Cours de mathématiques (MPSI)

Arnaud Girand

12 juillet 2021

Table des matières

Première partie

Premier semestre

Chapitre I

Rappels et compléments d'analyse

1.⁻ Fonctions à valeurs réelles

a) C'est quoi ?

Soit X un ensemble (pour le moment, un intervalle ou une réunion d'intervalles style $[0,1]$, $[0,12[$ ou $\mathbb{R}^*$). On appelle **fonction** de X à valeurs dans $\mathbb{R}$ tout "mécanisme" $f : X \to \mathbb{R}$ affectant à tout point $x \in X$ une valeur numérique $f(x) \in \mathbb{R}$.

➠ **Exemple I.1.** Vous en avez vu des centaines en terminale (j'espère) : $x \mapsto x^2$, sin, exp et j'en passe ...

Notation. Nous utiliserons la convention suivante pour définir une fonction :

$$f : X \to \mathbb{R}$$
$$x \mapsto f(x)$$

ou, lorsque X est clair, $f : x \mapsto f(x)$. Dans tous les cas, nous n'écrirons **PAS** d'horreurs du style "la fonction $f(x) = \ldots$" J'aime à croire que nous sommes des gens civilisés.

Définition I.1. L'ensemble X maximal des valeurs de x pour lesquelles l'expression $f(x)$ a un sens est appelé **ensemble de définition** de la fonction f.

➠ **Exemple I.2.** L'ensemble de définition de $x \mapsto \sqrt{x-1}$ est $[1,+\infty[$.

Définition I.2. Étant donné deux fonctions $f, g : X \mapsto \mathbb{R}$ nous pouvons définir :
- leur **somme** $f + g : x \mapsto f(x) + g(x)$ sur X ;
- leur **produit** $f \times g : x \mapsto f(x)g(x)$ sur X ;
- leur **quotient** $\frac{f}{g} : x \mapsto \frac{f(x)}{g(x)}$ sur l'ensemble des $x \in X$ tels que $g(x) \neq 0$.

De plus, si les valeurs d'une fonction g sont comprises dans l'ensemble de définition d'une fonction f, on peut définir la **composée** de f par g via :

$$f \circ g : x \mapsto f(g(x)) .$$

Nous aurons l'occasion de revenir en détails sur cette notion ultérieurement.

Exemple I.3. La fonction composée de $f : x \mapsto \sqrt{x}$ et $g : x \mapsto x^2$ est $f \circ g : x \mapsto \sqrt{x^2} = |x|$, définie sur $\mathbb{R}$ (pourquoi ?).

Définition I.3. Soient f, g deux fonctions définies sur un ensemble X. On dira que f est **inférieure ou égale** à g si, pour **tout** $x \in X$, $f(x) \leq g(x)$.

Notation. $f \leq g$

Remarque I.1.

— On définit de la même façon les relations $\geq, <, >$ et $=$ sur les fonctions.

— **Attention : il s'agit d'ordres partiels.** Comparer par exemple $x \mapsto x$ et $x \mapsto -x$.

— Graphiquement, $f \leq g$ si la courbe représentative de f est **toujours** au dessus de celle de g.

b) Fonctions bornées

Définition I.4. Soit $f : X \to \mathbb{R}$ une fonction. On dit que f est :

— **majorée** si il existe un réel M tel que pour tout $x \in X$; $f(x) \leq M$;

— **minorée** si il existe un réel m tel que pour tout $x \in X$; $f(x) \geq m$;

— **bornée** si elle est majorée **et** minorée.

Remarque I.2. Une fonction f est bornée si et seulement si la fonction $|f| : x \mapsto |f(x)|$ est majorée.

Exemple I.4. La fonction $x \mapsto x^2$ est bornée sur $[0, 1]$ mais pas sur $\mathbb{R}$. La fonction $x \mapsto \frac{1}{x}$ est bornée sur $[1, \infty[$.

c) Monotonie

Définition I.5. Soit $f : X \to \mathbb{R}$ une fonction. On dit que f est :

— **croissante** (resp. strictement croissante) si pour tous $x, y \in X$ tels que $x \leq y$ alors $f(x) \leq f(y)$ (resp. $f(x) < f(y)$) ;

— **décroissante** (resp. strictement décroissante) si pour tous $x, y \in X$ tels que $x \leq y$ alors $f(x) \geq f(y)$ (resp. $f(x) > f(y)$) ;

— **monotone** si elle est croissante **ou** décroissante.

Exemple I.5. La fonction $x \mapsto x^3 - 17$ est croissante sur $\mathbb{R}$. Elle y est donc monotone.

Remarque I.3.

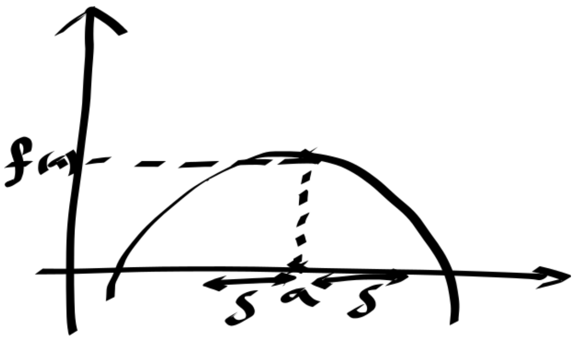
— La somme de deux fonctions croissantes est croissante. Par contre, on ne peut rien dire de la différence d'une fonction croissante et d'une fonction décroissante ...

— Soyez gentils, **ne confondez pas monotone** (croissant ou décroissant) **et constante** (toujours égale à la même valeur, donc *de facto* croissante et décroissante).

Proposition I.1. La composée de deux fonctions monotones suit la "règle des signes".

➠ **Exemple I.6.** Monotonie des fonctions $x \mapsto e^{x^2}$ (croissante) et $x \mapsto e^{-x^2}$ (décroissante).

d) Extrema



Définition I.6. Soit I un intervalle de $\mathbb{R}$ et soit $f : I \to \mathbb{R}$ une fonction. Si $a \in I$, on dit que :

— f admet un maximum local en a si :

$$\exists\delta > 0,\ (|x-a| \leq \delta) \Rightarrow (f(x) \leq f(a))\ ;$$

— f admet un minimum local en a si :

$$\exists\delta > 0,\ (|x-a| \leq \delta) \Rightarrow (f(x) \geq f(a))\ .$$

✌ **Remarque I.4.**

— Explication sans les quantificateurs, qui seront introduits au chapitre suivant. La cohérence, c'est compliqué.

— **Attention :** caractère **local** de ces notions ; le voir sur des exemples type cos.

— Si la dernière inégalité de la définition est vraie sur I, on parle d'extremum global. Jeter un oeil à $x \mapsto \sin(x)e^{-x}$.

e) Parité

Définition I.7. Soit I un intervalle de $\mathbb{R}$ **symétrique** par rapport à 0. On dit qu'une fonction $f : I \to \mathbb{R}$ est :

— **paire** si, pour tout $x \in I$, $f(-x) = f(x)$;

— **impaire** si, pour tout $x \in I$, $f(-x) = -f(x)$.

Exemple I.7. Puissances paires et impaires, cosinus, sinus.

Remarque I.5. La parité apporte des informations intéressantes relatives aux courbes représentatives : celle d'une fonction paire est symétrique par rapport à l'axe des ordonnées, celle d'une fonction impaire par rapport à l'origine. Ceci permet également de réduire le domaine d'étude.

2.– Dérivation

Ce paragraphe a vocation à n'être qu'un assemblage de recettes de cuisine pour débuter l'année. Tout ce qui suit sera démontré et mis en place proprement ultérieurement. Demain, on rase gratis.

a) Trucs de base à rappeler

Si $f : I \to \mathbb{R}$ est une fonction dérivable sur un intervalle I de $\mathbb{R}$, on note f' sa fonction dérivée. Si elle est dérivable plusieurs fois de suite, on notera $f'', f^{(3)}, \dots, f^{(k)}$ ses dérivées successives. On utilisera aussi quelques fois les notations de la physique, à savoir $\frac{\mathrm{d}f}{\mathrm{d}x}$, $\frac{\mathrm{d}^k f}{\mathrm{d}x^k}$. On n'écrire **PAS** d'abominations du type $(x^2+1)'$ ... Une dérivée est une fonction, traitons la comme telle.

Proposition I.2. Si $f : I \to \mathbb{R}$ est dérivable en $a \in I$, sa tangente a pour équation

$$y = f(a) + (x-a)f'(a) \, .$$

b) Opérations sur les dérivées

Logiquement, presque tout ceci est connu, mais il vaut mieux prévenir que guérir.

Proposition I.3. Soient f, g deux fonctions dérivables sur un intervalle I de $\mathbb{R}$ et $\lambda \in \mathbb{R}$. Alors :

(i) $f + \lambda g$ est dérivable sur I et $(f+\lambda g)' = f' + \lambda g'$;

(ii) $f \times g$ est dérivable sur I et $(f \times g)' = f' \times g + f \times g'$;

(iii) si g ne s'annule pas sur I, $\frac{f}{g}$ est dérivable sur I et

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2} \; ;$$

(iv) si g est à valeurs dans I, la composée $f \circ g$ est dérivable sur l'ensemble de dérivabilité de g et

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

Exemple I.8. La formule (iv) permet de retrouver les formules de terminales sur $\ln(u)'$ (notation abominable et à proscrire désormais) et consorts.

c) Monotonie et dérivées

Vous connaissez la musique en théorie : le signe de la dérivée est liée à la monotonie selon le schéma vu en terminale. Notons au passage que la monotonie est stricte lorsque f' ne s'annule qu'en des points isolés.

d) Réciproques

Définition I.8. Soient I, J deux intervalles de $\mathbb{R}$ et soient $f : I \to J$ et $g : J \to I$ deux fonctions. On dit que f et g sont **réciproques** si, pour tout $x \in J$ $f \circ g(x) = x$ et si, pour tout $x \in I$, $g \circ f(x) = x$.

Notation. g pourra alors être noté f^{-1} (et réciproquement, *pun intended*).

➠ **Exemple I.9.**
— exp et ln sont réciproques sur leurs ensembles de définition respectifs ;
— idem pour $x \mapsto x^2$ et $\sqrt{\cdot}$.

✌ **Remarque I.6.** Graphiquement, la courbe représentative de la réciproque d'une fonction s'obtient par symétrie relativement à la première bissectrice, *i.e* la droite d'équation $y = x$.

Le théorème qui suit est important, et sera démontré ultérieurement. Si vous avez l'impression que nous prenons de mauvaises habitudes, vous avez raison.

Théorème I.4 (Dérivée d'une réciproque)**.**
Soient I, J deux intervalles de $\mathbb{R}$ et soit $f : I \to J$ une fonction dérivable telle que f' **ne s'annule pas** sur I. Alors :
(i) f admet une réciproque f^{-1} dérivable sur J ;
(ii)
$$(f^{-1})' = \frac{1}{f' \circ f^{-1}} .$$

➠ **Exemple I.10.** On peut retrouver la dérivée de l'exponentielle. Ne me remerciez pas, c'est mon métier.

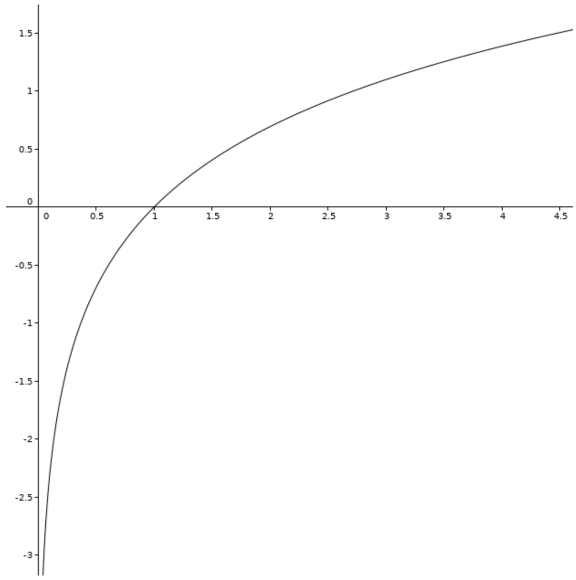
3.⁻ Logarithmes, exponentielles

a) Le cas néperien

Point culture : le nom commun néperien vient de John Napier (francisé en Jean Neper), qui fut théologien, physicien, astronome, mathématicien et écossais et vécu de 1550 à 1617.

Définition I.9. On appelle **logarithme néperien** l'unique fonction, notée ln, définie sur $\mathbb{R}_+^*$ telle que :
(i) $\forall x > 0, \ln'(x) = \frac{1}{x}$;
(ii) $\ln(1) = 0$.

L'allure de la courbe représentative de cette fonction devrait vous être bien connu, mais réiterons la au cas où :



Proposition I.5.

(i) ln est strictement croissante ;

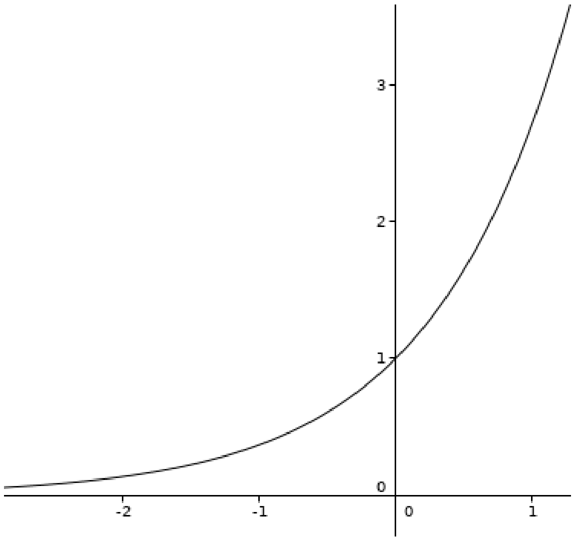
(ii) $\ln(x) \xrightarrow[x\to\infty]{} \infty$ et $\ln(x) \xrightarrow[x\to 0]{} -\infty$;

(iii) soient $x, y \in \mathbb{R}_+^*$. On a alors :
— $\ln(xy) = \ln(x) + \ln(y)$;
— $\ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y)$.

Ceux parmi vous qui suivent auront remarqué que la dérivée du logarithme néperien ne s'annule pas sur $\mathbb{R}_+^*$; le théorème I.4 entraîne donc que cette fonction admet une réciproque, dont la dérivée sera elle–même.

Définition I.10. La réciproque de la fonction ln, est appelée **exponentielle**.

Notation. $\exp : \mathbb{R} \to \mathbb{R}_+^*$



Proposition I.6.

(i) exp est strictement croissante, à valeurs strictement positives ;

(ii) $\exp(x) \xrightarrow[x\to\infty]{} \infty$ et $\exp(x) \xrightarrow[x\to-\infty]{} 0$;

(iii) si on pose $e = \exp(1)$, alors $\ln(e) = 1$;

(iv) soient $x, y \in \mathbb{R}$. On a alors :
- $\exp(x + y) = \exp(x)\exp(y)$;
- $\exp(x - y) = \frac{\exp(x)}{\exp(y)}$;
- $\exp(xy) = \exp(x)^y$.

(v) exp est dérivable et $\exp' = \exp$.

Remarque I.7. La troisième formule du point (iv) justifie l'écriture "$\exp(x) = e^x$" vue en terminale et que nous utiliserons désormais lorsque cela nous semblera pertinent.

b) En base quelconque

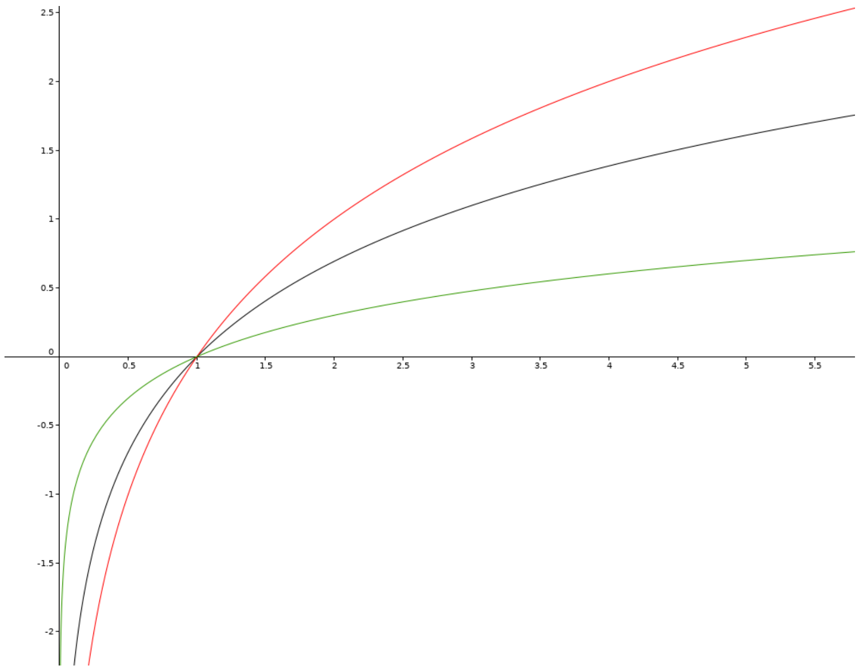
Définition I.11. Soit $a \in \mathbb{R}_+^*$ différent de 1. On appelle **logarithme en base** a la fonction

$$\begin{aligned} \log_a : \mathbb{R}_+^* &\to \mathbb{R} \\ x &\mapsto \frac{\ln(x)}{\ln(a)} \end{aligned} \quad .$$

Exemple I.11.
- Pour $a = 10$, on retrouve le logarithme décimal cher aux physiciens, souvent noté log.
- Pour $a = e$, pas de surprise, on obtient le logarithme néperien.

— Pour $a = 2$, on tombe sur le logarithme binaire, prisé des informaticiens, et parfois noté lg.

Les courbes de ces fonctions partagent une allure commune, modulo quelques différences d'inflexion. Plus la valeur de a est élevée, plus la courbe sera dominée à l'infini.



✌ **Remarque I.8.** Les logarithmes sont tous multiples du logarithme néperien : la proposition I.5 reste donc valide.

Proposition I.7. Soit $a > 0$ différent de 1. Alors la fonction $\log_a$ admet une réciproque, appelée **exponentielle en base** a.

Notation. $\exp_a : \mathbb{R} \to \mathbb{R}_+^*$

Proposition I.8. Soit $a > 0$ différent de 1 et soit $x \in \mathbb{R}$. Alors :

$$\exp_a(x) = e^{x\ln(a)} .$$

Démonstration. On sait que $\log_a(\exp_a(x)) = x$, ce qui signifie que $\ln(\exp_a(x)) = x\ln(a)$; on obtient le résultat en composant à droite et à gauche par exp. □

Notation. Tout ceci justifie la notation "$a^x = e^{x\ln(a)}$" que nous utiliserons désormais.

✌ **Remarque I.9.** Dans le doute, n'hésitez pas à repasser à l'écriture exponentielle dans tout exercice mettant en scène des puissances.

4.⁻ Puissances

a) C'est quoi ?

Définition I.12. Soit $a \in \mathbb{R}$. On appelle **fonction puissance** a la fonction

$$\begin{aligned} \phi_a : \mathbb{R}_+^* &\to \mathbb{R} \\ x &\mapsto e^{a\ln(x)} \end{aligned} .$$

Notation. Pour $x > 0$, on notera $x^a = \phi_a(x)$.

Remarque I.10.

— Si a est un entier positif, $\phi_a(x) = \underbrace{x \times \ldots \times x}_{a \text{ fois}}$. On retrouve la définition usuelle de la puissance, ce qui est bienvenu. La fonction ϕ_a est alors bien définie sur $\mathbb{R}$.

— Si a est un entier strictement négatif,

$$\phi_a(x) = \underbrace{\frac{1}{x} \times \ldots \times \frac{1}{x}}_{a \text{ fois}} .$$

La fonction ϕ_a est alors bien définie sur $\mathbb{R}^*$.

— Dans tous les cas, la fonction ϕ_a n'est *a priori* bien définie que pour x **strictement positif**.

Exemple I.12. Vous connaissez sans doute ϕ_2, ϕ_3 et leurs amies ; mais connaissez–vous ϕ_π ? Ou $\phi_{\frac{1}{2}}$?

Proposition I.9. Soient $x, y \in \mathbb{R}_+^*$ et soient $a, b \in \mathbb{R}$. Alors :

(i) $x^a y^a = (xy)^a$;

(ii) $x^a x^b = x^{a+b}$;

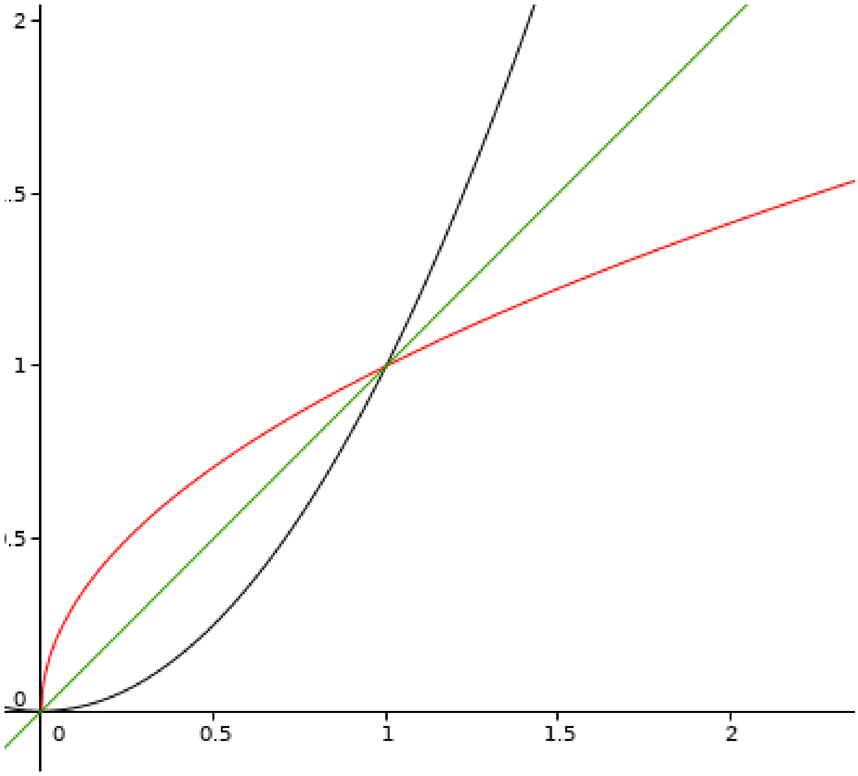
(iii) $(x^a)^b = x^{ab}$;

(iv) $1^a = 1 = x^0$;

(v) $\ln(x^a) = a\ln(x)$.

Démonstration. Utiliser la forme exponentielle. □

Concernant l'allure des courbes, on différencie aisément les exposants $a > 1$ (forme parabolique convexe) des exposants $0 < a < 1$ (forme concave, penser racine carrée). Nous justifierons ces variations dans le paragraphe suivant.



b) Variations

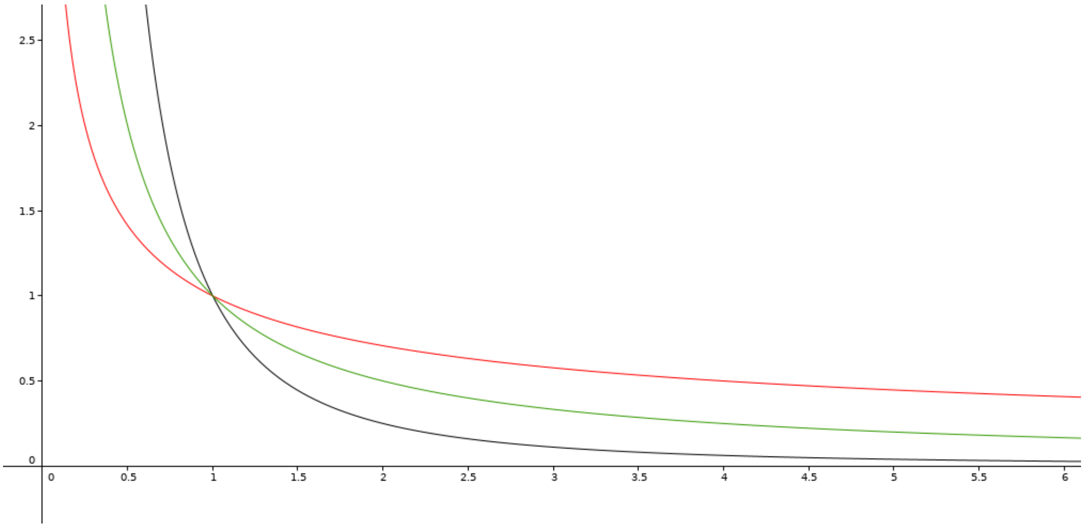
En utilisant la formule de dérivée d'une composée, on trouve, pour $x > 0$ et $a \in \mathbb{R}$:

$$\boxed{\phi_a'(x) = ax^{a-1}\ .}$$

On retrouve de cette façon les dérivées usuelles suivantes vues en terminale : puissances entières, mais aussi puissances inverses, racine carrée. Cette formule est un véritable couteau suisse de la dérivation.

Concernant les limites, on remarque aisément que, si $a > 0$, $\phi_a(x) \xrightarrow[x \to 0]{} 0$. Ceci nous permet de poser la convention suivante : $0^a = 0$ pour tout $a > 0$. Ceci est à **ne pas confondre** avec x^0, qui vaut 1 quel que soit x et ce sans aucune ambiguïté.

La formule de la dérivée ci–dessus nous permet de justifier les variations vues précédemment dans le cas $a > 0$; le cas $a < 0$ apporte quant à lui son lot de courbes décroissantes, comme illustré ci–ensuite. Cette fois ci, les courbes dominantes à l'infini sont celles pour $|a| < 1$.



c) Racines

Définition I.13. Soit $n \in \mathbb{N}^*$; la fonction $\phi_{\frac{1}{n}}$ est appelée **racine n–ième**.

Notation. $x \mapsto \sqrt[n]{x}$

Soit $x > 0$ et soit $n \geq 1$; effectuons un petit calcul :

$$\begin{aligned}\sqrt[n]{x^n} &= \phi_{\frac{1}{n}}(x^n) \\ &= \exp\left(\frac{1}{n}\ln(x^n)\right) \\ &= \exp(\ln(x)) \\ &= x\ .\end{aligned}$$

De la même façon, on vérifie que $(\sqrt[n]{x})^n = x$: **la fonction $\phi_{\frac{1}{n}}$ est donc la réciproque de ϕ_n.**

✌ **Remarque I.11.** Si n est impair, il est possible d'étendre la fonction $\phi_{\frac{1}{n}}$ à $\mathbb{R}$ tout entier.

d) Croissances comparées

✖ Ce qui suit est **absolument fondamental**, et sera démontré ultérieurement.

Proposition I.10. Soient $a, b > 0$; on a alors les limites suivantes :

(i)
$$\frac{(\ln(x))^b}{x^a} \xrightarrow[x\to\infty]{} 0 \;;$$

(ii)
$$x^a(\ln(x))^b \xrightarrow[x\to 0]{} 0 \;;$$

(iii)
$$\frac{e^{ax}}{x^b} \xrightarrow[x\to\infty]{} \infty \;;$$

(iv)
$$x^b e^{ax} \xrightarrow[x\to -\infty]{} 0 \;.$$

Chapitre II

Logique

1.⁻ Propositions

a) C'est quoi ?

Définition II.1. Une **assertion** est un énoncé, pouvant être vrai ou faux, mettant en relation des objets mathématiques.

➠ **Exemple II.1.**
- "$2+2=12$" est une assertion.
- " De toute suite réelle bornée on peut extraire une suite qui converge" également.

Convention. Une assertion est toujours vraie **ou** fausse, et jamais les deux à la fois.

Le lecteur averti aura remarqué que les termes "assertion vraie" et "assertion fausse" n'ont pas été définis…

Définition II.2. Faire des mathématiques, c'est énoncer (et démontrer !) des assertions vraies, appelées **propositions**.

Dans la suite de ce cours, nous parlerons de **valeur de vérité** d'une assertion. Cette valeur est de 1 si l'assertion associée est vraie et de 0 sinon.

Définition II.3. Un peu de vocabulaire, en vrac.
- Un **théorème** est une proposition importante.
- Un **lemme** est une proposition qui sert (dans le cadre où elle est énoncée) principalement à démontrer une autre proposition.
- Un **corollaire** est une proposition qui découle immédiatement d'une autre proposition.
- Un **axiome** est une assertion que l'on **décide** vraie.

b) Assertions équivalentes

Définition II.4. Soient P et Q deux assertions. On dira que P et Q sont équivalentes si elles ont exactement les mêmes valeurs de vérité.

Notation. $P \Leftrightarrow Q$

Exemple II.2.
- $(1 = 0) \Leftrightarrow 2$ est impair.
- Si $n \in \mathbb{Z}$, (n est pair) $\Leftrightarrow$ ($n + 1$ est impair).

2.– Connecteurs logiques

Les connecteurs jouent un rôle fondamental en logique puisqu'ils permettent de créer de nouvelles assertions à partir de assertions existantes (sans eux, les mathématiques seraient relativement bornées...).

a) Le connecteur "ET"

Définition II.5. Soient P et Q deux assertions. On définit l'assertion "P et Q" par la table de vérité suivante :

$P\backslash Q$	V	F
V	V	F
F	F	F

Notation. $P \wedge Q$

Exemple II.3.
- $(1 + 1 = 2) \wedge (2 + 1 = 3)$ est vraie.
- $(1 + 1 = 0) \wedge (2 + 1 = 3)$ est fausse.

Proposition II.1. Soient P, Q et R des assertions. Alors :
(i) $P \wedge P \Leftrightarrow P$ **[idempotence]** ;
(ii) $P \wedge Q \Leftrightarrow Q \wedge P$ **[commutativité]** ;
(iii) $P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R$ **[associativité]**.

Démonstration. Il s'agit uniquement de dresser les tables de vérités de chacun des membres de droite/gauche des équivalences. N'en démontrons qu'une, par exemple 2 :

P	Q	$P \wedge Q$	$Q \wedge P$
F	F	F	F
F	V	F	F
V	F	F	F
V	V	V	V

La troisième et la quatrième colonne sont les mêmes, donc les deux assertions associées sont équivalentes.

□

✌ **Remarque II.1.** Le troisième point de cette assertion donne un sens à la notation $P \wedge Q \wedge R$ qui désigne arbitrairement $P \wedge (Q \wedge R)$ ou $(P \wedge Q) \wedge R$.

b) Le connecteur "OU"

Définition II.6. Soient P et Q deux assertions. On définit l'assertion "P ou Q" par la table de vérité suivante :

$P \backslash Q$	V	F
V	V	V
F	V	F

Notation. $P \vee Q$.

✌ **Remarque II.2.** On parle aussi de "OU inclusif", par opposition au "OU exclusif".

➠ **Exemple II.4.** $(1+1=0) \vee (1+1=2)$ est vraie.

Proposition II.2. Soient P, Q et R des assertions. Alors on a :

(i) $P \vee P \Leftrightarrow P$;

(ii) $P \vee Q \Leftrightarrow Q \vee P$;

(iii) $P \vee (Q \vee R) \Leftrightarrow (P \vee Q) \vee R$;

(iv) $P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R)$ **[distributivité]** ;

(v) $P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$ **[distributivité]**.

Démonstration. En exercice. Il est conseillé à ce stade de la lecture de démontrer au moins l'une des assertions ci dessus. □

✌ **Remarque II.3.** Le troisième point de cette assertion donne un sens à la notation $P \vee Q \vee R$ qui désigne arbitrairement $P \vee (Q \vee R)$ ou $(P \vee Q) \vee R$.

c) Le connecteur "NON"

Définition II.7. Soit P une assertion. On définit l'assertion "Non P" par la table de vérité suivante :

P	Non P
F	V
V	F

Notation. $\overline{P}$

Proposition II.3. Soit P une assertion. Alors :

(i) $\overline{\overline{P}} \Leftrightarrow P$;

(ii) $\overline{(P \wedge \overline{P})}$ est vraie **[non–contradiction]** ;

(iii) $P \vee \overline{P}$ est vraie **[tiers exclus]**.

Le théorème suivant est du à Augustus de Morgan (1806—1871), mathématicien et logicien britannique.

Théorème II.4 (De Morgan)**.**
Soient P et Q deux assertions. Alors :

(i) $\overline{P \vee Q} = \overline{P} \wedge \overline{Q}$

(ii) $\overline{P \wedge Q} = \overline{P} \vee \overline{Q}$

Démonstration. (i) Voici la table de vérité incluant toutes les étapes intermédiaires necessaires à la démonstration. On déduit le résultat des colonnes 4 et 7.

P	Q	$P \vee Q$	$\overline{P \vee Q}$	$\overline{P}$	$\overline{Q}$	$\overline{P} \wedge \overline{Q}$
F	F	F	V	V	V	V
F	V	V	F	V	F	F
V	F	V	F	F	V	F
V	V	V	F	F	F	F

(ii) $P \wedge Q \Leftrightarrow \overline{\overline{P}} \wedge \overline{\overline{Q}} \Leftrightarrow \overline{\overline{P} \vee \overline{Q}}$ d'après le 1. Donc $\overline{P \wedge Q} \Leftrightarrow \overline{\overline{\overline{P} \vee \overline{Q}}} \Leftrightarrow \overline{P} \vee \overline{Q}$. D'où le résultat.

□

✍ **Exercice II.1.** Soient P et Q deux assertions. On définit l'assertion $P \otimes Q = (P \wedge \overline{Q}) \vee (\overline{P} \wedge Q)$ (ce nouveau connecteur est appelé "**OU exclusif**"). Dresser la table de vérité de $\otimes$.

d) Le connecteur "IMPLIQUE"

Définition II.8. Soient P et Q deux assertions. On définit l'assertion "P implique Q" comme étant fausse si et seulement si P est vraie alors que Q est fausse.

Notation. $P \Rightarrow Q$

✌ **Remarque II.4.** Ceci se traduit par la table de vérité suivante.

$P\backslash Q$	V	F
V	V	F
F	V	V

On notera que "faux" implique tout ce que l'on veut...

➠ **Exemple II.5.**

— $(1 = 0) \Rightarrow$ (J'ai mangé une pomme hier à 21h)
— (3 est pair) $\Rightarrow$ (2 est pair)
— Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, (f est dérivable) $\Rightarrow$ (f est continue)

Proposition II.5. Soient P, Q et R des assertions. Alors on a :

1. $(P \Rightarrow Q) \wedge (Q \Rightarrow R)) \Rightarrow (P \Rightarrow R)$ [Transitivité]
2. $(P \Leftrightarrow Q) \Leftrightarrow ((P \Rightarrow Q) \wedge (Q \Rightarrow P))$
3. $(P \Rightarrow Q) \Leftrightarrow \overline{P} \vee Q$
4. $(P \vee Q) \Leftrightarrow (\overline{P} \Rightarrow Q)$
5. $\overline{(P \Rightarrow Q)} \Leftrightarrow P \wedge \overline{Q}$
6. $(P \Rightarrow Q) \Leftrightarrow (\overline{Q} \Rightarrow \overline{P})$ [Contraposée]
7. $(\overline{P} \Rightarrow (Q \wedge \overline{Q})) \Rightarrow P$ [Démonstration par l'absurde]

Démonstration. Une fois n'est pas coutume, la preuve se fait facilement par tables de vérité (par exemple). □

3.⁻ Quantificateurs

a) Quantificateurs, phrases quantifiées

Définition II.9. On définit les deux **quantificateurs suivants** :

1. le **quantificateur universel**, noté "$\forall$"
2. le **quantificateur existentiel**, noté "$\exists$"

On appelle **phrase** (ou expression) **quantifiée** toute propriété faisant intervenir des quantificateurs.

Signification des quantificateurs :

Soit P une assertion portant sur les éléments d'un ensemble $\mathbb{E}$. On définit alors de nouvelles assertions :

1. $\forall x \in \mathbb{E}, P(x)$, qui signifie que tous les éléments de $\mathbb{E}$ vérifient la propriété P.
2. $\exists x \in \mathbb{E}, P(x)$, qui signifie qu'il existe un élément de $\mathbb{E}$ qui vérifie la propriété P.

De plus, si un **unique** élément de $\mathbb{E}$ vérifie P, on note : $\exists! x \in \mathbb{E}, P(x)$.

Définition II.10. On appelle **prédicat** d'une variable x toute assertion $P(x)$ dépendant de celle–ci.

b) Échange de quantificateurs, dépendances

◇ Échange de quantificateurs

Soit P une propriété dépendant de deux paramètres $x \in \mathbb{E}$ et $y \in \mathbb{F}$. Alors les deux propriétés suivantes sont équivalentes :

1. $\exists x \in \mathbb{E}, \exists y \in \mathbb{F}, P(x, y)$
2. $\exists y \in \mathbb{F}, \exists x \in \mathbb{E}, P(x, y)$

Il est ainsi possible au sein d'une phrase quantifiée d'intervertir deux quantificateurs existentiels sans en changer le sens. De la même façon, sont équivalents :

1. $\forall x \in \mathbb{E}, \forall y \in \mathbb{F}, P(x, y)$
2. $\forall y \in \mathbb{F}, \forall x \in \mathbb{E}, P(x, y)$

➠ **Exemple II.6.** Pour $x, y \in \mathbb{Z}$, on définit $P(x, y) = x|y$.

En revanche, il est **impossible** d'échanger la place d'un quantificateur universel et d'un quantificateur existentiel, comme nous le verrons ci-après. Les deux équivalences ci dessus donnent un sens aux notations "$\exists(x, y) \in \mathbb{E} \times \mathbb{F}, P(x, y)$" et "$\forall(x, y) \in \mathbb{E} \times \mathbb{F}, P(x, y)$".

◇ Dépendances

Le problème des dépendances dans les phrases quantifiées est très important en mathématiques. Par exemple, considérons les deux assertions suivantes, portant sur une fonction $f : \mathbb{R} \to \mathbb{R}$:

(1) $\forall x \in \mathbb{R}, \quad \exists y \in \mathbb{R}_+, \quad y = x^2$;

(2) $\exists y \in \mathbb{R}_+, \quad \forall x \in \mathbb{R}, \quad y = x^2$.

Ces deux assertions ne sont **absolument pas** équivalentes ! Pourtant, la seule différence entre ces deux assertions est la place occupée par le "$\forall x \in \mathbb{R}$"...

Le noeud du problème réside dans les **dépendances**. Dans l'assertion (1), le réel y dépend de x : en effet, ce que signifie cette phrase quantifiée, c'est que si on fixe x dans $\mathbb{R}$, **alors** on va pouvoir trouver un y tel que la propriété soit vérifiée. Tandis que (2) nous indique que l'on va pouvoir trouver y tel que la propriété soit vraie pour tous x.

Conséquence :
Lorsque l'on manipule des phrases quantifiées, il faut apporter une attention toute particulière aux dépendances des paramètres entre eux afin d'éviter de monumentales erreurs d'interprétation. Une règle efficace en pratique est la suivante : les "$\exists$" dépendent des "$\forall$" qui les précédent. Par exemple, dans $\exists z, \forall t \exists q, P(z,t,q)$ le paramètre q dépend de t alors que z et t ne dépendent d'aucun des autres paramètres (et ne sont pas dépendant entre eux).

c) Négation de phrases quantifiées

Proposition II.6. Soit P un prédicat dépendant d'un paramètre x. Alors :

1. $\overline{(\forall x, P(x))} \Leftrightarrow (\exists x, \overline{P(x)})$
2. $\overline{(\exists x, P(x))} \Leftrightarrow (\forall x, \overline{P(x)})$

➠ **Exemple II.7.**

$$\overline{\forall \varepsilon > 0, \quad \forall x \in \mathbb{R}, \quad \exists \delta > 0, \quad \forall y \in \mathbb{R}, \quad (|x-y| \leq \delta) \Rightarrow (|f(x)-f(y)| \leq \varepsilon)}$$
$$\Leftrightarrow \exists \varepsilon > 0, \quad \overline{\forall x \in \mathbb{R}, \quad \exists \delta > 0, \quad \forall y \in \mathbb{R}, \quad (|x-y| \leq \delta) \Rightarrow (|f(x)-f(y)| \leq \varepsilon)}$$
$$\Leftrightarrow \exists \varepsilon > 0, \quad \exists x \in \mathbb{R}, \quad \overline{\exists \delta > 0, \quad \forall y \in \mathbb{R}, \quad (|x-y| \leq \delta) \Rightarrow (|f(x)-f(y)| \leq \varepsilon)}$$
$$\vdots$$
$$\Leftrightarrow \exists \varepsilon > 0, \quad \exists x \in \mathbb{R}, \quad \forall \delta > 0, \quad \exists y \in \mathbb{R} \quad (|x-y| \leq \delta) \wedge (|f(x)-f(y)| > \varepsilon)$$

4.⁻ Méthodes de démonstration

Énoncer des assertions, c'est bien. Les démontrer c'est mieux. Les démontrer proprement, c'est encore mieux ! Voici une sélection de quelques-unes des méthodes les plus usitées dans tous les domaines des mathématiques...

a) Comment démontrer que $P \Rightarrow Q$?

Cette méthode de démonstration est absolument primordiale. On voit trop souvent des rédactions du genre "$P \Rightarrow R \Rightarrow S \Rightarrow Q \Rightarrow$ ok"... La première règle à respecter lorsque l'on veut démontrer une implication est de **ne pas utiliser le symbole "$\Rightarrow$"** ! En effet, comme l'on vient de le voir, **celui-ci est un connecteur logique qui n'a absolument rien à voir avec le mot "donc"**. Pour montrer que $P \Rightarrow Q$, le mieux est de suivre la méthode suivante :

1. Supposer que P est vraie.
2. Démontrer qu'alors Q est vraie.

Cette approche est en règle générale la seule valable...

Exemple "idiot" :
On veut démontrer "$(2+3=6) \Rightarrow$ (16 est un carré parfait)".

1. Supposons que $2+3=6$.
2. $16 = 4 \times 4$ donc 16 est un carré parfait. D'où le résultat.

Cet exemple n'a aucun intérêt mathématique étant donné que l'on a pas fait usage de l'hypothèse (qui par ailleurs à la facheuse propriété d'être fausse...) ! Lorsque l'on arrive à démontrer que $P \Rightarrow Q$ sans avoir utilisé la véracité de P, c'est que l'on a fort vraisemblablement commis une erreur...

b) Comment démontrer que $P \Leftrightarrow Q$?

Pour démontrer une équivalence, il faut et il suffit de montrer deux implications. Ainsi, on procède en 2 temps :

1. On montre que $P \Rightarrow Q$.
2. On montre que $Q \Rightarrow P$.

Ce type de rédaction à l'immense avantage d'être bien moins "casse-gueule" que les suites (plus ou moins logiques, hélas...) de symboles " $\Leftrightarrow$ "...

c) Comment démontrer que $P \vee Q$?

Sur le papier, ce genre de choses est excessivement simple. En pratique, il donne lieu à des raisonnements souvent assez brouillons. La meilleure méthode pour démontrer ce genre d'assertion est de procéder de la façon suivante : comme on a vu que $(P \vee Q) \Leftrightarrow (\overline{P} \Rightarrow Q)$, démontrer une assertion du type "soit P, soit Q" peut se faire en suivant ce plan :

1. Supposer que P est fausse.
2. Montrer qu'alors Q est vraie.

d) Comment démontrer que $P \wedge Q$?

Lorsque l'on doit démontrer "P et Q", le mieux reste encore... de démontrer P **et** de démontrer Q ! Attention cependant à bien séparer les deux démonstrations, il ne s'agit pas de montrer $P \Rightarrow Q$...

e) Contraposée

Comme $(P \Rightarrow Q) \Leftrightarrow (\overline{Q} \Rightarrow \overline{P})$, il est possible de démontrer $P \Rightarrow Q$ en démontrant $\overline{Q} \Rightarrow \overline{P}$.

f) Démonstration par l'absurde

On a vu que $(\overline{P} \Rightarrow (Q \wedge \overline{Q})) \Rightarrow P$. Ainsi, si on suppose "Non P" et que l'on aboutit à une contradiction, alors P est vraie. Il est recommandé de ne pas abuser de cette méthode.

g) Comment démontrer que $\forall x \in \mathbb{E}, P(x)$?

La procédure est la suivante :

1. On se donne $x \in \mathbb{E}$.
2. On montre $P(x)$.

L'erreur (gravissime !) serait ici de se donner un x "particulier" et de montrer $P(x)$...

h) Comment démontrer que $\exists x \in \mathbb{E}, P(x)$?

Certainement l'une des démonstrations les plus difficiles a priori... Les deux méthodes ci-après comptent parmi les seules efficaces (et correctes !) :

1. Exhiber un élément x vérifiant P.
2. Raisonner par l'absurde en supposant qu'aucun élément de $\mathbb{E}$ ne vérifie P.

i) Comment démontrer que $\exists! x \in \mathbb{E}, P(x)$?

1. On commence par montrer que $\exists x \in \mathbb{E}, P(x)$ (souvent l'étape la plus difficile).
2. On suppose qu'il existe x et $\hat{x}$ dans $\mathbb{E}$ vérifiant P.
3. On montre que $x = \hat{x}$.

Chapitre III

Ensembles

1.⁻ C'est quoi ?

a) Ensembles, éléments

Définition III.1. On appelle **ensemble** toute collection (non ordonnée) d'objets, appelés éléments.

Notation. $a \in E$ signifie que l'objet a est un élément de l'ensembe E.

La question qui va nous occuper dans un premier temps est la suivante : **comment décrire un ensemble ?** Plusieurs choix s'offrent à nous.

1. **Lister les éléments de l'ensemble.**

$$E = \{\clubsuit, \nabla, 4\}$$

2. **Caractériser l'ensemble par une propriété.**

$$E = \{ \underbrace{x \in \mathbb{R}}_{\text{cas de base}} \mid \underbrace{x > 0}_{\text{prédicat}} \}$$

$$F = \{n \in \mathbb{N} \,|\, \exists p \in \mathbb{N},\ n = 2p\}$$

3. **Décrire ses éléments comme image par une application**.

$$E = \{x^2 \,|\, x \in \mathbb{R}\} = \{y \in \mathbb{R} \,|\, \exists x \in \mathbb{R},\ y = x^2\}$$

Vocabulaire. Un ensemble ne contenant qu'un seul élément est appelé **singleton.**

b) Difficultés

Cette définition naïve n'est pas sans risques. L'une des complications rencontrées, appelée **paradoxe de Russel**, peut s'énoncer de la façon suivante : soit $\mathscr{E} = \{E \text{ ensemble} \,|\, E \notin E\}$; a–t–on $\mathscr{E} \in \mathscr{E}$? Peu importe que l'on suppose que ce soit vrai ou faux, on arrive à une contradiction.

Pour éviter ce type de difficulté, on prendra toujours gare à définir nos ensembles à partir d'un cas de base précis, clair et sans ambiguïté.

c) Égalité

Axiome A. Deux ensembles sont égaux si et seulement si ils ont les mêmes éléments.

Exemple III.1.
- $\{0,1,2\} = \{2,0,1\}$;
- $\{1,2\} \neq \{2,12\}$.

Par conséquent, la **seule** méthode valable pour démontrer que deux ensembles A et B sont égaux est la suivante :
- se donner $x \in A$ et démontrer que $x \in B$;
- réciproquement, se donner $x \in B$ et montrer que $x \in A$.

d) Inclusion

Définition III.2. On dit qu'un ensemble A est **inclus** dans un ensemble B si $\forall x \in A,\ x \in B$.

Notation. $A \subset B$

✖ **Attention :** ne pas confondre inclusion et appartenance : $1 \in \mathbb{R}$, $\{1\} \subset \mathbb{R}$.

Proposition III.1. Soient A, B, C des ensembles. Alors :
(i) $A \subset A$ **[reflexivité]** ;
(ii) $(A \subset B) \wedge (B \subset A) \Leftrightarrow (A = B)$ **[antisymétrie]** ;
(iii) $(A \subset B) \wedge (B \subset C) \Rightarrow (A \subset C)$ **[transitivité]**.

Démonstration. Immédiat. □

✌ **Remarque III.1.** Le point (ii) nous indique que pour démontrer une égalité d'ensembles, il faut démontrer deux inclusions. **À retenir.**

e) Zoologie des ensembles classiques

Axiome B. Il existe un ensemble ne contenant aucun élément, appelé **ensemble vide** et noté $\emptyset$.

✌ **Remarque III.2.** Si P est un prédicat faux sur un ensemble E, alors $\{x \in E \mid P(x)\} = \emptyset$. Par exemple, $\emptyset = \{x \in \mathbb{R} \mid x^2 < 0\}$.

Profitions de ce paragraphe pour rappeler quelques définitions logiquement vues dans une vie antérieure.
- $\mathbb{N}$ est l'ensemble des entiers naturels.
- $\mathbb{Z}$ est l'ensemble des entiers relatifs, *i.e*

$$\mathbb{Z} = \{a - b \mid a, b \in \mathbb{N}\}\ .$$

— $\mathbb{Q}$ est l'ensemble des nombres rationnels, soit :

$$\mathbb{Q} = \left\{ \frac{p}{q} \,\middle|\, p \in \mathbb{Z},\, q \in \mathbb{N}^* \right\} .$$

— $\mathbb{R}$ est l'ensemble des nombres réels.
— $\mathbb{C}$ est l'ensemble des nombres complexes, soit :

$$\mathbb{C} = \{a + ib \,|\, a, b \in \mathbb{R}\}$$

où i est tel que $i^2 = -1$.

Rappelons au passage que l'on a la chaine d'inclusions suivante :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} .$$

f) Parties d'un ensemble

Axiome C. Soit E un ensemble. Alors il existe un unique ensemble, noté $\mathscr{P}(E)$, tel que :

$$A \in \mathscr{P}(E) \Leftrightarrow A \subset E .$$

Cet ensemble est appelé **ensembles des parties de** E.

✌ **Remarque III.3.**
— Il s'agit donc d'un ensemble d'ensembles ...
— Quel que soit l'ensemble E, on a toujours $\emptyset \in \mathscr{P}(E)$ et $E \in \mathscr{P}(E)$.

➠ **Exemple III.2.**
— $\mathscr{P}(\emptyset) = \{\emptyset, \{\emptyset\}\}$.
— $\mathscr{P}(\{a, b, c\}) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$.

✌ **Remarque III.4.** La construction "moderne" de l'ensemble $\mathbb{N}$ des entiers naturels s'obtient via les "ensembles parties itérés" de l'ensemble vide, $i.e$ les $\mathscr{P}(\mathscr{P}(\ldots\ldots\mathscr{P}(\emptyset)))$.

2.⁻ Opérations sur les ensembles

a) Réunion

Définition III.3. Soient A, B deux sous–ensembles d'un ensemble E ; on appelle **réunion** de A et B l'ensemble

$$\{x \in E \,|\, (x \in A) \vee (x \in B)\} .$$

Notation. $A \cup B$

➠ **Exemple III.3.** $\{1, 2\} \cup \{2, 7\} = \{1, 2, 7\}$.

✌ **Remarque III.5.** De façon plus générale, si $(A_i)_{i\in I}$ est une famille d'ensembles indexée par un ensemble I, ce qui signifie qu'à chaque élément i de I on associe un unique sous–ensemble A_i d'un ensemble E, on pose :

$$\bigcup_{i\in I} A_i = \{x \in E \mid \exists i \in I,\, x \in A_i\}\ .$$

Il s'agit du plus petit sous–ensemble de E contenant tous les A_i. Notons que I peut tout à fait être infini ; on peut par exemple montrer assez aisément (démontrer deux inclusions) que

$$\bigcup_{n\in\mathbb{N}^*} \left[0, 1-\frac{1}{n}\right[= [0,1[\ .$$

Proposition III.2. Soient A, B, C, D quatre ensembles. Alors :

(i) $A \cup B = B \cup A$ **[commutativité]** ;

(ii) $(A\cup B)\cup C = A \cup (B\cup C)$ **[associativité]** ;

(iii) $A \cup \emptyset = A$ **[$\emptyset$ est neutre]** ;

(iv) $A \cup A = A$ **[idempotence]** ;

(v) $(A\cup B = A) \Leftrightarrow (B \subset A)$;

(vi) $(A \subset B)\wedge(C\subset D) \Rightarrow (A\cup C \subset B \cup D)$.

Démonstration. Immédiat en utilisant les propriétés du connecteur $\vee$. Un dessin peut être utile pour (v) et (vi). □

b) Intersection

Définition III.4. Soient A, B deux sous–ensembles d'un ensemble E ; on appelle **instersection** de A et B l'ensemble

$$\{x\in E \mid (x\in A)\wedge(x\in B)\}\ .$$

Notation. $A\cap B$

➠ **Exemple III.4.**
— $\{1,2\}\cap\{2,7\} = \{2\}$;
— $\{1,3\}\cap\{2,7\} = \emptyset$.

Vocabulaire. Deux ensembles A, B tels que $A\cap B=\emptyset$ sont dits **disjoints**.

✌ **Remarque III.6.** De façon plus générale, si $(A_i)_{i\in I}$ est une famille de sous–ensembles d'un ensemble E indexée par un ensemble I, on pose :

$$\bigcap_{i\in I} A_i = \{x \in E \mid \forall i \in I,\, x \in A_i\}\ .$$

Il s'agit du plus petit sous–ensemble de E contenant tous les A_i. Notons que I peut tout à fait être infini ; on peut par exemple montrer assez aisément (démontrer deux inclusions) que

$$\bigcap_{n\in\mathbb{N}^*}\left[0,1+\frac{1}{n}\right]=[0,1]\ .$$

Proposition III.3. Soient A, B, C, D quatre ensembles. Alors :

(i) $A\cap B=B\cap A$ **[commutativité]** ;

(ii) $(A\cap B)\cap C=A\cap(B\cap C)$ **[associativité]** ;

(iii) $A\cap\emptyset=\emptyset$ **[$\emptyset$ est absorbant]** ;

(iv) $A\cap A=A$ **[idempotence]** ;

(v) $(A\cap B=A)\Leftrightarrow(A\subset B)$;

(vi) $(A\subset B)\wedge(C\subset D)\Rightarrow(A\cap C\subset B\cap D)$.

Démonstration. Immédiat en utilisant les propriétés du connecteur $\wedge$. Un dessin peut une fois encore être utile pour (v) et (vi). □

Proposition III.4. Soient A, B, C trois ensembles. Alors :

(i) $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$;

(ii) $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$.

Démonstration. Faisons le premier :

$$\begin{aligned}
x\in A\cap(B\cup C) &\Leftrightarrow (x\in A)\wedge(x\in B\cup C)\\
&\Leftrightarrow (x\in A)\wedge((x\in B)\vee(x\in C))\\
&\Leftrightarrow ((x\in A)\wedge(x\in B))\vee((x\in A)\wedge(x\in C)) \qquad \text{par distributivité}\\
&\Leftrightarrow (x\in A\cap B)\vee(x\in A\cap C)\\
&\Leftrightarrow (x\in A\cap B)\cup(A\cap C)\ .
\end{aligned}$$

□

c) Différence, complémentaire

Définition III.5. Soient A, B deux parties d'un ensemble E. On appelle :

— **différence de** A **et** B l'ensemble $A\setminus B=\{x\in A\,|\,x\notin B\}$;

— **complémentaire de** A l'ensemble $\overline{A}=\{x\in E\,|\,x\notin A\}$.

✌ **Remarque III.7.** Si A est une partie de E on a :

— $\overline{A}=E\setminus A.$;

— $A\setminus A=\emptyset$;

— $A \setminus \emptyset = A$.

➠ **Exemple III.5.** $\{1,2,3\} \setminus \{1,4,5\} = \{2,3\}$.

Nous disposons d'une version ensembliste des lois de de Morgan via la proposition suivante.

Proposition III.5. Soient A, B, C trois parties d'un ensemble E. Alors :

(i) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$;

(ii) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Démonstration. Utiliser les lois de de Morgan logique en traduisant l'appartenance aux ensembles étudiants en termes de connecteurs. □

d) Produit cartésien

Définition III.6. On appelle **couple** toute paire ordonnée d'objets. On dira que deux couples (a, b) et (c, d) sont égaux si et seulement si $(a = c) \wedge (b = d)$.

➠ **Exemple III.6.** $(\spadesuit, \hbar) \neq (\hbar, \spadesuit)$.

Définition III.7. Soient A, B deux ensembles. On appelle **produit cartésien de** A **et** B l'ensemble

$$\{(a, b) \mid a \in A,\ b \in B\}\ .$$

Notation. $A \times B$

➠ **Exemple III.7.** Vous connaissez sans doute déjà $\mathbb{R}^2$. Notons que l'on peut généraliser : $\mathbb{R}^3$ est par exemple le produit cartésien (associatif !) de $\mathbb{R}^2$ par $\mathbb{R}$.

Chapitre IV

Entiers naturels, récurrence(s)

1.⁻ L'ensemble $\mathbb{N}$

L'ensemble $\mathbb{N}$ des entiers naturels est un ensemble **infini**, muni d'une **addition** et d'une **multiplication**. Il s'agit de lois de composition internes : la somme et le produit de deux entiers naturels restent des entiers naturels. On peut également le munir de deux relations d'ordre, à savoir :

— l'**ordre naturel** : on notera $a \leq b$ (pour $a, b \in \mathbb{N}$) si $\exists c \in \mathbb{N}$, $b = a + c$; c est alors **noté** $b - a$ (la soustraction n'est pas une opération admissible sur $\mathbb{N}$). Il s'agit d'un ordre total : si $a, b \in \mathbb{N}$ alors $(a \leq b) \vee (b \leq a)$;

— la **divisibilité** : on notera $a|b$ ("a divise b") si $\exists c \in \mathbb{N}$, $b = ac$; c est alors **noté**, lorsque $a \neq 0$, $\frac{b}{a}$. Il ne s'agit pas d'un ordre total : $2 \not| 3$ et $3 \not| 2$.

Notation. Dans toute la suite, pour $n \leq m$, on posera $[\![n, m]\!] = \{n, n+1, \ldots, m\}$.

2.⁻ Récurrences

a) Le théorème

Axiome D. Toute partie non vide de $\mathbb{N}$ possède un plus petit élément (minimum).

Cet axiome, bien qu'utile en lui-même comme nous le verons par la suite, nous amène la conséquence suivante, moins "évidente".

Proposition IV.1. Toute partie de $\mathbb{N}$ non vide et majorée admet un plus grand élément (maximum).

✌ **Remarque IV.1.**

— Une partie A de $\mathbb{N}$ est dite **majorée** (resp. **minorée**) si $\exists M \in \mathbb{N}$ tel que $\forall x \in \mathbb{N}$, $x \leq M$ (resp. $x \geq M$). Nous reviendrons sur cette notion dans le chapitre V.

— Remarquons que toute partie non vide de $\mathbb{N}$ est minorée.

Démonstration. Soit $A \subset \mathbb{N}$ non vide et majorée. Posons :

$$B = \{x \in \mathbb{N} \,|\, \forall a \in A,\, a \leq x\}$$

l'ensemble des majorants de A. Alors :

— $B \neq \emptyset$ car A est majorée ;
— $B \subset \mathbb{N}$.
Ainsi, d'après l'axiome D, il existe $b \in B$ tel que $\forall x \in B$, $b \leq x$ ($b = \min B$). Remarquons de plus que b majore A par construction.

Cas 1 : $b = 0$. Alors $A = \{0\}$ car A est majorée par 0. La preuve est terminée.

Cas 2 : $b \neq 0$. Alors $b - 1 \in \mathbb{N}$ et, comme $b = \min B$, ne majore pas A. Ainsi, il existe $a \in A$ tel que $a > b - 1$; ce qui signifie que $b - 1 < a \leq b$. La seule possibilité est alors que $a = b$; b est alors un majorant de A appartenant à l'ensemble : $b = \max A$.

□

Nous pouvons désormais énoncer le **principe de succession**, propriété fondamentale de l'ensemble des entiers naturels.

Proposition IV.2. Soit $A \subset \mathbb{N}$ telle que :
— $0 \in A$;
— $\forall n \in \mathbb{N}$, $(n \in A) \Rightarrow (n + 1 \in A)$.
Alors $A = \mathbb{N}$.

Démonstration. Supposons $A \neq \mathbb{N}$, *i.e* il existe $n \in \mathbb{N} \setminus A$. Posons $B = \mathbb{N} \setminus A$; alors :
— $B \neq \emptyset$;
— $B \subset \mathbb{N}$.
De fait, par axiome D, il existe $n_0 = \min B$. Par définition, $n_0 - 1 \in A$ et donc $n_0 = (n_0 - 1) + 1 \in A$, ce qui est absurde. □

Corollaire IV.2.a (Principe de récurrence)**.**
Soit $\mathscr{P}$ un prédicat dépendant d'une variable $n \in \mathbb{N}$ tel que :
— $\mathscr{P}(0)$ est vraie **[initialisation]** ;
— $\forall n \in \mathbb{N}$, $\mathscr{P}(n) \Rightarrow \mathscr{P}(n + 1)$ **[hérédité]**.
Alors $\forall n \in \mathbb{N}$, $\mathscr{P}(n)$ est vraie.

Démonstration. Appliquer la proposition précédente à l'ensemble $A = \{n \in \mathbb{N} \mid \mathscr{P}(n)\}$. □

✌ **Remarque IV.2.**
— Si $n_0 > 0$ est entier naturel que $\mathscr{P}(n_0)$ est vérifiée, alors l'hérédité entraîne que $\forall n \geq n_0$, $\mathscr{P}(n)$ est vraie. Ceci permet de faire des démonstrations par récurrence initialisées à $n = 1, 2, 3$ ou 42.
— Toute rédaction de récurrence doit faire apparaître clairement initialisation et hérédité.
— Il est absolument **proscrit** de faire débuter une hérédité par "Supposons la propriété vraie pour tout n" ...

✍ **Exercice IV.1.** Soit $n \geq 2$; démontrer que 10 divise $2^{2^n} - 6$.

➥ **Correction :** *Par récurrence ...*

— *Pour* $n = 2$, $2^{2^n} - 6 = 10$.
— *Si la propriété est vraie au rang* n, *on remarque que* $2^{2^{n+1}} - 6 = \left(2^{2^n}\right)^2 - 6$. *Or, par hypothèse de récurrence, il existe* $c \in \mathbb{N}$ *tel que* $2^{2^n} = 10c + 6$. *En réinjectant, on tombe sur*

$$\begin{aligned} 2^{2^{n+1}} - 6 &= (10c+6)^2 - 6 \\ &= 100c^2 + 120c + 36 - 6 \\ &= 10(10c^2 + 12c + 3) \end{aligned}$$

d'où le résultat.

Interlude sommatoire

Fixons nous une famille de nombres réels (ou complexes) $a_0, \ldots a_n$. On notera :

$$\sum_{k=0}^{n} a_k = a_0 + \ldots + a_k \ .$$

Toutes les propriétés attendues pour une somme restent valables sous cette notation ; en particulier la "relation de Chasles"

$$\sum_{k=0}^{n} a_k = \sum_{k=0}^{m} a_k + \sum_{k=m+1}^{n} a_k$$

est vérifiée pour tout $m \in [\![0, n-1]\!]$ et exprime le fait que

$$a_0 + \ldots + a_m + a_{m+1} + \ldots + a_n = (a_0 + \ldots + a_m) + (a_{m+1} + \ldots + a_n) \ .$$

De la même façon, si λ est un nombre complexe, on a naturellement :

$$\lambda \sum_{k=0}^{n} a_k = \sum_{k=0}^{n} \lambda a_k \ .$$

Légèrement plus technique, nous pouvons faire des changements d'indice, comme :

$$\sum_{k=0}^{n} a_k = \sum_{k=1}^{n+1} a_{k-1} \ .$$

Le paragraphe suivant contient plusieurs exemples de manipulation de sommes.

b) Exemples fondamentaux

Proposition IV.3. Soit $n \in \mathbb{N}$. Alors :

(i)

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \ ;$$

(ii)

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \ .$$

Démonstration. Faisons le (i), *i.e* démontrons par récurrence sur $n \in \mathbb{N}$ que :

$$\forall n \in \mathbb{N}, \quad \mathscr{P}(n) \ : \sum_{k=0}^{n} k = \frac{n(n+1)}{2} \ \text{ est vraie.}$$

— $\mathscr{P}(0)$ est trivialement vérifiée.
— Supposons $\mathscr{P}(n)$ vérifiée pour un certain $n \in \mathbb{N}$. Alors :

$$\begin{aligned}
\sum_{k=0}^{n+1} k &= n + 1 + \sum_{k=0}^{n} k \\
&= n + 1 + \frac{n(n+1)}{2} \quad \text{par hypothèse de récurrence} \\
&= \frac{2(n+1) + n(n+1)}{2} \\
&= \frac{(n+1)(n+2)}{2}
\end{aligned}$$

d'où $\mathscr{P}(n+1)$.

La propriété était initialisée et héréditaire, elle est vérifée sur $\mathbb{N}$. Le point (ii) se traite de façon similaire. □

✍ **Exercice IV.2.** Démontrer que si $n \geq 0$ on a :

$$\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}.$$

Proposition IV.4. Soit $x \in \mathbb{C}$ et soit $n \in \mathbb{N}$. Alors :

$$\sum_{k=0}^{n} x^k = \begin{cases} n+1 & \text{si } x = 1 \\ \dfrac{1 - x^{n+1}}{1 - x} & \text{sinon.} \end{cases}$$

Démonstration. Le cas $x = 1$ est trivial. Dans le cas contraire, démontrons par récurrence sur $n \in \mathbb{N}$ que :

$$\forall n \in \mathbb{N}, \quad \mathscr{P}(n) \ : \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} \ \text{ est vraie.}$$

— $\mathscr{P}(0)$ est trivialement vérifiée.
— Supposons $\mathscr{P}(n)$ vérifiée pour un certain $n \in \mathbb{N}$. Alors :

$$\begin{aligned}
\sum_{k=0}^{n+1} x^k &= x^{n+1} + \sum_{k=0}^{n} x^k \\
&= x^{n+1} + \frac{1 - x^{n+1}}{1 - x} \quad \text{par hypothèse de récurrence} \\
&= \frac{(1-x)x^{n+1} + 1 - x^{n+1}}{1 - x} \\
&= \frac{1 - x^{n+2}}{1 - x}
\end{aligned}$$

d'où $\mathscr{P}(n+1)$.

La propriété était initialisée et héréditaire, elle est vérifée sur $\mathbb{N}$. □

Proposition IV.5. Soient $a, b \in \mathbb{C}$ et $n \in \mathbb{N}$. Alors :

$$a^{n+1} - b^{n+1} = (a-b)\sum_{k=0}^{n} a^k\, b^{n-k} .$$

Démonstration. Il suffit de vérifier que :

$$\begin{aligned}(a-b)\sum_{k=0}^{n} a^k b^{n-k} &= a\sum_{k=0}^{n} a^k b^{n-k} - b\sum_{k=0}^{n} a^k b^{n-k}\\ &= \sum_{k=0}^{n} a^{k+1} b^{n-k} - \sum_{k=0}^{n} a^k b^{n+1-k}\\ &= \sum_{k=1}^{n+1} a^k b^{n+1-k} - \sum_{k=0}^{n} a^k b^{n+1-k} \text{ par changement d'indice}\\ &= b^{n+1} - a^{n+1} .\end{aligned}$$

□

c) Binôme de Newton

Définition IV.1. Soit $n \in \mathbb{N}$; on appelle **factorielle** de n l'entier $n!$ défini par récurrence de la façon suivante :
- $0! = 1$;
- si $n \geq 1$, $n! = n \times (n-1)!$.

✌ **Remarque IV.3.** Il découle de la définition que

$$n! = n \times (n-1) \times \ldots \times 2 \times 1 \times 1 = \prod_{k=1}^{n} k .$$

➠ **Exemple IV.1.** $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $14! = 87178291200$, $42! = 1.405006117752879898543142606244511569936384.10^{51}$: la croissance est rapide !

Définition IV.2. Soient $n, k \in \mathbb{N}$. On appelle **coefficient du binôme "k parmi n"** la quantité :

$$\binom{n}{k} = \begin{cases} 0 & \text{si } n < k \\ \dfrac{n!}{k!(n-k)!} & \text{sinon.}\end{cases}$$

✌ **Remarque IV.4.** Cette quantité prendra tout son sens lorsque nous parlerons de combinatoire dans le chapitre XVIII. Pour le moment, nous pouvons admettre qu'il s'agit du nombre de façons de choisir k éléments parmi n, sans ordre.

Proposition IV.6. Soient $k, n \in \mathbb{N}$ tels que $k \leq n$. Alors :

(i)
$$\binom{n}{0} = \binom{n}{n} = 1\,,\ \binom{n}{1} = n\ ;$$

(ii)
$$\binom{n}{k} = \binom{n}{n-k}$$

(iii)
$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \qquad \textbf{[identité de Pascal]}$$

Démonstration. Immédiat en passant par la définition. □

✌ **Remarque IV.5.** Le point (iii)fait écho à une contruction probablement vue en terminale : le triangle de Pascal (Blaise de son prénom, 1623—1662).

$n\backslash k$	0	1	2	3	...
0	1	0	0	0	...
1	1	1	0	0	...
2	1	2	1	0	...
3	1	3	3	1	...
⋮	⋮	⋮	⋮	⋮	

Comme son nom ne l'indique pas, la formule qui suit apparaît dés le X^e siècle dans des traités de mathématiciens arabes, indiens ou perses (Al–Karaji, Halayudha). Elle fut également démontrée indépendamment par Yang Hui en Chine au XIII^e siècle Elle fut toutefois démontrée sous sa forme moderne et généralisée par Isaac Newton en 1665.

Théorème IV.7 (Formule du binôme de Newton)**.**
Soient $a, b \in \mathbb{C}$ et soit $n \geq 0$. Alors :

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

✌ **Remarque IV.6.** L'addition étant commutative, les rôles de a et b sont interchangeables. Mettre ceci en parallèle avec l'égalité (ii) de la proposition IV.6.

Démonstration. Devinez quoi... Une récurrence ! Plus précisément, démontrons par récurrence sur $n \in \mathbb{N}$ que :

$$\forall n \in \mathbb{N}, \quad \mathscr{P}(n)\ :\ (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k} \ \text{ est vraie.}$$

— $\mathscr{P}(0)$ est trivialement vérifiée.
— Supposons $\mathscr{P}(n)$ vérifiée pour un certain $n \in \mathbb{N}$. Alors :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1}\, b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k+1} \\
&= \sum_{k'=1}^{n+1} \binom{n}{k'-1} a^{k'}\, b^{n-k'+1} + \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k+1} \\
&\text{via le changement d'indice } k = k'-1 \text{ dans la première somme} \\
&= \sum_{=1}^{n+1} \binom{n}{k-1} a^k\, b^{n-k+1} + \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k+1} \\
&= \binom{n}{n} a^{n+1} + \binom{n}{0} b^{n+1} + \sum_{k=1}^{n} \underbrace{\left(\binom{n}{k-1} + \binom{n}{k}\right)}_{=\binom{n+1}{k}} a^k\, b^{n-k+1} \\
&= a^{n+1} + b^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^k\, b^{n-k+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k\, b^{n-k+1}
\end{aligned}$$

d'où $\mathscr{P}(n+1)$.
La propriété était initialisée et héréditaire, elle est vérifée sur $\mathbb{N}$. □

➠ **Exemple IV.2.** En appliquant cette formule, on obtient immédiatement les deux (forts utiles) formules suivantes, pour $n \in \mathbb{N}^*$:

$$\sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n$$

et

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = (1-1)^n = 0\ .$$

3.⁻ Récurrences "alternatives"

a) Récurrence forte

Proposition IV.8. Soit $\mathscr{P}$ un prédicat dépendant d'une variable $n \in \mathbb{N}$ tel que :
— $\mathscr{P}(0)$ est vraie **[initialisation]** ;
— $\forall n \in \mathbb{N}$, $(\forall k \in [\![0, n]\!],\ \mathscr{P}(k)) \Rightarrow \mathscr{P}(n+1)$ **[hérédité forte]**.
Alors $\forall n \in \mathbb{N}$, $\mathscr{P}(n)$ est vraie.

Démonstration. Appliquer le principe de récurrence au prédicat $\mathscr{Q}(n) : \forall k \leq n,\ \mathscr{P}(n)$. □

➠ **Exemple IV.3.** Démontrons par récurrence forte sur $n \in \mathbb{N}$ que :

$$\forall n \in \mathbb{N} \setminus \{0, 1\},\ \mathscr{P}(n)\ :\ n \text{ admet un diviseur premier.}$$

— $n = 2$ est divisible par 2, qui est premier.
— Soit $n \in \mathbb{N}$; supposons que $\mathscr{P}(k)$ soit vérifiée pour tout $k \in [\![2, n]\!]$. Si $n+1$ est premier, c'est terminé ; dans le cas contraire, il existe $a, b \in [\![2, n]\!]$ tels que $n + 1 = ab$. Or, par hypothèse de récurrence, a (par exemple) admet un diviseur premier, d'où le résultat.

b) Récurrence à deux rangs

Proposition IV.9. Soit $\mathscr{P}$ un prédicat dépendant d'une variable $n \in \mathbb{N}$ tel que :
— $\mathscr{P}(0)$ et $\mathscr{P}(1)$ sont vraies **[initialisation double]** ;
— $\forall n \in \mathbb{N}$, $(\mathscr{P}(n) \wedge \mathscr{P}(n+1)) \Rightarrow \mathscr{P}(n+2)$ **[hérédité à deux rangs]**.
Alors $\forall n \in \mathbb{N}$, $\mathscr{P}(n)$ est vraie.

✍ **Exercice IV.3.** On définit la **suite de Fibonacci** (Leonardo, 1175—1250) $(F_n)_n$ de la façon suivante :

$$F_0 = 0, \quad F_1 = 1 \quad \text{et} \quad \forall n \geq 0,\ F_{n+2} = F_n + F_{n+1}\ .$$

On note ϕ et $\overline{\phi}$ les racines du polynôme $X^2 - X - 1$. Démontrer que :

$$\forall n \geq 0, \quad F_n = \frac{\phi^n - \overline{\phi}^n}{\phi - \overline{\phi}}\ .$$

➥ **Correction :** *L'initialisation double est immédiate. Si on suppose la propriété vraies aux rangs n et $n+1$ pour un certain $n \geq 0$ alors*

$$\begin{aligned} F_{n+2} &= F_n + F_{n+1} \\ &= \frac{\phi^n - \overline{\phi}^n}{\phi - \overline{\phi}} + \frac{\phi^{n+1} - \overline{\phi}^{n+1}}{\phi - \overline{\phi}} \text{ par hypothèse de récurrence} \\ &= \frac{1}{\phi - \overline{\phi}} \left(\left(\phi^n + \phi^{n+1} \right) + \left(\overline{\phi}^n + \overline{\phi}^{n+1} \right) \right)\ . \end{aligned}$$

Il nous suffit pour conclure de remarquer que

$$\phi^n + \phi^{n+1} = \phi^n(1 + \phi) = \phi^n \phi^2 = \phi^{n+2}$$

et que l'on a un résultat analogue pour $\overline{\phi}$.

Chapitre V

Applications, relations

1.⁻ Applications

a) C'est quoi ?

Définition V.1. Soient E, F deux ensembles. On appelle **application de E dans F** la donnée de tout triplet $f = (E, F, \mathscr{G})$, où $\mathscr{G} \subset E \times F$ vérifie la propriété suivante :

$$\forall x \in E, \exists! y \in F \quad (x, y) \in \mathscr{G} \ .$$

Vocabulaire. L'ensemble E est appelé **ensemble de départ** de f, F son **ensemble d'arrivée** et $\mathscr{G}$ son **graphe**.

➠ **Exemple V.1.** Cette définition n'est pas extrêmement maniable ; par exemple, la fonction $f : x \mapsto x^2$ que vous connaissez tous deviendrait

$$f = \big(\mathbb{R}, \mathbb{R}_+, \{(x, x^2) \,|\, x \in \mathbb{R}\}\big) \ \ldots$$

✌ **Remarque V.1.** Deux applications seront donc considérées comme égales si elles ont même ensemble de départ, même ensemble d'arrivée et même graphe. Cela n'est pas sans soulever quelques problèmes concernant l'ensemble d'arrivée : celui de la fonction ci–dessus est–il $\mathbb{R}$ ou $\mathbb{R}_+$?

Définition V.2. Soit $f = (E, F, \mathscr{G})$ une application et soit $x \in E$. Alors l'unique $y \in E$ tel que $(x, y) \in \mathscr{G}$ est appelé **image** de x par f.

Notation. y est alors noté $f(x)$. Ceci nous autorise à définir désormais nos applications via la notation

$$\begin{aligned} f : E &\to F \\ x &\mapsto f(x) \ , \end{aligned}$$

le graphe $\mathscr{G}$ étant sous–entendu.

Exemple V.2. La fonction vue dans l'exemple précédent (re)devient de fait

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}_+ \\ x &\mapsto x^2 \,, \end{aligned}$$

ce qui est bien plus raisonnable.

Notation. L'ensemble des fonctions partant d'un ensemble E et arrivant dans un ensemble F sera noté $\mathscr{F}(E, F)$ ou F^E.

Remarque V.2. Mettons en exergue quelques cas particuliers :

— l'ensemble $E^{\mathbb{N}}$ est l'ensemble des fonctions de $\mathbb{N}$ dans E, que vous connaissiez déjà sous le nom de **suites** (si, si). On utilisera la notation u_n (pour $n \in \mathbb{N}$) pour $u(n)$ et $(u_n)_{n\in\mathbb{N}}$ (ou simplement $(u_n)_n$) pour u ;

— plus généralement, une **famille** d'éléments de E indexées par un ensemble I est en réalité une application de I dans E.

b) Restrictions, prolongements

L'objectif de ce paragraphe est d'étudier deux méthodes nous permettant de modifier l'ensemble de départ d'une application en agissant de façon "naturelle" sur son graphe.

Définition V.3. Soit $f : E \to F$ une application et soit $E' \subset E$. On appelle **restriction de f à E'** l'application

$$\begin{aligned} f_{|_{E'}} : E' &\to F \\ x &\mapsto f(x) \,. \end{aligned}$$

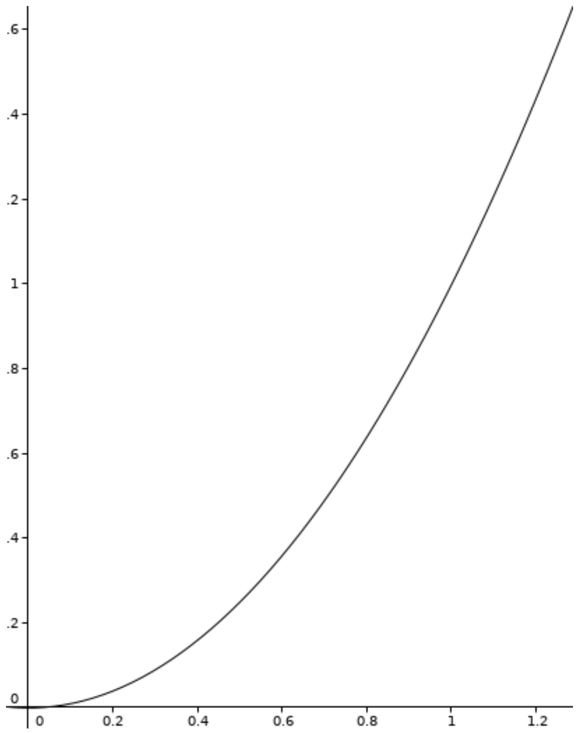
Remarque V.3.

— Il s'agit donc de l'application $(E', F, \mathscr{G}')$, avec (si $\mathscr{G}$ est le graphe de f)

$$\mathscr{G}' = \{(x, y) \in \mathscr{G} \,|\, x \in E'\} \,.$$

— La restriction de f à E' est unique.

Exemple V.3. La restriction de $x \mapsto x^2$ à $\mathbb{R}_+$ correspond au graphe ci–ensuite.



Définition V.4. Soit $f : E \to F$ une application et soit E' un ensemble contenant E. On appelle **prolongement de** f **à** E' toute application $g : E' \to F$ telle que $g_{|E} = f$.

✌ **Remarque V.4.** Il n'y a **PAS** unicité du prolongement de f à E'. Par exemple, la fonction

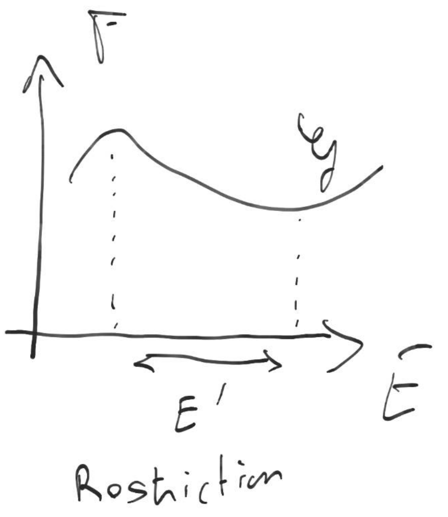
$$\begin{aligned} f : \mathbb{R}_+ &\to \mathbb{R}_+ \\ x &\mapsto x^2 \end{aligned}$$

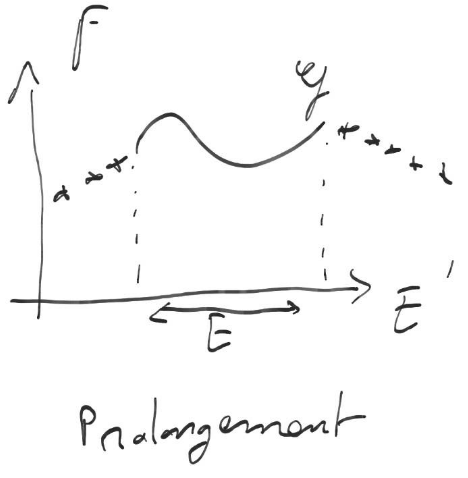
admet pour prolongements à $\mathbb{R}$ toutes les applications du type

$$\begin{aligned} f_\alpha : \mathbb{R} &\longrightarrow \mathbb{R}_+ \\ x &\mapsto \begin{cases} x^2 & \text{si } x \geq 0 \\ \alpha + 42 & \text{sinon} \end{cases} \end{aligned}$$

pour $\alpha \in \mathbb{R}_+$

Moralement, une restriction revient donc à "couper" le graphe pour n'en garder qu'une partie, alors qu'un prolongement est un extension du graphe par des points choisis aribitrairement.





c) Injections, surjections, bijections

Disons le d'entrée de jeu : ce paragraphe est **absolument fondamental**. Voyez le un peu un rhinocéros enragé courant dans votre direction : à ignorer à vos risques et périls.

Définition V.5. Soit $f : E \to F$ une application et soit $y \in F$. Un élément x de E est appelé **antécédent de y par f** si $f(x) = y$.

➠ **Exemple V.4.** Un point de l'ensemble de départ peut avoir n'importe quel nombre d'antécédent(s) par l'application : penser -1 pour $x \mapsto x^2$ (0 antécédent), 0 et 1 pour cette même fonction (1 et 2 antécédents respectivement), 14 pour la fonction constante égale à 14 (une infinité d'antécédents).

Définition V.6. Une application $f : E \to F$ est dite :
- **injective** si pour tout point F admettant un antécédent par f, ce dernier est unique ;
- **surjective** si tout point F admet au moins un antécédent par f ;
- **bijective** si tout point F admet exactement un antécédent par f.

✌ **Remarque V.5.**
- Une application est donc bijective si et seulement si elle est injective **et** bijective.
- Géométriquement, le caractère inj/surj/bijectif d'une application peut être observé *via* le nombre de points d'intersection entre sa courbe représentative et une droite horizontale.

Notation. L'ensemble des bijections de E dans F sera noté $\mathfrak{S}(E, F)$.

➠ **Exemple V.5.** "Par lecture graphique" (et rigoureusement, aussi), on peut remarquer que les fonctions périodiques type (co)sinus ne sont pas injectives, que $x \mapsto x$ est bijective et que $\exp : \mathbb{R} \to \mathbb{R}$ est injective non surjective.

Remarque V.6. Ces notions demandent de définir avec **précision et rigueur** les applications sur lesquelles nous travaillons. Par exemple, la fonction $x \mapsto x^2$ est :
- — surjective non injective de $\mathbb{R}$ dans $\mathbb{R}_+$;
- — injective non surjective de $\mathbb{R}_+$ dans $\mathbb{R}$;
- — bijective de $\mathbb{R}_+$ dans $\mathbb{R}_+$;
- — rien du tout de $\mathbb{R}$ dans $\mathbb{R}$.

Proposition V.1. Soit $f : E \to F$ une application. Alors :

$$f \text{ est injective}$$
$$\Longleftrightarrow$$
$$\forall x, x' \in E, \quad (f(x) = f(x')) \Rightarrow (x = x') .$$

Démonstration.

($\Uparrow$) Immédiat par définition d'antécédent.

($\Downarrow$) Soit $y \in F$ et soient x, x' deux antécédents de f. Alors $f(x) = f(x') = y$ donc $x = x'$. L'application f est donc injective.

□

Remarque V.7. Cela signifie que pour montrer que f est injective, il faut et il suffit de montrer que si x, x' sont tels que $f(x) = f(x')$ alors $x = x'$. **Cette méthode est à privilégier en pratique.**

Exemple V.6. Cette méthode permet aisément de démontrer que $x \mapsto x^3$ est injective sur $\mathbb{R}$ (passer à la racine cubique).

d) Composition

Définition V.7. Soient $f : E \to F$ et $g : F \to G$ deux applications. On appelle **composée** de g par f l'application

$$\begin{aligned} g \circ f : E &\longrightarrow G \\ x &\mapsto g(f(x)) . \end{aligned}$$

Remarque V.8.
- — **Attention au sens de composition :** la composée $g \circ f$ applique d'abord f, puis g.

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & F \arrow[r, "g"] & G \end{tikzcd}$$

— La composition des applications est associative ; la démonstration de ce résultat est un exercice aisé.
— La composition n'est **PAS** commutative ; regarder par exemple $f : x \mapsto x+1$ et $x \mapsto x^2$ (définies sur $\mathbb{R}$).

Définition V.8. Soit E un ensemble. On appelle **identité** de E l'application

$$\begin{aligned} \mathrm{id}_E : E &\to E \\ x &\mapsto x\ . \end{aligned}$$

✌ **Remarque V.9.** Pour toute application $f \in E^E$, on a $f \circ \mathrm{id}_E = \mathrm{id}_E \circ f = f$. On dit que id_E est un **élément neutre** pour la composition des applications.

Proposition V.2. La composée de deux injections (resp. surjections, bijections) est une injection (resp. surjection, bijection).

Démonstration. Immédiat via la définition. □

➠ **Exemple V.7.** Si f est une injection à valeurs réelles, la fonction $x \mapsto e^{f(x)}$ est injective.

e) Inverses directionnels, réciproque

✖ La proposition qui suit est **fondamentale**.

Proposition V.3. Soient E, F deux ensembles **non vides** et soit $f : E \to F$ une application. Alors :

(i) f est injective $\Leftrightarrow \exists\, g : F \to E$ telle que $g \circ f = \mathrm{id}_E$ **[inversibilité à gauche]** ;

(ii) f est surjective $\Leftrightarrow \exists\, g : F \to E$ telle que $f \circ g = \mathrm{id}_F$ **[inversibilité à droite]**.

Démonstration.

(i) **(⇒)** Supposons f injective et fixons un élément arbitraire $\mho \in E$. Nous pouvons alors définir g par le procédé suivant : pour tout $y \in F$, soit y admet un (unique) antécédent $x \in E$ et dans ce cas nous posons $g(y) = x$, soit ce n'est pas le cas et alors nous fixons $g(y) = \mho$.

(⇐) Supposons f inversible à gauche d'inverse $g : F \to E$ et soient $x, x' \in E$ tels que $f(x) = f(x')$. Alors, en composant cette égalité à gauche par g nous obtenons $g \circ f(x) = g \circ f(x')$, *i.e* $x = x'$. On en déduit que f est bien injective.

(ii) **(⇒)** Supposons f surjective. Alors, pour tout $y \in F$, il existe $x \in E$ tel que $y = f(x)$; il nous suffit de poser $g(y) = x$ pour avoir $f \circ g(y) = f(x) = y$.

(⇐) Supposons f inversible à droite d'inverse $g : F \to E$. Alors, pour tout $y \in F$, on a $f(g(y)) = y$; $g(y)$ est donc un antécédent de y, ce qui prouve la surjectivité de f.

□

Exemple V.8. La fonction $x \mapsto x^2$ définie sur $\mathbb{R}_+$ est injective car inversible à gauche d'inverse $x \mapsto \sqrt{|x|}$.

Proposition V.4. Soient E, F deux ensembles **non vides** et soit $f : E \to F$ une application. Alors :

(i) f est bijective $\Leftrightarrow$ f est inversible à gauche **et** à droite ;

(ii) dans ce cas, les inverses à gauche et à droite de f sont uniques et égaux. On les (l' ?) appelle réciproque(s) de f.

Notation. SI f est bijective, on note sa réciproque f^{-1}.

Remarque V.10. Au risque de me répéter, on n'utilise la notation f^{-1} **que** lorsque f est **bijective**.

Démonstration. Le point (i) est une conséquente immédiate du fait que la bijectivité équivaut à la combinaison injectivité–surjectivité. Pour le point (ii), notons g et g' deux inverses à gauche de f et h un inverse à droite de cette même application ; alors, comme $f \circ h = \mathrm{id}_F$, on a :

$$g \circ f \circ h = g' \circ f \circ h = h$$

ce qui donne, en simplifiant $g = g' = h$. □

Exemple V.9. L'application

$$\begin{aligned} f : \mathbb{N} &\to \mathbb{N}^* \\ n &\mapsto n+1 \end{aligned}$$

est bijective, de réciproque

$$\begin{aligned} g : \mathbb{N}^* &\to \mathbb{N} \\ n &\mapsto n-1 \end{aligned}$$

est bijective. Cet exemple est à rapprocher d'un problème attribué à David Hilbert (1862–1943) : "étant donné un hotel complet possèdant une infinité de chambres, comment accomoder un invité de plus ?".

Proposition V.5. Soient $f : E \to F$ et $g : F \to G$ deux bijections. Alors :

(i) f^{-1} est une bijection de F dans E, de réciproque $(f^{-1})^{-1} = f$;

(ii) $g \circ f$ est une bijection de E dans G, de réciproque

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} .$$

Démonstration. Comme les réciproques sont uniques, il nous suffit de vérifier que les applications proposées ci–dessus remplissent bien le contrat, ce qui est immédiat. □

f) Images directe et réciproque

Définition V.9. Soit $f : E \to F$ une application. Alors :

1. si $A \subset E$, on appelle **image (directe) de A par f** l'ensemble

$$f(A) = \{f(x) \mid x \in A\} ;$$

(ii) si $B \subset F$, on appelle **image réciproque de B par f** l'ensemble

$$f^{-1}(B) = \{x \in E \mid f(x) \in B\} .$$

Remarque V.11.

— Les esprits acérés auront remarqué que nous utilisons la notation $f^{-1}(B)$ alors que l'application n'est pas bijective … C'est triste, et cela signifie qu'il faudra être bien vigilant à ne pas confondre image directe par l'application réciproque (chose existant uniquement dans le cas bijectif) et image réciproque (machin existant quoi qu'il arrive). La seule bonne nouvelle dans tout cela est que si f est bijective, alors ces deux concepts coïncident. Il faut savoir se satisfaire de peu.
— Si $y \in F$, alors $f^{-1}(\{y\})$ est l'ensemble des antécédents de y.
— On remarque que f est surjective si et seulement si $f(E) = F$.

Exemple V.10.

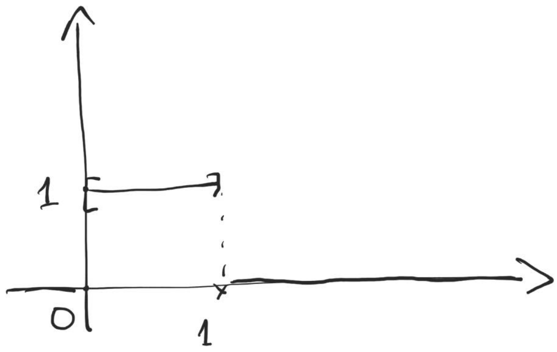
— Pour $f : x \mapsto x^2$, $f^{-1}([0, 4]) = [-2, 2]$;
— $\sin^{-1}(\{0\}) = \pi\mathbb{Z} = \{\pi k \mid k \in \mathbb{Z}\}$;
— $\exp(\mathbb{R}_+) = [1, \infty[$.

g) Indicatrices

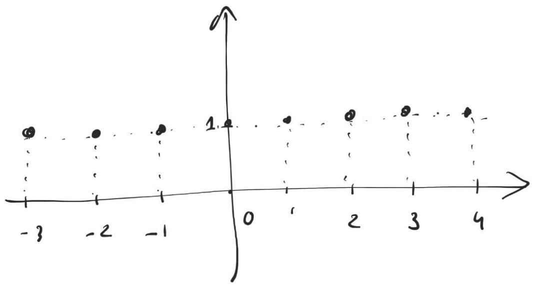
Définition V.10. Soit E un ensemble et soit $A \subset E$. On appelle **fonction indicatrice** (ou caractéristique) de A l'application

$$\mathbb{1}_A : E \longrightarrow \{0,1\}$$
$$x \mapsto \begin{cases} 0 & \text{si } x \notin A \\ 1 & \text{si } x \in A . \end{cases}$$

Exemple V.11. Pour $E = \mathbb{R}$, $\mathbb{1}_{\mathbb{R}}$ est la fonction constante égale à 1, $\mathbb{1}_{[0,1]}$ ressemble à ça :



et celle de $\mathbb{Z}$ à ceci.



Nous laissons l'entière responsabilité au lecteur et à son futur psychiatre d'imaginer l'indicatrice de $\mathbb{Q}$.

Proposition V.6. Soit E un ensemble et soient $A, B \in \mathscr{P}(E)$. Alors :

(i) $\mathbb{1}_{A\cap B} = \mathbb{1}_A \times \mathbb{1}_B$;

(ii) $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \times \mathbb{1}_B$;

(iii) $\mathbb{1}_{B\setminus A} = \mathbb{1}_B(\mathbb{1}_B - \mathbb{1}_A)$;

(iv) $\mathbb{1}_{\bar{A}} = 1 - \mathbb{1}_A$.

Démonstration.

(i) Soit $x \in E$; alors

$$\begin{aligned}\mathbb{1}_{A\cap B}(x) = 1 &\Leftrightarrow x \in A \cap B \\ &\Leftrightarrow (x \in A) \wedge (x \in B) \\ &\Leftrightarrow (\mathbb{1}_A(x) = 1) \wedge (\mathbb{1}_B(x) = 1)\,.\end{aligned}$$

De fait, on a bien $\mathbb{1}_{A\cap B} = \mathbb{1}_A \times \mathbb{1}_B$.

1. (ii)] De la même façon, on vérifie que, si $x \in E$:

$$\begin{aligned}\mathbb{1}_{A\cup B}(x) = 1 &\Leftrightarrow x \in A \cup B \\ &\Leftrightarrow (x \in A) \vee (x \in B) \\ &\Leftrightarrow (\mathbb{1}_A(x) = 1) \vee (\mathbb{1}_B(x) = 1)\,.\end{aligned}$$

Ici, il faut prendre garde au cas où $x \in A \cap B$, qui impose la formule suivante :

$$\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \times \mathbb{1}_B\,.$$

2. (iii)] Soit $x \in E$; alors

$$\begin{aligned} \mathbb{1}_{B\setminus A}(x) = 1 &\Leftrightarrow x \in B \setminus A \\ &\Leftrightarrow (x \in B) \wedge (x \notin A) \\ &\Leftrightarrow (\mathbb{1}_B(x) = 1) \wedge (\mathbb{1}_A(x) = 0) . \end{aligned}$$

   De fait, on a bien $\mathbb{1}_{B\setminus A} = \mathbb{1}_B(\mathbb{1}_B - \mathbb{1}_A)$.
3. (iv)] En appliquant la question précédente à $B = E$, on trouve $\mathbb{1}_{\bar{A}} = 1 - \mathbb{1}_A$.

□

2.⁻ Relations d'ordre

a) Relations binaires

Définition V.11. Soit E un ensemble. On appelle **relation** (binaire) sur E la donnée d'un couple $\mathscr{R} = (E, \mathscr{G})$, avec $\mathscr{G} \subset E \times E$.

Notation. Si $(x, y) \in \mathscr{G}$, on notera $x\mathscr{R}y$ et on dira que x **et** y **sont en relation via** $\mathscr{R}$.

Exemple V.12.
— $\mathscr{R}_0 = (\mathbb{R}, \{(x, -x) \,|\, x \in \mathbb{R}\})$ correspond à la relation "être égal au signe près" ;
— $\mathscr{R}_1 = (\mathbb{N}, \{(n, 2n) \,|\, n \in \mathbb{N}\})$ correspond à la relation "être le double de".

b) Ensembles ordonnés

Définition V.12. Soit E un ensemble et $\mathscr{R}$ une relation sur E. On dit que $\mathscr{R}$ est une **relation d'ordre** si :
— $\forall x \in E$, $x\mathscr{R}x$ **[reflexivité]** ;
— $\forall x, y \in E$, $(x\mathscr{R}y) \wedge (y\mathscr{R}x) \Rightarrow (x = y)$ **[antisymétrie]** ;
— $\forall x, y, z \in E$, $(x\mathscr{R}y) \wedge (y\mathscr{R}z) \Rightarrow (x\mathscr{R}z)$ **[transitivité]**.
La donnée du couple $(E, \mathscr{R})$ est appelée **ensemble ordonné**.

Exemple V.13.
— $(\mathbb{R}, \leq)$ est un ensemble ordonné, avec $\leq$ la relation $(\mathbb{R}, \{(x, y) \,|\, y - x \in \mathbb{R}_+\})$;
— $(\mathscr{P}(E), \subset)$ est un ensemble ordonné ;
— $(\mathbb{N}, |\,)$ est un ensemble ordonné ;
— si E et F sont ordonnés, F^E est ordonné (*cf.* chapitre I).

Définition V.13. Soit $(E, \mathscr{R})$ un ensemble ordonné. On dit que $\mathscr{R}$ est un **ordre total** si

$$\forall x, y \in E, \quad (x\mathscr{R}y) \vee (y\mathscr{R}x) .$$

Dans le cas contraire, on parle d'**ordre partiel**.

Exemple V.14. L'ordre classique sur $\mathbb{N}$ est total ; l'inclusion sur $\mathscr{P}(E)$ est partiel dès que E possède au moins deux éléments car $\mathscr{P}(E)$ contient alors deux singletons disjoints.

c) Majorants, minorants

Définition V.14. Soit $(E, \leq)$ un ensemble ordonné et soit $A \subset E$. On dit que $x \in E$ est un…
- …**majorant de** A si $\forall a \in A$, $a \leq x$;
- …**minorant de** A si $\forall a \in A$, $a \geq x$.

Un majorant (resp. minorant) de A appartenant à A est appelé **plus grand élément** (resp. **plus petit élément**) ou maximum (resp. minimum) de A.

Notation. $\max(A)$, $\min(A)$.

Remarque V.12. On démontre aisément que plus petit et plus grand élément sont uniques lorsqu'ils existent.

Exemple V.15.
- $(\mathbb{R}, \leq)$ n'est pas minoré ni majoré ;
- $([0,1[, \leq)$ est majoré par 1 qui n'est pas son maximum ;
- $(\mathbb{N}, |\,)$ admet pour minimum 1 et maximum 0 (en effet, tout entier n divise 0 car $0 = n \times 0$) ;
- $(\mathscr{P}(E), \subset)$ est minoré par $\emptyset$ et majoré par E. Ces deux derniers sont respectivement le plus petit et plus grand élément de $\mathscr{P}(E)$.

Remarque V.13. Merci de **ne pas confondre** minorant (resp. majorant) et minimum (resp. maximum). Penser à $]0,1[$ muni de l'ordre classique en cas de crise de foi.

Vocabulaire. Un ensemble majoré et minoré est dit **borné**.

3.⁻ Relations d'équivalence

a) C'est quoi ?

Définition V.15. Soit E un ensemble et $\mathscr{R}$ une relation sur E. On dit que $\mathscr{R}$ est une **relation d'équivalence** si :
- $\forall x \in E$, $x\mathscr{R}x$ **[reflexivité]** ;
- $\forall x, y \in E$, $(x\mathscr{R}y) \Rightarrow (y\mathscr{R}x)$ **[symétrie]** ;
- $\forall x, y, z \in E$, $(x\mathscr{R}y) \wedge (y\mathscr{R}z) \Rightarrow (x\mathscr{R}z)$ **[transitivité]**.

Exemple V.16.
- L'égalité est une relation d'équivalence sur tout ensemble.
- La relation triviale (E, E^2) est une relation d'équivalence sur tout ensemble E.

✍ **Exercice V.1.** Soit $f : E \to F$ une application. Démontrer que la relation $\mathscr{R}$ définie sur E par

$$x \mathscr{R} y \iff f(x) = f(y)$$

est une relation d'équivalence.

➥ **Correction :** *Un sage a dit : "Yaka écrire", et il avait raison.*

b) Congruences

◇ **Sur $\mathbb{Z}$**

Définition V.16. Soient $a, b, n \in \mathbb{Z}$. On dit que a **est congru à** b **modulo** n si il existe $k \in \mathbb{Z}$ tel que $b = a + kn$.

Notation. $a \equiv b\,[n]$.

➠ **Exemple V.17.** $3 \equiv 1\,[2]$, $9 \equiv -1\,[10]$.

✌ **Remarque V.14.**
- — La congruence modulo 0 est l'égalité.
- — La congruence modulo 1 est la relation triviale.
- — Remplacer n par $-n$ ne change absolument rien. On peut donc supposer $n \in \mathbb{N}$ sans perdre de généralité.

Proposition V.7. Soit $n \in \mathbb{N}$; alors la congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$.

Démonstration.
- — La reflexivité est immédiate.
- — Soient $a, b \in \mathbb{Z}$ tels que $a \equiv b\,[n]$. Alors il existe $k \in \mathbb{Z}$ tel que $b = a + kn$, ce qui entraîne $a = b + (-k)n$ et donc $b \equiv a\,[n]$, d'où la symétrie.
- — Soient $a, b, c \in \mathbb{Z}$ tels que $a \equiv b\,[n]$ et $b \equiv c\,[n]$. Ceci signifie qu'il existe $k, \ell \in \mathbb{Z}$ tels que $b = a + kn$ et $c = b + \ell n$; ainsi

$$c = b + \ell n = a + kn + \ell n = a + (k + \ell)n$$

  et donc $a \equiv c\,[n]$, d'où la transitivité.

□

◇ **Sur $\mathbb{R}$**

Définition V.17. Soient $a, b, \alpha \in \mathbb{R}$. On dit que a **est congru à** b **modulo** α si il existe $k \in \mathbb{Z}$ tel que $b = a + kn$.

Notation. $a \equiv b\,[\alpha]$.

✌ **Remarque V.15.**
— Remarquons que le "k" de la définition est **entier**.
— Les relations de congruence modulo π et 2π sont chères aux physiciens et utiles en trigonométrie. Nous y reviendrons.

Proposition V.8. Soit $\alpha \in \mathbb{R}$; alors la congruence modulo α est une relation d'équivalence sur $\mathbb{R}$.

Démonstration. Identique au cas entier. □

c) Classes d'équivalences

Définition V.18. Soit E un ensemble muni d'une relation d'équivalence $\mathscr{R}$ et soit $x \in E$. On appelle **classe de x modulo $\mathscr{R}$** l'ensemble

$$\overline{x} = \{y \in E \,|\, x\mathscr{R}y\} .$$

Notation. On notera $E/\mathscr{R}$ l'ensemble $\{\overline{x} \,|\, x \in E\}$. Cet ensemble est appelé **quotient de E par la relation $\mathscr{R}$**.

Proposition V.9. Soit E un ensemble muni d'une relation d'équivalence $\mathscr{R}$. Alors :
(i) $\forall C \in E/\mathscr{R}$, $C \neq \emptyset$;
(ii) si $C, C' \in E/\mathscr{R}$ sont telles que $C \neq C'$, alors $C \cap C' = \emptyset$;
(iii)
$$\bigcup_{C \in E/\mathscr{R}} C = E .$$

On dit que les classes d'équivalence modulo $\mathscr{R}$ forment une **partition** de l'ensemble E.

Démonstration. En exercice ; découle de la définition de relation d'équivalence. □

Cette notion sera approfondie plus tard (principalement en MP) autour de l'exemple suivant : si $n \in \mathbb{N}^*$, il existe n classes d'équivalences disjointes pour la congruence modulo n sur $\mathbb{Z}$: $\overline{0}, \overline{1}, \dots, \overline{n-1}$. L'ensemble quotient est alors noté $\mathbb{Z}/n\mathbb{Z}$.

Chapitre VI

Nombres réels

1.⁻ Le corps $\mathbb{R}$ des nombres réels

a) C'est quoi ?

L'ensemble $\mathbb{R}$ des nombres réels est muni d'une addition, d'une multiplication et d'une division (par des nombres non nuls). Un tel ensemble (en gros, *cf.* chapitre X) est appelé un **corps**.

b) Sous–ensembles remarquables

Théorème VI.1.
Il existe deux parties de $\mathbb{R}$, notées $\mathbb{R}_+$ et $\mathbb{R}_-$, telles que :
- (A) $\mathbb{R} = \mathbb{R}_+ \cup \mathbb{R}_-$;
- (B) $\mathbb{R}_+ \cap \mathbb{R}_- = \{0\}$;
- (C) $\mathbb{R}_+$ est stable par addition et multiplication ;
- (D) pour tout $x \in \mathbb{R}$, $(x \in \mathbb{R}_+) \Leftrightarrow (-x \in \mathbb{R}_-)$.

Démonstration. Nous ne disposons pas, hélas, de l'outillage technique requis pour aborder cette démonstration. Ce résultat sera donc admis. □

✌ **Remarque VI.1.**
- — Si on pose $\mathbb{R}^*_\pm = \mathbb{R}_\pm \setminus \{0\}$, alors $\mathbb{R}$ est égal à la réunion disjointe $\mathbb{R}^*_+ \sqcup \{0\} \sqcup \mathbb{R}^*_-$.
- — $\mathbb{R}_-$ est stable par addition ; en effet, si $x, y \in \mathbb{R}_-$ alors $x+y = -\underbrace{\left((-x)+(-y)\right)}_{\in \mathbb{R}_+} \in \mathbb{R}_-$.

Corollaire VI.1.a. Soient $x, y \in \mathbb{R}$. Alors :

(i) $(x \in \mathbb{R}_+) \wedge (y \in \mathbb{R}_-) \Rightarrow (xy \in \mathbb{R}_-)$;

(ii) $(x \in \mathbb{R}_-) \wedge (y \in \mathbb{R}_-) \Rightarrow (xy \in \mathbb{R}_+)$;

(iii) $x^2 \in \mathbb{R}_+$;

(iv) $1 \in \mathbb{R}_+$ et $-1 \in \mathbb{R}_-$;

(v) $(x \in \mathbb{R}_+^*) \Leftrightarrow \left(\frac{1}{x} \in \mathbb{R}_+^*\right)$.

Démonstration.

(i) Si $x \in \mathbb{R}_+$ et $y \in \mathbb{R}_-$ alors $xy = -\underbrace{x(-y)}_{\in \mathbb{R}_+} \in \mathbb{R}_-$.

(ii) Si $x, y \in \mathbb{R}_-$ alors $-x, -y \in \mathbb{R}_+$ et donc $xy = (-x)(-y) \in \mathbb{R}_+$.

(iii) Cela découle du (ii) et du (C) du théorème.

(iv) $1 = 1^2 \in \mathbb{R}_+$ donc $-1 \in \mathbb{R}_-$.

(v) Soit $x \in \mathbb{R}_+^*$; alors si $\frac{1}{x} \in \mathbb{R}_-$ nous aurions $1 = x \times \frac{1}{x} \in \mathbb{R}_-$, ce qui contedit le point (iv).

□

c) Ordre naturel sur $\mathbb{R}$

Définition VI.1. Soient $x, y \in \mathbb{R}$. On note :

— $x \leq y$ si $y - x \in \mathbb{R}_+$;

— $x < y$ si $y - x \in \mathbb{R}_+^*$.

Exemple VI.1. $3 < \pi$.

Proposition VI.2. La relation "$\leq$" est un ordre total sur $\mathbb{R}$.

Démonstration. Démontrons tout d'abord que nous avons affaire à une relation d'ordre.

Reflexivité. $\forall x \in \mathbb{R}$, $x - x = 0 \in \mathbb{R}_+$ donc $x \leq x$.

Antinsymétrie. Soit $x, y \in \mathbb{R}$ tels que $(x \leq y) \wedge (y \leq x)$. Alors $x - y \in \mathbb{R}_+$ et, comme $-(x-y) = y - x \in \mathbb{R}_+$, $x - y \in \mathbb{R}_+ \cap \mathbb{R}_- = \{0\}$. Au final, $x = y$.

Transitivité. Soient $x, y, z \in \mathbb{R}$ tels $x \leq y$ et $y \leq z$. Alors :

$$z - x = (z - y) + (y - x) \in \mathbb{R}_+$$

d'où $x \leq z$.

Enfin, pour vérifier que "$\leq$" est un ordre total, fixons $x, y \in \mathbb{R}$ et remarquons que $y - x \in \mathbb{R} = \mathbb{R}_+ \cup \mathbb{R}_-$. Ainsi, $(y - x \in \mathbb{R}_+) \vee (y - x \in \mathbb{R}_-)$, *i.e* $(x \leq y) \vee (y \leq x)$. □

✌ **Remarque VI.2.** Notons au passage que "<" ne risque pas d'être un ordre ; cette relation n'est pas réflexive !

Proposition VI.3. Soient $x, y, z \in \mathbb{R}$. Alors :

(i) $(x < y) \Leftrightarrow (x + z < y + z)$;

(ii) si $z \in \mathbb{R}_+^*$, $(x < y) \Leftrightarrow (xz < yz)$;

(iii) si $z \in \mathbb{R}_-^*$, $(x < y) \Leftrightarrow (xz > yz)$.

Démonstration.

(i)

$$\begin{aligned} x < y &\Longleftrightarrow y - x \in \mathbb{R}_+^* \\ &\Longleftrightarrow (y + z) - (x + z) = y - x \in \mathbb{R}_+^* \\ &\Longleftrightarrow x + z < y + z \ . \end{aligned}$$

(ii) Si $z > 0$ alors :

$$\begin{aligned} x < y &\Longleftrightarrow y - x \in \mathbb{R}_+^* \\ &\Longleftrightarrow z(y - x) = zy - zx \in \mathbb{R}_+^* \\ &\Longleftrightarrow xz < yz \ . \end{aligned}$$

(iii) Si $z < 0$ alors :

$$\begin{aligned} x < y &\Longleftrightarrow y - x \in \mathbb{R}_+^* \\ &\Longleftrightarrow z(y - x) = zy - zx \in \mathbb{R}_-^* \\ &\Longleftrightarrow xz > yz \ . \end{aligned}$$

□

Proposition VI.4. Soient $x, x', y, y' \in \mathbb{R}$ tels que $x \leq x'$ et $y \leq y'$. Alors :

$$x + y \leq x' + y' \ .$$

Démonstration. Il suffit de remarquer que $x' + y' - (x + y) = (x' - x) + (y' - y) \in \mathbb{R}_+$. □

✖ **Attention :** si on peut sommer des inégalités, il est **absolument exclus de les soustraire**. En effet $0 \leq 1$ et $-1 \leq 1$ mais, aux dernières nouvelles, $1 \not\leq 0$ ...

d) Valeur absolue

Définition VI.2. Soit $x \in \mathbb{R}$; on appelle **valeur absolue** de x le réel $\max(x, -x)$.

Notation. $|x|$

✌ **Remarque VI.3.** Notons que $x \in \mathbb{R}_+ \Leftrightarrow |x| = x$ et $x \in \mathbb{R}_- \Leftrightarrow |x| = -x$.

Proposition VI.5. Soient $x, y \in \mathbb{R}$. Alors :
(i) $|x| \in \mathbb{R}_+$;
(ii) $|x| = 0 \Leftrightarrow x = 0$;
(iii) $|xy| = |x||y|$;
(iv) $|x + y| \leq |x| + |y|$ **[inégalité triangulaire 1]** ;
(v) $||x| - |y|| \leq |x - y|$ **[inégalité triangulaire 2]**.

Démonstration.

(i), (ii) Conséquences immédiates de la remarque VI.3.

(iii) Laissé en exercice au lecteur ; procéder par disjonction de cas selon les signes respectifs de x et y et utiliser (pour changer) la remarque VI.3.

(iv) **Point méthode : comment gérer un maximum.**
Par définition, $|x + y| = \max(x + y, -x - y)$, or :
— $x + y \leq |x| + |y|$;
— $-x - y = (-x) + (-y) \leq |x| + |y|$.
Ainsi, $|x + y| = \max(x + y, -x - y) \leq |x| + |y|$.

(v) Échanger les rôles de x et y ne change rien à l'expression souhaitée : nous pouvons donc supposer, quitte à échanger x et y, que $|x| \geq |y|$. Remarquons alors que $x = x - y + y$; certes, me direz vous... Mais, en utilisant la première inégalité triangulaire, on obtient :

$$|x| = |x - y + y| \leq |x - y| + |y|$$

et donc $|x| - |y| \leq |x - y|$. Comme $|x| \geq |y|$, on a bien :

$$||x| - |y|| = |x| - |y| \leq |x - y| \ .$$

□

e) Parties positive, négative

Nous terminons ce paragraphe par un petit laïus dont nous ferons usage dans le chapitre XXI (et oui, il faut savoir être prévoyant !).

Définition VI.3. Soit $x \in \mathbb{R}$. On appelle :
- **partie positive** de x le réel
$$x^+ = \begin{cases} x & \text{si } x \in \mathbb{R}_+^* \\ 0 & \text{sinon} \end{cases} ;$$
- **partie négative** de x le réel
$$x^- = \begin{cases} -x & \text{si } x \in \mathbb{R}_-^* \\ 0 & \text{sinon} \end{cases} .$$

Exemple VI.2. $2^+ = 2$, $(-15)^- = 15$.

Remarque VI.4. Si $x \in \mathbb{R}$, alors $x = x^+ - x^-$ et $|x| = x^+ + x^-$.

2.⁻ Bornes supérieure, inférieure

a) C'est quoi ?

Définition VI.4. Soit $(E, \leq)$ un ensemble ordonné et soit $A \subset E$. On appelle **borne supérieure** (ou supremum) de A le plus petit élément (si il existe) de l'ensemble des majorants de A.

Notation. $\sup(A)$

En résumé, la borne supérieure est le **plus petit majorant de l'ensemble** A, lorsque ce dernier existe, *i.e* :
$$\sup(A) = \min\{x \in E \,|\, \forall a \in A,\, a \leq x\} .$$
On déduit de cette reformulation une caractérisation de la borne supérieure, à privilégier en pratique : pour tout $M \in E$,
$$M = \sup(A) \iff \begin{cases} \forall x \in A,\, x \leq M \\ \forall y < M,\, \exists\, x \in A\,,\; y < x \end{cases} . \qquad (\mathbf{E}\text{:VI.1})$$

Exemple VI.3. La borne supérieure de $[0,1[$ dans $\mathbb{R}$ est 1. En effet :
- pour tout $x \in [0,1[$, $x \leq 1$ donc 1 majore bien $[0,1[$;
- si $z < 1$, alors soit $z \leq 0$ et dans ce cas $\dfrac{1}{42} > z$ soit $z \in [0,1[$ et alors $\dfrac{z}{2} \in [0,1[$ et $\dfrac{z}{2} \leq z$. 1 est donc le plus petit majorant de $[0,1[$.

On peut définir de façon analogue le concept de borne inférieure d'un ensemble, qui sera alors son **plus grand minorant.**

Définition VI.5. Soit $(E, \leq)$ un ensemble ordonné et soit $A \subset E$. On appelle **borne inférieure** (ou infimum) de A le plus grand élément (si il existe) de l'ensemble des minorants de A.

Notation. $\inf(A)$

✖ **Attention :** l'existence de telles quantités est très loin d'être automatique : en effet, quelle pourrait être la borne supérieure de $\mathbb{Z}$ dans $\mathbb{R}$?

Dans toute la suite de ce chapitre, lorsque nous énoncerons un résultat relatif aux bornes supérieures, nous laissons le soin au lecteur d'en déduire un énoncé analogue s'appliquant aux bornes inférieures. Par exemple, la proposition suivante a pour pendant : lorsqu'un minimum existe, la borne inférieure existe et lui est égale.

Proposition VI.6. Soit $A \subset E$ une partie d'un ensemble ordonné **possèdant un plus grand élément**, alors :
- A admet une borne supérieure ;
- $\sup(A) = \max(A)$.

✌ **Remarque VI.5.** La réciproque est **FAUSSE** : penser à $[0, 1[$ qui admet une borne supérieure mais pas de maximum.

Démonstration. Posons $a = \max(A)$ et $A^+ = \{x \in E \,|\, \forall y \in A,\ y \leq x\}$ l'ensemble des majorants de A. Alors :
- il est clair que $a \in A^+$;
- $a \in A$ donc $\forall x \in A^+,\ x \geq a$.

Par caractérisation de la borne supérieure (**E** :VI.1), on a bien $a = \sup(A)$. □

✌ **Remarque VI.6.** Si A possède une borne supérieure, alors A est automatiquement non vide et majorée. La réciproque est, hélas fausse dans le cas général. En effet, posons $E = \mathbb{Q}$ et $A = \{x \in \mathbb{Q} \,|\, x^2 \leq 2\}$; alors, même si A est non vide ($1 \in A$) et majorée (par 3), elle n'admet pas de borne supérieure (il est possible, mais long et trop technique à ce stade du cours, de démontrer que cette dernière devrait avoir 2 pour carré, et donc être irrationnel).

b) Théorème fondamental

Même si la démonstration du théorème suivant dépasse le cadre du programme de MPSI, le résultat en est absolument remarquable et formera la pierre angulaire de la suite de ce chapitre.

Théorème VI.7.
Toute partie de $\mathbb{R}$ non vide et majorée admet une borne supérieure.

➠ **Exemple VI.4.** Si on voit la partie $A = \{x \in \mathbb{Q} \,|\, x^2 \leq 2\}$ comme un sous–ensemble de $\mathbb{R}$, il admet une borne supérieure, qui est $\sqrt{2}$.

Corollaire VI.7.a (Existence et unicité de la racine carrée positive)**.** Soit $x \in \mathbb{R}_+$. Alors il existe un unique $y \in \mathbb{R}_+$ tel que $y^2 = x$.

Afin de faciliter la vie de notre estimé lecteur, nous profitons de ce paragraphe pour anticiper sur un résultat qui sera démontré dans le chapitre IX.

Proposition VI.8. [Caractérisation séquentielle de la borne supérieure] Soit $A \subset \mathbb{R}$ et soit $M \in \mathbb{R}$; alors :

$$M = \sup(A) \Longleftrightarrow \begin{cases} \forall a \in A,\ a \leq M \\ \exists (a_n)_n \in A^{\mathbb{N}},\ a_n \xrightarrow[n\to\infty]{} M \end{cases} .$$

➠ **Exemple VI.5.** 0 est la borne inférieure de $\mathbb{Q}_+^*$ dans $\mathbb{R}$; en effet, il est clair qu'il minore cet ensemble et

$$\underbrace{\frac{1}{n}}_{\in\mathbb{Q}_+^*} \xrightarrow[n\to\infty]{} 0 .$$

c) Droite réelle achevée

Débutons ce paragraphe par fixer deux symboles, $-\infty$ et $+\infty$, sans signification particulière et posons :

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\} .$$

Cet ensemble est appelé **droite réelle achevée.**

Nous pouvons étendre l'ordre naturel de $\mathbb{R}$ à $\overline{\mathbb{R}}$ via les conventions suivantes :
— $-\infty < +\infty$;
— si $y \in \mathbb{R}$, alors $-\infty < y < +\infty$.

On en déduit en particulier que $\overline{\mathbb{R}}$ admet un maximum et un minimum, respectivement égaux à $+\infty$ et $-\infty$.

De même, les deux opérations classiques sur $\mathbb{R}$ peuvent (presque) être généralisées à $\mathbb{R}$ via les tables partielles suivantes.

$+$	$-\infty$	$y \in \mathbb{R}$	$+\infty$
$-\infty$	$-\infty$	$-\infty$	**?**
$x \in \mathbb{R}$	$-\infty$	$x+y$	$+\infty$
$+\infty$	**?**	$+\infty$	$+\infty$

$\times$	$-\infty$	$y \in \mathbb{R}_+^*$	0	$y \in \mathbb{R}_+^*$	$+\infty$
$-\infty$	$+\infty$	$+\infty$	**?**	$-\infty$	$-\infty$
$x \in \mathbb{R}_-^*$	$+\infty$	xy	0	xy	$-\infty$
0	**?**	0	0	0	**?**
$x \in \mathbb{R}_+^*$	$-\infty$	xy	0	xy	$+\infty$
$+\infty$	$-\infty$	$-\infty$	**?**	$+\infty$	$+\infty$

Notons que ces tables sont symétriques : les opérations étendues demeurent donc commutatives. Ces conventions nous seront, comme vous pouvez sans doute l'imaginer, très utiles pour nos futures études de limites.

✌ **Remarque VI.7.** Il serait possible sans grand risque de convenir du fait que $\dfrac{1}{\pm\infty} = 0$. Il n'est toutefois pas possible de convenir d'une valeur cohérente pour $\dfrac{1}{0}$ (quel serait son signe ?).

Proposition VI.9. Toute partie de $\overline{\mathbb{R}}$ admet une borne supérieure et une borne inférieure.

Démonstration. Soit $A \subset \overline{\mathbb{R}}$. Démontrons l'existence de la borne supérieure de A ; une démonstration similaire livrera alors celle de sa borne inférieure.

Cas 1 : $A = \emptyset$ ou $A = \{-\infty\}$. Alors dans ce cas l'ensemble des majorants de A est $\overline{\mathbb{R}}$ entier, ce qui entraîne que A admet pour borne supérieure $-\infty = \min(\overline{\mathbb{R}})$.

Cas 2 : $+\infty \in A$. Dans ce cas, $\sup(A) = \max(A) = +\infty$.

Cas 3 : $B = A \setminus \{-\infty\}$ est un sous ensemble non vide $\mathbb{R}$. Si l'ensemble B est majoré, d'après le théorème VI.7, il admet une borne supérieure qui sera également celle de A. Dans le cas contraire, $\sup(A) = +\infty$.

□

✌ **Remarque VI.8.** Une conséquence intéressante de tout ceci est que, si $A \subset \overline{\mathbb{R}}$ est **non vide**, alors :

$$\inf(A) \leq \sup(A) .$$

✖ **Attention :** l'inégalité est inversée dans le cas de l'ensemble vide, dont la borne supérieure est $-\infty$ et la borne inférieure $+\infty$.

3.⁻ Quelques résultats de topologie

a) Propriété d'Archimède

Comme ni son nom ni sa postérité ne l'indiquent, le théorème suivant se trouve dans le livre V des *Éléments* d'Euclide (∼300 av. J.-C.) bien qu'étant traditionnellement associé à Archimède (287 av. J.-C. — 212 av. J.-C.), qui en attribuait la paternité à Eudoxe de Cnide (408 av. J.-C. — 355 av. J.-C.). Sa formulation moderne est due à David Hilbert (1862—1943).

Théorème VI.10 (Archimède)**.**
Soient $a, b \in \mathbb{R}_+$ tels que $b \neq 0$. Alors :

$$\exists\, n \in \mathbb{N},\ a < nb .$$

Démonstration. Procédons par l'absurde en supposant que $\forall n \in \mathbb{N},\ a \geq nb$ et posons

$$\mathscr{E} = \{nb \,|\, n \in \mathbb{N}\} \subset \mathbb{R} .$$

Il s'agit d'une partie de $\mathbb{R}$ non vide et majorée : elle admet donc par théorème VI.7 une borne supérieure x_0.

Par définition de borne supérieure, $x_0 - b$ ne peut majorer $\mathscr{E}$ car il est strictement inférieur à x_0. Ainsi, il existe $n \in \mathbb{N}$ tel que $nb > x_0 - b$ *i.e* :

$$\underbrace{(n+1)b}_{\in \mathscr{E}} > \underbrace{x_0}_{=\sup \mathscr{E}}$$

ce qui est absurde. □

Corollaire VI.10.a. Soient $a, b \in \mathbb{R}$ tels que $b > 0$. Alors :

$$\exists!\, n \in \mathbb{Z}, \quad nb \leq a < (n+1)b\,.$$

Démonstration. Commençons par démontrer l'existence.

Cas 1 : $a \geq 0$. L'ensemble

$$\mathscr{E} = \{n \in \mathbb{N} \,|\, nb > a\} \subset \mathbb{N}$$

est non vide, par axiome D, admet un plus petit élément n_0. Il est alors clair que $n = n_0 - 1$ vérifie l'inégalité voulue.

Cas 2 : $a < 0$. En appliquant le cas précédent à $-a$, on obtient existence d'un entier relatif m tel que $mb \leq -a < (m+1)b$, ainsi :

$$-(m+1)b < a \leq -mb$$

ce qui permet de conclure en posant $n = -m-1$ si $a \neq mb$ et $n = -m$ sinon.

Pour démontrer l'unicité, supposons que nous disposions de deux entier n, p vérifiant l'égalité voulue, *i.e*

$$nb \leq a < (n+1)b \quad \text{et} \quad pb \leq a < (p+1)b$$

qui implique

$$nb \leq a < (p+1)$$

et donc $n < p+1$ (car $b > 0$), d'où $n \leq p$. De façon symétrique, on obtient $p \leq n$, *ergo* $n = p$. □

b) Partie entière

Définition VI.6. Soit $x \in \mathbb{R}$; on appelle **partie entière** (par défaut) de x l'unique entier n vérifiant :

$$n \leq x < n+1\,.$$

Notation. $\lfloor x \rfloor$, $E(x)$.

➠ **Exemple VI.6.** $\lfloor 1,25 \rfloor = 1$, $\lfloor -2,7 \rfloor = -3$.

✌ **Remarque VI.9.**

- Cette quantité est bien définie : appliquer le corollaire VI.10.a avec $a = x$ et $b = 1$.
- $\lfloor x \rfloor$ est caractérisée par :

$$\begin{cases} \lfloor x \rfloor \in \mathbb{Z} \\ \lfloor x \rfloor \leq x < \lfloor x \rfloor + 1 \end{cases}.$$

- $\lfloor x \rfloor$ est parfois appelée *plancher* de x ; il existe également un plafond, noté $\lceil x \rceil$ et caractérisé par :

$$\begin{cases} \lceil x \rceil \in \mathbb{Z} \\ \lceil x \rceil - 1 < x \leq \lceil x \rceil \end{cases}.$$

Une application classique de la notion de partie entière est l'**approximation d'un réel à** 10^{-n} près, pour $n \geq 0$. Posons, pour $n \in \mathbb{N}$ et $x \in \mathbb{R}$:

$$x_n = \frac{\lfloor 10^n x \rfloor}{10^n} \ .$$

Alors, comme

$$\lfloor 10^n x \rfloor \leq 10^n x < \lfloor 10^n x \rfloor + 1$$

on a bien :

$$x_n \leq x < x_n + 10^{-n} \quad i.e \quad |x - x_n| < 10^{-n} \ .$$

➠ **Exemple VI.7.** $\pi = 3.14$ à 10^{-2} près.

c) Résultats de densité

Proposition VI.11. Soient $x, y \in \mathbb{R}$ tels que $x < y$. Alors il existe un rationnel $\zeta \in \mathbb{Q}$ et un irrationnel $\tau \in \mathbb{R} \setminus \mathbb{Q}$ tels que :
- $x \leq \zeta \leq y$;
- $x \leq \tau \leq y$.

✌ **Remarque VI.10.** En itérant ce procédé, on en déduit qu'entre deux réels distincts il existe une infinité de rationnels et une infinité d'irrationnels. On dit que $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont **denses dans** $\mathbb{R}$.

Démonstration. Posons $\varepsilon = y - x$; alors par propriété d'Archimède (théorème VI.10), comme $\varepsilon > 0$, il existe $q \in \mathbb{N}$ tel que $q \times 1 > \frac{1}{\varepsilon}$, *i.e*

$$0 < \frac{1}{q} < \varepsilon = y - x \ .$$

Une application du corollaire VI.10.a à x et $\frac{1}{q}$ nous livre un entier naturel $p \in \mathbb{N}$ tel que :

$$\frac{p-1}{q} \leq x < \frac{p}{q}$$

et donc :

$$\begin{aligned} x < \frac{p}{q} &= \frac{p-1+1}{q} \\ &\leq x + \frac{1}{q} \\ &\leq x + \varepsilon \\ &= y \ . \end{aligned}$$

On en déduit que le rationnel $\zeta = \frac{p}{q}$ est contenu dans le segment $[x, y]$. Un raisonnement identique remplaçant q par $q\sqrt{2}$ lors de l'application du corollaire VI.10.a nous permet de démontrer que l'irrationnel $\tau = \frac{p}{q\sqrt{2}}$ appartient à ce même segment. □

4.⁻ Intervalles

a) C'est quoi ?

Rappellons que les intervalles de $\mathbb{R}$ sont les neuf familles d'ensembles suivants, pour $a, b \in \mathbb{R}$:

— $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ **[segment]** ;
— $[a, b[= \{x \in \mathbb{R} \mid a \leq x < b\}$;
— $]a, b] = \{x \in \mathbb{R} \mid a > x \leq b\}$;
— $]a, b[= \{x \in \mathbb{R} \mid a < x < b\}$;
— $[a, +\infty[= \{x \in \mathbb{R} \mid a \leq x\}$;
— $]a, +\infty[= \{x \in \mathbb{R} \mid a < x\}$;
— $]-\infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$;
— $]-\infty, b[= \{x \in \mathbb{R} \mid x < b\}$;
— $]-\infty, +\infty[= \mathbb{R}$.

✌ **Remarque VI.11.**
— Les quatre premier types d'intervalles ci–dessus sont bornés, les autres non.
— Si $a > b$, l'intervalle $[a, b]$ est vide. Ceci signifie en particulier que $\emptyset$ est un intervalle.
— **Certains** (pas tous !) intervalles sont dits **ouverts** ou **fermés**, comme résumé dans le tableau *infra*.

Ouvert	**Fermé**
$]a, b[$	$[a, b]$
$]a, +\infty[$	$[a, +\infty[$
$]-\infty, b[$	$]-\infty, b]$
$\emptyset$	$\emptyset$
$\mathbb{R}$	$\mathbb{R}$

b) Convexes

Définition VI.7. Une partie A de $\mathbb{R}$ est dite **convexe** si :

$$\forall x, y \in A, \quad [x, y] \subset A\,.$$

➠ **Exemple VI.8.** Les intervalles sont clairement connexes. L'ensemble $\mathbb{R}^*$ ne l'est pas car (par exemple) $[-1, 1] \not\subset \mathbb{R}^*$.

Proposition VI.12. Les convexes de $\mathbb{R}$ sont exactement les intervalles.

Démonstration. Ceci est une ébauche de démonstration. Nous venons de voir que les intervalles étaient convexes ; reste à montrer la réciproque. Soit donc $A \subset \mathbb{R}$ un ensemble convexe.

Cas 1 : $A = \emptyset$**.** Youpi.

Cas 2 : $A \neq \emptyset$**.**

Cas 2.1 : A n'est ni majoré ni minoré. A est alors égal à $\mathbb{R}$.

Cas 2.2 : A est majoré et minoré. A admet alors une borne supérieure b et une borne inférieure a et $A =]a, b[$, $[a, b[$,$]a, b]$ ou $[a, b]$.

Cas 2.3 : A est (par exemple) minorée et non majorée. Posons $a = \inf(A)$ et montrons que $A =]a, \infty[$ ou $[a, \infty[$. L'inclusion de gauche à droite est évidente ; prenons ensuite $z > a$ et montrons que $z \in A$. Comme z ne peut minorer A, il existe alors $x \in A$ tel que $x < z$ et comme A n'est pas majorée, il existe $y \in A$ tel que $z < y$. Le réel z est donc compris dans le segment $[x, y]$ qui est inclus dans A par convexité, donc $z \in A$.

□

Chapitre VII

Trigonométrie(s)

1.⁻ Fonctions circulaires

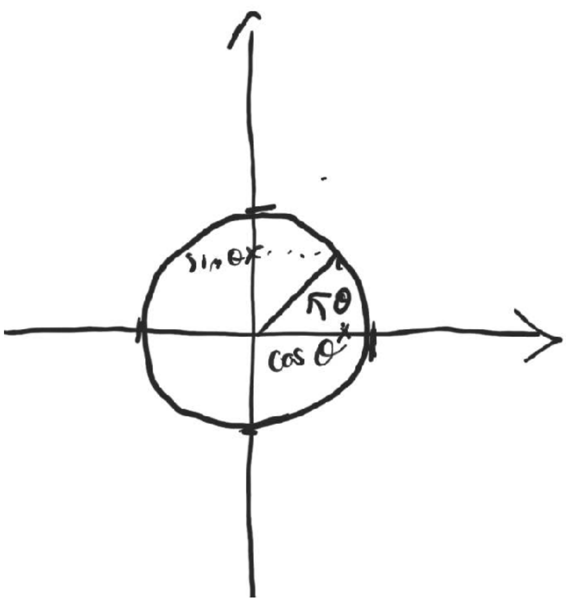
a) Cosinus, sinus, tangente

Définition VII.1. Soit $T \in \mathbb{R}$. Une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite T–**périodique** si :

$$\forall x \in \mathbb{R}, \quad f(x+T) = f(x) \ .$$

✌ **Remarque VII.1.** Il n'y pas unicité de T (appelé période) ; certains ouvrages préfèrent introduire une "période minimale", à savoir le plus petit T positif tel que f soit T–périodique.

Les fonctions cosinus et sinus ont normalement été définies en terminale à travers une construction géometrico–graphique que nous reproduisons ci–ensuite (le truc vaguement rond est le cercle unité de $\mathbb{R}^2$). Nous ne serons hélas guère en mesure de faire mieux.



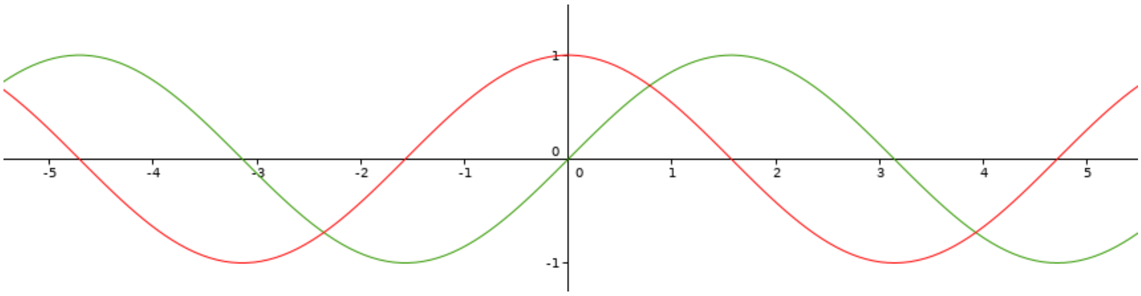
Proposition VII.1.

(i) Les fonctions cosinus et sinus sont 2π–périodiques et dérivables sur $\mathbb{R}$.

(ii) La fonction cosinus est paire, la fonction sinus impaire.

(iii) $\sin' = \cos$, $\cos' = -\sin$.

À toutes fins utiles, nous reproduisons ici l'allure courbes des fonctions cosinus (en rouge) et sinus (et vert) ci–ensuite.



Ces fonctions admettent de nombreuses valeurs remarquables, dont nous listons une selection dans le tableau *infra*.

θ	0	$\pi/2$	$\pi/3$	$\pi/4$	$\pi/6$	π	$3\pi/2$
$\cos(\theta)$	1	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	-1	0
$\sin(\theta)$	0	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	-1

Ces fonctions sont également remarquables de par leurs symétries : par exemple, si $\theta \in \mathbb{R}$, on a :

— $\cos(\theta + \pi) = -\cos(\theta)$;
— $\sin(\theta + \pi) = -\sin(\theta)$;
— $\cos(\pi - \theta) = -\cos(\theta)$;
— $\sin(\pi - \theta) = \sin(\theta)$.

Il est possible de "retrouver" ces formules via le cercle trigonométrique (faites un dessin !).

Arrive maintenant le moment que vous attendiez probablement avec impatience : le **formulaire de trigonométrie** ! Il est à connaître **impérativement et parfaitement**, mais ne paniquez pas (trop) : toutes les formules qui suivent peuvent être retrouvées à l'aide des cinq premières, à condition d'avoir une vague idée de ce que l'on cherche à obtenir.

Proposition VII.2 (Formulaire de trigonométrie, kit de survie)**.** Soient $a, b \in \mathbb{R}$; alors :

— $\cos^2(a) + \sin^2(a) = 1$;
— $\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$;
— $\sin(a - b) = \sin(a)\cos(b) - \sin(b)\cos(a)$;
— $\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b)$;
— $\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$;

Proposition VII.3 (Formulaire de trigonométrie, suite et fin)**.** Soient $a, b \in \mathbb{R}$; alors :
- $\cos(2a) = \cos^2(a) - \sin^2(a) = 2\cos^2(a) - 1 = 1 - 2\sin^2(a)$;
- $\sin(2a) = 2\sin(a)\cos(a)$;
- $\sin(a)\cos(b) = \frac{1}{2}(\sin(a+b) + \sin(a-b))$;
- $\cos(a)\cos(b) = \frac{1}{2}(\cos(a+b) + \cos(a-b))$;
- $\sin(a)\sin(b) = -\frac{1}{2}(\cos(a+b) - \cos(a-b))$.

Définition VII.2. On appelle **fonction tangente** l'application suivante :

$$\begin{aligned} \tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \,\middle|\, k \in \mathbb{Z} \right\} &\to \mathbb{R} \\ x &\mapsto \frac{\sin(x)}{\cos(x)} . \end{aligned}$$

Cette nouvelle fonction trigonométrique est définie partie sauf en les points d'annulation du cosinus. Elle admet des valeurs remarquables que l'on peut déduire de celles de cosinus et sinus, résumées (très partiellement) dans le tableau qui suit.

x	0	$\pi/4$	π
$\tan(x)$	0	1	0

Proposition VII.4. La fonction tangente est :
- impaire ;
- π–périodique ;
- dérivable sur son ensemble de définition, et :

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

Démonstration. L'imparité est immédiate. Pour la périodicité, remarquons que, si $x \in \mathbb{R}$, on a :

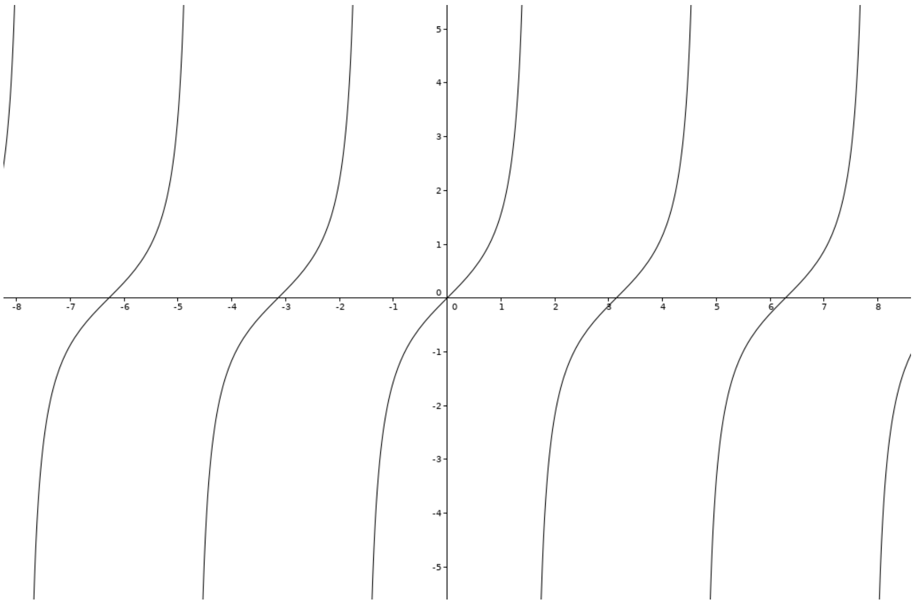
$$\begin{aligned} \tan(x+\pi) &= \frac{\sin(x+\pi)}{\cos(x+\pi)} \\ &= \frac{-\sin(x)}{-\cos(x)} \\ &= \tan(x) . \end{aligned}$$

La fonction tan est dérivable comme produit de fonctions dérivables partout ou le cosinus ne s'annule pas, et par dérivée d'un quotient on a :

$$\begin{aligned}\tan' &= \frac{\sin'\cos - \sin\cos'}{\cos^2}\\ &= \frac{\cos^2+\sin^2}{\cos^2}\\ &= \begin{cases} 1+\tan^2\\ \dfrac{1}{\cos^2}\end{cases} .\end{aligned}$$

□

Ce résultat nous permet de finir l'étude de la fonction tangente, en notant que sa dérivée est positive en tout point donc qu'elle est strictement croissante avec limites infinies en chaque $\frac{\pi}{2}+k\pi$, $\pi \in \mathbb{Z}$. Sa courbe a donc l'allure suivante.



Ajoutons à ceci une formule, dont la démonstration est un calcul immonde : pour tous $a, b \in \mathbb{R}$,

$$\tan(a \pm b) = \frac{\tan(a) \pm \tan(b)}{1 \mp \tan(a)\tan(b)} . \qquad (\mathbf{E}\text{:VII.1})$$

Proposition VII.5. Soit $x \in \mathbb{R}\setminus\{\pi + 2k\pi \mid k \in \mathbb{Z}\}$ et soit $t = \tan(x/2)$. On a alors les formules suivantes :

(i)
$$\cos(x) = \frac{1-t^2}{1+t^2} ;$$

(ii)
$$\sin(x) = \frac{2t}{1+t^2} ;$$

(iii)
$$\tan(x) = \frac{2t}{1-t^2} .$$

Démonstration.

(iii) Cette formule découle du fait que :

$$\begin{aligned}\tan(x) &= \tan\left(\frac{x}{2}+\frac{x}{2}\right)\\ &= \frac{2t}{1-t^2} \text{ par } (\mathbf{E}\ \text{:VII.1})\ .\end{aligned}$$

(ii)

$$\begin{aligned}\sin(x) &= \sin\left(2\frac{x}{2}\right)\\ &= 2\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)\\ &= 2t\cos^2\left(\frac{x}{2}\right)\\ &= 2t\times\frac{1}{1+\tan^2(x/2)}\\ &= \frac{2t}{1+t^2}\ .\end{aligned}$$

Le point (i) s'obtient comme corollaire des deux précédents. □

b) Fonctions circulaires réciproques

Commençons par énoncer un théorème sur lequel nous reviendrons dans le chapitre XI.

Théorème VII.6 (Théorème de la bijection)**.**
Soit I un intervalle de $\mathbb{R}$ et soit $f : I \to \mathbb{R}$ une application continue strictement monotone. Alors :

(i) f réalise une bijection de I vers $J = f(I)$;

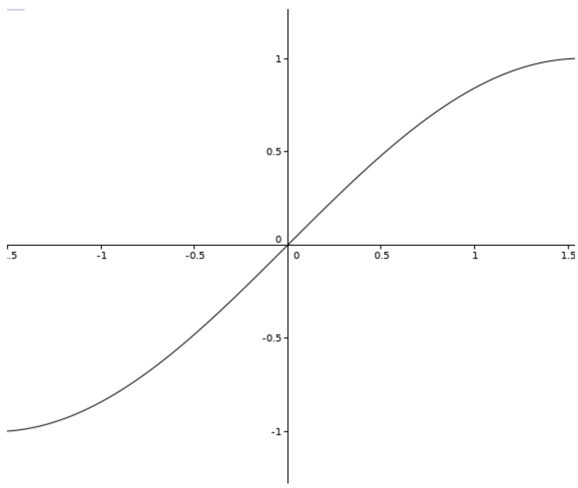
(ii) $f^{-1} : J \to I$ est de même monotonie que f.

✌ **Remarque VII.2.** Ce résultat est à ne pas confondre avec le théorème des valeurs intermédiaires ; affaire à suivre…

La question qui nous vient naturellement (ou pas) à l'esprit est la suivante : les fonctions cosinus, sinus et tangente sont–elles des bijections ? La réponse est évidemment **NON**, pour de bêtes raisons de non–injectivité liées à la périodicité. Par contre, la fonction suivante vérifie les hypothèses du théorème de la bijection :

$$\widehat{\sin} : \left[-\frac{\pi}{2},\frac{\pi}{2}\right] \to [-1,1]$$
$$x \mapsto \sin(x) \quad .$$

Pour s'en convaincre, il suffit de vérifier que sa dérivée est strictement positive (c'est le cas). La courbe représentative de $\widehat{\sin}$ est sans appel : cette fonction est strictement croissante ; elle est donc bijective de réciproque strictement croissante.



Définition VII.3. On appelle **fonction arc sinus** la réciproque de $\widehat{\sin}$.

Notation. arcsin : $[-1,1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

✖ **Attention :** arcsin n'est **PAS** la réciproque de la fonction sinus. En particulier, prenons garde à la quantification des égalités suivantes :
— si $x \in [-1,1]$, $\sin(\arcsin(x)) = x$;
— si $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $\arcsin(\sin(x)) = x$.
La première égalité ne posera (normalement) jamais problème ; la seconde par contre peut amener quelques pièges vicieux. Par exemple, $\sin(\pi/2 + 2\pi) = 1$ et pourtant $\arcsin(\sin(\pi/2 + 2\pi)) = \arcsin(1) = \pi/2$.

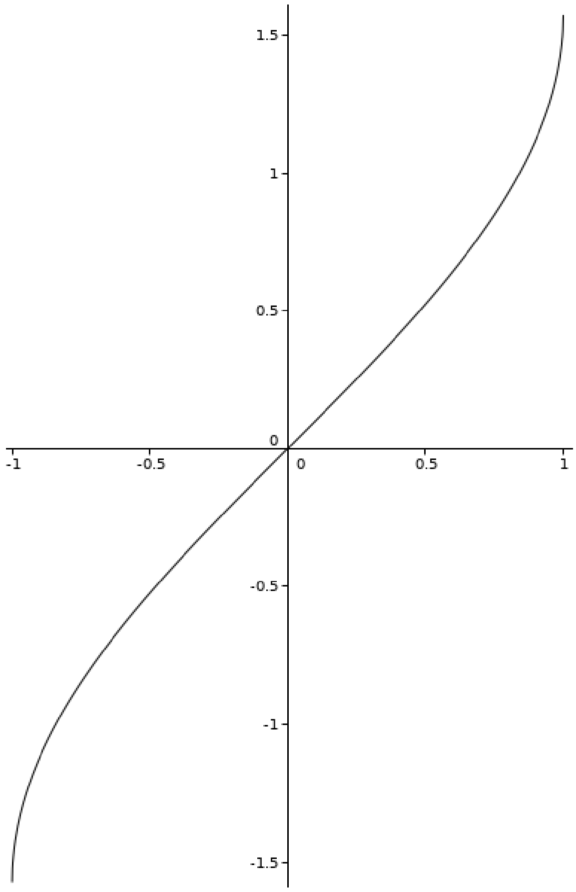
Proposition VII.7. La fonction arcsin est une bijection continue strictement croissante de $[-1,1]$ sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Il s'agit de plus d'une fonction impaire.

Démonstration. La seule chose qu'il nous reste à démontrer est l'imparité, qui est en fait un cas particulier d'un résultat général : si $f : E \to F$ est une bijection impaire, alors pour tout $x \in F$ on a :

$$\begin{aligned} f^{-1}(-x) &= f^{-1}(-f(f^{-1}(x)) \\ &= f^{-1} \circ f(-f^{-1}(x)) \text{ car } f \text{ est impaire} \\ &= -f^{-1}(x) \end{aligned}$$

et donc f^{-1} est également impaire. □

De tout ceci on peut déduire l'allure de la courbe représentative de arcsin ; il s'agit du symétrique de celle de $\widehat{\sin}$ par rapport à la première bissectrice.



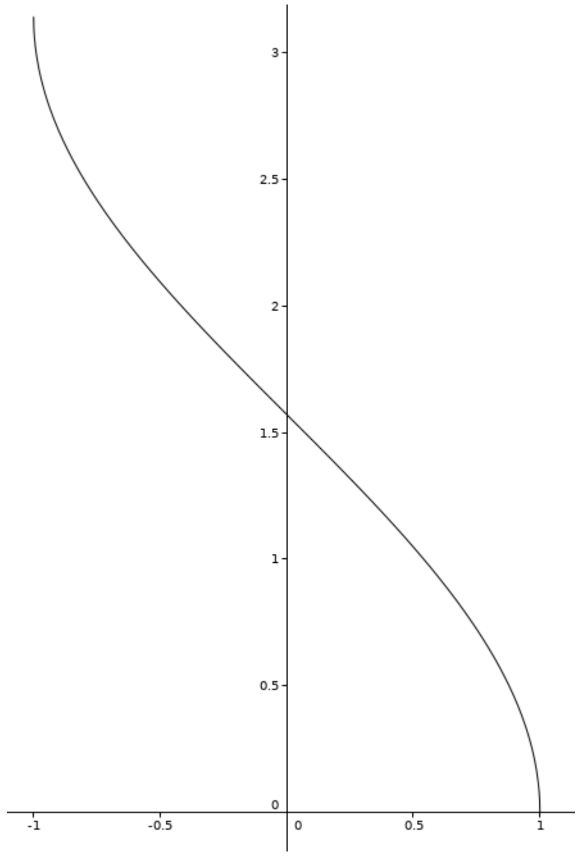
Définition VII.4. De façon analogue, on appelle **fonction arc cosinus** la réciproque de la bijection strictement décroissante

$$\widehat{\cos} : [0,\pi] \to [-1,1]$$
$$x \mapsto \cos(x)\ .$$

Il s'agit d'une bijection continue strictement décroissante qui n'est ni paire ni impaire.

Notation. arcos : $[-1,1] \to [0,\pi]$

Les mêmes précautions d'usage sont à observer pour pour arcos que pour arcsin. Le graphique ci–dessous présente la courbe de arcos.



✍ **Exercice VII.1.** Soit $x \in [-1, 1]$; déterminer une expression élémentaire de $\cos(\arcsin(x))$.

➥ **Correction :** *Remarquons tout d'abord que :*

$$\cos(\arcsin(x))^2 = 1 - \sin(\arcsin(x)^2 = 1 - x^2 \ .$$

Ainsi, on a :

$$|\cos(\arcsin(x)| = \sqrt{1 - x^2} \ .$$

Or, $\arcsin(x) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$*, donc* $\cos(\arcsin(x) \geq 0$*, ce qui permet de conclure in fine que :*

$$\cos(\arcsin(x) = \sqrt{1 - x^2} \ .$$

Proposition VII.8. Les fonctions arcsin et arcos sont dérivables sur $]-1, 1[$ et leurs dérivées sont données par, pour $x \in]-1, 1[$:

$$\arcsin'(x) = \frac{1}{\sqrt{1 - x^2}}$$

et

$$\operatorname{arcos}'(x) = -\frac{1}{\sqrt{1 - x^2}} \ .$$

Démonstration. La démonstration est laissée en exercice au lecteur : il s'agit d'une application du théorème I.4 (dérivée d'une réciproque) et de l'exercice précédent. □

Remarque VII.3. Le lecteur averti aura noté que ces fonctions ne sont pas dérivables en ± 1, qui sont les images des points d'annulation des dérivées de $\widehat{\sin}$ et $\widehat{\cos}$. Toute ceci sera approfondi au chapitre XII.

Proposition VII.9. La fonction suivante est une bijection strictement croissante et continue :

$$\widehat{\tan} : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \mathbb{R}$$
$$x \mapsto \tan(x) \ .$$

Démonstration. Il s'agit d'une conséquence immédiate du théorème de la bijection (théorème VII.6) et de l'étude précedemment menée de la fonction tangente. $\square$

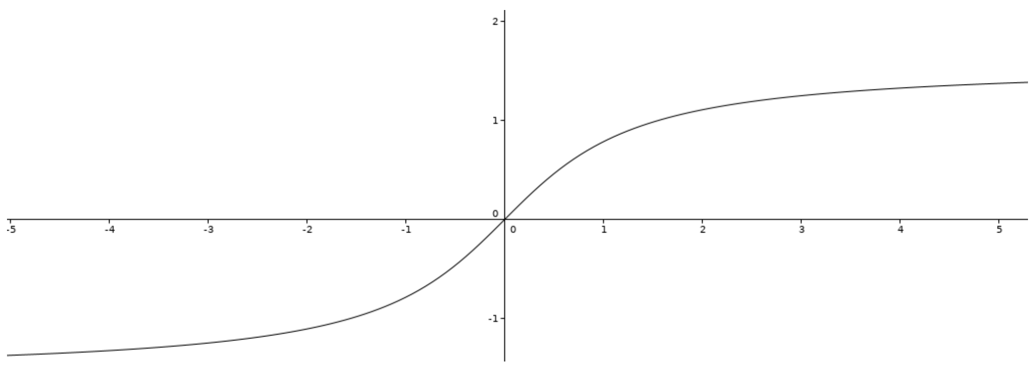
Définition VII.5. On appelle **fonction arc tangente** la réciproque de la fonction $\widehat{tan}$.

Notation. $\arctan : \mathbb{R} \to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$

Proposition VII.10. La fonction arctan est une bijection continue impaire strictement croissante. De plus :
- $\arctan(x) \xrightarrow[x\to\pm\pi/2]{} \pm\infty$;
- arctan est dérivable sur $\mathbb{R}$ et, pour tout $x \in \mathbb{R}$:

$$\arctan'(x) = \frac{1}{1+x^2} \ .$$

Pour conclure, nous reproduisons ci–ensuite l'allure de la courbe représentative de la fonction arctan.



2.⁻ Fonctions hyperboliques

a) Cosinus et sinus hyperboliques

Définition VII.6. Soit $x \in \mathbb{R}$; on appelle :
— **cosinus hyperbolique de** x le réel

$$\mathrm{ch}(x) = \frac{e^x + e^{-x}}{2} ;$$

— **sinus hyperbolique de** x le réel

$$\mathrm{sh}(x) = \frac{e^x - e^{-x}}{2} .$$

✌ **Remarque VII.4.** Ces formules nous permettent de définir deux fonctions $\mathrm{ch} : \mathbb{R} \to \mathbb{R}$ et $\mathrm{sh} : \mathbb{R} \to \mathbb{R}$.

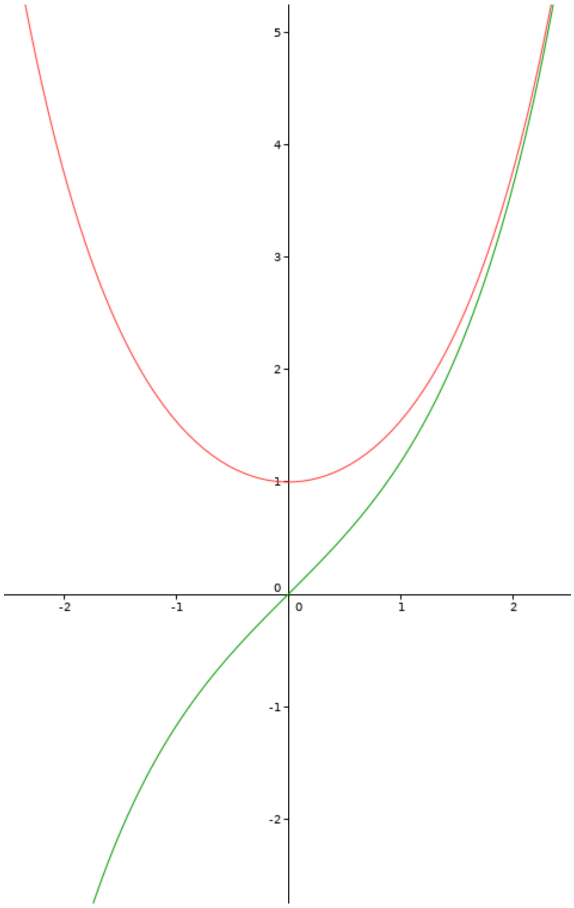
Proposition VII.11. On a les résultats suivants :
(i) la fonction ch est paire ;
(ii) la fonction sh est impaire ;
(iii) les fonctions *ch* et sh sont dérivables ; de plus $\mathrm{ch}' = \mathrm{sh}$ et $\mathrm{sh}' = \mathrm{ch}$;
(iv) pour tout $x \in \mathbb{R}$, on a :

$$\mathrm{ch}^2(x) - \mathrm{sh}^2(x) = 1 .$$

Démonstration. Il s'agit d'une succession de calculs immédiats. □

Pour terminer, une brève étude de fonction révèle de plus les propriétés suivantes :
— ch est strictement positive, de valeur minimale $1 = \mathrm{ch}(0)$;
— $\mathrm{ch}(x) \xrightarrow[x \to \pm\infty]{} \infty$;
— $\mathrm{sh}(x) \xrightarrow[x \to \pm\infty]{} \pm\infty$.

Ceci nous permet de déduire l'allure des courbes représentatives de ces deux fonctions, reproduite ci–ensuite (ch en rouge, sh en vert).



b) Tangente hyperbolique

Définition VII.7. Soit $x \in \mathbb{R}$; on appelle **tangente hyperbolique de** x la quantité

$$\frac{\mathrm{sh}(x)}{\mathrm{ch}(x)} = \frac{e^{2x}-1}{e^{2x}+1} .$$

Notation. $\mathrm{th} : \mathbb{R} \to \mathbb{R}$

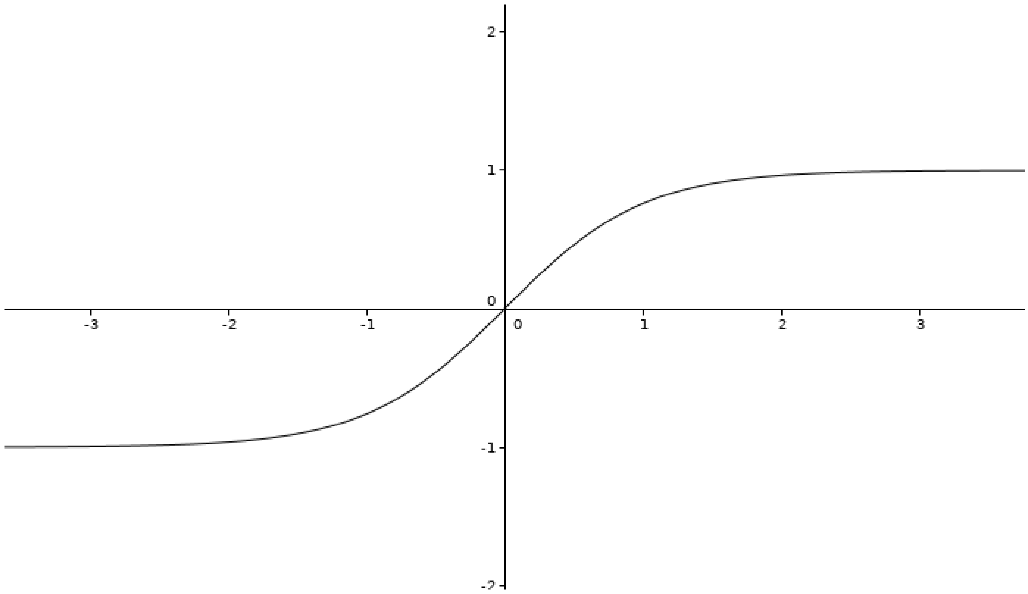
Proposition VII.12. La fonction th est impaire, strictement croissante et dérivable. Sa dérivée vérifie :

$$\mathrm{th}' = \frac{1}{\mathrm{ch}^2} = 1 - \mathrm{th}^2 .$$

Un rapide calcul utilisant l'expression $\mathrm{th}(x) = \dfrac{e^{2x}-1}{e^{2x}+1}$ nous permet de déduire que (mettre en facteur le terme dominant au numérateur et dénominateur) :

$$\mathrm{th}(x) \xrightarrow[x \to \pm\infty]{} \pm 1 .$$

Nous sommes donc en mesure de tracer l'allure de la courbe représentative de la fonction th.


2
1
0
-3
-2
-1
0
1
2
3
-1
-2

Chapitre VIII

Nombres complexes

1.⁻ Le corps $\mathbb{C}$ des nombres complexes

a) C'est quoi ?

Tout comme l'ensemble des nombres réels, $\mathbb{C} = \{a+ib \mid a,b \in \mathbb{R}\}$ est muni d'une structure de corps (*cf.* chapitre X). Il n'admet toutefois pas d'ordre "naturel" comme $\mathbb{R}$ ou $\mathbb{Q}$. Plus précisément, nous baserons ce chapitre sur le résultat (admis) suivant.

Théorème VIII.1 (Corps des nombres complexes)**.**
Il existe un "unique" (à isomorphisme près) corps $(\mathbb{C}, +, \times)$ tel que :

(A) $\mathbb{R} \subset \mathbb{C}$ avec préservation de l'addition et de la multiplication ;

(B) il existe $i \in \mathbb{C}$ tel que $i^2 = -1$;

(C) pour tout $z \in \mathbb{C}$, il existe un unique couple $(a,b) \in \mathbb{R}^2$ tel que $z = a + ib$.

Si $z = a+ib \in \mathbb{C}$, on appelle a (resp. b) sa **partie réelle** (resp. **partie imaginaire**), notée $\mathrm{Re}(z)$ (resp. $\mathrm{Im}(z)$).

Notation. On notera $i\mathbb{R}$ l'ensemble des **imaginaires purs**, à savoir $i\mathbb{R} = \{ib \mid b \in \mathbb{R}\}$.

✌ **Remarque VIII.1.**

— **Attention :** si $z \in \mathbb{C}$, $\mathrm{Re}(z), \mathrm{Im}(z) \in \mathbb{R}$.

— Si $z \in \mathbb{C}$, alors $z \in \mathbb{R} \Leftrightarrow \mathrm{Im}(z) = 0$ et $z \in i\mathbb{R} \Leftrightarrow \mathrm{Re}(z) = 0$.

— Le choix du nombre complexe i est arbitraire : on pourrait tout à fait choisir son opposé et développer exactement la même théorie.

b) Lien au plan $\mathbb{R}^2$

On peut identifier $\mathbb{C}$ au plan $\mathbb{R}^2$ via la bijection suivante :

$$\begin{aligned} \varphi : \mathbb{C} &\to \mathbb{R}^2 \\ z &\mapsto (\mathrm{Re}(z), \mathrm{Im}(z)) \end{aligned}$$

de réciproque

$$\begin{aligned} \psi : \mathbb{R}^2 &\to \mathbb{C} \\ (x,y) &\mapsto x + iy \end{aligned} \quad .$$

Vocabulaire.
- Si $A \in \mathbb{R}^2$, le nombre complexe $\psi(A)$ est appelé **affixe** de A.
- Si $z \in \mathbb{C}$, le point du plan $\varphi(z)$ est appelé **image** de z.

✍ **Exercice VIII.1.** Placer dans le plan les points d'affixes $1, i, 1+i, 1-i$ et $\dfrac{i}{2}$.

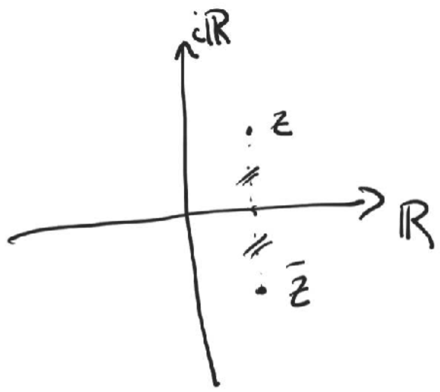
c) Conjugaison

Définition VIII.1. Soit $z \in \mathbb{C}$; on appelle **conjugé** de z le nombre complexe

$$\overline{z} = \mathrm{Re}(z) - i\mathrm{Im}(y)\,.$$

➠ **Exemple VIII.1.** $\overline{1+4i} = 1-4i$.

✌ **Remarque VIII.2.** Géométriquement, l'image de $\overline{z}$ est le symétrique de celle de z par rapport à la droite rélle.



Proposition VIII.2. Soit $z \in \mathbb{C}$; alors :

(i) $\mathrm{Re}(z) = \dfrac{z+\overline{z}}{2}$ et $\mathrm{Im}(z) = \dfrac{z-\overline{z}}{2i}$;

(ii) $\overline{\overline{z}} = z$;

(iii) $z \in \mathbb{R} \Leftrightarrow \overline{z} = z$;

(iv) $z \in i\mathbb{R} \Leftrightarrow \overline{z} = -z$

Démonstration. Il s'agit de calculs immédiats. Pour (iii) et (iv), utiliser (i). □

✌ **Remarque VIII.3.** Les points (iii) et (iv) sont particulièrement utiles en pratique pour démontrer qu'un nombre complexe est en fait réel ou imaginaire pur.

Proposition VIII.3. Soient $z, z' \in \mathbb{C}$. Alors :

(i) $\overline{z+z'} = \overline{z} + \overline{z'}$;

(ii) $\overline{zz'} = \overline{z}\,\overline{z'}$;

(iii) **si** $z' \neq 0$, $\overline{\left(\dfrac{z}{z'}\right)} = \dfrac{\overline{z}}{\overline{z'}}$.

Démonstration. Calculs, calculs, toujours des calculs... Le seul point légèrement technique est le (iii) ; il faut pour celui–ci remarquer que :

$$\begin{aligned}\overline{\left(\frac{z}{z'}\right)} &= \overline{\left(z \times \frac{1}{z'}\right)} \\ &= \overline{z} \times \overline{\left(\frac{1}{z'}\right)} \text{ par (ii)}\end{aligned}$$

Or, comme $z' \times \dfrac{1}{z'} = 1$, on déduit du point (ii) en conjugant à gauche et à droite de l'éqalité que

$$\overline{\left(\frac{1}{z'}\right)} = \frac{1}{\overline{z'}}$$

d'où le résultat. □

Exercice VIII.2. Donner une description explicite de l'ensemble

$$E = \left\{ z \in \mathbb{C} \setminus \{i\} \;\middle|\; \frac{z+i}{z-i} \in i\mathbb{R} \right\}.$$

Correction : *Soit $z \in \mathbb{C} \setminus \{i\}$; alors :*

$$\begin{aligned}\frac{z+i}{z-i} \in i\mathbb{R} &\Leftrightarrow \overline{\left(\frac{z+i}{z-i}\right)} = -\frac{z+i}{z-i} \\ &\Leftrightarrow \frac{\overline{z}-i}{\overline{z}+i} = -\frac{z+i}{z-i} \\ &\Leftrightarrow (\overline{z}-i)(z-i) = -(\overline{z}+i)(z+i) \\ &\Leftrightarrow z\overline{z} - i(z+\overline{z}) - 1 = -z\overline{z} - i(z+\overline{z}) + 1 \\ &\Leftrightarrow z\overline{z} = 1\,.\end{aligned}$$

Ainsi, E est l'ensemble des complexes $z = x + iy$ tels que $z\overline{z} = 1$, *i.e* $x^2 + y^2 = 1$. Il s'agit donc du cercle unité.

d) Module

Proposition VIII.4. Soit $z \in \mathbb{C}$; alors $z\overline{z} \in \mathbb{R}_+$.

Démonstration. $z\overline{z} = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2 \in \mathbb{R}_+$. □

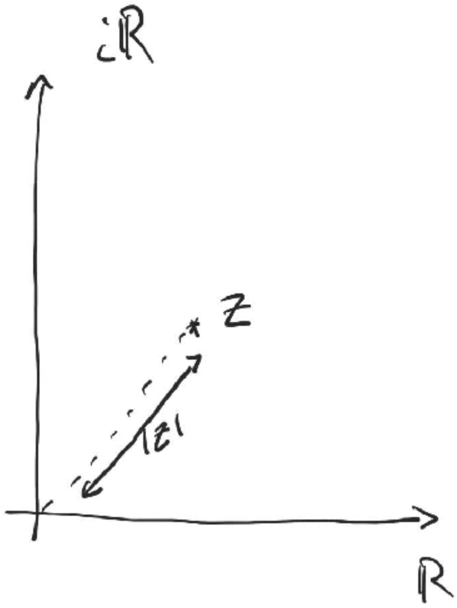
Définition VIII.2. Soit $z \in \mathbb{C}$; on appelle **module** de z la quantité $\sqrt{z\overline{z}}$.

Notation. $|z|$

Exemple VIII.2. $|1+i| = \sqrt{2}$.

✌ **Remarque VIII.4.**

— Géométriquement, le module de $z \in \mathbb{C}$ correspond à la distance entre le point d'affixe z et l'origine du plan $\mathbb{R}^2$.



Plus généralement, si $z, z' \in \mathbb{C}$, la quantité $|z - z'|$ est égale à la distance entre les points d'affixes respectives z et z' dans $\mathbb{R}^2$.

— Si $\omega \in \mathbb{C}$ et $R > 0$, une équation du cercle $\mathscr{C}(\omega, R)$ de centre ω et de rayon R est donc :

$$|z - \omega| = R \tag{E:VIII.1}$$

ou

$$(\mathrm{Re}(z) - \mathrm{Re}(\omega)^2 + (\mathrm{Im}(z) - \mathrm{Im}(\omega))^2 = R^2\,. \tag{E:VIII.2}$$

Proposition VIII.5. Soit $z \in \mathbb{C}$. Alors :

(i) $|z| = 0 \Leftrightarrow z = 0$;

(ii) $|z| = |\overline{z}|$;

(iii) $\mathrm{Re}(z) \leq |\mathrm{Re}(z)| \leq |z|$ et $\mathrm{Im}(z) \leq |\mathrm{Im}(z)| \leq |z|$

Démonstration. Les points (i) et (ii) découlent immédiatement de la définition. Pour le point (iii), posons $z = x + iy$; alors :

$$|z| = \sqrt{x^2 + y^2} \geq \sqrt{x^2} = |x| \geq x$$

et symétriquement avec la partie imaginaire, d'où le résultat. □

Proposition VIII.6. Soient $z, z' \in \mathbb{C}$. Alors :

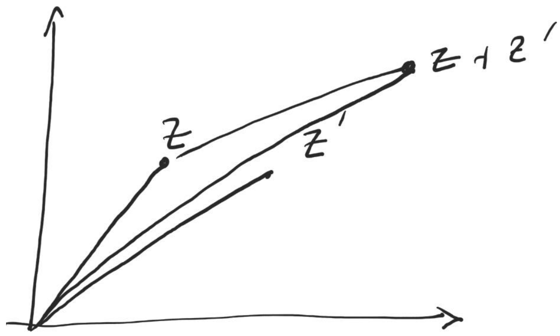
(i) $|zz'| = |z||z'|$;

(ii) si $z' \neq 0$, $\left|\dfrac{z}{z'}\right| = \dfrac{|z|}{|z'|}$;

(iii) $|z + z'| \leq |z| + |z'|$ **[inégalité triangulaire 1]** ;

(iv) $||z| - |z'|| \leq |z - z'|$ **[inégalité triangulaire 2]**.

✌ **Remarque VIII.5.** Géométriquement, l'inégalité triangulaire 1 nous indique que le plus court chemin entre deux points du plan est la ligne droite : la distance du parcours $(0, z, z + z')$ est infériure à celle du parcours $(0, z + z')$.



Démonstration.

(i) Découle de la définition et des propriétés de la conjugaison.

(ii) Idem.

(iii) Si $z' = 0$, l'inégalité est triviale. Sinon, posons $u = \dfrac{z}{z'}$ et montrons que $|1 + u| \leq 1 + |u|$. Pour ce faire, nous allons comparer les carrés de ces deux quantités :

$$\begin{aligned} |1+u|^2 - (1+|u|)^2 &= (1+u)(1+\overline{u}) - 1 - 2|u| - |u|^2 \\ &= u + \overline{u} - 2|u| \\ &= 2(\mathrm{Re}(u) - |u|) \\ &\leq 0\,. \end{aligned}$$

Ne nous reste plus qu'à multiplier cette inégalité par $|z'|$ pour obtenir le résultat.

(iv) Nous avons vu lors de la démonstration de la proposition VI.5 comment déduire la seconde inégalité triangulaire de la première. La même méthode est applicable ici.

□

✌ **Remarque VIII.6. Cas d'égalité dans la première inégalité triangulaire.** Pour caractériser les complexes z, z' tels que $|z + z'| = |z| + |z'|$, procédons par analyse–synthèse.

Analyse. Supposons trouvés $z, z' \in \mathbb{C}$ tels que $|z + z'| = |z| + |z'|$. Si $z' \neq 0$, on pose $u = \frac{z}{z'}$ et on obtient

$$|1 + u| = 1 + |u|$$

ce qui est équivalent, nous venons de le voir dans la démonstration de l'inégalité triangulaire, à :

$$\mathrm{Re}(u) = |u| .$$

Ceci entraine, comme $|u| = \sqrt{\mathrm{Re}(u)^2 + \mathrm{Im}(u)^2}$, que $u \in \mathbb{R}_+$, *i.e* $\frac{z}{z'} = u \in \mathbb{R}_+$.

Synthèse. Si $z' = 0$, l'égalité est trivialement vérifiée. Si $z = uz'$, avec $u \in \mathbb{R}_+$, alors :

$$\begin{aligned} |z' + z| &= |z' + uz'| \\ &= |(1 + u)z'| \\ &= |1 + u||z'| \\ &= (1 + u)|z'| \\ &= |z'| + |u||z| \\ &= |z'| + |z| . \end{aligned}$$

En conclusion, nous venons de démontrer le résultat suivant :

$$|z + z'| = |z| + |z'| \Longleftrightarrow (z' = 0) \vee (\exists\, u \in \mathbb{R}_+,\ z = uz') .$$

2.⁻ Trigonométrie, le retour

a) Nombres complexes de module 1

Ce paragraphe est, vous l'aurez deviné, dédié à l'étude de l'ensemble des nombres complexes de module 1, à savoir :

$$\mathscr{U} = \{z \in \mathbb{C} \,|\, |z| = 1\} \subset \mathbb{C}^*.$$

Proposition VIII.7. L'ensemble $\mathscr{U}$ est stable par multiplication et inverse.

Démonstration. Le module du produit est le produit des modules, idem pour l'inverse. Or $1 \times 1 = \frac{1}{1} = 1$ ... □

✌ **Remarque VIII.7.**

— Nous dirons dans le chapitre X que $\mathscr{U}$ possède, de par ces stabilités, une structure de groupe.

— Géométriquement, $\mathscr{U}$ correspond au cercle unité de $\mathbb{R}^2$.

— Si $z \in \mathscr{U}$, alors $|z|^2 = z\overline{z} = 1$ donc $\frac{1}{z} = \overline{z}$. En effet, si $z \in \mathbb{C}^*$, on a

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2} .$$

Définition VIII.3. Soit $\theta \in \mathbb{R}$; on **pose**

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)\,.$$

✖ **Attention :** Il s'agit la d'une **notation** ; n'imaginons pas que ces quantités héritent de toutes les propriétés de l'exponentielle. Pensez à la positivité …

➠ **Exemple VIII.3.**
- $e^{i\times 0} = \cos(0) + i\sin(0) = 1$; incidemment, notons que $1 = e^0$, ce qui est rassurant pour notre santé mentale à tous.
- $e^{2i\pi} = 1$;
- $e^{i\frac{\pi}{2}} = i$;
- $e^{i\frac{\pi}{4}} = \frac{\sqrt{2}}{2}(1+i)$.

Proposition VIII.8. On a l'égalité ensembliste suivante :

$$\mathscr{U} = \{e^{i\theta} \,|\, \theta \in \mathbb{R}\}\,.$$

✌ **Remarque VIII.8.** Ce résultat est parfois appelé **paramétrage** du cercle unité ; en effet, nous décrivons ce dernier comme l'image de la fonction

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{C} \\ \theta &\mapsto e^{i\theta} \end{aligned}\,.$$

Démonstration. Montrons deux inclusions.

($\supset$) Soit $\theta \in \mathbb{R}$; alors
$$|e^{i\theta}| = \sqrt{\cos(\theta)^2 + \sin(\theta)^2} = 1$$
donc $e^{i\theta} \in \mathscr{U}$.

($\subset$) Soit $z = x + iy \in \mathscr{U}$; alors $x^2 + y^2 = 1$, ce qui entraîne que le point (x, y) se situe sur le cercle trigonométrique (ou unité) de $\mathbb{R}^2$. De fait, il existe $\theta \in \mathbb{R}$ tel que $x = \cos(\theta)$ et $y = \sin(\theta)$. *In fine*, $z = e^{i\theta}$.

□

Proposition VIII.9. Soient $\theta, \theta' \in \mathbb{R}$. Alors :
(i) $e^{i(\theta+\theta')} = e^{i\theta}e^{i\theta'}$;
(ii) $e^{-i\theta} = \dfrac{1}{e^{i\theta}} = \overline{e^{i\theta}}$;
(iii) $e^{i(\theta-\theta')} = \dfrac{e^{i\theta}}{e^{i\theta'}}$;
(iv) $e^{i\theta} = 1 \Leftrightarrow \theta \equiv 0\,[2\pi]$;
(v) $e^{i\theta} = e^{i\theta'} \Leftrightarrow \theta \equiv \theta'\,[2\pi]$.

Démonstration.

(i)

$$\begin{aligned} e^{i\theta}e^{i\theta'} &= (\cos(\theta)+i\sin(\theta))(\cos(\theta')+i\sin(\theta')) \\ &= (\cos(\theta)\cos(\theta')-\sin(\theta)\sin(\theta'))+i(\cos(\theta)\sin(\theta')+\cos(\theta')\sin(\theta)) \\ &= \cos(\theta+\theta')+i\sin(\theta+\theta') \\ &= e^{i(\theta+\theta')}\,. \end{aligned}$$

(ii) L'égalité avec le conjugué est triviale ; le lien avec l'inverse découle du fait que si $z \in \mathscr{U}$, alors $|z|^2 = z\overline{z} = 1$ donc $\frac{1}{z} = \overline{z}$.

(iii) Découle des points (i) et (ii).

(iv)

$$\begin{aligned} e^{i\theta} = 1 &\Leftrightarrow \cos(\theta)+\sin(\theta) = 1 \\ &\Leftrightarrow (\cos(\theta)=1)\wedge(\sin(\theta)=0) \\ &\Leftrightarrow \theta \equiv 0\,[2\pi]\,. \end{aligned}$$

(v) Découle des points (iii) et (iv).

□

✌ **Remarque VIII.9.** En conséquence, l'application

$$\begin{aligned} \widehat{f} : [0,2\pi[&\to \mathscr{U} \\ \theta &\mapsto e^{i\theta} \end{aligned}\quad .$$

est une bijection. Topologiquement, $\mathscr{U}$ est obtenue en "raccordant" entre elles les extrémités de $[0,2\pi[$.

Le résultat qui suit est attribué à Leonhard Euler (mathématicien suisse, 1707—1783) et possède de nombreuses applications sur lesquelles nous reviendront ultérieurement.

Proposition VIII.10 (Formules d'Euler)**.** Soit $\theta \in \mathbb{R}$; alors :

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

et

$$\sin(\theta) = \frac{e^{i\theta}-e^{-i\theta}}{2i}\,.$$

Démonstration. Découle immédiatement des formules liant parties réelle et imaginaire au conjugué. □

✌ **Remarque VIII.10.** Toute ressemblance avec le (co)sinus hyperbolique est brutalement non fortuite.

La formule qui suit doit son nom au mathématicien français Abraham de Moivre (1667—1754), qui passa la majeure partie de sa vie exilé en Angleterre en raison de sa foi protestante.

Proposition VIII.11 (Formule de Moivre)**.** Soit $\theta \in \mathbb{R}$ et soit $n \in \mathbb{Z}$. Alors :

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)\,.$$

Démonstration. Triviale une fois que l'on a passé le terme de gauche sous forme exponentielle. □

La question que chacun d'entre vous doit (peut–être) se poser est la suivante : ces formules sont bien jolies (et encore...), mais quelle est donc leur utilité ? *Oh, sweet summer child...*

◇ Linéarisation

Le premier usage reconnu des ces formules est la linéarisation d'expressions trigonométriques. Par exemple, si $\theta \in \mathbb{R}$:

$$\begin{aligned}
\cos(\theta)^3 &= \left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right)^3 \\
&= \frac{e^{3i\theta}+3e^{i\theta}+3e^{-i\theta}+e^{-3i\theta}}{8} \\
&= \frac{2\cos(3\theta)+6\cos(\theta)}{8} \\
&= \frac{1}{4}\cos(3\theta)+\frac{3}{4}\cos(\theta)\,.
\end{aligned}$$

De façon plus générale, si $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$, on a :

$$\begin{aligned}
\cos(\theta)^n &= \left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right)^n \\
&= \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\,e^{ik\theta}e^{i(n-k)\theta} \\
&= \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\,e^{i(2k-n)\theta}\,.
\end{aligned}$$

Or, $\cos(\theta)^n \in \mathbb{R}$; il est donc égal à sa partie réelle, *i.e*

$$\begin{aligned}
\cos(\theta)^n &= \mathrm{Re}(\cos(\theta)^n) \\
&= \mathrm{Re}\left(\frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\,e^{i(2k-n)\theta}\right) \\
&= \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\,\mathrm{Re}(e^{i(2k-n)\theta}) \\
&= \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\cos((2k-n)\theta)\,.
\end{aligned}$$

Cette formule (et son analogue pour le sinus) n'est évidemment pas à apprendre par coeur, mais à savoir retrouver.

◇ Délinéarisation

À l'inverse, la formule de Moivre nous donne, si $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$, que :

$$\begin{aligned}\cos(n\theta) + i\sin(n\theta) &= (\cos(\theta) + i\sin(\theta))^n \\ &= \sum_{k=0}^{n} \binom{n}{k} \cos(\theta)^{n-k} i^k \sin(\theta)^k \,.\end{aligned}$$

En passant à la partie réelle, qui ne conservera ici que les termes pairs de la somme, nous obtenons que :

$$\cos(n\theta) = \sum_{0\leq 2p\leq n} \binom{n}{2p} \cos(\theta)^{n-2p}(-1)^p \sin(\theta)^{2p}$$

car $i^{2p} = (i^2)^p = (-1)^p$. Une formule analogue peut évidemment être obtenue pour le sinus en passant à la partie imaginaire.

◇ Calculs de sommes

Pour finir ce musée des horreurs calculatoires, intéressons nous au calcul, pour $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$, des sommes suivantes :

$$S_n = \sum_{k=0}^{n} \cos(k\theta) \quad \text{et} \quad T_n = \sum_{k=0}^{n} \sin(k\theta) \,.$$

L'idée naturelle ici est de calculer

$$\begin{aligned}S_n + iT_n &= \sum_{k=0}^{n} \cos(k\theta) + i\sin(k\theta) \\ &= \sum_{k=0}^{n} e^{ik\theta} \,.\end{aligned}$$

En effet, on reconnait la somme des termes d'une suite géométrique. À supposer que $\theta \not\equiv 0\,[2\pi]$ (auquel cas le calcul me semble abordable sans technique particulière), on a donc :

$$S_n + iT_n = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} \,.$$

Pour simplifier cette fraction, on **factorise par l'angle moitié**, *i.e* :

$$\begin{aligned}\frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} &= \frac{e^{i\frac{(n+1)\theta}{2}}}{e^{i\frac{\theta}{2}}} \frac{e^{-i\frac{(n+1)\theta}{2}} - e^{i\frac{(n+1)\theta}{2}}}{e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}} \\ &= e^{i\frac{n}{2}\theta} \frac{\sin\left(\frac{(n+1)\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)} \,.\end{aligned}$$

De fait :

$$S_n = \mathrm{Re}(S_n + iT_n) = \cos\left(\frac{n}{2}\theta\right) \frac{\sin\left(\frac{(n+1)\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}$$

et

$$T_n = \mathrm{Im}(S_n + iT_n) = \sin\left(\frac{n}{2}\theta\right) \frac{\sin\left(\frac{(n+1)\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)} \,.$$

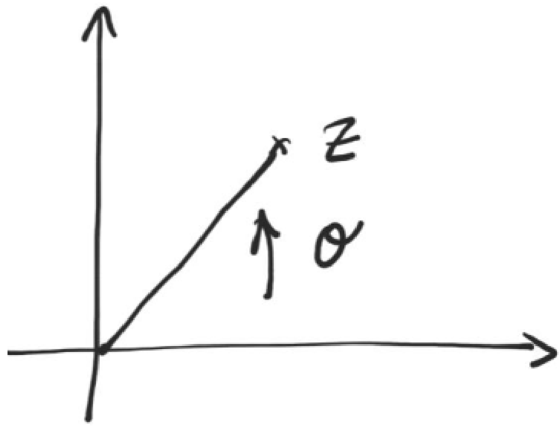
Youpi.

b) Argument

Définition VIII.4. Soit $z \in \mathbb{C}$; on appelle **argument** de z tout $\theta \in \mathbb{R}$ tel que

$$z = |z|e^{i\theta} .$$

✌ **Remarque VIII.11.**
- **Attention :** il n'y a **pas** unicité de l'argument ; par exemple $1 = e^{i\times 0} = e^{2i\pi}$.
- Géométriquement, un argument de $z \in \mathbb{C}$ d'image $A \in \mathbb{R}^2$ est une mesure de l'angle orienté entre l'axe des abscisses et la droite (OA).



- 0 admet tout nombre réel comme argument.

Proposition VIII.12. Soit $z \in \mathbb{C}^*$ et soit θ_0 un argument de z. Alors l'ensemble des arguments de z est :

$$\mathscr{A}_z = \{\theta_0 + 2k\pi \mid k \in \mathbb{Z}\} .$$

Démonstration. Démontrons deux inclusions.

($\supset$) Immédiat par périodicité.

($\subset$) Soit $\eta \in \mathscr{A}_z$; alors $z = |z|e^{i\eta} = |z|e^{i\theta_0}$ donc, comme $|z| \neq 0$, $e^{i\eta} = e^{i\theta_0}$ et donc $\eta \equiv \theta_0\,[2\pi]$, d'où le résultat.

□

✌ **Remarque VIII.12.**
- Un piège classique à éviter : si $z = ae^{i\theta}$, θ n'est pas nécessairement un argument de z. Il pourrait en effet être négatif... Mais pas (trop) de panique : $-1 = e^{i\pi}$.
- Sur un intervalle de longueur inférieure à 2π, l'argument est unique.

Notation. Soit $z \in \mathbb{C}$; on note $\arg(z)$ l'unique élément de $\mathscr{A}_z \cap [0, 2\pi[$. Cette notation est arbitraire ; certains préfèrent donc écrire des choses du style "$\arg(z) \equiv \theta\,[2\pi]$".

➠ **Exemple VIII.4.** Pour déterminer l'argument de $\sqrt{3}+i$, on le divise par son module et on essaie de reconnaître une exponentielle connue :

$$\frac{\sqrt{3}+i}{2} = e^{i\frac{\pi}{6}} .$$

Ainsi, $\arg(\sqrt{3}+i) = \dfrac{\pi}{6}$.

c) Forme trigonométrique

Définition VIII.5. On appelle **forme trigonométrique** d'un nombre complexe z toute écriture de la forme $z = re^{i\theta}$, avec $r \in \mathbb{R}_+$ et $\theta \in \mathbb{R}$.

➠ **Exemple VIII.5.** $-1 = e^{i\pi} = e^{3i\pi}$; notons au passage que cette écriture n'est **pas unique**.

✌ **Remarque VIII.13.**
- Travailler avec des complexes sous forme trigonométrique est en général une bonne idée lorsque l'on veut les multiplier et/ou diviser, mais pas pour les additionner.
- Dans l'écriture sous forme trigonométrique "$z = re^{i\theta}$", on a nécessairement $r = |z|$ et $\theta \in \mathscr{A}_z$.

Proposition VIII.13. Soient $z, z' \in \mathbb{C}$ et $n \in \mathbb{Z}$. Alors :
(i) $\arg(zz') \equiv \arg(z) + \arg(z')\ [2\pi]$;
(ii) $\arg\left(\dfrac{z}{z'}\right) \equiv \arg(z) - \arg(z')\ [2\pi]$;
(iii) $\arg(z^n) \equiv n\arg(z)\ [2\pi]$.

Démonstration. Mettre z et z' sous forme trigonométrique. □

d) Exponentielle complexe

Définition VIII.6. Soit $z \in \mathbb{C}$. On appelle **exponentielle** de z le nombre complexe

$$e^z = e^{\mathrm{Re}(z)}e^{i\mathrm{Im}(z)} .$$

✌ **Remarque VIII.14.** Remarquons que les deux "exponentielles" intervenant dans cette notation sont différentes : la première est un exponentielle réelle, la seconde est de type "$e^{i\theta}$".

Proposition VIII.14. Soient $z, z' \in \mathbb{C}$. Alors :

(i) $e^{\overline{z}} = \overline{e^z}$;

(ii) $|e^z| = e^{\mathrm{Re}(z)}$;

(iii) $e^{z+z'} = e^z e^{z'}$;

(iv) $e^z \neq 0$ et $\dfrac{1}{e^z} = e^{-z}$.

Démonstration. Immédiat en passant par la définition. □

Proposition VIII.15. Soit $z \in \mathbb{C}$; alors $e^z = 1 \Leftrightarrow z \in 2i\pi\mathbb{Z}$.

Remarque VIII.15.
- En particulier, l'exponentielle complexe n'est **PAS** bijective.
- On en déduit que, si $z, z' \in \mathbb{C}$, alors $e^z = e^{z'} \Leftrightarrow z \equiv z' \, [2i\pi]$.

Démonstration.

($\Leftarrow$) Immédiat.

($\Rightarrow$) Si $z = x + iy \in \mathbb{C}$ est tel que $e^z = 1$, alors $e^x e^{iy} = 1$ et donc $e^x = |e^z| = 1$, *ergo* $x = 0$ et $e^{iy} = 1$, d'où le résultat.

□

Proposition VIII.16. La fonction

$$\begin{aligned} \exp : \mathbb{C} &\to \mathbb{C}^* \\ z &\mapsto e^z \end{aligned}$$

est surjective. De plus, pour tout $a \in \mathbb{C}^*$ et $\gamma \in \mathbb{C}$ est tel que $e^\gamma = a$, on a :

$$\exp^{-1}(\{a\}) = \{\gamma + 2ik\pi \mid k \in \mathbb{Z}\} .$$

Démonstration. L'exponentielle complexe est à valeurs dans $\mathbb{C}^*$ d'après le point (iv) de la proposition VIII.14. Soit ensuite $a \in \mathbb{C}^*$, que l'on peut écrire sous forme trigonométrique $a = |a|e^{i\theta}$, avec $\theta \in \mathscr{A}_a$. Alors, si on pose $\gamma = \ln(|a|) + i\theta$, on a :

$$\exp(\gamma) = e^{\ln(|a|)} e^{i\theta} = |a|e^{i\theta} = a$$

donc exp est bien surjective. Enfin, pour tout $\tau \in \mathbb{C}$:

$$a = e^\tau \Leftrightarrow e^\gamma = e^\tau \Leftrightarrow \gamma \equiv \tau \, [2i\pi]$$

□

Exercice VIII.3. Soit $t \in \mathbb{R}$; factoriser $1 + e^{it}$.

➥ **Correction :** *On factorise par l'angle moitié :*

$$\begin{aligned} 1 + e^{it} &= (e^{-i\frac{t}{2}} + e^{i\frac{t}{2}})e^{i\frac{t}{2}} \\ &= 2\cos\left(\frac{t}{2}\right) e^{i\frac{t}{2}} \,. \end{aligned}$$

3.⁻ Équations algébriques

a) Trinômes du second degré

Vous avez logiquement une idée assez précise de la méthode de résolutions d'équations de la forme $ax^2 + bx + c = 0$ avec $a, b, c \in \mathbb{R}$. Le but de ce paragraphe est de mettre en place un analogue complexe de ce procédé.

Définition VIII.7. Soit $z \in \mathbb{C}$. On appelle **racine carrée** tout nombre complexe a tel que $a^2 = z$.

➠ **Exemple VIII.6.** 2 est une racine carrée de 4, i une racine carrée de -1.

Proposition VIII.17. Tout nombre complexe non nul possède exactement deux racines carrées distinctes.

Démonstration. Soit $z \in \mathbb{C}^*$ de forme trigonométrique $z = re^{i\theta}$ $(r > 0)$. Procédons par analyse–synthèse.

Analyse. Soit $a = \rho e^{i\varphi} \in \mathbb{C}$, avec $\rho > 0$ tel que $a^2 = z$, i.e $re^{i\theta} = \rho^2 e^{2i\varphi}$. Alors, en passant au module dans cette égalité, on obtient

$$\rho^2 = r \;\; i.e \;\; \rho = \sqrt{r} \,.$$

En passant ensuite à l'argument dans l'égalité initiale, on a :

$$2\varphi \equiv \theta \, [2\pi]$$

ce qui signifie qu'il existe $k \in \mathbb{Z}$ tel que $2\varphi = \theta + 2k\pi$ et donc, en divisant par 2 :

$$\varphi \equiv \frac{\theta}{2} \, [\pi] \,.$$

Ceci entraîne donc que φ est congru à $\frac{\theta}{2}$ ou $\frac{\theta}{2} + \pi$ modulo 2π. Or, comme $e^{i(\frac{\theta}{2}+\pi)} = -e^{i\frac{\theta}{2}}$ on a *in fine* que :

$$a \in \left\{ \sqrt{r}e^{i\frac{\theta}{2}}, -\sqrt{r}e^{i\frac{\theta}{2}} \right\} \,.$$

Synthèse. Il est immédiat que les deux valeurs proposées ci–dessus sont des racines carrées de z.

□

Exemple VIII.7. Les deux racines carrées de $i = e^{i\frac{\pi}{2}}$ sont donc $\pm e^{i\frac{\pi}{4}}$, *i.e* $\pm\frac{1}{\sqrt{2}}(1+i)$.

Remarque VIII.16. Il est donc aisé de trouver la racine carrée d'un nombre complexe dont on connaît la forme trigonométrique. Si on ne connait que sa forme algébrique $z = x + iy$, on peut faire usagde de la méthode suivante : si $a = X + iY$ est tel que $a^2 = z$ alors :

$$X^2 - Y^2 + 2iXY = x + iy$$

et, commme $|a|^2 = |z|$:

$$(X^2 + Y^2)^2 = x^2 + y^2 .$$

On obtient donc un système de trois équations qu'il nous sera généralement possible de résoudre :

$$\left\{\begin{array}{ccc} X^2 - Y^2 & = & x \\ 2XY & = & y \\ (X^2+Y^2)^2 & = & x^2+y^2 \end{array}\right. .$$

Exemple VIII.8. Pour déterminer les racines de $1 + i$, on tombe sur les équations :

$$\left\{\begin{array}{ccc} X^2 - Y^2 & = & 1 \\ 2XY & = & 1 \\ (X^2+Y^2)^2 & = & 2 \end{array}\right.$$

i.e

$$\left\{\begin{array}{ccc} X^2 - Y^2 & = & 1 \\ 2XY & = & 1 \\ X^2+Y^2 & = & \sqrt{2} \end{array}\right. .$$

En utilisant les deux équations quadratiques, on trouve :

$$X = \pm\sqrt{\frac{\sqrt{2}+1}{2}}, \quad Y = \pm\sqrt{\frac{\sqrt{2}-1}{2}} .$$

La seconde équation nous dit de plus que X et Y ont même signe, ainsi les racines de $1 + i$ sont :

$$\pm\left(\sqrt{\frac{\sqrt{2}+1}{2}} + i\sqrt{\frac{\sqrt{2}-1}{2}}\right) .$$

Proposition VIII.18. Soient $a, b, c \in \mathbb{C}$ tels que $a \neq 0$. On considère l'équation :

$$az^2 + bz + c = 0 \qquad (\mathbf{E}\text{:VIII.3})$$

d'inconnue $z \in \mathbb{C}$ et on pose $\Delta = b^2 - 4ac$. Alors :
- si $\Delta = 0$, l'équation (**E** :VIII.3) admet un unique solution, à savoir :

$$z_0 = \frac{-b}{2a}\,;$$

- si $\Delta \neq 0$, l'équation (**E** :VIII.3) admet exactement deux solutions distinctes, à savoir :

$$z_1 = \frac{b+\delta}{2a}, \quad z_2 = \frac{b-\delta}{2a}$$

  où δ est une racine carrée de Δ.

Remarque VIII.17. Notons que l'extension à $\mathbb{C}$ de la notion de racine carrée simplifie le cas réel en nous évitant de distinguer deux cas selon le signe du discriminant Δ.

Démonstration. Soit $z \in \mathbb{C}$; commençons par remarquer que :

$$\left(z + \frac{b}{2a}\right)^2 = z^2 + \frac{b}{a}z + \frac{b^2}{4a^2}$$

ce qui entraîne que :

$$\begin{aligned} az^2 + bz + c &= a\left(z^2 + \frac{b}{a}z + \frac{c}{a}\right) \\ &= a\left(z + \frac{b}{2a}\right)^2 - a\frac{b^2}{4a^2} + c \\ &= a\left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a}\,. \end{aligned}$$

Soit $\delta \in \mathbb{C}$ une racine carrée de Δ (qui sera donc nulle si Δ l'est). Alors :

$$\begin{aligned} z \text{ est solution de } (\mathbf{E}\text{ :VIII.3}) &\Leftrightarrow a\left(z + \frac{b}{2a}\right)^2 = \frac{\Delta}{4a} \\ &\Leftrightarrow \left(z + \frac{b}{2a}\right)^2 = \frac{\delta^2}{4a^2} \\ &\Leftrightarrow \left(z + \frac{b}{2a}\right)^2 = \left(\frac{\delta}{2a}\right)^2 \\ &\Leftrightarrow z + \frac{b}{2a} = \pm\frac{\delta}{2a} \end{aligned}$$

d'où le résultat. □

Exemple VIII.9. L'équation $z^2 - (1+2i)z + i - 1 = 0$ a pour discriminant $\Delta = 1$ et solutions $1+i$ et i.

✌ **Remarque VIII.18.** Le produit (resp. la somme) des solutions de (**E** :VIII.3) est égal à $\frac{c}{a}$ $\left(\text{resp. égale à } -\frac{b}{a}\right)$. Nous verrons dans le chapitre XV qu'il ne s'agit pas d'une coïncidence.

Proposition VIII.19. Soient $z, z', s, p \in \mathbb{C}$. Alors :

$$\begin{cases} z + z' &= s \\ zz' &= p \end{cases} \iff z \text{ et } z' \text{ sont solutions de } X^2 - sX + p = 0\,.$$

Démonstration. Le sens indirect provient de la proposition VIII.18. Pour le sens direct, il suffit de développer $(X - z)(X - z')$, qui a clairement pour racines z et z' et est égal à $X^2 - sX + p$ □

b) Racines n–ièmes

Dans tout ce paragraphe, nous considérons comme fixé un entier $n \in \mathbb{N}^*$ et nous intéressons à la notion de racine n–ième (complexe) définie ci–ensuite.

Définition VIII.8. Soit $z \in \mathbb{C}$; on appelle **racine n–ième** de z tout nombre complexe a tel que $a^n = z$.

Vocabulaire. Une racine n–ième de 1 est appelée **racine n–ième de l'unité**.

Notation. On note $\mathscr{U}_n$ l'ensemble des racines n–ièmes de l'unité ; *i.e*

$$\mathscr{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}\,.$$

✌ **Remarque VIII.19.** Il est aisé de démontrer que :
— $\mathscr{U}_n \subset \mathscr{U} \subset \mathbb{C}^*$ (passer au module) ;
— $\mathscr{U}_n$ est stable par produit et passage à l'inverse.

Proposition VIII.20. Posons, pour $k \in \mathbb{N}$, $\xi_k = e^{\frac{2ik\pi}{n}}$. Alors :

$$\mathscr{U}_n = \{\xi_0, \dots, \xi_{n-1}\}\,.$$

En particulier, $\mathscr{U}_n$ est un ensemble contenant exactement n éléments.

✌ **Remarque VIII.20.** Notons que $\xi_0 = \xi_n = 1$.

Démonstration.

($\supset$) Immédiat en remarquant que, pour tout $k \in \mathbb{N}$:

$$\xi_k^n = e^{2ik\pi} = 1\,.$$

($\subset$) Soit $z \in \mathscr{U}_n$, i.e $z^n = 1$. Ceci entraîne en particulier que $z \in \mathscr{U}$, *i.e* $|z| = 1$, donc il existe $\theta \in \mathbb{R}$ tel que $z = e^{i\theta}$. Le fait que $z^n = 1$ a pour conséquence que $n\theta \equiv 0\,[2\pi]$, *ergo* il existe $k \in [\![0, n-1]\!]$ tel que $e^{i\theta} = e^{\frac{2ik\pi}{n}}$. Il ne nous reste plus pour conclure qu'à remarquer que $\xi_0, \dots, \xi_{n-1}$ sont deux à deux distincts.

□

Remarque VIII.21.

— Si on pose $\xi e^{\frac{2i\pi}{n}}$, alors pour tout $k \in \mathbb{N}$, $\xi_k = \xi^k$.

— Ceci a pour conséquence que :

$$\sum_{k=0}^{n-1} \xi_k = \frac{\xi^n - 1}{\xi - 1} = 0$$

car $\xi^n = \xi_n = 1$.

Exemple VIII.10.

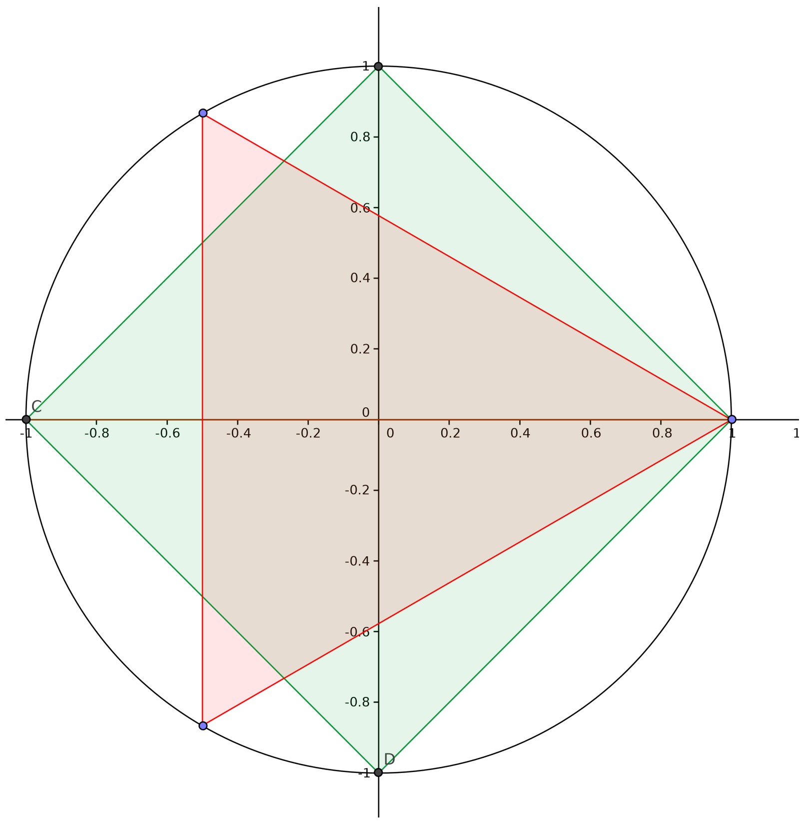
— $\mathscr{U}_1 = \{1\}$;

— $\mathscr{U}_2 = \{-1, 1\}$;

— en posant $j = e^{\frac{2i\pi}{3}}$, on a $\mathscr{U}_3 = \{1, j, j^2\}$. Remarquons que l'on a de fait $1 + j + j^2 = 0$ et $\overline{j} = j^2$;

— $\mathscr{U}_4 = \{-1, 1, -i, i\}$.

Géométriquement, les points de $\mathscr{U}_n$ forment un polygone régulier à n côtés, comme illustré sur la figure ci–ensuite pour $n = 3, 4$.



Proposition VIII.21. Soit $z \in \mathbb{C}^*$ de forme trigonométrique $z = re^{i\theta}$. Alors les racines n–ièmes de z sont exactement les

$$\sqrt[n]{r}e^{\frac{i\theta}{n}}\xi_k$$

pour $k \in [\![0, n-1]\!]$.

Démonstration. Cela provient du fait que si u et u' sont deux racines n–ièmes de z, alors $\frac{u}{u'} \in \mathscr{U}_n$. □

Exemple VIII.11. Les racines cubiques de 2 sont $\sqrt[3]{2}$, $j\sqrt[3]{2}$ et $j^2\sqrt[3]{2}$.

4.– Transformations du plan

Le but de ce paragraphe est de construire un "dictionnaire" reliant transformations géométriques du plan et applications de $\mathbb{C}$ dans $\mathbb{C}$. Rappelons que la correspondance entre ces deux ensembles est donnée par l'application

$$\begin{aligned} \varphi : \mathbb{C} &\to \mathbb{R}^2 \\ z &\mapsto (\mathrm{Re}(z), \mathrm{Im}(z)) \end{aligned}$$

de réciproque

$$\begin{aligned} \psi : \mathbb{R}^2 &\to \mathbb{C} \\ (x, y) &\mapsto x + iy \quad . \end{aligned}$$

a) Translations, homothéties

Soit $u = \varphi(z_o) \in \mathbb{R}^2$; on appelle **translation de vecteur** u l'application

$$\begin{aligned} \tau_u : \mathbb{R}^2 &\to \mathbb{R}^2 \\ \varphi(z) &\mapsto \varphi(z + z_0) \ . \end{aligned}$$

Si $\Omega = \varphi(\omega) \in \mathbb{R}^2$ et $\lambda \in \mathbb{R}$, on appelle **homothétie de centre Ω et de rapport λ** l'application

$$\begin{aligned} \mu_{\Omega,\lambda} : \mathbb{R}^2 &\to \mathbb{R}^2 \\ \varphi(z) &\mapsto \varphi(\omega + \lambda(z - \omega)) \ . \end{aligned}$$

b) Rotations

Soit $\theta \in \mathbb{R}$ et soit $\Omega = \varphi(\omega) \in \mathbb{R}^2$; on appelle **rotation de centre Ω et d'angle θ** l'application

$$\begin{aligned} \rho_{\Omega,\theta} : \mathbb{R}^2 &\to \mathbb{R}^2 \\ \varphi(z) &\mapsto \varphi\left(\omega + (z - \omega)e^{i\theta}\right) \ . \end{aligned}$$

c) Similitudes

Définition VIII.9. Soient $\theta \in \mathbb{R}$, $\lambda \in \mathbb{R}^*$ et $\Omega = \varphi(\omega) \in \mathbb{R}^2$. On appelle **similitude (directe) de centre Ω, de rapport λ et d'angle θ** l'application :

$$\begin{aligned} \mathscr{S}_{\Omega,\lambda,\theta} : \mathbb{R}^2 &\mapsto \mathbb{R}^2 \\ \varphi(z) &\mapsto \varphi\left(\omega + \lambda(z-\omega)e^{i\theta}\right) \ . \end{aligned}$$

✌ **Remarque VIII.22.** Une similitude directe est donc la composée d'une rotation et d'une homothétie. Il s'agit en fait des applications de la forme

$$\varphi(z) \mapsto \varphi(az+b)$$

avec $a, b \in \mathbb{C}$ tels que $a \neq 0$.

Proposition VIII.22. Les similitudes directes conservent les angles orientés.

Démonstration. Les rotations et les homothéties conservent les angles orientés, d'où le résultat. □

Chapitre IX

Suites numériques

1.$^-$ Généralités

a) Notion de suite numérique

Rappelons la définition suivante, à laquelle nous avions fait allusion dans le chapitre V.

Définition IX.1. On appelle **suite (numérique)** toute application $u : [\![n_0, \infty[\![\to \mathbb{K}$, où :
- $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$;
- $n_0 \in \mathbb{N}$ est appelé **rang initial** de la suite.

Notation. La suite u peut être notée $(u_n)_{n \geq n_0}$ (ou simplement $(u_n)_n$ si $n_0 = 0$) ; ses **termes** $u(n)$, pour $n \geq n_0$, peuvent être notés u_n. Il est **hors de question** d'écrire "la suite u_n", ou autres abominations du genre.

➠ **Exemple IX.1.** $(n^2 + 1)_{n \geq 0}$, $\left(\dfrac{1}{n}\right)_{n \geq 1}$.

Dans la suite du chapitre, nous énoncerons et démontrerons les résultats en partant du principe que $n_0 = 0$. Il s'agit d'une simplification arbitraire et aisée à généraliser à toute valeur du rang initial ; nous laissons le soin de le faire à notre bienveillant lecteur. Nous nous intéresserons également dans un premier temps **uniquement aux suites à valeurs réelles** ; la généralisation au cas complexe des résultats ne sera par contre en aucun cas automatique, et nous lui dédions la partie 4.$^-$ de ce chapitre.

b) Suites bornées

Définition IX.2. Une suite réelle $u \in \mathbb{R}^{\mathbb{N}}$ est dite :
- **majorée** si $\exists\, M \in \mathbb{R}$, $\forall n \in \mathbb{N}$, $u_n \leq M$;
- **minorée** si $\exists\, m \in \mathbb{R}$, $\forall n \in \mathbb{N}$, $u_n \geq m$;
- **bornée** si elle est majorée **et** minorée.

Remarque IX.1. Une suite $(u_n)_n \in \mathbb{R}^{\mathbb{N}}$ est donc bornée si et seulement si la suite $(|u_n|)_n$ est majorée.

Exemple IX.2. La suite $(n^2+1)_{n\geq 0}$ n'est pas majorée, mais minorée par 1, 0 ou même $-\dfrac{42\pi}{13}$.

c) Variations

Définition IX.3. Une suite réelle $u \in \mathbb{R}^{\mathbb{N}}$ est dite **croissante** (resp. **décroissante**) si :
$$\forall n, n' \in \mathbb{N} \text{ tels que } n \leq n',\ u_n \leq u_{n'} \text{ (resp. } u_n \geq u_{n'}).$$
Si l'inégalité ci–dessus est stricte, on parle de monotonie stricte.

Remarque IX.2. On peut (doit !) démontrer (par récurrence) que ceci est équivalent à l'étude du signe de $u_{n+1} - u_n$ pour tout $n \in \mathbb{N}$.

Exemple IX.3. Étudions la monotonie de la suite $\left(\dfrac{n^2+1}{2n+1}\right)_n$. Soit $n \in \mathbb{N}$; alors :

$$\begin{aligned}
u_{n+1} - u_n &= \frac{(n+1)^2+1}{2(n+1)+1} - \frac{n^2+1}{2n+1} \\
&= \frac{(2n+1)((n+1)^2+1) - (2n+3)(n^2+1)}{(2n+3)(2n+1)} \\
&= \frac{(2n+1)(n^2+1) + (2n+1)^2 - (2n+1)(n^2+1) - 2(n^2+1)}{(2n+3)(2n+1)} \\
&= \frac{(2n+1)^2 - 2(n^2+1)}{(2n+3)(2n+1)} \\
&= \frac{4n^2+4n+1-2n^2-2}{(2n+3)(2n+1)} \\
&= \frac{2n^2+4n-1}{(2n+3)(2n+1)} \,.
\end{aligned}$$

Le numérateur est un trinôme du second degré en n dont l'unique racine positive est $-1+\sqrt{\dfrac{3}{2}} \cong 0.225$. De fait, $(u_n)_n$ est croissante à partir de $n=1$.

Définition IX.4. Une suite $u \in \mathbb{R}^{\mathbb{N}}$ est dite **stationnaire** si
$$\exists\, c \in \mathbb{R},\ \exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = c\,.$$

Remarque IX.3. Une suite stationnaire est donc une suite constante à partir d'un certain rang.

➠ **Exemple IX.4.** La suite $\left(\left\lfloor \frac{1}{n} \right\rfloor\right)_{n\geq 1}$ est stationnaire ; elle en effet nulle à partir de $n = 2$.

Proposition IX.1. Toute suite décroissante d'entiers naturels est stationnaire.

Démonstration. Soit $u \in \mathbb{N}^{\mathbb{N}}$ une suite décroissante. On pose :

$$\mathscr{E} = \{u_n \,|\, n \geq 0\}$$

l'ensemble des termes de la suite. Il s'agit d'une partie de $\mathbb{N}$ non vide donc, par axiome D, $\mathscr{E}$ admet un minimum, soit $m_0 = u_{n_0}$. Par décroissance, pour tout $n \geq n_0$, $u_n \leq m_0$, ce qui entraîne par minimalité que $u_n = m_0$. La suite est donc bien stationnaire à partir du rang n_0 (au moins). □

2.⁻ Limite d'une suite

a) Limite réelle

Définition IX.5. Soit $u \in \mathbb{R}^{\mathbb{N}}$ et $\ell \in \mathbb{R}$; on dit que **la suite** $(u_n)_n$ **converge vers** ℓ si :

$$\forall \varepsilon > 0,\ \exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n - \ell| \leq \varepsilon\,.$$

Le réel ℓ est alors appelé **limite** de la suite u.

Notation. On utilisera indifférement $u_n \xrightarrow[n\to\infty]{} \ell$, $u_n \to \ell$, $\lim\limits_{n\to\infty} u_n = \ell$ ou $\lim u_n = \ell$.

Avant d'aller plus loin, analysons la définition de limite. Elle se découpe en trois parties :
- "$\forall\varepsilon > 0$" nous indique que ce qui va suivre sera vrai quel que soit le ε choisi. Ce dernier quantifiera la précision de notre approximation ;
- "$\exists\, N \in \mathbb{N},\ \forall n \geq N,$ " signifie que la propriété est vraie à partir d'un certain rang N ;
- "$|u_n - \ell| \leq \varepsilon$" nous dit que l'écart entre u_n et ℓ sera inférieur à la précision choisie.

In fine, la définition peut se reformuler de la façon suivante : **quelle que soit la précision choisie,** u_n **sera plus proche de** ℓ **que cette dernière pour** n **assez grand**.

➠ **Exemple IX.5.** Soit $\varepsilon > 0$; alors pour tout $n \geq \left\lfloor \frac{1}{\varepsilon} \right\rfloor + 1$, $\frac{1}{n} \leq \varepsilon$. Ainsi :

$$\frac{1}{n} \xrightarrow[n\to\infty]{} 0\,.$$

Définition IX.6. Une suite convergeant vers une limite **RÉELLE** est dite **convergente.** Dans le cas contraire, elle est dite **divergente**.

➠ **Exemple IX.6.** La suite $\left(\dfrac{1}{n}\right) n \geq 1$ est convergente (vers 0). Nous verrons que la suite $((-1)^n)_n$ est divergente (sans limite), de même que $(n^2+1)_n$ (limite infinie).

Proposition IX.2. Une suite convergente admet exactement une limite.

Démonstration. Soit $u \in \mathbb{R}^n$ une suite convergente dont on suppose qu'elle admet deux limites ℓ et ℓ'. Soit $\varepsilon > 0$; alors, par définition de limite :
- $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ ,|u_n - \ell| \leq \varepsilon$;
- $\exists N' \in \mathbb{N},\ \forall n \geq N',\ |u_n - \ell'| \leq \varepsilon$.

Ainsi, si on fixe $n \geq \max(N, N')$, on a :

$$\begin{aligned} |\ell - \ell'| &= |\ell - \ell' + u_n - u_n| \\ &\leq |u_n - \ell| + |u_n - \ell'| \\ &\leq 2\varepsilon\ . \end{aligned}$$

En posant $\varepsilon = \dfrac{1}{3}|\ell - \ell'|$ on obtient alors :

$$|\ell - \ell'| \leq \frac{2}{3}|\ell - \ell'|$$

ce qui n'est possible que si $|\ell - \ell'| = 0$, d'où le résultat.

□

Proposition IX.3. Une suite convergente est bornée.

Démonstration. Soit $u \in \mathbb{R}^{\mathbb{N}}$ convergeant vers $\ell \in \mathbb{R}$. Alors, en appliquant la définition de limite à $\varepsilon = 1$, on obtient :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n - \ell| \leq 1\ .$$

La seconde inégalité triangulaire nous livre alors que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ ||u_n| - |\ell|| \leq |u_n - \ell| \leq 1$$

et donc $|u_n| \in [1 - |\ell|, 1 + |\ell|]$. *In fine*, la suite $(u_n)_n$ est bornée par

$$\max(|u_0|, \ldots, |u_{N-1}|, 1 + |\ell|)\ .$$

□

✖ **Attention :** la réciproque est **fausse**, comme nous le verrons un peu plus tard. La suite $((-1)^n)_n$ est en effet bornée et divergente.

Proposition IX.4. Soit $u \in \mathbb{R}^{\mathbb{N}}$; alors :

(i) le produit d'une suite bornée par une suite convergeant vers 0 converge vers 0 ;

(ii) **si** $u_n \to \ell \in \mathbb{R}$, **alors** $|u_n| \to |\ell|$;

(iii) $u_n \to 0 \Leftrightarrow |u_n| \to 0$.

Démonstration.

(i) Supposons que $u_n \to 0$ et soit $v \in \mathbb{R}^{\mathbb{N}}$ une suite bornée par $M \in \mathbb{R}_+$. Alors, par définition de limite, pour tout $\varepsilon > 0$:

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n - 0| \leq \frac{\varepsilon}{M}$$

et donc

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n v_n| = |u_n||v_n| \leq M \times \frac{\varepsilon}{M} = \varepsilon$$

d'où $u_n v_n \to 0$.

(ii) Il suffit de remarquer que, pour tout $n \geq 0$, $||u_n| - |\ell|| \leq |u_n - \ell|$ et de réinjecter cela dans la définition de limite pour la suite $(u_n)_n$.

(iii) Immédiat via la définition de limite.

□

Proposition IX.5. Soient $u, v \in \mathbb{R}^{\mathbb{N}}$ telles que :
- $v_n \to 0$;
- $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n| \leq v_n$.

Alors $u_n \to 0$.

Démonstration. Soit $\varepsilon > 0$; alors par définition de limite, il existe $N' \in \mathbb{N}$ tel que pour tout $n \geq N$, $|v_n| = v_n \geq \varepsilon$. Ainsi, pour tout $n \geq \max(N, N')$, on a :

$$|u_n| \leq v_n \leq |v_n| \leq \varepsilon$$

d'où le résultat. □

b) "Limite" infinie

Définition IX.7. Soit $u \in \mathbb{R}^{\mathbb{N}}$; on dit que u **tend (ou diverge) vers** $+\infty$ **(resp.** $-\infty$**)** si :

$$\forall M \in \mathbb{R},\ \exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n \geq M \text{ (resp. } u_n \leq M).$$

Notation. $u_n \to \pm\infty$ ou $\lim u_n = \pm\infty$.

✖ **Attention :** Une suite tendant vers $\pm\infty$ n'est **pas** convergente.

➠ **Exemple IX.7.** Soit $q \in]1, \infty[$ et posons $h = q - 1$. Alors, par binôme de Newton, pour tout $n \in \mathbb{N}$:

$$q^n = (1+h)^n \geq 1 + nh\,.$$

Ainsi, si $M \in \mathbb{R}$, alors $q^n \geq M$ dés que $n \geq \dfrac{M-1}{h}$; de fait $q^n \to \infty$.

Proposition IX.6. Une suite admet au plus une limite, finie ou infinie.

Démonstration. Soit $u \in \mathbb{R}^{\mathbb{N}}$. Si u est convergente, elle est bornée et donc ne peut admettre une limite infinie et possède une unique limite finie par la proposition IX.2. Dans le cas où $u_n \to \infty$, alors elle ne peut être convergente par définition et $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n \geq 1$ donc u_n ne peut tendre vers $-\infty$. On traite de la même façon le cas $u_n \to -\infty$. □

c) Opérations sur les limites

Commençons ce paragraphe par rappeler les opérations "étendues" que nous avons mis en place sur $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ dans le chapitre VI ; elles nous seront fort utiles pour l'énoncé des résultats de ce paragraphe.

$+$	$-\infty$	$y \in \mathbb{R}$	$+\infty$
$-\infty$	$-\infty$	$-\infty$	**?**
$x \in \mathbb{R}$	$-\infty$	$x+y$	$+\infty$
$+\infty$	**?**	$+\infty$	$+\infty$

$\times$	$-\infty$	$y \in \mathbb{R}_+^*$	0	$y \in \mathbb{R}_+^*$	$+\infty$
$-\infty$	$+\infty$	$+\infty$	**?**	$-\infty$	$-\infty$
$x \in \mathbb{R}_-^*$	$+\infty$	xy	0	xy	$-\infty$
0	**?**	0	0	0	**?**
$x \in \mathbb{R}_+^*$	$-\infty$	xy	0	xy	$+\infty$
$+\infty$	$-\infty$	$-\infty$	**?**	$+\infty$	$+\infty$

Proposition IX.7. Soient $u, v \in \mathbb{R}^{\mathbb{N}}$ et $\ell, \ell' \in \overline{\mathbb{R}}$ tels que $u_n \to \ell$ et $v_n \to \ell'$. Alors, à condition de ne pas tomber dans un cas d'indétermination, on a :

(i) $u_n + v_n \to \ell + \ell'$;

(ii) pour tout $\lambda \in \mathbb{R}$, $\lambda u_n \to \lambda \ell$;

(iii) $u_n v_n \to \ell \ell'$.

Démonstration. Traitons par exemple le cas (i) lorsque $\ell, \ell' \in \mathbb{R}$. Soit $\varepsilon > 0$; alors, quitte a prendre le maximum des deux rangs apparaissant dans la définition de limite pour u et v comme nous l'avons déjà fait auparavant, on a :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ \left(|u_n - \ell| \leq \frac{\varepsilon}{2}\right) \wedge \left(|v_n - \ell'| \leq \frac{\varepsilon}{2}\right)$$

et donc, pour tout $n \geq N$:

$$\begin{aligned} |u_n + v_n - (\ell + \ell')| &\leq |u_n - \ell| + |v_n - \ell'| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

d'où le résultat. Les autres cas se traitent selon un procédé similaire. □

Proposition IX.8. Soit $u \in \mathbb{R}^{\mathbb{N}}$ et soit $\ell \in]0, \infty]$ tel que $u_n \to \ell$. Alors :

(i) $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n > 0$;

(ii) $\dfrac{1}{u_n} \to \dfrac{1}{\ell}$, avec la convention que $\dfrac{1}{\infty} = 0$.

Démonstration.

Cas 1 : $\ell = \infty$**.** Soit $\varepsilon > 0$; alors par définition de limite :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n \geq \frac{1}{\varepsilon}.$$

Ainsi, pour $n \geq N$, $u_n > 0$ et :

$$0 < \frac{1}{u_n} \leq \varepsilon$$

ergo $\dfrac{1}{u_n} \to 0$.

Cas 2 : $\ell \in \mathbb{R}_+^*$**.** Alors, par définition de limite :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n - \ell| \leq \frac{\ell}{2}\,.$$

Si $n \geq N$, u_n est alors strictement positif car compris entre $\dfrac{\ell}{2}$ et $\dfrac{3\ell}{2}$. En effet :

$$|u_n - \ell| \leq \frac{\ell}{2} \Longleftrightarrow u_n - \ell \in \left[-\frac{\ell}{2}, \frac{\ell}{2}\right]\,.$$

De plus, toujours pour $n \geq N$:

$$\begin{aligned}\left|\frac{1}{u_n} - \frac{1}{\ell}\right| &= \left|\frac{u_n - \ell}{u_n \ell}\right| \\ &= \frac{|u_n - \ell|}{|u_n|\ell} \\ &\leq \frac{2|u_n - \ell|}{\ell^2}\end{aligned}$$

d'où le résultat comme $u_n \to \ell$.

□

➠ **Exemple IX.8.** Ceci implique (entre autres) que si $q \in]0, 1[$, $q^n \to 0$; en effet $\dfrac{1}{q} > 1$ et donc $\left(\dfrac{1}{q}\right)^n \to \infty$.

Proposition IX.9. Soit $u \in \mathbb{R}^{\mathbb{N}}$ une suite telle que :
- $u_n \to 0$;
- $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n > 0$.

Alors :

$$\frac{1}{u_n} \to \infty\,.$$

Démonstration. Analogue à celle de la proposition précédente. □

✌ **Remarque IX.4.** Il existe un résultat analogue pour les suites strictement négatives à partir d'un certain rang convergeant vers 0.

✖ **ATTENTION :** l'hypothèse de stricte positivité est **essentielle**. En effet, la suite $\left(\frac{(-1)^n}{n}\right)_{n\geq 1}$ converge vers 0 en tant que produit d'une suite bornée par une suite de limite nulle (proposition IX.4), mais son inverse, la suite $((-1)^n n)_{n\geq 1}$ n'a pas de limite, comme nous le verrons un peu plus tard.

Le théorème suivant sera énoncé plus proprement et démontré au chapitre XI, mais nous en donnons une version simplifiée en avant–première afin de simplifier votre travail en exercices. Ne me remerciez pas...

Théorème IX.10 (Composition des limites)**.**
Soit $u \in \mathbb{R}^{\mathbb{N}}$ admettant pour limite $\ell \in \overline{\mathbb{R}}$ et soit f une fonction définie "autour de ℓ" telle que $f(x) \xrightarrow[x\to\ell]{} \ell' \in \overline{\mathbb{R}}$. Alors :

$$f(u_n) \to \ell' .$$

➠ **Exemple IX.9.**
- $e^{-n^2} \to 0$, $e^{\frac{1}{n}} \to 1$.
- Posons $u_0 = 0$ et, pour $n \in \mathbb{N}$, $u_{n+1} = \sqrt{1+u_n}$. Alors, à supposer que cette suite admette une limite $\ell \in \mathbb{R}$ (nous le démontrerons plus tard dans l'année), l'unicité de la limite (proposition IX.2) entraîne que $\ell = \sqrt{1+\ell}$ et donc $\ell = \dfrac{1+\sqrt{5}}{2}$ (ℓ doit être positive car les termes de la suite u le sont et $\ell^2 = 1+\ell$).

d) Limites et ordre

Proposition IX.11. Soient $u, v \in \mathbb{R}^{\mathbb{N}}$ convergeant respectivement vers $\ell, \ell' \in \mathbb{R}$ telles que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n \leq v_n .$$

Alors :

$$\ell \leq \ell' .$$

Démonstration. Soit $\varepsilon > 0$; pour n assez grand (marre d'écrire les max de choses et autres, vous me pardonnerez probablement ; dans le cas contraire, même tarif) $|u_n - \ell| \leq \varepsilon$ et $|v_n - \ell'| \leq \varepsilon$. Ainsi :

$$\ell - \varepsilon \leq u_n \leq v_n \leq \ell' + \varepsilon$$

toujours pour n suffisamment grand. On en déduit que $\ell \leq \ell' + 2\varepsilon$ pour tout $\varepsilon > 0$, ce qui implique que $\ell - \ell'$ minore $\mathbb{R}_+^*$ et donc $\ell - \ell' \leq \inf \mathbb{R}_+^* = 0$, d'où le résultat. □

✖ **Attention :** il n'existe **pas** de résultat analogue pour les inégalités strictes ; en effet, pour tout $n \geq 0$, $n < n+1$ et pourtant les limites de ces deux quantités sont égales.

Théorème IX.12 (Encadrement des limites)**.**
Soient $u, v, w \in \mathbb{R}^{\mathbb{N}}$ et soit $\ell \in \mathbb{R}$ tels que :
— u et w convergent vers ℓ ;
— $\exists N \in \mathbb{N}, \forall n \geq N, u_n \leq v_n \leq w_n$.
Alors :
$$v_n \to \ell \, .$$

Démonstration. Soit $\varepsilon > 0$, alors quitte à prendre le maximum des rangs associés à u et w :
$$\exists N' \in \mathbb{N}, \forall n \geq N', (|u_n - \ell| \leq \varepsilon) \wedge (|w_n - \ell| \leq \varepsilon) \, .$$
Ainsi, pour $n \geq \max(N, N')$, on a l'inégalité :
$$\ell - \varepsilon \leq u_n \leq v_n \leq w_n \leq \ell + \varepsilon$$
qui nous permet de conclure que $|u_n - \ell| \leq \varepsilon$ et d'obtenir la convergence désirée. □

➠ **Exemple IX.10.** La suite $\left(\frac{(-1)^n + 2}{n}\right)_{n \geq 1}$ converge vers 0 ; en effet, pour tout $n \geq 1$ on a :
$$\frac{1}{n} \leq \frac{(-1)^n + 2}{n} \leq \frac{3}{n} \, .$$

Nous disposons également d'une "version infinie" de ce théorème, que nous énonçons dans le cas "positif" (limite $+\infty$) en laissant au lecteur le soin d'énoncer et démontrer son analogue "négatif".

Proposition IX.13. Soient $u, v \in \mathbb{R}^{\mathbb{N}}$ telles que :
— $u_n \to \infty$;
— $\exists N \in \mathbb{N}, \forall n \geq N, u_n \leq v_n$.
Alors :
$$v_n \to \infty \, .$$

e) Suites extraites

Définition IX.8. Soit E un ensemble et soit $u \in E^{\mathbb{N}}$; on appelle **suite extraite** de u toute suite de la forme $(u_{\varphi(n)})_n$, où $\varphi : \mathbb{N} \to \mathbb{N}$ est une fonction strictement croissante.

✌ **Remarque IX.5.** Extraire une suite revient donc à ne "garder" que les termes indexés par l'image $\varphi(\mathbb{N})$.

➠ **Exemple IX.11.**

— $(u_{2n})_n$ et $(u_{2n+1})_n$ sont des suites extraites de u, ne contenant respectivement que ses termes d'indices pairs et impairs. De même, $(u_{n^2})_n$ est une suite extraite de u.
— Si $u = ((-1)^n)_n$, $(u_{2n})_n$ est la suite constante égale à 1.

Proposition IX.14. Soient $u \in \mathbb{R}^{\mathbb{N}}$ une suite admettant une limite $\ell \in \overline{\mathbb{R}}$; alors ses suites extraites tendent également vers ℓ.

Démonstration. Traitons par exemple le cas $\ell \in \mathbb{R}$. Soit $\varphi : \mathbb{N} \to \mathbb{N}$ est une fonction strictement croissante ; il est alors aisé de démontrer par récurrence que :

$$\forall n \in \mathbb{N},\ \varphi(n) \geq n\ .$$

Soit $\varepsilon > 0$, on a alors par définition de limite que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n - \ell| \leq \varepsilon$$

et donc

$$\forall n \geq N,\ |u_{\varphi(n)} - \ell| \leq \varepsilon$$

car $\phi(n) \geq n \geq N$, d'où le résultat. □

Exemple IX.12. Depuis le temps que nous vous le promettions : la suite $((-1)^n)_n$ n'admet pas de limite. En effet, la suite de ses termes d'indices pairs est constante égale à 1 et celle de ses termes impairs à -1 ; de fait elles n'ont pas même limite et donc la suite initiale ne peut admettre de limite.

Remarque IX.6. L'exemple ci–dessus est loin d'être anecdotique ; cette proposition nous servira principalement démontrer que des suites n'admettent **pas** de limite.

Proposition IX.15. Soit $u \in \mathbb{R}^{\mathbb{N}}$ une suite telle que $(u_{2n})_n$ et $(u_{2n+1})_n$ tendent vers une même limite ℓ. Alors :

$$u_n \to \ell\ .$$

Remarque IX.7. Le résultat reste vrai si $(u_{3n})_n$, $(u_{3n+1})_n$ et $(u_{3n+2})_n$ tendent vers la même limite (et ainsi de suite).

Démonstration. Traitons le cas $\ell \in \mathbb{R}$ et fixons $\varepsilon > 0$. Alors, quitte à remplacer N par le maximum des rangs associés aux extractions paire et impaire, on a :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ (|u_{2n} - \ell| \leq \varepsilon) \wedge (|u_{2n+1} - \ell| \leq \varepsilon)\ .$$

De fait, pour tout $n \geq 2N+1$, on a $|u_n - \ell| \leq \varepsilon$, d'où le résultat. □

3.⁻ Théorèmes d'existence de limites

a) Le cas monotone

Proposition IX.16. Soit $u \in \mathbb{R}^{\mathbb{N}}$ une suite croissante. Alors :
- si u est majorée, u converge vers $\sup\limits_{n\in\mathbb{N}}(u_n)$;
- sinon, u diverge vers $+\infty$.

✌ **Remarque IX.8.** Un résultat symétrique existe pour les suites décroissantes.

Démonstration.

Cas 1 : u **est majorée.** Alors l'ensemble

$$\mathscr{E} = \{u_n \,|\, n \in \mathbb{N}\}$$

est une partie de $\mathbb{R}$ non vide et majorée : elle admet donc une borne supérieure ℓ. Soit $\varepsilon > 0$; par définition de supremum, $\ell - \varepsilon$ ne majore pas $\mathscr{E}$ et donc il existe $N \in \mathbb{N}$ tel que $u_N \geq \ell - \varepsilon$. Or, $u_N \leq \ell$ donc $|u_N - \ell| \leq \varepsilon$ et comme la suite u est croissante, pour tout $n \geq N$, $u_n \geq u_N \geq \ell - \varepsilon$, *ergo*

$$\forall n \geq N,\ |u_n - \ell| \leq \varepsilon\,,$$

d'où le résultat.

Cas 2 : u **n'est pas majorée.** Soit $M \in \mathbb{R}$; comme u n'est pas majorée, il existe $N \in \mathbb{N}$ tel que $u_N \geq M$. Or, u est croissante donc :

$$\forall n \geq N,\ u_n \geq u_N \geq M\,,$$

d'où le résultat.

□

b) Adjacence

Définition IX.9. Deux suites $u, v \in \mathbb{R}^{\mathbb{N}}$ sont dites **adjacentes** si :
- l'une est croissante, l'autre décroissante ;
- $u_n - v_n \to 0$.

➠ **Exemple IX.13.** Posons, pour $n \geq 1$:

$$u_n = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k},\ v_n = u_{2n} \text{ et } w_n = u_{2n+1}\,.$$

Soit $n \geq 1$; alors :

$$\begin{aligned} v_{n+1} - v_n &= u_{2n+2} - u_{2n} \\ &= \sum_{k=1}^{2n+2} \frac{(-1)^{k+1}}{k} - \sum_{k=1}^{2n} \frac{(-1)^{k+1}}{k} \\ &= -\frac{1}{2n+2} + \frac{1}{2n+1} && \geq 0 \end{aligned}$$

donc v est croissante ; on montre de même que w est décroissante et

$$\begin{aligned}\forall n \geq 1,\ v_n - w_n &= \frac{-1}{2n+1} \\ &\to 0\end{aligned}$$

donc les suites u et v sont adjacentes.

Proposition IX.17. Deux suites adjacentes convergent vers la même limite.

Démonstration. Soient $(u_n)_n$ et $(v_n)_n$ deux suites adjacentes, la première étant croissante, la seconde décroissante. Il est clair que dans cette configuration, comme $u_n - v_n \to 0$, pour tout $n \in \mathbb{N}$, $u_0 \leq u_n \leq v_n \leq v_0$. La suite $(u_n)_n$ est donc croissante et majorée par v_0 ; elle converge vers $\ell \in \mathbb{R}$.

De la même v est décroissante et minorée par u_0 donc converge vers $\ell' \in \mathbb{R}$; ainsi $u_n - v_n \to \ell - \ell'$. Par unicité de la limite, $\ell = \ell'$. □

➠ **Exemple IX.14.** Les suites v et w de l'exemple précédent sont donc convergentes vers une même limite et donc, par proposition IX.15, la suite u converge vers cette même limite (nous verrons plus tard qu'il s'agit de $\ln(2)$).

Ce résultat a des conséquences topologiques profondes sur le corps des réels, dont nous présentons une à travers le théorème suivant.

Théorème IX.18 (Segments emboîtés)**.**
Soient $a, b \in \mathbb{R}^{\mathbb{N}}$ deux suites réelles telles que :
- $\forall n \geq 0\ [a_{n+1}, b_{n+1}] \subset [a_n, b_n]$;
- $b_n - a_n \to 0$.

Alors :

$$\exists\, x \in \mathbb{R},\ \bigcap_{n\in\mathbb{N}} [a_n, b_n] = \{x\}\ .$$

Démonstration. Il est clair que les suites a et b sont adjacentes ; elles convergent donc vers une même limite $x \in \mathbb{R}$. Montrons que

$$\exists\, x \in \mathbb{R},\ \bigcap_{n\in\mathbb{N}} [a_n, b_n] = \{x\}\ .$$

($\supset$) Par monotonie on a ,pour tout $n \in \mathbb{N}$ on a $a_n \leq x \leq b_n$ donc $x \in \bigcap_{n\in\mathbb{N}} [a_n, b_n]$.

($\subset$) Soit $y \in \bigcap_{n\in\mathbb{N}} [a_n, b_n]$, *i.e* $\forall n \in \mathbb{N}$, $y \in [a_n, b_n]$. Alors, en passant à la limite dans l'inégalité $a_n \leq y \leq b_n$, on obtient que $y = x$, ce qui conclut cette démonstration.

□

c) Théorème de Bolzano–Weierstrass

Le résultat central de ce paragraphe est nommé d'après Bernard Bolzano (autrichien, 1781—1848) et Karl Weierstass (allemand, 1815—1897).

Théorème IX.19 (Bolzano–Weierstrass)**.**
De toute suite réelle bornée on peut extraire une suite convergente.

La démonstration de ce théorème, bien qu'hors programme, nous est tout à fait accessible ; elle fera l'objet d'un devoir libre et repose sur une dichotomie et le théorème des segments emboîtés (IX.18).

➠ **Exemple IX.15.** Il existe une fonction $\varphi : \mathbb{N} \to \mathbb{N}$ telle que la suite $(\cos(\varphi(n)))_n$ converge. À méditer : les $\varphi(n)$ forment une **suite d'entiers strictement croissante**...

4.⁻ Suites à valeurs complexes

a) Considérations liminaires

Nous avons défini au début de ce chapitre la notion de suite numérique, englobant les suites réelles et complexes. En préambule au présent paragraphe, notons que si $(z_n)_n \in \mathbb{C}^{\mathbb{N}}$ est une suite à valeurs complexes, on peut lui associer les suites **réelles** $(\mathrm{Re}(z_n))_n$, $(\mathrm{Im}(z_n))_n$ et $(|z_n|)_n$.

✖ **Attention :** la suite $(\overline{z_n})_n$ est quant à elle à valeurs complexes.

Définition IX.10. Une suite complexe $z \in \mathbb{C}^{\mathbb{N}}$ est dite **bornée** si la suite $(|z_n|)_n$ est majorée.

✖ **Attention :** on ne peut pas parler de suite complexe majorée ou minorée...

➠ **Exemple IX.16.** Si $\theta \in \mathbb{R}$, la suite $(e^{in\theta})_n$ est bornée.

Proposition IX.20. Soit $z \in \mathbb{C}^{\mathbb{N}}$; alors :

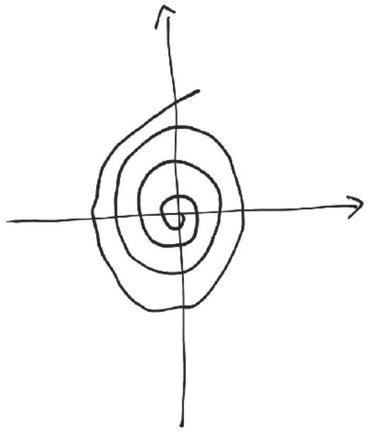
$$(z_n)_n \text{ est bornée} \iff (\mathrm{Re}(z_n))_n \text{ et } (Im(z_n))_n \text{ sont bornées.}$$

Démonstration. Il s'agit d'une conséquence des majorations des parties réelles et imaginaires par le module vues dans le chapitre VIII. □

b) Limites

La définition de limite finie se généralise sans problème aux suites complexes (lire la valeur absolue comme un module) ; la convergence d'une suite complexe est alors équivalente à la convergence des suites partie réelle et partie imaginaire associées.

✖ **Attention :** on ne parle **jamais** de limite infinie pour une suite complexe ; pour comprendre pourquoi, considérez la suite $(2^n e^{in})_n$ qui diverge "en spirale" : son module tend vers l'infini alors que son argument parcoure $[0, 2\pi[$ de façon périodique.



La plupart des résultats énoncés dans les paragraphes précédents se généralisent aux suites réelles sans avoir à en adapter les démonstrations, en particulier :
— les opératins sur les limites ;
— les théorèmes liés aux suites extraites, dont le théorème de Bolzano–Weierstrass (IX.19).
Par contre, en raison de l'absence d'ordre naturel sur $\mathbb{C}$, **aucun théorème en lien avec les inégalités ou la monotonie ne se généralise au cas complexe**.

5.⁻ Zoologie des suites usuelles

a) Suites arithmétiques, suites géométriques

Définition IX.11. Soit $u \in \mathbb{C}^{\mathbb{N}}$ et soit $a \in \mathbb{C}$. On dit que $(u_n)_n$ est une suite...
— ...**arithmétique de raison** a si $\forall n \geq 0,\ u_{n+1} = u_n + a$;
— ...**géometrique de raison** a si $\forall n \geq 0,\ u_{n+1} = au_n$.

Ces suites ont été étudiées en long, en large et en travers en terminale. Rappelons tout de même les résultats suivants :
— si $u \in \mathbb{C}^{\mathbb{N}}$ est une suite arithmétique de raison $a \in \mathbb{C}$, alors son terme général est donné par la formule suivante :

$$\forall n \geq 0,\ u_n = u_0 + an$$

et sa somme par :

$$\forall n \geq 0,\ \sum_{n=0}^{n} u_k = (n+1)u_0 + a\frac{n(n+1)}{2}\ .$$

— si $u \in \mathbb{C}^{\mathbb{N}}$ est une suite géométrique de raison $a \in \mathbb{C}$, alors son terme général est donné par la formule suivante :

$$\forall n \geq 0,\ u_n = u_0 a^n$$

et sa somme par :

$$\forall n \geq 0,\ \sum_{k=0}^{n} u_k = \begin{cases} u_0(n+1) & \text{si } a = 1 \\ u_0 \dfrac{1-a^{n+1}}{1-a} & \text{sinon.} \end{cases}$$

Proposition IX.21 (Nature des suites géométriques)**.** Soit $q \in \mathbb{C}$; alors :
— si $|q| < 1$, la suite $(q^n)_n$ converge vers 0 ;
— si $|q| > 1$, la suite $(q^n)_n$ diverge et $|q^n| \to \infty$;
— si $|q| = 1$ et $q \neq 1$, la suite $(q^n)_n$ diverge.

Démonstration. Les cas $|q| < 1$ et $|q| > 1$ ont déjà été traités dans le cas réel ; la démonstration s'adapte aisément au cas complexe. Supposons à présent que $|q| = 1$ et $q \neq 1$ ($i.e$ que $q \in \mathscr{U} \setminus \{1\}$) ; alors, en passant à la limite dans l'égalité $u_{n+1} = qu_n$, on obtient :

$$\ell = q\ell$$

ce qui est absurde car $q \neq 1$ et $|\ell| = \lim |q^n| = 1$ donc $\ell \neq 0$. □

Définition IX.12. Soient $a, b \in \mathbb{C}$ tels que $a \neq 1$ et $b \neq 0$; on appelle **suite arithmético–géométrique** associée à a et b toute suite $u \in \mathbb{C}^{\mathbb{N}}$ vérifiant :

$$\forall n \geq 0,\ u_{n+1} = au_n + b\,.$$

La méthode d'étude d'une suite arithmético–géométrique est relativement aisée à mettre en œuvre ; on cherche (par analyse–synthèse) à trouver les suites $(v_n)_n = (u_n - \nu)_n$ (avec $\nu \in \mathbb{C}$) géométriques de raison a. Cela signifie que, pour $n \geq 0$ on a :

$$\begin{aligned} v_{n+1} = av_n &\Leftrightarrow u_{n+1} - \nu = a(u_n - \nu) \\ &\Leftrightarrow au_n + b - \nu = au_n - a\nu \\ &\Leftrightarrow \nu = \frac{b}{1-a}\,. \end{aligned}$$

De cette façon, on obtient que, pour tout $n \geq 0$:

$$v_n = v_0 a^n$$

$i.e$

$$u_n = \left(u_0 - \frac{b}{1-a}\right) a^n + \frac{b}{1-a}\,.$$

b) Récurrences linéaires d'ordre 2

On se fixe dans cette partie $a, b \in \mathbb{C}^*$ et on cherche à déterminer les suites complexes $u \in \mathbb{C}^{\mathbb{N}}$ définies par récurrence de la façon suivante :

$$\begin{cases} u_0, u_1 \in \mathbb{C} \\ \forall n \geq 0,\ u_{n+2} = au_{n+1} + bu_n \end{cases}.$$

Commençons par supposer la suite u géométrique de raison $q \in \mathbb{C}^*$; on aurait alors l'équation, pour $n \geq 0$:

$$u_0 q^{n+2} = u_0 q^{n+1} + u_1 q^n$$

i.e

$$u_0 q^n (q^2 - aq - b) = 0 .$$

Il apparait ainsi que la raison q de notre suite doit être racine du polynôme $X^2 - aX - b$. Ceci nous donne l'intuition du résultat suivant.

Proposition IX.22. Soient $a, b \in \mathbb{C}^*$ et soit $u \in \mathbb{C}^{\mathbb{N}}$ une suite définie par récurrence de la façon suivante :

$$\begin{cases} u_0, u_1 \in \mathbb{C} \\ \forall n \geq 0,\ u_{n+2} = au_{n+1} + bu_n \end{cases}.$$

Considérons l'**équation caractéristique** associée à cette relation de récurrence, à savoir :

$$X^2 = aX + b . \qquad (\mathbf{E}\text{:IX.1})$$

Alors :

— si (**E** :IX.1) admet deux racines q et q', le terme général de la suite u est donné par :

$$\forall n \geq 0,\ u_n = \alpha q^n + \beta {q'}^n$$

avec $\alpha, \beta \in \mathbb{C}$ à déterminer à l'aide de u_0 et u_1.

— si (**E** :IX.1) admet un unique racine q, le terme général de la suite u est donné par :

$$\forall n \geq 0,\ u_n = \alpha q^n + \beta n q^n$$

avec $\alpha, \beta \in \mathbb{C}$ à déterminer à l'aide de u_0 et u_1.

Une démonstration de cette proposition à ce stade de l'année nous parait exagérément technique. Nous aborderons un outil permettant de prouver ce résultat efficacement dans le chapitre XX.

✌ **Remarque IX.9.** Les constantes α et β sont à déterminer via la système :

$$\begin{cases} u_0 &= \alpha + \beta \\ u_1 &= \alpha q + \beta q' \end{cases}$$

dans le premier cas et

$$\begin{cases} u_0 &= \alpha \\ u_1 &= (\alpha + \beta) q \end{cases}$$

dans le second.

Exemple IX.17.

— L'expression trouvée pour la suite de Fibonacci dans le chapitre IV peut être retrouvée par cette méthode. Rappelons qu'il s'agit de la suite définie par

$$F_0 = 0, \quad F_1 = 1 \quad \text{et} \quad \forall n \geq 0,\ F_{n+2} = F_n + F_{n+1}\ .$$

— Étudions la suite $(u_n)_n$ définie par :

$$u_0 = 1, \quad u_1 = 1 \quad \text{et} \quad \forall n \geq 0,\ u_{n+2} = -u_n\ .$$

Son équation caractéristique étant $X^2 + 1 = 0$, on sait que son terme général sera de la forme, pour $n \geq 0$

$$u_n = \alpha i^n + \beta(-i)^n\ .$$

À l'aide des valeurs de u_0 et u_1, on trouve que $\alpha = \dfrac{1-i}{2}$ et $\beta = \dfrac{1+i}{2}$.

6.– Retour sur la topologie du corps des réels

Dans ce paragraphe, nous revenons sur certains résultats évoqués dans le chapitre VI et les démontrons/complétons.

a) Caractérisation séquentielle des bornes

Proposition IX.23. Soit X une partie non vide de $\mathbb{R}$. Alors :
- si X est majorée, il existe une suite $u \in X^{\mathbb{N}}$ convergeant vers $\sup(X)$;
- sinon, il existe une suite $u \in X^{\mathbb{N}}$ divergeant vers $+\infty$.

Remarque IX.10. Une fois n'est pas coutume, il existe un résultat symétrique relatif aux bornes inférieures et aux parties minorées.

Démonstration.

Cas 1 : X est majorée. Alors $M = \sup(X)$ existe dans $\mathbb{R}$ et, pour tout $n \geq 1$, la définition de supremum entraîne que :

$$\exists,\ x_n \in X,\ M - x_n \leq \frac{1}{n}\ .$$

Comme M majore X, alors :

$$\forall n \geq 1,\ 0 \leq |x_n - M| \leq \frac{1}{n}$$

ce qui nous permet de déduire, par théorème d'encadrement des limites (IX.12), que $x_n \to M$.

Cas : X n'est pas majorée. Alors dans ce cas, pour tout $n \in \mathbb{N}$, il existe $x_n \in X$ tel que $n \leq x_n$; la suite $(x_n)_n$ diverge alors vers l'infini.

□

On déduit de ce résultat le corollaire suivant, déjà rencontré sous le nom de proposition VI.8.

Proposition IX.24. [Caractérisation séquentielle de la borne supérieure] Soit $A \subset \mathbb{R}$ et soit $M \in \mathbb{R}$; alors :

$$M = \sup(A) \iff \begin{cases} \forall a \in A,\ a \leq M \\ \exists (a_n)_n \in A^{\mathbb{N}},\ a_n \xrightarrow[n\to\infty]{} M \end{cases} .$$

b) Densité

Définition IX.13. Une partie X de $\mathbb{R}$ est dite **dense** si pour tout intervalle I non vide ni réduit à un point de $\mathbb{R}$, l'intersection $I \cap X$ est non vide.

➠ **Exemple IX.18.** Nous avons vu dans le chapitre VI que $\mathbb{Q}$ et $\mathbb{R}\setminus\mathbb{Q}$ étaient des parties denses de $\mathbb{R}$. Notons que c'est également le cas de $\mathbb{R}$...

Proposition IX.25. [Caractérisation séquentielle de la densité] Soit $X \subset \mathbb{R}$; alors :

$$X \text{ est une partie dense de } \mathbb{R}$$
$$\Leftrightarrow$$
$$\forall x \in \mathbb{R},\ \exists\, (x_n)_n \in X^{\mathbb{N}},\ x_n \to x.$$

Démonstration.

($\Downarrow$) Soit $x \in \mathbb{R}$ et soit $n \geq 1$; alors par densité

$$X \cap \left[x - \frac{1}{n}, x + \frac{1}{n}\right] \neq \emptyset$$

et donc il existe x_n dans cette intersection. Il est clair par encadrement que la suite $(x_n)_n$ converge vers x.

($\Uparrow$) Soit I un intervalle non vide de $\mathbb{R}$ et soit $x_0 \in I$ différent de ses bornes. Alors, par convexité de I, il existe $\varepsilon > 0$ tel

$$[x_0 - \varepsilon, x_0 + \varepsilon] \subset I \ :$$

Or, par hypothèse, il existe une suite $(x_n)_n$ d'éléments de X convergeant vers x_0, ce qui entraîne que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ x_n \in [x_0 - \varepsilon, x_0 + \varepsilon]$$

et donc, en particulier, $x_N \in I \cap X$.

□

Chapitre X

Groupes, anneaux et corps

0.⁻ Lois de composition interne

a) C'est quoi ?

Définition X.1. Soit E un ensemble non vide. On appelle **loi de composition interne** (LCI) sur E toute application $\star : E \times E \to E$.

➠ **Exemple X.1.** Nous en avons déjà rencontré un certain nombre ; par exemple

$$\begin{aligned} + : \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (m, n) &\mapsto m + n\ , \end{aligned}$$

$$\begin{aligned} - : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z} \\ (m, n) &\mapsto m - n \end{aligned}$$

ou

$$\begin{aligned} \circ : \mathbb{R}^{\mathbb{R}} \times \mathbb{R}^{\mathbb{R}} &\to \mathbb{R}^{\mathbb{R}} \\ (f, g) &\mapsto f \circ g\ . \end{aligned}$$

Il en existe d'autres, moins évidentes : l'intersection et la réunion d'ensembles, ainsi que la différence ensembliste, sont des LCI sur $\mathscr{P}(E)$.

Notation. Si $\star$ est une LCI sur E et $x, y \in E$, nous noterons $\star(x, y)$ "$x \star y$" ; cette convention s'appelle la **notation infixée** (par opposition à la notation préfixée $\star(x, y)$ et à la notation postfixée $(x, y)\star$).

✌ **Remarque X.1.** Dans le cas de certaines LCI (la multiplication vient à l'esprit), nous n'écrirons parfois même pas l'opérateur : $x \times y$ deviendra xy.

Définition X.2. Une partie A d'un ensemble E muni d'une LCI $\star$ est dite **stable par** $\star$ si :

$$\forall x, y \in A,\ x \star y \in A\ .$$

➠ **Exemple X.2.** $\mathbb{N}$ est une partie de $\mathbb{Z}$ stable par addition et multiplication.

b) Propriétés remarquables

Dans ce paragraphe, nous listons certaines propriétés intéressantes pouvant être possédées par une loi de composition interne $\star$ que nous supposerons fixée sur un ensemble arbitraire E.

◇ Associativité

Définition X.3. La LCI $\star$ est dite **associative** si :

$$\forall x, y, z \in E,\ (x \star y) \star z = x \star (y \star z)\ .$$

✌ **Remarque X.2.** Lorsque $\star$ est associative, on peut écrire des choses du style "$x \star y \star z$" sans ambiguïté.

➠ **Exemple X.3.**

- Nous avons vu que les opérations $+, \times, \circ$ étaient systématiquement associatives sur les ensembles *ad–hoc*.
- La soustraction n'est pas associative : $(2-3)-1 = -2 \neq 2-(3-1) = 0$. Cela signifie qu'il est logiquement insensé d'écrire des choses comme "$-2-3-1$" sans avoir préalablement décidé d'un sens prioritaire dans les opérations.
- De même, la division n'est pas associative.
- La réunion et l'intersection ensemblistes sont associatives, mais pas la différence.

◇ Élément neutre

Définition X.4. On dit que $e \in E$ est un **élément neutre** pour $\star$ si :

$$\forall x \in E,\ x \star e = e \star x = x\ .$$

➠ **Exemple X.4.**

- Dans les sous–ensembles de $\mathbb{C}$ appropriés, 1 est le neutre pour la multiplication et 0 le neutre pour l'addition.
- Dans E^E, l'application id_E est un élément neutre pour la composition.
- Dans $\mathscr{P}(E)$, $\emptyset$ est neutre pour la réunion et E est neutre pour l'intersection.

Proposition X.1. Un élément neutre, si il existe, est unique.

Démonstration. Soient $e, e' \in E$ deux éléments neutres pour $\star$. Alors, comme e est neutre, $e \star e' = e'$, et comme e' est neutre, $e \star e' = e$, donc $e = e'$. □

✌ **Remarque X.3.** Un élément $\tau \in E$ tel que $\forall x \in E$, $\tau \star x = x \star \tau = \tau$ est appelé **élément absorbant**. Penser à 0 et à la multiplication.

◇ Inversibles

Supposons dans ce paragraphe que la loi de composition interne $\star$ admette un élément neutre, que nous noterons e.

Définition X.5. Un élément $x \in E$ est dit **inversible** si :

$$\exists\, y \in E,\ x \star y = y \star x = e\ .$$

Notation. y est appelé **inverse de** x et noté x^{-1}. De plus :
— si $\star$ est l'addition, y sera noté $-x$;
— si $\star$ est la multiplication sur un sous–ensemble de $\mathbb{C}$, y sera noté $\dfrac{1}{x}$.

➠ **Exemple X.5.**
— Pour l'addition sur $\mathbb{R}$, $2^{-1} = -2$.
— Pour la multiplication sur ce même ensemble, $2^{-1} = \dfrac{1}{2}$

✖ **Attention :** il convient d'être extrêmement vigilant quant à la LCI considérée, et à rendre cette dernière limpide pour le correcteur au concours ou en exercices.

Qui sont les inversibles des LCI classiques ?
— pour l'addition sur $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$: tout le monde ;
— pour la multiplication sur $\mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$: tout le monde sauf 0 ;
— pour la composition sur E^E : les bijections ;
— pour la réunion sur $\mathscr{P}(E)$: uniquement $\emptyset$;
— pour l'intersection sur $\mathscr{P}(E)$: uniquement E.

Proposition X.2. Si la loi $\star$ est associative, il y a unicité de l'inverse lorsqu'il existe.

Démonstration. Soit $x \in E$ d'inverse(s) y et z. Alors, comme $x \star y = e$, $z \star (x \star y) = z$ *i.e* $(z \star x) \star y = z$. Or $z \star x = e$ donc *in fine* $y = z$. □

Corollaire X.2.a. Soit $x \in E$ un élément inversible. Alors $(x^{-1})^{-1} = x$.

Proposition X.3 (Simplification)**.** Supposons $\star$ associative et fixons $x, y, z \in E$ tels que x soit inversible. Alors :
(i) $(x \star y = x \star z) \Rightarrow (y = z)$;
(ii) $(y \star x = z \star x) \Rightarrow (y = z)$.

Démonstration. Il suffit de faire le produit (du bon côté et au sens de $\star$) par x^{-1} et d'utiliser l'associativité pour simplifier les égalités. □

✖ **Attention :** on ne peut simplifier **que** par un élément inversible. En effet, $0 \times 2 = 0 \times 1$ et pourtant $1 \neq 2$. Plus subtil : si on pose $f : x \mapsto x^2$, $g : x \mapsto \sqrt{|x|}$ et $h : x \mapsto -\sqrt{|x|}$ sur $\mathbb{R}$, on a $f \circ g = f \circ h$ et $g \neq h$.

Proposition X.4. On suppose $\star$ associative ; soient $x, y \in E$ deux éléments inversibles. Alors $x \star y$ est inversible et :

$$(x \star y)^{-1} = y^{-1} \star x^{-1} .$$

Démonstration. Vérifier que cet inverse convient et conclure par unicité. □

◇ Distributivité

Définition X.6. Soit Δ une LCE sur E ; on dit que $\star$ **est distributive par rapport à** Δ si on a, pour tous $x, y, z \in E$:

$$x \star (y \Delta Z) = (x \star y) \Delta (x \star z)$$

et

$$(y \Delta z) \star x = (y \star x) \Delta (z \star x) .$$

Exemple X.6.
- La multiplication est distributive par rapport à l'addition sur les sous–ensembles de $\mathbb{C}$.
- L'intersection et la réunion sont mutuellement distributives sur $\mathscr{P}(E)$.

◇ Commutativité

Définition X.7. La LCE $\star$ est dite **commutative** si :

$$\forall x, y \in E,\ x \star y = y \star x .$$

✌ **Remarque X.4.** Certaines des propriétés précédentes sont plus simples à énoncer dans le cas commutatif.

Exemple X.7.
- La multiplication et l'addition le sont sur les sous–ensembles de $\mathbb{C}$, mais pas la soustraction ou la division.
- L'intersection et la réunion le sont sur $\mathscr{P}(E)$, mais pas la différence ensembliste.

1.⁻ Groupes

a) C'est quoi ?

Définition X.8. Soit G un ensemble non vide muni d'une LCI $\star$. On dit que le couple $(G, \star)$ est un **groupe** si :

(A) la loi $\star$ est associative ;

(N) la loi $\star$ admet un élément neutre, noté e_G ;

(I) tous les éléments de G sont inversibles pour $\star$.

Si de plus la loi $\star$ est commutative, on parle de **groupe abélien** (du mathématicien norvégien Niels Henrik Abel, 1802—1829) ou commutatif.

Notation. Dans le cas général, nous utiliserons la notation multiplicative xy pour $x \star y$ afin d'alléger les énoncés des propositions. Nous réserverons la notation additive $x + y$ au cas des groupes abéliens.

Exemple X.8.
- si e est neutre pour $\star$, $(\{e\}, \star)$ est un groupe, appelé **groupe trivial** ;
- $(\mathbb{Z}, +)$ est un groupe abélien, mais pas $(\mathbb{N}, +)$ (mise en défaut de la propriété (I)) ;
- $(\mathbb{R}, +)$ et $(\mathbb{C}, +)$ sont des groupes, abélien mais pas $(\mathbb{R}, \times)$ (même problème) ;
- $(\mathbb{R}^*, \times)$ et $(\mathbb{C}^*, \times)$ sont des groupes abéliens ;
- si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $(\mathbb{K}^{\mathbb{N}}, +)$ est un groupe abélien ;
- $(\mathfrak{S}(\mathbb{R}, \mathbb{R}), \circ)$ est un groupe non abélien.

◇ **Un exemple fondamental : le groupe des permutations.**

Soit $n \geq 1$; on appelle **groupe des permutations de** $[\![1, n]\!]$ l'ensemble

$$\mathfrak{S}_n = \mathfrak{S}\left([\![1, n]\!]\right)$$

muni de la composition. Il s'agit d'un groupe non abélien dont les éléments sont les applications associant à chaque entier de $[\![1, n]\!]$ un autre tel nombre de façon bijective.

Exemple X.9.
- $\mathfrak{S}_1 = \{\mathrm{id}_{[\![1,n]\!]}\}$ est le groupe trivial ;
- $\mathfrak{S}_2 = \{\mathrm{id}_{[\![1,n]\!],\sigma}\}$, avec σ l'application permutant 1 et 2 ;
- $\mathfrak{S}_3$ possède 6 éléments : lesquels ?

Un élément $\sigma \in \mathfrak{S}_n$ peut être construit de la façon suivante : on dispose de n choix pour l'image de 1, puis (une fois celui–ci fixé), de $n - 1$ choix pour l'image de 2, $n - 2$ choix pour l'image de 3, etc. On en déduit que $\mathfrak{S}_n$ possède exactement $n!$ éléments.

✍ **Exercice X.1.** Déterminer $\mathfrak{S}_4$.

b) Puissances

Définition X.9. Soit G un groupe de neutre e. On définit, pour $x \in G$ et $n \in \mathbb{Z}$ l'élément **puissance n–ième de** x comme suit :
— $x^0 = e$;
— si $n \in \mathbb{N}$, $x^{n+1} = x^n x$;
— si $n < 0$, $x^n = (x^{-1})^{-n}$.

Exemple X.10.
— sur $(\mathbb{C}^*, \times)$, cela correspond aux puissances usuelles ;
— sur $(\mathbb{C}, +)$, pour $x \in \mathbb{C}$ et $n \geq 1$, $x^n = \underbrace{x + \ldots + x}_{n \text{ fois}} = nx$. On en déduit aisément que **les puissances additives sont les multiples**. Attention ici aux notations et à leur potentiel de confusion.
— Sur $\mathfrak{S}_n$, passer une permutation à la puissance $n \in \mathbb{N}^*$ revient à la composer n fois avec elle–même.

Proposition X.5. Soit G un groupe et soient $x \in G$ et $n, m \in \mathbb{Z}$. Alors :
(i) $x^{n+m} = x^n x^m = x^m x^n$;
(ii) $(x^n)^m) = x^{nm}$.

Remarque X.5. Ce résultat, comme tous ceux de ce paragraphe, peut se reformuler en notation additive (traditionnellement utilisée dans le cas abélien uniquement) :
(i) $(n + m)x = nx + mx$;
(ii) $m(nx) = (mn)x$.

Proposition X.6. Soit G un groupe et soient $x, y \in G$ tels que $xy = yx$. Alors, pour tous $m, n \in \mathbb{Z}$:
$$x^n y^m = y^m x^n .$$
De plus :
$$(xy)^n = x^n y^n = y^n x^n .$$

Démonstration. Ce résultat et le précédent se démontrent par récurrence en distinguant les cas d'exposants positifs et négatifs. □

c) Sous–groupes

Définition X.10. Soit $(G, \star)$ un groupe. On appelle **sous–groupe** de G tout ensemble H tel que :
— $H \subset G$;
— $(H, \star)$ est un groupe.

Exemple X.11.
- $(\mathbb{Q}, +)$ est un sous–groupe de $(\mathbb{R}, +)$;
- $\mathfrak{S}_3$ est un sous–groupe de $\mathfrak{S}_3$.

Proposition X.7. Soit G un groupe et H un sous–groupe de G. Alors :

(i) $e_H = e_G$;

(ii) si $x \in H$ alors son inverse est le même dans H et dans G.

Démonstration.

(i) Remarquons que $e_H \star e_H = e_H$ dans H donc dans G. Ainsi, toujours dans G, $e_H = e_H \star e_H^{-1}$ et donc $e_H = e_G$.

(ii) Notons y l'inverse de x dans G et z son inverse dans H. Alors, en posant $e = e_H = e_G$ (*cf.* (i)) on a :

$$y \star x = e = z \star x$$

et donc

$$y \star x \star z = z \star x \star z$$

ce qui entraîne que $y = z$ car $x \star z = e$.

□

Proposition X.8 (Caractérisation des sous–groupes)**.** Soit G un groupe de neutre e et soit $H \subset G$. Alors :

$$H \text{ est un sous–groupe de } G \iff \begin{cases} H \neq \emptyset \\ \forall x, y \in H,\ xy \in H \\ \forall x \in H,\ x^{-1} \in H \end{cases} .$$

Remarque X.6. Les deux derniers points de la caractérisation peuvent être remplacés par :

$$\forall x, y \in H,\ xy^{-1} \in H .$$

Exemple X.12. Cette caractérisation permet de démontrer, en utilisant des résultats vus dans le chapitre précédents, que :
- $\mathscr{U}$ est un sous–groupe de $(\mathbb{C}^*, \times)$;
- pour $n \in \mathbb{N}^*$, $\mathscr{U}_n$ est un sous–groupe de $\mathscr{U}$;
- pour $n \in \mathbb{Z}$, $n\mathbb{Z} = \{nk \,|\, k \in \mathbb{Z}\}$ est un sous–groupe de $(\mathbb{Z}, +)$.
- $\mathbb{R}_+^*$ est un sous–groupe de $(\mathbb{R}^*, \times)$, qui est lui–même un sous–groupe de $\mathbb{C}^*$.

Exercice X.2. Soit $n \geq 1$; on pose

$$H_n = \{\sigma \in \mathfrak{S}_n \,|\, \sigma(n) = n\} .$$

Démontrer que H_n est un sous–groupe de $(\mathfrak{S}_n, \circ)$.

➥ **Correction :** *H_n est non vide car il contient $\mathrm{id}_{[\![1,n]\!]}$. De même, il est clair que la composée et l'inverse de deux éléments de H_n sont dans H_n, d'où le résultat (composer par σ^{-1} dans $\sigma(n) = n$).*

Corollaire X.8.a. L'intersection de deux sous–groupes d'un même groupe en est un sous–groupe.

Démonstration. Soient H et K deux sous–groupes d'un groupe G de neutre e. Alors :
- $e \in H \cap K$ donc cet ensemble est non vide ;
- si $x, y \in H \cap K$, alors $xy^{-1} \in H$ et $xy^{-1} \in K$ car $x, y \in H$ et $x, y \in K$ respectivement.

D'où le résultat. □

✖ **Attention :** la réunion de deux sous–groupes n'est pas un sous groupe : en effet, $2, 3 \in 2\mathbb{Z} \cup 3\mathbb{Z}$ et pourtant $2 + 3 = 5 \notin 2\mathbb{Z} \cup 3\mathbb{Z}$.

d) Morphismes

Définition X.11. Soit $(G, \star)$ et (H, Δ) deux groupes. On appelle **morphisme de groupes** de G vers H toute application $f : G \to H$ telle que :

$$\forall x, y \in G,\ f(x \star y) = f(x)\Delta f(y)\ .$$

➠ **Exemple X.13.** La encore nous en avons déjà croisés quelques–un :
- $\exp : (\mathbb{R}, +) \to (\mathbb{R}_+^*, \times)$;
- $\ln : (\mathbb{R}_+^*, \times) \to (\mathbb{R}, +)$;
- $$\begin{array}{rl} \varphi : (\mathbb{R}, +) \to (\mathscr{U}, \times) \\ \theta \mapsto e^{i\theta} \end{array} .$$

Proposition X.9. Soit G, H deux groupes et $f : G \to H$ un morphisme de groupes. Alors :

(i) $f(e_G) = e_H$;

(ii) si $x \in G$, $f(x^{-1}) = f(x)^{-1}$;

(iii) si $x \in G$ et $n \in \mathbb{Z}$, $f(x^n) = f(x)^n$.

Démonstration.

(i) Il suffit de remarquer que $f(e_G) = f(e_G \cdot e_G) = f(e_G)f(e_G)$ et de simplifier.

(ii) Soit $x \in G$; alors :

$$f(x)f(x^{-1}) = f(x \cdot x^{-1}) = f(e_G) = e_H$$

d'où le résultat.

(iii) À démontrer par récurrence en distinguant les cas d'exposants positifs et négatifs.

□

Définition X.12. Soit G, H deux groupes. Un morphisme de groupes $f : G \to H$ est appelé…
- …**isomorphisme** s'il est bijectif ;
- …**endomorphisme** si $G = H$;
- …**automorphisme** si il est bijectif et que $G = H$.

2.⁻ Anneaux, corps

a) On prend les mêmes…

Définition X.13. Un ensemble $\mathbb{A}$ muni de **deux** LCI distinctes $+$ et $\times$ est appelé **anneau** (unitaire) si :

(G) $(\mathbb{A}, +)$ est un groupe **abélien** ;

(A) $\times$ est associative ;

(D) $\times$ est distributive par rapport à $+$;

(N) la loi $\times$ admet un élément neutre.

Si de plus la loi $\times$ est commutative, on parle d'**anneau commutatif**.

Notation. L'élément neutre pour $+$ est noté $0_{\mathbb{A}}$ (ou simplement 0), le neutre pour $\times$ est noté $1_{\mathbb{A}}$ (ou 1).

✌ **Remarque X.7.** Si $(\mathbb{A}, +, \times)$ est un anneau, alors pour tout $x \in \mathbb{A}$, $0 \cdot x = (1-1)x = x - x = 0$; 0 est donc absorbant pour la loi $\times$.

➠ **Exemple X.14.**
- les ensembles $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont des anneaux pour l'addition et la multiplication standards ;
- si E est un ensemble et $\mathbb{A}$ un anneau, $\mathbb{A}^E$ est un anneau pour les opérations terme à terme (*cf.* chapitre I) ;
- le seul anneau dans lequel $0 = 1$ est l'**anneau nul** $\{0\}$. En effet, si x est un élément d'un tel anneau alors $0 \cdot x = 0 = 1 \cdot x = x$.

Définition X.14. Un anneau est appelé **corps** si :
- il est commutatif ;
- il est différent de l'anneau nul ;
- tous ses éléments non nuls sont inversibles pour la multiplication.

➠ **Exemple X.15.** $\mathbb{R}$ et $\mathbb{C}$ sont des corps, mais pas $\mathbb{Z}$ ou $\mathbb{R}^{\mathbb{R}}$.

b) ... et on recommence !

On fixe dans ce paragraphe un anneau $(\mathbb{A}, +, \times)$. Alors la notation suivante a un sens, pour $a_1, \dots, a_h \in \mathbb{A}$:

$$\sum_{k=1}^{n} a_k = a_1 + \ldots + a_n \; .$$

De plus, si $\mathbb{A}$ est commutatif, on pose :

$$\prod_{k=1}^{n} a_k = a_1 \times \ldots \times a_n \; .$$

On généralise ces notations au cas de familles **finies** indexées par un ensemble quelconque.

Si $a \in \mathbb{A}$ et $n \in \mathbb{A}$, on peut définir le **multiple** na comme sa puissance additive. On ne peut toute fois pas définir a^n pour $n < 0$ lorsque a n'est pas inversible.

c) Sous–machins, inversibles

Définition X.15. Soit $\mathbb{A}$ un anneau et soit $\mathbb{B} \subset \mathbb{A}$. On dit que $\mathbb{B}$ est un **sous–anneau** de $\mathbb{A}$ si :
- $\mathbb{B}$ est un anneau ;
- $1_{\mathbb{B}} = 1_{\mathbb{A}}$.

Si de plus $\mathbb{A}$ et $\mathbb{B}$ sont des corps, on parle de sous–corps.

Exemple X.16. $\mathbb{Z}$ est un sous–anneau de $\mathbb{R}$, qui est un sous–corps de $\mathbb{C}$.

Comme dans le cas des sous–groupes, on dispose d'une caractérisation (analogue à a la proposition X.8) des sous–anneaux : si $\mathbb{A}$ est un anneau et $\mathbb{B} \subset \mathbb{A}$, alors

$$\mathbb{B} \text{ est un sous–anneau de } \mathbb{A} \Longleftrightarrow \begin{cases} 1 \in \mathbb{B} \\ \forall x, y \in \mathbb{B},\ x - y \in \mathbb{B} \\ \forall x, y \in \mathbb{B},\ xy \in \mathbb{B} \end{cases} .$$

Exemple X.17. À l'aide de cette caractérisation, il est aisé de montrer que l'ensemble des **entiers de Gauss**

$$\mathbb{Z}[i] = \{a + ib \,|\, a, b \in \mathbb{Z}\}$$

est un sous–anneau de $\mathbb{C}$.

Définition X.16. Soit $\mathbb{A}$ un anneau ; on dit que $x \in \mathbb{A}$ est **inversible** si il l'est pour la loi $\times$.

Notation. On note $\mathbb{A}^\times$ l'ensemble des inversibles de l'anneau $\mathbb{A}$.

Remarque X.8. Remarquons que $(\mathbb{A}^\times, \times)$ est naturellement un groupe.

Exemple X.18.
- $\mathbb{Z}^\times = \{-1, 1\}$;
- $\mathbb{C}^\times = \mathbb{C}^*$;
- $\{0\}^\times = \{0\}$ (et oui...).

d) Identités remarquables

Proposition X.10. Soit $\mathbb{A}$ un anneau et $a, b \in \mathbb{A}$ tels que $ab = ba$. Alors, pour tout $n \in \mathbb{N}$:

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

(ii)
$$a^{n+1} - b^{n+1} = (a-b) \sum_{k=0}^{n} a^k b^{n-k} .$$

Démonstration.

(i) Comme a et b commutent, on peut recopier la démonstration de la proposition IV.6.

(ii)
$$\begin{aligned}(a-b)\sum_{k=0}^{n} a^k b^{n-k} &= a\sum_{k=0}^{n} a^k b^{n-k} - b\sum_{k=0}^{n} a^k b^{n-k} \\ &= \sum_{k=0}^{n} a^{k+1} b^{n-k} - \sum_{k=0}^{n} a^k b^{n+1-k} \\ &= \sum_{k=1}^{n+1} a^k b^{n+1-k} - \sum_{k=0}^{n} a^k b^{n+1-k} \text{ par changement d'indice} \\ &= b^{n+1} - a^{n+1} .\end{aligned}$$

□

Corollaire X.10.a. Soit $\mathbb{A}$ un anneau et $a \in \mathbb{A}$. Alors, pour tout $n \in \mathbb{N}^*$:

$$1 - a^n = (1-a) \sum_{k=0}^{n-1} a^k .$$

Démonstration. Appliquer le (ii) de la proposition X.10 à $b = 1$. □

✌ **Remarque X.9.** On peut voir dans cette formule un analogue de la somme des termes d'une suite géométrique vue dans le chapitre IV.

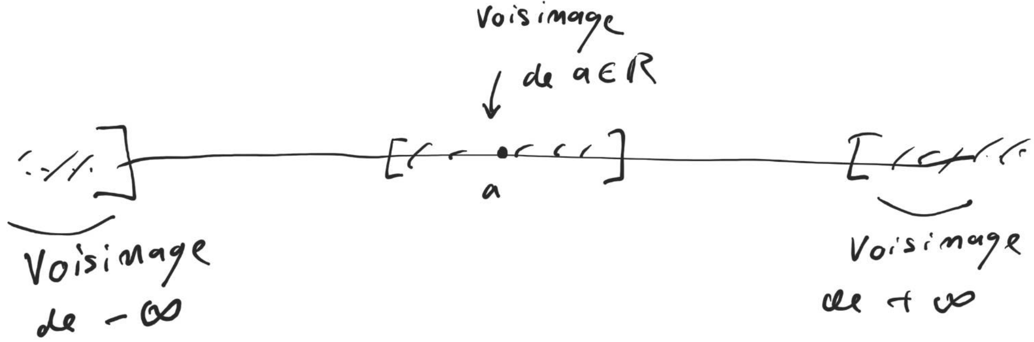
Chapitre XI

Limites, continuité

1.⁻ Étude locale d'une fonction

a) Voisinages

Définition XI.1. Soit $a \in \overline{\mathbb{R}}$. On appelle **voisinage de** a tout ensemble de la forme :
- $[a-\delta, a+\delta]$ avec $\delta > 0$ si $a \in \mathbb{R}$;
- $[M, +\infty[$ avec $M \in \mathbb{R}$ si $a = +\infty$;
- $]-\infty, m]$ avec $m \in \mathbb{R}$ si $a = -\infty$.



Notation. On notera $\mathscr{V}(a)$ l'ensemble des voisinages de $a \in \overline{\mathbb{R}}$.

➠ **Exemple XI.1.** $[-1,1] \in \mathscr{V}(0)$, $\mathbb{R}_+ \in \mathscr{V}(\infty)$.

Proposition XI.1. Soit $a \in \overline{\mathbb{R}}$. Alors :

(i) l'intersection de deux voisinages de a est un voisinage de a ;

(ii)

$$\bigcap_{V \in \mathscr{V}(a)} V = \begin{cases} \{a\} & \text{si } a \in \mathbb{R} \\ \emptyset & \text{sinon} \end{cases} .$$

Démonstration.

(i) Immédiat si l'on prend le temps de faire une disjonction de cas selon les valeurs de a.

(ii) ($\supset$) Immédiat.

($\subset$) Si $a \in \mathbb{R}$, alors si $x \in \bigcap_{V \in \mathscr{V}(a)} V$, pour tout $\varepsilon > 0$, $x \in [a-\varepsilon, a+\varepsilon]$, *i.e* $|x-a| \leq \varepsilon$. On en déduit que $|x-a| \leq \inf \mathbb{R}_+^* = 0$ donc $x = a$.
Si $a = +\infty$, alors si on suppose qu'il existe $x \in \bigcap_{V \in \mathscr{V}(\infty)}$, il existe $M > x$ et donc $x \notin [M, +\infty[\in \mathscr{V}(\infty)$, ce qui est absurde. On règle de la même façon le cas où $a = -\infty$.

□

Proposition XI.2. [Séparation de $\overline{\mathbb{R}}$] Soient $a, b \in \overline{\mathbb{R}}$ **distincts**. Alors il existe $V \in \mathscr{V}(a)$ et $W \in \mathscr{V}(b)$ tels que :

$$V \cap W = \emptyset .$$

✌ **Remarque XI.1.** Cela signifie qu'étant donné deux points distincts de la droite réelle achevée, on peut leur trouver deux voisinages ne se rencontrant pas. Topologiquement, on dit que $\overline{\mathbb{R}}$ est **séparé**.

Démonstration. Si a et b sont réels avec $a < b$, en posant $\delta = \dfrac{b-a}{3}$ on vérifie aisément que les voisinages $[a-\delta, a+\delta] \in \mathscr{V}(a)$ et $[b-\delta, b+\delta] \in \mathscr{V}(b)$ sont disjoints. Si a ou b est égal à $\pm\infty$ se traite de façon similaire. □

Définition XI.2. Soit $\mathscr{P}$ un prédicat dépendant d'une variable réelle x et soit $a \in \overline{\mathbb{R}}$. On dira que $\mathscr{P}(x)$ **est vraie au voisinage de** a si :

$$\exists V \in \mathscr{V}(a),\ \forall x \in V,\ \mathscr{P}(x) .$$

➠ **Exemple XI.2.** $x^2 \leq 1$ est vraie au voisinage de 0 ; $x \mapsto x^2$ est décroissante au voisinage de -42.

b) Limite et continuité en un point

Définition XI.3. Soit I un intervalle de $\mathbb{R}$ d'extrémités $a, b \in \overline{\mathbb{R}}$. On appelle :
- **intérieur de** I l'ensemble
$$\mathring{I} = I \setminus \{a, b\} \subset \mathbb{R}\,;$$
- **adhérence de** I l'ensemble
$$\bar{I} = I \cup \{a, b\} \subset \overline{\mathbb{R}}\,.$$

➠ **Exemple XI.3.** $\overset{\circ}{\overbrace{[0,2[}} =]0,2[$ et $\overline{[0,\infty[} = \overline{\mathbb{R}}_+$.

On fixe dans ce paragraphe un intervalle I de $\mathbb{R}$ d'intérieur non vide.

Définition XI.4. Soit $a \in \bar{I}$ et soit $\ell \in \overline{\mathbb{R}}$. On dit qu'une fonction $f : I \to \mathbb{R}$ **tend vers** ℓ **en** a lorsque :
$$\forall V \in \mathscr{V}(\ell),\ \exists W \in \mathscr{V}(a),\ f(W \cap I) \subset V\,.$$
La quantité ℓ est alors appelée **limite** de f en a.

Notation. Nous utiliserons indifféremment $f(x) \xrightarrow[x \to a]{} \ell$ et $\lim\limits_{x \to a} f(x) = \ell$.

✌ **Remarque XI.2.**
- Rappelons que l'assertion $f(W \cap I) \subset V$ signifie
$$\forall x \in W \cap I,\ f(x) \in V\,.$$
- Cette définition peut être traduite "en français" de la même façon que son analogue séquentiel vu dans le chapitre IX : pour tout voisinage de ℓ, si x est suffisamment proche de a, $f(x)$ appartiendra à ce voisinage.

Cette définition de limite possède l'avantage d'être universelle : peut importe que a ou ℓ soient réels, infinis ou un mélange des deux, elle restera valable. Pour aider dans nos considérations pratiques, nous listons malgré des considérations de type "dans le cambouis jusqu'au torse" ci–ensuite. Une fois de plus, remercier l'auteur est inutile.

$\diamond$ **$a, \ell \in \mathbb{R}$:**
$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in I,\ (|x - a| \leq \delta) \Rightarrow (|f(x) - \ell| \leq \varepsilon)\,.$$

$\diamond$ **$a \in \mathbb{R}$, $\ell = +\infty$:**
$$\forall M \in \mathbb{R},\ \exists \delta > 0,\ \forall x \in I,\ (|x - a| \leq \delta) \Rightarrow (f(x) \geq M)\,.$$

$\diamond$ **$a \in \mathbb{R}$, $\ell = -\infty$:**
$$\forall m \in \mathbb{R},\ \exists \delta > 0,\ \forall x \in I,\ (|x - a| \leq \delta) \Rightarrow (f(x) \leq m)\,.$$

$\diamond$ **$a = +\infty$, $\ell \in \mathbb{R}$:**

$$\forall \varepsilon > 0,\ \exists M \in \mathbb{R},\ \forall x \in I,\ (x \geq M) \Rightarrow (|f(x) - \ell| \leq \varepsilon)\ .$$

$\diamond$ **$a = -\infty$, $\ell \in \mathbb{R}$:**

$$\forall \varepsilon > 0,\ \exists m \in \mathbb{R},\ \forall x \in I,\ (x \leq m) \Rightarrow (|f(x) - \ell| \leq \varepsilon)\ .$$

$\diamond$ **$a = +\infty, \ell = -\infty$:**

$$\forall m \in \mathbb{R},\ \exists M \in \mathbb{R},\ \forall x \in I,\ (x \geq M) \Rightarrow (f(x) \leq m)\ .$$

$\diamond$ **$a = -\infty, \ell = +\infty$:**

$$\forall M \in \mathbb{R},\ \exists m \in \mathbb{R},\ \forall x \in I,\ (x \leq m) \Rightarrow (f(x) \geq M)\ .$$

$\diamond$ **$a = \ell = +\infty$:**

$$\forall M \in \mathbb{R},\ \exists R \in \mathbb{R},\ \forall x \in I,\ (x \geq R) \Rightarrow (f(x) \geq M)\ .$$

$\diamond$ **$a = \ell = -\infty$:**

$$\forall m \in \mathbb{R},\ \exists r \in \mathbb{R},\ \forall x \in I,\ (x \leq r) \Rightarrow (f(x) \leq m)\ .$$

Youpi tralala, comme dirait l'autre.

➠ **Exemple XI.4.** La fonction $x \mapsto e^x$ tend vers $+\infty$ quand $x \to +\infty$ car, pour tout $M > 0$, $(x \geq \ln(M)) \Rightarrow (e^x \geq M)$. De même, on peut montrer que $x^3 \xrightarrow[x \to -1]{} -1$.

Proposition XI.3. La limite d'une fonction en un point, lorsqu'elle existe, est unique.

Démonstration. Soit $f : I \to \mathbb{R}$ ayant deux limites ℓ et ℓ' en $a \in \bar{I}$. Soit $V \in \mathscr{V}(\ell)$ et $V' \in \mathscr{V}(\ell')$; alors par définition de limite il existe $W, W' \in \mathscr{V}(a)$ tels que

$$f(W \cap I) \subset V \text{ et } f(W' \cap I) \subset V'\ .$$

$W \cap W'$ est un voisinage de a intersectant I et, si $x \in W \cap W' \cap I$ alors $f(x) \in V \cap V'$. On en déduit que $V \cap V' \neq \emptyset$, ce qui entraîne par contraposée de la proposition XI.2 que $\ell = \ell'$. □

Proposition XI.4. Soit $f : I \to \mathbb{R}$ et soit $a \in \overline{\mathbb{R}}$. Si f admet une limite réelle en a, alors elle est bornée au voisinage de ce point.

Démonstration. Comme ℓ est réel, ses voisinages sont bordés, d'où le résultat par définition de limite. □

Définition XI.5. Une fonction $f : I \to \mathbb{R}$ est dite **continue** en un point $a \in I$ si elle admet une limite finie en a. Si cela est vrai en tout point de I, on dira que f est continue sur I.

Notation. L'ensemble des fonctions continues sur I sera noté $\mathscr{C}^0(I, \mathbb{R})$ ou simplement $\mathscr{C}^0(I)$.

➠ **Exemple XI.5.** La fonction $f : x \mapsto \sqrt{x}$ est continue sur $\mathbb{R}_+$.

Proposition XI.5. Soit $f : I \to \mathbb{R}$ et soit $a \in I$. Alors :

$$f \text{ est continue en } a \iff f(x) \xrightarrow[x \to a]{} f(a)\ .$$

Démonstration. ($\Uparrow$) Trivial.

($\Downarrow$) Supposons que f soit continue en a, *i.e* qu'il existe $\ell \in \mathbb{R}$ tel que $f(x) \xrightarrow[x \to a]{} \ell$.

Cela signifie que :

$$\forall V \in \mathscr{V}(\ell),\ \exists W \in \mathscr{V}(a),\ f(W \cap I) \subset V\ .$$

Or, pour tout $W \in \mathscr{V}(a)$, $a \in W \cap I$ ce qui entraîne que :

$$f(a) \in \bigcap_{V \in \mathscr{V}(\ell)} V = \{\ell\}$$

d'où le résultat.

□

➠ **Exemple XI.6.**

— Toutes les fonctions usuelles (chapitres I, VI et VII) sont continues sur leurs ensembles de définition respectifs.

— La fonction suivante n'est pas continue en 0 :

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{sinon} \end{cases} \end{aligned} .$$

En effet, si elle admettait une limite finie en 0, celle–ci devrait être égale à 1 et -1 à la fois, ce qui est, vous en conviendrez, compliqué.

Proposition XI.6. Soit $f : I \to \mathbb{R}$ et soit $a \in \bar{I}$. Alors :

(i) $f(x) \xrightarrow[x \to a]{} 0 \iff |f(x)| \xrightarrow[x \to a]{} 0$;

(ii) si g est une fonction bornée au voisinage de a et si $f(x) \xrightarrow[x \to a]{} 0$ alors

$$f(x)g(x) \xrightarrow[x \to a]{} 0\ .$$

Démonstration. Adapter la démonstration de propriétés similaires vues dans le chapitre IX. □

✌ **Remarque XI.3.** Quitte à "épointer" les voisinages (leur retirer le point a), nous pouvons définir une notion de limite en a pour $f : I \setminus \{a\} \to \mathbb{R}$ vérifiant toutes les propriétés vues ici.

c) Opérations sur les limites

Nous adaptons dans ce paragraphe les résultats démontrés sur les suites dans le chapitre IX. Nous laissons au lecteur le soin de transcrire les démonstrations *ad–hoc.*

Proposition XI.7. Soient $f, g : I \to \mathbb{R}$,$a \in \bar{I}$, $\lambda \in \mathbb{R}$, $\ell, \ell' \in \overline{\mathbb{R}}$ tels que $f(x) \xrightarrow[x \to a]{} \ell$ et $g(x) \xrightarrow[x \to a]{} \ell'$. Alors on a, sauf dans les cas d'indétermination :

(i) $f(x) + \lambda g(x) \xrightarrow[x \to a]{} \ell + \lambda \ell'$;

(ii) $f(x)g(x) \xrightarrow[x \to a]{} \ell\ell'$;

(iii) **si** $\ell \neq 0$:
$$\frac{1}{f(x)} \xrightarrow[x \to a]{} \frac{1}{\ell} \text{ avec la convention que } \frac{1}{\pm\infty} = 0\,.$$

Proposition XI.8 (Composition des limites)**.** Soient I, J deux intervalles de $\mathbb{R}$, $a \in \bar{I}$, $b \in \bar{J}$, $\ell \in \overline{\mathbb{R}}$ et soient $f : I \to \mathbb{R}$ et $g : J \to \mathbb{R}$ deux fonctions telles que :

— $f(I) \subset J$;

— $f(x) \xrightarrow[x \to a]{} b$;

— $g(x) \xrightarrow[x \to b]{} \ell$.

Alors :
$$g \circ f(x) \xrightarrow[x \to a]{} \ell\,.$$

✖ **ATTENTION :** il est impératif de bien vérifier toutes les hypothèses avant d'appliquer ce résultat.

Démonstration. Soit $V \in \mathscr{V}(\ell)$; alors il existe $W \in \mathscr{V}(b)$ tel que $g(W \cap J) \subset V$. Mais il existe $W' \in \mathscr{V}(a)$ tel que $f(W' \cap I) \subset W \cap J$ et donc :
$$g \circ f(W' \cap I) \subset g(W \cap J) \subset V$$
d'où le résultat. □

➠ **Exemple XI.7.** La fonction définie sur $\mathbb{R}_+^*$ par $x \mapsto \sin\left(\frac{1}{x}\right)$ peut être vue comme la composée $g \circ f$ avec :
$$f : \mathbb{R}_+^* \to \mathbb{R}$$
$$x \mapsto \frac{1}{x}$$

et

$$g : \begin{array}{l} \mathbb{R}_+^* \mapsto \mathbb{R} \\ x \mapsto \sin(x) \end{array} .$$

Or :
- $f(\mathbb{R}_+^*) \subset \mathbb{R}_+^*$;
- $f(x) \xrightarrow[x\to\infty]{} 0$;
- $g(x) \xrightarrow[x\to 0]{} 0$,

donc, par composition des limites :

$$\sin\left(\frac{1}{x}\right) \xrightarrow[x\to\infty]{} 0 .$$

Proposition XI.9. Soient $f, g : I \to \mathbb{R}$ et soit $a \in \overline{\mathbb{R}}$. On suppose que $f(x) \xrightarrow[x\to a]{} \ell$ et $g(x) \xrightarrow[x\to a]{} \ell'$, avec $\ell, \ell' \in \overline{\mathbb{R}}$. Alors :

(i) si $f \leq g$ au voisinage de a, $\ell \leq \ell'$;

(ii) si $\ell = \ell'$ et que h est une fonction définie au voisinage de a telle que $f \leq h \leq g$, alors $h(x) \xrightarrow[x\to a]{} \ell$.

✌ **Remarque XI.4.** Comme pour le théorème IX.12, on peut énoncer une "version infinie" du point (ii).

➠ **Exemple XI.8.** Au voisinage de $+\infty$, on a :

$$\frac{-1}{x} \leq \frac{\sin(x)}{x} \leq \frac{1}{x}$$

et donc

$$\frac{\sin(x)}{x} \xrightarrow[x\to\infty]{} 0 \ : .$$

d) Caractérisation séquentielle de la limite

La proposition suivante est une complétion et reformulation du théorème IX.10.

Proposition XI.10. Soit $f : I \to \mathbb{R}$, $a \in \bar{I}$ et $\ell \in \overline{\mathbb{R}}$. Alors :

$$f(x) \xrightarrow[x\to a]{} \ell$$
$$\Longleftrightarrow$$
pour toute suite $(x_n)_n \in I^{\mathbb{N}}$ convergeant vers a, $f(x_n) \to \ell$.

✌ **Remarque XI.5.** Cette caractérisation nous servira surtout à démontrer une absence de limite. Par exemple, sin n'admet pas de limite en $+\infty$ car si une telle limite existait elle serait égale à $\lim \sin(2n\pi) = 0$ et $\lim \sin\left(2n\pi + \frac{\pi}{2}\right) = 1$.

Démonstration. Faisons le pour a et ℓ réels.

($\Downarrow$) Soit $(x_n)_n \in I^{\mathbb{N}}$ convergeant vers a et soit $\delta > 0$; alors :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |x_n - a| \leq \delta$$

i.e

$$\forall W \in \mathscr{V}(a), \exists\, N \in \mathbb{N},\ \forall n \geq N,\ x_n \in W\ .$$

Soit $\varepsilon > 0$; alors $V = [\ell - \varepsilon, \ell + \varepsilon]$ est un voisinage de ℓ et donc il existe $W \in \mathscr{V}(a)$ tel que $f(W \cap I) \subset V$. On déduit de tout ceci que :

$$\forall \varepsilon > 0, \exists\, N \in \mathbb{N},\ \forall n \geq N,\ |x_n - \ell| \leq \varepsilon$$

d'où la convergence de $(f(x_n))_n$ vers ℓ.

($\Uparrow$) Démontrons le par contraposée. Si f ne tend pas vers ℓ en a, alors il existe $V = [\ell - \varepsilon, \ell + \varepsilon] \in \mathscr{V}(\ell)$ tel que

$$\forall W \in \mathscr{V}(a),\ f(W \cap I) \not\subset V\ .$$

En particulier, cela est vrai pour les voisinages $W_n = \left[a - \dfrac{1}{n}, a + \dfrac{1}{n}\right]$ (pour $n \geq 1$), et donc il existe une suite $(x_n)_n \in I^{\mathbb{N}}$ telle que :

— $\forall n \geq 1,\ |x_n - a| \leq \dfrac{1}{n}$;

— $\forall n \geq 1,\ |f(x_n) - \ell| > \varepsilon$.

Donc, par encadrement $x_n \to a$ et $(f(x_n))_n$ ne converge pas vers ℓ, d'où le résultat.

□

e) Limites directionnelles, fonctions monotones

Définition XI.6. Soit $f : I \to \mathbb{R}$ et soit $a \in \bar{I}$.

— Si $a \neq \inf I$, on appelle **limite à gauche de** f **en** a la quantité (si elle existe)

$$\lim_{x \to a^-} f(x) = \lim_{x \to a} f_{|I \cap]-\infty, a[}(x)\ .$$

— Si $a \neq \sup I$, on appelle **limite à droite de** f **en** a la quantité (si elle existe)

$$\lim_{x \to a^+} f(x) = \lim_{x \to a} f_{|I \cap]a, +\infty[}(x)\ .$$

Exemple XI.9. $\dfrac{1}{x} \xrightarrow[x \to 0^+]{} +\infty$.

Proposition XI.11. Soit $f : I \to \mathbb{R}$ et soit $a \in \mathring{I}$. Alors :

$$f \text{ admet une limite en } a$$
$$\iff$$
$$f \text{ admet une limite à gauche et à droite en } a \text{ qui sont égales à } f(a).$$

✖ **Attention :** La condition d'égalité à $f(a)$ n'est pas superfétatoire : l'indicatrice $\mathbb{1}_{\{0\}}$ n'admet pas de limite en 0 (utiliser la caractérisation séquentielle en se rappelant que la suite nulle converge vers 0) et pourtant y admet une limite à gauche et à droite.

➠ **Exemple XI.10.** Revenir à et méditer sur

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{sinon} \end{cases} \end{aligned} .$$

Proposition XI.12. Soit $f : I \to \mathbb{R}$ une fonction croissante et soit $a \in \bar{I}$. Alors :

(i) si $a \in \mathring{I}$ alors les limites directionnelles de f en a existent, avec les égalités

$$\lim_{x\to a^+} f(x) = \inf\{f(x) \,|\, x > a\}$$

et

$$\lim_{x\to a^-} f(x) = \sup\{f(x) \,|\, x < a\} ,$$

avec de plus

$$\lim_{x\to a^-} f(x) \leq f(a) \leq \lim_{x\to a^+} f(x) ;$$

(ii) si $a = \sup(I)$, et $a \in I$, alors f admet une limite à gauche en a donnée par

$$\lim_{x\to a^-} f(x) = \sup\{f(x) \,|\, x < a\} ,$$

avec de plus

$$\lim_{x\to a^-} f(x) \leq f(a) ;$$

(iii) si $a = \sup(I)$, et $a \in I$, alors f admet une limite à gauche en a donnée par

$$\lim_{x\to a^-} f(x) = \sup\{f(x) \,|\, x < a\} ,$$

cette borne supérieure étant prise dans $\overline{\mathbb{R}}$.

✌ **Remarque XI.6.** Ce résultat peut évidemment être adapté dans le cas décroissant.

Démonstration. Démontrons le point (i) ; pour ce faire, remarquons que l'ensemble

$$\mathscr{E} = \sup\{f(x) \,|\, x < a\}$$

est une partie de $\mathbb{R}$ non vide et majorée par $f(a)$; elle admet donc une borne supérieure $M \leq f(a)$. De plus, pour tout $\varepsilon > 0$, $M - \varepsilon$ ne majore pas $\mathscr{E}$ et donc il existe $x_0 < a$ tel que $f(x_0) > M - \varepsilon$, *i.e* $|M - f(x_0)| \leq \varepsilon$.

Si $x_1 \in]x_0, a[$, alors $f(x_1) \geq f(x_0) \geq M - \varepsilon$; ce qui signifie que :

$$f(]x_0, a[) \subset [f(a) - \varepsilon, f(a) + \varepsilon] .$$

On termine en remarquant que $]x_0, a[$ est l'intersection d'un voisinage de a avec $]-\infty, a[$, ce qui permet de conclure quant à la limite, et en raisonnant de façon similaire pour obtenir la seconde inégalité. □

2.– Fonctions continues

a) Opérations sur les fonctions continues

Les propriétés relatives aux opérations sur les limites se traduisent par les résultats suivants, muni des opérations présentées dans le chapitre I.

Proposition XI.13. $(\mathscr{C}^0(I), +, \times)$ est un anneau commutatif, de neutres $x \mapsto 0$ et $x \mapsto 1$. De plus, pour tous $\lambda \in \mathbb{R}$ et $f \in \mathscr{C}^0(I)$, $\lambda f \in \mathscr{C}^0(I)$.

Proposition XI.14. La composée de deux fonctions continues compatibles est continue.

Démonstration. Cela découle de la composition des limites. □

b) Prolongement par continuité

La question que nous nous posons dans ce paragraphe est la suivante : étant donné une fonction continue sur une intervalle "à trou", est–il possible de la prolonger de façon continue sur l'intervalle entier ?

Définition XI.7. Soit $a \in I$ et soit $f : I \setminus \{a\} \to \mathbb{R}$ une fonction continue. On dit que f est **prolongeable par continuité** en a si il existe $g \in \mathscr{C}(I)$ telle que

$$g_{|_{I\setminus\{a\}}} = f \ .$$

Proposition XI.15. Soit $a \in I$ et soit $f : I \setminus \{a\} \to \mathbb{R}$ une fonction continue. Alors :

$$f \text{ est prolongeable par continuité en } a$$
$$\Longleftrightarrow$$
$$f \text{ admet une limite réelle en } a.$$

Démonstration.

($\Uparrow$) Immédiat via :

$$g : x \mapsto \begin{cases} f(x) & \text{si } x \in I \\ \lim\limits_{x\to a} f(x) & \text{si } x = a \end{cases} .$$

($\Downarrow$) Si il existe un prolongement par continuité g de f en a, alors $g(x) \xrightarrow[x\to a]{} g(a)$ et donc, comme $g_{|_{I\setminus\{a\}}} = f$ alors $f(x) \xrightarrow[x\to a]{} g(a)$.

□

Exemple XI.11. La fonction définie sur $\mathbb{R}_+^*$ par $x \mapsto e^{-\frac{1}{x}}$ se prolonge par continuité en 0 par 0. Attention, elle n'est pas prolongeable par continuité sur $\mathbb{R}$ car

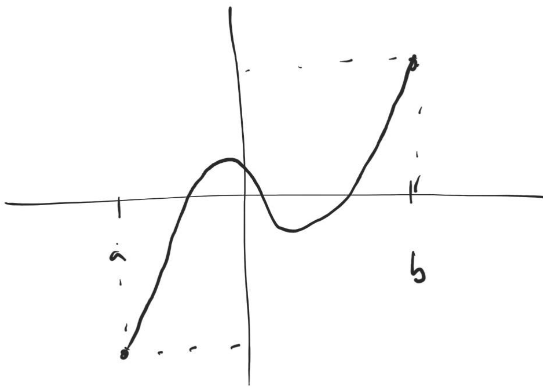
$$\lim_{x \to 0^-} e^{-\frac{1}{x}} = +\infty \, .$$

c) Images d'intervalles

Théorème XI.16 (Théorème des valeurs intermédiaires)**.**
Soit $f \in \mathscr{C}^0([a,b])$ telle que $f(a)f(b) < 0$. Alors :

$$\exists\, c \in]a,b[,\ f(c) = 0 \, .$$

Remarque XI.7.
— La condition $f(a)f(b) < 0$ impose que $f(a)$ et $f(b)$ soient de signes opposés.
— Il n'y a pas nécessairement unicité du point d'annulation c.



Démonstration. Pour fixer les idées, supposons $f(a) < 0$. L'ensemble

$$\mathscr{E} = \{x \in [a,b] \mid f(x) \leq 0\}$$

est une partie non vide de $\mathbb{R}$ majorée par b ; elle admet donc un supremum $c \in [a,b]$. De plus, par continuité, f est strictement négative (resp. strictement positive) au voisinage de a (resp. de b) donc $c \notin \{a,b\}$.

Notons ensuite que :

$$\exists\, N \in \mathbb{N}^*,\ \forall n \geq N,\ c + \frac{1}{n} \in [a,b]$$

et donc, comme $c + \dfrac{1}{n} \geq c$, on a

$$f\left(c + \frac{1}{n}\right) > 0$$

et donc, en passant à la limite (via la proposition XI.10), $f(c) \geq 0$. De plus, par caractérisation séquentielle de la borne supérieure (proposition IX.24), il existe une suite $(x_n)_n \in \mathscr{E}^{\mathbb{N}}$ convergeant vers c ; ne reste qu'à passer à la limite dans l'inégalité $f(x_n) \leq 0$ pour obtenir que $f(c) \leq 0$. *In fine*, $f(c) = 0$.

□

✍ **Exercice XI.1.** Démontrer que toute fonction continue de $[0, 1]$ dans lui même admet un point fixe, *i.e* un $x \in [0, 1]$ tel que $f(x) = x$.

➥ **Correction :** *Si $f(0) = 0$ ou $f(1) = 1$, c'est terminé. Sinon, posons $g : x \mapsto x - f(x)$; alors $g(0) = -f(0) < 0$ et $g(1) = 1 - f(1) > 0$; par TVI, il existe $x \in]0, 1[$ tel que $g(x) = 0$, i.e $f(x) = x$.*

Corollaire XI.16.a. Soit $f \in \mathscr{C}^0([a, b])$; posons $\alpha = f(a)$ et $\beta = f(b)$. Alors :

$$\forall \gamma \in [\alpha, \beta] \text{ (ou } [\beta, \alpha]), \ \exists c \in [a, b], \ f(c) = \gamma \, .$$

Corollaire XI.16.b. L'image d'un intervalle par une fonction continue est un intervalle.

Démonstration. Se rappeler que les intervalles de $\mathbb{R}$ sont exactement les convexes et utiliser le corollaire précédent. □

Corollaire XI.16.c. Soit $f \in \mathscr{C}^0([a, b])$. Alors :

(i) f admet un maximum M et un minimum m atteints sur $[a, b]$;

(ii) $f([a, b]) = [m, M]$.

En particulier, **l'image d'un segment par une fonction continue est un segment.**

La démonstration de ce dernier résultat est hors–programme et repose sur une combinaison du TVI et du théorème de Bolzano–Weierstras (IX.19).

➠ **Exemple XI.12.** Le TVI nous permet de démontrer que le cosinus est surjectif sur $[-1, 1]$.

◇ Application : méthode de dichotomie

Soit $f \in \mathscr{C}^0([a, b])$ telle que $f(a)f(b) < 0$. On peut alors définir par récurrence trois suites $(a_n)_n$, $(b_n)_n$ et $(c_n)_n$ selon le procédé suivant : posons $a_0 = a$, $b_0 = b$ et $c_0 = \frac{a+b}{c}$, puis, pour tout $n \geq 0$

— si $f(a_n)f(c_n) < 0$, on pose $a_{n+1} = a_n$ et $b_{n+1} = c_n$;

— dans le cas contraire, on pose $a_{n+1} c_n$ et $b_{n+1} = b_n$.

Enfin, dans tous les cas, on pose :

$$c_{n+1} = \frac{a_{n+1} + b_{n+1}}{2} \, .$$

On resserre donc petit à petit une "fenêtre" autour du point d'annulation de f, ce qui entraîne que la suite $(c_n)_n$ converge vers un point d'annulation de f. On peut implémenter cette procédure en `python` via le code suivant.

```python
def dicho(f,a,b,p):
    if f(a)*f(b) > 0:
        raise ValueError("f ne s'annule pas")
    while b-a>10**(-p):
        c=(a+b)/2
        if f(c) == 0:
            return c
        if f(a)*f(c)<0:
            b=c
        else:
            a=c
    return c
```

d) Continuité et (stricte) monotonie

Les résultats de ce paragraphe seront admis, leurs démonstrations dépassant largement nos capacités techniques à ce stade.

Proposition XI.17. Toute fonction continue et injective est strictement monotone.

Théorème XI.18 (Théorème de la bijection)**.**
Soit I un intervalle de $\mathbb{R}$ et soit $f : I \to \mathbb{R}$ une application continue strictement monotone. Alors :

(i) f réalise une bijection de I vers $J = f(I)$;

(ii) J est un intervalle du même type que I (*i.e* ouvert et fermé le même nombre de fois) ;

(iii) $f^{-1} : J \to I$ est de même monotonie que f.

✌ **Remarque XI.8.** Nous avons vu de nombreuses applications de ces résultats dans les chapitres I et VII.

3.⁻ Brêve extension aux fonctions à valeurs complexes

Une fois n'est pas coutume, tous les résultats de ce chapitre ne faisant pas intervenir d'inégalité se généralisent aux fonctions $f : I \to \mathbb{C}$ lorsque I est un intervalle de $\mathbb{R}$ (on ne parlera pas ici de fonction de variables complexes). En particulier, notons que le TVI ne se généralise pas. On peut par contre passer à la limite dans les parties réelle et imaginaire.

➠ **Exemple XI.13.** $e^{i\theta} \underset{\theta \to 0}{\longrightarrow} 1$.

Chapitre XII

Dérivation

Dans tout ce chapitre, nous fixons un intervalle I de $\mathbb{R}$ d'intérieur non vide.

1.⁻ Notion de dérivée

a) Qu'est–ce ?

Définition XII.1. Soit $f : I \to \mathbb{R}$ une fonction et $a \in I$. On dit que f est **dérivable en** a si son taux d'accroissement en a admet une limite finie, *i.e* si

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \text{ existe.}$$

Cette limite est alors appelée **nombre dérivé de** f **en** a.

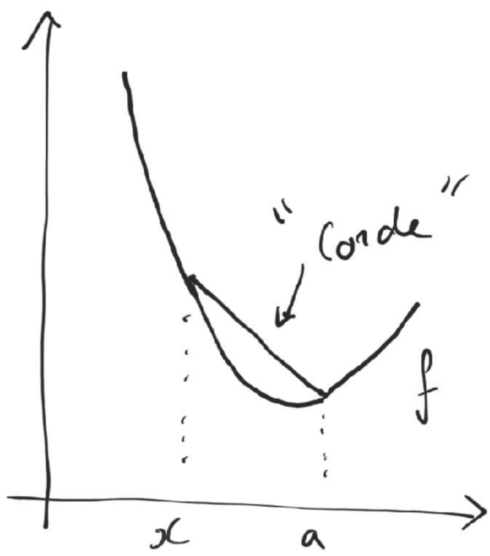
Notation. *Si* f est dérivable, on pose $f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$.

✌ **Remarque XII.1.**

— Il peut être parfois utile de se ramener à l'écriture alternative suivante (lorsque les quantités existent) :

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h} .$$

— Géométriquement, les taux d'accroissement correspondent aux pentes des "cordes" associées à la courbe représentative de f.



Exemple XII.1.

Pour la fonction $f : x \mapsto x^2$ on a, au voisinage de $a \in \mathbb{R}$ (excluant a) :

$$\begin{aligned}\frac{f(x)-f(a)}{x-a} &= \frac{x^2-a^2}{x-a} \\ &= x+a \\ &\xrightarrow[x\to a]{} 2a\end{aligned}$$

et donc f est dérivable en a et $f'(a) = 2a$.

Proposition XII.1. Soit $f : I \to \mathbb{R}$ une fonction et soit $a \in I$. Alors :

f est dérivable en a

$\iff$

il existe $\lambda \in \mathbb{R}$ et ε une fonction définie au voisinage de a et de limite nulle en a tels que $f(x) = f(a) + \lambda(x-a) + (x-a)\varepsilon(x)$ au voisinage de a.

Dans ce cas, on a de plus $f'(a) = \lambda$.

Remarque XII.2. On retrouve le résultat vu en première : la courbe représentative de f peut être approximée par une droite au voisinage de a ; celle–ci admet pour équation

$$y = (x-a)f'(a) + f(a)$$

et est appelée **tangente à la courbe de f en a**. Notons au passage que si $f'(a) = 0$, la tangente est **horizontale** et que si le taux d'accroissement tend vers l'infini, notre courbe représentative admet une **tangente verticale** (elle n'est alors évidemment pas dérivable).

Démonstration.

($\Downarrow$) Si f est dérivable en a, il suffit de poser $\varepsilon : x \mapsto \dfrac{f(x)-f(a)}{x-a} - f'(a)$. On a alors bien $\varepsilon(x) \xrightarrow[x\to a]{} 0$ et l'égalité voulue avec $\lambda = f'(a)$.

($\Uparrow$) On a, au voisinage de a (excluant a) :

$$\begin{aligned}\frac{f(x)-f(a)}{x-a} &= \lambda + \varepsilon(x)\\ &\xrightarrow[x\to a]{} \lambda\end{aligned}$$

d'où le résultat.

□

Proposition XII.2. Toute fonction dérivable en un point de I y est continue.

✖ **ATTENTION :** la réciproque est **FAUSSE**. La fonction valeur absolue est continue en 0 et pourtant :

$$\frac{|x|}{x} \xrightarrow[x\to 0^+]{} 1$$

et

$$\frac{|x|}{x} \xrightarrow[x\to 0^-]{} -1$$

ce qui signifie que cette fonction n'est pas dérivable en 0.

Démonstration. Soit $a \in I$ et soit f dérivable en a. Alors d'après la proposition XII.1 on a, au voisinage de a :

$$\begin{aligned}f(x) &= f(a) + (x-a)f'(a) + (x-a)\underbrace{\varepsilon(x)}_{\xrightarrow[x\to a]{} 0}\\ &\xrightarrow[x\to a]{} f(a)\end{aligned}$$

et donc f est continue en a. □

Définition XII.2. Une fonction dérivable en tout point de I est dite dérivable sur I. La fonction

$$\begin{aligned}f' : I &\to \mathbb{R}\\ x &\mapsto f'(x)\end{aligned}$$

est alors appelée **dérivée de** f.

Notation. On note $\mathscr{D}(I)$ l'ensemble des fonctions dérivables sur I. La dérivée pourra être noté f', $\dfrac{\mathrm{d}f}{\mathrm{d}x}$ ou même $\mathrm{D}f$, mais **surtout pas** $\dfrac{\partial f}{\partial x}$.

◇ **Dérivées usuelles**

Les fonctions vues dans les chapitres I et VII sont dérivables en presque tout point de leurs ensembles de définition respectifs. Nous rappelons dans le tableau ci–ensuite les valeurs de leurs nombres dérivés en les points *ad–hoc*.

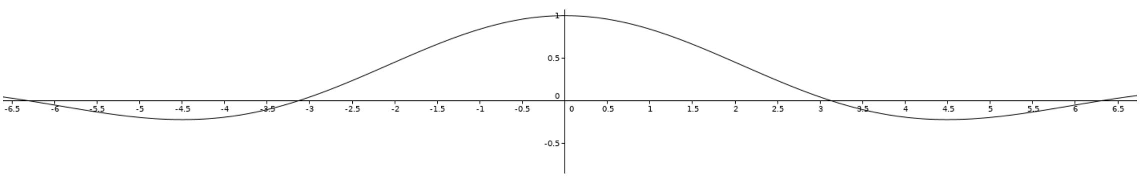
Valeur de $f(x)$	**Ensemble de dérivabilité**	**Valeur de** $f'(x)$
x^a	$\mathbb{R}_+^*$	ax^{a-1}
a^x $(a > 0,\ a \neq 1)$	$\mathbb{R}$	$\ln(a)a^x$
$\log_a(x)$ $(a > 0)$	$\mathbb{R}_+^*$	$\frac{1}{x\ln(a)}$
$\cos(x)$	$\mathbb{R}$	$-\sin(x)$
$\sin(x)$	$\mathbb{R}$	$\cos(x)$
$\tan(x)$	$\mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\right\}$	$1 + \tan^2(x) = \frac{1}{\cos^2(x)}$
$\text{arcos}(x)$	$]-1,1[$	$-\frac{1}{\sqrt{1-x^2}}$
$\arcsin(x)$	$]-1,1[$	$\frac{1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\mathbb{R}$	$\frac{1}{1+x^2}$
$\text{ch}(x)$	$\mathbb{R}$	$\text{sh}(x)$
$\text{sh}(x)$	$\mathbb{R}$	$\text{ch}(x)$
$\text{th}(x)$	$\mathbb{R}$	$1 - \text{th}^2(x) = \frac{1}{\text{ch}^2(x)}$

Un cas particulier mérite d'être souligné : les fonctions puissances $x \mapsto x^a$ sont dérivables sur $\mathbb{R}$ si $a \in \mathbb{N}$, $\mathbb{R}^*$ si $a \in \mathbb{Z} \setminus \mathbb{N}$. Les racines n–ièmes ne sont jamais dérivables en 0,mais sont dérivables sur $\mathbb{R}^*$ si n est impair.

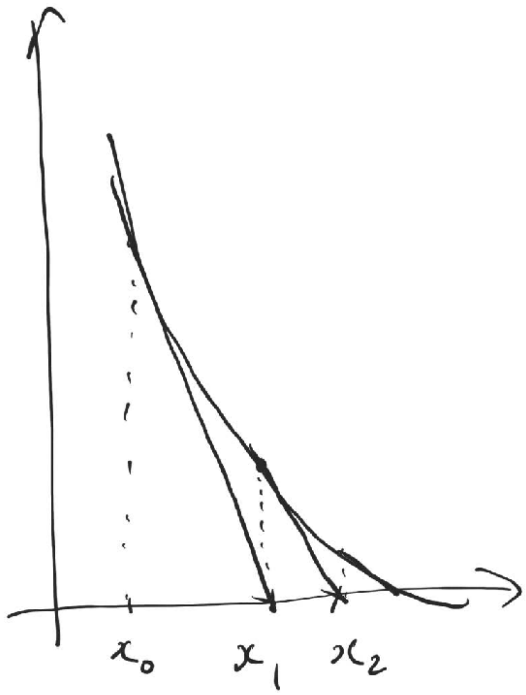
➠ **Exemple XII.2.** Comme la fonction sinus est dérivable en 0, on a

$$\frac{\sin(x)}{x} \xrightarrow[x\to 0]{} \cos(0) = 1$$

et donc la fonction sinus cardinal sinc : $x \mapsto \dfrac{\sin(x)}{x}$, définie sur $\mathbb{R}^*$ est prolongeable par continuité à $\mathbb{R}$.



◇ **Méthode de Newton**



Soit $f : I \to \mathbb{R}$ une fonction dérivable sympathique s'annulant en un point c de I. Alors, f sera localement croissante ou décroissante au voisinage de c (puisqu'on vous dit qu'elle est sympathique). Si on fixe x_0 pas trop loin de c, la tangente de f en x_0 va "pointer" vers le point d'annulation c au sens suivant : le point d'intersection x_1 entre cette dernière et l'axe des abscisses sera plus proche de c que x_0. Ce point x_1 doit vérifier :

$$0 = f(x_0) + (x_1 - x_0)f'(x_0)$$

et donc, si $f'(x_0)$ a la gentillesse de ne pas s'annuler trop fort $x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)}$.

À supposer que nous habitions dans le monde magique des fonctions gentiment localement monotones sans annulation intempestive de leurs fonctions dérivées, nous pouvons donc itérer ce procédé via la formule de récurrence :

$$\forall n \in \mathbb{N},\ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

et cette suite convergera vaguement vers le point c tel que $f(c) = 0$. En `python`, cela donne une procédure du style décrit *infra.*

```
def newton(f,fp,x0,N):
    x=x0
    for i in range(N):
        x = x - f(x)/fp(x)
    return x
```

Notons que cette fonction prend en argument la fonction `f` et sa dérivée `fp` et que nous avons utilisé un nombre d'itérations maximal comme condition d'arrêt ; cette méthode est en effet assez peu fiable dans le cas général et nous souhaitions éviter une boucle infinie ; il convient de toujours vérifier les résultats qu'elle renvoie.

b) Dérivées directionnelles

Définition XII.3. Soit $f : I \to \mathbb{R}$ et soit $a \in \bar{I}$.
- Si $a \neq \inf I$, on appelle **dérivée à gauche de f en a** la quantité (si elle existe)
$$f'_g(a) = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a}.$$
- Si $a \neq \sup I$, on appelle **dérivée à droite de f en a** la quantité (si elle existe)
$$f'_d(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a}.$$

Exemple XII.3. Soit $f : x \mapsto |x|$; alors f est dérivable à droite et à gauche en 0 et $f'_d(0) = 1$, $f'_g(0) = -1$.

Proposition XII.3. Soit $f : I \to \mathbb{R}$ et soit $a \in \mathring{I}$. Alors :

$$f \text{ est dérivable en } a$$
$$\iff$$
$$f \text{ admet une dérivée à gauche et à droite en } a \text{ qui sont égales.}$$

Démonstration. Cela découle du résultat analogue sur les limites directionnelles, à savoir la proposition XI.11 □

c) Classes de fonctions

Définition XII.4. Soit $k \geq 1$; une fonction $f : I \to \mathbb{R}$ est dite k–fois dérivable sur I si :
- $k = 1$ et f est dérivable sur I ; on note alors $f^{(1)} = f'$;
- $k \geq 2$ et la fonction $f^{(k-1)}$ est dérivable sur I ; on note alors $f^{(k)} = (f^{(k-1)})'$.

Notation. On note $\mathscr{D}^k(I)$ l'ensemble des fonctions k fois dérivables sur I. La fonction dérivée k–ième $f^{(k)}$ pourra également être notée $\mathrm{D}^k f$ ou $\dfrac{\mathrm{d}^k f}{\mathrm{d}x^k}$.

Exemple XII.4. La fonction $x \mapsto x^2$ est 7 fois dérivable sur $\mathbb{R}$, de dérivée septième nulle. Sa dérivée seconde est $x \mapsto 2$.

Convention. On notera, pour $f \in \mathscr{C}^0(I)$, $f^{(0)} = f$.

Remarque XII.3. Soit $k \geq 1$; alors si f est une fonction k–fois dérivable sur I, pour tout $0 \leq j < k$, la fonction $f^{(j)}$ est continue (car dérivable).

Définition XII.5. Soit $k \in \mathbb{N}^*$. On dira qu'une fonction $f : I \to \mathbb{R}$ est **de classe $\mathscr{C}^k$ sur I** si :
— $f \in \mathscr{D}^k(I)$;
— $f^{(k)} \in \mathscr{C}^0(I)$.
Si f est de classe $\mathscr{C}^k$ sur I pour tout $k \in \mathbb{N}^*$, on dira que f est **de classe $\mathscr{C}^\infty$ sur I**.

Notation. On notera, sans surprise, $\mathscr{C}^k(I)$ (resp. $\mathscr{C}^\infty$) l'ensemble des fonctions de classe $\mathscr{C}^k$ (resp. $\mathscr{C}^\infty$) sur I.

Exemple XII.5. Les fonctions usuelles sont de classe $\mathscr{C}^\infty$ dans les ensembles mentionnés dans le tableau du paragraphe a).

Remarque XII.4.
— On a la chaîne d'inclusions suivante :

$$\mathscr{C}^0(I) \supset \mathscr{D}^1(I) \supset \mathscr{C}^1(I) \supset \mathscr{D}^2(I) \supset \ldots \supset \mathscr{C}^{(k)}(I) \supset \mathscr{D}^{(k+1)}(I) \supset \ldots \supset \mathscr{C}^\infty(I)$$

et toutes ces inclusions sont **strictes**. Par exemple, $\sqrt{\cdot} \in \mathscr{C}^0(\mathbb{R}_+) \setminus \mathscr{D}^1(\mathbb{R}_+)$ et $x \mapsto x^{\frac{3}{2}} \in \mathscr{C}^1(\mathbb{R}_+) \setminus \mathscr{D}^2(\mathbb{R}_+)$.
— Si f est de classe $\mathscr{C}^k$ sur I, alors $f' \in \mathscr{C}^{(k-1}(I)$.
— On a

$$\mathscr{C}^\infty(I) = \bigcap_{k\in\mathbb{N}^*} \mathscr{D}^k(I) \ .$$

d) Opérations sur les dérivées

◇ Au premier ordre

Proposition XII.4. Soient $f, g \in \mathscr{D}^1(I)$ et soit $\lambda \in \mathbb{R}$. Alors :
(i) $f + \lambda g$ est dérivable sur I et $(f + \lambda g)' = f' + \lambda g'$;
(ii) $f \times g$ est dérivable sur I et $(f \times g)' = f' \times g + f \times g'$;
(iii) si g ne s'annule pas sur I, $\frac{f}{g}$ est dérivable sur I et

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2} \ ;$$

Démonstration. (i) Immédiat par opération sur les limites (passer aux taux d'accroissement).
(ii) Soient $x, a \in I$ tels que $x \neq a$. Alors :

$$\begin{aligned}\frac{f(x)g(x) - f(a)g(a)}{x-a} &= \frac{f(x)g(x) - f(x)g(a) + f(x)g(a) - f(a)g(a)}{x-a} \\ &= f(x)\frac{g(x)-g(a)}{x-a} + g(a)\frac{f(x)-f(a)}{x-a} \\ &\xrightarrow[x\to a]{} f(a)g'(a) + g(a)f'(a)\end{aligned}$$

d'où le résultat.

(iii) S'obtient de façon analogue au précédent, avec des calculs un tantinet plus crades.

□

✖ **ATTENTION :** il convient de prêter à la rédaction une attention particulière ; en particulier, on ne saurait dériver une fonction sans avoir préalablement justifié sa dérivabilité.

Proposition XII.5. Soient I, J deux intervalles d'intérieur non vide de $\mathbb{R}$ et soient $f \in \mathscr{D}^(I)$ et $g \in \mathscr{D}^1(J)$ telles que $f(I) \subset J$. Alors :

(i) $g \circ f \in \mathscr{D}^1(I)$;

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

Démonstration. Soit $a \in I$ et soit $b = f(a) \in J$. Alors, par proposition XII.1, il existe $W \in \mathscr{V}(b)$ tel que :

$$\forall y \in W \quad : g(y) = g(b) + (y-b)g'(b) + (y-b)\varepsilon(y)$$

avec $\varepsilon(y) \xrightarrow[y \to b]{} 0$. Comme $f(a) = b$ et que f est continue, il existe de plus $V \in \mathscr{V}(a)$ tel que $f(V) \subset W$ et donc, si $x \in V \cap I$ on a $f(x) \in W$ *ergo* :

$$g \circ f(x) = g \circ f(a) + (f(x) - f(a))g' \circ f(a) + (f(x) - f(a))\underbrace{\varepsilon(f(x))}_{\xrightarrow[x \to a]{} 0} \ .$$

Ainsi :

$$\begin{aligned} \frac{g \circ f(x) - g \circ f(a)}{x-a} &= \frac{f(x) - f(a)}{x-a} g' \circ f(a) + \frac{f(x) - f(a)}{x-a} \varepsilon(f(x)) \\ &\xrightarrow[x \to a]{} f'(a) g' \circ f(a) \end{aligned}$$

ce qui conclut la démonstration. □

➠ **Exemple XII.6.** La fonction définie sur $\mathbb{R}_+^*$ par $x \mapsto e^{\frac{1}{x}}$ est dérivable comme composée des fonctions dérivables

$$\begin{aligned} f : \mathbb{R}_+^* &\to \\ x &\mapsto \frac{1}{x} \end{aligned}$$

et $\exp : \mathbb{R} \to \mathbb{R}$. Sa dérivée vérifie, pour tout $x > 0$:

$$(\exp \circ f)'(x) = \frac{-1}{x^2} e^{\frac{1}{x}} \ .$$

Proposition XII.6. Soient I, J deux intervalles d'intérieur non vide de $\mathbb{R}$ et soit $f : I \to J$ une fonction **bijective et continue sur** I. Soit $a \in I$ tel que :
- f soit dérivable en a ;
- $f'(a) \neq 0$.

Alors
(i) f^{-1} est dérivable en $b = f(a)$;
(ii) $(f^{-1})'(b) = \dfrac{1}{f'(a)}$, *i.e* :

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))} .$$

Remarque XII.5.

- Ce théorème a donc **quatre** hypothèses à vérifier : deux globales (portant sur la fonction) et deux locales (relatives au point a).
- On retrouve le résultat énoncé dans le théorème I.4 : une bijection dérivable admet une réciproque dérivable en tout point de non–annulation de sa dérivée.
- Par la proposition XI.17, la fonction f est strictement monotone.

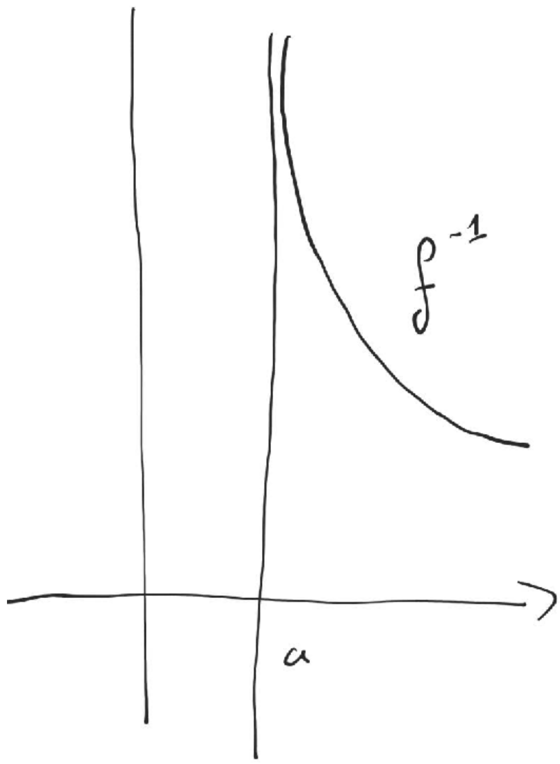
Exemple XII.7. Nous avons vu plusieurs applications de ce théorème dans le chapitre VII (fonctions trigonométriques réciproques).

Démonstration. Soit $V \in \mathscr{V}(a)$, $x \in V \setminus \{a\}$ et $y = f(x)$. Alors

$$\begin{aligned}\frac{f^{-1}(y) - f^{-1}(b)}{y - b} &= \frac{x - a}{f(x) - f(a)} \\ &\xrightarrow[x \to a]{} \frac{1}{f'(a)}\end{aligned}$$

d'où le résultat. □

Remarque XII.6. Si la dérivée de f s'annule en a, la fonction réciproque y admet une tangente verticale car le taux d'accroissement ci–dessus tend vers l'infini.



On retrouve, en reformulant, la proposition vue au chapitre I.

Corollaire XII.6.a. Soient I, J deux intervalles de $\mathbb{R}$ et soit $f : I \to J$ une fonction dérivable telle que f' **ne s'annule pas** sur I. Alors :

(i) f admet une réciproque f^{-1} dérivable sur J ;

(ii)
$$(f^{-1})' = \frac{1}{f' \circ f^{-1}} .$$

$\diamond$ Aux ordres supérieurs

Les choses se passent très bien concernant les dérivées d'ordre supérieur d'une combinaison linéaire ; les formules se compliquent (un peu) pour le produit, avec le résultat suivant du au mathématicien allemand Gottfried Wilhelm Leibniz (1646—1716).

Proposition XII.7 (Formule de Leibniz)**.** Soient $f, g \in \mathscr{C}^n(I)$ $(n \geq 1)$. Alors la fonction fg est de classe $\mathscr{C}^n$ sur I et :

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

Démonstration. Par récurrence. □

✍ **Exercice XII.1.** Le lecteur possédé de tendances masochistes pourra calculer les dérivées successives de $\sin \times \cos$.

La succession de résultats suivante est énoncée sans démonstration. Le lecteur pourra si il le souhaite combler ce manque à la sueur de son front.

Proposition XII.8.
(i) L'inverse d'une fonction de classe $\mathscr{C}^k$ ne s'annulant pas est de classe $\mathscr{C}^k$.
(ii) La composée de deux fonctions de classes $\mathscr{C}^k$ compatibles est de classe $\mathscr{C}^k$.
(iii) La réciproque d'une bijection de classe $\mathscr{C}^k$ dont **la dérivée première** ne s'annule pas est de classe $\mathscr{C}^k$.

Nous ne donnons pas dans les cas énoncés *supra* de formules pour le calcul explicite des dérivées successives, et nous suggérons au lecteur de nous en être reconnaissant.

2.⁻ Accroissements finis

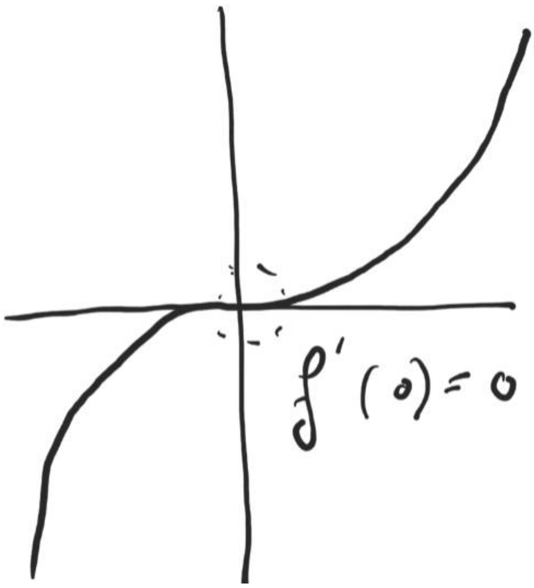
a) Extrema locaux

Nous pouvons reformuler les définitions vues au chapitre I en utilisant un langage un peu plus évolué. Notons le gain de place occasionné.

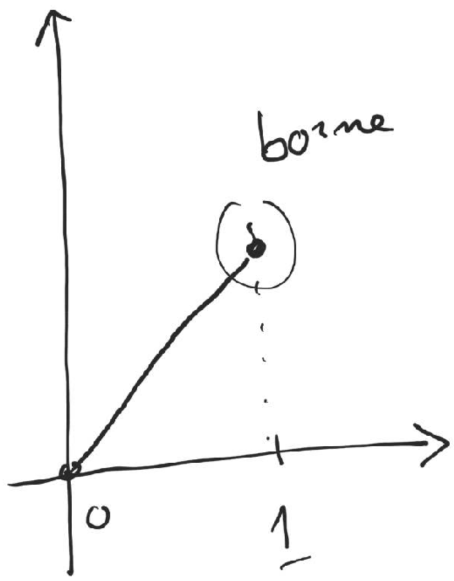
Définition XII.6. Soit $f : I \to \mathbb{R}$ une fonction. Si $a \in I$, on dit que :
— f admet un maximum local en a si $f(x) \leq f(a)$ au voisinage de a ;
— f admet un minimum local en a si $f(x) \geq f(a)$ au voisinage de a.

Proposition XII.9. Soit $f \in \mathscr{D}^1(I)$ et soit $a \in \mathring{I}$. **Si** f admet un extremum local en a, **alors** $f'(a) = 0$.

✖ **Attention :** la réciproque est **fausse**. En effet, la fonction $x \mapsto x^3$ admet pour dérivée $x \mapsto 3x^2$ qui s'annule en 0 sans que ce dernier ne soit un extremum local de la fonction initiale.



De même, il est **essentiel** que a ne soit pas une borne de I ; penser à la fonction $x \mapsto x$ sur $[0, 1]$...



Démonstration. Plaçons nous dans le cas où a est un maximum local de f. Alors, au voisinage de a :

— si $x < a$,
$$\frac{f(x) - f(a)}{x - a} \geq 0$$
et donc par passage à la limite $f'_g(a) \geq 0$;

— si $x > a$,
$$\frac{f(x) - f(a)}{x - a} \leq 0$$
et donc par passage à la limite $f'_d(a) \leq 0$.

Or f est dérivable en a, donc $f'_g(a) = f'_d(a) = f'(a)$. On en déduit que $f'(a) = 0$. □

✌ **Remarque XII.7.** Dans le cas $\mathscr{C}^1$, on peut démontrer ce résultat en appliquant le TVI (théorème XI.16) à f'.

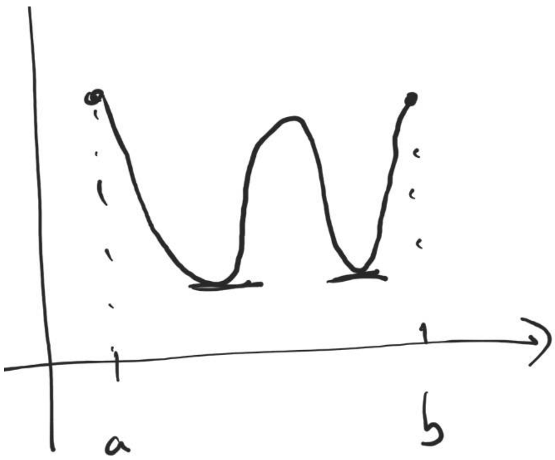
b) Théorème de Rolle

L'heuristique du résultat suivant se trouve dans les écrits de Michel Rolle (français, 1652—1719). Sa forme moderne, liée au théorème des accroissements finis, doit beaucoup aux écrits d'Augustin Louis Cauchy (français,1789—1857) et Joseph–Louis Lagrange (italien naturalisé français, 1736—1813). La démonstration proposée ici s'inspire de celle proposée par Joseph–Alfred Serret (français, 1819—1885).

Théorème XII.10 (Rolle)**.**
Soient $a, b \in \mathbb{R}$ tels que $a < b$ et soit $f : [a, b] \to \mathbb{R}$ telle que :
— $f \in \mathscr{C}^0([a, b])$;
— $f \in \mathscr{D}^1(]a, b[)$;
— $f(a) = f(b)$.
Alors :

$$\exists\, c \in]a, b[,\ f'(c) = 0\ .$$

✌ **Remarque XII.8.** L'idée derrière ce théorème n'a rien d'ésotérique (rétrospectivement) : si $f(a) = f(b)$, alors la courbe représentative de la fonction f admettra au moins une tangente horizontale entre a et b.



Démonstration. Si f est constante, c'est trivial. Sinon, il existe (par exemple) $y \in]a, b[$ tel que $f(y) < f(a)$ et donc f admet un minimum local en un point (non déterminé) $c \in]a, b[$ par le corollaire XI.16.c. Ainsi, $f'(c) = 0$ par proposition XII.9. □

➠ **Exemple XII.8.**
— La dérivée d'un polynôme admet une racine entre chaque paire de racines de ce dernier.
— En physique, un mobile unidimensionnel revenant à son point de départ devra annuler sa vitesse à un instant donné.

c) Théorème des accroissements finis

Le théorème qui suit, point central de ce chapitre, est du à Augustin Louis Cauchy (français,1789—1857) et Joseph–Louis Lagrange (italien naturalisé français, 1736—1813).

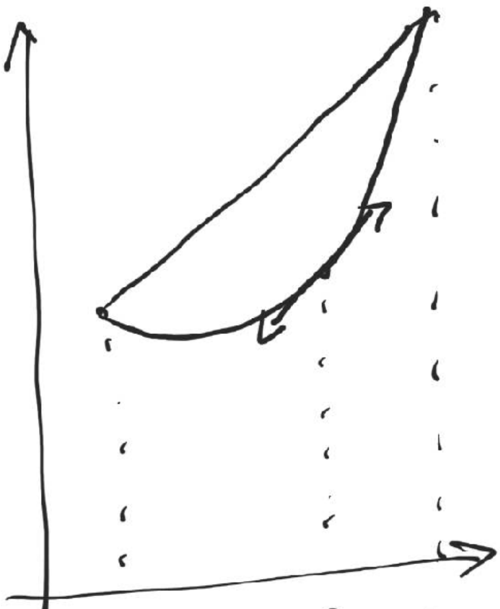
Théorème XII.11 (Accroissements finis)**.**
Soient $a, b \in \mathbb{R}$ tels que $a < b$ et soit $f : [a, b] \to \mathbb{R}$ telle que :
- — $f \in \mathscr{C}^0([a, b])$;
- — $f \in \mathscr{D}^1(]a, b[)$.

Alors :

$$\exists\, c \in]a, b[,\ f'(c) = \frac{f(b) - f(a)}{b - a}\ .$$

✌ **Remarque XII.9.**

- — Géométriquement, ce résultat se traduit de la façon suivante : il existe un point c en lequel la tangente à la courbe représentative de f est parallèle à la "corde" reliant les points $(a, f(a))$ et $(b, f(b))$.



a c b

- — En physique, ce résultat implique qu'il existe un instant dans tout mouvement en lequel vitesses instantanée et moyenne sont confondues.

Démonstration. Il suffit d'applique le théorème de Rolle à la fonction

$$\begin{aligned} g : [a, b] &\to \mathbb{R} \\ x &\mapsto f(x) - \frac{f(b) - f(a)}{b - a}(x - a)\ . \end{aligned}$$

□

Corollaire XII.11.a (Inégalité des accroissements finis)**.** Soient $a, b \in \mathbb{R}$ tels que $a < b$ et soit $f : [a, b] \to \mathbb{R}$ telle que :
- $f \in \mathscr{C}^0([a, b])$;
- $f \in \mathscr{D}^1(]a, b[)$;
- il existe $m, M \in \mathbb{R}$ tels que :

$$\forall x \in]a, b[,\ m \leq f'(x) \leq M\ .$$

Alors :

$$m \leq \frac{f(b) - f(a)}{b - a} \leq M\ .$$

◇ Une application : fonctions lipschitziennes

La classe de fonctions définies ci–ensuite est nommée en l'honneur du mathématicien allemand Rudolph Otto Sigismund Lipschitz (1832—1903).

Définition XII.7. Soit $M \in \mathbb{R}_+^*$. Une fonction $f : I \to \mathbb{R}$ est dite M–**lipschitzienne** si :

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

➠ **Exemple XII.9.** La fonction $x \mapsto x^2$ est 2–lipschitzienne sur $[0, 1]$. En effet, pour tous $x, y \in [0, 1]$:

$$|x^2 - y^2| = |(x + y)(x - y| \leq 2|x - y|\ .$$

Proposition XII.12. Toute application lipschitzienne est continue.

Démonstration. Soit $f : I \to \mathbb{R}$ une fonction M–lipschitzienne et soit $a \in I$. Alors, pour tout $x \in I$:

$$\begin{aligned} |f(x) - f(a)| &\leq M|x - a| \\ &\xrightarrow[x \to a]{} 0 \end{aligned}$$

d'où le résultat. □

✖ **Attention :** la réciproque est fausse. La fonction exponentielle n'est pas exemple pas lipschitzienne sur $\mathbb{R}$.

Proposition XII.13. Toute fonction dérivable dont la dérivée est bornée est lipschitzienne.

Démonstration. Il s'agit d'un corollaire de l'inégalité des accroissements finis. □

➠ **Exemple XII.10.** Les fonctions continues sur un segment, les fonctions sinus, cosinus sont lipschitziennes.

$\diamond$ **Une autre application : suites récurrentes et applications contractantes**

On se fixe une application $f : I \to I$ **contactante**, *i.e* lipschitzienne de rapport $K \in]0,1[$. On fixe $u_0 \in \mathring{I}$ et on pose, pour $n \geq 0$:

$$u_{n+1} = f(u_n) \, .$$

Si on suppose que f admet un point fixe c, alors pour tout $n \geq 0$:

$$\begin{aligned} |u_{n+1} - c| &= |f(u_n) - f(c)| \\ &\leq K|u_n - c| \\ &\vdots \\ &\leq K^n|u_1 - c| \\ &\to 0 \end{aligned}$$

et donc $u_n \to c$.

✌ **Remarque XII.10.** Il est possible de démontrer l'existence de ce point fixe dans le cas général.

✍ **Exercice XII.2.** Soit $\alpha \in \mathbb{R}_+$ et soit $(u_n)_n$ la suite définie par récurrence via :

$$\begin{cases} u_0 = \alpha \\ \forall n \in \mathbb{N}, \; u_{n+1} = \sqrt{1 + u_n} \end{cases} .$$

Démontrer que $(u_n)_n$ converge vers $\dfrac{1+\sqrt{5}}{2}$.

➥ **Correction :** *Nous avons déjà remarqué dans le chapitre IX que l'unicité de la limite (proposition IX.2) entraîne que $\ell = \sqrt{1+\ell}$ et donc $\ell = \dfrac{1+\sqrt{5}}{2}$ (ℓ doit être positive car les termes de la suite u le sont et $\ell^2 = 1 + \ell$). Par inégalité des accroissements finis, la fonction $f : x \mapsto \sqrt{1+x}$ est $\frac{1}{2}$–lispschitzienne sur $\mathbb{R}_+$. Elle admet de plus un point fixe (nous venons de le trouver !), d'où le résultat en appliquant l'heuristique vue supra.*

d) Monotonie

Théorème XII.14.
Soit $f \in \mathscr{C}^0(I) \cap \mathscr{D}^1(\mathring{I})$. Alors :

(i) f est croissante sur $I \iff f' \geq 0$ sur $\mathring{I}$;

(ii) f est décroissante sur $I \iff f' \leq 0$ sur $\mathring{I}$;

(iii) f est constante sur $I \iff f' = 0$ sur $\mathring{I}$.

✖ **Attention :** I est un **intervalle**. Considérer la fonction $x \mapsto |x|$ sur $\mathbb{R}^*$ qui est de dérivée seconde nulle sans que sa dérivée première ne sont constante.

Démonstration. (i)

($\Rightarrow$) Soient $a \in \mathring{I}$ et $x \in I$ tels que $x > a$. Alors :

$$\frac{f(x) - f(a)}{x - a} \geq 0$$

et donc, en passant à la limite $f'(a) = f'_d(a) \geq 0$.

($\Leftarrow$) Soient $x, y \in I$ tels que $x \geq y$. Alors, par TAF (théorème XII.11, applicable car $f \in \mathscr{C}^0([y,x]) \cap \mathscr{D}^1(]y,x[)$), il existe $c \in]y,x[$ tel que :

$$f(x) - f(y) = f'(c)(x - y) \geq 0$$

d'où le résultat.

(ii) Se traite de façon analogue.

(iii) Découle immédiatement de (i) et (ii).

□

Corollaire XII.14.a. Soit $f \in \mathscr{C}^0(I) \cap \mathscr{D}^1(\mathring{I})$ une fonction **croissante**. On pose :

$$\mathscr{E} = \{x \in \mathring{I} \mid f'(x) = 0\} .$$

Alors :

f est strictement croissante

$\Longleftrightarrow$

$\mathscr{E}$ non contient aucun intervalle d'intérieur non vide.

Démonstration. Cela découle du point (iii) de la proposition précédente. □

✌ **Remarque XII.11.**

— Cela signifie que tout va bien si $\mathscr{E}$ ne contient aucun convexe non trivial, *i.e* si il est "rempli de trous".

— Un résultat analogue existe évidemment pour les fonctions décroissantes.

e) Passages à la limite

La question qui nous intéresse dans ce paragraphe est la suivante : étant donné un point $a \in I$ et $f \in \mathscr{C}^0(I) \cap \mathscr{D}^1(I \setminus \{a\})$, comment vérifier que f est dérivable en a ? Cela nous sera en particulier utile pour tester la régularité de prolongements par continuité.

Proposition XII.15. Soit $a \in I$ et soit $f \in \mathscr{C}^0(I) \cap \mathscr{D}^1(I \setminus \{a\})$. Alors :

— si $f'(x) \xrightarrow[x \to a]{} \ell \in \mathbb{R}$, f est dérivable en a et $f'(a) = \ell$;

— si $f'(x) \xrightarrow[x \to a]{} \pm\infty$, f n'est pas dérivable en a et sa courbe y admet une tangente verticale ;

— sinon, on ne peut conclure.

Démonstration. Soit $x \in I \setminus \{a\}$, par exemple $x > a$. En appliquant le TAF au segment $[a, x]$, on obtient l'existence de $c_x \in]a, x[$ tel que

$$f'(c_x) = \frac{f(x) - f(a)}{x - a}$$

et $c_x \xrightarrow[x \to a]{} a$ par encadrement. On en déduit le résultat en examinant les différentes limites possibles pour $f'(c_x)$ quand $x \to a$. □

➠ **Exemple XII.11.** Le prolongement par continuité en 0 de la fonction $x \mapsto e^{-\frac{1}{x}}$ (initialement définie sur $\mathbb{R}_+^*$, *cf.* chapitre XI) est dérivable en 0 de dérivée nulle car

$$\frac{1}{x^2} e^{-\frac{1}{x}} \xrightarrow[x \to 0^+]{} 0 \, .$$

Proposition XII.16. Soient $a \in I$, $k \in \mathbb{N}$ et $f \in \mathscr{C}^k(I \setminus \{a\})$. Alors, si pour tout $i \leq k$, $f^{(i)}$ admet une limite finie en a, la fonction f est est prolongeable de façon $\mathscr{C}^k$ à I.

3.‾ Brêve extension aux fonctions à valeurs complexes

Il est aisé de vérifier qu'une fonction $f : I \to \mathbb{C}$ est dérivable si et seulement si ses parties réelle et imaginaire le sont. On en déduit par exemple que la fonction $x \mapsto e^{\alpha x}$, pour $\alpha \in \mathbb{C}$ est dérivable, de dérivée $x \mapsto \alpha e^{\alpha x}$.

✖ **ATTENTION :** le théorème de Rolle (XII.10) et le théorème des accroissements finis (XII.11) sont faux sur $\mathbb{C}$. On a cependant une version complexe de l'inégalité des accroissements finis (corollaire XII.11.a), dont nous admettrons la démonstration.

Théorème XII.17 ([Inégalité des accroissements finis, cas complexe)**.**]
Soit $f : [a, b] \to \mathbb{C}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que :

$$\exists\, K \in \mathbb{R}_+,\ \forall x \in]a, b[,\ |f'(x)| \leq K \, .$$

Alors :

$$|f(b) - f(a)| \leq K|b - a| \, .$$

Chapitre XIII

Entiers relatifs, arithmétique

1.⁻ Divisibilité

On rappelle que $(\mathbb{Z}, +, \times)$ est un anneau commutatif intègre (*i.e* dans lequel on peut simplifier par des éléments non nuls) sur lequel l'ordre naturel "$\leq^q$" est total.

a) Diviseurs, multiples

Définition XIII.1. Soient $a, b \in \mathbb{Z}$. On dit que a **divise** b si il existe $c \in \mathbb{Z}$ tel que $b = ac$. On dit alors que b est un **multiple** de a.

Notation. On notera $a|b$ si a divise b et on pose $a\mathbb{Z} = \{ak \mid k \in \mathbb{Z}\}$ l'ensemble des multiples de a.

✌ **Remarque XIII.1.**

— Si a divise b et que a et b ne sont pas tous les deux nuls, l'entier relatif c tel que $b = ac$ est unique par intégrité ; on l'appelle **quotient** de b par a et on le note, **si** $a \neq 0$, $\frac{b}{a}$. Cette notation est à utiliser avec parcimonie et à réserver au cas où $a|b$.

— On vérifie aisément que $a|b \iff b\mathbb{Z} \subset a\mathbb{Z}$.

— Pour tout $a \in \mathbb{Z}$, $0 = a \times 0$ et donc $a|0$: **tout le monde divise zéro**. À l'inverse, 0 ne divise que lui–même.

Proposition XIII.1. La relation "divise" est réflexive et transitive. De plus, si $a, b \in \mathbb{Z}$ alors :

$$(a|b) \wedge (b|a) \Leftrightarrow a = \pm b\ .$$

✌ **Remarque XIII.2.** Ceci implique que $a\mathbb{Z} = b\mathbb{Z} \Leftrightarrow a = \pm b$.

Démonstration. La réflexivité et la transitivité sont immédiates (*cf.* chapitre V). Il ensuite clair que si $a = \pm b$, alors $a = (\pm 1) \times b$ et donc $a|b$ et $b|a$. Réciproquement, si a et b se divisent mutuellement, il existe c et d dans $\mathbb{Z}$ tels que $b = ac$ et $a = bd$. On a alors $b = bdc$, ce qui entraîne par intégrité que $dc = 1$ et donc $d = c = \pm 1$. □

b) Division euclidienne

Théorème XIII.2.
Soient $a, b \in \mathbb{Z}$ tels que $b \neq 0$. Alors :

$$\exists!\,(q,r) \in \mathbb{Z}^2,\ \begin{cases} a = bq + r \\ 0 \leq r < |b| \end{cases}.$$

Vocabulaire. a est appelé **dividende**, b **diviseur**, q **quotient** et r **reste** de la division euclidienne.

Démonstration. Il suffit pour l'existence de poser $q = \left\lfloor \frac{a}{b} \right\rfloor$ et $r = a - bq$ lorsque $a, b \geq 0$. Dans le cas où a et b sont potentiellement négatifs, effectuer la division euclidienne de $|a|$ par $|b|$ et magouiller pour retomber sur le bon signe.

Pour l'unicité, si on suppose qu'il existe deux tels couples (q, r) et (q', r') alors $b(q - q') = r' - r$ et donc $b | r' - r$. Or $r, r' \in [\![0, |b|[\![$, *ergo* $r' - r \in]\!] - |b|, |b|[\![$, ce qui entraîne que $r' - r = 0$. Par conséquent, $q - q' = 0$, d'où le résultat. □

➠ **Exemple XIII.1.** $15 = 7 \times 2 + 1$ est une division euclidienne, mais pas $15 = 7 \times 3 - 6$.

Proposition XIII.3. [Sous–groupes de $\mathbb{Z}$] Soit G un sous–groupe de $(\mathbb{Z}, +)$. Alors il existe un unique $n \in \mathbb{N}$ tel que $G = n\mathbb{Z}$.

Démonstration.

Existence : Considérons l'ensemble $\mathscr{E} = G \cap \mathbb{N}^*$. S'il est vide, alors $G = \{0\} = 0\mathbb{Z}$; dans le cas contraire, il admet par axiome D un plus petit élement n_0. Un sous–groupe étant stable par puissances (ici multiples), on a alors $n_0\mathbb{Z} \subset G$.

Pour montrer l'inclusion réciproque, prenons $a \in G$ et effectuons sa division euclidienne par n_0 (qui est non nul) : par le théorème XIII.2, il existe un unique couple $(q, r) \in \mathbb{Z}^2$ tel que :

$$\begin{cases} a = n_0 q + r \\ 0 \leq r < |n_0| = n_0 \end{cases}.$$

Le reste vérifie $r = a - n_0 q$ et donc appartient à G comme différence de deux éléments du sous–groupe. Il ne peut donc pas être dans $\mathscr{E}$ et strictement inférieur à n_0, *ergo* $r = 0$ d'où $a \in n_0\mathbb{Z}$.

Unicité : S'il existe deux tels entiers naturels n, n' alors $G = n\mathbb{Z} = n'\mathbb{Z}$ et donc $n = n'$ car $n, n' \geq 0$.

□

2.⁻ PGCD, algorithme d'Euclide

a) Plus Grand Commun Diviseur

Proposition/définition XIII.2. Soient $a, b \in \mathbb{N}$ tels que $b \neq 0$. On appelle **plus grand commun diviseur** (PGCD) de a et b la quantité :

$$\max\{\delta \in \mathbb{N}^* \mid (\delta|a) \wedge (\delta|b)\} .$$

Notation. $\text{pgcd}(a, b)$, $a \wedge b$.

Démonstration. L'ensemble $\mathscr{E} = \{\delta \in \mathbb{N}^* \mid (\delta|a) \wedge (\delta|b)\}$ est une partie de $\mathbb{N}$ non vide (elle contient 1) et majorée par $\min(a, b)$ si $a \neq 0$, par b sinon, donc le PGCD est bien défini. □

Définition XIII.3. Soient $a, b \in \mathbb{Z}$ tels que $b \neq 0$. On appelle **PGCD** de a et b la quantité $|a| \wedge |b|$.

➠ **Exemple XIII.2.**
— $2 \wedge 3 = 1$;
— $2 \wedge (-3) = 1$;
— $(-2) \wedge (-4) = 2$.

✖ **ATTENTION :** le PGCD sera toujours, selon notre définition, un **entier naturel non nul**. D'autres conventions existent dans la littérature ; nous invitons le lecteur à faire preuve de vigilance.

b) Algorithme d'Euclide

Pour des raisons d'efficacité heuristique, nous utiliserons dans ce paragraphe la notation, pour $a \in \mathbb{Z}$:

$$\mathscr{D}(a) = \{k \in \mathbb{Z} \mid k|a\}$$

pour l'ensemble des diviseurs de a. Cet ensemble vérifie clairement que $\mathscr{D}(a) = \mathscr{D}(|a|)$ et, par définition :

$$a \wedge b = \max(\mathscr{D}(a) \cap \mathscr{D}(b)) \cap \mathbb{N}^* .$$

Remarquons que si $a, b \in \mathbb{Z}$ sont tels que $b \neq 0$ ont pour division euclidienne

$$a = bq + r$$

avec q, r vérifiant les conditions du théorème XIII.2, alors on peut vérifier rapidement (il s'agit d'un bon exercice pour le lecteur avisé) que :

$$\mathscr{D}(a) \cap \mathscr{D}(b) = \mathscr{D}(b) \cap \mathscr{D}(r)$$

i.e pour tout $k \in \mathbb{Z}$:

$$(k|a) \wedge (k|b) \Leftrightarrow (k|b) \wedge (k|r) .$$

En particulier, cela implique que $a \wedge b = b \wedge r$.

Proposition XIII.4. Soient $a, b \in \mathbb{Z}$ tels que $b \neq 0$ et soit $\delta \in \mathbb{N}^*$. Alors :

$$\delta = a \wedge b$$
$$\Longleftrightarrow$$
$$\begin{cases} (\delta|a) \wedge (\delta|b) \\ \forall d \in \mathbb{Z},\ (d|a) \wedge (d|b) \Rightarrow (d|\delta) \end{cases} .$$

✌ **Remarque XIII.3.** Cela signifie que le PGCD est en fait le "maximum", au sens de la relation "|" (qui n'est pas un ordre) des diviseurs de a et b.

Démonstration. Remarquons que cet énoncé est équivalent au suivant :

$$\delta = a \wedge b \Leftrightarrow \mathscr{D}(a) \cap \mathscr{D}(b) = \mathscr{D}(\delta)$$

et démontrons celui–ci.

($\Leftarrow$) Si d est un diviseur strictement positif de a et b alors $d|\delta$ et donc $d|\delta$. Comme δ divise a et b, on a bien $\delta = \max(\mathscr{D}(a) \cap \mathscr{D}(b)) \cap \mathbb{N}^* = a \wedge b$.

($\Rightarrow$) Démontrons par récurrence forte sur $|b| \in \mathbb{N}^*$ que $\mathscr{D}(a \wedge b) = \mathscr{D}(a) \cap \mathscr{D}(b)$.
— Si $|b| = 1$, alors

$$\mathscr{D}(a) \cap \mathscr{D}(b) = \mathscr{D}(a) \cap \mathscr{D}(1) = \mathscr{D}(a) = \mathscr{D}(a \wedge 1) .$$

— Si la propriété est vraie pour tout $k \leq |b| - 1$, alors, en posant $a = bq + r$ la division euclidienne de a par b, on a :

$$\begin{aligned} \mathscr{D}(a) \cap \mathscr{D}(b) &= \mathscr{D}(b) \cap \mathscr{D}(r) \\ &= \mathscr{D}(b \wedge r) \text{ par hypohtèse de récurrence} \\ &= \mathscr{D}(a \wedge b) \end{aligned}$$

d'où le résultat.

□

◇ Algorithme d'Euclide

Soient $a, b \in \mathbb{N}$ tels que $b \neq 0$; pour déterminer le PGCD de a et b on définit une suite récurrente $(r_n)_n$ selon le procédé suivant :
— $r_0 = a$;
— $r_1 = b$;
— pour $n \geq 0$, r_{n+2} sera le reste de la division euclidienne de r_n par r_{n+1} si ce dernier est non nul ; dans le cas contraire on pose $r_{n+2} = 0$.

Proposition XIII.5. La suite $(r_n)_n$ est décroissante et stationnaire à 0.

Démonstration. La décroissance est claire. Pour le côté stationnaire, commençons par remarquer que si il existe $N \geq 0$ tel que $r_N = 0$ alors $\forall n \geq N$, $r_n = 0$ par construction ; il nous suffit donc de démontrer l'existence d'un tel rang N.

Procédons par l'absurde en supposant que la suite $(r_n)_n$ ne s'annule jamais ; alors on a pour tout $n \geq 0$, $r_n < r_{n+1}$ par division euclidienne ; nous sommes donc en présence d'une suite décroissante d'entiers naturels ; d'après la proposition IX.1 cette dernière est stationnaire égale à $c \in \mathbb{N}$. Or, si $r_n = r_{n+1} = c \neq 0$ pour $n \geq 0$, $r_{n+2} = 0$, ce qui contredit notre hypothèse. □

Proposition XIII.6 (Euclide)**.** Posons :

$$N_0 = \min\{N \in \mathbb{N} \,|\, r_N = 0\} \;.$$

Alors :

$$a \wedge b = r_{N_0-1} \;.$$

Démonstration. Le minimum de l'énoncé existe par axiome D. D'après nos travaux préliminaires :

$$\begin{aligned}
\mathscr{D}(a) \cap \mathscr{D}(b) &= \mathscr{D}(r_0) \cap \mathscr{D}(r_1) \\
&= \mathscr{D}(r_1) \cap \mathscr{D}(r_2) \\
&\;\vdots \\
&= \mathscr{D}(r_{N_0-1}) \cap \underbrace{\mathscr{D}(0)}_{=\mathbb{Z}} \\
&= \mathscr{D}(r_{N_0-1})
\end{aligned}$$

et donc $r_{N_0-1} = a \wedge b$. □

Cet algorithme peut se retranscrire sans trop de difficulté en langage `python`.

```python
def euclide(a,b):
    s=a
    t=b
    while t!=0:
        s,t=t,s%t
    return s
```

➠ **Exemple XIII.3.** Pour 137 et 12, on passe par les étapes suivantes :
— $137 = 12 \times 11 + 5$;
— $12 = 5 \times 2 + 2$;
— $5 = 2 \times 2 + 1$;
— $2 = 1 \times 2 + 0$.
Ainsi $137 \wedge 12 = 1$.

✌ **Remarque XIII.4.** On peut étendre cet algorithme aux entiers relatifs via la relation $a \wedge b = |a| \wedge |b|$.

c) Relation de Bézout

Proposition XIII.7. Soient $a, b \in \mathbb{Z}$ tels que $b \neq 0$. Alors :

$$\exists\, u, v \in \mathbb{Z},\ au + bv = a \wedge b\ .$$

✖ **Attention :** : cette proposition n'est pas une équivalence : 4 n'est pas le PGCD de 2 et 3 et pourtant $2 \times 2 + 3 \times 0 = 4$.

Démonstration. La démonstration de ce résultat repose sur l'**algorithme d'Euclide étendu**. Conservons la suite $(r_n)_n$ définie au paragraphe précédent et posons :
— $u_0 = 1$, $u_1 = 0$;
— $v_0 = 0$, $v_1 = 1$.
Démontrons par récurrence double sur $n \in \mathbb{N}$ que :

$$\forall n \in \mathbb{N},\ \exists\, u_n, v_n \in \mathbb{Z},\ au_n + bv_n = r_n\ .$$

— Pour $n = 0, 1$ c'est immédiat.
— Supposons la propriété vérifiée aux rangs n et $n+1$ pour un certain $n \geq 0$. Alors, par définition de la suite $(r_n)_n$ il existe $q \in \mathbb{Z}$ tel que :

$$r_n = r_{n+1}q + r_{n+2}$$

et donc :

$$\begin{aligned} r_{n+2} &= r_n - r_{n+1}q \\ &= au_n + b_n - q(au_{n+1} + bv_{n+1}) \text{ par hypothèse de récurrence} \\ &= a(u_n - qu_{n+1}) + b(v_n - qv_{n+1})\ . \end{aligned}$$

Ne reste qu'à poser $u_{n+2} = u_n - qu_{n+1}$ et $v_{n+2} = v_n - qv_{n+1}$.
On conclut la démonstration en appliquant ce résultat à

$$n = \min\{N \in \mathbb{N} \,|\, r_N = 0\} - 1\ .$$

□

➠ **Exemple XIII.4.** Pour 33 et 21, l'algorithme d'Euclide livre les étapes suivantes :
— $33 = 21 \times 1 + 12$;
— $21 = 12 \times 1 + 9$;
— $12 = 9 \times 1 + 3$;
— $9 = 3 \times 3 + 0$.
En "remontant" ces opérations, on obtient :

$$\begin{aligned} 33 \wedge 21 = 3 &= 12 - 9 \times 1 \\ &= (33 - 21) - (21 - 12) \\ &= (33 - 21) - (21 - (33 - 21)) \\ &= 33 \times 2 + (-3) \times 21\ . \end{aligned}$$

d) Plus Petit Commun Multiple

Proposition/définition XIII.4. Soient $a, b \in \mathbb{Z} \setminus \{0\}$; on appelle **plus petit commun multiple** (PPCM) de a et b la quantité :

$$\min(a\mathbb{Z} \cap b\mathbb{Z}) \cap \mathbb{N}^* .$$

Démonstration. Le minimum existe par axiome D. □

Notation. $\mathrm{ppcm}(a, b), a \vee b$.

✌ **Remarque XIII.5.** Tout comme pour le PGCD, on dispose d'une caractérisation du PPCM ; si $\mu \in \mathbb{N}$ alors :

$$\mu = a \vee b$$
$$\Longleftrightarrow$$
$$\begin{cases} \mu \in a\mathbb{Z} \cap b\mathbb{Z} \\ \forall m \in \mathbb{Z},\ (m \in a\mathbb{Z} \cap b\mathbb{Z}) \Rightarrow (m \in \mu\mathbb{Z}) \end{cases}$$
$$\Longleftrightarrow$$
$$\begin{cases} \mu \geq 0 \\ \mu\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z} \end{cases} .$$

Proposition XIII.8. Soient $a, b \in \mathbb{Z}$. Alors :

$$|ab| = (a \vee b)(a \wedge b) .$$

Nous démontrerons ce résultat ultérieurement ; nous le mentionnons cependant à ce stade du cours pour la raison qu'il s'agit de la meilleure façon de déterminer un PPCM en pratique.

3.⁻ Entiers premiers entre eux

a) C'est quoi ?

Définition XIII.5. On dit que deux entiers entiers a et b sont **premiers entre eux** si $a \wedge b = 1$.

➠ **Exemple XIII.5.** 3 et 7 sont premiers entre eux.

Proposition XIII.9. Soient $a, b \in \mathbb{Z}$ tels que $b \neq 0$; alors $\dfrac{a}{a \wedge b}$ et $\dfrac{b}{a \wedge b}$ sont premiers entre eux.

Démonstration. Immédiat via la caractérisation du PGCD. □

b) Théorème de Bézout

Le théorème qui suit est du à Claude–Gaspard Bachet de Méziriac (français, 1581—1638). Étienne Bézout (français, 1730—1783) a généralisé cet énoncé à des structures plus générales, notamment les anneaux de polynômes (*cf.* chapitre XV).

Théorème XIII.10 (Bézout)**.**
Soient $a, b \in \mathbb{Z}$ tels que $b \neq 0$. Alors :

$$a \text{ et } b \text{ sont premiers entre eux}$$

$$\Leftrightarrow$$

$$\exists\, u, v \in \mathbb{Z},\ au + bv = 1\ .$$

Démonstration. ($\Downarrow$) Il s'agit d'un cas particulier de la proposition XIII.7.

($\Uparrow$) Soit $d \in \mathscr{D}(a) \cap \mathscr{D}(b)$. Alors, comme il existe $u, v \in \mathbb{Z}$ tel que $au + bv = 1$ on a $d|1$. On en déduit que $a \wedge b = 1$.

□

Corollaire XIII.10.a. Soient $a, b_1, \dots, b_n \in \mathbb{Z} \setminus \{0\}$. Alors :

$$\forall i \in [\![1, n]\!],\ a \wedge b_i = 1$$

$$\Leftrightarrow$$

$$a \wedge \left(\prod_{i=1}^{n} b_i\right) = 1\ .$$

Démonstration. Démontrons le par récurrence sur $n \in \mathbb{N}^*$; plus précisément, démontrons que, pour $a \neq 0$ fixé on a, pour tout n : "pour toute famille $b_1, \dots, b_n$ d'entiers relatifs non nuls, on a

$$\forall i \in [\![1, n]\!],\ a \wedge b_i = 1$$

$$\Leftrightarrow$$

$$a \wedge \left(\prod_{i=1}^{n} b_i\right) = 1\ ."$$

- $n = 1$: immédiat.
- Supposons la propriété vérifiée pour toute famille de n entiers $b_1, \dots b_n$. Alors, si $b_1, \dots, b_{n+1}$ est une famille de $n+1$ entiers relatifs non nuls :

$$\begin{aligned}\forall i \in [\![1, n+1]\!],\ a \wedge b_i = 1 &\Leftrightarrow (\forall i \in [\![1, n]\!],\ a \wedge b_i = 1) \wedge (a \wedge b_{n+1} = 1)\\ &\Leftrightarrow \left(a \wedge \left(\prod_{i=1}^{n} b_i\right) = 1\right) \wedge (a \wedge b_{n+1} = 1)\end{aligned}$$

D'après le théorème de Bézout, cette dernière propriété est équivalente à l'existence de $u, v, w, x \in \mathbb{Z}$ tels que

$$\begin{aligned} 1 &= au + \left(\prod_{i=1}^{n} b_i\right) v \\ &= aw + b_{n+1}x \end{aligned}$$

et alors

$$1 = Ua + V\left(\prod_{i=1}^{n+1} b_i\right)$$

avec

$$U = uw + vxb_{n+1} + wv\left(\prod_{i=1}^{n} b_i\right)$$

et $V = vx$, d'où le résultat.

□

c) Lemme de Gauss

Le résultat qui suit est une généralisation, nommée en l'honneur de Johann Carl Friedrich Gauss (allemand, 1777—1855) d'un lemme apparaissant dans le livre VII des *Éléments* d'Euclide. Par ailleurs, cet énoncé apparaît sous sa forme moderne dans un traité de Jean Prestet (mathématicien et prêtre français, 1648—1690).

Théorème XIII.11 (Lemme de Gauss)**.**
Soient $a, b, c \in \mathbb{Z}$ tels que :
— a et b sont premiers entre eux ;
— $a|bc$.
Alors $a|c$.

Démonstration. D'après le théorème de Bézout (XIII.10), il existe $u, v \in \mathbb{Z}$ tels que $au + bv = 1$. De plus, comme $a|bc$, il existe $w \in \mathbb{Z}$ tel que $bc = aw$. Alors :

$$\begin{aligned} c &= c \times 1 \\ &= uac + bvc \\ &= uav + awv \end{aligned}$$

d'où le résultat. □

Corollaire XIII.11.a. Si a et b sont deux entiers premiers entre eux, alors $a \vee b = |ab|$.

Démonstration. Soit $m \in a\mathbb{Z} \cap b\mathbb{Z}$; alors il existe $g, h \in \mathbb{Z}$ tel que $m = ah = bg$ et donc $b|ah$ ce qui entraîne par lemme de Gauss que $b|h$ et donc $ab|m$. □

Ce corollaire nous permet de démontrer la proposition XIII.8 ; en effet, on a alors (avec les hypothèse de l'énoncé de cette proposition) :

$$\frac{a}{a \wedge b} \vee \frac{b}{a \wedge b} = \frac{a \vee b}{a \wedge b} = |ab|$$

car $\frac{a}{a \wedge b}$ et $\frac{b}{a \wedge b}$ sont premiers entre eux.*In fine*, on a bien :

$$(a \wedge b)(a \vee b) = |ab| \ .$$

Corollaire XIII.11.b. Soient $a_1, \ldots, a_n \in \mathbb{Z}$ deux à deux premiers entre eux et $b \in \mathbb{Z}$. Alors :

$$\forall i \in [\![1, n]\!],\ a_i | b$$

$$\Leftrightarrow$$

$$\prod_{i=1}^{n} a_i | b \ .$$

Démonstration. Par récurrence... □

◇ **Application : équations diophantiennes**

Considérons dans ce paragraphe une équation de la forme suivante, pour $a, b, c \in \mathbb{Z}$ fixés tels que $ab \neq 0$ et d'inconnues $x, y \in \mathbb{Z}$:

$$ax + by = c \ . \qquad (\mathbf{E}\text{:XIII.1})$$

Commençons par poser $\delta = a \wedge b$ et par remarquer que si $\delta \not| c$, alors l'équation ne possède aucune solution. Dans le cas contraire, celle–ci est équivalente à la forme réduite suivante :

$$a'x + b'x = c' \qquad (\mathbf{E}\text{:XIII.2})$$

avec $a' = \frac{a}{\delta}$, $b' = \frac{b}{\delta}$ et $c' = \frac{c}{\delta}$.

Les entiers a' et b' étant premiers entre eux, l'algorithme d'Euclide étendu nous permet de trouver $u, v \in \mathbb{Z}$ tels que $a'u + b'v = 1$; une solution particulière de (**E** :XIII.2) est alors le couple

$$(x_0, y_0) = (uc, vc)$$

et

Supposons que nous disposions d'une autre solution $(x, y) \in \mathbb{Z}^2$ de (**E** :XIII.2) ; on obtient alors par soustraction des égalités $a'x + b'y = c'$ et $a'x_0 + b'y_0 = c'$ que :

$$a'(x - x_0) = b'(y_0 - y)$$

ce qui implique que $a'|b'(y_0 - y)$. Or, $a' \wedge b' = 1$ donc, par lemme de Gauss :

$$a'|y_0 - y$$

et donc il existe $k \in \mathbb{Z}$ tel que :

$$y = y_0 + ka' \ .$$

Revenant à l'égalité précédente on a désormais :

$$a'(x - x_0) = ka'b'$$

et donc, par intégrité :

$$x = x_0 - kb' \ .$$

En conclusion les solutions de (**E** :XIII.2) sont les éléments de l'ensemble :

$$\mathscr{S}_{\mathrm{red}} = \{(x_0 - kb', y_0 + ka') \mid k \in \mathbb{Z}\}$$

et donc les solutions de (**E** :XIII.1) parcourent cet ensemble, les deux équations étant équivalentes.

➠ **Exemple XIII.6.** L'équation $12x + 4y = 8$ est équivalente à $3x + y = 2$. On trouve comme solution particulière le couple $(1, -1)$ et donc les solutions sont les $(1 + k, -1 + 3k)$, pour $k \in \mathbb{Z}$.

◇ **Application : forme irréductible d'une fraction rationnelle**

Soit $x \in \mathbb{Q}$; alors il existe, par définition, $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ tels que $x = \dfrac{p}{q}$. Posons $\delta = p \wedge q$, $p' = \dfrac{p}{\delta}$ et $q' = \dfrac{q}{\delta}$; il est alors trivial que :

$$\frac{p'}{q'} = \frac{p}{q} = x.$$

Nous venons de trouver une écriture de x comme quotient de deux entiers premiers entre eux, le dénominateur étant strictement positif. Une telle écriture s'appelle **forme irréductible de** x et est unique. En effet si $x = \dfrac{a}{b}$ avec $(a, b) \in \mathbb{Z} \times \mathbb{N}^*$ tels que $a \wedge b = 1$ alors :

$$p'b = aq'$$

et donc $q'|p'b$, ce qui entraîne par lemme de Gauss que $q'|b$. On montre symétriquement que $b|q'$ et donc, par positivité, $q' = b$ *ergo* $p' = a$.

d) Généralisations diverses

Il est possible de définir le PGCD et le PPCM d'une famille d'entiers $a_1, \ldots a_n$ via les formules récursives :

$$\mathrm{pgcd}(a_1, \ldots, a_n) = \mathrm{pgcd}(\mathrm{pgcd}(a_1, \ldots, a_{n-1}), a_n)$$

et

$$\mathrm{ppcm}(a_1, \ldots, a_n) = \mathrm{ppcm}(\mathrm{ppcm}(a_1, \ldots, a_{n-1}), a_n) \ .$$

On a alors un analogue de la relation de Bézout, à savoir l'existence d'une famille $(u_1, \ldots, u_n) \in \mathbb{Z}^n$ telle que :

$$\sum_{i=1}^{n} u_i a_i = \mathrm{pgcd}(a_1, \ldots, a_n) \ .$$

4.– Nombres premiers

a) C'est quoi ?

Définition XIII.6. Un entier $p \in \mathbb{N}$ est dit **premier** si $p \neq 1$ et

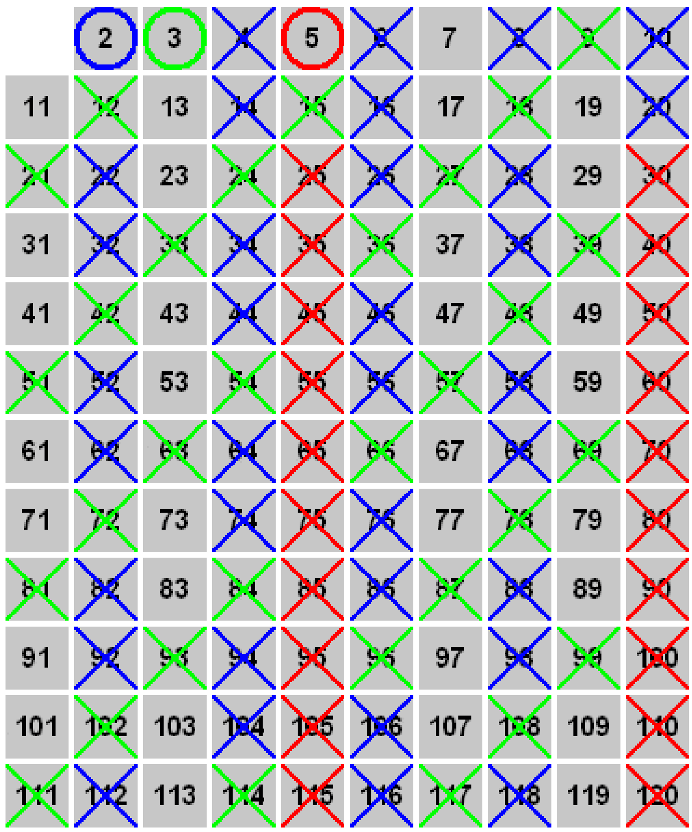
$$\mathscr{D}(p) = \{-1, -p, p, 1\} .$$

✖ **Attention :** : avec cette convention, les nombres premiers seront des entiers **naturels**.

➠ **Exemple XIII.7.** 1 n'est pas premier, 24 non plus ; 3 et 11 sont premiers.

Vocabulaire. Un entier naturel non premier différent de 1 sera dit **composé**.

✌ **Remarque XIII.6.** Pour dresser une table des premiers nombres premiers (ah ah), on peut utiliser un procédé appellé **crible d'Ératosthène** (grec, 276 av. J.-C. — 194 av. J.-C.) : il s'agit de supprimer d'une table des entiers tous les multiples d'un entier. En supprimant tous les multiples, à la fin il ne restera que les entiers qui ne sont multiples d'aucun entier, et qui sont donc les nombres premiers.



Proposition XIII.12.

(i) Deux nombres premiers distincts sont premiers entre eux ;

(ii) toute entier naturel supérieur ou égal à 2 admet un diviseur premier.

Démonstration. Le point (i) est immédiat par intersection des ensembles de diviseurs. Le point (ii) a été traité par récurrence forte au chapitre IV. □

Proposition XIII.13. Il existe une infinité de nombres premiers.

Démonstration. Supposons qu'il existe un nombre fini et notons les $p_1, \dots p_n$; posons :

$$p = \prod_{i=1}^{n} p_i + 1 .$$

L'entier p est supérieur ou égal à 2 donc admet un diviseur premier. Or, tous les p_i sont premiers avec p par Bézout, d'où absurdité. □

b) Décomposition en produits de facteurs premiers

Proposition XIII.14. Soit $a \geq 2$ un entier naturel. Alors, il existe une unique (à l'ordre près) famille de nombres premiers $p_1, \dots, p_n$ et d'exposants associés $\alpha_1, \dots, \alpha_n \in \mathbb{N}^*$ telle que :

$$a = \prod_{k=1}^{n} p_k^{\alpha_k} .$$

Démonstration. L'existence se démontre par récurrence forte sur le modèle de l'existence d'un diviseur premier. Pour l'unicité, supposons qu'il existe deux telles familles, *i.e* que a admette pour écritures

$$a = \prod_{k=1}^{n} p_k^{\alpha_k} \text{ et } a = \prod_{k=1}^{m} q_k^{\beta_k} .$$

Commençons par remarquer que les p_k et q_k sont en fait nécessairement les diviseurs premiers de a et donc $m = n$ et, quitte à les réorganiser, $\forall k \in [\![1, n]\!]$ $p_k = q_k$. De plus, à k fixé on a :

$$p_k^{\alpha_k} | \prod_{j=1}^{n} p_k^{\beta_j}$$

ce qui entraîne par théorème de Gauss (p_k est premiers avec les p_j pour $j \neq k$ que $p_k^{\alpha_k} | p_k^{\beta_k}$ et donc α_k. Il ne nous reste qu'à démontrer symétriquement l'inégalité inverse pour conclure. □

➠ **Exemple XIII.8.** $12 = 2^2 \times 4$, $15 = 3 \times 5$, $32 = 2^5$.

c) Valuation p–adique

Proposition/définition XIII.7. Soit $n \in \mathbb{N}^*$ et soit p un nombre premier. On appelle **valuation p–adique de n** l'entier :

$$\nu_p(n) = \max\{k \in \mathbb{N} \,|\, p^k | n\} \ .$$

Démonstration. Le maximum existe par axiome D (la partie est majorée d'après la proposition XIII.14). □

Exemple XIII.9. $\nu_5(250) = 3$.

Remarque XIII.7. La proposition XIII.14 entraîne l'égalité :

$$n = \prod_{p \text{ premier}} p^{\nu_p(n)} \ .$$

Notons que ce produit est fini car seul un nombre fini de valuations sont non nulles.

Proposition XIII.15. Soient $a, b \in \mathbb{N}^*$ et soit p un nombre premier. Alors :
 (i) $\nu_p(ab) = \nu_p(a) + \nu_p(b)$;
 (ii) $\nu_p(a+b) \geq \min(\nu_p(a), \nu_p(b))$.

Démonstration. Il s'agit d'une conséquence de l'unicité dans la proposition XIII.14. □

Proposition XIII.16. Soient $a, b \in \mathbb{N}^*$ et soit p un nombre premier. Alors :
 (i) $a|b \Leftrightarrow$ pour tout p premier, $\nu_p(a) \leq nu_p(b)$;
 (ii)
$$a \wedge b = \prod_{p \text{ premier}} p^{\min(\nu_p(a), \nu_p(b))}$$
 (iii)
$$a \vee b = \prod_{p \text{ premier}} p^{\max(\nu_p(a), \nu_p(b))} \ .$$

Démonstration. (i) Immédiat.
 (ii) Découle de la caractérisation : si d divise a et b alors pour tout p premier, $\nu_p(d) \leq \nu_p(a)$ et $\nu_p(d) \leq \nu_p(b)$ *ergo* $\nu_p(d) \leq \min(\nu_p(a), \nu_p(b))$ ce qui entraîne que :
$$d | \prod_{p \text{ premier}} p^{\min(\nu_p(a), \nu_p(b))} \ .$$
 (iii) Adapter la démonstration du point précédent.

□

5.⁻ Congruences

On fixe dans tout ce paragraphe $n \in \mathbb{N}$.

◇ Rappels (chapitre V)

Soient $a, b \in \mathbb{Z}$. On dit que a **est congru à** b **modulo** n (noté $a \equiv b\,[n]$) si il existe $k \in \mathbb{Z}$ tel que $b = a + kn$. Ceci définit une relation d'équivalence admettant exactement n classes $\bar{0}, \bar{1}, \ldots, \overline{n-1}$.

a) Opérations sur les congruences

Proposition XIII.17. La relation de congruence modulo n est compatible avec l'addition et la multiplication.

Démonstration. Soient $a, b, a', b' \in \mathbb{Z}$ tels que $a \equiv a'\,[n]$ et $b \equiv b'\,[n]$. Alors il existe $k, h \in \mathbb{Z}$ tels que $a = a' + kn$ et $b = b' + hn$. Ainsi :

$$\begin{aligned} a + b &= a' + b' + n(k + h) \\ &\equiv a' + b'\,[n] \end{aligned}$$

et

$$\begin{aligned} ab =&(a' + kn)(b' + hn) \\ &= a'b' + n(kb' + ha' + kh) \\ &\equiv a'b'\,[n] \end{aligned}$$

d'où le résultat. □

✌ **Remarque XIII.8.** Notons que l'on ne peut pas *a priori* inverser dans une congruence. Par exemple, $2 \times 2 \equiv 0\,[4]$.

b) Petit théorème de Fermat

Lemme XIII.1. Soit p un nombre premier et soient $a, b \in \mathbb{Z}$. Alors :

(i) $\forall k \in [\![1, p-1]\!]$, $p | \binom{p}{k}$;

(ii) $(a + b)^p \equiv a^p + b^p\,[p]$.

Démonstration. (i) Soit $k \in [\![1, p-1]\!]$; alors :

$$p! = k!(p - k)!\binom{p}{k}$$

donc $p|k!(p-k)!\binom{p}{k}$. Or, p n'apparaît pas dans le produit $k!(p-k)!$ d'entiers compris entre 1 et $p-1$ donc, comme p est premier, $p \wedge (k!(p-k)!) = 1$ d'où, par lemme de Gauss (théorème XIII.11) :

$$p|\binom{p}{k}\ .$$

(ii) Il suffit d'appliquer le binôme de Newton :

$$\begin{aligned}(a+b)^p &= \sum_{k=0}^{p} \binom{p}{k} a^k b^{n-k} \\ &= a^p + b^p + \underbrace{\sum_{k=1}^{p-1} \binom{p}{k} a^k b^{n-k}}_{\text{divisible par } p} \\ &\equiv a^p + b^p \, [p] \ .\end{aligned}$$

□

Le théorème suivant est souvent appelé "petit théorème de Fermat", du nom de Pierre de Fermat (français, ∼1610—1665), qui l'énonce dans une lettre à Bernard Frénicle de Bessy, autre mathématicien français et son contemporain.

Théorème XIII.18 (Fermat)**.**
Soit p un nombre premier. Alors :

$$\forall a \in \mathbb{Z},\ a^p \equiv a \, [p] \ .$$

Démonstration. Supposons dans un premier temps a positif ; nous pouvons alors démontrer ce résultat par récurrence sur a.

— $a = 0$: immédiat.

— Supposons la propriété vraie au range $a \geq 0$. Alors :

$$\begin{aligned}(a+1)^p &\equiv a^p + 1^p \, [p] \\ &\equiv a + 1 \, [p]\end{aligned}$$

d'où le résultat.

Dans le cas négatif, nous devons distinguer deux cas :

— si $p > 2$, alors p est impair et donc

$$\begin{aligned}a^p &= -(-a)^p \\ &\equiv -(-a) \, [p] \\ &\equiv a \, [p] \ ;\end{aligned}$$

— si $p = 2$ alors $-1 \equiv 1 \, [2]$ donc $a \equiv -a \, [2]$.

Nous pouvons donc dans les deux cas conclure via le cas $a \geq 0$. □

Corollaire XIII.18.a. Soit p un nombre premier ; alors :

$$\forall a \in \mathbb{Z} \setminus p\mathbb{Z},\ a^{p-1} \equiv 1 \, [p] \ .$$

Démonstration. Nous savons par le théorème de Fermat que $p | a^p - a$. Or $a^p - a = a(a^{p-1} - 1)$ et $p \wedge a = 1$ car $a \notin p\mathbb{Z}$ donc, par lemme de Gauss (théorème XIII.11) :

$$p | a^{p-1} - 1$$

d'où le résultat. □

Ce dernier résultat a été quelques temps utilisé comme test de primalité : si p vérifie la congruence indiquée, on le considérait comme premier. Malheureusement, la réciproque du corollaire est fausse, donc le test est faussé : les faux positifs, appelés nombres de Carmichael (Robert Daniel Carmichael, américain, 1879—1967) comprennent $561, 1105, 1729, 2465, 2821, 6601$ et 8911. Il a été démontré en 1994 par William Alford, Andrew Granville et Carl Pomerance qu'il en existe une infinité.

Chapitre XIV

Équations différentielles

On fixe dans ce paragraphe un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$ et un intervalle I de $\mathbb{R}$ d'intérieur non vide.

1.⁻ Primitives

a) C'est quoi ?

Définition XIV.1. Soit $f : I \to \mathbb{K}$. Une **primitive** de f est une application $F : I \to \mathbb{K}$ telle que :
- $F \in \mathscr{D}^1(I)$;
- $F' = f$.

➠ **Exemple XIV.1.** Soit $\alpha \in \mathbb{C} \setminus \{-1\}$. Alors la fonction

$$\begin{aligned} f : \mathbb{R}_+^* &\to \mathbb{R} \\ x &\mapsto x^\alpha \end{aligned}$$

admet pour primitive $F : x \mapsto \dfrac{x^{\alpha+1}}{\alpha+1}$.

Le résultat qui suit devra pour l'instant être admis. Nous le démontrerons au chapitre XXI.

Théorème XIV.1.
Soit $f \in \mathscr{C}^0(I)$. Alors :
- f admet une infinité de primitives ;
- si F et $\hat{F}$ sont deux primitives de f, la fonction $F - \hat{F}$ est constante.

Notation. Soit $f \in \mathscr{C}^0(I)$ et soit F une primitive de f. SI $a, b \in I$ on appellera **intégrale de f entre a et b** la quantité :

$$\int_a^b f(t)\,\mathrm{d}t := F(b) - F(a)\ .$$

Il est aisé de vérifier que cette dernière ne dépend pas du choix de la primitive F. De plus, la dérivation étant linéaire, l'intégrale l'est naturellement.

✌ **Remarque XIV.1.** Il découle de la "définition" donnée *supra* de l'intégrale que, pour tous $a, x \in I$:

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_a^x f(t)\,\mathrm{d}t = f(x) \; .$$

◇ Primitives usuelles

Il suffit de lire "à l'envers" le tableau des dérivées usuelles pour obtenir ceci. Nous verrons plus tard comment obtenir les primitives "moins évidentes" des fonctions usuelles omises *infra*.

Valeur de $f(x)$	**Ensemble de continuité**	**Valeur d'une primitive**
x^α $(\alpha \neq -1)$	$\mathbb{R}_+^*$	$\frac{x^{\alpha+1}}{\alpha+1}$
$\frac{1}{x}$	$\mathbb{R}_+^*$	$\ln(x)$
e^{ax} $(a \neq 0)$	$\mathbb{R}$	$\frac{1}{a}e^{ax}$
a^x $(a > 0,\, a \neq 1)$	$\mathbb{R}$	$\frac{a^x}{\ln(a)}$
$\cos(x)$	$\mathbb{R}$	$\sin(x)$
$\sin(x)$	$\mathbb{R}$	$-\cos(x)$
$-\frac{1}{\sqrt{1-x^2}}$	$]-1,1[$	$\arccos(x)$
$\frac{1}{\sqrt{1-x^2}}$	$]-1,1[$	$\arcsin(x)$
$\frac{1}{1+x^2}$	$\mathbb{R}$	$\arctan(x)$
$\mathrm{ch}(x)$	$\mathbb{R}$	$\mathrm{sh}(x)$
$\mathrm{sh}(x)$	$\mathbb{R}$	$\mathrm{ch}(x)$

b) Outils calculatoires

◇ Reconnaître une forme composée usuelle

Cette méthode est transparente : si on reconnaît la dérivée d'une composée, c'est gagné. Par exemple, la primitive de $x \mapsto \dfrac{\ln(x)}{x}$ est $x \mapsto \dfrac{1}{2}(\ln(x))^2$ et celle de $x \mapsto \dfrac{1}{x\ln(x)}$ est $\ln \circ \ln$. Même chose pour déterminer une primitive de tan et th.

◇ Intégration par parties (IPP)

Proposition XIV.2. Soient f et g deux fonctions **de classe** $\mathscr{C}^1$ sur I. Alors, pour tous $a, b \in I$:

$$\int_a^b f'(t)g(t)\,\mathrm{d}t = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)\,\mathrm{d}t \,,$$

avec

$$[f(t)g(t)]_a^b = f(b)g(b) - f(a)g(a) \; .$$

Démonstration. Il suffit de remarquer que la fonction fg est une primitive de $(fg)'$;

on a donc :

$$\begin{aligned}[f(t)g(t)]_a^b &= \int_a^b (fg)'(t)\,\mathrm{d}t \\ &= \int_a^b f'(t)g(t)\,\mathrm{d}t + \int_a^b f(t)g'(t)\,\mathrm{d}t\end{aligned}$$

d'où le résultat. □

✖ **Attention :** il est absolument **essentiel** de bien préciser le caractère $\mathscr{C}^1$ de f et g lorsque l'on fait une IPP.

Cette proposition est extrêmement utile pour calculer les primitives de produits dont l'un des termes est peu sensible à la dérivation et/ou contenant un terme polynomiale (qui disparaîtra après un certain nombre de dérivées). En particulier, toutes les constructions de la forme "polynôme × exponentielle" et "polynôme × cosinus ou sinus" devront être traitées de la sorte.

✍ **Exercice XIV.1.** Soit $n \geq 1$ et $x \in \mathbb{R}$. Déterminer

$$I_n(x) := \int_0^x t^n e^{-t}\,\mathrm{d}t\,.$$

➥ **Correction :** *Les fonctions* $t \mapsto t^n$ *et* $t \mapsto -e^{-t}$ *(primitive de* $t \mapsto e^{-t}$*) étant de classe* $\mathscr{C}^1$ *sur* $[0,x]$*, nous pouvons procéder par IPP (proposition XIV.2) :*

$$\begin{aligned}I_n(x) &= \left[t^n(-e^{-t})\right]_0^x - \int_0^x nt^{n-1}(-e^{-t})\,\mathrm{d}t \\ &= -x^n e^{-x} + nI_{n-1}(x)\,.\end{aligned}$$

Ceci constitue une relation de récurrence ; il est ensuite aisé de démontrer que :

$$\begin{aligned}&I_n(x) - x^n e^{-x} + nI_{n-1}(x) \\ &\quad= -x^n e^{-x} + n(-x^{n-1}e^{-x} + (n-1)I_{n-2}(x)) \\ &\quad\vdots \\ &\quad= -x^n e^{-x} - \sum_{k=2}^{n} n \times \ldots \times k\, x^{k-1}e^{-x} + n!I_0(x) \\ &\quad= --\, x^n e^{-x} - \sum_{k=2}^{n} \frac{n!}{(k-1)!}\, x^{k-1}e^{-x} + n!(1-e^{-x}) \\ &\quad= -\sum_{k=2}^{n+1} \frac{n!}{(k-1)!}\, x^{k-1}e^{-x} + n!(1-e^{-x})\end{aligned}$$

◇ Changement de variable

Théorème XIV.3.
Soit $\varphi \in \mathscr{C}^1(I)$ et soit $f \in \mathscr{C}^0(\varphi(I))$. Alors, pour tous $a, b \in I$ on a :

$$\int_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x = \int_a^b f \circ \varphi(t)\varphi'(t)\,\mathrm{d}t\,.$$

Démonstration. Soit F une primitive de f sur $\varphi(I)$ (une telle fonction existe par continuité, *cf.* théorème XIV.1). Alors :

$$\begin{aligned}\int_a^b f\circ\varphi(t)\varphi'(t)\,\mathrm{d}t &= \int_a^b F'\circ\varphi(t)\varphi'(t)\,\mathrm{d}t\\ &= \int_a^b (F\circ\varphi)'(t)\,\mathrm{d}t\\ &= [F\circ\varphi]_a^b\\ &= F(\varphi(b)) - F(\varphi(a))\\ &= \int_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x\,.\end{aligned}$$

□

Cette technique de calcul peut paraître un peu déroutante au début ; il est possible de la visualiser "à la physicienne" de la façon suivante : si $x=\varphi(t)$ alors :

$$\begin{aligned}\int_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x &= \int_a^b f(\varphi(t))\,\mathrm{d}(\varphi(t))\\ &= \int_a^b f\circ\varphi(t)\varphi'(t)\,\mathrm{d}t\,.\end{aligned}$$

Exemple XIV.2.

— Calculons l'intégrale

$$\int_0^1 \frac{e^x}{1+e^{2x}}\,\mathrm{d}x$$

par changement de variable. Pour ce faire, nous posons $\varphi : t\mapsto \ln(t)$; cette fonction est de classe $\mathscr{C}^1$ sur $[1,e]$ d'image $[0,1]$ (nous cherchons donc à "poser $x=\ln(t)$"). On a donc :

$$\begin{aligned}\int_0^1 \frac{e^x}{1+e^{2x}}\,\mathrm{d}x &= \int_1^e \frac{t}{1+t^2}\frac{\mathrm{d}t}{t}\\ &= \int_1^e \frac{1}{1+t^2}\,\mathrm{d}t\\ &= \arctan(e) - \frac{\pi}{4}\,.\end{aligned}$$

— De la même façon, on peut calculer (pour $u\in[-1,1]$)

$$\int_0^u \arccos(x)\,\mathrm{d}x$$

en posant $\varphi : t\mapsto \cos(t)$. On obtient :

$$\int_0^u \arccos(x)\,\mathrm{d}x = -\int_{\frac{\pi}{2}}^{\arccos(u)} t\sin(t)\,\mathrm{d}t$$

que nous pouvons ensuite calculer par IPP en remarquant que $t \mapsto t$ et $\cos(t)$ sont de classe $\mathscr{C}^1$ sur notre intervalle, *i.e*

$$\begin{aligned}
\int_0^u \operatorname{arcos}(x)\,\mathrm{d}x &= -\left[-t\cos(t)\right]_{\frac{\pi}{2}}^{\operatorname{arcos}(u)} + \int_{\frac{\pi}{2}}^{\operatorname{arcos}(u)} (-\cos(t))\,\mathrm{d}t \\
&= u\operatorname{arcos}(u) + \int_{\operatorname{arcos}(u)}^{\frac{\pi}{2}} \cos(t)\,\mathrm{d}t \\
&= u\operatorname{arcos}(u) + 1 - \sin(\operatorname{arcos}(u)) \\
&= u\operatorname{arcos}(u) + 1 - \sqrt{1-u^2}\,.
\end{aligned}$$

◇ **Primitives des inverses de trinômes du second degré**

L'objectif de ce paragraphe est de déterminer les primitives des fonctions de la forme

$$f : x \mapsto \frac{1}{ax^2+bx+c}$$

avec $a, b, c \in \mathbb{C}$.

Cas 1 : $a \neq 0$. Posons alors $\Delta = b^2 - 4ac$ et distinguons deux sous–cas.

Cas 1.1 : $\Delta \neq 0$. Alors il existe $r_1, r_2 \in \mathbb{C}$ tels que pour tous $x \in \mathbb{R}$, $ax^2 + bx + c = a(x - r_1)(x - r_2)$. Cherchons à déterminer $\lambda, \mu \in \mathbb{C}$ tels que :

$$\forall x \in \mathbb{R} \setminus \{r_1, r_2\},\ f(x) = \frac{\lambda}{x-r_1} + \frac{\mu}{x-r_2}\,.$$

Pour ce faire, nous multiplions cette égalité par $x - r_1$:

$$f(x)(x-r_1) = \lambda + \mu\frac{x-r_1}{x_2}$$

puis évaluons en $x = r_1$, ce qui nous donne

$$\lambda = \frac{1}{a(r_1-r_2)}\,.$$

En procédant de même pour μ, nous obtenons que :

$$\forall x \in \mathbb{R} \setminus \{r_1, r_2\},\ f(x) = \frac{1}{a(r_1-r_2)}\left(\frac{1}{x-r_1} - \frac{1}{x-r_2}\right)$$

et donc admet pour primitive, si $r_1, r_2 \in \mathbb{R}$:

$$\begin{aligned}
F :]\max(r_1, r_2), \infty[&\longrightarrow \mathbb{C} \\
x &\mapsto \frac{1}{a(r_1-r_2)}(\ln(x-r_1) - \ln(x-r_2))\,.
\end{aligned}$$

Si les racines sont complexes, il suffit de remarquer que, pour tout x raisonnable et $p, q \in \mathbb{R}$:

$$\begin{aligned}
\frac{1}{x-(p+iq)} &= \frac{1}{(x-p)-iq} \\
&= \frac{(x-p)+iq}{(x-p)^2+q^2}
\end{aligned}$$

et donc cette quantité admet pour primitive

$$x \mapsto \frac{1}{2}\ln((x-p)^2+q^2) + i\arctan\left(\frac{x-p}{q}\right) .$$

Cas 1.2 : $\Delta = 0$**.** Dans ce cas, il existe $r \in \mathbb{C}$ tel que f soit la fonction

$$f : x \mapsto \frac{1}{a(x-r)^2}$$

de primitive

$$\begin{aligned} F : \mathbb{R}\setminus\{r\} &\to \mathbb{C} \\ x &\mapsto -\frac{1}{a(x-r)} . \end{aligned}$$

Cas 2 : $a = 0$**.** Si $b = 0$, c'est trivial. Sinon, on a pour tout $x \neq \dfrac{-c}{b}$:

$$f(x) = \frac{1}{b}\frac{b}{bx+c}$$

et donc f admet pour primitive, si $b, c \in \mathbb{R}$:

$$\begin{aligned} F : \mathbb{R}\setminus\left\{-\frac{c}{b}\right\} &\to \mathbb{C} \\ x &\mapsto \frac{1}{b}\ln(bx+c) . \end{aligned}$$

Notons que si $-\frac{c}{b} \notin \mathbb{R}$, $\mathbb{R}\setminus\left\{-\frac{c}{b}\right\} = \mathbb{R}$. Dans le cas où b ou c n'est pas réel, on obtient en suivant la même méthode que dans le cas 2.1 une primitive combinaison linéaire d'un logarithme et d'une arc tangente.

2.‾ Équations différentielles linéaires du premier ordre

a) C'est quoi ?

Définition XIV.2. Soient $a, b \in \mathscr{C}^0(I,\mathbb{K})$; on appelle **équation différentielle linéaire d'ordre** 1 l'équation

$$y' + ay = b \qquad (\mathscr{E})$$

d'inconnue $y \in \mathscr{D}^1(I,\mathbb{K})$. L'**équation homogène associée** à $(\mathscr{E})$ est alors :

$$y' + ay = 0 \qquad (\mathscr{H}) .$$

Vocabulaire. La fonction a est appelée **coefficient** de l'équation, la fonction b est quand à elle dénommée **second membre** de celle–ci.

✌ **Remarque XIV.2.**

— **Résoudre** une équation différentielle linéaire d'ordre 1 revient à déterminer l'ensemble
$$\mathrm{Sol}_{\mathscr{E}} = \{y \in \mathscr{D}^1(I) \,|\, y' + ay = b\}$$
et similairement pour l'équation homogène associée.

— Si $y \in \mathrm{Sol}_{\mathscr{E}}$, alors comme $y' = -ay + b$, y est automatiquement de classe $\mathscr{C}^1$.

Convention. Nous avons longuement évoqué le fait qu'il était déraisonnable d'écrire des choses du genre $y' + x^2 y = e^x$. C'est cependant la norme dans l'étude des équations différentielles. Ne vous posez pas trop de questions …

➠ **Exemple XIV.3.** $y' + 13y = e^{-x^2+18}$ est une équation différentielle linéaire d'ordre 1 .

b) Résolution

Proposition XIV.4. Soit $(\mathscr{E})$ une équation différentielle linéaire d'ordre 1 d'équation homogène $(\mathscr{H})$ et soit $y_0 \in \mathrm{Sol}_{\mathscr{E}}$. Alors :
$$\mathrm{Sol}_{\mathscr{E}} = \{y_0 + y \,|\, y \in \mathrm{Sol}_{\mathscr{H}}\} .$$

Démonstration.

($\supset$) Immédiat, il suffit de réinjecter dans $(\mathscr{E})$.

($\subset$) Soit y une solution de $(\mathscr{E})$; alors on vérifie aisément en réinjectant que $y - y_0 \in \mathrm{Sol}_{\mathscr{H}}$.

$\square$

Cela signifie que la résolution d'une équation différentielle linéaire d'ordre 1 devra nécessairement se faire en trois temps :

1. résoudre l'équation homogène associée ;
2. déterminer une solution particulière y_0 de l'équation ;
3. combiner ces deux données via la proposition précédente.

◇ Résolution homogène

Proposition XIV.5. Soit $a \in \mathscr{C}^0(I)$; on considère l'équation différentielle linéaire d'ordre 1 homogène
$$y' + ay = 0 \qquad (\mathscr{H}) .$$
Alors, si A est une primitive de a :
$$\mathrm{Sol}_{\mathscr{H}} = \{x \mapsto Ce^{-A(x)} \,|\, C \in \mathbb{K}\} .$$

Démonstration.

($\supset$) Si y est de la forme $x \mapsto Ce^{-A(x)}$, on vérifie rapidement que $y' = -ay$.

($\subset$) Soit $y \in \mathrm{Sol}_{\mathscr{H}}$; alors, pour tout $x \in I$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(y(x)e^{A(x)}\right) &= y'(x)e^{A(x)} + a(x)y(x)e^{A(x)} \\ &= \left(y'(x) + a(x)y(x)\right)e^{A(x)} \\ &= 0\end{aligned}$$

et donc la fonction $x \mapsto y(x)e^{A(x)}$ est constante car de dérivée nulle sur l'intervalle I. Ceci entraîne le résultat.

□

✌ **Remarque XIV.3.** Si a est une fonction constante égale à $\kappa \in \mathbb{R}$, on a donc :

$$\mathrm{Sol}_{\mathscr{H}} = \{x \mapsto Ce^{-\kappa x} \mid C \in \mathbb{K}\}\ .$$

➠ **Exemple XIV.4.** Résolvons l'équation

$$y' + \frac{3}{x}y = 0 \qquad (\mathscr{H})$$

sur $\mathbb{R}_+^*$. Il s'agit d'une équation différentielle linéaire d'ordre 1 homogène de coefficient $a : x \mapsto \frac{3}{x}$ continu sur $\mathbb{R}_+^*$ et de primitive $x \mapsto \ln(x^3)$. Ainsi :

$$\begin{aligned}\mathrm{Sol}_{\mathscr{H}} &= \left\{x \mapsto Ce^{-\ln(x^3)} \;\middle|\; C \in \mathbb{C}\right\} \\ &= \left\{x \mapsto \frac{C}{x^3} \;\middle|\; C \in \mathbb{C}\right\}\ .\end{aligned}$$

◇ Recherche d'une solution particulière

La seconde étape du procédé de résolution d'une équation différentielle linéaire d'ordre 1 est un tantinet moins codifiée et repose parfois sur des techniques *ad–hoc*. Nous donnons ici quelques méthodes utilisables, à compléter en TD.

Recherche d'une solution évidente. Cela arrive, parfois. Par exemple, une solution particulière de $y' + 3y = 7$ est $x \mapsto \dfrac{7}{3}$ et une solution de $y' + xy = x$ est $x \mapsto 1$. De façon générale, toujours vérifier si il existe des solutions constantes à notre équation…

Nous verrons en TD que l'on peut appliquer des méthodes de ce type de façon heuristique si a est constante et que b est d'une certaine forme.

Méthode de variation de la constante. L'idée de cette méthode est relativement simple : on remplace le "C" dans l'expression des solutions homogène par une fonction, faisant ainsi "varier la constante". Plus rigoureusement, on recherche une solution particulière sous la forme

$$y_0 : x \mapsto \lambda(x)e^{-A(x)}$$

où A est une primitive du coefficient a de notre équation et $\lambda \in \mathscr{C}^1(I)$. On a alors, pour tout $x \in I$:

$$\begin{aligned}y_0'(x) + a(x)y_0(x) &= \lambda'(x)e^{-A(x)} - a(x)y_0(x) + a(x)y_0(x) \\ &= \lambda'(x)e^{-A(x)}\end{aligned}$$

et donc, si b est le second membre de l'équation :

$$\lambda'(x) = b(x)e^{A(x)} .$$

Nous sommes donc en capacité de déterminer une solution particulière de l'équation à condition de savoir "primitiver" la fonction $x \mapsto b(x)e^{A(x)}$ (ce qui n'est pas automatique !).

✖ **Attention :** ne pas oublier de multiplier λ par la solution homogène à la fin du procédé !

➠ **Exemple XIV.5.** En recherchant une solution particulière de

$$y' + xy = e^{-\frac{x^2}{2}} \qquad (\mathbf{E}:\text{XIV.1})$$

sous la forme $y_0 : x \mapsto e^{-\frac{x^2}{2}}$, on tombe sur $\lambda' = 1$ et donc $y_0 : x \mapsto xe^{-\frac{x^2}{2}}$ convient.

✍ **Exercice XIV.2.** Résoudre l'équation :

$$(x+1)y' - xy + 1 = 0 \qquad (\mathbf{E}:\text{XIV.2})$$

sur $I =]-1, \infty[$.

➥ **Correction :** *Il faut commencer par mettre cette équation sous forme normale :*

$$y' - \frac{x}{x+1}y = -\frac{1}{x+1} \qquad (\mathbf{E}:\text{XIV.3})$$

et d'appliquer les méthodes vu précédemment ; on peut alors reconnaître une équation différentielle linéaire d'ordre 1 à coefficient et second membre continus sur I. Pour trouver une primitive du coefficient, on remarque que :

$$\forall x \in I, \ \frac{x}{x+1} = \frac{x+1-1}{x+1} = 1 - \frac{1}{x+1}$$

et donc $x \mapsto x - \ln(x+1)$ convient. Les solutions homogènes sont donc les

$$x \mapsto C\frac{e^x}{x+1}$$

*pour $C \in \mathbb{C}$. La méthode de variation de la constante livre ensuite l'équation différentielle $\lambda' = -e^{-x}$ et donc les solutions de (**E** :XIV.2) sont les*

$$x \mapsto \frac{1}{x+1} + C\frac{e^x}{x+1}$$

pour $C \in \mathbb{C}$.

Principe de superposition. Notons au passage que si le second membre b d'une équation différentielle linéaire d'ordre 1 est de la forme $b = \sum_{i=1}^{n} b_i$, la linéarité de l'équation entraîne qu'il suffit de déterminer une solution particulière associée à chaque b_i et des les sommer.

c) Problèmes de Cauchy

Définition XIV.3. Soient $a, b \in \mathscr{C}^0(I)$ et soit $(x_0, y_0) \in I \times \mathbb{K}$. On appelle **problème de Cauchy linéaire du premier ordre** associé à ces données le système :

$$\begin{cases} y' + ay = b \\ y(x_0) = y_0 \end{cases}$$

d'inconnue $y \in \mathscr{D}^1(I)$.

Vocabulaire. Les données x_0 et y_0 sont appelées **conditions initiales du système**.

➠ **Exemple XIV.6.**

$$\begin{cases} y' + 3y = \pi e^{-18x^9} \\ y(0) = 127\gamma + 2i \end{cases}$$

est un problème de Cauchy sur $I = \mathbb{R}$.

Théorème XIV.6 (Cauchy–Lipschitz, version faible)**.**
Tout problème de Cauchy linéaire du premier ordre admet une unique solution.

Démonstration. Par variation de la constante, on sait qu'il existe une solution particulière y_p de l'équation différentielle sous–jacente. De fait, il existe une infinité de solutions de celle–ci, de la forme :

$$y : x \mapsto y_p(x) + Ce^{-A(x)}$$

avec $C \in \mathbb{K}$ et A une primitive de a. La constante C est alors déterminée par l'équation $y(x_0) = y_0$, qui entraîne que :

$$C = -y_p(x_0)e^{A(x_0)}$$

d'où le résultat. □

➠ **Exemple XIV.7.** Si on rajoute à l'équation **E** :XIV.2 la condition initiale $y(0) = 1$, l'unique solution du problème de Cauchy est :

$$y : x \mapsto \frac{1}{x+1}\,.$$

3.⁻ Équations différentielles linéaires du second ordre à coefficients constants

a) C'est quoi ?

Définition XIV.4. Soient $a, b \in \mathbb{K}$ et $f \in \mathscr{C}^0(I, \mathbb{K})$. On appelle **équation différentielle linéaire d'ordre** 2 **à coefficients constants** associée à ces données l'équation

$$y'' + ay' + by = f \qquad (\mathscr{E})$$

d'inconnue $y \in \mathscr{D}^2(I, \mathbb{K})$. L'**équation homogène associée** est alors

$$y'' + ay' + by = 0 \qquad (\mathscr{H})$$

Vocabulaire. Les **constantes** a et b sont appelées **coefficients** de l'équation. La **fonction** f est appelée **second membre** de celle–ci.

Remarque XIV.4. De la même façon que pour les équation différentielle linéaire d'ordre 1 , une solution d'une telle équation sera automatiquement de classe $\mathscr{C}^2$.

Exemple XIV.8.
— $y'' + iy' + \cos(13 + \sqrt{2}\pi)y = x^2$ est une équation différentielle linéaire d'ordre 2 à coefficients constants.
— $y'' + y = 0$ est une équation différentielle linéaire d'ordre 2 homogène à coefficients constants dont on connaît déjà deux solutions sur $\mathbb{R}$: cos et sin.

Définition XIV.5. Soit $y'' + ay' + by = 0$ une équation différentielle linéaire d'ordre 2 homogène à coefficients constants ; on appelle **équation caractéristique** associée l'équation polynomiale

$$X^2 + aX + b = 0 \qquad (\chi) .$$

Exemple XIV.9. L'équation caractéristique associée à $y'' + y = 0$ est $X^2 + 1 = 0$.

b) Résolution

Proposition XIV.7. Soit $(\mathscr{E})$ une équation différentielle linéaire d'ordre 2 à coefficients constants d'équation homogène $(\mathscr{H})$ et soit $y_0 \in \mathrm{Sol}_{\mathscr{E}}$. Alors :

$$\mathrm{Sol}_{\mathscr{E}} = \{y_0 + y \,|\, y \in \mathrm{Sol}_{\mathscr{H}}\} .$$

Ce résultat, analogue à celui vu pour le premier ordre, nous indique que la résolution d'une équation différentielle linéaire d'ordre 2 à coefficients constants se fera selon les mêmes "temps" que celle d'une équation différentielle linéaire d'ordre 1 .

◇ Résolution homogène

Proposition XIV.8. On considère une équation différentielle linéaire d'ordre 2 homogène à coefficients constants

$$y'' + ay' + bt = 0 \qquad (\mathscr{H})$$

dont l'équation caractéristique a pour discriminant $\Delta = a^2 - 4b$. Alors :

— si $\Delta \neq 0$, l'équation caractéristique admet deux solutions distinctes $r_1, r_2 \in \mathbb{C}$ et :

$$\mathrm{Sol}_{\mathscr{H}} = \{x \mapsto \lambda e^{r_1 x} + \mu e^{r_2 x} \mid \lambda, \mu \in \mathbb{C}\}\ ;$$

— si $\Delta = 0$, l'équation caractéristique admet une unique solution $r_0 \in \mathbb{C}$ et :

$$\mathrm{Sol}_{\mathscr{H}} = \{x \mapsto (\lambda x + \mu) e^{r_0 x} \mid \lambda, \mu \in \mathbb{C}\}\ .$$

Démonstration. Dans les deux cas, l'inclusion de droite à gauche est triviale. Pour la réciproque, fixons $\phi \in \mathrm{Sol}_{\mathscr{H}}$ et α une racine de l'équation caractéristique $X^2+aX+b = 0$. Alors, en posant

$$g : x \mapsto \phi(x) e^{-\alpha x}$$

on a :

$$\begin{aligned} 0 &= \phi'' + a\phi' + b\phi \\ &= g'' + (2\alpha + a)g' + \underbrace{(\alpha^2 + a\alpha + b)}_{=0} g \end{aligned}$$

car pour tout $x \in I$, $\phi(x) = g(x)e^{\alpha x}$. Au final, on obtient donc

$$g'' + (2\alpha + a)g' = 0 \qquad (\mathbf{E}\text{:XIV.4})$$

qui est une équation différentielle linéaire d'ordre 1 homogène en g'. En la resolvant, on obtient qu'il existe $C \in \mathbb{K}$ tel que :

$$\forall x \in I,\ g'(x) = Ce^{-(2\alpha+a)x}\ .$$

Cas 1 : $\Delta \neq 0$. Soit β l'unique racine de l'équation caractéristique distincte de α ; nous savons que $\alpha + \beta = -a \neq 2\alpha$ donc $2\alpha + a \neq 0$. Nous pouvons donc intégrer sans crainte l'expression de g' *supra* et obtenir qu'il existe $A, B \in \mathbb{K}$ tels que

$$\forall x \in I,\ g(x) = Ae^{-(2\alpha+a)x} + B$$

et donc

$$\begin{aligned} \forall x \in I,\ \phi(x) &= g(x)e^{\alpha x} \\ &= Ae^{-(\alpha+a)x} + Be^{\alpha x} \\ &= Ae^{\beta x} + Be^{\alpha x} \end{aligned}$$

d'où le résultat.

Cas 2 : $\Delta = 0$.Dans ce cas, $2\alpha = -a$ car α est l'unique racine double de l'équation caractéristique, donc g' est constante. Il suffit alors d'intégrer et de multiplier par $x \mapsto e^{\alpha x}$ comme dans le cas précédent pour obtenir le résultat voulu.

□

Remarque XIV.5.
— Notons que l'expression des solutions présente deux "degrés de liberté".
— Si Δ est un réel strictement négatif, les racines de l'équation caractéristique sont complexes conjuguées, *i.e* $r_1 = a + ib$ et $r_2 = a - ib$ pour $a, b \in \mathbb{R}$. De fait, les solutions **réelles** de $(\mathscr{H})$ sont de la forme

$$x \mapsto e^{ax}(\alpha \cos(bx) + \beta \sin(bx))$$

pour $\alpha, \beta \in \mathbb{R}$.

Exemple XIV.10. Les solutions complexes de $y' + y = 0$ sont les

$$x \mapsto Ae^{ix} + Be^{-ix}$$

pour $A, B \in \mathbb{C}$. En passant à la partie réelle, on obtient que ses solutions réelles sont les

$$x \mapsto C\cos(x) + D\sin(x)$$

pour $C, D \in \mathbb{R}$.

◇ **Recherche d'une solution particulière**

Recherche d'une solution évidente. Comme à l'ordre 1, il convient de vérifier si notre équation n'admet pas de solution "immédiate". Par exemple, $x \mapsto \frac{e^x}{3}$ est solution de $y'' + y' + y = e^x$.

Cas particuliers. La méthode de variations des constantes est hors–programme. Nous donnons cependant quelques "recettes" pour traiter les cas le plus couramment rencontrés aux concours. Plus précisément, si le second membre de l'équation étudiée est de la forme

$$x \mapsto Ae^{\alpha x}$$

avec $A, \alpha \in \mathbb{C}$, il convient de rechercher une solution particulière sous la forme

$$x \mapsto P(x)e^{\alpha x}$$

avec :
— si α n'est pas une racine de l'équation caractéristique, P constant ;
— si α est racine simple de l'équation caractéristique, P de degré 1 ;
— si α est racine double de l'équation caractéristique, P de degré 2.

Notons que cette méthode nous permet, via l'exponentielle complexe et un passage à la partie réelle ou imaginaire de résoudre les cas où le second membre est une fonction cos ou sin.

Exemple XIV.11. Cette méthode permet de résoudre rapidement les équations

$$y'' + y = e^x \qquad (\mathbf{E}:\text{XIV.5})$$

et

$$y'' - y = e^x \ . \qquad (\mathbf{E}\text{:XIV.6})$$

Pour les solutions particulières, la première en admet une évidente ($x \mapsto \frac{e^x}{2}$) et la seconde admet $x \mapsto \frac{x}{2}e^x$, obtenue via la méthode *supra* (1 est racine simple de $X^2 - 1 = 0$).

Remarque XIV.6. Le principe de superposition se généralise au second ordre.

c) Problèmes de Cauchy

Définition XIV.6. Soient $a, b \in \mathbb{K}$, $f \in \mathscr{C}^0(I)$ et soit $(x_0, y_0, y_1) \in I \times \mathbb{K} \times \mathbb{K}$. On appelle **problème de Cauchy linéaire du second ordre** associé à ces données le système :

$$\begin{cases} y'' + ay' + by = f \\ y(x_0) = y_0 \\ y'(x_0) = y_1 \end{cases}$$

d'inconnue $y \in \mathscr{D}^2(I)$.

Vocabulaire. Les données x_0, y_0 et y_1 sont appelées **conditions initiales du système**.

Exemple XIV.12.

$$\begin{cases} y'' + y = 0 \\ y(0) = 1 \\ y'(0) = 0 \end{cases}$$

est un problème de Cauchy sur $I = \mathbb{R}$.

Théorème XIV.9 (Cauchy–Lipschitz, version faible bis)**.**
Tout problème de Cauchy linéaire du second ordre admet une unique solution.

Exemple XIV.13. L'unique solution du problème de Cauchy de l'exemple précédent est :

$$y : x \mapsto \cos(x) \ .$$

Chapitre XV

Polynômes

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.⁻ L'algèbre $\mathbb{K}[X]$

a) Polynômes à une indéterminée

Définition XV.1. On appelle **polynôme à une indéterminé** toute suite $(a_n)_n \in \mathbb{K}^{\mathbb{N}}$ telle que :

$$\exists\, d \in \mathbb{N},\ \forall n \geq d,\ a_n = 0\,.$$

Les termes de cette suite sont appelés **coefficients** du polynôme.

Vocabulaire. Une suite vérifiant la propriété *supra* est dite **presque nulle** : seul un nombre fini de ses termes sont en effet différents de 0.

Notation.

— L'ensemble des polynômes à une indéterminée sur $\mathbb{K}$ est noté $\mathbb{K}[X]$.
— Le polynôme correspondant à la suite nulle sera noté 0.
— Le polynôme correspondant à la suite $(1, 0, \ldots)$ est noté 1 ou X^0.
— De façon générale, le polynôme correspondant à la suite $(\delta_{k,n})_n$ pour $k \geq 0$ est noté X^k, où $\delta_{k,n}$ est le **symbole de Kronecker** valant 0 si $k \neq n$ et 1 sinon.

Proposition XV.1. $(\mathbb{K}[X], +)$ est un sous–groupe de $(\mathbb{K}^{\mathbb{N}}, +)$. Il s'agit donc d'un groupe abélien de neutre $(0)_n$.

Démonstration. Trivial. □

➠ **Exemple XV.1.** $1 + X$ est le polynôme correspondant à la suite $(1, 1, 0 \ldots)$.

Remarquons ensuite que si $\lambda \in \mathbb{K}$ et $P = (a_n)_n$ est un polynôme, nous pouvons définir (de façon compatible avec l'addition des suites) le polynôme

$$\lambda P = (\lambda a_n)_n \ \in \mathbb{K}[X].$$

In fine, nous avons donc additionner les polynômes entre eux et les multiplier par des éléments de $\mathbb{K}$ (nous parlerons plus tard de **scalaires**, *cf.* chapitre XIX). De plus, nous pouvons avec ces conventions écrire, pour $P = (a_n)_n \in \mathbb{K}[X]$ que :

$$P = a_0X^0 + a_1X + a_2X^2 + \dots$$

soit

$$P = \sum_{k=0}^{\infty} a_k X^k .$$

✖ **Attention :** cette somme n'est **PAS** réellement infinie, étant donné que seul un nombre fini des termes a_k sont non nuls.

Définition XV.2. Soit $P \in \mathbb{K}[X]$. On appelle **degré** de P la quantité

$$\deg(P) = \begin{cases} \max\{N \in \mathbb{N} \,|\, a_N \neq 0\} \text{ si } P \neq 0 \\ -\infty \text{ si } P = 0 \end{cases} .$$

Un polynôme de degré nul ou $-\infty$ est dit **constant**.

✌ **Remarque XV.1.** Le degré est bien défini car toute partie de $\mathbb{N}$ non vide et majorée admet un plus grand élément (proposition IV.1).

➠ **Exemple XV.2.** $\deg(1) = 0$, $\deg(X^{17} - X) = 17$.

Notation. Pour $d \in \mathbb{N} \cup \{-\infty\}$, on note $\mathbb{K}_d[X]$ l'ensemble des polynômes de degré inférieur ou égal à d. Rappelons que par convention sur la droite réelle achevée, $-\infty$ est strictement inférieur à tout entier naturel.

Par conséquent, si $P = (a_n)_n \in \mathbb{K}[X]$ est de degré $d \geq 0$, nous pouvons écrire :

$$P = \sum_{k=0}^{d} a_k X^k .$$

✌ **Remarque XV.2.**

— Il découle de tout ceci que si $P = (a_n)_n, Q = (b_n)_n \in KK[X]$ alors :
$$P = Q$$
$$\Longleftrightarrow$$
$$(\deg(P) = \deg(Q)) \wedge (\forall k \leq \deg(P),\ a_k = b_k) .$$

— En `python`, un polynôme sera généralement représenté par la liste de ses coefficients, de façon similaire aux suites.

Définition XV.3. Soit $P \in \mathbb{K}[X]$. On appelle **coefficient dominant** de P son coefficient non nul d'indice le plus élevé. Si P est non nul de degré d, il s'agit du coefficient placé devant X^d dans l'écriture de P en tant que somme.
Si le coefficient dominant de P est égal à 1, le polynôme est dit **unitaire**.

Notation. $\mathrm{cd}(P)$

Exemple XV.3. $\mathrm{cd}(1 + X^2 + 2X^4) = 2$.

Proposition XV.2. Soient $P, Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}^*$. Alors :

(i) $\deg(\lambda P) = \deg(P)$;

(ii) $\deg(P + Q) \leq \max(\deg(P), \deg(Q)$ avec égalité si $\deg(P) \neq \deg(Q)$.

Démonstration.

(i) Trivial.

(ii) Posons $P = \sum\limits_{k=0}^{d} a_k X^k$ et $Q = \sum_{k=0}^{d'} b_k X^k$ avec $(d, d') = (\deg(P), \deg(Q))$.
Alors :
— si $d \neq d'$, par exemple $d < d'$ $P + Q$ admet $b_{d'}$ et son degré est clairement d'.
— si $d = d'$ alors pour tout $k > d$ on a $a_k + b_k = 0$ ce qui entraîne le résultat.
Pour un contre exemple à l'égalité dans le cas où les degrés sont égaux, additionner 1 et -1.

□

b) Produit de polynômes

Proposition/définition XV.4. Soient $P, Q \in \mathbb{K}[X]$, de coefficients respectifs $(a_i)_i$ et $(b_i)_i$. Alors :

(i) la suite de terme général (pour $k \geq 0$)

$$\sum_{j=0}^{k} a_j b_{k-j}$$

est un polynôme, appelé produit de P et Q et noté PQ ;

(ii) $\deg(PQ) = \deg(P) + \deg(Q)$.

Démonstration. Si P ou Q est le polynôme nul, inutile de trop se fatiguer. Dans le cas contraire, posons $d = \deg(P)$ et $d' = \deg(Q)$. Alors, pour tout $k > d + d'$ on a :

$$R_k = \sum_{j=0}^{k} a_j b_{k-j} = 0$$

car si $j > d$, $a_j = 0$ et si $j \leq d$ alors $k - j > d'$ et donc $b_{k-j} = 0$. Le produit PQ est donc bien un polynôme (suite presque nulle) et

$$\deg(PQ) \leq d + d' .$$

Pour obtenir l'égalité des degrés, il nous suffit de constater que

$$R_{d+d'} = a_d b_{d'} \neq 0 .$$

□

✌ **Remarque XV.3.** On a de fait la formule suivante :

$$\left(\sum_{k=0}^{d} a_k X^k\right)\left(\sum_{k=0}^{d'} b_k X^k\right) = \sum_{k=0}^{d+d'}\left(\sum_{j=0}^{k} a_j b_{k-j}\right) X^k .$$

Corollaire XV.2.a. $(\mathbb{K}[X], +, \times)$ est un anneau commutatif de neutres 0 et $1 = X^0$.

✌ **Remarque XV.4.** De plus, un produit de polynômes est nul si et seulement si l'un de ses facteurs est nul (on parle d'anneau intègre). Cela découle de la formule du degré d'un produit.

Proposition XV.3. Les inversibles de $\mathbb{K}[X]$ sont les polynômes constants non nuls.

Démonstration. Si $P \in \mathbb{K}[X]^\times$, alors il existe $Q \in \mathbb{K}[X]$ tel que $PQ = 1$. En passant au degré, on obtient que $\deg(P) + \deg(Q) = 0$, d'où le résultat. □

✌ **Remarque XV.5.** On vérifie par récurrence que, pour tout $k \in \mathbb{N}^*$, $\prod_{i=1}^{k} X = X^k$, ce qui est rassurant.

c) Composition

Définition XV.5. Soient $P, Q \in \mathbb{K}[X]$, avec $P = \sum_{k=0}^{d} a_k X^k$. On appelle composée de P par Q le polynôme :

$$P \circ Q = \sum_{k=0}^{d} a_k Q^k .$$

✌ **Remarque XV.6.** Il s'agit bien d'un polynôme par structure annelée sur $\mathbb{K}[X]$.

➠ **Exemple XV.4.** $X^2 \circ (X+1) = (X+1)^2$.

Proposition XV.4. Soient $P, Q \in \mathbb{K}[X]$ avec Q non constant. Alors :

$$\deg(P \circ Q) = \deg(P)\deg(Q) .$$

✖ **Attention :** cela est faux si $\deg(Q) = 0$; en effet, la composée de $X - 1$ par 1 est de degré $-\infty$.

Démonstration. Posons $d = \deg(P)$ et $P = \sum_{k=0}^{d} a_k X^k$. Alors

$$P \circ Q = \sum_{k=0}^{d} a_k Q^k$$

donc, par somme et comme tous les $\deg(Q^k)$ sont distincts car Q est non constant, $\deg(P \circ Q) = \max_{k \leq d} \deg(a_k Q^k)$, ce dernier étant égal à $\deg(Q^d)$ car $a_d \neq 0$. □

Proposition XV.5. La composition des polynômes est associative, non commutative, admet X pour neutre et est distributive **à droite** par rapport à l'addition.

Démonstration. Hastur, Hastur, Hastur. □

En résumé, il convient de retenir que :
— $X^2 \circ (X+1) = (X+1)^2 \neq (X+1) \circ X^2 = X^2 + 1$;
— si $P, Q, H \in \mathbb{K}[X]$ on a $P \circ (Q \circ H) = (P \circ Q) \circ H$ et $(P+Q) \circ H = P \circ Q + Q \circ H$ mais, comme vu *supra*, $P \circ (Q+H) \neq P \circ Q + P \circ H$.

2.‾ Arithmétique des polynômes

a) Multiples, diviseurs

Définition XV.6. Soient $A, B \in \mathbb{K}[X]$. On dira que A **divise** B, ou que B est un **multiple** de A, si il existe $C \in \mathbb{K}[X]$ tel que $B = AC$.

Notation.
— $A|B$.
— On notera $\mathscr{D}(A)$ l'ensemble des diviseurs de A et $A\mathbb{K}[X]$ l'ensemble de ses multiples.

✌ **Remarque XV.7.**
— Comme sur $\mathbb{Z}$, 0 est divisible par tout polynôme et $A|B \Leftrightarrow B\mathbb{K}[X] \subset A\mathbb{K}[X]$.
— Soit $P \in \mathbb{K}[X]$ et soit $\lambda \in \mathbb{K}^*$. Alors $P = (\lambda P) \times \frac{1}{\lambda}$ et donc $\lambda | P$.
— Par degré d'un produit, si $A|B$ alors $\deg(A) \leq \deg(B)$.

Proposition XV.6. La relation "|" est réflexive et transitive. De plus, pour tous $A, B \in \mathbb{K}[X]$ on a :

$$(A|B) \wedge (B|A)$$
$$\Leftrightarrow$$
$$\exists\, \lambda \in \mathbb{K}^*,\ A = \lambda B.$$

On dit alors que les polynômes A et B sont **associés**.

Démonstration. Le premier point se traite de façon analogue à ce que nous avons vu au chapitre XIII. Pour le second, le sens "bas vers haut" est trivial et si A, B sont tels que $(A|B) \wedge (B|A)$ alors il existe $P, Q \in \mathbb{K}[X]$ tels que $A = PB$ et $B = AQ$ et donc $A = PQA$, *i.e* $A(1 - PQ) = 0$, ce qui entraîne que soit $A = 0$ (et dans ce cas $B = 0$) soit $PQ = 1$ ce qui n'est possible que si $P, Q \in \mathbb{K}[X]^\times = \mathbb{K}^*$. □

Théorème XV.7 (Division euclidienne)**.**
Soient $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$. Alors il existe un unique couple $(Q, R) \in \mathbb{K}[X]^2$ tel que :
- $A = BQ + R$;
- $\deg(R) < \deg(B)$.

Vocabulaire. Comme dans le cas entier, on parle de quotient, reste, diviseur et dividende.

Exemple XV.5. $X^2 + X + 1 = X(X + 1) + 1$ est une division euclidienne.

Démonstration.

Existence. Si $\deg(A) < \deg(B)$ ou $A = 0$ alors $Q = 0$ et $R = A$ conviennent. Dans le cas contraire, démontrons l'existence du couple (Q, R) par récurrence forte sur $d = \deg(A)$. **Attention à la formulation de l'hypothèse de récurrence :** nous voulons montrer que, pour tout $d \geq 0$, pour tous polynômes A, B tels que $d = \deg(A) \geq deg(B)$ il existe un couple (Q, R) vérifiant les conclusions du théorème.
- Si $d = 0$, alors A et B sont inversibles ($\deg(B) \leq \deg(A)$) et donc $A = B \times \frac{A}{B} + 0$.
- Si on suppose l'hypothèse vérifiée jusqu'à un certain rang $d \geq 0$ et que l'on se donne $A = \sum_{k=0}^{d+1} a_k X^k$ et $B = \sum_{k=0}^{d'} b_k X^k$ deux polynômes tels que $d + 1 = \deg(A) \geq d' = \deg(B)$ alors :
$$A' = A - \frac{a_{d+1}}{b_{d'}} X^{d+1-d'} B$$
est un polynôme de degré au plus d. Par hypothèse de récurrence (ou, au pire, car $\deg(B) > \deg(A')$), il existe donc $(Q, R) \in \mathbb{K}[X]^2$ tels que $A' = BQ + R$ et $\deg(R) < \deg(B)$. Ceci entraîne que :
$$A = B\left(\frac{a_{d+1}}{b_{d'}} X^{d+1-d'} + Q\right) + R$$
d'où le résultat.

Unicité. Si il existe deux couples (Q, R) et (Q', R') vérifiant les conclusions du théorème alors $B(Q - Q') = R' - R$ et donc
$$\deg(B) + \deg(Q - Q') = \deg(R' - R) < \deg(B)$$
d'où $\deg(Q - Q') = -\infty$ et le résultat.

□

Remarque XV.8. La démonstration du théorème XV.7 nous livre un algorithme récursif implémentable en pratique.

b) PGCD

Proposition/définition XV.7. Soient $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$. On appelle **plus grand commun diviseur** de A et B tout élément de degré maximal de $\mathscr{D}(A) \cap \mathscr{D}(B)$.

Démonstration. Un tel élément existe car $\{\deg(P) \mid P \in \mathscr{D}(A) \cap \mathscr{D}(B)\}$ est une partie de $\mathbb{N}$ (car $B \neq 0$ donc $\mathscr{D}(B)$ ne contient pas $-\infty$) non vide (il contient 0 car les constantes non nulles divisent tout polynôme) et majorée par $\max(\deg(A), \deg(B))$. □

✖ **Attention :** il n'y a pas unicité : X^2 et $X^2 + X$ admettent (entre autres) X et $-X$ comme PGCD. En fait, c'est même bien pire que cela : si D est un PGCD de A et B alors λD l'est également pour tout $\lambda \in \mathbb{K}^*$ ...

Proposition XV.8. Soient $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$ et soit $\Delta \in \mathbb{K}[X]$. Alors :

$$\Delta \text{ est } \textbf{un} \text{ PGCD de } A \text{ et } B \iff \mathscr{D}(A) \cap \mathscr{D}(B) = \mathscr{D}(\Delta) \,.$$

Démonstration. Analogue au cas entier vu dans le chapitre XIII. □

Corollaire XV.8.a. Soient $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$ et soient Δ, Δ' deux PGCD de A et B. Alors Δ et Δ' sont associés.

✌ **Remarque XV.9.** Par conséquent, si Δ est **un** PGCD de A et B alors l'ensemble des PGCD de ces deux polynômes est

$$\{\lambda \Delta \mid \lambda \in \mathbb{K}^*\} \,.$$

Cet ensemble contient donc un unique polynôme unitaire.

Définition XV.8. On appelle PGCD de deux polynômes leur unique PGCD unitaire.

Notation. $A \wedge B$

Ceci nous permet d'énoncer une caractérisation du PGCD de $A, B \in \mathbb{K}[X]$ analogue à celle vue au chapitre XIII :

$$\Delta = A \wedge B \iff \begin{cases} \mathscr{D}(A) \cap \mathscr{D}(B) = \mathscr{D}(\Delta) \\ \operatorname{cd}(\Delta) = 1 \end{cases} .$$

✌ **Remarque XV.10.** Si $A \neq 0$ alors $A \wedge 1 = 1$ et $A \wedge A = \frac{A}{\operatorname{cd}(A)}$.

À la question "comment trouver le PGCD de deux polynômes ?", nous donnerons la réponse suivante : il faut commencer par suivre l'algorithme d'Euclide vu dans le chapitre XIII puis ensuite diviser le dernier reste non nul par son coefficient dominant. Il est **essentiel de ne pas omettre cette dernière étape.**

Exemple XV.6.
— $(X^2+3X+1)\wedge(X+1)=1$;
— $(X^2-3X+2)\wedge(X^3-2X^2+X-2)=X-2$.

Remarque XV.11. Comme dans le cas entier, l'algorithme d'Euclide étendu permet d'obtenir un couple U, V tel que $AU+BV=A\wedge B$.

c) Bézout et Gauss

Définition XV.9. Deux polynômes A et B sont dits **premiers entre eux** si $A\wedge B=1$.

Le théorème qui suit est, contrairement au théorème XIII.10, réellement du à Étienne Bézout(français, 1730—1783).

Théorème XV.9 (Bézout)**.**
Soient $A, B \in \mathbb{K}[X]$ tels que $B \neq 0$. Alors :

$$\begin{gathered}A \text{ et } B \text{ sont premiers entre eux}\\ \Longleftrightarrow\\ \exists\, U, V \in \mathbb{K}[X],\ AU+BV=1.\end{gathered}$$

Démonstration. Analogue au cas entier. □

De la même façon, le lemme de Gauss vu au chapitre XIII se généralise au cas d'anneaux de polynômes.

Théorème XV.10 (Lemme de Gauss)**.**
Soient $A, B, C \in \mathbb{K}[X]$ tels que $B \neq 0$. On suppose que :
— $A|BC$;
— $A\wedge B=1$.
Alors A divise C.

Démonstration. Analogue au cas entier. □

Exercice XV.1. Résoudre dans $\mathbb{K}[X]$ l'équation $(X^3-1)U+(X^2+1)V=2X^2$.

Correction : *Il s'agit d'adapter la méthode vue lors de l'étude des équations diophantiennes. Comme X^3-1 et X^2+1 sont premiers entre eux, on trouve par Euclide étendu que :*

$$(X^3-1)(X-1)+(1+X-X^2)(X^2+1)=2$$

et donc on trouve une solution particulière $U_0 = X^2(X-1)$, $V_0 = X^2(1+X-X^2)$ et on démontre à l'aide du lemme de Gauss que les seules solutions sont alors de la forme $(U_0 + (X^2+1)C, V_0 - (X^3-1)C)$, avec $C \in \mathbb{K}[X]$.

d) PPCM et généralisations

De la même façon, on définit le PPCM de deux polynômes A et B comme l'unique polynôme **unitaire** $A \vee B$ tel que $A\mathbb{K}[X] \cap B\mathbb{K}[X] = (A \vee B)\mathbb{K}[X]$. Il s'agit de l'unique élément unitaire de degré maximal de $A\mathbb{K}[X] \cap B\mathbb{K}[X]$. On retrouve alors l'égalité :

$$(A \wedge B)(A \vee B) = \frac{AB}{\mathrm{cd}(AB)} .$$

De plus, on peut définir le PGCD (resp. le PPCM) d'une famille de polynômes de la même façon que pour les entiers. On généralise également l'existence de relations de Bézout aux familles de n polynômes premiers entre eux dans leur ensemble.

3.– Fonctions polynomiales

a) C'est quoi ?

Définition XV.10. Soit $P = \sum_{k=0}^{d} a_k X^k \in \mathbb{K}[X]$. On appelle **fonction polynomiale associée à** P l'application

$$\begin{aligned} f : \mathbb{K} &\to \mathbb{K} \\ x &\mapsto \sum_{k=0}^{d} a_k x^k . \end{aligned}$$

Notation. On note $\mathbb{K}[x]$ l'ensemble des fonctions polynomiales sur $\mathbb{K}$.

✖ **Attention :** si $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$, la quantité "$P(a)$" n'a a priori aucun sens. Si f est la fonction polynomiale associée à P, $f(a)$ est par contre bien défini. On parle malgré tout d'**évaluation** du polynôme P en a.

➠ **Exemple XV.7.** La fonction polynomiale associée à X^2 est (normalement) bien connue du lecteur.

✌ **Remarque XV.12.**
— On montre aisément que $(\mathbb{K}[x], +, \times)$ est un sous–anneau de $\mathbb{K}^{\mathbb{K}}$.
— De plus, si $\lambda \in \mathbb{K}$ et $f \in \mathbb{K}[x]$ alors $\lambda f \in \mathbb{K}[x]$.

◇ Lien à la choucroute

On considère désormais l'application suivante, que l'on sait surjective par construction :

$$\varphi : \mathbb{K}[X] \longrightarrow \mathbb{K}[x]$$
$$\sum_{k=0}^{d} a_k X^k \mapsto \left(x \mapsto \sum_{k=0}^{d} a_k x^k \right) .$$

Ceci signifie, rappelons le, que toute fonction polynôme correspond à (au moins) un polynôme. Cool. Mais encore ? Eh bien on peut vérifier que, pour tous $P, Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$ on a :

$$\varphi(P + \lambda Q) = \varphi(P) + \lambda\varphi(Q) \quad \text{et} \quad \varphi(PQ) = \varphi(P)\varphi(Q) .$$

La question qui nous brûle les lèvres à ce stade est naturellement la suivante : la fonction φ est–elle injective ? À savoir, étant donné deux polynômes P, Q tels que $\varphi(P) = \varphi(Q)$, a–t–on $P = Q$? La réponse devra hélas attendre quelques temps ...

b) Racines

Définition XV.11. Soit $P \in \mathbb{K}[X]$ et soit $f = \varphi(P)$. On appelle **racine de** P tout scalaire $a \in \mathbb{K}$ vérifiant $f(a) = 0$.

Notation. On notera $\text{Rac}(P)$ l'ensemble des racines de P. Il sera parfois/souvent utile de spécifier le corps sur lequel nous travaillons ; on notera alors $\text{Rac}_{\mathbb{K}}(P)$.

➠ **Exemple XV.8.**
— $\text{Rac}_{\mathbb{C}}(X^2 + 1) = \{i, -i\}$;
— $\text{Rac}_{\mathbb{R}}(X^2 + 1) = \emptyset$;
— $\text{Rac}_{\mathbb{C}}(X^n - 1) = \mathscr{U}_n$ pour tout $n \geq 1$.

Proposition XV.11. Soit $P \in \mathbb{K}[X]$ et soit $a \in \mathbb{K}$. Alors :

$$a \in \text{Rac}(P)$$
$$\iff$$
$$X - a \text{ divise } P.$$

Démonstration.

($\Uparrow$) Immédiat : si $P = (X - a)Q$ avec $Q \in \mathbb{K}[X]$ alors $\varphi(P) = x \mapsto (x - a)\varphi(Q)(x)$.

($\Downarrow$) Supposons que $a \in \text{Rac}(p)$. Par division euclidienne, il existe un unique couple (Q, R) de polynômes tels que

$$P = (X - a)Q + R$$

et $\deg(R) < 1$. De fait, R est constant et comme $\varphi(P)(a) = 0$ on a $\varphi(R)(a) = 0$. Ainsi, $R = 0$ (car il est constant).

□

Corollaire XV.11.a. Soit $P \in \mathbb{K}[X]$ et soient $a_1, \dots a_n \in \mathrm{Rac}(P)$ deux à deux distincts. Alors $\prod_{k=1}^{n}(X - a_k)$ divise P.

Démonstration. Récurrence... □

On déduit de tout ceci la proposition fondamentale suivante, qui sera centrale à l'étude des racines de polynômes.

Proposition XV.12. Soit $P \in \mathbb{K}[X]$ de degré $n \geq 0$. Alors $\mathrm{Rac}(P)$ contient au plus n éléments.

Démonstration. Si $a_1, \dots, a_k$ sont des racines deux à deux distinctes de P on a que

$$\prod_{i=1}^{k}(X - a_i) \text{ divise } P$$

et donc

$$k = \deg\left(\prod_{i=1}^{k}(X - a_i)\right) \leq \deg(P) = n\,,$$

d'où le résultat. □

✌ **Remarque XV.13.** On déduit de ceci le résultat suivant, fort utile en pratique : **tout polynôme admettant plus de racines que son degré est nul**.

Théorème XV.13.
L'application φ est bijective.

Démonstration. Il ne reste qu'à montrer que φ est injective. Si on suppose que $P, Q \in \mathbb{K}[X]$ sont tels que $\varphi(P) = \varphi(Q)$ alors $\varphi(P - Q) = \varphi(P) - \varphi(Q) = 0$ et donc

$$\forall a \in \mathbb{K}, \quad \varphi(P - Q)(a) = 0\,.$$

Ceci entraîne, comme le corps $\mathbb{K}$ est infini, que $P - Q$ admet une infinité de racines. Il s'agit donc du polynôme nul, *ergo* $P = Q$. □

✌ **Remarque XV.14.** Il nous sera donc possible d'identifier (comme vous le faisiez probablement depuis un bon moment) fonctions polynomiales et polynômes. Ceci entraîne qu'il est désormais permis d'écrire des choses sulfureuses comme "$P(a)$" en toute impunité.

Définition XV.12. Soit $P \in \mathbb{K}[X]$ et soit $a \in \mathbb{K}$. On appelle **multiplicité de** a **relativement à** P la quantité

$$\mu_P(a) = \max\{k \in \mathbb{N} \,|\, (X-a)^k | P\}\,.$$

Vocabulaire. Si $\mu_P(a) = 1$, on parle de racine simple ; si $\mu_P(a) \geq 2$, on parle de racine multiple (double, triple,...).

✌ **Remarque XV.15.** Il apparaît clairement que $a \in \mathrm{Rac}(P) \Leftrightarrow \mu_P(a) \neq 0$.

➠ **Exemple XV.9.** On pourra faire le parallèle avec le cours de première et les racines simples, doubles, des trinômes du second degré.

Définition XV.13. Un polynôme $P \in \mathbb{K}[X]$ est dit **scindé sur** $\mathbb{K}$ si la relation suivante est vérifiée :

$$\sum_{a \in \mathrm{Rac}_{\mathbb{K}}(P)} \mu_P(a) = \deg(P)\,.$$

➠ **Exemple XV.10.** Tout polynôme de degré 2 est scindé sur $\mathbb{C}$.

✖ **Attention :** le caractère scindé dépend lui aussi du corps : comparer X^2+1 vu comme polynôme à coefficients complexes et son jumeau maléfique dans $\mathbb{R}[X]$...

Proposition XV.14. Soit $P \in \mathbb{K}[X]$. Alors :

$$P \text{ est scindé}$$

$$\Longleftrightarrow$$

$$P = \mathrm{cd}(P) \prod_{a \in \mathrm{Rac}(P)} (X-a)^{\mu_P(a)}\,.$$

Démonstration.

($\Uparrow$) Trivial.

($\Downarrow$) Par définition de la multiplicité, pour tout $a \in \mathrm{Rac}(P)$, $(X-a)^{\mu_P(a)}$ divise P ; ces polynômes étant premiers entre eux on a, par lemme de Gauss :

$$\prod_{a \in \mathrm{Rac}(P)} (X-a)^{\mu_P(a)} \mid P\,.$$

De plus, le degré de ces deux polynômes sont, comme P est scindé, égaux. On en déduit que leur quotient est de degré 0. Ce dernier est alors obligatoirement égal à $\mathrm{cd}(P)$ par identification des coefficients dominants.

□

c) Relations coefficients–racines

Autorisons nous à présent une (relativement) brève digression ; si $P = aX^2+bX+c \in \mathbb{K}[X]$ est un trinôme du second degré (avec donc $a \neq 0$), et que son discriminant est $\Delta = b^2 - 4ac$, il admet deux racine(s) (éventuellement égale(s)) donnée(s) par la formule

$$r_\pm = \frac{-b \pm \delta}{2a}$$

avec $\delta^2 = \Delta$. On en déduit que :

$$\begin{cases} r_+ + r_- = -\dfrac{b}{a} \\ r_+ r_- = \dfrac{b^2 - \Delta}{4a^2} = \dfrac{c}{a} \end{cases}$$

ce qui entraîne que :

$$\begin{aligned} aX^2 + bX + c &= a\left(X^2 + \frac{b}{a}X + \frac{c}{a}\right) \\ &= a(X^2 - (r_+ + r_-)X + r_+ r_-) \\ &= a(X - r_+)(X - r_-) \ . \end{aligned}$$

Ces relations entre coefficients et racines sont appelées, dans un élan d'originalité à faire pâlir un scénariste de chez Marvel, des **relations coefficients–racines**.

En dégré 3, on voit apparaître des choses similaires ; en effet, si $P = (X-x_1)(X-x_2)(X-x_3)$ est un polynôme scindé de degré 3, on obtient, en développant :

$$\begin{aligned} P &= (X - x_1)(X - x_2)(X - x_3) \\ &= X^3 + aX^2 + bX + c \end{aligned}$$

avec

$$\begin{cases} a = -(x_1 + x_2 + x_3) \\ b = x_1x_2 + x_1x_3 + x_2x_3 \\ c = -x_1x_2x_3 \end{cases} .$$

De façon plus générale, si $n, k \in \mathbb{N}^*$ sont tels que $k \leq n$, on appelle **k–ième fonction symétrique élémentaire à n variables** l'application

$$\begin{aligned} \sigma_k : \mathbb{K}^n &\longrightarrow \mathbb{K}^n \\ (x_1, \ldots, x_n) &\to \sum_{1\leq i_1<\ldots<i_k\leq n} x_{i_1} \ldots x_{i_k} \ . \end{aligned}$$

En particulier, pour tout $(x_1, \ldots, x_n) \in \mathbb{K}^n$, on a :

$$\sigma_1(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \ ,$$

$$\sigma_2(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} x_i x_j$$

et

$$\sigma_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i \ .$$

Ces fonctions ont un lien fort avec les relations que nous venons de voir pour les degrés 2 et 3. Celles–ci se généralisent via le théorème suivant.

Théorème XV.15 (Relations coefficients–racines)**.**
Soit $P = \sum_{k=1}^{n} a_k X^k$ un polynôme scindé de degré n dont les racines (comptées avec multiplicité) sont notées $r_1, \dots, r_n$. Alors :

$$\forall k \in [\![1, n]\!], \quad \sigma_k(r_1, \dots, r_n) = (-1)^k \frac{a_{n-k}}{a_n} .$$

✍ **Exercice XV.2.** Résoudre :

$$\begin{cases} x + y + z &= 2 \\ \frac{1}{x} + \frac{1}{y} + \frac{1}{z} &= \frac{5}{6} \\ xyz &= -6 \end{cases} .$$

➥ **Correction :** *En multipliant la ligne 2 par la ligne 3, on arrive au système :*

$$\begin{cases} x + y + z &= 2 \\ yz + xz + xy &= -5 \\ xyz &= -6 \end{cases} .$$

Les solutions sont donc les racines du polynôme $X^3 - 2X^2 - 5X + 6 = (X-1)(X^2 - X - 6)$.

4.– Dérivation

a) Dérivée formelle d'un polynôme

Définition XV.14. Soit $P = \sum_{k=0}^{d} a_k X^k \in \mathbb{K}[X]$. On appelle **dérivée** (formelle) de P le polynôme

$$P' = \sum_{k=1}^{d} k a_k X^{k-1} .$$

➠ **Exemple XV.11.** $(X^3 + X + 1)' = 3X^2 + 1$.

✌ **Remarque XV.16.**

— On définit par récurrence les dérivées formelles d'ordre supérieur d'un polynôme.

— On vérifie aisément que la fonction polynomiale (d'une variable réelle) associée à la dérivée formelle d'un polynôme est la dérivée de la fonction polynomiale associée à ce même polynôme. En particulier, les fonctions polynomiales sont de classe $\mathscr{C}^\infty(\mathbb{R})$.

— Il découle du point précédent que les formules usuelles de dérivation se généralisent à la dérivée formelle, quitte dans la cas complexe à travailler sur la restriction de la fonction polynomiale à $\mathbb{R}$. On peut également vérifier ces résultats par le calcul.

Proposition XV.16. Soit $P \in \mathbb{K}[X]$ un polynôme non constant. Alors :
(i) $\deg(P') = \deg(P) - 1$;
(ii) si $P \in \mathbb{R}[X]$, $\forall \alpha, \beta \in \mathrm{Rac}_{\mathbb{R}}(P)$, $\exists\, \gamma \in]\alpha, \beta[\cap \mathrm{Rac}_{\mathbb{R}}(P')$.

Démonstration. (i) Immédiat par définition.
(ii) Découle du théorème de Rolle appliqué à la fonction polynomiale associée à P.
□

➠ **Exemple XV.12.** La dérivée du polynôme $(X-2)(X-7)(X-18)(X-\pi)$ admet une racine dans $]2, \pi[$.

Corollaire XV.16.a. La dérivée d'un polynôme réel scindé à racines simples l'est également.

Démonstration. Un tel polynôme admet exactement autant de racines simples que son degré, et sa dérivée s'annule une fois entre chacune d'entre elles, d'où le résultat.
□

✌ **Remarque XV.17.** Ce résultat se généralise aux polynômes scindés simples sur $\mathbb{C}$.

b) Formule de Taylor polynomiale

Proposition XV.17. Soit $P \in \mathbb{K}[X]$ et soit $a \in \mathbb{K}$. Alors

$$P = \sum_{k=0}^{\infty} \frac{(X-a)^k}{k!} P^{(k)}(a) .$$

✌ **Remarque XV.18.**
— La notation "$P^{(k)}(a)$" est abusive, mais nous sommes entre gens de bonne compagnie.
— Comme de coutume, la somme apparaissant dans la formule est en faite finie, car (par décroissance du degré), $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ P^{(n)} = 0$.
— On peut donc déduire des valeurs successives des dérivées d'un polynôme en un point l'expression générale de celui–ci.

Démonstration. Vous l'aurez deviné, c'est reparti pour une récurrence sur le $n = \deg(P)$ …

— Si P est constant, tout va bien.
— Sinon, yaka appliquer l'hypothèse de récurrence à P' et "primitiver", puis être content.

□

➠ **Exemple XV.13.** La formule de Taylor permet de déterminer l'unique polynôme P de degré inférieur ou égal à deux vérifiant $P(0) = 0$, $P'(0) = 2$ et $P'(0) = 3$.

Nous verrons dans le chapitre XVI que la formule de Taylor permet d'approximer des fonctions à l'aide de polynômes.

Corollaire XV.17.a. Soit $a \in \mathbb{K}$, $k \in \mathbb{N}$ et $P \in \mathbb{K}[X]$. Alors :

$$\mu_p(a) = k \iff \begin{cases} P(a) = P'(a) = \ldots = P^{(k-1)}(a) = 0 \\ P^{(k)}(a) \neq 0 \end{cases}.$$

Démonstration.

($\Leftarrow$) Cela découle de la formule de Taylor : P est divisible par $(X-a)$, $(X-a)^2$, $\ldots$, $(X-a)^{k-1}$ mais pas par $(X-a)^k$.

($\Rightarrow$) On le fait par récurrence sur k et c'est la fête. Pour l'hérédité, remarquer que si $\mu_P(a) = k+1$ alors $P = (X-a)^{k+1}Q$ avec $Q(a) \neq 0$ et $P' = (X-a)^k R$ avec $R(a) \neq 0$ (c'est un calcul), ce qui permet de conclure.

□

5.⁻ Irréductibilité

a) C'est quoi ?

Définition XV.15. Un polynôme $P \in \mathbb{K}[X]$ est dit **irréductible** si :
— P est non constant ;
— si $P = QR$ avec $Q, R \in \mathbb{K}[X]$ alors Q ou R est constant.

➠ **Exemple XV.14.** Les polynômes de degré 1 sont constants sur $\mathbb{R}$ ou $\mathbb{C}$ et les polynômes de degré deux de discriminant négatif sur $\mathbb{R}$.

✖ **Attention :** l'irréductibilité dépend du corps sur lequel on travaille : $X^2+1 = (X-i)(X+i)$ sur $\mathbb{C}$.

Théorème XV.18.
Tout polynôme non constant s'écrit de façon unique (à l'ordre près des facteurs et à constante multiplicative près) comme produit de polynômes irréductibles.

✌ **Remarque XV.19.** Ce résultat est analogue à la proposition XIII.14 (décomposition en produits de facteurs premiers sur $\mathbb{Z}$). Ceci n'est pas une coïncidence.

Exemple XV.15. $X^3-1=(X-1)(X-j)(X-j^2)=(X-1)(X^2+X+1)$; on remarque que la décomposition dépend (sans surprise) du corps sur lequel on travaille.

Démonstration. **Existence.** Youpi, une récurrence forte sur le degré ! Attention à la formulation de l'hypothèse. . . .

— $d=1$. C'est plié.

— **Supposons la propriété vérifiée pour tout** $k \in [\![1,n]\!]$ **avec** $n \geq 1$ **fixé.** Si P est irréductible, c'est fini. Sinon, il existe deux polynômes $Q, R \in \mathbb{K}[X]$ non constants tels que $P=QR$. Nous pouvons appliquer l'hypothèse de récurrence à ceux–ci et voila, c'est plié.

Unicité. Nous en parlerons plus tard. Là, j'ai water–poney.

□

Exemple XV.16. Le polynôme X^n-1 (pour $n \geq 1$) admet exactement n racines sur $\mathbb{C}$: les éléments de $\mathscr{U}_n$. Ceci entraîne la factorisation suivante (**sur** $\mathbb{C}$) :

$$X^n-1=\prod_{\xi\in\mathscr{U}_n}(X-\xi)\,.$$

b) C'est qui (édition complexe) ?

Le théorème central de ce paragraphe est le suivant, parfois appelé théorème fondamental de l'algèbre. Son histoire est complexe (ah ah) et intrinsèquement lié à celle du corps $\mathbb{C}$ dont on peut considérer qu'il est la motivation première pour l'étude. On trouve des traces de ce résultat chez François Viète (français, 1540—1603), Albert Girard (français, 1595—1632) et Renée Descartes, que l'on ne présente plus (1596—1650), entre autres.

Une première démonstration de ce résultat est esquissée par Jean le Rond d'Alembert (français, 1717—1783). Celle–ci est incomplète, malgré quelques ajouts ultérieurs par Jean–Robert Agrand (amateur suisse, 1768—1822). Une démonstration complète n'apparaîtra que suite aux efforts (indépendants) de Lagrange, Euler et Gauss, qui en produit la première démonstration complète (et fort analytique) en 1816.

Théorème XV.19 (D'Alembert—Gauss).
Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine.

Démonstration. Admis. □

Corollaire XV.19.a. Les polynômes irréductibles de $\mathbb{C}$ sont les polynômes de degré un.

Démonstration. Il est clair que les polynômes de degré un sont irréductibles. Réciproquement, si $P \in \mathbb{C}[X]$ est irréductible, il admet une racine $a \in \mathbb{C}$ d'après le théorème XV.19 et donc est divisible par $X-a$. Par irréductibilité, P est de la forme $\lambda(X-a)$ avec $\lambda \in \mathbb{K}$. □

✌ **Remarque XV.20.** La décomposition apparaissant dans le théorème XV.18 est donc unique car déterminée par les racines. Plus précisément, si $P \in \mathbb{C}[X]$, on a :

$$P = \mathrm{cd}(P) \prod_{\alpha \in \mathrm{Rac}_{\mathbb{C}}(P)} (X-\alpha)^{\mu_P(\alpha)} .$$

Proposition XV.20. Soient $P, Q \in \mathbb{C}[X]$. Alors :

$$P|Q \iff \begin{cases} \mathrm{Rac}_{\mathbb{C}}(P) \subset \mathrm{Rac}_{\mathbb{C}}(Q) \\ \forall \alpha \in \mathbb{C},\ \mu_P(\alpha) \leq \mu_Q(\alpha) \end{cases} .$$

Démonstration. Ceci découle de la décomposition énoncée **supra**. □

➠ **Exemple XV.17.** Comparer les racines de X^2+X+1 et $(X^3-1)^2$.

c) C'est qui (retour au(x) réel(s)) ?

Les choses se compliquent : il est temps de dégainer les lemmes ...

Lemme XV.1. Soit $P \in \mathbb{R}[X]$. Alors, pour tout $\alpha \in \mathbb{C}$:

$$(\alpha \in \mathrm{Rac}_{\mathbb{C}}(P)) \Rightarrow (\overline{\alpha} \in \mathrm{Rac}_{\mathbb{C}}(P)) .$$

Démonstration. Si $P = \sum\limits_{k=0}^{d} a_k X^k$, on a :

$$\begin{aligned} 0 &= \overline{P(\alpha)} \\ &= \overline{\sum_{k=0} a_k \alpha^k} \\ &= \sum_{k=0} a_k \overline{\alpha}^k \\ &= P(\overline{\alpha}) \end{aligned}$$

d'où le résultat. □

Lemme XV.2. Soit $P \in \mathbb{R}[X]$ de degré supérieur ou égal à 3. Alors P est réductible sur $\mathbb{R}$.

Démonstration. D'après le théorème de d'Alembert Gauss (XV.19), il existe $\alpha \in \mathrm{Rac}_{\mathbb{C}}(P)$ et $\overline{\alpha}$ est également une racine de P (non nécessairement distincte de α d'ailleurs). De fait, il existe $Q \in \mathbb{C}[X]$ tel que :

$$P = (X-\alpha)(X-\overline{\alpha})Q .$$

Or $(X-\alpha)(X-\overline{\alpha}) = X^2 - 2\mathrm{Re}(\alpha)X + |\alpha|^2 \in \mathbb{R}[X]$ donc, par unicité dans la division euclidienne (théorème XV.7), $Q \in \mathbb{R}[X]$. Ce polynôme est de plus non constant car $\deg(P) \geq 3$, d'où le résultat. □

Théorème XV.21.
Les polynômes irréductibles de $\mathbb{R}$ sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant négatif.

Démonstration. Les polynômes cités sont clairement irréductibles. Réciproquement, si P est irréductible de degré 2, son discriminant est négatif (sans quoi il admet des racines réelles et donc une factorisation). Le cas du degré 3 et supérieur a été traité dans le lemme précédent. □

✌ **Remarque XV.21.** La décomposition apparaissant dans le théorème XV.18 est unique : cela découle du fait que la décomposition complexe est unique. On a de plus la formule suivante, pour tout $P \in \mathbb{R}[X]$:

$$P = \mathrm{cd}(P) \prod_{\alpha \in \mathrm{Rac}_{\mathbb{R}}(P)} (X-\alpha)^{\mu_P(\alpha)} \prod_{\beta \in \mathrm{Rac}_{\mathbb{C}}(P) \setminus \mathrm{Rac}_{\mathbb{R}}(P)} (X^2 - 2\mathrm{Re}(\beta)X + |\beta|^2)^{\mu_P(\beta)} .$$

➠ **Exemple XV.18.**

$$\begin{aligned} X^4 - 1 &= (X-1)(X+1)(X-i)(X+i) \\ &= (X-1)(X+1)(X^2+1) . \end{aligned}$$

6.$^-$ Interpolation de Lagrange

Dans ce paragraphe, nous nous intéressons à la question suivante : étant donné $(x_0, y_0), \dots, (x_n, y_n) \in \mathbb{K}^2$, comment trouver un polynôme $P \in \mathbb{K}[X]$ tel que :

$$\forall k \in [\![0, n]\!],\ P(x_k) = y_k \ ? \qquad (\mathbf{E}:\mathrm{XV}.1)$$

Pour $n = 0$, le polynôme constant égal à y_0 fera l'affaire. Pour $n = 1$, nous avons deux points à relier ; une droite (et donc un polynôme de degré 1) conviendra. Dans ces deux cas, la solution au problème semble unique.

Proposition XV.22. Soit $n \geq 0$ et soient $x_0, \dots x_n \in \mathbb{K}$ deux à deux distincts. Alors, pour toute famille $y_0, \dots, y_n \in \mathbb{K}$, il existe un unique polynôme $P \in \mathbb{K}_n[X]$ tel que :

$$\forall k \in [\![0, n]\!],\ P(x_k) = y_k .$$

Démonstration. **Unicité.** Si deux tels polynômes existent, leur différence admet $n+1$ racines (les x_k) et est donc nulle par argument de degré.

Existence. Posons, pour $k \in [\![0, n]\!]$:

$$L_k = \prod_{j \neq k} \frac{(X - x_j)}{(x_k - x_j} .$$

On a alors l'égalité $L_k(x_j) = \delta_{k,j}$ pour tous $k, j \in [\![0, n]\!]$ et donc

$$P = \sum_{k=0}^{n} y_k L_k$$

convient.

□

✌ **Remarque XV.22.** L'unicité n'est plus garantie si la condition de degré est supprimée ; il existe même une infinité de solutions possibles à l'équation **E** :XV.1.

➠ **Exemple XV.19.** $X+1$ interpole les couples $(0,1)$ et $(1,2)$. Mais $(X+1)+X(X-1)$ aussi ...

✌ **Remarque XV.23.** Tout ceci se code très bien en `python`. Une instabilité numérique est présente si n est très grand, car le polynôme "cherche" toujours à s'annuler autant de fois que son degré. La fonction suivante prend en argument les listes `X` et `Y` contenant respectivement les x_k et les y_k.

```
def lagrange(X,Y,a):
    def L(k,X,a):
        p=1
        for i in range(len(X)):
            if i != k:
                p*=(a-X[i])/(X[k]-X[i])
        return p

    if len(X) != len(Y):
        raise ValueError("Tailles␣incompatibles")
    n = len(X)
    res=0
    for k in range(n):
        res+=Y[k]*L(k,X,a)
    return res
```

Chapitre XVI

Analyse asymptotique

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.⁻ Comparaison des suites

a) Domination, négligeabilité, équivalence

Définition XVI.1. Soient $u, v \in \mathbb{K}^{\mathbb{N}}$. On dit que :

— u est **dominée** par v, noté $u_n = \mathcal{O}(v_n)$, si il existe une suite **bornée** $w \in \mathbb{K}^{\mathbb{N}}$ telle que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = w_n v_n\ ;$$

— u est **négligeable** devant v, noté $u_n = o(v_n)$, si il existe une suite **convergeant vers** 0 $w \in \mathbb{K}^{\mathbb{N}}$ telle que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = w_n v_n\ ;$$

— u est **équivalente** à v, noté $u_n \sim v_n$, si il existe une suite **convergeant vers** 1 $w \in \mathbb{K}^{\mathbb{N}}$ telle que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = w_n v_n\,.$$

✌ **Remarque XVI.1.**

— Si $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ v_n \neq 0$ alors :

(i) $u_n = \mathcal{O}(v_n) \Leftrightarrow \left(\dfrac{u_n}{v_n}\right)_n$ est bornée ;

(ii) $u_n = o(v_n) \Leftrightarrow \dfrac{u_n}{v_n} \xrightarrow[n\to\infty]{} 0$;

(iii) $u_n \sim v_n \Leftrightarrow \dfrac{u_n}{v_n} \xrightarrow[n\to\infty]{} 1$.

— Soit $u \in \mathbb{K}^{\mathbb{N}}$ une suite convergeant vers $\ell \in \mathbb{K}^*$. Alors, $\dfrac{u_n}{\ell} \to 1$ et donc $u_n \sim \ell$.

— Seule les suites stationnaires à 0 sont équivalentes à 0.

➠ **Exemple XVI.1.**

— $n^2 = o(n^3)$;

— $2n = \mathcal{O}(n)$;
— $n+1 \sim n$.

Proposition XVI.1. (i) Négligeabilité implique domination ;
(ii) les relations o et $\mathcal{O}$ sont transitives ;
(iii) l'équivalence est une relation... d'équivalence.

Démonstration. Trivial. □

✌ **Remarque XVI.2.** Notons que les croissances comparées vues au chapitre I peuvent se traduire par les résultats suivants, si $\alpha, \beta, \gamma \in \mathbb{R}_+$:
— $\ln(n)^\beta = o(n^\alpha)$;
— $n^\alpha = o(e^{\gamma n})$.

Proposition XVI.2. Soient $u, v \in \mathbb{K}^\mathbb{N}$. Alors :

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

✌ **Remarque XVI.3.** Ce résultat reste vrai en remplaçant le "o" par un "$\mathcal{O}$".

Démonstration.

($\Rightarrow$) On sait qu'il existe une suite $(w_n)_n$ convergeant vers 1 telle que :

$$\exists\, N \in \mathbb{N},\, \forall n \geq N,\, u_n = w_n v_n$$

donc on a :

$$\forall n \geq N,\, u_n - v_n = (\underbrace{w_n - 1}_{\to 0})v_n$$

d'où le résultat.

($\Leftarrow$) On sait qu'il existe une suite $(h_n)_n$ convergeant vers 0 telle que :

$$\exists\, N \in \mathbb{N},\, \forall n \geq N,\, u_n - v_n = h_n v_n$$

donc on a :

$$\forall n \geq N,\, u_n = (\underbrace{h_n + 1}_{\to 1})v_n$$

d'où le résultat.

□

b) Opérations et comparaisons

Proposition XVI.3. Soient $u, v, w \in \mathbb{K}^{\mathbb{N}}$ telles que $u_n = \mathcal{O}(w_n)$ et $v_n = \mathcal{O}(w_n)$. Alors :

(i) $u_n + v_n = \mathcal{O}(w_n)$;

(ii) $u_n v_n = \mathcal{O}(w_n^2)$.

✌ **Remarque XVI.4.** Ce résultat reste vrai en remplaçant "$\mathcal{O}$" par "o".

Démonstration. On sait, quitte à prendre le maximum entre les deux rangs où tout fonctionne (*cf.* chapitre IX pour de nombreux exemples de tels raisonnements) qu'il existe deux suites $h, k \in \mathbb{K}^{\mathbb{N}}$ bornées telles que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = h_n w_n \ \text{ et } \ v_n = k_n w_n\,.$$

De fait, pour tout $n \geq N$ on a :

$$u_n + v_n = (h_n + k_n) w_n = \mathcal{O}(w_n)$$

et

$$u_n v_n = (h_n k_n) w_n^2 = \mathcal{O}(w_n^2)$$

d'où le résultat. □

✌ **Remarque XVI.5.** Cette proposition pourra être (abusivement) résumée par : "$o(w_n) + o(w_n) = o(w_n)$ (et idem avec $\mathcal{O}$).

Proposition XVI.4. Soient $u, v, u', v' \in \mathbb{K}^{\mathbb{N}}$ telles que $u_n \sim u'_n$ et $v_n \sim v'_n$. Alors :

(i) $u_n v_n \sim u'_n v'_n$;

(ii) **si** v ne s'annule pas au delà d'un certain rang, c'est également le cas de v' et

$$\frac{u_n}{v_n} \sim \frac{u'_n}{v'_n}\ ;$$

(iii) pour tout $k \geq 1$, $u_n^k \sim (u'_n)^k$.

✖ **Attention :** on ne somme **pas** les équivalents : $n \sim n+1$ et $-n \sim -n+1$ et pourtant $0 \not\sim 2 \ldots$

Démonstration. Quitte à yada yada, il existe deux suites $h, k \in \mathbb{K}^{\mathbb{N}}$ convergeant vers 1 telles que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = h_n u'_n \ \text{ et } \ v_n = k_n v'_n$$

et donc, pour tout $n \geq N$:

$$u_n v_n = \underbrace{(h_n k_n)}_{\to 1} u' v_n$$

d'où le résultat pour le produit. Le reste se démontre de façon analogue. □

➠ **Exemple XVI.2.** $\dfrac{n^2+n}{n^3+n^2} \sim \dfrac{1}{n}$ car $n^2 + n \sim n^2$ et $n^3 + n^2 \sim n^3$.

c) Invariants

Proposition XVI.5. Soient $u, v \in \mathbb{K}^N N$ telles que $u_n \sim v_n$ et $v_n \to \ell \in \overline{\mathbb{R}}$. Alors $(u_n)_n$ converge également vers ℓ. Ce résultat se généralise au cas d'une "limite" infinie.

✖ **Attention :** la réciproque est **fausse** : deux suites de même limites ne sont pas nécessairement équivalentes. Il suffit de considérer $(n)_n$ et $(n^2)_n$ pour s'en convaincre (j'espère !).

Démonstration. Par définition d'équivalence, il existe une suite $w \in \mathbb{K}^{\mathbb{N}}$ convergeant vers 1 telle que $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = w_n v_n$. De fait, on a bien la convergence de u vers ℓ. □

Exemple XVI.3. $\dfrac{n^2+n}{n^3+n^2} \sim \dfrac{1}{n}$ donc $\dfrac{n^2+n}{n^3+n^2} \to 0$.

Proposition XVI.6. Soient $u, v \in \mathbb{K}^{\mathbb{N}}$ telles que $u_n \sim v_n$. Alors, si u est de signe constant à partir d'un certain rang, v vérifie la même propriété (avec le même signe donc).

Démonstration. Par définition d'équivalence, il existe une suite $w \in \mathbb{K}^{\mathbb{N}}$ convergeant vers 1 telle que $\exists\, N \in \mathbb{N},\ \forall n \geq N,\ u_n = w_n v_n$. La suite w sera, par définition de limite, strictement positive à partir d'un certain rang, d'où le résultat. □

2.⁻ Comparaison des fonctions

a) Négligeabilité, domination

Nous généralisons dans ce paragraphe les notions évoquées dans la partie précédente, tout en prenant garde qu'une fonction (contrairement à une suite) n'a pas la politesse de n'avoir qu'une seule limite possible.

Définition XVI.2. Soit $a \in \overline{\mathbb{R}}$ et soient f, g deux fonctions définies au voisinage de a (sauf peut–être en a). On dira que :
- f est **dominée** par g au voisinage de a si il existe une fonction h **bornée** au voisinage de a telle que $f = hg$;
- f est **négligeable** devant g au voisinage de a si il existe une fonction h **de limite nulle en** a telle que $f = hg$ au voisinage de a.

Exemple XVI.4. $3x^2 - 1 = o_\infty(x^2)$.

Remarque XVI.6. On peut obtenir, si g ne s'annule pas au voisinage de a, une caractérisation par le quotient analogue à celle vue pour les suites. De même, les propriétés concernant leur transitivité et compatibilité avec la somme et le produit vues précédemment se généralisent.

✖ **Attention :** le point où l'étude est effectuée est déterminant ; par exemple $x^2 = o_\infty(x^3)$ mais $x^3 = o_0(x^2)$.

b) Équivalence de fonctions

Définition XVI.3. Soient f, g deux fonctions définies au voisinage de $a \in \overline{\mathbb{R}}$ (sauf peut–être en a). On dit que f est **équivalente** à g en a si il existe une fonction λ définie au voisinage de a telle que $\lambda(x) \underset{x\to a}{\longrightarrow} 1$ et $f = \lambda g$.

Notation. $f(x) \sim_a g(x)$.

➠ **Exemple XVI.5.** $x^3 + x^2 + x + 1 \sim_\infty x^3$.

✌ **Remarque XVI.7.**

— Si g ne s'annule pas au voisinage de a, on a la caractérisation suivante :

$$f(x) \sim_a g(x) \quad \Longleftrightarrow \quad \frac{f(x)}{g(x)} \underset{x\to a}{\longrightarrow} 1\,.$$

— Si f admet une limite $\ell \in \mathbb{R}^*$ en a, alors $f(x) \sim_a \ell$.

— Comme pour les suites, l'équivalence de fonctions en un point est une relation . . . d'équivalence.

Les propositions qui suivent auront sans doute un goût de déjà vu. Cela n'est pas une coïncidence, et leurs démonstrations sont analogues à celles vues dans le paragraphe traitant de la comparaison des suites.

Proposition XVI.7. Soient f, g deux fonctions définies au voisinage de $a \in \overline{\mathbb{R}}$ telles que $f(x) \sim_a g(x)$. Alors :

(i) si $g(x) \underset{x\to a}{\longrightarrow} \ell \in \overline{\mathbb{R}}$, alors $f(x) \underset{x\to a}{\longrightarrow} \ell$;

(ii) si le signe de g est constant au voisinage de a, f vérifie la même propriété (avec un signe identique).

Proposition XVI.8. Soient f, g deux fonctions définies au voisinage de $a \in \overline{\mathbb{R}}$. Alors :

(i) $f(x) \sim_a g(x) \Leftrightarrow f(x) - g(x) = o_a(g(x))$, *i.e* $f(x) = g(x) + o_a(g(x))$;

(ii) l'équivalence des fonctions en un point est compatible avec le produit, le quotient et le passage à une puissance entière strictement positive.

✖ **Attention :** on ne somme **toujours pas** les équivalents. Genre, jamais.

c) Zoologie zérocentrique des équivalents usuels

Proposition XVI.9. On dispose des équivalents suivants en 0 :

(i) $e^x - 1 \sim_0 x$;

(ii) $\ln(1+x) \sim_0 x$;

(iii) $(1+x)^\alpha - 1 \sim_0 \alpha x$, pour $\alpha \in \mathbb{R}^*$;

(iv) $\sin(x) \sim_0 x$;

(v) $\tan(x) \sim_0 x$;

(vi) $1 - \cos(x) \sim_0 \dfrac{x^2}{2}$;

(vii) $\operatorname{sh}(x) \sim_0 x$;

(viii) $\operatorname{ch}(x) - 1 \sim_0 \dfrac{x^2}{2}$;

(ix) $\operatorname{th}(x) \sim_0 x$;

(x) $\arcsin(x) \sim_0 x$;

(xi) $\arctan(x) \sim_0 x$.

Démonstration. Cela découle dans presque tous les cas de la limite du taux d'accroissement des fonctions étudiées en 0 ; par exemple

$$\frac{e^x - 1}{x} = \frac{e^x - e^0}{x - 0} \xrightarrow[x \to 0]{} 1 = e^0 .$$

Le point (vii) sera traité ultérieurement, et le point (v) provient de l'égalité, pour $x \neq 0$:

$$\begin{aligned} 1 - \cos(x) &= 2\sin(x/2)^2 \\ &\sim_0 22\frac{x^2}{4} \\ &= \frac{x^2}{2} . \end{aligned}$$

□

d) Croissances comparées

Comme dans le cas des suites, nous pouvons reformuler les croissances comparées dans le langage des équivalents.

Proposition XVI.10. Soient $\alpha, \beta, \gamma \in \mathbb{R}_+^*$. Alors :

(i) $\ln(x)^\beta = o_\infty(x^\alpha)$;

(ii) $x^\alpha = o_\infty(e^{\gamma x})$.

À ce stade, ceux d'entre vous qui ne sont pas recroquevillés en position fœtale se demandent probablement quel peut bien être l'usage de toutes ces comparaisons.

En premier lieu et dans un premier temps, elles nous serviront à lever les indéterminations qui (facheusement) se produisent parfois lors de l'étude de limites.

Exemple XVI.6.

— $\frac{e^x}{x^3} \xrightarrow[x\to\infty]{} 0$ car $x^3 = o_\infty(e^x)$.

— Soit $f : x \mapsto x^2 \ln(x)^7$. Pour étudier la limite de f en 0^+, nous pouvons poser, pour $x > 0$, $t = \frac{1}{x}$ et remarquer qu'alors :

$$\begin{aligned} f(x) &= \frac{1}{t^2} \ln\left(\frac{1}{t}\right)^7 \\ &= -\frac{\ln(t)^7}{t^2}. \end{aligned}$$

Or, $\ln(t)^7 = o_\infty(t^2)$ donc $f(x) \xrightarrow[t\to\infty]{} 0$. Ceci nous permet de conclure que :

$$f(x) \xrightarrow[x\to 0^+]{} 0\,.$$

— De même, le "changement de variable" $t = -x$ nous permet de vérifier que $x^3 e^x \xrightarrow[x\to-\infty]{} 0$.

3.⁻ Développements limités

a) Quoi ?

Définition XVI.4. Soit $n \in \mathbb{N}$, $a \in \mathbb{R}$ et soit f une fonction définie au voisinage de a. On dira que f admet un **développement limité d'ordre n en a** (DL_n(a)) si il existe une famille $a_0, \dots, a_n$ d'éléments de $\mathbb{K}$ telle que :

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

au voisinage de a.

Vocabulaire. La somme apparaissant dans le membre de droite de cette égalité est appelée **partie régulière** du DL. Le reste est appelé **partie négligeable** de celui–ci.

Remarque XVI.8.

— Une fonction admettant un DL_n(a) peut donc être approximée au voisinage de a par une fonction polynomiale de degré n.

— On démontre aisément que si f est (im)paire et admet un $\mathrm{DL}_n(0)$, celui–ci ne peut contenir que des puissances (im)paires.

— Quitte à poser $h = x-a$, on peut écrire tout DL_n(a) sous **forme normalisée** :

$$f(a+h) = \sum_{k=0}^{n} a_k h^k + o_0(h^n)$$

soit, si $p = \min\{k \in \mathbb{N} \mid a_k \neq 0\}$:

$$f(a+h) = h^p(a_p + a_{p+1}h + \ldots + a_n h^{n-p} + o_0(h^{n-p}))\ .$$

Ceci nous permet de démontrer que $f(a+h) \sim_0 a_p h^p$ et, par conséquent, d'étudier le signe de f au voisinage de a.

Exemple XVI.7.

— $e^x = 1 + x + o_0(x)$; notons que tous les équivalents usuels vus précédemment nous livrent des DL d'ordre 1 ou 2.

— $\cos(x) = 1 - \dfrac{x^2}{2} + o_0(x^2)$.

— Un exemple fondamental est celui des sommes géométriques ; si $x \in]-1,1[$ et $n \geq 0$, on a :

$$\begin{aligned}\sum_{k=0}^{n} x^k &= \frac{1-x^{n+1}}{1-x} \\ &= \frac{1}{1-x} + \frac{x^{n+1}}{x-1} \\ &= \frac{1}{1-x} + o_0(x^n)\ .\end{aligned}$$

— Pour obtenir un DL d'un polynôme, il suffit de tronquer le degré à l'ordre voulu.

Proposition XVI.11. Si une fonction f admet un $\mathrm{DL}_n(\mathrm{a})$ pour $a \in \mathbb{R}$, alors ce dernier est unique.

Démonstration. Supposons qu'il existe deux familles $a_0, \ldots a_n$ et $b_0, \ldots, b_n$ de réels telles que :

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

Alors, par soustraction, on a :

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

Si il existe $k \in [\![0,n]\!]$ tel que $a_k \neq b_k$, on peut poser $k_0 = \min\{k \in [\![0,n]\!] \mid a_k \neq b_k\}$ et vérifier que

$$\frac{\sum_{k=0}^{n} (a_k - b_k)(x-a)^k}{(a_{k_0} - b_{k_0})(x-a)^{k_0}} \xrightarrow[x \to a]{} 1$$

i.e

$$\sum_{k=0}^{n} (a_k - b_k)(x-a)^k \sim_a (a_{k_0} - b_{k_0})(x-a)^{k_0}\ .$$

Ceci contredit le fait que cette somme soit un $o_a((x-a)^n)$ car $k_0 \leq n$, d'où le résultat. $\square$

Proposition XVI.12. Soit f une fonction définie au voisinage de $a \in \mathbb{R}$. Alors :

(i) f admet un $\mathrm{DL}_0(a) \Leftrightarrow f$ est continue en a ;

(ii) f admet un $\mathrm{DL}_1(a) \Leftrightarrow f$ est dérivable en a.

Démonstration. Le (i) découle de la définition de continuité et le (ii) a été démontré dans le chapitre XII (proposition XII.1). □

✖ **Attention :** ce résultat ne se généralise **pas** aux ordres supérieurs.

b) Formule de Taylor–Young

Ce résultat découle de la formule établie par Brook Taylor (anglais, 1685—1731) pour les polynômes (et vue au chapitre XV) et affinée par William Henry Young (anglais, 1863—1942). La démonstration de ce résultat est hors–programme. Il peut toutefois être vu comme une conséquence de la proposition XXI.19.

Théorème XVI.13 (Taylor–Young)**.**
Soit f une fonction de classe $\mathscr{C}^n$ au voisinage de $a \in \mathbb{R}$. Alors admet un DL_n(a) donné par la formule :

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

✌ **Remarque XVI.9.** Ceci entraîne qu'il y a existence et uncité du DL_npour les fonctions de classe $\mathscr{C}^n$.

◇ Développements limités usuels en 0

La formule de Taylor–Young permet de démontrer les DL suivants en 0, qui nous seront utiles par la suite.

$$\begin{aligned}\frac{1}{1-x} &= \sum_{k=0}^{n} x^k + \mathrm{o}_0(x^n)\\ &= 1 + x + x^2 + \ldots + x^n + \mathrm{o}_0(x^n)\end{aligned}$$

$$\begin{aligned}\ln(1+x) &= \sum_{k=1}^{n} (-1)^{k+1}\frac{x^k}{k} + \mathrm{o}_0(x^n)\\ &= x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots + (-1)^{n+1}\frac{x^n}{n} + \mathrm{o}_0(x^n)\end{aligned}$$

$$\begin{aligned}(1+x)^\alpha &= 1+\sum_{k=1}^{n}\left(\prod_{i=0}^{k-1}(\alpha-i)\right)\frac{x^k}{k!}+\mathrm{o}_0(x^n)\\ &= 1+\alpha x+\frac{\alpha(\alpha-1)}{2}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{6}x^3+\dots\\ &\quad \dots+\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n+\mathrm{o}_0(x^n)\end{aligned}$$

$$\begin{aligned}e^x &= \sum_{k=0}^{n}\frac{x^k}{k!}+\mathrm{o}_0(x^n)\\ &= 1+x+\frac{x^2}{2}+\dots+\frac{x^n}{n!}+\mathrm{o}_0(x^n)\end{aligned}$$

$$\begin{aligned}\cos(x) &= \sum_{k=0}^{n}(-1)^k\frac{x^{2k}}{(2k)!}+\mathrm{o}_0(x^{2n+1})\\ &= 1-\frac{x^2}{2}+\frac{x^4}{24}+\dots+(-1)^n\frac{x^{2n}}{(2n)!}+\mathrm{o}_0(x^{2n+1})\end{aligned}$$

$$\begin{aligned}\sin(x) &= \sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{(2k+1)!}+\mathrm{o}_0(x^{2n+2})\\ &= x-\frac{x^3}{6}+\frac{x^5}{120}+\dots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\mathrm{o}_0(x^{2n+2})\end{aligned}$$

$$\begin{aligned}\mathrm{ch}(x) &= \sum_{k=0}^{n}\frac{x^{2k}}{(2k)!}+\mathrm{o}_0(x^{2n+1})\\ &= 1+\frac{x^2}{2}+\frac{x^4}{24}+\dots+\frac{x^{2n}}{(2n)!}+\mathrm{o}_0(x^{2n+1})\end{aligned}$$

$$\begin{aligned}\mathrm{sh}(x) &= \sum_{k=0}^{n}\frac{x^{2k+1}}{(2k+1)!}+\mathrm{o}_0(x^{2n+2})\\ &= x+\frac{x^3}{6}+\frac{x^5}{120}+\dots+\frac{x^{2n+1}}{(2n+1)!}+\mathrm{o}_0(x^{2n+2})\end{aligned}$$

$$\begin{aligned}\arctan(x) &= \sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{2k+1}+\mathrm{o}_0(x^{2n+2})\\ &= x-\frac{x^3}{3}+\dots+(-1)^n\frac{x^{2n+1}}{2n+1}+\mathrm{o}_0(x^{2n+2})\end{aligned}$$

$$\tan(x) = x+\frac{x^3}{3}+\mathrm{o}_0(x^3)$$

c) Pourquoi ?

En pratique, les DL ont de nombreux usages, outre celui de générer de charmantes sessions de calcul. D'une façon générale, ils permettes d'étudier en finesse le comportement d'une fonction au voisinage d'un point donné.

◇ Limites

Une fonction est équivalente à la partie régulière de son DL : cela permet souvent de lever des indéterminations. Par exemple :

$$\frac{\sin(x)}{x} \sim_0 \frac{x}{x} = 1 \underset{x\to 0}{\longrightarrow} 1\,.$$

De même, certains taux d'accroissements voient leur vie considérablement simplifiée (d'aucun diraient limitée) par ce type d'étude.

$$\begin{aligned}\forall x \neq 0,\ \frac{\frac{\sin(x)}{x} - 1}{x - 0} &= \frac{\sin(x)}{x^2} - \frac{1}{x}\\ &= \frac{x - \frac{x^3}{6} + o_0(x^3)}{x^2} - \frac{1}{x}\\ &= -\frac{x}{6} + o_0(x)\\ &\underset{x\to 0}{\longrightarrow} 0\end{aligned}$$

et donc le sinus cardinal (*cf.* chapitre XII) est dérivable en 0 de dérivée nulle. En réalité, on peut itérer ce procédé et démontrer que cette fonction est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$.

◇ Extrema locaux

Nous avons vu dans le chapitre XII que toute fonction f dérivable admettant un extremum local en un point $a \in \mathbb{R}$ vérifiait $f'(a) = 0$ (on dit que a est un **point critique** pour f). La notion de développement limité nous permet d'affiner quelque peu cette étude et d'en obtenir une réciproque partielle via la proposition suivante.

Proposition XVI.14. Soit f une fonction admettant un développement limité d'ordre 2 en un point $a \in \mathbb{R}$ et vérifiant que $f'(a) = 0$. On a de fait, quitte à poser $\lambda = \dfrac{f''(a)}{2}$:

$$f(x) = f(a) + \lambda(x-a)^2 + o((x-a)^2)\,.$$

Alors :

(i) si $\lambda > 0$, f admet un minimum local en a ;

(ii) si $\lambda < 0$, f admet un maximum local en a.

Démonstration. Immédiat : au voisinage de a, $f(x) - f(a)$ est du signe de λ si ce dernier est non nul. □

➠ **Exemple XVI.8.** Ceci nous permet de vérifier que cos admet un maximul local en 0.

✌ **Remarque XVI.10.** Dans le cas où $\lambda = 0$, on ne peut conclure. Étudier le cas des fonctions $x \mapsto x^3$ et $x \mapsto x^2$ en 0.

d) Opérations sur les développements limités

Dans ce paragraphe, nous nous attachons à donner une série de "recettes" permettant de calculer efficacement avec les développements limités. L'objectif est ici presque purement opérationnel : nous ne nous attarderons pas plus que nécessaire sur le caractère général des résultats énoncé.

◇ Troncature

Pour le moment tout va bien : si une fonction admet un $\mathrm{DL}_n(\mathrm{a})$, on peut lui trouver des DL à des ordres inférieurs quitte à tronquer la partie régulière.

◇ Combinaison linéaire

Si $a \in \mathbb{R}$ et f et g sont deux fonctions admettant des $\mathrm{DL}_n(\mathrm{a})$, alors pour tout $\lambda \in \mathbb{R}$, $f + \lambda g$ admet un $\mathrm{DL}_n(\mathrm{a})$ obtenu en combinant linéairement les parties régulières de ceux de f et g.

➠ **Exemple XVI.9.** On retrouve de cette façon les DL de ch et sh à partir de celui de exp.

◇ Produit

Soit $a \in \mathbb{R}$ et soient f, g admettant les $\mathrm{DL}_n(\mathrm{a})$ suivants :

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

Alors, par relation sur les "o" et produit de polynômes, fg admet le $\mathrm{DL}_n(\mathrm{a})$ suivant :

$$f(x)g(x) = \sum_{k=0}^{n} \left(\sum_{j=0}^{k} a_j b_{k-j} \right) (x-a)^k + o_a((x-a)^n) .$$

En pratique, il s'agit ici de faire le produit des parties régulières en ne conservant que les puissances inférieures ou égales à n.

➠ **Exemple XVI.10.** Cherchons à obtenir un $\mathrm{DL}_3(0)$ de $x \mapsto e^x \sin(x)$. Pour cela, notons dans un premier temps que :

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + o_0(x^3)$$

et

$$\sin(x) = x - \frac{x^3}{6} + o_0(x^3)$$

et que, pour tout $x \in \mathbb{R}$:

$$\begin{aligned}(1 + x + \frac{x^2}{2} + \frac{x^3}{6})(x - \frac{x^3}{6}) &= x - \frac{x^3}{6} + x^2 + \frac{x^3}{2} + o_0(x^3) \\ &= x + x^2 + \frac{x^3}{3} + o_0(x^3) .\end{aligned}$$

In fine, on a donc :

$$e^x \sin(x) = x + x^2 + \frac{x^3}{3} + o_0(x^3)\,.$$

◇ **Composition**

Aucun résultat général n'est au programme : il s'agit ici de chercher à faire simple et efficace. Si f et g admettent un $\mathrm{DL}_n(0)$ (quitte à se ramener en ce point par un changement de variable) **et** que $g(0) = 0$ (ou *a minima* $g(x) \xrightarrow[x\to 0]{} 0$, mais je chipote) alors la composée $f \circ g$ admet un $\mathrm{DL}_n(0)$ (pour peu qu'elle soit bien définie). On trouve la partie régulière de ce dernier en priant très fort.

➠ **Exemple XVI.11.** Au voisinage de 0, on a :

$$\ln(\cos(x)) = \ln(1 + \underbrace{\cos(x) - 1}_{\text{vaut } 0 \text{ en } 0})\,.$$

Or :

$$\ln(1 + u) = u - \frac{u^2}{2} + o_0(u^2)$$

et

$$\cos(x) - 1 = -\frac{x^2}{2} + o_0(x^2)$$

ce qui permet de conclure que :

$$\ln(\cos(x)) = -\frac{x^2}{2} + o_0(x^2)$$

car $(\cos(x) - 1)^2 = o_0(x^2)$.

◇ **Inverse**

L'idée est ici de composer le DL par celui de $x \mapsto \dfrac{1}{1-x}$. Une fois de plus, un peu de courage et (éventuellement) un sac à vomi sont des atouts précieux.

➠ **Exemple XVI.12.** Développons $x \mapsto \dfrac{1}{\cos(x)}$ à l'ordre 4 en 0. Faites comme on vous dit et il n'y aura pas de blessé. Commençons par remarquer que :

$$\frac{1}{\cos(x)} = \frac{1}{1 - \underbrace{(1 - \cos(x))}_{\text{vaut } 0 \text{ en } 0}}\,.$$

On sait de plus que :

$$\frac{1}{1-u} = 1 + u + u^2 + u^3 + u^4 + o_0(u^4)\,.$$

En posant, au voisinage de 0, $u = 1 - \cos(x)$ (oui, je sais...), on obtient :

$$u = \frac{x^2}{2} - \frac{x^4}{24} + o_0(x^4)\,;$$

$$u^2 = \frac{x^4}{4} + o_0(x^4)$$

et

$$u^3,\ u^4 = o_0(x^4)\ .$$

Au final, on obtient :

$$\frac{1}{\cos(x)} = 1 + \frac{x^2}{2} + \frac{5x^4}{24} + o_0(x^4)\ .$$

Évidemment, pour obtenir le DL d'un quotient, nous passons par le produit de l'inverse. Le lecteur audacieux pourra utiliser l'exemple *supra* pour retrouver le DL de tan en 0.

◇ Primitives

Proposition XVI.15. Soit $a \in \mathbb{R}$ et soit f une fonction admettant un DL_n(a) donné par

$$f(x) = \sum_{k=0}^{n} a_k (x-a)^k + o_0((x-a)^n)\ .$$

Alors toute primitive F de f admet un $\mathrm{DL}_{n+1}(a)$ donné par la formule :

$$F(x) = F(a) + \sum_{k=0}^{n} \frac{a_k}{k+1}(x-a)^{k+1} + o_0((x-a)^{n+1})\ .$$

Démonstration. Immédiat. □

✌ **Remarque XVI.11.** Ceci nous permet de retrouver très rapidement les DL de arctan et $x \mapsto \ln(1+x)$.

✍ **Exercice XVI.1.** Déterminer un DL d'ordre 5 de arcsin en 0.

➥ **Correction :** *On sait que, pour tout* $x \in]-1,1[$ *:*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$$

et que :

$$\begin{aligned}\frac{1}{\sqrt{1-x^2}} &= (1+(-x^2))^{-1/2}\\ &= 1 + \frac{x^2}{2} + \frac{3x^4}{8} + o_0(x^4)\ .\end{aligned}$$

On en déduit que :

$$\arcsin(x) = x + \frac{x^3}{6} + \frac{3x^5}{40} + o_0(x^5)\ .$$

4.– Développements asymptotiques

a) Quoi maintenant ?

Jusque ici, nous avons fait des développements limités, à savoir des approximations de fonction en un point **réel** par un **polynôme**. Un peu contraignant n'est–il pas ? L'objectif de ce paragraphe est d'en finir avec ces limitations et de déverrouiller notre plein potentiel calculatoire. Ou quelque chose de ce genre. Nous nous contenterons de donner quelques exemples afin de nourrir la réflexion (et/ou les cauchemars) de notre lecteur.

➠ **Exemple XVI.13.** Développons la fonction $f : x \mapsto x^x$ en 0. Ne mentez pas, je sais que vous en avez toujours rêvé. On a, au voisinage de 0 :

$$\begin{aligned} f(x) &= e^{x\ln(x)} \\ &= 1 + x\ln(x) + \frac{x^2\ln(x)^2}{2} + o_0(x^2\ln(x)^2) \,. \end{aligned}$$

Notons que la composée de DL ci–dessus est valable car $x\ln(x) \xrightarrow[x\to 0]{} 0$. Ce résultat ne constitue pas un DL *stricto sensu*, mais nous apporte des informations assez précises sur le comportement asymptotique de f, ce qui est fort sympathique au demeurant : on parle de **développement asymptotique (DA) d'ordre** 2 **en** $x\ln(x)$ **au voisinage de** 0.

✍ **Exercice XVI.2.** Déterminer la limite en $+\infty$ de $x \mapsto \sqrt{x^2+1} - \sqrt{x^2-1}$.

➥ **Correction :** *Posons, pour $x > 0$, $t = \dfrac{1}{x}$. Alors on a :*

$$\begin{aligned} \sqrt{x^2-1} &= \sqrt{\frac{1}{t^2} - 1} \\ &= \sqrt{\frac{1-t^2}{t^2}} \\ &= \frac{1}{t}\sqrt{1-t^2} \end{aligned}$$

et, de façon analogue :

$$\sqrt{x^2+1} = \frac{1}{t}\sqrt{1+t^2} \,.$$

En faisant en DL (en t) à l'ordre 4 en 0, nous obtenons que :

$$\sqrt{1 \pm t^2} = 1 \pm \frac{t^2}{2} - \frac{t^4}{8} + o_0(t^4)$$

ce qui entraîne que :

$$\begin{aligned} \frac{1}{t}\sqrt{1+t^2} - \frac{1}{t}\sqrt{1-t^2} &= \frac{1}{t} + \frac{t}{2} - \frac{t^3}{8} - \left(\frac{1}{t} - \frac{t}{2} - \frac{t^3}{8}\right) + o_0(t^3) \\ &= t + o_0(t^3) \,. \end{aligned}$$

En conclusion, nous pouvons écrire (en repassant à la variable x) :

$$\sqrt{x^2+1} - \sqrt{x^2-1} = \frac{1}{x} + o_\infty\left(\frac{1}{x^3}\right) \xrightarrow[x\to\infty]{} 0 \,.$$

Notons qu'un DL à l'ordre 1 *suffisait (peut–être...). Mais bon, quand on aime, on ne compte pas.*

➠ **Exemple XVI.14.** Souvenons nous que, pour tout $x > 0$:

$$\arctan(x) = \frac{\pi}{2} - \arctan\left(\frac{1}{x}\right)$$

et donc, on peut établir que :

$$\arctan(x) = \frac{\pi}{2} - \frac{1}{x} + \frac{1}{3x^3} + o_\infty\left(\frac{1}{x^3}\right) .$$

b) Formule de Stirling

Achevons ce chapitre (et le lecteur) en énonçant la formule suivante, du à Abraham de Moivre (français, 1667—1754) et James Stirling (écossais, 1692—1770). Le premier démontra la formule sans calculer la constante apparaissant dans l'équivalent, injustice réparée par le second.

Théorème XVI.16 (Stirling)**.**

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n} .$$

Démonstration. Admis. □

✍ **Exercice XVI.3.** Déterminer la limite de $\left(\dfrac{e^n}{n!}\right)_n$.

Chapitre XVII

Fractions rationnelles

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.‾ Corps des fractions rationnelles

a) Notion de fraction rationnelle

Nous ne rentrerons pas cette fois–ci dans les détails ; les techniques nous manquent pour définir proprement tout ceci.

Théorème XVII.1.
Il existe un corps $\mathbb{K}(X)$, appelé **corps des fractions rationnelles sur** $\mathbb{K}$ tel que :
- $\mathbb{K}[X] \subset \mathbb{K}(X)$;
- tout corps contenant $\mathbb{K}[X]$ contient également $\mathbb{K}(X)$.

Un tel corps est de plus unique à isomorphisme près.

✌ **Remarque XVII.1.** Le corps $\mathbb{K}(X)$ contient donc :
- les polynômes ;
- leurs inverses (sauf pour 0, évidemment), notés $\dfrac{1}{P}$ pour $P \in \mathbb{K}[X] \setminus \{0\}$.

Il en découle par minimalité que les éléments de $\mathbb{K}[X]$ sont tous de la forme $P \times \dfrac{1}{Q}$ avec $P, Q \in \mathbb{K}[X]$ et $Q \neq 0$. Nous utiliserons la notation fractionnelle $\dfrac{P}{Q}$ pour de tels objets.

➠ **Exemple XVII.1.** $\dfrac{1}{X}, \dfrac{X^2+4X-1}{X-1} \in \mathbb{K}(X)$.

Opérations sur les fractions rationnelles
Si $P, Q, R, S \in \mathbb{K}[X]$ sont tels que $QS \neq 0$, alors on a :

(i) $\dfrac{P}{Q} = \dfrac{R}{S} \Leftrightarrow PS = QR$;

(ii) $\dfrac{P}{Q} + \dfrac{R}{S} = \dfrac{PS+QR}{QS}$;

(iii) $\dfrac{P}{Q} \times \dfrac{R}{S}$.

✖ **Attention :** la représentation d'une fraction rationnelle n'est pas unique :

$$\frac{X}{1} = \frac{X^2}{X} = X\,.$$

Théorème XVII.2 (Forme irréductible d'une fraction rationnelle)**.**
Soit $F \in \mathbb{K}(X)$. Alors il existe un unique (à constante multiplicative commune près) couple $(P, Q) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$ tel que :
(i) $P \wedge Q = 1$;
(ii) Q soit unitaire ;
(iii) $F = \dfrac{P}{Q}$.

Démonstration. Il suffit de simplifier n'importe quelle écriture de la fraction via les diviseurs communs du numérateur et du dénominateur. □

b) Degré

Proposition/définition XVII.1. Soit $F = \dfrac{P}{Q} \in \mathbb{K}(X)$. Alors la quantité $\deg(P) - \deg(Q) \in \mathbb{Z} \cup \{-\infty\}$ ne dépend pas de la représentation choisie et est appelée **degré** de F.

Notation. $\deg(F)$

Démonstration. Si F admet une autre écriture $F = \dfrac{R}{S}$, alors $PS = QR$ et donc $\deg(P) + \deg(S) = \deg(Q) + \deg(R)$, ce qui entraîne que :

$$\deg(P) - \deg(Q) = \deg(R) - \deg(S)\,.$$

□

➠ **Exemple XVII.2.** $\deg\left(\dfrac{X^2+1}{X^3}\right) = -1$.

✖ **Attention :** une fraction rationnelle de degré nulle peut ne pas être constante : considérer $\dfrac{X}{X+2}$.

✌ **Remarque XVII.2.**
— *deg*$(F) = -\infty \Leftrightarrow F = 0$;
— $\deg \leq 0 \Leftrightarrow \deg(P) \geq \deg(Q)$.

Proposition XVII.3. Soient $F, G \in \mathbb{K}(X)$. Alors :
(i) $\deg(F+G) \leq \max(\deg(F), \deg(G))$, avec égalité lorsque $\deg(F) \neq \deg(G)$;
(ii) $\deg(FG) = \deg(F) + \deg(G)$.

Démonstration. Posons, pour fixer les idées, $F = \dfrac{P}{Q}$ et $G = \dfrac{R}{S}$.

(i) On sait que $F + G = \dfrac{PS+QR}{QS}$ et donc :

$$\begin{aligned}\deg(F+G) &= \deg(PS+QR) - \deg(QS) \\ &\leq \max(\deg(PS), \deg(QR)) - (\deg(Q) + \deg(S)) \text{ par degrés polynomiaux} \\ &= \max(\deg(P) + \deg(S), \deg(Q) + \deg(R)) - (\deg(Q) + \deg(S)) \\ &= \max(\deg(P) - \deg(Q), \deg(R) - \deg(S)) \\ &= \max(\deg(F), \deg(G))\end{aligned}$$

avec égalité lorsque les degrés sont différents par résultat sur les degrés de polynômes.

(ii) Procéder de façon similaire en développant les degrés.

□

c) Racines, pôles

Définition XVII.2. Soit $F = \dfrac{P}{Q} \in \mathbb{K}(X)$ une fraction rationnelle **irréductible**.
On appelle :
— **racines** de F les racines de P sur $\mathbb{K}$;
— **pôles** de F les racines de Q sur $\mathbb{K}$.
La multiplicité d'une racine (resp. d'un pôle) de F est définie comme étant sa multiplicité en tant que racine de P (resp. de Q).

Notation. $\mathrm{Rac}_{\mathbb{K}}(F)$, $\mathrm{Pol}_{\mathbb{K}}(F)$.

✖ **Attention :** l'irréductibilité est ici essentielle : 0 n'est pas un pôle (ni une racine) de $\dfrac{X(X+1)}{X}$.

➠ **Exemple XVII.3.** Les pôles de $\dfrac{X}{X^2+1}$ sur $\mathbb{C}$ sont i et $-i$, et elle n'admet aucun pôle sur $\mathbb{R}$. Son unique racine sur les deux corps est 0.

✌ **Remarque XVII.3.**
— Toute fraction rationnelle admet un nombre fini de pôles.
— Toute fraction rationnelle **non nulle** admet un nombre fini de racines.

— De façon analogue aux fonctions polynomiales, il est possible étant donné $F = \dfrac{P}{Q} \in \mathbb{K}(X)$ de définir une fonction

$$\begin{aligned} f : \mathbb{K} \setminus \mathrm{Pol}_{\mathbb{K}}(F) &\longrightarrow \mathbb{K} \\ x &\mapsto \frac{P(x)}{Q(x)} . \end{aligned}$$

2.‾ Éléments simples

a) C'est quoi ?

Définition XVII.3. On appelle **élément simple** sur $\mathbb{K}$ toute fraction rationnelle du type $\dfrac{P}{Q^n} \in \mathbb{K}(X)$ avec :
- Q irréductible sur $\mathbb{K}$;
- $n \in \mathbb{N}^*$;
- $\deg(P) < \deg(Q)$.

✌ **Remarque XVII.4.** L'étude menée au chapitre XV nous permet de classifier les éléments simples sur $\mathbb{R}$ et $\mathbb{C}$.
- Sur $\mathbb{C}$, ce sont les fractions de la forme :
$$\frac{\alpha}{(X-\beta)^n}$$
avec $\alpha, \beta \in \mathbb{C}$ et $n \geq 1$.
- Sur $\mathbb{R}$ ce sont les fractions de la forme :
$$\frac{\alpha}{(X-\beta)^n}$$
avec $\alpha, \beta \in \mathbb{R}$ et $n \geq 1$ ainsi que celles de la forme :
$$\frac{aX+b}{(X^2+pX+q)^n}$$
avec $a, b, p, q \in \mathbb{R}$ et $n \geq 1$.

b) Partie entière d'une fraction rationnelle

Proposition/définition XVII.4. Soit $F \in \mathbb{K}(X)$. Alors il existe un unique polynôme $E \in \mathbb{K}[X]$ et une unique fraction rationnelle $G \in \mathbb{K}(X)$ tel que :
(i) $\deg(G) < 0$;
(ii) $F = E + G$.
Le polynôme E est appelé **partie entière** de F.

Notation. $E(F)$

Démonstration. **Existence.** Posons $F = \dfrac{A}{B}$; alors par division euclidienne on a l'existence d'un couple $(Q, R) \in \mathbb{K}[X]^2$ tel que $A = BQ+R$ et $\deg(R) < \deg(B)$. Ainsi, $F = Q + \dfrac{R}{B}$ avec $\deg\left(\dfrac{R}{B}\right) < 0$.

Unicité. Si il existe deux couples (E, G) et (E', G') vérifiant les conditions voulues, alors $E_1 - E_2 = G_2 - G_1$ ce qui n'est possible (car $deg(G_1 - G_2) < 0$ et $E_1, E_2 \in \mathbb{K}[X]$) que si les deux différences sont nulles.

□

Exemple XVII.4.

— $\dfrac{X+1}{X} = 1 + \dfrac{1}{X}$ donc $E\left(\dfrac{X+1}{X}\right) = 1$;

— $\dfrac{X}{X^2+1} = 0 + \dfrac{X}{X^2+1}$ donc $E\left(\dfrac{X}{X^2+1}\right) = 0$;

— Par division euclidienne, $2X^2 + 1 = (2X+2)(X+1) - 3$ donc

$$E\left(\frac{2X^2+1}{X+1}\right) = 2(X+1)\,.$$

c) Décomposition en éléments simples

◇ Le cas complexe

Théorème XVII.4 (Décomposition en éléments simples (cas complexe))**.**
Soit $F \in \mathbb{C}(X)$ admettant n exactement n pôles $a_1, \ldots, a_n$ de multiplicités respectives $k_1, \ldots, k_n$. Alors il existe une unique famille $\alpha_{1,1}, \ldots, \alpha_{k_1,1}, \alpha_{1,2}, \ldots, \alpha_{k_n,n}$ de nombres complexes telle que :

$$F = E(F) + \sum_{j=1}^{n} \sum_{\ell=1}^{k_j} \frac{\alpha_{\ell,j}}{(X-a_j)^\ell}\,.$$

Démonstration. Admis. □

Et en pratique ? La première étape est dégager la partie entière, ce qui nous permet de partir du principe que notre fraction $F \in \mathbb{C}(X)$ est de degré strictement négatif.

Si les pôles sont simples : F est donc de la forme $\dfrac{A}{\prod_{i=1}^{n}(X-a_i)}$ avec $A \in \mathbb{K}[X]$.

On fixe alors $j_0 \in [\![1, n]\!]$ et on remarque que $G = (X - a_{j_0})F$ n'a plus de pôle en a_{j_0} et

$$G(a_{j_0}) = \frac{A(a_{j_0})}{\prod_{j \neq j_0}(X-a_j)}\,.$$

Par unicité dans la DES, on obtient que si

$$F = \sum_{j=0} \frac{\alpha_j}{(X - a_j)}$$

alors

$$\alpha_{jj_0} = \frac{A(a_{j_0})}{\prod\limits_{j \neq j_0} (X - a_j)} .$$

➠ **Exemple XVII.5.** La DES de $\dfrac{X}{(X-1)(X-2)}$ est

$$\frac{X}{(X-1)(X-2)} = \frac{\lambda}{X-1} + \frac{\mu}{X-2} ,$$

avec $\lambda = -1$ et $\mu = 2$.

Dans le cas général : on utilise les dérivées successives de la fraction rationnelle pour déterminer les coefficients du DES.

➠ **Exemple XVII.6.** Considèrons la fraction rationnelle

$$F = \frac{X}{(X-1)(X-2)^2} .$$

Cette fraction est de degré -2 avec un pôle simple et un pôle double. Sa DES doit donc être de la forme :

$$F = \frac{\lambda}{X-1} + \frac{\mu}{X-2} + \frac{\nu}{(X-2)^2} .$$

Usant de la méthode vue *supra* relativement aux pôles simples, on trouve que $\lambda = 1$. Ensuite, en multipliant l'expression ci–dessus par $(X-2)^2$ et évaluant en "$X = 2$", on trouve :

$$\frac{2}{(2-1)} = \nu$$

et donc $\nu = 2$. Pour obtenir μ, on dérive l'expression ci dessus, puis on évalue en "$X = 2$", ce qui donne $\mu = -1$.

◇ **Le cas réel**

Théorème XVII.5 (Décomposition en éléments simples (cas réel))**.**
Soit $F = \dfrac{P}{Q} \in \mathbb{R}(X)$ une fraction irréductible et soit

$$Q = \prod_{i=1}^{n} (X - a_i)^{\alpha_i} \prod_{i=1}^{m} (X^2 + p_i X + q_i)^{\beta_i}$$

la décomposition en produit de facteurs irréductibles de Q sur $\mathbb{R}$, avec :
— les $\alpha_i, \beta_i \in \mathbb{N}^*$;
— les $a_i, p_i, q_i \in \mathbb{R}$ vérifiant $p_i^2 - q_i < 0$.
Alors il existe une unique famille de réels $\lambda_{i,j}, \mu_{i,j}$ et $\nu_{i,j}$ telle que :

$$F = E(F) + \sum_{i=1}^{n} \sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X - a_i)^j} + \sum_{i=1}^{m} \sum_{j=1}^{\beta_i} \frac{\mu_{i,j} X + \nu_{i,j}}{(X^2 + p_i X + q_i)^j} .$$

Démonstration. Admis. □

En pratique, il est souvent efficace de décomposer en éléments sur simples sur $\mathbb{C}$ puis de regrouper les pôles conjugués. On peut aussi, se ramener à un système linéaire, comme illustré *infra.*

➠ **Exemple XVII.7.** Pour déterminer la DES de $F = \dfrac{1}{X^3 - 1} = \dfrac{1}{(X-1)(X^2 + X + 1)}$, on part du théorème qui nous dit que cette dernière sera de la forme :

$$\frac{1}{(X-1)(X^2 + X + 1)} = \frac{\lambda}{X - 1} + \frac{\mu X + \nu}{X^2 + X + 1} .$$

Usant de la méthode usuelle, on trouve $\lambda = \dfrac{1}{3}$. De plus, on pose :

$$G = (X^2 + X + 1)F = \frac{1}{X - 1} = \lambda \frac{X^2 + X + 1}{X - 1} + \mu X + \nu$$

et on a, en évaluant G en 0 et 2 :

$$\begin{cases} -1 = -\lambda + \nu \\ 1 = 7\lambda + 2\mu + \nu \end{cases} ,$$

ce qui donne $\mu = -\dfrac{1}{3}$ et $\nu = -\dfrac{2}{3}$.

d) Dérivée logarithmique

Terminons ce chapitre par un exemple fondamental et fort utile. Si $P \in \mathbb{K}[X]$ est un polynôme non constant, alors il peut s'écrire (sur $\mathbb{C}$) sous la forme :

$$P = c \prod_{i=1}^{n} (X - a_i)^{k_i} ,$$

avec $c \in \mathbb{K}$, les $a_i \in \mathbb{C}$ et les $k_i \in \mathbb{N}^*$. De fait, si nous étions complètement malades, nous pourrions écrire la formule suivante :

$$\ln(P) = \ln(c) + \sum_{i=1}^{n} k_i \ln(X - a_i)$$

et "donc", en "dérivant" :

$$\frac{P'}{P} = \sum_{i=1}^{n} \frac{k_i}{X - a_i}.$$

Évidemment, on peut démontrer cette formule proprement en utilisant les méthodes de calcul de DES vues précédemment, mais cette recette de cuisine impie a l'avantage de rester en mémoire.

Exemple XVII.8.

— $\dfrac{2X+4}{(X+2)^2} = \dfrac{2}{(X+2)}$;

— $\dfrac{nX^{n-1}}{X^n - 1} = \sum\limits_{k=1}^{n} \dfrac{1}{X - e^{\frac{2ik\pi}{n}}}$ pour $n \geq 1$.

Addendum : calcul de primitives

Les résultats vus au chapitre précédent nous permettent de calculer efficacement les primitives de fonctions rationnelles ; il nous suffit en effet de savoir le faire pour les éléments simples.

◇ Éléments simples de première espèce

On entend par la les éléments simples de la forme

$$f : x \mapsto \frac{1}{(x-a)^n}$$

avec $n \geq 1$ et $a \in \mathbb{C}$.

Cas 1 : $n \geq 2$. Une primitive de f est aisée à trouver :

$$F : x \mapsto \frac{-1}{(n-1)(x-a)^{n-1}} .$$

Cas 2 : $n = 1$. Si $a \in \mathbb{R}$, une primitive de f est $F : x \mapsto \ln(|x-a|)$. Sinon, posons $a = p + iq$ et remarquons que, pour tout x raisonnable :

$$\begin{aligned} f(x) &= \frac{1}{(x-p)-iq} \\ &= \frac{(x-p)+iq}{(x-p)^2+q^2} \end{aligned}$$

et donc une primitive de f est donnée par

$$F : x \mapsto \frac{1}{2}\ln((x-p)^2+q^2) + i\arctan\left(\frac{x-p}{q}\right).$$

◇ Éléments simples de deuxième espèce

On entend par la les éléments simples de la forme

$$f : x \mapsto \frac{\alpha x + \beta}{x^2 + px + q}$$

avec $\alpha, \beta, p, q \in \mathbb{R}$ et $p^2 - 4q < 0$. En posant, pour $x \in \mathbb{R}$, $t = x + \frac{p}{2}$, on obtient que $f(x) = g(t)$, où g est de la forme :

$$g : t \mapsto \frac{ut+v}{t^2+a^2}$$

avec $u, v, a \in \mathbb{R}$ et $a \neq 0$. En effet :

$$\begin{aligned} x^2+px+q &= \left(t-\frac{p}{2}\right)^2 + p\left(t-\frac{p}{2}\right)+q \\ &= t^2 - pt + \frac{p^2}{4} + pt - \frac{p^2}{2} + q \\ &= t^2 - \frac{p^2-4q}{4} \end{aligned}$$

et donc $a = \sqrt{-\dfrac{p^2-4q}{4}}$ est bien défini et non nul car $p^2-4q<0$. Une primitive de g est donc donnée par :

$$G : t \mapsto \frac{u}{2}\ln(t^2+a^2) + \frac{v}{a}\arctan\left(\frac{t}{a}\right) .$$

◇ Éléments simples de deuxième espèce, mais en différent

Le dernier type d'éléments simples dont nous avons besoin de déterminer une primitive est celui représenté (modulo translation similaire à ce qui a été fait dans le paragraphe *supra* par les deux intégrales suivantes (pour x réel) :

$$I_n = \int_0^x \frac{t}{(t^2+a^2)^n}\,\mathrm{d}t \quad \text{et} \quad J_n = \int_0^x \frac{1}{(t^2+a^2)^n}\,\mathrm{d}t$$

avec $a \in \mathbb{R}^*$ et $n \geq 2$. Remarquons dans un premier temps que, par forme composée usuelle :

$$I_n = \frac{-1}{2(n-1)(x^2+a^2)^{n-1}}$$

et tentons de calculer J_n par IPP, appliquée aux fonctions $t :\mapsto t$ et $t \mapsto (t^2+a^2)^{-n}$, de classes $\mathscr{C}^1$. On a alors :

$$\begin{aligned} J_n =& x(x^2+a^2)^{-n} - n\int_0^x t\frac{-2t}{(t^2+a^2)^{n+1}}\,\mathrm{d}t \\ &= x(x^2+a^2)^{-n} + n\int_0^x \frac{2t^2}{(t^2+a^2)^{n+1}}\,\mathrm{d}t\,. \end{aligned}$$

Or, pour tout $t \in [0, x]$:

$$\begin{aligned} \frac{t^2}{(t^2+a^2)^{n+1}} &= \frac{t^2+a^2}{(t^2+a^2)^{n+1}} - \frac{a^2}{(t^2+a^2)^{n+1}} \\ &= \frac{1}{(t^2+a^2)^n} - \frac{a^2}{(t^2+a^2)^{n+1}} \end{aligned}$$

et donc

$$J_n = x(x^2+a^2)^{-n} + 2nJ_n - 2a^2nJ_{n+1}$$

i.e

$$J_{n+1} = \frac{x(x^2+a^2)^{-n}+J_n}{2a^2n}$$

ce qui permet un calcul itératif de J_n.

Deuxième partie

Second Semestre

Chapitre XVIII

Dénombrement, combinatoire

1.– Ensembles finis

a) Cardinal

Proposition XVIII.1. Soient $n, p \in \mathbb{N}$. Alors :

$$[\![1, p]\!] \text{ et } [\![1, n]\!] \text{ sont en bijection} \\ \iff \\ n = p.$$

Pour démontrer ce résultat, nous avons besoin du lemme suivant.

Lemme XVIII.1. Soient E, F deux ensembles non vides et soit $(a, b) \in E \times F$. Alors :

$$\text{il existe une bijection (resp. injection, surjection) de } E \text{ vers } F \\ \iff \\ \text{il existe une bijection (resp. injection, surjection) de } E \setminus \{a\} \text{ vers } F \setminus \{b\}.$$

Démonstration.

($\Downarrow$) Si $f : E \to F$ est une bijection, on pose :

$$\begin{aligned} \phi : F &\to F \\ x &\mapsto \begin{cases} b \text{ si } x = f(a) \\ f(a) \text{ si } x = b \\ x \text{ sinon} \end{cases}. \end{aligned}$$

Cette application est involutive (bijective de réciproque elle–même) et l'application $g = \phi \circ f$ est une bijection de E vers F telle que $g(a) = b$. Il suffit alors de la restreindre à $E \setminus \{a\}$.

($\Uparrow$) Immédiat : prolonger en envoyant a sur b. $\square$

Nous pouvons maintenant démontrer la proposition XVIII.1.

Démonstration.

($\Uparrow$) Si $n = p$, alors l'identité convient.

($\Downarrow$) On démontre par récurrence sur p que : pour tout $n \geq 0$ tel existe une bijection de $[\![1, n]\!]$ vers $[\![1, p]\!]$, alors $n = p$. L'hérédité découle du lemme précédent.

□

Proposition XVIII.2. De façon analogue, pour tous $n, p \in \mathbb{N}$:

(i) il existe une injection de $[\![1, n]\!]$ vers $[\![1, p]\!] \Leftrightarrow n \leq p$;

(ii) il existe une surjection de $[\![1, n]\!]$ vers $[\![1, p]\!] \Leftrightarrow n \geq p$.

Définition XVIII.1. Soit E un ensemble. Alors, on dit que E est **fini** si il existe $n \geq 0$ et $f : E \to [\![1, n]\!]$ bijective. L'entier n est alors unique et appelé **cardinal** de E.

Notation. Le cardinal sera noté indifféremment $|E|$, $\operatorname{card}(E)$ ou $\#E$.

✌ **Remarque XVIII.1.** L'unicité du cardinal découle de la proposition XVIII.1.

➠ **Exemple XVIII.1.**

— L'ensemble vide est fini de cardinal 0.

— L'ensemble $\{\spadesuit, \pitchfork, \clubsuit\}$ est fini de cardinal 3.

Notons donc que, de par cette définition, les bijections représentent un outil essentiel et central au dénombrement.

b) Propriétés et dénombrement

Proposition XVIII.3. Soient E, F deux ensembles finis. Alors :

(i) E et F sont équipotents (*i.e* il existe une bijection entre eux) $\Leftrightarrow$ $\operatorname{card}(E) = \operatorname{card}(F)$;

(ii) il existe une injection entre E et $F \Leftrightarrow \operatorname{card}(E) \leq \operatorname{card}(F)$;

(iii) il existe une surjection entre E et $F \Leftrightarrow \operatorname{card}(E) \geq \operatorname{card}(F)$.

Démonstration. Utiliser les règles de composition des applications (insérer truc ici)jectives.

□

Proposition XVIII.4. Soit E un ensemble fini et soit $E' \subset E$. Alors :

(i) E' est fini ;

(ii) $\operatorname{card}(E') \leq \operatorname{card}(E)$;

(iii) $E = E' \Leftrightarrow \operatorname{card}(E) = \operatorname{card}(E')$.

Démonstration. Les points (i) et (ii) découlent du fait que $x \mapsto x$ réalise une injection de E' dans E. Pour le cas d'égalité, procédons en deux temps.

($\Rightarrow$) ¯_(ツ)_/¯

($\Leftarrow$) Supposons $\mathrm{card}(E) = \mathrm{card}(E')$ et $E \neq E'$. Ceci signifie qu'il existe $e \in E \setminus E'$ et donc, en posant $n = \mathrm{card}(E')$ on a (la bijection est triviale à établir) :

$$\mathrm{card}(E' \cup \{a\}) = n + 1$$

et, comme $E \cup \{a\} \subset E$:

$$n = \mathrm{card}(E') < n + 1 \leq \mathrm{card}(E)$$

d'où le résultat.

□

La proposition suivante est au cœur de bien de ce chapitre et de plusieurs de ses applications ultérieures. Il peut donc être utile d'y accorder quelque attention.

Proposition XVIII.5. Soient E, F deux ensembles finis **de même cardinal** et soit $f : E \to F$ une application. Alors, les propriétés suivantes sont équivalentes :

(i) f est bijective ;

(ii) f est injective ;

(iii) f est surjective.

Démonstration. On procède de façon circulaire.

(i)$\Rightarrow$(ii) Trivial.

(ii)$\Rightarrow$(iii) Supposons f injective ; il en découle que $\mathrm{card}(f(E)) = \mathrm{card}(E)$. Par hypothèse, $\mathrm{card}(E) = \mathrm{card}(F)$ et donc $f(E)$ est une partie de F de cardinal égal à celui de F. *Ipso facto*, $f(E) = F$ et donc f est surjective.

(iii)$\Rightarrow$(i) Supposons f surjective ; alors il existe $g : F \to E$ telle que $f \circ g = \mathrm{id}_F$. L'application g est injective donc surjective (par le point précédent), ce qui entraîne que g est bijective et donc f également.

□

2.⁻ Zoologie cardinale

a) Opérations sur les ensembles finis

Proposition XVIII.6. Soient $E_1, \dots, E_n$ des ensembles finis **deux à deux disjoints**. Alors, la réunion de ces ensembles est finie et

$$\mathrm{card}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathrm{card}(E_k)\,.$$

Notation. Une telle réunion pourra être notée $\bigsqcup_{k=1}^{n} E_k$ pour en souligner le caractère disjoint.

✖ **Attention :** évidemment, ceci n'est vrai que pour les ensembles disjoints : le cardinal de la réunion $\{0, 1\} \cup \{0, 2\}$ n'est pas égal à 4.

Démonstration. Nous traitons le cas $n = 2$ car nous sommes de nature paresseuse. Le lecteur motivé et enthousiaste est libre de s'adonner à l'exercice hautement plaisant de démontrer le cas général par récurrence sur n. Regarder la peinture sécher sur un mur est aussi une option.

Donnons donc deux ensembles A et B finis de cardinaux respectifs p et q. Ceci nous livre deux bijections : $f : A \to [\![1, p]\!]$ et $g : B \to [\![1, q]\!]$. On pose :

$$\begin{aligned} h : A \sqcup B &\longrightarrow [\![1, p+q]\!] \\ x &\mapsto \begin{cases} f(x) \text{ si } x \leq p \\ p + g(x) \text{ sinon} \end{cases} \end{aligned}$$

et on vérifie que cette application est bien bijective, ce qui conclut notre démonstration. □

Proposition XVIII.7. Soient E et F deux ensembles finis. Alors :

(i) **si** $F \subset E$, $E \setminus F$ est fini et

$$\operatorname{card}(E \setminus F) = \operatorname{card}(E) - \operatorname{card}(F) \ ;$$

(ii) $E \cup F$ et $E \cap F$ sont finis, avec :

$$\operatorname{card}(E \cup F) = \operatorname{card}(E) + \operatorname{card}(F) - \operatorname{card}(E \cap F) \ .$$

✌ **Remarque XVIII.2.** La formule du crible de Poincaré, donnant le cardinal d'une réunion quelconque d'ensembles finis est hors programme. Le lecteur averti pourra se référer à la littérature, en prenant garde de porter des lunettes protectrices. En cas de contact direct avec la peau, consulter un médecin.

Démonstration. Il suffit de faire les deux remarques suivantes :

$$E = E \sqcup (E \setminus F)$$

et

$$E \cup F = (E \setminus (E \cap F)) \sqcup F \ .$$

□

✍ **Exercice XVIII.1.** Soient E, F, G trois ensembles finis. Calculer $\operatorname{card}(E \cup F \cup G)$.

➥ **Correction :** *On a :*

$$\begin{aligned} \operatorname{card}(E \cup F \cup G) &= \operatorname{card}(E \cup F) + \operatorname{card}(G) - \operatorname{card}((E \cup F) \cap G) \\ &= \operatorname{card}(E) + \operatorname{card}(F) + \operatorname{card}(G) - \operatorname{card}((E \cup F) \cap G) - \operatorname{card}(E \cap F) \\ &= \operatorname{card}(E) + \operatorname{card}(F) + \operatorname{card}(G) - \operatorname{card}((E \cap G) \cup (F \cap G)) - \operatorname{card}(E \cap F) \\ &= \operatorname{card}(E) + \operatorname{card}(F) + \operatorname{card}(G) \\ &\quad - \operatorname{card}(E \cap G) - \operatorname{card}(F \cap G)) - \operatorname{card}(E \cap F) \\ &\quad + \operatorname{card}(E \cap F \cap G) \ . \end{aligned}$$

Proposition XVIII.8. Soient E, F des ensembles finis. Alors :
(i) $\text{card}(E \times F) = \text{card}(E)\text{card}(F)$;
(ii) $\text{card}(E^p) = \text{card}(E)^p$ pour tout $p \in \mathbb{N}^*$.

Démonstration. Il suffit de remarquer pour le (i) que :

$$E \times F = \bigsqcup_{y \in F} E \times \{y\}\,.$$

Le point (ii) se traite ensuite par récurrence sur p. □

Exemple XVIII.2. Une expérience dans laquelle on a n choix puis m choix possède nm issues. Par exemple, un lancer de deux dés à 6 faces peut donner 36 résultats (pas somme) différents.

b) Parties d'un ensemble fini

Proposition XVIII.9. Soit E un ensemble fini de cardinal n. Alors :

$$\text{card}(\mathscr{P}(E)) = 2^n\,.$$

Démonstration. On procède par récurrence sur n. Le cas $n = 0$ est trivial car $\mathscr{P}(\emptyset) = \{\emptyset\}$. Si la propriété est vraie au rang n et que E est un ensemble fini de cardinal n, on fixe $a \in E$ et $E' = E \setminus \{a\}$.

Soit $A \subset E$; alors :
— soit $a \notin A$ et dans ce cas $A \subset E'$. Par hypothèse de récurrence il existe 2^n telles parties ;
— soit $a \in A$ et alors $A \setminus \{a\}E'$. Par hypothèse de récurrence il existe 2^n telles parties.

En conclusion, il existe $2^n + 2^n = 2^{n+1}$ parties distinctes de E. □

Exemple XVIII.3. $\text{card}(\mathscr{P}^(\llbracket 1, 3\rrbracket)) = 8$.

c) Applications

Proposition XVIII.10. Soient E et F deux ensembles finis. Alors :

$$\text{card}(F^E) = \text{card}(F)^{\text{card}(E)}\,.$$

Démonstration. On procède par récurrence sur $p = \text{card}(E)$.
— Si $p = 0$, E est vide et donc il existe une seule application dans F^E, donnée par le triplet $(\emptyset, F, \emptyset)$.

— Supposons la propriété vérifiée pour tous les ensembles de cardinal p, et fixons E un ensemble de cardinal $p+1$; on fixe $a \in E$ et $E' = E \setminus \{a\}$. Pour définir une application de E dans F il nous faut déterminer l'image de a ($\text{card}(F)$) possibilités puis celles des éléments de E' ($\text{card}(F)^p$ possibilités par hypothèse de récurrence). *In fine*,

$$\text{card}(F^E) = \text{card}(F) \times \text{card}(F)^p = \text{card}(F)^{\text{card}(E)} .$$

□

➠ **Exemple XVIII.4.** Il existe exactement 27 applications de $[\![1,3]\!]$ dans lui–même, dont 6 sont bijectives (les éléments de $\mathfrak{S}_3$).

3.– Un peu de combinatoire

a) Injections, bijections

Proposition XVIII.11. Soient E et F deux ensembles finis de cardinaux respectifs p et n. Alors :

(i) le nombre d'injections de E dans F est égal à $\dfrac{n!}{(n-p)!}$ si $n \geq p$ et 0 sinon ;

(ii) le nombre de bijections de E dans F est égal à $n!$ si $n = p$ et 0 sinon.

Démonstration. (i) Pour varier, procédons par récurrence sur $p = \text{card}(E)$. Si $p = 0$, on rigole, et si la propriété est vérifiée au rang p on fixe E un ensemble de cardinal $p+1$, $a \in E$ et $E' = E \setminus \{a\}$. Alors, pour définir une injection $f : E \hookrightarrow F$, il faut :

— choisir l'image de a (n possibilités) ;

— choisir (de façon injective) les images des éléments de E' dans $F \setminus \{f(a)\}$ $\left(\dfrac{(n-1)!}{(n-1-p)!} \text{ ou 0 possibilité(s) par hypothèse de récurrence}\right)$.

Ainsi, dans le cas ou la deuxième quantité mentionnée est non nulle, on a le nombre d'injections suivant :

$$n \times \frac{(n-1)!}{(n-1-p)!} = \frac{n!}{(n-(p+1))!}$$

d'où le résultat.

(ii) Déjà traité au chapitre X lors de l'étude du groupe symétrique. On peut aussi utiliser la proposition XVIII.5 et remarquer que si $n = p$ alors les injections de E dans F sont bijectives.

□

Notation. La quantité $\dfrac{n!}{(n-p)!}$ est notée $(n)_p$ et appelée **nombre d'arrangements** de p dans n.

Remarque XVIII.3. Notons que, si $n \geq p$:

$$(n)_p = n(n-1)\dots(n-p+1)\,.$$

Il s'agit du nombre de façon de classer p éléments parmi n.

Exemple XVIII.5. Il existe $\dfrac{9!}{2!} = 9\times 8\times 7\times 6\times 5\times 4\times 3 = 181\,440$ injections de $[\![1,7]\!]$ dans $[\![1,9]\!]$.

Corollaire XVIII.11.a. Soit F un ensemble de cardinal n et soit $p \leq n$. Alors l'ensemble des p–listes ordonnées (suites finies d'éléments deux à deux distincts) d'éléments de F est fini de cardinal $(n)_p$.

Démonstration. Il s'agit en fait de dénombrer les injections de $[\![1,p]\!]$ dans $[\![1,n]\!]$. □

Exemple XVIII.6. Si on effectue 4 tirages **successifs** sans remise dans une urne contenant 10 objets distinguables, il existe $(10)_4 = 5040$ tirages possibles.

b) Parties de taille fixée

Nous venons de voir comment dénombrer les choix de p éléments parmi n **avec** un ordre fixé. Dans ce paragraphe, nous nous intéressons à la combinatoire des parties (**non ordonnées**, donc) de taille fixée d'un ensemble fini.

Notation. Si E est un ensemble fini et $k \in \mathbb{N}$, on note $\mathscr{P}_k(E)$ l'ensemble des parties de E de cardinal k.

Proposition XVIII.12. Soient $n, k \in \mathbb{N}$ tels que $k \leq n$ et soit E un ensemble fini de cardinal n. Alors :

$$\mathrm{card}(\mathscr{P}_k(E)) = \binom{n}{k}\,.$$

Démonstration. Une récurrence sur n ? Joie !

— Si $n = 0$, la vie est belle.

— Supposons la propriété vérifiée pour $n \geq 0$. Soit E un ensemble de cardinal $n+1$ et soit $A \in \mathscr{P}_k(E)$; on a alors deux possibilités : soit A ne contient pas $n+1$ (et on dénombre par hypothèse de récurrence), soit A contient $n+1$ et alors on dénombre les possibilités pour $A \setminus \{n+1\}$ via cette même, fort urbaine, hypothèse. *In fine* :

$$\mathrm{card}(\mathscr{P}_k(E)) = \binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$$

par formule de Pascal.

□

Exemple XVIII.7. Si on tire 4 boules parmi 10 simultanément, il existe $\binom{10}{4} = 210$ issues possibles. Cela est très différent du tirage avec ordre vu précédemment.

Ces procédés combinatoires nous permettent de démontrer différemment certains résultats vus au chapitre IV ; nous offrons gracieusement au lecteur un florilège de ceux–ci.

◇ Sommes des coefficients binomiaux

Si n est un entier naturel, on a :

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n}{k} &= \sum_{k=0}^{n} \operatorname{card}(\mathscr{P}_k(\llbracket 1, n \rrbracket)) \\
&= \operatorname{card}\left(\bigsqcup_{k=0}^{n} \mathscr{P}_k(\llbracket 1, n \rrbracket) \right) \\
&= \operatorname{card}(\mathscr{P}(\llbracket 1, n \rrbracket)) \\
&= 2^n .
\end{aligned}$$

◇ Formule de Pascal

Soit E un ensemble de cardinal $n \geq 1$ et soit $k \leq n$. On fixe $a \in E$; alors pour construire une partie à k éléments de E, il faut :

— soit choisir k éléments parmi $E' = E \setminus \{a\}$;

— soit choisir $k - 1$ éléments parmi E' et ajouter a.

De fait,

$$\operatorname{card}(\mathscr{P}_k(E)) = \operatorname{card}(\mathscr{P}_k(E')) + \operatorname{card}(\mathscr{P}_{k-1}(E'))$$

ce qui se traduit par :

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

◇ Binôme de Newton

Soient $a, b \in \mathbb{C}$ et soit $n \geq 1$. Alors $(a + b)^n$ se développe en 2^n termes de la forme $a^k b^{n-k}$, où k est le nombre de facteurs dans lequel nous avons "choisi" a pour obtenir ce terme. Ce terme apparaîtra donc $\binom{n}{k}$ car ces k occurrences de a sont à "choisir" parmi n facteurs possibles.

$$(a+b)(a+b)(a+b)(a+b)$$

3 "choix" de a : $a^3 b^1$

Chapitre XIX

Espaces vectoriels

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.⁻ Structures linéaires

a) C'est quoi ?

Définition XIX.1. Soit E un ensemble muni d'un loi de composition interne $+ : E \times E \to E$ et d'une loi de composition externe $\cdot : \mathbb{K} \times E \to E$. On dit que le triplet $(E, +, \cdot)$ est un $\mathbb{K}$–**espace vectoriel** (abrégé $\mathbb{K}$–e.v) si :

(A) $(E, +)$ est un groupe abélien ;

(B) $\forall x \in E$, $1_{\mathbb{K}} \cdot x = x$;

(C) $\forall \lambda \in \mathbb{K}$, $\forall x, y \in E$, $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$;

(D) $\forall \lambda, \mu \in \mathbb{K}$, $\forall x \in E$, $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$;

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

Vocabulaire. Les éléments de E sont appelés **vecteurs**, ceux de $\mathbb{K}$ **scalaires**.

✌ **Remarque XIX.1.** La donnée des lois de composition pourra être sous entendue, tout comme la notation "$\cdot$" pour la multiplication externe.

➠ **Exemple XIX.1.**
- $\mathbb{K}$ est un $\mathbb{K}$–e.v ; la multiplication externe est alors interne ;
- l'ensemble $\mathscr{F}(\mathbb{K}, \mathbb{K}) = \mathbb{K}^{\mathbb{K}}$ est un $\mathbb{K}$–e.v pour l'addition des fonctions et leur multiplication par une constante ;
- $\{0_{\mathbb{K}}\}$ est un $\mathbb{K}$–e.v, appelé **espace nul** ;
- l'ensemble $\mathbb{K}^2$ muni de l'addition coordonnée par coordonnée et de la multiplication par un scalaire est un $\mathbb{K}$–e.v ;
- $\mathbb{K}[X]$ est un $\mathbb{K}$–e.v pour l'addition des polynômes et la multiplication par une constante. Il en va similairement de $\mathbb{K}(X)$.

Proposition XIX.1. Soit E un $\mathbb{K}$–e.v et soient $\lambda \in \mathbb{K}$ et $x \in E$. Alors :

$$(\lambda \cdot x = 0) \Leftrightarrow (\lambda = 0) \vee (x = 0) .$$

✖ **Attention :** il y a deux zéros différents dans la proposition *supra* : le premier est celui de $\mathbb{K}$, le second celui de E.

Démonstration. ($\Leftarrow$) Cela découle du fait que $\lambda \cdot 0 = \lambda \cdot (0+0)$ et donc $\lambda \cdot 0 = -\lambda \cdot 0$. De même, $(0+0) \cdot x = 0 \cdot x \ldots$

($\Rightarrow$) Supposons $\lambda \cdot x = 0$ et $\lambda \neq 0$. Alors, quitte à multiplier des deux côtés par $\frac{1}{\lambda}$ on obtient que $1 \cdot x = 0$, *i.e* $x = 0$.

□

Proposition XIX.2. Soit E un $\mathbb{K}$–e.v et soit $x \in E$. Alors $(-1) \cdot x = -x$.

✖ **Attention :** "$-x$" représente ici l'inverse de x pour "$+$".

Démonstration. On a $1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0$ d'où le résultat. □

b) Exemples fondamentaux

◇ $\mathbb{K}^n$

L'espace $\mathbb{K}^n$ est un $\mathbb{K}$–e.v pour l'addition terme à terme et la multiplication par un scalaire, *i.e* si $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in E$ et $\lambda \in \mathbb{K}$ on pose :

$$(x_1, \ldots x_n) + \lambda \cdot (y_1, \ldots, y_n) = (x_1 + \lambda y_1, \ldots, x_n + \lambda y_n) .$$

✖ **Attention :** le corps de base est évidemment fondamental : $\mathbb{R}^2$ est un $\mathbb{R}$–e.v et $\mathbb{C} = \mathbb{C}^1$ un $\mathbb{C}$–e.v et pourtant leurs structures sont différentes. En particulier, on ne multiplie pas par i dans $\mathbb{R}^2$......

◇ **Fonctions**

Soit X un ensemble et E un $\mathbb{K}$–e.v ; alors l'ensemble $\mathscr{F}(X, E) = E^X$ est un $\mathbb{K}$–e.v pour les opérations suivantes : si $f, g \in E^X$ et $\lambda \in \mathbb{K}$ alors

$$(f + \lambda g) : x \mapsto f(x) + \lambda g(x) .$$

En particulier, l'ensemble $E^{\mathbb{N}}$ des suites à valeurs dans E est un $\mathbb{K}$–e.v.

◇ **Produits**

Soient $E_1, \ldots, E_n$ des $\mathbb{K}$–e.v ; alors le produit $E_1 \times E_n$ est un $\mathbb{K}$–e.v pour l'addition terme à terme et la multiplication scalaire. On retrouve la structure vue sur $\mathbb{K}^n$ comme cas particulier.

c) Combinaisons linéaires

Définition XIX.2. Soit E un $\mathbb{K}$–e.v ; on appelle **combinaison linéaire sur** E toute expression de la forme

$$\sum_{i=1}^{n} \lambda_i x_i$$

avec les $\lambda_i \in \mathbb{K}$ (appelés **coefficients**) et les $x_i \in E$.

➠ **Exemple XIX.2.**

- Dans $\mathbb{K}^2$, tout élément est combinaison linéaire de $(0,1)$ et $(1,0)$.
- Dans $\mathbb{K}^3$, posons $e_1 = (1,2,1)$, $e_2 = (2,-1,0)$ et $e_3 = (3,1,1)$. Alors, $(x,y,z) \in \mathbb{K}^3$ est combinaison linéaire de e_1, e_2 et e_3 si et seulement si il existe λ, μ et ν tels que :

$$\lambda e_1 + \mu e_2 + \nu e_3 = (x,y,z)$$

i.e

$$\begin{cases} \lambda + 2\mu + 3\nu & = & x \\ 2\lambda - \mu + \nu & = & y \\ \lambda + \nu & = & z \end{cases}.$$

Ce concept de combinaisons linéaires (finies) peut se reformuler pour accepter des familles de vecteurs potentiellement infinies à condition que seul un nombre fini de coefficients soient non nuls. Nous avons vu un procédé similaire lors de l'étude des polynômes au chapitre XV.

Définition XIX.3. Une famille $(\lambda_i)_{i\in I}$ de scalaires sera dite **presque nulle** (ou à support fini) si il existe $J \subset I$ fini tel que :

$$\forall i \in I \setminus J,\ \lambda_i = 0\,.$$

Un tel ensemble J **minimal** est alors appelé **support** de la famille $(\lambda_i)_{i\in I}$.

Dans le cas où $(x_i)_{i\in I}$ est une famille de vecteurs et $(\lambda_i)_{i\in I}$ une famille presque nulle de scalaires de support J, on pourra noter :

$$\sum_{i\in I} \lambda_i x_i = \underbrace{\sum_{i\in J} \lambda_i x_i}_{\text{somme finie}}$$

et on parlera de "combinaison linéaire des x_i".

➠ **Exemple XIX.3.** Tout polynôme est combinaison linéaire de la famille $(X^n)_{n\in\mathbb{N}}$.

2.– Applications linéaires

a) C'est quoi ?

Définition XIX.4. Soient E et F deux $\mathbb{K}$–e.v ; une application $f : E \to F$ est dite **linéaire** si :
— $\forall x, y \in E,\ f(x+y) = f(x) + f(y)$;
— $\forall x \in E, \forall \lambda \in \mathbb{K},\ f(\lambda x) = \lambda f(x)$.

✌ **Remarque XIX.2.** Une telle application est donc un morphisme de groupes de $(E, +)$ vers $(F, +)$ ce qui entraîne que $f(0_E) = 0_F$. **Il est donc inutile de démontrer cette propriété lors de l'étude.**

➠ **Exemple XIX.4.** Les applications suivantes sont linéaires :

—
$$\begin{aligned} f : \mathbb{K}^2 &\longrightarrow \mathbb{K}^2 \\ (x, y) &\mapsto (x - y, x + y) \end{aligned} \qquad ;$$

—
$$\begin{aligned} g : \mathbb{K}^2 &\longrightarrow \mathbb{K} \\ (x, y) &\mapsto x - 14y \end{aligned} \qquad ;$$

—
$$\begin{aligned} h : \mathbb{K}^{\mathbb{N}} &\longrightarrow \mathbb{K}^{\mathbb{N}} \\ u &\mapsto (3u_n - u_0)_n \end{aligned} \qquad ;$$

—
$$\begin{aligned} \mathrm{d} : \mathbb{K}[X] &\longrightarrow \mathbb{K}[X] \\ P &\mapsto P' \end{aligned} \qquad ;$$

— pour $\lambda \in \mathbb{K}$, et E un $\mathbb{K}$–e.v l'homothétie de rapport λ sur E :
$$\begin{aligned} h_\lambda : E &\longrightarrow E \\ x &\mapsto \lambda x \end{aligned} \qquad ;$$

— si $a \in \mathbb{K}$, l'évaluation en a :
$$\begin{aligned} \mathrm{eval}_a : \mathbb{K}[X] &\longrightarrow \mathbb{K}[X] \\ P &\mapsto P(a) \end{aligned} \qquad ;$$

— si E est l'ensemble des suites convergentes de $\mathbb{K}^{\mathbb{N}}$ (vérifier qu'il s'agit bien d'un $\mathbb{K}$–e.v .) :
$$\begin{aligned} \lim : E &\longrightarrow \mathbb{K} \\ u &\mapsto \lim u_n \end{aligned} \qquad ;$$

— si I est un intervalle de $\mathbb{R}$, les espaces $\mathscr{C}(I, \mathbb{K})$ sont des $\mathbb{K}$–e.v et, pour tout $a \in I$,

$$I : \mathscr{C}^1(I, \mathbb{K}) \longrightarrow \mathscr{C}^0(I, \mathbb{K})$$
$$f \mapsto \left(x \mapsto \int_a^x f(t)\,\mathrm{d}t\right)$$

est linéaire.

✌ **Remarque XIX.3.** Une application linéaire conserve les combinaisons linéaires : ceci est **fondamental**. Si f est linéaire, $(\lambda_i)_{i\in I}$ est une famille presque nulle de scalaires et $(x_i)_{i\in I}$ une famille de vecteurs, alors :

$$f\left(\sum_{i\in I} \lambda_i x_i\right) = \sum_{i\in I} \lambda_i f(x_i)\,.$$

Notation.
— L'ensemble des applications linéaires de E vers F est noté $\mathscr{L}(E, F)$.
— Les applications linéaires de E dans E sont appelées **endomorphismes** ; leur ensemble est noté $\mathscr{L}(E)$.
— Les applications linéaires bijectives de E dans F sont appelées **isomorphismes** ; leur ensemble est noté $GL(E, F)$.
— Les applications linéaires bijectives de E dans E sont appelées **automorphismes** ; leur ensemble est noté $GL(E)$ et appelé **groupe linéaire de** E.
— Les applications linéaires de E dans $\mathbb{K}$ sont appelées **formes linéaires** ; leur ensemble est noté E^* et appelé **dual de** E.

◇ Endomorphismes du plan

Posons $e_1 = (1, 0)$, $e_2 = (0, 1) \in \mathbb{K}^2$. Alors, tout élément $u = (x, y) \in \mathbb{K}^2$ s'écrit de façon unique comme combinaison linéaire de e_1 et e_2, en l'occurrence :

$$u = xe_1 + ye_2\,.$$

Soit à présent $f \in \mathscr{L}(\mathbb{K}^2)$. Alors, pour tout $u = (x, y) \in \mathbb{K}^2$, on a :

$$f(u) = f(xe_1 + ye_2) = xf(e_1) + yf(e_2)$$

et donc l'application f est totalement déterminée par la donnée de $f(e_1)$ et $f(e_2)$. Plus précisément, si on pose $(a, c) = f(e_1)$ et $(b, d) = f(e_2)$ on a :

$$f(u) = (ax + by, cx + dy)\ \ :.$$

Réciproquement, les applications définies par ce type de formule sont bien des endomorphismes de $\mathbb{K}^2$, *ergo* :

$$\mathscr{L}(\mathbb{K}^2) = \{(x, y) \mapsto (ax + by, cx + dy)\,|\, a, b, c, d \in \mathbb{K}\}\,.$$

Ceci étant fait, intéressons nous au cas des automorphismes : si $f \in \mathscr{L}(\mathbb{K}^2)$ et que $u, u' \in \mathbb{K}^2$ vérifient $f(u) = f(u')$, on a :

$$f(u - u') = f(u) - f(u') = 0$$

et donc, on en déduit aisément que f est injective si et seulement si :

$$\{u \in \mathbb{K}^2 \mid f(u) = 0\} = \{0\} .$$

Si f est donnée par le mécanisme $f : (x, y) \mapsto (ax + by, cx + dy)$ avec $a, b, c, d \in \mathbb{K}$, alors :

$$\begin{aligned}\forall u = (x, y) \in \mathbb{K}^2,\ f(u) = 0 &\Leftrightarrow \left\{\begin{array}{lcl} ax + by &=& 0 \\ cx + dy &=& 0 \end{array}\right. \\ &\Leftrightarrow \left\{\begin{array}{lcl} (ad - bc)x &=& 0 \\ (ad - bc)y &=& 0 \end{array}\right.\end{aligned}$$

via les opérations élémentaires $L_1 \leftarrow dL_1 - bL_2$ et $L_2 \leftarrow aL_2 - cL_1$. Ce système n'admet de solution non nulle que si $ad - bc \neq 0$, ce qui nous livre une CNS d'injectivité pour les endomorphismes de $\mathbb{K}^2$. On peut adapter ce raisonnement pour démontrer que la même condition nous livre la surjectivité, ce qui entraîne que :

$$GL(\mathbb{K}^2) = \{(x, y) \mapsto (ax + by, cx + dy) \mid a, b, c, d \in \mathbb{K},\ ad - bc \neq 0\} .$$

b) Opérations sur les applications linéaires

Proposition XIX.3. Soient E et F deux $\mathbb{K}$–e.v ; alors $\mathscr{L}(E, F)$ est un $\mathbb{K}$–e.v .

Démonstration. C'est essentiellement immédiat une fois les opérations posées : il s'agit de celles déjà définies sur les ensembles de fonctions. □

✌ **Remarque XIX.4.** $\mathscr{L}(E, F) \subset F^E$: il s'agit donc d'un $\mathbb{K}$–e.v inclus dans un autre $\mathbb{K}$–e.v ; de là à parler de "sous–$\mathbb{K}$–e.v "...

Proposition XIX.4. Soient E, F, G des $\mathbb{K}$–e.v et soient $f \in \mathscr{L}(E, F)$ et $g \in \mathscr{L}(F, G)$. Alors :

(i) $g \circ f \in \mathscr{L}(E, G)$;

(ii) si $f \in GL(E, F)$, alors sa réciproque est linéaire.

Démonstration. (i) Trivial.

(ii) Soient $X, Y \in F$; alors, comme f est bijective, il existe un unique couple $(x, y) \in E^2$ tel que $f(x) = X$ et $f(y) = Y$. Ainsi, pour tout $\lambda \in \mathbb{K}$:

$$\begin{aligned} f^{-1}(X + \lambda Y) &= f^{-1}(f(x) + \lambda f(y)) \\ &= f^{-1} \circ f(x + \lambda y) \text{ par linéarité de } f \\ &= x + \lambda y \\ &= f^{-1}(X) + \lambda f^{-1}(Y) . \end{aligned}$$

□

Corollaire XIX.4.a. Soit E un $\mathbb{K}$–e.v . Alors $(\mathscr{L}(E), +, \circ)$ est un anneau.

Remarque XIX.5. Cet anneau est non commutatif dés que $\dim(E) \geq 2$ (*cf.* chapitre XX). On peut par exemple remarquer que $f : (x, y) \mapsto (0, y)$ et $g : (x, y) \mapsto (x + y, y)$ ne commutent pas car $g \circ f(0, 1) = (1, 1)$ et $f \circ g(0, 1) = (0, 1)$.

Corollaire XIX.4.b. Soit E un $\mathbb{K}$–e.v . Alors $(GL(E), \circ)$ est un groupe.

✖ **Attention :** $GL(E)$ n'est **PAS** un $\mathbb{K}$–e.v : si $f \in GL(E)$, $f - f \notin GL(E)$ (si $E \neq \{0\}$)...

3.⁻ Sous–espaces vectoriels

a) Quoi ?

Définition XIX.5. Soit $(E, +, \cdot)$ un $\mathbb{K}$–e.v et soit $F \subset E$. On dit que F est un **sous–espace vectoriel** (s–e.v) de E si $(F, +, \cdot)$ est un $\mathbb{K}$–e.v .

Exemple XIX.5.
- $\{0_E\}$ est un s–e.v de E pour tout $\mathbb{K}$–e.v E ;
- de même, tout $\mathbb{K}$–e.v est un s–e.v de lui–même ;
- si E et F sont deux $\mathbb{K}$–e.v , $\mathscr{L}(E, F)$ est un s–e.v de F^E ;
- pour tout $n \geq 0$, $\mathbb{K}_n[X]$ est un s–e.v de $\mathbb{K}[X]$;
- l'ensemble des suites convergentes sur $\mathbb{K}$ est un s–e.v de $\mathbb{K}^{\mathbb{N}}$;
- $\mathscr{C}^0(\mathbb{R})$ est un s–e.v de $\mathbb{R}^{\mathbb{R}}$;
- $\mathscr{C}^1(\mathbb{R})$ est un s–e.v de $\mathscr{C}^0(\mathbb{R})$ et donc de $\mathbb{R}^{\mathbb{R}}$.

Nous disposons, comme pour les sous–groupes/anneaux, d'une caractérisation des s–e.v qui nous sera précieuse en pratique. En l'occurrence, F est un s–e.v de E si et seulement si :

(i) $F \subset E$;

(ii) $0 \in F$;

(iii) $\forall x, y \in F$, $\forall \lambda \in \mathbb{K}$, $x + \lambda y \in F$.

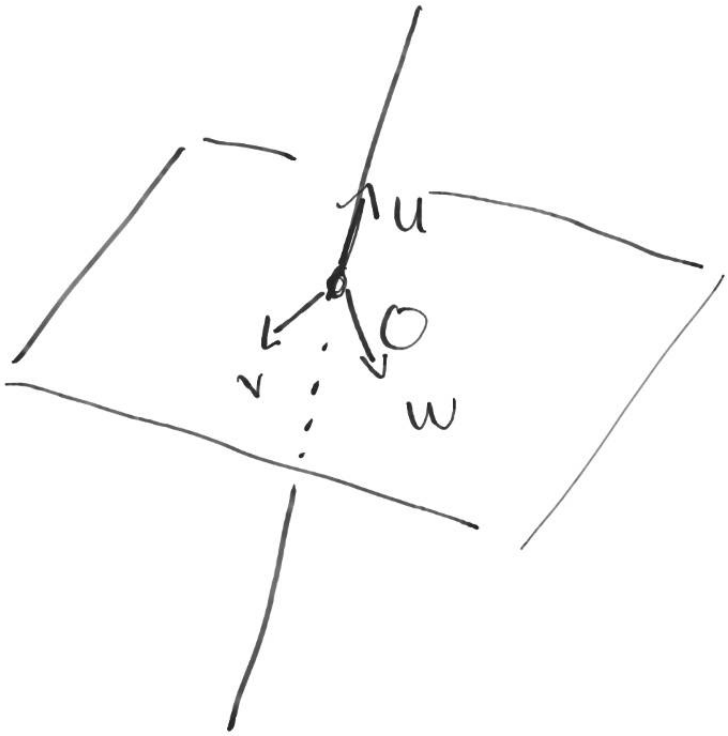
b) Exemples fondamentaux

◇ Objets "géométriques"

On fixe dans cette partie un $\mathbb{K}$–e.v E. On appelle **droite engendrée par un vecteur** $v \in E$ l'ensemble :

$$\mathscr{D}_v = \{\lambda v \mid \lambda \in \mathbb{K}\} .$$

Dans $\mathbb{R}^2$ et $\mathbb{R}^3$, on retrouve la notion de droite (vectorielle) vue dans les classes antérieures, ce qui est plutôt rassurant.



De la même façon, étant donné deux vecteurs $u, v \in E$ tels que $u \notin \mathscr{D}_v$ (on parle de vecteurs **non colinéaires**), on peut définir le **plan engendré par** u **et** v comme l'ensemble :

$$\mathscr{P}_{u,v} = \{\lambda u + \mu v \mid \lambda, \mu \in \mathbb{K}\}\ .$$

Il s'agit donc de l'ensemble des combinaisons linéaires de u et v ; ceci coïncide également avec la notion de plan vectoriel dans $\mathbb{R}^2$ et $\mathbb{R}^3$ avec laquelle nous nous permettons d'espérer que le lecteur soit familier. Dans le cas contraire, nous incluons un "joli" dessin.

➠ **Exemple XIX.6.** Pour tout $x \in \mathbb{R}$ et tout $\theta \in \mathbb{R}$, on a

$$\cos(x+\theta) = \cos(x)\cos(\theta) - \sin(x)\sin(\theta)$$

et donc $x \mapsto \cos(x+\theta)$ appartient au plan engendré par cos et sin.

Proposition XIX.5. Droites et plans sont des s–e.v des espaces concernés.

Démonstration. Trivial via la caractérisation. □

◇ **Image, image réciproque, noyau**

Proposition XIX.6. Soient E, F deux $\mathbb{K}$–e.v , E' (resp. F') un s–e.v de E (resp. F) et $f \in \mathscr{L}(E, F)$. Alors :

(i) $f(E')$ est un s–e.v de F ;

(ii) $f^{-1}(F')$ est un s–e.v de E.

Démonstration.

(i) Il est clair que $f(E') \subset F$; de plus $0 \in E'$ (car E' est un s–e.v de E) ce qui entraîne que $0 = f(0) \in f(E')$. Soient ensuite $X, Y \in f(E')$ et $\lambda \in \mathbb{K}$; alors il existe $x, y \in E'$ tels que $X = f(x)$ et $Y = f(y)$ donc :

$$\begin{aligned} X + \lambda Y &= f(x) + \lambda f(y) \\ &= f(\underbrace{x + \lambda y}_{\in E'}) \in f(E') . \end{aligned}$$

(ii) On sait que $f^{-1}(F') \subset E$; de plus $0 \in F'$ ce qui entraîne que $0 = f^{-1}(0) \in f^{-1}(F')$. Soient ensuite $x, y \in f^{-1}(F')$ et $\lambda \in \mathbb{K}$; alors :

$$f(x + \lambda y) = f(x) + \lambda f(y) \in F'$$

et donc $x + \lambda y \in f^{-1}(F')$.

$\square$

Proposition/définition XIX.6. Soient E, F deux $\mathbb{K}$–e.v et soit $f \in \mathscr{L}(E, F)$. On appelle :

— **noyau** de f l'ensemble

$$\mathrm{Ker}(f) = \{x \in E \,|\, f(x) = 0\} ;$$

— **image** de f l'ensemble

$$\mathrm{Im}(f) = \{f(x) \,|\, x \in E\} .$$

Ces deux ensembles sont de plus des s–e.v de E et F respectivement.

Démonstration. Il suffit de remarquer que $\mathrm{Ker}(f) = f^{-1}(\{0\})$ et $\mathrm{Im}(f) = f(E)$. $\square$

Exemple XIX.7.

— Pour tout $a \in \mathbb{K}$, le noyau de l'application $\mathrm{eval}_a \in \mathbb{K}[X]^*$ est l'ensemble des polynômes admettant a comme racine, *i.e* $(X - a)\mathbb{K}[X]$. Son image est $\mathbb{K}$ tout entier car, pour tout $x \in \mathbb{K}$, $x = \mathrm{eval}_a(x)$.

— Le noyau de l'application $\mathrm{d} \in \mathscr{L}(\mathscr{C}^1(\mathbb{R}), \mathscr{C}^0(\mathbb{R}))$ définie par $\mathrm{d} : f \mapsto f'$ est l'ensemble des fonctions constantes. Son image est l'ensemble des fonctions continues : en effet, si $f \in \mathscr{C}^0(\mathbb{R})$, alors

$$f = \mathrm{d}\left(x \mapsto \int_0^x f(t)\,\mathrm{d}t\right) .$$

— L'image de la fonction $x \mapsto (x, 2x)$ définie de $\mathbb{K}$ dans $\mathbb{K}^2$ est la droite engendrée par $(1, 2)$. Son noyau est nul.

Proposition XIX.7. Soient E, F deux $\mathbb{K}$–e.v et soit $f \in \mathscr{L}(E, F)$. Alors :

(i) f est injective $\Leftrightarrow \mathrm{Ker}(f) = \{0\}$;

(ii) f est surjective $\Leftrightarrow \mathrm{Im}(f) = F$.

Démonstration. (i) **(⇒)** Immédiat car $f(0) = 0$.

(⇔) Supposons que $\mathrm{Ker}(f) = \{0\}$ et fixons $x, x' \in E$ tels que $f(x) = f(x')$. Alors, $f(x - x') = f(x) - f(x') = 0$ et donc $x - x' \in \mathrm{Ker}(f)$, ce qui entraîne que $x = x'$.

(ii) Il s'agit de la définition de surjectivité.

□

➠ **Exemple XIX.8.** On peut déterminer le noyau de l'application

$$\begin{aligned} f : \mathbb{K}^2 &\to \mathbb{K}^2 \\ (x, y) &\mapsto (x + y, x - y) \end{aligned}$$

en résolvant le système

$$\left\{ \begin{array}{rcl} x + y &=& 0 \\ x - y &=& 0 \end{array} \right. .$$

On trouve $\mathrm{Ker}(f) = \{0\}$: f est injective. Le lecteur intrigué pourra vérifier qu'elle est même surjective.

c) Opérations sur les s–e.v

On fixe dans cette partie un $\mathbb{K}$–e.v E.

Proposition XIX.8. Soit $(F_i)_{i \in I}$ une famille (quelconque) de s–e.v de E. Alors :

$$\bigcap_{i \in I} F_i \text{ est un s–e.v de } E.$$

Démonstration. Immédiat via la caractérisation ; on rappelle que $x \in \bigcap_{i \in I} F_i$ si et seulement si $\forall i \in I$, $x \in F_i$. □

➠ **Exemple XIX.9.** L'intersection de deux droites est un s–e.v ; c'est même le sous–espace nul. Dans $\mathbb{R}^3$, on peut s'intéresser aux intersections de plans, qui sont des droites. Dans $\mathbb{R}^4$, deux plans peuvent ne s'intersecter qu'en 0, et cela peut causer des migraines.

✖ **Attention :** la réunion de deux s–e.v n'est a priori **pas** un s–e.v : regarder $\mathscr{D}_{e_1} \cup \mathscr{D}_{e_2}$ dans $\mathbb{K}^2$ qui contient e_1 et e_2 mais pas $(1, 1) = e_1 + e_2$.

Proposition/définition XIX.7. Soit $A \subset E$. Alors l'intersection de tout les s–e.v de E contenant A est un s–e.v de E, appelé **espace engendré par** A.

Notation. $\mathrm{Vect}(A)$ ou $\langle A \rangle$. Lorsque le corps de base sera une donnée essentielle, on notera $\mathrm{Vect}_{\mathbb{K}}(A)$.

Démonstration. Il s'agit d'une intersection de s–e.v . □

Remarque XIX.6. Le sous–espace engendré par A est donc le plus petit s–e.v de E contenant A.

Exemple XIX.10.
— $\mathrm{Vect}(\emptyset) = \{0\}$;
— si F est un s–e.v de E, $\mathrm{Vect}(F) = F$;
— si $e \in E \setminus \{0\}$, $\mathrm{Vect}(e)$ est la droite $\mathscr{D}_e$;
— de même, si $u, v \in E$ sont non colinéaires, $\mathrm{Vect}(u, v) = \mathscr{P}_{u,v}$.

Les deux derniers exemples se visualisent plus facilement une fois la proposition suivante démontrée.

Proposition XIX.9. Soit $A \subset E$. Alors $\mathrm{Vect}(A)$ est l'ensemble des combinaisons linéaires d'éléments de A.

Démonstration. Posons F l'ensemble des combinaisons linéaires d'éléments de A. Il s'agit naturellement d'un s–e.v de E contenant A et donc, par minimalité, $\mathrm{Vect}(A) \subset F$. De plus, tout s–e.v de E contenant A doit contenir F par stabilité, *ergo* $F \subset \mathrm{Vect}(A)$. □

Exercice XIX.1. Soient $u = (1, 1)$ et $v = (2, 2)$ deux vecteurs de $\mathbb{K}^2$. Déterminer l'espace $\mathrm{Vect}(u, v)$.

Correction : *u et v sont colinéaires, donc* $\mathrm{Vect}(u, v) = \mathrm{Vect}(u) = \mathscr{D}_u$.

Proposition XIX.10. Soient $A, B \subset E$. Alors :
(i) $A \subset \mathrm{Vect}(A)$ avec égalité si et seulement si A est un s–e.v de E ;
(ii) $\mathrm{Vect}(\mathrm{Vect}(A)) = \mathrm{Vect}(A)$;
(iii) $(A \subset B) \Rightarrow (\mathrm{Vect}(A) \subset \mathrm{Vect}(B))$;
(iv) $\mathrm{Vect}(A \cap B) \subset \mathrm{Vect}(A) \cap \mathrm{Vect}(B)$.

Démonstration. (i) L'inclusion est immédiate ; le cas d'égalité se traite en remarquant que si A est $\mathbb{K}$–e.v , alors il s'agit bien du plus petit s–e.v de E contenant A.

(ii) Trivial.

(iii) Si $A \subset B$, alors tout s–e.v de E contenant B contient A et donc $\mathrm{Vect}(A)$, d'où le résultat.

(iv) $A \cap B \subset A$ et donc $\mathrm{Vect}(A \cap B) \subset \mathrm{Vect}(A)$ par (iii). On procède symétriquement avec B.

□

ATTENTION : on n'a pas l'égalité $\mathrm{Vect}(A \cap B) = \mathrm{Vect}(A) \cap \mathrm{Vect}(B)$ dans le cas général. Prendre comme contre–exemple $A = \{(1, 2)\}$ et $B = \{(2, 4)\}$ dans $\mathbb{R}^2$: on a $\mathrm{Vect}(A) = \mathrm{Vect}(B)$ et $\mathrm{Vect}(A \cap B) = \mathrm{Vect}(\emptyset) = \{0\}$.

4.⁻ Sous–espaces supplémentaires

On fixe dans cette partie un $\mathbb{K}$–e.v E.

a) Somme de s–e.v

Soient F et G deux s–e.v de E. Nous avons vu que l'ensemble $F \cup G$ n'était pas en général un s–e.v de E ; cependant l'espace engendré $\text{Vect}(F \cup G)$ en est bien un. La question à laquelle nous devons répondre à présent est : certes, mais qui est–il ?

Proposition XIX.11. Soient F et G deux s–e.v de E. Alors l'espace $\text{Vect}(F \cup G)$ est égal à

$$F + G = \{x + y \,|\, (x, y) \in F \times G\} .$$

Démonstration. On procède par double inclusion pour montrer que $F+G = \text{Vect}(F\cup G)$.

($\supset$) Il est clair que $F \cup G \subset F + G$ car $0 \in F$ et $0 \in G$. Il est de plus aisé de vérifier que $F + G$ est un s–e.v de E, ce qui entraîne par minimalité que $\text{Vect}(F \cup G) \subset F + G$.

($\subset$) Soit H un s–e.v de E contenant $F \cup G$; alors pour tout $(x, y) \in F \times G$, $x, y \in F\cup G \subset H$ et donc $x+y \in H$. Ainsi $F+G \subset H$ et donc $F+G \subset \text{Vect}(F\cup G)$ par minimalité.

□

➠ **Exemple XIX.11.** Dans $\mathbb{K}^2$, $Vect((0, 1)) + \text{Vect}((1, 1)) = \mathbb{K}^2$ et

$$Vect((0, 1)) + \text{Vect}((0, -1)) = \mathscr{D}_{(0,1)} .$$

✌ **Remarque XIX.7.** On peut généraliser cette construction à une famille **finie** $(F_1, \ldots, F_n)$ de s–e.v de E et obtenir :

$$F_1 + \ldots + F_n = \text{Vect}\left(\bigcup_{k=1}^{n} F_k\right) .$$

b) Somme directe

Définition XIX.8. Soient F, G deux s–e.v de E. On dit que F et G sont en **somme directe** si :

$$\forall z \in F + G,\ \exists\,(x, y) \in F \times G,\ z = x + y .$$

Notation. L'ensemble $F + G$ est alors noté $F \oplus G$.

✌ **Remarque XIX.8.** La nouveauté est ici l'unicité de la décomposition selon F et G.

Proposition XIX.12. Soient F et G deux s–e.v de E. Alors :

$$F \text{ et } G \text{ sont en somme directe} \iff F \cap G = \{0\}.$$

Démonstration.

($\Downarrow$) Si $x \in F \cap G$, alors $x = 0 + x = x + 0$ et donc, par unicité de la décomposition, $x = 0$.

($\Uparrow$) Soit $z \in F + G$ que l'on suppose décomposable comme $z = x + y$ et $z = x' + y'$ avec $x, x' \in F$ et $y, y' \in G$. Alors $\underbrace{x - x'}_{\in F} = \underbrace{y' - y}_{\in G}$, ce qui entraîne que $x - x', y' - y \in F \cap G = \{0\}$ d'où le résultat.

□

Exemple XIX.12.
- Si $u, v \in E$ sont non colinéaires, alors $\mathscr{D}_u$ et $\mathscr{D}_v$ sont en somme directe. De plus, $\mathscr{D}_u \oplus \mathscr{D}_v = \mathscr{P}_{u,v}$.
- Posons, dans $\mathbb{R}^{\mathbb{R}}$, $F = \text{Vect}(\cos)$ et $G = \text{Vect}(\sin)$. Alors, si $f \in F \cap G$, il existe $\lambda, \mu \in \mathbb{R}$ tels que $f = \lambda \cos = \mu \sin$. De fait, $f(0) = \lambda$ et $f(0) = 0$ donc $\lambda = 0$ *ergo* $f = 0$. On en déduit que F et G sont en somme directe.

Définition XIX.9. Deux s–e.v F et G de E sont dit supplémentaires si $E = F \oplus G$.

Remarque XIX.9. En pratique, il faut donc démontrer que :
- F et G sont en somme directe ;
- $E \subset F + G$.

Exemple XIX.13. On vérifie aisément que $\mathbb{K}^2 = \text{Vect}((0,1)) \oplus \text{Vect}((1,0))$.

Exercice XIX.2. Soit F l'ensemble des fonctions impaires de $\mathbb{R}$ dans $\mathbb{R}$ et G celui des fonctions paires.

(a) Démontrer que F et G sont des s–e.v de $\mathbb{R}^{\mathbb{R}}$.

(b) Montrer que
$$\mathbb{R}^{\mathbb{R}} = F \oplus G\,.$$

Correction :

(a) Immédiat via la caractérisation des s–e.v.

(b) Commençons par remarquer que si $f \in \mathbb{R}^{\mathbb{R}}$ alors $g : x \mapsto \dfrac{f(x) + f(-x)}{2}$ est paire et $h : x \mapsto \dfrac{f(x) - f(-x)}{2}$ est impaire. Ainsi, $f = h + g \in F + G$ et donc $\mathbb{R}^{\mathbb{R}} = F + G$. ***Cette décomposition est souvent utile ; elle est à retenir.*** *Pour conclure, il suffit de noter que si $f \in F \cap G$ alors pour tout $x \in \mathbb{R}$, $f(x) = f(-x) = -f(x)$ et donc $f = 0$.*

Remarque XIX.10. Cette notion se généralise de la façon suivante : une famille **finie** $(F_1, \ldots, F_n)$ de s–e.v de E est dite en somme directe si :

$$\forall z \in F_1 + \ldots + F_n,\ \exists!\,(x_1, \ldots, x_n) \in F_1 \times \ldots F_n,\ \ z = \sum_{k=1}^{n} x_k\,.$$

On notera alors la somme $F_1 \oplus \ldots \oplus F_n$ ou $\bigoplus_{k=1}^{n} F_k$.

Exemple XIX.14. Dans $\mathbb{K}^n$, les vecteurs $e_i = (\delta_{i,j})_{j \in [\![1,n]\!]}$ engendrent des droites en somme directe. Elles sont même supplémentaires car, pour tout $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ on a :

$$x = \sum_{k=1}^{n} x_k e_k\,.$$

✖ **ATTENTION :** il ne suffit plus d'avoir la nullité des intersections des F_k deux à deux pour qu'ils soient supplémentaires. Prenons l'exemple de $u = (1, 0)$, $v = (0, 1)$ et $w = (1, 1)$ dans $\mathbb{K}^2$. Les droites $\mathscr{D}_u$, $\mathscr{D}_v$ et $\mathscr{D}_w$ sont deux à deux d'intersection nulle mais $w = 0 + 0 + w$ et $w = u + v + 0$: il n'y a donc pas unicité de la décomposition.

c) Projecteurs

Définition XIX.10. Un endomorphisme $p \in \mathscr{L}(E)$ est appelé un **projecteur** si

$$p \circ p = p\,.$$

Exemple XIX.15.
- id_E est un projecteur ;
- $(x, y) \mapsto (x, 0)$ est un projecteur sur $\mathbb{K}^2$;
- de façon plus générale, pour tout $i \in [\![1, n]\!]$, l'application

$$\begin{aligned} \pi_i : \mathbb{K}^n &\longrightarrow \mathbb{K}^n \\ (x_1, \ldots, x_n) &\mapsto (0, \ldots, 0, \underbrace{x_i}_{i}, 0, \ldots, 0) \end{aligned}$$

  est un projecteur.

Proposition XIX.13. Soit $p \in \mathscr{L}(E)$. Alors :

p est un projecteur

$\iff$

il existe deux s–e.v F et G **supplémentaires** dans E tels que :

(i) $\forall x \in F$, $p(x) = x$;

(ii) $\forall x \in G$, $p(x) = 0$.

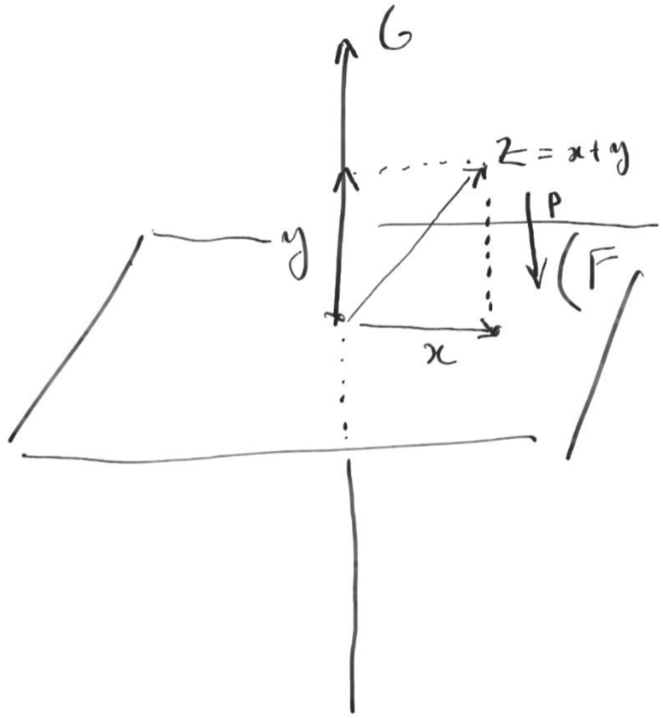
Dans ce cas, F et G sont de plus uniques.

Vocabulaire. On dit alors que p est **le projecteur sur F parallèlement à G.**

✌ **Remarque XIX.11.** Soient F et G deux s–e.v de E et soit p le projecteur sur F relativement à G. Alors, pour tout $z \in E$, il existe un unique couple $(x, y) \in F \times G$ tel que $z = x + y$ et :

$$\begin{aligned} p(z) = 0 &\Leftrightarrow p(x+y) = 0 \\ &\Leftrightarrow p(x) + p(y) = 0 \\ &\Leftrightarrow x + 0 = 0 \\ &\Leftrightarrow z \in G\,. \end{aligned}$$

De plus, $p \circ p(z) = z$ donc p$p(z) = p(x + 0) = p(x) = x$. *In fine*, on a :
($\star$) $F = \mathrm{Im}(p)$;
($\star$) $G = \mathrm{Ker}(p)$.
Ceci est **absolument fondamental** pour bien comprendre les projecteurs et leur action.



Démonstration.

($\Uparrow$) Soit $z \in E$, de décomposition $z = x+y \in F \oplus G$. Alors $p \circ p(z) = p(p(x)+p(y)) = p(x) = x$ et $p(z) = p(x) + p(y) = p(x) = x$, donc p est bien un projecteur.

($\Downarrow$) Supposons que p soit un projecteur et posons

$$G = \mathrm{Ker}(p) \quad \text{ et } \quad F = \mathrm{Ker}(p - \mathrm{id}_E) = \{x \in E \,|\, p(x) = x\}\,.$$

Il est clair que F et G vérifient les conditions (i) et (ii) de la proposition. De plus, si $x \in F \cap G$ alors $p(x) = x$ et $p(x) = 0$, donc $F \cap G = \{0\}$ *ergo* F et G sont en somme directe. Pour conclure quant au caractère supplémentaire de F et G, fixons $z \in E$ et notons que :

$$p(z - p(z)) = p(z) - p \circ p(z) = p(z) - p(z) = 0$$

et donc :

$$z = \underbrace{z - p(z)}_{\in G} + \underbrace{p(z)}_{\in F} \in F \oplus G \, .$$

Pour démontrer l'unicité de F et G, supposons trouvé deux autres candidats F' et G' en sus de ceux évoqués *supra*. Alors, comme p est nul sur G', on a $G' \subset \mathrm{Ker}(p) = G$. On montre similairement que $F' \subset \mathrm{Ker}(p - \mathrm{id}_E) = F$. Réciproquement, si $x \in F$ alors il existe un unique couple $(x_1, x_2) \in F' \times G'$ tel que $x = x_1 + x_2$ et $p(x) = x$, *i.e* $p(x_1 + x_2) = x$. Or, $p(x_1 + x_2) = p(x_1) + p(x_2) = x_1$ donc $x = x_1 \in F'$ d'où $F = F'$. On conclut similairement quant à G et G'. □

✌ **Remarque XIX.12.** On déduit de tout ceci que si p est un projecteur alors $\mathrm{Im}(p) = \mathrm{Ker}(p - \mathrm{id}_E)$. De plus, si p est bijectif alors $\mathrm{Ker}(p) = \{0\}$ et donc $p = \mathrm{id}_E$.

✍ **Exercice XIX.3.** Soient $u = (1, 1)$ et $v = (0, 1)$. Décrire le projecteur sur $\mathscr{D}_u$ parallèlement à $\mathscr{D}_v$ dans $\mathbb{K}^2$.

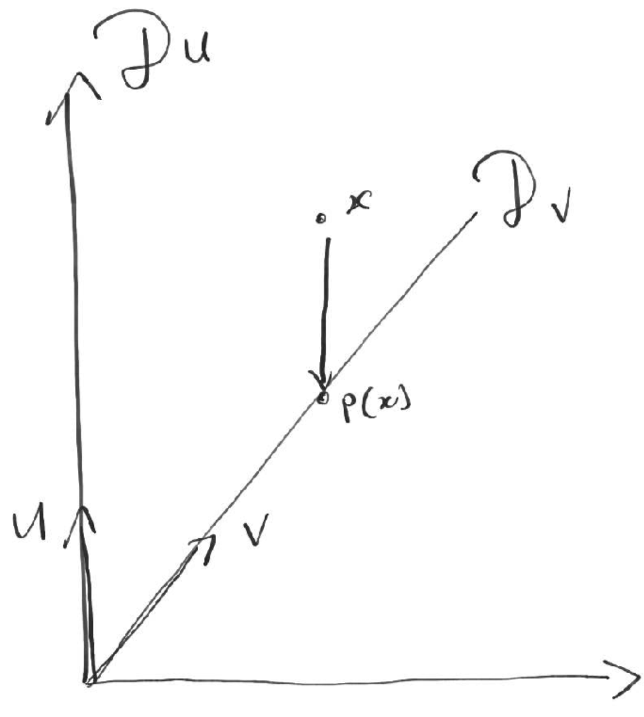
➥ **Correction :** *Commençons par remarquer que les deux droites en question sont bien supplémentaires : elles sont clairement en somme directe (leurs vecteurs directeurs sont non colinéaires) et, pour tout* $(x, y) \in \mathbb{K}^2$ *on a :*

$$(x, y) = x \cdot u + (y - x) \cdot v \, ,$$

expression que le lecteur averti pourra retrouver à l'aide d'un système linéaire en cherchant $\lambda, \mu \in \mathbb{K}$ *tels que* $(x, y) = \lambda u + \mu v$. *De fait, le projecteur* p *sur* $\mathscr{D}_u$ *parallèlement à* $\mathscr{D}_v$ *est donné par :*

$$p : (x, y) \mapsto x \cdot (1, 1) = (x, x) \, .$$

Il "suffit" de conserver la "coordonnée" selon u.



✍ **Exercice XIX.4.** De la même façon, déterminer le projecteur sur $\mathrm{Vect}((1, 0, 0), (0, 1, 0))$ parallèlement à $\mathrm{Vect}((0, 0, 1))$ dans $\mathbb{K}^3$.

d) Symétries

Définition XIX.11. Une application $s \in \mathscr{L}(E)$ est appelée **symétrie** si $s \circ s = \mathrm{id}_E$.

✌ **Remarque XIX.13.** Une symétrie est automatiquement un automorphisme. Il s'agit même d'une involution.

➠ **Exemple XIX.16.**
— id_E est une symétrie ;
— $x \mapsto -x$ est une symétrie.

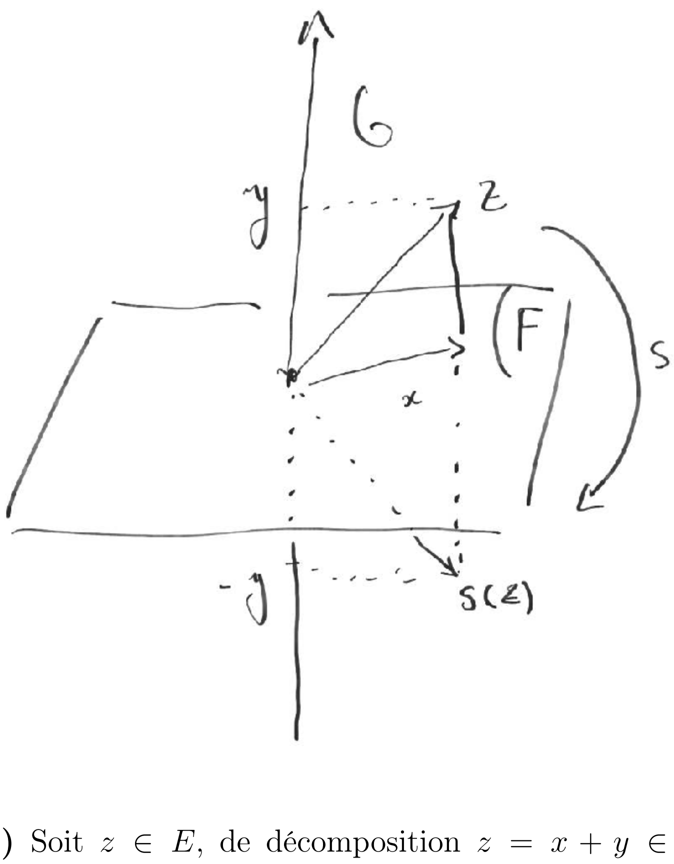
Proposition XIX.14. Soit $s \in \mathscr{L}(E)$. Alors :

s est une symétrie
$\iff$
il existe deux s–e.v F et G **supplémentaires** dans E tels que :

(i) $\forall x \in F,\ s(x) = x$;
(ii) $\forall x \in G,\ s(x) = -x$.

Dans ce cas, F et G sont de plus uniques.

Vocabulaire. On parle de **symétrie par rapport à** F **parallèlement à** G.



Démonstration. ($\Uparrow$) Soit $z \in E$, de décomposition $z = x + y \in F \oplus G$. Alors $s \circ s(z) = s(s(x) + s(y)) = s(x - y) = x + y = z$ donc s est bien une symétrie.

($\Downarrow$) Supposons que s soit une symétrie et posons

$$G = \mathrm{Ker}(s + \mathrm{id}_E) \quad \text{et} \quad F = \mathrm{Ker}(p - \mathrm{id}_E)\,.$$

Il est clair que F et G vérifient les conditions (i) et (ii) de la proposition. De plus, si $x \in F \cap G$ alors $s(x) = x$ et $s(x) = -x$, donc $F \cap G = \{0\}$ *ergo* F et G sont en somme directe. Pour conclure quant au caractère supplémentaire de F et G, fixons $z \in E$ et notons que :

$$z = \underbrace{\frac{z + s(z)}{2}}_{\in F} + \underbrace{\frac{z - s(z)}{2}}_{\in G} \in F + G\,.$$

L'unicité se traite de façon similaire au cas des projecteurs. □

Exemple XIX.17. On retrouve dans $\mathbb{R}^2$, $\mathbb{R}^3$ des procédés vus au collège et au lycée. On peut également chercher à déterminer dans $\mathbb{K}^2$ la symétrie par rapport à $\mathrm{Vect}((1,1))$ parallèlement à $\mathrm{Vect}((0,1))$; on trouve

$$s : (x, y) \mapsto (x, y) = x \cdot (1,1) - (y - x) \cdot (0,1) = (x, 2x - y)\,.$$

e) Hyperplans

Définition XIX.12. Un **hyperplan** dans un $\mathbb{K}$–e.v E est un s–e.v supplémentaire à une droite.

Proposition XIX.15. Soit E un $\mathbb{K}$–e.v et soit H un s–e.v de E. Alors :

$$H \text{ est un hyperplan} \iff \forall a \notin H, \quad H \oplus \mathscr{D}_a = E.$$

Démonstration.

($\Uparrow$) Immédiat par définition.

($\Downarrow$) Supposons que H soit un hyperplan de E. Alors, par définition, il existe $e \in E$ tel que $H \oplus \mathscr{D}_e = E$; il est de plus clair que $e \notin H$ par somme directe. Soit $a \notin H$: alors il existe un unique couple $(h, \lambda) \in H \times \mathbb{K}$ tel que $a = h + \lambda e$. De plus, $\lambda \neq 0$ car $a \notin H$, ce qui permet d'écrire que :

$$e = \frac{1}{\lambda}(a - h)\,.$$

Soit $x \in E$: alors il existe un unique couple $(y, \mu) \in H \times \mathbb{K}$ tel que :

$$x = y + \mu e = \underbrace{y - \frac{\mu}{\lambda} h}_{\in H} + \frac{\mu}{\lambda} a$$

et donc $E = H + \mathscr{D}_a$. Or, $a \notin H$ dont $H \cap \mathscr{D}_a = \{0\}$ ce qui entraîne que $E = H \oplus \mathscr{D}_a$.

□

Proposition XIX.16. Soit E un $\mathbb{K}$–e.v et soit $\varphi \in E^*$ **non nulle**. Alors :

(i) $\mathrm{Ker}(\varphi)$ est un hyperplan de E ;

(ii) $\mathrm{Im}(\varphi) = \mathbb{K}$.

Démonstration. Posons $H = \mathrm{Ker}(\varphi)$ et notons que comme $\varphi \neq 0$, il existe $a \notin H$. On a alors :

— $H \cap \mathscr{D}_a = \{0\}$ car $\mathscr{D}_a$ est une droite non comprise dans H ;

— pour tout $x \in E$:

$$x = \underbrace{x - \frac{\varphi(x)}{\varphi(a)}a}_{\in H} + \frac{\varphi(x)}{\varphi(a)}a$$

et donc $E = H + \mathscr{D}_a$.

En conclusion, H et $\mathscr{D}_a$ sont supplémentaire, d'où le point (i). Pour le point (ii), il nous suffit de remarquer que si $\lambda \in \mathbb{K}$ alors :

$$\lambda = \varphi\left(\frac{\lambda a}{\varphi(a)}\right) .$$

□

➠ **Exemple XIX.18.** L'ensemble $\{P \in \mathbb{K}[X] \,|\, P(\pi) = 0\}$ est un hyperplan de $\mathbb{K}[X]$, noyau de eval_π.

La proposition qui suit offre une réciproque du résultat *supra* en affirmant que tout hyperplan d'un $\mathbb{K}$–e.v est en fait le noyau d'une forme linéaire non nulle.

Proposition XIX.17. Soit H un hyperplan d'un $\mathbb{K}$–e.v E. Alors il existe $\varphi \in E^* \setminus \{0\}$ telle que $H = \mathrm{Ker}(\varphi)$.

Démonstration. Soit $a \notin H$; alors $E = H \oplus \mathscr{D}_a$ et donc l'unique application linéaire envoyant H sur $0_\mathbb{K}$ et a sur $1_\mathbb{K}$ convient. □

✌ **Remarque XIX.14.** Si $H = \mathrm{Ker}(\varphi)$ alors $\{\phi \in E^* \,|\, H = \mathrm{Ker}(\phi)\}$ est la droite $\mathscr{D}_\varphi \subset E^*$.

5.⁻ Familles remarquables de vecteurs

On fixe dans cette partie un $\mathbb{K}$–e.v E.

a) Familles libres, familles génératrices, bases

Définition XIX.13. Soit $\mathscr{F} = (e_i)_i i \in I$ une famille de vecteurs de E. On dit que $\mathscr{F}$ est :
- **libre** si pour tout élément de E s'écrivant comme combinaison linéaire des vecteurs de $\mathscr{F}$, cette écriture est unique ;
- **génératrice** si tout élément de E s'écrit d'**au moins** une façon comme combinaison linéaire des vecteurs de $\mathscr{F}$;
- une **base** de E si elle est libre et génératrice.

✌ **Remarque XIX.15.** $\mathscr{F}$ est une base si et seulement si tout vecteur de E s'écrit d'une unique façon comme combinaison linéaire des e_i.

➠ **Exemple XIX.19.**
- La famille vide $\mathscr{F} = ()$ est libre.
- Si $u \in E$, alors $((u)$ est libre$) \Leftrightarrow (u \neq 0)$.
- Si $u, v \in E$ sont non colinéaires, alors (u, v) est libre.
- Pour tout $A \subset E$, $(a)_{a \in A}$ est génératrice de $\text{Vect}(A)$.
- Dans $\mathbb{K}^n$, la famille $(\varepsilon_i)_{1 \leq i \leq n}$, où
$$\varepsilon_i = (\delta_{i,j})_{j \in [\![1,n]\!]} = (0, \ldots, 0, \underbrace{1}_{i}, 0 \ldots, 0)$$
est une base, appelée **base canonique** de $\mathbb{K}^n$.
- Par construction, la famille $(X^n)_{n \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$, également appelée **base canonique** de cet ensemble.
- Vous l'aurez deviné, $\mathbb{K}_n[X]$ a aussi une base canonique (pas de jaloux) : la famille $(1, X, X^2, \ldots, X^n)$.
- *Quid* de la décomposition en éléments simples, cher lecteur ?

Vocabulaire. Une famille non libre est dite **liée**. De plus, on parle parfois de vecteurs **linéairement indépendants** lorsque l'on considère une famille libre. Enfin, si $(e_i)_{i \in I}$ est une base de E et que $x \in E$ s'y décompose comme :
$$x = \sum_{i \in I} x_i e_i$$
alors les x_i sont appelés **coordonnées** de x dans cette base.

✌ **Remarque XIX.16.** Rappelons au passage que la somme *supra* est en fait finie.

Proposition XIX.18. Soit $\mathscr{F} = (e_i)_{i \in I}$ une famille de vecteurs de E. Alors :
$$\mathscr{F} \text{ est libre}$$
$$\Leftrightarrow$$
$$\text{pour toute famille } (\lambda_i)_{i \in I} \text{ presque nulle de scalaires,}$$
$$\left(\sum_{i \in I} \lambda_i e_i = 0 \right) \Rightarrow (\forall i \in I,\ \lambda_i = 0) .$$

Démonstration.

($\Downarrow$) Découle de l'unicité de la décomposition de 0 selon $\mathscr{F}$.

($\Uparrow$) Supposons que $x \in E$ admette deux décompositions selon $\mathscr{F}$: $x = \sum_{i\in I} \lambda_i e_i$ et $x = \sum_{i\in I} \mu_i e_i$. Alors :

$$0 = x - x = \sum_{i\in I} (\lambda_i - \mu_i) e_i$$

et donc, pour tout $i \in I$, $\lambda_i = \mu_i$.

□

Par conséquent, pour démontrer qu'une famille $(e_i)_{i\in I}$ est libre, on procède comme suit :

— on se donne une famille $(\lambda_i)_i$ de scalaires presque nulle telle que $\sum_{i\in I} \lambda_i e_i = 0$;

— on démontrer que $\forall i \in I,\ \lambda_i = 0$.

➠ **Exemple XIX.20.** La famille $(\cos, \sin)$ est libre dans $\mathbb{R}^{\mathbb{R}}$. En effet, si $\lambda, \mu \in \mathbb{R}$ sont tels que $\lambda \cos + \mu \sin = 0$, alors on a :

$$\forall x \in \mathbb{R},\ \lambda \cos(x) + \mu \sin(x) = 0$$

et donc, en évaluant en 0 (resp. $\frac{\pi}{2}$) on obtient que $\lambda = 0$ (resp. $\mu = 0$).

✌ **Remarque XIX.17.** Il découle de la proposition XIX.18 qu'une famille $\mathscr{F} = (e_i)_{i\in I}$ de vecteurs de E est liée si et seulement si il existe une famille presque nulle de scalaires $(\lambda_i)_i$ telle que :

— il existe $i_0 \in I$ tel que $\lambda_{i_0} \neq 0$;

— $\sum_{i\in I} \lambda_i e_i = 0$.

Ceci est équivalent au fait que :

$$e_{i_0} = \frac{-1}{\lambda_{i_0}} \sum_{i\neq i_0} \lambda_i e_i \ .$$

Une famille est donc liée si et seulement si l'un de ses vecteurs s'exprime en tant que combinaison linéaire des autres.

➠ **Exemple XIX.21.** La famille $((1,0,0),(0,2,0),(1,1,0))$ est liée dans $\mathbb{K}^3$.

✖ **Attention :** ne pas confondre famille liée (non libre) et famille génératrice.

Proposition XIX.19.

(i) Toute sous–famille d'une famille libre est libre ;

(ii) toute sur–famille d'une famille liée est liée ;

(iii) toute sur–famille d'une famille génératrice est génératrice.

Démonstration. Trivial. □

b) Lien aux applications linéaires

Proposition XIX.20. Soit F un $\mathbb{K}$–e.v et soit $f \in \mathscr{L}(E, F)$. Soit $\mathscr{F}$ une famille de vecteurs de E. Alors :

(i) $f(\mathrm{Vect}(\mathscr{F})) = \mathrm{Vect}(f(\mathscr{F}))$;

(ii) (f injective) $\wedge$ ($\mathscr{F}$ libre) $\Rightarrow$ ($f(\mathscr{F})$ libre) ;

(iii) (f surjective) $\wedge$ ($\mathscr{F}$ génératrice) $\Rightarrow$ ($f(\mathscr{F})$ génératrice).

Démonstration. On pose $\mathscr{F} = (e_i)_{i \in I}$. Dans ce cas, on a naturellement $f(\mathscr{F}) = (f(e_i))_{i \in I}$.

(i) Soit $y \in F$. Alors :

$$\begin{aligned} y \in f(\mathrm{Vect}(\mathscr{F})) &\Leftrightarrow \exists\, (\lambda_i)_i \in \mathbb{K}^I \text{ presque nulle telle que } y = f\left(\sum_{i \in I} \lambda_i e_i\right) \\ &\Leftrightarrow \exists\, (\lambda_i)_i \in \mathbb{K}^I \text{ presque nulle telle que } y = \sum_{i \in I} \lambda_i f(e_i) \\ &\Leftrightarrow y \in \mathrm{Vect}(f(\mathscr{F}))\ . \end{aligned}$$

(ii) Supposons que $\mathscr{F}$ soit libre et f injective. Si $(\lambda_i)_i$ est une famille presque nulle de scalaires telle que $\sum\limits_{i \in I} \lambda_i f(e_i) = 0$ alors

$$0 = f\left(\sum_{i \in I} \lambda_i e_i\right)$$

et donc, par injectivité :

$$\sum_{i \in I} \lambda_i e_i = 0\ .$$

De plus, $\mathscr{F}$ est libre, donc on peut conclure que les λ_i sont tous nuls.

(iii) Supposons $\mathscr{F}$ génératrice et f surjective. Alors :

$$\begin{aligned} \mathrm{Vect}(f(\mathscr{F})) &= f(\mathrm{Vect}(\mathscr{F})) \text{ par (i)} \\ &= f(E) \text{ car } \mathscr{F} \text{ est génératrice} \\ &= F \text{ car } f \text{ est surjective} \end{aligned}$$

et donc $f(\mathscr{F})$ est génératrice dans F.

□

Corollaire XIX.20.a. Soit $\mathscr{B}$ une base de E, F un $\mathbb{K}$–e.v et $f \in \mathscr{L}(E, F)$. Alors :

(i) f est injective $\Leftrightarrow$ $f(\mathscr{B})$ est libre ;

(ii) f est surjective $\Leftrightarrow$ $f(\mathscr{B})$ est génératrice ;

(iii) f est bijective $\Leftrightarrow$ $f(\mathscr{B})$ est une base.

Tout ceci nous permet de démontrer ce qui est sans doute le résultat élémentaire le plus important de la théorie des applications linéaires.

Proposition XIX.21. Soit F un $\mathbb{K}$–e.v , soit $\mathscr{B} = (e_i)_{i\in I}$ une base de E et $\mathscr{F} = (f_i)_{i\in I}$ une famille de vecteurs de F indexée par le même ensemble que $\mathscr{B}$. Alors :

$$\exists!\, f \in \mathscr{L}(E,F), \quad \forall i \in I,\ f(e_i) = f_j\,.$$

✌ **Remarque XIX.18.** Une conséquence de ceci est qu'UNE APPLICATION LINÉAIRE EST TOTALEMENT DÉTERMINÉE PAR LA DONNÉE DE L'IMAGE D'UNE BASE DE SON ENSEMBLE DE DÉPART. Ceci est absolument fondamental, et nous permettra de définir de telles applications uniquement par la donnée de leurs valeurs sur les vecteurs d'une base.

Démonstration. Soit $x \in E$; alors x se décompose dans la base $\mathscr{B}$, disons par $x = \sum_{i\in I} x_i e_i$. De fait, si f est application linéaire vérifiant la condition donnée :

$$f(x) = f\left(\sum_{i\in I} x_i e_i\right) = \sum_{i\in I} x_i f(e_i) = \sum_{i\in I} x_i f_i$$

et donc f est entièrement déterminée par la donnée des f_i. Réciproquement, le mécanisme

$$\sum_{i\in I} x_i e_i \mapsto \sum_{i\in I} x_i f_i$$

définit bien une application linéaire (en exercice pour notre lecteur motivé). □

➠ **Exemple XIX.22.**

— L'application linéaire

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (1,0,0) &\mapsto (1,0) \\ (0,1,0) &\mapsto (0,1) \\ (0,0,1) &\mapsto (0,2) \end{aligned}$$

est bien définie. Plus précsiément, si $(x,y,z) \in \mathbb{R}^3$ on a :

$$\begin{aligned} f((x,,y,z)) &= xf((1,0,0)) + yf((0,1,0)) + zf((0,0,1)) \\ &= x\cdot(1,0) + y\cdot(0,1) + z\cdot(0,2) \\ &= (x, y+2z)\,. \end{aligned}$$

Cette application est surjective car l'image de la base canonique de $\mathbb{R}^3$ par f est génératrice de $\mathbb{R}^2$. Elle n'est pas injective car

$$\operatorname{Ker}(f) = \{(0,-2z,z) | z \in \mathbb{R}\} = \operatorname{Vect}((0,-2,1))\,.$$

— L'application linéaire

$$\begin{aligned} g : \mathbb{R}^2 &\longrightarrow \mathbb{R}^{\mathbb{R}} \\ (1,0) &\mapsto \cos \\ (0,1) &\mapsto \sin \end{aligned}$$

est bien définie et injective car l'image de la base canonique de $\mathbb{R}^2$ par g est $(\cos,\sin)$, qui est libre dans $\mathbb{R}^{\mathbb{R}}$.

c) Maximalité, minimalité

Proposition XIX.22. Soit $\mathscr{F}$ une famille de vecteurs de E. Les trois propriétés suivantes sont alors équivalentes :

(i) $\mathscr{F}$ est libre et toutes ses sur–familles sont liées ;

(ii) $\mathscr{F}$ est génératrice et toutes ses sous–familles ne sont pas génératrices ;

(iii) $\mathscr{F}$ est une base de E.

✌ **Remarque XIX.19.** Ceci signifie que les bases sont exactement les familles libres maximales et les familles génératrices minimales. Ceci sera très important dans le chapitre XX.

Démonstration.

(i)⇒(ii) Si $\mathscr{F}$ est libre maximale, alors pour tout $e \in E$, $\mathscr{F} \cup \{e\}$ est liée et donc e s'écrit comme combinaison linéaire des autres vecteurs de $\mathscr{F} \cup \{e\}$, *i.e* des vecteurs de $\mathscr{F}$. La famille est donc génératrice. Pour montrer que $\mathscr{F}$ est génératrice minimale, il nous suffit de remarquer qu'aucun des vecteurs de $\mathscr{F}$ ne peut s'exprimer en fonction des autres (par liberté). De fait, aucune sous–famille de $\mathscr{F}$ ne peut être génératrice.

(ii)⇒(iii) Supposons $\mathscr{F}$ génératrice minimale. Alors, si $\mathscr{F}$ était liée, l'un de ses vecteurs serait combinaison linéaire des autres et donc $\mathscr{F}$ privée de celui–ci serait génératrice, ce qui est absurde. $\mathscr{F}$ est donc une base de E.

(iii)⇒(i) Si $\mathscr{F}$ est une base, elle est libre. De plus, tout vecteur $e \in E$ s'exprime dans $\mathscr{F}$ (elle est génératrice), donc $\mathscr{F} \cup \{e\}$ est liée, ce qui assure la maximalité.

□

d) Bases adaptées à des sous–espaces supplémentaires

Supposons trouvés $F_1, \dots, F_n$ des s–e.v supplémentaires de E, *i.e* tels que

$$E = \bigoplus_{k=1}^{n} F_k.$$

Alors, si on dispose de bases respectives $\mathscr{B}_1, \dots, \mathscr{B}_n$ des F_k, on démontre par une récurrence immédiate que $\bigcup_{k=1}^{n} \mathscr{B}_k$ est une base de E. On parle de **base adaptée** aux sous–espaces $F_1, \dots, F_n$.

➠ **Exemple XIX.23.** Méditer quant à l'expression d'un projecteur, d'une symétrie dans une base adaptée à leur décomposition intrinsèque.

Chapitre XX

Dimension finie

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.– Notion de dimension

a) Espaces de dimension finie

On fixe dans cette partie un $\mathbb{K}$–e.v E.

Définition XX.1. Un $\mathbb{K}$–e.v est dit **de dimension finie** si il admet une famille génératrice finie.

➠ **Exemple XX.1.** Les espaces $\mathbb{K}^n$, $\mathbb{K}_n[X]$ sont de dimension finie pour $n \geq 1$. L'espace $\mathbb{K}[X]$ ne l'est par contre pas car l'existence d'une famille génératrice finie entraînerait une borne sur le degré des polynômes.

Proposition XX.1. Soit $n \geq 1$. Si E admet une famille génératrice de cardinal n, alors toutes les familles libres de E sont finies de cardinal inférieur ou égal à n.

Pour démontrer cette proposition, nous allons (une fois n'est pas coutume) avoir recours à un lemme.

Lemme XX.1. Soit $\mathscr{F} = (e_i)_{1\leq i\leq n}$ et $\mathscr{F}' = (e'_i)_{1\leq i\leq n+1}$ deux familles de vecteurs de E vérifiant que :

$$\forall i \in [\![1, n+1]\!], \quad e'_i \in \mathrm{Vect}(\mathscr{F}) .$$

Alors $\mathscr{F}'$ est liée.

Démonstration. On démontre par récurrence sur $n \in \mathbb{N}^*$ la propriété suivante : pour tout $n \geq 1$, pour toutes familles $\mathscr{F} = (e_i)_{1\leq i\leq n}$ et $\mathscr{F}' = (e'_i)_{1\leq i\leq n+1}$ de vecteurs de E telles que $\forall i \in [\![1, n+1]\!]$, $e'_i \in \mathrm{Vect}(\mathscr{F})$, la famille $\mathscr{F}'$ est liée.

- $n = 1$: trivial.
- Supposons la propriété vraie au rang $n \geq 1$ et soit $\mathscr{F} = (e_i)_{1\leq i\leq n+1}$ et $\mathscr{F}' = (e'_i)_{1\leq i\leq n+2}$ deux familles de vecteurs de E telles que $\forall i \in [\![1, n+2]\!]$, $e'_i \in \mathrm{Vect}(\mathscr{F})$.

Ceci signifie qu'il existe une famille de scalaires $a_{i,j}$ telle que :

$$\forall i \in [\![1, n+2]\!], \quad e'_i = \sum_{j=1}^{n+1} a_{i,j} e_j \ .$$

Si tous les $a_{i,n+1}$ sont nuls, alors les vecteurs de $\mathscr{F}'$ sont dans l'espace engendré par $(e_1, \dots, e_n)$ et on peut conclure par hypothèse de récurrence. Dans le cas contraire, on peut supposer (quitte à réordonner les e_j) que $a_{n+2,n+1}$ est non nul.
Posons, pour $i \in [\![1, n+1]\!]$:

$$e''_i = e'_i - \frac{a_{i,n+1}}{a_{n+2,n+1}} e'_{n+2}$$

et remarquons que :

$$\begin{aligned}
e''_i &= \sum_{j=1}^{n+1} a_{i,j} e_j - \frac{a_{i,n+1}}{a_{n+2,n+1}} \sum_{j=1}^{n+1} a_{n+2,j} e_j \\
&= \sum_{j=1}^{n} a_{i,j} e_j + a_{i,n+1} e_{n+1} - \frac{a_{i,n+1}}{a_{n+2,n+1}} \sum_{j=1}^{n} a_{n+2,j} e_j - \frac{a_{i,n+1}}{a_{n+2,n+1}} a_{n+2,n+1} e_{n+1} \\
&= \sum_{j=1}^{n} \left(a_{i,j} - \frac{a_{i,n+1}}{a_{n+2,n+1}} \right) e_j \\
&\in \mathrm{Vect}(e_1, \dots, e_n) \ .
\end{aligned}$$

On peut ensuite conclure par hypothèse de récurrence que les e''_i sont liés, et en déduire que $\mathscr{F}'$ l'est également.

□

Nous sommes grâce à ce lemme en mesure de démontrer la proposition *supra*.

Démonstration. Soit $\mathscr{G}$ une famille génératrice de E de cardinal n et soit $\mathscr{F}$ une famille libre de cet espace. Si $\mathscr{F}$ admettait plus de n éléments, alors elle devrait être liée (en application du lemme) car $\mathrm{Vect}(\mathscr{G}) = E$. □

Corollaire XX.1.a. Toutes les bases d'un $\mathbb{K}$–e.v de dimension finie sont finies de même cardinal.

Démonstration. Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Comme ces deux familles sont libres, elles sont finies et comme $\mathscr{B}$ est libre et $\mathscr{B}'$ génératrice alors $\mathrm{card}(\mathscr{B}) \leq \mathrm{card}(\mathscr{B}')$. On conclut en échangeant les rôles de $\mathscr{B}$ et $\mathscr{B}'$. □

☞ Nous savons donc que si il y a existence de bases de E, elles sont de même cardinal. Mais existent–elles ?

b) Théorème de la base incomplète

Théorème XX.2 (Base incomplète)**.**
On suppose que E est de **dimension finie**. Soit $\mathscr{F}$ une famille libre de E. Alors il existe une base de E contenant $\mathscr{F}$.

Démonstration. Posons $\mathscr{F} = (e_1, \dots, e_k)$ et fixons $\mathscr{G}$ une famille génératrice finie de E. Alors, quitte à remplacer $\mathscr{G}$ par $\mathscr{G} \cup \mathscr{F}$, on peut supposer que $\mathscr{F}$ est contenue dans $\mathscr{G}$ et poser :

$$\mathscr{G} = (\underbrace{e_1, \dots, e_k}_{\mathscr{F}}, e_{k+1}, \dots e_m)\,.$$

L'ensemble

$$\mathscr{E} = \{\text{card}(A) \,|\, ([\![1, k]\!] \subset A \subset [\![1, m]\!]) \wedge ((e_i)_{i \in A} \text{ est libre })\}$$

est une partie de $\mathbb{N}$ non vide et majorée par $m = \text{card}(\mathscr{G})$. De fait, il est raisonnable de poser $n = \max \mathscr{E}$ associée à un ensemble A_0 contenant $[\![1, k]\!]$ et de remarquer que $\mathscr{B} = (e_i)_{i \in A_0}$ est une famille libre contenant $\mathscr{F}$. De plus, comme $\mathscr{G}$ est génératrice, on a pour tout $x \in E$ l'existence d'une famille $(x_i)_{i \in [\![1, m]\!]}$ de scalaires telle que :

$$\begin{aligned} x &= \sum_{i=1}^{n} x_i e_i \\ &= \sum_{i \in A_0} x_i e_i + \sum_{x \notin A_0} x_i e_i\,. \end{aligned}$$

Cependant, par maximalité de n dans $\mathscr{E}$, si $i \notin A_0$, alors $\mathscr{B} \cup \{e_i\}$ et donc $e_i \in \text{Vect}(\mathscr{B})$. On en déduite que $\mathscr{B}$ est génératrice : il s'agit bien d'une base de E contenant $\mathscr{F}$. □

✌ **Remarque XX.1.** Ce résultat se généralise en dimension infinie, mais ceci est bien au delà de notre niveau technique à ce stade.

Corollaire XX.2.a. Tout $\mathbb{K}$–e.v de dimension finie admet une base.

Démonstration. Appliquer le théorème XX.2 à la famille vide. □

De façon analogue, on peut démontrer un théorème "miroir" au théorème de la base incomplète, permettant d'amaigrir une famille génératrice.

Théorème XX.3 (Base extraite)**.**
On suppose que E est de **dimension finie**. Soit $\mathscr{G}$ une famille génératrice de E. Alors il existe une base de E contenue dans $\mathscr{G}$.

c) Bilan

Nous venons démontrer que tout espace de dimension finie admettait des bases, et que ces dernières étaient toutes de même cardinal. Ceci nous permet de poser sereinement la définition *infra*.

Définition XX.2. Supposons E de dimension finie. On appelle **dimension** de E le cardinal commun à toutes ses bases.

Notation. On notera la dimension de E $\dim(E)$ (ou $\dim_{\mathbb{K}}(E)$ lorsqu'il sera nécessaire de préciser le corps de base).

Exemple XX.2. Les dimensions qui suivent découlent trivialement des exemples de bases vues dans le chapitre XIX :
- $\dim(\{0\}) = 0$ (et il s'agit du seul espace vectoriel de dimension nulle) ;
- si $e \in E \setminus \{0\}$, $\dim(\mathscr{D}_e) = 1$;
- si $u, v \in E$ sont non colinéaires, $\dim(\mathscr{P}_{u,v}) = 2$;
- $\dim(\mathbb{K}^n) = n$;
- $\dim(\mathbb{K}_n[X]) = n + 1$.

✖ **Attention :** $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ mais $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Remarque XX.2. Si E est un $\mathbb{K}$–e.v de dimension finie de dimension n, on a donc naturellement que :
- les familles libres de E sont de cardinal inférieur ou égal à n ;
- tout famille de cardinal **strictement** supérieur à n est liée ;
- les familles génératrices de E sont de cardinal supérieur ou égal à n ou infinies.

On déduit de cette remarque et de la proposition XIX.22 sur les familles libres maximales et génératrices minimales le résultat suivant, qui sauve bien des vies en pratique.

Proposition XX.4. On suppose E de dimension finie $n \in \mathbb{N}$. Alors, pour toute famille $\mathscr{B}$ finie de vecteurs de E, les propriétés suivantes sont équivalentes :

(i) $\mathscr{B}$ est une base ;

(ii) $\mathscr{B}$ est libre de cardinal n ;

(iii) $\mathscr{B}$ est génératrice de cardinal n.

Exemple XX.3. La famille $((1,0,1),(1,0,0),(0,1,0))$ est une base de $\mathbb{K}^3$ car libre et de cardinal n.

✍ **Exercice XX.1.** Démontrer que toute famille de polynômes de degré échelonnée est libre dans $\mathbb{K}[X]$. En déduire que $(1, 2X, X^2 + 18)$ est une base de $\mathbb{K}_2[X]$.

➥ **Correction :** *Soit $\mathscr{F}$ une telle famille ; quitte à réordonner, on peut supposer que $\mathscr{F} = (P_n)_{n \in A}$, avec $A \subset \mathbb{N}$ pour tout $n \geq 0$, $\deg(P_n) < \deg(P_{n+1})$. On démontre ensuite par récurrence sur N que si $(\lambda_n)_{n \in [\![0,N]\!]}$ est une famille de scalaires telle que $\sum_{n=0}^{N} \lambda_n P_n = 0$ alors les λ_n sont tous nuls (s'intéresser au coefficient dominant de la somme).*

2.⁻ Sous–espaces et dimension

On fixe dans cette partie un $\mathbb{K}$–e.v de dimension finie E de dimension $n \in \mathbb{N}$.

a) Dimension d'un s–e.v

Proposition XX.5. Soit F un s–e.v de E. Alors :
(i) F est de dimension finie ;
(ii) $\dim(F) \leq \dim(E)$.
Il y a de plus égalité entre E et F si et seulement si leurs dimensions sont égales.

Démonstration. Toute famille libre de F est libre dans E. Il en découle donc que le cardinal de toute telle famille est inférieur ou égal à n et donc F est de dimension finie inférieure ou égale à n. De plus, $\dim(E) = \dim(F)$ si et seulement si F possède une base $\mathscr{B}$ de cardinal n. Or, dans ce cas, $\mathscr{B}$ est libre dans E et de cardinal n : il s'agit donc d'une base de E *ergo* $E = \mathrm{Vect}(\mathscr{B}) = F$. □

➠ **Exemple XX.4.** On retrouve le résultat "vu" au lycée concernant les dimensions possibles des sous–espaces vectoriels de $\mathbb{R}^2$ $(0, 1, 2)$ et $\mathbb{R}^3$ $(0, 1, 2, 3)$.

✌ **Remarque XX.3.** Ceci nous permet de dire que deux sous–espaces F et G de E sont supplémentaires **si et seulement si** :
(i) $F \cap G = \{0\}$;
(ii) $\dim(F + G) = \dim(E)$.

Définition XX.3. Si F est un s–e.v de E de dimension p, on appelle **base adaptée** à F toute base $(e_1, \ldots, e_n)$ de E telle que $(e_1, \ldots, e_p)$ soit une base de F.

➠ **Exemple XX.5.** $((1, 0, 1), (0, 0, 1), (0, 1, 0))$ est une base de $\mathbb{R}^3$ adaptée à $\mathscr{D}_{(1,0,1)}$.

Vocabulaire.
- Tout sous–espace vectoriel de dimension 1 est appelé **droite** ;
- tout sous–espace vectoriel de dimension 2 est appelé **plan**.

✌ **Remarque XX.4.** Les deux premiers points de vocabulaire *supra* correspondent aux définitions données pour ces objets dans le chapitre XIX : un s–e.v de dimension p de E est en effet un espace vectoriel engendré par une famille **libre** à p éléments de E.

b) Supplémentaires en dimension finie

Proposition XX.6. Soit F un s–e.v de E de dimension p. Alors :
(i) F admet un supplémentaire ;
(ii) tous les supplémentaires de F sont de dimension $n - p$.

Démonstration.

(i) Soit $\mathscr{F}$ une base de F ; alors par théorème de la base incomplète (XX.2), il existe une base $\mathscr{B} = (e_i)_{i\in[\![1,p]\!]}$ de E adaptée à F. Posons alors $G = \text{Vect}(e_{p+1}, \ldots, e_n)$: on a alors par définition $\dim(G) = n - p$ (la famille engendrant G est libre dans E) et $E = F \oplus G$ (car $(e_1, \ldots, e_p)$ est une base de F.

(ii) Si H est un supplémentaire de F de base $\mathscr{H}$, alors $Fr \cup \mathscr{H}$ est de cardinal n par somme directe (il s'agit d'une base de E adaptée à F et H). Il est découle que $\dim(H) = \text{card}(\mathscr{H}) = n - p$.

□

➠ **Exemple XX.6.** Dans $\mathbb{R}^3$, les supplémentaires de plans sont des droites. Dans $\mathbb{R}^2$, il s'agit de droites.

3.⁻ Zoologie dimensionnelle

a) Applications linéaires

Proposition XX.7. Soient E et F deux espaces vectoriels et soit $f \in \mathscr{L}(E, F)$. On se donne E' un s–e.v de E **de dimension finie**. Alors :

(i) $f(E')$ est de dimension finie ;

(ii) $\dim(f(E')) \leq \dim(E')$ avec égalité lorsque f est injective.

Démonstration. Fixons une base $\mathscr{B}$ de E'. $f(E') = \text{Vect}(f(\mathscr{B}))$, ce qui entraîne que $\dim(f(E'))$ est finie ; de plus $\dim(f(E')) = \text{card}(f(\mathscr{B})) \leq \text{card}\mathscr{B} = \dim(E')$ avec égalité lorsque f est injective (*cf.* chapitre XVIII). □

➠ **Exemple XX.7.**

— L'image d'une droite par $f \in \mathscr{L}(\mathbb{R}^2)$ est une droite ou un point ;

— il n'existe aucune surjection linéaire de $\mathbb{R}^2$ vers $\mathbb{R}^4$ (s'intéresser à la dimension de l'image pour obtenir une contradiction).

✌ **Remarque XX.5.** D'une façon générale, cette proposition entraîne que le passage "par" une application linéaire ne peut que **diminuer** la dimension d'un s–e.v, jamais l'augmenter.

Proposition XX.8. Deux $\mathbb{K}$–e.v de dimension finie sont isomorphes **si et seulement si** leurs dimensions sont égales.

Démonstration. Soient E et F deux $\mathbb{K}$–e.v de dimension finie et soit $f \in GL(E, F)$; alors, $\dim(E) = \dim(f(E))$ car f est injective. Or, f est surjective donc $f(E) = F$, d'où le résultat.

Réciproquement, si E et F sont deux $\mathbb{K}$–e.v de dimension finie de même dimension et de bases respectives $\mathscr{B}$ et $\mathscr{B}'$ alors l'unique application linéaire envoyant $\mathscr{B}$ sur $\mathscr{B}'$ est bijective (car elle envoie une base sur une base), d'où le résultat. □

Corollaire XX.8.a. Soit un E un $\mathbb{K}$–e.v de dimension finie et posons $n = \dim(E)$. Alors E est isomorphe à $\mathbb{K}^n$.

Ce résultat n'est pas anodin : il s'agit la d'une classification (à isomorphisme près) des $\mathbb{K}$–e.v de dimension finie par leur dimension : deux espaces vectoriels de même dimension sont de fait "fortement similaires", et l'espace $\mathbb{K}^n$ pourra être utilisé comme "prototype" de ces derniers.

Exemple XX.8. $\mathbb{K}_n[X]$ est isomorphe à $\mathbb{K}^{n+1}$. Un isomorphisme explicite peut même être déterminé par la méthode désormais usuelle du "jetons une base sur son homologue et prions" :

$$\begin{aligned} \varphi : \mathbb{K}_n[X] &\to \mathbb{K}^{n+1} \\ X^k &\mapsto (\delta_{i,k+1})_{1\leq i\leq n} \; . \end{aligned}$$

Proposition XX.9. Soient E et F deux $\mathbb{K}$–e.v de dimension finie. Alors :
(i) $\mathscr{L}(E,F)$ est de dimension finie ;
(ii) $\dim(\mathscr{L}(E,F)) = \dim(E) \times \dim(F)$.

Démonstration. Pour reprendre une expression tristement célèbre en politique, "il suffit de" construire une base de $\mathscr{L}(E,F)$. Le procédé est hélas relativement douloureux.

Commençons par fixer $\mathscr{B} = (e_1, \dots, e_p)$ et $\mathscr{B}' = (e'_1, \dots, e'_n)$ des bases respectives de E et F (et donc $p = \dim(E)$ et $n = \dim(F)$). Posons, pour $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$:

$$\begin{aligned} g_{i,j} : E &\to F \\ e_k &\mapsto \delta_{j,k} e'_i \end{aligned}$$

de façon à ce que $g_{i,j}(e_j) = e'_i$. Faites moi confiance, et il n'y aura pas de blessés. Posons $\mathscr{F} = (g_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ et démontrons que cette famille constitue une base de $\mathscr{L}(E,F)$.

Soit $f \in \mathscr{L}(E,F)$ et soit $x = \sum\limits_{j=1}^{p} x_j e_j \in E$. Alors :

$$\begin{aligned} f(x) &= f\left(\sum_{j=1}^{p} x_j e_j\right) \\ &= \sum_{j=1}^{p} x_j f(e_j) \; . \end{aligned}$$

Pour tout $j \in [\![1,p]\!]$, $f(e_j) \in F$: il existe donc une unique famille $(a_{i,j})_{i\in[\![1,n]\!]}$ de scalaires telle que :

$$f(e_j) = \sum_{i=1}^{n} a_{i,j} e'_i$$

et donc

$$\begin{aligned} f(x) &= \sum_{j=1}^{p} x_j \sum_{i=1}^{n} a_{i,j} e'_i \\ &= \sum_{j=1}^{p} \sum_{i=1}^{n} a_{i,j} x_j e'_i \,. \end{aligned}$$

Gardons cela en tête le temps de remarquer que, pour $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$, on a :

$$\begin{aligned} g_{i,j}(x) &= g_{i,j}\left(\sum_{k=1}^{p} x_k e_k\right) \\ &= \sum_{k=1}^{p} x_k g_{i,j}(e_k) \\ &= \sum_{k=1}^{p} x_k \delta_{j,k} e'_i \\ &= x_j e'_i \end{aligned}$$

ce qui entraîne que :

$$f(x) = \sum_{j=1}^{p} \sum_{i=1}^{n} a_{i,j} g_{i,j}(x)$$

i.e

$$f = \sum_{j=1}^{p} \sum_{i=1}^{n} a_{i,j} g_{i,j} \in \mathrm{Vect}(\mathscr{F}) \,.$$

On en déduit que $\mathscr{F}$ est génératrice dans $\mathscr{L}(E,F)$.

Donnons nous ensuite une famille $(\lambda_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ de scalaires telle que

$$\sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_{i,j} g_{i,j} = 0 \,.$$

Alors, pour tout $k \in [\![1,p]\!]$, on a :

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_{i,j} g_{i,j}(e_k) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_{i,j} \delta_{j,k} e'_i \\ &= \sum_{i=1}^{n} \lambda_{i,k} e'_i \,. \end{aligned}$$

Or, la famille $\mathscr{B}'$ est libre, ce qui entraîne que les $\lambda_{i,k}$ sont tous nuls. On en déduit que $\mathscr{F}$ est libre et donc une base, ce qui permet de conclure car $\mathrm{card}(\mathscr{F}) = p \times n = \dim(E) \times \dim(F)$. □

✌ **Remarque XX.6.** Soit E un $\mathbb{K}$–e.v de dimension $n \in \mathbb{N}$. Alors :

— $\dim(\mathscr{L}(E) = n^2$;

— $\dim(E^*) = n$ et donc E et E^* sont isomorphes.

➠ **Exemple XX.9.** $\dim(\mathscr{L}(\mathbb{K}_n[X], \mathbb{K}^n)) = (n+1)n$.

b) Somme et produit de s–e.v

Proposition XX.10. Soit E un $\mathbb{K}$–e.v de dimension finie et soient $F_1, \ldots, F_n$ des s–e.v de E. Alors :

$$\dim(\left(\sum_{k=1}^{n} F_k\right) \leq \sum_{k=1}^{n} \dim(F_k)$$

avec égalité **si et seulement si** les F_k sont en somme directe.

Démonstration. Pour tout $k \in [\![1, n]\!]$, fixons une base $\mathscr{B}_k$ de F_k. Alors la famille $\mathscr{F} = \bigcup_{k=1}^{n} \mathscr{B}_k$ est génératrice de $F_1 + \ldots + F_n$ ce qui entraîne que :

$$\begin{aligned} \dim\left(\sum_{k=1}^{n} F_k\right) &\leq \operatorname{card}\left(\bigcup_{k=1}^{n} \mathscr{B}_k\right) \\ &\leq \sum_{k=1}^{n} \operatorname{card}(\mathscr{B}_k) \\ &= \sum_{k=1}^{n} \dim(F_k) . \end{aligned}$$

Il y a de plus égalité si et seulement si les $\mathscr{B}_k$ sont deux à deux disjoints (*cf.* chapitre XVIII), ce qui est équivalent à la liberté de $\mathscr{F}$, et donc au fait que les F_k soient en somme directe. ☐

➠ **Exemple XX.10.** Deux droites sont en somme directe si et seulement si elles engendrent un plan. De même, un plan et une droite le sont à condition d'engendrer un espace de dimension 3.

Corollaire XX.10.a. Les hyperplans d'un $\mathbb{K}$–e.v de dimension finie de dimension $n \in \mathbb{N}^*$ sont exactement ses s–e.v de dimension $n - 1$.

Démonstration. Ceci découle du fait qu'un hyperplan est supplémentaire à une droite. ☐

Proposition XX.11. Soient E et F deux $\mathbb{K}$–e.v de dimension finie ; alors $\dim(E \times F) = \dim(E) + \dim(F)$.

Démonstration. Il suffit de vérifier que si $(e_1, \ldots, e_p)$ et $(e'_1, \ldots, e'_n)$ sont des bases respectives de E et F alors $((e_1, 0), \ldots, (e_p, 0), (0, e'_1), \ldots, (0, e'_n))$ est une base de $E \times F$. ☐

✌ **Remarque XX.7.**

— Ce résultat se généralise trivialement à un produit cartésien de n $\mathbb{K}$–e.v de dimension finie pour $n \geq 3$.
— Si $F_1, \ldots, F_n$ sont des s–e.v de dimension d'un $\mathbb{K}$–e.v E, alors l'application suivante est un isomorphisme :

$$\begin{aligned} \tau : F_1 \times \ldots \times F_n &\to \bigoplus_{k=1}^{n} F_k \\ (x_1, \ldots, x_n) &\to \sum_{k=1}^{n} x_k \ . \end{aligned}$$

En dimension finie, on retrouve ainsi l'égalité dimensionnelle évoquée *supra*.

4.⁻ Rang

a) C'est quoi ?

Définition XX.4. Soit E un $\mathbb{K}$–e.v et soit $\mathscr{F}$ une famille de vecteurs de E. On appelle **rang** de $\mathscr{F}$ la dimension (éventuellement infinie) de $\mathrm{Vect}(\mathscr{F})$.

Notation. $\mathrm{rg}(\mathscr{F})$

Exemple XX.11. Dans $\mathbb{R}^2$, $\mathrm{rg}((0,1),(0,2)) = 1$. Dans $\mathbb{K}[X]$, $\mathrm{rg}(X, X^3+X) = 2$.

Remarque XX.8. Si $\mathscr{F}$ est finie, alors $\mathrm{rg}(\mathscr{F}) \leq \mathrm{card}(\mathscr{F})$ avec égalité si et seulement si $\mathscr{F}$ est libre.

Définition XX.5. Soient E et F deux $\mathbb{K}$–e.v avec F de dimension finie et soit $f \in \mathscr{L}(E,F)$. On appelle **rang** de f la dimension de $\mathrm{Im}(f)$.

Notation. $\mathrm{rg}(f)$

Remarque XX.9. L'introduction simultanée de ces deux notions peut sembler étonnante et propice à confusion. Rassurons nous toutefois : si $\mathscr{B}$ est une base de E, alors on a, par définition(s) :

$$\mathrm{rg}(f) = \mathrm{rg}(f(\mathscr{B})) \ .$$

Je suis convaincu que le lecteur est désormais soulagé.

Exemple XX.12.
— $\mathrm{rg}(\mathrm{id}_E) = \dim(E)$;
— plus généralement, toute bijection est de rang égal à la dimension de son espace de départ/arrivée ;
— si p est le projecteur sur F parallèlement à G (ces espaces étant de dimension finie), alors $\mathrm{rg}(p) = \dim(\mathrm{Im}(p)) = \dim(F)$.

Remarque XX.10. f est surjective si et seulement si $\mathrm{rg}(f) = \dim(F)$. En effet, $\mathrm{rg}(f) = \dim(f(E)) \ldots$

◇ **Calcul pratique du rang d'une famille de vecteurs**

Soit E un $\mathbb{K}$–e.v de dimension $p \in \mathbb{N}$ et soit $\mathscr{B} = (e_1, \dots, e_p)$ une base de E. On fixe une famille $\mathscr{F} = (u_1, \dots, u_n)$ de vecteurs de E s'écrivant, pour tout $j \in [\![1, n]\!]$:

$$u_j = \sum_{i=1}^{p} a_{i,j} e_i$$

dans la base $\mathscr{B}$.

Afin d'aider au calcul du rang de $\mathscr{F}$, nous énonçons les faits suivants, qui seront démontrés au chapitre XXIV : on ne modifie pas le rang de $\mathscr{F}$ si on...

(A) ...on échange la position de deux vecteurs dans la famille $(u_i \leftrightarrow u_j)$;

(B) ...on multiplie un vecteur par un scalaire **non nul** $(u_i \leftarrow \lambda u_j)$;

(C) ...on ajoute à un vecteur une combinaison linéaire des **autres** $\left(u_i \leftarrow u_i + \sum\limits_{j \neq i} \lambda_j u_j\right)$.

Exemple XX.13. Pour déterminer le rang de la famille $((1,2,5),(2,1,4),(1,-1,-1))$ dans $\mathbb{K}^3$, on effectue les opérations élémentaires suivantes : partant de

$$\begin{array}{c|c|c} 1 & 2 & 1 \\ 2 & 1 & -1 \\ 5 & 4 & -1 \end{array}$$

on soustrait à la deuxième colonne deux fois la première $(C_2 \leftarrow C_2 - 2C_1)$ et à la troisième la première $(C_3 \leftarrow C_3 - C_1)$, obtenant

$$\begin{array}{c|c|c} 1 & 0 & 0 \\ 2 & -3 & -3 \\ 5 & -6 & -6 \end{array}$$

puis on effectue $C_3 \leftarrow C_3 - C_2$

$$\begin{array}{c|c|c} 1 & 0 & 0 \\ 2 & -3 & 0 \\ 5 & -6 & 0 \end{array}$$

ce qui permet de conclure que le rang de notre famille est égal à celui de $((1,2,5),(0,-3,-6),(0,0,0))$, à savoir 2.

b) Théorème du rang

Proposition XX.12. Soient E et F deux $\mathbb{K}$–e.v de dimension finie et soit $f \in \mathscr{L}(E,F)$. Alors, pour tout supplémentaire G de $\operatorname{Ker}(f)$, l'application

$$\begin{aligned} \varphi : G &\to \operatorname{Im}(f) \\ x &\mapsto f(x) \end{aligned}$$

est un isomorphisme.

Démonstration. Remarquons tout d'abord que l'application φ est bien définie et linéaire par construction. De plus :

$$\begin{aligned}\operatorname{Ker}(\varphi) &= \{x \in G \mid f(x) = 0\} \\ &= G \cap \operatorname{Ker}(f) \\ &= \{0\}\end{aligned}$$

donc φ est injective. Enfin, notons que si $y \in \operatorname{Im}(f)$ alors $\exists x \in E$, $f(x) = y$ et, comme $E = \operatorname{Ker}(f) \oplus G$ il existe un unique couple $(u, v) \in \operatorname{Ker}(f) \times G$ tel que : $y = f(u + v) = f(v)$ et donc φ est surjective. □

Théorème XX.13 (Théorème du rang)**.**
Soient E et F deux $\mathbb{K}$–e.v de dimension finie et soit $f \in \mathscr{L}(E, F)$. Alors :

$$\dim(E) = \dim(\operatorname{Ker}(f)) + \operatorname{rg}(f)\,.$$

Démonstration. Soit G un supplémentaire de $\operatorname{Ker}(f)$; alors $\dim(E) = \dim(\operatorname{Ker}(f)) + \dim(G)$. Or, la proposition XX.12 entraîne que $\dim(G) = \dim(\operatorname{Im}(f)) = \operatorname{rg}(f)$, d'où le résultat. □

✖ **Attention :** ce théorème, fort utile au demeurant, ne signifie pas que si $f \in \mathscr{L}(E)$ alors $\operatorname{Ker}(f)$ et $\operatorname{Im}(f)$ sont supplémentaires. Prendre par exemple

$$\begin{aligned} f : \mathbb{K}^2 &\longrightarrow \mathbb{K}^2 \\ (x, y) &\mapsto (y, 0)\end{aligned}$$

qui vérifie $\operatorname{Ker}(f) = \operatorname{Im}(f) = \{(x, 0) \mid x \in \mathbb{K}\}$.

Proposition XX.14. Soient E et F deux $\mathbb{K}$–e.v de dimension finie **de même dimension** et soit $f \in \mathscr{L}(E, F)$. Alors les trois propriétés suivantes sont équivalentes :

(i) f est injective ;

(ii) f est surjective ;

(iii) f est bijective.

Démonstration.

(i)⇒(ii) Si f est injective alors $\dim(E) = \operatorname{rg}(f)$ par théorème du rang (XX.13) et comme $\dim(E) = \dim(F)$ on a bien f surjective.

(ii)⇒(iii) On a dans ce cas $\dim(\operatorname{Ker}(f)) = \dim(E) - \operatorname{rg}(f) = \dim(F) - \operatorname{rg}(f) = 0$, d'où f est injective donc bijective.

(iii)⇒(i) Trivial.

□

➠ **Exemple XX.14.** Ce résultat facilite **considérablement** notre travail lorsque nous souhaitons démontrer le caractère bijectif d'une application linéaire pour peu que les étoiles (et les dimensions) soient alignées.

— l'application

$$\begin{aligned} f : \mathbb{K}^3 &\longrightarrow \mathbb{K}_2[X] \\ (a,b,c) &\mapsto aX^2 + bX + c \end{aligned}$$

est bijective : son noyau est aisé à déterminer et $\dim(\mathbb{K}_2[X]) = 3 = \dim(\mathbb{K}^3)$.

— De même, on montre aisément que l'application

$$\begin{aligned} \varphi : \mathbb{K}_n[X] &\longrightarrow \mathbb{K}^{n+1} \\ P &\mapsto (P(a_0), \dots, P(a_n)) . \end{aligned}$$

est bijective pour $a_0, \dots, a_n \in \mathbb{K}$ deux à deux distincts. Le lecteur avisé subira une impression massive de déjà vu.

La formule qui suit est attribuée à Hermann Günther Grassmann, mathématicien et indianiste (spécialiste des langues et civilisations du sous–continent indien) allemand (1809—1877).

Proposition XX.15. Soient F et G deux s–e.v d'un $\mathbb{K}$–e.v de dimension finie E. Alors :

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

Démonstration. Il suffit de remarquer que l'application

$$\begin{aligned} f : F \times G &\to F + G \\ (x,y) &\mapsto x + y \end{aligned}$$

est surjective, de noyau

$$\operatorname{Ker}(f) = \{(x, -x) \mid x \in F \cap G\}$$

isomorphe à $F \cap G$ et d'appliquer le théorème du rang (XX.13) en se rappelant que $\dim(F \times G) = \dim(F) + \dim(G)$. □

➠ **Exemple XX.15.** Soient $\mathscr{P}$ et $\mathscr{P}'$ deux plan de $\mathbb{R}^3$. Alors :

$$\dim(\mathscr{P} \cap \mathscr{P}') = \dim(\mathscr{P}) + \dim(\mathscr{P}') - \dim(\mathscr{P} + \mathscr{P}') = 4 - \dim(\mathscr{P} + \mathscr{P}') .$$

Or $\dim(\mathscr{P} + \mathscr{P}') = 3$ si les deux plans ne sont pas confondus et 2 sinon. Ceci entraîne que l'intersection de deux plans non confondus de $\mathbb{R}^3$ est une droite.

5.⁻ Dualité

Rappelons avant de débuter ce paragraphe que si E est un $\mathbb{K}$–e.v, son **dual** E^* est l'ensemble

$$E^* = \mathscr{L}(E, \mathbb{K})$$

des **formes linéaires** sur E. Nous avons également vu que si E est de dimension finie, alors $\dim(E) = \dim(E^*)$, ce qui entraîne que E et E^* sont isomorphes.

a) Base duale

Proposition/définition XX.6. Soit E $\mathbb{K}$–e.v de dimension finie de dimension $n \in \mathbb{N}$ et soit $\mathscr{B} = (e_i)_{i\in[\![1,n]\!]}$ une base de E. Alors, les formes linéaires suivantes forment une base de E^*, appelée **base duale de** $\mathscr{B}$:

$$e_i^* : E \to \mathbb{K}$$
$$e_j \mapsto \delta_{i,j}$$

pour $i \in [\![1,n]\!]$.

Notation. On note $\mathscr{B}^*$ la base duale de $\mathscr{B}$.

✌ **Remarque XX.11.** Si $x = \sum_{i=1}^{n} x_i e_i \in E$, on a, pour $j \in [\![1,n]\!]$:

$$e_j^*(x) = \sum_{i=1} x_i \underbrace{e_j^*(e_i)}_{=\delta_{i,j}} = x_j \, .$$

Les e_j^* sont de fait appelée **formes coordonnées** relativement à la base $\mathscr{B}$.

Démonstration. Il suffit de montrer que cette famille est libre, étant donné l'égalité $\dim(E) = \dim(E^*)$. Si l'on suppose trouvée une famille $(\lambda_i)_{i\in[\![1,n]\!]}$ de scalaires telle que $\sum_{i=1}^{n} \lambda_i e_i^* = 0$ alors, pour tout $j \in [\![1,n]\!]$:

$$0 = \sum_{i=1}^{n} \lambda_i e_i^*(e_j) = \lambda_j$$

d'où le résultat. □

➠ **Exemple XX.16.**

— Considérons la base $((1,1)(0,1))$ de $\mathbb{K}^2$; si $(x,y) \in \mathbb{K}^2$ alors

$$(x,y) = x \cdot (1,1) + (y-x) \cdot (0,1)$$

ce qui entraîne que la base duale de cette dernière est composée de $(x,y) \mapsto x$ et $(x,y) \mapsto y - x$.

— Soient $x_0, \ldots, x_n \in \mathbb{K}$ deux à deux distincts : on dispose alors de la famille des polynômes de Lagrange associée à ceux ci, en l'occurrence, pour $i \in [\![0,n]\!]$:

$$L_i = \prod_{j\neq i} \frac{X - x_j}{x_i - x_j} \in \mathbb{K}_n[X] \, .$$

Par existence du polynôme interpolateur de Lagrange (*cf.* chapitre XV), la famille $(L_0, \ldots, L_n)$ est génératrice de $\mathbb{K}_n[X]$ et de cardinal égal à la dimension de cet espace : il s'agit donc d'une base. Pour déterminer sa base duale, notons que pour tous $i, j \in [\![0,n]\!]$ nous devrions avoir :

$$L_i^*(L_j) = \delta_{i,j}$$

et donc, si $P \in \mathbb{K}_n[X]$ alors $P = \sum_{j=0}^{n} P(x_j)L_j$ ce qui entraîne que :

$$L_i^*(P) = \sum_{j=0}^{n} P(x_j)L_i^*(L_j) = P(x_i)\,.$$

La base duale des $(L_0, \dots, L_n)$ est donc composée des morphismes d'évaluation en les x_i.

b) Lien aux hyperplans

On fixe dans ce paragraphe un $\mathbb{K}$–e.v de dimension finie E de dimension $n \in \mathbb{N}$.

Si H est un hyperplan de E, alors H est le noyau d'une forme linéaire non nulle $\varphi \in E^*$. Fixons $(e_i)_{i\in[\![1,n]\!]}$ une base de E : alors il existe une unique famille $(\lambda_i)_{i\in[\![1,n]\!]}$ de scalaires telle que :

$$\varphi = \sum_{i=1}^{n} \lambda_i e_i^*$$

avec de plus, pour tout $j \in [\![1, n]\!]$:

$$\varphi(e_j) = \sum_{i=1}^{n} \lambda_i \underbrace{e_i^*(e_j)}_{=\delta_{i,j}} = \lambda_j$$

ce qui permet d'écrire :

$$\varphi = \sum_{i=1}^{n} \varphi(e_i)e_i^*\,.$$

Soit $x = \sum_{i=1}^{n} x_i e_i \in E$; alors on a :

$$\begin{aligned} x \in H &\Leftrightarrow \varphi(x) = 0 \\ &\Leftrightarrow \sum_{i=1}^{n} \varphi(e_i)e_i^*(x) = 0 \\ &\Leftrightarrow \sum_{i=1}^{n} \varphi(e_i)x_i = 0\,. \end{aligned}$$

Cette dernière égalité est appelée **équation cartésienne** de l'hyperplan $\mathscr{H}$.

➠ **Exemple XX.17.** $x + y - 3z = 0$ est l'équation d'un plan dans $\mathbb{R}^3$, noyau de $(x, y, z) \mapsto x + y - 3z$. Une base en est (par exemple) $(1, 2, 1), (1, -1, 0))$.

De fait, un hyperplan dans E correspond à l'espace des solutions d'une équation **linéaire** à n inconnues. On en déduit que les solutions d'un système linéaire est une intersection d'hyperplans : nous explorerons en détail cet aspect des choses dans le chapitre XXIV, mais nous donnons ici un résultat préliminaire en guise d'apéritif.

Proposition XX.16. Soient $H_1, \dots, H_m$ des hyperplans de E. Alors :

$$\dim\left(\bigcap_{i=1}^{m} H_i\right) \geq n - m\ .$$

Réciproquement, si F est un s–e.v de E de dimension $n-m$, alors il existe m hyperplans $\mathscr{H}_1, \dots, \mathscr{H}_m$ de E tels que :

$$F = \bigcap_{k=1}^{m} \mathscr{H}_k\ .$$

Démonstration. Soient $\varphi_1, \dots, \varphi_m \in E^*\backslash\{0\}$ telles que pour tout $i[\![1, m]\!]$, $H_i = \operatorname{Ker}(\varphi)$. On obtient le premier point en appliquant le théorème du rang à l'application

$$\begin{aligned} f : E &\to \mathbb{K}^m \\ x &\mapsto (\varphi_1(x), \dots, \varphi_m(x)) \end{aligned}$$

dont le noyau est égal à l'intersection des H_i et qui est surjective. Pour la réciproque, fixons $\mathscr{B} = (e_1, \dots, e_n)$ une base de E adaptée à F et posons, pour $k \in [\![1, m]\!]$ φ_k comme étant l'unique forme linéaire envoyant e_{n-m+k} sur 0 est les autres e_i sur 1. Il suffit alors de poser $\mathscr{H}_k = \operatorname{Ker}(\varphi_k)$. $\square$

➠ **Exemple XX.18.** Pour E un $\mathbb{K}$–e.v de dimension finie et $x \in E$, on pose

$$\begin{aligned} \phi_x : E^* &\to \mathbb{K} \\ f &\mapsto f(x)\ . \end{aligned}$$

On remarque que $\phi_x \in (E^*)^*$ (ou E^{**}, ensemble appelé **bidual de** E) puis on pose

$$\begin{aligned} \phi : E &\to E^{**} \\ x &\mapsto \phi_x\ . \end{aligned}$$

Alors :

$$\begin{aligned} x \in \operatorname{Ker}(\phi) &\Leftrightarrow \phi_x = 0 \\ &\Leftrightarrow \forall f \in E^*,\ \phi_x(f) = 0 \\ &\Leftrightarrow \forall f \in E^*,\ f(x) = 0 \end{aligned}$$

ce qui entraîne que x est dans tous les hyperplans de E et est donc nul. Comme $\dim(E^{**}) = \dim(E)$, on en déduit que E est (canoniquement) isomorphe à son bidual. Chose étonnante, pour obtenir un isomorphisme canonique entre E et son dual, il nous faudra nous restreindre au cas euclidien (*cf.* chapitre XXVI).

Chapitre XXI

Intégration

0.$^{-}$ Continuité uniforme

On fixe dans ce paragraphe un intervalle I d'intérieur non vide.

Définition XXI.1. Une fonction $f : I \to \mathbb{R}$ est dite **uniformément continue** (u.c) si :

$$\forall \varepsilon > 0,\, \exists\, \delta > 0,\, \forall x, y \in I,\, (|x - y| \leq \delta) \Rightarrow (|f(x) - f(y)| \leq \varepsilon)\,.$$

➠ **Exemple XXI.1.** Si $a \in \mathbb{R}^*$, la fonction $x \mapsto ax$ est u.c : il suffit, à ε fixé de poser $\delta = \dfrac{\varepsilon}{|a|}$.

✌ **Remarque XXI.1.** Notons la différence avec la définition de continuité sur I, qui est (rappelons le) :

$$\forall x \in I,\, \forall \varepsilon > 0,\, \exists\, \delta > 0,\, \forall y \in I,\, (|x - y| \leq \delta) \Rightarrow (|f(x) - f(y)| \leq \varepsilon)\,.$$

Dans le cas de la continuité "simple", δ dépend du point x où l'on étudie la continuité. Dans le cas uniforme, il est… uniforme relativement à celui–ci.

Proposition XXI.1. Toute fonction u.c sur I est continue sur cet intervalle.

Démonstration. Découle immédiatement de la remarque *supra*. □

✖ **Attention :** la réciproque est **fausse**. Pour un contre exemple, considérons la fonction (continue) $x \mapsto x^2$ sur $\mathbb{R}_+$ et fixons $\delta > 0$. Comme $(2x + \delta)\delta \xrightarrow[x\to\infty]{} \infty$, il existe $x \in \mathbb{R}_+$ tel que

$$(2x + \delta)\delta > 1\,.$$

En posant $y = x + \delta$ on a à la fois $|x - y| \leq \delta$ et

$$|x^2 - y^2| = |x^2 - x^2 - 2x\delta - \delta^2| = (2x + \delta)\delta > 1$$

ce qui contredit la définition de continuité uniforme pour $\varepsilon = 1$.

Proposition XXI.2. Toute fonction lipschitzienne est u.c.

Démonstration. Soit f une fonction lipschitzienne : il existe donc $M \in \mathbb{R}_+^*$ tel que :

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

Soit $\varepsilon > 0$; alors, en posant $\delta = \dfrac{\varepsilon}{M}$ on a bien

$$\forall x, y \in I, (|x - y| \leq \delta) \Rightarrow (|f(x) - f(y)| \leq \varepsilon)\ .$$

□

✖ **Attention :** une fois encore, la réciproque est fumeuse. L'exercice *infra* fournit un contre exemple (presque) gratuit.

✍ **Exercice XXI.1.** On considère la fonction

$$\begin{aligned} g : \mathbb{R}_+ &\to \mathbb{R} \\ x &\mapsto \sqrt{x} \end{aligned}\ .$$

1. (a) Démontrer que pour tous $x, y \in \mathbb{R}_+$,

   $$|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}\ .$$

   (b) En déduire que g est u.c. sur $\mathbb{R}_+$.
2. Démontrer g n'est pas lipschitzienne sur $\mathbb{R}_+$. On pourra considérer le taux d'accroissement de g en 0.

Le point culminant (et pertinent vis à vis du programme) est le résultat suivant, dont la démonstration est admise. Il est du au mathématicien allemand Eduard Heine (1821—1881).

Théorème XXI.3 (Heine)**.**
Toute fonction continue sur un segment y est uniformément continue.

✍ **Exercice XXI.2.** La démonstration de ce résultat est tout à fait abordable en MPSI. Le lecteur curieux pourra s'y essayer via cet exercice. Fixons deux réels a et b tels que $a < b$ et une fonction continue $f : [a, b] \to \mathbb{R}$ et procédons par l'absurde en supposant que f n'est pas u.c. sur $[a, b]$.

1. Démontrer qu'il existe dans ce cas un réel $\varepsilon > 0$ et deux suites $(x_n)_n, (y_n)_n$ d'éléments de $[a, b]$ tels que

   $$\forall n \geq 1, \qquad \left(|x_n - y_n| \leq \frac{1}{n}\right) \wedge\ (|f(x_n) - f(y_n)| > \varepsilon)\ .$$

2. (a) Montrer qu'il existe une fonction strictement croissante $\varphi : \mathbb{N} \to \mathbb{N}$ telle que $(x_{\varphi(n)})_n$ converge vers un nombre réel ℓ.

   (b) Justifier que $\ell \in [a, b]$.

   (c) Démontrer que la suite $(y_{\varphi(n)})_n$ converge également vers ℓ.

   (d) Déterminer la limite, quand n tend vers l'infini, de la quantité $|f(y_{\varphi(n)}) - f(x_{\varphi(n)})|$.
3. Conclure.

1.⁻ Intégrale des fonctions en escalier

On fixe dans ce paragraphe $a, b \in \mathbb{R}$ tels que $a < b$.

a) Subdivisions, fonctions en escalier

Définition XXI.2. On appelle **subdivision** du segment $[a, b]$ toute famille $\sigma = (x_0, \dots, x_n)$ de points de celui–ci tels que :

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

✌ **Remarque XXI.2.** Il s'agit donc de "délimiter" n sous–segments au segment $[a, b]$.

Notation. On notera $S(a, b)$ l'ensemble des subdivisions du segment $[a, b]$.

➠ **Exemple XXI.2.**
- $\left(0, \frac{1}{2}, 1\right)$ est une subdivision relativement aisée à visualiser du segment $[0, 1]$;
- pour $n \geq 1$, on appelle **subdivision régulière** du segment $[a, b]$ de rang n la subdivision
$$\sigma^n_{\mathscr{R}}(a, b) = \left(a + k\frac{b-a}{n}\right)_{k \in [\![0,n]\!]}\,.$$

Définition XXI.3. Soit $\sigma = (x_0, \dots, x_n) \in S(a, b)$. On appelle :
- **pas** de σ la quantité $\mu(\sigma) = \max\{x_{i+1} - x_i \,|\, i \in [\![0, n-1]\!]\}$;
- **support** de σ l'ensemble $\mathrm{supp}(\sigma) = \{x_i | i \in [\![0, n]\!]\}$.

➠ **Exemple XXI.3.** Le pas de $\sigma^n_{\mathscr{R}}(a, b)$ est $\dfrac{b-a}{n}$.

Définition XXI.4. Soient $\sigma, \sigma' \in S(a, b)$. On dit que σ est **plus fine** que σ' si $\mathrm{supp}(\sigma') \subset \mathrm{supp}(\sigma)$.

Notation. $\sigma' \prec \sigma$

✌ **Remarque XXI.3.** Cela signifie que σ contient *a minima* tous les points de σ'.

➠ **Exemple XXI.4.** $\left(0, \frac{1}{2}, 1\right) \prec \sigma^4_{\mathscr{R}}(0, 1)$.

✌ **Remarque XXI.4.** Il s'agit d'un ordre non total sur $S(a, b)$.

Proposition XXI.4. Soient $\sigma, \sigma' \in S(a, b)$. Alors il existe une subdivision σ'' telle que $\sigma \prec \sigma''$ et $\sigma' \prec \sigma''$.

Démonstration. Il suffit de choisir σ'' de façon à ce que $\mathrm{supp}(\sigma'') = \mathrm{supp}(\sigma) \cup \mathrm{supp}(\sigma')$. □

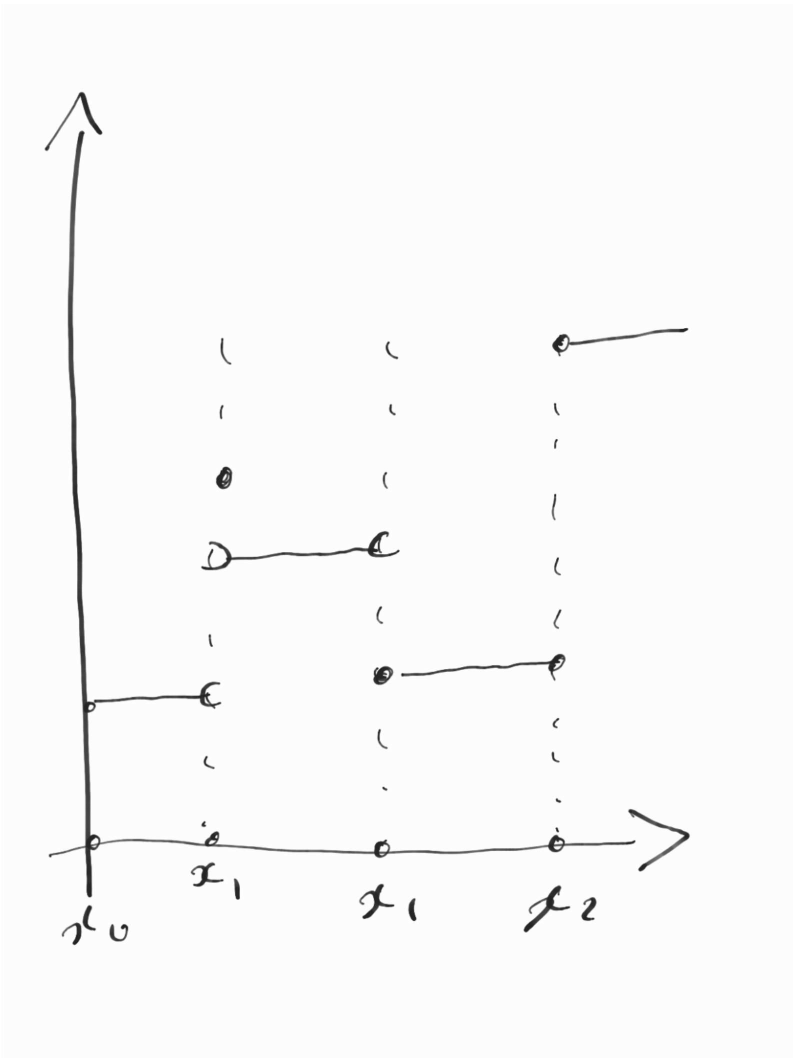
Ces définitions liminaires nous permettent (enfin) de rentrer dans le cœur de notre sujet à ce stade, en l'occurrence les fonctions en escalier.

Définition XXI.5. Une fonction $\phi : [a, b] \to \mathbb{R}$ est dite **en escaliers** si il existe une subdivision $\sigma = (x_0, \dots, x_n) \in S(a, b)$ telle que :

$$\forall i \in [\![0, n-1]\!],\ \text{la fonction } \phi_{|]x_i, x_{i+1}[} \text{ est constante.}$$

Notation. On notera $\mathscr{E}(a, b)$ l'ensemble des fonctions en escaliers sur $[a, b]$.

Vocabulaire. Une subdivision vérifiant les conditions de la définition XXI.5 est dite **adaptée** à la fonction ϕ. Notons que si σ est adaptée, alors toute subdivision σ' plus fine que σ l'est également.



Exemple XXI.5. La fonction partie entière est en escalier sur $[-5, 12]$ (par exemple), relativement à toute subdivision dont le support contient $\mathbb{Z}$.

✌ **Remarque XXI.5.** $(\mathscr{E}(a,b),+,.)$ (resp. $(\mathscr{E}(a,b),+,\times)$) est un s–e.v (resp. un sous–anneau) de $\mathbb{R}^{[a,b]}$. Lorsque ces deux structures sont présentes, on parle de **sous–algèbre**.

b) Intégrale d'une fonction en escalier

Proposition/définition XXI.6. Soit $\phi \in \mathscr{E}(a,b)$ et soit $\sigma \in S(a,b)$ une subdivision adaptée à ϕ. Pour $i \in [\![0,n-1]\!]$, on note y_i la valeur de la fonction (constante) $\phi_{|]x_i,x_{i+1}[}$. Alors la quantité

$$\sum_{i=0}^{n-1} y_i(x_{i+1}-x_i)$$

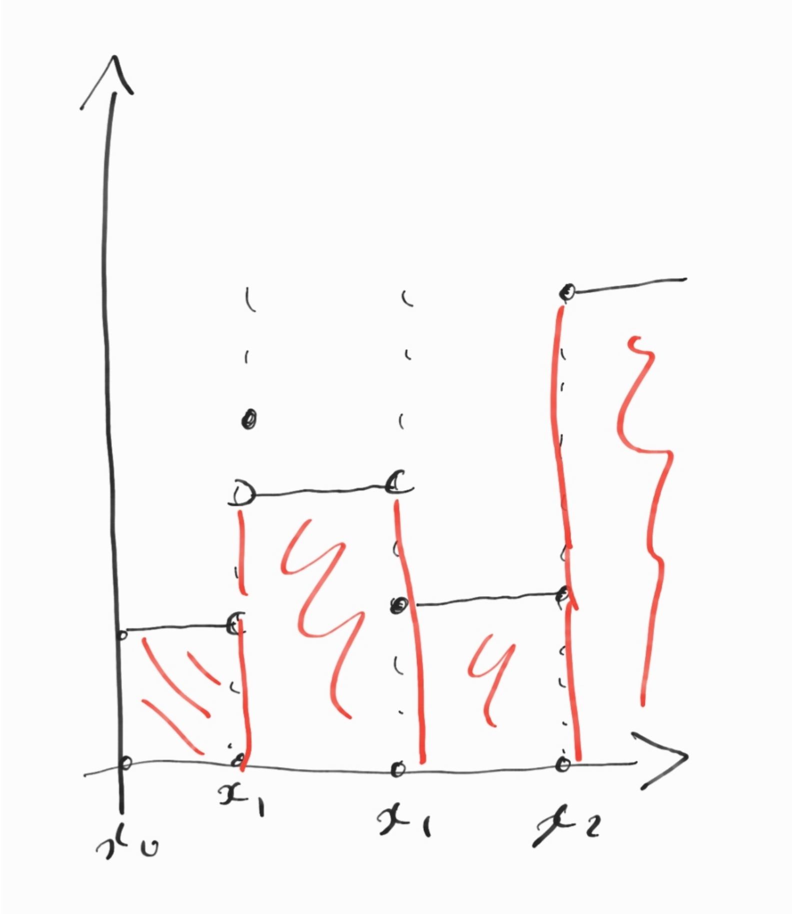
ne dépend pas de σ et est appelée **intégrale** de ϕ sur le segment $[a,b]$.

Démonstration. Il est clair que cette quantité est inchangée si l'on passe à une subdivision plus fine que σ. Si σ' est une autre subdivision adaptée à ϕ qui n'est pas plus fine que σ, on utilise la proposition XXI.4 pour régler la question. □

Notation. Les **seules** notations tolérées par le programme de MPSI sont :

$$\int_{[a,b]} \phi, \quad \int_a^b \phi \quad \text{et} \quad \int_a^b \phi(t)\,\mathrm{d}t\,.$$

Géométriquement, l'intégrale d'une fonction en escalier correspond à l'aire (algébrique) délimitée par la courbe de f (ignorant les discontinuités) et l'axe des abscisses.



➠ **Exemple XXI.6.**

$$\int_{-1}^{7} \lfloor x \rfloor \,\mathrm{d}x = \sum_{i=-1}^{6} i(i+1-i) = -1+0+1+2+3+4+5+6 = 20\,.$$

Proposition XXI.5. Soient $\phi, \psi \in \mathscr{E}(a,b)$ et soit $\lambda \in \mathbb{R}$. Alors :

(i)
$$\int_{[a,b]} \phi + \lambda\psi = \int_{[a,b]} \phi + \lambda \int_{[a,b]} \psi\,; \qquad \textbf{[linéarité]}$$

(ii)
$$(\phi \leq \psi) \Rightarrow \left(\int_{[a,b]} \phi \leq \int_{[a,b]} \psi \right)\;; \qquad \textbf{[croissance]}$$

(iii)
$$\left| \int_{[a,b]} \phi \right| \leq \int_{[a,b]} |\phi|\,. \qquad \textbf{[inégalité triangulaire]}$$

Démonstration.

(i) Trivial, quitte à utiliser la proposition XXI.4 pour matérialiser une subdivision adaptée à ϕ **et** ψ.

(ii) Par définition de l'intégrale, on a que si $\phi - \psi \leq 0$ alors $\int_a^b \phi - \psi \leq 0$. On conclut par linéarité.

(iii) On sait que $\phi \leq |\phi|$ et $-\phi \leq |\phi|$. De fait, par croissance et linéarité :

$$\int_a^b \phi \leq \int_a^b |\phi| \quad \text{et} \quad -\int_a^b \phi = \int_a^b (-\phi) \leq \int_a^b |\phi|$$

d'où le résultat.

□

La proposition suivante est souvent appelée "formule de Chasles", de part sa similarité avec la relation vectorielle éponyme. Elle n'a cependant rien à voir avec Michel Chasles, mathématicien français (1793—1880) connu pour ses travaux en géométrie projective et analyse harmonique.

Proposition XXI.6 ("Chasles")**.** Soit $c \in [a, b]$ et soit $\phi \in \mathscr{E}(a, b)$. Alors :

$$\int_a^b \phi = \int_a^c \phi + \int_c^b \phi\,.$$

Démonstration. Il s'agit d'un calcul rébarbatif et sans intérêt. Le lecteur masochiste pourra s'y essayer à l'envi. □

✌ **Remarque XXI.6.** Il est possible d'assouplir quelque peu notre notation en posant (rappelons que $a < b$) :

$$\int_b^a \phi = -\int_a^b \phi\,.$$

Les propriétés vues précédemment, formule de "Chasles" comprise, se généralisent sans guère de soucis à la condition de **prendre garde au sens des inégalités** : l'intégrale "à l'envers" est décroissante et non croissante. Notons au passage que l'on obtient :

$$\int_a^a \phi = \int_a^b \phi + \int_b^a \phi = 0\,.$$

2.⁻ Intégrale de Riemann

a) Fonctions continues par morceaux

Définition XXI.7. Une fonction $f : [a, b] \to \mathbb{R}$ est dite **continue par morceaux** si il existe $\sigma = (x_0, \ldots, x_{n-1}) \in S(a, b)$ telle que :
- $\forall i \in [\![0, , n-1]\!]$, $f \in \mathscr{C}^0(]x_i, x_{i+1}[)$;
- f admet des limites **réelles** (non nécessairement égales) à gauche et à droite en $x_1, \ldots, x_{n-1}$;
- f admet une limite **réelle** à gauche en b et à droite en a.

Vocabulaire. Toute subdivision satisfaisant ces propriétés est dite **adaptée** à f.

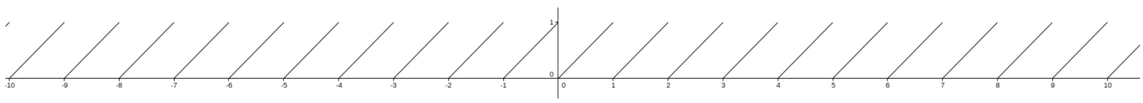
Notation. On note $\mathscr{C}_{\mathrm{pm}}([a,b])$ l'ensembles des fonctions continue par morceaux sur $[a,b]$.

Exemple XXI.7.

- les fonctions continues sont continue par morceaux ;
- les fonctions en escalier sont continue par morceaux ;
- la fonction

$$\begin{aligned} f : [-100, 100] &\to \mathbb{R} \\ x &\mapsto x - \lfloor x \rfloor \end{aligned}$$

  est continue par morceaux .



- La fonction

$$x \mapsto \begin{cases} \sin\left(\dfrac{1}{x}\right) & \text{si } x \neq 0 \\ 0 \text{ sinon.} \end{cases}$$

  n'est **pas** continue par morceaux sur $[0,1]$ car elle n'admet pas de limite à droite en 0.

Remarque XXI.7. $\mathscr{C}_{\mathrm{pm}}([a,b])$ est une sous–algèbre de $\mathbb{R}^{[a,b]}$ contenant $\mathscr{E}(a,b)$ et $\mathscr{C}^0([a,b])$.

Définition XXI.8. Une fonction est dite continue par morceaux sur I si elle l'est sur tout segment inclus dans I.

Exemple XXI.8. La fonction f *supra* est en fait continue par morceaux sur $\mathbb{R}$.

Théorème XXI.7 (Densité des fonctions en escaliers)**.**
Soit $f \in \mathscr{C}_{\mathrm{pm}}([a,b])$. Alors, pour tout $\varepsilon > 0$ il existe deux fonctions en escalier $\phi, \psi \in \mathscr{E}(a,b)$ telles que :

(i) $\phi \leq f \leq \psi$;

(ii) $0 \leq \psi - \phi \leq \varepsilon$.

Démonstration. Soit $\varepsilon > 0$. Commençons par remarquer qu'il nous suffit de définir ϕ et ψ sur chacun des intervalles de continuité de f : nous pouvons donc légitimement supposer f continue sur le segment $[a,b]$ et donc, par théorème de Heine (XXI.3) uniformément continue sur ce dernier. Nous obtenons de fait l'existence de $\delta > 0$ tel que :

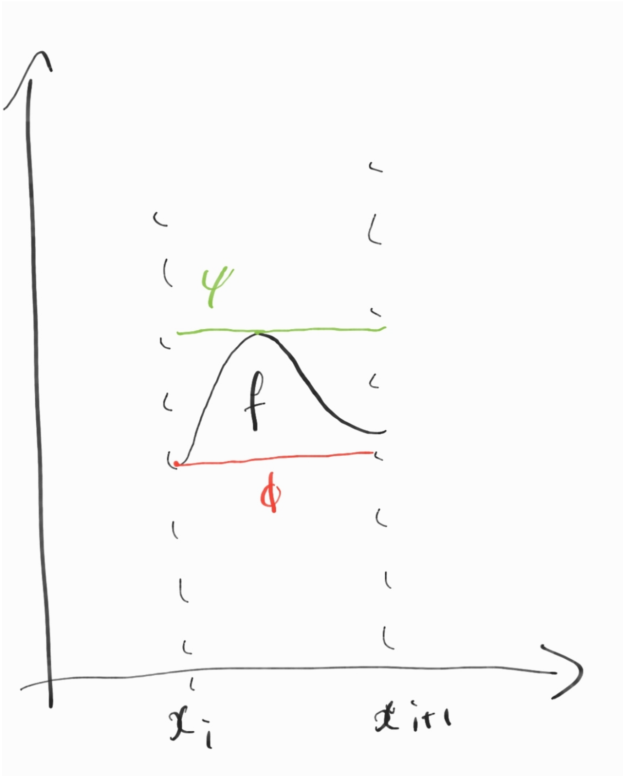
$$\forall x, y \in [a,b],\ (|x-y| \leq \delta) \Rightarrow (|f(x) - f(y)| \leq \varepsilon)\,.$$

Donnons nous à présent une subdivision $\sigma = (x_0, \ldots, x_n) \in S(a,b)$ telle que $\mu(\sigma) \leq \delta$ et posons :

$$\phi : [a,b] \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} \min f_{[x_i, x_{i+1}[} \text{ si } x \in [x_i, x_{i+1}[\text{ avec } i \in [\![0, n-1]\!] \\ f(b) \text{ si } x = b \end{cases}$$

ainsi que :

$$\psi : [a,b] \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} \max f_{[x_i, x_{i+1}[} \text{ si } x \in [x_i, x_{i+1}[\text{ avec } i \in [\![0, n-1]\!] \\ f(b) \text{ si } x = b \end{cases}$$



Notons que ϕ (resp. ψ) constitue une approximation de f par défaut (resp. excès) sur les intervalles $[x_i, x_{i+1}[$. On a donc immédiatement que $\phi \leq f \leq \psi$ sur $[a,b]$ et, pour tout $i \in [\![0, n-1]\!]$ et $x \in [x_i, x_{i+1}[$:

$$\psi(x) - \phi(x) = \max f_{[x_i, x_{i+1}[} - \min f_{[x_i, x_{i+1}[} \leq \varepsilon$$

par continuité uniforme ($|x_{i+1} - x_i| \leq \delta$ par construction). L'inégalité étant triviale en $x = b$, on a bien $0 \leq \psi - \phi \leq \varepsilon$ sur $[a,b]$, d'où le résultat.

□

b) Intégrale d'une fonction continue par morceaux

Notation. Soit $f \in \mathscr{C}_{\mathrm{pm}}([a,b])$. Comme f est bornée sur chacun des ses intervalles de continuité, elle l'est sur $[a,b]$; nous pouvons donc poser :

$$E_-(f) = \left\{ \int_{[a,b]} \phi \;\middle|\; \phi \in \mathscr{E}(a,b),\ \phi \leq f \right\}$$

et

$$E_+(f) = \left\{ \int_{[a,b]} \phi \,\middle|\, \phi \in \mathscr{E}(a,b),\ \phi \geq f \right\}.$$

Proposition XXI.8. Soit $f \in \mathscr{C}_{\mathrm{pm}}([a,b])$; alors :

(i) $E_-(f)$ admet une borne supérieure ;

(ii) $E_+(f)$ admet une borne inférieure ;

(iii) $\sup(E_-(f)) = \inf(E_+(f))$.

Démonstration. Les deux ensembles $E_+(f)$ et $E_-(f)$ sont des parties de $\mathbb{R}$ non vides : en effet, f est bornée donc il existe $M \in \mathbb{R}_+$ tel que $\forall x \in [a,b]$, $|f(x)| \leq M$ et donc $E_-(f)$ resp. ($E_+(f)$) contient la fonction $x \mapsto -M$ (resp. $x \mapsto M$).

Soient à présent $x \in E_-(f)$ et $y \in E_+(f)$. Alors, il existe $\phi, \psi \in \mathscr{E}(a,b)$ telles que $\phi \leq f \leq \psi$, $x = \displaystyle\int_{[a,b]} \phi$ et $y = \displaystyle\int_{[a,b]} \psi$. De fait, par croissance de l'intégrale $x \leq y$ et donc les bornes voulues existent (tout élément de $E_-(f)$ minore $E_+(f)$ et tout élément de $E_+(f)$ majore $E_-(f)$) et $\sup(E_-(f)) \leq \inf(E_+(f))$.

Pour obtenir l'égalité, fixons $n \geq 1$ et notons que nous avons l'existence, par théorème XXI.7, de deux fonctions en escalier $\phi, \psi \in \mathscr{E}(a,b)$ telles que $\phi \leq f \leq \psi$ et $\psi - \phi \leq \dfrac{1}{n(b-a)}$. On a, par croissance de l'intégrale sur les fonctions en escalier :

$$0 \leq \underbrace{\int_a^b \psi}_{\geq \inf(E_+(f))} - \underbrace{\int_a^b \phi}_{\leq \sup(E_-(f))} \phi \leq \int_a^b \frac{1}{n(b-a)} = \frac{1}{n}$$

ce qui entraîne que :

$$\inf(E_+(f)) - \sup(E_-(f)) \leq \frac{1}{n}$$

et donc, en faisant tendre n vers l'infini, on obtient l'inégalité désirée. □

Définition XXI.9. Soit $f \in \mathscr{C}_{\mathrm{pm}}([a,b])$; la quantité $\sup(E_-(f)) = \inf(E_+(f))$ est appelée **intégrale** de f sur le segment $[a,b]$.

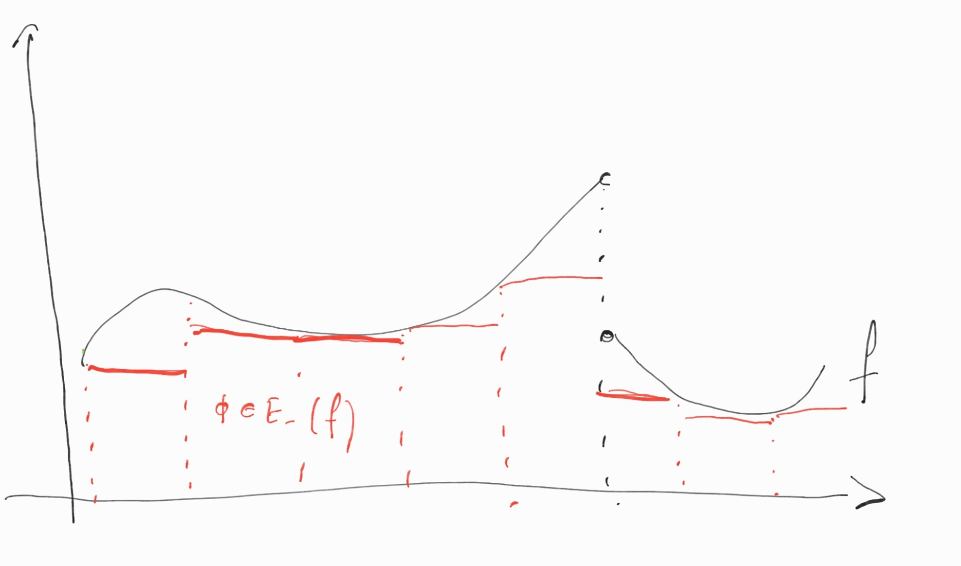
Notation. Une nouvelle fois, les **seules** notations tolérées par le programme de MPSI sont :

$$\int_{[a,b]} f, \quad \int_a^b f \quad \text{et} \quad \int_a^b f(t)\,\mathrm{d}t\,.$$

✌ **Remarque XXI.8.**

— Notons que si f est en escalier, les bornes de $E_+(f)$ et $E_-(f)$ sont atteintes : nos deux définitions d'intégrales coïncident bien. Notre psychologue peut s'autoriser un soupir de soulagement.

— Une fonction "pas trop mal choisie" dans $E_-(f)$ est une fonction en escalier "aussi proche que possible" de f tout en restant au–dessous du point de vue des courbes : approcher l'intégrale de f par celle de cette fonction revient donc à appliquer à f une méthode des rectangles par défaut. Même chose avec $E_+(f)$ et la méthode par excès.



— On appelle **valeur moyenne** de f la quantité

$$\langle f \rangle = \frac{1}{b-a} \int_a^b f(t)\,\mathrm{d}t\,.$$

✍ **Exercice XXI.3.** Déterminer $\displaystyle\int_0^1 x^2\,\mathrm{d}x$.

➥ **Correction :** *Soit $n \geq 1$; on pose $\sigma = \left(\frac{k}{n}\right)_{k\in[\![0,n]\!]}$ et*

$$\begin{aligned} \phi : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \left(\frac{k}{n}\right)^2 & \text{si } x \in \left[\frac{k}{n}, \frac{k+1}{n}\right[\quad \text{avec } k \in [\![0, n-1]\!] \\ 1 \ \text{si } x = b \end{cases} \end{aligned}$$

ainsi que :

$$\begin{aligned} \psi : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \left(\frac{k+1}{n}\right)^2 & \text{si } x \in \left[\frac{k}{n}, \frac{k+1}{n}\right[\quad \text{avec } k \in [\![0, n-1]\!] \\ 1 \ \text{si } x = b \end{cases} \end{aligned} \quad .$$

On a alors clairement $\forall x \in [0,1]$, $\phi(x) \leq x^2 \leq \psi(x)$ et :

$$\int_0^1 \phi(x)\,\mathrm{d}x = \frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{k}{n}\right)^2 = \frac{(n-1)(2n-1)}{6n^2}$$

et

$$\int_0^1 \psi(x)\,\mathrm{d}x = \frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{k+1}{n}\right)^2 = \frac{(n+1)(2n+1)}{6n^2}\,.$$

Ceci entraîne que

$$\frac{(n-1)(2n-1)}{6n^2} \leq \int_0^1 x^2 \,\mathrm{d}x \leq \frac{(n+1)(2n+1)}{6n^2}$$

et donc, par théorème d'encadrement des limites (IX.12) :

$$\int_0^1 x^2 \,\mathrm{d}x = \frac{1}{3}\,.$$

c) Propriétés de l'intégrale

Dans ce paragraphe, nous étendons à l'intégrale des fonctions continue par morceaux les propriétés observées dans le cas des fonctions en escalier. Sans surprise, tout ceci requiert un usage massif du théorème XXI.7.

◇ Linéarité

Proposition XXI.9. Soient $f, g \in \mathscr{C}_{\mathrm{pm}}([a,b]$ et soit $\lambda \in \mathbb{R}$. Alors :

$$\int_a^b (f(t) + \lambda g(t))\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t + \lambda \int_a^b g(t)\,\mathrm{d}t\,.$$

Démonstration. Soit $n \geq 1$. Alors, par théorème XXI.7, il existe $\phi_1, \phi_2, \psi_1, \psi_2 \in \mathscr{E}(a,b)$ telles que $\phi_1 \leq f \leq \psi_1$, $\phi_2 \leq g \leq \psi_2$, $\psi_1 - \phi_1 \leq \dfrac{1}{n}$ et $\psi_2 - \phi_2 \leq \dfrac{1}{n}$. De fait, par croissance et linéarité de l'intégrale sur les fonctions en escalier, on a :

$$\begin{aligned}\int_{[a,b]} (f+\lambda g) - \left(\int_{[a,b]} f + \lambda \int_{[a,b]} g\right) &\leq \int_{[a,b]} (\psi_1 + \lambda\psi_2) - \left(\int_{[a,b]} \phi_1 + \lambda \int_{[a,b]} \phi_2\right) \\ &= \int_{[a,b]} (\psi_1 - \phi_1) + \lambda \int_{[a,b]} (\psi_2 - \phi_2) \\ &\leq (b-a)\frac{1+\lambda}{n}\end{aligned}$$

et, de la même façon :

$$\begin{aligned}\int_{[a,b]} (f+\lambda g) - \left(\int_{[a,b]} f + \lambda \int_{[a,b]} g\right) &\geq \int_{[a,b]} (\phi_1 + \lambda\phi_2) - \left(\int_{[a,b]} \psi_1 + \lambda \int_{[a,b]} \psi_2\right) \\ &\geq (a-b)\frac{1+\lambda}{n}\,.\end{aligned}$$

On obtient le résultat voulu en faisant tendre n vers l'infini via le théorème d'encadrement des limites (IX.12).

□

◇ **Croissance, inégalité triangulaire**

Proposition XXI.10. Soit $f \in \mathscr{C}_{\mathrm{pm}}([a,b])$ une fonction **à valeurs positives**. Alors :

$$\int_a^b f(t)\,\mathrm{d}t \geq 0\,.$$

Démonstration. La fonction nulle est une fonction en escalier minorant f : elle appartient donc à $E_-(f)$. De fait :

$$\int_a^b f(t)\,\mathrm{d}t = \sup(E_+(f)) \geq \int_a^b 0\,\mathrm{d}t = 0\,.$$

□

✖ **Attention :** la réciproque est **fausse** : l'intégrale de $x \mapsto x$ entre -1 et 1 est positive sans que cette fonction ne le soit.

Corollaire XXI.10.a. Soient $f, g \in \mathscr{C}_{\mathrm{pm}}([a,b])$. Alors :

$$(f \leq g) \Rightarrow \left(\int_{[a,b]} f \leq \int_{[a,b]} g\right)\,.$$

Démonstration. Appliquer la proposition *supra* à $g - f$. □

✌ **Remarque XXI.9.** Cette proposition est apparente graphiquement : l'aire d'une fonction dont la courbe est située en permanence sous celle d'une autre risque de ne pas pouvoir excéder celle de son estimée collègue.

Proposition XXI.11. Soit $f \in \mathscr{C}_{\mathrm{pm}}([a,b])$. Alors :

(i) $x \mapsto |f(x)| \in \mathscr{C}_{\mathrm{pm}}([a,b])$;

(ii)

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x\,.$$

Démonstration. Identique à celle de l'inégalité triangulaire pour l'intégrale des fonctions en escalier. Pour le point(i), il suffit de remarquer que tout subdivision adaptée à f l'est aussi à $x \mapsto |f(x)|$. □

◇ **Relation de Chasles**

Proposition XXI.12 (Chasles)**.** Soit $c \in [a, b]$ et soit $f : [a, b] \to \mathbb{R}$. Alors :

(i) $(f \in \mathscr{C}_{\text{pm}}([a, b])) \Leftrightarrow (f \in \mathscr{C}_{\text{pm}}([a, c] \cap \mathscr{C}_{\text{pm}}([c, b]))$;

(ii) dans ce cas,
$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t\,.$$

Démonstration. Il s'agit d'une abomination calculatoire que nous ferions mieux de laisser en paix. □

✌ **Remarque XXI.10.** Il est possible, de la même façon et avec les **mêmes limites** que pour les fonctions en escalier, de généraliser l'intégrale ou cas où $a > b$.

◇ **Nullité**

Proposition XXI.13 (Nullité de l'intégrale)**.** Soit $f : [a, b] \to \mathbb{R}$ une fonction **continue**, **positive** et telle que $\displaystyle\int_a^b f(t)\,\mathrm{d}t = 0$. Alors f est la fonction nulle.

✖ **Attention :** : ce théorème est bien entendu **faux** si on l'ampute d'une (ou plusieurs hypothèses) : l'indicatrice de $\mathbb{Z}$ sur $[0, 1]$ donne un contre–exemple discontinu, tandis que l'identité est d'intégrale nulle sur $[-1, 1]$ mais non nulle (elle n'est pas positive sur ce segment).

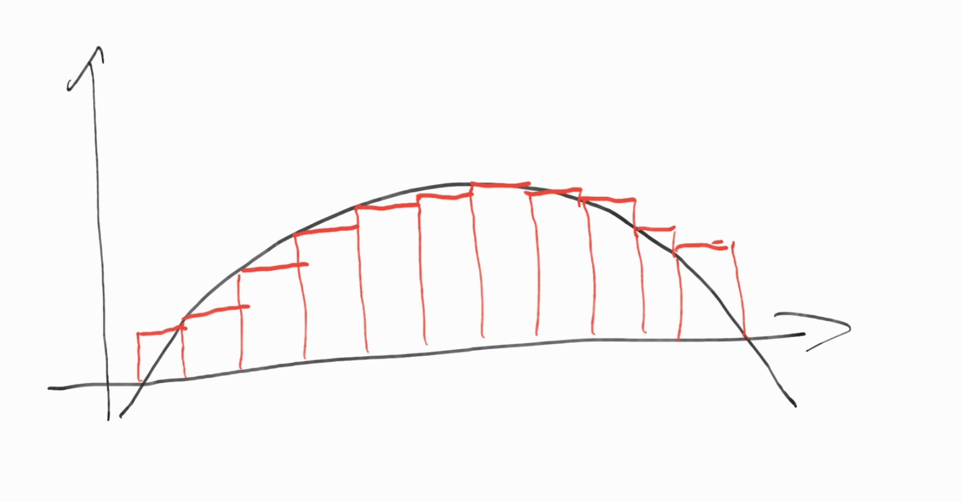
Démonstration. Supposons f non nulle : alors il existe $c \in [a, b]$ tel que $f(c) > 0$. Comme f est continue en c, il existe un voisinage $V \in \mathscr{V}(c)$ tel que $f > 0$ sur V ; de fait
$$\int_a^b f(t)\,\mathrm{d}t \geq \int_{[a,b]\cap V} f > 0$$
ce qui est absurde. □

✍ **Exercice XXI.4.** Soit $f \in \mathscr{C}^0([-a, a])$ une fonction impaire. Démontrer que $\int_{-a}^a f(t)\,\mathrm{d}t = 0$.

d) Sommes de Riemann

Proposition XXI.14. Soit $f \in \mathscr{C}_{\text{pm}}([a, b])$. Alors :
$$\frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) \xrightarrow[n\to\infty]{} \int_a^b f(t)\,\mathrm{d}t\,.$$

✌ **Remarque XXI.11.** Ce résultat est en fait une version séquentielle de la méthode des rectangles, appliquée aux subdivisions régulières $\sigma^n_{\mathscr{R}}(a,b)$.



Plus n sera "grand", plus nos rectangles seront "fins" et donc mieux l'intégrale sera approchée par la somme apparaissant dans la proposition, appelée **somme de Riemann**, du nom du mathématicien allemand Bernhard Riemann (1826—1866), que l'on ne présente plus.

Démonstration. **Nous nous limitons ici au cas où $f \in \mathscr{C}^1([a,b])$.** Soit $n \geq 1$ et posons, pour $k \in [\![0, n-1]\!]$ $a_k = a + k\dfrac{b-a}{n}$; alors :

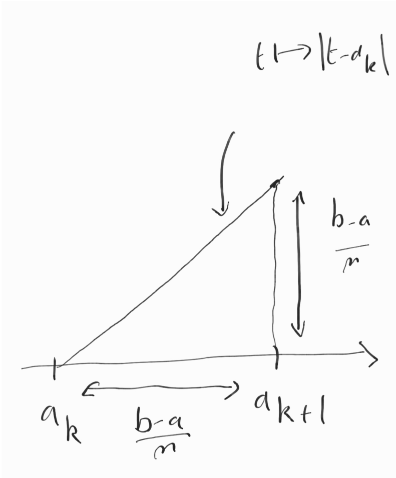
$$\begin{aligned}
\left|\int_a^b f(t)\,\mathrm{d}t - \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right)\right| &= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(t)\,\mathrm{d}t - \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a_k\right)\right| \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(t)\,\mathrm{d}t - \frac{b-a}{n} f\left(a_k\right)\right| \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(t)\,\mathrm{d}t - (a_{k+1}-a_k) f\left(a_k\right)\right| \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(t)\,\mathrm{d}t - \int_{a_k}^{a_{k+1}} f\left(a_k\right)\,\mathrm{d}t\right| \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(t) - f\left(a_k\right)\,\mathrm{d}t\right| \\
&\leq \sum_{k=0}^{n-1}\left|\int_{a_k}^{a_{k+1}} f(t) - f\left(a_k\right)\,\mathrm{d}t\right| \\
&\leq \sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} |f(t) - f(a_k)|\,\mathrm{d}t\,.
\end{aligned}$$

Or, comme f est de classe $\mathscr{C}^1$ sur le segment $[a,b]$, sa dérivée y est bornée par $M \in \mathbb{R}_+$, et, par théorème des accroissements finis (XII.11) on a :

$$\forall k \in [\![0, n-1]\!],\ \forall t \in [a_k, a_{k+1}]\ |f(t) - f(a_k)| \leq M|t - a_k|$$

ce qui entraîne que :

$$\sum_{k=0}^{n-1} \int_{a_k}^{a_{k+1}} |f(t) - f(a_k)| \,\mathrm{d}t \leq \sum_{k=0}^{n-1} \int_{a_k}^{a_{k+1}} M|t - a_k| \,\mathrm{d}t$$
$$= M \sum_{k=0}^{n-1} \int_{a_k}^{a_{k+1}} |t - a_k| \,\mathrm{d}t\ .$$



Or, l'intégrale de $t \mapsto |t - a_k|$ entre a_k et a_{k+1} correspond à l'aire du triangle rectangle de sommets $(a_k, 0)$, $(a_{k+1}, 0)$ et $(a_{k+1}, a_{k+1} - a_k)$. Le théorème de Pythagore nous livre alors que :

$$\int_{a_k}^{a_{k+1}} |t - a_k| \,\mathrm{d}t = \frac{1}{2} \left(\frac{b-a}{n} \right)^2$$

ce qui implique que :

$$M \sum_{k=0}^{n-1} \int_{a_k}^{a_{k+1}} |t - a_k| \,\mathrm{d}t = M \sum_{k=0}^{n-1} \frac{1}{2} \left(\frac{b-a}{n} \right)^2$$
$$= \frac{M}{2n}(b-a)^2$$
$$\xrightarrow[n\to\infty]{} 0$$

d'où le résultat.

□

➠ Exemple XXI.9.

—
$$\frac{1}{n} \sum_{k=0}^{n-1} \cos\left(\frac{k}{n}\right) \xrightarrow[n\to\infty]{} \int_0^1 \cos(t) \,\mathrm{d}t\ ;$$

— si $n \geq 1$ alors :

$$n \sum_{k=0}^{n-1} \frac{1}{n^2 + k^2} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{1 + \left(\frac{k}{n}\right)^2}$$
$$\xrightarrow[n\to\infty]{} \int_0^1 \frac{1}{1+t^2} \,\mathrm{d}t$$
$$= \frac{\pi}{4}\ .$$

3.⁻ Primitives

On fixe dans ce paragraphe un intervalle I de $\mathbb{R}$ d'intérieur non vide.

On rappelle que si f est une fonction définie sur I, on appelle **primitive** de f toute fonction F **dérivable sur I** telle que $F' = f$. Nous reprenons (et démontrons) dans ce paragraphe plusieurs résultats évoqués au chapitre XIV.

a) Existence

Proposition XXI.15. Soit $f \in \mathscr{C}^0(I)$ et soit $a \in I$. Alors la fonction

$$\begin{aligned} F : I &\to \mathbb{R} \\ x &\mapsto \int_a^x f(t)\,\mathrm{d}t \end{aligned}$$

est l'unique primitive de f telle que $F(a) = 0$.

Démonstration. La fonction F est bien définie car f est continue donc continue par morceaux. De plus si $x, y \in I$ sont tels que $x \neq y$ alors :

$$\begin{aligned} \frac{F(x)-F(y)}{x-y} &= \frac{1}{x-y}\left(\int_a^x f(t)\,\mathrm{d}t - \int_a^y f(t)\,\mathrm{d}t\right) \\ &= \frac{1}{x-y}\left(\int_a^x f(t)\,\mathrm{d}t + \int_y^a f(t)\,\mathrm{d}t\right) \\ &= \frac{1}{x-y}\int_y^x f(t)\,\mathrm{d}t\,. \end{aligned}$$

Nous savons en outre que f est continue en y ; ceci entraîne que :

$$\forall \varepsilon > 0,\ \exists\, \delta > 0, \forall t \in I\ (|t-y| \leq \delta) \Rightarrow (|f(t) - f(y)| \leq \varepsilon)\,.$$

De fait, si nous supposons $|x-y| \leq \delta$, on a :

$$\begin{aligned} \left|\frac{F(x)-F(y)}{x-y} - f(y)\right| &= \left|\frac{1}{x-y}\int_y^x f(t)\,\mathrm{d}t - f(y)\right| \\ &= \left|\frac{1}{x-y}\int_y^x f(t)\,\mathrm{d}t - \frac{1}{x-y}\int_y^x f(y)\,\mathrm{d}t\right| \\ &= \left|\frac{1}{x-y}\int_y^x f(t) - f(y)\,\mathrm{d}t\right| \\ &\leq \frac{1}{|x-y|}\int_{\underbrace{[x,y]}} |f(t)-f(y)|\,\mathrm{d}t \end{aligned}$$

où $\underbrace{[x,y]}$ désigne $[y,x]$ si $y < x$ et $[x,y]$ sinon. De fait, pour tout $t \in \underbrace{[x,y]}$, on a $|t-y| \leq \delta$ *ergo* :

$$\left|\frac{F(x)-F(y)}{x-y} - f(y)\right| \leq \frac{1}{|x-y|}\varepsilon|x-y| = \varepsilon$$

et donc

$$\frac{F(x)-F(y)}{x-y} \xrightarrow[x\to y]{} f(y)$$

d'où la dérivabilité de F sur I et le fait que $F' = f$.

Pour conclure, remarquons que si G est une primitive de f telle que $G(a) = 0$ alors F et G sont solutions du problème de Cauchy linéaire du premier ordre suivant :

$$\begin{cases} y' = f \\ y(a) = 0 \end{cases}$$

et sont donc égales par le théorème de Cauchy–Lipschitz (XIV.6). □

Théorème XXI.16 (Théorème fondamental du calcul intégral)**.**
Soit $f \in \mathscr{C}^0([a,b])$ et soit F une primitive de f sur $[a,b]$. Alors :

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a)\ .$$

Démonstration. La fonction $G : x \mapsto F(x) - F(a)$ est une primitive de f telle que $G(a) = 0$ donc, d'après la proposition XXI.15 on, pour tout $x \in [a,b]$:

$$F(x) - F(a) = \int_a^x f(t)\,\mathrm{d}t$$

d'où le résultat. □

✌ **Remarque XXI.12.** Ce théorème a plusieurs conséquences fondamentales :

— les primitives de f sont exactement les fonctions de la forme

$$x \mapsto \int_a^x f(t)\,\mathrm{d}t + C$$

pour C parcourant $\mathbb{R}$;

— l'intégrale est une fonction continue de ses bornes, *i.e* pour tout $\alpha \in \mathbb{R}$ on a :

$$\int_a^x f(t)\,\mathrm{d}t \xrightarrow[x\to\alpha]{} \int_a^\alpha f(t)\,\mathrm{d}t\ ;$$

— si f est de classe $\mathscr{C}^1$ sur $[a,b]$, on a :

$$f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t\ .$$

b) Retour sur quelques méthodes de calcul intégral

Nous revenons brièvement dans ce paragraphe sur des résultats déjà démontrés au chapitre XIV, afin d'offrir à notre estimé lecteur le plaisir de s'immoler les rétines une fois de plus. Les remerciements sont superflus, je vous assure.

Proposition XXI.17 (Intégration par parties)**.** Soient f et g deux fonctions **de classe** $\mathscr{C}^1$ sur I. Alors, pour tous $a, b \in I$:

$$\int_a^b f'(t)g(t)\,\mathrm{d}t = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)\,\mathrm{d}t\,,$$

avec

$$[f(t)g(t)]_a^b = f(b)g(b) - f(a)g(a)\,.$$

➠ **Exemple XXI.10.** Soit $x \in [-1,1]$; pour déterminer $\int_0^x \arcsin(t)\,\mathrm{d}t$ on applique l'IPP à $f : t \mapsto t$ et $g = \arcsin$, toutes deux de classe $\mathscr{C}^1$ sur $[0,x]$. On obtient :

$$\begin{aligned}\int_0^x \arcsin(t)\,\mathrm{d}t &= [t\arcsin(t)]_0^x - \int_0^x \frac{t}{\sqrt{1-t^2}}\,\mathrm{d}t \\ &= x\arcsin(x) - \sqrt{1-x^2} + 1\end{aligned}$$

car $t \mapsto \sqrt{1-t^2}$ est une primitive de $t \mapsto \dfrac{t}{\sqrt{1-t^2}}$.

Théorème XXI.18 (Changement de variable)**.**
Soit $\varphi \in \mathscr{C}^1(I)$ et soit $f \in \mathscr{C}^0(\varphi(I))$. Alors, pour tous $a, b \in I$ on a :

$$\int_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x = \int_a^b f \circ \varphi(t)\varphi'(t)\,\mathrm{d}t\,.$$

4.⁻ Formules de Taylor

On fixe dans ce paragraphe un intervalle I d'intérieur non vide.

Nous avons rencontré par le passé deux formules de Taylor : celle concernant les polynômes (proposition XV.17) et la formule de Taylor–Young (théorème XVI.13). Ces deux formules avaient pour point commun l'expression de la valeur d'une fonction en un point à partir du polynôme de Taylor de f en un point a de son ensemble de définition, *i.e*

$$T_n = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} X^k$$

et d'un éventuel reste négligeable (voire nul).

Nous donnons dans ce paragraphe une nouvelle formule de Taylor avec une expression explicite de ce reste, et étudions certaines conséquences d'un tel résultat.

a) Formule de Taylor avec reste intégral

Proposition XXI.19. Soit $n \in \mathbb{N}$, $a, b \in I$ et $f \in \mathscr{C}^{n+1}(I)$. Alors :

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

Démonstration. On le fait par récurrence (joie!) sur $n \geq 0$.

— Le cas $n = 0$ est trivial étant donné que l'on a :

$$f(b) = f(a) + \int_a^b f'(t)\, \mathrm{d}t$$

d'après le théorème XXI.16.

— Si on suppose la propriété vraie au rang n et que $f \in \mathscr{C}^{n+2}([a, b])$ alors on a, comme f est automatiquement de classe $\mathscr{C}^{n+1}$ sur ce même intervalle :

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

par hypothèse de récurrence. Les fonctions $f^{(n+1)}$ et $u : t \mapsto -\dfrac{(b-t)^{n+1}}{(n+1)!}$ sont de classes $\mathscr{C}^1$ sur I avec

$$\forall t \in I,\ u'(t) = \frac{(n+1)(b-t)^n}{(n+1)!} = \frac{(b-t)^n}{n!}$$

et donc, par intégration par parties (proposition XXI.17) :

$$\begin{aligned}\int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\, \mathrm{d}t &= \left[-\frac{(b-t)^{n+1}}{(n+1)!} f^{(n+1)}(t)\right]_a^b - \int_a^b -\frac{(b-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\, \mathrm{d}t \\ &= \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(a) + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\, \mathrm{d}t\,.\end{aligned}$$

En réinjectant, on obtient

$$\begin{aligned}f(b) &= \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(a) + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\, \mathrm{d}t \\ &= \sum_{k=0}^{n+1} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\, \mathrm{d}t\end{aligned}$$

d'où le résultat.

□

➠ **Exemple XXI.11.** Si $x \in \mathbb{R}$ et $n \geq 0$, on a :

$$\begin{aligned}e^x &= \sum_{k=0}^{n} \frac{x^k}{k!} \exp^{(k)}(0) + \int_0^x \frac{(x-t)^n}{n!} \exp^{(n+1)}(t)\, \mathrm{d}t \\ &= \sum_{k=0}^{n} \frac{x^k}{k!} + \int_0^x \frac{(x-t)^n}{n!} e^t\, \mathrm{d}t\end{aligned}$$

l'intégrale pouvant se calculer par IPP (*cf.* chapitre XIV).

b) Inégalité de Taylor–Lagrange

Proposition XXI.20. Soit $n \in \mathbb{N}$, $a, b \in I$ et $f \in \mathscr{C}^{n+1}(I)$. Alors :

$$f(b) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + R_n$$

avec

$$|R_n| \leq \frac{|b-a|^{n+1}}{(n+1)!} \sup_{\underbrace{[\,a,b\,]}} |f^{(n+1)}|$$

où $\underbrace{[\,a,b\,]}$ désigne le segment $[a,b]$ si $a \leq b$ et $[b,a]$ sinon.

Démonstration. D'après la formule de Taylor avec reste intégral (proposition XXI.19), on a :

$$\begin{aligned}
|R_n| &= \left| \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t \right| \\
&\leq \int_a^b \left| \frac{(b-t)^n}{n!} f^{(n+1)}(t) \right| \mathrm{d}t \\
&\leq \int_a^b \left| \frac{(b-t)^n}{n!} \right| \sup_{\underbrace{[\,a,b\,]}} |f^{(n+1)}|\,\mathrm{d}t \\
&= \sup_{\underbrace{[\,a,b\,]}} |f^{(n+1)}| \int_a^b \left| \frac{(b-t)^n}{n!} \right| \mathrm{d}t \\
&= \sup_{\underbrace{[\,a,b\,]}} |f^{(n+1)}| \frac{|b-a|^{n+1}}{(n+1)!}
\end{aligned}$$

d'où le résultat. □

✌ **Remarque XXI.13.** La borne supérieure est en réalité un maximum car $f^{(n+1)}$ est continue sur le segment concerné.

➠ **Exemple XXI.12.** En reprenant l'exemple précédent, nous avons, pour tout $n \in \mathbb{N}$ et $x \in \mathbb{R}$:

$$\left| e^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \leq \frac{|x|^{n+1}}{(n+1)!} \sup_{\underbrace{[\,0,x\,]}} \exp$$

et donc, si $|x| < 1$ on a ;

$$\sum_{k=0}^{n} \frac{x^k}{k!} \xrightarrow[n\to\infty]{} e^x .$$

5.– Brève extension aux fonctions à valeurs complexes

Si $a, b \in \mathbb{R}$ sont tels que $a < b$ et $f : [a, b] \to \mathbb{C}$ est une fonction, on dira que f est continue par morceaux si et seulement si sa partie réelle et sa partie imaginaire le sont. De cette façon, on peut poser :

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b \mathrm{Re}(f(t))\,\mathrm{d}t + i \int_a^b \mathrm{Im}(f(t))\,\mathrm{d}t\,.$$

Toutes les propriétés vues dans ce chapitre se généralisent aux intégrales de fonctions continue par morceaux à valeurs complexes **sauf la croissance**. L'inégalité triangulaire reste vérifiée, quitte à remplacer la valeur absolue par le module.

Chapitre XXII

Matrices

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.⁻ Généralités

a) C'est quoi ?

Définition XXII.1. Soient $n, m \in \mathbb{N}$; on appelle **matrice** à n lignes et m colonnes sur $\mathbb{K}$ toute application

$$M : [\![1, n]\!] \times [\![1, m]\!] \to \mathbb{K} .$$

Notation. Du point de vue pratique, nous représenterons une matrice sous la forme d'un tableau à n lignes et m colonnes, le coefficient en position (i, j) correspondant à la valeur de $M(i, j)$. Ceci donnera des objets ressemblant à :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \quad \text{ou} \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} ,$$

les deux notations étant usitées aux concours.
L'ensemble des matrices à n lignes et m colonnes sera quant à lui noté $\mathscr{M}_{n,m}(\mathbb{K})$. Si $M \in \mathscr{M}_{n,m}(\mathbb{K})$, on notera généralement $m_{i,j}$ son coefficient ligne i, colonne j.

Vocabulaire.

— Lorsque $n = m$, on parle de **matrices carrées** ; l'ensemble correspondant est simplement noté $\mathscr{M}_n(\mathbb{K})$.

— On appelle **matrice diagonale** toute matrice **carrée** $M = (m_{i,j})_{i,j \in [\![1,n]\!]} \in \mathscr{M}_n(\mathbb{K})$ telle que

$$\forall (i, j) \in [\![1, n]\!]^2,\ (i \neq j) \Rightarrow (m_{i,j} = 0) .$$

Une telle matrice est donc de la forme :

$$\mathrm{diag}(d_1, \ldots, d_n) = \begin{pmatrix} d_1 & & (0) & \\ & d_2 & & \\ & & \ddots & \\ & (0) & & d_n \end{pmatrix}$$

avec $d_1, \ldots, d_n \in \mathbb{K}$. L'ensemble de ces matrices sera noté $D_n(\mathbb{K})$.

— Une **matrice scalaire** est une matrice de la forme

$$\operatorname{diag}(\lambda,\dots,\lambda) = \begin{pmatrix} \lambda & & (0) \\ & \lambda & \\ & & \ddots & \\ (0) & & & \lambda \end{pmatrix} \in \mathscr{M}_n(\mathbb{K}) \ :,$$

Les plus célèbres étant la matrice nulle et la matrice identité, respectivement :

$$\begin{pmatrix} 0 & \dots & \dots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix} \quad \text{et} \quad I_n = \begin{pmatrix} 1 & & (0) \\ & 1 & \\ & & \ddots & \\ (0) & & & 1 \end{pmatrix} .$$

— On appelle **matrice triangulaire supérieure** toute matrice **carrée** $M = (m_{i,j})_{i,j\in[\![1,n]\!]} \in \mathscr{M}_n(\mathbb{K})$ telle que

$$\forall (i,j) \in [\![1,n]\!]^2,\ (i > j) \Rightarrow (m_{i,j} = 0)\ .$$

Une telle matrice est donc de la forme :

$$\begin{pmatrix} \star & & (\star) \\ & \star & \\ & & \ddots & \\ (0) & & & \star \end{pmatrix}$$

L'ensemble des matrices triangulaires supérieures est noté $T_n^+(\mathbb{K})$. On définit de la même façon l'ensemble $T_n^-(\mathbb{K})$ des matrices triangulaires inférieures ; on a alors naturellement :

$$T_n^+(\mathbb{K}) \cap T_n^-(\mathbb{K}) = D_n(\mathbb{K})\ .$$

b) Structure de $\mathbb{K}$–e.v

Il est aisé de définir deux lois de compositions sur les matrices ; en l'occurrence :

— la somme de deux matrices $M = (m_{i,j})_{i,j}$ et $N = (n_{i,j})_{i,j}$ de $\mathscr{M}_{n,m}(\mathbb{K})$ est définie comme étant la matrice $M + N = (a_{i,j})_{i,j}$ telle que :

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,m]\!], \quad a_{i,j} = m_{i,j} + n_{i,j}\ .$$

Ceci définit une loi de composition interne commutative sur $\mathscr{M}_{n,m}(\mathbb{K})$;

— si $M = (m_{i,j})_{i,j} \in \mathscr{M}_{n,m}(\mathbb{K})$ et $\lambda\mathbb{K}$, on définit la matrice $\lambda M = (a_{i,j})_{i,j}$ *via*

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,m]\!], \quad a_{i,j} = \lambda m_{i,j}\ .$$

Ceci définit une loi de composition externe sur $\mathscr{M}_{n,m}(\mathbb{K})$.

Proposition XXII.1. *$\mathscr{M}_{n,m}(\mathbb{K})$ est un $\mathbb{K}$–e.v de dimension finie de dimension nm.*

Démonstration. La structure d'espace vectoriel se vérifie dans la douleur. Concernant la dimension, posons, pour $(i,j) \in [\![1,n]\!] \times [\![1,m]\!]$:

$$E_{i,j} = (\delta_{i,k}\delta_{j,\ell})_{(k,\ell)\in[\![1,n]\!]\times[\![1,m]\!]} \ .$$

Ceci signifie que tous les coefficients de la matrice $E_{i,j}$ sont nuls, à l'exception de celui situé ligne i, colonne j, qui est égal à 1 ; *e.g* pour $n = m = 2$ nous avons :

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} .$$

Cette famille est clairement génératrice : en effet, si $A = (a_{i,j})_{i,j} \in \mathscr{M}_{n,m}(\mathbb{K})$ on a :

$$A = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{i,j}E_{i,j} \ .$$

De plus, si l'on suppose trouvée une famille $(\lambda_{i,j})_{i,j}$ de scalaires telle que

$$\sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_{i,j}E_{i,j} = 0$$

alors cela signifie que la matrice de coefficients $(\lambda_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ est nulle, *i.e* que tous les $\lambda_{i,j}$ le sont : la famille $\mathscr{F} = (E_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,m]\!]}$ est donc une base de $\mathscr{M}_{n,m}(\mathbb{K})$ *ergo* $\dim(\mathscr{M}_{n,m}(\mathbb{K})) = \mathrm{card}(\mathscr{F}) = nm$. □

2.⁻ Retour sur l'algèbre linéaire "classique"

a) Matrices et vecteurs

Soit E un $\mathbb{K}$–e.v de dimension finie de dimension n et soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de E. Fixons $x \in E$; alors il existe une unique famille $(x_1, \dots, x_n)$ de scalaires telle que :

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

Dans ces conditions, on appelle **matrice de x dans la base $\mathscr{B}$** la matrice

$$\mathrm{mat}_{\mathscr{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathscr{M}_{n,1}(\mathbb{K}) \ .$$

➠ **Exemple XXII.1.** Dans la base $\mathscr{B} = ((1,0),(1,1))$ de $\mathbb{R}^2$, on a, pour tous $x, y \in \mathbb{R}$:

$$(x,y) = (x-y)\cdot(1,0) + y\cdot(1,1)$$

et donc

$$\mathrm{mat}_{\mathscr{B}}((x,y)) = \begin{pmatrix} x-y \\ y \end{pmatrix} .$$

Proposition XXII.2. Soit E un $\mathbb{K}$–e.v de dimension finie de dimension $n \in \mathbb{N}$ et soit $\mathscr{B}$ une base de E. Alors l'application

$$\begin{aligned} \mathrm{mat}_{\mathscr{B}} : E &\to \mathscr{M}_{n,1}(\mathbb{K}) \\ x &\mapsto \mathrm{mat}_{\mathscr{B}}(x) \end{aligned}$$

est un isomorphisme.

Démonstration. La linéarité ne présente aucune difficulté. Notons ensuite que :

$$x \in \mathrm{Ker}(\mathrm{mat}_{\mathscr{B}}) \Leftrightarrow \forall i \in [\![1, n]\!],\ x_i = 0$$

ce qui entraîne l'injectivité. On conclut par égalité dimensionnelle. □

On définit de la même façon la matrice d'une famille de vecteurs $\mathscr{F} = (u_1, \dots, u_k)$ dans une base $\mathscr{B}$ comme la matrice $\mathrm{mat}_{\mathscr{B}}(\mathscr{F}) \in \mathscr{M}_{n,k}(\mathbb{K})$ dont les colonnes sont les matrices des u_i dans la base $\mathscr{B}$.

➠ **Exemple XXII.2.** La matrice dans la base canonique de $\mathbb{R}^3$ de la famille $(1, 0, 1), (0, 2, 3))$ est

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 3 \end{pmatrix} .$$

b) Matrices et applications linéaires

Définition XXII.2. Soit E (resp. F) un $\mathbb{K}$–e.v de dimension finie de dimension m (resp. n) et soit $\mathscr{B}$ (resp. $\mathscr{B}'$) une base de E (resp. de F). Alors, on appelle **matrice** d'une application linéaire $f \in \mathscr{L}(E, F)$ **dans les bases** $\mathscr{B}$ **et** $\mathscr{B}'$ la matrice

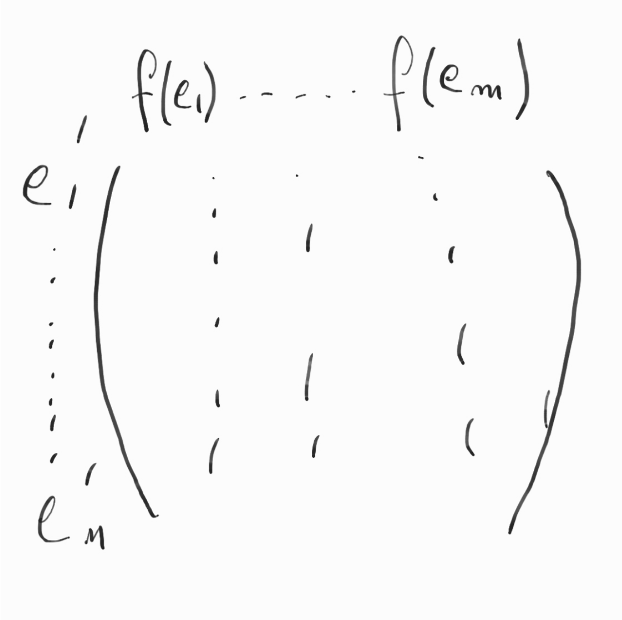
$$\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f) = \mathrm{mat}_{\mathscr{B}'}(f(\mathscr{B})) \in \mathscr{M}_{n,m}(\mathbb{K}) .$$

Notation. Si $f \in \mathscr{L}(E)$ et que $\mathscr{B} = \mathscr{B}'$, on notera $\mathrm{mat}_{\mathscr{B}}(f)$.

Concrètement, cela signifie que les colonnes de cette matrice sont les matrices des images par f des vecteurs de $\mathscr{B}$ dans la base $\mathscr{B}'$. Si on pose $\mathscr{B} = (e_1, \dots, e_m)$ et $\mathscr{B}' = (e'_1, \dots, e'_n)$, alors le coefficient en position (i, j) de la matrice $\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f)$ sera la coordonnée selon e'_i du vecteur $f(e_j)$; i.e en notant $m_{i,j}$ ces coefficients on aura :

$$\forall j \in [\![1, m]\!],\ f(e_j) = \sum_{i=1}^{n} m_{i,j} e'_i .$$

Ceci signifie que la matrice aura donc globalement la tronche suivante :



Exemple XXII.3. On considère l'application suivante :

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R}^3 \\ (x,y) &\mapsto (2x+y, x+y, 3y) . \end{aligned}$$

Il est clair que $f \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^3)$. Posons $\mathscr{B} = ((1,0),(1,1))$ et $\mathscr{B}' = ((1,0,0),(0,1,0),(0,1,1))$: il s'agit la de bases respectives de $\mathbb{R}^2$ et $\mathbb{R}^3$, avec :

$$f((1,0)) = (2,1,0) = 2 \cdot (1,0,0) + 1 \cdot (0,1,0) + 0 \cdot (0,1,1)$$

et

$$f((1,1)) = (3,2,3) = 3 \cdot (1,0,0) - 1 \cdot (0,1,0) + 3 \cdot (0,1,1) .$$

De fait, on a :

$$\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f) = \begin{pmatrix} 2 & 3 \\ 1 & -1 \\ 0 & 3 \end{pmatrix} .$$

Remarque XXII.1. Si E est un $\mathbb{K}$–e.v de dimension finie de dimension n et que $\mathscr{B}$ est une base de E, alors $\mathrm{mat}_{\mathscr{B}}(\mathrm{id}_E) = I_n$.

✖ **Attention :** l'application identité ne donne pas toujours naissance à la matrice identité ; par exemple, si $\mathscr{B} = ((1,0),(1,1))$ et $\mathscr{B}' = ((1,0),(0,1))$ on a :

$$\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} .$$

Proposition XXII.3. Soient E, F deux $\mathbb{K}$–e.v de dimension finie de dimensions respectives m et n. Soient $\mathscr{B}$ et $\mathscr{B}'$ des bases respectives de E et F. Alors l'application

$$\begin{aligned} \varphi : \mathscr{L}(E,F) &\to \mathscr{M}_{n,m}(\mathbb{K}) \\ f &\mapsto \mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f) \end{aligned}$$

est un isomorphisme.

Démonstration. La linéarité se démontre par un calcul rapide. De plus, cette application est bijective car toute application linéaire est totalement déterminée par la donnée de l'image d'une base de son ensemble de départ. □

✌ **Remarque XXII.2.** Soit $A \in \mathscr{M}_{n,m}(\mathbb{K})$; alors d'après la proposition *supra*, il existe une unique application $f \in \mathscr{L}(\mathbb{K}^m, \mathbb{K}^n)$ telle que A soit la matrice de f dans les bases canoniques respectives de $\mathbb{K}^m$ et $\mathbb{K}^n$. Il nous est donc possible de poser $\mathrm{Ker}(A) = \mathrm{Ker}(f)$, $\mathrm{Im}(A) = \mathrm{Im}(f)$ et $\mathrm{rg}(A) = \mathrm{rg}(f)$. Ce paradigme est régulièrement utilisé aux concours pour identifier matrices et applications linéaires : f est alors appelée **application linéaire canoniquement associée à la matrice** A.

Dans ces conditions, il est clair que les colonnes de A forment une famille génératrice de l'image de f.

3.– Produit matriciel

a) Composition et autres joyeusetés

Rappelons nous : si E est $\mathbb{K}$–e.v de dimension finie, l'espace $(\mathscr{L}(E), +, \circ)$ est un anneau. Nous savons que si $n = \dim(E)$, $\mathscr{L}(E)$ est isomorphe (en tant que $\mathbb{K}$–e.v) à $\mathscr{M}_n(\mathbb{K})$. Ceci amène fort naturellement la question suivante : existe–t–il une seconde LCI sur $\mathscr{M}_n(\mathbb{K})$ qui le munirait de la structure d'anneau ? Celle ci devrait par essence être un analogue de la composition ...

Donnons nous trois $\mathbb{K}$–e.v de dimension finie E, F et G, de dimensions respectives p, q et r et munis respectivement de bases $\mathscr{B} = (e_1, \dots, e_p)$, $\mathscr{B}' = (e'_1, \dots, e'_q)$ et $\mathscr{B}'' = (e''_1, \dots, e''_r)$. Fixons ensuite $u \in \mathscr{L}(E, F)$ et $v \in \mathscr{L}(F, G)$ et posons $A = (a_{i,j})_{i,j} = \mathrm{mat}_{\mathscr{B}',\mathscr{B}''}(v)$ ainsi que $B = (b_{i,j})_{i,j} = \mathrm{mat}_{\mathscr{B},\mathscr{B}'}(u)$. On a alors, pour tout $j \in [\![1, p]\!]$:

$$\begin{aligned}
v \circ u(e_j) &= v\left(\sum_{k=1}^{q} b_{k,j} e'_k\right) \\
&= \sum_{k=1}^{q} b_{k,j} v(e'_k) \\
&= \sum_{k=1}^{q} b_{k,j} \sum_{i=1}^{r} a_{i,k} e''_i \\
&= \sum_{i=1}^{r} \left(\sum_{k=1}^{q} a_{i,k} b_{k,j}\right) e''_i \,.
\end{aligned}$$

Ceci signifie que la matrice de $v \circ u$ dans les bases $\mathscr{B}$ et $\mathscr{B}''$ admet pour coefficient en position (i, j), pour $i \in [\![1, r]\!]$ et $j \in [\![1, p]\!]$:

$$\sum_{k=1}^{q} a_{i,k} b_{k,j} \,.$$

b) Définition et conséquences

Définition XXII.3. Soient $A = (a_{i,j})_{i,j} \in \mathscr{M}_{r,q}(\mathbb{K})$ et $B = (b_{i,j})_{i,j} \in \mathscr{M}_{q,p}(\mathbb{K})$. On appelle **produit** de A par B la matrice, notée $A \times B \in \mathscr{M}_{r,p}(\mathbb{K})$ dont le coefficient en position (i, j), pour $i \in [\![1, r]\!]$ et $j \in [\![1, p]\!]$, est :

$$\sum_{k=1}^{q} a_{i,k} b_{k,j} .$$

✖ **Attention :** il ne nous est possible de former le produit de deux matrices A et B que si leurs dimensions sont **compatibles** : il est nécessaire que le nombre de colonnes de A soit égal au nombre de lignes de B.

➠ **Exemple XXII.4.**

—

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} .$$

— Soient $(i, j) \in [\![1, r]\!] \times [\![1, q]\!]$ et $(k, \ell) \in [\![1, q]\!] \times [\![1, p]\!]$. Alors il est possible de former le produit $E_{i,j} \times E_{k,\ell} \in \mathscr{M}_{r,p}(\mathbb{K})$ et cette matrice admet pour coefficient en position $(a, b) \in [\![1, r]\!] \times [\![1, p]\!]$:

$$\sum_{c=1}^{q} \delta_{i,a}\delta_{j,c}\delta_{c,k}\delta_{\ell,b} = \delta_{i,a}\delta_{\ell,b} \sum_{c=1}^{q} \delta_{j,c}\delta_{c,k} .$$

La somme de droite *supra* est non nulle à la seule condition que $j = k$, et égale à 1 dans ce cas, entraînant que :

$$\sum_{c=1}^{q} \delta_{i,a}\delta_{j,c}\delta_{c,k}\delta_{\ell,b} = \delta_{i,a}\delta_{\ell,b}\delta_{j,k}$$

le produit des deux symboles de Kronecker de gauche étant reconnaissable (si, si) comme le coefficient en position (a, b) de la matrice $E_{i,\ell}$. On en déduit donc que :

$$E_{i,j} \times E_{k,\ell} = \delta_{j,k} E_{i,\ell} .$$

✌ **Remarque XXII.3.** Cette définition étant posée, il découle de notre abominable laïus calculatoire la précédant que cette nouvelle opération présente de forte similitude avec la composition des applications linéaires. Plus précisément, et en conservant les notations du début de cette section, nous avons :

(i) $\mathrm{mat}_{\mathscr{B},\mathscr{B}''}(v \circ u) = \mathrm{mat}_{\mathscr{B}',\mathscr{B}''}(v) \times \mathrm{mat}_{\mathscr{B},\mathscr{B}'}(u)$;

(ii) $\forall x \in E$, $\mathrm{mat}_{\mathscr{B}'}(u(x)) = \mathrm{mat}_{\mathscr{B},\mathscr{B}'}(u) \times \mathrm{mat}_{\mathscr{B}}(x)$.

Ceci nous permet de déduire que le produit matriciel est **associatif** et **bilinéaire** (si A, B, C sont des matrices raisonnables et $\lambda \in \mathbb{K}$, alors $(A+\lambda B)C = AC + \lambda BC$).

Proposition XXII.4. Soit $n \geq 0$. Alors $(\mathscr{M}_n(\mathbb{K}), +, \times, \cdot)$ est une algèbre. De plus, pour tout $\mathbb{K}$–e.v de dimension finie E de dimension n, l'algèbre $(\mathscr{L}(E), +, \circ, \cdot)$ est isomorphe à $\mathscr{M}_n(\mathbb{K})$.

Démonstration. Il s'agit d'une synthèse des propriétés évoquées précédemment dans ce chapitre. Notons que le neutre pour le produit matriciel dans $\mathscr{M}_n(\mathbb{K})$ est la matrice identité I_n. □

✖ **Attention :** cette algèbre n'est évidemment **pas** commutative dés que $n \geq 2$: en effet

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

et

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

◇ Abominations diverses

L'algèbre $\mathscr{M}_n(\mathbb{K})$ n'est pas un endroit particulièrement bien famé ; on y trouve divers objets inconnus dans des contrées civilisées, dont nous donnons quelques exemples ci–ensuite, de façon à encourager notre lecteur à une saine paranoïa en matière de calcul matriciel.

Diviseurs de zéro. La nullité d'un produit matriciel n'entraîne absolument pas celle de l'un des facteurs ; par exemple :

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

Nilpotents. Pire, certaines matrices non nulles ont la discourtoisie d'admettre une puissance nulle ; par exemple :

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

Compte tenu de ces exemples, il serait réellement de mauvais ton de considérer le produit matriciel comme intègre. À bon entendeur ...

◇ Binôme de Newton

Notons, en contraste avec le musée des horreurs *supra*, que si deux matrices $A, B \in \mathscr{M}_n(\mathbb{K})$ sont telles que $AB = BA$ (on dit qu'elles **commutent**), nous disposons de la formule du binôme de Newton (proposition X.10), pour $p \in \mathbb{N}$:

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

✍ **Exercice XXII.1.** Calculer les puissances successives de la matrice $A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

➥ **Correction :** *Remarquons que* $A = D + N$ *avec*

$$D = 3I_3 \quad \textit{et} \quad N = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} ,$$

ces deux matrices commutant. De plus, pour tout $k \geq 2$, on a $N^k = 0$, ce qui entraîne que, pour tout $p \in \mathbb{N}^*$:

$$\begin{aligned} A^p &= (D+N)^p \\ &= \sum_{k=0}^{p} \binom{p}{k} N^k D^{p-k} \\ &= N^0 D^p + pND^{p-1} \\ &= 3^p I_3 + p3^{p-1} N \\ &= \begin{pmatrix} 3^p & 0 & p3^{p-1} \\ 0 & 3^p & 0 \\ 0 & 0 & 3^p \end{pmatrix} \end{aligned}$$

c) Groupe linéaire

Proposition XXII.5. Soit $A \in \mathscr{M}_n(\mathbb{K})$; alors est inversible pour le produit matriciel à gauche si et seulement si elle est inversible à droite.

Démonstration. Soit $f \in \mathscr{L}(\mathbb{K}^n)$ l'application linéaire canoniquement associée à A. Alors, A est inversible à gauche si et seulement si il existe une matrice $B \in \mathscr{M}_n(\mathbb{K})$ telle que $BA = I_n$. Ceci signifie que si $g \in \mathscr{L}(\mathbb{K}^n)$ est l'application linéaire canoniquement associée à B on a $g \circ f = \mathrm{id}_{\mathbb{K}^n}$: f est donc injective et donc, par égalité de dimensions, surjective. Il en découle que $f \circ g = \mathrm{id}_{\mathbb{K}^n}$, *i.e* $AB = I_n$, d'où le résultat. □

Définition XXII.4. Le groupes des inversibles de l'anneau $(\mathscr{M}_n(\mathbb{K}), +, \times)$ est appelé **groupe linéaire** d'ordre n sur $\mathbb{K}$.

Notation. $GL_n(\mathbb{K}) = \mathscr{M}_n(\mathbb{K})^\times$.

✖ **Attention :** seules les matrices **carrées** peuvent être inversibles.

✌ **Remarque XXII.4.** Le groupe $(GL_n(\mathbb{K}), \times)$ est donc isomorphe au groupe $(GL(E), \circ)$ via restriction de l'isomorphisme entre $\mathscr{L}(E)$ et $\mathscr{M}_n(\mathbb{K})$. En particulier, si $f \in GL(E)$ et que $\mathscr{B}$ est une base de E on a $\mathrm{mat}_{\mathscr{B}}(f^{-1}) = \mathrm{mat}_{\mathscr{B}}(f)^{-1}$.

➠ **Exemple XXII.5.** Soient $a, b, c, d \in \mathbb{R}$; alors on a vu dans le chapitre XIX que l'application linéaire canoniquement associée à $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ étant dans $GL(E)$ si et seulement si $ad - bc \neq 0$. Il en va donc de même pour l'inversibilité de A.

✌ **Remarque XXII.5.** Tout ceci a pour conséquence la chose suivante : une matrice A est inversible si et seulement $\mathrm{Ker}(A) = \{0\}$.

Proposition XXII.6. Soit $M = (m_{i,j})_{i,j\in[\![1,n]\!]}$ une matrice triangulaire de $\mathscr{M}_n(\mathbb{K})$. Alors :

$$M \in GL_n(\mathbb{K}) \iff \forall k \in [\![1,n]\!],\ m_{k,k} \neq 0\,.$$

Dans ce cas, l'inverse de M est également triangulaire, dans le même "sens" que M.

Démonstration. Supposons pour fixer les idées que $M \in T_n^+(\mathbb{K})$. Soit $\mathscr{B} = (e_1,\dots,e_n)$ la base canonique de $\mathbb{K}^n$ et soit $f \in \mathscr{L}(\mathbb{K}^n)$ l'endomorphisme canoniquement associé à M : on a alors $\mathrm{mat}_{\mathscr{B}}(f) = M$. Posons ensuite :

— $E_0 = \{0\}$;
— $\forall k \in [\![1,n]\!]$, $E_k = \mathrm{Vect}(e_1,\dots,e_k)$.

Comme M est triangulaire supérieure, on a naturellement que $f(E_k) \subset E_k$ pour tout $k \geq 0$ car, pour tout $j \in [\![1,k]\!]$:

$$\begin{aligned} f(e_j) &= \sum_{i=1}^{n} \underbrace{m_{i,j}}_{=0 \text{ si } i>j} e_i \\ &= \sum_{i=1}^{j} m_{i,j} e_i \\ &\in E_j \subset E_k\,. \end{aligned}$$

Cas 1 : il existe $k_0 \in [\![1,n]\!]$ tel que $m_{k_0,k_0} = 0$. Ceci signifie que $f(e_{k_0})$ a une composante nulle selon e_{k_0} et donc $f(E_{k_0}) \subset E_{k_0-1}$. Or, $\dim(E_{k_0}) = k_0 > k_0-1 = \dim(E_{k_0-1})$: f ne peut être injective *ergo* $M \notin GL_n(\mathbb{K})$.

Cas 2 : $\forall k \in [\![1,n]\!]$, $m_{k,k} \neq 0$. Démontrons par récurrence sur $k \in [\![0,n]\!]$ que, pour tout tel k, $f(E_k) = E_k$.

— $k = 0$: trivial.
— Supposons la propriété vraie au rang $k \in [\![0,n-1]\!]$. Alors :

$$\begin{aligned} f(e_{k+1}) &= \sum_{i=1}^{k+1} m_{i,k+1} e_i \\ &= m_{k+1,k+1} e_{k+1} + \underbrace{\sum_{i=1}^{k} m_{i,k+1} e_i}_{\in E_k}\,. \end{aligned}$$

Or, $f(E_k) = E_k$ par hypothèse de récurrence et donc il existe $u \in E_k$ tel que :

$$\sum_{i=1}^{k} m_{i,k+1} e_i = f(u)$$

et donc

$$\begin{aligned} e_{k+1} &= \frac{1}{m_{k+1,k+1}}\left(f(e_{k+1}) - f(u)\right) \\ &= \frac{1}{m_{k+1,k+1}} f(e_{k+1} - u) \\ &\in f(E_{k+1})\,. \end{aligned}$$

On en déduit que $E_{k+1} \subset f(E_{k+1})$ et donc le résultat voulu.
En particulier, $f(E_n) = E_n$; comme $E_n = \mathbb{K}^n$ on a donc la surjectivité (et donc la bijectivité par égalité dimensionnelle) de f, d'où le résultat.

Dans le cas où f est bijective, on a de plus pour tout $k \in [\![1, n]\!]$,

$$f^{-1}(E_k) = f^{-1}(f(E_k)) = E_k$$

et donc M^{-1} est donc bien triangulaire supérieure. □

d) Produit par blocs

Pour manipuler des matrices de tailles plus ou moins déraisonnables, nous pourrons avoir recours à l'écriture "par blocs", *i.e* nous poserons des matrices sous la forme :

$$M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

avec A, B, C, D des matrices de tailles compatibles. Il nous sera possible d'utiliser les opérations classiques sur les matrices directement sur les blocs, pour peu que ceux–ci soient raisonnables (*i.e* que les dites opérations soient légales).

➠ **Exemple XXII.6.** Soit $M = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}$; alors on peut écrire M sous la forme

$$M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

avec $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 \end{pmatrix}$ et $D = (5)$. De fait :

$$\begin{aligned} M^2 &= \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \\ &= \left(\begin{array}{c|c} A^2 + BC & AB + BD \\ \hline CA + DC & CB + D^2 \end{array}\right) \\ &= \left(\begin{array}{c|c} A^2 & 0 \\ \hline 0 & D^2 \end{array}\right) \end{aligned}$$

ce qui permet de déduire que :

$$M^2 = \begin{pmatrix} 7 & 10 & 0 \\ 15 & 22 & 0 \\ 0 & 0 & 25 \end{pmatrix}.$$

4.⁻ Transposition

a) Qu'est–ce ?

Définition XXII.5. Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_{n,m}(\mathbb{K})$. On appelle **transposée** de A la matrice de $\mathscr{M}_{m,n}(\mathbb{K})$ dont le coefficient en position $(i,j) \in [\![1,m]\!] \times [\![1,n]\!]$ est $a_{j,i}$.

Notation. ${}^{\mathrm{t}}A$ ou A^T.

Exemple XXII.7.

—
$$ {}^{\mathrm{t}}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} ; $$

—
$$ {}^{\mathrm{t}}\begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 42 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 2 \\ 0 & 42 \end{pmatrix} . $$

Proposition XXII.7. Soient $A, B \in \mathscr{M}_{n,m}(\mathbb{K})$ et soit $\lambda \in \mathbb{K}$. Alors :

(i) ${}^{\mathrm{t}}(A + \lambda B) = {}^{\mathrm{t}}A + \lambda\, {}^{\mathrm{t}}B$;

(ii) ${}^{\mathrm{t}}({}^{\mathrm{t}}A) = A$;

(iii) si $C \in \mathscr{M}_{m,p}(\mathbb{K})$, ${}^{\mathrm{t}}(AC) = {}^{\mathrm{t}}C\, {}^{\mathrm{t}}A$.

Démonstration. Les points (i) et (ii) sont triviaux. Pour le point (iii), fixons $(i,j) \in [\![1,p]\!] \times [\![1,n]\!]$; alors le coefficient en position (i,j) de ${}^{\mathrm{t}}C\, {}^{\mathrm{t}}A$ est

$$ \sum_{k=0}^{m} c_{k,i} a_{j,k} . $$

Il s'agit donc bien du coefficient situé ligne j, colonne i de la matrice AC, d'où le résultat. □

Remarque XXII.6. Le point (i) entraîne que l'application $A \mapsto {}^{\mathrm{t}}A$ est linéaire de $\mathscr{M}_{n,m}(\mathbb{K})$ vers $\mathscr{M}_{m,n}(\mathbb{K})$.

Corollaire XXII.7.a. Soit $A \in GL_n(\mathbb{K})$. Alors ${}^{\mathrm{t}}A \in GL_n(\mathbb{K})$ et ${}^{\mathrm{t}}(A^{-1}) = ({}^{\mathrm{t}}A)^{-1}$

Exercice XXII.2. Soit $X \in \mathscr{M}_{n,1}(\mathbb{K})$. Exprimer de façon simple la quantité ${}^{\mathrm{t}}XX$.

➥ **Correction :** *Posons* $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. *Alors* ${}^{\mathrm{t}}XX \in \mathscr{M}_{1,1}(\mathbb{K})$ *aura pour seul coefficient la quantité :*

$$\sum_{k=1}^{n} x_k^2 \,.$$

b) Matrices (anti)symétriques

☞ ***Le contenu de ce paragraphe est hors–programme, mais fort utile pour la seconde année.***

Définition XXII.6. Soit $A \in \mathscr{M}_n(\mathbb{K})$ une matrice **carrée**. On dira que A est :
— **symétrique** si $A = {}^{\mathrm{t}}A$;
— **antisymétrique** si $A = -\,{}^{\mathrm{t}}A$.

Notation. On notera $S_n(\mathbb{K})$ l'ensemble des matrices symétriques $n \times n$ et $A_n(\mathbb{K})$ celui des matrices antisymétriques.

➠ **Exemple XXII.8.** Les matrices diagonales sont symétriques. La matrice $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est antisymétrique.

✌ **Remarque XXII.7.** Une matrice A est symétrique si et seulement si pour tout $(i,j) \in [\![1,n]\!]^2$ on a $a_{i,j} = a_{j,i}$. De même, une telle matrice sera antisymétrique à condition d'avoir $a_{i,j} = -a_{j,i}$; en particulier, $a_{i,i} = -a_{i,i}$. La diagonale d'une matrice antisymétrique est donc toujours nulle.

Proposition XXII.8. Les ensembles $S_n(\mathbb{K})$ et $A_n(\mathbb{K})$ sont des espaces vectoriels de dimensions respectives $\frac{n(n+1)}{2}$ et $\frac{n(n-1)}{2}$. De plus, on a :

$$\mathscr{M}_n(\mathbb{K}) = S_n(\mathbb{K}) \oplus A_n(\mathbb{K}) \,.$$

Démonstration. Considérons l'application

$$\begin{aligned} \phi : \mathscr{M}_n(\mathbb{K}) &\to \mathscr{M}_n(\mathbb{K}) \\ A &\mapsto {}^{\mathrm{t}}A \,. \end{aligned}$$

Il s'agit (proposition XXII.7) d'un endomorphisme de $\mathscr{M}_n(\mathbb{K})$ vérifiant la relation $\phi^2 = \mathrm{id}_{\mathscr{M}_n(\mathbb{K})}$: ϕ est donc une symétrie de $\mathscr{M}_n(\mathbb{K})$. De fait, par la proposition XIX.14, on a :

$$\mathscr{M}_n(\mathbb{K}) = \mathrm{Ker}(\phi - \mathrm{id}_{\mathscr{M}_n(\mathbb{K})}) \oplus \mathrm{Ker}(\phi + \mathrm{id}_{\mathscr{M}_n(\mathbb{K})}) \,.$$

De plus, si $M \in \mathscr{M}_n(\mathbb{K})$ alors :

$$\begin{aligned} M \in \mathrm{Ker}(\phi - \mathrm{id}_{\mathscr{M}_n(\mathbb{K})}) &\Leftrightarrow (\phi - \mathrm{id}_{\mathscr{M}_n(\mathbb{K})})(M) = 0 \\ &\Leftrightarrow {}^{\mathrm{t}}M - M = 0 \\ &\Leftrightarrow M \in S_n(\mathbb{K}) \end{aligned}$$

et donc $\mathrm{Ker}(\phi - \mathrm{id}_{\mathscr{M}_n(\mathbb{K})}) = S_n(\mathbb{K})$. Symétriquement, on vérifie que $\mathrm{Ker}(\phi + \mathrm{id}_{\mathscr{M}_n(\mathbb{K})}) = A_n(\mathbb{K})$.

Concernant la dimension, il suffit de remarquer que $S_n(\mathbb{K})$ est engendré par les matrices $E_{i,j} + E_{j,i}$ pour $i, j \in [\![1, n]\!]$ et que cette famille est libre ; de fait $\dim(S_n(\mathbb{K})) = \dfrac{n(n+1)}{2}$ et $\dim(A_n(\mathbb{K})) = \dim(\mathscr{M}_n(\mathbb{K})) - \dim(S_n(\mathbb{K})) = \dfrac{n(n-1)}{2}$. □

✌ **Remarque XXII.8.** En pratique, si $M \in \mathscr{M}_n(\mathbb{K})$, on a :

$$M = \underbrace{\frac{M + {}^{\mathrm{t}}M}{2}}_{\in S_n(\mathbb{K})} + \underbrace{\frac{M - {}^{\mathrm{t}}M}{2}}_{\in A_n(\mathbb{K})} .$$

5.– Bases, trace

a) Matrices de passage

On fixe dans ce paragraphe un $\mathbb{K}$–e.v de dimension finie E de dimension $n \in \mathbb{N}$.

Définition XXII.7. Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. On appelle **matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$** la matrice

$$P_{\mathscr{B}}^{\mathscr{B}'} = \mathrm{mat}_{\mathscr{B}}(\mathscr{B}') .$$

✌ **Remarque XXII.9.** On a donc $P_{\mathscr{B}}^{\mathscr{B}'} = \mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E)$.

➠ **Exemple XXII.9.** La matrice de passage de $((1,0),(0,1))$ à $((1,0),(1,1))$ est

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

et celle de $((1,0),(1,1))$ à $((1,0),(0,1))$ est

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

Proposition XXII.9. Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E ; alors :

(i) $P_{\mathscr{B}}^{\mathscr{B}'} \in GL_n(\mathbb{K})$;

(ii) $(P_{\mathscr{B}}^{\mathscr{B}'})^{-1} = P_{\mathscr{B}'}^{\mathscr{B}}$.

Démonstration. Nous savons que, d'après le lien produit–composition vu précédemment :

$$\begin{aligned} P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}} &= \mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E)\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_E) \\ &= \mathrm{mat}_{\mathscr{B}}(\mathrm{id}_E \circ \mathrm{id}_E) \\ &= \mathrm{mat}_{\mathscr{B}}(\mathrm{id}_E) \\ &= I_n \end{aligned}$$

d'où le résultat. □

L'usage principal à ce stade des matrices de passage sera, pour nous, le **changement de base**, *i.e* le procédé permettant de passer des coordonnées ou de la matrice d'un vecteur ou d'une application linéaire dans la base $\mathscr{B}$ à leur homologue dans $\mathscr{B}'$. Ceci générera de nombreux calculs hilarants.

Proposition XXII.10. Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E et soit $x \in E$. Alors :

$$\mathrm{mat}_{\mathscr{B}}(x) = P_{\mathscr{B}}^{\mathscr{B}'} \mathrm{mat}_{\mathscr{B}'}(x)\,.$$

Démonstration. Toujours d'après les propriétés élémentaires du produit matriciel, nous avons :

$$\begin{aligned}\mathrm{mat}_{\mathscr{B}}(x) &= \mathrm{mat}_{\mathscr{B}}(\mathrm{id}_E(x)) \\ &= \mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E)\mathrm{mat}_{\mathscr{B}'}(x) \\ &= P_{\mathscr{B}}^{\mathscr{B}'} \mathrm{mat}_{\mathscr{B}'}(x)\end{aligned}$$

d'où le résultat. □

Exemple XXII.10. Pour obtenir les coordonnées d'un vecteur $(x, y) \in \mathbb{R}^2$ dans la base $((1, 0), (1, 1))$ il nous suffit de calculer

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x-y \\ y \end{pmatrix}.$$

Proposition XXII.11. Soient E et F deux $\mathbb{K}$–e.v de dimension finie ; soient $\mathscr{B}_1$, $\mathscr{B}_1'$ deux bases de E et $\mathscr{B}_2$, $\mathscr{B}_2'$ deux bases de F. Alors, pour tout $f \in \mathscr{L}(E, F)$ on a :

$$\mathrm{mat}_{\mathscr{B}_1',\mathscr{B}_2'}(f) = P_{\mathscr{B}_2'}^{\mathscr{B}_2} \mathrm{mat}_{\mathscr{B}_1,\mathscr{B}_2}(f) P_{\mathscr{B}_1}^{\mathscr{B}_1'}\,.$$

Démonstration. Commençons par noter que $f = \mathrm{id}_F \circ f \circ \mathrm{id}_E$; ainsi, on a :

$$\begin{aligned}\mathrm{mat}_{\mathscr{B}_1',\mathscr{B}_2'}(f) =& \mathrm{mat}_{\mathscr{B}_1',\mathscr{B}_2'}(\mathrm{id}_F \circ f \circ \mathrm{id}_E) \\ &= \mathrm{mat}_{\mathscr{B}_2,\mathscr{B}_2'}(\mathrm{id}_F)\mathrm{mat}_{\mathscr{B}_1,\mathscr{B}_2}(f)\mathrm{mat}_{\mathscr{B}_1',\mathscr{B}_1}(\mathrm{id}_E) \\ &= P_{\mathscr{B}_2'}^{\mathscr{B}_2}\mathrm{mat}_{\mathscr{B}_1,\mathscr{B}_2}(f)P_{\mathscr{B}_1}^{\mathscr{B}_1'}\end{aligned}$$

d'où, pour changer, le résultat. □

Remarque XXII.10. Si $E = F$, $\mathscr{B}_1 = \mathscr{B}_2$ et $\mathscr{B}_1' = \mathscr{B}_2'$, on a alors, en posant $P = P_{\mathscr{B}_1}^{\mathscr{B}_1'}$:

$$\mathrm{mat}_{\mathscr{B}_1'}(f) = P^{-1}\mathrm{mat}_{\mathscr{B}_1}(f)P\,.$$

Exercice XXII.3. Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans une base $\mathscr{B} = (e_1, e_2, e_3)$ de $\mathbb{R}^3$ est

$$A = \begin{pmatrix} 15 & -11 & 5 \\ 20 & -15 & 8 \\ 8 & -7 & 6 \end{pmatrix}.$$

Montrer que les vecteurs

$$e'_1 = 2e_1 + 3e_2 + e_3, \quad e'_2 = 3e_1 + 4e_2 + e_3, \quad e'_3 = e_1 + 2e_2 + 2e_3$$

forment une base $\mathscr{B}'$ de $\mathbb{R}^3$ et calculer la matrice de f dans celle–ci.

➥ **Correction :** *On vérifie rapidement que $\mathscr{B}'$ est libre et de cardinal 3 : il s'agit bien d'une base. Soit $P = P_{\mathscr{B}}^{\mathscr{B}'}$; alors :*

$$P = \begin{pmatrix} 2 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 1 & 2 \end{pmatrix}$$

Trouver P^{-1} revient à trouver $a, b, c, d, e, f, g, h, i \in \mathbb{R}$ tels que :

$$\begin{pmatrix} 2 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Nous y parvenons en résolvant le système linéaire ainsi obtenu et trouvons :

$$P^{-1} = \begin{pmatrix} -6 & 5 & -2 \\ 4 & -3 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Ne reste qu'à calculer

$$B = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

qui est la matrice de f dans la base $\mathscr{B}'$.

b) Trace

Définition XXII.8. Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$ une matrice **carrée**. On appelle **trace** de A la quantité :

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i} .$$

➠ **Exemple XXII.11.** $\mathrm{Tr}(I_n) = n$, $\mathrm{Tr}\left(\begin{pmatrix} 1 & 0 \\ -12 & 3 \end{pmatrix}\right) = 4$.

Proposition XXII.12. Soient $A, B \in \mathscr{M}_n(\mathbb{K})$ et soit $\lambda \in \mathbb{K}$. Alors :

(i) $\mathrm{Tr}(A + \lambda B) = \mathrm{Tr}(A) + \lambda \mathrm{Tr}(B)$;

(ii) $\mathrm{Tr}({}^{\mathrm{t}}A) = \mathrm{Tr}(A)$;

(iii) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Démonstration. Les point (i) et (ii) découlent directement de la définition. Pour le point (iii), remarquons que si on note $a_{i,j}$ (resp. $b_{i,j}$) les coefficients de A (resp. de B) on a :

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,i}$$

et

$$\begin{aligned}\mathrm{Tr}(BA) &= \sum_{i=1}^{n}\sum_{k=1}^{n} b_{i,k}a_{k,i}\\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} a_{k,i}b_{i,k}\\ &= \mathrm{Tr}(AB)\end{aligned}$$

ce qui constitue le résultat voulu. □

✌ **Remarque XXII.11.** La trace est donc une forme linéaire sur $\mathscr{M}_n(\mathbb{K})$.

Proposition/définition XXII.9. Soit E un $\mathbb{K}$–e.v de dimension finie et soit $\mathscr{B}$ une base de E. Alors, si $f \in \mathscr{L}(E)$, la quantité $\mathrm{Tr}(\mathrm{mat}_{\mathscr{B}}(f))$ est indépendante du choix de la base $\mathscr{B}$ et est appelée **trace** de f.

Notation. $\mathrm{Tr}(f)$

Démonstration. Fixons $\mathscr{B}'$ une base de E et posons $M = \mathrm{mat}_{\mathscr{B}}(f)$, $M' = \mathrm{mat}_{\mathscr{B}'}(f)$ et $P = P_{\mathscr{B}}^{\mathscr{B}'}$. Alors :

$$\begin{aligned}\mathrm{Tr}(M') &= \mathrm{Tr}(P^{-1}MP)\\ &= \mathrm{Tr}(PP^{-1}M)\\ &= \mathrm{Tr}(M)\,.\end{aligned}$$

□

Exemple XXII.12.
— $\mathrm{Tr}(\mathrm{id}_E) = \dim(E)$.
— Soit $p \in \mathscr{L}(E)$ un projecteur ; et soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de E adaptée à la décomposition $E = \mathrm{Ker}(p) \oplus \mathrm{Im}(p)$. Alors, si $k = \dim(\mathrm{Ker}(p))$ on a :

(i) $\forall i \in [\![1, k]\!]$, $p(e_i) = 0$;

(ii) $\forall i \in [\![k+1, n]\!]$, $p(e_i) = e_i$.

On a donc :

$$\mathrm{mat}_{\mathscr{B}}(p) = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-k}\end{array}\right)$$

ce qui implique que $\mathrm{Tr}(p) = n - k = \mathrm{rg}(p)$.

c) Matrices semblables

Définition XXII.10. Soient $A, B \in \mathscr{M}_n(\mathbb{K})$. On dit que A et B sont **semblables** si il existe $P \in GL_n(\mathbb{K})$ telle que :

$$B = P^{-1}AP\,.$$

➠ **Exemple XXII.13.** Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases d'un $\mathbb{K}$–e.v de dimension finie E et soit $f \in \mathscr{L}(E)$. Alors, si on pose $P = P_{\mathscr{B}}^{\mathscr{B}'}$, on a :

$$\mathrm{mat}_{\mathscr{B}'}(f) = P^{-1}\mathrm{mat}_{\mathscr{B}}(f)P$$

et donc $\mathrm{mat}_{\mathscr{B}'}(f)$ et $\mathrm{mat}_{\mathscr{B}}(f)$ sont semblables.

Proposition XXII.13. La similitude des matrices est une relation d'équivalence.

Démonstration. Voici un fort bel exercice pour notre lecteur favori. Have fun. □

Proposition XXII.14. Soient $A, B \in \mathscr{M}_n(\mathbb{K})$ deux matrices semblables. Alors $\mathrm{Tr}(A) = \mathrm{Tr}(B)$.

Démonstration. Par similitude, il existe $P \in GL_n(\mathbb{K})$ telle que $B = P^{-1}AP$. On a donc :

$$\begin{aligned} \mathrm{Tr}(B) &= \mathrm{Tr}(P^{-1}AP) \\ &= \mathrm{Tr}(PP^{-1}A) \\ &= \mathrm{Tr}(A) \end{aligned}$$

youpi. □

✖ **Attention :** la réciproque est brutalement fausse : la matrice nulle et la matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ont la même trace sans être semblables.

6.⁻ Rang d'une matrice

a) C'est quoi ?

Définition XXII.11. Soit $A \in \mathscr{M}_{n,m}(\mathbb{K})$. On appelle **rang** de A le rang de la famille de ses vecteurs colonnes.

Notation. $\mathrm{rg}(A)$

Remarque XXII.12. Il s'agit *de facto* également du rang de l'application linéaire canoniquement associée à A. Il en découle qu'une matrice de $\mathscr{M}_n(\mathbb{K})$ est inversible si et seulement si son rang est égal à n.

Exemple XXII.14. $\mathrm{rg}(I_n) = n$, $\mathrm{rg}\left(\begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix}\right) = 1$.

Proposition/définition XXII.12 (Matrice canonique d'une application linéaire)**.** Soient E et F deux $\mathbb{K}$–e.v de dimension finie de dimensions respectives m et n. Soit $f \in \mathscr{L}(E,F)$ de rang $r \geq 0$; alors il existe une base $\mathscr{B}$ de E et une base $\mathscr{B}'$ de F telles que :

$$\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f) = J_r$$

où J_r est la **matrice canonique de rang** r donnée par :

$$J_r = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Démonstration. D'après le théorème du rang, nous savons que $\dim(\mathrm{Ker}(f)) = m-r$; fixons nous donc un supplémentaire G de ce dernier et une base $\mathscr{B} = (e_1, \dots, e_m)$ de E adaptée à la décomposition $E = G \oplus \mathrm{Ker}(f)$. La famille $\mathscr{F} = (e_1, \dots, e_r)$ est donc libre et engendre G. De fait, elle n'intersecte pas $\mathrm{Ker}(f)$, ce qui entraîne que $f(\mathscr{F})$ est libre car $f_{|G}$ est injective.

La famille $f(\mathscr{F})$ étant libre dans F, on peut la prolonger en une base de $\mathscr{B}'$ de F. Il est alors aisé de vérifier que $\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f) = J_r$. □

b) Matrices équivalentes

Définition XXII.13. Deux matrices $A, B \in \mathscr{M}_{n,m}(\mathbb{K})$ sont dites **équivalentes** si il existe $P \in GL_n(\mathbb{K})$ et $Q \in GL_m(\mathbb{K})$ tel que :

$$B = P^{-1}AQ\,.$$

Exemple XXII.15. Soient $\mathscr{B}_1, \mathscr{B}'_1$ deux bases d'un $\mathbb{K}$–e.v de dimension finie E et $\mathscr{B}_2, \mathscr{B}'_2$ deux bases d'un $\mathbb{K}$–e.v de dimension finie F. Alors, pour tout $f \in \mathscr{L}(E,F)$ on a :

$$\mathrm{mat}_{\mathscr{B}'_1,\mathscr{B}'_2}(f) = (P_{\mathscr{B}_2}^{\mathscr{B}'_2})^{-1}\mathrm{mat}_{\mathscr{B}_1,\mathscr{B}_2}(f)P_{\mathscr{B}_1}^{\mathscr{B}'_1}$$

et donc $\mathrm{mat}_{\mathscr{B}'_1,\mathscr{B}'_2}(f)$ et $\mathrm{mat}_{\mathscr{B}_1,\mathscr{B}_2}(f)$ sont équivalentes.

Remarque XXII.13. Deux matrices semblables sont équivalentes. La réciproque est évidemment fausse.

Proposition XXII.15. L'équivalence des matrices est une relation d'équivalence.

Démonstration. La démonstration de cette propriété fort surprenante est laissée en exercice au lecteur avide de savoir et d'eau fraîche. □

Proposition XXII.16. Soient $A, B \in \mathscr{M}_{n,m}(\mathbb{K})$. Alors les propriétés suivantes sont équivalentes :

(i) A est équivalente à B ;

(ii) A et B sont matrices d'une même application linéaire ;

(iii) $\mathrm{rg}(A) = \mathrm{rg}(B)$.

Démonstration.

(i)⇒(ii) Soit f l'application canoniquement associée à A et notons $\mathscr{B}$ la base canonique de $\mathbb{K}^m$ et $\mathscr{B}'$ celle de $\mathbb{K}^n$; nous avons donc $A = \mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f)$. Comme B est équivalente à A, il existe $P \in GL_n(\mathbb{K})$ et $Q \in GL_m(\mathbb{K})$ telles que $A = P^{-1}BQ$. Or P (resp. Q) peut être vue comme la matrice de passage $P_{\mathscr{B}'}^{\mathscr{C}}$ (resp. $P_{\mathscr{B}}^{\mathscr{D}}$), où $\mathscr{C}$ (resp. $\mathscr{D}$) est la famille des colonnes de P (resp. Q), ces familles étant des bases de leurs espaces respectives par inversibilité de P et Q. Nous avons donc, *in fine* :

$$\begin{aligned} B &= P_{\mathscr{C}}^{\mathscr{B}'} \mathrm{mat}_{\mathscr{B},\mathscr{B}'}(f) P_{\mathscr{B}}^{\mathscr{D}} \\ &= \mathrm{mat}_{\mathscr{D},\mathscr{C}}(f) \end{aligned}$$

d'où le résultat souhaité.

(ii)⇒(iii) Soit f l'application canoniquement associée à A. Alors B est également matrice de f donc $\mathrm{rg}(A) = \mathrm{rg}(f) = \mathrm{rg}(B)$.

(iii)⇒(i) Soit f (resp. g) l'application linéaire canoniquement associée à A (resp. à B). Comme $\mathrm{rg}(A) = \mathrm{rg}(B)$, on a $\mathrm{rg}(f) = \mathrm{rg}(g)$ et donc f et g admettent pour matrice dans un certain couple de bases (potentiellement différentes pour l'une et l'autre) J_r avec $r = \mathrm{rg}(A)$. De fait, A et B sont équivalentes à Jr. Par transitivité, A est donc équivalente à B.

□

✌ **Remarque XXII.14.** Une conséquence importante de cette proposition est qu'une matrice est de rang r si et seulement si elle est équivalente à J_r.

Proposition XXII.17. Soit $M \in \mathscr{M}_{n,m}(\mathbb{K})$. Alors $\mathrm{rg}({}^{\mathrm{t}}M) = \mathrm{rg}(M)$.

Démonstration. Soit $r = \mathrm{rg}(M)$; alors on sait qu'il existe $P \in GL_n(\mathbb{K})$ et $Q \in GL_m(\mathbb{K})$ telles que $M = P^{-1}J_rQ$ et donc :

$$\begin{aligned} {}^{\mathrm{t}}M &= {}^{\mathrm{t}}(P^{-1}J_rQ) \\ &= {}^{\mathrm{t}}Q\, {}^{\mathrm{t}}J_r({}^{\mathrm{t}}P^{-1}) \\ &= {}^{\mathrm{t}}Q J_r({}^{\mathrm{t}}P)^{-1} \end{aligned}$$

ce qui implique que ${}^{\mathrm{t}}M$ est équivalente à J_r et donc de rang r. □

✌ **Remarque XXII.15.** Toute matrice est donc équivalente à sa transposée.

c) Matrices extraites

Définition XXII.14. Soit $M \in \mathscr{M}_{n,m}(\mathbb{K})$; on appelle **matrice extraite** toute restriction de M à un ensemble de la forme $I \times J$, avec $I \subset [\![1, n]\!]$ et $J \subset [\![1, m]\!]$ **non vides**.

➠ **Exemple XXII.16.** Il s'agit en pratique de ne conserver que certaines lignes et colonnes de la matrice en question : les matrices $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ sont par exemple extraites de $\begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 2 \\ 4 & 0 & 0 \end{pmatrix}$.

✌ **Remarque XXII.16.** Dans la définition *supra*, nous disposons de $2^n - 1$ choix pour I et $2^m - 1$ choix pour J : toute matrice de $\mathscr{M}_{n,m}(\mathbb{K})$ admet donc $(2^n - 1)(2^m - 1)$ matrices extraites.

Proposition XXII.18. Soit $A \in \mathscr{M}_{n,m}(\mathbb{K})$ une matrice de rang $r \in \mathbb{N}$. Alors :

(i) toutes les matrices extraites de A sont de rang inférieur ou égal à r ;

(ii) il existe une matrice extraite de A appartenant à $GL_r(\mathbb{K})$.

Chapitre XXIII

Groupe symétrique, déterminant

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.⁻ Permutations d'un ensemble fini

a) Rappels et compléments

Nous avons eu l'occasion d'évoquer le **groupe symétrique** $(\mathfrak{S}_n, \circ)$, pour $n \geq 1$, dans le chapitre X. Rappelons ici qu'il s'agit du groupe $\mathfrak{S}([\![1,n]\!],[\![1,n]\!])$ des bijections de l'ensemble $[\![1,n]\!]$ vers lui–même (appelées **permutations**). Il s'agit d'un groupe fini possédant $n!$ éléments.

Notation. Soit $\sigma \in \mathfrak{S}_n$; nous utiliserons pour σ la notation "raccourcie" suivante :

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} .$$

➠ **Exemple XXIII.1.**

— $\mathfrak{S}_2$ est constitué des permutations $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

— De son côté, $\mathfrak{S}_3$ contient exactement 6 éléments, qui sont $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$.

Notation. Pour alléger les calculs, nous nous autoriserons à noter, pour $\sigma, \tau \in \mathfrak{S}_n$, la composition $\sigma \circ \tau$ simplement $\sigma\tau$.

➠ **Exemple XXIII.2.**

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} .$$

✖ **ATTENTION :** attention à l'ordre ; une composition se lit de droite à gauche.

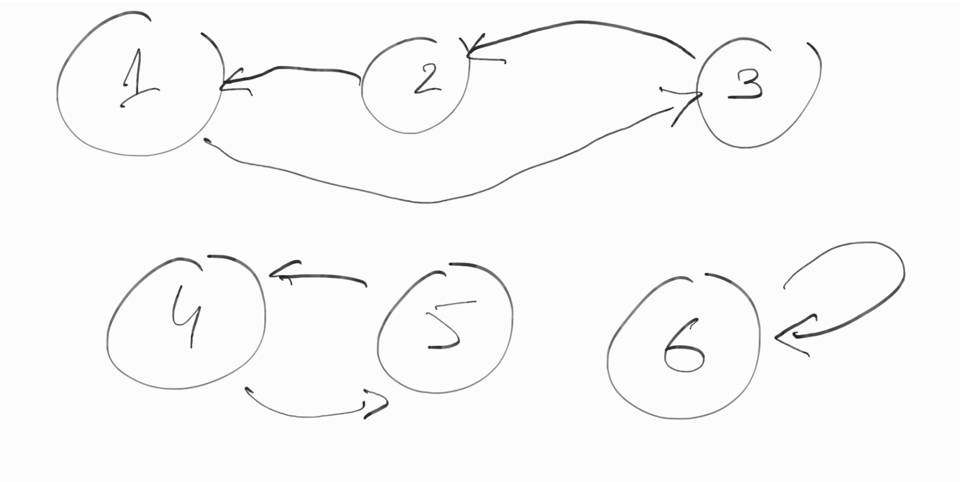
b) Cycles

Définition XXIII.1. Soit $\sigma \in \mathfrak{S}_n$ et soit $a \in [\![1, n]\!]$; on appelle **orbite de a selon σ** l'ensemble :

$$\mathscr{O}_\sigma(a) = \{\sigma^k(a) \,|\, k \in \mathbb{Z}\}$$

où σ^k désigne la composée $\underbrace{\sigma \circ \ldots \circ \sigma}_{k \text{ fois}}$.

Exemple XXIII.3. L'orbite d'un point est donc l'ensemble des points atteignables en partant de ce dernier *via* itérations de la permutation σ. Nous représentons ci–dessous les différentes orbites de la permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 5 & 4 & 6 \end{pmatrix}$.



Remarque XXIII.1. Les orbites disjointes selon une permutation $\sigma \in \mathfrak{S}_n$ partitionnent l'ensemble $[\![1, n]\!]$.

Notation. La remarque *supra* nous autorise à envisager de noter une permutation comme la liste de ses orbites. Pour calculer l'image d'un point, il nous "suffira" alors de "retracer son chemin" parmi ces dernières. Par exemple, la permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 5 & 4 & 6 \end{pmatrix}$ se notera :

$$\sigma = (1\ 3\ 2)(4\ 5)(6)\,.$$

Il est aussi conventionnel d'éliminer les orbites triviales de cette notation pour gagner en place et en efficacité. Dans le cas présent, nous noterions :

$$\sigma = (1\ 3\ 2)(4\ 5)\,.$$

Définition XXIII.2. On appelle **cycle** toute permutation admettant une unique orbite non triviale. Cette orbite est alors appelée **support** du cycle et son cardinal **longueur** de celui–ci.

Exemple XXIII.4. Le cycle $(1\ 3\ 4) \in \mathfrak{S}_5$ est en fait la permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 4 & 1 & 5 \end{pmatrix}.$$

Remarque XXIII.2. Si σ est un cycle de longueur k dans $\mathfrak{S}_n$ et que a est dans le support de σ, on a :

$$\sigma = (a\ \sigma(a)\ \sigma^2(a)\ \dots\ \sigma^{k-1}(a)).$$

Proposition XXIII.1. Deux cycles de supports disjoints commutent.

Démonstration. Ces deux cycles agissent (non trivialement) sur des ensembles disjoints : l'ordre de composition ne revêt donc aucune importance. □

Théorème XXIII.2.
Toute permutation différente de l'identité peut s'écrire de façon unique (à l'ordre près des facteurs) comme produit de cycles de supports disjoints.

Démonstration. Ce résultat est admis. □

Exemple XXIII.5. Pour déterminer la décomposition d'une permutation, il suffit de recopier ses orbites "dans l'ordre". Par exemple :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 4 & 3 \end{pmatrix} = (1\ 2)(3\ 5).$$

c) Signature

Définition XXIII.3. On appelle **transposition** tout cycle de longueur 2.

Exemple XXIII.6. Les permutations sont donc les $(i\ j) \in \mathfrak{S}_n$ avec $i \neq j$.

Proposition XXIII.3. Toute permutation de $\mathfrak{S}_n$ peut s'écrire sous la forme d'un produit de transpositions.

Démonstration. Faisons le par récurrence sur $n \geq 1$.

— Pour $n = 1$, c'est trivial : tout élément de $\mathfrak{S}_1$ est produit d'exactement 0 transposition(s). Oui, nous en sommes toujours la après 341 pages de cours...

— Supposons la propriété vérifiée au rang n et soit $\sigma \in \mathfrak{S}_n$; alors on a :

$$\sigma = (n+1\ \sigma(n+1))\sigma'$$

avec $\sigma' \in \mathfrak{S}_n$. Ne reste alors qu'à appliquer l'hypothèse de récurrence.

□

✖ **Attention :** il n'y a pas unicité. En effet, $(1\ 2\ 3) = (1\ 2)(2\ 3) = (2\ 3)(3\ 1)$.

➠ **Exemple XXIII.7.** Si $a_1, \dots a_k \in [\![1, n]\!]$, on a $(a_1\ a_2\ \dots\ a_k) = (a_1\ a_2)(a_2\ a_3)\dots(a_{k-1}\ a_k)$.

✖ **Attention :** si les orbites se lisent de gauche à droite, la composée, elle, se lit toujours de droite à gauche.

Théorème XXIII.4.
Il existe une unique application $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$ telle que :

(i) pour toute transposition τ, $\varepsilon(\tau) = -1$;

(ii) ε est un morphisme de groupes de $(\mathfrak{S}_n, \circ)$ vers $(\{-1, 1\}, \times)$.

Ce morphisme est appelé **signature** sur $\mathfrak{S}_n$.

Démonstration. Admis. □

➠ **Exemple XXIII.8.** Pour calculer la signature de $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 4 & 3 \end{pmatrix}$, on la décompose sous forme de produit de transpositions :

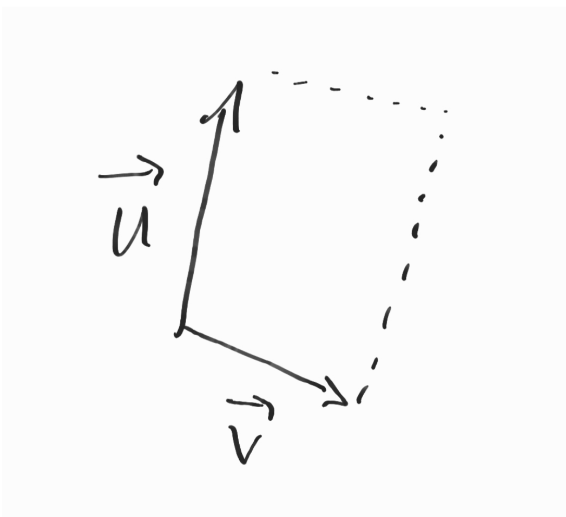
$$\begin{aligned}\varepsilon(\sigma) &= \varepsilon((1\ 2)(3\ 5)) \\ &= \varepsilon((1\ 2))\varepsilon((3\ 5)) \\ &= (-1)^2 \\ &= 1\,.\end{aligned}$$

✌ **Remarque XXIII.3.** Le noyau de ε est appelé **groupe alterné** et noté $\mathfrak{A}_n$. Il s'agit d'un sous–groupe de $\mathfrak{S}_n$ de cardinal $\dfrac{n!}{2}$ si $n \geq 2$.

2.⁻ Formes multilinéaires alternées

a) Laïus géométrique et définition

Étant donné deux vecteurs $\vec{u}$ et $\vec{v}$ dans $\mathbb{R}^2$, nous pouvons considérer l'aire **algébrique** $\mathscr{A}(\vec{u}, \vec{v})$ du parallélogramme engendré par ceux–ci.



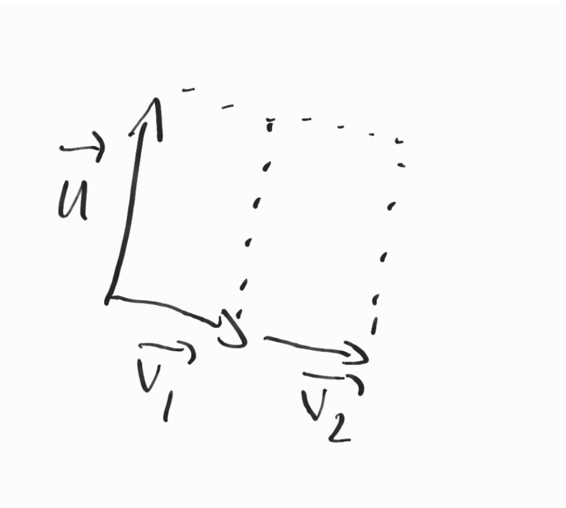
Si $\vec{u} = (a, b)$ et $\vec{v} = (c, d)$, nous avons la formule $\mathscr{A}(\vec{u}, \vec{v}) = ad - bc$. Les propriétés suivantes sont de plus vérifiées :

(i) $\mathscr{A}(\vec{v}, \vec{u}) = -\mathscr{A}(\vec{u}, \vec{v})$;

(ii) pour tout $\lambda \in \mathbb{R}$, $\mathscr{A}(\vec{u}, \lambda\vec{u}) = 0$;

(iii) pour tous vecteurs $\vec{v_1}$ et $\vec{v_2}$, on a

$$\mathscr{A}(\vec{u}, \vec{v_1} + \vec{v_2}) = \mathscr{A}(\vec{u}, \vec{v_1}) + \mathscr{A}(\vec{u}, \vec{v_2}) .$$



Nous pouvons développer la même théorie dans $\mathbb{R}^3$ via la notion de volume algébrique engendré par trois vecteurs. La définition suivante a pour but de proposer une généralisation de ces objets en dimension (éventuellement infinie) quelconque.

Définition XXIII.4. Soient $E_1, \ldots, E_n$ des $\mathbb{K}$–e.v ; on appelle **forme n–linéaire** sur ces espaces toute application $f : E_1 \times \ldots \times E_n \to \mathbb{K}$ telle que pour tout $i \in [\![1, \ldots, n]\!]$, pour tout $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in E_1 \times \ldots \times E_{i-1} \times E_{i+1} \times \ldots \times E_n$, l'application

$$f_i : E_i \to \mathbb{K}$$
$$u \mapsto f(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n)$$

est une forme linéaire sur E_i.

Notation. On note $\mathscr{L}(E_1, \ldots, E_n; \mathbb{K})$ l'ensemble des formes n–linéaires sur $E_1 \times \ldots \times E_n$. Si tous les E_i sont égaux à un même $\mathbb{K}$–e.v E, on notera cet ensemble $\mathscr{L}_n(E, \mathbb{K})$.

✌ **Remarque XXIII.4.** Une forme n–linéaire est donc linéaire selon chacune des coordonnées (au sens du produit cartésien) des vecteurs de son ensemble de départ.

➠ **Exemple XXIII.9.**

— $(x, y) \mapsto xy \in \mathscr{L}_2(\mathbb{K}, \mathbb{K})$;
— le produit scalaire (au sens vu en physique) est une forme 2–linéaire (on dit **bilinéaire**) sur $\mathbb{R}^2$, $\mathbb{R}^3$;
— l'aire et le volume algébriques évoqués dans le laïus géométrique *supra* sont des formes bilinéaire et trilinéaire respectivement.

Définition XXIII.5. Soit E un $\mathbb{K}$–e.v et soit $f \in \mathscr{L}_n(E, \mathbb{K})$. On dit que f est :

— **symétrique** si pour tout $(x_1, \ldots, x_n) \in E^n$ et tout couple $(i, j) \in [\![1, n]\!]$ tel que $i \neq j$ on a :

$$f(x_1, \ldots, \underbrace{x_i}_{i}, \ldots, \underbrace{x_j}_{j}, \ldots, x_n) = f(x_1, \ldots, \underbrace{x_j}_{i}, \ldots, \underbrace{x_i}_{j}, \ldots, x_n)\,;$$

— **antisymétrique** si pour tout $(x_1, \ldots, x_n) \in E^n$ et tout couple $(i, j) \in [\![1, n]\!]$ tel que $i \neq j$ on a :

$$f(x_1, \ldots, \underbrace{x_i}_{i}, \ldots, \underbrace{x_j}_{j}, \ldots, x_n) = -f(x_1, \ldots, \underbrace{x_j}_{i}, \ldots, \underbrace{x_i}_{j}, \ldots, x_n)\,;$$

— **alternée** si pour tout $(x_1, \ldots, x_n) \in E^n$ et tout couple $(i, j) \in [\![1, n]\!]$ tel que $i \neq j$ on a :

$$(x_i = x_j) \Rightarrow (f(x_1, \ldots, x_n) = 0)\,.$$

➠ **Exemple XXIII.10.**

— $(x, y) \mapsto xy$ est une forme bilinéaire symétrique ;

— l'aire algébrique est une forme bilinéaire antisymétrique alternée.

b) Quelques propriétés élémentaires

Proposition XXIII.5. Soit E un $\mathbb{K}$–e.v et soit $f \in \mathscr{L}_n(E, \mathbb{K})$. Alors :

$$f \text{ est alternée} \\ \Leftrightarrow \\ f \text{ est antisymétrique.}$$

Démonstration.

($\Uparrow$) Supposons f antisymétrique. Alors, pour tout $x_1, \ldots, x_n \in E$ et $i \neq j$ dans $[\![1, n]\!]$ on a :

$$f(x_1, \ldots, \underbrace{x_i}_{i}, \ldots, \underbrace{x_j}_{j}, \ldots, x_n) = -f(x_1, \ldots, \underbrace{x_j}_{i}, \ldots, \underbrace{x_i}_{j}, \ldots, x_n)$$

ce qui entraîne que si $x_i = x_j$

$$f(x_1, \ldots, x_n) = -f(x_1, \ldots, x_n)$$

et donc $f(x_1, \ldots, x_n) = 0$.

($\Downarrow$) Supposons f alternée. Alors, pour tout $x_1, \ldots, x_n \in E$ et $i \neq j$ dans $[\![1, n]\!]$ on a :

$$\begin{aligned}
0 &= f(x_1, \dots, \underbrace{x_i + x_j}_{i}, \dots, \underbrace{x_i + x_j}_{j}, \dots, x_n) \\
&= f(x_1, \dots, \underbrace{x_i}_{i}, \dots, \underbrace{x_i + x_j}_{j}, \dots, x_n) + f(x_1, \dots, \underbrace{x_j}_{i}, \dots, \underbrace{x_i + x_j}_{j}, \dots, x_n) \\
&= f(x_1, \dots, \underbrace{x_i}_{i}, \dots, \underbrace{x_i}_{j}, \dots, x_n) + f(x_1, \dots, \underbrace{x_i}_{i}, \dots, \underbrace{x_j}_{j}, \dots, x_n) \\
&\quad + f(x_1, \dots, \underbrace{x_j}_{i}, \dots, \underbrace{x_i}_{j}, \dots, x_n) + f(x_1, \dots, \underbrace{x_j}_{i}, \dots, \underbrace{x_j}_{j}, \dots, x_n) \\
&= f(x_1, \dots, \underbrace{x_i}_{i}, \dots, \underbrace{x_j}_{j}, \dots, x_n) + f(x_1, \dots, \underbrace{x_j}_{i}, \dots, \underbrace{x_i}_{j}, \dots, x_n)
\end{aligned}$$

par alternance, d'où le résultat.

□

Notation. On notera $\Lambda_n(E)$ l'ensemble des formes n–linéaires alternées/antisymétriques sur E.

✌ **Remarque XXIII.5.** On peut vérifier que $\Lambda_n(E)$ est un $\mathbb{K}$–e.v.

Proposition XXIII.6. Soit E un $\mathbb{K}$–e.v et soit $f \in \Lambda_n(E)$. Alors, pour toute famille **liée** $(x_1, \dots, x_n) \in E^n$ on a $f(x_1, \dots, x_n) = 0$.

Démonstration. Si la famille est liée, cela signifie que, quitte à modifier l'ordre des x_i (ce qui changera éventuellement le signe de $f(x_1, \dots, x_n)$) il existe $\lambda_1, \dots, \lambda_{n-1} \in \mathbb{K}$ tels que

$$x_n = \sum_{k=1}^{n-1} \lambda_k x_k$$

et donc :

$$\begin{aligned}
f(x_1, \dots, x_n) &= f\left(x_1, \dots, x_{n-1}, \sum_{k=1}^{n-1} \lambda_k x_k\right) \\
&= \sum_{k=1}^{n-1} \lambda_k f(x_1, \dots, x_k, , \dots, x_{n-1}, x_k) \\
&= 0\,.
\end{aligned}$$

□

Proposition XXIII.7. Soit E un $\mathbb{K}$–e.v et soit $f \in \Lambda_n(E)$. Alors, pour tout $\sigma \in \mathfrak{S}_n$ et tous $x_1, \dots, x_n \in E$, on a :

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) f(x_1, \dots, x_n)\,.$$

Démonstration. Écrire σ sous forme de produit de transpositions et développer. □

3.⁻ Déterminant

On fixe dans ce paragraphe un $\mathbb{K}$–e.v de dimension finie E de dimension $n \geq 1$ et une base $\mathscr{B} = (e_1, \dots, e_n)$ de celui–ci.

a) Déterminant d'une famille de vecteurs

Théorème XXIII.8.
Il existe une unique forme $f \in \Lambda_n(E)$ telle que $f(\mathscr{B}) = 1$.

Démonstration. La démonstration de l'existence n'est pas exigible. Pour l'unicité, fixons $f \in \Lambda_n(E)$ et $x_1, \dots, x_n \in E$. Alors, il existe une unique famille $(x_{1,1}, \dots, x_{n,1}, x_{1,2} \dots, x_{n,n})$ de scalaires telle que pour tout $j \in [\![1, n]\!]$ on ait

$$x_j = \sum_{i=1}^{n} x_{i,j} e_i \,.$$

et alors :

$$f(x_1, \dots, x_n) = f\left(\sum_{i=1}^{n} x_{i,1} e_i, \dots, \sum_{i=1}^{n} x_{i,n} e_i\right) \,.$$

Pour obtenir une intuition du résultat de cette horreur, plaçons nous temporairement dans le cas $n = 2$:

$$\begin{aligned} f(x_1, x_2) &= f(x_{1,1}e_1 + x_{2,1}e_2, x_{1,2}e_1 + x_{2,2}e_2) \\ &= x_{1,1} f(e_1, x_{1,2}e_1 + x_{2,2}e_2) + x_{2,1} f(e_2, x_{1,2}e_1 + x_{2,2}e_2) \\ &= x_{1,1}x_{1,2}(fe_1, e_1) + x_{1,1}x_{2,2} f(e_1, e2) + x_{2,1}x_{1,2} f(e_2, e_1) + x_{2,1}x_{2,2} f(e_2, e_2) \\ &= (x_{1,1}x_{1,2} - x_{2,1}x_{1,2}) f(e_1, e_2) \,. \end{aligned}$$

En développant l'expression multilinéaire de $f(x_1, \dots, x_n))$, nous "intuitons" que tous les termes contenant l'image d'une famille liée seront annulés : ceci entraîne que chaque terme sera de la forme :

$$\prod_{i=1}^{n} x_{\sigma(i),i} f(e_{\sigma(1)}, \dots, e_{\sigma(n)})$$

pour un certain $\sigma \in \mathfrak{S}_n$. Nous démontrons donc, par récurrence sur $n \in \mathbb{N}^*$, que :

$$\begin{aligned} f(x_1, \dots, x_n) &= \sum_{\sigma \in \mathfrak{S}_n} \prod_{i=1}^{n} x_{\sigma(i),i} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\ &= \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} x_{\sigma(i),i}\right) f(e_1, \dots, e_n) \end{aligned}$$

la dernière égalité provenant de la proposition XXIII.7. Ceci entraîne le résultat voulu. □

Définition XXIII.6. L'unique forme $f \in \Lambda_n(E)$ telle que $f(\mathscr{B}) = 1$ est appelée **déterminant dans la base $\mathscr{B}$.**

Notation. $\det_{\mathscr{B}}$

Remarque XXIII.6. On a donc, pour tout $x_1, \dots, x_n \in E$:

$$\det_{\mathscr{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} x_{\sigma(i),i}$$

avec les décompositions initales

$$x_j = \sum_{i=1}^{n} x_{i,j} e_i$$

pour $j \in [\![1, n]\!]$. De plus, il est immédiat que :

$$\det_{\mathscr{B}}(e_1, \dots, e_n) = 1$$

par définition.

Exemple XXIII.11. Si $x = (x_1, x_2)$ et $y = (y_1, y_2)$ sont deux vecteurs de $\mathbb{R}^2$ alors leur déterminant dans la bases canonique est :

$$\det(x, y) = x_1 y_2 - x_2 y_1 ,$$

ce qui correspond à la quantité $\mathscr{A}(x, y)$ décrite au début de ce chapitre.

Exercice XXIII.1. Donner une expression du déterminant dans la base canonique de trois vecteurs de $\mathbb{R}^3$. Interprétation géométrique ?

Remarque XXIII.7. La déterminant est une somme de $n!$ produits, dont exactement la moitié sont porteurs d'une signature négative.

Proposition XXIII.9. Soit $f \in \Lambda_n(E)$. Alors il existe un unique $\lambda \in \mathbb{K}$ tel que $f = \lambda \det_{\mathscr{B}}$.

Démonstration. Cela découle du calcul mené dans la démonstration du théorème XXIII.8. Le scalaire λ est en fait égal à $f(e_1, \dots, e_n)$. □

Remarque XXIII.8. Cela signifie que l'espace vectoriel $\Lambda_n(E)$ est de dimension 1 : il s'agit d'une droite.

Proposition XXIII.10. Soit $\mathscr{F} = (x_1, \dots, x_n)$ une famille de vecteurs de E. Alors :

$$\begin{gathered} \mathscr{F} \text{ est une base de } E \\ \Leftrightarrow \\ \det_{\mathscr{B}}(\mathscr{F}) \neq 0. \end{gathered}$$

Démonstration.

($\Downarrow$) Supposons que $\mathscr{F}$ est une base de E. Dans ce cas, on a $\det_{\mathscr{B}}, \det_{\mathscr{F}} \in \Lambda_n(E)$. Ce dernier ensemble étant une droite, il existe $\lambda \in \mathbb{K}$ tel que $\det_{\mathscr{B}} = \lambda \det_{\mathscr{F}}$. Ainsi :

$$\det_{\mathscr{B}}(\mathscr{F}) = \lambda \det_{\mathscr{F}}(\mathscr{F}) = \lambda\,.$$

De plus, $\lambda \neq 0$ car $\det_{\mathscr{B}}(\mathscr{B}) = 1 = \lambda \det_{\mathscr{F}}(\mathscr{B})$.

($\Uparrow$) Supposons désormais que la famille $\mathscr{F}$ ne soit pas une base de E. Comme $\operatorname{card}(\mathscr{F}) = \dim(E)$, elle ne peut donc pas être libre et donc $\det_{\mathscr{B}}(\mathscr{F}) = 0$ par la proposition XXIII.6.

□

✌ **Remarque XXIII.9.** En "bonus", nous tirons de la démonstration *supra* que pour toute base $\mathscr{B}'$ de E on a :

$$\det_{\mathscr{B}} = \det_{\mathscr{B}}(\mathscr{B}') \det_{\mathscr{B}'}\,.$$

b) Déterminant d'une matrice carrée

Définition XXIII.7. Soit $M \in \mathscr{M}_n(\mathbb{K})$. On appelle **déterminant** de M le déterminant de la famille de ses colonnes dans la base canonique de $\mathbb{K}^n$.

Notation. Le déterminant d'une matrice $M = (m_{i,j})_{i,j}$ sera noté $\det(M)$ ou

$$\begin{vmatrix} m_{1,1} & \dots & m_{1,n} \\ \vdots & \ddots & \vdots \\ m_{n,1} & \dots & m_{n,n} \end{vmatrix}\,.$$

✌ **Remarque XXIII.10.** Cette définition entraîne que si $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$ on a :

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}\,.$$

➠ **Exemple XXIII.12.**
- — $\det(I_n) = 1$;
- — si $a, b, c, d \in \mathbb{K}$, on a :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc\,.$$

✍ **Exercice XXIII.2.** Déterminer une formule explicite pour le déterminant sur $\mathscr{M}_3(\mathbb{K})$.

Proposition XXIII.11. Soit $A, B \in \mathscr{M}_n(\mathbb{K})$ et soit $\lambda \in \mathbb{K}$. Alors :
- (i) $\det({}^{\mathrm{t}}A) = \det(A)$;
- (ii) $A \in GL_n(\mathbb{K}) \Leftrightarrow \det(A) \neq 0$;
- (iii) $\det(AB) = \det(A)\det(B)$;
- (iv) $\det(\lambda A) = \lambda^n \det(A)$.

Démonstration.

(i)

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(i),i}\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(i),\sigma^{-1}(\sigma(i))}\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}\end{aligned}$$

via le changement d'indice $j = \sigma^{-1}(i)$ dans les produits. Enfin, $\mathfrak{S}_n$ étant un groupe, on a :

$$\{\sigma^{-1} \mid \sigma \in \mathfrak{S}_n\} = \mathfrak{S}_n$$

et donc :

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} a_{j,\sigma(j)}\\ &= \det({}^{\mathrm{t}}A)\,.\end{aligned}$$

(ii) Une matrice est inversible si et seulement si la famille de ses colonnes forme une base de $\mathbb{K}^n$. Ce fait est équivalent à la non–nullité du déterminant de cette famille dans la base mentionnée, d'où le résultat.

(iii) L'application

$$\begin{aligned}\phi : \mathscr{M}_n(\mathbb{K}) &\longrightarrow \mathbb{K}\\ M &\mapsto \det(MB)\ ,\end{aligned}$$

vue comme fonction des colonnes de M, est une forme n–linéaire alternée. De fait, elle est multiple du déterminant dans la base canonique, *i.e* il existe $\lambda \in \mathbb{K}$ tel que :

$$\forall M \in \mathscr{M}_n(\mathbb{K}),\ \det(MB) = \lambda \det(M)$$

En évaluant en $M = I_n$, on trouve que $\lambda = \det(B)$, d'où le résultat.

(iv) L'application det, vue comme fonction des colonnes de la matrice concernée, est n–linéaire.

□

Corollaire XXIII.11.a. Soit $A \in GL_n(\mathbb{K})$. Alors $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

c) Déterminant d'un endomorphisme

Proposition/définition XXIII.8. Soit $f \in \mathscr{L}(E)$. Alors la quantité $\det(\mathrm{mat}_{\mathscr{B}}(f))$ ne dépend pas du choix de la base $\mathscr{B}$ et est appelé **déterminant** de l'endomorphisme f.

Notation. $\det(f)$

Démonstration. Soit $\mathscr{B}'$ une base de E. Alors, en posant $P = P_{\mathscr{B}}^{\mathscr{B}'}$ on a

$$\mathrm{mat}_{\mathscr{B}'}(f) = P^{-1}\mathrm{mat}_{\mathscr{B}}(f)P$$

ergo

$$\begin{aligned}\det(\mathrm{mat}_{\mathscr{B}'}(f)) &= \det(P^{-1}\mathrm{mat}_{\mathscr{B}}(f)P)\\ &= \det(P^{-1})\det(\mathrm{mat}_{\mathscr{B}}(f))\det(P)\\ &= \frac{1}{\det(P)}\det(\mathrm{mat}_{\mathscr{B}}(f))\det(P)\\ &= \det(\mathrm{mat}_{\mathscr{B}}(f))\,.\end{aligned}$$

□

Proposition XXIII.12. Soient $f, g \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$. Alors :

(i) $\det(g \circ f) = \det(g)\det(f)$;

(ii) $\det(\lambda f) = \lambda^n \det(f)$;

(iii) $f \in GL(E) \Leftrightarrow \det(f) \neq 0$.

Démonstration. Il s'agit d'un analogue de la proposition XXIII.11. □

d) Orientation d'un $\mathbb{R}$–espace vectoriel

On se place dans ce paragraphe sur $\mathbb{K} = \mathbb{R}$.

Nous nous intéressons dans ce paragraphe à l'ensemble $\mathscr{E}$ des bases de l'espace E (dont nous rappelons que la dimension est strictement positive). Étant donné $\mathscr{B}, \mathscr{B}' \in \mathscr{E}$, nous noterons $\mathscr{B} \sim \mathscr{B}'$ si et seulement si $\det(P_{\mathscr{B}}^{\mathscr{B}'}) > 0$.

Proposition XXIII.13. La relation $\sim$ est une relation d'équivalence sur $\mathscr{E}$.

Démonstration. Laissée aux bons soins du lecteur. La transitivité découle du fait que pour $\mathscr{B}, \mathscr{B}'$ et $\mathscr{B}''$ trois bases de $\mathscr{E}$ on a :

$$P_{\mathscr{B}}^{\mathscr{B}''} = P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}''}$$

(pourquoi ?). □

Face à cette relation d'équivalence, une question survient : à combien de classes donne–t–elle naissance ? Pour y répondre, fixons $\mathscr{B}, \mathscr{B}'$ et $\mathscr{B}''$ trois bases de $\mathscr{E}$ et remarquons que si $\mathscr{B} \not\sim \mathscr{B}''$ alors :

$$\begin{aligned}\det(P_{\mathscr{B}}^{\mathscr{B}''}) &= \det(P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}''}) \\ &= \det(P_{\mathscr{B}}^{\mathscr{B}'})\det(P_{\mathscr{B}'}^{\mathscr{B}''})\end{aligned}$$

et donc, comme $\det(P_{\mathscr{B}}^{\mathscr{B}''}) < 0$ on obtient que $\det(P_{\mathscr{B}}^{\mathscr{B}'})$ et $\det(P_{\mathscr{B}'}^{\mathscr{B}''})$ sont de signes opposés. Ainsi, soit $\mathscr{B} \sim \mathscr{B}'$ soit $\mathscr{B}' \sim \mathscr{B}''$ et ce de façon exclusive. On en déduit qu'il existe **exactement** deux classes d'équivalence pour la relation $\sim$.

Définition XXIII.9. Orienter l'espace E consiste à choisir (arbitrairement) l'une des classes d'équivalences pour la relation $\sim$. Les éléments de celle–ci sont alors appelées **bases directes** de E.

Traditionnellement, lorsque $E = \mathbb{R}^n$, on a tendance à choisir la classe contenant la base canonique. Les bases directes obtenues sont alors celles utilisées en sciences physiques ou de l'ingénieur.

4.⁻ Calculs

Savoir définir un déterminant, c'est bien. Savoir le calculer, c'est. . . mieux ? Nous laissons le lecteur juge de ce fait.

a) Matrices triangulaires

Proposition XXIII.14. Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$ une matrice **triangulaire**. Alors :

$$\det(A) = \prod_{i=1}^{n} a_{i,i}\,.$$

Démonstration. Supposons, pour fixer les idées, que $A \in T_n^+(\mathbb{K})$. Soit $\sigma \in \mathfrak{S}_n$ différente de l'identité. Alors, il existe nécessairement $i \in [\![1, n]\!]$ tel que $\sigma(i) > i$ et donc $a_{\sigma(i),i} = 0$, ce qui entraîne que :

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \underbrace{\prod_{i=1}^{n} a_{\sigma(i),i}}_{=0 \text{ si } \sigma\neq\mathrm{id}} \\ &= \prod_{i=1}^{n} a_{i,i}\,.\end{aligned}$$

□

Exemple XXIII.13.

— pour tous $d_1, \dots, d_n \in \mathbb{K}$, $\det(\mathrm{diag}(d_1, \dots, d_n)) = d_1 \dots d_n$;

—

$$\begin{vmatrix} 1 & 2 & 51 & 22 & -\pi \\ 0 & 2 & 3 & 6 & 0 \\ 0 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & -1 & 42 \\ 0 & 0 & 0 & 0 & 8 \end{vmatrix} = 16\,.$$

✌ **Remarque XXIII.11.** Ceci se généralise aux matrices triangulaires par blocs **carrés** : si A, B, C sont des matrices alors

$$\det\left(\left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)\right) = \det(A)\det(C)$$

et même chose avec les matrices triangulaires par blocs admettant un plus grand nombre de blocs.

b) Impact des opérations élémentaires

Nous avions listé dans le chapitre XX une liste d'opérations élémentaires sur les familles de vecteurs. Appliquer celles–ci aux **colonnes** d'un déterminant modifie celui–ci de la façon suivante :

(A) échanger la position de deux colonnes ($C_i \leftrightarrow C_j$) multiplie le déterminant par -1 ;

(B) multiplier une colonne par un scalaire **non nul** ($C_i \leftarrow \lambda C_i$) multiplie le déterminant par ce dernier ;

(C) ajouter à une colonne une combinaison linéaire des **autres** $\left(C_i \leftarrow C_i + \sum\limits_{j \neq i} \lambda_j C_j\right)$ ne modifie pas le déterminant.

Il est également possible d'agir sur les **lignes** d'un déterminant de la même façon et avec les mêmes effets.

✍ **Exercice XXIII.3.** Soit $x \in \mathbb{C}$. Calculer le déterminant suivant :

$$D_n(x) = \begin{vmatrix} x & 1 & \dots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & x \end{vmatrix}\,.$$

➥ **Correction :** *Commençons par effectuer l'opération $C_1 \leftarrow C_1 + \ldots + C_n$; on* *obtient*

$$D_n(x) = \begin{vmatrix} x+n-1 & 1 & \dots & \dots & 1 \\ x+n-1 & x & \ddots & & \vdots \\ \vdots & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ x+n-1 & 1 & \dots & 1 & x \end{vmatrix} = (x+n-1) \begin{vmatrix} 1 & 1 & \dots & \dots & 1 \\ 1 & x & \ddots & & \vdots \\ \vdots & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 1 & 1 & \dots & 1 & x \end{vmatrix}.$$

Nous appliquons ensuite, pour tout $i \geq 2$, $L_i \leftarrow L_i - L_1$, ce qui donne :

$$D_n(x) = (x+n-1) \begin{vmatrix} 1 & 1 & \dots & \dots & 1 \\ 0 & x-1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & x-1 \end{vmatrix} = (x+n-1)(x-1)^{n-1}.$$

c) Développement suivant une ligne ou une colonne

Définition XXIII.10. Soit $A \in \mathscr{M}_n(\mathbb{K})$ et soient $i, j \in [\![1, n]\!]$. On appelle :
- **mineur** d'indices i et j la matrice $\mathrm{Min}_{i,j}(A)$ extraite de A suivant l'ensemble $([\![1, n]\!] \setminus \{i\}) \times ([\![1, n]\!] \setminus \{j\})$;
- **cofacteur** d'indices i et j la quantité $\mathrm{Cof}_{i,j}(A) = (-1)^{i+j}\det(\mathrm{Min}_{i,j}(A))$.

✌ **Remarque XXIII.12.** Le mineur d'indices i et j est donc la matrice obtenue en "rayant" la ligne i et la colonne j de la matrice A.

➠ **Exemple XXIII.14.** Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathscr{M}_2(\mathbb{K})$. Ses mineurs et cofacteurs sont donnés par le tableau récapitulatif *infra*.

i	j	$\mathrm{Min}_{i,j}(A)$	$\mathrm{Cof}_{i,j}(A)$
1	1	(d)	d
1	2	(c)	-c
2	1	(b)	-b
2	2	(a)	a

✍ **Exercice XXIII.4.** Lister les mineurs et les cofacteurs de la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

◇ **Petit calcul sympathique**

Soit $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$ et soient $A_1, \ldots, A_n \in \mathbb{K}^n$ ses colonnes (on aura donc $A = (A_1 \,|\, \ldots \,|\, A_n)$). On notera également $\mathscr{B} = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{K}^n$, de façon à avoir, pour tout $j \in [\![1, n]\!]$

$$A_j = \sum_{i=1}^{n} a_{i,j} e_i \, .$$

De fait, à $j \in [\![1, n]\!]$ fixé, on a :

$$\begin{aligned}
\det(A) &= \det_{\mathscr{B}}(A_1, \ldots, A_{j-1}, A_j, A_{j+1}, \ldots, A_n) \\
&= \det_{\mathscr{B}}\left(A_1, \ldots, A_{j-1}, \sum_{i=1}^{n} a_{i,j} e_i, A_{j+1}, \ldots, A_n\right) \\
&= \sum_{i=1}^{n} a_{i,j} \det_{\mathscr{B}}\left(A_1, \ldots, A_{j-1}, e_i, A_{j+1}, \ldots, A_n\right)
\end{aligned}$$

par n–linéarité du déterminant dans la base $\mathscr{B}$. Dit autrement, cela signifie que nous avons

$$\begin{vmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \ldots & a_{n,n} \end{vmatrix} = \sum_{i=1}^{n} a_{i,j} D_{i,j}$$

avec

$$D_{i,j} = \begin{vmatrix}
a_{1,1} & \ldots & a_{1,j-1} & 0 & a_{1,j+1} & \ldots & a_{1,n} \\
\vdots & \ddots & \vdots & \vdots & & & \vdots \\
\vdots & & a_{i-1,j-1} & 0 & & & \vdots \\
a_{i,1} & \ldots & a_{i,j-1} & 1 & a_{i,j+1} & \ldots & a_{i,n} \\
\vdots & & \vdots & 0 & \vdots & \ddots & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & a_{n-1,n} \\
a_{n,1} & \ldots & a_{n,j-1} & 0 & a_{n,j+1} & \ldots & a_{n,n}
\end{vmatrix}$$

la colonne "modifiée" étant en j–ième position.

Tout ceci est bien sympathique, mais sommes nous seulement capables de calculer les $D_{i,j}$? Bien entendu, mais rien n'est gratuit… Calculons (avec entrain) ! Quitte à effecteurs les opérations $C_j \leftrightarrow C_{j+1}$, …, $C_{n-1} \leftrightarrow C_n$ (dans cet ordre), on a :

$$D_{i,j} = (-1)^{n-j} \begin{vmatrix}
a_{1,1} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1,n} & 0 \\
\vdots & \ddots & \vdots & & & \vdots & \vdots \\
\vdots & & a_{i-1,j-1} & & & \vdots & 0 \\
a_{i,1} & \ldots & a_{i,j-1} & a_{i,j+1} & \ldots & a_{i,n} & 1 \\
\vdots & & \vdots & \vdots & \ddots & \vdots & 0 \\
\vdots & & \vdots & \vdots & & a_{n-1,n} & \vdots \\
a_{n,1} & \ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{n,n} & 0
\end{vmatrix} .$$

Nous opérons ensuite les échanges (ordonnés) suivants : $L_i \leftrightarrow L_{i+1}$, $L_{i+1} \leftrightarrow L_{i+2}$,

$\dots$, $L_{n-1} \leftrightarrow L_n$. On a *in fine* :

$$D_{i,j} = \underbrace{(-1)^{2n-(i+j)}}_{} = (-1)^{i+j} \begin{vmatrix} a_{1,1} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1,n} & 0 \\ \vdots & \ddots & \vdots & & & \vdots & \vdots \\ \vdots & & a_{i-1,j-1} & & & \vdots & 0 \\ \vdots & & \vdots & a_{j+1,j+1} & & \vdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & & \vdots \\ a_{n,1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} & 0 \\ a_{i,1} & \dots & a_{i,j-1} & a_{i,j+1} & \dots & a_{i,n} & 1 \end{vmatrix} .$$

Ce déterminant étant triangulaire par blocs, on a immédiatement :

$$D_{i,j} = (-1)^{i+j} \det(\mathrm{Min}_{i,j}(A)) = \mathrm{Cof}_{i,j}(A) .$$

Nous déduisons de cette plongée dans le stupre calculatoire, la proposition suivante, fort utile en pratique.

Proposition XXIII.15. Soit $A \in \mathscr{M}_n(\mathbb{K})$ et soit $j \in [\![1, n]\!]$. Alors :

$$\det(A) = \sum_{i=1}^{n} a_{i,j} \mathrm{Cof}_{i,j}(A) .$$

✌ **Remarque XXIII.13.** Ce procédé, appelé **développement selon une colonne** de A admet un pendant relatif aux lignes de celles ci, dont la démonstration est analogue :

$$\forall i \in [\![1, n]\!],\ \det(A) = \sum_{j=1}^{n} a_{i,j} \mathrm{Cof}_{i,j}(A) .$$

➠ **Exemple XXIII.15.** On retrouve aisément de cette façon la formule, pour $a, b, c, d \in \mathbb{K}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$. Le lecteur audacieux pourra également s'en servir pour obtenir une formule explicite pour le déterminant dans $\mathscr{M}_3(\mathbb{K})$.

✍ **Exercice XXIII.5.** Soit $x \in \mathbb{C}$ et $n \geq 0$. Calculer le déterminant :

$$D_n = \begin{vmatrix} 1+x^2 & x & 0 & \dots & \dots & 0 \\ x & 1+x^2 & x & 0 & \dots & 0 \\ 0 & x & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & x \\ 0 & \dots & \dots & 0 & x & 1+x^2 \end{vmatrix} .$$

➥ **Correction :** *Soit $n \geq 2$; on commence par opérer un développement par rapport à la première colonne, qui donne :*

$$D_n = (1 + x^2) D_{n-1} - x \Delta_{n-1}$$

où

$$\Delta_{n-1} = \begin{vmatrix} x & 0 & 0 & \dots & \dots & 0 \\ x & 1+x^2 & x & 0 & \dots & 0 \\ 0 & x & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & x \\ 0 & \dots & \dots & 0 & x & 1+x^2 \end{vmatrix}.$$

En développant Δ_{n-1} par rapport à sa première ligne, on trouve :

$$\Delta_{n-1} = xD_{n-2}$$

ce qui permet in fine d'écrire :

$$D_n = (1+x^2)D_{n-1} - x^2 D_{n-2}\,.$$

On reconnaît une relation de récurrence linéaire d'ordre 2 (cf. chapitre IX). Plus précisément, on a :

$$\begin{aligned} D_n - D_{n-1} &= x^2(D_{n-1} - D_{n-2}) \\ &\ \vdots \\ &= x^{2(n-1)}(D_1 - D_0) \end{aligned}$$

avec $D_0 = 1$ et $D_1 = 1 + x^2$ (en prenant la convention qu'un produit sans terme est égal à 1). Nous terminons donc avec :

$$D_n = D_{n-1} + x^{2n}$$

i.e

$$\begin{aligned} D_n &= \sum_{k=0}^{n} x^{2k} \\ &= \begin{cases} n+1 \text{ si } x = 1 \\ \dfrac{1-x^{2(n+1)}}{1-x^2} \text{ sinon.} \end{cases} \end{aligned}$$

◇ **Déterminant de Vandermonde**

Ce qui suit est un calcul de déterminant classique et exigible. Le lecteur est fortement encouragé à s'y confronter **réellement**.

Soit $n \geq 2$ et soient $a_0, \dots, a_{n-1} \in \mathbb{K}$. On appelle **déterminant de Vandermonde** (en l'honneur d'Alexandre–Théophile Vandermonde, mathématicien, économiste, musicien et chimiste français, 1735—1796) d'ordre n associé à ces scalaires le déterminant :

$$\begin{aligned} V_n(a_0, \dots, a_{n-1}) &= \det((a_i^j)_{i,j \in [\![0,n-1]\!]}) \\ &= \begin{vmatrix} 1 & a_0 & a_0^2 & \dots & a_0^{n-1} \\ 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \dots & a_{n-1}^{n-1} \end{vmatrix}. \end{aligned}$$

Pour calculer ce déterminant, nous effectuons, dans l'ordre **décroissant** de l'indice $k \in [\![2, n]\!]$, l'opération $C_k \leftarrow C_k - a_{n-1}C_{k-1}$, ce qui donne :

$$V_n(a_0, \ldots, a_{n-1}) = \begin{vmatrix} 1 & a_0 - a_{n-1} & a_0(a_0 - a_{n-1}) & \ldots & a_0^{n-2}(a_0 - a_{n-1}) \\ 1 & a_1 - a_{n-1} & a_1(a_1 - a_{n-1}) & \ldots & a_1^{n-2}(a_1 - a_{n-1}) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-2} - a_{n-1} & a_{n-2}(a_{n-2} - a_{n-1}) & \ldots & a_{n-2}^{n-2}(a_{n-2} - a_{n-1}) \\ 1 & 0 & 0 & \ldots & 0 \end{vmatrix}.$$

Ce déterminant se développe relativement à sa dernière ligne, livrant :

$$V_n(a_0, \ldots, a_{n-1}) = (-1)^{n+1} \begin{vmatrix} a_0 - a_{n-1} & a_0(a_0 - a_{n-1}) & \ldots & a_0^{n-2}(a_0 - a_{n-1}) \\ a_1 - a_{n-1} & a_1(a_1 - a_{n-1}) & \ldots & a_1^{n-2}(a_1 - a_{n-1}) \\ \vdots & \vdots & & \vdots \\ a_{n-2} - a_{n-1} & a_{n-2}(a_{n-2} - a_{n-1}) & \ldots & a_{n-2}^{n-2}(a_{n-2} - a_{n-1}) \end{vmatrix}$$

puis, nous pouvons factoriser, pour tout $k \in [\![1, n-1]\!]$, $a_{k-1} - a_{n-1}$ dans la ligne k, ce qui donne :

$$\begin{aligned} V_n(a_0, \ldots, a_{n-1}) &= (-1)^{n+1} \prod_{k=0}^{n-2} (a_k - a_{n-1}) \begin{vmatrix} 1 & a_0 & \ldots & a_0^{n-2} \\ 1 & a_1 & \ldots & a_1^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_{n-2} & \ldots & a_{n-2}^{n-2} \end{vmatrix} \\ &= (-1)^{n-1} \prod_{k=0}^{n-2} (a_k - a_{n-1}) V_{n-1}(a_0, \ldots, a_{n-2}) \\ &= \prod_{k=0}^{n-2} (a_{n-1} - a_k) V_{n-1}(a_0, \ldots, a_{n-2}). \end{aligned}$$

On obtient donc, par récurrence que :

$$V_n(a_0, \ldots, a_{n-1}) = \prod_{k=1}^{n-1} \prod_{j=0}^{k-1} (a_k - a_j).$$

Et donc :

$$V_n(a_0, \ldots, a_{n-1}) \neq 0$$
$$\Leftrightarrow$$
les a_i sont deux à deux distincts.

d) Inversion de matrices

Définition XXIII.11. Soit $A \in \mathscr{M}_n(\mathbb{K})$. On appelle **comatrice** de A la matrice de $\mathscr{M}_n(\mathbb{K})$ dont le coefficient en position $(i, j) \in [\![1, n]\!]^2$ est le cofacteur $\mathrm{Cof}_{i,j}(A)$.

Notation. $\mathrm{Com}(A)$

Vocabulaire. La transposée de la comatrice de A est appelée **transcomatrice** de A. On la note ${}^t\mathrm{Com}(A)$ ou $\tilde{A}$.

➠ **Exemple XXIII.16.** Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathscr{M}_2(\mathbb{K})$. Alors :

$$ {}^t\mathrm{Com}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. $$

Proposition XXIII.16. Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors :

$$ A\,{}^t\mathrm{Com}(A) = {}^t\mathrm{Com}(A)A = \det(A)I_n\,. $$

Démonstration. Notons $a_{i,j}$ les coefficients de A et $b_{i,j}$ ceux de sa transcomatrice. Alors, pour $i \in [\![1, n]\!]$, le coefficient en position (i, i) de la matrice $A\,{}^t\mathrm{Com}(A)$ est :

$$ \begin{aligned} \sum_{k=1}^n a_{i,k} b_{k,i} &= \sum_{k=1}^n a_{i,k} \mathrm{Cof}_{i,k}(A) \\ &= \det(A) \end{aligned} $$

via développement selon la ligne i. Si on fixe à présent $j \neq i$, le coefficient en position (i, j) de $A\,{}^t\mathrm{Com}(A)$ est :

$$ \sum_{k=1}^n a_{i,k} b_{k,j} = \sum_{k=1}^n a_{i,k} \mathrm{Cof}_{j,k}(A) $$

Il s'agit du déterminant de la matrice B définie comme suit : les lignes de B sont identigues à celles de A, à l'exception de la ligne j que l'on fixe égale à la ligne i de A. Cette matrice ayant deux lignes identiques, elle ne peut être inversible. De fait :

$$ \sum_{k=1}^n a_{i,k} \mathrm{Cof}_{j,k}(A) = 0 $$

d'où le résultat, quitte à raisonner similairement pour ${}^t\mathrm{Com}(A)A$. □

Corollaire XXIII.16.a. Soit $A \in GL_n(\mathbb{K})$. Alors :

$$ A^{-1} = \frac{1}{\det(A)}\,{}^t\mathrm{Com}(A)\,. $$

➠ **Exemple XXIII.17.** Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{K})$. Alors :

$$ A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. $$

✌ **Remarque XXIII.14.** Calculer l'inverse d'une matrice $n \times n$ de cette façon demande de calculer n^2 déterminants de taille $(n-1) \times (n-1)$, soit $n^2 n!$ multiplications. Ceci est déraisonnable : le calcul d'un inverse 20×20 demanderait plus de 150 ans sur une machine décente. Heureusement pour nous, le chapitre XXIV nous fournira une méthode bien plus efficace de calcul d'inverse…

Chapitre XXIV

Systèmes linéaires

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

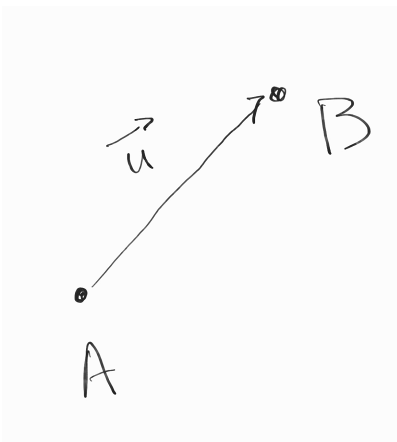
0.⁻ Sous–espaces affines d'un $\mathbb{K}$–e.v

L'objet de ce paragraphe est d'introduire, sans excès de technicité, la notion de sous–espace affine d'un $\mathbb{K}$–e.v. Nous fixons donc dans celui–ci un $\mathbb{K}$–e.v E.

La **structure affine d'un espace vectoriel** est, à notre niveau, un paradigme visant à distinguer au sein de celui–ci deux types d'objets essentiels : les **points** et les **vecteurs**. Nous appuyant sur les connaissances existence du lecteur, nous noterons, pour $A, B \in E$

$$B = A + \vec{u}$$

pour signifier que l'objet $\vec{u}$ ainsi (informellement) défini est le vecteur (au sens vu dans l'enseignement secondaire) $\overrightarrow{AB}$. Notons que l'ensemble $\vec{E}$ des vecteurs (affines) de E peut être identifié à $\ldots E$. Nous conserverons toutefois une notation différente pour celui–ci, afin de préserver la santé mentale du lecteur.



Définition XXIV.1. Soit $\vec{u}$ un vecteur (affine) de E. On appelle **translation** associée à $\vec{u}$ l'application

$$\begin{aligned} \tau_{\vec{u}} : E &\to E \\ A &\mapsto A + \vec{u}\,. \end{aligned}$$

➠ **Exemple XXIV.1.** Visualiser la translation de vecteur $\vec{u} = (1, 1)$ dans $\mathbb{R}^2$.

Définition XXIV.2. Soit $\vec{F}$ un s–e.v de $\vec{E}$ et soit $A \in E$. On appelle **sous–espace affine de direction $\vec{F}$ passant par** A l'ensemble :

$$A + \vec{F} = \{A + \vec{u} \,|\, \vec{u} \in \vec{F}\}.$$

✌ **Remarque XXIV.1.** On a donc :

$$A + \vec{F} = \{\tau_{\vec{u}}(A) \,|\, \vec{u} \in \vec{F}\}.$$

Vocabulaire. Un sous–espace affine dirigé par une droite (resp. un plan, un hyperplan) est appelé droite (resp. plan, hyperplan) affine. De façon générale, on appelle **dimension** d'un tel espace la dimension (éventuellement infinie) de sa direction.

➠ **Exemple XXIV.2.** Les sous–espaces affines de $\mathbb{R}^2$ et $\mathbb{R}^3$ ont été étudiés par le lecteur en terminale.

✌ **Remarque XXIV.2.** L'intersection de deux sous–espaces affines est soit vide, soit un sous–espace affine.

Proposition XXIV.1. Soit F un $\mathbb{K}$–e.v, $u \in \mathscr{L}(E, F)$ et soit $A \in F$. Alors l'ensemble

$$u^{-1}(\{A\}) = \{x \in E \,|\, u(x) = A\}$$

est soit vide, soit un sous–espace affine de E dirigé par $\mathrm{Ker}(u)$.

Démonstration. Supposons $u^{-1}(\{A\})$ non vide et fixons $x_0 \in u^{-1}(\{A\})$. Alors, pour tout $x \in E$:

$$\begin{aligned} x \in u^{-1}(\{A\}) &\Leftrightarrow u(x) = A \\ &\Leftrightarrow u(x) = u(x_0) \\ &\Leftrightarrow x - x_0 \in \mathrm{Ker}(u) \end{aligned}$$

d'où le résultat. □

✌ **Remarque XXIV.3.** Ceci entraîne que si $(\mathscr{E})$ est une équation différentielle linéaire sur un intervalle I, alors $\mathrm{Sol}_{\mathscr{E}}$ est un sous–espace affine de $\mathbb{K}^I$ de direction $\mathrm{Sol}_{\mathscr{H}}$ (où $(\mathscr{H})$ est l'équation homogène associée à $(\mathscr{E})$), passant par une solution particulière de $(\mathscr{E})$. La dimension de $\mathrm{Sol}_{\mathscr{E}}$ est de plus égale à l'ordre de l'équation.

Définition XXIV.3. On appelle **repère affine** de E la donnée d'une base $\vec{\mathscr{B}}$ de $\vec{E}$ et d'un point $O \in E$, appelé **origine** du repère.

➠ **Exemple XXIV.3.** Le lecteur saura sans nul doute donner quelques repères affines de $\mathbb{R}^2$ et $\mathbb{R}^3$ tiré de ses cours de sciences physiques et industrielles.

1.⁻ Notion de système linéaire

a) Définition

Définition XXIV.4. Soient $p, q \in \mathbb{N}^*$. On appelle **système linéaire** à p équations et q inconnues toute famille d'équations de la forme :

$$\forall i \in [\![1, p]\!], \quad \sum_{j=1}^{q} a_{i,j} x_j = b_i$$

avec les $a_{i,j}, b_i \in \mathbb{K}$ fixés et les x_i inconnus scalaires.

✌ **Remarque XXIV.4.** Un tel système est équivalent à une **équation matricielle** du type

$$AX = B$$

avec $A = (a_{i,j})_{i,j} \in \mathscr{M}_{p,q}(\mathbb{K})$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix} \in \mathscr{M}_{q,1}(\mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in \mathscr{M}_{p,1}(\mathbb{K})$.

Dans ce cas, A est appelée **matrice du système**, B son **second membre** et X son **inconnue**.

➠ **Exemple XXIV.4.** Le système linéaire

$$\begin{cases} 2x + 3y + z & = & 1 \\ x - y & = & 42 \end{cases}$$

peut s'écrire sous la forme

$$\underbrace{\begin{pmatrix} 2 & 3 & 1 \\ 1 & -1 & 0 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} 1 \\ 42 \end{pmatrix}}_{B}.$$

Définition XXIV.5. On considère un système linéaire $AX = B$. Alors :
- on appelle **rang** du système la quantité $\mathrm{rg}(A)$;
- on dira que le système est **compatible** si il admet au moins une solution $X \in \mathscr{M}_{q,1}(\mathbb{K})$;
- le système sera dit **homogène** si $B = 0$.

✌ **Remarque XXIV.5.**
- Tout système homogène est compatible : $X = 0$ est solution ;
- si le système possède p équations et q inconnues, alors son rang est inférieur ou égal à $\min(p, q)$.

Proposition XXIV.2 (Cas $p = 1$)**.** Soit $A \in \mathscr{M}_{1,q}(\mathbb{K})$ **non nulle** et soit $b \in \mathbb{K}$. Alors :

(i) le système $AX = (b)$ est compatible ;

(ii) l'ensemble des solutions du système homogène $AX = 0$ forme un hyperplan de $\mathbb{K}^q$;

(iii) les solutions du système $AX = (b)$ sont de la forme $X_0 + X_H$ avec X_0 une solution particulière du système et X_H une solution du système homogène $AX = 0$.

Démonstration. Si on note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix}$ l'inconnue du système $AX = (b)$ et $A = \begin{pmatrix} a_1 & \dots & a_q \end{pmatrix}$ sa matrice, ce dernier est équivalent à l'équation scalaire

$$\sum_{j=1}^{q} a_j x_j = b$$

d'où le résultat. □

✌ **Remarque XXIV.6.** La résolution d'un système linéaire à p équations et q inconnues peut donc être interprétée de quatre façons différentes :

— naïvement : on résout un paquet d'équations linéaires simultanément ;

— géométriquement : chaque équation du système décrit un hyperplan (éventuellement affine) de $\mathbb{K}^q$. Nous recherchons donc une intersection de p tels objets (voir aussi la proposition XX.16).

— matriciellement : on résout une équation matricielle à une inconnue, elle aussi matricielle ;

— s'appuyant sur le point précédent, on peut voir que résoudre un tel système s'apparente à résoudre une équation du type $f(x) = b$ avec $b \in \mathbb{K}^p$ et $f \in \mathscr{L}(\mathbb{K}^q, \mathbb{K}^p)$.

b) Structure de l'espace des solutions

Proposition XXIV.3. On considère un système linéaire homogène à p équations et q inconnues

$$(\mathscr{H}) \qquad AX = 0$$

de rang r. Alors l'espace $\mathrm{Sol}_{\mathscr{H}}$ des solutions de celui–ci est un s–e.v de $\mathbb{K}^q$ de dimension $q - r$.

Démonstration. Soit $f \in \mathscr{L}(\mathbb{K}^q, \mathbb{K}^p)$ l'application linéaire canoniquement associée à A. Alors $\mathrm{rg}(f) = \mathrm{rg}(A) = r$ et l'ensemble des solutions de $(\mathscr{H})$ est isomorphe (*via* l'application associant à un vecteur sa matrice dans la base canonique de $\mathbb{K}^q$) à $\mathrm{Ker}(f)$. Le résultat suit par théorème du rang (XX.13). □

Remarque XXIV.7. Si $p < q$ (moins d'équations que d'inconnues), alors $r \leq p < q$ et donc $\dim(\mathrm{Sol}_{\mathscr{H}}) > 0$. Le système $(\mathscr{H})$ admet donc automatiquement une infinité de solutions non triviales.

Proposition XXIV.4. On considère un système linéaire à p équations et q inconnues

$$(\mathscr{S}) \qquad AX = B$$

de rang r que l'on suppose **compatible**. Alors, l'espace $\mathrm{Sol}_{\mathscr{S}}$ des solutions de $(\mathscr{S})$ est un un sous–espace affine de dimension $q-r$ de $\mathbb{K}^q$ dirigé par $\mathrm{Sol}_{\mathscr{H}}$. Plus précisément, si $X_0 \in \mathrm{Sol}_{\mathscr{S}}$, on a :

$$\mathrm{Sol}_{\mathscr{S}} = X_0 + \mathrm{Sol}_{\mathscr{H}}\,.$$

Démonstration. Ceci découle de la proposition XXIV.1 appliquée à l'application canoniquement associée à A. □

2.– Systèmes de Cramer

Nommés ainsi en l'honneur du mathématicien et philosophe genevois Gabriel Cramer (1704—1752), les systèmes de Cramer représentent d'une certaine façon le "cas idéal" de l'étude des systèmes linéaires.

a) C'est quoi ?

Définition XXIV.6. Un système linéaire est dit **de Cramer** si :

(i) il admet autant d'équations que d'inconnues ;

(ii) sa matrice est inversible.

Remarque XXIV.8. En résumé, un système de Cramer est un système linéaire de la forme $AX = B$ avec $A \in GL_p(\mathbb{K})$. Reprenant les notations du paragraphe précédent, nous avons donc $p = q = r$.

Exemple XXIV.5. Le système

$$\begin{cases} 2x + y &= 1 \\ x + 2y &= 2 \end{cases}$$

est de Cramer.

Remarque XXIV.9. Soit $AX = B$ un système de Cramer. Alors il admet un unique solution, donnée par $X = A^{-1}B$.

Proposition XXIV.5. Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors les propriétés suivantes sont équivalentes :

(i) $A \in GL_n(\mathbb{K})$;

(ii) $\forall B \in \mathscr{M}_{n,1}(\mathbb{K})$, le système $AX = B$ est compatible ;

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

(iv) $\forall B \in \mathscr{M}_{n,1}(\mathbb{K})$, si le système $AX = B$ admet un solution, celle–ci est unique ;

(v) l'ensemble des solutions du système $AX = 0$ est réduit à $\{0\}$.

Démonstration. Soit f l'application canoniquement associée à A ; on peut alors traduire les propriétés énoncées comme suit :

(i) f est bijective ;

(ii) f est surjective ;

(iii) f est bijective ;

(iv) f est injective ;

(v) $\mathrm{Ker}(f) = \{0\}$.

Or, nous avons que ces propriétés sont équivalentes car f est un endomorphisme de $\mathbb{K}^n$, d'où le résultat. □

b) Formules de Cramer

Le résultat qui suit n'est pas exigible, et est uniquement donné pour parfaire la culture scientifique du lecteur.

Proposition XXIV.6. Soit $A \in GL_n(\mathbb{K})$ et soit $B \in \mathscr{M}_{n,1}(\mathbb{K})$. On note, pour $j \in [\![1, n]\!]$, A_j la matrice obtenue en remplaçant la j–ième colonne de A par B. Alors l'unique solution du système de Cramer $AX = B$ est donnée par :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

où, pour tout $j \in [\![1, n]\!]$, on pose :

$$x_j = \frac{\det(A_j)}{\det(A)}.$$

➠ **Exemple XXIV.6.** On considère, pour $a, b, c, d, \alpha, \beta \in \mathbb{K}$ tels que $ad - bc \neq 0$ le système de Cramer

$$\begin{cases} ax + by &= \alpha \\ cx + dy &= \beta \end{cases}.$$

On a alors :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ A_1 = \begin{pmatrix} \alpha & b \\ \beta & d \end{pmatrix} \text{ et } A_2 = \begin{pmatrix} a & \alpha \\ c & \beta \end{pmatrix}.$$

L'unique solution de ce système est donc $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ avec $x_1 = \dfrac{\alpha d - \beta b}{ad - bc}$ et $x_2 = \dfrac{a\beta - c\alpha}{ad - bc}$.

Ces formules peuvent paraître séduisantes, mais prenons un instant pour nous poser la question de leur coût : appliquer ces dernières nous demande en effet le calcul de n déterminants, chacun de ceux–ci étant une somme de $n!$ produits de n termes... À titre d'illustration, nous donnons quelques exemples de temps de calcul approximatifs sur une machine effectuant *grosso modo* 10^{10} opérations par seconde (ce qui est raisonnable).

n	Temps d'exécution
3	2.10^{-9}s
5	7.10^{-8}s
10	4.10^{-3}s
20	160 ans
25	10^9 ans
100	3.10^{142} ans
1000	10^{2553} ans

Pour l'érudition du lecteur, rappelons que l'on estime l'age de l'univers à 14 milliards d'années...

3.⁻ Opérations élémentaires

Nous avons déjà abordé dans les chapitres XX et XXIII la notion d'**opération élémentaire** sur les lignes et colonnes d'une matrice ou sur une famille de vecteurs. Nous donnons dans ce paragraphe une vision purement matricielle de ces dernières.

Nous fixons dans tout ce paragraphe une matrice $A = (a_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$.

a) Matrices élémentaires

Rappelons que pour $k, \ell \in [\![1, n]\!]$ nous avons défini au chapitre XXII la matrice $E_{k,l}$ dont le coefficient en position $(i, j) \in [\![1, n]\!]^2$ est $\delta_{i,k}\delta_{j,\ell}$. Ces matrices forment une base du $\mathbb{K}$–e.v $\mathscr{M}_n(\mathbb{K})$.

Alors, pour $(i, j, k, \ell) \in [\![1, n]\!]^4$, le coefficient situé en position (i, j) de la matrice $AE_{k,\ell}$ est :

$$\sum_{s=1}^{n} a_{i,s}\delta_{s,k}\delta_{j,\ell} = a_{i,k}\delta_{j,\ell}$$

et donc, si l'on note $C_1, \dots, C_n$ les colonnes de la matrice A on a :

$$AE_{k,\ell} = \left(\begin{array}{c|c|c|c|c|c|c} 0 & \dots & 0 & C_k & 0 & \dots & 0 \end{array} \right),$$

la colonne C_k se situant en position ℓ. De la même façon, on vérifie que, si on note

$L_1, \dots, L_n$ les lignes de A :

$$E_{k,\ell}A = \begin{pmatrix} 0 \\ \hline \vdots \\ \hline 0 \\ \hline L_\ell \\ \hline 0 \\ \hline \vdots \\ \hline 0 \end{pmatrix}$$

la ligne L_ℓ se situant en position k.

b) Multiplication par un scalaire

Soit $\lambda \in \mathbb{K}$ et $\ell \in [\![1, n]\!]$; posons :

$$M_\ell^\lambda = I_n + (\lambda - 1)E_{\ell,\ell}\,.$$

Les calculs du paragraphe précédent nous permettent d'affirmer que :

(i) l'opération élémentaire $C_\ell \leftarrow \lambda C_\ell$ sur A est équivalente au calcul du produit AM_ℓ^λ ;

(ii) l'opération élémentaire $L_\ell \leftarrow \lambda L_\ell$ sur A est équivalente au calcul du produit $M_\ell^\lambda A$.

✌ **Remarque XXIV.10.** Comme $\det(M_\ell^\lambda) = \lambda$ (il s'agit d'une matrice diagonale), on retrouve l'impact des opérations élémentaires *supra* sur le déterminant.

c) Échange

Pour $(i, j) \in [\![1, n]\!]^2$, on pose :

$$X_{i,j} = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}\,.$$

On vérifie alors que :

(i) l'opération élémentaire $C_i \leftrightarrow C_j$ est équivalente au calcul du produit $AX_{i,j}$;

(ii) l'opération élémentaire $L_i \leftrightarrow L_j$ est équivalente au calcul du produit $X_{i,j}A$

✌ **Remarque XXIV.11.** Dans la lignée de la remarque précédente, il est possible de vérifier que $\det(X_{i,j}) = -1$.

d) Combinaisons linéaires

Pour $(i, j) \in [\![1, n]\!]^2$ et $\lambda \in \mathbb{K}$, on pose :

$$\Omega_{i,j}^\lambda = I_n + \lambda E_{i,j}\,.$$

Il est alors aisé de vérifier que :

(i) l'opération élémentaire $C_i \leftarrow C_i + \lambda C_j$ sur A est équivalente au calcul du produit $A\Omega_{i,j}^\lambda$;

(ii) l'opération élémentaire $L_i \leftarrow L_i + \lambda L_j$ sur A est équivalente au calcul du produit $\Omega_{i,j}^\lambda A$.

✌ **Remarque XXIV.12.** Une fois n'est pas coutume, on peut aisément vérifier que $\det(\Omega_{i,j}^\lambda) = 1$; les opérations élémentaires *supra* n'ont donc aucun impact sur le déterminant.

e) So what ?

Cette vision matricielle des opérations élémentaires nous permet de démontrer la proposition suivante.

Proposition XXIV.7. Soient $A \in \mathscr{M}_p(\mathbb{K})$ et $B \in \mathscr{M}_q(\mathbb{K})$. Alors A et B sont équivalentes si et seulement si il est possible de passer de A à B grâce à des opérations élémentaires sur les lignes et/ou les colonnes.

✌ **Remarque XXIV.13.** Notons de plus que, par définition de la matrice d'une application linéaire, si A est la matrice d'une application linéaire f :

(i) les opérations élémentaires sur les colonnes de A préservent le rang et l'image de f ;

(ii) les opérations élémentaires sur les lignes de A préservent le rang et le noyau de f.

On en déduit que **pour résoudre un système linéaire, seules les opérations sur les lignes sont admissibles**.

4.⁻ Pivot de Gauss

a) Algorithme

◇ Cas d'étude

On se place dans ce paragraphe dans le cas d'un système linéaire de la forme $AX = B$ avec $A = (a_{i,j})_{i,j} \in \mathscr{M}_{p,q}(\mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in \mathscr{M}_{p,1}(\mathbb{K})$. Son inconnue est de la forme $X = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix} \in \mathscr{M}_{q,1}(\mathbb{K})$. Nous posons également $r = \mathrm{rg}(A)$: il s'agit de fait du rang du système étudié.

◇ Initialisation

Si $r = 0$, alors A est la matrice nulle. Nous avons entière confiance en la capacité de notre lecteur à résoudre le système linéaire étudié.

Dans le cas contraire, la matrice A admet *a minima* un coefficient non nul. Quitte à échanger des lignes (équations) ou colonnes (**inconnues** : prendre garde à ne pas oublier que l'on a procédé à cet échange !), il nous est possible de supposer que $a_{1,1} \neq 0$.

On effectue alors les opérations élémentaires suivantes (dans l'ordre) sur les matrices A **et** B :

— $L_1 \leftarrow \dfrac{1}{a_{1,1}} L_1$;

— $L_i \leftarrow L_i - a_{i,1} L_1$ pour $i \in [\![2, p]\!]$.

Nous venons, par équivalence matricielle, de transformer notre système en un autre, noté $A_1X = B_1$, tel que la matrice A_1 ait la forme suivante :

$$A_1 = \begin{pmatrix} 1 & & \\ 0 & (\star) & \\ \vdots & & \\ 0 & & \end{pmatrix}.$$

◇ **Hérédité**

L'objectif de cette étape est de donner un procédé permettant de transformer (*via* équivalence matricielle) un système linéaire de la forme $A_kX = B_k$ avec

$$A_k = \left(\begin{array}{c|c} I_k & (\star) \\ \hline 0 & (\star) \end{array}\right) \in \mathscr{M}_{p,q}(\mathbb{K})$$

et $k \in [\![1, r-1]\!]$ en un système de la forme $A_{k+1}X = B_{k+1}$, avec

$$A_{k+1} = \left(\begin{array}{c|c} I_{k+1} & (\star) \\ \hline 0 & (\star) \end{array}\right).$$

Quitte à opérer sur les lignes et (prudemment) sur les colonnes, nous pouvons supposer que $a_{k+1,k+1} \neq 0$ (car $k < r$). Il nous suffit alors d'effectuer les opérations élémentaires suivantes :

— $L_{k+1} \leftarrow \dfrac{1}{a_{k+1,k+1}} L_{k+1}$;

— $L_i \leftarrow L_i - a_{i,k+1}L_{k+1}$ pour $i \neq k+1$.

Une fois ce procédé terminé, si $k+1 < r$, on itère. Dans le cas contraire, on a $k+1 = r$ et donc la matrice A_r peut être transformée, quitte à opérer sur les lignes, en :

$$A' = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) = J_r.$$

Le système peut alors être résolu par remontée si il est compatible. Dans le cas contraire, une équation contradictoire apparaîtra, mettant fin à nos efforts. Notons que dans le cas compatible certaines inconnues peuvent ne pas être déterminées de façon unique : cela signifie que l'on dispose de familles de solution à paramètres, les valeurs de ces derniers pouvant être choisies arbitrairement.

✍ **Exercice XXIV.1.** Utiliser ce procédé pour résoudre les systèmes linéaires suivants :

$$\begin{cases} x + 2y - z &= 2 \\ x + z &= 2 \\ x + y + z &= 1 \end{cases},$$

$$\begin{cases} x + y + z &= 0 \\ x + 4z &= 2 \end{cases}$$

et

$$\begin{cases} x + y &= -1 \\ x - y &= 2 \\ 2x + 3y &= 1 \end{cases}.$$

b) À quel prix ?

Une application de l'algorithme du pivot de Gauss à un système de rang r nécessitera r étapes d'élimination. Dans chacune d'entre elles, nous devons normaliser une ligne ($\approx q$ divisions) et la soustraire à au plus p autres ($\approx q$ multiplications à chaque fois).

In fine, la complexité algorithmique du pivot de Gauss est en $\mathcal{O}(rpq)$. Dans le cas d'un système de Cramer, avec $A \in GL_n(\mathbb{K})$, nous avons affaire à du $\mathcal{O}(n^3)$. Ceci est **bien plus efficace** que les formules de Cramer, comme illustré dans le tableau *infra*, pour lequel nous avons repris notre "machine" effectuant 10^{10} opérations par seconde.

n	Cramer	Gauss
3	2.10^{-9}s	5.10^{-9}s
5	7.10^{-8}s	2.10^{-8}s
10	4.10^{-3}s	2.10^{-7}s
20	160 ans	2.10^{-6}s
25	10^9 ans	3.10^{-6}s
100	3.10^{142} ans	2.10^{-4}s
1000	10^{2553} ans	$0,2$s

c) Applications

◇ Calcul de déterminants

Le pivot permet de trigonaliser une matrice de façon effective. Quitte à tenir compte des opérations élémentaires effectuées, ceci permet un calcul de déterminant efficace.

◇ Inversion d'une matrice carrée

Pour déterminer l'inverse d'une matrice $A \in GL_n(\mathbb{K})$, on lui applique le pivot de Gauss jusqu'à obtenir la matrice I_n. Si l'on effectue les mêmes opérations élémentaires, dans le même ordre, sur la matrice identité, on obtient la matrice inverse A^{-1}.

➠ **Exemple XXIV.7.** Déterminons l'inverse de la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$.

Pour ce faire, nous partons de :

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 10 & 0 & 0 & 1 \end{array}$$

et effectuons les opérations $L_2 \leftarrow L_2 - 4L_1$ et $L_3 \leftarrow L_3 - 7L_1$, obtenant :

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -11 & -7 & 0 & 1 \end{array}.$$

Nous effectuons ensuite $L_2 \leftarrow \dfrac{-1}{3} L_2$:

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & -6 & -11 & -7 & 0 & 1 \end{array}$$

puis $L_1 \leftarrow L_1 - 2L_2$ et $L_3 \leftarrow L_3 + 6L_2$:

$$\begin{array}{ccc|ccc} 1 & 0 & -1 & -\frac{5}{3} & \frac{2}{3} & 0 \\ 0 & 1 & 2 & \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}$$

avant de conclure par $L_1 \leftarrow L_1 + L_3$ et $L_2 \leftarrow L_2 - 2L_3$:

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{3} & -\frac{4}{3} & 1 \\ 0 & 1 & 0 & -\frac{2}{3} & \frac{11}{3} & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} .$$

In fine, on trouve :

$$A^{-1} = \begin{pmatrix} -\frac{4}{3} & -\frac{4}{3} & 1 \\ -\frac{2}{3} & \frac{11}{3} & -2 \\ 1 & -2 & 1 \end{pmatrix} .$$

Chapitre XXV

Séries numériques

On fixe dans tout ce chapitre un corps $\mathbb{K}$ égal à $\mathbb{R}$ ou $\mathbb{C}$.

1.⁻ Qu'est–ce ?

a) Notion de série

Définition XXV.1. Soit $u \in \mathbb{K}^{\mathbb{N}}$. On appelle **série de terme général** u_n la suite

$$\left(\sum_{n=0}^{N} u_n\right)_{N\in\mathbb{N}} .$$

Pour $N \in \mathbb{N}$ fixé, on appelle **somme partielle de rang** N de la série la quantité

$$S_N = \sum_{n=0}^{N} u_n .$$

✌ **Remarque XXV.1.** La somme partielle de rang N d'une série est donc le N-ième terme de la suite sous–jacente à celle–ci.

Notation. On note $\sum u_n$ la série de terme général u_n. **Ceci n'est PAS une somme**, mais une suite de sommes.

✌ **Remarque XXV.2.** Il est possible de considérer des séries dont le terme général n'est défini qu'à partir d'un rang $n_0 \in \mathbb{N}$. On noter alors $\sum\limits_{(n\geq n_0)} u_n$. Les résultats énoncés dans ce chapitre se généralisent aisément à ce type de séries.

➠ **Exemple XXV.1.** $\sum \cos(n)$ est une série définie à partir du rang 0 ; $\sum\limits_{(n\geq 1)} \frac{1}{n}$ est définie à partir du rang 1.

⋄ **Exemples fondamentaux**

Définition XXV.2. Soit $q \in \mathbb{C}$. On appelle **série géométrique de raison** q la série $\sum q^n$.

✌ **Remarque XXV.3.** Il est aisé de déterminer les sommes partielles d'une série géométrique ; si $N \in \mathbb{N}$ et $q \in \mathbb{C}$ on a :

$$\sum_{n=0}^{N} q^n = \begin{cases} \dfrac{1-q^{N+1}}{1-q} \text{ si } q \neq 1 \\ N+1 \text{ sinon.} \end{cases}$$

Définition XXV.3. Soit $\alpha \in \mathbb{C}$. On appelle **série de Riemann de paramètre** α la série $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$.

➠ **Exemple XXV.2.** La série de Riemann de paramètre 1 est la série $\sum\limits_{(n\geq 1)} \frac{1}{n}$. On l'appelle **série harmonique**.

b) Somme, restes

Définition XXV.4. Une série $\sum u_n$ est dite **convergente** si la suite de ses sommes partielles converge. Dans le cas contraire, on dit que la série est **divergente**.

Notation. **Si** la série $\sum u_n$ est convergente, on pose :

$$\sum_{n=0}^{\infty} u_n = \lim_{N\to\infty} \sum_{n=0}^{N} u_n \,.$$

Cette quantité, que l'on appelle **somme** de la série $\sum u_n$, n'est **toujours pas** une somme : il s'agit de la limite d'une suite de sommes ...

➠ **Exemple XXV.3.**

— La série $\sum n$ diverge ; en effet, pour $N \geq 0$:

$$\sum_{n=0}^{N} n = \frac{N(N+1)}{2} \xrightarrow[N\to\infty]{} \infty \,.$$

— À l'inverse, la série $\sum\limits_{(n\geq 1)} \frac{1}{n(n+1)}$ converge. En effet, pour $N \geq 1$ on a :

$$\begin{aligned} \sum_{n=1}^{N} \frac{1}{n(n+1)} &= \sum_{n=1}^{N} \frac{1}{n} - \frac{1}{n+1} \\ &= \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=1}^{N} \frac{1}{n+1} \\ &= \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=2}^{N+1} \frac{1}{n} \\ &= 1 - \frac{1}{N+1} \\ &\xrightarrow[N\to\infty]{} 1 \end{aligned}$$

et donc
$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1\,.$$

Proposition XXV.1. Soit $\sum u_n$ une série convergente. Alors la suite $(u_n)_n$ converge vers 0.

Démonstration. Il suffit de remarque que, si on note $(S_N)_N$ la suite des sommes partielles de $\sum u_n$, on a, pour tout $N \geq 0$:

$$\begin{aligned} u_N &= S_N - S_{N-1} \\ &\xrightarrow[N\to\infty]{} \sum_{n=0}^{\infty} u_n - \sum_{n=0}^{\infty} u_n \\ &= 0 \end{aligned}$$

d'où le résultat. □

✖ **Attention :** la réciproque est **fausse**. Pour un contre–exemple, considérer la série $\sum \sqrt{n+1} - \sqrt{n}$: on vérifie (faire un DL) que $\sqrt{n+1} - \sqrt{n} \to 0$ et pourtant, pour tout $N \geq 0$ on a :

$$\sum_{n=0}^{N} \sqrt{n+1} - \sqrt{n} = \sqrt{N+1} \to \infty\,.$$

✌ **Remarque XXV.4.** Ce résultat sera toutefois fort utile pour démontrer la **divergence** d'une série : si la suite $(u_n)_n$ ne converge pas vers 0, alors la suite $\sum u_n$ diverge automatiquement : on parle alors de **divergence grossière**.

➠ **Exemple XXV.4.** Les séries $\sum \cos(n)$ et $\sum n^3 + \frac{2}{n+1}$ divergent grossièrement.

Proposition XXV.2. Soit $q \in \mathbb{C}$. Alors :

$$\text{la série géométrique } \sum q^n \text{ converge} \\ \iff \\ |q| < 1.$$

Dans ce cas, on a de plus :
$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}\,.$$

Démonstration. **Cas 1 : $|q| < 1$.** Comme nous le savons, si $N \geq 0$ on a :

$$\sum_{n=0}^{N} q^n = \frac{1-q^{N+1}}{1-q}\,.$$

Il est alors clair que la suite des sommes partielles de $\sum q^n$ converge vers $\dfrac{1}{1-q}$.

Cas 2 : $|q| \geq 1$. La suite $(q^n)_n$ ne convergeant pas vers 0, on a divergence grossière (proposition XXV.1) de la série $\sum q^n$.

□

Proposition XXV.3. Soient $\sum u_n$ et $\sum v_n$ deux séries **convergentes** et soit $\lambda \in \mathbb{K}$. Alors :

(i) la série $\sum u_n + \lambda v_n$ converge ;

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

Démonstration. Appliquer les opérations sur les limites vues dans le chapitre IX aux sommes partielles. □

✖ **Attention :** ce type d'opérations ne saurait être effectué que lorsque les **deux** séries considérées convergent.

Définition XXV.5. Soit $\sum u_n$ une série **convergente** de somme S. Pour $N \in \mathbb{N}$, on appelle **reste d'ordre** N de la série $\sum u_n$ la quantité $R_N = S - S_N$, où S_N est la somme partielle de rang N de celle–ci.

Notation. Le reste d'ordre N de $\sum u_n$ est noté

$$R_N = \sum_{n=N+1}^{\infty} u_n \,.$$

✌ **Remarque XXV.5.** Il est aisé de vérifier que, comme $\sum u_n$ converge, $R_N \xrightarrow[N\to\infty]{} 0$.

➠ **Exemple XXV.5.** Soit $q \in \mathbb{C}$ tel que $|q| < 1$ et soit $N \in \mathbb{N}$. Alors :

$$\sum_{n=N+1}^{\infty} q^n = \frac{q^{N+1}}{1-q} \,.$$

c) Lien suite–série

Proposition XXV.4. Soit $u \in \mathbb{K}^{\mathbb{N}}$. Alors :

$$\text{la suite } (u_n)_n \text{ converge} \iff \text{la série } \sum u_{n+1} - u_n \text{ est convergente.}$$

Dans ce cas, on a de plus :

$$\sum_{n=0}^{\infty} u_{n+1} - u_n = \lim u_n - u_0 \,.$$

Démonstration. Soit $N \in \mathbb{N}$. Alors :

$$\begin{aligned}\sum_{n=0}^{N} u_{n+1} - u_n &= \sum_{n=0}^{N} u_{n+1} - \sum_{n=0}^{N} u_n \\ &= \sum_{n=1}^{N+1} u_n - \sum_{n=0}^{N} u_n \\ &= u_{N+1} - u_0\end{aligned}$$

d'où le résultat. □

Vocabulaire. Les séries du type $\sum u_{n+1} - u_n$ sont appelées **séries télescopiques**.

Exemple XXV.6. La série $\sum_{(n\geq 1)} \ln\left(\frac{n+1}{n}\right)$ diverge. En effet, son terme général est égal, pour $n \geq 1$, à $\ln(n+1) - \ln(n)$ et $\ln(n) \to \infty$.

2.‾ Séries à termes positifs

On s'intéresse dans ce paragraphe à une famille particulière de séries, en l'occurrence celles dont le terme général ne prend que des valeurs réelles positives.

Définition XXV.6. Une **série à termes positifs** est une série $\sum u_n$ telle que :

$$\forall n \geq 0,\ u_n \in \mathbb{R}_+ .$$

a) Critère de convergence

Proposition XXV.5. Une série à termes positifs converge si et seulement si la suite de ses sommes partielles est majorée.

Démonstration. Ceci est immédiat car la suite des sommes partielles d'une série à termes positifs est une suite réelle croissante. □

✖ **Attention :** ce résultat est bien entendu faux lorsque la série n'est pas à termes positifs. La suite $\sum(-1)^n$ diverge grossièrement et pourtant ses sommes partielles sont comprises entre 0 et 1.

b) Comparaison des séries à termes positifs

Proposition XXV.6. Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs telles que, $\exists\, N \in \mathbb{N}$, $\forall n \geq N$, $0 \leq u_n \leq v_n$. Alors :

(i) la convergence de $\sum v_n$ implique celle de $\sum u_n$;

(ii) la divergence de $\sum u_n$ implique celle de $\sum v_n$.

Démonstration. Ceci découle du critère de convergence des séries à termes positifs (proposition XXV.6). □

De fait, ce résultat nous permet de déduire la convergence (ou divergence) de nombreuses séries à termes positifs en les comparant à des exemples connus (comme les séries géométriques de ce type) : nous nous constituons ainsi un "catalogue" de séries à termes positifs "témoins" pour établir la nature de leurs consœurs.

➠ **Exemple XXV.7.** La série de terme général $\dfrac{e^{-n!}}{2^n}$ converge par comparaison avec $\sum \frac{1}{2^n}$.

Proposition XXV.7. Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs telles que $u_n \sim v_n$. Alors $\sum u_n$ et $\sum v_n$ sont de même nature.

Démonstration. Par équivalence, il existe une suite $(h_n)_n$ convergeant vers 1 et un entier $N \geq 0$ tel que :

$$\forall n \geq N,\ u_n = h_n v_n\,.$$

On obtient l'existence d'un entier $N' \geq 0$ tel que pour $n \geq N'$, $h_n \leq \pi$ et donc, pour tout $n \geq \max(N, N')$ $u_n \leq \pi v_n$. Ceci entraîne que, par comparaison (proposition XXV.6) que la convergence de $\sum v_n$ entraîne celle de $\sum u_n$. On conclut en observant que les rôles de $(u_n)_n$ et $(v_n)_n$ sont symétriques. □

➠ **Exemple XXV.8.** La série $\sum \frac{n^2+1}{3n^2+2} e^{-n}$ est convergente car son terme général est équivalent à $\dfrac{1}{3}\left(\dfrac{1}{e}\right)^n$, ce dernier étant le terme général d'une série géométrique convergente (à facteur multiplicatif près).

✖ **Attention :** il n'y a pas en général égalité des sommes dans le cas convergent.

c) Critère de d'Alembert

Proposition XXV.8 (Critère de d'Alembert)**.** Soit $\sum u_n$ une série à termes positifs telle que :

$$\forall n \geq 0,\ u_n > 0\,.$$

On suppose que $\dfrac{u_{n+1}}{u_n} \xrightarrow[n\to\infty]{} \ell \in \overline{\mathbb{R}}$. Alors :

(i) si $\ell < 1$, la série $\sum u_n$ converge ;

(ii) si $\ell > 1$, la série $\sum u_n$ diverge grossièrement ;

(iii) si $\ell = 1$, on ne peut conclure.

Démonstration.

(i) Si $\ell < 1$ alors il existe $q < 1$ tel que :

$$\exists\, N \in \mathbb{N},\, \forall n \geq N,\, \frac{u_{n+1}}{u_n} \leq q$$

et donc, pour $n \geq N$:

$$\begin{aligned} u_{n+1} &\leq q u_n \\ &\leq q^2 u_{n-1} \\ &\ \ \vdots \\ &\leq q^{n+1-N} u_N \end{aligned}$$

et donc, par comparaison (proposition XXV.6) avec la série géométrique $\sum q^n$, la série $\sum u_n$ converge.

(ii) Si $\ell > 1$, la suite $(u_n)_n$ est strictement croissante et donc ne peut converger vers 0, d'où le résultat.

(iii) Dans le cas $\ell = 1$, considérer les séries $\sum 1$ et $\sum\limits_{(n\geq 1)} \frac{1}{n} - \frac{1}{n+1}$ pour mesurer la futilité de la chose.

□

➠ **Exemple XXV.9.** La série à termes positifs $\sum \frac{1}{n!}$ converge car :

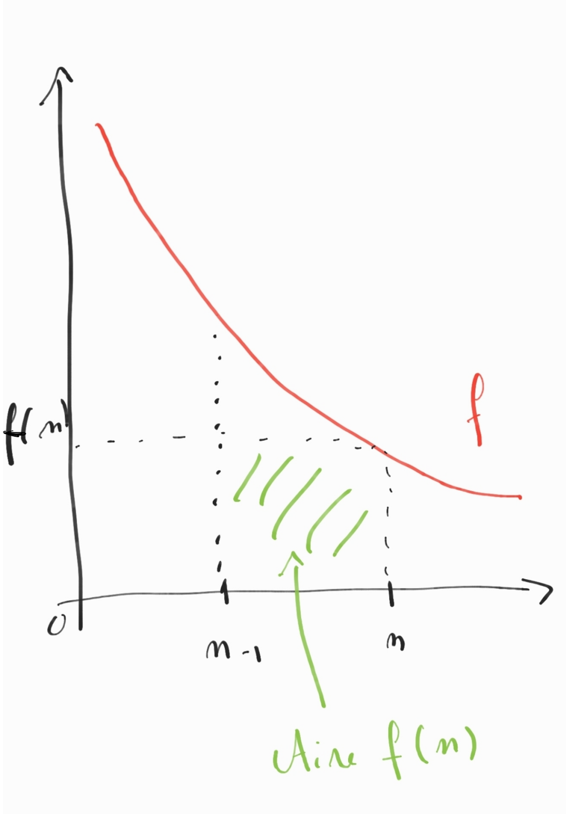
$$\frac{n!}{(n+1)!} = \frac{1}{n+1} \to 0 < 1\,.$$

3.⁻ Comparaison série–intégrale

a) Énoncé

Considérons une fonction $f : \mathbb{R}_+ \to \mathbb{R}$ continue par morceaux et **décroissante**. Alors, on vérifie par méthode des rectangles que :

$$\forall n \geq 1,\ f(n) \leq \int_{n-1}^{n} f(t)\,\mathrm{d}t\,.$$



Ceci nous permet, pour $M, N \in \mathbb{N}$ tels que $M \leq N$ de vérifier que :

$$\begin{aligned} \sum_{n=M}^{N} f(n) &= f(M) + \sum_{n=M+1}^{N} f(n) \\ &\leq f(M) + \sum_{n=M+1}^{N} \int_{n-1}^{n} f(t)\,\mathrm{d}t \\ &= f(M) + \int_{M}^{N} f(t)\,\mathrm{d}t\,. \end{aligned}$$

Ce résultat nous laisse entrevoir un encadrement possible des sommes partielles de la série $\sum f(n)$ par des intégrales, que nous formalisons *infra*.

Proposition XXV.9 (Comparaison série–intégrale)**.** Soit $f \in \mathscr{C}_{\mathrm{pm}}(\mathbb{R}_+)$ une fonction **décroissante**. Alors, pour tout $M, N \in \mathbb{N}$ tels que $M \leq N$ on a :

$$\int_{M}^{N+1} f(t)\,\mathrm{d}t \leq \sum_{n=M}^{N} f(n) \leq f(M) + \int_{M}^{N} f(t)\,\mathrm{d}t\,.$$

Démonstration. Nous avons déjà démontré l'inégalité de droite. Pour celle de gauche, appliquer une heuristique analogue en remarquant que, pour $n \in \mathbb{N}$:

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

□

✌ **Remarque XXV.6.** Un résultat analogue peut évidemment être obtenu pour les fonctions croissantes, quitte à changer le sens des inégalités.

Ce résultat a de nombreuses conséquences, dont nous étudierons certaines dans la suite de ce paragraphe. Dans un premier temps, intéressons nous aux séries de Riemann **de paramètre réel**.

Proposition XXV.10. Soit $\alpha \in \mathbb{R}$. Alors :

(i) si $\alpha \leq 0$, la série $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$ diverge grossièrement ;

(ii) si $\alpha \in]0,1]$, la série $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$ diverge ;

(iii) si $\alpha > 1$, la série $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$ converge.

Démonstration. Commençons par remarquer que, pour $\alpha \in \mathbb{R}$, la fonction $f : x \mapsto x^{-\alpha}$ est continue sur $\mathbb{R}_+^*$.

(i) Trivial.

(ii) **Cas 1 : $0 < \alpha < 1$.** Dans ce cas, une primitive de f est :

$$x \mapsto \frac{x^{1-\alpha}}{1-\alpha}$$

et donc, par comparaison série–intégrale (proposition XXV.9), on a pour $N \geq 1$:

$$\begin{aligned} \sum_{n=1}^{N} \frac{1}{n^\alpha} &\geq \int_1^{N+1} \frac{\mathrm{d}t}{t^\alpha} \\ &= \frac{(N+1)^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha} \\ &\xrightarrow[N\to\infty]{} \infty \end{aligned}$$

d'où le résultat.

Cas 2 : $\alpha = 1$. Une primitive de $f : x \mapsto \frac{1}{x}$ est :

$$x \mapsto \ln(x)$$

et donc, par comparaison série–intégrale, on a pour $N \geq 1$:

$$\begin{aligned} \sum_{n=1}^{N} \frac{1}{n} &\geq \int_1^{N+1} \frac{\mathrm{d}t}{t} \\ &= \ln(N+1) \\ &\xrightarrow[N\to\infty]{} \infty \end{aligned}$$

ce qui entraîne la divergence annoncée.

(iii) Dans ce cas, une primitive de f est donnée par

$$x \mapsto \frac{x^{1-\alpha}}{1-\alpha} = -\frac{1}{(\alpha-1)x^{\alpha-1}}.$$

De fait, par comparaison série–intégrale :

$$\begin{aligned}\forall N \geq 1, \sum_{n=1}^{N} \frac{1}{n^\alpha} &\leq 1 + \int_1^N \frac{\mathrm{d}t}{t^\alpha} \\ &= 1 + \frac{1}{\alpha-1} - \frac{1}{(\alpha-1)N^{\alpha-1}} \\ &\leq 1 + \frac{1}{\alpha-1}\end{aligned}$$

et donc, par critère de convergence des séries à termes positifs (proposition XXV.6), la série $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$ converge.

□

✍ **Exercice XXV.1.** Donner un équivalent du reste de la série $\sum\limits_{(n\geq 1)} \frac{1}{n^2}$.

➥ **Correction :** *On applique une comparaison série–intégrale (proposition XXV.9) pour $M, N \geq 1$ tels que $M \leq N$, obtenant :*

$$\int_M^{N+1} \frac{\mathrm{d}t}{t^2} \leq \sum_{n=M}^{N} \frac{1}{n^2} \leq \frac{1}{M^2} + \int_M^N \frac{\mathrm{d}t}{t^2}$$

i.e :

$$\frac{1}{M} - \frac{1}{N+1} \leq \sum_{n=M}^{N} \frac{1}{n^2} \leq \frac{1}{M^2} + \frac{1}{M} - \frac{1}{N}$$

et donc, en faisant tendre N vers $+\infty$ on obtient :

$$\frac{1}{M} \leq R_{M-1} \leq \frac{1}{M^2} + \frac{1}{M}.$$

Une application du théorème d'encadrement des limites (théorème IX.12) permet ensuite aisément de déduire que $R_M \sim \dfrac{1}{M}$.

b) Étude asymptotique de la série harmonique

Nous avons démontré plus haut que la série harmonique $\sum\limits_{(n\geq 1)} \frac{1}{n}$ diverge. Mais encore ? Nous nous attachons dans ce paragraphe à offrir à notre lecteur une vision plus raffinée de la situation. Commençons par poser, pour $N \geq 1$:

$$S_N = \sum_{n=1}^{N} \frac{1}{n}$$

et remarquer que, par comparaison série–intégrale (proposition XXV.9)

$$\ln(N+1) \leq S_N \leq 1 + \ln(N)$$

ce qui permet de démontrer, *via* le théorème d'encadrement des limites (théorème IX.12) que :

$$S_N \sim \ln(N)\,.$$

Ce premier résultat étant acquis, affinons cette approximation. Pour ce faire, posons, pour $n \geq 1$, $u_n = S_n - \ln(n)$. Alors :

$$\begin{aligned}
u_{n+1} - u_n &= \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n+1) - \sum_{k=1}^{n} \frac{1}{k} + \ln(n) \\
&= \frac{1}{n+1} - \ln\left(\frac{n+1}{n}\right) \\
&= \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right) \\
&= \frac{1}{n+1} - \left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\right)\,.
\end{aligned}$$

Or :

$$\begin{aligned}
\frac{1}{n+1} &= \frac{1}{n}\frac{1}{1+\frac{1}{n}} \\
&= \frac{1}{n}\left(1 - \frac{1}{n} + o\left(\frac{1}{n}\right)\right) \\
&= \frac{1}{n} - \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)
\end{aligned}$$

et donc, en réinjectant :

$$\begin{aligned}
u_{n+1} - u_n &= \frac{1}{n} - \frac{1}{n^2} - \frac{1}{n} + \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \\
&= -\frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\,.
\end{aligned}$$

Ceci nous permet de conclure que $u_n - u_{n+1} \sim \dfrac{1}{2n^2}$ et donc, par comparaison des séries à termes positifs avec la série de Riemann convergente $\sum\limits_{(n\geq 1)} \frac{1}{n^2}$, la série $\sum\limits_{(n\geq 1)} u_n - u_{n+1}$ converge. Par lien suite–série (proposition XXV.4), la suite $(u_n)_n$ converge vers un réel γ, appelée constante d'Euler—Mascheroni (le second, de son prénom Lorenzo, ayant été un mathématicien et abbé cisalpin ; 1750—1800).

Obtenir une valeur approchée de cette constante est aisé à l'aide de (par exemple) `python`. On obtient, en un temps raisonnable, $\gamma \approx 0,5772156649$.

```python
from numpy import log

def approx_gamma(N):
    S=-log(N)
    for i in range(N):
        S+=1/(i+1)
    return S
```

On obtient donc, *in fine*, le développement asymptotique suivant :

$$\sum_{n=1}^{N} \frac{1}{n} = \ln(N) + \gamma + o(1)\,. \tag{E:XXV.1}$$

4.⁻ Convergence absolue

a) What ?

Définition XXV.7. On dit qu'une série $\sum u_n$ **converge absolument** si la série à termes positifs $\sum |u_n|$ est convergente.

Exemple XXV.10.

— Les séries à termes positifs convergentes sont absolument convergentes...

— La série $\sum\limits_{(n\geq 1)} \frac{(-1)^n}{n^2}$ converge absolument.

— Soit $\alpha = a + ib \in \mathbb{C}$ et soit $n \geq 1$. Alors :

$$\begin{aligned}\left|\frac{1}{n^\alpha}\right| &= \frac{1}{|n^a n^{ib}|}\\ &= \frac{1}{n^a |e^{ib\ln(n)}|}\\ &= \frac{1}{n^a}\,.\end{aligned}$$

On en déduit que **la série de Riemann** $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$ **converge absolument si et seulement si** $\mathrm{Re}(\alpha) > 1$**.**

b) Lien à la convergence

Dans le chapitre IX, nous avons démontré que la convergence d'une suite $(u_n)_n$ entraînait celle de $(|u_n|)_n$ avec réciproque fausse. Ce résultat ne se généralise pas aux séries, comme nous le démontrons avec le contre–exemple suivant.

Posons, pour $N \geq 1$

$$H_N = \sum_{n=1}^{N} \frac{1}{n}$$

et intéressons nous à la **série harmonique alternée** $\sum\limits_{(n\geq 1)} \frac{(-1)^n}{n}$, dont nous noterons $(S_N)_N$ la suite des sommes partielles. Alors, pour $N \geq 1$:

$$\begin{aligned}S_{2N} &= \sum_{n=1}^{2N} \frac{(-1)^n}{n}\\ &= -\sum_{n=1}^{2N} \frac{1}{n} + 2\sum_{n=1}^{N} \frac{1}{2n}\\ &= -H_{2N} + 2\frac{1}{2}H_N\\ &= H_N - H_{2N}\end{aligned}$$

et donc, d'après (**E** :XXV.1) :

$$\begin{aligned} S_{2N} &= \ln(N) + \gamma - \ln(2N) - \gamma + o(1) \\ &= -\ln(2) + o(1) \end{aligned}$$

ce qui implique que $S_{2N} \to -\ln(2)$. De plus, on a :

$$\begin{aligned} S_{2N+1} &= S_{2N} + \frac{-1}{2N+1} \\ &\xrightarrow[N\to\infty]{} -\ln(2) \end{aligned}$$

ergo la série $\sum\limits_{(n\geq 1)} \frac{(-1)^n}{n}$ converge et

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln(2)\,.$$

Or, la série harmonique alternée ne converge pas absolument : la série de ses valeurs absolues est la série harmonique. **Convergence n'implique donc pas PAS convergence absolue pour les séries.**

Histoire de compliquer un peu plus les choses, la réciproque est, quant à elle, vraie...

Proposition XXV.11. Toute série absolument convergente est convergente.

Démonstration. Dans le cas réel, nous avons, pour tout $n \in \mathbb{N}$, $|u_n| = u_n^+ + u_n^-$ et $u_n = u_n^+ - u_n^-$ (rappelons que pour $x \in \mathbb{R}$, $x^{\pm} = \max(0, \pm x)$). De fait, les séries $\sum u_n^+$ et $\sum u_n^-$ sont à termes positifs et, pour tout $n \geq 0$, $u_n^{\pm} \leq |u_n|$: par comparaison des séries à termes positifs (proposition XXV.6), ces deux séries convergent si $\sum u_n$ converge absolument, ce qui entraîne la convergence de cette dernière par linéarité.

Dans le cas complexe, on procède de façon similaire avec les séries des parties réelles et imaginaires. □

➠ **Exemple XXV.11.** Si $\alpha \in \mathbb{C}$ est tel que $\mathrm{Re}(\alpha) > 1$, alors la série de Riemann $\sum\limits_{(n\geq 1)} \frac{1}{n^\alpha}$ converge. Ceci permet de définir la **fonction zêta de Riemann** :

$$\begin{aligned} \zeta : \{z \in \mathbb{C} \,|\, \mathrm{Re}(z) > 1\} &\longrightarrow \mathbb{C} \\ z &\mapsto \sum_{n=1}^{\infty} \frac{1}{n^z}\,. \end{aligned}$$

Proposition XXV.12. Soit $\sum u_n$ une série absolument convergente. Alors :

$$\left| \sum_{n=0}^{\infty} u_n \right| \leq \sum_{n=0}^{\infty} |u_n|\,.$$

Démonstration. Passer à la limite dans l'inégalité triangulaire appliquée aux sommes partielles. □

Proposition XXV.13. Soit $u \in \mathbb{C}^{\mathbb{N}}$ et soit $v \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que :
- $u_n = \mathcal{O}(v_n)$;
- $\sum v_n$ converge.

Alors $\sum u_n$ converge absolument.

Démonstration. Par domination, il existe une suite $(h_n)_n$ positive, bornée par $M \in \mathbb{R}_+$ telle que :

$$\exists\, N \in \mathbb{N},\ \forall n \geq N,\ |u_n| \leq h_n v_n$$

et donc, pour $n \geq N$, $|u_n| \leq M v_n$. Par comparaison des séries à termes positifs (proposition XXV.6), on obtient la convergence annoncée. □

Corollaire XXV.13.a. Soit $\alpha \in]1, +\infty[$ et soit $u \in \mathbb{C}^{\mathbb{N}}$ une suite telle que $(n^\alpha u_n)_n$ soit bornée. Alors la série $\sum u_n$ converge absolument.

➠ **Exemple XXV.12.** La série $\sum\limits_{(n\geq 1)} \frac{\sin(n)}{n^3}$ converge absolument.

Addendum : représentations décimales d'un réel

Soit $x \in \mathbb{R}$; alors on peut écrire x sous la forme :

$$x = a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots$$

avec les $a_k \in [\![0, 9]\!]$. Avec les conventions d'écriture vues dans ce chapitre, cela revient à écrire :

$$x = \sum_{k=0}^{\infty} \frac{a_k}{10^k}\,.$$

Une telle écriture est appelée **développement décimal** de x. Notons hélas que ce développement n'est pas unique ; en effet :

$$\begin{aligned} 0,99999\ldots &= \sum_{k=1}^{\infty} \frac{9}{10^k} \\ &= \frac{9}{10} \frac{1}{1 - \frac{1}{10}} \\ &= 1\,. \end{aligned}$$

Pour résoudre ce problème, nous appelons **développement décimal propre** de x tout développement décimal $(a_k)_k$ n'étant pas stationnaire à 9. On a alors le résultat (admis) suivant.

Théorème XXV.14.
Tout nombre réel admet un unique développement décimal propre.

Plus précisément, on peut démontrer que seuls les réels de l'ensemble

$$\mathscr{E} = \left\{ \frac{n}{10^k} \;\middle|\; n \in \mathbb{Z},\ k \in \mathbb{N} \right\}$$

admettent deux développements décimaux : l'un stationnaire à zéro (le développement propre) et l'autre stationnaire à 9.

Chapitre XXVI

Espaces préhilbertiens réels

1.⁻ Produits scalaires

a) C'est quoi ?

Définition XXVI.1. Soit E un $\mathbb{R}$–e.v ; on appelle **produit scalaire** sur E tout application $\phi : E \times E \to \mathbb{R}$ telle que :

(A) ϕ est une **forme bilinéaire**, *i.e* :

$$\forall x, y, z \in E, \forall \lambda \in \mathbb{R}, \quad \phi(x + \lambda y, z) = \phi(x, z) + \lambda\phi(y, z)$$

et

$$\forall x, y, z \in E, \forall \lambda \in \mathbb{R}, \quad \phi(z, x + \lambda y) = \phi(z, x) + \lambda\phi(z, y) ;$$

(B) ϕ est **symétrique**, *i.e* :

$$\forall x, y \in E, \quad \phi(x, y) = \phi(y, x) ;$$

(C) ϕ est **définie**, *i.e* :

$$\forall x \in E, \quad (\phi(x, x) = 0) \Rightarrow (x = 0) ;$$

(D) ϕ est **positive**, *i.e* :

$$\forall x \in E, \quad \phi(x, x) \geq 0 .$$

Notation. Le produit scalaire sur un $\mathbb{R}$–e.v E de deux vecteurs $x, y \in E$ sera souvent noté $\langle x\,,\,y\rangle$, $(x\,|\,y)$ ou $x \cdot y$.

➠ **Exemple XXVI.1.**

— Sur $E = \mathbb{R}^n$ avec $n \in \mathbb{N}^*$, on définit le **produit scalaire canonique** de deux vecteurs $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ comme :

$$\langle x\,,\,y\rangle = \sum_{k=1}^{n} x_k y_k .$$

Cette formule définit bien un produit scalaire sur $\mathbb{R}^n$ dont la définition coïncide avec celle vue en sciences appliquées pour les dimensions 2 et 3.

— Sur $E = \mathscr{C}^0([a,b])$, on a le produit scalaire suivant :

$$\phi : (f,g) \mapsto \int_a^b f(t)g(t)\,\mathrm{d}t\,.$$

Remarque XXVI.1. Si $x \in E$ et $\langle\cdot\,,\,\cdot\rangle$ est un produit scalaire sur E, on a :

$$\begin{aligned}\langle x\,,\,0\rangle &= \langle x\,,\,2\cdot 0\rangle\\ &= 2\,\langle x\,,\,0\rangle\end{aligned}$$

et donc $\langle x\,,\,0\rangle = 0 = \langle 0\,,\,x\rangle$.

Exercice XXVI.1. Démontrer que l'application suivante est un produit scalaire sur $\mathscr{M}_n(\mathbb{R})$:

$$\begin{aligned}\langle\cdot\,,\,\cdot\rangle : \mathscr{M}_n(\mathbb{R})^2 &\longrightarrow \mathbb{R}\\ (A,B) &\mapsto \mathrm{Tr}({}^{\mathrm{t}}AB)\ .\end{aligned}$$

Définition XXVI.2. On appelle **espace préhilbertien réel** la donnée d'un couple $(E, \langle\cdot\,,\,\cdot\rangle)$, où :
- E est un $\mathbb{R}$–e.v ;
- $\langle\cdot\,,\,\cdot\rangle$ est un produit scalaire sur E.

Si de plus E est de dimension finie, on parle d'**espace euclidien**.

Exemple XXVI.2. L'exemple précédent met en exergue une structure d'espace préhilbertien réel sur $\mathscr{C}^0([a,b])$ une structure d'espace euclidien sur $\mathbb{R}^n$.

Remarque XXVI.2. Lorsque la donnée du produit scalaire sera sous–entendue/"évident", nous l'omettrons.

b) Inégalité de Cauchy–Schwarz

L'inégalité qui suit est nommée en l'honneur du mathématicien français Augustin Louis Cauchy (1789—1857) et de son confrère allemand Hermann Amandus Schwarz (1843—1921), qu'il serait de bon ton de ne pas confondre avec Laurent Schwartz (mathématicien français, 1915—2002).

Proposition XXVI.1 (Cauchy–Schwarz)**.** Soit E un espace préhilbertien réel ; alors pour tous $x, y \in E$ on a :

$$\langle x\,,\,y\rangle^2 \leq \langle x\,,\,x\rangle\,\langle y\,,\,y\rangle$$

avec égalité si et seulement si x et y sont colinéaires.

Démonstration. Commençons par remarquer que si $y = 0$, la proposition est relativement aisée à démontrer (même en cas d'électroencéphalogramme plat). Dans le cas contraire, remarquons que si $\lambda \in \mathbb{R}$ on a :

$$0 \leq \langle x+\lambda y\,,\,x+\lambda y\rangle = \langle x\,,\,x\rangle + 2\lambda\,\langle x\,,\,y\rangle + \lambda^2\,\langle y\,,\,y\rangle\ .$$

Nous avons donc la un trinôme du second degré en λ de signe constant (positif) : son discriminant est donc négatif ou nul, *i.e*

$$4\langle x\,,\,y\rangle^2 - 4\langle x\,,\,x\rangle\langle y\,,\,y\rangle \leq 0$$

ce qui nous livre l'inégalité désirée.

Concernant le cas d'égalité, ce dernier se produit si et seulement si le discriminant *supra* est nul, et donc si et seulement si notre trinôme admet une racine double, *i.e* un réel α tel que :

$$\langle x + \alpha y\,,\,x + \alpha y\rangle = 0\,.$$

Par caractère défini du produit scalaire, on a donc que $x + \alpha y = 0$, d'où le résultat.

□

➠ **Exemple XXVI.3.** Nous pouvons appliquer ce résultat aux produits scalaires vus dans le paragraphe précédent, obtenant les inégalités suivantes :

— $\forall x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$,

$$\left(\sum_{k=1}^n x_k y_k\right)^2 \leq \left(\sum_{k=1}^n x_k^2\right)\left(\sum_{k=1}^n y_k^2\right)\ ;$$

— $\forall f, g \in \mathscr{C}^0([a,b])$,

$$\left(\int_a^b f(t)g(t)\,\mathrm{d}t\right)^2 \leq \left(\int_a^b f(t)^2\,\mathrm{d}t\right)\left(\int_a^b g(t)^2\,\mathrm{d}t\right)\ ;$$

— $\forall A, B \in \mathscr{M}_n(\mathbb{R})$,

$$\mathrm{Tr}({}^{\mathrm{t}}AB)^2 \leq \mathrm{Tr}({}^{\mathrm{t}}AA)\mathrm{Tr}({}^{\mathrm{t}}BB)\,.$$

c) Norme euclidienne

Définition XXVI.3. Soit $(E, \langle\cdot\,,\,\cdot\rangle)$ un espace préhilbertien réel ; on appelle **norme euclidienne** associée au produit scalaire $\langle\cdot\,,\,\cdot\rangle$ sur E la fonction :

$$\begin{aligned}\|\cdot\| : E &\longrightarrow \mathbb{R}_+\\ x &\mapsto \sqrt{\langle x\,,\,x\rangle}\,.\end{aligned}$$

➠ **Exemple XXVI.4.**

— Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$; alors la norme de x associée au produit scalaire canonique est :

$$\|x\| = \sqrt{\sum_{k=1}^n x_k^2}\ ;$$

— de même, si $f \in \mathscr{C}^0([a,b])$, on pose :

$$\|f\|_2 = \sqrt{\int_a^b f(t)^2\,\mathrm{d}t}\,.$$

Il s'agit de la norme associée au produit scalaire vu précédemment sur cet espace.

✌ **Remarque XXVI.3.** Notons (avec soulagement) que la norme associée à un produit scalaire est bien définie par positivité de ce dernier.

✍ **Exercice XXVI.2.** Exprimer à l'aide des coefficients de $A \in \mathscr{M}_n(\mathbb{R})$ sa norme associée au produit scalaire $(A, B) \mapsto \mathrm{Tr}({}^{\mathrm{t}}AB)$.

Proposition XXVI.2. Soit E un espace préhilbertien réel et soit $x, y \in E$. Alors :

(i) $\|x\| \geq 0$;

(ii) $(\|x\| = 0) \Leftrightarrow (x = 0)$;

(iii) $\forall \lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda|\,\|x\|$ **[homogénéité]** ;

(iv) $\|x + y\| \leq \|x\| + \|y\|$ **[inégalité triangulaire]**.

Démonstration. Tout est trivial sauf l'inégalité triangulaire, qui découle de l'inégalité de Cauchy–Schwarz (proposition XXVI.1) comme suit :

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y \,,\, x+y \rangle \\
&= \langle x \,,\, x \rangle + 2\langle x \,,\, y \rangle + \langle y \,,\, y \rangle \\
&= \|x\|^2 + 2\langle x \,,\, y \rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\sqrt{\langle x \,,\, x \rangle \langle y \,,\, y \rangle} + \|y\|^2 \\
&= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2 .
\end{aligned}$$

□

✍ **Exercice XXVI.3.** En s'inspirant des résultats analogues vus dans les chapitres VI et VIII, démontrer que si x, y sont deux points d'un espace préhilbertien réel on a :

$$\|x - y\| \geq |\,\|x\| - \|y\|\,| .$$

✌ **Remarque XXVI.4.** Le cas d'égalité dans l'inégalité triangulaire se produit lorsque il y a égalité dans l'inégalité de Cauchy–Schwarz, *i.e* si et seulement si x et y sont colinéaires. De plus, si $x = \alpha y$ (avec $y \neq 0$) alors

$$\alpha = \frac{|\langle x \,,\, y \rangle|}{\|y\|^2} = \frac{\|x\|}{\|y\|} \geq 0 .$$

Muni de la norme associée à un espace préhilbertien réel E, il nous est possible de reformuler l'inégalité de Cauchy–Schwarz (proposition XXVI.1) de la façon suivante :

$$\forall x, y \in E, \quad |\langle x \,,\, y \rangle| \leq \|x\|\|y\| .$$

On en déduit que, si $x, y \neq 0$:

$$\frac{\langle x \,,\, y \rangle}{\|x\|\|y\|} \in [-1, 1] ,$$

ce qui motive la définition suivante.

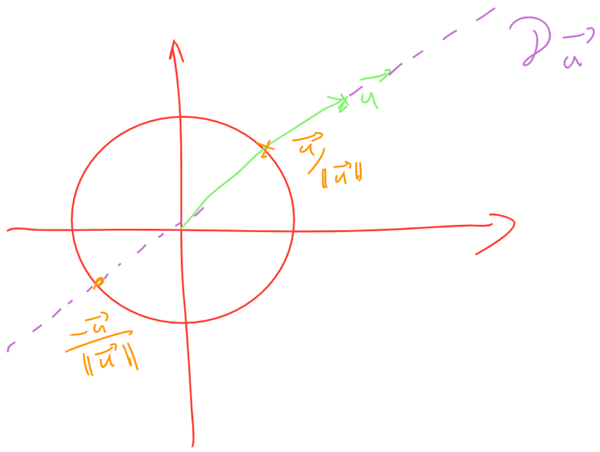
Définition XXVI.4. Soit E un espace préhilbertien réel et soient $x, y \in E \setminus \{0\}$. On appelle **écart angulaire** entre x et y la quantité :

$$\theta_{x,y} = \arccos\left(\frac{\langle x\,,\,y\rangle}{\|x\|\|y\|}\right) \in [0, \pi]\,.$$

✌ **Remarque XXVI.5.** En dimension 2 ou 3, ceci coïncide avec la notion éponyme vue en sciences appliquées.

Définition XXVI.5. Un vecteur de norme 1 est dit **unitaire**.

✌ **Remarque XXVI.6.** Un vecteur non nul u d'un espace préhilbertien réel est colinéaire à exactement deux vecteurs unitaires : $\frac{u}{\|u\|}$ et $-\frac{u}{\|u\|}$. Notons que ce résultat n'est vrai que parce que nous travaillons sur des $\mathbb{R}$–e.v.



d) Distance euclidienne

Définition XXVI.6. Soit E un espace préhilbertien réel ; on appelle **distance euclidienne** sur E l'application

$$\begin{aligned} d : E \times E &\to \mathbb{R}_+ \\ (x, y) &\mapsto \|y - x\|\,. \end{aligned}$$

➠ **Exemple XXVI.5.** On retrouve sur $\mathbb{R}^2$ et $\mathbb{R}^3$ la notion de distance vues en sciences physiques et de l'ingénieur.

Proposition XXVI.3. Soit E un espace préhilbertien réel et soient $x, y, z \in E$. Alors :

(i) $d(x, y) \geq 0$;

(ii) $(d(x, y) = 0) \Leftrightarrow (x = y)$;

(iii) $d(x, y) = d(y, x)$;

(iv) $d(x, z) \leq d(x, y) + d(y, z)$.

Démonstration. Tout ceci découle trivialement des propriétés de la norme euclidienne. □

✍ **Exercice XXVI.4.** Soit $f : x \mapsto x$ et $g : x \mapsto x^2$. Déterminer $d(f, g)$ dans $\mathscr{C}^0([0, 1])$.

e) Formulaire

Nous concluons ce paragraphe par une compilation de formules usuelles, dont la démonstration est laissée à la discrétion du lecteur, qui se révéleront sans nul doute fort utiles en pratique. Pour tout espace préhilbertien réel E et tous $x, y \in E$, on a :

(i)
$$\|x + y\|^2 = \|x\|^2 + 2\langle x\,,\,y\rangle + \|y\|^2\,;$$

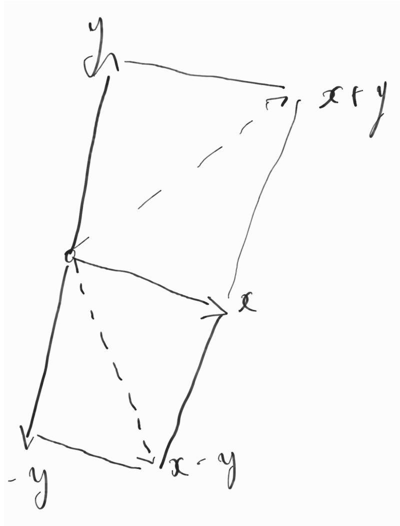
(ii)
$$\|x - y\|^2 = \|x\|^2 - 2\langle x\,,\,y\rangle + \|y\|^2\,;$$

(iii)
$$\langle x\,,\,y\rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2)\,;$$

(iv)
$$\langle x\,,\,y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)\,;$$

(v)
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)\,.$$

Les égalités (iii) et (iv) sont appelées **formules de polarisation** et établissent le fait que la donnée d'un produit scalaire est équivalente à celle de sa norme euclidienne associée. Le point (v) est connu sous le nom **d'identité du parallélogramme** et peut être visualisée géométriquement en dimension 2 (*cf. infra*).



2.– Orthogonalité

On fixe dans ce paragraphe un espace préhilbertien réel E.

a) Vecteurs orthogonaux

Définition XXVI.7. Deux vecteurs $x, y \in E$ sont dits **orthogonaux** si $\langle x\,,\,y\rangle = 0$.

Notation. Si x et y sont orthogonaux, on notera $x \perp y$.

➟ **Exemple XXVI.6.**
- Les vecteurs $(1,0)$ et $(0,1)$ sont orthogonaux dans $\mathbb{R}^2$;
- les fonctions $x \mapsto x^2$ et $x \mapsto x$ sont orthogonales dans $\mathscr{C}^0([-1,1])$ (mais pas dans $\mathscr{C}^0([0,1])$).

✌ **Remarque XXVI.7.**
- Le vecteur nul est orthogonal à tout vecteur de E. Réciproquement, si $x \in E$ est orthogonal à tout vecteur de E, alors $\langle x\,,\,x\rangle = 0$ et donc $x = 0$.
- Comme le produit scalaire est une forme bilinéaire symétrique on a, pour tous $x, y \in E$, $(x \perp y) \Leftrightarrow (y \perp x)$.

✍ **Exercice XXVI.5.** Soit E un espace euclidien.

1. Pour $a \in E$, montrer que l'application

$$\begin{aligned} \varphi_a : E &\to \mathbb{R} \\ x &\mapsto \langle a\,,\,x\rangle \end{aligned}$$

est une forme linéaire.

2. Démontrer que $a \mapsto \varphi_a$ réalise un isomorphisme (canonique) de E vers son dual E^*.

➥ **Correction :**

1. *Trivial*
2. *On vérifie que* $\mathrm{Ker}(\varphi) = \{0\}$ *car* 0 *est le seul vecteur orthogonal à tout élément de* E. *L'application* φ *est donc injective, ergo bijective car dim*$(E) = \dim(E^*)$.

Définition XXVI.8. Une famille $(x_i)_{i\in I}$ de vecteurs de E est dite **orthogonale** si :

$$\forall (i,j) \in I^2,\ (i \neq j) \Rightarrow (x_i \perp x_j)\,.$$

Si de plus les x_i sont tous unitaires, on parle de **famille orthonormée** (ou orthonormale).

✌ **Remarque XXVI.8.** Une famille $(x_i)_{i\in I} \in E^I$ est donc orthonormée si et seulement si :

$$\forall (i,j) \in I^2,\ \langle x_i\,,\,x_j\rangle = \delta_{i,j}\,.$$

➠ **Exemple XXVI.7.**

— La base canonique de $\mathbb{R}^n$ est orthonormée pour le produit scalaire canonique.

— La famille $(x \mapsto x^3, x \mapsto 1, x \mapsto x^7)$ n'est pas orthogonale dans $\mathscr{C}^0([-1,1])$.

Proposition XXVI.4. Toute famille orthogonale de vecteurs **non nuls** est libre.

Démonstration. Soit $(x_i)_{i\in I} \in E^I$ une famille orthogonale telle que $\forall i \in I,\ x_i \neq 0$. Alors, si il existe $\lambda_1,\dots,\lambda_n \in \mathbb{R}$ et $i_1,\dots,i_n \in I$ tels que $\sum_{k=1}^{n} \lambda_k x_{i_k} = 0$ on a, pour tout $j \in [\![1,n]\!]$:

$$\begin{aligned} 0 &= \left\langle \sum_{k=1}^{n} \lambda_k x_{i_k}\,,\, x_{i_j} \right\rangle \\ &= \sum_{k=1}^{n} \lambda_k \underbrace{\left\langle x_{i_k}\,,\,x_{i_j}\right\rangle}_{=0 \text{ si } k\neq j} \\ &= \lambda_j \|x_{i_j}\|^2 \end{aligned}$$

et donc $\lambda_j = 0$, ce qui livre le résultat voulu. □

Le théorème suivant est nommé en l'honneur du philosophe grec Pythagore de Samos (VI^e siècle avant J.–C.), auquel on attribue le cas particulier du triangle en dimension 2. On trouve cependant des traces d'un résultat similaire à ce dernier plus de mille ans auparavant en Mésopotamie. Concernant la démonstration, la plus ancienne qui nous soit connue est due à Euclide (300 av. J.–C.).

Théorème XXVI.5 (Pythagore)**.**
Soit $(x_1,\dots,x_n)$ une famille **orthogonale** de vecteurs de E. Alors :

$$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \sum_{k=1}^{n} \|x_k\|^2\,.$$

✌ **Remarque XXVI.9.** Si $n = 2$, la réciproque est vraie : en effet, pour toute famille $(x, y) \in E^2$ telle que $\|x + y\|^2 = \|x\|^2 + \|y\|^2$, on a :

$$\begin{aligned}\langle x\,,\,y\rangle &= \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2)\\ &= \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x\|^2 - \|y\|^2)\\ &= 0\end{aligned}$$

et donc (x, y) est orthogonale. Ceci devrait rappeler au lecteur quelques souvenirs collégiens.

Démonstration. Calculons :

$$\begin{aligned}\left\|\sum_{k=1}^{n} x_k\right\|^2 &= \left\langle \sum_{k=1}^{n} x_k\,,\,\sum_{j=1}^{n} x_j\right\rangle\\ &= \sum_{k=1}^{n}\sum_{j=1}^{n} \underbrace{\langle x_k\,,\,x_j\rangle}_{=0 \text{ si } j\neq k}\\ &= \sum_{k=1}^{n} \langle x_k\,,\,x_k\rangle\\ &= \sum_{k=1}^{n} \|x_k\|^2\,.\end{aligned}$$

□

b) Orthogonal d'une partie

Définition XXVI.9. Soit $A \subset E$ une **partie** de E. On appelle **orthogonal** de A l'ensemble

$$A^\perp = \{y \in E \,|\, \forall x \in A,\ \langle x\,,\,y\rangle = 0\}.$$

✌ **Remarque XXVI.10.** L'orthogonal de A est donc l'ensemble des vecteurs de E orthogonaux à tous ceux de A.

➠ **Exemple XXVI.8.** Posons $A = \{(1, 1, -1)\} \subset \mathbb{R}^3$. Alors, pour tout $(a, b, c) \in \mathbb{R}^3$, on a :

$$\begin{aligned}(a, b, c) \in A^\perp &\Leftrightarrow \langle (1, 1, -1)\,,\,(a, b, c)\rangle = 0\\ &\Leftrightarrow a + b - c = 0\end{aligned}$$

et donc $A^\perp$ est l'hyperplan de $\mathbb{R}^3$ d'équation $x+y-z = 0$. Une base de cet hyperplan est $((1, 0, 1), (1, 1, -2))$.

✍ **Exercice XXVI.6.** Déterminer l'orthogonal du plan d'équation $x + 2y - z = 0$ dans $\mathbb{R}^3$.

Proposition XXVI.6. Soient $A, B \in \mathscr{P}(E)$. Alors :
(i) $E^\perp = \{0\}$ et $\{0\}^\perp = E$;
(ii) $A^\perp$ est un s–e.v de E ;
(iii) $(A \subset B) \Rightarrow (B^\perp \subset A^\perp)$;
(iv) $A^\perp = \mathrm{Vect}(A)^\perp$;
(v) $A \subset (A^\perp)^\perp$.

Démonstration.

(i) Immédiat.

(ii) $A^\perp$ est bien une partie de A contenant 0. De plus, pour tous $x, y \in A^\perp$ et $\lambda \in \mathbb{R}$ on a :

$$\forall a \in A,\ \langle x + \lambda y\,,\, a\rangle = \langle x\,,\, a\rangle + \lambda \langle y\,,\, a\rangle$$
$$= 0$$

et donc $x + \lambda y \in A^\perp$, d'où le résultat.

(iii) Supposons $A \subset B$ et fixons $x \in B^\perp$. Alors, pour tout $y \in B$, $\langle x\,,\, y\rangle = 0$ et donc, comme $A \subset B$, ceci reste vrai pour tout $y \in A$. *In fine*, $B^\perp \subset A^\perp$.

(iv) Comme $A \subset \mathrm{Vect}(A)$, on par le point précédent que $\mathrm{Vect}(A)^\perp \subset A^\perp$. Réciproquement, tout vecteur orthogonal à A sera (par bilinéarité du produit scalaire) orthogonal à toute combinaison linéaires d'éléments de cet ensemble, d'où l'inclusion réciproque.

(v) Soit $x \in A$ et $y \in A^\perp$. Alors $\langle x\,,\, y\rangle = 0$ et donc $x \perp y$, ce qui entraîne que $x \in (A^\perp)^\perp$.

□

c) Algorithme de Gram–Schmidt

Le procédé exposé ici a été publié en 1883 par le mathématicien danois Jørgen Pedersen Gram (1850—1916) avant d'être reformulé par Erhard Schmidt (allemand, 1876—1959). On trouve un procédé analogue dans les écrits de Pierre–Simon de Laplace (mathématicien, astronome et physicien français, 1749—1827). Il permet, étant donné une base d'un espace euclidien, de la transformer en une base orthogonale ou orthonormée de ce même espace (on parle parfois de **procédé d'orthonormalisation**).

Théorème XXVI.7 (Gram–Schmidt)**.**
Soit E un espace euclidien et soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E. Alors il existe une base **orthogonale** $\mathscr{B}' = (u_1, \ldots, u_n)$ de E telle que :

$$\forall k \in [\![1, n]\!], \quad \mathrm{Vect}(e_1, \ldots e_k) = \mathrm{Vect}(u_1, \ldots, u_k)\,.$$

Démonstration. Posons $u_1 = e_1$, puis pour tout $k \in [\![1, n-1]\!]$:

$$u_{k+1} = e_{k+1} - \sum_{i=1}^{k} \frac{\langle e_{k+1}\,,\, u_i\rangle}{\|u_i\|^2} u_i\,.$$

Nous démontrons ensuite par récurrence sur $k \in [\![1, n]\!]$ que :

(i) $\mathrm{Vect}(e_1, \dots e_k) = \mathrm{Vect}(u_1, \dots, u_k)$;

(ii) $u_k \in \{u_1, \dots, u_{k-1}\}^{\perp}$ pour $k \geq 2$.

— $k = 1$: trivial.

— Supposons la propriété vraie au rang k ; le premier point est alors vérifié au rang $k+1$ par construction de u_{k+1}, il ne nous reste donc plus qu'à vérifier la condition d'orthogonalité. Pour ce faire, fixons $j \in [\![1, k]\!]$ et notons que :

$$\begin{aligned}
\langle u_{k+1}\,,\, u_j\rangle &= \left\langle e_{k+1} - \sum_{i=1}^{k} \frac{\langle e_{k+1}\,,\, u_i\rangle}{\|u_i\|^2} u_i\,,\, u_j \right\rangle \\
&= \langle e_{k+1}\,,\, u_j\rangle - \sum_{i=1}^{k} \frac{\langle e_{k+1}\,,\, u_i\rangle}{\|u_i\|^2} \underbrace{\langle u_i\,,\, u_j\rangle}_{=0 \text{ si } i\neq j} \\
&= \langle e_{k+1}\,,\, u_j\rangle - \frac{\langle e_{k+1}\,,\, u_j\rangle}{\|u_j\|^2}\|u_j\|^2 \\
&= 0
\end{aligned}$$

d'où le résultat. Notons que la famille $(u_1, \dots, u_n)$ est bien une base car orthogonale (donc libre) et génératrice car $\mathrm{Vect}(e_1, \dots e_n) = \mathrm{Vect}(u_1, \dots, u_n)$. □

➠ **Exemple XXVI.9.** Il est possible d'orthogonaliser la base $((1,0),(1,1))$ de $\mathbb{R}^2$ pour le produit scalaire canonique grace à ce procédé : on obtient $u_1 = (1,0)$ et

$$\begin{aligned}
u_2 &= (1,1) - \frac{\langle (1,1)\,,\,(1,0)\rangle}{\|(1,0)\|^2} \cdot (1,0) \\
&= (1,1) - (1,0) \\
&= (0,1)\,.
\end{aligned}$$

Corollaire XXVI.7.a. Tout espace euclidien admet une base orthonormée.

Démonstration. Nous savons par le théorème de la base incomplète (XX.2) qu'un tel espace admet une base. Via l'algorithme de Gram–Schmidt (théorème XXVI.7), nous obtenons une base orthogonale de E, notée $(u_1, \dots, u_n)$. Une base orthonormée de E est alors donnée par la famille $\left(\frac{u_1}{\|u_1\|}, \dots, \frac{u_n}{\|u_n\|}\right)$. □

✌ **Remarque XXVI.11.** Soit $\mathscr{B} = (e_1, \dots, e_n)$ une base **orthonormée** d'un espace euclidien E et soit $x, y \in E$; il existe de fait une unique famille $x_1, \dots, x_n, y_1, \dots y_n$ de réels telle que $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$. Avec ces notations, il est aisé de vérifier les formules (utiles) suivantes :

(i)
$$\forall i \in [\![1, n]\!], \quad x_i = \langle x \,,\, e_i \rangle \ ;$$

(ii)
$$\langle x \,,\, y \rangle = \sum_{i=1}^{n} x_i y_i \ ;$$

(iii)
$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2} \ .$$

✌ **Remarque XXVI.12.** Pour $i \in [\![1, n]\!]$, l'application : $x \mapsto \langle x \,,\, e_i \rangle$ est en fait la forme coordonnée e_i^*.

3.– Projecteurs orthogonaux, symétries orthogonales

On fixe dans ce paragraphe un espace **euclidien** E.

a) Supplémentaire orthogonal d'un s–e.v

Proposition XXVI.8. Soit F un s–e.v de E. Alors :

(i) $F \oplus F^{\perp} = E$;

(ii) $\dim(F^{\perp}) = \dim(E) - \dim(F)$;

(iii) $(F^{\perp})^{\perp} = F$.

Démonstration.

(i) Il est possible à l'aide d'une combinaison du théorème de la base incomplète (XX.2) et du procédé de Gram–Schmidt (théorème XXVI.7) de compléter une base orthonormée de F en une base orthonormée de E. Les vecteurs "ajoutés" de ce fait sont tous orthogonaux à F, ce qui permet de conclure, la nullité de $F \cap F^{\perp}$ étant triviale.

(ii) Découle immédiatement du point précédent.

(iii) Nous avons vu (proposition XXVI.6) que $F \subset (F^{\perp})^{\perp}$. De plus, on déduit du point précédent que $\dim((F^{\perp})^{\perp}) = \dim(F)$, ce qui permet de conclure.

□

Définition XXVI.10. Soit F un s–e.v de E. Alors $F^{\perp}$ est appelé **supplémentaire orthogonal** de F.

✌ **Remarque XXVI.13.** Il y a unicité du supplémentaire orthogonal.

➠ **Exemple XXVI.10.** Dans $\mathbb{R}^3$, le supplémentaire orthogonal d'un plan est l'unique droite vectorielle perpendiculaire (au sens vu dans le secondaire) à ce dernier.

b) Projecteurs, symétries : le retour (orthogonal)

Définition XXVI.11. Soit F un s–e.v de E ; alors on appelle :
- **projecteur orthogonal** sur F le projecteur sur F parallèlement à $F^\perp$;
- **symétrie orthogonale** par rapport à F la symétrie par rapport à F parallèlement à $F^\perp$.

✌ **Remarque XXVI.14.**
- On retrouve en dimension deux les symétries axiales vues en classe de 5^e : joie ! Il s'agit dans ce cas des symétries orthogonales par rapport à une droite de $\mathbb{R}^2$.
- Soit $\mathscr{F} = (u_1, \ldots, u_k)$ une base orthogonale de F, que l'on complète en une base $\mathscr{B} = (u_1, \ldots, u_n)$ orthogonale de E. On a alors, pour $x \in E$ la décomposition :
$$x = \sum_{i=1}^{n} \frac{\langle x\,,\, u_i\rangle}{\|u_i\|^2} u_i$$
et donc, si p est le projecteur orthogonal sur F, on a :
$$p(x) = \sum_{i=1}^{k} \frac{\langle x\,,\, u_i\rangle}{\|u_i\|^2} u_i\,.$$
Toute similarité avec les formules apparaissant dans l'algorithme de de Gram–Schmidt (théorème XXVI.7) est violemment non fortuite. Similairement, si $F = \mathscr{D}_e$ est une droite, on a :
$$\forall x \in E, \quad p(x) = \frac{\langle x\,,\, e\rangle}{\|e\|^2} e\,,$$
formule qui devrait titiller les souvenirs du lecteur relatifs aux sciences appliquées. Pour finir, on peut obtenir le projecteur p' sur l'hyperplan $H = \mathscr{D}_e^\perp$ via :
$$\forall x \in E, \quad p'(x) = x - \frac{\langle x\,,\, e\rangle}{\|e\|^2} e\,.$$

c) Distance à un s–e.v

Définition XXVI.12. Soit F un s–e.v de E et soit $x \in E$. On appelle **distance de x à F** la quantité
$$d(x, F) = \inf\{d(x, y) \,|\, y \in F\}\,.$$

Proposition XXVI.9. Soit F un s–e.v de E, soit $x \in E$ et soit p le projecteur orthogonal sur F. Alors on a :
$$\forall y \in F, \quad d(x, y) \geq d(x, p(x))$$
avec égalité si et seulement si $y = p(x)$.

Remarque XXVI.15. Cela signifie que le point $p(x)$ est l'unique vecteur de F minimisant la distance à x et que :

$$d(x, F) = d(x, p(x)).$$

Démonstration. Soit $y \in F$; alors :

$$\begin{aligned} d(x, y)^2 &= \|y - x\|^2 \\ &= \|(y - p(x)) + (p(x) - x)\|^2 \\ &= \|y - p(x)\|^2 + \|p(x) - x\|^2 \end{aligned}$$

d'après le théorème de Pythagore (XXVI.5). Ceci se reformule :

$$d(x, y)^2 = d(y, p(x))^2 + d(x, p(x))^2$$

d'où le résultat. □

Exemple XXVI.11.

— Si H est un hyperplan de E et que $e \in E$ est tel que $H^\perp = \mathscr{D}_e$, on a :

$$\forall x \in E, \quad d(x, H) = \frac{|\langle x\,,\,e\rangle|}{\|e\|}.$$

— Si $e \in E$ alors (calcul en exercice) :

$$\begin{aligned} \forall x \in E, \quad d(x, \mathscr{D}_e) &= \left\| x - \frac{\langle x\,,\,e\rangle}{\|e\|^2} e \right\| \\ &= \frac{\sqrt{\|x\|^2\|e\|^2 - \langle x\,,\,e\rangle^2}}{\|e\|}. \end{aligned}$$

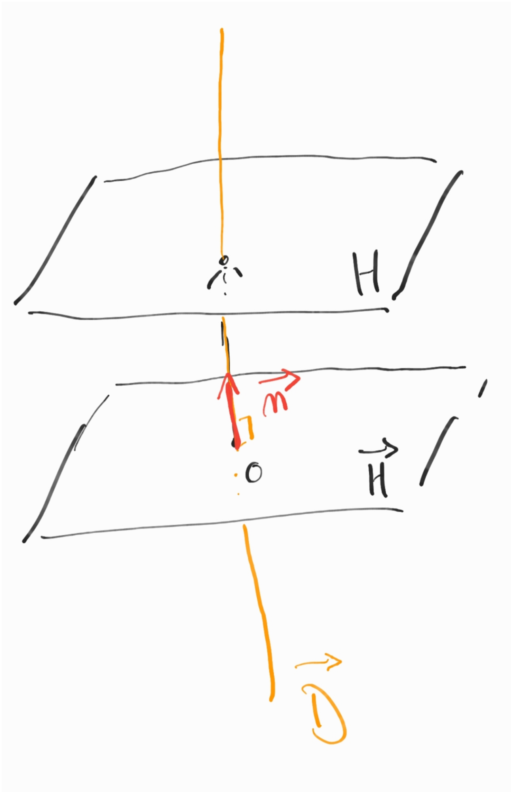
4.‾ Hyperplans affines d'un espace euclidien

On fixe dans ce paragraphe un espace **euclidien** E.

a) Vecteurs normaux à un hyperplan affine

Définition XXVI.13. Soit H un hyperplan affine de E de direction $\vec{H}$. On appelle **vecteur normal à** H tout vecteur directeur de la droite $\vec{H}^\perp$.

Remarque XXVI.16. Il s'agit donc d'un **vecteur** de $\vec{E}$.



✌ **Remarque XXVI.17.** Soit H un hyperplan affine de E de direction $\vec{H}$ et soit $\vec{n}$ un vecteur normal de H. On fixe une base **orthonormée** $(\vec{e_1}, \ldots, \vec{e_n})$ de E dans laquelle on peut écrire :

$$\vec{n} = \sum_{i=1}^{n} a_i \vec{e_i}\,.$$

Soit $\vec{x} = \sum_{i=1}^{n} x_i \vec{e_i} \in \vec{E}$; alors

$$(\vec{x} \in \vec{H}) \Leftrightarrow (\langle \vec{a}\,,\, \vec{x} \rangle = 0)$$

et donc une équation de $\vec{H}$ dans la base $(\vec{e_1}, \ldots, \vec{e_n})$ est donnée par :

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

et donc une équation de H sera de la forme :

$$\sum_{i=1}^{n} a_i x_i = b$$

avec $b \in \mathbb{R}$.

➠ **Exemple XXVI.12.** Ceci permet de déterminer des équations de droites (resp. plans) dans $\mathbb{R}^2$ (resp. $\mathbb{R}^3$), comme vu dans le secondaire.

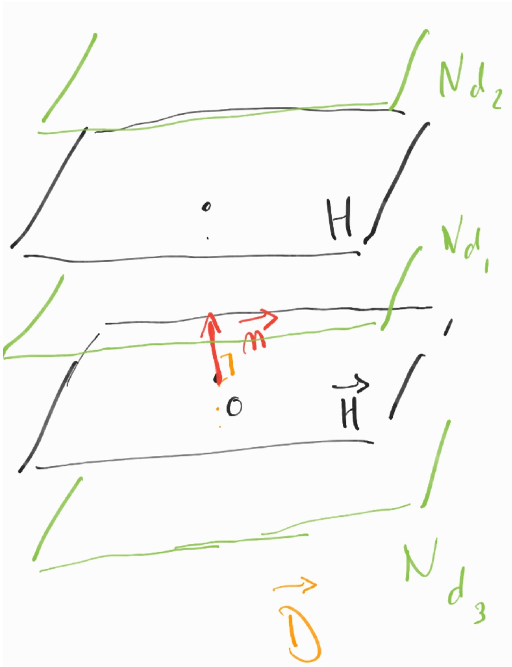
✌ **Remarque XXVI.18.** L'ajout d'un vecteur normal à une base d'un hyperplan vectoriel permet d'obtenir une base de $\vec{E}$. Si l'espace est orienté, celle–ci sera directe ou indirecte : il est donc possible d'orienter un hyperplan par le choix d'un vecteur normal.

b) Distance à un hyperplan affine

Proposition XXVI.10. Soit H un hyperplan affine de E passant par un point $A \in E$ et soit $\vec{n}$ un vecteur normal à H. Alors :

$$\forall M \in E, \quad d(M, H) = \left| \left\langle \vec{AM} \, , \, \vec{n} \right\rangle \right| .$$

✌ **Remarque XXVI.19.** Pour d parcourant $\mathbb{R}$, les ensembles $\mathcal{N}_d = \left\{ M \in E \,|\, \left\langle \vec{AM} \, , \, \vec{n} \right\rangle = d \right\}$ (appelés **lignes de niveaux** associées à A et $\vec{n}$) forment une famille d'hyperplans affines de même direction que H (on dit **parallèles**). Réciproquement, tout hyperplan affine parallèle à H est une ligne de niveau associée à A et $\vec{n}$.



Chapitre XXVII

Automorphismes orthogonaux

On fixe dans tout ce chapitre un espace euclidien $(E, \langle \cdot\,,\,\cdot \rangle)$ de dimension $n \geq 1$.

1.– Isométries vectorielles

a) Qu'est–ce ?

Définition XXVII.1. Une application linéaire $f \in \mathscr{L}(E)$ est appelée **isométrie vectorielle** ou **endomorphisme orthogonal** si :

$$\forall x \in E, \quad \|f(x)\| = \|x\| \,.$$

Notation. On note $\mathscr{O}(E)$ l'ensemble des endomorphismes orthogonaux sur E.

Exemple XXVII.1.
- $\mathrm{id}_E, -\mathrm{id}_E \in \mathscr{O}(E)$;
- $(x, y) \mapsto (x, -y)$ est une isométrie vectorielle de $\mathbb{R}^2$;
- $x \mapsto 2x$ n'est pas une isométrie vectorielle.

Proposition XXVII.1. L'image d'une base orthonormée par une isométrie vectorielle est une base orthonormée.

Démonstration. Fixons une base orthonormée $\mathscr{B} = (e_1, \ldots, e_n)$ de E et notons que pour tous $i, j \in [\![1, n]\!]$ on a, par une formule de polarisation :

$$\begin{aligned}
\langle f(e_i)\,,\, f(e_j)\rangle &= \frac{1}{2}(\|f(e_i) + f(e_j)\|^2 - \|f(e_i)\|^2 - \|f(e_j)\|^2) \\
&= \frac{1}{2}(\|f(e_i + e_j)\|^2 - \|f(e_i)\|^2 - \|f(e_j)\|^2) \\
&= \frac{1}{2}(\|e_i + e_j\|^2 - \|e_i\|^2 - \|e_j\|^2) \\
&= \langle e_i\,,\, e_j\rangle \\
&= \delta_{i,j} \,.
\end{aligned}$$

De fait la famille $f(\mathscr{B})$ est orthonormée et de cardinal n : il s'agit d'une base orthonormée de E. $\square$

✌ **Remarque XXVII.1.** Ceci entraîne que $\mathscr{O}(E) \subset GL(E)$: nous y reviendront.

Proposition XXVII.2. Soit $f : E \to E$ une application **non nécessairement linéaire**. Alors :

$$f \in \mathscr{O}(E)$$
$$\iff$$
$$\forall x, y \in E,\ \langle f(x)\,,\, f(y)\rangle = \langle x\,,\, y\rangle.$$

Démonstration.

($\Uparrow$) La conservation de la norme est triviale. Reste à démontrer la linéarité, ce qui se fait par un calcul laid et aussi passionnant qu'une rétrospective de l'œuvre de Patrick Sébastien.

($\Downarrow$) Supposons $f \in \mathscr{O}(E)$ et considérons $x, y \in E$. Alors, par une formule de polarisation :

$$\begin{aligned}\langle f(x)\,,\, f(y)\rangle &= \frac{1}{2}(\|f(x)+f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2)\\ &= \frac{1}{2}(\|f(x+y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2)\\ &= \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2)\\ &= \langle x\,,\, y\rangle\end{aligned}$$

d'où le résultat.

$\square$

✌ **Remarque XXVII.2.** Cette caractérisation permet de vérifier que les endomorphismes orthogonaux préservent les écarts angulaires.

Corollaire XXVII.2.a. Les symétries orthogonales sont des isométries vectorielles.

Démonstration. Soit $s \in \mathscr{L}(E)$ la symétrie orthogonale par rapport à un s–e.v F de E et soient $x, y \in E$. Alors il existe $(x_1, y_1) \in F^2$ et $(x_2, y_2) \in (F^\perp)^2$ tels que $x = x_1 + x_2$ et $y = y_1 + y_2$. Par caractérisation d'une symétrie (proposition XIX.14) on a de plus que $s(x) = x_1 - x_2$ et $s(y) = y_1 - y_2$ *ergo*

$$\begin{aligned}\langle s(x)\,,\, s(y)\rangle &= \langle x_1 - x_2\,,\, y_1 - y_2\rangle\\ &= \langle x_1\,,\, y_1\rangle - \langle x_1\,,\, y_2\rangle - \langle x_2\,,\, y_1\rangle + \langle x_2\,,\, y_2\rangle\\ &= \langle x_1\,,\, y_1\rangle + \langle x_2\,,\, y_2\rangle\ .\end{aligned}$$

Or

$$\begin{aligned}\langle x\,,\,y\rangle &= \langle x_1+x_2\,,\,y_1+y_2\rangle\\ &= \langle x_1\,,\,y_1\rangle+\langle x_1\,,\,y_2\rangle+\langle x_2\,,\,y_1\rangle+\langle x_2\,,\,y_2\rangle\\ &= \langle x_1\,,\,y_1\rangle+\langle x_2\,,\,y_2\rangle\end{aligned}$$

d'où le résultat. □

✌ **Remarque XXVII.3.** Les projecteurs orthogonaux n'étant pas en général bijectifs, ce ne sont pas des endomorphismes orthogonaux…

Définition XXVII.2. On appelle **réflexion** toute symétrie orthogonale par rapport à un hyperplan de E.

b) Groupe orthogonal

Proposition XXVII.3. L'ensemble $\mathscr{O}(E)$ est un sous–groupe de $GL(E)$.

Démonstration. Appliquons la caractérisation des sous–groupes (proposition X.8) à cet ensemble.

- $\mathscr{O}(E) \subset GL(E)$ car l'image d'une base orthonormée par un endomorphisme orthogonale est une base (orthonormée) d'après la proposition XXVII.1.
- $\mathscr{O}(E)$ est non vide car il contient id_E.
- Soit $f, g \in \mathscr{O}(E)$ et soit $x, y \in E$. Alors :

$$\begin{aligned}\langle f\circ g(x)\,,\,f\circ g(y)\rangle &= \langle f(g(x))\,,\,f(g(y))\rangle\\ &= \langle g(x)\,,\,g(y)\rangle\\ &= \langle x\,,\,y\rangle\end{aligned}$$

et donc $f \circ g \in \mathscr{O}(E)$. De même :

$$\begin{aligned}\left\langle g^{-1}(x)\,,\,g^{-1}(y)\right\rangle &= \left\langle g(g^{-1}(x))\,,\,g(g^{-1}(y))\right\rangle\\ &= \langle x\,,\,y\rangle\end{aligned}$$

et donc $g^{-1} \in \mathscr{O}(E)$.

□

✌ **Remarque XXVII.4.** Ceci signifie que tout endomorphisme orthogonal est en fait une bijection. On parlera donc à présent d'**automorphismes orthogonaux**. L'ensemble $\mathscr{O}(E)$ est appelé **groupe orthogonal**.

Proposition XXVII.4. Soit $f \in \mathscr{L}(E)$. Alors :

$$f \in \mathscr{O}(E)$$
$$\Longleftrightarrow$$
l'image d'une base orthonormée par f est une base orthonormée.

Démonstration. Le sens direct a déjà été traité (proposition XXVII.1) ; le sens indirect découle d'un rapide calcul en coordonnées. □

2.⁻ Matrices orthogonales

a) C'est quoi ?

Proposition XXVII.5. Soit $f \in \mathscr{L}(E)$ et soit $\mathscr{B}$ une base **orthonormée** de E. Si on pose $M = \mathrm{mat}_{\mathscr{B}}(f)$, alors les propriétés suivantes sont équivalentes :

(i) ${}^{\mathrm{t}}MM = I_n$;

(ii) $M\,{}^{\mathrm{t}}M = I_n$;

(iii) les lignes de M forment une base orthonormée de $\mathbb{R}^n$;

(iv) les colonnes de M forment une base orthonormée de $\mathbb{R}^n$;

(v) $f \in \mathscr{O}(E)$.

Démonstration.

(i)⇔(ii) Immédiat en passant à la transposée des deux côtés de l'égalité.

(ii)⇔(iii) Notons $(m_{i,j})_{i,j}$ les coefficients de la matrice M et $L_1, \dots, L_n$ ses lignes. Ceci signifie que, pour tout $i \in [\![1, n]\!]$, L_i peut être identifiée au vecteur $(m_{i,1}, \dots, m_{i,n}) \in \mathbb{R}^n$ et donc que pour tout $i, j \in [\![1, n]\!]$ on a, *via* cette convention :

$$\langle L_i \,,\, L_j \rangle = \sum_{k=1}^{n} m_{i,k} m_{j,k}$$

le produit scalaire *supra* désignant le produit scalaire canonique sur $\mathbb{R}^n$. On remarque alors que cette quantité est également le coefficient situé ligne i, colonne j de la matrice $M\,{}^{\mathrm{t}}M$. Celle–ci est donc égale à I_n si et seulement si :

$$\langle L_i \,,\, L_j \rangle = \delta_{i,j}$$

d'où le résultat.

(i)⇔(iv) Avec les mêmes notations que précédemment, on a, en notant $C_1, \dots C_n$ les colonnes de M, l'égalité suivante pour tous $i, j \in [\![1, n]\!]$:

$$\langle C_i \,,\, C_j \rangle = \sum_{k=1}^{n} m_{k,i} m_{k,j}$$

ce qui est égal au coefficient ligne i, colonne j de la matrice ${}^{\mathrm{t}}MM$. On conclut de la même façon que *supra*.

(iv)⇔(v) Les colonnes de M forment une base orthonormée de $\mathbb{R}^n$ si et seulement si l'image par f de $\mathscr{B}$ est une base orthonormée de E (le vérifier par un rapide calcul en coordonnées), d'où le résultat souhaité.

□

Définition XXVII.3. Une matrice $M \in \mathscr{M}_n(\mathbb{R})$ est dite **orthogonale** si ${}^{\mathrm{t}}MM = I_n$.

Notation. On note $O_n(\mathbb{R})$ ou $O(n)$ l'ensemble des matrices orthogonales de $\mathscr{M}_n(\mathbb{R})$.

Exemple XXVII.2.
- $I_n, -I_n \in O_n(\mathbb{R})$;
- pour tout $\theta \in \mathbb{R}$, $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in O_2(\mathbb{R})$;
- $\dfrac{1}{4}\begin{pmatrix} 3 & 1 & \sqrt{6} \\ 1 & 3 & -\sqrt{6} \\ -\sqrt{6} & \sqrt{6} & 2 \end{pmatrix} \in O_3(\mathbb{R})$.

Remarque XXVII.5.
- Il est clair que $O_n(\mathbb{R}) \subset GL_n(\mathbb{R})$; l'inverse d'une matrice orthogonale est de plus sa transposée.
- Les matrices orthogonales sont exactement les matrices **dans une même base orthogonale** d'automorphismes orthogonaux.

Proposition XXVII.6. L'ensemble $O_n(\mathbb{R})$ est un sous–groupe de $GL_n(\mathbb{R})$, appelé **groupe orthogonal d'ordre** n.

Démonstration. Le lecteur se fera un plaisir de vérifier ceci. □

Remarque XXVII.6. Si E est un espace euclidien de dimension n et que $\mathscr{B}$ est une base orthonormée de E, l'application $f \mapsto \mathrm{mat}_{\mathscr{B}}(f)$ réalise un isomorphisme de $\mathscr{O}(E)$ sur $O_n(\mathbb{R})$. Il en découle que si $\mathscr{B}$ et $\mathscr{B}'$ sont deux bases orthonormées de E, la matrice de passage $P_{\mathscr{B}}^{\mathscr{B}'}$ est orthogonale.

b) Rotations

Proposition XXVII.7. Soit $M \in \mathscr{O}_n(\mathbb{R})$. Alors $\det(M) \in \{-1, 1\}$.

Démonstration. On sait que ${}^{\mathrm{t}}MM = I_n$, *ergo* $\det({}^{\mathrm{t}}MM) = 1$. De plus, on a :

$$\begin{aligned}\det({}^{\mathrm{t}}MM) &= \det({}^{\mathrm{t}}M)\det(M) \\ &= \det(M)^2\end{aligned}$$

d'où le résultat. □

Proposition/définition XXVII.4. L'ensemble

$$SO_n(\mathbb{R}) = \{M \in \mathscr{O}_n(\mathbb{R}) \,|\, \det(M) = 1\}$$

est un sous–groupe de $\mathscr{O}_n(\mathbb{R})$, appelé **groupe spécial orthogonal d'ordre** n.

Démonstration. Il suffit d'appliquer la caractérisation de sous–groupe à cet ensemble (proposition X.8). □

Notation. L'ensemble $SO_n(\mathbb{R})$ est aussi noté $SO(n)$.

Remarque XXVII.7. De la même façon, on démontre que l'ensemble

$$\mathscr{SO}(E) = \{f \in \mathscr{O}(E) \,|\, \det(f) = 1\}$$

est un sous–groupe de $\mathscr{O}(E)$, lui aussi appelé **groupe spécial orthogonal**. Tout ceci fonctionne sans problème car $\det(f)$ ne dépend pas de la base dans laquelle il est calculé. Les éléments de $\mathscr{SO}(E)$ sont appelés **rotations** de E.

Exemple XXVII.3.

— id_E est une rotation. Le sort de $-\mathrm{id}_E$ dépend de la parité de $\dim(E)$ ...

— Pour tout $\theta \in \mathbb{R}$, $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in SO_2(\mathbb{R})$. On peut vérifier par un calcul que cette application réalise dans $\mathbb{R}^2$ la rotation d'angle θ au sens "collégien" du terme.

Les symétries orthogonales ne sont pas toujours des rotations : ceci dépend de la dimension de leur "axe". Les réflexions ne sont jamais (si $n \geq 1$) des rotations.

Vocabulaire. Plutôt que de parler de rotations, on parlera parfois de matrice orthogonale / isométrie positive. Leurs collègues de déterminant -1 seront affublés du qualificatif "négatif", ce qui est triste car ce sont des gens charmants.

c) Produit mixte

Dans ce paragraphe, l'espace E est supposé **orienté**.

Proposition XXVII.8. Soient $\mathscr{B}, \mathscr{B}'$ deux bases orthonormées **directes** de E. Alors $P_{\mathscr{B}}^{\mathscr{B}'} \in SO_n(\mathbb{R})$.

Démonstration. On a vu que $P_{\mathscr{B}}^{\mathscr{B}'} \in O_n(\mathbb{R})$. Nos deux bases étant directes, on a de plus $\det(P_{\mathscr{B}}^{\mathscr{B}'}) > 0$, ce qui entraîne que ce determinant est égal à 1, d'où le résultat. □

Corollaire XXVII.8.a. Soient $x_1, \ldots, x_n \in E$ et soient $\mathscr{B}, \mathscr{B}'$ deux bases orthonormées **directes** de E. Alors :

$$\det_{\mathscr{B}}(x_1, \ldots, x_n) = \det_{\mathscr{B}'}(x_1, \ldots, x_n)\,.$$

Démonstration. Il suffit d'écrire la formule de changement de base et d'utiliser la proposition précédente. □

Définition XXVII.5. Soit $x_1, \ldots, x_n \in E$. On appelle **produit mixte** des vecteurs $x_1, \ldots, x_n$ leur déterminant dans une base orthonormée directe.

Notation. $[x_1, \dots, x_n]$

✌ **Remarque XXVII.8.**
- Géométriquement, ceci correspond au volume orienté délimité par les x_i.
- Le produit mixte des vecteurs d'une base orthonormée directe (resp. indirect) est égal à 1 (resp. -1).
- Soit $f \in \mathscr{L}(E)$; alors on vérifie grâce aux propriétés du déterminant que :

$$[f(x_1), \dots, f(x_n)] = \det(f)[x_1, \dots, x_n]\,.$$

Dans le cas où $[x_1, \dots, x_n]$ est non nul, on a donc :

$$\det(f) = \frac{[f(x_1), \dots, f(x_n)]}{[x_1, \dots, x_n]}$$

ce qui permet de conclure que le déterminant d'une application linéaire traduit et quantifie l'action (dilatation, contraction) de celle–ci sur les volumes orientés de l'espace E.

3.⁻ Le groupe orthogonal du plan

On fixe dans ce paragraphe :
- $f \in \mathscr{O}(\mathbb{R}^2)$;
- $\mathscr{B}$ une base orthonormée **directe** de $\mathbb{R}^2$, que l'on suppose orienté relativement à sa base canonique.

On pose ensuite $M = \mathrm{mat}_{\mathscr{B}}(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et on cherche à déterminer (par analyse–synthèse) la nature de M et f.

a) Considérations liminaires

Comme la base $\mathscr{B}$ est orthonormée et que $f \in \mathscr{O}(\mathbb{R}^2)$, nous savons la matrice M orthogonale, ce qui entraîne que ${}^{t}MM = I_2$. Cette condition se traduit par le système (non linéaire) suivant :

$$\left\{\begin{array}{rcl} a^2 + c^2 & = & 1 \\ ab + cd & = & 0 \\ b^2 + d^2 & = & 1 \end{array}\right. . \qquad (\mathbf{E}\text{:XXVII.1})$$

Les lignes 1 et 3 *supra* nous permettent d'affirmer qu'il existe deux réels $\theta, \phi \in \mathbb{R}$ tels que $a = \cos(\theta)$, $c = \sin(\theta)$, $b = \cos(\phi)$ et $d = \sin(\phi)$. La ligne 2 devient alors :

$$\cos(\theta)\cos(\phi) + \sin(\theta)\sin(\phi) = 0$$

i.e

$$\cos(\theta - \phi) = 0\,. \qquad (\mathbf{E}\text{:XXVII.2})$$

Ceci entraîne que $\phi \equiv \theta \pm \frac{\pi}{2}\,[2\pi]$; nous devons donc traiter deux cas disjoints pour poursuivre notre étude.

b) Cas 1 : $\phi \equiv \theta - \dfrac{\pi}{2}\,[2\pi]$

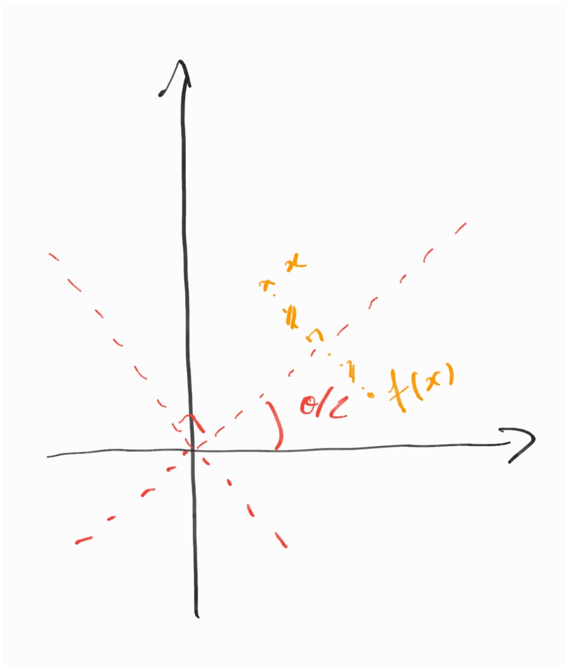
La matrice M est alors de la forme

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

Il s'agit bien d'une matrice orthogonale, et

$$\det(M) = -\cos(\theta)^2 - \sin(\theta)^2 = -1\,.$$

La matrice M est symétrique et orthogonal, ce qui implique que $M^2 = I_n$ et que f est une symétrie orthogonale. Mieux, un calcul rapide permet d'établir qu'il s'agit de la réflexion par rapport à la droite engendrée par le vecteur $\left(\cos\left(\frac{\theta}{2}\right), \sin\left(\frac{\theta}{2}\right)\right)$, dont l'orthogonale est la droite engendrée par $\left(-\sin\left(\frac{\theta}{2}\right), \cos\left(\frac{\theta}{2}\right)\right)$.



c) Cas 2 : $\phi \equiv \theta + \dfrac{\pi}{2}\,[2\pi]$

Dans ce cas, la matrice M est de la forme

$$\mathscr{M}(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Il s'agit bien d'une matrice orthogonale, et son déterminant est égal à

$$\cos(\theta)^2 + \sin(\theta)^2 = 1\,;$$

l'application f est donc une rotation. On en déduit que :

$$SO_2(\mathbb{R}) = \{\mathscr{M}(\theta) \mid \theta \in \mathbb{R}\}\,.$$

Un calcul aisé permet de démontrer que pour tous $\theta, \theta' \in \mathbb{R}$ on a $\mathscr{M}(\theta+\theta') = \mathscr{M}(\theta)\mathscr{M}(\theta')$, ce qui permet d'affirmer que l'application $\mathscr{M} : \theta \mapsto \mathscr{M}(\theta)$ réalise un morphisme de groupes surjectif de $\mathbb{R}$ vers $SO_2(\mathbb{R})$.

Ce morphisme n'est clairement pas injectif : on a pour tout $\theta \in \mathbb{R}$ $\mathscr{M}(\theta+2\pi) = \mathscr{M}(\theta)$. Nous pouvons toutefois construire un isomorphisme de groupes en remarquant que la restriction de $\mathscr{M}$ à $[0, 2\pi[$ est injective et en posant :

$$\begin{aligned} \psi : \mathscr{U} &\mapsto SO_2(\mathbb{R}) \\ e^{i\theta} &\mapsto \mathscr{M}(\theta) \, . \end{aligned}$$

Le groupe des rotations du plan est donc isomorphe à celui des nombres complexes de module 1.

✌ **Remarque XXVII.9.**

— Étant donné une rotation $\rho \in SO_2(\mathbb{R})$, l'unique $\theta \in [0, 2\pi[$ tel que $\mathrm{mat}_{\mathscr{B}}(\rho) = \mathscr{M}(\theta)$ ne dépend pas du choix de la base $\mathscr{B}$; on l'appelle **angle** de la rotation ρ.

— De plus, si $x, y \in \mathbb{R}^2$ sont deux vecteurs non nuls, il existe une unique rotation envoyant $\frac{x}{\|x\|}$ sur $\frac{y}{\|y\|}$; son angle θ correspond à l'écart angulaire entre x et y et on retrouve les formules vues en sciences appliquées :

$$\langle x \, , \, y \rangle = \|x\| \|y\| \cos(\theta)$$

et

$$[x, y] = \|x\| \|y\| \sin(\theta) \, .$$

Chapitre XXVIII

Probabilités

1.⁻ Notions liminaires

a) Univers

Ce cours sera axé autour de la notion d'**expérience aléatoire**, que nous définissons (informellement) comme une *expérience dont l'issue n'est pas déterminée à l'avance*. On pourra citer comme exemple le classique jet de dé(s), le remplissage d'un QCM, ...

Définition XXVIII.1. On appelle **univers** d'une expérience aléatoire l'ensemble, traditionnellement noté Ω, de ses issues possibles.

✖ **Attention :** nous nous limiterons cette année aux cas d'univers **finis**. Notons que cela ne signifie pas "petits"...

➠ **Exemple XXVIII.1.**

— Dans le cas d'un jet de dé à 6 faces, l'univers est $\Omega = [\![1, 6]\!]$;
— si on en lance deux, l'univers devient $\Omega = [\![1, 6]\!]^2$;
— l'univers associé à un lancer de pièce ("pile ou face") est $\Omega = \{\text{pile}, \text{face}\}$;
— l'univers associé à une étape d'un jeu de chaises musicales à n personnes est $\Omega = [\![1, n]\!] \times \mathfrak{S}_{n-1}$: il faut "choisir" le perdant, puis placer ses $n-1$ congénères sur les chaises restantes.

b) Événements

On fixe dans ce paragraphe un univers Ω **fini**.

Définition XXVIII.2. On appelle **événement** toute partie de l'univers Ω.

➠ **Exemple XXVIII.2.** Dans le cas de deux lancers "pile ou face" successifs, $\Omega = \{\text{pile}, \text{face}\}^2$ et l'événement "obtenir au moins une fois 'pile' " correspond à l'ensemble $\{(\text{pile}, \text{pile}), (\text{pile}, \text{face}), (\text{face}, \text{pile})\}$.

Vocabulaire. Plusieurs notions vues dans le chapitre II ont un pendant probabiliste rendu explicite par la définition *supra*. Nous les résumons dans le tableau suivant, en fixant deux parties $A, B \subset \Omega$ et un élément $\omega \in \Omega$.

Objet ou propriété	**Nom ensembliste**	**Nom probabiliste**
A	partie	événement
ω	élément	issue
$\{\omega\}$	singleton	événement élémentaire
$\Omega \setminus A$	complémentaire de A	événement contraire de A
$A \cup B$	réunion de A et B	A ou B
$A \cap B$	intersection de A et B	A et B
$\emptyset$	ensemble vide	événement impossible
Ω	tout	événement certain
$A \cap B = \emptyset$	ensembles disjoints	événements incompatibles

➠ **Exemple XXVIII.3.** Dans le cas de deux lancers "pile ou face" successifs, l'événement "obtenir au moins une fois 'pile' " est incompatible avec "n'obtenir que 'face' ".

Notation. Dans le cas des probabilités, il est courant et sans ambiguïté de noter $\bar{A}$ l'événement contraire $\Omega \setminus A$.

Définition XXVIII.3. On appelle **système complet d'événements** dans Ω toute famille $(A_1, \ldots, A_n)$ d'événements telle que :
- les A_i sont non impossibles (*i.e* non vides) ;
- les A_i sont deux à deux incompatibles (*i.e* deux à deux disjoints) ;
- $$\Omega = \bigsqcup_{i=1}^{n} A_i \, .$$

✌ **Remarque XXVIII.1.** Cela signifie que les parties $(A_i)_{i \in [\![1,n]\!]}$ partitionnent (*cf.* chapitre V) l'ensemble Ω.

➠ **Exemple XXVIII.4.**
- Si A est un événement différent de $\emptyset$ et Ω, alors $(A, \bar{A})$ est un système complet d'événements.
- Dans le cas du lancer successif de 14 dés à 6 faces, les événements A_i ="obtenir exactement i fois 1", pour $i \in [\![0, 14]\!]$, forment un système complet d'événements.

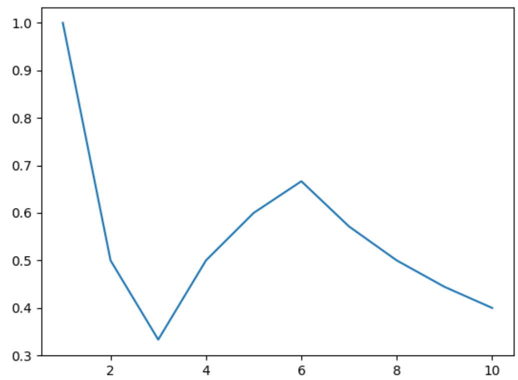
2.⁻ Espaces probabilisés finis

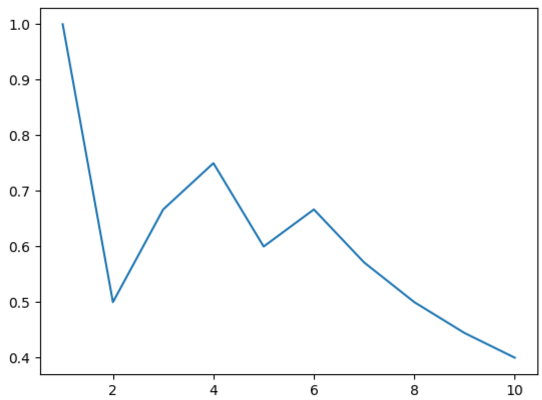
On fixe dans ce paragraphe un univers Ω **fini**.

Considérons l'experience aléatoire suivante : on lance une pièce de monnaie N fois (avec N "grand" et, n'en déplaise à H.P. Lovecraft, entier) et compte le nombre de fois où elle tombe sur "pile", que nous notons N_p. Intuitivement, il nous semble que le quotient $\frac{N_p}{N}$ (correspondant à la fréquence d'apparition de "pile") devrait, pour N suffisamment grand, être proche de $\frac{1}{2}$. Vérifions ceci en pratique, à l'aide du code `python` *infra*.

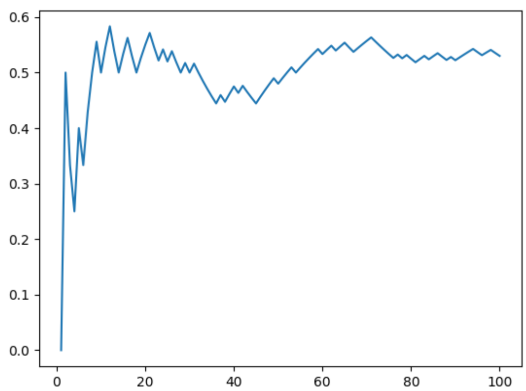
```
from numpy import *
from random import randint
from matplotlib.pyplot import *

def taux(N):
    S=0
    Y=[]
    for k in range(N):
        S+=randint(0,1)
        Y+=[S/(k+1)]
    plot(list(range(1,N+1)), Y)
    show()
```

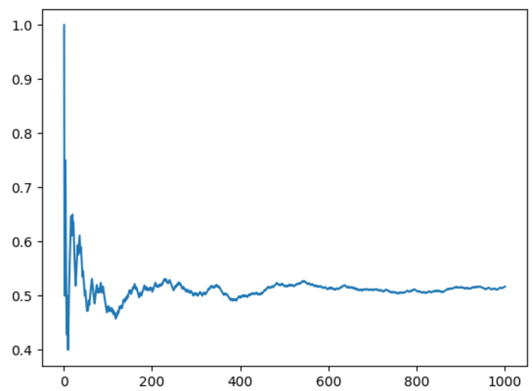
On remarque que les tracés obtenus pour de petits nombres de répétitions de l'expérience accusent une grande variabilité entre deux réalisations ; par exemple pour $N = 10$ on peut obtenir les deux graphes suivants :

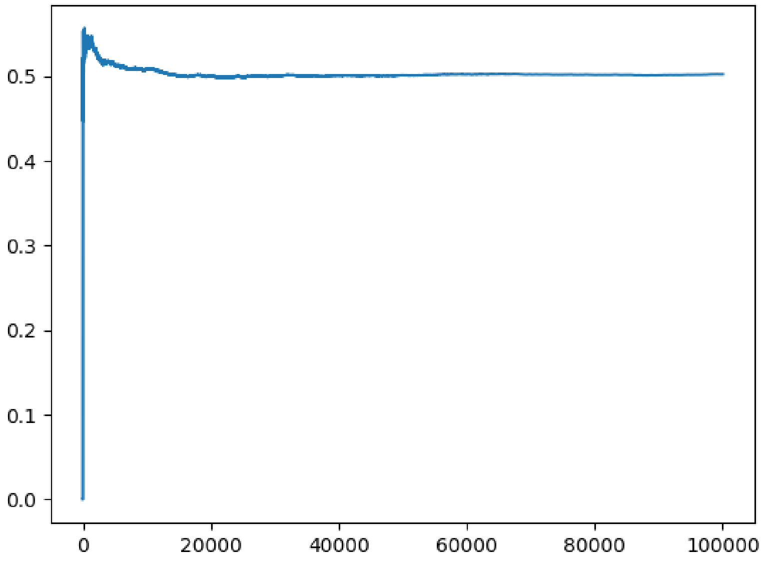




tandis qu'un grand nombre de répétitions a tendance à "lisser" le résultat, comme illustré *infra*.







Il semble donc en effet que la fréquence d'apparition de "pile" se stabilise, lorsque N tend vers l'infini, autour de la valeur $\frac{1}{2}$. Heuristiquement, nous avons la tentation d'avancer que la "chance" d'obtenir "pile" lors d'un tel lancer est de $\frac{1}{2}$; tentation à laquelle nous aller céder, non sans formaliser quelque peu notre cadre d'étude.

a) Notion de probabilité

Définition XXVIII.4. On appelle **probabilité** sur l'univers Ω toute application $\mathbb{P} : \mathscr{P}(\Omega) \to [0,1]$ telle que :
- $\mathbb{P}(\Omega) = 1$;
- pour tous événements **incompatibles** A et B on a :

$$\mathbb{P}(A \sqcup B) = \mathbb{P}(A) + \mathbb{P}(B)\,.$$

Le couple $(\Omega, \mathbb{P})$ est alors appelé **espace probabilisé** (fini).

✌ **Remarque XXVIII.2.**
- Si $A_1, \ldots, A_n$ sont des événements deux à deux incompatibles, on a (par récurrence) :
$$\mathbb{P}\left(\bigsqcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)\,.$$
- Une probabilité est caractérisée par ses valeurs en chaque événement élémentaire de Ω. En effet, si A est un événement, on a :
$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}\left(\bigsqcup_{\omega \in A} \{\omega\}\right)\\ &= \sum_{\omega \in A} \mathbb{P}(\{\omega\})\,.\end{aligned}$$

✖ **Attention :** ce raisonnement ne peut en aucun cas être généralisé au cas d'un univers infini.

Exemple XXVIII.5.
- La probabilité régissant un lancer de pièce de monnaie est caractérisée par $\mathbb{P}(\{\text{pile}\}) = \mathbb{P}(\{\text{face}\}) = \frac{1}{2}$;
- pour un lancer de dé à 6 faces, on a, pour $i \in [\![1, 6]\!]$:
$$\mathbb{P}(\text{obtenir } i) = \frac{1}{6} ;$$
- le lecteur amateur de magouilles pourra se pencher sur le cas d'un dé pipé.

Proposition XXVIII.1. Soit $\mathbb{P}$ une probabilité sur Ω. Alors $\mathbb{P}(\emptyset) = 0$.

Démonstration. L'événement impossible étant incompatible avec lui même (la chance ...), on a :
$$\mathbb{P}(\emptyset) = \mathbb{P}(\emptyset \sqcup \emptyset) = 2\mathbb{P}(\emptyset)$$
d'où le résultat. □

Remarque XXVIII.3. Il est possible d'avoir un événement A tel que $\mathbb{P}(A) = 1$ et $A \neq \Omega$. Prenons par exemple un pièce truquée pour toujours tomber sur "pile" et A l'événement "obtenir pile". De même, $\mathbb{P}(\text{"obtenir face"}) = 0$ et pourtant l'événement mentionné n'est pas impossible.

Proposition/définition XXVIII.5. Il existe une unique probabilité $\mathbb{P}$ sur Ω telle que :
$$\forall \omega, \omega' \in \Omega,\ \mathbb{P}(\{\omega\}) = \mathbb{P}(\{\omega'\})\,.$$
Celle–ci est alors appelée **probabilité uniforme** sur l'univers Ω.

Démonstration. Procédons par analyse–synthèse : si $\mathbb{P}$ est une telle probabilité alors, pour tout $\omega_0 \in \Omega$:
$$\begin{aligned} 1 &= \mathbb{P}(\Omega) \\ &= \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) \\ &= \operatorname{card}(\Omega)\mathbb{P}(\{\omega_0\}) \end{aligned}$$
ergo
$$\mathbb{P}(\{\omega_0\}) = \frac{1}{\operatorname{card}(\Omega)}\,.$$
En guise de synthèse, il nous suffit de vérifier que ceci caractérise bien une probabilité sur Ω. □

✌ **Remarque XXVIII.4.** Si $\mathbb{P}$ est la probabilité uniforme sur Ω et A est un événement, il est aisé de vérifier que :

$$\mathbb{P}(A) = \frac{\operatorname{card}(A)}{\operatorname{card}(\Omega)}\,.$$

✍ **Exercice XXVIII.1.** Quelle est la probabilité d'obtenir au moins un 4 en lançant 3 dés à 6 faces ?

➥ **Correction :** *Cette expérience est régie par la probabilité uniforme $\mathbb{P}$ sur $\Omega = [\![1,6]\!]^3$ car toutes ses issues sont équiprobables. Posons, pour $i \in [\![1,2,3]\!]$ l'événement $A_i =$"obtenir exactement i 4"; alors :*

$$\mathbb{P}(A_i) = \frac{\binom{3}{i}5^{3-i}}{6^3}$$

le coefficient binomial correspondant au choix des jets où le 4 est obtenu et le 5^{3-i} aux choix des valeurs affichées par les autres dés. Par incompatibilité on a alors :

$$\begin{aligned}\mathbb{P}(A_1 \sqcup A_2 \sqcup A_3) &= \frac{\binom{3}{1}5^2}{6^3} + \frac{\binom{3}{2}5}{6^3} + \frac{\binom{3}{3}}{6^3}\\ &= \frac{3\times 25 + 3\times 5 + 1}{216}\\ &= \frac{91}{216}\,.\end{aligned}$$

b) Propriétés générales

On fixe dans ce paragraphe une probabilité $\mathbb{P}$ sur Ω.

Proposition XXVIII.2 (Croissance)**.** Soient A, B deux événements tels que $A \subset B$. Alors :

(i) $\mathbb{P}(A) \leq \mathbb{P}(B)$;

(ii) $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.

Démonstration. Découle trivialement du fait que $B = A \sqcup (B \setminus A)$. □

Corollaire XXVIII.2.a. Soit A un événement. Alors $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.

✍ **Exercice XXVIII.2.** Refaire l'exercice XXVIII.1 en utilisant le corollaire XXVIII.2.a.

Proposition XXVIII.3. Soient A, B deux événements. Alors :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\,.$$

Démonstration. Découle, similairement au cas vu pour les cardinaux dans le chapitre XVIII, du fait que $A \cup B = (A \setminus A \cap B) \sqcup B$. □

✌ **Remarque XXVIII.5.** De façon analogue à ce qui a été fait au chapitre XVIII, si A, B et C sont trois événements on a :

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) =&\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ &- \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\ &+ \mathbb{P}(A \cap B \cap C)\end{aligned}$$

et on obtient, dans le cas de n événements, un analogue à la formule du crible de Poincaré évoquée dans ce même chapitre.

✌ **Remarque XXVIII.6.** On peut, similairement encore une fois à ce qui a été vu au chapitre XVIII, démontrer par une récurrence immédiate que pour tous événements $A_1, \dots, A_n$ on a :

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)\,.$$

✍ **Exercice XXVIII.3.** On lance 3 dés à 6 faces. Déterminer la probabilité que l'on obtienne au moins un 6 ou que la somme des résultats obtenus soit supérieure ou égale à 14.

3.⁻ Probabilités conditionnelles

On fixe dans ce paragraphe un espace probabilisé $(\Omega, \mathbb{P})$ **fini**.

a) Qu'est–ce ?

Définition XXVIII.6. Soient A, B deux événements tels que $\mathbb{P}(B) > 0$. On appelle **probabilité (conditionnelle) de** A **sachant** B la quantité

$$\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}\,.$$

Notation. $\mathbb{P}(A|B)$, $\mathbb{P}_B(A)$.

✌ **Remarque XXVIII.7.** Heuristiquement, cette quantité peut être interprétée comme une fréquence : on "compte" le "taux" de cas d'apparition de A et B parmi ceux de B.

➠ **Exemple XXVIII.6.** Pour citer une blague connue (parmi les mathématiciens, au moins) : *alors qu'un statisticien passe un contrôle de sécurité dans un aéroport, on découvre une bombe dans sa valise. En garde à vue, il s'explique : "La probabilité d'avoir une bombe dans un avion est certes faible, mais la chance d'avoir deux bombes dans un même avion est infime. Ainsi, je fais ma part pour la sécurité de tous."* Ce statisticien confond probabilité et probabilité conditionnelle.

Proposition XXVIII.4. Soit B un événement tel que $\mathbb{P}(B) > 0$. Alors l'application

$$\begin{aligned} \mathbb{P}_B : \mathscr{P}(\Omega) &\to [0,1] \\ A &\mapsto \mathbb{P}_B(A) \end{aligned}$$

est une probabilité.

Démonstration. Laissée en exercice à notre lecteur favori. □

✍ **Exercice XXVIII.4.** On considère une urne contenant r balles de couleur rouge, v balles de couleur verte et rien d'autre. On effectue deux tirages sans remise dans cette urne ; quelle est la probabilité de tirer deux balles vertes ?

➥ **Correction :** *Notons $N = r + v$ le nombre de balles initialement dans l'urne et A_i (pour $i = 1, 2$) l'événement "le i–ième tirage est une balle verte". Il est clair, par uniformité de la situation que*

$$\mathbb{P}(A_1) = \frac{v}{N}$$

et

$$\mathbb{P}_{A_1}(A_2) = \frac{v-1}{N-1}$$

ce qui entraîne que :

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}_{A_1}(A_2) = \frac{v(v-1)}{N(N-1)}.$$

b) Probabilités composées, probabilités totales

Théorème XXVIII.5 (Formules des probabilités composées)**.**
Soient $A_1, \dots, A_n$ des événements tels que $\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) > 0$. Alors :

$$\begin{aligned} \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) &= \mathbb{P}(A_1)\mathbb{P}_{A_1}(A_2)\mathbb{P}_{A_1\cap A_2}(A_3)\dots\mathbb{P}_{A_1\cap\dots\cap A_{n-1}}(A_n) \\ &= \prod_{i=1}^{n} \mathbb{P}\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right). \end{aligned}$$

Démonstration. Il suffit de remplacer chaque probabilité conditionnelle par leur définition et de simplifier le produit. Le lecteur masochiste trouvera sans doute également plaisant de le démontrer par récurrence sur n (prendre garde à bien quantifier les événements dans l'hypothèse de récurrence dans ce cas). □

✌ **Remarque XXVIII.8.** Le lecteur se souvenant encore de sa terminale aura sans doute des flash–backs avec des arbres n'ayant rien à voir avec le Vietnam. Ceci est non fortuit.

✍ **Exercice XXVIII.5.** Reprenons l'urne de l'exercice XXVIII.4 ; quelle est la probabilité de tirer les r balles rouges à la suite ?

Théorème XXVIII.6 (Formules des probabilités totales)**.**
Soit $(A_1, \dots, A_n)$ un système complet d'événements tel que $\forall i \in [\![1, n]\!]$, $\mathbb{P}(A_i) > 0$ et soit B un événement. Alors :

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}_{A_i}(B)\mathbb{P}(A_i)\,.$$

Démonstration. Remarquons que :

$$\begin{aligned} B &= B \cap \Omega \\ &= B \cap \left(\bigsqcup_{i=1}^{n} A_i \right) \\ &= \bigsqcup_{i=1}^{n} A_i \cap B \end{aligned}$$

et donc

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}\left(\bigsqcup_{i=1}^{n} A_i \cap B \right) \\ &= \sum_{i=1}^{n} \mathbb{P}(A_i \cap B) \\ &= \sum_{i=1}^{n} \mathbb{P}_{A_i}(B)\mathbb{P}(A_i)\,. \end{aligned}$$

□

✍ **Exercice XXVIII.6.** On considère trois usines fabriquant des boulons destiné à un usage bien précis, comme par exemple boulonner des trucs :
— l'usine 1 fournit 10% de la production et accuse un taux de boulons défectueux de 3% ;
— l'usine 2 fournit 50% de la production et accuse un taux de boulons défectueux de 12% ;
— l'usine 3 fournit 40% de la production et accuse un taux de boulons défectueux de 5%.

Démontrer que 8, 3% des boulons produits seront au final défectueux.

c) Formules de Bayes

Les résultats exposés dans ce paragraphe ont été formulés (sous une forme plus limitée) par le révérend anglais Thomas Bayes (∼1702—1761) et retrouvés indépendamment par Pierre-Simon de Laplace (français, 1749—1827), mathématicien, astronome, physicien et homme politique de son (ses ?) état(s).

Théorème XXVIII.7 (Bayes)**.**
Soient A, B deux événements et soit $(A_1, \dots, A_n)$ un système complet d'événements. On suppose que $\mathbb{P}(A)$, $\mathbb{P}(B)$ et les $\mathbb{P}(A_i)$ sont non nuls. Alors :

(i)
$$\mathbb{P}_B(A) = \frac{\mathbb{P}_A(B)\mathbb{P}(A)}{\mathbb{P}(B)} \ ;$$

(ii)
$$\forall j \in [\![1, n]\!] \ , \ \mathbb{P}_B(A_j) = \frac{\mathbb{P}_{A_j}(B)\mathbb{P}(A_j)}{\sum\limits_{i=1}^{n} \mathbb{P}_{A_i}(B)\mathbb{P}(A_i)} \ .$$

Démonstration.

(i) On a
$$\mathbb{P}(A \cap B) = \mathbb{P}_B(A)\mathbb{P}(B) = \mathbb{P}_A(B)\mathbb{P}(A)$$
d'où le résultat.

(ii) Soit $j \in [\![1, n]\!]$; alors, d'après le point (i), on a :
$$\mathbb{P}_B(A_j) = \frac{\mathbb{P}_{A_j}(B)\mathbb{P}(A_j)}{\mathbb{P}(B)} \ .$$
Or, d'après la formule des probabilités totales (théorème XXVIII.6) :
$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}_{A_i}(B)\mathbb{P}(A_i) \ .$$
ce qui permet de conclure.

□

✍ **Exercice XXVIII.7.** En reprenant le cadre de l'exercice XXVIII.6, démontrer qu'environ 72% des boulons défectueux proviennent de l'usine 2. *Quid* des usines 1 et 3 ?

d) Événements indépendants

Définition XXVIII.7. Deux événements A et B sont dits indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Notation. $A \perp\!\!\!\perp B$

✌ **Remarque XXVIII.9.** **Si** $\mathbb{P}(B) > 0$, alors

$$A \perp\!\!\!\perp B \iff \mathbb{P}_B(A) = \mathbb{P}(A)\,.$$

➠ **Exemple XXVIII.7.**
— Les résultats donnés par deux lancers de dés successifs sont indépendants ;
— il en va de même pour les résultats de tirages **avec remise** dans une urne.
— Nous invitons le lecteur à repenser au statisticien évoqué dans un paragraphe précédent et à ses bombes.

Définition XXVIII.8. On dit que des événements $A_1, \ldots, A_n$ sont **mutuellement indépendants** lorsque :

$$\forall J \subset [\![1, n]\!], \quad \mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)\,.$$

✌ **Remarque XXVIII.10.** Il est clair que l'indépendance mutuelle d'une famille implique l'indépendance deux à deux de ses éléments (prendre J de cardinal 2).

✖ **Attention :** la réciproque est cependant **fausse**. En effet, sur l'univers $\Omega = [\![1, 4]\!]$ muni de sa probabilité uniforme, les événements $A = \{1, 2\}$, $B = \{1, 3\}$ et $C = \{1, 4\}$ sont deux à deux indépendants tandis que :

$$\mathbb{P}(A \cap B \cap C) \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)\,.$$

Chapitre XXIX

Variables aléatoires

On fixe dans ce chapitre un espace probabilisé $(\Omega, \mathbb{P})$ **fini**.

1.⁻ Notion de variable aléatoire

a) C'est quoi ?

Nous cherchons dans ce chapitre à modéliser une "quantité" X prenant des valeurs aléatoires, dépendant des événements de l'univers Ω. Mathématiquement parlant, nous formalisons ceci par la définition *infra*.

Définition XXIX.1. On appelle **variable aléatoire** sur Ω toute application $X : \Omega \to E$, avec E un ensemble. Si $E = \mathbb{R}$, on parle de **variable aléatoire réelle**.

✌ **Remarque XXIX.1.** L'ensemble $X(\Omega)$ des valeurs prises par la variable aléatoire X est fini car Ω l'est.

➠ **Exemple XXIX.1.** Lançons deux dés à 6 faces ; l'univers nous intéressant est alors $\Omega = [\![1,6]\!]^2$. L'application

$$\begin{aligned} X : \Omega &\to \mathbb{N} \\ (x,y) &\mapsto x+y \end{aligned}$$

est alors une variable aléatoire sur Ω vérifiant $X(\Omega) = [\![2,12]\!]$.

Notation. Commettons quelques crimes contre la rigueur ; pour $X : \Omega \to E$ une variable aléatoire, $x \in E$ et $A \subset E$ on notera :
- $\{X \in A\}$ ou $(X \in A)$ l'événement $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$;
- $\{X = x\}$ ou $(X = x)$ l'événement $X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$;
- $\{X \leq x\}$ ou $(X \leq x)$ l'événement $\{\omega \in \Omega \mid X(\omega) \leq x\}$ si "$\leq$" est une relation sur E.

Ceci signifie que nous pourrons par exemple chercher à calculer la probabilité $\mathbb{P}(X \in A)$ de l'événement $X^{-1}(A)$; ces notations existent pour mettre l'accent que la partie intéressante de l'étude d'une variable aléatoire n'est pas celle de son mécanisme mais celle des valeurs qu'elle peut prendre, et avec quelle probabilité.

b) Loi d'une variable aléatoire

Définition XXIX.2. Soit $X : \Omega \to E$ une variable aléatoire ; on appelle **loi de** X l'application

$$\begin{aligned} \mathbb{P}_X : \mathscr{P}(E) &\to [0,1] \\ A &\mapsto \mathbb{P}(X \in A) \,. \end{aligned}$$

✌ **Remarque XXIX.2.** La loi $\mathbb{P}_X$ est caractérisée par les $\mathbb{P}_X(\{x\}) = \mathbb{P}(X = x)$ pour x parcourant $X(\Omega)$.

➠ **Exemple XXIX.2.** La loi de la somme X des résultats de deux dés à 6 faces est donnée par la tableau suivant.

x	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = x)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Proposition XXIX.1. Soit X une variable aléatoire ; alors $\mathbb{P}_X$ est une probabilité sur $X(\Omega)$.

Démonstration. Il est rapide de vérifier que $\mathbb{P}_X(X(\Omega)) = \mathbb{P}(X \in X(\Omega)) = 1$. De plus, si A et B sont deux événements incompatibles de $X(\Omega)$, alors :

$$\begin{aligned} \mathbb{P}_X(A \sqcup B) &= \mathbb{P}(X \in A \sqcup B) \\ &= \mathbb{P}(X^{-1}(A \sqcup B)) \\ &= \mathbb{P}(X^{-1}(A) \sqcup X^{-1}(B)) \\ &= \mathbb{P}(X^{-1}(A)) + \mathbb{P}(X^{-1}(B)) \\ &= \mathbb{P}_X(A) + \mathbb{P}_X(B) \end{aligned}$$

d'où le résultat. □

✌ **Remarque XXIX.3.**
- $X(\Omega)$ est donc lui aussi un univers (fini).
- Par définition de probabilité, si $A \subset E$ on a :

$$\mathbb{P}_X(A) = \sum_{x \in A} \mathbb{P}(X = x) \,.$$

c) Image d'une variable aléatoire par une application

Fixons dans ce paragraphe une variable aléatoire $X : \Omega \to E$ et une application $f : E \to F$, E et F étant deux ensembles quelconques. Alors l'application $f \circ X : \Omega \to F$ est elle aussi une variable aléatoire ; on la note (malheureusement) $f(X)$. Oui, je sais, nous avons affaire ici à une véritable hérésie notatoire.

Si $A \in \mathscr{P}(F)$ on a par ailleurs :

$$\begin{aligned}\mathbb{P}_{f(X)}(A) &= \mathbb{P}(f(X) \in A)\\ &= \mathbb{P}(X \in f^{-1}(A))\\ &= \mathbb{P}_X(f^{-1}(A))\,.\end{aligned}$$

✍ **Exercice XXIX.1.** Soit X la variable aléatoire correspondant à la somme des résultats d'un lancer de deux dés à 6 faces. Déterminer la loi de $X^2 - X$.

2.⁻ Zoologie des lois usuelles

a) Loi uniforme

Définition XXIX.3. Soit E un ensemble **fini**. On dit qu'une variable aléatoire $X : \Omega \to E$ suit une **loi uniforme** si :

$$\forall x \in E, \quad \mathbb{P}(X = x) = \frac{1}{\text{card}(E)}\,.$$

Notation. La loi $\mathbb{P}_X$ sera notée $\mathscr{U}(E)$ et on écrira $X \sim \mathscr{U}(E)$ pour indiquer que X suit une loi uniforme. Si $E = [\![1, n]\!]$, on écrira parfois $\mathscr{U}(n)$ au lieu de $\mathscr{U}(E)$.

✌ **Remarque XXIX.4.** Si $X \sim \mathscr{U}(E)$, on a, pour tout $A \subset E$,

$$\mathbb{P}_X(A) = \frac{\text{card}(A)}{\text{card}(E)}\,.$$

➠ **Exemple XXIX.3.** Un lancer de dé équilibré à 6 faces suit la loi $\mathscr{U}(6)$.

✍ **Exercice XXIX.2.** Soit $X \sim \mathscr{U}(n)$ pour $n \geq 0$. Démontrer que X^2 suit une loi uniforme sur un ensemble que l'on précisera.

b) Loi de Bernoulli

La loi décrite dans ce paragraphe est nommée en l'honneur du mathématicien suisse Jacques (ou Jakob selon son humeur) Bernoulli (1654—1705), qu'il ne faut pas confondre avec son frère Jean (1667—1748, mathématicien et physicien) ou ses neveux Daniel (1700—1782, médecin, physicien et mathématicien), Nicolas (1695–1726, mathématicien) et (encore) Jean (1710—1790, mathématicien). La prochaine fois que l'un de vos enseignants vous cite un "résultat de Bernoulli", demandez lui "lequel ?" ; effet (et représailles) assuré(es).

Définition XXIX.4. Soit $p \in [0, 1]$. On dit qu'une variable aléatoire $X : \Omega \to \{0, 1\}$ suit une **loi de Bernoulli** de paramètre p si

$$\mathbb{P}(X = 1) = p\,.$$

Notation. $X \sim \mathscr{B}(p)$

Exemple XXIX.4.
- un lancer de pièce équilibrée suit une loi $\mathscr{B}\left(\frac{1}{2}\right)$ (reste à décider qui de pile ou face correspond à 1 et qui correspond à 0) ;
- tirer une balle dans une urne en espérant qu'elle ait une couleur particulière suit une loi de Bernoulli.
- répondre à une question de QCM au hasard suit une loi de Bernoulli.

Remarque XXIX.5.
- Si $X \sim Br(p)$, comme $X(\Omega) \subset \{0,1\}$ alors on a nécessairement $\mathbb{P}(X=0) = 1-p$.
- Les lois de Bernoulli servent à modéliser les expériences aléatoires ayant exactement deux issues (on parle souvent, arbitrairement, de "succès" et "échec", mais cela reste une question de point de vue). Le paramètre p est alors la probabilité de "succès" de notre expérience. Encore une fois, avoir un flash–back arboricole ne semble pas déplacé.

Exercice XXIX.3. Soit $A \subset \Omega$; démontrer que $\mathbb{1}_A$ est variable aléatoire de loi $\mathscr{B}(\mathbb{P}(A))$.

c) Loi binomiale

Définition XXIX.5. Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$. On dit qu'un variable aléatoire $X : \Omega \to [\![0,n]\!]$ suit une **loi binomiale** de paramètres n et p si :

$$\forall k \in [\![0,n]\!], \quad \mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}\,.$$

Notation. $X \sim \mathscr{B}(n,p)$

Remarque XXIX.6. La formule *supra* définit bien une probabilité sur $[\![0,n]\!]$; en effet :

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1\,.$$

Exemple XXIX.5. On considère une urne contenant N balles, dont r rouges ($r \leq N$). Pour tout $n \in \mathbb{N}^*$, la variable aléatoire X correspondant au nombre de balles rouges tirées après n tirages **avec remise** suit la loi $\mathscr{B}\left(n, \dfrac{r}{N}\right)$.

Remarque XXIX.7. La loi $\mathscr{B}(n,p)$ modélise le nombre de succès lors de la répétition de n expériences indépendantes suivant une loi de Bernoulli $\mathscr{B}(p)$. D'ailleurs, $\mathscr{B}(1,p) = \mathscr{B}(p)$, ce qui n'est pas anodin. Nous y reviendrons un peu plus tard (*cf.* proposition XXIX.5).

3.⁻ Couples de variables aléatoires

Dans tout ce paragraphe, on fixe deux ensembles E et F et deux variables aléatoires $X : \Omega \to E$ et $Y : \Omega \to F$. On appelle **couple associé à** X **et** Y la variable aléatoire suivante :

$$\begin{aligned}(X,Y) : \Omega &\longrightarrow E \times F \\ \omega &\mapsto (X(\omega), Y(\omega)) .\end{aligned}$$

a) Loi conjointe, lois marginales

Définition XXIX.6. La loi $\mathbb{P}_{(X,Y)}$ du couple (X,Y) est appelée **loi conjointe** de ce dernier. Les lois $\mathbb{P}_X$ et $\mathbb{P}_Y$ sont quant à elles appelées **lois marginales** du couple (X,Y).

Exemple XXIX.6. Le tableau *infra* décrit la loi conjointe d'un couple (X,Y) de variables aléatoires à valeurs dans $\{0,1,2\}^2$.

$\mathbb{P}$	$X=0$	$X=1$	$X=2$
$Y=0$	0.1	0.3	0
$Y=1$	0	0.2	0.2
$Y=2$	0.1	0	0.1

On y lit par exemple que $\mathbb{P}((X,Y)=(0,2)) = 0.1$. On peut déduire de ces données les lois marginales en sommant les probabilités rencontrées sur les lignes et colonnes comme suit.

$\mathbb{P}$	$X=0$	$X=1$	$X=2$	$\mathbb{P}_Y$
$Y=0$	0.1	0.3	0	0.4
$Y=1$	0	0.2	0.2	0.4
$Y=2$	0.1	0	0.1	0.2
$\mathbb{P}_X$	0.2	0.5	0.3	$\times$

Remarque XXIX.8. De façon générale, il est toujours possible de déterminer les lois marginales à partir de la loi conjointe : si $a \in X(\Omega)$ on a

$$\{X=a\} = \bigsqcup_{b \in Y(\Omega)} (X=a, Y=b)$$

et donc :

$$\mathbb{P}(X=a) = \sum_{b \in Y(\Omega)} \mathbb{P}((X,Y)=(a,b)) .$$

Symétriquement, pour tout $b \in Y(\Omega)$ on a :

$$\mathbb{P}(Y=B) = \sum_{a \in X(\Omega)} \mathbb{P}((X,Y)=(a,b)) .$$

✖ **Attention :** il n'est par contre pas en général possible de reconstituer la loi conjointe à partir de la seule donnée des lois marginales. Se référer à l'exemple XXIX.6 pour un contre–exemple.

b) Loi conditionnelle

Proposition/définition XXIX.7. Soit $x \in E$ tel que $\mathbb{P}(X = x) > 0$; on appelle **loi conditionnelle de Y sachant que** $X = x$ la probabilité

$$\begin{aligned} \mathbb{P}_{Y|X=x} : \mathscr{P}(Y(\Omega)) &\longrightarrow [0,1] \\ B &\mapsto \mathbb{P}(B \,|\, X = x)\,. \end{aligned}$$

On définit symétriquement les lois conditionnelles de X sachant Y.

Exemple XXIX.7. Les lois conditionnelles sont données par les lignes et colonnes du tableau présenté dans l'exemple XXIX.6.

Remarque XXIX.9. Soit $y \in Y(\Omega)$. Alors

$$(Y = y) = \bigsqcup_{x \in X(\omega)} ((X, Y) = (x, y))$$

et donc

$$\begin{aligned} \mathbb{P}(Y = y) &= \sum_{x \in X(\Omega)} \mathbb{P}((X, Y) = (x, y)) \\ &= \sum_{x \in X(\Omega)} \mathbb{P}((X = x) \cap (Y = y)) \\ &= \sum_{x \in X(\Omega)} \mathbb{P}(Y = y \,|\, X = x)\mathbb{P}(X = x)\,. \end{aligned}$$

Ceci signifie que la donnée de $\mathbb{P}_X$ et des lois conditionnelles de Y par rapport à X permet de reconstituer $\mathbb{P}_Y$ (et symétriquement) ainsi que la loi conjointe $\mathbb{P}_{(X,Y)}$.

c) Généralisation aux n–uplets de variables aléatoires

Tout ce que nous venons d'exposer concernant les couples de variables aléatoires se généralise aux n–uplets de telles objets. On conserve la notion de loi conjointe, et on dispose désormais de n lois marginales. Les lois conditionnelles peuvent également être définies de façon analogue à ce qui a été fait *supra*. Les calculs auront une fâcheuse tendance à impliquer des doubles, triples, quadruples sommes et autres joyeusetés.

4.– Indépendance

a) Qu'est-ce ?

Définition XXIX.8. Soient X et Y deux variables aléatoires ; on dit que X et Y sont **indépendantes** si :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)\,.$$

Notation. $X \perp\!\!\!\perp Y$

➠ **Exemple XXIX.8.** Si on lance deux fois un dé, la variable aléatoire correspondant au premier résultat obtenu est indépendante de celle correspondant au second.

Proposition XXIX.2. Soient X, Y deux variables aléatoires **indépendantes**. Alors, pour tout $A \subset X(\Omega)$ et tout $B \subset Y(\Omega)$ on a :

$$\mathbb{P}((X,Y) \in A \times B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)\,.$$

✌ **Remarque XXIX.10.** Ceci se reformule par :

$$\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}_X(A)\mathbb{P}_Y(B)\,.$$

Démonstration. On a :

$$\begin{aligned}
\mathbb{P}((X,Y) \in A \times B) &= \sum_{a\in A}\sum_{b\in B} \mathbb{P}((X,Y) = (a,b)) \\
&= \sum_{a\in A}\sum_{b\in B} \mathbb{P}(X = a, Y = b) \\
&= \sum_{a\in A}\sum_{b\in B} \mathbb{P}(X = a)\mathbb{P}(Y = b) \\
&= \left(\sum_{a\in A} \mathbb{P}(X = a)\right)\left(\sum_{b\in B} \mathbb{P}(Y = b)\right) \\
&= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)
\end{aligned}$$

d'où le résultat. □

Corollaire XXIX.2.a. Soient X, Y deux variables aléatoires **indépendantes**. Alors pour tout $A \subset X(\Omega)$ et tout $B \subset Y(\Omega)$ les événements $(X \in A)$ et $(Y \in B)$ sont indépendants.

Proposition XXIX.3. Soient E, F, G, H quatre ensembles, $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires indépendantes et $f : E \to G$ et $g : F \to H$ deux applications. Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Démonstration. Laissé en exercice à notre cher lecteur. □

➠ **Exemple XXIX.9.** Si X et Y sont deux variables aléatoires réelles indépendantes, alors $(X^2 + 2) \perp\!\!\!\perp e^Y$.

b) Généralisation

Définition XXIX.9. Des variables aléatoires réelles $X_1, \dots, X_n$ sont dites **mutuellement indépendantes** si pour toute famille d'entiers deux à deux distincts $i_1, \dots, i_k \in [\![1, n]\!]$, tous $(x_{i_1}, \dots, x_{i_k}) \in X_{i_1}(\Omega) \times \dots \times X_{i_k}(\Omega)$ on a :

$$\mathbb{P}(X_{i_1} = x_{i_1}, \dots, X_{i_k} = x_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(X_{i_j} = x_{i_j})\,.$$

✌ **Remarque XXIX.11.** Une telle famille de variables peut modéliser une suite d'expériences aléatoires indépendantes. Ceci étant souvent modélisé par des arbres en Terminale...

Proposition XXIX.4. Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes. Alors pour toute famille $(A_1, \dots, A_n) \in \mathscr{P}(X_1(\Omega)) \times \dots \times \mathscr{P}(X_n(\Omega))$, les événements $(X_i \in A_i)$ sont mutuellement indépendants.

Démonstration. Trivial. □

Proposition XXIX.5. Soit $p \in [0, 1]$ et soient $X_1, \dots, X_n$ des variables aléatoires **mutuellement indépendantes** de loi $\mathscr{B}(p)$. Alors :

$$X_1 + \dots + X_n \sim \mathscr{B}(n, p)\,.$$

Démonstration. On le fait par récurrence sur n. Le cas $n = 1$ a déjà été traité et si on suppose tout ceci vrai au rang n alors pour toute famille $X_1, \dots, X_{n+1}$ de variables aléatoires mutuellement indépendantes de loi $\mathscr{B}(p)$, en remarquant que la variable X_{n+1} est à valeurs dans $\{0, 1\}$ on a, par incompatibilité, indépendance puis hypothèse de récurrence, pour tout $k \in [\![1, n+1]\!]$:

$$\begin{aligned}
\mathbb{P}(X_1 + \dots + X_{n+1} = k) &= \mathbb{P}(X_{n+1} = 0,\, X_1 + \dots + X_n = k) \\
&\quad + \mathbb{P}(X_{n+1} = 1,\, X_1 + \dots + X_n = k - 1) \\
&= \mathbb{P}(X_{n+1} = 0)\mathbb{P}(X_1 + \dots + X_n = k) \\
&\quad + \mathbb{P}(X_{n+1} = 1)\mathbb{P}(X_1 + \dots + X_n = k - 1) \\
&= (1 - p)\binom{n}{k} p^k (1 - p)^{n-k} + p\binom{n}{k-1} p^{k-1} (1 - p)^{n-k+1} \\
&= \left(\binom{n}{k} + \binom{n}{k-1}\right) p^k (1 - p)^{n+1-k} \\
&= \binom{n+1}{k} p^k (1 - p)^{n+1-k}\,.
\end{aligned}$$

Si $k = 0$, on vérifie par un calcul rapide que tout fonctionne, ce qui permet de conclure. □

Chapitre XXX

Espérance

On fixe dans ce chapitre un espace probabilisé $(\Omega, \mathbb{P})$ **fini**.

1.⁻ Mais qu'est–ce donc ?

a) Notion d'espérance

Définition XXX.1. Soit X une variable aléatoire réelle ; on appelle **espérance** de X la quantité

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\,\mathbb{P}(X = x)\,.$$

➠ **Exemple XXX.1.** Si $X \sim \mathscr{U}(n)$, on a :

$$\mathbb{E}(X) = \sum_{k=1}^{n} \frac{k}{n} = \frac{n+1}{2}\,.$$

✌ **Remarque XXX.1.** De façon générale, l'espérance de X est la moyenne de ses valeurs pondérée par leur fréquence (cette dernière étant donnée par la loi de X).

Vocabulaire. Si $\mathbb{E}(X) = 0$, on dit que la variable aléatoire X est **centrée**.

Proposition XXX.1. Soit X une variable aléatoire réelle ; alors :

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) X(\omega)\,.$$

Démonstration. On a :

$$\begin{aligned}
\sum_{\omega\in\Omega} \mathbb{P}(\{\omega\})X(\omega) &= \sum_{x\in X} \sum_{\omega\in X^{-1}(\{x\})} X(\omega)\mathbb{P}(\{\omega\}) \\
&= \sum_{x\in X} \sum_{\omega\in X^{-1}(\{x\})} x\,\mathbb{P}(\{\omega\}) \\
&= \sum_{x\in X} x \sum_{\omega\in X^{-1}(\{x\})} \mathbb{P}(\{\omega\}) \\
&= \sum_{x\in X} x\,\mathbb{P}(X^{-1}(\{x\}) \\
&= \sum_{x\in X(\Omega)} x\,\mathbb{P}(X = x)
\end{aligned}$$

d'où le résultat. □

✍ **Exercice XXX.1.** Soit $A \subset \Omega$. Démontrer que $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

Définition XXX.2. Soit X une variable aléatoire réelle et soit $k \in \mathbb{N}^*$. On appelle **moment d'ordre** k de X la quantité $\mathbb{E}(X^k)$.

b) Zoologie

◇ Variable (presque) constante

Définition XXX.3. Une variable aléatoire réelle est dite **constante presque sûrement** si il existe $m \in \mathbb{R}$ tel que $\mathbb{P}(X = m) = 1$.

✌ **Remarque XXX.2.** Il est aisé de démontrer que si X est constante presque sûrement égale à $m \in \mathbb{R}$ alors $\mathbb{E}(X) = m$. On en déduit que pour toute variable aléatoire réelle Y on a :

$$\mathbb{E}(\mathbb{E}(Y)) = \mathbb{E}(Y)\,.$$

◇ Loi uniforme

Proposition XXX.2. Soient $x_1, \dots, x_n \in \mathbb{R}$ et soit $X \sim \mathscr{U}(x_1, \dots, x_n)$. Alors :

$$\mathbb{E}(X) = \frac{1}{n}\sum_{k=1}^{n} x_k\,.$$

✌ **Remarque XXX.3.** Il s'agit de la moyenne arithmétique des valeurs de X.

◇ Loi de Bernoulli

Proposition XXX.3. Soit $p \in [0,1]$ et soit $X \sim \mathscr{B}(p)$. Alors :

$$\mathbb{E}(X) = p\,.$$

Démonstration. Il suffit de remarquer que :

$$\mathbb{E}(X) = 0 \times \mathbb{P}(X=0) + 1 \times \mathbb{P}(X=1) = p\,.$$

□

◇ Loi binomiale

Proposition XXX.4. Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$. Alors, si $X \sim \mathscr{B}(n,p)$ on a :

$$\mathbb{E}(X) = np\,.$$

Démonstration. On a :

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} n\frac{(n-1)!}{(k-1)!(n-1-(k-1))!}p^k(1-p)^{n-k} \\
&= np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-1-(k-1)} \\
&= np(p+1-p)^{n-1} \\
&= np
\end{aligned}$$

d'où le résultat. □

2.⁻ Propriétés de l'espérance

a) Propriétés élémentaires

Définition XXX.4. Un prédicat $P(A)$ dépendant d'un événement $A \subset \Omega$ sera dit **presque sûr** si $\mathbb{P}(P(A)) = 1$.

Exemple XXX.2. La variable aléatoire comptant le nombre de fois où une pièce équilibrée tombe sur la tranche parmi 42 lancers est nulle presque sûrement dans le modèle précédemment établi pour cette expérience aléatoire..

Proposition XXX.5. Soit X une variable aléatoire réelle telle que $X \geq 0$ presque sûrement ($i.e$ $\mathbb{P}(X \geq 0) = 1$) ; alors :

(i) $\mathbb{E}(X) \geq 0$;

(ii) $\mathbb{E}(X) = 0 \Leftrightarrow X$ est nulle presque sûrement ($i.e$ $\mathbb{P}(X = 0) = 1$).

Démonstration. Posons $X(\Omega) = \{x_1, \ldots, x_n\}$. Alors la quantité

$$\mathbb{E}(X) = \sum_{k=1}^{n} x_k \, \mathbb{P}(X = x_k)$$

est une somme de termes positifs, d'où le résultat. □

Corollaire XXX.5.a (Croissance de l'espérance)**.** Soient X, Y deux variables aléatoires réelles telles que $X \leq Y$ presque sûrement ; alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

Proposition XXX.6 (Linéarité de l'espérance)**.** Soient X, Y deux variables aléatoires réelles et soit $\lambda \in \mathbb{R}$. Alors :

$$\mathbb{E}(X + \lambda Y) = \mathbb{E}(X) + \lambda \mathbb{E}(Y) .$$

Démonstration. On a :

$$\mathbb{E}(X + \lambda Y) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\})(X(\omega) + \lambda Y(\omega))$$

d'où le résultat. □

Remarque XXX.4. On retrouve ainsi l'espérance d'une variable aléatoire de loi binomiale à l'aide de la proposition XXIX.5.

Exemple XXX.3. Soient $X, Y \sim \mathscr{U}(6)$; alors $\mathbb{E}(X + Y) = 7$. Nous venons de déterminer la somme moyenne des résultats de deux dés à 6 faces.

b) Lemme de transfert

Proposition XXX.7 (Lemme de transfert)**.** Soit $X : \Omega \to E$ une variable aléatoire **quelconque** et soit $\varphi : E \to \mathbb{R}$ une application. Alors :

$$\mathbb{E}(\varphi(X)) = \sum_{x \in E} \varphi(x) \mathbb{P}(X = x) .$$

✌ **Remarque XXX.5.** La somme *supra* est bien une somme finie, car si $x \notin X(\Omega)$, $\mathbb{P}(X = x) = 0$.

Démonstration. Posons $X(\Omega) = \{x_1, \ldots, x_n\}$. Alors, si l'on pose pour $k \in [\![1, n]\!]$ $A_k = \{X = x_k\}$ on a :

$$X = \sum_{k=1}^{n} x_k \mathbb{1}_{A_k}$$

et donc :

$$\varphi(X) = \sum_{k=1}^{n} \varphi(x_k) \mathbb{1}_{A_k} .$$

Ainsi :

$$\begin{aligned}\mathbb{E}(\varphi(X)) &= \mathbb{E}\left(\sum_{k=1}^{n} \varphi(x_k) \mathbb{1}_{A_k}\right) \\ &= \sum_{k=1}^{n} \varphi(x_k) \mathbb{E}(\mathbb{1}_{A_k}) \\ &= \sum_{k=1}^{n} \varphi(x_k) \mathbb{P}(A_k)\end{aligned}$$

d'après le résultat de l'exercice XXX.1. On en déduit le résultat. □

✌ **Remarque XXX.6.** Une conséquence importante du lemme de transfert est que l'espérance de $\varphi(X)$ est totalement déterminée par la loi de X.

➠ **Exemple XXX.4.** Soit $X \sim \mathscr{U}(n)$. Alors :

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{k=1}^{n} \frac{k^2}{n} \\ &= \frac{(n+1)(2n+1)}{6} .\end{aligned}$$

c) Inégalité de Markov

La proposition *infra* est ainsi nommée en l'honneur du mathématicien russe Andreï Andreïevitch Markov (1856—1922), connu comme l'un des pionniers de la théorie moderne des probabilités (on lui doit notamment la notion de processus stochastique).

Théorème XXX.8 (Inégalité de Markov)**.**
Soit X une variable aléatoire réelle telle que $X \geq 0$ presque sûrement et soit $t \in \mathbb{R}_+^*$. Alors :

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t} .$$

Démonstration. Posons $A = \{x \in X(\Omega) \,|\, x \geq t\}$. Alors :

$$\begin{aligned} t\,\mathbb{P}(X \geq t) &= t \sum_{x \in A} \mathbb{P}(X = x) \\ &= \sum_{x \in A} t\mathbb{P}(X = x) \\ &\leq \sum_{x \in A} x\mathbb{P}(X = x) \\ &\leq \mathbb{E}(X) \end{aligned}$$

d'où le résultat. □

✌ **Remarque XXX.7.** Soit X un variable aléatoire réelle. Alors on a, par exemple, que :

$$\mathbb{P}(X \geq 100\mathbb{E}(X)) \leq \frac{1}{100}\,.$$

Ceci nous permet d'estimer la répartition des valeurs de notre variable aléatoire X et de visualiser la façon dont la probabilité d'un écart à la moyenne donné évolue.

d) Lien à l'indépendance

Proposition XXX.9. Soient X et Y deux variables aléatoires réelles **indépendantes**. Alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

✖ **Attention :** l'indépendance est ici essentielle. Si $X \sim \mathscr{B}(p)$, on a $\mathbb{E}(X^2) = \mathbb{E}(X) = p$ car $X^2 = X$ presque sûrement. Pour peu que p ne soit pas égal à 0 ou 1, on a donc $\mathbb{E}(X^2) \neq \mathbb{E}(X)^2$.

Démonstration. Posons $X(\Omega) = \{x_1, \ldots, x_n\}$ et $Y(\Omega) = \{y_1, \ldots, y_m\}$. Alors, en appliquant le lemme de transfert au couple (X, Y) (nous demandons au lecteur de nous faire confiance ici : ce dernier se généralise au cas des fonctions de deux variables) on obtient :

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{k=1}^{n} \sum_{\ell=1}^{m} x_k y_\ell \mathbb{P}(X = x_k, Y = y_\ell) \\ &= \sum_{k=1}^{n} \sum_{\ell=1}^{m} x_k y_\ell \mathbb{P}(X = x_k)\mathbb{P}(Y = y_\ell) \end{aligned}$$

par indépendance. *In fine*, on a donc :

$$\begin{aligned} \mathbb{E}(XY) &= \left(\sum_{k=1}^{n} x_k \mathbb{P}(X = x_k) \right) \left(\sum_{\ell=1}^{m} y_\ell \mathbb{P}(Y = y_\ell) \right) \\ &= \mathbb{E}(X)\mathbb{E}(Y)\,. \end{aligned}$$

□

✖ **Attention :** la réciproque est **fausse** : si $X = m$ presque sûrement pour un certain $m \in \mathbb{R}$ on a $\mathbb{E}(X^2) = m^2 = \mathbb{E}(X)^2$ et pourtant X n'est pas indépendante d'elle même (avouez que ce serait sympathique, mais le monde est injuste).

3.⁻ Variance

a) C'est quoi ?

Connaître l'espérance d'une variable aléatoire réelle X nous permet d'avoir une idée de l'endroit moyen où se situent ses valeurs : c'est ce que l'on appelle un **paramètre de position**. Une telle donnée, bien qu'utile, n'est en général pas suffisante pour étudier finement le comportement de notre variable aléatoire malgré l'existence d'outils tels que l'inégalité de Markov (théorème XXX.8). La question qui nous préoccupe dans ce paragraphe est la suivante : comment diable sont réparties les valeurs de X ? Sont–elles resserrées autour de la moyenne ? Fortement dispersées ? Les quantités nous renseignant à ce sujet sont traditionnellement appelés **paramètres de dispersion**.

Une première idée pourrait être de déterminer l'espérance de la variable aléatoire réelle $|X - \mathbb{E}(X)|$; celle se révèle hélas en général d'une étude assez pénible. C'est pourquoi nous introduisons les quantités *infra*, dont nous dédierons cet ultime paragraphe de l'année à l'étude.

Définition XXX.5. Soit X une variable aléatoire réelle ; alors on appelle **variance** de X la quantité

$$V(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

et **écart–type** de X la quantité :

$$\sigma(X) = \sqrt{V(X)}\,.$$

Vocabulaire. Une variable aléatoire réelle de variance égale à 1 est dite **réduite**.

✌ **Remarque XXX.8.** Si nous souhaitions parler comme les physiciens, nous pourrions dire que $\sigma(X)$ est de même unité/dimension que X. Cette quantité représente l'écart moyen entre deux valeurs prises par la variable aléatoire X.

Le résultat *infra* est nommé en l'honneur du mathématicien (et élève de Jean Bernoulli, premier du nom) allemand Johann Samuel König (1712—1757), connu entre autres pour avoir traduit les éléments d'Euclide et du mathématicien, astronome et physicien néerlandais Christiaan Huygens (1629—1695), inventeur de la première horloge à pendule.

Théorème XXX.10 (König–Huygens)**.**
Soit X une variable aléatoire réelle ; alors :

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2\,.$$

✌ **Remarque XXX.9.** La variance est donc égale à la différence entre le moment d'ordre 2 et le carré de l'espérance.

Démonstration. On a :

$$\begin{aligned} V(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= E(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 . \end{aligned}$$

□

➠ **Exemple XXX.5.** Si X correspond à la somme des résultats de deux dés à 6 faces indépendants, on a $V(X) = \dfrac{35}{6}$ et $\sigma(X) \simeq 2.4$.

Proposition XXX.11. Soit X une variable aléatoire réelle et soient $a, b \in \mathbb{R}$. Alors :

(i) $V(aX + b) = a^2 V(X)$;

(ii) $\sigma(aX + b) = |a|\sigma(X)$.

Démonstration. Il s'agit d'un calcul direct. □

Corollaire XXX.11.a. Soit X une variable aléatoire réelle telle que $\sigma(X) > 0$; alors la variable aléatoire $\dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée réduite.

b) Zoologie

⋄ Variable presque constante

Soit X une variable aléatoire réelle telle que $X = m$ presque sûrement pour un certain $m \in \mathbb{R}$. Alors, il découle de la formule de König–Huygens (théorème XXX.10) que $V(X) = \sigma(X) = 0$. La réciproque est, pour une fois, vraie.

Proposition XXX.12. Soit X une variable aléatoire réelle telle que $V(X) = 0$. Alors X est constante presque sûrement.

Démonstration. Posons $Y = (X - \mathbb{E}(X))^2$; alors $\mathbb{E}(Y) = V(X) = 0$ et donc $Y = 0$ presque sûrement d'après la proposition XXX.5, d'où le résultat. □

⋄ Loi de Bernoulli

Proposition XXX.13. Soit $p \in [0, 1]$ et soit $X \sim \mathscr{B}(p)$. Alors $V(X) = p(1 - p)$.

Démonstration. Il suffit de faire le calcul en remarquant que $X^2 = X$ presque sûrement. □

◇ **Loi binomiale**

Proposition XXX.14. Soit $p \in [0,1]$ et soit $n \in \mathbb{N}^*$. Alors, si $X \sim \mathscr{B}(n,p)$ on a $V(X) = np(1-p)$.

Démonstration. On peut le faire par un calcul infect, mais nous le démontrerons un peu plus tard de façon non toxique *via* le corollaire XXX.18.a □

Remarque XXX.10. Si $X \sim \mathscr{B}(n,p)$ on a $\sigma(X) = \sqrt{p(1-p)}\sqrt{n} = o(n)$ lorsque $n \to \infty$. Cela signifie que les écarts entre les valeurs prises par X et l'espérance $\mathbb{E}(X)$ sont faibles devant sa moyenne. À titre d'exemple, pour $n = 10\,000$ et $p = \dfrac{1}{2}$ on a $\mathbb{E}(X) = 5000$ et $\sigma(X) = 50$ ...

c) Inégalité de Bienaymé–Tchebychev

Le résultat présenté dans ce paragraphe est nommé en l'honneur du mathématicien français Irénée-Jules Bienaymé (1796—1878), connu pour ses contributions à la théorie des probabilité et à la statistique, et de son homologue russe Pafnouti Lvovitch Tchebychev (1821—1894), dont les travaux ont porté sur les probabilités, la statistique et la théorie des nombres.

Théorème XXX.15 (Bienaymé–Tchebychev)**.**
Soit X une variable aléatoire réelle ; on pose $m = \mathbb{E}(X)$ et $\sigma = \sigma(X)$. Alors, pour tout $\lambda \in \mathbb{R}_+^*$ on a :

$$\mathbb{P}(|X - m| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}\,.$$

Démonstration. En appliquant l'inégalité de Markov (théorème XXX.8) à la variable aléatoire réelle $(X-m)^2$ avec $t = \lambda^2$, on obtient :

$$\mathbb{P}((X-m)^2 \geq \lambda^2) \leq \frac{\mathbb{E}((X-m)^2)}{\lambda^2}$$

soit

$$\mathbb{P}((X-m)^2 \geq \lambda^2) \leq \frac{V(X)}{\lambda^2}\,.$$

Il suffit alors pour conclure de remarquer que $V(X) = \sigma^2$ et que, par positivité de λ :

$$\mathbb{P}((X-m)^2 \geq \lambda^2) = \mathbb{P}(|X-m| \geq \lambda)\,.$$

□

Remarque XXX.11. Il peut être intéressant de formuler cette inégalité en remplaçant λ par le produit $\lambda\sigma$. On obtient alors :

$$\mathbb{P}(|X - m| \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$$

ce qui nous permet d'estimer l'écart de la variable aléatoire X à sa moyenne par incréments proportionnels à son écart–type.

➠ **Exemple XXX.6.** Lançons $n \geq 1$ dés à 6 faces et quantifions le nombre de 6 obtenus par une variable aléatoire X. Il est alors clair que $X \sim \mathscr{B}\left(n, \frac{1}{6}\right)$ et donc, par inégalité de Bienaymé–Tchebychev (théorème XXX.15), on obtient pour tout $\lambda > 0$ que :

$$\mathbb{P}\left(\left|X - \frac{n}{6}\right| \geq \lambda\right) \leq \frac{5n}{36\lambda^2}\,.$$

Pour $\lambda = \frac{n}{6}$, ceci devient

$$\mathbb{P}\left(\left|X - \frac{n}{6}\right| \geq \frac{n}{6}\right) \leq \frac{5}{n}\,.$$

On vérifie ensuite par un rapide calcul que $\left|X - \frac{n}{6}\right| \geq \frac{n}{6} \Leftrightarrow X \geq \frac{n}{3}$ et donc, *in fine* :

$$\mathbb{P}\left(X \geq \frac{n}{3}\right) \leq \frac{5}{n}$$

i.e

$$\mathbb{P}\left(X < \frac{n}{3}\right) \geq 1 - \frac{5}{n}\,.$$

Ceci signifie que lorsque $n = 10$, la probabilité que les 6 représentent moins d'un tiers des résultats obtenus est supérieure à $\frac{1}{2}$. Pour $n = 50$, elle dépasse $\frac{9}{10}$.

4.⁻ Covariance

a) Co–quoi ?

Définition XXX.6. Soient X, Y deux variables aléatoires réelles ; on appelle **covariance** de X et Y la quantité

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

✌ **Remarque XXX.12.** $\mathrm{Cov}(X,X) = V(X)$ pour toute variable aléatoire réelle X.

Proposition XXX.16. Soient X, Y deux variables aléatoires réelles ; alors :

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\,.$$

Démonstration. Il s'agit d'un calcul immédiat et remarquablement indolore (affirmation non contractuelle). □

Corollaire XXX.16.a. Soient X, Y deux variables aléatoires réelles **indépendantes** ; alors $\mathrm{Cov}(X, Y) = 0$.

✖ **Attention :** la réciproque est **fausse** (avouez, vous êtes surpris). Prenons $X \sim \mathscr{U}(-1, 0, 1)$ et posons $Y = X^2$: il est clair que X et Y ne sont pas indépendantes et pourtant $XY = X^3 = X$ presque sûrement donc

$$\mathbb{E}(XY) = \mathbb{E}(X) = 0 = \mathbb{E}(X)\mathbb{E}(Y)\,.$$

Proposition XXX.17. Soient X, Y et Z trois variables aléatoires réelles et soient $a, b, c, d \in \mathbb{R}$. Alors :

(i) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;

(ii) $\mathrm{Cov}(aX + b, cY + d) = ac\mathrm{Cov}(X, Y)$;

(iii) $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$.

Démonstration. Il s'agit d'une succession de calculs aussi enrichissants que la formidable activité consistant à regarder de la peinture sécher. □

b) Variance d'une somme de variables aléatoires

Proposition XXX.18. Soient X, Y deux variables aléatoires réelles. Alors :

(i) $V(X + Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y)$;

(ii) $V(X - Y) = V(X) + V(Y) - 2\mathrm{Cov}(X, Y)$.

Démonstration. Il suffit de remarquer que $V(X \pm Y) = \mathrm{Cov}(X \pm Y, X \pm Y)$ et de développer à l'aide des propriétés énoncées dans la proposition XXX.17. □

Démonstration. On lance deux dés à 6 faces, notant X le résultant du premier et Y celui du second. Alors la variance de la somme est :

$$V(X + Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y) = \frac{35}{12} + \frac{35}{12} + 2 \times 0 = \frac{35}{6}\,.$$

□

✌ **Remarque XXX.13.** On peut démontrer par récurrence (youpi) que pour toute famille $X_1, \ldots, X_n$ de variables aléatoires réelles, on a :

$$V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k) + 2\sum_{j=1}^{n}\sum_{i=1}^{j-1} \mathrm{Cov}(X_i, X_j)\,.$$

Corollaire XXX.18.a. Soient $X_1, \dots, X_n$ des variables aléatoires réelles **mutuellement indépendantes**. Alors :

$$V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k)\,.$$

✌ **Remarque XXX.14.** Ce résultat permet de retrouver aisément la variance d'une variable aléatoire réelle $X \sim \mathscr{B}(n,p)$ via la proposition XXIX.5.