaire.

L'un des intérêts des matrices carrées diagonales est la simplicité des calculs. En particulier, il est clair que: *si A et B sont deux matrices diagonales, alors on a: $AB = BA$.* En effet, on calcule:

$$AB = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \dots & 0 \\ 0 & a_2b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_nb_n \end{pmatrix} = \begin{pmatrix} b_1a_1 & 0 & \dots & 0 \\ 0 & b_2a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_na_n \end{pmatrix} = BA.$$

Il est clair par ailleurs que le produit de deux matrices triangulaires supérieures est encore une matrice triangulaire supérieure.

• Trace d'une matrice carrée

Par définition, la *trace* d'une matrice carrée A, notée $\operatorname{tr} A$, est la somme de ses coefficients diagonaux. Donc:

$$\text{si } A = (a_{i,j})_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{R}), \quad \operatorname{tr} A = a_{1,1} + a_{2,2} + \dots + a_{n,n} = \sum_{i=1}^{n} a_{i,i} \in \mathbb{R}.$$

Propriétés. On a les propriétés suivantes de la trace:

(i) Pour toutes $A, B \in \mathcal{M}_n(\mathbb{R})$ et tout $\alpha \in \mathbb{R}$, $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$ et $\operatorname{tr}(\alpha.A) = \alpha \operatorname{tr} A$.

(ii) Pour toutes $A, B \in \mathcal{M}_n(\mathbb{R})$, $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

(iii) Pour toute $A \in \mathcal{M}_n(\mathbb{R})$ et toute toute $P \in \mathcal{M}_n(\mathbb{R})$ inversible, $\operatorname{tr}(PAP^{-1}) = \operatorname{tr}(A)$.

Le point (iii), qui découle directement de (ii), traduit le fait que: *deux matrices carrées semblables ont toujours la même trace.*

2. Matrices et applications linéaires.

2.1 Applications linéaires et endomorphismes

D'une façon générale, une application $f : \mathbb{R}^n \to \mathbb{R}^p$ est dite *linéaire* si elle respecte l'addition des vecteurs et le produit externe d'un vecteur par un réel, c'est-à-dire $f(u + v) = f(u) + f(v)$ et $f(\alpha.u) = \alpha.f(u)$ pour tous $u, v \in \mathbb{R}^n$ et tout $\alpha \in \mathbb{R}$.

On s'intéressera dans la suite principalement au cas où $n = p$. On dit dans ce cas que f est un *endomorphisme*.

Un endomorphisme de $\mathbb{R}^n$ est une application $f : \mathbb{R}^n \to \mathbb{R}^n$ linéaire, c'est-à-dire telle que

$$f(u + v) = f(u) + f(v) \quad \text{et} \quad f(\alpha.u) = \alpha.f(u)$$

pour tous $u, v \in \mathbb{R}^n$ et tout $\alpha \in \mathbb{R}$

Observation fondamentale: pour connaître complètement un endomorphisme $f : \mathbb{R}^n \to \mathbb{R}^n$, il suffit de connaître les images par f des vecteurs d'une base de $\mathbb{R}^n$.

En effet: donnons-nous une base $\mathcal{B} = (u_1, u_2, \dots, u_n)$ de $\mathbb{R}^n$. Soit v un vecteur quelconque de $\mathbb{R}^n$. Grâce à ses composantes $(\alpha_1, \alpha_2, \dots, \alpha_n)$ dans la base $\mathcal{B}$, on sait qu'il se décompose en $v = \alpha_1.u_1 + \alpha_2.u_2 + \dots + \alpha_n.u_n$. On en tire $f(v) = f(\alpha_1.u_1 + \alpha_2.u_2 + \dots + \alpha_n.u_n)$. Mais la linéarité de f permet de transformer cette expression en $f(v) = \alpha_1.f(u_1) + \alpha_2.f(u_2) + \dots + \alpha_n.f(u_n)$. Il suffit donc de connaître les vecteurs $f(u_1), f(u_2), \dots, f(u_n)$ pour pouvoir déterminer $f(v)$ pour n'importe quel vecteur $v \in \mathbb{R}^n$.

2.2 Matrices d'un endomorphisme

• Exemple préliminaire. Soit $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$, avec donc $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$. Soit f un endomorphisme de $\mathbb{R}^3$. D'après ce qui précède, f est entièrement déterminée par les images de e_1, e_2 et e_3. Choisissons par exemple:

$$f(e_1) = (2, 1, 0) = 2.e_1 + e_2,\ f(e_2) = (-1, 2, 3) = -e_1 + 2.e_2 + 3.e_3 \text{ et } f(e_3) = (1, -1, -1) = e_1 - e_2 - e_3.$$

Pour tout vecteur $(x, y, z) \in \mathbb{R}^3$, on calcule: $f(x, y, z) = f(x.e_1 + y.e_2 + z.e_3) = x.f(e_1) + y.f(e_2) + z.f(e_3) = x.(2, 1, 0) + y.(-1, 2, 3) + z.(1, -1, -1) = (2x - y + z, x + 2y - z, 3y - z)$.

On peut donc écrire matriciellement: $$\begin{pmatrix} 2x - y + z \\ x + 2y - z \\ 3y - z \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 0 & 3 & -1 \end{pmatrix}}_{\text{soit } A} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

La matrice A s'appelle la matrice de f par rapport à la base $\mathcal{B}$. Elle définit complètement l'application f puisqu'elle permet de calculer l'image par f de tout vecteur de $\mathbb{R}^3$. Les trois colonnes de A sont formées des composantes dans la base $\mathcal{B} = (e_1, e_2, e_3)$ des vecteurs $f(e_1)$, $f(e_2)$ et $f(e_3)$, (dans le même ordre !).

• DÉFINITION.

Soit $\mathcal{B} = (u_1, \dots, u_n)$ une base de $\mathbb{R}^n$. Soit f un endomorphisme de $\mathbb{R}^n$. On appelle *matrice de f par rapport à la base* $\mathcal{B}$ la matrice carrée $M_{\mathcal{B}}(f)$ à n lignes et n colonnes dont les colonnes sont formées des composantes des vecteurs $f(u_1), f(u_2), \dots, f(u_n)$ décomposés suivant les vecteurs $u_1, u_2, \dots, u_n$ de la base $\mathcal{B}$.

$$M_{\mathcal{B}}(f) = \begin{array}{c} \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & & \vdots \\ * & * & \dots & * \end{pmatrix} \begin{array}{l} \leftarrow u_1 \\ \leftarrow u_2 \\ \quad\vdots \\ \leftarrow u_n \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ f(u_1) & f(u_2) & \dots & f(u_n) \end{array} \end{array}$$

En reprenant dans le cas général le raisonnement et les calculs faits sur un cas particulier dans l'exemple préliminaire ci-dessus, on obtient sans difficulté:

• PROPOSITION. *On reprend les données et notations précédentes; on note $A = M_{\mathcal{B}}(f)$. Soit u un vecteur quelconque de $\mathbb{R}^n$. Si l'on note X la matrice colonne des composantes de u dans la base $\mathcal{B}$, alors la matrice colonne Y des composantes de son image $f(u)$ dans la base $\mathcal{B}$ est donnée par le produit matriciel:*

$$\boxed{Y = AX}$$

EXEMPLE 1. Soit $\mathcal{B} = (u_1, u_2, u_3)$ une base de $\mathbb{R}^3$. Soit f l'endomorphisme de $\mathbb{R}^3$ défini par: $f(u_1) = u_1 + u_2 + u_3$, $f(u_2) = 2.u_1 - u_2$, $f(u_3) = 4.u_1 + u_3$, de sorte que: $A = M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 2 & 4 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

De plus, si $u \in \mathbb{R}^3$ a pour composantes (x, y, z) dans la base $\mathcal{B}$, les composantes (x', y', z') de $f(u)$ dans la base $\mathcal{B}$ sont:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{donc} \quad \begin{cases} x' = x + 2y + 4z \\ y' = x - y \\ z' = x + z \end{cases}$$

EXEMPLE 2. Soit f l'endomorphisme de $\mathbb{R}^3$ qui, à tout vecteur $u = (x, y, z)$ associe $f(u) = (x', y', z')$ défini par:

$$x' = x - y + 2z, \quad y' = -x + y + z, \quad z' = 3x + y - z.$$

Alors, en notant $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$, on a:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{donc} \quad M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix}, \quad \text{donc} \quad \begin{cases} f(e_1) = e_1 - e_2 + 3e_3, \\ f(e_2) = -e_1 + e_2 + e_3, \\ f(e_3) = 2e_1 + e_2 - e_3. \end{cases}$$

• REMARQUE. Il est aisé de vérifier que prendre la matrice d'un endomorphisme par rapport à une base est une méthode qui respecte l'addition et le produit externe naturellement défini sur les endomorphismes:

$$M_{\mathcal{B}}(f + g) = M_{\mathcal{B}}(f) + M_{\mathcal{B}}(g) \text{ et } M_{\mathcal{B}}(\alpha.f) = \alpha.M_{\mathcal{B}}(f).$$

Plus intéressante est la propriété suivante, qui explique la définition du produit matriciel:

$$M_{\mathcal{B}}(g \circ f) = M_{\mathcal{B}}(g) \times M_{\mathcal{B}}(f).$$

Rappelons aussi que l'on appelle *application identité* de $\mathbb{R}^n$ l'application qui, à tout vecteur $u \in \mathbb{R}^n$, associe u lui-même. Il est clair que sa matrice dans toute base est la matrice identité I_n définie en 1.3.

$$M_{\mathcal{B}}(\mathrm{Id}) = I_n.$$

Une application importante de ces résultats est la méthode de calcul de l'inverse d'une matrice carrée inversible.

2.3 Application au calcul de l'inverse d'une matrice

• PRINCIPE DE LA MÉTHODE.

Soit A une matrice carrée d'ordre n. On sait (voir 2.2) qu'on peut toujours la considérer comme la matrice d'un endomorphisme f par rapport à la base canonique $\mathcal{B}$ de $\mathbb{R}^n$. Cet endomorphisme f est défini par le fait que, pour tout vecteur $u \in \mathbb{R}^n$, si l'on note X la matrice colonne des composantes de u dans la base $\mathcal{B}$, alors la matrice colonne Y des composantes de $v = f(u)$ dans $\mathcal{B}$ est donnée par $Y = AX$.
Cette relation matricielle permet d'exprimer les termes y_i de Y en fonction des termes x_i de X.
Supposons que l'on parvienne, par le calcul, à "inverser formellement" le système de relations ainsi obtenu, c'està-dire à trouver une matrice carrée A' d'ordre n telle que la relation $Y = AX$ soit équivalente à $A'Y = X$. Alors cela prouve que A est inversible, et que l'on a $A' = A^{-1}$.

En effet. En terme d'endomorphismes l'existence de la matrice A' telle que $A'Y = X$ signifie qu'il existe un endomorphisme g, qui a A' pour matrice par rapport à $\mathcal{B}$, et qui vérifie que $u = g(v)$ lorsque $v = f(u)$. En d'autres termes $g(f(u)) = u$ pour tout $u \in \mathbb{R}^n$, c'est-à-dire $g \circ f = \mathrm{Id}$. On dit que f est bijectif et que sa bijection réciproque est g, que l'on note alors f^{-1}. Avec les propriétés rappelées à la fin de 2.2, on en tire $A' \times A = I_n$, d'où A inversible et $A' = A^{-1}$. En d'autres termes:

$$A \text{ inversible} \Leftrightarrow f \text{ bijectif}, \text{ et } M_{\mathcal{B}}(f^{-1}) = A^{-1}.$$

• Exemple. Soit $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$. L'endomorphisme f de $\mathbb{R}^3$ dont A est la matrice par rapport à la base canonique est celui qui, à un triplet quelconque $u = (x, y, z)$, associe $f(u) = (x', y', z')$ défini par:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad \text{On a : } \begin{cases} x' = x+2y \\ y' = y+z \\ z' = -x+z \end{cases} \Leftrightarrow \begin{cases} x' = x+2y \\ y' = y+z \\ z'+x' = 2y+z \end{cases} \Leftrightarrow \begin{cases} x' = x+2y \\ y' = y+z \\ z'+x'-2y' = -z \end{cases}$$

qui est encore équivalent, en remontant à partir de la dernière équation à:

$$\begin{cases} z = -x' + 2y' - z' \\ y = y' - z \\ x = x' - 2y \end{cases}, \quad \text{c'est-à-dire: } \begin{cases} z = -x' + 2y' - z' \\ y = x' - y' + z' \\ x = -x' + 2y' - 2z' \end{cases}, \quad \text{que l'on écrit: } \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 2 & -2 \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

On conclut que $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$ est inversible et que $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 2 & -2 \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix}$.

Remarques. Dans la pratique, il est prudent de vérifier le calcul en s'assurant que le produit de la matrice A de départ par la matrice A^{-1} que l'on vient de trouver est bien égal à la matrice identité.

On verra plus loin dans le cours d'une part une méthode permettant de tester si une matrice carrée est ou non inversible (par le déterminant), d'autre part une méthode beaucoup plus conceptuelle (utilisant aussi les déterminants) donnant une formule générale pour l'inverse (lorsqu'il existe). Cependant la méthode détaillée ci-dessus par inversion formelle d'un système d'équations linéaires, (qui consiste finalement à montrer qu'une matrice est inversible en trouvant son inverse), reste dans bien des cas concrets la plus rapide et efficace dans la pratique.

2.4 Noyau d'un endomorphisme

• Définition. Le *noyau* d'un endomorphisme f de $\mathbb{R}^n$ est l'ensemble des vecteurs dont l'image par f est égale au vecteur nul. On le note $\mathrm{Ker}\, f$. Il est facile de vérifier que c'est toujours un ss-e.v. de $\mathbb{R}^n$.

$$\boxed{\mathrm{Ker}\, f = \{\, u \in \mathbb{R}^n \,;\, f(u) = \vec{0}\,\}, \quad \text{ss-e.v. de } \mathbb{R}^n.}$$

Exemple 1. Soit f l'endomorphisme de $\mathbb{R}^3$ de matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -4 \\ 5 & 2 & -1 \end{pmatrix}$ par rapport à la base canonique.

Un vecteur $u = (x, y, z)$ appartient à $\mathrm{Ker}\, f$ si et seulement si $\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -4 \\ 5 & 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. On résout donc:

$$\begin{cases} x + y + z = 0 \\ 2x - y - 4z = 0 \\ 5x + 2y - z = 0 \end{cases} \xrightarrow[(3)-5\times(1)]{(2)-2\times(1)} \begin{cases} x + y + z = 0 \\ -3y - 6z = 0 \\ -3y - 6z = 0 \end{cases} \longrightarrow \begin{cases} y = -2z \\ x = -y - z = z \end{cases}$$

On conclut que $\mathrm{Ker}\, f = \{(z, -2z, z) \,;\, z \in \mathbb{R}\}$. C'est donc la droite de base (v_1), avec $v_1 = (1, -2, 1)$.

Exemple 2. Soit g l'endomorphisme de $\mathbb{R}^3$ de matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -4 \\ 5 & 2 & -2 \end{pmatrix}$ par rapport à la base canonique.

Un vecteur $u = (x, y, z)$ appartient à $\mathrm{Ker}\, f$ si et seulement si $\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -4 \\ 5 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. On résout donc:

$$\begin{cases} x + y + z = 0 \\ 2x - y - 4z = 0 \\ 5x + 2y - 2z = 0 \end{cases} \xrightarrow[(3)-5\times(1)]{(2)-2\times(1)} \begin{cases} x + y + z = 0 \\ -3y - 6z = 0 \\ -3y - 7z = 0 \end{cases} \xrightarrow{(3)-(2)} \begin{cases} x + y + z = 0 \\ -3y - 6z = 0 \\ -z = 0 \end{cases} \longrightarrow \begin{cases} z = 0 \\ y = -2z = 0 \\ x = -y - z = z = 0 \end{cases}$$

On conclut que $\mathrm{Ker}\, f = \{(0, 0, 0)\}$. C'est le sous-espace nul, c'est-à-dire le singleton formé du seul vecteur nul.

Le noyau d'un endomorphisme de $\mathbb{R}^3$ peut a priori être $\mathbb{R}^3$ tout entier (mais cela ne se produit que dans le cas trivial où f est l'endomorphisme nul, c'est-à-dire A est la matrice nulle), ou un plan vectoriel, ou une droite vectorielle, ou le sous-espace nul. Le théorème suivant exprime que ce dernier cas caractérise l'inversibilité de A.

• Théorème. *Soit f un endomorphisme de $\mathbb{R}^n$. Son noyau est réduit au vecteur nul si et seulement si sa matrice par rapport à une base quelconque de $\mathbb{R}^n$ est inversible.*

$$\text{Si } f \textit{ endomorphisme} \text{ de } \mathbb{R}^n, \text{ alors: } \boxed{\mathrm{Ker}\, f = \{\vec{0}\} \;\Leftrightarrow A = M_{\mathcal{B}}(f) \text{ est inversible.}}$$

Remarque. Rappelons, comme on l'a vu en 2.3, que ceci équivaut aussi à la bijectivité de l'endomorphisme f.

Changements de bases

1. Les principes du changement de bases.

1.1 Matrice d'une famille finie de vecteurs dans une base

• Définition. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de $\mathbb{R}^n$. Soit $\mathcal{C} = (v_1, v_2, \ldots, v_p)$ une famille finie de p vecteurs de $\mathbb{R}^n$. On appelle matrice de la famille $\mathcal{C}$ dans la base $\mathcal{B}$ la matrice à n lignes et p colonnes dont les colonnes sont constituées des composantes des vecteurs de $\mathcal{C}$ décomposés dans la base $\mathcal{B}$. C'est donc la matrice:

$$A = \begin{array}{cc} \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,p} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,p} \end{pmatrix} & \begin{array}{l} \leftarrow e_1 \\ \leftarrow e_2 \\ \vdots \\ \leftarrow e_n \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \ldots & v_p \end{array} & \end{array}$$

où l'on a noté
$v_i = a_{1,i}.e_1 + a_{2,i}.e_2 + \cdots + a_{n,i}.e_n$
pour tout $1 \leq i \leq p$.

• Remarques. Il est clair que réciproquement, toute matrice $A = (a_{i,j})_{1\leq i\leq n\,,\,1\leq j\leq p}$ donnée dans $\mathcal{M}_{n,p}(\mathbb{R})$ peut toujours être considérée comme la matrice dans la base canonique $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ de $\mathbb{R}^n$ de la famille $\mathcal{C} = (v_1, v_2, \ldots, v_p)$ de vecteurs de $\mathbb{R}^n$ définis par les *colonnes* de A:

$$v_1 = (a_{1,1}, a_{2,1}, \ldots, a_{n,1}),\ \ v_2 = (a_{1,2}, a_{2,2}, \ldots, a_{n,2}),\ \ \ldots,\ \ v_p = (a_{1,p}, a_{2,p}, \ldots, a_{n,p}).$$

Dans le cas où $p = n$, la matrice est carrée. On rappelle ci-dessous un critère permettant de reconnaître sur une telle matrice si une famille de n vecteurs dans $\mathbb{R}^n$ est ou non une base.

1.2 Matrice de passage

• Proposition. *Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de $\mathbb{R}^n$. Soit $\mathcal{B}' = (e'_1, e'_2, \ldots, e'_n)$ une famille de n vecteurs de $\mathbb{R}^n$. Soit P la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$, au sens de la définition ci-dessus. Alors:*

$$\boxed{\mathcal{B}' \text{ est une base de } \mathbb{R}^n} \iff \boxed{P \text{ est inversible.}}$$

• Définition. Dans ce cas, P est appelée la *matrice de passage* de la base $\mathcal{B}$ à la base $\mathcal{B}'$. On la note $\mathrm{Pass}_{\mathcal{B}\to\mathcal{B}'}$

C'est donc par définition même une matrice carrée d'ordre n inversible.

On peut en outre montrer que: $(\mathrm{Pass}_{\mathcal{B}\to\mathcal{B}'})^{-1} = \mathrm{Pass}_{\mathcal{B}'\to\mathcal{B}}$

$$P = \mathrm{Pass}_{\mathcal{B}\to\mathcal{B}'} = \begin{array}{cc} \begin{pmatrix} * & * & \ldots & * \\ * & * & \ldots & * \\ \vdots & \vdots & & \vdots \\ * & * & \ldots & * \end{pmatrix} & \begin{array}{l} \leftarrow e_1 \\ \leftarrow e_2 \\ \vdots \\ \leftarrow e_n \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ e'_1 & e'_2 & \ldots & e'_n \end{array} & \end{array}$$

1.3 Première formule de changement de bases

Elle a pour objet d'exprimer les composantes d'un vecteur dans une base en fonction des composantes du même vecteur dans une autre base à l'aide de la matrice de passage d'une base à l'autre. Plus précisément:

• Proposition. *Soient $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, e'_2, \ldots, e'_n)$ deux bases de $\mathbb{R}^n$. Soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Si, pour un vecteur u quelconque de $\mathbb{R}^n$, on note X la matrice colonne des composantes de u dans la base $\mathcal{B}$, et X' la matrice colonne des composantes de u dans la base $\mathcal{B}'$, alors on a:*

$$\boxed{X = PX'} \quad \text{ou encore} \quad \boxed{X' = P^{-1}X}.$$

• Remarque sur la démonstration. La formule $X = PX'$ ne fait que regrouper sous une forme "compacte" un calcul essentiellement élémentaire. Afin de bien comprendre cette formule, on va détailler ce calcul dans le cas particulier $n = 3$ (la preuve est exactement la même pour n quelconque).

Avec les données de l'énoncé, notons: $P = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$. On calcule:

$$\begin{aligned} u &= x'.e'_1 + y'.e'_2 + z'.e'_3 \quad \text{par définition de } X' \\ &= x'.(a.e_1 + a'.e_2 + a''.e_3) + y'.(b.e_1 + b'.e_2 + b''.e_3) + z'(c.e_1 + c'.e_2 + c''.e_3) \quad \text{par définition de } P \\ &= (x'a + y'b + z'c).e_1 + (x'a' + y'b' + z'c').e_2 + (x'a'' + y'b'' + z'c'').e_3 \quad \text{en calculant dans } \mathbb{R}^3. \end{aligned}$$

Comme par ailleurs $u = x.e_1 + y.e_2 + z.e_3$ par définition de X, on en déduit par unicité des composantes de u dans la base $\mathcal{B}$ que:

$$\begin{cases} x = ax' + by' + cz' \\ y = a'x' + b'y' + c'z' \\ z = a''x' + b''y' + c''z' \end{cases}, \quad \text{donc} \quad \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}}_{X'}.$$

1.4 Seconde formule de changement de bases

Elle a pour objet d'exprimer la matrice d'un endomorphisme par rapport à une base donnée en fonction de la matrice du même endomorphisme par rapport à une nouvelle base, et de la matrice de passage correspondante. Plus précisément:

• Proposition. *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de $\mathbb{R}^n$, et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Pour tout endomorphisme f de $\mathbb{R}^n$, on a dans $\mathcal{M}_n(\mathbb{R})$ les égalités:*

$$\boxed{M_{\mathcal{B}'}(f) = P^{-1}\, M_{\mathcal{B}}(f)\, P} \qquad \text{ou encore} \qquad \boxed{M_{\mathcal{B}}(f) = P\, M_{\mathcal{B}'}(f)\, P^{-1}}$$

Démonstration. Notons $A = M_{\mathcal{B}}(f)$ et $A' = M_{\mathcal{B}'}(f)$. Pour $u \in E$ quelconque, posons:

$$\left.\begin{array}{l} X \text{ la matrice colonne des composantes de } u \text{ dans la base } \mathcal{B}, \\ Y \text{ la matrice colonne des composantes de } f(u) \text{ dans la base } \mathcal{B}, \end{array}\right\} \text{ donc } Y = AX.$$

$$\left.\begin{array}{l} X' \text{ la matrice colonne des composantes de } u \text{ dans la base } \mathcal{B}', \\ Y' \text{ la matrice colonne des composantes de } f(u) \text{ dans la base } \mathcal{B}', \end{array}\right\} \text{ donc } Y' = A'X'.$$

Par ailleurs, il résulte de la première formule de changement de base de 1.3 que: $X = PX'$ et $Y = PY'$.

On déduit de ces quatre égalités que: $A'X' = Y' = P^{-1}Y = P^{-1}AX = P^{-1}APX'$.

L'égalité $A'X' = (P^{-1}AP)X'$ obtenue étant établie pour tout $u \in \mathbb{R}^n$, et donc pour toute matrice colonne X', on conclut que $A' = P^{-1}AP$.

• Exemple. Soit $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$. Soit $\mathcal{B}' = (e'_1, e'_2, e'_3)$ définie par:

$$e'_1 = e_1 + 2.e_2, \quad e'_2 = e_1 + e_2 + e_3, \quad e'_3 = 2.e_2 + e_3.$$

Soit f l'endomorphisme de $\mathbb{R}^3$ défini par son action sur les vecteurs de $\mathcal{B}$:

$$f(e_1) = e_1 + 2.e_2, \quad f(e_2) = -e_1 + e_2 + e_3 \quad \text{et} \quad f(e_3) = -3.e_2 + e_3.$$

On a donc $A = M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & -3 \\ 0 & 1 & 1 \end{pmatrix}$, et la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$ est $P = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$.

Suivant la technique usuelle, on montre que P est inversible avec $P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 & -2 \\ 2 & -1 & 2 \\ -2 & 1 & 1 \end{pmatrix}$.

Ceci prouve en particulier que $\mathcal{B}'$ est une base de $\mathbb{R}^3$ (d'après la proposition de 1.2), et donc P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. La seconde formule de changement de base donne alors:

$$A' = M_{\mathcal{B}'}(f) = P^{-1}AP = \frac{1}{3}\begin{pmatrix} 1 & 1 & -2 \\ 2 & -1 & 2 \\ -2 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & -3 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 & -2 & -5 \\ 0 & -1 & 5 \\ 0 & 4 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -1 & -4 & -9 \\ -2 & 4 & 3 \\ 8 & 2 & 6 \end{pmatrix}.$$

L'action de f sur les vecteurs de la base $\mathcal{B}'$ est donc:

$$f(e'_1) = -\tfrac{1}{3}.e'_1 - \tfrac{2}{3}.e'_2 + \tfrac{8}{3}.e'_3, \quad f(e'_2) = -\tfrac{4}{3}.e'_1 + \tfrac{4}{3}.e'_2 + \tfrac{2}{3}.e'_3 \quad \text{et} \quad f(e'_3) = -3.e'_1 + e'_2 + 2.e'_3.$$

2.1 Un principe général.

Rappelons d'abord que deux matrices carrées A et A' d'ordre n sont dites *semblables* lorsqu'il existe une matrice carrée d'ordre n *inversible* P telle que $A' = P^{-1}AP$.

D'après la proposition du 1.4 ci-dessus, cela signifie que A et A' représentent le même endomorphisme f de $\mathbb{R}^n$ dans deux bases différentes de $\mathbb{R}^n$, entre lesquelles la matrice de passage est P.

Dans de nombreuses situations concrètes où l'on est amené à travailler avec une matrice carrée A, on pourra avoir intérêt à introduire l'endomorphisme f de $\mathbb{R}^n$ dont A est la matrice par rapport à la base canonique $\mathcal{B}$, puis à chercher une autre base $\mathcal{B}'$ par rapport à laquelle la matrice A' du même f s'exprime sous une forme particulièrement simple (par exemple une matrice diagonale, ou au moins triangulaire...)

On travaille alors à la résolution de notre problème avec la matrice plus simple A', puis on revient à la base de départ pour interpréter le résultat, à l'aide des formules de changement de base.

Ce principe de se ramener grâce à un changement de bases à une forme matricielle plus simple, de type diagonal par exemple, n'est pas applicable à n'importe quelle matrice, et on verra plus tard dans le cours des méthodes permettant de déterminer systématiquement les cas où c'est possible, et de trouver les bases les plus adaptées.

Pour l'instant, on se limite à donner ci-dessous quelques exemples de type classique, en indiquant le changement de base à effectuer.

2.2 Exemple de calcul des puissances d'une matrice carrée.

On considère les deux suites de nombres réels $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$ définies par leurs premiers termes u_0 et v_0 et les relations de récurence:

$$u_{n+1} = -10u_n - 28v_n \text{ et } v_{n+1} = 6u_n + 16v_n \text{ pour tout } n \in \mathbb{N}.$$

Il s'agit de calculer les termes généraux u_n et v_n en fonction de u_0, v_0 et n.

On commence par écrire les deux relations de récurrence sous forme matricielle: $\left(\begin{smallmatrix} u_{n+1} \\ v_{n+1} \end{smallmatrix}\right) = \left(\begin{smallmatrix} -10 & -28 \\ 6 & 16 \end{smallmatrix}\right)\left(\begin{smallmatrix} u_n \\ v_n \end{smallmatrix}\right)$, ou encore:

$$X_{n+1} = AX_n \text{ avec } A = \left(\begin{smallmatrix} -10 & -28 \\ 6 & 16 \end{smallmatrix}\right) \text{ et } X_n = \left(\begin{smallmatrix} u_n \\ v_n \end{smallmatrix}\right) \text{ pour tout } n \in \mathbb{N}.$$

On en déduit que: $X_n = AX_{n-1} = A^2X_{n-2} = A^3X_{n-3} = \cdots = A^nX_0$.

On est ramené au problème de calculer A^n. Pour cela, on introduit l'endomorphisme f de $\mathbb{R}^2$ dont la matrice par rapport à la base canonique $\mathcal{B}$ est A. Considérons la famille $\mathcal{B}'$ constituée des deux vecteurs:

$$e_1' = -7.e_1 + 3.e_2, \quad e_2' = -2.e_1 + e_2.$$

On démontre par la méthode usuelle que $P = \left(\begin{smallmatrix} -7 & -2 \\ 3 & 1 \end{smallmatrix}\right)$ est inversible, d'inverse $P^{-1} = \left(\begin{smallmatrix} -1 & -2 \\ 3 & 7 \end{smallmatrix}\right)$, ce qui prouve que $\mathcal{B}'$ est une autre base de $\mathbb{R}^2$, et que P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. La seconde formule de changement de bases permet de calculer la matrice de f par rapport à $\mathcal{B}'$:

$$D = P^{-1}AP = \left(\begin{smallmatrix} -1 & -2 \\ 3 & 7 \end{smallmatrix}\right)\left(\begin{smallmatrix} -10 & -28 \\ 6 & 16 \end{smallmatrix}\right)\left(\begin{smallmatrix} -7 & -2 \\ 3 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 4 \end{smallmatrix}\right).$$

Parce que D est diagonale, on peut immédiatement calculer $D^n = \left(\begin{smallmatrix} 2^n & 0 \\ 0 & 4^n \end{smallmatrix}\right)$. Or, suivant une “astuce” que l'on a déjà rencontrée, la formule de changement de base $A = PDP^{-1}$ conduit à:

$$\begin{aligned} A^n &= (PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)D(P^{-1}P)\cdots D(P^{-1}P)DP^{-1} = PD^nP^{-1} \end{aligned}$$

On calcule donc: $A^n = PD^nP^{-1} = \left(\begin{smallmatrix} -7 & -2 \\ 3 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 2^n & 0 \\ 0 & 4^n \end{smallmatrix}\right)\left(\begin{smallmatrix} -1 & -2 \\ 3 & 7 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 7\times 2^n - 6\times 4^n & 7\times 2^{n+1} - 14\times 4^n \\ -3\times 2^n + 3\times 4^n & -3\times 2^{n+1} + 7\times 4^n \end{smallmatrix}\right)$

En reportant dans la relation de départ $X_n = A^nX_0$, on obtient finalement:

$$\begin{cases} u_n &= (7\times 2^n - 6\times 4^n)u_0 + (7\times 2^{n+1} - 14\times 4^n)v_0 \\ v_n &= (-3\times 2^n + 3\times 4^n)u_0 + (-3\times 2^{n+1} + 7\times 4^n)v_0 \end{cases}$$

2.3 Exemple de résolution de système différentiel linéaire.

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel:

$$(S)\ \begin{cases} x'(t) &=& -2x(t) &-& 2y(t) &+& 2z(t) \\ y'(t) &=& 3x(t) &+& 5y(t) && \\ z'(t) &=& x(t) &+& 2y(t) &+& 3z(t) \end{cases}.$$

On note matriciellement ce système sous la forme:

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} -2 & -2 & 2 \\ 3 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

On introduit l'endomorphisme f de $\mathbb{R}^3$ dont la matrice par rapport à la base canonique $\mathcal{B}$ est A. Considérons la famille $\mathcal{B}'$ constituée des trois vecteurs:

$$e'_1 = 2.e_1 - e_2, \quad e'_2 = e_1 - e_2 + e_3, \quad e'_3 = e_2 + e_3$$

On démontre par la méthode usuelle que $P = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ est inversible, d'inverse $P^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & 1 \end{pmatrix}$, ce qui prouve que $\mathcal{B}'$ est une autre base de $\mathbb{R}^3$, et que P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. La seconde formule de changement de bases permet de calculer la matrice de f par rapport à $\mathcal{B}'$:

$$D = P^{-1}AP = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & 1 \end{pmatrix}\begin{pmatrix} -2 & -2 & 2 \\ 3 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

On pose: $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où: $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$.

On a alors:

$$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = DU(t).$$

On est donc ramené à résoudre le système (beaucoup plus simple):

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}\begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \quad \text{c'est-à-dire:} \quad \begin{cases} u'(t) = -u(t) \\ v'(t) = 2v(t) \\ w'(t) = 5w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions de ce dernier système sont:

$$u(t) = \alpha e^{-t},\ v(t) = \beta e^{2t},\ w(t) = \gamma e^{5t}, \text{ pour } \alpha, \beta, \gamma \text{ décrivant } \mathbb{R}.$$

En revenant à $X(t) = PU(t) = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} \alpha e^{-t} \\ \beta e^{2t} \\ \gamma e^{5t} \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont:

$$\begin{cases} x(t) &=& 2\alpha e^{-t} &+& \beta e^{2t} && \\ y(t) &=& -\alpha e^{-t} &-& \beta e^{2t} &+& \gamma e^{5t}, \\ z(t) &=& && \beta e^{2t} &+& \gamma e^{5t} \end{cases} \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

Remarque. Sur cet exemple, la matrice A du système était particulièrement simple car semblable à une matrice diagonale. On verra plus loin dans le cours que c'est loin d'être toujours le cas. Donnons un autre exemple, un peu plus compliqué, où la matrice n'est pas semblable à une matrice diagonale, mais est quand même semblable à une matrice triangulaire.

2.4 Un autre exemple de résolution de système différentiel linéaire.

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel:

$$(S)\ \begin{cases} x'(t) &=& -y(t) &-& 2z(t) \\ y'(t) &=& x(t) &+& z(t) \\ 5z'(t) &=& -6x(t) &+& 8y(t) \end{cases}.$$

On note matriciellement ce système sous la forme:

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ -\frac{6}{5} & \frac{8}{5} & 0 \end{pmatrix}.$$

On introduit l'endomorphisme f de $\mathbb{R}^3$ dont la matrice par rapport à la base canonique $\mathcal{B}$ est A. Considérons la famille $\mathcal{B}'$ constituée des trois vecteurs:

$$e_1' = 3.e_1 - 4.e_2 + 5.e_3, \quad e_2' = 3.e_1 + e_2 - 2.e_3, \quad e_3' = -e_1 - 2.e_2$$

On démontre par la méthode usuelle que $P = \begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix}$ est inversible, d'inverse $P^{-1} = -\frac{1}{45}\begin{pmatrix} -4 & 2 & -5 \\ -10 & 5 & 10 \\ 3 & 21 & 15 \end{pmatrix}$, ce qui prouve que $\mathcal{B}'$ est une autre base de $\mathbb{R}^3$, et que P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. La seconde formule de changement de bases permet de calculer la matrice de f par rapport à $\mathcal{B}'$:

$$T = P^{-1}AP = -\frac{1}{45}\begin{pmatrix} -4 & 2 & -5 \\ -10 & 5 & 10 \\ 3 & 21 & 15 \end{pmatrix}\begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ -\frac{6}{5} & \frac{8}{5} & 0 \end{pmatrix}\begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

On pose: $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où: $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$. Donc:

$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = TU(t)$.

On est ainsi ramené à résoudre le système (beaucoup plus simple):

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \quad \text{c'est-à-dire:} \quad \begin{cases} u'(t) = -2u(t) \\ v'(t) = v(t) + w(t) \\ w'(t) = w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions des équations (1) et (3) sont $u(t) = \alpha e^{-2t}$ et $w(t) = \gamma e^t$ pour α, γ décrivant $\mathbb{R}$. La seconde équation devient $v'(t) = v(t) + \gamma e^t$, dont la solution générale est $v(t) = (\gamma t + \beta)e^t$ pour β décrivant $\mathbb{R}$.

En revenant à $X(t) = PU(t) = \begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix}\begin{pmatrix} \alpha e^{-2t} \\ (\gamma t+\beta)e^t \\ \gamma e^t \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont:

$$\begin{cases} x(t) = 3\alpha e^{-2t} + (3\gamma t + 3\beta - \gamma)e^t \\ y(t) = -4\alpha e^{-2t} + (\gamma t + \beta - 2\gamma)e^t \\ z(t) = 5\alpha e^{-2t} + (-2\gamma t - 2\beta)e^t \end{cases}, \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

Déterminants

1. Définition et propriétés.

1.1 Exemple préliminaire des matrices carrées d'ordre 2.

• Définition. *Pour toute matrice carrée* $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{M}_2(\mathbb{R})$, *on appelle déterminant de* A, *noté* $\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right|$, *ou encore* $\det A$, *le nombre réel défini par:*

$$\det A = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| = ad - bc \in \mathbb{R}.$$

• Quelques propriétés. Les observations suivantes découlent immédiatement de la définition précédente:

(i) Si on échange les deux lignes ou les deux colonnes, on multiplie le déterminant par -1:

$$\left|\begin{smallmatrix} c & d \\ a & b \end{smallmatrix}\right| = cb - da = -(ad-bc) = -\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| \quad \text{et} \quad \left|\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right| = bc - ad = -(ad-bc) = -\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right|.$$

(ii) Le déterminant est linéaire par rapport à chaque ligne ou chaque colonne:

$$\left|\begin{smallmatrix} \lambda a+\mu a' & b \\ \lambda c+\mu c' & d \end{smallmatrix}\right| = (\lambda a+\mu a')d - b(\lambda c+\mu c') = \lambda(ad-bc) + \mu(a'd-bc') = \lambda\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| + \mu\left|\begin{smallmatrix} a' & b \\ c' & d \end{smallmatrix}\right|,$$

de même: $\left|\begin{smallmatrix} a & \lambda b+\mu b' \\ c & \lambda d+\mu d' \end{smallmatrix}\right| = \lambda\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| + \mu\left|\begin{smallmatrix} a & b' \\ c & d' \end{smallmatrix}\right|$,

et encore: $\left|\begin{smallmatrix} \lambda a+\mu a' & \lambda b+\mu b' \\ c & d \end{smallmatrix}\right| = \lambda\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| + \mu\left|\begin{smallmatrix} a' & b' \\ c & d \end{smallmatrix}\right|$, et $\left|\begin{smallmatrix} a & b \\ \lambda c+\mu c' & \lambda d+\mu d' \end{smallmatrix}\right| = \lambda\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| + \mu\left|\begin{smallmatrix} a & b \\ c' & d' \end{smallmatrix}\right|$.

(iii) En particulier, si on multiplie une ligne ou une colonne par un scalaire λ, on multiplie le déterminant par λ:

$$\left|\begin{smallmatrix} \lambda a & \lambda b \\ c & d \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & b \\ \lambda c & \lambda d \end{smallmatrix}\right| = \left|\begin{smallmatrix} \lambda a & b \\ \lambda c & d \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & \lambda b \\ c & \lambda d \end{smallmatrix}\right| = \lambda\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right|.$$

Il en résulte que si A a une ligne ou une colonne formée de 0, alors $\det A$ est nul:

$$\left|\begin{smallmatrix} 0 & 0 \\ c & d \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & b \\ 0 & 0 \end{smallmatrix}\right| = \left|\begin{smallmatrix} 0 & b \\ 0 & d \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & 0 \\ c & 0 \end{smallmatrix}\right| = 0.$$

Il en résulte aussi que, si on multiplie la matrice A par un scalaire λ, on multiplie le déterminant par λ^2:

$$\det(\lambda A) = \left|\begin{smallmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{smallmatrix}\right| = \lambda^2\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| = \lambda^2 \det A.$$

(iv) Si les deux lignes ou les deux colonnes sont égales, ou plus généralement si les deux lignes ou les deux colonnes sont multiples l'une de l'autre par un scalaire, alors le déterminant est nul:

$$\left|\begin{smallmatrix} a & b \\ a & b \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & a \\ c & c \end{smallmatrix}\right| = 0. \qquad \left|\begin{smallmatrix} a & b \\ \lambda a & \lambda b \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & \lambda a \\ c & \lambda c \end{smallmatrix}\right| = 0.$$

(v) D'après le deuxième et le quatrième point, on ne change pas la valeur du déterminant si on ajoute à une ligne (ou une colonne) un multiple de l'autre ligne (ou colonne) par un scalaire λ:

$$\left|\begin{smallmatrix} a & b \\ c+\lambda a & d+\lambda b \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| + \lambda\left|\begin{smallmatrix} a & b \\ a & b \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| \quad \text{et} \quad \left|\begin{smallmatrix} a & b+\lambda a \\ c & d+\lambda c \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| + \lambda\left|\begin{smallmatrix} a & a \\ c & c \end{smallmatrix}\right| = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right|.$$

• Multiplicativité.

Soient $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ deux matrices de $\mathcal{M}_2(\mathbb{R})$. Alors: $\boxed{\det(AB) = \det A \times \det B.}$

En effet: $AB = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) = \left(\begin{smallmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{smallmatrix}\right)$,

donc $\det(AB) = \left|\begin{smallmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{smallmatrix}\right| = (aa'+bc')(cb'+dd') - (ab'+bd')(ca'+dc')$

$= aa'dd' - adb'c' - bca'd' + bcb'c' = (ad-bc)(a'd'-b'c') = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| \times \left|\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right|$.

• Inversibilité.

Une matrice $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ de $\mathcal{M}_2(\mathbb{R})$ est inversible si et seulement si $\det A \neq 0$ dans $\mathbb{R}$, et dans ce cas:

$$\boxed{\det(A^{-1}) = \tfrac{1}{\det A}} \quad \text{et} \quad \boxed{A^{-1} = \tfrac{1}{\det A}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right).}$$

En effet: si $ad - bc \neq 0$, on vérifie immédiatement que:

$$\tfrac{1}{ad-bc}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \tfrac{1}{ad-bc}\left(\begin{smallmatrix} ad-bc & 0 \\ 0 & ad-bc \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = I_2.$$

Réciproquement, si A est inversible, on a $A^{-1} \times A = I_2$, donc :
$\det(A^{-1})\det A = \det I_2 = 1$, d'où $\det A \neq 0$ et $\det(A^{-1}) = \frac{1}{\det A}$.

1.2 Déterminant d'une matrice carrée d'ordre n.

On a défini le déterminant pour les matrices carrées d'ordre 2. Par convention, pour une matrice carrée d'ordre 1 (une telle matrice est réduite à un coefficient), c'est-à-dire $A = (a)$ avec $a \in \mathbb{R}$, on pose $\det A = a$. On va maintenant définir le déterminant pour les matrices carrées d'ordre n quelconque, par récurrence sur n, de la façon suivante.

• Définition du déterminant. (*Développement par rapport à la première ligne*)

Supposons que l'on ait défini le déterminant des matrices carrées d'ordre $n-1$, pour un certain $n \geq 3$. Fixons dans $\mathcal{M}_n(\mathbb{R})$ une matrice carrée:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n} \end{pmatrix}. \text{ On veut définir: } \det A = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n} \end{vmatrix}.$$

Pour tous $1 \leq i, j \leq n$, notons $A_{i,j}$ la matrice extraite de A obtenue en supprimant dans A la i-ième ligne et la j-ième colonne; c'est donc une matrice carrée de $\mathcal{M}_{n-1}(\mathbb{R})$, pour laquelle la notion de déterminant est supposée être définie. On appelle alors *déterminant* de A le nombre réel:

$$\det A = a_{1,1} \det A_{1,1} - a_{1,2} \det A_{1,2} + a_{1,3} \det A_{1,3} + \ldots + (-1)^{1+n} a_{1,n} \det A_{1,n} = \sum_{k=1}^{n} (-1)^{1+k} a_{1,k} \det A_{1,k}.$$

Cette formule est dite *développement du déterminant par rapport à la première ligne.*

• Proposition. (*Permutation des lignes*)

Si on échange deux lignes d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$, on multiplie son déterminant par -1.

C'est le cas général de la propriété montrée au (i) de 1.1 pour les matrices d'ordre 2.

• Conséquence. (*Développement par rapport à une ligne quelconque*)

Il résulte du point précédent que, pour tout $1 \leq i \leq n$, on a :

$$\boxed{\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j}.}$$

Cette formule est dite *développement du déterminant par rapport à la i-ième ligne.*

Pour tous $1 \leq i, j \leq n$, le réel: $\boxed{\Delta_{i,j} = (-1)^{i+j} \det A_{i,j}}$ est appelé le *cofacteur* associé au coefficient $a_{i,j}$.

• Exemple du déterminant d'une matrice carrée d'ordre 3.

Par exemple, pour $n = 3$, le déterminant $\delta = \begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}$ est égal à chacun des trois développements:

$$\delta = a \begin{vmatrix} b' & b'' \\ c' & c'' \end{vmatrix} - a' \begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} + a'' \begin{vmatrix} b & b' \\ c & c' \end{vmatrix} = a(b'c'' - b''c') - a'(bc'' - b''c) + a''(bc' - b'c).$$

$$\delta = -b \begin{vmatrix} a' & a'' \\ c' & c'' \end{vmatrix} + b' \begin{vmatrix} a & a'' \\ c & c'' \end{vmatrix} - b'' \begin{vmatrix} a & a' \\ c & c' \end{vmatrix} = -b(a'c'' - a''c') + b'(ac'' - a''c) - b''(ac' - a'c).$$

$$\delta = c \begin{vmatrix} a' & a'' \\ b' & b'' \end{vmatrix} - c' \begin{vmatrix} a & a'' \\ b & b'' \end{vmatrix} + c'' \begin{vmatrix} a & a' \\ b & b' \end{vmatrix} = c(a'b'' - a''b') - c'(ab'' - a''b) + c''(ab' - a'b).$$

• Attention !

(i) Ainsi le développement par rapport à une ligne quelconque d'un déterminant d'ordre n ramène au calcul de n déterminants d'ordre $n-1$. On comprend que ceci aboutit très vite (dès que n est un peu grand...) à des calculs laborieux, voire inextricables. D'où l'importance des méthodes de réduction que l'on va voir plus loin, et qui aboutissent dans la pratique au principe suivant: faire sur le déterminant des transformations faisant apparaître le maximum de zéros, et ne développer que lorsqu'on ne peut plus rien faire d'autre!

(ii) On veillera particulièrement à l'alternance des signes dans l'apparition des cofacteurs (ce signe étant $+$ si la somme de l'indice de ligne et de l'indice de colonne est paire, et $-$ si elle est impaire).

$$\begin{pmatrix} + & - \\ - & + \end{pmatrix} \qquad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \qquad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

(iii) On gardera bien à l'esprit que la notion de déterminant n'est définie que pour des matrices <u>carrées</u> !

(iv) Tout ce que l'on présente ici pour les matrices carrées à coefficients réels reste vrai sans aucun changement pour des matrices à coefficients <u>complexes</u>.

1.3 Opérations sur les lignes

On résume dans la proposition suivante les règles de calculs qui permettent, dans la pratique, le calcul effectif des déterminants. Les deux premières sont des conséquences de la définition donnée en 1.2. On pourra à titre d'exercice vérifier que les autres découlent simplement des deux premières.

PROPOSITION

1. *Si on échange deux lignes d'une matrice carrée, on multiplie son déterminant par* -1.
2. *Le déterminant est linéaire par rapport à chaque ligne, ce qui signifie que*
   a) *si l'on multiplie par* $\lambda \in \mathbb{R}$ *une ligne de la matrice* A, *on multiplie son déterminant par* λ;
   b) *si* A, A', A'' *sont trois matrices carrées d'ordre* n *telles que, pour un certain* $1 \le i \le n$, *la* i*-ième ligne de* A *soit la somme de la* i*-ième ligne de* A' *et de la* i*-ième ligne de* A'', *et telles que toutes les autres lignes soient les mêmes pour les trois matrices, alors* $\det A = \det A' + \det A''$.
3. *Si une matrice a une ligne entière de zéros, alors son déterminant est nul.*
4. *Si l'on multiplie par un réel* λ *une matrice carrée d'ordre* n, *on multiplie son déterminant par* λ^n.
5. *Si une matrice carrée a deux lignes identiques, alors son déterminant est nul.*
6. *Si dans une matrice carrée, on ajoute à une ligne une combinaison linéaire des autres lignes, on ne change pas la valeur de son déterminant.*
7. *Si dans une matrice carrée, une ligne est combinaison linéaire des autres lignes, alors son déterminant est nul.*

1.4 Opérations sur les colonnes

• TRANSPOSITION. Pour toute matrice carrée $A = (a_{i,j})_{1\le i,j\le n}$ d'ordre n, on appelle *transposée de* A, et on note ${}^tA$, la matrice carrée obtenue en échangeant les lignes et les colonnes de A, c'est-à-dire la matrice ${}^tA = (a'_{i,j})_{1\le i,j\le n}$ dont le terme général est $a'_{i,j} = a_{j,i}$.

Par exemple: si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors ${}^tA = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, et si $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$, alors ${}^tA = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & k \end{pmatrix}$.

• THÉORÈME. *Toute matrice carrée* $A \in \mathcal{M}_n(\mathbb{R})$ *a le même déterminant que sa transposée:* $\boxed{\det({}^tA) = \det A.}$

Parmi les conséquences pratiques de ce théorème, dont nous ne rappelons pas ici la preuve, figure le fait que tout ce que l'on peut dire sur les lignes d'un déterminant reste vrai pour les colonnes. En particulier on a:

• COROLLAIRE 1. *Toutes les règles de calcul de 1.3 restent vraies en remplaçant "ligne" par "colonne".*

• COROLLAIRE 2. *Soit* $A = (a_{i,j})_{1\le i,j\le n}$ *une matrice carrée d'ordre* n. *Pour tout* $1 \le j \le n$, *on a la formule suivante, dite développement de* A *par rapport à la* j*-ième colonne:*

$$\det A = \sum_{k=1}^{n} (-1)^{k+j} a_{k,j} \det A_{k,j},$$

où $(-1)^{k+j} \det A_{k,j}$ *désigne toujours le cofacteur associé au coefficient* $a_{k,j}$.

1.5 Exemples de calculs

On commence par un résultat général sur les matrices triangulaires, qui s'applique donc aussi en particulier aux matrices diagonales, et qui est très utile dans la pratique.

• EXEMPLE FONDAMENTAL: DÉTERMINANT D'UNE MATRICE TRIANGULAIRE.

Proposition. Le déterminant d'une matrice triangulaire est égal au produit de ses coefficients diagonaux.

Preuve: Dans le cas d'une matrice triangulaire supérieure, d'ordre n, on développe le déterminant par rapport à la première colonne, et on réitère n fois:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n-1} & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & a_{n,n} \end{vmatrix} = a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} & \cdots & a_{2,n-1} & a_{2,n} \\ 0 & a_{3,3} & \cdots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{n,n} \end{vmatrix} = \cdots = a_{1,1}a_{2,2}\cdots a_{n,n}.$$

Le cas d'une matrice triangulaire inférieure s'en déduit à l'aide des résultats de 1.4. □

• Exemples de calcul par triangularisation.

Principe. En appliquant les règles de 1.3 ou 1.4, on cherche à faire apparaître le maximum de zéros, et à se ramener à une forme triangulaire pour développer à l'aide de la proposition ci-dessus.

Exemple: calculer dans $\mathbb{R}$ les déterminants: $D_1 = \begin{vmatrix} a&0&3&5\\0&0&b&2\\1&c&2&3\\0&0&0&d \end{vmatrix}$ et $D_2 = \begin{vmatrix} a&a&a&a\\a&b&b&b\\a&b&c&c\\a&b&c&d \end{vmatrix}$.

Pour D_1, on échange les lignes l_1 et l_3, puis les lignes l_2 et l_3, puis les colonnes c_1 et c_2:

$$D_1 = \begin{vmatrix} a&0&3&5\\0&0&b&2\\1&c&2&3\\0&0&0&d \end{vmatrix} = -\begin{vmatrix} 1&c&2&3\\0&0&b&2\\a&0&3&5\\0&0&0&d \end{vmatrix} = \begin{vmatrix} 1&c&2&3\\a&0&3&5\\0&0&b&2\\0&0&0&d \end{vmatrix} = -\begin{vmatrix} c&1&2&3\\0&a&3&5\\0&0&b&2\\0&0&0&d \end{vmatrix} = -abcd.$$

Pour D_2, on remplace d'abord c_2 par $c_2 - c_1$, c_3 par $c_3 - c_1$ et c_4 par $c_4 - c_1$; ensuite on remplace c_3 par $c_3 - c_2$ et c_4 par $c_4 - c_2$; enfin, on remplace c_4 par $c_4 - c_3$.

$$D_2 = \begin{vmatrix} a&0&0&0\\a&b-a&b-a&b-a\\a&b-a&c-a&c-a\\a&b-a&c-a&d-a \end{vmatrix} = \begin{vmatrix} a&0&0&0\\a&b-a&0&0\\a&b-a&c-b&c-b\\a&b-a&c-b&d-b \end{vmatrix} = \begin{vmatrix} a&0&0&0\\a&b-a&0&0\\a&b-a&c-b&0\\a&b-a&c-b&d-c \end{vmatrix} = a(b-a)(c-b)(d-c).$$

• Exemple de calcul utilisant la linéarité par rapport aux colonnes

Calculons dans $\mathbb{C}$ le déterminant: $D_3 = \begin{vmatrix} a+i\alpha & \alpha+ia & a+\alpha\\ b+i\beta & \beta+ib & b+\beta\\ c+i\gamma & \gamma+ic & c+\gamma \end{vmatrix}$. On applique la linéarité par rapport à chaque colonne:

$$D_3 = \begin{vmatrix} a & \alpha+ia & a+\alpha\\ b & \beta+ib & b+\beta\\ c & \gamma+ic & c+\gamma \end{vmatrix} + i\begin{vmatrix} \alpha & \alpha+ia & a+\alpha\\ \beta & \beta+ib & b+\beta\\ \gamma & \gamma+ic & c+\gamma \end{vmatrix} = \begin{vmatrix} a & \alpha & a+\alpha\\ b & \beta & b+\beta\\ c & \gamma & c+\gamma \end{vmatrix} + i\begin{vmatrix} a & a & a+\alpha\\ b & b & b+\beta\\ c & c & c+\gamma \end{vmatrix} + i\begin{vmatrix} \alpha & \alpha & a+\alpha\\ \beta & \beta & b+\beta\\ \gamma & \gamma & c+\gamma \end{vmatrix} + i^2\begin{vmatrix} \alpha & a & a+\alpha\\ \beta & b & b+\beta\\ \gamma & c & c+\gamma \end{vmatrix}.$$

Le 2ième et le 3ième déterminants sont nuls car ils ont deux colonnes identiques; le 1er et le 4ième déterminants sont nuls car la colonne c_3 est somme des colonnes c_1 et c_2. On conclut $D_3 = 0$.

1.6 Multiplicativité.

La propriété de multiplicativité démontrée au 1.1 dans le cas particulier des matrices carrées d'ordre 2 peut en fait être prouvée pour des matrices carrées d'ordre n quelconque.

Théorème.

(i) *Pour toutes matrices carrées $A, B \in \mathcal{M}_n(\mathbb{R})$, on a:* $\boxed{\det(AB) = \det A \times \det B}$.

(ii) *Une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si et seulement si* $\det A \neq 0$, *et alors* $\boxed{\det(A^{-1}) = \dfrac{1}{\det A}}$.

Corollaire. *Deux matrices carrées semblables ont le même déterminant.*

En effet: A et B semblables signifie que: $A = PBP^{-1}$ avec P carrée d'ordre n inversible. On a alors: $\det A = \det P \times \det B \times \det(P^{-1}) = \det P \times \det B \times (\det P)^{-1} = \det P \times (\det P)^{-1} \times \det B = \det B$.

2. Applications des déterminants.

2.1 Calcul de l'inverse d'une matrice carrée inversible.

• Définition. Pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$, on appelle *comatrice* de A, notée $\operatorname{Com} A$ la matrice:

$$\operatorname{Com} A = (\Delta_{i,j})_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{R})$$

dont le coefficient général est, sur la i-ième ligne et la j-ième colonne, le cofacteur:

$$\Delta_{i,j} = (-1)^{i+j} \det A_{i,j}$$

de la matrice A tel qu'on l'a défini plus haut en 1.2

• Proposition. *Soit A une matrice carrée d'ordre n inversible; alors:* $\boxed{A^{-1} = \dfrac{1}{\det A}\ {}^t(\operatorname{Com} A).}$

• Exemples.

Soit $A = \begin{pmatrix} a&b\\c&d \end{pmatrix}$, donc $\operatorname{Com} A = \begin{pmatrix} d&-c\\-b&a \end{pmatrix}$, donc $A^{-1} = \frac{1}{\det A}\ {}^t(\operatorname{Com} A) = \frac{1}{ad-bc}\begin{pmatrix} d&-b\\-c&a \end{pmatrix}$; on retrouve 1.1.

Soit $A = \begin{pmatrix} a&b&c\\d&e&f\\g&h&k \end{pmatrix}$, donc $\operatorname{Com} A = \begin{pmatrix} +(ek-fh) & -(dk-fg) & +(dh-eg)\\ -(bk-ch) & +(ak-cg) & -(ah-bg)\\ +(bf-ce) & -(af-cd) & +(ae-bd) \end{pmatrix}$,

$$\text{donc } A^{-1} = \frac{1}{\det A}\ {}^t(\operatorname{Com} A) = \frac{1}{\det A}\begin{pmatrix} ek-fh & ch-bk & bf-ce\\ fg-dk & ak-cg & cd-af\\ dh-eg & bg-ah & ae-bd \end{pmatrix}.$$

Choisissons par exemple $A = \left(\begin{smallmatrix} 1 & 4 & 2 \\ 0 & 3 & 1 \\ 2 & 9 & 4 \end{smallmatrix}\right)$; on a: $\det A = 1 \times (12 - 9) - 4 \times (0 - 2) + 2 \times (0 - 6) = -1$.

On calcule $\mathrm{Com}\, A = \left(\begin{smallmatrix} 3 & 2 & -6 \\ 2 & 0 & -1 \\ -2 & -1 & 3 \end{smallmatrix}\right)$; on conclut $A^{-1} = -\,{}^t(\mathrm{Com}\, A) = \left(\begin{smallmatrix} -3 & -2 & 2 \\ -2 & 0 & 1 \\ 6 & 1 & -3 \end{smallmatrix}\right)$.

Remarque. La proposition ci-dessus a un intérêt théorique, mais, dans la pratique, elle conduit souvent à des calculs nombreux et complexes (par exemple calculer l'inverse d'une matrice 4×4 revient à calculer 16 déterminants 3×3). C'est pourquoi, sur des exemples concrets de matrices carrées à coefficients numériques explicites, il est souvent beaucoup plus rapide d'utiliser la méthode vue précédemment d'inversion formelle d'un système d'équations.

2.2 Déterminant d'un endomorphisme.

• Lemme et définition. *Soit f un endomorphisme de $\mathbb{R}^n$. Le déterminant de la matrice de f par rapport à une base $\mathcal{B}$ de $\mathbb{R}^n$ garde la même valeur quelle que soit la base choisie. Cette valeur commune du déterminant de la matrice de f par rapport à une base quelconque de E est appelée le déterminant de f, noté* $\det f$.

En effet, prenons $\mathcal{B}$ et $\mathcal{B}'$ deux bases de $\mathbb{R}^n$, et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. D'après la seconde formule de changement de base, en notant $A = M_{\mathcal{B}}(f)$ et $A' = M_{\mathcal{B}'}(f)$, on a $A = PA'P^{-1}$. Donc $\det A = \det A'$ d'après le corollaire du 1.6 ci-dessus.

• Conséquence. *Pour tout endomorphisme f de $\mathbb{R}^n$, on a:* $\boxed{\mathrm{Ker}\, f = \{\vec{0}\} \;\Leftrightarrow \det f \neq 0.}$

En effet, on a vu que le noyau de f est réduit au vecteur nul si et seulement si la matrice de f par rapport à une base quelconque de $\mathbb{R}^n$ est inversible. On applique alors le point (ii) du théorème de 1.6.

2.3 Caractérisation des bases.

• Proposition. *Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de $\mathbb{R}^n$. Soit $\mathcal{B}' = (e'_1, e'_2, \ldots, e'_n)$ une famille de n vecteurs de $\mathbb{R}^n$. Soit P la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$. Alors:*

$$\boxed{\mathcal{B}' \text{ est une base de } \mathbb{R}^n} \Longleftrightarrow \boxed{\det P \neq 0.}$$

En effet, on a vu que au début de l'étude des changements de bases que $\mathcal{B}'$ est une base de $\mathbb{R}^n$ si et seulement si P est inversible.

2.4 Formules de Cramer pour les sytèmes d'équations linéaires.

• Données. Considérons un système linéaire (S) de n équations à n inconnues dans $\mathbb{R}$:

$$(S):\ \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \ldots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n \end{cases}, \quad \text{que l'on note matriciellement:} \quad A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

en notant $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$ la matrice carrée des coefficients.
Un n-uplet $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ qui satisfait chacune des n équations de (S) s'appelle une *solution* de (S).
Un système (S) tel que sa matrice A soit inversible s'appelle un *système de Cramer*.
On peut démontrer sur les systèmes de Cramer le théorème suivant.

• Théorème. *Avec les données et notations ci-dessus:*

$$\boxed{\text{(S) admet une unique solution}} \Longleftrightarrow \boxed{\text{(S) est un système de Cramer}} \Longleftrightarrow \boxed{\det A \neq 0}$$

De plus, cette unique solution $(x_1, \ldots, x_n)$ est donnée par les formules de Cramer:

$$\boxed{x_i = \frac{\det(A|B)_i}{\det A} \qquad \textit{pour tout } 1 \leq i \leq n,}$$

où $(A|B)_i$ désigne la matrice carrée d'ordre n obtenue en remplaçant dans A la i-ième colonne par la colonne du second membre de (S).

• Exemple des systèmes 2×2.

On considère un système $(S) \begin{cases} ax+by=e \\ cx+dy=f \end{cases}$, avec donc $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, et $\det A = ad - bc$

(S) a une unique solution si et seulement si $ad - bc \neq 0$, et cette solution (x_0, y_0) est donnée par:

$$x_0 = \frac{1}{\det A} \begin{vmatrix} e & b \\ f & d \end{vmatrix} = \frac{ed - bf}{ad - bc} \quad \text{et} \quad y_0 = \frac{1}{\det A} \begin{vmatrix} a & e \\ c & f \end{vmatrix} = \frac{af - ce}{ad - bc}.$$

• un exemple de système 3×3.

On considère le système: $(S) : \begin{cases} (2a+1)x - ay + (a+1)z = a-1 & (1) \\ (a-2)x + (a-1)y + (a-2)z = a & (2) \\ (2a-1)x + (a-1)y + (2a-1)z = a & (3) \end{cases}$, $\quad$ ($a \in \mathbb{R}$ paramètre fixé).

Pour la matrice A de (S), on calcule $\det A = a(a-1)(a+1)$. D'où quatre cas:

Premier cas: si $a \neq 0, a \neq 1, a \neq -1$; alors $\det A \neq 0$ donc (S) est de Cramer et son unique solution (x_0, y_0, z_0) dans $\mathbb{R}$ est donnée par:

$$x_0 = \frac{1}{\det A} \begin{vmatrix} a-1 & -a & a+1 \\ a & a-1 & a-2 \\ a & a-1 & 2a-1 \end{vmatrix}, \quad y_0 = \frac{1}{\det A} \begin{vmatrix} 2a+1 & a-1 & a+1 \\ a-2 & a & a-2 \\ 2a-1 & a & 2a-1 \end{vmatrix}, \quad z_0 = \frac{1}{\det A} \begin{vmatrix} 2a+1 & -a & a-1 \\ a-2 & a-1 & a \\ 2a-1 & a-1 & a \end{vmatrix},$$

que l'on simplifie après calculs en: $x_0 = \frac{2a^2-2a+1}{a(a-1)}$, $y_0 = \frac{a}{a-1}$, $z_0 = -\frac{2a^2-2a+1}{a(a-1)}$.

Deuxième cas: si $a = 0$, on réécrit le système: $\begin{cases} x+z = -1 \\ -2x-y-2z = 0 \\ -x-y-z = 0 \end{cases}$;

les équations (1), (2) et (3) sont incompatibles; (S) n'a pas de solution.

Troisième cas: si $a = 1$, on réécrit le système: $\begin{cases} 3x-y+2z = 0 \\ -x-z = 1 \\ x+z = 1 \end{cases}$;

les équations (2) et (3) sont incompatibles; (S) n'a pas de solution.

Quatrième cas: si $a = -1$, les équations (2) et (3) sont équivalentes; (S) devient: $\begin{cases} -x+y = -2 \\ -3x-2y-3z = -1 \end{cases}$;

l'ensemble des solutions de (S) est $\{(-\frac{3}{5}z + 1, -\frac{3}{5}z - 1, z)\,;\, z \in \mathbb{R}\}$ dépend d'un paramètre.

Diagonalisation

1. Valeurs propres et polynôme caractéristique.

1.1 Remarques préliminaires et motivations

On a déjà vu précédemment que les calculs sur les matrices triangulaires, et en particulier sur les matrices diagonales, sont particulièrement simples.

RAPPELONS QUE:

1. le produit de deux matrices triangulaires est une matrice triangulaire; en particulier le produit de deux matrices diagonales est une matrice diagonale, et ses coefficients diagonaux s'obtiennent en multipliant terme à terme les coefficient diagonaux de chacune des matrices:
$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_n \end{pmatrix} = \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_n \end{pmatrix} \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \dots & 0 \\ 0 & a_2b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_nb_n \end{pmatrix}$$
2. le déterminant d'une matrice triangulaire (et donc en particulier d'une matrice diagonale) est égal au produit de ses coefficients diagonaux; il en résulte qu'une matrice triangulaire (et donc en particulier une matrice diagonale) est inversible si et seulement si tous ses coefficients diagonaux sont non-nuls.

On comprend donc que, pour représenter un endomorphisme par sa matrice A dans une base $\mathcal{B}$, on a tout intérêt à choisir $\mathcal{B}$ telle que A soit diagonale, ou à défaut triangulaire.

Les applications de ce principe (en analyse et en géométrie en particulier) sont extrêmement importantes: nous en avons déjè détaillé certaines lors de l'étude des changements de bases.

Encore faut-il que cela soit possible, c'est-à-dire qu'une telle base existe. C'est ce problème (pour un endomorphisme f donné, déterminer une base dans laquelle la matrice de f soit la plus simple possible (diagonale, ou à défaut triangulaire) que l'on va aborder maintenant.

1.2 Valeurs propres, vecteurs propres, sous-espaces propres

• DÉFINITIONS. Soit f un endomorphisme de $\mathbb{R}^n$.

1. Une *valeur propre* de f dans $\mathbb{R}$ est un réel $\lambda \in \mathbb{R}$ pour lequel il existe un vecteur $v \in \mathbb{R}^n$ non-nul tel que $f(v) = \lambda.v$. On utilisera l'abréviation v.p. pour valeur propre.
2. Si λ est une v.p. de f, on appelle *vecteur propre* de f associé à λ tout vecteur $v \in \mathbb{R}^n$ non-nul tel que $f(v) = \lambda.v$.
3. Pour toute v.p. λ de f, on appelle *sous-espace propre* de f associé à la v.p. λ le sous-espace vectoriel $E_\lambda = \mathrm{Ker}(f - \lambda.\mathrm{Id})$ de $\mathbb{R}^n$.

• REMARQUE 1. Si v est un vecteur propre associé à une valeur propre λ, alors v ne peut pas être vecteur propre pour une autre valeur propre $\mu \neq \lambda$. En effet, $f(v) = \lambda.v = \mu.v$ implique $(\lambda - \mu).v = \vec{0}$, ce qui, puisque $v \neq \vec{0}$, conduit nécessairement à $\lambda = \mu$.

• REMARQUE 2. Soit $\lambda \in \mathbb{R}$. considérons $E_\lambda = \mathrm{Ker}(f - \lambda.\mathrm{Id})$.
Par définition, E_λ est l'ensemble des vecteurs $v \in \mathbb{R}^n$ tels que $f(v) = \lambda.v$. Ce ss-e.v. E_λ est donc l'ensemble des vecteurs propres associés à λ (qui par définition sont tous non-nuls), plus le vecteur nul (qui n'est pas un vecteur propre, mais qui appartient quand même au sous-espace propre). Dire que λ est une v.p. de f signifie que le ss-e.v. E_λ n'est pas réduit à $\{\vec{0}\}$. Par définition, on a alors: $1 \leq \dim E_\lambda \leq n$.

1.3 Polynôme caractéristique

• DÉFINITIONS. Soit f un endomorphisme de $\mathbb{R}^n$. On appelle *polynôme caractéristique de f* le polynôme:
$$\boxed{P_f(x) = \det(f - x.\mathrm{Id})}.$$
Soit A une matrice carrée dans $\mathcal{M}_n(\mathbb{R})$. On appelle *polynôme caractéristique* de la matrice A le polynôme :
$$\boxed{P_A(x) = \det(A - x.I_n)}.$$

Remarque. Soit f un endomorphisme de $\mathbb{R}^n$. 0n choisit une base quelconque $\mathcal{B}$ de $\mathbb{R}^n$. En notant A la matrice de f par rapport à la base $\mathcal{B}$, on sait que la matrice par rapport à $\mathcal{B}$ de l'endomorphisme $(f - x.\mathrm{Id})$ est $(A - x.I_n)$. On a alors $\det(f - x.\mathrm{Id}) = \det(A - x.I_n)$, et ceci quelle que soit la base $\mathcal{B}$ choisie. Donc $P_f(x) = P_A(x)$.

Réciproquement, toute matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ peut être considérée comme la matrice d'un certain endomorphisme f de $\mathbb{R}^n$ par rapport à la base canonique, et on alors $P_A(x) = P_f(x)$.

Le théorème suivant est un résultat fondamental permettant le calcul effectif des v.p. d'un endomorphisme.

• Théorème. *Soit f un endomorphisme de $\mathbb{R}^n$. Les valeurs propres de f dans $\mathbb{R}$ sont exactement les zéros dans $\mathbb{R}$ de son polynôme caractéristique $P_f(x)$.*

Preuve. Soit f un endomorphisme de $\mathbb{R}^n$. Soit A sa matrice dans une base $\mathcal{B}$ quelconque de $\mathbb{R}^n$. On a vu à la remarque 2 de 1.2 qu'un réel λ est une v.p. de f si et seulement si le noyau de l'endomorphisme $(f - \lambda.\mathrm{Id})$ n'est pas réduit à $\{\vec{0}\}$. Comme on l'a vu à la fin de l'étude des matrices, ceci est équivalent à dire que sa matrice $A - \lambda.I_n$ n'est pas inversible, c'est-à-dire que son déterminant est nul.

• Conséquences pratiques.

1. Ainsi, pour déterminer les v.p. de f, on fixe une base quelconque $\mathcal{B}$, on introduit la matrice A de f par rapport à $\mathcal{B}$, et on détermine par un calcul de déterminant le polynôme $P_f(x) = P_A(x) = \det(A - x.I_n)$. Les zéros de ce polynômes sont les v.p. de l'endomorphisme f, que l'on appelle aussi les v.p. de la matrice A.

2. De plus, comme il est clair que $P_f(x)$ est de degré n, il admet forcément au plus n zéros dans $\mathbb{R}$. Donc f admet au plus n valeurs propres dans $\mathbb{R}$.

3. *Attention !* Le polynôme $P_f(x)$ n'a pas forcément des zéros dans $\mathbb{R}$ (par exemple s'il est de degré 2 et de discriminant < 0). Ce qui signifie que l'endomorphisme f ou la matrice A *n'a pas forcément de valeurs propres dans* $\mathbb{R}$ *!*

4. En revanche, le polynôme à coefficients réels $P_f(x)$ a forcément des zéros dans $\mathbb{C}$, et donc f ou A ont forcément des v.p. dans $\mathbb{C}$. On verra sur des exemples qu'il est parfois très utile de considérer les v.p. complexes d'une matrice même si celle-ci est au départ une matrice à coefficients réels.

5. A fortiori, tout ce que l'on détaille ici sur les endomorphismes de $\mathbb{R}^n$ et les matrices de $M_n(\mathbb{R})$ s'applique aussi aux endomorphismes de $\mathbb{C}^n$ et aux matrices de $M_n(\mathbb{C})$

• Quelques précisions sur le polynôme caractéristique.

Considérons l'exemple élémentaire d'une matrice $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(\mathbb{R})$. Son polynôme caractéristique est:

$$P_A x) = \left|\begin{smallmatrix} a-x & b \\ c & d-x \end{smallmatrix}\right| = x^2 - x(a+d) + (ad - bc).$$

On reconnaît en $(a+d)$ la trace de A et en $(ad - bc)$ son déterminant. C'est le cas le plus simple du résultat général suivant, que nous citons ci-dessous sans démonstration.

Proposition. Soit $A \in M_n(\mathbb{R})$ une matrice carrée d'ordre n. Son polynôme caractéristique $P_A(x) = \det(A - x.I_n)$ est un polynôme de degré n, à coefficients dans $\mathbb{R}$, tel que:

le coefficient de x^n est $(-1)^n$; le coefficient de x^{n-1} est $(-1)^{n-1}\,\mathrm{tr}\,A$; le coefficient constant est $\det A$.

En d'autres termes: $P_A(x) = (-1)^n x^n + (-1)^{n-1}(\mathrm{tr}\,A)x^{n-1} + \cdots + \det A$.

1.5 Multiplicités d'une valeur propre

• Un rappel sur les zéros d'un polynôme. Soit $P(x)$ un polynôme à coefficients réels de degré n. Soit α un zéro de $P(x)$ dans $\mathbb{R}$, c'est-à-dire un nombre $\alpha \in \mathbb{R}$ tel que $P(\alpha) = 0$. On peut alors mettre $(x - \alpha)$ en facteur dans $P(x)$. Il existe donc un unique entier $1 \leq q \leq n$, appelé la *multiplicité* de α, tel que:

$$P(x) = (x - \alpha)^q Q(x), \quad \text{avec } Q(x) \text{ polynôme de degré } n - q \text{ vérifiant } Q(\alpha) \neq 0.$$

En d'autres termes, cela signifie que l'on ne peut plus mettre $(x - \alpha)$ en facteur dans $Q(x)$, ou encore que q est l'exposant de la puissance maximale de $(x - \alpha)$ qu'il est possible de mettre en facteur dans $P(x)$.

• Définition. Soit λ une v.p. d'un endomorphisme f de $\mathbb{R}^n$ (ou d'une matrice A carrée d'ordre n à coefficients dans $\mathbb{R}$). On appelle *multiplicité algébrique* de λ sa multiplicité en tant que zéro du polynôme caractéristique de f (ou de A).

• Proposition. *Avec les données ci-dessus, et en notant E_λ le sous-espace propre associé à la v.p. λ, on a:*

$$\boxed{1 \;\leq\; \dim E_\lambda \;\leq\; (\text{multiplicité algébrique de } \lambda) \;\leq\; n.}$$

• TERMINOLOGIE. La dimension du sous-espace propre E_λ est parfois appelée la *multiplicité géométrique* de la v.p. λ. Elle est donc toujours $\leq$ à sa multiplicité algébrique.

Une v.p. est dite *simple* si sa multiplicité algébrique est 1 (sa multiplicité géométrique est alors 1).
Une v.p. est dite *double* si sa multiplicité algébrique est 2 (sa multiplicité géométrique est alors 1 ou 2).
Une v.p. est dite *triple* si sa multiplicité algébrique est 3 (sa multiplicité géométrique est alors 1 ou 2 ou 3).

1.6 Quelques exemples de calculs

• PREMIER EXEMPLE. Soit $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$.

On calcule: $P_A(x) = \begin{vmatrix} 1-x & 2 & 0 \\ 0 & 1-x & 1 \\ 1 & 1 & -1-x \end{vmatrix} = (1-x)(x^2-2)+2 = -x^3+x^2+2x = -x(x-2)(x+1)$.

Donc A admet trois valeurs propres simples qui sont 0, 2 et -1. Déterminons les sous-espaces propres correspondants.

$$(x,y,z) \in E_2 \Leftrightarrow \begin{pmatrix} 1-2 & 2 & 0 \\ 0 & 1-2 & 1 \\ 1 & 1 & -1-2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -x + 2y = 0 \\ -y + z = 0 \\ x + y - 3z = 0 \end{cases}.$$

Donc $E_2 = \{(2y, y, y) \in \mathbb{R}^3 \,;\, y \in \mathbb{R}\}$ est la droite de base (u_2), où $u_2 = (2,1,1)$.

$$(x,y,z) \in E_{-1} \Leftrightarrow \begin{pmatrix} 1-(-1) & 2 & 0 \\ 0 & 1-(-1) & 1 \\ 1 & 1 & -1-(-1) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Leftrightarrow \begin{cases} 2x + 2y = 0 \\ 2y + z = 0 \\ x + y = 0 \end{cases}.$$

Donc $E_{-1} = \{(-y, y, -2y) \in \mathbb{R}^3 \,;\, y \in \mathbb{R}\}$ est la droite de base (u_{-1}), où $u_{-1} = (1,-1,2)$.

$$(x,y,z) \in E_0 \Leftrightarrow \begin{pmatrix} 1-0 & 2 & 0 \\ 0 & 1-0 & 1 \\ 1 & 1 & -1-0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Leftrightarrow \begin{cases} x + 2y = 0 \\ y + z = 0 \\ x + y - z = 0 \end{cases}.$$

Donc $E_0 = \{(-2y, y, -y) \in \mathbb{R}^3 \,;\, y \in \mathbb{R}\}$ est la droite de base (u_0), où $u_0 = (2,-1,1)$.

• SECOND EXEMPLE. Soit $A = \begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}$.

On calcule: $P_A(x) = \begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ -2 & 4 & 5-x \end{vmatrix} = \begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ 0 & 9-x & 9-x \end{vmatrix} = (9-x) \begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ 0 & 1 & 1 \end{vmatrix} = (9-x) \begin{vmatrix} 8-x & 2 & -4 \\ 2 & 5-x & x-1 \\ 0 & 1 & 0 \end{vmatrix}$

$= (9-x)(-1) \begin{vmatrix} 8-x & -4 \\ 2 & x-1 \end{vmatrix} = (9-x)(x^2-9x) = -x(x-9)^2$.

Donc 0 est v.p. simple et 9 est v.p. double. On sait qu'alors E_0 est forcément une droite, mais E_9 peut être soit une droite, soit un plan. Déterminons-les.

$$(x,y,z) \in E_0 \Leftrightarrow \begin{pmatrix} 8-0 & 2 & -2 \\ 2 & 5-0 & 4 \\ -2 & 4 & 5-0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 8x + 2y - 2z = 0 \\ 2x + 5y + 4z = 0 \\ -2x + 4y + 5z = 0 \end{cases}.$$

Après calculs, on trouve que E_0 est la droite de base (u_0), où $u_0 = (1,-2,2)$.

$$(x,y,z) \in E_9 \Leftrightarrow \begin{pmatrix} 8-9 & 2 & -2 \\ 2 & 5-9 & 4 \\ -2 & 4 & 5-9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -x + 2y - 2z = 0 \\ 2x - 4y + 4z = 0 \\ -2x + 4y - 4z = 0 \end{cases}.$$

Donc E_9 est le plan d'équation $x - 2y + 2z = 0$, dont une base est (u_9, v_9), où $u_9 = (2,1,0)$ et $v_9 = (2,0,-1)$. Donc, sur cet exemple, la multiplicité de la v.p. double 9 est égale à la dimension du sous-espace propre correspondant.

• TROISIÈME EXEMPLE. Soit $A = \begin{pmatrix} 3 & 2 & 4 \\ -1 & 3 & -1 \\ -2 & -1 & -3 \end{pmatrix}$.

On calcule: $P_A(x) = \begin{vmatrix} 3-x & 2 & 4 \\ -1 & 3-x & -1 \\ -2 & -1 & -3-x \end{vmatrix} = \begin{vmatrix} -1-x & 2 & 4 \\ 0 & 3-x & -1 \\ x+1 & -1 & -3-x \end{vmatrix} = (x+1) \begin{vmatrix} -1 & 2 & 4 \\ 0 & 3-x & -1 \\ 1 & -1 & -3-x \end{vmatrix} = (x+1) \begin{vmatrix} -1 & 2 & 4 \\ 0 & 3-x & -1 \\ 0 & 1 & 1-x \end{vmatrix}$

$= -(x+1)(x^2-4x+4) = -(x+1)(x-2)^2$.

Donc -1 est v.p. simple et 2 est v.p. double. On sait qu'alors E_{-1} est forcément une droite, mais E_2 peut être soit une droite, soit un plan. Déterminons-les.

$$(x,y,z) \in E_{-1} \Leftrightarrow \begin{pmatrix} 3-(-1) & 2 & 4 \\ -1 & 3-(-1) & -1 \\ -2 & -1 & -3-(-1) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 4x + 2y + 4z = 0 \\ -x + 4y - z = 0 \\ -2x - y - 2z = 0 \end{cases}.$$

Après calculs, on trouve que E_{-1} est la droite de base (u_{-1}), où $u_{-1} = (1,0,-1)$.

$$(x,y,z)\in E_2 \Leftrightarrow \begin{pmatrix} 3-2 & 2 & 4 \\ -1 & 3-2 & -1 \\ -2 & -1 & -3-2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Leftrightarrow \begin{cases} x + 2y + 4z = 0 \\ -x + y - z = 0 \\ -2x - y - 5z = 0 \end{cases}.$$

Après calculs, on trouve que E_2 est la droite de base (u_2), où $u_2 = (2,1,-1)$. Donc, sur cet exemple, la multiplicité de la v.p. double 2 est strictement supérieure à la dimension du sous-espace propre correspondant.

• Quatrième exemple. Soit $A = \begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 1 & -1 \\ 5 & -3 & -2 & 5 \\ 4 & -2 & -2 & 3 \end{pmatrix}$. Après calculs, $P_A(x) = (x+1)(x-1)^3$.

-1 est v.p. simple. On sait qu'alors E_{-1} est forcément une droite. Après calculs, on trouve que E_{-1} est la droite de base (u_{-1}), où $u_{-1} = (1,0,0,-1)$.

1 est v.p. triple. $(x,y,z)\in E_1 \Leftrightarrow \begin{cases} -2x + y + z = 0 \\ -x + y + z - t = 0 \\ 5x - 3y - 3z + 5t = 0 \\ 4x - 2y - 2z + 2t = 0 \end{cases} \Leftrightarrow \cdots \Leftrightarrow \begin{cases} x = t = 0 \\ y + z = 0 \end{cases}$

Donc E_1 est la droite de base (u_1), où $u_1 = (0,1,-1,0)$. Sur cet exemple, le sous-espace propre associé à la v.p. triple 1 est seulement de dimension 1.

• Cinquième exemple. Soit $A = \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}$. Après calculs, $P_A(x) = x^2 + x + 1$. Donc A n'admet pas de v.p. dans $\mathbb{R}$. En revanche, dans $\mathbb{C}$, elle admet deux v.p. simples distinctes qui sont j et j^2. On vérifie aisément (par la même méthode que sur les exemples précédents) que E_j est la droite de base (u_j), où $u_j = (1, 1-j)$ et E_{j^2} est la droite de base (u_{j^2}), où $u_{j^2} = (1, 1-j^2)$.

2. Diagonalisation.

2.1 Notion de matrice ou d'endomorphisme diagonalisable.

• Définition. Un endomorphisme f de $\mathbb{R}^n$ est dit *diagonalisable* (sur $\mathbb{R}$) lorsqu'il existe une base $\mathcal{C}$ de $\mathbb{R}^n$ telle que la matrice de f dans la base $\mathcal{C}$ soit une matrice diagonale D. Une matrice carrée d'ordre n est dite *diagonalisable* (sur $\mathbb{R}$) lorsqu'elle est semblable à une matrice diagonale.

Si A désigne la matrice de f dans une base quelconque $\mathcal{B}$ de $\mathbb{R}^n$, il résulte de la seconde formule de changement de base que:

(l'endomorphisme f est diagonalisable) $\Leftrightarrow$ (la matrice A est diagonalisable).

Dans ce cas, on a: $\boxed{A = PDP^{-1}, \quad \text{avec } A = M_{\mathcal{B}}(f),\ D = M_{\mathcal{C}}(f) \text{ diagonale, et } P = \mathrm{Pass}_{\mathcal{B}\to\mathcal{C}}.}$

1.2 Conditions de diagonalisabilité.

• Théorème fondamental. *Soit f un endomorphisme de $\mathbb{R}^n$. Les conditions suivantes sont équivalentes:*

(i) *f est diagonalisable (sur $\mathbb{R}$);*

(ii) *il existe une base $\mathcal{C}$ de $\mathbb{R}^n$ qui est formée de vecteurs propres de f ;*

(iii) *$\mathbb{R}^n$ est la somme directe des sous-espaces propres de f ;*

(iv) *le polynôme caractéristique $P_f(x)$ admet n zéros réels (comptés avec leur multiplicité), et la multiplicité algébrique q_i de chaque valeur propre λ_i de f est égale à la dimension du sous-espace propre E_{λ_i};*

$$D = M_{\mathcal{C}}(f) = \begin{pmatrix} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{bmatrix}}_{q_1} & & & \\ & \underbrace{\begin{bmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_2 \end{bmatrix}}_{q_2} & & \\ & & \ddots \ \cdots & \\ & & & \underbrace{\begin{bmatrix} \lambda_s & & \\ & \ddots & \\ & & \lambda_s \end{bmatrix}}_{q_s} \end{pmatrix}$$

$\lambda_1, \lambda_2 \dots \lambda_s \in \mathbb{R}$ sont les v.p. distinctes de f ou A,

$P_f(x) = (-1)^n (x-\lambda_1)^{q_1}(x-\lambda_2)^{q_2}\cdots(x-\lambda_s)^{q_s}$,

$q_i = \dim E_{\lambda_i}$ pour toute v.p. λ_i,

$\mathbb{R}^n = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_s}$.

(tous les coefficients hors de la diagonale étant des zéros).

En appliquant tout ce qui précède au cas particulier où $n = s$, c'est-à-dire $q_i = 1$ pour tout $1 \leq i \leq n$, on obtient immédiatement le corollaire suivant:

• Corollaire. *Soit f un endomorphisme de $\mathbb{R}^n$. Si f admet n v.p. distinctes (chacune est nécessairement une v.p. simple), alors f est diagonalisable.*

Attention ! Ce corollaire ne donne qu'une condition suffisante pour qu'un endomorphisme soit diagonalisable; elle n'est nullement nécessaire (voir les exemples ci-dessous). Si un endomorphisme de $\mathbb{R}^n$ admet n valeur propres simples distinctes, alors il est diagonalisable, mais il peut très bien être diagonalisable en admettant des v.p. doubles (il faut alors que le sous-espace propre associé soit de dimension 2), ou triples (il faut alors que le sous-espace propre associé soit de dimension 2), ou plus....

2.3. Exemples. Reprenons, dans le même ordre, les exemples traités en 1.6

• Premier exemple. La matrice A considérée est diagonalisable, d'après le corollaire ci-dessus. Une base de vecteurs propres est $\mathcal{C} = (u_2, u_{-1}, u_0)$, donc:

$$A = PDP^{-1}, \text{ avec } P = \begin{pmatrix} 2 & 1 & 2 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

• Deuxième exemple. La matrice A considérée est diagonalisable, d'après le (iv) du théorème fondamental ci-dessus. Une base de vecteurs propres est $\mathcal{C} = (u_0, u_9, v_9)$, donc:

$$A = PDP^{-1}, \text{ avec } P = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

• Troisième exemple. La matrice A considérée n'est pas diagonalisable, car le sous-espace propre associé à la valeur propre double 2 est seulement de dimension 1.

• Quatrième exemple. La matrice A considérée n'est pas diagonalisable, car le sous-espace propre associé à la valeur propre triple 1 est seulement de dimension 1.

• Cinquième exemple. La matrice A considérée n'est pas diagonalisable sur $\mathbb{R}$ mais est diagonalisable sur $\mathbb{C}$ (car admet deux v.p. simples distinctes complexes). Une base de vecteurs propres est $\mathcal{C} = (u_j, u_{j^2})$, donc: $A = PDP^{-1}$, avec $P = \begin{pmatrix} 1 & 1 \\ 1-j & 1-j^2 \end{pmatrix}$ et $D = \begin{pmatrix} j & 0 \\ 0 & j^2 \end{pmatrix}$.

2.4 Application à la résolution des systèmes différentiels linéaires à coefficients constants.

On a déjà indiqué lors de l'étude des changements de bases comment la recherche d'une matrice diagonale éventuellement semblable à une matrice donnée permet de résoudre certains systèmes différentiels linéaires à coefficients constants.

• Exemple: cas diagonalisable avec v.p. simples.

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel:

$$(S) \begin{cases} x'(t) = -2x(t) - 2y(t) + 2z(t) \\ y'(t) = 3x(t) + 5y(t) \\ z'(t) = x(t) + 2y(t) + 3z(t) \end{cases}, \quad \text{dont la matrice est } A = \begin{pmatrix} -2 & -2 & 2 \\ 3 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

De la même façon qu'on l'a fait sur les exemples de 1.6, on détermine son polynôme caractéristique, ses v.p. , et les sous-espaces propres associés. Après calculs, on obtient:

$$P_A(x) = -(x+1)(x-2)(x-5).$$

Donc A admet trois v.p. distinctes -1, 2 et 5, pour lesquels (après calculs) les sous-espaces propres sont les droites dirigées respectivement par les vecteurs $e_1' = 2.e_1 - e_2$, $e_2' = e_1 - e_2 + e_3$ et $e_3' = e_2 + e_3$. La matrice A est donc diagonalisable, et le changement de bases est donné par:

$$A = PDP^{-1} \text{ où } D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \text{ et } P = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

On achève alors la résolution comme on l'a vu au paragraphe 2.3 de l'étude des changements de bases. Rappelons que l'on a obtenu:

$$\begin{cases} x(t) = 2\alpha e^{-t} + \beta e^{2t} \\ y(t) = -\alpha e^{-t} - \beta e^{2t} + \gamma e^{5t} \\ z(t) = \beta e^{2t} + \gamma e^{5t} \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

• EXEMPLE: CAS DIAGONALISABLE AVEC V.P. MULTIPLES.

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel:

$$(S)\ \begin{cases} x'(t) = 2x(t) + 3y(t) - 6z(t) \\ y'(t) = -6x(t) - 7y(t) + 12z(t) \\ z'(t) = -3x(t) - 3y(t) + 5z(t) \end{cases}.$$

On note matriciellement ce système sous la forme:

$$X'(t) = AX(t),\ \text{avec}\ X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} 2 & 3 & -6 \\ -6 & -7 & 12 \\ -3 & -3 & 5 \end{pmatrix}.$$

Suivant les méthodes vues ci-dessus, on montre que A est diagonalisable et on réalise la diagonalisation:

$$P_A(x) = -(x+1)^2(x-2),\ \text{et}\ A = PDP^{-1}\ \text{où}\ D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\ \text{et}\ P = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

On pose: $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où: $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$. On a alors:

$$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = DU(t).$$

On est donc ramené à résoudre le système (beaucoup plus simple):

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix},\quad \text{c'est-à-dire:}\quad \begin{cases} u'(t) = 2u(t) \\ v'(t) = -v(t) \\ w'(t) = -w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions de ce dernier système sont:

$$u(t) = \alpha e^{2t},\ v(t) = \beta e^{-t},\ w(t) = \gamma e^{-t},\ \text{pour}\ \alpha, \beta, \gamma\ \text{décrivant}\ \mathbb{R}.$$

En revenant à $X(t) = PU(t) = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha e^{2t} \\ \beta e^{-t} \\ \gamma e^{-t} \end{pmatrix}$, on conclut que les solutions de (S) sont:

$$\begin{cases} x(t) = -\alpha e^{2t} + (\beta+\gamma)e^{-t} \\ y(t) = 2\alpha e^{2t} + (-\beta+\gamma)e^{-t} \\ z(t) = \alpha e^{2t} + \gamma e^{-t} \end{cases},\qquad \text{avec}\ \alpha, \beta, \gamma \in \mathbb{R}.$$

• EXEMPLE: CAS DIAGONALISABLE SUR $\mathbb{C}$ MAIS PAS SUR $\mathbb{R}$.

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel:

$$(S)\ \begin{cases} x'(t) = 2y(t) - 2z(t) \\ y'(t) = -2x(t) + z(t) \\ z'(t) = 2x(t) - y(t) \end{cases}.$$

On note matriciellement ce système sous la forme:

$$X'(t) = AX(t),\ \text{avec}\ X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} 0 & 2 & -2 \\ -2 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}.$$

On calcule: $P_A(x) = -x(x^2+9)$.

1. *On résout d'abord sur* $\mathbb{C}$. $P_A(x) = -x(x-3i)(x+3i)$, donc A est diagonalisable, et après calculs:

$$A = PDP^{-1}\ \text{où}\ D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3i & 0 \\ 0 & 0 & -3i \end{pmatrix}\ \text{et}\ P = \begin{pmatrix} 1 & 4 & 4 \\ 2 & -1+3i & -1-3i \\ 2 & -1-3i & -1+3i \end{pmatrix}.$$

En raisonnant comme sur les deux exemples précédents, on montre que les solutions à valeurs complexes du système différentiel (S) sont:

$$\begin{cases} x(t) = \alpha + 4\beta e^{3it} + 4\gamma e^{-3it} \\ y(t) = 2\alpha + (-1+3i)\beta e^{3it} + (-1-3i)\gamma e^{-3it} \\ z(t) = 2\alpha + (-1-3i)\beta e^{3it} + (-1+3i)\gamma e^{-3it} \end{cases},\qquad \text{avec}\ \alpha, \beta, \gamma \in \mathbb{C}.$$

2. *On cherche ensuite les solutions réelles.*

D'après ce qui précède, les trois fonctions vectorielles $\mathbb{R} \to \mathbb{C}^3$ définies par:

$$u_0 : t \mapsto \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix},\quad v : t \mapsto \begin{pmatrix} 4e^{3it} \\ (-1+3i)e^{3it} \\ (-1-3i)e^{3it} \end{pmatrix},\quad \overline{v} : t \mapsto \begin{pmatrix} 4e^{-3it} \\ (-1-3i)e^{-3it} \\ (-1+3i)e^{-3it} \end{pmatrix},$$

sont des solutions de (S), et toute solution complexe de (S) est combinaison linéaire à coefficients complexes de ces trois solutions u_0, v et $\overline{v}$.

On en déduit que les solutions réelles de (S) sont les combinaisons linéaires à coefficients réels des trois fonctions vectorielles $\mathbb{R} \to \mathbb{C}^3$ définies par:

$$u_0 : t \mapsto \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad w_1 = \tfrac{1}{2}(v + \overline{v}) : t \mapsto \begin{pmatrix} 4\cos 3t \\ -\cos 3t - 3\sin 3t \\ -\cos 3t + 3\sin 3t \end{pmatrix}, \quad w_2 = \tfrac{1}{2i}(v - \overline{v}) : t \mapsto \begin{pmatrix} 4\sin 3t \\ -\sin 3t + 3\cos 3t \\ -\sin 3t - 3\cos 3t \end{pmatrix}.$$

On conclut que les solutions réelles du système différentiel (S) sont:

$$\begin{cases} x(t) = \alpha + 4\beta\cos 3t + 4\gamma\sin 3t \\ y(t) = 2\alpha + (-\beta+3\gamma)\cos 3t + (-3\beta-\gamma)\sin 3t \\ z(t) = 2\alpha + (-\beta-3\gamma)\cos 3t + (3\beta-\gamma)\sin 3t \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

2.5 Quelques mots sur la trigonalisation.

On a traité à la section précédente la situation la plus favorable: celle où la matrice carrée A considérée est semblable à une matrice diagonale. Bien sûr, on l'a vu, ce n'est pas toujours le cas ! Dans bien des cas, on sera déjà très content si on sait écrire $A = PTP^{-1}$ pour une matrice T triangulaire. C'est le problème de la trigonalisation (on dit aussi triangulation, ou triangularisation), sur lequel on donne ci-dessous sans démonstration quelques brèves indications.

• Définition. Un endomorphisme f de $\mathbb{R}^n$ est dit *trigonalisable* (on dit aussi triangularisable) sur $\mathbb{R}$ lorsqu'il existe une base $\mathcal{C}$ de $\mathbb{R}^n$ telle que la matrice de f dans la base $\mathcal{C}$ soit une matrice triangulaire (supérieure) T. Une matrice carrée d'ordre n à coefficients dans $\mathbb{R}$ est dite *trigonalisable* sur $\mathbb{R}$ lorsqu'elle est semblable à une matrice triangulaire (supérieure).
Si A désigne la matrice de f dans une base quelconque $\mathcal{B}$ de $\mathbb{R}^n$, on a donc:

(l'endomorphisme f est trigonalisable) $\Leftrightarrow$ (la matrice A est trigonalisable).

Alors: $A = PTP^{-1}$ avec $A = M_{\mathcal{B}}(f)$, $T = M_{\mathcal{C}}(f)$ triangulaire, $P = \mathrm{Pass}_{\mathcal{B}\to\mathcal{C}}$.

• Proposition (trigonalisation sur $\mathbb{R}$). *Un endomorphisme f de $\mathbb{R}^n$ est trigonalisable sur $\mathbb{R}$ si et seulement si le polynôme caractéristique $P_f(x)$ est produit de n facteurs de degré 1 sur $\mathbb{R}$.*

• Proposition (trigonalisation sur $\mathbb{C}$). *Tout endomorphisme de $\mathbb{C}^n$ est trigonalisable sur $\mathbb{C}$.*

Remarque. Soit A une matrice carrée d'ordre n à coefficients dans $\mathbb{R}$. On peut toujours la considérer comme un élément de $\mathcal{M}_n(\mathbb{C})$. Elle est donc toujours trigonalisable sur $\mathbb{C}$, mais cela ne signifie bien sûr pas qu'elle est trigonalisable sur $\mathbb{R}$.
Si par exemple $P_A(x) = -(x-2)(x^2+1)$, A n'est pas trigonalisable sur $\mathbb{R}$ mais l'est sur $\mathbb{C}$ puisque $P_A(x) = -(x-2)(x-i)(x+i)$.

• Exemples. Reprenons, dans le même ordre, les cinq exemples du paragraphe 1.6.

On a déjà vu en 2.3 que le premier et le deuxième exemples correspondent à des matrices diagonalisables sur $\mathbb{R}$, et que le cinquième exemple est diagonalisable sur $\mathbb{C}$ (il n'est sur $\mathbb{R}$ ni diagonalisable ni trigonalisable puisque qu'il n'admet pas de v.p. réelles.)

Dans les troisième et quatrième exemples, les matrices considérées ne sont pas diagonalisables, mais elles sont trigonalisables sur $\mathbb{R}$ puisque le polynôme caractéristique se décompose en produit de facteurs de degré 1. Détaillons:

- Reprenons la matrice A du troisième exemple de 1.6. Soit f l'endomorphisme de $\mathbb{R}^3$ dont A est la matrice par rapport à la base canonique. Le sous-espace propre associé à la valeur simple -1 est la droite de base (u_{-1}) avec $u_{-1} = (1, 0, -1)$. Le sous-espace propre associé à la valeur double 2 est la droite de base (u_2) avec $u_2 = (2, 1, -1)$. Quelle que soit la façon de compléter la famille libre (u_{-1}, u_2) en une base $\mathcal{C}$ par l'adjonction d'un vecteur quelconque qui n'est pas combinaison linéaire de u_{-1} et u_2, la matrice de f par rapport à la base $\mathcal{C}$ est triangulaire. Prenons par exemple $\mathcal{C} = (u_{-1}, u_2, e_1)$ avec $e_1 = (1, 0, 0)$. Notons P la matrice de passage de la base canonique à la base $\mathcal{C}$. Donc

$$P = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \quad T = \mathrm{Mat}_{\mathcal{C}}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

- Reprenons la matrice A du quatrième exemple de 1.6. Soit f l'endomorphisme de $\mathbb{R}^4$ dont A est la matrice par rapport à la base canonique. Le sous-espace propre associé à la valeur simple -1 est la droite de base (u_{-1}) avec $u_{-1} = (1, 0, 0, -1)$. Le sous-espace propre associé à la valeur triple 1 est la droite de base (u_1) avec $u_1 = (0, 1, -1, 0)$. On sait que l'on peut compléter (u_{-1}, u_1) en une base $\mathcal{C} = (u_{-1}, u_1, v, w)$ de $\mathbb{R}^4$ telle que la matrice de f par rapport à la base $\mathcal{C}$ est triangulaire. Mais ici, le choix des deux vecteurs v, w n'est pas indifférent ; il existe une méthode générale pour les déterminer (voir remarque finale ci-dessous). Contentons-nous de noter ici que, pour $v = (1, 0, 2, 0)$ et $w = (0, 0, 1, 1)$, on a:

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 2 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ 2 & -1 & -1 & 2 \end{pmatrix}, \quad T = \mathrm{Mat}_{\mathcal{C}}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

• Exemple d'application à la résolution des systèmes différentiels. On a traité au paragraphe 2.4 de l'étude des changements de base un tel exemple.

• Remarque finale. Une étude plus approfondie des questions de trigonalisation conduit à la notion de réduction de Jordan, méthode beaucoup plus élaborée de recherche d'une forme canonique standard à laquelle ramener par changement de bases toute matrice carrée. Ce processus repose sur d'autres notions que le polynôme caractéristique et les sous-espaces propres (à savoir le polynôme minimal et les sous-espaces caractéristiques), qui ne peuvent être développées ici.